

ERRATA

AO NÚMERO 74 DO BOLETIM DA SPM
DE NOVEMBRO DE 2016

No número 74 do Boletim da SPM, publicado em Novembro de 2016, mais concretamente na listagem das teses de doutoramento em Ciências Matemáticas defendidas nas universidades portuguesas em 2015, incluída no artigo *Teses de doutoramento em Ciências Matemáticas defendidas nas universidades portuguesas em 2015*, há um lapso da inteira responsabilidade dos autores do artigo.

Na primeira tese da lista, defendida por Teresa de Jesus Costa Pereira Caracol Clain, foi omitido o nome da Co-Orientadora, a Professora Maryvonne Spiesser, da Université Paul-Sabatier, Toulouse, França.

Pelo lapso cometido, apresentamos o nosso pedido de desculpas à Professora Maryvonne Spiesser, à Doutora Teresa de Jesus Costa Pereira Caracol Clain e ao Professor Helmuth Robert Malonek, bem como aos leitores do Boletim.

Agradecemos ao Professor Helmuth Malonek, pela chamada de atenção.

A informação completa e corrigida da tese em causa é a seguinte.

1 Universidade de Aveiro

1.1 Programa Doutoral em Matemática

- Nome: Teresa de Jesus Costa Pereira Caracol Clain
Título: *A Matemática e o comércio em Portugal através das obras de aritméticos do século XVI: Gaspar Nicolas, Ruy Mendes e Bento Fernandes*
Data: 22 de setembro de 2015
Orientadores: Helmuth Robert Malonek (DM-UA), Maryvonne Spiesser (Université Paul-Sabatier, Toulouse, França)
URI/URL: <https://ria.ua.pt/handle/10773/15228>

O RISCO E A RUÍNA NA ATIVIDADE SEGURADORA

Irene Brito

Centro de Matemática da Universidade do Minho
Braga, Portugal
e-mail: ireneb@math.uminho.pt

Patrícia Gonçalves

Instituto Superior Técnico
Lisboa, Portugal
e-mail: patricia.goncalves@math.tecnico.ulisboa.pt

Pedro Lima Ramos

Escola de Psicologia da Universidade do Minho
Braga, Portugal
e-mail: pedro.ramos@psi.uminho.pt

Resumo: Neste artigo é apresentado um modelo teórico para a evolução temporal do capital de uma seguradora, nomeadamente o modelo clássico de risco de Crámer-Lundberg e o Teorema Fundamental do Risco que dá uma expressão explícita para a probabilidade de uma seguradora arruinar como função do capital inicial. É também apresentado um programa em linguagem R que permite obter estimativas para a probabilidade de ruína. Com este programa são feitas simulações considerando o modelo clássico de risco com diferentes distribuições para as indemnizações individuais e é feito um estudo comparativo das respetivas probabilidades de ruína.

Abstract: This paper presents a theoretical model for the temporal evolution of the capital of an insurance company, namely, the classical Crámer-Lundberg risk model and also the fundamental risk theorem, which provides an explicit expression for the ruin probability of the insurance company as a function of the initial capital. We also present a program in the R code which allows to get estimates for the ruin probability. With this program we perform some computational simulations considering the classical risk model with different distributions for the individual claims and we also present a comparison study of the respective ruin probabilities.

palavras-chave: Modelo clássico de risco de Crámer-Lundberg; probabilidade de ruína; teorema fundamental do risco.

keywords: Classical Crámer-Lundberg risk model; ruin probability; fundamental risk theorem.

1 Introdução

O atuário, como especialista na aplicação de cálculos matemáticos e estatísticos a operações financeiras na área dos seguros, e o conceito de seguro remontam ao tempo dos romanos. Entre os romanos, o atuário era um escriba que redigia as atas do senado. No decorrer dos séculos o atuário passou pelas funções de organizador de modelos estatísticos, como tábuas de vida, de gestor e começou por se preocupar com o cálculo de prémios e a venda de seguros de vida.

No século XV, o sistema de seguros europeu faliu, graças a técnicas de gestão de risco intuitivas e pouco elaboradas. Daí em diante, alguns matemáticos começaram a debruçar-se neste assunto, o que gradualmente conduziu à origem do atuariado que conhecemos hoje em dia. Pelo facto de se lidar com conceitos e técnicas de cariz diverso, desde conceitos matemáticos e estatísticos a económicos e financeiros, frequentemente se recorre à designação mais geral de “Ciências Atuariais” para o ramo de conhecimento sobre a análise de riscos e expectativas financeiras, principalmente na administração de seguros e fundos de pensão.

Os seguros dividem-se em dois ramos: vida e não-vida. O primeiro trata das questões de longo prazo, como reformas, pensões e seguros. O segundo está relacionado com características de curto prazo, como os seguros de automóveis e de responsabilidade civil. Note-se que no seguro de ramo vida a indemnização apenas ocorre uma vez, enquanto que no seguro de ramo não-vida, a indemnização pode ocorrer várias vezes. Por este e outros motivos, a abordagem e modelação teórica destes tipos de seguros exige abordagens e técnicas diferentes.

Este trabalho explora alguns conceitos, técnicas e resultados no âmbito das ciências atuariais não-vida, e apresenta o modelo clássico de risco de Crámer-Lundberg para o cálculo da probabilidade de ruína de uma seguradora como função do seu capital inicial, podendo este último ser visto como um montante de investimento ou o montante da seguradora num instante conhecido.

Este artigo está dividido da seguinte forma: na secção 2 apresenta-se o modelo clássico de risco e introduz-se a notação necessária para a compreensão das secções seguintes, na secção 3 apresentam-se alguns teoremas clássicos e fundamentais sobre a probabilidade ruína e na secção 4 apresentam-se

os resultados obtidos para a probabilidade de ruína através da realização de simulações computacionais. As estimativas para a probabilidade de ruína são obtidas considerando o modelo clássico de risco em que as indemnizações individuais seguem algumas distribuições conhecidas. É feito um estudo comparativo destes valores e é analisada a influência dos parâmetros das distribuições no valor da probabilidade de ruína. A secção 5 consiste num anexo onde são apresentadas as demonstrações dos resultados enunciados na secção 3.

2 Modelo clássico de risco

Ao longo do tempo, uma seguradora vai, por um lado, arrecadando os montantes referentes a prémios previamente estabelecidos, e, por outro, disponibilizando, montantes relativos às indemnizações associadas a sinistros envolvendo os seus clientes. Por uma questão de simplicidade, considera-se que o valor da taxa de prémio é invariável no tempo, ainda que, na realidade, nem sempre tal se suceda. Desta forma, as quantias que a seguradora vai auferindo ao longo do tempo, pelos prémios recebidos, são valores determinísticos, em nada dependentes de fenómenos aleatórios. Pelo contrário, a soma concedida em indemnizações desde um instante considerado inicial até um determinado momento é o resultado de acontecimentos que não podem ser previstos com exatidão absoluta, ou seja, tem de ser calculada recorrendo à utilização de variáveis aleatórias, um instrumento capaz de lidar eficazmente com o risco e a incerteza de eventos futuros.

O denominado *modelo clássico de risco* ou *modelo de Crámer-Lundberg* pretende, precisamente, modelar a situação descrita acima. Para tal considera-se um momento como sendo o inicial e, para cada instante t subsequente, representa-se por $U(t)$ a quantia que, então, se encontra na posse da seguradora, ou seja, o capital da seguradora no instante t , para um determinado ramo de seguros. A verba no instante inicial, por seu turno, é simbolizada por u e é uma quantidade determinística, sendo neste caso uma constante positiva; enquanto que o somatório do montante de todas as indemnizações relativas a sinistros ocorridos até ao momento t é representado por $S(t)$ e é uma quantidade aleatória. O montante arrecadado em prémios por unidade de tempo é designado por taxa de prémio e é denotado por c , uma constante estritamente positiva, e assume-se que o valor arrecadado pela seguradora em prémios até ao instante t é dado por ct . Note-se que, para cada $t \geq 0$, $U(t)$ e $S(t)$ são variáveis aleatórias. Este modelo não

tem em conta qualquer tipo de dividendos providos da gestão administrativa.

Em suma, o modelo clássico de risco em tempo contínuo para a atividade seguradora é um processo estocástico indexado no tempo $\{U(t) : t \geq 0\}$, onde, para cada $t \geq 0$,

$$U(t) = u + ct - S(t). \quad (1)$$

Note-se que $U(t)$ tem uma parte determinística e linear, nomeadamente $u + ct$ e uma parte aleatória $S(t)$. Agora, ir-se-á precisar o conceito das indemnizações agregadas até ao instante t , ou seja, $S(t)$. Para se poder tratar o modelo matematicamente, é assumido que as variáveis aleatórias representativas das indemnizações individuais são independentes e identicamente distribuídas a uma variável aleatória X . Ora, $S(t)$ é o somatório de X tantas vezes quanto o número de indemnizações ocorridas até ao instante t . O número de indemnizações é, também, um valor aleatório, representado pela variável aleatória $N(t)$, pelo que $S(t)$ é uma soma aleatória de variáveis aleatórias, ou seja, uma *variável aleatória composta*. É também assumido que as variáveis aleatórias $\{X_i\}_{i \geq 1}$ são independentes de $N(t)$ para todo $t > 0$. Conclui-se, então, que para $t \geq 0$,

$$S(t) = \sum_{i=1}^{N(t)} X_i, \quad (2)$$

onde para $i \geq 1$, X_i representa o montante da i -ésima indemnização. Finalmente, considera-se que a sequência de variáveis aleatórias, ou processo estocástico, $\{N(t) : t \geq 0\}$, goza das seguintes propriedades importantes (ver [3], [4] e [8] para mais detalhes):

1. Para $t = 0$, $N(t) = 0$, ou seja, no instante inicial não há indemnizações.
2. O processo $\{N(t) : t \geq 0\}$ tem incrementos estacionários e independentes.
3. Para cada instante t , $N(t)$ segue uma distribuição de Poisson de parâmetro λt , ou seja, $\forall t > 0$

$$\mathbb{P}(N(t) = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots \quad (3)$$

Há uma descrição equivalente do processo acima em [4], que consiste em assumir que $\{N(t) : t \geq 0\}$ satisfaz as propriedades 1., 2. descritas acima e ainda a seguinte propriedade:

(*) Para $h \rightarrow 0^+$, $\mathbb{P}(N(h) = 1) = \lambda h + o(h)$ e $\mathbb{P}(N(h) \geq 2) = o(h)$, onde $o(h)$ significa que $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$. Isto significa que em cada instante de tempo ocorre, no máximo, uma indemnização, e por essa razão, o tamanho dos saltos do processo $N(t)$ é 1 (ver a Figura 1).

Pelas propriedades descritas, $\{N(t) : t \geq 0\}$ é um *processo de Poisson homogéneo*. O processo diz-se homogéneo uma vez que o parâmetro λ não depende de t e diz-se de Poisson uma vez que para cada $t \geq 0$ a variável aleatória $N(t)$ tem distribuição de Poisson. Uma vez que $N(t)$ é um *processo de Poisson homogéneo*, neste caso, diz-se que $S(t)$ tem distribuição de Poisson composta ou que $\{S(t) : t \geq 0\}$ é um *processo de Poisson composto*. Obviamente que se poderia considerar outros tipos de processos estocásticos mais gerais do que o processo de Poisson homogéneo, como, por exemplo, o processo de Poisson não homogéneo, o processo de Poisson misto ou ainda considerar $\{N(t) : t \geq 0\}$ como um processo de Markov, veja-se [4].

De seguida, uma trajetória ou concretização de um processo de Poisson homogéneo $\{N(t) : t \geq 0\}$ é representada na Figura 1. Os processos $\{S(t) : t \geq 0\}$ e $\{U(t) : t \geq 0\}$ correspondentes encontram-se representados nas Figuras 2 e 3, respetivamente.

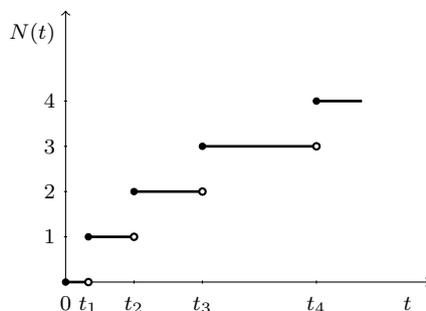


Figura 1: Processo de Poisson $N(t)$.

A título de exemplo, considere-se o instante t_2 . Até esse instante inclusivé a seguradora teve de suportar duas indemnizações, uma vez que $N(t_2) = 2$. Pelo que, por (2), $S(t_2) = \sum_{i=1}^2 X_i = X_1 + X_2$, como se pode constatar na Figura 2. Por último, por (1), $U(t_2) = u + ct_2 - S(t_2) = u + ct_2 - X_1 - X_2$.

Note-se que todos os processos $\{N(t) : t \geq 0\}$, $\{U(t) : t \geq 0\}$ e $\{S(t) : t \geq 0\}$ são processos de saltos, uma vez que as funções $N(t), U(t)$ e

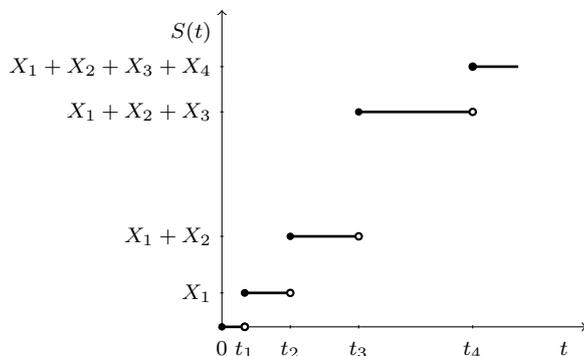


Figura 2: Indemnizações agregadas $S(t)$.

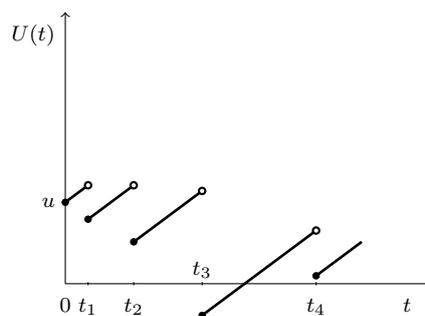


Figura 3: Modelo clássico de risco $U(t)$.

$S(t)$ são descontínuas, mais precisamente, contínuas à direita e com limite à esquerda. O tamanho do salto no processo $\{N(t) : t \geq 0\}$ é igual a 1, uma vez que em cada instante de tempo ocorre, no máximo, uma indemnização. Por outro lado, no instante de tempo em que ocorre uma indemnização a seguradora procede ao pagamento da respetiva indemnização e portanto cada salto de $S(t)$ tem tamanho igual ao valor desse montante.

No exemplo, t_3 é o chamado *instante de ruína*, ou seja, é o primeiro instante no qual o processo $\{U(t) : t \geq 0\}$ assume um valor negativo. De um modo geral, o tempo de ruína é representado por T e corresponde ao primeiro instante de tempo no qual a seguradora tem um capital negativo. Nesse instante, a seguradora está em dívida para com outrém. Todavia, não é forçoso que tal momento exista. A probabilidade de que exista, denomina-se *probabilidade de ruína*, depende naturalmente da quantia inicial u com

que a seguradora se inicia na sua exposição ao risco. A probabilidade de ruína em horizonte infinito representa-se por $\psi(u)$ e, desta forma, $\psi(u) = \mathbb{P}(T < \infty | U(0) = u)$, onde

$$T = \inf\{t > 0 : U(t) < 0\}.$$

No caso em que, $U(t) \geq 0$ para todo $t > 0$ convencionam-se que $T = +\infty$. Ao longo do trabalho, para facilitar a exposição, é usado $\psi(u) = \mathbb{P}(T < \infty)$, ficando desde já sub-entendido para o leitor a dependência de ψ em u .

Há uma restrição no valor do prémio c que é bastante intuitiva e fácil de demonstrar. Para fixar a notação chama-se μ ao valor médio das indemnizações individuais, ou seja, $\mu = \mathbb{E}[X]$. Lembre que as variáveis aleatórias $\{X_i\}_{i \geq 1}$ são independentes e identicamente distribuídas a uma variável aleatória X , que tem valor médio representado por μ ; obviamente $\mu > 0$. Lembre também que, como $N(t)$ tem distribuição de Poisson de parâmetro λt , então $\mathbb{E}[N(t)] = \lambda t$. É natural esperar que o valor médio do capital da seguradora no tempo t , nomeadamente $\mathbb{E}[U(t)]$, seja superior ao capital inicial u , ou seja, $\mathbb{E}[U(t)] > u$. Veja-se a que corresponde essa desigualdade e que informação advém dela. Primeiro terá que se calcular $\mathbb{E}[U(t)]$. Por (1) tem-se que $\mathbb{E}[U(t)] = u + ct - \mathbb{E}[S(t)]$. Lembre agora que $S(t)$ tem distribuição composta e para calcular o seu valor médio, é necessário recorrer à chamada *equação de Wald*, veja-se por exemplo [5], [9] ou [12]. A equação de Wald dita que o valor esperado de uma variável aleatória composta, como, por exemplo, a variável das indemnizações agregadas $S(t)$, satisfaz a seguinte igualdade $\mathbb{E}[S(t)] = \mathbb{E}[X]\mathbb{E}[N(t)]$. Nas condições do modelo definido acima, tem-se que $\mathbb{E}[S(t)] = \mu\lambda t$. Voltando a $\mathbb{E}[U(t)] > u$, tem-se que $u + ct - \mu\lambda t > u$, o que impõe a condição $c > \mu\lambda$. Caso esta condição não seja satisfeita, pode provar-se que, o processo estocástico $\{U(t) : t \geq 0\}$ incorre em ruína quase certamente, veja-se por exemplo [8] e [12]. Daqui em diante exige-se que $c > \mu\lambda$. Neste sentido, aplica-se (veja-se [3]) um *coeficiente de segurança* $\theta > 0$ de tal forma que

$$c = (1 + \theta)\lambda\mu. \tag{4}$$

Pretende-se obter o máximo de informação sobre a probabilidade de ruína $\psi(u)$ tendo em conta a modelação descrita acima. Tal será feito na secção seguinte, mas para tal ainda é necessário apresentar os seguintes conceitos. A *função geradora de momentos* de $S(t)$, definida por $M_{S(t)}(r) = \mathbb{E}[e^{rS(t)}]$, pode ser calculada a partir da fórmula (para mais detalhes veja-se [10])

$$M_{S(t)}(r) = M_{N(t)}(\ln(M_X(r))),$$

onde $M_X(r)$ denota a função geradora de momentos de X , e assume-se que satisfaz a seguinte condição

$$\lim_{r \rightarrow \infty} M_X(r) = +\infty. \quad (5)$$

Considerando que $S(t)$ tem distribuição de Poisson composta, onde $\{N(t) : t \geq 0\}$ é um processo de Poisson com distribuição dada por (3), então obtém-se

$$M_{S(t)}(r) = e^{\lambda t(M_X(r)-1)}. \quad (6)$$

É necessário, agora, introduzir o conceito de *coeficiente de ajustamento*, ajustamento, no caso, entre os prémios requeridos pela seguradora e as indemnizações que vai garantindo aos seus clientes. Este coeficiente define-se (veja-se [9]) como a única raiz positiva $r = R$ de

$$1 + (1 + \theta)\mu r = M_X(r), \quad (7)$$

que é equivalente à equação $M_{S(t)}(r) = e^{rct}$. Note-se que a unicidade do coeficiente de ajustamento assenta na hipótese (5).

3 Teorema Fundamental do Risco

Nesta secção enuncia-se o Teorema Fundamental do Risco (ver, por exemplo, [1], [4] ou [7]) e algumas consequências deste teorema. Apresenta-se também o cálculo do valor exato da probabilidade de ruína nos casos em que as indemnizações seguem uma distribuição exponencial e também no caso em que $u = 0$. As demonstrações dos resultados desta secção podem ser consultadas no anexo da secção 5.

Teorema 3.1 (*Teorema Fundamental do Risco*)

Para o processo (1), em que $S(t)$ é um processo de Poisson composto, com $c > \lambda\mu$, tem-se, para $u \geq 0$, que

$$\psi(u) = \frac{e^{-Ru}}{\mathbb{E}[e^{-RU(T)} \mid T < \infty]}, \quad (8)$$

onde R representa o coeficiente de ajustamento.

Este teorema pode ser de aplicação complexa, uma vez que o cálculo de $\mathbb{E}[e^{-RU(T)} \mid T < \infty]$ nem sempre é possível, pelo menos de forma exata. Daí que frequentemente se determine somente um majorante para a probabilidade de ruína, a saber, $\psi(u) \leq e^{-Ru}$, a denominada *Desigualdade de*

Lundberg. Esta desigualdade pode facilmente ser obtida a partir do Teorema Fundamental do Risco, tendo em conta que o denominador excede o valor 1, porque, dado que a ruína acontece em tempo finito, isto é, $T < \infty$, o capital da seguradora no momento da ruína, $U(T)$, é necessariamente negativo.

Há algumas consequências do Teorema Fundamental do Risco que se podem obter facilmente:

1. Se $\theta \rightarrow 0$, então $R \rightarrow 0$, logo $\psi(u) \rightarrow 1$.

Para provar esta observação note-se que para $g(r) = f(r) - M_X(r)$, onde $f(r) = 1 + (1 + \theta)\mu r$, tem-se que $g(0) = 0$, $g'(0) = \theta\mu$ e $g''(r) = -M_X''(r) < 0$. Quando $\theta \rightarrow 0$, tem-se que $g'(0) = \theta\mu \rightarrow 0$. Daqui resulta que próximo de $r = 0$ a função g é constante e nula, e como para além disso é côncava, ela não possui nenhum outro zero para além de $r = 0$. Logo o ponto de interseção de $M_X(r)$ e $f(r)$ é único, ou seja, $R \rightarrow 0$; e pela Desigualdade de Lundberg, uma vez que $u \geq 0$, tem-se que $\psi(u) \rightarrow 1$.

2. Se $\theta \leq 0$, então $\psi(u) = 1$ (a ruína é quase certa).

3. Se b é uma cota superior para as indemnizações particulares, então $\psi(u) > e^{-R(u+b)}$. Para provar esta observação, note-se que como $F_X(b) = 1$, tem-se que $0 \leq X_i \leq b$, e assim $-\sum_{i \geq 1}^{N(T)} X_i = -X_1 - \dots - X_{N(T)} \geq -b$, ou seja, $-S(T) \geq -b$. Como $u > 0$ e $c > 0$, então conclui-se que $U(T) \geq -b$. Assim, $e^{-RU(T)} \leq e^{Rb}$ e portanto, de (8) obtém-se $\psi(u) > \frac{e^{-Ru}}{e^{Rb}}$, como pretendido.

4. Fixado $u > 0$, tem-se que $\lim_{R \rightarrow \infty} \psi(u) = 0$, ou seja, quanto maior for o coeficiente de ajustamento, menor será a probabilidade de ruína.

5. Fixado $R > 0$, tem-se que $\lim_{u \rightarrow \infty} \psi(u) = 0$, ou seja, quanto maior for o capital inicial, menor será a probabilidade de ruína.

As observações 4. e 5. decorrem facilmente da Desigualdade de Lundberg.

O cálculo, de forma exata, da probabilidade de ruína pelo Teorema Fundamental do Risco é possível no caso em que as indemnizações individuais seguem uma distribuição exponencial e no caso em que $u = 0$.

De seguida vai ser apresentado o cálculo exato da probabilidade de ruína nestes dois casos, onde se vai aplicar o seguinte resultado (ver, por exemplo, [1], [4] ou [7]).

Teorema 3.2 *Para $u \geq 0$, tem-se que*

$$\psi'(u) = \frac{\lambda}{c}(\psi(u) - \int_0^u \psi(u-x)dF_X(x) - (1 - F_X(u))). \quad (9)$$

Integrando a expressão (9) (ver, por exemplo, [2]), obtém-se para $u \geq 0$

$$\psi(u) = \frac{\lambda}{c} \int_u^{+\infty} (1 - F_X(x))dx + \frac{\lambda}{c} \int_0^u \psi(u-x)(1 - F_X(x))dx. \quad (10)$$

As expressões para a probabilidade de ruína e para a sua derivada podem ser reescritas usando a *probabilidade de sobrevivência* $\delta(u) = 1 - \psi(u)$, da seguinte forma:

$$\delta(u) = \delta(0) + \frac{\lambda}{c} \int_0^u \delta(u-x)(1 - F_X(x))dx, \quad (11)$$

$$\delta'(u) = \frac{\lambda}{c} \left[\delta(u) - \int_0^u \delta(u-x)dF_X(x) \right]. \quad (12)$$

Agora apresenta-se o valor exato da probabilidade de ruína no caso em que $u = 0$. Tome $u = 0$ em (10) e recorde que $\mu = \mathbb{E}[X] = \int_0^{+\infty} (1 - F_X(x))dx$. Daqui resulta que $\psi(0) = \frac{\lambda\mu}{c}$. Usando (4) obtém-se

$$\psi(0) = \frac{1}{1 + \theta}. \quad (13)$$

Como corolário do resultado anterior, pode obter-se uma expressão para a função geradora de momentos de $\psi(u)$, que será útil para obter a expressão exata de $\psi(u)$ no caso em que for possível identificar a função geradora de momentos obtida. Abaixo irá-se exemplificar este método no caso das indemnizações terem distribuição exponencial.

Corolário 3.1 *Para $u \geq 0$*

$$\int_0^{\infty} e^{ru}(-\psi'(u)) du = \frac{\theta}{1 + \theta} \frac{M_X(r) - 1}{1 + (1 + \theta)\mu r - M_X(r)},$$

onde θ é o coeficiente de segurança e $M_X(\cdot)$ denota a função geradora de momentos de X .

A prova deste resultado, geralmente, faz uso de um processo designado perda agregada máxima, cuja definição, por si só, é bastante complicada. Para tornar este artigo o mais simples possível, decidiu-se apresentar a prova do resultado anterior de forma diferente da habitual usando apenas o cálculo de integrais [2], veja-se o anexo na secção 5.

Para o cálculo da probabilidade de ruína no caso de X ter distribuição exponencial será usado o Teorema Fundamental do Risco e o corolário anterior.

Considere-se o processo $U(t) = u + ct - S(t)$, em que as indemnizações individuais seguem uma distribuição exponencial de parâmetro $\beta > 0$. No intuito de determinar $\psi(u)$ pelo Teorema Fundamental do Risco, vai-se calcular em primeiro lugar o coeficiente de ajustamento R usando (7). Ora, como X tem distribuição exponencial de parâmetro β , tem-se que $\mathbb{E}[X] = \frac{1}{\beta}$ e

$$M_X(r) = \mathbb{E} \left[e^{rX} \right] = \int_0^\infty e^{rx} \beta e^{-\beta x} dx = \frac{\beta}{\beta - r}, \quad \beta > r,$$

logo,

$$1 + (1 + \theta)\mu r = M_X(r) \iff 1 + \frac{(1 + \theta)}{\beta} r = \frac{\beta}{\beta - r},$$

donde se obtém

$$R = \frac{\beta\theta}{1 + \theta}.$$

Note-se que R está bem definido porque $\beta, \theta > 0$.

Uma vez calculado o coeficiente de ajustamento, há necessidade de caracterizar a variável aleatória $-U(T)$. Para tal, veja-se que $-U(T)$ também tem distribuição exponencial de parâmetro β . Para tal, seja T o instante em que ocorre a ruína; \hat{u} o capital imediatamente antes de acontecer a ruína, ou seja, imediatamente antes de T e seja $y > 0$ um certo montante. Note-se que os acontecimentos $-U(T) > y$ e $X > \hat{u} + y | X > \hat{u}$ são equivalentes. Ora, se $-U(T) > y$, então $U(T) < -y$. Logo a indemnização que deu origem à ruína tem que ser superior ao valor imediatamente antes da ruína, nomeadamente \hat{u} , mais y , ou seja, tem que ter ocorrido ruína, portanto (tem-se a condição) $X > \hat{u}$, e condicionando a esse facto tem que se ter $X > \hat{u} + y$ (veja-se a Figura 4 para uma melhor compreensão).

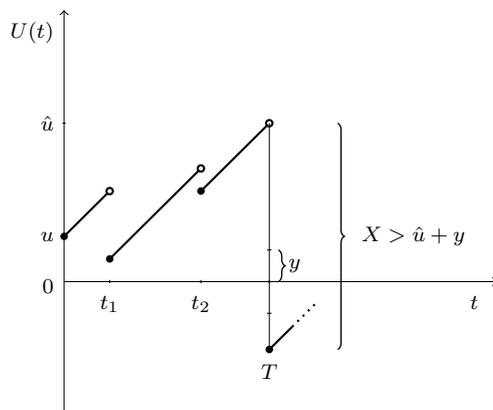


Figura 4: Processo $U(t)$ com $X > \hat{u} + y$.

Assim, para $y > 0$, tem-se

$$\begin{aligned} \mathbb{P}(-U(T) > y | T < +\infty) &= \mathbb{P}(X > \hat{u} + y | X > \hat{u}) = \frac{\mathbb{P}(X > \hat{u} + y, X > \hat{u})}{\mathbb{P}(X > \hat{u})} \\ &= \frac{\mathbb{P}(X > \hat{u} + y)}{\mathbb{P}(X > \hat{u})} = \frac{e^{-\beta(\hat{u}+y)}}{e^{-\beta\hat{u}}} = e^{-\beta y} = \mathbb{P}(X > y). \end{aligned}$$

Logo condicionando a $T < +\infty$, $-U(T)$ tem distribuição Exponencial de parâmetro β e portanto, a sua função geradora de momentos é dada por $\mathbb{E}[e^{-rU(T)} | T < +\infty] = \frac{\beta}{\beta-r}$, $r < \beta$.

O coeficiente de segurança θ , definido a partir da condição (4), é dado por $\theta = \frac{c}{\lambda\mu} - 1$ e recorde que $R = \frac{\beta\theta}{1+\theta}$. Então, tem-se

$$\psi(u) = \frac{e^{-Ru}}{\mathbb{E}[e^{-RU(T)} | T < +\infty]} = \frac{e^{(-\beta\theta u/(1+\theta))}}{\left(\frac{\beta}{\beta-R}\right)} = \frac{1}{1+\theta} e^{(-\beta\theta u/(1+\theta))}. \quad (14)$$

Agora, irá apresentar-se uma demonstração alternativa de (14) usando o corolário anterior.

Lembre que neste caso se tem $M_X(r) = \frac{\beta}{\beta-r}$, se $r < \beta$. Logo,

$$\int_0^\infty e^{ru} (-\psi'(u)) du = \frac{\theta}{1+\theta} \frac{\frac{\beta}{\beta-r} - 1}{1 + (1+\theta)\frac{r}{\beta} - \frac{\beta}{\beta-r}} = \frac{\gamma\delta}{\gamma-r},$$

onde $\gamma = \beta\theta/(1+\theta)$ e $\delta = 1/(1+\theta)$ e, então, como

$$\int_0^\infty e^{ru} \left(\frac{-\psi'(u)}{\delta}\right) du = \frac{\gamma}{\gamma-r},$$

resulta que a função geradora de momentos de $-\psi'(u)/\delta$ é igual a $\frac{\gamma}{\gamma-r}$. Ora, esta também é a função geradora de momentos de uma variável aleatória com distribuição exponencial de parâmetro γ . Logo, pela unicidade da função geradora de momentos, veja-se [5] ou [12], resulta que $\frac{-\psi'(u)}{\delta}$ tem que coincidir com a função densidade de probabilidade da exponencial de parâmetro γ , ou seja,

$$\frac{-\psi'(u)}{\delta} = \gamma e^{-\gamma u} \mathbf{1}_{\{u \geq 0\}}.$$

Integrando ambos os membros da igualdade anterior entre x e ∞ e usando o facto de que $\lim_{u \rightarrow \infty} \psi(u) = 0$, obtém-se

$$\frac{\psi(u)}{\delta} = -e^{-\gamma u} \mathbf{1}_{\{u \geq 0\}},$$

donde resulta que $\psi(u) = \delta e^{-\gamma u} = \frac{1}{1+\theta} e^{(-\beta\theta u/(1+\theta))}$, que coincide com o valor obtido anteriormente, tendo em conta que $u \geq 0$.

4 Simulações Computacionais

Nesta secção apresenta-se um programa desenvolvido em linguagem R que permite obter aproximações para a probabilidade de ruína $\psi(u)$, uma vez que nem sempre é possível determinar $\psi(u)$ de forma explícita. Com este programa obtêm-se estimativas para a probabilidade de ruína considerando o modelo clássico de risco com diferentes distribuições para as indemnizações individuais e analisam-se os resultados obtidos. São ainda considerados dois casos particulares de distribuições com o intuito de estudar a influência dos parâmetros das distribuições no valor da probabilidade de ruína. Num dos casos faz-se variar o parâmetro para uma dada distribuição das indemnizações individuais, nomeadamente a distribuição de Pareto, e no outro caso fazem-se variar os parâmetros da distribuição para o número de indemnizações individuais, nomeadamente a distribuição de Poisson.

O programa seguinte, escrito na linguagem de programação R, destina-se a calcular uma estimativa para a probabilidade de ruína, neste caso para situações em que as indemnizações individuais seguem ou uma distribuição Gama ou uma distribuição exponencial (a distribuição exponencial é um caso particular da distribuição Gama) ou uma distribuição de Pareto. Note-se que a distribuição Gama tem dois parâmetros: $\alpha > 0$, representando a

forma, “shape”; $\beta > 0$, representando a taxa, “rate”. A função densidade de probabilidade é dada por

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}.$$

O valor esperado é igual a $\frac{\alpha}{\beta}$ e a variância é igual a $\frac{\alpha}{\beta^2}$. A distribuição de Pareto, por seu turno, tem dois parâmetros: $\alpha > 0$, “scale”, escala; $\beta > 0$, “shape”, forma. A sua função densidade de probabilidade é dada por

$$f(x) = \frac{\beta \alpha^\beta}{x^{\beta+1}},$$

$x \geq \alpha$. O valor esperado é igual a $\frac{\alpha \beta}{\beta-1}$, se $\beta > 1$, a variância é igual a $\frac{\beta \alpha^2}{(\beta-1)^2(\beta-2)}$, $\beta > 2$. Contudo, o programa em questão é facilmente adaptável a outras distribuições.

Este programa recebe como dados de entrada o parâmetro `claims_type`, que identifica a distribuição das indenizações individuais, o coeficiente `lambda` do processo de Poisson $N(t)$, a quantia inicial `u` com que a seguradora se expõe ao risco, os coeficientes `alpha` e `beta` da distribuição das indenizações individuais, o coeficiente de segurança `theta`, o número de indenizações `nclaims` que o programa permite que ocorram dentro de cada simulação e, por fim, o número de simulações `nsim` a executar. Sublinhe-se que, para todo $t > 0$, $N(t)$ segue uma distribuição de Poisson de parâmetro λt , cuja função de probabilidade é dada por (3). O programa devolve como dados de saída, a probabilidade de ruína e, de entre o número de indenizações que tiveram de se suceder para se dar a ruína nas diversas simulações, devolve o maior, a média deles e o respetivo desvio padrão.

• O código em R

```
Simulation<-function(claims_type,lambda, u, alpha, beta, theta, nclaims, nsim){
#Se claims_type=1, X-Gama(alpha,beta); se claims_type=2, X-Pareto(alpha,beta);
# theta- coeficiente de segurança;
#Se claims_type=2, então beta tem de ser superior a 1.

library(stats4)
library(splines)
library(VGAM)

if (claims_type==1){
```

```

        c=(1+theta)*lambda*(alpha/beta)
    }else{
        c=(1+theta)*lambda*beta*alpha/(beta-1)
    }

    Ruin_pos=rep(Inf, nsim)

    #Cada iteração do loop principal representa uma simulação;
    #Cada simulação compreende exatamente nclaims indemnizações.

    for(i in c(1:nsim)){
        T_entre_claims=rexp(nclaims, lambda)
        T_espera=c()
        T_espera[1]=T_entre_claims[1]
        for(j in c(2:nclaims)){
            T_espera[j]=T_espera[j-1]+T_entre_claims[j]
        }
        if (claims_type==1){
            Claims=rgamma(nclaims,alpha,beta)
        }else{
            Claims=rpareto(nclaims,alpha,beta)
        }
        S=c()
        S[1]=Claims[1]
        for(j in c(2:nclaims)){
            S[j]=S[j-1]+Claims[j]
        }
        U=c()
        for(j in c(1:nclaims)){
            U[j]=u+c*T_espera[j]-S[j]
        }
        aux1=which(U<0)
        if (length(aux1)!=0){
            Ruin_pos[i]=min(aux1)
        }
    }

    aux2=which(Ruin_pos<Inf)
    Sol=Ruin_pos[aux2]
    cat(length(Sol), mean(Sol), sd(Sol), max(Sol), (length(Sol)/nsim))
}

```

Com este programa, e outros análogos, podem executar-se múltiplos cálculos. Considere-se a seguinte situação hipotética. A unidade temporal considerada é de um dia. Em média, ocorre um sinistro por cada cinco dias decorridos, ou seja, $\lambda = 1/5$. Suponha que as indemnizações particulares têm distribuição Gama de parâmetros $(1, \frac{1}{900})$, que coincide com a distribuição Exponencial de parâmetro $\frac{1}{900}$. Suponha que o capital inicial seja $u = 600$. Considere o coeficiente de segurança $\theta = 0.3$. Desta escolha resulta por (4) que $c = 234$. Note-se que a condição $c > \lambda\mu$ neste caso é satisfeita pois $234 > 900/5 = 180$ e portanto a ruína não ocorre quase

certamente. Pretende-se determinar uma aproximação para a probabilidade de ruína do seguinte modelo:

$$U(t) = 600 + (1 + 0.3) \frac{900}{5} t - \sum_{i=1}^{N(t)} X_i,$$

onde X_i tem distribuição Gama de parâmetros $(1, \frac{1}{900})$ ($\mathbb{E}[X_i] = 900$) para todo $i \geq 1$ e $N(t)$ tem distribuição de Poisson de parâmetro $t/5$. Fazem-se 10000 simulações, sendo que em cada uma delas existem 200 indenizações, obtendo-se:

```
R> Simulation(1,0.2,600,1,1/900,0.3,200,10000)
```

```
6540    6.690826  13.07788  167    0.654
```

Estes resultados podem ser interpretados da seguinte forma. Em 6540 das 10000 simulações ocorreu a ruína, pelo que $\psi(600) \approx 0.654$. Nestas simulações, em média, a ruína surgiu (pela primeira vez) na 7ª indenização e no máximo ocorreu na 167ª indenização. Não é, então, muito provável que a ruína surja após a indenização nº 200. Por seu turno, a solução positiva da equação $1 + 1170r = \frac{1}{1-900r}$, $r < \frac{1}{900}$, é o coeficiente de ajustamento. Para obter tal igualdade, está-se a aplicar (7) e o termo do lado direito corresponde à função geradora de momentos de uma distribuição Gama de parâmetros $(1, \frac{1}{900})$. Resolvendo esta igualdade, vem que $R \approx 2.5641 \times 10^{-4}$. Pelo que, pela Desigualdade de Lundberg, $\psi(600)$ será no máximo $e^{-600 \times 2.5641 \times 10^{-4}} \approx 0.8574$.

Considerem-se, agora, quatro tipos de modelos, com $u \geq 0$ e $j = 1, 2, 3, 4$:

$$U_j(t) = u + (1 + 0.3) \frac{900}{5} t - \sum_{i=1}^{N(t)} X_{j,i},$$

onde $N(t)$ é como acima e, para todo $i \geq 1$, $X_{1,i}$ tem distribuição Gama de parâmetros $(1, \frac{1}{900})$, ou Exponencial de parâmetro $\frac{1}{900}$, $X_{2,i}$ tem distribuição Gama de parâmetros $(90, \frac{1}{10})$, $X_{3,i}$ tem distribuição Gama de parâmetros $(900, 1)$ e $X_{4,i}$ tem distribuição de Pareto de parâmetros $(870.9827, 31.016)$. Desta forma, todas as distribuições das indenizações particulares têm valor esperado igual a 900. A variância da distribuição de Pareto e a variância da distribuição Gama de parâmetros $(900, 1)$ igualam 900. A variância da distribuição exponencial é 900^2 ; a da distribuição Gama de parâmetros

Tabela 1: Probabilidade de ruína associada a U_1 .

u	200	600	1250	5000
$\psi_{200;U_1}(u)$	0.7204	0.654	0.5572	0.2154
$\psi_{U_1}(u)$	0.7308	0.6595	0.558	0.2134
$\psi_{L;U_1}(u)$	0.95	0.8574	0.7258	0.2775

Tabela 2: Probabilidade de ruína associada a U_2 .

u	200	600	1250	5000
$\psi_{200;U_2}(u)$	0.7166	0.6216	0.4219	0.0538
$\psi_{L;U_2}(u)$	0.8954	0.7178	0.5013	0.0631

$(90, \frac{1}{10})$ é 9000.

Os valores da Tabela 1, por um lado, advêm da utilização do programa exibido atrás, dando origem aos valores representados por $\psi_{200;U_1}(u)$. Por outro lado, determinam-se por (14), uma vez que dizem respeito a modelos do tipo $U_1(t)$, isto é, modelos em que as indemnizações individuais seguem uma distribuição exponencial de parâmetro $1/900$. Trata-se da probabilidade de ruína exata, representada por $\psi_{U_1}(u)$. Por fim, $\psi_{L;U_1}(u)$ é o valor máximo da probabilidade de ruína pela Desigualdade de Lundberg.

Na Tabela 2, figuram os valores relativos a modelos do tipo $U_2(t)$; na Tabela 3, os valores de modelos do tipo $U_3(t)$; na Tabela 4, os valores para a probabilidade de ruína de quatro modelos do tipo $U_4(t)$ com diferentes valores de capital inicial. Note-se que, neste caso, não se calculam valores máximos pela Desigualdade de Lundberg. De facto, tais são inexistentes, uma vez que a função geradora de momentos de uma variável aleatória seguindo uma distribuição de Pareto somente se encontra definida para valores negativos, o que, neste contexto de indemnizações individuais, não faz sentido.

Como é óbvio a probabilidade de ruína é tanto menor quanto maior for o capital inicial. Repare-se que, quando o desfazamento entre o capital inicial e a média das indemnizações particulares não é grande, a probabilidade de ruína não se aproxima de valores extremos. Pelo contrário, quando $u = 200$, a probabilidade de ruína em todos os casos é de cerca

Tabela 3: Probabilidade de ruína associada a U_3 .

u	200	600	1250	5000
$\psi_{200;U_3}(u)$	0.7292	0.6191	0.4277	0.0534
$\psi_{L;U_3}(u)$	0.8943	0.7153	0.4975	0.0613

Tabela 4: Probabilidade de ruína associada a U_4 .

u	200	600	1250	5000
$\psi_{200;U_4}(u)$	0.7315	0.6237	0.4402	0.0523

de 70%. Inclusivamente, na simulação para a distribuição Exponencial, quando a ruína ocorre, em média, é logo na 4.983904^a indemnização, o que é compreensível uma vez que em média essa primeira indemnização será de 900, com o processo a iniciar-se em $u = 200$. Quando $u = 5000$ a ruína é menos provável, verificando-se apenas, no caso da distribuição Exponencial, em 2154 das 10000 simulações.

Claramente, os valores da probabilidade de ruína associada aos modelos do tipo $U_1(t)$ são significativamente mais elevados do que os relativos aos demais tipos de modelos, uma vez que, nos modelos do tipo $U_1(t)$, a variância das indemnizações é bastante superior à que se verifica nos outros tipos de modelos (variância em torno de um valor esperado comum). Portanto a probabilidade de ocorrência de indemnizações elevadas é significativamente mais elevada nos modelos do tipo $U_1(t)$, o que poderá aumentar a probabilidade do processo incorrer em ruína.

Na Figura 5 encontram-se representados de forma exata: o gráfico da probabilidade de ruína teórica para diferentes valores do capital inicial quando as indemnizações individuais seguem uma distribuição exponencial, $\psi_{U_1}(u)$; o gráfico com o valor máximo da probabilidade de ruína para a mesma situação, resultante da aplicação da Desigualdade de Lundberg, $\psi_{L;U_1}(u)$. Figuram igualmente quatro aproximações para quatro curvas. As curvas da probabilidade de ruína para diferentes valores do capital inicial quando as indemnizações particulares seguem: a distribuição Exponencial de parâmetros $(1, \frac{1}{900})$, $\psi_{200;U_1}(u)$; a distribuição Gama $(90, \frac{1}{10})$, $\psi_{200;U_2}(u)$; a distribuição Gama $(900, 1)$, $\psi_{200;U_3}(u)$; a distribuição de Pareto, $\psi_{200;U_4}(u)$.

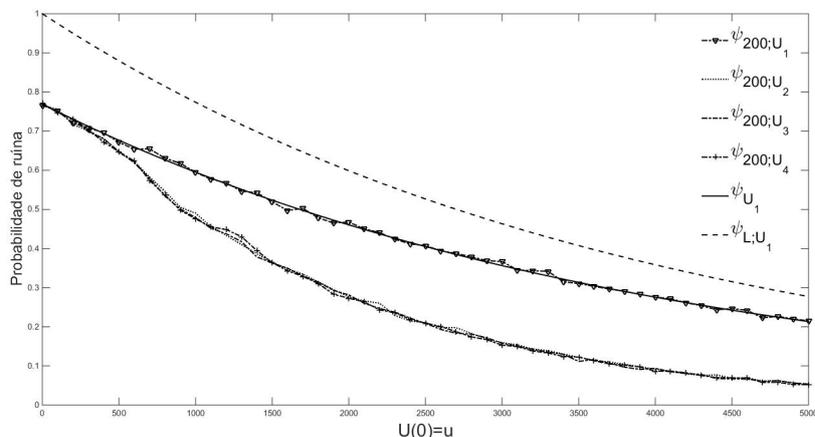


Figura 5: Probabilidade de ruína para $U_1(t)$, $U_2(t)$, $U_3(t)$ e $U_4(t)$.

De salientar ainda três aspetos. Em primeiro lugar, a distribuição de Pareto, no geral, parece exibir valores ligeiramente superiores para a probabilidade de ruína relativamente aos apresentados pela distribuição Gama(900, 1), o que é, provavelmente, consequência do facto de esta última ter a cauda menos pesada do que a primeira. Noutras palavras, a distribuição de Pareto mais facilmente assume valores considerados elevados em relação ao que acontece com a distribuição Gama(900, 1), não obstante possuírem o mesmo valor esperado e a mesma variância. Em segundo lugar, sublinhe-se a relativa proximidade entre os valores de $\psi_{U_1}(u)$ e $\psi_{L;U_1}(u)$, patente na Tabela 1 e na Figura 5. Essa proximidade é tanto maior quanto maior for o valor de u . Por fim, saliente-se a proximidade entre os valores teóricos $\psi_{U_1}(u)$ e os valores $\psi_{200;U_1}(u)$, provenientes da simulação computacional, o que confirma a eficácia deste método.

No seguinte exemplo de simulações são considerados modelos em que as indemnizações individuais X_i seguem distribuições de Pareto com diferentes parâmetros:

- 1) $X_i \sim \text{Pareto}(800, 9)$, com $\mathbb{E}[X_i] = 900$ e $\text{Var}[X_i] \approx 0.012867 \times 10^6$;
- 2) $X_i \sim \text{Pareto}(600, 2.8229)$, com $\mathbb{E}[X_i] \approx 900$ e $\text{Var}[X_i] \approx 0.36 \times 10^6$;

Tabela 5: Probabilidade de ruína

u	200	600	1250	5000
$\psi_{P_1}(u)$	0.7306	0.6192	0.4236	0.0541
$\psi_{P_2}(u)$	0.7293	0.6268	0.4603	0.1282
$\psi_{P_3}(u)$	0.7255	0.6129	0.4896	0.1823

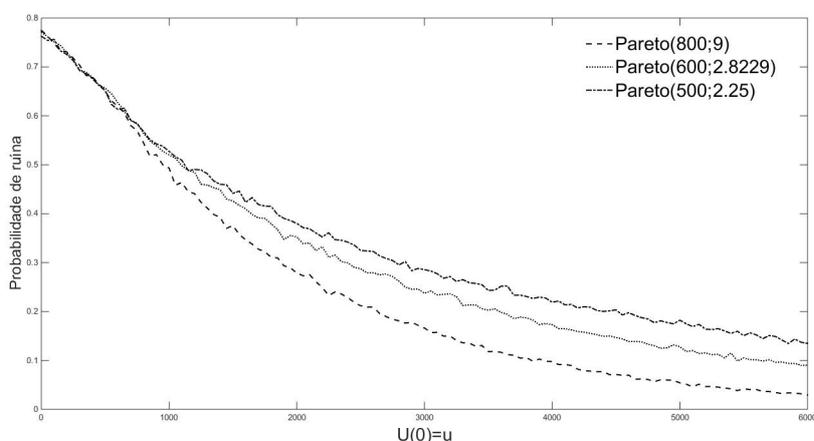


Figura 6: Probabilidade de ruína

3) $X_i \sim \text{Pareto}(500, 2.25)$, com $\mathbb{E}[X_i] = 900$ e $\text{Var}[X_i] = 1.44 \times 10^6$.

As probabilidades de ruína associadas aos modelos com indemnizações dadas por 1), 2) e 3) serão representadas por $\psi_{P_1}(u)$, $\psi_{P_2}(u)$ e $\psi_{P_3}(u)$, respetivamente. A Tabela 5 contém as probabilidades de ruína para os valores do capital inicial já considerados anteriormente e na Figura 6 estão representados os gráficos das probabilidades de ruína para os três modelos em função do capital inicial u .

Analisando a Tabela 5 e observando a Figura 6 conclui-se que a probabilidade de ruína aumenta com a variância das indemnizações e que esta diferença se torna mais notável com o aumento do capital inicial e quanto maior for a diferença entre as variâncias. De facto, os gráficos das probabilidades de ruína apresentam um decrescimento exponencial, que é mais acentuado em $\psi_{P_1}(u)$. Note-se que a variância das indemnizações do

Tabela 6: Probabilidade de ruína

u	200	600	1250	5000
$\lambda = 0.05$	0.7149	0.6083	0.4836	0.1852
$\lambda = 0.2$	0.7318	0.629	0.5084	0.2191
$\lambda = 1$	0.7272	0.639	0.5179	0.2277

modelo 2) é aproximadamente 28 vezes superior à do modelo 1) enquanto que a variância das indenizações do modelo 3) é aproximadamente 4 vezes superior à do modelo 2).

Por fim, o seguinte exemplo de simulações corresponde a um modelo em que as indenizações individuais seguem uma distribuição de Pareto(500, 2.25), considerando três processos de Poisson com parâmetros distintos:

- 1) $N(t) \sim \text{Poisson}(0.05t)$;
- 2) $N(t) \sim \text{Poisson}(0.2t)$;
- 3) $N(t) \sim \text{Poisson}(t)$.

O número total de unidades temporais é igual a 1000 para cada modelo. Nesse intervalo de tempo ocorrem em média: 50 indenizações no modelo 1), o que significa que num intervalo de 20 unidades temporais ocorre 1 indenização; 200 indenizações no modelo 2), o que significa que num intervalo de 5 unidades temporais ocorre 1 indenização; 1000 indenizações no modelo 3), o que significa que num intervalo de uma unidade temporal ocorre 1 indenização. As probabilidades de ruína associadas aos modelos com processo de Poisson definido em 1), 2) e 3) serão representadas por $\psi_{\lambda_1}(u)$, $\psi_{\lambda_2}(u)$ e $\psi_{\lambda_3}(u)$, respectivamente. A Tabela 6 contém as probabilidades de ruína para os valores do capital inicial considerados anteriormente e na Figura 7 estão representados os gráficos das probabilidades de ruína para os três modelos em função do capital inicial u .

Os valores na Tabela 6 e os gráficos na Figura 7 parecem indicar que a probabilidade de ruína aumenta com a frequência das indenizações e que a diferença entre as probabilidades aumenta com o capital inicial u .

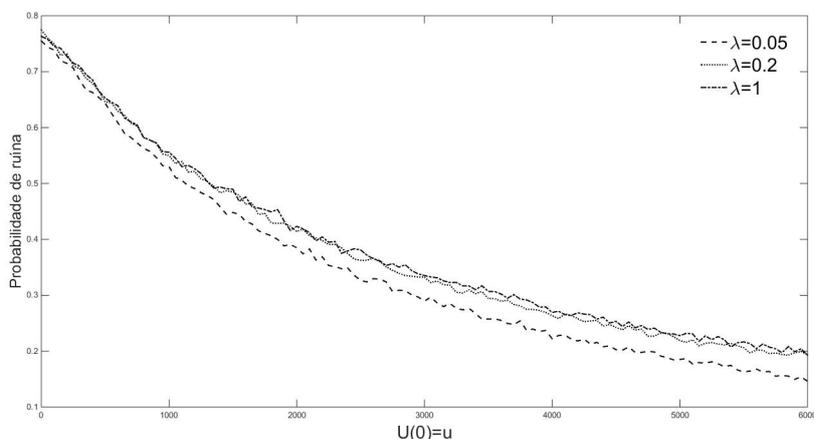


Figura 7: Probabilidade de ruína

Os valores de $\psi_{\lambda_1}(u)$ são sempre inferiores aos de $\psi_{\lambda_2}(u)$ e $\psi_{\lambda_3}(u)$, mas os gráficos destes dois últimos estão muito próximos e nem sempre se verifica que $\psi_{\lambda_2}(u) < \psi_{\lambda_3}(u)$, por exemplo para alguns valores de u no intervalo $u \in [0, 1000]$.

Agradecimentos

Irene Brito e Pedro Ramos agradecem ao Centro de Matemática da Universidade do Minho pelo apoio através da Fundação para a Ciência e a Tecnologia, no âmbito do projeto PEstOE/MAT/UI0013/2014. Patrícia Gonçalves agradece à FCT pelo suporte financeiro concedido através do projecto UID/MAT/04459/2013.

5 Anexo

Neste anexo são apresentadas as demonstrações dos resultados enunciados na secção 3, nomeadamente, o Teorema 3.1, o Teorema 3.2 e o Corolário 3.1.

5.1 Demonstração do Teorema Fundamental do Risco

A demonstração deste teorema pode ser sumariada da seguinte forma. Uma vez que se pretende obter uma relação entre $\psi(u)$ e a esperança condicional da variável aleatória $e^{-RU(T)}$, começa-se por fixar r e t arbitrários e por considerar a esperança da variável aleatória $e^{-rU(t)}$. De seguida, discretiza-

se o espaço tendo em conta o tempo de ruína, ou seja, tendo em conta se a ruína ocorreu antes ou depois do tempo t . Posteriormente, estima-se cada termo da esperança de $e^{-rU(t)}$ i.e. quando a ruína ocorre antes e depois do tempo t . Para completar a demonstração basta provar que a esperança no caso em que a ruína ocorre depois do tempo t é nula. Abaixo apresentam-se todos os passos detalhados deste argumento.

Demonstração. Fixe-se $t > 0$ e $r > 0$ e discretize-se o espaço tendo em conta a posição do tempo t em relação ao tempo de ruína T , ou seja, considerando os conjuntos $T \leq t$ e $T > t$. Tem-se que

$$\begin{aligned} \mathbb{E}[e^{-rU(t)}] &= \mathbb{E}[e^{-rU(t)}(\mathbf{1}_{\{T \leq t\}} + \mathbf{1}_{\{T > t\}})] \\ &= \mathbb{E}[e^{-rU(t)}\mathbf{1}_{\{T \leq t\}}] + \mathbb{E}[e^{-rU(t)}\mathbf{1}_{\{T > t\}}], \end{aligned}$$

e condicionando, obtém-se

$$\mathbb{E}[e^{-rU(t)}] = \mathbb{E}[e^{-rU(t)}|T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[e^{-rU(t)}|T > t]\mathbb{P}(T > t). \quad (15)$$

Agora começa-se por simplificar a expressão $\mathbb{E}[e^{-rU(t)}|T \leq t]$. Para $T \leq t$, tem-se que $U(t) = U(T) + c(t-T) - [S(t) - S(T)]$ e como $U(t) = u + ct - S(t)$, resulta que

$$\mathbb{E}[e^{-rU(t)}|T \leq t] = e^{-rc(t-T)}\mathbb{E}[e^{-rU(T) + r[S(t) - S(T)]}|T \leq t]. \quad (16)$$

Agora, usando o facto de que: se $T \leq t$ então $U(T)$ e $S(t) - S(T)$ são independentes, pois o que acontece num certo intervalo de tempo é independente do que acontece noutro intervalo de tempo disjunto do primeiro; e $S(t) - S(T)$ tem distribuição de Poisson composta com parâmetro $\lambda(t - T)$, a expressão (16) fica igual a

$$e^{-rc(t-T)}\mathbb{E}[e^{-rU(T)}|T \leq t]e^{\lambda(t-T)[M_X(r)-1]},$$

e para $r = R$ o coeficiente de ajustamento, a expressão (16) simplifica para

$$\mathbb{E}[e^{-RU(T)}|T \leq t].$$

Como consequência, a expressão (15) para $r = R$ fica escrita como

$$\mathbb{E}[e^{-RU(t)}] = \mathbb{E}[e^{-RU(T)}|T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[e^{-RU(t)}|T > t]\mathbb{P}(T > t). \quad (17)$$

Agora, é fácil ver que $\mathbb{E}[e^{-RU(t)}] = e^{-Ru}$. Para tal, note-se que por definição de $U(t)$, tem-se que $\mathbb{E}[e^{-RU(t)}] = e^{-Ru}e^{-Rct}\mathbb{E}[e^{RS(t)}] = e^{-Ru}e^{-Rct}M_{S(t)}(R)$, e por (4), (7) e (6) resulta que $\mathbb{E}[e^{-RU(t)}] = e^{-Ru}$.

Juntando as observações acima, a expressão (15) pode escrever-se como

$$e^{-Ru} = \mathbb{E}[e^{-RU(T)}|T \leq t]\mathbb{P}(T \leq t) + \mathbb{E}[e^{-RU(t)}|T > t]\mathbb{P}(T > t).$$

Notando que $\lim_{t \rightarrow \infty} \mathbb{P}(T \leq t) = \mathbb{P}(T < \infty)$ e tomando o limite quando $t \rightarrow \infty$ na igualdade anterior, tem-se que

$$e^{-Ru} = \mathbb{E}[e^{-RU(T)}|T < \infty]\mathbb{P}(T < \infty) + \lim_{t \rightarrow \infty} \mathbb{E}[e^{-RU(t)}|T > t]\mathbb{P}(T > t).$$

Para terminar a prova basta provar que o segundo termo do lado direito da expressão anterior é nulo, uma vez que a igualdade que resta é exatamente a igualdade que se pretende provar. Para tal, note-se que se $T > t$ então $U(t) \geq 0$, logo, discretizando o espaço tendo em conta a posição do montante $U(t)$ em relação a um montante $u_0(t)$, tem-se que

$$\begin{aligned} \mathbb{E}[e^{-RU(t)}|T > t]\mathbb{P}(T > t) &= \mathbb{E}[e^{-RU(t)}, T > t] \\ &= \mathbb{E}[e^{-RU(t)}, T > t, 0 \leq U(t) \leq u_0(t)] + \mathbb{E}[e^{-RU(t)}, T > t, U(t) > u_0(t)] \\ &= \mathbb{E}[e^{-RU(t)}|T > t, 0 \leq U(t) \leq u_0(t)]\mathbb{P}(T > t, 0 \leq U(t) \leq u_0(t)) \\ &\quad + \mathbb{E}[e^{-RU(t)}|T > t, U(t) > u_0(t)]\mathbb{P}(T > t, U(t) > u_0(t)) \\ &\leq \mathbb{P}(U(t) \leq u_0(t)) + \mathbb{E}[e^{-Ru_0(t)}]. \end{aligned}$$

Observe que, para já, não se impôs nenhuma condição no montante $u_0(t)$. Mas se $\lim_{t \rightarrow \infty} u_0(t) = \infty$, então $\lim_{t \rightarrow \infty} \mathbb{E}[e^{-Ru_0(t)}] = 0$. Sendo assim, só resta verificar que a probabilidade $\mathbb{P}(U(t) \leq u_0(t))$ se anula quando $t \rightarrow \infty$. Para tal, note-se que se obtém $\mathbb{E}[U(t)] = u + ct - \lambda\mu t$ e $Var[U(t)] = t\lambda\mathbb{E}[X^2]$. A última igualdade resulta do facto de que uma variável de Poisson tem média e variância iguais, e neste caso, igual a λt . Logo, escolhe-se, por exemplo, $u_0(t) = \mathbb{E}[U(t)] - t^{2/3}\lambda\mathbb{E}[X^2]$, que satisfaz a condição imposta acima. Finalmente, pela desigualdade de Chebychev, tem-se que

$$\begin{aligned} \mathbb{P}(U(t) \leq u_0(t)) &= \mathbb{P}(U(t) - \mathbb{E}[U(t)] \leq -t^{2/3}\lambda\mathbb{E}[X^2]) \\ &\leq \mathbb{P}(|U(t) - \mathbb{E}[U(t)]| \geq t^{2/3}\lambda\mathbb{E}[X^2]) \\ &\leq \frac{Var[U(t)]}{t^{4/3}(\mathbb{E}[X^2])^2\lambda^2} \leq \frac{1}{t^{1/3}\mathbb{E}[X^2]\lambda}, \end{aligned}$$

que tende a 0, quando $t \rightarrow \infty$.

□

5.2 Demonstração do Teorema 3.2

A demonstração deste resultado baseia-se no seguinte argumento. Uma vez que se pretende obter uma expressão exata para a derivada de $\psi(u)$, começa-se por escrever $\psi(u)$ à custa do número de indemnizações que ocorrem num intervalo de tempo infinitesimal $(0, dt)$. Usando as propriedades do processo de Poisson $N(t)$ e dos processos $S(t)$ e $U(t)$, obtêm-se expressões para a probabilidade de ruína no caso em que não ocorrem indemnizações nesse intervalo de tempo e quando ocorre apenas uma indemnização. A prova decorre da manipulação dessas expressões.

Demonstração. Fixado um intervalo de tempo infinitesimal $(0, dt)$, decompõe-se o evento $\{T < +\infty\}$ mediante o número de indemnizações que ocorrem nesse intervalo, ou seja, o número de indemnizações que levaram a seguradora à ruína

$$\begin{aligned} \psi(u) &= \\ &= \mathbb{P}(T < \infty) = \mathbb{P}(T < +\infty, N(dt) = 0, N(dt) = 1, N(dt) \geq 2) \\ &= \mathbb{P}(T < \infty, N(dt) = 0) + \mathbb{P}(T < \infty, N(dt) = 1) + \mathbb{P}(T < \infty, N(dt) \geq 2) \\ &= \mathbb{P}(T < \infty | N(dt) = 0) \mathbb{P}(N(dt) = 0) + \mathbb{P}(T < \infty | N(dt) = 1) \mathbb{P}(N(dt) = 1) \\ &\quad + \mathbb{P}(T < \infty | N(dt) \geq 2) \mathbb{P}(N(dt) \geq 2). \end{aligned}$$

Como $N(dt)$ tem distribuição de Poisson de parâmetro λt e pela expansão de Taylor da exponencial tem-se que

$$\mathbb{P}(N(dt) = 0) = 1 - \lambda dt, \quad \mathbb{P}(N(dt) = 1) = \lambda dt, \quad \mathbb{P}(N(dt) \geq 2) = o(dt),$$

ignorando os termos de ordem igual ou superior a $(dt)^2$. Note-se que as igualdades acima também podem ser obtidas usando a descrição alternativa do processo de Poisson acima, veja-se a propriedade (*) (cf. Teorema 2.1 de [4]). Por outro lado, tendo ainda em conta que

$$\begin{aligned} \mathbb{P}(T < \infty | N(dt) = 0) &= \mathbb{P}(T < \infty | S(dt) = 0) \\ &= \mathbb{P}(T < \infty | U(dt) = u + cdt) = \psi(u + cdt) \\ \mathbb{P}(T < \infty | N(dt) = 1) &= \mathbb{P}(T < \infty, X > u + cdt | N(dt) = 1) \\ &\quad + \mathbb{P}(T < \infty, X \leq u + cdt | N(dt) = 1) \\ &= (1 - F_X(u + cdt)) + \left(\int_0^{u+cdt} \psi(u + cdt - x) dF_X(x) \right) \\ \mathbb{P}(T < \infty | N(dt) \geq 2) &= o(dt), \end{aligned}$$

obtéem-se

$$\begin{aligned} \psi(u) &= (1 - \lambda dt) \psi(u + cdt) + \lambda dt \int_0^{u+cdt} \psi(u + cdt - x) dF_X(x) \\ &\quad + \lambda dt (1 - F_X(u + cdt)) + o(dt) \end{aligned}$$

ou seja,

$$\begin{aligned} \psi(u) &= \psi(u + cdt) \\ &\quad - \lambda dt \left[\psi(u + cdt) - \int_0^{u+cdt} \psi(u + cdt - x) dF_X(x) - [1 - F_X(u + cdt)] \right] \\ &\quad + o(dt). \end{aligned}$$

Daqui conclui-se que

$$\frac{\psi(u + cdt) - \psi(u)}{cdt}$$

se escreve como

$$\frac{\lambda}{c} \left[\psi(u + cdt) - \int_0^{u+cdt} \psi(u + cdt - x) dF_X(x) - [1 - F_X(u + cdt)] \right] + \frac{o(dt)}{cdt}.$$

Fazendo $dt \rightarrow 0$, segue que

$$\psi'(u) = \frac{\lambda}{c} \left[\psi(u) - \int_0^u \psi(u - x) dF_X(x) - [1 - F_X(u)] \right].$$

□

5.3 Demonstração do Corolário 3.1

Optou-se por apresentar a prova deste resultado usando apenas cálculos simples de integrais, nos quais se usa a fórmula de integração por partes, o Teorema de Fubini e propriedades da probabilidade de ruína listadas na secção 3 como consequências do Teorema 3.1.

Demonstração. Pelo Teorema 3.2 tem-se que

$$\begin{aligned} \int_0^\infty e^{ru} (-\psi'(u)) du &= \underbrace{\frac{-\lambda}{c} \int_0^\infty e^{ru} \psi(u) du}_{I(r)} + \underbrace{\frac{\lambda}{c} \int_0^\infty e^{ru} \int_0^u \psi(u - x) dF_X(x) du}_{II(r)} \\ &\quad + \underbrace{\frac{\lambda}{c} \int_0^\infty e^{ru} (1 - F_X(u)) du}_{III(r)}. \end{aligned}$$

Agora note-se que

$$\begin{aligned} III(r) &= \frac{\lambda}{c} \int_0^\infty e^{ru} \int_u^\infty dF_X(y) du = \frac{\lambda}{c} \int_0^\infty \left(\int_0^y e^{ru} du \right) dF_X(y) \\ &= \frac{\lambda}{cr} \int_0^\infty (e^{ry} - 1) dF_X(y) = \frac{\lambda}{cr} (M_X(r) - 1) = \frac{1}{(1 + \theta)\mu r} (M_X(r) - 1). \end{aligned}$$

Por outro lado, fazendo integração por partes e usando o facto de que $\lim_{u \rightarrow \infty} \psi(u) = 0$, tem-se que

$$\int_0^\infty e^{ru} (-\psi'(u)) du = \psi(0) + r \int_0^\infty e^{ru} \psi(u) du = \psi(0) + rI(r). \quad (18)$$

Sendo assim, tem-se que $II(r) + III(r) = \psi(0) + (r + \lambda/c)I(r)$. Agora vai-se calcular $I(r)$. Por (10) tem-se que

$$\begin{aligned} I(r) &= \underbrace{\frac{\lambda}{c} \int_0^\infty e^{ru} \int_u^\infty (1 - F_X(x)) dx du}_{A(r)} \\ &\quad + \underbrace{\frac{\lambda}{c} \int_0^\infty e^{ru} \left(\int_0^u \psi(u-x)(1 - F_X(x)) dx \right) du}_{B(r)}. \end{aligned}$$

Agora note-se que, aplicando o Teorema de Fubini duas vezes, obtém-se

$$\begin{aligned} A(r) &= \frac{\lambda}{c} \int_0^\infty \left(\int_0^x e^{ru} du \right) (1 - F_X(x)) dx = \frac{\lambda}{c} \int_0^\infty \frac{e^{rx} - 1}{r} (1 - F_X(x)) dx \\ &= \frac{\lambda}{c} \int_0^\infty \frac{e^{rx} - 1}{r} \left(\int_x^\infty dF_X(y) \right) dx = \frac{\lambda}{cr} \int_0^\infty \left(\int_0^y (e^{rx} - 1) dx \right) dF_X(y) \\ &= \frac{\lambda}{cr} \int_0^\infty \left(\frac{e^{ry} - 1}{r} - y \right) dF_X(y) = \frac{\lambda}{cr} \left(\frac{1}{r} (M_X(r) - 1) - \mu \right). \end{aligned}$$

Daqui conclui-se que

$$I(r) = \frac{\lambda}{cr} \left(\frac{1}{r} (M_X(r) - 1) - \mu \right) + B(r).$$

Finalmente calcula-se $B(r)$. Ora, aplicando o Teorema de Fubini e uma

mudança de variáveis, obtém-se

$$\begin{aligned}
 B(r) &= \frac{\lambda}{c} \int_0^\infty e^{ru} \int_0^u \psi(u-x)(1-F_X(x)) dx du \\
 &= \frac{\lambda}{c} \int_0^\infty (1-F_X(x)) \left(\int_x^\infty e^{ru} \psi(u-x) du \right) dx \\
 &= \frac{\lambda}{c} \int_0^\infty (1-F_X(x)) \left(\int_0^\infty e^{r(y+x)} \psi(y) dy \right) dx \\
 &= \frac{\lambda}{c} \int_0^\infty e^{rx} (1-F_X(x)) \left(\int_0^\infty e^{ry} \psi(y) dy \right) dx \\
 &= I(r) \frac{\lambda}{c} \int_0^\infty e^{rx} \left(\int_x^\infty dF_X(x) \right) dx \\
 &= I(r) \frac{\lambda}{c} \int_0^\infty \left(\int_0^y e^{rx} dx \right) dF_X(y) \\
 &= I(r) \frac{\lambda}{c} \int_0^\infty \frac{e^{ry} - 1}{r} dF_X(y) \\
 &= I(r) \frac{\lambda}{cr} (M_X(r) - 1).
 \end{aligned}$$

Sendo assim, conclui-se que

$$I(r) = \frac{\lambda}{cr} \left(\frac{1}{r} (M_X(r) - 1) - \mu \right) + I(r) \frac{\lambda}{cr} (M_X(r) - 1),$$

ou seja,

$$I(r) = \frac{\lambda}{r} \frac{M_X(r) - 1 - \mu r}{cr - \lambda(M_X(r) - 1)}.$$

Logo, substituindo $I(r)$ em (18) e usando (4) e (13), obtém-se o resultado. \square

Referências

- [1] N. Bowers, H. Gerber, J. Hickman, D. Jones e C. Nesbitt, *Actuarial Mathematics*, 2nd ed., Schaumburg, The Society of Actuaries, 1997.
- [2] I. Brito e P. Gonçalves, *Introdução à teoria do risco*, Universidade do Minho, 2015.
- [3] H. Bühlmann, *Mathematical Methods in Risk Theory*, 2nd ed., New York, Springer, 1996.

- [4] M. Centeno, *Teoria do Risco na Atividade Seguradora*, Oeiras, Celta Editora, 2003.
- [5] K. Chung, *A Course in Probability Theory*, 2nd ed., New York, Academic Press, 1974.
- [6] D. Dickson, *Insurance Risk and Ruin*, Cambridge, Cambridge University Press, 2005.
- [7] R. Kaas, M. Goovaerts, J. Dhaene e R. Denuit, *Modern Actuarial Risk Theory*, 2nd ed., Springer, 2009.
- [8] S. Karlin e H. Taylor, *A first course in stochastic processes*, Academic Press, 1975.
- [9] S. Klugman. H. Panjer e G. Willmot, *Loss Models*, 4th ed., Wiley, 2012.
- [10] P. Ramos, *Princípios de cálculo de prémios e medidas de risco em modelos atuariais*, Tese de Mestrado, Universidade do Minho, Portugal, 2014.
- [11] A. Reis, *Teoria da Ruína*, ISEG, Dezembro de 2001.
- [12] A. N. Shiriyayev, *Probability*, New York, Springer Verlag, 1984.

CÁLCULO FRACCIONÁRIO

Rui A. C. Ferreira

Grupo Física-Matemática, Universidade de Lisboa
Av. Prof. Gama Pinto, 2, 1649-003 Lisboa, Portugal
e-mail: raferreira@fc.ul.pt

Resumo: Nesta nota, pretendemos fazer uma breve exposição do denominado *Cálculo Fraccionário*. Introduzimos alguns conceitos, bem como explicitamos e comentamos algumas de suas propriedades. Algumas perspectivas sobre o futuro da área são também mencionadas. Finalmente, os principais livros escritos sobre o assunto, são referenciados para conveniência do leitor.

Abstract: In this note we intend to make a brief exposition of the so-called *Fractional Calculus*. We introduce some concepts as well, made explicit and make some comments on some of its properties. Some perspectives on the future of the area are also mentioned. Finally, the main books written on the subject, are referenced to the reader's convenience.

palavras-chave: Cálculo fraccionário; derivada de Riemann–Liouville, derivada de Caputo .

keywords: Fractional calculus; Riemann–Liouville derivative; Caputo's derivative.

1 A origem do Cálculo Fraccionário

A notação $d^n y/dx^n$, $n \in \mathbb{N}$, para a derivada de uma função é atribuída originalmente a Leibniz. Em 1695, L'Hôpital numa carta escrita a Leibniz, formulara a *simples* e ao mesmo tempo elegante questão: e se $n = \frac{1}{2}$?

Havia nascido o Cálculo Fraccionário!

Na sua carta de resposta a L'Hôpital, Leibniz escreve profeticamente, citando:

...This is an apparent paradox from which, one day, useful consequences will be drawn.

Subsequentemente, vários matemáticos contribuíram directa ou indirectamente para o estabelecimento (e desenvolvimento) desta teoria; destacam-se nomes como Euler, Lagrange ou Riemann e, mais recentemente, G. H. Hardy, J. E. Littlewood e H. Weyl (o leitor pode consultar [3] onde a história do cálculo fraccionário é descrita pormenorizadamente).

2 Integrais e derivadas de ordem arbitrária

O título desta secção não é inocente. De facto, a ordem dos operadores que iremos definir mais adiante poderá ser um número real qualquer e não apenas um número fraccionário. No entanto, a nomenclatura *Cálculo Fraccionário* ainda hoje é usada tendo em consideração a pergunta que conduziu ao início do estudo das derivadas de ordem não-inteira (cf. Secção 1).

Comecemos então por definir um integral de ordem real: sejam $a \in \mathbb{R}$ e f uma função real de variável real e integrável para $t \geq a$. Se $n \in \mathbb{N}$ e

$$I_a^n f(t) = \int_a^t dt_1 \int_a^{t_1} dt_2 \dots \int_a^{t_{n-1}} f(s) ds,$$

então é fácil demonstrar por indução que

$$I_a^n f(t) = \frac{1}{(n-1)!} \int_a^t (t-s)^{n-1} f(s) ds. \quad (1)$$

Ora, como $(n-1)! = \Gamma(n)$, onde

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0,$$

é a função *Gamma*, então o lado direito de (1) pode ter sentido para valores não inteiros de n .

Define-se assim o *Integral de Riemann–Liouville* de ordem $\alpha > 0$ de uma função f por

$$I_a^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (t-s)^{\alpha-1} f(s) ds, \quad (2)$$

desde que o lado direito de (2) exista.

Impondo condições bastante razoáveis à função f (e.g. continuidade), é possível demonstrar que $\lim_{\alpha \rightarrow 0} I_a^\alpha f(t) = f(t)$ (cf. [4, Secção 2.3.2]). Por essa razão, define-se $I_a^0 f(t) = f(t)$.

Calculamos a título de exemplo o integral de Riemann–Liouville (de ordem $\alpha > 0$) da função potência: $f(t) = (t-a)^\beta$, $\beta > -1$. Temos que:

$$\begin{aligned} I_a^\alpha f(t) &= \frac{1}{\Gamma(\alpha)} \int_a^t (t-s)^{\alpha-1} (s-a)^\beta ds \\ &= \frac{1}{\Gamma(\alpha)} \int_0^1 (t-a-r(t-a))^{\alpha-1} r^\beta (t-a)^\beta (t-a) dr \\ &= \frac{1}{\Gamma(\alpha)} (t-a)^{\alpha+\beta} \int_0^1 r^\beta (1-r)^{\alpha-1} dr \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\Gamma(\alpha)}(t-a)^{\alpha+\beta} B(\beta+1, \alpha) \\
&= \frac{\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}(t-a)^{\alpha+\beta},
\end{aligned}$$

onde $B(x, y)$ é a denominada função beta (ver [3, pag. 299–300]).

Tendo definido um integral de ordem arbitrária, definimos agora as duas derivadas *fraccionárias* mais utilizadas na literatura (para outras definições o leitor poderá consultar e.g. [5]). A *Derivada de Riemann–Liouville* de ordem $\alpha > 0$ de uma função f é definida por

$${}_a D^\alpha f(t) = D^n I_a^{n-\alpha} f(t), \quad \text{onde } n = \lceil \alpha \rceil, \quad (3)$$

e a *Derivada de Caputo* de ordem $\alpha > 0$ de uma função f é definida por

$${}_a^C D^\alpha f(t) = I_a^{n-\alpha} D^n f(t), \quad \text{onde } n = \lceil \alpha \rceil, \quad (4)$$

sempre que o lado direito de (3) e de (4) estejam bem definidos, respectivamente.

Assumindo que (3) e (4) existem, então as derivadas de Riemann–Liouville e de Caputo estão relacionadas pela seguinte fórmula:

$${}_a^C D^\alpha f(t) = {}_a D^\alpha f(t) - \sum_{k=0}^{n-1} \frac{D^k f(a)}{\Gamma(k-\alpha+1)}(t-a)^{k-\alpha}, \quad n = \lceil \alpha \rceil. \quad (5)$$

Podemos concluir imediatamente que estas derivadas coincidem quando $D^k f(a) = 0$, $k = 0, \dots, n-1$. De facto, uma das principais razões¹ para um investigador optar pela escolha do operador (3) ou do operador (4) é a definição das condições iniciais para uma determinada equação diferencial (cf. e.g. Secção 5 e Secção 6 em [1]).

Vejamos agora algumas propriedades destes operadores (para a derivada usaremos a de Caputo). Assumindo que para uma determinada função f os operadores estão bem definidos, temos que:

1. $I_a^\alpha I_a^\beta f(t) = I_a^{\alpha+\beta} f(t) = I_a^\beta I_a^\alpha f(t)$.
2. $I_a^{\alpha C} D^\alpha f(t) = f(t) - \sum_{k=0}^{n-1} \frac{D^k f(a)}{k!} (t-a)^k$.

¹A propósito, note-se que a derivada de Riemann–Liouville da função constante não é zero enquanto que a derivada de Caputo da mesma é.

$$3. {}^C D_a^\alpha I_a^\alpha f(t) = f(t).$$

A fórmula em 1. é designada por *propriedade de semigrupo* (demonstraremos esta propriedade no final desta secção); as igualdades em 2. e 3. permitem, por exemplo, transformar um problema de valor inicial para a sua (equivalente) forma integral (cf. e.g. [1, Secção 6.1]). Mais propriedades podem ser consultadas em qualquer dos livros referenciados [1, 2, 3, 4, 5], pois não pretendemos de modo nenhum ser exaustivos nesta apresentação. Concluimos esta secção alertando o leitor de que, mesmo para classes de funções f e g bastante gerais, a fórmula de Leibniz (derivada do produto de funções) não é válida em geral para estes operadores, i.e. em geral, ${}_a D^\alpha f g(t) \neq {}_a D^\alpha f(t) g(t) + f(t) {}_a D^\alpha g(t)$, onde D^α representa um operador de ordem α (o leitor pode consultar a fórmula de Leibniz em [1, Theorem 2.18]). Isto apresenta um claro inconveniente aquando da tentativa de provar alguns resultados da teoria de equações diferenciais usando ao invés derivadas fraccionárias (um claro exemplo deste facto é a falta de resultados (fortes) em sistemas dinâmicos usando derivadas fraccionárias).

Demonstração de propriedade 1.:

$$\begin{aligned} I_a^\alpha I_a^\beta f(t) &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_a^t (t-s)^{\alpha-1} \int_a^s (s-r)^{\beta-1} f(r) dr ds \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_a^t f(r) \int_r^t (t-s)^{\alpha-1} (s-r)^{\beta-1} ds dr \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_a^t f(r) (t-r)^{\alpha+\beta-1} dr \int_0^1 (1-s)^{\alpha-1} s^{\beta-1} ds \\ &= \frac{1}{\Gamma(\alpha+\beta)} \int_a^t (t-r)^{\alpha+\beta-1} f(r) dr, \end{aligned}$$

onde na segunda igualdade usamos o Teorema de Fubini.

3 A actualidade e o futuro do cálculo fraccionário

Presentemente, a teoria e as aplicações do cálculo fraccionário constituem uma área muito activa de investigação. Cientistas ditos “mais puros” bem como outros ditos “mais aplicados” fazem investigação em EDOs ou PDEs e suas aplicações a problemas que emergem do mundo real. Basta para isso fazer uma breve consulta de algumas revistas onde os artigos são publicados (note-se também a qualidade das próprias revistas!): *Advances in Mathematics*, *Archive for Rational Mechanics and Analysis*, *Discrete and Continuous Dynamical Systems–Series A*, *Nonlinear Analysis–Theory, Methods*

& Applications, Proceedings of the American Mathematical Society, SIAM Journal on Mathematical Analysis, SIAM Journal on Numerical Analysis e Transactions of the American Mathematical Society. Mais ainda, gostaríamos de referir que a revista de maior relevância e expressividade na área, *Fractional Calculus and Applied Analysis*, tem um *impact factor* de 2.245, estando classificada em 11^o lugar (no ranking dos *impact factor*) na área *Mathematics*, de acordo com a ISI Web of Science.

Tendo em conta o que foi escrito anteriormente, assim como, levando em consideração a opinião de alguns especialistas da área na actualidade, expressa no artigo [6], pensamos poder afirmar que o cálculo fraccionário continuará a ser uma área de grande potencial. Em particular, parece-nos que será bastante importante para a ciência em geral que matemáticos fazendo investigação em equações com derivadas fraccionárias colaborem mais de perto com cientistas provenientes de outras áreas, como por exemplo, biologia, física e engenharias.

Referências

- [1] K. Diethelm, *The analysis of fractional differential equations*, Lecture Notes in Mathematics, 2004, Springer, Berlin, 2010.
- [2] A. A. Kilbas, H. M. Srivastava e J. J. Trujillo, *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies, 204, Elsevier, Amsterdam, 2006.
- [3] K. S. Miller e B. Ross, *An introduction to the fractional calculus and fractional differential equations*, Wiley, New York, 1993.
- [4] I. Podlubny, *Fractional differential equations*, Mathematics in Science and Engineering, 198, Academic Press, San Diego, CA, 1999.
- [5] S. G. Samko, A. A. Kilbas e O. I. Marichev, *Fractional integrals and derivatives*, translated from the 1987 Russian original, Gordon and Breach, Yverdon, 1993.
- [6] J. T. Machado, F. Mainardi e V. Kiryakova, “Fractional calculus: quo vadimus? (Where are we going?)”, *Fract. Calc. Appl. Anal.*, Vol. 18, No. 2 (2015), pp. 495–526.

CONSTRUÇÃO DE MATRIZES INTEIRAS COM SOMAS DAS LINHAS E DAS COLUNAS PRESCRITAS

Rosário Fernandes

Departamento de Matemática, Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa, Caparica, Portugal
e-mail: mrff@fct.unl.pt

Henrique F. da Cruz

Departamento de Matemática da Universidade da Beira Interior
Covilhã, Portugal
e-mail: hacruz@ubi.pt

Resumo: Descrevemos um algoritmo de construção de matrizes, que só têm números inteiros não negativos e não superiores a determinado inteiro positivo p , cujas somas dos elementos de cada linha e de cada coluna da matriz são prescritas.

Abstract: We describe an algorithm for the construction of matrices, with nonnegative integers and no greater than a positive integer p , that have the sum of the integers in each row and in each column of the matrix known.

palavras-chave: Matrizes com entradas inteiras; Partições; Algoritmo.

keywords: Integral matrices; Partitions; Algorithm.

1 Introdução

No Boletim da Sociedade Portuguesa da Matemática [3] encontra-se um artigo onde é apresentado um algoritmo para a construção de matrizes de zeros e uns com somas de linhas e de colunas prescritas. Como foi mencionado, as matrizes de zeros e uns são muito utilizadas tanto em áreas da Matemática como também em áreas da Física, Química e Biologia. Em certos problemas de Investigação Operacional, surgem não só as matrizes de zeros e uns, mas também matrizes cujas posições são números inteiros não negativos, não superiores a um determinado número positivo p fixo, que designaremos por p -matrizes. Tal como as matrizes de zeros e uns, as p -matrizes parecem ser matrizes relativamente simples de estudar. Porém esta ideia inicial não é de todo verdadeira, e inúmeros problemas podem ser colocados em relação a estas matrizes. Por conseguinte, as p -matrizes têm sido o alvo de um

intenso estudo feito por imensos matemáticos, [1, 2, 4, 6]. Com este texto pretendemos estender o texto de [3] para p -matrizes, e descrever um algoritmo de construção destas matrizes, verificando certas condições ligadas às somas dos elementos das linhas e das colunas das matrizes. O algoritmo por nós apresentado estende para as p -matrizes alguns algoritmos já conhecidos para a construção de matrizes de zeros e uns, como o algoritmo de Ryser, seguindo ideias similares.

2 Partições e p -majoração

Uma *partição* de peso $t \geq 0$ é uma sequência não crescente de números inteiros não negativos, chamados *partes* da partição, cuja soma é igual a t . O número de elementos não nulos na partição λ é o *comprimento* de λ e é denotado por $l(\lambda)$. Quando λ é uma partição de peso t , representamos habitualmente λ por uma sequência finita $(\lambda_1, \lambda_2, \dots, \lambda_{l(\lambda)})$, onde $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{l(\lambda)} > 0$. Por vezes pode ser necessário que a partição tenha um número de partes superior ao seu comprimento. Nesse caso acrescentamos, depois da última parte não nula, uma sequência de zeros.

Definição 2.1 *Seja $(\lambda_1, \dots, \lambda_{l(\lambda)})$ uma partição de peso t . A partição conjugada de $(\lambda_1, \dots, \lambda_{l(\lambda)})$ é a partição λ^* , de peso t , cuja coordenada j é*

$$\lambda_j^* = |\{i : l(\lambda) \geq i \geq 1, \lambda_i \geq j\}|, \quad \text{com } 1 \leq j \leq \lambda_1.$$

Exemplo 2.2 *Consideremos a partição λ de peso 15, tal que $\lambda = (4, 3, 3, 2, 2, 1)$. A partição conjugada desta partição tem as coordenadas*

$$\begin{aligned} \lambda_1^* &= |\{i : \lambda_i \geq 1\}| = |\{1, 2, 3, 4, 5, 6\}| = 6, \\ \lambda_2^* &= |\{i : \lambda_i \geq 2\}| = |\{1, 2, 3, 4, 5\}| = 5, \\ \lambda_3^* &= |\{i : \lambda_i \geq 3\}| = |\{1, 2, 3\}| = 3, \\ \lambda_4^* &= |\{i : \lambda_i \geq 4\}| = |\{1\}| = 1. \end{aligned}$$

Donde, $\lambda^* = (6, 5, 3, 1)$. □

Definição 2.3 *Sejam α e β duas partições com o mesmo peso. Dizemos que α é dominado ou majorado por β , $\alpha \preceq \beta$, quando*

$$\alpha_1 + \dots + \alpha_i \leq \beta_1 + \dots + \beta_i, \quad \text{para } i \geq 1.$$

Existe uma relação muito útil entre a majoração de duas partições e a majoração das respetivas partições conjugadas:

Proposição 2.4 *Sejam α e β duas partições com o mesmo peso. Então,*

$$\alpha \preceq \beta, \text{ se, e só se, } \beta^* \preceq \alpha^*.$$

Em [4], generalizámos a relação de majoração entre duas partições com o mesmo peso:

Definição 2.5 *Sejam α e β duas partições de peso t e seja p um inteiro positivo. Dizemos que α é p -dominado (ou p -majorado) por β , e escrevemos $\alpha \preceq_p \beta$, se*

$$\sum_{i=1}^k \alpha_i \leq \sum_{i=1}^{pk} \beta_i, \text{ para } k \geq 1.$$

Note-se que $\alpha \preceq_1 \beta$ é o mesmo que $\alpha \preceq \beta$.

Exemplo 2.6 *Consideremos as partições $\lambda = (5, 2)$ e $\mu = (3, 3, 1)$ de peso 7. Uma vez que*

$$\begin{aligned} \mu_1 = 3 &\leq 5 = \lambda_1 \\ \mu_1 + \mu_2 = 6 &\leq 7 = \lambda_1 + \lambda_2 \\ \mu_1 + \mu_2 + \mu_3 = 7 &\leq 7 = \lambda_1 + \lambda_2 + \lambda_3, \end{aligned}$$

concluimos que $\mu \preceq \lambda$. Porém, $\lambda \not\preceq \mu$ porque

$$\lambda_1 = 5 \not\leq 3 = \mu_1.$$

No entanto,

$$\begin{aligned} \lambda_1 = 5 &\leq 6 = \mu_1 + \mu_2 \\ \lambda_1 + \lambda_2 = 7 &\leq 7 = \mu_1 + \mu_2 + \mu_3 + \mu_4, \end{aligned}$$

e então $\lambda \preceq_2 \mu$. □

Como era de esperar, a Proposição 2.4 pode ser generalizada para esta relação:

Proposição 2.7 [4] *Sejam α and β partições, com o mesmo peso e p um inteiro positivo. Então,*

$$\alpha \preceq_p \beta, \text{ se, e só se, } \beta^* \preceq_p \alpha^*.$$

3 Partições e p -matrizes

Fixemos um inteiro positivo, p . Recordemos que uma matriz diz-se uma p -matriz se todas as suas posições forem números inteiros não negativos e não superiores a p . Dada uma p -matriz, $A = [a_{ij}]$, do tipo m por n (com m linhas e n colunas), podemos calcular a soma da linha i da matriz A , que designamos por R_i , i.e.,

$$\sum_{t=1}^n a_{it} = R_i,$$

e a soma da coluna j da matriz A , que designamos por S_j , i.e.,

$$\sum_{t=1}^m a_{tj} = S_j.$$

Obtemos assim a sequência das somas das linhas de A e a sequência das somas das colunas de A , i.e.,

$$(R_1, \dots, R_m) \text{ e } (S_1, \dots, S_n).$$

Nem sempre estas sequências são partições.

Exemplo 3.1 *Sendo A a 3-matriz do tipo 3 por 5, tal que*

$$A = \begin{bmatrix} 2 & 1 & 3 & 0 & 0 \\ 1 & 2 & 0 & 2 & 2 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

então as somas das linhas de A são

$$R_1 = 6, R_2 = 7, R_3 = 1,$$

pelo que a sequência das somas das linhas de A é $(6, 7, 1)$, não sendo portanto uma partição. Se calcularmos a sequência das somas das colunas de A obtemos a sequência $(4, 3, 3, 2, 2)$, a qual já é uma partição. \square

Repare que se as sequências das somas das linhas e das somas das colunas de uma p -matriz A forem duas partições, porque a soma dos inteiros de cada uma destas partições é o número de entradas não nulas da matriz A , então podemos afirmar que as duas partições são do mesmo peso.

Tal como vimos para as matrizes de zeros e uns, mais complicado é o problema inverso. Sejam R e S duas partições com o mesmo peso, tais que

$R = (R_1, \dots, R_m)$ e $S = (S_1, \dots, S_n)$. Denotamos por $\mathcal{A}^{(p)}(R, S)$ o conjunto das p -matrizes, $[a_{i,j}]$, do tipo m por n , que satisfazem

$$\sum_{t=1}^n a_{i,t} = R_i \text{ e } \sum_{t=1}^m a_{t,j} = S_j,$$

para todo $i = 1, \dots, m$ e $j = 1, \dots, n$.

Uma questão que naturalmente se coloca é a de saber em que condições o conjunto $\mathcal{A}^{(p)}(R, S)$ é não vazio. Esta questão, para $p = 1$ foi resolvida, independentemente, por Gale e por Ryser na década de 50 do século XX, [5, 8].

Teorema 3.2 [1, 5, 8] (*Teorema de Gale-Ryser*) *Sejam R e S duas partições do mesmo peso. Então, $\mathcal{A}(R, S) \neq \emptyset$ se e só se $S \preceq R^*$.*

Quando p é um número inteiro positivo, há vários resultados que nos levam a concluir se $\mathcal{A}^{(p)}(R, S)$ é não vazio, [1, 2, 6, 7]. Recentemente, [4], aplicando a relação p -majoração descrita na secção anterior, obtivemos uma generalização mais natural do teorema de Gale-Ryser, uma vez que o nosso resultado usa o conceito de p -majoração que é uma generalização imediata do conceito de majoração, usado no teorema de Gale-Ryser .

Teorema 3.3 [4] *Sejam R e S partições, com o mesmo peso. Então,*

$$\mathcal{A}^{(p)}(R, S) \text{ é não vazio, se, e só se, } S \preceq_p R^*.$$

Exemplo 3.4 *Vejam uma aplicação deste resultado. Sejam R e S as partições de peso 10, tais que $R = (3, 3, 2, 2)$ e $S = (5, 5)$.*

Temos $R^ = (4, 4, 2)$. Como,*

$$S_1 = 5 \not\leq 4 = R_1^*,$$

então $S \not\preceq R^$ e pelo teorema de Gale-Ryser, não existe nenhuma matriz de zeros e uns com vetor soma das linhas igual a R e vetor soma das colunas igual a S .*

No entanto, se quisermos saber se $\mathcal{A}^{(2)}(R, S)$ é não vazio, poderemos recorrer ao último resultado. Tal como para o caso das matrizes de zeros e uns bastava-nos ver que

$$\begin{aligned} S_1 = 5 &\leq 8 = R_1^* + R_2^* \\ S_1 + S_2 = 10 &\leq 10 = R_1^* + R_2^* + R_3^* + R_4^*, \end{aligned}$$

para concluirmos que $S \preceq_2 R^*$ e, pelo Teorema 3.3, que existe pelo menos uma matriz em $\mathcal{A}^{(2)}(R, S)$. De facto,

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \in \mathcal{A}^{(2)}(R, S).$$

□

4 Algoritmo da divisão e uma matriz de $\mathcal{A}^{(p)}(R, S)$

Sejam p um inteiro positivo, R e S partições de peso t , tais que $R = (R_1, \dots, R_m)$, $S = (S_1, \dots, S_n)$, e $S \preceq_p R^*$. Portanto, $\mathcal{A}^{(p)}(R, S)$ é um conjunto não vazio. Antes de descrever um algoritmo para a construção de uma matriz de $\mathcal{A}^{(p)}(R, S)$, vamos usar o algoritmo da divisão para determinar uma matriz de $\mathcal{A}^{(p)}(F, S)$, para certa partição F , [4]. Esta matriz será depois usada para a construção de uma matriz de $\mathcal{A}^{(p)}(R, S)$.

Definição 4.1 Para cada $1 \leq l \leq n$, sejam d_l e b_l , inteiros não negativos, tais que

$$S_l = d_l p + b_l, \quad \text{com } 0 \leq b_l < p.$$

Seja $\overline{A_S} = [a_{i,j}]$ a matriz, de tipo m por n , tal que, para $1 \leq i \leq m$ e $1 \leq j \leq n$,

$$a_{i,j} = \begin{cases} p & \text{se } i \leq d_j \\ b_j & \text{se } i = d_j + 1 \\ 0 & \text{nos restantes casos} \end{cases}.$$

Repare que o vetor soma das colunas da matriz $\overline{A_S}$ é igual a S . Mas nem sempre o vetor soma das linhas da matriz $\overline{A_S}$ é igual a R .

Exemplo 4.2 Sejam $R = (7, 7, 6, 5)$, $S = (6, 5, 5, 5, 4)$ e $p = 3$. Então temos

$$S \preceq_3 R^* = (4, 4, 4, 4, 3, 2).$$

Como $S_1 = 6 = 2 \times 3 + 0 = d_1 p + b_1$, obtemos $d_1 = 2$ e $b_1 = 0$.

Como $S_2 = 5 = 1 \times 3 + 2 = d_2 p + b_2$, obtemos $d_2 = 1$ e $b_2 = 2$.

Como $S_3 = 5 = 1 \times 3 + 2 = d_3 p + b_3$, obtemos $d_3 = 1$ e $b_3 = 2$.

Como $S_4 = 5 = 1 \times 3 + 2 = d_4 p + b_4$, obtemos $d_4 = 1$ e $b_4 = 2$.

Como $S_5 = 4 = 1 \times 3 + 1 = d_5 p + b_5$, obtemos $d_5 = 1$ e $b_5 = 1$.

Consequentemente,

$$\begin{array}{llll}
 a_{1,1} = 3 = p & \text{porque} & 1 \leq d_1 = 2, \\
 a_{2,1} = 3 = p & \text{porque} & 2 \leq d_1 = 2, \\
 a_{3,1} = 0 = b_1 & \text{porque} & 3 = d_1 + 1, \\
 a_{4,1} = 0 & \text{porque} & 4 \not\leq d_1 = 2, \quad 4 \neq d_1 + 1 = 3, \\
 a_{1,2} = 3 = p & \text{porque} & 1 \leq d_2 = 1, \\
 a_{2,2} = 2 = b_2 & \text{porque} & 2 = d_2 + 1, \\
 a_{3,2} = 0 & \text{porque} & 3 \not\leq d_2 = 1, \quad 3 \neq d_2 + 1 = 2, \\
 a_{4,2} = 0 & \text{porque} & 4 \not\leq d_2 = 1, \quad 4 \neq d_2 + 1 = 2,
 \end{array}$$

e assim sucessivamente.

Então,

$$\overline{A}_S = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Repare que a partição soma das colunas de A_S é S e a partição soma das linhas de A_S é $F = (15, 10, 0, 0)$. \square

5 Construção de uma matriz em $\mathcal{A}^{(p)}(R, S)$

Tal como na seção anterior, sejam p um inteiro positivo, R e S partições de peso t , tais que $R = (R_1, \dots, R_m)$, $S = (S_1, \dots, S_n)$, e $S \preceq_p R^*$. Portanto, $\mathcal{A}^{(p)}(R, S)$ é um conjunto não vazio. A matriz m por n apresentada na definição 4.1 é o ponto de partida do nosso algoritmo, e a matriz de $\mathcal{A}^{(p)}(R, S)$ é construída linha por linha começando na última linha, [4].

Algoritmo para construir uma matriz em $\mathcal{A}^{(p)}(R, S)$

Para tornar o algoritmo mais compreensível, antes da descrição de cada passo do algoritmo, damos uma pequena ideia do objetivo desse passo.

Passo-1. Começamos com a matriz \overline{A}_S , dada na definição 4.1, e $R = (R_1, \dots, R_m)$;

Passo-2. Seja $E = [e_{ij}] = \overline{A}_S$;

Passo-3. Seja z o número de linhas de E ;

Passo-4. (Neste passo, comparamos a soma dos elementos da linha z da matriz E , última linha de E , e o valor de R_z .)

Se $\sum_{j=1}^n e_{zj} = R_z$, então vamos para o Passo-9., com $C = E$.

Caso contrário, ou seja, se $\sum_{j=1}^n e_{zj} < R_z$, então vamos para o Passo-5.;

Passo-5. (Neste passo, usando a relação de ordem lexicográfica \gg , ou seja

$$(x_1, x_2) \gg (y_1, y_2) \text{ se, e só se, } x_1 > y_1 \text{ ou } x_1 = y_1 \text{ e } x_2 > y_2,$$

para determinarmos qual o elemento da matriz E que não está na última linha de E e que irá ser alterado.)

Seja (r, l) o maior par (pela relação de ordem lexicográfica) que satisfaz

$$1 \leq r \leq z - 1, \quad 1 \leq l \leq n, \quad e_{rl} \neq 0, \quad \text{e } e_{zl} \neq p;$$

Passo-6. (Neste passo estabelecemos o valor que irá afetar o elemento da matriz E , encontrado no passo anterior e o elemento de E que está na mesma coluna deste, mas na última linha de E .)

Seja

$$f_{r,l} = \min \left\{ p - e_{zl}, \quad e_{rl}, \quad R_z - \sum_{j=1}^n e_{zj} \right\};$$

Passo-7. (Neste passo, usando o valor que encontramos no passo anterior, modificamos a matriz E , alterando unicamente duas posições, a encontrada no Passo-5. e a que está na mesma coluna, mas na última linha de E .)

Seja $C = [c_{ij}]$ a matriz z por n tal que

$$c_{ij} = \begin{cases} e_{ij} & \text{se } (ij) \notin \{(rl), (zl)\} \\ e_{rl} - f_{r,l} & \text{se } (ij) = (rl) \\ e_{zl} + f_{r,l} & \text{se } (ij) = (zl) \end{cases};$$

Passo-8. (Neste passo, com a matriz C , comparamos a soma dos elementos da linha z com o R_z .)

Se $\sum_{j=1}^n c_{zj} < R_z$, repetimos o processo começando no Passo-5., com $E = [e_{ij}] = C$.

Caso contrário, ou seja, se $\sum_{j=1}^n c_{zj} = R_z$, vamos para o Passo-9.;

Passo-9. (Neste passo, limitamo-nos a remover a última linha da matriz, para reiniciarmos ou terminarmos o algoritmo.)

Se $z > 2$, construímos a matriz G , do tipo $(z - 1)$ por n obtida de C removendo a última linha. Repetimos o processo, começando no Passo-3., com $E = [e_{ij}] = G$.

Caso contrário, ou seja, se $z = 2$, colocamos em C as linhas removidas em cada aplicação do Passo-9. e o algoritmo termina. \square

Para terminar, vamos ilustrar este algoritmo com um exemplo.

Exemplo 5.1 Tal como no Exemplo 4.2, seja $R = (7, 7, 6, 5)$, $S = (6, 5, 5, 5, 4)$ e $p = 3$. Como vimos,

$$\overline{A_S} = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Iniciemos o algoritmo com esta matriz. Temos,

$$\sum_{j=1}^5 a_{4,j} = 0, \quad R_4 = 5.$$

De acordo com o algoritmo, o par $(2, 5)$ é o maior par (pela relação de ordem lexicográfica) tal que $a_{25} = 1 \neq 0$ e $a_{45} = 0 \neq p = 3$. Então,

$$f_{2,5} = \min\{3 - 0, 1, 5 - 0\} = 1,$$

e obtemos a matriz

$$C = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Pelo Passo-8., e porque $\sum_{j=1}^5 c_{4j} = 1 < R_4 = 5$, repetimos o processo, começando no Passo-5., com a matriz $E = C$.

Com esta nova aplicação do algoritmo, temos $f_{2,4} = 2$ e uma nova matriz C surge,

$$C = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \end{bmatrix}.$$

Pelo Passo-8., e uma vez que $\sum_{j=1}^5 c_{4j} = 3 < R_4 = 5$, repetimos novamente este processo, começando no Passo-5., com a matriz $E = C$.

Desta vez, obtemos $f_{2,3} = 2$ e a nova matriz

$$C = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 1 \end{bmatrix}.$$

Pelo Passo-8., e uma vez que $\sum_{j=1}^5 c_{4j} = 5 = R_4 = 5$, vamos para o Passo-9., ou seja, removemos a última linha de C e repetimos o processo, começando no Passo-3., com a matriz

$$E = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Repetindo três vezes o processo de Passo-5. a Passo-8., obtemos a matriz

$$C = \begin{bmatrix} 3 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{bmatrix}.$$

Pelo Passo-8., e porque $\sum_{j=1}^5 c_{3j} = 6 = R_3 = 6$, vamos para o Passo-9., que nos leva a remover a última linha de C e a repetirmos o processo, começando no Passo-3., com a matriz

$$E = \begin{bmatrix} 3 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Repetindo três vezes o processo de Passo-5. a Passo-8., obtemos a matriz

$$C = \begin{bmatrix} 3 & 3 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 2 \end{bmatrix}.$$

Pelo Passo-8., e porque $\sum_{j=1}^5 c_{2j} = 7 = R_2 = 7$ vamos para o Passo-9., onde uma vez que o número de linhas desta matriz é dois, o algoritmo termina. Colocando nesta última matriz as linhas removidas em cada aplicação do Passo-9., obtemos a matriz

$$A = \begin{bmatrix} 3 & 3 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 2 \\ 3 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{bmatrix},$$

que é uma matriz de $\mathcal{A}^{(3)}(R, S)$.

Esquematizando a construção desta matriz A , temos

$$\begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 2 & 1 \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} 3 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 3 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 3 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 & 2 \\ 3 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{bmatrix}$$

Pelo que a matriz de $\mathcal{A}^{(3)}(R, S)$ construída por este algoritmo é a matriz

$$A = \begin{bmatrix} 3 & 3 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & 2 \\ 3 & 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 & 1 \end{bmatrix}$$

□

Referências

- [1] R.A. Brualdi, Combinatorial matrix classes, *Encyclopedia of Mathematics and its Applications*, vol. 108: Cambridge University Press, Cambridge (2006).
- [2] J.A. Dias da Silva e A. Fonseca, Constructing integral matrices with given line sums, *Linear Algebra App.*, 431: 1553-1563 (2009).

- [3] R. Fernandes and H.F. da Cruz, Algoritmo de construção de matrizes de zeros e uns com soma das filas prescritas, *Boletim da Sociedade Portuguesa da Matemática*, 70: 1-14 (2014).
- [4] R. Fernandes and H.F. da Cruz, A canonical construction for nonnegative integral matrices with given line sums, *Linear Algebra App.*, 484; 304-321 (2015).
- [5] D. Gale, A theorem on flows in networks, *Pacific J. Math*, 7: 1073-1082 (1957).
- [6] L. Mirsky, Combinatorial theorems and integral matrices, *J. Combin. Theory*, 5: 30-44 (1968).
- [7] L. Mirsky, Transversal Theory, *Academic Press*, New York (1971).
- [8] H. J. Ryser, Combinatorial properties of matrices of zeros and ones, *Canad. J. Math*, 9: 371-377 (1957).

A NOTE ON REVERSIBILITY AND PELL EQUATIONS

Mário Bessa

Departamento de Matemática
Universidade da Beira Interior
Covilhã, Portugal
e-mail: bessa@ubi.pt

Maria Carvalho, Alexandre A. P. Rodrigues

Centro de Matemática da Universidade do Porto
Faculdade de Ciências da Universidade do Porto
Porto, Portugal
e-mail: mpcarval@fc.up.pt
alexandre.rodrigues@fc.up.pt

Resumo: Esta nota é sobre automorfismos hiperbólicos do toro que são reversíveis relativamente a uma involução linear conservativa. Por estarmos em dimensão dois, será possível associar à reversibilidade uma equação de Pell generalizada de cujo conjunto de soluções retiraremos informação adicional sobre aquela propriedade. Provaremos ainda que a reversibilidade é uma propriedade rara neste contexto e descreveremos o caso genérico.

Abstract: This note concerns hyperbolic toral automorphisms which are reversible with respect to a linear area-preserving involution. Due to the low dimension, we will be able to associate the reversibility with a generalized Pell equation from whose set of solutions we will infer further information. Additionally, we will show that reversibility is a rare feature and will characterize the generic setting.

palavras-chave: Automorfismos hiperbólicos do toro; reversibilidade; equações de Pell generalizadas.

keywords: Hyperbolic toral automorphisms; reversibility; generalized Pell equations.

1 Introduction

Let M be a compact, connected, smooth, Riemannian two-dimensional manifold without boundary and μ its normalized area. Denote by $\text{Diff}_\mu^1(M)$ the set of all area-preserving C^1 -diffeomorphisms of M endowed with the C^1 -topology. A diffeomorphism $f : M \rightarrow M$ is said to be Anosov if M is a

hyperbolic set for f , that is, if the tangent bundle of M admits a splitting $E^s \oplus E^u$ such that there exist an adapted norm $\| \cdot \|$ and a constant $\sigma \in (0, 1)$ satisfying $\|Df_x(v)\| \leq \sigma$ and $\|Df_x^{-1}(u)\| \leq \sigma$ for every $x \in M$ and any unitary vectors $v \in E_x^s$ and $u \in E_x^u$. It is known, after [6, 8], that among the 2-dimensional manifolds only the torus (we will denote by \mathbb{T}^2) may support this type of diffeomorphisms. Moreover, each Anosov diffeomorphism on \mathbb{T}^2 is topologically conjugate to a linear model, that is, to a diffeomorphism induced on the torus by an element of the linear group $SL(2, \mathbb{Z})$ of the 2×2 matrices with integer entries, determinant equal to ± 1 and whose eigenvalues do not belong to the unit circle (the so called toral automorphisms). In this note, we consider linear Anosov diffeomorphisms in $\text{Diff}_\mu^1(\mathbb{T}^2)$ which exhibit a certain kind of symmetry. More precisely, let $R : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ be a diffeomorphism such that $R \circ R$ is the identity map of \mathbb{T}^2 (R is called an involution); our attention will be turned to the subset $\text{Diff}_{\mu, R}^1(\mathbb{T}^2)$ of maps $f \in \text{Diff}_\mu^1(\mathbb{T}^2)$, said to be R -reversible, such that R conjugates f and f^{-1} , that is,

$$R \circ f = f^{-1} \circ R.$$

Reversibility plays a fundamental role in physics and dynamical systems. In view of the many known applications, the references [7] and [11] present a thorough survey on reversible systems. Regarding this subject, important work was done in [12] and [1]. In the former, the authors derived necessary conditions for local reversibility within mappings of the plane with a symmetric fixed point, expressing them through the quadratic and cubic coefficients of the Taylor expansion about the fixed point. This way, they established an efficient negative criteria to show that a mapping is not reversible. On the other hand, on Section 2.2 of [1] we may read a detailed characterization of the kind of groups that are admissible as groups of reversible symmetries associated with unimodular matrices. Using algebraic techniques, the authors unveiled the possible structure of those groups, showing that it is completely resolvable when the matrices belong to $GL(2, \mathbb{Z})$ or $PG(2, \mathbb{Z})$, as summarized in [1, Theorem 3].

Our work has followed a different, not purely algebraic, approach and addressed other questions, thus complementing the information disclosed by the two previous references. First, we will present a simple method to construct a (non unique) Anosov diffeomorphism that anti-commutes with a given linear involution (Section 4). Conversely, after concluding that an automorphism that reverses orientation is not reversible, we will settle (cf. Section 5) a connection between the existence of reversible symmetries for a hyperbolic toral automorphism and the set of solutions of a generalized

Pell equation [9]. Moreover, we will make clear in Section 6 why generically, in the C^1 topology, the r -centralizer (definition (1)) of a toral Anosov diffeomorphism is trivial, summoning a similar result obtained in [2] within the conservative context for the centralizer. We believe that an extension of these results may be proved for reversing symmetries that are not involutions, and on higher dimensions. Yet, those are so far open problems.

2 Overview

Whereas it is often useful to check if there is a dynamical system which is reversible under a given involution, we may also be interested in ascertaining if there exists an involution under which a given dynamics is reversible, and to what extent this is a common feature. The latter query is hard to answer in general, and that is why we will confine our study to hyperbolic toral automorphisms, benefiting both from the linear structure and the low dimensional setting. In what follows we will address three questions.

Q₁. Given a linear involution $R \in \text{Diff}_\mu^1(\mathbb{T}^2)$, is there an R -reversible linear area-preserving Anosov diffeomorphism f ?

The answer is obvious (and no) if $R = \pm Id$, the so called trivial case. Concerning the other possible involutions, we will look for a linear Anosov diffeomorphism f whose derivative at any point of \mathbb{T}^2 is a fixed linear map with matrix $L \in SL(2, \mathbb{Z})$. Observe that, as R is induced by a matrix $A \in SL(2, \mathbb{Z})$ as well, if we lift the equality $R \circ f = f^{-1} \circ R$ by differentiating it at any point of \mathbb{T}^2 , we obtain $A \circ L = L^{-1} \circ A$; as we will verify in Remark 3.2, this equality implies that $\det L = 1$. Analyzing the entries of the matrices involved we will answer positively to question Q_1 .

Proposition 2.1 *Let \mathcal{L} be the set of linear Anosov diffeomorphisms on \mathbb{T}^2 . If $R \in \text{Diff}_\mu^1(\mathbb{T}^2) \setminus \{\pm Id\}$ is a linear involution, then $\text{Diff}_{\mu,R}^1(\mathbb{T}^2) \cap \mathcal{L} \neq \emptyset$.*

The proof of this result will be presented on Section 4. The argument is straightforward once the list of involutions in $SL(2, \mathbb{Z})$ is determined. Afterwards we will deal with the converse query.

Q₂. Given a linear Anosov diffeomorphism $f \in \text{Diff}_\mu^1(\mathbb{T}^2)$, is there a linear involution $R \in \text{Diff}_\mu^1(\mathbb{T}^2)$ such that f is R -reversible?

If such an f reverses orientation (meaning that f is induced by a linear transformation $L \in SL(2, \mathbb{Z})$ such that $\det(L) = -1$), then f is never R -reversible for every linear involution $R \in \text{Diff}_\mu^1(\mathbb{T}^2)$. We will show (cf. Section 5) that, if f preserves orientation, then the reversibility of f with respect to a linear involution R determines a generalized Pell equation (see [4, 10] for this concept) whose set of solutions displays the entries of the R 's matrix.

Theorem A *Let $f \in \text{Diff}_\mu^1(\mathbb{T}^2)$ be a linear Anosov diffeomorphism induced by a transformation*

$$L = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z})$$

such that $\det(L) = 1$. Then:

- (i) *b divides $a - d$ if and only if f is R -reversible, where R is the projection on \mathbb{T}^2 of either*

$$A = \begin{pmatrix} 1 & 0 \\ \frac{d-a}{b} & -1 \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} -1 & 0 \\ \frac{a-d}{b} & 1 \end{pmatrix}.$$

- (ii) *c divides $a - d$ if and only if f is R -reversible, where R is the projection on \mathbb{T}^2 of either*

$$A = \begin{pmatrix} 1 & \frac{d-a}{c} \\ 0 & -1 \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} -1 & \frac{a-d}{c} \\ 0 & 1 \end{pmatrix}.$$

- (iii) *Given $\alpha, \beta \in \mathbb{Z} \setminus \{0\}$ such that $1 - \alpha^2 \neq 0$ and β divides $1 - \alpha^2$, consider the involution R obtained by projecting on \mathbb{T}^2 the matrix*

$$A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix} \in SL(2, \mathbb{Z}).$$

Then f is R -reversible if and only if the generalized Pell equation $x^2 - Dy^2 = N$, where $D = (a+d)^2 - 4$ and $N = 4b^2$, has among its solutions

$$x = 2b\alpha + (d - a)\beta \quad \text{and} \quad y = \beta.$$

Observe that, if $f, g \in \text{Diff}_{\mu, R}^1(\mathbb{T}^2)$, then $R \circ f^{-1} = f \circ R$ and $R \circ g^{-1} = g \circ R$ but $R \circ (f \circ g) = (f^{-1} \circ R) \circ g = (f^{-1} \circ g^{-1}) \circ R = (g \circ f)^{-1} \circ R$; this means that

the set $\text{Diff}_{\mu,R}^1(\mathbb{T}^2)$, endowed with the composition of maps is, in general, not a group. Yet, $\text{Diff}^1(\mathbb{T}^2)$, $\text{Diff}_{\mu}^1(\mathbb{T}^2)$ and $\text{Diff}_{\mu,R}^1(\mathbb{T}^2)$ are Baire spaces (cf. [5] and Subsection 3.2), so it is natural to adopt another perspective towards reversibility.

Q₃. Is reversibility generic?

Given a diffeomorphism $f : \mathbb{T}^2 \rightarrow \mathbb{T}^2$, the r -centralizer of f is the set

$$\mathcal{Z}_r(f) = \{R \in \text{Diff}^1(\mathbb{T}^2) : R \circ f = f^{-1} \circ R\}. \quad (1)$$

$\mathcal{Z}_r(f)$ is said to be trivial if it is either empty or reduces to a set $\{R \circ f^n : n \in \mathbb{Z}\}$ for some $R \in \text{Diff}^1(\mathbb{T}^2)$. We will verify on Section 6 that, for any diffeomorphism f in a residual subset of $\text{Diff}_{\mu}^1(\mathbb{T}^2)$, the equation $R \circ f = f^{-1} \circ R$ has only trivial solutions.

Theorem B *C^1 -generically in $\text{Diff}_{\mu}^1(M)$, the r -centralizer is trivial.*

Notice that, in spite of the fact that, on the torus \mathbb{T}^2 , each Anosov diffeomorphism is conjugate to a hyperbolic toral automorphism, the conclusions we have drawn cannot be extended to all Anosov diffeomorphisms because the R -reversibility, for a fixed R , is not preserved by conjugacy. In fact, if $f \in \text{Diff}_{\mu,R}^1(\mathbb{T}^2)$ is conjugate to $g \in \text{Diff}_{\mu}^1(\mathbb{T}^2)$ through a homeomorphism h , then, although we have $(R \circ h) \circ g = f^{-1}(R \circ h)$, the map g may be not R -reversible. Nevertheless, the last equation indicates that $(h^{-1} \circ R \circ h) \circ g = g^{-1} \circ (h^{-1} \circ R \circ h)$, so a diffeomorphism conjugate to an R -reversible linear Anosov diffeomorphism is reversible as well, although with respect to another involution, which is conjugate to R but may be neither linear nor even differentiable.

3 Linear involutions on \mathbb{T}^2

We start characterizing the linear involutions $R : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ of the torus, induced by matrices A in $SL(2, \mathbb{Z})$. Given this, we will describe the set of fixed points of such involutions.

3.1 Classification

After differentiating the equality $R^2 = Id_{\mathbb{T}^2}$ at any point of \mathbb{T}^2 , we obtain $A^2 = Id_{\mathbb{R}^2}$. Comparing the entries of the matrices in this equality, we conclude that:

Lemma 3.1 *A matrix $A \in SL(2, \mathbb{Z}) \setminus \{\pm Id\}$ is an involution if and only if it belongs to the following list:*

- $A = \begin{pmatrix} \pm 1 & 0 \\ \gamma & \mp 1 \end{pmatrix}$ or its transpose, for some $\gamma \in \mathbb{Z}$.
- $A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$ for $\alpha \in \mathbb{Z} \setminus \{\pm 1\}$ and $\beta \in \mathbb{Z} \setminus \{0\}$ such that β divides $1 - \alpha^2$.

Let A be a matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2, \mathbb{Z})$ such that $A^2 = Id$. These properties imply that $\alpha\delta - \beta\gamma = \pm 1$ and

$$\begin{cases} \alpha^2 + \beta\gamma = 1 \\ \beta(\alpha + \delta) = 0 \\ \gamma(\alpha + \delta) = 0 \\ \gamma\beta + \delta^2 = 1 \end{cases}$$

so

$$\begin{cases} \beta = 0 \vee \alpha = -\delta \\ \gamma = 0 \vee \alpha = -\delta. \end{cases}$$

1st case: $\beta = 0$. One must have $\alpha = \pm 1$ and $\delta = \pm 1$. If $\alpha = \delta = 1$ or $\alpha = \delta = -1$, we conclude that $\gamma = 0$ and so $A = \pm Id$. Therefore, $-\alpha = \delta = 1$ or $\alpha = -\delta = 1$, and there are no restrictions on the value of γ .

Hence $A = \begin{pmatrix} \pm 1 & 0 \\ \gamma & \mp 1 \end{pmatrix}$ for $\gamma \in \mathbb{Z}$.

2nd case: $\gamma = 0$. Again $\alpha = \pm 1$ and $\delta = \pm 1$, and so either $-\alpha = \delta = 1$ or $\alpha = -\delta = 1$. Therefore $A = \begin{pmatrix} \pm 1 & \beta \\ 0 & \mp 1 \end{pmatrix}$ for $\beta \in \mathbb{Z}$.

3rd case: $\beta \neq 0$ and $\gamma \neq 0$. We must have $\alpha = -\delta$, hence $\alpha^2 + \beta\gamma = 1$ and so $\gamma = \frac{1-\alpha^2}{\beta}$. Moreover, for A to belong to $SL(2, \mathbb{Z})$, the entry β must divide $1 - \alpha^2$. Thus $A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$. To prevent repetitions with the preceding case, α has to be different from ± 1 .

Remark 3.2 *From the previous argument we also conclude that:*

1. A linear map $A \in SL(2, \mathbb{Z}) \setminus \{\pm Id\}$ is an involution if and only if $\det A = -1$ and $\text{trace } A = 0$.¹ Therefore, the eigenvalues of both A and $-A$ are precisely $-1, 1$.
2. Given a linear involution $A \in SL(2, \mathbb{Z}) \setminus \{\pm Id\}$ and a matrix $L \in SL(2, \mathbb{Z}) \setminus \{\pm Id\}$, then L is A -reversible (that is, $A \circ L = L^{-1} \circ A$) if and only if $(LA)^2 = Id$. If the spectrum of the matrix L does not contain $-1, 1$, then the involution LA cannot be $\pm Id$ (otherwise, $L = A$ or $L = -A$ and so $-1, 1$ would be eigenvalues of L). Therefore, there is a linear involution $B \in SL(2, \mathbb{Z}) \setminus \{\pm Id\}$ such that $L = BA$. In particular, we must have $\det L = 1$.

3.2 Fixed points

From the previous description of the matrices A and by solving the equation $A(x, y) = (x, y)$ in \mathbb{R}^2 , we deduce that the fixed point set of a linear non-trivial involution R of the torus is a finite union of smooth closed curves obtained by projecting subspaces of \mathbb{R}^2 with dimension one. Firstly, we know that such an R is induced by a matrix $A \in SL(2, \mathbb{Z})$ as specified by Lemma 3.1, so we get

A	<i>Fixed points subspace of A</i>
$\begin{pmatrix} 1 & 0 \\ \gamma & -1 \end{pmatrix}$	$y = \frac{\gamma}{2}x$
$\begin{pmatrix} 1 & \gamma \\ 0 & -1 \end{pmatrix}$	$y = 0$
$\begin{pmatrix} -1 & 0 \\ \gamma & 1 \end{pmatrix}$	$x = 0$
$\begin{pmatrix} -1 & \gamma \\ 0 & 1 \end{pmatrix}$	$x = \frac{\gamma}{2}y$
$\begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}, \beta \neq 0$	$y = \frac{1-\alpha}{\beta}x$

Moreover, if $\pi : \mathbb{R}^2 \rightarrow \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ denotes the usual projection and $A(x, y) = (x, y)$, then $R(\pi(x, y)) = \pi(x, y)$, so the projections on the torus of the subspaces of the previous table are sets of fixed points by R . As the slopes of these lines are rational numbers, these projections are closed curves on \mathbb{T}^2 .

Conversely, observe that $R(\pi(x, y)) = \pi(x, y)$ if and only if $\pi(A(x, y)) = \pi(x, y)$, and this implies that there exists $(n, m) \in \mathbb{Z}^2$ such that $A(x, y) =$

¹We thank the referee for calling our attention to this property.

$(x, y) + (n, m)$. So, for instance, if $A = \begin{pmatrix} 1 & 0 \\ \gamma & -1 \end{pmatrix}$, then

$$A(x, y) = (x, y) + (n, m) \Leftrightarrow x \in \mathbb{R}, n = 0, y = \frac{\gamma}{2}x - \frac{m}{2}$$

and therefore, for the involution R induced by such a matrix A , we have

$$\text{Fix}(R) = \pi \left(\left\{ (x, \frac{\gamma}{2}x) : x \in \mathbb{R} \right\} \right) \cup \pi \left(\left\{ (x, \frac{\gamma}{2}x - \frac{1}{2}) : x \in \mathbb{R} \right\} \right).$$

This is the union of two closed curves on \mathbb{T}^2 . The other cases are analogous.

4 Answer to question Q_1

Recall that, if a linear Anosov diffeomorphism f is induced by a matrix $L(x, y) = (ax + by, cx + dy)$ of $SL(2, \mathbb{Z})$, then the entries of L must satisfy the conditions:

- (IL) (Integer lattice invariance) $a, b, c, d \in \mathbb{Z}$ and $ad - bc = \pm 1$.
- (H1) (Hyperbolicity) If $ad - bc = 1$, then $(a + d)^2 - 4 > 0$.
- (H2) (Hyperbolicity) If $ad - bc = -1$, then $(a + d)^2 + 4$ is not a perfect square.

These requirements explain why a linear Anosov diffeomorphism is never $\pm Id$ -reversible. Indeed, if $R = Id$ and f is R -reversible, then $f^2 = Id$; however, this equality does not hold among Anosov diffeomorphisms, whose periodic points are hyperbolic and so isolated. If $R = -Id$ and f is induced by a matrix $L \in SL(2, \mathbb{Z})$ and is R -reversible, then the equality $(-Id) \circ L = L^{-1} \circ (-Id)$ yields

$$\begin{cases} a = d, b = c = 0, & \text{if } ad - bc = 1 \\ a = -d, & \text{if } ad - bc = -1 \end{cases}$$

which contradicts one of the properties (H1) or (H2).

Going through the available matrices A , given by Lemma 3.1, we will determine, for each R , an orientation-preserving R -reversible and linear Anosov diffeomorphism f .

(a) $A = \begin{pmatrix} 1 & 0 \\ \gamma & -1 \end{pmatrix}$. We start noticing that the equality $A \circ L = L^{-1} \circ A$, with $\det(L) = 1$, is equivalent to $b\gamma = d - a$. If $\gamma = 0$, we may take $L = \begin{pmatrix} a & b \\ c & a \end{pmatrix}$ with integer entries such that $a^2 - bc = 1$ and $4a^2 > 4$ (so $b \neq 0$ and $c \neq 0$). For instance, $a = d = 3$, $b = 4$ and $c = 2$. If $\gamma \neq 0$, it must divide $d - a$ and L has to be $\begin{pmatrix} a & \frac{d-a}{\gamma} \\ c & d \end{pmatrix}$, with integer entries such that $ad - bc = 1$, $(a + d)^2 > 4$, $d - a \neq 0$ and $c \neq 0$. For example, $a = \gamma \in \mathbb{Z} \setminus \{0\}$, $b = 1$, $c = 2\gamma^2 - 1$ and $d = 2\gamma$. As the reversibility condition $A \circ L = L^{-1} \circ A$, with $\det(L) = 1$, is equivalent to $A^T \circ L^T = (L^T)^{-1} \circ A^T$, with $\det(L^T) = 1$, the case of the transpose matrix is equally solved.

(b) $A = \begin{pmatrix} -1 & 0 \\ \gamma & 1 \end{pmatrix}$. As in the previous case, the equality $A \circ L = L^{-1} \circ A$, with $\det(L) = 1$, is equivalent to $b\gamma = a - d$. So, if $\gamma = 0$, we may take $L = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$. If $\gamma \neq 0$, we may choose, for instance, $a = \gamma \in \mathbb{Z} \setminus \{0\}$, $b = -1$, $c = 1 - 2\gamma^2$ and $d = 2\gamma$. Again, for $A^T = \begin{pmatrix} -1 & \beta \\ 0 & 1 \end{pmatrix}$, we may just pick the Anosov diffeomorphism induced by L^T .

(c) $A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$, with $\alpha, \beta, 1 - \alpha^2 \neq 0$ and β a divisor of $1 - \alpha^2$. The equality

$$\begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$$

is equivalent to the equation

$$\alpha\beta a + (1 - \alpha^2)b + \beta^2 c - \alpha\beta d = 0. \quad (2)$$

To ease our task, we may try to find a matrix satisfying $a = d$. Under this assumption, equation (2) becomes $(1 - \alpha^2)b + \beta^2 c = 0$. As c must also comply with the equality $a^2 - bc = 1$ and b cannot be zero, we must have $c = \frac{\alpha^2 - 1}{b}$. In addition, we know that β divides $1 - \alpha^2$ and that $\alpha^2 \neq 1$, so $4\alpha^2 - 4 > 0$. Therefore, a convenient choice is $a = d = \alpha$, $b = \pm\beta$ and $c = \frac{\alpha^2 - 1}{\pm\beta}$. Namely,

$$L = \begin{pmatrix} \alpha & \beta \\ \frac{\alpha^2 - 1}{\beta} & \alpha \end{pmatrix} \quad \text{or} \quad L = \begin{pmatrix} \alpha & -\beta \\ \frac{1 - \alpha^2}{\beta} & \alpha \end{pmatrix}.$$

5 Answer to question Q_2

Let f be a linear Anosov diffeomorphism, induced by a matrix $L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of $SL(2, \mathbb{Z})$. As seen in Remark 3.2, we must have $ad - bc = 1$. Take a linear involution R given by the projection on the torus of a matrix A as described by Lemma 3.1.

Case 1: $A = \begin{pmatrix} 1 & 0 \\ \gamma & -1 \end{pmatrix}$ or $A = \begin{pmatrix} -1 & 0 \\ \gamma & 1 \end{pmatrix}$. The reversibility equality is equivalent to $b\gamma = d - a$ or $b\gamma = a - d$. So there is such an involution A if and only if b divides $d - a$, in which case only one valid γ exists, namely, $\gamma = \frac{d-a}{b}$ or $\gamma = \frac{a-d}{b}$, respectively.

Case 2: $A = \begin{pmatrix} 1 & \gamma \\ 0 & -1 \end{pmatrix}$ or $A = \begin{pmatrix} -1 & \gamma \\ 0 & 1 \end{pmatrix}$. Dually, the reversibility condition is equivalent to $c\gamma = d - a$ or $c\gamma = a - d$. So there is such an involution A if and only if c divides $d - a$, and then we get a unique value for γ .

Case 3: $A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$, where $\alpha, \beta, 1 - \alpha^2 \neq 0$ and β divides $1 - \alpha^2$. The pairs $(\alpha, \beta) \in \mathbb{Z}^2$ for which f is R -reversible are the integer solutions of the equation, in the variables α and β , given by

$$b\alpha^2 + \alpha\beta(d - a) - \beta^2c = b \quad (3)$$

which satisfy the mentioned constraints. This quadratic form defines a conic whose kind depends uniquely on the sign of

$$\Delta = (d - a)^2 + 4bc = (a + d)^2 - 4$$

which we know to be always positive within our setting. So the conic is a hyperbola. After the change of variables

$$x = 2b\alpha + (d - a)\beta \quad \text{and} \quad y = \beta$$

the equation of the conic becomes

$$x^2 - Dy^2 = N \quad (4)$$

where $D = \Delta = (a + d)^2 - 4 > 0$ and $N = 4b^2$. Thus the problem of finding the intersections of the conic (3) with the integer lattice is precisely the one

of existence of solutions of the generalized Pell equation (4). According to [3, 4, 10], this sort of Pell equation either has no integer solutions or has infinitely many, and there are several efficient algorithms to determine which one holds in each particular case. However, if they exist, the solutions we are interested in have also to fulfill the other requirements, namely $\alpha, \beta, 1 - \alpha^2 \neq 0$ and β divides $1 - \alpha^2$.

5.1 Examples

We may test the previous information in a few examples.

<i>Anosov</i>			<i>Involutions</i>		
	$\begin{pmatrix} 1 & 0 \\ \gamma & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & \gamma \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ \gamma & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & \gamma \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$
$\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$	$\gamma = 0$	$\gamma = 0$	$\gamma = 0$	$\gamma = 0$	e.g. $\begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}$
$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$	$\gamma = -1$	$\gamma = -1$	$\gamma = 1$	$\gamma = 1$	e.g. $\begin{pmatrix} 5 & 3 \\ -8 & -5 \end{pmatrix}$
$\begin{pmatrix} 4 & 9 \\ 7 & 16 \end{pmatrix}$	–	–	–	–	–

When $L = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$, the generalized Pell equation is $x^2 - 12y^2 = 4$ and there are infinitely many matrices A of the third kind which correspond to linear involutions R such that f is R -reversible. Similarly, for $L = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ the generalized Pell equation is $x^2 - 5y^2 = 4$ and there are infinitely many involutions of type 3. The third example in this table has no linear involutions, although its Pell equation $x^2 - 396y^2 = 324$ has infinitely many solutions.

Remark 5.1 Notice that, if R is an involution such that $R \circ f = f^{-1} \circ R$, then, for each $n \in \mathbb{Z}$, the diffeomorphism $R \circ f^n$ is also an involution, since $(R \circ f^n)^2 = (R \circ f^n) \circ (f^{-n} \circ R) = Id$; and f is $R \circ f^n$ -reversible due to the equality

$$(R \circ f^n) \circ f = (R \circ f) \circ f^n = (f^{-1} \circ R) \circ f^n = f^{-1} \circ (R \circ f^n).$$

Therefore, once such an involution R is found for an Anosov diffeomorphism f , then we have infinitely many involutions with respect to which f

is reversible. This is so because, as no non-trivial power of an Anosov diffeomorphism is equal to the Identity, we have $R \circ f^k \neq R \circ f^m$ for every $k \neq m \in \mathbb{Z}$.

The next table shows examples of a similar construction of generalized Pell equations and conics in the orientation-reversing setting (although no involution exists). If $A = \begin{pmatrix} \alpha & \beta \\ \frac{1-\alpha^2}{\beta} & -\alpha \end{pmatrix}$, where $\alpha, \beta, 1-\alpha^2 \neq 0$ and β divides $1-\alpha^2$, then the pairs $(\alpha, \beta) \in \mathbb{Z}^2$ satisfying the R -reversibility condition for f should belong to the set of integer solutions, in the variables α and β , of the equations $\alpha b + \beta d = 0$, $\alpha c - \frac{a}{\beta}(1-\alpha^2) = 0$ and $\alpha a + \beta c = -\alpha d + \frac{b}{\beta}(1-\alpha^2)$. The last equality describes a (possibly degenerate) conic

$$b\alpha^2 + \alpha\beta(a+d) + \beta^2c = b$$

whose type is determined by the sign of

$$\Delta = (a+d)^2 - 4bc = (a-d)^2 - 4.$$

Once again, the problem of finding the points of this conic in the integer lattice is linked to the existence of solutions of the generalized Pell equation $x^2 - Dy^2 = N$, where $x = 2b\alpha + (a+d)\beta$, $y = \beta$, $D = \Delta = (a-d)^2 - 4$ and $N = 4b^2$.

<i>Anosov</i>	Δ	<i>Pell eq.</i>	$\#$ of solutions	<i>Conic</i>	<i>Involutions</i>
$\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$	-3	$x^2 + 3y^2 = 36$	6	Ellipse	-
$\begin{pmatrix} 3 & 4 \\ 1 & 1 \end{pmatrix}$	0	$x^2 = 64$	∞	Two lines	-
$\begin{pmatrix} 4 & 5 \\ 1 & 1 \end{pmatrix}$	5	$x^2 - 5y^2 = 100$	∞	Hyperbola	-

6 Answer to question Q_3

Given an area-preserving diffeomorphism f , if $f^2 = Id$, then f^n belongs to $\mathcal{Z}_r(f)$ for all $n \in \mathbb{Z}$; and conversely. However, a generic $f \in \text{Diff}_\mu^1(M)$ does not satisfy the equality $f^n = Id$, for any integer $n \neq 0$. This is due to the fact that, by Kupka-Smale Theorem for area-preserving diffeomorphisms [5], given $k \in \mathbb{N}$, generically in the C^1 topology the periodic orbits of period less or equal to k are isolated.

The existence of hyperbolic toral diffeomorphisms (which are structurally stable) without reversible symmetries shows that there are non-empty open subsets of $\text{Diff}_\mu^1(M)$ where reversibility is absent. Moreover, if $R \neq S$ are in $\mathcal{Z}_r(f)$, then $R \circ S$ belongs to the centralizer of f , as shown by the equalities $(R \circ S) \circ f = R \circ (S \circ f) = R \circ (f^{-1} \circ S) = f \circ (R \circ S)$. Now, according to [2], for a C^1 -generic $f \in \text{Diff}_\mu^1(M)$, the centralizer of f is trivial, meaning that it reduces to the powers of f . Therefore, there must exist $n \in \mathbb{Z}$ such that $S = R \circ f^n$. Thus, if $\mathcal{Z}_r(f) \neq \emptyset$, then its elements are obtained from the composition of one of them with the powers of f . And so, C^1 -generically, $\mathcal{Z}_r(f)$ is trivial.

Acknowledgements

M. Bessa and A. Rodrigues are grateful to António Machiavelo for useful discussions. M. Bessa was partially supported by National Funds through Fundação para a Ciência e a Tecnologia (FCT) project UID/MAT/00212/2013. M. Carvalho and A. Rodrigues were funded by CMUP (UID/MAT/00144/2013), which is financially supported by FCT (Portugal) with national (MEC) and European structural funds through the programs FEDER, under the partnership agreement PT2020. A. Rodrigues has also benefited from the FCT grant SFRH/BPD/84709/2012.

References

- [1] M. Baake and J.A.G. Roberts, “Reversing symmetry of $Gl(2, \mathbb{Z})$ and $PGL(2, \mathbb{Z})$ matrices with connections to cat maps and trace maps”, *J. Phys. A: Math. Gen.*, Vol. 30 (1997), 1549-1573.
- [2] C. Bonatti, S. Crovisier and A. Wilkinson, “ C^1 -generic conservative diffeomorphisms have trivial centralizer”, *J. Mod. Dyn.*, Vol. 2 (2008), 359-373.
- [3] D.A. Buell, *Binary quadratic forms: classical theory and modern computations*, Springer-Verlag, NY, 1989.
- [4] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993.
- [5] R. Devaney, “Reversible diffeomorphisms and flows”, *Trans. Amer. Math. Soc.*, Vol. 218 (1976), 89-113.

- [6] J. Franks, “Anosov diffeomorphisms”, *Global Analysis, Proc. Sympos. Pure Math., Vol. XIV, Berkeley, California. AMS Proceedings*, Vol. 1070, Providence, R.I. (1968), 61-93.
- [7] J. Lamb and J. Roberts, “Time-reversal symmetry in dynamical systems: A survey”, *Physica D*, Vol. 112 (1998), 1-39.
- [8] A. Manning, “There are no new Anosov diffeomorphisms on tori”, *Amer. J. Math.*, Vol. 96, No. 3 (1974), 422-429.
- [9] K. Matthews, “The diophantine equation $x^2 - Dy^2 = N$, $D > 1$, in integers”, *Expo. Math.*, Vol. 18 (2000), 323-331.
- [10] R.E. Mollin, *Fundamental Number Theory with Applications*, CRC Press, Boca Raton, 1998.
- [11] J.A.G. Roberts and G.R.W. Quispel, “Chaos and time-reversal symmetry. Order and chaos in reversible dynamical systems”, *Phys. Rep.*, Vol. 216 (1992), 63-177.
- [12] G.R. Quispel and H.W. Capel, “Local reversibility in dynamical systems”, *J. Phys. A: Math. Gen.*, Vol. 30 (1997), 1549-1573.

Opinião – Cartas ao Director

Editor:
Ana Jacinta Soares

CARTA AO DIRECTOR

Pedro R. S. Antunes e Hugo Tavares

Faculdade de Ciências da Universidade de Lisboa
Lisboa, Portugal
e-mail: prantunes@fc.ul.pt
hrtavares@ciencias.ulisboa.pt

Nuno Costa Dias

Escola Superior Náutica Infante D. Henrique
Paço d’Arcos, Portugal
e-mail: ncdias@meo.pt

Caro Director do Boletim da Sociedade Portuguesa de Matemática,

Com esta carta pretendemos dar um contributo para o debate sobre a qualidade dos processos e dos resultados dos concursos públicos para recrutamento de pessoal docente nas Universidades e Institutos Politécnicos, em particular na área da Matemática.

É notório que alguns destes concursos continuam a decorrer de forma francamente incorreta, como se evidencia no texto, sobre um concurso recente, que vem publicado neste número do Boletim. Infelizmente, não é um caso único. Todos já tivemos conhecimento de concursos cujo edital, composição do júri e grelha de avaliação não satisfazem critérios mínimos de qualidade e isenção, conduzindo facilmente a listas de ordenação dos candidatos que são injustas, não servem o interesse da instituição, nem promovem, como deveriam fazer, a qualidade de ensino e da investigação científica no país.

Há alguns anos atrás, eram frequentes os editais onde o júri, quase na sua totalidade, era constituído por docentes da instituição responsável pela abertura do concurso. Não menos usual, em concursos abertos por certas instituições, eram definidos perfis preferenciais, que valorizavam aspetos menores do currículo, como o tema específico de uma tese de Mestrado ou de Doutoramento ou a frequência de uma dada ação de formação. Eram formas habituais de viciar *a priori* os resultados dos concursos.

Estes casos eram bem conhecidos da comunidade, sendo de assinalar que, na revisão do Estatuto da Carreira Docente Universitária (ECDU) e do Estatuto da Carreira do Pessoal Docente do Ensino Superior Politécnico (ECPDESP) operada pelo Decreto-Lei n.º 205/2009 de 31 de Agosto, tenha sido introduzido um conjunto de princípios, os quais se encontram destacados na introdução do próprio documento:

“A obrigatoriedade de concursos para professores, com júris maioritariamente externos à instituição.”

“O reforço da transparência nos concursos, desde a proibição da adopção de especificações que estreitem de forma inadequada o universo dos candidatos, à publicidade alargada de todas as fases do processo.”

“A valorização, nos concursos, de todas as componentes das funções dos docentes, com expressa consideração do desempenho científico, da capacidade pedagógica e de outras actividades relevantes para a missão da instituição de ensino superior.”

Apesar das melhorias que resultaram dos novos estatutos, ao consultar as grelhas de avaliação de alguns editais mais recentes, percebe-se imediatamente que em certas instituições ainda subsistem muitos problemas. Segundo a lei, a componente científica tem de ser expressamente valorizada nos concursos, mas é comum os editais contabilizarem como sendo “ciência” elementos sem importância, ignorando aquilo que é de facto atividade científica. A título de exemplo, num edital recente não é atribuída nenhuma pontuação pela publicação de artigos científicos. Noutra, esta componente vale 5% da classificação final, sendo a classificação máxima atribuída a quem tenha publicado um ou dois artigos científicos (não é explicitado, como raramente o é, algum critério que permita avaliar a qualidade científica do trabalho). Parece inacreditável que critérios deste género possam ser usados para seleccionar candidatos a posições permanentes no ensino superior, mas são bastante comuns nos concursos de certas instituições.

Poder-se-ia também falar do método de escolha do júri (que é da responsabilidade da instituição promotora), bem como da falta de consequências práticas que advêm para o júri e para a instituição pela escolha de um mau

candidato, ou pela aprovação de um mau edital. Mais importante ainda é a falta de escrutínio das decisões do júri. Pela dificuldade que alguém externo à nossa comunidade tem em compreender uma avaliação feita a currículos e matérias científicas tão específicas, a sociedade em geral não tem uma consciência clara do que se passa. E por mais incorreto que seja o resultado de um concurso (com a notória exceção dos casos em que existe algum erro processual) é praticamente impossível contestá-lo, sendo tipicamente alegada a soberania do júri. Por outro lado, a comunidade académica, na sua larga maioria, opta por não se pronunciar sobre as decisões dos pares. A maioria das discussões que são tidas a este respeito são mantidas “nos corredores”, não são abordadas de forma sistemática nem factual, e não são tornadas públicas, pelo que se perdem no tempo, sem que se tenha retirado grande proveito das mesmas.

Este problema assume ainda uma dimensão maior face ao contexto particular que Portugal atravessa. Houve em anos recentes um grande esforço do país para investimento na ciência, apesar de todas as restrições orçamentais. O país encorajou uma larga faixa da sua juventude mais capaz e competente a prosseguir uma formação científica avançada, e não está a ser capaz de integrar nas instituições nem uma pequena parte desta nova geração. E não é verdade que não existam lugares disponíveis, nem que as instituições não abram concursos. Mas parece haver em muitas instituições uma espécie de “carreira paralela”, que passa pela contratação direta de docentes convidados, muitos dos quais ex-alunos das próprias instituições, que mais tarde são os candidatos naturais a vencer os concursos para posições permanentes. De fora ficam os candidatos sem associação às instituições, que estão a realizar Pós-Doutoramentos, têm posições de Investigador FCT ou estão fora do país. Tipicamente aqueles que são cientificamente mais capazes e produtivos e com carreiras mais promissoras. Fazem muita falta às instituições, mas vão sendo sistematicamente excluídos dos concursos por procedimentos como os que descrevemos.

Alguns acreditarão que uma nova legislação pode promover algumas mudanças. Por exemplo, impondo, como em alguns países, que uma parte das novas posições seja preenchida por candidatos sem ligação anterior à instituição. Outra hipótese será a de atuar *a posteriori*, penalizando as instituições que contratem docentes e investigadores que venham a revelar-se pouco produtivos ou incompetentes. Seja qual for a proposta, a nossa convicção é que a redução deste tipo de situações depende em larga medida de ter uma comunidade científica mais atenta e interventiva, nomeadamente através da pressão dos pares. Este não é um problema para quem está a participar

como candidato em concursos, mas de toda a nossa comunidade: para quem elabora editais, para quem é convidado e integra júris, para quem toma conhecimento destas situações. Teremos todos de, em conjunto, assumir de forma ativa a responsabilidade de tornar o sistema mais justo e exigente.

Para retirar a conversa “dos corredores”, e não discutirmos apenas aspectos gerais, escrevemos em maio de 2016, em conjunto com alguns colegas, um texto que vem publicado neste número do Boletim. Nesse documento, partimos de um caso concreto e de dados concretos, e descrevemos de forma factual e documentada como decorreu um concurso público recente cujo resultado nos parece inaceitável. Nenhum de nós é já parte interessada neste concurso, que se encontra encerrado e com resultado homologado, nem nada nos move contra a instituição, o candidato vencedor, ou o júri em particular. A motivação que nos levou a escrever este documento foi a de ilustrar com factos que um concurso pode respeitar todas os requisitos legais no que diz respeito à constituição do júri, ao edital, etc. mas ainda assim conduzir a resultados que podem ser classificados, no mínimo, de “estranhos”.

Acreditamos que o Boletim é um espaço adequado para que possa ser levada a cabo uma discussão sobre este assunto, se possível com outros artigos de opinião de diferentes intervenientes em concursos públicos, como sendo candidatos, docentes com experiência enquanto júris de concursos, ou pessoas responsáveis pela elaboração dos editais e cujo contributo possa alimentar uma discussão saudável e produtiva.

RECRUTAMENTO DE DOCENTES DE MATEMÁTICA NO ENSINO SUPERIOR: UM CASO NO ISCAL¹

*Pedro R. S. Antunes, Gilda Ferreira, Rui Ferreira,
José Agapito Ruiz, Hugo Tavares*

Faculdade de Ciências da Universidade de Lisboa
Lisboa, Portugal
e-mail: prantunes@fc.ul.pt
gmferreira@fc.ul.pt
raferreira@fc.ul.pt
jaruiz@ciencias.ulisboa.pt
hrtavares@ciencias.ulisboa.pt

Nuno Costa Dias

Escola Superior Náutica Infante D. Henrique
Paço d'Arcos, Portugal
e-mail: ncdias@meo.pt

Alexandre Rodrigues

Centro de Matemática da Universidade do Porto
Porto, Portugal
e-mail: alexandre.rodrigues@fc.up.pt

João Lita da Silva

Departamento de Matemática
Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa
Caparica, Portugal
e-mail: jfls@fct.unl.pt

Jorge Tiago

Instituto Superior Técnico
Lisboa, Portugal
e-mail: jftiago@math.ist.utl.pt

Com esta carta aberta, endereçada a toda a comunidade do Ensino Superior, e em particular à comunidade dos matemáticos, pretendemos divulgar alguns acontecimentos relativos a um concurso público, aberto em janeiro de 2015, para preenchimento de uma vaga de Professor Adjunto na área de Matemática Aplicada no Instituto Superior de Contabilidade e Administração de Lisboa (ISCAL) do Instituto Politécnico de Lisboa (IPL).

¹Texto escrito a 17 de maio de 2016.

Os autores desta carta fizeram parte dos 38 candidatos ao referido concurso, não tendo nenhum de nós sido vencedor. Entre os candidatos, houve um entendimento bastante generalizado de que os resultados foram muito incorretos, tendo sido apresentadas 14 contestações à lista de ordenação provisória proposta pelo júri. Nenhuma destas contestações foi aceite, pelo que 3 dos autores desta carta interpuseram ainda recurso tutelar junto do presidente do IPL. Novamente sem sucesso. Os resultados do concurso foram homologados em dezembro de 2015.

Em nossa opinião, este caso mostra claramente que um concurso pode satisfazer todas as normas em vigor (nomeadamente as que dizem respeito à constituição do júri e à grelha de avaliação), pode ser disputado por um grande número de candidatos com currículos de qualidade e, ainda assim, não determinar uma seleção minimamente aceitável do candidato vencedor. Por entendermos que este é um caso exemplar, decidimos divulgar os detalhes do processo, para que a comunidade universitária, e esperamos o próprio legislador, possam ter uma ideia mais precisa da forma como alguns concursos ainda decorrem. O nosso objetivo é dar um contributo para aumentar a transparência destes processos e ajudar a repensar as regras que enquadram estes concursos.

Este é um assunto que a todos diz respeito. É sabido que o país investiu muito na formação de novos doutorados e investigadores, e que tem havido muita dificuldade em integrar esta nova geração nas instituições de ensino superior. É um lugar-comum que isto constitui um entrave à modernização e progresso das próprias instituições. Simultaneamente, e principalmente, é também um problema para os mais jovens, alguns com percursos de muito mérito, e que veem o país fechar-lhes as portas a uma vida profissional satisfatória.

O concurso público em questão foi aberto pelo Edital n.º 19/2015 publicado no Diário da República, 2ª série, n.º 6, de 9 de janeiro de 2015, e o Júri teve a seguinte composição:

Presidente do Júri:

- Professora Ana Cristina Arrabaça Miranda Perdigão, Vice-Presidente do Instituto Politécnico de Lisboa e Professora Coordenadora do Instituto Superior de Contabilidade e Administração de Lisboa;

Vogais:

- **Prof. Doutor Jorge Sebastião de Lemos Carvalhão Buescu**, Professor Associado com Agregação da Faculdade de Ciências da Universidade de Lisboa;

- **Prof. Doutor Fernando Manuel Pestana da Costa**, Professor Associado com Agregação da Universidade Aberta;
- **Prof. Doutor Fernando José Malheiro de Magalhães**, Professor Coordenador do Instituto Superior de Contabilidade e Administração do Porto do Instituto Politécnico do Porto;
- **Prof. Doutor Carlos Alberto Martins André**, Professor Associado da Faculdade de Ciências da Universidade de Lisboa;
- **Prof.^a Doutora Sandra Cristina Dias Nunes**, Professora Coordenadora da Escola Superior de Ciências Empresariais do Instituto Politécnico de Setúbal.

Consta do referido Edital a abertura de um concurso para **Professor Adjunto na área de Matemática Aplicada** com as seguintes ponderações, previamente aprovadas pelo júri do concurso, tendo por objetivo a ordenação dos candidatos:

- a) Desempenho Técnico-Científico e Profissional na área (30%).
- b) Avaliação da Componente Pedagógica (50%).
- c) Outras atividades relevantes para a missão da instituição (20%).

Destacamos a distribuição das ponderações relativas ao desempenho Técnico-Científico (alínea a) que se divide da seguinte forma:

- “a1) Atividades de divulgação científica e tecnológica e de valorização económica e social do conhecimento (moderador em palestras, seminários ou congressos, membro de corpo editorial ou de revisão/arbitragem científica de revistas técnicas, serviços à comunidade, etc.) — (15%).
- a2) Coordenação de projetos científicos/desenvolvimento, orientação, arguição e participação em júris de investigação, dissertações, relatórios de estágio e projetos e organização de eventos científicos — (10%).
- a3) Livros, capítulos de livros, comunicações científicas, artísticas e técnicas — (5%).”

Para além destes critérios, é de notar que foi avaliada, na alínea c), a “Participação em órgãos e nas atividades de gestão **da** instituição” (10%), e a “Participação em grupos/comissões de trabalho institucionais, ações de divulgação **da** instituição” (10%).

Alguns aspetos desta grelha são surpreendentes. Em particular, não encontramos nenhuma ponderação para a classificação da produção científica, nomeadamente relativa à autoria de publicações em revistas internacionais da área científica do concurso, sendo este um dos principais critérios para avaliar o mérito científico de um candidato, em qualquer concurso para docente do ensino superior. É bom lembrar que o objetivo da lei é o de seleci-

onar candidatos com um percurso científico de qualidade. Consultando por exemplo o Decreto-Lei n.º 207/2009 de 31 de Agosto, citam-se os Artigos 15.º-A (finalidade dos concursos): “Os concursos (...) destinam-se a averiguar o mérito dos candidatos, da sua capacidade profissional, da sua atividade científica, técnica e de investigação (...)” e o Artigo 30.º-A (deveres do pessoal docente): “d) Manter atualizados e desenvolver os seus conhecimentos culturais e científicos e efetuar trabalhos de investigação, numa procura constante do progresso científico e técnico e da satisfação das necessidades sociais”. Para lá do que diz a lei é isso que, reconhecidamente, melhor serve os interesses dos utentes do ISCAL e também o desenvolvimento da ciência no país. Parece-nos óbvio que esta grelha não cumpre esse objetivo.

Concorreram 38 candidatos ao referido concurso. Da aplicação que o júri fez da grelha de avaliação resultou como primeira classificada uma candidata doutorada em **Ciências da Educação** que já lecionava na Instituição responsável pelo concurso há cerca de quatro anos, com a categoria de Professora Adjunta Convidada. A referida candidata não é autora de **nenhum** artigo científico em revista internacional na área do concurso, não participou em nenhum projeto de investigação na área da Matemática ou Matemática Aplicada (ou tão pouco em algum projeto financiado pela FCT), não é autora de nenhum livro, nem coordenou projetos ou realizou atividade editorial **nas áreas da Matemática ou Matemática Aplicada**.

Na Tabela 1 são apresentados os dados que constam nas bases de dados Web of Science e MathSciNet relativos aos trabalhos científicos da candidata vencedora do concurso, e de uma amostra dos restantes candidatos, bem como a respetiva posição na lista de ordenação final do concurso. O artigo da candidata vencedora que é mencionado na tabela é uma publicação em ata de conferência na área de Ciências da Educação.

Para além do número de artigos e de citações destes concorrentes, gostaríamos ainda de salientar que uma pesquisa na base de dados Web of Science, revela igualmente que alguns destes publicaram em revistas de prestígio como *Advances in Mathematics*, *Communications on Pure and Applied Mathematics*, *SIAM Journal on Scientific Computing*, *Annals of Pure and Applied Logic*, *SIAM Journal on Control and Optimization*, *Journal de Mathématiques Pures et Appliquées*, *Transactions of the AMS*, *Journal of the EMS*, *Proceedings of the LMS*, *Communications in Mathematical Physics*, entre muitas outras.

Poderíamos ser levados a pensar que a candidata venceu o concurso porque as suas pontuações na componente pedagógica e em outras atividades compensaram pontuações muito fracas na componente científica. Mas isso

Posição na lista de ordenação final	N.º artigos Web of Science	N.º artigos MathSciNet	N.º citações Web of Science
1	1	0	0
2	45	34	377
5	9	9	107
7	12	16	11
8	20	17	108
10	11	11	136
11	13	18	4
18	9	8	22
26	26	32	437

Tabela 1: Posição dos candidatos na lista de ordenação final e alguns indicadores de desempenho científico.

não foi verdade. Na Tabela 2 são apresentadas as pontuações atribuídas na componente científica pelos vários membros do júri (com exceção da presidente do júri, que não atribuiu classificações) aos candidatos que constam da tabela anterior (as pontuações foram normalizadas de acordo com o procedimento definido no edital do concurso, tendo por referência uma pontuação total máxima de 100 pontos).

Na coluna “Média Componente Científica” da Tabela 2 é apresentada a pontuação final efetivamente atribuída a cada candidato na *componente científica* (alínea (a) da grelha de avaliação). A esta pontuação foram depois somadas as pontuações na *componente pedagógica* (alínea (b)) e em *outras atividades* (alínea (c)) para obter a pontuação final que determinou a lista de ordenação proposta pelo júri. A título de exemplo, a candidata vencedora teve uma pontuação final de 92,6/100 onde 22,8/100 dizem respeito à avaliação da componente científica. Já o segundo classificado teve 88,9/100 de pontuação final, com 26,6/100 na componente científica, etc.

Fazemos a ressalva que não é o nosso entendimento que o processo de classificação dos candidatos deva ser uma tradução literal da sua performance nos indicadores de produção e impacto das publicações. Há certamente muitos outros aspetos a considerar, como a área de trabalho, a idade dos candidatos, a afinidade com o trabalho dos docentes da instituição que lançou o concurso, para mencionar apenas alguns. Muitos destes aspetos são subjetivos e a sua ponderação é naturalmente do âmbito de discricionariedade do júri. No entanto, no caso desta tabela classificativa, parece-nos

Posição	Aval. 1	Aval. 2	Aval. 3	Aval. 4	Aval. 5	Média Componente Científica
1	18,0	19,4	28,8	20,3	27,4	22,8
2	30,0	24,2	31,2	21,5	26,0	26,6
5	27,0	21,0	31,9	16,5	41,1	27,5
7	23,0	22,6	27,4	17,7	26,7	23,5
8	18,0	33,9	28,9	20,3	31,5	26,5
10	30,0	41,9	27,4	24,1	30,8	30,8
11	11,0	22,6	34,2	15,2	15,8	19,8
18	7,0	17,7	21,9	11,4	20,5	15,7
26	17,0	21,0	26,4	17,7	16,4	19,7

Tabela 2: Pontuações atribuídas na componente científica por cada avaliador.

que foram ultrapassados todos os limites razoáveis: afinal, um curriculum científico inexistente (para a área do concurso), é sempre inexistente.

Estes argumentos foram apresentados ao júri nas 14 reclamações à lista provisória de ordenação final. Também se contestou o facto de a área científica da candidata (doutorada em Ciências da Educação) ter sido considerada pelo júri como afim da Matemática Aplicada.

O júri do concurso decidiu por unanimidade manter a ordenação dos candidatos e a pontuação atribuída à candidata vencedora. Em resposta à reclamação que argumentava que a tese de doutoramento da candidata não era da área do concurso, o júri declarou que “decidiu admitir a concurso todos os detentores do grau de doutor que, na sua tese de doutoramento e na sua investigação, tenham desenvolvido ou feito uso relevante de métodos matemáticos ou estatísticos, independentemente da área científica formal em que a referida tese foi originalmente classificada”. Por entendermos que era importante, fomos consultar a tese de doutoramento da candidata vencedora. Constatamos facilmente que no seu trabalho foram usados apenas métodos matemáticos e estatísticos de nível elementar, igualmente presentes em muitos trabalhos de outras áreas científicas como a Sociologia, História, Economia, Medicina ou Relações Internacionais. Na maioria das teses de doutoramento nestas áreas faz-se “uso relevante” de métodos matemáticos e estatísticos, porquanto são calculadas médias, quartis e percentagens, apresentadas tabelas de dados numéricos e gráficos de séries temporais, e efetuadas análises de variâncias, com o propósito de estudar o assunto da

tese. Exatamente aquilo, e apenas aquilo, que é “feito uso” na tese da candidata. Seguindo este raciocínio chegamos à conclusão que quase todos os doutorados em Ciências Sociais e Humanas (e por maioria de razão em todas as outras áreas do saber) são doutorados em área afim da Matemática Aplicada, o que é claramente um absurdo.

Por estas razões, na sequência da decisão do júri, três dos subscritores desta carta apresentaram ainda recurso tutelar ao Presidente do Instituto Politécnico de Lisboa. Foram apontadas, novamente, a falta de produção científica e a incorreta avaliação da área de trabalho da candidata. A 18 de dezembro de 2015, a resposta do presidente do IPL, sustentada por pareceres do departamento jurídico, alegou a soberania do júri e homologou os resultados.

E é assim que em 2015, uma candidata doutorada em Ciências da Educação, sem um único artigo científico publicado, nem atividade científica conhecida na área da Matemática e Matemática Aplicada, vence um concurso para Professor Adjunto de Matemática, um concurso com 38 candidatos, alguns com prémios Gulbenkian, vários Investigadores FCT, alguns com currículos longos (tanto científica como pedagogicamente), outros extremamente promissores e muitos, muitos matemáticos que fazem investigação nestas áreas. Estes resultados foram possíveis não obstante um grande número de candidatos ter apresentado reclamações, e tentado tudo o que a lei permite para evitar este desenlace.

Em conclusão, não nos parece que o interesse público tenha sido salvaguardado neste concurso, começando com os critérios do edital, e terminando na escolha da candidata vencedora. Em particular, tudo isto nos parece ir contra os objetivos de todo o investimento feito pelo país para desenvolver a Ciência e o Ensino Superior.

Cabe a todos nós, fazendo ou não parte do sistema universitário, contribuir para o melhor ensino superior público possível. Assim, e por último, apelamos a todos que ajudem a alertar para outras situações do mesmo género, tornando o sistema mais transparente, e aos docentes, em particular, que não aceitem colaborar em concursos com estes critérios, ou que ajam ativamente numa tentativa de os alterar.

Problemas

Editor:
Jorge Nuno Silva

NOTAS SOBRE O PROBLEMA ANTERIOR E NÃO JULGUE UM NÚMERO PELO ASPECTO!

Jorge Nuno Silva

Os leitores são convidados a enviar, para eventual publicação, soluções, comentários, propostas de problemas, etc. Essa correspondência deve ser enviada para a SPM, ao cuidado do editor desta secção. Há livros para sortear entre as soluções recebidas em cada número.

Relembremos o problema do número anterior.

Estética e crítica

No último encontro anual *MathsJam*, que se realiza habitualmente perto de Manchester, no Reino Unido, tomámos conhecimento de um *blog* focado em *puzzles* matemáticos. O seu nome, *Puzzle Critic*, diz ao que se propõe: a promoção de elegantes desafios matemáticos. Percorrer os seus arquivos é um prazer, damos com questões muito bem concebidas, de enunciado claro e conteúdo substancial.

Deste *site* vem nossa selecção para hoje, que se resume a três questões sobre números.

1. Seja n um número inteiro positivo qualquer. De cada um dos números $n+1, \dots, 2n$ seleccione o maior divisor ímpar. Prove que a soma desses divisores é n^2 .
2. Seja n um número inteiro positivo qualquer. Mostre que n tem um múltiplo cuja soma dos dígitos é ímpar.
3. As dízimas infinitas $0.abab\dots$ e $0.abcabc\dots$ verificam

$$0.abab\dots + 0.abcabc\dots = \frac{33}{37}.$$

Determine os dígitos a , b , e c .

1. Note que se dois números diferentes tiverem o mesmo maior divisor ímpar, então um deles é pelo menos o dobro do outro. Assim, os maiores divisores ímpares dos n números $n+1, \dots, 2n$ são os primeiros n números ímpares, e é sabido que a respectiva soma é n^2 .

Como o autor do *blog* refere, este problema reúne várias boas qualidades: é bonito, o seu enunciado é simples, a resolução também é simples (embora difícil de encontrar), o resultado é surpreendente e parece uma amostra genuína de boa matemática.

2. Dado um inteiro k , considere os números $A = 10^n - k$ e $B = 10^{n+1} - k$. Para n suficientemente grande, A e B são múltiplos de k e diferem um do outro por B ter mais um dígito 9 do que A . Assim, as paridades das somas dos dígitos são distintas, pelo que uma delas tem de ser par.
3. Notando que

$$0.010101\dots = \frac{1}{99} \quad \text{e} \quad 0.001001001\dots = \frac{1}{999}$$

a equação dada pode escrever-se

$$\frac{10a + b}{99} + \frac{100a + 10b + c}{999} = \frac{33}{97}$$

donde se deduz $2210a + 221b = 9801 - 11c$. Notando que o lado direito deve ser múltiplo de 221 concluímos que $c = 7$. Um pequeno cálculo dá-nos agora $a = b = 4$.

Não julgue um número pelo aspecto!

Os programas doutorais contêm, em algumas universidades, exames de acesso, constituídos por problemas que, ao longo dos anos, vão criando belas coleções, que os estudantes usam para se prepararem. Isto sucede com UC Berkeley, de cujos *Preliminary examination (2016)*, retiramos a nossa proposta de hoje. Este era o primeiro de nove problemas propostos, de que os candidatos deviam resolver seis em três horas.

Mostre que o número

$$\int_4^9 \sqrt{-6 + 5\sqrt{-6 + 5\sqrt{-6 + 5\sqrt{-6 + 5\sqrt{-6 + 5\sqrt{x}}}}} dx$$

é racional.

Notícias

SPM vence Prémio Gulbenkian – Conhecimento

Instituído neste ano, o Prémio Gulbenkian – Conhecimento, subordinado ao tema “Promoção do Sucesso Escolar”, foi atribuído à Sociedade Portuguesa de Matemática pela criação e organização das Olimpíadas Portuguesas de Matemática, que contou em 2017 com a sua 35.^a edição.

De acordo com o júri do Prémio Gulbenkian, na base desta distinção está a promoção de uma “iniciativa educativa de referência e de grande impacto nacional, que há mais de três décadas promove o gosto por esta disciplina”. Os Prémios Gulbenkian foram entregues numa cerimónia dirigida pelo Presidente da República, Professor Marcelo Rebelo de Sousa, na Fundação Calouste Gulbenkian, no dia 20 de Julho.

Portugal alcança o bronze nas Olimpíadas Internacionais de Matemática

A 58.^a edição das Olimpíadas Internacionais de Matemática, que decorreu até dia 22 de Julho, no Rio de Janeiro, saldou-se em duas medalhas de bronze e duas menções honrosas para a equipa portuguesa.

Kevin Pucci, 15 anos, residente em Chaves, alcançou a medalha de bronze, com 18 pontos, ficando a um ponto apenas do medalhado com a prata. Henrique Navas (17 anos, Lisboa) conquistou também uma medalha de bronze, enquanto que Manuel Cabral (18 anos, Lisboa) e Pedro Fernandes (16 anos, Pombal) receberam uma menção honrosa.

A equipa contou ainda com a participação de Duarte Nascimento (18 anos, Amadora) e Matilde Silva (16 anos, Coimbra).

VII Olimpíadas de Matemática da CPLP

A sétima edição das Olimpíadas de Matemática da Comunidade de Países de Língua Portuguesa (OMCPLP) decorreu na Faculdade de Ciências da Universidade do Porto, de 23 a 30 de Julho, numa organização conjunta Departamento de Matemática da Universidade do Porto e da Sociedade Portuguesa de Matemática. O principal objectivo das OMCPLP é unir os países lusófonos através da matemática, incentivar o desenvolvimento da disciplina em todos eles e aprofundar a cooperação nesta área. O evento contou com a

participação de jovens oriundos dos oito países de expressão portuguesa, Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal, São Tomé e Príncipe e Timor Leste.

Rodrigo Andrade, do 10.º ano da Escola Secundária Senhora da Hora, em Matosinhos, e Rui Wang, do 9.º ano da Escola 2, 3 Dr. Joaquim de Magalhães, em Faro, conquistaram medalhas de ouro. Pedro Dias, do 10.º ano da Escola Alemã, em Lisboa, e David Nassauer, do 9.º ano na Escola Secundária D. Filipa de Lencastre, em Lisboa, alcançaram medalhas de prata.

O pódio foi dividido com os brasileiros Gabriel Paiva e Pedro Cabral, no ouro, e com Eduardo Oliveira e André Koga, na prata, e ainda com o são-tomense Adolfo Boa Morte Luiz, também na prata.

As restantes medalhas, de bronze, foram distribuídas por Jordy Jorge Silva, Pablo Daloyce Veiga e Tiago Rocha Pires, de Cabo Verde, por Etianeth Creusa Miguel e Tanaco da Conceição Cariri, de Angola, por Yunus Nazir Amuji e Delson Augusto, de Moçambique, e por Cemilton Boa Morte, de São Tomé e Príncipe.

Feira da Matemática regressa em Novembro

O Museu Nacional de História Natural e da Ciência, em Lisboa, receberá a IV Feira da Matemática nos dias 10 e 11 de Novembro. Tal como nos anos anteriores, o primeiro dia é exclusivo para as escolas e o segundo está aberto ao público em geral.

Exposições, sessões práticas, jogos e desafios, demonstrações, Circo Matemático, palestras e concursos são apenas algumas das muitas actividades da Feira.

As pré-inscrição para as escolas deverão ser efetuadas através do endereço de correio electrónico geral@museus.ulisboa.pt.

“Compreender os números na Matemática Escolar”, de Hung-Hsi Wu, na loja SPM

Traduzida e adaptada com o patrocínio da Sociedade Portuguesa de Matemática, a obra “Compreender os números na Matemática Escolar”, do investigador Hung-Hsi Wu, apresenta-se como um manual de referência para professores dos Ensinos Básico e Secundário.

Através de uma linguagem matemática precisa, este livro abarca todos os tópicos relativos a números e operações leccionados até ao 7.º ano de escolaridade. “Compreender os Números na Matemática Escolar” está disponível para compra na Loja da SPM online, em <https://www.spm.pt/store/list/novidades>.