

# Um método não convencional para a diagonalização de matrizes $2 \times 2$

Adilson Francisco da Silva 

Josimar da Silva Rocha 

Thiago Pinguello de Andrade 

## Resumo

Neste trabalho, apresentamos o conceito de matrizes diagonalizáveis e encontramos condições necessárias e suficientes para que uma matriz em  $\mathcal{M}_2(\mathbb{R})$  seja diagonalizável, bem como explicitamos a matriz diagonal correspondente e a matriz conjugadora, sem que para isso sejam utilizados conceitos de espaços vetoriais. Como aplicações, mostramos como calcular potências de matrizes diagonalizáveis em  $\mathcal{M}_2(\mathbb{R})$  e também um método para a resolução de um caso particular de equações de recorrências lineares de segunda ordem.

**Palavras-chave:** matrizes; diagonalização; recorrências lineares.

## Abstract

In this work, we present the concept of diagonalizable matrices and find necessary and sufficient conditions for a matrix in  $\mathcal{M}_2(\mathbb{R})$  to be diagonalizable, as well as explaining the corresponding diagonal matrix and the conjugating matrix, without using vector spaces concepts. As applications, we show how to calculate powers of diagonalizable matrices in  $\mathcal{M}_2(\mathbb{R})$  and we also show a method for solving a particular case of second order linear recurrence equations.

**Keywords:** matrices; diagonalization; linear recurrences.

## 1. Introdução

O método geralmente utilizado para diagonalização de matrizes requer conhecimento sobre espaço vetorial para a sua aplicação (ver, por exemplo, as referências [1], [2] e [4]), o que restringe sua utilização àqueles que possuem algum conhecimento da matemática do Ensino Superior. Assim, como alternativa, mostraremos uma técnica não usual de diagonalização de matrizes em  $\mathcal{M}_2(\mathbb{R})$  em que são utilizados apenas conhecimentos de matemática elementar.

Neste artigo, apresentaremos o conceito de matrizes semelhantes de  $\mathcal{M}_2(\mathbb{R})$  e, como aplicação, classificaremos as matrizes  $2 \times 2$  diagonalizáveis sem utilizar o conceito de autovetores. Apresentaremos o cálculo de potência de matrizes, e resolveremos recorrências lineares de segunda ordem.

## 2. Matrizes semelhantes

Nesta seção apresentaremos o conceito de semelhança de matrizes e suas principais propriedades. Veremos que matrizes semelhantes formam uma relação de equivalência.

**Definição 1.** Seja  $A$  uma matriz  $2 \times 2$ . O polinômio

$$p(\lambda) = \det(A - \lambda I),$$

em que  $I$  denota a matriz identidade, é chamado de **polinômio característico** de  $A$ . Além disso, as raízes do polinômio característico de  $A$  são chamadas de **autovalores** de  $A$ .

**Definição 2.** Sejam  $A$  e  $B$ , matrizes de ordem  $2 \times 2$ . Dizemos que  $A$  é **semelhante** a  $B$ , e denotaremos  $A \sim B$ , se existir  $M \in \mathcal{M}_2(\mathbb{R})$  invertível, tal que

$$B = M^{-1}AM.$$

**Exemplo 1.** As matrizes  $A = \begin{bmatrix} -22 & -28 \\ 18 & 23 \end{bmatrix}$  e  $B = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$  são semelhantes, pois, para  $M =$

$\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$  temos que

$$\begin{aligned} M^{-1}BM &= \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \\ &= -\frac{1}{2} \begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \\ &= \begin{bmatrix} -22 & -28 \\ 18 & 23 \end{bmatrix} = A \end{aligned}$$

**Proposição 1.** Sejam  $A$ ,  $B$  e  $C$  matrizes em  $\mathcal{M}_2(\mathbb{R})$ . A relação de semelhança de matrizes é uma relação de equivalência, ou seja, satisfaz as seguintes propriedades:

1.  $A \sim A$  (propriedade reflexiva);
2. Se  $A \sim B$ , então  $B \sim A$  (propriedade simétrica);
3. Se  $A \sim B$  e  $B \sim C$ , então  $A \sim C$  (propriedade transitiva).

*Demonstração.* 1. De fato, como  $A = I^{-1}AI$ , temos pela definição de matriz semelhante que  $A \sim A$ .

2. Sendo  $A \sim B$  temos, pela definição, que existe  $M$  invertível tal que  $B = M^{-1}AM$ . Assim, denotando  $N = M^{-1}$ , temos que  $N$  é invertível e multiplicando a igualdade  $B = M^{-1}AM$ , por  $N^{-1}$  à esquerda e por  $N$  à direita, obtemos

$$N^{-1}BN = N^{-1}M^{-1}AMN = MM^{-1}AMM^{-1} = A,$$

ou seja,  $B \sim A$ .

3. Sendo  $A \sim B$  e  $B \sim C$  existem matrizes invertíveis  $M$  e  $N$ , tais que  $B = M^{-1}AM$  e  $C = N^{-1}BN$ . Logo, considerando  $P = MN$  temos, por propriedades de multiplicação de matrizes e matrizes inversas, que  $P$  é invertível e  $P^{-1} = N^{-1}M^{-1}$ . Assim,

$$P^{-1}AP = N^{-1}M^{-1}AMN = N^{-1}BN = C,$$

ou seja  $A \sim C$ .

**Proposição 2.** *Sejam A e B matrizes semelhantes em  $\mathcal{M}_2(\mathbb{R})$ . Então:*

(i)  $\det A = \det B$ ;

(ii) A e B possuem os mesmos autovalores.

*Demonstração.*

(i) Como  $A \sim B$ , existe uma matriz invertível M tal que  $B = M^{-1}AM$ . Assim, utilizando as propriedades do determinante, temos

$$\det B = \det (M^{-1}AM) = (\det M^{-1})(\det A)(\det M) = \left(\frac{1}{\det M}\right) (\det A)(\det M) = \det A.$$

(ii) Como  $A \sim B$ , existe  $M \in \mathcal{M}_2(\mathbb{R})$  invertível tal que  $B = M^{-1}AM$ . Assim, para  $\lambda \in \mathbb{R}$ , temos

$$\begin{aligned} B - \lambda I &= M^{-1}AM - \lambda I \\ &= M^{-1}AM - \lambda M^{-1}M \\ &= M^{-1}(AM - \lambda M) \\ &= M^{-1}(AM - \lambda IM) \\ &= M^{-1}(A - \lambda I)M, \end{aligned}$$

isto é,  $(A - \lambda I) \sim (B - \lambda I)$ . De sorte que, por (i), temos  $\det (A - \lambda I) = \det (B - \lambda I)$ . Portanto, pela definição de autovalor, A e B possuem os mesmos autovalores. □

**Definição 3.** Seja  $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$  uma matriz quadrada de ordem  $2 \times 2$ . O **traço** de A, denotado por  $\text{tr}(A)$ , é definido por

$$\text{tr}(A) = a_{11} + a_{22}.$$

**Lema 1.** *Para quaisquer matrizes A e B em  $M_2(\mathbb{R})$ , tem-se  $\text{tr}(AB) = \text{tr}(BA)$ .*

*Demonstração.* Para quaisquer matrizes  $A = (a_{ij})$  e  $B = (b_{ij})$  em  $M_2(\mathbb{R})$ , temos que

$$AB = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

e

$$BA = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix}.$$

Assim,

$$\text{tr}(AB) = a_{11}b_{11} + a_{12}b_{21} + a_{21}b_{12} + a_{22}b_{22} = b_{11}a_{11} + b_{12}a_{21} + b_{21}a_{12} + b_{22}a_{22} = \text{tr}(BA).$$

□

**Teorema 1.** Se  $A \sim B$ , então  $\text{tr}(A) = \text{tr}(B)$ .

*Demonstração.* Como  $A \sim B$ , existe uma matriz invertível  $M$ , tal que  $B = M^{-1}AM$ .

Portanto, pelo Lema 1,

$$\text{tr}(B) = \text{tr}(M^{-1} \cdot (AM)) = \text{tr}((AM) \cdot M^{-1}) = \text{tr}(A \cdot (MM^{-1})) = \text{tr}(A).$$

□

### 3. Diagonalização de matrizes

Nesta seção, apresentaremos o conceito de matrizes diagonalizáveis e demonstraremos as principais propriedades das matrizes diagonalizáveis em  $\mathcal{M}_2(\mathbb{R})$ . Apresentaremos também um método não convencional para diagonalização de matrizes em  $\mathcal{M}_2(\mathbb{R})$  que não utiliza o conceito de espaço vetorial.

**Definição 4.** Dizemos que uma matriz  $A$  de  $\mathcal{M}_2(\mathbb{R})$  é diagonalizável se existir uma matriz diagonal  $D \in \mathcal{M}_2(\mathbb{R})$  tal que  $A \sim D$ , ou seja, se existir uma matriz diagonal  $D$  tal que

$$D = M^{-1}AM$$

para alguma matriz invertível  $M \in \mathcal{M}_2(\mathbb{R})$ .

**Exemplo 2.** Toda matriz diagonal é uma matriz diagonalizável. De fato, se  $A$  é uma matriz diagonal, como  $A = I^{-1}AI$ , temos que  $A$  é diagonalizável.

**Proposição 3.** Se  $A$  é uma matriz em  $\mathcal{M}_2(\mathbb{R})$  diagonalizável e  $D$  é uma matriz diagonal semelhante a  $A$ , então, os elementos da diagonal principal de  $D$  são os autovalores de  $A$ .

*Demonstração.* Sejam  $D$  uma matriz diagonal semelhante a  $A$ , e  $d_{11}$  e  $d_{22}$  os elementos da diagonal principal de  $D$ . Como  $D$  é uma matriz diagonal, então o polinômio característico de  $D$  é

$$p(\lambda) = \det(D - \lambda I) = (d_{11} - \lambda)(d_{22} - \lambda).$$

Assim,  $p(d_{11}) = 0$  e  $p(d_{22}) = 0$ , ou seja,  $d_{11}$  e  $d_{22}$  são autovalores de  $D$ . Logo, como  $d_{11}$  e  $d_{22}$  são autovalores de  $D$ , e  $D$  é semelhante a  $A$ , segue do item (ii) da Proposição 2, que  $d_{11}$  e  $d_{22}$  são os autovalores de  $A$ . □

**Proposição 4.** Seja  $(A, D, M)$  uma terna de matrizes de  $\mathcal{M}_2(\mathbb{R})$  tais que  $M$  é invertível e  $D$  é diagonal. A terna  $(A, D, M)$  satisfaz a equação matricial  $D = M^{-1}AM$  se, e somente se, exatamente uma das condições abaixo ocorre.

(i)  $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ ,  $D = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$  e  $M$  com  $\det(M) \neq 0$ ;

(ii)  $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ ,  $M = \begin{bmatrix} x & 0 \\ 0 & w \end{bmatrix}$  e  $D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ , com  $a \neq d$ ,  $x \neq 0$  e  $w \neq 0$ ;

(iii)  $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ ,  $M = \begin{bmatrix} 0 & y \\ z & 0 \end{bmatrix}$  e  $D = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix}$ , com  $y \neq 0$ ,  $z \neq 0$  e  $a \neq d$ ;

$$(iv) A = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}, M = \begin{bmatrix} 0 & y \\ z & \frac{cy}{a-d} \end{bmatrix} \text{ e } D = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix}, \text{ com } c \neq 0, z \neq 0, y \neq 0 \text{ e } a \neq d;$$

$$(v) A = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}, M = \begin{bmatrix} x & 0 \\ \frac{cx}{a-d} & w \end{bmatrix} \text{ e } D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, \text{ com } c \neq 0, x \neq 0, w \neq 0 \text{ e } a \neq d;$$

$$(vi) A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, M = \begin{bmatrix} \frac{bz}{d-a} & y \\ z & 0 \end{bmatrix} \text{ e } D = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix}, \text{ com } b \neq 0, z \neq 0, y \neq 0 \text{ e } a \neq d;$$

$$(vii) A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, M = \begin{bmatrix} x & \frac{bw}{d-a} \\ 0 & w \end{bmatrix} \text{ e } D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, \text{ com } b \neq 0, x \neq 0, w \neq 0 \text{ e } a \neq d;$$

$$(viii) A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, M = \begin{bmatrix} x & y \\ \frac{(\lambda_1 - a)x}{b} & \frac{(\lambda_2 - a)y}{b} \end{bmatrix} \text{ e } D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \text{ com } b \neq 0, c \neq 0, x \neq 0 \text{ e } y \neq 0, \text{ desde que } \lambda_1 \text{ e } \lambda_2 \text{ sejam soluções reais da equação } (\lambda - a)(\lambda - d) = bc \text{ com } \lambda_1 \neq \lambda_2.$$

*Demonstração.* Sejam

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, M = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \text{ e } D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

satisfazendo  $D = M^{-1}AM$ . Assim,  $MD = AM$ , ou seja,

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x & y \\ z & w \end{bmatrix}.$$

Logo

$$\begin{bmatrix} x\lambda_1 & y\lambda_2 \\ z\lambda_1 & w\lambda_2 \end{bmatrix} = \begin{bmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{bmatrix}$$

donde obtemos o seguinte sistema de equações:

$$\begin{cases} 0 = (a - \lambda_1)x + bz \\ 0 = (a - \lambda_2)y + bw \\ 0 = cx + (d - \lambda_1)z \\ 0 = cy + (d - \lambda_2)w. \end{cases} \quad (1)$$

**Caso 1.** Se  $b = 0$  e  $c = 0$ , então o sistema (1) torna-se

$$\begin{cases} 0 = (a - \lambda_1)x & (2) \\ 0 = (a - \lambda_2)y & (3) \\ 0 = (d - \lambda_1)z & (4) \\ 0 = (d - \lambda_2)w. & (5) \end{cases}$$

**Subcaso 1.1.** Se  $a = d$ , então  $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = aI$ . Sendo  $A = aI$ , para toda matriz  $M$  inversível, tem-se que  $D = M^{-1}AM = M^{-1}(aI)M = a(M^{-1}M) = aI = A$ . Com isso, obtemos (i), que é o caso em que  $A = aI$  é uma matriz escalar e os autovalores de  $A$  são  $\lambda_1 = \lambda_2 = a$ .

**Subcaso 1.2.** Se  $a \neq d$  e  $x \neq 0$ , então por (2),  $\lambda_1 = a$ . Com isso, por (4) temos  $z = 0$ . Como  $M$  é invertível temos que  $\det M = xw - zy \neq 0$ , donde  $w \neq 0$  e por (5),  $\lambda_2 = d$ . Além disso, como  $\lambda_1 \neq \lambda_2$ , temos por (3) que  $y = 0$ . Reciprocamente, se

$$A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, M = \begin{bmatrix} x & 0 \\ 0 & w \end{bmatrix} \text{ e } D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix},$$

com  $x \neq 0$ ,  $w \neq 0$  e  $a \neq d$ , então  $MD = AM$ . Desse modo, obtemos (ii), que é o caso em que  $A$  é a matriz diagonal e possui dois autovalores reais distintos, a saber,  $\lambda_1 = a$  e  $\lambda_2 = d$ .

**Subcaso 1.3.** Se  $a \neq d$  e  $x = 0$ , então, por  $M$  ser invertível, temos que  $\det M \neq 0$  e  $\det M = yz$ , ou seja,  $y \neq 0$  e  $z \neq 0$ . Assim, de (3) e (4) temos que  $\lambda_2 = a$  e  $\lambda_1 = d$ . Além disso, por (5) temos que  $w = 0$ . Reciprocamente, se

$$A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, M = \begin{bmatrix} 0 & y \\ z & 0 \end{bmatrix} \text{ e } D = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix},$$

com  $y \neq 0$ ,  $z \neq 0$  e  $d \neq a$ , então  $MD = AM$ . Com isso, obtemos (iii), que é o caso em que  $A$  é a matriz diagonal com dois autovalores distintos, a saber,  $\lambda_1 = d$  e  $\lambda_2 = a$ .

**Caso 2.** Se  $b = 0$  e  $c \neq 0$ , então o sistema (1) torna-se

$$\begin{cases} 0 = (a - \lambda_1)x & (6) \\ 0 = (a - \lambda_2)y & (7) \\ 0 = cx + (d - \lambda_1)z & (8) \\ 0 = cy + (d - \lambda_2)w. & (9) \end{cases}$$

**Subcaso 2.1.** Se  $x = 0$ , então, por  $M$  ser invertível, temos que  $\det M = yz$  com  $y \neq 0$  e  $z \neq 0$ . Assim, por (7) e (8) temos que  $\lambda_2 = a$  e  $\lambda_1 = d$ . Com isso, temos por (9) que  $(d - \lambda_2)w \neq 0$ . Assim,  $w \neq 0$  e  $\lambda_2 \neq d$ , ou seja  $a \neq d$ , onde podemos escrever

$$w = \frac{cy}{\lambda_2 - d} = \frac{cy}{a - d}.$$

Reciprocamente, se

$$A = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}, M = \begin{bmatrix} 0 & y \\ z & \frac{cy}{a-d} \end{bmatrix} \text{ e } D = \begin{bmatrix} d & 0 \\ 0 & a \end{bmatrix},$$

com  $c \neq 0$ ,  $z \neq 0$ ,  $y \neq 0$  e  $a \neq d$ , então,  $MD = AM$ . Com isso, obtemos (iv), que é o caso em que  $A$  é uma matriz triangular inferior e possui dois autovalores distintos, a saber,  $\lambda_1 = d$  e  $\lambda_2 = a$ .

**Subcaso 2.2.** Se  $x \neq 0$ , então, por (6),  $\lambda_1 = a$ . Como  $c \neq 0$ , por (8), temos  $(d - \lambda_1)z \neq 0$ , ou seja,  $z \neq 0$  e  $a \neq d$ . Pelo Teorema 1 temos que  $\text{tr}(A) = \text{tr}(D)$ , ou seja  $a + d = \lambda_1 + \lambda_2$ , donde  $\lambda_2 = d$ . Assim, por (9),  $y = 0$ , e, como  $M$  é invertível, temos que  $\det M = xw - yz \neq 0$ , com  $w \neq 0$ . Reescrevendo (8), temos

$$z = \frac{cx}{\lambda_1 - d} = \frac{cx}{a - d}.$$

Reciprocamente, se

$$A = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}, M = \begin{bmatrix} x & 0 \\ \frac{cx}{a-d} & w \end{bmatrix} \text{ e } D = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix},$$

com  $c \neq 0$ ,  $x \neq 0$ ,  $w \neq 0$  e  $a \neq d$ , então,  $MD = AM$ . Com isso, obtemos (v), que é o caso em que  $A$  é uma matriz triangular superior e possui dois autovalores reais distintos, a saber,  $\lambda_1 = a$  e  $\lambda_2 = d$ .

**Caso 3.** Se  $b \neq 0$  e  $c = 0$ , então, obtemos (vi) e (vii) aplicando transposição nas relações obtidas nos casos (iv) e (v), respectivamente.

**Caso 4.** Se  $b \neq 0$  e  $c \neq 0$ , então, por (1) obtemos

$$\begin{cases} x = \frac{(\lambda_1 - d)z}{c} & (10) \end{cases}$$

$$\begin{cases} y = \frac{(\lambda_2 - d)w}{c} & (11) \end{cases}$$

$$\begin{cases} w = \frac{(\lambda_2 - a)y}{b} & (12) \end{cases}$$

$$\begin{cases} z = \frac{(\lambda_1 - a)x}{b}, & (13) \end{cases}$$

o que implica que

$$xz = \frac{(\lambda_1 - a)(\lambda_1 - d)xz}{bc}$$

e

$$wy = \frac{(\lambda_2 - a)(\lambda_2 - d)wy}{bc}.$$

Como  $M$  é invertível, temos que  $xw - yz \neq 0$ , ou seja,  $xw \neq 0$  ou  $yz \neq 0$ . Se  $xw = 0$ , por (11) e (13) teríamos  $yz = 0$  e  $M$  não seria invertível. Reciprocamente, se  $yz = 0$ , por (10) e (12) teríamos  $xw = 0$  e  $M$  não seria invertível. Assim, temos que  $xw \neq 0$  e  $yz \neq 0$ , ou seja,  $x \neq 0$ ,  $y \neq 0$ ,  $z \neq 0$  e  $w \neq 0$ . Logo,

$$\begin{cases} (\lambda_1 - a)(\lambda_1 - d) = bc \\ (\lambda_2 - a)(\lambda_2 - d) = bc, \end{cases} \quad (14)$$

ou seja,  $\lambda_1$  e  $\lambda_2$  são raízes da equação

$$(\lambda - a)(\lambda - d) = bc.$$

Além disso, como  $M$  é invertível, por (10), (11), (12) e (13), temos

$$xw - yz = \frac{(\lambda_1 - d)(\lambda_2 - a)yz}{bc} - \frac{(\lambda_2 - d)(\lambda_1 - a)xw}{bc}.$$

Se  $\lambda_1 = \lambda_2$ , temos

$$xw - yz = \frac{(\lambda_2 - d)(\lambda_2 - a)}{bc}(yz - xw).$$

Como  $xw - yz \neq 0$

$$(\lambda_2 - d)(\lambda_2 - a) = -bc,$$

o que contradiz (14). Logo,  $\lambda_1 \neq \lambda_2$ .

Reciprocamente, se

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, M = \begin{bmatrix} x & y \\ \frac{(\lambda_1 - a)x}{b} & \frac{(\lambda_2 - a)y}{b} \end{bmatrix} \text{ e } D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

com  $b \neq 0$ ,  $c \neq 0$ ,  $x \neq 0$  e  $y \neq 0$ , desde que  $\lambda_1$  e  $\lambda_2$  sejam soluções da equação  $(\lambda - a)(\lambda - d) = bc$  com  $\lambda_1 \neq \lambda_2$ , então,  $MD = AM$ . Com isso, obtemos (viii), que é o caso em que a matriz  $A$  não possui entradas nulas na diagonal secundária e possui dois autovalores reais distintos definidos pela equação característica.  $\square$

**Corolário 1.** Uma matriz  $A$  de  $\mathcal{M}_2(\mathbb{R})$  é diagonalizável se, e somente se, uma das seguintes situações é verdadeira:

- (i)  $A$  possui dois autovalores reais distintos;
- (ii)  $A$  possui apenas um autovalor real e  $A$  é uma matriz escalar.

*Demonstração.* Se  $A$  é diagonalizável, então, pela Proposição 4,  $A$  possui dois autovalores reais iguais se, e somente se,  $A$  é a matriz escalar. Portanto, se  $A$  possui dois autovalores iguais e  $A$  não é a matriz escalar, então  $A$  não é diagonalizável. Além disso, segue também da Proposição 4 que se  $A$  possui autovalores reais distintos, então  $A$  é diagonalizável.  $\square$

**Definição 5.** Uma matriz  $A$  de  $\mathcal{M}_2(\mathbb{R})$  é **simétrica** se  $A = A^t$ , onde  $A^t$  é a transposta da matriz  $A$ .

**Corolário 2.** Seja  $A$  uma matriz  $2 \times 2$ . Se  $A$  é simétrica, então  $A$  é diagonalizável.

*Demonstração.* Se  $A$  é uma matriz  $2 \times 2$  simétrica, então,  $A$  pode ser escrita na forma

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

Se  $b = 0$ ,  $A$  é uma matriz diagonal. Se  $b \neq 0$ , então, os autovalores de  $A$  são

$$\lambda_1 = \frac{(a + d) + \sqrt{(a - d)^2 + 4b^2}}{2} \quad \text{e} \quad \lambda_2 = \frac{(a + d) - \sqrt{(a - d)^2 + 4b^2}}{2},$$

que são dois autovalores distintos, e  $A$  satisfaz o item (viii) da Proposição 4. Logo,  $A$  é diagonalizável.  $\square$

**Exemplo 3.** O Corolário 1 mostra-nos que as matrizes  $A = \begin{bmatrix} -1 & 4 \\ 0 & -1 \end{bmatrix}$  e  $B = \begin{bmatrix} -3 & 4 \\ -1 & 1 \end{bmatrix}$  não são diagonalizáveis, pois  $A$  e  $B$  não são matrizes escalares e possuem o mesmo polinômio característico  $p(\lambda) = (\lambda + 1)^2$ , que admite apenas uma raiz (dupla)  $\lambda = -1$ .

#### 4. Potência de Matriz

Utilizando a definição e as propriedades da multiplicação de matrizes, se  $n \in \mathbb{N}$ , a  $n$ -ésima potência de uma matriz  $A$  é definida por

$$A^n = \underbrace{A \cdot A \cdots A}_{n\text{-fatores}}.$$

Se  $n$  for um número consideravelmente grande, calcular  $A^n$  envolve um esforço computacional elevado. Nesta seção, apresentaremos um método para calcular a  $n$ -ésima potência de uma matriz  $A$  diagonalizável com um baixo custo computacional em comparação com o cálculo de  $A^n$  utilizando as  $n - 1$  multiplicações de matrizes em  $\underbrace{A \cdot A \cdots A}_{n\text{-fatores}}$ .

**Proposição 5.** Sejam  $A, D$  e  $M$  matrizes em  $\mathcal{M}_2(\mathbb{R})$ , sendo  $M$  invertível e  $D$  uma matriz diagonal, tais que  $A = MDM^{-1}$ . Então,

$$A^n = MD^nM^{-1}.$$

A demonstração da Proposição 5 pode ser feita utilizando-se o *Princípio de Indução* sobre o expoente  $n$ , com  $n \in \mathbb{N}$ .

**Exemplo 4.** Determine a  $n$ -ésima potência da matriz  $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ .

De fato, como os autovalores de  $A$  satisfazem  $p(\lambda) = (1 - \lambda)(-\lambda) = 2$ , temos que

$$\lambda_1 = -1 \quad \text{e} \quad \lambda_2 = 2.$$

Pelo item (viii) da Proposição 4,  $A$  é diagonalizável e satisfaz  $A = MDM^{-1}$  com  $M = \begin{bmatrix} 1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix}$  e

$D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ . Logo, pela Proposição Proposição 5, obtemos

$$\begin{aligned} A^n &= \begin{bmatrix} 1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^n \begin{bmatrix} 1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^n \frac{2}{3} \begin{bmatrix} \frac{1}{2} & -1 \\ 1 & 1 \end{bmatrix} \\ &= \frac{2}{3} \begin{bmatrix} 1 & 1 \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2^{n+1} - (-1)^{n+1}}{3} & \frac{2[2^n - (-1)^n]}{3} \\ \frac{2^{n-1} - (-1)^n}{3} & \frac{2[2^{n-1} - (-1)^{n-1}]}{3} \end{bmatrix} \end{aligned} \quad (15)$$

De modo análogo ao que foi feito no Exemplo 4, o leitor poderá demonstrar o seguinte resultado mais geral:

**Proposição 6.** Se  $A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$ , com  $a^2 > -4b$ , então,

$$A^n = \begin{bmatrix} \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n+1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n+1}}{\sqrt{a^2+4b}} & b \left[ \left(\frac{a+\sqrt{a^2+4b}}{2}\right)^n - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^n \right] \\ \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^n - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^n}{\sqrt{a^2+4b}} & b \left[ \frac{\sqrt{a^2+4b}}{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-1}} \right] \end{bmatrix}. \quad (16)$$

## 5. Recorrências Lineares de Segunda Ordem

Uma relação de recorrência é uma equação que expressa cada elemento de uma sequência  $(x_n)$  como função de elementos precedentes. Relações da forma  $x_{n+1} = ax_n + bx_{n-1}$ , com  $a, b \in \mathbb{R}$ , são chamadas *recorrências lineares de segunda ordem sobre  $\mathbb{R}$* . Dados  $x_1$  e  $x_2$ , se, a partir de  $n = 3$  todos os elementos  $x_n$  de uma sequência satisfazem a relação  $x_{n+1} = ax_n + bx_{n-1}$ , dizemos que a sequência  $(x_n)$  é uma solução para a recorrência linear.

Uma sequência importante que aparece na natureza é a Sequência de Fibonacci. Tal sequência está relacionada ao crescimento populacional de coelhos, ao crescimento do número de galhos de árvores, e aparece também na construção de estruturas naturais. O seguinte exemplo define a Sequência de Fibonacci por meio de uma relação de recorrência:

**Exemplo 5.** A solução da relação de recorrência  $F_{n+1} = F_n + F_{n-1}$ , com  $F_1 = 1$  e  $F_2 = 1$  é conhecida como Sequência de Fibonacci.

Nesta seção, encontraremos soluções para recorrências lineares de segunda ordem para o caso particular em que  $x_1$  e  $x_2$  são conhecidos e  $x_{n+1} = ax_n + bx_{n-1}$ , para  $n \geq 2$  e  $a^2 > -4b$ . Para isso, utilizaremos a seguinte Proposição:

**Proposição 7.** A solução geral para a equação de recorrência  $x_{n+1} = ax_n + bx_{n-1}$  para  $n \geq 2$ ,  $a^2 > -4b$  com os valores iniciais  $x_1$  e  $x_2$  conhecidos é

$$x_n = x_2 \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-1}}{\sqrt{a^2+4b}} + x_1 \frac{b \left[ \left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-2} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-2} \right]}{\sqrt{a^2+4b}}.$$

*Demonstração.* Como  $x_1$  e  $x_2$  são dados, temos  $x_3 = ax_2 + bx_1$  e  $x_4 = ax_3 + bx_2$ . Na forma matricial podemos reescrever essas igualdades como

$$\begin{bmatrix} x_3 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

e

$$\begin{bmatrix} x_4 \\ x_3 \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \left( \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \right) = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}.$$

Utilizando argumento de Indução, obtemos

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix} \left( \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \right) = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}^{n-1} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}.$$

Pela Proposição 6 e Exemplo 4, temos

$$\begin{bmatrix} x_{n+1} \\ x_n \end{bmatrix} = \begin{bmatrix} \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^n - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^n}{\sqrt{a^2+4b}} & \frac{b \left[ \left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-1} \right]}{\sqrt{a^2+4b}} \\ \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-1}}{\sqrt{a^2+4b}} & \frac{b \left[ \left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-2} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-2} \right]}{\sqrt{a^2+4b}} \end{bmatrix} \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}.$$

Portanto,

$$x_n = x_2 \frac{\left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-1} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-1}}{\sqrt{a^2+4b}} + x_1 \frac{b \left[ \left(\frac{a+\sqrt{a^2+4b}}{2}\right)^{n-2} - \left(\frac{a-\sqrt{a^2+4b}}{2}\right)^{n-2} \right]}{\sqrt{a^2+4b}}.$$

□

**Proposição 8.** O termo geral  $F_n$  da Sequência de Fibonacci é dado por

$$F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{1-\sqrt{5}}{2} \right)^n \right].$$

De fato, pela Proposição 7, a relação de recorrência definida por  $F_{n+1} = F_n + F_{n-1}$ , com  $F_1 = 1$  e  $F_2 = 1$ , possui como solução a sequência definida por

$$\begin{aligned} F_n &= \frac{\left( \frac{1+\sqrt{5}}{2} \right)^{n-1} - \left( \frac{1-\sqrt{5}}{2} \right)^{n-1}}{\sqrt{5}} + \frac{\left( \frac{1+\sqrt{5}}{2} \right)^{n-2} - \left( \frac{1-\sqrt{5}}{2} \right)^{n-2}}{\sqrt{5}} \\ &= \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^{n-1} - \left( \frac{1-\sqrt{5}}{2} \right)^{n-1} + \left( \frac{1+\sqrt{5}}{2} \right)^{n-2} - \left( \frac{1-\sqrt{5}}{2} \right)^{n-2} \right] \\ &= \frac{1}{\sqrt{5}} \left\{ \left( \frac{1+\sqrt{5}}{2} \right)^{n-1} \left[ 1 + \left( \frac{1+\sqrt{5}}{2} \right)^{-1} \right] - \left( \frac{1-\sqrt{5}}{2} \right)^{n-1} \left[ 1 + \left( \frac{1-\sqrt{5}}{2} \right)^{-1} \right] \right\} \\ &= \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^{n-1} \left( \frac{1+\sqrt{5}}{2} \right) - \left( \frac{1-\sqrt{5}}{2} \right)^{n-1} \left( \frac{1-\sqrt{5}}{2} \right) \right] \\ &= \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{1-\sqrt{5}}{2} \right)^n \right]. \end{aligned}$$

Por exemplo, para calcular  $F_4$  utilizando a Fórmula obtida na Proposição 8, temos

$$\begin{aligned} F_4 &= \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^4 - \left( \frac{1-\sqrt{5}}{2} \right)^4 \right] \\ &= \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^2 - \left( \frac{1-\sqrt{5}}{2} \right)^2 \right] \left[ \left( \frac{1+\sqrt{5}}{2} \right)^2 + \left( \frac{1-\sqrt{5}}{2} \right)^2 \right] \\ &= \frac{1}{\sqrt{5}} \cdot \frac{4\sqrt{5}}{4} \cdot \frac{2 \cdot (1+5)}{4} = 3. \end{aligned}$$

Calculando  $F_4$  utilizando a Relação de Recorrência, teríamos

$$F_4 = F_3 + F_2 = F_1 + 2F_2 = 1 + 2 \cdot 1 = 3.$$

Podemos notar que utilizar a Fórmula obtida na Proposição 8 demanda mais cálculos quando feitos manualmente comparado com a utilização da Relação de Recorrência quando  $n$  é pequeno. No entanto, a Proposição 8 é vantajosa, pois permite o cálculo de  $F_n$  com  $n$  grande, sem necessidade de calcular todos os anteriores. Além disso, a fórmula obtida serve para explicar que a sequência de Fibonacci cresce exponencialmente, já que podemos utilizar essa fórmula para mostrar que

$$\frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n - 1 < F_n < \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n + 1.$$

## Agradecimentos

Agradecemos aos revisores pelos comentários e sugestões que contribuíram para o aprimoramento deste trabalho.

## Referências

- [1] Nicholson, K. *Linear Algebra with Applications*. 1ª edição. Toronto: McGraw-Hill, 2006.
- [2] Poole, D. *Álgebra Linear: uma introdução moderna*. 2ª edição. São Paulo: Cengage Learning, 2016.
- [3] SILVA, A. F. *Recorrências lineares, isometrias, criptografia e outras aplicações envolvendo matrizes 2 por 2*. Dissertação (Mestrado), Universidade Tecnológica Federal do Paraná. Programa de Mestrado Profissional em Matemática em Rede Nacional. Cornélio Procopio/PR. 96pp. 2017.
- [4] Strang, G. *Álgebra Linear e suas aplicações*. 1ª edição. São Paulo: Cengage Learning, 2010.

Adilson Francisco da Silva  
Universidade Tecnológica Federal do Paraná  
<[adilson554@gmail.com](mailto:adilson554@gmail.com)>

Josimar da Silva Rocha  
Universidade Tecnológica Federal do Paraná  
<[jrocha@utfpr.edu.br](mailto:jrocha@utfpr.edu.br)>

Thiago Pinguello de Andrade  
Universidade Tecnológica Federal do Paraná  
<[thiagoandrade@utfpr.edu.br](mailto:thiagoandrade@utfpr.edu.br)>

Recebido: 20/04/2020  
Publicado: 30/06/2020

# A matemática e a epidemia

Selma Helena de Jesus Nicola 

## Resumo

Para a dedução de uma Curva Epidêmica são utilizadas várias ferramentas matemáticas que podem ser exploradas no ensino. Este artigo tem como objetivo ser um material de apoio teórico para o professor responder perguntas de alunos sobre aspectos da epidemia e a Curva Epidêmica, destacando os tópicos de matemática utilizados.

**Palavras-chave:** Curva Epidêmica; limiar de uma epidemia; ensino; Modelo SIR.

## Abstract

Many elementary mathematical tools are used to get an Epidemic Curve. This article aims to be a theoretical support for teachers to answer student questions about epidemic features and the Epidemic Curve pointing out some useful mathematical topics.

**Keywords:** Epidemic curve; epidemic threshold; teaching; SIR Model.

## 1. Introdução

Nos dias atuais estamos vivendo a crise de uma epidemia e ouvindo muito sobre o pico da curva epidêmica. Mostramos como chegar a uma representação dessa curva em um modelo matemático de uma doença com determinadas características de infecção, cura, mortalidade e condições sobre a população, que é dividida em indivíduos suscetíveis, infectados e removidos, como a Covid-19. Trata-se do bem conhecido Modelo SIR que é, na verdade, um sistema de equações diferenciais ordinárias aqui apresentado na linha do livro de Martin Braun [2], nossa referência básica, também utilizada para o Trabalho de Conclusão de Curso [3].

Para uma solução  $(S(t), I(t), R(t))$  do Sistema SIR, a Curva Epidêmica é representada pelo gráfico da componente  $I(t)$ , que descreve a evolução da classe de indivíduos infectados na população.

Entre os conhecimentos de matemática necessários, além da Álgebra Elementar, estão ferramentas básicas de Cálculo Diferencial e Integral e Equações Diferenciais Ordinárias, para os quais sugerimos [4] e [1]. Assim, podemos aproveitar a crise para motivar os alunos no estudo dessas disciplinas. Destacamos os seguintes tópicos: fatoração de polinômio, resolução de equação de segundo grau, função hiperbólica, limite, derivada, Regra da Cadeia, Teorema Fundamental do Cálculo, Série de Taylor, análise qualitativa de gráfico, integral, mudança de variável, modelagem, ponto de equilíbrio, órbita, plano de fase, equação de variáveis separáveis.

## 2. O Modelo

Consideramos que uma população tenha sido afetada pelo aparecimento de um microrganismo que provoca uma doença contagiosa. Vamos dividi-la nas classes

- (S) dos suscetíveis, que consiste daqueles indivíduos que não foram infectados mas estão sujeitos a contrair a doença,
- (I) dos infectados, que consiste daqueles indivíduos que são capazes de transmitir a doença para outros,
- (R) dos removidos, que consiste daqueles indivíduos que se recuperaram, estão se recuperando em isolamento ou foram a óbito.

Supomos as seguintes hipóteses:

- A população permanece num valor fixo  $N$  no intervalo de tempo considerado, isto é, desprezamos nascimentos, mortes por causas não relacionadas à doença em questão, imigração e emigração.
- A doença confere imunidade a todo indivíduo que foi infectado e tenha se recuperado.
- Um indivíduo da classe  $S$  que contrai a doença passa imediatamente para a classe  $I$ .
- A taxa de variação da população suscetível é proporcional ao produto do número de indivíduos da classe  $S$  e o número de indivíduos da classe  $I$ , ou seja, é proporcional ao número de encontros de indivíduos dessas classes.
- Indivíduos são removidos a uma taxa proporcional ao tamanho da classe  $I$ .

As funções  $S(t)$ ,  $I(t)$ ,  $R(t)$  dadas pelo número de indivíduos no instante  $t$  das classes  $S$ ,  $I$  e  $R$ , respectivamente, satisfazem então o sistema de equações diferenciais

$$\begin{aligned}\frac{dS}{dt} &= -rSI \\ \frac{dI}{dt} &= rSI - \gamma I \\ \frac{dR}{dt} &= \gamma I,\end{aligned}\tag{1}$$

onde  $r$  e  $\gamma$  são as constantes positivas de contágio e de remoção, respectivamente.

## 3. Ocorrência ou não de epidemia

Analisamos as soluções do Sistema (1) para saber como será a propagação da doença e, como as primeiras equações não dependem de  $R$ , concentramo-nos no sistema

$$\begin{aligned}\frac{dS}{dt} &= -rSI \\ \frac{dI}{dt} &= rSI - \gamma I.\end{aligned}\tag{2}$$

Observamos que a função  $R(t)$  pode ser obtida da terceira equação do Sistema (1) ou, como  $S(t) + I(t) + R(t) = N$ , temos  $R(t) = N - S(t) - I(t)$ .

Uma *órbita* do Sistema (2) é a curva  $(S(t), I(t))$  descrita pela solução no chamado *plano de fase* SI. Uma das vantagens de considerar a órbita é a de ser frequentemente possível descrevê-la sem o conhecimento da correspondente solução. Os *pontos de equilíbrio* do Sistema (2) são as soluções constantes, ou seja, os pontos que anulam os segundos membros. Note que os pontos de equilíbrio do Sistema (2) são os pontos  $(S, 0)$  do eixo S. Além disso, o semieixo I positivo é uma órbita, pois a solução que passa por qualquer ponto  $(0, I_0)$ ,  $I_0 > 0$ , é  $(0, I_0 e^{-\gamma t})$ ,  $-\infty < t < \infty$ . Assim, como duas órbitas nunca se cruzam, todas as órbitas com condição inicial no primeiro quadrante permanecem nele. Esse fato revela que o sistema é coerente com a realidade, pois em nenhum instante a população de uma classe pode ser negativa.

De agora em diante supomos o instante inicial  $t_0 = 0$  e que o número de indivíduos suscetíveis, infectados e removidos nesse instante sejam  $S_0, I_0 > 0$  e  $R_0 = 0$ , respectivamente. Ou seja, vamos considerar uma condição inicial  $(S(0), I(0), R(0)) = (S_0, I_0, 0)$ , portanto  $S(t), I(t)$  e  $R(t)$  positivos para  $t > 0$ . De acordo com (2), a componente  $S(t)$  é decrescente e, portanto, vamos nos ater a  $S < S_0$ .

Como estamos restritos ao quadrante  $S, I > 0$ , podemos utilizar a *regra da derivação da função inversa* e a *regra da cadeia* para obter as órbitas do Sistema (2) como as curvas soluções da *equação diferencial ordinária de primeira ordem*

$$\frac{dI}{dS} = \frac{dI}{dt} \frac{dt}{dS} = \frac{rSI - \gamma I}{-rSI} = -1 + \frac{\rho}{S}, \quad (3)$$

onde definimos  $\rho = \gamma/r$ .

Integrando a Equação (3) de  $S_0$  a  $S$  obtemos

$$I(S) = I_0 + S_0 - S + \rho \ln \frac{S}{S_0}. \quad (4)$$

A função  $I(S)$  está definida para todo  $S > 0$ , mas é solução da Equação (3) somente no intervalo em que  $I(S) > 0$ . Para analisar o comportamento das curvas (4), observamos que

- $I(S)$  é *decrescente* se  $dI/dS = -1 + \rho/S < 0$ , ou seja,  $S > \rho$ ,
- $I(S)$  é *crescente* se  $dI/dS = -1 + \rho/S > 0$ , ou seja,  $S < \rho$ .

Portanto,  $S = \rho$  é o único ponto de máximo da função  $I(S)$ , ou seja, de qualquer solução  $I(S)$  da Equação (3).

Como  $\lim_{S \rightarrow 0} I(S) = -\infty$  e  $I(\rho) \geq I(S_0) = I_0 > 0$ , sendo  $I(S)$  contínua e crescente no intervalo  $(0, \rho)$ , segue do *Teorema do Anulamento* que existe um único ponto  $S_\infty \in (0, \rho]$  tal que  $I(S_\infty) = 0$ .

O ponto  $(S_\infty, 0)$  é um *ponto de equilíbrio* do Sistema (2), e, como as coordenadas  $S(t)$  e  $I(t)$  da órbita são *decrescentes* e *limitadas* para  $S < \rho$ , então,  $\lim_{t \rightarrow \infty} (S(t), I(t))$  deve ser um *ponto de equilíbrio*. Logo,  $\lim_{t \rightarrow \infty} (S(t), I(t)) = (S_\infty, 0)$ .

Portanto, as órbitas do Sistema (2) têm a forma que se indica na Figura 1, e, quando  $t$  varia de 0 até  $\infty$ , o ponto  $(S(t), I(t))$  percorre a órbita movendo-se na direção de  $S$  decrescente. Assim,

- se  $S_0 < \rho$ ,  $S(t)$  decresce a  $S_\infty$  e  $I(t)$  decresce à extinção. Nesse caso não ocorre epidemia.
- se  $S_0 > \rho$ ,  $I(t)$  cresce enquanto  $S(t)$  decresce para  $\rho$  e atinge um valor máximo quando  $S = \rho$ . A partir desse instante  $I(t)$  e  $S(t)$  decrescem com  $I(t)$  tendendo à extinção. Este é o caso em que ocorre uma epidemia.

Concluimos que ocorre uma epidemia se, e somente se, o número inicial  $S_0$  de indivíduos suscetíveis for maior que o número  $\rho$ . Seja qual for a condição inicial  $(S_0, I_0)$ , com  $S_0 > \rho$ , a órbita correspondente atinge um máximo para  $S = \rho$ ; portanto,  $\rho$  é uma característica da enfermidade. O número  $\rho$  é chamado *limiar* (*threshold*) da epidemia.

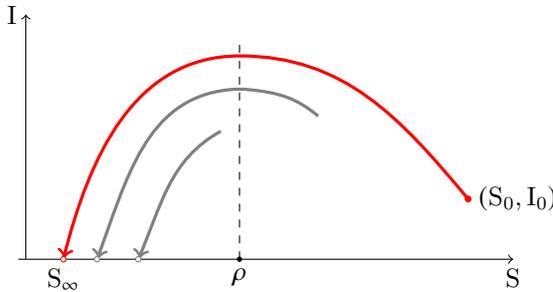


Figura 1: Órbita por  $(S_0, I_0)$  no plano de fase SI.

Uma estratégia para controlar uma epidemia é aumentar o limiar  $\rho$  aproximando-o do valor inicial  $S_0$  de indivíduos suscetíveis. Isso é possível aumentando  $\gamma$  e ou diminuindo  $r$ . Na prática, o aumento da constante  $\gamma$  pode ser obtido melhorando as condições hospitalares e investindo em pesquisas médicas. Para diminuir a constante  $r$  pode-se praticar o afastamento social e estimular condutas de proteção pessoal. Veremos mais adiante que a diminuição de  $r$  tem também a contribuição de diminuir o valor máximo  $I(\rho)$  de indivíduos infectados.

No caso de existir uma vacina, a estratégia seria diminuir o  $S_0$  visando tornar  $S_0 < \rho$ .

#### 4. Os valores finais $S_\infty$ e $R_\infty$

O valor  $I_\infty$  foi excluído do título desta seção, pois já sabemos que  $I_\infty = \lim_{t \rightarrow \infty} I(t) = 0$ .

A função  $S(t)$  é decrescente, mas ao final de uma epidemia sempre restará uma quantidade  $S_\infty = \lim_{t \rightarrow \infty} S(t)$  de indivíduos suscetíveis. Tal fato torna provável ressurgimentos de epidemias da mesma enfermidade.

A função  $R(t)$  é crescente, e portanto,  $R_\infty$  é maior que  $R(t)$  em qualquer momento da epidemia. Observamos que a quantidade total  $R_\infty = \lim_{t \rightarrow \infty} R(t)$  de indivíduos recuperados é fornecida por  $R_\infty = I_0 + S_0 - S_\infty$ , e para obtê-la basta encontrar  $S_\infty$ .

Fazendo  $t \rightarrow \infty$  em (4) temos  $0 = I_0 + S_0 - S_\infty + \rho \ln(S_\infty/S_0)$ , o que implica

$$\ln \frac{S_\infty}{S_0} = \frac{1}{\rho} (S_\infty - S_0 - I_0).$$

Da primeira equação do Sistema (2) temos

$$\frac{dS}{S} = -rI dt \Rightarrow \ln \frac{S}{S_0} = -rIt \Rightarrow S = S_0 e^{-rIt} \quad \text{e} \quad It = -\frac{1}{r} \ln \frac{S}{S_0}.$$

Substituindo  $It$  em  $S = S_0 e^{-rIt}$  e fazendo  $t \rightarrow \infty$  obtemos  $S_\infty = S_0 e^{(S_\infty - S_0 - I_0)/\rho}$ , ou seja,  $S_\infty$  é raiz da equação

$$S = S_0 e^{(S - S_0 - I_0)/\rho}. \tag{5}$$

Essa equação não pode ser resolvida algebricamente, mas afirmamos que ela possui duas raízes  $S_\infty$  e  $S'$ . Buscar as soluções da Equação (5) é equivalente a encontrar as raízes da função

$$f(S) = S - S_0 e^{(S - S_0 - I_0)/\rho}.$$

Observando que  $d^2f/dS^2 = -(S_0/\rho^2)e^{(S - S_0 - I_0)/\rho}$  é negativa, podemos garantir que o gráfico de  $f$  tem a concavidade voltada para baixo. Como  $\lim_{S \rightarrow \pm\infty} f(S) = -\infty$  a  $f$  tem um ponto de máximo e, sendo  $f(S_0) = S_0(1 - e^{-I_0/\rho})$  positivo, segue que  $f$  possui duas raízes, a primeira igual a  $S_\infty < S_0$  e a segunda  $S'$  que será desprezada por satisfazer  $S' > S_0$ .

### 5. A Curva Epidêmica

Nesta seção obtemos as funções coordenadas  $I(t)$  e  $R(t)$ . A Curva Epidêmica é o gráfico da função  $I(t)$ . Note que, conhecidos  $I(t)$  e  $R(t)$ , obtemos  $S(t) = N - I(t) - R(t)$ .

Enquanto está ocorrendo uma epidemia, conseguimos registrar o número de pessoas que procuram ajuda médica e o número de óbitos, ou seja, conseguimos registrar somente o número de novos indivíduos removidos. Sendo assim, podemos conhecer  $dR/dt$ .

Pelas regras da cadeia e da derivação da função inversa,  $dS/dR = (dS/dt)(dt/dR) = (dS/dt)(dR/dt)^{-1}$  ou seja,

$$\frac{dS}{dR} = \frac{-rSI}{\gamma I} = \frac{-S}{\rho},$$

e, resolvendo esta equação diferencial ordinária, obtemos

$$S(R) = S_0 e^{-R/\rho}.$$

Como

$$\frac{dR}{dt} = \gamma I = \gamma(N - R - S),$$

temos

$$\frac{dR}{dt} = \gamma(N - R - S_0 e^{-R/\rho}). \tag{6}$$

A epidemia é em geral pequena em relação a  $N$ . Assim,  $S$  é sempre um número grande em comparação com  $R$ , e, como  $\rho$  é um valor de  $S$ , temos  $R/\rho$  muito pequeno. Logo, é possível truncar a *Série de Taylor*

$$e^{-R/\rho} = 1 - \frac{R}{\rho} + \frac{1}{2} \left(\frac{R}{\rho}\right)^2 + \dots$$

depois do terceiro termo.

Com tal aproximação obtemos

$$\frac{dR}{dt} = \gamma \left[ N - R - S_0 \left( 1 - \frac{R}{\rho} + \frac{1}{2} \left( \frac{R}{\rho} \right)^2 \right) \right] = \gamma \left[ N - S_0 + \left( \frac{S_0}{\rho} - 1 \right) R - \frac{S_0}{2} \left( \frac{R}{\rho} \right)^2 \right], \quad (7)$$

ou seja,

$$\frac{dR}{dt} = \gamma \left[ (N - S_0) + \left( \frac{S_0}{\rho} - 1 \right) R - \left( \frac{S_0}{2\rho^2} \right) R^2 \right]. \quad (8)$$

Definindo as variáveis

$$A = -\frac{S_0}{2\rho^2}, \quad B = \frac{S_0}{\rho} - 1, \quad C = N - S_0$$

e, substituindo em (8), obtemos a equação de *variáveis separáveis*

$$\frac{dR}{AR^2 + BR + C} = \gamma dt. \quad (9)$$

Observando que  $B^2 - 4AC > 0$ , pois  $N > S_0$ , podemos definir

$$\alpha = \sqrt{B^2 - 4AC} = \left[ \left( \frac{S_0}{\rho} - 1 \right)^2 + \frac{2S_0(N - S_0)}{\rho^2} \right]^{1/2} \quad (10)$$

e a equação do segundo grau  $AR^2 + BR + C = 0$ , possui, então as raízes

$$R = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} = \frac{-B \pm \alpha}{2A} = \rho - \frac{\rho^2}{S_0} \pm \frac{\rho^2}{S_0} \alpha,$$

e, definindo as variáveis

$$D = \rho - \frac{\rho^2}{S_0} \quad \text{e} \quad E = \frac{\rho^2}{S_0},$$

tornam-se

$$R = D \pm E\alpha.$$

Podemos *fatorar*  $AR^2 + BR + C$  da seguinte maneira

$$A[R - (D + E\alpha)][R - (D - E\alpha)] = A[(R - D) + E\alpha][(R - D) - E\alpha] = A[(R - D)^2 - E^2\alpha^2]. \quad (11)$$

Assim,

$$\int \frac{dR}{AR^2 + BR + C} = \frac{1}{A} \int \frac{dR}{(R - D)^2 - E^2\alpha^2}.$$

Fazendo a *mudança de variável*  $x = R - D$ , e definindo  $a^2 = E^2\alpha^2$ , temos

$$\frac{1}{A} \int \frac{dR}{(R - D)^2 - E^2\alpha^2} = \frac{1}{A} \int \frac{dx}{x^2 - a^2},$$

e, com uma nova *mudança de variável*  $x = a \tanh u$ , temos

$$\frac{1}{A} \int \frac{dx}{x^2 - a^2} = -\frac{1}{A} \int \frac{a \operatorname{sech}^2 u}{a^2 \operatorname{sech}^2 u} du = -\frac{1}{Aa} \int du = -\frac{1}{Aa} u + K = -\frac{1}{Aa} \operatorname{arctanh} \frac{x}{a} + K,$$

onde  $K$  é uma constante. Nas variáveis  $R, D, E$  temos

$$\frac{1}{A} \int \frac{dR}{(R-D)^2 - E^2 \alpha^2} = -\frac{1}{AE\alpha} \operatorname{arctanh} \frac{R-D}{E\alpha} + K$$

e, como

$$R-D = R - \left( \rho - \frac{\rho^2}{S_0} \right), \quad E\alpha = \frac{\rho^2}{S_0} \alpha, \quad AE\alpha = -\frac{\alpha}{2},$$

temos

$$-\frac{1}{AE\alpha} \operatorname{arctanh} \frac{R-D}{E\alpha} + K = \frac{2}{\alpha} \operatorname{arctanh} \left\{ \frac{1}{\alpha} \left[ \frac{S_0}{\rho^2} R - \left( \frac{S_0}{\rho} - 1 \right) \right] \right\} + K. \quad (12)$$

Voltando à Equação (9), temos

$$\frac{2}{\alpha} \operatorname{arctanh} \left\{ \frac{1}{\alpha} \left[ \frac{S_0}{\rho^2} R - \left( \frac{S_0}{\rho} - 1 \right) \right] \right\} + K = \gamma t.$$

Observamos que para  $t = 0$  temos  $R = 0$  e

$$K = \frac{2}{\alpha} \operatorname{arctanh} \left[ \frac{1}{\alpha} \left( \frac{S_0}{\rho} - 1 \right) \right].$$

Então,

$$\operatorname{arctanh} \left\{ \frac{1}{\alpha} \left[ \frac{S_0}{\rho^2} R - \left( \frac{S_0}{\rho} - 1 \right) \right] \right\} = \frac{1}{2} \alpha \gamma t - \operatorname{arctanh} \left[ \frac{1}{\alpha} \left( \frac{S_0}{\rho} - 1 \right) \right],$$

ou seja,

$$\frac{1}{\alpha} \left[ \frac{S_0}{\rho^2} R - \left( \frac{S_0}{\rho} - 1 \right) \right] = \tanh \left( \frac{1}{2} \alpha \gamma t - \phi \right),$$

onde

$$\phi = \operatorname{arctanh} \left[ \frac{1}{\alpha} \left( \frac{S_0}{\rho} - 1 \right) \right].$$

Logo,

$$R(t) = \frac{\rho^2}{S_0} \left[ \frac{S_0}{\rho} - 1 + \alpha \tanh \left( \frac{1}{2} \alpha \gamma t - \phi \right) \right]. \quad (13)$$

Derivando  $R$  em relação a  $t$  na expressão (13), obtemos

$$\frac{dR}{dt} = \frac{\gamma \alpha^2 \rho^2}{2S_0} \operatorname{sech}^2 \left( \frac{1}{2} \alpha \gamma t - \phi \right). \quad (14)$$

Como  $dR/dt = \gamma I$  temos finalmente

$$I(t) = \frac{\alpha^2 \rho^2}{2S_0} \operatorname{sech}^2 \left( \frac{1}{2} \alpha \gamma t - \phi \right). \quad (15)$$

Essa função tem um único *ponto crítico*  $\bar{t} = 2\phi/\alpha\gamma$ , que obtemos resolvendo em  $t$  a equação  $dI/dt = 0$  e, como  $(d^2I/dt^2)(\bar{t}) < 0$ , temos que  $\bar{t}$  é *ponto de máximo*, sendo  $I(\bar{t}) = \alpha^2 \rho^2 / 2S_0$  o *valor máximo*. Como a função *secante hiperbólica* é par, a função  $I(t)$  é simétrica em relação à reta  $t = \bar{t}$  e tem dois *pontos de inflexão* também simétricos em relação a essa, os quais obtemos resolvendo em  $t$  a equação  $d^2I/dt^2 = 0$ . Como  $I(t) > 0$ , para todo  $t$ , e  $\lim_{t \rightarrow \pm\infty} I(t) = 0$ , o *gráfico* da função  $I(t)$

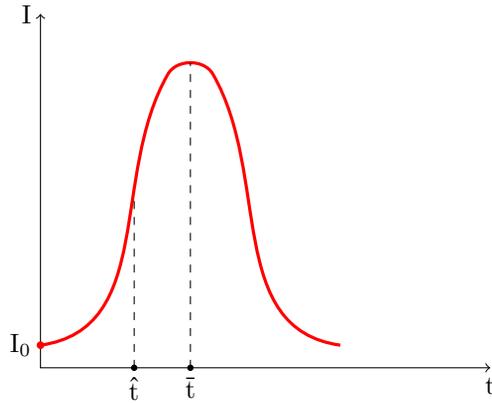


Figura 2: A Curva Epidêmica.

define uma curva simétrica com forma de sino conhecido como *Curva Epidêmica* da enfermidade; veja Figura 2.

Tal *curva* confirma a observação comum de que, em muitas epidemias reais, o número de novos casos reportados cada dia aumenta até alcançar o valor máximo  $I(\bar{t})$  no instante  $\bar{t}$ .

De acordo com (10) temos

$$\alpha = \frac{1}{\rho} \sqrt{(S_0 - \rho)^2 + 2S_0(N - S_0)^2}$$

e, portanto, o valor máximo  $I(\bar{t})$ , o chamado “*pico da epidemia*”, satisfaz

$$I(\bar{t}) = \frac{\alpha^2 \rho^2}{2S_0} = \frac{1}{2S_0} [(S_0 - \rho)^2 + 2S_0(N - S_0)^2] \quad (16)$$

no instante

$$\bar{t} = \frac{2\phi}{\alpha\gamma} = \frac{2\phi}{r\sqrt{(S_0 - \rho)^2 + 2S_0(N - S_0)^2}}. \quad (17)$$

Como, para que haja epidemia, é necessário  $S_0 > \rho$ , e, lembrando que  $\rho = \gamma/r$ , se diminuirmos a constante de contágio  $r$ , aumentamos  $\rho$ , e, portanto, de acordo com (16) e (17), diminuimos o valor do pico  $I(\bar{t})$  e retardamos o instante  $\bar{t}$ .

Analisando a Curva Epidêmica podemos entender o porquê de as autoridades de saúde falarem em “virada do jogo” antes do pico da epidemia. Isso se justifica porque a Curva Epidêmica começa com um crescimento acentuado e o gráfico tendo a convexidade voltada para cima (a derivada é crescente). Se, porém, em um instante  $\hat{t}$  anterior ao instante  $\bar{t}$  do pico há uma inflexão, a concavidade volta-se para baixo (a derivada passa a decrescer). Por um lapso de tempo a epidemia ainda cresce, mas o crescimento começa a diminuir. A partir de  $\hat{t}$ , o instante da “virada”, a epidemia tende de modo sustentável à extinção, passando a decrescer somente depois de  $\bar{t}$ .

## 6. Comentários Finais

A utilização de modelos matemáticos tem sido uma ferramenta importante para o entendimento de uma epidemia. O conhecimento do Modelo SIR é importante para o sistema de saúde no

planejamento do combate a uma epidemia, pois ele fornece a evolução da população de indivíduos infectados, a qual é, na verdade, a evolução da epidemia.

O contágio da Covid-19 tem um grau de nocividade maior do que as enfermidades mais conhecidas cujas epidemias são descritas pelo Modelo SIR. Um indivíduo infectado, por onde passa, vai deixando um rastro de focos de infecção que se desvanece em um tempo considerável. Assim, um indivíduo suscetível pode ser infectado por ele sem nunca tê-lo encontrado. Ou seja, a Covid-19 é mais contagiosa mas isso não inviabiliza o Modelo SIR para descrevê-la, basta tomar uma constante  $r$  maior.

A Curva Epidêmica real não tem as simetrias da forma de um sino, mas é uma deformação da curva estudada aqui. No mesmo sentido o limiar é apenas uma estimativa, mas é um parâmetro extremamente importante para o conhecimento da epidemia, pois a partir dele podemos estimar o valor e o instante do pico. É preciso entender que um modelo matemático é sempre uma aproximação da realidade, uma simplificação. Se não fosse assim, ele seria um retrato fiel, em geral tão complexo quanto a própria realidade.

Finalmente, na inexistência de vacina e considerando as dificuldades naturais de implementar rapidamente um tratamento clínico, resta uma única estratégia eficaz de combater uma epidemia. Trata-se de diminuir o contágio pelo distanciamento social e a difusão da proteção pessoal.

## Agradecimentos

A autora agradece a P.Z. Táboas por algumas profícuas discussões.

## Referências

- [1] Boyce, W.E.; DiPrima, R.C. *Equações Diferenciais Elementares e Problemas de Valores de Contorno*, 7<sup>a</sup> edição, LTC - Livros Técnicos e Científicos Editora, Rio de Janeiro, RJ, 2002.
- [2] Braun, M. *Differential Equations and their Applications*, 4<sup>a</sup> edição, Springer, New York, NY, USA, 1993.
- [3] Rodrigues, J.V.S. *Um Modelo Matemático em Epidemiologia*, Trabalho de Conclusão de Curso, Departamento de Matemática, UFSCar - Universidade Federal de São Carlos, São Carlos, SP, 2019.
- [4] Táboas, P.Z. *Cálculo em uma Variável Real*, Edusp - Editora da Universidade de São Paulo, São Paulo, SP, 2008.

Selma Helena de Jesus Nicola  
Universidade Federal de São Carlos  
<[selma.nicola@ufscar.br](mailto:selma.nicola@ufscar.br)>

Recebido: 21/04/2020  
Publicado: 02/07/2020

# O GeoGebra como recurso na promoção da interdisciplinaridade na educação básica: a geometria das bandeiras

Inês Naves Cunha de Oliveira <sup>1</sup> 

Thiago Porto de Almeida Freitas 

## Resumo

O presente artigo apresenta uma proposta de atividade matemática desenvolvida em contexto interdisciplinar para estudantes do sexto ano do Ensino Fundamental de uma escola municipal da cidade de Uberlândia no estado de Minas Gerais. A atividade subsidiou uma investigação ocorrida no âmbito do Mestrado Profissional em Matemática em Rede Nacional - Profmat e abordou aspectos históricos, geométricos, artísticos e de legislações das bandeiras: do Brasil, do estado de Minas Gerais e do município de Uberlândia. Em especial, os elementos geométricos foram explorados com o *software* GeoGebra no laboratório de informática da escola. A partir dessa experiência, observou-se uma mudança no comportamento dos estudantes e na receptividade dos conteúdos matemáticos explorados. Ademais, o envolvimento de professores de outras áreas contribuiu para o melhor entendimento dos educandos acerca das bandeiras e de aspectos correlacionados.

**Palavras-chave:** Bandeiras; GeoGebra; Geometria; Interdisciplinaridade.

## Abstract

This article presents a proposal for mathematical activity developed in an interdisciplinary context for students in the sixth year of elementary school at a municipal school in the city of Uberlândia in the state of Minas Gerais. The activity supported an investigation that took place within the scope of the Professional Master's in Mathematics in the National Network - PROFMAT and covered historical, geometric, artistic and legislative aspects of the flags: of Brazil, the State of Minas Gerais and municipality of Uberlândia. In particular, the geometric elements were explored with the GeoGebra software in the school's informatics laboratory. From this experience, there was a change in the students' behavior and in the receptivity of the mathematical contents explored. In addition, the involvement of teachers from other areas contributed to a better understanding of students about the flags and related aspects.

**Keywords:** Flags; GeoGebra; Geometry; Interdisciplinarity.

<sup>1</sup>Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes durante o Profmat.

## 1. Introdução

A matemática é uma das disciplinas que mais provocam queixas por estudantes e professores. De um lado, estudantes reclamam da dificuldade de aprender tópicos abstratos, sobre os quais, *a priori*, não enxergam a relação com seu cotidiano. Por outro lado, professores ficam insatisfeitos com o índice de aproveitamento da turma, que não atinge sua expectativa. Não são raras as vezes em que, dentro de sala de aula, o professor ouve determinadas expressões: “Não entendi nada!”, “Esse negócio é muito difícil!”, “O que é que eu vou fazer com isso?”, “Eu não vou seguir na área de Matemática, então pra que eu estou estudando isso?”.

Segundo Miorim e Fiorentini,

*As dificuldades encontradas por alunos e professores no processo de ensino-aprendizagem da matemática, em todas as suas áreas, são muitas e conhecidas. Por um lado, o aluno não consegue entender a matemática que a escola lhe ensina, muitas vezes é reprovado nesta disciplina, ou, então, mesmo que aprovado, sente dificuldade em utilizar o conhecimento matemático “adquirido”; em síntese, não consegue efetivamente ter acesso a esse saber de fundamental importância.* (Miorim e Fiorentini, 1990, p1)

No âmbito do ensino de geometria, Lorezato aponta

*A segunda causa da omissão geométrica deve-se à exagerada importância que, entre nós, desempenha o livro didático, quer devido à má formação de nossos professores, quer devido à estafante jornada de trabalho a que estão submetidos. E como a Geometria neles aparece? Infelizmente em muitos deles a Geometria é apresentada apenas como um conjunto de definições, propriedades, nomes e fórmulas, desligado de quaisquer aplicações ou explicações de natureza histórica ou lógica; noutros a Geometria é reduzida a meia dúzia de formas banais do mundo físico. Como se isso não bastasse, a Geometria quase sempre é apresentada na última parte do livro, aumentando a probabilidade de não vir a ser estudada por falta de tempo letivo.* (Lorenzato, 1995, p4)

A aprendizagem de conceitos geométricos no Ensino Fundamental é de extrema importância à continuidade dos estudos matemáticos, uma vez que tais conceitos entrelaçam-se com outras áreas da Matemática e do conhecimento. Dessa maneira, é importante relacionar atividades geométricas para além daquelas que estão nos livros didáticos. Assim, ações com outras disciplinas e recursos podem ser uma ótima alternativa, isto é, promoção de momentos interdisciplinares na escola.

Segundo Berger,

*interdisciplinaridade é a interação existente entre duas ou mais disciplinas. Essa interação pode ir desde a simples comunicação das ideias até à integração mútua dos conceitos diretos, da epistemologia, da terminologia, da metodologia, dos procedimentos, dos dados e da organização da investigação e do ensino correspondentes.* (Berger, 1972, p24)

Nesse contexto, em uma escola municipal da cidade de Uberlândia, em Minas Gerais, foi desenvolvida, com alunos do sexto ano do Ensino Fundamental, uma experiência interdisciplinar. A partir

de atividades inseridas nas aulas de Artes, História, Matemática e Português, onde foram explorados diversos contextos da temática das bandeiras, que se constituem símbolos representativos de países e estados.

Foram escolhidas as bandeiras do Brasil, do estado de Minas Gerais e da cidade de Uberlândia. A disciplina de Artes promoveu discussão em torno de elementos artísticos das referidas bandeiras. Nas aulas de História ocorreu o estudo de cada uma das bandeiras sob a perspectiva histórica de sua idealização e dos significados das cores. A disciplina de Português aprofundou-se na linguagem utilizada para redigir a legislação de cada uma das bandeiras. Nas aulas de Matemática, no laboratório de informática da escola, foram realizadas as construções geométricas das bandeiras com o apoio do *software* GeoGebra.

Pormenorizar as atividades das disciplinas de Artes, de História e Língua Portuguesa não constitui o escopo deste artigo. Na próxima seção, com respaldo das respectivas legislações, apresentamos, respectivamente, as construções geométricas das bandeiras do Brasil, de Minas Gerais e de Uberlândia, que ocorreram nas aulas de Matemática. Em caixas coloridas exibimos comentários para que professores de matemática potencializem a realização de experiência semelhante em sua sala de aula: caixa amarela informa uma estimativa de tempo para a construção da bandeira, caixa vermelha apresenta a conexão da atividade com as habilidades da Base Nacional Comum Curricular e a caixa verde traz reflexões relativas ao GeoGebra e aos conceitos matemáticos abordados na atividade.

## 2. Construção de Bandeiras

As construções geométricas das bandeiras foram baseadas nas legislações e realizadas no *software* GeoGebra *Classic 5*. As construções tiveram ênfase em conteúdos do sexto ano do Ensino Fundamental, mas podem ser adaptadas para aprofundar em conteúdos de outras séries da educação básica. Em sala de aula, as construções podem se tornar mais atraentes se realizadas de forma dialogada com os alunos e paralela ao estudo dos conteúdos matemáticos relacionados.

### 2.1. A Bandeira do Brasil

⊗ Para a construção da bandeira do Brasil, com o apoio do GeoGebra, estimam-se 2 aulas, tendo em vista que não serão criadas as estrelas e a frase da bandeira.

O formato da bandeira do Brasil é descrito no Art. 5º e no anexo II da Lei Nacional N º 5.700, de 1º de setembro de 1971.

O supracitado artigo é taxativo em seus dois primeiros incisos:

*I- Para cálculo das dimensões, tomar-se-á por base a largura desejada, dividindo-se esta em 14 (quatorze) partes iguais. Cada uma das partes será considerada uma medida ou módulo.*

*II- O comprimento será de vinte módulos (20M).*

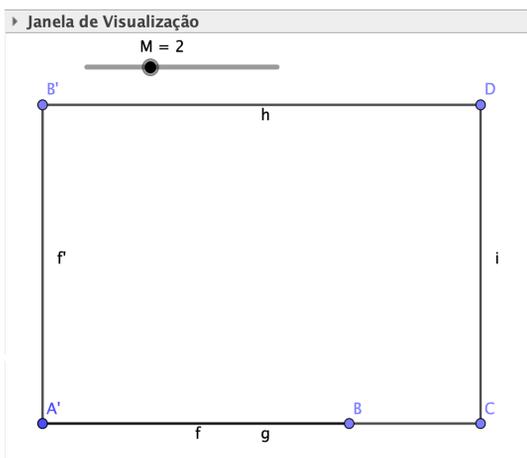
(BRASIL, 1971, p1)

A partir de uma análise do referido texto legal, constata-se que a bandeira nacional possui formato geométrico de um retângulo com comprimento e largura, respectivamente,  $20M$  e  $14M$ , onde  $M$  é uma medida a ser adotada para a construção. Para a construção desse, apresentamos uma possibilidade de roteiro de rotinas a serem executadas no ambiente do GeoGebra, onde *Eixos* e *Malha* devem ser ocultados na *Janela de Visualização*.

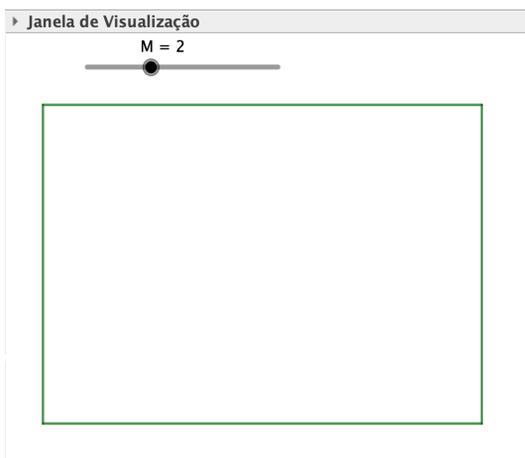
☞ Habilidades da Base Nacional Comum Curricular: **EF06MA19, EF06MA20, EF06MA21 e EF06MA22.**

- **Passo 1 - BRA:** Crie um *Controle Deslizante*  $M$  e de parâmetros de intervalo: mínimo 1, máximo 5 e incremento 1;
- **Passo 2 - BRA:** Com o recurso *Segmento de Comprimento Fixo*, crie o segmento  $AB$  de comprimento  $14M$ . Este segmento será criado horizontalmente, entretanto precisa se apresentar na vertical, como a largura da bandeira. Para isso, com a ferramenta *Rotação em Torno de um Ponto*, gire  $90$  graus, no sentido anti-horário, o segmento  $AB$  em torno do ponto  $A$  e obtenha o segmento  $A'B'$ ;
- **Passo 3 - BRA:** Com *Segmento de Comprimento Fixo*, obtenha os segmentos  $AC$  e  $B'D$  de comprimento  $20M$ ;
- **Passo 4 - BRA:** Com *Segmento*, construa o segmento  $CD$  para obter o retângulo  $A'CDB'$ ;
- **Passo 5 - BRA:** Com botão direito do *mouse*, apague o segmento  $AB$  e o ponto  $B$ . Na sequência, em *Propriedades*, fixe e desabilite o *Exibir Rótulo* de todos os elementos da *Janela de Visualização*, e desabilite a *Exibição do Objeto* dos vértices do retângulo  $A'CDB'$ ;
- **Passo 6 - BRA:** Em *Propriedades*, coloque os objetos da *Janela de Visualização*, os lados do retângulo da bandeira, na cor verde.

☞ A alteração do valor do *Controle Deslizante* permitirá a visualização de vários tamanhos de retângulo com as propriedades construídas. Para a atividade sugerimos adotar  $M = 2$ , para ilustração completa do retângulo na *Janela de Visualização*.



(a) Após Passo 4 - BRA.



(b) Após Passo 6 - BRA.

Figura 1: Construção do Retângulo da Bandeira do Brasil.

O terceiro inciso do Art. 5º da legislação da Bandeira Nacional explicita que

*III- A distância dos vértices do losango amarelo ao quadro externo será de um módulo e sete décimos (1,7M).*

(BRASIL, 1971, p1)

Conforme trecho legal supracitado, para continuidade da construção é preciso determinar os vértices do losango amarelo e, como consequência, obter seus lados. Para isso, sugerem-se os seguintes passos de construção.

- **Passo 7 - BRA:** Com a ferramenta *Ponto Médio ou Centro*, encontre os pontos médios de todos os lados do retângulo construído;
- **Passo 8 - BRA:** Una, por meio da ferramenta *Segmento*, os pontos médios dos lados paralelos do retângulo e obtenha dois segmentos perpendiculares;
- **Passo 9 - BRA:** A partir da ferramenta *Interseção de Dois Objetos*, exiba o ponto de interseção dos segmentos perpendiculares construídos no passo anterior;
- **Passo 10 - BRA:** Por meio de raciocínio análogo ao Passo 2, encontre os vértices do losango, com o recurso *Segmento com Comprimento Fixo* com comprimento de  $1,7M$ . Em *Propriedades*, fixe os vértices do losango;
- **Passo 11 - BRA:** Una os pontos encontrados no passo anterior, com a ferramenta *Segmento* e obtenha o losango da bandeira.
- **Passo 12 - BRA:** Desabilite as marcações de *Exibir Rótulo* dos lados do losango obtido no Passo 11. Desabilite *Exibir Objeto* dos segmentos criados e que não são lados do retângulo e do losango, de modo a manter visíveis à interseção o centro da bandeira, os pontos médios dos lados

do retângulo e os vértices do losango e retângulo da bandeira. Por fim, atribua a cor amarelo para os lados do losango.

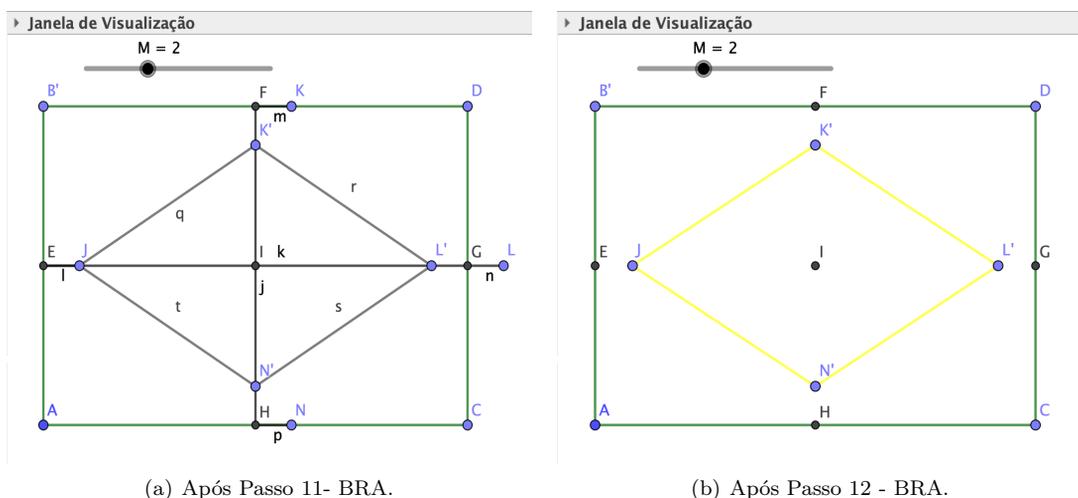
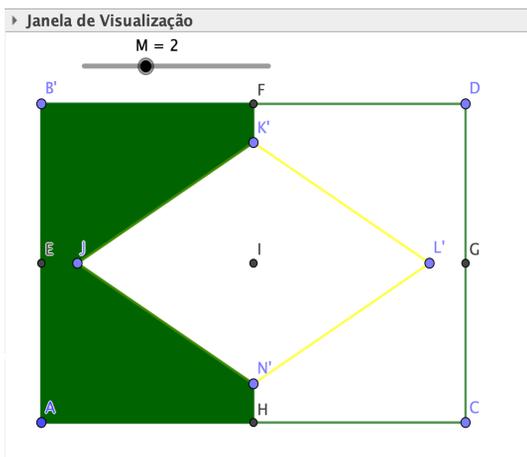


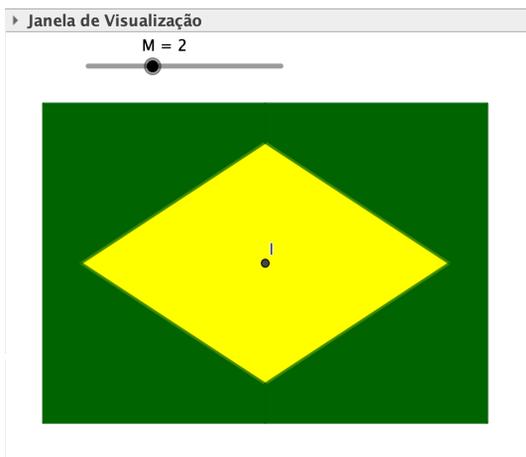
Figura 2: Construção do Losango da Bandeira do Brasil.

Para executar o preenchimento em verde e amarelo, respectivamente, do retângulo e do losango da bandeira, tem-se:

- **Passo 13 - BRA:** Utilize a ferramenta *Polígono*. Como o losango está sobreposto ao retângulo, para colorir a região descoberta, construa os heptágonos  $AHN'JK'FB'$  e  $HCDFK'L'N'$ , e, a partir de *Propriedades*, faça a escolha da cor verde-escuro com *Transparência 100*;
- **Passo 14 - BRA:** Com a mesma ferramenta do passo anterior, construa o losango  $N'L'K'J$  e preencha ele com a cor amarelo e *Transparência 100*.



(a) Durante a execução do Passo 13 - BRA.



(b) Após Passo 14 - BRA.

Figura 3: Retângulo e Losango da Bandeira do Brasil.

O quarto inciso do Art. 5º da legislação da Bandeira Nacional define que

*IV- O círculo azul no meio do losango amarelo terá o raio de três módulos e meio ( $3,5M$ ).*

(BRASIL, 1971, p1)

Com o auxílio da ferramenta *Círculos Dados Centro e Raio* construa um círculo centrado no ponto I com raio  $3,5M$ . Em *Propriedades*, escolha a cor azul com *Transparência* igual a 100%. Por fim, desabilite o *Exibir Rótulo* do círculo construído e o *Exibir Objeto* do ponto I.



Figura 4: Construção Parcial da Bandeira do Brasil.

Os quinto e sexto incisos do Art. 5º da legislação da Bandeira Nacional traz:

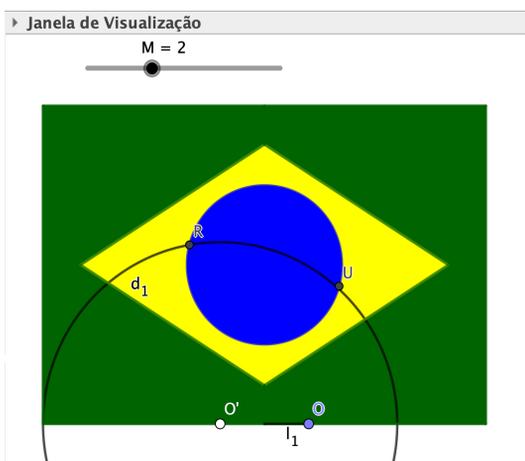
V- O centro dos arcos da faixa branca estará dois módulos ( $2M$ ) à esquerda do ponto do encontro do prolongamento do diâmetro vertical do círculo com a base do quadro externo (ponto  $C$  indicado no Anexo nº 2).

VI- O raio do arco inferior da faixa branca será de oito módulos ( $8M$ ); o raio do arco superior da faixa branca será de oito módulos e meio ( $8,5M$ ).

(BRASIL, 1971, p1)

Para a construção da faixa, esboce os arcos de raios  $8M$ ,  $8,1M$ ,  $8,2M$ ,  $8,3M$ ,  $8,4M$  e  $8,5M$ , com centro no ponto que dista  $2M$  à esquerda do ponto médio do lado inferior do retângulo da bandeira. Por se tratar de rotinas semelhantes, indicamos neste espaço a construção do arco da faixa de  $8M$ .

- **Passo 15 - BRA:** Com a ferramenta *Segmento com Comprimento Fixo*, construa um segmento, de origem o ponto médio do lado inferior do retângulo da bandeira, com comprimento  $2M$ ;
- **Passo 16 - BRA:** Utilize o recurso *Rotação em Torno de um Ponto* e encontre a rotação de  $180$  graus do ponto obtido no passo anterior em torno do ponto médio do segmento;
- **Passo 17 - BRA:** Com *Círculos dados Centro e Raio*, construa um círculo centrado no ponto obtido em Passo 16 e de raio igual a  $8M$ ;
- **Passo 18 - BRA:** Determine a interseção do círculo do Passo 17 com o círculo azul da bandeira. Observe que essa interseção dar-se-á em dois pontos distintos. Em seguida, oculte o círculo do passo anterior, de modo a deixar exibidos o centro e os pontos de interseção encontrados;
- **Passo 19 - BRA:** Com a ferramenta *Arco Circular* construa o arco de centro igual ao da circunferência ocultada no Passo 18 e que passe pelos pontos de interseção. Em *Propriedades* atribua branco ao traço do arco circular. Em *Estilo* coloque a *Opacidade do Traço* igual a  $100$ .



(a) Durante a execução do Passo 18 - BRA.



(b) Após Passo 19 - BRA.

Figura 5: Construção da Faixa da Bandeira do Brasil.

Após a construção de todos os arcos anteriormente mencionados, por meio de raciocínio análogo conforme feito para o de raio  $8M$ , obtém-se a Bandeira Nacional.



Figura 6: Bandeira do Brasil.

☞ Conforme apontado por Giraldo, Caetano e Mattos (2012, p81) é importante explorar não só as potencialidades como também as limitações técnicas do computador. Assim, após o **Passo 19 - BRA**, o professor fica diante de uma dessas limitações, pois dependendo da partição feita na faixa, após o final da construção dos arcos, poderá ser exibida uma faixa com “falhas”, que dependem da escolha de  $M$  e do *zoom* dado na *Janela de Visualização*.

## 2.2. A Bandeira do Estado de Minas Gerais

☹ Para a construção da bandeira de Minas Gerais, com o apoio do GeoGebra, estima-se 1 aula, visto que não serão criadas as letras da bandeira.

☞ Habilidades da Base Nacional Comum Curricular: **EF06MA18**, **EF06MA19**, **EF06MA20**, **EF06MA21**, **EF06MA23**, **EF06MA25** e **EM06MA27**.

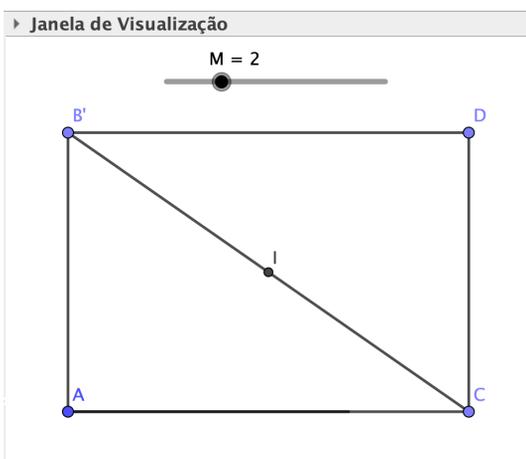
O formato da bandeira do estado de Minas Gerais está descrito no Art. 2º da Lei Estadual de Minas Gerais nº 2.793, de 8 de janeiro de 1963, que traz,

*A Bandeira do Estado de Minas Gerais tem os seguintes desenho e forma: um retângulo em branco com 20 (vinte) módulos de comprimento e 14 (quatorze) módulos de largura; ao centro, um triângulo equilátero em vermelho com 8 (oito) módulos de cada lado, tendo no lado superior esquerdo a palavra ‘LIBERTAS’, no lado superior direito as palavras ‘QUAE SERA’ e na base a palavra ‘TAMEN’, as quais palavras*

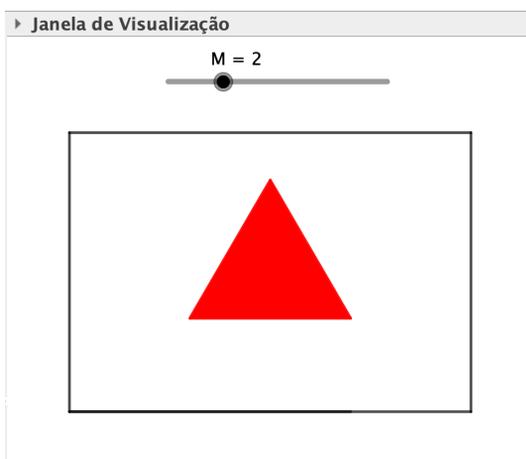
são em TIPO ROMANO, com letras de  $2/3$  de módulo em altura e separadas do triângulo por  $1/3$  do módulo, formando no conjunto a frase “*Libertas quae sera tamen*”, que é a divisa da Inconfidência Mineira. (MINAS GERAIS, 1963, p1)

O retângulo da bandeira do estado de Minas Gerais possui as mesmas dimensões da bandeira do Brasil, assim omitimos a sua construção, pois está nos passos de 1 a 6 da subseção anterior, exceto a cor pois, no contexto, é branco. Desse modo, falta construir no centro do retângulo um triângulo equilátero vermelho com 8 módulos de lado. Para isso:

- **Passo 1 - MG:** Para localizar o centro da bandeira, trace com *Segmento* uma das diagonais do retângulo, e, na sequência, com o recurso *Ponto Médio ou Centro*, aplicado à diagonal criada, para obter o ponto desejado. Em *Propriedades* desabilite *Exibir Objeto* relacionado à diagonal criada. Atente-se de deixar exibido no interior do retângulo apenas o centro do retângulo;
- **Passo 2 - MG:** Com o recurso *Segmento com Comprimento Fixo*, construa um segmento, com origem no centro do retângulo, de comprimento  $\frac{8\sqrt{3}M}{3}$ . Em seguida, com *Rotação em Torno de um Ponto*, gire 90 graus, no sentido anti-horário, o segmento em torno do centro do retângulo. O vértice do segmento obtido, que não é o centro, será o vértice superior do triângulo da bandeira.
- **Passo 3 - MG:** A partir do vértice obtido no Passo 2, com *Segmento com Comprimento Fixo*, crie um segmento de tamanho 8M. Com a ferramenta *Rotação em Torno de um Ponto*, faça duas rotações do segmento criado, no sentido horário, de 60 e 120 graus, de modo a obter os demais vértices do triângulo da bandeira.
- **Passo 4 - MG:** Com *Polígonos*, una os pontos obtidos nos Passos 2 e 3 e crie o triângulo da bandeira. Em *Propriedades* desabilite *Exibir Objeto* dos pontos e segmentos que auxiliaram a construção do triângulo equilátero. Ademais, atribua a cor vermelho com *Transparência* 100.



(a) Durante a execução do Passo 1 - MG.



(b) Após Passo 4 - MG.

Figura 7: Construção da Bandeira de Minas Gerais.

### 2.3. A Bandeira da Cidade de Uberlândia

☉ Para a construção da bandeira de Uberlândia, com o apoio do GeoGebra, estimam-se 4 aulas.

O desenho e a forma da bandeira do município de Uberlândia estão no memorial descritivo do anexo da Lei Municipal de Uberlândia nº 3.191, de 22 de setembro de 1980.

☞ Habilidades da Base Nacional Comum Curricular: **EF06MA16, EF06MA18, EF06MA19, EF06MA21, EF06MA22, EF06MA25 e EF06MA27.**

O memorial descritivo traz os significados das cores e dos elementos geométricos da bandeira. Em especial, os primeiro e segundo parágrafos explicitam que

*A Bandeira do Município de Uberlândia tem as dimensões de 1,80m × 1,20m.*

*A dimensão maior, de 1,80m de comprimento da bandeira, é dividida em duas partes iguais, de 90cm cada, no sentido vertical. A outra dimensão, ou seja, a menor, de 1,20m é, também, dividida em duas partes proporcionais, desiguais, no sentido horizontal, sendo uma correspondente a 2/3 da dimensão, 80 cm, parte superior, e a outra, 1/3 da dimensão, 40cm, parte inferior; a parte de cima, a maior porção, tem a cor azul e representa o nosso céu, a parte de baixo, a porção menor, tem a cor verde e representa o nosso campo, nossa terra, o Município de Uberlândia. (UBERLÂNDIA, 1980, p1)*

Em virtude do modo de explicação da bandeira de Uberlândia, optou-se por sua construção no plano cartesiano, com a inserção de pontos por meio de suas coordenadas. Nesse sentido, na *Janela de Visualização* do GeoGebra devem estar os *Eixos* e a *Malha*. Adotou-se que a bandeira seria construída no primeiro quadrante do plano coordenado, com dois de seus lados sobre os eixos  $x$  e  $y$ .

☞ Os pontos foram inseridos no GeoGebra via digitação na *Entrada* de suas coordenadas.

- **Passo 1 - UDI:** Marque os pontos A(0,0), B(18,0), C(18,12) e D(0,12) que consistirão dos vértices da bandeira. Em *Propriedades*, de cada ponto, habilite *Fixar Objeto*;
- **Passo 2 - UDI:** Com a ferramenta *Segmento* construa os lados da bandeira, a saber: AB, BC, CD e AD. Em *Propriedades*, de cada segmento, desabilite a função *Exibir Rótulo*;
- **Passo 3 - UDI:** Marque os pontos E(9,0) e F(9,12) que serão suporte para a reta que dividirá verticalmente a dimensão maior da bandeira. Em *Propriedades*, de E e F, habilite *Fixar Objeto*;
- **Passo 4 - UDI:** Com a ferramenta *Reta* construa a reta que passe pelos pontos E e F. Em *Propriedades* desabilite *Exibir Rótulo*;
- **Passo 5 - UDI:** Marque os pontos G(0,4) e H(18,4) que definirão a reta que dividirá o retângulo na proporção de  $\frac{2}{3}$  da parte superior e  $\frac{1}{3}$  da parte inferior. Em *Propriedades*, de G e H, habilite *Fixar Objeto*;

- **Passo 6 - UDI:** Com a ferramenta *Reta* construa a reta definida pelos pontos G e H. Em *Propriedades* desabilite a função *Exibir Rótulo*;
- **Passo 7 - UDI:** Com a ferramenta *Polígono* construa o quadrilátero GHCD. Em *Propriedades*, atribua a cor azul com *Transparência* 100 ao polígono. Em *Propriedades* desabilite a função *Exibir Rótulo* do quadrilátero e seus lados;
- **Passo 8 - UDI:** Com a ferramenta *Polígono* construa o quadrilátero ABHG. Em *Propriedades*, atribua a cor verde com *Transparência* 100 ao polígono. Em *Propriedades* desabilite a função *Exibir Rótulo* do quadrilátero e seus lados.

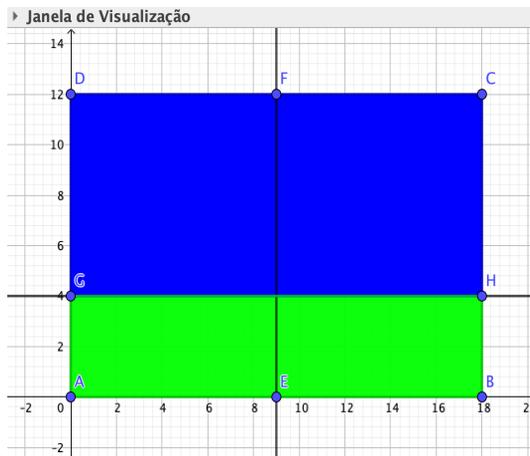


Figura 8: Retângulo da Bandeira de Uberlândia.

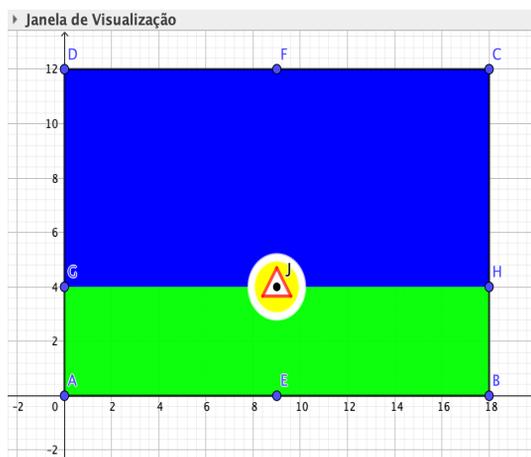
Na sequência do memorial descritivo da legislação da bandeira de Uberlândia tem-se

*O centro ideal da bandeira é encontrado pelo cruzamento das linhas vertical com a horizontal. Determinado o centro ideal da bandeira têm-se, também, o centro do triângulo isósceles de altura igual a 10,40cm e 12cm de lado, sendo a sua base paralela à maior dimensão, distância de 3,4cm; o triângulo é de cor branca e em seus lados existe uma faixa de 0,8cm, de cor vermelha.*

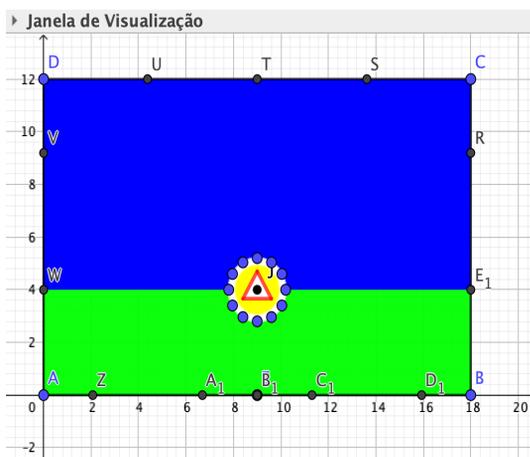
*Do centro ideal da bandeira e do triângulo isósceles partem 12 eixos para formação de uma estrela de 12 pontas, de cor branca, e que representa o sistema viário, as vias de comunicação para todo o Brasil e que determinam o progresso de Uberlândia. O Triângulo lembra a região e o Estado de Minas Gerais.*

*Ainda, base no centro ideal da bandeira e do triângulo isósceles, traçam-se duas circunferências com raios de 9 cm e 12 cm, uma circunscrevendo o triângulo e a outra que determinará a origem das 12 pontas da estrela; a determinação das pontas das estrelas se faz dividindo a circunferência maior, em 12 partes iguais de 30° cada. A circunferência menor de cor amarelo representa as riquezas do município. (UBERLÂNDIA, 1963, p1)*

- **Passo 9 - UDI:** Com *Interseção de Dois Objetos* encontre a interseção das retas construídas nos Passos 4 e 6. Em *Propriedades* desabilite *Exibir Objeto* das retas envolvidas nesse passo e deixe apenas o ponto de interseção. Para esse ponto atribua, em *Propriedades*, a *Camada* 8;
- **Passo 10 - UDI:** Com *Círculo dados Centro e Raio*, crie dois círculos, concêntricos com centro no ponto obtido no passo anterior, com raios iguais a 1.2 e 0.9 (nesta ordem). Em *Propriedades* atribua a cor branco com *Transparência* 100 para o círculo de maior raio e desabilite o rótulo. Em relação ao outro círculo, atribua a cor amarelo com *Transparência* 100 e desabilite o rótulo;
- **Passo 11 - UDI:** Como a base do triângulo isósceles tem 12cm e está distante 3,4cm do centro ideal da bandeira, marque os pontos  $K(9, 3.66)$ ,  $L(9.6, 3.66)$  e  $M(8.4, 3.66)$ ;
- **Passo 12 - UDI:** Como a altura do triângulo isósceles é de 10,5cm, marque o ponto  $N(9, 4.7)$ ;
- **Passo 13 - UDI:** Com *Polígonos*, esboce o triângulo MLN. Em *Propriedades* atribua a cor branco, *Transparência* 100 e *Camada* 9. E ainda, desabilite os rótulos dos elementos do triângulo;
- **Passo 14 - UDI:** Para ter a faixa vermelha que contorna o triângulo, com a ferramenta *Segmentos* obtenha os segmentos ML, LN e MN. Em *Propriedades* atribua, para cada um dos segmentos, a cor vermelho e a *Espessura da Linha* 7. Ademais, desabilite os rótulos dos segmentos;
- **Passo 15 - UDI:** Desabilite a opção *Exibir Objeto* nas propriedades dos pontos K, L, M e N;
- **Passo 16 - UDI:** Marque o ponto  $P(10.2, 4)$ . Os demais pontos que nortearão na construção das pontas são obtidos pela rotação sucessiva, em 30 graus, no sentido anti-horário, do ponto P em relação ao ponto J (centro da circunferência), com o recurso *Rotação em Torno de um Ponto*;
- **Passo 17 - UDI:** Crie, com *Semirreta*, as 12 semirretas com origem em J e que passem em cada um dos pontos criados no passo anterior. Os inícios das pontas da estrela consistirão da interseção da semirreta criada com a borda da bandeira. Encontre tais pontos com o recurso *Interseção de Dois Objetos*. Em *Propriedades*, desabilite *Exibir Objeto* das semirretas criadas no passo anterior e *Exibir Rótulos* dos pontos que consistem as pontas da estrela da bandeira.



(a) Após Passo 15 - UDI.



(b) Após Passo 17 - UDI.

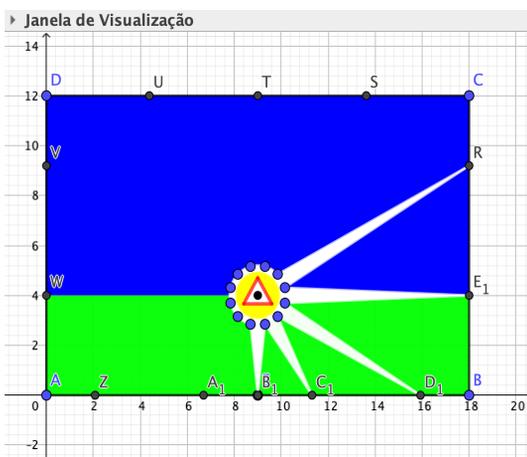
Figura 9: Construção da Bandeira de Uberlândia.

Do memorial descritivo da bandeira de Uberlândia tem-se ainda que

*As pontas das estrelas, em número de 12, partem da linha que divide a maior dimensão em duas partes, no sentido horário, formando entre eixos das pontas um ângulo de  $30^\circ$  e com os seguintes comprimentos,  $0,80 - 0,92 - 1,04 - 0,90 - 0,80 - 0,46 - 0,40 - 0,46 - 0,80 - 0,90 - 1,04 - 0,92$ , com o seu interior de branco. As pontas das estrelas, em cor branco, representam, além do sistema viário, a paz, a tranquilidade, a ordem que reinam em Uberlândia. (UBERLÂNDIA, 1963, p1)*

Assim, para finalizar a construção da estrela da bandeira, segue um roteiro de rotinas no GeoGebra:

- **Passo 18 - UDI:** Em *Propriedades* desabilite *Exibir Objeto* dos pontos criados na circunferência no Passo 16, com exceção do ponto P;
- **Passo 19 - UDI:** Com *Rotação em Torno de um Ponto*, faça a rotação do ponto P, de 15 graus, no sentido anti-horário, em torno do centro da circunferência. Desabilite a opção *Exibir Objeto* do Ponto P.
- **Passo 20 - UDI:** A partir do ponto obtido no passo anterior, faça sucessivas rotações de 30 graus, no sentido anti-horário, em torno do centro da circunferência. Desabilite a opção *Exibir Rótulo* dos pontos criados;
- **Passo 21 - UDI:** Com a ferramenta *Polígono*, crie triângulos cujos vértices são um ponto criado no Passo 17 e os dois pontos mais próximos do Passo 20 desse. Em *Propriedades* atribua a cor branco com *Transparência* 100 e desabilite o rótulo dos elementos envolvidos ao triângulo criado.



(a) Durante a Execução do Passo 21 - UDI.



(b) Após Passo 21 - UDI.

Figura 10: Bandeira de Uberlândia.

### 3. Considerações Finais

A experiência interdisciplinar exigiu dos professores envolvidos um planejamento conjunto das atividades desenvolvidas nas disciplinas de Artes, História, Língua Portuguesa e Matemática relacionadas às bandeiras do Brasil, de Minas Gerais e de Uberlândia.

No que tange à disciplina de Matemática, ao final das construções geométricas das referidas bandeiras, observou-se nos alunos: aumento no nível de curiosidade e motivação quanto à aula de matemática; melhoria na relação e interação com o professor, devido à organização diferenciada da aula que ocorria no laboratório de informática da escola; melhoria na aprendizagem dos conteúdos matemáticos envolvidos nas construções.

Em relação ao grupo de professores, a experiência potencializou a prática do diálogo e da partilha de informações para o desenvolvimento de um trabalho colaborativo, o que auxiliou na resolução de situações de conflito próprios do ambiente escolar.

Por fim, cabe destacar que atividades de construções geométricas de bandeiras podem ser inseridas nas aulas de matemática, e essas terão novas nuances e desafios a depender das escolhas de bandeiras realizadas pelo professor ou alunos. Em nosso contexto, com alunos do sexto ano do Ensino Fundamental, constitui-se um desafio a utilização de um número irracional no Passo 2 da construção da bandeira de Minas Gerais, visto que a construção deste conceito dar-se-á pela BNCC em habilidades vivenciadas a partir do oitavo ano do Ensino Fundamental. Adotou-se um comando do GeoGebra para o processo de construção do triângulo, mas para séries mais avançadas uma discussão mais aprofundada do número envolvido precisaria ser realizada em sala de aula pelo professor de matemática.

### Agradecimentos

Os autores agradecem o(a) parecerista pelas contribuições apresentadas ao trabalho, à Secretaria Municipal de Educação de Uberlândia/MG e ao Programa de Mestrado Profissional em Matemática em Rede Nacional - Profmat da Universidade Federal de Goiás - Regional Catalão.

### Referências

- [1] BERGER, G. “*Conditions d’une problématique de l’interdisciplinarité*”. In: CERI (Ed). *L’Interdisciplinarité: problèmes d’enseignement et de recherche dans les Universités*. Paris: UNESCO/OCDE, 1972. pp21-24, apud POMBO, O. Contribuição para um vocabulário sobre interdisciplinaridade. In: POMBO, O.; GUIMARÃES, H.; LEVY, T. Interdisciplinaridade: reflexão e experiência. 2ª ed. Lisboa: Texto, 1994.
- [2] BRASIL. *Lei nº 5.700*, de 01 de Setembro de 1971. Dispõe sobre a forma e a apresentação dos Símbolos Nacionais, e dá outras providências. Brasília, 02 set. 1971. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/l5700.htm](http://www.planalto.gov.br/ccivil_03/leis/l5700.htm)>. Acesso em: 20 de abril de 2020.
- [3] GIRALDO, V., CAETANO, P. e MATTOS, F. *Recursos Computacionais no Ensino de Matemática*. Rio de Janeiro: Editora SBM. (2012), 415p. (Coleção Profmat)
- [4] LORENZATO, Sérgio. *Por que não Ensinar Geometria? A Educação Matemática em Revista*. Recife, v. 3, n. 4, pp3-13, 1995.
- [5] MINAS GERAIS. *Lei nº 2.793*, de 08 de Janeiro de 1963. Institui a bandeira de Minas Gerais. Belo Horizonte, 08 jan. 1963. Disponível em: <<https://leisestaduais.com.br/mg/>>

[lei-ordinaria-n-2793-1963-minas-gerais-institui-a-bandeira-do-estado-de-minas-gerais](#)>. Acesso em: 20 de abril de 2020.

- [6] MIORIM, M. A.; FIORENTINI, D. *Uma reflexão sobre o uso de materiais concretos e jogos no Ensino da Matemática*. Boletim da SBEM-SP, São Paulo, v. 4, nº 7, pp5-10, 1990.
- [7] UBERLÂNDIA. *Lei nº 3.191, de 22 de Setembro de 1980*. Cria a bandeira do município de Uberlândia. Uberlândia, 22 set. 1980. Disponível em: <<https://leismunicipais.com.br/a/mg/u/uberlandia/lei-ordinaria/1980/320/3191/lei-ordinaria-n-3191-1980-cria-bandeira-do-municipio-de-uberlandia?q=3191>>. Acesso em: 20 de abril de 2020.

Inês Naves Cunha de Oliveira  
Escola Municipal Professor Otávio Batista Coelho Filho  
<[inesncoliveira@outlook.com](mailto:inesncoliveira@outlook.com)>

Thiago Porto de Almeida Freitas  
Universidade Federal de Goiás  
<[tportto@ufg.br](mailto:tportto@ufg.br)>

Recebido: 20/04/2020  
Publicado: 02/07/2020

# Quadrados e polinômios do segundo grau com o uso do Algeplan

Karina Rodriguez de Fraga 

Leandro Sebben Bellicanta 

## Resumo

Neste artigo é apresentada uma proposta de atividade didática utilizando o material concreto chamado Algeplan. Mais especificamente, o artigo explora a relação existente entre os coeficientes de alguns polinômios do segundo grau em duas variáveis e a quantidade de peças do Algeplan necessárias para se construir um quadrado. No desenvolvimento dessa atividade, os estudantes poderão descobrir uma maneira sistemática de contar e saber quais são os quadrados que podem ser construídos com as peças do Algeplan.

**Palavras-chave:** matemática, ensino, algeplan, material didático-pedagógico, álgebra.

## Abstract

This paper presents a proposal for didactic activity using the concrete material called Algeplan. More specifically, the paper explores the relationship between the coefficients of some second-degree polynomials in two variables and the amount of Algeplan pieces needed to construct a square. In developing this activity students will be able to discover a systematic way to count and to know which squares can be constructed with Algeplan tiles.

**Keywords:** mathematics, teaching, algeplan, didactic-pedagogical material, algebra.

## 1. Introdução

Diversas pesquisas têm mostrado que o uso de materiais concretos no estudo de matemática tem um efeito potencializador no aprendizado e ajudam a estabelecer as bases para as fases subsequentes de abstração envolvidas no desenvolvimento do saber matemático do estudante [1].

Neste trabalho examina-se uma atividade prática, de fácil implementação, utilizando o bem conhecido material didático chamado Algeplan, que é constituído por peças retangulares de diferentes formatos e cores. O Algeplan usado neste trabalho está ilustrado na Figura 1 e possui um total 40 peças de madeira contendo:

- Quatro peças quadradas de cor amarela com lados medindo um certo valor que, neste trabalho, será denotado por  $x$ .
- Quatro peças quadradas de cor azul com lados medindo um certo valor denotado por  $y$ . Observamos que os quadrados azuis possuem lado menor que os quadrados amarelos, de forma que  $y < x$ .

- Doze quadrados de cor vermelha com lados medindo menos que  $y$ . Os lados desses quadrados serão tomados como a unidade de comprimento, ou seja, os lados serão denotados por 1.
- Oito peças retangulares de cor rosa de lados medindo  $y$  e 1.
- Oito peças retangulares de cor laranja de lados medindo  $x$  e 1.
- Quatro peças retangulares de cor verde de lados medindo  $x$  e  $y$ .

Figura 1: Peças do Algeplan



O Algeplan é um material particularmente adaptado ao estudo de polinômios do segundo grau pois, ao convencionarmos que a medida dos lados das peças são denotadas por  $x$ ,  $y$  e 1, a área de cada uma dessas peças possui uma expressão algébrica que coincide com os monômios presentes em um polinômio de segundo grau nas variáveis  $x$  e  $y$ . Vários artigos e trabalhos acadêmicos exploram essa relação entre as peças do Algeplan e polinômios, conforme pode ser visto, por exemplo, em [2] e [3].

A Tabela 1 apresenta todas as peças do Algeplan com suas respectivas representações algébricas.

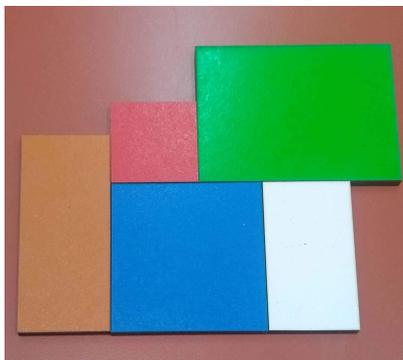
Tabela 1: Representações algébricas das peças do Algeplan

Quantidade de peças	Forma	Cor das peças	Representação algébrica
4	quadrado	amarelas	$x^2$
4	quadrado	azuis	$y^2$
4	retângulo	verdes	$xy$
8	retângulo	laranjas	$x$
8	retângulo	rosa	$y$
12	quadrado	vermelhas	1

Dessa forma, pode-se associar a uma justaposição de peças do Algeplan uma expressão algébrica polinomial de grau dois, como mostram os seguintes exemplos:

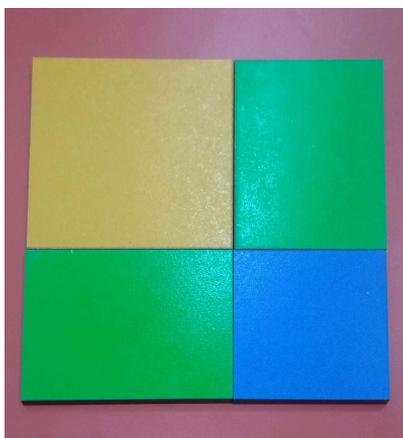
**Exemplo 1.** O polinômio  $y^2 + xy + y + x + 1$  pode ser representado no Algeplan justapondo-se uma peça azul ( $y^2$ ), uma peça verde ( $xy$ ), uma peça rosa ( $y$ ), uma peça laranja ( $x$ ) e uma peça vermelha (1) que correspondem aos termos que estão sendo somados, conforme pode ser visualizado na Figura 2.

Figura 2: Representação, com as peças do Algeplan, do polinômio  $y^2 + xy + y + x + 1$  citado no exemplo 1



**Exemplo 2.** O polinômio  $x^2 + 2xy + y^2$  pode ser representado no Algeplan justapondo-se uma peça amarela ( $x^2$ ), duas peças verdes ( $xy$ ) e uma peça azul ( $y^2$ ), conforme pode ser visualizado na Figura 3.

Figura 3: Representação, com as peças do Algeplan, do polinômio  $x^2 + 2xy + y^2$  citado no exemplo 2



Tal associação entre a justaposição de peças do Algeplan com polinômios de grau dois permite-nos usar um material concreto no estudo dos polinômios, o que constitui um valioso recurso didático.

## 2. O problema a ser abordado nesta atividade

Ao manipularmos o Algeplan, algumas justaposições são especialmente atraentes pela regularidade alcançada na reunião das peças, como é o caso do quadrado que apareceu no exemplo 2.

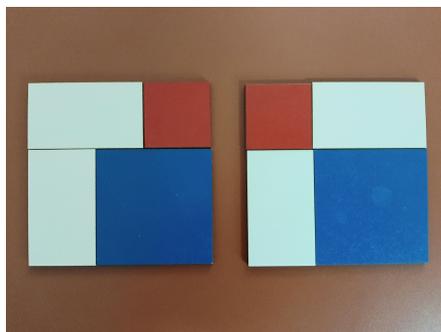
Uma questão de enunciado fácil que pode despertar a curiosidade dos alunos que estejam usando o Algeplan em alguma atividade é a seguinte:

**Questão 1.** Quantos quadrados diferentes é possível montar usando as peças do Algeplan ?

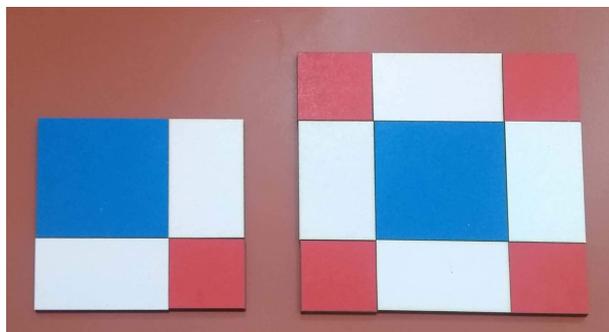
Para facilitar a resolução da questão 1, sugere-se que o significado da expressão “diferentes” seja interpretado como “não congruentes” de forma que os dois quadrados mostrados na Figura 4a serão contados apenas uma

vez, pois são congruentes entre si, já que ambos possuem lados com medida  $y + 1$ . Dessa forma, os quadrados diferentes a que se refere a questão acima serão apenas os que possuem lados distintos, conforme os ilustrados na Figura 4b.

Figura 4: Quadrados usando o Algeplan.



(a) Quadrados congruentes serão considerados iguais e contados apenas uma vez.



(b) Dois quadrados distintos montados com as peças do Algeplan.

### 3. Quadrados geométricos e polinômios quadrados perfeitos.

Para responder de forma definitiva à questão 1, pode-se usar da relação que existe entre as justaposições de peças do Algeplan com os polinômios de grau 2.

Afim de tornar mais clara e precisa essa relação entre peças do material concreto com os polinômios de grau dois, este trabalho usa os termos *quadrados geométricos* e *polinômios quadrados perfeitos* conforme as definições que seguem.

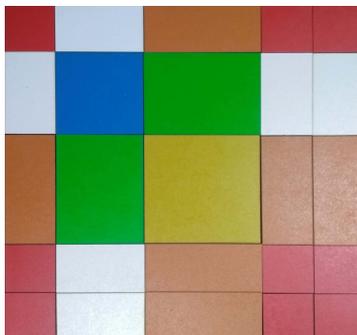
**Definição 1** (Quadrado Geométrico). Um quadrado formado pela justaposição de uma ou mais peças do Algeplan será chamado de *quadrado geométrico*.

As figuras 4 e 5 mostram exemplos de quadrados geométricos.

Devido ao fato de as peças do Algeplan medirem  $x$ ,  $y$  ou  $1$ , o lado de um quadrado geométrico terá como medida possível  $\alpha x + \beta y + \gamma$ , onde  $\alpha$  representa a quantidade de peças com lado  $x$ ,  $\beta$  representa a quantidade de peças com lado  $y$ , e  $\gamma$  representa a quantidade de peças com lado  $1$  que compõe o lado desse quadrado.

Por exemplo, o quadrado da figura 5 possui o comprimento do lado igual a  $x + y + 3$ , pois um de seus lados é formado pela união dos lados de: uma peça rosa com medida  $y$ ; uma peça laranja com medida  $x$  e três peças vermelhas com medidas iguais a  $1$ .

Figura 5: Quadrado de lado  $x + y + 3$



Note que a medida da área do quadrado da Figura 5 é  $(x + y + 3)^2$  e, usando-se as regras usuais da multiplicação, tem-se a seguinte relação:

$$(x + y + 3)^2 = x^2 + y^2 + 2xy + 6x + 6y + 9 \tag{1}$$

Observe que os coeficientes do polinômio do lado direito da Equação 1 indicam exatamente quantas peças de cada tipo foram utilizadas na construção do quadrado. É precisamente essa relação entre o quadrado (modelado algebricamente pelo lado esquerdo da equação 1) e os coeficientes da sua forma não fatorada (lado esquerdo da equação 1) que será usada para responder definitivamente à questão 1.

A Tabela 2 destaca a relação entre a quantidade de peças e os coeficientes do polinômio representado pela Figura 5.

Tabela 2: Relação entre os termos do polinômio 1 e a quantidade de peças.

Termo do polinômio	Quantidade de peças
$x^2$	uma peça amarela
$y^2$	uma peça azul
$2xy$	duas peças verdes
$6x$	seis peças laranjas
$6y$	seis peças rosa
9	nove peças vermelhas

A fim de tornar mais precisa a relação entre quadrados geométricos e polinômios, faremos uso da seguinte definição:

**Definição 2** (Polinômio Quadrado Perfeito). Um polinômio de segundo grau nas variáveis  $x$  e  $y$  do tipo  $Ax^2 + By^2 + Cxy + Dx + Ey + F$  é um *polinômio quadrado perfeito* quando existe  $\alpha$ ,  $\beta$  e  $\gamma$  tal que

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = (\alpha x + \beta y + \gamma)^2. \tag{2}$$

De acordo com a definição 2, um quadrado geométrico cujo lado mede  $\alpha x + \beta y + \gamma$  está associado a um polinômio quadrado perfeito de forma que, para contar quantos quadrados é possível construir usando as peças do Algeplan, pode-se contar quantos polinômios quadrados perfeitos consegue-se construir levando em conta que, no caso do Algeplan, os coeficientes  $\alpha$ ,  $\beta$  e  $\gamma$  são todos inteiros não negativos e estão sujeitos às restrições impostas pela quantidade limitada de peças.

Para explicitar quais são as relações entre a quantidade de peças do Algeplan e os coeficientes  $\alpha$ ,  $\beta$  e  $\gamma$ , pode-se expandir o lado direito da equação 2 e estudar a seguinte igualdade:

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = \alpha^2x^2 + \beta^2y^2 + 2\beta\alpha xy + 2\alpha\gamma x + 2\beta\gamma y + \gamma^2. \quad (3)$$

A equação 3 evidencia que os coeficientes A, B, C, D, E e F de um polinômio quadrado perfeito devem satisfazer as seguintes equações:

$$A = \alpha^2 \quad (4) \qquad C = 2\beta\alpha \quad (6) \qquad E = 2\beta\gamma \quad (8)$$

$$B = \beta^2 \quad (5) \qquad D = 2\alpha\gamma \quad (7) \qquad F = \gamma^2 \quad (9)$$

para  $\alpha$ ,  $\beta$  e  $\gamma$  convenientes.

Por exemplo, o polinômio  $4x^2 + xy + 1$  possui  $A = 4, B = 0, C = 1, D = 0, E = 0$  e  $F = 1$  e não é um quadrado perfeito pois as equações 4, 5 e 6 são incompatíveis. Já o polinômio  $x^2 + y^2 + 2xy + 6x + 6y + 9$  possui  $A = 1, B = 1, C = 2, D = 6, E = 6$  e  $F = 9$  e é um quadrado perfeito já que com  $\alpha = 1, \beta = 1$  e  $\gamma = 3$  as igualdades 4, 5, 6, 7, 8, 9 são todas verdadeiras.

#### 4. Solução da questão 1

Para se descobrir qual é o número máximo de quadrados que se pode formar com as peças, faremos uso da Tabela 3, na qual a primeira coluna contém os coeficientes dos termos de um polinômio quadrado perfeito genérico; na segunda coluna estão representadas as restrições que esses termos devem satisfazer de acordo com a quantidade de peças de cada tipo que compõe o Algeplan usado neste trabalho; na terceira e última coluna tem-se os valores (inteiros não negativos) de  $\alpha$ ,  $\beta$  e  $\gamma$  que podem ser aplicados satisfazendo as restrições da segunda coluna.

Tabela 3: Valores de  $\alpha$ ,  $\beta$  e  $\gamma$  - condição de existência

Coeficientes do polinômio.	Restrições devido a quantidade de peças.	Valores possíveis para $\alpha$ , $\beta$ e $\gamma$ .
$\alpha^2$	$0 \leq \alpha^2 \leq 4$	$0 \leq \alpha \leq 2$
$\beta^2$	$0 \leq \beta^2 \leq 4$	$0 \leq \beta \leq 2$
$\gamma^2$	$0 \leq \gamma^2 \leq 12$	$0 \leq \gamma \leq 3$
$2\beta\alpha$	$0 \leq 2\beta\alpha \leq 4$	$0 \leq \beta\alpha \leq 2$
$2\alpha\gamma$	$0 \leq 2\alpha\gamma \leq 8$	$0 \leq \alpha\gamma \leq 4$
$2\beta\gamma$	$0 \leq 2\beta\gamma \leq 8$	$0 \leq \beta\gamma \leq 4$

Analisando as três primeiras linhas da Tabela 3 pode-se observar que os valores de  $\alpha$ ,  $\beta$  e  $\gamma$  ficam restritos a, no máximo, 3 possibilidades para  $\alpha$ , 3 para  $\beta$  e 4 para  $\gamma$ . Sendo assim, as três primeiras linhas delimitam  $3 \times 3 \times 4 = 36$  casos a serem considerados. No entanto, 8 desses casos não satisfazem as restrições explicitadas nas três últimas linhas da Tabela 3. Por exemplo,  $\alpha = 1, \beta = 2$  e  $\gamma = 3$  não satisfaz a última linha da tabela já que  $\beta\gamma = 6 > 4$ , o que significa que faltariam peças para construir um quadrado com lado igual a  $x + 2y + 3$ . São portanto  $36 - 8 = 28$  os quadrados possíveis de se construir com as peças da Tabela 1.

A seguir, pode-se ver na Tabela 4, quais são os quadrados que podem ser construídos, bem como sua representação algébrica.

Tabela 4: Tabela com todas possibilidades de quadrados construtíveis com as 40 peças do Algeplan

$\alpha$	$\beta$	$\gamma$	Expressão Algébrica	Imagem
0	0	0	0	
		1	1	6a
		2	4	6b
		3	9	6c
	1	0	$y^2$	7a
		1	$y^2 + 2y + 1$	7b
		2	$y^2 + 4y + 4$	7c
		3	$y^2 + 6y + 9$	7d
	2	0	$4y^2$	8a
		1	$4y^2 + 4y + 1$	8b
		2	$4y^2 + 8y + 4$	8c
		3	Seriam necessárias 12 peças do tipo y.	
1	0	0	$x^2$	9a
		1	$x^2 + 2x + 1$	9b
		2	$x^2 + 4x + 4$	9c
		3	$x^2 + 6x + 9$	9d
	1	0	$x^2 + 2xy + y^2$	10a
		1	$x^2 + 2xy + y^2 + 2x + 2y + 1$	10b
		2	$x^2 + 2xy + y^2 + 4x + 4y + 4$	10c
		3	$x^2 + 2xy + y^2 + 6x + 6y + 9$	10d
	2	0	$x^2 + 4xy + 4y^2$	11a
		1	$x^2 + 4xy + 4y^2 + 2x + 4y + 1$	11b
		2	$x^2 + 4xy + 4y^2 + 4x + 8y + 4$	11c
		3	Seriam necessárias 12 peças do tipo y.	
2	0	0	$4x^2$	12a
		1	$4x^2 + 4x + 1$	12b
		2	$4x^2 + 8x + 4$	12c
		3	Seriam necessárias 12 peças do tipo x.	
	1	0	$4x^2 + 4xy + y^2$	13a
		1	$4x^2 + 4xy + y^2 + 4x + 2y + 1$	13b
		2	$4x^2 + 4xy + y^2 + 8x + 4y + 4$	13c
		3	Seriam necessárias 12 peças do tipo x.	
	2	0	Seriam necessárias 8 peças do tipo xy.	
		1	Seriam necessárias 8 peças do tipo xy.	
		2	Seriam necessárias 8 peças do tipo xy.	
		3	Seriam necessárias 12 peças do tipo x, 12 do tipo y e 8 do tipo xy.	

Figura 6: Quadrados perfeitos com  $\alpha = 0; \beta = 0$

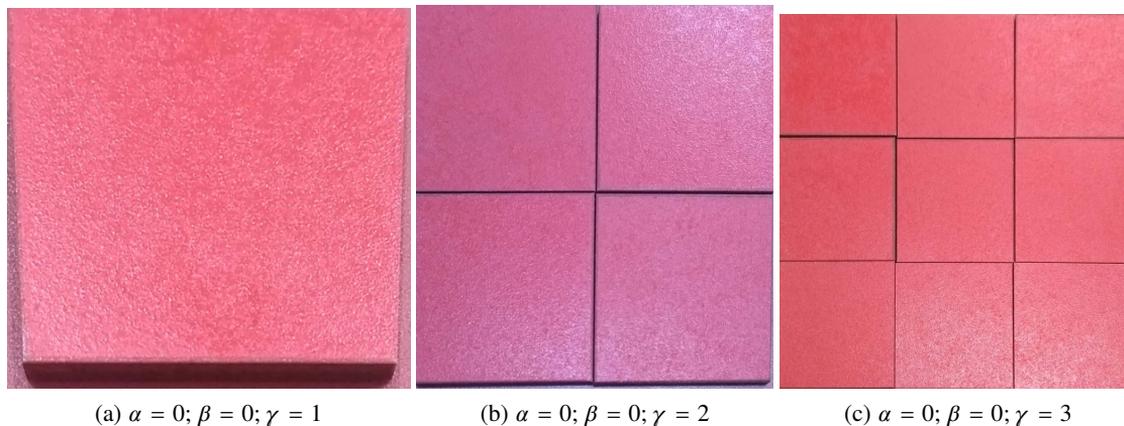


Figura 7: Quadrados perfeitos com  $\alpha = 0; \beta = 1$

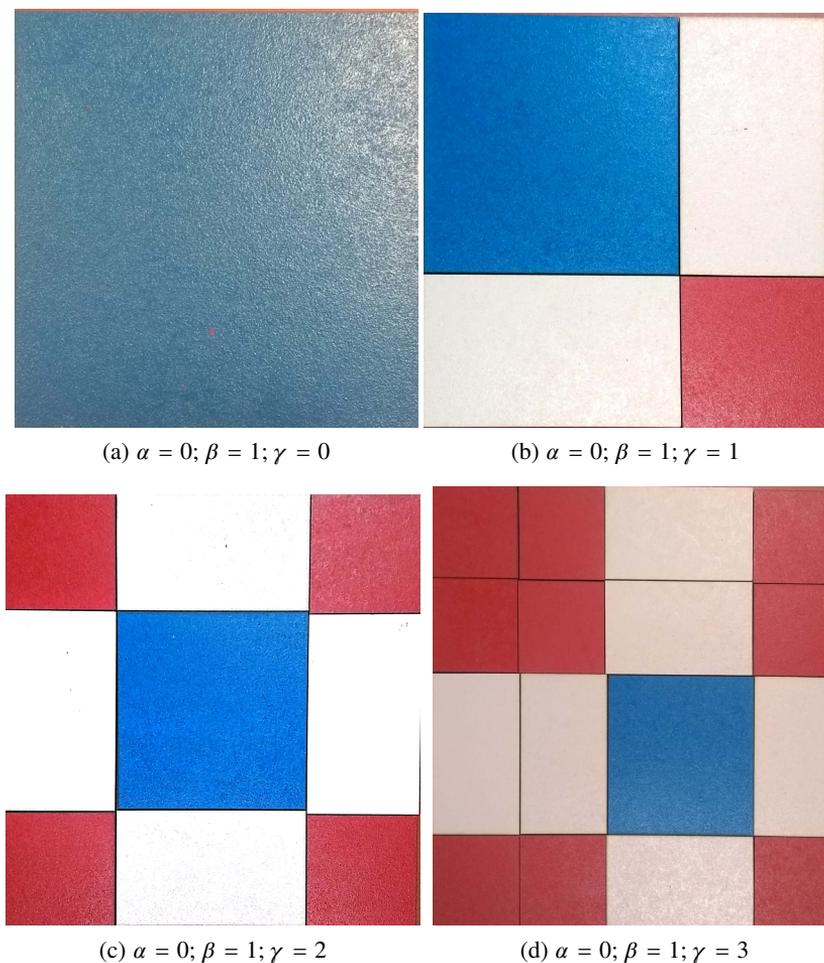


Figura 8: Quadrados perfeitos com  $\alpha = 0; \beta = 2$

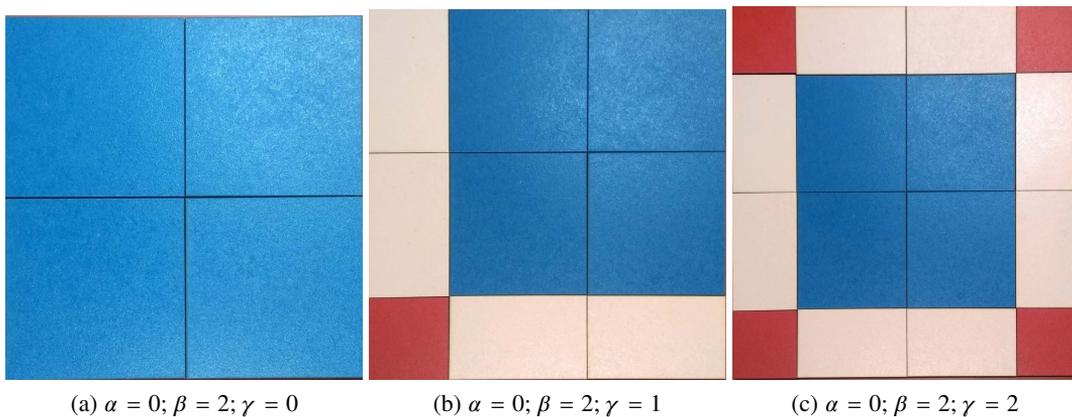


Figura 9: Quadrados perfeitos com  $\alpha = 1; \beta = 0$

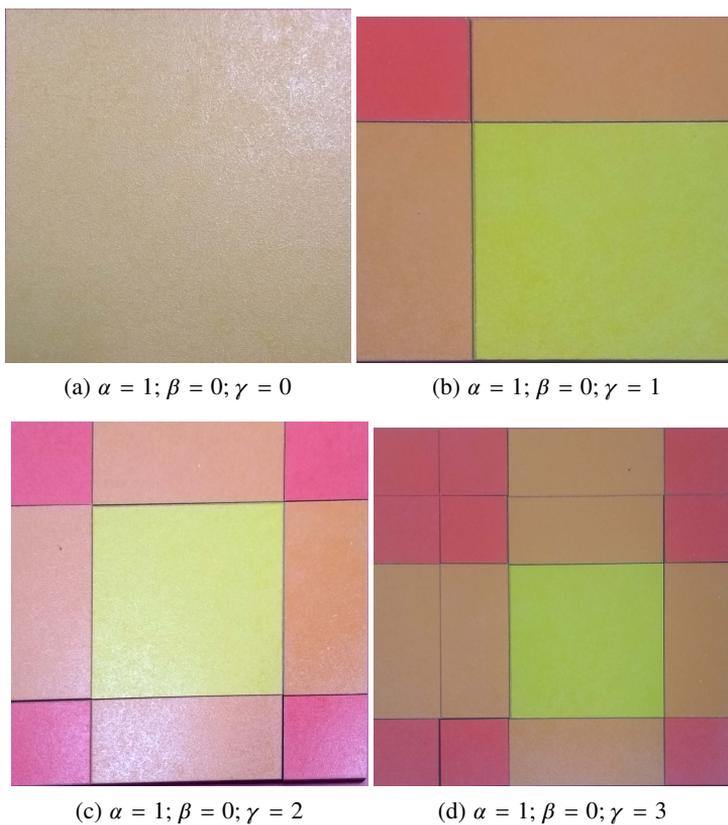
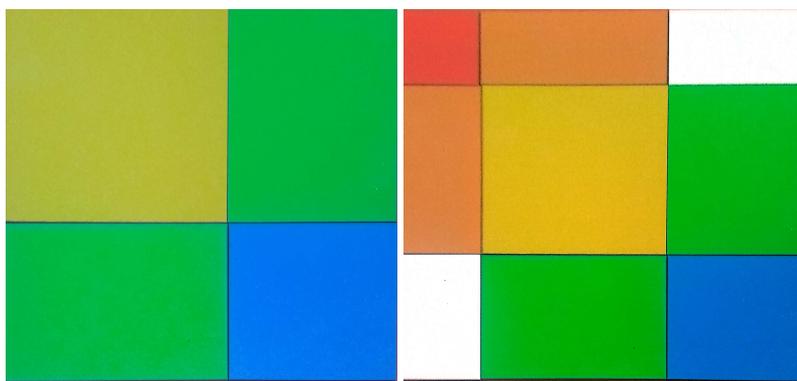
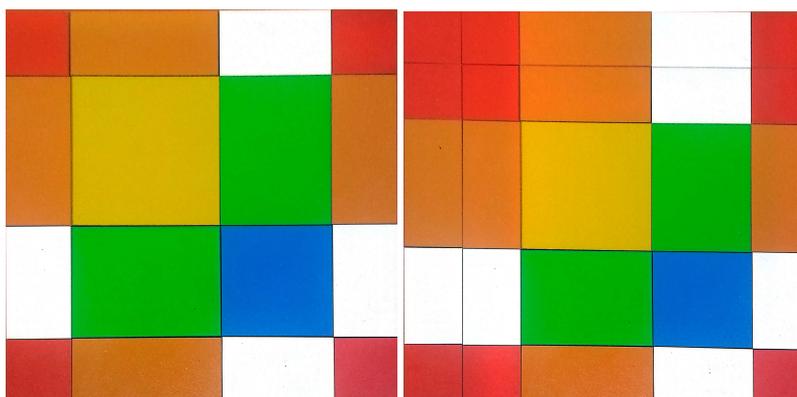


Figura 10: Quadrados perfeitos com  $\alpha = 1; \beta = 1$



(a)  $\alpha = 1; \beta = 1; \gamma = 0$

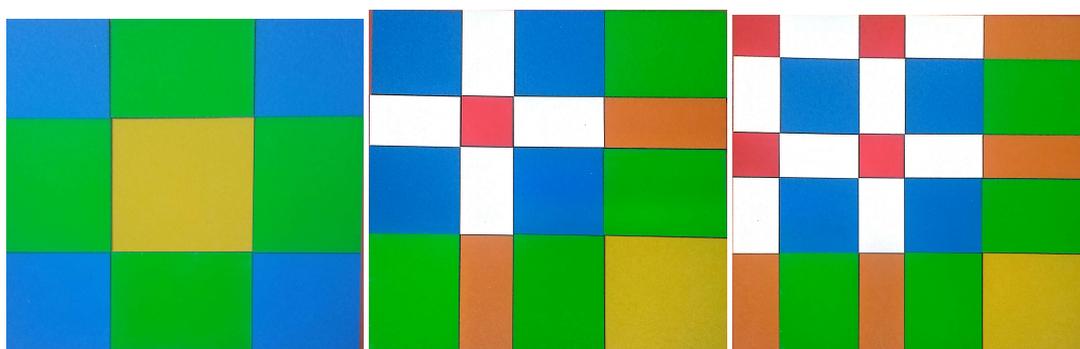
(b)  $\alpha = 1; \beta = 1; \gamma = 1$



(c)  $\alpha = 1; \beta = 1; \gamma = 2$

(d)  $\alpha = 1; \beta = 1; \gamma = 3$

Figura 11: Quadrados perfeitos com  $\alpha = 1; \beta = 2$

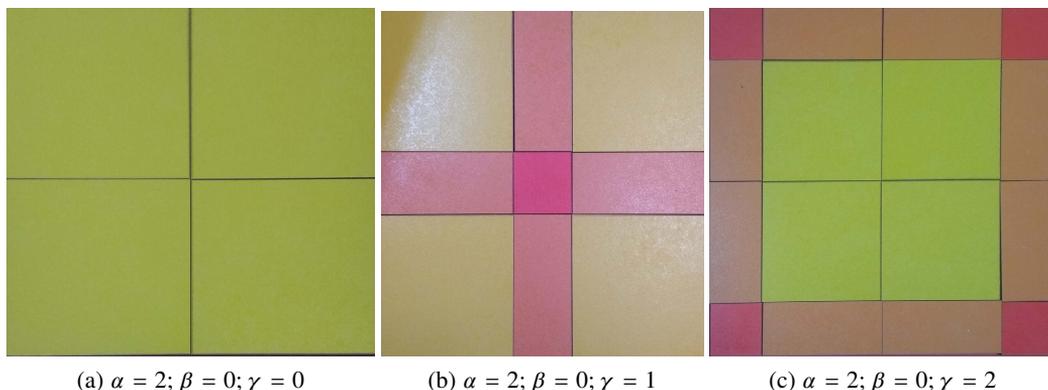


(a)  $\alpha = 1; \beta = 2; \gamma = 0$

(b)  $\alpha = 1; \beta = 2; \gamma = 1$

(c)  $\alpha = 1; \beta = 2; \gamma = 2$

Figura 12: Quadrados perfeitos com  $\alpha = 2; \beta = 0$

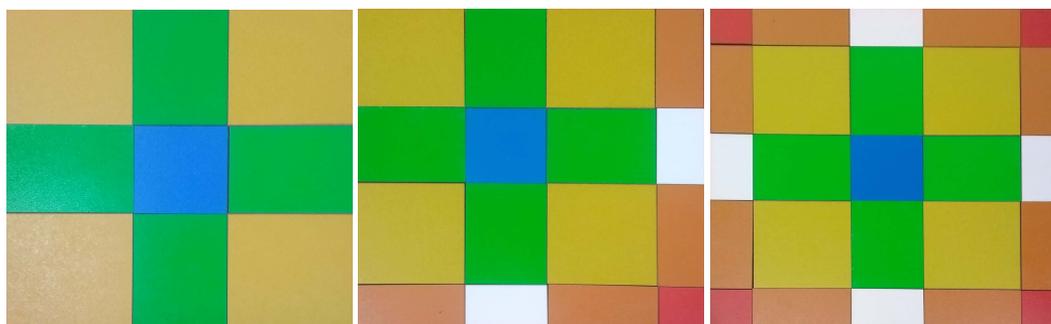


(a)  $\alpha = 2; \beta = 0; \gamma = 0$

(b)  $\alpha = 2; \beta = 0; \gamma = 1$

(c)  $\alpha = 2; \beta = 0; \gamma = 2$

Figura 13: Quadrados perfeitos com  $\alpha = 2; \beta = 1$



(a)  $\alpha = 2; \beta = 1; \gamma = 0$

(b)  $\alpha = 2; \beta = 1; \gamma = 1$

(c)  $\alpha = 2; \beta = 1; \gamma = 2$

### 5. Sugestões para uso em sala de aula

A fim de auxiliar o professor que deseja utilizar esta atividade em suas aulas, sugerimos alguns procedimentos baseados na nossa experiência com o uso deste material.

Quando se propõe a questão “quantos quadrados podem-se formar com a justaposição das peças do Algeplan” os alunos, na sua maioria, põem-se a trabalhar montando os quadrados usando, inicialmente, o método de tentativa e erro.

Apenas depois de certo tempo manipulando as peças é que algumas ideias de sistematização da contagem começam a surgir, mas ainda sem envolver a modelagem algébrica do problema.

Sendo assim, sugerimos que no início da atividade, logo após a apresentação e descrição do Algeplan, sejam fixadas as nomenclaturas  $x, y$  e  $1$  para os lados das peças e também a relação de cada tipo de peça com sua representação algébrica, conforme a Tabela 1. Recomendamos que alguns exemplos sejam expostos na lousa, enfatizando a relação entre os coeficientes do polinômio de grau 2 com a quantidade de peças utilizadas na justaposição e, no caso específico de a justaposição resultar em um quadrado, a relação da quantidade de peças com os coeficientes da forma fatorada  $(\alpha x + \beta y + \gamma)^2$  do polinômio.

Tabela 5: Tabela para os alunos preencherem durante a aplicação

	AMARELO	VERDE	AZUL	VERMELHO	ROSA	LARANJA	EXPRESSION ALGÉBRICA
1							
2							
∴							

Uma maneira de chamar a atenção do aluno para a relação entre a quantidade de peças usadas em um quadrado e a expressão algébrica correspondente é solicitar que, em uma tabela similar à Tabela 5, a cada quadrado encontrado, seja anotada a quantidade de peças de cada cor que foram utilizadas bem como a expressão algébrica a ele relacionada. Dessa forma, mesmo que algum estudante não descubra por si só um argumento que justifique que a quantidade exata de quadrados é 28, ficará mais fácil a compreensão da explicação dada posteriormente pelo professor.

No entanto, a maneira de desenvolver a atividade dependerá do nível de maturidade dos alunos envolvidos e do tempo disponível, cabendo ao professor mediar e conduzir a execução do trabalho.

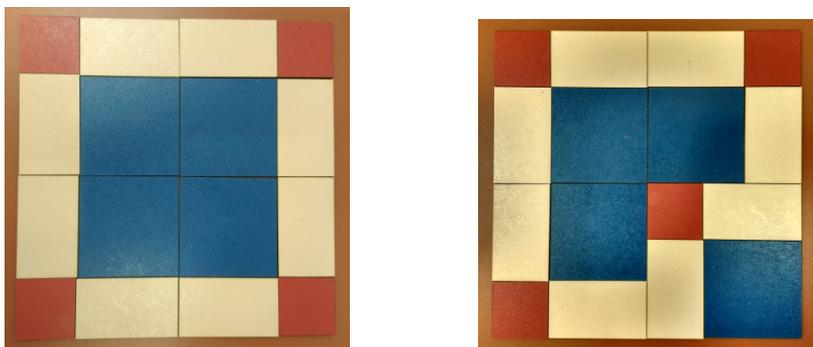
### Considerações finais

Vale notar que, ao se considerar uma quantidade de peças diferente da apresentada na Tabela 1, uma nova análise de possibilidades precisa ser feita, seguindo o raciocínio da Tabela 3, usando as novas restrições. Este fato abre a possibilidade do uso de tal atividade em mais de uma ocasião, modificando a quantidade de um ou mais tipos de peças do Algeplan.

Ainda: mesmo que o professor não tenha acesso a kits do material concreto, existem versões virtuais do Algeplan que podem ser usadas em computadores e *smartphones*, conforme [4], [5]. O uso da versão virtual pode ser uma outra alternativa de implementação desta atividade, inclusive com diferentes quantidades de peças.

Por fim, quando se observa a Figura 14, percebe-se que existe mais de uma maneira de rearranjar as peças do Algeplan para se construir um quadrado. Descobrir, para cada quadrado, quantas “maneiras diferentes” as peças podem ser rearranjadas, é também um outro problema de contagem que pode ser trabalhado com o auxílio deste material.

Figura 14: O quadrado de lado  $2y + 2$  montado de duas maneiras diferentes



## Agradecimentos

Agradecemos ao Laboratório de Ensino e Pesquisa em Educação Matemática - Furg-SAP pelo empréstimo do material usado neste trabalho, e as valiosas sugestões da Editora responsável pelo parecer.

## Referências

- [1] Ausubel, D. P. *Aquisição e retenção de conhecimentos: uma perspectiva cognitiva*. Lisboa: Plátano, 1. 2003.
- [2] Fanti, E. L. C; Rosa, R. A., Dias, F. M.; Medeiros, L. T.; *O Algeplan como um recurso didático na exploração de expressões algébricas e fatoração*. In: III Biental da SBM, 2006, Goiânia. Pôsteres da III Biental da SBM, 2006.
- [3] Poletto, C. S., *Algeplan, Álgebra e Geometria : entendendo práticas matemáticas como jogos de linguagem*. Trabalho de conclusão de curso - UFRS, Porto Alegre, 2010.
- [4] Silva, P. L.; Basso, M. V. A. “*Algeplan Virtual: Um Recurso para o Ensino de Operações com Polinômios*”. Anais do XI Encontro Nacional de Educação Matemática ”– ISSN 2178-034X, 2013.
- [5] Massante, K. A. S. C. C.; Barbosa, A. C. M.; Gonçalves, I. M. “*Atividades Utilizando o Algeplan no Software Geogebra*”. XII Encontro Nacional de Educação Matemática”, 2016.

Karina Rodriguez de Fraga  
Universidade Federal do Rio Grande  
<[fragak14@gmail.com](mailto:fragak14@gmail.com)>

Leandro Sebben Bellicanta  
Universidade Federal do Rio Grande  
<[leandroblicanta@furg.br](mailto:leandroblicanta@furg.br)>

Recebido: 29/10/2019  
Publicado: 08/07/2020

# Desenvolvimento da disciplina de Cálculo Numérico tendo um modelo em Ecologia como eixo norteador

Bruno Telch dos Santos 

Carolina Cominato 

Rodrigo Cardoso dos Santos 

## Resumo

Neste artigo mostramos parte do desenvolvimento da disciplina de Cálculo Numérico tendo como eixo norteador um problema em ecologia matemática. É introduzido ao aluno um problema de valor inicial com uma equação diferencial tendo significado biológico, significado que é absorvido conforme o aluno avança nas técnicas do curso. Com o avanço das técnicas, o aluno obtém resultados numéricos e os apresenta em formato de texto científico. A parte apresentada cobre os métodos numéricos para resolução de problemas de valores iniciais começando em Euler-implícito e Euler-implícito, onde, neste último, aplicamos as técnicas para obter zeros de função. Numa etapa final refinamos o problema de valor inicial e fazemos novos experimentos numéricos com auxílio de métodos numéricos mais eficientes, como o clássico método de Runge-Kutta de quarta ordem.

**Palavras-chave:** Cálculo numérico, Euler-implícito, Euler-explícito, Métodos de Runge-Kutta, Ecologia matemática, Modelos populacionais.

## Abstract

In this article, we show part of the development of the Numerical Calculus course, with a problem in mathematical ecology as its guiding axis. An initial value problem with a differential equation with biological meaning is introduced to the student, that meaning is absorbed according to the student who is learning about course techniques. With the advancement of techniques, the student obtains numerical results and presents them in a scientific text format. The part presented covers the numerical methods for solving initial value problems starting with Euler-explicit and Euler-implicit, where in the latter method we apply the techniques to obtain function zeros. In a final step, we refine the initial value problem and make new numerical experiments with the aid of more efficient numerical methods, such as the classic fourth-order Runge-Kutta method.

**keywords:** Numerical Calculus, Euler-explicit, Euler-implicit, Runge-Kutta methods, Ecology, population models.

## 1. Introdução

Este trabalho tem por objetivo apresentar o desenvolvimento do curso de Cálculo Numérico tendo um problema de valor inicial com uma equação diferencial ordinária (PVI) como o eixo norteador. Essa ideia faz com que as resoluções numéricas desses problemas sejam o principal foco do curso,

onde há um espaço importante para o aluno apresentar os resultados que vem obtendo e treinar sua escrita científica. O curso foi ministrado por um dos autores, e o PVI sugerido é bastante simples e visível, mas teve o cuidado de, durante sua construção, as soluções dos problemas intermediários terem soluções exatas, onde conseguimos exibir de forma simples o que é a solução numérica de um PVI, uma vez que conseguimos comparar com a solução exata do problema. Assim, para o estudante, fica o objetivo de escrever um texto descontraído/agradável conforme o professor avança com o conteúdo em sala de aula. Acreditamos que tal abordagem seja interessante antes ou depois de um primeiro curso de equações diferenciais ordinárias, já que as modelagens são bastante simples e as análises qualitativas do problema ficam bastante evidentes. São os exemplos abordados meramente ilustrativos, sem qualquer compromisso com a realidade, levando em consideração que Cálculo Numérico é o primeiro curso nesse ambiente.

Nota-se, através de pesquisas como as de Soares e Sauer em [13], que há um certo padrão a respeito da apresentação de conteúdos, enfoques, bibliografia e técnicas de ensino nos cursos de Cálculo Numérico de algumas universidades brasileiras, sendo o programa clássico dessa disciplina resumido a: erros computacionais, raízes de funções, sistemas lineares, interpolação, integrais, ajuste de curvas e equações diferenciais, levando em consideração que esses autores notaram que em alguns planos de ensino não constam equações diferenciais. Além disso, a metodologia utilizada nos processos de ensino-aprendizagem é composta de explicações teóricas, seguidas de uma bateria de exercícios a serem resolvidos com o uso de calculadoras e/ou de *softwares* computacionais, além de trabalhos que envolvem a produção de programas com as técnicas apresentadas. Dessa forma, o ensino dos conteúdos a serem apresentados durante a disciplina pode tornar-se desestimulante e sem sentido para os aprendizes, visto que esses buscam um significado das ferramentas a serem aprendidas no curso em sua área de formação. Assim, nos propomos a trabalhar com um texto em ecologia matemática de dinâmica populacional, onde é estudado um modelo de crescimento populacional logístico com um termo de predação, que pode sofrer acentuações (como uma epidemia, guerra, migração etc.) ao longo do tempo. O texto proposto é clássico na literatura, por exemplo no primeiro capítulo de [11], e a proposta é basicamente a construção do texto com as ferramentas que vamos absorvendo ao longo do semestre. No decorrer do curso exibimos experimentos numéricos a fim de entender a dinâmica estudada.

Além disso, vale a pena comentar que, apesar dos métodos de Euler implícito ou explícito não serem de uso comum, foram utilizados, pois acreditamos que são extremamente didáticos para o entendimento do que é uma solução numérica para um problema de valor inicial com uma Equação Diferencial Ordinária. Após a apresentação desses métodos, a aprendizagem dos alunos em relação aos métodos mais sofisticados pode ser facilitada. Foi explorada também neste texto uma porção considerável na disciplina de Cálculo Numérico, a saber, os zeros de funções e os métodos de Runge-Kutta. Fica claro que métodos de interpolação ou ajustes de curva caberiam na discussão, mas limitamos esse trabalho apenas às soluções numéricas dos PVIs. No desenvolver do trabalho, escolhemos utilizar os métodos da bissecção e Newton-Raphson para aproximação de zeros de funções e método de Runge-Kutta de quarta ordem por ser o de uso mais comum.

Uma ferramenta tecnológica que nos auxiliou no desenvolvimento do curso foi o *software* Matlab, porém, outros *softwares* clássicos de programação podem ser utilizados a fim de atingir a metodologia aplicada. Por fim, foi importante perceber que o interesse dos estudantes por avançar na disciplina teve uma avaliação satisfatória, visto que parte substancial desse artigo é o trabalho de uma das alunas do curso, o qual aqui refinamos para dar um enfoque didático. Como afirmam Barbosa e Canalli em [3], acreditamos que as ações associadas no conjunto composto por alunos e professor têm por objetivo a estimulação para assimilação de conteúdos e técnicas facilmente aplicáveis aos

aprendizes em diversas situações de maneira criativa e consciente.

A seguir, apresentaremos o desenvolvimento do modelo composto e os momentos em que as técnicas de Cálculo Numérico foram sendo apresentadas e trabalhadas.

## 2. Apresentação do problema

A Ecologia é o estudo científico da distribuição e abundância dos seres vivos e das interações que determinam a sua distribuição. As interações podem ser entre seres vivos e/ou com o meio ambiente. Ciência ampla e complexa, a Ecologia preocupa-se com o entendimento do funcionamento de toda a natureza. Um exemplo disso é a Ecologia de populações, onde modelos matemáticos realistas e úteis na biologia populacional são um forte objeto de estudos, uma vez que ajudam a entender os processos dinâmicos envolvidos e também a fazer previsões práticas. Talvez o mais comum no dia a dia sejam modelos com população humana, porém também é comum lidar nesses modelos com população de uma espécie ameaçada de extinção, crescimento bacteriano ou viral, entre outras.

Modelos de crescimento populacional são clássicos na literatura e são introduzidos desde os primeiros cursos de Cálculo. Esses modelos também costumam ser úteis e eficientes para descrever as sequências pelas quais as sociedades são transformadas. Denotamos por  $N(t)$  a densidade de uma certa população no instante de tempo  $t$  e consideramos que essa segue um crescimento logístico, seguindo [11] (Cap. 1, pp1-43, e ex. 4), dada por

$$\begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) \\ N(0) = n_0, \end{cases} \quad (1)$$

onde  $r$  representa um coeficiente de crescimento,  $K$  é entendido como a capacidade de carga do ambiente e  $n_0$  é a população inicial. Essa população pode passar pelo que chamaremos de efeitos de predação – por exemplo, se considerarmos a população humana, essa pode enfrentar períodos de guerras ou pandemias, ou também efeitos de imigração/emigração, tendo consequência direta na sua taxa de crescimento. Nessa ideia, completamos o PVI (1) para

$$\begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) - P(N)F(t) \\ N(0) = n_0, \end{cases} \quad (2)$$

onde  $P$  representa uma taxa de predação (que depende da população  $N$ ),  $F(t)$  é dada e representa se há ou não a presença do efeito predador. O PVI (2) diz que a população  $N$  cresce segundo um crescimento logístico e tem seu crescimento afetado quando há a presença do termo de predação  $P(N)F(t)$ . Na literatura é comum  $F \equiv 1$ , e algumas escolhas para  $P$  clássicas são

$$P(N) = P_0 \left(1 - e^{-\frac{N^2}{A}}\right), \quad P(N) = \frac{P_0 N^2}{A^2 + N^2},$$

onde os parâmetros  $A$  e  $P_0$  são constantes positivas. Sendo assim, espera-se desse trabalho um estudo construtivo via métodos numéricos básicos das soluções numéricas e das construções dos PVI's

$$\begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) - P \left(1 - e^{-\frac{N^2}{A}}\right) F(t) \\ N(0) = n_0, \end{cases} \quad (3)$$

e

$$\begin{cases} \frac{dN}{dt} = rN \left( 1 - \frac{N}{K} \right) - \frac{P_0 N^2 F(t)}{A^2 + N^2} \\ N(0) = n_0. \end{cases} \quad (4)$$

A proposta é estudarmos numericamente o desenvolvimento histórico (1), (3) e (4) de acordo com o desenvolvimento cronológico da disciplina de Cálculo Numérico, onde começaremos com os métodos mais simples para os casos iniciais e utilizaremos métodos mais sofisticados a fim de obter a solução para escolhas de  $F$  sendo constante (ou aproximadamente constante) na maior parte e ter certos picos de crescimento. Exemplos clássicos de funções com esse comportamento são as gaussianas  $F(t) = 1 + ce^{-\frac{(t-t_0)^2}{d}}$ , com parâmetros  $c \in \mathbb{R}$  e  $d > 0$  que podem ter significados absorvidos observando alguns exemplos nas figuras 1 e 2.

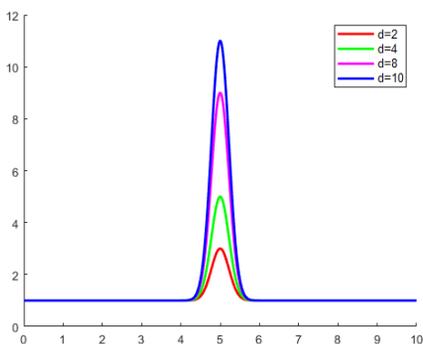


Figura 1: Propostas para  $F$ , onde consideramos  $F(t) = 1 + ce^{-\frac{(t-t_0)^2}{d}}$  com  $c = 0.1$ ,  $x_0 = 5$  e  $d$  como na figura.

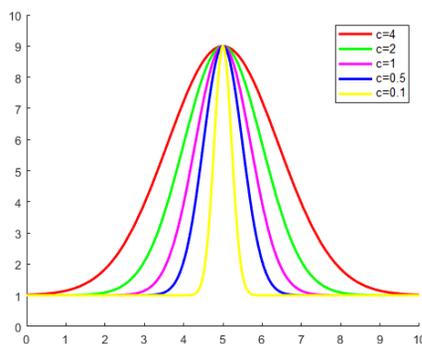


Figura 2: Propostas para  $F$ , onde consideramos  $F(t) = 1 + ce^{-\frac{(t-t_0)^2}{d}}$  com  $d = 8$ ,  $x_0 = 5$  e  $c$  como na figura.

Como exibido nessas figuras, o parâmetro  $c$  representa a intensidade adicional que o termo de predação apresentará, e  $d$  podemos entender como o tempo em que esse efeito acentuado ocorre.

### 3. Desenvolvimento da proposta

Representamos por  $N(t)$  a população da espécie no tempo  $t$ , então temos a relação

$$\text{Variação de } N = \text{Nascimentos} - \text{Mortes} + \text{Migração} \quad (5)$$

que é uma equação de conservação para a população. A forma dos termos do lado direito de (5) requer modelar a situação com a qual estamos preocupados. O modelo mais simples nesse cenário não tem migração e é razoável supor que os termos de nascimento e morte são proporcionais à população  $N$ , ou seja,

$$\begin{cases} \frac{d}{dt} N = bN - dN \\ N(0) = n_0, \end{cases} \Rightarrow N(t) = n_0 e^{(b-d)t} = \text{solução exata}, \quad (6)$$

onde  $b$  e  $d$  são constantes positivas que representam a taxa de nascimento e morte da população, respectivamente, e a população inicial  $N(0) = n_0$  que representa o número de indivíduos no instante

inicial. Assim, se  $b > d$ , a população cresce exponencialmente, enquanto decresce exponencialmente se  $b < d$ .

Aproximar numericamente a solução de um problema de valor inicial pode ser entendido, basicamente, por obter uma quantidade de pontos  $(t_1, y_1), \dots, (t_n, y_n)$  de modo que a curva formada ao ligar pontos  $(t_i, y_i)$  a  $(t_{i+1}, y_{i+1})$ ,  $i = 1, \dots, n - 1$ , via alguma interpolação (por exemplo, um segmento de reta), ganhe alguma suavidade.

Os métodos mais simples para ilustrar são os métodos de Euler-explícito (EE) e Euler-implícito (EI), que são oriundos, respectivamente, das seguintes discretizações:

$$N'(t) \approx \frac{N(t+h) - N(t)}{h} \quad \text{e} \quad N'(t) \approx \frac{N(t) - N(t-h)}{h}. \quad (7)$$

Apesar de serem raramente usados, são muito eficientes para ilustrar a ideia. A fim de aproximar numericamente a solução de (6) no intervalo  $[0, T_{\text{final}}]$ , onde  $T_{\text{final}} > 0$  é um valor a ser escolhido, consideramos uma grade de tempos  $\{t_1, t_2, \dots, t_n\}$ , onde  $t_1 = 0$ ,  $t_{i+1} = t_i + h$  para  $i = 1, \dots, n - 1$ , com  $t_n = T_{\text{final}}$ . Assim,  $n$  e  $h$  são relacionados por  $h = T_{\text{final}}/n$ . Denotamos, em ambos os casos,  $N_i = N(t_i)$ . Também,  $a = b - d$ . Com essa notação, podemos fazer a leitura das discretizações de (7) da seguinte forma:

$$\text{Euler-explícito: } N'(t) \approx \frac{N(t+h) - N(t)}{h} = \frac{N \text{ avaliado um passo a frente} - N \text{ atual}}{h}$$

e

$$\text{Euler-implícito: } N'(t) \approx \frac{N(t) - N(t-h)}{h} = \frac{N \text{ atual} - N \text{ avaliado um passo atrás}}{h}$$

A aplicação do método de Euler explícito resulta

$$\frac{d}{dt} N = aN \Rightarrow \frac{N_{i+1} - N_i}{h} = aN_i \Rightarrow \begin{cases} N_{i+1} = N_i + haN_i, & i = 1, \dots, n - 1 \\ N_1 = n_0. \end{cases} \quad (\text{EE})$$

Para Euler-implícito, temos

$$\frac{d}{dt} N = aN \Rightarrow \frac{N_{i+1} - N_i}{h} = aN_{i+1} \Rightarrow \begin{cases} N_{i+1} = \frac{N_i}{1 - ah}, & i = 1, \dots, n - 1 \\ N_1 = n_0, \end{cases} \quad (\text{EI})$$

A nomenclatura explícito-implícito pode ser entendida preliminarmente da seguinte forma: para obtermos  $N_{i+1}$  conhecendo  $N_i$  via (EE) basta efetuar a operação descrita no algoritmo, já para obtermos  $N_{i+1}$  conhecendo  $N_i$  via (EI) precisamos fazer uma manipulação adicional, que nesse exemplo ultras-simplificado ainda pode não estar clara, que é calcular o zero de uma função. Desenvolveremos isso mais à frente; por ora, para fixar a ideia, consideremos para o PVI os dados

$$n_0 = 1 \quad a = 0.45, \quad T_{\text{final}} = 4, \quad \text{e} \quad h = 0.5 \quad (8)$$

isto é, aproximaremos numericamente a solução de

$$\begin{cases} \frac{d}{dt} N = 0.45N \\ N(0) = 1, \end{cases} \Rightarrow N(t) = e^{0.45t} = \text{solução exata.} \quad (9)$$

Aproximaremos a solução em  $[0, 4]$  com  $h = 1$ . Com algoritmos apresentados em **(EE)** e **(EI)** obtemos o resultado apresentado na Tabela 1. Com esse resultado, em Matlab, implementamos como na Figura 3.

```

1 - T=[0 0.5000 1.0000 1.5000 2.0000 2.5000 3.0000 3.5000 4.0000];
2 - EE=[1.0000 1.2250 1.5006 1.8383 2.2519 2.7585 3.3792 4.1395 5.0709];
3 - EI=[1.0000 1.2903 1.6649 2.1483 2.7720 3.5768 4.6152 5.9551 7.6840];
4 - hold on % Plotando as soluções aproximadas e solução exata
5 - plot(T,EI,'bo','markerfacecolor','r');
6 - plot(T,EE,'bo','markerfacecolor','b');
7 - fplot(@(x) exp(0.45*x),[0 4],'LineWidth',1)
8 - hold off
9 - legend('Solução aproximada pelo método implícito','Solução aproximada pelo método explícito','Solução Exata')

```

Figura 3: Plotando em Matlab o resultado obtido na Tabela 1

t	EE	EI
0	1.0000	1.0000
0.5	1.2250	1.2903
1.0	1.5006	1.6649
1.5	1.8383	2.1483
2.0	2.2519	2.7720
2.5	2.7585	3.5768
3.0	3.3792	4.6152
3.5	4.1395	5.9551
4.0	5.0709	7.6840

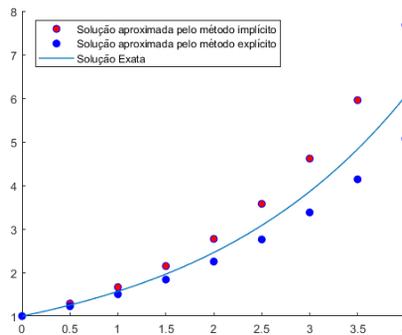


Tabela 1: Aproximações numéricas via **(EE)** e **(EI)** com dados (8).

Figura 4: Soluções exata e numérica (Euler explícito e implícito) para (9) com dados (8).

Intuitivamente, quanto menor  $h$ , melhor a aproximação numérica para o PVI. Implementamos então em Matlab o desenvolvimento da Figura 5, onde agora ligaremos os pontos  $(t_i, y_i)$  a  $(t_{i+1}, y_{i+1})$ ,  $i = 1, \dots, n - 1$ , por segmentos de reta.

```

Tfinal= ; h= ; % Determinar o Tempo final e o espaçamento h
T= 0:h:Tfinal; % Grade de tempos
n0= ; a= ; % Determinar a população inicial e o coeficiente a
EI(1)=n0; EE(1)=n0; % Os primeiros elementos da lista
n=Tfinal/h; % A quantidade de pontos a ser determinada
for i= 1 : n
    EI(i+1)= EI(i)/(1-a*h); % (*)
    EE(i+1)=(1+a*h)*EE(i);
end
hold on % Plotando as soluções aproximadas e solução exata
plot(T,EI,'LineWidth',3);
plot(T,EE,'LineWidth',3);
fplot(@(x) N0*exp(a*x),[0 4],'LineWidth',3)
hold off
legend('Solução aproximada pelo método implícito','Solução aproximada pelo método explícito','Solução Exata')
axis([0 Tfinal 0 6]);

```

Figura 5: Algoritmo geral para **(EE)** e **(EI)** para o problema (6). Precisamos determinar  $T_{final}$ ,  $h$ ,  $n_0$  e  $a := b - d$ .

Apresentamos nas figuras 6 e 7 alguns exemplos abaixo com alguns outros valores.

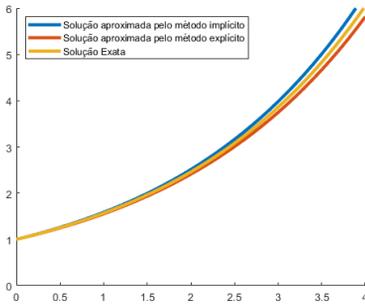


Figura 6: Aplicação do Algoritmo da Figura 5 com  $T_{final} = 4$ ,  $h = 0.1$ ,  $n_0 = 1$  e  $a = 0.45$ .

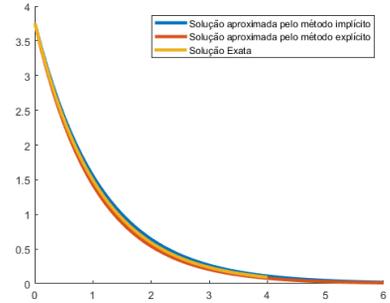


Figura 7: Aplicação do Algoritmo da Figura 5 com  $T_{final} = 6$ ,  $h = 0.1$ ,  $n_0 = 3.76$  e  $a = -0.925$ .

Voltando ao problema populacional, podemos observar que essa abordagem não corresponde à realidade e sugere que modelos mais realistas devam ser propostos. No entanto, um clássico exemplo é considerar as estimativas de crescimento passadas e previstas para a população mundial total do século XVII ao século XXI, que mostram que tal abordagem não é tão ruim e pode ser adequada para o início da dinâmica, como podemos ver nas figuras 8 e 9, que mostram tais dados para as populações americanas e francesas num período catalogado.

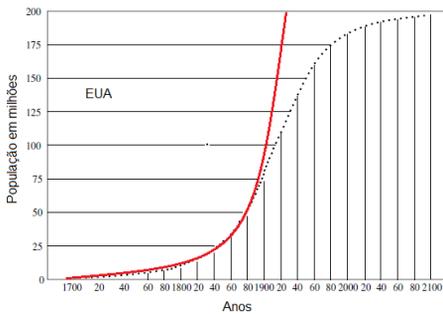


Figura 8: Crescimento populacional da EUA ao longo de alguns anos e previsões futuras (Figura adaptada de [11])

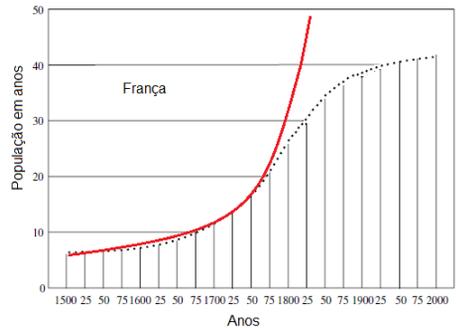


Figura 9: Crescimento populacional da França ao longo de alguns anos e previsões futuras (Figura adaptada de [11])

O modelo (6) e a teoria da transição demográfica de Warren Thompson (1887-1973) ajudam-nos a entender os modelos populacionais mais sofisticados. Thompson, em sua teoria, observou as mudanças que tinham experimentado nos últimos duzentos anos as sociedades industrializadas do seu tempo com respeito às taxas de natalidade e de mortalidade. De acordo com essas observações, expôs os quatro estágios da teoria da transição demográfica pelos quais a sociedade pós-industrial passa:

- Processo que acontece antes da industrialização, a fase 1 é caracterizada com elevadas taxas de natalidade compensadas por altas taxas de mortalidade. Muitas crianças morreram antes de

atingir a idade adulta, e as doenças e fomes foram generalizadas. A expectativa de vida era menor do que hoje, mas, em média, mais crianças nasceram para cada mãe. Nessa época não existiam métodos contraceptivos e as condições de saúde eram precárias;

- A melhoria nas condições sanitárias, a evolução da medicina e a urbanização aumentam a expectativa de vida na fase 2, onde os índices de mortalidade iniciam importante descida. Surgem problemas como a explosão demográfica, superpopulação e aumento do desemprego;
- Na fase 3 ocorre um declínio na taxa de natalidade devido ao acesso a métodos anticoncepcionais, o elevado custo de vida nas grandes metrópoles e o acesso à educação, fazendo com que o planejamento familiar fique mais difundido;
- Na fase 4 as taxas de natalidade e mortalidade encontram-se muito baixas, e é criada uma estabilização no crescimento vegetativo, tendo por consequência uma taxa de crescimento natural nula ou negativa.

Atualmente se fala em uma quinta fase, onde a mortalidade superará a natalidade, devido ao alto custo de se criar filhos (principalmente em países desenvolvidos) e famílias optam por ter um número muito reduzido (entre 1 e nenhum) de filhos para manter o padrão de vida. Alemanha e Itália são exemplos de países nessa fase, onde se cria a preocupação quanto ao rombo que isso pode acarretar na previdência.

As figuras 8 e 9 mostram os resultados colhidos para a população dos EUA e da França entre 1790 e 1910, junto à previsão para os próximos anos, segundo modelos mais realistas. Há um bom ajuste para a população no período catalogado, que representas as fases 1 e 2 da teoria de Thompson. No entanto, o resto da curva solução de (5) não está nem perto dos dados reais da população, uma vez que na fase 3 há a redução do número de nascimentos. O ponto principal não é que as previsões sejam tão imprecisas usando (5), mas sim que a curva ajuste apenas parte dos dados, e particularmente a parte que não cobre a maior parte da curva de crescimento torne a comparação com dados e previsões futuras extremamente não confiáveis. *Grosso modo*, um modelo mais realista que a abordagem feita em (5) seria dizer que a variação da população  $N$  tem crescimento exponencial caso essa população seja pequena (fases 1 e 2), que essa desacelera a partir de uma certa quantidade de indivíduos (fase 3) e que há mortalidade quando a população  $N$  for muito grande (fase 4). Visando isso, a longo prazo, deve haver algum ajuste a esse crescimento exponencial e uma proposta que se adequa seria

$$\begin{cases} \frac{d}{dt} N = rN(1 - N/K) \\ N(0) = n_0, \end{cases} \Rightarrow N(t) = \frac{n_0 K e^{rt}}{K + n_0(e^{rt} - 1)}, \quad (10)$$

onde  $r$  e  $K$  são constantes positivas. É interessante observar que  $N(t) \xrightarrow{t \rightarrow +\infty} K$ . Esse modelo é conhecido por crescimento logístico em uma população. Neste modelo, a taxa de natalidade *per capita* é  $r(1 - N/K)$ , isto é, depende de  $N$ . Essa taxa de crescimento indica que, conforme a densidade populacional  $N$  aproxima-se de  $K$  a taxa de crescimento vai se aproximando de zero e estabilizando. Assim, a constante  $K$  num contexto geral é a capacidade de carga do ambiente, que geralmente é determinada pelos recursos de sustentação disponíveis.

Trabalhando da mesma forma como feito em (6), temos as abordagens por Euler-explícito dadas

por

$$\frac{d}{dt} N = rN(1-N/K) \Rightarrow \frac{N_{i+1} - N_i}{h} = rN_i(1-N_i/K) \Rightarrow \begin{cases} N_{i+1} = N_i + hrN_i(1 - N_i/K), \\ i = 1, \dots, n-1, \\ N_1 = n_0, \end{cases} \quad (\text{EE})$$

e para Euler-implícito, temos

$$\begin{aligned} \frac{d}{dt} N = rN(1 - N/K) &\Rightarrow \frac{N_{i+1} - N_i}{h} = rN_{i+1}(1 - N_{i+1}/K) \\ &\Rightarrow \begin{cases} N_{i+1} - N_i - rhN_{i+1}(1 - N_{i+1}/K) = 0, \\ i = 1, \dots, n-1, \\ N_1 = n_0. \end{cases} \end{aligned} \quad (\text{EI})$$

Notemos que agora o problema (EI) não é totalmente trivial. Em cada passo, a fim de determinar quem é o valor  $N_{i+1}$ , precisamos obter o zero de uma função apropriada, isto é, para cada  $i = 1, \dots, n-1$ ,

$$N_{i+1} - N_i - rhN_{i+1}(1 - N_{i+1}/K) = 0 \Rightarrow F(N_{i+1}) = 0, \text{ com } F(x) = x - N_i - rhx(1 - x/K),$$

onde é preciso que se veja bem: em cada passo que queremos determinar  $N_{i+1}$  já temos conhecido o valor  $N_i$ . Por exemplo, a fim de determinar  $N_2$ , isto é,

$$N_2 - N_1 - rhN_2(1 - N_2/K) = 0 \stackrel{N_1=n_0}{\Rightarrow} F(N_2) = 0, \text{ com } F(x) = x - n_0 - rhx(1 - x/K).$$

Obtendo o zero de  $F$ , temos o  $N_2$ . De forma indutiva, para determinar o  $N_3$ , temos

$$N_3 - N_2 - rhN_3(1 - N_3/K) = 0 \stackrel{\text{Agora já conhecemos } N_2}{\Leftrightarrow} F(N_3) = 0, \text{ com } F(x) = x - N_2 - rhx(1 - x/K),$$

e assim por diante. Claro que não é nenhum grande desafio obter os candidatos a  $N_i$ ,  $i = 1, \dots, n-1$ , já que em cada passo acima, para determinar o zero de cada  $F$  podemos simplesmente usar a fórmula de Bhaskara. A fim, porém, de se dar mais complexidade ao modelo, eventualmente, surgirá a necessidade de obter zeros de função em geral de forma eficiente.

**Zeros de funções:** Dos métodos clássicos para aproximar zeros de funções vamos usar bissecção (intervalar) e o método de Newton-Raphson (método de ponto fixo), porém, os desenvolvimentos podem ser estendidos para os demais casos. Da teoria elementar, de forma bastante resumida, temos os métodos intervalares sendo, em geral, convergentes, porém lentos. Enquanto isso, os métodos de ponto fixo são mais eficientes, desde que o pontapé inicial esteja próximo da solução. Os métodos intervalares, em geral, são bem empregados para iniciarmos algum método de ponto fixo, e é essa técnica que iremos utilizar. Todos esses métodos citados podem ser vistos em [4], por exemplo.

Para fixar as ideias, usaremos primeiramente o método da bissecção e consideramos o problema (10) com dados

$$n_0 = 0.36, \quad r = 1.925, \quad K = 3, \quad T_{\text{final}} = 4, \quad h = 0.25, \quad (11)$$

isto é, aproximaremos numericamente a solução de

$$\begin{cases} \frac{d}{dt} N = 1.925N(1 - N/3) \\ N(0) = 0.36, \end{cases} \Rightarrow N(t) = \frac{0.36 \times 3e^{1.925t}}{3 + 0.36(e^{1.925t} - 1)} = \text{Solução exata}, \quad (12)$$

Apresentamos o algoritmo do método da bissecção como na Figura 10, com o esclarecimento de cada entrada apresentada na legenda.

```
function c=bissec(f,a,b,tol)
d=1;
if subs(f,a)*subs(f,b)>0
    fprintf('Pode ser que não haja raiz no intervalo desejado! =')
    error('Não obtivemos raiz, melhorar o intervalo!')
end
while d>tol
    c=(a+b)/2;
    if subs(f,a)*subs(f,c)<0
        b=c;
    end
    if subs(f,c)*subs(f,b)<0
        a=c;
    end
    d=abs(subs(f,c));
end
```

Figura 10: Algoritmo da Função Bissecção: tem como entrada a função  $f$  que queremos determinar o zero, o intervalo  $[a, b]$  onde acreditamos estar um zero de  $f$  e  $tol$  é a tolerância desejada.

Temos  $N_1 = n_0 = 0.36$ , e obtemos  $N_2$  via  $F_2(N_2) = 0$ , com  $F_2(x) = x - 0.36 - (0.25)(1.925)x(1 - x/3)$ :

$$N_2 = \text{bissec}(F_2, 0.36 - 0.4, 0.36 + 0.4, 0.001) \Rightarrow N_2 = 0.5885. \quad (13)$$

Temos  $N_2 = 0.5885$  e obtemos  $N_3$  via  $F_3(N_3) = 0$  com  $F_3(x) = x - 0.5885 - (0.25)(1.925)x(1 - x/3)$ :

$$N_3 = \text{bissec}(F_3, 0.5885 - 0.4, 0.5885 + 0.4, 0.001) \Rightarrow N_3 = 0.8903. \quad (14)$$

Temos  $N_3 = 0.8903$  e obtemos  $N_4$  via  $F_4(N_4) = 0$  com  $F_4(x) = x - 0.8903 - (0.25)(1.925)x(1 - x/3)$ :

$$N_4 = \text{bissec}(F_4, 0.8903 - 0.4, 0.8903 + 0.4, 0.001) \Rightarrow N_4 = 1.2404. \quad (15)$$

E assim sucessivamente. Temos os resultados para  $N_i$ ,  $i = 1, \dots, 16$  apresentados na Tabela 2 na coluna referente a Euler-implícito (**EI**). Para obter a coluna referente ao método de Euler-explícito (**EE**) basta seguir o algoritmo com os dados (8).

No exemplo acima vimos uma forma bastante simples de obter os pontos referentes ao método de Euler-implícito (**EI**) utilizando o método da bissecção, e essa forma pode ser resumida em dois pontos abordados na Figura 12:

**1** O método de Euler-implícito consiste em obter o zero de uma função; então, em cada passo precisamos calcular o zero de uma função adequada, ou seja, a  $f$  é atualizada em cada passo, como podemos ver em (13), (14) e (15).

**2** O método da bissecção pede um intervalo  $[a, b]$  onde haja uma raiz da função  $f$ . Como esperamos que em cada passo o ponto obtido esteja razoavelmente próximo ao anterior, a fim de obter

$EI(i + 1)$  fizemos a proposta de que, como esse é a raiz de uma  $f$  adequada,  $EI(i + 1)$  deve estar no intervalo  $[EI(i) - 0.4, EI(i) + 0.4]$ . Dependendo do crescimento/decrescimento da solução, a proposta de um intervalo centrado em  $EI(i)$  de raio 0.4 pode não ser um intervalo adequado, precisando de um ajuste para cima. Da mesma forma, considerando  $h$  pequeno, esse intervalo pode ser diminuído.

t	EE	EI
0	0.3600	0.3600
0.25	0.5125	0.5885
0.5	0.7170	0.8903
0.75	0.9795	1.2404
1.0	1.2970	1.5993
1.25	1.6513	1.9303
1.5	2.0086	2.2101
1.75	2.3280	2.4313
2.0	2.5790	2.5983
2.25	2.7532	2.7199
2.5	2.8622	2.8063
2.75	2.9255	2.8678
3.0	2.9604	2.9103
3.25	2.9792	2.9396
3.5	2.9892	2.9586
3.75	2.9944	2.9718
4.0	2.9971	2.9806

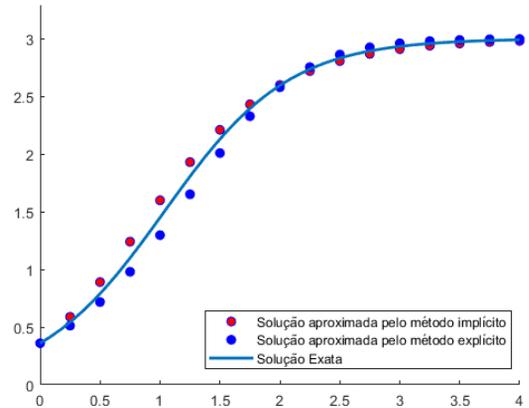


Figura 11: Soluções exata e numérica (Euler explícito e implícito) para (12) com  $h = 0.25$

$$EI(i+1) = \text{bissec}(F, EI(i) - 0.5, EI(i) + 0.5, 0.001);$$



Figura 12: Construção do algoritmo.

Tabela 2: Aproximações numéricas via (EE) e (EI) utilizando a bissecção com dados (11).

Com essa discussão em mente, atualizamos o laço do algoritmo da Figura 5, como apresentado na Figura 13. Assim, podemos tomar  $h$  menor para obter melhor precisão.

```

n0= ; r= ; K= ;
EE(1)=n0;
for i= 1 : n
    EE(i+1)= EE(i)+h*r*EE(i)*(1-EE(i)/K);
end

EI(1)=n0; syms x; F(x)=x-n0-r*h*x*(1-x/K);
for i= 1 : n
    EI(i+1)=bissec(F,EI(i)-0.5,EI(i)+0.5,0.001);
    F(x)=x-EI(i+1)-r*h*x*(1-x/K);
end
    
```

Figura 13: À esquerda temos (EE) e à direita a atualização do algoritmo do método de Euler-implícito (EI). Precisamos determinar  $n0$ ,  $r$  e  $K$ , além de mantermos a grade de tempo relativas a  $h$  e  $T_{final}$  apresentadas na Figura 5. Note que na Figura 5 tem-se que (\*) é o passo em que precisamos determinar o zero de uma função.

Em vez de tomarmos  $h$  pequeno nesse algoritmo, vamos considerar o fato de que esse, de repente, ficou lento. Como lembrado anteriormente, o método de bissecção é lento, e é geralmente usado para iniciarmos outros métodos, como por exemplo, o método de Newton-Raphson, onde a teoria elementar diz que se nosso pontapé inicial for suficientemente próximo do zero da função,

então haverá convergência, e essa costuma ser muito mais rápida. Na Figura 14 apresentamos à esquerda da Figura 10 o algoritmo do método de Newton-Raphson e à direita a utilização do método da bissecção de uma forma mais eficiente: definimos a função bissec\_NR com entradas  $(f, a, b, tol\_1, tol\_2)$ , onde  $f$ ,  $a$  e  $b$  têm o mesmo significado que no método da bissecção,  $tol\_1$  significa o quanto iremos diminuir o intervalo original  $[a, b]$  via o método da bissecção clássico até esse ficar com comprimento menor que  $tol\_1$ , e então aplicamos o método de Newton-Raphson clássico no ponto intermediário  $c$  desse intervalo reduzido.

```
function c=NR(f,x0,tol,itermax)
F=diff(f); i=1; c=x0; A=abs(subs(f,x0));
while A>tol
c=c-sub(F,c)/subs(F,c);
A=abs(subs(f,c));
i=i+1;
if i>itermax
fprintf('Numero de iteracoes maximas atingidas');
error('Não obtivemos raiz, melhorar o chute inicial!')
end
end

function c=bissec_NR(f,a,b,tol_1,tol_2)
d=1;
if subs(f,a)*subs(f,b)>0
error('Melhorar o intervalo!')
end
while d>tol_1
c=(a+b)/2;
if subs(f,a)*subs(f,c)<0
b=c;
end
if subs(f,c)*subs(f,b)<0
a=c;
end
d=abs(subs(f,c));
end
c=NR(f,c,tol_2,100);
```

Figura 14: Algoritmo do método de Newton-Raphson/Algoritmo da bissecção melhorado.

Esses são dois exemplos de métodos que podem ser mais eficientes que o método da bissecção. Vale a pena observar que, dependendo do caso em que estamos estudando, poderíamos usar  $EI(i)$  diretamente como pontapé inicial para obter  $E(i + 1)$  ao utilizar o método de Newton-Raphson, porém, dependendo da natureza da solução, isso pode ser problemático, levando em consideração que tal método tem algumas armadilhas. Assim, o algoritmo como escrito tem seu valor.

```
for i= 1 : n
EI(i+1)=NR(F,EI(i),0.0001,20);
F(x)=x-EI(i+1)-r*h*x*(1-x/K);
end
```

Figura 15:  $(EI)$  usando Newton-Raphson.

```
for i= 1 : n
EI(i+1)=bissec_NR(F,EI(i)-0.5,EI(i)+0.5,0.01,0.001);
F(x)=x-EI(i+1)-r*h*x*(1-x/K);
end
```

Figura 16:  $(EI)$  usando a bissecção melhorada.

Vamos usar esses desenvolvimentos para discutir a solução  $N(t)$  de (10). Exibimos na Figura 17 a aplicação do algoritmo da Figura 15 para a solução em azul e vermelho; já o algoritmo da Figura 16 para as soluções verde e rosa, onde os dados iniciais  $n_0$  e taxas de crescimento  $r$  indicados na legenda; e em todos os casos consideramos  $K = 10$ . Em cada caso onde o dado inicial é positivo, vemos que  $K$  representa o estado estacionário da solução. Nos casos em que  $0 < n_0 < K$ , a solução cresce no limite indo para  $K$ , enquanto que para dados  $n_0 > K$  a solução decresce para  $K$ . Na Figura 18 vemos que  $r$ , representa a taxa com que se dá esse crescimento quando  $0 < n_0 < K$ . Quando maior o valor  $r$  mais rápida a solução aproxima-se de  $K$ .

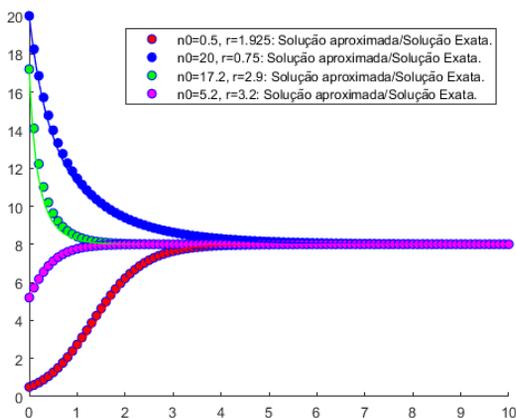


Figura 17: Soluções exatas/numéricas pelo método de Euler-implícito para  $K = 10$  e  $h = 0.1$  no intervalo de tempo  $[0, 10]$  para alguns valores de  $r$  e  $n_0$ .

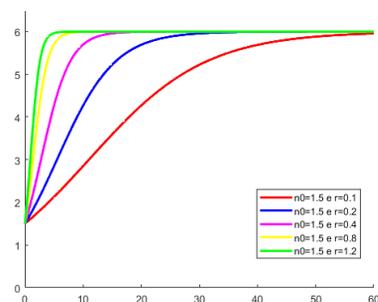


Figura 18: Soluções numéricas pelo método de Euler-implícito para  $K = 8$  e  $h = 0.1$  no intervalo de tempo  $[0, 60]$  para valores de  $r = 0.1, 0.2, 0.4, 0.8, 1.2$  e  $n_0 = 1.5$ .

A teoria da transição demográfica junto com modelos matemáticos eficientes ajudam a descrever a sequência pela qual as sociedades são transformadas. Não há nenhuma contradição entre o fato de a taxa de natalidade decair e o fato de a população crescer, mesmo que em um ritmo menos acelerado. Isso ocorre porque, devido ao aumento da contracepção as taxas de natalidade devem cair, porém, devido aos efeitos de uma população jovem, o crescimento ainda deve acontecer por mais algumas décadas. Há uma considerável preocupação com as consequências do crescimento da população humana para o meio ambiente e para o desenvolvimento social e econômico. É concluído em [9] que esse crescimento provavelmente chegará ao fim em um futuro previsível, mostrando que há cerca de 85% de chance de que a população mundial pare de crescer antes do final do século e que existe a probabilidade de 60% de que a população mundial não exceda 10 bilhões de pessoas antes de 2100, e cerca de 15% de probabilidade de que a população mundial no final do século seja menor do que 6 bilhões. Para diferentes regiões, a data e o tamanho da população de pico variam consideravelmente. Dados que chamam mais a atenção estão voltados para a região da China, onde a probabilidade de atingir um pico nas próximas duas décadas ainda é baixa devido à sua estrutura etária relativamente jovem, e na África Subsaariana, onde apesar da prevalência do HIV, há uma baixa probabilidade de atingir um pico antes da metade do século. A probabilidade chega a 25% em 2070, 50% em 2085 e quase 75% em 2100, devido a supostas reduções na fertilidade.

Estudos mais recentes (2019, Organização das Nações Unidas) mostram que, num contexto geral, desde 1900 a população mundial vem crescendo exponencialmente. Estima-se que, no início da década de 50, o número de pessoas no planeta era de aproximadamente 2,6 bilhões de pessoas. Seguindo o crescimento, a população mundial chegou em 5 bilhões de habitantes em julho de 1987, e 6 bilhões em outubro de 1999. Já em 2010 esse número já estaria em aproximadamente 7 bilhões de pessoas. Para o futuro, a projeção é que no ano de 2030 ter-se-iam 8,5 bilhões de pessoas no planeta. Fazendo análises sucessivas, para o ano de 2050 são previstos 9,7 bilhões de habitantes e, por fim, no ano de 2100 serão 11,2 bilhões de pessoas. Assim como toda projeção, tem-se um certo grau de incerteza quanto a concretização de tais fatos. Os dados acima citados são baseados em uma projeção média variante, ou seja, existem alguns fatores que podem vir a influenciar nesses valores. Alguns deles são, por exemplo, a taxa de fertilidade, onde projeta-se que a quantidade de

filhos por mulher passe de 2,5 em 2019 para 2,2 em 2050; a expectativa de vida, onde espera-se que passe de 72,6 anos em 2019 para 77,1 anos em 2050; e as taxas de migração que, ainda não tão significante como os dois anteriormente citados, também pode gerar uma mudança no número populacional do planeta.

O crescimento exacerbado vem causando sérias conseqüências à população do planeta. Cada vez mais faltam recursos à saúde, ocorrem migrações de refugiados para países próximos, acelera-se a urbanização, aumenta a demanda por habitações, e ocorre a precariedade no abastecimento de água potável e alimentos, assim como a falta de saneamento básico, entre outros. Ainda em [9], também é feita uma análise semelhante, onde ressalta que na ausência de uma crise de mortalidade catastrófica, um aumento adicional na população mundial de 7 bilhões para quase 9 bilhões é inevitável, o que sugere que políticas apropriadas devam ser adotadas e programas implementados a fim de acelerar a transição demográfica. Além dos benefícios ambientais de uma transição acelerada para a estabilização da população na Ásia e na África, grandes benefícios de saúde, sociais e econômicos seriam alcançados. O número de mulheres que morrem no parto seria cortado, e a saúde e a educação das crianças melhorariam. A proporção de adultos para crianças dependentes aumentaria, criando oportunidade de investir mais na agricultura e na indústria. A promoção da contracepção e legitimação de famílias menores precisa ser uma das principais prioridades nessa região.

Durante a maior parte da história da humanidade, a população mundial cresceu de forma lenta dado o fato de os coeficientes de mortalidade serem muito elevados. Tais coeficientes foram, sobretudo, influenciados pelo frequente surgimento de epidemias e pandemias, que acometeram grandes parcelas da população, além de guerras que dizimaram milhões de pessoas, como é o caso das duas grandes guerras que ocorreram durante o Século XX. Alguns estudos (ver [14], por exemplo) apontaram que, geralmente, essas epidemias e pandemias foram ocasionadas em sua maioria por precariedade de higiene pessoal e saneamento básico, falta de conhecimentos médicos e baixos padrões de condições de vida. Assim, com uma grande baixa da população mundial, seja por doenças ou por guerras, o crescimento da mesma, que seguia um ritmo exponencial, pode ter um declínio considerável. Nos dias de hoje, as grandes epidemias e pandemias tornaram-se fatos de interesse mundial frente a situação de epidemia de Coronavírus (Covid-19), oficialmente declarada como pandemia em Março de 2020. O que muitos não sabem, no entanto, é que desde a antiguidade tem sido relativamente comum se ter registros desses eventos. Podem-se citar grandes epidemias e pandemias que foram surgindo durante toda a história e que influenciaram no decréscimo da população, tais como a Peste Negra (1346-1353) que assolou a Europa Medieval, matando entre 20 e 25 milhões de pessoas, o que era equivalente a 25% da população europeia na época; a “Gripe Espanhola” (1918-1920), que foi considerada uma das pandemias mais mortais já constatadas, sendo que cerca de 3% a 5% da população mundial foi morta pelo vírus gripal da H1N1 (ver [6]); a epidemia do Ebola (2007 e 2013-2016), que ocorreu sobretudo nos países da África, sendo um vírus altamente infeccioso e que pode acabar matando mais de 90% das pessoas que a contraem<sup>1</sup>, e, atualmente, o Covid-19 (também chamado de Coronavírus) preocupa a população mundial por ser um vírus de fácil transmissão e que tem, cada dia mais, mortes contabilizadas por sua conta.

Então, como proposto em (3) e (4), o modelo (10) requer tal ajuste, e comum na literatura (ver

<sup>1</sup>Ver <https://www.msf.org.br/o-que-fazemos/atividades-medicas/ebola>.

[11, 1, 12, 5, 2, 8, 10]) é esses serem dados por

$$\begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) - P \left(1 - e^{-\frac{N^2}{A}}\right) \\ N(0) = n_0, \end{cases} \quad \text{ou} \quad \begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) - \frac{BN^2}{A^2 + N^2} \\ N(0) = n_0, \end{cases}$$

onde estamos considerando em (3) e (4) o caso  $F \equiv 1$ . No caso de um efeito de predação acentuado, consideraremos  $F$  como nas figuras 1 e 2.

Os métodos de Euler-explícito e Euler-implícito são bastante interessantes para fixar as ideias, porém não são de uso comum para resolver um problema de valor inicial com uma EDO, esse da forma

$$\begin{cases} \frac{dN}{dt} = f(N, t) \\ N(0) = n_0, \end{cases} \quad (16)$$

já que na literatura existem métodos mais acurados e que não têm problemas com erros acumulados. O estudo do método implícito, sobretudo, é interessante também para fixar as ideias tendo em vista que em problemas de ordem maior ou em problemas de valor inicial com uma EDP são predominantes os métodos implícitos, veja [4] para vários exemplos. Dentre os métodos que apresentam melhorias aos métodos implícito/explicito de Euler podemos citar o método de Heun, o método do ponto médio ou o método de Ralston, que são todos estendidos pelos métodos de Runge-Kutta:

A fim de aproximar numericamente a solução de (16) utilizaremos aqui o clássico **método de Runge-Kutta de quarta ordem**, que é o de uso mais comum no dia a dia, que, segundo [4], é dado por

$$N_{i+1} = N_i + \frac{h}{6} (K_1 + 2K_2 + 2K_3 + K_4),$$

onde

$$\begin{aligned} K_1 &= f(t_i, N_i), & K_2 &= f\left(t_i + \frac{h}{2}, N_i + \frac{K_1 h}{2}\right), \\ K_3 &= f\left(t_i + \frac{h}{2}, N_i + \frac{K_2 h}{2}\right) & \text{e} & \quad K_4 = f(t_i + h, N_i + K_3 h). \end{aligned}$$

No que segue utilizaremos então o método para a análise numérica dos problemas de valores iniciais propostos, onde o algoritmo está destacado na Figura 19.

```
Tfinal= 25; h=0.1;
T= 0:h:Tfinal; n=Tfinal/h;
K= ; n0= ; r= ; N(1)=n0;
syms t y; f(t,y)=
for i=1 : n
    K1=f(i*h,N(i));
    K2=f(i*h+h/2, N(i)+K1*h/2);
    K3=f(i*h+h/2,N(i)+K2*h/2);
    K4=f(i*h+h,N(i)+K3*h);
    N(i+1)=N(i)+(h/6)*(K1+2*K2+2*K3+K4);
end
plot(T,N,'r','LineWidth',2)
```

Figura 19: Algoritmo utilizado para Runge-Kutta de 4ª ordem.

\* Utilizaremos então para

$$f(t, y) = r * (1 - y/K) * y - F(t) * P * (1 - \exp(-y^2/A))$$

e

$$f(t, y) = r * (1 - y/K) * y - F(t) * P * N^2 / (A^2 + N^2)$$

e testaremos alguns valores para  $A$  e  $P_0$  e  $F(t)$  conforme apresentadas nas figuras 1 e 2.

Discutiremos numericamente nesse trabalho apenas o modelo (3), o modelo (4) segue de maneira similar. Consideremos então o problema de valor inicial

$$\begin{cases} \frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right) - PF(t) \left(1 - e^{-\frac{N^2}{A}}\right) \\ N(0) = n_0, \end{cases} \quad (17)$$

Num primeiro exemplo exibimos os resultados numéricos utilizando o método de Runge-Kutta considerando  $F(t) \equiv 1$ . Nas figuras 20 e 21 exibimos essa análise para alguns valores de A e  $P_0 = P$ , onde consideramos para isso  $K = 8$ ,  $r = 1.1$  e  $n_0 = 1$ .

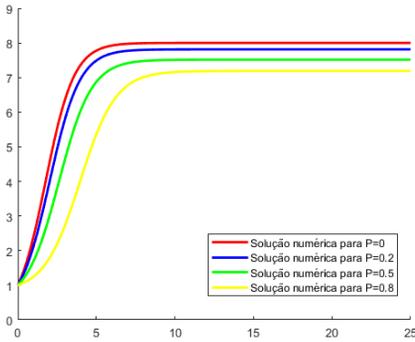


Figura 20: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando  $K = 8$ ,  $r = 1.1$ ,  $n_0 = 1$ ,  $A = 0.1$  e P como na figura.

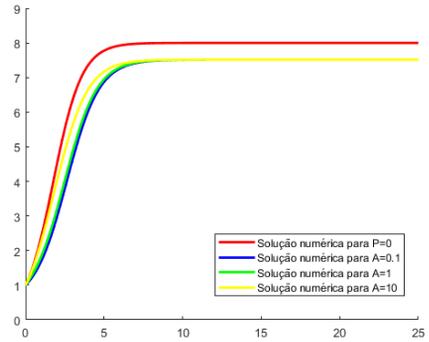


Figura 21: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando  $K = 8$ ,  $r = 1.1$ ,  $n_0 = 1$ ,  $P = 0.5$  e A como na figura.

Vale a pena observar que o resultado obtido com esse modelo ainda pode ser visto como um bom ajuste para dados como abordados nas figuras 8 e 9 para a população americana e população francesa ao longo dos últimos séculos. Como um exemplo ilustrativo, consideremos os dados colhidos e fornecidos por *Statistisches Bundesamt (Federal Statistical Office of Germany)* para a população alemã e pelo *Istituto Nazionale di Statistica* para a população italiana ao longo dos anos de 1900 a 2020. Exibimos nas figuras 22 e 23 os gráficos de ambas as populações ao longo desses anos. É possível ver os efeitos estudados influenciando na dinâmica populacional desses países nos períodos da Primeira Guerra Mundial (1914-1918) e Segunda Guerra Mundial (1939-1945).

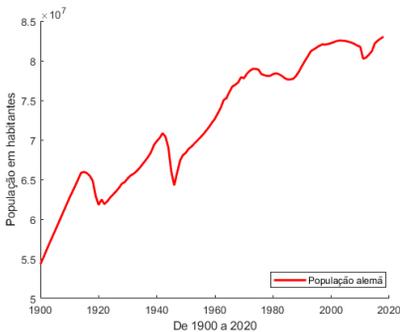


Figura 22: Gráfico da população alemã de 1900 a 2020.

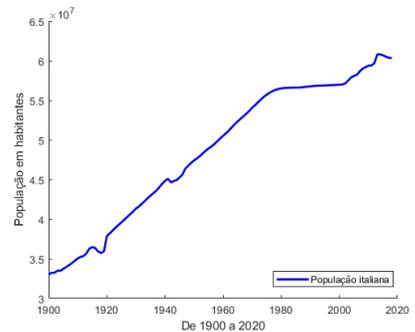


Figura 23: Gráfico da população italiana de 1900 a 2020.

Talvez, isso fique mais claro se fizermos uma aproximação grosseira para um ajuste usando o modelo logístico (10), onde para Alemanha consideremos  $r = 0.055$ ,  $n_0 = 54.326.000$  e  $K = 80.500.000$ , e para a Itália consideremos  $r = 0.047$ ,  $n_0 = 33.004.000$  e  $K = 57.000.000$ .

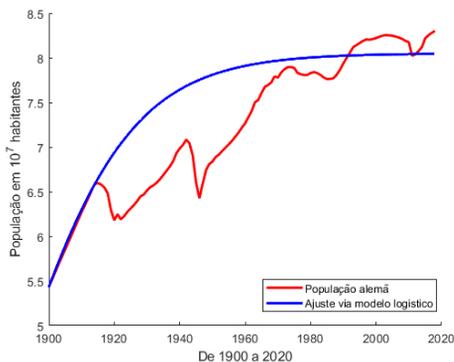


Figura 24: Ajuste ilustrativo para a população alemã com o modelo (10).

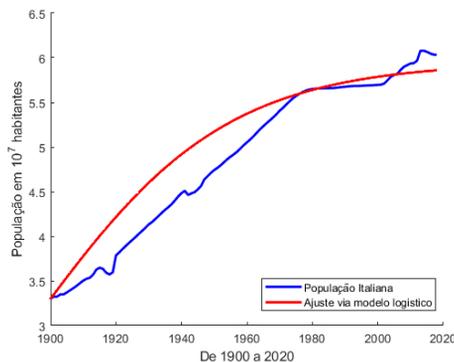


Figura 25: Ajuste ilustrativo para a população italiana com o modelo (10).

Provavelmente, essas análises ficariam ainda mais interessantes se considerássemos fatores de migração e imigração, que são muito comuns na história da Itália, por exemplo. A imigração italiana teve como ápice o período entre 1880 e 1930, enquanto que recebeu imigrantes devido às profundas mudanças econômicas e sociais induzidas pela industrialização do pós-guerra, incluindo baixas taxas de natalidade e envelhecimento da população. Consequentemente, com uma força de trabalho encolhida, durante os anos 80, a Itália passou a atrair fluxos crescentes de imigrantes estrangeiros. Podemos observar esses efeitos na Figura 25. Desde a queda do Muro de Berlim em 1989 e, mais recentemente, as ampliações da União Europeia em 2004 e 2007, as principais ondas de migração vieram dos antigos países da Europa Oriental, como Romênia, Albânia, Ucrânia e Polônia. O segundo fator mais importante de migração para a Itália é ser vizinho ao norte da África, recebendo muitos imigrantes do Marrocos, Egito e Tunísia, por exemplo. Além disso, nos últimos anos, foram registrados fluxos crescentes de migração do Extremo Oriente, como China e Filipinas e também de países da América Latina, como Equador e Peru.

Então, a fim de ver os efeitos que esses impactos na taxa de predação podem fazer no crescimento populacional de uma comunidade, faremos então a abordagem considerando  $F$  da seguinte forma:

$$F(t) = 1 + ce^{-\frac{(t-t_0)^2}{d}},$$

com parâmetros  $c \in \mathbb{R}$ ,  $d > 0$ , onde já abordamos alguns exemplos nas figuras 1 e 2. Esperamos que com o efeito acentuado dado por  $F$  tendo ponto de máximo ( $c > 0$ ) haja retardo no crescimento da população ou até a reduza, e com ponto de mínimo ( $c < 0$ ) aumente a população, como acontece quando há imigração, por exemplo. Lembramos que o parâmetro  $c$  representa a intensidade adicional que o termo de predação apresentará, e  $d$  podemos entender como o tempo em que esse efeito acentuado ocorre. No primeiro exemplo exibido na Figura 26 consideramos  $F(t) = 7.2e^{-(t-3.9)^2/0.1}$ . Observe que em azul está a solução com  $F \equiv 1$ , donde podemos perceber que o efeito acentuado resultou num retardo no crescimento em certo período de tempo próximo a  $t = 3.9$  e com raio 0.1. Logo em seguida, na Figura 27, aumentamos o período de acentuação

considerando  $F(t) = 7.2e^{-(t-3.9)^2/0.8}$ , e, na Figura 28, exibimos o resultado do aumento da intensidade para  $F(t) = 11.6e^{-(t-3.9)^2/0.8}$ . Em todos esses exemplos consideramos os mesmos parâmetros  $r = 1.1$ ,

$$K = 8, \quad n_0 = 1, \quad A = 0.1, \quad P = 0.2, \quad (18)$$

No quarto experimento, exibido na Figura 29, consideramos também os dados (18),  $r = 1.1$  e para a acentuação  $F$  consideramos  $F(t) = 1 + 12e^{-(t-0.4)^2/0.001} + 4e^{-(t-4)^2/0.4} + 4e^{-(t-10)^2/2}$  onde indicamos que ao longo do tempo aconteceram três eventos que influenciaram diretamente no crescimento populacional. Nas figuras 30 e 31 exibimos os últimos dois experimentos com dados (18) e, no primeiro caso consideramos  $r = 0.825$  e  $F(t) = 1 + 12e^{-(t-0.4)^2/0.01} + 8e^{-(t-4)^2/0.4} + 4e^{-(t-10)^2/2} - 4.5e^{-(t-25)^2/15}$  e no segundo consideramos  $r = 0.685$  e  $F(t) = 1 + e^{-(t-5)^2/10} - 2e^{-(t-22)^2/30}$ , onde finalmente consideramos os efeitos de migração aumentando a população.

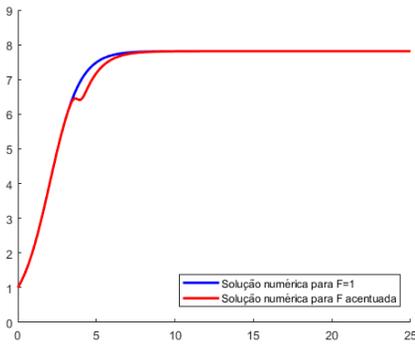


Figura 26: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando (18),  $r = 1.1$  e  $F(t) = 1 + 7.2e^{-(t-3.9)^2/0.1}$ .

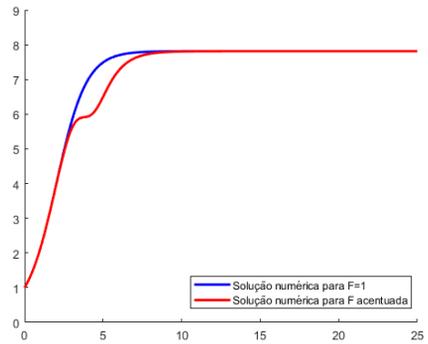


Figura 27: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando (18),  $r = 1.1$  e  $F(t) = 1 + 7.2e^{-(t-3.9)^2/0.8}$ .

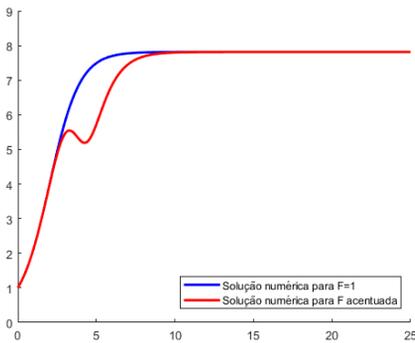


Figura 28: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando (18),  $r = 1.1$  e  $F(t) = 1 + 11.6e^{-(t-3.9)^2/0.8}$ .

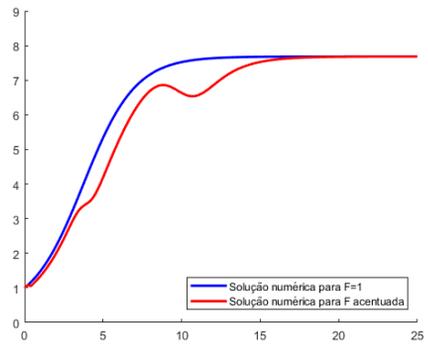


Figura 29: Solução numérica de (17) pelo método de Runge-Kutta de 4ª ordem considerando (18),  $r = 0.66$  e  $F(t) = 1 + 12e^{-(t-0.4)^2/0.001} + 4e^{-(t-4)^2/0.4} + 4e^{-(t-10)^2/2}$ .

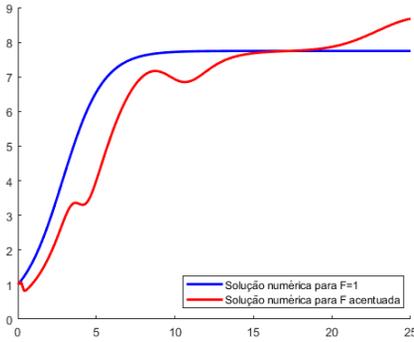


Figura 30: Solução numérica de (17) pelo método de Runge-Kutta de 4<sup>ª</sup> ordem considerando (18),  $r = 0.825$  e  $F(t) = 1 + 12e^{-(t-0.4)^2/0.01} + 8e^{-(t-4)^2/0.4} + 4e^{-(t-10)^2/2} - 4.5e^{-(t-25)^2/15}$ .

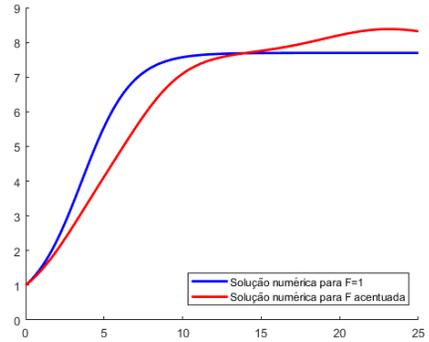


Figura 31: Solução numérica de (17) pelo método de Runge-Kutta de 4<sup>ª</sup> ordem considerando (18),  $r = 0.685$  e  $F(t) = 1 + e^{-(t-5)^2/10} - 2e^{-(t-22)^2/30}$ .

Vale a pena observar que, apesar de conseguirmos incluir tais efeitos no nosso modelo, essas técnicas não são de comum uso e é muito mais conveniente fazer uma aproximação do crescimento populacional de uma comunidade via o clássico modelo logístico (10). Assim, um ótimo exercício para concluir a atividade é fazer um ajuste como feito nas figuras 24 e 25 usando técnicas semelhantes a de quadrados mínimos com a solução exata apresentada à direita de (10), onde o objetivo seria encontrar obter os valores  $K$  e  $r$  adequados. O modelo logístico (10) foi introduzido em 1837 por Pierre F. Verhulst. Esse modelo foi comprovado como adequado para algumas espécies populacionais em modelos populacionais estáveis e experiências de laboratório. Citamos de [7] o exemplo do caso da população de focas no estado de Washington (EUA). No início do século passado, esses animais eram caçados por um programa do governo americano que via a espécie como um perigo, o que conseqüentemente reduziu o número de focas. Desde o fim do programa, a população de focas do estado de Washington recuperou-se seguindo um padrão aproximadamente logístico, sendo que o tamanho da população varia quando se atinge a capacidade de carga, com leves flutuações para cima e para baixo. Assim, vemos que apesar de seus problemas de limitações, ainda assim o uso desses métodos pode ser de grande utilidade para o desenvolvimento de teorias para a Ecologia.

## Referências

- [1] R. Angelini. *Ecossistemas e modelagem ecológica. Perspectivas da Limnologia no Brasil*, v. 1, pp1-19, 1999.
- [2] A. Hastings, *Population Biology - Concepts and Models*, Springer 1996
- [3] F.R.M. BARBOSA, M.P. CANALLI, “Qual a importância da relação professor-aluno no processo ensino-aprendizagem?” *EFDeportes.com, Revista Digital*. Buenos Aires, Ano 16, n. 160, 2011.
- [4] L.R. Burden, J.D. Faires. *Análise Numérica*, 8<sup>a</sup> edição. [S.l.]: São Paulo, 2003.
- [5] C. Cannings, F. Hoppensteadt, *Mathematical methods of population biology*. Cambridge University Press 1982
- [6] J.N. Hays. *Epidemics and pandemics. Their impacts on Human History*. Austin, Texas: Fundação Kahle, 2005.

- [7] H. Huber, J. Laake. Trends and Status of Harbor Seals in Washington State: 1978–99. *AFSC Quarterly Report*, USA, pp1–13, 1 dez. 2002. (Archive)
- [8] M. Kot, *Elements of Mathematical Ecology*, Cambridge University Press, 2001
- [9] W. Lutz, W.C. Sanderson, S.Scherbov, *The end of world population growth in the 21st century: new challenges for human capital formation and sustainable development*. 1st Edition, 2004.
- [10] P. Magal, S. Ruan, *Structured Population Models in Biology and Epidemiology*. Springer, 2008
- [11] J.D. Murray, *Mathematical Biology I - An Introduction*. Third Edition, Springer, 2000.
- [12] S.P. Otto, T.Day. *A biologist's guide to mathematical modeling in ecology and evolution*. New Jersey: Princeton University Press, 2007.
- [13] E.M.S. SOARES,; L.Z. SAUER, “Um novo olhar sobre a aprendizagem de matemática para a engenharia”. In: Cury, H.N. *Disciplinas matemáticas em cursos superiores: reflexões, relatos, propostas*. Porto Alegre: EDIPUCRS, 2004.
- [14] J. Yunes. “A dinâmica populacional dos países desenvolvidos e subdesenvolvidos”. *Revista de Saúde Pública*, [s.l.], v. 5, n. 1, pp129–150, jun. 1971. Fap-Unifesp (SciELO).

Bruno Telch dos Santos  
Universidade do Estado de Santa Catarina  
<[telchbruno@gmail.com](mailto:telchbruno@gmail.com)>

Carolina Cominato  
Universidade do Estado de Santa Catarina  
<[cominatoc@gmail.com](mailto:cominatoc@gmail.com)>

Rodrigo Cardoso dos Santos  
Universidade Federal do Rio de Janeiro  
<[rodrigo\\_cardoso\\_dos\\_santos@hotmail.com](mailto:rodrigo_cardoso_dos_santos@hotmail.com)>

Recebido: 13/04/2020  
Publicado: 17/07/2019

# Aplicação do Teorema do Valor Médio ao Ensino Médio

Edney Freitas Gregorio

Nícolas Alcântara de Andrade

## Resumo

Neste trabalho temos por objetivo apresentar o teorema do valor médio como recurso a ser utilizado em provas e demonstrações de algumas afirmações matemáticas. A relevância do trabalho reside na consequência extremamente simples do teorema, ancorada a demonstrações interessantes para alguns teoremas famosos e importantes do ensino médio.

**Palavras-chave:** Teorema do Valor Médio; Binômio de Newton; Lei de Moivre; Diferença de Potências de Mesmo Expoente; Desigualdade de Bernoulli; Desigualdade de Napier.

## Abstract

In this work we aim to present the mean value theorem as a resource to be used in proofs and demonstrations of some mathematical statements. The relevance of the work lies in the fact that an extremely simple consequence of the theorem, anchored to interesting demonstrations for some famous and important high school theorems.

**Keywords:** Mean Value Theorem; Newton's Binomial; Moivre's Law; Difference of Powers with Same Exponents; Bernoulli Inequality; Napier Inequality

## 1. Introdução

O Teorema do Valor Médio é um dos principais resultados estudados nos cursos de cálculo e análise no ensino superior.

Por sua vez, no ensino médio, dentre outros resultados de destaques, temos:

- (i) O Teorema do Binômio de Newton, que permite calcular a expansão binomial e possui aplicações notadamente na teoria da probabilidade.
- (ii) A Lei de Moivre, que nos fornece uma maneira prática e rápida de calcular as potências de números complexos.
- (iii) A fatoração da diferença de potências de mesmo expoente.
- (iv) Desigualdade de Bernoulli, que permite obter uma cota inferior para potências de  $(1 + x)$ .
- (v) Desigualdade de Napier, que nos dá informações sobre a regularidade da função  $\ln x$ .

Neste artigo vamos apresentar uma demonstração de (i), (ii), (iii), (iv) e (v) a partir do Teorema do Valor Médio, mostrando, assim, que esses resultados, os quais aparentemente não possuem relação, estão ligados entre si.

## 2. Teorema do Valor Médio

O Teorema do Valor Médio, ou Teorema de Lagrange, é um dos principais resultados abordados nos cursos de Cálculo. Apesar de ser um assunto geralmente estudado no ensino superior, ele também costuma ser estudado em turmas do 3º ano do ensino médio voltadas às olimpíadas e preparatórias para vestibulares como ITA e IME. Como o teorema do valor médio aborda conceitos da geometria analítica, como coeficiente angular e equação da reta, dessa forma ele pode ser aplicado no ensino médio, na resolução de alguns exercícios específicos que envolvem tais conceitos de Geometria Analítica.

Esse teorema foi provado pela primeira vez por Lagrange e, posteriormente, Cauchy apresentou uma forma mais geral dele. Tal resultado, obtido a partir do Teorema de Rolle, foi peça fundamental para a demonstração do Teorema Fundamental do Cálculo.

Podemos enunciá-lo da seguinte forma:

**Teorema 1** (Teorema do Valor Médio (TVM)). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função contínua em  $[a, b]$  e diferenciável em  $(a, b)$ . Então, existe pelo menos um  $c \in (a, b)$  tal que:*

$$\frac{f(b) - f(a)}{b - a} = f'(c). \quad (1)$$

Tal resultado apresenta a seguinte interpretação geométrica: se uma função  $y = f(x)$  é contínua em  $[a, b]$  e derivável em  $(a, b)$ , então existe pelo menos um ponto  $c$  entre os pontos  $a$  e  $b$  onde a reta tangente à curva é paralela à reta suporte da corda que une os pontos  $A(a, f(a))$  e  $B(b, f(b))$ , conforme a figura abaixo.

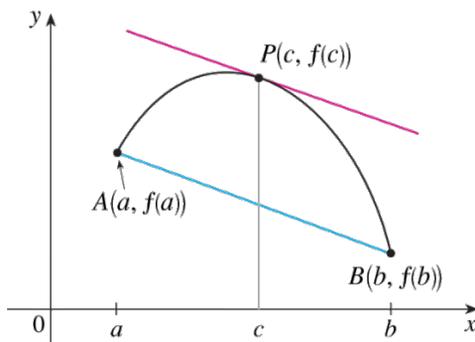


Figura 1: TVM (Teorema do Valor Médio).

Como o teorema do Valor Médio está relacionado à inclinação da reta tangente a uma curva qualquer, vê-se que uma das condições necessárias é que a função que define a curva seja diferenciável

apenas no intervalo aberto  $(a, b)$ , o que equivale a afirmar que a função tem que ser diferenciável apenas nos pontos interiores do intervalo  $[a, b]$ , ou seja, diferenciável no interior de  $[a, b]$ . Vale observar ainda que o ponto  $c$  cuja existência é garantida pelo teorema pode não ser único, como vemos na imagem abaixo.

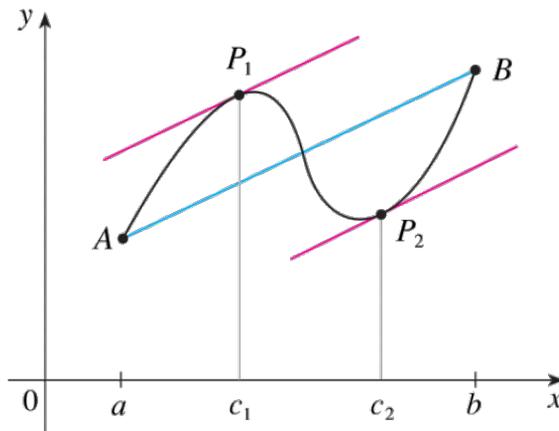


Figura 2: TVM (Teorema do Valor Médio).

O Teorema de Lagrange também pode ser interpretado em termos físicos: supondo que um objeto encontra-se em movimento, e se a sua velocidade média for  $v$  durante um determinado percurso, digamos que do ponto  $a$  ao ponto  $b$ ; então, durante esse percurso, ou seja, no intervalo  $[a, b]$ , existe (pelo menos) um momento em que a sua velocidade instantânea (daqui o conceito de derivada) também é  $v$ .

Tal teorema possui várias aplicações importantes, em particular destacamos o seguinte corolário:

**Corolário 1.** *Seja  $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$  uma função diferenciável. Se  $f'(x) = 0, \forall x \in (a, b)$ , então  $f$  é constante.*

No que segue, usaremos o corolário acima para dar demonstrações alternativas para resultados clássicos do ensino médio.

### 3. Aplicação ao Ensino Médio

#### 3.1. Binômio de Newton

O desenvolvimento de potências com expoente natural de um binômio é conhecido como Binômio de Newton ou Teorema Binomial. Dessa forma, podem-se desenvolver expressões do tipo  $(x + y)^n$ , com  $n$  natural. Tal resultado é amplamente utilizado em conexão com cálculos de probabilidade envolvendo eventos mutuamente exclusivos, onde não importa a ordem de ocorrência de tais eventos.

**Teorema 2** (Binômio de Newton). *Sejam  $x, y \in \mathbb{R}$  e  $n \in \mathbb{N}$ , então:*

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}. \quad (2)$$

*Demonstração.* Inicialmente, provaremos que para todo  $x$  real vale:

$$(x + 1)^n = \sum_{k=0}^n \binom{n}{k} x^k. \quad (3)$$

De fato, defina  $f : \mathbb{R} \rightarrow \mathbb{R}$  dada por

$$f(x) = (x + 1)^n - \sum_{k=0}^n \binom{n}{k} x^k. \quad (4)$$

Observe que  $f$  é um polinômio; portanto,  $f \in C^\infty$ . Assim, derivando  $f(x)$  obtemos para todo  $i$  com  $1 \leq i \leq n$ :

$$f^{(i)}(x) = \frac{n!}{(n-i)!} (x + 1)^{n-i} - \sum_{k=i}^n \binom{n}{k} \frac{k!}{(k-i)!} x^{k-i}.$$

Onde  $f^{(i)}$  é a derivada  $n$ -ésima de  $f$  em relação a  $x$ .

Daí, fazendo  $i = n$  na equação acima, obtemos, para todo  $x$  real:

$$\begin{aligned} f^{(n)}(x) &= \frac{n!}{(n-n)!} (x + 1)^{n-n} - \sum_{k=n}^n \binom{n}{k} \frac{k!}{(k-n)!} x^{k-n} \\ &= n! - \binom{n}{n} \frac{n!}{0!} x^0 \\ &= 0. \end{aligned}$$

Além disso, para todo  $i$ , com  $1 \leq i < n$ , temos que:

$$\begin{aligned} f^{(i)}(0) &= \frac{n!}{(n-i)!} (0 + 1)^{n-i} - \sum_{k=i}^n \binom{n}{k} \frac{k!}{(k-i)!} 0^{k-i} \\ &= \frac{n!}{(n-i)!} - \binom{n}{i} \frac{i!}{(i-i)!} 0^0 \\ &= 0. \end{aligned}$$

Onde aqui usamos que  $\lim_{x \rightarrow 0^+} x^x = 1$

*Afirmção 1.* Se  $f^{(i)}(x) \equiv 0$ , então,  $f^{(i-1)}(x) \equiv 0$ .

De fato, se  $f^{(i)}(x) = 0$  para todo  $x$ , então, pelo corolário 1 do TVM,  $f^{(i-1)}$  é constante. Como, porém,  $f^{(i-1)}(0) = 0$ , segue que  $f^{(i-1)}(x) = 0$  para todo  $x$  real.

Dessa forma, como  $f^{(n)}(x) \equiv 0$ , segue que  $f^{(i)}(x) \equiv 0$  para todo  $i$  entre 1 e  $n$ . Em particular,  $f'(x) = 0$  para todo  $x$  real, o que nos diz que  $f$  é constante. Como, porém,  $f(0) = 0$ , segue que  $f(x) \equiv 0$ . Finalmente temos:

$$0 = f(x) = (x + 1)^n - \sum_{k=0}^n \binom{n}{k} x^k.$$

E, portanto:

$$(x + 1)^n = \sum_{k=0}^n \binom{n}{k} x^k.$$

Para o caso geral, basta fazer a substituição de  $x$  por  $\frac{x}{y}$ , onde  $y \neq 0$ . O que nos dá:

$$\left(\frac{x}{y} + 1\right)^n = \sum_{k=0}^n \binom{n}{k} \left(\frac{x}{y}\right)^k$$

$$\left(\frac{x + y}{y}\right)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{-k}.$$

O que nos diz que

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$$

□

*Observação 1.* O caso particular  $n = 2$  possui uma interpretação geométrica interessante.

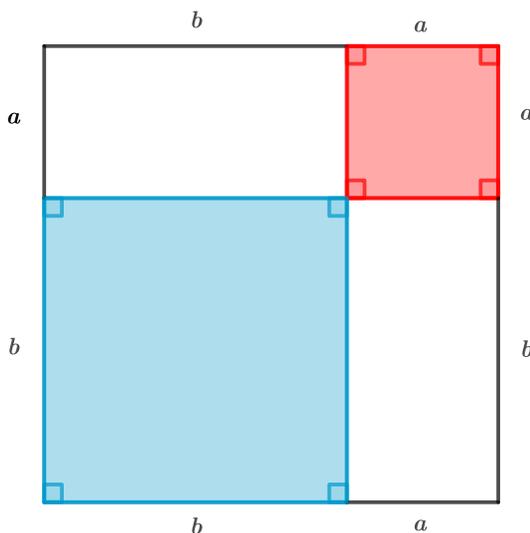


Figura 3: Quadrado da soma de dois termos positivos

Como a figura é um quadrado de lado  $a + b$ , então, sua área é  $(a + b)^2$ , por outro lado sua área também é igual à soma das áreas do quadrado de lado  $b$ , do quadrado de lado  $a$  e de dois retângulos congruentes de lados  $a$  e  $b$ . Totalizando  $a^2 + 2ab + b^2$ , por fim, sabemos que os dois resultados calculam a mesma área, logo,

$$(a + b)^2 = a^2 + 2ab + b^2.$$

### 3.2. Lei de Moivre

A Primeira Lei de Moivre estabelece as vantagens computacionais da representação polar de números complexos e, além disso, simplifica bastante a obtenção do resultado de uma potência  $z^n = (x + iy)^n$ , já que, para expoentes grandes, resultaria numa decepcionante distribuição de um binômio de Newton.

**Teorema 3** (Fórmula de Moivre). *Seja  $z = r(\cos x + i \sin x) \in \mathbb{C}$ , com  $r > 0$ ,  $x \in \mathbb{R}$ . Então, para todo  $n \in \mathbb{Z}$ , temos que:*

$$z^n = r^n(\cos(nx) + i \sin(nx)). \quad (5)$$

*Demonstração.* Basta provar que

$$(\cos x + i \sin x)^n = (\cos(nx) + i \sin(nx)).$$

Defina  $f: \mathbb{R} \rightarrow \mathbb{C}$  dada por

$$f(x) = (\cos(x) + i\sin(x))^n - (\cos(nx) + i\sin(nx)).$$

Temos que  $f$  é diferenciável, pois é combinação de funções diferenciáveis. Assim, calculando a derivada de  $f$ , obtemos:

$$\begin{aligned}
 f'(x) &= n(\cos x + i\sin x)^{n-1}(-\sin x + i\cos x) - n(-\sin(nx) + i\cos(nx)) \\
 &= n(\cos x + i\sin x)^{n-1}(i\sin x + \cos x) - ni(\sin(nx) + i\cos(nx)) \\
 &= ni(\cos x + i\sin x)^{n-1}(\cos x + i\sin x) - ni(\cos(nx) + i\sin(nx)) \\
 &= ni(\cos x + i\sin x)^n - ni(\cos(nx) + i\sin(nx)) \\
 &= ni[(\cos x + i\sin x)^n - (\cos(nx) + i\sin(nx))] \\
 &= nif(x).
 \end{aligned}$$

*Afirmção 2.* Afirmamos que  $f(x) = 0$  para todo  $x \in \mathbb{R}$ .

De fato, pelo visto acima:

$$\begin{aligned}
 f'(x) &= nif(x) \\
 e^{-nix}f'(x) &= e^{-nix}nif(x) \\
 0 &= e^{-nix}f'(x) - e^{-nix}nif(x) \\
 0 &= (e^{-nix}f(x))'.
 \end{aligned}$$

Para todo  $n \in \mathbb{Z}$  e para todo  $x \in \mathbb{R}$ . Assim, pelo corolário 1 do teorema do valor médio aplicado à parte real  $\text{Re}(f)$  e imaginária  $\text{Im}(f)$  de  $f$ , temos que  $\text{Re}(f)$  e  $\text{Im}(f)$  são constantes e, portanto, a própria  $f$  é constante. Daí, como  $f(0) = 0$ , segue que  $e^{-nix}f(x) = 0$  para todo  $n \in \mathbb{Z}$  e para todo  $x \in \mathbb{R}$ . Como  $e^{-nix}$  nunca se anula, segue que  $f(x) \equiv 0$ .

Finalmente,

$$\begin{aligned}
 0 &= f(x) \\
 &= (\cos x + i\sin x)^n - (\cos(nx) + i\sin(nx)) \\
 (\cos x + i\sin x)^n &= (\cos(nx) + i\sin(nx)).
 \end{aligned}$$

Para todo  $n \in \mathbb{Z}$  e para todo  $x \in \mathbb{R}$ . □

*Observação 2.* No caso particular para  $n = 2$ ,  $(\cos x + i\sin x)^2 = \cos(2x) + i\sin(2x)$  podemos obter as fórmulas para o arco duplo do seno e do cosseno. Como as partes reais e complexas de ambos os lados da igualdade devem ser iguais, então, após os cálculos obtemos

$$\sin(2x) = 2\sin x \cos x \quad \text{e} \quad \cos(2x) = \cos^2 x - \sin^2 x.$$

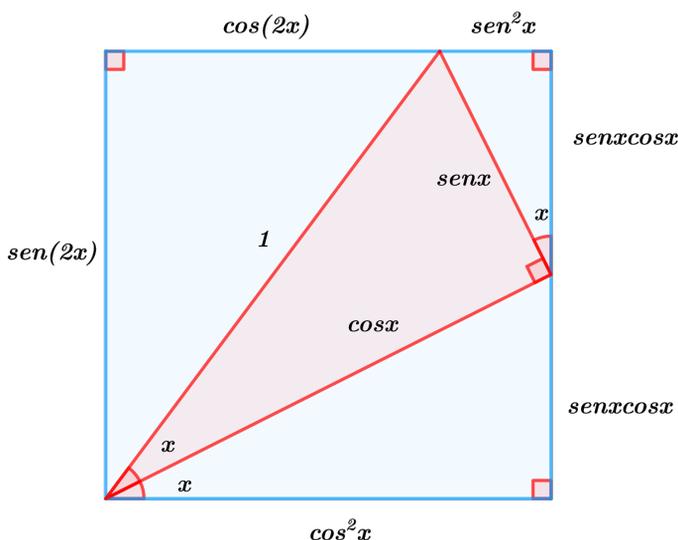


Figura 4: Seno e cosseno do arco duplo para  $0 < x < \frac{\pi}{2}$ .

### 3.3. Diferença de Potências de Mesmo Expoente

As fatorações constituem uma ferramenta imprescindível para a Matemática, em se tratando de cálculos algébricos. As aplicações envolvem desde o ensino básico até o superior com fatorações importantes para o cálculo diferencial e integral e para a Teoria dos Números.

**Teorema 4.** *Sejam  $x, y \in \mathbb{R}$  e  $n \in \mathbb{N}$ , então:*

$$x^n - y^n = (x - y) \sum_{k=0}^{n-1} x^k y^{n-1-k}. \quad (6)$$

*Demonstração.* Provaremos inicialmente que para todo  $x$  real vale:

$$x^n - 1 = (x - 1) \sum_{k=0}^{n-1} x^k. \quad (7)$$

Para isso, defina  $f: \mathbb{R} \rightarrow \mathbb{R}$  dada por  $f(x) = x^n - 1 - (x - 1) \sum_{k=0}^{n-1} x^k$ . Note que  $f$  é diferenciável, pois é um polinômio, daí, derivando  $f(x)$   $i$  vezes ( $1 \leq i \leq n - 1$ ), obtemos:

$$f^{(i)}(x) = \frac{n!}{(n-i)!} x^{n-i} - i \sum_{k=i-1}^{n-1} \frac{k!}{(k-(i-1))!} x^{k-(i-1)} - (x-1) \sum_{k=i}^{n-1} \frac{k!}{(k-i)!} x^{k-i}.$$

Note que  $f^{n-1}(x) \equiv 0$ . De fato, para todo  $x$  real:

$$f^{n-1}(x) = n!x - (n-1)((n-2)! + (n-1)!x) - (x-1)(n-1)! = 0.$$

Além disso, observe que  $f^{(i)}(0) = 0$  para  $1 \leq i \leq n-1$ . De fato:

$$f^{(i)}(0) = -i(i-1)! - (-1)i! = -i! + i! = 0.$$

Finalmente, perceba que se  $f^{(i)}(x) = 0$ , então,  $f^{(i-1)}(x) \equiv 0$ . Ora, tal fato é imediato, pois  $f^{(i)}(x) = 0$  nos diz que  $f^{(i-1)}$  é constante pelo corolário 1 do teorema do valor médio. Como, porém, vimos que  $f^{(i-1)}(0) = 0$ , segue que  $f^{(i-1)}(x) \equiv 0$ . Dessa forma, temos que  $f', f'', \dots, f^{(n-1)}$  são todas identicamente nulas; portanto, a própria  $f$  é identicamente nula, uma vez que, em particular,  $f'(x) \equiv 0$ , o que nos diz que  $f$  é constante. Como, porém,  $f(0) = 0$ , segue que  $f(x) = 0$  para todo  $x$  real. O que nos diz que:

$$0 = f(x)$$

$$0 = x^n - 1 - (x-1) \sum_{k=0}^{n-1} x^k$$

$$x^n - 1 = (x-1) \sum_{k=0}^{n-1} x^k.$$

Para o que falta, substitua  $x$  por  $\frac{x}{y}$ , onde  $y \neq 0$ , em (7), obtendo:

$$\left(\frac{x}{y}\right)^n - 1 = \left(\frac{x}{y} - 1\right) \sum_{k=0}^{n-1} \left(\frac{x}{y}\right)^k$$

$$x^n - y^n = (x-y) \sum_{k=0}^{n-1} x^k y^{n-1-k}.$$

□

*Observação 3.* No caso particular para  $n = 2$ ,  $x^2 - y^2 = (x-y) \sum_{k=0}^1 x^k y^{1-k}$  representa a fórmula para a diferença de quadrados.

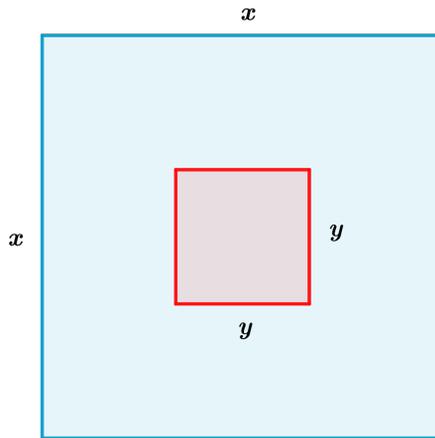


Figura 5: Diferença de dois quadrados de termos positivos.

A Figura 5 é formada por dois quadrados concêntricos de lados paralelos medindo  $x$  e  $y$  ( $x > y > 0$ ). A área da região destacada em azul é a diferença entre a área do quadrado maior e do menor que é igual a  $x^2 - y^2$ ; por outro lado, ela também é igual à soma das áreas de quatro trapézios azuis congruentes que são formados ligando os vértices dos lados paralelos dos quadrados. Como cada trapézio tem altura  $\frac{x-y}{2}$ , base maior  $x$  e base menor  $y$ , segue que a área deles quatro é

$$4 \frac{\left(\frac{x-y}{2}\right)(x+y)}{2} = (x-y)(x+y).$$

Por fim, sabemos que os dois resultados representam a mesma área, logo.

$$x^2 - y^2 = (x-y)(x+y).$$

### 3.4. Desigualdade de Bernoulli

A desigualdade de Bernoulli é uma relação que permite obter aproximações de potências de  $(1+x)$  por um polinômio do primeiro grau. Tal relação foi publicada pela primeira vez por Bernoulli em 1689, e, desde então, tem sido usada frequentemente em várias áreas, como combinatória ou no estudo da convergência de seqüências.

**Teorema 5.** Se  $x > -1$ , então,  $(1+x)^r \geq 1+xr$  para todo  $r \in \mathbb{N}$ .

*Demonstração.*

- (i) Caso  $x \geq 0$ : Considere  $f(t) = (1+t)^n$ , para  $t \in [0, x]$ . Assim,  $f$  satisfaz as hipóteses do teorema do valor médio, e isso garante a existência de um  $c \in (0, x)$  com  $f(x) - f(0) = (x-0)f'(c)$ . Assim, temos  $(1+x)^n - 1 = nx(1+c)^{n-1} \geq nx$
- (ii) Caso  $-1 < x < 0$ : É feito de forma análoga, considerando  $f(t) = (1+t)^n$ , para  $t \in [x, 0]$ .

□

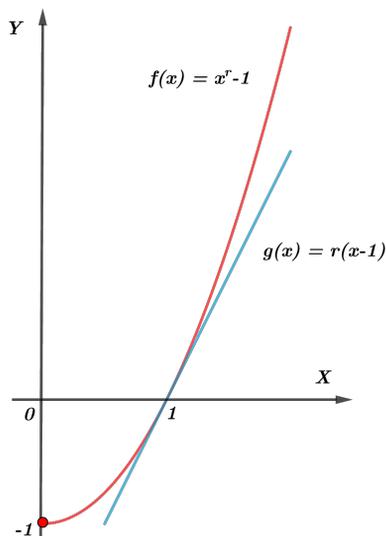


Figura 6: Desigualdade de Bernoulli para  $x > 0$ ,  $x \neq 1$ ,  $r > 1 \rightarrow x^r - 1 > r(x - 1)$

### 3.5. Desigualdade de Napier

A desigualdade de Napier constitui uma ferramenta não muito conhecida, porém bastante útil. Ela foi descoberta por John Napier, reconhecido por seus trabalhos envolvendo logaritmos naturais. Dessa forma não é de se espantar que a desigualdade de Napier possa ser usada para dar uma demonstração bem elementar da diferenciabilidade, e conseqüente continuidade, da função  $f(x) = \ln x$ .

**Teorema 6.** Se  $b > a > 0$ , então,  $\frac{1}{b} < \frac{\ln b - \ln a}{b - a} < \frac{1}{a}$ .

*Demonstração.* Considere  $f(x) = \ln x$ , para  $x \in [a, b]$ . Assim,  $f$  satisfaz as hipóteses do teorema do valor médio, e isso garante a existência de  $c \in (a, b)$ , ou seja,  $\frac{1}{b} < \frac{1}{c} < \frac{1}{a}$ , com  $f(b) - f(a) = f'(c)(b - a)$ . Assim, temos:

$$\frac{\ln b - \ln a}{b - a} = \frac{1}{c} \quad \Rightarrow \quad \frac{1}{b} < \frac{\ln b - \ln a}{b - a} < \frac{1}{a}.$$

□

Para uma interpretação geométrica, considere as retas  $L_1, L_2, L_3$  onde seus coeficientes angulares são  $m_{L_1} < m_{L_2} < m_{L_3}$

$$m_{L_1} = \frac{1}{b}, \quad m_{L_2} = \frac{\ln b - \ln a}{b - a}, \quad m_{L_3} = \frac{1}{a}.$$

### 4. Considerações Finais

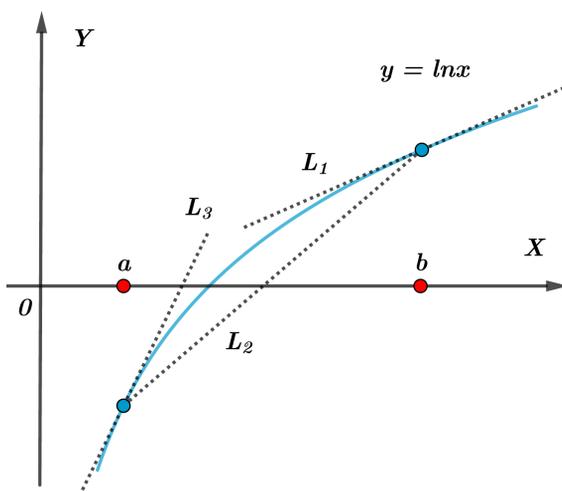


Figura 7: Desigualdade de Napier.

Como vimos, graças ao corolário 1 do Teorema do valor médio, foram apresentadas uma demonstração não indutiva para a primeira lei de Moivre, uma demonstração alternativa para o Binômio de Newton, diferença de potências de mesmo expoente, desigualdade de Bernoulli e desigualdade de Napier. Ressaltamos que um conjunto muito variado de identidades matemáticas pode ser demonstrado por meio desse corolário, em particular, e um exercício simples e interessante para treinar o método abordado aqui seria usá-lo para demonstrar a famosa relação trigonométrica fundamental. Por fim, esperamos que este trabalho possa suscitar outros estudos nesse sentido.

## Referências

- [1] T. Apostol, *Calculus, Volume 1*, John Wiley & Sons, Nova Iorque, 1967.
- [2] E. L. Lima, *Curso de Análise, Volume 1*, Impa, Rio de Janeiro, 2000.
- [3] R.B. Nelsen, *Proofs without words exercises in visual thinking. Volume 1*, MAA, 1993.
- [4] P.K. Sahoo, *Mean Value Theorems and functional equations*, World Scientific, 1998
- [5] M. Soares, *Cálculo em Uma Variável Complexa, Quinta Edição*, SBM, Rio de Janeiro, 2016.

Edney Freitas Gregorio  
Universidade Estadual do Ceará  
<[edney.gregorio@uece.br](mailto:edney.gregorio@uece.br)>

Nícolas Alcântara de Andrade  
Universidade Estadual do Ceará  
<[nicolas.andrade@uece.br](mailto:nicolas.andrade@uece.br)>

Recebido: 11/05/2020  
Publicado: 30/07/2020

# Aplicação de modelos não lineares na precificação de passagens aéreas

Juliana de Melo 

Roberto Molina de Souza 

## Resumo

Devido à autonomia dos usuários de transporte aéreo no que diz respeito à compra de bilhetes, a necessidade de intermediários, como agências de turismo, tornou-se desnecessária em muitos casos, sendo do usuário a tomada de decisão quanto à companhia aérea e tempo de antecedência da compra de sua passagem. Adquirir um bilhete aéreo com antecedência pode tornar a compra mais barata. Nesse sentido, este trabalho estuda o comportamento dos preços praticados pelas três maiores empresas aéreas brasileiras nos trechos Londrina/PR - Curitiba/PR e Londrina/PR - São Paulo/SP quanto ao tempo de antecedência na aquisição do bilhete. São apresentados gráficos do preço médio ao longo de trinta e sessenta dias de estudo, nos quais se observa que a relação entre trinta dias de antecedência e o valor cobrado é não linear, motivando assim o estudo de modelos não lineares para a modelagem dos dados obtidos, visto que com sessenta dias de antecedência, no geral, os preços estabilizam. Logo, para os trechos em estudo recomenda-se a compra da passagem com trinta dias de antecedência.

**Palavras-chave:** Modelos não lineares; Preços de passagens aéreas; Dias de antecedência.

## Abstract

Due to the autonomy of air transport users with regard to ticket purchase, the need for intermediaries, such as travel agencies, has become unnecessary in many cases, being the user's decision regarding the airline and time in advance of your ticket purchase. Acquiring an air ticket in advance can make the purchase cheaper. In this sense, this work studies the behavior of the prices practiced by the three largest Brazilian airlines in the Londrina / PR - Curitiba / PR and Londrina / PR - São Paulo / SP flight stretch regarding the time in advance in ticket acquisition. Average price graphs are presented during thirty and sixty days. This study shows that the relation between thirty days in advance and the amount charged is non-linear, thus motivating the study of non-linear statistical models for the modeling of the data obtained, since at sixty days in advance prices generally stabilize. Therefore, for the flight stretch under study it is recommended to purchase the ticket thirty days in advance.

**Keywords:** Nonlinear models; airline tickets; days in advance

## 1. Introdução

A partir do surgimento da rede mundial de computadores como relevante canal de distribuição e comercialização dos bilhetes aéreos, diminuiu-se a necessidade da intermediação de agentes de viagens, reduzindo os custos de pesquisa (*search costs*) e também as assimetrias de informação no mercado. Observou-se, assim, um aumento na competitividade do setor aéreo brasileiro [15].

Segundo Evans e Kessides (1993) [10] existem duas categorias nos estudos de análise empírica, que são aproximações para o estudo dos elementos que influenciam o nível das tarifas aéreas. A primeira categoria é a análise *inter-rota*, em que se inferem os efeitos das variáveis de estrutura de mercado em preços *cross-sectional* nas rotas, ou seja, os dados estão desagregados no nível da rota e referem-se a variáveis como distância de voo, número de companhias presentes, entre outras. A segunda categoria é a de análise *intra-rota*, em que se introduz a heterogeneidade das empresas que operam em uma rota como as características específicas das companhias, quantidades de assentos, entre outros.

A gestão de oferta também conhecida como *Yield Management (YM)* é uma técnica de gestão baseada em modelos matemáticos utilizada para calcular a menor política de preços, por meio da otimização e maximização dos lucros gerados pela venda de um determinado produto ou serviço. O *YM* surgiu na década de 80, nos Estados Unidos, com a desregulamentação da indústria aérea americana e a tecnologia disponível a um custo aceitável, ou seja, as companhias aéreas puderam precificar seus assentos como bem entendessem. Além disso, houve abertura para novas companhias aéreas de custo mais baixo entrarem no mercado, aumentando a concorrência [14].

As primeiras empresas que investiram fortemente em *YM* como sua principal arma para manter a lucratividade continuaram no mercado, sendo elas: *American Airlines* e a *Delta Airlines*. Essas empresas criaram grandes sistemas centrais nas quais tinham o potencial de reduzir a concorrência, baseando-se em três situações: a dificuldade de empresas aéreas que não possuem sistemas centrais teriam para visualizar a tela de um agente de viagem; as cláusulas restritivas de contrato entre as agências de viagem, e as empresas proprietárias do sistema [22].

A primeira empresa da América Latina com o modelo de negócios "*Low Cost, Low Fare*" foi a Gol Linhas Aéreas, com entrada no mercado doméstico em janeiro de 2001, dando início a uma trajetória de crescimento impressionante do transporte aéreo mundial. A forte expansão da empresa provocou um incremento da contestabilidade ao mercado aéreo e uma maior competitividade do setor. Fatores como preços menores, publicidade e acesso ao aeroporto de Congonhas desde o primeiro ano de operações (2001) e ao aeroporto de Santos Dumont e à Ponte Aérea Rio de Janeiro - São Paulo desde o segundo ano (2002) contribuíram para o crescimento acelerado da Gol [16].

Em geral, existem fatores que influenciam na variação dos preços cobrados pelos bilhetes, como dia da semana, horário do voo, a empresa que se escolhe para voar e a antecedência da compra da passagem em relação ao voo etc.

Quanto à antecedência da compra, algumas indefinições podem ser consideradas [5]. A expectativa mais comum é de que os preços cobrados cresçam conforme a data do embarque se aproxima. É, porém, pertinente observar que a aproximação da data do voo coloca a empresa aérea em uma situação de risco.

Nesse caminho, o objetivo deste trabalho é estudar o menor valor diário de passagens aéreas em um período de 30 e 60 dias de antecedência da compra do bilhete, dos trechos Londrina/PR - Curitiba/PR (LDB-CWB) e Londrina/PR - São Paulo/SP (LDB-CGH), considerando as três maiores empresas aéreas que fazem esses trechos. Os dados coletados para a rota Londrina - São Paulo consideraram o aeroporto de Congonhas.

Este trabalho está organizado da seguinte maneira: na Seção 2 são apresentados os conceitos de modelos de regressão linear e não linear, e, nas subseções, um resumo dos métodos de estimativa dos parâmetros e critério de comparação dos modelos; na Seção 3 são apresentados gráficos que resumem o conjunto de dados obtido, os modelos escolhidos para a estimação dos parâmetros, as aplicações e os resultados, e, finalmente, na Seção 4 as considerações finais.

## 2. Modelos de Regressão

A análise de regressão tem o objetivo de verificar a existência de uma relação funcional entre uma variável dependente com uma ou mais variáveis independentes. Em outras palavras, a análise de regressão consiste na obtenção de uma equação que busca explicar a variável dependente pela variação dos níveis das variáveis independentes [7]. Modelos estatísticos são capazes de explicar essa relação por meio de uma equação que relaciona tais variáveis.

### 2.1. Regressão linear

Na análise de regressão linear simples, os dados são modelados utilizando uma função linear do tipo  $f(x) = \alpha + \beta x + \varepsilon$ , sendo os parâmetros  $\alpha$  e  $\beta$  desconhecidos e estimados a partir dos dados  $\varepsilon$  representa o que não é explicado pela variável independente, conhecido como erro aleatório, seguindo alguma distribuição de probabilidade.

Na análise de regressão linear múltipla, mais de uma variável independente pode ser considerada. Essa relação pode ser analisada como um processo no qual os valores das variáveis  $X_1, X_2, \dots, X_n$  podem ser chamados de variáveis regressoras, e  $Y$  de variável resposta [7]. Logo, tais modelos são da forma  $f(x) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$ .

### 2.2. Regressão não linear

Os modelos lineares são bastante utilizados para um grande número de aplicações em regressão, cujo propósito é relacionar uma variável resposta a uma ou mais variáveis explicativas. Todavia, em algumas situações um modelo não linear pode ser mais adequado, uma vez que muitos fenômenos geralmente não são lineares.

O surgimento dos modelos não lineares deu-se por meio da regressão geral na década de 20 com Ronald Fisher Aylmer e Winifred Mackenzie. Nos anos 70, esses modelos ganharam mais popularidade e passaram a ser mais bem estudados devido aos avanços computacionais. Com exceção de alguns trabalhos mais antigos, a maior parte da literatura referente a modelos não lineares surgiu depois dos anos 80, e, desde então, tem sido uma área de pesquisa bastante explorada [21].

Com o surgimento de novos algoritmos para ajustes de modelos de regressão não lineares, houve o aumento da popularidade dos modelos não lineares e a valorização dos processos inferenciais por eles gerados. Os modelos não lineares são usados por pesquisadores para modelar variáveis nas mais diversas áreas do conhecimento, como econometria, agricultura, biologia, química etc [21].

Dentre as diferenças entre as classes de modelos lineares e não lineares, a principal delas está relacionada as suas formulações. Por meio de um conjunto de observações, o modelo linear busca o modelo que melhor explique a relação, se existir alguma, entre as variáveis inerentes a um dado fenômeno [13]. Já no caso não linear, quase sempre, as formulações de possíveis modelos são baseadas em considerações teóricas referentes ao fenômeno de interesse em modelar [3, 19].

Um modelo de regressão é dito não linear se ao menos um dos seus parâmetros aparece de forma não linear. Por exemplo, os modelos:

$$y = \exp(\alpha + \theta x) \tag{1}$$

e

$$y = \alpha + \beta \exp(\theta x) \tag{2}$$

são não lineares. No modelo (1)  $\alpha$  e  $\theta$  são parâmetros não lineares, e no modelo (2)  $\alpha$  e  $\beta$  são parâmetros lineares, enquanto  $\theta$  é não linear.

Os modelos não lineares podem ser escritos da seguinte forma:

$$y_i = f(x_i, \theta) + \varepsilon_i, \tag{3}$$

em que,  $y_i$  e  $x_i$ , para  $i = 1, 2, \dots, n$ , são os vetores de variável resposta e variável explicativa, respectivamente, para uma amostra de tamanho  $n$ ;  $f(x_i, \theta)$  é uma função das variáveis regressoras e dos parâmetros, chamada função de regressão,  $\theta$  é o vetor dos parâmetros desconhecidos e  $\varepsilon_i$  é o vetor de erros aleatórios. Em geral, atribui-se que os erros são variáveis aleatórias independentes e identicamente distribuídos, seguindo distribuição normal com média zero e variância constante denotada por  $\sigma^2$ .

### 2.3. Método dos mínimos quadrados

O Método dos Mínimos Quadrados (MMQ) é uma técnica utilizada no método iterativo de Gauss-Newton, em que seu objetivo é encontrar o melhor ajuste para um conjunto de dados, cujo objetivo é minimizar a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Considerando que a função de regressão é contínua e admite derivadas de primeira e segunda ordens com relação aos parâmetros, define-se a soma de quadrados dos erros de (3) por:

$$SQRes(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2, \tag{4}$$

em que  $SQRes(\theta)$  depende exclusivamente de  $\theta$ . Denota-se por  $\hat{\theta}$  os estimadores de mínimos quadrados de  $\theta$ , ou seja, os valores de  $\hat{\theta}$  que minimizam  $SQRes(\theta)$ . Para encontrar esses estimadores, é necessário fazer a diferenciação da equação (4) com relação a cada  $\theta_i$  com  $i = 1, 2, \dots, p$  e igualar as equações a zero, isto é:

$$\frac{\partial SQRes(\theta)}{\partial \theta_i} = \sum_{i=1}^n -2(Y_i - f(X_i, \theta)) \left[ \frac{\partial f(X_i, \theta)}{\partial \theta_i} \right]. \tag{5}$$

Igualando as derivadas parciais definidas em (5), a zero, tem-se:

$$\sum_{i=1}^n Y_i \left[ \frac{\partial f(X_i, \theta)}{\partial \theta_i} \right] - \sum_{i=1}^n f(X_i, \theta) \left[ \frac{\partial f(X_i, \theta)}{\partial \theta_i} \right] = 0. \tag{6}$$

Dessa forma, determina-se  $p$  equações denominadas de equações normais ou homogêneas, sendo definidas por:

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \left[ \frac{\partial f(x_i, \theta)}{\partial \theta_i} \right]_{\theta=\hat{\theta}} = 0, \quad (7)$$

e quando  $\partial f(x_i, \theta)$  não depender de  $\theta$ , ou seja,  $\left[ \frac{\partial^2 f(x_i, \theta)}{\partial \theta_i^2} \right] = 0$ , tem-se as equações de um modelo de regressão linear. Para funções não lineares,  $\theta$  estará presente em pelo menos uma das derivadas parciais de  $f(x_i, \theta)$ .

Para modelos lineares esse sistema normalmente possui um vetor único de soluções para os parâmetros  $\theta$ , em que cada parâmetro depende unicamente dos dados observados. Já para modelos não lineares, os estimadores  $\hat{\theta}$  resultantes da ultima equação, em geral, não apresentam forma analítica, e as soluções precisam ser obtidas por algum método de aproximação iterativo. Dentre os principais, pode-se citar o método de Gauss-Newton [21] descrito a seguir.

### 2.3.1 Método iterativo de Gauss-Newton

Para minimizar  $SQRes(\theta)$  descrito em (4), pode ser necessário o uso de algum método numérico iterativo. O método iterativo de Gauss-Newton permite estabelecer uma relação linear entre a função não linear e os dados observados. Essa relação é feita com base na expansão em série de Taylor.

Suponha que  $\theta^{(k)}$  seja uma aproximação da estimativa de mínimos quadrados  $\theta$  de  $f(X_i, \theta)$ . Tomando  $\theta \approx \theta^{(k)}$ , tem-se que a expansão em série de Taylor de primeira ordem é dada por:

$$f(X_i, \theta) \approx f(X_i, \theta^{(k)}) + \frac{\partial f(X_i, \theta)}{\partial \theta_j} (\theta - \theta^{(k)}), \quad (8)$$

em que a matriz com as derivadas da função não linear em relação a cada um dos parâmetros,  $\frac{\partial f(X_i, \theta)}{\partial \theta_j}$ , será denotada por  $F(\theta)$ . Seja  $r(\theta) = Y_i - f(X_i, \theta)$  sendo o vetor de resíduos, na qual podem ser aproximados por:

$$r(\theta) \approx r(\theta^{(k)}) - F(\theta^{(k)}) (\theta - \theta^{(k)}). \quad (9)$$

Como o objetivo é minimizar a  $SQRes = r(\theta)' r(\theta)$  pelo método de mínimos quadrados, em que  $r(\theta)'$  é o vetor transposto de  $r(\theta)$ , tem-se:

$$SQRes = [r(\theta)' r(\theta)] \approx [r(\theta^{(k)}) - F(\theta^{(k)}) (\theta - \theta^{(k)})]' [r(\theta^{(k)}) - F(\theta^{(k)}) (\theta - \theta^{(k)})], \quad (10)$$

$$\begin{aligned}
 SQRes &= r(\theta^{(k)})' r(\theta^{(k)}) - r(\theta^{(k)})' F(\theta^{(k)}) (\theta - \theta^{(k)}) - \\
 & r(\theta^{(k)}) [F(\theta^{(k)}) (\theta - \theta^{(k)})]' + (\theta - \theta^{(k)})' F(\theta^{(k)})' F(\theta^{(k)}) (\theta - \theta^{(k)}).
 \end{aligned}$$

Dessa forma,  $SQRes$  será minimizada quando:

$$(\theta - \theta^{(k)}) = [F(\theta^{(k)})' F(\theta^{(k)})]^{-1} F(\theta^{(k)})' r(\theta^{(k)}). \quad (11)$$

A cada iteração é gerada uma nova aproximação  $\theta^{(k+1)}$  para o vetor de parâmetros  $\theta^{(k)}$ , através de:

$$\theta^{(k+1)} = \theta^{(k)} + [F'(\theta^{(k)}) F(\theta^{(k)})]^{-1} F'(\theta^{(k)}) r(\theta^{(k)}), \quad (12)$$

resultando no processo iterativo conhecido como Método de Gauss-Newton. O processo (12) é repetido até obter a convergência.

Para a estimação dos parâmetros e ajuste dos modelos, pode ser utilizada a função *nls* do software livre R, em que o algoritmo padrão é Gauss-Newton. Essa função tem como estrutura os principais argumentos:

$$\text{nls}(\text{formula}; \text{data}; \text{start}; \text{algorithm}),$$

em que *formula* é a fórmula matemática do modelo que será utilizado; *data* é o conjunto de dados a ser utilizado; *start* é uma lista com valores iniciais para cada um dos parâmetros, e *algorithm* define o tipo de algoritmo utilizado. O algoritmo padrão é o de Gauss-Newton, já o algoritmo de Golub-Pereyra é utilizado para modelos parcialmente lineares, e a opção *port* é para um algoritmo que considera restrições nos parâmetros do modelo.

## 2.4. Critérios de informação

Os modelos estatísticos são simplificações da realidade e, ao fazer sua seleção, é preciso ter em mente que não existem modelos perfeitos, há apenas modelos aproximados da realidade na qual causam perda de informações. Dessa forma, para fazer a seleção do melhor modelo, dentre aqueles que foram ajustados, são utilizados os critérios de informação [9].

### 2.4.1 Critério de Akaike - AIC

Para a comparação dos ajustes de modelos pode ser utilizado o critério de Informação de Akaike [2]. O critério de Informação de Akaike foi formulado pelo estatístico japonês Hirotugu Akaike (1927-2009). Esse critério é um estimador da qualidade relativa de modelos estatísticos para um dado conjunto de dados. A partir de uma coleção de modelos para os dados, o AIC estima a qualidade de cada modelo, em relação a cada um dos outros modelos fornecendo assim um meio para a seleção dos mesmos. Entre vários modelos candidatos, deve ser escolhido aquele que apresentar o menor valor de AIC.

Akaike (1974) [2] mostrou que o viés é dado assintoticamente por  $p$ , em que  $p$  é o número de parâmetros a serem estimados no modelo, e definiu seu critério de informação como

$$\text{AIC} = -2 \log L(\hat{\theta}) + 2(p), \quad (13)$$

em que  $\log L(\hat{\theta})$  representa o logaritmo da função de verossimilhança e  $p$  o número de parâmetros desconhecidos e estimados. A obtenção completa do resultado acima pode ser encontrada, por exemplo, em Konishi e Kitagawa (2007)[12] e Emiliano (2009)[8].

De acordo com Sugiura (1978) [20] e Sakamoto *et. al* (1986) [17], o critério de Akaike pode ter um comportamento ruim se existirem muitos parâmetros, em comparação com o tamanho da amostra. Burnham (2002) [6] recomendam usar AIC para selecionar modelos quando o número de observações,  $n$ , for maior ou igual do que pelo menos 40 vezes o número de parâmetros  $p$  [4].

## 2.4.2 Critério de informação de Akaike corrigido - AICc

Hurvich e Tsai (1989) [11] desenvolveram o critério Critério de Akaike Corrigido - (AICc) que é indicado para seleção de modelos nos casos de respostas com distribuição normal e pequenas amostras, ou seja, quando o número de observações,  $n$ , for menor do que pelo menos 40 vezes o número de parâmetros, denotado por  $p$  [4]:

$$AIC_c = -2 \log L(\hat{\theta}) + 2(p) + 2 \frac{p(p+1)}{n-p-1}. \quad (14)$$

## 2.4.3 Critério de informação Bayesiano - BIC

O Critério de informação Bayesiano, também chamado de Critério de Schwarz, proposto por Schwarz (1978)[18], é um critério de avaliação de modelos definido em termos de probabilidade [8].

$$BIC = -2 \log L(\hat{\theta}) + p \log n, \quad (15)$$

em que  $\log L(\hat{\theta})$  representa o logaritmo da função de verossimilhança, e  $p$  o número de parâmetros desconhecidos e estimados.

## 3. Aplicação

O conjunto de dados foi obtido considerando um período ininterrupto de 30 e 60 dias de antecedência. No período 30 dias de antecedência, os dados foram coletados no mês de setembro e novembro de 2017, para os trechos Londrina/PR - Curitiba/PR (LDB-CWB) e Londrina/PR - São Paulo/SP (LDB-CGH) respectivamente, e setembro de 2018, para o período de 60 dias de antecedência, em que a coleta ocorreu simultaneamente para ambos os trechos.

Os dados foram coletados por meio dos *websites* das três maiores companhias aéreas nacionais que fazem os trechos em estudo, denominadas por A, B e C. Considerou-se sempre a passagem com menor valor. Para 30 dias de antecedência, a coleta ocorreu de maneira prospectiva, em que sempre considerava o próximo dia até 30 dias à frente, gerando um total de 30 observações. Para 60 dias de antecedência, a coleta foi feita de 5 em 5 dias, ou seja, coletava-se o menor valor referente ao dia seguinte, e, depois, de 5 dias à frente, até completar 60 dias de antecedência, totalizando 13 observações.

### 3.1. Modelos não lineares

No intuito de explicar a relação entre a quantidade de dias de antecedência do bilhete adquirido e o valor cobrado, foram considerados os cinco modelos apresentados a seguir [1]. Cada modelo é uma função com seus respectivos parâmetros, em que  $x$  é a variável explicativa, ou seja, dias de antecedência, e  $y$  representa a média do menor valor diário, conhecida também como variável resposta.

**Modelo 1** - Modelo de regressão linear simples:

$$y = \alpha + \beta x,$$

em que:  $\alpha$  é coeficiente linear; e  $\beta$  é coeficiente angular

**Modelo 2** - Exponencial:

$$y = \exp(\alpha + \beta x),$$

em que:  $\alpha$  é o deslocamento no eixo x; e:

se  $\beta > 0$

$$x \rightarrow \infty$$

$$y \rightarrow 0,$$

se  $\beta < 0$

$$x \rightarrow -\infty$$

$$y \rightarrow 0.$$

**Modelo 3** - Modelo de crescimento exponencial:

$$y = \alpha + \beta \exp(\gamma x),$$

em que:  $\alpha$  é o deslocamento no eixo y;  $\beta$  a concavidade; e  $\gamma$  a rotação.

**Modelo 4** - Soma de exponenciais:

$$y = \alpha \exp(\beta x) + \gamma \exp(\delta x),$$

em que:  $\alpha$  é o deslocamento no eixo y e a concavidade;  $\beta$  é a rotação;  $\gamma$  é o deslocamento no eixo y, e  $\delta$  é o fechamento da concavidade.

**Modelo 5** - Logístico:

$$y = \alpha + \frac{\beta - \alpha}{1 + \exp\left[\frac{x - \gamma}{\delta}\right]}$$

em que:  $\alpha$  é o máximo da função;  $\beta$  é a assintota horizontal;  $\gamma$  é a assintota vertical; e

se  $\delta > 0$

$$x \rightarrow \infty$$

$$y \rightarrow 0,$$

se  $\delta < 0$

$$x \rightarrow -\infty$$

$$y \rightarrow 0.$$

### 3.2. Resultados

Para exemplificar a estimação dos parâmetros da Tabela 1, tomando como exemplo o Modelo 2,  $y = \exp(\alpha + \beta x)$ , os dois parâmetros ( $\alpha$  e  $\beta$ ) são estimados utilizando o método iterativo de

Gauss-Newton descrito na seção 2.3.1. Logo, seja

$$r_i = y_i - \exp(\alpha + \beta x_i), \quad (16)$$

em que  $i = 1, 2, \dots, 30$ . Na equação (16),  $y_i$  representa a média no menor valor diário em reais para  $x_i$  dias de antecedência.

Denotando-se por  $\theta = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$  o vetor dos parâmetros a serem estimados, tem-se:

$$\theta^{(k+1)} = \theta^{(k)} + \left[ \mathbf{J}^{(k)} \right]^T \mathbf{J}^{(k)} \left[ \mathbf{J}^{(k)} \right]^{-1} \mathbf{J}^{(k)} \mathbf{r}^{(k)}$$

em que  $k$  denota cada iteração do método até a sua convergência.  $\mathbf{J}^{(k)}$  é a matriz jacobiana, dada por:

$$\mathbf{J}^{(k)} = \begin{bmatrix} \frac{\partial r_1}{\partial \alpha} & \frac{\partial r_1}{\partial \beta} \\ \frac{\partial r_2}{\partial \alpha} & \frac{\partial r_2}{\partial \beta} \\ \vdots & \vdots \\ \frac{\partial r_{30}}{\partial \alpha} & \frac{\partial r_{30}}{\partial \beta} \end{bmatrix}$$

em que, nesse caso,

$$\begin{aligned} \frac{\partial r_i}{\partial \alpha} &= \exp(\alpha + \beta x_i) \\ \frac{\partial r_i}{\partial \beta} &= x_i \exp(\alpha + \beta x_i). \end{aligned}$$

A implementação das expressões acima, de forma iterativa, pode ser feita no *software* livre R conforme listagem abaixo:

#### Listagem 1: Script - Método iterativo de Gauss-Newton

```

y<- c(689.77, 675.27, 675.62, 582.04, 586.42, 598.77, 486.61, 488.44, 481.47, 408.01,
      389.36, 394.89, 406.19, 336.89, 333.67, 320.17, 310.47, 311.26, 319.21, 325.06,
      249.57, 237.09, 229.02, 225.92, 230.91, 244.14, 246.54, 191.88, 190.44, 194.46)
x<- seq(1:30)

a<- 6.85 #valor inicial para alfa
b<- -0.1 #valor inicial para beta

est.a<- vector()
est.b<- vector()

for(i in 1:10)
{
  r<- y-exp(a+b*x)
  da<- exp(a+b*x)
  db<- x*exp(a+b*x)
  Jac<- matrix(c(da,db),30,2)

  theta<- matrix(c(a,b),2,1) + ginv(t(Jac)%*%Jac) %*% (t(Jac)%*%r)

  a<- theta[1]
  b<- theta[2]

  est.a[i]<-a
  est.b[i]<-b
}
  
```

Após 10 iterações, pode-se observar a convergência do método a partir da terceira iteração, e os valores das estimativas para  $\alpha$  (a) e  $\beta$  (b) obtidos são apresentados abaixo:

k	1	2	3	4	5	6	7	8	9	10
$\alpha$	6,550	6,577	6,590	6,590	6,590	6,590	6,590	6,590	6,590	6,590
$\beta$	-0.033	-0.045	-0.048	-0.048	-0.048	-0.048	-0.048	-0.048	-0.048	-0.048

De forma mais simples, e como citado na subseção 2.3.1, o *software* R disponibiliza a função *nls* que dispensa a implementação do algoritmo, e pode ser utilizada de forma bastante simples conforme exemplificado na listagem abaixo.

### Listagem 2: Script - Uso da função *nls*

```
#Modelo 1:
MK<- lm(dados$A ~ dados$Dias.consecutivos)

#Modelo 2:
M2= nls(y ~ exp(a + b*x), data = dados, start = list(a =6.85, b = -0.1))

#Modelo 3:
M3= nls(y ~ a + b*exp(gama*x), data = dados, start = list(a =689, b = 40, gama=-0.09))

#Modelo 4:
M4= nls(y ~ a*exp(b*x) + gama*exp(delta*Dias.consecutivos), data = dados,
start = list(a=1000, b=-0.09, gama=100, delta=-0.01))

#Modelo 5:
M5= nls(y ~ a+(b-a)/(1+exp((x-gama)/delta)), data = dados,
start = list(a = 500, b = 1350, gama=0.5, delta=10))
```

A Tabela 1 apresenta um exemplo das estimativas obtidas para cada modelo com seu respectivo erro padrão e valor de AIC, AICc e BIC. Para todos os dois trechos e as três companhias consideradas foram construídas tabelas como esta, porém as mesmas foram omitidas do texto.

Modelos	Parâmetros	Estimativas	AIC	AICc	BIC
1	$\alpha$	641,97(5,01)	21822,28	21822,72	21838,77
	$\beta$	-16,99(0,28)			
2	$\alpha$	6,59(0,009)	21584,37	21584,81	21600,86
	$\beta$	-0,05(0,0007)			
3	$\alpha$	111,48(16,35)	<b>21561,76</b>	<b>21562,68</b>	<b>21583,74</b>
	$\beta$	647,45(12,88)			
	$\gamma$	-0,07(0,004)			
4	$\alpha$	698,99(144,64)	21563,64	21565,24	21591,12
	$\beta$	-0,06(0,01)			
	$\gamma$	58,13(152,17)			
	$\delta$	0,015(0,06)			
5	$\alpha$	149,88(26,02)	21562,46	21564,06	21589,93
	$\beta$	1940,09(1180,6)			
	$\gamma$	-7,28(11,37)			
	$\delta$	10,55(2,39)			

Tabela 1: Estimativa dos parâmetros (erro padrão) dos modelos para o trecho LDB - CWB, empresa A, 30 dias de antecedência.

Optou-se pelo menor valor de AICc para representar os modelos melhores ajustados devido aos tamanhos pequenos de amostra, sendo que esses são apresentados na Tabela 2.

Comp. Aérea/Dias de antec.	Trecho	Modelo
A/30 dias	LDB - CWB	3
A/60 dias	LDB - CWB	4
B/30 dias	LDB - CWB	3
B/60 dias	LDB - CWB	5
C/30 dias	LDB - CWB	2
C/60 dias	LDB - CWB	5
A/30 dias	LDB - CHG	5
A/60 dias	LDB - CHG	5
B/30 dias	LDB - CHG	5
B/60 dias	LDB - CHG	5
C/30 dias	LDB - CHG	3
C/60 dias	LDB - CHG	5

Tabela 2: Modelos mais bem ajustados segundo AICc.

Os modelos mais bem ajustados foram sobrepostos aos valores médios dos trechos em estudo para cada companhia aérea, como mostram as Figuras 1 a 6, em que os gráficos à esquerda são referentes a 30 dias de antecedência, e os da direita, 60 dias de antecedência. O nome da empresa e os respectivos modelos estão na legenda de cada figura.

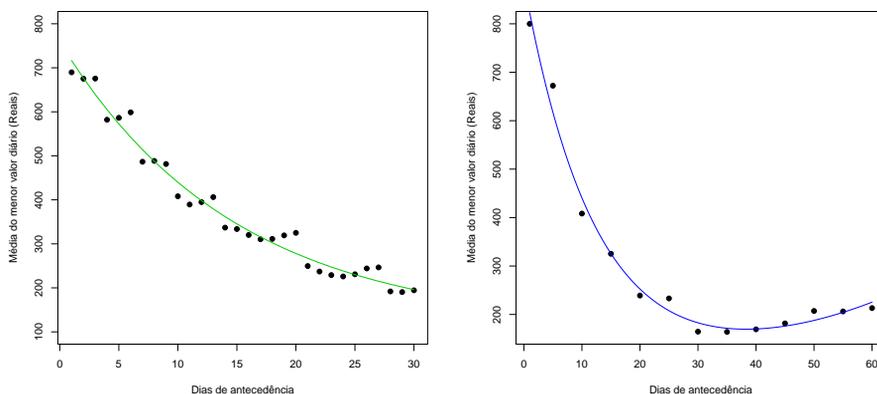


Figura 1: LDB-CWB, empresa A, modelos M3 e M4.

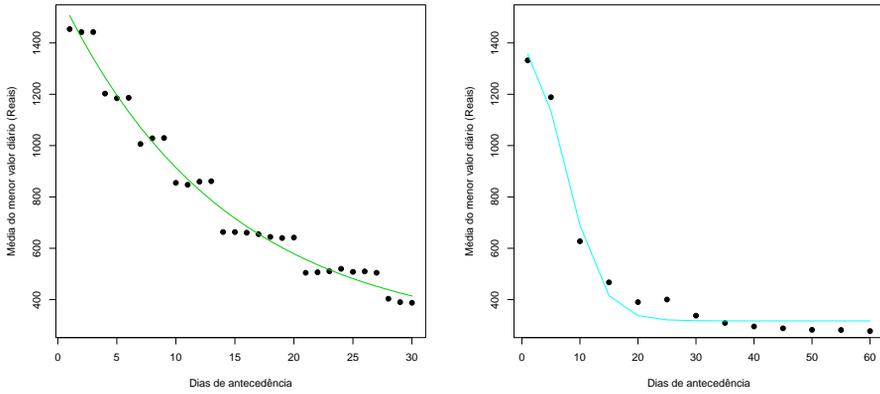


Figura 2: LDB-CWB, empresa B, modelos M3 e M5.

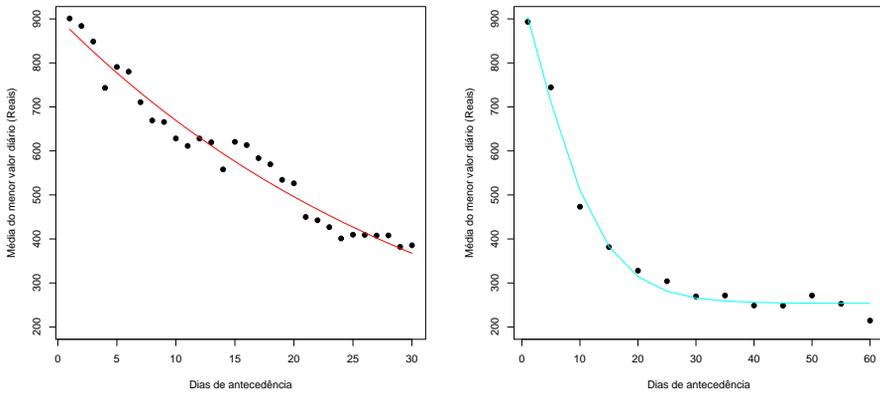


Figura 3: LDB-CWB, empresa C, modelos M2 e M5.

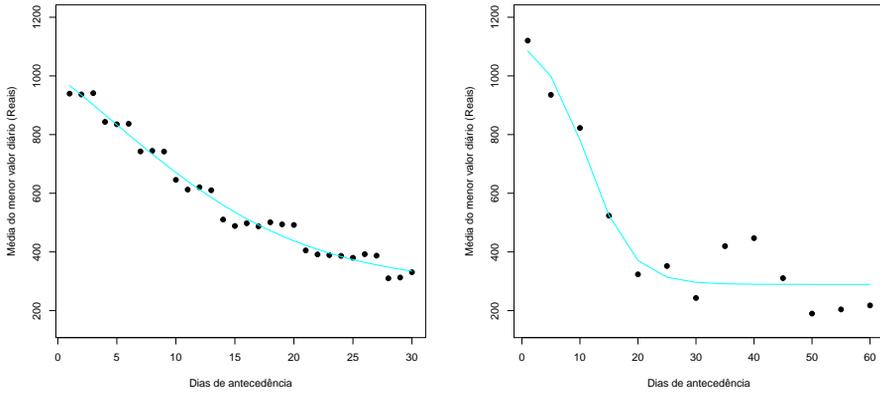


Figura 4: LDB-CGH, empresa A, modelos M5 e M5.

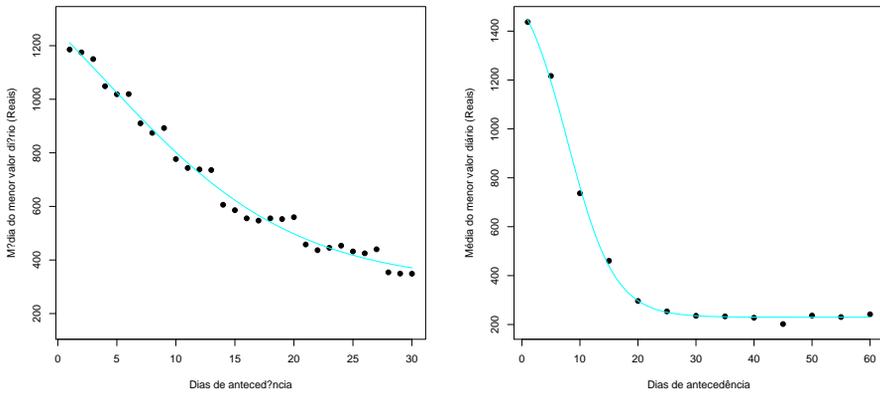


Figura 5: LDB-CGH, empresa B, modelos M5 e M5.

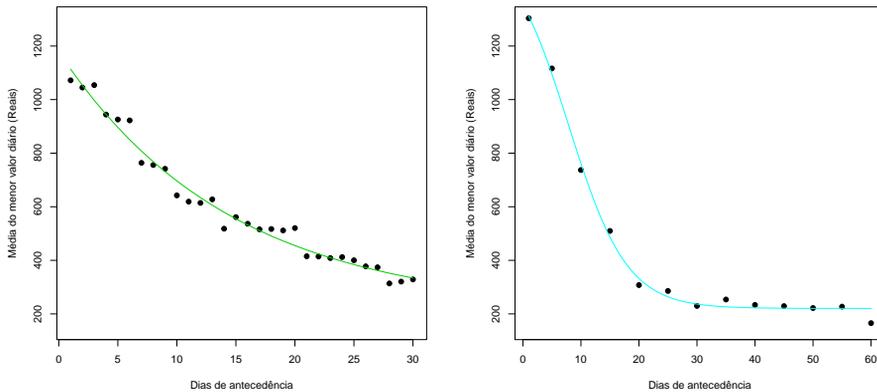


Figura 6: LDB-CGH, empresa C, modelos M3 e M5.

Dessa forma, considerando um período de 30 dias de antecedência, os valores das passagens comportam-se de forma não linear, ou seja, o valor pago por uma passagem adquirida com 15 dias de antecedência está mais próximo do valor pago na mesma com 30 dias de antecedência do que com 1 dia de antecedência da data da viagem; após esse intervalo, no geral, os preços das passagens não sofrem alterações significativas.

#### 4. Considerações Finais

Embora existam poucos estudos disponíveis na literatura nacional sobre as formas de precificação dos trechos para companhias aéreas, entender como isso funciona pode trazer alguma economia na compra desses bilhetes.

Esse trabalho mostrou que, em um período de 30 dias de antecedência da compra do bilhete, os valores comportam-se de forma não linear, e, conseqüente a essa data, não há alterações significativas do mesmo, sendo assim, comprar passagem de forma antecipada pode trazer economia ao consumidor, além de pesquisar companhias que ofereçam os melhores preços.

Contudo, caso o consumidor não consiga adquirir a passagem com 30 dias de antecedência ainda é possível economizar comprando com até 15 dias antes da data da viagem. Do contrário, a compra realizada posterior a 30 dias de antecedência, não garante ao consumidor uma expressiva economia. Logo, recomenda-se a compra de passagens aéreas para o trecho Londrina - Curitiba ou Londrina - São Paulo com 30 dias de antecedência da viagem. Além disso, por meio dos ajustes, é possível prever o valor pago ao longo de 60 dias de antecedência, utilizando os modelos da Tabela 3 em que  $x$  é o número de dias de antecedência para a compra do bilhete. Obviamente, deve-se considerar uma possível inflação com o passar do tempo.

Trecho	Aérea	Melhor Modelo Ajustado
	A	$\hat{y} = 832,76 \exp(-0,080x) + 53,15 \exp(0,024x)$
LDB-CWB	B	$\hat{y} = 316,88 + \frac{1473,1-316,9}{1+\exp(\frac{x-7,73}{3,07})}$
	C	$\hat{y} = 253,51 + \frac{1454,4-253,5}{1+\exp(\frac{x-2,03}{6,13})}$
	A	$\hat{y} = 289,37 + \frac{1143,6-289,4}{1+\exp(\frac{x-11,21}{3,91})}$
LDB-CGH	B	$\hat{y} = 229,92 + \frac{1666,6-229,9}{1+\exp(\frac{x-7,88}{4,03})}$
	C	$\hat{y} = 220,34 + \frac{1589,1-220,3}{1+\exp(\frac{x-7,88}{5,02})}$

Tabela 3: Tabela de predição dos preços para 60 dias de antecedência.

## Referências

- [1] Achcar, J. A.; Oliveira, B. R. G. “Gompertz growth curves assuming stable distributions: An application to intrauterine growth for preterm infants”. *Revista Brasileira de Biometria*, v. 33, n.º. 2, pp. 170-183, 2015.
- [2] Akaike, H. *A new look at the statistical model identification*. *Automatic Control, IEEE Transactions on*, v. 19, n.º. 6, pp. 716-723, 1974.
- [3] Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*. Wiley, (Wiley Series in Probability and Statistics), 2007
- [4] Bello, L. H. A. *Modelagem em Experimentos Mistura-Processo para Otimização de Processos Industriais*. Rio de Janeiro, 2010
- [5] Bilotkach, V. “Understanding Price Dispersion in the Airline Industry: Capacity Constraints and Consumer Heterogeneity”. Darin Lee, Org. *Advances in Airline Economics*, Vol. I. Elsevier. *Advances in Airline Economics*, 2006.
- [6] Burnham, K. P.; Anderson, D.R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer New York, 2002.
- [7] Draper, N. R.; Smith, H. *Applied regression analysis*. New York, U.S.A, 1998
- [8] Emiliano, P. C. *Fundamentos e aplicações dos critérios de informação: Akaike e Bayesiano*. Dissertação Mestrado em Estatística e Experimentação Agropecuária. Universidade Federal de Lavras, 2009
- [9] Emiliano, P. C. *Crítérios de informação de akaike versus bayesiano: análise comparativa*. 19º Simpósio Nacional de Probabilidade e Estatística, 2010
- [10] Evans, W.; Kessides, I. “Localized market power in the U.S. airline industry”. *The Review of Economics and Statistics*, v. 75, n.º. 1, pp. 66-75, 1993.
- [11] Hurvich, C. M.; Tsai, C.L. “Regression and time series model selection in small samples”. *Biometrika*, v. 76, n.º. 2, pp. 297-307, 1989.
- [12] Konishi, S.; Kitagawa, G. *Information Criteria and Statistical Modeling*. Springer Publishing Company, Incorporated, 2007.
- [13] Mazucheli, J.; Achcar, J. A. *Algumas considerações em regressão não linear*. *Acta Scientiarum : Technology*, v. 24, n.º. 6, pp. 1761-1770, 2002.

- [14] Morais, L. G. A.; Huse, C. *Yield Management nos hotéis do Rio de Janeiro: levantamento e análise*. Dissertação Mestrado em Engenharia de Produção. Pontifícia Universidade Católica, 2002.
- [15] Oliveira, A. V. M.; Huse, C. *Localized Competitive Advantage and Price Reactions to Entry: Full-Service vs. Low-Cost Airlines in the Brazilian Market*. Acervo científico do núcleo de estudos em competição e regulação do transporte aéreo (Nectar), 2005.
- [16] Oliveira, A. V. M. *Estudo dos Determinantes dos Preços das Companhias Aéreas no Mercado Doméstico*. Anac - Agência Nacional de Aviação Civil, 2009.
- [17] Sakamoto, Y.; Ishiguro, M.; Kitagawa, G. *Akaike Information Criterion Statistics*. Reidel Publishing Company, 1986
- [18] Schwarz, G.; "Estimating the dimension of a model". *The Annals of Statistics*, v. 6, pp. 461-464, 1978.
- [19] Seber, G. A. F.; Wild, C. J. *Nonlinear Regression*. Wiley, (Wiley Series in Probability and Statistics), 2003
- [20] Sugiura, N. "Further analysts of the data by Akaike's information criterion and the infinite corrections". *Communications in Statistics-Theory and Methods.*, Ontario, v. 7, n°. 1, pp. 13-26, 1978.
- [21] Thomas, G. *Regressão não Linear*. Trabalho de conclusão do curso de pós-graduação em estatística e experimentação agropecuária. Escola superior de Agricultura Luiz de Queiroz - USP, 2016.
- [22] Weatherford, L. R. *Perishable Asset Revenue Management in General Business Situations*. Ph.D. thesis, Darden Graduate School of Business Administration, University of Virginia, Charlottesville, VA, 1991.

Juliana de Melo  
Universidade Tecnológica Federal do Paraná  
<[jumello@hotmail.com](mailto:jumello@hotmail.com)>

Roberto Molina de Souza  
Universidade Tecnológica Federal do Paraná  
<[rmolinasouza@utfpr.edu.br](mailto:rmolinasouza@utfpr.edu.br)>

Recebido: 11/05/2020  
Publicado: 06/08/2020

# Funções auxiliares nos teoremas clássicos do valor médio

German Lozada-Cruz 

## Resumo

Neste trabalho mostramos algumas variantes dos teoremas clássicos do valor médio. As ferramentas principais para mostrarmos estes resultados são funções auxiliares elementares.

**Palavras-chave:** Teorema do Valor Médio de Lagrange; Teorema do Valor Médio de Cauchy; Teorema do Valor Médio para Integrais.

## Abstract

In this work, some variants of classical mean value theorems are proved. The main tools to prove these results are some elementary auxiliary functions.

**Keyword:** Lagrange's mean value theorem; Cauchy's mean value theorem; Integral mean value theorem.

## 1. Introdução

Sabemos que o primeiro contato de nossos alunos dos cursos de graduação em Matemática ou de Engenharia com os teoremas do tipo valor médio é em um curso de Cálculo Diferencial e Integral ([11] e [17] por exemplo) ou em um primeiro curso de Análise Real ([12]).

O primeiro teorema do valor médio é o conhecido Teorema do Valor Médio de Lagrange, o qual relaciona a taxa média da variação de uma função nos extremos de um intervalo com o valor da derivada da função em um ponto do mesmo intervalo.

O segundo teorema de valor médio é o Teorema do Valor Médio de Cauchy, o qual é uma generalização do Teorema do Valor Médio de Lagrange, que estabelece uma relação entre as derivadas de duas funções e a variação dessas funções em um intervalo finito.

O terceiro teorema de valor médio é o Teorema do Valor Médio para Integrais ou simplesmente Teorema do Valor Médio Integral, que relaciona a área da região sob o gráfico de uma função real  $f$  definida num intervalo fechado  $[a, b]$  com a área de um retângulo de base  $b - a$  e altura  $f(c)$  para algum  $c$  no intervalo  $[a, b]$ .

Nas demonstrações desses teoremas (veja por exemplo [3, Teorema 2.3] e [12, Teorema 4.12]) algumas funções auxiliares são usadas e então aplica-se diretamente o Teorema de Rolle: Seja  $f : [a, b] \rightarrow \mathbb{R}$  contínua em  $[a, b]$  e diferenciável em  $(a, b)$ . Se  $f(a) = f(b)$ , então, existe  $c \in (a, b)$  tal que  $f'(c) = 0$ . Alguns autores chamam isso de “*ideia feliz*” ([7, p.24]).

O objetivo principal deste trabalho é usar funções auxiliares elementares para demonstrar algumas variações tanto dos Teoremas do Valor Médio de Lagrange, de Cauchy e Integral.

## 2. Funções auxiliares para o Teorema do Valor Médio de Lagrange

Nesta seção vamos usar algumas funções auxiliares elementares para demonstrar algumas variações do Teorema do Valor Médio de Lagrange.

**Teorema 1** (Teorema de Lagrange). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$  e derivável em  $(a, b)$ , então existe  $\eta \in (a, b)$  tal que*

$$f(\eta) - f(a) = f'(\eta)(\eta - a). \quad (1)$$

Geometricamente, o Teorema do Valor Médio de Lagrange diz-nos que existe um ponto  $\eta$  dentro do intervalo  $(a, b)$ , onde a reta tangente ao gráfico de  $f$  é paralela à reta secante que passa pelos pontos  $A = (a, f(a))$  e  $B = (b, f(b))$ , ou seja  $f'(\eta) = \frac{f(b)-f(a)}{b-a}$ .

Se  $f(t)$  representa a posição de um corpo que se move ao longo de uma reta, dependendo do tempo  $t$ , então o quociente  $\frac{f(b)-f(a)}{b-a}$  é a velocidade média do corpo no período de tempo  $b-a$ . Como  $f'(t)$  é a velocidade instantânea, fisicamente o Teorema do Valor Médio de Lagrange diz que existe um instante de tempo  $\eta$ , para o qual a velocidade instantânea é igual à velocidade média.

A demonstração do Teorema de Lagrange (veja [12, Teorema 4.12]) consiste em aplicar o Teorema de Rolle à função auxiliar

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a). \quad (2)$$

Neste ponto, cabe fazer a seguinte pergunta: Qual é a origem da função auxiliar  $\phi$  usada na demonstração do Teorema do Valor Médio de Lagrange?

A explicação mais simples da escolha de  $\phi$  provavelmente seja a dada por R.C. Yates ([23]). A função  $\phi$  é a diferença da ordenada do ponto  $P = (x, f(x))$  no gráfico de  $f$  e da ordenada do ponto  $Q = (x, h(x))$  na reta secante ao gráfico de  $f$ . Como as funções  $f$  e  $h$  se encontram nos pontos  $(a, f(a))$  e  $(b, f(b))$  temos  $\phi(a) = 0 = \phi(b)$ . Assim, temos uma função para a qual podemos aplicar o Teorema de Rolle; conseqüentemente, obtemos o Teorema do Valor Médio de Lagrange.

Existem outras maneiras de obter  $\phi$  (veja [10], [15] por exemplo). M.R. Spiegel em [16] observou que podemos encontrar  $\phi$  da seguinte forma: vamos procurar uma aproximação linear de  $f$ ,  $\alpha + \beta x$ , no intervalo  $[a, b]$ , onde  $\alpha$  e  $\beta$  são constantes a serem determinadas. Agora consideremos a diferença,  $\phi(x)$ , entre  $f(x)$  e sua aproximação linear  $\alpha + \beta x$ , i.e.,

$$\phi(x) = f(x) - (\alpha + \beta x), \quad (3)$$

onde  $\alpha$  e  $\beta$  são números reais a serem determinados satisfazendo a condição  $\phi(a) = \phi(b)$ . Dessa última condição obtemos  $\beta = \frac{f(b)-f(a)}{b-a}$ . Para acharmos  $\alpha$ , fazemos  $\phi(a) = 0$ , assim obtemos  $\alpha = f(a) - \frac{f(b)-f(a)}{b-a}a$ . Logo,

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Se fizermos uma ligeira modificação da aproximação linear de  $f$  da forma

$$\phi(x) = f(x) + \lambda x, \quad (4)$$

onde  $\lambda \in \mathbb{R}$  tal que a condição  $\phi(a) = \phi(b)$  seja satisfeita, obtemos a seguinte função auxiliar

$$\phi(x) = f(x) - \frac{f(b) - f(a)}{b - a}x.$$

J.Tong em [18] introduziu a função auxiliar da forma

$$H(x) = f(x) - \frac{f(b) - f(a)}{b - a} \left( x - \frac{a + b}{2} \right), \quad (5)$$

a qual satisfaz a condição  $H(a) = \frac{f(a) + f(a)}{2} = H(b)$ .

Se mudarmos ligeiramente a função auxiliar  $H$  dada em (5) pela função auxiliar

$$H(x) = f(x) - \frac{f(a) + f(b)}{2} - \frac{f(b) - f(a)}{b - a} \left( x - \frac{a + b}{2} \right). \quad (6)$$

Facilmente vemos que  $H(a) = 0 = H(b)$ .

J.Tong em [19, Teorema 1] exhibe uma família de funções auxiliares que envolvem dois parâmetros  $\alpha$  e  $\beta$ , os quais satisfazem o Teorema de Rolle. Conseqüentemente, obtemos uma família de Teoremas de Valor Médio. Em particular, quando  $\alpha = \beta$ , obtém-se o Teorema do Valor Médio de Lagrange. Isso leva a concluir que existem uma infinidade de funções auxiliares.

Em 1958 T.M. Flett em [2] enunciou e demonstrou uma variação do Teorema de Lagrange.

**Teorema 2** (Teorema de Flett [2]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função derivável em  $[a, b]$  e  $f'(a) = f'(b)$ , então, existe  $\eta \in (a, b)$  tal que*

$$f(\eta) - f(a) = f'(\eta)(\eta - a). \quad (7)$$

Geometricamente, o Teorema de Flett diz que se uma curva  $(t, f(t))$  é suave no intervalo  $[a, b]$  e as retas tangentes nos extremos  $(a, f(a))$  e  $(b, f(b))$  são paralelas, então, existe um ponto  $\eta \in (a, b)$  de modo que a reta tangente ao gráfico de  $f$  que passa por  $(\eta, f(\eta))$  também passa por  $(a, f(a))$ .

Fisicamente, o Teorema de Flett diz que, se as velocidades inicial e final de uma partícula com trajetória  $(t, f(t))$  suave no intervalo de tempo  $[a, b]$  forem iguais, então, existe um momento  $\eta \in (a, b)$  tal que a velocidade instantânea da partícula nesse instante, é exatamente a velocidade média do percurso até o instante  $\eta$ .

Em 1977, R.E. Myers [9] demonstrou uma variação do Teorema de Flett.

**Teorema 3** (Teorema de Myers [9]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função derivável em  $[a, b]$  e  $f'(a) = f'(b)$ , então, existe  $\eta \in (a, b)$  tal que*

$$f(b) - f(\eta) = f'(\eta)(b - \eta). \quad (8)$$

Em 1998, P.K. Sahoo e T. Riedel ([13, Teorema 5.2]) deram uma variação do Teorema de Flett (Teorema 2) onde eles removeram a condição de fronteira na derivada de  $f$ , i.e.,  $f'(a) = f'(b)$ .

**Teorema 4** (Teorema de Sahoo-Riedel [13]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função derivável em  $[a, b]$ ,*

então, existe  $\eta \in (a, b)$  tal que

$$f(\eta) - f(a) = f'(\eta)(\eta - a) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (\eta - a)^2. \quad (9)$$

*Demonstração.* Aplicar o Teorema de Flett à função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (x - a)^2. \quad (10)$$

□

*Observação 1.* A função auxiliar  $\psi$  usada no Teorema de Sahoo-Riedel é obtida considerando a diferença de  $f$  com uma aproximação quadrática de  $f$ ,

$$a + \beta(x - a) + \gamma(x - a)^2$$

numa vizinhança de  $a$ , i.e.,  $\psi(x) = f(x) - [a + \beta(x - a) + \gamma(x - a)^2]$  e impondo a condição de fronteira na derivada de  $\psi$ ,  $\psi'(a) = \psi'(b)$ . Assim,

$$\begin{aligned} \psi'(a) = \psi'(b) &\Leftrightarrow f'(a) - \beta = f'(b) - \beta - 2\gamma(b - a) \\ &\Leftrightarrow \gamma = \frac{1}{2} \frac{f'(b) - f'(a)}{b - a}. \end{aligned}$$

As constantes  $\alpha$  e  $\beta$  são arbitrárias e por conveniência podemos tomar  $\alpha = 0 = \beta$ .

Da Observação 1 podemos considerar uma função auxiliar adequada para mostrar uma variação do Teorema de Sahoo-Riedel.

**Teorema 5** ([4]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função diferenciável em  $[a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$f(\eta) - f(a) = f'(\eta)(\eta - a) - \frac{(n-1)}{n} \frac{f'(b) - f'(a)}{(b-a)^{n-1}} (\eta - a)^n, \quad n \in \mathbb{N}. \quad (11)$$

*Demonstração.* Seja  $n \in \mathbb{N}$  e considere a função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) + \lambda(x - a)^n,$$

onde  $\lambda \in \mathbb{R}$  que vamos escolher  $\lambda$  tal que a condição  $\psi'(a) = \psi'(b)$  seja satisfeita.

Facilmente vemos que  $\psi$  é derivável em  $[a, b]$  e  $\psi'(x) = f'(x) + n\lambda(x - a)^{n-1}$ . Então,

$$\begin{aligned} \psi'(a) = \psi'(b) &\Leftrightarrow f'(a) = f'(b) + n\lambda(b - a)^{n-1} \\ &\Leftrightarrow \lambda = -\frac{1}{n} \frac{f'(b) - f'(a)}{(b - a)^{n-1}}. \end{aligned}$$

Assim, temos a função auxiliar

$$\psi(x) = f(x) - \frac{1}{n} \frac{f'(b) - f'(a)}{(b - a)^{n-1}} (x - a)^n$$

, a qual satisfaz as hipóteses do Teorema de Flett (Theorem 2). Então, existe  $\eta \in (a, b)$  tal que

$$\psi(\eta) - \psi(a) = \psi'(\eta)(\eta - a).$$

Dessa igualdade segue (11). □

D.Çakmak e A.Tiryaki ([1, Teorema 2.1]) mostrou uma ligeira variação do Teorema de Sahoo-Riedel (Teorema 4) e esse se reduz ao Teorema de Myers (Teorema 3) quando  $f'(a) = f'(b)$ .

**Teorema 6** (Teorema de Çakmak-Tiryaki [1]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável em  $[a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$f(b) - f(\eta) = f'(\eta)(b - \eta) + \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (b - \eta)^2. \quad (12)$$

*Demonstração.* Aplicar o Teorema de Myers à função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  definida por

$$h(x) = f(x) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (x - b)^2. \quad (13)$$

□

Escolhendo um função auxiliar adequada obtemos a seguinte variação do Teorema de Çakmak-Tiryaki.

**Teorema 7** ([4]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função derivável em  $[a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$f(b) - f(\eta) = f'(\eta)(b - \eta) + \frac{n-1}{n} \frac{f'(b) - f'(a)}{(b-a)^{n-1}} (b - \eta)^n, \quad n \in \mathbb{N}. \quad (14)$$

*Demonstração.* Seja  $n \in \mathbb{N}$  e considere a função auxiliar  $\phi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\phi(x) = f(x) + \lambda(x - b)^n,$$

onde  $\lambda \in \mathbb{R}$ . Vamos escolher  $\lambda$  tal que a condição  $\phi'(a) = \phi'(b)$  esteja satisfeita.

A função  $\phi$  é derivável em  $[a, b]$  e  $\phi'(x) = f'(x) + n\lambda(x - b)^{n-1}$ . Então,

$$\begin{aligned} \phi'(a) = \phi'(b) &\Leftrightarrow f'(a) + n\lambda(a - b)^{n-1} = f'(b) \\ &\Leftrightarrow \lambda = \frac{1}{n} \frac{f'(b) - f'(a)}{(a - b)^{n-1}}. \end{aligned}$$

Assim, temos a função auxiliar

$$\phi(x) = f(x) + \frac{1}{n} \frac{f'(b) - f'(a)}{(a - b)^{n-1}} (x - b)^n$$

, a qual satisfaz as condições do Teorema Myers (Theorem 3). Então, existe  $\eta \in (a, b)$  tal que

$$\phi(b) - \phi(\eta) = \phi'(\eta)(b - \eta).$$

Dessa última igualdade segue (14). □

### 3. Funções auxiliares para o Teorema do Valor Médio de Cauchy

Nesta seção vamos usar algumas funções auxiliares elementares para demonstrar algumas variações do Teorema do Valor Médio de Cauchy.

**Teorema 8** (Teorema de Cauchy). *Se  $f, g : [a, b] \rightarrow \mathbb{R}$  são funções contínuas em  $[a, b]$  e deriváveis em  $(a, b)$ , então, existe  $\eta \in (a, b)$  tal que*

$$[f(b) - f(a)]g'(\eta) = [g(b) - g(a)]f'(\eta). \quad (15)$$

Se  $g'(x) \neq 0$  para todo  $x \in (a, b)$ , então,

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\eta)}{g'(\eta)}. \quad (16)$$

Geometricamente, o Teorema de Cauchy diz que existe uma reta tangente à curva  $\gamma : [a, b] \rightarrow \mathbb{R}^2$  dada por  $\gamma(t) = (f(t), g(t))$ , a qual é paralela à reta que passa pelos pontos  $A = (f(a), g(a))$  e  $B = (f(b), g(b))$  da curva  $\gamma$ .

Se  $g(x) = x$  no Teorema de Cauchy (Teorema 8), então obtemos o Teorema de Lagrange (Teorema 1).

A demonstração do Teorema de Cauchy [12, Teorema 4.14] consiste em aplicar o Teorema de Rolle à função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  dada por

$$h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}[g(x) - g(a)]. \quad (17)$$

Como na demonstração do Teorema de Lagrange cabe aqui a seguinte pergunta: Qual é a origem da função auxiliar  $\phi$  usada na demonstração do Teorema de Cauchy?

Uma ligeira modificação da função auxiliar (17) é usada em [8, Teorema 5.11] para demonstrar o Teorema de Cauchy:

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g(x). \quad (18)$$

M. Poliferno [10] considerou a função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  dada por

$$h(x) = \frac{f(x) - mg(x)}{\sqrt{m^2 + 1}},$$

onde  $m = \frac{f(b)-f(a)}{g(b)-g(a)}$  é a inclinação de reta  $L$  que passa pelos extremos  $A = (f(a), g(a))$  e  $B = (f(b), g(b))$  da curva  $\gamma(x) = (f(x), g(x))$ . A função  $h$  é a distância de ponto da curva  $\gamma$  à reta  $L$ , então, geometricamente,  $h(a) = 0 = h(b)$ . Logo, pelo Teorema de Rolle, segue a conclusão do Teorema de Cauchy.

J. Tong [20] considerou a seguinte função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  dada por

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} \left[ g(x) - \frac{g(a) + g(b)}{2} \right] \quad (19)$$

Claramente vemos que  $h(a) = \frac{f(a)+f(b)}{2} = h(b)$ , logo, o Teorema de Rolle dá a conclusão do Teorema de Cauchy.

*Observação 2.* Modificando ligeiramente a função auxiliar  $h$  dada em (19) por J. Tong pela função auxiliar

$$h(x) = f(x) - \frac{f(a) + f(b)}{2} - \frac{f(b) - f(a)}{g(b) - g(a)} \left[ g(x) - \frac{g(a) + g(b)}{2} \right],$$

também obtemos a conclusão do Teorema do Valor Médio de Cauchy.

M. Spiegel [16] considerou a função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  da forma

$$h(x) = f(x) - [\alpha + \beta g(x)], \quad (20)$$

onde  $\alpha$  e  $\beta$  são números reais a serem determinados satisfazendo a condição  $h(a) = h(b)$ . Dessa condição, obtemos  $\beta = \frac{f(b)-f(a)}{g(b)-g(a)}$ . Para determinar a constante  $\alpha$  fazemos  $h(a) = 0$ , assim  $\alpha = f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}g(a)$ . Logo, a função auxiliar tem a forma

$$h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} [g(x) - g(a)]. \quad (21)$$

Note que a função dada em (21) é a mesma função auxiliar considerada por M. Protter e C. Morrey (veja (17)) e, uma forma reduzida da função dada em (21) é

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g(x), \quad (22)$$

a qual é obtida quando tomamos  $\alpha = 0$ .

Seguindo a ideia de M. Spiegel em [16], consideramos a função auxiliar  $h : [a, b] \rightarrow \mathbb{R}$  dada por

$$h(x) = f(x) + \lambda g(x),$$

onde  $\lambda \in \mathbb{R}$ , o qual será escolhido tal que  $h(a) = h(b)$ . Dessa condição, obtemos  $\lambda = -\frac{f(b)-f(a)}{g(b)-g(a)}$ . Assim, temos a função auxiliar

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g(x).$$

Veja que essa função auxiliar é a mesma obtida em (22).

Em 2000, E.Wachnicki ([21, Teorema 3]) demonstrou a seguinte variação do Teorema do Valor Médio de Cauchy.

**Teorema 9** (Teorema de Wachnicki [21]). *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$  e*

$$\frac{f'(a)}{g'(a)} = \frac{f'(b)}{g'(b)}, \quad (23)$$

então, existe  $\eta \in (a, b)$  tal que

$$\frac{f(\eta) - f(a)}{g(\eta) - g(a)} = \frac{f'(\eta)}{g'(\eta)}. \quad (24)$$

E. Wachnicki em [21, Teorema 4] removeu a condição (23) e obteve o seguinte resultado o qual é um análogo ao Teorema de Sahoo-Riedel.

**Teorema 10.** *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$\frac{f(\eta) - f(a)}{g(\eta) - g(a)} = \frac{f'(\eta)}{g'(\eta)} - \frac{1}{2} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{g(\eta) - g(a)}{g(b) - g(a)}. \quad (25)$$

A demonstração desse Teorema (veja [21, Teorema 4]) consiste em aplicar o Teorema de Wachnicki à função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) - \frac{1}{2} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{(g(x) - g(a))^2}{g(b) - g(a)}.$$

Em [5] mostramos a seguinte variação do Teorema 10.

**Teorema 11** ([5, Teorema 2.3]). *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$\frac{f(\eta) - f(a)}{g(\eta) - g(a)} = \frac{f'(\eta)}{g'(\eta)} - \frac{n-1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \left( \frac{g(\eta) - g(a)}{g(b) - g(a)} \right)^{n-1}, \quad n \in \mathbb{N}. \quad (26)$$

*Demonstração.* Seja  $n \in \mathbb{N}$  e consideremos a função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) + \lambda (g(x) - g(a))^n, \quad \lambda \in \mathbb{R}.$$

Vamos escolher  $\lambda$  de modo que a seguinte condição esteja satisfeita

$$\frac{\psi'(a)}{g'(a)} = \frac{\psi'(b)}{g'(b)}. \quad (27)$$

A função  $\psi$  é derivável em  $[a, b]$  e  $\psi'(x) = f'(x) + \lambda n (g(x) - g(a))^{n-1} g'(x)$ . De (27) obtemos

$$\begin{aligned} \frac{\psi'(a)}{g'(a)} = \frac{\psi'(b)}{g'(b)} &\Leftrightarrow \frac{f'(a)}{g'(a)} = \frac{f'(b)}{g'(b)} + \lambda n [g(b) - g(a)]^{n-1} \\ &\Leftrightarrow \lambda = -\frac{1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{1}{(g(b) - g(a))^{n-1}}. \end{aligned}$$

Assim, a função auxiliar  $\psi$  tem a forma

$$\psi(x) = f(x) - \frac{1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{(g(x) - g(a))^n}{(g(b) - g(a))^{n-1}}.$$

Agora, aplicando o Teorema de Wachnicki (Teorema 9) às funções  $\psi$  e  $g$ , existe  $\eta \in (a, b)$  tal que

$$\frac{\psi(\eta) - \psi(a)}{g(\eta) - g(a)} = \frac{\psi'(\eta)}{g'(\eta)}.$$

De onde segue a igualdade (26) □

*Observação 3.* Se  $g(x) = x$ , então, os Teoremas 9 e 10 transformam-se nos Teoremas de Flett (Teorema 2) e de Sahoo-Riedel (Teorema 4), respectivamente.

*Observação 4.* Se  $g(x) = x$ , então, o Teorema 11 transform-se em uma variação do Teorema de Sahoo-Riedel (veja [4]), i.e.,

$$f(\eta) - f(a) = f'(\eta)(\eta - a) - \frac{n-1}{n} \frac{f'(b) - f'(a)}{(b-a)^{n-1}} (\eta - a)^n, \quad n \in \mathbb{N}. \quad (28)$$

Em [5], provamos a seguinte variação do Teorema de Wachnicki.

**Teorema 12** ([5, Teorema 2.4]). *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$  e*

$$\frac{f'(a)}{g'(a)} = \frac{f'(b)}{g'(b)}, \quad (29)$$

então, existe  $\eta \in (a, b)$  tal que

$$\frac{f(b) - f(\eta)}{g(b) - g(\eta)} = \frac{f'(\eta)}{g'(\eta)}. \quad (30)$$

Removendo a condição (29), obtivemos uma variação do Teorema 12, o qual é um análogo do Teorema Çakmak-Tiryaki (Teorema 6).

**Teorema 13.** *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$\frac{f(b) - f(\eta)}{g(b) - g(\eta)} = \frac{f'(\eta)}{g'(\eta)} + \frac{1}{2} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{g(b) - g(\eta)}{g(b) - g(a)}. \quad (31)$$

Para demonstrar esse teorema, usamos a função auxiliar  $\phi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\phi(x) = f(x) - \frac{1}{2} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{(g(x) - g(b))^2}{g(b) - g(a)}$$

junto com a função  $g$ , e logo aplicamos o Teorema 12.

Em [5] obtivemos a seguinte variação do Teorema 13.

**Teorema 14** ([5, Teorema 2.5]). *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções deriváveis em  $[a, b]$ . Se  $g'(x) \neq 0$  para todo  $x \in [a, b]$ , então, existe  $\eta \in (a, b)$  tal que*

$$\frac{f(b) - f(\eta)}{g(b) - g(\eta)} = \frac{f'(\eta)}{g'(\eta)} + \frac{n-1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \left( \frac{g(b) - g(\eta)}{g(b) - g(a)} \right)^{n-1}, \quad n \in \mathbb{N}. \quad (32)$$

*Demonstração.* Seja  $n \in \mathbb{N}$ , e consideremos a função auxiliar  $\phi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\phi(x) = f(x) + \lambda(g(x) - g(b))^n, \lambda \in \mathbb{R}.$$

Vamos escolher  $\lambda$  de modo que a condição seja satisfeita

$$\frac{\phi'(a)}{g'(a)} = \frac{\phi'(b)}{g'(b)}. \quad (33)$$

A função  $\phi$  é derivável em  $[a, b]$  e  $\phi'(x) = f'(x) + \lambda n(g(x) - g(b))^{n-1}g'(x)$ . Assim,

$$\begin{aligned} \frac{\phi'(a)}{g'(a)} = \frac{\phi'(b)}{g'(b)} &\Leftrightarrow \frac{f'(a)}{g'(a)} + n\lambda[g(a) - g(b)]^{n-1} = \frac{f'(b)}{g'(b)} \\ &\Leftrightarrow \lambda = \frac{1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{1}{(g(a) - g(b))^{n-1}}. \end{aligned}$$

Logo, a função auxiliar  $\phi$  tem a forma

$$\phi(x) = f(x) + \frac{1}{n} \left( \frac{f'(b)}{g'(b)} - \frac{f'(a)}{g'(a)} \right) \frac{(g(x) - g(b))^n}{(g(a) - g(b))^{n-1}}.$$

Como as funções  $\phi$  e  $g$  satisfazem as hipóteses do Teorema 12, existe  $\eta \in (a, b)$  tal que

$$\frac{\phi(b) - \phi(\eta)}{g(b) - g(\eta)} = \frac{\phi'(\eta)}{g'(\eta)}.$$

Dessa última igualdade segue o resultado. □

#### 4. Funções auxiliares para o Teorema do Valor Médio Integral

Nesta seção vamos usar algumas funções auxiliares elementares para demonstrar algumas variações do Teorema do Valor Médio Integral.

**Teorema 15** (Teorema do Valor Médio Integral). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua, então, existe  $\xi \in (a, b)$  tal que*

$$(b - a)f(\xi) = \int_a^b f(x)dx. \quad (34)$$

Geometricamente, para uma função  $f$  não negativa, o Teorema do Valor Médio Integral diz que a área compreendida entre o gráfico de  $f$ , as retas  $x = a$ ,  $x = b$  e o eixo  $x$  é igual a  $(b - a)f(\xi)$ , ou seja, a área do retângulo de lados  $b - a$  e  $f(\xi)$ .

Em 1970, S.G. Wayment ([22]) demonstrou a primeira variação do Teorema do Valor Médio Integral, o qual é a versão integral do Teorema de Flett ([2]).

**Teorema 16** (Teorema de Wayment). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$  e  $f(a) = f(b)$ , então, existe  $\xi \in (a, b)$  tal que*

$$(\xi - a)f(\xi) = \int_a^\xi f(x)dx. \quad (35)$$

O seguinte resultado é uma ligeira modificação do Teorema de Wayment.

**Teorema 17** ([14, Teorema 3.2]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$  e  $f(a) = f(b)$ , então, existe  $\xi \in (a, b)$  tal que*

$$(b - \xi)f(\xi) = \int_{\xi}^b f(x)dx. \quad (36)$$

Em 2007, P.K. Sahoo [14, Teorema 2.1] removeu a condição de fronteira,  $f(a) = f(b)$ , do Teorema de Wayment (Teorema 16), e obteve o seguinte resultado:

**Teorema 18** (Teorema de Sahoo). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$(\xi - a)f(\xi) - \frac{1}{2} \frac{f(b) - f(a)}{b - a} (\xi - a)^2 = \int_a^{\xi} f(x)dx. \quad (37)$$

Para demonstrar o Teorema de Sahoo aplicamos o Teorema de Wayment à função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) - \frac{f(b) - f(a)}{b - a} (x - a).$$

Em [6] demonstramos a seguinte variante do Teorema de Sahoo.

**Teorema 19.** *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$(\xi - a)f(\xi) - \frac{n}{n + 1} \frac{f(b) - f(a)}{(b - a)^n} (\xi - a)^{n+1} = \int_a^{\xi} f(x)dx, \quad n \in \mathbb{N}. \quad (38)$$

*Demonstração.* Seja  $n \in \mathbb{N}$  e considere a função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por  $\psi(x) = f(x) + \lambda(x - a)^n$ , onde  $\lambda \in \mathbb{R}$ . Vamos escolher  $\lambda$  de modo que a condição  $\psi(a) = \psi(b)$  seja satisfeita. Então,

$$f(a) = f(b) + \lambda(b - a)^n \Leftrightarrow \lambda = -\frac{f(b) - f(a)}{(b - a)^n}.$$

Assim, temos a função auxiliar  $\psi(x) = f(x) - \frac{f(b) - f(a)}{(b - a)^n} (x - a)^n$ , a qual é contínua em  $[a, b]$ . Aplicando o Teorema de Wayment à função  $\psi$ , existe  $\xi \in (a, b)$  tal que

$$\psi(\xi)(\xi - a) = \int_a^{\xi} \psi(x)dx.$$

Dessa igualdade segue (38). □

*Observação 5.* A equação (38) no Teorema 19 dá uma família de Teoremas de valor médio para integrais.

Se consideramos a sequência de números reais  $\{I_n\}_{n \in \mathbb{N}}$  onde

$$I_n = (\xi - a)f(\xi) - \frac{n}{n + 1} \frac{f(b) - f(a)}{(b - a)^n} (\xi - a)^{n+1},$$

temos

$$\lim_{n \rightarrow \infty} I_n = (\xi - a)f(\xi).$$

Note que:

- Se  $n = 1$  em (38), obtemos o Teorema de Sahoo (Teorem 18). Nesse caso, a função auxiliar é da forma

$$\psi(x) = f(x) - \frac{f(b) - f(a)}{(b - a)}(x - a).$$

- Se  $n = 2$  em (38), obtemos uma variação de (37)

$$(\xi - a)f(\xi) - \frac{2}{3} \frac{f(b) - f(a)}{(b - a)^2} (\xi - a)^3 = \int_a^\xi f(x) dx.$$

Nesse caso, temos a função auxiliar

$$\psi(x) = f(x) - \frac{f(b) - f(a)}{(b - a)^2} (x - a)^2.$$

Quando removemos a condição de fronteira  $f(a) = f(b)$  no Teorema 17, obtemos o seguinte resultado:

**Teorema 20.** *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$(b - \xi)f(\xi) + \frac{1}{2} \frac{f(b) - f(a)}{b - a} (b - \xi)^2 = \int_\xi^b f(x) dx. \quad (39)$$

Na demonstração desse teorema, aplicamos o Teorema 17 à função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\psi(x) = f(x) + \frac{f(b) - f(a)}{b - a} (b - x).$$

Em [6] mostramos a seguinte variação do Teorem 20.

**Teorema 21.** *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$(b - \xi)f(\xi) + \frac{n}{n + 1} \frac{f(b) - f(a)}{(b - a)^n} (b - \xi)^{n+1} = \int_\xi^b f(x) dx, \quad n \in \mathbb{N}. \quad (40)$$

*Demonstração.* Seja  $n \in \mathbb{N}$  e considere a função auxiliar  $\psi : [a, b] \rightarrow \mathbb{R}$  dada por  $\psi(x) = f(x) + \lambda(x - b)^n$ , onde  $\lambda \in \mathbb{R}$ . Vamos escolher  $\lambda$  de modo que a condição  $\psi(a) = \psi(b)$  seja satisfeita. Então,

$$\psi(a) = \psi(b) \Leftrightarrow f(a) + \lambda(a - b)^n = f(b) \Leftrightarrow \lambda = \frac{f(b) - f(a)}{(a - b)^n}.$$

Assim, temos a função auxiliar  $\psi(x) = f(x) + \frac{f(b) - f(a)}{(b - a)^n} (x - b)^n$  a qual é contínua em  $[a, b]$ . Aplicando

o Teorema 17 à função  $\psi$ , existe  $\xi \in (a, b)$  tal que

$$(b - \xi) \psi(\xi) = \int_{\xi}^b \psi(x) dx,$$

a qual implica a igualdade (40). □

*Observação 6.* A equação (40) no Teorema 21 dá uma família de teoremas de valor médio para integrais.

Se consideramos a sequência de números reais  $\{I'_n\}_{n \in \mathbb{N}}$ , onde

$$I'_n = (b - \xi)f(\xi) + \frac{n}{n+1} \frac{f(b) - f(a)}{(b-a)^n} (b - \xi)^{n+1},$$

temos

$$\lim_{n \rightarrow \infty} I'_n = (b - \xi)f(\xi).$$

Note que:

- Se  $n = 1$  em (40), obtemos o Teorema 20. Nesse caso, temos a função auxiliar

$$\psi(x) = f(x) + \frac{f(b) - f(a)}{(b-a)}(x-b).$$

- Se  $n = 2$  in (40), obtemos uma variação de (39)

$$(b - \xi)f(\xi) + \frac{2}{3} \frac{f(b) - f(a)}{(b-a)^2} (b - \xi)^3 = \int_{\xi}^b f(x) dx.$$

Nesse caso, a função auxiliar é da forma

$$\psi(x) = f(x) + \frac{f(b) - f(a)}{(b-a)^2} (x-b)^2.$$

## Referências

- [1] Çakmak, D.; Tiryaki, A. *Mean value theorem for holomorphic functions*. Electron. J. Differential Equations, N°. 34, 6pp., 2012.
- [2] Flett, T.M. *A mean value problem*, Math. Gazette. 42, pp.38-39, 1958.
- [3] Lang, S., *Undergraduate Analysis*. Second edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2005.
- [4] Lozada-Cruz, G., *Some variants of Lagrange's mean value theorem*. Accepted for publication, 2020.
- [5] Lozada-Cruz, G., *Some variants of Cauchy's mean value theorem*. Int. J. Math. Ed. Sci. Tech. Accepted for publication, 2019.
- [6] Lozada-Cruz, G., *Some variants of Integral mean value theorem*. Submitted for publication, 2020.

- [7] Martínez de la Rosa, F., *Panorámica de los Teoremas de Valor Medio*. Miscelánea Matemática 47, pp.23-38, 2008.
- [8] Mercer, Peter R., *More Calculus of a single variable*. Undergraduate Texts in Mathematics. Springer, New York, 2014.
- [9] Myers, R.E. *Some elementary results related to the mean value theorem*, The Two-Year College Mathematics Journal, Vol. 8, N°.1, pp.51-53, 1977.
- [10] Poliferno, M. J.; *A natural auxiliary function for the Mean Value Theorem*. Amer. Math. Monthly 69, N° .1, pp.45-47, 1962.
- [11] Protter, M.H.; Protter, P.E., *Calculus with Analytic Geometric*. Fourth Edition. Jones and Bartlett Publishers, Inc., 1988.
- [12] Protter, M.H.; Morrey Jr., C.B., *A First Course in Real Analysis*. Second Edition. Undergraduate Texts in Mathematics, Springer-Verlag New York, Inc., 1991.
- [13] Sahoo, P.K.; Riedel, T. *Mean Value Theorems and Functional Equations*, World Scientific, River Edge, NJ, 1998.
- [14] Sahoo, P.K., Some results related to the integral mean value theorem, *Int. J. Math. Ed. Sci. Tech.* 38(6), 818-822, 2007.
- [15] Silverman, H., *A simple auxiliary function for the mean value theorem*. Coll. Math. J. 20, p. 323, 1989.
- [16] Spiegel, Murray R., *Mean value theorems and Taylor series*. Mathematics Magazine 29, pp.263-266, 1956.
- [17] Stewart, J., *Calculus Early Transcendentals*. Sixth Edition, Thomson Brooks/Cole, 2008.
- [18] Tong, J., *A generalization of the mean value theorem for Integrals*, The College Mathematics Journal, Vol. 33, pp.408-409, 2002.
- [19] Tong, J., *The mean value theorem generalised to involve two parameters*. Math. Gazette, vol. 88, N° 513, pp.538-540, 2004.
- [20] Tong, J., *A new auxiliary function for the mean value theorem*. Journal of the North Carolina Academy of Science Vol. 121, N° 4, pp.174-176, 2005.
- [21] Wachnicki, E., *Une variante du théorème de Cauchy de la valeur moyenne*, Demonstratio Math. 33. 4, pp.737-740, 2000.
- [22] Wayment, S.G., An integral mean value theorem, *Math. Gazette*. 54, 300-301, 1970.
- [23] Yates, R. C., *The law of the mean*. Amer. Math. Monthly. 66, pp.579-580, 1959.

German Lozada-Cruz  
Departamento de Matemática,  
Instituto de Biociências, Letras e Ciências Exatas (IBILCE),  
Universidade Estadual Paulista (UNESP),  
São José do Rio Preto, Brasil.  
<[german.lozada@unesp.br](mailto:german.lozada@unesp.br)>

Recebido: 28/02/2020

Publicado: 11/08/2020

# Sequências numéricas e cálculo de probabilidades associadas a modelos de urnas

Valdivino Vargas Júnior 

Tiago Moreira Vargas 

Divaldo P. Fernandes Júnior 

## Resumo

Considere uma urna contendo bolas diferindo apenas por suas cores. Dessa urna, bolas são extraídas ao acaso. No caso mais simples, a urna contém bolas de duas cores. Essas são extraídas de forma aleatória, podendo ou não haver reposição. No mais geral, podem-se estabelecer diversas regras: acrescentar bolas ao longo do processo de acordo com algum critério, trabalhar com bolas de diversas cores, fazer intercâmbio de bolas de uma urna para outra etc. Neste artigo, apresentamos algumas situações nas quais o cálculo probabilístico é feito via propriedades de sequências numéricas.

**Palavras-chave:** Modelos de Urnas; Sequências Numéricas; Probabilidade.

## Abstract

Consider an urn containing balls differing only in their colors. From this urn, balls are extracted at random. In the simplest case, the urn contains balls of two colors. These are extracted at random and may or may not be replaced. In general, several rules can be established: adding balls along the process according to some criteria, working with balls of different colors, exchanging balls from one urn to another, etc. In this article, we present some situations in which the probabilistic calculation is done through number sequence properties.

**keywords:** Urns Models; Numeric Sequences; Probability.

## 1. Introdução

A maioria dos processos que vivenciamos no nosso cotidiano envolvem algum tipo de aleatoriedade. Com base nisso, torna-se fundamental que estudantes dominem conceitos básicos envolvendo experimentos aleatórios. Em termos de Ensino Básico, o ensino de probabilidade é encarado como a competência na qual deve ser desenvolvida a compreensão da natureza aleatória da maioria dos fenômenos naturais e sociais. O cálculo probabilístico provê ferramentas adequadas para mensurar as incertezas associadas aos fenômenos aleatórios. Os modelos de urna apresentam problemas cujas situações descritas auxiliam a compreensão básica de conceitos de probabilidade. Na versão mais simples de um modelo de urnas, temos uma urna com bolas que diferem apenas por suas cores. Extrações são realizadas ao acaso. Numa versão, anota-se a cor da bola retirada e, logo a seguir é devolvida à urna antes da próxima extração. Noutra versão, as bolas extraídas não são devolvidas. O interesse, nesses casos, é a distribuição de probabilidade do número de bolas extraídas de uma

das cores. A partir dessas versões básicas surgem modelos mais elaborados. Em alguns, há acréscimo de bolas entre as retiradas de acordo com alguma regra específica. E, em geral, trabalha-se com bolas de diversas cores.

Heubeck (1974) cita que a primeira referência ao uso de urnas em problemas de probabilidade apareceu nos trabalhos de Huygens (1629-1695). Embora ele não cite a expressão modelo de urnas, muitas de suas contas podem ser expressas no contexto de problemas de urnas. Alguns anos depois, menções a modelos de urnas aparecem no trabalho de J. Bernoulli (1654-1705), em *Ars Conjectandi* (livro III), publicado postumamente em 1713. Alguns anos depois, D. Bernoulli (1700-1782), por diversas vezes, recorre a modelos de urnas. Em particular, em um trabalho de 1768, D. Bernoulli propôs um modelo, hoje conhecido como modelo de Ehrenfest. Modelos de urnas aparecem ainda em outros trabalhos de importantes matemáticos da época, dos quais podemos destacar: De Moivre (1667-1754) e Montmort (1678-1719). Algum tempo mais tarde, Laplace (1749-1827) estudou em detalhes o modelo de Ehrenfest. No século XIX modelos de urnas apareceram em diversos trabalhos de probabilistas e estatísticos da época. Dentre os autores que voltaram a atenção para esse tipo de modelo podemos citar: Quetelet (1794-1874) e Poisson (1781-1840). No início do século XX, mais especificamente em 1906, Markov (1856-1922) introduziu os modelos que hoje são chamados modelos de Pólya-Eggenberger. Para maiores detalhes sobre modelagem com urnas e sua evolução histórica veja Johnson e Kotz (1977).

Dentre as modelagens com urnas, têm destaque processos onde se estudam as distribuições do número de bolas em urnas. Esses processos, chamados modelos de ocupação, são definidos em diferentes vertentes no que tange às regras de distribuição das bolas, número de urnas, tipos de urnas etc. Sua origem ainda é razão de investigação, pois remontam a vários séculos atrás. Sprott (1956) cita que esse tipo de modelo já aparecia nos trabalhos de De Moivre (1718), e em versões mais generalizadas em conexão com problemas da física teórica nos trabalhos de Clausius (1822-1888), Boltzmann (1844- 1906) e Maxwell (1831-1879) sobre teoria dos gases. Mais detalhes sobre modelos de ocupação podem ser encontrados em Charalambides (2005).

Os modelos de urna possuem muitas aplicações práticas. Pode-se imaginar que as bolas constituem uma população de indivíduos. As características de interesse naquela população são representadas pelas cores das bolas. As extrações constituem o processo de amostragem. Assim, esses modelos ajudam na compreensão do funcionamento básico de um processo de amostragem. Em situações práticas os processos de amostragem antecedem os procedimentos de inferência estatística. Basicamente, fazer inferência estatística é, a partir de dados coletados, estimar e tomar decisões acerca de uma ou mais características da população em estudo. Por outro lado, muitos jogos de loterias também têm seus sorteios realizados a partir de processos cujas probabilidades de ganho são obtidas através do estudo de modelo de urnas. Um exemplo muito famoso no Brasil é a Mega-sena. No mais, a gama de problemas de probabilidade que pode ser pensada em termos de modelos de urnas é muito grande. Polya (1954) chegou a dizer que todo problema de probabilidade parece comparável a um problema de urnas contendo bolas. Nos dias atuais, além das evidentes aplicações na Física teórica e na Estatística, tem surgido uma imensidão de outras aplicações, das quais podemos destacar as aplicações nas áreas de tecnologia e mineração de dados. Para mais detalhes, veja Clarke, Fokoue e Zhang (2009).

Os problemas envolvendo modelos de urnas neste texto têm soluções baseadas em propriedades de sequências numéricas. Definimos, em cada problema, uma relação de recorrência, a qual é uma relação que determina cada termo de uma sequência como função de termos anteriores, e obtemos os resultados a partir de propriedades de sequências numéricas. Para que uma sequência seja descrita por uma equação de recorrência, é preciso estipular como se dá essa relação com termos anteriores

e também um ou mais termos iniciais. Embora seja um assunto muito relevante e abrangente, em termos de ensino básico, poucas sequências são estudadas nessa etapa. Normalmente, trabalha-se com progressões aritméticas e geométricas. Outros tipos de sequência aparecem, quando muito, de maneira discreta e normalmente apenas como exemplos. Um exemplo clássico é a sequência de Fibonacci. Detalhes sobre sequências e progressões podem ser encontrados em Morgado, Wagner e Zani(2015) e Hazzan e Iezzi (2012). Para compreensão de problemas envolvendo modelos de urnas é importante estudar as técnicas básicas de contagem antes de atacar os modelos em si. Os conceitos básicos de contagem podem ser encontrados em Carvalho, Fernandez, Morgado e Pitombeira (2016), dos Santos, Melo e Murari (2008), Dante (2016), Almeida, Degenszajn, Dolce, Iezzi e Périgo (2016), Diniz e Smole (2016) e Hazzan(2013). Os problemas deste artigo podem ser encontrados em livros-textos de probabilidade como em Ross(2010) e Stirzaker (2003).

## 2. Modelos

Nesta seção apresentamos resultados acerca de alguns modelos de urnas cujas soluções são baseadas em propriedades de sequências numéricas. Os resultados são obtidos a partir da teoria básica de probabilidade. Basicamente evocamos a fórmula da probabilidade total para dois eventos (dada pela Proposição 1) e os resultados seguem a partir de propriedades de séries e sequências após obtenção de fórmulas de recorrência adequadas.

**Proposição 1.** (*Fórmula da probabilidade total*) Sejam  $A$  e  $B$  dois eventos aleatórios para o espaço amostral  $\Omega$ . Então,

$$\mathbb{P}(A) = \mathbb{P}(B)\mathbb{P}(A|B) + \mathbb{P}(B^C)\mathbb{P}(A|B^C).$$

*Demonstração.* Aqui,  $B$  e  $B^C$  formam uma partição do espaço amostral  $\Omega$  (Ver Figura 1).

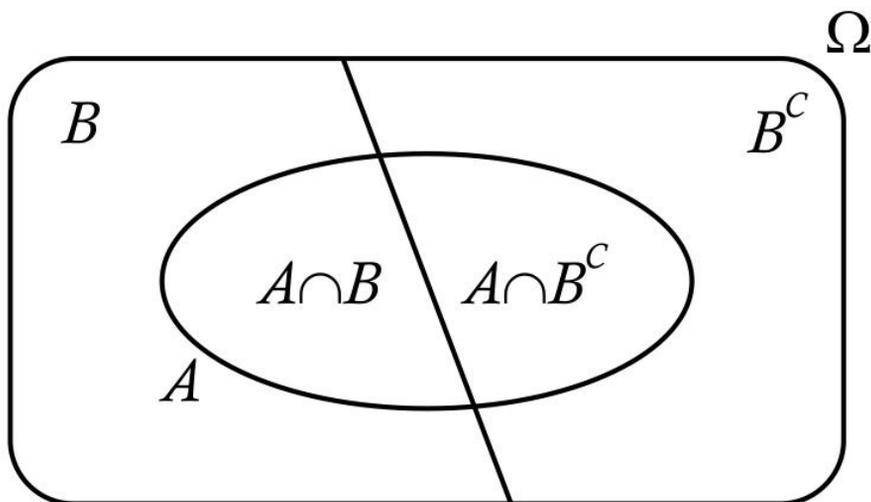


Figura 1: Partição do espaço amostral  $\Omega$ .

Assim,

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^C).$$

Como  $\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B)$  e  $\mathbb{P}(A \cap B^C) = \mathbb{P}(B^C)\mathbb{P}(A|B^C)$ , segue o resultado. □

A fórmula da probabilidade total pode ser estendida para qualquer quantidade finita ou, mais geralmente, enumerável de eventos formadores de uma partição do espaço amostral  $\Omega$  (Veja Ross (2010)). Intuitivamente falando, usamos a fórmula da probabilidade total em experimentos envolvendo duas ou mais etapas para calcular probabilidade de um evento  $A$ , o qual não podemos obter o valor de forma direta. Esse resultado permite-nos, em algumas situações, obter indiretamente a probabilidade de um dado evento  $A$ , e, em outras situações, montar equações de recorrência para encontrar a probabilidade de um evento relativo à  $n$ -ésima etapa de um experimento composto (experimento com duas ou mais etapas). Neste trabalho, usaremos a fórmula da probabilidade total com esse intuito.

No que se segue, apresentaremos cinco modelos de urnas os quais exemplificam a modelagem de problemas probabilísticos via equações de recorrência.

Inicialmente, considere a seguinte dinâmica realizada com duas urnas. A urna I contém  $a_1$  bolas azuis e  $v_1$  bolas vermelhas, enquanto que a urna II contém  $a_2$  bolas azuis e  $v_2$  bolas vermelhas. Bolas são extraídas sequencialmente das urnas e recolocadas de acordo com a seguinte regra: com probabilidade  $\alpha$ , retira-se inicialmente uma bola da urna I, e, com probabilidade  $1 - \alpha$ , retira-se uma bola da urna II. As escolhas posteriores são então realizadas de acordo com o esquema de que sempre que uma bola azul for extraída (e devolvida), a próxima bola é sacada da mesma urna. Por outro lado, se a bola extraída (e devolvida) for vermelha, a próxima bola é retirada da outra urna.

Escreva

$$p_1 = \frac{a_1}{a_1 + v_1} \text{ e } p_2 = \frac{a_2}{a_2 + v_2}.$$

Aqui,  $p_1$  é a probabilidade de se extrair ao acaso uma bola azul da urna I. Analogamente,  $p_2$  é a probabilidade de se extrair ao acaso uma bola azul da urna II.

**Teorema 1.** *Seja  $\alpha_n$  a probabilidade de que a  $n$ -ésima bola seja escolhida da urna I. Então,*

$$\alpha_n = \frac{1 - p_2}{2 - p_1 - p_2} + \left[ \alpha - \frac{(1 - p_2)}{2 - p_1 - p_2} \right] (p_1 + p_2 - 1)^{n-1}$$

e

$$\lim_{n \rightarrow \infty} \alpha_n = \frac{1 - p_2}{2 - p_1 - p_2}.$$

*Seja  $\beta_n$  a probabilidade de que a  $n$ -ésima bola escolhida seja azul. Então,*

$$\beta_n = (p_1 - p_2) \left[ \frac{1 - p_2}{2 - p_1 - p_2} + \left( \alpha - \frac{(1 - p_2)}{2 - p_1 - p_2} \right) (p_1 + p_2 - 1)^{n-1} \right] + p_2$$

e

$$\lim_{n \rightarrow \infty} \beta_n = \frac{p_1 - 2p_1p_2 + p_2}{2 - p_1 - p_2}.$$

*Demonstração.* Defina  $B_n$  como o evento onde a  $n$ -ésima bola sorteada vem da urna I. Logo, pela fórmula da probabilidade total,

$$\mathbb{P}(B_{n+1}) = \mathbb{P}(B_n)\mathbb{P}(B_{n+1}|B_n) + \mathbb{P}(B_n^C)\mathbb{P}(B_{n+1}|B_n^C),$$

então,

$$\alpha_{n+1} = p_1\alpha_n + (1-p_2)(1-\alpha_n).$$

Daí,

$$\alpha_{n+1} = (p_1 + p_2 - 1)\alpha_n + (1-p_2).$$

Escreva  $c = p_1 + p_2 - 1$  e  $s = 1 - p_2$ . Então, fica  $\alpha_{n+1} = c\alpha_n + s, n \geq 1$  e  $\alpha_1 = \alpha$ . Daí, temos

$$\begin{aligned} \alpha_n &= c\alpha_{n-1} + s \\ c\alpha_{n-1} &= c^2\alpha_{n-2} + sc \\ &\vdots \\ c^{n-2}\alpha_2 &= c^{n-1}\alpha_1 + c^{n-2}s. \end{aligned}$$

Somando essas equações:

$$\alpha_n = c^{n-1}\alpha_1 + (1 + c + \dots + c^{n-2})s.$$

Logo, usando a fórmula para soma de termos de uma progressão geométrica finita:

$$\alpha_n = c^{n-1}\alpha + \left(\frac{1-c^{n-1}}{1-c}\right)s,$$

e, assim,

$$\alpha_n = \frac{1-p_2}{2-p_1-p_2} + \left[\alpha - \frac{(1-p_2)}{2-p_1-p_2}\right](p_1+p_2-1)^{n-1}.$$

Tomando  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \alpha_n = \lim_{n \rightarrow \infty} \left\{ \frac{1-p_2}{2-p_1-p_2} + \left[\alpha - \frac{(1-p_2)}{2-p_1-p_2}\right](p_1+p_2-1)^{n-1} \right\} = \frac{1-p_2}{2-p_1-p_2}.$$

Agora, seja o evento  $A_n$  onde a  $n$ -ésima bola sorteada é azul. Temos:

$$\mathbb{P}(A_n) = \mathbb{P}(B_n)\mathbb{P}(A_n|B_n) + \mathbb{P}(B_n^C)\mathbb{P}(A_n|B_n^C),$$

assim,

$$\beta_n = p_1\alpha_n + p_2(1-\alpha_n) = (p_1-p_2)\alpha_n + p_2$$

Logo,

$$\beta_n = (p_1-p_2) \left[ \frac{1-p_2}{2-p_1-p_2} + \left(\alpha - \frac{(1-p_2)}{2-p_1-p_2}\right)(p_1+p_2-1)^{n-1} \right] + p_2.$$

Tomando  $n \rightarrow \infty$

$$\lim_{n \rightarrow \infty} \beta_n = \lim_{n \rightarrow \infty} [\alpha_n(p_1-p_2) + p_2] = (p_1-p_2) \lim_{n \rightarrow \infty} \alpha_n + p_2 = (p_1-p_2) \frac{1-p_2}{2-p_1-p_2} + p_2.$$

$$\lim_{n \rightarrow \infty} \beta_n = \frac{p_1 - 2p_1p_2 + p_2}{2-p_1-p_2}.$$

□

Como aplicação do Teorema 1 considere o caso onde  $a_1 = 795, v_1 = 5, a_2 = 598$  e  $v_2 = 2$ . Nesse caso, as duas urnas têm uma proporção muito grande de bolas azuis. Os valores das probabilidades  $\alpha_n$  e  $\beta_n$  bem como seus limites são:

$$\alpha_n = \frac{8}{23} + \frac{7}{46} \left( \frac{2377}{2400} \right)^{n-1}, n \geq 1 \text{ e } \lim_{n \rightarrow \infty} \alpha_n = \frac{8}{23},$$

$$\beta_n = \frac{229}{230} - \frac{49}{109342} \left( \frac{2377}{2400} \right)^n, n \geq 1 \text{ e } \lim_{n \rightarrow \infty} \beta_n = \frac{229}{230}.$$

Os gráficos a seguir apresentam as convergências das funções  $\alpha_n$  e  $\beta_n$ , respectivamente.

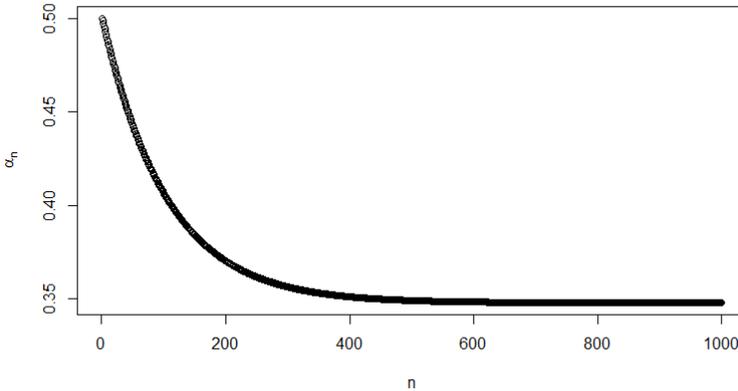


Figura 2: Probabilidade de que a n-ésima bola seja escolhida da urna I em função de n.

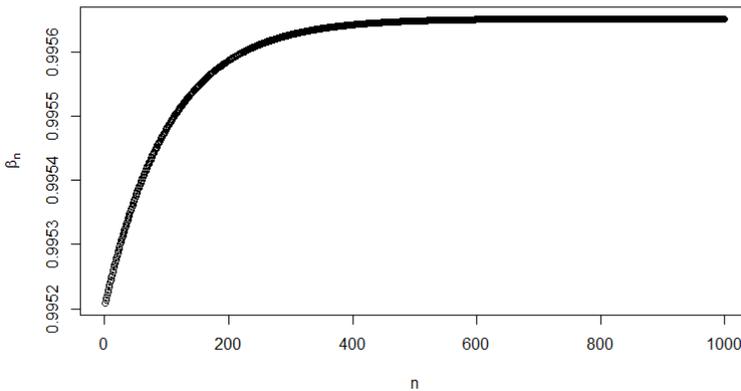


Figura 3: Probabilidade de que a n-ésima bola extraída seja azul em função de n.

Agora, considere uma única urna contendo bolas diferindo unicamente pela cor. Inicialmente há na urna  $a$  bolas azuis e  $v$  bolas vermelhas, onde  $a > v$ . Bolas são extraídas uma a uma ao acaso e sem reposição da urna até que essa fique vazia.

Seja  $p_{a,v}$  a probabilidade de que, ao longo da contagem das bolas extraídas, o número de bolas azuis sempre seja maior do que o número de bolas vermelhas. O teorema a seguir é conhecido como Teorema de Ballot.

**Teorema 2.** *Temos*

$$p_{a,v} = \frac{a-v}{a+v}.$$

*Demonstração.* Defina os eventos  $A$  onde a última bola retirada é azul, e  $B$  onde ao longo da contagem das bolas extraídas, o número de bolas azuis sempre é maior do que o número de bolas vermelhas. Pela fórmula da probabilidade total:

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^C)\mathbb{P}(B|A^C) \\ p_{a,v} &= \frac{a}{(a+v)}p_{a-1,v} + \frac{v}{(a+v)}p_{a,v-1}. \end{aligned}$$

Faremos a prova por indução. Primeiro supomos que para  $a+v = k-1$  o resultado vale, isto é,

$$p(k-1) = p_{a,v} = \frac{a-v}{a+v}.$$

Logo,

$$\begin{aligned} p(k) &= \frac{a}{(a+v)}p_{a-1,v} + \frac{v}{(a+v)}p_{a,v-1} = \frac{a}{(a+v)} \frac{a-1-v}{(a-1+v)} + \frac{v}{(a+v)} \frac{(a-(v-1))}{(a+v-1)} \\ p(k) &= \frac{a}{(a+v)} \frac{(a-1-v)}{(a-1+v)} + \frac{v}{(a+v)} \frac{(a-v+1)}{(a+v-1)} = \frac{a^2 - a - av + av - v^2 + v}{(a+v)(a-1+v)} \\ p(k) &= \frac{(a-v)(a+v) - (a-v)}{(a+v)(a-1+v)} = \frac{a-v}{a+v}. \end{aligned}$$

□

Acabamos de provar uma conjectura, mas como chegar a ela? A ideia é obter primeiro valores de  $p_{a,v}$  para  $a$  e  $v$  pequenos. Com algum esforço obtém-se:

$$p_{2,1} = \frac{1}{3}, p_{3,1} = \frac{2}{4}, p_{3,2} = \frac{1}{5}, p_{4,1} = \frac{3}{5}, p_{4,2} = \frac{2}{6}, \text{ etc.}$$

Daí, fixamos  $v$  e podemos obter

$$p_{a,1} = \frac{a-1}{a+1}, p_{a,2} = \frac{a-2}{a+2}.$$

Com alguma imaginação conjecturamos  $p_{a,v} = \frac{a-v}{a+v}$ .

A tabela a seguir apresenta as probabilidades  $p_{a,v}$  de que, ao longo da contagem das bolas extraídas, o número de bolas azuis seja sempre maior do que o número de bolas vermelhas em função do número de bolas azuis ( $a$ ) e vermelhas ( $v$ ).

a	v	$P_{a,v}$
2	1	0.3333333
3	1	0.5000000
4	1	0.6000000
5	1	0.6666667
3	2	0.2000000
4	2	0.3333333
5	2	0.4285714
4	3	0.1428571
5	3	0.2500000
5	4	0.1111111

Tabela 1: Probabilidades  $p_{a,v}$  de que, ao longo da contagem das bolas extraídas, o número de bolas azuis seja sempre maior do que o número de bolas vermelhas em função do número de bolas azuis (a) e vermelhas (v).

Cabe ressaltar que fixado o número de bolas vermelhas, o valor da probabilidade  $p_{a,v}$  aumenta em função da quantidade de bolas azuis a mais que bolas vermelhas. De fato, podemos escrever

$$h(a, v) := p_{a,v} = 1 - \frac{2v}{a + v}.$$

Para ilustrar a evolução dessas probabilidades em função do número de bolas azuis a, fixamos  $v = 3$  bolas vermelhas. Nesse caso,  $h(a, 3) = 1 - \frac{6}{a + 3}$ .

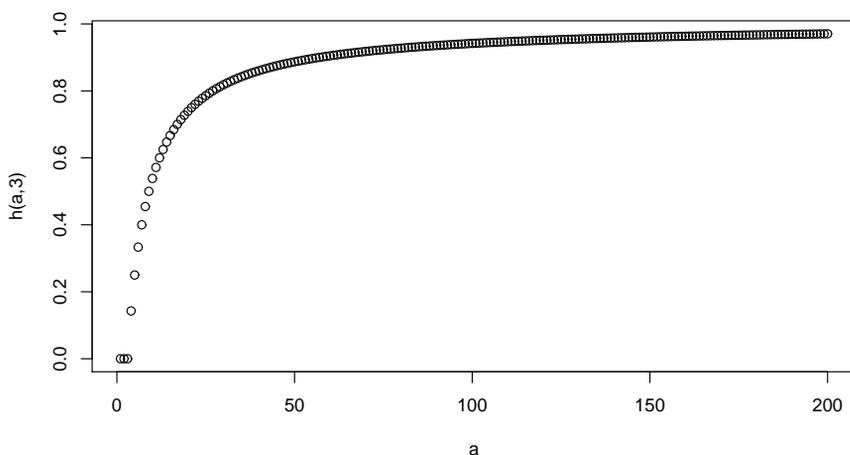


Figura 4: Evolução das Probabilidades  $h(a, v) = p_{a,v}$  para  $v = 3$ .

A seguir considere uma urna contendo bolas diferindo unicamente pela cor. Inicialmente há na urna

a bolas azuis e  $v$  bolas vermelhas. Dois jogadores,  $J_A$  e  $J_V$ , apostam nos resultados de sucessivas extrações dessa urna. Em cada extração, se der bola azul,  $J_A$  ganha 1 unidade monetária de  $J_V$ . Se der bola vermelha,  $J_V$  ganha uma unidade monetária de  $J_A$ . Eles continuam jogando indefinidamente a menos que um deles fique sem dinheiro. Após cada extração, a bola extraída é devolvida.

**Teorema 3.** *Seja  $p_i$  a probabilidade que  $J_A$  termine com todo o dinheiro, tendo começado com  $i$  unidades monetárias e  $J_V$  com  $N - i$  unidades. Temos  $v \neq a$*

$$p_i = \frac{1 - \left(\frac{v}{a}\right)^i}{1 - \left(\frac{v}{a}\right)^N}$$

e se  $v = a$ ,  $p_i = \frac{i}{N}$ . *Seja  $q_i$  a probabilidade de que  $J_V$  termine com todo o dinheiro, quando  $J_A$  começa com  $i$  unidades monetárias e  $J_V$  com  $N - i$  unidades temos para  $a \neq v$ :*

$$q_i = \frac{1 - \left(\frac{a}{v}\right)^{N-i}}{1 - \left(\frac{a}{v}\right)^N}$$

e  $q_i = \frac{N-i}{N}$  para  $a = v$ . *Por fim, se  $r_i$  é a probabilidade de que o jogo nunca termina quando  $J_A$  começa com  $i$  unidades e  $J_V$  com  $N - i$  unidades, então,  $r_i = 0$ .*

*Demonstração.* Suponha que  $A$  seja o evento onde  $J_A$  termina com todo o dinheiro, tendo começado com  $i$  unidades e  $J_V$  com  $N - i$  unidades. Escreva  $p_i = \mathbb{P}(A)$ . Defina  $Z$  como o evento em que a extração da primeira bola resulta em bola azul. Condicionado na extração da primeira bola, temos:

$$p_i = \mathbb{P}(A) = \mathbb{P}(Z)\mathbb{P}(A|Z) + \mathbb{P}(Z^C)\mathbb{P}(A|Z^C) = \frac{a}{a+v}\mathbb{P}(A|Z) + \frac{v}{a+v}\mathbb{P}(A|Z^C).$$

Note que, uma vez que a primeira extração resulta em bola azul, após a primeira aposta  $J_A$  fica com  $i + 1$  unidades enquanto  $J_V$  fica com  $N - (i + 1)$ . Por outro lado, se a extração resultar em bola vermelha, ao fim da primeira aposta,  $J_A$  fica com  $i - 1$  unidades, enquanto  $J_V$  fica com  $N - (i - 1)$ . Como os resultados das apostas sucessivas são independentes, temos

$$\mathbb{P}(A|Z) = p_{i+1} \text{ e } \mathbb{P}(A|Z^C) = p_{i-1}.$$

Logo, podemos escrever

$$p_i = \frac{a}{a+v}p_{i+1} + \frac{v}{a+v}p_{i-1}, i = 1, 2, \dots, N - 1.$$

Temos ainda as condições de contorno  $p_0 = 0$  e  $p_N = 1$ .

Assim,

$$p_{i+1} - p_i = \frac{v}{a}(p_i - p_{i-1}), i = 1, 2, \dots, N - 1.$$

Logo, como  $p_0 = 0$ :

$$\begin{aligned}
 p_2 - p_1 &= \frac{v}{a}(p_1 - p_0) = \frac{v}{a} \\
 p_3 - p_2 &= \frac{v}{a}(p_2 - p_1) = \left(\frac{v}{a}\right)^2 p_1 \\
 &\vdots \\
 p_i - p_{i-1} &= \frac{v}{a}(p_{i-1} - p_{i-2}) = \left(\frac{v}{a}\right)^{i-1} p_1 \\
 &\vdots \\
 p_N - p_{N-1} &= \frac{v}{a}(p_{N-1} - p_{N-2}) = \left(\frac{v}{a}\right)^{N-1} p_1.
 \end{aligned}$$

Somando as  $i - 1$  primeiras equações,

$$p_i - p_1 = p_1 \left[ \left(\frac{v}{a}\right) + \left(\frac{v}{a}\right)^2 + \dots + \left(\frac{v}{a}\right)^{i-1} \right].$$

Assim, se  $a \neq v$ , aplicamos a soma dos termos de uma progressão geométrica finita e obtemos

$$p_i = \left[ \frac{1 - \left(\frac{v}{a}\right)^i}{1 - \frac{v}{a}} \right] p_1.$$

Como  $p_N = 1$ , temos

$$p_1 = \frac{1 - \frac{v}{a}}{1 - \left(\frac{v}{a}\right)^N},$$

e, assim,

$$p_i = \frac{1 - \left(\frac{v}{a}\right)^i}{1 - \left(\frac{v}{a}\right)^N}.$$

Por outro lado, se  $a = v$ ,  $p_i = ip_1$ , donde  $p_1 = \frac{1}{N}$ , e, assim,  $p_i = \frac{i}{N}$ .

De modo inteiramente similar, se escrevermos  $q_i$  como a probabilidade de que  $J_V$  termine com todo o dinheiro, quando  $J_A$  começa com  $i$  unidades e  $J_V$  com  $N - i$  unidades, obtemos para  $a \neq v$ :

$$q_i = \frac{1 - \left(\frac{a}{v}\right)^{N-i}}{1 - \left(\frac{a}{v}\right)^N},$$

e  $q_i = \frac{N-i}{N}$  para  $a = v$ . Por fim, note que existem três possibilidades:  $J_A$  vence,  $J_V$  vence ou o jogo continua indefinidamente sem que ninguém vença. Seja  $r_i$  a probabilidade de que o jogo nunca termine. Para  $a \neq v$ , de  $p_i + q_i + r_i = 1$  temos  $r_i = 0$ , pois

$$p_i + q_i = \frac{1 - \left(\frac{v}{a}\right)^i}{1 - \left(\frac{v}{a}\right)^N} + \frac{1 - \left(\frac{a}{v}\right)^{N-i}}{1 - \left(\frac{a}{v}\right)^N} = 1.$$

Para  $a = v$ :

$$p_i + q_i = \frac{i}{N} + \frac{N-i}{N} = 1.$$

E, nesse caso, também,  $r_i = 0$ .

□

O Teorema 3 apresenta uma propriedade muito interessante para o caso onde  $v < a$ . Mesmo que o jogador  $J_V$  inicie com uma quantidade muito grande de dinheiro, ele tem chance de sair derrotado da disputa e perder todo seu dinheiro. De fato, se  $v < a$ , fazendo  $N \rightarrow \infty$  temos:

$$\lim_{N \rightarrow \infty} p_i = \lim_{N \rightarrow \infty} \frac{1 - \left(\frac{v}{a}\right)^i}{1 - \left(\frac{v}{a}\right)^N} = 1 - \left(\frac{v}{a}\right)^i \text{ e}$$

$$\lim_{N \rightarrow \infty} q_i = \lim_{N \rightarrow \infty} \frac{1 - \left(\frac{a}{v}\right)^{N-i}}{1 - \left(\frac{a}{v}\right)^N} = \left(\frac{v}{a}\right)^i.$$

Agora, considere uma urna contendo bolas diferindo unicamente pela cor. Inicialmente há na urna  $a$  bolas azuis e  $v$  bolas vermelhas. Bolas são extraídas uma a uma ao acaso e com reposição.

**Teorema 4.** *Seja  $p_n$  a probabilidade de que em  $n$  extrações saiam um número par de bolas azuis. Então,*

$$p_n = \frac{1 + \left(\frac{v-a}{v+a}\right)^n}{2}, n = 1, 2, \dots$$

Além disso,

$$\lim_{n \rightarrow \infty} p_n = \frac{1}{2}.$$

*Demonstração.* Defina  $A$  o evento em que a primeira retirada resulta em bola azul e  $B$  o evento o qual em  $n$  extrações saia um número par de bolas azuis. Temos  $p_n = \mathbb{P}(B)$ . Pela fórmula da probabilidade total temos

$$\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^C)\mathbb{P}(B|A^C).$$

Note que  $\mathbb{P}(B|A) = 1 - p_{n-1}$ . De fato, para  $B$  ocorrer, dado que a primeira retirada resultou em bola azul, é preciso que nas próximas  $n-1$  extrações saia um número ímpar de bolas azuis, evento cuja probabilidade é  $1 - p_{n-1}$ . Por outro lado,  $\mathbb{P}(B|A^C) = p_{n-1}$  pois para  $B$  ocorrer, dado que a primeira retirada resultou em bola vermelha, é preciso que nas próximas  $n-1$  extrações saia um número par de bolas azuis, evento cuja probabilidade é  $p_{n-1}$ . Logo,

$$p_n = \frac{a}{v+a}(1 - p_{n-1}) + \frac{v}{v+a}p_{n-1},$$

$$p_n = \frac{a}{(v+a)} + \frac{(v-a)}{(v+a)}p_{n-1}.$$

Escreva  $c = \frac{a}{v+a}$  e  $b = \frac{v-a}{v+a}$ .

$$\begin{aligned}
 p_n &= c + bp_{n-1} \\
 b \cdot p_{n-1} &= bc + b^2 p_{n-2} \\
 b^2 \cdot p_{n-2} &= b^2 c + b^3 p_{n-3} \\
 &\vdots \\
 b^{n-2} \cdot p_2 &= b^{n-2} c + b^{n-1} p_1.
 \end{aligned}$$

Somando essas equações

$$p_n = c + cb + cb^2 + \dots + cb^{n-2} + b^{n-1}p_1$$

Aplicando a soma dos termos de uma progressão geométrica finita

$$p_n = c \left( \frac{1 - b^{n-1}}{1 - b} \right) + b^{n-1}p_1.$$

Como  $p_1 = \frac{v}{v+a}$ ,

$$p_n = \frac{1 + \left(\frac{v-a}{v+a}\right)^n}{2}, n = 1, 2, \dots$$

Por fim,

$$\lim_{n \rightarrow \infty} p_n = \lim_{n \rightarrow \infty} \frac{1}{2} + \frac{1}{2} \left( \frac{v-a}{v+a} \right)^n = \frac{1}{2} + \frac{1}{2} \lim_{n \rightarrow \infty} \left( \frac{v-a}{v+a} \right)^n = \frac{1}{2}.$$

□

Para exemplificar o resultado do Teorema 4 considere o caso o qual temos  $a = 995$  bolas azuis e  $v = 5$  bolas vermelhas. O gráfico a seguir mostra o comportamento da função  $p(n) = p_n$  que dá a probabilidade de que em  $n$  extrações saia um número par de bolas azuis em função do número de extrações  $n$ . Nesse caso,

$$p(n) = \frac{1}{2} + \frac{1}{2} \left( \frac{-99}{100} \right)^n, n \geq 1.$$

Note que a função  $p(n)$  assume valores maiores que  $\frac{1}{2}$  para  $n$  par e valores menores que  $\frac{1}{2}$  para  $n$  ímpar. Essa interessante propriedade vale em geral para todos os casos em que o número de bolas azuis é maior que o número de bolas vermelhas. Por outro lado, se o número de bolas vermelhas é maior que o número de bolas azuis,  $p(n)$  é uma função decrescente que converge para  $\frac{1}{2}$ .

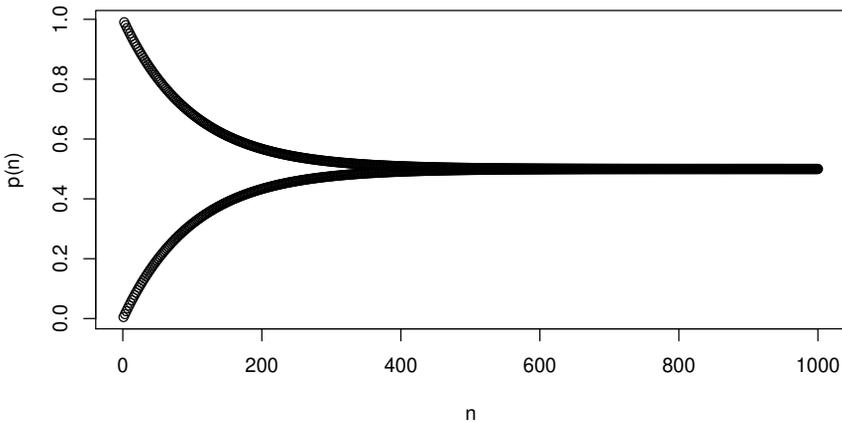


Figura 5: Evolução das Probabilidades  $h(a, v) = p_{a,v}$  para  $v = 3$ .

Por fim, considere uma sequência de urnas numeradas de 1 até  $n$  tal que a urna de número  $i$  ( $i = 1, 2, \dots, n$ ) possui  $2i$  bolas azuis e 1 bola vermelha. Uma bola é extraída ao acaso de cada urna.

**Teorema 5.** *Seja  $p_n$  a probabilidade de que o número de bolas vermelhas extraídas seja ímpar. Então,*

$$p_n = \frac{n}{2n+1}, n = 1, 2, \dots$$

*Demonstração.* Defina  $A$  o evento onde a bola retirada na urna  $n$  é azul e  $B$  o evento onde nas  $n$  extrações saia um número ímpar de bolas vermelhas. Temos  $p_n = \mathbb{P}(B)$ . Pela fórmula da probabilidade total temos

$$\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^C)\mathbb{P}(B|A^C).$$

Note que  $\mathbb{P}(B|A^C) = 1 - p_{n-1}$ . De fato, para  $B$  ocorrer, dado que a retirada na  $n$ -ésima urna resultou em bola vermelha, é preciso que nas próximas  $n-1$  extrações anteriores tenha saído um número par de bolas vermelhas, evento cuja probabilidade é  $1 - p_{n-1}$ . Por outro lado,  $\mathbb{P}(B|A) = p_{n-1}$ , pois para  $B$  ocorrer, dado que a retirada na  $n$ -ésima urna resultou em bola azul, é preciso que nas  $n-1$  extrações anteriores tenha saído um número ímpar de bolas vermelhas, evento cuja probabilidade é  $p_{n-1}$ . Logo,

$$p_n = \left(\frac{2n}{2n+1}\right)p_{n-1} + \left(\frac{1}{2n+1}\right)(1 - p_{n-1}).$$

E, assim,

$$(2n+1)p_n - (2n-1)p_{n-1} = 1.$$

Daí,

$$\begin{aligned} (2n+1)p_n - (2n-1)p_{n-1} &= 1 \\ (2n-1)p_{n-1} - (2n-3)p_{n-2} &= 1 \\ (2n-3)p_{n-2} - (2n-5)p_{n-3} &= 1 \\ &\vdots \\ 5p_2 - 3p_1 &= 1. \end{aligned}$$

Como  $p_1 = \frac{1}{3}$ , somando essas  $n-1$  equações obtemos

$$p_n = \frac{n}{2n+1}, n = 1, 2, \dots$$

□

Uma consequência do Teorema 5 é que a probabilidade de que o número de bolas vermelhas extraídas seja ímpar converge rapidamente para  $\frac{1}{2}$ .

### 3. Considerações Finais

Este trabalho apresenta modelos de simples compreensão com resultados bastante interessantes. Ele pode servir de motivação para estudantes que queiram se dedicar ao estudo da teoria de probabilidade. Modelos de urnas fornecem problemas de fácil entendimento. As abordagens estão

relacionadas com problemas presentes no nosso cotidiano e fornecem base para cálculos probabilísticos envolvendo loterias bem como problemas associados à amostragem. A abordagem apresentada por nós torna-se interessante também por associar cálculo probabilístico ao uso de propriedades de seqüências numéricas. Assim, mostramos uma das muitas relevantes aplicações do estudo de seqüências e séries.

## Referências

- [1] Almeida, N., Degenszajn, D., Dolce, O., Iezzi, G., Périgo, R.; *Matemática: ciência e aplicações*, Saraiva, Volume 2, nona edição, 2016.
- [2] Carvalho, P.C.P., Fernandez, P., Morgado, A.C.O., Pitombeira, J.B.; *Análise Combinatória e Probabilidade*. SBM, décima edição, 2016.
- [3] Charalambides, A. C.; *Combinatorial Methods in Discrete Distributions*. Wiley, primeira edição, 2005.
- [4] Clarke, B.; Fokoue, E.; Zhang, H. H.; *Principles and theory for data mining and machine learning*. Springer Science & Business Media, 2009.
- [5] Dante, L.R.; *Matemática: contexto & aplicações: ensino médio*, Ática, Volume 2, terceira edição, 2016.
- [6] Diniz, M.I., Smole, K.S. *Matemática para compreender o mundo*, Saraiva, Volume 2, primeira edição, 2016.
- [7] Hazzan, S. *Fundamentos de matemática elementar: combinatória e probabilidade*. Atual editora, Volume 5, 2013.
- [8] Heubeck, K.; *Urnenmodelle und ihre Anwendung in der Versicherungsmathematik*, BI. Deut. Ges. Versicher. math., 11 (3), 37 1-429, 1974.
- [9] Hazzan, S.; Iezzi, G.; *Fundamentos de matemática elementar: Sequências, matrizes, determinantes e sistemas*. Atual editora, Volume 4, 2012.
- [10] Johnson, N.L.; Kotz, S.; *Urn Models and their Application: An Approach to Modern Discrete Probability Theory*. New York and London, Wiley, 1977. Wiley Series in Probability and Mathematical Statistics, 1977.
- [11] dos Santos, J. P. O.; Mello, M. P. ; Murari, I.; *Introdução à análise combinatória*. Editora Ciência Moderna, primeira edição, 2008.
- [12] Polya, G. *Patterns of Plausible Inference*. Princeton, N.J.: Princeton University Press, 1954.
- [13] Ross, S.M. *Probabilidade- Um Curso Moderno com Aplicações*. Bookman, 2010.
- [14] Sprott, D. A.; *Probability distributions associated with distinct hits on targets*, Bull. Math. Biophys., 19, 163-170, 1957.
- [15] Stirzaker, D.; *Elementary Probability*. Cambridge University Press , 2003.
- [16] Morgado, A.C.O.; Wagner, E.; Zani, S.; *Progressões e Matemática Financeira*. SBM, sexta edição, 2015.

Valdivino Vargas Júnior  
Universidade Federal de Goiás  
<[vvjunior@ufg.br](mailto:vvjunior@ufg.br)>

Tiago Moreira Vargas  
Universidade Federal de Goiás  
<[vargas@ufg.br](mailto:vargas@ufg.br)>

Divaldo P. Fernandes Júnior  
Colégio Fractal  
<[profjr77@gmail.com](mailto:profjr77@gmail.com)>

Recebido: 14/02/2020  
Publicado: 21/08/2019

# Um truque com cartas usando a base numérica 3

Miguel V. S. Frasson

Lee Y. Sheng

## Resumo

Com habilidade em representar números na base 3, um mágico pode executar um impressionante truque com cartas apresentado neste trabalho.

**Palavras-chave:** Base numérica três; truque de mágica; cartas.

## Abstract

With the ability to represent numbers in base 3, a magician can perform a spectacular card trick revealed in this work.

**Keywords:** Number base three; magic trick; cards.

## 1. Introdução

A mágica sempre encantou o homem, mesmo os truques simples. Entre tantos truques, há os que se baseiam em conceitos matemáticos, os quais podem despertar o interesse pela Matemática uma vez que o truque seja embasado matematicamente [1, 2]. Entre os truques que têm conceitos matemáticos estão o truque do “número pensado” [3], “do calendário” [4], aqueles que usam cartas/baralho [5, 6], entre vários outros [4, 7].

Apresentamos aqui um truque mágico que utiliza a decomposição na base 3 para transportar a carta desconhecida a posição escolhida, montando um “truque mágico” que certamente encantará a plateia, já apresentado em [8]. O truque requer apenas um pouco de prática mental de se escrever números pequenos na base 3. Ao final, apresentamos uma generalização para qualquer base.

## 2. A matemática de uma distribuição de cartas

Tomemos um baralho de 27 cartas, viradas para baixo e vamos numerá-las sequencialmente de 0 a 26, de cima para baixo. Vamos distribuí-las em três montes em sequência. Veja a Figura 1. Os montes receberão as seguintes cartas:

1º monte:  $0, 3, 6, \dots, 3q, \dots, 24;$

2º monte:  $1, 4, 7, \dots, 3q + 1, \dots, 25;$

3º monte:  $2, 5, 8, \dots, 3q + 2, \dots, 26.$

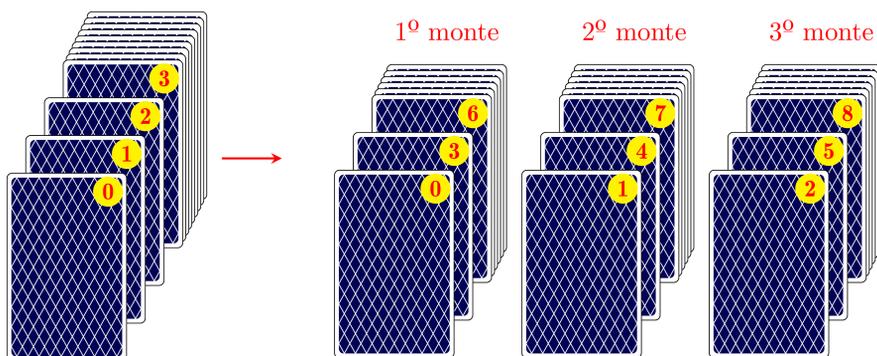


Figura 1: Distribuição das cartas em seus respectivos montes. Os montes recebem cartas com numeração da forma  $3q$ ,  $3q + 1$  e  $3q + 2$ , respectivamente.

Observe que, em cada monte, a carta  $3q + r$  passará a ocupar a numeração  $q$  (ao reiniciarmos a numeração em cada monte), isto é, o quociente da antiga numeração por 3.

Escolha um dos montes e volte a empilhá-los. Se o monte escolhido ficar no topo, a carta originalmente na posição  $3q + r$  ficará na posição  $q = 0 \cdot 9 + q$ . Se o monte escolhido ficar no meio, como os montes têm nove cartas, a carta originalmente na posição  $3q + r$  ficará na posição  $1 \cdot 9 + q$ . Se o monte escolhido for posto embaixo, a carta originalmente na posição  $3q + r$  ficará na posição  $2 \cdot 9 + q$ . Assim, marcaremos as posições do monte escolhido com  $p = 0$ ,  $p = 1$  ou  $p = 2$  de acordo com o posicionamento no topo, no meio ou embaixo, respectivamente.

Vejam como essa distribuição de cartas reordena os números das cartas escritos na base 3. Seja um número  $n$  entre 0 e 26 e escreva-o na base 3, com dígitos  $a, b, c$ :  $n = (abc)_3$ . A separação em montes coloca  $n$  na posição de seu quociente por três num dos montes:

$$(abc)_3 = 9a + 3b + c \mapsto q = 3a + b = (0ab)_3.$$

Se o monte que contém essa carta for parar na posição  $p$ , soma-se  $9p$  a  $(0ab)_3$  e o posicionamento na pilha final será  $(pab)_3$ , ou seja:

**Uma distribuição que coloca o monte escolhido na posição  $p$ , acrescenta  $p$  à esquerda e descarta o dígito final!**

$$(abc)_3 \mapsto (pab)_3$$

Recorde que a numeração das cartas começa em zero. Se desejarmos deslocar a carta que começa na numeração  $n = (abc)_3$  para a posição  $m$ , que corresponde à numeração  $m - 1 = (xyz)_3$ , basta repetir a distribuição três vezes, colocando o monte em que aparece a carta escolhida nas posições correspondendo à ordem inversa dos dígitos de  $m - 1$ , ou seja,  $z, y$  e  $x$ :

$$n = (abc)_3 \xrightarrow{1^{\text{a}}: z} (zab)_3 \xrightarrow{2^{\text{a}}: y} (yza)_3 \xrightarrow{3^{\text{a}}: x} (xyz)_3 = m - 1$$

que corresponde à  $m$ -ésima posição do baralho, de cima para baixo. Isso explica o truque que descreveremos adiante.

### 3. Fazendo uma distribuição na prática

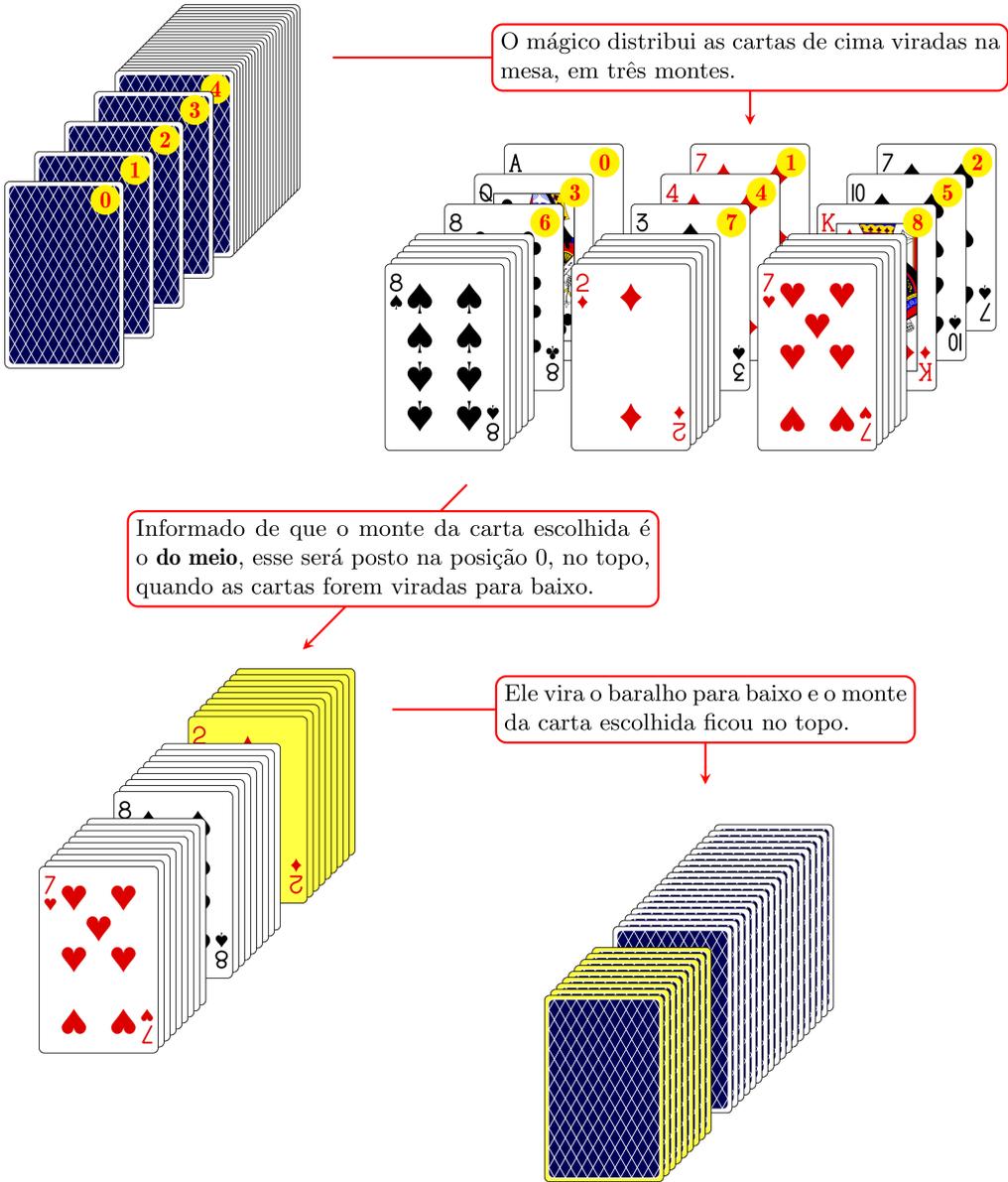


Figura 2: Uma distribuição para colocar o monte da carta escolhida na posição 0 = no topo. O participante informou que a carta escolhida está no monte do meio.

Para efetuar uma distribuição, com o baralho virado para baixo em suas mãos, o mágico deve retirar as cartas de cima e pô-las na mesa, empilhando-as nos três montes enquanto o participante observa, de forma que esse pode ver todas as cartas e apontar em que monte a carta que escolheu

está. Além disso, a ordenação da numeração das cartas é mantida.

Ao final, tendo sido revelado o monte, o mágico deve lembrar-se que as posições 0 = no topo, 1 = no meio e 2 = embaixo referem-se ao baralho virado para baixo quando for posicionar os montes com as cartas ainda sobre a mesa viradas para cima. Veja na Figura 2 a demonstração de uma distribuição.

#### 4. Decompondo um número na base 3 mentalmente

É fácil decompor mentalmente números pequenos na base 3. Um bom método é pelos sucessivos restos da divisão por 3, sendo o próximo dividendo o quociente da primeira divisão. Com efeito, o resto por 3 é o quanto o dividendo excede um múltiplo de 3 que está na tabuada e o quociente é imediato dela. Esse método tem a vantagem de dar os dígitos na ordem para executar o truque. Por exemplo, para o número 17, 17 excede 15 em 2,  $15 \div 3 = 5$ , e 5 excede 3 em 2, com quociente 1. Portanto, o truque deve ser feito com as posições 2, 2 e 1 ( $17 = (122)_3$ ).

#### 5. Execução do truque

1. Selecione um participante da plateia. Mostre um baralho com 27 cartas e este deve escolher, em segredo, uma das cartas, podendo revelá-la à plateia. O baralho pode ser rapidamente embaralhado pelo próprio participante.
2. O mágico pergunta um número (que chamaremos aqui de  $m$ ) – digamos até 25 – a ele ou a outra pessoa da plateia. O número pode ser anotado numa lousa.
3. Mentalmente o mágico decompõe o antecessor  $m - 1$  na base 3:  $m - 1 = (xyz)_3$ .
4. O mágico deve executar três distribuições como explicadas acima para as posições  $z$ ,  $y$  e  $x$ , nessa ordem, perguntando ao participante ao final de cada distribuição em qual monte está a carta que escolheu.
5. Ao final das três distribuições, com o baralho em mãos virado para baixo, o mágico retira as cartas de cima uma a uma, contando de 1 a  $m$ , virando apenas a  $m$ -ésima carta, perguntando no ato se é esta a carta escolhida, o que o espantado participante confirmará.

A Figura 3 ilustra uma execução completa do truque.

#### 6. Uma generalização para $b^k$ cartas

O truque pode estender-se para qualquer base. A base  $b$  é o número de montes, e, para um baralho de  $b^k$  cartas, precisamos de  $k$  distribuições. O desafio é decompor números na base  $b$  mentalmente.

Um caso interessante é fazer o truque com  $b^2$  cartas, com  $b$  montes e apenas duas distribuições, uma vez que decompor números de apenas dois dígitos na base  $b$  é uma questão de quociente e resto na divisão por  $b$ :

$$0 \leq n < b^2 : \quad n = bq + r, \quad 0 \leq r < b \iff n = (qr)_b.$$

Como exercício, demonstre a generalização e aprenda também o truque com 16 ou 25 cartas e duas distribuições.

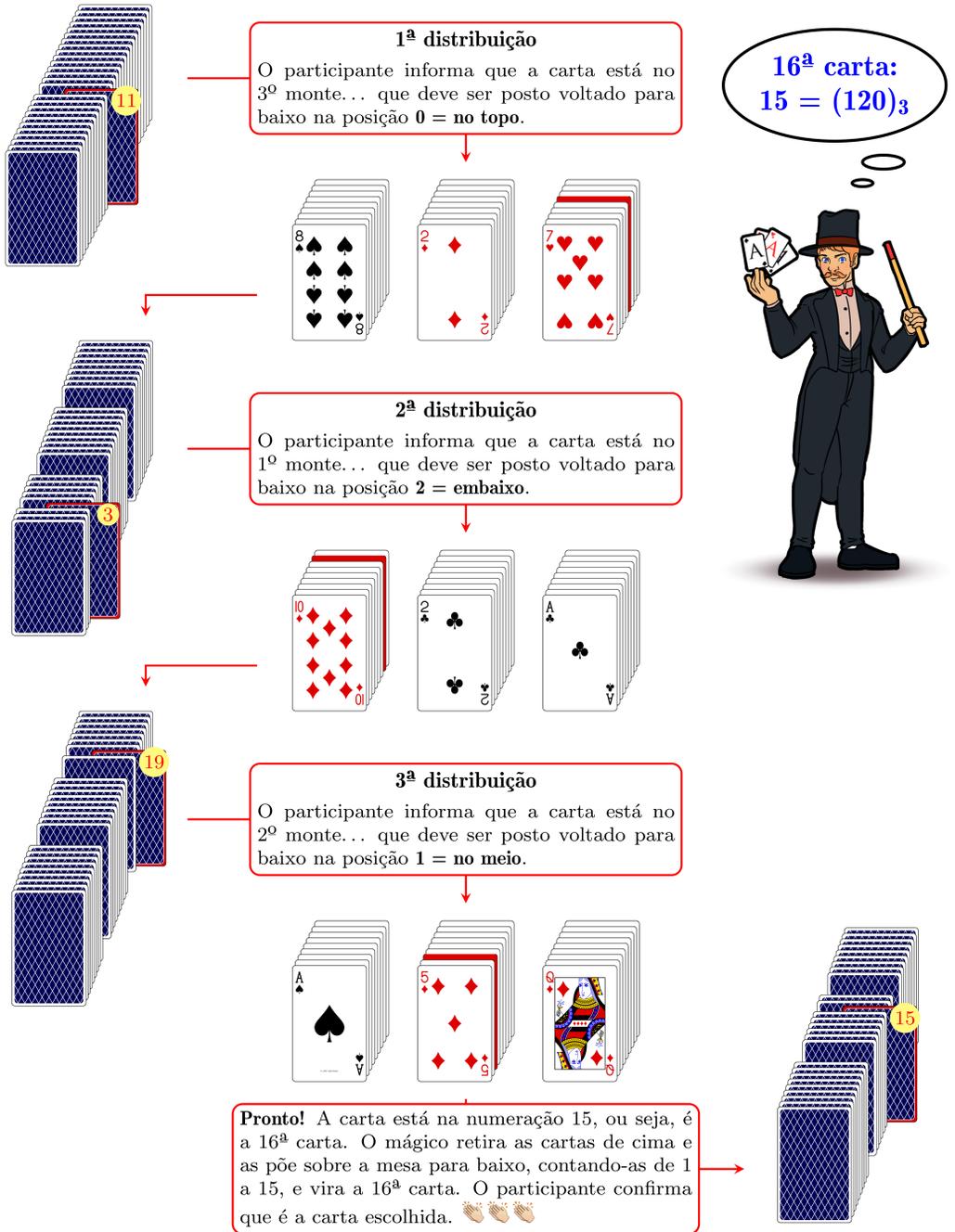


Figura 3: Um exemplo do truque em sua totalidade. Supomos que foi escolhido o número 16. O mágico decompõe  $15 = (120)_3$  e deve fazer as distribuições para as posições 0, 2 e 1, nessa ordem. Supondo que a carta escolhida iniciou na posição  $11 = (102)_3$ , as posições ocupadas por ela após cada distribuição serão  $11 = (102)_3 \xrightarrow{1^a: 0} (010)_3 = 3 \xrightarrow{2^a: 2} (201)_3 = 19 \xrightarrow{3^a: 1} (120)_3 = 15$ .

## Referências

- [1] CARVALHO, M. P. de; SANTOS, J. M. "Dinâmica de um truque de cartas." *Boletim da Sociedade Portuguesa de Matemática*, n° 70, 2014.
- [2] FIALHO, P. M. S. *Matemática e Embaralhamento de Cartas: de Mágicas a Cadeias de Markov*. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, 2016.
- [3] RODRIGUES, J. C. F. "Adivinhando o número pensado: uma contribuição para atividades em sala de aula." *Educação Matemática em Revista*, pp.152–165, 2017.
- [4] SAMPAIO, J. C.; MALAGUTTI, P. L. *Mágicas, matemática e outros mistérios*. São Carlos: EdUFSCar, 2008.
- [5] ALVES, A. G. *Mágicas matemáticas como metodologia de ensino*. Dissertação (ProfMat) — Universidade Federal de São Carlos, São Carlos, 2015.
- [6] SHENG L. Y.; FRASSON, M. V. S. "Uma mágica com cartas usando a base dois." *Revista do Professor de Matemática*, v. 99, 2019.
- [7] MCOWAN, P.; PARKER, M. *The manual of Mathematical Magic*. [S.l.]: [www.mathematicalmagic.com/](http://www.mathematicalmagic.com/), 2010.
- [8] GARDNER, M. *Mathematics, Magic and Mystery*. Nova York: Dover, 1956.

Miguel V. S. Frasson  
ICMC – Universidade de São Paulo  
<[frasson@icmc.usp.br](mailto:frasson@icmc.usp.br)>

Lee Y. Sheng  
Universidade Federal do Mato Grosso  
<[leeufmt@yahoo.com.br](mailto:leeufmt@yahoo.com.br)>

Recebido: 08/06/2020  
Publicado: 28/08/2020