CADEIAS DE MARKOV: APLICAÇÕES NO COTIDIANO

Marcos André Pereira Almeida[1]
Douglas Alves[2]
Ian Rodrigues Santos[3]
Kleber Passos Silva[4]
Ana Clesia da Silva Melo[5]

Matemática Licenciatura

cadernos de graduação
ciências exatas e tecnológicas

ISSN IMPRESSO 1980-1777
ISSN ELETRÔNICO 2316-3135

RESUMO

Esta pesquisa aborda a aplicação das Cadeias de Markov em nosso cotidiano. Um processo especial estocástico criado pelo Suíço Andrei Andreyevich Markov. O estudo teve por objetivo analisar e discutir problemas relacionados com probabilidade, com o propósito de estudar um evento futuro de acordo com outro presente. O trabalho contém inicialmente um esclarecimento da teoria de Markov, logo em seguida, alguns exemplos de como a mesma é utilizada no dia a dia da sociedade. Os resultados foram favoráveis, por ter simplificado o nosso entendimento parcial do conteúdo presente, pois o estudo é uma amostra básica da importância das Cadeias de Markov para humanidade, além do mais, por causa do curto prazo não foi possível mostrar todos os pontos, porém, citamos com certeza os triviais para um rápido dos conceitos básicos.

PALAVRAS-CHAVE

Cadeias de Markov. Probabilidade. Matemática estatística.

ABSTRACT

This research deals with the application of Markov chains in our daily lives. A special stochastic process created by the Swiss Andrei Andreyevich Markov. The study aimed to analyze and discuss problems related to probability, in order to study a future event according to another present. The work initially contains a clarification of the Markov theory, soon after, some examples of how it is used in everyday life of society. Results were favorable, has simplified our partial understanding of this content because the study is a basic sample of the importance of Markov chains to mankind, moreover, because of the short period it was not possible to show all points, but we quote certainly trivial for a quick basic concepts.

KEYWORDS

Markov Chains. Probability. Mathematical Statistics.

1 INTRODUÇÃO

A matemática, desde o início de sua existência, esteve sempre presente na vida do homem, é fácil observamos o quanto ela foi, é e sempre será necessária em nosso cotidiano. A imensidão que é a matemática pode ser observada em seus ramos – geometria, trigonometria, cálculo e álgebra – que podemos aplicá-los em nosso benefício.

Cientificamente, estudamos matemática desde os primeiros anos escolares, mesmo às vezes sem saber alguma de suas diversas aplicações para resolução de nossos problemas. Hoje elaboraremos um trabalho sobre álgebra linear, apresentando uma de suas diversas aplicações para o cálculo de probabilidade, evolução de população etc.

Algumas das ferramentas da álgebra linear, como a solução de sistemas lineares, são datados do século II d.C., mesmo que não fosse algo tratado de forma separada. O método dos mínimos quadrados, técnica de otimização matemática, foi usado por Carl Friedrich Gauss no final do século XVIII, esta foi uma aplicação inicial e significante da álgebra linear.

O assunto começou a tomar forma por volta do século XIX, onde se viu muitas noções e métodos dos séculos anteriores abstraídas e generalizadas, daí se deu o início da álgebra abstrata. Matrizes e tensores foram introduzidos como objetos matemáticos abstratos na virada do século XX. Dado a importância do conteúdo, matrizes é estudada, ainda que em seus conceitos mais básicos, na 3ª etapa da educação básica.

O trabalho a seguir mostra apenas uma pequena parte das vastas aplicações da álgebra linear, utilizando o método das Cadeias de Markov basicamente, definindo este método e exemplificando.

2 CADEIAS DE MARKOV

Andrei Andreyevich Markov foi um matemático russo, formado pela Universidade Estatal de São Petersburgo em 1878, onde se tornou professor em 1886. Seus primeiros trabalhos foram sobre limites de integrais e teoria da aproximação. Depois de 1900 Markov aplicou métodos de frações contínuas, que havia sido iniciada por Pafnuty Chebyshev, na teoria da probabilidade. Ele também provou o teorema do limite central.

Mesmo tendo contribuído para a matemática com esses estudos, o nome de Markov ficou marcado pelos seus estudos sobre as cadeias de markov, método usado para resolução de problemas com probabilidades. Os processos de Markov é um tipo especial de processo estocástico, onde as distribuições de probabilidade para o passos futuros do processo dependem somente do estado presente, desconsiderando como o processo chegou a tal estado.

2.1 PROCESSO ESTOCÁSTICO

Um processo estocástico é uma família de variáveis aleatórias {X(t),t € t} , definidas em um espaço de probabilidade, indexado por um parâmetro t, onde t varia no conjunto T.

Considere um processo estocástico X(t). Para um tempo fixo t_1, $X(t_1) = X_1$ é uma variável aleatória e a sua função de distribuição acumulada $F_x(x_1;t_1)$ é definida por:

$$F_x(x_1;t_1) = P_r[X(t_1) \leq x_1]$$

$\mathbf{F_x(x_1;t_1)}$ é conhecida como a distribuição de primeira ordem de X(t). De forma geral, define-se a distribuição de n-ésima ordem de X(t) é dada por:

$$F_x(x_1,\ldots,x_n;t_1,\ldots,t_n) = P_r[X(t_1) \leq x_1,\ldots,X(t_n) \leq x_n]$$

Para uma caracterização completa do processo estocástico X(t) é preciso saber as distribuições de todas as ordens ($n \rightarrow \infty$).

2.2 MATRIZ DE TRANSIÇÃO

Considere a Cadeia de Markov com espaço de estados E = {1,2,3}, teremos a seguinte matriz de transição:

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & P_{1,3} \\ P_{2,1} & P_{2,2} & P_{2,3} \\ P_{3,1} & P_{3,2} & P_{3,3} \end{bmatrix}$$

Entende-se que $P_{1,1}$ é do estado **1** para o estado **1**.

2.3 INTERPRETAÇÃO DA DISTRIBUIÇÃO LIMITE

Existem duas interpretações para essa distribuição:

1) Após o processo está sendo executado por um longo período a probabilidade de encontrarmos a cadeia em um dado estado **j** é $P_{j.}$

2) P_j significa a fração média do tempo em que a cadeia se encontra no estado **j**.

2.4 CADEIAS DE MARKOV

A probabilidade de X_{n+1} estar no estado **j** dado que X_n está no estado **i** é chamada probabilidade de transição de um passo é denotada por $P_{ij}^{n,n+1} = \Pr\{\frac{Xn+1=j}{Xn=i}\}$.

Esta notação enfatiza que, em geral, as probabilidades de transição são funções não somente dos estados inicial e final, mas também do tempo de transição. Quando as probabilidades de transição em um passo são independentes da variável do tempo n, onde elas não se alterarão mais, dizemos que a Cadeia de Markov tem probabilidades de transição estacionárias.

Então $P_{ij}^{n,n+1} = P_{ij}$ é independente de n e P_{ij} é a probabilidade condicional que o valor do estado transite de **i** a **j**. Podem-se visualizar as quantidades Pij de forma matricial, num arranjo quadrático infinito:

$$\boldsymbol{P} = |\boldsymbol{Pij}|$$

As Cadeias de Markov são caracterizadas pelo cálculo das probabilidades de transições em n passos. São fundamentais, portanto as matrizes de probabilidades de transição em n passos, denotados por,

$$\boldsymbol{P}\,(\mathbf{n}) = |\boldsymbol{Pij}^{(n)}|$$

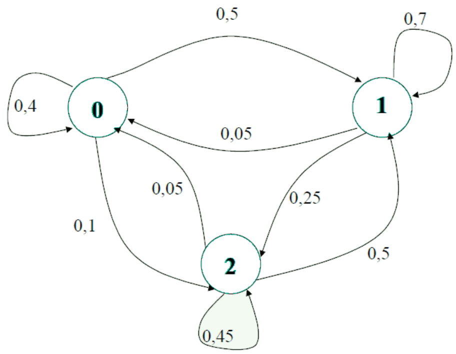
sabendo que $P_{ij}^{(n)}$ é a probabilidade que o processo vá do estado **i** para o estado **j** em **n** transições.

2.5 EXEMPLOS

Exemplo I

Observe o diagrama de mudança de estados, denotados por: $X_{0,}$ X_1 e X_2.

Figura 1 – Diagrama de mudança de estados



Fonte: Próprio autor.

A matriz de transição da probabilidade é dada por:

$$\begin{array}{cc} \begin{array}{ccc} X_0 & X_1 & X_2 \end{array} & \\ \begin{bmatrix} 0,4 & 0,5 & 0,1 \\ 0,05 & 0,7 & 0,25 \\ 0,05 & 0,5 & 0,45 \end{bmatrix} & \begin{array}{c} X_0 \\ X_1 \\ X_2 \end{array} \end{array}$$

Os determinantes das probabilidades limites são:

$X_0 = 0,4X_0 + 0,05X_1 + 0,05X_2$
$X_1=0,5X_0 + 0,7X_1 + 0,5X_2$
$X_2= 0,1X_0 + 0,25X_1 + 0,45X_2$
$X_0 + X1 + X2= 1$

Calculando os resultados para X_0, X_1 e X_2, temos:

$$X_0= \frac{1}{13}, \quad X_1 = \frac{5}{8} \text{ e } X_2 = \frac{31}{104}$$

Exemplo II

Podemos observar que um fenômeno que não é naturalmente um processo de Markov pode ser modelado, incluindo parte da história passada em cada estado, vejamos. Suponha, por exemplo, que o clima em qualquer dia depende do clima nos dois dias anteriores. Especificamente, suponha que se hoje e ontem o clima foi ensolarado entã oamanhã o clima será ensolarado com probabilidade 0,8. Se foi ensolarado hoje e chuvoso ontem, então amanhã será ensolarado com probabilidade 0,6. Se for chuvoso hoje, mas ensolarado ontem, então amanhã será ensolarado com probabilidade

0,4. E se os dois últimos dias foram chuvosos, então amanhã será ensolarado com probabilidade 0,1. Então, definiremos os estados assim:

0 - Ensolarado nos últimos dois dias.
1 - Ensolarado ontem, mas chuvoso hoje.
2 - Chuvoso ontem, mas ensolarado hoje.
3 - Chuvoso nos dois últimos dias.

Fazendo quatro Tabelas para saber a porcentagem de o dia de amanha ser ensolarado ou chuvoso de acordo com os dias de hoje e ontem, temos:

As siglas das Tabelas e para o fim deste exemplo são: **S** = Sol, **C** = Chuva, **EA** = Estado atual, **EF** = Estado final.

0 – Ontem sol, hoje sol.

	EF	
EA	S	C
S, S	0,8	0,2

1 – Sol ontem, chuva hoje.

	EF	
EA	S	C
S, S	0,4	0,6

Fonte: Próprio autor.

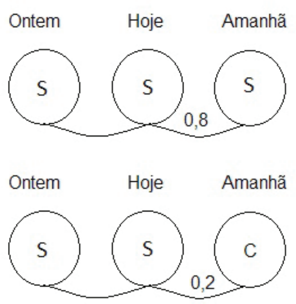
2 – Chuva ontem, sol hoje.

	EF	
EA	S	C
S, S	0,6	0,4

3 – Chuva ontem e hoje.

	EF	
EA	S	C
S, S	0,1	0,9

Para montar a matriz de transição teremos que entender e interpretar as Tabelas acima que mostram a chance de amanhã chover ou não. Cada tabela apresenta duas transições, tomando como exemplo o estado **0** (Sol ontem e hoje), temos o diagrama:



Fonte: Próprio autor.

A matriz de transição A_{ij} denota um estado que vai de **i** para **j**, o que implica dizer que as linhas equivalem ao estado atual e as colunas o estado final. Denotamos assim a matriz de Transição A:

s,ss,cc,sc,c

$$A=\left[\begin{array}{cccc} \quad & \quad & \quad & \quad \\ \\ \\ \\ \end{array}\right]\begin{array}{l} s,s\\ s,c\\ c,s\\ c,c\end{array}$$

Ainda observando o diagrama do estado **0**, podemos encontrar a sua posição na matriz de transição observando a primeira transição (s,s) e a segunda (s,s) temos que colocar o valor na linha "s,s" e coluna "s,s", pois as linhas são equivalentes ao estado inicial e as colunas os estados finais. O outro diagrama mostra que a primeira transição é "s,s" e a segunda é "s,c", então o valor achado será colocado na linha "s,s" e na coluna "s,c". Fazendo isso para todas as tabelas montadas anteriormente e colocando o número 0 não há transições, temos a matriz de transição dada por:

$$\begin{bmatrix} 0{,}8 & 0{,}2 & 0 & 0 \\ 0 & 0 & 0{,}4 & 0{,}6 \\ 0{,}6 & 0{,}4 & 0 & 0 \\ 0 & 0 & 0{,}1 & 0{,}9 \end{bmatrix}$$

Com a matriz de transição completa, as equações determinantes das propriedades limites são:

$$x = 0{,}8x + 0{,}6z$$

$$y = 0{,}2x + 0{,}4z$$

$$z = 0{,}4y + 0{,}1ty$$

$$t = 0{,}6y + 0{,}9t$$

$$x + y + z + t = 1$$

Resolvendo o sistema teremos os seguintes resultados para cada estado:

$$x = \frac{3}{11}, y = \frac{1}{11}, z = \frac{1}{11} \text{ e } t = \frac{6}{11}$$

Pode-se perceber que os estados finais (colunas) que mostram a transição para o dia ensolarado são os estados **0** e **2**, no caso acima, ***x*** e ***z***, somando os dois, temos uma probabilidade de $\frac{4}{11}$.

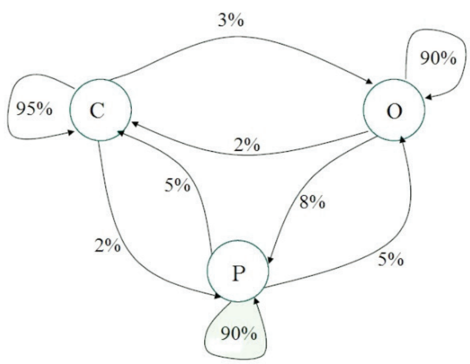
Exemplo III

Imagine um serviço de aluguel/empréstimo de bicicletas públicas. Para conseguir o empréstimo de alguma, a pessoa precisa ir a alguma das três estações: Centro (C), Orla (O) e Parque (P). Suponhamos, também, que as devoluções das bicicletas devem ser obrigatoriamente em algum dos locais citados acima. Os dados obtidos revelam que a movimentação das bicicletas ocorre da seguinte maneira: 95% das bicicletas que são pegas no centro são devolvidas no centro, 3% das bicicletas retiradas do centro são devolvidas na Orla e 2% delas são devolvidas no parque.

Em relação às bicicletas pegas na Orla, temos que 2% são devolvidas no centro, 90% são deixadas na Orla, e 8% são devolvidas no Parque. Sobre as bicicletas retiradas no Parque, 5% são deixadas no Centro, 5% no Orla e 90% deixadas no mesmo local. Sabendo que a distribuição inicial, no primeiro dia, do total de bicicletas foi de 50% no Centro, 30% na Orla e 20% no Parque. Qual será a distribuição ao final deste dia?

Verificamos que o diagrama de transição ficou assim:

Figura 2 – Diagrama de mudança de estados



Fonte: Próprio autor

Do diagrama retiramos a matriz de transição:

$$\begin{array}{cc} \begin{array}{ccc} \text{C} & \text{O} & \text{P} \end{array} & \\ \begin{bmatrix} 0{,}95 & 0{,}02 & 0{,}05 \\ 0{,}03 & 0{,}9 & 0{,}05 \\ 0{,}02 & 0{,}08 & 0{,}9 \end{bmatrix} & \begin{array}{c} \text{C} \\ \text{O} \\ \text{P} \end{array} \end{array}$$

A matriz de transição acima será chamada de M. A distribuição inicial é dada pelo vetor:

$$X_0 = \begin{array}{c} \begin{bmatrix} 0,5 \\ 0,3 \\ 0,2 \end{bmatrix} \end{array} \begin{array}{c} C \\ O \\ P \end{array}$$

A quantidade final de bicicletas em cada uma das estações será dada por:

$$M^*X_0 = 0,5 \begin{bmatrix} 0,95 \\ 0,03 \\ 0,02 \end{bmatrix} + 0,3 \begin{bmatrix} 0,02 \\ 0,9 \\ 0,08 \end{bmatrix} + 0,2 \begin{bmatrix} 0,05 \\ 0,05 \\ 0,9 \end{bmatrix} = \begin{bmatrix} 0,491 \\ 0,295 \\ 0,214 \end{bmatrix} \left\{ \begin{array}{l} \text{Essa é a quantidade de} \\ \text{distribuição de bicicleta} \\ \text{ao fim do primeiro dia.} \end{array} \right.$$

Caso seja necessário o cálculo do dia seguinte, é só trocar o vetor de distribuição X_0 pelo X_1 que representa o dia seguinte. Assim, é necessário atentar a propriedade da falta de memória das Cadeias de Markov, pois o cálculo depende apenas do estado anterior ao seguinte, como no exemplo. O cálculo de X_2, terceiro dia, depende apenas do X_1, segundo dia, desconsiderando o X_0, primeiro dia.

3 CONCLUSÕES FINAIS

Procuramos, por meio deste trabalho, mostrar algumas das aplicações das Cadeias de Markov no nosso cotidiano em relação a probabilidades. Ressaltamos também, que este não é um estudo completo do conteúdo, há vários outros aspectos que podem ser estudados por meio das Cadeias de Markov.

Notamos, durante o desenvolvimento, que para o estudo destas aplicações é necessário o conhecimento sobre sistemas lineares, geometria analítica, matrizes e conceitos básicos de probabilidade, para o seu entendimento de maneira completa e sua aplicação de forma mais simples possível.

Além disto, este trabalho foi uma oportunidade de fazer a relação do conteúdo da álgebra linear e outras disciplinas, com exemplos práticos, mostrando a importância de cada uma para um profissional que pretende se formar na área de exatas, já que estas relações são pouco vistas em sala de aula.

Portanto, na medida de nossos objetivos com o trabalho, consideramos que obtivemos êxito na busca do mesmo de forma plena, ressaltando que devido o curto prazo, não foi possível elaborar um trabalho mais detalhado, podendo o mesmo ser completado posteriormente.

REFERÊNCIAS

LOPES, Hélio. **Processos Estocásticos**. INF2035 – Introdução a Simulação Estocátisca. Notas de aula. Disponível em: http://www-di.inf.puc-rio.br/~lopes//courses/inf2035---simulacao-estocas.html. Acesso em: 12 dez. 2014

GARCIA, Nancy L. **Inferência para Cadeias de Markov.** Unicamp. 2012
NOGUEIRA, Fernando. **Cadeias de Markov**. Notas de aula. Disponível em: <http://www.ufjf.br/epd042/files/2009/02/cadeiaMarkov1.pdf>. Acesso em: 12 dez. 2014

SILVA, Tallyta Carolyne Martins da. **Cadeias de Markov**: Conceitos e Aplicações em Modelos de Difusão de Informação. 2010. Trabalho de Conclusão de Curso (Graduação em Estatística) – Universidade Federal de Goiás, 2010.

SPECK, Ana Lice Caroline. **Coletânea de aplicações da álgebra linear**. 2006. Trabalho de Conclusão de Curso (Graduação em Matemática) – Centro de Ciências Físicas e Matemáticas, Universidade Federal de Santa Catarina, Florianópolis-SC, 2006.

Data do recebimento: 20 de Maio de 2015
Data da avaliação: 19 de junho de 2015
Data de aceite: 29 de junho de 2015

1. Graduando em Licenciatura em Matemática – UNIT. E-mail: marcinho399@hotmail.com
2. Graduando em Licenciatura em Matemática – UNIT. E-mail: gigamega_04@hotmail.com
3. Graduando em Licenciatura em Matemática – UNIT. E-mail: ian_rs15@hotmail.com
4. Graduando em Licenciatura em Matemática – UNIT. E-mail: kleber_cross15@hotmail.com
5. Graduanda em Licenciatura em Matemática – UNIT. .