

Olzemir J. da S. Lobato

**Modelos de atendimento ao público via
cadeias de Markov**

Abaetetuba, Pará – BR

2019

Olzemir J. da S. Lobato

Modelos de atendimento ao público via cadeias de Markov

Trabalho de Conclusão de Curso apresentado a Faculdade de Ciências Exatas e Tecnologia da Universidade Federal do Pará, Campus Universitário do Baixo Tocantins, como requisito final para obtenção do grau de Licenciado Pleno em Matemática, sob orientação do Prof. Me. Raimundo das Graças Carvalho de Almeida.

Universidade Federal do Pará – UFPA

Faculdade de Ciências Exatas e Tecnologia – FACET

Orientador: Prof. Me. Raimundo das Graças Carvalho de Almeida

Abaetetuba, Pará – BR

2019

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

L796m Lobato, Olzimir Jeffrey da Silva
Modelos de atendimento ao público via cadeias de Markov /
Olzimir Jeffrey da Silva Lobato. — 2019.
63 f. : il. color.

Orientador(a): Prof. Me. Raimundo das Graças Carvalho de Almeida

Trabalho de Conclusão de Curso (Graduação) - Curso de Matemática, Campus Universitário de Abaetetuba, Universidade Federal do Pará, Abaetetuba, 2019.

1. Cadeias de Markov . 2. Matrizes de probabilidades . 3. Processo de Poisson . 4. Modelo de fila . I. Título.

CDD 519



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DO PARÁ
CAMPUS UNIVERSITÁRIO DE ABAETETUBA
SECRETARIA GERAL DE GRADUAÇÃO

ATA DA DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Aos 11 dias do mês de JULHO do ano de dois mil e dezenove, na SALA 13, no Campus Universitário de Abaetetuba, reuniram-se os Membros da Banca Examinadora, abaixo-assinados, sob a presidência do(a) Professor(a) RAIMUNDO DAS GRAÇAS PARVALHO DE Almeida com a finalidade de examinar em forma final o Trabalho de Conclusão de Curso do(a) aluno(a) OLZEMIR JEFFREY DA SILVA LOBATO da Turma/curso de 2013-LIC. EM MATEMÁTICA do supramencionado Campus, sob o título MODELOS DE ATENDIMENTO AO PÚBLICO VIA CADEIAS DE MARKOV. A sessão teve início às 18 horas e 5 minutos e foi encerrada às 18 e 55 minutos. Após a exposição do aluno, houve arguição dos Membros da Banca Examinadora seguido de resposta do(a) aluno(a). Em seguida, a Banca reuniu-se para deliberar sobre o CONCEITO atribuído ao TCC. Por consenso, a Banca Examinadora decidiu:

() Não aprovar

Aprovar com o conceito BOM.

Aprovar com o conceito BOM e recomendações para a revisão dos seguintes pontos: JUSTES GERAIS DO TRABALHO

Obs.: Caso o trabalho seja aprovado com recomendações de revisão, o (a) aluno (a) dispõe de ___ dias para dar forma final ao trabalho e entregar em CD no formato digital em PDF na secretaria Geral de Graduação, após anuência da banca examinadora.

Abaetetuba (PA), 11 de JULHO de 2019.

[Assinatura]
Presidente/Orientador(a)

[Assinatura]
Membro/Examinador(a)

[Assinatura]
Membro/Examinador(a)

*Dedico este trabalho aos meus pais,
Otacílio, em memória,
e Zelice.*

Agradecimentos

Em primeiro lugar, quero agradecer a Deus, pois acredito que Ele concede sabedoria a todos que lhe pedem com fé (citação curta da Bíblia Sagrada, Tiago 1:1-9). Também sou muito grato à minha mãe, que sempre viu potencial em mim. Nos momentos mais difíceis da minha vida, ela me apoiou com uma força sobre-humana. Me ensinou minha fé e meus princípios e agora, com o nível de ensino superior, tenho constatado o quanto eles são excelentes.

Também foram valiosas as orientações dispensadas a mim pelo Professor Mestre Raimundo de Almeida. Seu modo polido tanto me motivou como ensinou.

Agradeço muito à minha irmã, Daniela, por ter me ajudado a resolver contra-tempos que poderiam ter impedido a conclusão deste trabalho.

Também é meu desejo reconhecer a ajuda que obtive nos fóruns de \LaTeX , dentre eles, <https://tex.stackexchange.com> e, principalmente, o grupo \LaTeX -br <https://groups.google.com/group/latex-br>.

*“Deus não só joga dados com o Universo,
também os jogou onde ninguém consegue encontrar.”*

*Paráfrase de uma fala
do personagem Stephen Hawking,
no filme “A Teoria de Tudo”.*

Resumo

Cadeias de Markov são uma poderosa ferramenta probabilística de previsão do comportamento de diversos sistemas, cite-se como exemplo a gestão de quadro de funcionários para atendimento em estabelecimentos comerciais (1); em ciências físicas, provê um modelo com propriedades globais de interações locais de acordo com o artigo de Aharonov et al. (2). As aplicações são muitas; cite-se ainda os modelos de PageRank, uma aplicação fundamental para motores de busca da internet mais modernos como o Google (3, 4). Todavia, o foco deste trabalho são modelos de atendimento de pessoas em estabelecimentos, como bancos, repartições públicas, aeroportos, postos de combustível, por exemplo, ou por telefone, via cadeias de Markov, como sugerido pelo título do trabalho. Interessante notar também que existe uma variedade de situações diversas que podem ser descritas com o mesmo modelo. O trabalho está dividido em duas partes maiores, a primeira relembra conhecimentos básicos de Probabilidade e um pouco de Álgebra Linear; somente o necessário. Não se faz estritamente necessária a leitura da parte 1; mas em caso de dúvidas ou dificuldades, é possível que uma rápida consulta ajude. A contribuição deste trabalho está em estimar o tamanho futuro de uma fila, nas situações relatadas acima, após um tempo t indicado, com auxílio de um programa de computador em linguagem C construído neste trabalho de pesquisa. Também foi calculada a complexidade computacional do algoritmo.

Palavras-chave: cadeias de Markov, matrizes de probabilidades, processo de Poisson, modelo de fila.

Abstract

Markov chains are a powerful probability tool of prediction to behavior of sundry systems. For example, staff management for attendance in stores (1); in physic sciences, they provide a model with global properties of local interactions, following Aharonov et al article (2). The applications are several; also the PageRank models, one fundamental application to web search engines most smart, like the Google (3, 4). Nevertheless, the focus of this work is the attendance models of people in establishments, like a banks, public repartitions, airports, for example, or attendance by phone, by using Markov chains, as suggested by work title. Interesting also to note there is a manifold of different situations that may to be described with the same model. The work is divided in two grander parts, the first remembers basic knowledge of probability and a little of linear algebra; only the necessary. It is not strictly necessary the reading of part 1; though, in case of doubts or difficulties a quick consultation is likely to help. The work contribution is in estimating the posterior queue's size, after a time given t , on mentioned situations, with computational help of a program built in C language, for this search work. Too, it was evaluated the algorithm's computational complexity.

Key-words: Markov chains, matrices of probabilities, Poisson process, queue model.

Sumário

	Notação	11
	Breve histórico	12
	Introdução	13
I	NOÇÕES BÁSICAS DE PROBABILIDADE E ÁLGEBRA LINEAR	14
1	PROBABILIDADE	15
1.1	Experimento aleatório	15
1.2	Conceitos de probabilidade	15
2	ÁLGEBRA LINEAR	20
2.1	Vetores e matrizes	20
II	CADEIAS DE MARKOV	22
3	INTRODUÇÃO ÀS CADEIAS DE MARKOV	23
3.1	Exemplos	24
3.1.1	Problema de mercado	24
4	CONSTRUÇÃO DO MODELO DE MARKOV	27
4.1	Algoritmo para obterem-se as probabilidades de transições	27
4.1.1	Exemplo com dados empíricos	27
5	DISTRIBUIÇÃO ESTACIONÁRIA	29
5.1	Exemplo de convergência	30
5.2	Sistemas Lineares	30
5.2.1	Método de Jacobi	31
5.2.2	Método de Gauss-Seidel	33
6	CADEIA DE MARKOV COM TEMPO CONTÍNUO DE DOIS ESTADOS	35
6.1	Processo de Poisson	35
6.1.1	Distribuição exponencial	36
6.1.2	Fórmula de Poisson	38
6.2	Uma solução para um problema prático	40

7	CADEIAS DE MARKOV COM TEMPO CONTÍNUO	44
7.1	Modelagem do tamanho de filas	44
7.2	A função de probabilidade $P_{ij}(t)$	45
7.2.1	Método iterativo de computar $P_{ij}(t)$	48
7.2.1.1	Complexidade computacional	49
7.3	A matriz R no modelo de filas	51
7.4	O tamanho da fila após um tempo t	54
8	OBTENÇÃO DOS PARÂMETROS DOS PROCESSOS	55
	Conclusões	56
	O programa de computador criado	56
	Método analítico para estimação de fila	57
	ANEXO A – CÓDIGO-FONTE DO PROGRAMA	58
	REFERÊNCIAS	62

Notação

π	Distribuição estacionária da cadeia de Markov
δ_{ij}	Delta de Kroenecker
$\mathcal{O}(p)$	Complexidade computacional de ordem polinomial p
$E[X]$	Esperança matemática, ou ainda, valor esperado
F_{ij}	Frequência de transições do estado i ao estado j observada
$M_{m \times n}$	M é uma matriz de m linhas e n colunas
$o(h)$	Função que tende a zero mais rapidamente do que h
$P(A)$	Probabilidade do evento A
$P(A B)$	Probabilidade de A condicionado que B ocorre
$P_0(t)$	Probabilidade de o atendente estar desocupado no tempo t , ou, de nenhum consumidor chegar até o tempo t , conforme o contexto
$P_1(t)$	Probabilidade de o atendente estar ocupado no tempo t , ou, de chegar um consumidor até o tempo t , conforme o contexto
P_{ij}	Probabilidade de passar do estado i ao estado j em uma unidade de tempo discreto, ou, conforme o contexto, probabilidade de haver uma transição ao estado j condicionado que haverá uma transição do estado i
$P_{ij}(t)$	Probabilidade $P\{X(0) = i \cap X(t) = j\}$
$P_{ij}^{(n)}$	Probabilidade de, partindo do estado i , estar no estado j depois de n tempos discretos
$P_n(t)$	Chegada de n consumidores até o tempo t
S	Espaço de estados da cadeia de Markov
U	Espaço amostral, ou conjunto universo, relativo a um experimento aleatório

Breve histórico

Cadeias de Markov receberam esta designação em homenagem a Andrei Andreyevich Markov, um matemático russo que viveu de 1856 a 1922. Viveu a maior parte de sua vida em Petrogrado, atual São Petesburgo, na Rússia. Seu irmão mais novo, Vladimir, estava ganhando prestígio internacional como matemático quando infelizmente faleceu, à idade de apenas 25 anos acometido de tuberculose. Andrei Markov teve sua formação acadêmica na prestigiada Universidade de São Petesburgo. Fez importantes contribuições à Matemática no campo de Teoria dos Números – suas ideias foram tão revolucionárias que se passaram em torno de 40 anos, desde a publicação de sua tese de mestrado **Sobre as formas binárias quadráticas com determinante positivo**, para que esse trabalho fosse totalmente aceito na Europa Ocidental. Todavia, ele é mais lembrado por seu trabalho em cadeias de Markov; acredita-se até mesmo que esta nova teoria inaugurou um novo ramo na Matemática, o campo de estudo de **processos estocásticos**. Markov desenvolveu sua teoria de cadeias como um trabalho de Matemática Pura inicialmente. Todavia ele era interessado em poesia, e utilizou cadeias de Markov de dois estados para estudar a distribuição de vogais e consoantes em textos literários. Em julho de 1922, Markov veio a falecer, porém deixando um grande legado para a Matemática. Interessante ressaltar também que ele teve um filho, de mesmo nome, que se tornou um renomado matemático na geração seguinte. A fonte é a referência (5).

Introdução

As cadeias de Markov são um objeto de estudo deveras fértil em aplicações. Este fato contribuiu para que a literatura relacionada se multiplicasse, conquistando assim, pesquisadores que trabalham com variadas abordagens em problemas probabilísticos. Este trabalho se propõe a melhorar o aproveitamento de recursos humanos em situações de atendimento ao público, em particular, o atendimento pessoal, embora não se descarte a possibilidade de técnicas apresentadas neste trabalho serem estendidas à situação de servidores de internet que mantêm um site acessível na rede.

O trabalho é dedicado a dois modelos de atendimento ao público. O primeiro calcula a probabilidade de um atendente de loja estar ocupado. Isto é interessante porque, se todos os atendentes em uma loja estão simultaneamente ocupados com frequência, os clientes que não são prontamente atendidos podem deixar a loja sem realizar qualquer compra, e isto representa prejuízo para a empresa; por outro lado, se há atendentes ociosos por um tempo longo, isto significa desperdício de recursos humanos que poderiam ser aproveitados em outras ocupações.

Admite-se porém que a aplicabilidade do citado modelo requer um esforço adicional de criar um aplicativo de computador ou celular que automatize os cálculos, e que seja facilmente manuseável. Não obstante, o segundo modelo conta com um aplicativo de computador que faz o mencionado, sendo necessário apenas que o usuário insira ao programa os dados observados no contexto do trabalho de atendimento ao público; o código-fonte do programa foi escrito em linguagem C, e está anexo a este trabalho.

Este segundo modelo é dedicado à situações de atendimento ao público onde o fato de o atendente estar ocupado não afasta o interesse de ser atendido, ensejando a formação de filas, como acontece em agências bancárias, repartições e hospitais públicos, centros e postos de saúde.

Em regra geral, instituições financeiras possuem altos rendimentos, e optam por comprar softwares de empresas especializadas. Olhando por outro lado, diversos serviços públicos estão afetados por problemas de má administração. Acredita-se que a parte deste trabalho mencionada agora, em cooperação com a boa vontade dos administradores públicos, possam contribuir para o melhor atendimento da população interessada, mediante planejamento para um tempo de espera razoável.

Parte I

Noções básicas de Probabilidade e Álgebra Linear

1 Probabilidade

Os conteúdos deste capítulo são destinados à consulta, conforme o necessário, para suporte ao desenvolvimento do tema; e tiveram como fonte o livro de Meyer (6).

1.1 Experimento aleatório

Um conceito essencial para compreender a Teoria das Probabilidades é a noção de experimento aleatório. Uma definição precisa deste conceito é um pouco complexa, e dispensável para os propósitos deste trabalho; uma compreensão intuitiva é preferível.

Experimentos constituem parte essencial à construção de um modelo científico, ou projeto de engenharia. Para que um experimento esclareça o fenômeno estudado, no entanto, é preciso que as condições de realização do experimento sejam controladas; por outro lado, não é possível o controle total dessas condições. Algumas vezes, esses fatores que não são totalmente controlados não influenciam significativamente nos resultados dos experimentos; como consequência, a cada repetição do experimento, os resultados são praticamente os mesmos. De outro lado, existem experimentos que são sensivelmente afetados por fatores que não podem ser totalmente controlados, disto resulta uma variação significativa nos resultados quando o experimento é repetido muitas vezes; neste caso, o experimento é chamado **experimento aleatório**; caso contrário, é denominado um **experimento determinístico**.

Exemplo 1 (Experimento determinístico). Lançamento vertical de um projétil de massa considerável. A altura h é determinada pela equação:

$$h = v_0 \cdot t + \frac{g \cdot t^2}{2},$$

onde v_0 é a velocidade no momento do lançamento, g é a aceleração da gravidade e t é o tempo após o lançamento.

Exemplo 2 (Experimento aleatório). Vida-útil de uma lâmpada.

1.2 Conceitos de probabilidade

Definição 1.1. É definido **espaço amostral**, ou **conjunto universo**, o conjunto de todos os resultados possíveis de um experimento aleatório, e denotamos seu conjunto por U .

Definição 1.2. Define-se um **evento** A como um subconjunto do espaço amostral U . Isto é, $A \subset U$.

Definição 1.3. Define-se o **complemento** de um evento A o evento A^c que possui a propriedade $A \cup A^c = U$.

Definição 1.4. Define-se **probabilidade de um evento** A o número $P(A)$ que possui as seguintes propriedades

- a) $0 \leq P(A) \leq 1$;
- b) $P(U) = 1$;
- c) Se A e B forem eventos mutuamente excludentes, então $P(A \cup B) = P(A) + P(B)$.

Da definição 1.3 combinada com a definição 1.4 é possível demonstrar o seguinte teorema.

Teorema 1.1. Seja A um evento e A^c seu complementar. Então $P(A) = 1 - P(A^c)$.

Uma probabilidade de interesse é a **probabilidade condicional**. Para dois eventos A e B , a probabilidade de A ocorrer supondo que B ocorrerá é denotada por $P(A | B)$. Isto é interessante quando se tem dois eventos relacionados, e a ocorrência de um evento B influencia a probabilidade de ocorrência do evento A . Isto conduz ao seguinte teorema fundamental da Probabilidade.

Teorema 1.2. A probabilidade de A condicionado B , denotada por $P(A | B)$, é calculada pela fórmula

$$P(A | B) = \frac{P(A \cap B)}{P(B)}.$$

Demonstração. Utiliza-se o diagrama de Venn. Veja a figura 1.

Da definição de probabilidade 1.4, item 3, tem-se que a probabilidade de um evento X , $P(X)$, depende do tamanho do evento X ou, de outro modo, de quantos elementos o conjunto X possui. Esta medida do conjunto está relacionada a uma **estrutura teórica de medida**, que faz uma contagem em conjuntos finitos, e uma estrutura algébrica chamada σ -álgebra \mathcal{A} , onde \mathcal{A} denota o conjunto de todos os subconjuntos de um evento qualquer A , isto permite medir a probabilidade de um evento A , de acordo com a obra de Knill (7), página 7. A Teoria da Probabilidade vai muito longe para explicar as operações algébricas contidas nesta demonstração, deste modo não foi possível incluí-las aqui. De todo modo, deixando de lado o rigor e os aspectos mais profundos da teoria, vê-se que a demonstração faz sentido. Denote-se o tamanho de X por $|X|$. Observe agora o diagrama de Venn. Suponha que B ocorre, procura-se saber qual a probabilidade de A ocorrer condicionado que B ocorre. Se A e B ocorrem, então o resultado do experimento pertence ao conjunto $A \cap B$; por outro lado, é condicionado que B ocorre, mas não é condicionado que A ocorre, portanto o resultado pertence ao conjunto B por hipótese, mas não necessariamente, pertence a A . Toma-se então B como conjunto universo.

Perceba que $0 < |A \cap B| < |B|$, portanto

$$0 < \frac{|A \cap B|}{|B|} < 1,$$

satisfazendo portanto o axioma 1 da definição de probabilidade; tem-se que $P(B | B) = 1$, que é o segundo axioma; e, se $C, D \subset B$ e $C \cap D = \emptyset$, então

$$P(C \cup D | B) = P(C | B) + P(D | B)$$

verificando assim o terceiro axioma; portanto

$$P(X | B), \quad X \subset B,$$

é uma probabilidade, e é calculada por $\frac{|X|}{|B|}$. Considere agora o conjunto universo U , contendo todos os eventos possíveis do experimento. Foi mostrado que

$$P(A | B) = \frac{|A \cap B|}{|B|}.$$

Com uma pequena manipulação deste resultado, chega-se à conclusão do teorema em seguinte.

$$\begin{aligned} P(A | B) &= \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|U|}}{\frac{|B|}{|U|}} \\ &= \frac{P(A \cap B)}{P(B)}. \end{aligned}$$

É fácil ver que, para $X \subset U$,

$$\frac{|X|}{|U|} = P(X),$$

$P(X)$ satisfazendo os axiomas de probabilidade. □

Definição 1.5. Sejam A e B dois eventos independentes. Afirma-se que $P(A | B) = P(A)$, e $P(B | A) = P(B)$.

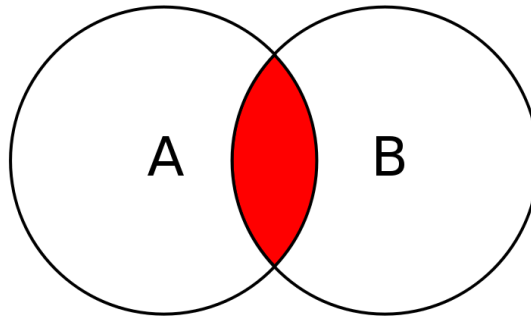
Definição 1.6 (Função densidade de probabilidade). A função densidade de probabilidade, $f(x)$, de uma variável aleatória X , é definida por

$$\int_a^b f(x)dx = P(a \leq X \leq b),$$

onde $P(X)$ é a probabilidade de X .

Definição 1.7 (Esperança matemática). Define-se média, ou esperança matemática, ou expectativa, de uma variável aleatória X , o valor esperado de X durante a realização de um experimento

Figura 1 – Diagrama de Venn.



Fonte: adaptado de Meyer (6).

aleatório associado a X . Isto é, se X é uma variável aleatória discreta que assume os valores $X = X_i$, então a média de X será

$$E[X] = \sum_{i=0}^{\infty} X_i \cdot P(X_i);$$

ou, se $X = x$ é uma variável aleatória contínua, com função densidade de probabilidade $f(x)$, define-se a média por

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

De acordo com Meyer (6) e Spiegel (8), tem-se as proposições 1.1 e 1.2.

Proposição 1.1. Se k é uma constante, então $E[k] = k$.

Proposição 1.2. Da definição de esperança matemática, dada acima, tem-se a seguinte propriedade. Se X e Y são duas variáveis aleatórias quaisquer, e C é uma constante, então

$$E[C \cdot X + Y] = C \cdot E[X] + E[Y].$$

A demonstração da proposição acima é um pouco complexa, portanto não será demonstrada no trabalho atual. Por outro lado pode ser conveniente retomar este resultado em trabalhos futuros fazendo a sua demonstração, com o intuito de agregar conhecimentos mais sólidos à aplicação pretendida. Ver as conclusões.

Definição 1.8 (Distribuição exponencial). Define-se distribuição de probabilidade exponencial a distribuição que tem função densidade de probabilidade $f(x)$,

$$f(x) = \alpha e^{-\alpha x}, \quad x \geq 0.$$

Proposição 1.3. Uma variável aleatória X , com distribuição de probabilidade exponencial de parâmetro α , possui média

$$E[X] = 1/\alpha.$$

Demonstração.

$$\begin{aligned} E[X = x] &= \int_0^{\infty} x\alpha e^{-\alpha x} dx; \\ E[X = x] &= -xe^{-\alpha x} \Big|_0^{\infty} - \int_0^{\infty} -e^{-\alpha x} dx; \\ E[X = x] &= 0 + \int_0^{\infty} e^{-\alpha x} dx; \\ E[X = x] &= -\frac{e^{-\alpha x}}{\alpha} \Big|_0^{\infty}; \\ E[X = x] &= -0 + \frac{e^0}{\alpha} = \frac{1}{\alpha}. \end{aligned}$$

□

Definição 1.9 (Distribuição de Poisson). A distribuição de probabilidade de Poisson possui probabilidades dadas por

$$P(X = k) = \frac{e^{-\alpha} \alpha^k}{k!}, \quad k = 0, 1, 2, \dots$$

Proposição 1.4. Uma variável aleatória com distribuição de probabilidade de Poisson, X , de parâmetro α , possui média

$$E[X] = \alpha.$$

Demonstração.

$$\begin{aligned} E[X = k] &= X_i \cdot P(X_i) = \sum_{k=0}^{\infty} k \frac{e^{-\alpha} \alpha^k}{k!}; \\ E[X = k] &= \sum_{k=1}^{\infty} \frac{e^{-\alpha} \alpha^k}{(k-1)!}; \end{aligned}$$

fazendo $s = k - 1$ tem-se

$$\begin{aligned} E[X = k] &= \sum_{s=0}^{\infty} \frac{e^{-\alpha} \alpha^{s+1}}{s!}; \\ E[X = k] &= \alpha \sum_{s=0}^{\infty} \frac{e^{-\alpha} \alpha^s}{s!}. \end{aligned}$$

Desde que $\sum_{s=0}^{\infty} \frac{e^{-\alpha} \alpha^s}{s!}$ é uma soma de todas as possibilidades, então

$$E[X = k] = \alpha \sum_{s=0}^{\infty} \frac{e^{-\alpha} \alpha^s}{s!} = \alpha \cdot 1 = \alpha.$$

□

2 Álgebra Linear

Os livros de referência deste capítulo são o de Coelho (9) e o de Dante (10). Para o estudo das cadeias de Markov, é fundamental ter conhecimento de conceitos da Álgebra Linear, principalmente de vetores e matrizes e as operações definidas nos espaços vetoriais.

2.1 Vetores e matrizes

Definição 2.1 (Produto matricial). Seja $A_{m \times n}$ uma matriz e $B_{n \times p}$ também uma matriz, define-se o **produto matricial** de A por B do seguinte modo

$$A \cdot B = C$$

$$[C]_{ik} = c_{ik} = \sum_{j=1}^n a_{ij} \cdot b_{jk},$$

com $[A]_{ij} = a_{ij}$ e $[B]_{jk} = b_{jk}$.

Observação 2.1. Perceba da definição acima que a matriz resultante $C = A \cdot B$ tem o número de linhas de A , e o mesmo número de colunas de B . Ou seja, C é uma matriz $C_{m \times p}$. Perceba também que o produto matricial requer que o número de colunas de A seja igual ao de linhas de B . O produto matricial não é comutativo.

Um vetor coluna $v \in \mathbb{R}^m$ pode ser visto como uma matriz $M_{m \times n}$ cuja primeira coluna tem as entradas de v , e todas as outras colunas são nulas, como segue

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} := \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ v_m & 0 & \cdots & 0 \end{pmatrix}.$$

Um vetor linha v , do mesmo modo, pode ser entendido como uma matriz cuja primeira linha é o vetor v , e os demais elementos são nulos,

$$v = (v_1, \dots, v_m) := \begin{pmatrix} v_1 & \cdots & v_m \\ 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Teorema 2.1. Dadas duas matrizes $A_{m \times n}$ e $B_{n \times p}$, tem-se que

$$(A \cdot B)^\dagger = B^\dagger \cdot A^\dagger,$$

onde \dagger sobrescrito denota a transposta da matriz.

O conjunto das matrizes $M_{n \times n}$ possui um elemento identidade, denotado neste trabalho por I .

Definição 2.2 (Matriz identidade). As matrizes quadradas $M_{n \times n}$ possuem um elemento identidade $I_{n \times n}$ tal que I tem todas as entradas da diagonal principal iguais a 1, e as demais são nulas. Por exemplo, tem-se as matrizes $I_{2 \times 2}$ e $I_{3 \times 3}$, respectivamente dadas por

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Definição 2.3 (Multiplicação matriz por escalar). Define-se multiplicação de uma matriz A por um escalar α assim

$$[A]_{ij} = a_{ij} \quad \text{é equivalente a} \quad [\alpha A]_{ij} = \alpha a_{ij}.$$

Definição 2.4 (Autovalores e autovetores). Um vetor \mathbf{x} é um autovetor associado ao autovalor λ , de uma matriz A , se

$$\mathbf{x}A = \lambda\mathbf{x}, \quad \lambda \text{ é um escalar.}$$

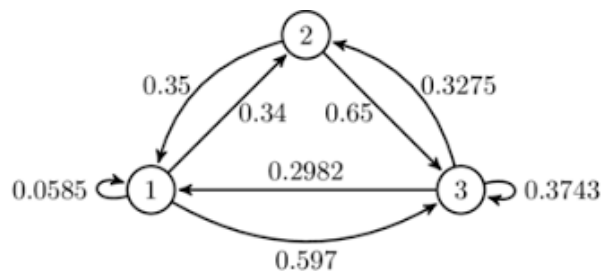
Parte II

Cadeias de Markov

3 Introdução às cadeias de Markov

Cadeias de Markov se tratam de uma espécie de processo estocástico sem memória, isto é, a probabilidade de chegar a um estado j da cadeia só depende de seu estado atual i , ou seja, os próximos estados da cadeia não dependem dos estados passados, somente do estado presente. Uma forma de interpretar um processo de cadeia de Markov é vendo-o como um caminho aleatório sobre um grafo, com pesos nas arestas de acordo com a dissertação de Santos (11). Veja a figura 2. Com pesos nas arestas significa que a probabilidade de ir por uma aresta pode ser diferente das outras arestas. Na figura, temos os estados da cadeia, 1, 2 e 3, e as arestas com as respectivas probabilidades. A probabilidade associada à transição de um estado i a um estado j , no grafo, está associada à uma aresta; essa probabilidade deve ser fixa para poder caracterizar uma cadeia de Markov. Tome como exemplo uma variável aleatória $X^{(t)}$ que é a condição climática no tempo t e X pode assumir os três valores seguintes: $S = \{\text{ensolarado, nublado, chuvoso}\}$. Neste caso, S é chamado **espaço de estados** da cadeia. Supondo que a condição climática no dia seguinte depende somente de seu estado hoje, tem-se uma cadeia de Markov, com variável aleatória X , espaço de estados S e transição de estados para o tempo $t = \{0, 1, 2, \dots\}$. Como t é uma variável discreta, no caso o dia, chamamos o processo de **cadeia de Markov com tempo discreto**. Além disso, como foi mencionado acima, a cada transição de um estado i a um estado j está associada uma probabilidade, denote-se esta probabilidade por P_{ij} . É bastante útil organizar as probabilidades de transição em uma matriz, a qual será chamada **matriz de probabilidades de transição** P .

Figura 2 – Representação em forma de grafo.



Fonte: Wikipédia, <https://pt.m.wikipedia.org/wiki/Cadeias_de_Markov>

3.1 Exemplos

Veja um exemplo de uma cadeia de Markov com 3 estados. Tem-se a seguinte matriz de probabilidades de transição P :

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} \quad (3.1)$$

Nesta matriz, cada probabilidade P_{ij} está associada a uma transição de um estado presente i , a um estado seguinte j . Considerando, novamente, o grafo da figura 2, tem-se por exemplo que $P_{12} = 0.34$, ao passo que $P_{22} = 0$ já que a aresta ligando o estado 2 ao estado 2 não existe. Portanto, quando não há aresta conectando dois estados i, j , define-se $P_{ij} = 0$.

Sendo uma cadeia de Markov atrelada ao transcorrer do tempo, denote-se uma variável aleatória de uma cadeia de Markov $X^{(t)}$ como sendo o estado da cadeia no tempo t . Desta forma temos uma sequência de estados, ou também, uma sequência de variáveis aleatórias $X = \{X^{(0)}, X^{(1)}, \dots, X^{(n)}\}$. Neste contexto, de tempo discreto, a matriz de probabilidades de transição P , dada acima, é dita matriz de probabilidades de transição **1 passo**, por estar relacionada à transição em um tempo discreto. Por definição, a matriz de probabilidades 0 passo, denotada por $P^{(0)}$, é uma matriz identidade I , definida em 2.2. Isto porque $P_{ii}^{(0)} = 1$ e $P_{ij}^{(0)} = 0$, com $i \neq j$, se não há tempo para haver uma transição. Assim, $P_{ij}^{(0)} = \delta_{ij}$. Veja um exemplo retirado de Ching et al (1).

3.1.1 Problema de mercado

Suponha que em uma cidade existam duas lojas, a loja **Bem-vindo** e a loja **Disponha**. Suponha também que um comprador da loja Bem-vindo tem uma probabilidade α de fazer sua próxima compra em Bem-vindo, e uma probabilidade $1 - \alpha$ de fazê-la em Disponha, conforme o teorema 1.1. Analogamente, vamos supor que um comprador de Disponha tenha uma probabilidade β de fazer sua próxima compra em Disponha, e uma probabilidade $1 - \beta$ de fazê-la em Bem-vindo. Vemos então que onde o cliente faz sua t -ésima compra é uma variável aleatória, a qual denotemos por $X^{(t)}$. Tem-se, então, a seguinte matriz de probabilidades de transição:

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \beta & \beta \end{pmatrix}.$$

Como se nota da matriz 3.1, podemos relacionar um estado a um número natural; um cliente em Bem-vindo está em estado 1, enquanto um cliente em Disponha está no estado 2; assim, facilmente podemos relacionar uma transição entre dois estados a uma linha e uma coluna da matriz. A probabilidade de sair do estado i e chegar ao estado j é a probabilidade localizada na linha i e coluna j da matriz de transição P .

Exemplo 3. Qual a probabilidade de um cliente fazer uma compra em Bem-vindo, e em seguida em Disponha?

Solução: O que procuramos é a probabilidade $P(\{X^{(1)} = 1\} \cap \{X^{(2)} = 2\})$. Se o cliente fez sua última compra em Bem-vindo, ou seja, o seu estado atual é 1, então a probabilidade de fazer sua próxima compra em Bem-vindo é

$$P(X^{(1)} = 1 \mid X^{(0)} = 1) = \alpha.$$

Por outro lado, se $X^{(0)} = 2$, temos a probabilidade

$$P(X^{(1)} = 1 \mid X^{(0)} = 2) = 1 - \beta.$$

Sabendo que a probabilidade de $X^{(1)} = 1$ é

$$\begin{aligned} P(X^{(1)} = 1) &= P(X^{(1)} = 1 \mid X^{(0)} = 1) + P(X^{(1)} = 1 \mid X^{(0)} = 2) \\ &= \alpha + (1 - \beta), \end{aligned}$$

e, utilizando o teorema da probabilidade condicional, tem-se que

$$\begin{aligned} P(\{X^{(1)} = 1\} \cap \{X^{(2)} = 2\}) &= P(X^{(1)} = 1) \cdot P(X^{(2)} = 2 \mid X^{(1)} = 1) \\ &= (1 + \alpha - \beta)(1 - \alpha). \end{aligned}$$

Portanto, a solução é

$$P(\{X^{(1)} = 1\} \cap \{X^{(2)} = 2\}) = 1 - \alpha(\alpha - \beta) - \beta$$

◁

No exemplo anterior calculou-se a probabilidade de uma compra em uma loja seguida de uma compra na concorrente; para tanto, considerou-se o estado anterior, que é a **propriedade markoviana**, característica do processo de Markov, a qual define que um estado futuro depende somente do estado presente. Assim é possível fazer previsões sem conhecer os estados passados, bastando conhecer o estado presente. Sabendo que o processo utiliza probabilidades condicionadas pelo estado atual, o mesmo deve ser incorporado ao problema, uma distribuição de probabilidades iniciais.

Deste modo, se $X^{(0)} = (x_1, x_2)$, $x_1 + x_2 = 1$, é a proporção de clientes para as lojas 1 e 2 em algum momento, e P é a matriz de probabilidades, então as probabilidades no tempo discreto seguinte são dadas por

$$X^{(0)} \cdot P.$$

De fato, veja que as probabilidades no tempo seguinte estão corretas. A probabilidade de estar em estado j é condicionada pelo estado anterior i , deste modo, pelo produto matricial, definido em 2.1, tem-se

$$x_{1j}^{(2)} = x_{1h}^{(1)} \cdot P_{hj}.$$

Sabendo que $X^{(1)}$, $X^{(2)}$ são vetores linha, sejam denotados apenas por

$$X^{(2)} = (x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)}) \quad \text{e} \quad X^{(1)} = (x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}),$$

tem-se então

$$x_j^{(2)} = \sum_{h=1}^n x_h^{(1)} \cdot P_{hj}.$$

Logo, a proporção de clientes pertencentes à loja j , no tempo seguinte a t , é

$$\begin{aligned} x_j^{(t+1)} &= \sum_{h=1}^n x_h^{(t)} \cdot P_{hj} \\ &= \sum_{h=1}^n P(X^{(t)} = h) \cdot P(X^{(t+1)} = j \mid X^{(t)} = h). \end{aligned}$$

Ao utilizar o teorema 1.2 da probabilidade condicional, a equação anterior pode ser reescrita como segue

$$\begin{aligned} &\sum_{h=1}^n P(\{X^{(t)} = h\} \cap \{X^{(t+1)} = j\}) \\ &= x_j^{(t+1)} = [X^{(t+1)}]_j, \end{aligned}$$

onde $[X^{(t+1)}]_j$ é a j -ésima componente do vetor $X^{(t+1)}$.

Ou seja, dado um vetor de probabilidades iniciais, $X^{(0)}$ por exemplo, a probabilidade de um estado específico j no tempo seguinte é a soma das probabilidades de todas as combinações possíveis que levam ao estado j no tempo seguinte. Usando o **princípio de indução finita**, este resultado pode ser estendido a qualquer número inteiro positivo de tempos discretos.

Sendo assim,

$$X^{(0)} \cdot P = X^{(1)},$$

e

$$X^{(0)} \cdot P^n = X^{(n)}.$$

Um pouco além disso, tem-se que

$$[P^n]_{ij} = P_{ij}^{(n)},$$

onde $P_{ij}^{(n)}$ é a probabilidade de, estando em estado i no tempo t , estar no estado j no tempo discreto $t + n$. Do que foi exposto, segue a

Proposição 3.1. Se P é uma matriz de probabilidades de transição 1 passo, então a potência com relação ao produto matricial é $P^n = P^{(n)}$, onde $P^{(n)}$ denota a matriz de probabilidades de transição n passos.

Observação 3.1. Foi utilizada a noção de **distribuição de probabilidade**, a qual é razoavelmente intuitiva para dispensar uma definição formal; todavia é oportuno notar que, se $X = (x_1, x_2, \dots, x_n)$ é um vetor distribuição de probabilidade, então $\sum_{h=1}^n x_h = 1$.

4 Construção do modelo de Markov

Dado um fenômeno que é objeto de estudo, e é um suposto processo de Markov, é possível obter as probabilidades de transição mediante observação das transições de estados. Neste contexto, tem-se as frequências F_{ij} de transições de um estado i a um estado j . Em primeiro lugar, se observa a sequência de estados $X = X^{(0)}, X^{(1)}, X^{(2)}, \dots$, isto é, a sequência de valores assumidos pela variável aleatória X com o decorrer do tempo, no experimento aleatório estudado. Suponha que a cadeia possua n estados, suponha também que foram observados 20 valores para a variável X , o próximo passo é contar cada transição de estados $X^{(t)} = i, X^{(t+1)} = j; i, j = 1, 2, \dots, n$. Assim, se destes dados observados 3 estados i 's foram seguidos, cada um, por um estado j , então $F_{ij} = 3$. Note que, se o número total de estados é n , então $F_{ij} = 0$ se $i > n$ ou $j > n$.

4.1 Algoritmo para obterem-se as probabilidades de transições

Segue aqui um algoritmo para se obter as probabilidades de transição.

$$P_{ij} = \begin{cases} \frac{F_{ij}}{\sum_{j=1}^n F_{ij}} & \text{se } \sum_{j=1}^n F_{ij} > 0 \\ 0 & \text{se } \sum_{j=1}^n F_{ij} = 0. \end{cases} \quad (4.1)$$

Veja a seguir um exemplo com alguns dados empíricos.

4.1.1 Exemplo com dados empíricos

Suponha que foram observados os seguintes dados da variável aleatória X durante um período de 20 tempos discretos.

$$X = 1, 1, 3, 2, 2, 1, 2, 2, 3, 2, 1, 3, 1, 2, 1, 3, 3, 3, 2, 1.$$

Comece por construir a matriz F com as frequências das transições; o número da linha indique o estado presente, e o da coluna, o seguinte.

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}. \quad (4.2)$$

Portanto, a matriz P no exemplo dado é

$$P = \begin{pmatrix} 1/6 & 1/3 & 1/2 \\ 4/7 & 2/7 & 1/7 \\ 1/6 & 1/2 & 1/3 \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix}.$$

5 Distribuição estacionária

A distribuição estacionária é uma propriedade interessante que algumas cadeias de Markov possuem.

Definição 5.1. Define-se distribuição estacionária de uma matriz P um vetor $\pi = (\pi_1, \pi_2, \dots, \pi_k)$ que satisfaz as duas seguintes condições:

(i)

$$\pi_i \geq 0 \quad \text{e} \quad \sum_{i=1}^k \pi_i = 1$$

(ii)

$$\pi P = \pi, \quad \text{ou seja,} \quad \sum_{i=1}^k \pi_i P_{ij} = \pi_j \quad (5.1)$$

Uma cadeia de Markov **irredutível** e com espaço de estados finito tem, no máximo, um único vetor distribuição estacionária. Mas, em alguns casos, ela pode não ter nenhum. Para uma cadeia de Markov finita (com espaço de estados finito) e irredutível ter uma distribuição estacionária, é suficiente que ela seja **aperiódica**. Estes dois termos são definidos a seguir.

Definição 5.2 (Irredutibilidade). Diz-se que uma cadeia de Markov é irredutível se existirem inteiros positivos n tais que $P_{ij}^{(n)} > 0$ para todo i, j pertencente ao espaço de estados S .

Definição 5.3 (Periodicidade). Um estado i é dito periódico, de período d , se d for finito e $P_{ii}^{(n)} > 0$ sempre que d divide n e $P_{ii}^{(n)} = 0$ caso contrário, e d for o maior inteiro com esta propriedade. Se $d = 1$, diz-se que i é aperiódico.

Se uma cadeia de Markov finita for irredutível e aperiódica, então ela possui uma única distribuição estacionária π ; pode ser provado que, para qualquer distribuição inicial X , $\lim_{n \rightarrow \infty} X P^n = \pi$. Esta propriedade é interessante, pois ela diz que o processo perde a memória com o tempo; isto significa dizer que a distribuição inicial perde a influência sobre o resultado final se este acontece depois de um longo intervalo de tempo. Todavia, há casos em que é possível reverter a cadeia, e assim recuperar a memória dos estados iniciais. Um tratamento sobre métodos para recuperar a distribuição de estados inicial da cadeia de Markov pode ser encontrado no livro de Meyn e Tweedie (12) por exemplo. Por isso, tem-se o seguinte teorema, retirado de Ching et al (1), p. 16.

Teorema 5.1. Se P for a matriz de probabilidades de uma cadeia de Markov finita, irredutível e aperiódica, então, para qualquer distribuição inicial X , tem-se:

$$\lim_{n \rightarrow \infty} X P^n = \pi \quad \iff \quad \lim_{n \rightarrow \infty} \|X P^n - \pi\| = 0.$$

A demonstração deste teorema requer a apresentação de diversos outros conceitos pouco conhecidos, de tal modo que um tratamento minucioso ampliaria muito o volume deste trabalho. Contudo, o resultado deste teorema está bem difundido na literatura de Probabilidade, e pode ser checado em diversas fontes. A demonstração deste teorema pode ser encontrado no livro de Meyer (13), p. 674, 687-694. Esta última fonte utiliza, de forma equivalente, o conceito de primitividade em lugar de aperiodicidade, para demonstrar o teorema acima. Periodicidade e primitividade são relacionados, ver a p. 694.

Observação 5.1. A distribuição estacionária é uma das medidas de interesse, dentre outras relacionadas a uma cadeia de Markov, importantes aplicações são obtidas a partir dela. Um conhecimento prévio do comportamento futuro de um processo permite um melhor planejamento de atividades econômicas, de modo a evitar uma utilização de recursos desproporcional à demanda real; como por exemplo o tamanho do quadro de funcionários em uma agência bancária.

5.1 Exemplo de convergência

Considere a seguinte matriz $P_{3 \times 3}$:

$$P = \begin{pmatrix} 0.2 & 0.1 & 0.7 \\ 0.4 & 0.25 & 0.35 \\ 0.12 & 0.38 & 0.5 \end{pmatrix}.$$

Visivelmente esta cadeia é irredutível, porque $P_{ij} > 0$ para todo $i, j \in S$. Também é aperiódica pois $P_{ii} > 0$ para todo $i \in S$. Considere então a distribuição de probabilidade inicial $X = (1, 0, 0)$; utilizou-se um auxílio computacional para se obter as distribuições ao longo do tempo, observe as figuras 3 e 4.

Nas imagens, se percebe a sintaxe do software *Mathematica*, onde “%” representa o **resultado anterior**, e “.P” representa o **produto matricial com P à direita**. O vetor $X = (1, 0, 0)$ é então sucessivamente multiplicado, à direita, pela matriz P . Como se nota nos resultados, as distribuições convergem para o vetor $\pi = (0.216458, 0.282648, 0.500894)$. π é o vetor distribuição estacionária da matriz P .

5.2 Sistemas Lineares

Um **sistema linear** pode ser escrito como uma equação do tipo

$$\mathbf{x}A = \mathbf{b},$$

onde A é uma matriz dada, \mathbf{b} é um vetor linha dado, e \mathbf{x} também é um vetor linha, cujas entradas são incógnitas em primeiro grau, que se desejam encontrar. Conforme a equação 5.1, vê-se que o problema de encontrar a distribuição estacionária consiste em resolver um sistema linear,

Figura 3 – Sequência gerada no software *Mathematica*.

```

MatrixForm[P]

$$\begin{pmatrix} 0.2 & 0.1 & 0.7 \\ 0.4 & 0.25 & 0.35 \\ 0.12 & 0.38 & 0.5 \end{pmatrix}$$

X = {1, 0, 0}
{1, 0, 0}
X.P
{0.2, 0.1, 0.7}
%.P
{0.164, 0.311, 0.525}
%.P
{0.2202, 0.29365, 0.48615}
%.P
{0.219838, 0.280169, 0.499992}

```

especificamente um do tipo $\mathbf{x}A = \mathbf{x}$ tal que $A \neq I$, I é a matriz identidade. Ou seja, \mathbf{x} é um autovetor de A associado ao autovalor 1, definição 2.4.

Resolver um sistema linear é uma tarefa complexa para um ser humano, na melhores hipóteses. Para uma aplicação de cadeias de Markov, faz-se necessário a utilização de um auxílio computacional no momento de resolver sistemas lineares. Serão abordados aqui, portanto, dois métodos numéricos iterativos para solução de sistemas lineares, retirados de (14).

5.2.1 Método de Jacobi

Este é um método numérico iterativo que fornece uma solução $\mathbf{x}^{(k)}$, sucessivamente mais próxima de \mathbf{x} a cada iteração, que resolve a equação $\mathbf{x}A = \mathbf{b}$. Isso significa que, se $\mathbf{x}^{(k)}$ é a aproximação da k -ésima iteração, há uma norma vetorial $\|\cdot\|$ tal que $\lim_{k \rightarrow \infty} \|\mathbf{x} - \mathbf{x}^{(k)}\| = 0$. O método é dado pelo seguinte

Teorema 5.2. Dada uma aproximação inicial $\mathbf{x}^{(0)}$ para a solução exata \mathbf{x} , do sistema $\mathbf{x}A = \mathbf{b}$,

Figura 4 – Sequência gerada no software *Mathematica*.

```
%P  
{0.216034, 0.282023, 0.501942}  
  
%P  
{0.216249, 0.282847, 0.500903}  
  
%P  
{0.216497, 0.28268, 0.500823}  
  
%P  
{0.21647, 0.282632, 0.500897}  
  
%P  
{0.216455, 0.282646, 0.500899}  
  
%P  
{0.216457, 0.282649, 0.500894}  
  
%P  
{0.216458, 0.282648, 0.500894}  
  
%P  
{0.216458, 0.282648, 0.500894}
```

tem-se o seguinte algoritmo

$$x_j^{(1)} = - \left(\sum_{\substack{i=1 \\ i \neq j}}^n \frac{a_{ij} x_i^{(0)}}{a_{jj}} \right) + \frac{b_j}{a_{jj}},$$

e, de uma forma geral, tem-se

$$x_j^{(k+1)} = - \left(\sum_{\substack{i=1 \\ i \neq j}}^n \frac{a_{ij} x_i^{(k)}}{a_{jj}} \right) + \frac{b_j}{a_{jj}}.$$

Demonstração. Dado um produto matricial $\mathbf{x}A = \mathbf{b}$, pela definição de produto matricial 2.1, tem-se que

$$[\mathbf{b}]_j = b_j = \sum_{i=1}^n a_{ij} x_i = \sum_{\substack{i=1 \\ i \neq j}}^n a_{ij} x_i + a_{jj} x_j;$$

de forma equivalente,

$$a_{jj} x_j = - \left(\sum_{\substack{i=1 \\ i \neq j}}^n a_{ij} x_i \right) + b_j,$$

e finalmente, dividindo-se ambos os lados desta última equação por a_{jj} , obtém-se a equação

$$x_j = - \left(\sum_{\substack{i=1 \\ i \neq j}}^n \frac{a_{ij} x_i}{a_{jj}} \right) + \frac{b_j}{a_{jj}}.$$

A convergência para a solução vem do fato de que o vetor \mathbf{x}^{k+1} pode ser escrito como

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}T + \mathbf{c},$$

onde \mathbf{c} é um valor que não depende da aproximação escolhida. Assim, o erro na aproximação $\mathbf{x}^{(k)}$ diminui quando ela é combinada com \mathbf{c} para obter uma aproximação $\mathbf{x}^{(k+1)}$. \square

Observação 5.2. Dado o fato de que a_{jj} é um denominador no método de Jacobi, é necessário que $a_{jj} \neq 0$ para todo j . Isto pode ser obtido, caso o sistema seja não-singular, com uma permutação de colunas, ou linhas se tratarem-se de vetores coluna \mathbf{x} e \mathbf{b} , cf. (14).

5.2.2 Método de Gauss-Seidel

Este é um outro método iterativo muito interessante, por ser semelhante ao método de Jacobi porém, com um aprimoramento. Enquanto que no método de Jacobi se utilizam as componentes de $\mathbf{x}^{(k)}$ para obter as componentes de $\mathbf{x}^{(k+1)}$, o método de Gauss-Seidel utiliza as componentes $x_i^{(k+1)}$ para obter as componentes $x_j^{(k+1)}$, $i < j$. Matematicamente é

$$x_j^{(k+1)} = - \left(\sum_{i=1}^{j-1} \frac{a_{ij} x_i^{(k+1)}}{a_{jj}} + \sum_{i=j+1}^n \frac{a_{ij} x_i^{(k)}}{a_{jj}} \right) + \frac{b_j}{a_{jj}}. \quad (5.2)$$

Isto acelera a convergência, reduzindo desta forma o tempo de computação, uma vez que o método de Gauss-Seidel utiliza valores mais apurados que o método de Jacobi, sempre que estes estão disponíveis, em cada iteração.

6 Cadeia de Markov com tempo contínuo de dois estados

Até aqui só foram tratadas cadeias de Markov com tempo sendo uma variável discreta. Esta categoria bem específica de cadeias, tratadas neste capítulo, tem em vista uma aplicação; está relacionada a atendimentos de diversas espécies, nos setores de vendas de bens e de serviços, em estabelecimentos ou com atendentes de telefone.

6.1 Processo de Poisson

O processo de Poisson é apropriado para descrever fenômenos onde há uma contagem de eventos durante um período de tempo.

Exemplo 4. O número de vezes que um telefone toca, em cinco minutos.

De acordo com o livro de Ross (15), p. 305, tem-se a seguinte definição para o processo de Poisson:

Definição 5.3 O processo de contagem $\{N(t), t \geq 0\}$ é dito ser um processo de Poisson tendo parâmetro λ , $\lambda > 0$, se

- (i) $N(0) = 0$.
- (ii) o processo tem incrementos estacionários e independentes.
- (iii) $P\{N(h) = 1\} = \lambda h + o(h)$.
- (iv) $P\{N(h) \geq 2\} = o(h)$.

(Tradução nossa, número da definição da supracitada obra). A afirmação de que os incrementos são estacionários significa que, para um intervalo de tempo fixo h , a probabilidade de uma contagem $N(h) = k$, ou seja, $P\{N(h) = k\}$, é sempre a mesma para o mesmo k ; analogamente, incrementos independentes significam que a contagem não depende do momento em que é feita a contagem, por outro lado, depende do intervalo de tempo que é feita a contagem. Portanto, o item (ii) é equivalente a

$$(ii') P\{N(h+t) - N(t)\} = P\{N(h)\} \text{ para todo } h, t \geq 0.$$

Desta definição pode ser obtida a fórmula da distribuição de Poisson, o que será feito a seguir com o propósito de desenvolver uma aplicação prática. A demonstração é baseada principalmente no trabalho de Ching et al (1).

6.1.1 Distribuição exponencial

O propósito desta seção é deduzir a fórmula da distribuição de Poisson, e desta subseção, a relação do processo de Poisson com a distribuição exponencial. Para tanto se fará uso da definição do processo de Poisson dada na citação de Ross.

Comece por definir a função $o(h)$, nos itens (iii) e (iv) da citação de Ross. Na literatura matemática, uma função denotada por $o(h)$, ou uma outra notação assemelhada, costuma ter a seguinte característica

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0.$$

Em outras palavras, o item (iii) da definição citada diz que, para $h \rightarrow 0^+$, a probabilidade de uma contagem $N(h) = 1$ é linear; por outro lado, de acordo com o item (iv), da mesma citação direta, a probabilidade de ocorrência de uma contagem $N(h) \geq 2$ é menos do que linear se $h \rightarrow 0$.

Observação 6.1. Evite-se pensar erroneamente que $\lambda h + o(h)$ é a probabilidade de ocorrer uma contagem de “pelo menos 1”, e que λh é a probabilidade de ocorrer exatamente 1, baseando-se em que $o(h)$ é a probabilidade de ocorrer uma contagem maior do que 1. O contexto exposto é outro. Na realidade, a função $o(h)$ que aparece nos itens (iii) e (iv) do excerto de Ross não são exatamente a mesma função; por outro lado, elas são denotadas pelo mesmo símbolo citado por possuírem ambas a propriedade $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$. Uma definição mais geral do processo de Poisson é aquela encontrada no livro de Knill (7), páginas 319, 320.

Agora, denote $P_n(t)$ a probabilidade de uma contagem igual a n até o tempo t . Um exemplo disto é a chegada de n consumidores a uma loja até o tempo t . Considere também que o processo de chegada de consumidores, isto é, o processo de contagem, é um processo de Poisson com parâmetro λ . Desta forma tem-se a seguinte equação

$$P_n(t + \delta t) = \underbrace{(1 - \lambda \delta t - o(\delta t))P_n(t)}_{(a)} + \underbrace{(\lambda \delta t + o(\delta t))P_{n-1}(t)}_{(b)} + \underbrace{o(\delta t)}_{(c)}. \quad (6.1)$$

- a) contagem de n no intervalo $[0, t]$, nenhum em $[t, t + \delta t]$;
- b) contagem de $n - 1$ no intervalo $[0, t]$, e 1 em $[t, t + \delta t]$;
- c) contagem maior do que 1 em $[t, t + \delta t]$.

Um estudo da equação 6.1 leva à conclusão de que há três casos possíveis para $P_n(t + \delta t)$, os três são (a), (b) e (c), mutuamente excludentes, por isso a soma das probabilidades, de acordo com a definição 1.4 item (c).

Observação 6.2. Note que na equação 6.1 foi utilizado o teorema 1.2 e a definição 1.5.

Prosseguindo os cálculos, tem-se

$$P_n(t + \delta t) = P_n(t) - \lambda \delta t P_n(t) - o(\delta t) P_n(t) + \lambda \delta t P_{n-1}(t) + o(\delta t) P_{n-1}(t) + o(\delta t)$$

$$\begin{aligned}
&= P_n(t) - \lambda\delta t P_n(t) + \lambda\delta t P_{n-1}(t) + (-P_n(t) + P_{n-1}(t) + 1)o(\delta t) \\
&= P_n(t + \delta t) \\
\iff P_n(t + \delta t) - P_n(t) &= -\lambda\delta t P_n(t) + \lambda\delta t P_{n-1}(t) + (-P_n(t) + P_{n-1}(t) + 1)o(\delta t) \\
\iff \frac{P_n(t + \delta t) - P_n(t)}{\delta t} &= -\lambda P_n(t) + \lambda P_{n-1}(t) + (-P_n(t) + P_{n-1}(t) + 1)\frac{o(\delta t)}{\delta t} \\
\iff \lim_{\delta t \rightarrow 0} \frac{P_n(t + \delta t) - P_n(t)}{\delta t} &= -\lambda P_n(t) + \lambda P_{n-1}(t) + \lim_{\delta t \rightarrow 0} (-P_n(t) + P_{n-1}(t) + 1)\frac{o(\delta t)}{\delta t} \\
&= \frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t) + 0 \\
&= \frac{dP_n(t)}{dt} = -\lambda P_n(t) + \lambda P_{n-1}(t); \tag{6.2} \\
&\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \lambda P_{-1}(t).
\end{aligned}$$

Dado que um processo de Poisson é um processo de contagem do número de vezes que um determinado evento de interesse acontece, pode-se supor que $P_{-1}(t) = 0$, conforme feito no supracitado trabalho de Ching et al (1). Desta forma, tem-se a seguinte equação diferencial

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t);$$

esta é uma equação diferencial bem simples e seu conjunto solução, conforme descrito em Guidorizzi (16), p. 286, é

$$\frac{dy}{dx} = \alpha y \iff y = ke^{\alpha x}.$$

A hipótese $N(0) = 0 \implies P_0(0) = 1$ conduz a um problema de valor inicial

$$\begin{aligned}
\frac{dP_0(t)}{dt} &= -\lambda P_0(t) \quad \text{com} \quad P_0(0) = 1. \\
P_0(0) = ke^{-\lambda 0} = 1 &\iff k = 1 \quad \text{portanto,} \quad P_0(t) = e^{-\lambda t}.
\end{aligned}$$

Isto significa que $e^{-\lambda t}$ é a probabilidade de que nenhum evento ocorra no intervalo de tempo t . Então $1 - e^{-\lambda t}$ é a probabilidade de que o primeiro evento ocorra no tempo t .

Ao se derivar uma probabilidade se obtém a função densidade de probabilidade.

$$\frac{d}{dt}(1 - e^{-\lambda t}) = \lambda e^{-\lambda t} \tag{6.3}$$

O membro direito de 6.3 é a expressão algébrica da distribuição de probabilidade exponencial, de parâmetro λ , conforme definida em 1.8. Isto significa que o tempo de espera de chegada de um consumidor a uma loja segue uma distribuição de probabilidade exponencial, de parâmetro λ .

6.1.2 Fórmula de Poisson

Nesta subsecção será demonstrado que a distribuição de Poisson relacionada a um processo de Poisson, de parâmetro λ , é dada por

$$P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!} \quad n = 0, 1, 2, \dots$$

Mostre-se $P_1(t)$, e a partir daí os demais casos serão demonstrados por indução.

Primeiro note que 6.2 é uma equação diferencial ordinária linear em $P_n(t)$, de primeira ordem; o livro de Guidorizzi (16), p. 456, fornece uma fórmula, e sua demonstração, para calcular equações deste tipo. A fórmula para uma equação do tipo

$$\frac{dx}{dt} = xg(t) + f(t) \quad \acute{e} \quad x = ke^{\int g(t)dt} + e^{\int g(t)dt} \int f(t)e^{-\int g(t)dt} dt, \quad (6.4)$$

$k \in \mathbb{R}$ é uma constante.

Proposição 6.1. Um processo de Poisson de parâmetro λ está relacionado a uma distribuição de Poisson de parâmetro λt . Isto é, $P_n(t)$, solução da equação 6.2, é

$$P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}.$$

Demonstração. Utilize-se o método de indução finita. Primeiro especifique-se os valores utilizados na fórmula 6.4. $x = P_1(t)$, $g(t) = -\lambda$ e $f(t) = \lambda e^{-\lambda t}$. Portanto,

$$P_1(t) = ke^{-\int \lambda dt} + e^{-\int \lambda dt} \int \lambda e^{-\lambda t} e^{-\int -\lambda dt} dt,$$

$$P_1(t) = ke^{-\lambda t} + e^{-\lambda t} \int \lambda e^{-\lambda t} e^{\lambda t} dt,$$

$$P_1(t) = ke^{-\lambda t} + e^{-\lambda t} \int \lambda dt,$$

$$P_1(t) = ke^{-\lambda t} + e^{-\lambda t} \lambda t. \quad (6.5)$$

$$P_1(0) = 0 = ke^{-\lambda \cdot 0} + e^{-\lambda \cdot 0} \lambda \cdot 0,$$

$$P_1(0) = k \cdot 1 + 0 = 0.$$

Portanto, $k = 0$. Então, substituindo k em 6.5, obtém-se

$$P_1(t) = e^{-\lambda t} \lambda t.$$

Agora, se fará uso do princípio de indução finita para demonstrar os demais casos. Suponha que $P_n(t) = \frac{e^{-\lambda t} (\lambda t)^n}{n!}$, prove-se que $P_{n+1}(t) = \frac{e^{-\lambda t} (\lambda t)^{n+1}}{(n+1)!}$. Será utilizada a fórmula 6.4, e também o princípio de indução, para se resolver a equação 6.2, para todo $n \in \mathbb{N}$.

Aplicando a fórmula, tem-se os seguintes valores

$$\begin{aligned}
 x &= P_{n+1}(t), \quad g(t) = -\lambda, \quad f(t) = \lambda \frac{e^{-\lambda t} (\lambda t)^n}{n!}. \\
 P_{n+1}(t) &= k e^{\int -\lambda dt} + e^{\int -\lambda dt} \int \frac{\lambda e^{-\lambda t} (\lambda t)^n}{n!} e^{-\int -\lambda dt} dt; \\
 P_{n+1}(t) &= k e^{-\lambda t} + e^{-\lambda t} \int \frac{\lambda e^{-\lambda t} (\lambda t)^n}{n!} e^{\lambda t} dt; \\
 P_{n+1}(t) &= k e^{-\lambda t} + e^{-\lambda t} \int \frac{\lambda (\lambda t)^n}{n!} dt.
 \end{aligned} \tag{6.6}$$

Para resolver $\int \frac{\lambda (\lambda t)^n}{n!} dt$, faz-se a seguinte mudança de variável,

$$\lambda t = u, \quad \text{o que implica} \quad \lambda dt = du \iff dt = \frac{du}{\lambda}.$$

Portanto

$$\begin{aligned}
 \int \frac{\lambda (\lambda t)^n}{n!} dt &= \int \frac{\lambda u^n}{n!} \frac{du}{\lambda} = \frac{1}{n!} \int u^n du = \frac{1}{n!} \frac{u^{n+1}}{n+1} = \frac{u^{n+1}}{(n+1)!} \\
 &= \frac{(\lambda t)^{n+1}}{(n+1)!}.
 \end{aligned} \tag{6.7}$$

Substituindo 6.7 em 6.6, tem-se

$$P_{n+1}(t) = k e^{-\lambda t} + e^{-\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!}. \tag{6.8}$$

$$P_{n+1}(0) = 0 = k e^{-\lambda \cdot 0} + \frac{e^{-\lambda \cdot 0} (\lambda \cdot 0)^{n+1}}{(n+1)!};$$

$$P_{n+1}(0) = k \cdot 1 + 0 = 0,$$

portanto $k = 0$. Substituindo $k = 0$ em 6.8, tem-se que

$$P_{n+1}(t) = \frac{e^{-\lambda t} (\lambda t)^{n+1}}{(n+1)!},$$

que é o que se queria provar. □

Por tudo que foi exposto nesta seção, tem-se a seguinte

Proposição 6.2. As seguintes afirmações são equivalentes,

- a) A chegada de consumidores é um processo de Poisson de parâmetro λ ;
- b) a probabilidade de n chegadas até o tempo t é dada por

$$P\{N(t) = n\} = \frac{e^{-\lambda t} (\lambda t)^n}{n!};$$

- c) o tempo entre chegadas tem distribuição exponencial de parâmetro λ .

6.2 Uma solução para um problema prático

A seção anterior, que tratou um pouco sobre processo de Poisson, fornece cálculos que serão utilizados nesta seção para solução de um problema prático, **a otimização do tamanho de quadro de pessoal de empresas**. Para esta finalidade se fará uso da notação utilizada na seção anterior, $P_1(t)$ e $P_0(t)$, com outros significados. Imagine um atendente em uma empresa; $P_0(t)$ denote a probabilidade de o atendente estar desocupado no tempo t , e $P_1(t)$ denote a probabilidade de ele estar ocupado no tempo t . Será abordada primeiro a construção da equação de $P_0(t)$, depois $P_1(t)$.

Suponha que o processo de chegada de consumidores à loja é um processo de Poisson, de parâmetro λ ; suponha também que o tempo de atendimento de um cliente segue uma distribuição exponencial de parâmetro μ , portanto o processo de atendimentos é um processo de Poisson de parâmetro μ . Considere que $P_0(t + \delta t)$ seja a probabilidade de o atendente estar desocupado no tempo $t + \delta t$, e que $P_1(t + \delta t)$ seja a probabilidade de estar ocupado no tempo $t + \delta t$. Suponha também a não existência de filas, de forma que a probabilidade de um atendimento e a chegada de um cliente em um tempo δt é uma função $o(\delta t)$, porque $P\{N(\delta t) \geq 2\} = o(\delta t)$. Isto significa que, enquanto se atende um cliente, não há probabilidade de chegada de outro. Desta forma tem-se a seguinte equação para $P_0(t + \delta t)$.

$$P_0(t + \delta t) = \underbrace{(1 - \lambda\delta t - o(\delta t))P_0(t)}_{(a)} + \underbrace{(\mu\delta t + o(\delta t))P_1(t)}_{(b)} + \underbrace{o(\delta t)}_{(c)}. \quad (6.9)$$

- a) Probabilidade de estar desocupado no tempo t e não chegar ninguém no intervalo $[t, t + \delta t]$;
- b) probabilidade de estar ocupado no tempo t e terminar o atendimento em $[t, t + \delta t]$;
- c) probabilidade de mais de um evento, que pode ser tanto chegada de clientes como atendimentos, em $[t, t + \delta t]$.

Veja que em (a) e (b) utilizaram-se o teorema 1.2 e a definição 1.5. Perceba também que em (b) foi utilizada a relação da distribuição exponencial com o processo de Poisson, conforme afirmado na proposição 6.2; $\lambda\delta t + o(\delta t)$ é a probabilidade de chegada de uma pessoa, ou seja, $P\{N(\delta t) = 1\} = \lambda\delta t + o(\delta t)$; de forma análoga, $P\{N(\delta t) = 1\} = \mu\delta t + o(\delta t)$ é a probabilidade de saída de uma pessoa, depois de ser atendida.

Considere, agora, a equação da probabilidade de o atendente estar ocupado no tempo $t + \delta t$, de modo semelhante ao que foi feito em 6.9.

$$P_1(t + \delta t) = \underbrace{(\lambda\delta t + o(\delta t))P_0(t)}_{(a)} + \underbrace{(1 - \mu\delta t - o(\delta t))P_1(t)}_{(b)} + \underbrace{o(\delta t)}_{(c)}. \quad (6.10)$$

- a) Probabilidade de o atendente estar desocupado no tempo t e chegar uma pessoa em $[t, t + \delta t]$;

- b) probabilidade de estar ocupado no tempo t , e não terminar de atender a pessoa em $[t, t + \delta t]$;
- c) probabilidade de mais de um evento em $[t, t + \delta t]$, que pode ser tanto pessoas atendidas quanto clientes que chegam, todavia somente aqueles nos quais o atendente está ocupado no tempo $t + \delta t$.

Operando aritmeticamente a equação de $P_0(t + \delta t)$, tem-se

$$\begin{aligned}
 P_0(t + \delta t) &= P_0(t) - \lambda \delta t P_0(t) - o(\delta t) P_0(t) + \mu \delta t P_1(t) + o(\delta t) P_1(t) + o(\delta t); \\
 \iff P_0(t + \delta t) - P_0(t) &= -\lambda \delta t P_0(t) + \mu \delta t P_1(t) + (1 - P_0(t) + P_1(t)) o(\delta t); \\
 \iff \lim_{\delta t \rightarrow 0} \frac{P_0(t + \delta t) - P_0(t)}{\delta t} &= -\lambda P_0(t) + \mu P_1(t) + \lim_{\delta t \rightarrow 0} (1 - P_0(t) + P_1(t)) \frac{o(\delta t)}{\delta t}; \\
 &= \frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t).
 \end{aligned}$$

De forma semelhante se faz com $P_1(t + \delta t)$.

$$\begin{aligned}
 P_1(t + \delta t) &= (1 - \mu \delta t - o(\delta t)) P_1(t) + (\lambda \delta t + o(\delta t)) P_0(t) + o(\delta t); \\
 \iff P_1(t + \delta t) &= P_1(t) - \mu \delta t P_1(t) - o(\delta t) P_1(t) + \lambda \delta t P_0(t) + o(\delta t) P_0(t) + o(\delta t); \\
 \iff P_1(t + \delta t) - P_1(t) &= -\mu \delta t P_1(t) + \lambda \delta t P_0(t) + (1 + P_0(t) - P_1(t)) o(\delta t); \\
 \iff \lim_{\delta t \rightarrow 0} \frac{P_1(t + \delta t) - P_1(t)}{\delta t} &= -\mu P_1(t) + \lambda P_0(t) + \lim_{\delta t \rightarrow 0} (1 + P_0(t) - P_1(t)) \frac{o(\delta t)}{\delta t}; \\
 &= \frac{dP_1(t)}{dt} = -\mu P_1(t) + \lambda P_0(t).
 \end{aligned}$$

Desta forma, tem-se o seguinte sistema de equações diferenciais.

$$\begin{cases} P_0'(t) = -\lambda P_0(t) + \mu P_1(t) \\ P_1'(t) = \lambda P_0(t) - \mu P_1(t) \end{cases}$$

Com o propósito de resolver este sistema, note que

$$\begin{aligned}
 P_1(t) &= 1 - P_0(t), \quad \text{logo,} \\
 P_0'(t) &= -\lambda P_0(t) + \mu(1 - P_0(t)).
 \end{aligned}$$

Prosseguindo, tem-se que

$$\begin{aligned}
 P_0'(t) &= -\lambda P_0(t) + \mu - \mu P_0(t); \\
 P_0'(t) &= -(\lambda + \mu) P_0(t) + \mu.
 \end{aligned}$$

Foi encontrada, então, uma **equação diferencial ordinária** (EDO) linear de primeira ordem, cuja incógnita é $P_0(t)$. Para este tipo de equação, pode-se utilizar a fórmula 6.4, com

$$x = P_0(t), \quad g(t) = -(\lambda + \mu) \quad \text{e} \quad f(t) = \mu.$$

$$P_0(t) = ke^{-\int(\lambda+\mu)dt} + e^{-\int(\lambda+\mu)dt} \int \mu e^{-\int-(\lambda+\mu)dt} dt;$$

$$P_0(t) = ke^{-(\lambda+\mu)t} + e^{-(\lambda+\mu)t} \mu \int e^{(\lambda+\mu)t} dt;$$

$$P_0(t) = ke^{-(\lambda+\mu)t} + e^{-(\lambda+\mu)t} \cdot \frac{\mu e^{(\lambda+\mu)t}}{\lambda + \mu};$$

$$P_0(t) = ke^{-(\lambda+\mu)t} + \frac{\mu}{\lambda + \mu}.$$

Utilizando o valor inicial $P_1(0) = 1$, e portanto $P_0(0) = 0$, tem-se

$$P_0(0) = ke^{-(\lambda+\mu) \cdot 0} + \frac{\mu}{\lambda + \mu} = 0;$$

$$P_0(0) = ke^0 + \frac{\mu}{\lambda + \mu} = 0;$$

$$P_0(0) = k + \frac{\mu}{\lambda + \mu} = 0;$$

$$\Leftrightarrow k = -\frac{\mu}{\lambda + \mu}.$$

Portanto, a probabilidade $P_0(t)$ é dada por

$$P_0(t) = -\frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda + \mu};$$

$$P_0(t) = \frac{1}{\lambda + \mu} (\mu - \mu e^{-(\lambda+\mu)t});$$

ou, de forma equivalente,

$$P_0(t) = \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda+\mu)t}). \quad (6.11)$$

Para encontrar $P_1(t)$ basta utilizar novamente a relação $P_1(t) = 1 - P_0(t)$.

$$P_1(t) = 1 - P_0(t) = 1 - \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda+\mu)t});$$

$$P_1(t) = \frac{\lambda + \mu}{\lambda + \mu} - \frac{\mu(1 - e^{-(\lambda+\mu)t})}{\lambda + \mu};$$

$$P_1(t) = \frac{\lambda}{\lambda + \mu} - \frac{\mu(-e^{-(\lambda+\mu)t})}{\lambda + \mu};$$

$$P_1(t) = \frac{\lambda + \mu e^{-(\lambda+\mu)t}}{\lambda + \mu};$$

$$P_1(t) = \frac{1}{\lambda + \mu} (\mu e^{-(\lambda+\mu)t} + \lambda). \quad (6.12)$$

As soluções 6.11 e 6.12 são chamadas de soluções transitórias, uma vez que esta cadeia converge para uma distribuição estacionária. Ela possui um vetor distribuição estacionária

$\pi = (\pi_0, \pi_1)$, dado por

$$\begin{aligned}\pi_0 &= \lim_{t \rightarrow \infty} \frac{\mu}{\lambda + \mu} (1 - e^{-(\lambda + \mu)t}); \\ \pi_0 &= \frac{\mu}{\lambda + \mu}; \\ \pi_1 &= \lim_{t \rightarrow \infty} \frac{1}{\lambda + \mu} (\mu e^{-(\lambda + \mu)t} + \lambda); \\ \pi_1 &= \frac{\lambda}{\lambda + \mu}.\end{aligned}$$

Este modelo normalmente é apropriado para atendimentos em lojas, e em alguns casos, atendimentos por telefone, por exemplo. Todavia ele não é apropriado para algumas outras situações semelhantes, como atendimento em lojas em dias de saldão, agências bancárias ou alguns atendimentos por telefone, isto porque ele não leva em consideração a existência de filas. Em situações onde há filas, o próximo modelo é mais apropriado. A seguir, será apresentado um modelo para situações em que há uma fila à espera de atendimento, e que tenta prever o aumento ou diminuição da fila. O livro de Ross (15) denomina este modelo de **processo de nascimento e morte**, porque o mesmo modelo se aplica ao crescimento e decrescimento de populações.

7 Cadeias de Markov com tempo contínuo

No capítulo anterior foram estudadas cadeias de Markov de tempo contínuo com dois estados. Este capítulo, como sugere o título, estende o que foi estudado no anterior a um caso mais geral, o espaço de estados infinito.

7.1 Modelagem do tamanho de filas

Suponha que clientes chegam à uma agência bancária em um tempo distribuído exponencialmente, de parâmetro λ , e que clientes são atendidos também em um tempo com distribuição de probabilidade exponencial, esta com parâmetro μ . Será feito um estudo do crescimento (ou decréscimo) da fila de clientes em espera.

Em primeiro lugar, considere as **taxas de transição instantâneas** ν_i e q_{ij} . ν_i é a probabilidade em que a cadeia faz uma transição de estado, dividida pelo tempo, quando em estado i , e q_{ij} , $i \neq j$, é a probabilidade que a cadeia faz uma transição do estado i ao estado j , também dividida pelo tempo. Dado que ν_i e q_{ij} são taxas de transição instantâneas, elas dão as probabilidades de transição, quando multiplicadas pelo tempo, do estado i , e de transição do estado i ao estado j , respectivamente, para intervalos de tempo bem curtos. Uma definição mais matemática e menos verbal, portanto mais desejável, é dada pelo lema 7.1. Estas taxas se relacionam do seguinte modo

$$q_{ij} = \nu_i P_{ij}, \quad \nu_i = \sum_j \nu_i P_{ij} = \sum_j q_{ij}, \quad (7.1)$$

e também

$$P_{ij} = \frac{q_{ij}}{\nu_i} = \frac{q_{ij}}{\sum_j q_{ij}}, \quad (7.2)$$

onde P_{ij} é a probabilidade de, condicionado que haverá uma transição do estado i em um tempo bem curto, que ela seja ao estado j .

Embora ν_i e q_{ij} sejam chamadas de taxas de transição instantâneas, no exemplo de cadeias que agora está sendo estudado elas coincidem com as taxas de transição em um intervalo de tempo longo. Isto porque, de acordo com a proposição 1.4, uma variável de Poisson X de parâmetro λt tem média $E[X] = \lambda t$, portanto $\lambda = E[X]/t$. Neste processo que está sendo explicado, tem-se $q_{i,i+1} = \lambda$, $q_{i,i-1} = \mu$. Todavia, da definição de taxas **instantâneas** decorre que, se $|i - j| > 1$, então $q_{ij} = 0$, dado que da definição de processo de Poisson a probabilidade de mais de uma transição em um intervalo de tempo muito curto é desprezível, e para sair do estado i e chegar ao estado $i + 2$, por exemplo, neste processo que está sendo explicado agora, é preciso passar pelo estado $i + 1$ antes, isto é, são necessárias no mínimo 2 transições.

Observação 7.1. A média, ou valor esperado, $E[X]$, é o valor que se espera medir quando um experimento aleatório é realizado muitas vezes, no caso discreto, ou durante um longo intervalo de tempo no caso de tempo contínuo. Por exemplo, ao lançar uma moeda honesta, a probabilidade de obter cara é $1/2$; mas ao lançar a moeda 2 vezes, por exemplo, pode se obter 2 caras, o que não reflete a probabilidade $1/2$. Por outro lado, se o experimento for repetido muitas vezes, ou durante um longo tempo, afirma-se que os resultados convergirão para média $E[X]$.

Como foi suposto que o processo de chegada é um processo de Poisson com parâmetro λ , e o processo de saída de clientes, idem com parâmetro μ ; percebe-se que a taxa de transição quando em estado i é $\nu_i = \lambda + \mu$, e $q_{ij} = \nu_i P_{ij}$, onde P_{ij} é a probabilidade de, condicionado que haverá uma transição do estado i em um tempo bem curto, que seja ao estado j . Neste modelo específico de filas que está sendo definido nesta seção, tem-se

$$\nu_i = \lambda + \mu, \quad P_{i,i+1} = \frac{\lambda}{\lambda + \mu}, \quad P_{i,i-1} = \frac{\mu}{\lambda + \mu}.$$

Dado que no processo de Poisson a probabilidade de mais de uma transição $o(t)$ tende a zero mais rapidamente que t , então, se $|i - j| > 1$, $P_{ij} = 0$ e portanto $q_{ij} = 0$. Desde que o avanço da cadeia de um estado i a um estado $i + 1$, ou o retrocesso a um estado $i - 1$, são processos de Poisson com parâmetros λ e μ , respectivamente, vê-se que as taxas de transição são $q_{i,i+1} = \lambda$, $q_{i,i-1} = \mu$ e $\nu_i = \lambda + \mu$. Isto é perceptível da definição do processo de Poisson dada.

$$P\{N(t) = 1\} = \lambda t + o(t).$$

Visto que ν_i e q_{ij} são taxas instantâneas, ou seja, com $t \rightarrow 0$, então $o(t) \rightarrow 0$.

7.2 A função de probabilidade $P_{ij}(t)$

No exemplo que está sendo trabalhado, tem-se $\nu_i = \lambda + \mu$ para todo i . Mas em outros exemplos, poderia se ter $\nu_i \neq \nu_j$, para alguns i, j ; $i \neq j$. Note que as probabilidades P_{ij} , nas equações 7.1 e 7.2, não consideram o tempo da transição. Nestas equações, P_{ij} é apenas a probabilidade de, condicionado que haverá uma transição do estado i , que essa transição seja ao estado j . Não se tem uma função de probabilidade adequada, uma vez que q_{ij} apenas aproxima a probabilidade de **uma** transição de i para j em um tempo infinitesimal ou suficientemente pequeno. Logo a seguir, serão tecidas considerações extraídas do livro de Ross (15) que permitirão encontrar $P_{ij}(t)$, sendo t uma variável contínua. Para tanto, será necessário ver dois lemas.

Lema 7.1. Para as taxas de transição ν_i e q_{ij} , e $t > 0$, tem-se que

a)

$$\lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} = q_{ij};$$

b)

$$\lim_{t \rightarrow 0} \frac{1 - P_{ii}}{t} = \nu_i.$$

Demonstração. Para provar-se os itens (a) e (b), utiliza-se a definição de processo de Poisson. A probabilidade de que haja uma transição até o tempo t é igual a probabilidade de que haja uma transição do estado i , ou seja, $\nu_i t + o(t)$, multiplicada pela probabilidade de que esta transição seja ao estado j , P_{ij} .

$$\begin{aligned} \lim_{t \rightarrow 0} P_{ij}(t) &= \lim_{t \rightarrow 0} (\nu_i t + o(t)) P_{ij} \\ \iff \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} &= \nu_i P_{ij} + \lim_{t \rightarrow 0} \frac{o(t)}{t} P_{ij} \\ &= \nu_i P_{ij} + 0 = \nu_i P_{ij} = q_{ij}. \end{aligned}$$

De forma semelhante se prova a parte (b). Sabendo que $1 - P_{ii}(t)$ é a probabilidade de que haja uma transição até o tempo t , segue que

$$\begin{aligned} \lim_{t \rightarrow 0} 1 - P_{ii}(t) &= \lim_{t \rightarrow 0} (t\nu_i + o(t)) \\ \iff \lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} &= \nu_i + 0 = \nu_i. \end{aligned}$$

Ambas as partes estão provadas. □

Lema 7.2 (Equações de Chapman-Kolmogorov). A demonstração deste lema é bem elementar. Para todo $t, s \geq 0$, tem-se

$$P_{ij}(t + s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s).$$

Demonstração. Para sair do estado i , no tempo 0, e chegar a j no tempo $t + s$, a cadeia deve estar em algum estado no tempo t . Deste modo,

$$\begin{aligned} P_{ij}(t + s) &= P\{X(t + s) = j | X(0) = i\} \\ &= \sum_{k=0}^{\infty} P\{X(t + s) = j \cap X(t) = k | X(0) = i\}. \end{aligned}$$

Utilizando o teorema 1.2 nesta última equação, segue que

$$P_{ij}(t + s) = \sum_{k=0}^{\infty} P\{X(t + s) = j | X(t) = k, X(0) = i\} \cdot P\{X(t) = k | X(0) = i\};$$

e, devido à perda de memória da cadeia (propriedade markoviana), segue

$$\begin{aligned} P_{ij}(t + s) &= \sum_{k=0}^{\infty} P\{X(t + s) = j | X(t) = k\} \cdot P\{X(t) = k | X(0) = i\} \\ &= \sum_{k=0}^{\infty} P_{kj}(s) P_{ik}(t) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s). \end{aligned}$$

E a prova está completa. □

As equações de Chapman-Kolmogorov são a chave para encontrar a função de probabilidade $P_{ij}(t)$. Destas equações derivam as chamadas **equações de Kolmogorov para trás**, e também as equações de Kolmogorov para frente; estas últimas serão deixadas em segundo plano, até o ponto em que não haja prejuízo de compreensão.

Teorema 7.1 (Kolmogorov para trás).

$$P'_{ij}(t) = q_{ik}P_{kj}(t) - \nu_i P_{ij}(t).$$

Demonstração. Fazendo uso das equações de Chapman-Kolmogorov, tem-se

$$\begin{aligned} P_{ij}(h+t) &= \sum_{k=0}^{\infty} P_{ik}(h)P_{kj}(t) \\ \iff P_{ij}(h+t) - P_{ij}(t) &= \sum_{k=0}^{\infty} P_{ik}(h)P_{kj}(t) - P_{ij}(t) \\ P_{ij}(h+t) - P_{ij}(t) &= \sum_{\substack{k=0 \\ k \neq i}}^{\infty} P_{ik}(h)P_{kj}(t) + P_{ii}(h)P_{ij}(t) - P_{ij}(t) \\ &= \sum_{\substack{k=0 \\ k \neq i}}^{\infty} P_{ik}(h)P_{kj}(t) - (1 - P_{ii}(h))P_{ij}(t) \\ \iff \lim_{h \rightarrow 0} \frac{P_{ij}(h+t) - P_{ij}(t)}{h} &= \lim_{h \rightarrow 0} \left\{ \sum_{\substack{k=0 \\ k \neq i}}^{\infty} \frac{P_{ik}(h)}{h} P_{kj}(t) - \frac{1 - P_{ii}(h)}{h} P_{ij}(t) \right\} \\ &= P'_{ij}(t) = \lim_{h \rightarrow 0} \sum_{\substack{k=0 \\ k \neq i}}^{\infty} \frac{P_{ik}(h)}{h} P_{kj}(t) - \lim_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} P_{ij}(t). \end{aligned}$$

Aplicando a parte (b) do lema 7.1, chega-se a

$$P'_{ij}(t) = \lim_{h \rightarrow 0} \sum_{\substack{k=0 \\ k \neq i}}^{\infty} \frac{P_{ik}(h)}{h} P_{kj}(t) - \nu_i P_{ij}(t).$$

Supondo que seja correto permutar o limite com o somatório, chega-se a

$$P'_{ij}(t) = \sum_{\substack{k=0 \\ k \neq i}}^{\infty} \lim_{h \rightarrow 0} \frac{P_{ik}(h)}{h} P_{kj}(t) - \nu_i P_{ij}(t).$$

Aplicando agora a parte (a) do mesmo lema, segue

$$P'_{ij}(t) = \sum_{\substack{k=0 \\ k \neq i}}^{\infty} q_{ik}P_{kj} - \nu_i P_{ij}.$$

□

Ainda de acordo com obra de Ross, a mesma afirma que esta permuta de limite com somatório nem sempre é correta; todavia é justificável em todos os processos de nascimento e morte, como o citado autor chama o processo que está sendo usado neste trabalho para modelar o tamanho de filas; e também em todos os modelos de cadeias de Markov de estados finitos. Cabe ressaltar que este processo de nascimento e morte (tamanho de filas) é uma cadeia de Markov, já que o próximo estado da cadeia só depende do estado presente. E também que o livro de Cálculo do Guidorizzi (16) respalda o que foi feito; na seção de propriedades de limite, tem-se

$$\lim_{x \rightarrow p} [f(x) + g(x)] = \lim_{x \rightarrow p} f(x) + \lim_{x \rightarrow p} g(x).$$

Por outro lado, se a cadeia de Markov tiver espaço de estados S infinito, o somatório da demonstração acima pode eventualmente também ser infinito. Neste caso, não há garantia de que o limite da soma será igual à soma dos limites.

O modelo de filas introduzido na seção 7.1, em equação de Kolmogorov para trás tem a seguinte forma.

$$\begin{aligned} P'_{ij}(t) &= (\lambda + \mu) \left[\frac{\lambda}{\lambda + \mu} P_{i+1,j}(t) + \frac{\mu}{\lambda + \mu} P_{i-1,j}(t) \right] - \nu_i P_{ij}(t) \\ &= \lambda P_{i+1,j}(t) + \mu P_{i-1,j}(t) - \nu_i P_{ij}(t). \end{aligned}$$

Isto porque, neste modelo específico, se $|i - j| > 1$ é preciso mais de uma transição para sair de i e chegar a j , portanto $q_{ij} = 0$ neste caso.

7.2.1 Método iterativo de computar $P_{ij}(t)$

Considere os números r_{ij} dados por

$$r_{ij} = \begin{cases} q_{ij} & \text{se } i \neq j \\ -\nu_i & \text{se } i = j \end{cases}.$$

Com esta notação as equações de Kolmogorov para trás 7.1 se tornam

$$P'_{ij}(t) = \sum_k r_{ik} P_{kj}(t).$$

Fazendo $P'(t)$ denotar uma matriz quadrada dos elementos $P'_{ij}(t)$, tem-se a seguinte equação diferencial

$$P'(t) = R \cdot P(t),$$

onde R é uma matriz quadrada dos elementos r_{ij} e $P(t)$ é a matriz de probabilidades de transição no intervalo de tempo $[0, t]$ que se procura.

A solução de uma equação diferencial do tipo $f'(x) = cf(x)$ é $f(x) = ke^{cx}$, onde k é uma constante a ser determinada. O mesmo raciocínio é aplicado a funções de uma variável real a uma matriz, em Ross. Deste modo, $P(t) = Ke^{Rt}$. Utilizando o valor inicial $P(0) = I$,

de forma semelhante ao explicado na seção 3.1 para as cadeias de Markov com tempo discreto, pode-se determinar a constante K . Desta maneira,

$$\begin{aligned} P(0) &= Ke^{R \cdot 0} = I \\ &= Ke^0 = I \\ &= K \cdot I = I, \quad \text{portanto} \quad K = I \end{aligned}$$

Feito isto, vê-se que

$$P'(t) = RP(t), \quad P(0) = I, \quad \iff P(t) = e^{Rt}.$$

Um significado para e^{Rt} pode ser encontrado a partir das séries de Taylor. Nestas séries, tem-se a função exponencial sob a forma

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}.$$

De forma análoga

$$e^{Rt} = \sum_{i=0}^{\infty} R^i \frac{t^i}{i!}.$$

Todavia, a equação acima tem um custo computacional além do necessário; um método melhor pode ser utilizar um limite fundamental do Cálculo; o seguinte

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Portanto,

$$e^{Rt} = \lim_{n \rightarrow \infty} \left(I + R \frac{t}{n}\right)^n. \quad (7.3)$$

7.2.1.1 Complexidade computacional

É interessante saber a complexidade computacional de um algoritmo para se estimar o tempo de execução do cômputo, e também decidir se os parâmetros escolhidos são viáveis, sejam eles a ordem m das matrizes, e o número de iterações n . Portanto, aqui será estimada a complexidade computacional do método 7.3; para tal, serão contadas as multiplicações e as somas.

Para começar, pode-se fazer

$$e^{Rt} = \lim_{n \rightarrow \infty} \left[I + R \left(\frac{t}{n} \right) \right]^n,$$

o que é um pouco mais econômico pelo que segue; então, ao fazer t/n , tem-se uma multiplicação. Deve ser feita, agora, a multiplicação de t/n por R ; supondo que R seja uma matriz quadrada de ordem m , então ela tem m^2 entradas portanto, são mais m^2 multiplicações. Se fosse feito $R \cdot t$ e $R \cdot 1/n$, seriam $2m^2$ multiplicações. Até aqui, $m^2 + 1$ multiplicações. Em seguida, vem a soma no argumento da potência, $I + Rt/n$, tem-se uma soma.

Seja $H = \left(I + R \frac{t}{n} \right)$. Tem-se que $H^s = H^{s-1} \cdot H$; e sejam $[H^s]_{ij} = h_{ij}^{(s)}$ e $[H]_{ij} = h_{ij}$, destas considerações segue que

$$[H^s]_{ij} = h_{ij}^{(s)} = \sum_{k=1}^m h_{ik}^{(s-1)} h_{kj}.$$

Ou seja, a cada multiplicação de H por ela própria, obtem-se uma matriz resultante de m^2 entradas, sendo que cada entrada é calculada por m multiplicações, e $m - 1$ somas. Portanto, ao multiplicar H por ela mesma, tem-se mais m^3 multiplicações, e $m^2(m - 1)$ somas. Até aqui são $m^3 + m^2 + 1$ multiplicações, e $m^2(m - 1) + 1$ somas.

Perto de terminar, veja que ao fazer $H \cdot H$ tem-se, como foi dito, m^3 multiplicações e $m^2(m - 1)$ somas; no entanto, na equação 7.3 tem-se $n - 1$ produtos iguais a este. O que conduz a um total de $m^3(n - 1) + m^2 + 1$ multiplicações, e $m^2(m - 1)(n - 1) + 1$ somas. Em resumo

$$\begin{cases} m^3(n - 1) + m^2 + 1 & \text{multiplicações} \\ m^2(m - 1)(n - 1) + 1 & \text{somas} \end{cases}.$$

Considerando que o número de somas é inferior ao de multiplicações, e também que multiplicações são exageradamente mais complexas que somas, é natural desprezar as somas na estimativa do tempo de cômputo. Outro ponto importante é que n é um parâmetro elevado, uma vez que ele tem um peso pequeno no custo computacional, e também porque é ele que dá o grau de refinamento do cômputo. Sendo assim, o custo computacional se resume a m^3n quando $n \rightarrow \infty$. Diz-se, então, que a complexidade do método 7.3 é da ordem $\mathcal{O}(m^3n)$, onde m é a ordem das matrizes utilizadas e n é o grau de refinamento do cálculo computacional.

Observação 7.2. Todavia, o que acontece na prática é diferente. Em uma implementação computacional, o mais comum é fazer o produto da matriz $H = I + Rt/n$ por ela mesma, e assim obtendo H^2 , em seguida $H^2 \cdot H^2 = H^4$, $H^4 \cdot H^4 = H^8 \dots$. Assim, se n é um número da forma $n = 2^k$, $k \in \mathbb{N}$, o número de produtos matriciais é k . Então o correto é dizer que o custo computacional é $\mathcal{O}(m^3 \log_2 n + m^2 + 1)$, portanto pouco depende de n , depende quase que exclusivamente de m . Deste modo, os números de multiplicações e somas são dados a seguir.

$$\begin{cases} m^2(m - 1) \log_2 n + 1 & \text{somas} \\ m^3 \log_2 n + m^2 + 1 & \text{multiplicações} \end{cases}.$$

Exemplo 5. Se um algoritmo baseado na equação 7.3, com matrizes de ordem $m_0 = 4$ e grau de refinamento $n_0 = 100$, levou um tempo de computação τ_0 , quanto tempo mais de cômputo se espera duplicando a ordem das matrizes, e aumentando em 50% o parâmetro n de refinamento da resposta.

Solução: A complexidade do método é da ordem $\mathcal{O}(m^3 \log_2 n + m^2 + 1)$, então duplicando a ordem das matrizes m , tem-se $m = 2m_0$; aumentando n em 50%, tem-se $n = 1,5n_0$. Portanto

o tempo do segundo cômputo será da ordem

$$\mathcal{O}[(2m_0)^3 \log_2 1, 5n_0 + (2m_0)^2 + 1] = \mathcal{O}(8m_0^3 \log_2 1, 5n_0 + 4m_0^2 + 1).$$

Se n_0 for grande, então $\log_2 1, 5n_0 \approx \log_2 n_0$, o que conduz a um tempo τ que é dado por

$$\tau/\tau_0 = \frac{8m_0^3 \log_2 n_0 + 4m_0^2 + 1}{m_0^3 \log_2 n_0 + m_0^2 + 1} = \frac{8 \cdot 64 \log_2 100 + 4 \cdot 16 + 1}{64 \log_2 100 + 17} = \frac{512 \log_2 100 + 65}{64 \log_2 100 + 17} \approx 7,84.$$

Por tudo isso vê-se que o tempo do segundo cômputo será quase 8 vezes o do primeiro; e também, que neste caso o que determina o tempo de computação é a ordem da matriz de probabilidades a ser calculada.

◁

7.3 A matriz R no modelo de filas

A atenção retorna à aplicação da modelagem de filas, principalmente, filas em agências bancárias, onde há constantes chegadas e saídas de clientes, alterando o tamanho da fila de espera.

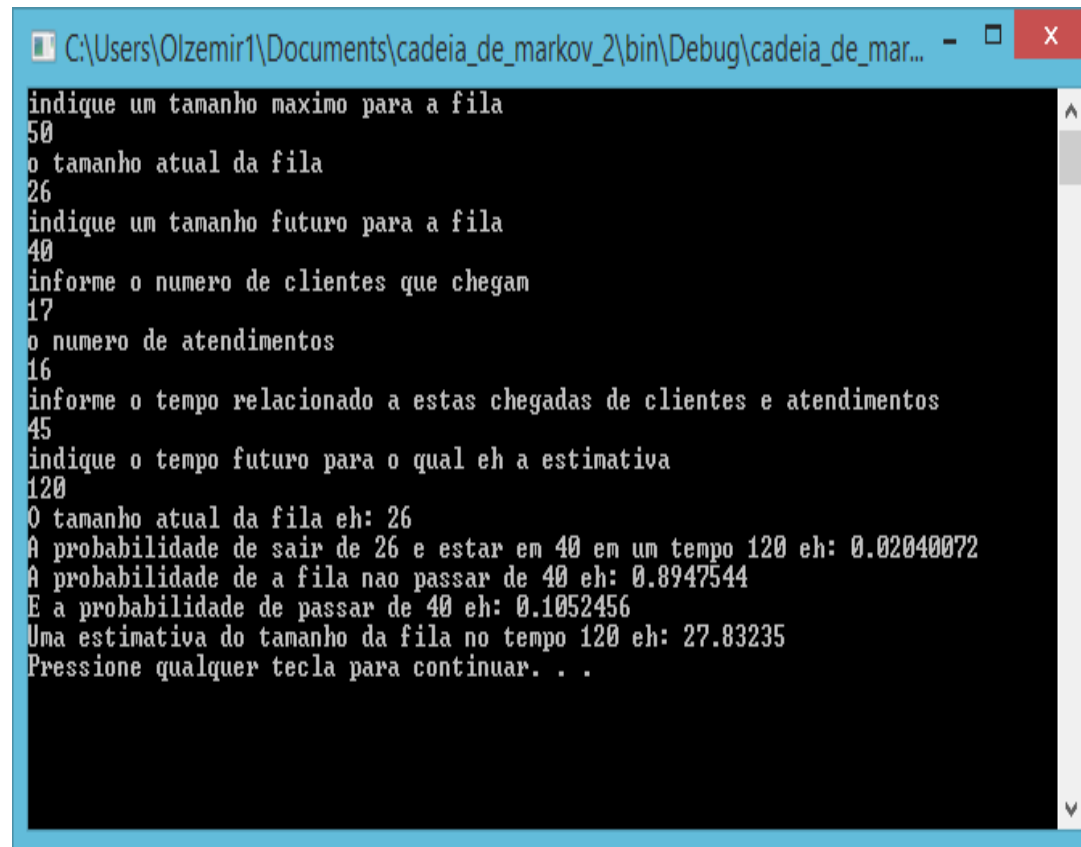
Como explicado anteriormente, a probabilidade de mais de uma transição em um tempo pequeno é desprezível; também foram definidas as taxas de transição q_{ij} como a probabilidade de **uma** transição em um intervalo de tempo pequeno, e assim incorporadas às equações de Kolmogorov para trás, o teorema 7.1, a partir do lema 7.1. Sendo assim, se são necessárias mais de uma transição para sair de i e chegar a j , então $q_{ij} = 0$. Segue que a matriz R deste modelo, de qualquer ordem, tem a seguinte forma.

$$R = \begin{pmatrix} -\nu_i & q_{12} & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ q_{21} & -\nu_i & q_{23} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & q_{32} & -\nu_i & q_{34} & 0 & 0 & \cdots & 0 & \vdots \\ 0 & 0 & q_{43} & -\nu_i & q_{45} & 0 & \cdots & 0 & \vdots \\ \vdots & \cdots & \cdots & \ddots & \ddots & \ddots & \cdots & 0 & \vdots \\ 0 & 0 & \cdots & q_{l,l-1} & -\nu_i & q_{l,l+1} & 0 & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & \cdots & \vdots & \ddots & 0 & q_{n-1,n-2} & -\nu_i & q_{n-1,n} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & q_{n,n-1} & -\nu_i \end{pmatrix}$$

$$= \begin{pmatrix} -(\lambda + \mu) & \lambda & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \mu & -(\lambda + \mu) & \lambda & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & 0 & 0 & 0 & 0 & \mu & -(\lambda + \mu) & \lambda \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \mu & -(\lambda + \mu) \end{pmatrix}.$$

Veja um exemplo de simulação em computador, utilizando um programa compilado em linguagem de programação C.

Figura 5 – Estimação de fila feita em computador.



```
C:\Users\Olzemir1\Documents\cadeia_de_markov_2\bin\Debug\cadeia_de_mar... - □ ×
indique um tamanho maximo para a fila
50
o tamanho atual da fila
26
indique um tamanho futuro para a fila
40
informe o numero de clientes que chegam
17
o numero de atendimentos
16
informe o tempo relacionado a estas chegadas de clientes e atendimentos
45
indique o tempo futuro para o qual eh a estimativa
120
O tamanho atual da fila eh: 26
A probabilidade de sair de 26 e estar em 40 em um tempo 120 eh: 0.02040072
A probabilidade de a fila nao passar de 40 eh: 0.8947544
E a probabilidade de passar de 40 eh: 0.1052456
Uma estimativa do tamanho da fila no tempo 120 eh: 27.83235
Pressione qualquer tecla para continuar. . .
```

Utilizou-se dados empíricos a título de experimentação, os quais aparecem na figura, abaixo de suas respectivas requisições. O código-fonte do programa exibido na figura está no anexo A.

Observação 7.3. É vital destacar a importância de ter alguns cuidados ao se manipular o programa acima. Veja que o programa informa a soma de probabilidades de uma linha de $P(t)$, desde o estado 0 até o estado j indicado. Denotando esta soma por $P(X)$, isto significa que é exibida na tela a probabilidade $P(X)$ tal que

$$P(X) = \sum_{k=0}^j P_{ik}(t).$$

Afirma-se que $P(X)$ é um indicador importante para o correto uso do programa. Este ajuda o usuário a saber a distribuição de probabilidades em torno dos estados i e j , que são o tamanho atual da fila, e um palpite do usuário sobre o tamanho futuro, respectivamente. É portanto necessário que ao informar o campo “tamanho máximo da fila”, o valor atribuído a este campo seja grande o suficiente para incluir nos cálculos as probabilidades mais relevantes; do contrário, alguns dados de saída do programa serão imprecisos, por exemplo, o tamanho estimado da fila após um tempo t indicado.

7.4 O tamanho da fila após um tempo t

O tamanho da fila em um tempo futuro pode ser estimado de acordo com os dados obtidos no presente. O programa de computador elaborado faz essa estimativa para o tempo t indicado. Para fazer isso, é utilizada a definição de média 1.7.

Por outro lado, é bem provável que seja possível estimar o tamanho da fila em um tempo futuro fazendo uso de um método analítico. É natural pensar que o tamanho da fila seja estimado pela quantidade de pessoas que chegam menos as que são atendidas. Portanto, se X_t é a quantidade de pessoas que chegam em um intervalo de tempo t , e Y_t as que são atendidas no intervalo de tempo t , e se k é o tamanho inicial, então **supõe-se** que o tamanho estimado da fila, de acordo com as proposições 1.1 e 1.2, seja

$$E[X_t - Y_t + k] = E[X_t] - E[Y_t] + k.$$

Como os processos de chegada e saída de clientes são processos de Poisson, então fixando o tempo t , tem-se variáveis aleatórias X_t e Y_t com distribuições de probabilidade de Poisson de parâmetros λt e μt , respectivamente, de acordo com a proposição 6.2. Portanto, se X_t e Y_t são variáveis de Poisson, então suas médias são conhecidas, e são, respectivamente, λt e μt , de acordo com a proposição 1.4. Por tudo isto, tem-se a estimativa do tamanho da fila assim

$$E[X_t - Y_t + k] = E[X_t] - E[Y_t] + k = \lambda t - \mu t + k.$$

Esta última equação está de acordo com todas as simulações feitas no programa.

8 Obtenção dos parâmetros dos processos

Foram tratadas cadeias de Markov de tempo contínuo nos capítulos anteriores, e os processos de transições utilizados foram os processos de Poisson. Daí segue que é preciso um meio para encontrar os parâmetros nos processos de Poisson. Esta é uma tarefa trivial, pelo menos no que diz respeito aos conhecimentos matemáticos.

Utilizando-se alguma técnica estatística de amostragem, deve-se observar a média de chegada de clientes; do mesmo modo faz-se com a média de atendimento de clientes. É preciso lembrar a proposição 1.4; de acordo com ela, uma variável de Poisson X , de parâmetro α , tem média $E[X] = \alpha$. Em um processo de Poisson, fixando o tempo t , tem-se uma distribuição de Poisson com parâmetro λt ; portanto, a média de chegada de clientes é $E[X_\lambda] = \lambda t$, ou de forma equivalente, tem-se $E[X_\lambda]/t = \lambda$, o que permitiu encontrar a taxa de transição instantânea $q_{i,i+1} = \lambda$. O mesmo vale para o processo de decrescimento da fila, $E[X_\mu] = \mu t \iff E[X_\mu]/t = \mu = q_{i,i-1}$.

Observação 8.1. Como foi mencionado anteriormente, tomando um tempo t suficientemente grande, as amostras ou informações coletadas relacionadas a uma variável aleatória X terão convergido para a média $E[X]$.

Conclusões

Por tudo que foi mostrado, vê-se que a teoria de cadeias de Markov é um tema de muito interesse prático. Não somente pela importância de algumas de suas aplicações, também pela variedade de fenômenos que podem ser modelados a partir do conceito de cadeias de Markov. O livro de Meyen e Tweedie (12), p. 4, elenca vários fenômenos de elevado interesse prático que podem ser modelados via cadeias de Markov. Cite-se dois exemplos do citado livro.

- a) Sistemas de controle cruzado de motores de veículos modernos possuem um monitor que faz leituras de informações, como fluxo de combustível, velocidade, e outros semelhantes, e em resposta ajusta a aceleração. O ajuste da aceleração em resposta às leituras produz novas leituras. Se X é um vetor cujas entradas são as informações de leitura, e Y é um vetor de resposta do algoritmo de controle, então $\phi_n = (X_n, Y_n)$ é uma cadeia de Markov, já que o estado futuro só depende do estado presente;
- b) A taxa de câmbio X_n entre duas moedas pode ser representada como uma função de vários valores passados seus X_n, \dots, X_{n-k} , modificada pela volatilidade do mercado, que é incorporada em um termo de perturbação W_n . O modelo autoregressivo

$$X_n = \sum_{j=1}^k \alpha_j X_{n-j} + W_n$$

central em análise de séries de tempo captura a essência do conceito relacionado a tal sistema. Considerando o comprimento do vetor $\phi_n = (X_n, \dots, X_{n-k+1})$, métodos markovianos podem ser trazidos para a análise de tais modelos de séries de tempo.

Ainda que algumas coisas mudem, talvez acabem as filas nos bancos, é bem provável que algumas situações continuem a apresentar filas, como aeroportos por exemplo. Além disto, vale continuar os estudos de cadeias de Markov porque, com algumas poucas modificações, é possível fazer modelos de outros fenômenos.

O programa de computador criado

Além do que foi dito acima, vale destacar que a aplicação de computador apresentada neste trabalho pode ser uma ferramenta de trabalho bem recepcionada em ambientes de atendimento ao público, tanto os atendimentos via telefone, *chat*, . . . , como os presenciais; o que se diz por reparar a simplicidade de operar tal programa, que não exige formação além de conhecimentos corriqueiros de informática básica. Além da simplicidade do programa, um exame rápido do mesmo parece mostrar uma alta performance, tanto no sentido computacional, que é devida aos progressos da Engenharia da Computação em colocar no mercado computadores pessoais sucessivamente mais avançados e a baixo custo financeiro, e também à própria linguagem C, uma

linguagem de alta performance, assim também a praticidade de se poder fazer estimativas pouco depois de observar o fluxo de clientes, sem ter que se fazer uso de roteiros estatísticos complexos. Isto porque os tempos declarados ao programa são arbitrariamente escolhidos, inclusive em relação à unidade de medida do tempo, que pode ser outra quando o programa é reiniciado.

Método analítico para estimação de fila

Conforme brevemente falado na seção 7.4, é provável uma solução analítica para estimar o tamanho futuro de uma fila. No entanto, chegando a este ponto já com o prazo de conclusão do trabalho se esgotando, não há outra alternativa senão esperar que tudo que foi escrito até agora seja o suficiente para este momento. E, com sorte, uma “semente” para o futuro.

Por tudo isto, espera-se que este trabalho venha contribuir para melhor aproveitamento de recursos em geral, planejamento eficiente de atividades econômicas e, conseqüentemente, aumento de produtividade.

ANEXO A – Código-fonte do programa

Segue anexo o código-fonte de um programa em linguagem de programação científica C, este mostrado na figura 5, que simula o tamanho de uma fila, utilizando um modelo de cadeias de Markov que foi delineado no capítulo 7.

```

1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <math.h>
4
5  /**< Este projeto é um código para calcular uma cadeia de Markov de tempo contínuo */
6  /* Existem alguns comandos inativos dentro do código-fonte. Eles podem ser
7  ativados e têm o papel de mostrar o que acontece durante a execução do programa,
8  com o intuito de mostrar se o programa está funcionando corretamente. */
9
10 int main()
11 {
12     int g,i,j,h,n,s,k = 20,x,y,m;
13     double e_lambda,e_mu,tempo,lambda,mu,t,J,M = 0.0;
14     printf("indique um tamanho maximo para a fila\n");
15     scanf("%d", &g);
16     printf("o tamanho atual da fila\n");
17     scanf("%d", &x);
18     printf("indique um tamanho futuro para a fila\n");
19     scanf("%d", &y);
20     printf("informe o numero de clientes que chegam\n");
21     scanf("%lf", &e_lambda);
22     printf("o numero de atendimentos\n");
23     scanf("%lf", &e_mu);
24     printf("informe o tempo relacionado a estas chegadas de clientes e atendimentos\n");
25     scanf("%lf", &tempo);
26     printf("indique o tempo futuro para o qual eh a estimativa\n");
27     scanf("%lf", &t);
28     m = g + 1;
29     double R[m][m],I[m][m],H[m][m],L[m][m];
30     lambda = e_lambda/tempo;
31     mu = e_mu/tempo;
32     n = pow(2,k);
33
34     for(i = 0;i < m;i++){
35         for(j = 0;j < m;j++){
36             if(j - i == 1){          /**< Preenchendo a matriz R */
37                 R[i][j] = lambda;
38             }
39             else{
40                 if(i - j == 1){
41                     R[i][j] = mu;
42                 }
43
44                 else{
45                     if(i == j){
46                         R[i][j] = - lambda - mu;
47                     }
48                     else{
49                         R[i][j] = 0.0;
50                     }
51                 }
52             }
53         }
54     }
55
56     /* for(i = 0;i < m;i++){
57         for(j = 0;j < m;j++){
58             printf("%.15g ",R[i][j]);
59         }
60         printf("\n");
61     }
62
63     system("pause");          /**< Programar a matriz identidade I*/
64
65     for(i = 0;i < m;i++){
66         for(j = 0;j < m;j++){

```

```

67         if(i == j){
68             I[i][j] = 1.0;
69         }
70         else{I[i][j] = 0.0;
71         }
72     }
73 }
74
75 /*
76     for(i = 0;i < m;i++){
77         for(j = 0;j < m;j++){
78             printf("%.7g ",I[i][j]);
79         }
80     printf("\n");
81     }
82
83     system("pause");
84
85                                     /**< Começando a programar o algoritmo */
86     for(i = 0;i < m;i++){
87         for(j = 0;j < m;j++){
88             H[i][j] = R[i][j] * (t/((double)n));
89         }
90     }
91 /*
92     for(i = 0;i < m;i++){
93         for(j = 0;j < m;j++){
94             printf("%.7g ",H[i][j]);
95         }
96     printf("\n");
97     }
98
99     system("pause"); /* até aqui ok! */
100
101     for(i = 0;i < m;i++){
102         for(j = 0;j < m;j++){ /**< Para 'n' muito grande a precisão da máquina é extrapolada aqui */
103             H[i][j] = H[i][j] + I[i][j];
104         }
105     }
106 /*
107     for(i = 0;i < m;i++){
108         for(j = 0;j < m;j++){
109             printf("%.7g ",H[i][j]);
110         }
111     printf("\n");
112     }
113
114
115     system("pause");
116
117                                     /**< Agora, a potência matricial */
118     for(s = 1;s <= k;s++){
119         for(i = 0;i < m;i++){
120             for(j = 0;j < m;j++){
121                 J = 0;
122                 for(h = 0;h < m;h++){
123                     J = H[i][h] * H[h][j] + J; /*Começa o produto matricial*/
124                 }
125                 L[i][j] = J; /*Calcula cada entrada da matriz produto */
126             }
127         }
128     for(i = 0;i < m;i++){
129         for(j = 0;j < m;j++){
130             H[i][j] = L[i][j];          /*Atualiza a matriz H para o próximo produto */
131         }
132     }

```

```

133     }
134
135     /* for(i = 0;i < m;i++){
136         for(j = 0;j < m;j++){
137             printf("%.7g ",H[i][j]);
138         }
139         printf("\n");
140     }
141     system("pause");                Potência ok! agora é imprimir na tela */
142
143                                     /**< O tamanho médio da fila */
144     for(j = 0;j < m;j++){
145         M = H[x][j] * j + M;
146     }
147
148     if(m < 7){                        /**< Se a matriz for pequena, dá para imprimir tudo */
149         for(i = 0;i < m;i++){
150             for(j = 0;j < m;j++){
151
152                 printf("%.4g ",H[i][j]);
153             }
154             printf("\n");
155         }
156         printf("O resultado esta acima\n");
157         system("pause");
158     }
159     else{
160         for(J = 0,s = 0;s <= y;s++){
161             J = H[x][s] + J;
162         }
163         printf("O tamanho atual da fila eh: %d\n",x);
164         printf("A probabilidade de sair de %d e estar em %d em um tempo %.7g eh: %.7g\n",x,y,t,H[x][y]);
165         printf("A probabilidade de a fila nao passar de %d eh: %.7g\n",y,J);
166         printf("E a probabilidade de passar de %d eh: %.7g\n",y,1 - J);
167         printf("Uma estimativa do tamanho da fila no tempo %.7g eh: %.7g\n",t,M);
168     }
169     system("pause");
170     return 0;}

```

Referências

- 1 CHING, W. et al. *Markov Chains*. 2. ed. New York: Springer, 2013. 1-22 p. Citado 6 vezes nas páginas 7, 8, 24, 29, 35 e 37.
- 2 AHARONOV, D. et al. Quantum walks on graphs. p. 1, 2002. Disponível em: <[arXiv:quant-ph/0012090](http://arxiv.org/abs/quant-ph/0012090)>. Citado 2 vezes nas páginas 7 e 8.
- 3 BRIN, S. et al. The pagerank citation ranking: Bringing order to the web. p. 3–6, 1998. Disponível em: <<http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>>. Citado 2 vezes nas páginas 7 e 8.
- 4 BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, Elsevier, v. 30, n. 1-7, p. 107–117, 1998. Citado 2 vezes nas páginas 7 e 8.
- 5 O’CONNOR, J. J.; ROBERTSON, E. F. *Andrei Andreyevich Markov*. Disponível em: <<http://www-history.mcs.st-andrews.ac.uk/Biographies/Markov.html>>. Acesso em: 27 de julho de 2018. Citado na página 12.
- 6 MEYER, P. L. *Probabilidade: aplicações à estatística*. 2. ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos, 1983. Citado 2 vezes nas páginas 15 e 18.
- 7 KNILL, O. *Probability and Stochastic Processes*. [S.l.: s.n.], 2011. 7, 319, 320 p. Citado 2 vezes nas páginas 16 e 36.
- 8 SPIEGEL, M. R. *Probabilidade e Estatística*. São Paulo: McGraw-Hill do Brasil, 1978. Citado na página 18.
- 9 COELHO, F. U.; LOURENÇO, M. L. *Curso de Álgebra Linear, Um*. 2. ed. São Paulo: Edusp – Editora da Universidade de São Paulo, 2010. Citado na página 20.
- 10 DANTE, L. R. *Matemática: contexto e aplicações*. 1. ed. São Paulo: Ática, 2010. v. 2. Citado na página 20.
- 11 SANTOS, R. A. M. *Cadeias de Markov Quânticas: Capítulo 2, caminhos aleatórios*. 10 p. Dissertação (Mestrado em Modelagem Computacional) — Laboratório Nacional de Computação Científica, Petrópolis, RJ – Brasil, 2010. Citado na página 23.
- 12 MEYN, S. P.; TWEEDIE, R. L. *Markov Chains and Stochastic Stability*. New York: Springer-Verlag, 2005. 4 p. Citado 2 vezes nas páginas 29 e 56.
- 13 MEYER, C. D. *Matrix Analysis and Applied Linear Algebra*. [S.l.]: SIAM, 2000. 674,687-694 p. Citado na página 30.
- 14 BURDEN, R. L.; FAIRES, J. D. *Numerical Analysis*. 9. ed. Boston: Brooks/Cole – Cengage Learning, 2011. 450-454 p. Citado 2 vezes nas páginas 31 e 33.
- 15 ROSS, S. M. *Introduction to Probability Models*. 10. ed. New York: Elsevier, 2010. 305, 388-436 p. Citado 3 vezes nas páginas 35, 43 e 45.

16 GUIDORIZZI, H. L. *Curso de Cálculo, Um. 5.* ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2001. v. 1. 286,456 p. Citado 3 vezes nas páginas 37, 38 e 48.

17 WIKIPÉDIA, A ENCICLOPÉDIA LIVRE. *Cadeias de Markov*. Disponível em: <https://pt.m.wikipedia.org/wiki/Cadeias_de_Markov>. Acesso em: 29 de maio de 2019. Nenhuma citação no texto.