Simpósio de Integração Acadêmica

"Ciências Básicas para o Desenvolvimento Sustentável"

SIA UFV 2023

UFV Universidade Federal de Viçosa

CADEIA DE MARKOV E APLICAÇÕES

UNIVERSIDADE FEDERAL DE VIÇOSA - CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS - DEPARTAMENTO DE MATEMÁTICA

Fábio Ferreira de Freitas Júnior - fabio.f.junior@ufv.br Lais Moreira dos Santos - lais.msantos@ufv.br

Pesquisa-Cadeias de Markov-Matemática

Palavras chaves: Cadeias de Markov, Probabilidade.

1. Introdução

Processos estocásticos descrevem sistemas aleatórios que evoluem com o tempo e são uma importante ferramenta na compreensão de diversos fenômenos observados em diferentes áreas do conhecimento. Flutuações no mercado de ações e nas taxas de câmbio, variações no campo magnético da Terra e mutações que ocorrem no DNA de uma espécie são exemplos de fenômenos que podem ser melhor compreendidos através de processos estocásticos. Dentro desta área, destaca-se os processos markovianos, cujo o estudo teve como precursor o matemático russo Andrei Andreyevich Markov (1856-1922). Dizemos que um processo estocástico satisfaz a propriedade de Markov se os estados passados são irrelevantes para prever os estados futuros, desde que o estado atual seja conhecido. Nesse sentido, toda a informação necessária para predizer sobre o futuro se concentra no presente. Atualmente, as cadeias de Markov tem sido usadas em diversas áreas como economia, computação, biologia, etc. Prever o nível de volatilidade dos retornos de ativos, o PageRank de uma página da web usado no Google e simulações da função cerebral são alguns exemplos de aplicação de cadeias de Markov.

2. Objetivos

O objetivo deste presente projeto é o estudo da teoria matemática das cadeias de Markov e suas aplicações. Mais precisamente, estudamos cadeias de Markov, com enfoque principal no comportamento assintótico de tais processos. Como aplicação da teoria estudada, apresentamos o algoritmo PageRank.

3. Material e Métodos

O desenvolvimento desse projeto se deu através de pesquisa bibliográfica do tema, discussões semanais e apresentações de seminários para orientadora.

4. Cadeias de Markov

Um processo estocástico é uma coleção de variáveis aleatórias, denotado por $\{X_t, t \in T\}$, em que t é um parâmetro real, normalmente conhecido como tempo, pertencente a um conjunto T, que em geral é dado por $T = \mathbb{N}$ ou $T = [0, \infty)$. Neste trabalho, estudamos processos estocásticos que possuem a forma $\{X_n, n \in \mathbb{N}\}$ com espaço de estados discreto, em que $T = \mathbb{N}$, e que satisfazem a propriedade de Markov, cuja a definição é dada a seguir.

Definição 1 (Propriedade de Markov). *Dizemos que uma família de variáveis aleatórias $\{X_n, n \in \mathbb{N}\}$ possui a propriedade de Markov ou "perda de memória", se*

$$P(X_{n+1} = j \mid X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j \mid X_n = i),$$

para quaisquer que sejam os estados $i_0, i_1, ..., i_{n-1}, i, j$ e todo $n \geq 0$. Um processo estocástico $\{X_n, n \in \mathbb{N}\}$ com essa propriedade é chamado de cadeia de Markov.

Se um processo estocástico goza da propriedade de Markov, então o futuro do processo depende somente do estado presente, ou seja, os estados anteriores são irrelevantes para a predição do futuro da cadeia, desde que o estado atual seja conhecido.

As probabilidades condicionais $P(X_{n+1} = j \mid X_n = i)$ são denominadas probabilidades de transição e representam a probabilidade do processo estar no estado j no tempo $n+1$, dado que o processo está no estado i no tempo n.

Definição 2. *Dizemos que uma cadeia de Markov $\{X_n, n \in \mathbb{N}\}$ é homogênea se as probabilidades de transição não mudam com o tempo, ou seja,*

$$P(X_{n+1} = j \mid X_n = i) = P(X_1 = j \mid X_0 = i), \forall n \in \mathbb{N}.$$

Seja $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov homogênea. Denotaremos a probabilidade de transição $P(X_1 = j \mid X_0 = i) = P_{ij}$. Nesse caso, a matriz dada por $P = (P_{ij})$ será chamada de matriz de transição da cadeia. Usando as propriedades da probabilidade condicional, podemos mostrar que:

i) $P_{ij} \geq 0, \forall i, j \in \mathbb{S}$

ii) $\sum_{j \in \mathbb{S}} P_{ij} = 1.$

Além disso, vamos indicar por P^n a n-ésima potência da matriz de transição P e por $P_{ij}^{(n)}$ a entrada (i, j) dessa matriz.

Vamos fazer uma análise mais detalhada sobre a relação entre o comportamento de uma cadeia de Markov e sua matriz de probabilidade de transição de estados. Também estudaremos como uma cadeia de Markov se comporta a longo prazo e como isso se relaciona com os tipos de suas classes de estados.

Definição 3. *Seja $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov. Dizemos que o estado j é acessível desde i, e denotamos $i \to j$, se existe $n \in \mathbb{N}$ tal que*

$$P_{ij}^{(n)} = P(X_n = j \mid X_0 = i) > 0.$$

Além disso, se j é acessível desde i e i é acessível desde j, ou seja, se existem $n, m \in \mathbb{N}$ tais que $P_{ij}^{(n)} > 0$ e $P_{ji}^{(m)} > 0$, diremos que os estados i e j se comunicam e denotaremos isso por $i \leftrightarrow j$.

Em outras palavras, dizer que j é acessível desde de i, significa que, uma vez que a cadeia partiu do estado i, com probabilidade positiva ela atingirá o estado j em um tempo finito.

Definição 4. *Se dois estados se comunicam, dizemos que eles pertencem à mesma classe. Ademais, se nenhum estado fora da classe é acessível desde qualquer estado pertencente a classe, então dizemos que essa é uma classe fechada. Por fim, se todos os estados de uma cadeia de Markov são comunicantes, então todos os estados pertencem a uma única classe e essa cadeia de Markov é chamada de irredutível.*

Definição 5. *Um estado é dito ser transiente se, entrando neste estado, o processo pode nunca retornar para este estado. Isto é, o estado i é transiente se, e somente se, existe um estado j, com $j \neq i$, que é acessível a partir do estado i mas i não é acessível desde j. Caso contrário, o estado é dito ser recorrente.*

Definição 6. *Definimos o tempo médio de primeiro retorno ao estado i por*

$$\mu(i) = E(t_i \mid X_0 = i) = \sum_{n=0}^{\infty} n P(t_i = n \mid X_0 = i),$$

onde t_i é o número de transições até a primeira visita ao estado i.

Definição 7. *Seja $\{X_n\}$ uma cadeia de Markov e $i \in \mathbb{S}$ um estado recorrente, dizemos que*

- *i é recorrente nulo se $\mu(i) = \infty$;*
- *i é recorrente positivo se $\mu(i) < \infty$.*

O nosso objetivo é entender como se comporta $P_{ij}^{(n)}$ para n suficientemente grande.

Definição 8. *Seja $P = (P_{ij})$ a matriz de transição de uma cadeia de Markov. Dizemos que a cadeia $\{X_n\}$ é ergódica se para todo $j \in \mathbb{S}$ o limite $\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$ existe, é independente de i e $\sum_{j \in \mathbb{S}} \pi_j = 1$.*

Os seguintes resultados descrevem este comportamento para o caso em que j é recorrente.

Teorema 1. *Seja $\{X_n\}$ uma cadeia de Markov irredutível, aperiódica e recorrente positiva. Para cada $j \in \mathbb{S}$, temos $\lim_{n\to\infty} P_{ij}^{(n)} = \dfrac{1}{\mu(j)}, \forall i \in \mathbb{S}$, e portanto*

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{1}{\mu(0)} & \frac{1}{\mu(1)} & \frac{1}{\mu(2)} & \cdots \\ \frac{1}{\mu(0)} & \frac{1}{\mu(1)} & \frac{1}{\mu(2)} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ \frac{1}{\mu(0)} & \frac{1}{\mu(1)} & \frac{1}{\mu(2)} & \cdots \end{bmatrix}.$$

Embora o resultado anterior nos forneça informações desejáveis sobre o comportamento assintótico da cadeia, o cálculo de $\mu(j)$ nem sempre é trivial. Veremos a seguir uma maneira alternativa para determinar o limite $\lim_{n\to\infty} P_{ij}^{(n)}$.

Definição 9. *Seja $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov com espaço de estados $\mathbb{S}$ e matriz de transição $P = (P_{ij})$. Um vetor linha $\pi = (\pi_i)_{i\in\mathbb{S}}$ é dito ser uma distribuição estacionária para a cadeia de Markov, se satisfaz: $\pi P = \pi$ e $\sum_{i\in\mathbb{S}} \pi_i = 1$.*

Teorema 2. *Seja $\{X_n, n \in \mathbb{N}\}$ uma cadeia de Markov irredutível, aperiódica e recorrente positiva. Então existe uma única distribuição estacionária para a cadeia, que é dada por*

$$\pi_j = \frac{1}{\mu(j)}, j \in \mathbb{S}.$$

5. PageRank

O PageRank foi um algoritmo criado em 1997 por Larry Page e Sergey Brinn, com o objetivo de ordenar as páginas da Web por importância. Essa classificação revolucionou os métodos de busca até então existentes por otimizar o tempo de busca do usuário, fornecendo-o as páginas mais relevantes primeiro. Esse algoritmo serviu de base para a criação do Google, que é o principal buscador Web da atualidade.

Para apresentar o PageRank, considere um usuário que pesquisa algo na Web. O buscador inicialmente agrupa todas as páginas encontradas com as palavras-chaves inseridas na pesquisa. Em seguida, esses resultados serão ordenados por importância. Para introduzir o conceito de importância, admita que existam n webpages, onde a importância da i-ésima página será denotada por π_i. É natural supor que a quantidade de hyperlinks que uma página recebe influencie a sua pontuação. Entretanto, este não pode ser o único critério para medir a popularidade de um site, pois os sites que apontam para ele podem ser irrelevantes. Assim, a importância da i-ésima página, será da dada pela soma da importância das páginas que a indicam, ponderada pelas probabilidades de que tais páginas indiquem i. Desta forma, se a página j possuir k_j links e um desses links levar j para a página i, então j passará $\dfrac{1}{k_j}$ de importância para i. Logo, o PageRank da página i é dado por

$$\pi_i = \sum_j A_{ji} \frac{\pi_j}{k_j},$$

onde A_{ji} é 1 se j acessa i e 0 caso contrário. Podemos representar a relação acima por $\pi = P\pi$, ou equivalentemente $\pi^T = \pi^T P^T$, onde $\pi = (\pi_1, \pi_2, \cdots, \pi_n)$ e $P^T = \left(\dfrac{A_{ji}}{k_j}\right)^T$. Além disso, o vetor π pode sempre ser normalizado para que a soma de suas importâncias totalize 1. Observe que P^T seria uma matriz estocástica, caso não pudesse admitir linhas nulas. Entretanto, na maioria dos casos isso não é possível. Para contornar essa situação, vamos considerar o seguinte ajuste em $P = (P_{ij})$:

$$H_{ij} = \begin{cases} P_{ij}, & \text{se } \sum_i P_{ij} = 1, \\ \frac{1}{n}, & \text{se } \sum_i P_{ij} = 0. \end{cases}$$

Logo, temos que $H^T = (H_{ij})^T$ é a matriz de probabilidade de transição de uma cadeia de Markov. Assim, vamos buscar por distribuições estacionárias de H^T. Para tanto, considere o seguinte processo iterativo

$$\pi^{(k+1)} = H\pi^k, \text{ onde } \pi_i^{(0)} = \frac{1}{n}$$

Para que o processo iterativo anterior convirja para a única distribuição estacionária da cadeia de Markov, cuja a matriz de transição é H^T, precisamos garantir que os estados da cadeia satisfaçam as condições de ergodicidade, e isso nem sempre é possível. Para contornar essa situação, vamos considerar o seguinte ajuste ergódico em $H = (H_{ij})$:

$$G = \alpha H + (1 - \alpha)\frac{1}{n}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

onde α é a probabilidade de navegar em sequência.

Repare que quando chegamos em um estado recorrente, a cadeia fica presa em loop, pois a página i leva à página j e vice versa. No entanto, a matriz G foi definida de tal forma que quando a cadeia entra em um loop, ela consegue mudar para outra página com probabilidade $\dfrac{1-\alpha}{n}$. Portanto, a matriz G é uma matriz irredutível e aperiódica, ou seja, G é ergódica. Assim, a única distribuição estacionária de G nos fornecerá o vetor de importâncias das páginas.

6. Conclusões

Com esse projeto, além de aprofundar meu conhecimento em áreas tão essenciais da Matemática como Probabilidade e Álgebra Linear, tive a oportunidade de iniciar meus estudos em uma área de pesquisa tão profícua e multidisciplinar, que são as cadeias de Markov.

7. Apoio Financeiro

Este projeto contou com o apoio financeiro da CNPq.

8. Referências Bibliográficas

[1] D. Kannan, *An Introduction to Stochastic Processes*, North Holland, New York, 1979.

[2] H. M.Taylor, S. Karlin, *An Introduction to Stochastic Modeling*, 3^a ed., Academic Press, New York, 1998.

[3] B. R. James, *Probabilidade: Um Curso em Nível Intermediário*, 4^a ed., IMPA, 2015.