

SOBRE A CONVERGÊNCIA DAS CADEIAS DE MARKOV

F. S. BATISTA C. C. SANTOS L. ANDRADE¹
J. C. ARAÚJO R. G. MÁRQUEZ²

Resumo

As cadeias de Markov são representadas por matrizes quadradas P_{ij} de ordem r e indicam a probabilidade de transição do estado e_j para o estado e_i . O objetivo deste artigo é provar o teorema da convergência de matrizes estocásticas regulares usando tão somente o produto matricial e conhecimentos elementares de análise real, em particular o uso de desigualdades na reta. Uma aplicação prática é apresentada a luz dos três métodos utilizados nesse estudo para a obtenção do vetor estacionário.

1. Introdução

A modelagem matemática utiliza equações matemáticas com vistas a estudar o comportamento do sistema físico analisado. Um sistema físico pode mudar com o tempo de um estado a outro de tal modo que, sua modelagem pode se tornar uma tarefa difícil, ou até mesmo impraticável. Caso o estado desse sistema em qualquer observação não puder ser predito com certeza, mas se for possível prever a sua probabilidade de ocorrer unicamente a partir do conhecimento do estado do sistema na observação precedente, então, esse processo de mudança (probabilística) de um estado para o outro, é conhecido por cadeia (ou processo) de Markov. Várias são as aplicações das cadeias de Markov, entre as quais estão a Física, que utiliza os sistemas Markovianos em termodinâmica e mecânica e a Ciência da Informação que também utiliza a cadeia de Markov em todo o processamento da informação. A Cadeia de Markov também é bastante utilizada na Economia para estudar diferentes fenômenos, como preços ativos e falhas no mercado [5].

Certamente, um resultado importante nessa teoria, é a convergência dessas cadeias. Recentemente, Silva e Rota [7] analisaram esse problema com modificações em relação ao desenvolvimento descrito por [2] e [4] para calcularem $T^n v$ (em vez de vT^n), onde T representa uma matriz quadrada de ordem r de vetores de probabilidade, como fizeram os autores reportados anteriormente.

Entretanto, a demonstração apresentada por [7] mostrou-se, no nosso entendimento, demasiadamente “onerada”. Deve ser observado que o desenvolvimento proposto por [2] foi apresentado de forma concisa, isto é, omitindo passagens de crucial importância no desenvolvimento de sua demonstração. Nesse sentido, é apresentada uma demonstração detalhada do teorema da convergência da cadeia de Markov com as

Palavras-chave: Matriz de transição; Vetor de Estado Estacionário; Convergência.

¹ FFP-UERJ, fabiana_s_b_@hotmail.com; cris_campos@hotmail.com; lorrene.alc@hotmail.com

²DMAT/FFP-UERJ, jcaraujo_55@yahoo.com.br; rosagmarquez@yahoo.com.br

alterações necessárias em relação ao desenvolvimento descrito por [2] para ser adaptada ao produto $T^n v$. Uma aplicação é proposta e discutida à luz dos diferentes métodos para a obtenção do vetor de estado estacionário.

2. Método

Sejam $X_1, X_2, \dots, X_n, n \in \mathbb{N}$ variáveis aleatórias discretas de probabilidade definidas em um conjunto $E = \{e_1, \dots, e_k\}$, onde $e_i, i = 1, \dots, k$ são os estados probabilísticos possíveis de um evento ε . Uma cadeia de Markov é definida pela probabilidade condicional dada por:

$$P(X_n = x_n | X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0). \quad (1)$$

A probabilidade de o sistema estar no estado i (ou e_i) em qualquer observação, se na observação precedente estava no estado j (ou e_j), é denotada por P_{ij} , e é chamada a probabilidade de transição do estado e_j ao estado e_i . A equação (1) indica que tudo que acontece no futuro dependerá do último estado presente e não do passado [6].

A matriz quadrada $P = (P_{ij})_{k \times k}$, de ordem k , com $P_{ij} \geq 0$ e $\sum_{i=1}^k P_{ij} = 1$, para $j = 1, \dots, k$ é chamada matriz de transição ou matriz estocástica da cadeia de Markov.

A matriz de transição P é regular se existe $n \in \mathbb{N}$, tal que, $P_{ij}^n > 0$, para todo $i, j = 1, \dots, k$, isto é P^n tem todas as entradas estritamente positivas.

2.1. Formulação analítica

Proposição

O produto de matrizes estocásticas é estocástica.

Demonstração: Sejam A e B matrizes estocásticas de ordem k . Do produto matricial [1], um elemento genérico $(AB)_{ij}$ desse produto é dado por

$$(AB)_{ij} = \sum_{r=1}^k A_{ir} B_{rj}, \quad (2)$$

onde, A_{ir} e B_{rj} são respectivamente a i -ésima linha de A e a j -ésima coluna de B .

Queremos provar que a soma dos elementos de cada coluna da matriz AB é igual a um, isto é, $\sum_{i=1}^k (AB)_{ij} = 1$.

$$\begin{aligned} \sum_{i=1}^k (AB)_{ij} &= \sum_{i=1}^k \left(\sum_{r=1}^k A_{ir} B_{rj} \right) \\ &= \sum_{r=1}^k \left(\sum_{i=1}^k A_{ir} B_{rj} \right) \\ &= \sum_{r=1}^k \left(\sum_{i=1}^k A_{ir} \right) B_{rj} \\ &= \sum_{r=1}^k (1) \cdot B_{rj} = 1. \end{aligned}$$

Essa proposição, aparentemente ingênua, mostra que se P é regular, P^n é também estocástica para algum n natural. De fato, sendo P uma matriz de transição regular, P é uma matriz estocástica e existe n natural, tal que $P_{ij}^n > 0$, para todo $i, j = 1, \dots, k$. Daí, e da proposição anterior, resulta que a matriz P_{ij}^n é estocástica.

O próximo Teorema nos permitirá demonstrar o Teorema de Convergência de Matrizes Regulares (tais teoremas foram adaptados para o produto Px em vez de xP , como desenvolvido por [2]).

Teorema 1

Seja $P_{r \times r}$ uma matriz de transição de ordem r não tendo entradas nulas. Sejam $\varepsilon = \min_{i,j=1,\dots,r} (P_{ij})$ e x um vetor linha com $m_o = \min_{i=1,\dots,r} (x_i)$ e $M_o = \max_{i=1,\dots,r} (x_i)$. Analogamente, sejam $m_1 = \min_{j=1,\dots,r} (xP)$ e $M_1 = \max_{j=1,\dots,r} (xP)$. Então,

$$M_1 \leq M_o, \quad m_1 \geq m_o, \quad (3)$$

$$M_1 - m_1 \leq (1 - 2\varepsilon)(M_o - m_o). \quad (4)$$

Demonstração: Seja x' o vetor linha obtido de x onde todas as suas componentes foram substituídas por M_o , exceto uma que continuou com m_o .

Sem perda de generalidade, podemos supor que o m_o ocupe a primeira coordenada de x' e que $P_{1j} = a$. Assim, a j -ésima componente de $x'P$, é dada pelo produto do vetor x' com a j -ésima coluna da matriz P , indicada por $P(., j)$.

$$(x'P)_j = x'P(., j) = am_o + \sum_{i=2}^r M_o P_{ij} = am_o + M_o \sum_{i=2}^r P_{ij}. \quad (5)$$

Como $\sum_{j=1}^r P_{ij} = 1$, tem-se que $\sum_{j=2}^r P_{ij} = 1 - a$.

Daí,

$$(x'P)_j = am_o + M_o(1 - a) = M_o - a(M_o - m_o). \quad (6)$$

Desde que $a \geq \varepsilon$, pois $a \in P$, obtemos:

$$-a(M_o - m_o) \leq -\varepsilon(M_o - m_o) \quad (7)$$

ou

$$(x'P)_j \leq M_o - \varepsilon(M_o - m_o). \quad (8)$$

Em particular, como M_1 é a maior das coordenadas do vetor $x'P$ satisfaz a desigualdade

$$M_1 \leq M_o - \varepsilon(M_o - m_o). \quad (9)$$

Como, $\varepsilon(M_o - m_o) > 0$, temos que

$$M_1 \leq M_o. \quad (10)$$

Se o vetor linha x for multiplicado por -1 , a desigualdade (9) fica na forma:

$$-m_1 \leq -m_o - \varepsilon(M_o - m_o). \quad (11)$$

Para verificar a desigualdade (11), basta usar a desigualdade (9) observando que:

$$\max_{i=1,\dots,r}(-x') = -m_o \quad (= "M_o")$$

$$\max_{j=1,\dots,r}(-x'P) = -m_1 \quad (= "M_1")$$

$$\min_{i=1,\dots,r}(-x') = -M_o \quad (= "m_o")$$

$$\min_{j=1,\dots,r}(-x'P) = -M_1 \quad (= "m_1")$$

Da desigualdade (11) obtemos

$$m_1 \geq m_o. \quad (12)$$

Adicionando $-m_1$ às desigualdades (9) e (10), temos:

$$M_1 - m_1 \leq M_o - \varepsilon(M_o - m_o) - m_o - \varepsilon(-m_o + M_o)$$

ou

$$M_1 - m_1 \leq (1 - 2\varepsilon)(M_o - m_o). \quad (13)$$

Além disso, como $M_o \geq M_1$ e $m_o \leq m_1$, tem-se

$$M_1 - m_1 \leq M_o - m_1 \leq M_o - m_o. \quad (14)$$

Desse modo,

$$M_o - m_o \geq M_o - m_1. \quad (15)$$

Da desigualdade (14) resulta:

$$\frac{M_1 - m_1}{M_o - m_o} \leq 1. \quad (16)$$

Da desigualdade (13) tem-se $1 - 2\varepsilon \geq 0$ e como $\varepsilon > 0$ segue de (16) que

$$0 < \frac{M_1 - m_1}{M_o - m_o} \leq 1 - 2\varepsilon < 1.$$

Logo,

$$0 < 1 - 2\varepsilon < 1. \quad (17)$$

O Teorema 1 é uma extensão do lema 4 de Silva e Rota [7] no sentido dos resultados adicionais (10), (12) e (17) obtidos. Também a demonstração aqui apresentada para a desigualdade (13) possui diferenças significativas em relação à realizada pelos referidos autores. Por exemplo, a igualdade (6) não foi demonstrada claramente em [2] e [7] e, a desigualdade (11) foi provada usando uma metodologia diferente da adotada por [7]. Além disso, a desigualdade (17) foi realizada de forma direta em [7], quando no nosso entendimento, haveria de se exigir que $r \geq 2$.

Teorema 2

Se P é uma matriz de transição regular, então $P^n \rightarrow A$, sendo A uma matriz de probabilidade de ordem r , onde cada coluna é formada pelo mesmo vetor de probabilidade $a = (a_1, a_2, \dots, a_r)^T$.

Demonstração: Suponhamos primeiramente que $P_{ij} \neq 0$, para todo $i, j = 1, \dots, r$ e que $\varepsilon = \min (P_{i,j})$. Seja $e_j = (0, 0, \dots, 0, 1, 0, \dots, 0)$, com a unidade na j -ésima posição. Seja $M_n = \max_{j=1, \dots, r} e_j P^n$ e $m_n = \min_{j=1, \dots, r} e_j P^n$, $n \geq 1$. Isto é, M_n e m_n são respectivamente a máxima e a mínima componentes do vetor linha $e_j P^n = P^n(j, \cdot)$, onde $e_j P^n = (e_j P^{n-1})P$.

Dos resultados (3) e (4) do teorema anterior, tem-se a cadeia não crescente $M_1 \geq M_2 \geq \dots$ e não decrescente $m_1 \leq m_2 \leq \dots$ de modo que usando (13) no estágio n tem-se

$$M_n - m_n \leq (1 - 2\varepsilon)(M_{n-1} - m_{n-1})$$

ou

$$M_n - m_n \leq (1 - 2\varepsilon)^n (M_0 - m_0). \quad (18)$$

Seja $d_n = M_n - m_n$, $n \geq 1$. Como $0 \leq d_0 \leq 1$ tem-se

$$(1 - 2\varepsilon)^n d_0 \leq (1 - 2\varepsilon)^n.$$

Daí, e de (18) resulta:

$$0 \leq d_n \leq (1 - 2\varepsilon)^n d_0 \leq (1 - 2\varepsilon)^n.$$

$$d_n \rightarrow 0 \text{ quando } n \rightarrow \infty, \text{ logo } M_n \rightarrow m_n$$

Assim, a máxima componente do vetor linha, $e_j P^n$ é igual à mínima componente deste mesmo vetor linha, digamos que a_j seja esse valor em comum, logo a linha $e_j P^n = (a_j \dots a_j)$.

Repetindo sucessivas vezes o cálculo, $e_j P^n$, $j = i, \dots, r$, tem-se

$$e_1 P^n = (a_1, \dots, a_1), e_2 P^n = (a_2, \dots, a_2), \dots, e_r P^n = (a_r, \dots, a_r)$$

ou

$$P^n = A = \begin{bmatrix} a_1 & \dots & a_1 \\ a_2 & \dots & a_2 \\ \dots & & \dots \\ a_r & \dots & a_r \end{bmatrix}, \quad (19)$$

isto é, P^n tende a uma matriz A cujas colunas são todas iguais ao vetor $a = (a_1, a_2, \dots, a_r)^T$. Como a_j é o valor comum através da aproximação entre m_n e M_n , é claro que

$$m_n \leq a_j \leq M_n, \text{ para todo } n \geq 1.$$

Em particular,

$$0 < m_1 \leq a_j \leq M_1 < 1.$$

Então cada componente a_j é positiva. Como P^n é estocástica, $\sum_{j=1}^r a_j = 1$.

Caso P seja regular, existe $N \in \mathbb{N}$, tal que $P_{ij}^N > 0$, daí e usando a prova realizada para o caso do Teorema 1, em que $P_{ij} \neq 0$, tem-se uma sequência não crescente $0 \leq d_{kN} \leq (1 - 2\varepsilon')^k$ onde $d_{kN} \rightarrow 0$, $k \rightarrow \infty$ e $P^n \rightarrow A$, $n \rightarrow \infty$, onde cada coluna de A é um mesmo vetor de probabilidade.

O Teorema 2, é equivalente ao Teorema 5, itens (i, ii) de Silva e Rota [7]. Entretanto, a prova dos referidos autores envolveu o lema 4 e também o lema 3 não utilizado nesse estudo. Além disso, a metodologia adotada fez uso de limite de sequências de vetores, enquanto a descrita nesse artigo baseou-se tão somente nas desigualdades (3), (4) e em (13), todas decorrentes do Teorema 1. Nesse sentido, acreditamos que nossa abordagem ficou mais detalhada quando comparada com [2] e [7].

O próximo teorema mostrará que $P^n x_0$ se aproxima do vetor estacionário a independente do vetor de probabilidade inicial x_0 , e que a é único no sentido de satisfazer a equação, $P^n a = a$. Além disso, a matriz P comuta com A .

Teorema 3

Se P é uma matriz de transição regular e A e a são dados como no teorema 2, então:

- (a) Para qualquer vetor de probabilidade x_0 , $P^n x_0$ tende a a quando n tende ao infinito.
- (b) O vetor a é o único vetor de probabilidade tal que $Pa = a$.
- (c) $PA = AP$.

Demonstração: Como P é regular, pelo teorema 2, $P^n \rightarrow A$ quando $n \rightarrow \infty$.

$$P^n x_0 = Ax_0 = \left(a_1 \sum_{i=1}^r x_i, a_2 \sum_{i=1}^r x_i, \dots, a_r \sum_{i=1}^r x_i \right)^T,$$

onde, $x_0 = (x_1, x_2, \dots, x_r)^T$. Como x_0 é um vetor de probabilidade, $\sum_{i=1}^r x_i = 1$. Daí, $P^n x_0 = Ax_0 = a$.

Seja β um vetor de probabilidade tal que, $P\beta = \beta$, então usando o item (a), $P^n \beta = a$ quando n tende a infinito. Por outro lado, $P^2 \beta = P(P\beta) = \beta, \dots, P^n \beta = \beta$. Logo, $\beta = a$.

$P^{n+1} = P^n P = AP$, quando n tende ao infinito. Mas $PA = PP^n$, logo $AP = PA$.

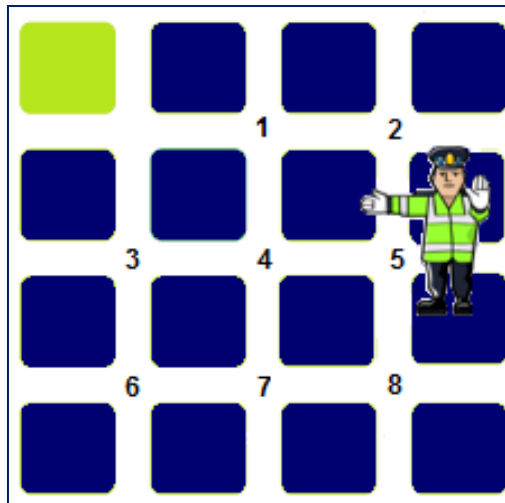
O Teorema 3 é equivalente ao Teorema 5, itens (iii e iv) de Silva e Rota [7], mas a prova que fizemos para o item (b) que corresponde ao item (iv) de [7] não fez uso do teorema da unicidade do limite.

2.2. Aplicação

A seguir veremos uma aplicação da cadeia de Markov baseada no problema proposto por [1] e aqui ampliada, utilizando três diferentes métodos para a obtenção do vetor estacionário.

Um guarda de trânsito é designado para controlar o tráfego nos oito cruzamentos indicados na Figura 1. Ele é instruído a permanecer em cada cruzamento por uma hora e, em seguida, ou permanecer no mesmo cruzamento ou seguir para um cruzamento adjacente.

Figura 1. Diagrama do cruzamento de ruas e posição inicial do guarda.



Fonte: Os autores (2017).

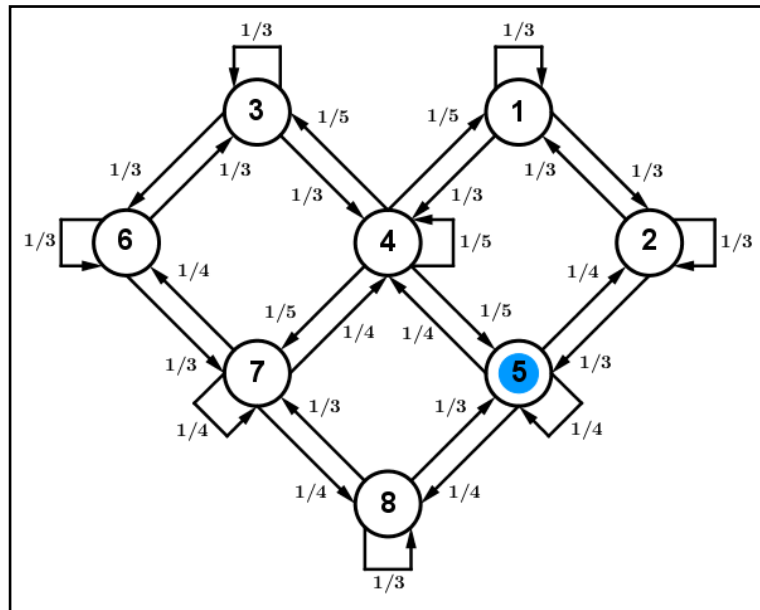
Para evitar que ele estabeleça um padrão, ele deve escolher o novo cruzamento de maneira aleatória, com qualquer escolha igualmente provável. Por exemplo, se ele está no cruzamento 5, seu próximo cruzamento pode ser 2, 4, 5 ou 8, cada um com igual probabilidade $1/4$. Cada dia ele recomeça no cruzamento em que parou no dia anterior.

A matriz de transição desta cadeia de Markov com o guarda iniciando seu trabalho na posição 5 é dada abaixo por:

$$P = \begin{matrix} & \begin{matrix} \text{Cruzamento Velho} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} \end{matrix} \\ \begin{matrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \\ \mathbf{8} \end{matrix} & \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{5} & 0 & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & \frac{1}{4} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{3} \end{bmatrix} & \begin{matrix} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \\ \mathbf{8} \end{matrix} \end{matrix} \quad \text{Cruzamento Novo}$$

Podemos representar as probabilidades de transição utilizando a representação por grafo, que para este fim é denominado como Diagrama de Transição de Estado. Neste os sentidos das flechas indicam a probabilidade de transição de um estado i para um estado j . Para a matriz de transição P dada acima o diagrama fica da seguinte forma conforme pode ser visto na Figura 2.

Figura 2. Diagrama de transição de estado da matriz P . Nota a posição 5 pintada de azul indicado a posição inicial.



Fonte: Os autores (2017).

Primeiro Método: Utilizando a equação matricial $x^{(n+1)} = Px^{(n)}$ decorrente do Teorema 3 item (a). Supondo que o guarda inicialmente começa no cruzamento 5, suas prováveis localizações, hora à hora, são dadas pelos vetores representados pelas colunas da Tabela 1.

Para todos os valores de n maiores do que 18, todos os vetores-estado são iguais a $x^{(18)}$ até três casas decimais. Assim, os vetores-estado convergem a um vetor fixo à medida que n cresce. Esse vetor fixo é o vetor estacionário

$$a = [0.108 \ 0.108 \ 0.106 \ 0.178 \ 0.144 \ 0.106 \ 0.142 \ 0.107]^T,$$

que corresponde ao vetor probabilidade.

Tabela 1. Vetores resultantes desde $x^{(0)}$, indicando a posição inicial 5, até o vetor o vetor estacionário $x^{(18)}$, indicando a convergência a partir dele.

Vetor	Iteração (n)									
	0	1	2	3	4	5	10	15	18	19
$x_1^{(n)}$	0	0.000	0.133	0.115	0.130	0.122	0.113	0.109	0.108	0.108
$x_2^{(n)}$	0	0.250	0.146	0.163	0.140	0.138	0.115	0.110	0.108	0.108
$x_3^{(n)}$	0	0.000	0.050	0.0391	0.0668	0.0732	0.100	0.106	0.106	0.106
$x_4^{(n)}$	0	0.250	0.112	0.186	0.162	0.178	0.178	0.179	0.178	0.178
$x_5^{(n)}$	1	0.250	0.279	0.190	0.190	0.168	0.149	0.145	0.144	0.144
$x_6^{(n)}$	0	0.000	0.000	0.0499	0.0556	0.0733	0.0988	0.105	0.106	0.106
$x_7^{(n)}$	0	0.000	0.133	0.104	0.130	0.125	0.138	0.142	0.142	0.142
$x_8^{(n)}$	0	0.250	0.146	0.152	0.124	0.121	0.108	0.107	0.107	0.107

Segundo Método: Usando a equação $P^{(n)} \rightarrow A$, decorrente do Teorema 2, obtivemos as matrizes abaixo. Para todos os valores de n maiores do que 19, todos os vetores-estado são iguais a $P^{(19)}$ até três casas decimais. Assim, os vetores-estado de cada coluna da matriz convergem a um vetor fixo à medida que n cresce, resultando numa matriz estacionária w . Esse vetor fixo é o vetor estacionário

$$a = [0.107 \ 0.108 \ 0.107 \ 0.179 \ 0.143 \ 0.107 \ 0.143 \ 0.107]^T,$$

que corresponde ao vetor probabilidade. Note que neste caso não foi necessário uma posição inicial x_0 .

$$P^2 = \begin{bmatrix} 0.289 & 0.222 & 0.066 & 0.107 & 0.133 & 0.000 & 0.050 & 0.000 \\ 0.222 & 0.306 & 0.000 & 0.117 & 0.146 & 0.000 & 0.000 & 0.083 \\ 0.066 & 0.000 & 0.289 & 0.107 & 0.050 & 0.222 & 0.133 & 0.000 \\ 0.178 & 0.194 & 0.178 & 0.273 & 0.112 & 0.194 & 0.112 & 0.167 \\ 0.178 & 0.194 & 0.066 & 0.090 & 0.279 & 0.000 & 0.133 & 0.194 \\ 0.000 & 0.000 & 0.222 & 0.117 & 0.000 & 0.306 & 0.146 & 0.083 \\ 0.066 & 0.000 & 0.178 & 0.090 & 0.133 & 0.194 & 0.279 & 0.194 \\ 0.000 & 0.083 & 0.000 & 0.100 & 0.146 & 0.083 & 0.146 & 0.278 \end{bmatrix}$$

...

$$P^{19} = \begin{bmatrix} 0.108 & 0.108 & 0.106 & 0.107 & 0.108 & 0.106 & 0.107 & 0.107 \\ 0.109 & 0.109 & 0.107 & 0.108 & 0.108 & 0.107 & 0.108 & 0.108 \\ 0.106 & 0.106 & 0.108 & 0.107 & 0.107 & 0.108 & 0.108 & 0.107 \\ 0.180 & 0.179 & 0.179 & 0.180 & 0.178 & 0.178 & 0.178 & 0.178 \\ 0.143 & 0.144 & 0.142 & 0.142 & 0.144 & 0.142 & 0.142 & 0.143 \\ 0.106 & 0.106 & 0.108 & 0.107 & 0.107 & 0.108 & 0.108 & 0.107 \\ 0.142 & 0.142 & 0.144 & 0.143 & 0.143 & 0.144 & 0.144 & 0.143 \\ 0.107 & 0.107 & 0.107 & 0.107 & 0.108 & 0.107 & 0.108 & 0.108 \end{bmatrix}$$

$$P^{20} = \begin{bmatrix} 0.108 & 0.108 & 0.106 & 0.107 & 0.108 & 0.106 & 0.107 & 0.107 \\ 0.109 & 0.109 & 0.107 & 0.108 & 0.108 & 0.107 & 0.108 & 0.108 \\ 0.106 & 0.106 & 0.108 & 0.107 & 0.107 & 0.108 & 0.108 & 0.107 \\ 0.180 & 0.179 & 0.179 & 0.180 & 0.178 & 0.178 & 0.178 & 0.178 \\ 0.143 & 0.144 & 0.142 & 0.142 & 0.144 & 0.142 & 0.142 & 0.143 \\ 0.106 & 0.106 & 0.108 & 0.107 & 0.107 & 0.108 & 0.108 & 0.107 \\ 0.142 & 0.142 & 0.144 & 0.143 & 0.143 & 0.144 & 0.144 & 0.143 \\ 0.107 & 0.107 & 0.107 & 0.107 & 0.108 & 0.107 & 0.108 & 0.108 \end{bmatrix}$$

Terceiro Método: Usando o item (b) do Teorema 3, tem-se a equação $Pw = w$, que é equivalente a resolver o sistema linear $(P - I)w = 0$.

$$P - I = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & -\frac{2}{3} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & -\frac{2}{3} & \frac{1}{5} & 0 & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & -\frac{4}{5} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & -\frac{3}{4} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & 0 & -\frac{2}{3} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 & \frac{1}{3} & -\frac{3}{4} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{4} & -\frac{2}{3} \end{bmatrix}$$

Resolvendo $(P - I)w = 0$, chegamos à matriz ampliada usando o método da eliminação de Gauss completo [1] na forma

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -\frac{5}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -\frac{4}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\frac{4}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

onde $x_1 = x_8$, $x_2 = x_8$, $x_3 = x_8$, $x_4 = \frac{5}{3}x_8$, $x_5 = \frac{4}{3}x_8$, $x_6 = x_8$, $x_7 = \frac{4}{3}x_8$ e x_8 é a variável livre.

Escolhemos $x_8 = 3$ para obtermos uma base simples, sem frações. Assim,

$$w = [3 \ 3 \ 3 \ 5 \ 4 \ 3 \ 4 \ 3]^T.$$

Multiplicando o vetor estacionário w pelo escalar $1/28$, onde o denominador é obtido somando as coordenadas de w , obtemos o vetor probabilidade estacionário

$$a = [0.107 \ 0.107 \ 0.107 \ 0.179 \ 0.143 \ 0.107 \ 0.143 \ 0.107]^T.$$

O denominador do escalar, $1/28$ pode ser obtido somando as coordenadas de w .

3. Resultados e Discussões

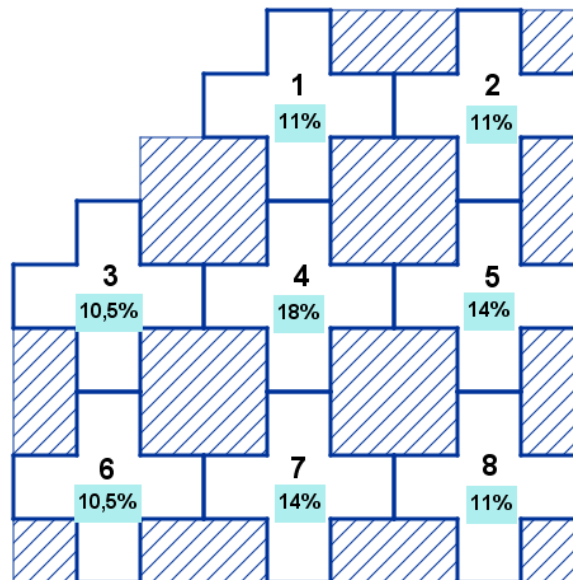
A Tabela 2 mostra o vetor estacionário calculado pelos três procedimentos descritos anteriormente. As pequenas variações podem ser creditadas aos erros de arredondamento e truncamento realizados pela máquina. No terceiro método não é possível observar os estados probabilísticos da transição da cadeia de Markov. Além do mais, devido à facilidade de programação dos métodos 1 e 2 eles devem ser preferíveis devido a sua praticidade em relação ao método 3 de característica mais algébrica.

Tabela 2. Vetor estacionário resultante de cada método. Em 1 e 2, n representa o número da iteração na qual ocorreu a convergência.

Métodos	Vetor probabilidade (a)
1 $x^{(n+1)} = Px^{(n)}$	$[0.108 \ 0.108 \ 0.106 \ 0.178 \ 0.144 \ 0.106 \ 0.142 \ 0.107]^T$
2 $Px^{(n)} = w$	$[0.107 \ 0.108 \ 0.107 \ 0.179 \ 0.143 \ 0.107 \ 0.143 \ 0.107]^T$
3 $Pw = w$	$[0.107 \ 0.107 \ 0.107 \ 0.179 \ 0.143 \ 0.107 \ 0.143 \ 0.107]^T$

Retomando a análise do problema proposto, as coordenadas desses vetores expõem a probabilidade da posição do guarda após passadas 19 horas, tendo como início a posição 5. Utilizando os resultados do primeiro método, e considerando arredondamentos da máquina, temos a seguinte representação de probabilidade das posições em porcentagem:

Figura 3. Representação dos cruzamentos das ruas com as respectivas probabilidades em porcentagem da posição do guarda de pois de 19 horas.



Fonte: Os autores (2017).

Consideramos também uma análise geral utilizando o primeiro método, que nos permite melhor observar os estados probabilísticos da transição da convergência da cadeia de Markov, considerando o guarda em cada um dos outros sete cruzamentos ampliando desse modo, o projeto da aplicação proposto em [1]. Os resultados mostraram pequenas variações no número de interações para alcançar a convergência, porém os vetores estacionários resultante de cada posição inicial mostraram atingir probabilidades concordantes entre eles como mostra a Tabela 3, o que era de se esperar, visto que todos partiram da mesma matriz de transição P . Novamente, as pequenas variações nas coordenadas do vetor estacionário pelos diferentes métodos podem ser creditadas aos erros de arredondamento e truncamento realizados pela máquina.

Tabela 3. Vetores estacionários resultantes utilizando o Método 1 de cada posição inicial do guarda e o número de iterações (horas) necessárias para a convergência.

Posição inicial	Número de iterações	Vetor probabilidade
1	20	[0.109 0.109 0.107 0.180 0.144 0.107 0.144 0.108] ^T
2	21	[0.109 0.109 0.107 0.180 0.144 0.107 0.144 0.108] ^T
3	19	[0.106 0.106 0.108 0.178 0.141 0.108 0.142 0.106] ^T
4	9	[0.107 0.107 0.107 0.178 0.142 0.107 0.142 0.106] ^T
5	18	[0.108 0.108 0.106 0.178 0.144 0.106 0.142 0.107] ^T
6	21	[0.107 0.107 0.109 0.180 0.144 0.109 0.144 0.108] ^T
7	18	[0.106 0.106 0.108 0.178 0.142 0.108 0.144 0.108] ^T
8	12	[0.106 0.106 0.106 0.178 0.142 0.106 0.142 0.107] ^T

Nosso objetivo foi, portanto, seguir exclusivamente os passos de [2], adaptando-os para o caso $T^n v$ de modo que, os principais resultados fossem demonstrados de forma mais compreensível do que a versão sucinta encontrada em [2] e desse modo, pudessem ser acessíveis a um público mais geral. A descrição de três métodos para obter o vetor estacionário de uma cadeia de Markov foi aqui detalhada e utilizada em uma aplicação prática.

Referências

- [1] ANTON, H.; RORRES, C. Álgebra Linear com Aplicações, trad. Claus Ivo Doering. 8. ed. Porto Alegre: Bookman, 2001.
- [2] KEMENY, J. G.; SNELL, J. L. Finite Markov chains. New York: Springer-Verlag, 1960.
- [3] LAY, D. C. Álgebra Linear e Suas Aplicações. 2. ed. Editora LTC, 1997.
- [4] MANOEL, M. R. Cadeias de Markov: uma abordagem voltada para o ensino médio. 2016. 69 f. Dissertação (Mestrado Profissional) - Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas. Campinas, 2016.
- [5] PRASAD, N. R.; ENDER, R. C.; REILY S. T. e NESGOS G. IEEE Conference Publications: Allocation of resources on a minimized cost basis, 1974 IEEE Conference on Decision and including the 13th Symposium on Adaptive Processes.
- [6] RUFFINO, P. R. C. Uma iniciação aos sistemas dinâmicos estocásticos. 2. ed. Rio de Janeiro: IMPA, 2009.
- [7] SILVA, F. B.; ROTA, I. S. Convergência de matrizes estocásticas regulares. C.Q.D. – Revista Eletrônica Paulista de Matemática, Bauru, v. 8, p. 4-14, dez. 2016. Edição Iniciação Científica.