Revista Eletrônica
Paulista de Matemática

ISSN 2316-9664

Volume 11, dez. 2017

Edição Iniciação Científica

**Amanda Silvieri Leite de
Oliveira**

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”, Bauru/SP.
amandasilvieri@hotmail.com

Thaís Saes Giuliani Ribeiro

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”, Bauru/SP
thais_saes@hotmail.com

Fabiano Borges da Silva

UNESP - Universidade Estadual
Paulista “Júlio de Mesquita
Filho”, Bauru/SP
fabiano@fc.unesp.br

Cadeia de Markov: modelo probabilístico e convergência das distribuições de probabilidade

Markov chain: probabilistic model and convergence of
probability distributions

Resumo

Neste artigo, mostramos como construir um processo estocástico de Markov e seu espaço de probabilidade a partir das probabilidades de transição e da distribuição inicial. Além disso, mostramos a convergência das distribuições de probabilidade para uma cadeia com dois estados e probabilidades de transição positivas, usando técnicas de resolução de recorrências lineares não-homogêneas para sequências.

Palavras-chave: Cadeia de Markov, processos estocásticos, espaço de probabilidade, convergência, probabilidade condicional.

Abstract

In this article, we show how to construct a stochastic Markov process and its probability space from the transition probabilities and the initial distribution. In addition, we show the convergence of probability distributions to a chain with two states and positive transition probabilities, using non-homogeneous linear recurrence resolution techniques for sequences.

Keywords: Markov chain, stochastic processes, probability space, convergence, conditional probability.

1 Introdução

Processos de Markov descrevem a evolução de sistemas dinâmicos aleatórios sem memória. Mais precisamente, considere um espaço de estados com um número finito (ou enumerável) de elementos $\mathcal{E} = \{e_1, \dots, e_n\}$. Um processo estocástico discreto $(X_n)_{n \in \mathbb{N}}$ é uma cadeia (ou processo) de Markov se a probabilidade condicional satisfizer

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

para todo $n \geq 1$ e para toda sequência $x_0, x_1, \dots, x_{n+1}$ de elementos do espaço de estados $\mathcal{E}$. Essa condição (1) significa, em linguagem natural, pensando que n indica o tempo, que o futuro do processo, uma vez conhecido o estado presente, é independente do passado.

As probabilidades condicionais

$$\mathbb{P}(X_{n+1} = e_i | X_n = e_j),$$

são chamadas probabilidades de transição. E se para cada i, j

$$\mathbb{P}(X_{n+1} = e_i | X_n = e_j) = \mathbb{P}(X_1 = e_i | X_0 = e_j),$$

para todo natural n , a cadeia de Markov é dita estacionária e as probabilidades de transição, que não mudam ao longo do tempo, são denotadas por p_{ij} .

Um processo de Markov está completamente definido a partir do momento em que se especifica as probabilidades de transição e a distribuição inicial de probabilidades dos estados, como pode ser visto em Brezezniak e Zastawniak (1999, p.85). Ao processo associa-se uma matriz de probabilidades de transição T , onde as entradas da matriz são dadas pelas probabilidades de transição p_{ij} , ou seja,

$$T = [p_{ij}]_{r \times r}.$$

As entradas da matriz T^n correspondem à probabilidade de, saindo do estado e_j , chegar-se ao estado e_i depois de n passos. Desta maneira, dada uma distribuição inicial, representada matricialmente por

$$v_0 = [v_1 \dots v_r]^t,$$

a distribuição do processo no tempo $n \geq 1$ é dada por

$$v_n = T^n v_0.$$

Neste artigo estamos interessados basicamente em dois tópicos:

1. Estudar convergência das distribuições de probabilidade v_n , que evoluem em tempos discretos $n \in \mathbb{N}$, para uma cadeia finita com dois estados (Teorema 4). Para isto, usaremos técnicas de resolução para recorrências lineares não-homogêneas.
2. Mostrar como construir um espaço de probabilidade e um processo estocástico X_n de Markov, isto é, que verifica (1), a partir de um modelo em que é apenas dado as probabilidades de transição e uma distribuição inicial (Teorema 5).

2 Espaço de probabilidade e processos estocásticos

Nesta seção, daremos alguns resultados que serão necessários, na construção do processo estocástico a partir das probabilidades de transição.

O conjunto de todos os resultados possíveis de um experimento é o espaço amostral Ω , e um subconjunto $A \subset \Omega$ deste espaço é chamado de evento aleatório. Os eventos Ω , $\emptyset$ são chamados de evento certo e evento impossível, respectivamente. Uma maneira de definir uma probabilidade $\mathbb{P}$ em um determinado evento A , é a maneira “frequentista” ou “estatística”. Mais precisamente,

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \frac{1}{n} \times \{\text{número de ocorrências de } A \text{ em } n \text{ “ensaios” independentes}\}.$$

Esta maneira, a qual não usaremos neste trabalho, não é única (ver por exemplo James (1996)). Neste artigo, usaremos a definição que se deve a Kolmogorov. Para isto, iremos admitir que a classe $\mathcal{F}$ dos eventos aleatórios possui as seguintes propriedades:

1. $\Omega \in \mathcal{F}$;
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
3. Se $A \in \mathcal{F}$ e $B \in \mathcal{F}$, então $A \cup B \in \mathcal{F}$.

Esta classe $\mathcal{F}$ de subconjuntos de Ω é chamada de álgebra. E quando a terceira propriedade acima vale para uniões enumeráveis, ou seja, se $A_n \in \mathcal{A}$ para $n = 1, 2, \dots$, temos que

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F},$$

então, neste caso, a classe $\mathcal{F}$ é chamada de σ -álgebra, e o par ordenado $(\Omega, \mathcal{F})$ de espaço mensurável. Neste espaço definimos uma medida de probabilidade como sendo uma função $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$, tal que

1. $\mathbb{P}(\emptyset) = 0$;
2. Se $(A_n)_{n \geq 1}$ é uma sequência de subconjuntos disjuntos, $A_n \in \mathcal{F}$, então

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbb{P}(A_k).$$

Sendo $\mathcal{F}$ uma σ -álgebra, a tripla $(\Omega, \mathcal{F}, \mathbb{P})$ é chamada de espaço de probabilidade. Se $B \in \mathcal{F}$ e $\mathbb{P}(B) > 0$, a probabilidade condicional do evento A dado B é definida por

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad A \in \mathcal{F}.$$

Proposição 1 *Se a sequência (finita ou enumerável) de eventos aleatórios $A_1, A_2, \dots$ formam uma partição de Ω , então para todo $B \in \mathcal{F}$ temos que*

$$\begin{aligned} \mathbb{P}(B) &= \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(A_i) \mathbb{P}(B|A_i) \end{aligned}$$

A demonstração do resultado acima, pode ser visto por exemplo, em James (1996, p.17).

2.0.1 Processo estocástico com tempo discreto

Vamos assumir primeiramente que Ω é um espaço amostral finito e $\mathcal{F}$ é uma σ -álgebra para este espaço amostral. Uma função $X : \Omega \rightarrow \mathbb{R}$ é chamada $\mathcal{F}$ -mensurável ou variável aleatória em $(\Omega, \mathcal{F})$ se os conjuntos

$$\{X = x_i\} = \{\omega \in \Omega : X(\omega) = x_i\}, \quad i = 1, 2, \dots, k,$$

pertencem a $\mathcal{F}$, onde $\{x_1, x_2, \dots, x_k\}$ são elementos da imagem da função X . Isto significa que, se temos a informação descrita por $\mathcal{F}$, isto é, sabemos que o evento ocorreu, então sabemos qual o valor de X ocorreu. É fácil ver, por exemplo, que se $\mathcal{F} = 2^\Omega$ (conjunto das partes de Ω), então qualquer função em Ω é uma variável aleatória. Um *processo estocástico* é uma sequência de variáveis aleatórias X_n , ou seja, para cada tempo $n \in \mathbb{N}$, X_n é uma variável aleatória em $(\Omega, \mathcal{F})$.

Para ilustrar tal conceito, considere uma sequência de experimentos: o lançamento de uma moeda não viciada em dois instantes, $t = 1, 2$. Denotemos por α e β , os resultados obtidos para cara e coroa, respectivamente. Neste caso temos que nosso espaço amostral é dado por

$$\Omega = \{\omega_1 = (\alpha, \beta), \omega_2 = (\alpha, \alpha), \omega_3 = (\beta, \alpha), \omega_4 = (\beta, \beta)\}.$$

Seja A o evento onde se obtém no primeiro experimento cara, isto é, $A = \{\omega_1, \omega_2\}$. Além disso, considere as σ -álgebras $\mathcal{F}_1 = \{\emptyset, \Omega, A, A^C\}$, $\mathcal{F}_2 = 2^\Omega$, e as seguintes funções:

(a) $X : \Omega \rightarrow \mathbb{R}$ dada por

$$X(\omega_1) = X(\omega_2) = 15; \quad X(\omega_3) = X(\omega_4) = 45.$$

(b) $Y : \Omega \rightarrow \mathbb{R}$ dada por

$$Y(\omega_1) = 17, \quad Y(\omega_2) = Y(\omega_3) = 36, \quad Y(\omega_4) = 42.$$

Temos então que X é variável aleatória em $(\Omega, \mathcal{F}_1)$. De fato, temos que

$$\{X = 15\} = A \in \mathcal{F}_1;$$

$$\{X = 45\} = A^C \in \mathcal{F}_1.$$

Enquanto que a função Y não é $\mathcal{F}_1$ -mensurável, uma vez que, por exemplo,

$$\{Y = 36\} = \{\omega_2, \omega_3\} \notin \mathcal{F}_1.$$

Por outro lado, Y é $\mathcal{F}_2$ -mensurável já que todo subconjunto de Ω pertence a $\mathcal{F}_2 = 2^\Omega$, por definição.

Uma *filtração* $\mathbb{F}$ é uma coleção de σ -álgebras,

$$\mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n, \dots, \mathcal{F}_N\}, \quad \mathcal{F}_n \subset \mathcal{F}_{n+1},$$

a qual é usada para modelar um fluxo de informações do processo. Em geral, se toma $\mathcal{F}_0 = \{\emptyset, \Omega\}$, ou seja, no instante $n = 0$ não temos nenhuma informação. Conforme o tempo passa,

um observador consegue saber mais detalhes sobre os acontecimentos dos experimentos, ou seja, partições “mais finas” de Ω . Na ilustração acima, temos que

$$\mathbb{F} = \{\mathcal{F}_0, \mathcal{F}_A, 2^\Omega\}$$

é uma filtração.

Dizemos que um processo estocástico X_n é adaptado a filtração $\mathbb{F}$, se cada $X_n : \Omega \rightarrow \mathbb{R}$ é variável aleatória em $(\Omega, \mathcal{F}_n)$, ou equivalentemente, X_n é $\mathcal{F}_n$ -mensurável. Por exemplo, a sequência de variáveis aleatórias (X, Y) , formada com as mesmas funções mencionadas anteriormente, itens (a) e (b), é um processo estocástico adaptado a filtração $\mathbb{F} = \{\mathcal{F}_1, \mathcal{F}_2\}$.

Seja $(\Omega, 2^\Omega)$ um espaço amostral com a álgebra de todos os eventos, e X uma variável aleatória com valores $x_i, i = 1, 2, \dots, k$. Considere

$$A_i = \{\omega : X(\omega) = x_i\} \subseteq \Omega. \quad (2)$$

A álgebra gerada pela partição $\{A_0, A_1, \dots, A_n\}$, isto é, via uniões e intersecções destes conjuntos, é chamada álgebra gerada por X . Ela é a menor álgebra que contém todos os conjuntos da forma $A_i = \{X = x_i\}$ e é denotada por $\mathcal{F}_X$ ou $\sigma(X)$. A álgebra gerada por X representa a informação que podemos extrair observando X .

Ao leitor interessado em mais detalhes sobre processos estocásticos (discretos ou contínuos), sugerimos, entre outros, Ruffino (2009).

2.0.2 Probabilidade no espaço das sequências

Para uma sequência infinita (enumerável) de variáveis aleatórias, podemos construir o espaço de probabilidade da seguinte forma. Seja Ω o conjunto de todas as sequências de elementos do espaço de estados $\mathcal{E}$. Um elemento $\omega \in \Omega$ pode então ser escrito da forma

$$\omega = (e_0, e_1, e_2, \dots),$$

onde cada $e_i \in \mathcal{E}$. A função $X_n : \Omega \rightarrow \mathcal{E}$, dada por

$$X_n(e_0, e_1, e_2, \dots) = e_n,$$

é chamada de *função saída* ou *avaliação* da trajetória ω .

Fixado n , seja $\mathcal{F}_n$ a família de todas as uniões de Ω da forma

$$\{\omega : X_0(\omega) \in \mathcal{E}_0, X_1(\omega) \in \mathcal{E}_1, \dots, X_n(\omega) \in \mathcal{E}_n\},$$

onde $\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_n$ são subconjuntos do espaço de estados $\mathcal{E}$. Neste caso, não é difícil verificar que cada $\mathcal{F}_n$ é uma σ -álgebra. Além disso, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ forma uma filtração natural na qual o processo X_n é adaptado. Considere agora $\mathcal{F}$ a família de conjuntos definida por

$$\mathcal{F} = \bigcup_{n=0}^{\infty} \mathcal{F}_n.$$

Cada elemento em $\mathcal{F}$ é um conjunto de trajetórias para as quais um número finito de entradas da sequência são restritas a pertencer a certos subconjuntos de $\mathcal{E}$, e as demais infinitas entradas são irrestritas.

Um conjunto de $\mathcal{F}$ é chamado de *cilindro*. Apesar de $\mathcal{F}$ ser uma álgebra, não é uma σ -álgebra. Porém, existe a menor σ -álgebra $\mathcal{G}$, tal que $\mathcal{F} \subset \mathcal{G}$. Associado a $\mathcal{G}$ existe uma única medida de probabilidade μ tal que

$$\mu(C_i^n) = \mu\{\omega : X_0(\omega) = e_0, X_1(\omega) = e_1, \dots, X_n(\omega) = e_n\}$$

é dado pelo produto das probabilidades condicionais entre os estados $e_0, e_1, \dots, e_n$. Cada conjunto C_i^n é chamado de *cilindro básico* de $\mathcal{F}$. Para maiores detalhes ver Kemeny, Snell e Knapp (1976, p.43).

2.1 Cadeia de Markov e matriz de transição

Uma cadeia de Markov, como mencionamos na introdução deste artigo, é um processo estocástico $(X_n)_{n \in \mathbb{N}}$, associado a um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ que satisfaz a equação (1). Daremos a seguir algumas definições e propriedades básicas das probabilidades de transição e suas matrizes associadas. Para maiores detalhes ver Allen (2003), Brezezniak e Zastawniak (1999) e Kemeny, Snell e Knapp (1976).

Definição 2 A probabilidade de transição em um passo, denotada por $p_{ij}(k)$, é definida como a seguinte probabilidade condicional:

$$p_{ij}(k) = \mathbb{P}(X_{k+1} = i | X_k = j).$$

Isto é, a probabilidade de estar no estado i no tempo $k+1$, dado que estava no estado j no momento anterior k , para $i, j = 1, 2, \dots$

Se a probabilidade de transição $p_{ij}(k)$ numa cadeia não depende do tempo k , dizemos que ela é *homogênea*. Neste caso, usaremos a notação p_{ij} . Ao longo deste artigo somente iremos trabalhar com cadeias homogêneas.

Para uma cadeia de Markov com um número finito de estados, $\mathcal{E} = \{1, 2, \dots, m\}$, associa-se uma *matriz de transição* T , que é dada pelas probabilidades de transição, isto é,

$$T = (p_{ij}).$$

Definição 3 A probabilidade de transição em n -passos ($n \geq 0$), denotada por $p_{ij}^{(n)}$, é a probabilidade de transferência do estado j para o estado i em n etapas de tempo discreto, isto é,

$$p_{ij}^{(n)} = \mathbb{P}\{X_n = i | X_0 = j\}.$$

Novamente, para uma quantidade de estados finitos, podemos associar uma matriz $T^{(n)}$, onde a ij -ésima posição é dada por $p_{ij}^{(n)}$. Note que $T^{(0)}$ é a matriz identidade, uma vez que $p_{ii}^{(0)} = 1$ e $p_{ij}^{(0)} = 0$ quando $i \neq j$.

Existe uma relação entre as probabilidades de transição em n -passos, s -passos e $(n-s)$ -passos. Essas relações são conhecidas como as *equações de Chapman-Kolmogorov*:

$$p_{ij}^{(n)} = \sum_{k=1}^{\infty} p_{ik}^{(n-s)} p_{kj}^{(s)}, \quad 0 < s < n.$$

Em termos matriciais, essas equações podem ser escritas da forma

$$T^{(n)} = T^{(n-s)}T^{(s)}.$$

Como $T^{(1)} = T$, segue-se então que

$$\begin{aligned} T^{(2)} &= T^{(2-1)}T^{(1)} \\ &= T^{(1)}T^{(1)} \\ &= TT \\ &= T^2. \end{aligned}$$

E fazendo este processo sucessivamente tem-se que $T^{(n)} = T^n$, para todo $n \geq 0$. Portanto, uma maneira fácil de se obter as probabilidades $p_{ij}^{(n)}$ é por meio da matriz T^n .

Além disso, um outro aspecto interessante em conhecer T^n , é que o vetor v_n de distribuição de probabilidades do processo, no tempo n , é igual ao produto matricial $T^n v_0$, com a distribuição inicial v_0 escrita de forma transposta.

3 Modelo probabilístico e convergência da sequência de probabilidades

Nesta seção, inicialmente iremos estudar convergência das distribuições de probabilidade para uma cadeia finita $\mathcal{E} = \{1, 2\}$, cujas probabilidades de transição são todas positivas. Para tal estudo usaremos técnicas de resolução para recorrências lineares não-homogêneas como aparece em Morgado e Carvalho (2013, p.73).

A seguir, apresentamos um modelo que descreve a dinâmica de uma partícula de elétron que salta entre dois átomos. Neste modelo é dado as probabilidades de transição e a distribuição inicial da partícula. Apesar de simples, ele é suficiente e interessante para os propósitos deste artigo.

Exemplo 1 *Suponha que uma determinada partícula de elétron, salta em tempos discretos $n = 0, 1, 2, \dots$, entre dois átomos, que representaremos por 1 e 2, com as seguintes condições:*

- (a) *Se a partícula está no átomo 1 em um período de tempo n , então com probabilidade p , ela salta para o átomo 2, onde $0 < p < 1$;*
- (b) *Se a partícula está no átomo 2 no tempo n , então ela salta com probabilidade q para o estado 1, onde $0 < q < 1$;*
- (c) *A partícula se encontra no átomo 1 no instante inicial ($n = 0$).*

Neste modelo acima, no contexto das cadeias de Markov, podemos interpretar os átomos como estados, portanto, $\mathcal{E} = \{1, 2\}$, e a matriz de transição é dada por

$$T = \begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix}.$$

Na próxima seção construiremos um processo estocástico X_n com estas probabilidades de transição e provaremos que verifica (1).

Ainda com relação ao modelo acima, seja x_n a probabilidade da partícula estar no estado **1** no tempo n . Neste caso, temos que $(1 - x_n)$ é a probabilidade da partícula estar no estado **2** no tempo n . Ou seja, a distribuição da partícula no tempo n , em notação vetorial, é dada por

$$v_n = (x_n, 1 - x_n),$$

sendo $v_0 = (1, 0)$ a distribuição inicial, conforme escolhida no item (c) do exemplo acima.

Teorema 4 A sequência x_n converge para $\frac{q}{p+q}$ quando n tende ao infinito.

Demonstração. Denotemos por A_n o evento em que a partícula está no estado **1** no tempo n e seja $B_n = \Omega - A_n$, isto é, o evento em que a partícula está no estado **2** no tempo n . Sendo assim, nas condições do modelo apresentado no Exemplo 1, temos

$$\mathbb{P}(B_{n+1}|A_n) = p; \quad \mathbb{P}(A_{n+1}|B_n) = q; \quad \mathbb{P}(A_0) = 1.$$

Além disso, como A_n e B_n fazem uma partição em Ω segue da Proposição 1 que

$$\begin{aligned} x_{n+1} &= \mathbb{P}(A_{n+1}) \\ &= \mathbb{P}(A_{n+1}|A_n)\mathbb{P}(A_n) + \mathbb{P}(A_{n+1}|B_n)\mathbb{P}(B_n) \\ &= (1 - p)x_n + q(1 - x_n) \\ &= q + (1 - p - q)x_n. \end{aligned} \tag{3}$$

Note que a igualdade (3) é uma recorrência linear não-homogênea de primeira ordem, ou seja, é uma recorrência do tipo $x_{n+1} = g(n)x_n + h(n)$, sendo as funções $h(n) = 1 - p - q$ e $f(n) = q$.

Para resolver esta recorrência, iremos transformá-la em uma outra não homogênea da forma $x_{n+1} = x_n + f(n)$, que é fácil de resolvê-la. Com efeito, temos

$$\begin{aligned} x_1 &= x_0 + f(0) \\ x_2 &= x_1 + f(1) \\ &\vdots \\ x_n &= x_{n-1} + f(n-1) \end{aligned}$$

Somando ambos os membros, obtemos $x_n = x_0 + \sum_{k=0}^{n-1} f(k)$.

Para isto, considere a_n uma solução não nula da recorrência $x_{n+1} = g(n)x_n$. A substituição $x_n = a_n y_n$, transforma

$$x_{n+1} = g(n)x_n + h(n)$$

em

$$a_{n+1}y_{n+1} = g(n)a_n y_n + h(n).$$

Mas, $a_{n+1} = g(n)a_n$, pois a_n é solução de $x_{n+1} = g(n)x_n$. Portanto, a equação se transforma em

$$g(n)a_n y_{n+1} = g(n)a_n y_n + h(n),$$

ou seja,

$$y_{n+1} = y_n + h(n)[g(n)a_n]^{-1}.$$

Agora, resolvendo a equação $y_{n+1} = y_n + h(n)[g(n)a_n]^{-1}$, que está na forma $y_{n+1} = y_n + f(n)$, como mencionado anteriormente, basta depois tomar $x_n = a_n y_n$.

Voltando à recorrência dada pela equação (3), vamos resolvê-la utilizando os passos mencionados no parágrafo anterior. Inicialmente, devemos encontrar uma solução não nula da recorrência

$$x_{n+1} = (1 - p - q)x_n. \quad (4)$$

Temos então que

$$x_1 = (1 - p - q)x_0$$

$$x_2 = (1 - p - q)x_1$$

$$\vdots$$

$$x_n = (1 - p - q)x_{n-1}.$$

Multiplicando todos os termos de cada lado das igualdades, resulta em

$$x_n = (1 - p - q)^n x_0.$$

Logo, tomando a condição inicial $x_0 = 1$ temos que $a_n = (1 - p - q)^n$ é uma solução não nula da recorrência (4). Fazemos a substituição de $x_n = (1 - p - q)^n y_n$ em (3). Obtemos então que

$$(1 - p - q)^{n+1} y_{n+1} = q + (1 - p - q)(1 - p - q)^n y_n,$$

e, portanto,

$$y_{n+1} = \frac{q}{(1 - p - q)^{n+1}} + y_n.$$

Como $x_0 = (1 - p - q)^0 y_0$ segue que $y_0 = 1$. Temos então que

$$y_1 = \frac{q}{(1 - p - q)^1} + 1$$

$$y_2 = \frac{q}{(1 - p - q)^2} + y_1$$

$$\vdots$$

$$y_n = \frac{q}{(1 - p - q)^n} + y_{n-1}$$

Somando os termos de cada lado das igualdades, resulta em

$$y_n = 1 + \frac{q}{(1 - p - q)} + \frac{q}{(1 - p - q)^2} + \frac{q}{(1 - p - q)^3} + \dots + \frac{q}{(1 - p - q)^n}.$$

Note que os termos do segundo membro da igualdade acima, após a condição inicial $y_0 = 1$, é a soma dos n primeiros termos de uma progressão geométrica de razão $\frac{1}{1-p-q}$. Logo

$$\begin{aligned} y_n &= 1 + \frac{q}{1-p-q} \left[\frac{\left(\frac{1}{1-p-q}\right)^n - 1}{\frac{1}{1-p-q} - 1} \right] \\ &= 1 + \frac{q}{1-p-q} \left[\frac{1 - (1-p-q)^n}{\frac{p+q}{1-p-q}} \right] \\ &= 1 + \frac{q(1 - (1-p-q)^n)}{(p+q)(1-p-q)^n} \end{aligned}$$

Como $x_n = (1-p-q)^n y_n$, segue que

$$\begin{aligned} x_n &= (1-p-q)^n + \frac{(1-p-q)^n q(1 - (1-p-q)^n)}{(p+q)(1-p-q)^n} \\ &= (1-p-q)^n + \frac{q(1 - (1-p-q)^n)}{p+q} \\ &= (1-p-q)^n + \frac{q - q(1-p-q)^n}{p+q} \\ &= \frac{q}{p+q} + \frac{(p+q)(1-p-q)^n - q(1-p-q)^n}{p+q} \\ &= \frac{q}{p+q} + \frac{p(1-p-q)^n}{p+q}. \end{aligned}$$

Agora, como $0 < p < 1$ e $0 < q < 1$ segue que $|1-p-q| < 1$, e assim $(1-p-q)^n \rightarrow 0$ para $n \rightarrow \infty$. Portanto,

$$x_n \rightarrow \frac{q}{p+q},$$

quando $n \rightarrow \infty$.

□

Portanto, a distribuição da partícula no tempo n , dada pelo vetor de probabilidade $v_n = (x_n, 1-x_n)$ converge para

$$\left(\frac{q}{p+q}, \frac{p}{p+q} \right)$$

quando $n \rightarrow \infty$.

Uma abordagem diferente desta demonstração acima, pode ser visto, por exemplo, em Silva e Rota (2016). Como a matriz de transição T é regular, por uma versão do Teorema de Perron-Frobenius, é possível mostrar que a distribuição dada por $v_n = T^n v_0$ converge para um único vetor w , tal que $Tw = w$, independentemente da distribuição inicial v_0 . Além disso, caso o leitor tenha curiosidade, em Brezezniak e Zastawniak (1999, p.86), também tem uma outra demonstração em que não se usa soluções de recorrências lineares não-homogêneas (como apresentamos neste artigo) e nem Perron-Frobenius.

O objetivo principal agora, é mostrar que de fato o modelo apresentado no Exemplo 1 é uma cadeia de Markov. Para isto, considere o processo estocástico X_n como na Seção 2.0.2.

Teorema 5 O modelo apresentado no Exemplo 1 é uma cadeia de Markov, isto é, verifica (1).

Demonstração. Considere o espaço de estados $\mathcal{E} = \{1, 2\}$. Tomaremos Ω como sendo o conjunto de todas as sequências como mencionamos na Seção 2.0.2.

A fim de construir a probabilidade nos cilindros, tomemos ν_0 como sendo alguma probabilidade em $\mathcal{E}$. Apenas, para simplificar, vamos escolher $\nu_0(1) = 1$ e $\nu_0(2) = 0$. A medida ν_0 escolhida, corresponde a distribuição inicial do processo estocástico que iremos definir. Podemos então definir a probabilidade $\mathbb{P}$ da seguinte forma. Se $\omega = (e_0, e_1, \dots) \in \Omega$, tomemos

$$\mathbb{P}(\{\omega \in \Omega : \omega_0 = e_0\}) = \nu_0(\{e_0\});$$

$$\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n+1\}) = p(e_{n+1}|e_n)\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n\}), \quad (5)$$

onde $p(e_{n+1}|e_n)$ é a probabilidade de transição do estado e_n para e_{n+1} , podendo ser $p(1|1) = 1 - p$, $p(1|2) = q$, $p(2|1) = p$ ou $p(2|2) = 1 - q$.

Notemos que $\mathbb{P}$ está definida por meio de um processo indutivo que só depende da distribuição inicial e das probabilidades de transição. Por exemplo, a medida para o conjunto das sequências em que restringimos os dois primeiros estados, e_0 e e_1 , é dada pelo produto da probabilidade inicial de e_0 pela probabilidade de transição de e_0 para e_1 , isto é

$$\mathbb{P}(\{\omega \in \Omega : \omega_0 = e_0, \omega_1 = e_1\}) = p(e_1|e_0)\nu_0(\{e_0\}).$$

Analogamente, para o caso em que restringimos aos estados e_0, e_1 e e_2 , temos

$$\begin{aligned} \mathbb{P}(\{\omega \in \Omega : \omega_0 = e_0, \omega_1 = e_1, \omega_2 = e_2\}) &= p(e_2|e_1)\mathbb{P}(\{\omega \in \Omega : \omega_0 = e_0, \omega_1 = e_1\}) \\ &= p(e_2|e_1)p(e_1|e_0)\nu_0(\{e_0\}). \end{aligned}$$

E fazendo isso, sucessivamente, para mais estados, nota-se que de fato a fórmula (5) só depende das probabilidades de transição e da distribuição inicial. Além disso, como

$$\Omega = \{\omega \in \Omega : \omega_0 = 1\} \cup \{\omega \in \Omega : \omega_0 = 2\},$$

segue que $\mathbb{P}(\Omega) = \nu_0(\Omega) = 1$.

Como na Seção 2.0.2, considere o processo estocástico $X_n : \Omega \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, dado por

$$X_n(\omega) = \omega_n,$$

onde $\omega = (\omega_0, \omega_1, \dots, \omega_n, \dots)$. Primeiro vamos mostrar que as probabilidades de transição de X_n são o que deveriam ser, isto é,

$$\mathbb{P}(X_{n+1} = 2|X_n = 1) = p, \quad (6)$$

$$\mathbb{P}(X_{n+1} = 1|X_n = 2) = q. \quad (7)$$

Da definição de $\mathbb{P}$ temos que

$$\begin{aligned} \mathbb{P}(X_{n+1} = 2, X_n = 1) &= \mathbb{P}(\{\omega \in \Omega : \omega_n = 1, \omega_{n+1} = 2\}) \\ &= \sum_{e_0, \dots, e_{n-1} \in \mathcal{E}} \mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n-1, \omega_n = 1, \omega_{n+1} = 2\}) \\ &= \sum_{e_0, \dots, e_{n-1} \in \mathcal{E}} p(2|1)\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n-1, \omega_n = 1\}) \\ &= p \sum_{e_0, \dots, e_{n-1} \in \mathcal{E}} \mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n-1, \omega_n = 1\}) \\ &= p\mathbb{P}(X_n = 1). \end{aligned} \quad (8)$$

A segunda igualdade, em que aparece a soma sobre os estados, segue da Proposição 1 tomando uma partição em Ω , por meio de conjuntos formados de trajetórias em que se fixa os primeiros n estados $e_0, e_1, \dots, e_{n-1}$.

Pela definição de probabilidade condicional e da igualdade (8) segue que

$$\mathbb{P}(X_{n+1} = \mathbf{2} | X_n = \mathbf{1}) = \frac{\mathbb{P}(X_{n+1} = \mathbf{2}, X_n = \mathbf{1})}{\mathbb{P}(X_n = \mathbf{1})} = p.$$

Pelos mesmos argumentos segue que $\mathbb{P}(X_{n+1} = \mathbf{1} | X_n = \mathbf{2}) = q$.

Vamos agora verificar que

$$\mathbb{P}(X_{n+1} = e_{n+1} | X_0 = e_0, \dots, X_n = e_n) = \mathbb{P}(X_{n+1} = e_{n+1} | X_n = e_n).$$

De fato,

$$\begin{aligned} \mathbb{P}(X_{n+1} = e_{n+1} | X_0 = e_0, \dots, X_n = e_n) &= \frac{\mathbb{P}(X_0 = e_0, \dots, X_n = e_n, X_{n+1} = e_{n+1})}{\mathbb{P}(X_0 = e_0, \dots, X_n = e_n)} \\ &= \frac{\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n+1\})}{\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n\})} \\ &= \frac{p(e_{n+1} | e_n) \mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n\})}{\mathbb{P}(\{\omega \in \Omega : \omega_i = e_i, i = 0, \dots, n\})} \\ &= p(e_{n+1} | e_n). \end{aligned}$$

Por outro lado, de (6) e (7) temos que

$$\mathbb{P}(X_{n+1} = e_{n+1} | X_n = e_n) = p(e_{n+1} | e_n).$$

□

Apesar de trabalhosa, a demonstração acima poderia ser adaptada para um processo com mais de dois estados, desde que $\mathcal{E}$ seja finito. Boa parte da demonstração acima foi baseada em técnicas que aparecem em Brezezniak e Zastawniak (1999, p.88).

4 Agradecimentos

Agradecemos as contribuições dadas pela Comissão Científica e Editorial da revista C.Q.D..

A primeira autora agradece a FAPESP, processo 2016/21006-5, pelo suporte financeiro para desenvolver as atividades de Iniciação Científica.

A segunda autora agradece a bolsa de estudo fornecida pela CAPES durante a vigência do programa de pós-graduação em matemática (PROFMAT), período de realização deste trabalho.

5 Referências Bibliográficas

ALLEN, L. J. S. **An introduction to stochastic processes with applications to biology**. Upper Saddle River: Pearson Education Inc., 2003.

JAMES, B. R. **Probabilidade**: Um curso em nível intermediário. 2. ed. Rio de Janeiro: IMPA, 1996.

BRZEZNIAK, Z.; ZASTAWNIAK, T. **Basic stochastic processes**: a course through exercises. London: Springer, 1999. (Springer undergraduate mathematics series).

KEMENY, J. G.; SNELL, J. L.; KNAPP, A. W. **Denumerable Markov chains**. New York: Springer-Verlag, 1976.

MORGADO, A. C.; CARVALHO, P. C. P. **Matemática discreta**. Rio de Janeiro: SBM, 2013.

RUFFINO, P. R. C. **Uma iniciação aos sistemas dinâmicos estocásticos**. 2. ed. Rio de Janeiro: IMPA, 2009.

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.

Artigo recebido em jul. 2017 e aceito em set. 2017.