50 grandi idee matematica



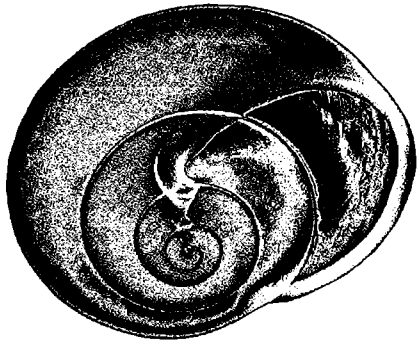
Tony Crilly

edizioni Dedalo

Chi ha inventato lo zero? Quanto è grande l'infinito? Dove si incontrano le rette parallele? Ed è vero che il battito d'ali di una farfalla può provocare una tempesta dall'altra parte del mondo? A queste e ad altre domande risponde un testo originale che presenta in 50 brevi saggi, chiari e rigorosi, i concetti matematici antichi e moderni, teorici e pratici, quotidiani ed esoterici che ci consentono di comprendere e modificare il mondo che ci circonda. In un percorso che procede dall'invenzione dello zero all'ultimo grande problema matematico rimasto insoluto, l'ipotesi di Riemann, Crilly guida il lettore alla scoperta della matematica che non si impara a scuola: spiega le potenzialità del calcolo e della statistica, chiarisce i concetti della relatività e della teoria del caos, svela la logica nascosta del Sudoku e il meccanismo dei giochi d'azzardo. I capitoli, ciascuno di quattro pagine, sono completati da una serie di interessanti inserti: aneddoti, linee del tempo, diagrammi esplicativi e approfondimenti.

Una perfetta introduzione alla matematica moderna che spiega anche i concetti più complessi in modo chiaro e accessibile a tutti.

50 grandi idee di matematica



Tony Crilly

edizioni Dedalo

© Tony Crilly 2007

Pubblicato su licenza Quercus Publishing Plc, London

Titolo originale: *50 mathematical ideas you really need to know*
Edited by Keith Mansfield

Traduzione di Laura Bussotti

Editing scientifico di Elena Ioli

Illustrazioni e diagrammi di Tony Crilly e Patrick Nugent

© settembre 2009 Edizioni Dedalo srl
Prima ristampa: ottobre 2009
Viale Luigi Jacobini 5, 70123 Bari - www.edizionidedalo.it

Volume di pagine 208, carta Bianco offset, gr. 120
Finito di stampare nell'ottobre 2009 dalla Dedalo litostampa srl, Bari

Tutti i diritti sono riservati.
Riproduzione vietata ai sensi di legge (art. 171 della legge 22 aprile 1941, n. 633)

Introduzione

La matematica è una disciplina molto vasta e non si può pensare di conoscerla tutta. Ciò che invece si può fare è esplorarla individuando un proprio percorso: le possibilità che ci si aprono davanti ci condurranno verso altre epoche, culture differenti dalla nostra e concetti che hanno stimolato la curiosità dei matematici per molti secoli.
La matematica è allo stesso tempo antica e moderna, ed è stata edificata a partire da un'ampia varietà di contributi politici e culturali. Il nostro attuale sistema di numerazione è di origine indiana e araba, ma è stato poi contaminato da una serie di stratificazioni storiche. La «base 60» dei Babilonesi di due o tre millenni a.C. è ancora presente nella nostra cultura, perché un minuto è costituito da 60 secondi e un'ora da 60 minuti; d'altra parte, un angolo retto continua a essere uguale a 90 gradi e non a 100, com'era stato stabilito nel periodo della Rivoluzione francese in un primo passo in direzione del sistema decimale.
I successi trionfali della tecnologia moderna si fondano sulla matematica e il fatto di non aver capito niente di matematica ai tempi della scuola non può certo essere considerato un motivo di vanto. Naturalmente la matematica che si impara a scuola è un po' diversa, perché spesso viene insegnata in funzione degli esami; né aiutano i ritmi forzati del calendario scolastico, perché in matematica la velocità non è un merito in sé. Ci vuole tempo per assorbire nuove idee. Alcuni dei più grandi matematici hanno proceduto con terribile lentezza, sforzandosi di comprendere in profondità i concetti su cui si basava il loro campo di indagine.
Questo non è un libro da leggere in fretta, ma da consultare e leggiucchiare con calma. Prendetevi il vostro tempo per scoprire il vero significato di idee che magari conoscete solo per sentito dire. Partite dallo zero, o da dove più vi piace, e veleggiate nell'arcipelago delle idee matematiche: ad esempio, potreste farvi una cultura sulla teoria dei giochi e poi mettervi a leggere come funzionano i quadrati magici, oppure passare dai rettangoli aurei al celeberrimo ultimo teorema di Fermat, o scegliere un qualunque altro percorso.
Quella in cui viviamo è un'epoca esaltante per la matematica. Di recente sono stati risolti alcuni dei problemi più famosi di tutti i tempi; in certi casi, i moderni computer si sono rivelati preziosi, in altri del tutto impotenti. Il problema dei quattro colori è stato risolto con l'ausilio di un computer, ma l'ipotesi di Riemann, trattata nell'ultimo capitolo del libro, resta irrisolta, con o senza computer.
La matematica è alla portata di tutti. Il successo del Sudoku dimostra che chiunque è in grado di fare matematica (anche senza saperlo), per di più divertendosi un sacco. La matematica, come l'arte o la musica, ha avuto i suoi geni, ma non si esaurisce in quelli. Vedrete eminenti personaggi entrare e uscire di scena in alcuni capitoli, per poi riapparire in altri. Eulero, di cui nel 2007 si è festeggiato il tricentenario della nascita, è un ospite assiduo di queste pagine, ma il reale progresso della matematica è frutto dei tanti studiosi che si sono succeduti nel corso dei secoli. Una selezione di 50 argomenti è necessariamente soggettiva, ma ho tentato di operare una scelta il più possibile equilibrata: troverete argomenti quotidiani e di livello superiore, matematica pura e applicata, idee astratte ed esempi concreti, cose vecchie e novità. La matematica è una disciplina unitaria, perciò la difficoltà non è stata scegliere alcuni argomenti, quanto lasciarne fuori altri. Queste idee avrebbero potuto essere 500, ma 50 sono sufficienti per celebrare il vostro ingresso nel mondo matematico.

01 Lo zero

Il nostro primo ingresso nel mondo dei numeri avviene con una certa gradualità. In un primo momento impariamo che 1 è il primo elemento dell'«alfabeto dei numeri» e che dà inizio alla serie dei numeri naturali 1, 2, 3, 4, 5... che servono per contare gli oggetti concreti, come mele, arance, banane, pere. Solo in seguito scopriremo che è possibile contare quante mele ci sono in un contenitore che non ne contiene nessuna.

Né gli antichi Greci, che fecero compiere passi da gigante alla fisica e alla matematica, né i Romani, celebri per le loro grandiose opere di ingegneria, disponevano di un metodo efficace per esprimere il numero di mele presenti in un contenitore vuoto. I Romani contavano combinando in vari modi i simboli, I, V, X, L, C, D e M, ma non conoscevano lo 0. Non sapevano contare «il nulla».

Come si è affermato lo zero? Si pensa che l'introduzione di un simbolo per indicare «il nulla» risalga a migliaia di anni fa. La civiltà Maya, che fiorì nell'attuale Messico, usava varie forme di zero. Qualche tempo dopo l'astronomo Claudio Tolomeo, influenzato dai Babilonesi, utilizzò un simbolo analogo al moderno 0 come «segnaposto» nel suo sistema di numerazione posizionale; in questo modo era possibile ad esempio distinguere con chiarezza (nella notazione moderna) 75 da 705, invece di riconoscerli dal contesto come facevano i Babilonesi. È un processo analogo all'introduzione della virgola nel linguaggio, perché entrambi aiutano a interpretare correttamente il significato di ciò che si legge; come la virgola, anche lo zero richiede alcune regole di utilizzo.

Nel VII secolo, il matematico indiano Brahmagupta non si limitò a usare lo zero come simbolo posizionale, ma lo trattò come un numero a tutti gli effetti; inoltre stabilì una serie di norme che ne regolavano l'utilizzo, come «la somma di un numero positivo e zero è positiva» e «la somma di zero e zero è zero». L'idea di Brahmagupta che lo zero fosse un numero e non un mero indicatore di posizione era un'importante novità concettuale. Il sistema di numerazione indo-arabico che comprendeva lo zero in questa accezione fu diffuso in Occidente da Leonardo Fibonacci (Leonardo da Pisa) che, nel 1202, pubblicò

linea del tempo

700 a.C.
I Babilonesi usano lo zero come «segnaposto» nel loro sistema numerico posizionale

628
Brahmagupta adotta lo zero e stabilisce delle regole per utilizzarlo in combinazione con le altre cifre

il *Liber Abaci (Il libro dell'abaco)*. Fibonacci, che era cresciuto in Africa settentrionale e aveva studiato l'aritmetica indo-arabica, aveva potuto constatare l'efficacia dell'utilizzo del simbolo 0 in combinazione con i simboli indiani 1, 2, 3, 4, 5, 6, 7, 8 e 9.

L'ingresso dello zero nel sistema di numerazione sollevò un problema che in parte si era posto anche Brahmagupta: come andava trattato quell'«intruso»? Il matematico indiano aveva iniziato ad affrontare la questione, ma non era giunto a conclusioni certe. Bisognava trovare il modo di integrare lo zero nel sistema aritmetico preesistente in modo sistematico. Alcune modifiche erano immediate, perché con l'addizione e la moltiplicazione lo zero funzionava benissimo, ma le operazioni di sottrazione e divisione non andavano molto d'accordo con la nuova cifra. Perché lo zero «si sposasse» con l'aritmetica già consolidata bisognava attribuirgli un significato.

Come funziona lo zero? Fare addizioni e moltiplicazioni che coinvolgono lo zero è semplice e non crea alcun problema; è bene precisare che, se si può pensare di aggiungere 0 a 10 in modo da ottenere 100, per «addizione» qui si intende banalmente l'operazione aritmetica. Se si somma 0 a un numero, questo resta invariato, mentre se si moltiplica un numero per 0 il risultato è sempre 0; ad esempio, $7 + 0 = 7$ e $7 \times 0 = 0$. Anche la sottrazione è un'operazione semplice, ma il risultato può essere un numero negativo ($7 - 0 = 7$ e $0 - 7 = -7$), mentre la divisione per 0 pone qualche problema.

Immaginiamo di dover misurare una lunghezza con un righello e supponiamo che quest'ultimo sia lungo esattamente 7 unità. Vogliamo sapere quanti righelli stanno nella lunghezza considerata. Se la lunghezza da misurare corrisponde precisamente a 28 unità, il risultato si ottiene dividendo 28 per 7, ovvero, in forma simbolica, $28 \div 7 = 4$. Una notazione migliore per rappresentare questa divisione è

$$\frac{28}{7} = 4$$

dopo di che si possono moltiplicare entrambi i membri per 7, trasformando la divisione in una moltiplicazione: $28 = 7 \times 4$. Se si vuole dividere 0 per 7, la faccenda si complica. Per trovare una soluzione, indichiamo il risultato della divisione con a:

$$\frac{0}{7} = a$$

830
Mahavira studia l'interazione dello zero con le altre cifre

1100
Bhaskara usa lo zero come simbolo algebrico e cerca di dimostrare come va trattato

1202
Fibonacci incorpora il simbolo 0 nel sistema di numerazione indo-arabico costituito dalle cifre 1, ..., 9 ma non lo considera un numero a tutti gli effetti

Moltiplicando per 7 si ottiene $0 = 7 \times a$. Se ciò è vero, a può assumere solo il valore 0. Infatti, se il prodotto di due numeri è uguale a 0, uno di essi deve essere 0; ovviamente non può essere il 7 e quindi $a = 0$.

La difficoltà maggiore, comunque, non è questa: la questione davvero spinosa è la divisione per 0. Se provassimo a trattare 7/0 come avevamo fatto con 0/7, otterremmo l'equazione

$$\frac{7}{0} = b$$

Moltiplicando entrambi i membri per 0 si otterrebbe $0 \times b = 7$ e quindi $0 = 7$, che è assurdo. Se si ammette che 7/0 possa essere un numero, ci si avvia verso un'enorme baraonda numerica. Non potendo dare un senso alla divisione di 7 (o di qualsiasi altro numero diverso da zero) per 0 si proibisce questa operazione, così come non si può inserire una virgola in *mez,zo* a una parola senza scrivere un'assurdità.

Seguendo le orme di Brahmagupta, il matematico indiano del XII secolo di nome Bhaskara prese in esame la divisione per zero e ipotizzò che dividendo un numero per 0 si ottenesse un valore infinito. È un'affermazione ragionevole, perché quando si divide un numero per un altro numero molto piccolo si ottiene un risultato molto grande; ad esempio, 7 diviso un decimo fa 70, 7 diviso un centesimo fa 700. Al diminuire del divisore il risultato aumenta sempre di più e, quando il divisore è il numero più piccolo possibile, lo zero, diventa infinitamente grande. Questo ragionamento, però, porta a dover spiegare un concetto ancora più strano: l'infinito. Combattere con l'infinito (indicato con il simbolo ∞) non serve a niente, perché non rispetta le consuete regole aritmetiche e non è un numero nel senso comune del termine.

Se 7/0 rappresentava un problema, che dire di una frazione ancora più strana, 0/0? Partendo da $0/0 = c$, e moltiplicando entrambi i termini per 0 si arriva all'equazione $0 = 0 \times c$, e alla conclusione che $0 = 0$. Non è un risultato particolarmente illuminante, ma non è neanche assurdo. Infatti, l'equazione scritta sopra è valida *per qualsiasi c*; ciò significa che 0/0 può essere qualunque cosa, ovvero, come si dice nei raffinati circoli matematici, è «indeterminato».

A proposito della divisione per zero si arriva quindi rapidamente alla conclusione che il meglio che si possa fare è escluderla dai calcoli. L'aritmetica può farne tranquillamente a meno.

A cosa serve lo zero? Lo zero è assolutamente indispensabile e ha svolto un ruolo essenziale nel progresso della scienza. Se in ambito scientifico si parla comu-

nemente di zero gradi di longitudine, zero gradi di temperatura e «gravità zero», questa cifra è entrata anche nel linguaggio comune, con espressioni come «ora zero» e «tolleranza zero».

Ma il suo campo di applicazione potrebbe essere ancora più ampio. Chi entra nell'Empire State Building, a New York, si ritrova nel maestoso atrio di ingresso situato al Piano 1. In questo caso i numeri indicano l'ordine dei piani: 1 sta per «primo», 2 per «secondo» e così via fino a 102 per «centoduesimo». In Europa esiste anche il piano 0, ma in genere si preferisce chiamarlo in altro modo.

La matematica non potrebbe funzionare senza lo zero, perché questa cifra è il perno dei concetti matematici che fanno girare il sistema di numerazione, l'algebra e la geometria. Oltre a occupare una posizione privilegiata sulla retta dei numeri, dove separa i numeri positivi da quelli negativi, lo zero agisce da indicatore di posizione nel sistema decimale, consentendoci di utilizzare numeri enormi o microscopici.

Nel corso dei secoli lo zero si è affermato ed è diventato di uso comune, collocandosi fra le invenzioni più importanti dell'umanità. Il matematico americano del XIX secolo G.B. Halsted, parafrasando il *Sogno di una notte di mezza estate* di Shakespeare, lo ha descritto come il motore del progresso che conferisce «a un aereo nulla, non solo una concreta dimora e un nome, un'immagine e un simbolo, ma anche un'efficacia pratica, com'è caratteristico della stirpe indiana che l'ha creato».

Quando è stato introdotto, lo zero deve essere sembrato una stranezza, ma i matematici hanno il vizio di attaccarsi a concetti bizzarri che si rivelano utili solo dopo molto tempo. L'equivalente moderno dello zero si trova nella teoria degli insiemi (dove per *insieme* si intende un raggruppamento di elementi): è l'insieme che non contiene nessun elemento, il cosiddetto «insieme vuoto», che si indica con Ø. Anche questa è un'idea strana, ma è assolutamente indispensabile, proprio come lo zero.

Tutto sul nulla

La somma di zero e di un numero positivo è positiva

La somma di zero e di un numero negativo è negativa

La somma di un numero positivo e di uno negativo corrisponde alla loro differenza oppure, se sono uguali, è zero

Zero diviso per un numero negativo o positivo è uguale a zero o si può scrivere come una frazione che ha zero al numeratore e la quantità finita al denominatore

Brahmagupta, 628

idea chiave

Tanto rumore per nulla

02 I sistemi di numerazione

Un sistema di numerazione è un metodo che consente di trattare il concetto del «quanti sono». In periodi diversi, culture diverse hanno adottato metodi di vario tipo, dal semplice «uno, due, tre, molti» alla sofisticata notazione decimale posizionale che utilizziamo oggi.

I Sumeri e i Babilonesi che circa 4000 anni fa popolavano le regioni oggi corrispondenti a Iraq, Siria e Giordania utilizzavano quotidianamente un sistema di numerazione «posizionale», ovvero un sistema in cui i numeri erano determinati dalla posizione dei simboli. L'unità di base era il numero 60 e per questo motivo oggi lo definiamo un sistema «in base 60». Alcune vestigia della base 60 si ritrovano nel fatto che un'ora è composta da 60 minuti e un minuto da 60 secondi; inoltre, nella misura degli angoli continuiamo a pensare che un angolo giro sia costituito da 360 gradi, sebbene con l'introduzione del sistema metrico si fosse tentato di uguagliarlo a 400 gradi (in modo che ogni angolo retto corrispondesse a 100 gradi).

Se i nostri progenitori usavano i numeri a fini essenzialmente pratici, esistono varie testimonianze del fatto che quelle antiche civiltà nutrissero interesse per la matematica in sé e sottraessero tempo alle attività pratiche della vita quotidiana per indagarne le caratteristiche. Le loro indagini riguardavano le proprietà delle figure geometriche e ciò che potremmo definire «algebra».

Il sistema di numerazione egizio risalente al XIII secolo a.C. era decimale e si basava su un insieme di geroglifici. Gli Egizi avevano anche sviluppato un metodo di calcolo frazionario, ma l'attuale notazione decimale posizionale è stata elaborata dai Babilonesi e perfezionata dagli indiani. Il vantaggio di un sistema di questo tipo è che consente di rappresentare allo stesso tempo numeri molto piccoli e molto grandi; inoltre le cifre indo-arabiche 1, 2, 3, 4, 5, 6, 7, 8 e 9 consentono di svolgere i calcoli con relativa facilità. Per convincercene, basta confrontare questo sistema di nume-

linea del tempo

30 000 a.C. — Le civiltà paleolitiche europee tracciano sulle ossa dei segni corrispondenti a numeri

2000 a.C. — I Babilonesi rappresentano i numeri per mezzo di simboli

razione con quello usato dai Romani: anche se rispondeva ai loro scopi, solo pochi specialisti erano in grado di usarlo per effettuare dei calcoli.

Il sistema di numerazione romano I Romani usavano una serie di simboli fondamentali per indicare le unità, le decine, le centinaia, le migliaia (*I*, X, C e M) e le loro metà (*V*, *L* e *D*), e poi li combinavano in modo da formare altri numeri. È stato ipotizzato che le cifre *I*, *II*, *III* e *IIII* derivino dalle dita di una mano e *V* dalla sua forma complessiva; l'unione di questo simbolo e di quello che si ottiene capovolgendolo dà luogo a X, che corrisponde a due mani, ovvero a dieci dita. C deriva da *centum* e M da *mille*. Inoltre i Romani usavano *S* per indicare «un mezzo» e disponevano di un sistema di frazioni in base 12.

Sistema di numerazione romano

Impero romano		aggiunte medievali	
S	un mezzo		
I	uno		
V	cinque	$\bar{V}$	cinquemila
X	dieci	$\bar{X}$	diecimila
L	cinquanta	$\bar{L}$	cinquantamila
C	cento	$\bar{C}$	centomila
D	cinquecento	$\bar{D}$	cinquecentomila
M	mille	$\bar{M}$	un milione

I Romani utilizzavano un sistema basato sull'importanza della posizione «prima e dopo» dei simboli fondamentali (per esempio IX e XI), ma pare che tale metodo non fosse uniformemente applicato. Gli antichi Romani preferivano scrivere *IIII*, mentre *IV* fu introdotto in seguito. Sembra che fosse in uso la combinazione *IX*, ma *SIX* sarebbe stato interpretato come 8½! I numeri *fondamentali* del sistema romano, con qualche aggiunta risalente al periodo medievale, sono riportati nel riquadro a fianco.

Le cifre romane non sono affatto facili da utilizzare. Ad esempio, il significato di *MMMCDXLIIII* diventa chiaro solo a condizione di introdurre idealmente delle parentesi, in modo da leggere (*MMM*)(*CD*)(*XL*)(*IIII*) come 3000 + 400 + 40 + 4 = 3444. Proviamo ora a sommare *MMMCDXLIIII* e *CCCXCIIII*. Un antico romano esperto di questa arte probabilmente conoscerebbe i trucchetti giusti, ma per noi è arduo venirne a capo se non svolgendo il calcolo nel sistema decimale e poi ritraducendo il numero ottenuto in notazione romana:

Addizione		
3444	→	*MMMCDXLIIII*
+ 394	→	*CCCXCIIII*
= 3838	→	*MMMDCCCXXXVIII*

La moltiplicazione di due numeri è ancora più difficile e all'interno del sistema di base potrebbe rivelarsi impossibile anche per un romano! Per moltiplicare 3444 × 394 bisogna ricorrere alle aggiunte medievali.

600 In India entra in uso il precursore della moderna notazione decimale

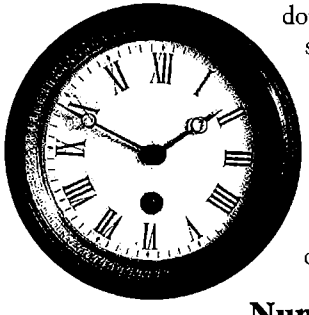
1200 Si diffonde il sistema indo-arabico che utilizza le cifre 1, ..., 9 più lo zero

1600 I simboli del sistema decimale assumono le loro forme attuali

Moltiplicazione

	3444	→	*MMMCDXLIIII*
×	394	→	*CCCXCIIII*
=	1 356 936	→	$\overline{M}\,\overline{C}\,\overline{C}\,\overline{C}\,\overline{L}\,\overline{V}MCMXXXVI$

I Romani non possedevano un simbolo specifico per indicare lo zero. Se avessimo chiesto a un cittadino romano vegetariano di annotare quante bottiglie di vino aveva bevuto nel corso della giornata avrebbe potuto scrivere III, ma se gli avessimo domandato quanti polli aveva mangiato non avrebbe scritto 0. Le vestigia del sistema numerico romano sono ancora visibili nella numerazione delle pagine di certi libri (ma non di questo) e nelle iscrizioni sulle fondamenta degli edifici. Alcune strutture, come *MCM* per indicare 1900, non sono mai state usate dai Romani ma sono state introdotte in tempi recenti per motivi stilistici (i Romani avrebbero scritto *MDCCCC*). Il quattordicesimo re di Francia di nome Luigi, ora universalmente noto come Luigi *XIV*, in realtà preferiva farsi chiamare Luigi *XIIII* e aveva ordinato che nei quadranti degli orologi le quattro fossero indicate come *IIII*.



Orologio Luigi XIII

Numeri interi decimali Per noi, i «numeri» per antonomasia sono i numeri decimali. Il sistema decimale è basato sul 10 e sulle sue potenze (decine, centinaia, migliaia e così via), e utilizza le cifre 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9. Ad esempio, è possibile esplicitare il significato decimale del numero **394**, dicendo che è costituito da **3** centinaia, **9** decine e **4** unità:

$$394 = 3 \times 100 + 9 \times 10 + 4 \times 1$$

Questa espressione può essere riscritta in termini di potenze di 10:

$$394 = 3 \times 10^2 + 9 \times 10^1 + 4 \times 10^0$$

in cui $10^2 = 10 \times 10$, $10^1 = 10$ e si assume che $10^0 = 1$. In questa forma risalta con maggiore chiarezza la base *decimale* del nostro sistema di numerazione, in cui l'addizione e la moltiplicazione risultano più trasparenti.

La virgola decimale Finora abbiamo preso in esame la rappresentazione dei numeri interi. Come se la cava il sistema decimale quando deve rappresentare una parte di un numero, come 572/1000? Ciò significa

$$\frac{572}{1000} = \frac{5}{10} + \frac{7}{100} + \frac{2}{1000}$$

Scrivendo i reciproci di 10, 100 e 1000 sotto forma di potenze *negative* di 10, si ottiene

$$\frac{572}{1000} = 5 \times 10^{-1} + 7 \times 10^{-2} + 2 \times 10^{-3}$$

che può essere scritto nella forma 0,572, dove la virgola decimale indica l'inizio delle potenze negative.

Per numeri molto grandi la notazione decimale diventa troppo lunga e si torna a utilizzare la «notazione scientifica». Ad esempio, 1 356 936 892 può essere scritto come $1{,}356936892 \times 10^9$. Sullo schermo delle calcolatrici e dei computer, in genere questo numero compare nella forma «1,356936892 × 10E9», in cui l'esponente 9 corrisponde al numero di cifre del numero meno una e la lettera E sta per «esponenziale». Può capitare di utilizzare numeri ancora più grandi, come il numero di atomi di idrogeno presenti nell'Universo noto, il cui valore è stimato approssimativamente in $1{,}7 \times 10^{77}$. La notazione scientifica è più comoda anche quando si vogliono esprimere numeri estremamente piccoli, come $1{,}7 \times 10^{-77}$ (con una potenza negativa). Non si può neanche pensare di esprimere questi numeri in simboli romani.

Serie di zero e uno Se la base 10 è moneta corrente nella vita quotidiana, alcune applicazioni richiedono l'utilizzo di basi diverse. Il sistema binario (in base 2) è il segreto della potenza dei moderni computer. Il vantaggio di questo sistema di numerazione è che consente di esprimere qualsiasi numero utilizzando solo i simboli 0 e 1; il rovescio della medaglia è che le espressioni numeriche binarie possono essere estremamente lunghe.

Come si esprime **394** in notazione binaria? Questa volta entrano in gioco le potenze di 2; dopo qualche calcolo, si può scrivere l'espressione completa nella forma

$$\mathbf{394} = 1 \times 256 + \mathbf{1} \times 128 + \mathbf{0} \times 64 + \mathbf{0} \times 32 + \mathbf{0} \times 16 + 1 \times 8 + \mathbf{0} \times 4 + \mathbf{1} \times 2 + \mathbf{0} \times 1$$

L'espressione di **394** in forma binaria si ottiene leggendo solo la sequenza di 0 e 1, ovvero **110001010**.

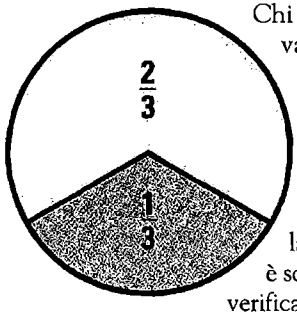
La lunghezza delle espressioni binarie fa sì che in ambito informatico si utilizzino comunemente anche altre basi, come il sistema ottale (base 8) e il sistema esadecimale (base 16). Il sistema ottale comprende solo i simboli 0, 1, 2, 3, 4, 5, 6, 7, mentre quello esadecimale fa uso di 16 simboli, che in genere sono 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F. Dato che 10 corrisponde alla lettera A, la rappresentazione esadecimale del numero **394** è 18A. È semplice come l'ABC, che peraltro corrisponde al numero decimale 2748!

Potenze di 2	Scrittura in base 10
2^0	1
2^1	2
2^2	4
2^3	8
2^4	16
2^5	32
2^6	64
2^7	128
2^8	256
2^9	512
2^{10}	1024

idea chiave
La scrittura dei numeri

03 Le frazioni

Una frazione è un «numero fratto», cioè «spezzato». Il modo migliore per spezzare un numero intero consiste nell'utilizzare una frazione. Consideriamo l'esempio classico, la famosa torta, e dividiamola in tre parti.



Chi riceve due delle tre parti della torta ottiene una frazione equivalente a 2/3; chi è meno fortunato, invece, ne otterrà solo 1/3. Combinando le due porzioni si ricostituisce la torta intera; in forma frazionaria ciò si esprime scrivendo 1/3 + 2/3 = 1, dove 1 rappresenta la totalità della torta.

Vediamo un altro esempio. È tempo di saldi e una camicia viene venduta ai quattro quinti del suo prezzo originario: in questo caso la frazione è 4/5 . Si potrebbe anche dire che il prezzo della camicia è scontato di un quinto, scrivendo lo sconto nella forma 1/5 . È facile verificare che 1/5 + 4/5 = 1, dove 1 rappresenta il prezzo di partenza.

Una frazione è sempre costituita da un numero intero «su» un altro numero intero. Il numero che sta sotto si chiama «denominatore», perché indica quante sono le parti che costituiscono l'intero, mentre quello che sta sopra è detto «numeratore», perché specifica il numero delle unità frazionarie. Quindi, nella notazione convenzionale una frazione assume sempre la forma

$$\frac{\text{numeratore}}{\text{denominatore}}$$

Nel caso della torta, la frazione che ciascuno di noi spera di mangiare è 2/3, dove 3 è il denominatore e 2 è il numeratore. 2/3 è costituita da 2 unità frazionarie, ciascuna uguale a 1/3.

Esistono anche frazioni del tipo 14/5 (dette «frazioni improprie»), in cui il numeratore è maggiore del denominatore. Dividendo 14 per 5 si ottiene 2 con un resto di 4; questo risultato può essere scritto sotto forma del «numero misto» 2 + 4/5, costituito dal numero intero 2 e dalla frazione «propria» 4/5; alcuni autori del

linea del tempo

1800 a.C. I Babilonesi usano le frazioni

1650 a.C. Gli Egizi usano frazioni unitarie

passato scrivevano le frazioni miste come 4/5 + 2. In genere le frazioni vengono rappresentate nella forma in cui il numeratore e il denominatore (la parte «di sopra» e quella «di sotto») non possiedono fattori comuni. Ad esempio, il numeratore e il denominatore di 8/10 hanno come fattore comune 2, perché 8 = 2 × 4 e 10 = 2 × 5. Nella frazione 8/10 = 2×4/2×5 si può «semplificare» il 2 ottenendo 8/10 = 4/5, una forma più semplice corrispondente allo stesso valore. In linguaggio matematico, le frazioni sono dette «numeri razionali» (dal latino *ratio* che significa «calcolo» o «rapporto») perché si ottengono dal rapporto di due numeri interi. I numeri razionali sono i numeri che i Greci erano in grado di «misurare».

Addizione e moltiplicazione Una caratteristica singolare delle frazioni è che è più facile moltiplicarle che sommarle. La moltiplicazione dei numeri interi è così complicata da aver richiesto lo sviluppo di metodi di calcolo ingegnosi; nel caso delle frazioni, invece, è l'addizione a creare maggiori difficoltà e a richiedere un po' di riflessione.

Cominciamo dalla moltiplicazione. Chi acquista una camicia per i quattro quinti del prezzo originario di 30 euro, pagherà un prezzo scontato pari a 24 euro. Infatti, immaginando di suddividere i 30 euro in cinque parti da 6 euro ciascuna e di prenderne quattro, si ottiene 4 × 6 = 24.

Quando il padrone del negozio scopre che le camicie sono rimaste per buona parte invendute, decide di abbassare ulteriormente il prezzo, portandolo a 1/2 del prezzo scontato. Ora chi entra nel negozio può acquistare la camicia per 12 euro, perché 1/2 × 4/5 × 30 = 12. Per moltiplicare due frazioni basta moltiplicare fra loro i numeratori e i denominatori:

$$\frac{1}{2} \times \frac{4}{5} = \frac{1 \times 4}{2 \times 5} = \frac{4}{10}$$

Se il negoziante avesse effettuato i due sconti in una volta sola, avrebbe annunciato che il prezzo delle camicie corrispondeva ai quattro decimi del prezzo originario di 30 euro; infatti, 4/10 × 30 = 12.

Sommare due frazioni è tutta un'altra storia. La somma 1/3 + 2/3 è immediata, perché il denominatore è lo stesso: basta sommare i numeratori per ottenere 3/3 = 1. Se però vogliamo sommare due terzi del dolce ad altri due quinti, la faccenda si complica. Come si calcola 2/3 + 4/5? Sarebbe bello poter scrivere 2/3 + 4/5 = 2 + 4/3 + 5 = 6/8, ma purtroppo non si può.

100 I cinesi concepiscono un metodo per effettuare i calcoli con le frazioni

1202 Leonardo Fibonacci diffonde l'uso della linea di separazione fra numeratore e denominatore

1585 Simon Stevin (Stevino) propone una teoria delle frazioni decimali

1700 La linea frazionaria – diventa di uso generale (a/b)

La somma tra frazioni richiede un approccio diverso. Per sommare 2/3 e 4/5, bisogna prima trasformarle in frazioni con lo *stesso* denominatore. Si moltiplicano quindi il numeratore e il denominatore di 2/3 per 5, ottenendo 10/15, e poi il numeratore e il denominatore di 4/5 per 3, ottenendo 12/15. Ora che le due frazioni hanno un denominatore comune (15), è possibile sommare i numeratori:

$$\frac{2}{3} + \frac{4}{5} = \frac{10}{15} + \frac{12}{15} = \frac{22}{15}$$

Conversione in numeri decimali In ambito scientifico e applicativo si preferisce esprimere le frazioni in forma decimale. La frazione 4/5 è equivalente a 8/10; in questo caso il denominatore è 10, perciò si può trasformare la frazione nel numero decimale 0,8.

Convertire in forma decimale le frazioni che hanno come denominatore 5 o 10 è semplice, ma come si procede nel caso di una frazione come 7/8? Basta ricordare che la divisione fra due numeri interi può produrre un risultato esatto oppure un risultato più qualcosa che rimane, il cosiddetto *resto*.

Vediamo come si trasforma una frazione in un numero decimale, prendendo come esempio 7/8:

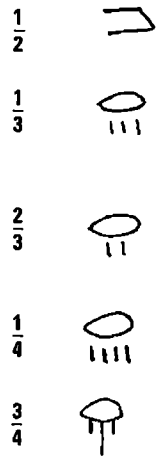
- Si guarda quante volte l'8 sta nel 7. Si può dire che «non ci sta», cioè che la divisione dà 0 con il resto di 7. Si scrive uno zero seguito dalla virgola decimale: «0,».
- Si guarda quante volte l'8 sta nel 70 (il resto del passaggio precedente moltiplicato per 10): ci sta 8 volte, perché 8 × 8 = 64, perciò il risultato è 8 con il resto di 6 (70 – 64). Si scrive 8 accanto al risultato ottenuto dal primo passaggio: «0,8».
- Si guarda quante volte l'8 sta nel 60 (il resto del passaggio precedente moltiplicato per 10). Poiché 7 × 8 = 56, si ottiene 7 con il resto di 4. Si scrive anche questo risultato di seguito al precedente «0,87».
- Si guarda quante volte l'8 sta nel 40 (il resto del passaggio precedente moltiplicato per 10). Ci sta 5 volte *precise*, con il resto di *zero*. Quando si ottiene un resto uguale a zero il procedimento è finito. Il risultato finale è «0,875».

In alcuni casi questa procedura non conduce a un risultato finale, ma può andare avanti in eterno; ad esempio, quando si cerca di trasformare 2/3 in un numero decimale si scopre che ogni passaggio implica di dividere 20 per 3, da cui si ottiene 6 con un resto di 2. A questo punto bisogna dividere un'altra volta 20 per 3, e così via, senza mai ottenere un resto nullo. In questo caso il risultato è il numero

decimale illimitato 0,66666..., che si scrive nella forma $0,\bar{6}$ per indicare che si tratta di un «numero periodico».

Esistono molte frazioni che si comportano in questo modo; una particolarmente interessante è 5/7. In questo caso si ottiene 5/7 = 0,714285714285714285..., in cui si nota il ripetersi della sequenza 714285. Le frazioni che producono una sequenza che si ripete non possono essere scritte sotto forma di un numero decimale finito, ma deve entrare in gioco la notazione «con il trattino». Nel caso di 5/7 si scrive $5/7 = 0,\overline{714285}$.

Frazioni egizie Il papiro di Rhind, conservato al British Museum di Londra, ci ha rivelato che il sistema frazionario degli Egizi del secondo millennio a.C. era basato su una serie di geroglifici che rappresentavano frazioni *unitarie* (il cui numeratore è uguale a 1). Il sistema frazionario egizio era talmente complicato che solo pochi iniziati ne conoscevano i segreti ed erano in grado di svolgere correttamente i calcoli.



Frazioni egizie

Gli Egizi esprimevano tutte le frazioni, a parte qualche caso privilegiato come 2/3, in termini di frazioni unitarie, come 1/2, 1/3, 1/11 o 1/168: queste erano le «frazioni di base» da cui era possibile ricavare tutte le altre. Ad esempio, 5/7 non è unitaria, ma può essere espressa in termini di frazioni di quel tipo:

$$\frac{5}{7} = \frac{1}{3} + \frac{1}{4} + \frac{1}{8} + \frac{1}{168}$$

in cui le frazioni unitarie devono essere tutte diverse. Una caratteristica di questo metodo è che in certi casi una frazione può essere espressa in modi diversi, alcuni dei quali più brevi degli altri. Ad esempio

$$\frac{5}{7} = \frac{1}{2} + \frac{1}{7} + \frac{1}{14}$$

L'utilità pratica dell'«espansione egizia» è limitata, ma si tratta di un metodo che ha ispirato intere generazioni di matematici puri e ha prodotto molti stimolanti problemi, alcuni dei quali sono tuttora insoluti. Ad esempio, l'analisi dei metodi che consentono di trovare le espansioni egizie *più brevi* resta ancora per buona parte inesplorata.

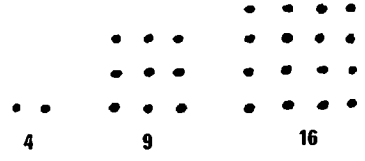
idea chiave

sopra un

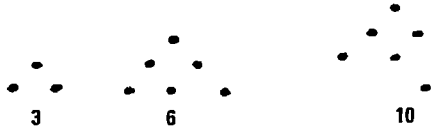
04 I quadrati e le radici quadrate

Chi si diverte a comporre quadrati con una serie di punti ragiona in modo analogo ai Pitagorici. Infatti questa attività era tenuta in gran conto dai seguaci di Pitagora, l'inventore dell'omonimo teorema. Pitagora nacque sull'isola greca di Samo e creò una setta segreta di carattere religioso che si diffuse nell'Italia meridionale. I Pitagorici erano convinti che la matematica permettesse di comprendere la natura dell'Universo.

Se si contano i punti che compongono i «quadrati» riportati a margine, si vede che il primo è costituito da un solo punto. Per i Pitagorici l'uno era il numero più importante ed era impregnato di un'essenza spirituale. È un buon inizio. Continuando a contare i punti dei quadrati successivi, si ottengono i «numeri quadrati» 1, 4, 9, 16, 25, 36, 49, 64,... detti anche «quadrati perfetti». Per calcolare un numero quadrato basta sommare il numero di punti della forma esterna ⌐ a quello del quadrato precedente; ad esempio, 9 + 7 = 16. I Pitagorici non si limitarono ai quadrati, ma presero in considerazione anche altre forme, come triangoli, pentagoni (figure con cinque lati) e altri poligoni (figure con molti lati).



I numeri triangolari assomigliano a tanti mucchietti di sassi; contando i punti che li costituiscono si ottiene 1, 3, 6, 10, 15, 21, 28, 36... Per calcolare un numero triangolare basta sommare a quello precedente il numero dei punti della riga aggiuntiva. Ad esempio, qual è il numero triangolare che viene dopo il 10? La sua ultima riga sarà costituita da 5 punti, perciò bisogna sommare 10 + 5 = 15.



Confrontando i numeri quadrati e quelli triangolari si vede che il numero 36 compare in entrambi gli elenchi. Ma esiste un collegamento più importante. Se si

linea del tempo

1750 a.C.
I Babilonesi stilano tabelle di radici quadrate

525 a.C.
I Pitagorici studiano le disposizioni geometriche corrispondenti a numeri quadrati perfetti

circa 300 a.C.
Il quinto libro degli *Elementi* di Euclide contiene la teoria dei numeri irrazionali di Eudosso

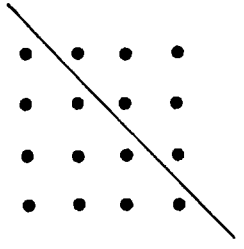
sommano due numeri triangolari *successivi*, che cosa si ottiene? Proviamo a farlo, riportando i risultati in una tabella.

Somma di due numeri triangolari successivi

1 + 3	4
3 + 6	9
6 + 10	16
10 + 15	25
15 + 21	36
21 + 28	49
28 + 36	64

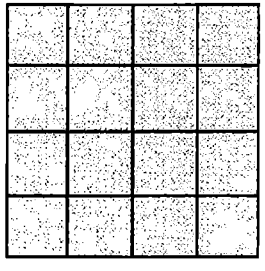
Esatto! Sommando due numeri triangolari successivi si ottiene un numero quadrato, come si può verificare con una «dimostrazione visiva». Consideriamo un quadrato costituito da 4 righe di 4 punti in cui è tracciata una diagonale: i punti situati sopra la retta (vedi figura a lato) formano un numero triangolare, quelli posti al di sotto formano il successivo; ciò è vero per ogni quadrato, indipendentemente dalle sue dimensioni. Da questi «diagrammi a punti» alla misura delle aree il passo è breve. L'area di un quadrato di lato 4 è uguale a $4 \times 4 = 4^2 = 16$ unità quadratiche. Generalizzando, se si indica il lato con x l'area è data da x^2.

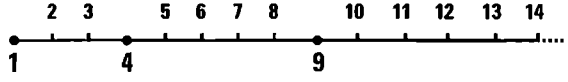
La rappresentazione grafica della funzione x^2 è una curva dalla forma parabolica che caratterizza, ad esempio, le antenne satellitari e gli specchi riflettori dei fari delle automobili. Una parabola è definita rispetto a un punto specifico, detto «fuoco». In un'antenna parabolica, i raggi paralleli del segnale proveniente dallo spazio colpiscono la superficie curva e vengono fatti convergere nel fuoco, dove si trova un sensore in grado di rivelarli.



Nel faro di un'automobile, invece, i raggi luminosi emessi da una lampadina posta nel fuoco del riflettore escono dallo specchio sotto forma di fascio di luce parallelo. La forma parabolica si ritrova anche in ambito sportivo: come ben sanno gli atleti che lanciano il peso, il giavellotto e il martello, ogni oggetto scagliato in aria ricade sulla Terra seguendo una traiettoria parabolica.

Radici quadrate Ribaltiamo il problema. Se si parte da un quadrato di area 16 e si vuole ricavare la lunghezza del lato, è facile intuire che il risultato è 4. Si dice allora che la «radice quadrata» di 16 è 4 e si scrive $\sqrt{16} = 4$. L'uso del simbolo $\sqrt{\ }$ per indicare la radice quadrata risale al 1500. La radice quadrata di un numero quadrato è un numero intero, come si può verificare considerando $\sqrt{1} = 1$, $\sqrt{4} = 2$, $\sqrt{9} = 3$, $\sqrt{16} = 4$, $\sqrt{25} = 5$ e così via. Sulla retta dei numeri, però, i quadrati perfetti sono intervallati da altri numeri: 2, 3, 5, 6, 7, 8, 10, 11,...





Per le radici quadrate esiste una splendida notazione alternativa. Se x^2 rappresenta un numero quadrato, la radice quadrata di un numero può essere scritta come $x^{1/2}$, una forma che si rivela utilissima quando si vuole ricorrere al trucco di moltiplicare dei numeri sommando i loro esponenti. È il concetto che sta alla base del logaritmo,

630
Brahmagupta elabora dei metodi per il calcolo delle radici quadrate

1550
Viene introdotto il simbolo $\sqrt{\ }$ per le radici quadrate

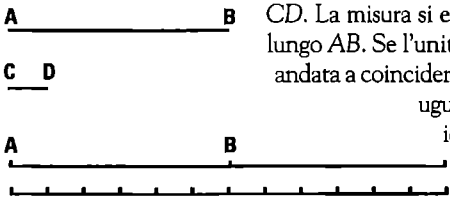
1872
Richard Dedekind elabora una teoria dei numeri irrazionali

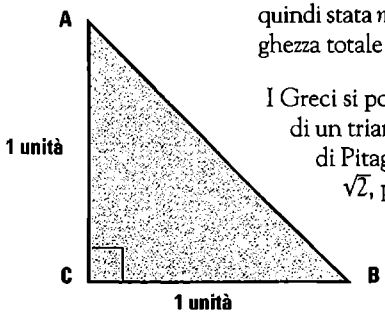
che fu inventato dopo la scoperta (avvenuta verso il 1600) della possibilità di trasformare un problema moltiplicativo in uno additivo. Ma questa è un'altra storia. Anche gli altri numeri presenti sulla retta hanno una radice quadrata, ma non è un numero intero. Quasi tutte le calcolatrici hanno un tasto con il simbolo $\sqrt{\ }$ che consente di calcolare, ad esempio, $\sqrt{7} = 2{,}645751311$.

Consideriamo $\sqrt{2}$. I Pitagorici attribuivano un significato particolare al numero 2, in virtù del fatto che è il primo numero pari (i Greci pensavano che i numeri pari fossero femminili, quelli dispari maschili e che i numeri piccoli fossero dotati di una personalità ben definita). Utilizzando una calcolatrice con un numero sufficiente di cifre decimali, si trova che $\sqrt{2}$ è uguale a 1,414213562. Ciò significa che questo numero è la radice quadrata di 2? Per verificarlo calcoliamo 1,414213562 × 1,414213562. Il risultato è 1,999999999. Non è esattamente uguale a 2, perché 1,414213562 è solo il valore approssimato della radice quadrata di 2.

Va sottolineato che un'approssimazione è il massimo che si possa ottenere: anche un'espansione decimale di $\sqrt{2}$ con milioni di cifre decimali resterà sempre e comunque un'approssimazione. Il numero $\sqrt{2}$ riveste un ruolo rilevante in matematica; non sarà illustre come π o e (vedi capitoli 5 e 6), ma è abbastanza importante da aver meritato un nome proprio, ed è anche detto «costante di Pitagora».

Le radici quadrate sono frazioni? Per rispondere a questa domanda bisogna rifarsi alla teoria della misura degli antichi Greci. Supponiamo di avere un segmento *AB* di cui si vuole misurare la lunghezza e un'«unità di misura» indivisibile *CD*. La misura si effettua riportando più volte consecutivamente l'unità di misura *CD* lungo *AB*. Se l'unità di misura è stata ripetuta *m* volte e l'ultima volta la sua estremità è andata a coincidere esattamente con quella di *AB* (nel punto B), la lunghezza di *AB* sarà uguale a *m*. In caso contrario si può prolungare *AB* con una sua copia identica e proseguire l'operazione di misura (vedi figura). I Greci pensavano che a un certo punto, usando *n* copie di *AB* e ripetendo *m* volte l'unità di misura, questa sarebbe andata a coincidere *esattamente* con il punto finale dell'*n*-esimo *AB*. La lunghezza di *AB* sarebbe quindi stata *m*/*n*. Ad esempio, se si ripetono consecutivamente 3 copie di *AB* e sulla lunghezza totale l'unità di misura viene riportata 29 volte, la lunghezza di *AB* sarà 29/3.





I Greci si posero anche il problema di misurare la lunghezza del lato *AB* (ipotenusa) di un triangolo in cui ciascuno degli altri due lati è lungo 1 «unità». Per il teorema di Pitagora, la lunghezza di *AB* potrebbe essere espressa in forma simbolica come $\sqrt{2}$, perciò il problema diventa quello di stabilire se $\sqrt{2} = m/n$.

La calcolatrice ci ha già detto che l'espressione decimale per $\sqrt{2}$ è potenzialmente infinita e ciò (il fatto che la sua espressione decimale non finisca mai) sembra indicare che $\sqrt{2}$ non sia una

frazione. D'altra parte, anche il numero decimale 0,3333333... non finisce mai, però rappresenta la frazione 1/3. Bisogna trovare un argomento più convincente.

$\sqrt{2}$ è una frazione? Questa domanda conduce a una delle dimostrazioni più famose della matematica, che si sviluppa secondo il ragionamento prediletto dai Greci: il metodo della *reductio ad absurdum*, o «dimostrazione per assurdo». In primo luogo si assume che $\sqrt{2}$ non possa essere contemporaneamente una frazione e «una non frazione». Si tratta del cosiddetto «principio del terzo escluso», una legge logica che non ammette la possibilità dell'esistenza di una via di mezzo fra due alternative. I Greci svilupparono un procedimento ingegnoso. Partirono dall'assunzione che $\sqrt{2}$ fosse una frazione e, mediante una serie di passaggi rigorosamente logici, giunsero a una contraddizione, un'«assurdità». Vediamo in che modo. Supponiamo che

$$\sqrt{2} = \frac{m}{n}$$

Si può supporre anche qualcos'altro: si può supporre che m e n non abbiano fattori comuni. Ciò è senz'altro vero perché, se li avessero, potremmo semplificarli preliminarmente (ad esempio, semplificando il fattore comune 7 la frazione 21/35 diventa uguale a 3/5, che è ridotta ai minimi termini).

È possibile elevare al quadrato entrambi i membri dell'uguaglianza $\sqrt{2} = m/n$, ottenendo $2 = m^2/n^2$ e quindi $m^2 = 2n^2$. A questo punto si può fare una prima osservazione: se m^2 è uguale a 2 moltiplicato per qualcosa, deve essere un numero pari. Quindi neanche m può essere dispari (perché il quadrato di un numero dispari è dispari), ma deve necessariamente essere pari.

Per ora la logica del ragionamento è impeccabile. Se m è pari, deve essere il doppio di qualcosa, perciò può essere scritto nella forma $m = 2k$. Elevando al quadrato entrambi i membri si ottiene $m^2 = 4k^2$. Combinando questo risultato con $m^2 = 2n^2$ si trova che $2n^2 = 4k^2$ e, semplificando il 2, si arriva alla conclusione che $n^2 = 2k^2$. Siamo tornati al punto di partenza! Anche in questo caso, concludiamo che n^2 è pari e che lo è anche n. Abbiamo così dedotto per via puramente logica che sia m che n sono pari e quindi hanno un fattore 2 in comune. Ma ciò contraddice l'assunzione iniziale che m e n non avessero fattori comuni: la conclusione è che $\sqrt{2}$ non può essere una frazione.

Si può anche dimostrare che nessun numero del tipo $\sqrt{n}$ (eccetto i casi in cui n è un quadrato perfetto) può essere una frazione. I numeri non esprimibili in forma frazionaria sono detti «numeri irrazionali»; abbiamo quindi constatato che esistono infiniti numeri irrazionali.

idea chiave

Verso i numeri irrazionali

05 π

π è il numero più famoso della matematica. Per le altre costanti naturali non c'è speranza: π sarà sempre il primo della lista. Se esistesse l'Oscar per i numeri, π lo vincerebbe ogni anno.

π, o «pi greco», è uguale alla lunghezza esterna di un cerchio (la circonferenza) divisa per la lunghezza della corda che passa per il centro (il diametro). Il suo valore, cioè il rapporto fra queste due lunghezze, non dipende dalle dimensioni del cerchio: quest'ultimo può essere grande o piccolo, ma π resta costante. Sebbene il cerchio sia il suo habitat naturale, π compare un po' ovunque in matematica, anche in luoghi che con il cerchio non hanno niente a che fare.

Per un cerchio di diametro *d* e raggio *r*:

circonferenza = $\pi d = 2\pi r$

area = πr^2

Per una sfera di diametro *d* e raggio *r*:

superficie = $\pi d^2 = 4\pi r^2$

volume = $\frac{4}{3}\pi r^3$

Archimede di Siracusa Il rapporto fra la circonferenza e il diametro di un cerchio ha suscitato un grande interesse fin dall'antichità. Verso il 2000 a.C. i Babilonesi osservarono che la lunghezza della circonferenza era circa uguale al triplo del diametro.

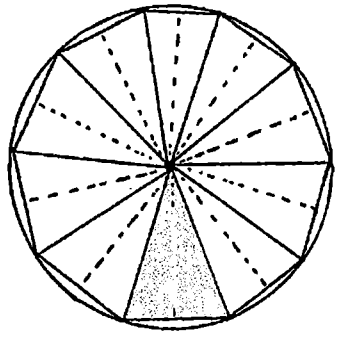
La teoria matematica di π ha avuto inizio intorno al 225 a.C. a opera di Archimede di Siracusa, uno dei grandi protagonisti della matematica. I matematici, che amano stilare graduatorie, lo collocano alla pari di Carl Friedrich Gauss (detto «il principe dei matematici») e di Sir Isaac Newton. Senza entrare nel merito del giudizio, è chiaro che Archimede troverebbe posto in qualunque Pantheon della matematica. Non bisogna pensare che fosse rinchiuso in una torre d'avorio: oltre a fornire importanti contributi all'astronomia, alla matematica e alla fisica, Archimede progettò varie macchine da guerra, come catapulte, leve e «specchi ustori», per respingere gli attacchi dei Romani. Tutti i resoconti, però, concordano nel dire che, da buon scienziato, aveva la testa fra le nuvole; solo così si può spiegare il fatto che, dopo aver scoperto la legge idrostatica del galleggiamento, sia saltato fuori dalla vasca da bagno e si sia messo a correre completamente nudo per strada gridando «Eureka»! In che modo abbia festeggiato i risultati ottenuti su π non ci è stato tramandato.

linea del tempo

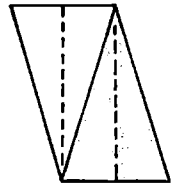
2000 a.C. I Babilonesi osservano che π è circa uguale a 3

225 a.C. Archimede approssima π in modo soddisfacente con 22/7

Se π è definito come il rapporto fra la circonferenza e il corrispondente diametro, cosa c'entra con l'*area* di un cerchio? Il fatto che l'area di un cerchio di raggio r sia uguale a πr^2 in realtà è solo una deduzione, anche se probabilmente è più nota della definizione di π come rapporto fra circonferenza e diametro. È comunque notevole che π sia coinvolto sia nel calcolo dell'area che della circonferenza del cerchio.



Vediamo come si dimostra la formula dell'area del cerchio. Si suddivide il cerchio in un gran numero di triangolini tutti uguali, di base b e di altezza pari al raggio r; i triangoli formano un poligono che è inscritto nel cerchio e la cui area approssima quella del cerchio. Si parte con 1000 triangoli e si procede con una serie di approssimazioni successive. Ogni coppia di triangoli adiacenti può essere unita in modo da formare un rettangolo di area (circa) uguale a $b \times r$, cosicché l'area totale del poligono sarà data da $500 \times b \times r$. Ma $500 \times b$ corrisponde approssimativamente a metà della circonferenza, cioè a πr, perciò l'area del poligono è $\pi r \times r = \pi r^2$. Maggiore è il numero di triangoli considerati, migliore sarà l'approssimazione; passando al limite, si può concludere che l'area del cerchio è πr^2.



Archimede stimava che il valore di π fosse compreso fra 223/71 e 220/70, perciò è a lui che si deve l'approssimazione di questa costante con il consueto valore di 22/7. L'onore di aver introdotto il simbolo π spetta invece al molto meno noto William Jones, un matematico gallese che nel XVIII secolo fu vicepresidente della Royal Society di Londra. Il matematico e fisico Leonhard Euler (Eulero) rese noto che π corrisponde al rapporto fra circonferenza e diametro.

Il valore esatto di π Non è possibile conoscere il valore *esatto* di π perché (come dimostrò Johann Lambert nel 1768) π è un numero irrazionale; il suo sviluppo decimale è infinito e non presenta una regolarità prevedibile. Le prime 20 cifre decimali sono 3,14159265358979323846... I matematici cinesi lo approssimavano con il valore $\sqrt{10}$ = 3,16227766016837933199, che intorno all'anno 500 fu adottato da Brahmagupta. In realtà questo valore funziona poco meglio di un semplice 3 e differisce da π già a partire dalla seconda cifra decimale.

π può essere calcolato a partire da una *serie* di numeri. La serie

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \ldots$$

1706 William Jones introduce il simbolo π

1761 Lambert dimostra che π è irrazionale

1882 Lindemann dimostra che π è trascendente

converge molto lentamente al valore di π e quindi richiede lunghi e faticosi calcoli. Una serie molto interessante che converge al valore di π fu trovata da Eulero:

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \ldots$$

Srinivasa Ramanujan, un geniale autodidatta, concepì alcune formule spettacolari in grado di approssimare π, fra cui anche una in cui compare solo la radice quadrata di 2:

$$\frac{9801}{4412}\sqrt{2} = 3.14159273001330566031399618 90\ldots$$

I matematici sono sempre stati affascinati da π. Se Lambert ha dimostrato che non corrispondeva a una frazione, nel 1882 il matematico tedesco Ferdinand von Lindemann ha risolto il problema principale associato a questa costante dimostrando che π è «trascendente», cioè che non può costituire la soluzione di un'equazione algebrica (un'equazione in cui compaiono solo potenze intere di x). La soluzione di questo enigma dette anche risposta al problema della «quadratura del cerchio», cioè della costruzione di un quadrato con la stessa area di un cerchio dato (usando solo una riga e un compasso). Lindemann dimostrò in modo conclusivo che ciò non era possibile e oggi l'espressione «quadratura del cerchio» è passata a indicare un compito impossibile da realizzare.

Il calcolo di π fu portato avanti a ritmo serrato. Nel 1853 William Shanks affermò di averne trovato il valore corretto fino alla 607ª cifra decimale (in realtà lo era solo fino alla 527ª). Più recentemente, la caccia a un numero sempre maggiore di cifre decimali di π ha acquistato maggiore impulso grazie all'avvento dei computer. Nel 1949, dopo un calcolo di 70 ore su un computer ENIAC, il valore di π fu calcolato fino alla 2037ª cifra decimale. Nel 2002 il numero di cifre decimali note è stato portato a 1 241 100 000 000, un risultato davvero impressionante, ma la «coda» continua ad allungarsi. Se ci trovassimo all'equatore e iniziassimo a scrivere l'espansione di π, il calcolo di Shanks occuperebbe 14 metri, mentre l'espansione del 2002 farebbe 62 volte il giro del pianeta!

Riguardo a π sono state poste varie domande e sono state date varie risposte. È costituito da cifre casuali? È possibile individuare una sequenza predeterminata nella sua espansione? Si può pensare di trovarvi, ad esempio, la sequenza 0123456789? Negli anni '50 del Novecento questa domanda sembrava destinata a non trovare risposta. Nessuno aveva individuato una sequenza di questo tipo nelle 2000 cifre allora note di π. L'eminente matematico olandese L.E.J. Brouwer affermò che si trattava di una domanda priva di senso, perché riteneva che fosse impossibile fornirne una dimostrazione sperimentale. In realtà una sequenza di cifre fu scoperta nel 1997 a partire dalla 17 387 594 880ª cifra decimale, ovvero, nella metafora dell'equatore, circa 4800 km prima di aver compiuto un giro. Dieci sei di fila

π in poesia

Per ricordare le prime cifre dell'espansione di π può essere d'aiuto una poesia. Seguendo la tradizione di insegnare matematica in modo mnemonico, Michael Keith ha prodotto un'eccellente variante de *Il corvo* di Edgar Allan Poe.

La poesia originale di Poe inizia così:

The Raven

Once upon a midnight dreary,
while I pondered weak and weary,
Over many a quaint and curious volume
of forgotten lore

La variante di Keith inizia così:

Poe, E.: Near a Raven

Midnights so dreary, tired and weary.
Silently pondering volumes extolling
all by-now obsolete lore

Contando il numero delle lettere che compongono le parole della versione di Keith si ottengono le prime 740 cifre di π. In italiano, un esempio potrebbe essere:

Noi e loro, a volte / bisognamo di notare / cifre fra molte, / affinché calcolare / possiam lunghezze. / Con il mio versetto / quel numero si arreca. / Dici che son prodezze? / Fai un risetto, / letterina greca!

compaiono entro i primi 1000 km, mentre per trovare dieci sette consecutivi bisogna aspettare di aver effettuato un giro intero, più altri 5800 km.

L'importanza di π A che serve conoscere tutte queste cifre decimali di π? Dopo tutto, per la maggior parte dei calcoli ne bastano molte meno; probabilmente nessuna applicazione pratica ne richiede più di dieci e nella maggior parte dei casi è sufficiente l'approssimazione di Archimede 22/7. Questi calcoli lunghi e laboriosi, però, non vengono effettuati per puro divertimento. Sono utili per verificare i limiti della capacità di calcolo dei computer, oltre a esercitare un grande fascino sul gruppo di matematici che si definiscono «amici di pi greco».

Forse l'episodio più strano nella storia di π è stato il tentativo effettuato dallo stato dell'Indiana (Stati Uniti) di fissarne il valore per legge. Accadde alla fine del XIX secolo, quando un medico di nome E.J. Goodwin propose una legge di questo tipo per rendere π «meno indigesto». Sorse però un problema pratico: il promotore della legge non era in grado di stabilire quale fosse il valore da adottare. Fortunatamente per la reputazione dell'Indiana, prima dell'approvazione definitiva della legge apparve evidente a tutti quanto fosse folle l'idea di legiferare sul valore di π: da allora, i politici hanno sempre girato alla larga da π.

idea chiave

Un π sempre più lungo

06 e

Rispetto al suo unico rivale π, il numero *e* è un vero pivello. Se il primo ha un glorioso passato risalente ai Babilonesi, *e* è molto meno appesantito dalle incrostazioni della storia. Questa costante giovane e vivace è sempre presente quando è in ballo una «crescita»: che si tratti di popolazione, denaro o grandezze fisiche, se qualcosa aumenta, invariabilmente salta fuori *e*.

e è il numero il cui valore approssimato è 2,71828. Cos'ha di tanto speciale? Non si tratta di un numero pescato a caso, ma di una delle costanti più importanti della matematica che ha fatto la sua comparsa all'inizio del XVII secolo, quando molti matematici si dedicarono a chiarire il concetto del logaritmo (un'invenzione geniale che consentiva di trasformare la moltiplicazione di numeri molto grandi in un'addizione).

In realtà la storia inizia con un «*e-business*» del XVII secolo. Jacob Bernoulli, un membro dell'illustre famiglia svizzera specializzata nel fornire al mondo una stirpe di matematici, nel 1683 affrontò il problema dell'interesse composto.

Soldi, soldi e ancora soldi Consideriamo un periodo di un anno, uno stratosferico tasso di interesse del 100% e un deposito iniziale di 1 euro. Nella vita reale difficilmente riusciremo a spuntare un tasso del 100%, ma si tratta di un valore comodo per i nostri scopi e lo stesso concetto vale anche per tassi d'interessi più realistici, come il 6% o il 7%. Se invece si parte da un capitale maggiore, ad esempio 10 000 euro, basta moltiplicare il tutto per 10 000.

Al termine dell'anno con un tasso del 100%, incassiamo la somma iniziale più gli interessi maturati, in questo caso 1 euro, e ci ritroviamo con una somma principesca: 2 euro. Supponiamo ora che il tasso di interesse sia la metà (al 50%), ma che venga applicato separatamente a ciascun semestre. Per la prima metà dell'anno si guadagnano 50 centesimi di interessi, perciò alla fine di questo periodo il capitale sarà passato a 1,50 euro. Alla fine dell'anno potremo incassare questa somma più i relativi interessi, che ammontano a 75 centesimi: in un anno l'euro iniziale si è trasformato in 2,25 euro! Calcolando gli interessi su base semestrale abbiamo guadagnato 25

linea del tempo

1618 — Mentre studia i logaritmi, John Napier (Nepero) si imbatte in una costante, *e*

1727 — Eulero usa la notazione *e* in relazione alla teoria dei logaritmi. *e* viene anche chiamata «numero di Eulero»

centesimi in più. Può sembrare una sciocchezza, ma se avessimo investito 10 000 euro avremmo ricavato 2250 euro di interessi invece di 2000; calcolando separatamente le due metà dell'anno, il guadagno sarebbe aumentato di 250 euro.

Attenzione, però: facendo i calcoli su base semestrale si guadagna di più, ma anche la banca guadagna di più sul denaro che le devono i clienti! Supponiamo ora che l'anno sia suddiviso in quattro trimestri e che a ciascuno di essi venga applicato un tasso di interesse del 25%. Con un calcolo analogo, si trova che in questo caso l'euro iniziale si è trasformato in 2,44141 euro. Il capitale aumenta e, disponendo di 10 000 euro, sembrerebbe conveniente suddividere l'anno in tante parti e applicare tassi di interesse minori per intervalli temporali minori.

Potremmo pensare che il denaro aumenti indefinitamente, rendendoci milionari, ma purtroppo non è così. Suddividendo l'anno in periodi sempre più brevi (vedi tabella a lato), si vede che «al limite» la quantità di denaro si stabilizza su un valore costante. L'unica base su cui realisticamente si possono capitalizzare gli interessi è quella giornaliera (ed è ciò che fanno le banche). Il significato matematico di questo limite, che i matematici chiamano *e*, è che rappresenta la quantità fino a cui cresce 1 euro se la capitalizzazione avviene in modo continuo. È un bene o un male? Conosciamo già la risposta: per i risparmiatori è un bene, per i debitori è un male.

Capitalizzazione ogni...	Interessi maturati
anno	€ 2,00000
sei mesi	€ 2,25000
tre mesi	€ 2,44141
mese	€ 2,61304
settimana	€ 2,69260
giorno	€ 2,71457
ora	€ 2,71813
minuto	€ 2,71828
secondo	€ 2,71828

Il valore esatto di *e* Anche *e*, come π, è un numero irrazionale, perciò anche in questo caso non è possibile conoscere il valore esatto della costante. Il valore limitato alle prime venti cifre decimali è 2,71828182845904523536...

Usando solo le frazioni e limitandosi a quelle in cui il numeratore e il denominatore sono numeri a due cifre, la migliore approssimazione del valore di *e* è 87/32. Stranamente, se numeratore e denominatore sono numeri a tre cifre la frazione migliore è 878/323, una sorta di estensione palindroma della frazione precedente (la matematica riserva spesso sorprese di questo tipo!). Una nota espansione in serie di *e* è

$$e = 1 + \frac{1}{1} + \frac{1}{2\times1} + \frac{1}{3\times2\times1} + \frac{1}{4\times3\times2\times1} + \frac{1}{5\times4\times3\times2\times1} + \ldots$$

In questo caso conviene usare la notazione fattoriale con il punto esclamativo, per cui, ad esempio, $5! = 5 \times 4 \times 3 \times 2 \times 1$. In questa notazione *e* assume un aspetto più familiare:

1748
Eulero calcola *e* fino alle prime 23 cifre decimali. A lui è attribuita anche la scoperta della celebre formula $e^{i\pi} + 1 = 0$, più o meno in questo periodo

1873
Hermite dimostra che *e* è un numero trascendente

2007
Vengono calcolate le prime 10^{11} cifre di *e*

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \ldots$$

Quindi, il numero *e* sembra mostrare una certa regolarità: dal punto di vista delle proprietà matematiche, appare più «simmetrico» di π.

Per ricordare le prime cifre, possiamo tentare con: «Ai modesti o vanitosi, ai violenti o timorosi do, cantando gaio ritmo, logaritmo...» dove il numero delle lettere di ciascuna parola ci dà una cifra di *e*. Chi conosce la storia americana può anche ricordare che *e* è «2,7 Andrew Jackson Andrew Jackson», perché Andrew Jackson, il settimo presidente degli Stati Uniti, fu eletto nel 1828. Esistono molti altri trucchetti per ricordare il valore di *e*, ma sono interessanti a puro titolo di curiosità, non certo dal punto di vista matematico.

Il carattere irrazionale di *e* (il fatto che non è equivalente a una frazione) fu dimostrato da Eulero nel 1737. Nel 1840 il matematico francese Joseph Liouville dimostrò che *e* non costituiva la soluzione di nessuna equazione quadratica e nel 1873, in un'opera pionieristica, il suo connazionale Charles Hermite provò che è trascendente (non può essere la soluzione di *nessuna* equazione algebrica, sempre a coefficienti interi). L'elemento più importante di questa dimostrazione era il metodo usato da Hermite; nove anni dopo Ferdinand von Lindemann lo utilizzò, con le dovute modifiche, per dimostrare che π era trascendente, un problema di profilo molto più alto.

Appena fu data risposta a questa domanda, ne sorsero altre. Anche *e* elevato a *e* è trascendente? Dev'esserlo per forza: è un'espressione così strana! Però una dimostrazione rigorosa del fatto che lo sia non è ancora stata fornita e quindi, in base ai rigidi criteri della matematica, questa affermazione deve essere classificata come una congettura. Si sono fatti dei progressi ed è stato dimostrato che è impossibile che *e* ed e^{e^2} siano entrambi trascendenti. Ci siamo quasi, ma manca l'ultimo passo.

Fra π ed *e* esistono collegamenti affascinanti. I valori e^{π} e π^{e} sono molto vicini, ma è facile dimostrare (senza bisogno di calcolarli) che $e^{\pi} > \pi^{e}$. Se «bariamo» sbirciando la calcolatrice, vediamo che i valori approssimati sono $e^{\pi} = 23{,}14069$ e $\pi^{e} = 22{,}45916$.

Il numero e^{π} è detto «costante di Gelfond» (dal matematico russo Aleksandr Gelfond) ed è stato dimostrato che è trascendente, mentre di π^{e} si sa molto meno; non è ancora stato dimostrato che è irrazionale (ammesso che lo sia).

Perché *e* è importante? Il numero *e* compare principalmente nei problemi in cui c'è qualcosa che aumenta (come una crescita economica o un au-

mento di popolazione); a questo andamento sono collegate anche le curve dipendenti da *e* che descrivono il decadimento radioattivo.

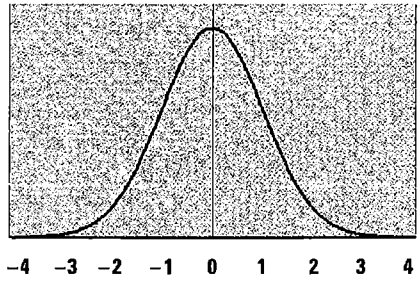
D'altra parte, *e* compare anche in problemi che non riguardano la crescita. Nel XVIII secolo Pierre Montmort ha trattato un problema di probabilità che è stato oggetto di molti studi successivi. Nella sua versione più semplice, l'enunciato del problema dice che alcune persone vanno a pranzo insieme e al momento di uscire ciascuno prende un cappello a caso. Qual è la probabilità che a nessuno tocchi il proprio cappello?

Si può dimostrare che questa probabilità è $1/e$ (circa il 37%), cosicché la probabilità che almeno una persona prenda il suo cappello è $1 - 1/e$ (63%). Questo esempio costituisce solo una delle mille applicazioni della teoria della probabilità; un'altra è la distribuzione di Poisson, che vale per gli eventi rari. Questi primissimi esempi non sono casi isolati. James Stirling ha ricavato un'ottima approssimazione del valore $n!$ in cui compaiono e e π; in ambito statistico *e* compare nella famosa «curva a campana», o gaussiana, della distribuzione normale, mentre in ingegneria è noto che la curva del cavo di un ponte sospeso dipende da *e*. L'elenco potrebbe continuare all'infinito.

Un'identità sconvolgente *e* compare anche nella formula più importante della matematica. Quando pensiamo ai numeri più famosi pensiamo a 0, 1, π ed *e*, oltre che al numero immaginario $i = \sqrt{-1}$. Sembra impossibile che valga la relazione

$$e^{i\pi} + 1 = 0$$

eppure è proprio così! Si tratta di un risultato attribuito a Eulero.



La distribuzione normale

Forse l'importanza di *e* consiste proprio nell'aura di mistero con cui ha affascinato generazioni di matematici. Di fatto, *e* è imprescindibile. Ci si può chiedere perché E.V. Wright si sia preso la briga di scrivere un romanzo privo della lettera *e*, ma con *Gadsby* è riuscito nel suo intento. È molto più difficile immaginare che un matematico decida di scrivere un manuale senza *e* o che, anche volendo, sia in grado di farlo.

idea chiave

Il numero naturale

07 L'infinito

Quanto è grande l'infinito? In sintesi, si può dire che ∞ (il simbolo dell'infinito) è grandissimo. Si può immaginare una retta su cui giacciono numeri sempre più grandi e che si estende «all'infinito»: per ogni numero, anche gigantesco come 10^{1000}, ne esiste sempre uno più grande, come $10^{1000} + 1$.

È il concetto tradizionale dell'infinito: una serie di numeri che procede indefinitamente. I matematici usano l'infinito in tutti i modi possibili e immaginabili, ma è bene essere cauti nel trattarlo come un numero qualsiasi, perché non lo è.

Contare Il matematico tedesco Georg Cantor ha proposto un concetto di infinito completamente diverso, mentre creava da solo una teoria che ha dato impulso a gran parte della matematica moderna. La teoria di Cantor deriva da un'idea collegata alla forma primitiva del concetto del contare, una nozione più semplice di quella che usiamo quotidianamente.

Immaginiamo che un allevatore non sappia contare con i numeri. Come fa a sapere quante pecore ha? Semplice! Al mattino le fa uscire e la sera capisce se sono tornate tutte o meno associando a ogni animale che rientra all'ovile un sasso preso da un mucchio che tiene accanto all'entrata. Se manca una pecora, avanzerà un sasso. Anche se non ricorre ai numeri, l'allevatore procede in modo matematico: sfrutta l'idea della corrispondenza biunivoca fra pecore e sassi. Da questo concetto primitivo derivano conseguenze sorprendenti.

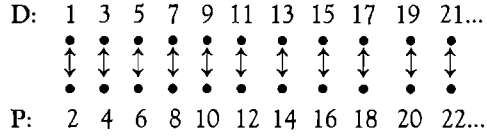
La teoria di Cantor riguarda gli *insiemi* (un insieme è una collezione di oggetti). Ad esempio, $\mathbf{N} = \{1, 2, 3, 4, 5, 6, 7, 8, \ldots\}$ corrisponde all'insieme dei numeri interi (positivi). Se si ha un insieme si possono definire dei sottoinsiemi, che sono insiemi più piccoli contenuti in quello più grande. I sottoinsiemi più ovvi dell'insieme **N** scelto come esempio sono $D = \{1, 3, 5, 7, \ldots\}$ e $P = \{2, 4, 6, 8, \ldots\}$, ovvero, rispettivamente, l'insieme dei numeri dispari e quello dei numeri pari. Ci si può domandare se esistano tanti numeri pari quanti numeri dispari. Sebbene non sia possibile contare gli elementi di ciascun insieme e confrontare i risultati ottenuti, la risposta a questa domanda è certamente «sì». Su cosa si basa tale convinzione? Proba-

linea del tempo

350 a.C. Aristotele esclude l'esistenza dell'infinito

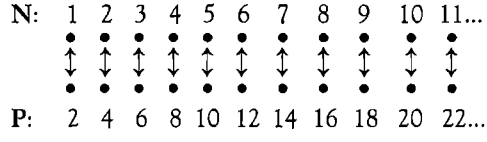
1639 Girard Desargues introduce il concetto di infinito in geometria

bilmente su un ragionamento di questo tipo: «I numeri interi sono per metà pari e per metà dispari». Cantor concorderebbe con questa risposta, ma avanzerebbe un'altra motivazione: direbbe che ogni numero dispari possiede un «compagno» adiacente che è pari. Il concetto che gli insiemi D e P siano costituiti dallo stesso numero di elementi si basa sulla possibilità di abbinare un numero pari a ogni numero dispari.



Se poi qualcuno domandasse: «Esistono tanti numeri interi quanti numeri *pari*?», la risposta sarebbe «no», con la motivazione che l'insieme **N** possiede il doppio degli elementi dell'insieme dei numeri pari.

L'affermazione che un insieme possiede «più elementi» di un altro, però, diventa assai confusa quando si ha a che fare con insiemi costituiti da un numero indefinito di elementi. Potrebbe andarci meglio con il concetto della corrispondenza biunivoca. Strano ma vero, esiste una corrispondenza biunivoca fra **N** e l'insieme dei numeri pari:



Si giunge quindi alla sorprendente conclusione che esiste lo «stesso numero» di numeri interi e di numeri pari. Ciò contraddice in pieno la «nozione comune» enunciata dagli antichi Greci: all'inizio degli *Elementi*, Euclide di Alessandria afferma che «il tutto è maggiore della parte».

Cardinalità Il numero di elementi di un insieme è detto «cardinalità». Nel caso delle pecore, la cardinalità registrata dal contabile dell'allevamento sarà, per esempio, 42; la cardinalità dell'insieme $\{a, b, c, d, e\}$ è 5 e si indica con *card* $\{a, b, c, d, e\} = 5$. Quindi la cardinalità misura le «dimensioni» di un insieme. Per indicare la cardinalità dell'insieme dei numeri interi **N** e di ogni insieme in corrispondenza biunivoca con **N**, Cantor introdusse il simbolo $\aleph_0$ ($\aleph$, o «alef», è una lettera dell'alfabeto ebraico e il simbolo $\aleph_0$ si legge «alef zero»). In linguaggio matematico, *card* (**N**) = *card* (**D**) = *card* (**P**) = $\aleph_0$.

1655
Si ritiene che John Wallis sia stato il primo a usare il simbolo del «nodo d'amore» ∞ per indicare l'infinito

1874
Cantor descrive il concetto di infinito in modo rigoroso, specificando vari ordini di infinità

anni '60
Abraham Robinson elabora l'analisi non standard, un'aritmetica basata sul concetto dell'infinitesimo e dell'infinito

Ogni insieme che può essere posto in corrispondenza biunivoca con N è detto «insieme infinito numerabile»: ciò significa che è possibile elencarne gli elementi. Ad esempio, l'elenco dei numeri dispari è 1, 3, 5, 7, 9,... ed è noto quale elemento viene per primo, quale per secondo e così via.

Le frazioni formano un insieme infinito numerabile? L'insieme delle frazioni Q è più grande di N, nel senso che N può essere considerato un sottoinsieme di Q. È possibile elencare tutti gli elementi di Q? Si può pensare di costruire un elenco che contenga qualsiasi frazione (anche negativa)? L'idea di stabilire una corrispondenza biunivoca fra un insieme così grande e N può sembrare assurda. Eppure è possibile.

In primo luogo, scriviamo una riga contenente tutti i numeri interi, alternativamente positivi e negativi. Nella riga sottostante scriviamo tutte le frazioni il cui denominatore è 2, omettendo quelle che compaiono nella riga sopra (come 6/2 = 3). Sotto questa riga scriviamo le frazioni il cui denominatore è 3, continuando a omettere quelle che sono già state registrate, e così via. Naturalmente non si arriva mai alla fine della costruzione, ma si sa in quale punto esatto dello schema compare ciascuna frazione. Ad esempio, 209/67 si trova nella 67ª riga, circa 200 posizioni più a destra di 1/67.

$1\ -1\ 2\ -2\ 3\ -3\ 4\ldots$

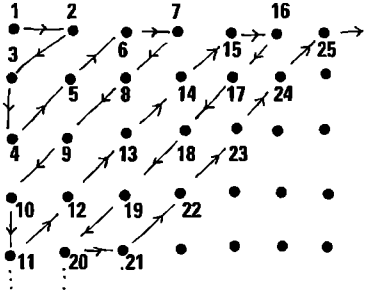
$\frac{1}{2}\ -\frac{1}{2}\ \frac{3}{2}\ -\frac{3}{2}\ \frac{5}{2}\ -\frac{5}{2}\ \frac{7}{2}\ldots$

$\frac{1}{3}\ -\frac{1}{3}\ \frac{2}{3}\ -\frac{2}{3}\ \frac{4}{3}\ -\frac{4}{3}\ \frac{5}{3}\ldots$

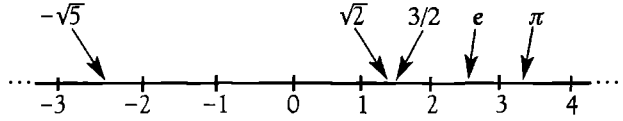
$\frac{1}{4}\ -\frac{1}{4}\ \frac{3}{4}\ -\frac{3}{4}\ \frac{5}{4}\ -\frac{5}{4}\ \frac{7}{4}\ldots$

$\frac{1}{5}\ -\frac{1}{5}\ \frac{2}{5}\ -\frac{2}{5}\ \frac{3}{5}\ -\frac{3}{5}\ \frac{4}{5}\ldots$

Rappresentando (almeno in linea di principio) tutte le frazioni in questo modo, si costruisce un elenco di dimensione 1. Partendo dalla prima riga e spostandosi a ogni passo verso destra non si arriva mai alla seconda riga; per raggiungerla bisogna seguire un percorso a zig-zag. Partendo da 1, il preannunciato elenco lineare comincia con: 1, –1, 1/2, 1/3, –1/2, 2, –2, e così via, seguendo le frecce. Ogni frazione, positiva o negativa che sia, compare nell'elenco e, inversamente, la sua posizione ci dice qual è la sua «compagna» nell'elenco delle frazioni. Si può quindi concludere che l'insieme delle frazioni Q è un insieme infinito numerabile e si può scrivere *card* (Q) = $\aleph_0$.



L'elenco dei numeri reali Se molti degli elementi presenti sulla retta dei numeri reali appartengono all'insieme delle frazioni, esistono anche numeri reali come $\sqrt{2}$, *e*, π che *non* sono frazioni. Si tratta dei numeri irrazionali, che «riempiono i buchi» fra una frazione e l'altra producendo la retta dei numeri reali R.



Una volta colmate le lacune, si parla dell'insieme **R** come del «continuo». Come si costruisce un elenco dei numeri reali? Con una mossa assolutamente geniale, Cantor dimostrò che il tentativo di stilare un elenco dei numeri reali compresi anche solo fra 0 e 1 era destinato a fallire. Questa notizia sarà uno shock per gli amanti delle liste, che si chiederanno perché mai un insieme di numeri non possa essere elencato scrivendo un numero dopo l'altro.

Supponiamo di non credere a Cantor. Sappiamo che ogni numero compreso fra 0 e 1 può essere espresso mediante un'espansione decimale, come 1/2 = 0,50000000000000000... e $1/\pi$ = 0,318309886183790671 53... Dobbiamo essere in grado di dire a Cantor: «Ecco qui l'elenco di *tutti* i numeri compresi fra 0 e 1», che indichiamo con r_1, r_2, r_3, r_4, r_5,... Se non ci riusciamo, ha ragione Cantor.

Immaginiamo che Cantor esamini l'elenco che abbiamo prodotto ed evidenzi in **grassetto** i numeri sulla diagonale:

r_1: $0, \mathbf{a_1} a_2 a_3 a_4 a_5 \ldots$

r_2: $0, b_1 \mathbf{b_2} b_3 b_4 b_5 \ldots$

r_3: $0, c_1 c_2 \mathbf{c_3} c_4 c_5$

r_4: $0, d_1 d_2 d_3 \mathbf{d_4} d_5 \ldots$

Potrebbe dire: «Benissimo, ma dov'è il numero $x = x_1x_2x_3x_4x_5\ldots$ in cui x_1 è diverso da a_1, x_2 è diverso da b_2, x_3 è diverso da c_3 e così via scendendo lungo la diagonale?». Il numero x differisce per una cifra decimale da *ogni* numero dell'elenco e quindi non può comparirvi. Cantor ha ragione.

In effetti non si può scrivere un elenco dell'insieme dei numeri reali **R**; ciò significa che si tratta di un insieme infinito «più grande», dotato di «un ordine di infinità maggiore» rispetto all'insieme delle frazioni **Q**. Il grande è diventato ancora più grande.

idea chiave

Grande, grandissimo, anzi infinito

08 I numeri immaginari

Chiunque è in grado di immaginare dei numeri ma, se immaginiamo che il nostro conto in banca contenga un milione di euro, si tratterà quasi certamente di un «numero immaginario». In matematica, però, il termine «immaginario» assume un significato che non ha niente a che vedere con i sogni a occhi aperti.

Sembra che il termine «immaginario» sia stato coniato dal filosofo e matematico René Descartes quando si rese conto che alcune equazioni ammettevano soluzioni bizzarre che non erano certamente numeri ordinari. La possibilità dell'esistenza dei numeri immaginari è una questione su cui i filosofi hanno rimuginato a lungo, perché si concentravano sul termine «immaginario». Per i matematici l'esistenza dei numeri immaginari non è affatto un problema: questi numeri appartengono alla loro quotidianità quanto 5 o π. I numeri immaginari non servono per andare a far spese, ma basta parlare con un progettista di aeroplani o un ingegnere elettrico per scoprire che sono fondamentali. Se poi si sommano un numero reale e un numero immaginario, si ottiene un «numero complesso», che dal punto di vista filosofico suona molto meno problematico. La teoria dei numeri complessi è incentrata sulla radice quadrata di -1. Qual è il numero che, elevato al quadrato, dà -1?

Moltiplicando per se stesso (cioè elevando al quadrato) qualunque numero diverso da zero si ottiene un numero positivo. Se ciò è plausibile per un numero positivo, nel caso di un numero negativo questa conclusione è meno intuitiva. Per verificarla, consideriamo il prodotto $(-1) \times (-1)$. Chi ha dimenticato la regola scolastica che dice «il prodotto di due numeri negativi è positivo», forse ricorderà che il risultato può essere -1 o $+1$. Se $(-1) \times (-1)$ fosse uguale a -1 si potrebbero dividere entrambi i membri per -1, ma si arriverebbe all'equivalenza $-1 = 1$, che è un'assurdità. Quindi $(-1) \times (-1) = 1$, che è positivo. Estendendo il ragionamento ad altri numeri negativi si giunge alla conclusione che, elevando al quadrato un numero reale, non si

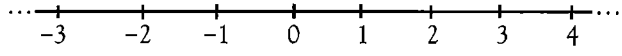
linea del tempo

1572
Rafael Bombelli effettua calcoli con i numeri immaginari

1777
Eulero introduce il simbolo *i* per rappresentare la radice quadrata di -1

ottiene *mai* un risultato negativo. Quando i numeri immaginari fecero la loro comparsa, nel XVI secolo, ciò portò a un'*impasse*, ma il risultato che consentì di uscirne affrancò la matematica dalla schiavitù dei numeri ordinari e spalancò campi di ricerca fino ad allora inimmaginabili. Lo sviluppo dei numeri complessi porta al «completamento dei numeri reali» per produrre un sistema più perfetto.

La radice quadrata di –1 Come abbiamo visto, se ci si limita a considerare la retta dei numeri reali



la radice quadrata di –1 non esiste, perché il quadrato di un numero deve necessariamente essere positivo. Se continuiamo a pensare solo ai numeri della retta reale possiamo anche darci subito per vinti, continuare a chiamarli numeri immaginari, andare a prendere un caffè con i filosofi e lavarcene le mani. In alternativa, possiamo fare una mossa coraggiosa e accettare $\sqrt{-1}$ come una nuova entità matematica, che indicheremo con i.

L'ingegnerizzazione di $\sqrt{-1}$

Anche gli ingegneri, noti per la loro concretezza, hanno trovato alcune applicazioni per i numeri complessi. Quando Michael Faraday scoprì la corrente alternata, fra il 1830 e il 1840, i numeri immaginari acquistarono una realtà fisica. In questo ambito $\sqrt{-1}$ è rappresentata dalla lettera j, perché i è il simbolo della corrente elettrica.

Già solo per effetto di questo atto mentale, i numeri immaginari esistono. Cosa siano ancora non lo sappiamo, ma siamo convinti della loro esistenza. Quanto meno sappiamo che $i^2 = -1$. In questo nuovo sistema numerico si ritrovano tutti i vecchi amici, come i numeri reali 1, 2, 3, 4, π, e, $\sqrt{2}$ e $\sqrt{3}$, e si incontra qualche nuovo arrivato che contiene i, come $1 + 2i$, $-3 + i$, $2 + 3i$, $1 + i\sqrt{2}$, $\sqrt{3} + 2i$, $e + \pi i$ e altri.

Questo eccezionale sviluppo della matematica ebbe luogo all'inizio del XIX secolo, quando si sfuggì alla monodimensionalità della retta dei numeri per passare in un nuovo e bizzarro piano numerico bidimensionale.

Somme e moltiplicazioni Ora che abbiamo in mente i numeri complessi di forma $a + bi$, vediamo come possiamo utilizzarli. Innanzi tutto è possibile sommarli e moltiplicarli, esattamente come i numeri reali. La somma si effettua addizionando le parti corrispondenti; perciò sommando $2 + 3i$ e $8 + 4i$ si ottiene $(2 + 8) + (3 + 4)i = 10 + 7i$.

La moltiplicazione è quasi altrettanto semplice. Se si vuole moltiplicare $2 + 3i$ per $8 + 4i$, si moltiplica ogni coppia di simboli

1806
La rappresentazione dei numeri complessi in un grafico prende il nome di «piano di Argand»

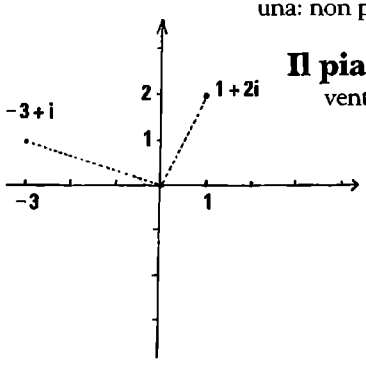
1811
Carl Friedrich Gauss usa funzioni di variabili complesse

1837
William R. Hamilton tratta i numeri complessi come coppie ordinate di numeri reali

$$(2 + 3i) \times (8 + 4i) = (2 \times 8) + (2 \times 4i) + (3i \times 8) + (3i \times 4i)$$

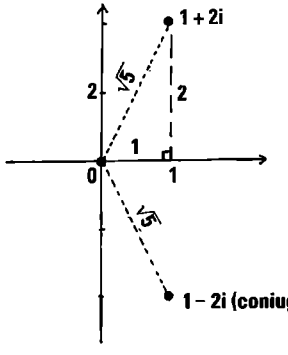
e poi si sommano i termini risultanti, 16, 8i, 24i e 12i^2 (nell'ultimo si sostituisce i^2 con –1). Il risultato della moltiplicazione è (16 –12) + (8i + 24i), che è uguale al numero complesso 4 + 32i.

I numeri complessi soddisfano tutte le regole aritmetiche ordinarie. La sottrazione e la divisione sono sempre possibili (tranne la divisione per il numero complesso 0 + 0i, ma non è consentita neanche la divisione per zero nel caso dei numeri reali). Di fatto i numeri complessi godono di tutte le proprietà dei numeri reali, tranne una: non possono essere suddivisi in numeri positivi e negativi.



Il piano di Argand Il carattere bidimensionale dei numeri complessi diventa evidente quando si rappresentano su un grafico, cioè quando si riportano, ad esempio, i numeri complessi –3 + 2i e 1 + 2i nel cosiddetto «piano di Argand». Questa rappresentazione dei numeri complessi prende il nome dal matematico svizzero Jean Robert Argand, sebbene anche altri abbiano avuto un'idea analoga più o meno nello stesso periodo.

Ogni numero complesso possiede un «compagno», che ufficialmente è il suo «coniugato». Il coniugato di 1 + 2i è 1 – 2i, il numero che si ottiene cambiando di segno il secondo termine. Per lo stesso motivo, il coniugato di 1 – 2i è 1 + 2i.



Sommando e moltiplicando due numeri coniugati si ottiene un numero reale. La somma di 1 + 2i e 1 – 2i è uguale a 2, il loro prodotto è 5. Il caso della moltiplicazione è particolarmente interessante. Il risultato ottenuto, 5, è il quadrato della «lunghezza» (o, in termini matematici, del «modulo») del numero complesso 1 + 2i (che è equivalente a quella del suo coniugato). Invertendo la definizione, il modulo di un numero complesso w può essere definito come:

$$\textit{modulo di } w = \sqrt{(w \times \textit{coniugato di } w)}.$$

Verificando l'espressione per –3 + i, si trova che il modulo di (–3 + i) = $\sqrt{(-3 + i) \times (-3 - i)} = \sqrt{(9 + 1)} = \sqrt{10}$.

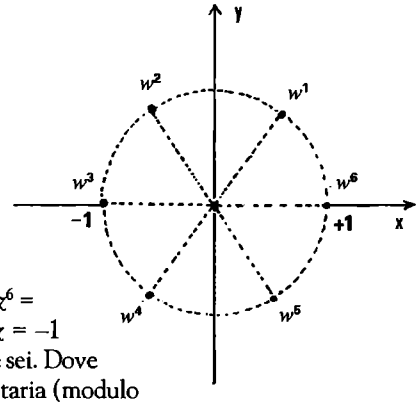
L'uscita dei numeri complessi dall'ambito mistico si deve in buona parte a Sir William Rowan Hamilton, il più eminente matematico irlandese del XIX secolo. Hamilton si rese conto che i non rispondeva a un'effettiva esigenza teorica, ma

agiva solo da indicatore di posizione e quindi poteva essere eliminato. Trattò allora i numeri complessi come «coppie ordinate» di numeri reali (a, b), rivelandone la natura bidimensionale e rinunciando alla mistica $\sqrt{-1}$. Privata di i, la somma diventa

$$(2, 3) + (8, 4) = (10, 7)$$

e, per quanto sia un po' meno ovvio, la moltiplicazione è

$$(2, 3) \times (8, 4) = (4, 32)$$



La completezza del sistema dei numeri complessi diventa più evidente se si pensa alle cosiddette «radici n-esime dell'unità» (nel linguaggio matematico «unità» significa «uno»), cioè alle soluzioni dell'equazione $z^n = 1$. A titolo di esempio consideriamo $z^6 = 1$. Sulla retta dei numeri reali si trovano le due radici $z = 1$ e $z = -1$ (perché $1^6 = 1$ e $(-1)^6 = 1$), ma in totale le radici dovrebbero essere sei. Dove sono finite le altre? Come le due radici reali, hanno lunghezza unitaria (modulo unitario) e si trovano sulla circonferenza di raggio unitario centrata nell'origine.

Ma c'è dell'altro. Partendo da $w^1 = 1/2 + \sqrt{3}/2\, i$, la radice che si trova nel primo quadrante, e procedendo in senso antiorario si trovano successivamente le radici w^2, w^3, w^4, w^5, $w^6 = 1$, che occupano i vertici di un esagono regolare. In generale, le n radici dell'unità si trovano tutte sulla circonferenza e occupano gli angoli, o «vertici», di una figura regolare con n lati (un poligono).

Estensione dei numeri complessi Una volta ottenuti i numeri complessi, i matematici hanno istintivamente cercato di generalizzarli. Se i numeri complessi sono bidimensionali, non è detto che il numero di dimensioni debba per forza essere uguale a 2. Hamilton tentò per anni di costruire dei numeri tridimensionali e di elaborare un metodo per sommarli e moltiplicarli, ma riuscì nel suo intento solo dopo essere passato a quattro dimensioni. Poco dopo, i numeri tetradimensionali vennero a loro volta generalizzati in otto dimensioni (numeri di Cayley). Sono stati effettuati vari tentativi di ottenere numeri a sedici dimensioni, ma cinquant'anni dopo la storica impresa di Hamilton è stato dimostrato che è impossibile.

idea chiave

09 I numeri primi

La matematica attraversa tutti gli ambiti dell'attività umana ed è talmente vasta che ci si può sentire smarriti; perciò di tanto in tanto è bene rifarsi ai suoi princìpi fondamentali. Invariabilmente, ciò si traduce in un ritorno ai numeri naturali 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12... Più fondamentali di così!

Com'è noto, $4 = 2 \times 2$, perciò 4 può essere scomposto in fattori primi. Il quattro non è l'unico numero scomponibile in questo modo; ad esempio, $6 = 2 \times 3$, $8 = 2 \times 2 \times 2$, $9 = 3 \times 3$, $10 = 2 \times 5$, $12 = 2 \times 2 \times 3$. Sono tutti numeri composti, perché sono stati costruiti a partire dai numeri *fondamentali* 2, 3, 5, 7, 11, 1... I numeri non ulteriormente scomponibili si chiamano «numeri primi» e sono divisibili solo per 1 e per se stessi. Ci si può chiedere allora se 1 sia un numero primo. In base alla definizione dovrebbe esserlo e in effetti molti autorevoli matematici del passato lo hanno trattato come tale, ma oggi i numeri primi vengono fatti partire dal 2, perché in questo modo gli enunciati dei teoremi assumono una forma più elegante. Consideriamo quindi che 2 sia il primo numero di questo tipo.

Sottolineiamo i numeri naturali che sono anche primi: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 21, 22, 23,... Lo studio dei numeri primi ci riporta ai fondamenti della fisica. L'importanza di questi numeri risiede nel fatto che sono gli «atomi» della matematica; come gli elementi chimici costituiscono la base da cui si ottengono tutti gli altri composti chimici, così i numeri primi possono essere combinati per creare i composti matematici.

Il risultato matematico che unifica tutto ciò porta il nome altisonante di «teorema fondamentale dell'aritmetica» e afferma che ogni numero intero maggiore di 1 può essere scomposto come prodotto di numeri primi in un unico modo. Ad esempio, $2 \times 2 \times 3$ è l'unica scomposizione in fattori primi possibile per il numero 12. In molti casi si preferisce usare la notazione esponenziale: $12 = 2^2 \times 3$. Un altro esempio è 6 545 448, che è scomponibile in $2^3 \times 3^5 \times 7 \times 13 \times 37$.

La scoperta dei numeri primi Purtroppo non esistono formule prestabilite che ci consentano di identificare i numeri primi e non sembra esistere alcuna

linea del tempo

300 a.C.

Negli *Elementi* di Euclide compare la dimostrazione dell'esistenza di infiniti numeri primi

230 a.C.

Eratostene di Cirene descrive un metodo che consente di isolare i numeri primi dai numeri interi

regolarità nel modo in cui essi compaiono fra i numeri interi. Uno dei primi metodi per scovarli fu ideato da Eratostene di Cirene, un contemporaneo un po' più giovane di Archimede che visse quasi sempre ad Atene. Ai suoi tempi Eratostene suscitò molta ammirazione per aver calcolato la lunghezza dell'equatore con una buona precisione, mentre oggi è ricordato soprattutto per il suo «crivello» (setaccio), un procedimento che consente di individuare i numeri primi. Raffigurandosi la successione dei numeri naturali, Eratostene immaginò di sottolineare 2 e di cancellare tutti i suoi multipli, per poi passare a 3, sottolinearlo, eliminare tutti i multipli, e così via. In questo modo tutti i numeri composti venivano scartati e i numeri che erano «rimasti nel setaccio» (sottolineati) erano i numeri primi (nella tabella sono evidenziati in grigio).

0	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49
50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69
70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89
90	91	92	93	94	95	96	97	98	99

Se è possibile prevedere quali sono i numeri primi, stabilire se un determinato numero sia primo o meno è un'altra questione. Ci si potrebbe chiedere, ad esempio, se 19 071 o 19 073 siano primi. Tranne 2 e 5, i numeri primi terminano sempre con 1, 3, 7, o 9, ma il fatto che un numero soddisfi questo requisito non garantisce che sia primo. È difficile stabilire se un numero grande che termina con 1, 3, 7, o 9 sia un numero primo senza provare a scomporlo (per inciso, $19\,071 = 3^2 \times 13 \times 163$ e quindi non è un numero primo, mentre 19 073 lo è).

Un'altra impresa di notevole difficoltà è consistita nell'individuare una regolarità nella distribuzione dei numeri primi. Vediamo quanti numeri primi ci sono in ogni intervallo di 100 numeri fra 1 e 1000.

Intervallo	1-100	101-200	201-300	301-400	401-500	501-600	601-700	701-800	801-900	901-1000	1-1000
Numeri primi	25	21	16	16	17	14	16	14	15	14	168

Nel 1792, a soli 15 anni, Carl Friedrich Gauss elaborò una formula $P(n)$ che consentiva di calcolare il numero di primi minori di un certo n (il cosiddetto «teorema dei numeri primi»). Per $n = 1000$ la formula dà un valore approssimato di 172, ma in realtà i numeri primi sono 168, un valore minore della stima teorica. Si era sempre dato per scontato che ciò fosse vero per ogni n, ma i numeri primi riservano continue

1742

Goldbach ipotizza che ogni numero pari (maggiore di 2) sia equivalente alla somma di due numeri primi

1896

Viene dimostrato il teorema dei numeri primi riguardante la loro distribuzione

1966

Chen Jingrun si avvicina alla dimostrazione della congettura di Goldbach

sorprese ed è stato dimostrato che per $n = 10^{371}$ (scritto per esteso è un numero enorme, un 1 seguito da 371 zeri) sono più del previsto. Infatti, in alcune regioni dei numeri naturali la stima teorica può essere inferiore o superiore al numero effettivo di numeri primi minori di un dato valore.

Quanti sono? Esistono infiniti numeri primi. Nei suoi *Elementi* (libro IX, proposizione 20) Euclide affermò che «esistono [sempre] numeri primi in numero maggiore di quanti numeri primi si voglia proporre». La bella dimostrazione di Euclide dice più o meno così:

> Supponiamo che P sia il più grande numero primo e consideriamo $N = (2 \times 3 \times 5 \times \quad \times P) + 1$. O N è primo o non lo è. Se N è primo, si è ottenuto un numero primo maggiore di P, in contraddizione con l'ipotesi iniziale. Se non è primo, deve essere divisibile per un numero primo p, che può essere 2, 3, 5, ..., P. Ciò significa che p è un divisore di $N - (2 \times 3 \times 5 \times ... \times P)$ e, poiché questa differenza è uguale a 1, che p è un divisore di 1. Ma ciò è impossibile, perché ogni numero primo è maggiore di 1. Quindi, qualunque sia la natura di N, si giunge a una contraddizione. Ciò significa che l'assunzione iniziale che esista un numero primo P più grande di tutti gli altri è falsa. Conclusione: esistono infiniti numeri primi.

Il fatto che i numeri primi «si estendano all'infinito» non ha posto fine alla ricerca del più grande numero primo conosciuto. Un detentore recente del record è il gigantesco numero primo di Mersenne $2^{25\,964\,951} - 1$, che è uguale a 1 seguito da 7 816 230 cifre.

Mondi ignoti Due ambiti ancora poco chiari, ma molto importanti, che riguardano i numeri primi sono il «problema dei primi gemelli» e la famosa «congettura di Goldbach».

I numeri primi gemelli sono coppie di numeri primi consecutivi separati solo da un numero pari. I primi gemelli compresi fra 1 e 100 sono: 3, 5; 5, 7; 11, 13; 17, 19; 29, 31; 41, 43; 59, 61; 71, 73. Dal punto di vista numerico si sa che esistono 27 412 679 numeri primi gemelli minori di 10^{10}; ciò significa che i numeri pari associati ai gemelli, come 12 (che separa i gemelli 11 e 13), costituiscono appena lo 0,274% dei numeri compresi in questo intervallo. Esiste un numero infinito di primi gemelli? Sarebbe strano se non fosse così, ma per il momento nessuno è stato in grado di dimostrarlo.

Christian Goldbach ipotizzò che:

ogni numero pari maggiore di 2 può essere scritto come somma di due numeri primi.

Ad esempio, 42 è un numero pari e può essere scritto come 5 + 37. Il fatto che sia equivalente anche ad altre somme di numeri primi, come 11 + 31, 13 + 29 o 19 + 23, non

ha alcuna importanza; l'unico requisito è che esista una somma. Questa ipotesi è vera per moltissimi numeri, ma non è mai stata dimostrata in generale; comunque sono stati fatti dei progressi e si pensa che una dimostrazione sia imminente. Il matematico cinese Chen Jingrun ha fatto un importante passo avanti: il suo teorema afferma che ogni numero pari sufficientemente grande può essere scritto come somma di due numeri primi o, in alternativa, come la somma di un numero primo e di uno semiprimo (uguale al prodotto di due numeri primi).

Il grande studioso di teoria dei numeri Pierre de Fermat dimostrò che i numeri primi di forma $4k + 1$ potevano essere espressi come somma di due quadrati in un unico modo (ad esempio, $17 = 1^2 + 4^2$), mentre quelli di forma $4k + 3$ (come il 19) non erano esprimibili come somma di due quadrati. Anche Joseph Lagrange dimostrò un celebre teorema matematico riguardante i numeri quadrati: *ogni* numero intero positivo è dato dalla somma di quattro quadrati; per esempio, $19 = 1^2 + 1^2 + 1^2 + 4^2$. Sono state prese in esame anche le potenze superiori e si sono riempiti libri e libri di teoremi, ma molte questioni restano ancora aperte.

Il sogno del numerologo

Uno dei settori più ardui della teoria dei numeri riguarda il «problema di Waring». Nel 1770 Edward Waring, un professore di Cambridge, enunciò alcuni problemi riguardanti la possibilità di esprimere i numeri interi come somme di potenze. Quando si parla di numeri primi, somme di quadrati e somme di cubi si entra in un ambito in cui l'oggettività scientifica della matematica incontra le arti magiche della numerologia. In questo campo, nessun numero può aspirare alla venerazione di cui gode il 666, che viene citato come «numero della bestia» nel libro biblico dell'Apocalisse ed è dotato di proprietà insolite. Infatti, 666 è esprimibile come somma dei quadrati dei primi sette numeri primi:

$$666 = 2^2 + 3^2 + 5^2 + 7^2 + 11^2 + 13^2 + 17^2$$

I numerologi ci tengono anche a sottolineare che 666 è equivalente a una somma palindroma di cubi; come se non bastasse, il termine centrale di tale somma è 6^3, ovvero la notazione abbreviata di $6 \times 6 \times 6$:

$$666 = 1^3 + 2^3 + 3^3 + 4^3 + 5^3 + 6^3 + 5^3 + 4^3 + 3^3 + 2^3 + 1^3$$

Non c'è che dire, 666 è davvero il sogno di ogni numerologo.

Abbiamo detto che i numeri primi sono gli «atomi della matematica». Si potrebbe obiettare che la fisica si è spinta oltre la frontiera dell'atomo, scoprendo costituenti ancora più fondamentali, come i quark. La matematica si è fermata ai numeri primi? Se ci si limita ai numeri naturali, 5 è un numero primo e sempre lo sarà. Ma Gauss fece una scoperta di vasta portata: per alcuni numeri primi, come 5, si può scrivere $5 = (1 - 2i) \times (1 + 2i)$ dove i è l'unità immaginaria uguale a $\sqrt{-1}$. In quanto prodotti di due numeri interi gaussiani, 5 e altri numeri dello stesso tipo diventano scomponibili, contrariamente a quanto si era sempre pensato.

idea chiave

Gli atomi della matematica

10 I numeri perfetti

La ricerca della perfezione matematica ha condotto in varie direzioni coloro che vi aspiravano. Esistono i quadrati perfetti, ma in questo caso la perfezione non ha una valenza estetica: serve più che altro a ricordarci che esistono anche quadrati imperfetti. In ambito diverso, esistono numeri che hanno pochi divisori, altri che ne hanno tanti e altri ancora che hanno proprio i divisori «giusti»: un numero che è uguale alla somma dei propri divisori è «perfetto».

Secondo il filosofo greco Speusippo, il nipote di Platone che gli succedette alla guida dell'Accademia, i Pitagorici ritenevano che il numero 10 presentasse tutte le caratteristiche della perfezione. Infatti, fra 1 e 10 erano compresi tanti numeri primi (2, 3, 5, 7) quanti non primi (4, 6, 8, 9) e il dieci era il numero più piccolo a possedere questa proprietà. Certa gente ha una strana idea di perfezione...

In realtà, sembra che l'ideale di perfezione numerica dei Pitagorici fosse molto più articolato. Le proprietà matematiche dei numeri perfetti furono descritte da Euclide nei suoi *Elementi* e studiate in modo approfondito da Nicomaco quattro secoli dopo. In base alle relazioni esistenti fra i numeri e i loro divisori furono definiti i numeri «amichevoli» e «socievoli», finché, con l'elaborazione della teoria dei numeri abbondanti e difettivi, la scuola pitagorica trovò il suo ideale di perfezione.

Il fatto che un numero sia abbondante o meno è determinato dai suoi divisori e gioca sul collegamento fra moltiplicazione e addizione. Consideriamo il numero 30 e i suoi divisori, cioè tutti i numeri che lo dividono esattamente e che sono *minori* di 30. Per un numero piccolo come 30 è facile verificare che i divisori sono 1, 2, 3, 5, 6, 10 e 15, e che la loro somma è uguale a 42. Il numero 30 è «abbondante», perché la somma dei divisori (42) è maggiore del numero stesso.

Un numero è «difettivo» nel caso opposto, quando la somma dei divisori è minore del numero in questione. Quindi 26 è difettivo, perché sommando i suoi divisori

linea del tempo

525 a.C.
In vari scritti, si legge che i Pitagorici si interessavano ai numeri perfetti e a quelli abbondanti

300 a.C.
Il nono libro degli *Elementi* di Euclide prende in esame i numeri perfetti

100
Nicomaco di Gerasa fornisce una classificazione dei numeri basata sui numeri perfetti

I primi numeri perfetti

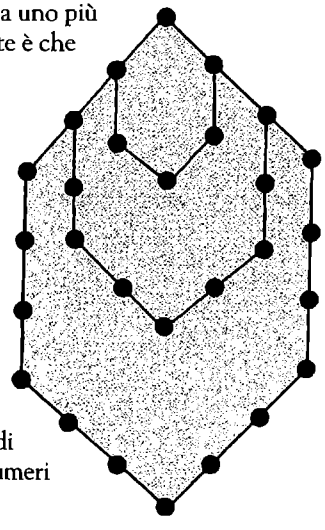
Posizione nella sequenza	1	2	3	4	5	6	7
Numero perfetto	6	28	496	8128	33 550 336	8 589 869 056	137 438 691 328

1, 2 e 13 si ottiene 16, che è minore di 26. I numeri primi sono molto difettivi, perché la somma dei loro divisori è sempre e solo 1.

Un numero che non è né abbondante né difettivo è perfetto: la somma dei divisori di un numero perfetto è uguale al numero stesso. Il primo numero perfetto è 6 perché i suoi divisori sono 1, 2 e 3, dalla cui somma si ottiene 6. I Pitagorici erano talmente affascinati dal numero 6 e dal modo in cui si combinavano i suoi fattori da chiamarlo «matrimonio, salute e bellezza». Un'altra storia che riguarda il 6 viene riferita da Sant'Agostino (354-430). Agostino era convinto che la perfezione del 6 precedesse l'esistenza del mondo e che questo fosse stato creato in sei giorni proprio *perché* quel numero era perfetto.

Il numero perfetto successivo è 28: i suoi divisori sono 1, 2, 4, 7 e 14, la cui somma fa esattamente 28. I primi due numeri di questo tipo, 6 e 28, occupano un ruolo particolare nella scienza dei numeri perfetti, perché è possibile dimostrare che ogni numero perfetto pari termina con un 6 o con un 28. Dopo 28 non si incontrano altri numeri perfetti fino a 496 che, come si può facilmente verificare, è uguale alla somma dei suoi divisori: 496 = 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248. Per trovare il termine successivo bisogna avventurarsi fra numeri stratosferici. I primi cinque numeri perfetti sono noti dal XVI secolo, ma si ignora se ne esista uno più grande di tutti o se i numeri perfetti siano illimitati. L'opinione prevalente è che vadano avanti all'infinito, come i numeri primi.

I Pitagorici avevano una vera passione per le rappresentazioni geometriche. Se si possiede un numero perfetto di perle, è possibile disporle in modo da formare una collana esagonale; nel caso del 6 si tratta di un semplice esagono ai cui vertici si trovano le perle, ma per numeri perfetti maggiori bisogna aggiungere altre collane più piccole all'interno di quella principale.



I numeri di Mersenne La chiave per costruire i numeri perfetti è costituita dai numeri che prendono il nome da padre Marin Mersenne, un monaco francese che studiò nello stesso collegio di gesuiti di René Descartes. Entrambi si interessarono alla ricerca dei numeri perfetti. I numeri di Mersenne si ottengono sottraendo 1 dalle potenze di 2, cioè dai numeri (che raddoppiano sempre) 2, 4, 8, 16, 32, 64, 128, 256, ..., e quindi hanno la forma $2^n - 1$. Sono tutti dispari, ma non sempre sono primi: i numeri

1603

Pietro Cataldi trova il sesto e il settimo numero perfetto, $2^{16}(2^{17} - 1) = 8\,589\,869\,056$ e $2^{18}(2^{19} - 1) = 137\,438\,691\,328$

2006

Il grande progetto di ricerca dei numeri primi scopre il 44° primo di Mersenne (composto da quasi dieci milioni di cifre), ma si sa che è possibile generare un numero perfetto ancora più grande

Esponente	Risultato	Si sottrae 1 (numero di Mersenne)	Numero primo?
2	4	3	primo
3	8	7	primo
4	16	15	non primo
5	32	31	primo
6	64	63	non primo
7	128	127	primo
8	256	255	non primo
9	512	511	non primo
10	1024	1023	non primo
12	4096	4095	non primo
13	8192	8191	primo
14	16384	16383	non primo
15	32768	32767	non primo

di Mersenne che sono anche primi possono essere utilizzati per costruire i numeri perfetti.

Mersenne sapeva che, se l'esponente della potenza di 2 non era un numero primo, non poteva esserlo neanche il corrispondente numero di Mersenne, come si può constatare osservando i primi esponenti non primi 4, 6, 8, 9, 10, 12, 14 e 15 in tabella. Quindi, i numeri di Mersenne potevano essere primi solo se l'esponente era un numero primo; restava da capire se questa condizione era sufficiente. Nei primi casi si ottiene 3, 7, 31 e 127, che sono tutti primi. Si può affermare che, in generale, un numero di Mersenne derivante da 2 elevato a un esponente primo è a sua volta primo?

Molti matematici dell'antichità (all'incirca fino al 1500) ritenevano che fosse così. Ma i numeri primi non si piegano a un'esigenza di semplicità e si scoprì che, nel caso dell'esponente 11 (un numero primo), $2^{11} - 1 = 2047 = 23 \times 89$, che ovviamente non è primo. Sembra che non esistano regole. I numeri di Mersenne $2^{17} - 1$ e $2^{19} - 1$ sono entrambi primi, mentre $2^{23} - 1$ non è primo, perché

$$2^{23} - 1 = 8\,388\,607 = 47 \times 178\,481$$

Costruzioni Combinando i risultati ottenuti da Euclide e da Eulero si ottiene una formula che consente di generare i numeri perfetti pari: n è un numero perfetto pari se e solo se $n = 2^{p-1}(2^p - 1)$, dove p è un numero primo e $2^p - 1$ è un numero primo di Mersenne.

Solo buoni amici

La mente razionale dei matematici è poco sensibile al fascino mistico dei numeri, ma la numerologia non è ancora morta. I numeri amichevoli sono arrivati dopo i numeri perfetti, ma non è escluso che fossero già noti ai Pitagorici, e in seguito sono stati utilizzati per stilare oroscopi sentimentali in cui le loro proprietà matematiche venivano tradotte nella natura del vincolo amoroso. I due numeri 220 e 284 sono amichevoli. Infatti, sommando i divisori di 220 (1, 2, 4, 5, 10, 11, 20, 22, 44, 55 e 110) si ottiene 284. L'avrete già capito: la somma dei divisori di 284 è uguale a 220. Se questa non è amicizia!

Alcuni esempi sono $6 = 2^1(2^2-1)$, $28 = 2^2(2^3-1)$ e $496 = 2^4(2^5-1)$. Con questa formula si possono generare i numeri perfetti pari, purché si sappiano individuare i numeri primi di Mersenne. I numeri perfetti hanno messo a dura prova uomini e macchine, e continueranno a farlo al di là di ogni previsione iniziale. Quando il grande compilatore di tabelle Peter Barlow scrisse, all'inizio del XIX secolo, che a suo parere nessuno sarebbe andato oltre il calcolo del numero perfetto di Eulero

$$2^{30}(2^{31}-1) = 2\,305\,843\,008\,139\,952\,128$$

perché aveva poco senso, faceva i conti senza la capacità di calcolo dei moderni computer e l'esigenza imprescindibile dei matematici di affrontare nuove sfide.

Numeri perfetti dispari Nessuno può sapere se verrà mai scoperto un numero perfetto dispari. Descartes pensava che fosse impossibile, ma anche gli esperti possono sbagliare. Il matematico inglese James Joseph Sylvester dichiarò che l'esistenza di un numero perfetto dispari «sarebbe stato poco meno di un miracolo», perché tale numero avrebbe dovuto soddisfare un gran numero di condizioni. Non c'è da stupirsi che Sylvester nutrisse dei dubbi; dopo tutto si tratta di uno dei problemi più antichi della matematica. Quello che è certo, però, è che se un numero perfetto dispari esiste, sappiamo già diverse cose sul suo conto: dovrebbe avere almeno otto divisori primi diversi, uno dei quali maggiore di un milione, ed essere composto da almeno 300 cifre.

I numeri primi di Mersenne

Individuare i numeri primi di Mersenne non è affatto facile. Nel corso dei secoli molti matematici hanno contribuito a stilarne l'elenco, con un'alternanza di errori e successi. Nel 1732 il grande Eulero ha trovato l'ottavo numero primo di Mersenne, $2^{31} - 1 = 2\,147\,483\,647$, e la scoperta (avvenuta nel 1963) del 23° primo di Mersenne, $2^{11213} - 1$, è stata una tale fonte di orgoglio per i matematici dell'Università dell'Illinois da indurli ad annunciare l'evento al mondo scrivendo il numero sul bollo postale dell'università. Con l'avvento di computer sempre più potenti la fabbrica dei numeri primi di Mersenne ha fatto progressi: alla fine degli anni '70 del Novecento due liceali, Laura Nickel e Landon Noll, hanno scoperto il 25° numero primo di Mersenne e in seguito Noll ha scoperto da solo il 26°. A tutt'oggi sono noti 45 numeri primi di Mersenne.

idea chiave
La mistica dei numeri

11 I numeri di Fibonacci

Nel *Codice da Vinci* l'autore Dan Brown immaginava che Jacques Saunière, il curatore del Louvre assassinato, lasciasse come indizio i primi otto termini di una successione numerica. Solo l'abilità della crittografa Sophie Neveu consentiva di riordinare i numeri 13, 3, 2, 21, 1, 1, 8 e 5 in modo da coglierne il significato. Ecco a voi la successione più famosa della matematica.

La successione dei numeri interi di Fibonacci

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584,...

è famosissima, perché presenta molte proprietà interessanti. Quella principale (che poi è la caratteristica che la definisce) è il fatto che ogni termine è uguale alla somma dei due precedenti. Ad esempio, 8 = 5 + 3, 13 = 8 + 5,..., 2584 = 1597 + 987, e così via. Basta ricordarsi di partire con 1 e 1 per generare il resto della successione in un battibaleno. La successione di Fibonacci si ritrova nelle proporzioni delle stanze e degli edifici progettati dagli architetti e, in natura, nel numero delle spirali successive orarie e antiorarie formate dai semi nei dischi dei girasoli (ad esempio, 34 in un senso, 55 nell'altro). Inoltre ha ispirato anche vari compositori e si pensa che la *Dance Suite* di Bartók sia collegata alla successione. Per quanto riguarda la musica contemporanea, l'album *The Binary Universe* di Brian Transeau (in arte BT) contiene un brano chiamato 1,618 in omaggio al limite dei rapporti fra i numeri di Fibonacci, di cui parleremo più avanti.

Le origini La successione di Fibonacci ha fatto la sua comparsa nel *Liber Abaci* pubblicato da Leonardo Fibonacci nel 1202, ma è probabile che questi numeri fossero già noti ai matematici indiani. Fibonacci enunciò il seguente problema riguardante la riproduzione dei conigli:

linea del tempo

1202 Leonardo Fibonacci pubblica il *Liber Abaci* in cui compaiono i numeri di Fibonacci

1724 Daniel Bernoulli esprime i numeri della successione di Fibonacci in termini della sezione aurea

Ogni mese una coppia di conigli adulti genera una coppia di coniglietti. All'inizio dell'anno è presente solo una coppia di coniglietti: alla fine del primo mese i due conigli sono diventati adulti, alla fine del secondo mese avranno generato una coppia di coniglietti. Il processo di raggiungimento dell'età adulta e di procreazione continua in questo modo; miracolosamente, non muore nessuna coppia di conigli.

○ = coppia giovane

● = coppia adulta

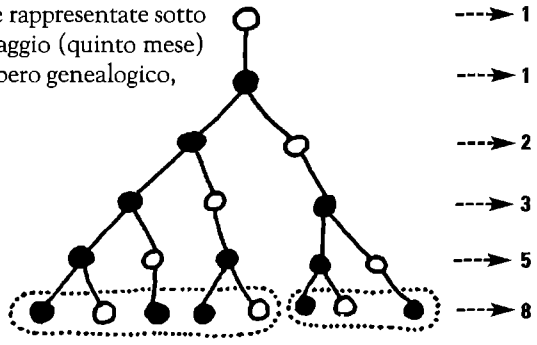
Fibonacci si chiedeva quante sarebbero state le coppie di conigli alla fine dell'anno. Le varie generazioni possono essere rappresentate sotto forma di un «albero genealogico». Alla fine di maggio (quinto mese) sono presenti otto coppie; in questo livello dell'albero genealogico, la parte di sinistra

● ○ ● ● ○

è uguale alla riga soprastante, mentre quella di destra

● ○ ●

è identica alla riga ancora precedente. Ciò dimostra che la nascita delle coppie di conigli segue l'equazione fondamentale di Fibonacci:



Popolazione dei conigli

numero dopo n mesi = numero dopo $(n-1)$ mesi + numero dopo $(n-2)$ mesi

Proprietà Vediamo cosa succede se si sommano i termini della successione:

$$1 + 1 = 2$$
$$1 + 1 + 2 = 4$$
$$1 + 1 + 2 + 3 = 7$$
$$1 + 1 + 2 + 3 + 5 = 12$$
$$1 + 1 + 2 + 3 + 5 + 8 = 20$$
$$1 + 1 + 2 + 3 + 5 + 8 + 13 = 33$$

Il risultato di ciascuna di queste somme forma a sua volta una successione, che si può scrivere sotto quella originaria, traslata di alcune posizioni:

1923

Bartók compone la *Dance Suite* che si ritiene sia stata ispirata dai numeri di Fibonacci

1963

Viene fondato il «Fibonacci Quarterly», una rivista dedicata alla teoria dei numeri della successione di Fibonacci

2007

Lo scultore Peter Randall-Page crea la scultura da settanta tonnellate *Seed* basata sulla successione di Fibonacci per l'Eden Project in Cornovaglia (Regno Unito)

Fibonacci	1	1	2	3	5	8	13	21	34	55	89
Somma				2	4	7	12	20	33	54	88

La somma di n termini della sequenza di Fibonacci è uguale al numero di Fibonacci che compare due posizioni dopo, diminuito di uno. Se si vuole calcolare la somma 1 + 1 + 2 + + 987 basta sottrarre 1 da 2584 per ottenere 2583. Sommando alternativamente i numeri della successione (cioè saltando ogni volta un termine) si ottiene 1 + 2 + 5 + 13 + 34, che è uguale a 55, un altro numero di Fibonacci. Sommando invece i numeri intermedi, la somma diventa 1 + 3 + 8 + 21 + 55 e si ottiene 88, corrispondente a un numero di Fibonacci diminuito di 1.

I quadrati dei numeri di Fibonacci sono altrettanto interessanti. Moltiplicando ciascun numero della successione per se stesso e sommando i risultati ottenuti, si ottiene un'altra successione:

Fibonacci	1	1	2	3	5	8	13	21	34	55
Quadrati	1	1	4	9	25	64	169	441	1156	3025 .
Somma dei quadrati	1	2	6	15	40	104	273	714	1870	4895

In questo caso, la somma di tutti i quadrati fino all'n-esimo termine della successione è equivalente al prodotto dell'n-esimo numero di Fibonacci per il successivo. Ad esempio

$$1 + 1 + 4 + 9 + 25 + 64 + 169 = 273 = 13 \times 21$$

I numeri di Fibonacci saltano fuori nei momenti più inaspettati. Immaginiamo di avere un borsellino contenente alcune monete da 1 euro e da 2 euro: vogliamo sapere in quanti modi si possono prelevare le monete dal borsellino per ottenere una determinata somma in euro. In un problema di questo tipo, l'ordine in cui si effettuano le azioni è importante. Per ottenere 4 euro si possono prelevare le monete in uno dei seguenti modi: 1 + 1 + 1 + 1, 2 + 1 + 1, 1 + 2 + 1, 1 + 1 + 2 e 2 + 2. In totale esistono cinque possibilità, che corrispondono al quinto numero di Fibonacci. Se si vogliono ottenere 20 euro si possono prelevare le monete da 1 e 2 euro in 6765 modi, corrispondenti al ventunesimo numero di Fibonacci. Questo esempio fornisce una dimostrazione dell'efficacia dei concetti matematici semplici.

La sezione aurea Osservando i rapporti che si ottengono dividendo ogni termine della successione di Fibonacci per il precedente, si scopre un'altra proprietà straordinaria di questi numeri. Proviamo con i termini 1, 1, 2, 3, 5, 8, 13, 21, 34, 55:

1/1	2/1	3/2	5/3	8/5	13/8	21/13	34/21	55/34
1,000	2,000	1,500	1,333	1,600	1,625	1,615	1,619	1,617

Il valore del rapporto si avvicina molto rapidamente a un valore chiamato «sezione aurea», un celebre numero della matematica che viene comunemente indicato con la lettera greca ϕ. La sezione aurea si colloca fra le principali costanti matematiche, come π ed e. Il suo valore esatto è

$$\phi = \frac{1+\sqrt{5}}{2}$$

che può essere approssimato con il numero decimale 1,618033988... Con qualche passaggio in più è possibile dimostrare che ogni numero di Fibonacci può essere espresso in termini di ϕ.

Dei numeri di Fibonacci si sanno molte cose, tuttavia numerose domande non hanno ancora trovato risposta. Ad esempio, si sa che la serie di Fibonacci contiene i numeri primi 2, 3, 5, 13, 89, 233, 1597, ma si ignora se ne contenga un numero infinito.

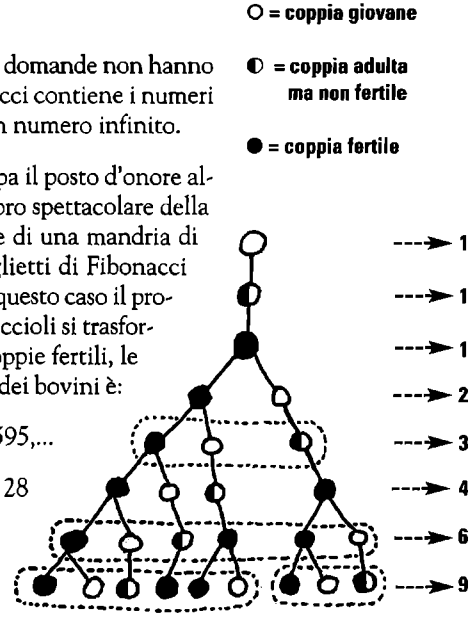
Un'aria di famiglia La successione di Fibonacci occupa il posto d'onore all'interno di un'ampia famiglia di sequenze analoghe. Un membro spettacolare della famiglia è associato a un problema riguardante la popolazione di una mandria di bovini. A differenza di ciò che accadeva alla coppia di coniglietti di Fibonacci (che nel giro di un mese diventa adulta e inizia a riprodursi), in questo caso il processo di crescita presenta uno stadio intermedio: le coppie di cuccioli si trasformano prima in coppie adulte ma non ancora fertili, e poi in coppie fertili, le uniche in grado di riprodursi. La successione della popolazione dei bovini è:

1, 1, 1, 2, 3, 4, 6, 9, 13, 19, 28, 41, 60, 88, 129, 189, 277, 406, 595,...

In questo caso la generazione salta un valore; ad esempio, 41 = 28 + 13 e 60 = 41 + 19. Questa successione ha proprietà analoghe a quelle dei numeri di Fibonacci. Nel caso della successione dei bovini, i rapporti che si ottengono dividendo ciascun termine per il precedente tendono al limite indicato dalla lettera greca ψ («psi»), in cui

$$\psi = 1,46557123187676802665\ldots$$

Si tratta della cosiddetta «sezione superaurea».



Popolazione di bovini

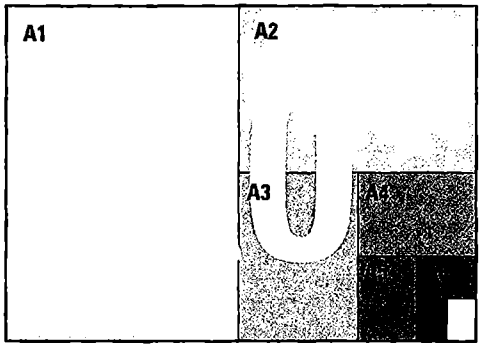
idea chiave

Il Codice da Vinci nell'ordine giusto

12 I rettangoli aurei

Gran parte di ciò che ci circonda ha una forma rettangolare: edifici, finestre, porte e perfino questo libro. Anche il mondo dell'arte è popolato da rettangoli di vario tipo, come quelli usati da Piet Mondrian, Ben Nicholson e altri esponenti dell'astrattismo. Qual è, allora, il rettangolo più bello? Uno lungo e sottile «alla Giacometti» o uno compatto, più simile a un quadrato? Oppure uno di un terzo tipo, intermedio fra questi due estremi?

Ha senso domandarsi quale sia il rettangolo più bello? Alcuni pensano di sì, perché ritengono che certi rettangoli corrispondano più di altri a un modello ideale; fra questi, quello che ha riscosso più successo è il rettangolo aureo. Infatti, fra tutti i rettangoli che si possono scegliere in virtù delle loro proporzioni (perché di questo si tratta, in fin dei conti), il rettangolo aureo rappresenta un caso molto particolare a cui si sono ispirati artisti, architetti e matematici. Ma prima vediamo qualche altro rettangolo.



Una piega matematica In un foglio di carta di formato A4, le cui dimensioni sono 210 mm (lato minore) per 297 mm (lato maggiore), il rapporto fra lunghezza e larghezza è 297/210, che è circa uguale a 1,4142. Un foglio di formato internazionale A e di larghezza b sarà sempre lungo 1,4142 × b; nel caso del formato A4, b = 210 mm, mentre per A5 b = 148 mm. Un foglio di carta appartenente alla serie di formati A definita da questa formula ha una proprietà molto interessante, che un foglio di dimensioni arbitrarie non possiede: piegandolo a metà si ottengono due rettangoli più piccoli che sono proporzionali al rettangolo di partenza. Si tratta cioè di due versioni più piccole dello *stesso* rettangolo.

linea del tempo

circa 300 a.C.
Negli *Elementi* Euclide parla di «rapporto estremo» e «rapporto medio»

1202
Leonardo Fibonacci pubblica il *Liber Abaci*

Piegando a metà un foglio A4 si ottengono due fogli A5, ciascuno dei quali può essere suddiviso nello stesso modo per produrre due A6. Andando verso dimensioni maggiori, un foglio di carta A3 è composto da due A4. Più piccolo è il numero che compare nel nome del formato, più grande è il foglio. Come facevamo a sapere che il numero 1,4142 avrebbe funzionato? Pieghiamo in due un rettangolo, ma questa volta supponiamo di non conoscerne la lunghezza, che indichiamo con x; se si ammette che abbia larghezza unitaria, il rapporto fra lunghezza e larghezza sarà $x/1$. Se ora lo pieghiamo a metà, il rapporto fra lunghezza e larghezza del rettangolo più piccolo è $1/(x/2)$, che equivale a $2/x$. Il vantaggio del formato A consiste nel fatto che la proporzione fra i due rapporti deve rimanere costante, perciò si ottiene l'equazione $x/1 = 2/x$, ovvero $x^2 = 2$. Il valore effettivo di x è $\sqrt{2}$, che si può approssimare con 1,4142.

Matematica aurea Il rettangolo aureo è diverso, ma non di molto. Questa volta si piega il rettangolo lungo la linea *RS* (vedi figura) in modo che i punti *MRSQ* siano i vertici di un quadrato.

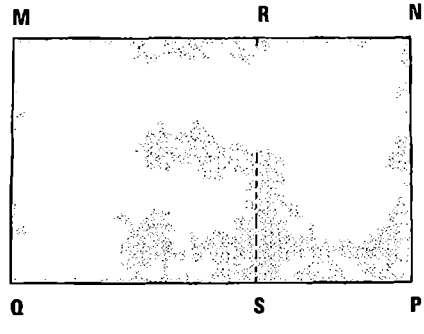
La proprietà fondamentale del rettangolo aureo è che il rettangolo più piccolo, *RNPS*, è proporzionale a quello più grande, *MNPQ*: il «residuo» dell'operazione di piegatura dovrebbe essere una versione in miniatura del rettangolo di partenza.

Come nel caso precedente, si impone che la larghezza $MQ = MR$ del rettangolo più grande sia unitaria e si indica la lunghezza del rettangolo (MN) con x. Anche in questo caso il rapporto lunghezza/larghezza è $x/1$. La larghezza del rettangolo più piccolo *RNPS* è $MN - MR$, che è uguale a $x - 1$, perciò il rapporto fra la lunghezza e la larghezza di questo rettangolo è $1/(x - 1)$. Uguagliando i due termini si ottiene l'equazione

$$\frac{x}{1} = \frac{1}{x-1}$$

Moltiplicando entrambi i termini per $x - 1$ si trova che $x^2 = x + 1$. Si può verificare facilmente che 1,618 è una soluzione approssimata di questa equazione: infatti, digitando 1,618 su una calcolatrice e moltiplicandolo per se stesso si ottiene 2,618, che è come dire che $x + 1 = 2{,}618$. Questo numero è la celebre «sezione aurea», che si indica con la lettera greca ϕ. La definizione e il valore approssimato della sezione aurea sono

$$\phi = \frac{1+\sqrt{5}}{2} = 1{,}61803398874989484820...$$



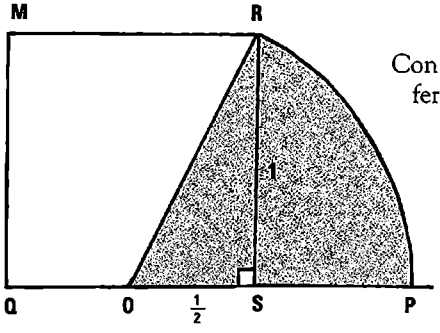
1509 Pacioli pubblica il *De divina proportione*

1876 Nei suoi scritti Fechner parla della possibilità di determinare le proporzioni del rettangolo più «estetico» per mezzo di esperimenti psicologici

1975 L'International Organization for Standardization (ISO) definisce il formato di carta A

Si tratta di un numero collegato alla successione di Fibonacci e al problema della riproduzione dei conigli (vedi capitolo 11).

La corsa all'oro Vediamo ora come si costruisce un rettangolo aureo. Partiamo dal quadrato di lato unitario $MQSR$ e indichiamo con O il punto medio di QS. Poiché il segmento OS è lungo $1/2$, applicando il teorema di Pitagora (vedi capitolo 21) al triangolo ORS si ricava che la lunghezza di OR è



$$\sqrt{\left(\frac{1}{2}\right)^2 + 1^2} = \frac{\sqrt{5}}{2}$$

Con un compasso centrato in O si può tracciare l'arco di circonferenza RP, dimostrando che $OP = OR = \sqrt{5}/2$. Si ottiene allora

$$QP = \frac{1}{2} + \frac{\sqrt{5}}{2} = \phi$$

che è la «sezione aurea», cioè la lunghezza del rettangolo aureo desiderato.

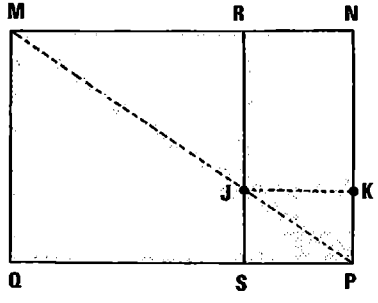
Un po' di storia Della sezione aurea ϕ si dicono molte cose. Dopo averne compreso le affascinanti proprietà matematiche può capitare di scorgerla nei luoghi più impensati, anche dove non c'è. Peggio ancora, si può affermare che abbia preceduto l'oggetto in cui si riscontra, cioè che musicisti, architetti e artisti ne abbiano tenuto conto nel creare le loro opere e questa idea può diventare una vera mania. In generale, trasformare una constatazione numerica in un'affermazione generale non suffragata da prove è un tipo di ragionamento estremamente rischioso.

Consideriamo l'esempio del Partenone di Atene: quando fu costruito, la sezione aurea era nota, ma ciò non significa che il tempio sia basato su questo numero. È vero che il rapporto fra la larghezza e l'altezza (compreso il frontone triangolare) della facciata è 1,74, un valore prossimo a 1,618, ma lo è a sufficienza per sostenere che rispetta la sezione aurea? Qualcuno sostiene che il frontone dovrebbe essere escluso dal calcolo; in questo caso il rapporto fra larghezza e altezza è uguale al numero intero 3.

Nel suo libro *De divina proportione* pubblicato nel 1509, Luca Pacioli «scoprì» alcuni collegamenti fra le proprietà di Dio e le caratteristiche della proporzione determinata da ϕ, che chiamò «proporzione divina». Pacioli, un monaco francescano autore di importanti testi di matematica, può essere considerato il padre della contabilità per aver contribuito alla diffusione del metodo della partita doppia usato dai mercanti veneziani. Viene anche ricordato perché insegnò matematica a Leonardo da Vinci. Nel Rinascimento la sezione aurea era oggetto di una venerazione quasi mistica e l'astronomo Johannes Kepler (Keplero) la definì un «gioiello prezioso». In seguito Gustav Fechner, uno psicologo sperimentale tedesco, effettuò migliaia di misure di figure rettangolari (carte da gioco, libri, finestre), scoprendo che in molti casi il rapporto fra i loro lati era prossimo a ϕ.

Le Corbusier era affascinato dal rettangolo, e in particolare dal rettangolo aureo, come elemento centrale del progetto architettonico e guardava l'architettura con gli occhi di un matematico: la matematica appagava la sua aspirazione all'ordine e all'armonia, due caratteristiche a cui attribuiva molta importanza. Uno dei punti di forza di Le Corbusier era il sistema «modulare», una teoria delle proporzioni che gli consentiva di generare una miriade di rettangoli aurei da utilizzare nei suoi progetti. In questo, Le Corbusier si ispirava a Leonardo da Vinci, che a sua volta aveva imparato dall'architetto romano Vitruvio a tenere in grande considerazione le proporzioni del corpo umano.

Altre figure Esiste anche un rettangolo «superaureo», la cui costruzione presenta alcune analogie con quella del rettangolo aureo.



Vediamo come si costruisce il rettangolo «superaureo» $MQPN$. Si parte anche in questo caso da un quadrato $MQSR$ di lato unitario. Dopo aver tracciato la diagonale MP e aver indicato con J il suo punto di intersezione con RS, si traccia una retta JK, parallela a RN, che interseca NP nel punto K. Chiamiamo y la lunghezza del segmento RJ e x quella di MN. Per ogni rettangolo, vale la relazione $RJ/MR = NP/MN$ (perché i triangoli MRJ e MNP sono simili), perciò $y/1 = 1/x$ ovvero $x \times y = 1$: si dice che x e y sono l'uno il «reciproco» dell'altro. Il rettangolo «superaureo» si ottiene imponendo la condizione che il rettangolo $RJKN$ sia proporzionale al rettangolo di partenza $MQPN$, cioè che $y/(x-1) = x/1$. Sfruttando la relazione $xy = 1$, si può concludere che la lunghezza x del rettangolo «superaureo» è una soluzione dell'equazione cubica $x^3 = x^2 + 1$, analoga all'equazione $x^2 = x + 1$ trovata nel caso del rettangolo aureo. L'equazione cubica ammette una soluzione reale positiva ψ (sostituendo x con il simbolo di uso comune ψ) il cui valore è

$$\psi = 1{,}465571231876768026\ldots$$

che è il numero associato con la successione dei bovini (vedi capitolo 11). Se il rettangolo aureo può essere costruito con riga e compasso, nel caso del rettangolo «superaureo» ciò è impossibile.

idea chiave

Proporzioni divine

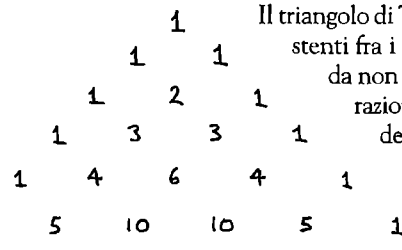
13 Il triangolo di Tartaglia

Il numero 1 è senz'altro molto importante, ma anche 11 è interessante, così come 11 × 11 = 121, 11 × 11 × 11 = 1331 e tutte le potenze superiori del numero 11. Scrivendole per esteso si ottiene

11
121
1331
14 641
15 101 051

Si tratta del triangolo di Tartaglia. Ma dove si trova?

Dopo averci aggiunto per sicurezza anche $11^0 = 1$, conviene spaziare un po' di più le cifre. In questo modo, 14 641 diventa 1 4 6 4 1.



Il triangolo di Tartaglia

Il triangolo di Tartaglia è famoso per la sua simmetria e per le relazioni nascoste esistenti fra i suoi termini; nel 1653 Blaise Pascal riteneva che fossero così tante da non poter essere trattate in una sola opera. Sebbene sia oggetto di venerazione soprattutto per i suoi numerosi collegamenti con altre branche della matematica, questo triangolo ha origini molto più antiche: non è stato inventato da Tartaglia, come si potrebbe pensare, ma era già noto agli studiosi cinesi del XIII secolo.

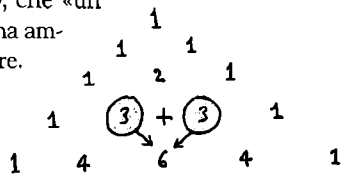
Lo schema di Tartaglia si costruisce a partire dall'alto. Si comincia scrivendo un 1 e se ne aggiungono altri due nella riga sottostante, a destra e a sinistra del primo. Per formare le righe successive si aggiungono altri 1 alle estremità e si ricava ciascun numero interno sommando i due numeri immediatamente soprastanti. Ad esempio, il 6 presente nella quinta riga si ottiene sommando i due 3 della

linea del tempo

circa 500 a.C.
Testi frammentari in sanscrito che parlano del triangolo di Tartaglia

circa 1070
Omar Khayyam scopre il triangolo che in alcuni paesi porta il suo nome

riga sopra. Se è vero, come disse il matematico inglese G.H. Hardy, che «un matematico, come il pittore e il poeta, è un creatore di forme», bisogna ammettere che quella del triangolo di Tartaglia è particolarmente regolare.

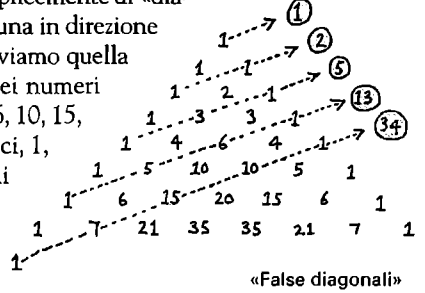


Collegamenti algebrici Il triangolo di Tartaglia è basato sulla matematica dura e pura. Ad esempio, svolgendo il calcolo $(1 + x) \times (1 + x) \times (1 + x) = (1 + x)^3$ si ottiene $1 + 3x + 3x^2 + x^3$. Se si guarda attentamente, si nota che i coefficienti che precedono i simboli in questa espressione sono uguali ai numeri presenti nella corrispondente riga del triangolo di Tartaglia. Lo schema è:

$(1 + x)^0$ 1
$(1 + x)^1$ 1 1
$(1 + x)^2$ 1 2 1
$(1 + x)^3$ 1 3 3 1
$(1 + x)^4$ 1 4 6 4 1
$(1 + x)^5$ 1 5 10 10 5 1

Sommando i numeri di ogni riga del triangolo di Tartaglia si ottiene sempre una potenza di 2; ad esempio, nella quinta riga dall'alto si ha $1 + 4 + 6 + 4 + 1 = 16 = 2^4$. Lo stesso risultato si ottiene sostituendo $x = 1$ nella colonna riportata qui sopra a sinistra.

Proprietà La prima e più evidente proprietà del triangolo di Tartaglia è la sua simmetria. Se si traccia una linea verticale che passa per il centro si può constatare che il triangolo presenta una «simmetria speculare», cioè è identico a destra e a sinistra della linea. Questa caratteristica consente di parlare semplicemente di «diagonali», perché una diagonale in direzione nord-est è uguale a una in direzione nord-ovest. Sotto la diagonale composta interamente da 1 troviamo quella costituita dai numeri naturali 1, 2, 3, 4, 5, 6... e poi quella dei numeri triangolari (rappresentabili con triangoli formati da punti) 1, 3, 6, 10, 15, 21,... Nella diagonale sottostante compaiono i numeri tetraedrici, 1, 4, 10, 20, 35, 56... che possono essere visualizzati come «triangoli tridimensionali» o come le palle di cannone che è possibile ammucchiare su basi triangolari di dimensioni crescenti. Prendiamo ora in esame le «false diagonali».



«False diagonali» nel triangolo di Tartaglia

Sommando i numeri del triangolo lungo le linee che lo attraversano, ma che non sono né righe né vere e proprie diagonali, si trova la successione 1, 2, 5, 13, 34, ...,

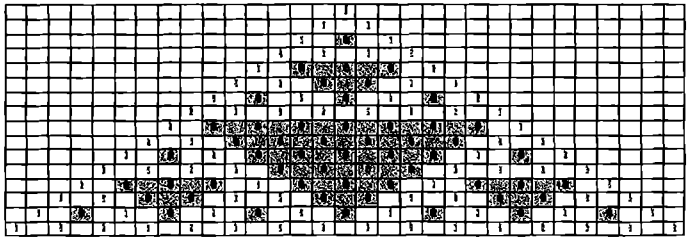
1303 Zhu Shijie definisce il triangolo di Tartaglia e dimostra come si sommano alcune successioni

1560 Niccolò Fontana (Tartaglia) pubblica un trattato in cui applica il triangolo a problemi di probabilità

1714 Leibniz prende in esame il triangolo armonico

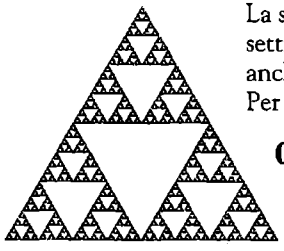
in cui ciascun numero si ottiene sottraendo dal triplo di quello che lo precede il numero ancora precedente; ad esempio, $34 = 3 \times 13 - 5$. Procedendo con lo stesso meccanismo, si trova che il numero successivo è $3 \times 34 - 13 = 89$. Abbiamo saltato le «false diagonali» alternative che cominciano con 1, $1 + 2 = 3$, ma queste ci danno la successione 1, 3, 8, 21, 55,... i cui numeri vengono prodotti con la stessa regola. Siamo quindi in grado di creare il numero successivo della sequenza, cioè $3 \times 55 - 21 = 144$. Ma c'è dell'altro. Se si uniscono, intervallandole, le due successioni di «false diagonali» si ottengono i numeri di Fibonacci:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144,...



Numeri pari e dispari nel triangolo di Tartaglia

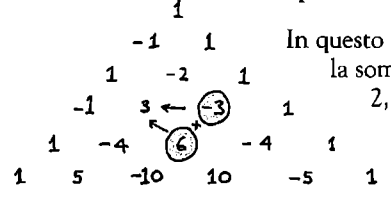
Combinazioni I numeri del triangolo di Tartaglia forniscono la soluzione di alcuni problemi di conteggio. Immaginiamo che in una stanza ci siano sette persone di nome Anna, Caterina, Emma, Guido, Luca, Matteo e Tommaso. In quanti modi si può scegliere un gruppo di tre persone fra queste sette? Una combinazione potrebbe essere A, C, E, un'altra A, C, T. Il numero che occupa la r-esima posizione (a partire da $r = 0$) della n-esima riga (a partire da $n = 0$) del triangolo di Tartaglia viene rappresentato dai matematici con la notazione $C(n, r)$. La soluzione del problema è $C(7, 3)$: il numero che occupa la terza posizione della settima riga è 35. Scegliendo un gruppo di tre persone, automaticamente si definisce anche un gruppo di quattro persone «non scelte», perciò vale anche $C(7, 4) = 35$. Per la simmetria speculare del triangolo di Tartaglia, $C(n, r) = C(n, n - r)$.



Il triangolo di Sierpiński

0 e 1 I numeri interni del triangolo di Tartaglia formano un «motivo» diverso a seconda che siano pari o dispari. Sostituendo i numeri dispari con 1 e i numeri pari con 0 si ottiene una rappresentazione che ha lo stesso schema di un celebre frattale, il «triangolo di Sierpiński» (vedi capitolo 25).

Introduzione dei segni Scriviamo il triangolo di Tartaglia che corrisponde alle potenze di $(-1 + x)$, cioè $(-1 + x)^n$.



Introduzione dei segni

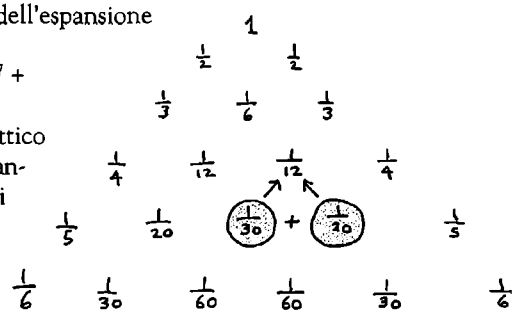
In questo caso il triangolo non presenta una simmetria speculare perfetta e la somma degli elementi di ciascuna riga non equivale a una potenza di 2, ma a zero. In questo caso, però, la cosa più interessante sono le diagonali. Gli elementi della diagonale nord-occidentale 1, –1, 1, –1, 1, –1, 1, –1,... sono i coefficienti dell'espansione

$$(1 + x)^{-1} = 1 - x + x^2 - x^3 + x^4 - x^5 + x^6 - x^7 +$$

mentre i termini della diagonale successiva sono quelli dell'espansione

$$(1 + x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + 5x^4 - 6x^5 + 7x^6 - 8x^7 +$$

Il triangolo armonico di Leibniz L'eclettico studioso tedesco Gottfried Leibniz scoprì un insieme triangolare di numeri dalle proprietà sorprendenti. Anche i numeri di Leibniz sono speculari rispetto a una linea verticale passante per il centro, ma a differenza di ciò che accade nel triangolo di Tartaglia, ogni numero si ottiene sommando due numeri della riga *sottostante*. Ad esempio, 1/30 + 1/20 = 1/12. Il triangolo si costruisce a partire dall'alto, procedendo da sinistra a destra per sottrazione: da 1/12 e 1/30 si ottiene 1/12 – 1/30 = 1/20, il numero a destra di 1/30. Avrete forse notato che la diagonale esterna è la famosa «serie armonica»



Il triangolo armonico di Leibniz

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \cdots$$

mentre la seconda diagonale è la cosiddetta «serie di Leibniz»

$$\frac{1}{1\times 2} + \frac{1}{2\times 3} + \cdots + \frac{1}{n\times(n+1)}$$

che dopo opportuni rimaneggiamenti risulta essere uguale a $n/(n + 1)$. Anche i numeri di Leibniz possono essere espressi nella notazione $B(n, r)$, dove $B(n, r)$ rappresenta l'r-esimo numero dell'n-esima riga, e sono collegati ai numeri di Tartaglia $C(n, r)$ dalla formula

$$B(n, r) \times C(n, r) = \frac{1}{n+1}$$

Il triangolo di Tartaglia è intimamente connesso con molte branche della matematica (come la geometria moderna, il calcolo combinatorio e l'algebra, per citarne solo tre), ma soprattutto illustra chiaramente come procede l'attività matematica: mediante una continua ricerca di schemi regolari e di corrispondenze che consentono di approfondire la comprensione dell'argomento.

idea chiave

La fontana dei numeri

14 L'algebra

Per risolvere i problemi algebrici si utilizza un particolare metodo deduttivo basato sul «ragionamento inverso». Se il problema consiste nel sommare 25 e 17 per ottenere 42, si segue un ragionamento diretto: basta prendere i dati del problema e sommarli. Se invece il risultato è noto e il problema richiede di trovare il numero che sommato a 25 dà 42, il ragionamento è inverso: si vuole trovare il valore di x che risolve l'equazione $25 + x = 42$ e per ricavarlo bisogna sottrarre 25 da 42.

Da secoli e secoli gli scolari si trovano a dover affrontare problemi algebrici con enunciati di questo tipo:

La mia nipotina Michela ha 6 anni e io ne ho 40.
Quando avrò il triplo della sua età?

In questo caso si potrebbe procedere per tentativi, ma operando per via algebrica si risparmia tempo. Fra x anni a partire da ora Michela avrà $6 + x$ anni e io ne avrò $40 + x$. Avrò il triplo della sua età quando

$$3 \times (6 + x) = 40 + x$$

Svolgendo il prodotto nel membro di sinistra dell'equazione si ottiene $18 + 3x = 40 + x$; se si raggruppano tutti i termini contenenti x da una parte e tutti quelli numerici dall'altra, si trova $2x = 22$, cioè $x = 11$. Quando Michela avrà 17 anni io ne avrò 51, il triplo della sua età. Fantastico!

Il problema potrebbe anche chiedere di stabilire quale sarà la mia età quando avrò il *doppio* dei suoi anni. Si può usare lo stesso approccio, questa volta risolvendo l'equazione

$$2 \times (6 + x) = 40 + x$$

da cui si ricava $x = 28$. Quando Michela avrà 34 anni, io ne avrò 68. Le equazioni riportate qui sopra sono della tipologia più semplice, perché non contengono

linea del tempo

1950 a.C. I Babilonesi utilizzano le equazioni quadratiche

250 Diofanto di Alessandria pubblica l'*Arithmetica*

termini come x^2 o x^3 che ne complicano la risoluzione, e sono dettte «equazioni lineari». Le equazioni in cui compaiono termini del tipo x^2 si chiamano «quadratiche» e quelle con termini del tipo x^3 sono dette «cubiche», perché un tempo x^2 veniva rappresentato con un quadrato e x^3 con un cubo.

Nel passaggio dall'aritmetica, o scienza dei numeri, all'algebra, o scienza dei simboli, la matematica ha subìto una profonda trasformazione. L'evoluzione da numeri a lettere rappresenta un notevole salto concettuale, ma vale la pena di farlo.

Le origini L'algebra occupava un posto rilevante nell'attività degli studiosi islamici del IX secolo. Al-Khwarizmi scrisse un manuale di matematica che conteneva la parola araba *al-jabr*: il termine «algebra» ci è quindi stato fornito dalla «scienza delle equazioni» con cui al-Khwarizmi esprimeva problemi concreti in termini di equazioni lineari e quadratiche. Alcuni secoli dopo, Omar Khayyam è diventato celebre per aver scritto le *Quartine* (*Rubaiyyàt*), con i versi immortali che dicono:

Se aver puoi sol per te un pane di bianco frumento,
due colme pinte di vino, un coscio d'agnello sugoso,
e qualcuna, dolce al cuore, in un paesaggio deserto;

ma nel 1070, a 22 anni, aveva scritto un libro di algebra in cui prendeva in esame le soluzioni delle equazioni cubiche.

La grande opera di matematica pubblicata nel 1545 da Girolamo Cardano, l'*Ars Magna*, ha segnato una svolta nella teoria delle equazioni, perché conteneva una grande messe di risultati sulle equazioni cubiche e quartiche (in cui compare un termine del tipo x^4). Da quella intensa attività di ricerca è scaturita la dimostrazione che le equazioni quadratiche, cubiche e quartiche potevano essere risolte per mezzo di formule in cui comparivano solo le operazioni $+, -, \times, \div, \sqrt[n]{}$ (dove l'ultimo simbolo indica la radice n-esima). Ad esempio, le equazioni quadratiche, la cui forma generale è $ax^2 + bx + c = 0$, possono essere risolte mediante la formula

Una questione italiana

La teoria delle equazioni cubiche fu sviluppata durante il Rinascimento, ma è associata a un episodio non proprio edificante. Quando l'insegnante veneziano Niccolò Fontana (soprannominato «Tartaglia» per la sua balbuzie) venne a sapere che Scipione Del Ferro aveva trovato la soluzione di alcune forme particolari di equazione cubica, pubblicò i propri risultati sull'algebra senza rendere noti i metodi utilizzati. Il milanese Girolamo Cardano convinse Tartaglia a rivelarglieli, giurando di mantenere il segreto, ma le informazioni finirono per trapelare e fra i due matematici scoppiò una lite quando Tartaglia scoprì che le sue ricerche erano state incluse nell'*Ars Magna*, il libro di Cardano pubblicato nel 1545.

825 Al-Khwarizmi parla di *al-jabr*, introducendo nella matematica il termine «algebra»

1591 François Viète scrive un testo matematico rappresentando termini noti e incognite per mezzo di lettere

anni '20 Emmy Noether pubblica alcuni saggi sullo sviluppo dell'algebra astratta moderna

1930 Bartel van der Waerden pubblica il suo famoso testo *Algebra moderna*

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Per trovare le soluzioni dell'equazione $x^2 - 3x + 2 = 0$ basta applicare la formula, sostituendo i valori $a = 1$, $b = -3$ e $c = 2$.

Le formule che consentono di risolvere le equazioni cubiche e quartiche sono lunghe e ingombranti, però esistono. Ciò che stupiva i matematici del passato era il fatto di non riuscire a trovare una formula analoga per le equazioni che contenevano x^5, le equazioni «di quinto grado». Cos'aveva di particolare quella potenza?

Nel 1826, tre anni prima della sua prematura scomparsa, Niels Abel risolse brillantemente l'enigma dell'equazione di quinto grado. In realtà fornì una dimostrazione negativa, che in genere è più difficile da produrre di una dimostrazione affermativa: dimostrò che non poteva esistere una formula in grado di risolvere qualsiasi equazione di quinto grado e concluse che ogni ulteriore ricerca sarebbe stata vana. L'*élite* dei matematici se ne convinse rapidamente, ma trascorse molto tempo prima che la notizia si diffondesse nella comunità matematica. Comunque, non tutti accettarono la validità del risultato di Abel e in pieno XIX secolo si continuavano a pubblicare opere che annunciavano la scoperta della formula inesistente.

I tempi moderni Per cinquecento anni l'algebra è stata «la teoria delle equazioni», ma nel XIX secolo ha subìto degli sviluppi che le hanno fatto prendere una nuova piega. In quel periodo si comprese infatti che, oltre ai numeri, i simboli algebrici potevano rappresentare qualcos'altro: ad esempio, potevano rappresentare delle «proposizioni», stabilendo un collegamento fra algebra e logica, oppure oggetti di dimensioni superiori, come quelli di cui si occupa l'algebra lineare (vedi capitolo 39). O forse, come molti non matematici sospettano da tempo, potevano anche non rappresentare un bel niente e non essere altro che simboli spostati di qua e di là in base a determinate regole (di natura formale).

Un evento importante per l'algebra moderna si verificò nel 1843, quando l'irlandese William Rowan Hamilton scoprì i quaternioni. Hamilton cercava un sistema simbolico in grado di estendere i numeri complessi bidimensionali in un numero maggiore di dimensioni e per molti anni si concentrò sui simboli tridimensionali, senza trovare alcun sistema soddisfacente. Tutte le mattine, quando si presentava a fare colazione e i figli gli chiedevano: «Allora, papà, sei riuscito a *moltiplicare* i tripletti?», Hamilton era costretto ad ammettere che era in grado solo di sommarli e di sottrarli.

La svolta arrivò in modo inaspettato: Hamilton capì che la ricerca tridimensionale era un vicolo cieco e che avrebbe dovuto dedicarsi ai simboli quadrimensionali. La folgorazione gli venne mentre passeggiava con la moglie lungo il Royal Canal, a Dublino. In preda all'euforia per la scoperta, quel vandalo trentottenne, Astronomo Reale d'Irlanda e Cavaliere del Regno, non esitò un istante a incidere le relazioni fondamentali nella pietra del Brougham Bridge, in un punto oggi indicato da una targa. Segnato da quell'evento memorabile, Hamilton sviluppò una vera ossessione per i simboli quadrimensionali. Tenne conferenze per anni e anni, e pubblicò due grossi tomi sull'argomento.

Una particolarità dei quaternioni è che nella loro moltiplicazione l'ordine dei termini è di vitale importanza, al contrario di ciò che avviene in aritmetica. Nel 1844 il linguista e matematico Hermann Grassmann pubblicò, con molto meno pathos, un altro sistema algebrico che all'epoca fu ignorato, ma che poi si è rivelato di ampia portata. Oggi sia i quaternioni che l'algebra di Grassmann trovano applicazione in geometria, in fisica e nella grafica computerizzata.

L'astrazione Nel corso del XX secolo, il paradigma dominante dell'algebra è stato il metodo assiomatico, che aveva costituito il fondamento della geometria euclidea, ma la cui applicazione all'algebra risale a tempi relativamente recenti.

Emmy Noether è stata la regina dell'algebra astratta, una moderna versione dell'algebra improntata allo studio strutturale, in cui le specificità dei singoli casi sono subordinate al concetto astratto generale. Esempi specifici dotati della stessa struttura ma espressi in una notazione diversa sono detti «isomorfi».

La struttura algebrica fondamentale è il «gruppo», definito da una lista di assiomi (vedi capitolo 38). Esistono strutture definite da un numero di assiomi minore (come monoidi, semigruppi e quasigruppi) e altre che ne richiedono un numero maggiore (come anelli, anelli di divisione, domini di integrità e campi). Tutti questi termini sono stati introdotti nella matematica all'inizio del XX secolo, quando l'algebra si è trasformata in una scienza astratta chiamata «algebra moderna».

idea chiave

15 L'algoritmo di Euclide

Oltre a dare al mondo l'«algebra», Al-Khwarizmi ha introdotto nel suo testo di aritmetica del IX secolo l'«algoritmo», un concetto utile tanto ai matematici quanto agli informatici. Cos'è un algoritmo? La risposta a questa domanda ci mette sulla buona strada per comprendere l'algoritmo euclideo della divisione.

Un algoritmo è innanzi tutto un programma, cioè una serie di istruzioni del tipo «fai questo e poi fai quello». È più che comprensibile che i computer amino così tanto gli algoritmi, visto che seguire le istruzioni senza mai sgarrare è la loro specialità. Gli algoritmi possono sembrare noiosi per via della loro ripetitività, ma scrivere un algoritmo e tradurlo in centinaia di righe di codice contenente istruzioni matematiche non è un'impresa da poco, e il rischio di compiere errori devastanti è altissimo. La scrittura di un algoritmo richiede molta creatività: capita spesso di dover operare una scelta fra metodi diversi per effettuare una certa azione e può accadere che alcuni algoritmi non «vadano bene» o che siano talmente tortuosi da risultare inefficienti, oppure che siano rapidi ma forniscano il risultato sbagliato. È un po' come cucinare: esistono centinaia di ricette (algoritmi) per preparare il ragù e non vogliamo certo sceglierne una scadente proprio quando abbiamo a cena un ospite di riguardo! Si parte da una serie di ingredienti e una lista di istruzioni. Una ricetta semplificata potrebbe essere:

- preparare un battuto di cipolla, sedano e carota
- farlo soffriggere nell'olio
- unire la carne macinata e farla rosolare a fuoco vivo
- aggiungere pomodoro, sale, pepe e lasciare cuocere a fuoco basso per due ore

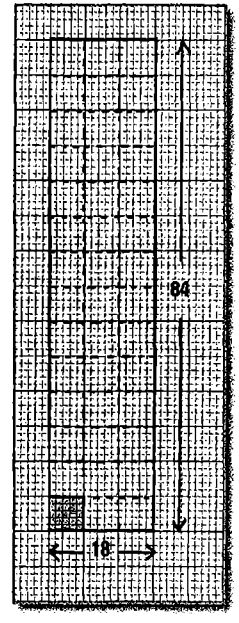
Non dobbiamo fare altro che eseguire l'algoritmo in una serie di passi successivi. L'unica cosa che manca in questa ricetta, e che in genere è presente in un algoritmo matematico, è un *loop*, un mezzo che consente di introdurre la ricorsione. Si spera di non dover preparare il ragù due volte di seguito.

linea del tempo

circa 300 a.C.
L'algoritmo di Euclide viene pubblicato nel Libro VII degli *Elementi*

circa 300
Sun Tzu scopre il teorema cinese del resto

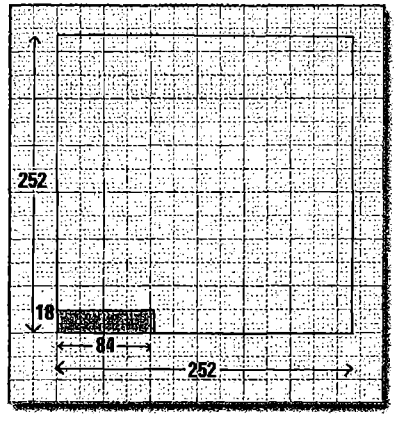
Anche la matematica ha i suoi ingredienti: i numeri. L'algoritmo della divisione di Euclide è concepito per calcolare il massimo comun divisore (MCD) di due numeri, ovvero il numero più grande che divide entrambi. A titolo di esempio, scegliamo come ingredienti i numeri 18 e 84.



Il massimo comun divisore Il MCD è il numero più grande che divide esattamente sia 18 che 84. Il numero 2 li divide entrambi, ma ciò vale anche per il 3, perciò anche 6 sarà un divisore dei due numeri. Esiste un numero ancora più grande in grado di dividerli? Potremmo provare con 9 o con 18, ma è facile verificare che non sono divisori di 84, quindi il numero più grande che divide entrambi i numeri di partenza è 6. La conclusione che 6 è il massimo comun divisore di 18 e 84 può essere espressa nella forma MCD(18, 84) = 6.

Si può fornire un'interpretazione del MCD facendo l'esempio delle piastrelle: il massimo comun divisore è il lato della piastrella quadrata più grande con cui è possibile rivestire una parete rettangolare larga 18 e alta 84 senza dover tagliare nessuna piastrella. È facile verificare che una piastrella di dimensioni 6 × 6 va benissimo.

Il massimo comun divisore è chiamato anche «massimo fattore comune» ed è collegato al concetto del «minimo comune multiplo» (*mcm*). Il *mcm* di 18 e 84 è il più piccolo numero divisibile sia per 18 che per 84. Il collegamento fra *mcm* e MCD risulta evidente quando si considera che moltiplicare il *mcm* di due numeri per il loro MCD equivale a moltiplicare fra loro i numeri stessi. Nel caso in questione, $mcm(18, 84) = 252$ ed è facile verificare che $6 \times 252 = 1512 = 18 \times 84$.



Rivestimento di una parete quadrata con piastrelle rettangolari di dimensioni 18 x 84

Geometricamente, il *mcm* corrisponde al lato del quadrato più piccolo che può essere rivestito di piastrelle rettangolari di dimensioni 18 × 84. Poiché $mcm(a, b) = ab \div MCD(a, b)$, ci concentreremo sul calcolo del MCD. Abbiamo già visto che MCD(18, 84) = 6, ma per ottenere questo risultato abbiamo dovuto trovare i divisori sia di 18 che di 84.

Ricapitolando, prima di tutto si scompongono in fattori entrambi i numeri: $18 = 2 \times 3 \times 3$ e $84 = 2 \times 2 \times 3 \times 7$. Confrontandoli, si vede che 2 è un fattore comune e che è la più alta potenza di 2 in grado di dividerli. Allo stesso modo, 3 è un

810 Al-Khwarizmi introduce il termine «algoritmo»

1202 Fibonacci pubblica nel *Liber Abaci* i suoi risultati sulle congruenze

anni '70 Il teorema cinese del resto viene applicato alla codifica dei messaggi

fattore comune ed è la potenza di 3 più alta per cui possono essere divisi, mentre 7 è un divisore di 84, ma non di 18, perciò non può essere un fattore del MCD. In conclusione, $2 \times 3 = 6$ è il numero più grande che divide entrambi. Si può evitare di fare questo gioco delle tre carte con i fattori? È facile immaginare quali calcoli saltano fuori quando si cerca di calcolare il MCD di (17 640, 54 054) con questo metodo: innanzi tutto si devono scomporre in fattori entrambi i numeri, e non parliamo del seguito... Deve esistere un metodo più semplice.

L'algoritmo In effetti, un metodo più semplice esiste. L'algoritmo di Euclide compare negli *Elementi*, libro VII, proposizione 2: «Dati due numeri che non siano primi fra loro, trovare il loro massimo comun divisore».

L'algoritmo fornito da Euclide è molto efficiente e sostituisce la penosa ricerca dei fattori comuni con una semplice sottrazione. Vediamo come funziona.

Si vuole calcolare $d = \text{MCD}(18, 84)$. Si inizia dividendo 84 per 18; la divisione non è esatta e si ottiene 4 con il resto di 12:

$$84 = \mathbf{4} \times 18 + 12$$

Essendo un divisore di 84 e 18, d deve dividere anche il resto 12, perciò $d = \text{MCD}(12, 18)$. Ripetendo il procedimento, si divide 18 per 12

$$18 = 1 \times 12 + 6$$

e si ottiene il resto 6, cosicché $d = \text{MCD}(6, 12)$. La divisione di 12 per 6 dà resto 0, perciò $d = \text{MCD}(0, 6)$. Il numero più grande che divide sia 0 che 6 è 6: è il risultato che cercavamo.

Nel calcolo di $d = \text{MCD}(17\,640, 54\,054)$, i resti successivi sono 1134, 630, 504, 126 e 0, cosicché $d = 126$.

Applicazioni di MCD Il MCD può essere utilizzato per risolvere le equazioni che ammettono come soluzioni dei numeri interi. Queste equazioni sono dette «diofantee», dal nome dell'antico matematico Diofanto di Alessandria.

Immaginiamo che la vecchia zia Cristina sia in procinto di partire per la sua consueta vacanza alle Canarie. Il domestico Giovanni porta all'aeroporto le valigie, ognuna delle quali pesa 18 o 84 kilogrammi, e viene informato che il peso complessivo dei bagagli ammonta a 652 kilogrammi. Quando torna a casa, il figlioletto Mario di nove anni obietta: «Non è possibile, perché 652 non è divisibile per 6», e suggerisce che il peso reale potrebbe essere 642 kilogrammi.

Mario sa che l'equazione $18x + 84y = c$ ammette come soluzione dei numeri interi se e solo se il MCD 6 è un divisore del numero c: per $c = 652$ questa condizione non è soddisfatta, mentre lo è per $c = 642$. Mario non ha bisogno di sapere quante valigie x, y dei due pesi intende portare zia Cristina.

Il teorema cinese del resto Quando il MCD di due numeri è 1 si dice che i numeri sono «primi tra loro». Non è necessario che siano primi, ma che siano primi l'uno rispetto all'altro; ad esempio, MCD(6, 35) = 1, anche se né 6 né 35 sono primi. Questa definizione è utile per il teorema cinese del resto.

Consideriamo un altro problema. Alberto non sa quante bottiglie di vino ha in casa, ma quando le abbina a due a due gliene resta una. Quando poi le ripone nell'apposita rastrelliera in file di cinque, nota che ne avanzano tre. Quante bottiglie di vino possiede Alberto? Si sa che dividendo il numero incognito per 2 si ottiene un resto di 1 e dividendolo per 5 si ottiene un resto di 3. La prima condizione consente di escludere tutti i numeri pari. Scorrendo l'elenco dei numeri dispari, si scopre che 13 soddisfa tutte le condizioni (anche 3 le soddisfa, ma si può ragionevolmente supporre che Alberto possieda più di tre bottiglie). Però esistono altri numeri che andrebbero ugualmente bene; di fatto, tutta la serie che comincia con 13, 23, 33, 43, 53, 63, 73, 83, ...

Aggiungiamo ora un'altra condizione, ovvero che il numero desiderato deve dare un resto di 3 anche quando viene diviso per 7 (le bottiglie sono arrivate in confezioni da 7, più 3 singole). Esaminando la serie 13, 23, 33, 43, 53, 63,... si vede che 73 soddisfa tutte le condizioni, ma che lo stesso vale per 143, 213 e per tutti gli altri numeri che si ottengono sommando un multiplo di 70 a questi numeri.

In termini matematici, si può dire di aver trovato le soluzioni previste dal teorema cinese del resto, che afferma anche che ogni coppia di soluzioni differisce per un multiplo di $2 \times 5 \times 7 = 70$. Se Alberto possiede un numero di bottiglie compreso fra 150 e 250, il teorema consente di stabilire che la soluzione è 213. Niente male per un teorema scoperto nel IV secolo.

idea chiave

Puntare massimo

16 La logica

«Se ci fossero meno auto in circolazione, l'inquinamento sarebbe accettabile. Quindi si potrebbe ridurre il numero di auto in circolazione o imporre il pagamento di un pedaggio stradale, o entrambe le cose. Se imponiamo un pedaggio, l'estate sarà incredibilmente "calda". Ma l'estate si sta rivelando piuttosto tranquilla, perciò la conclusione è certa: l'inquinamento è accettabile».

Questo ragionamento presente nell'editoriale di un quotidiano è «valido» o è illogico? Non ci interessa appurare se espone una proposta sensata ai fini della riduzione del traffico o se è un buon articolo; ci interessa solo capire se si tratta di un ragionamento valido dal punto di vista logico. Per stabilirlo bisogna ricorrere alla logica, che si occupa precisamente di valutare il rigore dei ragionamenti.

Due premesse e una conclusione Il brano in questione è piuttosto complicato, perciò prima di affrontarlo sarà meglio esaminare qualche ragionamento più semplice, rifacendoci nientemeno che al filosofo greco Aristotele di Stagira, considerato il padre della logica. L'approccio di Aristotele si basava sulle diverse forme assunte dal sillogismo, un ragionamento formato da tre affermazioni: due premesse e una conclusione. Un esempio di sillogismo può essere:

Ogni cocker è un cane
Ogni cane è un animale

Ogni cocker è un animale

dove la linea separa le premesse dalla conclusione. In questo caso la conclusione suona in una certa misura inevitabile, indipendentemente dal significato attribuito alle parole «cocker», «cani» e «animale». Lo stesso sillogismo, ma espresso con parole diverse è

linea del tempo

circa 335 a.C. Aristotele formalizza la logica del sillogismo

1847 Boole pubblica *L'analisi matematica della logica*

Ogni mela è un'arancia
Ogni arancia è una banana

Ogni mela è una banana

Se si attribuisce alle parole il loro significato comune, ciascuna di queste tre affermazioni è priva di senso. Tuttavia entrambi gli esempi presentano la stessa struttura ed è la struttura a rendere valido il sillogismo. Se A, B e C rispettano questa struttura (vedi qui a lato), non è possibile che le premesse siano vere ma la conclusione sia falsa. È questa caratteristica a far sì che un ragionamento valido sia anche utile.

Ogni A è B
Ogni B è C

Ogni A è C

Un ragionamento valido

Utilizzando vari quantificatori, come «tutti», «alcuni» e «nessuno» (ad esempio «nessun A è B»), si ottengono molti sillogismi diversi. Un altro esempio potrebbe essere

Alcuni A sono B
Alcuni B sono C

Alcuni A sono C

È un ragionamento valido? È applicabile a tutti i possibili A, B e C, o da qualche parte si annida un controesempio, un caso in cui le premesse sono vere ma la conclusione è falsa? Verifichiamo, ad esempio, se il sillogismo ottenuto sostituendo A con «cocker», B con «oggetti marroni» e C con «tavoli» è plausibile:

Alcuni cocker sono marroni
Alcuni oggetti marroni sono tavoli

Alcuni cocker sono tavoli

Questo controesempio dimostra che il sillogismo di partenza *non* è valido. Le possibili tipologie di sillogismi sono così tante che per ricordarle gli studiosi medievali hanno introdotto delle parole mnemoniche, le cui tre vocali corrispondono alle «quantità» delle premesse e delle conclusioni. Ad esempio, al primo tipo di sillogismo corrisponde la parola BARBARA, perché vi compare tre volte «ogni» (il caso più affermativo, indicato con l'iniziale della parola latina *Affirmo*). Questi metodi di analisi dei ragionamenti sono rimasti in vigore per più di duemila anni e hanno occupato un posto importante nei corsi di studi delle università medievali. Nel XIX secolo si continuava a ritenere che la logica di Aristotele (la sua teoria del sillogismo) fosse una scienza perfetta.

1910
Russell e Whitehead cercano di ridurre la matematica alla logica

1965
Lofti Zadeh sviluppa la logica *fuzzy*

1987
Il sistema di frenata di una metropolitana giapponese si basa sulla logica *fuzzy*

Logica proposizionale Un tipo di logica diversa, che supera i sillogismi, è quella che si occupa delle proposizioni (affermazioni semplici) e delle loro combinazioni. Per analizzare il brano dell'articolo bisogna conoscere almeno i rudimenti di questa «logica proposizionale». In passato veniva chiamata «algebra della logica», perché George Boole aveva capito che si poteva trattarla come un nuovo tipo di algebra (e questo ci dice qualcosa riguardo alla sua struttura). Fra il 1840 e il 1850 la logica conobbe un grande sviluppo, grazie ai risultati ottenuti da Boole, Augustus De Morgan e altri matematici.

a	b	a ∨ b
V	V	V
V	F	V
F	V	V
F	F	F

Tabella di verità dell'OR

Vediamo come funziona la logica proposizionale. Consideriamo una proposizione **a**, dove **a** sta per «Ugo è un cocker». La proposizione **a** può essere Vera o Falsa: se penso al mio cane Ugo, che in effetti è un cocker, la proposizione è vera (V), ma se la riferisco a mio cugino, anche lui di nome Ugo, è falsa (F). Il fatto che una proposizione sia vera o falsa dipende da ciò a cui si riferisce.

a	b	a ∧ b
V	V	V
V	F	F
F	V	F
F	F	F

Tabella di verità dell'AND

Se abbiamo un'altra proposizione **b** del tipo «Stella è un gatto», possiamo combinarla con la prima in vari modi. Una possibile combinazione si indica con **a** ∨ **b**, dove il connettivo ∨ si legge «OR» (la congiunzione disgiuntiva inglese corrispondente al nostro «o»). In logica, però, OR assume un significato un po' diverso rispetto all'uso comune. Infatti, **a** ∨ **b** è vera se *una delle due proposizioni* «Ugo è un cocker» e «Stella è un gatto» è vera *oppure* se sono vere *entrambe*, ed è falsa solo se sono false sia **a** che **b**. I vari casi possibili per questa congiunzione di proposizioni possono essere riassunti in una tabella di verità.

a	¬ a
V	F
F	V

Tabella di verità del NOT

Per combinare le proposizioni si possono usare anche i connettivi «AND», indicato con **a** ∧ **b**, e «NOT», che si scrive ¬ **a**. L'aspetto algebrico di questo tipo di logica diventa evidente quando si combinano le proposizioni **a**, **b** e **c** con vari connettivi, come nel caso **a** ∧ (**b** ∨ **c**). Si può ottenere anche un'equazione che chiamiamo «identità»:

a	b	a → b
V	V	V
V	F	F
F	V	V
F	F	V

Tabella di verità dell'implicazione

$$\mathbf{a} \wedge (\mathbf{b} \vee \mathbf{c}) \equiv (\mathbf{a} \wedge \mathbf{b}) \vee (\mathbf{a} \wedge \mathbf{c})$$

Il simbolo ≡ indica l'equivalenza fra enunciati logici che si verifica quando entrambi i membri dell'equivalenza hanno la stessa tabella di verità. Esiste un parallelismo fra l'algebra della logica e l'algebra ordinaria, perché i simboli ∧ e ∨ agiscono in modo analogo ai simboli × e + dell'algebra, dove si ha $x \times (y + z) = (x \times y) + (x \times z)$; tuttavia si tratta di un parallelismo imperfetto, che ammette varie eccezioni.

A partire da questi connettivi fondamentali se ne possono definire molti altri. Uno particolarmente utile è il connettivo «di implicazione» **a** → **b**, che per definizione è equivalente a ¬ **a** ∨ **b** e per cui si può scrivere la tabella di verità riportata a margine. Tornando all'articolo del quotidiano, ora siamo in grado di esprimerlo in forma simbolica in modo da ottenere il ragionamento scritto a lato:

A = meno Auto sulle strade
I = l'Inquinamento sarà accettabile
P = dovrebbe esserci un Piano per introdurre un pedaggio
C = l'estate sarà terribilmente Calda

$$\begin{array}{r} A \rightarrow I \\ A \vee P \\ P \rightarrow C \\ \neg C \\ \hline I \end{array}$$

Cerchiamo di stabilire se si tratta di un ragionamento valido o meno. Assumiamo che la conclusione I sia falsa, ma che *tutte* le premesse siano vere. Se si riesce a dimostrare che ciò conduce a una contraddizione, il ragionamento deve essere valido, perché si sarà dimostrata l'impossibilità del fatto che le premesse siano vere e la conclusione falsa. Se I è falsa, dalla premessa iniziale $A \rightarrow I$ risulta che A deve essere falsa. Poiché $A \vee P$ è vera, il fatto che A sia falsa significa che P è vera. Per la terza premessa $P \rightarrow C$, ciò significa anche che C è vera, cioè che $\neg C$ è falsa: ciò contraddice l'assunzione che $\neg C$, l'ultima premessa, fosse vera. Il contenuto delle affermazioni dell'editoriale resta molto discutibile, ma la struttura del ragionamento è valida.

Logiche diverse Gottlob Frege, C.S. Peirce e Ernst Schröder introdussero la quantificazione nella logica proposizionale, costruendo una logica «predicativa del primo ordine» e del «secondo ordine» (con l'introduzione di variabili). Questa logica fa uso del quantificatore universale $\forall$, che significa «per ogni», e del quantificatore esistenziale $\exists$, che si legge «esiste».

$\vee$	or
$\wedge$	and
$\neg$	not
$\rightarrow$	implica
$\forall$	per ogni
$\exists$	esiste

Un altro sviluppo recente è la logica *fuzzy* («sfumata»). Il nome può far pensare che si tratti di un modo di ragionare piuttosto confuso, ma in realtà si tratta di un ampliamento dei confini tradizionali della logica. La logica tradizionale si basa sugli insiemi, o collezioni di elementi; esiste l'insieme dei cocker, l'insieme dei cani e l'insieme degli oggetti marroni, e non c'è margine di dubbio riguardo all'appartenenza o meno di un elemento a un certo insieme: se al parco vediamo un levriero afgano, siamo certi che non appartiene all'insieme dei cocker.

La teoria degli insiemi *fuzzy* si occupa di insiemi definiti in modo vago. Se esistesse l'insieme dei cocker pesanti, quanto dovrebbe pesare un cocker per farne parte? Nel caso degli insiemi *fuzzy* esiste un «valore di appartenenza» e il confine che separa ciò che è incluso nell'insieme da ciò che è escluso è sfumato. La matematica, comunque, consente di precisare anche le sfumature. La logica è un campo tutt'altro che arido: rispetto ai tempi di Aristotele ha subìto una notevole evoluzione e oggi è una branca della matematica estremamente dinamica sia dal punto di vista scientifico che applicativo.

idea chiave

Un ragionamento limpido e rigoroso

17 La dimostrazione

I matematici cercano sempre di avvalorare ciò che affermano per mezzo di dimostrazioni. La continua ricerca di ragionamenti ferrei è la forza trainante della matematica pura: mediante una concatenazione di deduzioni corrette il matematico procede da ciò che sa, o assume, verso una conclusione che poi entra a far parte del patrimonio della matematica.

Non è facile ottenere una dimostrazione, perché in genere si presenta solo dopo un bel po' di ricerche e di giri a vuoto, e lo sforzo per produrla occupa una posizione centrale nell'attività dei matematici. Una dimostrazione riuscita porta il sigillo di autenticità del matematico e distingue un teorema affermato da una congettura; un'idea brillante o una prima supposizione.

Le qualità che si ricercano in una dimostrazione sono il rigore, la limpidezza e (non ultima) l'eleganza, a cui va aggiunta la capacità di fornire nuovi elementi di comprensione. Una buona dimostrazione è «una dimostrazione che ci insegna qualcosa», ma è sempre meglio avere una dimostrazione di un qualche tipo che non averne alcuna. Procedendo sulla base di fatti non dimostrati si corre il pericolo di costruire le proprie teorie sull'equivalente matematico della sabbia.

Non si deve pensare, però, che una dimostrazione duri necessariamente in eterno: può darsi che debba essere corretta a causa dell'evoluzione dei concetti a cui fa riferimento.

Cos'è una dimostrazione? Cosa ci induce a credere (o a non credere) a un nuovo risultato matematico che leggiamo o di cui sentiamo parlare? Una risposta potrebbe essere: ci crediamo se esiste un ragionamento logico inoppugnabile che procede dalle idee che consideriamo vere all'affermazione riguardo alla quale ci stiamo interrogando. È ciò che i matematici chiamano una dimostrazione, che in genere è un miscuglio di linguaggio quotidiano e logica stringente. In base alla qualità della dimostrazione si può essere pienamente convinti della sua validità o restare scettici.

linea del tempo

circa 300 a.C.
Gli *Elementi* di Euclide forniscono un modello di dimostrazione matematica

1637
Nel suo *Discorso sul metodo*, Descartes fa assurgere il rigore matematico a modello

In matematica si usano vari tipi di dimostrazione: il metodo del controesempio, il metodo diretto, il metodo indiretto e il metodo dell'induzione matematica.

Il controesempio Supponiamo di essere scettici; è uno dei metodi per dimostrare che un'affermazione non è corretta. A titolo di esempio, consideriamo un'affermazione specifica: qualcuno ha affermato che moltiplicando qualunque numero per se stesso si ottiene un numero pari. Ci crediamo? Prima di precipitarci a rispondere sarà bene fare qualche verifica. Consideriamo, ad esempio, il numero 6: moltiplicandolo per se stesso si ottiene 36, che in effetti è un numero pari. Ma una rondine non fa primavera. L'affermazione parlava di *qualunque* numero e i numeri sono un'infinità. Per prendere confidenza con il problema bisogna considerare qualche altro esempio. Provando con 9, si trova che $9 \times 9 = 81$, che è un numero dispari. Ciò significa che l'affermazione che *ogni* numero moltiplicato per se stesso produce un numero pari è falsa. Un esempio che è in grado di contraddire l'affermazione di partenza si chiama «controesempio». Un controesempio dell'affermazione «tutti i cigni sono bianchi» potrebbe consistere nell'avvistamento di un cigno nero. Uno degli aspetti divertenti della matematica è la ricerca di un controesempio in grado di demolire un potenziale teorema.

Se non si riesce a trovare un controesempio si può pensare che l'affermazione sia corretta. A questo punto il matematico cambia strategia: è il momento di costruire una dimostrazione, e il metodo più semplice è quello diretto.

Il metodo diretto Il metodo diretto consiste nel procedere con un ragionamento logico da ciò che è stato accertato, o che si assume come vero, alla conclusione. Se ciò è possibile, si è ottenuto un teorema. Non è possibile dimostrare che moltiplicando ogni numero per se stesso si ottiene un numero pari, perché abbiamo già dimostrato che ciò è falso. Ma forse si può salvare qualcosa: la differenza fra il numero del primo esempio, 6, e quello del controesempio, 9, è che il primo è pari, mentre il secondo è dispari. Niente ci vieta di modificare l'ipotesi, perciò ora l'affermazione diventa: moltiplicando un numero *pari* per se stesso, si ottiene un numero pari.

Provando con altri numeri, si scopre che questa affermazione è sempre verificata e non è possibile trovare un controesempio. Allora si cambia linea di condotta e si tenta di dimostrare che è vera. Ma da dove conviene partire? Si potrebbe iniziare con un generico numero pari n, ma forse è un po' troppo astratto; vediamo invece come si può costruire una dimostrazione considerando un numero specifico, ad esempio il 6. Com'è noto, un numero pari è un multiplo di 2, cioè $6 = 2 \times 3$. Poiché

1838
De Morgan introduce il termine «induzione matematica»

1967
Bishop dimostra alcuni risultati con metodi esclusivamente costruttivi

1976
Imre Lakatos pubblica l'importante opera *Dimostrazioni e confutazioni*

$6 \times 6 = 6 + 6 + 6 + 6 + 6 + 6$, si può scrivere $6 \times 6 = 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3$ oppure, utilizzando le parentesi

$$6 \times 6 = 2 \times (3 + 3 + 3 + 3 + 3 + 3)$$

Ciò significa che 6×6 è un multiplo di 2 e, come tale, è un numero pari. Questo ragionamento però non contiene niente che riguardi specificamente il numero 6; avremmo potuto benissimo cominciare con $n = 2 \times k$, in modo da ottenere

$$n \times n = 2 \times (k + k + \quad + k)$$

e giungere alla conclusione che $n \times n$ è pari. A questo punto la dimostrazione è completa. I matematici del passato avevano l'abitudine di scrivere «Q.E.D.» alla fine di una dimostrazione, per dire «è fatta!»: si tratta di un'abbreviazione della frase latina *quod erat demonstrandum* («come volevasi dimostrare», o C.V.D.). Oggi si usa anche un quadratino pieno ■ che qualcuno chiama «halmos», dal nome di Paul Halmos che l'ha introdotto.

Il metodo indiretto Questo metodo consiste nel fingere che la conclusione sia falsa, per poi dimostrare con un ragionamento logico che ciò contraddice l'ipotesi. Utilizziamolo per dimostrare il risultato trovato con il metodo diretto.

L'ipotesi è che n sia pari: si finge quindi che $n \times n$ sia dispari. Se scriviamo $n \times n = n + n + \quad + n$, la somma è composta da n termini; ciò significa che n non può essere pari (altrimenti anche $n \times n$ sarebbe pari). Perciò n è dispari, il che contraddice l'ipotesi. ■

In realtà questa è una versione piuttosto blanda del metodo indiretto. Nella sua forma completa, il metodo indiretto prende il nome di «metodo della *reductio ad absurdum*» (dimostrazione per assurdo). La *reductio ad absurdum* era il ragionamento prediletto dai Greci; nell'Accademia di Atene, Socrate e Platone amavano dimostrare una questione controversa invischiando i loro avversari in un groviglio di contraddizioni, al di fuori del quale si trovava la conclusione a cui miravano. La dimostrazione classica del fatto che la radice quadrata di 2 è un numero irrazionale ha proprio questa forma: si parte dall'assunzione che sia razionale per poi arrivare a contraddire tale assunzione.

Il metodo dell'induzione matematica L'induzione matematica è un metodo molto efficace per dimostrare che tutte le affermazioni che costituiscono una sequenza P_1, P_2, P_3,... sono vere. L'induzione matematica (da non confondere con l'induzione utilizzata nelle scienze sperimentali) fu definita in modo rigoroso fra il 1830 e il 1840 da Augustus De Morgan, che formalizzò un procedimento già utilizzato da secoli, anche se in modo confuso. Si tratta di una tecnica di uso comune per dimostrare le affermazioni che riguardano i numeri *interi* ed è particolarmente utile nella teoria dei grafi, nella teoria dei numeri e più in generale in informatica. Per fare un esempio pratico, supponiamo di voler calcolare la somma dei numeri

dispari. La somma dei primi tre numeri dispari è $1 + 3 + 5 = 9$, quella dei primi quattro è $1 + 3 + 5 + 7 = 16$. Se 9 è uguale a $3 \times 3 = 3^2$ e 16 è uguale a $4 \times 4 = 4^2$, si può ipotizzare che la somma dei primi n numeri dispari sia uguale a n^2. Provando a verificare l'ipotesi per un valore di n, ad esempio per $n = 7$, in effetti si trova che la somma dei primi sette numeri dispari è $1 + 3 + 5 + 7 + 9 + 11 + 13 = 49$, che è uguale a 7^2. Ma chi ci assicura che questa regola valga per ogni valore di n? È un bel problema, perché non possiamo verificare un numero infinito di casi.

A questo punto entra in gioco l'induzione matematica. Volendone dare una descrizione informale, si può dire che è un metodo di dimostrazione che sfrutta l'«effetto domino». La metafora si riferisce a una fila di tessere del domino ritte sulle loro estremità: cadendo, ogni tessera abbatte la successiva e basta far cadere la prima perché cadano *tutte*. Lo stesso ragionamento può essere applicato al problema dei numeri dispari. L'affermazione P_n dice che la somma dei primi n numeri dispari è uguale a n^2. L'induzione matematica innesca una reazione a catena per cui P_1, P_2, P_3,... sono *tutte* vere. L'affermazione P_1 è sicuramente vera, perché $1 = 1^2$. Quindi, P_2 è vera perché $1 + 3 = 1^2 + 3 = 2^2$, P_3 è vera perché $1 + 3 + 5 = 2^2 + 5 = 3^2$ e P_4 è vera perché $1 + 3 + 5 + 7 = 3^2 + 7 = 4^2$. Si usa il risultato di ogni stadio per passare al successivo. Questo procedimento può essere formalizzato fino a legittimare il metodo di dimostrazione per induzione, o per ricorrenza.

Difficoltà di dimostrazione Esistono dimostrazioni di tutti i tipi. Alcune sono brevi ed eleganti, come quelle che compaiono nei manuali, mentre altre, che descrivono in dettaglio i risultati più recenti, occupano interi numeri di riviste e sono costituite da migliaia di pagine. In questi casi, solo una piccola minoranza di matematici è in grado di seguire il ragionamento per intero.

Esistono anche questioni fondazionali. Ad esempio, alcuni matematici non sono d'accordo nell'utilizzare il metodo della *reductio ad absurdum* per dimostrare l'esistenza di qualcosa. Se ipotizzando che un'equazione non ammetta soluzione si arriva a un'assurdità, non è detto che ciò sia sufficiente a dimostrare l'esistenza di una soluzione. I critici di questo metodo sostengono che la logica è semplicemente un gioco di prestigio e non ci dice come costruire una soluzione concreta. Questi matematici «costruttivisti» (di varie gradazioni) sostengono che il metodo di dimostrazione per assurdo è privo di «significato numerico» e irridono i matematici classici che lo considerano uno strumento fondamentale. D'altra parte, i matematici più legati alla tradizione obiettano che bandire questo tipo di ragionamento equivale a operare con una mano sola e che, se si rinnegassero tutti i risultati dimostrati con questo metodo indiretto, l'arazzo della matematica apparirebbe assai sfilacciato.

idea chiave

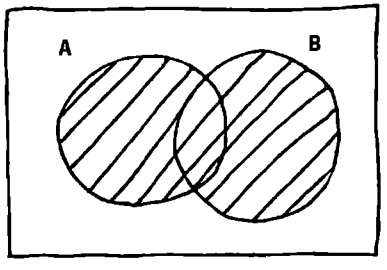
Come volevasi dimostrare

18 Gli insiemi

Nicholas Bourbaki era lo pseudonimo di un gruppo autocostituito di accademici francesi che intendevano riscrivere la matematica da cima a fondo «nel modo giusto» e affermavano arditamente che doveva basarsi solo sulla teoria degli insiemi. Tutto ruotava attorno al metodo assiomatico e i loro testi seguivano una struttura rigorosa costituita da «definizione, teorema e dimostrazione». Ciò costituiva anche la sostanza della matematica degli anni '60 del Novecento.

Georg Cantor creò la teoria degli insiemi per dotare la teoria dei numeri reali di una base solida. Nonostante i pregiudizi e le critiche di cui fu oggetto inizialmente, all'inizio del XX secolo la teoria degli insiemi era ampiamente riconosciuta come una branca della matematica.

Che cosa sono gli insiemi? Un insieme può essere considerato una collezione di oggetti. È una definizione informale, ma esprime bene il concetto di fondo. Gli oggetti sono detti «elementi» o «membri» dell'insieme. Se A è un insieme che ha come membro a, Cantor ci dice che si può scrivere $a \in A$. Un esempio può essere l'insieme $A = \{1, 2, 3, 4, 5\}$; l'appartenenza all'insieme si indica con $1 \in A$, mentre la non appartenenza si esprime con $6 \notin A$.



Unione di A e di B

Gli insiemi possono essere combinati in due modi fondamentali. Se A e B sono due insiemi, l'insieme costituito dagli elementi che sono membri di A o di B (o di entrambi) è detto «unione» dei due insiemi. In forma matematica l'unione si scrive $A \cup B$ e può essere descritta anche per mezzo di un diagramma di Venn. I diagrammi di questo tipo prendono il nome dal reverendo John Venn, un logico dell'epoca vittoriana, ma Eulero li aveva utilizzati già molto tempo prima.

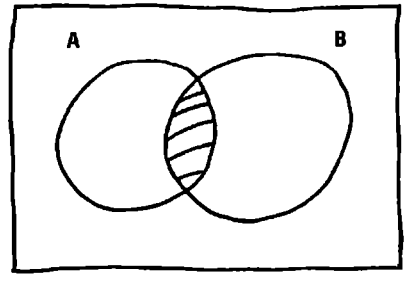
L'insieme $A \cap B$ è costituito da elementi che appartengono *sia* ad A che a B ed è detto «intersezione» dei due insiemi.

linea del tempo

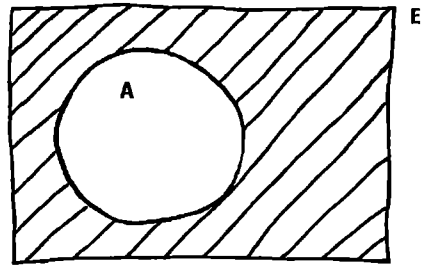
1872 Cantor fa un primo passo verso la creazione della teoria degli insiemi

1881 Venn diffonde l'uso degli omonimi diagrammi per rappresentare gli insiemi

Se $A = \{1, 2, 3, 4, 5\}$ e $B = \{1, 3, 5, 7, 10, 21\}$, l'unione $A \cup B$ è $\{1, 2, 3, 4, 5, 7, 10, 21\}$ e l'intersezione $A \cap B$ è $\{1, 3, 5\}$. Ammettendo che A faccia parte di un insieme universale E, si può definire l'insieme complemento $\neg A$ come l'insieme costituito dagli elementi di E che *non* appartengono ad A.



Intersezione di A e di B



Complemento di A

Le operazioni fra insiemi $\cap$ e $\cup$ sono analoghe alle operazioni algebriche $\times$ e $+$. Aggiungendo anche l'operazione complemento $\neg$, si ottiene un'«algebra degli insiemi». Il matematico britannico di origine indiana Augustus De Morgan formulò delle leggi che dimostravano in che modo interagiscono queste tre operazioni. Nella notazione attuale, le leggi di De Morgan sono:

$$\neg (A \cup B) = (\neg A) \cap (\neg B)$$

e

$$\neg (A \cap B) = (\neg A) \cup (\neg B)$$

I paradossi Quando si ha a che fare con insiemi finiti non sorge alcun problema, perché se ne possono sempre elencare gli elementi come abbiamo fatto con $A = \{1, 2, 3, 4, 5\}$, ma all'epoca di Cantor gli insiemi infiniti apparivano più interessanti.

Cantor definì gli insiemi come collezioni di elementi dotati di una determinata proprietà. Consideriamo l'insieme $\{11, 12, 13, 14, 15,...\}$ costituito da tutti i numeri interi maggiori di 10. Trattandosi di un insieme infinito è impossibile elencarne tutti gli elementi, ma lo si può comunque definire sulla base della proprietà condivisa da tutti i suoi elementi. Seguendo le orme di Cantor, possiamo scrivere l'insieme come $A = \{x : x \text{ è un numero intero} > 10\}$, dove i due punti si leggono «tale che».

Nella sua forma originaria, la teoria degli insiemi ammetteva anche la possibilità di avere un insieme di oggetti astratti $A = \{x : x \text{ è un oggetto astratto}\}$. In questo caso anche A sarebbe un oggetto astratto, perciò potrebbe valere $A \in A$. Se si ammette che valga questa relazione, però, sorgono dei problemi gravissimi. Il filosofo britannico Bertrand Russell ideò un insieme S che conteneva tutti gli oggetti che non contenevano se stessi. In forma simbolica, ciò si traduce in $\{x : x \notin x\}$.

1931
Gödel dimostra che ogni sistema matematico assiomatico formale contiene delle proposizioni indecidibili

1939
Adozione dello pseudonimo Bourbaki da parte dei matematici francesi

1964
Cohen dimostra l'indipendenza dell'ipotesi del continuo

Poi si pose la domanda: «$S \in S$?». Se la risposta è «sì», S deve soddisfare la condizione che definisce S, perciò $S \notin S$. D'altra parte, se la risposta è «no» e $S \notin S$, S *non* soddisfa la definizione di $S = \{x : x \notin x\}$, perciò $S \in S$. La domanda posta da Russell portò alla seguente affermazione, che costituisce il paradosso di Russell:

$$S \in S \text{ se e solo se } S \notin S$$

Si tratta di un risultato analogo al «paradosso del barbiere», in cui il barbiere di un villaggio annuncia agli abitanti che raderà soltanto coloro che non si radono da soli. Il problema è: il barbiere dovrà radersi da solo? Se non si rade dovrebbe farlo, se si rade no.

I paradossi di questo tipo, chiamati eufemisticamente «antinomie», vanno evitati a ogni costo, perché i matematici considerano inammissibile che un sistema generi delle contraddizioni. Russell creò una teoria dei tipi in cui la relazione $a \in A$ era ammessa solo se a era di tipo inferiore ad A; in questo modo si evitava che comparissero espressioni come $S \in S$.

Un altro modo per evitare queste antinomie consisteva nel formalizzare la teoria degli insiemi. In questo approccio non ci si preoccupa della natura degli insiemi, ma si elencano gli assiomi formali che specificano le regole che consentono di trattarli. Gli antichi Greci tentarono di fare qualcosa di analogo con uno dei loro problemi: non si preoccuparono di spiegare che cosa fossero le rette, ma solo come bisognava trattarle.

Nel caso della teoria degli insiemi ciò dette origine agli assiomi di Zermelo-Fraenkel, che impedivano la comparsa di insiemi troppo «grandi» nel sistema. In questo modo si scongiurava il pericolo di veder spuntare entità pericolose come l'insieme di tutti gli insiemi.

Il teorema di Gödel Il matematico austriaco Kurt Gödel sferrò un colpo decisivo a coloro che cercavano di sfuggire ai paradossi rifugiandosi nei sistemi assiomatici formali. Nel 1931 Gödel dimostrò che, anche per i sistemi formali più semplici, esistevano delle proposizioni che non potevano essere né dimostrate né confutate internamente al sistema. In parole povere, tali proposizioni si collocavano fuori della portata degli assiomi del sistema: erano indecidibili. Per questo motivo il teorema di Gödel è detto anche «teorema di incompletezza». Questo risultato valeva sia per il sistema di Zermelo-Fraenkel che per altri sistemi.

I numeri cardinali Il numero di elementi di un insieme finito può essere determinato facilmente; ad esempio, $A = \{1, 2, 3, 4, 5\}$ ha 5 elementi. Si può anche dire che la sua «cardinalità» è 5 e scrivere $card\ (A) = 5$. In sostanza, la cardinalità misura le «dimensioni» di un insieme.

In base alla teoria degli insiemi di Cantor, l'insieme delle frazioni Q e quello dei numeri reali R sono molto diversi. Se è possibile elencare gli elementi dell'insieme Q, nel caso dell'insieme R ciò è impossibile (vedi capitolo 7). Entrambi sono infiniti, ma R ha un ordine di infinità superiore a Q. I matematici indicano *card* (Q) con $\aleph_0$, l'«alef zero» ebraica e *card* (R) con *c*, perciò $\aleph_0 < c$.

L'ipotesi del continuo L'ipotesi del continuo fu proposta da Cantor nel 1878 e afferma che il livello di infinità successivo a quello di Q è quello dei numeri reali *c*. In altre parole, questa ipotesi dice che non esiste alcun insieme di cardinalità strettamente compresa fra $\aleph_0$ e *c*. Cantor tentò a lungo di dimostrarla, ma, benché fosse convinto della sua veridicità, non ci riuscì. Per confutarla avrebbe dovuto trovare un sottoinsieme X di R per cui $\aleph_0 < card(X) < c$, ma non riuscì a fare neanche questo.

Questo problema era così importante che il matematico tedesco David Hilbert lo mise al primo posto nella famosa lista dei ventitré problemi da risolvere nel secolo a venire che presentò al Congresso internazionale di matematica di Parigi del 1900.

Gödel era assolutamente convinto che l'ipotesi del continuo fosse falsa, ma non riuscì a dimostrarlo; dimostrò invece (nel 1938) che era compatibile con gli assiomi di Zermelo-Fraenkel della teoria degli insiemi. Venticinque anni dopo, Paul Cohen sbigottì Gödel e i logici dimostrando che l'ipotesi del continuo non poteva essere dedotta dagli assiomi di Zermelo-Fraenkel. Ciò equivaleva a dimostrare che gli assiomi e la negazione dell'ipotesi non erano in contraddizione. Il risultato di Cohen, insieme a quello ottenuto da Gödel nel 1938, dimostrava che l'ipotesi del continuo era indipendente dagli assiomi della teoria degli insiemi.

Tale indipendenza è di natura analoga all'indipendenza del postulato geometrico delle parallele (vedi capitolo 27) dagli altri assiomi euclidei. Quella scoperta portò alla fioritura delle geometrie non-euclidee che, fra le altre cose, hanno reso possibile l'elaborazione della teoria della relatività di Einstein. Allo stesso modo, l'ipotesi del continuo può essere accettata o rifiutata senza perturbare gli altri assiomi della teoria degli insiemi. Dal pionieristico risultato di Cohen è nato un nuovo campo di studi capace di affascinare intere generazioni di matematici, che hanno adottato le tecniche da lui utilizzate per dimostrare l'indipendenza dell'ipotesi del continuo.

idea chiave

Tanti trattati come uno solo

19 Il calcolo differenziale

In senso lato, un calcolo è un procedimento che serve a calcolare qualcosa, perciò i matematici parlano di «calcolo logico», «calcolo della probabilità», e così via. Tutti però concordano sul fatto che il calcolo in sé, puro e semplice, è il calcolo differenziale, ed è unico.

Il calcolo differenziale è un caposaldo della matematica. Le sue applicazioni sono talmente varie che oggi è praticamente impossibile che uno scienziato, un ingegnere o un economista quantitativo non vi si siano imbattuti. Storicamente il calcolo differenziale è associato a Isaac Newton e Gottfried Leibniz, che nel XVII secolo ne furono i pionieri. I due grandi scienziati elaborarono teorie simili e si disputarono l'attribuzione della paternità del calcolo differenziale, ma in realtà ciascuno era giunto alle proprie conclusioni in modo indipendente e utilizzando un metodo diverso.

Da allora il calcolo differenziale si è ampliato enormemente. Ogni generazione elabora alcune tecniche che ritiene di dover trasmettere alla successiva, perciò oggi i manuali superano tranquillamente il migliaio di pagine e sono corredati da una grande quantità di materiale aggiuntivo. Ma, aggiunte a parte, gli elementi davvero indispensabili sono la *derivazione* (o *differenziazione*) e l'*integrazione*, le vette gemelle del calcolo istituito da Newton e Leibniz. Questi termini derivano, rispettivamente, dal *calculus differentialis* (per il calcolo delle «differenze») e dal *calculus integralis* (o *sommatorius*, per il calcolo delle «somme») di Leibniz.

Detto in linguaggio specialistico, la derivazione ha a che fare con la misura di una *variazione* e l'integrazione con la misura di un'*area*, ma la perla del calcolo differenziale, il risultato principe, è il fatto che si tratta di due facce della stessa medaglia: la derivazione e l'integrazione sono l'una l'inverso dell'altra. Il calcolo differenziale è un unico argomento, perciò bisogna conoscerle entrambe.

linea del tempo

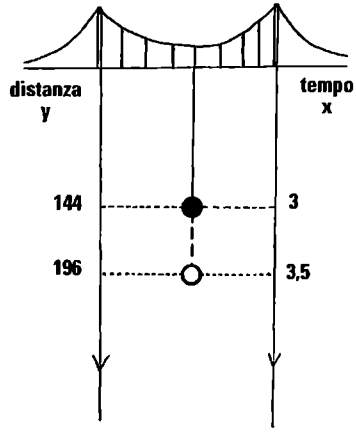
circa 450 a.C.
Zenone ridicolizza le quantità infinitesimali con un paradosso

1660-1670
Newton e Leibniz iniziano a sviluppare il calcolo differenziale

Derivazione Gli scienziati hanno una vera passione per gli «esperimenti mentali» (Einstein li amava in modo particolare). Immaginiamo allora di trovarci su un ponte sospeso su una profonda gola e di lasciar cadere un sasso. Cosa succede? Il vantaggio di effettuare un esperimento mentale è che non c'è bisogno di trovarsi realmente sul posto; inoltre si possono fare cose che normalmente sono impossibili, come fermare il sasso a mezz'aria o seguirlo al rallentatore per un breve intervallo di tempo.

Come previsto dalla teoria della gravitazione di Newton, il sasso cade. Fin qui, niente di strano: il sasso viene attratto verso il suolo e cade sempre più velocemente all'avanzare della lancetta del cronometro. Un altro vantaggio di un esperimento di questo tipo è che si possono ignorare le complicazioni, come la resistenza dell'aria.

Vogliamo sapere qual è la velocità del sasso in un certo istante, ad esempio quando il cronometro segna *esattamente* 3 secondi dall'avvio. Vediamo come si calcola. È senz'altro possibile misurare la velocità *media*, ma a noi interessa quella *istantanea*. Dato che si tratta di un esperimento mentale, possiamo fermare la pietra in volo per poi farla scendere per un tratto molto piccolo, aspettando una frazione di secondo in più: dividendo la distanza aggiuntiva per il tempo aggiuntivo si ottiene la velocità media calcolata sull'intervallino di tempo in più. Se si considerano intervalli temporali sempre più brevi, la velocità media tende a coincidere con quella istantanea nel punto in cui abbiamo fermato il moto del sasso. Questo «passaggio al limite» è il concetto di base del calcolo differenziale.



Potremmo essere tentati di uguagliare a zero il tempo aggiuntivo. In questo caso, però, il sasso dell'esperimento mentale non si sarebbe mosso affatto: non avrebbe percorso alcuna distanza e non avrebbe impiegato nessun tempo per farlo! Come velocità media si otterrebbe allora la quantità 0/0 che il filosofo irlandese Bishop Berkeley definì «fantasmi di quantità defunte». Si tratta di un'espressione che non può essere determinata, perché *non ha senso*. Seguendo questa via si finisce per impantanarsi in una palude numerica.

Per procedere oltre bisogna ricorrere ai simboli. La formula che collega la distanza percorsa durante la caduta y e il tempo x necessario per percorrerla fu ricavata da Galileo:

1734	1820-1830	1854	1902
Berkeley ne sottolinea le debolezze fondazionali	Cauchy formalizza la teoria in modo rigoroso	Riemann introduce l'omonimo integrale	Lebesgue elabora la teoria dell'omonimo integrale

$$y = 4{,}903 \times x^2$$

Il valore del fattore numerico (4,903) dipende dal fatto di aver scelto come unità di misura i metri e i secondi. Per calcolare, ad esempio, quale distanza ha percorso il sasso in 3 secondi di caduta libera, basta sostituire $x = 3$ nella formula e svolgere il calcolo: $y = 4{,}903 \times 3^2 = 44{,}127$ metri. Ma come si calcola la velocità del sasso al tempo $x = 3$?

Aspettiamo altri 0,5 secondi e vediamo che distanza ha percorso il sasso nel tempo compreso fra 3 e 3,5 secondi. In 3,5 secondi ha percorso $y = 4{,}903 \times 3{,}5^2 = 60{,}062$ metri, perciò nel tempo *compreso fra* 3 e 3,5 secondi è precipitato per $60{,}062 - 44{,}127 = 15{,}935$ metri. Poiché la velocità si ottiene dividendo la distanza per il tempo, la velocità media in questo intervallo di tempo è $15{,}935/0{,}5 = 31{,}87$ metri al secondo. Questo valore sarà prossimo a quello della velocità istantanea per $x = 3$, ma si potrebbe obiettare che 0,5 secondi non è un intervallo di tempo sufficientemente piccolo. Ripetendo il ragionamento con 0,05 secondi, ad esempio, si vede che il sasso ha percorso $45{,}610 - 44{,}127 = 1{,}483$ metri a una velocità media di $1{,}483/0{,}05 = 29{,}66$ metri al secondo. Questo valore si avvicina di più alla velocità istantanea del sasso dopo 3 secondi (quando $x = 3$).

Ora è il momento di prendere il toro per le corna e calcolare la velocità media del sasso nel tempo compreso fra x secondi e un tempo immediatamente successivo, $x + h$ secondi. Con qualche rimaneggiamento, si ottiene

$$4{,}903 \times (2x) + 4{,}903 \times h$$

Riducendo progressivamente h, come abbiamo fatto quando siamo passati da 0,5 a 0,05 secondi, si vede che il primo termine resta inalterato (perché non contiene h), mentre il secondo diventa sempre più piccolo. Si può quindi concludere che

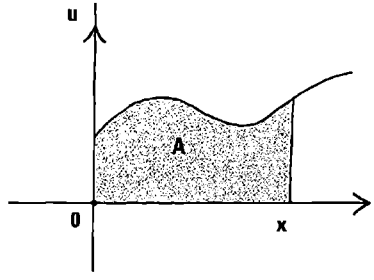
$$v = 4{,}903 \times (2x)$$

in cui v è la velocità istantanea del sasso al tempo x. Ad esempio, la velocità istantanea dopo 1 secondo (quando $x = 1$) è $4{,}903 \times (2 \times 1) = 9{,}806$ metri al secondo; dopo 3 secondi è $4{,}903 \times (2 \times 3)$, ovvero 29,418 metri al secondo.

u	du/dx
x^2	$2x$
x^3	$3x^2$
x^4	$4x^3$
x^5	$5x^4$
...	...
x^n	nx^{n-1}

Confrontando la formula della distanza di Galileo $y = 4{,}903 \times x^2$ con quella della velocità $v = 4{,}903 \times (2x)$, si nota che la differenza principale consiste nella trasformazione di x^2 in $2x$. È l'effetto della derivazione, che trasforma $u = x^2$ nella sua *derivata* $\dot{u} = 2x$. Newton chiamava $\dot{u} = 2x$ «flussione» e la variabile x «fluente», perché ragionava in termini di flussi di quantità. Oggi in genere si scrive $u = x^2$ mentre la derivata assume la forma $du/dx = 2x$. Il fatto che la notazione introdotta da Leibniz sia tuttora in uso rappresenta l'affermazione delle «d» di Leibniz sui «punti» di Newton. La caduta del sasso è solo uno dei tanti esempi possibili; se disponessimo di

altre espressioni corrispondenti a u potremmo calcolarne la derivata allo stesso modo e ciò potrebbe tornarci utile in altre occasioni. La regola per calcolare la derivata delle funzioni x^n è: si diminuisce di 1 l'esponente della potenza di x e si moltiplica tutto per n.



Integrazione Inizialmente l'integrazione veniva usata soprattutto per la misura delle aree. Per misurare l'area sottesa da una curva si suddivide l'area stessa in tante striscioline più o meno rettangolari, ciascuna di larghezza dx. Misurando l'area di ognuna e addizionandole si ottiene la «somma», ovvero l'area complessiva. La notazione della S allungata, $\int$, per rappresentare la somma fu introdotta da Leibniz. L'area di ogni strisciolina rettangolare è $u dx$, perciò l'area A sottesa dalla curva fra 0 e x è

$$A = \int_0^x u\,dx$$

Se la curva in questione è $u = x^2$, l'area si ricava disegnando tante striscioline rettangolari sotto la curva, sommandole per calcolare l'area approssimata e riducendo la larghezza delle striscioline fino a passare al limite, in modo da ottenere il valore esatto dell'area. Il risultato è

$$A = x^3/3$$

È possibile calcolare l'integrale di molte altre curve (e quindi di molte altre espressioni di u). Anche in questo caso, come in quello della derivata, esiste una regola per calcolare l'integrale delle funzioni x^n: si aumenta di 1 l'esponente della potenza di x e si divide tutto per questo numero.

	$\int_0^x u\,dx$
x^2	$x^3/3$
x^3	$x^4/4$
x^4	$x^5/5$
x^5	$x^6/6$
	...
x^n	$x^{n+1}/(n+1)$

Il risultato principe Derivando l'integrale $A = x^3/3$ si ottiene l'espressione di partenza, $u = x^2$. La derivazione è l'inverso dell'integrazione: questa osservazione costituisce il «teorema fondamentale del calcolo differenziale», uno dei teoremi più importanti di tutta la matematica.

Se il calcolo differenziale non esistesse, non ci sarebbero satelliti in orbita attorno al pianeta, non esisterebbero le teorie economiche e la statistica sarebbe una disciplina completamente diversa. Ogni volta in cui entra in gioco una variazione, c'è di mezzo il calcolo differenziale.

idea chiave

Al limite

20 Le costruzioni geometriche

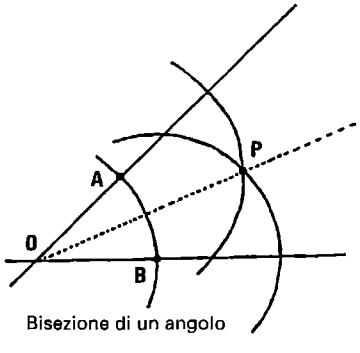
Dimostrare una negazione in genere è molto difficile, ma è così che sono stati ottenuti alcuni dei più importanti risultati matematici. In pratica si tratta di dimostrare l'impossibilità di fare qualcosa. La quadratura del cerchio è un'impresa impossibile: come si può dimostrare?

Gli antichi Greci si ponevano quattro importanti problemi di costruzione geometrica:

- la trisezione dell'angolo (dividere un angolo in tre angoli uguali più piccoli)
- la duplicazione del cubo (costruire un cubo di volume doppio rispetto a uno dato)
- la quadratura del cerchio (costruire un quadrato di area equivalente a quella di un dato cerchio)
- la costruzione di poligoni (costruire poligoni regolari con lati e angoli uguali)

che affrontavano con gli strumenti strettamente necessari:

- una riga con cui tracciare delle rette (ma non misurare delle lunghezze),
- un compasso con cui disegnare delle circonferenze.



Bisezione di un angolo

Chi ama scalare montagne senza corde, ossigeno, cellulari e altre diavolerie apprezzerà sicuramente questi problemi. In mancanza degli strumenti di misura moderni la dimostrazione dei problemi di costruzione classici richiedeva l'impiego di tecniche matematiche sofisticate, perciò tali problemi sono stati risolti solo nel XIX secolo grazie alle tecniche dell'analisi moderna e dell'algebra astratta.

Trisezione dell'angolo Vediamo ora un metodo che consente di suddividere un angolo in due angoli uguali, cioè di «bisecarlo». Si punta un compasso nel punto O e, con un'apertura a piacere, si individuano i punti OA e OB. Poi si sposta la punta del compasso

linea del tempo

450 a.C.
Mentre si trova in prigione, Anassagora tenta di quadrare il cerchio

1672
Mohr dimostra che per realizzare una qualsiasi costruzione euclidea è sufficiente un compasso

in A e si traccia un arco di circonferenza. Si ripete l'operazione nel punto B, indicando con P il punto di intersezione delle due circonferenze, quindi si uniscono O e P con una riga. I triangoli AOP e BOP sono simili e quindi gli angoli AÔP e BÔP sono uguali. La retta OP è la bisettrice desiderata, cioè la retta che suddivide l'angolo in due angoli uguali.

È possibile suddividere un angolo di ampiezza arbitraria in *tre* angoli uguali compiendo una serie di azioni di questo tipo? È il problema della trisezione dell'angolo.

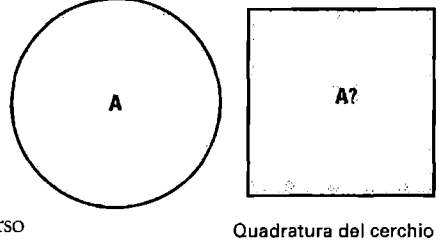
La trisezione di un angolo di 90 gradi (angolo retto) non presenta alcuna difficoltà, perché è possibile costruire l'angolo di 30 gradi. Ma un angolo di 60 gradi, ad esempio, non può essere trisecato. È evidente che si otterrebbero tre angoli di 20 gradi, ma non esiste alcun metodo che consenta di costruire un angolo di questa ampiezza usando solo una riga e un compasso. Quindi, riassumendo:

- è *sempre* possibile bisecare *qualsiasi* angolo
- è *sempre* possibile trisecare *determinati* angoli, ma
- non è *mai* possibile trisecare *altri* angoli.

Un problema analogo è quello della duplicazione del cubo, noto anche come «problema di Delo». Narra la leggenda che quando gli abitanti di Delo, in Grecia, interrogarono l'oracolo sul modo di liberarsi dalla pestilenza da cui erano afflitti, questi rispose di costruire un nuovo altare di volume doppio rispetto a quello esistente.

Supponiamo che inizialmente l'altare di Delo fosse un cubo tridimensionale di lato a. Si trattava di costruire un altro cubo, di lato b e di volume doppio. I volumi dei due cubi sono, rispettivamente, a^3 e b^3 e sono collegati dalla relazione $b^3 = 2a^3$ ovvero $b = \sqrt[3]{2} \times a$ (dove $\sqrt[3]{2}$ è la radice cubica di 2, cioè il numero che moltiplicato tre volte per se stesso dà 2). Se il lato del cubo di partenza era $a = 1$, gli abitanti di Delo avrebbero dovuto segnare su una retta la lunghezza $\sqrt[3]{2}$. Purtroppo per loro, era impossibile farlo utilizzando solo una riga non graduata e un compasso, indipendentemente dall'ingegno dimostrato nella costruzione geometrica.

La quadratura del cerchio È di gran lunga il più noto dei problemi di costruzione classici ed è leggermente diverso dagli altri:



Quadratura del cerchio

1801 Gauss pubblica le *Disquisizioni aritmetiche*

1837 Wantzel dimostra che i problemi classici della duplicazione del cubo e della trisezione dell'angolo non possono essere risolti con riga e compasso

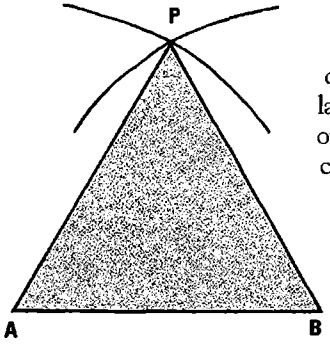
1882 Lindemann dimostra l'impossibilità della quadratura del cerchio

Costruire un quadrato di area equivalente a quella di un cerchio dato.

L'espressione «quadratura del cerchio» è diventata di uso comune per indicare un'impresa impossibile. L'equazione algebrica $x^2 - 2 = 0$ ammette come soluzioni $x = \sqrt{2}$ e $x = -\sqrt{2}$. Si tratta di due numeri irrazionali (non possono essere scritti come frazioni), ma dimostrare che non è possibile quadrare il cerchio equivale a dimostrare che π non può essere soluzione di *nessuna* equazione algebrica. I numeri irrazionali dotati di questa proprietà sono detti «numeri trascendenti», perché possiedono un «grado maggiore» di irrazionalità rispetto ai loro cugini irrazionali, come $\sqrt{2}$.

Era opinione comune dei matematici che π fosse trascendente, ma la questione restò un enigma difficile da dimostrare finché Ferdinand von Lindemann non vi riuscì utilizzando una variante della tecnica elaborata da Charles Hermite per dimostrare che la base dei logaritmi naturali, e, era trascendente (vedi capitolo 6).

Si potrebbe pensare che a seguito del risultato di Lindemann il flusso di pubblicazioni da parte degli indomiti «quadratori del cerchio» fosse cessato. Niente affatto. Nelle zone d'ombra della matematica, continuarono ad applicarsi a questo problema tutti coloro che erano restii ad accettare la logica della dimostrazione o che non l'avevano mai sentita nominare.



Costruzione di un triangolo equilatero

Costruzione di poligoni Euclide pose il problema della costruzione di un poligono regolare, ovvero di una figura simmetrica dotata di molti lati, come un quadrato o un pentagono, in cui tutti i lati hanno la stessa lunghezza e i lati adiacenti formano angoli uguali. Nella sua celebre opera intitolata gli *Elementi* (libro IV), Euclide dimostrò che era possibile costruire poligoni con 3, 4, 5 e 6 lati utilizzando solo riga e compasso.

Il poligono con tre lati è comunemente chiamato «triangolo equilatero» ed è molto semplice da costruire. Indichiamo con A e B due punti separati dalla distanza desiderata come lunghezza del lato del triangolo, qualunque essa sia. Puntando il compasso in A si traccia un arco di circonferenza di raggio AB, quindi si ripete l'operazione puntando il compasso in B (con lo stesso raggio). Chiamiamo P il punto di intersezione dei due archi. Poiché $AP = AB$ e $BP = AB$, tutti e tre i lati del triangolo APB sono uguali. La realizzazione del triangolo vero e proprio viene completata unendo AB, AP e BP con la riga.

Chi pensa che usare una riga sia già un lusso, sappia che non è il solo: lo pensava anche il danese Georg Mohr. La costruzione del triangolo equilatero è completa

quando si trova il punto P e fino a quel punto è sufficiente un compasso; la riga serve solo per unire *fisicamente* i punti. Mohr dimostrò che ogni costruzione ottenibile con riga e compasso poteva essere ottenuta anche solo con il compasso. L'italiano Lorenzo Mascheroni dimostrò lo stesso risultato 125 anni dopo; l'unica caratteristica originale del suo testo del 1797 dedicato a Napoleone, *La geometria del compasso*, era quella di essere in versi.

È nato un principe

Carl Friedrich Gauss rimase così colpito dalla sua scoperta che era possibile costruire un poligono con 17 lati che accantonò il progetto di studiare le lingue straniere e decise di darsi alla matematica. Il resto è storia: Gauss è stato definito «il principe dei matematici». La base del suo monumento commemorativo, a Gottinga (Germania), è costituita da un poligono a 17 lati: un tributo appropriato al suo genio.

Per quanto riguarda il problema generale, rivestono una particolare importanza i poligoni con p lati in cui p è un numero primo. Abbiamo visto come si costruisce il poligono con tre lati; Euclide costruì quello che ne aveva cinque, ma non riuscì a costruire il poligono con sette lati (eptagono). Studiando questo problema a soli diciassette anni, Carl Friedrich Gauss elaborò una dimostrazione negativa, deducendo che non era possibile costruire un poligono con p lati per $p = 7, 11$ o 13.

Del resto, Gauss sviluppò anche una dimostrazione positiva, concludendo che era possibile costruire un poligono con 17 lati. In realtà si spinse ancora oltre e dimostrò che un poligono con p lati può essere costruito se e solo se il numero primo p ha la forma

$$p = 2^{2^n} + 1$$

I numeri di questo tipo si chiamano «numeri di Fermat». Calcolandoli per $n = 0$, 1, 2, 3 e 4 si scopre che sono i numeri primi $p = 3, 5, 17, 257$ e $65\,537$, e che corrispondono a un poligono costruibile con p lati.

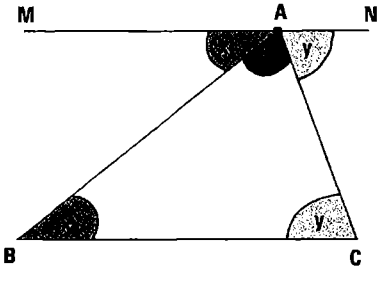
Il numero di Fermat corrispondente a $n = 5$ è $p = 2^{32} + 1 = 4\,294\,967\,297$. Pierre de Fermat aveva ipotizzato che tali numeri fossero tutti primi, ma questo non lo è, perché $4\,294\,967\,297 = 641 \times 6\,700\,417$. Sostituendo nella formula $n = 6$ o 7 si ottengono due numeri di Fermat giganteschi, che però, come quello ottenuto per $n = 5$, non sono primi.

idea chiave

Prendiamo una riga e un compasso...

21 I triangoli

La definizione più ovvia che si può dare di un triangolo è che si tratta di una figura geometrica con tre lati e tre angoli (tri-angolo). La trigonometria è la parte della matematica che consente di «misurare il triangolo», che si tratti delle dimensioni degli angoli, della lunghezza dei lati o dell'area. Pur essendo una delle figure geometriche più semplici, il triangolo costituisce una fonte inesauribile di interesse.



La storia del triangolo La somma degli angoli interni di un triangolo è sempre uguale a due angoli retti, ovvero a 180 gradi. Questo enunciato può essere dimostrato con un ragionamento semplice ed elegante; il primo passo consiste nel tracciare una retta *MAN* passante per la punta (o «vertice») *A* di un triangolo e parallela alla sua base *BC*.

Gli angoli AB̂C (che indicheremo con *x*) e BÂM sono uguali, perché sono angoli alterni interni formati dalle rette parallele *MN* e *BC*. Indichiamo con *y* gli altri due angoli alterni interni AĈB e CÂN. L'angolo MÂN ha un'ampiezza di 180 gradi (metà di 360 gradi) ed è uguale a $x + y + z$, ovvero alla somma degli angoli interni del triangolo. Q.E.D., come scriveva Euclide al termine delle sue dimostrazioni. Naturalmente tutto ciò vale nell'assunzione che il triangolo sia disegnato su una superficie piatta come questa pagina, perché la somma degli angoli di un triangolo tracciato su una superficie sferica (un triangolo sferico) non corrisponde a 180 gradi; ma questa è un'altra storia.

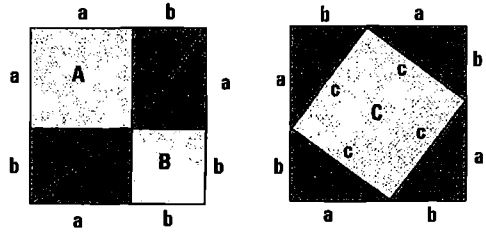
Euclide dimostrò molti teoremi sui triangoli, procedendo sempre in modo rigorosamente deduttivo. Fra le altre cose, dimostrò che «la somma di due lati di un triangolo è sempre maggiore del terzo», un risultato che oggi è molto importante in matematica astratta, dove viene chiamato «disuguaglianza triangolare». Con il consueto pragmatismo, gli epicurei affermarono che non c'era alcun bisogno di dimostrarlo, perché risultava evidente anche a un somaro: se si fosse posta una balla di fieno in un vertice del triangolo e il somaro in un altro, l'animale non avrebbe certo percorso due lati prima di soddisfare il proprio appetito.

linea del tempo

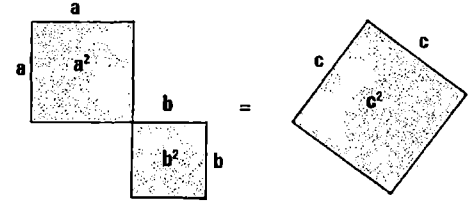
1850 a.C. I Babilonesi conoscono il «teorema di Pitagora»

1335 Richard di Wallingford scrive un pionieristico trattato di trigonometria

Il teorema di Pitagora Il più importante teorema sui triangoli è il teorema di Pitagora, che gioca un ruolo di primo piano nella matematica moderna (ma non è certo che Pitagora sia stato il primo a scoprirlo). Sebbene il suo enunciato più noto sia espresso in forma algebrica, $a^2 + b^2 = c^2$, Euclide si riferiva alle figure geometriche vere e proprie: «Nei triangoli rettangoli il quadrato del lato che sottende l'angolo retto è uguale alla somma dei quadrati dei lati che contengono l'angolo retto».



La dimostrazione di Euclide costituisce la Proposizione 47 del libro I degli *Elementi* e ha angustiato intere generazioni di scolari che sono stati costretti a impararla a memoria. Esistono centinaia di dimostrazioni di questo teorema; una delle più apprezzate è più nello stile di Bhāskara (XII secolo) che in quello della dimostrazione euclidea del 300 a.C.

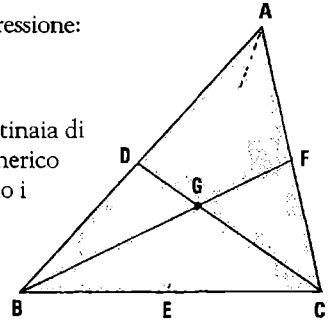


Si tratta di una dimostrazione «visiva». Nella figura a margine si vede che il quadrato di lato $a + b$ può essere suddiviso in due modi diversi.

Poiché i quattro triangoli uguali (ombreggiatura scura) sono comuni a entrambi i quadrati, è possibile eliminarli conservando l'equivalenza fra le aree delle due figure. Osservando come sono costituite queste ultime, si vede saltare fuori la celebre espressione:

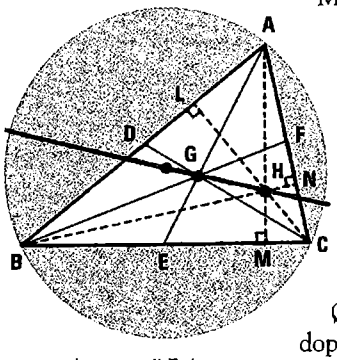
$$a^2 + b^2 = c^2$$

La retta di Eulero Riguardo ai triangoli si possono enunciare centinaia di teoremi. In primo luogo prendiamo in esame i punti medi dei lati. In un generico triangolo ABC, indichiamo i punti medi dei lati con D, E, e F; poi tracciamo i segmenti BF e CD, e chiamiamo G il loro punto di intersezione. Infine uniamo A ed E: il fatto che questa retta passi per G non è scontato e richiederebbe altre deduzioni, ma di fatto ci passa. Il punto G è detto «baricentro» del triangolo, perché rappresenta il suo centro di gravità.



In un triangolo si possono individuare (letteralmente) centinaia di «centri». Oltre al baricentro, i più importanti sono l'«ortocentro» H, ovvero il punto di intersezione delle altezze (le perpendicolari ai lati tracciate a partire dal vertice opposto, indicate con una linea tratteggiata nella figura a pagina 86) e il «circocentro» O, che è il punto in cui si incontrano gli assi dei lati (cioè il punto equidistante dai vertici del triangolo). Questo punto è il centro della circonferenza passante per A, B e C.

1571	1822	1873
François Viète pubblica un testo di trigonometria contenente delle tavole trigonometriche	Karl Feuerbach definisce per un triangolo il cerchio dei nove punti	Brocard produce uno studio esaustivo sulle proprietà del triangolo

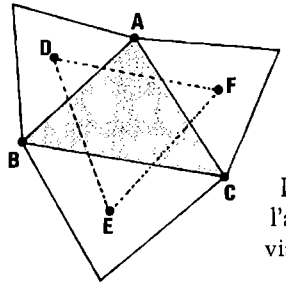


La retta di Eulero

Ma c'è dell'altro. In *ogni* triangolo ABC i centri G, H e O (il baricentro, l'ortocentro e il circocentro, rispettivamente) giacciono su una stessa retta, detta «*retta di Eulero*»; nei triangoli equilateri (in cui i lati hanno tutti la stessa lunghezza) questi tre centri coincidono in un punto che rappresenta l'unico centro del triangolo.

Il teorema di Napoleone Se si costruisce un triangolo equilatero su ciascun lato di un triangolo generico ABC, i baricentri dei nuovi triangoli formano a loro volta un triangolo DEF. Il teorema di Napoleone afferma che, per *qualunque* triangolo ABC, il triangolo DEF è *equilatero*.

Questo teorema fu pubblicato in una rivista inglese nel 1825, pochi anni dopo la morte di Napoleone (avvenuta a Sant'Elena nel 1821). L'abilità matematica dimostrata a scuola avrà sicuramente giovato a Napoleone ai fini dell'ammissione alla scuola di artiglieria e in qualità di imperatore il condottiero corso ebbe senz'altro modo di conoscere a Parigi i matematici più importanti dell'epoca, ma non esiste nessun'altra prova che lo colleghi all'omonimo teorema; come spesso avviene in matematica, il teorema di Napoleone è attribuito a una persona che c'entra poco o niente con la sua scoperta o la sua dimostrazione. A dire il vero, questo teorema è soggetto a continue riscoperte ed estensioni.



Il teorema di Napoleone

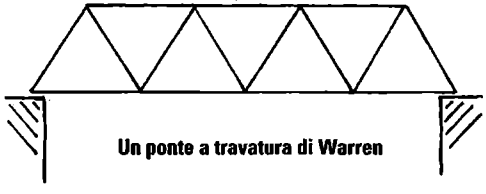
I dati essenziali per la definizione di un triangolo sono la lunghezza di un lato e l'ampiezza di due angoli; tutte le altre grandezze possono essere determinate per via trigonometrica.

Quando si effettuano i rilevamenti topografici necessari per disegnare una carta geografica, conviene dimenticare che la Terra è rotonda e assumere invece che i triangoli siano piatti. Il procedimento di «triangolazione» consiste nel costruire un reticolo di triangoli: a partire da una base BC di lunghezza nota, si sceglie un punto lontano A (il punto di triangolazione) e si misurano gli angoli $A\hat{B}C$ e $A\hat{C}B$ con un teodolite. Dopo aver ricavato per via trigonometrica tutte le informazioni relative al triangolo ABC, il topografo si sposta, fissa un nuovo punto di triangolazione sulla nuova linea di base AB o AC e ripete l'operazione finché non ha ottenuto la rete di triangoli desiderata. Il vantaggio di questo metodo è che consente di mappare anche territori non misurabili direttamente a causa di ostacoli naturali come stagni, paludi, sabbie mobili e fiumi.

Il metodo di triangolazione costituiva la base del *Great Trigonometrical Survey*, il rilevamento trigonometrico del subcontinente indiano che ebbe inizio nei primi anni dell'Ottocento e durò quarant'anni. Gli obiettivi erano quelli di ricavare una carta dettagliata dell'India e di misurare la lunghezza dell'arco di meridiano compreso fra Capo Comorin (a sud) e l'Himalaya (a nord), per una distanza di quasi 2500 kilometri. Per ottenere la massima accuratezza nella misura degli angoli Sir George Everest fece costruire a Londra due giganteschi teodoliti del peso di una tonnellata,

Costruire con i triangoli

In campo edile il triangolo è assolutamente indispensabile. Il principale vantaggio del suo utilizzo consiste nella caratteristica che lo rende insostituibile nei rilevamenti topografici: la sua rigidità. Se si possono deformare un quadrato o un rettangolo fino a trasformarli in una figura diversa, ciò è impossibile nel caso del triangolo. La travatura reticolare usata nelle costruzioni è costituita dall'unione di vari triangoli; il modulo strutturale oltre a essere visibile nei tetti, trova impiego nella costruzione dei ponti, dove ha segnato una svolta davvero epocale.



Un ponte a travatura di Warren

La trave di Warren è una struttura in grado di sostenere un carico notevole rispetto al suo peso. Fu brevettata da James Warren nel 1848 e due anni dopo venne impiegata per la prima volta nella costruzione di un ponte (presso la stazione del London Bridge, a Londra). L'impiego di triangoli equilateri garantiva una maggiore stabilità rispetto a progetti analoghi basati sui triangoli isosceli (che hanno solo due lati uguali).

il cui spostamento richiedeva squadre di decine di uomini. L'accuratezza nelle misure degli angoli era fondamentale e costituì l'oggetto di molte discussioni, ma il cuore del procedimento era l'umile triangolo. Gli scienziati dell'epoca vittoriana dovevano cavarsela senza GPS; in realtà disponevano anche loro di «calcolatori», ma erano calcolatori umani! Se si conoscono tutte le lunghezze di un triangolo, il calcolo dell'area diventa immediato. Fra le varie formule che consentono di calcolare l'area A di un triangolo, la più straordinaria è quella di Erone di Alessandria:

$$A=\sqrt{p\times(p-a)\times(p-b)\times(p-c)}$$

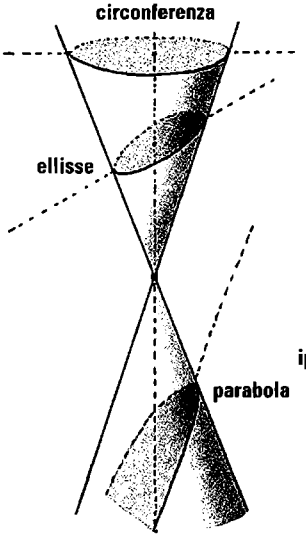
Si tratta di una formula applicabile a qualsiasi triangolo, senza bisogno di conoscere nessun angolo. Il simbolo p rappresenta il semiperimetro del triangolo di lati a, b e c; ad esempio, se i lati di un triangolo sono lunghi 13, 14 e 15 il suo perimetro è 13 + 14 + 15 = 42, e p = 21. Completando il calcolo, si ottiene $A = \sqrt{(21 \times 8 \times 7 \times 6)} = \sqrt{7056} =$ 84. Il triangolo è un oggetto familiare tanto ai bambini che giocano con le figure geometriche più semplici quanto ai ricercatori che ogni giorno utilizzano la disuguaglianza triangolare nell'ambito della matematica astratta. Le proprietà dei triangoli sono ricavabili per mezzo della trigonometria, le cui funzioni seno, coseno e tangente consentono di effettuare calcoli con l'accuratezza necessaria per varie applicazioni pratiche. Il triangolo ha costituito l'oggetto di molti studi, ma su questa figura così semplice formata da tre rette restano ancora da scoprire molte cose.

idea chiave

I tre lati della questione

22 Le curve

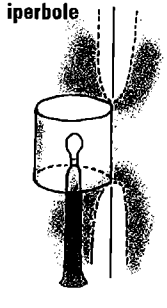
È facile disegnare una curva: gli artisti lo fanno continuamente, mentre gli architetti progettano complessi residenziali che seguono la curvatura di una mezzaluna o di altre forme sinuose. In ambito sportivo, il lanciatore del baseball può eseguire un lancio curvo e i calciatori risalgono il campo di gioco descrivendo un'ampia curva; quando poi tirano in rete, il pallone segue una traiettoria curva. Però alla domanda: «Che cos'è una curva?» non è facile rispondere.



Le sezioni coniche

I matematici hanno studiato le curve per molti secoli e da vari punti di vista. I primi sono stati i Greci, che hanno preso in esame le curve oggi definite «classiche».

Curve classiche La prima famiglia di curve classiche è costituita dalle cosiddette «sezioni coniche», che comprendono la circonferenza, l'ellisse, la parabola e l'iperbole. Le coniche si ottengono tagliando un doppio cono (due coni gelato uniti per la punta) con un piano: in base all'inclinazione del piano rispetto all'asse verticale del cono, la curva di intersezione può essere una circonferenza, un'ellisse, una parabola o un'iperbole.



Una conica può essere considerata la proiezione di una circonferenza su uno schermo. I raggi di luce emessi da una lampada da tavolo cilindrica formano un doppio cono luminoso delimitato dalle proiezioni dei bordi circolari (superiore e inferiore) del paralume. Sul soffitto si forma l'immagine di una circonferenza, ma basta inclinare la lampada per trasformarla in un'ellisse; sulla parete si delinea l'immagine di una curva costituita da due rami, l'iperbole.

linea del tempo

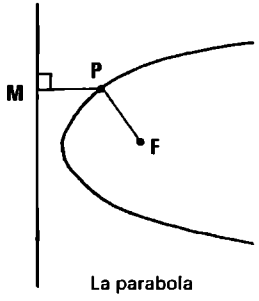
circa 300 a.C. Euclide definisce le sezioni coniche

circa 250 a.C. Archimede studia le spirali

circa 225 a.C. Apollonio di Perga pubblica *Le coniche*

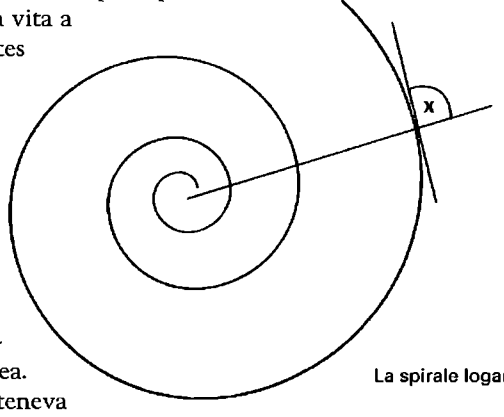
Le coniche possono essere descritte anche in base al movimento di un punto nel piano: si tratta del metodo del «luogo geometrico» tanto caro ai Greci. A differenza della definizione proiettiva, questo metodo si basa sulle distanze. Un punto che si muove in modo da mantenere costante la sua distanza da *un unico* punto descrive una circonferenza; uno che si muove in modo da tenere fissa la somma delle sue distanze da *due* punti fissi (detti «fuochi») genera un'ellisse (se i due fuochi coincidono l'ellisse diventa una circonferenza). L'ellisse fornì la spiegazione del moto dei pianeti: nel 1609 l'astronomo tedesco Keplero ripudiò la vecchia idea delle orbite circolari e annunciò che i pianeti ruotavano attorno al Sole lungo traiettorie ellittiche.

Un caso meno ovvio è quello del punto che si muove in modo che la sua distanza da un punto fisso (il fuoco *F*) sia equivalente alla lunghezza della sua perpendicolare rispetto a una determinata retta (chiamata «direttrice»): in questo caso si ottiene una parabola. La parabola possiede molte proprietà interessanti. Se una sorgente luminosa viene posta nel fuoco *F* di una parabola, i raggi emessi sono tutti paralleli a *PM*. Inversamente, un'antenna parabolica concentra i segnali che riceve dal satellite nel proprio fuoco, da dove poi verranno inviati all'apparecchio televisivo.



La parabola

Quando un'asta viene fatta ruotare attorno a un punto fisso ogni suo punto descrive una circonferenza, ma se un punto può scorrere lungo l'asta mentre questa è in rotazione si ottiene una spirale. Pitagora amava particolarmente la spirale; molto tempo dopo, Leonardo da Vinci passò dieci anni della sua vita a studiarne le diverse varietà e René Descartes dedicò all'argomento un intero trattato. La spirale logaritmica viene chiamata anche «spirale equiangolare», perché l'angolo compreso fra il raggio e la tangente alla curva (nel punto di intersezione fra raggio e spirale) è costante in ogni suo punto.

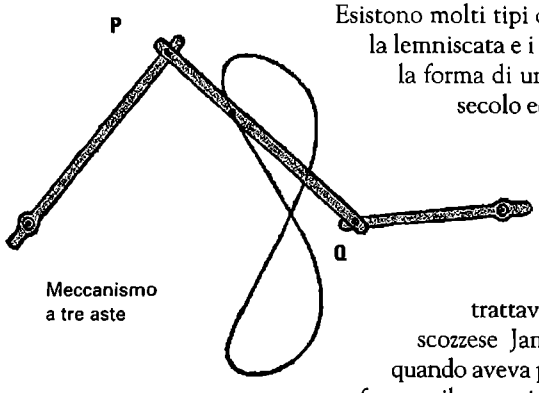


La spirale logaritmica

Jacob Bernoulli, appartenente alla celebre famiglia di matematici svizzeri, era talmente innamorato della spirale logaritmica da chiedere che venisse incisa sulla sua tomba, a Basilea. Il «rinascimentale» Emanuel Swedenborg riteneva

1704	1890	1920-1930
Newton classifica le curve cubiche	Peano dimostra che un quadrato pieno è una curva (curva di Peano)	Menger e Urysohn definiscono le curve nell'ambito della topologia

che la spirale fosse la forma più perfetta. Una spirale tridimensionale che si avvolge attorno a un cilindro prende il nome di «elica»; la molecola di DNA è costituita da due eliche e quindi ha una struttura «a doppia elica».



Meccanismo a tre aste

Esistono molti tipi di curve classiche, come la chiocciola di Pascal, la lemniscata e i vari ovali. La cardioide si chiama così perché ha la forma di un cuore. La catenaria è stata studiata nel XVIII secolo ed è definita come la curva formata da una catena sospesa fra due punti; è la curva creata, ad esempio, da un ponte sospeso fra due piloni verticali.

Nel XIX secolo furono studiate anche le curve generate dal movimento di aste meccaniche. Si trattava di un'estensione del problema che l'ingegnere scozzese James Watt aveva risolto in forma approssimata quando aveva progettato delle barre incernierate in grado di trasformare il moto circolare in moto lineare. Questi meccanismi rappresentarono un notevole passo avanti ai fini dello sfruttamento delle macchine a vapore.

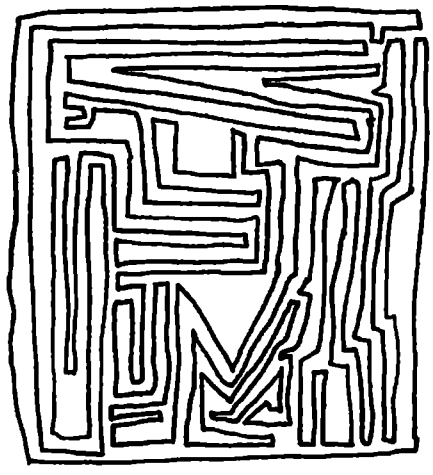
Il più semplice di questi congegni è costituito da tre aste mobili incernierate a due a due, con le estremità esterne fisse. Quando la «barra di accoppiamento» PQ si muove in una delle due direzioni il luogo geometrico di uno dei suoi punti è una curva di sesto grado, o «sestica».

Curve algebriche Dopo che Descartes ebbe rivoluzionato la geometria con l'introduzione delle coordinate x, y e z e degli assi cartesiani a cui dà il nome, fu possibile studiare le coniche esprimendole come equazioni algebriche. Ad esempio, una circonferenza di raggio unitario e centrata nell'origine degli assi è descritta dall'equazione $x^2 + y^2 = 1$; si tratta di un'equazione di secondo grado, come quelle di tutte le coniche. Si sviluppò così una nuova branca della geometria, la «geometria analitica».

In un testo di importanza capitale, Newton classificò le curve descritte dalle equazioni algebriche di terzo grado, le cosiddette «curve cubiche». A fronte delle quattro coniche fondamentali, in questo caso furono individuati 78 diversi tipi di curve, raggruppati in cinque classi. Il numero di tipologie continua ad aumentare vertiginosamente quando si passa alle curve quartiche, le cui varietà sono talmente numerose da non essere mai state classificate per intero.

Tuttavia lo studio delle curve in termini di equazioni algebriche non risolve tutti i problemi, perché molte curve, come le catenarie, le cicloidi (le curve tracciate da una punta fissata a una ruota che gira) e le spirali, non sono facilmente esprimibili in forma algebrica.

Una definizione In realtà i matematici ambivano a dare una definizione della curva in sé, non solo di alcuni esempi specifici. Camille Jordan propose una teoria in cui le curve venivano definite in termini di punti variabili.



Una curva semplice chiusa di Jordan

Vediamo un esempio. Se valgono le relazioni $x = t^2$ e $y = 2t$, al variare di t si origina una serie di punti esprimibili mediante coppie di coordinate (x, y). Ad esempio, per $t = 0$ si ottiene il punto $(0, 0)$, per $t = 1$ il punto $(1, 2)$ e così via. Riportando in grafico questi punti sul piano cartesiano e «unendo i punti», si ottiene una parabola. Jordan perfezionò poi questo concetto della rappresentazione dei punti, che costituiva la sua definizione di curva.

Le curve di Jordan possono essere intricate anche quando sono «semplici» (cioè non si incrociano) e «chiuse» (non hanno inizio né fine), come la circonferenza. Il celebre teorema di Jordan afferma che una curva semplice chiusa possiede un lato interno e uno esterno; a dispetto della sua apparente «banalità», questo teorema ha un significato profondo.

Nel 1980 l'italiano Giuseppe Peano fece scalpore dimostrando che, in base alla definizione di Jordan, un quadrato pieno può essere considerato una curva: infatti Peano riuscì a disporre i punti di un quadrato in modo che fossero *tutti* «rappresentati», rispettando allo stesso tempo la definizione di Jordan. Questo tipo di curva, detta «curva di Peano», aprì una falla nella definizione di Jordan, perché è evidente che un quadrato non è una curva nel senso tradizionale del termine.

Le curve di Peano e altri casi «patologici» costrinsero i matematici a tornare alla lavagna per correggere i fondamenti della teoria delle curve. Nacque così l'esigenza di elaborare una definizione migliore per la curva, che all'inizio del XX secolo spinse i matematici ad avventurarsi in un campo ancora inesplorato: la topologia.

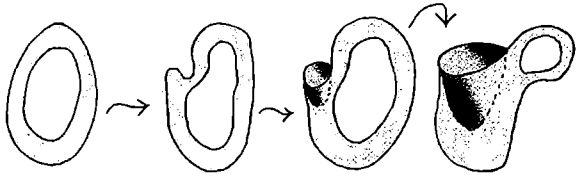
idea chiave

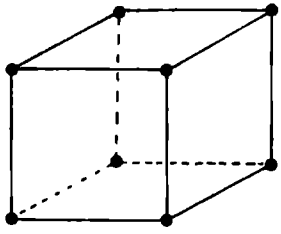
Una lunga serie di curve

23 La topologia

La topologia è la branca della geometria che studia le proprietà delle superfici e delle figure in generale, senza preoccuparsi di misurarne lunghezze o angoli. Una delle sue priorità è l'individuazione delle caratteristiche che restano invariate quando una figura viene trasformata in un'altra. Infatti la topologia ammette la possibilità di deformare le figure in tutte le direzioni, tanto da essere anche chiamata la «geometria del foglio di gomma». I topologi sono persone assai strane: non distinguono una ciambella da una tazzina da caffè!

Una ciambella è una superficie in cui è presente un buco. Lo stesso si può dire di una tazzina da caffè: in questo caso il buco definisce il manico. Per trasformare una ciambella in una tazzina da caffè si procede così:





Classificazione dei poliedri Le figure più semplici prese in esame dai topologi sono i poliedri (una parola di origine greca che significa «con molte facce»). Un esempio di poliedro è il cubo, che possiede 6 facce quadrate, 8 vertici (i punti di incontro fra le facce) e 12 spigoli (i segmenti che collegano i vertici). Il cubo è un poliedro regolare, perché

- tutte le facce sono uguali e hanno una forma regolare
- tutti gli angoli delimitati dagli spigoli che si incontrano in un vertice sono uguali.

linea del tempo

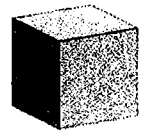
circa 300 a.C. Euclide dimostra che esistono cinque poliedri regolari

circa 250 a.C. Archimede studia i poliedri troncati

1752 Eulero elabora l'omonima formula che collega il numero dei vertici, degli spigoli e delle facce di un poliedro

La topologia è una branca della matematica relativamente nuova, ma le sue origini risalgono ai tempi dell'antica Grecia; infatti, il risultato cruciale degli *Elementi* di Euclide consiste nella dimostrazione che esistono *esattamente* cinque poliedri regolari, i solidi platonici:


Tetraedro

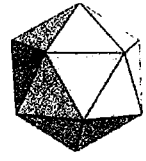

Cubo

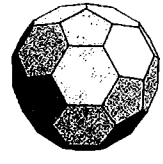
- il tetraedro (4 facce triangolari)
- il cubo (6 facce quadrate)
- l'ottaedro (8 facce triangolari)
- il dodecaedro (12 facce pentagonali)
- l'icosaedro (20 facce triangolari).


Ottaedro


Dodecaedro

Se si rinuncia alla condizione dell'uguaglianza delle facce si ottengono i solidi archimedei, che sono solo semiregolari; vediamo alcuni esempi, ricavandoli dai solidi platonici. Tagliando via (troncando) i vertici dell'icosaedro si ottiene la figura a cui si ispira il moderno pallone da calcio: le 32 facce che costituiscono i riquadri corrispondono a 12 pentagoni e 20 esagoni, uniti da 90 spigoli e 60 vertici. È la stessa forma delle molecole di buckminsterfullerene (o «fullerene»), così chiamato in onore di Richard Buckminster Fuller, il visionario padre della cupola geodetica. Il fullerene, una forma del carbonio scoperta recentemente, è costituito da molecole di formula C_{60} in cui ogni vertice dell'icosaedro troncato è occupato da un atomo di carbonio.


Icosaedro

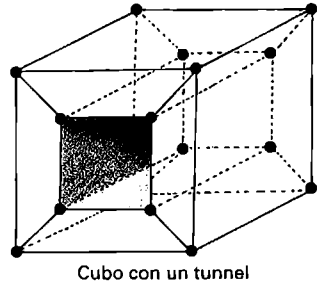

Icosaedro troncato

La formula di Eulero La formula di Eulero stabilisce che il numero dei vertici (V), il numero degli spigoli (S) e il numero delle facce (F) di un poliedro sono collegati dalla formula

$$V - S + F = 2$$

Nel caso di un cubo $V = 8$, $S = 12$ e $F = 6$, perciò $V - S + F = 8 - 12 + 6 = 2$, mentre per il fullerene si ha $V - S + F = 60 - 90 + 32 = 2$. In realtà, questo teorema mette in discussione il concetto stesso di poliedro.

Ci si può domandare, ad esempio, se un cubo attraversato da un «tunnel» possa essere considerato un poliedro a tutti gli effetti. In questo caso la


Cubo con un tunnel

1858	1961	1982	2002
Möbius e Listing presentano il nastro di Möbius	Stephen Smale dimostra la congettura di Poincaré per un numero di dimensioni maggiore di quattro	Michael Freedman dimostra la congettura di Poincaré per quattro dimensioni	Perelman dimostra la congettura di Poincaré per tre dimensioni

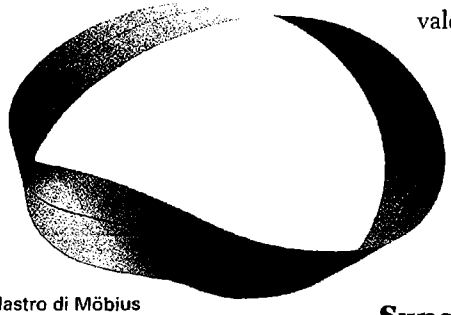
formula di Eulero non vale, perché $V = 16$, $S = 32$, $F = 16$ e $V - S + F = 16 - 32 + 16 = 0$. Per salvaguardare la validità della formula si può pensare allora di limitarne l'applicazione ai poliedri *privi di* tunnel oppure di generalizzarla in modo da includere questo caso particolare.

Classificazione delle superfici Se a un topologo una ciambella e una tazzina da caffè appaiono identiche, quale superficie gli apparirà *diversa* dalla ciambella? Una buona candidata potrebbe essere una sfera di gomma. Infatti non c'è modo di trasformare la ciambella in una sfera, perché queste due superfici presentano una differenza sostanziale: la prima ha un buco e la seconda no. Quindi le superfici possono essere classificate in base al numero di buchi che contengono.

Consideriamo una superficie con r buchi e suddividiamola in regioni delimitate da spigoli che uniscono i vertici presenti sulla superficie stessa. A questo punto si può contare il numero di vertici, di spigoli e di facce della superficie. Per qualunque suddivisione, l'espressione $V - S + F$ assume sempre lo stesso valore, detto «caratteristica di Eulero» della superficie:

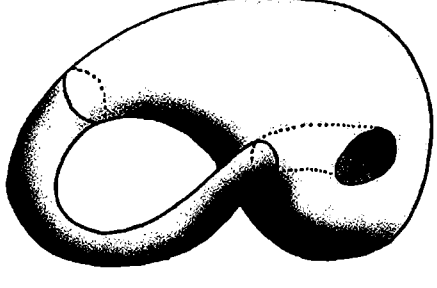
$$V - S + F = 2 - 2r$$

Se la superficie non contiene buchi ($r = 0$), come avviene nel caso dei poliedri ordinari, si ritrova la formula di Eulero $V - S + F = 2$; se invece è presente un buco ($r = 1$), come nel caso del cubo con un tunnel, $V - S + F = 0$.



Nastro di Möbius

Superfici a una faccia In generale, ogni superficie ha due facce. L'esterno di una sfera è diverso dal suo interno e l'unico modo per passare dall'uno all'altro consiste nel praticare un foro nella sfera, un'operazione di taglio non consentita in topologia (si può «stirare» una superficie, ma non tagliarla). Un altro esempio di superficie a due facce è un foglio di carta, in cui i due lati del foglio si incontrano solo lungo la linea di confine formata dai suoi spigoli.



Bottiglia di Klein

L'idea di una superficie dotata di un'unica faccia sembra assurda, eppure una celebre superficie di questo tipo fu scoperta nel XIX secolo dal matematico e astronomo tedesco August Möbius. Per costruirla basta prendere una striscia rettangolare di carta, effettuare un mezzo giro di torsione e unire le sue estremità. Il risultato è un «nastro di Möbius», una superficie che è dotata di una sola faccia e di

un solo bordo, come si può facilmente verificare seguendo il centro del nastro con una matita: in men che non si dica ci si ritrova al punto di partenza.

Esiste anche una superficie a una sola faccia priva di bordi: è la «bottiglia di Klein», che prende il nome dal matematico tedesco Felix Klein. La caratteristica più rilevante di questa figura è che non presenta intersezioni, anche se ciò è strettamente vero solo nello spazio a quattro dimensioni, il suo ambito di esistenza; qualsiasi modello tridimensionale della bottiglia di Klein presenta necessariamente un'intersezione fisica.

Queste superfici sono due esempi di ciò che i topologi chiamano «varietà», ovvero superfici geometriche che appaiono come fogli di carta bidimensionali quando se ne osservano *piccole* porzioni isolate. Essendo priva di bordi, la bottiglia di Klein è definita come una varietà bidimensionale «chiusa».

La congettura di Poincaré Un importante problema che ha angustiato i topologi per oltre un secolo è la famosa «congettura di Poincaré», formulata da Henri Poincaré e incentrata sul collegamento fra algebra e topologia.

La parte della congettura che è rimasta insoluta fino a tempi recenti riguarda le varietà tridimensionali, che sono piuttosto complesse: possiamo immaginarle come una bottiglia di Klein con una dimensione in più. Poincaré ipotizzò che alcune varietà tridimensionali chiuse che possedevano tutte le caratteristiche geometriche delle sfere tridimensionali fossero effettivamente delle sfere. Supporre il contrario equivaleva a girare attorno a una gigantesca palla e, a dispetto di tutti gli indizi favorevoli, dubitare che fosse una palla per l'impossibilità di vederla per intero.

Per molto tempo nessuno è riuscito a dimostrare la congettura di Poincaré per le varietà tridimensionali. Era vera o falsa? Anche se era stata dimostrata per tutte le altre dimensioni, il caso della varietà tridimensionale opponeva una strenua resistenza. I falsi allarmi si sono succeduti, finché nel 2002 è stato riconosciuto che Grigori Perelman dell'Istituto Steklov di San Pietroburgo era riuscito a dimostrarla. Com'è avvenuto anche per altri grandi problemi matematici, i metodi impiegati per risolvere la congettura di Poincaré non appartengono strettamente al suo ambito, ma riguardano una tecnica collegata alla diffusione del calore.

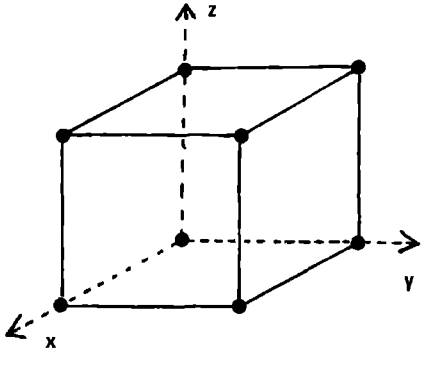
idea chiave

Da una ciambella a una tazzina di caffè

24 Le dimensioni

Leonardo da Vinci scrisse nel suo taccuino: «Il principio della scienza della pittura è il punto, il secondo è la linea, il terzo è la superficie, il quarto è il corpo che si veste di tal superficie». Nella gerarchia stabilita da Leonardo, il punto è privo di dimensioni, la retta è monodimensionale, il piano è bidimensionale e lo spazio è tridimensionale. Niente di più ovvio. È così che il punto, la retta, il piano e la geometria dei solidi sono stati tramandati dal geometra greco Euclide, di cui Leonardo ha seguito l'esposizione.

Il fatto che lo spazio fisico sia tridimensionale è un'opinione comune accettata da millenni. Nello spazio fisico è possibile uscire *fuori* da questa pagina lungo l'asse x o *attraversarla* orizzontalmente lungo l'asse y o *salire* verticalmente lungo l'asse z, oppure effettuare una qualunque combinazione di questi spostamenti. Prendendo come riferimento l'origine (il punto in cui si incontrano i tre assi), ogni punto è individuato da tre coordinate spaziali che assumono i valori x, y e z, e può essere scritto nella forma (x, y, z).



Lo spazio tridimensionale

È evidente che un cubo possiede queste tre dimensioni, come tutto ciò che è solido. In genere, a scuola si insegna la geometria nel piano (bidimensionale), quindi si passa alla geometria in tre dimensioni (la «geometria dei solidi») e poi ci si ferma lì.

All'inizio del XIX secolo i matematici iniziarono a occuparsi di quattro dimensioni e poi di un numero di dimensioni n ancora maggiore. Molti filosofi e matematici si domandarono allora se tali dimensioni esistessero o meno.

Dimensioni fisiche superiori Molti autorevoli matematici del passato ritenevano che quattro dimensioni fossero inconcepibili e ne mettevano in dubbio la reale esistenza, sfidando chiunque a giustificarla.

linea del tempo

circa 300 a.C.
Euclide descrive il mondo in termini tridimensionali

1877
Cantor è stupito dalla portata delle controversie che suscitano le sue scoperte nell'ambito della teoria delle dimensioni

Un metodo molto usato per giustificare l'esistenza di quattro dimensioni consisteva nel tornare momentaneamente a considerarne soltanto due. Nel 1884 un insegnante e teologo inglese, Edwin Abbott, pubblicò un libro di grande successo che parlava di «Flatlandia» (in inglese *Flatland*, «terra piatta»), un mondo interamente bidimensionale. Non potendo allontanarsi dal piano lungo una terza dimensione, gli abitanti di Flatlandia non potevano vedere i triangoli, i quadrati o i cerchi presenti nel loro mondo e quindi ne avevano una visione estremamente limitata. In poche parole, nel concepire una terza dimensione incontravano le stesse difficoltà che proviamo noi nel concepirne una quarta; per questo motivo, leggendo il libro di Abbott si entra nella disposizione d'animo più opportuna per accettare l'esistenza della quarta dimensione.

L'esigenza di ammettere la realtà di uno spazio a quattro dimensioni diventò più pressante quando comparve sulla scena Einstein. La sua geometria a quattro dimensioni era plausibile e perfino comprensibile, perché la dimensione aggiuntiva del suo modello era il *tempo*. A differenza di Newton, Einstein riteneva infatti che il tempo fosse unito allo spazio in un continuo spazio-temporale a quattro dimensioni e decretò che il mondo in cui viviamo è quadridimensionale ed è descritto dalle coordinate (x, y, z, t), dove t indica il tempo.

Oggi il mondo a quattro dimensioni di Einstein ci appare scontato e realistico. Un modello più recente descrive la realtà fisica in termini di «stringhe»; secondo questa teoria, le particelle subatomiche che conosciamo, come gli elettroni, non sono altro che la manifestazione di minuscole corde (stringhe) in continua vibrazione. La teoria delle stringhe propone di sostituire il continuo spazio-temporale a quattro dimensioni con una versione di dimensioni maggiori. I risultati ottenuti finora indicano che il continuo spazio-temporale in migliore accordo con la teoria delle stringhe dovrebbe avere 10, 11 o 26 dimensioni (in casi definiti da assunzioni e punti di vista diversi).

Un contributo alla soluzione del problema potrebbe venire dall'enorme magnete da 2000 tonnellate del CERN di Ginevra, in Svizzera, progettato per realizzare collisioni di particelle ad alta velocità. Sebbene le ricerche siano mirate primariamente a rivelare la struttura della materia, in via secondaria potrebbero indicare una teoria migliore e fornire la risposta «giusta» al problema della dimensionalità. L'opinione degli esperti è che viviamo in un Universo a undici dimensioni.

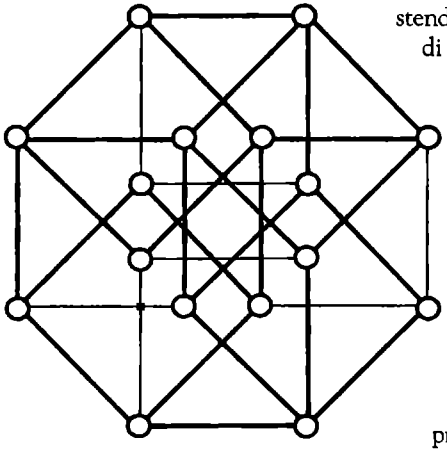
Iperspazio A differenza di quanto accade per dimensioni *fisiche* di ordine superiore, uno spazio *matematico* di più di tre dimensioni non pone alcun problema.

1909
Le ricerche di Brouwer modificano il concetto di dimensione

1919
Hausdorff introduce il concetto della «dimensione di Hausdorff» (frazionaria)

1970
La teoria delle stringhe prevede che il nostro Universo abbia 10, 11 o 26 dimensioni

Infatti lo spazio matematico può assumere un numero arbitrario di dimensioni e i matematici hanno lavorato tranquillamente con n variabili fin dall'inizio del XIX secolo. George Green, un mugnaio di Nottingham che studiò la teoria matematica dell'elettricità, e i matematici puri A.L. Cauchy, Arthur Cayley e Hermann Grassmann espressero le loro trattazioni matematiche in un iperspazio n-dimensionale: sembrava infatti che non ci fossero motivi validi per porre dei limiti alla matematica e che ci fosse invece tutto da guadagnare in termini di eleganza e di chiarezza nel considerare un numero maggiore di dimensioni.



Un cubo quadridimensionale

In sostanza, l'utilizzo di n dimensioni consiste semplicemente nell'estendere le coordinate tridimensionali (x, y, z) a un numero imprecisato di variabili. Se una circonferenza bidimensionale è descritta dall'equazione $x^2 + y^2 = 1$ e una sfera tridimensionale dall'equazione $x^2 + y^2 + z^2 = 1$, viene spontaneo pensare che l'equazione di un'ipersfera quadridimensionale sia $x^2 + y^2 + z^2 + w^2 = 1$.

Gli otto vertici di un cubo tridimensionale sono definiti da coordinate del tipo (x, y, z), dove x, y e z possono essere 0 o 1; il cubo ha sei facce quadrate e $2 \times 2 \times 2 = 8$ vertici. Ogni vertice di un cubo quadridimensionale sarà allora rappresentato da coordinate del tipo (x, y, z, w), in cui x, y, z e w sono 0 o 1, per un totale di $2 \times 2 \times 2 \times 2 = 16$ possibili vertici e otto facce, ciascuna delle quali è un cubo. Un cubo quadridimensionale di questo tipo è impossibile da visualizzare, ma si può rappresentare sulla pagina la proiezione che hanno in mente i matematici (vedi figura a margine), considerando che le facce cubiche sono difficili da distinguere.

Uno spazio matematico multidimensionale è un'entità in cui i matematici puri si imbattono spesso. Nessuno sostiene che esista realmente, sebbene si possa assumerne l'esistenza in un mondo ideale platonico. Nella problematica della classificazione dei gruppi (vedi capitolo 38), il «gruppo mostro» consente di misurare la simmetria in uno spazio matematico a 196 883 dimensioni. Anche se non possiamo «vedere» questo spazio nello stesso modo in cui vediamo lo spazio tridimensionale, possiamo sempre immaginarlo e trattarlo in modo preciso per mezzo dell'algebra moderna.

L'interesse dei matematici per le dimensioni non ha niente a che vedere con il significato attribuito dai fisici all'analisi dimensionale. Per accertarsi che le loro equazioni abbiano un senso, i fisici esprimono tutte le unità di misura in termini di massa (M), lunghezza (L) e tempo (T) e poi effettuano un'analisi dimensionale, cioè verificano che in entrambi i membri di ogni equazione compaiano le stesse dimensioni.

Ad esempio, l'analisi dimensionale rivela che l'uguaglianza forza = velocità non è valida. Infatti la velocità è espressa in metri al secondo, perciò ha le dimensioni di una lunghezza divisa per il tempo (L/T), che si può scrivere nella forma LT^{-1}. La forza invece è data dal prodotto della massa per l'accelerazione, che si misura in metri al secondo quadrato, per cui le sue dimensioni sono MLT^{-2}.

Coordinate personali

L'essere umano è un'entità multidimensionale perché possiede un numero di «coordinate» molto maggiore di tre; ad esempio, si potrebbero indicare con (*a, b, c, d, e, f, g, h*) l'età, l'altezza, il peso, il sesso, il numero di scarpa, il colore degli occhi, il colore dei capelli, la nazionalità e così via di ciascuna persona e rappresentarla come se fosse un punto geometrico. In questo «spazio» a otto dimensioni popolato da persone, Mario Rossi avrebbe coordinate del tipo (43 anni, 1,65 m, 83 kg, maschio, 42, castani, castani, italiana), mentre quelle di Mary Smith potrebbero essere (26 anni, 1,57 m, 56 kg, femmina, 37, castani, neri, britannica).

Topologia La teoria delle dimensioni fa parte della topologia generale. In termini di spazi matematici astratti si possono dare molte definizioni indipendenti del concetto di dimensione: un compito importante che i matematici si trovano a svolgere è quello di mostrarne i collegamenti. Il significato delle dimensioni è stato sviscerato dagli esponenti di spicco di vari settori della matematica, fra cui Henri Lebesgue, L.E.J. Brouwer, Karl Menger, Paul Urysohn e Leopold Vietoris (morto nel 2002 a 110 anni, dopo essere stato per un certo tempo il più vecchio cittadino austriaco).

Un testo fondamentale sull'argomento è *Dimension theory*; si continua ancora oggi a pensare che la sua pubblicazione, avvenuta nel 1948 a opera di Witold Hurewicz ed Henry Wallman, abbia segnato una svolta nel concetto di dimensione.

Dimensioni in tutte le loro forme Da quando i Greci hanno introdotto le tre dimensioni spaziali, il concetto di dimensione è stato sottoposto a una grande varietà di analisi critiche e di estensioni.

Se l'introduzione delle n dimensioni dello spazio matematico è stata del tutto indolore, in fisica esistono teorie basate sullo spazio-tempo (a quattro dimensioni) e recenti versioni della teoria delle stringhe (vedi p. 97) che richiedono 10, 11 e 26 dimensioni. Inoltre, la scoperta dei frattali (vedi capitolo 25) ha portato a spingersi nell'ambito delle dimensioni frazionarie, misurando varie grandezze. Infine, non va dimenticato che Hilbert ha introdotto uno spazio matematico a infinite dimensioni che costituisce una struttura portante della matematica pura. Come abbiamo visto, il discorso sulle dimensioni va ben oltre l'un, due, tre della geometria euclidea.

idea chiave

Oltre la terza dimensione

25 I frattali

Nel marzo del 1980 il modernissimo elaboratore centrale del centro di ricerca dell'IBM a Yorktown Heights, nello Stato di New York, stava comunicando le sue istruzioni a una vecchia stampante Tektronix quando l'apparecchio cominciò a disseminare una serie di punti nelle zone più impensate del foglio di carta. Quando ebbe finito di sferragliare, il risultato assomigliava a una manciata di polvere sparsa sulla pagina. Benoît Mandelbrot non credeva ai propri occhi: doveva essere un risultato importante, ma non capiva che cosa fosse. L'immagine prendeva forma lentamente davanti a lui nello stesso modo in cui una stampa in bianco e nero emerge da un bagno di sviluppo fotografico. Era la prima apparizione di quello che sarebbe diventato l'emblema del mondo dei frattali: l'insieme di Mandelbrot.

Si trattava di matematica sperimentale per eccellenza, con un approccio che prevedeva per i matematici l'utilizzo di banchi di laboratorio analoghi a quelli dei fisici e dei chimici: anche loro potevano finalmente effettuare degli esperimenti. Si prospettavano, letteralmente, nuove visioni. Era in atto un'emancipazione dall'arido mondo della triade «definizione, teorema, dimostrazione», anche se in seguito si sarebbe verificato un ritorno al rigore del ragionamento razionale.

Il lato negativo di questo approccio sperimentale era che le immagini visive precedevano i fondamenti teorici: gli scienziati sperimentali navigavano a vista. Mandelbrot coniò il termine «frattali», ma che cos'erano queste entità? Si poteva pensare di darne una definizione precisa in termini tradizionali? Inizialmente Mandelbrot non aveva intenzione di farlo, perché temeva di distruggere il fascino dell'esperimento affinando una definizione stringente che avrebbe poi potuto rivelarsi inadeguata e riduttiva. Pensava che il concetto di frattale, «come un buon vino, richiedesse un po' di invecchiamento prima di essere "imbottigliato"».

L'insieme di Mandelbrot Mandelbrot e i suoi colleghi non amavano la matematica astrusa, ma preferivano manipolare le formule più semplici. La loro

linea del tempo

1879 Cayley studia un precursore dei moderni frattali

1904 Von Koch genera la sua curva a fiocco di neve

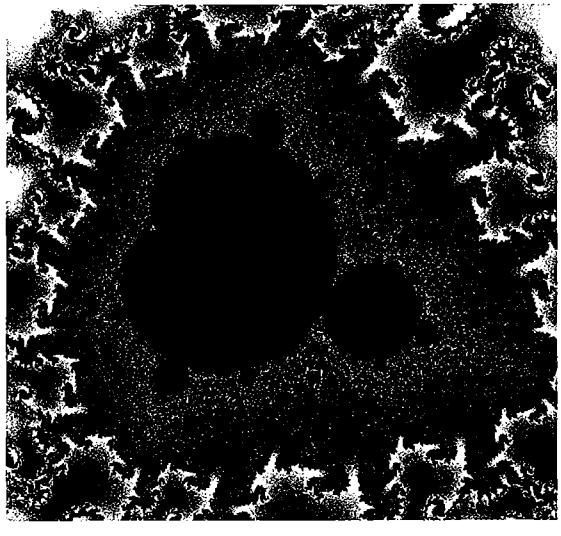
teoria è interamente basata sull'iterazione, un procedimento che consiste nell'applicare più volte di seguito la stessa formula. Quella che ha originato l'insieme di Mandelbrot è semplicemente $x^2 + c$.

Si comincia scegliendo un valore di c, ad esempio $c = 0{,}5$. Nella formula $x^2 + 0{,}5$ si sostituisce quindi $x = 0$, ottenendo 0,5. Si usa questo risultato come nuovo valore di x, sostituendolo in $x^2 + 0{,}5$, da cui si ottiene $(0{,}5)^2 + 0{,}5 = 0{,}75$. Continuando in questo modo, il terzo stadio sarà $(0{,}75)^2 + 0{,}5 = 1{,}0625$, come si può verificare con una semplice calcolatrice. Continuando ulteriormente, si ottengono risultati sempre più grandi.

Scegliamo ora un altro valore di c, ad esempio $c = -0{,}5$. Come nel caso precedente, si sostituisce $x = 0$ in $x^2 - 0{,}5$, ottenendo $-0{,}5$. Il passaggio successivo produce $-0{,}25$; questa volta però il risultato della formula non diventa sempre più grande, ma dopo alcune fluttuazioni si stabilizza su un valore prossimo a $-0{,}3660...$

Quindi, per $c = 0{,}5$ la successione che parte da $x = 0$ diverge all'infinito, mentre per $c = -0{,}5$ converge a un valore prossimo a $-0{,}3660$. L'insieme di Mandelbrot è costituito da tutti i valori di c per i quali la successione che parte da $x = 0$ *non* diverge all'infinito.

L'insieme di Mandelbrot



La storia non finisce qui. Infatti, fino ad ora ci siamo limitati a considerare numeri reali monodimensionali, ma l'insieme di Mandelbrot prodotto da questi numeri è a sua volta monodimensionale e non ci dice un granché. Vogliamo ora considerare la stessa formula $z^2 + c$, nell'assunzione che z e c siano numeri complessi bidimensionali (vedi capitolo 8): in questo caso si ottiene un insieme di Mandelbrot bidimensionale.

Per alcuni valori di c all'interno dell'insieme di Mandelbrot la successione dei valori di z può abbandonarsi a ogni tipo di stramberia, come saltellare da un punto all'altro, ma non diverge mai all'infinito. L'insieme di Mandelbrot presenta un'altra caratteristica fondamentale dei frattali, l'autosomiglianza: osservando

1918 Hausdorff introduce il concetto di dimensione frazionaria

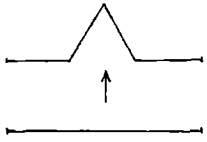
1919 Julia e Fatou studiano le strutture frattali nel campo del complesso

1975 Mandelbrot conia il termine «frattale»

l'insieme a distanza sempre più ravvicinata non si riesce a stabilire qual è il fattore di ingrandimento, perché si vedono sempre e solo altri insiemi di Mandelbrot.

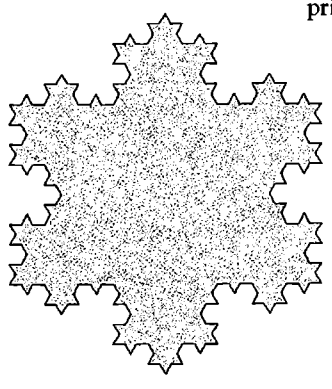
Prima di Mandelbrot In matematica, difficilmente le scoperte sono novità assolute. Guardando al passato, Mandelbrot scoprì che la sua stessa idea era già balenata un secolo prima a matematici del calibro di Henri Poincaré e Arthur Cayley, che però non disponevano della capacità di calcolo necessaria per approfondirne lo studio.

Fra le forme scoperte dalla prima generazione di teorici dei frattali c'erano le curve frastagliate e le «curve mostruose», che fino a quel momento erano state bollate come patologiche e quindi, essendo patologiche, erano state chiuse in un armadio e ignorate. A quel tempo erano molto più apprezzate le comuni curve «lisce» che potevano essere studiate con il calcolo differenziale. Quando i frattali acquistarono popolarità furono riesumati i risultati ottenuti da Gaston Julia e Pierre Fatou, che avevano studiato strutture frattali nel piano complesso negli anni successivi alla Prima guerra mondiale. Naturalmente quelle curve non si chiamavano ancora frattali, e Julia e Fatou non avevano gli strumenti tecnologici necessari per visualizzarne la forma.



L'elemento che genera il fiocco di neve di Koch

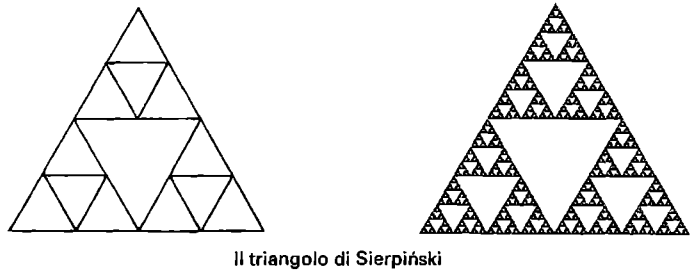
Altri celebri frattali La celeberrima curva di Koch prende il nome dal matematico svedese Niels Fabian Helge von Koch ed è stata sostanzialmente la prima a essere descritta. Questa curva frattale, detta anche «fiocco di neve», si ottiene suddividendo ciascun lato di un triangolo equilatero in tre parti uguali e sostituendo ogni segmento centrale con i due lati di un altro triangolo equilatero.



Il fiocco di neve di Koch

La proprietà curiosa della curva di Koch è che possiede un'area finita, perché può essere sempre inscritta in una circonferenza, ma a ogni stadio di generazione diventa più lunga: quindi è una curva che racchiude un'area finita, ma possiede un perimetro «infinito»!

Un altro celebre frattale è quello che prende il nome dal matematico polacco Wacław Sierpiński. Il «triangolo di Sierpiński» si



Il triangolo di Sierpiński

costruisce sottraendo progressivamente dei triangoli da un triangolo equilatero di partenza (ma nel capitolo 13 è stato generato in modo diverso).

Una dimensione frazionaria Felix Hausdorff guardava le dimensioni con occhi nuovi. Era una questione di scala. Se una retta viene ingrandita di un fattore 3, diventa 3 volte più lunga; poiché $3 = 3^1$, si dice che la retta ha dimensione 1. Se si ingrandisce un quadrato di un fattore 3, la sua *area* aumenta di $9 = 3^2$ volte, perciò la sua dimensione è 2. Se un cubo viene ingrandito dello stesso fattore, il volume diventa $27 = 3^3$ volte maggiore e la sua dimensione sarà 3. In tutti questi casi, la dimensione di Hausdorff assume il valore che ci si aspetta per una retta, un quadrato o un cubo.

Se l'unità di base della curva di Koch viene ingrandita di un fattore 3, diventa 4 volte più lunga; per analogia con lo schema precedente, si definisce «dimensione di Hausdorff» il valore D per cui $4 = 3^D$. In alternativa, la dimensione di Hausdorff può essere calcolata mediante la formula

$$D = \frac{\log 4}{\log 3}$$

Nel caso della curva di Koch, D è circa uguale a 1,262. Quando si ha a che fare con i frattali accade spesso che la dimensione di Hausdorff sia maggiore della dimensione ordinaria, che nel caso della curva di Koch è 1.

La dimensione di Hausdorff costituisce il cuore della definizione di frattale enunciata da Mandelbrot: un frattale è un insieme di punti il cui valore di D non è un numero intero. La dimensione frazionaria diventa così la proprietà caratteristica dei frattali.

Le applicazioni dei frattali Le potenziali applicazioni dei frattali spaziano in una grande varietà di ambiti, ad esempio, potrebbero essere utilizzati per descrivere matematicamente l'evoluzione di eventi naturali come la crescita delle piante o la formazione delle nuvole.

Del resto, lo sviluppo di organismi marini come coralli e spugne è gia stato descritto in termini frattalici ed è stato dimostrato che l'espansione delle città moderne è analoga alla crescita dei frattali. In medicina i frattali sono stati usati per modellare l'attività cerebrale, mentre in ambito economico si è indagata la natura frattale dell'andamento dei titoli e dei mercati valutari esteri. I risultati di Mandelbrot hanno spalancato le porte di un nuovo mondo che per buona parte è ancora inesplorato.

idea chiave

Figure di dimensione frazionaria

26 Il caos

Ha senso parlare di una teoria del caos? Il caos non si manifesta precisamente in assenza di qualsiasi teoria? Per rispondere a queste domande bisogna risalire al 1812. Mentre Napoleone avanzava verso Mosca, un suo compatriota, il marchese Pierre-Simon de Laplace, pubblicò un saggio sull'universo deterministico. Secondo Laplace, se si conoscessero le posizioni e le velocità di tutti gli oggetti dell'Universo nonché tutte le forze a cui sono sottoposti in un dato istante, sarebbe possibile calcolare esattamente queste grandezze per ogni tempo successivo: l'Universo e tutto ciò che contiene sarebbero quindi completamente determinati. La teoria del caos dimostra però che il mondo è di gran lunga più complesso.

Nella realtà non si possono conoscere *esattamente* tutte le posizioni, le velocità e le forze, ma il corollario dell'idea di Laplace è che, se conoscessimo i valori approssimati di queste grandezze in un certo istante, l'Universo non sarebbe poi molto diverso. Si tratta di un'assunzione ragionevole, perché è indubbio che, se un velocista partisse con un decimo di secondo di ritardo, taglierebbe il traguardo un decimo di secondo dopo. Il presupposto su cui si basava la teoria di Laplace era che a piccole differenze nelle condizioni iniziali dovessero corrispondere piccole discrepanze nei risultati finali. La teoria del caos distrusse questa convinzione.

L'effetto farfalla Con l'espressione «effetto farfalla» si fa riferimento al concetto che una piccola variazione nelle condizioni iniziali possa tradursi in un risultato completamente diverso da quello atteso. Il battito d'ali di una farfalla in Sud America potrebbe scatenare un uragano in Europa, dove inizialmente era previsto bel tempo, perché l'infinitesima variazione di pressione prodotta dal movimento delle ali potrebbe generare uno schema meteorologico completamente diverso.

Questo concetto può essere illustrato con un semplice esperimento di meccanica. Se si lascia cadere una sferetta metallica nell'apertura presente sulla sommità di una

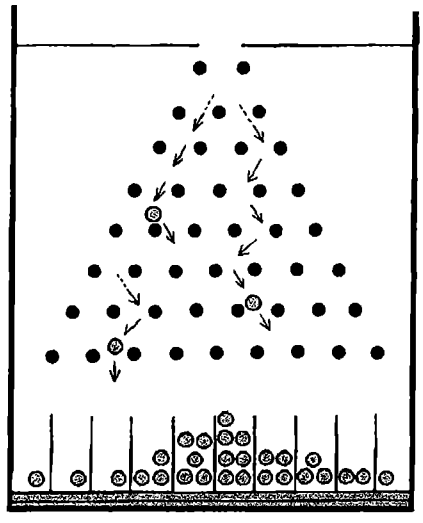
linea del tempo

1812 Laplace pubblica il saggio in cui espone la sua visione deterministica del mondo

1889 Poincaré si imbatte nel caos durante lo studio del problema dei tre corpi per cui riceve un premio dal re Oscar di Svezia

macchina di Galton, la pallina scenderà subendo una serie di deviazioni da parte dei chiodi in cui si imbatte, fino a raggiungere una delle scanalature praticate sul fondo del contenitore. Si potrebbe poi tentare di far cadere un'altra sferetta dalla stessa posizione iniziale e con la stessa velocità. Se la velocità e la posizione di partenza fossero *esattamente* le stesse, il marchese de Laplace avrebbe ragione: la pallina seguirebbe esattamente lo stesso percorso. Se la prima sfera fosse caduta nella terza scanalatura da destra, la seconda farebbe lo stesso.

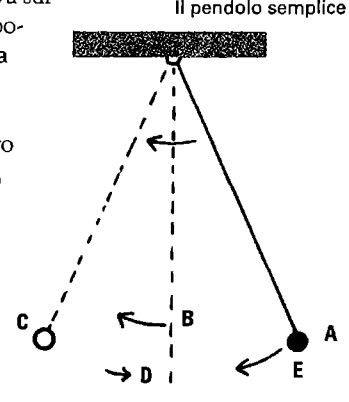
Ovviamente, però, è impossibile lasciar cadere la seconda sfera esattamente dalla stessa posizione, con la stessa velocità e la stessa forza: in pratica esisterà una differenza, magari talmente piccola da non essere neppure misurabile. La conseguenza è che la sfera potrà seguire un percorso del tutto diverso e probabilmente finirà in un'altra scanalatura.



L'esperimento della macchina di Galton

Il pendolo semplice Il pendolo semplice è uno dei sistemi meccanici più semplici da analizzare. Mentre oscilla avanti e indietro, il pendolo perde gradualmente energia; l'ampiezza dello spostamento rispetto alla verticale e la velocità (angolare) della massa diminuiscono sempre più, finché non si raggiunge uno stato di quiete.

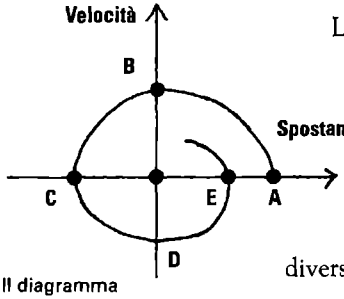
È possibile rappresentare il movimento della massa in un diagramma di fase, riportando sull'asse orizzontale lo spostamento (angolare) e sull'asse verticale la velocità. Il punto in cui il pendolo viene lasciato libero (A) si trova sull'asse orizzontale positivo del diagramma, perché inizialmente lo spostamento è massimo e la velocità è nulla. Quando la massa passa dall'asse verticale il suo spostamento è nullo e la velocità è massima (punto B), mentre in C, all'altro estremo dell'oscillazione, lo spostamento è negativo e la velocità è nulla. La massa poi torna indietro passando per D (dove si muove in senso opposto rispetto all'andata, cosicché la velocità è negativa) e completa l'oscillazione in E. Nel diagramma di fase il movimento completo corrisponderebbe a una rotazione di 360°, ma fra l'andata e il ritorno l'ampiezza dell'oscillazione si riduce, perciò il punto E cade *internamente* ad A. Man mano che l'oscillazione del pendolo si smorza, la rappresentazione della fase si avvicina spiraleggiando all'origine, finché il pendolo si ferma.



1961 Lorenz osserva l'effetto farfalla

1971 Robert May studia il caos a proposito della dinamica di popolazione

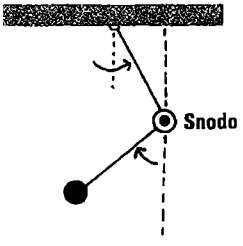
2004 La teoria del caos diventa nota al grande pubblico grazie al film *The butterfly effect*



Il diagramma di fase per il pendolo semplice

Le cose vanno diversamente nel caso del pendolo doppio, in cui la massa è posta all'estremità di una coppia di aste incernierate. Se lo spostamento è piccolo, il moto del pendolo doppio è analogo a quello del pendolo semplice, ma se è grande la massa oscilla, ruota e sbanda, muovendosi in modo apparentemente casuale attorno allo snodo intermedio. Se il moto non è forzato, la massa finirà per raggiungere anche in questo caso uno stato di quiete, ma la curva che descrive il suo moto è ben diversa dalla spirale regolare del pendolo semplice.

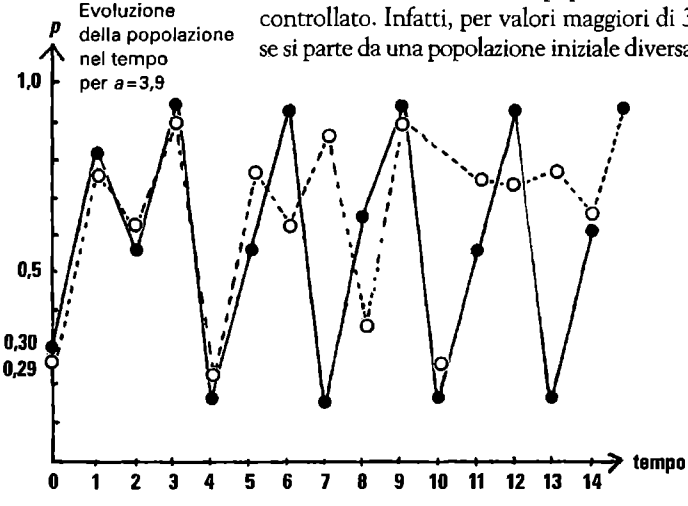
Moto caotico La caratteristica del caos è che il comportamento casuale sembra scaturire da un sistema deterministico. Vediamo un altro esempio. Consideriamo la formula ripetitiva (o iterativa) $a \times p \times (1 - p)$, dove p rappresenta una popolazione relativa misurata su una scala che va da 0 a 1. Perché il valore di p resti all'interno di questo intervallo, il valore di a deve essere compreso fra 0 e 4.



Il movimento del pendolo doppio

Calcoliamo il valore della popolazione per $a = 2$. Se come valore iniziale al tempo $t = 0$ si sceglie $p = 0{,}3$, per trovare la popolazione al tempo $t = 1$ bisogna sostituire $p = 0{,}3$ in $a \times p \times (1 - p)$, da cui con una semplice calcolatrice si ottiene 0,42. Si ripete poi l'operazione sostituendo nella formula il valore $p = 0{,}42$ e si ottiene il numero successivo (0,4872). Continuando in questo modo, si ricava la popolazione per tempi via via successivi. In questo caso il valore della popolazione si stabilizza rapidamente su $p = 0{,}5$. In effetti, per valori di a inferiori a 3 il risultato si stabilizza sempre su un valore di p.

Se a partire dalla stessa popolazione iniziale $p = 0{,}3$ si sceglie $a = 3{,}9$ (un valore prossimo al massimo consentito) la popolazione non si stabilizza, ma oscilla in modo incontrollato. Infatti, per valori maggiori di 3,57, a cade nella «zona caotica». Inoltre, se si parte da una popolazione iniziale diversa, come $p = 0{,}29$ (un valore prossimo a 0,3), la crescita della popolazione segue lo stesso andamento del caso precedente durante i primi cicli, ma poi diverge. Questo comportamento fu osservato sperimentalmente da Edward Lorenz nel 1961 (vedi riquadro).



Previsioni del tempo Come è noto, pur disponendo di computer estremamente potenti non siamo tuttora in grado di prevedere che tempo farà con un anticipo maggiore di due o tre giorni: basta qualche giorno in più perché le previsioni meteoro-

logiche ci riservino brutte sorprese. Il motivo è che le equazioni che regolano il clima non sono lineari, cioè non coinvolgono solo le variabili in questione, ma anche il loro prodotto.

La teoria matematica su cui si basa la meteorologia fu elaborata in modo indipendente dall'ingegnere francese Claude Navier nel 1821 e dal fisico matematico britannico George Gabriel Stokes nel 1845. Le risultanti equazioni di Navier-Stokes sono di alto interesse scientifico, tanto che il Clay Mathematics Institute di Cambridge (Stati Uniti) ha messo in palio un premio da un milione di dollari per chi compirà progressi sostanziali nello sviluppo di una teoria matematica in grado di svelarne i segreti. L'applicazione di queste equazioni al flusso dei fluidi ha consentito di comprendere molti aspetti dei moti stazionari nell'alta atmosfera, ma in prossimità della superficie terrestre il flusso d'aria produce turbolenze e dà luogo a un regime caotico di cui è impossibile prevedere l'evoluzione.

Dalla meteorologia alla matematica

L'effetto farfalla fu scoperto casualmente nel 1961 da Edward Lorenz, un meteorologo del MIT (Massachusetts Institute of Technology). Mentre il suo vecchio computer stampava dei grafici, Lorenz andò a prendere un caffè, ma al suo ritorno trovò un risultato inatteso: invece delle mappe meteorologiche che voleva ripetere aveva ottenuto dei grafici completamente diversi. La cosa lo stupì profondamente, perché aveva inserito gli stessi dati iniziali e quindi si aspettava di ottenere le stesse immagini. Era tempo di rottamare il suo vecchio computer e di procurarsi un calcolatore più affidabile? Ripensandoci, Lorenz si rese conto che in realtà esisteva una differenza fra i valori iniziali inseriti nei due casi: la prima volta aveva usato sei cifre decimali, la seconda ne aveva riportate solo tre. Per indicare l'effetto prodotto da questa piccola differenza, Lorenz coniò l'espressione «effetto farfalla». A seguito di questa scoperta il suo interesse scientifico si rivolse alla matematica.

Se la teoria dei sistemi lineari di equazioni è in gran parte nota, le equazioni di Navier-Stokes contengono termini non lineari che le rendono intrattabili; di fatto, l'unico modo per risolverle consiste nel procedere per via numerica utilizzando computer estremamente potenti.

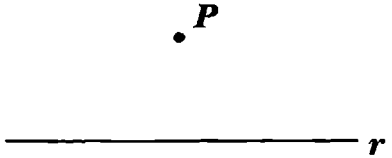
Attrattori strani Si può pensare che i diagrammi di fase dei sistemi dinamici contengano degli «attrattori». Nel caso del pendolo semplice l'attrattore è il singolo punto posto nell'origine verso cui è diretto il moto. Nel caso del pendolo doppio le cose si complicano, ma anche in questo caso la rappresentazione di fase del sistema presenta delle regolarità e viene attratta verso un insieme di punti del diagramma. Per sistemi di questo tipo l'insieme di punti può formare un frattale (vedi capitolo 25); in questo caso l'attrattore è dotato di una struttura matematica ben definita e viene detto «attrattore strano». Quindi, niente è perduto: da questa nuova teoria del caos non emerge il caos «caotico», ma il caos «regolare».

idea chiave

Un'imprevedibile regolarità

27 Il postulato delle parallele

Questa storia emozionante inizia con una semplice disposizione geometrica. Immaginiamo una retta *r* e un punto *P* esterno ad essa. Quante rette passanti per *P* e parallele alla retta *r* si possono tracciare? È evidente che esiste una sola retta passante per P che non incontrerà mai *r*, indipendentemente dalla sua estensione in entrambe le direzioni. Una variante di questa affermazione, che ci appare ovvia e dettata dal buon senso, costituisce uno dei postulati inclusi da Euclide di Alessandria in quel caposaldo della geometria che sono i suoi *Elementi*.



Non sempre ci si può fidare del buon senso, perciò è bene verificare se l'assunzione di Euclide è convincente dal punto di vista matematico.

Gli *Elementi* di Euclide La geometria euclidea è esposta nei tredici volumi degli *Elementi*, che furono scritti verso il 300 a.C. Gli *Elementi* di Euclide sono uno dei più importanti testi matematici di tutti i tempi. I matematici greci li consideravano la prima codificazione sistematica della geometria; gli scienziati di età più tarda li hanno studiati e tradotti dai manoscritti ancora esistenti, tramandandoli ai posteri con la convinzione universalmente condivisa che costituissero il paradigma perfetto della geometria.

Con la diffusione degli *Elementi* a livello scolastico, l'insegnamento della geometria si ridusse a una serie di letture dal «sacro testo», che però si dimostrò poco adatto agli allievi più giovani; come osservò con sarcasmo il poeta A.C. Hilton: «lo riscrivevano tutto senza un briciolo di esitazione, ma con una montagna di errori». D'altra parte, si potrebbe obiettare che era un testo destinato agli adulti, non ai

linea del tempo

circa 300 a.C.
Euclide include il postulato delle parallele nei suoi *Elementi*

1829-1831
Lobačevskij e Bolyai pubblicano i loro risultati sulla geometria iperbolica

ragazzi. In molti paesi gli *Elementi* hanno toccato l'apice dell'importanza didattica nel XIX secolo, ma restano una pietra miliare anche per i matematici odierni.

L'importanza del testo di Euclide risiede soprattutto nello stile in cui è scritto: il merito principale del libro è quello di presentare la geometria sotto forma di una serie di proposizioni dimostrate. Sherlock Holmes ne avrebbe apprezzato il metodo deduttivo, che consiste nel procedere in modo logico a partire da postulati enunciati con chiarezza, e avrebbe forse bacchettato Watson per il fatto di non considerarlo un «sistema freddo e spassionato».

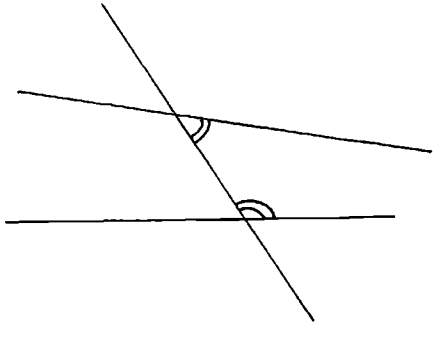
Pur costituendo le fondamenta dell'edificio della geometria euclidea, i postulati (ora chiamati «assiomi») non erano sufficienti, ed Euclide aggiunse una serie di «proposizioni» e di «nozioni comuni». Le proposizioni comprendono affermazioni del tipo «punto è ciò che non ha parti» e «linea è lunghezza senza larghezza», mentre le nozioni comuni stabiliscono, fra l'altro, che «il tutto è maggiore della parte» e che «le cose che sono uguali a una stessa cosa sono anche uguali fra loro». Solo alla fine del XIX secolo si è compreso che Euclide aveva fatto ricorso ad assunzioni implicite.

I postulati di Euclide

Una caratteristica della matematica è che un numero limitato di assunzioni può dare luogo a teorie di portata molto ampia. Un ottimo esempio è costituito dai cinque postulati di Euclide, che hanno rappresentato un modello per tutti i sistemi assiomatici successivi:

1. Risulti postulato: che si possa condurre una retta [un segmento di retta] da un qualsiasi punto ad ogni altro punto.
2. E che una retta terminata [finita] si possa prolungare continuamente in linea retta.
3. E che si possa descriver un cerchio con qualsiasi centro ed ogni distanza [raggio].
4. E che tutti gli angoli retti siano uguali fra loro.
5. E che se una retta, venendo a cadere su due rette, forma gli angoli interni e dalla stessa parte minori di due angoli retti, le due rette, prolungate indefinitamente, verranno a incontrarsi dalla parte in cui sono gli angoli minori di due retti.

Il quinto postulato A suscitare molte controversie, oltre duemila anni dopo la comparsa degli *Elementi*, è stato il quinto postulato di Euclide, che appare fuori luogo già solo per la verbosità e la goffaggine del suo enunciato. Lo stesso Euclide ne era insoddisfatto, ma fu costretto a includerlo per dimostrare alcune proposizioni dopo aver tentato inutilmente di derivarlo dagli altri postulati.



I matematici delle epoche successive cercarono di dimostrare il quinto postulato o di sostituirlo con uno più semplice.

1854 Lezioni di Riemann sui fondamenti della geometria

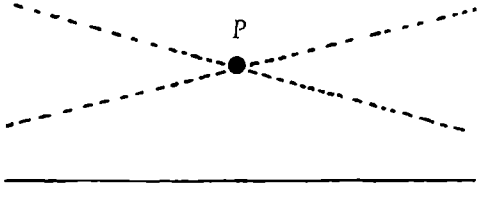
1872 Klein unifica la geometria mediante la teoria dei gruppi

1915 La teoria della relatività generale di Einstein si basa sulla geometria riemanniana

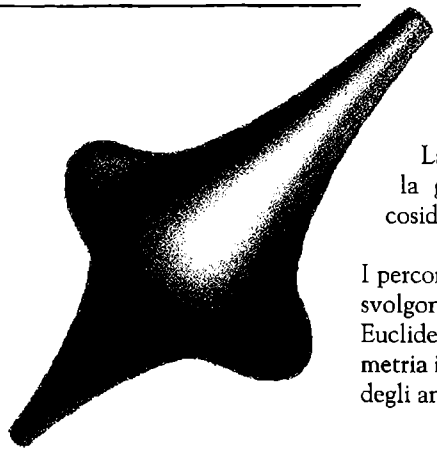
Nel 1795 John Playfair lo enunciò in una forma destinata ad avere successo: dati una retta r e un punto P esterno a essa, esiste *una e una sola* retta passante per P e parallela a r. Nello stesso periodo Adrien Marie Legendre lo sostituì con una versione equivalente, affermando che la somma degli angoli interni di un triangolo è uguale a due angoli retti, ovvero a 180 gradi. Questi enunciati risultavano più accettabili della farraginosa versione di Euclide ed entro certi limiti rispondevano all'accusa di artificiosità mossa al quinto postulato.

Un'altra linea di azione consisteva nel ricercare la sfuggente dimostrazione del postulato. Questo approccio piaceva particolarmente ai suoi sostenitori: se si fosse trovata una dimostrazione, il postulato sarebbe diventato un teorema, uscendo dall'occhio del ciclone. Purtroppo i tentativi di dimostrarlo si rivelarono fulgidi esempi di ragionamento circolare, un'argomentazione basata precisamente sull'assunto che intende dimostrare.

Geometrie non euclidee La svolta si verificò con le ricerche di Carl Friedrich Gauss, János Bolyai e Nikolaj Ivanovič Lobačevskij. Gauss non pubblicò i propri risultati, ma sembra appurato che li avesse ottenuti nel 1817. Bolyai pubblicò il suo lavoro nel 1831 e Lobačevskij nel 1829 (in modo indipendente), cosa che fece nascere fra i due una disputa per la paternità dei risultati. Questi tre scienziati di indubbia genialità riuscirono a dimostrare che il quinto postulato era indipendente dagli altri quattro e che, aggiungendo a questi ultimi la sua negazione, si otteneva un sistema coerente.



Bolyai e Lobačevskij costruirono una nuova geometria ammettendo la possibilità che esistesse *più di una* retta passante per P che non incontrava la retta r. Com'è possibile? È evidente che le rette tratteggiate (vedi figura a margine) incontreranno r. Se si accetta questa affermazione, però, si ricade inconsapevolmente nella visione euclidea. L'immagine inganna, perché Bolyai e Lobačevskij proponevano un nuovo tipo di geometria che non concordava con quella di Euclide, più rispondente al senso comune. La loro geometria non euclidea può essere considerata la geometria che si ha sulla superficie curva della cosiddetta «pseudosfera».



I percorsi più brevi che uniscono i punti di una pseudosfera svolgono la stessa funzione delle rette nella geometria di Euclide. Questa nuova geometria non euclidea è detta «geometria iperbolica» e una delle sue peculiarità è che la somma degli angoli di un triangolo è minore di 180 gradi.

Un'altra alternativa al quinto postulato afferma che *ogni* retta passante per P incontra la retta r; in altre parole, nessuna retta passante per P è «parallela» a r. Si tratta a tutti gli effetti di una nuova geometria, per quanto diversa da quella di Bolyai e Lobačevskij. Un esempio è costituito dalla geometria che si ha su una superficie sferica: in questo caso le rette della geometria euclidea sono sostituite dai cerchi massimi (di circonferenza uguale a quella della sfera) e la somma degli angoli di un triangolo è *maggiore* di 180 gradi. Questa «geometria ellittica» è associata al matematico tedesco Bernhard Riemann che l'ha studiata fra il 1850 e il 1860.

La geometria euclidea, che fino a quel momento era stata considerata l'unica vera geometria (secondo Immanuel Kant era «incisa nell'uomo»), veniva sbalzata dal suo piedistallo e diventava uno dei tanti sistemi possibili, stretto fra la geometria iperbolica e quella ellittica. Tutte queste versioni della geometria furono unificate in una teoria generale da Felix Klein nel 1872. L'avvento della geometria non euclidea sconvolse la matematica e spianò la strada alla geometria della relatività generale di Einstein (vedi capitolo 48). Infatti, la teoria della relatività generale richiedeva un nuovo tipo di geometria, la geometria dello spazio-tempo curvo, o «riemanniana»: la caduta dei corpi non veniva più attribuita all'effetto della forza di attrazione gravitazionale di Newton, ma veniva spiegata in termini di questa geometria non euclidea. Il concetto è che la presenza nello spazio di corpi massivi come la Terra e il Sole deforma lo spazio-tempo in modo analogo a un oggetto appoggiato su un sottile foglio di gomma: se una biglia crea una rientranza appena accennata, una palla da bowling produce un avvallamento molto più pronunciato.

Misurando la curvatura dello spazio-tempo per mezzo della geometria riemanniana è possibile prevedere correttamente la deviazione subita dai raggi luminosi in presenza di corpi spaziali massivi. Il comune spazio euclideo, di cui il tempo è una componente indipendente, non è più sufficiente per la relatività generale. Uno dei motivi è che lo spazio euclideo è piatto, privo di curvatura, come un foglio di carta appoggiato su un tavolo: si può dire che in ogni suo punto la curvatura è zero. Lo spazio-tempo riemanniano presuppone invece che la curvatura vari continuamente, come varia da un punto all'altro la curvatura di un pezzo di stoffa spiegazzato. È come osservare qualcosa in uno specchio deformante: l'immagine che si vede dipende dal punto dello specchio in cui si guarda.

Visti questi risultati, non c'è da stupirsi che negli anni 1850-1860 Gauss fosse rimasto così colpito dal giovane Riemann da affermare che la «metafisica» dello spazio sarebbe stata rivoluzionata dalle sue intuizioni.

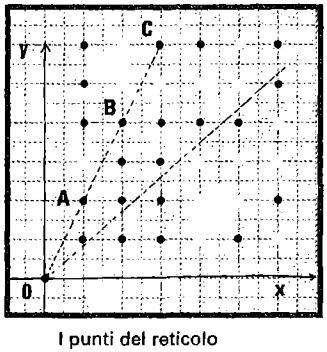
idea chiave

E se le parallele si incontr sfer

28 La geometria discreta

La geometria, che letteralmente significa «misura della Terra», è una delle attività più antiche dell'uomo. La geometria ordinaria studia le rette continue e le figure che possono essere descritte come una serie di punti posti «l'uno accanto all'altro». La matematica discreta non utilizza la successione continua dei numeri reali, ma i numeri interi e prende in esame reticoli costituiti da un numero finito di punti e di rette: il continuo viene sostituito da un insieme di elementi isolati.

Tipicamente, un reticolo (o griglia) è definito come un insieme di punti le cui coordinate sono numeri interi. La geometria discreta solleva problemi interessanti e trova applicazione negli ambiti più disparati, come la teoria della codifica e la progettazione degli esperimenti scientifici.



I punti del reticolo degli assi *x*/*y*

Osserviamo un fascio luminoso emesso da un faro. Immaginiamo che i raggi luminosi partano dall'origine O e abbiano un'inclinazione variabile fra l'asse orizzontale e quello verticale. Ci si può chiedere quali raggi colpiscano i vari punti del reticolo (che potrebbero essere, ad esempio, barche ancorate in porto secondo una disposizione uniforme).

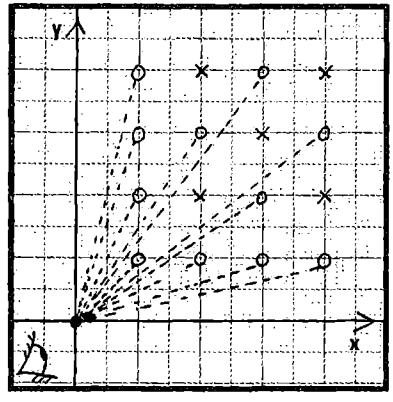
L'equazione di un raggio che parte dall'origine è $y = mx$, corrispondente all'equazione di una retta che passa per l'origine di coefficiente angolare m. Il raggio descritto da $y = 2x$ colpirà il punto di coordinate $x = 1$ e $y = 2$, perché questi valori soddisfano la sua equazione. Se il raggio colpisce un punto del reticolo di coordinate $x = a$ e $y = b$, il coefficiente angolare m è dato dalla frazione b/a; di conseguenza, se m non è una vera frazione (ad esempio se è uguale a $\pi\sqrt{2}$), il raggio luminoso mancherà tutti i punti del reticolo.

linea del tempo

1639
Ad appena sedici anni, Pascal scopre l'omonimo teorema

1806
Brianchon scopre il duale del teorema di Pascal

Il raggio luminoso $y = 2x$ colpisce il punto A di coordinate $x = 1$ e $y = 2$, ma non B di coordinate $x = 2$ e $y = 4$, né tutti gli altri punti che si trovano «nascosti dietro» A (come C, descritto da $x = 3$ e $y = 6$, e D, per cui $x = 4$ e $y = 8$). Possiamo immaginare di trovarci nell'origine e di stabilire quali punti sono visibili e quali invece sono nascosti.



I punti «visibili» dall'origine (O) e quelli «nascosti» (X)

Si può dimostrare che dall'origine sono visibili i punti le cui coordinate $x = a$ e $y = b$ sono numeri primi tra loro, cioè non ammettono altri divisori comuni a parte 1; un esempio è il punto $x = 2$ e $y = 3$, che nasconderà alla vista i punti le cui coordinate sono multiple delle sue (come $x = 4$ e $y = 6$, o $x = 6$ e $y = 9$, e così via).

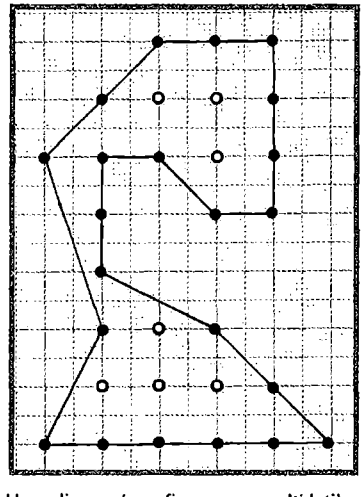
Il teorema di Pick Il matematico austriaco Georg Pick può aspirare alla fama per due motivi. Il primo è che era molto amico di Albert Einstein e contribuì a portarlo all'università tedesca di Praga nel 1911, quando Einstein era ancora un giovane scienziato; il secondo è che ha scritto un breve saggio sulla geometria «reticolare» pubblicato nel 1899. Nel corso della sua vita Pick ha effettuato una grande quantità di studi su un'ampia gamma di argomenti, ma è famoso soprattutto per l'affascinante teorema che porta il suo nome. E che teorema!

Il teorema di Pick consente di calcolare l'area racchiusa da un poligono (una figura con molti lati) ottenuto dall'unione di punti a coordinate intere. È la matematica del flipper.

Per ricavare l'area di questa figura bisogna contare il numero dei punti ● presenti sul perimetro e quello dei punti interni ○; nel nostro esempio, il primo è $b = 22$ e il secondo è $c = 7$. Conoscendo questi dati si può applicare il teorema di Pick:

$$\text{area} = b/2 + c - 1$$

Applicando questa formula si trova che l'area del poligono è uguale a $22/2 + 7 - 1 = 17$ unità quadrate. Tutto qui. Il teorema di Pick vale per qualunque figura costituita dall'unione di punti discreti a coordinate intere; l'unica condizione è che il bordo non si incroci mai.

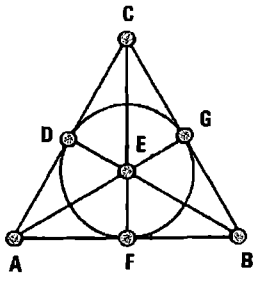


Un poligono (una figura con molti lati)

1846
Kirkman anticipa la scoperta dei sistemi tripli di Steiner (STS)

1892
Fano scopre l'omonimo piano, il più semplice esempio di geometria proiettiva

1899
Pick pubblica il suo teorema sull'area dei poligoni



Il piano di Fano

Il piano di Fano La geometria del piano di Fano fu scoperta più o meno contemporaneamente alla formula di Pick, ma non riguarda alcuna operazione di misura. Il piano che prende il nome dal matematico italiano Gino Fano, uno dei primi a studiare la geometria finita, è il più semplice esempio di geometria «proiettiva». È costituito solo da sette punti e sette rette.

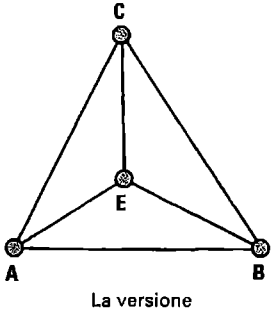
Nella figura a margine, i sette punti sono indicati con *A*, *B*, *C*, *D*, *E*, *F* e *G*. Sei rette sono chiaramente visibili, ma la settima è molto più difficile da individuare. Le proprietà della geometria e il modo in cui è costruito il diagramma portano a identificarla con *DFG*, la circonferenza passante per *D*, *F* e *G*. Ciò non crea alcun problema, perché nell'ambito della geometria discreta non è necessario che le linee siano «rette» nel senso tradizionale del termine.

Questa piccola geometria è dotata di molte proprietà, come:

- da ogni coppia di punti passa una e una sola retta
- per ogni coppia di rette esiste uno e un solo punto che appartiene a entrambe.

Queste due proprietà illustrano la straordinaria dualità delle geometrie di questo tipo. La seconda proprietà non è altro che la prima in cui sono stati invertiti i termini «punto» e «retta», così come la prima proprietà è uguale alla seconda tranne che per lo scambio dei due termini.

Il fatto che, scambiando di posto due termini e apportando qualche piccola modifica linguistica a un'affermazione vera, si ottenga un'altra affermazione vera dimostra che la geometria proiettiva è altamente simmetrica. Quella euclidea invece non lo è. Nella geometria di Euclide esistono rette parallele, ovvero coppie di rette che non si incontrano mai, e si può tranquillamente invocare il concetto di parallelismo, mentre nella geometria proiettiva tutte le coppie di rette si incontrano in un punto. Per i matematici ciò significa che la geometria euclidea è di natura inferiore.



La versione euclidea del piano di Fano

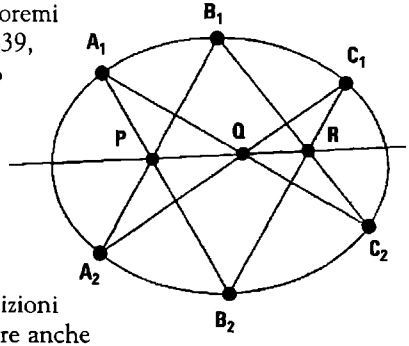
Eliminando dalla geometria di Fano una retta e i suoi punti ci si ritrova nell'ambito dell'asimmetrica geometria euclidea, che prevede l'esistenza delle rette parallele. Supponiamo di eliminare la retta «circolare» *DGF* in modo da ottenere un grafico euclideo.

Avendo eliminato una retta, ne restano sei: *AB*, *AC*, *AE*, *BC*, *BE* e *CE*. Ora esistono tre coppie di rette «parallele», ovvero *AB* e *CE*, *AC* e *BE*, *BC* e *AE*; infatti per presentare questo tipo di parallelismo le rette non devono avere punti in comune, come *AB* e *CE*.

Il piano di Fano è oggetto di grande venerazione in ambito matematico, perché è collegato a un gran numero di concetti e applicazioni: ad esempio, è una chiave del problema delle scolarette di Thomas Kirkman (vedi capitolo 41). Nella teoria della progettazione degli esperimenti, il piano di Fano si manifesta sotto forma di un esempio versatile, un sistema triplo di Steiner (STS). Un STS è un metodo che consente di suddividere un numero finito n di oggetti in blocchi di tre, in modo che ogni coppia di oggetti appartenga a uno e un solo blocco. Dati i sette oggetti A, B, C, D, E, F e G, i blocchi del STS corrispondono alle rette del piano di Fano.

A	F	B
B	G	C
C	A	D
D	B	E
E	C	F
F	D	G
G	E	A

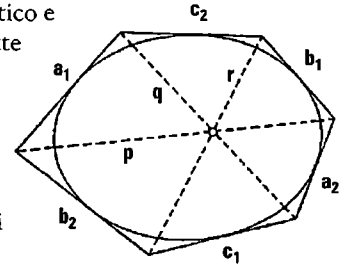
Teoremi duali Il teorema di Pascal e quello di Brianchon si collocano al confine fra geometria discreta e continua; si tratta di due teoremi diversi, ma collegati. Il primo fu scoperto da Blaise Pascal nel 1639, quando aveva solo sedici anni. Su una circonferenza allungata, o ellisse (vedi capitolo 22), individuiamo sei punti che indicheremo con A_1, B_1 e C_1, e A_2, B_2 e C_2. Chiamiamo P il punto di intersezione fra le rette A_1B_2 e A_2B_1, Q il punto di intersezione fra A_1C_2 e A_2C_1, e R il punto di intersezione fra B_1C_2 e B_2C_1. Il teorema afferma che i punti P, Q e R giacciono sulla stessa retta.



Il teorema di Pascal

Il teorema di Pascal è valido indipendentemente dalle posizioni occupate dai punti sull'ellisse. Addirittura continuerebbe a valere anche se si sostituisse l'ellisse con un'altra conica, come un'iperbole, un cerchio, una parabola o due rette parallele (nel qual caso si ottiene una configurazione chiamata «esagramma mistico»).

Il teorema di Brianchon fu scoperto molto tempo dopo dal matematico e chimico francese Charles-Julien Brianchon. Si tracciano sei rette tangenti a un'ellisse, indicandole con a_1, b_1 e c_1, e a_2, b_2 e c_2. Si definiscono poi tre diagonali p, q e r rispetto ai punti di intersezione fra le rette: p è la retta che unisce il punto in cui a_1 interseca b_2 a quello in cui a_2 interseca b_1, q è la retta che collega i punti di intersezione fra a_1 e c_2 e fra a_2 e c_1, e r è la retta che unisce il punto in cui b_1 interseca c_2 a quello in cui b_2 interseca c_1. Il teorema di Brianchon afferma che le rette p, q e r si incontrano in un punto.



Il teorema di Brianchon

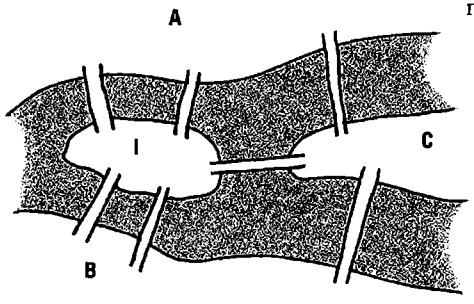
Il fatto che questi teoremi siano l'uno il duale dell'altro rappresenta un altro esempio del principio di dualità in geometria proiettiva.

idea chiave
Punti singolarmente interessanti

29 I grafi

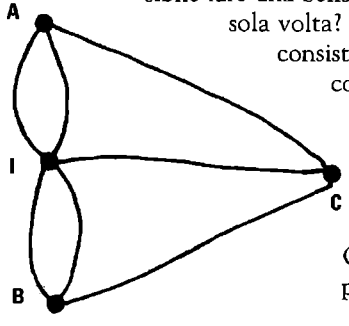
In matematica esistono i «grafici» e i «grafi». A scuola impariamo a disegnare i grafici, tracciando le curve che esprimono la relazione esistente fra le variabili *x* e *y*. I grafi sono un'invenzione più recente costituita da un insieme di punti collegati da linee.

Königsberg (l'attuale Kaliningrad) era una città della Prussia orientale famosa per i suoi sette ponti sul fiume Pregel. Oltre ad aver dato i natali all'illustre filosofo Immanuel Kant, Königsberg resta legata (con i suoi ponti) al famoso matematico Leonhard Euler, noto anche come Eulero.



Nel XVIII secolo fu posta una strana domanda: è possibile camminare per Königsberg attraversando ciascun ponte una e una sola volta? Non viene richiesto che la passeggiata si concluda nel punto di partenza; l'unico requisito è che ogni ponte venga attraversato una sola volta.

Nel 1735 Eulero espose all'Accademia Russa la sua soluzione, in cui oggi si individua l'origine della moderna teoria dei grafi. Nel diagramma semiastratto riportato a margine, l'isola fluviale è indicata con *I* e le sponde del fiume con *A*, *B* e C. È possibile fare una bella passeggiata domenicale che attraversi ogni ponte una e una sola volta? Proviamo a tracciare un percorso con una matita. Il segreto consiste nel rinunciare alla semiastrazione per giungere all'astrazione completa di un grafico costituito solo da punti e linee, detto «grafo». Le sponde del fiume sono rappresentate da «punti», i ponti che le uniscono da «linee». Il fatto che le linee non siano dritte o abbiano lunghezze diverse è del tutto ininfluente; ciò che importa sono i collegamenti.



Come osservò Eulero, per poter soddisfare il requisito del problema la passeggiata deve presentare una caratteristica: a

linea del tempo

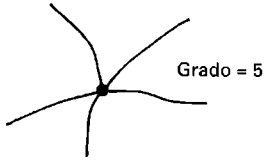
1735 Eulero risolve il problema dei ponti di Königsberg

1874 Carl Schorlemmer stabilisce un collegamento fra la chimica e gli «alberi»

eccezione del punto di partenza e di arrivo del percorso, ogni volta che si attraversa un ponte per arrivare in una regione sulla terraferma bisogna essere in grado di lasciarla per mezzo di un ponte mai percorso prima.

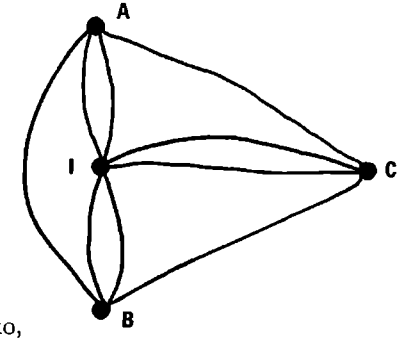
Trasferendo questo concetto nell'immagine astratta, si può dire che le linee che si incontrano in un punto devono presentarsi a coppie. A parte i due punti che rappresentano l'inizio e la fine della passeggiata, è possibile attraversare i ponti se, e solo se, in ogni punto converge un numero pari di linee.

Il numero di linee che si incontrano in un punto definisce il «grado» del punto in questione.



Il teorema di Eulero stabilisce che:

I ponti di una città possono essere attraversati una e una sola volta se tutti i punti a parte due (al massimo) sono di grado pari.



Osservando il grafo che rappresenta Königsberg, si vede che ogni punto è di grado dispari: ciò significa che in questa città è impossibile fare una passeggiata che attraversi ogni ponte una e una sola volta. Diventerebbe possibile solo a condizione di variare la disposizione dei ponti; se si costruisse un ponte aggiuntivo fra l'isola *I* e C, ad esempio, entrambi i punti diventerebbero di grado pari e quindi si potrebbe camminare da *A* a *B* percorrendo ciascun ponte una e una sola volta. Se si costruisse ancora un altro ponte, questa volta fra *A* e *B* (vedi figura a margine), si potrebbe partire da un punto qualsiasi e arrivare nello *stesso* punto, perché in questo caso *tutti* i punti sarebbero di grado pari.

Il lemma della stretta di mano Se ci venisse chiesto di disegnare un grafo contenente tre punti di grado dispari ci troveremmo in difficoltà. In effetti è impossibile, perché

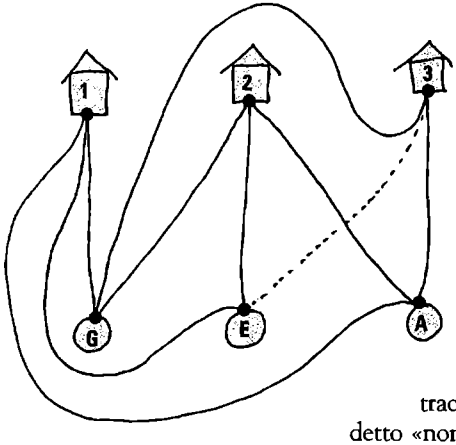
in ogni grafo, il numero di punti che hanno grado dispari deve essere pari.

1930
Kuratowski dimostra il suo teorema dei grafi planari

1935
George Pólya sviluppa alcune tecniche algebriche per il conteggio dei grafi

1999
Eric Rains e Neil Sloane estendono il conteggio delle strutture ad albero

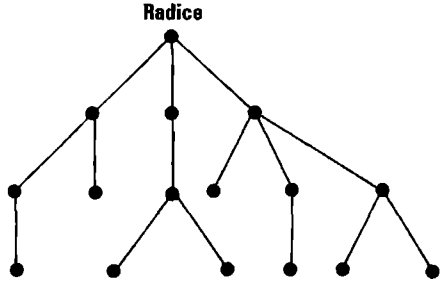
Si tratta del cosiddetto «lemma della stretta di mano», il primo teorema della teoria dei grafi. Ogni linea di un grafo ha un inizio e una fine (che è come dire che per stringersi la mano ci vogliono due persone), perciò la somma dei gradi di tutti i punti del grafo deve essere un numero pari, che chiameremo N. Ammettiamo che ci siano x punti di grado dispari e y di grado pari. Sommando tutti i gradi dispari si ottiene N_x e sommando tutti i gradi pari si ottiene N_y, che è pari. Si ha allora che $N_x + N_y = N$, e quindi $N_x = N - N_y$. Di conseguenza N_x è pari. Ma allora x è pari, poiché altrimenti N_x sarebbe dispari dato che la somma di un numero dispari di gradi dispari sarebbe a sua volta dispari.



Grafi non planari Il problema delle tre utenze è un vecchio rompicapo. Immaginiamo di dover fornire tre servizi pubblici (gas, elettricità e acqua) a tre case collegando ciascuna abitazione a ciascuna centrale, con un inghippo: i tubi non devono incrociarsi.

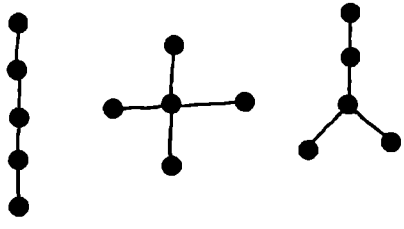
In realtà, è impossibile (ma niente impedisce di sottoporre il problema a un amico ignaro): il grafo che descrive il collegamento di tre punti con altri tre punti secondo tutte le combinazioni possibili (usando solo nove linee) non può essere tracciato nel piano senza produrre incroci. Un grafo di questo tipo è detto «non planare». Il grafo delle tre utenze e quello di tutte le linee che collegano cinque punti occupano un posto particolare nella teoria dei grafi. Nel 1930 il matematico polacco Kazimierz Kuratowski dimostrò il sorprendente teorema che afferma che un grafo è planare se, e solo se, non contiene nessuno di questi due grafi come sottografo (un grafo più piccolo contenuto all'interno del principale).

Alberi Un «albero» è un tipo particolare di grafo molto diverso da quello delle utenze o di Königsberg. Nel problema dei ponti di Königsberg era contemplata la possibilità di partire da un punto e di farvi ritorno seguendo un percorso diverso: un percorso che fa ritorno nel punto di partenza è detto «ciclo». Un albero è un grafo privo di cicli.



Un grafo ad albero noto a tutti è costituito dalla disposizione delle directory di un computer; le directory sono ordinate secondo una gerarchia che prevede una directory principale (detta *root*, ovvero «radice») da cui deriva una serie di subdirectory. Poiché non esistono cicli, per trasferirsi da un ramo all'altro è necessario passare dalla *root*, come ben sa chiunque usi il computer.

Contare gli alberi Quanti alberi diversi si possono realizzare con un dato numero di punti? Il problema di contare gli alberi fu affrontato da Arthur Cayley, un matematico inglese del XIX secolo. Ad esempio, con cinque punti si possono costruire esattamente tre alberi diversi:



Non disponendo di un computer, Cayley era in grado di contare il numero di alberi ottenibili solo da un numero molto limitato di punti. Prima che la complessità del calcolo diventasse eccessiva riuscì a considerare alberi costituiti da meno di 14 punti. Da allora i metodi di calcolo hanno fatto grandi progressi e si è potuto contare il numero di alberi composti da 22 punti, che è dell'ordine dei milioni.

Le ricerche di Cayley trovarono un'applicazione pratica già ai suoi tempi. Infatti il conteggio degli alberi è importante in chimica, perché le caratteristiche dei composti chimici dipendono anche dalla disposizione spaziale degli atomi che li costituiscono; composti formati dagli stessi atomi disposti diversamente hanno proprietà chimiche differenti. Questo tipo di analisi dei grafi ha consentito di prevedere «sulla carta» l'esistenza di alcuni composti chimici prima che venissero scoperti in laboratorio.

idea chiave

Attraverso il ponte e fra gli alberi

30 Il problema dei quattro colori

Chi avrebbe potuto regalare a Tiny Tim del *Racconto di Natale* di Dickens quattro matite colorate e una carta geografica politica bianca dell'Inghilterra? Forse il suo vicino cartografo che di tanto in tanto gli faceva qualche regalino o il matematico Augustus De Morgan, uno strano personaggio che viveva nei dintorni e passava la giornata con il padre di Tim. Certo non Scrooge.

La famiglia Cratchits viveva in una grigia casa a schiera di Bayham Street, nel quartiere londinese di Camden Town, appena più a nord dell'allora nuovissimo University College dove insegnava De Morgan. La provenienza del dono si sarebbe chiarita all'inizio del nuovo anno, quando il professore passò a trovarli per vedere se Tim aveva colorato la mappa.

De Morgan aveva le idee molto chiare sul procedimento da seguire: «Devi colorare la carta in modo che due contee confinanti siano di colore diverso».

«Ma non mi bastano i colori!» obiettò Tim senza riflettere. De Morgan sorrise, lasciandolo al suo compito. Poco tempo prima uno dei suoi studenti, Frederick Guthrie, lo aveva interpellato a quel proposito, accennando al fatto di essere riuscito a colorare la carta dell'Inghilterra servendosi solo di quattro colori e il problema aveva stuzzicato la curiosità matematica di De Morgan.

È possibile colorare *qualsiasi* carta geografica politica con quattro colori in modo da distinguere le varie regioni? I cartografi ne sono stati convinti per secoli, ma il problema consiste nel dimostrarlo in modo rigoroso. Oltre alla carta dell'Inghilterra, si può pensare alla mappa di un qualsiasi paese del mondo, ad esempio gli stati americani o i dipartimenti francesi, o anche a carte geografiche fittizie costituite da regioni e confini arbitrari. Tre colori, però, non sono sufficienti.

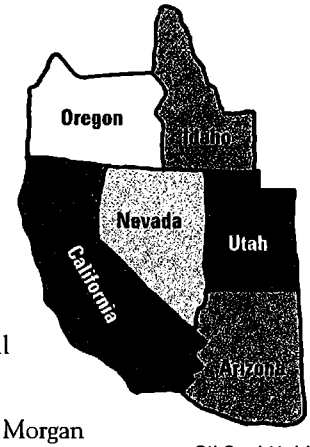
linea del tempo

1852
Guthrie, uno studente di De Morgan, gli sottopone il problema dei quattro colori

1879
Si pensa che Kempe abbia risolto il problema

1890
Heawood individua gli errori presenti nella dimostrazione di Kempe e dimostra un problema analogo con cinque colori

Esaminiamo la carta degli Stati Uniti occidentali. Avendo a disposizione solo blu, verde e rosso, si potrebbero iniziare a colorare il Nevada e l'Idaho; il colore da cui si parte è indifferente, perciò ammettiamo di scegliere il blu per il Nevada e il verde per l'Idaho. Fin qui tutto bene. Ciò significa che lo Utah deve essere rosso, e di conseguenza l'Arizona verde, la California rossa e l'Oregon verde: in questo modo, però, sia l'Oregon che l'Idaho sono verdi e quindi non sono distinguibili. Se fossero disponibili quattro colori (se ad esempio avessimo anche il giallo), potremmo usare il quarto colore per l'Oregon e tutto funzionerebbe benissimo. Quattro colori (blu, verde, rosso e giallo) sarebbero sufficienti per *qualunque* carta geografica? Questa domanda costituisce il problema dei quattro colori.



Gli Stati Uniti occidentali

Il dilagare del problema Nel giro di vent'anni da quando De Morgan ne riconobbe l'importanza, il problema dei quattro colori diventò famoso nella comunità matematica europea e americana. Negli anni 1860-1870 il matematico e filosofo americano Charles Sanders Peirce credette di averlo dimostrato, ma il suo ragionamento non ci è pervenuto.

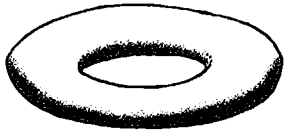
Il problema acquistò risalto grazie all'intervento dello scienziato vittoriano Francis Galton, che vi scorse un potenziale pubblicitario con cui allettò il celebre matematico di Cambridge Arthur Cayley, inducendolo a scrivere un articolo sull'argomento nel 1878. Purtroppo Cayley dovette ammettere di non essere riuscito a elaborare una dimostrazione, ma osservò che bastava limitarsi a considerare carte cubiche (in cui tre e tre soli paesi si incontrano in un punto). Ciò indusse il suo studente Alfred Bray Kempe a cimentarsi nella ricerca di una soluzione. Dopo appena un anno Kempe annunciò di aver trovato una dimostrazione: Cayley si congratulò caldamente con lui, la dimostrazione venne pubblicata e Kempe fu eletto membro della Royal Society di Londra.

E poi? La dimostrazione di Kempe era lunga e tecnicamente impegnativa e, a eccezione di un paio di persone che erano rimaste scettiche, venne accettata da tutti. Dieci anni dopo, però, ci fu un colpo di scena: Percy Heawood dell'Università di Durham scoprì una carta geografica che evidenziava un errore nel ragionamento di Kempe. Anche se non riuscì a fornire una propria dimostrazione, Heawood dimostrò che la sfida del problema dei quattro colori era ancora aperta. Il quesito sarebbe tornato sulle lavagne dei matematici, dando la possibilità ad altri novellini di apportare il loro contributo. Usando alcune tecniche di Kempe, Heawood

1976
Appel e Haken ottengono al computer una dimostrazione di validità generale

1994
La dimostrazione viene semplificata, ma è sempre assistita da computer

dimostrò il teorema dei cinque colori, che afferma che cinque colori sono sufficienti per colorare qualsiasi carta. Si sarebbe rivelato un ottimo risultato se qualcuno fosse riuscito a disegnare una carta che non poteva essere colorata con quattro colori. Nel frattempo, i matematici si trovavano ad affrontare un dilemma: quattro colori o cinque?



Ciambella semplice o «toro»

Nella sua forma più semplice, il problema dei quattro colori riguardava mappe tracciate su una superficie piatta o sferica, ma ci si poteva domandare cosa sarebbe successo se le carte fossero state disegnate su una superficie a forma di ciambella (naturalmente i matematici erano molto più interessati alla forma che al sapore della ciambella). Heawood dimostrò che sette colori erano necessari e sufficienti per colorare ogni mappa disegnata su una superficie di quel tipo e determinò il numero di colori che garantivano di poter colorare qualunque mappa nel caso di una ciambella a più buchi (con un numero di buchi pari a b), sebbene non avesse dimostrato che si trattava del numero minimo di colori. I primi valori di b trovati da Heawood sono riportati nella seguente tabella:



Un «toro» con due buchi

numero di buchi, b	1	2	3	4	5	6	7	8
numero di colori sufficienti, C	7	8	9	10	11	12	12	13

La formula generale è $C = [1/2(7 + \sqrt{(1 + 48b)})]$. Le parentesi quadre stanno a indicare che si considera solo la parte intera del termine contenuto all'interno; ad esempio, per $b = 8$, $C = [13{,}3107...] = 13$. La formula di Heawood è stata ricavata imponendo la condizione stringente che il numero di buchi fosse maggiore di zero; il fatto che sostituendo il valore proibito $b = 0$ si ottenga $C = 4$ appare molto promettente.

Problema risolto? Dopo cinquant'anni, il problema emerso nel 1852 era ancora irrisolto: all'inizio del XX secolo le migliori menti dell'*élite* matematica mondiale erano assai confuse.

In realtà, qualche passo avanti era stato fatto: un matematico aveva dimostrato che quattro colori erano sufficienti per una carta con 27 paesi, un altro aveva rilanciato con 31 e un terzo si era inserito proponendone 35. Si sarebbe potuto continuare in eterno a rosicchiare un valore dopo l'altro, ma le osservazioni contenute nei primissimi articoli di Kempe e Cayley indicavano che si poteva seguire un procedimento migliore: si scoprì che bastava controllare determinate disposizioni che potevano verificarsi nelle mappe per garantire che quattro colori fossero sufficienti. Il guaio era che tali disposizioni erano tantissime; nella fase iniziale di questi tentativi di dimostrazione se ne contavano migliaia. Un controllo di questo tipo non poteva essere effettuato manualmente, ma per fortuna il matematico tedesco Wolfgang Haken, che aveva studiato il problema per anni, riuscì ad assicurarsi la collaborazione di Kenneth Appel, un matematico americano esperto di informatica. Il numero di configurazioni da con-

trollare fu ridotto, con alcuni stratagemmi, a meno di 1500 e alla fine di giugno del 1976, dopo molte notti insonni, la procedura fu completata: Haken e Appel, in collaborazione con il fido computer IBM 370, avevano risolto quel formidabile problema.

Gli scienziati del dipartimento di matematica dell'Università dell'Indiana avevano un altro motivo di vanto e il bollo postale che celebrava la «scoperta del più grande numero primo» fu sostituito dall'annuncio «bastano quattro colori». Orgoglio locale a parte, come reagì la comunità matematica mondiale? Sebbene fosse un quesito di vecchia data comprensibile anche a un pubblico dell'età di Tiny Tim, era pur sempre un problema che per oltre un secolo aveva rappresentato una spina nel fianco dei più importanti matematici del pianeta.

Le reazioni furono miste. Alcuni ammisero a malincuore che il problema era stato risolto, ma molti altri rimasero scettici. Il guaio era che si trattava di una dimostrazione assistita dal computer e quindi si discostava dalla classica forma della dimostrazione matematica: che la dimostrazione del problema dei quattro colori sarebbe stata lunga e difficile da seguire era ampiamente prevedibile, ma una dimostrazione al computer andava oltre. In particolare, si poneva la questione della sua «verificabilità»: era impossibile verificare le migliaia di righe di codice da cui dipendeva, ma si sapeva anche che quando si scrive un codice gli errori sono all'ordine del giorno. E un solo errore poteva rivelarsi fatale.

Cosa ancora più grave, mancava del tutto il fattore «ah, ecco!». Non si poteva leggere la dimostrazione e apprezzare la sottigliezza del problema, o ricostruire la parte cruciale del ragionamento sperimentando il momento dell'«ah, ecco!». Uno dei critici più agguerriti del risultato ottenuto era il celebre matematico Paul Halmos, che attribuiva a una dimostrazione effettuata al computer la stessa attendibilità del vaticinio di una cartomante. D'altra parte, molti altri matematici riconoscono la validità del risultato e bisognerebbe essere molto audaci o molto sciocchi per dedicare una parte del proprio prezioso tempo di ricerca al tentativo di scovare un controesempio, individuando una mappa che richieda cinque colori. Qualcuno avrebbe forse potuto farlo prima della conclusione di Appel e Haken, ma non dopo.

Dopo la dimostrazione Rispetto al 1976 il numero di configurazioni da verificare si è dimezzato e i computer sono diventati molto più veloci e potenti, ma la comunità matematica aspetta ancora una dimostrazione più breve e di forma più tradizionale del teorema dei quattro colori. Nel frattempo, questo teorema ha generato importanti problemi di teoria dei grafi e ha contribuito a mettere in discussione il concetto stesso di dimostrazione matematica.

idea chiave

Bastano quattro colori

31 La probabilità

Quante probabilità ci sono che domani nevichi? Qual è la probabilità che il signor Bianchi riesca a prendere il primo treno del mattino? E che vinca alla lotteria? Oltre a essere invocata quotidianamente per esprimere la verosimiglianza di un certo scenario, la probabilità costituisce l'oggetto di una teoria matematica.

La teoria della probabilità è molto importante, perché è collegata all'incertezza ed è fondamentale per compiere la valutazione di un rischio. Se riguarda l'incertezza, però, è difficile pensare di poterla quantificare: dopo tutto, la matematica è una scienza esatta!

Il problema consiste appunto nel quantificare la probabilità.

Consideriamo un esempio semplicissimo, il lancio di una moneta. Qual è la probabilità che esca testa? Si potrebbe rispondere senza pensarci due volte che è 1/2 (espresso anche in forma decimale 0,5, o percentuale, 50%). Osservando la moneta ci si accerta del fatto che sia equilibrata, cioè che la probabilità di ottenere testa sia uguale a quella di ottenere croce: quindi la probabilità che esca testa è 1/2.

Tutti i casi in cui si ha a che fare con monete, palline all'interno di contenitori e altri esempi «meccanici» sono relativamente semplici. In generale, il calcolo delle probabilità viene affrontato seguendo due teorie fondamentali: un approccio consiste nel prendere in considerazione la simmetria a due facce della moneta, l'altro (il cosiddetto «approccio della frequenza relativa») nel ripetere l'esperimento un gran numero di volte e contare i casi in cui esce testa. Ma quanto è grande «un gran numero»? È facile convincersi che il rapporto fra il numero di volte in cui esce testa e quello in cui esce croce è circa 50 50, ma se l'esperimento venisse prolungato questa proporzione potrebbe variare.

Ci si può chiedere, ad esempio, se sia possibile effettuare una misura sensata della probabilità che domani nevichi. Anche in questo caso gli esiti possibili sono due (può nevicare o non nevicare), ma non è affatto ovvio che siano ugualmente probabili come avveniva nel caso della moneta.

linea del tempo

circa 1650-1660
Pascal e Huygens pongono le basi della teoria della probabilità

1785
Condorcet applica la probabilità all'analisi delle giurie e dei sistemi elettorali

Per stimare la probabilità che domani nevichi bisogna prendere in considerazione le condizioni atmosferiche attuali e una miriade di altri fattori, ma anche facendolo non siamo comunque in grado di determinare un valore preciso per la probabilità. Può però essere utile attribuire un «grado di credenza» al fatto che la probabilità sia bassa, media o alta. In matematica la probabilità si misura su una scala compresa fra 0 e 1: se un evento è impossibile la sua probabilità è 0, se è certo la probabilità è 1. Un valore di probabilità uguale a 0,1 indica che è poco probabile che l'evento si verifichi, un valore di 0,9 significa che è molto probabile.

Le origini della probabilità La teoria matematica della probabilità è nata nel XVII secolo con le discussioni fra Blaise Pascal, Pierre de Fermat e Antoine Gombaud (noto anche come Cavaliere di Méré) sui problemi riguardanti il gioco d'azzardo. In particolare, questi personaggi si scervellavano su un gioco molto semplice. Il quesito del Cavaliere di Méré era il seguente: è più probabile che esca un «sei» in quattro lanci di un dado o una «coppia di sei» in 24 lanci di due dadi? Su quale opzione scommettereste la camicia?

L'opinione prevalente a quel tempo era che convenisse scommettere sulla coppia di sei, perché era consentito un maggior numero di lanci. Questa convinzione però andò in frantumi quando si analizzarono le probabilità relative ai due casi. Vediamo come si effettua il calcolo.

Lancio di un dado: la probabilità che *non* esca un sei in un unico lancio è 5/6, perciò in quattro lanci la probabilità sarebbe $5/6 \times 5/6 \times 5/6 \times 5/6$, ovvero $(5/6)^4$. Infatti, poiché gli esiti dei lanci non si influenzano a vicenda o, come si dice, sono «indipendenti», la probabilità totale si ottiene moltiplicando la probabilità associata a ciascun lancio. Quindi la probabilità che esca almeno un sei è

$$1 - (5/6)^4 = 0{,}517746\ldots$$

Lancio di due dadi: la probabilità che *non* esca una coppia di sei in un lancio è 35/36; in 24 lanci diventa $(35/36)^{24}$.

Quindi la probabilità che esca almeno una coppia di sei è

$$1 - (35/36)^{24} = 0{,}491404\ldots$$

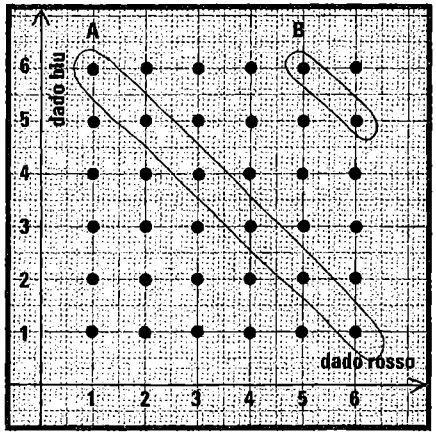
Spingiamoci un po' oltre in questo esempio.

1812 Laplace pubblica il testo in due volumi *Teoria analitica delle probabilità*

1912 Keynes pubblica il *Trattato sulla probabilità*, che influenza le sue teorie economiche e statistiche

1933 Kolmogorov elabora una trattazione assiomatica della probabilità

Giocare a craps L'esempio con i due dadi costituisce la base del moderno *craps*, un gioco d'azzardo con i dadi praticato nei casinò e in Internet. Lanciando due dadi distinguibili (uno rosso e uno blu) si producono 36 possibili esiti che è possibile registrare come coppie di dati nella forma (x, y) e rappresentare con 36 punti in un sistema di assi xy; l'insieme di tutti i possibili esiti è detto «spazio campionario» (o insieme universo) dell'esperimento.



Spazio campionario (per due dadi)

Supponiamo che l'«evento» A consista nell'ottenere due dadi di somma pari a 7. Poiché esistono 6 combinazioni di questo tipo, si può descrivere l'evento come

$$A = \{(1, 6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$

e cerchiarlo sul diagramma. La probabilità che si verifichi A è pari a 6 su 36, e può essere espressa come $Pr(A) = 6/36 = 1/6$. Se l'evento B corrisponde a ottenere due dadi di somma uguale a 11, si ha che $B = \{(5, 6), (6, 5)\}$ e $Pr(B) = 2/36 = 1/18$.

Il gioco del *craps* consiste nel lanciare due dadi su un tavolo: si vince o si perde al primo turno, ma se escono determinati punteggi c'è ancora speranza e si può procedere a un secondo turno. Per vincere al primo lancio bisogna ottenere un «natural», cioè deve verificarsi l'evento A o B; la probabilità di un natural è data dalla somma delle probabilità dei due eventi: $6/36 + 2/36 = 8/36$. Si perde al primo turno se il risultato del lancio è 2, 3 o 12 (detti «craps»). Con un calcolo analogo a quello già visto si trova che la probabilità di perdere al primo turno è 4/36. Se dal lancio si ottiene una somma di dadi pari a 4, 5, 6, 8, 9 o 10 si procede a un secondo turno; la probabilità che ciò avvenga è $24/36 = 2/3$.

Nel mondo dei giochi d'azzardo le probabilità vengono espresse come rapporto fra eventi sfavorevoli e favorevoli. Su 36 giochi di *craps* in media si vince al primo turno 8 volte e non si vince 28 volte, perciò le probabilità contrarie a una vittoria al primo lancio sono di 28 a 8, ovvero di 3,5 a 1.

La scimmia alla macchina da scrivere Alfredo, uno scimpanzé che vive allo zoo, ha una macchina da scrivere vecchia e scassata con 26 tasti per le lettere dell'alfabeto, più uno per il punto, uno per la virgola, uno per il punto interrogativo e uno per lo spazio, per un totale di 30 tasti. Se ne sta seduto in un angolo, pieno di ambizione letteraria, ma ha uno strano modo di scrivere: colpisce i tasti a caso.

Poiché ogni sequenza di lettere ha una probabilità non nulla di essere scritta, in principio esiste la possibilità che Alfredo scriva le opere teatrali di Shakespeare alla perfezione, dalla prima all'ultima parola.

Inoltre esiste anche una certa probabilità (per quanto piccolissima) che alle opere originali faccia seguire una loro traduzione in italiano, poi una in francese e una in spagnolo; per sicurezza, si può ammettere anche la possibilità che continui con la *Divina Commedia*. La probabilità che avvenga tutto ciò è minuscola, ma certamente non è nulla: il punto cruciale del ragionamento è questo. Vediamo quanto tempo impiegherà Alfredo per battere a macchina il monologo dell'*Amleto*, a cominciare dall'incipit «*To be or* [not to be]». Immaginiamo di avere 8 caselle contenenti queste 8 lettere (compresi gli spazi):

T	o		b	e		o	r

Ci sono 30 possibilità per la prima posizione, 30 per la seconda e così via, quindi le 8 caselle possono essere riempite in $30 \times 30 \times 30 \times 30 \times 30 \times 30 \times 30 \times 30$ modi diversi: la probabilità che Alfredo riesca a scrivere «*To be or*» è 1 su $6{,}561 \times 10^{11}$. Se colpisce un tasto ogni secondo, impiegherà circa 20 000 anni (rivelandosi fra l'altro un primate straordinariamente longevo), quindi non è il caso di restare con il fiato sospeso in attesa delle opere complete di Shakespeare: per la maggior parte del tempo Alfredo scriverà solo cose prive di senso come «xo,h?yt?».

Come si è sviluppata la teoria? I risultati derivanti dall'applicazione della teoria della probabilità possono essere discutibili, ma le sue basi matematiche sono ragionevolmente solide. Nel 1933 Andrej Nikolaevič Kolmogorov contribuì in modo determinante a elaborare una definizione assiomatica della probabilità, servendosi di un metodo molto simile a quello utilizzato due millenni prima per definire i princìpi della geometria.

La probabilità è definita dai seguenti assiomi:

1. la probabilità complessiva di tutti gli eventi è 1;
2. il valore della probabilità è maggiore o uguale a 0;
3. quando gli eventi sono incompatibili, è possibile sommarne le probabilità.

Da questi assiomi, enunciati in un linguaggio specialistico, si possono dedurre le proprietà matematiche della probabilità. Il concetto di probabilità trova applicazione in moltissimi ambiti ed è indispensabile per molti aspetti della vita moderna: analisi dei rischi, sport, sociologia, psicologia, progettazione ingegneristica, finanza... l'elenco è infinito. Chi poteva immaginare che i problemi sul gioco d'azzardo del XVII secolo, da cui sono nati questi concetti, avrebbero dato luogo a una disciplina così vasta? Quante probabilità c'erano?

idea chiave
Il segreto del giocatore

32 La teoria di Bayes

Della prima giovinezza del reverendo Thomas Bayes si sa pochissimo. Nato nel sud-est dell'Inghilterra, probabilmente nel 1702, Bayes diventò un ministro del culto non conformista, ma acquistò fama anche come matematico e nel 1742 fu eletto membro della Royal Society of London. Nel celebre *Saggio sulla soluzione di un problema della dottrina delle probabilità*, pubblicato nel 1763, due anni dopo la sua morte, Bayes ha presentato una formula per il calcolo della probabilità inversa e ha contribuito a elaborare un concetto fondamentale per la filosofia bayesiana: la probabilità condizionata.

Thomas Bayes ha dato il nome ai «bayesiani», i sostenitori di un tipo di statistica in contrasto con quella propugnata dagli statistici tradizionali, o «frequentisti». Infatti, mentre questi ultimi hanno una concezione della probabilità di tipo puramente numerico, le idee dei bayesiani sono fondate sulla celebre formula di Bayes e sul principio che i gradi di credenza soggettivi possono essere trattati come una probabilità matematica.

Probabilità condizionata Supponiamo che il dottor Ics debba diagnosticare il morbillo nei suoi pazienti. La comparsa di macchie rosse sulla pelle è un sintomo molto indicativo, ma la diagnosi non è affatto scontata: alcuni pazienti possono avere il morbillo senza presentare un'eruzione cutanea, altri possono avere un'eruzione ma non il morbillo. La probabilità che un paziente presenti un'eruzione *dato che* ha il morbillo è una probabilità condizionata. Nelle formule dei bayesiani «dato che» si indica con una linea verticale, perciò l'espressione

$$P(\text{un paziente presenta un'eruzione} \mid \text{il paziente ha il morbillo})$$

rappresenta la probabilità che un paziente abbia un'eruzione dato che ha il morbillo. Il valore di P(un paziente presenta un'eruzione | il paziente ha il mor-

linea del tempo

1763 Viene pubblicato il saggio di Bayes sulla probabilità

1937 De Finetti propugna la probabilità soggettiva in alternativa alla teoria della frequenza

billo) non è equivalente a P(il paziente ha il morbillo | un paziente presenta un'eruzione): in termini relativi, le due probabilità sono l'una l'inversa dell'altra e ciascuna può essere calcolata a partire dall'altra per mezzo della formula di Bayes. I matematici adorano rappresentare le cose con i simboli; assecondiamoli indicando con M l'evento che un paziente abbia il morbillo e con E l'evento che presenti un'eruzione. Il simbolo $\tilde{E}$ corrisponde all'evento che un paziente non abbia un'eruzione, mentre $\tilde{M}$ è l'evento che non abbia il morbillo. Queste informazioni possono essere visualizzate in un diagramma di Venn.

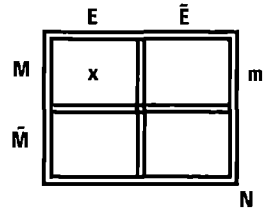


Diagramma di Venn che illustra la struttura logica della comparsa dell'eruzione e del morbillo

Il dottor Ics deduce quindi che x pazienti hanno il morbillo e un'eruzione, m hanno solo il morbillo e che il numero complessivo di pazienti è N. Dal diagramma, il medico capisce anche che la probabilità che qualcuno abbia il morbillo e un'eruzione è semplicemente x/N, mentre la probabilità che qualcuno abbia solo il morbillo è m/N. La probabilità condizionata, ovvero la probabilità che qualcuno abbia un'eruzione dato che ha il morbillo, si scrive $P(E \mid M)$ ed è uguale a x/m. Combinando queste informazioni, il dottor Ics ricava la probabilità che qualcuno abbia sia il morbillo che l'eruzione

$$P(M \cap E) = \frac{x}{N} = \frac{x}{m} \times \frac{m}{N}$$

ovvero

$$P(M \cap E) = P(E \mid M) \times P(M)$$

e analogamente

$$P(M \cap E) = P(M \mid E) \times P(E)$$

La formula di Bayes Uguagliando le espressioni per $P(M \cap E)$ si ottiene la formula di Bayes, che esprime la relazione esistente fra la probabilità condizionata e la sua inversa. Il dottor Ics si è fatto un'idea abbastanza precisa sul valore di $P(E \mid M)$, ovvero la probabilità che un paziente abbia un'eruzione dato che ha il morbillo, ma ciò a cui è realmente interessato è la probabilità condizionata inversa: vuole stimare la probabilità che un paziente abbia il morbillo dato che presenta un'eruzione. Questo problema di tipo inverso viene affrontato da Bayes nel suo saggio. Per calcolare le probabilità bisogna inserire nelle espressioni qualche valore numerico; si tratterà di valori arbitrari, ma l'importante è capire come si combinano. La probabilità che, se i pazienti

$$P(M|E) = \frac{P(M)}{P(E)} \times P(E|M)$$

La formula di Bayes

1950
Jimmy Savage e Dennis Lindley sono a capo della scuola bayesiana moderna

1950-1960
Viene introdotto il termine «bayesiano»

1992
Viene fondata la Società internazionale per l'analisi bayesiana

hanno il morbillo, presentino un'eruzione, cioè $P(E \mid M)$, è alta (ad esempio 0,9) e il dottor Ics, in questo modo, ha un'idea della situazione. Inoltre il medico sarà in grado di stimare la percentuale di popolazione affetta dal morbillo; ad esempio, potrebbe essere il 20% ovvero $P(M) = 0,2$. L'ultima informazione di cui abbiamo bisogno è $P(E)$, la percentuale di popolazione che presenta un'eruzione. La probabilità che qualcuno abbia un'eruzione si ottiene sommando la probabilità che qualcuno abbia sia il morbillo che un'eruzione e la probabilità che qualcuno non abbia il morbillo ma abbia un'eruzione. In base alle relazioni fondamentali viste sopra, $P(E) = 0,9 \times 0,2 + 0,15 \times 0,8 = 0,3$. Sostituendo questi valori nella formula di Bayes si ottiene:

$$P(M \mid E) = \frac{0,2}{0,3} \times 0,9 = 0,6$$

La conclusione è che, fra tutti i pazienti visitati che presentano un'eruzione, il medico diagnostica correttamente il morbillo nel 60% dei casi. Supponiamo ora che il dottor Ics riceva altre informazioni sui sintomi del morbillo che facciano aumentare la probabilità di riconoscerlo: la probabilità di presentare un'eruzione causata dal morbillo, $P(E \mid M)$, sale da 0,9 a 0,95, mentre quella che l'eruzione sia dovuta a una causa diversa, $P(E \mid \bar{M})$, scende da 0,15 a 0,1. Di quanto migliora la percentuale di casi di morbillo diagnosticati correttamente a seguito di questa variazione? Qual è il nuovo valore di $P(M \mid E)$? Con le nuove informazioni,

$$P(E) = 0,95 \times 0,2 + 0,1 \times 0,8 = 0,27;$$

sostituendo nella formula di Bayes si trova che

$$P(M \mid E) = \frac{0,2}{0,27} \times 0,95 = 0,704$$

Grazie alle informazioni supplementari che ha ricevuto, il dottor Ics è in grado di diagnosticare il 70% dei casi di morbillo. Se le probabilità diventassero, rispettivamente, 0,99 e 0,01, la probabilità di individuare la malattia, $P(M \mid E)$, sarebbe uguale a 0,961; in questo caso la probabilità di fare una diagnosi corretta sarebbe all'incirca del 96%.

I bayesiani attuali Gli statistici tradizionali non trovano niente da ridire sull'utilizzo della formula di Bayes nei casi in cui la probabilità è misurabile. Il punto controverso riguarda l'interpretazione della probabilità in termini di gradi di credenza, ovvero, come viene anche definita, di «probabilità soggettiva».

In tribunale, a volte, la colpevolezza o l'innocenza di un imputato viene stabilita in base al «bilancio delle probabilità». Mentre lo statistico frequentista non attribuisce alcun significato alla probabilità che un imputato sia colpevole di un crimine, il bayesiano non ha niente in contrario a farsi carico delle opinioni. In che modo? Vediamo con un esempio come si applica il metodo del bilancio delle probabilità per discriminare fra colpevolezza e innocenza.

Dopo aver assistito alla discussione di una causa nell'aula di un tribunale, un giurato ha deciso che la probabilità che l'accusato sia colpevole è circa 1 su 100. Dalla Camera di consiglio, il giurato viene poi richiamato in aula per assistere alla presentazione di altre prove da parte dell'accusa. A casa dell'imputato è stata rinvenuta un'arma e il pubblico ministero afferma che la probabilità di trovarla è uguale a 0,95 in caso di colpevolezza dell'imputato, ma si riduce a 0,1 se questo è innocente: quindi la probabilità di trovare un'arma è molto maggiore nel caso in cui l'imputato è colpevole rispetto al caso in cui è innocente. A questo punto il giurato si domanda se sia giusto o meno modificare la propria opinione riguardo alla colpevolezza dell'imputato alla luce di questo aggiornamento. Usando la consueta notazione, C rappresenta l'evento che l'imputato sia colpevole e N è l'evento corrispondente alla disponibilità di nuove prove. La probabilità «a priori», cioè la stima iniziale del giurato, è $P(C) = 1/100 = 0{,}01$. La probabilità $P(C \mid N)$, ovvero la probabilità della colpevolezza aggiornata alla luce delle nuove prove N, è detta «probabilità a posteriori». La formula di Bayes

$$P(C \mid N) = \frac{P(C \mid N)}{P(N)} \times P(C)$$

illustra il concetto dell'aggiornamento della probabilità a priori per dare la probabilità a posteriori $P(C \mid N)$. Come nell'esempio del medico avevamo calcolato $P(E)$, così ora possiamo calcolare $P(N)$:

$$P(C \mid N) = \frac{0{,}95}{0{,}95 \times 0{,}01 + 0{,}1 \times 0{,}99} \times 0{,}01 = 0{,}088$$

Ciò pone un dilemma al giurato, perché dalla stima iniziale dell'1% la probabilità della colpevolezza è aumentata fin quasi al 9%. Se l'accusa avesse affermato che la probabilità di trovare l'arma incriminante era 0,99 in caso di colpevolezza dell'imputato e 0,01 in caso della sua innocenza, applicando la formula di Bayes il giurato avrebbe dovuto aggiornare la sua stima dall'1% al 50%.

L'utilizzo della formula di Bayes in situazioni di questo tipo ha suscitato molte polemiche, riguardanti soprattutto il modo in cui si ricava la probabilità a priori. A favore dell'analisi bayesiana c'è il fatto che fornisce un metodo per trattare le probabilità soggettive e aggiornarle sulla base di nuove prove. Il metodo bayesiano trova applicazione negli ambiti più disparati, come la scienza, la meteorologia e le scienze forensi. I suoi fautori ne sostengono la validità intrinseca, oltre a esaltare il pragmatismo con cui tratta l'incertezza. In effetti, questo metodo ha molti punti a suo favore.

idea chiave

33 Il problema del compleanno

Immaginiamo di trovarci su un autobus e di non avere niente di meglio da fare che contare i nostri compagni di viaggio. È probabile che tutti i passeggeri siano indipendenti l'uno dall'altro, perciò possiamo presumere che i loro compleanni siano distribuiti in modo casuale in tutto il corso dell'anno. In totale i passeggeri sono 23; non è un numero molto alto, ma è comunque sufficiente per affermare che esiste una probabilità superiore al 50% che due passeggeri compiano gli anni nello stesso giorno. Incredibile, vero? Eppure è così. Perfino un grande esperto di teoria della probabilità, William Feller, la considerava una conclusione stupefacente.

In realtà, l'autobus è troppo piccolo per i nostri scopi, perciò riprendiamo il ragionamento spostandoci in una grande stanza: quante persone vi si devono trovare perché si abbia la certezza che due di esse compiano gli anni nello stesso giorno? Poiché un anno ha 365 giorni (per semplicità ignoriamo gli anni bisestili), se nella stanza ci fossero 366 persone, almeno due compirebbero *sicuramente* gli anni nello stesso giorno: sarebbe matematicamente impossibile che li compissero tutti in un giorno diverso.

È il principio della colombaia: se $n + 1$ piccioni sono distribuiti in n nicchie, una nicchia deve contenere più di un piccione. Se invece le persone fossero 365, non potremmo essere certi che due di esse compiano gli anni nello stesso giorno, perché i compleanni potrebbero cadere ciascuno in un giorno diverso dell'anno. Tuttavia, scegliendo 365 persone a caso ciò non accade quasi mai: la probabilità che non siano presenti neanche due persone che compiono gli anni nello stesso giorno è minima. Anche se la stanza contiene solo 50 persone, c'è una probabilità del 95% che due di esse condividano il giorno del compleanno.

linea del tempo

1654 — Blaise Pascal pone le basi della teoria della probabilità

1657 — Christiaan Huygens pubblica la prima opera riguardante la probabilità

1718 — Abraham de Moivre pubblica *La dottrina degli azzardi*, con nuove edizioni ampliate nel 1738 e nel 1756

Se il numero dei presenti si riduce ulteriormente, la probabilità che due persone compiano gli anni nello stesso giorno diminuisce. Si trova quindi che 23 è il numero di persone per cui tale probabilità è appena maggiore di 1/2 e 22 è quello per cui è appena minore di 1/2: il numero 23 rappresenta il valore critico. La risposta al problema classico del compleanno è sorprendente, ma non paradossale.

Si può dimostrare? Per convincercene, scegliamo una persona a caso. La probabilità che un'altra persona compia gli anni nello stesso giorno è 1/365, perciò la probabilità che queste due persone non condividano il giorno del compleanno è 1 – 1/365, cioè 364/365. La probabilità che il compleanno di una terza persona scelta a caso coincida con quello di una delle altre due è 2/356, mentre la probabilità dell'evento opposto è 1 – 2/365, cioè 363/365. La probabilità che nessuno dei tre condivida il giorno del compleanno con un altro si ottiene moltiplicando queste due probabilità, cioè (364/365) × (363/365) = 0,9918.

Continuando il ragionamento per 4, 5, 6,... persone, si risolve il paradosso del problema del compleanno. Quando, calcolatrice alla mano, arriviamo a svolgere il calcolo per 23 persone, troviamo che la probabilità che nessuno compia gli anni nello stesso giorno è 0,4927. La negazione di «nessuno compie gli anni nello stesso giorno» è «almeno due persone compiono gli anni nello stesso giorno», la cui probabilità è 1 – 0,4927 = 0,5073, appena maggiore del fatidico 1/2.

Per $n = 22$, la probabilità che due persone condividano il giorno del compleanno è 0,4757, che è minore di 1/2. La natura apparentemente paradossale del problema del compleanno è collegata al linguaggio. La conclusione di questo problema consiste nell'affermare che due persone compiono gli anni nello stesso giorno, senza specificare chi sono queste persone: non sappiamo chi riguarderà l'abbinamento. Se nella stanza è presente il signor Rossi, che compie gli anni l'8 marzo, si potrebbe formulare un'altra domanda.

Quanti compleanni coincidono con quello del signor Rossi?

In questo caso il calcolo è diverso. La probabilità che il signor Rossi compia gli anni in un giorno differente da quello di tutti gli altri è 364/365, perciò la probabilità che il suo compleanno non coincida con quello di nessuna delle altre $n - 1$ persone presenti nella stanza è $(364/365)^{n-1}$. Quindi la probabilità che il signor Rossi condivida il giorno del compleanno con qualcun altro è $1 - (364/365)^{n-1}$.

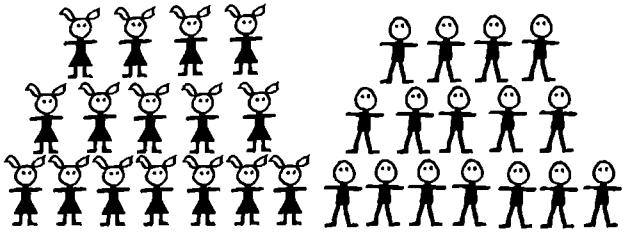
Svolgendo il calcolo per $n = 23$ si trova che il valore assunto da questa probabilità è 0,061151: la probabilità che qualcun altro oltre al signor Rossi compia gli anni l'8

marzo è quindi solo del 6%. All'aumentare del valore di n aumenta anche quello della probabilità, ma perché diventi maggiore di 1/2 bisogna arrivare a $n = 254$ (compreso il signor Rossi), quando è uguale è 0,5005. Siamo al punto di svolta, perché con $n = 253$ si ottiene 0,4991, che è minore di 1/2. La probabilità che il signor Rossi compia gli anni nello stesso giorno di qualcun altro diventa maggiore di 1/2 quando nella stanza sono presenti 254 persone. Probabilmente questa conclusione è più intuitiva della sorprendente affermazione che conclude il classico problema del compleanno.

Altri problemi relativi ai compleanni Esistono molte generalizzazioni del problema del compleanno. Un approccio consiste nel considerare tre persone che compiono gli anni nello stesso giorno: perché la probabilità di questo evento sia maggiore di 1/2, complessivamente devono essere presenti 88 persone. Se il requisito diventa che il compleanno sia comune a quattro persone, a cinque persone, ecc., sono necessari gruppi proporzionalmente più numerosi. Nel caso di una folla di 1000 persone, ad esempio, c'è una probabilità maggiore del 50% che nove dei presenti compiano gli anni nello stesso giorno.

Altri tentativi di affrontare il problema del compleanno hanno preso in esame compleanni vicini. In un problema di questo tipo, l'abbinamento riguarda compleanni che cadono entro un certo numero di giorni di distanza. La conclusione a cui si arriva è che sono sufficienti 14 persone perché la probabilità che i compleanni di due di esse cadano nello stesso giorno o a un giorno di distanza l'uno dall'altro sia maggiore del 50%.

Una variante del problema del compleanno che richiede l'uso di tecniche matematiche più sofisticate è il problema del compleanno di maschi e femmine: se una classe scolastica è costituita da un ugual numero di ragazzi e di ragazze, qual è il gruppo più piccolo per cui la probabilità che un ragazzo e una ragazza compiano gli anni nello stesso giorno raggiunge il 50%?



La risposta è che il gruppo più piccolo sarebbe costituito da una classe di 32 ragazzi (16 maschi e 16 femmine), rispetto alle 23 persone necessarie nella versione classica del problema del compleanno.

Variando leggermente il quesito si ottengono altri casi (a cui però non è facile dare risposta). Supponiamo che si sia creata una lunga coda all'entrata del concerto di Bob Dylan e che continuino ad aggiungersi persone in modo casuale; visto che ci interessano i compleanni, escludiamo la possibilità che arrivino insieme coppie o

triplette di gemelli. All'entrata, a ogni fan viene chiesto quando compie gli anni. Il quesito matematico è il seguente: quante persone vengono fatte entrare prima che ne capitino due *di seguito* che compiono gli anni nello stesso giorno? Altra domanda: quante persone entrano prima che se ne presenti una che compie gli anni lo stesso giorno del signor Rossi (8 marzo)?

I calcoli che riguardano i compleanni si basano sul presupposto che i giorni in cui cadono siano uniformemente distribuiti e che una persona scelta a caso abbia la stessa probabilità di compiere gli anni in un qualunque giorno. I risultati sperimentali dimostrano che ciò non è propriamente vero (nei mesi estivi la natalità è più alta), ma si avvicina a sufficienza alla realtà da produrre una soluzione applicabile in pratica.

I quesiti sui compleanni sono esempi di problemi di occupazione, in cui i matematici riflettono sui modi in cui è possibile disporre un certo numero di palline all'interno di alcune caselle. Nel problema del compleanno il numero di caselle a disposizione è 365 (corrispondenti ai possibili compleanni) e le palline da disporre in modo casuale al loro interno sono le persone. È possibile semplificare il problema riducendolo al calcolo della probabilità che due palline vadano a occupare la stessa casella; nel caso del problema dei ragazzi e delle ragazze le palline sono di due colori diversi.

I matematici non sono gli unici a interessarsi al problema del compleanno. Satyendra Nath Bose era affascinato dalla teoria della luce basata sui fotoni proposta da Albert Einstein; discostandosi dalle linee di ricerca tradizionali, Bose tradusse il sistema fisico in un problema di occupazione in cui le caselle non rappresentavano i giorni, come nel problema del compleanno, ma i livelli energetici dei fotoni; invece di disporre le persone nelle caselle vi distribuì un gran numero di fotoni. I problemi di occupazione trovano numerose applicazioni anche in altre scienze: ad esempio, in biologia la diffusione delle epidemie può essere trattata matematicamente come un problema di questo tipo, in cui le caselle sono aree geografiche, le palline sono malattie e il problema consiste nel calcolare come sono raggruppate le malattie.

Il mondo è pieno di coincidenze sorprendenti, ma solo la matematica ci consente di calcolarne la probabilità. Il classico problema del compleanno è solo la punta dell'iceberg e costituisce un'ottima introduzione a un tipo di matematica più impegnativa dotata di importanti applicazioni.

idea chiave

Il calcolo delle coincidenze

34 Le distribuzioni di probabilità

Ladislaus J. Bortkiewicz era affascinato dalle tabelle di mortalità: lungi dall'intristirlo, ai suoi occhi rappresentavano un inesauribile oggetto di indagine scientifica. Fra i suoi studi, resta celebre il conteggio del numero di cavalieri dell'esercito prussiano uccisi dal calcio di un cavallo. È stata poi la volta di Frank Benford, un ingegnere elettrico che ha contato la prima cifra di dati numerici di vario tipo per stabilire in quanti casi era uguale a uno, in quanti a due, e così via. George Kingsley Zipf, un filologo che insegnava tedesco ad Harvard, ha invece analizzato le occorrenze delle parole in vari brani di testo.

Tutti gli esempi citati riguardano la misura della probabilità che si verifichi un certo evento. Quante probabilità ci sono che in un anno x cavalieri muoiano per il calcio di un cavallo? L'insieme dei valori assunti dalla probabilità al variare del valore di x è detto «distribuzione di probabilità». Inoltre si tratta di una distribuzione *discreta*, nel senso che x può assumere solo valori isolati: i valori di interesse sono separati da intervalli, perché i cavalieri prussiani stroncati da un calcio potranno essere 3 o 4, ma non 3,5. Come vedremo, nel caso della distribuzione di Benford si prende in esame soltanto la comparsa delle cifre 1, 2, 3,... mentre in quello della distribuzione di Zipf la parola inglese *it* può occupare l'ottava posizione nella graduatoria delle parole più frequenti, ma non la posizione 8,23.

Questioni di vita o di morte nell'esercito prussiano Bortkiewicz registrò gli eventi di interesse per dieci corpi d'armata su un periodo di vent'anni, raccogliendo complessivamente i dati relativi a 200 corpi-anno. Bortkiewicz annotava il numero dei decessi (ciò che i matematici chiamano «variabile») e il numero di corpi-anno in cui si verificava. Ad esempio, in 109 corpi-anno non si era verificato alcun decesso, mentre in un corpo-anno ne erano avvenuti quattro: in un determinato anno, il corpo d'armata C aveva subìto quattro perdite.

linea del tempo

1837 Siméon-Denis Poisson descrive la distribuzione che porta il suo nome

1881 Newcomb scopre la legge che prenderà il nome di Benford

1898 Bortkiewicz analizza i dati sui decessi dei cavalieri dell'esercito prussiano

Com'è distribuito il numero dei decessi? La raccolta di informazioni di questo tipo, cioè la registrazione dei dati sul campo, è uno dei compiti degli statistici. Bortkiewicz ottenne i seguenti dati:

Numero di decessi	0	1	2	3	4
Frequenza	109	65	22	3	1

Fortunatamente, morire per il calcio di un cavallo è un evento raro. Lo strumento teorico più adatto per descrivere matematicamente la frequenza con cui si verificano eventi rari è la cosiddetta «distribuzione di Poisson». Se Bortkiewicz avesse utilizzato questa tecnica, avrebbe potuto prevedere i risultati senza bisogno di recarsi di persona nelle stalle? In base alla distribuzione teorica di Poisson, la probabilità che il numero di decessi (che chiameremo x) assuma il valore r è data dalla formula di Poisson, dove e è quel particolare numero ($e \approx 2{,}718...$) associato alla crescita che abbiamo già incontrato (vedi capitolo 06) e il punto esclamativo indica il fattoriale, ovvero il prodotto fra un numero e tutti i numeri interi compresi fra 1 e il numero stesso. La lettera greca lambda (λ) rappresenta il numero medio di decessi. Per calcolare questa media su 200 corpi-anno bisogna moltiplicare 0 decessi per 109 morti-anno (che fa 0), 1 decesso per 65 corpi-anno (che fa 65), 2 decessi per 22 corpi-anno (44), 3 decessi per 3 corpi-anno (9) e 4 per 1 corpo-anno (4), poi sommare tutti questi risultati (122) e dividere per 200. In conclusione, il numero medio di decessi per corpo-anno è 122/200 = 0,61.

$$e^{-\lambda}\lambda^{x}/x!$$

Formula di Poisson

Le probabilità teoriche (che indicheremo con p) si calcolano sostituendo i valori $r = 0, 1, 2, 3$ e 4 nella formula di Poisson. I risultati sono:

Numero di decessi	0	1	2	3	4
Probabilità, p	0,543	0,331	0,101	0,020	0,003
Numero atteso di decessi, $200 \times p$	108,6	66,2	20,2	4,0	0,6

La distribuzione teorica sembra riprodurre bene i dati sperimentali raccolti da Bortkiewicz.

Le prime cifre Se esaminiamo l'ultima cifra dei numeri riportati in una pagina dell'elenco telefonico, ci aspettiamo di trovare che 0, 1, 2,..., 9 presentano una distribuzione uniforme. Le cifre compaiono in modo casuale, perciò ciascuna ha la stessa probabilità di saltare fuori. Nel 1938 l'ingegnere elettrico Frank Benford si accorse che

1938 Benford formula diversamente la legge di distribuzione delle prime cifre

1950 Zipf ricava una formula che collega l'utilizzo delle parole all'estensione del vocabolario

2003 La distribuzione di Poisson viene applicata all'analisi delle colonie di pesci dell'Atlantico settentrionale

ciò non valeva per la prima cifra di alcuni insiemi di dati, riscoprendo di fatto una legge osservata per la prima volta nel 1881 dall'astronomo Simon Newcomb.

Facciamo un piccolo esperimento. Osserviamo i dati riguardanti il cambio della valuta estera su un quotidiano. Un tasso di cambio pari a 2,119 può indicare, ad esempio, che ci vogliono 2,119 dollari USA per acquistare 1 sterlina. Allo stesso modo, per comprare 1 sterlina sono necessari 1,59 euro e 15,390 dollari di Hong Kong. Passando in rassegna tutti i dati e contando il numero di volte in cui un certo numero compare come prima cifra, si ottiene la seguente tabella:

Prima cifra	1	2	3	4	5	6	7	8	9	Totale
Numero di occorrenze	18	10	3	1	3	5	7	2	1	50
Percentuale, %	36	20	6	2	6	10	14	4	2	100

Questi risultati confermano la legge di Benford, secondo la quale per alcuni insiemi di dati il numero 1 appare come prima cifra all'incirca nel 30% dei casi, il numero 2 nel 18% e così via. È un risultato ben diverso dalla distribuzione uniforme dell'ultima cifra dei numeri telefonici.

Non è chiaro il motivo per cui tanti insiemi di dati seguono la legge di Benford. Nel XIX secolo, quando notò questa distribuzione utilizzando le tavole dei logaritmi, Simon Newcomb non poteva certo immaginare che si sarebbe rivelata così diffusa.

Alcuni insiemi di dati che seguono la distribuzione di Benford sono i punteggi degli eventi sportivi, i dati azionari, i numeri civici, la popolazione di vari paesi e la lunghezza dei fiumi. È interessante notare che le unità di misura sono ininfluenti: non ha alcuna importanza se le lunghezze dei fiumi sono misurate in metri o in miglia. La legge di Benford trova varie applicazioni pratiche; in particolare, la scoperta che i dati contabili seguono tale legge ha consentito di individuare più facilmente i dati falsificati, smascherando un gran numero di frodi.

Parole Fra i suoi mille interessi, G.K. Zipf annoverava anche la singolare abitudine di contare le parole. Le dieci parole che compaiono più spesso nella lingua inglese sono i monosillabi riportati nella tabella qui sotto con il corrispondente rango:

Rango	1	2	3	4	5	6	7	8	9	10
Parola	the	of	and	to	a	in	that	it	is	was

Zipf ottenne questo risultato semplicemente contando le parole che costituivano ampi brani di una grande varietà di opere scritte. A quella che ricorreva più di frequente attribuì rango 1, alla successiva rango 2, e così via. Se si analizzassero molti altri testi si potrebbero riscontrare piccole variazioni in questo ordine, ma nell'insieme non cambierebbe molto.

Il fatto che *the* sia la parola più diffusa e che *of* occupi il secondo posto non è certo sorprendente. La lista continua e magari potrà interessarci sapere che *among* occupa il 500° posto e *neck* il 1000°, ma ai nostri fini ci limiteremo a considerare le prime dieci parole. Se prendiamo un qualsiasi testo e contiamo quante volte vi compaiono questi termini, otterremo più o meno la stessa classifica. La cosa davvero sorprendente, invece, è che esiste un collegamento fra il rango e il numero di volte in cui una parola compare effettivamente nel testo. La parola *the* compare due volte più spesso di *of*, tre volte più spesso di *and*, e così via. Il numero delle occorrenze effettive viene fornito da una celebre formula, una legge empirica che Zipf ricavò dai dati raccolti. La legge di Zipf afferma che la percentuale di occorrenze della parola di rango r è data da

$$\frac{k}{r} \times 100$$

dove il numero k dipende solo dall'estensione del vocabolario dell'autore del testo. Se un autore utilizzasse tutte le parole della lingua inglese (si stima che siano circa un milione) il valore di k sarebbe circa 0,0694. In base alla legge di Zipf, in questo caso la parola *the* rappresenterebbe il 6,94% circa delle parole di un testo e *of* costituirebbe la metà di quella percentuale, ovvero all'incirca il 3,47% delle parole. Un saggio di 3000 parole scritto da questo bravissimo autore conterrebbe quindi 208 volte *the* e 104 volte *of*.

Per scrittori che si servono solo di 20 000 parole il valore di k sale a 0,0954, quindi in un testo di questo tipo *the* comparirebbe 286 volte e *of* 143 volte. Più ridotto è il vocabolario, più spesso comparirà la parola *the*.

Guardando nella sfera di cristallo Le distribuzioni di Poisson, di Benford e di Zipf consentono di effettuare delle previsioni; magari non saremo in grado di prevedere un evento con certezza assoluta, ma sapere come si distribuiscono le probabilità è sempre meglio che fare un salto nel buio. Con queste tre distribuzioni e altre di tipo diverso (binomiale, binomiale negativa, geometrica, ipergeometrica e molte altre) gli statistici dispongono di un armamentario tecnico che consente loro di analizzare efficacemente un'ampia gamma di attività umane.

idea chiave

La previsione delle quantità

35 La curva di distribuzione normale

La curva di distribuzione «normale» svolge un ruolo fondamentale in statistica, tanto che è stata anche definita l'equivalente statistico della retta. Per quanto sia dotata di importanti proprietà matematiche, se ci mettessimo ad analizzare blocchi di dati grezzi scopriremmo che non seguono quasi mai una distribuzione perfettamente normale.

La distribuzione normale è descritta da una particolare formula matematica che genera una curva «a campana», cioè una curva costituita da una gobba che diminuisce da entrambi i lati fino a svanire. L'importanza della curva di distribuzione normale non risiede tanto nella sua natura, quanto nella teoria su cui è basata e in questo senso si inserisce in una gloriosa tradizione. Nel 1733 Abraham de Moivre, un ugonotto francese fuggito in Inghilterra per scampare alle persecuzioni religiose, introdusse questa distribuzione nel contesto dei suoi studi sulla probabilità. Pierre Simon de Laplace pubblicò alcuni risultati al riguardo, mentre Carl Friedrich Gauss la utilizzò in ambito astronomico, dove è anche nota come «legge gaussiana degli errori».

Nei suoi studi sociologici pubblicati nel 1835, Adolphe Quetelet si servì della curva normale per misurare la deviazione rispetto all'«uomo medio»; nel corso di altri esperimenti misurò l'altezza delle reclute francesi e la circonferenza toracica dei soldati scozzesi e ipotizzò che seguissero questo tipo di distribuzione. A quel tempo vigeva la ferrea convinzione che la maggior parte dei fenomeni seguisse questo tipo di comportamento «normale».

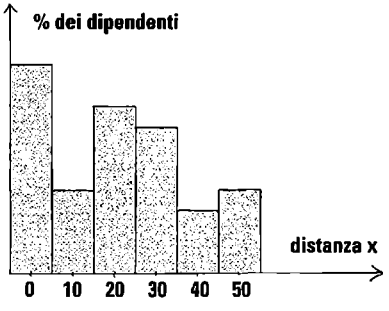
La festa Supponiamo che Gloria sia andata a una festa e che Stefano, il padrone di casa, le abbia chiesto se abitava lontano. In seguito Gloria si è resa conto che è un'ottima domanda da fare a una festa: si adatta a chiunque, non è difficile e richiede una risposta, perciò consente di rompere il ghiaccio se la conversazione stenta ad avviarsi.

linea del tempo

1733
De Moivre pubblica alcuni risultati sulla curva normale come approssimazione della distribuzione binomiale

1820
Gauss utilizza la distribuzione normale (detta anche «gaussiana») in astronomia come legge degli errori

Il giorno dopo, mentre andava in ufficio (con un leggero cerchio alla testa) Gloria si è domandata se i suoi colleghi abitassero lontano. Alla mensa aziendale è venuta a sapere che alcuni vivevano dietro l'angolo e altri a 50 kilometri di distanza; c'era un ampio grado di variabilità. In qualità di responsabile delle risorse umane dell'azienda, Gloria ha potuto aggiungere una domanda al questionario che ogni anno sottopone a ciascun dipendente: «Che distanza hai percorso oggi per venire al lavoro?». Il suo scopo era quello di calcolare la distanza media percorsa ogni giorno dai dipendenti dell'azienda. Quando Gloria ha tracciato l'istogramma dei risultati, non ha ottenuto una distribuzione con una forma particolare, ma almeno è stata in grado di calcolare la distanza media percorsa.



Istogramma di Gloria che rappresenta la distanza percorsa dai suoi colleghi per andare al lavoro

La media ottenuta è 20 kilometri. In matematica la media si indica con la lettera greca μ (mu), perciò in questo caso $\mu = 20$. La variabilità all'interno della popolazione, o «deviazione standard», si indica invece con la lettera greca σ (sigma). La deviazione standard è piccola quando i dati sono molto simili e poco variabili, mentre è grande quando i dati assumono valori molto diversi. L'analista di marketing dell'azienda (che ha studiato statistica) ha dimostrato a Gloria che avrebbe ottenuto la stessa media effettuando un campionamento: non era necessario porre la domanda a *tutti* i dipendenti. Questa tecnica di valutazione si basa sul teorema del limite centrale.

Consideriamo un campione di dipendenti scelto in modo casuale nella totalità del personale. L'ideale sarebbe avere un campione il più grande possibile, ma 30 dipendenti vanno già benissimo. Se il campione è stato scelto casualmente, è probabile che contenga persone che abitano dietro l'angolo e altre che vengono da molto lontano; nel calcolo della distanza media relativa al campione le distanze più lunghe si medieranno con quelle più brevi. I matematici indicano la media del

1835

Quetelet utilizza la curva normale per misurare la deviazione rispetto all'uomo medio

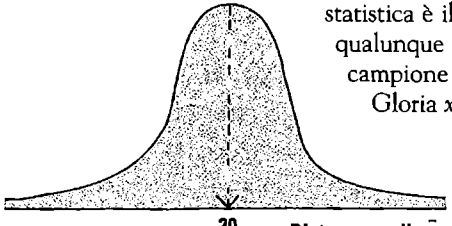
1870-1880

Questo tipo di distribuzione viene chiamata «normale»

1901

Aleksandr Lyapunov dimostra il teorema del limite centrale in modo rigoroso utilizzando funzioni caratteristiche

campione con $\bar{x}$. Nel caso di Gloria è estremamente probabile che il valore di $\bar{x}$ sia prossimo a 20, la media della popolazione. Per quanto sia senz'altro possibile, è improbabile che la media del campione sia molto piccola o molto grande.



Distribuzione media del campione

Uno dei motivi per cui la curva di distribuzione normale è tanto importante in statistica è il teorema del limite centrale. Questo teorema afferma che, per qualunque distribuzione di x, la distribuzione effettiva delle medie del campione $\bar{x}$ approssima una curva normale. Cosa significa? Nel caso di Gloria x rappresenta la distanza dal luogo di lavoro e $\bar{x}$ è la media di un campione. Anche se la distribuzione di x nell'istogramma di Gloria non assomiglia affatto a una curva a campana, la distribuzione di $\bar{x}$ ha questa forma ed è centrata attorno a $\mu = 20$.

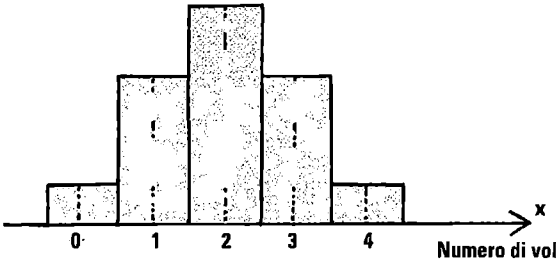
È per questo motivo che la media di un campione $\bar{x}$ costituisce una stima della media della popolazione μ. Un altro motivo di interesse sono le informazioni relative alla variabilità delle medie del campione $\bar{x}$. Se la dispersione dei valori di x corrisponde alla deviazione standard σ, la dispersione di $\bar{x}$ è $\sigma/\sqrt{n}$, dove n è la dimensione del campione scelto. Maggiore è la dimensione del campione, più stretta sarà la curva normale e migliore sarà la stima di μ.

Altre distribuzioni normali Facciamo un semplice esperimento: lanciamo in aria una moneta per quattro volte. Per ogni lancio la probabilità che esca «testa» è 1/2. Si può registrare l'esito dei quattro lanci indicando «testa» con T e «croce» con C, e riportando le facce ottenute nell'ordine in cui si presentano. In totale gli esiti possibili sono 16. Ad esempio, potrebbe uscire «testa» per tre volte secondo la sequenza CTTT, però esistono quattro possibili esiti che contengono tre volte «testa» (le altre combinazioni sono TCTT, TTCT, TTTC), quindi la probabilità che si ottenga tre volte testa è 4/16 = 0,25.

Se il numero di lanci è ridotto, le relative probabilità sono facili da calcolare e possono essere riassunte in una tabella; inoltre si può anche calcolarne la distribuzione. La riga del numero di combinazioni può essere estratta dal triangolo di Tartaglia (vedi capitolo 13):

Numero di teste	0	1	2	3	4
Numero di combinazioni	1	4	6	4	1
Probabilità	0,0625	0,25	0,375	0,25	0,0625
	(= 1/16)	(= 4/16)	(= 6/16)	(= 4/16)	(= 1/16)

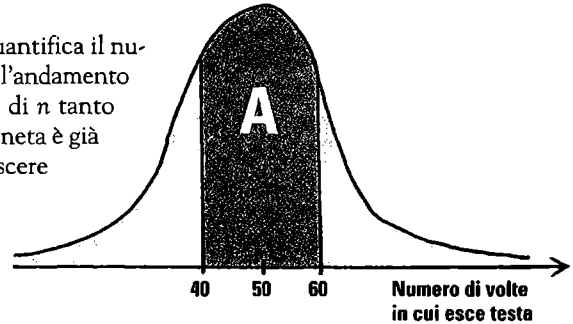
Questa «distribuzione binomiale di probabilità» ha luogo quando possono verificarsi due esiti (in questo caso, testa e croce). Le probabilità possono essere rappresentate per mezzo di un istogramma formato da rettangoli le cui dimensioni forniscono i dati possibili.



Numero di volte in cui esce testa in quattro lanci di una moneta in base alla distribuzione binomiale

Considerare solo quattro lanci è un po' riduttivo. Cosa succede se si lancia la moneta un numero di volte maggiore, per esempio 100? Per $n = 100$ si può applicare la distribuzione binomiale di probabilità, ma è anche possibile approssimarla con una curva normale a campana con media $\mu = 50$ (perché ci aspettiamo che lanciando 100 volte una moneta esca 50 volte testa) e dispersione (deviazione standard) $\sigma = 5$, come ha scoperto de Moivre nel XVI secolo.

All'aumentare del valore di n, la variabile x che quantifica il numero di eventi favorevoli riproduce sempre meglio l'andamento della curva normale: quanto maggiore è il valore di n tanto migliore è l'approssimazione e 100 lanci di una moneta è già un valore grande. Ammettiamo ora di voler conoscere la probabilità di ottenere da 40 a 60 volte testa. La regione a cui siamo interessati è l'area A, che rappresenta la probabilità che esca da 40 a 60 volte testa, ovvero $P(40 \leq x \leq 60)$. Il valore numerico deve essere ricavato da tabelle matematiche precalcolate, da cui si trova che $P(40 \leq x \leq 60) = 0{,}9545$. Ciò dimostra che la probabilità che su 100 lanci di una moneta esca testa un numero di volte compreso fra 40 e 60 è del 95,45%: è molto probabile che ciò avvenga.



Distribuzione della probabilità che esca testa in 100 lanci di una moneta

L'area residua è $1 - 0{,}9545 = 0{,}0455$. Poiché la curva normale è simmetrica rispetto al centro, la probabilità che in 100 lanci di una moneta si ottenga più di 60 volte testa corrisponde alla metà di questo valore, ovvero ad appena il 2,275%. È un valore di probabilità davvero molto basso: se fosse associato a una scommessa, sarebbe meglio lasciar perdere.

idea chiave

36 La correlazione fra i dati

Qual è il legame esistente tra due insiemi di dati? Gli statistici pensano di aver trovato una risposta già da un secolo. La correlazione e la regressione si comportano come cavallo e carrozza, si muovono di pari passo, ma come questi svolgono compiti diversi. La correlazione misura quanto due grandezze, ad esempio il peso e l'altezza, dipendano l'una dall'altra. La regressione serve invece a prevedere i valori assunti da una grandezza, come il peso, a partire da quelli dell'altra, in questo caso l'altezza.

La correlazione di Pearson Il termine «correlazione» fu introdotto da Francis Galton verso il 1870-1880 nella forma iniziale «co-relazione», che aveva il pregio di chiarirne meglio il significato. Galton, vero gentleman della scienza dell'epoca vittoriana, aveva la pretesa di misurare tutto e applicò la correlazione alle sue ricerche sulle coppie di variabili, come l'apertura alare e la lunghezza della coda degli uccelli. Il coefficiente di correlazione di Pearson, che prende il nome dal suo pupillo e biografo Karl Pearson, si misura su una scala che va da –1 a + 1. Se il suo valore numerico è alto, ad esempio + 0,9, si dice che tra le variabili esiste una forte correlazione. Il coefficiente di correlazione misura la tendenza dei dati ad allinearsi lungo una retta: se è vicino a zero, di fatto i dati non sono correlati.

Stabilire se due variabili sono correlate e qual è il loro grado di correlazione è un problema che si pone di frequente. A titolo di esempio si possono considerare le vendite di occhiali da sole e vedere come sono collegate a quelle dei gelati. Come luogo dove condurre la ricerca scegliamo Roma: i dati verranno raccolti ogni mese. Se si riportano in un grafico in cui la coordinata x (orizzontale) rappresenta la vendita di occhiali da sole e la coordinata y (verticale) la vendita dei gelati, per ogni mese si ottiene un punto (x, y) che rappresenta entrambi i dati. Il punto (3, 4), per esempio, potrebbe indicare che, a maggio, dalla vendita degli occhiali da sole si sono ricavati 30 000 euro, mentre nello stesso mese, in città, i gelati hanno fruttato 40 000 euro.

linea del tempo

1806 Adrien-Marie Legendre usa il metodo dei minimi quadrati per interpolare i dati

1809 Carl Friedrich Gauss applica il metodo dei minimi quadrati a problemi di astronomia

1885-1888 Galton introduce la regressione e la correlazione

Si possono poi rappresentare i punti corrispondenti ai dati mensili (x, y) registrati su un anno intero in un diagramma di dispersione. Il valore del coefficiente di correlazione di Pearson per questo esempio sarebbe uguale all'incirca a + 0,9, corrispondente a una forte correlazione. Il valore del coefficiente è positivo perché la retta ha un'inclinazione positiva, cioè punta verso nord-est.

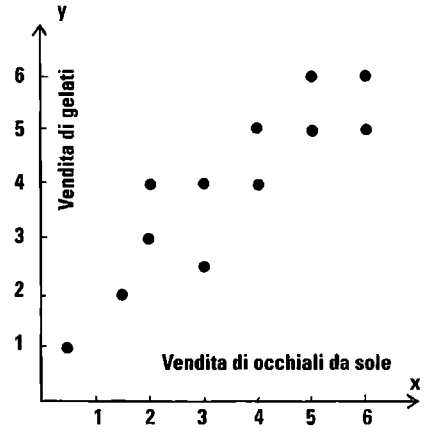


Diagramma di dispersione

Correlazione e causa Il fatto che due variabili siano fortemente correlate non è sufficiente per affermare che l'una sia la causa dell'altra. Può anche darsi che siano legate da una relazione di causa-effetto, ma non è lecito dedurlo da indizi puramente numerici. Per sciogliere il nodo fra causa e correlazione di solito si ricorre al termine «associazione»: è meglio andarci piano ed evitare le affermazioni troppo forti.

Nell'esempio visto sopra, esiste una forte correlazione tra le vendite di occhiali da sole e quelle di gelati: all'aumentare delle une tendono ad aumentare anche le altre. Sarebbe tuttavia ridicolo affermare che l'aumento della spesa per gli occhiali da sole sia avvenuto *a causa* dell'incremento nelle vendite dei gelati. Quando si ha a che fare con la correlazione entra spesso in gioco, come intermediaria, una variabile nascosta. Nel caso in questione, il legame tra le vendite degli occhiali da sole e dei gelati dipende da fattori stagionali, ovvero il caldo d'estate e il freddo d'inverno. Ma nell'utilizzare la correlazione si corre anche un altro rischio: le variabili possono essere fortemente correlate anche in assenza di qualunque nesso logico o scientifico. Ad esempio, può darsi che i numeri civici siano fortemente correlati all'età complessiva degli abitanti delle case, ma non sarebbe certamente il caso di leggervi un significato.

La correlazione di Spearman La correlazione trova altri impieghi: il coefficiente di correlazione può essere adattato per il trattamento di dati ordinati, dei quali si vuole sapere quale viene per primo, quale per secondo, per terzo, ecc., senza indagarne altri valori numerici.

In alcuni casi i dati sono costituiti unicamente dalle posizioni all'interno di un elenco. Consideriamo Alberto e Zaccaria, i due inflessibili giudici del merito artistico di una gara di pattinaggio sul ghiaccio. Si tratta di una valutazione soggettiva. Entrambi hanno vinto la medaglia d'oro alle Olimpiadi e devono giudicare il gruppo delle finaliste, che è stato ridotto a cinque concorrenti: Anna, Barbara,

1896

Pearson pubblica i propri studi riguardanti regressione e correlazione

1904

Spearman applica ai propri studi di psicologia la correlazione del rango

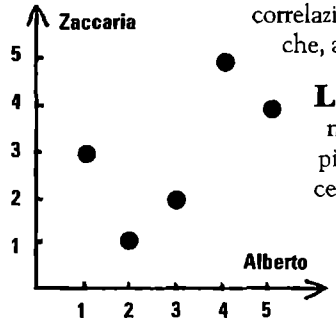
Carla, Diana ed Emma. L'ideale sarebbe che le classifiche di Alberto e Zaccaria fossero perfettamente coincidenti, ma la vita non è così semplice. D'altra parte, non ci aspettiamo neanche che la classifica di Alberto e quella di Zaccaria siano esattamente l'una l'inverso dell'altra. È probabile che le loro classifiche si trovino in una situazione intermedia tra questi due estremi. Alberto ha classificato le concorrenti da 1 a 5 nell'ordine per cui Anna è la migliore, seguita da Emma, Barbara, Carla e Diana in quinta posizione. Per Zaccaria la migliore è stata Emma, seguita da Barbara, Anna, Diana e Carla. Riportiamo le classifiche in una tabella.

Pattinatrice	Classifica di Alberto	Classifica di Zaccaria	Differenza tra le posizioni, d	d^2
Anna	1	3	-2	4
Emma	2	1	1	1
Barbara	3	2	1	1
Carla	4	5	-1	1
Diana	5	4	1	1
$n = 5$			Somma	8

$$1 - \frac{6 \times \Sigma d^2}{n \times (n^2 - 1)}$$

Formula di Spearman

Come si può quantificare il grado di accordo tra i giudici? Lo strumento usato dai matematici per insiemi di dati ordinati è il coefficiente di correlazione di Spearman: nel caso in questione ha un valore di + 0,8, che indica un alto grado di accordo tra Alberto e Zaccaria. Considerando che ogni coppia di posizioni in classifica corrisponda a un punto, si può ottenere una rappresentazione grafica dell'accordo fra i due giudici.



Misura dell'accordo tra due giudici

La formula che fornisce questo coefficiente di correlazione (detto «coefficiente di correlazione del rango») fu sviluppata nel 1904 dallo psicologo Charles Spearman che, al pari di Pearson, fu influenzato dal lavoro di Francis Galton.

La retta di regressione Siamo più bassi o più alti di entrambi i nostri genitori, o abbiamo un'altezza compresa tra le loro? Se fossimo tutti più alti dei nostri genitori e ciò si verificasse per ogni generazione, a un certo punto l'umanità sarebbe composta da individui alti più di tre metri, il che è impossibile. Ma è altrettanto improbabile che l'altezza tenda a diminuire nel tempo, come accadrebbe se fossimo tutti più bassi dei nostri genitori. La verità risiede altrove.

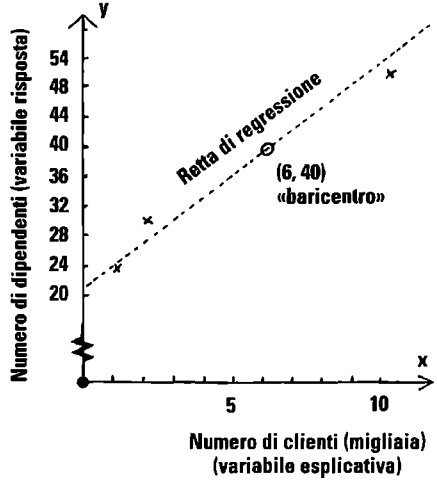
Negli anni '80 dell'Ottocento Francis Galton svolse una ricerca comparativa sulle altezze dei giovani adulti, confrontandole con quelle dei loro genitori: per ogni valore della variabile x che misurava l'altezza dei genitori (o meglio quella di un «genitore intermedio» tra madre e padre, ottenuto combinandone le altezze) registrò le altezze dei figli. Stiamo parlando di uno scienziato pratico, che tirava fuori penna e carta quadrettata per rappresentare i dati. Partendo da 205 genitori intermedi e 928 figli, Galton scoprì che l'altezza media di entrambi gli insiemi

era di 173,4 cm, e chiamò questo valore «mediocrità». Galton osservò che in genere i figli di genitori intermedi molto alti superavano la mediocrità, ma non tanto da superare il proprio genitore intermedio, mentre i figli più bassi erano, in genere, più alti del genitore intermedio, ma senza superare la mediocrità. In altre parole, le altezze dei figli presentavano una regressione verso la mediocrità. È un po' come il numero di reti segnate da Francesco Totti per la Roma: se in una stagione fa registrare un numero di gol eccezionale, è probabile che in quella successiva ne segni qualcuno in meno, restando comunque ben al di sopra della media degli altri giocatori della squadra. Le reti segnate sono regredite verso la media.

La regressione è una tecnica potente e ha un campo d'applicazione molto vasto. Supponiamo che, per un'analisi di mercato, il gruppo di ricerca applicata di una famosa catena di negozi di abbigliamento scelga cinque tra i suoi punti vendita, partendo dalle piccole boutique (con 1000 clienti al mese) per arrivare ai grandi magazzini (con 10 000 clienti al mese). Il gruppo di ricerca analizza il numero di dipendenti di ciascun esercizio, usando la regressione per ottenere una stima del personale necessario a tutti gli altri negozi della catena.

Numero di clienti (migliaia)	1	4	6	9	10
Numero di dipendenti	24	30	46	47	53

Rappresentiamo i dati della tabella in un grafico, scegliendo come coordinata x il numero di clienti (la cosiddetta «variabile esplicativa») e come coordinata y il numero di dipendenti (la «variabile risposta»). È il numero di clienti che spiega quello dei dipendenti, e non viceversa. Come numero medio di clienti dei negozi si riporta 6 (cioè 6000 clienti), mentre il numero medio di dipendenti è 40. La retta di regressione passa sempre dal «baricentro», che in questo caso ha coordinate (6, 40). La retta di regressione (detta anche «retta dei minimi quadrati») è quella che si adatta meglio ai dati e può essere ricavata per mezzo di apposite formule. In questo caso l'equazione della retta è $y = 20{,}8 + 3{,}2\,x$, perciò la pendenza è positiva ed è uguale a 3,2 (la retta «sale» da sinistra a destra); la retta incontra l'asse verticale nel punto 20,8. Il termine $\hat{y}$ è la stima del valore y ottenuta sulla base della retta; per sapere quanti dipendenti dovrebbe avere un negozio frequentato da 5000 clienti al mese bisogna sostituire il valore $x = 5$ nell'equazione di regressione, ottenendo il valore stimato $\hat{y} = 37$. Ciò dimostra quanto la regressione possa essere utile ai fini pratici.



idea chiave

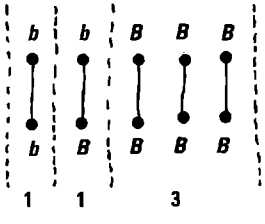
L'interazione fra i dati

37 La genetica

Che cosa ci fa la genetica in un libro di matematica, visto che è una branca della biologia? La risposta è che si tratta di due discipline che si alimentano e si arricchiscono a vicenda. La matematica serve a risolvere i problemi di genetica, ma dal canto suo la genetica ha dato impulso a nuovi settori dell'algebra. Un protagonista indiscusso della genetica, la scienza che studia l'ereditarietà umana, è Gregor Mendel. I caratteri ereditari, come il colore degli occhi e dei capelli, il daltonismo, il mancinismo e il gruppo sanguigno, sono determinati da alcuni fattori (gli alleli) caratteristici dei geni. Mendel fu il primo ad affermare che questi fattori vengono trasmessi alla generazione successiva in modo indipendente.

Come viene trasmesso alla generazione successiva il fattore responsabile del colore degli occhi? Nel modello più semplice si ammette l'esistenza di due fattori (alleli), *b* e *B*.

b è il fattore per gli occhi azzurri
B è il fattore per gli occhi castani



Proporzione 1 : 1 : 3 dei genotipi, bb, bB e BB in una popolazione

In ciascun individuo i fattori si presentano a coppie, dando luogo ai tre possibili genotipi *bb*, *bB* e *BB* (*bB* è equivalente a *Bb*), perciò ogni essere umano possiede uno di questi tre genotipi, che determinano il colore dei suoi occhi. Ad esempio, una popolazione potrebbe essere composta da un quinto di persone con il genotipo *bb*, un altro quinto con il genotipo *bB* e i restanti tre quinti con il genotipo *BB*; in termini percentuali questi genotipi rappresenterebbero, rispettivamente, il 20%, il 20% e il 60% della popolazione. Tutto ciò può essere rappresentato per mezzo di un diagramma.

Il fattore *B*, che codifica il colore castano degli occhi, è il fattore dominante, mentre *b*, il fattore responsabile degli occhi azzurri, è quello recessivo. Una persona il cui genotipo sia puro, cioè *BB*, avrà gli occhi castani, ma saranno castani anche gli occhi di un individuo dal genotipo ibrido *bB*, perché *B* è dominante. Gli unici che possono avere gli occhi azzurri sono gli individui con il genotipo puro *bb*.

linea del tempo

1718 Abraham de Moivre pubblica *La dottrina degli azzardi*

1865 Mendel suggerisce l'esistenza dei geni e le leggi sull'ereditarietà

Agli inizi del XIX secolo si pose una questione scottante per la biologia: gli occhi castani avrebbero finito per prendere il sopravvento, facendo scomparire quelli azzurri? La risposta fu un «no» secco.

La legge di Hardy-Weinberg La spiegazione arrivò con la legge di Hardy-Weinberg, che è una semplice applicazione della matematica alla genetica: questa legge afferma che nella teoria dell'ereditarietà di Mendel il gene dominante non prende completamente il sopravvento e il gene recessivo non scompare.

G.H. Hardy era un matematico inglese che si vantava del fatto che la matematica fosse priva di applicazioni. Sebbene fosse un valente ricercatore nell'ambito della matematica pura, è passato alla storia per il suo unico contributo alla genetica, che si concretizzò sotto forma di una serie di numeri scarabocchiati sul retro di una busta dopo una partita di cricket. La formazione di Weinberg fu molto diversa: medico generico in Germania, si dedicò per tutta la vita alla genetica, e scoprì la legge in questione contemporaneamente a Hardy, intorno al 1908.

La legge di Hardy-Weinberg si applica a un'ampia popolazione in cui gli incroci avvengono casualmente. Ciò significa che non ci sono preferenze: ad esempio, le persone con gli occhi azzurri non preferiscono incrociarsi con altre dello stesso tipo. Dopo l'incrocio, il figlio riceve un allele da ciascun genitore. Ad esempio, l'incrocio tra due genotipi ibridi *bB* può produrre una qualsiasi coppia di fattori fra *bb*, *bB* e *BB*, mentre l'incrocio tra un genotipo *bb* e uno *BB* genera solo ibridi *bB*. Con quale probabilità si trasmette il fattore di questo tipo? Contando il numero di fattori di questo tipo, si trova che ne esistono due per ciascun genotipo *bb* e uno per ciascun genotipo *bB*: in proporzione, ci sono tre fattori *b* su 10 (per una popolazione come la nostra in cui i genotipi sono in rapporto 1 1 : 3). La probabilità di trasmissione di un fattore *b* nel genotipo di un figlio è quindi 3/10, ovvero 0,3, mentre quella del fattore *B* è di 7/10, cioè 0,7. La probabilità di ottenere un genotipo *bb* nella generazione successiva, ad esempio, risulta 0,3 × 0,3 = 0,09. Il quadro completo delle probabilità è riassunto in tabella.

	b		*B*	
b	*bb*	0,3 × 0,3 = 0,09	*bB*	0,3 × 0,7 = 0,21
B	*BB*	0,3 × 0,7 = 0,21	*BB*	0,7 × 0,7 = 0,49

I genotipi ibridi *bB* e *Bb* sono identici, perciò la probabilità che si presentino è data dalla somma 0,21 + 0,21 = 0,42. In termini percentuali, le proporzioni tra i genotipi *bb*, *bB* e *Bb* nella generazione successiva saranno, rispettivamente, del 9%, del 42% e del 49%. Dato che *B* è il fattore dominante, la percentuale della prima gene-

1908 Hardy e Weinberg spiegano perché i geni dominanti non soppiantano quelli recessivi

1918 Fisher riconcilia la teoria di Darwin con quella di Mendel sull'ereditarietà

1953 Viene scoperta la doppia elica del DNA

razione dotata di occhi castani sarà del 42% + 49% = 91%. Solo gli individui con il genotipo *bb* manifesteranno le caratteristiche del fattore *b*, pertanto solo il 9% della popolazione avrà gli occhi azzurri.

La distribuzione iniziale dei genotipi era 20%, 20% e 60%, mentre nella nuova generazione la distribuzione risulta essere 9%, 42% e 49%. Cosa accadrà in seguito? Vediamo cosa succede in un'ulteriore generazione prodotta dagli incroci casuali degli individui di quella appena vista. La proporzione degli alleli *b* è 0,21 + 1/2 × 0,42 = 0,3, mentre la proporzione *B* è 1/2 × 0,42 + 0,49 = 0,7. Queste probabilità di trasmissione sono *identiche* a quelle trovate per i fattori *b* e *B* nel caso precedente. La distribuzione dei genotipi *bb*, *bB* e *Bb* nella seconda generazione sarà quindi uguale a quella della prima generazione, perciò il genotipo *bb* responsabile degli occhi azzurri non scomparirà, ma si manterrà stabile e continuerà ad appartenere al 9% della popolazione. Le proporzioni successive dei genotipi generate da una sequenza di incroci casuali saranno dunque

$$20\%, 20\%, 60\% \rightarrow 9\%, 42\%, 49\% \rightarrow \quad \rightarrow 9\%, 42\%, 49\%$$

Il risultato è in accordo con la legge di Hardy-Weinberg: dopo la prima generazione, le proporzioni dei genotipi si mantengono costanti di generazione in generazione, così come le probabilità di trasmissione.

Il ragionamento di Hardy Per verificare se la legge di Hardy-Weinberg è applicabile a *qualsiasi* popolazione iniziale, e non solo a quella con le proporzioni 20%, 20% e 60% che abbiamo scelto come esempio, la cosa migliore è rifarsi proprio al ragionamento seguito da Hardy nell'articolo che inviò nel 1908 alla rivista americana «Science».

Hardy comincia indicando le distribuzioni iniziali dei genotipi *bb*, *bB* e *BB* rispettivamente con p, $2r$ e q: da queste derivano le probabilità di trasmissione $p + r$ e $r + q$. Nel caso del nostro esempio (20%, 20%, 60%), $p = 0{,}2$, $2r = 0{,}2$ e $q = 0{,}6$. Le probabilità di trasmissione dei fattori *b* e B sono quindi $p + r = 0{,}2 + 0{,}1 = 0{,}3$ e $r + q = 0{,}1 + 0{,}6 = 0{,}7$. Cosa succederebbe se si partisse da una distribuzione iniziale diversa dei genotipi, ad esempio 10%, 60% e 30%? Come funziona la legge di Hardy-Weinberg in questo caso? I valori di p, $2r$ e q sarebbero, rispettivamente, 0,1, 0,6 e 0,3, e le probabilità di trasmissione dei fattori *b* e *B* risulterebbero uguali a $p + r = 0{,}4$ e $r + q = 0{,}6$. Quindi la distribuzione dei genotipi nella generazione successiva sarà uguale al 16%, 48% e 36%, e le proporzioni successive dei genotipi *bb*, *bB* e *BB* a seguito di incroci casuali saranno

$$10\%, 60\%, 30\% \rightarrow 16\%, 48\%, 36\% \rightarrow \quad \rightarrow 16\%, 48\%, 36\%$$

Le proporzioni si stabilizzano nel giro di una generazione, come nel caso precedente, e le probabilità di trasmissione si mantengono costanti sui valori 0,4 e 0,6. In questo caso, il 16% della popolazione avrà gli occhi azzurri e il 48% + 36% = 84% avrà gli occhi castani, dato che nel genotipo *bB* il fattore dominante è *B*.

La legge di Hardy-Weinberg implica che queste proporzioni dei genotipi *bb*, *bB* e *BB* restino costanti di generazione in generazione, a prescindere dalla distribuzione iniziale degli alleli nella popolazione. Il gene dominante *B* non prenderà il sopravvento e le proporzioni dei genotipi resteranno intrinsecamente stabili.

Hardy sottolineò che il modello da lui proposto era solo un'approssimazione, dato che la sua semplicità e la sua eleganza dipendevano da una serie di assunzioni che non corrispondevano esattamente alla realtà. Infatti, questo modello non tiene conto della possibilità di mutazioni genetiche o di cambiamenti all'interno dei geni stessi: se si ammette che le probabilità di trasmissione si mantengano costanti, l'evoluzione diventa inesplicabile. Nel mondo reale esiste una «deriva genetica» che fa sì che le probabilità di trasmissione degli alleli possano cambiare, determinando una variazione delle proporzioni complessive e consentendo l'evoluzione di nuove specie. La legge di Hardy-Weinberg stabilisce un legame intrinseco tra la teoria di Mendel (la «meccanica quantistica» della genetica) e la selezione naturale di Darwin, ma si dovette attendere il genio di R.A. Fisher per conciliare la teoria mendeliana dell'ereditarietà con la teoria della variabilità continua, che ammette l'evoluzione dei caratteri.

L'elemento che mancava alla genetica fino agli anni 1950-1960 era la conoscenza della struttura del materiale genetico. Il progresso decisivo a questo proposito fu compiuto da Francis Crick, James Watson, Maurice Wilkins e Rosalind Franklin. Se il protagonista di questo successo scientifico è l'acido desossiribonucleico, o DNA, ci vuole la matematica per ottenere un modello della famosa doppia elica (una coppia di spirali avvolte intorno a un cilindro), i cui segmenti costituiscono i geni.

Per studiare la genetica è necessario ricorrere alla matematica. Dalla semplice geometria di base delle spirali di DNA e dalla sofisticata legge di Hardy-Weinberg sono stati ricavati modelli in grado di gestire, oltre al colore degli occhi, anche molti altri caratteri fisici, differenze maschio-femmina e incroci non casuali. La genetica ha restituito il favore alla matematica suggerendo nuovi e affascinanti settori di ricerca per l'algebra astratta.

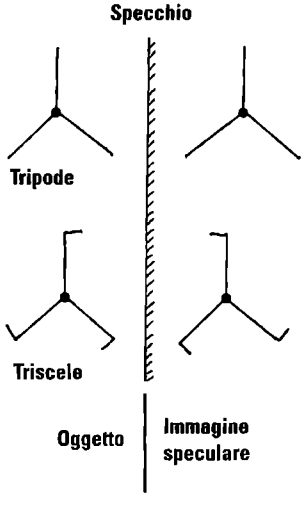
idea chiave

Le incertezze del codice genetico

38 I gruppi

Evariste Galois morì in duello a vent'anni, lasciandosi però alle spalle una quantità di idee sufficiente a tenere occupati i matematici per secoli. Una di esse è la teoria dei gruppi, particolari costruzioni matematiche che servono a rappresentare la simmetria in termini quantitativi. Al di là del fascino artistico, la simmetria è un elemento essenziale per gli scienziati che sognano di elaborare una teoria del tutto: la teoria dei gruppi è il collante che tiene insieme questo «tutto».

La simmetria è dappertutto: si trova nei vasi greci, nei fiocchi di neve, spesso anche negli edifici e in alcune lettere dell'alfabeto. Esistono vari tipi di simmetria, ma i principali sono la simmetria assiale e quella rotazionale. Qui ci limiteremo a considerare la simmetria bidimensionale, studiando solo gli oggetti che possono popolare la superficie piana di questa pagina.



Simmetria assiale È possibile orientare uno specchio in modo da far coincidere l'immagine di un oggetto con il suo riflesso? Le parole OHO e AHA possiedono un asse di simmetria, CON e UNA no: infatti, l'immagine di AHA nello specchio coincide con l'originale, mentre UNA diventa ANU. Un tripode possiede un asse di simmetria, ma il triscele (il tripode con i piedi) no: il triscele che sta davanti allo specchio ruota in senso orario (è destrogiro), mentre la sua immagine speculare ruota in senso antiorario (è levogira).

Simmetria rotazionale Ci si può anche domandare se esista un asse perpendicolare alla pagina intorno a cui si possa far ruotare un oggetto sul piano della pagina di un certo angolo fino a riportarlo nella posizione di partenza. Sia il tripode che il triscele presentano una simmetria rotazionale di questo tipo. Il triscele (un termine che significa «tre gambe») ha una forma particolarmente interessante. La versione destrorsa è il simbolo dell'Isola di Man e

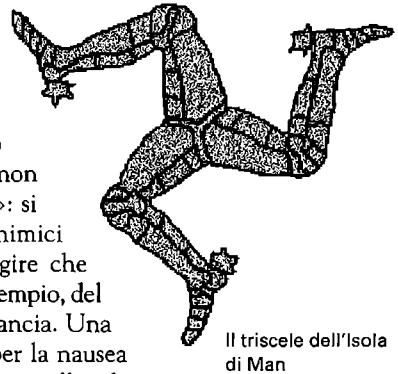
linea del tempo

1832 Galois espone l'idea dei gruppi di permutazioni

1854 Tentativo di Cayley di generalizzare il concetto di gruppo

1872 Felix Klein avvia un programma di classificazione della geometria per mezzo dei gruppi

compare sulla bandiera della Sicilia. Ruotandolo di 120 o 240 gradi, si ottiene una figura coincidente con se stessa: se chiudiamo gli occhi prima di ruotarlo e poi li riapriamo, ci ritroveremo davanti lo stesso triscele.



Il triscele dell'Isola di Man

La caratteristica più curiosa del triscele, però, consiste nel fatto che non è possibile convertire un triscele destrogiro in uno levogiro per mezzo di una rotazione nel piano. Gli oggetti che non coincidono con la loro immagine speculare sono detti «chirali»: si assomigliano molto, ma non sono uguali. Alcuni composti chimici possono avere strutture molecolari tridimensionali sia destrogire che levogire, perciò sono degli esempi di oggetti chirali. È il caso, ad esempio, del limonene: una forma chirale ha un gusto di limone, l'altra di arancia. Una delle due forme del farmaco talidomide è un rimedio efficace per la nausea mattutina delle gestanti, ma l'altra forma produce effetti drammatici sulla salute.

Misurare la simmetria Nel caso del triscele le operazioni di simmetria fondamentali sono le rotazioni (in senso orario) R di 120 gradi e S di 240 gradi. La trasformazione I ruota il triscele di 360 gradi, che è come dire che lo lascia inalterato. Si può creare una tabella riportando le combinazioni di queste rotazioni, seguendo lo stesso procedimento che useremmo per fare una tavola di moltiplicazione.

La tabella è analoga a una tavola di moltiplicazione normale, cioè contenente dei numeri, con la sola differenza che in questo caso si «moltiplicano» dei simboli. Secondo la convenzione più diffusa, la moltiplicazione $R \circ S$ implica prima una rotazione del triscele in senso orario di 240 gradi secondo S, quindi una di 120 gradi secondo R per produrre una rotazione di 360 gradi, che equivale a restare fermi. Ciò può essere espresso in termini simbolici come $R \circ S = I$, il risultato che si legge all'intersezione tra la penultima riga e l'ultima colonna della tabella.

∘	I	R	S
I	I	R	S
R	R	S	I
S	S	I	R

Tavola di Cayley per il gruppo di simmetria triscele

Il gruppo di simmetria del triscele è composto da I, R e S e dalla loro tabella di combinazione. Dato che contiene tre elementi, la sua dimensione (o «ordine») è tre. Queste tabelle si chiamano anche «tavole di Cayley» (dal matematico Arthur Cayley, lontano cugino del pioniere dell'aviazione inglese Sir George Cayley).

Anche il tripode possiede la simmetria rotazionale, ma in più ha quella assiale, perciò appartiene a un gruppo di simmetria più vasto. Indichiamo con U, V e W le riflessioni rispetto ai tre assi di simmetria.

1891

Evgraf Fedorov e Arthur Schönflies classificano in modo indipendente i 230 gruppi cristallografici

1983

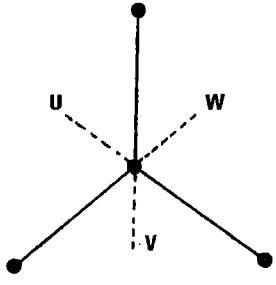
Viene completata la classificazione dei gruppi finiti semplici e viene dimostrato il teorema enorme

	I	R	S	U	V	W
I	I	R	S	U	V	W
R	R	S	I	V	W	U
S	S	I	R	W	U	V
U	U	W	V	I	S	R
V	V	U	W	R	I	S
W	W	V	U	S	R	I

Tavola di Cayley per il gruppo di simmetria del tripode

Il gruppo di simmetria del tripode, più ampio, è di ordine 6, ed è composto dalle trasformazioni I, R, S, U, V e W che si combinano secondo la tavola di moltiplicazione riportata a margine.

Una trasformazione interessante è quella che si ottiene combinando due riflessioni rispetto ad assi distinti, come $U \circ W$ (dove si applica prima la riflessione W, poi la U): il risultato è una rotazione di 120 gradi del tripode, ovvero, in forma simbolica, $U \circ W = R$. Se le riflessioni vengono combinate nell'altro senso si ottiene una rotazione di 240 gradi. Quindi $U \circ W \neq W \circ U$. Si tratta della differenza principale tra la tavola di moltiplicazione di un gruppo e una normale tavola di moltiplicazione tra numeri.



Riflessioni in un tripode

Un gruppo in cui l'ordine di combinazione tra gli elementi non ha alcuna importanza è detto «gruppo abeliano», dal nome del matematico norvegese Niels Abel. Il gruppo di simmetria del tripode è il più piccolo gruppo non abeliano esistente.

Gruppi astratti Nel XX secolo l'algebra è diventata sempre più astratta; per definire un gruppo in algebra astratta sono sufficienti alcune semplici regole, dette «assiomi». Da questo punto di vista il gruppo di simmetria del triangolo non è altro che un esempio di sistema astratto. Esistono sistemi algebrici che sono più semplici di un gruppo e necessitano di un minor numero di assiomi, come sistemi più complessi che ne richiedono di più, ma in ogni caso l'utilità del concetto di gruppo è indiscutibile: si tratta del sistema algebrico più importante. L'ampiezza delle conoscenze ottenute a partire da un numero così ristretto di assiomi è davvero impressionante. Il vantaggio del metodo astratto consiste nel fatto che si possono dedurre i teoremi generali per tutti i gruppi e poi applicarli, quando serve, ai casi specifici.

Una caratteristica della teoria dei gruppi è che si possono individuare gruppi più piccoli all'interno di quelli più grandi; ad esempio, il gruppo di simmetria di ordine 3 del triscele è un sottogruppo del gruppo di simmetria di ordine 6 del tripode. J.L. Lagrange dimostrò una proprietà fondamentale dei sottogruppi: il loro ordine deve essere un divisore esatto dell'ordine del gruppo di appartenenza. Ciò automaticamente comporta che il gruppo di simmetria del tripode non possa avere sottogruppi di ordine quattro o cinque.

Classificare i gruppi La classificazione di tutti i possibili gruppi finiti ha richiesto un programma di ricerca intensivo. In realtà non è necessario elencarli tutti, perché alcuni gruppi si costruiscono a partire dai gruppi semplici, che sono

quelli davvero utili. Il principio seguito è analogo a quello su cui si basa la classificazione degli elementi chimici, in cui non si prendono in esame i composti ottenibili per combinazione degli elementi. Il gruppo di simmetria dei sei elementi del tripode è un «composto» che deriva dalla combinazione del gruppo delle rotazioni (di ordine tre) e di quello delle riflessioni (di ordine due).

Assiomi di gruppo

Un insieme di elementi G con una legge di combinazione $\circ$ è un gruppo se

1. Esiste un elemento 1 tale che $a \circ 1 = 1 \circ a = a$ per tutti gli elementi a del gruppo G (1 è detto «elemento neutro»)
2. Per ogni elemento a di G esiste un elemento $\bar{a}$ di G tale che $\bar{a} \circ a = a \circ \bar{a} = 1$ ($\bar{a}$ è l'«elemento inverso» di a).
3. Dati tre elementi qualsiasi a, b e c di G vale $a \circ (b \circ c) = (a \circ b) \circ c$ (proprietà associativa).

Quasi tutti i gruppi semplici possono essere classificati in classi note. L'ottenimento della classificazione completa, detta «teorema enorme», fu annunciato nel 1983 da Daniel Gorenstein, che vi era giunto raccogliendo e studiando i risultati di trenta anni di ricerche e di pubblicazioni matematiche. Si tratta di un atlante di tutti i gruppi noti. I gruppi semplici possono essere suddivisi in quattro tipi principali, anche se esistono 26 «gruppi sporadici» che non rientrano in nessuna categoria.

I gruppi sporadici sono poco ortodossi e in genere sono di ordine elevato. Cinque tra i più piccoli erano già noti a Emile Mathieu negli anni 1860-1870, ma gran parte dell'attività di ricerca più recente ha avuto luogo tra il 1965 e il 1975. Il gruppo sporadico più piccolo è di ordine $7920 = 2^4 \times 3^2 \times 5 \times 11$, ma alla fine della lista si trova il cosiddetto «gruppo mostro» che è di ordine $2^{46} \times 3^{20} \times 5^9 \times 7^6 \times 11^2 \times 13^3 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times 59 \times 71$; in termini decimali ciò equivale a circa 8×10^{53} o, se si preferisce, a un 8 seguito da 53 zeri: è un numero davvero gigantesco. Si può dimostrare che 20 dei 26 gruppi sporadici sono sottogruppi del «mostro», mentre i sei gruppi che sfuggono a ogni classificazione sono chiamati «i sei paria».

Anche se in matematica le dimostrazioni agili e brevi vanno per la maggiore, la dimostrazione della classificazione dei gruppi finiti occupa qualcosa come 10 000 pagine di argomentazioni simboliche fitte fitte. Il progresso matematico non sempre è il prodotto del lavoro di un formidabile genio solitario.

idea chiave

Il cuore della simmetria

39 Le matrici

Raccontiamo qui la storia dell'«algebra non ordinaria», ovvero di una rivoluzione matematica che si è verificata verso la metà del XIX secolo. I matematici hanno giocato con i blocchi di numeri per secoli, ma l'idea di trattare quei blocchi come un singolo oggetto si è diffusa 150 anni fa grazie a un gruppetto di matematici che seppe comprenderne le potenzialità.

Quando si parla di algebra ordinaria ci si riferisce all'algebra classica, in cui simboli come a, b, c, x e y rappresentano dei numeri. È un concetto che a molti risulta ostico, ma per i matematici ha rappresentato un notevole passo avanti. In confronto, però, l'algebra «non ordinaria» ha provocato una svolta clamorosa: per alcune applicazioni particolarmente sofisticate della matematica, il progresso dall'algebra monodimensionale a quella multidimensionale si è rivelato incredibilmente efficace.

Numeri a più dimensioni Nell'algebra ordinaria a può rappresentare un numero, come 7, e si scrive $a = 7$; nella teoria delle matrici, una *matrice* A è un «numero a più dimensioni», come il blocco

$$A = \begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix}$$

Questa matrice ha tre righe e quattro colonne (o, come si usa dire, è una «matrice 3 per 4»), ma in principio si possono avere matrici con un numero arbitrario di righe e di colonne; ad esempio, esisterà anche una gigantesca matrice «100 per 200» con 100 righe e 200 colonne. Un vantaggio fondamentale dell'algebra delle matrici è che consente di trattare ampi gruppi di numeri, ad esempio un insieme di dati statistici, come una sola entità. Cosa ancora più importante, permette di maneggiare questi blocchi di dati in modo semplice ed efficace. Per sommare o moltiplicare fra loro due insiemi di dati costituiti ciascuno da 1000 numeri non è necessario effettuare 1000 calcoli, ma basta farne uno solo, che consiste nel sommare o nel moltiplicare le due matrici.

linea del tempo

200 a.c.
I matematici cinesi usano blocchi ordinati di numeri

1850
J.J. Sylvester introduce il termine «matrice»

1858
Cayley pubblica *Memoir on the Theory of Matrices*

Un esempio pratico Supponiamo che la matrice A rappresenti la produzione della società AJAX in una data settimana. Questa azienda possiede tre fabbriche in parti diverse del paese, la cui produzione viene misurata in numero di unità (ad esempio, in migliaia di pezzi) dei quattro articoli prodotti. A titolo di esempio, scegliamo come quantità i valori che costituiscono la matrice A della pagina precedente:

	Prodotto 1	Prodotto 2	Prodotto 3	Prodotto 4
Fabbrica 1	7	5	0	1
Fabbrica 2	0	4	3	7
Fabbrica 3	3	2	0	2

Può darsi che la settimana successiva il piano di produzione sia diverso, ma potrà comunque essere espresso sotto forma di una matrice, che indichiamo con B. Ad esempio, B potrebbe essere

$$B = \begin{pmatrix} 9 & 4 & 1 & 0 \\ 0 & 5 & 1 & 8 \\ 4 & 1 & 1 & 0 \end{pmatrix}$$

Qual è la produzione complessiva per le due settimane in questione? L'esperto di teoria delle matrici dice che la produzione complessiva è data dalla matrice $A + B$, che si ottiene sommando i numeri corrispondenti delle due matrici A e B

$$A + B = \begin{pmatrix} 7+9 & 5+4 & 0+1 & 1+0 \\ 0+0 & 4+5 & 3+1 & 7+8 \\ 3+4 & 2+1 & 0+1 & 2+0 \end{pmatrix} = \begin{pmatrix} 16 & 9 & 1 & 1 \\ 0 & 9 & 4 & 15 \\ 7 & 3 & 1 & 2 \end{pmatrix}$$

Fin qui, nessun problema. Purtroppo la moltiplicazione fra matrici è meno immediata. Tornando alla società AJAX, supponiamo che i profitti unitari dei suoi quattro prodotti siano **3**, **9**, **8**, **2**. A questo punto niente ci impedisce di calcolare il profitto complessivo della fabbrica 1, in cui le produzioni dei quattro articoli sono 7, 5, 0, 1. Svolgendo il calcolo, si ottiene $7 \times \mathbf{3} + 5 \times \mathbf{9} + 0 \times \mathbf{8} + 1 \times \mathbf{2} = 68$.

Invece di occuparci di una sola fabbrica possiamo anche calcolare il profitto totale T di *tutte* le fabbriche:

$$T = \begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix} \times \begin{pmatrix} \mathbf{3} \\ \mathbf{9} \\ \mathbf{8} \\ \mathbf{2} \end{pmatrix} = \begin{pmatrix} 7\times\mathbf{3}+5\times\mathbf{9}+0\times\mathbf{8}+1\times\mathbf{2} \\ 0\times\mathbf{3}+4\times\mathbf{9}+3\times\mathbf{8}+7\times\mathbf{2} \\ 3\times\mathbf{3}+2\times\mathbf{9}+0\times\mathbf{8}+2\times\mathbf{2} \end{pmatrix} = \begin{pmatrix} 68 \\ 74 \\ 31 \end{pmatrix}$$

1878

Georg Frobenius dimostra alcuni dei più importanti risultati dell'algebra delle matrici

1925

Heisenberg utilizza la meccanica delle matrici nell'ambito della teoria quantistica

Osservando con attenzione il calcolo effettuato, si nota una caratteristica fondamentale della moltiplicazione fra matrici: il prodotto viene effettuato secondo uno schema «*riga* per *colonna*». Se oltre ai profitti unitari vengono forniti i *volumi* unitari **7, 4, 1, 5** di ciascuna unità dei prodotti, si possono calcolare i profitti e le esigenze di magazzino per le tre fabbriche in un colpo solo, per mezzo di un'unica moltiplicazione fra matrici:

$$\begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix} \times \begin{pmatrix} 3 & 7 \\ 9 & 4 \\ 8 & 1 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 68 & 74 \\ 74 & 54 \\ 31 & 39 \end{pmatrix}$$

Il magazzino complessivo è fornito dalla seconda colonna della matrice risultante, ovvero 74, 54 e 39. La teoria delle matrici è estremamente efficace. Immaginiamo una situazione in cui un'azienda ha centinaia di fabbriche e migliaia di prodotti, con esigenze di magazzino e profitti unitari diversi in settimane differenti: per mezzo dell'algebra delle matrici, i calcoli e la comprensione del problema diventano immediati.

Algebra delle matrici e algebra ordinaria Si possono effettuare vari paralleli fra l'algebra delle matrici e quella ordinaria. La differenza principale si riscontra nella moltiplicazione fra matrici. Se si moltiplica la matrice A per la matrice B e poi si ripete l'operazione invertendo l'ordine dei fattori, si trova che

$$A \times B = \begin{pmatrix} 3 & 5 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 7 & 6 \\ 4 & 8 \end{pmatrix} = \begin{pmatrix} 3\times7+5\times4 & 3\times6+5\times8 \\ 2\times7+1\times4 & 2\times6+1\times8 \end{pmatrix} = \begin{pmatrix} 41 & 58 \\ 18 & 20 \end{pmatrix}$$

$$B \times A = \begin{pmatrix} 7 & 6 \\ 4 & 8 \end{pmatrix} \times \begin{pmatrix} 3 & 5 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 7\times3+6\times2 & 7\times5+6\times1 \\ 4\times3+8\times2 & 4\times5+8\times1 \end{pmatrix} = \begin{pmatrix} 33 & 41 \\ 28 & 28 \end{pmatrix}$$

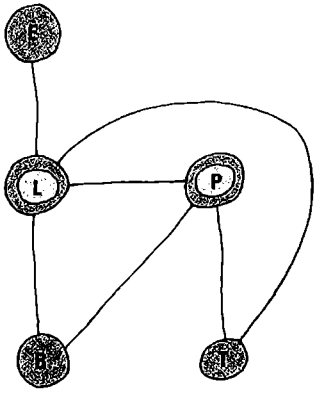
Nell'algebra delle matrici $A \times B$ e $B \times A$ possono essere diversi, al contrario di ciò che accade nell'algebra ordinaria, dove l'ordine dei fattori di una moltiplicazione è ininfluente ai fini del risultato.

Un'altra differenza riguarda l'inversione. Nell'algebra ordinaria, l'inverso di un numero può essere calcolato facilmente: se $a = 7$ il suo inverso è $1/7$, perché vale la proprietà $1/7 \times 7 = 1$. L'inverso viene anche scritto nella forma $a^{-1} = 1/7$, da cui $a^{-1} \times a = 1$.

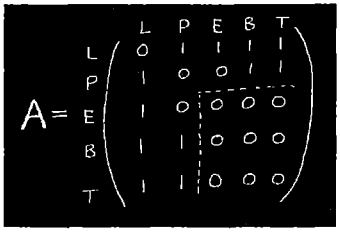
Scegliendo come esempio la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}$, si può verificare che $A^{-1} = \begin{pmatrix} 7 & -2 \\ -3 & 1 \end{pmatrix}$

perché $A^{-1} \times A = \begin{pmatrix} 7 & -2 \\ -3 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

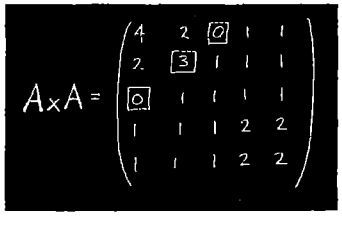
dove $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ è detta «matrice identità» ed è l'equivalente matriciale del numero 1 dell'algebra ordinaria; molte matrici non ammettono un inverso, mentre nell'algebra ordinaria solo 0 ne è privo.



Progetti di viaggio Un'altra potenziale applicazione delle matrici riguarda l'analisi di una rete di voli per conto di una compagnia aerea. Una rete di questo tipo comprenderà sia aeroporti principali che secondari e nella realtà sarà costituita da centinaia di destinazioni, ma qui ci limiteremo a considerare un esempio ristretto: gli aeroporti principali di Londra (L) e Parigi (P), e gli aeroporti secondari di Edimburgo (E), Bordeaux (B) e Tolosa (T), indicando nella rete solo i voli *diretti*. Perché una rete di questo tipo possa essere analizzata da un computer, deve prima essere codificata sotto forma di una matrice. Se due aeroporti sono collegati da un volo diretto, si scrive un 1 nel punto di intersezione fra la riga e la colonna corrispondenti agli aeroporti in questione (ad esempio, fra Londra ed Edimburgo). Chiamiamo A la «matrice di connessione» che descrive la rete che stiamo esaminando.



La sottomatrice di A che compare in basso a destra (delimitata dalle linee tratteggiate) rivela che fra i tre aeroporti più piccoli non esistono collegamenti diretti, mentre il prodotto $A \times A = A^2$ fornisce il numero di viaggi possibili fra due aeroporti *con uno e un solo scalo*. Così, ad esempio, esistono tre possibili viaggi di andata e ritorno a partire da Parigi che passano da altre città, ma nessun viaggio da Londra a Edimburgo implica scali. Il numero di percorsi che sono diretti oppure implicano uno scalo è dato dagli elementi della matrice $A + A^2$. Si tratta di un ulteriore esempio della capacità delle matrici di rappresentare la sostanza di un'enorme quantità di dati in un singolo calcolo di carattere generale.



Quando un gruppetto di matematici creò la teoria delle matrici, negli anni 1850-1860, l'obiettivo era quello di risolvere problemi di matematica pura. Dal punto di vista applicativo, la teoria delle matrici era sostanzialmente «una soluzione in cerca di un problema». Come accade spesso, nacquero effettivamente dei «problemi» che avevano bisogno di quella teoria emergente. Una delle prime applicazioni si presentò negli anni '20 del Novecento, quando Werner Heisenberg studiò la «meccanica delle matrici», una formulazione della teoria quantistica. Un'altra pioniera fu Olga Taussky-Todd, che per un periodo si occupò della progettazione di aeroplani utilizzando l'algebra delle matrici. A chi le domandò come avesse scoperto quell'argomento, rispose che era successo l'inverso: erano state le matrici a trovare lei. Così vanno le cose in matematica.

idea chiave

Combinare blocchi di numeri

40 I codici

Cos'hanno in comune Giulio Cesare e la trasmissione dei moderni segnali digitali? Per dirla in due parole: i codici e la codifica. Se si vuole inviare un segnale digitale a un computer o a un apparecchio televisivo, è necessario codificare immagini e parole in un flusso di 0 e di 1, cioè in un codice binario, perché è l'unico linguaggio che quei dispositivi sono in grado di comprendere. Cesare si serviva di codici per comunicare con i suoi generali e rendeva incomprensibili i propri messaggi spostandone le lettere secondo una chiave nota solo a lui e ai destinatari della comunicazione.

Se l'accuratezza era essenziale per Cesare, è necessaria anche per garantire che i segnali digitali vengano trasmessi in modo efficace. Inoltre Cesare voleva tenere segreti i suoi codici, esattamente come le compagnie televisive che trasmettono via cavo e via satellite vogliono che solo gli abbonati siano in grado di decifrare i loro segnali.

Consideriamo in primo luogo la questione dell'accuratezza. L'errore umano o il «rumore sulla linea» sono sempre in agguato e bisogna farci i conti: il ragionamento matematico permette di costruire dei sistemi di codificazione in grado di rivelare automaticamente gli errori e di correggerli.

Rivelazione e correzione degli errori Uno dei primi sistemi di codificazione binaria mai ideati è il codice Morse, che utilizza due simboli, punti • e linee −. L'inventore americano Samuel F.B. Morse inviò il primo messaggio interurbano per mezzo del suo codice nel 1844, da Washington a Baltimora. Si trattava di un codice concepito per il telegrafo elettrico di metà Ottocento, perciò non si era pensato a ottimizzarne l'efficienza. Nell'alfabeto Morse la lettera A è codificata come • −, B come − • • •, C come − • − • e le altre lettere come sequenze diverse di punti e di linee. Se un telegrafista vuole inviare «CAB», spedirà la stringa − • − • / • − / − • • •. Senza niente togliere ai suoi meriti, il codice Morse non se la cava affatto bene nel rivelare gli errori, figuriamoci poi nel correggerli. Se l'operatore Morse volesse inviare «CAB» ma scrivesse per sbaglio un punto al

linea del tempo

55 a.C.
Nella campagna militare in Britannia, Giulio Cesare comunica con i suoi generali per mezzo di messaggi in codice

circa 1750
Il teorema di Eulero pone le basi per la crittografia a chiave pubblica

posto di una linea in C, dimenticasse la linea in A e il rumore sulla linea telegrafica sostituisse un punto con una linea in B, il ricevitore otterrebbe la sequenza • • – • / • / – – • • e la interpreterebbe come «FEZ», senza trovarci niente di strano.

A uno stadio più primitivo, possiamo pensare a un sistema di codificazione costituito solo da 0 e 1, dove 0 rappresenta una parola e 1 un'altra. Supponiamo che il comandante di un esercito debba trasmettere alle sue truppe un messaggio che può essere «invadete» o «non invadete»; l'istruzione «invadete» è codificata come 1, «non invadete» come 0. Se un 1 o uno 0 venissero trasmessi in modo errato, il ricevitore non se ne accorgerebbe e verrebbe impartito l'ordine sbagliato, con conseguenze disastrose.

Per migliorare le cose si possono usare parole di codice di lunghezza due. Se si codifica l'istruzione «invadete» come 11 e «non invadete» come 00, va già meglio: un errore in una cifra trasformerebbe il messaggio in 01 o 10, ma poiché solo 11 e 00 sono parole di codice ammesse il ricevitore si accorgerebbe subito che c'è stato un errore. Il vantaggio di questo metodo è che un eventuale errore verrebbe sicuramente rivelato. Però non verrebbe corretto, perché chi riceve 01 non può sapere se il messaggio originale era 00 o 11.

La soluzione per migliorare il sistema consiste nel combinare un meccanismo logico più sofisticato con parole di codice più lunghe. Se si codifica l'istruzione «invadete» come 111 e «non invadete» come 000, un errore in una cifra verrebbe sicuramente rivelato anche in questo caso; se poi fosse noto che il messaggio non può contenere più di un errore (e si tratta di un'assunzione ragionevole, perché la probabilità di fare due errori nella stessa parola è bassa) il ricevitore sarebbe in grado di correggerlo. Ad esempio, se il messaggio ricevuto fosse 110, vorrebbe dire che quello inviato era 111. In base alle regole enunciate non potrebbe essere 000, perché questa parola di codice differisce da 110 per due errori. Le uniche due parole di codice ammesse, 000 e 111, sono abbastanza diverse da consentire di rivelare e correggere gli errori.

Lo stesso principio viene sfruttato nella correzione automatica di un programma di videoscrittura. Se digitiamo «animule», il programma individua l'errore e lo corregge scegliendo la parola più simile, «animale». Va notato, però, che non sempre è possibile apportare una correzione; se la parola digitata è «cira», ad esempio, non esiste un'unica parola più simile delle altre: in termini di errori singoli, «cera», «cara», «cura» sono tutte equidistanti da «cira».

Le parole dei moderni codici binari sono blocchi di zero e di uno. Come abbiamo visto, se le parole di codice ammesse sono sufficientemente diverse diventa possibile

1844 Morse trasmette il primo messaggio utilizzando l'omonimo codice

1920-1930 Viene sviluppata la macchina Enigma

1950 Richard Hamming pubblica un saggio fondamentale sulla rivelazione e la correzione degli errori nei codici

1970-1980 Viene sviluppata la crittografia a chiave pubblica

sia rivelare che correggere gli errori. Le parole del codice Morse sono troppo simili, ma nei sistemi di codificazione moderni utilizzati per la trasmissione di dati satellitari è sempre possibile inserire l'opzione di correzione automatica. Parole di codice lunghe danno ottimi risultati in termini di correzione degli errori ma richiedono più tempo per essere trasmesse, perciò di volta in volta bisogna trovare un compromesso fra lunghezza e velocità di trasmissione. Le spedizioni della NASA nello spazio profondo hanno fatto uso di codici in grado di correggere tre errori che si sono dimostrati soddisfacenti anche per quanto riguarda l'abbattimento del rumore sulla linea.

Messaggi segreti Per rendere impenetrabili i suoi messaggi, Giulio Cesare ne spostava le lettere secondo una chiave nota solo a lui e ai suoi generali; se la chiave fosse caduta nelle mani sbagliate, però, i messaggi sarebbero stati decifrati dai suoi nemici. In epoca medievale, dalla cella della prigione in cui era rinchiusa, la regina Maria di Scozia inviò dei messaggi segreti in codice nel tentativo di detronizzare la cugina, la regina Elisabetta, ma i messaggi vennero intercettati. I suoi codici si basavano su una serie di sostituzioni e quindi erano più sofisticati del metodo adottato dai romani che consisteva nello spostare tutte le lettere del messaggio in base a una chiave; le sostituzioni erano però facilmente individuabili analizzando la frequenza delle lettere e dei simboli impiegati. Anche il codice Enigma, concepito dai tedeschi durante la Seconda guerra mondiale, fu decifrato quando ne fu scoperta la chiave; in questo caso si trattava di un'impresa difficilissima, ma il codice era intrinsecamente vulnerabile, perché la chiave era inclusa nel messaggio.

Negli anni '70 del Novecento si è verificato uno sviluppo sorprendente nelle tecniche di criptaggio dei messaggi. Si scoprì infatti che, contrariamente a quanto si era sempre creduto, era possibile divulgare la chiave segreta senza pregiudicare la segretezza del messaggio: si tratta della cosiddetta «crittografia a chiave pubblica». Il metodo impiegato si basa su un teorema scoperto due secoli fa in una branca della matematica di cui in genere viene esaltata la totale inutilità.

Crittografia a chiave pubblica Il signor John Sender, un agente segreto noto nell'ambiente dello spionaggio come «J», è appena arrivato in città e vuole comunicarlo al suo contatto, il dottor Rodney Receiver, per mezzo di un messaggio segreto. A questo punto fa una cosa strana: si procura un elenco telefonico e cerca il nome del dottor Receiver. Nell'elenco trova due numeri associati al nome di Rodney Receiver: uno lungo, 247, e uno più corto, 5. Queste informazioni sono accessibili a tutti e sono sufficienti a John Sender per criptare il suo messaggio, che per semplicità è il suo «nome in codice», una J. Questa lettera compare nella 74ª posizione di un elenco di parole, anch'esso di dominio pubblico.

Sender codifica 74 calcolando 74^5 (modulo 247), cioè ricavando il resto della divisione di 74^5 per 247. È possibile calcolare 74^5 con una normale calcolatrice, ma il calcolo deve essere preciso:

$$74^5 = 74 \times 74 \times 74 \times 74 \times 74 = 2\,219\,006\,624$$

e

$$2\,219\,006\,624 = 8\,983\,832 \times 247 + 120$$

Dividendo questo numero enorme per 247 si ottiene un resto di 120. In questo modo Sender ha prodotto il messaggio criptato (120), che trasmette a Receiver. Poiché i numeri 247 e 5 erano di dominio pubblico, chiunque avrebbe potuto criptare un messaggio. Però non tutti sarebbero stati in grado di decrittarlo. Infatti, il dottor Receiver ha un altro asso nella manica: ha ottenuto il suo numero personale 247 moltiplicando due numeri primi, $p = 13$ e $q = 19$, ma è l'unico a saperlo.

È arrivato il momento di rispolverare il vecchio teorema di Eulero. Il dottor Receiver utilizza le informazioni $p = 13$ e $q = 19$ per calcolare un valore di a tale che: $5 \times a \equiv 1 (\mathrm{mod}(p-1)(q-1))$, dove il simbolo $\equiv$ significa «uguale» in aritmetica modulare. Per quale valore di a il prodotto $5 \times a$ diviso per $12 \times 18 = 216$ dà resto 1? Sorvolando sui passaggi del calcolo, diciamo che $a = 173$.

Essendo l'unico a conoscere i numeri primi p e q, il dottor Receiver è anche l'unico in grado di calcolare il numero 173, che utilizza per ricavare il resto della divisione fra il gigantesco numero 120^{173} e 247. Questo calcolo non è alla portata di una calcolatrice, ma può essere facilmente effettuato con un computer: il risultato è 74, come sapeva Eulero già duecento anni fa. Receiver cerca nell'elenco la parola numero 74 e capisce che J è tornato in città.

Si può obiettare che un hacker potrebbe scoprire che $247 = 13 \times 19$, decifrando il codice. È vero. Ma il principio della cifratura resterebbe identico anche se il dottor Receiver utilizzasse un altro numero invece di 247, perciò potrebbe scegliere due numeri primi molto grandi e moltiplicarli in modo da ottenere un numero molto maggiore di 247.

Trovare i due fattori primi che scompongono un numero molto grande è virtualmente impossibile; ad esempio, quali sono i fattori primi di 24 812 789 922 307? Del resto, si potrebbe scegliere un numero ancora più grande. Il sistema della chiave pubblica è intrinsecamente sicuro; se poi, utilizzando una combinazione di computer potentissimi, si riuscisse a scomporre in fattori primi un determinato numero di codificazione, il dottor Receiver non dovrebbe fare altro che sceglierne uno ancora più grande. In definitiva, è molto più facile per il dottor Receiver «mischiare sabbia nera e bianca» di quanto non lo sia per un hacker separarle.

idea chiave

41 Il calcolo combinatorio

Il calcolo combinatorio è una branca della matematica che richiede tecniche avanzate di calcolo, ma non si tratta di sommare a mente lunghe colonne di numeri. Ciò che importa non è tanto la quantità degli oggetti, quanto il numero delle loro possibili combinazioni. Di solito le questioni sono poste in modo semplice, senza la pesante sovrastruttura della teoria matematica: basta rimboccarsi le maniche, anche se non si dispone di grandi conoscenze preliminari. È questo il bello dei problemi di calcolo combinatorio. Attenzione alla salute, però: danno dipendenza e possono turbare il sonno.

La filastrocca di Camogli I bambini possono intraprendere il calcolo combinatorio fin dalla più tenera età. Infatti c'è una filastrocca tradizionale che pone una domanda di natura combinatoria:

> Per una strada che mena a Camogli
> vidi un uomo che aveva sette mogli
> e ogni moglie aveva sette sacche
> e in ogni sacco aveva sette gatte
> e ogni gatta sette gattini
> fra gatti e gatte e sacche e mogli
> in quanti andavano, dite, a Camogli?

L'ultimo verso in realtà è un tranello (la risposta è uno). Ma ogni domanda può essere rovesciata: in quanti erano a *venire* da Camogli?

L'interpretazione è importante. Chi ci dice che l'uomo e le sue sette mogli stessero *provenendo* da Camogli? Al momento dell'incontro le mogli erano col marito o no?

linea del tempo

circa 1800 a.C.
In Egitto viene scritto il papiro di Rhind

circa 1100
Bhaskara si occupa di permutazioni e combinazioni

Il primo requisito di un problema combinatorio è che il suo enunciato sia chiaro e comprensibile.

Daremo per scontato che il gruppo provenisse dalla cittadina della Riviera Ligure dalla stessa strada, e che gattini, gatte, sacche e mogli fossero tutti presenti. In quanti erano a venire da Camogli? La soluzione può essere ricavata dalla seguente tabella.

Marito	1	1
Mogli	7	7
Sacche	7×7	49
Gatte	$7 \times 7 \times 7$	343
Gattini	$7 \times 7 \times 7 \times 7$	2401
Totale		**2801**

Alexander Rhind, un antiquario scozzese in viaggio a Luxor, nel 1858 trovò e acquistò un papiro lungo 5 metri tutto pieno di formule matematiche risalenti al 1800 a.C. Qualche anno dopo il papiro fu comprato dal British Museum, che ne tradusse i geroglifici. Il problema 79 del Papiro di Rhind riguarda case, gatti, topi e grano ed è analogo a quello di gattini, gatte, sacche e mogli di Camogli: entrambi sono legati alle potenze del 7 e allo stesso tipo di analisi. A quanto pare, il calcolo combinatorio ha una lunga tradizione.

Numeri fattoriali Il problema delle persone in coda ci porta a far conoscenza con la prima arma dell'arsenale del calcolo combinatorio: il numero fattoriale. Ammettiamo che Aldo, Bruno, Carla, Daniele ed Emma si dispongano in coda nell'ordine

E C A B D

con Emma in testa, seguita da Carla, Aldo, Bruno e Daniele a chiudere la fila. Se le persone si scambiano di posto, si formano code diverse: quante code diverse si possono ottenere?

In questo problema tutta l'arte del conteggio dipende dalla *scelta*. Esistono 5 opzioni per il primo della fila e, stabilito questo, altre 4 per il secondo posto, e così via. Per l'ultima posizione la scelta è obbligata, perché sarà occupata dall'unica persona rimasta. Pertanto si possono avere $5 \times 4 \times 3 \times 2 \times 1 = 120$ code diverse. Se si parte da 6 persone il numero di code arriva a $6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$ e con 7 persone sale a $7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$.

1850
Kirkman enuncia il problema delle 15 scolarette

1930
Frank Ramsey studia il calcolo combinatorio

1971
Ray-Chaudhuri e Wilson dimostrano l'esistenza di sistemi di Kirkman generali

Un numero ottenuto dalla moltiplicazione di numeri interi consecutivi è detto «numero fattoriale». I numeri fattoriali sono talmente frequenti in matematica da meritare una notazione apposita: 5! (leggi «5 fattoriale») sta per $5 \times 4 \times 3 \times 2 \times 1$. Basta dare un'occhiata ai primi fattoriali (ponendo $0! = 1$ per convenzione) per rendersi conto che configurazioni piuttosto «piccole» generano numeri fattoriali «grandi». Anche se n è piccolo, $n!$ può essere enorme.

numeri	fattoriale
0	1
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40320
9	362880

Il ragionamento è più o meno lo stesso se lo scopo resta quello di mettere in coda 5 persone, scegliendole però all'interno di un gruppo di 8 membri: **A, B, C, D, E, F, G e H**. Esistono 8 possibilità per la posizione di testa, 7 per il secondo posto, e così via. Questa volta per l'ultima posizione ci sono 4 opzioni, quindi il numero di code possibili è

$$8 \times 7 \times 6 \times 5 \times 4 = 6720$$

che può essere scritto a sua volta in notazione fattoriale, perché

$$8 \times 7 \times 6 \times 5 \times 4 = 8 \times 7 \times 6 \times 5 \times 4 \times \frac{3 \times 2 \times 1}{3 \times 2 \times 1} = \frac{8!}{3!}$$

Combinazioni In una coda *l'ordine* è importante. Le due code

C E B A D D A C E B

sono diverse, per quanto formate dalle stesse lettere. Come già sappiamo, il numero di code ottenibili con queste lettere è 5! Per calcolare il numero di code che è possibile formare scegliendo 5 persone da un gruppo di 8 *senza prendere in considerazione l'ordine* bisogna dividere $8 \times 7 \times 6 \times 5 \times 4 = 6720$ per 5!. Quindi il numero di modi in cui si possono scegliere 5 persone da un gruppo di 8 risulta essere

$$\frac{8 \times 7 \times 6 \times 5 \times 4}{5 \times 4 \times 3 \times 2 \times 1} = 56.$$

Il numero di modi viene scritto anche nella forma $C_{8,5}$, dove C sta per combinazione e corrisponde a

$$C_{8,5} = \frac{8!}{3!5!} = 56.$$

Quando si gioca al Superenalotto, si devono scegliere 6 numeri su 90. Quante sono le combinazioni possibili?

$$C_{90,6} = \frac{90!}{84!6!} = \frac{90 \times 89 \times 88 \times 87 \times 86 \times 85}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 622\,614\,630$$

La probabilità di vincere il montepremi giocando una sola sestina è circa 1 su 600 milioni.

Il problema di Kirkman L'ambito del calcolo combinatorio è vasto e studiato da tempo, ma si è sviluppato rapidamente negli ultimi quarant'anni per l'importanza che riveste in informatica. I problemi riguardanti la teoria dei grafi, i quadrati latini e simili possono essere considerati una parte del calcolo combinatorio moderno.

La vera essenza del calcolo combinatorio fu colta da un asso della materia, il Reverendo Thomas Kirkman, in un'epoca in cui l'interesse principale dell'argomento era connesso soprattutto alla matematica ludica. Kirkman fornì molti contributi originali alla geometria discreta, alla teoria dei gruppi e al calcolo combinatorio, pur non ricoprendo alcun ruolo accademico. L'enigma che consolidò la sua reputazione di matematico ludico è quello per cui sarà sempre ricordato. Infatti nel 1850 Kirkman presentò «il problema delle 15 scolarette» che vanno in chiesa in 5 file di 3 tutti i giorni della settimana. Chi ne ha abbastanza del Sudoku può provare a risolverlo. Per evitare che due ragazze procedano insieme per più di una volta, bisogna organizzare un programma giornaliero. Usando di proposito maiuscole e minuscole, le ragazze saranno: anna, benedetta, costanza, diana, emma, francesca, giovanna, Adele, Bianca, Carla, Daniela, Eva, Flavia, Gloria e Vittoria, indicate rispettivamente con **a, b, c, d, e, f, g, A, B, C, D, E, F, G** e **V**.

Il problema di Kirkman ammette sette soluzioni distinte e quella riportata è ciclica, cioè viene generata procedendo in circolo. È qui che si capisce l'importanza delle etichette assegnate alle ragazze.

Lunedì			Martedì			Mercoledì			Giovedì			Venerdì			Sabato			Domenica		
a	A	V	b	B	V	c	C	V	d	D	V	e	E	V	f	F	V	g	G	V
b	E	D	c	F	E	d	G	F	e	A	G	f	B	A	g	C	B	a	D	C
c	B	G	d	C	A	e	D	B	f	E	C	g	F	D	a	G	E	b	A	F
d	f	g	e	g	a	f	a	b	g	b	c	a	c	d	b	d	e	c	e	f
e	F	C	f	G	D	g	A	E	a	B	F	b	C	G	c	D	A	d	E	B

È una soluzione ciclica, perché ogni giorno da **a** si passa a **b**, da **b** a **c**, per poi tornare da **g** ad **a**, e lo stesso vale per le lettere maiuscole, da **A** a **B**, da **B** a **C** e così via, mentre Vittoria rimane al suo posto.

Ma il vero motivo per cui si è scelta questa notazione consiste nel fatto che le righe corrispondono alle rette nel piano di Fano (vedi capitolo 28). Il problema di Kirkman non è solo un gioco da salotto, ma fa parte a pieno titolo della matematica più recente.

idea chiave

Quante combinazioni ci sono?

42 I quadrati magici

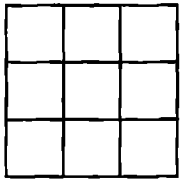
«I matematici sono come i pittori o i poeti», scrisse G.H. Hardy, «inventano forme». I quadrati magici presentano forme strane anche per la matematica: si collocano al confine tra la pesante simbologia della matematica e le affascinanti sequenze amate dagli enigmisti.

Un quadrato magico è una griglia quadrata in cui vengono collocati numeri interi distinti, in modo che ogni colonna orizzontale, verticale e *anche* diagonale abbia la stessa somma.

I quadrati con una sola riga e una sola colonna in linea di principio sono magici, ma sono anche tremendamente noiosi, perciò non li tratteremo. Non esistono quadrati magici con due righe e due colonne, ma se esistessero avrebbero la struttura illustrata a margine. Dato che la somma delle righe e delle colonne deve essere uguale, deve valere l'uguaglianza $a + b = a + c$. Ciò implica che $b = c$, contraddicendo la condizione che ciascun elemento debba essere distinto.

a	b
c	d

Il quadrato di Lo Shu Visto che non esistono quadrati 2 × 2, ci dedicheremo a quelli di dimensioni 3 × 3 e tenteremo di costruirne uno usando una griglia. Partiamo da un quadrato magico *normale*, dove la griglia viene riempita con i numeri consecutivi 1, 2, 3, 4, 5, 6, 7, 8 e 9.



Con una griglia così piccola si può procedere per tentativi, ma per semplificarci il compito è bene fare alcune deduzioni preliminari. Ad esempio, sommando tutti i numeri della tabella si ottiene

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 45$$

che deve corrispondere alla somma dei totali delle tre righe. Ciò dimostra che la somma di ciascuna riga, colonna e diagonale deve essere uguale a 15. Concentriamoci ora sulla casella centrale, che chiameremo c. Questa cella è contenuta nelle due diagonali, nonché nella riga e nella colonna centrali. Sommando i numeri che costituiscono queste quattro linee si ottiene $15 + 15 + 15 + 15 = 60$, che deve equi-

linea del tempo

circa 2800 a.C. Nasce la leggenda del quadrato di Lo Shu

circa 1690 De la Loubère introduce il metodo siamese per la costruzione di quadrati magici

valere alla somma di tutti i numeri più tre volte c. Dall'equazione $3c + 45 = 60$ si deduce che *c deve* essere 5; inoltre si può dimostrare che le celle d'angolo non possono contenere un 1. Ora che con queste indicazioni siamo arrivati a buon punto, è possibile procedere per tentativi. Provateci!

Sarebbe bello se esistesse un metodo completamente automatico per costruire i quadrati magici. Simon de la Loubère, ambasciatore francese presso il Regno del Siam nel tardo XVII secolo, ne scoprì uno. Loubère era interessato alla matematica cinese ed elaborò un metodo per costruire quadrati magici con un numero dispari di righe e colonne. Si parte scrivendo un 1 al centro della prima riga e poi «si segue la diagonale ascendente a destra, rientrando dall'altro lato del quadrato, se necessario» per inserire il 2 e i numeri successivi. Se la casella è già occupata, si passa al numero successivo inserendolo sotto quello in questione.

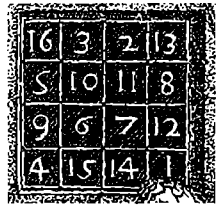
8	1	6
3	5	7
4	9	2

Una soluzione del quadrato 3 × 3

È interessante notare che questo quadrato magico normale è sostanzialmente l'unico ottenibile con tre righe e tre colonne. Tutti gli altri quadrati magici di dimensioni 3 × 3 possono essere ottenuti per rotazione dei numeri intorno al centro o per riflessione rispetto alla colonna o alla riga centrale. Si tratta del «quadrato di Lo Shu», che era noto in Cina già nel 3000 a.C. La leggenda vuole che sia apparso per la prima volta sul dorso di una tartaruga che emergeva dalle acque del fiume Lo; gli abitanti del luogo lo considerarono il segno che gli Dei non li avrebbero liberati dalla peste se non avessero aumentato le offerte votive.

Se esiste un solo quadrato 3 × 3, quanti quadrati magici 4 × 4 ci saranno? La risposta è sbalorditiva: ne esistono 880 (reggetevi forte: i quadrati magici di ordine cinque sono 2 202 441 792). Non è nota alcuna formula generale che consenta di prevedere il numero dei quadrati magici di ordine *n*.

I quadrati di Dürer e Franklin Se il quadrato magico di Lo Shu è famoso in virtù della sua antichità e della sua unicità, esiste un quadrato magico 4 × 4 che è diventato emblematico per il fatto di essere associato a un celebre artista. Inoltre possiede molte più proprietà dei quadrati magici ordinari che concorrono a formare le sue 880 varianti. Si tratta del quadrato 4 × 4 che compare nell'incisione *Melancolia I* di Albrecht Dürer, realizzata nel 1514.



Nel quadrato di Dürer, la somma di ciascuna riga è 34, come quella di ciascuna colonna e diagonale, nonché dei quadrati più piccoli 2 × 2 da cui è costituito il quadrato complessivo. Dürer riuscì addirittura a siglare il suo capolavoro con la data, che compare al centro della riga inferiore.

1693
Bernard Frénicle de Bessy elenca tutti gli 880 quadrati magici 4 × 4 possibili

1770
Eulero realizza un quadrato di quadrati perfetti

1986
Sallows crea un quadrato basato sul numero di lettere

52	61	4	13	20	29	36	45
14	3	62	51	46	35	30	19
53	60	5	12	21	28	37	44
11	6	59	54	43	38	27	22
55	58	7	10	23	26	39	42
9	8	57	56	41	40	25	24
50	63	2	15	18	31	34	47
16	1	64	49	48	33	32	17

Lo scienziato e diplomatico americano Benjamin Franklin pensava che la costruzione di quadrati magici fosse uno strumento utile per affinare la mente. Ne era un grande esperto e a tutt'oggi i matematici sanno ben poco dei metodi che utilizzava: i quadrati magici di grandi dimensioni non si ottengono certo per caso. Franklin ammise di aver sprecato molto tempo della sua gioventù nel costruirli, sebbene da ragazzo non avesse affatto il pallino dell'aritmetica. Uno dei quadrati ottenuti dal giovane Franklin è riportato a margine.

Questo quadrato magico normale presenta simmetrie di tutti tipi. La somma di ciascuna riga, colonna e diagonale è 260, così come quella delle «righe spezzate» analoghe a quella evidenziata; inoltre il quadrato nasconde molti altri segreti, come il fatto che la somma del quadratino centrale 2 × 2 e delle quattro celle d'angolo è uguale a 260. Osservandolo attentamente, si nota una proprietà interessante di ciascun quadrato 2 × 2.

Quadrati di quadrati Si possono costruire anche quadrati magici le cui celle sono occupate da numeri diversi elevati al quadrato. Il problema della loro realizzazione fu posto dal matematico Edouard Lucas nel 1876. A tutt'oggi non è stato trovato nessun quadrato di quadrati di dimensioni 3 × 3, anche se in un caso ci siamo andati vicino.

127^2	46^2	58^2
2^2	113^2	94^2
74^2	82^2	97^2

Ciascuna riga, ciascuna colonna e *una* diagonale del quadrato a fianco danno la somma magica 21 609, ma la somma dell'altra diagonale non torna, perché $127^2 + 113^2 + 97^2 = 38\,307$. Se volessimo costruircene uno da soli, dovremmo tener conto di un fatto appurato: il valore centrale deve essere maggiore di $2{,}5 \times 10^{25}$, perciò è inutile andare in cerca di numeri piccoli! Qui siamo nel campo della matematica «seria» collegata allo studio delle curve ellittiche, che sono state utilizzate per dimostrare l'ultimo teorema di Fermat. È stato dimostrato che non esistono quadrati magici 3 × 3 ottenibili a partire da cubi o quarte potenze.

La ricerca di quadrati di quadrati è stata fruttuosa per quadrati più grandi: esistono quadrati magici di quadrati perfetti di ordine 4 × 4 e 5 × 5. Nel 1770 Eulero ne produsse un esempio senza rivelarne il metodo di costruzione e da allora ne sono state trovate intere famiglie in relazione allo studio dell'algebra dei quaternioni, i numeri immaginari a quattro dimensioni.

Quadrati magici esotici I quadrati magici di grandi dimensioni possono presentare proprietà stupefacenti. L'esperto di quadrati magici William Benson ha

realizzato una griglia 32 × 32 che è un quadrato magico e che resta tale anche se si elevano i numeri che contiene al quadrato e al cubo. Nel 2001 è stato realizzato un quadrato di ordine 1024 × 1024 in cui tutte le potenze degli elementi, fino alla quinta, formano un quadrato magico. Esistono molti altri risultati analoghi a questi.

Basta rendere le regole più elastiche per ottenere un'ampia varietà di quadrati magici. I quadrati magici normali sono pane per tradizionalisti. È sufficiente rimuovere la condizione che la somma delle diagonali debba eguagliare quella di ciascuna riga e ciascuna colonna per generare una miriade di risultati diversi. Ad esempio, si può andare alla ricerca di quadrati di soli numeri primi, o di figure diverse dal quadrato dotate di proprietà magiche. Aumentando il numero delle dimensioni siamo indotti a immaginare cubi o ipercubi magici.

Ma il premio per il quadrato più straordinario, se non altro dal punto di vista della curiosità, va a un semplice quadrato 3 × 3 costruito dall'enigmista e ingegnere elettronico olandese Lee Sallows.

5	22	18
28	15	2
12	8	25

Cosa ha di tanto straordinario? Scriviamo i numeri in lettere (in inglese)

five	twenty-two	eighteen
twenty-eight	fifteen	two
twelve	eight	twenty-five

Contando le lettere di ciascuna parola, si ottiene

4	9	8
11	7	3
6	5	10

Stranamente, anche questo è un quadrato magico formato dai numeri consecutivi 3, 4, 5 fino a 11. Inoltre si nota che (sempre in inglese) il numero di lettere delle somme magiche di ciascun quadrato 3 × 3 (21, *twenty-one* e 45, *forty-five*) è 9, in accordo col fatto che 3 × 3 = 9.

idea chiave
Stregoneria matematica

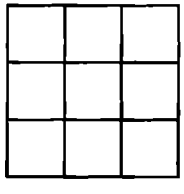
43 I quadrati latini

Da qualche anno impazza la moda del Sudoku. Da una parte all'altra del globo, in attesa dell'ispirazione vincente per il numero da incasellare, milioni di penne e matite vengono masticate nervosamente. Sarà un 4 o un 5? Forse è un 9. Ogni mattina dai treni sbarcano pendolari che hanno impiegato nel Sudoku uno sforzo mentale maggiore di quello che metteranno in campo nel resto della giornata. E la sera le cene vengono dimenticate a carbonizzarsi nel forno. Sarà un 5, un 4 o un 7? Tutte queste persone in realtà giocano con i quadrati latini: fanno i matematici.

	4		8		3			
		7						3
		9	7			2	6	
3				1		7		9
			6	9	8			
1		5		2				6
	2	3			6	5		
6						1		
			5		2		8	

Il segreto del Sudoku Nel Sudoku ci si trova davanti una griglia 9 × 9 in cui sono già riportati alcuni numeri. Lo scopo è quello di riempirla basandosi sui numeri dati. Ogni riga e ogni colonna devono contenere una e una sola volta ciascuna delle cifre 1, 2, 3, ..., 9, e lo stesso vale per ogni quadratino 3 × 3.

Si pensa che il Sudoku (una parola che significa «singole cifre») sia stato inventato alla fine degli anni '70 del Novecento. Ha preso piede negli anni '80 in Giappone e nel 2005 è diventato un fenomeno di massa. L'attrattiva del gioco consiste nel non richiedere una particolare cultura, al contrario dei cruciverba, pur essendo altrettanto irresistibile. In effetti, i fanatici di queste forme di tortura autoinflitta hanno molto in comune.



Quadrati latini 3 × 3 Un quadrato latino è una griglia quadrata in cui ciascun simbolo compare una sola volta in ciascuna riga e ciascuna colonna. Il numero di simboli è uguale alle dimensioni del quadrato e si chiama «ordine». Siamo in grado di riempire una griglia vuota 3 × 3 in modo tale che ogni riga e ogni colonna contengano una e una sola volta tutti i simboli *a*, *b* e *c*? Se la risposta è affermativa, abbiamo creato un quadrato latino di ordine 3.

linea del tempo

1779 Eulero studia la teoria dei quadrati latini

1900 Tarry dimostra che non esistono quadrati latini ortogonali di ordine 6

Quando Eulero introdusse il concetto di quadrato latino, ne parlò come di un nuovo tipo di quadrato magico. Ma, a differenza dei quadrati magici, quelli latini non hanno niente a che vedere con l'aritmetica: i simboli possono anche non essere dei numeri. Il loro nome è dovuto al fatto che Eulero li riempì utilizzando le lettere dell'alfabeto latino, destinando le lettere greche ad altri tipi di quadrato.

a	*b*	*c*
b	*c*	*a*
c	*a*	*b*

La costruzione di un quadrato latino 3 × 3 è immediata.

Se si immagina che *a*, *b* e *c* rappresentino tre giorni della settimana, ad esempio Lunedì, Mercoledì e Venerdì, si può utilizzare il quadrato per programmare una serie di incontri tra due gruppi di persone. Il Gruppo 1 è composto da Lorenzo, Maria e Nadia, il Gruppo 2 da Roberto, Sara e Tommaso.

	R	S	T
L	*a*	*b*	*c*
M	*b*	*c*	*a*
N	*c*	*a*	*b*

Maria del Gruppo 1, per esempio, Lunedì incontrerà Tommaso del Gruppo 2 (perché l'intersezione fra la riga M e la colonna T corrisponde ad *a* = Lunedì). La disposizione secondo il quadrato latino garantisce che si verifichi un incontro per ogni coppia di membri di due squadre diverse e che le date non si sovrappongano.

Questo non è l'unico quadrato latino 3 × 3 che si può ottenere. Se si assume che A, B e C rappresentino gli argomenti discussi nel corso degli incontri tra il Gruppo 1 e il Gruppo 2, si può costruire un quadrato latino che garantisca che ciascuna persona parli di un argomento diverso con ciascun membro dell'altro gruppo.

	R	S	T
L	A	B	C
M	C	A	B
N	B	C	A

Così Maria della Squadra Uno affronterà l'argomento C con Roberto, l'argomento A con Sara e l'argomento B con Tommaso.

Ma *chi* dovrebbe discutere, *quando* e su *quale* argomento? Come dovrebbe essere programmata un'organizzazione così complessa? Per fortuna possiamo combinare simbolo per

1925
Fisher propone di impiegare i quadrati latini in esperimenti di tipo statico

1960
Bose, Parker e Shrikhande dimostrano che la congettura di Eulero sull'inesistenza di alcune coppie di quadrati latini è falsa

1979
Invenzione a New York di giochi ispirati al Sudoku

simbolo i due quadrati latini in modo da generare un quadrato latino composto, in cui ciascuna delle nove possibili coppie di giorni e argomenti ricorre in una sola posizione.

	R	S	T
L	*a*, A	*b*, B	*c*, C
M	*b*, C	*c*, A	*a*, B
N	*c*, B	*a*, C	*b*, A

Questo quadrato può essere utile anche per interpretare lo storico «problema dei nove ufficiali», in cui nove ufficiali appartenenti a tre reggimenti *a*, *b* e *c* e di tre ranghi A, B e C devono schierarsi sul campo di parata in modo tale che ciascuna riga e ciascuna colonna contenga un ufficiale per ciascun reggimento e con ciascun rango. I quadrati latini che possono essere combinati in questo modo sono detti «ortogonali». Nel caso 3×3 la soluzione si trova facilmente, ma per ordini maggiori è tutt'altro che facile trovare coppie di quadrati latini ortogonali. Anche questo fu scoperto da Eulero.

Nel caso del quadrato latino 4×4, il «problema dei 16 ufficiali» potrebbe consistere nel tentativo di disporre tutte le 16 figure di un mazzo di carte in un quadrato in modo che ciascuna figura (asso, re, regina o fante) e ciascun seme (cuori, quadri, fiori, picche) compaia una sola volta in ogni riga e in ogni colonna. Nel 1782 Eulero si pose il problema dei «36 ufficiali»: in sostanza, era alla ricerca di due quadrati ortogonali 6×6. Non essendo stato in grado di trovarli, ipotizzò che non potessero esistere quadrati latini ortogonali di ordine, 6, 10, 14, 18, 22... Ma era possibile dimostrarlo?

Le orme di Eulero furono seguite da Gaston Tarry, un matematico dilettante che lavorava come funzionario pubblico in Algeria. Studiando approfonditamente vari esempi, nel 1900 Tarry riuscì a verificare la congettura di Eulero in un caso: non esistevano coppie di quadrati latini ortogonali di ordine 6. Ovviamente i matematici pensavano che Eulero avesse ragione anche per tutti gli altri ordini (10, 14, 18, 22...).

Nel 1960 tre matematici misero in subbuglio la comunità scientifica dimostrando con uno sforzo congiunto che in tutti gli altri casi Eulero aveva torto. Raj Bose, Ernest Parker e Sharadchandra Shrikhande dimostrarono l'esistenza di coppie di quadrati latini di ordine 10, 14, 18, 22...; l'unico caso in cui non esistevano (oltre ai casi banali di ordine 1 e 2) era proprio quello di ordine 6.

Come abbiamo visto, esistono due quadrati latini ortogonali di ordine 3; per l'ordine 4 si possono produrre tre quadrati ortogonali a due a due. È possibile dimostrare che non esistono mai più di $n - 1$ quadrati latini reciprocamente ortogonali, perciò per $n = 10$ il numero di quadrati ortogonali a due a due è limitato a nove. Trovarli, però, è un'altra storia. A tutt'oggi, nessuno è stato capace di presentare neanche tre quadrati latini di ordine 10 che siano reciprocamente ortogonali.

Ma i quadrati latini hanno una qualche utilità? L'eminente statistico R.A. Fisher comprese l'utilità pratica dei quadrati latini, impiegandoli per rivoluzionare le tecniche agricole durante la sua permanenza presso la Stazione di Ricerca Rothamsted nello Hertfordshire, nel Regno Unito.

L'intento di Fisher era quello di verificare l'efficacia dei fertilizzanti sulla resa del raccolto. In condizioni ideali la semina dovrebbe avvenire su suoli identici, per evitare che la qualità del terreno influisca sulla resa del raccolto. In questo caso si potrebbero applicare i vari fertilizzanti con la certezza di avere eliminato il «disturbo» dovuto alla qualità del suolo. L'unico modo per ottenere condizioni di terreno assolutamente identiche consisterebbe nell'impiegare sempre lo stesso suolo; ciò però comporta di dover arare e riseminare in continuazione. E, se anche fosse possibile, un'altra fonte di disturbo potrebbe derivare dalle diverse condizioni meteorologiche.

Questo problema può essere aggirato utilizzando i quadrati latini. Supponiamo di voler sottoporre a test quattro trattamenti diversi. Suddividendo un campo quadrato in 16 lotti, si può visualizzare un quadrato latino in cui la qualità del suolo varia «orizzontalmente» e «verticalmente».

I quattro fertilizzanti verranno quindi applicati casualmente secondo lo schema *a*, *b*, *c* e *d*, in modo tale che ogni tipo di fertilizzante venga distribuito una sola volta per ciascuna riga e per ciascuna colonna, nel tentativo di rimuovere l'effetto della qualità del suolo. Se poi si sospetta che la resa del raccolto sia influenzata anche da altri fattori, si può provare a tenerne conto. Ammettiamo che l'ora della giornata in cui viene applicato il trattamento costituisca un fattore rilevante: si possono indicare quattro fasce temporali della giornata con *A*, *B*, *C* e *D* e utilizzare i quadrati latini ortogonali per ottenere uno schema di raccolta dati. In questo modo si può essere sicuri che ciascun tipo di trattamento e ciascun orario venga applicato a un solo lotto.

L'esperimento può essere rappresentato così:

a, orario *A*	*b*, orario *B*	*c*, orario *C*	*d*, orario *D*
b, orario *C*	*a*, orario *D*	*d*, orario *A*	*c*, orario *B*
c, orario *D*	*d*, orario *C*	*a*, orario *B*	*b*, orario *A*
d, orario *B*	*c*, orario *A*	*b*, orario *D*	*a*, orario *C*

Volendo, si possono prendere in considerazione anche ulteriori fattori, realizzando schemi a quadrato latino ancora più elaborati. Eulero non poteva certo immaginare che il suo problema degli ufficiali avrebbe trovato applicazione nelle tecniche agronomiche.



44 La matematica finanziaria

Quando si tratta di piazzare biciclette, Roberto è davvero un gran venditore. Pensa che sia suo preciso dovere far salire tutti in sella a una bici, e quando vede un cliente entrare nel suo negozio e comprare senza esitazione una bici a 99 euro, fa salti di gioia. Il cliente paga con un assegno da 150 euro e, dato che le banche sono chiuse, Roberto va a chiedere a un vicino di incassarlo. Al ritorno dà il resto al cliente, che inforca la bici e se ne va di gran carriera. Ma succede un disastro. L'assegno è protestato, il vicino reclama i suoi soldi e Roberto deve rivolgersi a un amico per prenderli in prestito. Se la bici gli era costata 79 euro, quanto ci ha rimesso in totale?

La sostanza di questo rompicapo fu proposta dal grande enigmista Henry Dudeney. Possiamo considerarlo un quesito di matematica finanziaria, anche se si tratta solo di un indovinello che ha a che vedere con i soldi. Fa però capire come il valore del denaro dipenda dal tempo e che l'inflazione fa la sua parte. Infatti, dato che Dudeney scriveva negli anni '20 del Novecento, nel suo problema la bicicletta costava al cliente 15 euro. Un modo per combattere l'inflazione è imporre un interesse sul denaro: siamo nel campo della serissima matematica dei moderni mercati finanziari.

L'interesse composto L'interesse può essere di due tipi, semplice e composto. Concentriamo la nostra curiosità matematica su due fratelli, Carlo Composto e Simone Semplice, che vanno a depositare in banca 1000 euro (a testa) avuti in dono dal padre. Carlo Composto sceglie sempre un conto corrente che applica l'interesse composto, mentre Simone Semplice, da buon tradizionalista, preferisce i conti che applicano l'interesse semplice. Storicamente, l'interesse composto veniva associato all'usura e come tale era malvisto, ma al giorno d'oggi è molto diffuso e

linea del tempo

3000 a.C. I Babilonesi usano un sistema di numerazione sessagesimale per le transazioni finanziarie

1494 Luca Pacioli introduce l'uso del bilancio e della partita doppia

riveste un'importanza centrale nei moderni sistemi monetari. L'interesse composto viene calcolato anche sugli interessi ed è per questo che Carlo lo preferisce. L'interesse semplice non ha questa caratteristica e viene calcolato su un ammontare fisso detto «capitale». Simone riesce a comprenderne il meccanismo senza alcuno sforzo, perché ogni anno il capitale genera gli stessi interessi.

In matematica godere dell'appoggio di Albert Einstein è sempre buon segno, anche se è piuttosto improbabile che Einstein abbia davvero affermato, come viene sostenuto diffusamente, che l'interesse composto è la più grande scoperta di tutti i tempi. Ciò che è indubbio è che la formula dell'interesse composto è più immediata rispetto a $E = mc^2$. Quando si tratta di risparmi, di prestiti, di carte di credito, di mutui o di rendite, la formula dell'interesse composto (vedi a margine) è sempre dietro le quinte a lavorare a nostro vantaggio (o svantaggio). Qual è il significato dei simboli? C rappresenta il capitale (il denaro depositato o preso in prestito), i è il tasso di interesse percentuale diviso per 100 e n conta i periodi di tempo.

$$A = C \times (1 + i)^n$$

Formula dell'interesse composto

Carlo deposita i suoi 1000 euro su un conto che riconosce il 7% annuo di interessi. Quanto maturerà in tre anni? In questo caso C = 1000, i = 0,07 e n = 3. Il simbolo A rappresenta la somma maturata: usando la formula dell'interesse composto si ottiene A = 1225,04 €.

Il conto di Simone applica lo stesso tasso di interesse, il 7%, come interesse semplice. Quanto avrà accumulato Simone fra tre anni? Il primo anno guadagna 70 euro di interessi, e lo stesso varrà per il secondo ed il terzo anno. Perciò in totale maturerà 3 × 70 euro di interessi, arrivando ad avere sul conto 1210 euro. L'investimento migliore è quello di Carlo.

Le somme che maturano interessi composti possono aumentare molto rapidamente, il che va bene quando si risparmia, non certo quando si prende un prestito. Il fattore chiave dell'interesse composto è il periodo di tempo che fa scattare gli interessi. Carlo ha sentito parlare di un fondo che rende l'1% alla settimana, cioè un centesimo per ogni euro. Quanto può pensare di guadagnarci?

Simone crede di conoscere la risposta: suggerisce di moltiplicare il tasso di interesse dell'1% per 52 (il numero di settimane in un anno) per ottenere un tasso percentuale annuo del 52%. Ciò comporterebbe 520 euro di interessi che concorrerebbero a dare un totale di 1520 euro. Carlo gli ricorda, però, che la formula dell'interesse composto può

1718
Abraham de Moivre studia le statistiche sulla mortalità e la teoria delle rendite vitalizie

1756
James Dodson pubblica *First Lectures on Insurances*

1848
Viene fondato a Londra l'Istituto degli Attuari

fare miracoli. Con C = 1000, i = 0,01 e n = 52, calcola un ammontare di $1000 \times (1{,}01)^{52}$ = 1677,69 euro, molto maggiore della somma calcolata da Simone Semplice. Il tasso equivalente annuo percentuale di Carlo è del 67,769% e supera di gran lunga il 52% calcolato da Simone.

Simone resta davvero colpito, ma i suoi soldi sono già in banca a regime di interesse semplice. Si domanda quanto tempo occorrerà per raddoppiare la somma iniziale di 1000 €. Dato che ogni anno riceve 70 euro di interessi, non deve fare altro che dividere 1000 per 70: il risultato è 14,29, quindi può star sicuro che tra 15 anni avrà in banca più di 2000 euro. Certo, l'attesa è lunga. Per dimostrargli la superiorità dell'interesse composto, Carlo si mette a calcolare il tempo di raddoppio del proprio capitale. È un po' più complicato, ma per fortuna un amico gli svela la regola del 72.

La regola del 72 Dato un tasso percentuale qualsiasi, la regola del 72 consente di stimare il numero di cicli necessari a far raddoppiare una somma di denaro. Il problema di Carlo è espresso in termini di anni, ma la regola del 72 può essere applicata anche ai giorni o ai mesi. Per trovare il tempo di raddoppio Carlo deve semplicemente dividere 72 per il tasso di interesse: il calcolo è 72/7 = 10,3, perciò Carlo può riferire al fratello che il suo investimento raddoppierà in 11 anni, molto più rapidamente dei 15 anni di Simone. Si tratta di una regola approssimata, che però consente di prendere rapidamente una decisione.

Il valore attuale Il padre di Carlo Composto resta talmente impressionato dall'assennatezza del figlio da prenderlo da parte e dirgli: «Ti regalerò 100 000 euro». Carlo ne è felicissimo, ma suo padre aggiunge la condizione che il regalo arriverà solo in occasione del suo 45° compleanno, cioè fra dieci anni. Carlo non è più tanto contento.

Carlo vorrebbe spendere i soldi subito, ma ovviamente non può. Allora va in banca, dove si impegna a restituire i 100 000 euro nel giro di dieci anni. La banca risponde che il tempo è denaro e che fra dieci anni 100 000 euro non varranno quanto valgono oggi. Inoltre, stima quale investimento possa equivalere al valore di 100 000 euro tra dieci anni: sarà la cifra che potrà dargli in prestito. La banca calcola che un tasso di crescita del 12% potrà garantirle un profitto sicuro. Qual è la somma che in 10 anni raggiungerà il valore di 100 000 € al 12% di interessi? La formula dell'interesse composto consente di risolvere anche questo problema. Questa volta conosciamo A = 100 000 €, e dobbiamo calcolare C, cioè il valore attuale del capitale. Dati n = 10 e i = 0,12, la banca si prepara ad anticipare a Carlo la somma di $100\,000/1{,}12^{10}$ = 32 197,32 €. Carlo è sconvolto dall'esiguità della somma, ma potrà comunque permettersi una macchina nuova.

Un piano di risparmio ragionevole Ora che il padre di Carlo ha promesso di dare al figlio 100 000 euro tra dieci anni, deve cominciare a mettere il denaro da parte. Pianifica quindi una serie di versamenti sempre uguali su un conto di deposito, da effettuare alla fine di ogni anno. Alla fine di questo periodo sarà in grado di consegnare i soldi a Carlo nella data prevista e Carlo potrà restituire il denaro alla banca per estinguere il prestito.

$$S = R \times \frac{((1+i)^n - 1)}{i}$$

Formula del piano di risparmio

Il padre di Carlo riesce a trovare un conto adatto, che gli garantisce un interesse dell'8% per tutto il decennio. Il compito di calcolare il piano dei versamenti annuali tocca a Carlo. Carlo sa utilizzare la formula dell'interesse composto nel caso di un pagamento (che corrisponde al capitale iniziale), ma ora deve trovare un risultato per dieci pagamenti che avvengono in momenti successivi. Se viene versata regolarmente una somma R alla fine di ogni anno e il tasso di interesse è dell'8%, il montante dei risparmi al temine di un periodo di n anni si calcola con la formula del piano di risparmio.

Sapendo che S = 100 000 €, n = 10 e i = 0,08, Carlo trova che R = 6902,95 €. Ora che ha una bella macchina nuova di zecca, gentilmente concessa dalla banca, ha bisogno di un garage dove custodirla. Decide allora di accendere un mutuo di 300 000 € per l'acquisto di una casa, contando di restituire la somma con una serie di versamenti annuali di uguale importo. Rendendosi conto che in questo problema i 300 000 euro rappresentano il valore attuale del capitale, Carlo può calcolare facilmente l'ammontare dei versamenti annuali che dovrà effettuare. Il padre, profondamente colpito, decide di ricorrere nuovamente all'abilità del figlio: ha appena riscosso 150 000 euro di liquidazione e ne vuole ricavare una rendita vitalizia. «Nessun problema», dice Carlo, «dato che il ragionamento è lo stesso, basta usare la stessa formula. Ma invece di essere io a farmi anticipare i soldi dalla compagnia finanziaria per poi restituirli a rate, sarai tu a dar loro dei soldi e saranno loro a effettuare dei versamenti regolari a tuo favore».

A proposito: la risposta al rompicapo di Henry Dudeney è 130 euro, che ammontano alla somma dei 51 euro dati di resto al cliente e del prezzo pagato da Roberto per la bicicletta.

idea chiave

L'interesse composto conviene

45 Il problema della dieta

Sandra Rossi prende molto sul serio l'atletica leggera. Va tutti i giorni in palestra e segue una dieta ferrea. Sandra si guadagna da vivere lavorando part-time e deve fare attenzione ai soldi. Per mantenersi in salute e in piena forma è di vitale importanza che assuma ogni mese il quantitativo di vitamine e di sali minerali stabilito dal suo allenatore, il quale ritiene che una futura campionessa olimpica debba assumere almeno 120 milligrammi (mg) di vitamine e 880 mg di sali minerali al mese. Per raggiungere questo obiettivo, Sandra confida in due integratori alimentari. Il primo viene venduto in forma solida col nome di Solido, mentre il secondo è liquido ed è in commercio come Liquex. Sandra ha un problema: deve stabilire quali quantità deve acquistare dell'uno e dell'altro per soddisfare le richieste dell'allenatore.

Il classico problema della dieta consiste nell'architettare una dieta salutare con la minima spesa. È stato un problema di base della programmazione lineare, una disciplina sviluppata negli anni '40 del Novecento e ora impiegata in un'ampia gamma di applicazioni.

All'inizio di marzo Sandra fa un salto al supermercato e dà un'occhiata al Solido e al Liquex. Osservando il retro delle confezioni, scopre che un pacchetto di Solido contiene 2 mg di vitamine e 10 mg di sali minerali, mentre un cartone di Liquex contiene 3 mg di vitamine e 50 mg di sali. Riempie quindi scrupolosamente il carrello con 30 pacchetti di Solido e 5 cartoni di Liquex, che dovranno

	Solido	Liquex	Quantità richieste
Vitamine	2 mg	3 mg	120 mg
Sali	10 mg	50 mg	880 mg

linea del tempo

1826 Fourier precorre la programmazione lineare; Gauss risolve sistemi di equazioni lineari col metodo dell'eliminazione gaussiana

1902 Farkas propone una soluzione per sistemi di disequazioni

bastarle per tutto il mese. Mentre si fa strada verso la cassa si domanda se le saranno sufficienti, perciò fa il conto di quante vitamine ha nel carrello: i 30 pacchetti di Solido contengono $2 \times 30 = 60$ mg di vitamine, mentre il Liquex ne contiene $3 \times 5 = 15$ mg. In tutto, Sandra ha $2 \times 30 + 3 \times 5 = 75$ mg di vitamine. Quindi ripete il calcolo per i sali minerali e si accorge di averne $10 \times 30 + 50 \times 5 = 550$ mg.

Dato che l'allenatore le ha prescritto di assumere almeno 120 mg di vitamine e 880 mg di sali minerali, Sandra deve aggiungere al carrello ulteriori pacchetti e cartoni. Il problema di Sandra è far quadrare i quantitativi di Solido e Liquex con le vitamine e i sali richiesti. Torna verso il reparto salute del supermercato e mette nel carrello altre confezioni. Adesso ha 40 pacchetti e 15 cartoni: basteranno? Ripetendo i calcoli, scopre di avere a disposizione $2 \times 40 + 3 \times 15 = 125$ mg di vitamine e $10 \times 40 + 50 \times 15 = 1150$ mg di sali minerali. Avendo addirittura superato le quantità richieste, Sandra è sicura di riuscire a rispettare le prescrizioni del suo allenatore.

Soluzioni ammissibili Quindi la combinazione di integratori (40, 15) consentirà a Sandra di rispettare la dieta. Si tratta di una possibile soluzione, che in genere si definisce un risultato «ammissibile». Abbiamo già visto, invece, che (30, 5) non è una soluzione ammissibile. Esiste una linea di demarcazione tra due tipi di combinazioni: quelle ammissibili consentono di rispettare la dieta, quelle non ammissibili no.

In realtà, Sandra ha più di due opzioni. Potrebbe decidere di riempire il carrello soltanto con il Solido; in tal caso le servirebbero almeno 88 pacchetti. L'acquisto (88, 0) soddisfa entrambe le richieste, perché con questa combinazione si ottengono $2 \times 88 + 3 \times 0 = 176$ mg di vitamine e $10 \times 88 + 50 \times 0 = 880$ mg di sali minerali. Ma potrebbe comprare anche soltanto il Liquex e in questo caso le servirebbero almeno 40 cartoni. Infatti la soluzione ammissibile (0, 40) soddisfa sia la richiesta di vitamine che quella di sali minerali, dato che $2 \times 0 + 3 \times 40 = 120$ mg di vitamine e $10 \times 0 + 50 \times 40 = 2000$ mg di sali. Nessuna di queste combinazioni contiene esattamente il quantitativo richiesto di vitamine e sali, ma per l'allenatore sarà sufficiente che Sandra ne assuma almeno la quantità stabilita.

Soluzioni ottimali È giunto il momento di tener conto dei costi: quando Sandra arriverà alla cassa le sarà richiesto di pagare i suoi acquisti. Sandra osserva che pacchetti e cartoni costano entrambi 5 euro. Il conto per ciascuna delle soluzioni possibili finora esaminate (40, 15), (88, 0) e (0, 40), sarà, rispettivamente, di 275, 440 e 200 euro, perciò l'acquisto di 40 cartoni di Liquex e nessun pacchetto di Solido rappresenta la soluzione più conveniente. Si tratta della più economica

1945 Stigler risolve il problema della dieta con il metodo euristico

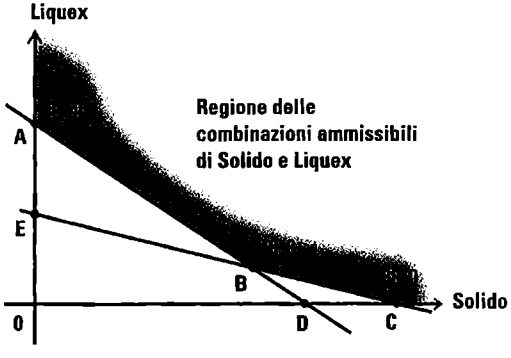
1947 Dantzig formula l'algoritmo del simplesso e risolve il problema della dieta con la programmazione lineare

1984 Karmarkar ottiene un nuovo algoritmo per risolvere i problemi di programmazione lineare

tra tutte le combinazioni che consentono di rispettare le richieste della dieta. Questa soluzione, però, è stata trovata casualmente: Sandra ha tentato tre combinazioni di Solido e Liquex scelte d'impulso e ha calcolato il costo solo in questi tre casi. Si può fare di meglio. Esiste una combinazione di Solido e Liquex che possa, allo stesso tempo, soddisfare il suo allenatore e costare il meno possibile? Sandra decide di andare a casa per analizzare il problema con carta e penna.

Problemi di programmazione lineare Da sempre, Sandra si è allenata a visualizzare i propri obiettivi. Se questa tecnica funziona nella corsa per un oro olimpico, perché non applicarla anche alla matematica? Quindi Sandra fa un grafico della regione di ammissibilità: è in grado di farlo, perché sta prendendo in considerazione solo due integratori. La retta AD rappresenta le combinazioni di Solido e Liquex che contengono esattamente 120 mg di vitamine, perciò le combinazioni al di sopra di questa retta contengono più di 120 mg di vitamine. La retta EC rappresenta le combinazioni che contengono almeno 880 mg di sali minerali. Le combinazioni di integratori che stanno al di sopra di entrambe le rette costituiscono la regione di ammissibilità e rappresentano tutte le combinazioni ammissibili che potevano essere acquistate da Sandra.

I problemi che rientrano nello stesso schema di quello della dieta sono problemi di programmazione lineare. Per «programmazione» si intende una procedura codificata (nell'uso che si faceva del termine prima dell'avvento dei computer), mentre «lineare» si riferisce all'utilizzo di linee rette. Per trovare una soluzione al problema di Sandra con la programmazione lineare, i matematici hanno dimostrato che è sufficiente calcolare la spesa nei punti d'angolo del grafico di Sandra. Sandra ha scoperto che esiste un'altra soluzione ammissibile in corrispondenza del punto B di coordinate (48, 8), che equivale all'acquisto di 48 pacchetti di Solido e 8 cartoni di Liquex. In questo caso la dieta verrebbe rispettata in modo esatto, perché questa combinazione corrisponde a 120 mg di vitamine e 880 mg di sali minerali. Calcolando un prezzo di 5 € sia per il pacchetto di Solido che per il cartone di Liquex, la spesa totale per questa combinazione sarebbe di 280 €. L'acquisto ottimale resta quello già visto: Sandra dovrà acquistare 40 cartoni di Liquex e nessun pacchetto di Solido per un prezzo complessivo di 200 €, nonostante la combinazione fornisca 1120 mg di vitamine, ben più degli 880 mg richiesti.



È il costo relativo degli integratori a determinare, in ultima analisi, la soluzione ottimale. Se il prezzo del Solido scendesse a 2 € e quello del Liquex salisse a 7 €, le combinazioni corrispondenti ai punti d'angolo A (0, 40), B (48, 8) e C (88, 0) costerebbero, rispettivamente, 280 €, 152 € e 176 €.

In questo regime di prezzi la soluzione più conveniente per Sandra consisterebbe nell'acquisto di 48 pacchetti di Solido e di 8 cartoni di Liquex, per un totale di 152 €.

Storia Nel 1947 il matematico americano George Dantzig, che al tempo lavorava per l'aviazione, formulò l'algoritmo del simplesso, un metodo di soluzione dei problemi di programmazione lineare che ebbe un successo tale da far conoscere Dantzig in tutto l'Occidente come il padre della programmazione lineare. Nella Russia sovietica Leonid Kantorovič, che era tagliato fuori dalla comunità internazionale a causa della guerra fredda, sviluppò una teoria indipendente sulla programmazione lineare. Nel 1975 Kantorovič e il matematico olandese Tjalling Koopmans ricevettero il premio Nobel per l'economia per gli studi sull'allocazione delle risorse, basati anche sulla programmazione lineare.

Sandra prendeva in considerazione solo due integratori, cioè due variabili, ma al giorno d'oggi esistono molti problemi che coinvolgono migliaia di variabili. Quando Dantzig mise a punto il suo metodo c'erano ben pochi computer, ma era attivo il Progetto Tavole Numeriche, un progetto di lavoro decennale che ebbe inizio nel 1938 a New York. Furono necessari dodici giorni di lavoro di una squadra di dieci calcolatori umani per risolvere un problema della dieta con nove «vitamine» richieste e 77 variabili.

Nonostante l'algoritmo del simplesso e le sue varianti hanno avuto un successo fenomenale, sono stati sperimentati anche altri metodi. Nel 1984 il matematico indiano Narendra Karmarkar ottenne un nuovo algoritmo di notevole importanza pratica, mentre il russo Leonid Khachiyan ne propose uno di fondamentale rilevanza teorica.

Il modello di base della programmazione lineare è stato applicato a molte altre situazioni, oltre alla scelta di una dieta. Una di queste è il problema dei trasporti, che ha a che vedere col trasporto delle merci dalle fabbriche ai magazzini. Il problema ha una struttura particolare e ha dato vita a un campo a sé: in questo caso l'obiettivo consiste nel minimizzare il costo del trasporto. In altri problemi di programmazione lineare, invece, l'obiettivo è quello di massimizzare una grandezza (ad esempio il profitto). In altre situazioni, le variabili possono assumere valori interi oppure soltanto i due valori 0 e 1. Si tratta di problemi molto diversi che necessitano di apposite strategie risolutive.

Resta da vedere se Sandra Rossi riuscirà a vincere la medaglia d'oro ai giochi olimpici: se ce la farà, si tratterà di un ulteriore successo della programmazione lineare.

idea chiave

Mantenersi in salute a ogni costo

46 Il commesso viaggiatore

James Cook vive a Bismarck (North Dakota, USA), ed è uno dei migliori venditori della Electra, una ditta che produce aspirapolvere. Per capire quanto sia abile, basta ricordare che è stato eletto venditore dell'anno per tre anni consecutivi. La sua regione di vendita comprende le città di Albuquerque, Chicago, Dallas e El Paso. James, seguendo uno schema a rotazione, le visita tutte una volta al mese. Il problema che deve porsi è come completare il giro facendo il minor numero di kilometri possibile. Si tratta del classico problema del commesso viaggiatore.

James ha tracciato una tabella delle distanze chilometriche fra le città. Ad esempio, la distanza tra Bismarck e Dallas è di 1641 kilometri, come si legge all'intersezione (ombreggiata) tra la colonna di Bismarck e la riga di Dallas.

Albuquerque				
1421	Bismarck			
1831	1136	Chicago		
933	1641	1263	Dallas	
380	1770	2021	948	El Paso

Il metodo goloso James Cook è un tipo pratico, perciò nel fare uno schizzo della sua regione di vendita non si preoccupa di essere preciso: si accontenta di avere un'idea di massima delle posizioni e delle distanze tra le città. Uno dei suoi percorsi abituali parte da Bismarck, passa da Chicago, Albuquerque, Dallas ed El Paso, per poi tornare a Bismarck. Si tratta del percorso BCADEB, ma Cook si rende conto che questo giro di 6618 kilometri non è economico in termini di distanze percorse. Si può fare di meglio.

Il fatto che pianifichi la regione di vendita non deve trarci in inganno: James non è tipo da programmazioni dettagliate, il suo scopo è vendere. Dando un'occhiata ad

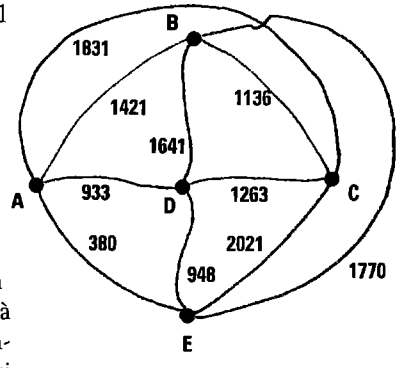
linea del tempo

circa 1810 — Charles Babbage si interessa al problema del commesso viaggiatore

1831 — Si pone il problema pratico del commesso viaggiatore

1926 — Borůvka elabora l'algoritmo goloso

una mappa nel suo ufficio di Bismarck, vede che la città più vicina è Chicago, che si trova a 1136 kilometri di distanza, contro i 1421 di Albuquerque, i 1641 di Dallas e i 1770 di El Paso. Parte quindi subito per Chicago senza aver fatto alcun piano generale. Solo dopo essere giunto a Chicago e aver portato a termine i suoi affari comincia a pensare alla tappa successiva. Sceglie Dallas, rispetto a Albuquerque ed El Paso, poiché si trova a 1263 kilometri da Chicago e quindi è l'opzione più vicina.



Arrivato a Dallas, James ha totalizzato 1136 + 1263 kilometri. Ora deve scegliere tra Albuquerque ed El Paso, e sceglie la prima perché è più vicina. Per concludere la visita a tutte le città dovrà raggiungere El Paso da Albuquerque, dopo di che, a lavoro completato, potrà tornare a Bismarck. Il conto totale dei kilometri ammonta ora a 1136 + 1263 + 933 + 380 + 1770 = 5482 kilometri. La strada BCDAEB è stata molto più breve di quella vista in precedenza e ha risparmiato all'ambiente un bel po' di emissioni di anidride carbonica.

Questo metodo per trovare la strada più breve viene spesso chiamato «metodo goloso», perché James Cook prende sempre decisioni su scala locale: si trova in una determinata città e cerca il percorso più breve a partire da quella. Con questo metodo non guarda mai oltre il passo successivo e non segue alcuna strategia generale per trovare la strada migliore. Infatti, concludendo il giro a El Paso sarà costretto a fare un lungo viaggio di ritorno verso Bismarck. A questo punto James è curioso di sapere se il percorso che ha seguito è davvero il più breve possibile.

James si rende conto di poter sfruttare il fatto che ci sono solo cinque città; infatti è possibile elencare tutti i possibili percorsi e scegliere il più breve. Nel caso di cinque città ci sono solo 24 percorsi da esaminare, che si riducono a 12 se si tiene conto dell'equivalenza tra una strada e quella inversa, cosa ammissibile visto che i kilometri percorsi sarebbero gli stessi. Il metodo funziona più che bene e James Cook scopre che il percorso ottimale è BAEDCB (o il suo inverso BCDEAB), lungo soltanto 5147 kilometri.

Albuquerque				
12 (auto)	Bismarck			
6 (aereo)	2 (aereo)	Chicago		
2 (aereo)	4 (aereo)	3 (aereo)	Dallas	
4 (auto)	3 (aereo)	5 (aereo)	1 (aereo)	El Paso

Una volta tornato a Bismarck, però, si rende conto che il suo viaggio dura troppo a lungo. Ora non vuole risparmiare in termini di di-

1954

Dantzig e Dijkstra propongono alcuni metodi per affrontare il problema del commesso viaggiatore

1971

Cook introduce le categorie *P* versus *NP* per gli algoritmi

2004

David Applegate risolve il problema per tutte le 24 978 città della Svezia

stanza, ma è interessato solo al tempo. Allora traccia una nuova tabella che tiene conto del tempo di percorrenza fra le diverse città della sua zona.

Quando era in ballo il numero di kilometri, James poteva star certo che la somma delle distanze percorse lungo i due lati di un triangolo superasse sempre la lunghezza percorsa lungo l'altro lato. In questo caso il grafico è euclideo e sono noti molti metodi risolutivi. Ma ciò non vale per questioni legate al tempo. Volare lungo rotte principali spesso fa risparmiare tempo rispetto a intraprendere rotte secondarie e James Cook osserva che per andare in aereo da El Paso a Chicago conviene passare da Dallas. La cosiddetta diseguaglianza triangolare in questo caso non vale.

Il metodo goloso applicato al problema del tempo fornisce un tempo complessivo di 22 ore per il percorso BCDEAB, mentre esistono due percorsi ottimali da 14 ore ciascuno, BCADEB e BCDAEB. Fra questi, il primo è lungo 6618 kilometri e il secondo 5482, quindi James Cook prende atto che il risparmio complessivo maggiore si ottiene scegliendo il percorso BCDAEB. Per il futuro si ripropone di studiare il percorso più economico.

Dai secondi ai secoli Ma la vera difficoltà del problema del commesso viaggiatore si presenta quando sono coinvolte molte città. Dato che James Cook è un impiegato così brillante, non passerà molto tempo prima che venga promosso supervisore: le città che dovrà visitare a partire da Bismarck saranno 13 invece delle 4 precedenti. In questo caso, non può accontentarsi di usare il metodo goloso e vorrebbe dare un'occhiata alla lista completa dei percorsi possibili, perciò si mette d'impegno a elencare le strade percorribili per raggiungere le sue 13 città. Ben presto, però, si accorge che i percorsi da esaminare sarebbero dell'ordine di $3{,}1 \times 10^9$. Vista da un'altra ottica, se un computer impiegasse un secondo per stampare ciascuna strada, impiegherebbe circa un secolo per stamparle tutte. Un problema con 100 città potrebbe impegnare il computer per millenni.

Il problema del commesso viaggiatore è stato affrontato con metodi sofisticati. Sono stati trovati metodi esatti che funzionano fino a 5000 città e si è venuti a capo perfino di un problema con 33 810 città, usando un computer con una potenza eccezionale. Metodi non esatti producono percorsi che hanno una certa probabilità di trovarsi nell'intorno di quello ottimale; il vantaggio di questi tipi di metodo è che consentono di gestire casi riguardanti milioni di città.

Complessità computazionale Se guardiamo il problema dal punto di vista del computer, basta misurare il tempo necessario a trovare una soluzione. Lo scenario più sfavorevole è dato dal semplice elenco di tutti i percorsi possibili e James ha già scoperto che questo metodo brutale richiede quasi un secolo. Basta aggiungerci un altro paio di città perché il tempo diventi circa 20 000 anni!

È chiaro che queste stime dipendono dal tipo di computer usato in pratica, ma si può dire che, date n città, il tempo cresce come n fattoriale ($n!$, il numero che si

ottiene moltiplicando tutti i numeri interi da uno a n). Per 13 città si calcolano $3{,}1 \times 10^9$ percorsi. Stabilire ogni volta se l'ultimo percorso trovato è davvero il più breve diventa questione di tempo *fattoriale*: un tempo lunghissimo.

Per affrontare il problema sono disponibili anche altri metodi, in cui il tempo per n città aumenta secondo 2^n (cioè 2 moltiplicato n volte per se stesso), quindi il numero di decisioni da prendere per 13 città è dell'ordine di 8192 (8 volte maggiore che per 10 città). Un metodo con questo grado di complessità è chiamato «algoritmo *esponenziale*». L'uovo di Colombo in questi «problemi di ottimizzazione combinatoria» sarebbe scovare un algoritmo che non dipenda dalla potenza n-esima di 2, ma da una potenza fissa di n, meglio se piccola. Ad esempio, se l'algoritmo variasse secondo n^2, nel caso di 13 città ci troveremmo di fronte ad appena 169 decisioni e si impiegherebbe meno del doppio del tempo necessario per 10 città. Un metodo di questo tipo si dice «algoritmo polinomiale» e i problemi che si risolvono seguendo questi algoritmi sono «veloci»; possono richiedere 3 minuti, invece di secoli.

La classe di problemi che un computer può *risolvere* in tempo polinomiale viene indicata con P. Non sappiamo se il problema del commesso viaggiatore possa rientrare in questa classe: nessuno ha infatti ancora trovato un algoritmo a tempo polinomiale in grado di risolverlo, ma non è stato neanche dimostrato che non possa esistere.

Una classe più ampia, detta NP, comprende i problemi che hanno soluzioni che possono essere *verificate* in un tempo polinomiale. Il problema del commesso viaggiatore ne fa parte, perché è sicuramente possibile verificare in un tempo polinomiale se un percorso *dato* è più breve di una qualsiasi distanza fissa: basta sommare le distanze percorse e confrontare il totale con il numero dato. Ma *trovare* e *verificare* sono operazioni ben diverse: è semplicissimo verificare che $167 \times 241 = 40247$, mentre proporsi di trovare i fattori di 40247 è una questione ben più ardua.

Si può affermare che *qualsiasi* problema verificabile in tempo polinomiale può anche essere risolto in tempo polinomiale? Se ciò fosse vero le classi P e NP coinciderebbero e potremmo scrivere $P = NP$. La questione attualmente più scottante per chi si occupa di scienza dei computer è proprio quella di stabilire se $P = NP$. Più della metà degli esperti pensa che ciò non sia vero, ritenendo che esistano problemi verificabili in tempo polinomiale che non possono però essere risolti in tempo polinomiale. È un problema talmente eccezionale che l'Istituto Matematico Clay ha messo in palio un premio di un milione di dollari per chi sarà in grado di dimostrare se $P = NP$ o $P \neq NP$.

idea chiave

47 La teoria dei giochi

Erano in molti a pensare che il piccolo Johnny fosse la persona più intelligente del mondo. John von Neumann fu un bambino prodigio e divenne una leggenda nella comunità matematica: nessuno fece una piega quando si seppe che aveva scribacchiato il «teorema minimax» mentre era in taxi diretto a un convegno. Era proprio il tipo di cose che von Neumann era solito fare. Visto che aveva dato un contributo alla meccanica quantistica, alla logica, all'algebra, come avrebbe potuto sfuggirgli la teoria dei giochi? Infatti fu autore, assieme a Oskar Morgenstern, dell'autorevole *Teoria dei giochi e comportamento economico*. In senso lato, la teoria dei giochi è materia antica, ma il contributo di von Neumann fu fondamentale per affinare la teoria dei giochi «a due giocatori e a somma zero».

Giochi a due giocatori e a somma zero Sembra complicato, ma un gioco a due persone e somma zero si «gioca» semplicemente tra due persone, due aziende o due squadre, e quanto viene vinto da un partecipante corrisponde esattamente a quanto viene perso dall'altro: se A vince 200 € allora B ne perde 200. Non c'è motivo per cui A e B debbano cooperare: si tratta di pura competizione tra vincenti e perdenti. Se ci mettiamo nell'ottica di considerare solo vincite, A vince 200 € e B vince – 200 €, perciò la somma è 200 + (– 200) = 0. Ecco perché si parla di «somma zero».

Immaginiamo che le due compagnie televisive ATV e BTV stiano formulando un'offerta per un nuovo programma di informazione in Scozia o in Inghilterra. Ciascuna compagnia potrà fare un'offerta per una sola delle due nazioni e baserà la propria scelta sulle proiezioni sull'incremento degli ascolti. Entrambe le aziende hanno infatti accesso a una ricerca, condotta da esperti di comunicazione di massa,

linea del tempo

1713 Waldegrave fornisce la prima soluzione matematica di un gioco a due giocatori

1944 Von Neumann e Morgenstern pubblicano la *Teoria dei giochi e comportamento economico*

che stima l'aumento degli ascolti nelle diverse situazioni. I dati della ricerca (espressi in milioni di ascoltatori) sono stati disposti in una «tabella dei vantaggi».

		BTV	
		Scozia	Inghilterra
ATV	Scozia	+ 5	– 3
	Inghilterra	+ 2	+ 4

Nel caso in cui sia ATV che BTV decidano di operare in Scozia, ATV acquisterà 5 milioni di spettatori, perciò BTV ne perderà 5 milioni. Dove compare un valore negativo ci si riferisce a una perdita di ascoltatori da parte di ATV, perciò il segno + indica un vantaggio per ATV, mentre il segno – corrisponde a un vantaggio per BTV.

Ammetteremo che le due compagnie prendano una decisione esclusivamente in base alla «tabella dei vantaggi», presentando le offerte simultaneamente e in busta chiusa. Ovviamente ciascuna azienda agirà nel proprio interesse.

Se ATV sceglie la Scozia, la situazione peggiore che potrà capitarle sarà di perdere 3 milioni di telespettatori; se fa la propria offerta per l'Inghilterra, la condizione più sfavorevole sarà un guadagno di 2 milioni. Quindi la strategia ovvia per ATV sarà scegliere l'Inghilterra, visto che, a prescindere dalla scelta operata da BTV, il quadro più negativo sarà rappresentato da un incremento di 2 milioni di ascoltatori. Se guardiamo i numeri, la ATV ha individuato il minimo di ciascuna riga (–3 e 2) e ha scelto la riga che corrisponde al massimo fra i due.

La BTV si trova in una posizione più debole, tuttavia può elaborare una strategia che minimizzi le proprie perdite potenziali, sperando in una migliore tabella dei vantaggi per l'anno successivo. Se BTV sceglie la Scozia (colonna 1) la situazione peggiore che le può capitare è una perdita di 5 milioni; se sceglie l'Inghilterra la perdita più cospicua è di 4 milioni. La strategia più sicura per la BTV è scegliere l'Inghilterra, visto che è preferibile perdere 4 milioni di ascoltatori ri-

A beautiful mind

John F. Nash (n. 1928), la cui vita tormentata è stata descritta nel film del 2001 *A Beautiful Mind*, nel 1994 ha vinto il premio Nobel per l'economia per i suoi contributi alla teoria dei giochi.

Insieme ad altri studiosi, Nash ha esteso la teoria dei giochi a un numero di giocatori maggiore di due e a situazioni in cui i giocatori possono collaborare, ad esempio per coalizzarsi ai danni di un terzo. L'equilibrio di Nash (che può essere considerato un equilibrio di punto di sella) ha ampliato le prospettive rispetto all'ottica di von Neumann, consentendo di raggiungere una comprensione più profonda delle situazioni economiche.

1950
Tucker enuncia il dilemma del prigioniero e Nash propone l'equilibrio di Nash

1982
Maynard Smith pubblica *Evoluzione e teoria dei giochi*

1994
Nash riceve il premio Nobel per l'economia per le sue ricerche sulla teoria dei giochi

spetto a 5. Indipendentemente dalla scelta che farà ATV, non potrà capitarle di peggio che una perdita di 4 milioni.

Quelle appena viste sono le strategie più accorte per ciascun giocatore; se entrambi le seguono, ATV otterrà 4 milioni di spettatori in più, mentre BTV li perderà.

Quando un gioco è determinato? L'anno successivo le due compagnie televisive hanno un'opzione in più: operare nel Galles. Dato che le circostanze sono cambiate, c'è una nuova tabella dei vantaggi.

		BTV			
		Galles	Scozia	Inghilterra	minimo di riga
ATV	Galles	+3	+2	[illegible]	+1
	Scozia	+4	−1	0	−1
	Inghilterra	−3	+5	−2	−3
	massimo di colonna	+4	+5	+1	

Come abbiamo visto, la strategia più sicura per la ATV consiste nello scegliere la riga che corrisponde al migliore tra gli scenari negativi. Il massimo di $\{+1, -1, -3\}$ equivale a scegliere il Galles (prima riga). Il piano più sicuro per la BTV consiste nello scegliere la colonna che minimizza $\{+4, +5, +1\}$, ovvero l'Inghilterra (colonna 3).

Scegliendo il Galles la ATV ha la *certezza* di guadagnare non meno di 1 milione di spettatori, a prescindere dalle scelte di BTV, mentre per BTV optare per l'Inghilterra significa avere la *garanzia* che le perdite non supereranno 1 milione di spettatori, qualunque sia il comportamento di ATV. Queste scelte sono le strategie *migliori* per ciascuna delle due aziende, pertanto il gioco è deciso (anche se resta sfavorevole a BTV). In questo gioco si ha che

$$\text{massimo di } \{+1, -1, -3\} = \text{minimo di } \{+4, +5, +1\}$$

e in entrambi i membri dell'equazione compare il valore +1. A differenza del primo gioco, questa volta esiste una situazione di equilibrio con un «punto di sella» pari a +1.

Giochi ripetuti La morra cinese (sasso, carta, forbice) è l'emblema dei giochi ripetitivi. A differenza di quello delle compagnie televisive, che si gioca una volta per tutte, questo gioco in genere viene giocato una mezza dozzina di volte, o anche qualche centinaio di volte in occasione dei Campionati Mondiali annuali.

	carta	forbici	sasso	minimo di riga
carta	pari = 0	perde = −1	vince = +1	−1
forbici	vince = +1	pari = 0	perde = −1	−1
sasso	perde = −1	vince = +1	pari = 0	−1
massimo di colonna	+1	+1	+1	

Nella morra cinese i due giocatori rappresentano (rispettivamente) carta, forbici o sasso mostrando alternativamente il palmo della mano, due dita o il pugno.

Si conta fino a tre e i giocatori fanno la loro mossa simultaneamente. La carta pareggia con la carta, perde con le forbici (perché le forbici possono tagliare la carta) e vince con il sasso (perché la carta può avvolgere la pietra): se si gioca «carta» i vantaggi sono 0, – 1, + 1, come si legge nella prima riga della tabella dei vantaggi.

In questo gioco non ci sono punti di sella, quindi non è facile trovare una strategia logica. Se un giocatore fa sempre la stessa mossa, ad esempio «carta», il suo avversario se ne accorgerà e sceglierà «forbici», finendo per vincere sempre. Secondo il «teorema minimax» di von Neumann esiste una strategia mista, ovvero un modo per scegliere mosse diverse sulla base del calcolo delle probabilità. Per i matematici la strategia migliore consiste in una scelta casuale delle mosse, purché in totale le scelte di carta, forbici e sasso si mantengano nella proporzione relativa di un terzo. Scegliere completamente a caso, però, non sempre è la strada più azzeccata e i campioni del mondo aggiungono alle loro strategie un pizzico di astuzia psicologica. In parole povere, riescono a prevedere le intenzioni dei loro avversari.

Quando un gioco non è a somma zero? Non tutti i giochi sono a somma zero; a volte ciascun giocatore ha una sua tavola dei vantaggi distinta. Un celebre esempio è il «dilemma del prigioniero» enunciato da A.W. Tucker.

Due uomini, Andrea e Bernardo, vengono arrestati dalla polizia con l'accusa di aver commesso una rapina e, dato che vengono tenuti in celle separate, non hanno la possibilità di comunicare. I vantaggi, che in questo caso corrispondono a condanne penali, non dipenderanno solo dalle risposte che ciascuno fornirà individualmente, ma dal quadro congiunto che ne deriverà. Se A confessa e B no, A prenderà solo 1 anno di prigione (vedi la tavola dei profitti di A), ma B sarà condannato a 10 anni (vedi tavola dei profitti di B). Se invece A non parla, ma B sì, le sentenze saranno le stesse, solo su fronti rovesciati. Se confessano entrambi, prenderanno 4 anni ciascuno, mentre se nessuno canta e ciascuno sostiene la propria innocenza se la fileranno liberi come l'aria!

A		B	
		confessa	non confessa
A	confessa	+ 4	+ 1
	non confessa	+ 10	0

B		B	
		confessa	non confessa
A	confessa	+ 4	+ 10
	non confessa	+ 1	0

Se i prigionieri potessero cooperare, sceglierebbero la linea di condotta ottimale, cioè quella di non confessare: entrambi sarebbero accontentati!

idea chiave

La matematica che accontenta tutti

48 La relatività

Il moto di un corpo può essere misurato solo in riferimento ad altri corpi. Se due automobili procedono affiancate su una strada a due corsie alla stessa velocità di 70 chilometri all'ora (km/h) e nella stessa direzione, la loro velocità relativa è zero, anche se rispetto al suolo si stanno muovendo entrambe a 70 km/h. Non solo: rispetto a un veicolo che viaggia a 70 km/h nella carreggiata opposta, la velocità relativa di ciascuna è pari a 140 km/h. Ma questo semplice modo di vedere le cose è stato sconvolto dalla teoria della relatività ristretta.

La teoria della relatività ristretta fu elaborata dal fisico olandese Hendrik Lorentz alla fine del XIX secolo, ma fu Albert Einstein a darne la versione definitiva nel 1905. Il celebre articolo di Einstein sulla relatività ristretta rivoluzionò lo studio del moto dei corpi, riducendo a un semplice caso particolare la teoria classica di Newton, che a suo tempo aveva costituito un risultato straordinario.

Come ai tempi di Galileo Per spiegare la teoria della relatività possiamo servirci degli esempi utilizzati proprio dal suo autore: infatti, Einstein adorava parlare di treni e ferrovie, e ne faceva ampio uso nei suoi ragionamenti. Nel nostro caso, Ugo Diamanti si trova a bordo di un treno che viaggia a 60 km/h. Dal suo posto in fondo al treno si incammina con una velocità di 2 km/h verso il vagone ristorante. Quindi la velocità di Ugo rispetto al suolo sarà 62 km/h, mentre nel percorso di ritorno verso il suo posto diventerà 58 km/h, visto che si sta muovendo in direzione opposta a quella di marcia del treno. Questo è il succo della teoria di Newton: la velocità è un concetto relativo ed è la direzione del moto di Ugo a stabilire se si deve effettuare una somma o una sottrazione.

Dato che ogni moto è relativo, si introduce il concetto di «sistema di riferimento», ovvero di un punto di vista rispetto al quale misurarlo. Se prendiamo in esame il moto unidimensionale del treno che corre lungo un binario dritto, possiamo pensare a un sistema di riferimento fisso posto all'interno di una stazione, nonché a una distanza x e a un tempo t misurati rispetto a tale sistema di riferimento. La

linea del tempo

circa 1632 Galileo elabora le «trasformazioni di Galileo» per la caduta dei corpi

1676 Römer ricava la velocità della luce in base alle osservazioni compiute sulle lune di Giove

1687 I *Principia* di Newton descrivono le leggi classiche del moto

posizione zero è individuata da un punto segnato sulla pensilina e il tempo zero dall'ora indicata dall'orologio della stazione, perciò le coordinate spazio/tempo relative al sistema di riferimento della stazione sono (x, t).

Ma anche a bordo del treno esiste un sistema di riferimento. Misurando lo spazio percorso a partire dalla coda del treno e il tempo in base all'orologio di Ugo, si ottiene un altro insieme di coordinate $(\bar{x}, \bar{t})$. Due sistemi di coordinate possono essere sincronizzati: quando il treno oltrepassa il punto segnato sulla pensilina, lo spazio corrisponde a $x = 0$ e il tempo indicato dall'orologio della stazione è $t = 0$. Se in questo momento anche Ugo pone $\bar{x} = 0$ e registra il suo orologio su $\bar{t} = 0$, si stabilisce una relazione fra i due sistemi di coordinate.

Ugo si incammina verso il vagone ristorante nell'istante in cui il treno passa dalla stazione. Proviamo allora a calcolare la sua distanza dalla stazione dopo cinque minuti. Sappiamo che il treno sta viaggiando a 1 km al minuto, perciò in questo lasso di tempo ha percorso 5 km e Ugo ha camminato per $\bar{x} = 10/60$ km (ottenuti moltiplicando la velocità di 2 km/h per il tempo 5/60); quindi la distanza totale (x) percorsa da Ugo è $(5 + 10/60)$ km. La relazione tra x e $\bar{x}$ è data pertanto da $x = \bar{x} + v \times t$ (dove $v = 60$ km/h). Da questa equazione si può ricavare la distanza percorsa rispetto al sistema di riferimento del treno:

$$\bar{x} = x - v \times t$$

Nella teoria classica di Newton il tempo è concepito come un flusso unidimensionale che va dal passato al futuro e ha lo stesso valore universale per tutti, indipendente dallo spazio. Dato che è una grandezza assoluta, il tempo di Ugo sul treno coincide con quello del capostazione sulla pensilina, perciò vale

$$\bar{t} = t$$

Queste due formule per $\bar{x}$ e $\bar{t}$, che Galileo fu il primo a ricavare, sono note come trasformazioni di Galileo, proprio perché trasformano le grandezze da un sistema di riferimento a un altro. Secondo la teoria classica di Newton, la velocità della luce dovrebbe obbedire alle trasformazioni galileiane per $\bar{x}$ e $\bar{t}$.

Già prima del 1700 era noto che la luce era dotata di una velocità propria: il suo valore approssimato venne misurato nel 1676 dall'astronomo danese Ole Römer. Nel 1881 Albert Michelson ottenne una misura più accurata della velocità della luce, pari a 299 821 km/s, rendendosi conto inoltre che la trasmissione della luce

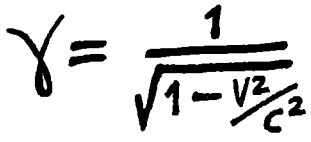
1881 Michelson misura con grande accuratezza la velocità della luce

1887 Prima versione delle trasformazioni di Lorentz

1905 Einstein pubblica *Sull'elettrodinamica dei corpi in movimento*, l'articolo che descrive la relatività ristretta

1915 Einstein pubblica *Le equazioni di campo per la gravitazione*, presentando la relatività generale

seguiva leggi del tutto diverse da quelle della trasmissione del suono. Michelson scoprì che, a differenza di quanto accade alla velocità dell'osservatore sul treno in movimento, la direzione di un fascio di luce non provocava alcun effetto sulla velocità della luce. Questo risultato paradossale richiedeva una spiegazione.



Il fattore di Lorentz

La teoria della relatività ristretta Lorentz elaborò le equazioni matematiche che regolano il collegamento fra spazio e tempo in un sistema di riferimento che si muove a velocità costante v rispetto a un altro. Si tratta di trasformazioni molto simili a quelle che abbiamo già ricavato, che contengono però un fattore (detto «fattore di Lorentz») che dipende da v e dalla velocità della luce, c.

Entra in gioco Einstein I risultati ottenuti da Michelson sulla velocità della luce furono assunti da Einstein come postulato:

la velocità della luce è uguale per tutti gli osservatori e non dipende dalla direzione.

Se, mentre il suo treno sfrecciava attraverso la stazione, Ugo Diamanti avesse acceso una torcia e avesse rivolto il fascio di luce lungo la carrozza in direzione del moto del treno, la velocità della luce della torcia sarebbe risultata pari a c. Il postulato di Einstein afferma che la velocità del fascio di luce misurata dal punto di osservazione del capotreno sul binario sarà ugualmente c, e non c + 60 km/h. Einstein postulò anche un secondo principio:

i sistemi di riferimento si muovono a velocità costante gli uni rispetto agli altri.

Uno dei motivi per cui l'articolo di Einstein del 1905 fu considerato geniale era l'eleganza matematica del suo approccio. Le onde sonore si propagano sotto forma di vibrazioni delle molecole del mezzo attraversato dal suono. Molti fisici avevano ipotizzato che anche per la luce fosse necessario un mezzo di propagazione e, sebbene nessuno ne conoscesse la natura, gli fu comunque attribuito un nome: fu chiamato *etere luminifero*.

Einstein non fu costretto a postulare l'esistenza dell'etere per giustificare la trasmissione della luce, ma si limitò a ricavare le trasformazioni di Lorentz a partire dai due semplici principi di relatività, procedendo così allo sviluppo della teoria completa. In particolare, dimostrò che l'energia E di una particella è determinata dall'equazione $E = \gamma \times mc^2$. Per un corpo a riposo (quando $v = 0$ e quindi $\gamma = 1$) si ottiene la famosa equazione che dimostra l'equivalenza di massa ed energia:

$$E = mc^2$$

Lorentz e Einstein furono entrambi candidati al premio Nobel nel 1912. Lorentz lo aveva già vinto nel 1902, ma Einstein dovette aspettare fino al 1921, quando lo

ottenne per le ricerche sull'effetto fotoelettrico (anch'esse pubblicate nel 1905). Fu un anno cruciale per l'oscuro impiegato dell'ufficio brevetti di Berna.

Einstein contro Newton Se ci limitiamo a osservare il movimento lento dei treni, le differenze tra la teoria della relatività di Einstein e la teoria newtoniana classica risultano del tutto trascurabili. In questo caso la velocità relativa v è talmente piccola rispetto alla velocità della luce che il fattore di Lorentz assume un valore circa uguale a 1 e le equazioni di Lorentz sono virtualmente coincidenti con le trasformazioni galileiane classiche. Dunque, per velocità piccole, Einstein e Newton si sarebbero trovati d'accordo. Anche il velocissimo treno francese TGV è ben lontano dall'aver raggiunto velocità confrontabili a quella della luce e l'ingegneria ferroviaria dovrà percorrere una lunga strada prima che le teorie di Newton possano essere abbandonate in favore di quelle di Einstein. Saranno i viaggi spaziali a obbligarci a seguire Einstein.

La teoria della relatività generale Nel 1925 Einstein pubblicò una teoria generale che si applica a sistemi di riferimento che possono accelerare l'uno rispetto all'altro e che collega gli effetti dell'accelerazione con quelli della gravità.

Grazie a questa teoria generale, Einstein fu in grado di prevedere fenomeni fisici come la deviazione dei fasci luminosi da parte dei campi gravitazionali dei corpi di grande massa (ad esempio, il Sole). La sua teoria riuscì inoltre a spiegare il moto di precessione dell'asse di rotazione di Mercurio, che nel quadro della teoria newtoniana della gravitazione non risultava attribuibile all'effetto delle forze esercitate su Mercurio da parte degli altri pianeti. Era un problema che assillava gli astronomi fin dagli anni '40 dell'Ottocento.

Il sistema di riferimento appropriato per la teoria della relatività generale possiede quattro dimensioni spazio-temporali. Lo spazio euclideo è piatto, cioè ha curvatura zero, ma la geometria dello spazio-tempo quadridimensionale di Einstein (detta anche «geometria di Riemann») è curva. Per spiegare l'attrazione reciproca fra i corpi non si può più invocare la forza di gravità newtoniana; nell'ambito della teoria della relatività generale di Einstein la spiegazione va ricercata nella curvatura dello spazio-tempo. Nel 1925 Einstein ha dato inizio a una nuova rivoluzione scientifica.

idea chiave

La velocità della luce è assoluta

49 L'ultimo teorema di Fermat

È senz'altro possibile prendere due numeri interi, sommarne i quadrati e ottenere il quadrato di un altro numero intero, come dimostra l'esempio $5^2 + 12^2 = 13^2$, ma è altrettanto possibile sommare due cubi e trovarne un terzo? E cosa succede per potenze maggiori? La risposta, davvero sorprendente, è che in questi casi è impossibile. L'ultimo teorema di Fermat afferma che, dati quattro numeri interi x, y, z e n, l'equazione $x^n + y^n = z^n$ non ammette soluzione quando n è maggiore di 2. Fermat si vantò di aver trovato una splendida dimostrazione, dando filo da torcere alle successive generazioni di matematici. Fra questi c'era anche un ragazzino di dieci anni, che venne a sapere di questa caccia al tesoro matematica curiosando nella biblioteca della sua città.

L'ultimo teorema di Fermat riguarda le equazioni diofantee: queste equazioni, fra le più ardue da risolvere, ammettono come soluzioni dei numeri interi e prendono il nome da Diofanto di Alessandria, la cui *Arithmetica* ha costituito una pietra miliare della teoria dei numeri. Pierre de Fermat era un avvocato e funzionario governativo francese che visse a Tolosa nel XVII secolo. Era un matematico versatile e acquistò un'ottima reputazione nel campo della teoria dei numeri soprattutto per l'enunciato del grande teorema che costituì il suo ultimo contributo alla matematica. Pare che Fermat lo avesse dimostrato, o quantomeno pensasse di averlo fatto, perché annotò sulla sua copia dell'*Arithmetica* di Diofanto: «Ho scoperto una dimostrazione davvero meravigliosa, ma i margini del libro sono troppo stretti per contenerla».

Fermat riuscì a risolvere molti problemi di estrema difficoltà, ma è probabile che il suo ultimo teorema non fosse tra quelli. Si tratta di un teorema che ha impegnato intere schiere di matematici per trecento anni e solo di recente ne è stata data una

linea del tempo

1665 Fermat muore senza lasciare traccia della sua meravigliosa dimostrazione

1753 Eulero dimostra il caso $n = 3$

1825 Legendre e Dirichlet dimostrano indipendentemente il caso $n = 5$

1839 Lamé prova il caso $n = 7$

dimostrazione; però tale dimostrazione è tale da non poter essere contenuta in alcun margine e il fatto che sia stata generata ricorrendo a tecniche moderne suscita profondi dubbi sull'annuncio di Fermat.

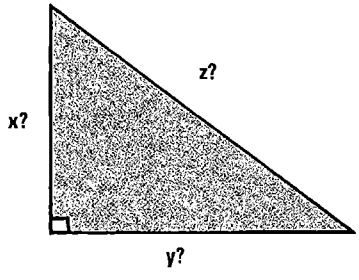
L'equazione $x + y = z$ Come si risolve un'equazione in tre variabili x, y e z? In genere un'equazione contiene un'unica incognita x, ma qui ce ne sono addirittura tre. Per la verità, è proprio questo a facilitarne la risoluzione. Infatti, possiamo scegliere i valori di x e y che più ci piacciono e sommarli per produrre z, ottenendo così una terna di valori che forma una soluzione valida. Niente di più semplice.

Se, ad esempio, si scelgono $x = 3$ e $y = 7$, è immediato verificare che i valori $x = 3$, $y = 7$ e $z = 10$ costituiscono una soluzione dell'equazione. Però esistono anche terne di valori di x, y e z che non soddisfano l'equazione: infatti $x = 3$, $y = 7$ e $z = 9$ non sono una soluzione, perché il primo membro dell'equazione $(x + y)$ non risulta uguale al secondo (z).

L'equazione $x^2 + y^2 = z^2$ Ora passiamo ai quadrati. Il quadrato di un numero è uguale al numero moltiplicato per se stesso e si indica con x^2; se $x = 3$, il quadrato è $x^2 = 3 \times 3 = 9$. In questo caso l'equazione di riferimento non è più $x + y = z$, bensì

$$x^2 + y^2 = z^2$$

È possibile risolverla usando il metodo visto sopra, cioè calcolando z a partire da valori scelti per x e per y? Ponendo $x = 3$ e $y = 7$, ad esempio, il primo membro dell'equazione risulta $3^2 + 7^2$, cioè $9 + 49 = 58$. Perciò z dovrà essere uguale alla radice quadrata di 58 ($z = \sqrt{58}$), che corrisponde approssimativamente a 7,6158. Possiamo senza dubbio affermare che $x = 3$, $y = 7$ e $z = \sqrt{5}$ formano una soluzione di $x^2 + y^2 = z^2$, ma la caratteristica delle equazioni diofantee è quella di ammettere come soluzioni solo numeri interi. Dato che $\sqrt{58}$ non è un numero intero, la terna $x = 3$, $y = 7$ e $z = \sqrt{58}$ non fa al caso nostro.



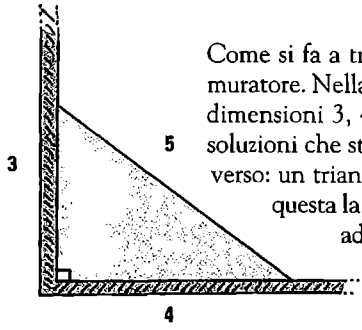
L'equazione $x^2 + y^2 = z^2$ è legata ai triangoli. Se consideriamo x, y e z come le misure dei lati di un triangolo rettangolo, l'equazione è soddisfatta. Inversamente, se x, y e z rispettano l'equazione il triangolo è rettangolo. Visto il collegamento con il teorema di Pitagora, le soluzioni x, y e z sono dette «terne pitagoriche».

1843 Kummer dichiara di aver dimostrato il teorema, ma Dirichlet coglie un errore nel procedimento

1907 Von Lindemann pubblica una dimostrazione, ma deve ammettere di aver commesso un errore

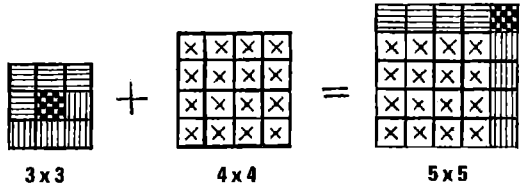
1908 Wolfskehl offre un premio per chi troverà la soluzione entro 100 anni

1994 Finalmente Wiles dimostra il teorema



Come si fa a trovare una terna pitagorica? Qui può venirci in aiuto un semplice muratore. Nella sua cassetta degli attrezzi non mancherà certamente un triangolo di dimensioni 3, 4, 5: i valori $x = 3$, $y = 4$ e $z = 5$ costituiscono una delle terne di soluzioni che stiamo cercando, perché $3^2 + 4^2 = 9 + 16 = 5^2$. Ma è vero anche l'inverso: un triangolo con dimensioni 3, 4 e 5 deve avere un angolo retto. È proprio questa la ragione matematica per cui i muratori lo utilizzano per costruire muri ad angolo retto.

In questo caso possiamo «aprire» un quadrato 3×3 e sistemare i suoi pezzi attorno a un quadrato 4×4 in modo da ottenerne uno 5×5.



L'equazione $x^2 + y^2 = z^2$ ammette come soluzioni altri numeri interi: per esempio, $x = 5$, $y = 12$ e $z = 13$ costituiscono una soluzione valida, perché $5^2 + 12^2 = 13^2$. Esistono infinite soluzioni di questo tipo. La soluzione più famosa è quella del muratore ($x = 3$, $y = 4$ e $z = 5$), perché è costituita dai numeri più piccoli, oltre ad essere l'unica composta da tre numeri interi consecutivi. In molte soluzioni due dei tre numeri interi sono consecutivi, come $x = 20$, $y = 21$ e $z = 29$, oppure $x = 9$, $y = 40$ e $z = 41$, ma in nessun'altra lo sono tutti e tre.

Dal banchetto al digiuno A questo punto, passare da $x^2 + y^2 = z^2$ a $x^3 + y^3 = z^3$ sembra un gioco da ragazzi. Tornando all'esempio in cui si scomponeva un quadrato per disporlo attorno a un altro e ottenerne un terzo, ci si può domandare se sia possibile ricorrere allo stesso trucchetto anche nel caso di un cubo. È possibile smontarlo e rimontarlo attorno a un altro per ottenere un terzo cubo? La risposta è no. Non si può. L'equazione $x^2 + y^2 = z^2$ ammette infinite soluzioni, ma per $x^3 + y^3 = z^3$ Fermat non fu capace di trovarne nemmeno una che fosse costituita da numeri interi. Proseguire risultò ancora più difficile, tanto che Eulero, non trovando alcun risultato, enunciò l'ultimo teorema:

non esiste alcuna soluzione costituita da numeri interi per l'equazione $x^n + y^n = z^n$ *per qualsiasi valore di* n *maggiore di 2.*

Per trovare una soluzione si potrebbe partire da valori piccoli di n e andare ad aumentare. In effetti, fu questo l'approccio seguito da Fermat che, con tutta probabilità, doveva aver ricavato una dimostrazione per il caso $n = 4$, più semplice di $n = 3$. Nel corso del XVIII e del XIX secolo, Eulero completò il caso $n = 3$, Adrien-Marie Legendre dimostrò il caso $n = 5$ e Gabriel Lamé quello $n = 7$. In un primo momento Lamé credette di aver trovato una dimostrazione del teorema generale, ma purtroppo si sbagliava.

Uno dei contributi più importanti fu fornito da Ernst Kummer, che nel 1843 presentò un manoscritto in cui sosteneva di aver trovato una dimostrazione generale del teorema; tuttavia Dirichlet individuò una falla nel suo ragionamento. L'Accademia francese delle Scienze aveva offerto un premio di 3000 franchi per chi avesse fornito una dimostrazione corretta, ma alla fine dovette concederlo a Kummer per il suo nobile tentativo. Kummer aveva dimostrato il teorema per tutti i numeri primi minori di 100 (e per altri valori), fatta però eccezione per i numeri primi irregolari 37, 59 e 67. Ad esempio, non riuscì a dimostrare che non esistono numeri interi che rispettano l'equazione $x^{67} + y^{67} = z^{67}$. Per quanto fosse fallito, il suo tentativo di dimostrare la forma generale del teorema spianò la strada a efficaci tecniche di algebra astratta e ciò fornì alla matematica un contributo forse più importante della risoluzione stessa del problema.

Nel 1907 Ferdinand von Lindemann, il matematico celebre per aver dimostrato l'impossibilità della quadratura del cerchio (vedi capitolo 05), si vantò di aver dimostrato il teorema, ma fu colto in fallo. Nel 1908 Paul Wolfskehl lasciò in eredità 100 000 marchi al primo che avesse trovato una dimostrazione, mantenendo il premio disponibile per 100 anni. Nel corso degli anni sono pervenute oltre 5000 dimostrazioni, ma sono risultate tutte false e quindi sono state puntualmente rispedite allo speranzoso mittente.

La dimostrazione Sebbene il collegamento al teorema di Pitagora si applichi solo nel caso $n = 2$, è stato proprio il legame con la geometria a condurre alla prova conclusiva. I due matematici giapponesi Yutaka Taniyama e Goro Shimura proposero una congettura basata sul collegamento con la teoria delle curve ellittiche. Nel 1993 Andrew Wiles tenne una lezione a Cambridge su questa teoria e vi incluse la propria dimostrazione per il teorema di Fermat: purtroppo era sbagliata.

Il matematico francese André Weil (dal nome curiosamente simile a quello di Wiles) liquidò questi tentativi sostenendo che dimostrare un teorema è come scalare l'Everest: se si manca la vetta di 100 metri non ci si può vantare di averlo conquistato. La tensione saliva. Wiles si ritirò in isolamento a lavorare incessantemente sul problema, ma molti ritenevano che avrebbe finito per ingrossare le file di quelli che ci erano arrivati soltanto vicini.

Con l'aiuto dei colleghi, tuttavia, Wiles riuscì ad eliminare l'errore e a sostituirlo con un ragionamento corretto. Questa volta riuscì a convincere gli esperti e a provare il teorema. La sua dimostrazione fu pubblicata nel 1995, giusto in tempo per consentirgli di reclamare il premio Wolfskehl e diventare una celebrità matematica. Ne aveva fatta di strada, quel ragazzino di dieci anni che tanti anni prima se ne stava seduto nella biblioteca pubblica di Cambridge incuriosito dal problema...

idea chiave

La soluzione di un caso difficile

50 L'ipotesi di Riemann

L'ipotesi di Riemann rappresenta una delle sfide più ardue della matematica pura. L'assalto alla congettura di Poincaré e all'ultimo teorema di Fermat è riuscito, quello all'ipotesi di Riemann ancora no. Quando questa sarà risolta, in un senso o nell'altro, le sfuggenti questioni sulla distribuzione dei numeri primi troveranno risposta e si aprirà la strada a una nuova serie di problemi su cui i matematici dovranno riflettere.

La storia inizia con la somma di frazioni del tipo

$$1+\frac{1}{2}+\frac{1}{3}$$

In questo caso, il risultato è 1 + 5/6 (circa 1,83). Ma cosa succede se continuiamo ad aggiungere frazioni via via più piccole, ad esempio fermandoci a dieci?

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}$$

Basta una semplice calcolatrice per verificare che la somma di queste frazioni, espressa in forma decimale, è uguale a circa 2,9. La tabella a margine mostra come aumenta il totale per aggiunta di ulteriori termini.

Numero di termini	Totale (approssimato)
1	1
10	2,9
100	5,2
1 000	7,5
10 000	9,8
100 000	12,1
1 000 000	14,4
1 000 000 000	21,3

La serie numerica

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\ldots$$

prende il nome di «serie armonica». L'aggettivo deriva dalla teoria dei pitagorici secondo la quale le note musicali essenziali per l'armonia si ottenevano per suddivisione successiva di una corda a metà, poi in tre parti, poi in quattro.

linea del tempo

1854 Riemann inizia a lavorare sulla funzione zeta $\zeta(s)$

1859 Riemann dimostra che gli zeri non banali di $\zeta(s)$ si trovano in una striscia critica e formula la sua congettura

1896 De la Vallée-Poussin e Hadamard dimostrano che tutti gli zeri non banali cadono all'interno della striscia critica di Riemann

Le frazioni che costituiscono la serie armonica diventano progressivamente più piccole, ma cosa succede al totale? Aumenta indefinitamente, oppure prima o poi si trova davanti una barriera, un limite che non può superare? Per rispondere a questa domanda il trucco consiste nel raggruppare i termini, raddoppiando via via che si procede il numero di frazioni associate. Ad esempio, sommando i primi 8 termini (e ricordando che $8 = 2 \times 2 \times 2 = 2^3$) si ottiene

$$S_{2^3} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)$$

(dove S sta per «somma»). Dato che 1/3 è maggiore di 1/4, 1/5 è maggiore di 1/8 e così via, il tutto è maggiore di

$$1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}$$

Perciò possiamo dire che

$$S_{2^3} > 1 + \frac{3}{2}$$

e, più in generale, che

$$S_{2^k} > 1 + \frac{k}{2}$$

Per $k = 20$, 2^k vale 1 048 576 (più di un milione di termini), mentre la somma della serie è appena maggiore di 11 (vedi tabella). È una crescita terribilmente lenta, ma è sempre possibile scegliere un valore di k che porti la somma della serie a superare qualunque numero stabilito, per grande che sia: si usa dire che la serie «diverge all'infinito». Ciò non si osserva invece nel caso della serie a termini quadrati

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \ldots$$

Anche in questo caso il procedimento è lo stesso: si tratta di sommare termini via via più piccoli. Stavolta però un limite viene raggiunto, ed è inferiore a 2. La serie converge vertiginosamente a $\pi^2/6 = 1{,}64493\ldots$

In questa ultima serie l'esponente dei termini è pari a 2, mentre nella serie armonica è implicitamente uguale a 1, un valore critico: infatti, se l'esponente aumenta di una quantità minuscola fino a superare di poco 1, la serie converge, ma se dimi-

1900 — Hilbert inserisce l'ipotesi di Riemann nella sua lista dei grandi problemi matematici ancora da risolvere

1914 — Hardy dimostra che esistono infinite soluzioni lungo la retta di Riemann

2004 — Si verifica che i primi mille miliardi di zeri (di soluzioni) si trovano sulla retta critica

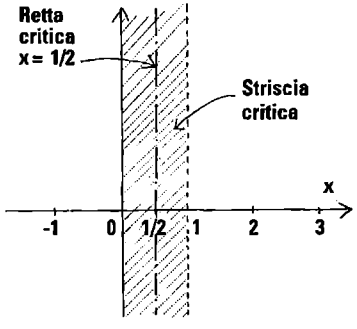
nuisce di una quantità altrettanto minuscola fino a un valore appena minore di 1, la serie diverge. Quindi la serie armonica sta a cavallo del confine tra convergenza e divergenza.

La funzione zeta di Riemann La celebre funzione zeta (ζ) di Riemann in realtà, era già nota a Eulero nel XVIII secolo, ma fu Bernard Riemann a comprenderne appieno l'importanza. La ζ rappresenta la lettera greca zeta, e la funzione si scrive così:

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} +$$

Fra i numerosi valori che sono stati calcolati per questa funzione, il più importante è senz'altro $\zeta(1) = \infty$, dato che $\zeta(1)$ è la serie armonica. Il valore di $\zeta(2)$ è $\pi^2/6$, ovvero il risultato scoperto da Eulero. È stato dimostrato che tutti i valori di $\zeta(s)$ comprendono π quando s è pari, mentre la teoria di $\zeta(s)$ per valori dispari di s risulta molto più complicata. Roger Apéry dimostrò l'importante risultato che $\zeta(3)$ è un numero irrazionale, ma il suo metodo non è applicabile a $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ e così via.

L'ipotesi di Riemann La variabile s della funzione zeta di Riemann rappresenta un numero reale ma, per estensione, può essere considerata una variabile complessa: ciò consente di applicare alla funzione le potenti tecniche dell'analisi complessa (vedi capitolo 08).



La funzione zeta di Riemann ha un'infinità di zeri, cioè esistono infiniti valori di s per i quali $\zeta(s) = 0$. In un articolo presentato nel 1859 all'Accademia delle Scienze di Berlino, Riemann dimostrò che tutti gli zeri non banali erano numeri complessi situati all'interno della striscia critica delimitata da $x = 0$ e $x = 1$. Inoltre formulò la sua celebre ipotesi:

tutti gli zeri non banali della funzione zeta di Riemann $\zeta(s)$ stanno sulla retta critica $x = 1/2$ che divide a metà la striscia critica.

Il primo passo concreto verso la conferma di questa ipotesi fu compiuto nel 1896, in maniera indipendente, da Charles de la Vallée-Poussin e Jacques Hadamard, che dimostrarono la necessità che gli zeri cadano all'interno della striscia (cioè che $\zeta(s)$ non si annulla sulle rette $x = 0$ e $x = 1$). Nel 1914 il matematico francese G.H. Hardy dimostrò che sulla retta $x = 1/2$ si trovano infiniti zeri, tuttavia, ciò non impedisce che ne esistano infiniti altri al di là di essa.

I calcoli numerici dicono che tutti gli zeri non banali calcolati fino al 1986 (1 500 000 000 risultati) si trovano sulla retta $x = 1/2$, mentre le elaborazioni più recenti hanno confermato che ciò è vero per i primi 100 miliardi di zeri. Anche se questi

risultati sperimentali fanno pensare che la congettura sia ragionevole, esiste comunque la possibilità che possa rivelarsi falsa. La congettura afferma che *tutti* gli zeri si trovano all'interno della striscia critica, ma ciò resta in attesa di una conferma o di una smentita.

Perché l'ipotesi di Riemann è importante? Esiste un collegamento inaspettato fra la funzione zeta di Riemann $\zeta(s)$ e la teoria dei numeri primi (vedi capitolo 09). I numeri primi sono 2, 3, 5, 7, 11, e così via, cioè sono i numeri divisibili solo per 1 e per se stessi. Utilizzando i numeri primi si può scrivere l'espressione

$$\left(1-\frac{1}{2^s}\right)\times\left(1-\frac{1}{3^s}\right)\times\left(1-\frac{1}{5^s}\right)\times\ldots$$

che è semplicemente una forma alternativa di $\zeta(s)$, la funzione zeta di Riemann. Ciò sembra indicare che una conoscenza più approfondita della funzione zeta di Riemann potrebbe consentirci di far luce sulla distribuzione dei numeri primi, migliorando il nostro livello di comprensione dei fondamenti della matematica.

Nel 1900 David Hilbert presentò i famosi ventitré problemi che i matematici avrebbero dovuto risolvere nel corso del secolo. Riguardo all'ottavo disse: «Se dovessi risvegliarmi dopo un sonno lungo cinquecento anni, la mia prima domanda sarebbe: è stata confermata l'ipotesi di Riemann?».

Hardy usò l'ipotesi di Riemann come assicurazione sulla traversata del Mare del Nord, mentre tornava da una visita estiva all'amico Harald Bohr, in Danimarca. Prima che la nave uscisse dal porto, spedì all'amico una cartolina in cui annunciava di aver dimostrato l'ipotesi di Riemann. Era una scommessa furba, in cui poteva solo vincere: se la nave fosse affondata si sarebbe garantito l'onore postumo di aver risolto quell'importante problema, ma d'altra parte se Dio fosse esistito non avrebbe mai consentito a un ateo come Hardy di assicurarsi un tale onore, e quindi avrebbe fatto in modo che la nave non affondasse.

Il Clay Mathematics Institute mette in palio un premio da un milione di dollari per chi saprà risolvere la questione in modo rigoroso. Ma non è il denaro la forza trainante per i matematici: la maggior parte di loro si accontenterebbe di ottenere, nel contempo, il risultato e un posto di primo piano nel Pantheon dei grandi matematici.

idea chiave
L'ultima sfida

Glossario

Algebra Nata come estensione dell'aritmetica basata sull'utilizzo di lettere al posto dei numeri, col tempo è diventata un metodo generale valido per tutte le branche e le applicazioni della matematica. La parola «algebra» deriva da *al-jabr*, un termine usato in un testo arabo del IX secolo.

Algoritmo «Ricetta matematica» consistente in una sequenza prestabilita di istruzioni per la risoluzione di un problema.

Assi cartesiani (*x*-*y*) Rappresentazione dei punti con una coordinata *x* (asse orizzontale) e una *y* (asse verticale) ideata da René Descartes.

Assioma Affermazione, di cui non si cerca di dimostrare la validità, utilizzata come base di una teoria matematica. I Greci usavano il sinonimo «postulato», nell'accezione di una verità tanto evidente da non richiedere dimostrazione.

Base La base di un sistema di numerazione. Il sistema babilonese era in base 60 (sessagesimale), quello attuale è in base 10 (decimale).

Cardinalità Il numero degli elementi che costituiscono un insieme. La cardinalità dell'insieme $\{a, b, c, d, e\}$ è 5, ma si può attribuire un significato anche alla cardinalità di insiemi infiniti.

Commutativo In algebra la moltiplicazione è commutativa, cioè $a \times b = b \times a$, come nell'aritmetica ordinaria. In molte branche dell'algebra moderna (ad esempio, nell'algebra delle matrici) questa proprietà non vale.

Controesempio Singolo esempio sufficiente a confutare un'affermazione. L'esistenza di un cigno nero costituisce un controesempio dell'affermazione «tutti i cigni sono bianchi».

Corollario Conseguenza secondaria di un teorema.

Corrispondenza biunivoca Natura della relazione per cui ciascun elemento di un insieme corrisponde a uno e un solo elemento di un altro insieme, e viceversa.

Denominatore La parte inferiore di una frazione. Nella frazione 3/7, il numero 7 è il denominatore.

Derivazione Operazione fondamentale del calcolo differenziale che dà luogo alla derivata. Nel caso di un'espressione che descrive come varia la distanza in funzione del tempo, la derivata rappresenta la velocità. La derivata dell'espressione della velocità corrisponde all'accelerazione.

Diagramma di Argand Metodo grafico che consente di rappresentare il piano bidimensionale dei numeri complessi.

Diagramma di Venn Metodo grafico (diagrammi a cerchi) usato nella teoria degli insiemi.

Discreto Termine usato in opposizione a «continuo». I valori discreti sono separati da un certo intervallo, come quello esistente fra i numeri interi 1, 2, 3, 4,...

Distribuzione La gamma di probabilità degli eventi che si verificano in un esperimento o in una certa situazione. Ad esempio, la distribuzione di Poisson fornisce la probabilità che un evento raro si verifichi *x* volte, per ogni valore di *x*.

Divisore Un numero intero che divide esattamente un altro numero intero. Il numero 2 è un divisore di 6 perché $6 \div 2 = 3$. Quindi anche 3 lo è, perché $6 \div 3 = 2$.

Equazione diofantea Equazione a coefficienti e soluzioni intere o frazionarie. Prende il nome dal matematico greco Diofanto di Alessandria (250 circa).

Esponente Notazione aritmetica. Il prodotto di un numero con se stesso, 5×5, si scrive 5^2, dove 2 è l'esponente; l'espressione $5 \times 5 \times 5$ si scrive 5^3, e così via. Questa notazione può essere estesa ad altri casi: ad esempio, $5^{1/2}$ rappresenta la radice quadrata di 5. Viene anche chiamato impropriamente «potenza».

Frazione Un numero intero diviso per un altro, come 3/7.

Frazione unitaria Frazione la cui parte superiore (numeratore) è uguale a 1. Il sistema di numerazione degli antichi Egizi era parzialmente basato sulle frazioni unitarie.

Geometria Disciplina che prende in esame le proprietà di rette, figure e spazi, formalizzata nel III secolo a.C. da Euclide negli *Elementi*. Rispetto al suo ambito tradizionale, oggi si è ampliata fino a permeare tutta la matematica.

Insieme Raggruppamento di oggetti. Un esempio è l'insieme costituito da alcuni oggetti di arredamento: F = {sedia, tavolo, divano, sgabello, armadio}.

Insieme vuoto L'insieme che non contiene alcun elemento, convenzionalmente indicato con il simbolo Ø. È un concetto molto utile nella teoria degli insiemi.

Integrazione Operazione fondamentale del calcolo differenziale utile per la misura delle aree. Si può dimostrare che è l'operazione inversa della derivazione.

Ipotesi Affermazione che attende di essere dimostrata o confutata. In matematica è equivalente a una congettura.

Iterazione Procedimento che consiste nel ripetere più volte di seguito la stessa operazione, a partire da un valore dato. Ad esempio, partendo da 3 e sommando ripetutamente 5 si ottiene la sequenza iterativa 3, 8, 13, 18, 23,...

Lemma Affermazione dimostrata come passo preliminare per la risoluzione di un problema importante.

Massimo comun divisore, MCD Il MCD di due numeri è il numero più grande che li divide entrambi esattamente. Ad esempio, 6 è il MCD di 18 e 84.

Matrice Blocco di numeri o simboli disposti a formare un quadrato o un rettangolo. Le matrici possono essere sommate e moltiplicate, e formano un sistema algebrico.

Numeratore La parte superiore di una frazione. Nella frazione 3/7, il numero 3 è il numeratore.

Numeri immaginari Numeri che contengono l'unità immaginaria $i = \sqrt{-1}$. In combinazione con i numeri ordinari (o «reali») formano i numeri complessi.

Numeri irrazionali Numeri non esprimibili sotto forma di frazione.

Numeri primi gemelli Due numeri primi separati al massimo da un numero. Ad esempio, 11 e 13 sono numeri primi gemelli. Non è noto se ne esistano infiniti.

Numeri razionali Numeri interi o frazioni.

Numero primo Numero intero divisibile solo per 1 e per se stesso. Ad esempio, 7 è un numero primo, ma 6 non lo è (perché $6 \div 2 = 3$). Convenzionalmente, la successione dei numeri primi viene fatta partire dal 2.

Numero quadrato (quadrato perfetto) Il risultato del prodotto di un numero per se stesso. Il 9 è un quadrato perché $9 = 3 \times 3$. I quadrati sono 1, 4, 9, 16, 25, 36, 49, 64,...

Numero trascendente Numero che non può essere soluzione di un'equazione algebrica: π è un numero trascendente.

Poliedro Figura solida dotata di molte facce. Ad esempio, un tetraedro ha quattro facce triangolari e un cubo ha sei facce quadrate.

Quadratura del cerchio Problema consistente nel costruire un quadrato che ha la stessa area di un cerchio dato, usando solo una riga per tracciare delle rette e un compasso per disegnare delle circonferenze. È impossibile da realizzare.

Quaternioni Numeri immaginari a quattro dimensioni scoperti da W.R. Hamilton.

Radice quadrata Numero che, moltiplicato per se stesso, è uguale a un numero dato. Ad esempio, 3 è la radice quadrata di 9, perché $3 \times 3 = 9$.

Resto Ciò che rimane dalla divisione di un numero intero per un altro dello stesso tipo. Ad esempio, dividendo 17 per 3 si ottiene 5, con il resto di 2.

Serie Somma (possibilmente infinita) di numeri o simboli.

Sezione conica Il termine con cui si indica la famiglia di curve classiche che comprende cerchi, rette, ellissi, parabole e iperboli. Ciascuna di queste curve è ottenibile come sezione di un doppio cono.

Simmetria La regolarità presentata da una figura: se può essere ruotata in modo da tornare a coincidere con la sua immagine originaria, ha una simmetria rotazionale; se il suo riflesso coincide con l'immagine di partenza, ha una simmetria speculare.

Sistema di numerazione binario Sistema di numerazione basato su due soli simboli, 0 e 1, fondamentale in ambito informatico.

Sistema esadecimale Sistema di numerazione in base 16 che utilizza i simboli 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F. È molto usato in informatica.

Sistema posizionale Sistema di numerazione in cui il valore dei numeri dipende dalla posizione delle cifre che li compongono. Nel numero 73, il 7 corrisponde a «7 decine» e il 3 a «3 unità».

Soluzione ottimale Molti problemi richiedono di trovare la soluzione ottimale, cioè la migliore. Può trattarsi di una soluzione che minimizza i costi o massimizza i profitti, come avviene nella programmazione lineare.

Successione Sequenza (possibilmente infinita) di numeri o simboli.

Teorema di Pitagora Se le lunghezze dei lati di un triangolo rettangolo sono x, y e z, vale l'equazione $x^2 + y^2 = z^2$, dove z è la lunghezza del lato più lungo (ipotenusa), opposto all'angolo retto.

Teorema Termine che indica un fatto appurato di una certa importanza.

Teoria del caos Teoria dei sistemi dinamici che appaiono casuali, ma possiedono una regolarità di fondo.

Indice analitico

I numeri di pagina in neretto si riferiscono alla corrispondente voce del glossario.

A
Abel Niels, 58, 154.
Alberi, 118-119.
Algebra, 8, 56-59, **204**;
  curve algebriche, 90;
  delle matrici, 58, 156-159, **205**;
  e genetica, 148;
  e topologia, 95;
  gruppi astratti, 154;
  triangolo di Tartaglia, 53;
  ultimo teorema di Fermat, 196-199.
Algoritmi, 60-63, 183-184, 187, **204**.
Angoli:
  misura degli, 8;
  postulati di Euclide, 109;
  trisezione dell'angolo, 80-81.
Archimede di Siracusa, 20-21, 88, 92.
Area:
  del cerchio, 21;
  di un poligono, 113;
  di un triangolo, 87;
  massimo comun divisore, 61;
  sottesa da una curva, 79.
Argand, piano di, 33-35, **204**.
Aristotele, 28, 64-65.
Assiomi, 59, 72, 74, 109, 127, 154-155, **204**.

B
Base dieci, 8, 10, 11, **204**;
  sessanta, 8, **204**.
Bayes, teoria di, 128-131.
Benford Frank, 136-138.
Bernoulli Jacob, 24, 89.
Bourbaki Nicholas, 72-73.
Brianchon, teorema di, 112,115.
Buckminsterfullerene (fullerene), 93.

C
Calcolo combinatorio, 164-167, 187.
Calcolo differenziale, 76-79.
Cantor Georg, 28-31, 72-73, 75, 96.
Caos, teoria del, 104-107, **205**.
Cardinalità, 29-30, 74-75, **204**.
Carta, formato, 48-49.
Cayley Arthur, 35, 98, 100, 102, 119, 121-122, 153-154, 156.
Cerchio, 109, 115;
  dei nove punti, 85;
  quadratura del, 22, 80-82, 199, 205.
Chiocciola di Pascal, 90.
Ciambella, 92, 94-95,122.
Circonferenza, 88-91, 98;
  pi greco (π), 20-21.
Codice Morse, 160, 162.
Codici, 160-163.
Codificazione, 108, 160-163.
Coincidenze, 135.
Colore:
  problema dei quattro colori, 120-123;
  degli occhi (ereditarietà), 148-149.
Commesso viaggiatore, 184-187.
Commutativo, **204**.
Contare, 4, 28-30, 119, 132, 138, 164-167.
Controesempio, 65, 69, 123, **204**.
Correlazione fra i dati, 144-147.
Corrispondenza biunivoca, 28-30, **204**.
Corollario, **204**.
Costruzioni geometriche, 80-83.
Curve, 22, 88-91, 102, 116, 170, 199;
  algebriche 90;
  calcolo differenziale 79;
  catenaria 90;
  di distribuzione normale, 27, 140-143;
  fiocco di neve di Koch, 100, 102-103.

D
De Morgan Augustus, 66, 69-70, 73, 120-121.
Denominatore, 12-14, **204**.
Derivazione, 76-79, **204**.
Descartes René, 32, 41-43, 68, 89-90.
Dilemma del prigioniero, 189, 191.
Dimensioni, 29, 35, 58, 93, 95-99, 156, 170-172, 195;
  frazionarie, 99, 103.
Dimostrazione, 17, 19, 22, 58, 68-72, 80, 85, 100, 123.
Discreto, **204**.
Distribuzioni di probabilità, 136-143, 150, **204**.
Divisione, 14, 34;
  algoritmo di Euclide, 61-63;
  per zero, 5-7.
Divisore, **204**.
DNA, 90, 149, 151.
Dodecaedro, 93.
Dudeney Henry, 176, 179.
Dürer Albrecht, 169.

E
e, 24-27.
Edilizia:
  sezione aurea, 50-51;
  strutture triangolari, 87.
Effetto farfalla, 104-105, 107.
Egizi, 8, 15, 165.
Einstein Albert, 75, 77, 97, 109, 111, 113, 177, 192-195.
Ellisse, 88-89, 115.
Equazioni, 56-58, 62-63, 90, 107, 196-197;
  diofantee, 62-63, 196-197, 204;
  lineari, 57, 180;
  quadratiche, 56-58.
Esponente, **204**.
Euclide di Alessandria:
  algoritmo di, 60-63;
  costruzione dei poligoni, 82-83;
  gli *Elementi*, 16, 29, 36, 38, 40, 48, 60, 62, 68, 82, 85, 93, 108-109;
  numeri perfetti, 40, 43;
  numeri primi, 38;
  postulati, 108-111;
  teoremi sui triangoli, 84-85.
Euler Leonhard (Eulero), 32, 72, 160;
  e, 24-27;
  formula di, 92-94, 163;
  grafi, 116-117;
  numeri perfetti, 42-43;
  pi greco (π), 21-22;
  quadrati di quadrati, 169-170;
  quadrato latino, 172-174;
  retta di, 85-86;
  ultimo teorema di Fermat, 196, 198, 202.

F
Fermat Pierre de:
  numeri primi, 39, 83;
  probabilità, 125;
  ultimo teorema di, 170, 196-200.
Fibonacci Leonardo (Leonardo da Pisa), 4-5, 13, 44-48, 54, 61.
Fibonacci, successione di, 44-47, 50, 54.
Fontana Niccolò (Tartaglia), 53, 57.
Franklin Benjamin, 170.
Frattali, 54, 99,100-103,107.
Frazioni, 8-9, 12-15, 75, 82, **204**;
  conteggio delle, 30-31;
  conversione in numeri decimali, 14-15;
  ipotesi di Riemann, 200-203;
  radici quadrate, 18-19;
  unitarie, **204**.
Funzione zeta, 200, 202-203.

G
Galilei Galileo, 77-78, 192-193.
Galton Francis, 105, 121, 144, 146-147.
Gauss Carl Friedrich, 20, 33, 37, 39, 81, 83, 110-111, 140, 144, 180.
Gelfond, costante di, 26.
Genetica, 148-151.
Geometria, **204**;
  dimensioni, 96-99;
  discreta, 112-115;
  ellittica, 111;
  Euclidea, 108-110, 114;
  iperbolica, 108, 110;
  postulato delle parallele, 75, 108-111;
  proiettiva, 113-115;
  topologia, 92-95, 99.
Geroglifici, 8, 15, 165.
Gioco d'azzardo, 125-127, 143.
Giulio Cesare, 160, 162.
Godel, teorema di, 73-75.
Goldbach, congettura di, 37-38.
Grafi, 116-119, 123, 167;
  non-planari, 118.
Grassmann Hermann, 59, 98.
Gravità (gravitazione), 77-79, 193, 195.
Gruppi, teoria dei, 59, 109, 152-155, 167.

H
Halmos Paul, 70, 123.
Hamilton William Rowan (Sir), 33-35, 58-59.
Hardy Godfrey H., 53, 149-151, 168, 201-203.
Heawood Percy, 120-122.
Hilbert David, 75, 99, 201, 203.

I
i, 33-34.
Icosaedro, 93.
Induzione matematica, 69-71.
Infinito (∞), 6, 28-31.
Insieme vuoto, 7, **204**.
Insiemi, 7, 28-29, 67, 72-75, **204**.
Integrazione, 7, 79, **204**.
Interesse, 24-25, 176-179;
  composto, 24, 176-179.
Iperbole, 88, 115.
Iperspazio, 97-98.
Ipotesi, **204**.
Ipotesi del continuo, 73, 75.
Iterazione, 101, **205**.

J
Jordan Camille, 91.

K
Kirkman Thomas (Rev)., 113-115, 165, 167.

Klein, bottiglia di, 94-95.
Koch, fiocco di neve di, 100, 102-103.

L
Lagrange Joseph, 39, 154.
Laplace Pierre-Simon de (Marquis), 104, 125, 140.
Legendre Adrien Marie, 110, 144, 196, 198.
Leibniz Gottfried, 53, 55, 76, 78-79.
Lemma, **205**;
della stretta di mano, 117-118.
Lemniscata, 90.
Leonardo da Vinci, 50-51, 89, 96.
Logaritmo, 17, 24, 82.
Logica, 19, 64-68, 71, 188;
fuzzy, 65, 67.
Lo Shu, quadrato di, 168-169.
Luce, velocità della, 192-195.

M
Mandelbrot, insieme di, 100-102.
Mappe a quattro colori, 120-123.
Massimo comun divisore (MCD), 61-63, **205**.
Massimo fattore comune, 61.
Matematica finanziaria, 176-179;
interesse, 176-179;
problema della dieta, 180-183.
Matrici, 156-159, **205**.
Meccanismo a tre aste, 90.
Media, 141-142, 146-147.
Memorizzazione cifre di π, 23.
Mendel Gregor, 148-149, 151.
Mersenne, numeri di, 38, 41-43.
Meteorologia, 104, 106-107.
Metodo diretto, 69-70;
indiretto, 69-70.
Minimax, teorema, 188, 191.
Minimo comune multiplo (*mcm*), 61.
Möbius, nastro di, 93-94.
Moltiplicazione, 9-10, 24, 59;
di frazioni, 13-14;
di matrici, 157-158;
di numeri complessi, 33-34;
di quaternioni, 59;
per zero, 5;
tavole di moltiplicazione, 153-154.
Moto, 77-79, 89-90, 106-107, 192-195.

N
Napoleone, teorema di, 86.
Nash John F., 189.
Navier-Stokes, equazioni di, 107.
Newton Isaac, 76, 78, 89-90, 97, 192-193, 195.
Notazione scientifica, 11.
Numeratore, 12-13, **205**.
Numeri:
abbondanti, 40-41;
amichevoli, 40, 42;
complessi, 32-35, 58;
cubici, 33, 81, 170-171, 196, 198;
decimali, 10-11;
conversione delle frazioni, 14-15;
origini dei, 8;
difettivi, 40-41;
fattoriali, 165-166;
immaginari, 32-35, **205**;
irrazionali, 16-17, 19, 21, 25, 30-31, **205**;
negativi, 32-33;
perfetti, 40-43;
primi, 36-39, 40-43, 83, 113, 163, 171, 189, 200, 203, **205**;
primi gemelli, 38, **205**;
quadrati (quadrati perfetti), 16-17, 39-40, 169-170, 196-197, **205**;
razionali, 13, **205**;
reali, 30-31, 33-35, 72, 75, 112;
triangolari, 16, 53.
Numerologia, 39, 42.

O
Ottaedro, 93.

P
Parabola, 17, 88-89, 91, 115.
Pascal Blaise, 52, 90, 124-125, 132;
teorema di, 112, 115.
Pearson, correlazione di, 144-145.
Pendoli, 105-107.
Piano di Fano, 114-115, 167.
Pick, teorema di, 113-114.
Pi greco (π), 20-23.
Pitagora, 16, 18, 40-42, 89, 200;
teorema di, 16, 18, 50, 84-85, 197, 199, **205**.
Poincaré Henri, 95, 102, 104.
Poisson, distribuzione di, 27, 136-137, 139.
Poliedri, 92-94, **205**.
Poligoni, 16, 21, 80, 82-83, 113.
Postulato delle parallele, 75, 108-111.
Previsioni, 106, 139.
Probabilità, 53, 124-127, 166, 186, 191;
condizionata, 128-129;
curva normale, 142-143;
distribuzioni di, 136-139;
e, 27;
genetica, 148-151.
Problema del compleanno, 132-135;
teoria di Bayes, 128-131.
Problema dei quattro colori, 120-123;
del compleanno, 132-135;
della dieta, 180-183.
Programmazione lineare, 180-183.

Q
Q.E.D., 70, 84.
Quadrati:
latini, 167, 172-175;
magici, 168-171, 173.
Quadratura del cerchio, 22, 80-82, 199, **205**.
Quaternioni, 58-59, 170, **205**.

R
Radice quadrata, 17-19, 22, **205**;
di -1, 32-33.
Ragionamento, 19, 50 56, 64-65, 67-68, 70, 110, 113, 133, 150, 160.
Rettangoli aurei, 48-51.
Rettangolo superaureo, 51.
Regressione, 144-147.
Relatività, 75, 109, 111, 192-195.
Resto, 14, **205**.
Rette parallele, 84, 108, 114-115.
Riemann Bernhard, 77, 109, 111, 200;
geometria ellittica, 109, 111, 195;
ipotesi di, 200-203.
Rivelazione degli errori nei codici, 123, 160-162.
Russell Bertrand, 65, 73-74.

S
Serie, **205**.
Sezione aurea (ϕ), 44, 46-47, 49, 50;
superaurea, 47, 51.
Sezioni coniche, 88-90, 115, **204**.
Sfere, 20, 94-95, 98, 110-111.
Sierpiński, triangolo di, 54, 102.
Sillogismo, 64-65.
Simmetria, 52-54, 98, 152-155, **205**;
rotazionale, 152-153;
speculare, 53-54, 152-153.
Sistema Triplo di Steiner (STS), 113, 115.
Sistemi di numerazione, 4-5, 7-11, 176;
binario, 11, 160-162, **204**;
esadecimale, 11, **205**;
indo-arabico, 4-5, 8-9;
posizionale, 4, 8, **205**.
Soluzione ottimale, **205**.
Somma:
con zero, 5, 7;
di frazioni, 13-14;
di matrici, 156-157;
di numeri complessi, 33-34.
Sottrazione con zero, 5.
Spazio-tempo, 97, 99, 111, 195.
Spearman, correlazione di, 145-146.
Spirale logaritmica, 89.
Statistica:
correlazione fra i dati, 144-147;
curva normale, 140-143;
probabilità, 128, 136-139.
Stokes George Gabriel, 107.
Stringhe, teoria delle, 97, 99.
Struttura ponti (trave di Warren), 87.
Successione, **205**.
Sudoku, 167, 172-173, 175.

T
Tartaglia, triangolo di, 53-55, 142.
Tempo polinomiale, 187.
Teorema, 68-70, 72, 100, 110, **205**;
cinese del resto, 60-61, 63;
del limite centrale, 141-142.
Teoria dei giochi, 188-191.
Tetraedro, 93.
Topologia, 89, 91-95, 99.
Toro, 122.
Triangoli, 16, 21, 50-51, 81-82, 84-87, 110-111, 186, 197-198;
costruzione di un triangolo, 82-83;
di Sierpiński, 54, 102;
di Tartaglia, 53-55, 142;
geometria ellittica, 111;
piano di Fano, 114-115, 167;
simmetria 154.
Trigonometria, 84-85, 87.
Tripode, 152-155.
Triscele, 152-154.

U
Unità, 35.

V
Varietà, 95.
Venn, diagramma di, 72, 129, **204**.
Viaggi/trasporti, 159, 183.
Von Lindemann Ferdinand, 21-22, 26, 81-82, 197-199.
Von Neumann John, 188-189, 191.

W
Weinberg Wilhelm, 149-151.
Wiles Andrew, 196-197, 199.

Z
Zermelo-Fraenkel, assiomi di, 74-75.
Zero, 4-7, 9-11, 34, 202.

Indice

Introduzione 3

01 Lo zero 4
02 I sistemi di numerazione 8
03 Le frazioni 12
04 I quadrati e le radici quadrate 16
05 π 20
06 *e* 24
07 L'infinito 28
08 I numeri immaginari 32
09 I numeri primi 36
10 I numeri perfetti 40
11 I numeri di Fibonacci 44
12 I rettangoli aurei 48
13 Il triangolo di Tartaglia 52
14 L'algebra 56
15 L'algoritmo di Euclide 60
16 La logica 64
17 La dimostrazione 68
18 Gli insiemi 72
19 Il calcolo differenziale 76
20 Le costruzioni geometriche 80
21 I triangoli 84
22 Le curve 88
23 La topologia 92
24 Le dimensioni 96
25 I frattali 100
26 Il caos 104
27 Il postulato delle parallele 108
28 La geometria discreta 112
29 I grafi 116
30 Il problema dei quattro colori 120
31 La probabilità 124
32 La teoria di Bayes 128
33 Il problema del compleanno 132
34 Le distribuzioni di probabilità 136
35 La curva di distribuzione normale 140
36 La correlazione fra i dati 144
37 La genetica 148
38 I gruppi 152
39 Le matrici 156
40 I codici 160
41 Il calcolo combinatorio 164
42 I quadrati magici 168
43 I quadrati latini 172
44 La matematica finanziaria 176
45 Il problema della dieta 180
46 Il commesso viaggiatore 184
47 La teoria dei giochi 188
48 La relatività 192
49 L'ultimo teorema di Fermat 196
50 L'ipotesi di Riemann 200

Glossario 204
Indice analitico 206