Matemática Recreativa

Editor:
António Machiavelo

O Problema do Totobola

António Machiavelo e **Rogério Reis**

Departamentos de Matemática Pura e de Ciência dos Computadores

da Faculdade de Ciências da Universidade do Porto

`ajmachia@fc.up.pt` `rvr@ncc.up.pt`

1. Uma revelação a meio da noite

Há já um quarto de século, um de nós foi acordado a meio da noite por um seu familiar tomado por uma frenética vontade de partilhar a sua constatação que, num totobola com três partidas, o número mínimo de apostas simples para garantir dois resultados certos, é menor que nove! No dia seguinte, a horas muito mais decentes, e apesar de em estados diferentes de repouso, já ambos os autores desta coluna estavam cientes deste facto. E desde essa altura o «problema do totobola» não deixa de nos apoquentar o sono...

Mas, comecemos pelo princípio... O totobola, um jogo de apostas hoje quase em vias de extinção, consiste no desafio de tentar adivinhar o resultado qualitativo de um certo número de partidas de futebol*. Assim, para cada uma dessas partidas, que se disputa entre as equipas A e B, o apostador tem de escolher 1, X ou 2, consoante aposte que a equipa A ganha, empata ou perde. Para que a leitura óptica das apostas seja facilitada, coloca-se uma

*Eram 13 no enquadramento histórico do parágrafo anterior, mais tarde passaram a 14 e o actual «super 14» corresponde, para todos os efeitos, a um desafio com 15 partidas.

	1	X	2
A-B		X	
C-D		X	
E-F	X		

	1	X	2
A-B			X
C-D	X		
E-F		X	

	1	X	2
A-B	X		
C-D			X
E-F	X		

Figura 1: Algumas apostas simples num totobola com 3 partidas

cruz (de facto um X) na coluna correspondente à aposta pretendida. Esta forma de assinalar as apostas permite, para além disso, abreviar apostas múltiplas (duplas e triplas) colocando mais do que um X na mesma linha. Na figura 1 dá-se exemplo de três apostas simples para um totobola envolvendo três partidas. A um conjunto de resultados para as partidas constituintes de um jogo chama-se uma chave.

O problema do totobola pode então ser formulado da seguinte maneira:

> *Num jogo de totobola envolvendo n partidas, qual o número mínimo de apostas (simples) que é necessário fazer para garantir, pelo menos, $n-1$ resultados certos?*

Designe-se por d_n esse número mínimo de apostas simples que garantem $n-1$ resultados correctos num totobola envolvendo n partidas. Observe-se que o número total de possíveis chaves é 3^n e que cada aposta simples cobre $2n+1$ chaves: aquela em que se acertam em todos os resultados e as que diferem dessa aposta no resultado de uma só partida, o que, por sua vez, pode acontecer de duas maneiras. Tem-se então:

$$d_n \geq \frac{3^n}{2n+1}.$$

É claro que $d_{n+1} \leq 3d_n$, pois se se tiver um conjunto que garanta $n-1$ resultados num totobola com n partidas, basta a cada uma dessas apostas acrescentar uma tripla para obter um conjunto (com 3 vezes mais elementos) que garante n resultados num totobola com $n+1$ partidas. Assim, para $n \geq 3$,

$$\frac{3^n}{2n+1} \leq d_n \leq 3d_{n-1}.$$

Em particular, $d_n \leq 3^{n-1}$, para todo n. Apesar de esta majoração trivial não poder ser melhorada para $n=2$ (porquê?), já para $n=3$ é possível fazer bastante melhor. Sugerimos que antes de prosseguir, o leitor tente encontrar um conjunto de menos de 9 apostas que assegure dois em três resultados no totobola com três partidas.

2. Uma formulação geométrica

É claro que os símbolos 1X2 usados no totobola para codificar vitória-empate-derrota são inteiramente arbitrários, podendo ser substituídos por outros quaisquer. Usando os símbolos 012, uma aposta no totobola com 3 partidas corresponde naturalmente a um ponto de $\mathbb{R}^3$ e o conjunto das chaves, $\{(a, b, c) : a, b, c \in \{0, 1, 2\}\}$, é o conjunto dos vértices, pontos médios das arestas e pontos médios das faces de um cubo de lado 2, assim como o ponto médio do próprio cubo. A figura 2 representa assim o espaço de todas

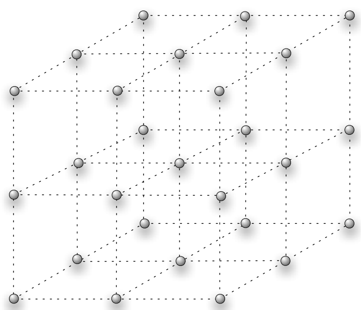

Figura 2: O espaço das chaves do totobola com 3 partidas

as chaves do totobola com 3 partidas. Por exemplo, as apostas da figura 1 correspondem aos pontos $(1, 1, 0)$, $(2, 0, 1)$ e $(0, 2, 0)$, respectivamente.

Uma chave é coberta por uma dada aposta, ou seja, pelo menos dois dos seus resultados coincidem com os da aposta, se pelo menos duas das coordenadas dos respectivos pontos forem iguais. Isto corresponde a dizer que a recta que une os dois pontos, no caso de serem distintos, é paralela a um dos eixos (ver figura 3). Deste modo é relativamente fácil encontrar a

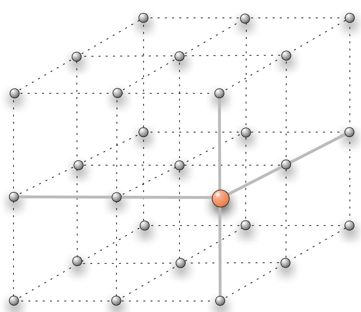

Figura 3: Uma aposta e as chaves que ela cobre

solução para o caso $n = 3$. Consegue o leitor fazê-lo? Uma solução óptima é apresentada no fim deste artigo, ficando como exercício a prova que essa solução é minimal.

Para prosseguir esta abordagem para totobolas com mais do que três partidas, temos de mergulhar em espaços de dimensão superior. Não deixa de ser curioso que a análise de um totobola com umas meras quatro partidas conduza assim à quarta dimensão! Desafiámos aqui o leitor a encontrar uma solução para $n = 4$ usando uma figura em que represente o respectivo espaço das chaves como um cubo quadri-dimensional, que poderá ser pensado, por exemplo, como três cópias da figura 2, em que há rectas (que é preferível não desenhar) ligando pontos que têm as mesmas coordenadas internas em cada cubo.

Mas esta formulação geométrica do problema não é, obviamente, a única. De facto uma outra formulação integra este problema num conjunto muito mais vasto de problemas similares. Consideremos o grafo cujo conjunto de vértices é constituído pelas chaves de um jogo de totobola e em que existe uma aresta entre dois vértices se e só se as chaves correspondentes diferem somente no resultado de uma partida. O problema do totobola, nesta versão, corresponde em saber qual o número mínimo de vértices para os quais, juntamente com os seus vizinhos directos (aqueles que a eles estão ligados por uma aresta) cobrem todo o grafo. Este tipo de problemas, chamados problemas de dominância em grafos, é, em geral, de solução muito difícil, e para muitos deles não se conhece solução apesar de profusamente estudados na Teoria de Grafos (ver, por exemplo, [Hay98]).

3. Códigos de Hamming sobre $\mathbb{F}_3$

Considerar o problema do totobola para n partidas como um problema geométrico em $\mathbb{R}^n$ tem óbvias vantagens, como se pode constatar no caso $n = 3$. Mas, o espaço mais natural para lidar com esta questão é $\mathbb{F}_3^n$, onde $\mathbb{F}_3$ denota o (único, a menos de isomorfismo) corpo com três elementos. Fica-se assim com uma estrutura de espaço vectorial no espaço das chaves, o que permite usar as poderosas ferramentas da Álgebra Linear para atacar o problema.

Por uma daquelas coincidências que evidenciam a natureza multifacetada das boas abordagens, Richard Hamming, em 1947–48, ao introduzir códigos correctores de erros para «ensinar» um computador a corrigir erros simples e desse modo tornar o seu trabalho mais eficiente (ver [Tho83]), resolveu, sem o saber, uma infinidade de casos do problema do totobola†. Para descrevermos como, comecemos por introduzir as noções de *peso* e *distância de Hamming* em $\mathbb{F}_3^n$: o *peso de Hamming* de um vector é simplesmente o número de entradas não nulas deste, enquanto a *distância de Hamming* de dois vectores é o número de coordenadas distintas que estes têm, ou, o que é o mesmo, é o peso de Hamming da diferença desses vectores. É fácil ver que esta noção fornece de facto uma distância em $\mathbb{F}_3^n$. Tem-se:

Lema 1 *Se A é uma matriz com entradas em $\mathbb{F}_3$ tal que quaisquer r das suas colunas são linearmente independentes, então a distância de Hamming de dois quaisquer vectores do seu núcleo é maior ou igual a $r+1$.*

Demonstração. Como o núcleo é um subespaço linear, basta mostrar que o peso de Hamming de um seu qualquer vector é maior ou igual a $r+1$. Mas um vector do núcleo é constituído pelos coeficientes de uma combinação linear das colunas da matriz A e portanto, por hipótese, tem de ter pelo menos $r+1$ coordenadas não nulas. □

Agora, dado $m \in \mathbb{N}$, e uma vez que dois vectores de $\mathbb{F}_3^m$ são linearmente independentes se e só se não são múltiplos um do outro, a matriz cujas colunas são todos os vectores de $\mathbb{F}_3^m$ cuja primeira entrada não-nula é 1,

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 2 & 2 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 2 & 2 & 2 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & \cdots & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & \cdots & 0 & 1 & 2 \end{pmatrix}$$

satisfaz as condições da Proposição anterior. É fácil ver que o número de colunas desta matriz é $1+3+3^2+\cdots+3^{m-1}=\frac{3^m-1}{2}$. Daqui resulta:

Proposição 1 *Se n é um número da forma $\frac{3^m-1}{2}$, para algum $m \in \mathbb{N}$, então $d_n = 3^{n-m}$.*

†Aparentemente o problema do totobola foi pela primeira vez investigado por Olga Taussky e John Todd em 1948 (ver [Tho83], p. 84), em [Tau48]. Este artigo levou à descoberta do resultado descrito na proposição abaixo, de modo independente e no contexto da teoria dos grupos, por B. Kuttner, E. Mattioli, J. G. Mauldon e S. K. Zaremba (ver [Zar52]). Curiosamente, o chamado código de Golay $\mathcal{G}_{11}$ foi descoberto em 1947, na Finlândia e no contexto do problema do totobola, por Juhani Virtakallio, ano e meio antes de ter sido descoberto por Marcel Golay (ver [Bar93], p. 25).

Demonstração. Seja então $n = \frac{3^m-1}{2}$. O núcleo da aplicação sobrejectiva $\mathbb{F}_3^n \to \mathbb{F}_3^m$ dada pela matriz acima explicitada tem dimensão $n - m$ e portanto tem 3^{n-m} elementos, que formam um conjunto de elementos de $\mathbb{F}_3^n$ tais que a distância de Hamming entre quaisquer dois é pelo menos 3, pelo Lema anterior. Mas então, as bolas unitárias (para a distância de Hamming) centradas nesses pontos são disjuntas. Ora, cada uma dessas bolas assegura $2n + 1 = 3^m$ apostas. Resulta que todas as chaves ficam asseguradas por uma e uma só aposta do referido núcleo. □

Observe-se que, como $13 = \frac{3^3-1}{2}$, resulta que para o problema «clássico» do totobola com 13 partidas há uma solução perfeita (i.e. por bolas unitárias disjuntas), consistindo de 3^{10} apostas simples. O caso anterior em que tal acontece é $n = 4$ e o seguinte $n = 40$. Reciprocamente, é fácil ver que só existe uma cobertura disjunta para o totobola com n partidas quando n tem a forma prescrita na proposição anterior (o que é deixado como exercício para o leitor).

O núcleo da matriz acima explicitada constitui aquilo a que se chama um «código corrector de erros», nomeadamente um código ternário de Hamming (para mais detalhes, ver [Hil03]).

4. Problemas em aberto

A atestar a dificuldade do problema do totobola, em geral, está o facto de, embora o caso $n = 5$ do problema do totobola ter sido resolvido em 1967, por H. J. L. Kamps e Jacobus H. van Lint [Kam67], o caso $n = 6$ continuar em aberto. Observe-se que, muito embora 6 seja um número pequeno, a procura de uma solução pelo «método da força bruta» envolve procurar subconjuntos de um conjunto com 3^6 elementos, ou seja, envolve lidar com um conjunto com 2^{3^6} elementos!

A sequência cujo n-ésimo termo é o menor número de apostas simples que garantem $n-1$ resultados correctos num totobola de n partidas, é a sequência A004044 da «*The Online Encyclopedia of Integer Sequences*» mantida por Neil J. A. Sloane em `http://www.research.att.com/~njas/sequences`. Para além dos 5 primeiros termos, são conhecidos apenas os termos da forma $n = \frac{3^m-1}{2}$, graças aos códigos ternários de Hamming, acima descritos.

Há 25 anos atrás, sem acesso a revistas de investigação e numa era pré-internet, desconhecíamos as ligações entre o totobola e a teoria de códigos correctores de erros, assim como a literatura que havia sobre o assunto, e muito menos imaginávamos o quanto se continuaria a escrever (ver [Bar93, Ber04, Ham95] e respectivas referências) sobre o assunto!

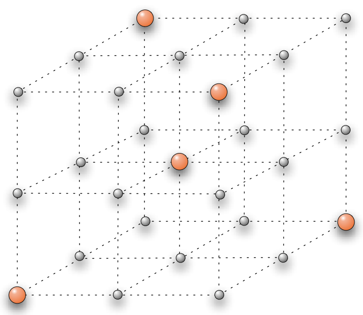

Figura 4: Uma solução óptima para $n = 3$

Referências

[Bar93] Alexander Barg, *At the Dawn of the Theory of Codes*, The Mathematical Intelligencer **15** (1993) 20–26.

[Ber04] Riccardo Bertolo, Patric Östergård, William D. Weakley, *An Updated Table of Binary/Ternary Mixed Covering Codes*, Journal of Combinatorial Designs **12** (2004) 157–176.

[Ham95] Heikki Hämäläinen, Iiro Honkala, Simon Litsyn, Patric Östergård, *Football Pools — A Game for Mathematicians*, American Mathematical Monthly **102** (1995) 579–588.

[Hay98] Teresa W. Haynes, Stephen Hedetniemi, Peter Slate, **Domination in Graphs**, CRC Press, 1998.

[Hil03] Raymond Hill, **A First Course in Coding Theory**, Oxford University Press, 2003.

[Kam67] H. J. L. Kamps, J. H. van Lint, *The Football Pool Problem for* 5 *Matches*, Journal of Combinatorial Theory **3** (1967) 315–325.

[Tau48] Olga Taussky, John Todd, *Covering Theorems for Groups*, Annales de la Société Polonaise de Mathématique **21** (1948) 303–305.

[Tho83] T. M. Thompson, **From Error-Correcting Codes Through Sphere Packings to Simple Groups**, Mathematical Association of America, 1983.

[Zar52] S. K. Zaremba, *Covering Problems Concerning Abelian Groups*, Journal of the London Mathematical Society **27** (1952) 242–246.