50 ideias matemática

que precisa mesmo de saber

Tony Crilly

50 Ideias de Matemática
que precisa mesmo de saber

Tony Crilly

Tradução de Jorge Nuno Silva

D. QUIXOTE

Índice

Introdução

A matemática é um campo vasto, e ninguém pode conhecê-lo completamente. O que todos podem fazer é explorá-lo e encontrar o seu próprio caminho. As possibilidades que aqui são apresentadas conduzir-nos-ão a outras épocas e culturas e a ideias que durante séculos intrigaram os matemáticos.

A matemática é simultaneamente antiga e moderna, e constrói-se a partir de muitas e variadas influências culturais e políticas. O nosso sistema de numeração veio da Índia e da Arábia, mas tem várias outras raízes. A «base 60», ou sistema sexagesimal, da Babilónia de dois ou três milénios a.C. marca presença na nossa cultura – há 60 segundos num minuto e 60 minutos numa hora; um ângulo recto ainda tem 90 graus e não 100 grados, como a França ditou num primeiro movimento em direcção à decimalização.

Os triunfos tecnológicos da era moderna dependem da matemática, e já não é motivo de orgulho anunciar que não se foi um bom aluno na matéria. É evidente que a matemática leccionada na escola é diferente, muitas vezes ensinada com os olhos postos no exame. A pressão do tempo na escola também não ajuda, pois a matemática é um assunto em que não há mérito em se ser rápido. É necessário tempo para permitir que as ideias se instalem. Alguns dos melhores matemáticos são penosamente lentos no seu esforço por compreender os conceitos profundos do seu objecto de estudo.

Não há pressa neste livro. Ele pode ser lido como pura forma de lazer. Com calma, descubram o que significam realmente estas ideias, de que possivelmente já ouviram falar. Começando pelo zero, ou qualquer outro assunto que queiram, podem fazer uma viagem entre ilhas de ideias matemáticas. Podem, por exemplo, adquirir conhecimentos sobre a Teoria dos Jogos e a seguir ler sobre Quadrados Mágicos. Ou podem passar dos Rectângulos de Ouro para o famoso Último Teorema de Fermat, ou escolher qualquer outro caminho.

Este é um momento emocionante para a matemática. Alguns dos seus maiores problemas foram recentemente resolvidos. Desenvolvimentos recentes na área da computação ajudaram a resolver alguns, mas revelaram-se impotentes na resolução de outros. O Problema das Quatro Cores foi resolvido com recurso a um computador, mas a hipótese de Riemann, último capítulo deste livro, permanece sem solução – quer utilizando computadores quer utilizando outros meios.

A matemática é para todos. A popularidade do *sudoku* mostra que as pessoas podem fazer matemática (sem que disso se apercebam) e gostar de a fazer. Em matemática, como na arte ou na música, tem havido génios, mas eles não constituem toda a história. Alguns deles aparecerão e desaparecerão nuns capítulos, apenas para reaparecerem noutros. Leonhard Euler, cujo tricentenário ocorreu em 2007, será um visitante frequente destas páginas. Porém, o progresso em matemática é trabalho de muitos, acumulado ao longo de séculos. A escolha dos 50 tópicos é pessoal, mas tentei manter o equilíbrio. Existem questões do dia-a-dia e outras mais avançadas, matemática pura e aplicada, abstracta e concreta, antiga e recente. No entanto, a matemática constitui um todo, e o difícil na escrita do livro não foi escolher os tópicos, mas sim deixar alguns de fora.

Poderiam ter sido 500 ideias, mas 50 são as suficientes para um bom início da vossa carreira matemática.

01 Zero

Ainda muito jovens, fazemos uma entrada hesitante no mundo dos números. Aprendemos que o 1 é o primeiro no «alfabeto numérico» e que introduz os números naturais: 1, 2, 3, 4, 5, etc. Os números naturais são exactamente isso: contam coisas naturais – maçãs, laranjas, bananas, peras. Só mais tarde conseguimos contar o número de maçãs numa caixa vazia.

Mesmo os antigos gregos, que avançaram na ciência e na matemática por saltos quânticos, e os romanos, famosos pelas suas obras de engenharia, não conseguiram encontrar uma forma eficaz de lidar com o número de maçãs numa caixa vazia. Não conseguiram dar um nome a «nada». Os romanos tinham as suas formas de combinar I, V, X, L, C, D e M, mas onde estava o 0? Eles não contavam «nada».

Como é que o zero passou a ser aceite? Pensa-se que o uso de um símbolo que designa a «não-existência» teve a sua origem há milhares de anos. A civilização maia do actual México usou o zero de várias formas. Um pouco mais tarde, o astrónomo Cláudio Ptolomeu, influenciado pelos babilónios, usou um símbolo semelhante ao nosso actual zero como marcador no seu sistema de numeração. Como marcador, o zero pode ser usado para distinguirmos entre exemplos (em notação moderna) como 75 e 705, em vez de nos basearmos no contexto, como os babilónios fizeram. A introdução do zero pode ser comparada à introdução da vírgula na linguagem – ambos ajudam a *ler* o significado correcto. Mas, tal como à vírgula se associa um conjunto de regras de utilização, há regras de utilização do zero.

Brahmagupta, um matemático indiano do século VII, tratou o zero como um número e não apenas como marcador, e definiu regras para a sua utilização. Entre elas, contam-se «A soma de um número positivo com zero é positiva» e «A soma de zero com zero é zero». Estava bastante avançado, ao pensar no zero como um número e não como marcador. O sistema de numeração indo-árabe

Cronologia

700 a.C.
Os babilónios usam o zero como marcador no seu sistema de numeração

628
Brahmagupta usa o zero e enuncia regras para a sua utilização com outros números

que assim incluía o zero foi dado a conhecer no Ocidente por Leonardo de Pisa – Fibonacci – no seu *Liber Abaci* (*O Livro do Cálculo*), publicado pela primeira vez em 1202. Tendo crescido no Norte de África e frequentado a escola aritmética indo-árabe, reconheceu o poder do recurso ao símbolo extra 0 combinado com os símbolos hindus 1, 2, 3, 4, 5, 6, 7, 8 e 9.

A introdução do zero no sistema de numeração colocou um problema que Brahmagupta abordara brevemente: Como deveria o «intruso» ser tratado? O matemático começara a resolver o problema, mas com respostas vagas. Como poderia o zero ser integrado no sistema aritmético existente de uma forma mais precisa? Alguns ajustamentos eram simples. No que respeitava à adição e à multiplicação, o 0 não oferecia problemas, mas o «estrangeiro» não se encaixava bem nas operações de subtracção e divisão. Eram necessários significados para garantir que o 0 se harmonizava com a aritmética aceite.

Como funciona o zero? Somar e multiplicar com zero é simples e incontroverso – pode somar-se 0 a 10 para obter uma centena –, tomando contudo aqui «somar» no sentido menos imaginativo da operação numérica. Somar 0 a um número deixa o número inalterado, enquanto multiplicar qualquer número por 0 conduz sempre à resposta 0. Por exemplo, $7 + 0 = 7$ e $7 \times 0 = 0$. A subtracção é uma operação simples, mas pode conduzir a números negativos, $7 - 0 = 0$ e $0 - 7 = -7$, enquanto a divisão envolvendo o zero levanta dificuldades.

Pensemos num comprimento que deve ser medido com uma vareta de medição. Vamos supor que a vareta de medição tem na realidade 7 unidades de comprimento. Estamos interessados em saber quantas varas podemos estender ao longo do dado comprimento. Se o comprimento a medir tiver 28 unidades, a resposta será 28 dividido por 7 ou, simbolicamente, $28 \div 7 = 4$. Uma melhor notação para representar esta divisão é

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

e, multiplicando em cruzado, pode escrever-se em termos de multiplicação como $28 = 74 \times 4$. Que resultado terá então o 0 divido por 7? Para ajudar a sugerir uma resposta, chame-se a à resposta para que

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

Pela multiplicação cruzada, isto é equivalente a $0 = 7 \times a$. Neste caso, o único valor possível para a é o próprio 0, porque, se a multiplicação de dois números dá 0, um deles tem de ser 0. Como, obviamente, não é o 7, a tem de ser 0.

830
Mahavira elabora raciocínios sobre a forma como o zero interage com os outros números

1100
Bhaskara usa o 0 como um símbolo em álgebra e tenta mostrar como é possível manipulá-lo

1202
Fibonacci usa o símbolo extra 0 juntamente com o sistema indo-árabe de numerais 1, ..., 9, mas não enquanto número com a mesma natureza

Esta não é a principal dificuldade do zero. A questão difícil é a divisão por 0. Se tentarmos tratar $\frac{7}{0}$ da mesma forma que $\frac{0}{7}$, teremos a equação

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

Por multiplicação cruzada, $0 \times b = 7$ e chegamos ao absurdo de que $0 = 7$. Se admitimos a possibilidade de $\frac{7}{0}$ ser um número, criamos o potencial para a confusão numérica em grande escala. A saída é dizer que $\frac{7}{0}$ é indefinido. Não é admissível obter um sentido da operação de dividir 7 (ou qualquer outro número diferente de 0) por 0 e, sendo assim, simplesmente não se permite que esta operação tenha lugar. De forma semelhante, não é permitido colocar uma vírgula no me,io de uma palavra sem perda de sentido.

Bhaskara, um matemático indiano do século XII, seguindo as pisadas de Brahmagupta, pensou sobre a divisão por 0 e sugeriu que um número a dividir por 0 era infinito. A conclusão é razoável porque, se dividirmos um número por outro muito pequeno, o resultado é muito grande. Por exemplo, 7 dividido por uma décima é 70 e por uma centésima, 700. Tornando o denominador cada vez mais pequeno, o resultado obtido é cada vez maior. No mais pequeno de todos, o próprio 0, o resultado deve ser infinito. Adoptando esta forma de raciocinar, ficamos em condições de explicar um conceito ainda mais bizarro – o infinito. Lutar com o infinito não ajuda; o infinito (com a notação padrão ∞) não obedece às regras aritméticas habituais e não é um número no sentido usual do termo.

Se $\frac{7}{0}$ representa um problema, o que fazer com o ainda mais bizarro $\frac{0}{0}$? Se $\frac{0}{0} = c$, por multiplicação cruzada, chega-se à equação $0 = 0 \times c$, e ao facto de que $0 = 0$. Nada de particularmente interessante, mas também não absurdo. De facto, c pode ser *qualquer número* e não se obtém nenhuma impossibilidade. Chega-se à conclusão de que $\frac{0}{0}$ pode ser qualquer coisa. Nos círculos matemáticos chama-se-lhe «indeterminado».

Tomando tudo em consideração, quando se pensa na divisão por zero chega-se à conclusão de que o melhor é excluir essa operação da forma habitual de fazer cálculos. A aritmética pode passar perfeitamente sem ela.

Para que serve o zero? Simplesmente não podemos passar sem ele. O progresso da ciência tem dependido dele. Fala-se sobre zero graus de longitude, zero graus de temperatura, e ainda de energia zero e gravidade zero. O zero entrou na linguagem não científica com expressões como a «hora zero» e a «tolerância zero».

No entanto, poderíamos dar-lhe mais uso. Se sair do passeio da 5.ª Avenida em Nova Iorque e entrar no Empire State Building, vê-se na majestosa entrada do piso número 1. Utilizou-se a capacidade de ordenação dos números, 1 para «primeiro», 2 para «segundo», etc., até 102 para «centésimo segundo» piso. Na Europa, há alguns pisos 0, mas existe alguma relutância em lhes chamar assim.

A matemática não pode funcionar sem o zero. Ele está no âmago dos conceitos matemáticos que fazem o sistema de numeração, a álgebra e a geometria funcionarem. Na sequência de números, o 0 é o que separa os positivos dos negativos, ocupando assim uma posição privilegiada. No sistema decimal, o zero é um marcador que nos permite a utilização de números muito grandes, bem como de números microscópicos.

Ao longo de centenas de anos o zero tornou-se aceite e passou a ser utilizado como uma das maiores invenções do Homem. G. B. Halsted, um matemático americano do século XIX, influenciado por o *Sonho de Uma Noite de Verão* de Shakespeare, escreveu que o zero é o motor do progresso que fornece «ao nada não só uma morada e um nome, uma imagem, um símbolo, mas também um poder útil, característico da raça hindu que o originou».

Quando foi introduzido, o 0 deve ter parecido estranho, mas os matemáticos têm o hábito de se fixarem em conceitos estranhos que se provam úteis muito mais tarde. O equivalente actual é a teoria de conjuntos, em que o conceito de conjunto é uma colecção de elementos. Nessa teoria, ∅ designa o conjunto sem elemento algum, o chamado «conjunto vazio». É uma ideia estranha, mas, tal como o 0, indispensável.

Tudo sobre nada

A soma de zero com um número positivo é positiva.

A soma de zero com um número negativo é negativa.

A soma de um número positivo com um negativo é a sua diferença; ou, se são iguais, zero.

Zero dividido por um número negativo ou positivo ou é zero ou é expresso por uma fracção de numerador zero e com uma quantidade finita como denominador.

Brahmagupta, 628

a ideia resumida
Nada é muita coisa

02 Sistemas de numeração

Um sistema de numeração é um método de lidar com o conceito «quantos». Diferentes culturas em diferentes períodos adoptaram vários métodos, desde o básico «um, dois, três, muitos» até à muito sofisticada notação decimal de posição que se utiliza actualmente.

Os sumérios e os babilónios, que viviam nos actuais Síria, Jordânia e Iraque há cerca de 4000 anos, usavam um sistema de valor de posição para utilização prática no dia-a-dia. Chama-se sistema de valor de posição porque pode dizer-se o «número» pela posição dum símbolo. Os dois povos também utilizavam o 60 como unidade-base – o que se chama hoje em dia um sistema sexagesimal. Ainda existem vestígios do sistema sexagesimal: 60 segundos num minuto, 60 minutos numa hora. Quando medimos ângulos, ainda dizemos que o ângulo giro tem 360 graus, apesar da tentativa do sistema métrico de igualá-lo a 400 grados (para que cada ângulo recto seja igual a 100 grados).

Embora os nossos mais remotos antepassados tenham usado os números para fins práticos, existem indícios de que essas primeiras culturas se sentiram intrigadas pela própria matemática e tiraram algum tempo às utilizações práticas para fazerem explorações. Estas incluíam o que se pode chamar «álgebra» e também as propriedades de figuras geométricas.

O sistema egípcio do século XIII a.C. usava uma base decimal com sinais hieroglíficos. Os egípcios desenvolveram um sistema notável para lidar com fracções, mas a notação actual do valor de posição decimal veio dos babilónios e foi mais tarde melhorada pelos hindus. A vantagem está na forma pela qual se podem representar números muito pequenos e números muito grandes. Usando apenas os números indo-árabes 1, 2, 3, 4, 5, 6, 7, 8 e 9, podem-se fazer cálculos com relativa facilidade. Para verificar que assim é, veja-se o sistema romano.

Cronologia

30 000 a.C.
Povos europeus do Paleolítico fazem marcas de números em ossos

2000 a.C.
Os babilónios usam símbolos para números

Servia as suas necessidades, mas só os especialistas no sistema eram capazes de efectuar cálculos com ele.

O sistema romano Os símbolos básicos usados pelos romanos eram os «dez» (*I*, *X*, *C* e *M*), e as «metades» deles (*V*, *L* e *D*). Os símbolos combinam-se para formar outros. Tem sido apontado que as formas *I*, *II*, *III* e *IIII* derivam da aparência dos dedos, *V* da forma da mão, e o *X* da inversão do *V* e sua junção a outro *V*, obtendo assim duas mãos ou dez dedos. C deriva de *centum* e *M* de *mille*, as palavras latinas para cem e mil respectivamente. Os romanos também usavam *S* para «metade» e um sistema de fracções baseado em 12.

Os romanos utilizavam um método de «anterior e posterior» para obterem os símbolos necessários, mas este não parece ter sido uniformemente adoptado. Os antigos romanos preferiam escrever *IIII*, tendo o *IV* sido introduzido mais tarde. A combinação *IX* também foi usada, mas os romanos queriam dizer 8½ quando escreviam *SIX*! No quadro, estão os números básicos do sistema romano com alguns acrescentos medievais.

Sistema de numeração romano

Império Roman		Acrescentos medievais	
S	um meio		
I	um		
V	cinco	$\bar{V}$	cinco mil
X	dez	$\bar{X}$	dez mil
L	cinquenta	$\bar{L}$	ciquenta mil
C	cem	$\bar{C}$	cem mil
D	quinhentos	$\bar{D}$	quinhentos mil
M	mil	$\bar{M}$	um milhão

Não é fácil trabalhar com os números romanos. Por exemplo, o significado de *MMMCDXLIIII* só se torna claro quando se introduzem mentalmente parênteses por forma a que (*MMM*)(*CD*)(*XL*)(*IIII*) possa ser lido como 3000 + 400 + 40 + 4 = 3444. Mas tente-se somar *MMMCDXLIIII* + *CCCXCIIII*. Um romano treinado nesta arte teria atalhos e truques, mas para nós é difícil obter a resposta correcta sem calcular primeiro no sistema decimal e depois transferir o resultado para o sistema romano:

Adição

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

600

Utilização do precursor do nosso sistema moderno de numeração decimal na Índia

1200

Difusão do sistema indo-árabe de escrever os algarismos de 1 a 9 e um 0

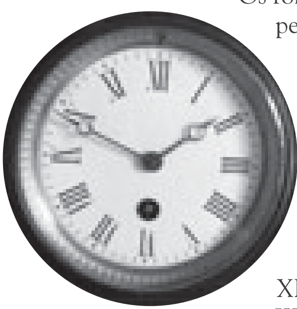
1600

Os símbolos do sistema decimal tomam as formas actuais

A multiplicação de dois números é muito mais difícil e pode ser impossível no sistema básico, mesmo para os romanos! Para multiplicar 3444 por 394, são necessários os acréscimos medievais.

Multiplicação

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


Um relógio Luís XIIII

Os romanos não utilizavam nenhum símbolo específico para o zero. Se pedíssemos a um cidadão vegetariano de Roma para registar quantas garrafas de vinho tinha consumido nesse dia, ele poderia escrever *III*, mas se lhe perguntássemos quantas galinhas tinha comido, não poderia escrever 0. Ainda há vestígios do sistema romano na paginação de alguns livros (embora neste não) e nas pedras dos fundamentos de edifícios. Algumas construções nunca foram usadas pelos romanos, como *MCM* para 1900, mas foram introduzidas por razões de estilo nos tempos modernos. Os romanos teriam escrito *MDCCCC*. O décimo quarto rei Luís de França, universalmente conhecido por Luís XIV, preferia ser conhecido como Luís XIIII e estabeleceu a regra de que os relógios deviam mostrar as 4 horas como *IIII* horas.

Números inteiros decimais Identificamos naturalmente a palavra «números» com números decimais. O sistema decimal é baseado em dez, usando os algarismos 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9. De facto, é baseado em décimos *e* em unidades, mas as unidades podem ser absorvidas pela base decimal. Quando escrevemos o número **394**, podemos explicar o seu significado decimal dizendo que é composto por **3** centenas, **9** dezenas e **4** unidades, e podemos escrever

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

Isto pode ser escrito usando potências de dez (também conhecidas por exponenciais ou índices),

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

em que $10^2 = 10 \times 10$, $10^1 = 10$ e se convenciona que $10^0 = 1$. Nesta expressão, vemos claramente a base decimal do nosso sistema de numeração actual, que torna a adição e a multiplicação bastante transparentes.

A vírgula decimal Até agora, só vimos representações de números inteiros. Será que o sistema decimal consegue lidar com partes de um número, como $^{572}/_{1000}$? Isto significa

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

Podemos tratar os «recíprocos» de 10, 100, 1000 como potências *negativas* de 10, tal que

$$\frac{572}{1000} = 5\times10^{-1} + 7\times10^{-2} + 2\times10^{-3}$$

e isto pode ser escrito **,572**, em que a vírgula decimal indica o início das potências negativas de 10. Se adicionarmos isto à expressão decimal de 394, obtemos a expansão decimal para o número $349\,{}^{572}\!/\!{}_{1000}$, que é simplesmente **394,572**.

Potências de 2	Decimal
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

Para números muito grandes, a notação decimal pode ser muito extensa, portanto neste caso passamos para a «notação científica». Por exemplo, 1 356 936 892 pode escrever-se $1{,}356936892 \times 10^9$, que aparece muitas vezes como «1,356936892 × 10E9» nas calculadoras ou nos computadores. Aqui, a potência 9 é o número de dígitos do número menos um e a letra E significa «exponencial». Por vezes, podemos querer usar números ainda maiores, como por exemplo se estivermos a falar do número de átomos de hidrogénio existentes no universo conhecido. A estimativa desse número é $1{,}7 \times 10^{-77}$. De igual modo $1{,}7 \times 10^{-77}$, com expoente negativo, é um número muito pequeno e também ele pode ser facilmente tratado usando notação científica. Não conseguimos sequer imaginar estes números com os símbolos romanos.

Zeros e uns Embora a base 10 seja comum, algumas aplicações exigem outras bases. O sistema binário que usa a base 2 está por trás do poder do moderno computador. A beleza do sistema binário é a de qualquer número poder ser representado usando apenas os símbolos 0 e 1. A desvantagem é que as representações numéricas podem ser muito extensas. Como podemos representar **394** em notação binária? Desta vez, estamos a trabalhar com potências de 2 e, depois de algum trabalho, podemos obter a expressão completa:

$$\mathbf{394} = \mathbf{1}\times256 + \mathbf{1}\times128 + \mathbf{0}\times64 + \mathbf{0}\times32 + \mathbf{0}\times16 + \mathbf{1}\times8 + \mathbf{0}\times4 + \mathbf{1}\times2 + \mathbf{0}\times1$$

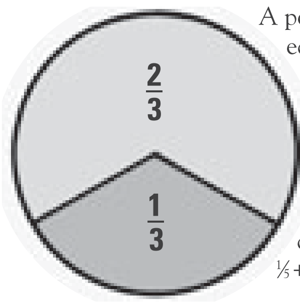
de modo que, anotando os zeros e uns, **394** em notação binário é **110001010**.

Como as representações binárias podem ser muito extensas, na computação aparecem frequentemente outras bases. São o sistema octal (base 8) e o sistema hexadecimal (base 16). No sistema octal só precisamos dos símbolos 0, 1, 2, 3, 4, 5, 6 e 7, enquanto o hexadecimal usa 16 símbolos. No caso do sistema de base 16, usamos normalmente 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E e F. Assim como 10 corresponde à letra A, o número **394** é representado em hexadecimal por 18A. É tão simples como o ABC, que, tenhamos em mente, equivale a 2784 em notação decimal!

a ideia resumida
Escrevendo números

03 Fracções

Uma fracção é um número literalmente fraccionado. Se quisermos partir um número inteiro, a forma correcta de o fazer é recorrer às fracções. Tomemos o exemplo tradicional, o célebre bolo, e partamo-lo em três partes.



A pessoa que fica com duas das três partes fica com uma fracção equivalente a $2/3$. A pessoa com menos sorte fica com $1/3$. Juntando os dois pedaços do bolo, obtemos de novo o bolo inteiro, ou em fracções $1/3 + 2/3 = 1$, em que 1 representa o bolo inteiro.

Aqui está outro exemplo. Decerto já viram nos saldos uma camisa anunciada a quatro quintos do preço original, ou $4/5$. Também se pode dizer que a camisa tem um desconto de um quinto do preço original, que se escreverá $1/5$, e vimos que $1/5 + 4/5 = 1$, em que 1 representa o preço original.

Uma fracção tem sempre a forma de um número inteiro «sobre» um número inteiro. O número de baixo é o «denominador», porque nos diz quantas partes constituem o todo. O número de cima é o «numerador», porque nos diz quantas unidades de fracção existem. Assim, uma fracção na notação habitual é sempre do tipo

$$\frac{\text{numerador}}{\text{denominador}}$$

No caso do bolo, a fracção que quereremos comer é $2/3$, em que o denominador é 3 e o numerador é 2. Os $2/3$ são compostos por duas fracções unitárias de $1/3$.

Também podemos ter fracções como $14/5$ (chamadas fracções impróprias), em que o numerador é maior do que o denominador. Dividindo 14 por 5, obtemos 2 e sobram 4, o que pode ser representado como o número «misto» $2\,4/5$, que inclui o número inteiro 2 e a fracção «própria» $4/5$. Alguns escritores mais antigos escrevem $4/5\,2$. As fracções são habitualmente representadas por forma a que o numerador e o denominador (o número de cima e o número de baixo)

Cronologia

1800 a.C.
Utilização das fracções em culturas babilónicas

1650 a.C.
Os egípcios utilizam fracções unitárias

não tenham factores comuns. Por exemplo, o numerador e o denominador de $\frac{8}{10}$ têm em comum o factor 2, porque 8 = 2 × 4 e 10 = 2 × 5. Se escrevermos a fracção $\frac{8}{10} = \frac{2\times 4}{2\times 5}$ podemos «cortar» os 2 e então $\frac{8}{10} = \frac{4}{5}$, uma forma mais simples com o mesmo valor. Os matemáticos referem-se às fracções como «números racionais» por elas serem razões de dois números. Os números racionais são os números que os gregos conseguiam «medir».

Somar e multiplicar O mais curioso acerca das fracções é que elas são mais fáceis de multiplicar do que somar. A multiplicação de números inteiros é tão difícil, que tiveram de ser inventadas formas engenhosas de a efectuar. Porém, com as fracções, é a adição que é mais difícil e exige algum raciocínio.

Comecemos por multiplicar fracções. Se comprar uma camisa por quatro quintos do preço original de €30, pagará o preço de saldo de €24. Os €30 são divididos em cinco partes de €6 cada e quatro dessas cinco partes é 4 × 6 = 24, o montante que pagará pela camisa.

Posteriormente, o gerente da loja percebe que as camisas não se vendem nada bem e baixa ainda mais o preço, anunciando-as a $\frac{1}{2}$ do preço de saldo. Se for à loja, pode agora comprar a camisa por €12, ou $\frac{1}{2} \times \frac{4}{5} \times 30$, o que é igual a 12. Para multiplicar duas fracções uma pela outra basta multiplicar os denominadores um pelo outro e os numeradores um pelo outro:

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

Se tivesse feito as reduções de uma só vez, o gerente teria anunciado as camisas a quatro décimos do preço original de €30, ou $\frac{4}{10} \times 30$, o que é €12.

Somar duas fracções é um problema diferente. A soma $\frac{1}{3} + \frac{2}{3}$ é fácil porque os denominadores são os mesmos. Limitamo-nos a somar os dois numeradores para obter $\frac{3}{3}$, ou 1. Mas como poderemos somar dois terços de um bolo com quatro quintos desse bolo? Como é que solucionamos $\frac{2}{3} + \frac{4}{5}$? Se ao menos pudéssemos dizer que $\frac{2}{3} + \frac{4}{5} = \frac{2+4}{3+5} = \frac{6}{8}$... mas, infelizmente, não podemos.

A soma de fracções exige outra abordagem. Para somar $\frac{2}{3}$ e $\frac{4}{5}$ temos primeiro de expressar cada uma como fracções com o mesmo denominador. Em primeiro lugar, multiplicamos $\frac{2}{3}$ em cima e em baixo por 5 para obtermos $\frac{10}{15}$. Depois, multiplicamos $\frac{4}{5}$ em cima e em baixo por 3 para obtermos $\frac{12}{15}$. Agora, ambas as

100	1202	1585	1700
Os chineses inventam um sistema para calcular com fracções	Leonardo de Pisa (Fibonacci) populariza a notação de barra nas fracções	Simon Stevin estabelece uma teoria sobre fracções decimais	A barra (/) é de uso comum em fracções (como em a/b)

fracções têm 15 como denominador comum e, para as somar, basta somar os novos numeradores um com o outro:

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

Conversão para decimais No mundo da ciência e na maioria das aplicações da matemática, os números decimais são a forma preferida de representar fracções. A fracção $\frac{4}{5}$ é o mesmo que a fracção $\frac{8}{10}$, que tem 10 como denominador e que podemos escrever como o decimal 0,8.

As fracções que têm 5 ou 10 como denominador são fáceis de converter. Mas como podemos converter, por exemplo, $\frac{7}{8}$, para a forma decimal? Tudo o que precisamos de saber é que, quando dividimos um número inteiro por outro, ou ele cabe exactamente no outro, ou cabe um certo número de vezes e sobra qualquer coisa, a que se chama «resto».

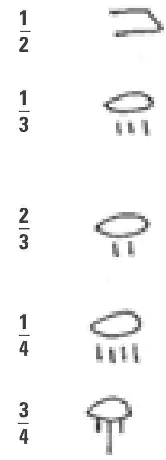
Usando $\frac{7}{8}$ como exemplo, a receita para converter fracções em decimais é a seguinte:

- Tentemos dividir 7 por 8. O 8 não cabe, ou pode dizer-se que cabe 0 vezes com o resto 7. Registamos isto escrevendo zero seguido de vírgula: «0,»
- Agora dividimos 70 por 8 (o resto do passo anterior multiplicado por 10). Cabe 8 vezes, dado que $8 \times 8 = 64$, portanto a resposta é 8 com resto 6 (70 – 64). Acrescentamos isto ao primeiro passo e temos «0,8».
- Agora dividimos 60 por 8 (o resto do passo anterior multiplicado por 10). Como $7 \times 8 = 56$, a resposta é 7 com resto 4. Registamos e até agora temos «0,87».
- Dividimos 40 por 8 (o resto do passo anterior multiplicado por 10). A resposta é *exactamente* 5 com resto *zero*. Quando obtemos o resto 0, a receita está completa. Acabámos. A resposta final é «0,875».

Se aplicarmos esta receita de conversão a outras fracções, é possível que nunca acabemos! Podemos continuar para sempre. Se tentarmos converter $\frac{2}{3}$ em números decimais, por exemplo, verificamos que em cada patamar o resultado de dividir 20 por 3 é 6 com resto 2. Então, temos de dividir novamente 20 por 3, e nunca chegaremos ao ponto em que o resto é zero. Neste caso, temos um número decimal infinito 0,666666... Escreve-se $0{,}\dot{6}$ para indicar «decimal periódico».

Há muitas fracções que continuam infinitamente, como esta. A fracção $\frac{5}{7}$ é interessante. Neste caso, obtemos $\frac{5}{7}$=0,714285714285714285... e constatamos que a sequência 714285 continua a repetir-se. Se *qualquer* fracção resultar numa sequência de repetição, nunca poderemos escrevê-la como um número decimal finito e usaremos a notação com os pontos. No caso de $\frac{5}{7}$, escrevemos $\frac{5}{7} = \dot{7}\dot{1}\dot{4}\dot{2}\dot{8}\dot{5}$.

Fracções egípcias Os egípcios do segundo milénio a.C. basearam o seu sistema de fracções em hieróglifos que designavam por fracções *unitárias* – as fracções cujo numerador é 1. Sabemo-lo pelo papiro de Rhind, conservado no Museu Britânico. Era um sistema de tal forma complicado, que só aqueles treinados no seu uso podiam conhecer os seus segredos ocultos e fazer os cálculos correctos.



Fracções egípcias

Os egípcios usavam algumas fracções privilegiadas, tais como $\frac{2}{3}$, mas todas as outras eram expressas em termos de fracções unitárias como $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{11}$ ou $\frac{1}{168}$. Eram as «fracções básicas», a partir das quais todas as outras podiam ser expressas. Por exemplo $\frac{5}{7}$ não é uma fracção unitária, mas pode ser escrita em termos de fracções unitárias:

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

em que têm de ser usadas fracções unitárias *diferentes*. Uma característica do sistema é poder existir mais de uma maneira de escrever a fracção e algumas fracções serem mais curtas do que outras. Por exemplo,

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

A «expansão egípcia» deve ter tido uma utilização prática limitada, mas o sistema tem inspirado gerações de matemáticos puros e originou muitos problemas desafiantes, alguns dos quais ainda se encontram sem solução. A análise completa dos métodos para encontrar a *menor* «expansão egípcia», por exemplo, aguarda um intrépido explorador matemático.

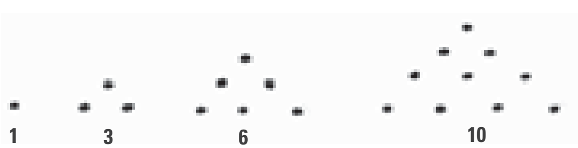
A ideia resumida
Um número sobre outro

04 Quadrados e raízes quadradas

Se gosta de fazer quadrados com pontos, os seus padrões mentais são semelhantes aos dos pitagóricos. Esta actividade era valorizada pelo grupo que seguia o líder Pitágoras, um homem mais recordado pelo teorema a que deu nome. Nasceu na Grécia, na ilha de Samos, e a sua sociedade religiosa secreta prosperou no Sul de Itália. Os pitagóricos acreditavam que a matemática era a chave para a natureza do universo.

Contando os pontos, vemos que o primeiro «quadrado» à esquerda é constituído por um ponto. Para os pitagóricos, o 1 era o número mais importante, impregnado de uma existência espiritual. Começamos bem, portanto. Continuando a contar os pontos dos quadrados seguintes, obtemos os números «quadrados» 1, 4, 9, 16, 25, 36, 49, 64, etc. São os chamados quadrados «perfeitos». Podemos calcular um número quadrado adicionando os pontos da forma ⏋ exterior ao quadrado anterior, por exemplo 9 + 7 = 16. Os pitagóricos não se detiveram nos quadrados. Consideraram outras formas, tais como triângulos, pentágonos (a figura com cinco lados) e outras formas poligonais (muitos lados).

Os números triangulares lembram uma pilha de pedras. Contando os seus pontos, obtemos 1, 3, 6, 10, 15, 21, 28, 36, etc. Se quisermos calcular um número triangular, podemos tomar o anterior e somar-lhe o número de pontos que terá na última fila. Por exemplo, qual é o número triangular que vem a seguir a 10? Terá 5 pontos na última fila, logo limitamo-nos a somar 10 + 5 = 15.



Cronologia

1750 a.C.
Os babilónios elaboram tábuas de raízes quadradas

525 a.C.
Os pitagóricos estudam quadrados de números distribuídos geometricamente

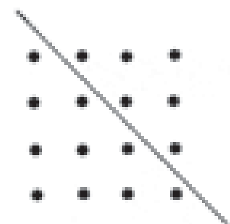
cerca de 300 a.C.
Publicação da teoria dos números irracionais de Eudoxo no Livro V dos *Elementos* de Euclides

Se compararmos os números quadrados e triangulares, veremos que o número 36 aparece em ambas as listas. Mas existe uma ligação ainda mais surpreendente. Se tomarmos números triangulares sucessivos e os somarmos, o que é que obtemos? Tentemos pôr isto numa tabela.

Soma de dois números triangulares sucessivos

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

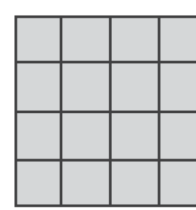
É verdade! Quando somamos dois números triangulares sucessivos, obtemos um número quadrado. Também podemos verificá-lo sem palavras. Consideremos um quadrado constituído por 4 linhas de 4 pontos com uma diagonal desenhada de uma ponta a outra. Os pontos acima da linha (ver figura) formam um número triangular e abaixo da linha está o número triangular seguinte. Esta observação é válida para um quadrado de qualquer dimensão. Destes diagramas de pontos à medição de áreas vai um pequeno passo. A área do quadrado cujo lado é 4 é $4 \times 4 = 4^2 = 16$ unidades quadradas. Genericamente, se o lado é x, então a área será x^2.

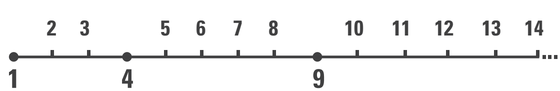


O quadrado x^2 é a base para a forma parabólica. Esta é a forma das antenas de satélite (parabólicas) ou dos espelhos reflectores dos faróis de um carro. Uma parábola tem um ponto focal. Numa antena parabólica, um sensor colocado no foco recebe os sinais reflectidos quando os raios paralelos vindos do espaço chegam à antena curva e ressaltam em direcção ao ponto focal.

Nos faróis de um carro, a lâmpada no foco envia raios paralelos de luz. No desporto, os lançadores de peso, dardo ou martelo reconhecem todos a parábola como a forma do trajecto curvo que qualquer objecto segue antes de cair no chão.

Raízes quadradas Se virarmos a questão ao contrário e pretendermos determinar o comprimento do lado do quadrado que tem área 16, a resposta é simplesmente 4. A raiz quadrada de 16 é 4 e escreve-se $\sqrt{16} = 4$. O símbolo $\sqrt{\ }$ para raízes quadradas é utilizado desde o século XVI. Todos os números quadrados têm números inteiros como raízes quadradas. Por exemplo, $\sqrt{1} = 1$, $\sqrt{4} = 2$, $\sqrt{9} = 3$, $\sqrt{16} = 4$, $\sqrt{25} = 5$, etc. No entanto, existem muitas lacunas ao longo da linha de números entre estes quadrados perfeitos. Estes são 2, 3, 5, 6, 7, 8, 10, 11, etc.





630
Brahmagupta indica métodos para calcular raízes quadradas

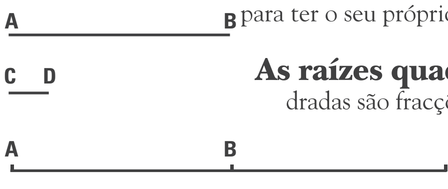
1550
Introdução do símbolo √ para raízes quadradas

1872
Richard Dedekind estabelece uma teoria dos números irracionais

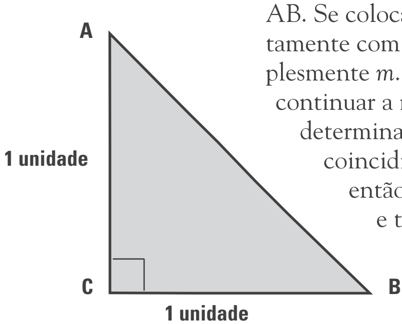
Existe uma alternativa brilhante de notação para raízes quadradas. Tal como x^2 significa um número ao quadrado, podemos escrever a raiz quadrada de um número como $x^{1/2}$, o que se encaixa com a multiplicação de números pela soma dos expoentes. Esta é a base dos logaritmos, inventada depois de termos aprendido, cerca do século XVII, que o problema da multiplicação podia ser vertido num problema de adição. Mas isso é outra história. Estes números têm todos raízes quadradas, mas elas não são iguais a números inteiros. Praticamente todas as calculadoras têm um botão √ e, utilizando-o, verificamos por exemplo que $\sqrt{7} = 2{,}645751311$.

Atentemos a $\sqrt{2}$. O número 2 tem um significado especial para os pitagóricos, porque é o primeiro número par (os gregos consideravam os números pares como femininos e os ímpares como masculinos – e os números pequenos como possuidores de personalidades distintas). Se determinarmos $\sqrt{2}$ com a calculadora, obteremos 1,414213562, presumindo que a calculadora mostra este número de casas decimais. É esta a raiz quadrada de 2? Para verificar, calculamos $1{,}414213562 \times 1{,}414213562$. O resultado é 1,999999999. Não é exactamente 2 porque 1,414213562 é apenas uma aproximação da raiz quadrada de 2.

Tudo o que conseguimos obter nunca é senão uma aproximação! A aproximação decimal de $\sqrt{2}$ com milhões de casas decimais continuará a ser uma aproximação. O número $\sqrt{2}$ é importante em matemática, talvez não tão célebre como o π ou *e* (ver páginas 20-27) mas suficientemente importante para ter o seu próprio nome – é muitas vezes chamado «o número pitagórico».



As raízes quadradas são fracções? Perguntar se as raízes quadradas são fracções liga-se à teoria da medida, como era conhecida pelos antigos gregos. Suponhamos que temos a linha AB, cujo comprimento pretendemos medir, e uma «unidade» indivisível CD para efectuar a medição. Para fazer a medição, colocamos a unidade CD sequencialmente em AB. Se colocarmos a unidade *m* vezes *e* o fim da última unidade coincide exactamente com o fim de AB (no ponto B), o comprimento de AB será simplesmente *m*. Senão, podemos colocar uma cópia de AB junto do original e continuar a medir com a unidade (ver figura). Os gregos acreditavam que, a determinada altura, usando *n* cópias de AB e *m* unidades, a unidade iria coincidir com o ponto final do *m*-ésimo AB. O comprimento de AB seria então *m*/*n*. Por exemplo, se forem colocadas lado a lado 3 cópias de AB e tivermos 29 unidades que cabem exactamente nelas, o comprimento de AB será 29/3.



Os gregos também pensaram na forma de medir o comprimento do lado AB (a hipotenusa) de um triângulo em que os outros dois lados tivessem uma «unidade» de comprimento. Pelo teorema de Pitágoras, o comprimento de AB pode ser representado simbolicamente como $\sqrt{2}$, portanto a questão é se $\sqrt{2} = m/n$.

Com a calculadora, já verificámos que a expressão decimal de $\sqrt{2}$ é potencialmente infinita, e este facto (a expressão decimal não ter fim) talvez indique que $\sqrt{2}$ é uma fracção. Mas não há fim para o decimal 0,3333333… e ele representa a fracção $\frac{1}{3}$. Precisamos de argumentos mais convincentes.

Será $\sqrt{2}$ uma fracção? A pergunta leva-nos a uma das mais famosas provas da matemática. É o tipo de prova de que os gregos mais gostavam: o método de *reductio ad absurdum*. Em primeiro lugar, presume-se que $\sqrt{2}$ não pode ser uma fracção e uma «não-fracção» ao mesmo tempo. É a lei da lógica chamada «terceiro excluído». Não há terceira hipótese nesta lógica. Os gregos foram, portanto, engenhosos. Presumiram que $\sqrt{2}$ era uma fracção e, por pura lógica em cada passo, obtiveram uma contradição, um «absurdo». Vamos fazê-lo. Suponhamos que

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

Também podemos presumir um pouco mais. Podemos presumir que m e n não têm factores comuns. Até aqui, é fácil, porque se tivessem podiam ser cancelados antes de começarmos. (Por exemplo, a fracção $\frac{21}{35}$ é equivalente à fracção $\frac{3}{5}$ cortando o factor comum 7.) Podemos elevar ao quadrado ambos os termos de $\sqrt{2} = \frac{m}{n}$ para obter $2 = \frac{m^2}{n^2}$ e então $m^2 = 2n^2$. Aqui fazemos a primeira observação: dado que m^2 é 2 vezes qualquer coisa, tem de ser um número par. A seguir verifica-se que o próprio m não pode ser ímpar (porque o quadrado de um número ímpar é um número ímpar), logo m é também um número par.

Até agora, a lógica é impecável. Como m é par, deve ser 2 vezes qualquer coisa, o que podemos representar por $m = 2k$. Elevando ambos os termos ao quadrado, temos $m^2 = 4k^2$. A combinação disto com o facto $m^2 = 2n^2$ significa que $2n^2 = 4k^2$ e, cortando o 2, concluímos que $n^2 = 2k^2$. Mas já tínhamos chegado aqui! E, como anteriormente, concluímos que n^2 é par e o próprio n é par. Deduzimos então, por pura lógica, que m e n são ambos pares e têm o factor 2 em comum. Isto é contraditório com a assunção de que m e n não têm factores comuns. A conclusão é então que $\sqrt{2}$ não pode ser uma fracção.

Também se pode provar que a sequência de números $\sqrt{n}$ (excepto quando n é um quadrado perfeito) não pode ser uma fracção. Os números que não podem ser expressos por fracções chamam-se números «irracionais», pelo que acabámos de observar que existe um número infinito de números irracionais.

a ideia resumida
O caminho para os números irracionais

05 π

O π é o número mais famoso da matemática. Esqueçam todas as outras constantes; o π estará sempre no topo da lista. Se houvesse um Óscar para os números, o π ganharia o prémio todos os anos.

O π, ou pi, é o comprimento do exterior de um círculo (a circunferência) dividido pelo comprimento através do centro (diâmetro). O seu valor, a razão destes dois comprimentos, não depende do tamanho do círculo. Quer o círculo seja grande ou pequeno, o π é uma constante matemática. O círculo é o seu ambiente natural, mas o π aparece por todo o lado na matemática, e em lugares que não estão sequer remotamente relacionados com o círculo.

Para um círculo de diâmetro *d* e raio *r* :

circunferência $= \pi d = 2\pi r$

área $= \pi r^2$

Para uma esfera de diâmetro *d* e raio *r*:

área da superfície $= \pi d^2 = 4\pi r^2$

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

Arquimedes de Siracusa A relação entre a circunferência e o diâmetro de um círculo é um assunto que há muito suscita interesse. Cerca de 2000 a.C., os babilónios observaram que a circunferência era aproximadamente 3 vezes mais comprida que o seu diâmetro.

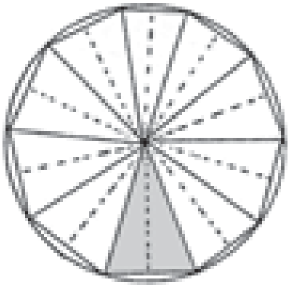
Foi Arquimedes de Siracusa quem iniciou realmente a teoria matemática do π, por volta de 225 a.C. Arquimedes é um dos grandes. Os matemáticos gostam de classificar os seus colegas de trabalho e colocaram-no ao nível de Carl Friedrich Gauss (o «Príncipe dos Matemáticos») e Sir Isaac Newton. Quaisquer que sejam os méritos desta opinião, é claro que Arquimedes figuraria em qualquer quadro de honra. No entanto, dificilmente seria uma figura numa torre de marfim – além de contribuir para a astronomia, a matemática e a física, desenhou armas de guerra, como catapultas, alavancas e «espelhos de incendiar», todas utilizadas para repelir os romanos. Arquimedes tinha, não obstante, algo de professor distraído, como demonstra o salto que deu do banho e a sua corrida todo nu pela rua abaixo gritando «Eureca!», quando descobriu a lei da flutuabilidade em hidrostática. A história não registou a forma como comemorou o seu trabalho sobre o π.

Cronologia

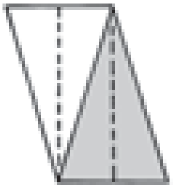
2000 a.C.
Os babilónios observam que o π é aproximadamente 3

250 a.C.
Arquimedes faz um cálculo muito razoável do π em 22/7

Se o π é definido como a razão entre a circunferência e o seu diâmetro, o que é que isso tem a ver com a *área* do círculo? É uma *dedução* que a área de um círculo de raio r seja πr^2, embora provavelmente seja mais conhecida do que a definição de circunferência/diâmetro do π. É notável que o π sirva tanto para as áreas como para as circunferências.



Como é que isto pode ser mostrado? O círculo pode ser dividido num número de triângulos estreitos e iguais com base de comprimento b cuja altura é aproximadamente o raio r. Formam um polígono dentro do círculo que se aproxima da área do círculo. Comecemos com 1000 triângulos. Todo o processo é um exercício de aproximações. Podemos juntar cada *par* adjacente desses triângulos para formar um rectângulo (aproximadamente) com área $b \times r$, logo a área total do polígono será $500 \times b \times r$. Como $500 \times b$ é cerca de metade da circunferência, o seu comprimento é πr e a área do polígono é $\pi r \times r = \pi r^2$. Quanto mais triângulos considerarmos, mais próxima será a aproximação e no limite concluiremos que a área do círculo é πr^2.



Arquimedes estimou que o valor de π estaria compreendido entre $^{223}/_{71}$ e $^{220}/_{70}$. Assim, é a Arquimedes que devemos a familiar aproximação 22/7 para o valor do π. A honra da designação do símbolo como π é do pouco conhecido William Jones, um matemático galês que se tornou vice-presidente da Royal Society of London no século XVIII. Foi o matemático e físico Leonhard Euler que popularizou o π no contexto das relações do círculo.

O valor exacto do π Nunca podemos saber o valor exacto de π visto que ele é um número irracional, facto provado por Lambert em 1768. A expansão decimal é infinita, sem nenhum padrão previsível. As primeiras 20 casas decimais são 3,14159265358979323846... O valor de $\sqrt{10}$ usado pelos matemáticos chineses é $3{,}1\underline{6}227766016837933199$ e foi adoptado por volta do ano 500 por Brahmagupta. Na verdade, esse valor é pouco melhor do que o valor bruto de 3 e difere na segunda casa decimal de π.

O π pode ser calculado a partir de uma série de números. Uma muito conhecida é

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

1706
William Jones introduz o símbolo π

1761
Lambert prova que o π é irracional

1882
Lindemann prova que o π é transcendental

embora seja penosamente lenta na sua convergência para o π e bastante ineficaz para o cálculo. Euler encontrou uma série notável que converge para π:

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

O autodidacta genial Srinivasa Ramanujan imaginou algumas fórmulas espectaculares de aproximação ao π. Uma envolve apenas a raiz quadrada de 2:

$$\frac{9801}{4412}\sqrt{2} = 3{,}1415927300133056603139961890\ldots$$

Os matemáticos são fascinados pelo π. Enquanto Lambert provou que ele não podia ser uma fracção, em 1882 o matemático alemão von Lindemann resolveu o mais fascinante problema associado ao π. Mostrou que o π é «transcendente», isto é, o π não pode ser solução de uma equação algébrica (uma equação que só envolve potências de x). Ao resolver este «enigma secular», Lindemann concluiu o problema da «quadratura do círculo». Dado um círculo, o desafio era construir um quadrado com a mesma área usando apenas um compasso e uma régua. Lindemann provou de forma conclusiva que isso não pode ser feito. Actualmente, a expressão «quadratura do círculo» é sinónimo de uma impossibilidade.

O cálculo efectivo do π continuou a bom ritmo. Em 1853, William Shanks reivindicou o valor correcto com 607 casas (na realidade correcto apenas até à 527.ª). Hoje em dia, a pesquisa para calcular o π com mais e mais casas decimais ganhou ímpeto através dos computadores modernos. Em 1949, o π foi calculado com 2037 casas decimais, o que demorou 70 horas a concluir num computador ENIAC. Em 2002, o π foi calculado com umas impressionantes 1 241 100 000 000 casas decimais, mas o número é uma cauda sempre crescente. Se estivéssemos no equador e começássemos a escrever a expansão do π, os cálculos de Shank ocupariam 14 metros, mas o comprimento de 2002 levar-nos-ia a dar 62 voltas à volta do mundo!

Muitas questões sobre o π têm sido postas e respondidas. Os algarismos do π são aleatórios? É possível encontrar uma sequência predeterminada na expansão? Por exemplo, é possível encontrar a sequência 0123456789 na expansão? Em 1950, isto parecia irresolúvel. Ninguém tinha encontrado tal sequência nos algarismos conhecidos do π. L. E. J. Brouwer, um influente matemático holandês, declarou que a questão era desprovida de sentido, dado estar convencido de que não podia ser verificada. Na realidade, os dígitos foram descobertos em 1997 começando na posição 17 387 594 880, ou, usando a metáfora do equador, a cerca de 3000 milhas de uma volta completa ao mundo. Encontraríamos dez seis seguidos antes de completarmos 600 milhas, mas teríamos de esperar até que uma volta estivesse completa e continuar mais 3600 milhas para encontrar dez setes seguidos.

A importância do π De que adianta saber tantas casas decimais do π? Afinal, a maioria dos cálculos necessita apenas de algumas; provavelmente, não são necessárias mais de dez numa aplicação prática, e a aproximação de Arquimedes de 22/7 é em geral suficiente. Mas os longos

O π na poesia

Se realmente nos quisermos lembrar dos primeiros valores da expansão do π, a poesia pode ajudar. Seguindo a tradição de ensinar matemática pela utilização de mnemónicas, existe uma variação brilhante de Michael Keith do poema «O Corvo», de Edgar Allan Poe.

O poema de Poe começa

The raven E. A. Poe

Once upon a midnight dreary, while I pondered weak and weary,

Over many a quaint and curious volume of forgotten lore,

A variante de Keith para o π começa

Poe, E. Near A Raven

Midnights so dreary, tired and weary.

Silently pondering volumes extolling all by-now obsolete lore.

O número de letras de cada palavra sucessiva na versão de Keith fornece os primeiros 740 dígitos do π.

cálculos não são feitos apenas por prazer. São utilizados para testar os limites dos computadores, para além de serem um fascínio para o grupo de matemáticos que se intitulou de «amigos do pi».

Talvez o mais estranho episódio na história do π tenha sido uma tentativa da Assembleia Legislativa do Indiana (nos Estados Unidos da América) de fazer aprovar uma lei que fixasse o seu valor. Foi no final do século XIX, quando o médico E. J. Goodwin introduziu a lei para tornar o π «digerível». O problema prático da proposta foi a incapacidade do proponente de fixar o valor que queria. Felizmente para o estado do Indiana, a insensatez de legislar sobre o π foi percebida antes de a lei ser ratificada. Desde esse dia, os políticos têm deixado o π em sossego.

a ideia resumida
Achar o π

06 *e*

O *e* é um bebé, quando comparado com o seu único rival π . Enquanto o π é mais solene e tem um grande passado que remonta aos babilónios, o *e* não é tão sobrecarregado pela História. A constante *e* é jovem e vibrante e encontra-se sempre presente quando se analisa o «crescimento». Quer se trate de populações, dinheiro ou outras quantidades físicas, o crescimento envolve invariavelmente *e*.

O *e* é um número cujo valor aproximado é 2,71828. Porque é que é tão especial? Não é um número escolhido ao acaso, mas sim uma das grandes constantes matemáticas. Surgiu no início do século XVII quando muitos matemáticos dedicavam os seus esforços à clarificação da ideia de logaritmo, a brilhante invenção que permitia que a multiplicação de números grandes fosse convertida em adição.

Mas, na realidade, a história começa com um negócio do século XVII. Jacob Bernoulli foi um dos célebres Bernoullis suíços, uma família que se dedicou a dar ao mundo uma dinastia de matemáticos. Jacob começou a trabalhar o problema dos juros compostos em 1683.

Dinheiro, dinheiro, dinheiro Consideremos um período de um ano, uma taxa de juro exorbitante de 100%, e um depósito inicial (chamado montante «principal») de €1. Claro que raramente obtemos 100% pelo nosso dinheiro, mas esse número adapta-se ao nosso objectivo e o conceito pode ser adaptado a taxas realistas, como 6% ou 7%. Da mesma forma, se tivermos montantes principais maiores, como €10 000, podemos multiplicar tudo por 10 000.

No final do ano, à taxa de 100%, teremos o montante principal e o montante dos juros ganhos, que neste caso também são de €1. Assim, teremos a principesca soma de €2. Suponhamos agora que a taxa de juro é reduzida a metade, para 50%, mas é aplicada a cada meio ano separadamente.

Cronologia

1618
John Napier descobre uma constante *e* relacionada com logaritmos

1727
Euler utiliza a notação *e* em relação com a teoria dos logaritmos; é chamada por vezes a constante de Euler.

No primeiro meio ano, ganhamos um juro de 50 cêntimos e o nosso montante principal cresceu para €1,50 no final do primeiro meio ano. Assim, no final do ano teremos este montante e 75 cêntimos de juros. O nosso €1 cresceu para €2,25 no final do ano! Compondo (ou capitalizando) os juros em cada meio ano, fizemos mais 25 cêntimos. Pode não parecer muito mas, se tivermos €10 000 para investir, ganharíamos €2250 em vez de €2000. A composição a cada meio ano resulta em mais €250.

Mas se a composição a cada meio ano significa que ganhamos poupanças, o banco também ganha em qualquer quantia que estivermos a dever, portanto temos de ser cuidadosos! Suponhamos agora que o ano está dividido em quatro quartos e são aplicados 25% em cada um. Fazendo cálculos semelhantes, constatamos que o nosso €1 cresceu para €2,44141. O nosso dinheiro está a crescer e, com os nossos €10 000, parece vantajoso dividir o ano e aplicar percentagens menores de taxas de juro a períodos de tempo menores.

Será que o nosso dinheiro crescerá desmesuradamente e nos fará milionários? Se continuarmos a dividir o ano em unidades cada vez menores, como se vê na tabela, este «processo de limitação» mostra que o montante parece estabilizar numa constante. Claro que o único período de composição realista é o dia (é o que os bancos fazem). A mensagem matemática é de que este limite, a que os matemáticos chamam *e*, é o montante que 1 euro cresce se a composição se verificar continuamente. É uma coisa boa ou má? Sabemos a resposta: se estivermos a poupar, boa; se devermos dinheiro, má.

Composição a cada …	Montante acumulado
ano	€2,00000
meio ano	€2,25000
quarto	€2,44141
mês	€2,61304
semana	€2,69260
dia	€2,71457
hora	€2,71813
minuto	€2,71828
segundo	€2,71828

O valor exacto de *e* Tal como o π, *e* é um número irracional, pelo que, como com o π, não podemos saber o seu valor exacto. Com 20 casas decimais, o valor é 2,71828182845904523536...

Usando apenas fracções, a melhor aproximação do valor de *e* é 87/32, se o numerador e o denominador da fracção estiverem limitados a dois dígitos. Curiosamente, se o numerador e o denominador estiverem limitados a três dígitos, a melhor fracção é 878/323. Esta segunda fracção é um tipo de extensão palindrómica da primeira – os matemáticos têm o hábito de oferecer estas pequenas surpresas. Uma expansão em série de *e* muito conhecida é dada por

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

1748

Euler determina 23 algarismos do *e;* é-lhe atribuído crédito pela descoberta da famosa fórmula $e^{i\pi} + 1 = 0$ na mesma altura

1873

Hermite prova que *e* é um número transcendental

2007

Determinação do *e* até à ordem de 10^{11} dígitos

Aqui, é conveniente a notação de factorial com um ponto de exclamação. Como nesta, por exemplo $5! = 5 \times 4 \times 3 \times 2 \times 1$. Usando esta notação, *e* toma um aspecto mais familiar

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

Desta forma, o número *e* parece ter indubitavelmente um padrão. Nas suas propriedades matemáticas, *e* parece mais «simétrico» do que o π.

Se quiserem uma forma de lembrar os primeiros algarismos do *e*, tentem o seguinte: «We attempt a mnemonic to remember a strategy to memorize this count...», em que o número de letras de cada palavra dá o algarismo seguinte de *e*. Para quem tiver conhecimentos de História norte-americana, o *e* é «2,7 Andrew Jackson Andrew Jackson», dado que Andrew Jackson, o sétimo presidente dos Estados Unidos, foi eleito em 1828. Há muitas estratégias semelhantes para recordar o valor de *e*, mas o seu interesse reside mais na sua singularidade do que em qualquer vantagem matemática.

Foi Leonhard Euler que, em 1737, provou que o *e* é irracional (não é uma fracção). Em 1840, o matemático francês Joseph Liouville mostrou que *e* não é solução de nenhuma equação quadrática e, em 1873, num trabalho pioneiro, o seu compatriota Charles Hermite provou que o *e* é transcendental (não pode ser solução de nenhuma equação algébrica). O importante aqui é o método usado por Hermite. Nove anos mais tarde, Ferdinand von Lindemann adaptou o método de Hermitage para provar que o π era transcendental, um problema muito mais complicado.

Resolveu-se um problema, mas outros apareceram. Será que *e* elevado a *e* é transcendental? É uma expressão muito bizarra, mas como poderia ser de outra forma? No entanto, isto ainda não foi provado com rigor e, pelos estritos padrões da matemática, deve continuar a ser classificado como conjectura. Os matemáticos aproximaram-se de uma prova e mostraram que é impossível que tanto *e* como *e* elevado a e^2 sejam transcendentais. Quente, mas não o suficiente.

As relações entre π e *e* são fascinantes. Os valores de e^{π} e de π^{e} são próximos, mas é fácil provar (sem chegar a calcular os seus valores) que $e^{\pi} > \pi^{e}$. Se fizeram batota e consultaram a calculadora, verão que os valores aproximados são $e^{\pi} = 23{,}14069$ e $\pi^{e} = 22{,}45916$.

O número e^{π} é conhecido como constante de Gelfond (do matemático russo Aleksandr Gelfond) e provou-se que era transcendental. Sabemos muito menos coisas sobre o π^{e}; se é irracional, ainda não o conseguimos provar.

O *e* é importante? É principalmente no crescimento que o *e* nos interessa, por exemplo, no crescimento económico e no crescimento populacional. Também relacionadas com *e* estão as curvas que dependem de dele e que são usadas para modelar a decomposição radioactiva.

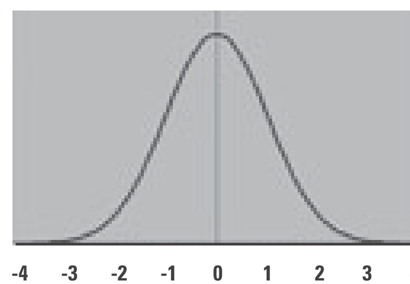
O número *e* também ocorre em problemas não relacionados com crescimento. Pierre Montmort investigou um problema de probabilidade no século XVIII quem tem sido estudado intensivamente desde então. Na versão mais simples, um grupo de pessoas vão almoçar juntas, após o que cada uma leva um chapéu de forma aleatória. Qual é a probabilidade de ninguém levar o próprio chapéu?

Pode provar-se que a probabilidade é $1/e$ (cerca de 37%), donde a probabilidade de pelo menos uma pessoa levar o seu próprio chapéu é $1 - 1/e$ (cerca de 63%). Esta aplicação na teoria das probabilidades é uma de muitas. A distribuição de Poisson, que trata de acontecimentos raros, é outra. Estes foram os exemplos iniciais, mas de forma alguma isolados: James Stirling conseguiu uma notável aproximação ao valor factorial $n!$ envolvendo o e (e o π); em estatística, a conhecida curva de Bell da distribuição normal envolve o e; e em engenharia a curva de suspensão dos cabos de uma ponte depende do e. A lista é infindável.

Uma identidade avassaladora O prémio de fórmula mais notável de toda a matemática envolve o e. Quando pensamos nos números mais famosos da matemática, pensamos em 0, 1, π, e e o número imaginário $i = \sqrt{1}$. Como pode ser que

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

É mesmo! Este resultado é atribuído a Euler.



Distribuição normal

Talvez a real importância do e esteja no enigma que cativou gerações de matemáticos. O e é, com efeito, inevitável. É exactamente esta a razão pela qual um autor como E. V. Wright se deu ao trabalho de escrever um romance sem a letra e – provavelmente também tinha um pseudónimo –, mas o seu *Gadsby* é apenas e tão-só isso. É difícil imaginar um matemático começar a escrever um manual sem o número e, ou ser sequer capaz de o fazer.

a ideia resumida
O mais natural dos números

07 Infinito

Qual é o tamanho do infinito? A resposta mais simples é que o ∞ (o símbolo de infinito) é muito grande. Pensemos numa linha recta com números cada vez maiores que se estende «até ao infinito». Para cada grande número obtido, digamos 10^{1000}, existe sempre um maior, como $10^{1000} + 1$.

Esta é a ideia tradicional de infinito, números continuamente sucessivos. Os matemáticos usam o infinito de várias formas, mas há que ter cuidado ao tratar o infinito como um número normal. Não é.

Contagem O matemático alemão Georg Cantor deu-nos um conceito de infinito completamente diferente. Criou sozinho uma teoria que tem impulsionado imenso a matemática moderna. A ideia da qual depende a teoria de Cantor tem a ver com a noção primitiva de contagem, mais simples do que aquela que usamos no dia-a-dia.

Imaginemos um camponês que não soubesse contar com números. Como saberia ele quantas ovelhas tinha? Simples – depois de as deixar sair pela manhã, saberia dizer se todas tinham voltado à tarde emparelhando cada ovelha com uma pedra de uma pilha no portão do seu campo. Se faltasse uma ovelha, sobraria uma pedra. Mesmo sem usar números, o camponês pensa de forma assaz matemática. Utiliza o conceito da correspondência «um para um» entre ovelhas e pedras. Esta noção primitiva tem algumas consequências surpreendentes.

A teoria de Cantor implica conjuntos (um conjunto é muito simplesmente uma colecção de objectos). Por exemplo, **N** = {1, 2, 3, 4, 5, 6, 7, 8, ...} significa o conjunto dos números inteiros (positivos). Uma vez que temos um conjunto, podemos falar de subconjuntos, que são conjuntos mais pequenos dentro do conjunto maior. Os subconjuntos mais óbvios relacionados com o nosso conjunto **N** são os subconjuntos **O** = {1, 3, 5, 7, ...} e **E** = {2, 4, 6, 8, ...}, que são

Cronologia

350
Aristóteles rejeita um verdadeiro infinito

1639
Girard Desargues introduz o conceito de infinito na geometria

os conjuntos dos ímpares e dos pares, respectivamente. Qual será a nossa resposta à pergunta «existe o mesmo número de números pares e ímpares?»? Embora não nos seja possível contar os elementos de cada conjunto e comparar os resultados, a resposta será seguramente «sim». Em que se baseia esta confiança? Provavelmente qualquer coisa como «metade dos inteiros são pares e metade são ímpares». Cantor concordaria com a resposta, mas daria uma razão diferente. Diria que, para cada par, temos um «emparelhado» par. A noção de que os conjuntos **O** e **E** têm o mesmo número de elementos baseia-se no emparelhamento de cada par com um ímpar:

O:	1	3	5	7	9	11	13	15	17	19	21...
	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
E:	2	4	6	8	10	12	14	16	18	20	22...

Se fizermos uma nova pergunta, «existirá o mesmo número de números inteiros e números *pares*?», a resposta será «não», seguindo o raciocínio de que o conjunto **N** tem o dobro dos números do conjunto dos pares.

No entanto, a noção de «mais» é bastante obscura quando se trata de conjuntos com um número infinito de elementos. Podemos fazer melhor com o conceito de correspondência «um para um». Surpreendentemente, existe uma correspondência «um para um» entre **N** e o conjunto dos números pares **E**:

N:	1	2	3	4	5	6	7	8	9	10	11...
	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕	↕
E:	2	4	6	8	10	12	14	16	18	20	22...

Chegamos à surpreendente conclusão de que existe o «mesmo número» de números inteiros e números pares! Isto é exactamente o oposto da «noção comum» declarada pelos antigos gregos; o início dos *Elementos* de Euclides de Alexandria diz que «o todo é maior do que as partes».

Cardinalidade Chama-se cardinalidade ao número de elementos de um conjunto. No caso das ovelhas, a cardinalidade registada pelas contas do camponês é de 42. A cardinalidade do conjunto $\{a, b, c, d, e\}$ é 5 e escreve-se $card\{a, b, c, d, e\} = 5$. Logo, a cardinalidade é a medida do «tamanho» do conjunto. Para a cardinalidade do conjunto dos números inteiros **N**, e qualquer

1655
É atribuído a John Wallis o crédito de ter sido o primeiro a usar o símbolo do nó, ∞, para infinito

1874
Cantor faz um tratamento rigoroso da noção de infinito, especificando diferentes ordens de infinito

anos 1960
Abraham Robison inventa uma aritmética não-padrão baseada na noção de infinitesimal

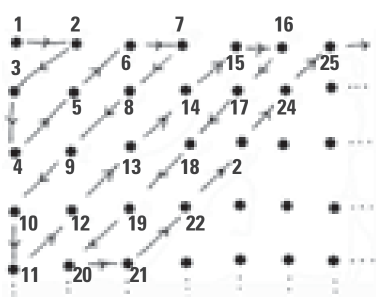
conjunto com uma correspondência «um para um» com **N**, Cantor usou o símbolo $\aleph_0$ ($\aleph$ ou *aleph*, do alfabeto hebraico; o símbolo $\aleph_0$ lê-se «alefe zero»), Assim, em linguagem matemática, podemos escrever *card*(**N**) = *card*(**O**) = *card*(**E**) = $\aleph_0$.

A qualquer conjunto que possa ser posto em correspondência «um para um» com **N** chama-se um «infinito enumerável», o que significa que podemos escrever os seus elementos numa lista. Por exemplo, a lista dos números ímpares é simplesmente 1, 3, 5, 7, 9, ... e sabemos qual é o primeiro elemento, qual é o segundo, etc.

Serão as fracções um infinito contável? O conjunto das fracções **Q** é um conjunto maior do que **N**, no sentido em que se pode pensar em **N** como um subconjunto de **Q**. Podemos escrever todos os elementos de **Q** numa lista? Será possível inventar uma lista tal, que todas as fracções (incluindo as negativas) se encontram nela? A ideia de que um conjunto tão grande como esse possa ser posto em correspondência «um para um» com **N** parece impossível. No entanto, pode ser feito.

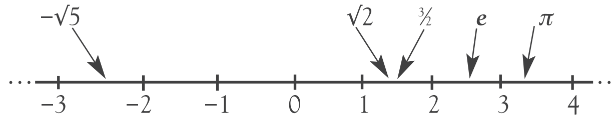
1	-1	2	-2	3	-3	4	...
$\frac{1}{2}$	$\frac{-1}{2}$	$\frac{3}{2}$	$\frac{-3}{2}$	$\frac{5}{2}$	$\frac{-5}{2}$	$\frac{7}{2}$	...
$\frac{1}{3}$	$\frac{-1}{3}$	$\frac{2}{3}$	$\frac{-2}{3}$	$\frac{4}{3}$	$\frac{-4}{3}$	$\frac{5}{3}$	...
$\frac{1}{4}$	$\frac{-1}{4}$	$\frac{3}{4}$	$\frac{-3}{4}$	$\frac{5}{4}$	$\frac{-5}{4}$	$\frac{7}{4}$	...
$\frac{1}{5}$	$\frac{-1}{5}$	$\frac{2}{5}$	$\frac{-2}{5}$	$\frac{3}{5}$	$\frac{-3}{5}$	$\frac{4}{5}$	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	

Para o fazer, comecemos a pensar em termos bidimensionais. Para começar, escrevemos uma linha com todos os números inteiros positivos e negativos alternadamente. Abaixo, escrevemos outra linha com todas as fracções com denominador 2, omitindo aquelas que aparecem na linha de cima (como $6/2$ = 3). Abaixo desta, escrevemos uma linha das fracções cujo denominador é 3, omitindo de novo aquelas que já tenham sido escritas. Continua-se desta forma, obviamente nunca acabando, mas sabendo exactamente onde cada fracção aparece neste diagrama. Por exemplo, $209/67$ está na linha 67.ª, cerca de 200 posições à direita de $1/67$.



Exibindo as todas as fracções desta forma, pelo menos em teoria, podemos construir uma lista unidimensional. Se começarmos com a linha de cima e começarmos a andar para a direita, nunca chegaremos à segunda linha. Contudo, escolhendo um tortuoso caminho em ziguezague, podemos ter êxito. Começando com 1, a prometida lista começa: 1, −1, ½, ⅓, −½, 2, −2, e vamos seguindo as setas. Cada fracção, positiva ou negativa, está algures nesta lista linear e reciprocamente a sua posição dá o seu «emparelhamento» na lista bidimensional de fracções. Logo, podemos concluir que o conjunto de fracções **Q** é um infinito enumerável e escrever *card*(**Q**) = $\aleph_0$.

Listar os números reais Embora o conjunto de fracções seja responsável por muitos elementos na recta dos números reais, também há números reais como √2, *e* e π que *não são* fracções. São números irracionais – preenchem os hiatos para nos fornecer a recta dos números reais **R**.



Quando os hiatos são preenchidos, o conjunto **R** é referido como *continuum*. Assim, como poderemos fazer uma lista dos números reais? Num gesto de pura genialidade, Cantor mostrou que mesmo uma tentativa de colocar os números reais *entre 0 e 1* numa lista estava condenada ao fracasso. A afirmação surgiu indubitavelmente como um choque para quem se dedicava à elaboração da lista, que podia de facto perguntar como é que um conjunto de números não podia ser escrito um a seguir ao outro.

Suponhamos que não acreditamos em Cantor. Sabemos que cada número entre 0 e 1 pode ser expresso de forma decimal, por exemplo, ½ = 0,50000000000000000. . . e $1/\pi$ = 0,31830988618379067153. . . e teremos de dizer a Cantor «aqui está a nossa lista de todos os números entre 0 e 1», a que chamaremos r_1, r_2, r_3, r_4, r_5, . . . Se não a conseguirmos elaborar, Cantor está correcto.

Suponhamos que Cantor olha para a lista e marca a **negrito** os números na diagonal:

$$r_1\colon\ 0,\boldsymbol{a_1}a_2a_3a_4a_5\ldots$$

$$r_2\colon\ 0,b_1\boldsymbol{b_2}b_3b_4b_5\ldots$$

$$r_3\colon\ 0,c_1c_2\boldsymbol{c_3}c_4c_5\ldots$$

$$r_4\colon\ 0,d_1d_2d_3\boldsymbol{d_4}d_5\ldots$$

Cantor perguntaria: «Pois, mas onde está o número $x = x_1x_2x_3x_4x_5. . .$ em que x_1 é diferente de a_1, x_2 é diferente de b_2, x_3 é diferente de c_3 continuando ao longo da diagonal?» O x dele difere de todos os números da lista numa casa decimal e portanto não pode estar lá. Cantor tem razão.

De facto, não é possível uma lista para o conjunto dos números reais **R**, que assim é um conjunto infinito «maior», com uma «ordem de infinito maior», do que o infinito do conjunto de fracções **Q**. O que era grande tornou-se ainda maior.

a ideia resumida
Uma chuva de infinitos

08 Números imaginários

Podemos certamente imaginar números. Às vezes imagino que tenho um milhão de euros na minha conta bancária, e não há dúvida de que esse é um número imaginário. Mas a utilização matemática do imaginário nada tem a ver com estes devaneios.

Pensa-se que o rótulo «imaginário» se deve ao filósofo e matemático René Descartes, em reconhecimento de algumas curiosas soluções de equações que não eram decididamente números habituais. Os números imaginários existem? Era esta pergunta que os filósofos faziam quando se concentravam na palavra imaginário. Para os matemáticos, a existência de números imaginários não é um problema. Eles fazem tanto parte do dia-a-dia como o número 5 ou o π. Os números imaginários podem não ajudar nas idas às compras, mas perguntem a qualquer projectista de aeronaves ou engenheiro electrotécnico e constatarão que eles são de vital importância. E, somando um número real com um imaginário, obtemos o que se chama um «número complexo», que parece filosoficamente menos incómodo. A teoria dos números complexos gira à volta da raiz quadrada de *menos 1*. Então qual é o número que, elevado ao quadrado, dá -1?

Se tomarmos qualquer número diferente de zero e o multiplicarmos por si próprio (elevá-lo ao quadrado), obteremos sempre um número positivo. Isto é plausível quando se elevam ao quadrado números positivos, mas será verdade quando se elevam ao quadrado números negativos? Podemos usar -1×-1 como teste. Mesmo que nos tivéssemos esquecido da regra básica de que «menos com menos dá mais», devemos lembrar-nos de que a resposta ou é -1 ou $+1$. Se pensarmos que -1×-1 dá -1, podemos dividir ambos por -1 e concluímos que $-1 = 1$, o que é absurdo. Logo, temos de concluir que $-1 \times -1 = 1$, que é

Cronologia

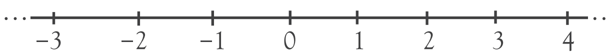
1572
Rafael Bombelli faz cálculos com números imaginários

1777
Euler é o primeiro a utilizar o símbolo *i* para representar a raiz quadrada de -1

positivo. O mesmo raciocínio pode ser feito para outros números negativos além do –1, e assim, quando qualquer número real é elevado ao quadrado, o resultado nunca pode ser negativo.

Isto causou um impasse no início dos números complexos, no século XVI. Quando foi ultrapassado, a resposta libertou os matemáticos dos grilhões dos números habituais e abriu um vasto campo de investigação inimaginável até então. O desenvolvimento dos números complexos é a «completação dos números reais» para um sistema naturalmente mais perfeito.

A raiz quadrada de –1 Já vimos que, limitando-nos à recta dos números reais,



não há raiz quadrada de –1, dado que nenhum quadrado de um número pode ser negativo. Se continuarmos a pensar nos números apenas na recta dos números reais, podemos desistir, continuar a chamá-los imaginários, ir tomar chá com os filósofos, e não ter mais nada que fazer com eles. Ou podemos dar o passo arrojado de aceitar $\sqrt{-1}$ como uma nova entidade, que representamos por i.

Com este único acto mental, os números imaginários existem. O que são, não sabemos, mas acreditamos na sua existência. Pelo menos, sabemos que $i^2 = -1$. Logo, no nosso novo sistema de números temos todos os nossos antigos amigos, como os números reais 1, 2, 3, 4, π, e, $\sqrt{2}$ e $\sqrt{3}$, bem como alguns novos, envolvendo i como $1 + 2i$, $-3 + i$, $2 + 3i$, $1 + i\sqrt{2}$, $\sqrt{3} + 2i$, e $+\pi i$, etc.

Este importante passo em matemática aconteceu no início do século XIX, quando nos libertámos da recta numérica unidimensional num novo e invulgar plano bidimensional.

Somar e multiplicar Agora que temos números complexos na cabeça, números com a forma $a + bi$, o que é que podemos fazer com eles? Tal como com os números reais, eles podem ser somados e multiplicados entre si. Somamo-los somando as suas partes. Assim, $2 + 3i$ somado com $8 + 4i$ dá $(2 + 8) + (3 + 4)i$ com o resultado $10 + 7i$.

A multiplicação é quase tão simples quanto isso. Se quisermos multiplicar $2 + 3i$ por $8 + 4i$, multiplicamos primeiro cada par de símbolos um pelo outro

Construir $\sqrt{-1}$

Até os engenheiros, uma espécie muito prática, encontraram utilização para os números complexos. Quando Michael Faraday descobriu a corrente alternada na década de 1830, os números imaginários ganharam uma realidade física. Neste caso, a letra j é usada para representar $\sqrt{-1}$, em vez de i porque i significa corrente eléctrica.

1806
A representação de Argand utilizando um diagrama conduz ao nome «diagrama de Argand»

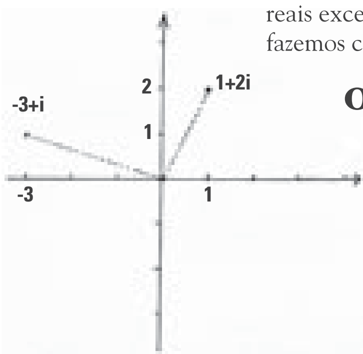
1811
Carl Friedrich Gauss trabalha com funções de variável complexa

1837
William R. Hamilton trata os números complexos como pares ordenados de números reais

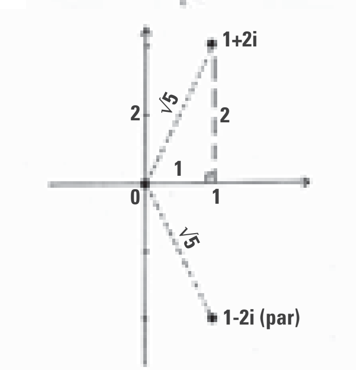
$$(2 + 3i) \times (8 + 4i) = (2 \times 8) + (2 \times 4i) + (3i \times 8) + (3i \times 4i)$$

e somamos os termos resultantes, 16, $8i$, $24i$ e $12i^2$ (neste último termo, substituímos i^2 por -1). O resultado desta multiplicação é então $(16 - 12) + (8i + 24i)$, que é o número complexo $4 + 32i$.

Com os números complexos, todas as regras aritméticas são satisfeitas. A subtracção e a divisão são sempre possíveis (excepto para o número complexo $0 + 0i$, mas também não é possível para 0 nos números reais). De facto, os números complexos gozam de todas as propriedades dos números reais excepto uma. Não podemos dividi-los em positivos e negativos como fazemos com os números reais.



O diagrama de Argand A bidimensionalidade dos números complexos é claramente vista se os representarmos num diagrama. Os números complexos $-3 + i$ e $1 + 2i$ podem ser representados no chamado diagrama de Argand. Esta forma de representar números complexos recebeu o nome do matemático suíço Jean Robert Argand, embora outros tenham tido ideias semelhantes mais ou menos ao mesmo tempo.



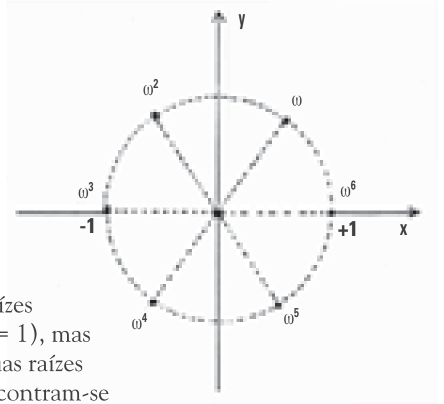
Mesmo os números complexos têm um «par» oficialmente chamado «conjugado». O par de $1 + 2i$ é $1 - 2i$, determinado por troca de sinal do segundo componente. O par de $1 - 2i$, pela mesma razão, é $1 + 2i$, logo é um verdadeiro emparelhamento.

Somando e multiplicando pares uns pelos outros, obtemos sempre um número real. No caso de somarmos $1 + 2i$ e $1 - 2i$, obtemos 2 e, multiplicando, obtemos 5. Esta multiplicação é mais interessante. A resposta 5 é o quadrado do «comprimento» do número complexo $1 + 2i$ e é igual ao comprimento do seu par. Posto de outra forma, podemos definir o comprimento de um número complexo como:

$$\textit{comprimento de } w = \sqrt{(w \times \textit{par de } w)}$$

Verificando-o para $-3 + i$, constatamos que o comprimento de $(-3 + i) = \sqrt{(-3 + i \times -3 - i)} = \sqrt{(9 + 1)}$, e assim o comprimento de $(-3 + i) = \sqrt{10}$

O fim do misticismo dos números complexos deve-se em grande parte a Sir William Rowan Hamilton, famoso matemático irlandês do século XIX. Hamilton reconheceu que i não é realmente necessário para a teoria e que funciona apenas como marcador e pode ser deitado fora. Hamilton considerou um número complexo como um «par ordenado» de números reais (a, b) exibindo a sua qualidade bidimensional e não invocando o místico $\sqrt{-1}$. Livre do i, a adição torna-se



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

e, um pouco menos óbvia, a multiplicação é

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

A perfeição do sistema de números complexos torna-se clara quando pensamos nas chamadas «n raízes da unidade» (para os matemáticos, «unidade» significa «um»). São as soluções da equação $z^n = 1$. Tomemos $z^6 = 1$ como exemplo. Existem duas raízes $z = 1$ e $z = -1$ na recta dos números reais (porque $1^6 = 1$ e $(-1)^6 = 1$), mas onde estão as outras, se devem ser seguramente seis? Como as duas raízes reais, todas as seis raízes têm uma unidade de comprimento e encontram-se numa circunferência centrada na origem e com uma unidade de raio.

Há mais. Se olharmos para $w = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ que é a raiz no primeiro quadrante, as raízes sucessivas (movendo-se no sentido anti-horário) são w^2, w^3, w^4, w^5, $w^6 = 1$ e encontram-se nos vértices de um hexágono regular. De uma forma geral, as n raízes da unidade estarão na circunferência e serão os cantos ou «vértices» de uma forma ou polígono regular de n lados.

Estender os números complexos Logo que tiveram os números complexos, os matemáticos procuraram instintivamente generalizações. Os números complexos são bidimensionais, mas o que é que o 2 tem de especial? Durante anos, Hamilton procurou construir números tridimensionais trabalhando numa forma de os somar e multiplicar, mas só teve êxito quando passou para quatro dimensões. Pouco depois, estes números com 4 dimensões foram generalizados para 8 dimensões (os chamados números de Cayley). Muitos se questionaram sobre números com 16 dimensões como uma possível continuação da história, mas, 50 anos depois do feito memorável de Hamilton, provou-se que eles eram uma impossibilidade.

a ideia resumida
Números não reais com utilizações reais

09 Primos

A matemática é uma matéria de tal maneira extensa, atravessando todas as avenidas da iniciativa humana, que por vezes pode parecer esmagadora. Ocasionalmente, temos de voltar às bases. Isto significa invariavelmente voltar aos números naturais, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, ... Será que podemos ser mais básicos do que isto?

Bem, 4 = 2 × 2, por isso podemos decompô-lo em componentes primários. Podemos decompor qualquer outro número? Na realidade, aqui estão mais alguns: 6 = 2 × 3, 8 = 2 × 2 × 2, 9 = 3 × 3, 10 = 2 × 5, 12 = 2 × 2 × 3. Estes números são compostos porque são construídos a partir dos muito básicos 2, 3, 5, 7, ... Os «números não decomponíveis» são os números 2, 3, 5, 7, 11, 13, ... Estes são os números primos, ou simplesmente primos. Um primo é um número que só é divisível por 1 e por si próprio. Poderíamos pensar que o próprio 1 é um número primo. De acordo com a definição, devia ser, e na realidade muitos eminentes matemáticos do passado trataram o 1 como primo, mas os matemáticos de hoje iniciam os primos com o 2. Isso permite que os teoremas sejam enunciados de forma elegante. Para nós, o número 2 também é o primeiro primo.

Podemos sublinhar os primos entre os primeiros números naturais: 1, $\underline{2}$, $\underline{3}$, 4, $\underline{5}$, 6, $\underline{7}$, 8, 9, 10, $\underline{11}$, 12, $\underline{13}$, 14, 15, 16, $\underline{17}$, 18, $\underline{19}$, 20, 21, 22, $\underline{23}$, . . . O estudo dos números primos conduz-nos à base de tudo. Os números primos são importantes porque são os «átomos» da matemática. Tal como os elementos básicos da química de que todos os outros compostos químicos derivam, os números primos podem construir os compostos matemáticos.

O resultado matemático que consolida tudo isto tem o grandioso nome de «teorema da decomposição em números primos». Segundo ele, todos os números inteiros maiores do que 1 podem ser representados pelo produto de números primos de uma única forma. Vimos que 12 = 2 × 2 × 3 e não há outra maneira de fazê-lo com primos. Isto é muitas vezes representado em notação de

Cronologia

300 a.C.
Os *Elementos* de Euclides provam que há infinitos números primos

230 a.C.
Eratóstenes de Cirene descreve um método de distinguir os números primos entre os números inteiros

potências: $12 = 2^2 \times 3$. Outro exemplo: 6 545 448 pode escrever-se $2^3 \times 3^5 \times 7 \times 13 \times 37$.

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

Descobrir os primos Infelizmente não há nenhuma fórmula para identificar os números primos, e parece não existir um padrão na sua colocação entre os números inteiros. Um dos primeiros métodos para os determinar foi desenvolvido por Eratóstenes de Cirene, um jovem contemporâneo de Arquimedes que passou grande parte da sua vida em Atenas. A sua determinação precisa do comprimento do equador foi muito admirada no seu tempo. Hoje é recordado pelo seu crivo para encontrar números primos. Eratóstenes imaginou os números naturais estendidos diante de si. Sublinhou 2 e eliminou todos os múltiplos de 2. Avançou para o 3, sublinhou-o e eliminou todos os múltiplos de 3. Prosseguindo desta forma, eliminou todos os números compostos. Os números sublinhados que ficaram para trás no crivo são os primos.

Podemos então predizer os primos, mas como é que decidimos se um dado número é ou não primo? Que dizer sobre 19 071 ou 19 073? Excepto para os primos 2 e 5, um número primo tem de terminar em 1, 3, 7 ou 9, mas esse requisito não é suficiente para que um número seja primo. É difícil saber quando é que um número grande terminado em 1, 3, 7 ou 9 é primo ou não sem se tentarem os factores possíveis. A propósito, $19\,071 = 3^2 \times 13 \times 163$ não é primo, mas 19 073 é.

Outro desafio tem sido descobrir padrões na distribuição dos primos. Vejamos quantos números primos existem em cada intervalo de 100 entre 1 e 1000.

Intervalo	1–100	101–200	201–300	301–400	401–500	501–600	601–700	701–800	801–900	901–1000	1–1000
Número de primos	25	21	16	16	17	14	16	14	15	14	168

Em 1792, apenas com 15 anos, Carl Friedrich Gauss sugeriu a fórmula $P(n)$ para estimar o número de números primos menores do que um dado número n (actualmente conhecida como teorema dos números primos). Para $n = 1000$, a fórmula dá um valor aproximado de 172. O verdadeiro número de números primos, 168, é menor do que esta estimativa. Presumiu-se sempre que esta era a situação para qualquer valor de n, mas é frequente que os primos nos surpreendam e já foi mostrado que, para

1742
Goldbach conjectura que cada número par (maior que 2) é a soma de dois primos

1896
O teorema dos números primos sobre a distribuição dos primos é provado

1966
Chen Jingrun quase confirma a conjectura de Goldbach

$n = 10^{371}$ (um número enorme representado com um 1 seguido de 371 zeros), o real número de primos *excede* a estimativa. De facto, nalgumas zonas dos números naturais, a diferença entre a estimativa e o número real oscila entre defeito e excesso.

Quantos? Existem infinitos números primos. Euclides declarou nos seus *Elementos* (Livro 9, Proposição 20) que «os números primos são mais do que qualquer número grande atribuído aos números primos». A magnífica prova de Euclides desenvolve-se assim:

> Suponha-se que P é o maior primo, e considere-se o número $N = (2 \times 3 \times 5 \times \ldots \times P) + 1$. N ou é primo ou não é. Se N é primo, obtivemos um primo maior do que P, o que contradiz a suposição. Se N não é primo, então tem de ser divisível por algum primo, digamos p, que é um de $2, 3, 5, \ldots, P$. Isto significa que p divide $N - (2 \times 3 \times 5 \times \ldots \times P)$. Mas este número é igual a 1, logo p divide 1. Isto não é possível, visto que todos os primos são maiores do que 1. Então, qualquer que seja a natureza de N, chegamos a uma contradição. A nossa suposição inicial de existir um primo P maior que todos os outros é falsa. *Conclusão:* O número de primos não tem limite.

O facto de os primos se estenderem até ao infinito não impediu que se procurasse arduamente o maior primo conhecido. Um que obteve recentemente o recorde é o enorme primo de Mersenne $2^{24036583} - 1$, que é aproximadamente $10^{7235332}$ ou um número começando com 1 seguido de 7 235 732 zeros.

O desconhecido «O problema dos primos gémeos» e a famosa «conjectura de Goldbach» são áreas proeminentes e desconhecidas relativas aos primos.

Os números primos gémeos são pares de primos consecutivos separados apenas por um número par. Os primos gémeos entre 1 e 100 são 3, 5; 5, 7; 11, 13; 17, 19; 29, 31; 41, 43; 59, 61; 71, 73. Do ponto de vista numérico, é sabido que existem 27 412 679 gémeos menores do que 10^{10}. Isto significa que números pares com gémeos, como 12 (tendo os gémeos 11 e 13) constituem apenas 0,274% dos números do intervalo. Existe um número infinito de primos gémeos? Seria curioso que não, mas ninguém até agora conseguiu prová-lo.

Christian Goldbach conjecturou que:

> *Todo o número par maior do que 2 é a soma de dois números primos*.

Por exemplo, 42 é um número par e podemos representá-lo como 5 + 37. O facto de também podermos representá-lo por 11 + 31, 13 + 29 ou 19 + 23 não é importante – só precisamos de *uma* forma. A conjectura é verdadeira para um enorme intervalo de números, mas nunca foi provada na generalidade. No entanto, têm sido feitos progressos, e há quem pressinta que a prova não está longe. O matemático chinês Chen Jingrun deu um grande passo. O seu teorema enuncia que qualquer número par suficientemente grande pode ser representado como a soma de dois primos *ou* como a soma de um primo com um *semi*-primo (um número que é o produto de dois primos).

O grande teórico da teoria dos números Pierre de Fermat provou que os primos da forma $4k + 1$ podem ser expressos como a soma de dois quadrados de uma e uma só maneira (*i.e.*, $17 = 1^2 + 4^2$), enquanto os da forma $4k + 3$ (como 19) não podem de todo ser representados como a soma de dois quadrados. Joseph Lagrange também demonstrou um famoso teorema matemático sobre quadrados: *todo* o número inteiro positivo é a soma de quatro quadrados. Assim, por exemplo, $19 = 1^2 + 1^2 + 1^2 + 4^2$. Têm sido exploradas potências mais elevadas e escrito livros cheios de teoremas, mas muitos problemas mantêm-se.

Descrevemos os números primos como «átomos da matemática».
Mas poderemos perguntar-nos se, ao contrário dos físicos, que foram além dos átomos para unidades mais fundamentais, como os quarks, a matemática estagnou. Se nos limitarmos aos números naturais, o 5 é um número primo e sempre o será.
Mas Gauss fez uma descoberta de longo alcance, que para alguns primos, como o 5, $5 = (1 - 2i) \times (1 + 2i)$, em que $i = \sqrt{-1}$ do sistema de números imaginários. Como produto de dois números inteiros gaussianos, o 5 e números como ele não são tão indecomponíveis como se supunha.

O número do numerólogo

Uma das áreas mais desafiantes da teoria dos números é o «problema de Waring». Em 1770 Edward Waring, professor em Cambridge, colocou problemas respeitantes à representação de que números inteiros como somas de potências. Neste cenário, a arte mágica da numerologia mistura-se com a ciência clínica da matemática sob a forma dos números primos, somas de quadrados e somas de cubos.
Em numerologia, pensemos no culto sem rival do número 666, o «número da besta» no livro bíblico do Apocalipse, que tem algumas propriedades inesperadas. É a soma dos quadrados dos primeiros 7 primos:

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

Os numerólogos também são entusiásticos a apontar que é a soma palindrómica de cubos e, se isso não for suficiente, a pedra angular 6^3 no centro é a forma abreviada de 6 x 6 x 6:

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

O número 666 é verdadeiramente o «número do numerólogo».

a ideia resumida
Os átomos da matemática

10 Números perfeitos

Em matemática, a busca da perfeição tem conduzido os seus pretendentes a lugares distintos. Há quadrados perfeitos, e utilizamos a expressão, não no sentido estético, mas sobretudo para avisar que existem quadrados imperfeitos. Alguns números têm poucos divisores e outros têm muitos. Mas, como na história dos três ursos, em que só uma cama servia na perfeição, há números que são «simplesmente certos». Quando a adição dos divisores de um número iguala o próprio número, diz-se que ele é perfeito.

O filósofo grego Espeusipo, que recebeu a direcção da Academia do seu tio Platão, declarou que os pitagóricos acreditavam que o número 10 tinha as credenciais certas para a perfeição. Porquê? Porque o número de primos entre 1 e 10 (nomeadamente 2, 3, 5, 7) igualava o dos não-primos (4, 6, 8, 9) e este era o menor número com esta propriedade. Há quem tenha uma ideia singular da perfeição.

Parece que na realidade os pitagóricos tinham um conceito mais rico de número perfeito. As propriedades matemáticas de um número perfeito foram delineadas por Euclides nos *Elementos* e estudadas em profundidade por Nicómaco 400 anos mais tarde, conduzindo aos números amigos e mesmo aos números sociáveis. Estas categorias foram definidas em termos das relações entre eles e os seus divisores. A determinada altura, surgiu-lhes a ideia de números superabundantes e deficientes e isso conduziu-os ao seu conceito de perfeição.

Determina-se se um número é superabundante pelos seus divisores e consoante desempenha um papel na relação entre multiplicação e adição. Tomemos

Cronologia

525 a.C.
Os pitagóricos são associados aos números perfeitos e aos números abundantes

300 a.C.
O Livro 9 dos *Elementos* de Euclides discute os números perfeitos

100
Nicómaco de Gerasa dá uma classificação dos números baseada nos números perfeitos

o número 30 e consideremos os seus divisores, ou seja todos os números em que ele se divide exactamente e são menores do que 30. Para um número tão pequeno como 30, podemos constatar que os divisores são 1, 2, 3, 5, 6, 10 e 15. Somando estes divisores, obtemos 42. O número 30 é superabundante porque a soma dos seus divisores (42) é maior que o próprio 30.

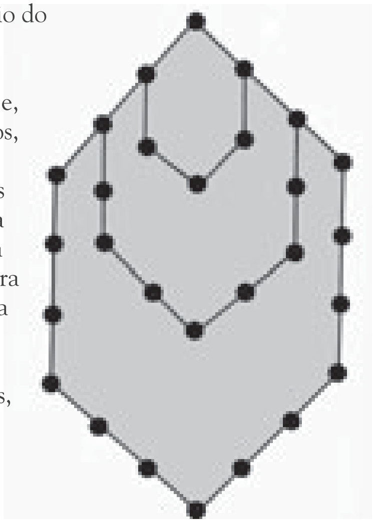
Classe	1	2	3	4	5	6	7
Número perfeito	6	28	496	8128	33 550 336	8 589 869 056	137 438 691 328

Os primeiros números perfeitos

Um número é deficiente se o oposto for verdade – se a soma dos seus divisores for menor do que ele próprio. Logo, o número 26 é deficiente porque os seus divisores 1, 2 e 13 somam apenas 16, o que é menor do que 26. Os números primos são muito deficientes, porque a soma dos seus divisores é sempre apenas 1.

Um número que não seja superabundante nem deficiente é perfeito. A soma dos divisores de um número perfeito é igual ao próprio número. O primeiro número perfeito é o 6. Os seus divisores são 1, 2 e 3 e, quando os somamos, obtemos 6. Os pitagóricos ficaram tão encantados com o número 6 e com a forma como as suas partes se ajustavam, que lhe chamaram «saúde, casamento e beleza». Há outra história relacionada com o 6, contada por Santo Agostinho (354-430). Este acreditava que a perfeição do 6 já existia antes do início do mundo, que teria sido criado em 6 dias porque o número era perfeito.

O número perfeito seguinte é o 28. Os seus divisores são 1, 2, 4, 7 e 14 e, quando os somamos, obtemos 28. Estes dois primeiros números perfeitos, 6 e 28, são bastante especiais para o conhecimento dos números perfeitos, dado que pode provar-se que todos os números perfeitos pares terminam em 6 ou em 28. Depois do 28, há que esperar até ao 496 para o número perfeito seguinte. É fácil verificar que ele é realmente a soma dos seus divisores: 496 = 1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248. Para o número perfeito seguinte, teremos de começar a entrar na estratosfera numérica. Os primeiros cinco já eram conhecidos no século XVI, mas ainda não se sabe se existe um que seja maior que todos, ou se eles continuam infinitamente. A opinião geral é de que, tal como os primos, eles continuam para sempre.



1603

Pietro Cataldi determina o sexto e o sétimo números perfeitos, $2^{16}(2^{17}-1) = 8\,589\,869\,056$ e $2^{17}(2^{18}-1) = 137\,438\,691\,328$

2006

O grande projecto de pesquisa de primos encontra o 44.º primo de Mersenne (com quase um milhão de dígitos) e pode ser gerado um novo número perfeito

Expoente	Resultado	Subtrair 1 (número de Mersenne)	Número primo?
2	4	3	primo
3	8	7	primo
4	16	15	não primo
5	32	31	primo
6	64	63	não primo
7	128	127	primo
8	256	255	não primo
9	512	511	não primo
10	1024	1023	não primo
11	2048	2047	não primo
12	4096	4095	não primo
13	8192	8191	primo
14	16384	16383	não primo
15	32768	32767	não primo

Os pitagóricos gostavam das ligações geométricas. Se tivermos um número perfeito de contas, elas podem ser dispostas num colar hexagonal. No caso do 6, é o hexágono mais simples com as contas colocadas nos seus cantos, mas para números perfeitos mais elevados temos de adicionar subcolares menores dentro do grande.

Números de Mersenne A chave para construir números perfeitos é uma colecção de números cujo nome vem do padre Marin Mersenne, um monge francês que estudou num colégio jesuíta com René Descartes. Ambos estavam interessados em encontrar números perfeitos. Os números de Mersenne são construídos a partir das potências de 2, os números duplicados 2, 4, 8, 16, 32, 64, 128, 256, ..., subtraindo depois 1. Um número de Mersenne tem a forma $2^n - 1$. Embora sejam sempre ímpares, nem sempre são primos. Mas são estes números de Mersenne que também são primos, que podem ser usados para construir os números perfeitos.

Mersenne sabia que, se a potência *não* fosse um número primo, o número de Mersenne também não podia ser primo, como se vê para os expoentes 4, 6, 8, 9, 10, 12, 14 e 15 na tabela. Os números de Mersenne só poderão ser primos se o expoente for um número primo. Mas será suficiente? Para os primeiros casos, obtemos 3, 7, 31 e 127, que são todos primos. Então será genericamente verdadeiro que um número de Mersenne formado com um expoente primo também é primo?

Muitos matemáticos anteriores ao ano 1500 pensaram que assim era. Mas os números primos não estão aprisionados pela simplicidade, e descobriu-se que, para o expoente 11 (um número primo), $2^{11} - 1 = 2047 = 23 \times 89$; consequentemente, não é primo. Esta parece ser a regra. Os números de

Apenas bons amigos

O matemático não costuma ser dado à mística dos números, mas a numerologia não morreu. Os números amigos vieram depois dos números perfeitos, embora possam ter sido conhecidos pelos pitagóricos. Mais tarde, tornaram-se úteis na elaboração dos horóscopos sentimentais, em que as suas propriedades matemáticas se traduzem na natureza da ligação etérea. Os números 220 e 284 são números amigos. Porquê? Bem, os divisores de 220 são 1, 2, 4, 5, 10, 11, 20, 22, 44, 55 e 110 e, se os somarmos, obtemos 284. Adivinharam. Se determinarmos os divisores de 284 e os somarmos, obtemos 220. É uma verdadeira amizade.

Mersenne $2^{17} - 1$ e $2^{19} - 1$ são ambos primos, mas $2^{23} - 1$ não é primo, porque

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

Construção A combinação dos trabalhos de Euclides e de Euler fornece uma fórmula que permite que sejam gerados os números perfeitos pares: n é qualquer número perfeito par se e só se $n = 2^{p-1}(2^p - 1)$, em que $2^p - 1$ é um primo de Mersenne.

Por exemplo, $6 = 2^1(2^2 - 1)$, $28 = 2^2(2^3 - 1)$ e $496 = 2^4(2^5 - 1)$. Esta fórmula para calcular números perfeitos pares significa que podemos gerá-los se determinarmos os primos de Mersenne. Os números perfeitos têm desafiado tanto pessoas como máquinas e continuarão a fazê-lo de uma forma que os antigos não teriam previsto. Peter Barlow, que elaborou as tabelas de Barlow, escreveu, no início do século XIX, que pensava que ninguém iria além da determinação do número perfeito de Euler

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

por isso não fazer grande sentido. Barlow não podia prever o poder dos modernos computadores ou a insaciável necessidade dos matemáticos de encontrarem novos desafios.

Os números primos de Mersenne

Determinar os primos de Mersenne não é fácil. Ao longo dos séculos, muitos matemáticos contribuíram para a lista, que tem uma história matizada feita de uma combinação de erros e êxitos. O grande Leonhard Euler contribuiu com o oitavo primo de Mersenne, $2^{31} - 1 = 2\,147\,483\,647$, em 1732. Encontrar o 23.º primo de Mersenne, $2^{11213} - 1$, em 1963, foi motivo de orgulho para o departamento de matemática da Universidade de Illinois, nos EUA, que o anunciou ao mundo no seu selo de correio. Mas, com os computadores potentes, a indústria dos números primos de Mersenne avançou e, no final dos anos 70, Laura Nicklel e Landon Noll, dois estudantes de liceu, descobriram em conjunto o 25.º primo de Mersenne, e Noll o 26.º. Até agora, foram descobertos 45 números primos de Mersenne.

Números perfeitos ímpares Ninguém sabe se alguma vez descobriremos um número perfeito ímpar. Descartes supunha que não, mas os peritos podem estar errados. O matemático inglês James Joseph Sylvester declarou que a existência de um número perfeito ímpar «seria como um milagre», porque teria de satisfazer muitas condições. Não é surpreendente que Sylvester estivesse duvidoso. Este é um dos problemas matemáticos mais antigos, mas, se existir um número perfeito ímpar, muito já sabemos sobre ele. Terá de ter pelo menos 8 divisores primos e distintos, um dos quais é maior que um milhão, e ter pelo menos 300 dígitos de comprimento.

a ideia resumida
A mística dos números

11 Os números de Fibonacci

No *Código da Vinci*, Dan Brown fez o seu curador assassinado Jacques Saunière deixar os primeiros oito termos de uma sequência de números como pista do seu destino. Foram necessárias as aptidões da criptógrafa Sophie Neveu para reorganizar os números 13, 3, 2, 21, 1, 1, 8 e 5 para lhes descobrir o significado. Bem-vindos à mais famosa sequência de números de toda a matemática.

A sequência de números inteiros de Fibonacci é:

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

A sequência é amplamente conhecida pelas suas múltiplas e intrigantes propriedades. A mais básica – na realidade o traço característico que a define – é que cada termo é a soma dos dois anteriores. Por exemplo, 8 = 5 + 3, 13 = 8 + 5, ..., 2584 = 1587 + 98, etc. Basta começarmos com os dois números 1 e 1 e conseguimos gerar a sequência de imediato. A sequência de Fibonacci encontra-se na natureza no número de espirais formadas pelo número de sementes nas espirais dos girassóis (por exemplo, 34 numa direcção, 55 na outra), e nas proporções das divisões e nas construções desenhadas pelos arquitectos. Os compositores de música clássica usaram-na como inspiração e pensa-se que a *Dance Suite* de Bartók está relacionada com a sequência. Na música contemporânea, Brian Transeau tem uma faixa no seu álbum *This Binary Universe* chamada «1618», em saudação à última razão dos números de Fibonacci, um número que discutiremos adiante.

Origens A sequência de Fibonacci aparece no *Liber Abaci* publicado por Leonardo de Pisa (Fibonacci) em 1202, mas os números já eram

Cronologia

1202
Leonardo de Pisa publica o *Liber Abaci* e os números de Fibonacci

1724
Daniel Bernoulli expressa os números da sequência de Fibonacci em termos do número de ouro

provavelmente conhecidos na Índia. Fibonacci colocou o seguinte problema da procriação de coelhos:

Os casais de coelhos adultos geram novos casais de coelhos jovens todos os meses. No início do ano, existe um casal jovem de coelhos. No final do primeiro mês, tornam-se adultos, no final do segundo mês o par adulto gerou um casal jovem. O processo de se tornarem adultos e gerarem novos casais continua. Miraculosamente, nenhum dos casais de coelhos morre.

Fibonacci queria saber quantos casais de coelhos haveria no fim do ano. As gerações podem ser vistas numa «árvore genealógica». Pensemos no número de casais no final de Maio (o quinto mês). Vemos que o número de casais é 8. Nesse nível da árvore genealógica, o grupo da esquerda

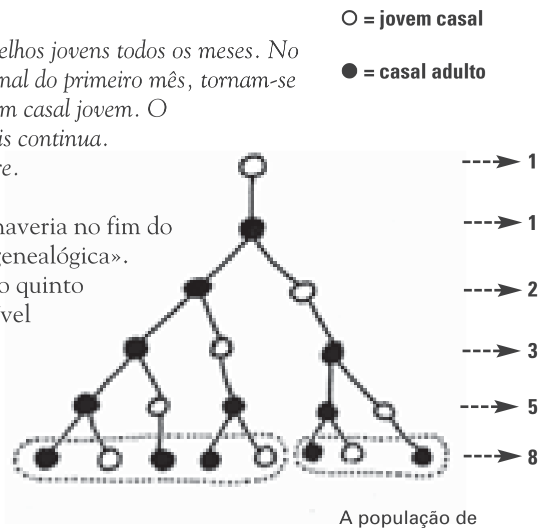
● ○ ● ● ○

é um duplicado da coluna completa em cima, e o grupo do lado direito

● ○ ●

é um duplicado da linha acima dessa. Isto mostra que o nascimento dos casais de coelhos segue a equação básica de Fibonacci:

$$\text{número depois de } n \text{ meses} = \text{número depois de } (n-1) \text{ meses} + \text{número depois de } (n-2) \text{ meses}$$



A população de coelhos

Propriedades Vejamos o que acontece se somarmos os termos da sequência:

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

1923
Bartók compõe a sua *Dance Suite*, que se acredita ter sido inspirada nos números de Fibonacci

1963
Fundação do *Fibonacci Quarterly*, jornal dedicado à teoria dos números da sequência de Fibonacci

2007
O escultor Peter Randall-Page cria a escultura «Seed», de 70 toneladas, baseada na sequência de Fibonacci, para o projecto Eden na Cornualha, no Reino Unido

O resultado de cada uma destas somas também formará uma sequência, que podemos colocar debaixo da sequência original, mas deslocada

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

A soma de *n* termos da sequência de Fibonacci acaba por ser o seguinte mais 1 número de Fibonacci menos 1, ou seja, se quisermos conhecer o resultado da soma 1 + 1 + 2 + ... + 987, basta-nos subtrair 1 a 2584 para obtermos 2583. Se os números forem somados alternadamente falhando alguns termos, como 1 + 2 + 5 + 13 + 34, obtemos 55, também ele um número de Fibonacci. Se usarmos outra alternância, tal como 1 + 3 + 8 + 21 + 55, a resposta é 88, que é um número de Fibonacci menos 1.

Os quadrados da sequência de Fibonacci também são interessantes. Obtemos uma nova sequência multiplicando cada número de Fibonacci por si próprio e somando-os.

Fibonacci	1	1	2	3	5	8	<u>13</u>	<u>21</u>	34	55 . . .
Quadrados	1	1	4	9	25	64	169	441	1156	3025 . . .
Soma de quadrados	1	2	6	15	40	104	<u>273</u>	714	1870	4895 . . .

Neste caso, somar todos os quadrados até ao *n*-ésimo termo é o mesmo que multiplicar o *n*-ésimo termo da sequência original de Fibonacci pelo seguinte. Por exemplo,

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

Os números de Fibonacci também aparecem onde menos se espera. Suponhamos que temos na carteira várias moedas de €1 e €2 e queremos contar o número de maneiras em que podemos tirar as moedas da carteira para perfazer um determinado montante expresso em euros. Neste problema, a ordem das acções é importante. O valor de €4, conforme tiramos as moedas da carteira, pode ser qualquer das seguintes maneiras: 1 + 1 + 1 + 1; 2 + 1 + 1; 1 + 2 + 1; 1 + 1 + 2; e 2 + 2. Ao todo, existem 5 maneiras – e isto corresponde ao quinto número de Fibonacci. Se tirarmos €20, existem 6765 maneiras de tirarmos moedas de €1 e €2, correspondendo ao 21.º número de Fibonacci! Isto mostra o poder de ideias matemáticas simples.

O número de ouro Se observarmos a razão dos termos construída a partir da sequência de Fibonacci dividindo cada termo pelo antecedente, encontramos outra propriedade notável dos números de Fibonacci. Vamos fazê-lo para alguns termos: 1, 1, 2, 3, 8, 13, 21, 34, 55.

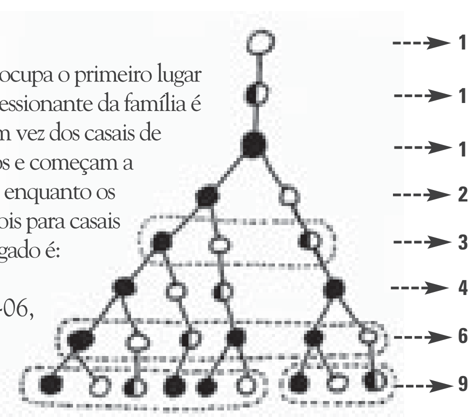
1/1	2/1	3/2	5/3	8/5	13/8	21/13	34/21	55/34
1000	2000	1500	1333	1600	1625	1615	1619	1617

Rapidamente as razões se aproximam de um valor conhecido como número de ouro, um número famoso designado pela letra grega ϕ, que está entre as mais importantes constantes matemáticas, como π e e, e tem o valor exacto de

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

e pode ser aproximado ao decimal 1,618033988... Na verdade, é possível mostrar que cada número de Fibonacci pode representar-se em termos de ϕ.

Apesar da variedade de conhecimentos que se tem sobre a sequência de Fibonacci, ainda existem muitas questões por responder. Os primeiros primos na sequência de Fibonacci são 2, 3, 5, 13, 89, 233, 1597, mas não se sabe se existem infinitos primos na sequência de Fibonacci.



Cabeças de gado

Semelhanças de família A sequência de Fibonacci ocupa o primeiro lugar numa vasta família de sequências semelhantes. Um membro impressionante da família é um que podemos associar com o problema das cabeças de gado. Em vez dos casais de coelhos de Fibonacci que evoluem num mês de jovens para adultos e começam a reproduzir-se, há um estado intermédio no processo de maturação, enquanto os casais de gado evoluem de casais jovens para casais imaturos e depois para casais adultos. Só os casais adultos podem reproduzir-se. A sequência do gado é:

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

Por conseguinte, a geração salta um valor, para que, por exemplo, 41 = 28 + 13 e 60 = 41 + 19. Esta sequência tem propriedades semelhantes à sequência de Fibonacci. Para a sequência do gado, a razão obtida pela divisão entre um termo e o precedente aproxima-se de um limite representado pela letra grega *psi*, ψ, em que

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

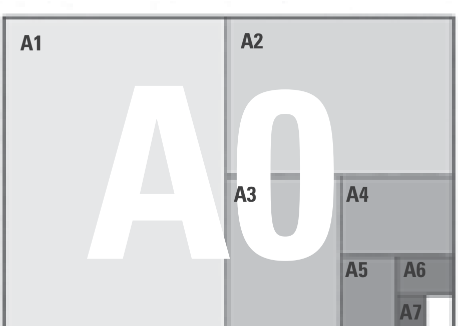
Este é conhecido como o «supernúmero de ouro»

a ideia resumida
O *Código da Vinci* descodificado

12 Os rectângulos de ouro

Os rectângulos estão por todo o lado à nossa volta – edifícios, fotografias, janelas, portas, até este livro. Os rectângulos estão presentes na comunidade artística – Piet Mondrian, Ben Nicholson e outros, que avançaram para a abstracção, todos usaram um tipo ou outro de rectângulo. Qual será o mais bonito de todos? Um longo e estreito «rectângulo de Giacometti» ou aquele que é quase um quadrado? Ou um rectângulo entre esses extremos?

Será que a questão chega a fazer sentido? Alguns pensam que sim, e acreditam que alguns triângulos em particular são mais «ideais» do que outros. Destes, talvez o rectângulo de ouro tenha tido maior preferência. De entre todos os rectângulos que se podem escolher, pelas suas diferentes proporções – porque é disso que se trata –, o rectângulo de ouro é um rectângulo muito especial que tem inspirado artistas, arquitectos e matemáticos. Comecemos por observar outros rectângulos.



Papel matemático Se pegarmos numa folha de papel A4, cujas dimensões são um lado menor de 210 mm e um lado maior de 297 mm, a razão entre o comprimento e a largura será 297/210, ou aproximadamente 1,4142. Qualquer papel internacional de tamanho A com o lado menor igual a b, o lado maior será sempre $1{,}4142 \times b$. Assim para A4, $b = 210$ mm, enquanto para A5, $b = 148$ mm. O sistema da fórmula A usado para as dimensões do papel tem uma propriedade muito desejável, que não ocorre para dimensões de papel arbitrárias. Se uma folha de papel de tamanho A

Cronologia

cerca de 300 a.C.
As razões extrema e média são publicadas nos *Elementos* de Euclides

1202
Leonardo de Pisa publica o *Liber Abaci*

for dobrada ao meio, os dois rectângulos menores que se obtêm estão em proporção directa com o rectângulo maior. São duas versões menores do mesmo rectângulo.

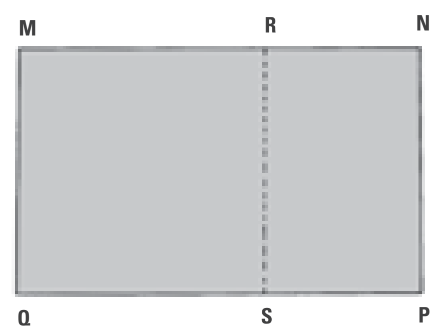
Desta forma, uma folha A4 dobrada em duas partes gera duas folhas A5. De forma semelhante, uma folha A5 gera duas folhas A6. Ao inverso, uma folha de papel A3 é feita com duas folhas A4. Quanto menor for o número de A, maior será a folha de papel. Como sabemos que o número 1,4142 fará esta habilidade? Vamos dobrar um rectângulo, mas desta vez um rectângulo cujo comprimento do lado maior não conhecemos. Se tomarmos a largura do rectângulo como 1 e escrevermos o comprimento como x, a razão entre o comprimento e a largura será $x/1$. Se agora dobrarmos o rectângulo, a razão entre o comprimento e a largura será $1/\tfrac{1}{2}x$, que é o mesmo que $2/x$. O objectivo dos tamanhos A é que as razões mantenham a mesma proporção, logo obtemos a equação $x/1 = 2/x$ ou $x2 = 2$. O valor de x é portanto $\sqrt{2}$, que é aproximadamente 1,4142.

Matemática de ouro O rectângulo de ouro é diferente, mas apenas *ligeiramente* diferente. Desta vez, o rectângulo é dobrado pela linha *RS* do diagrama, de forma que os pontos *MRSQ* sejam os cantos de um *quadrado*.

A propriedade fundamental do rectângulo de ouro é que o rectângulo que sobra, *RNPS*, é proporcional ao rectângulo maior – o que sobra deve ser uma mini-réplica do rectângulo maior.

Como anteriormente, diremos que a largura $MQ = MR$ do rectângulo maior tem 1 unidade de comprimento, enquanto chamamos x ao comprimento do lado maior *MN*. A razão entre o comprimento e a largura é de novo $x/1$. Desta vez, a largura do rectângulo menor *RNPS* é *MN* – MR, o que é $x - 1$, donde a razão entre o comprimento e a largura deste rectângulo é $1(x - 1)$. Igualando, obtemos a equação

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



que pode ser simplificada para $x^2 = x + 1$. Uma solução aproximada é 1,618. Podemos verificá-la facilmente. Se utilizarmos uma calculadora e multiplicarmos 1,618 por si próprio, obtemos 2,618, que é o mesmo que $x + 1 = 2{,}618$. Este número é a famosa proporção de ouro e é designado pela letra grega *phi*, ϕ. A sua definição e aproximação é dada por

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

e este número está relacionado com a sequência de Fibonacci e o problema dos coelhos (ver página 44).

1509
Paciola publica
A Proporção Divina

1876
Fechner escreve as suas experiências psicológicas para determinar o rectângulo mais estético

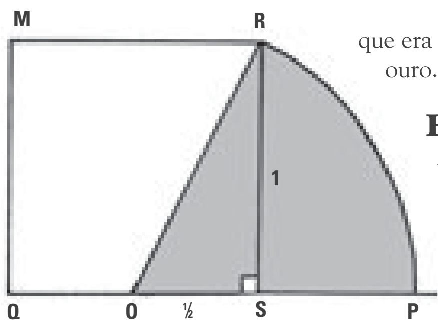
1975
A International Organization for Standardization (ISO) define o tamanho de papel A

Em busca do ouro Vejamos agora se conseguimos construir um rectângulo de ouro. Começamos com o nosso quadrado *MQSR* com lados iguais a 1 unidade e marcamos o ponto médio de *QS* como O. O comprimento de *OS* = ½ e, pelo teorema de Pitágoras (ver página 84), no triângulo *ORS*, *OR* =

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

Usando um compasso centrado em O, podemos desenhar o arco *RP* e teremos *OP* = *OR* = $\sqrt{5}/2$. Desta forma, chegamos a

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



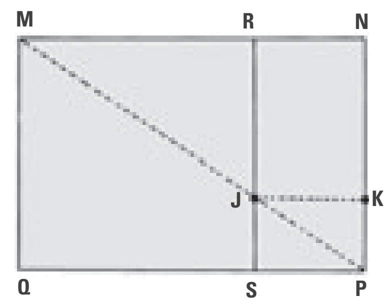
que era o pretendido: a «secção de ouro» ou o lado do rectângulo de ouro.

História Muito se tem dito sobre a proporção dourada ϕ. Assim que as suas propriedades matemáticas são percebidas, é possível encontrá-la em lugares inesperados, mesmo em lugares onde não está. Pior do que isto é o risco de afirmar que a proporção de ouro já lá estava antes do artefacto – que os músicos, arquitectos e artistas a tinham em mente na altura da criação. Este ponto fraco é chamado «numerismo de ouro». Passar de números para afirmações gerais sem qualquer evidência é um raciocínio perigoso de se fazer.

Pensemos no Pártenon, em Atenas. No tempo da sua construção, a proporção de ouro era conhecida, mas isso não significa que o Pártenon tenha sido baseado nela. De facto, na vista frontal do Pártenon a razão entre o comprimento e a altura (incluindo o frontão triangular) é 1,74, o que é próximo de 1,618, mas será suficientemente próximo para afirmar que a proporção de ouro foi a motivação? Alguns defendem que o frontão deve ser excluído do cálculo e, se isso for feito, a razão entre o comprimento e a altura será na realidade o número inteiro 3.

Em 1509, no livro *De divina proportione*, Luca Pacioli «descobriu» relações entre as características de Deus e as propriedades da proporção determinada por ϕ. Baptizou-a como «divina proporção». Pacioli foi um frade franciscano que escreveu livros de matemática influentes. É tido por alguns como o «pai da contabilidade», por ter popularizado o método da dupla entrada na contabilidade usada pelos mercadores venezianos. Outro crédito seu foi o de ter ensinado matemática a Leonardo da Vinci. Na Renascença, a secção de ouro adquiriu uma posição quase mística – o astrónomo Johannes Kepler descreveu-a como a «jóia preciosa» da matemática. Mais tarde, Gustav Fechner, um psicólogo experimentalista alemão, efectuou milhares de medições de formas rectangulares (cartas de jogar, livros, janelas) e verificou que a razão entre os lados que ocorria mais frequentemente era próxima de ϕ.

Le Corbusier era fascinado pelo rectângulo como elemento central nos projectos de arquitectura, particularmente pelo rectângulo de ouro. Colocava grande ênfase na harmonia e na ordem e encontrou-as na matemática. Via a arquitectura com olhos de matemático. Uma das suas bases era o sistema «modelador», a teoria das proporções. Na realidade era uma forma de gerar cadeias de rectângulos de ouro, formas que usava nos seus projectos. Le Corbusier inspirou-se em Leonardo da Vinci, que, por sua vez, tomara cuidadosas notas do arquitecto romano Vitrúvio, que forneceu as bases para as proporções que se encontram na figura humana.



Outras formas Também existe um rectângulo «superdourado», cuja construção tem semelhanças com a do rectângulo de ouro.

É assim que se constrói um rectângulo superdourado $MQPN$. Como antes, $MQSR$ é um quadrado cujo lado tem comprimento 1. Junta-se a diagonal MP e marca-se a intersecção com RS como o ponto J. A seguir, constrói-se a linha JK, que é paralela a RN com K em NP. Seja o comprimento de RJ y e o comprimento de MN x. Para qualquer rectângulo, $RJ/MR = NP/MN$, porque os triângulos MRJ e MNP são semelhantes, logo $y/1 = 1/x$, o que significa que $x \times y = 1$ e dizemos que x e y são «recíprocos» um do outro. Obtemos o rectângulo superdourado fazendo o rectângulo $RJKN$ proporcional ao rectângulo original $MQPN$, ou seja $y/(x - 1) = x/1$. Usando o facto de que $xy = 1$, podemos concluir que o comprimento do triângulo superdourado x é determinado resolvendo a equação «cúbica» $x^3 = x^2 + 1$, que é claramente semelhante à equação $x^2 = x + 1$ (a equação que determina o rectângulo de ouro). A equação cúbica tem uma solução real e positiva ψ (substituindo x pelo símbolo mais usual ψ) cujo valor é

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

o número associado à sequência do gado (ver página 47). Enquanto o rectângulo de ouro pode ser construído com régua e compasso, o rectângulo superdourado não pode.

a ideia resumida
Proporções divinas

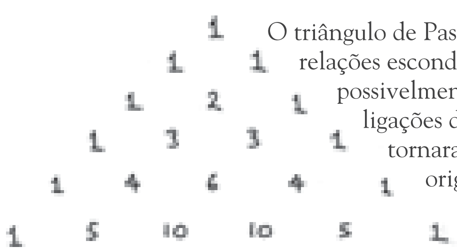
13 O triângulo de Pascal

O número 1 é importante, mas que dizer do 11? Também é interessante. Como também o é 11 × 11 = 121, 11 × 11 × 11 = 1331 e 11 × 11 × 11 × 11 = 14 641. Organizando-os, temos

11
121
1331
14641

Estas são as primeiras linhas do triângulo de Pascal. Mas onde é que o encontramos?

Colocando $11^0 = 1$ para finalizar, a primeira coisa a fazer é introduzir espaços entre os algarismos. Assim, 14 641 fica 1 4 6 4 1.



Triângulo de Pascal

O triângulo de Pascal é famoso em matemática pela sua simetria e as suas relações escondidas. Em 1653, Blaise Pascal assim pensou e observou que possivelmente não podia colocá-las todas num só trabalho. As muitas ligações do triângulo de Pascal com outros ramos da matemática tornaram-no um venerável objecto matemático, mas as suas origens podem ser encontradas muito antes. De facto, Pascal não inventou o triângulo que tem o seu nome – ele era conhecido dos sábios chineses do século XIII.

O padrão de Pascal é gerado a partir de cima. Começamos com 1 e colocamos dois 1 de cada lado na linha de baixo. Para construir as linhas seguintes, continuamos a colocar 1 nas extremidades de cada linha, enquanto os números interiores são obtidos pela soma dos dois números imediatamente acima. Para obter 6 na quinta linha, por exemplo, somamos 3 + 3 da linha acima.

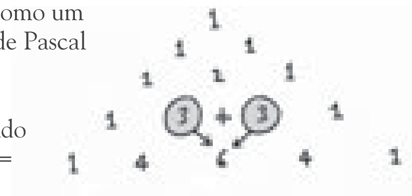
Cronologia

cerca de 500 a.C.

Fragmentos que evidenciam a existência do triângulo de Pascal em sânscrito

cerca de 1070

Omar Khayyam descobre o triângulo (que nalguns países tem o seu nome)



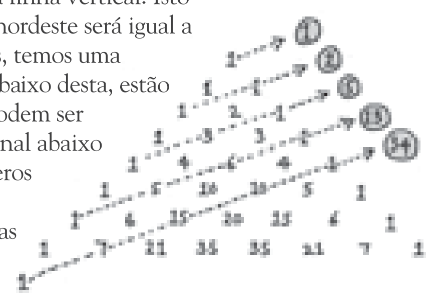
O matemático inglês G. H. Hardy disse que «um matemático, como um pintor ou um poeta, é um construtor de padrões» e o triângulo de Pascal tem padrões em alto grau.

Ligações com a álgebra O triângulo de Pascal é baseado na matemática real. Se resolvermos $(1 + x) \times (1 + x) \times (1 + x) = (1 + x)^3$, por exemplo, obtemos $1 + 3x + 3x^2 + x^3$. Se observarmos atentamente, verificamos que os números antes dos símbolos desta expressão coincidem com os números da linha correspondente do triângulo de Pascal. O esquema seguido é:

$(1 + x)^0$	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

Se somarmos os números de uma qualquer linha do triângulo de Pascal obtemos sempre uma potência de 2. Por exemplo, na quinta linha temos $1 + 4 + 6 + 4 + 1 = 16 = 2^4$. Isto pode ser obtido da coluna da esquerda fazendo $x = 1$.

Propriedades A primeira e a mais óbvia propriedade do triângulo de Pascal é a sua simetria. Se desenharmos uma linha vertical pelo meio do triângulo, ele tem uma «simetria em espelho» – é igual quer à esquerda quer à direita da linha vertical. Isto permite-nos falar de «diagonais» simples, porque uma diagonal nordeste será igual a uma diagonal noroeste. Por baixo da diagonal constituída por 1s, temos uma diagonal constituída pelos números naturais 1, 2, 3, 4, 5, 6, ... Abaixo desta, estão os números triangulares 1, 3, 6, 10, 15, 21, ... (os números que podem ser construídos a partir de pontos na forma de triângulos). Na diagonal abaixo temos os números tetraédricos 1, 4, 10, 20, 35, 56, ... Estes números correspondem ao tetraedro («triângulos tridimensionais», ou, se quisermos, o número de balas de canhão que podem ser colocadas em bases triangulares de tamanho crescente). E as «quase-diagonais»?



Quase-diagonais no triângulo de Pascal

Se somarmos os números nas linhas oblíquas do triângulo (que não são linhas horizontais nem diagonais), obtemos a sequência 1, 2, 5, 13, 34, ...

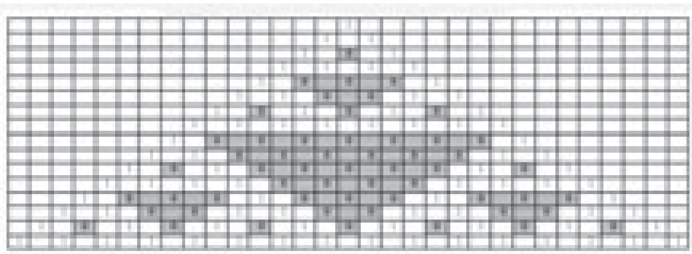
1303
Zhu Shijie define o triângulo de Pascal e mostra como somar determinadas sequências

1664
Publicação póstuma do trabalho de Pascal sobre as propriedades do triângulo

1714
Leibniz analisa o triângulo harmónico

Cada número é três vezes o anterior com o anterior a esse subtraído. Por exemplo, 34 = 3 × 13 – 5. Com base nisto, o próximo número na sequência será 3 × 34 – 13 = 89. Falhamos a outra «quase-diagonal» que começa com 1, 1 + 2 = 3, mas esta dá-nos a sequência 1, 3, 8, 21, 55, ... e estes são gerados pela mesma regra «3 vezes menos 1». Podemos portanto gerar o número seguinte nesta sequência, 3 × 55 – 21 = 144. Mas há mais. Se intercalarmos estas duas sequências de «quase-diagonais», obtemos os números de Fibonacci:

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



Números pares e ímpares no triângulo de Pascal

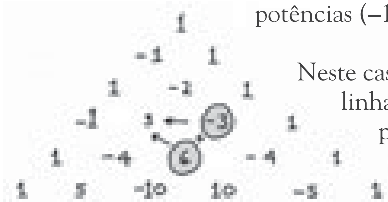
As combinações de Pascal

Os números de Pascal respondem a alguns problemas de contagem. Pensemos em 7 pessoas numa sala. Chamemos-lhes **A**lison, **C**atherine, **E**mma, **G**ary, **J**ohn, **M**atthew e **T**homas. Quantas maneiras existem de escolher diferentes grupos de 3? Uma hipótese será **A**, **C**, **E**; outra será **A**, **C**, **T**. Os matemáticos consideram útil escrever C (n,r) para representar o número na n-ésima linha, na n-ésima posição (contando a partir de $r = 0$) do triângulo de Pascal. A resposta à nossa questão é C (7,3). O número na 7.ª linha do triângulo, na 3.ª posição, é 35. Se escolhermos um grupo de 3, seleccionamos automaticamente um grupo «não escolhido» de 4 pessoas. Isto explica-se pelo facto de que C (7,4) = 35 também. Genericamente, C (n, r) = C $(n, n - r)$, o que é consequência da simetria do triângulo de Pascal.

O triângulo de Serpińskii

Zeros e Uns No triângulo de Pascal, vemos que os números interiores formam um padrão, dependendo de serem pares ou ímpares. Se substituirmos os números ímpares por 1 e os números pares por 0, obtemos uma representação com o mesmo padrão do notável fractal conhecido como triângulo de Serpiński (ver página 102).

Somar sinais Podemos escrever o triângulo de Pascal que corresponde às potências $(-1 + x)$, ou seja $(-1 + x)^n$.



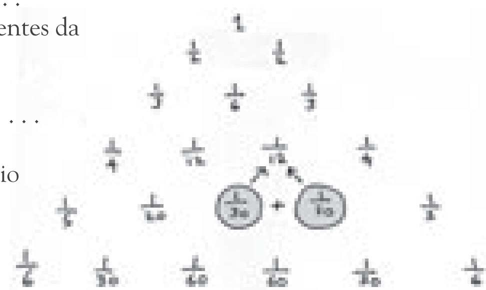
Somar sinais

Neste caso, o triângulo não é completamente simétrico relativamente à linha vertical, e as linhas, em vez de terem uma soma que é uma potência de 2, somam 0. Contudo, são as diagonais que são interessantes aqui. Os números da diagonal sudoeste 1, –1, 1, –1, 1, –1, 1, –1, ... são os coeficientes da expansão

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

enquanto os termos da diagonal seguinte são os coeficientes da expansão

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



O triângulo harmónico de Leibniz

O triângulo harmónico de Leibniz O sábio alemão Gottfried Leibniz descobriu um notável conjunto de números na forma de um triângulo. Os números de Leibniz são simétricos relativamente a uma linha vertical. Mas, ao contrário do triângulo de Pascal, o número numa linha é obtido somando os dois números *abaixo* dele. Por exemplo, 1/30 + 1/20 = 1/12. Para construir este triângulo, podemos começar em cima e avançar da esquerda para a direita por subtracção: Conhecemos 1/12 e 1/30 e então 1/12 – 1/30 = 1/20, será o número a seguir a 1/30. Provavelmente já notaram que a diagonal exterior é a famosa série harmónica

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

mas a segunda diagonal é aquela que é conhecida pela série de Leibniz

$$\frac{1}{1\times2} + \frac{1}{2\times3} + \cdots + \frac{1}{n \times (n+1)}$$

que, depois de alguma manipulação habilidosa, acaba por ser igual a $n/(n + 1)$. Tal como fizemos antes, podemos escrever estes números de Leibniz como $B\,(n,r)$ para serem o n-ésimo número na r-ésima linha. Estão relacionados com os números de Pascal $C\,(n,r)$ pela fórmula

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

É um pouco como diz uma canção em inglês antiga: «o osso do joelho está ligado ao osso da coxa, e o osso da coxa está ligado ao osso da anca».
O triângulo de Pascal tem relações íntimas com muitas áreas matemáticas, como a geometria moderna, a combinatória e a álgebra, para nomear apenas três. Mas, mais do que isso, é um excelente exemplo do trabalho matemático – a busca constante de padrões e harmonia que reforça a nossa compreensão do próprio objecto.

a ideia resumida
A fonte dos números

14 Álgebra

A álgebra oferece-nos uma forma distinta de resolvermos problemas, um método dedutivo peculiar, que é o «pensamento às arrecuas». Consideremos o problema de pegar no número 25, somar-lhe 17, e obter 42. Isto é pensar adiante. São-nos dados os números e somamo-los simplesmente. Mas, em vez disso, suponhamos que nos é dada a resposta 42, e feita outra pergunta. Agora queremos saber o número que, quando somado com 25, dá 42. É aqui que entra o pensamento às arrecuas. Queremos o valor de x, que resolve a equação $25 + x = 42$ e subtraímos 25 a 42 para o obter.

Há séculos que as crianças resolvem problemas cuja solução requer o recurso à álgebra:

A minha sobrinha tem 6 anos e eu tenho 40.
Quando é que terei o triplo da idade dela?

Podemos achar o resultado pelo método da tentativa e erro, mas a álgebra é mais económica. Daqui a x anos, a sobrinha terá $6 + x$ anos e eu terei $40 + x$. Terei o triplo da idade dela quando

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

Efectuando o produto no lado esquerdo da equação, obtemos $18 + 3x = 40 + x$ e passando todos os x para um lado da equação e os números para o outro, obtemos $2x = 22$, o que significa que $x = 11$. Quando eu tiver 51 anos, a minha sobrinha terá 17 anos. Magia!

E se quiséssemos saber quando é que eu terei o dobro da idade dela? Podemos utilizar a mesma abordagem, resolvendo agora

Cronologia

1950 a.C.
Os babilónios trabalham com equações quadráticas

250
Diofanto de Alexandria publica *Arithmetica*

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

para obtermos $x = 28$. Ela terá 34 quando eu tiver 68 anos. Todas as equações acima são do tipo mais simples – as chamadas equações «lineares». Não têm termos como x^2 ou x^3, o que torna as equações mais difíceis de resolver. Equações com termos como x^2 chamam-se «quadráticas» e aquelas com termos como x^3, equações «cúbicas». Antigamente, x^2 era representado como um quadrado e, por um quadrado ter quatro lados, era usado o termo quadrático; x^3 era representado por um cubo.

A matemática atravessou uma grande mudança quando passou de ciência da aritmética para ciência dos símbolos ou álgebra. Avançar de números para letras é um salto mental, mas o esforço vale a pena.

Origens A álgebra era um elemento relevante no trabalho dos académicos islâmicos no século IX. Al-Khwarizmi escreveu um livro de matemática que continha a palavra árabe *al-jabr*. Tratando de problemas práticos em termos de equações lineares e quadráticas, a «ciência das equações» de Al-Khwarizmi deu-nos a palavra «álgebra». Mais tarde, Omar Khayyam ficou famoso por ter escrito as *Rubaiyat* e as linhas imortais

> *Um Jarro de Vinho, um Pedaço de Pão – e Tu a meu lado cantando na Imensidão*

mas em 1070, com 22 anos, escreveu um livro de álgebra em que estudou a resolução das equações cúbicas.

O grande trabalho matemático de Girolamo Cardano, publicado em 1545, marcou uma mudança importante na teoria das equações, porque continha uma grande riqueza de resultados sobre as equações cúbicas e as equações quárticas – aquelas que envolvem um termo do tipo x^4. Esta intensa actividade de pesquisa mostrou que as

A ligação italiana

A teoria das equações cúbicas foi completamente desenvolvida durante a Renascença. Infelizmente, resultou de um episódio em que os matemáticos nem sempre mostraram o melhor dos comportamentos. Scipione Del Ferro encontrou uma solução para vários tipos específicos de equações cúbicas e, sabendo disso, Niccolò Fontana – apelidado *Tartaglia* ou *o Gago* –, professor de Veneza, publicou os seus próprios resultados em álgebra, mas fez segredo dos seus métodos. Girolamo Cardano de Milão convenceu *Tartaglia* a contar-lhe os seus métodos mas jurou segredo. O método foi divulgado, e os dois entraram em contenda quando *Tartaglia* descobriu que o seu trabalho tinha sido publicado no livro de Cardano *Ars Magna,* em 1545.

825
Al-Khwarizmi oferece a palavra «álgebra», derivada de *al-jabr*, à matemática

1591
François Viète escreve um texto matemático utilizando letras para conhecidos e desconhecidos

anos 1920
Emmy Noether publica artigos sobre o desenvolvimento da álgebra moderna

1930
Bartel van der Waerden publica a sua famosa *Moderne Algebra*

equações quadráticas, cúbicas e quárticas podiam ser todas resolvidas por fórmulas que envolviam apenas as operações $+$, $-$, $\times$, $\div$, $\sqrt[q]{\ }$ (a última operação significa raiz q-ésima). Por exemplo, a equação quadrática $ax^2 + bx + c = 0$ pode ser resolvida usando a seguinte fórmula:

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

Se quisermos resolver a equação $x^2 - 3x + 2$, tudo o que há a fazer é colocar os valores $a = 1$, $b = -3$ e $c = 2$ na fórmula.

As fórmulas para resolver as equações cúbicas e quárticas são longas e de difícil resolução, mas é certo que existem. O que intrigava os matemáticos era não conseguirem produzir uma fórmula que pudesse ser genericamente aplicada a equações envolvendo x^5, as equações do quinto grau. O que haveria de tão especial na potência de expoente 5?

Em 1826, Niels Abel, que teve uma curta vida, apresentou uma resposta notável para a difícil questão das equações do quinto grau. Na realidade, provou um conceito contrário, quase sempre mais difícil do que provar que qualquer coisa pode ser feita. Abel provou que não pode existir uma fórmula para resolver todas as equações do quinto grau, e concluiu que qualquer pesquisa futura deste Santo Graal seria fútil. Abel convenceu os matemáticos dos níveis mais altos, mas as notícias demoram muito tempo a espalhar-se no vasto mundo matemático. Alguns matemáticos recusaram aceitar o resultado e, no século XIX, ainda havia quem publicasse trabalhos proclamando ter encontrado a fórmula inexistente.

O mundo moderno Durante 500 anos, álgebra significou «teoria das equações», mas o seu desenvolvimento deu um salto no século XIX. Percebeu-se que os símbolos da álgebra podiam representar mais do que apenas números – podiam representar «proposições» e, assim, a álgebra podia ser relacionada com o estudo da lógica. Podiam mesmo representar objectos de muitas dimensões, como os que se encontram nas matrizes algébricas (ver página 156). E, como muitos não-matemáticos há muito suspeitavam, podiam mesmo representar coisa nenhuma e ser apenas símbolos movendo-se de acordo com certas regras (formais).

Um acontecimento importante na álgebra moderna ocorreu em 1843, quando o irlandês William Rowan Hamilton descobriu os quaterniões. Hamilton procurava um sistema de símbolos que expandissem os complexos bidimensionais para dimensões mais altas. Durante muitos anos, tentou símbolos tridimensionais, mas sem obter um sistema satisfatório. Todas as manhãs quando descia para o pequeno-almoço, os filhos perguntavam-lhe «Então, pai, já consegues *multiplicar* triplos?» e ele era obrigado a responder que só conseguia somá-los e subtraí-los.

O sucesso surgiu de uma forma inesperada. A pesquisa tridimensional era um beco sem saída – Hamilton devia concentrar-se nos símbolos de quatro dimensões. Este momento de inspiração surgiu-lhe enquanto caminhava com a mulher ao longo do canal de Dublin. Ficou estático com a sensação de descoberta. Sem hesitação, o vândalo de 38 anos, real astrónomo da Irlanda e cavaleiro do Reino, gravou as relações da definição na pedra da ponte de Brougham, num local hoje assinalado por uma placa. Com a descoberta no espírito, o assunto tornou-se a obsessão de Hamilton, que fez conferências sobre ele ano após ano e publicou dois livros de peso sobre o seu «sonho místico de quatro flutuando para oeste».

Uma peculiaridade dos quarterniões é que, quando são multiplicados uns pelos outros, a ordem por que o fazem é vital, contrariamente às regras da vulgar aritmética. Em 1844, o linguista e matemático germânico Herman Grassmann publicou outro sistema algébrico com muito menos drama. Ignorado na altura, acabou por ser muito mais abrangente. Hoje, tanto os quarterniões como a álgebra de Grassmann têm aplicações na geometria, na física e na computação gráfica.

O abstracto No século XX, o paradigma dominante da álgebra era o método axiomático. Tinha sido usado como base para a geometria de Euclides, mas não foi aplicado na álgebra senão recentemente.

Emmy Noether foi o campeão do modelo abstracto. Nesta álgebra moderna, a ideia central é o estudo da estrutura em que os exemplos individuais são subservientes à noção abstracta geral. Se os exemplos individuais têm a mesma estrutura mas talvez notação diferente, são chamados isomórficos.

A estrutura algébrica mais fundamental é o grupo, que é definido por uma lista de axiomas (ver página 155). Há estruturas com menos axiomas (como grupóides, semigrupos e quase-grupos) e estruturas com mais axiomas (como anéis, anéis de divisão, corpos). Todas estas palavras novas foram importadas para a matemática no início do século XX, enquanto a álgebra se transformava numa ciência abstracta conhecida como «álgebra moderna».

a ideia resumida
Resolver o desconhecido

15 O algoritmo de Euclides

Al-Khwarizmi deu-nos a palavra «álgebra», mas foi o seu livro do século IX sobre aritmética que nos deu a palavra «algoritmo», um conceito útil tanto para matemáticos como para cientistas da computação. Mas o que é? Se conseguirmos responder, estamos perto de compreender o algoritmo da divisão de Euclides.

Em primeiro lugar, um algoritmo é uma rotina. É uma lista de instruções, tais como «faça isto e depois faça aquilo». Percebe-se, portanto, porque é que os computadores gostam de algoritmos: são muito bons a seguir instruções e nunca se desviam do caminho. Alguns matemáticos pensam que os algoritmos são maçadores, por serem repetitivos, mas escrever um algoritmo e depois traduzi-lo para centenas de linhas de código de computador contendo instruções matemáticas não é tarefa fácil. Há um risco considerável de tudo correr muito mal. Escrever um algoritmo é um desafio criativo. Muitas vezes, há vários métodos para executar a mesma função e temos de escolher o melhor. Alguns algoritmos podem não ser «adequados ao objectivo» e alguns podem ser francamente ineficientes porque serpenteiam. Alguns podem ser rápidos, mas produzir a resposta errada. É um pouco como cozinhar. Devem existir centenas de receitas (algoritmos) para fazer peru recheado. De certeza que não queremos arriscar, fazendo a refeição seguindo um algoritmo fraco. Assim, temos os ingredientes e as instruções. O início da (abreviada) receita pode ser qualquer coisa como:

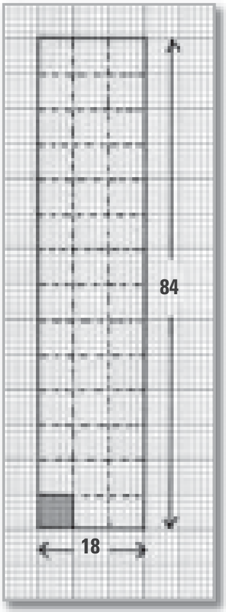
- Encher a cavidade do peru com o recheio
- Esfregar a pele do peru com manteiga
- Temperar com sal, pimenta e pimentão
- Assar a 335 graus durante 3½ horas
- Deixar o peru cozinhado repousar durante ½ hora

Cronologia

cerca de 300 a.C.
Publicação do algoritmo de Euclides no Livro *7 dos Elementos*

cerca de 300
Sun Tzu descobre o teorema chinês dos restos

Tudo o que temos de fazer é executar o algoritmo em passos sequenciais. A única coisa que falta nesta receita, normalmente presente nos algoritmos matemáticos, é um *loop*, uma ferramenta para lidar com a recursão. Felizmente, só temos de cozinhar o peru uma vez.



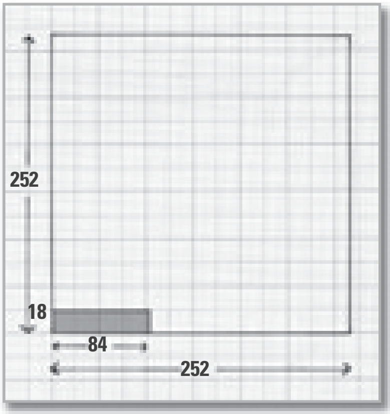
Em matemática, também temos ingredientes – os números. O algoritmo da divisão de Euclides é concebido para calcular o máximo divisor comum (*mdc*). O *mdc* de dois números inteiros é o maior número que é divisor de ambos. Como exemplos de ingredientes, escolhamos os números 18 e 84.

O máximo divisor comum O *mdc* no nosso exemplo é o maior número que divide exactamente tanto o 18 como o 84. O número 2 divide tanto o 18 e como o 84, mas o número 3 também. Então, 6 também dividirá os dois números. Será este o maior número que os divide? Podemos tentar o 9 ou o 18. Constatamos que estes candidatos não dividem 84, logo 6 é o maior número que os divide. Podemos concluir que 6 é o *mdc* de 18 e 84, representando *mdc*(18, 24) = 6.

O *mdc* pode ser interpretado em termos de azulejos de cozinha. É o lado do maior azulejo quadrado que cobrirá uma parede rectangular com uma largura de 18 e um comprimento de 84, e não é permitido cortar azulejos. Neste caso, constatamos que o melhor azulejo é o 6 × 6.

O máximo divisor comum também é conhecido por «maior factor comum» ou «maior divisor comum». Há um conceito relacionado, o mínimo múltiplo comum (*mmc*). O *mmc* de 18 e 84 é o menor número divisível tanto por 18 como por 84. A ligação entre o *mdc* e o *mmc* é realçada pelo facto de o *mmc* de dois números multiplicado pelo seu *mdc* ser igual ao produto dos próprios números. Aqui, *mdc*(18, 24) = 252 e podemos verificar que 6 × 252 = 1512 = 18 × 84.

Geometricamente, o *mmc* é o comprimento do lado do menor quadrado que pode ser coberto com azulejos rectangulares 18 × 84. Porque *mmc* (*a*, *b*) = *ab* ÷ *mdc*(*a*,*b*), vamos concentrar-nos em determinar o *mdc*. Já calculámos o *mdc*(18, 84) = 6, mas para o fazermos precisamos de conhecer os divisores tanto de 18 como de 24.



Pavimentação do quadrado com azulejos rectangulares 18 x 84

810
Al-Khwarizmi fornece a palavra algoritmo à matemática

1202
Fibonacci publica um trabalho sobre congruências no *Liber Abaci*

anos 1970
Utilização do teorema chinês dos restos na encriptação de mensagens

Recapitulando, primeiro decompomos ambos os números nos seus factores: $18 = 2 \times 3 \times 3$ e $84 = 2 \times 2 \times 3 \times 7$. Depois, comparando-os, o número 2 é comum a ambos e é a potência de maior expoente que dividirá ambos. Da mesma forma, 3 é comum e é a potência de maior expoente que divide ambos, mas embora 7 divida 84, não divide 18, logo não pode entrar no *mdc* como factor. Concluímos que $2 \times 3 = 6$ é o maior número que divide ambos. Poderá este malabarismo ser evitado? Pensem nos cálculos se quiséssemos determinar *mdc*(17640, 54054). Primeiro, teríamos de factorizar estes números, e isso seria apenas o início. Tem de haver uma forma mais simples.

O algoritmo Existe uma forma melhor. O algoritmo de Euclides dado nos *Elementos*, Livro 7, Proposição 2: «Dados dois números não primos entre si, para encontrar a sua maior medida comum.»

O algoritmo de Euclides é maravilhosamente eficiente e de facto substitui o esforço de determinar factores pela simples subtracção. Vejamos como funciona.

O objectivo é calcular $d = mdc(18, 84)$. Começamos por dividir 84 por 18. Não é uma divisão exacta, mas cabe 4 vezes com 12 (o resto) a sobrar:

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

Dado que d tem de dividir 84 e 18, tem de dividir o resto 12. Logo, $d = mdc(12, 18)$. Assim, podemos repetir o processo e dividir 18 por 12:

$$18 = \mathbf{1} \times 12 + 6$$

e obtemos resto 6, logo $d = mdc(6, 12)$. Dividindo 12 por 6, obtemos o resto 0, logo $d = mdc(0, 6)$. O 6 é o maior número que divide 0 e 6, logo é a resposta.

Se calcularmos $d = mdc(17640, 54054)$, os sucessivos restos serão 1134, 630, 504 e 0, dando-nos $d = 126$.

Utilizações do *mdc* O *mdc* pode ser usado na resolução de equações quando as soluções tenham de ser números inteiros. São as chamadas equações diofantina, que receberam o nome do matemático grego Diofanto de Alexandria.

Suponhamos que a nossa tia-avó está de partida para as suas férias anuais em Barbados. Manda o seu mordomo John ao aeroporto com o conjunto das malas, que pesam ou 18 ou 84 quilogramas, e é informada de que o peso total da bagagem é 652 quilogramas. Quando John regressa a casa, o seu filho de 9 anos

James diz: «Não pode estar certo, porque o *mdc* 6 não divide 652.» James sugere que de facto o peso total deverá ser 642 quilogramas.

James sabe que há uma solução nos números inteiros para a equação $18x + 84y = c$ se, e só se, o *mdc* 6 dividir o número c. Não divide para $c = 652$, mas divide para 642. James nem necessita de saber quantas malas x, y de cada peso a nossa tia-avó tenciona levar para Barbados.

O teorema chinês dos restos Quando o *mdc* de dois números é 1, dizemos que são «relativamente primos». Eles próprios não têm de ser primos, mas apenas primos entre si, por exemplo *mdc*(6, 35) = 1, embora nem 6 nem 35 sejam primos. Precisaremos disto para o teorema chinês dos restos.

Observemos outro problema: Angus não sabe quantas garrafas de vinho tem, mas, quando as põe aos pares, resta uma. Quando as arruma em filas de cinco na sua prateleira de vinhos, sobram 3. Quantas garrafas tem? Sabemos que, na divisão por 2, temos resto 1 e, na divisão por 5, temos resto 3. A primeira condição permite-nos excluir todos os números pares. Pesquisando ao longo dos ímpares, rapidamente descobrimos que o 13 satisfaz as condições (podemos presumir que Angus tem mais de 3 garrafas, um número que também satisfaz as condições). Mas há outros números que estão correctos – de facto, toda uma sequência que começa em 13, 23, 33, 43, 53, 63, 73, 83, ...

Vamos juntar uma nova condição: que o número deve dar resto 3 na divisão por 7 (as garrafas chegaram em caixas de 7 com 3 extras). Pesquisando ao longo da sequência 13, 23, 33, 43, 53, 63, ..., para considerar este facto, verificamos que 73 é uma solução, mas note-se que 143 também, tal como 213 ou qualquer outro número determinado pela soma de múltiplos de 70 a estes.

Em termos matemáticos, encontramos soluções garantidas pelo teorema chinês dos restos, que também diz que quaisquer duas soluções diferem por um múltiplo de $2 \times 5 \times 7 = 70$. Se Angus tiver entre 150 e 250 garrafas, o teorema fixa a solução em 213 garrafas. Nada mal para um teorema do século III.

a ideia resumida
Um caminho para o maior

16 Lógica

«Se existissem menos carros nas estradas, a poluição seria aceitável. Ou temos menos carros na estrada ou temos de impor taxas rodoviárias, ou ambos. Se houver taxas rodoviárias, o Verão será insuportavelmente quente. Na realidade, o Verão está muito fresco. A conclusão é inevitável: a poluição é aceitável.»
Este argumento do editorial de um jornal diário é «válido» ou é ilógico? Não nos interessa se faz sentido como política para o tráfego rodoviário ou se é bom jornalismo. Só estamos interessados na sua validade como argumento racional. A lógica pode ajudar-nos a decidir a questão, porque ela diz respeito a uma verificação rigorosa do raciocínio.

Duas premissas e uma conclusão Tal como está, a passagem do jornal é bastante complicada. Vejamos primeiro alguns argumentos mais simples, voltando ao filósofo grego Aristóteles de Estagira, que é considerado o fundador da ciência da lógica. A sua abordagem baseou-se em diferentes formas de silogismo, um estilo de argumento baseado em três afirmações: duas premissas e uma conclusão. Um exemplo:

Todos os *spaniels* são cães
Todos os cães são animais

Todos os *spaniels* são animais

Acima da linha temos as premissas, e abaixo dela, a conclusão. Neste exemplo, a conclusão tem uma inevitabilidade certa, seja qual for o significado que se dê às palavras «*spaniels*», «cães» e «animais». O mesmo silogismo, mas usando palavras diferentes, será

Cronologia

cerca de 335 a.C.
Aristóteles formaliza a lógica do silogismo

1847
Boole publica *The Mathematical Analysis of Logic*

Todas as maçãs são laranjas
Todas as laranjas são bananas

Todas as maçãs são bananas

Neste caso, nenhuma das afirmações individuais faz sentido se usarmos as conotações habituais das palavras. No entanto, os exemplos de silogismos têm a mesma estrutura e é a estrutura que torna o silogismo válido. Não é simplesmente possível encontrar um exemplo de As, Bs e Cs com esta estrutura em que as premissas sejam verdadeiras e a conclusão falsa. É isto que torna útil um argumento válido.

Todos os As são Bs
Todos os Bs são Cs

Todos os As são Cs

Um argumento válido

É possível uma variedade de silogismos, se variarmos os quantificadores, como «Todos», «Alguns» e «Nenhum» (como em «Nenhum A é B). Por exemplo:

Alguns As são Bs
Alguns Bs são Cs

Alguns As são Cs

Será este argumento válido? Aplicar-se-á a *todos* os casos de As, Bs e Cs, ou existe algum contra-exemplo escondido, um exemplo para o qual as premissas são verdadeiras mas a conclusão é falsa? E se substituirmos A por *spaniels*, B por objectos castanhos, e C por mesas? Será o seguinte convincente?

Alguns *spaniels* são castanhos
Alguns objectos castanhos são mesas

Alguns *spaniels* são mesas

O nosso contra-exemplo mostra que este silogismo *não* é válido. Existem tantos tipos diferentes de silogismos, que os académicos medievais inventaram uma mnemónica para os recordarem. O nosso primeiro exemplo era conhecido como BARBARA, porque contém três utilizações de «All» (Todos). Estes métodos de analisar argumentos duraram mais de 2000 anos e desempenharam um papel importante nos estudos nas universidades medievais. A lógica de Aristóteles – a sua teoria dos silogismos – foi considerada uma ciência perfeita até boa metade do século XIX.

1910
Russel e Whitehead tentam reduzir a matemática à lógica

1965
Lofti Zadec desenvolve a lógica difusa

1987
O sistema de metropolitano japonês é baseado na lógica difusa

a	b	a ∨ b
T	T	T
T	F	T
F	T	T
F	F	F

Tabela de verdade «ou»

a	b	a ∧ b
T	T	T
T	F	F
F	T	F
F	F	F

Tabela de verdade «e»

a	¬ a
T	F
F	T

Tabela de verdade «não»

a	b	a → b
T	T	T
T	F	F
F	T	T
F	F	T

Tabela de verdade «implicação»

Lógica proposicional Outro tipo de lógica vai além dos silogismos. Trabalha com proposições ou afirmações simples e a combinação entre elas. Para analisar o editorial do jornal, precisaremos de conhecer minimamente esta «lógica proposicional». É geralmente conhecida por «álgebra da lógica», o que nos dá uma pista sobre a sua estrutura, desde que George Boole percebeu que podia ser tratada como uma forma de álgebra. Na década de 1840 havia muito trabalho desenvolvido em lógica por matemáticos como Boole e Augustus De Morgan.

Consideremos a proposição **a**, em que **a** significa «O Freddy é um *spaniel*». A proposição **a** pode ser Verdadeira ou Falsa. Se estiver a pensar no meu cão chamado *Freddy*, que é de facto um *spaniel*, a afirmação é verdadeira (**V**), mas se estiver a pensar no meu primo cujo nome também é Freddy a afirmação é falsa (**F**). A verdade ou falsidade de uma proposição depende da sua referência.

Se tivermos uma proposição **b**, tal como «A Ethel é uma gata», podemos combinar as duas proposições de várias formas. Uma das combinações escreve-se **a ∨ b**. O conector **∨** corresponde a «ou», mas o seu uso na lógica é ligeiramente diferente do «ou» da linguagem corrente. Em lógica, **a ∨ b** é verdade se «O Freddy é um *spaniel*» for verdade, ou se «A Ethel é uma gata» for verdade, ou ambas forem verdade, e é falsa apenas se tanto **a** como **b** forem falsas. Esta conjunção de proposições pode ser resumida numa tabela de verdade.

Também podemos combinar as proposições usando «e», que se escreve **a ∧ b**, e «não», que se escreve **¬a**. A álgebra da lógica torna-se clara quando se combinam estas proposições usando uma mistura de conectores com **a**, **b** e **c**, como **a ∧(b ∨ c)**. Podemos obter uma equação a que chamamos identidade:

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

O símbolo ≡ significa equivalência entre afirmações lógicas em que ambos os lados da equivalência têm a mesma tabela de verdade. Isto é um paralelo entre a álgebra da lógica e a álgebra ordinária, porque os símbolos ∧ e ∨ agem como o × e o + na álgebra ordinária, em que se tem $x \times (y + z) = (x \times y) + (x \times z)$. Contudo, o paralelo não é absoluto e há excepções.

Outros conectores lógicos podem ser definidos em termos destes conectores básicos. Um bastante útil é o conector da implicação **a→b**, que é definido como equivalente a **¬a ∨ b** e tem a tabela de verdade que se mostra.

Se voltarmos agora a ler o editorial do jornal, podemos escrevê-lo em forma simbólica para produzir o argumento.
C = menos **C**arros nas estradas
P = a **P**oluição será aceitável
E = há um **E**squema de taxas rodoviárias
Q = o Verão será insuportavelmente **Q**uente

$$\frac{\begin{array}{c} C \rightarrow P \\ C \vee E \\ E \rightarrow Q \\ \neg Q \end{array}}{P}$$

O argumento é válido? Vamos supor que a conclusão **P** é falsa, mas todas as premissas são verdadeiras. Se pudermos mostrar que isto leva a uma contradição, o argumento será válido. Será então impossível ter premissas verdadeiras e uma conclusão falsa. Se **P** é falso, então, da primeira premissa **C → P**, **C** tem de ser falso. Como **C v E** é verdadeiro, o facto de **C** ser falso significa que **E** é verdadeiro. Da terceira premissa **E → Q** decorre que **Q** seja verdadeiro. Ou seja, **¬Q** é falso. Isto contradiz o facto de **¬Q,** a última premissa, ter sido assumida como verdadeira. O conteúdo das afirmações do editorial pode ser objecto de discussão, mas a estrutura do argumento é válida.

∨	**ou**
∧	**e**
¬	**não**
→	**implica**
∀	**para todos**
∃	**existe**

Outras lógicas Gottlob Frege, C. S. Peirce, e Ernst Schröder introduziram a quantificação na lógica proposicional e construíram uma «lógica de primeira ordem de predicados» (por ser predito em variáveis). Esta usa o quantificador universal, ∀, para significar «para todos», e o quantificador existencial, ∃, para «existe».

Outro novo desenvolvimento em lógica é a ideia de lógica difusa, que sugere raciocínios confusos, mas na realidade trata o alargamento das tradicionais fronteiras da lógica. A lógica tradicional baseia-se em colecções ou conjuntos. Assim, temos o conjunto dos *spaniels*, o dos cães, e o dos objectos castanhos. Temos a certeza do que está incluído ou não em cada um deles. Se virmos um dálmata no parque, temos a certeza absoluta de que não se inclui no conjunto dos *spaniels*.

A teoria dos conjuntos difusos trabalha com o que parecem ser conjuntos definidos de forma imprecisa. Imaginemos que tínhamos um conjunto de *spaniels* pesados. Que peso deveria ter um *spaniel* para ser incluído no conjunto? Com conjuntos difusos, existe uma *gradação* de participação e uma fronteira quanto ao que está dentro e ao que está fora. A matemática permite-nos ser precisos em relação ao difuso. A lógica está longe de ser um assunto acabado. Tem progredido desde Aristóteles e agora é um domínio produtivo nas pesquisas e aplicações modernas.

a ideia resumida
A linha evidente da razão

17 Prova

Os matemáticos tentam justificar as suas afirmações com provas. A tentativa de tornarem infalíveis os seus argumentos racionais é a força motriz da matemática pura. As cadeias de deduções correctas a partir do que é conhecido ou presumido conduzem os matemáticos a uma conclusão que entra para o inventário matemático do provado.

As provas não se obtêm de forma fácil – muitas vezes chegam depois de muita investigação sobre pistas falsas. A luta para as conseguir ocupa o centro da vida de um matemático. Uma prova contém o selo de autenticidade do matemático, separando teoremas demonstrados de conjecturas, ideias brilhantes e primeiras suposições.

As qualidades procuradas numa prova são o rigor, a transparência e, não menos importante, a elegância. A isto soma-se conhecimento. Uma boa prova é «uma prova que nos torna mais conhecedores» – mas é melhor haver uma prova do que não haver nenhuma. O progresso com base em factos não provados transporta o perigo de essas teorias poderem ser construídas no equivalente matemático da areia.

O que não quer dizer que uma prova dure para sempre, porque pode ter de ser revista à luz dos desenvolvimentos que com ela se relacionam.

O que é uma prova? Quando lemos ou ouvimos falar de um resultado matemático, acreditamos nele? O que é que nos faz acreditar nele? Uma resposta será um argumento lógico sólido que avançou de ideias que aceitamos para a afirmação em que pensamos. É a isto que os matemáticos chamam prova, na sua forma usual de linguagem corrente e lógica estrita. Dependendo da qualidade da prova, acreditam ou continuam cépticos.

Os principais tipos de prova utilizados pelos matemáticos são: o método do contra-exemplo; o método directo; o método indirecto; e o método da indução matemática.

Cronologia

cerca de 300 a.C.
Os *Elementos* de Euclides fornecem o modelo para a prova matemática

1637
Descartes promove o rigor matemático como modelo no seu *Discurso do Método*

O contra-exemplo Comecemos por ser cépticos – é um método de provar que uma afirmação é incorrecta. Pensemos numa afirmação específica como contra-exemplo. Suponhamos que ouvimos a alegação de que todos os números multiplicados por si próprios resultam num número par. Acreditamos? Antes de nos precipitarmos na resposta, devemos tentar alguns exemplos. Se tivermos um número como o 6 e o multiplicarmos por si próprio para obter $6 \times 6 = 36$, verificamos que 36 é realmente um número par. Mas uma andorinha não faz a Primavera. A alegação respeitava a *todos* os números, e há infinitos. Tentando o 9, verificamos que $9 \times 9 = 81$. E 81 é um número ímpar. Isto significa que a afirmação de que *todos* os números quando multiplicados por si próprios dão um número par é falsa. A um exemplo que é contrário à afirmação original chama-se contra-exemplo. Um contra-exemplo da alegação de que «todos os cisnes são brancos» será ver um cisne negro. Parte do prazer da matemática é procurar um contra-exemplo para derrubar um pretenso teorema.

Se não encontrarmos um contra-exemplo, podemos sentir que o teorema está correcto. Aqui, os matemáticos têm de jogar um jogo diferente. Tem de ser construída uma prova e o tipo mais simples é o método directo de prova.

O método directo No método directo, prosseguimos com os argumentos lógicos a partir do que já está provado, ou tem sido presumido, para a conclusão. Se conseguirmos fazê-lo, temos um teorema. Não podemos provar que todos os números multiplicados por si próprios resultam num número par, porque já demonstrámos que isso não é verdade. Mas podemos ser capazes de salvar alguma coisa. A diferença entre o nosso primeiro exemplo, 6, e o contra-exemplo, 9, é que o primeiro número é par e o do contra-exemplo é ímpar. Alterar a hipótese é uma coisa que podemos fazer. A nossa afirmação é: se multiplicarmos qualquer número *par* por si próprio, o resultado é um número par.

Em primeiro lugar, ensaiamos outros exemplos numéricos e verificamos que nem sempre conseguimos encontrar um contra-exemplo. Mudando de rumo, tentamos prová-lo, mas como começar? Podemos iniciar-nos com um número genérico par n, mas, como isto parece um pouco abstracto, veremos como a prova pode continuar examinando um número concreto, digamos o 6. Ora, um número par é aquele que é múltiplo de 2, isto é $6 = 2 \times 3$. Como $6 \times 6 = 6 + 6 + 6 + 6 + 6 + 6$ ou, escrito de outra forma, $6 \times 6 = 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3 + 2 \times 3$ ou, rescrevendo usando parênteses,

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

Isto quer dizer que 6×6 é múltiplo de 2 e, como tal, um número par. Mas neste argumento não há nada que seja particular ao 6, e poderíamos ter começado com $n = 2 \times k$ para obter

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

1838
De Morgan introduz a expressão «indução matemática»

1967
Bishop prova resultados usando exclusivamente métodos construtivistas

1976
Imre Lakatos publica o influente *Provas e Refutações*

e concluir que $n \times n$ é par. A nossa prova está agora completa. Traduzindo os *Elementos* de Euclides, os matemáticos mais recentes escreviam «QED» no fim de uma prova, para dizer «trabalho concluído» – QED é uma abreviatura do latim *quod erat demonstrandum* (que era para ser demonstrado). Hoje utiliza-se um quadrado preenchido, ■, chamado halmos, devido ao seu introdutor, Paul Halmos.

O método indirecto Neste método, supõe-se que a conclusão é falsa e por argumentação lógica demonstra-se que isso contradiz a hipótese. Provemos o resultado anterior com este método.

A nossa hipótese é de que n é par e supomos que $n \times n$ é ímpar. Podemos escrever $n \times n = n + n + \ldots + n$ e existem n destes. Isto significa que n não pode ser par (porque, se fosse, $n \times n$ seria par). Logo n é ímpar, o que contradiz a hipótese. ■

Na realidade, esta é uma forma suave de método indirecto. O método indirecto extremo é conhecido por *reductio ad absurdum* (redução ao absurdo) e era muito apreciado pelos gregos. Na Academia de Atenas, Sócrates e Platão adoravam provar qualquer questão em debate embrulhando os adversários numa rede de contradições e deixando fora dela o ponto que estavam a tentar provar. A prova clássica de que a raiz quadrada de 2 é um número irracional é uma das que têm esta forma, em que se começa por supor que a raiz quadrada de 2 é um número racional e resulta uma contradição dessa suposição.

O método de indução matemática A indução matemática é uma forma poderosa de demonstrar que uma sequência de afirmações $P_1, P_2, P_3, \ldots$ é toda verdadeira. Isto foi reconhecido por Augustus De Morgan na década de 1830, quando formalizou aquilo que já era conhecido há centenas de anos. Esta técnica específica (não confundir com indução científica) é largamente usada para provar afirmações que envolvem números *inteiros*. É especialmente útil na teoria dos grafos, na teoria dos números, e genericamente em ciências da computação. Como exemplo prático, pensemos no problema de somar números ímpares. Por exemplo, a adição dos primeiros três números ímpares $1 + 3 + 5$ é 9, enquanto a soma dos primeiros quatro $1 + 3 + 5 + 7$ é 16. Mas, se 9 é $3 \times 3 = 3^2$ e 16 é $4 \times 4 = 42$, será que a adição dos primeiros n números ímpares será igual a n^2? Se tentarmos escolher aleatoriamente um valor para n, digamos $n = 7$, verificamos que realmente a soma dos primeiros sete é $1 + 3 + 5 + 7 + 9 + 11 + 13 = 49$, que é 7^2. Mas será que *todos* os valores de n seguem este padrão? Como é que podemos ter a certeza? Temos um problema, porque não podemos verificar individualmente um número infinito de casos.

É aqui que entra a indução matemática. Informalmente, é um método de prova em dominó. Esta metáfora aplica-se a uma linha de dominós colocados ao alto. Se um dos dominós cair, fará cair o seguinte. Isto é óbvio. Para os fazer cair *a todos*, só precisamos de que o primeiro caia. Podemos aplicar este raciocínio ao problema dos números primos. A afirmação P_n diz que a soma dos n primeiros ímpares é n^2. A indução matemática provoca uma reacção em cadeia em que P_1, P_2, P_3, . . . serão *todos* verdadeiros. A afirmação P_1 é trivialmente verdadeira, dado que $1 = 1^2$. A seguir, P_2 é verdadeira porque $1 + 3 = 1^2 + 3 = 2^2$, P_3 é verdadeira porque $1 + 3 + 5 = 2^2 + 5 = 3^2$ e P_4 é verdadeira porque $1 + 3 + 5 + 7 = 3^2 + 7 = 4^2$. Usamos o resultado de um passo para saltar para o seguinte. Este processo pode ser formalizado para conceber o método de indução matemática.

Dificuldades com a prova Há provas de todos os estilos e tamanhos. Algumas são curtas e sucintas, particularmente as que se encontram nos livros de estudo. Outras detalham as últimas pesquisas e ocupam toda uma edição de um jornal e têm milhares de páginas. Poucas pessoas haverá que dominem o argumento completo nesses casos.

Também existem questões fundamentais. Por exemplo, um pequeno número de matemáticos não aprecia o método de *reductio ad absurdum* da prova indirecta quando aplicado à existência. Se a assunção de que a solução de uma equação não existe leva a uma contradição, será ela suficiente para provar que a solução existe? Os opositores deste método de prova reclamarão que a lógica é apenas um truque e não nos diz como realmente se constrói uma solução concreta. São os «Construtivistas» (de vários matizes), que dizem que o método de prova falha em proporcionar «significado numérico». Desprezam os matemáticos clássicos que encaram o método de *reductio ad absurdum* como arma essencial do arsenal matemático. Por outro lado, os matemáticos mais tradicionais dirão que a exclusão deste tipo de argumento nos deixa de mãos atadas e, além disso, a negação de tantos resultados provados por este método indirecto deixa a tapeçaria da matemática bastante puída.

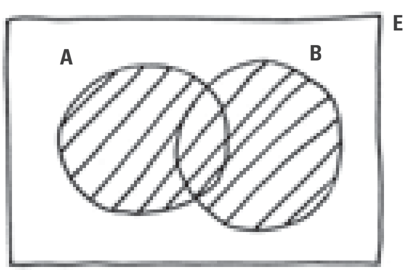
a ideia resumida
A prova provada

18 Conjuntos

Nicholas Bourbaki era o pseudónimo de um grupo auto--seleccionado de académicos franceses que queriam reescrever a matemática de alto a baixo e na «forma correcta». A sua principal reivindicação era de que tudo se deveria basear na teoria dos conjuntos. O método axiomático era central para eles e os livros que produziram foram escritos no rigoroso estilo de «definição, teorema e prova». Este foi também o impulso do movimento da matemática moderna na década de 1960.

Georg Cantor criou a teoria dos conjuntos a partir do seu desejo de assentar a teoria dos números reais em bases sólidas. Apesar do preconceito e da censura iniciais, a teoria dos conjuntos foi perfeitamente consagrada como um ramo da matemática no virar do século XX.

O que são conjuntos? Um conjunto pode ser encarado como uma colecção de objectos. É uma definição informal, mas dá-nos a ideia principal. Aos objectos chamava-se «elementos» ou «membros» do conjunto. Se escrevermos um conjunto A que tem um membro *a*, podemos escrever $a \in A$, como fez Cantor. Um exemplo é $A = \{1, 2, 3, 4, 5\}$ e podemos escrever $1 \in A$ para significar que é membro, e $6 \notin A$ para significar que não é.



A união de *A* e *B*

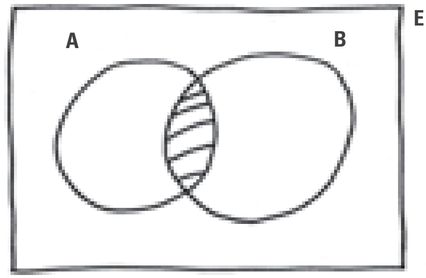
Os conjuntos podem ser combinados de duas formas importantes. Se *A* e *B* são dois conjuntos, ao conjunto que tem os elementos que são membros de *A* ou *B* (ou ambos) chama-se a «união» dos dois conjuntos. Os matemáticos escrevem isto como $A \cup B$, que também se pode representar por um diagrama de Venn, assim denominado em homenagem ao reverendo John Venn. Euler usou diagramas do género ainda mais cedo.

O conjunto $A \cap B$ consiste nos elementos que são membros de *A* e de B e chama-se «intersecção» dos dois conjuntos.

Cronologia

1872
Cantor dá o primeiro passo na criação da teoria dos conjuntos

1881
Venn populariza «os diagramas de Venn» para conjuntos



A intersecção de A e B

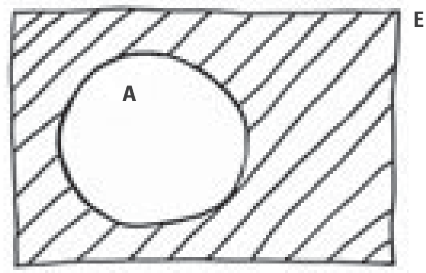
Se $A = \{1, 2, 3, 4, 5\}$ e $B = \{1, 3, 5, 7, 10, 21\}$, a união é $A \cup B = \{1, 2, 3, 4, 5, 7, 10, 21\}$ e a intersecção é $A \cap B = \{1, 3, 5\}$. Se olharmos para o conjunto A como parte de um conjunto universal E, podemos definir o conjunto complementar $\neg A$ como o que consiste naqueles elementos de E que não estão em A.

As operações de $\cap$ e $\cup$ sobre conjuntos são análogas ao $\times$ e ao $+$ na álgebra. Juntamente com a operação de complementar $\neg$, existe uma «álgebra de conjuntos». Augustus De Morgan, um matemático britânico nascido na Índia, formulou leis para mostrar como as três operações funcionavam em conjunto. Na nossa notação moderna, as leis de De Morgan são:

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

e

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



O complemento de A

Os paradoxos Não existe problema com os conjuntos finitos, porque podemos listar os seus elementos, como em $A = \{1, 2, 3, 4, 5\}$, mas, no tempo de Cantor, os conjuntos infinitos eram mais desafiantes.

Cantor definiu conjuntos como colecções de elementos com uma propriedade específica. Pensemos no conjunto $\{11, 12, 13, 14, 15, \ldots\}$ de todos os números inteiros maiores do que 10. Como o conjunto é infinito, não podemos escrever todos os seus elementos, mas ainda assim podemos especificá-lo atendendo à propriedade que todos os seus membros têm em comum. Seguindo as pisadas de Cantor, podemos escrever o conjunto como $A = \{x: x \text{ é um número inteiro} > 10\}$, em que os dois pontos significam «tais que».

Na teoria de conjuntos mais antiga, também poderíamos ter um conjunto de coisas abstractas, $A = \{x: x \text{ é uma coisa abstracta}\}$. Neste caso, A é ele próprio uma coisa abstracta, logo é possível ter $A \in A$. Mas, permitindo esta relação, levantam-se sérios problemas. O filósofo britânico Bertrand Russell teve a ideia

1931
Gödel prova que qualquer sistema matemático axiomático formal contém afirmações indecidíveis

1939
O pseudónimo Bourbaki é usado pela primeira vez em França por matemáticos

1964
Cohen prova a independência da hipótese do *continuum*

de um conjunto S que continha todas as coisas que *não* se continham a si próprias. Simbolicamente, $S = \{x\colon x \notin x\}$.

Fazemos então a pergunta «está $S \in S$?». Se a resposta for «sim», S tem de satisfazer a afirmação que o define, logo $S \notin S$. Por outro lado, se a resposta for «não» e $S \notin S$, *S não* satisfaz a relação que o define $S = \{x\colon x \notin x\}$ e assim $S \in S$. A questão de Russell acabou com esta afirmação, a base do paradoxo de Russel,

$$S \in S \text{ se, e só se, } S \notin S$$

É semelhante ao «paradoxo do barbeiro»: o barbeiro de uma aldeia anuncia aos habitantes que só barbeará aqueles que não se barbearem a si próprios. A questão surge: barbear-se-á o barbeiro a si próprio? Se não se barbear, devia. Se se barbear, não devia.

É indispensável evitar estes paradoxos, educadamente chamados «antinomias». Para os matemáticos, simplesmente não é admissível existirem sistemas que gerem contradições. Russell criou uma teoria de tipos e só permitia que $a \in$ A se a fosse de um tipo inferior a A, evitando assim expressões como $S \in S$.

Outra forma de evitar estas antinomias era formalizar a teoria dos conjuntos. Nesta abordagem, não nos preocupamos com a natureza dos conjuntos em si, mas listamos os axiomas formais que especificam as regras para os tratarmos. Os gregos tentaram algo semelhante com um problema próprio – não tinham de explicar o que eram as linhas rectas, mas apenas como lidar com elas.

No caso da teoria dos conjuntos, foi a origem dos axiomas de Zermelo-Fraenkel para a teoria dos conjuntos que impediu que os conjuntos no seu sistema fossem demasiado «grandes». Impediu eficazmente que aparecessem criaturas tão perigosas como o conjunto de todos os conjuntos.

O teorema de Gödel O matemático austríaco Kurt Gödel deu um soco no estômago daqueles que queriam escapar aos paradoxos refugiando-se nos sistemas axiomáticos formais. Em 1931, Gödel provou que mesmo para o mais simples dos sistemas formais existiam afirmações cuja verdade ou falsidade não podia ser deduzida dentro desses sistemas. Informalmente, existem afirmações tais, que os axiomas do sistema não as conseguem alcançar. São afirmações indecidíveis. Por esta razão, o teorema de Gödel é parafraseado como «o teorema da incompletude». Este resultado aplica-se ao sistema de Zermelo-Fraenkel, como também a outros sistemas.

Números cardinais O número de elementos dum conjunto finito é fácil de contar, por exemplo A = {1, 2, 3, 4, 5} tem 5 elementos, ou dizemos que a sua «cardinalidade» é 5 e escrevemos *card*(A)=5. Falando em termos simples, a cardinalidade mede o «tamanho» de um conjunto.

Segundo a teoria dos conjuntos de Cantor, o conjunto de fracções **Q** e os números reais **R** são muito diferentes. O conjunto **Q** pode ser posto numa lista, mas o conjunto **R** não (ver página 31). Embora ambos os conjuntos sejam infinitos, o conjunto **R** tem uma ordem de infinito superior à de **Q**. Os matemáticos representam *card*(*Q*) por $\aleph_0$, em hebraico «aleph zero», e *card*(*R*) = *c*. Assim, isto significa $\aleph_0 < c$.

A hipótese do *continuum* Trazida à luz por Cantor em 1878, a hipótese do *continuum* diz que o nível seguinte de infinito depois do infinito de **Q** é o infinito dos números reais *c*. Posto de outra forma, a hipótese do *continuum* afirmou que não existe nenhum conjunto cuja cardinalidade esteja estritamente entre $\aleph_0$ e *c*. Cantor debateu-se com esta ideia e, embora acreditasse que ela era verdadeira, não conseguiu prová-la. Refutá-la significaria encontrar um subconjunto *X* de **R** com $\aleph_0 < card(X) < c$, mas também não o conseguiu fazer.

Este problema era tão importante, que o matemático alemão David Hilbert o colocou no topo da sua famosa lista de 23 problemas pendentes para o século seguinte, apresentada no Congresso Internacional de Matemática em Paris, em 1900.

Gödel acreditava profundamente que a hipótese era falsa, mas não o provou. Provou (em 1938) que a hipótese era compatível com os axiomas de Zermelo-Fraenkel para a teoria dos conjuntos. Um quarto de século mais tarde, Paul Cohen surpreendeu Gödel e os lógicos provando que a hipótese do *continuum* não podia ser deduzida dos axiomas de Zermelo-Fraenkel. É o equivalente a mostrar que axiomas e a negação da hipótese são consistentes. Combinado com o resultado de Gödel de 1938, Cohen mostrou que a hipótese do *continuum* era independente dos restantes axiomas da teoria dos conjuntos.

Este estado de coisas é semelhante na sua natureza à forma como o postulado das paralelas em geometria (ver página 108) é independente dos outros axiomas de Euclides. Esta descoberta levou ao florescimento das geometrias não-euclidianas, que, entre outras coisas, tornaram possível o avanço da teoria da relatividade de Einstein. De forma semelhante, a hipótese do *continuum* pode ser aceite ou rejeitada sem perturbar os outros axiomas para a teoria dos conjuntos. Depois do resultado pioneiro de Cohen, abriu-se todo um novo campo que atraiu gerações de matemáticos a adoptar as técnicas que ele usou para provar a independência da hipótese do *continuum*.

a ideia resumida
Muitos, tratados como um

19 Cálculo

Um cálculo é uma forma de calcular, daí os matemáticos por vezes referirem o «cálculo da lógica», o «cálculo das probabilidades», etc. Porém, todos concordam que na realidade há apenas um Cálculo, puro e simples, escrito com C maiúsculo.

O Cálculo é o ponto principal da matemática. Será agora raro um cientista, engenheiro ou economista não deparar com o Cálculo, ou as suas múltiplas aplicações. Historicamente, está relacionado com Isaac Newton e Gottfried Leibniz, os seus pioneiros no século XVII. As teorias semelhantes dos dois resultaram numa disputa de primazia sobre quem foi o inventor do Cálculo. De facto, ambos chegaram às suas conclusões de modo independente e os seus métodos foram inteiramente diferentes.

Desde então, o Cálculo tornou-se um assunto muito vasto. Cada geração cria técnicas que pensa que devem ser aprendidas pela geração mais jovem, e os manuais actuais têm mais de mil páginas e muitos extras. Para todos esses complementos, o que é absolutamente essencial é a *diferenciação* e a *integração*, os dois picos gémeos do Cálculo, como estabelecidos por Newton e Leibniz. Os termos derivam dos *differentialis* (resolver diferenças ou «desmontar») e *integralis* (somar as partes, ou «juntar») de Leibniz.

Em linguagem técnica, a diferenciação diz respeito à medida da *mudança* e a integração à medida da *área*, mas a jóia da coroa do Cálculo é o «resultado estrela» de que eles são as duas faces da mesma moeda – diferenciação e integração são inversas uma da outra. O Cálculo é realmente um só tópico, e é necessário conhecer os dois lados. Não é de admirar que «o próprio modelo de um moderno major-general» d'*Os Piratas de Penzance*, de Gilbert e Sullivan, proclamasse orgulhosamente:

> *Com muitos factos alegres sobre o quadrado da hipotenusa.*
> *Sou muito bom em cálculo integral e diferencial.*

Cronologia

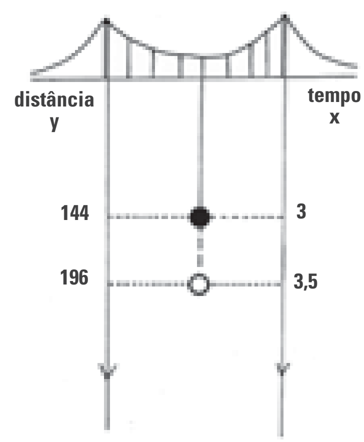
cerca de 450 a.C.
Zenão ridiculariza os infinitesimais com um paradoxo

anos 1660-1670
Newton e Leibniz dão os primeiros passos no Cálculo

Diferenciação Os cientistas apreciam fazer «experiências em pensamento» – Einstein apreciava-o especialmente. Suponhamos que estamos numa ponte bastante acima de um desfiladeiro e estamos prestes a deixar cair uma pedra. O que acontecerá? A vantagem da experiência em pensamento é que não é necessário lá estar em pessoa. Podemos fazer coisas impossíveis como parar a pedra no ar a meio do caminho ou observá-la em movimento lento durante um curto intervalo de tempo.

De acordo com a teoria da gravidade de Newton, a pedra cairá. Até aí, não há nada de surpreendente; a pedra é atraída para a Terra e cairá cada vez mais depressa, enquanto o nosso cronómetro faz tiquetaque. Outra vantagem da experiência em pensamento é que podemos ignorar factores que complicam, como a resistência do ar.

Qual é a velocidade da pedra num dado instante, digamos quando o cronómetro marca exactamente 3 segundos depois de ter sido largada? Como poderemos saber? Podemos certamente medir a velocidade *média*, mas o nosso problema é a velocidade *instantânea*. Como é uma experiência em pensamento, porque não paramos a pedra a meio do ar e depois a deixamos cair uma pequena distância durante mais uma fracção de segundo? Se dividirmos esta distância extra pelo tempo extra, teremos a velocidade média durante o pequeno intervalo de tempo. Registando intervalos de tempo cada vez menores, estaremos cada vez mais perto da velocidade instantânea no lugar onde paramos a pedra. Este processo de limitação é a ideia básica por detrás do Cálculo.



Podemos ser tentados a tornar o pequeno tempo extra igual a zero. Mas, na nossa experiência em pensamento, a pedra não se moveu de todo. Não andou distância nenhuma e não demorou tempo nenhum! Isto dá-nos a velocidade média 0/0, que o filósofo irlandês Bishop Berkeley descreveu como «fantasmas das quantidades mortas». Esta expressão não pode ser determinada – na realidade *não tem sentido*. Seguindo este caminho, somos conduzidos a um pântano numérico.

1734
Berkeley chama a atenção para as fraquezas fundamentais

anos 1820
Cauchy formaliza a teoria de forma rigorosa

1854
Riemann introduz o integral de Riemann

1902
Lebesgue estabelece a teoria do integral de Lebesgue

Para avançarmos, precisamos de alguns símbolos. A fórmula exacta que relaciona a distância de queda y e o tempo x gasto para lá chegar deriva de Galileu:

$$y = 16 \times x^2$$

O factor 16 aparece porque os pés e segundos são as unidades de medida escolhidas. Se quiséssemos saber, digamos, quanto tinha a pedra caído durante os 3 segundos, substituíamos simplesmente $x = 3$ na fórmula e calculávamos a resposta $y = 16 \times 3^2 = 144$ pés. Mas como podemos calcular a *velocidade* da pedra no momento $x = 3$?

Vamos andar mais 0,5 segundos e ver até onde a pedra andou entre os 3 e os 3,5 segundos. Aos 3,5 segundos, deslocava-se $y = 16 \times (3{,}5)^2 = 196$ pés, logo *entre* os 3 e os 3,5, ela caiu 196 – 144 = 52 pés. Dado que a velocidade é a distância a dividir pelo tempo, a velocidade média durante este intervalo de tempo é 52/0,5 = 104 pés por segundo. Este resultado estará próximo da velocidade instantânea no instante $x = 3$, mas também podemos dizer que 0,5 segundos não é suficientemente pequeno. Repetindo o argumento com um intervalo de tempo menor, digamos 0,05 segundos, veremos que a distância de queda é 148,44 – 144 = 4,84 pés, o que dá uma velocidade média de 4,84/0,05 = 96,8 pés por segundo. Esta será decerto mais próxima da velocidade instantânea da pedra aos 3 segundos (quando $x = 3$).

Enfrentemos agora o problema e calculemos a velocidade média da pedra entre x segundos e ligeiramente mais tarde aos $x + h$ segundos. Depois de algum trabalho com os símbolos, verificamos

$$16 \times (2x) + 16 \times h$$

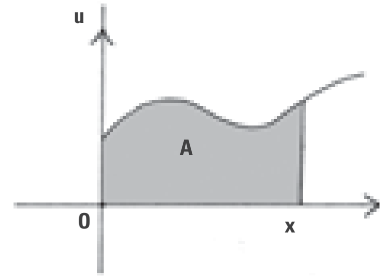
u	du/dx
x^2	$2x$
x^3	$3x^2$
x^4	$4x^3$
x^5	$5x^4$
. . .	. . .
x^n	nx^{n-1}

Conforme tornamos o h cada vez menor, como fizemos passando de 0,5 para 0,05, vemos que o primeiro termo não é afectado (porque não envolve h) e o segundo termo se torna cada vez menor. Concluímos que

$$v = 16 \times (2x)$$

em que v é a velocidade instantânea da pedra no momento x. Por exemplo, a velocidade instantânea da pedra depois de 1 segundo (quando $x = 1$) é 16 x (2 x 1) = 32 pés por segundo; depois de 3 segundos, é $16 \times (2 \times 3)$, o que dá 96 pés por segundo.

Se compararmos a fórmula da distância de Galileu $y = 16 \times x^2$ com a fórmula da velocidade $v = 16 \times (2x)$, a diferença essencial é a mudança de x^2 para $2x$. Este é o efeito da diferenciação, passando de $u = x^2$ para a *derivada* $\dot{u} = 2x$. Newton chamava a $\dot{u} = 2x$ uma «fluxão» e à variável x uma fluente, porque pensava em termos de quantidades a fluir. Hoje escrevemos frequentemente $u = x^2$ e a sua *derivada* como $du/dx = 2x$. Originalmente introduzida por Leibniz, o uso continuado desta notação representa o sucesso da «frescura da Leibniz sobre a caducidade de Newton».



A queda da pedra é um exemplo, mas se tivermos outras expressões representadas por u também podemos calcular a derivada, que pode ser útil noutros contextos. Há um padrão nisto: a derivada é formada pela multiplicação pelo expoente da potência anterior e pela subtracção de 1 para formar uma nova potência.

Integração A primeira aplicação da integração foi a medição de áreas. A medida de uma área debaixo de uma curva é dada pela divisão em tiras aproximadamente rectangulares, cada uma com o comprimento dx. Medindo a área de cada uma e somando-as, obtemos a «soma» e portanto a área total. A notação S para a soma foi introduzida por Leibniz numa forma alongada $\int$. A área de cada uma das tiras rectangulares é udx, logo a área A debaixo da curva de 0 a x é

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

Se a curva for $u = x^2$, a área será encontrada desenhando estreitas tiras rectangulares debaixo da curva, somando-as para calcular a área aproximada, e aplicando um processo de limitação aos seus comprimentos para obtermos a área exacta. Esta resposta dá-nos a área

$$A = x^3/3$$

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

Também podemos calcular o integral para outras curvas (e portanto para outras expressões para u). Como na derivada, há um padrão regular para o integral das potências de x. O integral é formado dividindo pelo «expoente da potência anterior + 1» e somando-lhe 1 para obter a nova potência.

O resultado estrela Se diferenciarmos o integral $A = x^3/3$, obtemos o original $u = x^2$. Se integrarmos a derivada $du/dx = 2x$, também obtemos o original $u = x^2$. A diferenciação é o inverso da integração, uma observação conhecida por Teorema Fundamental do Cálculo, um dos mais importantes teoremas de toda a matemática.

Sem o Cálculo, não haveria satélites nem teoria económica, e a estatística seria muito diferente. Sempre que há mudança, encontramos o Cálculo.

a ideia resumida
Ir ao limite

20 Construções

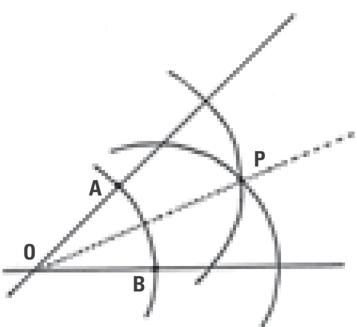
Provar uma negação é muitas vezes difícil, mas alguns dos maiores triunfos em matemática fazem exactamente isso, ou seja, provar que qualquer coisa não pode ser feita. A quadratura do círculo é impossível, mas como podemos prová-lo?

Os antigos gregos tinham quatro grandes problemas de construção:

- trissecção do ângulo (dividir um ângulo em três ângulos menores e iguais);
- duplicação do cubo (construir um segundo cubo com o dobro do volume do primeiro);
- quadratura do círculo (criar um quadrado com a mesma área de um determinado círculo);
- construção de polígonos (construir formas regulares com lados e ângulos iguais).

Para executar estas tarefas, usavam apenas o essencial:

- uma régua para desenhar linhas rectas (e *não* para medir comprimentos);
- um compasso para desenhar círculos.



Bissecção de um ângulo

Se o leitor gosta de escalar montanhas sem cabos, oxigénio, telemóvel e outra parafernália, estes problemas ser-lhe-ão certamente atraentes. Sem o moderno equipamento de medição, as técnicas matemáticas necessárias para provar estes resultados eram sofisticadas e os clássicos problemas de construção da antiguidade só foram resolvidos no século XIX, com técnicas da análise moderna e álgebra abstracta.

Trissecar o ângulo Aqui está uma forma de dividir o ângulo em dois ângulos menores e iguais ou, por outras palavras, bissectá-lo. Primeiro, colocamos o compasso no ponto O e, com qualquer raio,

Cronologia

450 a.C.
Anaxágoras tenta resolver a quadratura do círculo na prisão

1672
Mohr mostra que todas as construções euclidianas podem fazer-se só com um compasso

marcamos *OA* e *OB*. Movendo o compasso para *A*, desenhamos um arco. Fazemos o mesmo em *B*. Chamamos ao ponto de intersecção destes círculos *P*, e com a régua juntamos *O* a *P*. Os triângulos *AOP* e *BOP* são idênticos na forma e, portanto, os ângulos $A\hat{O}P$ e $B\hat{O}P$ são iguais. A recta *OP* é o bissector procurado, dividindo o ângulo em dois ângulos iguais.

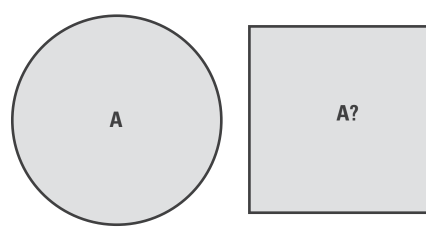
Será que podemos usar uma sequência de acções como estas para dividir um ângulo arbitrário em *três* ângulos iguais? Este é o problema da trissecção do ângulo.

Se o ângulo tem 90 graus, um ângulo recto, não há problema, porque podemos construir o ângulo de 30 graus. Mas, se tiver 60 graus, o ângulo *não* pode ser trissecado. Sabemos que a resposta é 20 graus, mas não há maneira de construir este ângulo só com régua e compasso. Então, sumariando:

- podem *sempre* bissectar-se *todos* os ângulos,
- podem *sempre* trissecar-se *alguns* ângulos, mas
- *alguns* ângulos *nunca* se podem trissecar.

A duplicação do cubo é um problema semelhante conhecido por «problema deliano». Conta a história que os nativos de Delos, na Grécia, consultaram o oráculo por causa de uma praga que estavam a sofrer. Foi-lhes dito que construíssem um altar com o dobro do volume do existente.

Suponhamos que o altar de Delos começou como um cubo tridimensional com todos os lados iguais em comprimento, digamos a. Precisaríamos de construir outro cubo de comprimento b com o dobro do volume. O volume de cada um é a^3 e b^3 estão relacionados por $b^3 = 2a^3$ ou $b^3 = 2a^3$ ou $b = \sqrt[3]{2} \times a$ em que $\sqrt[3]{2}$ é o número que, multiplicado três vezes por si próprio, dá 2 (a raiz cúbica). Se o lado do cubo original for $a = 1$, os nativos de Delos terão de marcar o comprimento $\sqrt[3]{2}$ numa recta. Infelizmente para eles, isto é impossível com régua e compasso, independentemente do engenho utilizado para suportar a pretensa construção.



Quadratura do círculo

Quadratura do círculo Este problema é um pouco diferente e é o mais famoso dos problemas de construção:

> *Construir um quadrado cuja área seja igual à área de um dado círculo.*

1801
Gauss publica *Discourses on Arithmetic*, incluindo uma secção sobre a construção do polígono regular de 17 lados com régua e compasso

1837
Wantzel prova que os problemas clássicos da duplicação do cubo e da trissecção do ângulo não podem ser resolvidos com régua e compasso

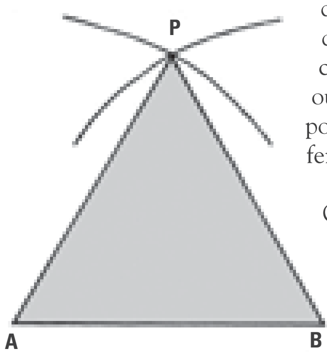
1882
Lindemann prova que não se pode fazer a quadratura do círculo

A expressão «quadratura do círculo» é frequentemente usada para expressar o impossível. A equação algébrica $x^2 - 2 = 0$ tem duas soluções específicas: $x = \sqrt{2}$ ou $x = -\sqrt{2}$. Estes são números irracionais (não podem ser escritos como fracções), mas mostrar que não é possível fazer a quadratura do círculo equivale a mostrar que π não pode ser solução de *nenhuma* equação algébrica. Os números irracionais com esta propriedade chamam-se números transcendentes, porque têm uma irracionalidade «mais alta» do que os seus primos irracionais, como $\sqrt{2}$.

Os matemáticos acreditavam geralmente que o π era um transcendental, mas este «mistério secular» foi difícil de provar até Ferdinand von Lindermann fazer uma modificação a uma técnica de Charles Hermite. Hermite usou-a para tratar o problema menor de provar que a base dos logaritmos naturais, e, era transcendente (ver página 26).

Depois do resultado de Lindermann, poderíamos pensar que o fluxo de artigos do indomável bando dos «quadratores do círculo» pararia. Nem por sombras. Os mais relutantes em aceitar a lógica da prova e outros que nunca tinham ouvido falar dela continuaram a dançar nas margens da matemática.

Construir polígonos Euclides colocou o problema da construção de um polígono regular, uma figura simétrica de muitos lados, como um quadrado ou um pentágono, na qual os lados são todos iguais em comprimento e os lados adjacentes formam ângulos iguais uns aos outros. Na sua famosa obra *Elementos* (Livro 4), Euclides mostrou como podiam construir-se os polígonos de 3, 4, 5 e 6 lados usando apenas ferramentas básicas.



Construir um triângulo equilátero

O polígono com 3 lados, a que normalmente chamamos triângulo equilátero, é particularmente simples de construir. Seja qual for o tamanho que queiramos que o triângulo tenha, chamemos a um ponto A e a outro B, com a distância desejada entre eles. Colocamos o compasso no ponto A e desenhamos um arco de raio AB. Repetimos, pondo o compasso no ponto B e usando o mesmo raio. O ponto de intersecção dos dois arcos é P. Como $AP = AB$ e $BP = AB$, todos os lados do triângulo são iguais. O triângulo real é obtido juntando AB, AP e BP usando uma régua.

Se o leitor pensa que ter uma régua é um luxo, não está sozinho – Dane Georg Mohr também o pensava. O triângulo equilátero é construído encontrando o ponto P e, para isso, só o compasso é necessário – a régua só é utilizada para unir *fisicamente* os pontos. Mohr mostrou que qualquer construção susceptível de se fazer com régua e compasso pode ser feita apenas com o compasso. O italiano Lorenzo Mascheroni provou os mesmos resultados 125 anos depois. Uma característica inovadora do seu livro de 1797 *Geometria del Compasso*, dedicado a Napoleão, foi tê-lo escrito em verso.

Para o problema geral, os polígonos com p lados em que p é um número primo são especialmente importantes. Já construímos o polígono de 3 lados, e Euclides construiu o de 5 lados, mas *não* conseguiu construir o polígono de 7 lados (heptágono). Investigando este problema, um jovem de 17 anos chamado Carl Friedrich Gauss provou a negativa. Deduziu que não é possível construir um polígono de p lados se $p = 7$, 11 ou 13.

Nasceu um príncipe

Carl Friedrich Gauss ficou tão impressionado com a sua demonstração de que o polígono de 17 lados podia ser construído, que pôs de lado o seu projecto de estudar línguas e se tornou matemático. O resto é história – Gauss tornou-se conhecido como «o príncipe dos matemáticos». O polígono de 17 lados é a forma da base do seu memorial em Göttingen, na Alemanha, uma justa homenagem ao seu génio.

Mas Gauss também fez uma prova positiva: concluiu que é possível construir um polígono de 17 lados. Na realidade, Gauss foi mais longe e provou que é possível construir um polígono de p lados se, e só se, o número primo p tiver a forma

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

Os números com esta forma chamam-se números de Fermat. Se os determinarmos para $n = 0, 1, 2, 3$ e 4, concluímos que são os números primos $p = 3, 5, 17, 257$ e 65 537, e estes correspondem a polígonos construíveis com p lados.

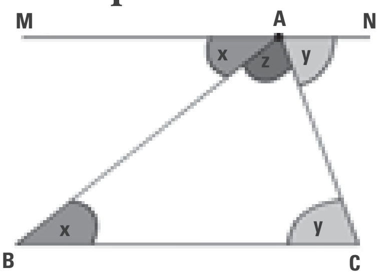
Quando experimentamos com $n = 5$, o número de Fermat é $p = 2^{32} + 1 = 4\,294\,967\,297$. Pierre de Fermat conjecturou que seriam todos números primos, mas infelizmente este não é um número primo, porque $4\,294\,967\,297 = 641 \times 6\,700\,417$. Se utilizarmos $n = 6$ ou 7 na fórmula, os resultados serão números de Fermat enormes mas, como com o 5, nenhum deles é primo.

Haverá outros números de Fermat que sejam primos? A sensatez diz-nos que não, mas ninguém sabe ao certo.

a ideia resumida
Com uma régua e um compasso . . .

21 Triângulos

O facto mais óbvio acerca de um triângulo é que é uma figura com três lados e três ângulos (donde tri-ângulo). A trigonometria é a teoria que nos permite tirar as «medidas do triângulo», quer seja a amplitude dos ângulos, o comprimento dos lados ou a área. Esta figura – uma das mais simples de todas – desperta em nós um interesse contínuo.



A lenda do triângulo Há um argumento engenhoso para mostrar que os ângulos de qualquer triângulo somam dois ângulos rectos ou 180 graus. A partir do ponto ou «vértice» *A* de qualquer triângulo, desenhamos uma recta *MAN* paralela à base *BC*.

O ângulo $A\hat{B}C$ a que chamaremos *x*, é igual ao ângulo $B\hat{A}M$, porque são ângulos alternos e *MN* e *BC* são paralelos. Os outros dois ângulos alternos são iguais a *y*. O ângulo no ponto *A* é igual a 180 graus (metade de 360 graus) e é $x + y + z$, o que é a soma dos ângulos do triângulo. No final desta explicação, escreve-se muitas vezes QED. Claro que supomos que o triângulo está desenhado numa superfície plana como esta folha de papel. Os ângulos de um triângulo desenhado numa bola (um triângulo esférico) não somam 180 graus, mas isso é outra história.

Euclides provou muitos teoremas sobre triângulos, tendo sempre a certeza de o fazer de forma dedutiva. Mostrou, por exemplo, que, «em qualquer triângulo, quaisquer dois lados são maiores do que o outro». Hoje chama-se a isso «desigualdade triangular», algo importante em matemática abstracta. Os epicuristas, com a sua abordagem realista à vida, reclamam que isso não necessita de prova, dado que é evidente até para um burro. Se um fardo de palha fosse colocado num vértice e o burro no outro, argumentam, o animal dificilmente atravessaria dois lados para satisfazer a sua fome.

O teorema de Pitágoras O maior de todos os teoremas do triângulo é o teorema de Pitágoras, que ainda é utilizado na matemática moderna –

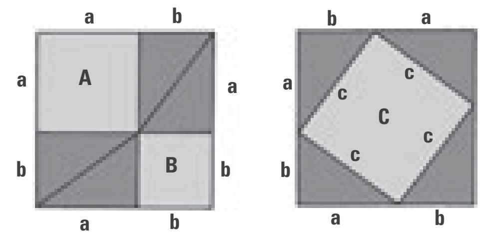
Cronologia

1850 a.C.
Os babilónios conhecem o «teorema de Pitágoras»

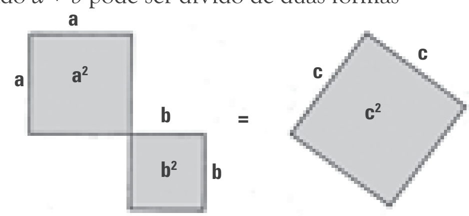
1335
Richard de Wallingford escreve um tratado revolucionário sobre trigonometria

embora haja algumas dúvidas de que tenha sido Pitágoras o primeiro a descobri-lo. A afirmação mais conhecida do teorema respeita à álgebra, $a^2 + b^2 = c^2$, mas na realidade Euclides refere formas quadradas: «Em triângulos rectângulos, o quadrado no lado oposto ao ângulo recto é igual aos quadrados dos lados que limitam o ângulo recto.»

A prova de Euclides é a proposição 47, no Livro 1 dos *Elementos*, uma prova que se tornou causa de ansiedade para gerações de alunos que tentaram a custo decorá-la. Existem centenas de provas. A favorita segue mais o espírito de Bhaskara, do século XII, do que uma prova de Euclides, de 300 a.C.



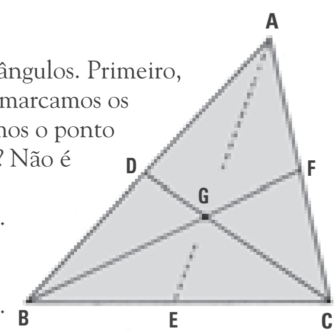
É uma prova «sem palavras». Na figura, o quadrado com lado $a + b$ pode ser divido de duas formas diferentes.



Dado que os quatro triângulos iguais (sombreados a escuro) são comuns a ambos os quadrados, podemos removê-los e continuamos com uma igualdade de área. Se olharmos para as áreas das figuras restantes, emerge a expressão familiar

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

A recta de Euler São possíveis centenas de proposições sobre triângulos. Primeiro, consideremos os pontos médios dos lados. Em qualquer triângulo *ABC* marcamos os pontos médios *D*, *E* e *F* nos seus lados. Ligamos *B* a *F* e *C* a *D* e marcamos o ponto em que se cruzam como G. Depois, ligamos A a *E*. A linha passa por G? Não é óbvio que seja necessariamente assim sem mais nenhum raciocínio. Na realidade, ela passa por G, a que se chama «centróide» do triângulo. Este é o centro de gravidade do triângulo.

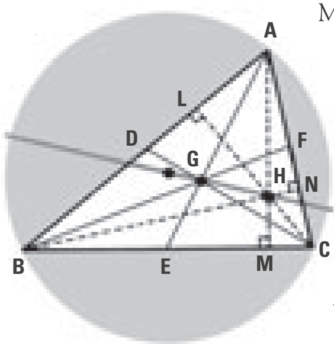


Existem literalmente centenas de centros relacionados com o triângulo. Outro é o ponto *H*, em que as alturas (as linhas desenhadas perpendicularmente de um vértice para a base – mostradas a tracejado na página 86) se cruzam. Esse ponto é o «ortocentro». Há outro centro, chamado «circuncentro», O, em que cada uma das linhas (conhecidas como «perpendiculares») em *D*, *E*, e *F* se encontram (não mostrado). Este é o centro da circunferência que passa por *A*, *B* e C.

1571
Richard Viète publica um livro sobre trigonometria e tábuas trigonométricas

1822
Karl Feuerbach descreve a circunferência dos nove pontos

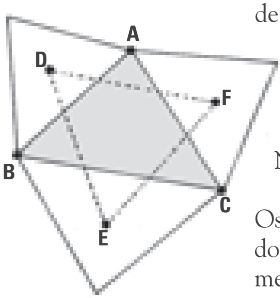
1873
Brocard produz o seu trabalho exaustivo sobre o triângulo



A recta de Euler

Mas há mais. Em qualquer triângulo *ABC*, os centros G, *H* e O, respectivamente o centróide, o ortocento e o circuncentro, estão sobre a mesma recta, chamada a «*recta de Euler*». No caso de um triângulo equilátero (todos os lados são iguais), estes três pontos coincidem e o ponto resultante é indubitavelmente o centro do triângulo.

O teorema de Napoleão Em qualquer triângulo *ABC*, os triângulos equiláteros podem ser construídos em cada lado e, a partir dos seus centros, pode ser construído um novo triângulo *DEF*. O teorema de Napoleão afirma que, para *qualquer* triângulo *ABC*, o triângulo *DEF* é um triângulo *equilátero*.



O teorema de Napoleão

O teorema de Napoleão foi publicado num jornal inglês em 1825, alguns anos após a morte daquele em Santa Helena, em 1821. As capacidades matemáticas de Napoleão ajudaram-no indubitavelmente a entrar na escola de artilharia. Mais tarde, já imperador, viria a conhecer os matemáticos de topo em Paris. Infelizmente, não existem evidências que nos possam levar mais longe e o «teorema de Napoleão» é, como muitos outros resultados matemáticos, atribuído a alguém que pouco teve a ver com a sua descoberta ou a sua prova. Na realidade, é um teorema que é frequentemente reinventado e ampliado.

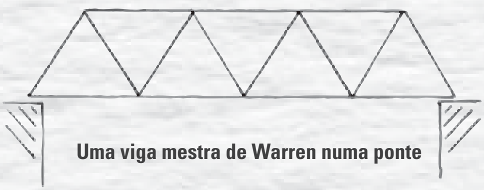
Os dados necessários para determinar um triângulo consistem no conhecimento do comprimento de um lado e dois ângulos. Usando a trigonometria, podemos medir tudo o resto.

No levantamento de áreas de terreno para desenhar mapas, é muito útil ter um «mundo plano» e presumir que os triângulos são planos. Constrói-se uma rede de triângulos começando com uma linha de base *BC* de comprimento conhecido, escolhendo um ponto afastado *A* (o ponto de triangulação) e medindo os ângulos $A\hat{B}C$ e $A\hat{C}B$ com um teodolito. Com a trigonometria, sabemos tudo sobre o triângulo ABC e o agrimensor continua, fixando o novo ponto de triangulação na nova linha de base *AB* ou *AC* e repetindo a operação para obter uma rede de triângulos. O método tem a vantagem de permitir o mapeamento de zonas inóspitas envolvendo obstáculos como terrenos pantanosos, lamaçais, areias movediças e rios.

Foi usado como base para o Grande Levantamento Trigonométrico da Índia que começou na década de 1800 e durou 40 anos. O objectivo era pesquisar e mapear ao longo do Grande Arco Meridional desde o cabo Comorin, no Sul, até aos Himalaias no Norte, uma distância de mais ou menos 1500 milhas. Para garantir a máxima precisão na medida dos ângulos, Sir George Everest providenciou o fabrico de dois teodolitos gigantes em Londres, pesando

Construir com triângulos

Os triângulos são indispensáveis na construção. A sua utilidade e força baseiam-se no facto que os tornou indispensáveis nos levantamentos – a sua rigidez. Podemos forçar um quadrado ou um rectângulo a mudar de forma, mas não um triângulo. Em construção, tira-se partido da junção de triângulos, como se pode ver na construção de telhados. Uma importante descoberta ocorreu na construção de pontes.



Uma viga mestra de Warren numa ponte

Um suporte de Warren pode suportar uma grande carga, considerando o seu peso. Foi patenteado em 1848 por James Warren, e a primeira ponte desenhada segundo o seu modelo foi construída na estação de London Bridge dois anos mais tarde. O uso de triângulos equiláteros provou ser mais sólido do que desenhos semelhantes baseados em triângulos isósceles, os triângulos em que só é necessário que dois dos lados sejam iguais.

conjuntamente uma tonelada e exigindo equipas de doze homens para o transporte. Era vital que os ângulos fossem bem medidos. A precisão da medida era obrigatória e muito discutida, mas foi o humilde triângulo que ocupou o centro das operações. Os vitorianos tiveram de o fazer sem GPS, embora tivessem computadores – computadores humanos. Uma vez calculados todos os comprimentos do triângulo, o cálculo da área é simples. Mais uma vez, o triângulo é a unidade. Existem muitas fórmulas para calcular a área A de um triângulo, mas a mais notável é a fórmula de Herão de Alexandria:

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

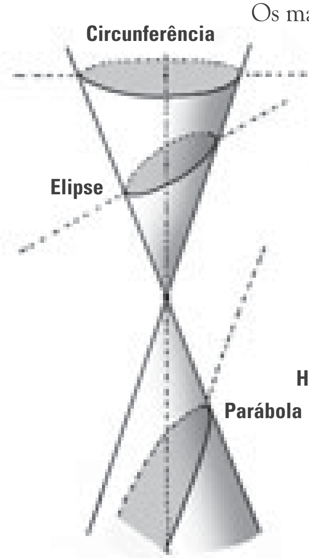
Pode aplicar-se a qualquer triângulo e nem sequer é necessário saber os ângulos. O símbolo *s* significa semiperímetro de um triângulo cujos lados têm o comprimento *a*, *b* e c. Por exemplo, se um triângulo tem lados 13, 14 e 15, o perímetro é 13 + 14 + 15 = 42, logo *s* = 21. Completando o cálculo, $A = \sqrt{(21 \times 8 \times 7 \times 6)} = \sqrt{7056} = 84$. O triângulo é um objecto familiar, sejamos crianças que brincam com formas simples ou investigadores que lidam no dia-a-dia com a desigualdade triangular em matemática abstracta. A trigonometria é a base para se fazerem cálculos com triângulos e as funções de seno, co-seno e tangente são ferramentas para os descrevermos, permitindo-nos fazer cálculos precisos com aplicações práticas. O triângulo recebe muitas atenções, mas é surpreendente que ainda tanto esteja por descobrir sobre as três linhas que formam esta figura tão básica.

a ideia resumida
Os três lados de uma história

22 Curvas

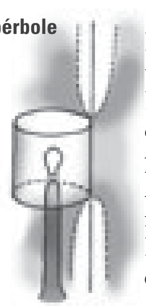
É fácil desenhar uma curva. Os artistas fazem-no a toda a hora. Os arquitectos desenham novos edifícios ou quarteirões modernos numa curva em crescente. Os lançadores de basebol lançam a bola em curva. Os lançadores de peso fazem um movimento em curva antes do lançamento, e quando os futebolistas chutam para golo, a bola segue uma curva. Contudo, não é fácil responder à pergunta «O que é uma curva?».

Os matemáticos têm vindo a estudar as curvas desde há séculos e de vários pontos de vista. Tudo começou com os gregos e as curvas que eles estudaram, hoje chamadas curvas «clássicas».



As secções cónicas

Curvas clássicas A primeira família no domínio das curvas clássicas são as chamadas «secções cónicas». Os membros desta família são a circunferência, a elipse, a parábola e a hipérbole. A secção cónica é formada por um cone duplo, como dois cones de gelado juntos com um de cabeça para baixo. Cortando através dela com uma superfície plana, as curvas da intersecção serão a circunferência, a elipse, a parábola, ou a hipérbole, dependendo da inclinação do plano de corte relativamente ao eixo vertical do cone.



Podemos pensar numa secção cónica como a projecção de uma circunferência num ecrã. Os raios de luz de uma lâmpada de um candeeiro de mesa com um abajur cilíndrico formam um duplo cone de luz de onde a luz se projecta através dos rebordos superior e inferior. A imagem no tecto será uma circunferência, mas, se inclinarmos o candeeiro, a circunferência torna-se uma elipse. Por outro lado, a imagem na parede dará a curva com duas partes, a hipérbole.

Cronologia

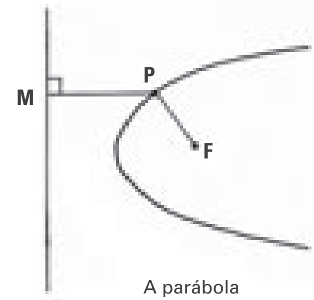
cerca de 300 a.C.
Euclides define as secções cónicas

cerca de 250 a.C.
Arquimedes investiga as espirais

cerca de 225 a.C.
Apolónio de Perga publica *Cónicas*

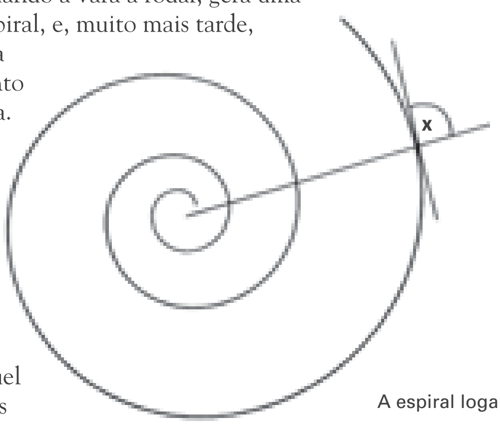
As cónicas também podem ser descritas pela forma como os pontos se movimentam no plano. Este é o método do *locus* muito apreciado pelos gregos, e ao contrário da definição projectiva, envolve comprimento. Se um ponto se move de forma que a sua distância de um ponto fixo é sempre a mesma, obtemos uma circunferência. Se um ponto se move de forma que a soma das suas distâncias a dois pontos fixos (os focos) é constante, obtemos uma elipse (quando os dois focos são o mesmo, a elipse torna-se uma circunferência). A elipse foi a chave para o movimento dos planetas. Em 1609, o astrónomo alemão Johannes Kepler anunciou que os planetas se moviam à volta do Sol em elipses, rejeitando a antiga ideia de órbitas circulares.

Não é tão óbvio o ponto que se move de forma que a sua distância de um ponto (foco *F*) seja a mesma que a sua distância perpendicular a uma dada recta (a directriz). Nesse caso, obtemos uma parábola. A parábola tem uma panóplia de propriedades úteis. Se uma fonte de luz for colocada no foco *F*, os raios de luz emitidos são paralelos a *PM*. Noutro plano, se forem enviados por satélite e recebidos numa antena parabólica, os sinais de televisão juntam-se no foco e são enviados para o televisor.



A parábola

Se uma vara for rodada sobre qualquer ponto fixo nela mesma, desenha uma circunferência, mas se o ponto puder mover-se ao longo da vara em direcção ao exterior continuando a vara a rodar, gera uma espiral. Pitágoras gostava especialmente da espiral, e, muito mais tarde, Leonardo da Vinci passou dez anos da sua vida a estudar os diferentes tipos de espiral, enquanto René Descartes escreveu um tratado sobre esta. A espiral logarítmica também é chamada espiral equiangular, porque faz o mesmo ângulo com o raio e a tangente no ponto onde o raio intersecta a espiral.



A espiral logarítmica

Jacob Bernoulli, do famoso clã de matemáticos suíços, estava tão enamorado da espiral logarítmica, que a quis gravada no seu túmulo em Basileia. O «renascentista» Emanuel Swedenborg considerava a espiral como a mais perfeita das formas. Uma espiral tridimensional que gira

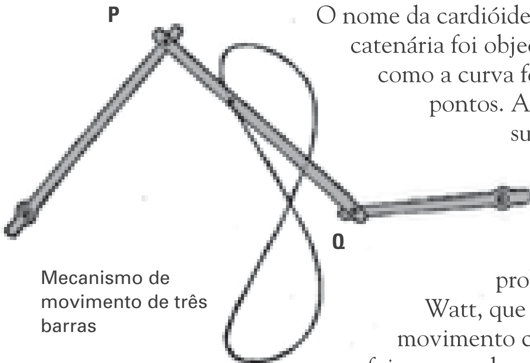
1704
Newton classifica as curvas cúbicas

1890
Peano prova que um quadrado sólido é uma curva (a curva de preenchimento de espaço)

anos 1920
Menger e Urysohn definem as curvas como parte da topologia

à volta de um cilindro chama-se hélice. Duas – uma dupla hélice – formam a estrutura básica do ADN.

Há muitas curvas clássicas, tais como a limaçon, a lemniscata e várias ovais. O nome da cardióide deriva da sua forma em coração. A curva catenária foi objecto de pesquisa no século XVIII e foi identificada como a curva formada por uma corrente pendurada entre dois pontos. A parábola é a curva que existe numa ponte suspensa entre dois pilares verticais.



Mecanismo de movimento de três barras

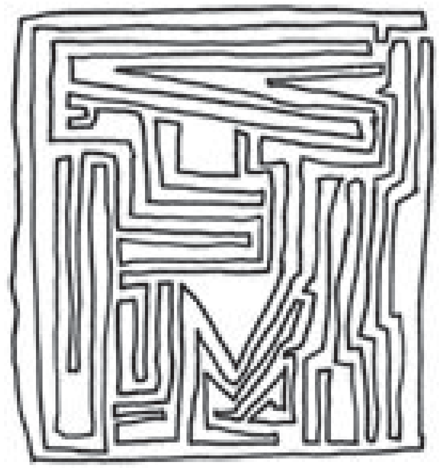
Um aspecto da pesquisa sobre curvas do século XIX era as curvas geradas por barras mecânicas. A questão era uma extensão do problema resolvido pelo engenheiro escocês James Watt, que desenhou barras articuladas para transformar movimento circular em movimento linear. Na era do vapor, foi um grande passo em frente.

A mais simples dessas ferramentas mecânicas é o mecanismo de três barras, em que as barras são ligadas entre si e têm posições fixas nos seus extremos. Se a «barra de acoplagem» *PQ* se move de qualquer maneira, o lugar geométrico de um ponto nela é uma curva de grau seis, uma «curva sêxtica».

Curvas algébricas Depois de Descartes, que revolucionou a geometria com a introdução das coordenadas x, y e z e os eixos cartesianos a que deu nome, as cónicas puderam ser estudadas como equações algébricas. Por exemplo a circunferência de raio 1 tem por equação $x^2 + y^2 = 1$, que é uma equação do segundo grau como são *todas* as cónicas. Surgiu um novo ramo da geometria chamado geometria algébrica.

Num estudo profundo, Isaac Newton classificou as curvas descritas por equações de grau três, ou curvas cúbicas. Comparadas com as quatro cónicas básicas, descobriram-se 78 tipos, agrupados em cinco classes. A explosão do número de tipos diferentes continua com as curvas quárticas, com tantos tipos que a classificação completa nunca chegou a ser feita.

A história não acaba com o estudo das curvas enquanto equações algébricas. Muitas curvas, como as catenárias, as ciclóides (curvas traçadas por um ponto numa roda giratória) e as espirais, não são fáceis de expressar como equações algébricas.



Uma simples curva fechada de Jordan

Uma definição O que os matemáticos perseguiam era uma definição da própria curva, e não apenas exemplos específicos. Camille Jordan propôs uma teoria de curvas baseada na definição de curva em termos de pontos variáveis.

Eis um exemplo. Se $x = t^2$ e $y = 2t$, para diferentes valores de t, obtemos muitos pontos diferentes que podemos escrever como coordenadas (x, y). Por exemplo, se $t = 0$, obtemos o ponto (0,0), $t = 1$ dá-nos o ponto (1, 2), e por aí adiante. Se marcarmos estes pontos nos eixos x-y e «ligarmos os pontos», obtemos uma parábola. Jordan redefiniu esta ideia de como os pontos podem ser traçados. Para ele, esta era a definição de curva.

As curvas de Jordan podem ser intricadas, mesmo quando são como a circunferência, no sentido em que são «simples» (não se cruzam a si próprias) e «fechadas» (não têm princípio nem fim). O famoso teorema de Jordan tem sentido. Segundo ele, uma curva simples fechada tem um interior e um exterior. A sua aparente «evidência» é um engano.

Em Itália, Giuseppe Peano causou sensação quando, em 1890, mostrou que, de acordo com a definição de Jordan, um quadrado preenchido é uma curva. Podia organizar os pontos no quadrado para que *todos* pudessem ser «traçados» e, ao mesmo tempo, obedecer à definição de Jordan. A isto se chamou uma curva de preenchimento de espaço, que abriu uma brecha na definição de Jordan – claramente um quadrado não é uma curva no sentido convencional.

Exemplos de curvas que preenchem espaços e outros exemplos patológicos provocaram o regresso dos matemáticos aos estiradores e à reflexão sobre as bases da teoria das curvas. Levantou-se o problema de desenvolver uma melhor definição de curva. No início do século XX, esta tarefa conduziu os matemáticos ao novo campo da topologia.

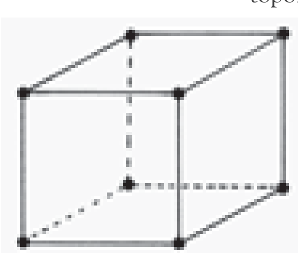
a ideia resumida
Andar às curvas

23 Topologia

A topologia é um ramo da geometria que lida com as propriedades das superfícies e formas genéricas, mas é indiferente à medição de comprimentos ou ângulos. No topo da agenda estão as qualidades que não se alteram quando as formas são transformadas noutras. É-nos permitido puxar e empurrar em qualquer direcção e, por essa razão, a topologia é por vezes descrita como a «geometria da folha de borracha». Os topólogos são pessoas que não conseguem estabelecer a diferença entre um *donut* e uma chávena de café!

Um *donut* é uma superfície com um único buraco. Uma chávena de café é o mesmo, tomando o buraco a forma de pega. Eis como um *donut* pode ser transformado numa chávena de café.

Classificação dos poliedros As formas mais básicas estudadas pelos topólogos são os poliedros («poli» significa «muitos» e «edro», «face»). Um exemplo de um poliedro é um cubo, com 6 faces, 8 vértices (pontos de confluência das faces) e 12 arestas (as linhas que ligam os vértices). O cubo é um poliedro regular porque:



- todas as faces têm a mesma forma regular,
- todos os ângulos entre as arestas que se encontram num vértice são iguais.

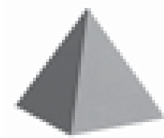
Cronologia

cerca de 300 a.C.
Euclides mostra que existem 5 poliedros regulares

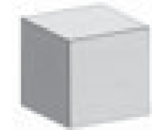
cerca de 250 a.C.
Arquimedes investiga os poliedros truncados

1752
Euler fornece a sua fórmula para o número de vértices, arestas e faces num poliedro

A topologia é um campo relativamente novo, mas mesmo assim remonta aos gregos e, na realidade, o resultado último dos *Elementos* de Euclides é mostrar que existem *exactamente* cinco poliedros regulares. Estes são os «sólidos platónicos»:



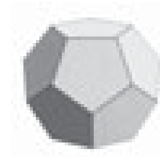
Tetraedro



Cubo

- tetraedro (com 4 faces triangulares),
- cubo (com 6 faces quadradas),
- octaedro (com 8 faces triangulares),
- dodecaedro (com 12 faces pentagonais),
- icosaedro (com 20 faces triangulares).

Octaedro



Dodecaedro

Se deixarmos cair a condição de cada face ter de ser igual, estamos no domínio dos sólidos de Arquimedes, que são semi-regulares. Podem gerar-se exemplos a partir dos sólidos platónicos. Se cortarmos (truncarmos) alguns cantos do icosaedro, temos a forma usada numa bola de futebol. As 32 faces que formam os painéis consistem em 12 pentágonos e 20 hexágonos. Existem 90 arestas e 60 vértices. Também é a forma das moléculas de buckminsterfulereno, assim chamadas em homenagem ao visionário Buck Minster Fuller, criador das cúpulas geodésicas. Os fulerenos são uma forma de carbono recentemente descoberta, C_{60}, com um átomo de carbono em cada vértice.

Icosaedro

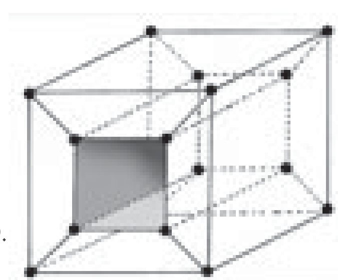
Icosaedro truncado

Fórmula de Euler A fórmula de Euler diz que o número de vértices V, arestas A e faces F de um poliedro está relacionado pela seguinte fórmula

$$V - A + F = 2$$

Por exemplo, para um cubo, $V = 8$, $A = 12$ e $F = 6$ logo $V - A + F = 8 - 12 + 6 = 2$ e, para os buckminsterfulerenos, $V - A + F = 60 - 90 + 32 = 2$. Este teorema na realidade desafia a própria noção de poliedro.

Se um cubo for atravessado por um túnel, é realmente um poliedro? Para esta forma, $V = 16$, $A = 32$, $F = 16$ e $V - A + F = 16 - 32 + 16 = 0$. A fórmula de Euler não funciona. Para reivindicar a correcção da



O cubo com um túnel

1858
Möbius e Listing introduzem a tira de Möbius

1961
Stephen Smale prova a conjectura de Poincaré para dimensões superiores a 4

1982
Michael Freedman prova a conjectura de Poincaré para a dimensão 4

2002
Perelman prova a conjectura de Poincaré para a dimensão 3

fórmula, o tipo de poliedro podia ser limitado aos que não tivessem túneis. Em alternativa, a fórmula podia ser generalizada para incluir esta peculiaridade.

Classificação de superfícies Um topólogo pode ver o *donut* e a chávena de café como idênticos, mas que tipo de superfície é *diferente* da do *donut*? Uma candidata é a bola de borracha. Não há forma de transformar o *donut* numa bola, dado que o *donut* tem um buraco e a bola não. Esta é a principal diferença entre as duas superfícies. Assim, uma forma de classificar superfícies é o número de buracos que elas contêm.

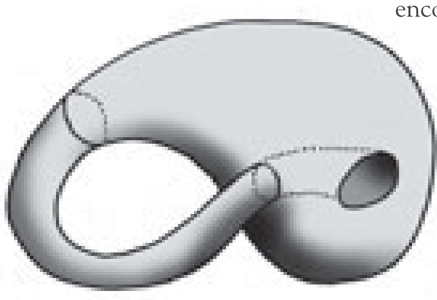
Consideremos uma superfície com r buracos e dividamo-la em regiões delimitadas pelas arestas que juntam vértices situados na superfície. Uma vez isto feito, podemos contar o número de vértices, arestas e faces. Para qualquer divisão, a expressão de Euler $V - A + F$ tem sempre o mesmo valor, chamado «característica de Euler para a superfície»:

Tira de Möbius

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

Se a superfície não tem buracos ($r = 0$), como no caso dos poliedros comuns, a fórmula reduz-se à de Euler $V - A + F = 2$. No caso de um buraco ($r = 1$), como no caso com o cubo com um túnel, $V - A + F = 0$.

Superfícies unilaterais Normalmente, uma superfície terá dois lados. O exterior de uma bola é diferente do seu interior e a única forma de atravessar de um lado para o outro é fazer um furo na bola – uma operação de corte que não é permitida em topologia (pode-se esticar, mas não cortar). Uma folha de papel é outro exemplo de uma superfície com dois lados. O único lado onde um lado encontra o outro é ao longo da curva delimitadora formada pelas arestas do papel.



Garrafa de Klein

A ideia de uma superfície unilateral parece rebuscada. Contudo, August Möbius, matemático e astrónomo alemão do século XIX, descobriu uma superfície unilateral famosa. A forma de construir uma tal superfície é segurar uma fita de papel, dar-lhe uma volta e colar as pontas. O resultado é uma «tira de Möbius», uma superfície unilateral com uma delimitação curva. Podemos pegar num lápis e desenhar uma linha ao longo do seu meio. Pouco tempo depois, estaremos onde começámos!

É mesmo possível ter uma superfície unilateral que não tenha uma curva delimitadora. É a «garrafa de Klein», assim chamada em homenagem ao matemático alemão Felix Klein. O que é particularmente impressionante nesta garrafa é que ela não se intersecta a si própria. Contudo, não é possível fazer um modelo de uma garrafa de Klein no espaço tridimensional sem que exista uma intersecção física, porque ela existe em quatro dimensões, onde não terá intersecções.

Ambas estas superfícies são exemplos do que os topólogos chamam «variedades» – superfícies geométricas que parecem folhas de papel bidimensional quando se vêem separadamente porções *pequenas*. Visto que a garrafa de Klein não tem fronteira, é chamada uma 2-variedade fechada.

A conjectura de Poincaré Durante mais de um século, um conhecido problema em topologia foi a famosa conjectura de Poincaré, assim chamada em homenagem a Henri Poincaré. A conjectura centra-se na relação entre álgebra e topologia.

A parte da conjectura que ficou por resolver até recentemente aplicava-se a 3-variedades fechadas. Isto pode ser complicado – imaginemos uma garrafa de Klein com uma dimensão extra. Poincaré conjecturou que algumas 3-variedades que tinham todas as características algébricas para serem esferas tridimensionais, na realidade, eram esferas. É como se andássemos à volta de uma bola gigante e todas as pistas que recebêssemos indicassem que era uma esfera, mas, como não conseguíssemos ver todo o conjunto, nos perguntássemos se seria realmente uma esfera.

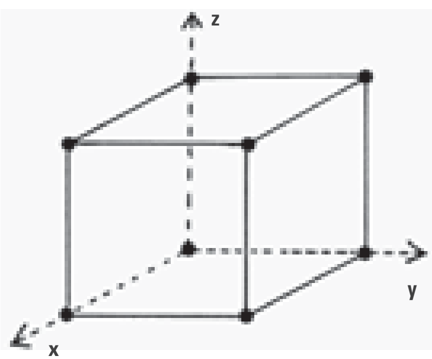
Ninguém conseguiu provar a conjectura de Poincaré para 3-variedades. Era verdadeira ou falsa? Tinha sido provada para todas as outras dimensões, mas o caso das 3-variedades resistia. Existiram muitas provas falsas, até 2002, quando se reconheceu que Grigori Perelmen, do Instituto de Stekov de São Petersburgo, a tinha finalmente provado. Como a solução de todos os outros grandes problemas em matemática, as técnicas de resolução da conjectura de Poincaré estavam fora da sua área próxima, numa técnica relacionada com a difusão de calor.

a ideia resumida
Dos *donuts* às chávenas de café

24 Dimensão

Leonardo da Vinci escreveu no seu bloco de notas: «A ciência da pintura começa com o ponto, depois a linha, o plano vem em terceiro, e o quarto é o corpo nas suas vestes de plano.» Na hierarquia de Da Vinci, o ponto tem dimensão zero, a recta a dimensão um, o plano é bidimensional e o espaço é tridimensional. O que poderia ser mais óbvio? Era a forma como o ponto, a recta, o plano e a geometria sólida tinham sido difundidos pelo geómetra grego Euclides, e Da Vinci estava a seguir a apresentação de Euclides.

Há milénios que encaramos o espaço físico como tridimensional. No espaço físico, podemos mover-nos para fora desta página ao longo do eixo x, ou atravessá-la horizontalmente ao longo do eixo y, ou subirmos ao longo do eixo z, ou fazer qualquer combinação destes movimentos. Relativamente à origem (onde os eixos se encontram), cada ponto tem um conjunto de coordenadas espaciais especificadas pelos valores de x, y e z e representado por (x, y, z).



O espaço de três dimensões

Um cubo, bem como tudo aquilo que é sólido, tem claramente estas três dimensões. Na escola, é-nos normalmente ensinada a geometria do plano, que é bidimensional, e a seguir as três dimensões – a «geometria sólida» – e paramos aí.

Por volta do início do século XIX, os matemáticos começaram a interessar-se pelas quatro dimensões e mesmo pela mais elevada matemática n-dimensional. Muitos filósofos e matemáticos começaram a perguntar-se se existiam mais dimensões.

Dimensões físicas mais elevadas Muitos matemáticos importantes do passado pensavam que não podiam ser imaginadas quatro dimensões. Interrogavam-se sobre a realidade das quatro dimensões, e tornou-se um desafio explicá-las.

Cronologia

cerca de 300 a.C.
Euclides descreve um mundo tridimensional

1877
Cantor é surpreendido pelas suas controversas descobertas na teoria da dimensão

Uma forma usual de explicar porque é que as quatro dimensões podiam ser possíveis era o regresso às duas dimensões. Em 1884, Edwin Abbott, professor e teólogo inglês, publicou um livro muito popular sobre «planolandenses», que viviam no plano bidimensional. Não conseguiam ver os triângulos, os quadrados nem os círculos que existiam na «Planolândia», porque não conseguiam sair para a terceira dimensão para os ver. A sua visão era fortemente limitada. Tinham a mesma dificuldade em pensar numa terceira dimensão que nós temos em pensar numa quarta. Ler Abbott coloca-nos no estado de espírito para aceitar a quarta dimensão.

A necessidade de ponderar a real existência de um espaço quadridimensional tornou-se mais premente quando Einstein entrou em cena. A geometria quadridimensional tornou-se mais plausível, e mesmo compreensível, por a dimensão extra no modelo de Einstein ser o *tempo*. Ao contrário de Newton, Einstein concebeu o tempo como unido com o espaço num *continuum* espaço--tempo quadridimensional. Einstein decretou que vivemos num mundo quadridimensional com quatro coordenadas (x, y, z, t) em que t designa tempo.

Hoje, o mundo quadridimensional de Einstein parece bastante inofensivo e trivial. Um modelo mais recente da realidade física baseia-se na teoria das «cordas». Nela, as familiares partículas como os electrões são manifestações de cordas extremamente minúsculas em vibração. A teoria das cordas sugere uma substituição do *continuum* espaço-tempo quadridimensional por uma versão de dimensão superior. As pesquisas em curso sugerem que as dimensões do *continuum* espaço-tempo para a teoria das cordas deverão ser 10, 11 ou 26, dependendo de futuras suposições e diferenças de pontos de vista.

Um enorme íman de 2000 toneladas no CERN, perto de Genebra, na Suíça, concebido para testar colisão de partículas a altas velocidades, pode ajudar a resolver a questão. O íman tem como objectivo revelar a estrutura da matéria e, como subproduto, pode apontar para uma melhor teoria e para a resposta «correcta» da dimensionalidade. A melhor resposta parece ser que vivemos num universo de 11 dimensões.

Hiperespaço Ao contrário das dimensões físicas mais elevadas, não há absolutamente problema nenhum com um *espaço matemático* de mais de três dimensões. O espaço matemático pode ter qualquer número de dimensões. Desde o início do século XIX que os matemáticos têm vindo a usar n variáveis

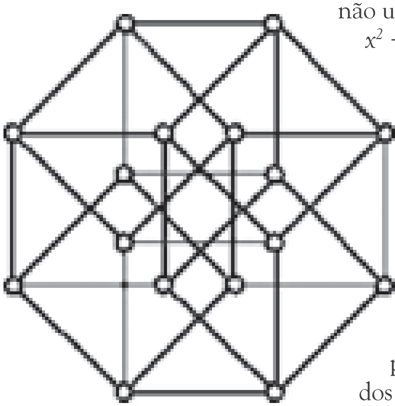
1909
O trabalho de Brouwer modifica a nossa noção de dimensão

1919
Hausdorff introduz a noção de «dimensão de Hausdorff» fraccionária

1970
A teoria das cordas concebe que o nosso universo tem 10, 11 ou 26 dimensões

no seu trabalho. George Green, um moleiro de Nottingham que explorou a matemática da electricidade, e os matemáticos puros A. L. Cauchy, Arthur Cayley e Hermann Grassmann, todos descreveram a sua matemática em termos do hiperespaço *n*-dimensional. Parece não haver nenhuma boa razão para limitar a matemática e todas as razões para ganhar em elegância e claridade.

A ideia por trás das *n* dimensões é meramente uma extensão das coordenadas tridimensionais (x, y, z) para um número não especificado de variáveis. Uma circunferência com duas dimensões tem por equação $x^2 + y^2 = 1$, uma esfera com três dimensões tem por equação $x^2 + y^2 + z^2 = 1$, logo porque não uma hiperesfera de quatro dimensões com a equação $x^2 + y^2 + z^2 + w^2 = 1$?



O cubo quadridimensional

Os oito cantos de um cubo em três dimensões têm coordenadas da forma (x, y, z), em que cada um de x, y, z é 0 ou 1. O cubo tem seis faces, cada uma das quais é um quadrado e há $2 \times 2 \times 2 = 8$ cantos. E o cubo em quatro dimensões? Terá coordenadas da forma (x, y, z, w), em que cada um de x, y, z e w é 0 ou 1. Logo, há $2 \times 2 \times 2 \times 2 = 16$ possíveis cantos para o cubo em quatro dimensões, e 8 faces, cada uma das quais é um cubo. Na realidade, não podemos ver este cubo quadrimensional, mas podemos ilustrá-lo nesta folha de papel. A ilustração mostra a projecção do cubo quadridimensional que existe na imaginação dos matemáticos. As faces cúbicas só podem ser imaginadas.

Um espaço matemático de muitas dimensões é uma ocorrência muito comum para os matemáticos puros. Nenhuma reivindicação é feita sobre a sua real existência, embora se possa assumir que existe num mundo platónico ideal. No grande problema de classificação de grupos, por exemplo (ver página 155), o «grupo monstro» é uma forma de medir a simetria num espaço de 196 883 dimensões. Não podemos «ver» esse espaço da mesma forma que podemos ver o espaço a três dimensões habitual, mas, apesar disso, ele pode ser imaginado e tratado de uma forma precisa pela álgebra moderna.

A preocupação dos matemáticos pela dimensão é completamente independente do significado que os físicos dão à análise dimensional. As unidades usuais da física são medidas em termos de massa M, comprimento C e tempo T. Assim, fazendo a sua análise dimensional, um físico pode verificar que as equações fazem sentido desde que ambos os lados de uma equação tenham as mesmas dimensões.

Não faz sentido ter força = velocidade. Uma análise dimensional dá velocidade como metros por segundo, logo tem dimensão de comprimento dividido pelo tempo ou C/T, o que se escreve CT^{-1}. A força é a massa vezes a aceleração, e como a aceleração é metros por segundo por segundo, o resultado final é que a força terá dimensões MCT^{-2}.

Pessoas em coordenadas

Os próprios seres humanos são coisas com muitas dimensões. Um ser humano tem muito mais do que três coordenadas. Poderíamos usar (*a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*) para idade, peso, altura, sexo, número do sapato, cor dos olhos, cor do cabelo, nacionalidade, etc. Em vez de pontos geométricos, podemos ter pessoas. Se nos limitarmos a este espaço octodimensional, o senhor Rasmussen pode ter as coordenadas (43 anos, 83 kg, 165 cm, masculino, 42, azul, louro, dinamarquês) e as coordenadas da senhora Smith podem ser (26 anos, 56 kg, 157 cm, feminino, 37, castanho, morena, britânica).

Topologia A teoria das dimensões faz parte da topologia geral. Outros conceitos de dimensão podem ser definidos independentemente em termos de espaços matemáticos. Uma tarefa crucial é mostrar como se relacionam umas com as outras. Figuras de destaque em muitos ramos da matemática têm conduzido investigações no domínio das dimensões, incluindo Henri Lebesgue, L. E. J. Brouwer, Karl Menger, Paul Urysohn e Leopold Vietoris (até recentemente a pessoa mais velha na Áustria, morreu em 2002 com 110 anos).

O principal livro sobre o assunto foi *Teoria da Dimensão*. Publicado em 1948 por Witold Hurewicz e Henry Wallman, ainda é considerado um marco na nossa compreensão do conceito de dimensão.

A dimensão em todas as suas formas Desde as três dimensões introduzidas pelos gregos, o conceito de dimensão tem sido criticamente analisado e ampliado.

As *n* dimensões do espaço matemático foram introduzidas sem esforço, enquanto os físicos têm baseado as suas teorias no espaço-tempo (de 4 dimensões) e versões recentes da teoria das cordas (ver página 97) exigem 10, 11 e 26 dimensões. Tem havido incursões em dimensões fraccionárias e sido estudadas formas fractais (ver página 100) com diversas medidas. Hilbert introduziu um espaço matemático infinito-dimensional que é agora o enquadramento básico dos matemáticos puros. A dimensão é muito mais do que o um-dois-três da geometria euclidiana.

a ideia resumida
Para lá da terceira dimensão

25 Fractais

Em Março de 1980, o mais avançado computador de grande porte no centro de pesquisas da IBM em Yorktown Heights, no estado de Nova Iorque, enviava as suas instruções para uma antiga impressora Tektronix. Esta imprimiu obedientemente pontos em lugares curiosos de uma página branca, e quando acabou de fazer barulho, o resultado parecia uma mão-cheia de pó espalhado por toda a folha. Benoît Mandelbrot esfregou os olhos incrédulo. Percebeu que era importante, mas o que seria? A imagem que apareceu lentamente à sua frente foi como uma fotografia a preto e branco emergindo de um tanque de revelação de fotografia. Foi um primeiro vislumbre do ícone do mundo dos fractais – o conjunto de Mandelbrot.

Isto foi matemática experimental por excelência, uma introdução ao campo em que os matemáticos ocupam bancadas de laboratório, tal como os físicos e os químicos. Também eles podiam agora fazer experiências. Abriram-se, literalmente, novas perspectivas. Foi uma libertação dos climas áridos da «definição, teorema, prova», embora houvesse um regresso ulterior ao rigor do argumento racional.

A desvantagem desta abordagem experimental era que as imagens visuais precediam uma sustentação teórica. Os experimentalistas navegavam sem mapa. Embora Mandelbrot tivesse criado a palavra «fractais», o que eram eles? Haveria uma definição precisa na forma usual da matemática? No início, Mandelbrot não quis dá-la. Não quis destruir a magia da experiência desenvolvendo uma definição forte que pudesse ser inadequada e limitadora. Sentiu a noção de fractal, «como um bom vinho que exige um pouco de envelhecimento antes de ser engarrafado».

Cronologia

1879
Cayley trabalha num precursor de fractais modernos

1904
Von Koch cria a sua curva de floco de neve

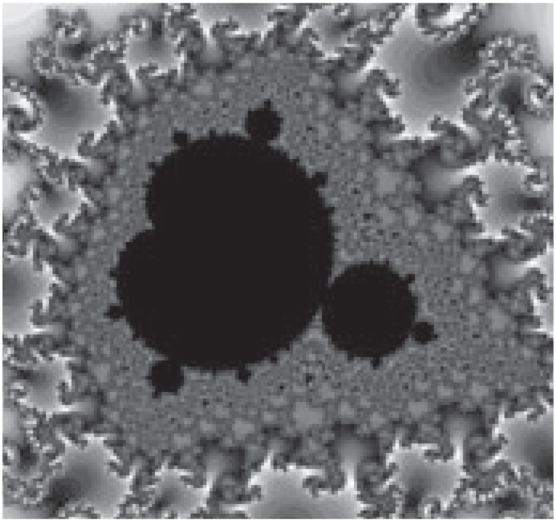
O conjunto de Mandelbrot Mandelbrot e os seus colegas não eram matemáticos particularmente abstrusos. Brincavam com a mais simples das fórmulas. Toda a ideia se baseava na iteração – a prática de aplicar uma fórmula uma e outra vez. A fórmula que gerou o conjunto de Mandelbrot era simplesmente $x^2 + c$.

A primeira coisa que fazemos é escolher um valor para c. Escolhamos $c = 0,5$. Começando com $x = 0$, substituímos na fórmula $x^2 + 0,5$. O primeiro cálculo dá 0,5 outra vez. Usemos agora este como x, substituindo-o $x^2 + 0,5$ para obter um segundo cálculo: $(0,5)^2 + 0,5 = 0,75$. Continuamos e, na terceira etapa, este será $(0,75)^2 + 0,5 = 1,0625$. Todos estes cálculos podem ser feitos com uma calculadora simples. Se continuarmos, verificamos que a resposta vai sendo cada vez maior.

Tentemos outro valor para c, desta vez $c = -0,5$. Como anteriormente, começamos com $x = 0$ e substituímos em $x^2 - 0,5$ para obtermos –0,5. Continuando, obtemos –0,25, mas desta vez os valores não se tornam cada vez maiores; antes, depois de algumas oscilações, estacionam num valor próximo de –0,3660...

Então, escolhendo $c = 0,5$, a sequência que começa em $x = 0$ aumenta para infinito, mas escolhendo $c = -0,5$, verificamos que a sequência que se inicia em $x = 0$ na realidade converge para um valor próximo de –0,3660. O conjunto de Mandelbrot consiste em todos estes valores de c para os quais a sequência que começa em $x = 0$ *não* foge para infinito.

O conjunto de Mandelbrot



A história não acaba aqui, porque até agora considerámos apenas os números reais unidimensionais – dando um conjunto de Mandelbrot unidimensional, donde não termos conseguido ver muito. O que precisa de ser considerado é a mesma fórmula $z^2 + c$, mas com z e c como números complexos bidimensionais. (ver página 32). Isto dar-nos-á um conjunto de Mandelbrot bidimensional.

Para alguns valores de c no conjunto de Mandelbrot, a sequência de zs pode fazer todo o tipo de coisas estranhas, como dançar entre um número de pontos, mas estes não escaparão para o infinito. No conjunto de Mandelbrot vemos outra propriedade crucial nos fractais,

1918
Hausdorff introduz o seu conceito de dimensão fraccionária

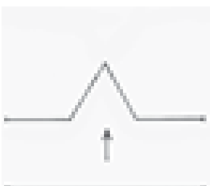
1919
Julia e Fatou investigam estruturas fractais no plano complexo

1975
Mandelbrot introduz o termo fractal

que é a auto-semelhança. Se ampliarmos um conjunto, não saberemos ao certo o nível de amplificação porque veremos apenas mais conjuntos de Mandelbrot.

Antes de Mandelbrot Como a maioria das coisas em matemática, as descobertas raramente são completamente novas. Investigando a História, Mandelbrot descobriu que matemáticos como Henri Poincaré e Arthur Cayley tinham tido alguns vislumbres da ideia cem anos antes de si. Infelizmente, não tinham usufruído do poder dos computadores para investigar mais o assunto.

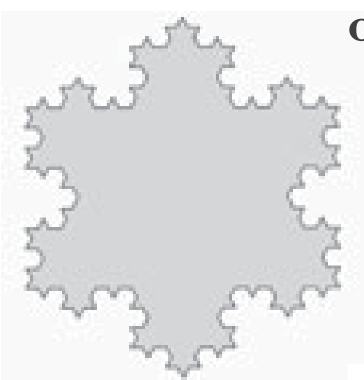
As formas descobertas pela primeira vaga de teóricos dos fractais incluíam curvas enrugadas e «curvas monstros» que tinham sido previamente recusadas como exemplos patológicos de curvas. Como eram demasiado patológicas, ficaram guardadas no armário dos matemáticos e pouca atenção lhes foi dada. O que se pretendia então eram curvas «suaves» normais que pudessem ser tratadas com o cálculo diferencial. Com a popularidade dos fractais, outros matemáticos cujo trabalho foi ressuscitado foram Gaston Julia e Pierre Fatou, que trabalharam com estruturas semelhantes a fractais no plano complexo a seguir à Primeira Guerra Mundial. As suas curvas não se chamavam, como é óbvio, fractais, nem Julia e Fatou tinham o equipamento tecnológico para verem as suas formas.



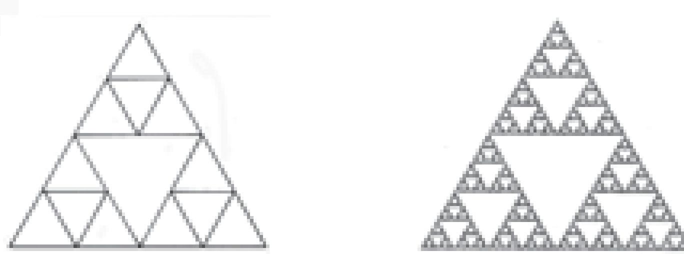
O elemento gerador do floco de neve de Koch

Outros fractais famosos A famosa curva de Koch deve o nome ao matemático sueco Niels Fabien Helge von Koch. A curva do floco de neve é praticamente a primeira curva fractal. É gerada a partir do lado de um triângulo tratado como um elemento, dividindo-o em três partes, cada uma com o comprimento ⅓ e juntando um triângulo na posição central.

A propriedade curiosa da curva de Koch é ela ter uma área finita, porque se mantém sempre dentro de um círculo, mas a cada passo na sua geração o comprimento aumenta. É uma curva que cerca uma área finita mas tem uma circunferência «infinita»!



O floco de neve de Koch



Triângulo de Sierpiński

Outro fractal famoso recebeu o nome do matemático polaco Waclaw Sierpiński. Este fractal é determinado retirando triângulos de um triângulo equilátero; e, continuando o processo, obtém-se um triângulo de Sierpiński (gerado por um processo diferente na página 54).

Dimensão fraccionária A forma como Felix Hausdorff pensava as dimensões era inovadora. Tinha tudo a ver com escala. Se uma linha for ampliada pelo factor 3, fica 3 vezes mais comprida que anteriormente. Se um quadrado for ampliado pelo factor 3, a sua *área* é 9 vezes o valor anterior, ou 3^2, e assim a dimensão é dois. Se um cubo for ampliado por este factor, o seu volume é 27 ou 3^3 vezes o seu valor anterior, logo a sua dimensão é 3. Estes valores da dimensão de Hausdorff coincidem todos com as nossas expectativas para uma linha, quadrado ou cubo.

Se for ampliada por 3, a unidade básica da curva de Koch fica 4 vezes mais comprida do que anteriormente. Seguindo o mesmo esquema, a dimensão de Hausdorff é o valor de D para o qual $4 = 3^D$. Um cálculo alternativo é

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

o que significa que D para a curva de Koch é aproximadamente 1,262. Com fractais, é frequente que a dimensão de Hausdorff seja maior do que a dimensão ordinária, que é 1 no caso da curva de Koch.

A dimensão de Hausdorff influenciou a definição de fractal de Mandelbrot – um conjunto de pontos cujo valor de D não é um número inteiro. A dimensão fraccionária tornou-se a propriedade crucial dos fractais.

As aplicações dos fractais O potencial de aplicação dos fractais é grande. Os fractais podem muito bem ser o meio matemático que modela objectos naturais, como o crescimento das plantas ou a formação das nuvens.

Os fractais já foram aplicados ao crescimento de organismos marinhos como corais ou esponjas. Já se provou que o crescimento das cidades modernas tem semelhanças com o crescimento fractal. Na medicina foi encontrada uma aplicação dos fractais na modelação da actividade cerebral. E a natureza fractal dos movimentos das acções e obrigações e dos mercados de câmbio também foi investigada. O trabalho de Mandelbrot abriu uma nova janela e ainda há muito para ser descoberto.

a ideia resumida
Formas com dimensões fraccionárias

26 Caos

Como é possível existir uma teoria do caos? O caos não acontece sem teoria? A história recua a 1812. Enquanto Napoleão avançava sobre Moscovo, o seu compatriota marquês Pierre-Simon de Laplace publicava um ensaio sobre o universo determinista: se, num dado instante, as posições e velocidades de todos os objectos no universo, bem como as forças que actuam sobre eles, fossem conhecidas, essas quantidades podiam ser calculadas com exactidão para todo o futuro. O universo e todos os objectos nele seriam completamente determinados. A teoria do Caos mostra-nos que o mundo é mais complicado do que isto.

No mundo real não podemos conhecer *exactamente* todas as posições, velocidades e forças, mas o corolário da convicção de Laplace era que, se conhecêssemos valores aproximados num dado instante, o universo não seria em todo o caso muito diferente. Era razoável afirmá-lo, pois quaisquer velocistas que começassem um décimo de segundo depois do tiro de partida chegariam à meta apenas um décimo de segundo depois do seu tempo habitual. A convicção era de que pequenas discrepâncias nas condições iniciais significariam pequenas discrepâncias nos resultados. A teoria do caos destruiu esta ideia.

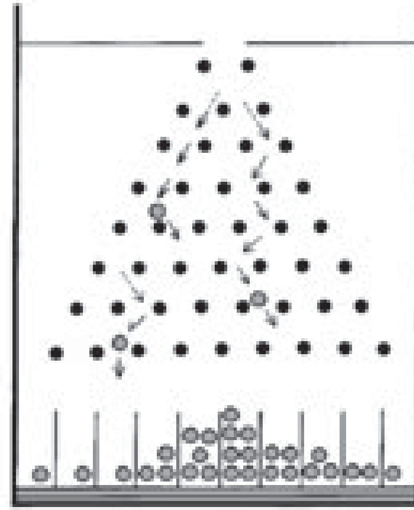
O efeito de borboleta O efeito de borboleta mostra como condições iniciais ligeiramente diferentes das dadas podem produzir na realidade um resultado muito diferente do previsto. Se existir previsão de bom tempo para um dia na Europa, mas uma borboleta bater as asas na América do Sul, o seu movimento pode pressagiar temporais no outro lado do mundo – porque o bater de asas muda muito lentamente a pressão, causando um padrão de tempo completamente diferente da previsão original.

Cronologia

1812
Laplace publica o seu ensaio sobre um mundo determinista

1889
Poincaré encontra o caos no seu trabalho sobre o problema dos três corpos, pelo qual recebeu um prémio do rei Óscar da Suécia

Podemos ilustrar a ideia com uma experiência mecânica simples. Se deixarmos cair um rolamento de esferas através de uma abertura no topo de um painel com pinos, ele progredirá para baixo, sendo deflectido de uma forma ou de outra consoante os diferentes pinos que encontrar pelo caminho até alcançar finalmente as ranhuras no fundo. Podemos ser tentados a deixar cair um rolamento idêntico, na mesma posição, com exactamente a mesma velocidade. Se conseguirmos fazê-lo *exactamente*, o marquês de Laplace estaria correcto e o caminho seguido pelo rolamento seria exactamente o mesmo. Se o primeiro caísse na terceira ranhura, o segundo também cairia nela.

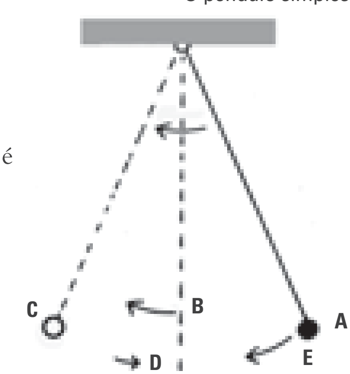


Experiência do painel com pinos

É, contudo, óbvio que não conseguimos deixar cair o rolamento exactamente da mesma posição com exactamente a mesma velocidade e força. Na realidade, haverá uma diferença muito pequena que provavelmente nem seremos capazes de medir. O resultado é que o rolamento pode tomar um caminho diferente e provavelmente terminará numa ranhura diferente.

Um pêndulo simples O pêndulo livre é dos sistemas mecânicos mais simples de analisar. Enquanto balança para a frente e para trás, o pêndulo perde gradualmente energia. O desvio da vertical e a velocidade (angular) da esfera do pêndulo decrescem até o pêndulo ficar estacionário.

O pêndulo simples



O movimento da esfera do pêndulo pode ser traçado num diagrama de fases. No eixo horizontal é medido o desvio (angular) e no eixo vertical é medida a velocidade. O ponto de lançamento é marcado no ponto *A* no lado positivo do eixo horizontal. No ponto *A*, o desvio está no máximo e a velocidade é zero. Conforme a esfera do pêndulo se desloca até ao eixo vertical (onde o desvio é zero), a velocidade está no máximo, marcado no diagrama como *B*. Em *C*, quando a esfera do pêndulo está na outra extremidade da sua oscilação, o desvio é negativo e a velocidade é zero. O prumo baloiça então para trás através de *D* (onde se move na direcção oposta e portanto com velocidade negativa) e completa uma

1961
Lorenz observa o efeito de borboleta

1971
Robert May investiga o caos no modelo populacional

2004
A teoria do caos chega à cultura popular no filme *O Efeito-Borboleta*

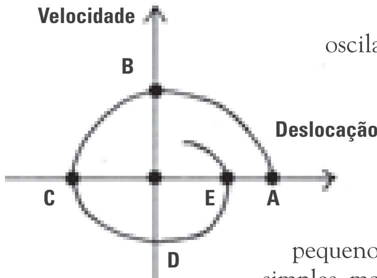
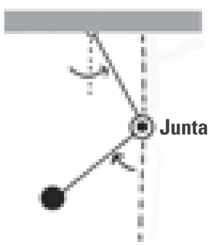


Diagrama de fases do pêndulo simples

oscilação em *E*. No diagrama de fases, isto é representado por uma rotação de 360 graus, mas, como o balanço é reduzido, o ponto *E* é mostrado dentro de *A*. Conforme baloiça menos, o pêndulo descreve uma espiral para a origem até finalmente parar.

Esse não é o caso do pêndulo duplo, em que as esferas dos pêndulos estão no final de duas varas articuladas. Se o desvio for pequeno, o movimento do pêndulo duplo é semelhante ao do pêndulo simples, mas se o desvio for maior, a massa baloiça, roda e muda de direcção e o desvio sobre a junta intermédia é aparentemente aleatório. Se o movimento não for forçado, o pêndulo acabará por parar, mas a curva que o descreve estará longe da bem comportada espiral do pêndulo simples.

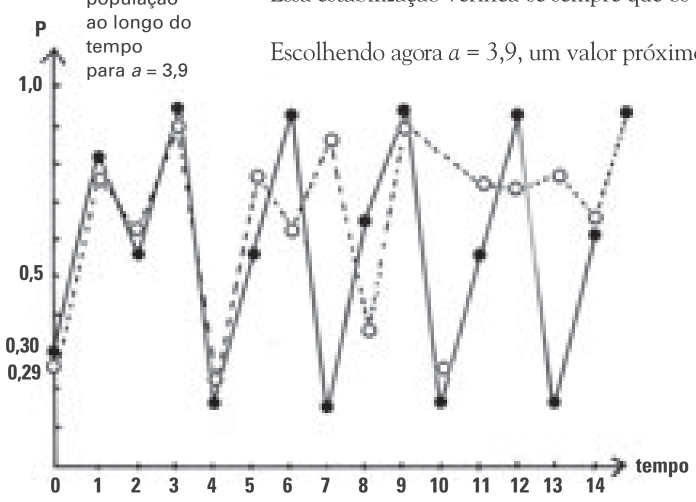


Movimento do duplo pêndulo

Movimento caótico A característica do caos é que um sistema determinista pode parecer gerar um comportamento aleatório. Observemos outro exemplo, a repetitiva, ou iterativa, fórmula $a \times p \times (1 - p)$ em que p significa a população, medida como uma proporção na escala de 0 a 1. O valor de a deve estar entre 0 e 4 para garantir que o valor de p fica entre 0 e 1.

Modelemos a população quando $a = 2$. Se tomarmos um valor inicial de, digamos $p = 0{,}3$ no instante *tempo* = 0, para conhecer a população no *tempo* = 1, substituímos $p = 0{,}3$ em $a \times p \times (1 - p)$ para dar 0,42. Usando uma calculadora simples, podemos repetir a operação, desta vez com $p = 0{,}42$, para nos dar o número seguinte (0,4872). Avançando desta forma, determinamos a população em momentos posteriores. Neste caso, a população estaciona rapidamente em $p = 0{,}5$. Essa estabilização verifica-se sempre que os valores de a são menores do que 3.

Alteração da população ao longo do tempo para $a = 3{,}9$



Escolhendo agora $a = 3{,}9$, um valor próximo do máximo permitido, e usando a mesma população inicial $p = 0{,}3$, a população não estabiliza mas oscila descontroladamente. Isto deve-se ao facto de o valor de a estar na «região caótica», ou seja, a é um número maior que 3,57. Mais ainda, se escolhermos um valor inicial diferente para a população, $p = 2{,}9$, um valor próximo de 0,3, o crescimento fica sobre o padrão anterior de crescimento dos primeiros passos, mas depois começa a divergir completamente. Este foi o comportamento observado por Lorenz em 1961.

Prever o tempo Mesmo com computadores potentes, sabemos que não é possível prever o tempo senão com uns dias de antecedência. Mesmo assim, o tempo prega-nos partidas desagradáveis, porque as equações que o determinam não são lineares – envolvem as variáveis multiplicadas em conjunto e não apenas as variáveis individualmente consideradas.

A teoria por trás da matemática da previsão do tempo foi trabalhada independentemente pelo engenheiro francês Claude Navier em 1821 e pelo físico--matemático George Gabriel Stokes em 1845. As equações resultantes de Navier-Stokes são de um enorme interesse para os cientistas. O Instituto de Matemática Clay em Cambridge, no Massachusetts, oferece um prémio de um milhão de dólares a quem fizer progressos em direcção a uma teoria matemática que desvende os segredos dessas equações. Aplicadas ao problema do fluxo de fluidos, muito é conhecido sobre os movimentos estáveis da atmosfera superior. Mas o fluxo de ar junto da superfície cria turbulência que resulta em caos, com um subsequente comportamento largamente desconhecido.

Da meteorologia à matemática

A descoberta do efeito de borboleta aconteceu, por acaso, por volta de 1961. Quando o meteorologista do MIT Edward Lorenz foi tomar um café, deixou o seu velho computador a imprimir e, quando regressou, encontrou algo inesperado. Tinha estado a tentar recapturar alguns gráficos interessantes sobre o tempo e encontrou um novo gráfico irreconhecível, o que era estranho, uma vez que tinha usado os mesmos valores iniciais e portanto devia ter sido desenhado o mesmo gráfico. Estaria na altura de trocar de computador?

Depois de alguma reflexão, Lorenz encontrou uma diferença na forma como fornecera os dados iniciais: antes tinha usado seis casas decimais mas, na repetição, só inseriu três. Para explicar a disparidade, usou a expressão «efeito de borboleta». Depois da descoberta, os seus interesses intelectuais migraram para a matemática.

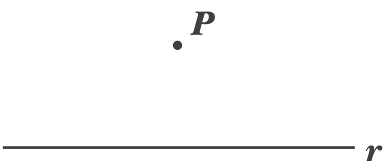
Embora muito se saiba sobre a teoria dos sistemas lineares de equações, as equações de Navier-Stokes contêm termos não lineares que as tornam intratáveis. Na prática, a única maneira de as resolver é a numérica, usando computadores potentes.

Atractores estranhos Pode pensar-se que os sistemas dinâmicos possuem «atractores» nos seus diagramas de fases. No caso do pêndulo simples, o atractor é o único ponto na origem relativamente ao qual o movimento é dirigido. Com o pêndulo duplo é mais complicado, mas mesmo aqui o retrato da fase mostrará alguma regularidade e o pêndulo será atraído para um conjunto de pontos no diagrama de fases. Para sistemas como este, o conjunto de pontos pode ser um fractal (ver página 100) chamado atractor «estranho» que terá uma estrutura matemática definida. Assim, nem tudo está perdido. Na nova teoria do caos, não é assim tão «caótico» um caos que resulta num caos «regular».

a ideia resumida
A incontrolabilidade da regularidade

27 O postulado das paralelas

Esta história dramática começa com um cenário geométrico muito simples. Imaginemos uma recta *r* e um ponto *P* que não está na recta. Quantas rectas podemos fazer passar por *P* paralelas à recta *r*? Parece óbvio que há exactamente uma recta que passa por *P* e que nunca encontra *r*, independentemente da sua extensão em qualquer direcção.



Isto parece evidente e em perfeita concordância com o senso comum. Euclides de Alexandria incluiu uma variante deste cenário como um dos seus postulados, na base da geometria que são os *Elementos*.

O senso comum nem sempre é um guia fidedigno. Vejamos se a assunção de Euclides tem sentido matemático.

Os *Elementos* de Euclides A geometria de Euclides está estabelecida nos 13 livros dos *Elementos*, escritos por volta de 300 a.C. A obra é um dos mais influentes textos de matemática alguma vez escritos e os matemáticos gregos referem-na constantemente como a primeira codificação matemática da geometria. Mais tarde, os académicos estudaram e traduziram o texto a partir de manuscritos existentes, e a obra foi transmitida e elogiada universalmente como o próprio modelo daquilo que a geometria deve ser.

Os *Elementos* infiltraram-se na escola e as leituras do «livro sagrado» tornaram-se a forma de ensinar a geometria. Contudo, a obra demonstrou-se inadequada para os alunos mais novos. Como o poeta A. C. Hilton satirizou, «embora o

Cronologia

cerca de 300 a.C.
Euclides inclui o postulado das paralelas nos seus *Elementos*

1829-31
Lobachevsky e Bolyai publicam o seu trabalho em geometria hiperbólica

escrevessem por hábito, não o escreviam bem». Pode dizer-se que Euclides era para homens, não para meninos. Nas escolas inglesas, a sua obra atingiu o zénite da sua influência durante o século XIX, mas mantém-se uma pedra de toque para os matemáticos de hoje.

É o estilo dos *Elementos* de Euclides que os torna tão dignos de nota – o seu feito é apresentar a geometria como uma sequência de proposições provadas. Sherlock Holmes teria admirado o seu sistema dedutivo que avança logicamente de postulados claramente enunciados e teria castigado o Dr. Watson por não os ver como «um sistema frio e sem emoção».

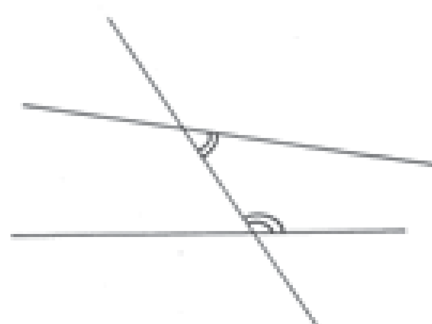
Embora o edifício da geometria de Euclides assentasse nos postulados (hoje chamados axiomas; ver caixa), estes não eram suficientes. Euclides juntou-lhes «definições» e «noções comuns». As definições incluem declarações como «um *ponto* é o que não tem parte» e «uma *recta* é um comprimento sem largura». As noções comuns incluem afirmações como «o todo é maior do que as partes» e «coisas que são iguais à mesma coisa também são iguais entre si». Só no final do século XIX se reconheceu que Euclides tinha feito assunções tácitas.

Os postulados de Euclides

Uma das características da matemática é que algumas assunções podem gerar extensas teorias. Os postulados de Euclides são um excelente exemplo e determinaram o modelo dos sistemas axiomáticos posteriores. Os seus cinco postulados são:

1. Um segmento de recta pode ser traçado unindo quaisquer dois pontos.
2. Um segmento de recta pode ser prolongado indefinidamente para construir uma recta.
3. Um círculo pode ser construído com qualquer centro e qualquer raio.
4. Todos os ângulos rectos são iguais entre si.
5. Se uma linha recta cortar outras duas de modo que a soma dos dois ângulos internos do mesmo lado seja menor do que dois ângulos rectos, essas duas rectas, se indefinidamente prolongadas, encontram-se no mesmo lado em que os ângulos são inferiores a dois ângulos rectos.

O quinto postulado Foi o quinto postulado que causou controvérsia mais de 2000 anos depois de os *Elementos* aparecerem pela primeira vez. O próprio estilo parece estranho, pela verbosidade e falta de jeito. O próprio Euclides não estava satisfeito com o quinto postulado, mas necessitava dele para provar proposições e teve de o incluir. Tentou prová-lo a partir de outros postulados, mas falhou.



Matemáticos posteriores tentaram prová-lo ou substituí-lo por um mais simples. Em 1795, John Playfair enunciou-o de uma

1854
Riemann dá conferências sobre os fundamentos da geometria

1872
Klein unifica a geometria através da teoria de grupos

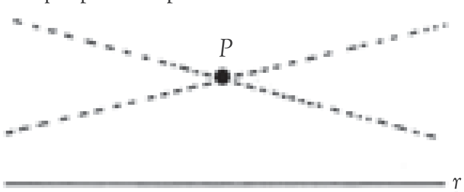
1915
A teoria geral da relatividade de Einstein é baseada na geometria de Riemann

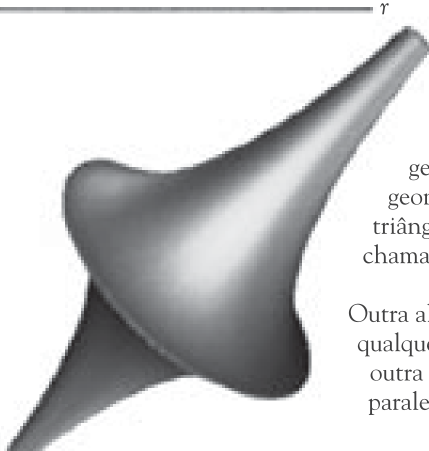
forma que ganhou popularidade: para uma recta r e um ponto P que não esteja na recta r, há uma *única* recta que passa por P que é paralela a r. Pela mesma altura, Adrien Marie Legendre substituiu outra versão equivalente, quando afirmou a existência de um triângulo cujos ângulos somavam 180 graus. Estas novas formas do quinto postulado iam de alguma forma ao encontro da objecção de artificialidade. Eram mais aceitáveis do que a pesada versão dada por Euclides.

Outra linha de ataque foi procurar a elusiva prova do quinto postulado, busca que exerceu uma poderosa atracção entre os matemáticos. Se a prova fosse encontrada, o postulado tornar-se-ia um teorema e sairia da linha de fogo. As tentativas malogradas de a encontrar acabaram por se tornar exemplos excelentes de raciocínio circular, argumentos que presumem exactamente o mesmo do que tentam provar.

Geometria não euclidiana Os trabalhos de Carl Friedrich Gauss, János Bolyai e Nikolai Ivanovich Lobachevsky ofereceram um importante avanço. Gauss não publicou o seu trabalho, mas parece claro que chegou às suas conclusões em 1817. Bolyai publicou a sua obra em 1831 e Lobachevsky, independentemente, em 1829, o que causou uma disputa de prioridade entre os dois. Não há dúvida do brilhantismo destes homens. Todos mostraram que o quinto postulado era independente dos outros quatro postulados. Ao juntarem esta negação aos outros quatro postulados, mostraram que era possível um sistema consistente.

Bolyai e Lobachevsky construíram uma nova geometria permitindo que existisse mais de uma recta que passasse por P e não encontrasse a recta r. Como é isto possível? Certamente que as linhas ponteadas encontram r. Se o aceitarmos, estamos a cair inconscientemente na visão de Euclides. O diagrama é portanto um artifício de confiança, pois o que Bolyai e Lobachevsky propunham era um novo tipo de geometria que não é conforme com o senso comum de Euclides. De facto, a sua geometria não euclidiana pode ser considerada como a geometria da superfície curva daquilo que é conhecido como pseudoesfera.





O caminho mais curto entre dois pontos numa pseudoesfera tem o mesmo papel que a linha recta na geometria euclidiana. Uma das curiosidades desta geometria não euclidiana é que a soma dos ângulos de um triângulo é *menor* do que 180 graus. Esta geometria é chamada geometria hiperbólica.

Outra alternativa ao quinto postulado de Euclides diz que qualquer recta que passe por P encontra a recta r. Posto de outra forma, não existem rectas que passem por P e sejam paralelas a r. Esta geometria é diferente da de Bolyai e Lobachevsky, mas é apesar de tudo uma geometria genuína. Um modelo dela é o da geometria na superfície de uma esfera. Aí, os grandes círculos (os que têm a mesma circunferência da própria esfera) tomam o papel das rectas da geometria euclidiana. Nesta geometria não euclidiana, a soma dos ângulos de um triângulo é maior do que 180 graus. É a chamada geometria elíptica e está associada ao matemático alemão Bernhard Riemann, que a investigou na década de 1850.

A geometria de Euclides, que se pensava ser a única verdadeira geometria – de acordo com Immanuel Kant, a geometria «instalada nos circuitos humanos» –, tem sido afastada do seu pedestal. A geometria euclidiana é agora um de muitos sistemas, entalada entre a geometria hiperbólica e a geometria elíptica. As diferentes versões foram unificadas sob a mesma cúpula por Felix Klein em 1872. O advento das geometrias não euclidianas foi um acontecimento sísmico na matemática e abriu caminho para a geometria da relatividade geral de Einstein (ver página 192). É a teoria *geral* da relatividade que exige um novo tipo de geometria – a geometria do espaço-tempo curvo, ou geometria riemanniana. Foi esta geometria não euclidiana, e não a força de atracção gravítica entre objectos de Newton, que passou a explicar porque é que as coisas caem. A presença de enormes objectos no espaço, como a Terra e o Sol, faz que o espaço-tempo seja encurvado. Um berlinde numa folha de papel provocará uma pequena mossa, mas tente-se colocar uma bola de *bowling* em cima dela e a distorção será enorme.

Essa curvatura medida pela geometria de Riemann prediz como é que os raios de luz se dobram na presença de grandes objectos espaciais. A geometria euclidiana ordinária, com o tempo como componente independente, não chegará para a relatividade geral. Uma razão é o espaço euclidiano ser plano – não há curvatura. Imaginemos uma folha de papel em cima de uma mesa: podemos dizer que em qualquer ponto do papel a curvatura é zero. Subjacente ao espaço-tempo de Riemann está um conceito de curvatura que varia continuamente – tal como a curvatura de uma peça de roupa amarrotada varia de ponto para ponto. É como ver um objecto num recinto de espelhos curvos – a imagem depende de onde se olha para o espelho.

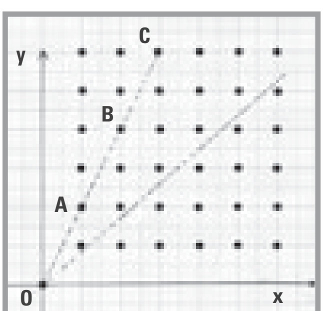
Não admira que Gauss tenha ficado tão impressionado pelo jovem Riemann na década de 1850 e tenha mesmo sugerido que a «metafísica» do espaço seria revolucionada pela sua visão.

a ideia resumida
E se as linhas paralelas se encontrassem?

28 Geometria discreta

A geometria é um dos ofícios mais antigos do mundo – significa literalmente terra (*geo*) e medida (*metria*). Na geometria ordinária, há linhas contínuas e formas sólidas para investigar, todas as quais podemos considerar compostas de pontos «próximos» uns dos outros. A matemática discreta trata números inteiros em oposição aos números reais contínuos. A geometria discreta pode envolver um número finito de pontos e linhas ou reticulados de pontos – o contínuo é substituído pelo isolado.

Uma grade ou grelha é tipicamente um conjunto de pontos cujas coordenadas são números inteiros. Esta geometria coloca problemas interessantes e tem aplicações em áreas tão díspares como a teoria da codificação e a concepção de experiências científicas.



A grelha de pontos dos eixos *x* e *y*

Observemos um farol que projecta um feixe de luz. Suponhamos que os raios de luz começam na origem O e fazem um varrimento entre a horizontal e a vertical. Podemos perguntar que raios atingem que pontos da grelha (que podem ser barcos atracados no porto numa disposição muito uniforme).

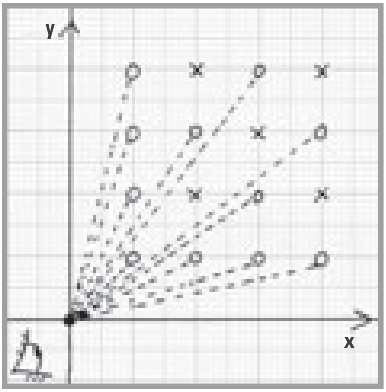
A equação do raio pela origem é $y = mx$. Esta é a equação de uma recta que passa pela origem e tem declive m. Se o raio é $y = 2x$, atingirá o ponto de coordenadas $x = 1$ e $y = 2$, porque estes valores satisfazem a equação. Se o raio atinge a grelha no ponto com $x = a$ e $y = b$, o declive m é a fracção b/a. Consequentemente, se m não é uma fracção genuína (pode ser $\sqrt{2}$, por exemplo), o raio de luz falhará todos os pontos da grelha.

Cronologia

1639
Pascal, com apenas 16 anos, descobre o seu teorema

1806
Brianchon descobre o teorema dual do teorema de Pascal

O raio de luz $y = 2x$ atinge o ponto A com coordenadas $x = 1$ e $y = 2$, mas não atinge o ponto B com coordenadas $x = 2$ e $y = 4$, e todos os outros pontos «atrás» de A (tais como C, com coordenadas $x = 3$ e $y = 6$, e D com $x = 4$ e $y = 8$) estarão obscurecidos. Podemos imaginar-nos na origem, identificando os pontos que conseguimos ver de lá e aqueles que ficarão obscurecidos.



Os pontos ○ «visíveis» da origem, e os pontos obscuros ×

Podemos mostrar que esses pontos com coordenadas $x = a$ e $y = b$ que podem ser vistos são aqueles que são primos entre si. São pontos com coordenadas, tais como $x = 2$ e $y = 3$, em que nenhum número além do 1 divide nem x nem y. Os números para trás deste serão múltiplos, tais como $x = 4$ e $y = 6$, ou $x = 6$ e $y = 9$, etc.

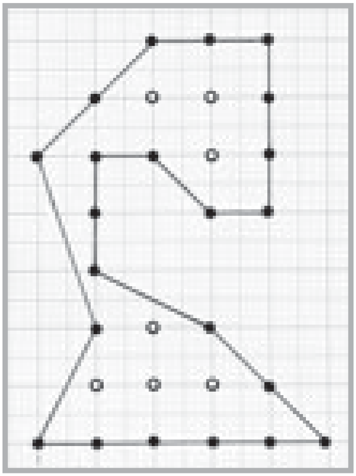
Teorema de Pick O matemático austríaco Georg Pick é famoso por duas coisas. Uma é ter sido amigo próximo de Albert Einstein e ter contribuído para levar o jovem cientista para a Universidade Alemã em Praga em 1911. A outra é ter escrito um curto artigo, publicado em 1899, sobre geometria «reticular». Entre todo o trabalho de uma longa vida, cobrindo um vasto leque de tópicos, é recordado pelo apaixonante teorema de Pick – e que teorema!

O teorema de Pick dá-nos os meios para calcular a área de uma forma com muitos lados (poligonal) constituída pela ligação de pontos cujas coordenadas são números inteiros. É uma matemática de *pinball*.

Para determinar a área da forma, temos de contar os pontos ● na fronteira e o número de pontos interiores ○. No nosso exemplo, o número de pontos na fronteira é $b = 22$ e o número de pontos interiores é $c = 7$. Isto é tudo o que precisamos para usar o teorema de Pick:

$$\text{área} = b/2 + c - 1$$

Segundo esta fórmula, a área é $22/2 + 7 - 1 = 17$. A área é 17 unidades quadradas. É tão simples como isto. O teorema de

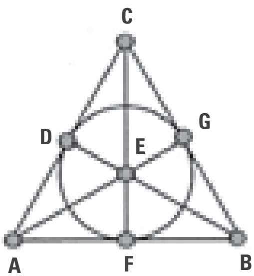


Uma forma com muitos lados ou poligonal

1846
Kirkman antecipa a descoberta dos Sistemas Triplos de Steiner

1892
Fano descobre o plano de Fano, o mais simples exemplo de geometria projectiva

1899
Pick publica o seu teorema sobre áreas de polígonos



O plano de Fano

Pick pode ser aplicado a qualquer forma que una pontos discretos com números inteiros como coordenadas, sendo a única condição que a fronteira não se cruze a si própria.

O plano de Fano A geometria do plano de Fano foi descoberta mais ou menos ao mesmo tempo que a fórmula de Pick, mas não tem nada a ver com medir o que quer que seja. Assim denominado em homenagem ao matemático italiano Gino Fano, o precursor do estudo da geometria finita, o plano de Fano é o mais simples exemplo de uma geometria «projectiva». Tem apenas sete pontos e sete linhas.

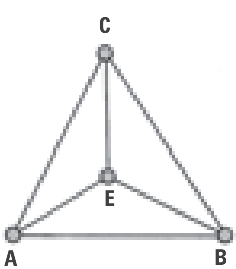
Os sete pontos são *A*, *B*, *C*, *D*, *E*, *F* e G. É fácil escolher seis das sete linhas, mas onde está a sétima? As propriedades da geometria e a forma como o diagrama é construído tornam necessário tratar a sétima linha como *DFG* – a circunferência que passa por *D*, *F* e G. Isto não é problema, visto que as linhas em geometria discreta não têm de ser «rectas» no sentido convencional.

Esta pequena geometria tem muitas propriedades, por exemplo:

- cada par de pontos determina uma linha que passa por ambos,
- cada par de linhas determina um ponto que está em ambas.

Estas duas propriedades ilustram a notável dualidade das geometrias deste tipo. A segunda propriedade é exactamente a primeira com as palavras «ponto» e «linha» trocadas, e a primeira é exactamente a segunda com as mesmas trocas.

Se, em qualquer afirmação verdadeira, trocarmos as duas palavras e fizermos alguns ajustamentos para corrigir a linguagem, obtemos outra afirmação verdadeira. A geometria projectiva é muito simétrica. A geometria euclidiana nem tanto. Na geometria euclidiana há rectas paralelas, ou seja rectas que nunca se encontram. Podemos falar do conceito de paralelismo na geometria euclidiana. Tal não é possível na geometria projectiva. Em geometria projectiva, todos os pares de linhas se encontram num ponto. Para os matemáticos, isto significa que a geometria euclidiana é um tipo inferior de geometria.



O plano de Fano tornado euclidiano

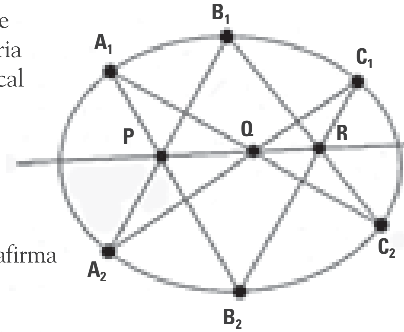
Se retirarmos uma linha *e* os seus pontos do plano de Fano, regressamos ao domínio da geometria euclidiana assimétrica e da existência de paralelas. Suponhamos que retiramos a linha «circular» *DFG* para fazer um diagrama euclidiano.

Com uma linha a menos, existem agora seis linhas: *AB*, *AC*, *AE*, *BC*, *BE* e *CE*. Temos agora pares de linhas que são «paralelas», nomeadamente *AB* e *CE*, *AC* e *BE*, e *BC* e *AE*. As linhas são paralelas neste sentido, se não têm pontos comuns – como as linhas *AB* e *CE*.

O plano de Fano ocupa uma posição icónica na matemática pelas suas relações com tantas ideias e aplicações. É uma chave para o problema das alunas da escola de Thomas Kirkman (ver página 167). Na teoria da concepção de experiências, o plano de Fano aparece como um exemplo versátil, um Sistema Triplo de Steiner (STS). Dado um número finito de *n* objectos, um STS é uma maneira de os dividir em blocos de três de forma que qualquer par retirado dos *n* objectos esteja num único bloco. Dados os sete objectos *A*, *B*, *C*, *D*, *E*, *F* e G, os blocos no STS correspondem às linhas do plano de Fano.

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

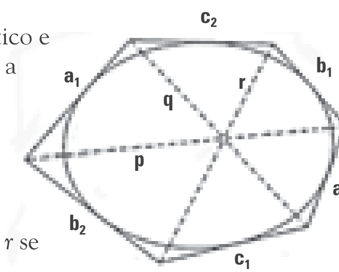
Um par de teoremas O teorema de Pascal e o teorema de Brianchon estão na fronteira entre a geometria discreta e a geometria contínua. São diferentes mas estão relacionados. O teorema de Pascal foi descoberto por Blaise Pascal em 1639, quando tinha apenas 16 anos. Pensemos num círculo estendido chamado elipse (ver página 89) e marquemos seis pontos sobre ela, a que chamaremos A_1, B_1 e C_1 e A_2, B_2 e C_2. Chamemos *P* ao ponto onde a linha A_1B_2 encontra a linha A_2B_1; *Q* ao ponto onde A_1C_2 encontra a linha A_2C_1; e *R* ao ponto onde B_1C_2 encontra a linha B_2C_1. O teorema afirma que os pontos *P*, *Q* e *R* estão todos numa única linha recta.



Teorema de Pascal

O teorema de Pascal é verdadeiro quaisquer que sejam as posições dos diferentes pontos na elipse. De facto, podíamos substituir a elipse por outra cónica, tal como a hipérbole, o círculo, a parábola ou mesmo um par de rectas (neste caso a configuração chama-se «cama de gato») e o teorema continuaria a ser verdadeiro.

O teorema de Brianchon foi descoberto muito mais tarde pelo matemático e químico francês Charles-Julien Brianchon. Desenhemos seis tangentes, a que chamaremos as linhas a_1, b_1 e c_1 e a_2, b_2 e c_2, tocando a elipse. De seguida, podemos definir três diagonais, as linhas *p*, *q* e *r*, pelos pontos de encontro das linhas, tal que: *p* é a linha entre os pontos onde a_1 encontra b_2 e a_2 encontra b_1; *q* é a linha entre os pontos onde a_1 encontra c_2 e a_2 encontra c_1; e *r* é a linha entre os pontos onde b_1 encontra c_2 e b_2 encontra c_1. O teorema de Brianchon afirma que *p*, *q* e *r* se encontram num ponto.



Teorema de Brianchon

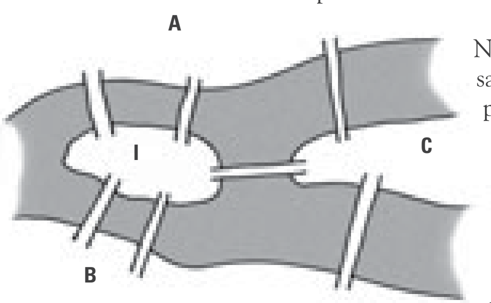
Estes dois teoremas são duais um do outro, e este é outro exemplo de teoremas de geometria projectiva que ocorrem aos pares.

a ideia resumida
Pontos individuais de interesse

29 Grafos

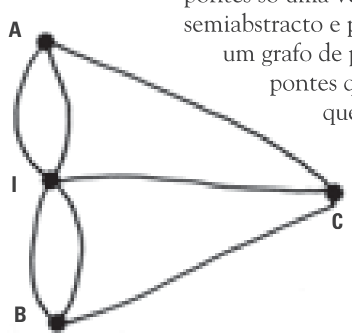
Há dois tipos de grafos em matemática. Na escola desenhamos curvas que mostram as relações entre as variáveis *x* e *y*. Noutro tipo mais recente, os pontos são juntos por linhas onduladas.

Königsberg é uma cidade na Prússia Oriental famosa pelas suas sete pontes sobre o rio Pregel. Terra natal do ilustre filósofo Immanuel Kant, a cidade e as suas pontes também estão ligadas ao famoso matemático Leonhard Euler.



No século XVIII, fez-se uma curiosa pergunta: seria possível sair e passear à volta de Königsberg atravessando cada ponte exactamente uma vez? O passeio não exige que se acabe no mesmo sítio em que se começou – apenas que se atravesse cada ponte apenas uma vez.

Em 1735, Euler apresentou a sua solução na Academia Russa, uma solução que agora é vista como o início da teoria dos grafos. No nosso diagrama semiabstracto, a ilha no meio do rio é I e as margens do rio são *A*, *B* e *C*. Podemos planear um passeio para uma tarde de domingo que atravesse as pontes só uma vez? Peguemos num lápis e tentemos. O passo crucial é sair do semiabstracto e progredir para a completa abstracção. Se o fizermos, obtemos um grafo de pontos e linhas. A terra é representada por «pontos» e as pontes que os ligam são representadas por «linhas». Não nos interessa que as linhas não sejam rectas ou que difiram em comprimento. Isso são coisas sem importância. Só as relações nos interessam.



Euler fez uma observação sobre um passeio bem-sucedido. Excluindo o início e o fim do passeio, cada vez que uma ponte é atravessada para terra, deve ser possível sair desta

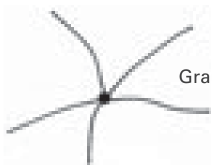
Cronologia

1735
Euler resolve o problema das pontes de Königsberg

1874
Carl Schorlemmer liga a química com «árvores»

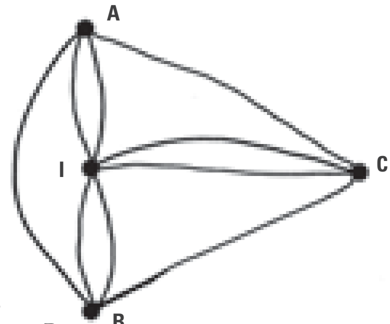
numa ponte ainda não atravessada. Traduzindo este pensamento na figura abstracta, podemos dizer que as linhas que se encontram num ponto devem ocorrer aos pares. Para lá dos dois pontos representando o início e o fim do passeio, as pontes podem ser atravessadas se, e só se, cada ponto tiver um número par de linhas incidindo sobre ele.

O número de linhas que se encontram num ponto é o chamado «grau» do ponto.



O teorema de Euler afirma que

> *As pontes de uma vila ou cidade podem ser atravessadas exactamente uma vez se, excepto no máximo dois, todos os pontos tiverem grau par.*



Olhando para o grafo que representa Königsberg, todos os pontos têm grau ímpar. Isto significa que o passeio atravessando cada ponte uma só vez não é possível em Königsberg. Se a instalação das pontes fosse alterada, o passeio tornar-se-ia possível. Se fosse construída outra ponte entre a ilha *I* e *C*, os graus de *I* e de *C* seriam ambos pares. Isto significa que poderíamos começar o passeio em *A* e terminá-lo em *B* tendo passado por cada ponte exactamente uma vez. Ainda se outra ponte fosse construída, desta vez entre *A* e *B* (mostrado à direita), podíamos começar em qualquer lado e acabar no *mesmo* sítio porque *todos* os pontos teriam grau par nesse caso.

O teorema do aperto de mão Se nos pedissem para desenhar um grafo que contivesse três pontos de grau ímpar, teríamos um problema. Experimente. É impossível, porque

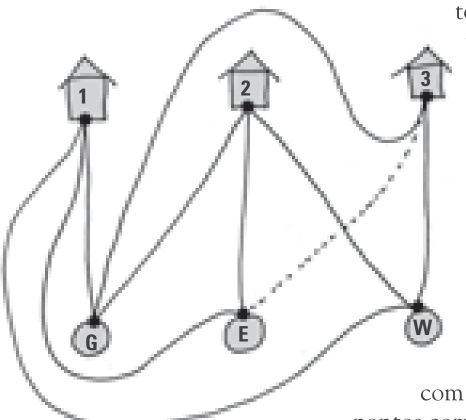
> *Em qualquer grafo, o número de pontos com grau ímpar deve ser par.*

1930
Kuratowski prova o seu teorema sobre grafos planares

1935
Georg Pólya desenvolve técnicas de contagem para grafos como álgebra

1999
Eric Rains e Neil Sloane ampliam a contagem de árvores

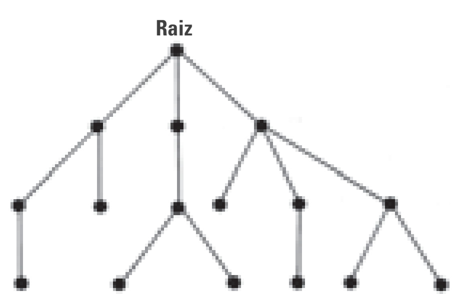
Este é o teorema do aperto de mão – o primeiro teorema da teoria dos grafos. Em qualquer grafo, todas as linhas têm um princípio e um fim, ou por outras palavras são necessárias duas pessoas para haver um aperto de mão. Se juntarmos os graus de cada ponto para todo o grafo, devemos obter um número ímpar, digamos N. A seguir, dizemos que há x pontos com grau ímpar e y pontos com grau par. Somando todos os graus dos pontos ímpares, teremos N_x e, somando todos os graus dos pontos pares, teremos N_y, que é par. Assim temos $N_x + N_y = N$ e portanto $N_x = N - N_y$. Segue-se que N_x é par. Mas o próprio x não pode ser ímpar, porque a soma de um número ímpar de graus ímpares seria um número ímpar. Logo, segue-se que x deve ser par.



Grafos não planares O problema dos serviços é um quebra-cabeças antigo. Consideremos três casas e três serviços – água, electricidade e água. Temos de ligar cada uma das casas a cada um dos serviços, mas há um problema – as ligações não podem cruzar-se.

Na verdade, isto não é possível, mas pode tentá-lo fazer com os amigos mais confiantes. O grafo descrito, ligando três pontos com outros três pontos de todas as formas possíveis (com apenas nove linhas), não pode ser desenhado no plano sem cruzamentos. Um tal grafo é chamado não planar. Este grafo de serviços, juntamente com o grafo feito com todas as linhas ligando cinco pontos, tem um lugar especial na teoria dos grafos. Em 1930, o matemático polaco Kazimierz Kuratowski provou o surpreendente teorema de que um grafo é planar se, e só se, não contém um destes dois como subgrafo, um grafo menor contido no principal.

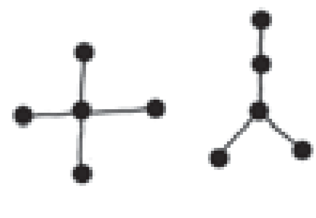
Árvores Uma «árvore» é um tipo particular de grafo, muito diferente do grafo dos serviços ou do grafo de Königsberg. No problema das pontes de



Königsberg, havia a possibilidade de começar em qualquer ponto e voltar a ele por uma via diferente. O caminho entre o ponto e o regresso a ele é chamado ciclo. Uma árvore é um grafo sem ciclos.

Um exemplo familiar de um grafo de árvore é a forma como os directórios estão dispostos num computador. Estão organizados de forma hierárquica com um directório-raiz e subdirectórios dependentes. Porque não há ciclos, não há forma de cruzar de um ramo para outro a não ser através da raiz – uma manobra familiar aos utilizadores de computadores.

Contar árvores Quantas árvores diferentes podem ser feitas com um número específico de pontos? O problema de contar árvores foi abordado por Arthur Cayley, um matemático inglês do século XIX. Por exemplo, existem exactamente três tipos diferentes de árvores com cinco pontos:



Cayley conseguiu contar o número de diferentes tipos de árvores para pequenos números de pontos. Foi tão longe quanto contar árvores com menos de 14 pontos antes de a pura complexidade dos cálculos ser demasiada para um homem sem computador. Desde então, os cálculos têm avançado até árvores com 22 pontos. Há milhões de tipos possíveis para estas.

Mesmo no seu tempo, a pesquisa de Cayley teve aplicações práticas. Contar árvores é relevante em química, campo em que a distinção de alguns compostos depende da forma como os átomos estão distribuídos nas suas moléculas. Os compostos com o mesmo número de átomos mas com diferente organização têm propriedades químicas diferentes. Usando a análise de Cayley, foi possível prever «com a ponta da caneta» a existência de químicos subsequentemente encontrados em laboratório.

a ideia resumida

Atravessar pontes e chegar às árvores

30 O problema das quatro cores

Quem poderia ter dado ao Pequeno Tim quatro lápis de cera e um mapa em branco dos condados de Inglaterra como presente de Natal? Talvez o vizinho cartógrafo que enviava ocasionalmente pequenos presentes, ou o matemático excêntrico Augustus de Morgan, que vivia por perto e passava os dias com o pai do Tim. O Mr. Scrooge não seria certamente.

Os Cratchits viviam numa moradia em banda sem graça em Bayham Street, Camden Town, mesmo a norte da recente universidade onde De Morgan dava aulas. A origem da prenda tornou-se clara no ano novo, quando o professor telefonou para saber se Tim tinha colorido o mapa.

De Morgan tinha ideias definidas sobre como a coloração devia ser feita: «Tens de colorir o mapa de forma que dois condados com uma fronteira comum tenham cores diferentes.»

«Mas eu não tenho cores suficientes», respondeu Tim sem pensar. De Morgan teria sorrido e deixado o rapaz entregue à tarefa. Contudo, muito recentemente, Frederick Guthrie, um dos seus alunos, tinha-lhe perguntado sobre o assunto, e mencionado uma bem-sucedida coloração do mapa de Inglaterra apenas com quatro cores. O problema agitou a imaginação de De Morgan.

Cronologia

1852
Guthrie, aluno de De Morgan, coloca-lhe o problema

1879
Acredita-se que Kempe tenha resolvido o problema

1890
Heawood expõe erros na prova de Kempe e prova um teorema de cinco cores

Os estados do Oeste dos EUA

Será possível colorir qualquer mapa com apenas quatro cores de forma e distinguir ainda assim? Os cartógrafos podem ter acreditado nisto durante séculos, mas haverá uma prova rigorosa? Podemos pensar em qualquer mapa no mundo para além do mapa dos condados ingleses, como o mapa dos estados norte-americanos ou dos departamentos franceses, e mesmo mapas artificiais, construídos por fronteiras e regiões arbitrárias. Três cores, no entanto, não são suficientes.

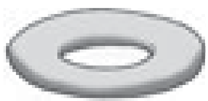
Observemos o mapa dos estados do Oeste dos EUA. Se só tivéssemos azul, verde e vermelho, podíamos começar por colorir o Nevada e o Idaho. Não interessa por que cor começamos, portanto escolhamos azul para o Nevada e verde para o Idaho. Até agora, tudo bem. Esta escolha significa que o Utah tem de ser colorido de vermelho, e por sua vez o Arizona de verde, a Califórnia de vermelho, e o Óregão de verde. Isto significa que tanto o Óregão como o Idaho estão coloridos de verde, logo não podem ser distinguidos. Mas, se tivéssemos quatro cores, juntando por exemplo o amarelo, podíamos usar a quarta cor para o Óregão e tudo seria satisfatório. Serão estas quatro cores – azul, verde, vermelho e amarelo – suficientes para qualquer mapa? Esta questão é conhecida como problema das quatro cores.

A propagação do problema No espaço de 20 anos desde o reconhecimento da importância do problema por parte de De Morgan, o problema tornou-se conhecido na comunidade matemática da Europa e da América do Norte. Na década de 1860, Charles Sanders Pierce, um matemático e filósofo norte-americano, pensou que o tinha provado, mas não existem vestígios do seu argumento.

O problema adquiriu maior proeminência pela intervenção de Francis Galton, um homem de ciência vitoriano. Galton viu nele valor publicitário e convenceu o eminente matemático de Cambridge Arthur Cayley a escrever um artigo sobre o problema em 1878. Infelizmente, Cayley foi forçado a admitir que tinha falhado na obtenção da prova, mas observou que era suficiente considerar apenas mapas cúbicos (onde três países se encontram exactamente num ponto). Esta contribuição estimulou o seu aluno Alfred Bray Kempe a procurar uma solução. Apenas um ano mais tarde, Kempe anunciou que tinha encontrado a prova. Cayley congratulou-o, a sua prova foi publicada, e Kempe entrou na Real Sociedade de Londres.

1976
Appel e Haken dão uma prova baseada num computador para o resultado geral

1994
A prova por computador é simplificada, mas mantém-se uma prova baseada num computador



O *donut* simples ou «toro»

Que aconteceu a seguir? A prova de Kempe era longa e tecnicamente exigente, e embora uma ou duas pessoas não estivessem convencidas, ela foi geralmente aceite. Foi uma surpresa quando dez anos mais tarde Percy Heawood, de Durham, encontrou um exemplo de um mapa que revelava uma falha no argumento de Kempe. Embora não tenha conseguido encontrar a sua própria prova, Heawood mostrou que o desafio do problema das quatro cores ainda estava em aberto. Voltaria às secretárias dos matemáticos e um novo principiante teria a oportunidade de deixar a sua marca. Usando algumas das técnicas de Kempe, Heawood provou um problema de cinco cores – que qualquer mapa pode ser colorido com cinco cores. Este teria sido um grande resultado se alguém conseguisse construir um mapa que não pudesse ser colorido com quatro cores. Os matemáticos têm assim um dilema: serão quatro ou cinco?

Um toro com dois buracos

O problema básico das quatro cores respeitava mapas desenhados numa superfície plana ou esférica. E que dizer de mapas desenhados numa superfície como um *donut* (uma superfície mais interessante para os matemáticos pela sua forma do que pelo seu sabor)? Heawood provou que eram necessárias e suficientes sete cores para colorir qualquer mapa desenhado nessa superfície. Provou ainda um resultado para um *donut* com muitos buracos (com um número, b, de buracos), no qual contou o número de cores que garantiam que qualquer mapa fosse colorido – embora não tivesse provado que esse era o número mínimo de cores. Uma tabela dos primeiros valores de Heawood para b é:

número de buracos, b	1	2	3	4	5	6	7	8
número de cores suficientes, C	7	8	9	10	11	12	12	13

e genericamente, $C = [\frac{1}{2}(7 + \sqrt{(1 + 48b)})]$. Os parênteses rectos indicam que só se toma a parte inteira do termo dentro deles. Por exemplo, quando $b = 8$, $C = [13{,}3107...] = 13$. A fórmula de Heawood derivou do estrito entendimento de que o número de buracos é maior do que zero. Se se forçar a fórmula, tem-se a resposta $C = 4$ quando o valor proibido de $b = 0$ é substituído.

Problema resolvido? Passados 50 anos, o problema que tinha emergido em 1852 permanecia por provar. No século XX, a elite matemática mundial continuava perplexa.

Fizeram-se alguns progressos e um matemático provou que quatro cores eram suficientes para até 27 países num mapa, outro melhorou para 31 países e ainda outro avançou 35 países. Este processo duraria para sempre, se continuado. De facto, as observações feitas por Kempe e Cayley nos seus artigos iniciais proporcionavam um melhor caminho, e os matemáticos perceberam que só tinham de verificar algumas configurações de mapas para garantir que as quatro cores eram suficientes. O problema era que existia um largo número delas – nos estados iniciais destas tentativas de prova, existem milhares para verificar. Esta verificação não podia ser feita à mão, mas felizmente o matemático alemão Wolfgang Haken, que tinha trabalhado no problema durante muitos anos, conseguiu os serviços de Kenneth Appel, um matemático e perito informático norte-americano. Métodos engenhosos baixaram o número de configurações para menos de 1500. No fim de Junho de 1976, depois de muitas noites sem dormir, o trabalho estava feito e, em parceria com o seu fidedigno computador IBM 370, tinham resolvido o grande problema.

O departamento de matemática da Universidade de Indiana tinha uma nova trombeta para soprar. Substituíram o seu selo do «maior primo descoberto» por um novo que dizia «quatro cores bastam». Foi um orgulho local, mas onde estava o aplauso geral da comunidade mundial de matemáticos? Afinal de contas, era um problema venerável que podia ser compreendido por qualquer pessoa com a tenra idade do Pequeno Tim, mas que por mais de um século tinha aborrecido e torturado alguns dos maiores matemáticos do mundo.

O aplauso foi desigual. Alguns aceitaram de má vontade que o trabalho tivesse sido concluído, mas continuaram cépticos. O incómodo era que a prova fosse baseada num computador, o que não era a forma tradicional da prova matemática. Seria de esperar que a prova fosse difícil de seguir, e possivelmente longa, mas uma prova de computador era ir longe demais. Levantava a questão da «verificabilidade». Como podia alguém verificar as milhares de linhas de código de computador das quais a prova dependia? Podiam decerto ser cometidos erros na codificação. Um erro podia demonstrar-se fatal.

Isto não era tudo. O que de facto faltava era o factor «uau!». Como poderia alguém ler a prova e apreciar a subtileza do problema, ou vivenciar a parte crucial do argumento, o momento «uau»? Um dos mais ferozes críticos foi o eminente matemático Paul Halmos. Halmos achava que uma prova de computador tinha tanta credibilidade como uma prova dum respeitável vidente. Mas muitos reconheceram a conquista, e seria preciso uma pessoa corajosa ou tola para gastar o seu precioso tempo de pesquisa tentando encontrar um contra-exemplo de um mapa que necessitasse de cinco cores. Podia tê-lo feito antes de Appel e Haken, mas não depois.

Depois da prova Desde 1976, o número de configurações a ser verificadas foi reduzido a metade e os computadores têm-se tornado mais rápidos e mais potentes. Dito isto, o mundo matemático ainda espera por uma prova mais curta no estilo tradicional. Entretanto, o teorema das quatro cores tem gerado problemas importantes na teoria dos grafos e tem um efeito suplementar de desafiar os matemáticos sobre aquilo que constitui a própria noção de prova.

a ideia resumida
Bastarão quatro cores

31 Probabilidade

Qual é a hipótese de nevar amanhã? Qual é a possibilidade de eu apanhar o primeiro comboio da manhã? Qual é a probabilidade de eu ganhar a lotaria? Probabilidade, possibilidade, hipótese são palavras que usamos todos os dias quando queremos conhecer as respostas. Também são palavras da teoria matemática das probabilidades.

A teoria das probabilidades é importante. Tem uma relação com a incerteza e é um ingrediente essencial na avaliação de risco. Mas como é que uma teoria que envolve incerteza pode ser quantificada? Afinal, a matemática não é uma ciência exacta?

O verdadeiro problema é *quantificar* a probabilidade.

Consideremos o problema mais simples do planeta, o lançamento de uma moeda. Qual é a probabilidade de sair caras? Podemos apressar-nos a dizer que é ½ (algumas vezes expressa como 0,5 ou 50%). Olhando para a moeda, presumimos que não é viciada, o que significa que a hipótese de obtermos uma cara iguala a hipótese de obtermos uma coroa, e portanto a probabilidade de uma cara é de ½.

As situações que envolvem moedas, bolas em caixas e exemplos «mecânicos» são relativamente simples. Há duas teorias principais de atribuição de probabilidades. Uma abordagem é analisar os dois lados simétricos de uma moeda. Outra é a frequência relativa, quando fazemos a experiência um grande número de vezes e contamos o número de caras. Mas o que quer dizer «grande»? É fácil acreditar que o número de caras relativamente ao número de coroas é grosseiramente 50:50, mas é possível que a proporção mude se continuarmos a experiência.

Cronologia

anos 1650
Os fundamentos da probabilidade são lançados por Pascal e Huygens

1785
Condorcet aplica as probabilidades à análise de júris e sistemas eleitorais

Mas o que dizer quando se trata da sensível medida da probabilidade de nevar amanhã? Existirão de novo dois resultados: ou neva ou não neva, mas não é de todo claro que elas sejam igualmente possíveis como no caso da moeda. Uma avaliação da probabilidade de nevar amanhã terá de ter em conta as condições do tempo na altura e um grande número de outros factores. Mas, mesmo assim, não é possível apontar um número exacto para esta probabilidade. Embora não cheguemos a um número efectivo, podemos atribuir um «grau de confiança» de que a probabilidade será baixa, média ou alta. Na matemática, a probabilidade é medida numa escala de 0 a 1. A probabilidade de um acontecimento impossível é 0 e a da certeza é 1. Uma probabilidade de 0,1 significa uma probabilidade baixa e uma probabilidade de 0,9 significa uma probabilidade alta.

Origens da probabilidade A teoria matemática da probabilidade nasceu no século XVII com discussões sobre problemas de jogos entre Blaise Pascal, Pierre Fermat e Antoine Gombaud (também conhecido por Chevalier de Méré). Encontraram um jogo incrivelmente simples. A questão do Chevalier de Méré é: o que é mais verosímil: tirar um seis em quatro lançamentos de um dado, ou tirar um «duplo seis» em 24 lançamentos de dois dados? Em qual das opções apostava?

A sabedoria predominante da altura considerou que a melhor opção era apostar no duplo seis, porque o número de lançamentos permitidos era muito maior. Esta visão foi destruída quando as probabilidades foram analisadas. Aqui estão os cálculos:

Lançamento um dado: a probabilidade de não obter seis num só lançamento é $5/6$, e em quatro lançamentos a probabilidade será $5/6 \times 5/6 \times 5/6 \times 5/6$ o que é $(5/6)^4$. Porque os resultados dos lançamentos não se afectam entre si, são «independentes» e podemos multiplicar as probabilidades. A probabilidade de sair pelo menos um seis é portanto

$$1 - (5/6)^4 = 0{,}517746\ldots$$

Lançamento de dois dados: a probabilidade de não sair um duplo seis num lançamento é $35/36$ e em 24 lançamentos $(35/36)^{24}$.

1812

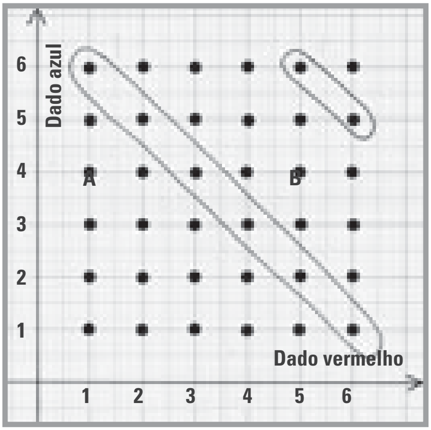
Laplace publica a sua *Teoria Analítica das Probabilidades*, em dois volumes

1912

Keynes publica o seu *Tratado sobre a Probabilidade*, que influencia as suas teorias de economia e estatística

1933

Kolmogorov apresenta a probabilidade de forma axiomática



Espaço amostral
(para dois dados)

A probabilidade de sair pelo menos um duplo seis é portanto

$$1 - (35/36)^{24} = 0{,}491404\ldots$$

Podemos levar este exemplo mais longe.

Jogar *craps* O exemplo dos dois dados é a base do moderno jogo de *craps* jogado nos casinos e nas apostas *online*. Quando dois dados distinguíveis (vermelho e azul) são lançados, existem 36 possíveis resultados, que podem ser registados como pares (x, y) e representados como 36 pontos num conjunto de eixos x/y – chamado o «espaço amostral»

Consideremos o acontecimento A, de obter os dados cuja soma é 7. Existem 6 combinações que somam 7, logo podemos descrever o acontecimento como

$$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$

e cercá-lo no diagrama. A probabilidade de A é 6 em 36, o que pode ser escrito $\Pr(A) = 6/36 = 1/6$. Se B for o acontecimento de obter dados cuja soma é 11, temos o acontecimento $B = \{(5, 6), (6, 5)\}$ e $\Pr(B) = 2/36 = 1/18$.

No jogo de dados *craps*, em que se lançam dois dados numa mesa, podemos ganhar ou perder na primeira jogada, mas para algumas pontuações nem tudo está perdido e podemos passar para a segunda jogada. Ganhamos na primeira jogada se o acontecimento A ou B ocorrer – uma «natural». A probabilidade de conseguir uma natural é obtida somando as probabilidades individuais, 6/36 + 2/36 = 8/36. Perdemos na primeira jogada se lançarmos 2, 3 ou 12 («*craps*»). Um cálculo como o acima dá uma probabilidade de perder na primeira jogada de 4/36. Se a soma no lançamento for 4, 5, 6, 8, 9 ou 10, passamos para a segunda jogada, e a probabilidade de isto acontecer é 24/36 = 2/3.

No *craps*, por cada 36 jogos que jogarmos, em média ganharemos ao primeiro lançamento 8 vezes e não ganharemos 28 vezes, logo a probabilidade de não ganhar no primeiro lançamento é de 28 para 8, o que é o mesmo de 3,5 para 1.

O macaco e a máquina de escrever O *Alfredo* é um macaco do jardim zoológico. Tem uma velha e maltratada máquina de escrever com 26 teclas com as letras do alfabeto, uma tecla com o ponto final, outra com a vírgula, outra com o ponto de interrogação e outra com o espaço – 30 teclas ao todo. Senta-se num canto, cheio de ambições literárias, mas o seu método de escrever é curioso – ele bate as teclas aleatoriamente.

Qualquer sequência de letras escrita terá uma hipótese não nula de ocorrer, logo existe a possibilidade de ele escrever as peças de Shakespeare palavra por palavra. Mais do que isso, existe uma hipótese (embora pequena) de que continue com a tradução para francês, e depois espanhol, e depois alemão. Em boa medida, poderíamos permitir a possibilidade de ele continuar com os poemas de William Wordsworth. A hipótese de tudo isto é mínima, mas não é certamente zero. Este é o ponto crucial. Vejamos quanto tempo demoraria ele para escrever o solilóquio de *Hamlet*, começando com o início «*To be or*» («Ser ou»). Consideremos 8 caixas que conterão 8 letras incluindo espaços.

T	o		b	e		o	r

O número de possibilidades para a primeira posição é 30, para a segunda, 30, etc. Assim, o número de maneiras de encher as 8 caixas é 30 × 30 × 30 × 30 × 30 × 30 × 30 × 30. A probabilidade de o *Alfredo* chegar tão longe quanto «To be or» é 1 em $6{,}561 \times 10^{11}$. Se o Alfredo bater uma tecla a cada segundo, há uma expectativa de que tenha escrito «*to be or*» daqui a 20 000 anos e provado a si próprio ser um primata com uma vida particularmente longa. Portanto, não sustenhamos a respiração à espera das obras completas de Shakespeare. O *Alfredo* produzirá disparates como «xo,h?yt?» durante muito tempo.

Como se desenvolveu a teoria? Quando a teoria das probabilidades é aplicada, os resultados podem ser controversos, mas pelo menos as bases matemáticas são razoavelmente seguras. Em 1933, Andrey Nikolaevich Kolmogorov contribuiu para definir a probabilidade numa base axiomática – de forma muito parecida àquela com que os princípios da geometria foram definidos há dois milénios.

A probabilidade é definida pelos seguintes axiomas:

1. A probabilidade de todas as ocorrências é 1
2. A probabilidade tem um valor que é maior ou igual a zero
3. Quando as ocorrências não podem coincidir, as suas probabilidades podem ser somadas

A partir destes axiomas, escritos em linguagem técnica, podem deduzir-se as propriedades matemáticas da probabilidade. O conceito de probabilidade pode ser largamente aplicado. Muita da vida moderna não podia passar sem ela. A análise de risco, o desporto, a sociologia, a psicologia, a engenharia de projecto, as finanças, etc. – a lista é interminável. Quem imaginaria que os problemas de jogos que iniciaram estas ideias no século XVII gerariam uma disciplina tão grande? Qual era a probabilidade de isso acontecer?

a ideia resumida
O sistema secreto do jogador

32 Teoria de Bayes

Os primeiros anos do reverendo Thomas Bayes são obscuros. Nascido no Sudeste de Inglaterra, provavelmente em 1702, tornou-se um ministro religioso inconformista, mas também ganhou reputação como matemático e foi eleito para a Real Sociedade de Londres em 1742. O seu famoso *Essay towards solving a problem in the doctrine of chances* foi publicado em 1763, dois anos depois da sua morte. Bayes deu a fórmula para encontrar a probabilidade inversa, a probabilidade «ao contrário», que ajudou a criar o conceito central da filosofia bayesiana – probabilidade condicional.

Thomas Bayes emprestou o nome aos bayesianos, os adeptos de um ramo da estatística que discorda dos estatísticos tradicionais ou «frequencistas». Os frequencistas adoptam uma visão da probabilidade baseada em dados numéricos pesados. A visão bayesiana é centrada na famosa fórmula de Bayes e no princípio de que se podem tratar graus subjectivos de confiança como probabilidades matemáticas.

Probabilidade condicional Suponhamos que o enérgico Dr. Porquê tem a tarefa de diagnosticar sarampo nos seus doentes. O surgimento de manchas é um indicador usado para a detecção, mas o diagnóstico não é simples. Um doente pode ter sarampo sem ter manchas e alguns doentes podem ter manchas sem ter sarampo. A probabilidade de que um doente tenha manchas se tiver sarampo é uma probabilidade condicional. Os bayesianos usam uma barra vertical na sua fórmula para significar «se», logo

prob(um doente tem manchas | o doente tem sarampo)

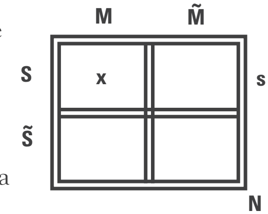
significa a probabilidade de o doente ter manchas se tiver sarampo. O valor de *prob*(um doente tem manchas | o doente tem sarampo) não é a mesma que *prob*(um doente tem sarampo | o doente tem manchas). Em relação uma

Cronologia

1763
Publicação do ensaio de Bayes sobre probabilidades

1937
De Finetti defende a probabilidade subjectiva como uma alternativa à teoria da frequência

à outra, uma é a probabilidade ao contrário. A fórmula de Bayes é a fórmula de calcular uma a partir da outra. Não há nada de que os matemáticos mais gostem do que usar notação para designar coisas. Assim, digamos que o acontecimento de ter sarampo é S e o acontecimento de um doente ter manchas, M. O símbolo $\tilde{S}$ é o acontecimento de o doente *não* ter sarampo e $\tilde{M}$ o acontecimento de o doente *não* ter manchas. Podemos ver isto num diagrama de Venn.



O diagrama de Venn mostrando a estrutura lógica do sarampo e do aparecimento de manchas

Isto diz ao Dr. Porquê que existem x doentes que têm sarampo e manchas, s é o número de doentes que têm sarampo enquanto o número total de doentes é N. A partir do diagrama, podemos ver que a probabilidade de alguém ter sarampo e manchas é simplesmente x/N, enquanto a probabilidade de alguém ter sarampo é s/N. A probabilidade condicional, a probabilidade de alguém ter manchas se tiver sarampo, representada por $prob(M \mid S)$, é x/s. Juntando tudo, o Dr. Porquê tem a probabilidade de alguém ter quer sarampo quer manchas:

$$prob(S \,\&\, M) = \frac{x}{N} = \frac{x}{s} \times \frac{s}{N}$$

ou

$$prob(S \,\&\, M) = prob(M \mid S) \times prob(S)$$

e de forma semelhante

$$prob(S \,\&\, M) = prob(S \mid M) \times prob(M)$$

A fórmula de Bayes Igualando as expressões de $prob(S \,\&\, M)$, obtém-se a fórmula de Bayes, a relação entre a probabilidade condicional e a sua inversa. O Dr. Porquê terá uma ideia razoável do valor da $prob(M \mid S)$, a probabilidade de o doente ter manchas se tiver sarampo. Ao contrário, a probabilidade condicional é aquela em que o Dr. Porquê está realmente interessado, é a sua avaliação de um doente ter sarampo se tiver manchas. Determiná-lo é o problema inverso e o tipo de problema tratado por Bayes no seu ensaio.

$$prob(S \mid M) = \frac{prob(S)}{prob(M)} \times prob(M \mid S)$$

A fórmula de Bayes

Para calcular a probabilidade, precisamos de alguns números. Serão subjectivos, mas o importante é ver como se combinam. A probabilidade de, se tiverem sarampo, os doentes terem manchas, $prob(M \mid S)$ será elevada, digamos 0,9, e se os doentes não tiverem sarampo, a probabilidade de terem manchas $prob(M \mid \tilde{S})$ será baixa, digamos 0,15. Em ambas as situações, o Dr. Porquê terá uma ideia do valor destas probabilidades. O enérgico médico também terá uma ideia da

1950
Jimmy Savage e Dennis Lindley lideram o moderno movimento bayesiano

anos 1950
Começo do uso do termo «bayesiano»

1992
Fundação da Sociedade Internacional de Análise Bayesiana

percentagem de pessoas que têm sarampo entre a população, digamos 20%. Isto é expresso como *prob*(*S*) = 0,2. A única parcela de informação de que ainda precisamos é *prob*(M), a percentagem da população que tem manchas. Ora a probabilidade de alguém ter manchas é a probabilidade de alguém ter sarampo *e* manchas mais a probabilidade de alguém não ter sarampo mas ter manchas. Das nossas relações principais, *prob*(M) = 0,9 × 0,2 + 0,15 × 0,8 = 0,3. Substituindo estes valores na fórmula de Bayes, temos:

$$prob(S \mid M) = \frac{0{,}2}{0{,}3} \times 0{,}9 = 0{,}6$$

A conclusão é que, de todos os doentes com manchas que vê, o médico identifica correctamente o sarampo em 60% dos casos. Suponhamos agora que o médico recebe mais informação sobre a estirpe de sarampo de modo que a probabilidade de detecção aumenta, ou seja a *prob*(M | *S*), a probabilidade de ter manchas provocadas por sarampo, sobe de 0,9 para 0,95 e *prob*(M | $\tilde{S}$), a probabilidade de ter manchas por outra causa, diminui de 0,15 para 0,1. Como é que esta mudança melhora a sua taxa de detecção de sarampo? Qual é a nova *prob*(*S* | M)? Com esta nova informação, *prob*(M) = 0,95 × 0,2 + 0,1 × 0,8 = 0,27, logo na fórmula de Bayes, *prob*(*S* | *M*) é 0,2 dividido por *prob*(M) = 0,27 e, tudo multiplicado por 0,95, dá 0,704. Logo, o Dr. Porquê pode agora detectar 70% dos casos com esta informação melhorada. Se as probabilidades mudarem para 0,99 e 0,01, respectivamente, a probabilidade da detecção, *prob*(*S* | M), passa a ser 0,961, donde a sua possibilidade de diagnóstico correcto nesse caso ser de 96%.

Bayesianos dos dias de hoje Os estatísticos tradicionais não se oporão ao uso da fórmula de Bayes quando a probabilidade possa ser medida. O ponto de discórdia é a interpretação da probabilidade como graus de confiança ou, como às vezes se define, a probabilidade subjectiva.

Num tribunal, as questões de culpa ou inocência são por vezes decididas por um «balanço de probabilidades». Em bom rigor, este critério só se aplica a casos civis, mas podemos conceber um cenário em que também se aplique a casos criminais. O estatístico frequencista tem um problema em atribuir significado à probabilidade de o prisioneiro ser culpado de um crime. O mesmo não acontece com os bayesianos, que não se importam de abarcar os sentimentos. Como funciona? Se usarmos o método do «balanço de probabilidades» para julgar a culpa ou a inocência, poderemos ver como as probabilidades podem ser enganadoras. Eis um cenário possível.

Um jurado acabou de ouvir um caso no tribunal e decidiu que a probabilidade de o acusado ser culpado é cerca de 1 em 100. Durante a deliberação na sala

dos jurados, estes são chamados de volta ao tribunal para ouvirem mais provas apresentadas pela acusação. Uma arma foi encontrada na casa do prisioneiro e a acusação alega que a probabilidade de a encontrarem é 0,95 se o prisioneiro for culpado, mas 0,1, se ele for inocente.
A probabilidade de encontrar uma arma em casa do prisioneiro é muito maior se ele for culpado do que se for inocente. A questão para os jurados é saber como devem modificar a sua opinião sobre o prisioneiro à luz da nova informação. Usando de novo a nossa notação, C é o acontecimento do prisioneiro ser culpado e *P* o acontecimento de obtenção de nova prova. Os jurados tinham feito uma avaliação inicial de que *prob*(C) = 1/100, ou 0,01. Esta probabilidade é chamada probabilidade *a priori*. A reavaliação da probabilidade *prob*(C | *P*) é a probabilidade de culpa revista dada a nova prova *P*, e é chamada probabilidade *a posteriori*. A fórmula de Bayes na forma

$$prob(C \mid P) = \frac{prob\ (P \mid C)}{prob\ (P)} \times prob(C)$$

mostra a ideia de a probabilidade a *priori* ser actualizada para a probabilidade a *posteriori prob*(C | *P*). Assim como se resolveu *prob*(M) no exemplo médico, podemos resolver *prob*(*P*) e temos

$$prob(C \mid P) = \frac{0{,}95}{0{,}95 \times 0{,}01 + 0{,}1 \times 0{,}99} \times 0{,}01 = 0{,}088$$

Isto representa um dilema para os jurados, porque a avaliação inicial de 1% de hipóteses de culpa aumentou para quase 9%. Se a acusação tivesse alegado que a probabilidade de encontrar a arma incriminatória era 0,99 se o prisioneiro fosse culpado, mas apenas 0,01 se ele fosse inocente, repetindo a fórmula de Bayes, os jurados teriam de reavaliar a sua opinião de 1% para 50%.

O uso da fórmula de Bayes nestas situações tem sido exposto a críticas. O ponto principal tem sido a forma de chegar à probabilidade *a priori*. A análise de Bayes tem a seu favor apresentar uma forma de tratar probabilidades subjectivas e de como elas podem ser actualizadas com base nas provas.
O método bayesiano tem aplicações em áreas tão diversas como a ciência, a previsão meteorológica e a justiça criminal. Os seus proponentes defendem o seu carácter robusto e pragmático no tratamento da incerteza. Têm muito a seu favor.

a ideia resumida
Actualizar crenças recorrendo à evidências

33 O problema do aniversário

Suponha que está no piso superior de um autocarro de dois andares sem nada de particular para fazer a não ser contar os seus companheiros de viagem a deslocarem-se para o trabalho de manhã. Como é provável que todos os passageiros sejam independentes uns dos outros, podemos seguramente presumir que os seus aniversários estejam dispersos aleatoriamente ao longo do ano. Há apenas 23 passageiros, contando consigo. Não são muitos, mas são os suficientes para afirmar que há uma hipótese de mais de 50% de dois passageiros terem a mesma data de aniversário. Acredita? Milhões de pessoas não acreditam, mas é absolutamente verdade. Mesmo William Feller, um perito com experiência em probabilidades, o achou espantoso.

O autocarro de dois andares tornou-se demasiado pequeno, portanto retomamos o argumento numa sala grande. Quantas pessoas se devem juntar na sala para termos a certeza de que existem duas pessoas com a mesma data de aniversário? Há 365 dias num ano normal (vamos ignorar os anos bissextos para tornar as coisas mais simples), logo se estiverem 366 pessoas na sala, pelo menos um par terá *definitivamente* o mesmo dia de aniversário. Não pode acontecer que todos os tenham diferentes.

Este é o princípio dos compartimentos dos pombos: Se existirem $n + 1$ pombos que ocupam n compartimentos, um compartimento deve conter mais do que um pombo. Se existirem 365 pessoas na sala, não podemos ter a certeza de que haja um aniversário comum, porque os aniversários podem distribuir-se por

Cronologia

1654
Blaise Pascal estabelece as bases da teoria das probabilidades

1657
Christiaan Huygens escreve o primeiro trabalho publicado sobre probabilidades

1718
Abraham de Moivre publica *The Doctrine of Chance*, com edições ampliadas em 1738 e 1756

cada um dos dias do ano. Contudo, se escolhermos 365 pessoas aleatoriamente isso será extremamente improvável e a probabilidade de duas pessoas não partilharem a data de aniversário será ínfima. Mesmo se houver apenas 50 pessoas na sala, há uma probabilidade de 96,5% de que duas pessoas tenham a mesma data de aniversário. Se o número de pessoas for ainda mais reduzido, a probabilidade de partilha de aniversário reduz-se. Constatamos que 23 pessoas é o número para o qual a probabilidade é ligeiramente superior a ½ e para 22 pessoas a probabilidade de partilha do aniversário é ligeiramente inferior a ½. O número 23 é o valor crítico. Embora a resposta ao problema do aniversário seja surpreendente, não é um paradoxo.

Podemos prová-lo? Como podemos ser convencidos? Vamos escolher uma pessoa aleatoriamente. A probabilidade de que outra pessoa tenha o mesmo dia de aniversário da primeira é 1/365 e, logo, a probabilidade de esses dias serem diferentes é 1 menos esta (ou 364/365). A probabilidade de que ainda outra pessoa escolhida ao acaso partilhe o dia de aniversário com as duas primeiras é 2/365 e, assim, a probabilidade de que não partilhe o dia de aniversário com nenhuma das primeiras é de 1 menos esta (ou 363/365). A probabilidade de nenhuma destas três partilhar o dia de aniversário é a multiplicação destas duas probabilidades, ou (364/365) × (363/365), que dá 0,9918.

Continuando esta linha de pensamento para 4, 5, 6, ... pessoas, deslinda-se o paradoxo do problema do aniversário. Quando chegamos a 23 pessoas, com uma calculadora simples, obtemos a resposta 0,4927 como a probabilidade de nenhuma delas partilhar o dia de aniversário. A negação de «nenhuma delas partilhar o dia de aniversário» é «pelo menos duas pessoas partilham o dia de aniversário» e a sua probabilidade é 1 – 0,4927 = 0,5073, ligeiramente superior ao crucial ½.

Se $n = 22$, a probabilidade de duas pessoas partilharem o dia de aniversário é 0,4757, o que é menor do que ½. A aparente natureza de paradoxo do problema do aniversário está ligada à língua. O resultado do aniversário faz uma afirmação acerca de duas pessoas que partilham o dia de anos. Mas não nos diz quem são elas. Não sabemos onde é que as igualdades se manifestarão. Se o Sr. Trevor Thomson, cujo aniversário é a 8 de Março, estivesse na sala, poderíamos fazer uma pergunta diferente.

anos 1920
Bose considera a teoria da luz de Einstein como um problema de ocupação

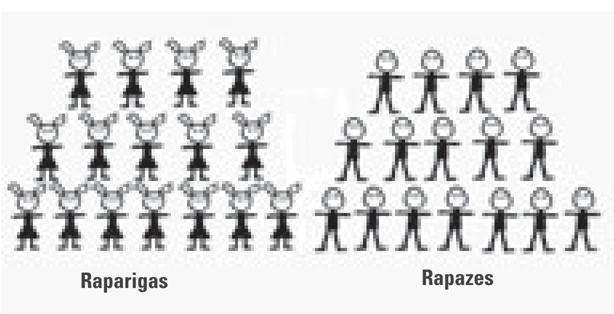
1939
Richard von Mises propõe o problema do aniversário

Quantos aniversários coincidem com o do Sr. Thomson? Para esta questão, o cálculo seria diferente. A probabilidade de o Sr. Thomson não partilhar o seu dia de aniversário com outra pessoa é (364/365), logo a probabilidade de não partilhar o dia de aniversário com qualquer outra das $n - 1$ pessoas na sala é $(364/365)^{n-1}$. Logo a probabilidade de o Sr. Thomson partilhar o seu dia de aniversário com alguém será 1 menos este valor.

Se calcularmos para $n = 23$, esta probabilidade é apenas 0,061151, logo há apenas 6% de hipóteses de que alguém mais faça anos a 8 de Março, a mesma data do aniversário do Sr. Thomson. Se aumentarmos o valor de n, a probabilidade aumenta. Mas temos de ir tão longe como $n = 254$ (que inclui o Sr. Thomson) para a probabilidade ser maior do que ½. Para $n = 254$, o seu valor é de 0,5005. Este é o ponto de corte, porque $n = 253$ dará o valor 0,4991, que é menor do que ½. Haverá que juntar 254 pessoas na sala para existir uma hipótese superior a ½ de que o Sr. Thomson partilhe o seu dia de aniversário com mais alguém. Este resultado está, talvez, mais em sintonia com a nossa intuição do que a surpreendente solução do clássico problema do aniversário.

Outros problemas de aniversários O problema do aniversário tem sido generalizado de muitas formas. Uma abordagem é considerar três pessoas que partilham o dia de aniversário. Nesse caso, seriam necessárias 88 pessoas antes de existir uma hipótese maior do que ½ de que três pessoas partilhassem o aniversário. Há grupos correspondentes maiores se forem quatro pessoas, cinco pessoas, etc., quantas sejam necessárias para partilharem o dia de aniversário. Num conjunto de 1000 pessoas, por exemplo, existe uma possibilidade maior do que ½ de que nove partilhem o dia de aniversário.

Uma das incursões no problema do aniversário tem sido a investigação dos aniversários próximos. Neste problema, considera-se que foi conseguido um par se um aniversário está afastado do outro menos de um certo número de dias. Acontece que apenas 14 pessoas numa sala darão uma hipótese superior a ½ de duas pessoas terem um aniversário comum ou fazerem anos em dias seguidos.



Uma variante do problema do aniversário que requer ferramentas matemáticas mais sofisticadas é o problema do aniversário de rapazes e raparigas: se uma turma consiste num número igual de raparigas e de rapazes, qual será o grupo mínimo que dará uma hipótese maior do que ½ de que um rapaz e uma rapariga partilhem o dia de aniversário?

O resultado é que uma turma de 32 (16 raparigas e 16 rapazes) faria um grupo mínimo. O número pode ser comparado ao 23 no problema clássico do aniversário.

Mudando ligeiramente a questão, podemos obter outras novidades (embora pouco fáceis de responder). Suponhamos que se forma uma longa fila à porta de um concerto do Bob Dylan e as pessoas se juntam a ela ao acaso. Como estamos interessados em aniversários, podemos pôr de parte a hipótese de gémeos ou trigémeos chegarem juntos. Enquanto os fãs vão entrando, é-lhes perguntada a data de aniversário. A questão matemática é: Quantas pessoas se espera que entrem até duas pessoas *consecutivas* terem a mesma data de aniversário? Outra questão: Quantas pessoas vão entrar para o concerto antes de aparecer uma pessoa com a mesma data de aniversário do Sr. Thomson (8 de Março)?

O cálculo do aniversário presume que os aniversários são uniformemente distribuídos e cada aniversário tem a mesma hipótese de se repetir se escolhermos uma pessoa ao acaso. Os resultados experimentais mostram que isto não é exactamente assim (nascem mais pessoas nos meses de Verão), mas está perto o suficiente para que a solução seja relevante.

Os problemas de aniversário são exemplos de problemas de ocupação, nos quais os matemáticos pensam na colocação de bolas em células. No problema do aniversário, o número de células é 365 (estas são identificadas como aniversários possíveis) e as bolas a serem colocadas ao acaso nas células são as pessoas). O problema pode ser simplificado para investigar a probabilidade de duas bolas caírem na mesma célula. Para o problema das raparigas e dos rapazes, as bolas são de duas cores.

Não são só os matemáticos que se interessam pelo problema do aniversário. Satyendra Nath Bose sentiu-se atraído pela teoria da luz de Einstein baseada nos fotões. Saiu das linhas tradicionais de pesquisa e considerou a configuração física como um problema de ocupação. Para ele, as células não eram os dias do ano do problema do aniversário, mas os níveis de energia dos fotões. Em vez de pessoas postas em células, como no problema do aniversário, Bose distribuiu fotões. Existem muitas aplicações de problemas de ocupação noutras ciências. Em biologia, por exemplo, a propagação de epidemias pode ser modelada como um problema de ocupação – as células neste caso são as áreas geográficas, as bolas são as doenças e o problema é perceber como é que as doenças se agrupam.

O mundo está cheio de assombrosas coincidências, mas só a matemática nos dá uma forma de calcular as suas probabilidades. O clássico problema do aniversário é só a ponta do icebergue nesta matéria e é uma grande entrada na matemática, com importantes aplicações.

a ideia resumida
O cálculo das coincidências

34 Distribuições

Ladislaus J. Bortkiewicz era fascinado pelas tabelas de mortalidade. Para ele, não era um tópico sombrio, era um campo de investigação científica estável. Ficou famoso por contar o número de cavaleiros do exército prussiano mortos por um coice. Depois dele, houve Frank Benford, um engenheiro electrotécnico que contou os primeiros dígitos de diferentes tipos de dados numéricos para ver quantos eram uns, dois, etc. E George Kingsley Zipf, que ensinava alemão em Harvard, interessava-se por filologia e analisou as ocorrências de palavras em blocos de texto.

Todos estes exemplos envolvem a medição da probabilidade de acontecimentos. Quais são as probabilidades de x cavaleiros num ano receberem um coice fatal de um cavalo? À listagem de todas as probabilidades para cada valor de x chama-se distribuição de probabilidades ou distribuição de probabilidade. É também uma distribuição *discreta*, porque os valores de x só podem tomar valores isolados – há lacunas entre os valores com interesse. Pode ter havido três ou quatro cavaleiros prussianos atingidos por um coice fatal, mas não 3½. Como veremos, no caso da distribuição de Benford, só estamos interessados no aparecimento dos dígitos 1, 2, 3, ... e, para a distribuição de Zipf, podemos ter a palavra «se» na oitava posição da lista de principais palavras, mas não na posição, digamos, 8,23.

Vida e morte no exército prussiano Bortkiewicz reuniu registos de 10 batalhões durante um período de 20 anos, obtendo dados para 200 batalhões-anos. Observou o número de mortes (a que os matemáticos chamariam a variável) e o número de batalhões-anos quando as mortes aconteceram. Por exemplo, existiam 109 batalhões-anos quando não ocorreram mortes, enquanto num batalhão-ano, houve 4 mortes. Nos quartéis, o batalhão C (digamos) num ano em particular sofreu quatro mortes.

Cronologia

1837
Siméon-Denis Poisson descreve a distribuição que tem o seu nome

1881
Newcomb descobre o que ficou conhecido pela lei de Benford

1898
Bortkiewicz analisa as mortes dos cavaleiros prussianos

Como está o número de mortes distribuído? Recolher esta informação é um lado do trabalho do estatístico – sair para o campo e registar resultados. Bortkiewicz obteve os seguintes dados:

Número de mortes	0	1	2	3	4
Frequência	109	65	22	3	1

Felizmente, ser morto por um coice de cavalo é um acontecimento raro. A técnica teórica mais adequada para modelar a frequência de acontecimentos raros é usar a chamada distribuição de Poisson. Com esta técnica, poderia Bortkiewicz ter previsto o resultado sem visitar os estábulos? A distribuição teórica de Poisson diz que a probabilidade de o número de mortes (a que chamaremos X) ter o valor x é dado pela fórmula de Poisson, em que e é o número especial já discutido, que está relacionado com crescimento (ver página 24), e o ponto de exclamação significa factorial, o número multiplicado por todos os outros números inteiros entre si próprio e 1 (ver página 26). A letra grega *lambda*, λ, é a média do número de mortes. Precisamos de encontrar essa média dos nossos 200 batalhões-anos, logo multiplicamos 0 mortes por 109 batalhões-anos (que dá 0), 1 morte por 65 batalhões-anos (que dá 65), 2 mortes por 22 batalhões-anos (que dá 44), 3 mortes por 3 batalhões-anos (que dá 9) e 4 mortes por 1 batalhão-anos (que dá 4) e agora somamos todos (que dá 122) e dividimos por 200. Logo, a nossa média de mortes por batalhões-anos é 122/200 = 0,61.

$$e^{-\lambda}\lambda^{x}/x!$$

A fórmula de Poisson

As probabilidades teóricas (a que chamaremos p) podem ser encontradas substituindo os valores r = 1, 2, 3 e 4 na fórmula de Poisson. Os resultados são:

Número de mortes	0	1	2	3	4
Probabilidades, p	0,543	0,331	0,101	0,020	0,003
Número de mortes esperadas, $200 \times p$	108,6	66,2	20,2	4,0	0,6

Parece que a distribuição teórica é um bom ajuste para os dados experimentais reunidos por Bortkiewicz.

Os primeiros números Se analisarmos os últimos dígitos dos números de telefone numa coluna da lista telefónica, prevemos que 0, 1, 2, ..., 9 estejam uniformemente distribuídos. Estes aparecem ao acaso e cada número tem a mesma probabilidade de aparecer. Em 1938, o engenheiro electrotécnico Frank Benford verificou que isso não era verdade para os primeiros dígitos dos mesmos

1938
Benford volta a apresentar a lei da distribuição dos primeiros dígitos

1950
Zipf deduz uma fórmula relacionando o uso das palavras com o vocabulário

2003
A distribuição de Poisson é usada na análise de *stocks* de peixe no Atlântico Norte

conjuntos de dados. De facto, redescobriu uma lei observada pela primeira vez pelo astrónomo Simon Newcomb em 1881.

Ontem realizei uma pequena experiência. Espreitei os dados sobre o câmbio das moedas estrangeiras num jornal nacional. Existiam taxas de câmbio como 2,119 para significar que precisaríamos de $ 2,119 para comprar £ 1. Da mesma forma, precisaríamos de €1,59 para comprar £ 1 e HK$ 15,390 para comprar £ 1. Revendo o resultado dos dados e registando o número de vezes que cada primeiro dígito apareceu, construí a seguinte tabela:

Primeiro dígito	1	2	3	4	5	6	7	8	9	Total
Número de ocorrências	18	10	3	1	3	5	7	2	1	**50**
Percentagem, %	36	20	6	2	6	10	14	4	2	**100**

Estes resultados suportam a lei de Benford, que diz que, para algumas classes de dados, o número 1 aparece em aproximadamente 30% dos dados, o número 2 em 18%, etc. Não é certamente a distribuição uniforme como a que acontece com o último dígito dos números de telefone.

Não é óbvio porque é que tantos conjuntos de dados seguem a lei de Benford. No século XIX, quando Simon Newcomb a observou na utilização de tabelas matemáticas dificilmente adivinharia que seria tão difundida.

Os casos em que a distribuição de Benford pode ser detectada incluem as pontuações dos acontecimentos desportivos, os dados dos mercados de acções, os números das moradas, a população dos países e o comprimento dos rios. As unidades de medida não são importantes – não interessa se o comprimento dos rios é medido em metros ou milhas. A lei de Benford tem aplicações práticas. Uma vez reconhecido que a informação contabilística seguia esta lei, tornou-se mais fácil detectar as informações falsas e expor as fraudes.

Palavras Um dos vastos interesses de G. K. Zipf era a prática pouco usual de contar palavras. Acontece que as dez palavras mais populares na língua inglesa são as pequenas palavras classificadas abaixo:

Classe	1	2	3	4	5	6	7	8	9	10
Palavra	the	of	and	to	a	in	that	it	is	was

Os valores são determinados tomando uma larga amostra de uma grande variedade de trabalho escrito e contando apenas as palavras. À palavra mais comum é dada a posição 1, à seguinte a posição 2, etc. Pode haver pequenas diferenças na popularidade, conforme o conjunto de textos analisados, mas não variarão muito.

Não é surpresa que «the» seja a mais comum e «of» a segunda. A lista continua e talvez lhe interesse saber que «among» está na posição 500.ª e «neck» na 1000.ª. Consideremos apenas as dez primeiras. Se pegarmos num texto ao acaso e contarmos essas palavras, obteremos mais ou menos as mesmas palavras na mesma posição. O facto surpreendente é que as posições têm influência no número efectivo de ocorrências das palavras num texto. A palavra «the» ocorrerá duas vezes mais do que a palavra «of» e três vezes mais do que «and», etc. O número efectivo é dado por uma fórmula bem conhecida. É uma lei experimental e foi descoberta por Zipf a partir de dados. A lei teórica de Zipf diz que a percentagem de ocorrências da palavra na posição r é dada por

$$\frac{k}{r} \times 100$$

em que o número k depende apenas da extensão do vocabulário do autor. Se um autor tiver domínio sobre todas as palavras da língua inglesa, cerca de um milhão segundo algumas estimativas, o valor de k será 0,0694. Na fórmula da lei de Zipf, a palavra «the» contará então como 6,94% de todas as palavras do texto. Da mesma forma, «of» contará metade disso, ou seja 3,47% das palavras. Um ensaio de 3000 palavras de um autor tão talentoso conterá portanto 208 ocorrências de «the» e 104 ocorrências da palavra «of».

Para escritores que dominem apenas 20 000 palavras, o valor de k eleva-se para 0,0954, logo existirão 286 ocorrências de «the» e 143 ocorrências da palavra «of». Quanto menor o vocabulário, mais frequentemente se verá «the» aparecer.

Contemplar a bola de cristal Quer seja a de Poisson, Benford ou Zipf, todas estas distribuições nos permitem fazer previsões. Podemos não ser capazes de prever uma morte certa, mas saber como as probabilidades se distribuem é muito melhor do que dar um tiro no escuro. Acrescente-se, a estas três, outras distribuições como a binomial, a binomial negativa, a geométrica, a hipergeométrica, e muitas outras. O estatístico tem uma lista de ferramentas eficazes para analisar um vasto leque da actividade humana.

a ideia resumida
Prever a quantidade

35 A curva normal

A curva «normal» desempenha um papel central em estatística. Tem sido chamada o equivalente da linha recta na matemática. Tem certamente propriedades matemáticas importantes, mas, se começarmos a trabalhar analisando um bloco de dados em bruto, raramente verificaremos que eles seguem exactamente uma curva normal.

A curva normal é determinada por uma fórmula matemática específica que cria uma curva em forma de sino, uma curva com uma bossa que vai diminuindo de cada um dos lados. A importância da curva normal está menos na natureza e mais na teoria, e nisso tem uma longa linhagem. Em 1733, Abraham de Moivre, um huguenote francês que se refugiou em Inglaterra para escapar à perseguição religiosa, introduziu-a em ligação com a sua análise do acaso. Pierre Simon Laplace publicou resultados acerca dela e Carl Friedrich Gauss usou-a em astronomia, campo em que é por vezes referida como a lei gaussiana do erro.

Adolphe Quetelet usou a curva normal nos seus estudos sociológicos publicados em 1835, em que mediu a divergência do «homem médio» pela curva normal. Noutras experiências, mediu as alturas dos seus recrutas e a medida do peito dos soldados escoceses e supôs que estes seguiam a curva normal. Na época, havia uma forte convicção de que a maioria dos fenómenos era «normal» neste sentido.

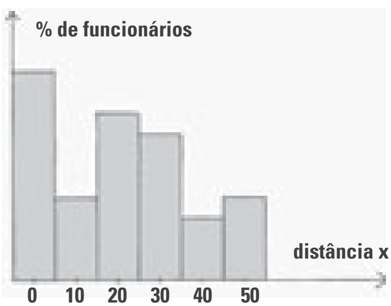
O *cocktail* Suponhamos que a Georgina foi a um *cocktail* e o anfitrião, Sebastian, lhe perguntou se ela tinha vindo de longe. Ela percebeu mais tarde que a pergunta era útil em *cocktails* – aplicava-se a toda a gente e convidava a uma resposta. Não tem que ver com impostos e inicia a conversa, se tal for difícil.

Cronologia

1733
De Moivre publica um trabalho sobre a curva normal como aproximação à distribuição binomial

1820
Gauss usa a distribuição normal (como a gaussiana) em astronomia como uma lei de erro

No dia seguinte, com uma ligeira ressaca, Georgina foi para o trabalho perguntando-se se os seus colegas teriam vindo de longe. Na cantina dos funcionários, soube que alguns viviam ao virar da esquina e outros viviam a 50 quilómetros dali – uma grande variabilidade. Georgina tirou vantagem do facto de ser directora dos Recursos Humanos de uma grande empresa para acrescentar a pergunta no fim do seu questionário anual aos funcionários: «Que distância percorreu para chegar ao trabalho hoje?» Queria calcular a distância média de viagem dos funcionários da companhia. Quando Georgina desenhou um histograma dos resultados, a distribuição representada não mostrou nenhuma forma particular, mas pelo menos permitiu-lhe calcular a média da distância percorrida.



Histograma de Georgina da distância percorrida pelos seus colegas de trabalho

Essa média revelou ser os 20 quilómetros. Os matemáticos indicam-na com a letra grega *mu*, escrita μ, e assim aqui $\mu = 20$. A variabilidade na população é indicada pela letra grega *sigma*, escrita σ, a que por vezes se chama desvio-padrão. Se o desvio-padrão é pequeno, os dados estão próximos uns dos outros e têm pequena variabilidade, mas se ele é grande os dados estão espalhados. O analista de *marketing* da empresa, que tinha formação em estatística, mostrou a Georgina que ela podia ter chegado ao mesmo valor por amostragem. Não havia necessidade de perguntar a *todos* os funcionários. Esta técnica de estimação depende do Teorema do Limite Central.

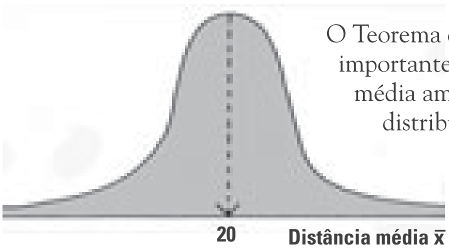
Tomemos uma amostra aleatória dos funcionários da empresa. Quanto maior a amostra, melhor, mas 30 funcionários bastará. Seleccionando esta amostra ao acaso, é provável que existam funcionários que vivam ao virar da esquina e outros que viajam longas distâncias. Quando calculamos a média das distâncias da amostra, o efeito das longas distâncias equilibrará as curtas distâncias.

1835
Quetelet usa a curva normal para medir as divergências relativamente ao homem médio

anos 1870
A distribuição adquire o nome «normal»

1901
Aleksandr Lyapunov prova o Teorema do Limite Central usando apenas funções características

Os matemáticos escrevem a média da amostra como $\bar{x}$, que se lê como «x barra». No caso de Georgina, é muito provável que o valor de $\bar{x}$ seja perto de 20, a média da população. Embora seja certamente possível, não é provável que a média da amostra seja demasiado pequena ou demasiado grande.



O Teorema do Limite Central é uma razão pela qual a curva normal é importante para os estatísticos. Ele afirma que a real distribuição da média amostral $\bar{x}$ se aproxima da curva normal para qualquer distribuição de x. O que é que isto significa? No caso da Georgina, x representa a distância ao local de trabalho, e $\bar{x}$ a média da amostra. A distribuição de x no histograma da Georgina não é minimamente semelhante a uma curva em sino, mas a distribuição de $\bar{x}$ é, e é centrada em $\mu = 20$.

Distribuição da média amostral

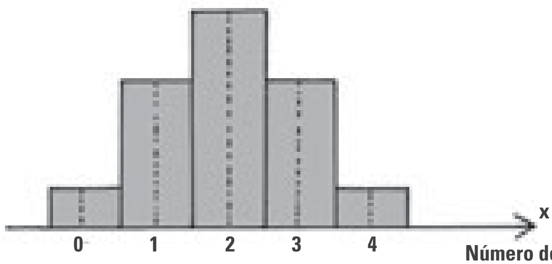
É por isto que podemos usar a média amostral $\bar{x}$ como uma estimativa da média da população μ. A variabilidade da média amostral $\bar{x}$ é uma vantagem adicional. Se a variabilidade dos valores de x é o desvio-padrão σ, a variabilidade de $\bar{x}$ é $\sigma/\sqrt{n}$, em que n é a dimensão da amostra que escolhemos. Quanto maior a amostra, mais estreita será a curva normal, e melhor será a estimativa de μ.

Outras curvas normais Façamos uma experiência simples. Atiremos uma moeda ao ar quatro vezes. A probabilidade de que saia sempre caras é $p = ½$. O resultado dos quatro lançamentos pode ser registado usando *H* para caras e *T* para coroas, segundo a ordem por que ocorrem. Ao todo, há 16 possíveis resultados. Por exemplo, podemos ter três caras no resultado *THHH*. Existem de facto quatro possíveis resultados com três caras (os outros são *HTHH*, *HHTH*, *HHHT*), donde a probabilidade de três caras 4/16 = 0,25.

Com um pequeno número de lançamentos, as probabilidades são facilmente calculadas e expostas numa tabela, e podemos calcular como é que as probabilidades estão distribuídas. A linha do número de combinações pode ser encontrada no triângulo de Pascal (ver página 52):

Número de caras	0	1	2	3	4
Número de combinações	1	4	6	4	1
Probabilidade	0,0625	0,25	0,375	0,25	0,0625
	(= 1/16)	(= 4/16)	(= 6/16)	(= 4/16)	(= 1/16)

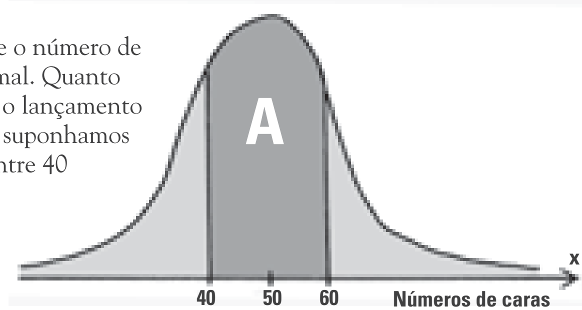
Esta é a chamada distribuição binomial de probabilidades, que ocorre quando há dois possíveis resultados (caras ou coroas). Estas probabilidades podem ser representadas num diagrama em que tanto as alturas quanto as áreas as descrevem.



O número de caras em quatro lançamentos de uma moeda, de acordo com a distribuição binomial

Lançar a moeda ao ar quatro vezes é um pouco limitativo. O que acontece se a lançarmos, digamos, 100 vezes? A distribuição binomial das probabilidades pode ser aplicada quando $n = 100$, mas pode ser proveitosamente aproximada pela curva normal em forma de sino com média $\mu = 50$ (porque há que esperar 50 caras, quando se lança uma moeda 100 vezes) e variabilidade (desvio-padrão) de $\sigma = 5$. Foi isto que De Moivre descobriu no século XVI.

Para grandes valores de n, a variável x que mede o número de êxitos enquadra-se perfeitamente na curva normal. Quanto maior for o valor de n, melhor a aproximação e o lançamento da moeda 100 vezes é tido como grande. Agora suponhamos que queríamos saber a probabilidade de obter entre 40 e 60 caras. A área A mostra a região em que estamos interessados e dá-nos a probabilidade de obter entre 40 e 60 caras, o que escrevemos como $prob(40 \leq x \leq 60)$. Para encontrar o valor numérico, precisamos de usar tabelas matemáticas pré-calculadas, e, uma vez isso feito, verificamos que $prob(40 \leq x \leq 60) = 0{,}9545$. Isto mostra que a probabilidade de obter entre 40 e 60 caras em 100 lançamentos é 95,45%, ou seja, muito alta.



Distribuição da probabilidade de sair caras em 100 lançamentos de uma moeda

A área restante é 1 – 0,9545, tão-só 0,0455. Como a curva normal é simétrica em relação ao seu meio, metade disto dará a probabilidade de obter mais de 60 caras em 100 lançamentos de uma moeda. Isto é apenas 2,275% e representa realmente uma probabilidade muito pequena. Se visitar Las Vegas, esta será uma aposta a não fazer.

a ideia resumida
A omnipresente curva em forma de sino

36 Relacionando dados

Como é que dois conjuntos de dados se relacionam? Os estatísticos de há cem anos pensavam saber a resposta. A correlação e a regressão andam juntas como um cavalo e uma carruagem, mas, tal como este par, são diferentes e têm as suas missões a cumprir. A correlação mede em que medida duas quantidades como o peso e a altura se relacionam entre si. A regressão pode ser usada para prever valores de uma propriedade (digamos o peso) a partir da outra (neste caso, a altura).

Correlação de Pearson O termo correlação foi introduzido por Francis Galton na década de 1880. Originalmente, Galton chamou-lhe «co-relação», para exprimir melhor o seu significado. Galton, um homem de ciência vitoriano, desejava medir tudo e aplicava a correlação na sua investigação a pares de variáveis: o comprimento das asas e da cauda das aves, por exemplo. O coeficiente da correlação de Pearson, que recebeu o nome de Karl Pearson, biógrafo e protegido de Galton, é medido numa escala entre menos um e mais um. Se o seu valor for alto, digamos +0,9, diz-se que existe uma forte correlação entre as variáveis. O coeficiente de correlação mede a tendência dos dados para se distribuírem ao longo de uma recta. Se for próximo de zero, a correlação é praticamente inexistente.

Frequentemente, desejamos descobrir a correlação entre duas variáveis para verificar com que intensidade elas se relacionam. Pensemos num exemplo como a venda de óculos de sol e vejamos como ela se relaciona com a venda de gelados. São Francisco será um bom local para levar a efeito o nosso estudo, pelo que recolheremos dados mensais nessa cidade. Se marcarmos os pontos num gráfico em que a coordenada x (horizontal) representa as vendas dos óculos de sol e a coordenada y (vertical) as vendas de gelados, em cada mês

Cronologia

1806
Adrien-Marie Legendre ajusta dados pelos mínimos quadrados

1809
Carl Friedrich Gauss usa o método dos mínimos quadrados em problemas de astronomia

1885-88
Galton introduz regressão e correlação

teremos um ponto (x, y) que representa ambas as informações. Por exemplo, o ponto (3, 4) pode significar que, no mês de Maio, as vendas de óculos de sol foram de \$ 30 000, enquanto a venda de gelados foi \$ 40 000 nesse mesmo mês. Podemos marcar mensalmente estes pontos (x, y) para um ano inteiro num diagrama de dispersão. Para este exemplo, o valor do coeficiente de correlação de Pearson será por volta de +0,9, o que indica uma forte relação. Os dados têm tendência para seguir uma linha recta. Isso é positivo, porque a recta tem um declive positivo – está apontada na direcção nordeste.

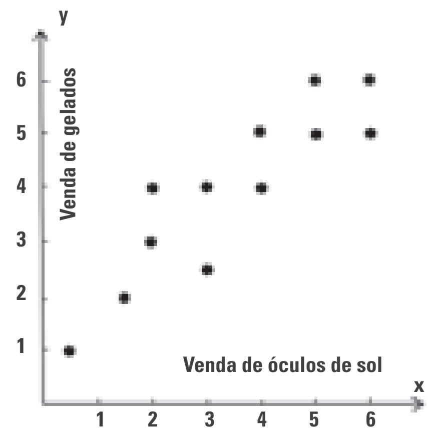


Diagrama de dispersão

Causa e correlação Encontrar uma forte relação entre duas variáveis não é suficiente para afirmar que uma causa a outra. Pode haver uma relação de causa e efeito entre duas variáveis, mas isso não pode ser afirmado apenas com base em evidências numéricas. Na questão da causa/correlação, é usual utilizar a palavra «associação» e sensato e prudente não afirmar mais do que isso.

No exemplo dos óculos de sol e dos gelados, há uma forte correlação entre as vendas dos óculos de sol e as dos gelados. Conforme as vendas dos óculos de sol aumentam, o número de gelados vendidos tende a aumentar. Seria absurdo afirmar que o consumo de óculos de sol causa a venda de mais gelados. Com a correlação, pode haver uma variável intermediária escondida em acção. Por exemplo, a despesa em óculos de sol e a despesa em gelados estão ligadas em resultado dos efeitos sazonais (tempo quente nos meses de Verão, tempo frio no Inverno). Há outro perigo no uso da correlação. Pode haver uma alta correlação entre duas variáveis mas nenhuma conexão lógica ou científica. Pode haver uma alta correlação entre o número das portas e as idades combinadas dos habitantes das casas, mas ler qualquer significado nisso seria lamentável.

Correlação de Spearman A correlação pode ser utilizada para outros fins. O coeficiente de correlação pode ser adaptado para tratar dados ordenados – dados em que queremos saber o primeiro, o segundo, o terceiro, etc., mas não necessariamente outros valores numéricos.

Ocasionalmente, só temos os dados das posições. Vejamos Albert e Zac, dois juízes não influenciáveis de patinagem no gelo, numa competição em que têm de avaliar o mérito artístico dos patinadores. Será uma avaliação subjectiva. Albert e Zac têm ambos medalhas olímpicas e foram chamados para julgar o grupo final, que foi reduzido a cinco competidores: Ann, Beth, Charlotte,

1896

Pearson publica contribuições sobre a correlação e a regressão

1904

Spearman usa a correlação como ferramenta em estudos psicológicos

Dorothy e Ellie. Seria óptimo que Albert e Zac as classificassem exactamente da mesma forma, mas a vida não é assim. Por outro lado, não se espera que Albert as classifique de uma determinada forma e que Zac as classifique exactamente na ordem inversa. A realidade é que as classificações devem estar entre esses dois extremos. Albert classificou-as de 1 a 5, começando por Ann (a melhor), seguida de Ellie, Beth, Charlotte e finalmente Dorothy, na 5.ª posição. Zac classificou Ellie como a melhor, seguida de Beth, Ann, Dorothy e Charlotte. Essas classificações podem ser resumidas na tabela.

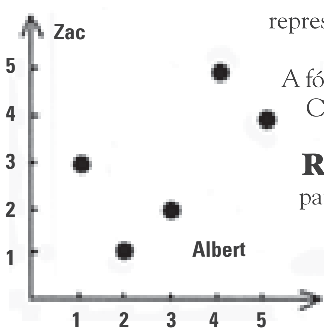
Patinadora	Classificações de Albert	Classificações de Zac	Diferença entre as classificações, d	d^2
Ann	1	3	– 2	4
Ellie	2	1	1	1
Beth	3	2	1	1
Charlotte	4	5	– 1	1
Dorothy	5	4	1	1
$n = 5$			Total	8

$$1 - \frac{6 \times total}{n \times (n^2-1)}$$

A fórmula de Spearman

Como podemos medir o nível de concordância entre os juízes? O coeficiente de correlação de Spearman é o instrumento que os matemáticos usam para o fazer para dados ordenados. O seu valor aqui é +0,6, o que indica que existe uma medida de concordância limitada entre Albert e Zac. Se tratarmos os pares de classificações como pontos, podemos marcá-los num gráfico, para obtermos uma representação visual do grau de concordância dos juízes.

A fórmula para este coeficiente de correlação foi desenvolvida em 1904 pelo psicólogo Charles Spearman que, como Pearson, foi influenciado por Francis Galton.



Medir a concordância entre os dois juízes

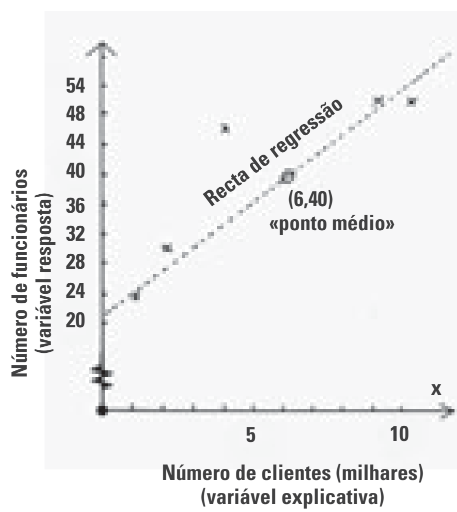
Rectas de regressão O leitor é mais baixo ou mais alto do que os seus pais? Ou está entre um e outro? Se fôssemos mais altos que os nossos pais e isso se verificasse em todas as gerações, um dia a população seria composta de pessoas com três metros ou mais, o que é impossível. Se fôssemos todos mais baixos que os nossos pais, a população diminuiria gradualmente, o que é igualmente pouco provável. A verdade estará noutro lugar.

Francis Galton realizou experiências na década de 1880 em que comparou as alturas de adultos jovens com as alturas dos seus pais. Para cada valor da variável x que media a altura dos pais (na realidade combinando a altura da mãe e do pai numa altura «média de pais»), Galton observou as alturas dos filhos. Sendo um cientista prático, pegou em lápis e folhas de papel divididas em quadrados que registou os dados. Para 205 médias de pais e 928 filhos, Galton determinou a média de alturas de ambos os conjuntos em 5 pés e 8¼ polegadas (173,4 cm), valor a que chamou a mediocridade. Galton verificou que os filhos com médias de pais muito altas eram geralmente mais altos do que essa mediocridade, mas não tão altos como os seus pais, enquanto os filhos mais baixos eram mais altos que os seus pais, mas mais baixos do que a mediocridade. Por outras palavras, as alturas dos filhos regridem para a mediocridade. É um pouco com o desempenho de um jogador de futebol excepcional. A sua média de golos numa época excepcional é provavelmente seguida por uma média inferior na seguinte, no entanto continuará a ser melhor que a média dos outros jogadores. Dizemos que a média de golos regrediu para a média.

A regressão é uma ferramenta poderosa e largamente aplicável. Suponhamos que, para um inquérito, a equipa de pesquisas de uma popular rede de lojas escolhe cinco pequenas lojas (com 1000 clientes por mês) até megalojas (com 10 000 clientes por mês). A equipa observa o número de funcionários que trabalha em cada uma. Planeia usar a regressão para estimar quantos funcionários serão necessários noutras lojas.

Número de clientes (milhares)	1	4	6	9	10
Número de funcionários	24	30	46	47	53

Registemos os dados num gráfico, em que faremos da coordenada x o número de clientes (variável explicativa), enquanto o número de funcionários é marcado na coordenada y (variável resposta). É o número de clientes que explica o número de funcionários necessários e não o contrário. O número médio de clientes por loja é marcado como 6 (isto é 6000 clientes) e o número médio de funcionários é 40. A recta de regressão passa sempre pelo «ponto médio», aqui (6, 40). Existem fórmulas para calcular a recta de regressão, a linha que melhor se ajusta aos dados (também conhecida como recta dos mínimos quadrados). No nosso caso, a recta é $\hat{y} = 20{,}8 + 3{,}2x$, logo o declive é 3,2 e é positivo (subindo da esquerda para a direita). A recta cruza o eixo vertical y no ponto 20,8. O termo $\hat{y}$ é a estimativa do valor de y obtido da recta. Logo, se quisermos saber quantos funcionários devemos colocar numa loja que recebe 5000 clientes por mês, podemos substituir o valor $x = 5$ na equação da regressão e obter a estimativa $\hat{y} = 37$ funcionários, mostrando como a regressão tem objectivos muito práticos.



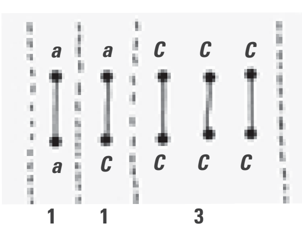
a ideia resumida
A interacção dos dados

37 Genética

A genética é um ramo da biologia, logo porque estará num livro de matemática? A resposta é que estes dois assuntos se fecundam e enriquecem um ao outro. Os problemas de genética requerem matemática, mas a genética também sugeriu novos ramos na álgebra. Gregor Mendel é central no tema da genética, o estudo da herança humana. As características hereditárias, tais como a cor dos olhos e do cabelo, o daltonismo, o destrimanismo ou canhotismo, ou o grupo sanguíneo, são todas determinadas por factores (alelos) de um gene. Mendel disse que estes factores passam de forma independente à geração seguinte.

Como pode a cor dos olhos ser transmitida à geração seguinte? No modelo básico há dois factores, *a* e C:

a é o factor olhos azuis
C é o factor olhos castanhos



População representando as proporções 1:1:3 dos genótipos *aa*, *a*C, CC

Nos indivíduos, os factores aparecem aos pares e dão origem aos possíveis genótipos *aa*, *a*C e CC (porque *a*C é o mesmo que C*a*). Qualquer pessoa tem um destes três genótipos, que determinam a cor dos seus olhos. Por exemplo, uma população pode consistir em um quinto de pessoas com o genótipo *aa*, outro quinto com o genótipo *a*C e os restantes três quintos com o genótipo CC. Em termos percentuais, estes genótipos serão 20%, 20% e 60% da população. Isto pode ser representado num diagrama mostrando esta proporção de genótipos.

O factor C, que representa olhos castanhos, é o factor dominante e *a*, a cor de olhos azul, o factor recessivo. Uma pessoa com genes puros de genótipo CC terá olhos castanhos, mas também os terá uma pessoa com os factores mistos,

Cronologia

1718
Abraham de Moivre publica a *Doctrine of Chances*

1865
Mendel propõe a existência de genes e leis de hereditariedade

ou seja, aqueles com um genótipo híbrido *a*C, porque C é dominante. Uma pessoa com genes puros de genótipo *aa* será a única a ter olhos azuis.

No início do século XIX, surgiu uma pergunta urgente no campo da biologia. Será que os olhos castanhos acabarão por assumir o controle e os azuis desaparecerão? Será que os olhos azuis se extinguirão? A resposta foi um retumbante «Não».

A lei de Hardy-Weinberg A resposta explica-se pela lei de Hardy-Weinberg, uma aplicação de matemática básica à genética que explica como, na teoria mendeliana da hereditariedade, um gene dominante não assumirá completamente o controle e um gene recessivo não morrerá.

G. H. Hardy era um matemático inglês que se orgulhava da não-aplicabilidade da matemática. Foi um grande investigador em matemática pura, mas é provavelmente mais conhecido pela sua única contribuição para a genética – que começou por ser um fragmento de matemática escrito no verso de um envelope depois de um jogo de *bridge*. Wilhelm Weinberg vem de uma área muito diferente. Médico de clínica geral na Alemanha, foi geneticista toda a sua vida. Descobriu a lei ao mesmo tempo que Hardy, por volta de 1908.

A lei relaciona-se com uma grande população na qual o acasalamento acontece aleatoriamente. Não há pares preferidos para que, por exemplo, pessoas de olhos azuis não prefiram acasalar com pessoas de olhos azuis. Depois do acasalamento, a criança recebe um factor de cada um dos progenitores. Por exemplo, um genótipo híbrido *a*C acasalando com um híbrido *a*C pode produzir qualquer um de *aa*, *a*C, CC, mas um *aa* acasalando com um CC só pode produzir um híbrido *a*C. Qual é a probabilidade de o factor *a* ser transmitido? Contando o número dos factores a, há dois factores *a* para cada genótipo *aa* e um factor *a* para cada genótipo *a*C, dando, em proporção, um total de três factores *a* em 10 (no nosso exemplo com as proporções 1:1:3 dos três genótipos). A probabilidade de a transmissão de um factor *a* estar incluída no genótipo de uma criança é portanto 3/10 ou 0,3. A probabilidade de a transmissão de um factor C estar incluída é 7/10 ou 0,7. A probabilidade de o genótipo *aa* estar incluído na geração seguinte, por exemplo, é portanto 0,3 × 0,3 = 0,09. O conjunto completo das probabilidades está resumido na tabela.

	a		C	
a	*aa*	0,3 × 0,3 = 0,09	*a*C	0,3 × 0,7 = 0,21
C	C*a*	0,3 × 0,7 = 0,21	CC	0,7 × 0,7 = 0,49

Os genótipos *a*C e C*a* são idênticos, logo a probabilidade de ocorrerem é 0, 21 + 0,21 = 0,42. Expressos em percentagens, os rácios dos genótipos *aa*, *a*C e CC na nova geração são 9%, 42% e 49%. Porque C é o factor dominante, 42% + 49% = 91% dos seres da primeira geração terão olhos

1908
Hardy e Weinberg mostram porque é que os genes dominantes não suplantam os genes recessivos

1918
Fisher reconcilia a teoria de Darwin com a teoria da hereditariedade mendeliana

1953
Descoberta da estrutura de dupla hélice do ADN

castanhos. Só um indivíduo com genótipo *aa* terá características observáveis do factor *a*, logo só 9% da população terão olhos azuis.

A distribuição inicial dos genótipos era 20%, 20% e 60% e na nova geração a distribuição de genótipos é 9%, 42% e 49%. O que acontece a seguir? Vamos ver o que acontece se nascer uma nova geração a partir desta por acasalamentos aleatórios. A proporção de factores *a* é 0,9 + ½ × 0,42 = 0,3, a proporção de factores C é ½ × 0,42 + 0,49 = 0,7. Estes são idênticos às anteriores transmissões de probabilidades dos factores *a* e C. A distribuição dos genótipos *aa*, *a*C e CC na geração seguinte é a mesma da geração anterior, e em particular o genótipo *aa* que dá os olhos azuis permanece estável nos 9% da população. As proporções sucessivas de genótipos durante uma sequência de acasalamentos aleatórios são portanto

$$20\%, 20\%, 60\% \rightarrow 9\%, 42\%, 49\% \rightarrow \ldots \rightarrow 9\%, 42\%, 49\%$$

Isto está de acordo com a lei de Hardy-Weinberg: depois de uma geração, as proporções do genótipo permanecem constantes de geração para geração, e as probabilidades de transmissão também.

O argumento de Hardy Para vermos a lei de Hardy-Weinberg a funcionar para qualquer população inicial, e não apenas para os 20%, 20% e 60% que seleccionámos no nosso exemplo, não podemos fazer melhor do que referir o argumento do próprio Hardy, escrito ao editor do jornal americano *Science* em 1908.

Hardy começa com a distribuição inicial de genótipos *aa*, *a*C e CC como p, $2r$ e q e as probabilidades de transmissão $p + r$ e $r + q$. No nosso exemplo numérico (de 20%, 20%, 60%), $p = 0{,}2$, $2r = 0{,}2$ e $q = 0{,}6$. As probabilidades de transmissão dos factores *a* e C são $p + r = 0{,}2 + 0{,}1 = 0{,}3$ e $r + q = 0{,}1 + 0{,}6 = 0{,}7$.
E se houver uma distribuição inicial diferente dos genótipos *aa*, *a*C e CC e começarmos com, digamos, 10%, 60% e 30%? Como funciona a lei de Hardy-Weinberg neste caso? Aqui teremos $p = 0{,}1$, $2r = 0{,}6$ e $q = 0{,}3$, e a probabilidade de transmissão dos factores *a* e C é respectivamente $p + r = 0{,}4$ e $r + q = 0{,}6$. Assim, a distribuição da geração seguinte de genótipos é 16%, 48% e 36%. As proporções sucessivas dos genótipos *aa*, *a*C e CC depois de acasalamentos aleatórios são

$$10\%, 60\%, 30\% \rightarrow 16\%, 48\%, 36\% \rightarrow \ldots \rightarrow 16\%, 48\%, 36\%$$

e as proporções estabilizam depois de uma geração, como antes, e as probabilidades das transmissões de 0,4 e 0,6 permanecem constantes.

Com estes números, 16% da população terão os olhos azuis e 48% + 36% = 84% terão olhos castanhos, porque C é dominante no genótipo *a*C.

Assim, a lei de Hardy-Weinberg implica que estas proporções de genótipos *aa*, *a*C e CC permanecem constantes de geração para geração qualquer que seja a distribuição inicial dos factores na população. O gene dominante C não assume o controle e as proporções dos genótipos são intrinsecamente estáveis.

Hardy sublinhou que o seu modelo era apenas aproximado. A sua simplicidade e elegância dependiam de muitas assunções que não se mantêm na vida real. No modelo, a probabilidade de mutação genética ou de modificações nos próprios genes não foi considerada, e a consequência de a transmissão de proporções ser constante significa que nada tem a dizer acerca da evolução. Na vida real há «desvios genéticos» e as probabilidades de transmissão dos factores não se mantêm constantes. Isto provocará variações nas proporções e o desenvolvimento de novas espécies.

A lei de Hardy-Weinberg juntou a teoria de Mendel – a «teoria quântica» da genética – e o darwinismo e selecção natural de uma forma intrínseca. Foi preciso esperar pelo génio de R. A. Fisher para reconciliar a teoria da hereditariedade mendeliana com a teoria contínua em que as características evoluem.

O que faltou na ciência da genética até à década de 1950 foi o entendimento físico do próprio material genético. Deu-se então um avanço notável com as contribuições de Francis Crick, James Watson, Maurice Wilkins e Rosalind Franklin. O meio foi o ácido desoxirribonucleico ou ADN. A matemática é necessária para modelar a famosa dupla hélice (ou par de espirais enroladas à volta de um cilindro). Os genes localizam-se em segmentos dessa dupla hélice.

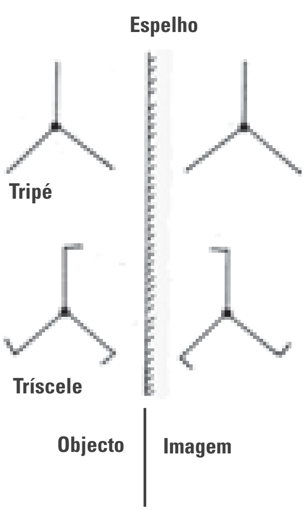
A matemática é indispensável no estudo da genética. A partir da geometria básica das espirais de ADN e da potencialidade sofisticada da lei de Hardy-Weinberg, desenvolveram-se modelos matemáticos que trabalham com muitas características (não só a cor dos olhos), incluindo diferenças entre homem e mulher e também acasalamentos não aleatórios. A ciência da genética sempre retribuiu à matemática sugerindo novos ramos de álgebra abstracta de interesse para as suas intrigantes propriedades matemáticas.

a ideia resumida
Incerteza na piscina dos genes

38 Grupos

Evariste Galois morreu num duelo com a idade de 20 anos, mas deixou ideias que chegam para manter os matemáticos ocupados durante séculos. Estas envolviam a teoria dos grupos, construções matemáticas que podem ser usadas para quantificar simetrias. Para lá da sua atracção artística, a simetria é um ingrediente essencial para os cientistas que sonham com uma futura teoria de tudo. A teoria dos grupos é a cola que une o «tudo» num só.

A simetria está por todo o lado à nossa volta. Nos vasos gregos, nos cristais de neve, muitas vezes nos edifícios e nalgumas letras do alfabeto. Há muitos tipos de simetria: as principais são a simetria de espelho e a simetria de rotação. Só analisaremos a simetria em duas dimensões – todos os nossos objectos de estudo vivem na superfície plana desta página.



Simetria de espelho Poderemos construir um espelho tal que um objecto pareça o mesmo em frente do espelho como no espelho? A palavra MIM tem simetria de espelho mas HEM não; MIM em frente do espelho é o mesmo MIM no espelho, enquanto HEM se torna MEH. Um tripé tem simetria de espelho, mas o tríscele (tripé com pernas) não. O tríscele como objecto à frente do espelho é destro, mas na sua imagem no espelho, na chamada imagem plana, é canhoto.

Simetria rotacional Também podemos perguntar se existe um eixo perpendicular à página, tal que o objecto possa ser rodado na página de um ângulo e ser trazido de volta à sua posição. Tanto o tripé como o tríscele têm simetria rotacional. O tríscele, que significa «com três pernas», é uma forma interessante. A sua versão destra é o símbolo da ilha de Man e também aparece na bandeira da Sicília.

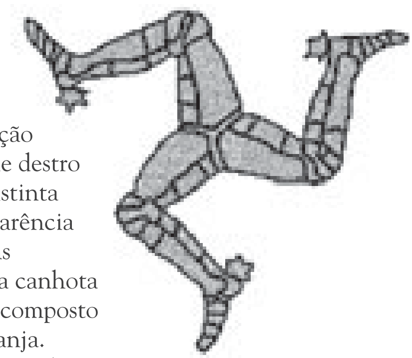
Cronologia

1832
Galois propõe a ideia de grupos de permutações

1854
Cayley tenta generalizar o conceito de grupo

1872
Felix Klein começa um programa de classificação da geometria usando grupos

Se a rodarmos 120 graus ou 240 graus, a figura rodada coincidirá consigo própria; se fecharmos os olhos antes de rodarmos, veremos o mesmo tríscele quando os voltarmos a abrir depois da rotação.



O tríscele da ilha de Man

O curioso acerca da figura com três pernas é que nenhuma rotação mantendo a figura no plano transformará alguma vez um tríscele destro num tríscele canhoto. Os objectos cuja imagem no espelho é distinta do objecto em frente do espelho são chamados quirais – têm aparência semelhante, mas não são iguais. A estrutura molecular de alguns compostos químicos tanto pode ter a forma destra como a forma canhota em três dimensões e ser exemplo de objecto quiral. É o caso do composto limoseno, que, numa forma, sabe a limão e, na outra, sabe a laranja. O medicamento talidomida, numa das suas formas, é eficaz na cura dos enjoos matinais da gravidez, mas, na outra forma, tem consequências trágicas.

Medir a simetria No caso do nosso tríscele, as operações básicas de simetria são as rotações (no sentido dos ponteiros do relógios) R de 120 graus e S de 240 graus. A transformação I é a que roda de 360 graus ou, alternativamente, não faz coisa nenhuma. Podemos criar uma tabela baseada nas combinações destas rotações, da mesma forma que podemos criar uma tabela de multiplicação.

Esta tabela é como uma tabela vulgar de multiplicação com números, com a diferença de estarmos a «multiplicar» símbolos. De acordo com a convenção mais usada, a multiplicação $R \circ S$ significa rodar primeiro o tríscele 240 graus com S e depois 120 graus com R, sendo o resultado uma rotação de 360 graus, como se não tivéssemos feito nada. Isto pode ser expresso como $R \circ S = I$, o resultado encontrado no cruzamento da penúltima linha com a última coluna da tabela.

$\circ$	I	R	S
I	I	R	S
R	R	S	I
S	S	I	R

A tabela de Cayley para o grupo de simetria do tríscele

O grupo de simetria do tríscele é constituído por I, R e S e a tabela de multiplicação de como os combinar. Como o grupo contém três elementos, o seu tamanho (ou «ordem») é três. A tabela também é chamada tabela de Cayley (recebeu o nome do matemático Arthur Cayley, primo distante de Sir George Cayley, um pioneiro da aviação).

1891

Evgraf Fedorov e Arthur Schönflies classificam independentemente os 230 grupos cristalográficos

1983

A classificação de grupos simples finitos é completada e o enorme teorema, provado

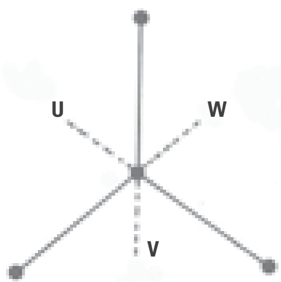
∘	I	R	S	U	V	W
I	I	R	S	U	V	W
R	R	S	I	V	W	U
S	S	I	R	W	U	V
U	U	W	V	I	S	R
V	V	U	W	R	I	S
W	W	V	U	S	R	I

A tabela de Cayley para o grupo de simetria do tripé

Como o tríscele, o tripé tem simetria rotacional. Mas também tem simetria de espelho e portanto tem um grupo de simetria maior. Chamaremos U, V e W às reflexões nos três eixos de espelho.

O grupo de simetria maior do tripé, que é de ordem 6, é composto de seis transformações I, R, S, U, V e W e tem a tabela de multiplicação mostrada.

Uma transformação interessante é conseguida combinando duas reflexões em eixos diferentes, tal como $U \circ W$ (em que a reflexão W é aplicada primeiro e seguida pela reflexão U). Esta é na realidade a rotação do tripé de 120 graus, simbolicamente $U \circ W = R$. A combinação das reflexões na ordem inversa $W \circ U = S$, dá uma rotação de 240 graus. Em particular $U \circ W \neq W \circ U$. Esta é uma enorme diferença entre uma tabela de multiplicação de um grupo e uma tabela da multiplicação vulgar com números.



Reflexões de um tripé

Um grupo no qual a ordem da combinação dos elementos é imaterial chama-se grupo abeliano, por ter recebido o nome do matemático norueguês Niels Abel. O grupo de simetria do tripé é o menor grupo que não é abeliano.

Grupos abstractos A tendência na álgebra no século XX tem sido a abstracção, em que um grupo é definido por algumas regras básicas conhecidas por axiomas. Com esta visão, o grupo de simetria do triângulo torna-se apenas um exemplo de um sistema abstracto. Existem sistemas em álgebra que são mais básicos do que um grupo e requerem menos axiomas; outros são mais complexos e requerem mais axiomas. No entanto, o conceito de grupo é correcto e é o mais importante sistema algébrico de todos. É notável que tenha emergido um corpo de conhecimento tão grande de tão poucos axiomas. A vantagem do método abstracto é que os teoremas gerais podem ser deduzidos para todos os grupos e aplicados, se tal for necessário, nalguns específicos.

Uma característica da teoria dos grupos é a de poderem existir pequenos grupos dentro de grupos maiores. O grupo de simetria do tríscele de ordem três é um subgrupo do grupo de simetria do tripé de ordem seis. J. L. Lagrange provou um facto básico sobre subgrupos. O teorema de Lagrange diz que a ordem do subgrupo deve dividir a ordem do grupo. Assim, sabemos automaticamente que o grupo de simetria do tripé não tem subgrupos de ordem quatro ou cinco.

Classificar grupos Houve um extenso programa para classificar todos os grupos finitos possíveis. Não há necessidade de os listar todos, porque alguns grupos são construídos a partir de outros básicos, e é desses básicos que precisamos. O princípio de classificação é quase o mesmo que na química, em que o interesse está focado nos elementos químicos básicos e não nos compostos que se podem fazer com eles. O grupo de simetria do tripé de seis elementos é um «composto» constituído pelo grupo das rotações (de ordem três) e reflexões (ordem dois).

Axiomas de um grupo

Uma colecção de elementos G com «multiplicação» $\circ$ é um grupo se

1. Há um elemento 1 em G, tal que $1 \circ a = a \circ 1 = a$ para todos os elementos a no grupo G (o elemento especial 1 é chamado elemento identidade).

2. Para cada elemento a em G existe um elemento $\bar{a}$ em G com $\bar{a} \circ a = a \circ \bar{a} = 1$ (o elemento $\bar{a}$ é chamado elemento inverso de a).

3. Para todos os elementos a, b e c em G é verdade que $a \circ (b \circ c) = (a \circ b) \circ c$ (a chamada lei associativa).

Praticamente todos os grupos básicos podem ser classificados em classes conhecidas. A classificação completa, chamada «o enorme teorema», foi anunciada por Daniel Gorenstein em 1983 e concluída pelo trabalho acumulado ao longo de 30 anos meritórios de pesquisa e publicações pelos matemáticos. É um atlas de todos os grupos conhecidos. Os grupos básicos caem num de quatro tipos principais, embora se tivessem encontrado 26 grupos que não estão em nenhuma das categorias. São conhecidos como grupos esporádicos.

Os grupos esporádicos são dissidentes e tipicamente de ordem muito alta. Cinco dos menores eram conhecidos de Emile Mathieu na década de 1860, mas muita da actividade moderna teve lugar entre 1965 e 1975. O menor grupo esporádico é de ordem $7920 = 2^4 \times 3^2 \times 5 \times 11$, mas no extremo superior estão o «bebé monstro» e o puro «monstro» com a ordem $2^{46} \times 3^{20} \times 5^9 \times 7^6 \times 11^2 \times 13^3 \times 17 \times 19 \times 23 \times 29 \times 31 \times 41 \times 47 \times 59 \times 71$, o que em numeração decimal é cerca de 8×10^{53} ou, se quiserem, um 8 seguido de 53 zeros – um número realmente muito grande. Pode mostrar-se que 20 dos 26 grupos esporádicos são representados como subgrupos dentro do «monstro» – os seis grupos que desafiam todos os sistemas de classificação são conhecidos como os «seis párias».

Ainda que em matemática se procurem provas curtas e duras, a prova da classificação dos grupos finitos é qualquer coisa como 10 000 páginas de símbolos fortemente intrincados. O progresso em matemática nem sempre se deve ao trabalho de um génio único e excepcional.

a ideia resumida
Medir a simetria

39 Matrizes

Esta é a história da «álgebra extraordinária», uma revolução na matemática que teve lugar em meados do século XIX. Há séculos que os matemáticos brincavam com blocos de números, mas a ideia de tratar os blocos como um número único arrancou há 150 anos com um pequeno grupo de matemáticos que reconheceu o seu potencial.

A álgebra ordinária é a álgebra tradicional em que símbolos como *a*, *b*, *c*, *x* e *y* representam números simples. Muitas pessoas acham difícil de compreender, mas para os matemáticos foi um grande passo em frente. Em comparação, a «álgebra extraordinária» gerou uma mudança sísmica. Para aplicações sofisticadas, este progresso de uma álgebra unidimensional para uma álgebra multidimensional revelar-se-ia incrivelmente importante.

Números multidimensionais Na álgebra ordinária, *a* pode significar um número como 7, e escreveremos $a = 7$, mas na teoria das matrizes uma matriz A será um «número multidimensional», por exemplo o bloco

$$A = \begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix}$$

A matriz tem três linhas e quatro colunas (é uma matriz «3 por 4»), mas em princípio podemos ter matrizes com qualquer número de linhas e de colunas – mesmo uma matriz gigante «100 por 200», com 100 linhas e 200 colunas. Uma vantagem crítica da álgebra de matrizes é podermos pensar numa vasta colecção de números, como um conjunto de dados em estatística, como uma única entidade. Mais do que isto, podemos manipular estes blocos de forma simples e eficiente. Se quisermos somar ou multiplicar todos os números de dois conjuntos de dados, cada um consistindo em 1000 números, não temos de executar 1000 cálculos – só temos de executar um (somar ou multiplicar as duas matrizes).

Cronologia

200 a.C.
Os matemáticos chineses usam matrizes de números

1850
J. J. Sylvester introduz o termo «matriz»

1858
Cayley publica *Memoir on the Theory of Matrices*

Um exemplo prático Suponhamos que a matriz A representa a produção da empresa AJAX numa semana. A empresa AJAX tem três fábricas localizadas em diferentes sítios no país e a sua produção é medida em unidades (digamos milhares de produtos) dos quatro produtos que produz. No nosso exemplo, as quantidades, correspondendo com a nossa matriz A, são

	produto 1	produto 2	produto 3	produto 4
fábrica 1	7	5	0	1
fábrica 2	0	4	3	7
fábrica 3	3	2	0	2

Na semana seguinte, a produção programada pode ser diferente, mas pode ser escrita noutra matriz *B*. Por exemplo, *B* pode ser dada por

$$B=\begin{pmatrix} 9 & 4 & 1 & 0 \\ 0 & 5 & 1 & 8 \\ 4 & 1 & 1 & 0 \end{pmatrix}.$$

Qual é o total de produção em ambas as semanas? O teórico das matrizes dirá que é a matriz $A + B$, em que os números correspondentes são somados,

$$A+B=\begin{pmatrix} 7+9 & 5+4 & 0+1 & 1+0 \\ 0+0 & 4+5 & 3+1 & 7+8 \\ 3+4 & 2+1 & 0+1 & 2+0 \end{pmatrix}=\begin{pmatrix} 16 & 9 & 1 & 1 \\ 0 & 9 & 4 & 15 \\ 7 & 3 & 1 & 2 \end{pmatrix}$$

Simples. Infelizmente, a multiplicação de matrizes é menos óbvia. Voltando à empresa AJAX, suponhamos que o lucro unitário dos seus produtos é **3**, **9**, **8**, **2**. Podemos com certeza calcular o lucro total da fábrica 1 com a produção 7, 5, 0, 1 dos seus quatro produtos. Resolve-se como $7 \times \mathbf{3} + 5 \times \mathbf{9} + 0 \times \mathbf{8} + 1 \times \mathbf{2} = 68$.

Mas, em vez de tratarmos apenas uma fábrica, podemos com a mesma facilidade calcular os lucros totais *T* de *todas* as fábricas.

$$T=\begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix}\times\begin{pmatrix} \mathbf{3} \\ \mathbf{9} \\ \mathbf{8} \\ \mathbf{2} \end{pmatrix}=\begin{pmatrix} 7\times\mathbf{3}+5\times\mathbf{9}+0\times\mathbf{8}+1\times\mathbf{2} \\ 0\times\mathbf{3}+4\times\mathbf{9}+3\times\mathbf{8}+7\times\mathbf{2} \\ 3\times\mathbf{3}+2\times\mathbf{9}+0\times\mathbf{8}+2\times\mathbf{2} \end{pmatrix}=\begin{pmatrix} 68 \\ 74 \\ 31 \end{pmatrix}$$

1878

Greg Frobenius prova alguns resultados cruciais da álgebra das matrizes

1925

Heisenberg usa mecânica matricial na teoria quântica

Se observarmos com atenção, veremos a multiplicação da *linha* pela *coluna*, uma característica essencial da multiplicação de matrizes, Se além disso nos forem dados os *volumes* unitários **7**, **4**, **1**, **5** de cada unidade de produtos, podemos calcular de uma assentada os lucros e as necessidades de armazenamento das três fábricas numa única multiplicação de matrizes:

$$\begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix} \times \begin{pmatrix} \mathbf{3} & \mathbf{7} \\ \mathbf{9} & \mathbf{4} \\ \mathbf{8} & \mathbf{1} \\ \mathbf{2} & \mathbf{5} \end{pmatrix} = \begin{pmatrix} 68 & 74 \\ 74 & 54 \\ 31 & 39 \end{pmatrix}$$

O armazenamento total é fornecido pela segunda coluna da matriz resultante, ou seja 74, 54 e 39. A teoria das matrizes é muito influente. Imaginemos uma empresa com centenas de fábricas, milhares de produtos, e diferentes lucros unitários e necessidades de armazenagem em semanas diferentes. Com a álgebra das matrizes, os cálculos e a nossa compreensão, são relativamente imediatos, sem necessidade de preocupação com os detalhes que estão a ser tratados.

Álgebra das matrizes *vs.* álgebra ordinária Há muitos paralelos a fazer entre a álgebra das matrizes e a álgebra ordinária. A diferença mais notória ocorre na multiplicação de matrizes. Se multiplicarmos a matriz *A* pela matriz *B* e depois tentarmos o inverso:

$$A \times B = \begin{pmatrix} 3 & 5 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 7 & 6 \\ 4 & 8 \end{pmatrix} = \begin{pmatrix} 3\times7+5\times4 & 3\times6+5\times8 \\ 2\times7+1\times4 & 2\times6+1\times8 \end{pmatrix} = \begin{pmatrix} 41 & 58 \\ 18 & 20 \end{pmatrix}$$

$$B \times A = \begin{pmatrix} 7 & 6 \\ 4 & 8 \end{pmatrix} \times \begin{pmatrix} 3 & 5 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 7\times3+6\times2 & 7\times5+6\times1 \\ 4\times3+8\times2 & 4\times5+8\times1 \end{pmatrix} = \begin{pmatrix} 33 & 41 \\ 28 & 28 \end{pmatrix}$$

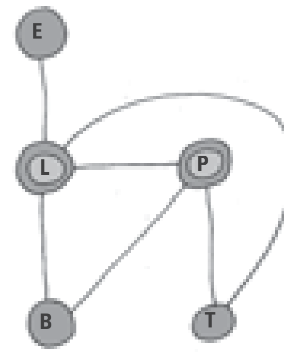
Assim, na álgebra de matrizes temos que $A \times B$ e $B \times A$ são diferentes, uma situação que não se levanta na álgebra ordinária, em que a ordem da multiplicação de dois números não tem influência no resultado.

Outra situação ocorre com os inversos. Na álgebra ordinária, os inversos podem ser facilmente calculados. Se $a = 7$, o seu inverso é $1/7$, porque tem a propriedade $1/7 \times 7 = 1$. Por vezes, escrevemos este inverso como $a^{-1} = 1/7$ e temos $a^{-1} \times a = 1$.

Um exemplo na teoria das matrizes é $A = \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}$ e podemos verificar que $A^{-1} = \begin{pmatrix} 7 & -2 \\ -3 & 1 \end{pmatrix}$

porque $A^{-1} \times A = \begin{pmatrix} 7 & -2 \\ -3 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

em que $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ é chamada a matriz identidade e é a matriz correspondente ao 1 na álgebra ordinária. Em álgebra ordinária, só o 0 não tem inverso, mas em álgebra das matrizes muitas matrizes não têm inversos.



Planos de viagens Outro exemplo da utilização das matrizes é a análise da rede de voos das companhias aéreas. Isto envolverá aeroportos centrais e aeroportos mais pequenos. Na prática, pode envolver centenas de destinos – analisemos um pequeno exemplo: os aeroportos centrais **L**ondres (**L**) e **P**aris (**P**), e os mais pequenos **E**dimburgo (**E**), **B**ordéus (**B**) e **T**oulouse (**T**) e a rede mostrando os voos directos possíveis. Para uma análise informática dessas redes, primeiro elas têm de ser codificadas com o recurso a matrizes. Se há um voo directo entre dois aeroportos, regista-se um 1 na intersecção da linha e da coluna respeitantes a esses aeroportos (como **L**ondres e **E**dimburgo). A matriz da «conectividade» que descreve a rede é A.

A submatriz mais baixa (marcada a tracejado) mostra que não há ligações directas entre os três aeroportos menores. O produto da matriz $A \times A = A^2$ da matriz por si própria pode ser interpretado dando o número de possíveis itinerários entre dois aeroportos com *exactamente um transbordo*. Assim, por exemplo, há três possíveis viagens para Paris via outras cidades, mas não há viagens de Londres para Edimburgo que envolvam transbordo. O número de itinerários que são directos ou envolvem um transbordo são os elementos da matriz $A + A^2$. Este é outro exemplo da capacidade das matrizes para capturarem a essência dum vasto conjunto de dados sob um só cálculo.

Quando um pequeno grupo de matemáticos criou a teoria das matrizes na década de 1850, o objectivo foi resolver problemas de matemática pura. De uma perspectiva aplicada, a teoria das matrizes era uma «solução à procura de um problema». Como tão frequentemente acontece, apareceram mesmo «problemas» que precisavam da teoria que começava a nascer. Uma primeira aplicação ocorreu na década de 1920, quando Werner Heisenberg investigou «mecânica matricial», uma parte da teoria quântica. Outra pioneira foi Olga Taussky-Todd, que trabalhou durante algum tempo num projecto de aeronaves e usou álgebra matricial. Quando lhe perguntaram como tinha descoberto aquele campo, Taussky-Todd respondeu que era ao contrário: a teoria das matrizes tinha-a encontrado a ela. Assim é o jogo matemático.

a ideia resumida
Combinar blocos de números

40 Códigos

O que tem Júlio César em comum com a transmissão de sinais digitais modernos? A resposta curta é códigos e codificação. Para enviar sinais digitais para um computador ou um televisor digital, é essencial a codificação das imagens e da voz numa corrente de zeros e uns – um código binário –, porque é a única linguagem que esses dispositivos compreendem. César utilizava códigos para comunicar com os seus generais e mantinha as suas mensagens secretas trocando as letras da mensagem de lugar, segundo uma chave que só ele e os seus generais conheciam.

A exactidão era essencial para César e também é necessária para a transmissão eficaz de sinais digitais. César também queria manter os códigos só para si, tal como as companhias de transmissão de sinal digital de televisão querem que só os assinantes que pagam sejam capazes de interpretar os seus sinais.

Comecemos por analisar a exactidão. O erro humano ou «o ruído» podem sempre ocorrer, e têm de ser tratados. O raciocínio matemático permite-nos construir sistemas de codificação que detectam automaticamente erros e até fazem correcções.

Detecção e correcção de erros Um dos primeiros códigos binários do mundo foi o código Morse, que utiliza dois símbolos, o ponto • e o traço –. O inventor norte-americano Samuel F. B. Morse enviou a primeira mensagem usando este código de Washington para Baltimore em 1844. Foi um código concebido para telégrafos eléctricos em meados do século, sem grandes preocupações de eficiência de concepção. No código Morse, a letra A é codificada como • –, a *B* como – •••, a C como – • – • e as outras letras têm diferentes sequências de pontos e traços. Se o operador do telégrafo desejasse enviar a mensagem «CAB», enviaria o conjunto – • – • / • – / – •••. A despeito dos seus muitos méritos, o código Morse não era bom a detectar erros,

Cronologia

55 a.C.
Júlio César invade a Bretanha e usa códigos para comunicar com os seus generais

cerca de 1750
O teorema de Euler fornece as bases à criptografia de chave pública

muito menos a corrigi-los. Se o operador quisesse digitar «CAB», mas pusesse um ponto em vez de um traço no C, esquecesse o traço no A e o ruído na linha substituísse um traço por um ponto em *B*, o receptor leria •• – • / • / – – ••, e não veria nada de errado na sua interpretação de «FEZ».

A um nível mais primitivo, poderíamos analisar um sistema de codificação consistindo apenas em 0 e 1, em que 0 representa uma palavra e 1 outra. Imaginemos que o comandante de uma armada tem de transmitir uma mensagem para as suas tropas, que tanto pode ser «invadir» como «não invadir». A instrução «invadir» é codificada com «1» e a instrução «não invadir» é codificada com «0». Se o 1 ou o 0 fossem incorrectamente transmitidos, o receptor nunca o saberia – e seria dada a instrução errada, com consequências desastrosas.

Podemos melhorar a situação usando palavras de código de comprimento duplo. Se desta vez codificarmos «invadir» com o código 11 e o «não invadir» com 00, será melhor. Um erro num dígito resultaria na recepção de 01 ou 10. Como só o 11 ou o 00 são palavras de código legítimas, o receptor saberia de certeza que tinha sido cometido um erro. A vantagem deste sistema é que seria possível detectar erros, mas continuaríamos sem saber como os corrigir. Se recebermos 01, como saberemos se tinha sido enviado 00 ou 11?

O caminho para um sistema melhor é a combinação de desenhos com palavras de código maiores. Se codificarmos «invadir» com 111 e «não invadir» com 000, um erro num dígito seria decerto detectado como anteriormente. Se soubéssemos que podia ser cometido no máximo um erro (uma assunção razoável, dado que a probabilidade de dois erros numa palavra de código é pequena), a correcção poderia de facto ser feita pelo receptor. Por exemplo, se se receber 110, a mensagem correcta teria sido 111. Com as nossas regras, não podia ser 000, dado que essa palavra de código teria dois erros relativamente a 110. Neste sistema há apenas duas palavras de código 000 e 111, mas estão suficientemente longe para tornarem possível a detecção e a correcção do erro.

É este princípio que é usado quando o processador de texto está no modo de autocorrecção. Se escrevermos «animul», o processador de texto detecta o erro e corrige-o pela palavra mais próxima, «animal». A língua inglesa não é completamente corrigível, porque, se escrevermos «*lomp*», não há uma única

1844
Morse transmite a primeira mensagem usando o seu código

anos 1920
Desenvolvimento da máquina Enigma

1950
Richard Hamming publica um artigo crucial sobre detecção e correcção de erros

anos 1970
Desenvolvimento da criptografia de chave pública

palavra mais próxima; as palavras *lamp*, *limp*, *lump*, *pomp* e *romp* distam todas de *lomp* em apenas uma letra.

Um código binário moderno consiste em palavras codificadas que são blocos consistindo em zeros e uns. Ao escolhermos palavras codificadas legítimas suficientemente apartadas, a detecção e a correcção são possíveis. As palavras codificadas do código de Morse são demasiado próximas, mas os sistemas de codificação modernos usados para transmitir dados de satélites podem sempre ser colocados em modo de autocorrecção. As palavras codificadas longas com alto desempenho em termos de correcção de erros levam mais tempo a transmitir, logo há um compromisso entre o comprimento e a velocidade de transmissão. As viagens espaciais da NASA têm usado códigos com correcção de três erros que se têm revelado satisfatórios no combate ao ruído na linha.

Tornar as mensagens secretas Júlio César mantinha as suas mensagens secretas trocando as letras da mensagem de acordo com uma chave que só ele e os seus generais conheciam. Se a chave caísse nas mãos erradas, a mensagem podia ser decifrada pelos seus inimigos. Na Idade Média, a rainha Mary Stuart da Escócia mandou mensagens em código da sua cela na prisão. A rainha tinha em mente derrubar a sua prima, a rainha Isabel I, mas as suas mensagens codificadas foram interceptadas. Mais sofisticado que o método romano de distribuir as letras por uma chave, os seus códigos eram baseados na substituição, que podia contudo ser facilmente revelada pela análise da frequência das letras e símbolos usados. Durante a Segunda Guerra Mundial, o código alemão Enigma foi quebrado pela descoberta da sua chave. Neste caso, o desafio formidável era enorme, mas o código estava sempre vulnerável porque a chave era transmitida como parte da mensagem.

Um desenvolvimento sensacional na encriptação de mensagens foi descoberto na década de 1970. Contrariamente a tudo o que se acreditava, dizia-se que a chave secreta podia ser transmitida a todos e, apesar disso, a mensagem ser mantida completamente segura. É a chamada criptografia de chave pública. O método depende de um teorema com 200 anos de um ramo da matemática glorificado por ser o mais inútil de todos.

Encriptação de chave pública John Emissor, um agente secreto conhecido no mundo da espionagem como «J», tinha acabado de chegar à cidade e queria mandar ao seu superior Dr. Rodney Receptor uma mensagem secreta, para anunciar a sua chegada. O que fez é bastante curioso. Foi a uma biblioteca pública, tirou uma lista telefónica da cidade da prateleira e procurou Dr. R. Receptor. Encontrou dois números junto do nome Receptor – um longo, 247, e um curto, 5. Esta informação estava disponível para toda a gente, e era toda a informação de que John Emissor precisava para encriptar a sua mensagem, que era, para simplicidade, o seu cartão telefónico, J. Esta letra é o número 74 numa lista de palavras, também publicamente disponível.

Emissor encripta 74 calculando 74^5 (módulo 247), ou seja, quer saber o resto quando se divide 74^5 por 247. Calcular 74^5 é possível com uma calculadora simples, mas tem de ser feito exactamente:

$$74^5 = 74 \times 74 \times 74 \times 74 \times 74 = 2\ 219\ 006\ 624$$

e

$$2\ 219\ 006\ 624 = 8\ 983\ 832 \times 247 + 120$$

logo, dividindo este número enorme por 247, obtém o resto 120. A mensagem encriptada de Emissor é 120, e é isso que ele transmite a Receptor. Dado que os números 247 e 5 estão disponíveis para todos, qualquer um podia encriptar a mensagem. Mas nem todos a podiam desencriptar. O Dr. Receptor tem mais informação na manga. Chegou ao seu número pessoal 247 multiplicando dois números primos. Neste caso, obteve o número 247 multiplicando $p = 13$ e $q = 19$, mas só ele o sabe.

É aqui que o velho teorema atribuído a Leonhard Euler é recuperado e utilizado. O Dr R. Receptor usa o conhecimento de $p = 13$ e $q = 19$ para encontrar um valor de a em que $5 \times a \equiv 1$ módulo $(p - 1)\ (q - 1)$ em que o símbolo $\equiv$ significa igual, em módulo aritmético. Qual é o a, tal que dividindo $5 \times a$ por $12 \times 18 = 216$ dá resto 1? Saltando os cálculos, ele chega a $a = 173$.

Dado que é o único que conhece os números primos p e q, o Dr. Receptor é o único que pode calcular o número 173. Com ele determina o resto, quando divide o número enorme 120^{173} por 247. Isto está fora das capacidades de uma calculadora simples, mas é facilmente determinado por um computador. A resposta é 74, como Euler sabia há já 200 anos. Com esta informação, Receptor procura a palavra 74 e vê que J chegou à cidade.

Podemos dizer, decerto, que um *hacker* podia descobrir o facto que $247 = 13 \times 19$ e que o código podia ser quebrado. Estaríamos correctos. Mas o princípio da encriptação e da decriptação é o mesmo se o Dr. Receptor tivesse usado outro número em vez de 247. Podia escolher dois números primos muito grandes e multiplicá-los para obter um número muito maior que 247.

Encontrar os dois factores primos de um número muito grande é virtualmente impossível – quais são os factores de 24 812 789 922 307, por exemplo? Mas também podem ser escolhidos números muito maiores do que este. O sistema de chave pública é seguro e, se o poder de supercomputadores a trabalhar em conjunto conseguir factorizar um número de encriptação, tudo o que o Dr. Receptor terá de fazer é aumentar ainda mais o seu tamanho. Afinal, é consideravelmente mais fácil para o Dr. Receptor «misturar caixas de areia preta com caixas de areia branca» do que para qualquer *hacker* voltar a dividir as areias pelas respectivas caixas.

a ideia resumida
Manter as mensagens secretas

41 Contagem avançada

O ramo da matemática chamado análise combinatória é muitas vezes conhecido por contagem avançada. Não tem que ver com somar uma coluna de números de cabeça. Além de «quanto é?», a pergunta é «como podem os objectos ser combinados?». Os problemas são muitas vezes enunciados de forma simples, sem a pesada superestrutura da teoria matemática – não é necessário conhecer uma grande quantidade de trabalhos preliminares antes de arregaçar as mangas. Isto torna os problemas de análise combinatória interessantes. Contudo, estes deviam vir acompanhados de um aviso de saúde: podem causar dependência e insónias.

Uma história de St Ives As crianças podem começar a fazer análise combinatória em tenra idade. Uma canção de embalar inglesa tradicional coloca um problema de análise combinatória:

Ia eu a caminho de St Ives,
Encontrei um homem com sete esposas;
Cada esposa tinha sete sacos,
Cada saco tinha sete gatos,
Cada gato tinha sete gatinhos.
Gatinhos, gatos, sacos e esposas,
Quantos iam a caminho de St Ives?

A última linha contém a pergunta traiçoeira. Uma assunção implícita é que o narrador é a única pessoa a ir «a caminho de» St Ives, logo a resposta é «um». Alguns excluem o narrador e, para eles, a resposta é «nenhum».

Cronologia

cerca de 1800 a.C.
Escrita do papiro de Rhind no Egipto

cerca de 1100
Bhaskara trata permutações e combinações

O encanto do poema está na sua ambiguidade e nas várias questões que pode gerar. Podemos perguntar: Quantos vinham *de* St Ives? A interpretação é novamente importante. Podemos ter a certeza de que o homem com as suas sete mulheres está a *afastar-se* de St Ives? As esposas acompanhavam o homem quando se deu o encontro ou estavam noutro local qualquer? O primeiro requisito dum problema de análise combinatória são as assunções previamente acordadas.

Assumamos que a comitiva viajava na única estrada que saía da cidade marítima de St Ives e que os «gatinhos, gatos, sacos e esposas» estavam todos presentes. Quantos vinham de St Ives? A tabela seguinte dá-nos a solução.

homem	1	1
esposas	7	7
sacos	7 × 7	49
gatos	7 ×7 × 7	343
gatinhos	7 × 7 × 7 × 7	2401
Total		**2801**

Em 1858 o antiquário escocês Alexander Rhind visitou Luxor e deparou com um papiro de 5 metros de comprimento cheio de matemática egípcia de cerca de 1800 a.C. Comprou-o. Alguns anos mais tarde, o papiro foi adquirido pelo Museu Britânico e os seus hieróglifos traduzidos. O problema 79 do papiro de Rhind é um problema de casas, gatos, ratos e trigo muito semelhante ao dos gatinhos, gatos, sacos e esposas de St Ives. Ambos envolvem potências de 7 e o mesmo tipo de análise. Parece que a análise combinatória tem uma longa história.

Números factoriais O problema das filas introduz-nos à primeira arma da análise combinatória – o *factorial* de um número. Suponhamos que **A**lan, **B**rian, **C**harlotte, **D**avid e **E**llie formam uma fila

E C A B D

com Ellie à cabeça da fila, seguida de Charlotte, Alan, Brian e David no final. Trocando as pessoas de lugar, formam-se outras filas; quantas filas diferentes são possíveis?

A arte da contagem neste problema depende da escolha. Há 5 escolhas para quem vai ser colocado no primeiro lugar da fila, quatro para o segundo lugar, etc. Quando chegamos à última posição, não há nenhuma escolha a fazer, porque só pode ser preenchida, com a pessoa que falta. Logo, existem $5 \times 4 \times 3 \times 2 \times 1 = 120$ possíveis filas. Se começarmos com 6 pessoas, o número de filas diferentes será $6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$ e, para 7 pessoas, será $7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$ filas possíveis.

1850
Kirkman coloca o problema das 15 alunas

1930
Frank Ramsey trabalha em análise combinatória

1971
Ray-Chaudhuri e Wilson provam a existência de sistemas de Kirkman gerais

Um número obtido por multiplicação sucessiva de números inteiros chama-se factorial. Estes números ocorrem tão frequentemente em matemática, que são escritos usando a notação 5! (lê-se «5 factorial»), em vez de $5 \times 4 \times 3 \times 2 \times 1$. Vejamos os primeiros factoriais (definimos 0! como igual a 1). Constatamos imediatamente que configurações bastante «pequenas» dão origem a «grandes» factoriais. O número *n* pode ser pequeno, mas *n!* pode ser enorme.

número	factorial
0	1
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40 320
9	362 880

Se ainda quisermos formar filas com 5 pessoas, mas agora pudermos recorrer a um conjunto de 8 pessoas, **A**, **B**, **C**, **D**, **E**, **F**, **G**, e **H**, a análise é quase igual. Há 8 escolhas para a primeira pessoa da fila, 7 para a segunda, etc. Mas, desta vez, há 4 escolhas para a última posição. O número de possíveis filas é

$$8 \times 7 \times 6 \times 5 \times 4 = 6720$$

Isto pode ser escrito com a notação para números factoriais, porque

$$8 \times 7 \times 6 \times 5 \times 4 = 8 \times 7 \times 6 \times 5 \times 4 \times \frac{3 \times 2 \times 1}{3 \times 2 \times 1} = \frac{8!}{3!}$$

Combinações Numa fila, a *ordem* é importante. As duas filas

C E B A D **D A C E B**

são constituídas pelas mesmas letras, mas são diferentes. Já sabemos que podemos fazer 5! filas com estas letras. Se quisermos contar as formas de seleccionar 5 pessoas de 8 *por não importa que ordem*, devemos dividir $8 \times 7 \times 6 \times 5 \times 4 = 6720$ por 5!. Logo, o número de maneiras de seleccionar 5 pessoas de 8 é

$$\frac{8\times7\times6\times5\times4}{5\times4\times3\times2\times1} = 56.$$

Este número, usando C para combinações, escreve-se ${}^{8}C_{5}$ e é

$${}^{8}C_{5} = \frac{8!}{3!\,5!} = 56$$

Na lotaria inglesa, as regras exigem uma selecção de seis números de 49 possíveis – quais são as possibilidades de ganhar?

$${}^{49}C_{6} = \frac{49!}{43!\,6!} = \frac{49\times48\times47\times46\times45\times44}{6\times5\times4\times3\times2\times1} = 13\,983\,816$$

Só uma combinação ganha, logo há aproximadamente uma hipótese em 14 milhões de ganhar o prémio.

O problema de Kirkman A análise combinatória é um campo vasto e, embora antigo, tem-se desenvolvido rapidamente ao longo dos últimos 40 anos, dada a sua relevância para as ciências da computação. Os problemas que envolvem teoria dos grafos, quadrados latinos e assim por diante podem ser considerados como parte da análise combinatória.

A essência da análise combinatória é capturada pelo mestre no assunto, o rev. Thomas Kirkman, que trabalhou numa época em que a análise combinatória estava sobretudo ligada à matemática recreativa. Kirkman fez muitas contribuições originais para a geometria discreta, a teoria dos grupos e a análise combinatória, mas nunca exerceu um cargo universitário. Ainda hoje é conhecido devido à adivinha que reforçou a sua reputação como o matemático do absurdo. Em 1850, Kirkman apresentou o «problema das 15 alunas», no qual 15 alunas caminham para a igreja em 5 linhas de 3 em cada dia da semana. Se estão aborrecidos com o *sudoku*, podem tentar resolver este problema. Temos de organizar um calendário diário tal, que cada par não caminhe junto mais do que uma vez. Usando deliberadamente letras minúsculas e maiúsculas, as alunas são: **a**bigail, **b**eatrice, **c**onstance, **d**orothy, **e**mma, **f**rances, **g**race, **A**gnes, **B**ernice, **C**harlotte, **D**anielle, **E**dith, **F**lorence, **G**wendolyn e **V**ictoria, representadas por **a**, **b**, **c**, **d**, **e**, **f**, **g**, **A**, **B**, **C**, **D**, **E**, **F**, **G** e **V**, respectivamente.

Existem na realidade sete soluções para o problema de Kirkman. A que daremos é «cíclica» – gerada «andando à volta». É aqui que a representação das alunas é importante.

segunda			terça			quarta			quinta			sexta			sábado			domingo		
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**

A solução diz-se cíclica, dado que no dia seguinte o calendário altera **a** para **b**, **b** para **c**, até **g**, que é substituído por **a**. O mesmo se aplica às letras maiúsculas **A** para **B**, **B** para **C**, etc, mas **V**ictoria mantém-se inalterado.

A razão subjacente à escolha da notação é que as linhas correspondem às rectas na geometria de Fano (ver página 115). O problema de Kirkman não só é um jogo de salão, como faz parte da corrente principal da matemática.

a ideia resumida
Quantas combinações?

42 Quadrados mágicos

«Um matemático», escreveu G. H. Hardy, «como um pintor ou um poeta, é um criador de padrões.» Os quadrados mágicos têm padrões curiosos, mesmo segundo os critérios matemáticos. Situam-se na fronteira entre a matemática excessivamente simbólica e os padrões fascinantes dos amantes de *puzzles*.

Um quadrado mágico é uma grelha quadrada com números inteiros distintos escritos em cada célula, de tal forma que cada linha horizontal e cada coluna vertical e cada diagonal somem o mesmo número.

a	*b*
c	*d*

Os quadrados só com uma linha e uma coluna são tecnicamente quadrados mágicos, mas são muito aborrecidos, pelo que vamos esquecê-los. Não existe tal coisa como quadrados mágicos com duas linhas e duas colunas. Se houvesse, teria a forma mostrada. Dado que as somas das linhas e das colunas terá de ser igual, então $a + b = a + c$. Isto significa que $b = c$, contradizendo o facto de todas as entradas deverem ser diferentes.

O quadrado de Lo Shu Como os quadrados 2 × 2 não existem, observemos os 3 × 3 e tentemos construir um com uma grelha. Começaremos com um quadrado mágico *normal*, cuja grelha é preenchida com os números consecutivos 1, 2, 3, 4, 5, 6, 7, 8 e 9.

Para um quadrado tão pequeno, é possível construir um quadrado mágico 3 × 3, pelo método de «tentativa e erro», mas podemos fazer primeiro algumas deduções, para facilitação. Se somarmos todos os números da grelha, temos

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = 45$$

Cronologia

cerca de 2800 a.C.

Nascimento da lenda do quadrado de Lo Shu

cerca de 1690

De la Loubère produz o método Siamese para construir quadrados mágicos e este total terá de ser igual ao total das três linhas. Isto mostra que cada linha (e coluna e diagonal) deve somar 15. Vamos agora olhar para a célula do meio – chamar-lhe-emos *c*. As duas diagonais, tal como a linha do meio e a coluna do meio, envolvem *c*. Se somarmos os números destas quatro linhas, obtemos 15 + 15 + 15 + 15 = 60, e este valor deve igualar *todos* os números somados mais 3 *c* extras. Da equação $3c + 45 = 60$, vemos que *c tem de ser* 5. Também podem ser estudados outros factos, como não ser possível colocar 1 numa célula de canto. Depois de algumas pistas reunidas, estamos preparados para usar o método da tentativa e erro. Experimente!

8	1	6
3	5	7
4	9	2

Uma solução para o quadrado 3 x 3 pelo método Siamese

Claro que gostaríamos de um *método* totalmente sistemático para construir quadrados mágicos. Simon de la Loubère, embaixador francês do rei de Sião no final do século XVII, descobriu um. Loubère interessava-se pela matemática chinesa e escreveu um método para construir quadrados mágicos contendo um número ímpar de linhas e colunas. Este método começa pela colocação de 1 no meio da primeira linha e «indo para cima e para o lado e rodando se necessário» para colocar 2 e os números subsequentes. Em caso de bloqueio, usa-se o número seguinte abaixo do número corrente.

Extraordinariamente, este quadrado mágico normal é essencialmente o único com 3 linhas e 3 colunas. Todos os outros quadrados mágicos 3 × 3 podem ser obtidos deste por rotação dos números sobre o centro e/ou reflectindo números do quadrado na coluna do meio ou na linha do meio. É o chamado quadrado de «Lo Shu», conhecido na China por volta de 3000 a.C. Diz a lenda que foi visto pela primeira vez nas costas de uma tartaruga que saía do rio Lo. A população local tomou-o como um sinal divino de que não se veria livre da peste, a não ser que aumentasse as suas ofertas.

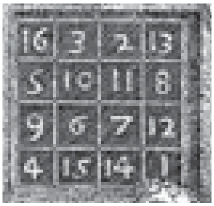
Se existe um quadrado mágico 3 × 3, quantos quadrados mágicos 4 × 4 existem? A impressionante resposta é 880 (e, prepare-se, há 2 202 441 792 quadrados mágicos de ordem 5). Não se sabe quantos existem para um valor geral de *n*.

Os quadrados de Dürer e de Franklin O quadrado de Lo Shu é bem conhecido pela sua antiguidade e unicidade, mas houve um quadrado 4 × 4 que se tornou um ícone pela sua associação a um artista famoso. Também tem muito mais propriedades do que os habituais quadrados mágicos que constituem as 880 versões. É o quadrado mágico 4 × 4 da gravura *Melancholia* de Albrecht Dürer, feita em 1514.

1693
Bernard Frénicle de Bessy lista todos os 880 possíveis quadrados mágicos 4 x 4

1770
Euler produz um quadrado quadrado

1986
Sallows cria o seu quadrado baseado em letras



No quadrado de Dürer, todas as linhas somam 34, bem como as colunas, as diagonais e os quadrados pequenos 2 × 2 que formam o quadrado completo 4 × 4. Dürer até conseguiu «assinar» a sua obra-prima com a data em que o terminou no meio da linha inferior.

O cientista e diplomata norte-americano Benjamin Franklin verificou que a construção de quadrados mágicos era uma ferramenta útil para aguçar a mente. Era um perito nestas construções, e até hoje os matemáticos têm pouca ideia de como ele o fazia. Os quadrados mágicos grandes não podem ser feitos ao acaso. Franklin confessou que, na sua juventude, tinha gastado muito tempo com eles, apesar de não se interessar muito por «aritmética» enquanto jovem. Aqui está um quadrado que ele descobriu na juventude.

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

Neste quadrado mágico normal, há todos os tipos de simetria. Todas as linhas, colunas e diagonais somam 260, tal como as «linhas dobradas», uma das quais está assinalada. Há muitas outras coisas para descobrir – como a soma do central 2 × 2 mais os quatro cantos, que também é 260. Se observarmos com atenção, veremos um resultado interessante para cada quadrado 2 × 2.

Quadrados quadrados Alguns quadrados mágicos podem ter células ocupadas por diferentes números quadrados. O problema de os construir foi colocado pelo matemático francês Edouard Lucas em 1876. Até à data, não se encontrou nenhum quadrado 3 × 3 de quadrados, embora um tenha ficado perto.

127^2	46^2	58^2
2^2	113^2	94^2
74^2	82^2	97^2

Todas as linhas e colunas e *uma* diagonal deste quadrado têm a soma mágica de 21 609, mas a outra diagonal falha, dado que $127^2 + 113^2 + 97^2 = 38\,307$. Se estiver tentado a encontrar um por si próprio, deve tomar nota de um resultado já provado: a célula central deve ser maior do que $2{,}5 \times 10^{25}$, logo não faz sentido procurar um quadrado com números pequenos! Isto é matemática séria, que tem conexão com curvas elípticas, o tópico utilizado para provar o Último Teorema de Fermat. Foi provado que não existem quadrados mágicos 3 × 3 cujas entradas sejam cubos ou quartas potências.

A procura de quadrados quadrados tem, contudo, sido bem-sucedida para quadrados grandes. Existem quadrados mágicos 4 × 4 e 5 × 5. Em 1770, Euler produziu um exemplo sem mostrar o método de construção. Já foram encontradas famílias inteiras ligadas ao estudo da álgebra dos quaterniões, os números imaginários de quatro dimensões.

Quadrados mágicos exóticos Os grandes quadrados mágicos têm propriedades espectaculares. William Benson, especialista em quadrados mágicos, produziu um quadrado 32 × 32 em que os números, os seus quadrados e os seus cubos formam todos quadrados mágicos. Em 2001, produziu-se um quadrado 1024 × 1024 em que todas as potências de elementos até à quinta potência fazem quadrados mágicos. Há muitos resultados como este.

Podemos criar toda uma variedade de outros quadrados mágicos, se os requisitos forem atenuados. Os quadrados mágicos normais estão na corrente principal. Removendo a condição de que a soma dos elementos da diagonal tem de ser igual à soma das linhas e das colunas, abre-se uma infinidade de resultados especializados. Podemos procurar quadrados cujos elementos sejam apenas números primos, ou podemos considerar outras formas para além dos quadrados que tenham «propriedades mágicas». Indo para dimensões mais altas, podemos considerar cubos mágicos e hipercubos.

Mas o prémio para o mais notável quadrado mágico de todos, sem dúvida pelo valor da curiosidade, tem de atribuir-se ao humilde quadrado 3 × 3 produzido pelo engenheiro electrónico e escritor holandês Lee Sallows:

5	22	18
28	15	2
12	8	25

Porque é que é tão notável? Primeiro, escrevamos os números em inglês:

five	twenty-two	eighteen
twenty-eight	fifteen	two
twelve	eight	twenty-five

Depois, se contarmos o número de letras que formam cada palavra, obtemos:

4	9	8
11	7	3
6	5	10

Notavelmente, este é um quadrado mágico que consiste nos números consecutivos 3, 4, 5 até 11. Também verificamos que o número de letras das somas mágicas de ambos os quadrados 3 × 3 (21 e 45) é 9 e, adequadamente, 3 × 3 = 9.

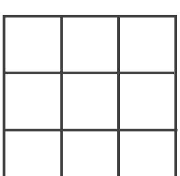
a ideia resumida
Magia matemática

43 Quadrados latinos

Há alguns anos que o mundo anda louco com o *sudoku*. Em todos os países, as canetas e os lápis são mastigados enquanto se espera a inspiração certa para o número a colocar na caixa. É um 4 ou um 5? Talvez seja 9. Os passageiros saem dos seus comboios matinais tendo despendido um maior esforço mental do que aquele que despenderão durante o resto do dia. À noite, o jantar queima-se no forno. É 5, 4 ou 7? Todas estas pessoas brincam com quadrados latinos – tal como os matemáticos.

	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	

***Sudoku* aberto** No *sudoku*, há uma grelha 9 × 9 com alguns números preenchidos. O objectivo é preencher o resto da grelha, usando os números dados como pistas. Cada linha e cada coluna devem conter exactamente um dos números 1, 2, 3, ..., 9, tal como os quadrados 3 × 3 que a constituem.



Supõe-se que o *sudoku* («um só dígito») foi inventado no fim da década de 1970. Ganhou popularidade no Japão na década seguinte, antes de se tornar imensamente popular em 2005. O apelo deste quebra-cabeças é que, ao contrário das palavras cruzadas, não é necessário termos feito muitas leituras para o solucionar. Contudo, como as palavras cruzadas, o *sudoku* pode ser muito atraente. Os viciados em ambas as formas de autotortura têm muito em comum.

Cronologia

1779
Euler explora a teoria dos quadrados latinos

1900
Tarry mostra que não há quadrados latinos ortogonais de ordem 6

Quadrados latinos 3 × 3 Um quadrado que contém exactamente um símbolo em cada linha e em cada coluna chama-se um quadrado latino.
O número de símbolos iguala o tamanho do quadrado e chama-se a sua «ordem». Conseguimos preencher uma grelha limpa 3 × 3, tal que cada linha e cada coluna contenham exactamente um dos símbolos *a*, *b* e *c*? Se conseguirmos, temos um quadrado latino de ordem 3.

Introduzindo o conceito de quadrado latino, Leonhard Euler chamou-lhe uma «nova espécie de quadrado mágico». No entanto, ao contrário dos quadrados mágicos, os quadrados latinos não estão preocupados com a aritmética e os símbolos não têm de ser números. A razão do nome é simplesmente os símbolos usados para o formarem serem os do alfabeto latino, enquanto Euler usou o grego com outros quadrados.

a	*b*	*c*
b	*c*	*a*
c	*a*	*b*

Um quadrado latino 3 × 3 pode ser facilmente escrito.

Se pensarmos em *a*, *b* e *c* como os dias da semana *segunda*, *quarta* e *sexta*, o quadrado pode ser usado para calendarizar reuniões entre duas equipas de pessoas. A Equipa Um é constituída por **L**arry, **M**ary e **N**ancy, e a Equipa Dois, por **R**oss, **S**ophie e **T**om.

	R	**S**	**T**
L	*a*	*b*	*c*
M	*b*	*c*	*a*
N	*c*	*a*	*b*

Por exemplo, **M**ary da Equipa Um tem uma reunião com **T**om da Equipa Dois na segunda-feira (a intersecção da linha **M** com a coluna **T** é *a* = segunda. A disposição do quadrado latino garante que existe uma reunião para cada par dos membros das equipas e que não há conflitos de datas.

Este não é o único quadrado latino 3 × 3 possível. Se interpretarmos A, *B* e C como os tópicos discutidos nas reuniões entre a Equipa Um e a Equipa Dois, podemos produzir um quadrado latino que garanta que cada pessoa discute um tópico diferente com um membro da outra equipa.

	R	**S**	**T**
L	A	*B*	C
M	C	A	*B*
N	*B*	C	A

Assim, **M**ary da Equipa Um discute o tópico C com **R**oss da Equipa Dois, o tópico A com **S**ophie e o tópico *B* com Tom.

1925
Fisher sugere o uso de quadrados latinos nos projectos de experiências estatísticas

1960
A conjectura de Euler acerca da não-existência de alguns pares de quadrados latinos é refutada por Bose, Parker e Shrikhande

1979
Invenção de jogos semelhantes ao *sudoku* em Nova Iorque

Mas *quando* deveriam essas discussões ter lugar, entre *quem*, e sobre qual tópico? Qual seria o calendário para esta complexa organização? Felizmente, os dois quadrados latinos podem ser combinados símbolo a símbolo, para produzir um quadrado latino composto em que cada um dos possíveis nove pares de dias e tópicos ocorre exactamente numa posição.

	R	**S**	**T**
L	*a*,A	*b*,*B*	*c*,C
M	*b*,C	*c*,A	*a*,*B*
N	*c*,*B*	*a*,C	*b*,A

Outra interpretação do quadrado é o histórico «problema dos nove oficiais», no qual nove oficiais pertencendo a três regimentos *a*, *b* e *c* e de três patentes A, *B* e C são colocados na parada, de forma que cada linha e cada coluna tenham um oficial de cada regimento e patente. Os quadrados latinos que se combinam desta forma chamam-se «ortogonais». O caso 3 × 3 é simples, mas encontrar pares de quadrados latinos ortogonais para alguns maiores está longe de ser fácil, como Euler descobriu.

No caso de um quadrado latino 4 × 4, um «problema dos 16 oficiais» seria arrumar as 16 figuras e ases de um baralho de cartas num quadrado, de tal forma que haja uma patente (Ás, Rei, Dama ou Valete) e um naipe (espadas, paus, copas e ouros) em cada linha e coluna. Em 1872, Euler colocou o mesmo problema para «36 oficiais». No essencial, Euler estava à procura de dois quadrados ortogonais de ordem 6. Não conseguiu encontrá-los e conjecturou que não havia pares de quadrados latinos ortogonais das ordens 6, 10, 14, 18, 22… Poderá isto provar-se?

A seguir, apareceu Gaston Tarry, um matemático amador que trabalhava como funcionário público na Argélia. Tarry fez um escrutínio dos exemplos e em 1900 tinha verificado a conjectura de Euler num caso: não há nenhum par de quadrados latinos ortogonais de ordem 6. Os matemáticos presumiram naturalmente que Euler estava correcto nos casos de ordem 10, 14, 18, 22...

Em 1960, os esforços combinados de três matemáticos chocaram o mundo matemático, provando que Euler estava errado em *todos* os outros casos. Raj Bose, Ernest Parker e Sharadchandra Shrikhande provaram que existem realmente pares de quadrados latinos de ordens 10, 14, 18, 22... O *único* caso em que os quadrados latinos não existem (excluindo os triviais de ordem 1 e 2) é a ordem 6.

Já vimos que há dois quadrados latinos de ordem 3 mutuamente ortogonais. Para a ordem 4, podemos produzir três quadrados mutuamente ortogonais uns aos outros. Pode mostrar-se que nunca há mais de $n - 1$ quadrados latinos mutuamente ortogonais de ordem n, logo para $n = 10$, por exemplo, não podem existir mais do que nove quadrados mutuamente ortogonais. Mas encontrá-los é outra história. Até à data, ninguém conseguiu produzir nem sequer três quadrados latinos de ordem 10 que fossem mutuamente ortogonais entre si.

Os quadrados latinos são úteis? R. A. Fisher, um eminente estatístico, percebeu o uso prático dos quadrados latinos. Usou-os para revolucionar os métodos agrícolas durante o tempo em que esteve na Estação de Pesquisa Rothamsted em Hertfordshire, na Inglaterra.

O objectivo de Fisher era investigar a eficácia dos fertilizantes na produtividade agrícola. O plano era plantar em solos com idênticas condições, para que a qualidade dos solos não fosse um factor indesejado que influenciasse a produtividade da cultura. Podíamos então aplicar os diferentes fertilizantes de forma segura, com o conhecimento de que o «transtorno» da qualidade do solo tinha sido eliminado. A única forma de conseguir condições de solo iguais seria usar o mesmo solo, mas é impraticável continuar a cavar e a replantar culturas. Mesmo que isso fosse possível, as condições meteorológicas seriam um novo estorvo.

Uma maneira de contornar o problema é usar quadrados latinos. Vejamos os testes para quatro tratamentos. Se marcarmos um campo quadrado com 16 parcelas, podemos encarar um quadrado latino como a descrição de um campo onde a qualidade do solo varia «verticalmente» e «horizontalmente».

Os quatro fertilizantes são então aplicados aleatoriamente no esquema com as designações *a*, *b*, *c* e *d*, para que seja aplicado exactamente um em cada linha e coluna numa tentativa de eliminar a variação da qualidade do solo. Se suspeitarmos que outro factor pode influenciar a colheita, também podemos lidar com ele. Suponhamos que a altura do dia em que se aplica o fertilizante é um factor. Designamos quatro zonas temporais ao longo do dia como *A*, *B*, *C* e *D* e usamos quadrados latinos ortogonais para o desenho do esquema de recolha de dados. Isto garante que cada tratamento e zona temporal é aplicado numa das parcelas. O plano da experiência seria:

a, tempo *A*	*b*, tempo *B*	*c*, tempo *C*	*d*, tempo *D*
b, tempo *C*	*a*, tempo *D*	*d*, tempo *A*	*c*, tempo *B*
c, tempo *D*	*d*, tempo *C*	*a*, tempo *B*	*b*, tempo *A*
d, tempo *B*	*c*, tempo *A*	*b*, tempo *D*	*a*, tempo *C*

Podem eliminar-se outros factores continuando a criar quadrados latinos ainda mais elaborados. Euler nunca teria sonhado que a solução do seu problema dos oficiais fosse aplicada a experiências agrícolas.

a ideia resumida
Sudoku resolvido

44 A matemática do dinheiro

Norman é um supervendedor quando se trata de bicicletas. Além do mais, acha que é seu dever pôr toda a gente a andar de bicicleta, logo fica deliciado quando um cliente entra na sua loja e, sem hesitar, compra uma bicicleta de € 99. O cliente paga com um cheque de € 150 e, como os bancos estão fechados, Norman pede ao vizinho para o descontar. Volta, dá ao cliente € 51 e este desaparece a grande velocidade. Segue--se a calamidade. O cheque é rejeitado, o vizinho pede o seu dinheiro de volta, e Norman tem de pedir a um amigo que lhe empreste dinheiro. A bicicleta tinha-lhe custado € 79, mas quanto é que Norman perdeu ao todo?

O conceito desta pequena adivinha foi proposto pelo criador de *puzzles* Henry Dudeney. É uma espécie de matemática do dinheiro, mas mais precisamente um *puzzle* relacionado com dinheiro. Na altura em que foi escrito, na década de 1920, a bicicleta de Dudeney custava €15. Uma maneira de combater a inflação são os juros. Estes envolvem matemática séria e estão no cerne dos modernos mercados financeiros.

Juros compostos Há dois tipos de juros: simples e compostos. Concentremos a nossa atenção matemática em dois irmãos, Charlie Composto e Simon Simples. O pai dá a cada um €1000, que ambos depositam num banco. Charlie Composto escolhe uma conta que aplica juros compostos, mas Simon Simples é mais tradicional e prefere contas que usem juros simples. Os juros compostos são historicamente relacionados com a usura e a desaprovação. Hoje em dia, os juros compostos são um facto da vida, central nos sistemas monetários modernos. Os juros compostos são juros compostos sobre juros,

Cronologia

3000 a.C.
Os babilónios usam um sistema numérico sexagesimal para questões financeiras

1494
Luca Pacioli publica tabelas financeiras e uma contabilidade de registo com duas entradas

e é por isso que Charlie gosta deles. Os juros simples não têm essa característica e são calculados sobre determinado montante conhecido como «principal». Simon pode compreendê-lo facilmente, dado que o montante rende a mesma quantia em juros todos os anos.

Quando se fala de matemática, é sempre bom ter Albert Einstein ao lado – mas a ideia generalizada de que ele afirmou que a maior descoberta de todos os tempos eram os juros compostos é demasiado forçada.
É, contudo, inegável que a fórmula dos juros compostos tem uma relação imediata com a sua $E = mc^2$. Se poupar dinheiro, pedir um empréstimo, usar cartão de crédito, fizer uma hipoteca ou comprar uma anuidade, a fórmula do juro composto estará por trás a trabalhar a seu favor (ou contra si). O que significam os símbolos? O termo P é o montante principal (o dinheiro poupado ou o empréstimo), j é a taxa de juro percentual dividida por 100 e n é o número de períodos de tempo.

$$A = P \times (1 + j)^n$$

Fórmula do juro composto

Charlie coloca os seus €1000 numa conta que rende anualmente 7%. Quanto é que obterá ao fim de três anos? Aqui $P = 100$, $i = 0{,}07$ e $n = 3$. O símbolo A representa o montante obtido e pela fórmula A = €1225,04.

A conta de Simon rende o mesmo juro, 7%, mas em juro simples. Como é que os seus ganhos se comparam ao fim de três anos? No primeiro ano, ganhará €70 de juro, tal como no segundo e terceiro anos. Portanto, Simon receberá 3 × €70 de juros, num total de €1210. O investimento de Charlie foi a melhor decisão.

As somas de dinheiro que podem crescer por composição podem aumentar muito rapidamente. Isto é óptimo se estamos a poupar, mas não é assim tão bom se estamos a dever. A principal componente do juro composto é o período no qual a composição tem lugar. Charlie ouviu falar de um esquema que rende 1% por semana, um cêntimo por cada euro. Quanto teria de esperar para ganhar com este esquema?

Simon pensa conhecer a resposta: sugere multiplicar o juro de 1% por 52 (o número de semanas no ano) para obter a percentagem da taxa anual de 52%. Isto significa um juro de €520, num total de €1520 na conta. Porém,

1718
Abraham de Moivre investiga estatísticas de mortalidade e os fundamentos da teoria das anuidades

1756
James Dodson publica *First Lectures on Insurances*

1848
É fundado em Londres o Instituto de Actuários

Charlie lembra-o do mágico juro composto e da fórmula do juro composto. Com $P = 100$, $j = 0{,}01$ e $n = 52$, Charlie calcula que o aumento é €1000 x $(1{,}01)^{52}$. Usando uma calculadora, Charlie chega ao valor de €1677,69, muito mais que o resultado da soma de Simon Simples. O equivalente da taxa anual é 67,769% e é muito maior que os 52% do cálculo de Simon.

Simon está impressionado, mas o seu dinheiro já está no banco num regime de juro simples. Pergunta-se quanto tempo levará a duplicar os seus €1000 originais. A cada ano, recebe €70 de juro, logo tudo o que tem a fazer é dividir 1000 por 70. O total é 14,29 e, assim, pode ter a certeza de que, ao fim de 15 anos, terá mais de €2000 no banco. É muito tempo à espera. Para mostrar a superioridade do juro composto, Charlie começa a calcular o seu período de duplicação. Isto é mais complicado, mas um amigo fala-lhe da regra dos 72.

A regra dos 72 Para uma dada taxa, a regra dos 72 é uma regra de ouro para estimar o número de períodos necessário para o dinheiro duplicar. Embora Charlie esteja interessado em anos, a regra dos 72 também se aplica a dias ou meses. Para encontrar o período de duplicação, tudo o que Charlie tem de fazer é dividir 72 pela taxa. O cálculo é $72/7 = 10{,}3$, logo Charlie pode dizer ao irmão que o seu investimento duplicará ao fim de 11 anos, muito mais rápido que os 15 de Simon. A regra é uma aproximação, mas é útil quando têm de ser tomadas decisões rápidas.

Valor actual O pai de Charlie Composto está tão impressionado com o bom senso do filho, que o chama à parte e diz: «Proponho dar-te €100 000.» Charlie fica muito animado. A seguir, o pai junta a condição de que só lhe dará os €100 000 quando ele tiver 45 anos, o que só será daqui a dez anos. Charlie não fica muito satisfeito.

Charlie quer gastar o dinheiro agora, mas obviamente não pode. Vai ao seu banco e promete-lhes €100 000 daí a dez anos. O banco responde que tempo é dinheiro e que €100 000 daí a dez anos não são o mesmo que €100 000 agora. O banco tem de estimar o investimento actual que realizará €100 000 daqui a dez anos. Esse será o montante que emprestará a Charlie. O banco acredita que uma taxa de crescimento de 12% lhe dará um bom lucro. Qual será o montante que crescerá para €100 000 em dez anos, a 12% de juro? A fórmula do juro composto também pode ser usada para este problema. Desta vez, temos A = €100 000 e precisamos calcular P, o valor actual de A. Com $n = 10$ e $j = 0{,}12$, o banco estará preparado para avançar a Charlie o montante de $100\,000/1{,}12^{10}$. Charlie fica chocado com este número pequeno, mas ainda poderá comprar o Porsche novo.

Como tratar os pagamentos regulares? Agora que prometeu dar €100 000 ao filho daqui a dez anos, o pai de Charlie tem de poupar. Planeia fazê-lo com um fluxo de pagamentos iguais da conta-poupança no fim de cada ano durante dez anos. No fim deste período, será capaz de entregar o dinheiro a Charlie no dia prometido, e Charlie pode levar o dinheiro ao banco para pagar o empréstimo.

O pai de Charlie consegue encontrar um banco que paga uma taxa de juro de 8% para o ano inteiro. Dá a Charlie a tarefa de calcular os pagamentos anuais. Com a fórmula do juro composto, Charlie estava preocupado com um pagamento (o montante original), mas agora está preocupado com dez pagamentos feitos em alturas diferentes. Se os pagamentos regulares R são feitos no final de cada ano numa conjuntura em que a taxa de juro é j, o montante poupado depois de n anos pode ser calculado pela fórmula dos pagamentos regulares.

$$S = R \times \frac{((1+j)^n - 1)}{j}$$

Fórmula dos pagamentos regulares

Charlie sabe que S = €100 000, n = 10 e j = 0,08 e calcula que R = €6902,95.

Agora que Charlie tem o seu Porsche novinho em folha, cortesia do banco, necessita de uma garagem para o guardar. Decide fazer uma hipoteca de €300 000 para comprar uma casa, uma soma que pagará num fluxo de pagamentos anuais e iguais ao longo de 25 anos. Charlie identifica a questão como um problema no qual os €300 000 são o valor actual de um fluxo de pagamentos a ser feito e calcula os pagamentos anuais com facilidade. O pai fica impressionado e faz mais uma utilização das proezas de Charlie. Acabou de lhe ser atribuído um montante fixo de aposentação de €150 000 e o pai quer comprar uma anuidade. Está bem, diz Charlie, «podemos usar a mesma fórmula, dado que a matemática é a mesma. Em vez da sociedade hipotecária me avançar o dinheiro que pagarei em prestações regulares, o pai dá-lhes o dinheiro e eles fazem os pagamentos regulares a si».

A propósito, a resposta para o quebra-cabeças de Henry Dudeney é €130, constituídos pelos €51 que deu ao cliente e os €79 que pagou pela bicicleta.

a ideia resumida
Os juros compostos funcionam melhor

45 O problema da dieta

Tanya Smith leva a sua forma física muito a sério. Vai ao ginásio todos os dias e toma muita atenção à sua dieta. Tanya ganha a vida com empregos em tempo parcial e tem de vigiar onde gasta o dinheiro. É crucial que tome as quantidades necessárias de minerais e vitaminas todos os meses para continuar em boa forma e saudável. As quantidades foram determinadas pelo seu treinador, que sugere que os futuros campeões olímpicos devem ingerir pelo menos 120 mg de vitaminas e pelo menos 880 mg de minerais todos os meses. Para ter a certeza de que segue o regime, Tanya conta com dois suplementos alimentares. Um tem forma sólida e o nome comercial de Solido e o outro é líquido e é comercializado como Liquex. O problema é decidir que quantidade deve comprar de cada um todos os meses para satisfazer o seu treinador.

O clássico problema da dieta é organizar uma dieta saudável e pagar o menos possível por ela. Foi um protótipo de problemas de programação linear, um assunto desenvolvido na década de 1940 e actualmente utilizado numa grande variedade de aplicações.

No início de Março, Tanya foi ao supermercado e analisou o Solido e o Liquex. No verso do pacote, descobriu que Solido continha 2 mg de vitaminas e 10 mg de minerais, enquanto uma embalagem de Liquex continha 3 mg de vitaminas e 50 mg de minerais. Na dúvida, colocou no seu carrinho de compras 30 pacotes de Solido e 5 embalagens de Liquex para

	Solido	Liquex	Necessidades
Vitaminas	2 mg	3 mg	120 mg
Minerais	10 mg	50 mg	880 mg

Cronologia

1826

Fourier antecipa a programação linear; Gauss resolve equações lineares pela eliminação gaussiana

1902

Farkas dá uma solução para sistemas de inequações um mês. Enquanto caminhava para a saída, interrogou-se se teria as quantidades certas. Primeiro, calculou quantas vitaminas tinha no carrinho. Nos 30 pacotes de Solido tinha $2 \times 30 = 60$ mg de vitaminas e nos de Liquex $3 \times 5 = 15$. Ao todo, tinha $2 \times 30 + 3 \times 5 = 75$ mg de vitaminas. Repetindo os cálculos para os minerais, tinha $10 \times 30 + 50 \times 5 = 550$ mg de minerais.

Como o treinador pretendia que tomasse pelo menos 120 mg de vitaminas e 880 mg de minerais, Tanya precisava de mais pacotes e embalagens no carrinho. O problema de Tanya era combinar as quantidades certas de Solido e Liquex com as vitaminas e os minerais necessários. Voltou à secção de saúde do supermercado e pôs mais pacotes e embalagens no carrinho. Tinha agora 40 pacotes e 15 embalagens. Chegará? Tanya recalculou e verificou que tinha $2 \times 40 + 3 \times 15 = 125$ mg de vitaminas e $10 \times 40 + 50 \times 15 = 1150$ mg de minerais. Agora certamente cumpria as recomendações do treinador e até excedia as quantidades exigidas.

Soluções viáveis A combinação (40, 15) de suplementos possibilitará a Tanya cumprir a dieta. Chama-se a isto uma combinação possível, ou solução «viável». Já vimos que a solução (30, 5) não é uma solução viável, logo há uma demarcação entre os dois tipos de combinações – soluções viáveis pelas quais a dieta é cumprida e soluções pelas quais não é.

Tanya tem mais opções. Podia encher o carrinho só com Solido. Se o fizesse, precisaria de pelo menos 88 pacotes. A compra (88, 0) satisfaz ambas as necessidades, porque a combinação conterá $2 \times 88 + 3 \times 0 = 176$ mg de vitaminas e $10 \times 88 + 50 \times 0 = 880$ mg de minerais. Se só comprasse Liquex, precisaria de pelo menos 40 embalagens, a solução viável (0, 40) satisfaz a necessidade tanto de vitaminas como de minerais, porque $2 \times 0 + 3 \times 40 = 120$ mg de vitaminas e $10 \times 0 + 50 \times 40 = 2000$ mg de minerais. Podemos notar que a quantidade de vitaminas e minerais não é encontrada *exactamente* em nenhuma destas possíveis combinações, embora o treinador ficasse certamente satisfeito por Tanya estar a tomar o suficiente.

Soluções óptimas Agora, introduzimos o dinheiro no problema. Quando Tanya chega à saída, tem de pagar as compras. Ela nota que os pacotes e as embalagens têm o mesmo preço de €5 cada. Das combinações viáveis, encontramos até agora (40, 15), (88, 0), (0, 40), cujo custo seria €275, €440 e €200, respectivamente, donde a melhor solução até agora será não comprar

1945
Stigler resolve o problema da dieta por um método heurístico

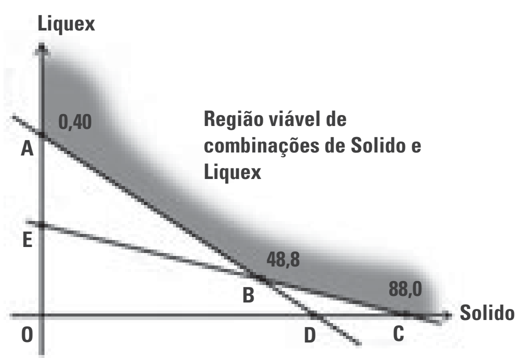
1947
Dantzig formula o método simplex e resolve o problema da dieta por programação linear

1984
Karmarkar deduz um novo algoritmo para resolver problemas de programação linear

nenhum Solido e comprar 40 embalagens de Liquex. Este será o menor custo, e as necessidades da dieta são conseguidas. Mas a quantidade de suplementos a comprar tem sido uma questão de tentativa e erro. No impulso do momento, Tanya tentou várias combinações de Solido e Liquex e calculou os custos apenas nesses casos. Poderia melhorar a situação? Existe alguma combinação de Solido e Liquex que satisfaça o treinador e ao mesmo tempo lhe custe menos? O que ela gostaria de fazer era ir para casa e analisar o problema com um lápis e um papel.

Problemas de programação linear Tanya tem sido sempre treinada para visualizar os seus objectivos – se pode aplicar essa faculdade à conquista do ouro olímpico, porque não à matemática? Assim, ela desenha um gráfico com a região viável, o que é possível porque só está a considerar dois suplementos. A recta AD representa a combinação de Solido e Liquex que contém exactamente 120 mg de vitaminas. As combinações acima desta linha têm mais de 120 mg de vitaminas. A recta EC representa as combinações que contêm exactamente 880 mg de minerais. A combinação dos dois suplementos que estão acima de ambas as linhas é a região viável e representa todas as combinações viáveis que Tanya podia tentar.

Aos problemas com a estrutura do problema da dieta chama-se problemas de programação linear. A palavra «programação» significa procedimento (a sua utilização é anterior à palavra se ter tornado sinónimo de computadores), enquanto «linear» se refere ao uso de rectas. Para resolver o problema de Tanya com programação linear, os matemáticos mostraram que tudo o que é preciso fazer é determinar o valor do custo dos suplementos nos pontos dos cantos do gráfico de Tanya. Tanya descobriu uma nova solução viável no ponto B de coordenadas (48, 8), o que significa que pode comprar 48 pacotes de Solido e 8 embalagens de Liquex. Se o fizer, satisfará *exactamente* a sua dieta, porque nesta combinação há 120 mg de vitaminas e 880 mg de minerais. A €5 por cada pacote e cada embalagem, esta combinação custaria €280. Logo, a compra óptima continua a ser a anterior, ou seja, a compra de 40 embalagens de Liquex e nenhuma de Solido, com um custo total de €200, embora tenha mais 1120 mg de vitaminas do que os 880 mg que são necessários.



Em última análise, a combinação óptima depende dos custos relativos dos suplementos. Se o custo por pacote de Solido descer para €2 e o de Liquex subir para €7, os custos para os pontos nos cantos *A* (0, 40), *B* (48, 8) e C (88, 0) serão respectivamente €280, €152 e €176.

A melhor compra de Tanya, com estes preços, é 48 pacotes de Solido e 8 embalagens de Liquex, com um custo de €152.

História Em 1947, o matemático norte-americano George Dantzig, que na altura trabalhava para a Força Aérea dos EUA, formulou um método para resolver problemas de programação linear chamado «método simplex». Este foi tão bem-sucedido, que Dantzig se tornou conhecido no Ocidente como o pai da programação linear. Na Rússia Soviética, excluída durante a Guerra Fria, Leonid Kantorovich formulou independentemente uma teoria de programação linear. Em 1975, Kantorovich e o matemático holandês Tjalling Koopmans ganharam o Prémio Nobel da Economia pelo trabalho em alocação de recursos, que incluía técnicas de programação linear.

Tanya só considerou dois suplementos – duas variáveis –, mas os problemas de hoje que envolvem milhares de variáveis são comuns. Quando Dantzig instituiu o seu método, havia poucos computadores, mas existia o Projecto das Tabelas Matemáticas – um trabalho de uma década na criação de um programa, iniciado em Nova Iorque em 1938. Foi necessária uma equipa de dez calculadoras humanas a trabalhar durante 12 dias com calculadoras de mão para resolver o problema da dieta com nove requisitos de «vitaminas» e 77 variáveis.

Embora o método simplex e as suas variantes tenham sido espantosamente bem-sucedidos, outros métodos têm sido tentados. Em 1984 o matemático indiano Narendra Karmarkar deduziu um novo algoritmo de importância prática, e o russo Leonid Khachiyan propôs um de grande importância teórica.

O modelo da programação linear básica tem sido aplicado a muitas situações além da escolha de uma dieta. Um tipo de problema é o do transporte de bens das fábricas para os armazéns.
O problema tem uma estrutura especial e tornou-se uma área por direito próprio, cujo objectivo é minimizar os custos de transporte. Nalguns problemas de programação linear o objectivo é a maximização, por exemplo, do lucro. Noutros problemas, as variáveis só podem tomar valores inteiros ou apenas dois valores 0 ou 1, mas estes problemas são bastantes diferentes e requerem os seus próprios procedimentos de resolução.

Falta saber se Tanya Smith ganha a sua medalha de ouro nos Jogos Olímpicos. Se conseguir, a programação linear terá outro triunfo.

a ideia resumida
Manter a saúde ao menor preço

46 O caixeiro--viajante

James Cook, com base em Bismark (Dakota do Norte, EUA), é um supervendedor da companhia Electra, um fabricante de produtos de limpeza de tapetes. O facto de ter sido eleito vendedor do ano em três anos seguidos é uma evidência da sua capacidade. A sua área de vendas inclui as cidades de Albuquerque, Chicago, Dallas e El Paso, e Cook visita cada uma delas numa viagem de ida e volta uma vez por mês. A questão que põe a si próprio é como fazer a viagem e, ao mesmo tempo, minimizar a distância viajada. É o problema clássico do caixeiro-viajante.

James tinha elaborado uma tabela de distâncias entre as cidades. Por exemplo, a distância entre Bismark e Dallas é 1020 milhas, encontrada na intersecção (sombreada) da coluna Bismark com a linha Dallas.

Albuquerque				
883	Bismarck			
1138	706	Chicago		
580	1020	785	Dallas	
236	1100	1256	589	El Paso

O método ganancioso (*greedy*) Sendo uma pessoa prática, James Cook desenha um mapa da área de vendas, mas não se preocupa com a precisão, desde que o mapa lhe diga onde ficam as cidades e as distâncias entre elas. Um caminho que faz muitas vezes começa em Bismark para Chicago, Albuquerque, Dallas e El Paso antes de regressar a Bismark. Este é o caminho BCADEB, mas Cook dá-se conta de que esta viagem de 4113 milhas no total é dispendiosa em termos de distância percorrida. Poderá fazer melhor?

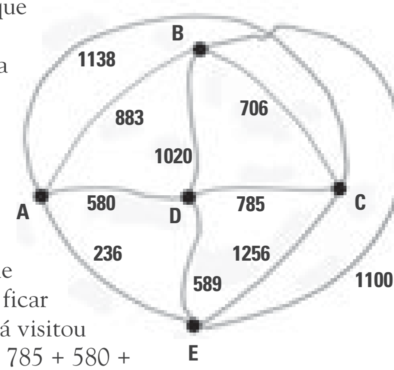
Fazer um plano da área de vendas não deve disfarçar o facto de James não estar com disposição para um planeamento detalhado – ele quer é chegar e vender.

Cronologia

cerca de 1810
Charles Babbage menciona o problema como um problema interessante

1831
O problema do caixeiro-viajante aparece como um problema prático

1926
Borůvka introduz o algoritmo ganancioso (*greedy*)



Olhando para um mapa no seu escritório em Bismark, Cook vê que a cidade mais próxima é Chicago. Está a 706 milhas, enquanto Albuquerque fica a 883, Dallas a 1020, e El Paso a 1100. Começa de imediato por Chicago sem um plano global. Quando chega a Chicago e completa os seus negócios, procura a cidade seguinte. Escolhe Dallas em detrimento de Albuquerque e El Paso, porque esta fica a 785 milhas de Chicago, e é mais perto do que as outras opções.

Uma vez em Dallas, já acumulou 706 + 785 milhas. Tem então de escolher entre Albuquerque e El Paso. Escolhe Albuquerque por ficar mais perto. De Albuquerque tem de ir para El Paso, após o qual já visitou todas as cidades e volta a Bismark. O seu total de milhas é 706 + 785 + 580 + 236 + 1100 = 3407. O caminho BCDAEB é mais curto do que o anterior e também poupou em emissões de carbono.

Esta forma de pensar é muitas vezes referida como o método ganancioso de encontrar um caminho mais curto. Isto porque a decisão de James Cook foi sempre uma decisão local – estar numa certa cidade e procurar o melhor caminho para sair dela. Com este método, Cook nunca tenta olhar mais longe do que um passo de cada vez. Não é um método estratégico, porque não tem uma consideração global sobre o melhor caminho. O facto de Cook ter terminado em El Paso significou que foi forçado a fazer uma longa viagem de regresso a Bismark. Desta forma, encontrou um caminho mais curto, mas será o mais curto? James está intrigado.

James vê como pode tirar vantagem de haver apenas cinco cidades envolvidas. Com tão poucas, é possível listar todos os caminhos possíveis e escolher o mais curto. Com cinco cidades, existem apenas 24 caminhos para examinar ou apenas 12, se contarmos um caminho e o seu inverso como equivalentes, o que se pode fazer dado ambos terem o mesmo total de milhas. O método serve a James Cook, que aprende que o caminho BAEDCB (ou o seu inverso BCDEAB) é na realidade óptimo, com apenas 3199 milhas.

De volta a Bismark, James percebe que a sua viagem está a demorar muito tempo. Não é na distância que ele quer poupar, é no tempo. Elabora uma nova tabela que lhe dá os tempos de viagem entre as diferentes cidades da sua área.

Albuquerque				
12 (estrada)	Bismarck			
6 (ar)	2 (ar)	Chicago		
2 (ar)	4 (ar)	3 (ar)	Dallas	
4 (estrada)	3 (ar)	5 (ar)	1 (ar)	El Paso

1954

Dantzig e Dijkstra propõem métodos para resolver o problema do caixeiro-viajante

1971

Cook formula o conceito de *P versus NP* para algoritmos

2004

David Applegate resolve o problema para as 24 978 cidades da Suécia

Quando o problema estava concentrado nas milhas, James sabia que a soma das distâncias ao longo de dois lados de um triângulo é sempre maior do que o comprimento do terceiro lado; neste caso, o gráfico chama-se euclidiano e sabemos muito sobre os métodos de solução. Não é o caso quando o problema é de tempo. Voar em rotas principais é muitas vezes mais rápido do que seguir caminhos secundários e James Cook repara que a viagem entre El Paso e Chicago é mais rápida passando por Dallas. A chamada desigualdade triangular não se aplica.

O método ganancioso aplicado ao problema do tempo produz um total de 22 horas no caminho BCDEAB, enquanto há dois caminhos distintos óptimos BCADEB e BCDAEB que totalizam 14 horas cada um. Destes dois caminhos, o primeiro tem 4113 milhas e o segundo 3407 milhas. James Cook está feliz porque, escolhendo o BCDAEB, poupou o máximo. No futuro irá considerar o caminho com o menor custo.

De segundos a séculos A real dificuldade associada ao problema do caixeiro-viajante ocorre quando há um grande número de cidades. Como é um funcionário brilhante, James Cook não tarda a ser promovido a supervisor. Tem agora de visitar 13 cidades a partir de Bismark, em vez das anteriores 4. Não está satisfeito com o método ganancioso e prefere analisar uma lista completa de caminhos. Prepara-se para listar os possíveis caminhos para as suas 13 cidades. Cedo descobre que existiriam tantos como $3{,}1 \times 10^9$ caminhos para analisar. Posto de outra forma, se um computador levar um segundo para imprimir um caminho, demorará cerca de um século para os imprimir a todos. Um problema com 100 cidades ocuparia o computador por milénios.

Têm sido aplicados alguns métodos sofisticados ao problema do caixeiro--viajante. Há métodos exactos que se aplicam a 5000 cidades ou menos e um trata mesmo com sucesso um problema particular de 33 810 cidades, embora o poder computacional requerido nesse caso tenha sido colossal. Métodos não exactos produzem caminhos que estão num intervalo óptimo com uma probabilidade específica. Os métodos deste tipo têm a vantagem de poderem lidar com milhões de cidades.

Complexidade computacional Analisando o problema do ponto de vista do computador, vamos pensar apenas no tempo que demoraríamos a encontrar uma solução. Listar simplesmente todos os possíveis caminhos é o pior cenário. James descobriu que o método da força bruta para 13 cidades demoraria quase um século para se completar. Se acrescentarmos 2 cidades, o tempo subirá para 20 000 anos!

Claro que estas estimativas dependerão do computador usado, mas para *n* cidades o tempo aumenta em linha com *n* factorial (o número que se obtém multiplicando todos os inteiros de 1 a *n*). Calculamos $3{,}1 \times 10^9$ caminhos para 13 cidades. Decidir se cada caminho é o menor já encontrado torna-se um problema de tempo *factorial* – e será muito tempo.

Há outros métodos para atacar o problema em que o tempo para *n* cidades aumenta 2^n (2 multiplicado por si próprio *n* vezes), logo para 13 cidades será da ordem de 8192 decisões (8 vezes mais do que para 10 cidades). Um método com esta complexidade é chamado um algoritmo de tempo *exponencial*. O cálice sagrado destes «problemas de optimização combinatória» é encontrar um algoritmo que dependa, não da potência *n* de 2, mas de uma potência fixa de *n*. Quanto menor a potência, melhor; por exemplo, se o algoritmo variar de acordo com n^2, no caso das 13 cidades, isso equivaleria a apenas 169 decisões – menos de metade do tempo necessário para 10 cidades. Um método com esta «complexidade» diz-se ser conduzido em tempo *polinomial* – os problemas resolvidos desta maneira são «problemas rápidos» e podem demorar 3 minutos, em vez de séculos.

A classe de problemas que podem ser *resolvidos* por um computador em tempo polinomial é representada por *P*. Não sabemos se o problema do caixeiro-viajante é um deles. Ninguém apresentou ainda um algoritmo de tempo polinomial para o resolver, mas também ninguém foi capaz de provar que não existe nenhum.

Uma classe mais vasta representada por *NP* consiste em problemas cujas soluções podem ser *verificadas* em tempo polinomial. O problema do caixeiro-viajante é definitivamente um deles, porque pode verificar-se em tempo polinomial se um *dado* caminho tem a menor distância relativamente a qualquer outro. Somam-se simplesmente as distâncias ao longo de um caminho dado e comparam-se com o número dado. *Encontrar* e *verificar* são duas operações diferentes: por exemplo, é fácil verificar que $167 \times 241 = 40\,247$ mas encontrar os factores de 40 247 é uma proposição diferente.

Será que *todos* os problemas verificáveis por tempo polinomial podem ser resolvidos em tempo polinomial? Se isto fosse verdade, as classes *P* e *NP* seriam idênticas e poderíamos escrever $P = NP$. Se $P = NP$ é a pergunta mais feita entre os cientistas da computação. Mais de metade dos profissionais pensa que não são iguais: acreditam que existem problemas que podem ser verificados em tempo polinomial, mas não podem ser resolvidos em tempo polinomial. É um problema de tal forma excepcional, que o Instituto de Matemática Clay oferece um prémio de $ 1 000 000 a quem provar que $P = NP$ ou $P \neq NP$.

a ideia resumida
Encontrar o melhor caminho

47 Teoria dos jogos

Alguns diziam que o pequeno John era a pessoa mais inteligente viva. John Neumann foi uma criança-prodígio que se tornou uma lenda no mundo matemático. Quando souberam que ele tinha chegado a uma reunião de táxi e acabado de rabiscar o seu «teorema minimax» da teoria dos jogos, não ficaram surpreendidos. Era exactamente o tipo de coisa que von Neumann fazia. Newmann contribuiu para a mecânica quântica, a lógica, a álgebra, logo porque deveria a teoria dos jogos escapar-lhe? Não escapou – com Oskar Morgenstern, foi co-autor da influente *Teoria dos Jogos e Comportamento Económico*. No seu sentido mais lato, a teoria dos jogos é um assunto antigo, mas von Neumann foi a chave para aprimorar o «jogo de duas pessoas de soma zero».

Jogos de duas pessoas de soma zero Parece complicado, mas um jogo de duas pessoas de soma zero é simplesmente um jogo «jogado» por duas pessoas, empresas ou equipas, em que um dos lados ganha e o outro perde. Se A ganha €200, B perde €200: é isto que significa soma zero. Não tem sentido A cooperar com B – é uma competição pura, apenas com vencedores e perdedores. Na linguagem «todos ganham», A ganha €200 e B ganha €-200 e a soma 200 + (–200) = 0. Esta é a origem da expressão «soma zero».

Imaginemos duas empresas de televisão, ATV e BTV, que licitam a operação de um serviço noticioso extra ou na Escócia ou na Inglaterra. Cada empresa deve fazer uma licitação para um único país e basear a sua decisão no aumento de audiências previsto. Os analistas dos *media* estimaram o aumento das audiências e ambas as empresas têm acesso ao estudo. O aumento é convenientemente previsto numa «tabela de retorno» e medido em unidades de um milhão de espectadores.

Cronologia

1713
Waldegrave dá a primeira solução matemática de um jogo de dois jogadores

1944
Von Neumann e Morgenstern publicam *Teoria dos Jogos e Comportamento Económico*

		BTV Escócia	BTV Inglaterra
ATV	Escócia	+5	−3
ATV	Inglaterra	+2	+4

Se tanto a ATV como a BTV decidirem operar na Escócia, a ATV ganhará 5 milhões de espectadores, mas a BTV perderá 5 milhões de espectadores. O significado do sinal de menos, como no retorno -3, é que a ATV *perderá* uma audiência de 3 milhões. Os retornos + são bons para a ATV e os retornos – são bons para a BTV.

Presumamos que as empresas tomam as suas decisões *únicas* com base na tabela de retorno e que fazem as suas ofertas, simultaneamente, em carta fechada. Obviamente que ambas agem no seu melhor interesse.

Se a ATV escolhe a Escócia, o pior que poderá acontecer será a perda de 3 milhões; se licitar por Inglaterra, o pior será um ganho de 2 milhões. A estratégia óbvia para a ATV será escolher Inglaterra (linha 2). Não pode acontecer-lhe pior do que ganhar 2 milhões de espectadores, seja o que for que a BTV escolha. Analisando numericamente, a ATV consegue o -3 e 2 (os mínimos das linhas) e escolhe a linha que corresponde ao máximo deles.

A BTV está numa posição mais fraca, mas ainda pode conseguir uma estratégia que limite o seu potencial de perdas e esperar por uma melhor tabela de retorno no ano seguinte. Se a BTV escolher a Escócia (coluna 1), o pior que pode acontecer é uma perda de 5 milhões; se escolher Inglaterra, o pior será uma perda de 4 milhões. A estratégia mais segura para a BTV é escolher Inglaterra (coluna 2), porque perderá uma audiência de 4 milhões em vez de 5 milhões. Não poderá acontecer-lhe pior do que perder 4 milhões, seja o que for que a ATV decida.

Estas serão as estratégias mais seguras para cada jogador, e, se seguidas, a ATV ganhará 4 milhões de espectadores, enquanto a BTV os perderá.

Uma mente brilhante

John F. Nash (n. 1928), cuja vida conturbada foi retratada em 2001 no filme *Uma Mente Brilhante*, ganhou o Prémio Nobel da Economia em 1994, pelas suas contribuições para a teoria dos jogos.

Nash e outros alargaram a teoria dos jogos ao caso de mais de dois jogadores e a jogos em que ocorre a cooperação entre jogadores, incluindo o acrescento um terceiro jogador. O «equilíbrio de Nash» (como um equilíbrio de ponto de sela) deu uma perspectiva muito mais alargada do que a estabelecida por von Neumann, resultando numa maior compreensão das situações económicas.

1950
Tucker coloca o dilema do prisioneiro e Nash propõe o equilíbrio de Nash

1982
Maynard Smith publica *Evolução e a Teoria dos Jogos*

1994
Nash recebe o Prémio Nobel da Economia pelo seu trabalho sobre a teoria dos jogos

Quando é que um jogo está determinado? No ano seguinte, as duas empresas de televisão têm outra opção – operar no País de Gales. Como as circunstâncias mudaram, existe uma nova tabela de retorno.

		BTV			
		P. Gales	Escócia	Inglaterra	mínimo das linhas
ATV	País de Gales	+3	+2	+1	+1
	Escócia	+4	−1	0	−1
	Inglaterra	−3	+5	−2	−3
	máximo das colunas	+4	+5	+1	

Como antes, a estratégia segura para a ATV é escolher a linha que maximiza o pior que pode acontecer. O máximo de {+1, –1, –3} é a escolha do País de Gales (linha 1). A estratégia segura para a BTV é escolher a coluna que minimiza {+4, +5, +1}, ou seja, Inglaterra (coluna 3).

Escolhendo o País de Gales, a ATV *garante* ganhar não menos de 1 milhão de espectadores, qualquer que seja a escolha da BTV, e, escolhendo Inglaterra (coluna 3), a BTV *garante* perder não mais de 1 milhão de espectadores, seja o que for que a ATV faça. Estas escolhas representam portanto as melhores estratégias para cada empresa, e é neste sentido que o jogo está determinado (mas ainda é injusto para a BTV). Neste jogo, o

$$\text{máximo de } \{+1, -1, -3\} = \text{mínimo de } \{+4, +5, +1\}$$

e ambos os lados da equação têm o valor comum +1. Ao contrário do primeiro jogo, esta versão tem um «ponto de sela» de equilíbrio de +1.

Jogos repetitivos O jogo repetitivo icónico é o jogo tradicional «pedra, papel e tesoura». Ao contrário do jogo das empresas de televisão, que é um jogo de uma só jogada, o «pedra, papel e tesoura» é normalmente jogado meia dúzia de vezes, ou algumas centenas de vezes pelos competidores do seu campeonato mundial anual.

No jogo do «papel, tesoura, pedra», dois jogadores mostram uma mão, dois dedos ou um punho, cada um simbolizando papel, tesoura e pedra, respectivamente. Jogam simultaneamente à contagem de três: o papel empata com o papel, perde com a tesoura (dado que a tesoura corta o papel), mas ganha à pedra (porque pode embrulhá-la). Ao jogar «papel», os retornos são então 0, –1, +1, na coluna de cima da tabela completa de retornos.

	papel	tesoura	pedra	mínimo linhas
papel	empate = 0	perde = −1	ganha = +1	−1
tesoura	ganha = +1	empate = 0	perde = −1	−1
pedra	perde = −1	ganha = +1	empate = 0	−1
máximo colunas	+1	+1	+1	

Não há nenhum ponto de sela para este jogo nem nenhuma estratégia *pura* óbvia a adoptar.

Se um jogador escolher sempre a mesma acção, digamos papel, o oponente detecta-o e simplesmente escolherá sempre tesoura. Pelo «teorema minimax» de von Neumann, há uma «estratégia mista» ou uma forma de escolher diferentes acções, com base em probabilidades.

De acordo com a matemática, os jogadores devem escolher aleatoriamente, mas, no geral, apostar um terço das vezes em cada escolha: papel, tesoura ou pedra. Contudo, a aleatoriedade «cega» pode nem sempre ser o melhor caminho, dado que os campeões do mundo têm sempre maneiras de escolher a sua estratégia com uma pequena volta «psicológica». São bons a prever as escolhas dos seus oponentes.

Quando é que um jogo *não é* de soma zero? Nem todos os jogos são de soma zero – cada jogador tem por vezes a sua própria tabela de retorno. Um exemplo famoso é o «dilema do prisioneiro» criado por A. W. Tucker.

Duas pessoas, **A**ndrew e **B**ertie, são apanhados pela polícia por suspeita de roubo na auto-estrada e fechados em celas separadas, para não poderem conferenciar um com o outro. Os retornos, neste caso as penas de prisão, dependem não só das suas respostas individuais ao interrogatório policial, mas também do conjunto das respostas. Se **A** confessar e **B** não, **A** só cumprirá 1 ano de prisão (da tabela de retorno de **A**), mas **B** sofrerá uma sentença de 10 anos (da tabela de retorno de **B**). Se **A** não confessar mas **B** sim, as sentenças serão ao contrário. Se ambos confessarem, servirão 4 anos de cadeia cada um, mas se nenhum confessar e ambos mantiverem a sua declaração de inocência, sairão impunes!

A		B	
		confessa	não confessa
A	confessa	+4	+1
	não confessa	+10	0

B		B	
		confessa	não confessa
A	confessa	+4	+10
	não confessa	+1	0

Se os prisioneiros pudessem cooperar, tomariam a decisão óptima e não confessariam – a situação em que «todos ganham».

a ideia resumida
Matemática em que todos ganham

48 Relatividade

Quando um objecto se move, o seu movimento é medido relativamente a outros objectos. Se conduzirmos ao longo de uma estrada principal a 110 km/h e vier outro carro a 110 km/h na mesma direcção, a nossa velocidade relativamente a ele é 0. No entanto, viajamos ambos a 110 km/h relativamente ao solo. E a nossa velocidade é 220 km/h relativamente a um carro viajando a 110 km/h na faixa de rodagem contrária. A teoria da relatividade mudou esta forma de pensar.

A teoria da relatividade foi estabelecida pelo físico alemão Hendrik Lorentz no final do século XIX, mas o avanço definitivo foi feito por Albert Einstein em 1905. O famoso artigo de Einstein sobre a relatividade revolucionou o estudo de como os objectos se movem, reduzindo a teoria clássica de Newton, uma magnífica conquista, a um caso especial.

Regresso a Galileu Para descrever a teoria da relatividade, vamos aproveitar uma deixa do próprio mestre: Einstein adorava falar sobre comboios e experiências de pensamento. No nosso exemplo, Jim Diamond viaja num comboio a 60 km/h. Do seu lugar na parte de trás do comboio, caminha até à cafetaria a 2 km/h. A sua velocidade é 62 km/h relativamente ao chão. No regresso, a velocidade de Jim relativa ao chão será de 58 km/h, porque estará a deslocar-se em sentido contrário ao do comboio. Isto é o que nos diz a teoria de Newton. A velocidade é um conceito relativo e o sentido do movimento determina se se soma ou subtrai.

Porque todo o movimento é relativo, chamamos «referencial» ao ponto de observação a partir do qual um movimento em particular é medido.
No movimento unidimensional de um comboio que se move ao longo de uma trajectória rectilínea, podemos pensar num referencial fixo posicionado numa

Cronologia

cerca de 1632
Galileu apresenta as «transformações de Galileu» para a queda dos corpos

1676
Römer calcula a velocidade da luz a partir de observação das luas de Júpiter

1687
Os *Principia* de Newton descrevem as leis clássicas do movimento

estação e a uma distância x e um tempo t em termos deste referencial. A posição 0 é determinada por um ponto marcado na plataforma e o tempo lido no relógio da estação. As coordenadas distância/tempo relativas a esse referencial na estação são (x, t).

Também há um referencial dentro do comboio. Se medirmos a distância do fim do comboio e o tempo pelo relógio de pulso do Jim, teremos outro conjunto de coordenadas $(\bar{x},\bar{t})$. Também é possível sincronizar esses dois sistemas de coordenadas. Quando o comboio passa a marca na plataforma, $x = 0$ e o relógio da estação está em $t = 0$. Se Jim colocar $\bar{x} = 0$ nesse ponto, e puser $\bar{t} = 0$ no seu relógio, passará a existir uma ligação entre as coordenadas.

Quando o comboio passa na estação, Jim vai para a cafetaria. Podemos calcular a que distância se encontra da estação passados cinco minutos. Sabemos que o comboio viaja a 1 km por minuto, logo naquele tempo viajou 5 km *e* Jim andou $\bar{x} = {}^{10}\!/\!_{60}$ de um quilómetro (da sua velocidade de 2 km/h multiplicada pelo tempo ${}^{5}\!/\!_{60}$). Assim, no total Jim está à distância (x), que é $5\,{}^{10}\!/\!_{60}$ km da estação. A relação entre $\bar{x}$ e x é dada, portanto, por $x = \bar{x} + v \times t$ (aqui $v = 60$). Invertendo a equação para saber a distância que Jim viajou relativamente ao referencial na estação, obtemos

$$\bar{x} = x - v \times t$$

O conceito de tempo na teoria clássica newtoniana é um fluxo unidimensional do passado para o futuro. É universal para todos e é independente do espaço. Dado que é uma quantidade absoluta, o tempo de Jim dentro do comboio é o mesmo que o da estação na plataforma t, logo

$$\bar{t} = t$$

Estas duas fórmulas para $\bar{x}$ e $\bar{t}$, dadas em primeiro lugar por Galileu, são tipos de equações chamadas transformações, dado que transformam quantidades de um referencial noutro. De acordo com a teoria clássica de Newton, é espectável que a velocidade da luz obedeça a estas duas transformações de Galileu para $\bar{x}$ e $\bar{t}$.

No século XVII, foi reconhecido que a luz tinha velocidade, e o seu valor aproximado foi medido em 1676 pelo astrónomo dinamarquês Ole Römer. Quando Albert Michelson mediu a velocidade da luz com maior precisão em 1881, determinou que ela era de 300 000 km por segundo. Mais do que isso, tomou consciência de que a transmissão da luz era muito diferente da transmissão do som. Michelson verificou que, ao contrário do nosso observador no comboio em movimento, a direcção dos raios de luz não tem qualquer influência na velocidade da luz. Este resultado paradoxal tinha de ser explicado.

1881
Michelson mede a velocidade da luz com grande precisão

1887
As transformações de Lorentz são escritas pela primeira vez

1905
Einstein publica *Sobre a Electrodinâmica dos Corpos em Movimento*, o artigo que descreve a relatividade especial

1915
Einstein publica *As Equações de Campo da Gravitação*, descrevendo a relatividade geral

$$\alpha = \frac{1}{\sqrt{1 - v^2/c^2}}$$

O factor de Lorentz

A teoria da relatividade especial Lorenz estabeleceu as equações matemáticas que explicam a relação entre distância e tempo quando um referencial se move a uma velocidade constante v relativamente a outro. Estas transformações são muito semelhantes àquelas com que já trabalhámos, mas envolvem um factor (Lorentz) dependente de v e da velocidade da luz, c.

Entra Einstein Einstein lidou com as descobertas de Michelson acerca da velocidade da luz admitindo-as como um postulado

A velocidade da luz tem o mesmo valor para todos os observadores e é independente da direcção.

Se Jim Diamond ligar e desligar uma lanterna enquanto passa pela estação no seu comboio em movimento, dirigindo os raios de luz na direcção em que o comboio se move, medirá a sua velocidade como c. O postulado de Einstein diz que o chefe da estação observando na plataforma também medirá a velocidade dos raios de luz como c, e não como $c + 60$. Einstein também assumiu um segundo princípio:

Um referencial move-se com velocidade constante em relação a outro.

O brilhantismo do artigo de Einstein de 1905 foi em parte devido à forma como o cientista abordou o seu trabalho, motivado pela elegância matemática. As ondas sonoras viajam como vibrações de moléculas no meio através do qual o som é transportado. Outros físicos presumiam que a luz também precisava de algum meio para se transmitir. Ninguém sabia qual era, mas deram-lhe um nome – o éter luminífero.

Einstein não sentiu necessidade da existência do éter luminífero como meio de transmissão da luz. Em vez disso, deduziu as transformações de Lorentz dos dois simples princípios da relatividade e toda a teoria se revelou. Em particular, Einstein mostrou que a energia de uma partícula E é determinada pela equação $E = \alpha \times mc^2$. Para a energia de um corpo em repouso (quando $v = 0$ e portanto $\alpha = 1$), isto conduz à icónica equação que mostra que a massa e a energia são equivalentes:

$$E = mc^2$$

Lorentz e Einstein foram ambos propostos para o Prémio Nobel de 1912. Lorentz já o tinha recebido em 1902, mas Einstein teve de esperar até 1921, quando finalmente recebeu o prémio pelo seu trabalho sobre o efeito fotoeléctrico (que também tinha publicado em 1905). Foi um ano e tanto para o funcionário do Serviço de Registo de Patentes suíço.

Einstein *versus* Newton No que respeita a observações sobre comboios em movimento lento, há apenas uma pequena diferença entre a teoria de relatividade de Einstein e a teoria clássica newtoniana. Nestas situações, a velocidade relativa v é tão pequena comparada com a velocidade da luz, que o valor do factor de Lorentz α é quase 1. Neste caso, as equações de Lorentz são virtualmente as mesmas que as clássicas transformações de Galileu. Logo, para velocidades pequenas, Einstein e Newton concordarão um com o outro. As velocidades e distâncias têm de ser muito grandes para que as diferenças entre as duas teorias se tornem evidentes. Até o TGV francês que bateu recordes não atingiu ainda estas velocidades e demorará muito tempo de desenvolvimento do transporte ferroviário antes de se abandonar a teoria de Newton a favor da de Einstein. As viagens espaciais forçar-nos-ão a preferir Einstein.

A teoria geral da relatividade Einstein publicou a sua teoria geral em 1915. Esta teoria aplica-se ao movimento quando é possível que os referenciais *acelerem* uns em relação aos outros e liga os efeitos da aceleração com os da gravidade.

Usando a teoria geral, Einstein conseguiu prever fenómenos físicos tais como a deflexão de raios de luz por campos gravitacionais de objectos grandes como o Sol. A sua teoria também explicou o movimento dos eixos de rotação de Mercúrio. A precessão não podia ser completamente explicada pela teoria gravitacional de Newton da força exercida sobre Mercúrio pelos outros planetas. Era um problema que preocupava os astrónomos desde a década de 1840.

O referencial apropriado para a teoria geral é o espaço-tempo de quatro dimensões. O espaço euclidiano é plano (tem curvatura zero), mas a geometria do espaço--tempo de quatro dimensões de Einstein (ou geometria de Riemann) é curva. Afasta a força de gravidade de Newton como explicação da atracção dos objectos uns pelos outros. Com a teoria geral da relatividade de Einstein, é a curvatura do espaço-tempo que explica essa atracção. Em 1915 Einstein iniciou outra revolução científica.

a ideia resumida
A velocidade da luz é absoluta

49 O último teorema de Fermat

Podemos somar dois quadrados para obtermos um terceiro quadrado. Por exemplo, $5^2 + 12^2 = 13^2$. Mas poderemos somar dois cubos para obtermos outro cubo? E as potências mais altas? Notavelmente, não podemos. O último teorema de Fermat diz que para quaisquer quatro números inteiros x, y, z e n, não há soluções para a equação $x^n + y^n = z^n$, quando n é maior do que 2. Fermat afirmava ter encontrado uma «prova maravilhosa», atormentando as gerações de matemáticos que lhe sucederam, incluindo um rapaz de dez anos que um dia leu sobre esta caça ao tesouro matemático na biblioteca municipal.

O último teorema de Fermat é acerca de uma equação diofantina, um tipo de equação que coloca um dos mais duros desafios da matemática. Estas equações exigem que as suas soluções sejam números inteiros. Receberam o nome de Diofanto de Alexandria, cuja *Aritmética* se tornou um marco na teoria dos números. Pierre de Fermat era um advogado e funcionário público em Toulouse, em França. Matemático versátil, gozava de alta reputação na teoria dos números, e é mais notavelmente lembrado pela afirmação do seu último teorema, a sua contribuição final para a matemática. Fermat provou-o ou pensava que o tinha provado, e escreveu na sua cópia da *Aritmética* de Diofanto: «Descobri uma prova verdadeiramente maravilhosa, mas a margem é muito pequena para a escrever.»

Fermat resolveu muitos problemas bem conhecidos, mas parece que o último teorema de Fermat não foi um deles. O teorema tem ocupado legiões de matemáticos desde há 300 anos e só foi provado recentemente. Esta prova não

Cronologia

1665
Fermat morre sem deixar registo da sua «maravilhosa prova»

1753
Euler prova o caso de $n = 3$

1825
Legendre e Dirichlet provam independentemente o caso de $n = 5$

1839
Lamé prova o caso de $n = 7$

pode ser escrita em nenhuma margem e as técnicas modernas requeridas para a gerar lançam muitas dúvidas sobre a reivindicação de Fermat.

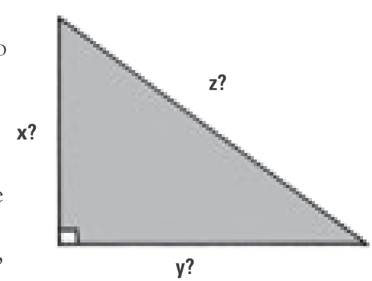
A equação $x + y = z$ Como podemos resolver esta equação em três variáveis x, y e z? Numa equação, temos normalmente um desconhecido x, mas aqui temos três. Na realidade, isto torna a equação $x + y = z$ bastante fácil de resolver. Podemos escolher os valores de x e de y da forma que quisermos, somá-los para obter z e estes três valores darão uma solução. É tão simples quanto isto.

Por exemplo, se escolhermos $x = 3$ e $y = 7$, os valores $x = 3$, $y = 7$ e $z = 10$ são uma solução para a equação. Também podemos constatar que alguns valores de x, y e z não são soluções da equação. Por exemplo, $x = 3$, $y = 7$ e $z = 9$ não são solução, porque não tornam o lado esquerdo da equação $x + y$ igual ao lado direito z.

A equação $x^2 + y^2 = z^2$ Pensemos agora em quadrados. O quadrado de um número é o número multiplicado por si próprio, um número que escrevemos como x^2. Se $x = 3$, $x^2 = 3 \times 3 = 9$. A equação em que estamos a pensar agora é, não $x + y = z$, mas

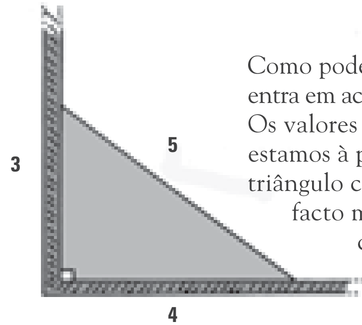
$$x^2 + y^2 = z^2$$

Podemos resolver esta equação como a anterior, escolhendo valores para x e y e calculando z. Com os valores $x = 3$ e $y = 7$, por exemplo, o lado esquerdo da equação é $3^2 + 7^2$, o que é $9 + 49 = 58$. Para isto, z terá de ser a raiz quadrada de 58 ($z = \sqrt{58}$), o que é aproximadamente 7,6158. Temos decerto o direito de reclamar que $x = 3$, $y = 7$ e $z = \sqrt{58}$ é uma solução de $x^2 + y^2 = z^2$, mas infelizmente as equações diofantina concentram-se, em primeiro lugar, nas soluções de números inteiros. Como $\sqrt{58}$ não é um número inteiro, a solução $x = 3$, $y = 7$ e $z = \sqrt{58}$ não serve.



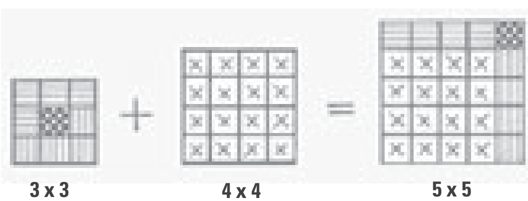
A equação $x^2 + y^2 = z^2$ está relacionada com os triângulos. Se x, y e z representarem os comprimentos dos três lados de um triângulo rectângulo, satisfazem a equação. Inversamente, se x, y e z satisfazem a equação, o ângulo entre x e y é um ângulo recto. Pelas suas relações com o teorema de Pitágoras, as soluções para x, y e z são chamadas triplos pitagóricos.

1843	1907	1908	1994
Kummer reivindica que provou o teorema, mas Dirichlet expõe uma falha	Von Lindemann reivindica uma prova que se mostrou, contudo, errada	Wolfskehl oferece um prémio para soluções durante os 100 anos seguintes	Wiles prova finalmente o problema



Como poderemos encontrar triplos pitagóricos? É aqui que o construtor local entra em acção. Parte do equipamento é o omnipresente triângulo 3-4-5. Os valores $x = 3$, $y = 4$ e $z = 5$ acabam por ser uma solução do tipo da que estamos à procura, porque $3^2 + 4^2 = 9 + 16 = 5^2$. Do lado inverso, um triângulo com as dimensões 3, 4 e 5 tem de incluir um ângulo recto. Este é o facto matemático que o construtor usa para construir as suas paredes com ângulos rectos.

Neste caso, podemos dividir um quadrado 3 × 3, e colocá-lo à volta de um quadrado 4 × 4 para fazer um quadrado 5 × 5.



Há outros números inteiros que são soluções de $x^2+ y^2 = z^2$. Por exemplo $x = 5$, $y = 5$ e $z = 13$ é outra solução, porque $5^2 + 12^2 = 13^2$ e, de facto, existe um número infinito de soluções para a equação. A solução do construtor ocupa o primeiro lugar dado que é a menor solução, e é a única solução constituída por números inteiros consecutivos. Há muitas soluções em que dois dos números são consecutivos, tais como $x = 20$, $y = 21$ e $z = 29$ ou $x = 9$, $y = 40$ e $z = 41$, mas nenhuma com os três.

Da fartura à fome Parece um pequeno passo ir de $x^2+ y^2 = z^2$ para $x^3+ y^3 = z^3$. Logo, seguindo a ideia de recolocar um quadrado à volta de outro para construir um terceiro, podemos usar o mesmo estratagema com um cubo? Podemos recolocar um cubo à volta de outro para obter um terceiro? Acontece que isto não pode ser feito. A equação $x^2+ y^2 = z^2$ tem um número infinito de soluções, mas Fermat foi incapaz de encontrar um único exemplo de números inteiros de $x^3+ y^3 = z^3$. O pior estava para vir, e a ausência de soluções de Leonhard Euler levou-o a enunciar o último teorema:

Não há solução nos números inteiros para a equação $x^n + y^n = z^n$ para todos os valores de n superiores a 2.

Uma maneira de resolver o problema da prova é começar com os valores baixos de n e avançar. Foi desta maneira que Fermat iniciou o trabalho. O caso de $n = 4$ é na realidade mais simples do que o de $n = 3$ e é provável que Fermat o tenha provado. Nos séculos XVIII e XIX, Euler juntou o caso de $n = 3$, Adrien--Marie Legendre completou o caso de $n = 5$ e Gabriel Lamé provou o caso de $n = 7$. Lamé pensou inicialmente que tinha provado o teorema geral, mas infelizmente estava errado.

Ernst Kummer foi um enorme contribuinte e, em 1843, submeteu um manuscrito afirmando que tinha provado o teorema em geral, mas Dirichlet apontou uma falha no argumento. A Academia Francesa ofereceu um prémio de 3000 francos por uma prova válida, finalmente concedendo-o a Kummer pelo seu trabalho meritório. Kummer provou o teorema para todos os números primos menores do que 100 (e outros valores), mas excluindo os primos irregulares 37, 59 e 67. Por exemplo, não conseguiu provar que não existem números inteiros que satisfaçam $x^{67} + y^{67} = z^{67}$. A sua incapacidade de provar genericamente o teorema abriu caminho para valiosas técnicas em álgebra abstracta. Foi talvez uma maior contribuição para a matemática que a resolução da própria questão.

Ferdinand von Lindemann, que provou a impossibilidade da quadratura do círculo (ver página 22), reivindicou ter provado o teorema em 1907. Provou-se, contudo, que ele errara. Em 1908, Paul Wolfskehl legou 100 000 marcos para um prémio a ser concedido ao primeiro que fornecesse uma prova, um prémio disponível por 100 anos. Ao longo dos anos, qualquer coisa como 5000 provas foram submetidas e verificadas, e todas regressaram aos candidatos como falsas.

A prova Embora a ligação ao teorema de Pitágoras só se aplique para $n = 2$, a ligação com a geometria provou ser a chave para esta prova final. A relação foi feita com a teoria das curvas e uma conjectura apresentada pelos japoneses Yukata Taniyama e Goro Shimura. Em 1993, Andrew Wiles deu uma conferência sobre a teoria em Cambridge e incluiu a sua prova do teorema de Fermat. Infelizmente, a prova estava errada.

O matemático francês André Weil, cujo nome é semelhante, ignorou tais tentativas. Comparou a prova do teorema à escalada do Evereste e acrescentou que, se um homem ficasse aquém do topo por 100 jardas, não subira o Evereste. A pressão estava instalada. Wiles fechou-se e trabalhou no problema incessantemente. Muitos pensaram que Wiles se juntaria à multidão de pessoas que quase o tinham resolvido.

No entanto, com a ajuda de colegas, Wiles foi capaz de estripar o erro e substituí-lo por um argumento correcto. Desta vez, convenceu os peritos e provou o teorema. A sua prova foi publicada em 1995 e reclamou o Prémio Wolfskehl mesmo no limite do tempo para se tornar uma celebridade matemática. O rapaz de dez anos que se sentara numa biblioteca pública de Cambridge a ler sobre o problema anos antes tinha feito um longo caminho.

a ideia resumida
Provar um ponto marginal

50 A hipótese de Riemann

A hipótese de Riemann representa um dos mais difíceis desafios da matemática pura. A conjectura de Poincaré e o último teorema de Fermat foram conquistados, mas não a hipótese de Riemann. Uma vez decididas, de uma forma ou de outra, as questões elusivas serão resolvidas e abrir-se-á uma série de novas questões para os matemáticos reflectirem.

A história começa com a adição de fracções do tipo

$$1+\frac{1}{2}+\frac{1}{3}$$

A resposta é $1\frac{5}{6}$ (aproximadamente 1,83). Mas o que acontece se continuarmos a somar fracções cada vez menores, digamos dez?

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}$$

Número de termos	Total (aproximado)
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

Usando apenas uma calculadora simples, estas fracções somam aproximadamente 2,9, em numeração decimal. A tabela mostra como o total cresce conforme mais e mais fracções são adicionadas.

A série de números

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

é chamada série harmónica. O termo «harmónica» teve origem em Pitágoras, que acreditava que uma corda musical dividida a meio, a um terço e a um quarto, produziria as notas essenciais para a harmonia.

Cronologia

1854
Riemann começa a estudar a função zeta

1859
Riemann prova que há soluções cruciais numa faixa crítica e apresenta a sua conjectura

1896
De la Vallée-Poussin e Hadamard mostram que todos os zeros importantes se situam *dentro* da faixa crítica de Riemann

Na série harmónica, vão sendo somadas fracções cada vez menores, mas o que acontece com o total? Cresce para lá de todos os números, ou existirá uma barreira algures, um limite que essa soma nunca ultrapassa? Para responder a isto, o artifício é agrupar os termos, duplicando o número deles enquanto avançamos. Se somarmos os primeiros 8 termos (reconhecendo que $8 = 2 \times 2 \times 2 = 2^3$), por exemplo

$$S_{2^3} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)$$

(em que S significa soma) e, porque ⅓ é maior do que ¼ e ⅕ é maior do que ⅛ (e assim por diante), isto é maior do que

$$1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}$$

Logo podemos dizer

$$S_{2^3} > 1 + \frac{3}{2}$$

e de uma forma geral

$$S_{2^k} > 1 + \frac{k}{2}$$

Se tomarmos $k = 20$, então $n = 2^{20} = 1\,048\,576$ (mais de um milhão de termos), a soma da série só terá excedido 11 (ver tabela). Está a aumentar de uma forma penosamente lenta, mas pode ser escolhido um valor de k para fazer com que o total vá para além de *qualquer* número predefinido, ainda que muito grande. Diz-se que a série diverge para o infinito. Por outro lado, isto não acontece com a série de termos quadrados

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

Continuemos a usar o mesmo processo: somar números cada vez menores, mas desta vez atinge-se um limite, e esse limite é menor do que 2. De forma bastante dramática, esta série converge para $\pi^2/6 = 1{,}64493\ldots$

Na última série, a potência dos termos era 2. Na série harmónica, a potência dos denominadores era silenciosamente igual a 1 e este valor é crítico. Se a potência aumentar um valor minúsculo para um número pouco acima de 1, a série converge, mas se a potência diminuir um valor minúsculo imediatamente abaixo de 1, a série diverge. A série harmónica está na fronteira entre a divergência e a convergência.

1900
Hilbert coloca a hipótese na sua lista de problemas principais para os matemáticos resolverem

1914
Hardy prova que há soluções infinitas ao longo da recta de Riemann

2004
Verifica-se que os primeiros 10 biliões de zeros estão na recta crítica

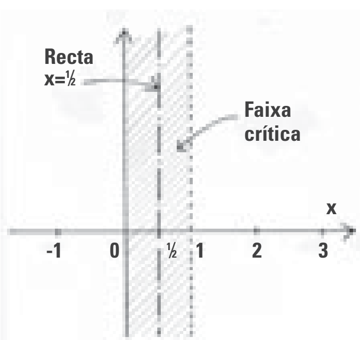
A função zeta de Riemann A famosa função zeta de Riemann $\zeta(s)$ na realidade já era conhecida de Euler no século XVIII, mas Bernhard Riemann reconheceu a sua total importância. O ζ é a letra grega *zeta*, enquanto a função é representada como:

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \ldots$$

Vários valores da função zeta foram calculados, sobretudo $\zeta(1) = \infty$ porque $\zeta(1)$ é a série harmónica. O valor de $\zeta(2)$ é $\pi^2/6$, resultado descoberto por Euler. Tem sido mostrado que os todos os valores de $\zeta(s)$ envolvem π quando s é um numero par, enquanto a teoria de $\zeta(s)$ para números ímpares de s é bastante mais difícil. Roger Apéry provou o resultado importante de $\zeta(3)$ ser um número irracional, mas o seu método não é extensível a $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, etc.

A hipótese de Riemann A variável s na função zeta de Riemann representa uma variável real, mas pode ser alargada para representar um número complexo (ver página 32). Isto permite que as potentes técnicas de análise complexa lhe possam ser aplicadas.

A função zeta de Riemann tem uma infinidade de zeros, ou seja, uma infinidade de valores de s para os quais $\zeta(s) = 0$. Num artigo apresentado na Academia das Ciências de Berlim em 1859, Riemann mostrou que todos os zeros importantes eram números complexos que se situavam numa faixa crítica delimitada por $x = 0$ e $x = 1$. Também colocou a sua famosa hipótese



Todos os zeros da função zeta de Riemann $\zeta(s)$ se situam na recta $x = ½$; a recta ao meio da faixa crítica.

O primeiro passo para realmente resolver esta hipótese foi dado em 1896, de forma independente, por Charles de la Vallée-Poussin e Jacques Hadamard. Ambos mostraram que os zeros devem situar-se no interior da faixa (logo x não pode ser igual a 0 nem a 1).

Em 1914, o matemático inglês G. H. Hardy provou que havia uma infinidade de zeros na recta $x = ½$, embora isso não garanta que não possa haver uma infinidade de zeros que se situem fora dela.

No que respeita aos resultados numéricos, os zeros não triviais calculados até 1986 (1 500 000 000 deles) situam-se na recta $x = ½$, enquanto cálculos actualizados verificaram que isto também se verifica para os primeiros 100 mil milhões de zeros. Embora estes resultados experimentais sugiram que a conjectura é razoável, ainda existe a possibilidade de ser falsa. A conjectura diz que todos os zeros estão nesta recta crítica, mas aguarda-se prova positiva ou negativa.

Porque é que a hipótese de Riemann é importante? Existe uma inesperada relação entre a função zeta de Riemann $\zeta(s)$ e a teoria dos números primos (ver página 36). Os números primos são 2, 3, 5, 7, 11, etc., números só divisíveis por 1 e por si próprios. Usando primos, podemos formar a expressão

$$\frac{1}{\left(1-\frac{1}{2^s}\right)}\times\frac{1}{\left(1-\frac{1}{3^s}\right)}\times\frac{1}{\left(1-\frac{1}{5^s}\right)}\times\ldots$$

e esta é outra forma de escrever $\zeta(s)$, a função zeta de Riemann. Isto diz-nos que o conhecimento da função zeta de Riemann irá lançar luz sobre a distribuição dos números primos e aumentar a nossa compreensão sobre os blocos básicos da construção da matemática.

Em 1900, David Hilbert enunciou os seus famosos 23 problemas para os matemáticos resolverem. Sobre o seu oitavo problema, disse: «Se acordasse depois de ter dormido durante 500 anos, a minha primeira pergunta seria: A hipótese de Riemann já foi provada?»

Hardy usou a hipótese de Riemann como seguro de vida, quando atravessou o mar do Norte depois da sua visita de Verão ao seu amigo Harald Bohr na Dinamarca. Antes de deixar o porto, enviaria um postal ao seu amigo reivindicando que tinha acabado de provar a hipótese de Riemann. Era uma maneira inteligente de aposta dupla. Se o barco se afundasse, teria a honra póstuma de ter resolvido o grande problema. Por outro lado, se Deus de facto existisse, não deixaria que um ateu como Hardy tivesse tal honra e portanto impediria o barco de se afundar.

Quem conseguir resolver de forma rigorosa a questão ganhará um prémio de um milhão de dólares oferecido pelo Instituto de Matemática Clay. Mas o dinheiro não é o chamariz – a maioria dos matemáticos contentar-se-ia em obter o resultado e, com ele, uma posição altíssima no panteão dos grandes matemáticos.

a ideia resumida
O derradeiro desafio

Glossário

Álgebra Trabalhando com letras em vez de números para ampliar a aritmética, a álgebra é hoje um método geral aplicável a toda a matemática e suas aplicações. A palavra «álgebra» deriva de *al-jabr*, forma usada num texto árabe do século IX.

Algoritmo Uma receita matemática; uma rotina definida para resolver um problema.

Axioma Afirmação para a qual não se procura justificação, usada para definir um sistema. O termo «postulado» servia o mesmo propósito para os gregos, mas para eles era uma verdade evidente em si mesma.

Base A base de um sistema de numeração. Os babilónios basearam o seu sistema em 60, enquanto a base moderna é 10 (decimal).

Cardinalidade Número de objectos num conjunto. A cardinalidade do conjunto $\{a, b, c, d, e\}$ é 5, mas a cardinalidade também tem significado nos conjuntos infinitos.

Comutativa A multiplicação em álgebra é comutativa se $a \times b = b \times a$, como na aritmética ordinária (p., ex, $2 \times 3 = 3 \times 2$). Em muitos ramos de álgebra moderna, este não é o caso (p. ex., álgebra matricial).

Conjunto Uma colecção de objectos: por exemplo, o conjunto de peças de mobiliário pode ser F = {*cadeira*, *mesa*, *sofá*, *banco*, *armário*}.

Conjunto vazio Conjunto sem elementos. Tradicionalmente representado por ϕ, é um conceito útil na teoria dos conjuntos.

Contra-exemplo Exemplo simples que refuta uma afirmação. Prova-se que a afirmação «todos os cisnes são brancos» é falsa mostrando um cisne preto como contra-exemplo.

Corolário Consequência secundária de um teorema.

Correspondência «um-para-um» A natureza da relação em que cada objecto num conjunto corresponde exactamente a um objecto no outro e vice-versa.

Denominador Parte inferior de uma fracção. Na fracção $3/7$ o número 7 é o denominador.

Diagrama de Argand Método visual de representar o plano bidimensional dos números complexos.

Diagrama de Venn Método pictórico (diagrama de balões) usado na teoria de conjuntos.

Diferenciação Operação básica do cálculo que dá a derivada ou a taxa de variação. Para uma expressão que descreva como a distância depende do tempo, por exemplo, a derivada representa a velocidade. A derivada da expressão da velocidade representa a aceleração.

Discreto Termo oposto a «contínuo». Há lacunas entre os valores discretos, tais como as lacunas entre os números inteiros 1, 2, 3, 4, ...

Distribuição Intervalo de probabilidade de acontecimentos que ocorrem numa experiência ou numa situação. Por exemplo, a distribuição de Poisson dá a probabilidade de x ocorrências de um determinado acontecimento raro se verificar para cada valor de x.

Divisor Número inteiro que divide exactamente outro número inteiro. O número 2 é divisor de 6 porque $6 \div 2 = 3$. Logo 3 é outro porque $6 \div 3 = 2$.

Eixos x-y Ideia devida a René Descartes de marcar pontos tendo uma coordenada x (eixo horizontal) e uma coordenada y (eixo vertical).

Equação diafontina Equação cujas soluções têm de ser números inteiros ou possivelmente fracções.

Expoente Notação utilizada em aritmética. Multiplicar um número por si próprio, 5×5 escreve-se 5^2 com expoente 2. A expressão $5 \times 5 \times 5$ escreve-se 5^3, etc. A notação pode ser ampliada: por exemplo, o número $5^{1/2}$ significa a raiz quadrada de 5. A potência e o índice são termos equivalentes.

Fracção Número inteiro dividido por outro, por exemplo $3/7$.

Fracções unitárias Fracções com parte superior (numerador) igual a 1. Os antigos egípcios basearam em parte o seu sistema de numeração em fracções unitárias.

Geometria Trabalhando com as propriedades das linhas, formas e espaços, a geometria foi formalizada nos *Elementos* de Euclides no século III a.C. A geometria atravessa toda a matemática e hoje em dia já perdeu o seu restrito significado histórico.

Hipótese Afirmação provisória que espera ser provada ou refutada. Tem a mesma condição de uma conjectura.

Integração Operação básica de cálculo que mede a área. Pode provar-se que é a operação inversa da diferenciação.

Iteração Partir de um valor a e repetir uma operação. Por exemplo, começando com 3 e somando 5 repetidamente, temos a sequência iterada 3, 8, 13, 18, 23, ...

Lema Afirmação provada que abre caminho à prova de um teorema mais abrangente.

Matriz Conjunto de números ou símbolos dispostos num quadrado ou num rectângulo. As matrizes podem ser somadas e multiplicadas e formam um sistema algébrico.

Máximo divisor comum, *mdc* O *mdc* de dois números é o maior número que divide ambos exactamente. Por exemplo, 6 é o *mdc* dos números 18 e 84.

Numerador Parte superior de uma fracção. Na fracção 3/7, 3 é o numerador.

Número primo Número inteiro que tem como divisores apenas ele próprio e 1. Por exemplo, 7 é primo, mas 6 não é (porque 6 ÷ 2 = 3). É habitual começar a sequência dos números primos com 2.

Número quadrado Resultado de multiplicar um número inteiro por si próprio. O número 9 é um quadrado, porque 9 = 3 × 3. Os números quadrados são 1, 4, 9, 16, 25, 36, 49, 64, ...

Número transcendental Número que não pode ser solução de uma equação algébrica, como $ax^2 + bx + c = 0$ ou uma em que x tenha um expoente ainda mais elevado. O π é um número transcendental.

Números imaginários Números que envolvem o «imaginário» $i = \sqrt{-1}$. Contribuem para a formação dos números complexos quando combinados com os números ordinários (ou «reais»).

Números irracionais Números que não podem ser representados por uma fracção. (p. ex., a raiz quadrada de 2).

Números racionais Números que são números inteiros ou fracções.

Poliedro Forma sólida com muitas faces. Por exemplo, um tetraedro tem quatro faces triangulares e um cubo, seis faces quadradas.

Primos gémeos Dois números primos separados no máximo por um número. Por exemplo, os gémeos 11 e 13. Não se sabe se haverá uma infinidade destes primos.

Quadratura do círculo Problema de construir um quadrado com a mesma área de um dado círculo, usando apenas uma régua para desenhar linhas rectas e um compasso para desenhar circunferências. Impossível de fazer.

Quaterniões Números imaginários de quatro dimensões descobertos por W. R. Hamilton.

Raiz quadrada Número que, quando multiplicado por si próprio, iguala um número dado. Por exemplo, 3 é a raiz quadrada de 9 porque 3 × 3 = 9.

Resto Se um número inteiro é divido por outro número inteiro, o número que sobra é o resto. O número 17 dividido por 3 dá 5 com resto 2.

Secção cónica Nome genérico para a família clássica de curvas que incluem circunferências, rectas, elipses, parábolas e hipérboles. Cada uma destas curvas é determinada por cortes de um cone.

Sequência Fila (possivelmente infinita) de números ou símbolos.

Série Fila (possivelmente infinita) de números ou símbolos somados.

Simetria Regularidade de uma forma. Se a forma pode ser rodada de forma a coincidir com a sua forma original, diz-se que tem simetria rotacional. Uma figura tem uma simetria de espelho, se a sua reflexão coincide com a sua forma original.

Sistema de numeração binário Sistema de numeração baseado em dois símbolos, 0 e 1, fundamental para cálculos com computador.

Sistema de posição A grandeza de um número depende da posição dos seus dígitos. Em 73, a posição de 7 significa «7 dez» e 3, «3 unidades».

Sistema hexadecimal Sistema de numeração de base 16 baseado em 16 símbolos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, e F. É largamente utilizado em computação.

Solução óptima Muitos problemas requerem a melhor solução ou a solução óptima. Pode ser uma solução que minimiza custos ou maximiza lucros, como acontece na programação linear.

Teorema de Pitágoras Se os lados de um triângulo rectângulo têm por comprimentos x, y e z, então $x^2 + y^2 = z^2$, em que z é o comprimento do lado maior (a hipotenusa) oposto ao ângulo recto.

Teorema Nome dado a um facto comprovado de certa importância.

Teoria do caos Teoria de sistemas dinâmicos que parecem aleatórios mas têm uma regularidade subjacente.

Índice remissivo

Título: *50 Ideias de Matemática Que Precisa mesmo de Saber*
Título original: *50 Mathematical Ideas You Really Need to Know*

Revisão: Rita Almeida Simões

Adaptação da capa: Transfigura.design

1.ª edição: Outubro de 2011

ISBN: 9789722048507

Publicações D. Quixote
Uma editora do Grupo Leya
Rua Cidade de Córdova, n.º 2
2610-038 Alfragide – Portugal
www.dquixote.pt
www.leya.com