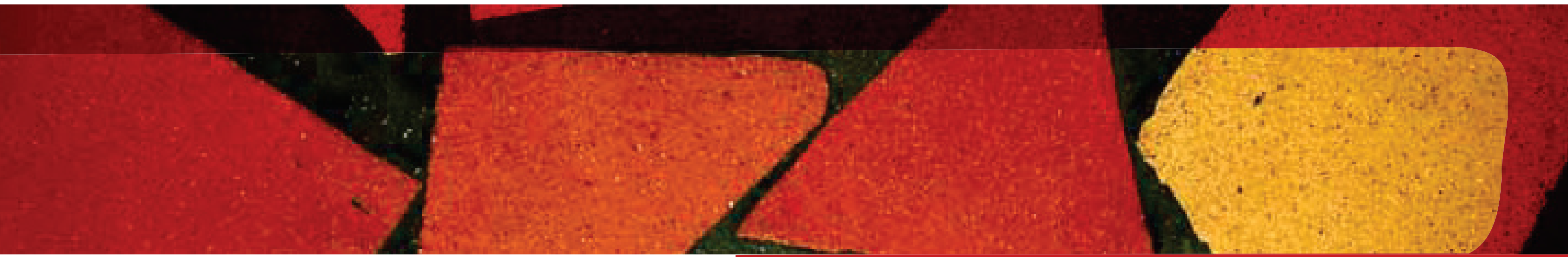

Manuel Silva
Universidade Nova de Lisboa
mnas@fct.unl.pt

Pedro J. Freitas
Universidade de Lisboa
pedro@ptmat.fc.ul.pt

UM PROBLEMA PARA POLIMATEMÁTICOS

Estamos habituados a pensar nos grandes avanços científicos como fruto da inspiração e do suor de alguns indivíduos geniais que conseguem ver mais além trabalhando isoladamente. Talvez no futuro esta situação possa mudar e o progresso seja obtido coletivamente como resultado da soma de inúmeras pequenas contribuições.

1. A ARTE DE BEM DISTRIBUIR

A teoria da discrepância teve origem numa questão formulada por Van der Corput em 1935: será possível construir uma sequência de números reais que esteja bem distribuída no intervalo [0,1]? Teremos em primeiro lugar de definir o que se entende por "estar bem distribuída". Para isso, podemos por exemplo verificar quantos dos primeiros n elementos da sequência estão contidos num dado intervalo $[a, b] \subset [0, 1]$. Se a sequência estiver bem distribuída este valor deverá ser aproximadamente proporcional ao comprimento do intervalo, mais precisamente $(b-a)n$. O conceito de número normal introduzido por Émile Borel em 1909 está de algum modo implícito nesta discussão. Recorde-se que um número real se diz normal na base 10 se na sua representação decimal aparecem todas as possíveis sequências finitas de dígitos com a frequência esperada. Por exemplo, a sequência 666 deverá aparecer com frequência $\frac{1}{1000}$. Para saber mais recomendamos a consulta de [1], onde se mostra entre outras coisas a relação da teoria da discrepância com o estudo da complexidade computacional.

A noção de discrepância que vamos discutir aqui, embora relacionada com a descrita no parágrafo anterior, tem natureza discreta. Denominamos por *hipergrafo* um par $\mathcal{H} = (V, \mathcal{A})$, em que V é um conjunto (eventualmente infinito), cujos elementos são denominados vértices, e onde $\mathcal{A} = \{A_i, i = 1, \cdots, m\}$ onde $A_i \subset V$ são subconjuntos finitos de vértices, designados por (hiper)arestas. Suponhamos que, dado um hipergrafo $\mathcal{H}$, pretendemos colorir os seus vértices com duas cores, azul e vermelho, de modo a que todas as arestas contenham aproximadamente o mesmo número de elementos de cada cor. Escolher uma coloração do conjunto de vértices V é equivalente a definir uma função $\chi : V \to \{-1, 1\}$.

A *discrepância* de um conjunto $B \subset V$ relativamente à coloração χ é dada pela diferença em valor absoluto entre o número de elementos azuis e vermelhos. A discrepância total do hipergrafo $\mathcal{H}$ é definida pela expressão

$$disc(\mathcal{H}) = \min_{\chi} \max_{i=1,\cdots,m} \Big|\sum_{x \in A_i} \chi(x)\Big|.$$

O mínimo é tomado sobre todas as 2^n colorações possíveis, onde $n = |V|$.

Nem sempre é possível escolher uma coloração dos vértices para a qual todas as arestas tenham discrepância baixa.

O resultado a seguir permite obter um majorante para $disc(\mathcal{H})$ no caso em que nenhum vértice aparece repetido em demasiadas arestas.

Teorema 1 (Beck-Fiala). *Se todo o vértice $v \in V$ aparece no máximo em t arestas, então $disc(\mathcal{H}) < 2t$.*

Este resultado tem uma demonstração simples que usa álgebra linear. Seria interessante melhorar o majorante linear n por uma função com um crescimento mais lento, como por exemplo $\sqrt{n}$.

2. O PROBLEMA DE DISCREPÂNCIA DE ERDÖS

Por volta de 1930, o matemático húngaro Paul Erdös formulou um problema envolvendo a discrepância nos conjuntos de múltiplos dos números naturais que permanece ainda hoje em aberto. O problema proposto por Erdös é neste momento objeto de um projeto de colaboração na Internet aberto a toda comunidade matemática, com o intuito de o resolver. Estamos habituados a pensar nos grandes avanços científicos como fruto da inspiração e do suor de alguns indivíduos geniais que conseguem ver mais além trabalhando isoladamente. Talvez no futuro esta situação possa mudar e o progresso seja obtido coletivamente como resultado da soma de inúmeras pequenas contribuições. O leitor interessado poderá encontrar na Internet uma página Wiki dedicada à questão proposta por Erdös.

O hipergrafo relevante no problema proposto por Erdös tem como conjunto de vértices o conjunto dos números naturais e como arestas os conjuntos de múltiplos: $P(d,k) = \{d, 2d, \cdots, kd\}, d, k \in \mathbb{N}$. Erdös conjeturou que usando duas cores não seria possível evitar que algum destes conjuntos de múltiplos tivesse discrepância arbitrariamente grande.

Conjetura (Erdös). *Para qualquer sequência $x_n \in \{-1, 1\}$, $n \in \mathbb{N}$ e $L > 0$, existem $d, k \in \mathbb{N}$ tais que*

$$|\sum_{i=1}^{k} x_{id}| > 1?$$

3. EXEMPLOS DE SEQUÊNCIAS

Como construir uma sequência x_n com discrepância baixa nos conjuntos de múltiplos? Uma ideia que certamente não funciona é escolher uma coloração periódica, isto porque facilmente constatamos existirem conjuntos de múltiplos arbitrariamente grandes todos da mesma cor. A procura de uma sequência equilibrada, i.e., com discrepância total baixa, terá de evitar regularidades como as repetições periódicas. A sequência pretendida não parece portanto ter uma estrutura simples.

Será boa ideia gerar a sequência procurada lançando a moeda ao ar: azul se sair cara e vermelho caso contrário? A ideia de usar um processo aleatório pode parecer à primeira vista estranha, mas muitas vezes este método permite construir objetos para os quais uma construção explícita se revelou inacessível.

No caso da sequência pretendida para o problema de discrepância de Erdös é possível mostrar que não é boa ideia lançar a moeda ao ar para decidir a cor de cada número natural. Para verificar a ineficiência do processo aleatório basta verificar que o máximo entre os valores médios das discrepâncias nos conjuntos de múltiplos pode ser muito grande.

Que outras construções explícitas de sequências binárias $x_n \in \{0, 1\}$, $n \in \mathbb{N}$ não periódicas faz sentido tentar? Um exemplo de uma sequência binária não periódica fácil de definir é a sequência de Thue-Morse $01101001 \cdots$. Esta sequência é determinada pela paridade do número de uns na expansão binária de cada natural n. Podemos definir um processo iterativo simples para obter esta sucessão: em cada passo justapomos ao bloco inicial x o bloco contrário $\bar{x}$ que se obtém substituindo em x cada zero por um e cada um por zero. Obtemos assim sucessivamente: 0, 01, 0110, 01101001. O leitor interessado pode encontrar mais informação sobre esta curiosa sequência redescoberta diversas vezes em vários contextos. Se eu escolher um bloco de comprimento arbitrariamente grande algures no meio desta sequência, não será possível adivinhar qual é o elemento que irá aparecer a seguir. A sequência de Thue-Morse tem no entanto discrepância $n^{\log_4(3)}$ nos conjuntos de múltiplos. Podemos experimentar outras sequências com este tipo de comportamento pseudo-aleatório. Por exemplo, a sequência binária dos restos da divisão por 3 dos números primos $p \geq 5$: $2, 1, 2, 1, 2, 1, 2, 2, 1, \cdots$. Também neste caso, adivinhar qual o próximo resto é algo que não pode ser obtido facilmente a partir dos restos anteriores.

Podemos, para simplificar o problema, considerar sequências $x_n \in \{-1, 1\}$ (*completamente*) *multiplicativas*, para as quais $x_{mn} = x_m x_n$ para qualquer $n, m \in \mathbb{N}$. Neste caso para avaliar a discrepância nos conjuntos de múltiplos basta determinar as somas parciais $\sum_{i=1}^{n} x_i$. Isto porque

$$|\sum_{i=1}^{n} x_{id}| = |x_d||\sum_{i=1}^{n} x_i| = |\sum_{i=1}^{n} x_i|.$$

Existe além disso alguma evidência empírica de que a haver uma sequência equilibrada esta deverá ser aproximadamente multiplicativa. Para definir uma sequência multiplicativa basta escolher a cor x_p onde p é um número primo. Um exemplo clássico de função completamente multiplicativa é a sequência de Liouville, a qual toma o valor –1 no conjunto dos números primos. Para cada natural n o valor da sequência de Liouville depende da paridade do número de fatores primos contados com a respetiva multiplicidade.

Encontrar um majorante para as somas parciais da função de Liouville é um problema clássico em teoria de números. É possível mostrar que o módulo da soma dos primeiros n primeiros termos da sequência de Liouville excede $C.\sqrt{n}$ para infinitos valores de n.

Podemos tentar usar um algoritmo guloso (*greedy*) para construir uma sequência multiplicativa: escolhendo para cada número primo entre os valores 1 e –1 aquele que garante menor discrepância nos conjuntos de múltiplos.

A melhor sequência conhecida para o problema de Erdös tem discrepância da ordem de $\ln n$. Começamos por escrever o número n na base 3 (onde usamos 3 dígitos 0, 1 e 2). Por exemplo, $13 = 2.3^2 + 1.3 + 2$, logo escreve-se 212 na base 3. Para definir a nossa sequência definimos $x_n = 1$ se o último dígito não nulo da representação ternária de n é 1 e $x_n = -1$ se for 2. Esta sequência é completamente multiplicativa e é possível mostrar que as somas parciais até n não excedem $\ln_3(n) + 1$, o que corresponde à sua discrepância.

Tratando-se de um problema em aberto que já resistiu 80 anos, não podemos saber quais as ferramentas para o resolver.

Se não conseguimos resolver um problema, podemos tentar resolver um outro mais simples ou pelo menos relacionado com o primeiro. Existem diversas variantes naturais para o problema proposto por Erdös. Podemos considerar três ou mais cores em vez de apenas duas. A definição de discrepância neste caso deverá ser o afastamento relativamente a $1/3$ da frequência relativa dos elementos de cada uma das cores nos conjuntos de múltiplos. Uma outra variante seria considerar funções f que tomam valores em $z \in \mathbb{C}$ com $|z| = 1$. Ao considerar variantes do problema procuramos desta forma obter um melhor entendimento do problema original. Até aqui a questão sugerida por Erdös resiste a todos os ataques. Convidamos o leitor a juntar-se na busca coletiva da solução deste problema.

REFERÊNCIAS

[1] Bernard Chazelle, *The Discrepancy method: Randomness and Complexity*