UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José do Rio Preto

Thaís Saes Giuliani Ribeiro

Processos de Markov discretos: exemplos voltados para o Ensino Médio

Bauru
2017

Thaís Saes Giuliani Ribeiro

Processos de Markov discretos: exemplos voltados para o Ensino Médio

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, Polo de Bauru.

Financiadora: CAPES

Orientador: Prof. Dr. Fabiano Borges da Silva

Bauru
2017

Ribeiro, Thaís Saes Giuliani.
Processos de Markov discretos: exemplos voltados para o ensino médio / Thaís Saes Giuliani Ribeiro. -- São José do Rio Preto, 2017
73 f. : il.

Orientador: Fabiano Borges da Silva
Dissertação (Mestrado profissional) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Letras e Ciências Exatas

1. Matemática (Ensino Médio) - Estudo e ensino. 2. Markov, Processos de. 3. Processo estocástico. 4. Probabilidades. 5. Álgebra linear. 6. Matemática – Metodologia. I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 51(07)

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Thaís Saes Giuliani Ribeiro

Processos de Markov discretos: exemplos voltados para o Ensino Médio

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, Polo de Bauru.

Financiadora: CAPES

Comissão Examinadora

Prof. Dr. Fabiano Borges da Silva
UNESP Bauru
Orientador

Profª. Drª. Tatiana Miguel Rodrigues de Souza
UNESP Bauru

Prof. Dr. Leandro Batista Morgado
Universidade Federal de Santa Catarina

Bauru
30 de novembro de 2017

Agradecimentos

Agradeço à Deus, primeiramente, por me propiciar a oportunidade de concluir este trabalho.

À minha família que, mesmo de longe, acredita e torce muito por mim.

Ao meu marido Octávio, que me apoiou, compreendendo minhas ausências e me tranquilizou nos momentos difíceis.

Aos professores do PROFMAT de Bauru , por transmitirem todo conhecimento e incentivo a cada etapa, nos motivando a sempre prosseguir.

Aos colegas do PROFMAT 2015, pelo convívio, amizade e companheirismo.

Ao meu orientador Prof. Dr. Fabiano Borges da Silva, pela dedicação e incentivo na execução deste trabalho.

À CAPES, pelo apoio financeiro.

Resumo

Neste trabalho, mostramos como construir um processo estocástico de Markov e seu espaço de probabilidade a partir das probabilidades de transição e da distribuição inicial. Além disso, mostramos a convergência das matrizes de transição utilizando como ferramenta conhecimentos de Álgebra Linear. A aplicação das cadeias de Markov num contexto voltado para o Ensino Médio é mostrado no último capítulo, onde procuramos oferecer aos alunos a oportunidade de ter uma visão mais ampla de como a Matemática pode ser aplicada em outras áreas do conhecimento.

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

Abstract

In this work, we show how to construct a stochastic Markov process and its probability space from the transition probabilities and the initial distribution. In addition, we show to investigate the convergence of the transition matrices using Linear Algebra knowledge as a tool. Application of Markov chains in a context focused on High School, it is shown in the last chapter, where we try to offer the students the opportunity to have a view of how mathematics can be applied in other areas of knowledge.

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

Sumário

Introdução

Um dos maiores desafios no ensino de matemática é a contextualização dos conteúdos, isto é, trabalhar a matemática de maneira prática e com aplicações em situações cotidianas. Muitas vezes, nós professores, enfrentamos certa dificuldade em trabalhar os conteúdos de matemática, principalmente conceitos do Ensino Médio, mostrando aos alunos uma aplicação prática sobre aquele assunto.

A grosso modo, cadeia de Markov é um processo estocástico sem memória, caracterizado por seu estado futuro depender apenas do seu estado no presente, sabendo que os estados passados não influenciam no estado futuro. Para o estudo das cadeias de Markov, em geral utilizamos como ferramenta matrizes e probabilidades de transição, que nos permite, com certa precisão, prever acontecimentos futuros de certos fenômenos, e com isso tomarmos decisões para atingirmos os resultados desejados. Assim, a cadeia de Markov é um instrumento que propicia ao professor uma abordem do conteúdo de matrizes, sistemas lineares e probabilidade de maneira contextualizada e de forma eficaz, pois dessa maneira, os alunos poderão assimilar os conceitos de forma prática e constatar que o conteúdo que se aprende na escola pode ser utilizado de forma efetiva em problemas reais.

Este trabalho busca oferecer ao leitor um embasamento teórico sobre espaço de probabilidade para cadeias de Markov e dar exemplos que possam ser utilizados no Ensino Médio.

No Capítulo 1, buscamos retomar os principais conceitos da teoria de probabilidade que serão de grande importância para o desenvolvimento posterior deste trabalho, especialmente as definições de álgebra, σ-álgebra, probabilidade e variável aleatória.

No segundo capítulo, fizemos uma introdução ao espaço amostral de um processo estocástico discreto e sua interpretação como espaço das sequências, definimos a álgebra de eventos nesse espaço e probabilidade para o espaço das sequências.

No Capítulo 3, apresentamos o conceito de cadeias de Markov, a partir da definição das probabilidades de transição p_{ij}. Fizemos um exemplo onde buscamos estudar a convergência das distribuições de probabilidade via recorrência linear não-homogênea de primeira ordem para sequências. Ainda, para este exemplo, verificamos que ele é de fato uma cadeia de Markov.

Uma cadeia de Markov é um processo que está completamente definido a partir do momento em que se conhece as probabilidades de transição e a distribuição inicial de probabilidades dos estados. Veremos no Capítulo 4 que associamos ao processo de Markov uma matriz de probabilidades de transição T, onde as entradas dessa matriz são dadas pelas probabilidades de transição p_{ij}. Como nosso objetivo é obter previsões a longo prazo, veremos que, com base na aproximação dos elementos da matriz T^n quando T é regular, podemos encontrar uma convergência a uma matriz fixa M a medida que os valores de n aumentam. Portanto, iremos verificar sob quais condições uma matriz de probabilidades de transição se aproximará de uma determinada matriz fixa M. Para isso, utilizaremos como ferramenta conhecimentos de álgebra linear a fim de estudarmos a convergência das matrizes de transição via autovetores e autovalores até chegarmos no importante Teorema de Perron - Frobenius que nos dá uma condição de convergência para matrizes de transição.

No último capítulo, buscamos apresentar alguns exemplos do contexto de cadeias de Markov para serem trabalhados com alunos do Ensino Médio. Inicialmente, apresentamos ao leitor uma versão do Teorema de Perron - Frobenius, no qual não utiliza conhecimentos avançados de álgebra linear para determinar previsões a longo prazo. Fazendo uso deste Teorema, poderemos encontrar tais previsões utilizando apenas conceitos trabalhados no Ensino Médio: matrizes, probabilidade e sistemas lineares. Em seguida, oferecemos uma proposta de plano de aula para a aplicação dos problemas apresentados anteriormente.

Capítulo 1

Uma introdução à teoria de probabilidade

Neste capítulo, apresentamos ao leitor uma abordagem de alguns tópicos da teoria de probabilidade, especialmente as definições de álgebra, σ-álgebra, probabilidade e variável aleatória. Estes tópicos serão necessários para o desenvolvimento do conceito de cadeias de Markov que veremos adiante. Aqui, seguimos como referência o livro [1].

Definição 1.0.1. *O conjunto de todos os resultados possíveis de um experimento aleatório será chamado espaço amostral e denotado por Ω. Todo subconjunto $A \subset \Omega$ será chamado evento. Um evento A ao qual atribuímos uma probabilidade será chamado evento aleatório.*

No lançamento de um dado, temos por exemplo

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

E o evento: “a face voltada para cima é uma número par”é dado por

$$A = \{2, 4, 6\}.$$

Note que, pela definição acima, Ω é o evento certo e $\emptyset$ o evento impossível.

Iremos supor que os eventos aleatórios possuem certas propriedades básicas e intuitivas, que serão essenciais para o desenvolvimento posterior da teoria e do cálculo de probabilidades.

Definição 1.0.2. *Seja* $\mathbb{A}$ *um conjunto de eventos aleatórios de* $\Omega \neq \emptyset$ *, satisfazendo as seguintes propriedades :*

$P1$. $\Omega \in \mathbb{A}$.

$P2$. *Se* $A \in \mathbb{A}$*, então* $A^c \in \mathbb{A}$.

$P3$. *Se* $A \in \mathbb{A}$ *e* $B \in \mathbb{A}$*, então* $A \cup B \in \mathbb{A}$ *.*

O conjunto $\mathbb{A}$ *é chamado álgebra de subconjuntos de* Ω*.*

Proposição 1.0.3. *Seja* $\mathbb{A}$ *uma álgebra de subconjuntos de* Ω*. Então valem as seguintes propriedades:*

$P4$. $\emptyset \in \mathbb{A}$ *;*

$P5$. *Para todo* $n \in \mathbb{N}$ *e para toda sequência de subconjuntos* $A_1, A_2, ..., A_n \in \mathbb{A}$*, temos* $\bigcup_{i=1}^{n} A_i \in \mathbb{A}$ *e* $\bigcap_{i=1}^{n} A_i \in \mathbb{A}$*.*

Demonstração. Por $P1$ temos que $\Omega \in \mathbb{A}$ e por $P2$ $\Omega^c = \emptyset \in \mathbb{A}$, mostrando a validade de $P4$.

Pelo princípio da indução finita, mostraremos que $\bigcup_{i=1}^{n} A_i \in \mathbb{A}$.

Para n=2, temos por $P3$ que $A_1 \cup A_2 \in \mathbb{A}$.

Suponhamos a validade para n, ou seja, $\bigcup_{i=1}^{n} A_i \in \mathbb{A}$ e mostremos a validade pra n+1. Novamente, por $P3$, como $\bigcup_{i=1}^{n} A_i \in \mathbb{A}$ e $A_{n+1} \in \mathbb{A}$ segue que $\bigcup_{i=1}^{n+1} A_i \in \mathbb{A}$.

Agora, para mostrar que $\bigcap_{i=1}^{n} A_i \in \mathbb{A}$, observemos que

$$\bigcap_{i=1}^{n} A_i = \left(\bigcup_{i=1}^{n} A_i^c \right)^c.$$

Como $A_i \in \mathbb{A}, i = 1, 2, ..., n$, segue que $A_i^c \in \mathbb{A}, i = 1, 2, ..., n$. Logo, $\bigcup_{i=1}^{n} A_i^c \in \mathbb{A}$ pela parte já provada de $P5$ e assim, $\bigcap_{i=1}^{n} A_i = \left(\bigcup_{i=1}^{n} A_i^c \right)^c \in \mathbb{A}$. $\square$

Esta proposição diz que uma álgebra é fechada para um número finito de aplicações das operações $\cup, \cap$ *e* ^{c}.

Considere a seguinte propriedade para eventos aleatórios:

$P3'$. Se $A_n \in \mathbb{A}$ para n = 1,2,3,..., então $\bigcup_{n=1}^{\infty} A_n \in \mathbb{A}$.

Definição 1.0.4. *Uma classe $\mathbb{A}$ de subconjuntos de um conjunto não vazio Ω satisfazendo $P1, P2$ e $P3'$ é chamada σ-álgebra de subconjuntos de Ω.*

Observação 1.0.5. *Uma σ-álgebra é sempre uma álgebra, pois $P3$ é consequência de $P3'$, já que $A \cup B = A \cup B \cup B \cup B ... \in \mathbb{A}$ se $\mathbb{A}$ é σ-álgebra.*

Proposição 1.0.6. *Seja $\mathbb{A}$ uma σ-álgebra de subconjuntos de Ω. Se $A_1, A_2, ... \in \mathbb{A}$, então $\bigcap_{n=1}^{\infty} A_n \in \mathbb{A}$.*

Demonstração. Basta observar que

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c$$

e aplicar $P2$ e $P3'$. □

Podemos dizer, então, que uma σ-álgebra é fechada para um número enumerável de aplicações com as operações $\cup, \cap$ *e* ^{c}.

Por exemplo, se $\mathbb{A} = 2^{\Omega}$ é o conjunto das partes de Ω. Temos que 2^{Ω} é uma σ-álgebra de subconjuntos de Ω. De fato, pela própria definição de conjunto das partes segue que as propriedades $P1$, $P2$ e $P3'$ são satisfeitas, isto é

- $\Omega \in 2^{\Omega}$;
- Se $A \in 2^{\Omega}$ então $A^c \in 2^{\Omega}$;
- Se $A_n \in 2^{\Omega}$ para $n = 1, 2, 3, ...$, então $\bigcup_{n=1}^{\infty} A_n \in 2^{\Omega}$.

1.1 Probabilidade

Uma maneira de definir uma probabilidade P no conjunto dos eventos aleatórios é a chamada "frequentista"ou "estatística". Neste contexto, definimos a probabilidade de um evento aleatório A, $P(A)$, como o limite da frequência relativa da ocorrência de A em n repetições independentes do experimento, com n tendendo ao infinito, isto é,

$$P(A) = \lim_{n\to\infty} \frac{1}{n} \times \ \{\text{número de ocorrências de A em } n \text{ ``ensaios''independentes}\}.$$

Esta maneira "frequentista" não é única para se definir probabilidade. Em nosso trabalho, usaremos uma definição axiomática para a probabilidade que se deve a Kolmogorov. Iremos tomar as probabilidades na σ-álgebra $\mathbb{A}$ de eventos, com isso, iremos associar a todo $A \in \mathbb{A}$ um número real $P(A)$. Mais precisamente temos:

Definição 1.1.1. *Uma função $P : \mathbb{A} \to \mathbb{R}$ satisfazendo os axiomas 1,2 e 3 a seguir é chamada probabilidade finitamente aditiva.*

Axioma 1 $P(A) \geq 0$*;*

Axioma 2 $P(\Omega) = 1$*;*

Axioma 3 *Se $A_1, ..., A_n \in \mathbb{A}$ são disjuntos dois a dois, então*

$$P\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} P(A_k).$$

Se ainda a função P satisfaz,

Axioma 3' *Se $A_1, A_2, ... \in \mathbb{A}$ são disjuntos, então*

$$P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} P(A_n),$$

então P é chamada de medida de probabilidade ou simplesmente uma probabilidade. Quando P não satisfaz apenas o Axioma 2 dizemos apenas que P é uma medida em $\mathbb{A}$.

Definição 1.1.2. *Um espaço de probabilidade é uma tripla $(\Omega, \mathbb{A}, P)$, onde*

i) Ω é um conjunto não-vazio;

ii) $\mathbb{A}$ é uma σ-álgebra de subconjuntos de Ω;

iii) P é uma probabilidade em $\mathbb{A}$.

Exemplo 1. *Temos a seguir alguns exemplos de espaços de medida (de probabilidade):*

i) Seja Ω um conjunto qualquer e tome $\mathbb{A} = \{\emptyset, \Omega\}$. Defina $P(\emptyset) = 0$ e $P(\Omega) = 1$. Então $(\Omega, \mathbb{A}, P)$ é um espaço de probabilidade trivial.

ii) Considere Ω sendo um conjunto dos 6 possíveis resultados no lançamento de um dado. Escolhe-se peso de $\frac{1}{6}$ para cada um dos 6 pontos, e para qualquer subconjunto de Ω escolhe-se como medida a soma dos pesos dos pontos do conjunto. Então $\mathbb{A}$ é a família de todos os subconjuntos de Ω, e $(\Omega, \mathbb{A}, P)$ é um espaço de probabilidade.

iii) Seja $\Omega = \{0, 1, 2, 3, ...\}$ e π um vetor da forma $\pi = (\pi_0, \pi_1, \pi_2, ...)$ com $\pi_i \geqslant 0$. E defina o peso de cada elemento i de Ω como $m(i) = \pi_i$. Para cada subconjunto de Ω escolha como peso deste conjunto a soma de todos os pesos dos pontos neste conjunto. Então $\mathbb{A}$ é a família de todos os subconjuntos de Ω e $(\Omega, \mathbb{A}, m)$ é um espaço de medida com $m(\Omega) = \sum_{i=0}^{\infty} \pi_i$. Quando $m(\Omega) = 1$, m é uma medida de probabilidade.

1.1.1 Propriedades da probabilidade

Seja P uma probabilidade em uma σ-álgebra $\mathbb{A}$. Suponhamos que todo conjunto A abaixo pertença a $\mathbb{A}$. Então as seguintes propriedades são consequências dos axiomas acima:

Propriedade 1: $P(A^c) = 1 - P(A)$.

De fato, pelo Axioma 2, $P(\Omega) = 1$ e $\Omega = A \cup A^c$.

Como A e A^c são disjuntos, segue do Axioma 3 que

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

Logo, $P(A^c) = 1 - P(A)$.

Em particular, $P(\emptyset) = 1 - P(\Omega) = 0$.

Propriedade 2: $0 \leq P(A) \leq 1$.

Esta propriedade é uma consequência do Axioma 1 e da Propriedade 1.

Propriedade 3: Se $A_1 \subset A_2$ então $P(A_1) \leq P(A_2)$.

Note que, $A_2 = A_1 \cup (A_2 - A_1)$. Pelo Axioma 3, $P(A_2) = P(A_1) + P(A_2 - A_1) \geqslant P(A_1)$.

Propriedade 4: $P\Big(\bigcup\limits_{i=1}^{n} A_i\Big) \leq \sum\limits_{i=1}^{n} P(A_i)$.

Vamos utilizar o princípio da indução finita para verificar a desigualdade acima.

Para $n = 2$, temos:

Pelo Axioma 3, $P(A_1 \cup A_2) = P(A_1) + P(A_2 \cap A_1^c) \leq P(A_1) + P(A_2)$, por P_3, já que $(A_2 \cap A_1^c) \subset A_2$. Agora, suponhamos a validade para n, ou seja, $P\Big(\bigcup\limits_{i=1}^{n} A_i\Big) \leq \sum\limits_{i=1}^{n} P(A_i)$ e mostremos a validade para $n+1$.

Observe que

$$P\Big(\bigcup_{i=1}^{n+1} A_i\Big) = P\Big(\Big(\bigcup_{i=1}^{n} A_i\Big) \cup A_{n+1}\Big).$$

Pelo caso $n = 2$ e pela hipótese de indução segue que:

$$\begin{aligned} P\Big(\bigcup_{i=1}^{n+1} A_i\Big) &= P\Big(\Big(\bigcup_{i=1}^{n} A_i\Big) \cup A_{n+1}\Big) \\ &\leq P\Big(\bigcup_{i=1}^{n} A_i\Big) + P(A_{n+1}) \\ &\leq \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) \\ &= \sum_{i=1}^{n+1} P(A_i). \end{aligned}$$

Assim, pelo princípio da indução finita temos a validade para todo n.

Propriedade 5: $P\Big(\bigcup\limits_{i=1}^{\infty} A_i\Big) \leq \sum\limits_{i=1}^{\infty} P(A_i)$.

Note que

$$A_1 \cup A_2 \cup A_3 \cup ... = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup (A_4 \setminus (A_1 \cup A_2 \cup A_3)) \cup ...,$$

onde $A_1, (A_2 \setminus A_1), (A_3 \setminus (A_1 \cup A_2)), (A_4 \setminus (A_1 \cup A_2 \cup A_3))$,... são disjuntos.

Assim,

$$
\begin{aligned}
P\Big(\bigcup_{i=1}^{\infty} A_i\Big) &= P(A_1 \cup A_2 \cup A_3 \cup ...) \\
&= P(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup ...) \\
&\underbrace{=}_{Axioma3'} P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus (A_1 \cup A_2)) + ... \\
&\leq P(A_1) + P(A_2) + P(A_3) + ... \\
&= \sum_{i=1}^{\infty} P(A_i).
\end{aligned}
$$

Onde a desigualdade se deve ao fato de que $(A_2 \setminus A_1) \subset A_2, (A_3 \setminus (A_1 \cup A_2)) \subset A_3...$ e pela Propriedade 3.

Como iremos trabalhar com cadeias de Markov, a definição a seguir será bastante utilizada nos demais capítulos.

Definição 1.1.3. *Seja $(\Omega, \mathbb{A}, P)$ um espaço de probabilidade. Se $B \in \mathbb{A}$ e $P(B) > 0$, a probabilidade condicional de A dado B é definida por*

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

Teorema 1.1.4 (Probabilidade total)**.** *Se a sequência (finita ou enumerável) de eventos aleatórios $A_1, A_2, ...$ formar uma partição de Ω, então*

$$P(B) = \sum_i P(A_i)P(B|A_i),$$

para todo $B \in \mathbb{A}$.

Demonstração. Suponhamos que A_1,A_2,... formam uma partição de Ω, isto é, os A_i são disjuntos e $\bigcup_{i=1}^{\infty} A_i = \Omega$.

Para todo evento $B \in \mathbb{A}$, temos $B = \bigcup_i (B \cap A_i)$. Como os A_i são disjuntos, então os $B \cap A_i$ são disjuntos e

$$P(B) = P\Big(\bigcup_i (B \cap A_i)\Big) = \sum_i P(B \cap A_i) = \sum_i P(A_i)P(B|A_i).$$

□

1.2 Variável aleatória

Definição 1.2.1. *Uma função $X : \Omega \longrightarrow \mathbb{R}$ é denominada variável aleatória, se o conjunto $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}$ é um evento aleatório (isto é, pertence a σ-álgebra $\mathbb{A}$ de eventos de Ω), para todo $x \in \mathbb{R}$.*

Note que se X é uma variável aleatória, então $\{X = x\} = \{\omega \in \Omega : X(\omega) = x\}$ pertence a σ-álgebra $\mathbb{A}$ de eventos em Ω, para todo $x \in \mathbb{R}$. De fato,

se X é uma variável aleatória, então

$$\left\{X \leq x + \frac{1}{n}\right\} \in \mathbb{A} \;\text{ e }\; \left\{X \leq x - \frac{1}{n}\right\} \in \mathbb{A};\; n \in \mathbb{N}.$$

Assim, como $\mathbb{A}$ é uma σ-álgebra de eventos,

$$\left\{X \leq x - \frac{1}{n}\right\}^c = \left\{X > x - \frac{1}{n}\right\} \in \mathbb{A}$$

e

$$\left\{X > x - \frac{1}{n}\right\} \cap \left\{X \leq x + \frac{1}{n}\right\} = \left\{x - \frac{1}{n} < X \leq x + \frac{1}{n}\right\} \in \mathbb{A}.$$

Logo,

$$\bigcap_{n=1}^{\infty} \left\{x - \frac{1}{n} < X \leq x + \frac{1}{n}\right\} = \{X = x\} \in \mathbb{A}.$$

Além disso, se Ω é um conjunto finito, vale notar que pela própria definição de variável aleatória, se a $\sigma-$álgebra de eventos em Ω for o conjunto das partes de Ω, então qualquer função em Ω é uma variável aleatória, pois $\{X \leq x\} \in 2^{\Omega}$ (conjunto das partes de Ω), para todo $x \in \mathbb{R}$.

Exemplo 2. *Considere que uma moeda é lançada duas vezes sucessivamente.*
Seja C cara e K coroa. O espaço amostral deste experimento é:

$$\Omega = \{(C, C); (C, K); (K, C); (K, K)\}.$$

Tomemos a $\sigma-$álgebra deste exemplo como sendo o conjunto das partes 2^{Ω}. Podemos definir a função X como sendo o "número de caras obtidas nos dois lançamentos". Por exemplo, temos que X((C,C)) = 2 e X((K,C))=1. Note que, $X(\omega) \in \{0, 1, 2\}, \forall \omega \in \Omega$.

A função $X : \Omega \longrightarrow \{0, 1, 2\}$ é tal que $\{\omega : X(\omega) \leq x\} \in 2^{\Omega}$ para todo $x \in \mathbb{R}$. Por exemplo, se x = 1 temos que $\{\omega : X(\omega) \leq 1\} = \{(C, K), (K, C), (K, K)\} \in 2^{\Omega}$ e se x = -2, temos $\{\omega : X(\omega) \leq -2\} = \emptyset \in 2^{\Omega}$.

Logo, X é uma variável aleatória.

Definição 1.2.2. *Seja X uma variável aleatória. Se o número de valores possíveis de X for enumerável (finito ou infinito), dizemos que X é uma variável aleatória discreta. Ou seja, a imagem da função é $Im(X) = \{x_1, x_2, ...\} \subseteq \mathbb{R}$.*

O exemplo anterior é um caso de variável aleatória discreta. Vejamos a seguir outro exemplo.

Exemplo 3. *Após um exame médico, as pessoas podem ser diagnosticadas com zika vírus (Z) ou não (N). Admita que três pessoas sejam escolhidas ao acaso e classificadas de acordo com esse esquema. Temos que o espaço amostral é dado por:*

$$\Omega = \{ZZZ, ZZN, ZNN, ZNZ, NZN, NZZ, NNZ, NNN\}.$$

Considere a $\sigma-$álgebra deste exemplo como sendo o conjunto das partes 2^{Ω}. Então a função X, sendo o "número de pessoas diagnosticadas com o zika vírus", é uma variável aleatória. O conjunto dos possíveis valores de X é $\{0, 1, 2, 3\}$, ou seja, X é uma variável aleatória discreta.

Definição 1.2.3. *Seja X uma variável aleatória. Suponha que os valores possíveis de X seja qualquer valor numérico em um determinado intervalo, ou, coleção de intervalos. Então, diremos que X é uma variável aleatória contínua.*

Exemplo 4. *Escolher um ponto ao acaso em [0, 1], e seja X o valor do resultado. Então $\Omega = [0, 1]$ e $X(\omega) = \omega$. Note que os valores possíveis (resultados) para X não é enumerável, pois é o intervalo $[0, 1]$. No caso contínuo, em geral, usa-se a $\sigma-$álgebra de Borel. A grosso modo, a σ-álgebra de Borel na reta é a menor σ-álgebra contendo todos os intervalos, que são gerados pelas operações $\cup, \cap$ e ^{c} de intervalos abertos da reta. É possível mostrar, mas foge dos objetivos dessa dissertação, que a função X é uma variável aleatória contínua com a σ-álgebra de Borel, que toma valores em $[0, 1]$.*

Capítulo 2

Espaço de probabilidade para cadeias de Markov

Neste capítulo, buscamos construir um espaço de probabilidade para uma sequência de variáveis aleatórias em tempo discreto, que será útil para o cálculo de probabilidades envolvendo cadeias de Markov. Esta sequência de variáveis aleatórias será denominada como processo estocástico discreto.

Considere o lançamento de uma moeda não viciada no tempo $n = 1, 2, 3, ...T$. As faces da moeda são: cara(C) ou coroa(K). Denotando por Ω o conjunto de todos os possíveis resultados dos lançamentos até o tempo T, temos:

$$\Omega = \{\omega : \omega = (e_1, e_2, e_3, ..., e_T)\} \quad \text{onde, } e_n = C \; ou \; K, \; para \; n = 1, 2, ..., T.$$

O conjunto $\mathcal{E} = \{C, K\}$ que representa os possíveis estados da moeda (possíveis resultados em cada lançamento) será chamado de espaço de estados.

Para modelar incertezas sobre os resultados, nós listamos todos os resultados possíveis, e o chamamos de estados possíveis. Conforme o tempo passa, mais e mais informações são reveladas sobre os estados que já ocorreram. No tempo $n = 2$, por exemplo, nós conhecemos os resultados e_1 e e_2. Então, a informação que temos é

$$A = \{(e_1, e_2, *, ..., *)\} \subset \Omega,$$

onde “*” indica os estados que ainda poderão ocorrer com o decorrer do tempo.

Esta ideia de informação disponível no tempo n, iremos denotar por F_n, que consiste nos resultados que ocorreram antes e no tempo n. F_n é chamada Álgebra de Eventos.

Voltando ao exemplo anterior, para $T = 2$, temos que no tempo $n = 0$, antes de realizarmos o experimento, nós não temos informações sobre e_1 e e_2, e portanto $F_0 = \{\emptyset, \Omega\}$. Tudo que nós sabemos é que os estados possíveis de ocorrer estão em Ω. Considere agora o caso $n = 1$. Suponha que em $n = 1$ o resultado obtido foi cara (C). Então sabemos que os estados possíveis de ocorrer (nossa informação) estão em A, e não estão no seu complementar A^c, onde

$$A = \{(C, e_2), e_2 = C \text{ ou } K\} = \{(C, C), (C, K)\}.$$

Assim, no tempo $n = 1$, temos a seguinte álgebra de eventos

$$F_1 = \{\emptyset, \Omega, A, A^c\}.$$

Note que $F_0 \subset F_1$.

Observação 2.0.1. *F_n é uma álgebra de subconjuntos de Ω, ou seja, tem as seguintes propriedades:*

$P1$. $\Omega \in F_0 \subset F_n$, para $n = 1, 2, ...T$.

$P2$. Se $A \in F_n$, então $A^c \in F_n$.

$P3$. Se $A \in F_n$ e $B \in F_n$, então $A \cup B \in F_n$.

Exemplo 5. *A seguir temos alguns exemplos de álgebras de eventos:*

1. *$F_0 = \{\Omega, \emptyset\}$ é chamada álgebra trivial.*

2. *$\{\emptyset, \Omega, A, A^c\}$ é chamada álgebra gerada por A e denotada por F_A.*

3. *$\{A : A \in \Omega\}$ é chamada a álgebra de todos os subconjuntos de Ω e é denotada por 2^Ω.*

Se F é uma álgebra de eventos, então todo conjunto de F é chamado de conjunto mensurável. Se $F = 2^\Omega$, por exemplo, então todo subconjunto de Ω é mensurável.

Definição 2.0.2. *Uma filtração $\mathbb{F}$ é a coleção de álgebras de eventos F_n, com*

$$\mathbb{F} = \{F_0, F_1, ..., F_n, ..., F_T\}; \quad F_n \subset F_{n+1}.$$

A filtração $\mathbb{F}$ é utilizada para modelar um fluxo de informações que são obtidas com o decorrer dos experimentos. Conforme o tempo passa, teremos as informações mais detalhadas, isto é, partições de Ω cada vez mais finas.

Exemplo 6. *Temos que* $\mathbb{F} = \{F_0, F_A, 2^\Omega\}$ *é uma filtração, pois* F_0, F_A *e* 2^Ω *são álgebras de eventos, e ainda* $F_0 \subset F_A \subset 2^\Omega$*.*

2.1 Processo estocástico

Como vimos no capítulo anterior, se a função X é uma variável aleatória em (Ω, F), com Ω sendo um conjunto finito, então todos os conjuntos $\{X = x_i\} = \{\omega \in \Omega : X(\omega) = x_i\}$, i = 1,...,k pertencem a F. Isto significa que, se temos a informação descrita por F, isto é, conhecemos qual evento da forma $\{X = x_i\}$ ocorreu em F, então sabemos qual valor de X ocorreu. Lembramos que se $F = 2^\Omega$, então qualquer função em Ω é uma variável aleatória.

A fim de definir processo estocástico, daremos o seguinte exemplo.

Exemplo 7. *Considere o exemplo do lançamento de uma moeda não viciada, para* $n = 1, 2$*. Então* $\Omega = \{\omega_1 = (C, C), \omega_2 = (C, K), \omega_3 = (K, C), \omega_4 = (K, K)\}$*.*

Tome $A = \{\omega_1, \omega_2\}$*, o qual representa o evento em que no tempo* $n = 1$ *o resultado obtido foi cara. Assim,* $F_1 = \{\emptyset, \Omega, A, A^c\}$ *e tome* $F_2 = 2^\Omega$*.*

Considere as seguintes funções em Ω*:*

a) $X(\omega_1) = X(\omega_2) = 15$ *e* $X(\omega_3) = X(\omega_4) = 5$*.*

X *é uma variável aleatória em* F_1*, já que* $\{\omega : X(\omega) = 15\} = \{\omega_1, \omega_2\} = A \in F_1$ *e* $\{\omega : X(\omega) = 5\} = \{\omega_3, \omega_4\} = A^c \in F_1$*.*

b) $Y(\omega_1) = 15, Y(\omega_2) = 75, Y(\omega_3) = 75,$ *e* $Y(\omega_4) = 5$*.*

Y *não é uma variável aleatória em* F_1*, pois* $\{\omega : Y(\omega) = 75\} = \{\omega_2, \omega_3\} \notin F_1$*. Por outro lado, é claro que* Y *é uma variável aleatória em* F_2*.*

Definição 2.1.1. *Um processo estocástico adaptado a filtração* $\mathbb{F} = \{F_0, F_1, ..., F_T\}$ *é uma sequência de variáveis aleatórias* (X_n) *em* (Ω, F_n)*, para qualquer* n *fixado,* $n = 0, 1, ..., T$*. Note que, em particular, para cada* n*,* X_n *é uma variável aleatória em* (Ω, F_T)*.*

Exemplo 8. *Seja* $X_1 = X$ *e* $X_2 = Y$ *como apresentados no Exemplo 7. A sequência* $\{X, Y\}$ *é um processo estocástico adaptado a filtração* $\mathbb{F} = \{F_1, F_2\}$.

Definição 2.1.2. *Seja* $(\Omega, 2^{\Omega})$ *um espaço amostral com a álgebra de todos os eventos, e* X *uma variável aleatória com valores* x_i, i = *1,2,...k. Considere os conjuntos*

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

Esses conjuntos formam uma partição de Ω, *e a álgebra gerada por essa partição é chamada álgebra gerada por* X. *Esta é a menor álgebra que contém todos os conjuntos da forma* $A_i = \{X = x_i\}$ *e é denotada por* F_X. *A álgebra gerada por* X *representa a informação que podemos extrair observando* X.

Exemplo 9. *Considerando o Exemplo 7, temos* $\{\omega : X(\omega) = 15\} = \{\omega_1, \omega_2\} = A$ *e* $\{\omega : X(\omega) = 5\} = \{\omega_3, \omega_4\} = A^c$. *Logo,* $F_X = F_1 = \{\emptyset, \Omega, A, A^c\}$.

2.2 Probabilidade no espaço das sequências

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

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

onde cada $e_i \in \mathcal{E}$. Os pontos $\omega \in \Omega$ são chamados de caminhos, o espaço Ω é chamado de espaço das sequências, e o valor e_n em um caminho é chamado n-ésimo "resultado"em ω. Por exemplo, se $\mathcal{E} = \{1, 2\}$, então $\Omega = \{(e_0, e_1, e_2, e_3, \dots); e_i = 1 \ ou \ 2\}$ e o elemento $(1, 2, 1, 1, 1, 2, \dots) \in \Omega$ é um caminho ou realização de experimentos.

A função $X_n : \Omega \to \mathcal{E}$, dada por

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

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

Fixado n, seja F_n a família de todas as uniões de Ω da forma

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

onde $\mathcal{E}_0, \mathcal{E}_1, \ldots \mathcal{E}_n$ são subconjuntos do espaço de estados $\mathcal{E}$. É possível verificar que cada F_n é uma σ-álgebra, mas foge dos objetivos deste trabalho. Além disso, $F_n \subset F_{n+1}$ forma uma filtração natural na qual o processo X_n é adaptado. Considere agora $\mathcal{F}$ a família de conjuntos definida por

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

Cada elemento em $\mathcal{F}$ é um conjunto de trajetórias para as quais um número finito de entradas da sequência são restritas a pertencer a certos subconjuntos de $\mathcal{E}$, e as demais infinitas entradas são irrestritas. Um conjunto de $\mathcal{F}$ é chamado de *cilindro*. Apesar de $\mathcal{F}$ ser uma álgebra, não é uma σ-álgebra. Porém, existe a menor σ-álgebra $\mathcal{G}$, tal que $\mathcal{F} \subset \mathcal{G}$. Associado a $\mathcal{G}$ existe uma única medida de propabilidade μ, tal que

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

é dado pelo produto das probabilidades condicionais entre os estados $e_0, e_1, \ldots, e_n$. Cada conjunto C_i^n é chamado de *cilindro básico* de $\mathcal{F}$ e vale notar que cada $C \in F_n$ é da forma $C = \bigcup_i C_i^n$. Ou seja, construímos um espaço de probabilidade $(\Omega, \mathcal{G}, \mu)$, onde μ é dada por

$$\mu(e_0, e_1, \ldots, e_n, *, *, *, \ldots) = p_0 p_{10} p_{21} \ldots p_{n(n-1)},$$

onde p_0 é a probabilidade do processo iniciar no estado e_0 e $p_{ij} = \mu\{\omega \in \Omega : X_i = e_i | X_j = e_j\}$ são probabilidades condicionais. Para maiores detalhes ver por exemplo, [2, p.43].

Capítulo 3

Cadeias de Markov

Uma cadeia de Markov é um processo estocástico $(X_n)_{n\in\mathbb{N}}$, associado a um espaço de probabilidade $(\Omega, \mathbb{A}, P)$ que satisfaz uma determinada propriedade (ver equação 3.2.0.1) que iremos apresentar mais adiante neste capítulo. Daremos a seguir algumas definições e propriedades básicas das probabilidades de transição que serão úteis para a compreensão de tal propriedade markoviana.

3.1 Probabilidade de transição

Definição 3.1.1. *A probabilidade de transição em um passo, denotada por $p_{ij}(k)$, é definida como a seguinte probabilidade condicional:*

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

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

Se a probabilidade de transição $p_{ij}(k)$ numa cadeia não depende do tempo k, dizemos que ela é *homogênea*. Neste caso, usaremos a notação p_{ij}. Se as probabilidades de transição dependem do tempo k, dizemos que elas são *não-estacionárias* ou *não homogêneas*. Neste trabalho, estamos interessados em estudar somente os casos homogêneos.

Para uma cadeia de Markov com um número finito de estados, associa-se uma matriz de transição T, que será dada pelas probabilidades de transição, isto é,

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

Definição 3.1.2. *A probabilidade de transição em n-passos ($n \geq 0$), denotada por $p_{ij}^{(n)}$, é a probabilidade de transferência do estado j para o estado i em n etapas de tempo discreto, isto é,*

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

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

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

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

Em termos matriciais, essas equações podem ser escritas da forma

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

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

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

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

Além disso, um outro aspecto interessante em conhecer T^n, é que o vetor v^n de distribuição de probabilidade do processo, no tempo n, é igual ao produto matricial $T^n v^0$, com a distribuição inicial v^0 escrita de forma transposta. No próximo capítulo iremos explorar a relação entre a distribuição v^n e $T^n v^0$ com mais detalhes.

3.2 Cadeias de Markov

Vejamos alguns exemplos, que se encontram em [3] e [4], que servirão de ilustração para compreendermos o contexto de cadeia de Markov.

Considere um interruptor com dois estados: ligado e desligado. No começo do experimento, o interruptor está ligado. A cada minuto depois, jogamos um dado. Se o resultado do dado mostrar 6, mudamos o estado do interruptor, caso contrário, deixamos como está. O estado do interruptor em função do tempo é um processo de Markov. Este simples exemplo nos permite explicar o significado: "não tem nenhuma memória". Se conhecemos o estado do interruptor no tempo n, nós podemos prever sua evolução (em termos de variável aleatória) para todos os tempos futuros, sem requerer conhecimento sobre o estado do interruptor em tempos menores que n. Em outras palavras, o futuro do processo depende do presente, mas independe do passado.

Denote por $x_1^{(n)}$ a probabilidade do interruptor estar ligado no tempo n. Similarmente, seja $x_2^{(n)}$ a probabilidade do interruptor estar desligado no tempo n. A partir das informações anteriores, temos a seguinte relação de recorrência:

$$x_1^{(n+1)} = \frac{5}{6}x_1^{(n)} + \frac{1}{6}x_2^{(n)}, \quad x_2^{(n+1)} = \frac{1}{6}x_1^{(n)} + \frac{5}{6}x_2^{(n)},$$

com $x_1^{(0)} = 1$ e $x_2^{(0)} = 0$. A primeira igualdade vem do fato de que o interruptor ficará ligado no tempo $n+1$, se ele estava ligado no tempo n e o dado não mostrou um 6, ou ele estava desligado no tempo n e o dado mostrou o número 6. A segunda igualdade é análoga.

A equação anterior pode ser escrita na forma matricial:

$$x^{(n+1)} = \begin{pmatrix} x_1^{(n+1)} \\ x_2^{(n+2)} \end{pmatrix} = \begin{pmatrix} \frac{5}{6}x_1^{(n)} + \frac{1}{6}x_2^{(n)} \\ \frac{1}{6}x_1^{(n)} + \frac{5}{6}x_2^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{pmatrix} \cdot \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix} . x^{(n)}.$$

Considere novamente o interruptor que tem dois estados e no começo do experimento está ligado. E vamos supor que é jogado um dado a todo minuto. Contudo, desta vez, nós mudamos o estado do interruptor somente se o dado mostrar o número 6 e não ter ocorrido o número 6 no tempo anterior. Assim, o futuro do processo depende do que ocorreu no passado.

A cadeia de Markov é um processo estocástico caracterizado por seu estado futuro depender apenas do seu estado no presente, sendo que os estados passados não influenciam no estado futuro. Dizemos então que um processo de Markov "não tem nenhuma memória". Analisando os exemplos anteriores, podemos dizer que o primeiro exemplo, onde o estado do interruptor depende apenas do resultado do dado lançado naquele minuto, é uma cadeia de Markov, enquanto que o segundo exemplo não é uma cadeia de Markov, pois o estado do interruptor depende do resultado do dado nos tempos anteriores.

Formalmente temos o seguinte:

Definição 3.2.1. *Considere um espaço de estados com um número finito (ou enumerável) de elementos $\mathcal{E} = \{e_1, ..., e_n\}$. Um processo estocástico discreto $(X_n)_{n\in\mathbb{N}}$ é uma cadeia (ou processo) de Markov se a probabilidade condicional satisfizer*

$$P(X_{n+1} = x_{n+1}|X_0 = x_0, ..., X_n = x_n) = P(X_{n+1} = x_{n+1}|X_n = x_n), \tag{3.2.0.1}$$

para todo $n \geq 1$ e para toda sequência $x_0, x_1, ..., x_{n+1}$ de elementos do espaço de estados $\mathcal{E}$. Essa condição (3.2.0.1) significa, em linguagem natural, pensando que n indica o tempo, que o futuro do processo, uma vez conhecido o estado presente, é independente do passado.

As probabilidades condicionais, como já mencionamos anteriormente,

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

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

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

para todo natural n, a cadeia de Markov é dita estacionária.

Um processo de Markov está completamente definido a partir do momento em que se especifica as probabilidades de transição e a distribuição inicial de probabilidades dos estados.

A propriedade (3.2.0.1) será geralmente referida como a propriedade da cadeia de Markov. A mesma propriedade também pode ser estendida para processos de Markov com tempo contínuo e espaço de estados sendo um intervalo da reta real, como pode ser visto, por exemplo, em [5].

Agora, iremos estudar convergência das distribuições de probabilidade para uma cadeia finita, cujas probabilidades de transição são todas positivas. Para tal estudo usaremos técnicas de resolução de recorrências lineares não-homogêneas (ver [6, p.73]).

Para tal estudo, apresentamos um modelo que descreve a dinâmica de um telefone que possui dois estados: livre e ocupado. Neste modelo é dado as probabilidades de transição e a distribuição inicial do telefone. Apesar de simples, ele é suficiente e interessante para os propósitos deste capítulo.

Exemplo 10. *Em alguns casos o uso do telefone pode se tornar uma questão bastante interessante. Suponha que se o telefone está livre em um período de tempo, digamos no n-ésimo minuto, com probabilidade p, onde $0 < p < 1$, ele estará ocupado no próximo minuto. Se o telefone está ocupado no n-ésimo minuto, ele estará livre no próximo minuto com probabilidade q, onde $0 < q < 1$. Assuma que o telefone está livre no minuto inicial $(n = 0)$. Estamos interessados em responder as seguintes questões:*

a)Qual é a probabilidade x_n do telefone estar livre no n-ésimo minuto?

b) Qual é o $\lim_{n\to\infty} x_n$, se existir?

Afim de responder as perguntas acima, denotemos por A_n o evento em que o telefone está livre no n-ésimo minuto e seja $B_n = \Omega$ - A_n o seu complementar, isto é, o evento em que o telefone está ocupado no n-ésimo minuto. As condições do exemplo nos dá:

$$P(B_{n+1}|A_n) = p \qquad e \qquad P(A_{n+1}|B_n) = q.$$

Assumimos também que $P(A_0) = 1$, isto é $x_0 = 1$. Utilizando esta notação, temos que $P(A_n) = x_n$. Logo,

$$\begin{aligned} x_{n+1} &= P(A_{n+1}) \\ &= P(A_{n+1}|A_n)P(A_n) + P(A_{n+1}|B_n)P(B_n) \qquad (3.2.0.2) \\ &= (1-p)x_n + q(1-x_n) \\ &= q + (1-p-q)x_n \qquad (3.2.0.3) \end{aligned}$$

A igualdade (3.2.0.2) segue do Teorema 1.1.4.

Note que a igualdade (3.2.0.3) é uma recorrência linear não-homogênea de primeira ordem, ou seja, é uma recorrência do tipo $x_{n+1} = g(n)x_n + h(n)$. Vamos resolvê-la utilizando o seguinte fato.

As recorrências lineares não-homogêneas de primeira ordem que mais facilmente se resolvem, são da forma $x_{n+1} = x_n + f(n)$.

Com efeito, temos

$$x_1 = x_0 + f(0)$$

$$x_2 = x_1 + f(1)$$

$$x_3 = x_2 + f(2)$$

$$x_4 = x_3 + f(3)$$

$$\vdots$$

$$x_n = x_{n-1} + f(n-1)$$

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

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

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

em

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

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

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

ou seja,

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

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

Voltando à recorrência

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

encontrada em nosso problema , vamos resolvê-la utilizando os comentários anteriores.

Inicialmente, vamos encontrar uma solução não nula da recorrência

$$x_{n+1} = (1 - p - q)x_n. \tag{3.2.0.4}$$

Temos:

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

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

$$x_3 = (1-p-q)x_2$$

$$\vdots$$

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

Multiplicando todos os termos de cada lado das igualdades, resulta

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

Logo, tomando uma condição inicial $x_0 = 1$ *temos que* $a_n = (1-p-q)^n$ *é uma solução não nula da recorrência (3.2.0.4). Façamos a substituição de* $x_n = (1-p-q)^n y_n$ *em (3.2.0.3). Obtemos então que*

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

e assim

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

onde $x_0 = (1-p-q)^0 y_0$ *e, portanto,* $y_0 = x_0 = 1$*. Temos então que*

$$y_1 = \frac{q}{(1-p-q)^1} + y_0$$

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

$$y_3 = \frac{q}{(1-p-q)^3} + y_2$$

$$\vdots$$

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

Somando os termos de cada lado das igualdades, resulta em

$$y_n = y_0 + \frac{q}{(1-p-q)} + \frac{q}{(1-p-q)^2} + \frac{q}{(1-p-q)^3} + \ldots + \frac{q}{(1-p-q)^n}$$

$$y_n = 1 + \underbrace{\frac{q}{(1-p-q)} + \frac{q}{(1-p-q)^2} + \frac{q}{(1-p-q)^3} + \ldots + \frac{q}{(1-p-q)^n}}_{\textit{soma dos n primeiros termos de uma P.G. de razão}\frac{1}{1-p-q}}$$

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

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

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

Substituindo em $x_n = (1-p-q)^n y_n$, *temos*

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

Logo, a probabilidade x_n *para que o telefone esteja livre no n-ésimo minuto é dada por*

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

Agora, como $0 < p < 1$ *e* $0 < q < 1$ *segue que* $|1-p-q| < 1$, *e assim* $(1-p-q)^n \to 0$, *quando* $n \to \infty$. *Isto nos dá a resposta do item (b) deste exemplo, isto é,* $\lim_{n\to\infty} x_n = \frac{q}{p+q}$.

Uma abordagem diferente da demonstração acima, poderá ser vista mais adiante no Exemplo 20 do Capítulo 5, na qual utiliza uma versão do Teorema de Perron-Frobenius para estudar a existência do limite (convergência da matriz de transição). Além disso, caso o leitor tenha curiosidade, em [4, p.86], também há uma outra demonstração em que não se usa soluções de recorrências lineares não-homogêneas (como apresentamos neste trabalho), nem Perron-Frobenius.

O objetivo principal agora, é mostrar que de fato o modelo apresentado no Exemplo 10 é uma cadeia de Markov. Para isto, considere o processo estocástico (X_n) como visto no Capítulo 2.

Proposição 3.2.2. *O modelo apresentado no Exemplo 10 é uma cadeia de Markov, isto é, verifica (3.2.0.1).*

Demonstração.

Considere o espaço de estados $\mathcal{E} = \{\mathbf{1}, \mathbf{2}\}$, sendo $\mathbf{1}$ o estado livre e $\mathbf{2}$ o estado ocupado. Tomaremos Ω como sendo o conjunto de todas as sequências.

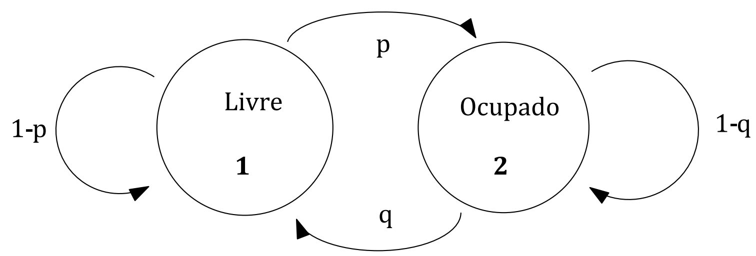

Figura 3.1: Diagrama do Exemplo 10

A fim de construir a probabilidade nos cilindros, tomemos ν^0 como sendo alguma probabilidade em $\mathcal{E}$. Pelas informações do problema temos que $\nu^0(\mathbf{1}) = 1$ *e* $\nu^0(\mathbf{2}) = 0$, pois o telefone está livre no minuto inicial. A medida de probabilidade ν^0 escolhida, corresponde a distribuição inicial do processo estocástico que iremos definir. Podemos então definir a probabilidade P da seguinte forma.

Se $\omega = (e_0, e_1, ...) \in \Omega$, tomemos

$$P(\{\omega \in \Omega : \omega_0 = e_0\}) = \nu^0(\{e_0\});$$

$$P(\{\omega \in \Omega : \omega_i = e_i, i = 0, ..., n+1\}) = p(e_{n+1}|e_n)P(\{\omega \in \Omega : \omega_i = e_i, i = 0, ..., n\}), \tag{3.2.0.5}$$

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

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

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

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

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

E fazendo isso, sucessivamente, para mais estados, nota-se que de fato a fórmula (3.2.0.5) só depende das probabilidades de transição e da distribuição inicial. Além disso, como

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

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

Como no Capítulo 2, considere o processo estocástico $X_n : \Omega \longrightarrow \mathbb{R},\ n \in I\!N$, dado por

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

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

$$P(X_{n+1} = \mathbf{2}|X_n = \mathbf{1}) = p, \tag{3.2.0.6}$$

$$P(X_{n+1} = \mathbf{1}|X_n = \mathbf{2}) = q. \tag{3.2.0.7}$$

Da definição de P temos que

$$\begin{aligned} P(X_{n+1} = \mathbf{2}, X_n = \mathbf{1}) &= P(\{\omega \in \Omega : \omega_n = \mathbf{1}, \omega_{n+1} = \mathbf{2}\}) \\ &= \sum_{e_0,...,e_{n-1} \in E} P(\{\omega \in \Omega : \omega_i = e_i, i = 0, ..., n-1, \omega_n = \mathbf{1}, \omega_{n+1} = \mathbf{2}\}) \\ &= \sum_{e_0,...,e_{n-1} \in E} p(\mathbf{2}|\mathbf{1})P(\{\omega \in \Omega : \omega_i = e_i, i = 0, ..., n-1, \omega_n = \mathbf{1}\}) \\ &= p \sum_{e_0,...,e_{n-1} \in E} P(\{\omega \in \Omega : \omega_i = e_i, i = 0, ..., n-1, \omega_n = \mathbf{1}\}) \\ &= pP(X_n = \mathbf{1}). \end{aligned} \tag{3.2.0.8}$$

A segunda igualdade, em que aparece a soma sobre os estados, segue da Proposição (1.1.4) tomando uma partição em Ω, por meio de conjuntos formados de trajetórias em que se fixa os primeiros n estados $e_0, e_1, \ldots, e_{n-1}$.

Pela definição de probabilidade condicional e da igualdade (3.2.0.8) segue que

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

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

Vamos agora verificar que

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

De fato,

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

Por outro lado, de (3.2.0.6) e (3.2.0.7) temos que

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

□

Apesar de trabalhosa, a demonstração acima poderia ser adaptada para um processo com mais de dois estados, desde que $\mathcal{E}$ seja finito. Boa parte da demonstração acima foi baseada em técnicas que aparecem em [4, p.88].

Capítulo 4

Matrizes de transição agindo em vetores de distribuição

A fim de introduzir as propriedades de vetor de probabilidade, matrizes de transição e suas relações, iniciaremos com o seguinte exemplo.

Considere que em um habitat natural existem 1000 passarinhos e que este habitat foi dividido em quatro regiões denotadas por R_1, R_2, R_3 e R_4. Vamos supor ainda que, a cada dia, um pesquisador anota as informações sobre as posições dos pássaros em cada região, com o intuito de observar a migração em busca de alimento. Não estando interessados em entender o comportamento de cada passarinho, mas a distribuição deles nas regiões R_1, R_2, R_3 e R_4, podemos considerar um vetor com 4 coordenadas indicando a quantidade de passarinhos em cada região em um determinado dia. Por exemplo, se em um determinado dia temos R_1 com 150 passarinhos, R_2 com 225 passarinhos, R_3 com 250 passarinhos e R_4 com 375 passarinhos podemos escrever esta informação como

$$\begin{pmatrix} 150 \\ 225 \\ 250 \\ 375 \end{pmatrix}.$$

Podemos também escrever este vetor indicando a porcentagem de pássaros em cada região, se não estivermos interessados na quantidade total de passarinhos por região. A partir dos números acima, temos que 15% dos passarinhos estão em R_1, 22,5% dos passarinhos estão em R_2, 25% dos passarinhos estão em R_3 e 37,5% dos passarinhos estão em R_4. Podemos escrever esta informação como

$$\begin{pmatrix} 15\% \\ 22,5\% \\ 25\% \\ 37,5\% \end{pmatrix} \mathit{ou} \begin{pmatrix} 0,15 \\ 0,225 \\ 0,25 \\ 0,375 \end{pmatrix}.$$

Se para cada dia associarmos um vetor como este último, descrevendo a porcentagem de passarinhos de cada região, podemos observar que todos estes vetores possuem quatro coordenadas não negativas cuja soma resulta em 1.

Definição 4.0.1. *Dizemos que $v \in \mathbb{R}^n$ é um vetor de probabilidade se suas coordenadas $v_1, ..., v_n$ satisfazem:*

a) $v_i \geq 0, i \in \{1, ..., n\}$;

b) $v_1 + v_2 + ... + v_n = 1$.

4.1 Matrizes estocásticas de transição

Suponhamos agora que para este modelo, independente da informação conhecida a mais de um dia e da quantidade de dias já observados, as porcentagens p_{ij}'s de passarinhos que migram para a região R_i saindo da região R_j , ao compararmos dias consecutivos, seja estabelecida em $T = (p_{ij})_{4\times4}$, onde

$$T = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} 0,2 & 0,1 & 0,1 & 0,15 \\ 0,3 & 0,3 & 0,3 & 0,25 \\ 0,2 & 0,4 & 0,5 & 0,2 \\ 0,3 & 0,2 & 0,1 & 0,4 \end{pmatrix}.$$

Assim, observando por exemplo a terceira coluna desta matriz, temos como previsão:

- 10% dos passarinhos em R_3 na fotografia atual estarão em R_1 na próxima fotografia.
- 30% dos passarinhos em R_3 na fotografia atual estarão em R_2 na próxima fotografia.
- 50% dos passarinhos em R_3 na fotografia atual estarão em R_3 na próxima fotografia.
- 10% dos passarinhos em R_3 na fotografia atual estarão em R_4 na próxima fotografia.

As entradas da matriz dada acima são não negativas e a soma dos elementos de cada coluna resulta em 1. Assim, as colunas da matriz são vetores de probabilidade.

Definição 4.1.1. *Dizemos que uma matriz quadrada $T_{n\times n}$ com entradas $\{p_{ij}\}$, $1 \leq i, j \leq n$ é (coluna) estocástica se satisfaz:*

a) $p_{ij} \geq 0$, para quaisquer $i, j \in \{1, ..., n\}$;

b) $p_{1j} + ... + p_{nj} = 1$ para qualquer $j \in \{1, ..., n\}$.

Denotemos por X_n a variável aleatória que assume valores em $\{1, 2, 3, 4\}$, que indica a região onde se localiza um determinado passarinho no n-ésimo dia. Ou seja, o processo X_n indica a trajetória realizada por um determinado passarinho (escolhido pelo pesquisador) ao longo do tempo n. As probabilidades de deslocamento entre as regiões podem ser descritas como uma probabilidade condicional:

$$P(X_{n+1} = i | X_n = j) = p_{ij}.$$

Como já vimos anteriormente, denota-se por $T^{(n)} = (p_{ij}^{(n)})$ a matriz de probabilidades de sair do estado j e chegar em i em n etapas. Para $n = 0$ e $n = 1$,

$$p_{ij}^{(1)} = p_{ij} \quad e \quad p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

Isto é, $T^{(1)} = T$ e $T^{(0)} = I$, onde I é a matriz identidade. Além disso, vimos que $T^{(n)} = T^n$.

Obtemos uma matriz das probabilidades de transição T a partir da tabela de probabilidades, onde o elemento na i-ésima linha e j-ésima coluna indica a probabilidade de transição do j-ésimo estado para o i-ésimo estado. Notemos que, desta maneira, a soma dos elementos de cada coluna será sempre igual a 1, pois essa soma representa a probabilidade do espaço amostral, em que cada entrada é a probabilidade de um evento disjunto deste espaço amostral. Podemos obter as probabilidades de transição p_{ij} em cada passo de forma matricial, representada da seguinte forma para uma matriz 2×2:

$$T = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}.$$

Neste caso, por exemplo, a probabilidade de sair do estado 1 e continuar no estado 1 após um passo é:

$$p_{11}^{(1)} = P(X_1 = 1|X_0 = 1) = p_{11}.$$

Da mesma forma, a probabilidade de sair do estado 1 e chegar no estado 2 após um passo é:

$$p_{21}^{(1)} = P(X_1 = 2|X_0 = 1) = p_{21}.$$

E assim sucessivamente.

Notemos então que a matriz $T = (p_{ij})_{4\times4}$ do exemplo dos passarinhos é uma matriz de transição.

Para calcular as probabilidades de transição em mais de um passo, basta elevarmos a matriz ao passo desejado. Por exemplo, para sabermos a possibilidade de sair do estado 1 e continuar no estado 1 em 2 passos fazemos:

$$T^2 = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \cdot \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} p_{11}p_{11} + p_{12}p_{21} & p_{11}p_{12} + p_{11}p_{22} \\ p_{21}p_{11} + p_{22}p_{21} & p_{21}p_{12} + p_{22}p_{22} \end{bmatrix},$$

onde o elemento a_{11} de T^2 nos dará o resultado desejado.

Pode-se observar esse resultado de outra forma, conhecida como árvore das probabilidades.

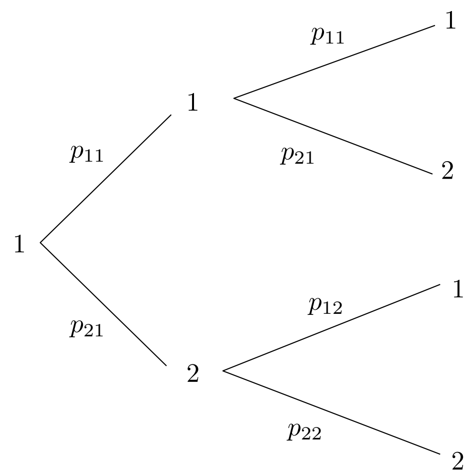

Figura 4.1: Árvore das Probabilidades

Para encontrarmos a probabilidade de sair do estado 1 e continuar no estado 1 em 2 passos através da árvore, basta multiplicar os "ramos da árvore"para obter:

$$p_{11}p_{11} + p_{21}p_{12},$$

onde este resultado nada mais é que $p_{11}^{(2)}$, como esperado.

Observe que em ambos os métodos, tanto o método da árvore quanto o da multiplicação de matrizes, obtemos o mesmo resultado. O mesmo vale para $p_{12}^{(2)}$, $p_{21}^{(2)}$ e $p_{22}^{(2)}$.

Podemos concluir então que o enésimo-passo da matriz de transição T é apenas a enésima potência de T.

Tendo em conta a distribuição da probabilidade associada com X_n, veremos mais adiante que a distribuição da probabilidade associada com X_{n+1} pode ser obtida multiplicando a matriz de transição T por v^n.

Exemplo 11. *Considerando o vetor de probabilidade encontrado no exemplo dos passarinhos*

$$\begin{pmatrix} 0,15 \\ 0,225 \\ 0,25 \\ 0,375 \end{pmatrix}$$

para o dia atual, qual previsão podemos fazer para o vetor associado ao dia seguinte?

Denotemos por v^0 o vetor que indica a distribuição dos passarinhos no tempo 0 (dia atual) e buscamos uma previsão de v^1, que indica a distribuição dos passarinhos no tempo 1 (dia seguinte). Qual a probabilidade de um dado passarinho estar em R_1 no tempo 1?

Pelas informações acima, temos que

$$v^0 = \begin{pmatrix} 0,15 \\ 0,225 \\ 0,25 \\ 0,375 \end{pmatrix}.$$

Assim, observando os "ramos da árvore"da Figura 4.2 *temos que*

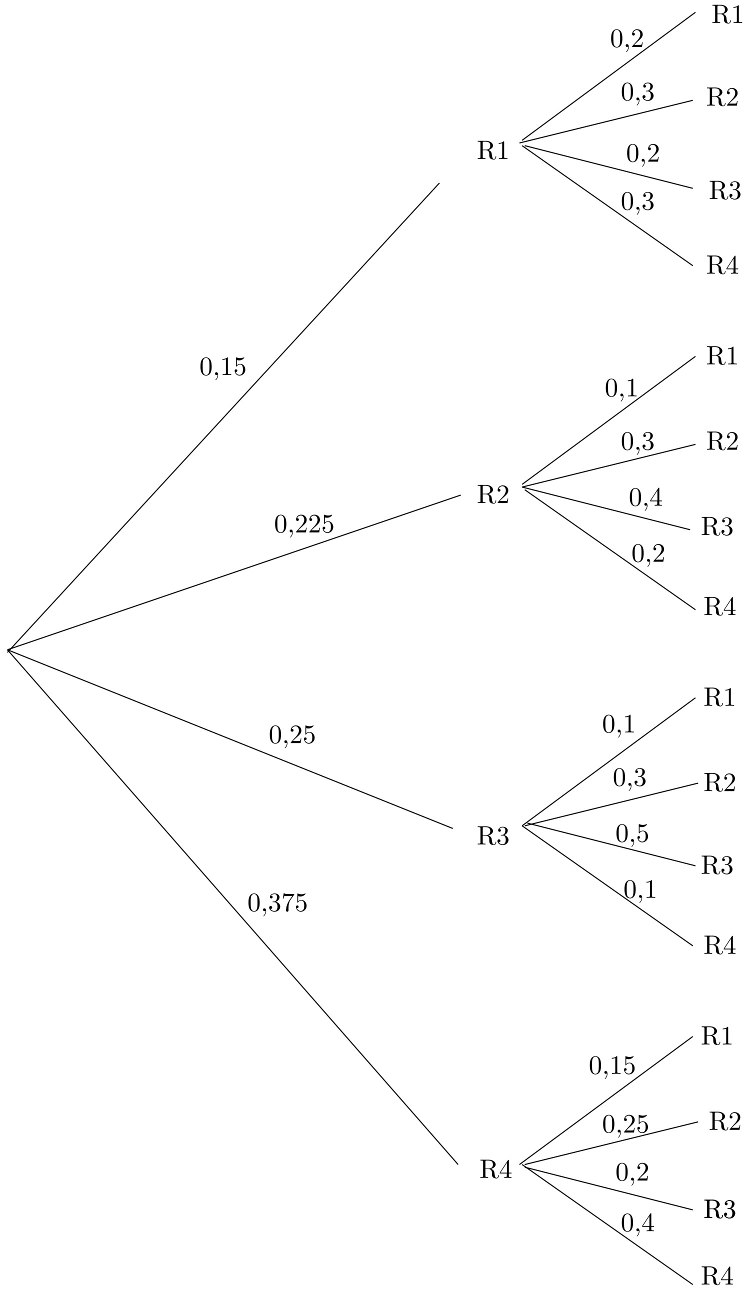

Figura 4.2: Árvore de probabilidade.

$$\begin{aligned} P(X_1 = 1) = {} & P(X_1 = 1|X_0 = 1)P(X_0 = 1) + P(X_1 = 1|X_0 = 2)P(X_0 = 2) + \\ & P(X_1 = 1|X_0 = 3)P(X_0 = 3) + P(X_1 = 1|X_0 = 4)P(X_0 = 4) = \\ & = (0,2)(0,15) + (0,1)(0,225) + (0,1)(0,25) + (0,15)(0,375) = \\ & = 0,03 + 0,0225 + 0,025 + 0,05625 = \\ & = 0,13375. \end{aligned}$$

Analogamente temos,

$$P(X_1 = 2) = (0,3)(0,15) + (0,3)(0,225) + (0,3)(0,25) + (0,25)(0,375) =$$
$$= 0,045 + 0,0675 + 0,075 + 0,09375 = 0,28125.$$
$$P(X_1 = 3) = (0,2)(0,15) + (0,4)(0,225) + (0,5)(0,25) + (0,2)(0,375) =$$
$$= 0,03 + 0,09 + 0,125 + 0,075 = 0,32.$$
$$P(X_1 = 4) = (0,3)(0,15) + (0,2)(0,225) + (0,1)(0,25) + (0,4)(0,375) =$$
$$= 0,045 + 0,045 + 0,025 + 0,15 = 0,265.$$

Assim, nossa previsão v^1 *é dada por*

$$v^1 = \begin{pmatrix} 0,13375 \\ 0,28125 \\ 0,32 \\ 0,265 \end{pmatrix}.$$

Note que esses cálculos podem ser obtidos de maneira mais "fácil":

$$v^1 = \begin{pmatrix} 0,13375 \\ 0,28125 \\ 0,32 \\ 0,265 \end{pmatrix} = \begin{pmatrix} 0,2 & 0,1 & 0,1 & 0,15 \\ 0,3 & 0,3 & 0,3 & 0,25 \\ 0,2 & 0,4 & 0,5 & 0,2 \\ 0,3 & 0,2 & 0,1 & 0,4 \end{pmatrix} . \begin{pmatrix} 0,15 \\ 0,225 \\ 0,25 \\ 0,375 \end{pmatrix} = T.v^0$$

Para a previsão de v^1 *podemos, como visto acima, calcular o produto da matriz coluna estocástica T, pelo vetor de probabilidade* v^0*. Note que neste exemplo o produto de uma matriz coluna estocástica por um vetor de probabilidade resultou em um novo vetor de probabilidade.*

Proposição 4.1.2. *Se* $T_{n\times n}$ *é uma matriz estocástica e* $v \in \mathbb{R}$ *é um vetor de probabilidade, então* Tv *é um vetor de probabilidade.*

Demonstração. Seja $w = Tv$. Temos que $T = (p_{i,j})_{1\leq i,j\leq n}$, $v = (v_i)_{1\leq i\leq n}$ e $w = (w_i)_{1\leq i\leq n}$. Assim,

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = Tv = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} p_{11}v_1 + p_{12}v_2 + \cdots + p_{1n}v_n \\ p_{21}v_1 + p_{22}v_2 + \cdots + p_{2n}v_n \\ \vdots \\ p_{n1}v_1 + p_{n2}v_2 + \cdots + p_{nn}v_n \end{pmatrix}$$

Assim,

$$w_i = \sum_{k=1}^{n} p_{ik}v_k, \quad i = 1, 2, ..., n.$$

Como T é coluna estocástica, segue que:

$$\sum_{i=1}^{n} p_{ik} = 1, \quad k = 1, 2, ..., n \quad e \quad p_{ik} \geq 0, \quad \text{para todo } i, k = 1, 2, ..., n.$$

Como v é um vetor de probabilidade segue que:

$$\sum_{k=1}^{n} v_k = 1 \quad e \quad v_k \geq 0, \quad \forall k = 1, 2, ..., n.$$

Logo,

$$w_i = \sum_{k=1}^{n} p_{ik}v_k \geq 0, \quad \forall i = 1, 2, ..., n \quad e$$

$$w_1 + w_2 + ... + w_n = \sum_{i=1}^{n} w_i = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} p_{ik}v_k\right) = \sum_{k=1}^{n}\sum_{i=1}^{n} p_{ik}v_k = \sum_{k=1}^{n} v_k\left(\sum_{i=1}^{n} p_{ik}\right) = \sum_{k=1}^{n} v_k = 1.$$

Portanto, $w = Tv$ é um vetor de probabilidade. □

Proposição 4.1.3. *Se $A_{n\times n}$ e $B_{n\times n}$ são matrizes estocásticas, então AB é uma matriz estocástica.*

Demonstração. Basta mostrar que se $A = (a_{ij})_{r\times r}$ e $B = (b_{ij})_{r\times r}$ são matrizes estocásticas, então a matriz AB também é estocástica.

Cada entrada ij da matriz AB é dada por

$$(AB)_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}.$$

E somando os elementos da *j-ésima* coluna, temos que:

$$\begin{aligned}
\sum_{i=1}^{r}(AB)_{ij} &= \sum_{i=1}^{r}\Big(\sum_{k=1}^{r} a_{ik}b_{kj}\Big)\\
&= \sum_{k=1}^{r}\underbrace{\Big(\sum_{i=1}^{r} a_{ik}\Big)}_{1}b_{kj}\\
&= \sum_{k=1}^{r} b_{kj}\\
&= 1.
\end{aligned}$$

□

A fim de mostrar que $v^n = T^n v^0$, iniciamos com o seguinte exemplo.

Exemplo 12. *Retomando o exemplo dos passarinhos, seja v^n o vetor da previsão no tempo n. Qual será o vetor v^{n+1} previsto no tempo n + 1?*

Temos

$$v^n = \begin{pmatrix} v_1^n \\ v_2^n \\ v_3^n \\ v_4^n \end{pmatrix}.$$

Seja $P(X_{n+1} = i|X_n = j) = p_{ij}$, onde

$$T = (p_{ij}) = \begin{pmatrix} 0,2 & 0,1 & 0,1 & 0,15 \\ 0,3 & 0,3 & 0,3 & 0,25 \\ 0,2 & 0,4 & 0,5 & 0,2 \\ 0,3 & 0,2 & 0,1 & 0,4 \end{pmatrix}.$$

Então,

$$v^{n+1} = \begin{pmatrix} P(X_{n+1} = 1) \\ P(X_{n+1} = 2) \\ P(X_{n+1} = 3) \\ P(X_{n+1} = 4) \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{4} P(X_{n+1} = 1|X_n = j)P(X_n = j) \\ \sum_{j=1}^{4} P(X_{n+1} = 2|X_n = j)P(X_n = j) \\ \sum_{j=1}^{4} P(X_{n+1} = 3|X_n = j)P(X_n = j) \\ \sum_{j=1}^{4} P(X_{n+1} = 4|X_n = j)P(X_n = j) \end{pmatrix}$$

$$= \begin{pmatrix} 0,2v_1^n + 0,1v_2^n + 0,1v_3^n + 0,15v_4^n \\ 0,3v_1^n + 0,3v_2^n + 0,3v_3^n + 0,25v_4^n \\ 0,2v_1^n + 0,4v_2^n + 0,5v_3^n + 0,2v_4^n \\ 0,3v_1^n + 0,2v_2^n + 0,1v_3^n + 0,4v_4^n \end{pmatrix}$$

$$= \begin{pmatrix} 0,2 & 0,1 & 0,1 & 0,15 \\ 0,3 & 0,3 & 0,3 & 0,25 \\ 0,2 & 0,4 & 0,5 & 0,2 \\ 0,3 & 0,2 & 0,1 & 0,4 \end{pmatrix} \begin{pmatrix} v_1^n \\ v_2^n \\ v_3^n \\ v_4^n \end{pmatrix} = Tv^n.$$

Por outro lado, como $v^1 = Tv^0$, segue que

$$v^2 = Tv^1 = T(Tv^0) = T^2v^0,$$

e, portanto,

$$v^3 = Tv^2 = T(T^2v^0) = T^3v^0,$$

e, assim por diante. Temos então que $v^n = T^n v^0$.

Com base neste raciocínio, conhecida uma matriz estocástica (transição) $T_{n\times n}$ e um vetor de probabilidade $v^0 \in \mathbb{R}^n$, podemos tentar estimar o comportamento do vetor v^n "a longo prazo". Sendo assim, queremos verificar se as coordenadas de v^n se aproximam das coordenadas de algum vetor w, quando $n \to \infty$.

Daremos agora um exemplo, utilizando o fato que $v^{n+1} = Tv^n$, para estudar uma convergência dos vetores de probabilidade conforme o valor de n (tempo) aumenta.

Exemplo 13. *Em uma determinada região, verifica-se que se chover em um dia, a probabilidade de que chova no dia seguinte é $\frac{1}{2}$, e por consequência a probabilidade de não chover (seca) também é $\frac{1}{2}$. Se, no entanto, tivermos seca em um dia, temos que a probabilidade de chuva para o próximo dia será de $\frac{1}{4}$ e de seca, $\frac{3}{4}$. Suponhamos que estas probabilidades não mudarão ao decorrer do tempo e que no dia atual choveu.*

Assim, temos dois possíveis estados a cada dia: Chuva (C) e Seca (S).

O vetor de probabilidade inicial v^0 é dado por:

$$v^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

A partir das informações fornecidas pelo problema, cujas possibilidades são chuva e seca, compomos uma tabela de probabilidades de transição:

	Chuva	*Seca*
Chuva	$\frac{1}{2}$	$\frac{1}{4}$
Seca	$\frac{1}{2}$	$\frac{3}{4}$

Utilizando os dados acima, temos a seguinte matriz de transição:

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}.$$

Qual será a probabilidade de chuva e seca após um longo período?

Vamos verificar, por exemplo, a probabilidade após um período de $n = 5$ dias.

Temos:

$$v^1 = Tv^0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0,5 \\ 0,5 \end{pmatrix};$$

$$v^2 = Tv^1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{8} \\ \frac{5}{8} \end{pmatrix} = \begin{pmatrix} 0,375000000000000 \\ 0,625000000000000 \end{pmatrix};$$

$$v^3 = Tv^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{8} \\ \frac{5}{8} \end{pmatrix} = \begin{pmatrix} \frac{11}{32} \\ \frac{21}{32} \end{pmatrix} = \begin{pmatrix} 0,343750000000000 \\ 0,656250000000000 \end{pmatrix};$$

$$v^4 = Tv^3 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} \frac{11}{32} \\ \frac{21}{32} \end{pmatrix} = \begin{pmatrix} \frac{43}{128} \\ \frac{85}{128} \end{pmatrix} = \begin{pmatrix} 0,335937500000000 \\ 0,664062500000000 \end{pmatrix};$$

$$v^5 = Tv^4 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix} \cdot \begin{pmatrix} \frac{43}{128} \\ \frac{85}{128} \end{pmatrix} = \begin{pmatrix} \frac{171}{512} \\ \frac{341}{512} \end{pmatrix} = \begin{pmatrix} 0,333984375000000 \\ 0,666015625000000 \end{pmatrix}.$$

A partir dos cálculos acima, podemos perceber que, após um período de 5 dias, a probabilidade de chuva será, aproximadamente, $0,334 = 33,4\%$ *e de fazer seca será* $0,666 = 66,6\%$.

Além disso, podemos usar o fato de que $v^n = T^n v^0$ *e o software "calculadora de matrizes", disponível em [7], no qual é uma ferramenta para cálculos envolvendo matrizes, para calcular as distribuições* v^n *para valores de* n *cada vez maiores.*

Considerando uma aproximação com 15 casas decimais, temos,

Para $n = 10$*:*

$$v^{10} = T^{10}v^0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}^{10} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,333333969116211 \\ 0,666666030883789 \end{pmatrix}.$$

Para $n = 50$*:*

$$v^{50} = T^{50}v^0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}^{50} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,333333333333333 \\ 0,666666666666667 \end{pmatrix}.$$

Para $n = 100$*:*

$$v^{100} = T^{100}v^0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}^{100} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,333333333333333 \\ 0,666666666666667 \end{pmatrix}.$$

Para $n = 200$*:*

$$v^{200} = T^{200}v^0 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{3}{4} \end{pmatrix}^{200} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,333333333333333 \\ 0,666666666666667 \end{pmatrix}.$$

Podemos notar que parece haver uma convergência nas probabilidades conforme o valor de n *aumenta ou seja,*

$$v^n \longrightarrow \begin{pmatrix} 0,333333333333333 \\ 0,666666666666667 \end{pmatrix}, \quad \textit{quando} \quad n \longrightarrow \infty.$$

Veremos mais adiante que, a longo prazo, o comportamento do clima desta dada região poderá ser previsto com base na aproximação dos elementos das matrizes $T^n(n = 1, 2, ...)$ a uma matriz fixa M, quando os valores de n aumentam $(n \longrightarrow \infty)$.

Se T^n se aproxima de uma matriz M, então haverá previsão concisa a longo prazo, não sofrendo mudanças significativas a cada passo do processo. Portanto, faz-se necessário identificar sob quais condições uma matriz das probabilidades de transição se aproximará de uma determinada matriz fixa M.

4.1.1 Convergência de matrizes de transição diagonalizáveis

Como vimos anteriormente, estamos interessados em estudar sob quais condições uma matriz estocástica se aproximará de uma determinada matriz M. Mostraremos nesta seção um caso especial de convergência de matrizes estocásticas diagonalizáveis.

Seja $T = [f]_{can}^{can}$, onde *can* representa a base canônica de $\mathbb{R}^k$, uma matriz de transição. Da álgebra linear, sabemos que podemos associar a matriz T a uma transformação linear $f : \mathbb{R}^k \longrightarrow \mathbb{R}^k$. Se a matriz T é diagonalizável, significa que existe uma base $\beta = \{v_1, v_2, ..., v_k\}$ de autovetores para $\mathbb{R}^k$, tais que:

$$[f]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix},$$

onde $\lambda_1, \lambda_2, ..., \lambda_k$ são os autovalores. Lembrando que, um vetor não nulo $v_i \in \mathbb{R}^k$ é um autovetor de $\mathbb{R}^k$ se existe um número real λ_i, tal que $f(v_i) = \lambda_i v_i$. O escalar λ_i é denominado um autovalor de f associado a $\mathbb{R}^k$. Para maiores detalhes sobre autovalores e autovetores ver, por exemplo, [8].

É conhecido da álgebra linear que existe uma matriz de mudança de base $[I]_{can}^\beta$, tal que

$$[f]_{can}^{can} = [I]_{can}^\beta \cdot [f]_\beta^\beta \cdot [I]_\beta^{can}, \quad onde \quad [I]_\beta^{can} = ([I]_{can}^\beta)^{-1}.$$

Note que a multiplicação das matrizes acima pode ser interpretada como uma composição de transformações lineares: $I \circ f \circ I : \mathbb{R}^k(can) \longrightarrow \mathbb{R}^k(can)$.

Considerando $\{e_1, e_2, ..., e_k\}$, a base canônica para R^k, temos

$$I(v_1) = v_1 = v_{11}e_1 + v_{21}e_2 + ... + v_{k1}e_k$$

$$I(v_2) = v_2 = v_{12}e_1 + v_{22}e_2 + ... + v_{k2}e_k$$

$$\vdots$$

$$I(v_k) = v_k = v_{1k}e_1 + v_{2k}e_2 + ... + v_{kk}e_k$$

ou seja,

$$[I]^{\beta}_{can} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1k} \\ v_{21} & v_{22} & \cdots & v_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ v_{k1} & v_{k2} & \cdots & v_{kk} \end{bmatrix},$$

onde em cada coluna de $[I]^{\beta}_{can}$ temos as coordenadas do vetores v_i, para $i = 1, 2, ..., k$. Note que $B = [I]^{\beta}_{can}$ é uma matriz formada a partir dos autovetores de $[f]^{can}_{can}$ e $[f]^{\beta}_{\beta}$ é a matriz diagonal dos autovalores.

Além disso,

$$([f]^{can}_{can})^2 = ([I]^{\beta}_{can} \cdot [f]^{\beta}_{\beta} \cdot \underbrace{[I]^{can}_{\beta}) \cdot ([I]^{\beta}_{can}}_{[I \circ I]^{\beta}_{\beta} = [I]^{\beta}_{\beta}} \cdot [f]^{\beta}_{\beta} \cdot [I]^{can}_{\beta}) =$$

$$= [I]^{\beta}_{can} \cdot [f]^{\beta}_{\beta} \cdot [I]^{\beta}_{\beta} \cdot [f]^{\beta}_{\beta} \cdot [I]^{can}_{\beta} = [I]^{\beta}_{can} \cdot ([f]^{\beta}_{\beta})^2 \cdot [I]^{can}_{\beta}.$$

A última igualdade se deve ao fato de que :

$$I(v_1) = 1v_1 + 0v_2 + \cdots + 0v_k$$

$$I(v_2) = 0v_1 + 1v_2 + \cdots + 0v_k$$

$$\vdots$$

$$I(v_1) = 0v_1 + 0v_2 + \cdots + 1v_k$$

Assim,

$$[I]^{\beta}_{\beta} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} (matriz\ identidade).$$

Seguindo este raciocínio sucessivamente para n, temos

$$([f]^{can}_{can})^n = [I]^{\beta}_{can} \cdot ([f]^{\beta}_{\beta})^n \cdot [I]^{can}_{\beta} \ , \ onde\ [I]^{can}_{\beta} = ([I]^{\beta}_{can})^{-1}$$

e dependendo dos autovalores de $[f]^{can}_{can}$, $([f]^{\beta}_{\beta})^n$ pode possuir um limite finito ou não.

Exemplo 14. *Considere que a nossa matriz de transição seja a matriz encontrada no exemplo do interruptor visto no Capítulo 3:*

$$T = [f]_{can}^{can} = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{bmatrix}.$$

Esta matriz está associada à transformação linear $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *dada por*

$$f(x,y) = \Big(\frac{5}{6}x + \frac{1}{6}y, \frac{1}{6}x + \frac{5}{6}y\Big).$$

Inicialmente, vamos encontrar os autovalores e autovetores de f*. Para o cálculo dos autovalores resolvemos a equação característica* $det(T - \lambda I) = 0$.

$$det(T - \lambda I) = \begin{vmatrix} \frac{5}{6} - \lambda & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} - \lambda \end{vmatrix} = 0.$$

Temos que,

$$\begin{aligned} \Big(\frac{5}{6} - \lambda\Big)^2 - \frac{1}{36} &= 0 \\ \Big(\frac{5}{6} - \lambda\Big)^2 &= \frac{1}{36} \\ \frac{5}{6} - \lambda &= \pm\frac{1}{6} \\ -\lambda &= \pm\frac{1}{6} - \frac{5}{6} \\ -\lambda = \frac{-2}{3} \ &ou \ -\lambda = -1. \end{aligned}$$

Logo, os autovalores são $\lambda_1 = 1$ *e* $\lambda_2 = \frac{2}{3}$.

Para cada autovalor λ *encontrado, resolvemos o sistema linear* $(T - \lambda I)v = 0$ *para encontrar os autovetores* v_1 *e* v_2.

Para $\lambda_1 = 1$ *e* $v_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, *temos*

$$(T - \lambda_1 I)v_1 = \begin{bmatrix} \frac{5}{6} - 1 & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} - 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 0.$$

Obtemos então o seguinte sistema linear

$$\begin{cases} \frac{-1}{6}x_1 & + & \frac{1}{6}y_1 & = & 0 \\ \frac{1}{6}x_1 & + & \frac{-1}{6}y_1 & = & 0 \end{cases}$$

Assim, $\frac{1}{6}x_1 = \frac{1}{6}y_1 \Rightarrow x_1 = y_1$. *Então,* $v_1 = (y_1, y_1)$ *e um de seus representantes é o vetor* $v_1 = (1,1)$.

Para $\lambda_2 = \frac{2}{3}$ *e* $v_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, *temos*

$$(T - \lambda_2 I)v_2 = \begin{bmatrix} \frac{5}{6} - \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} - \frac{2}{3} \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = 0.$$

Obtemos então o seguinte sistema linear

$$\begin{cases} \frac{1}{6}x_2 & + & \frac{1}{6}y_2 & = & 0 \\ \frac{1}{6}x_2 & + & \frac{1}{6}y_2 & = & 0 \end{cases}$$

Assim, $\frac{1}{6}x_2 = \frac{-1}{6}y_2 \Rightarrow x_2 = -y_2$. *Então,* $v_2 = (-y_2, y_2)$ *e um de seus representantes é o vetor* $v_2 = (-1,1)$.

Logo, $\beta = \{(1,1), (-1,1)\}$ *é uma base de autovetores para* $\mathbb{R}^2$. *Portanto,* f *é diagonalizável. Desse modo, temos que*

$$[f]_\beta^\beta = \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{3} \end{bmatrix} \text{ e } [I]_{can}^\beta = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Queremos encontrar agora a matriz $[I]_\beta^{can}$. *Temos*

$$I(1,0) = (1,0) = a(1,1) + b(-1,1).$$

Obtemos o sistema linear

$$\begin{cases} a & - & b & = & 1 \\ a & + & b & = & 0 \end{cases}$$

Assim, $a = \frac{1}{2}$ *e* $b = \frac{-1}{2}$.

$$I(0,1) = (0,1) = c(1,1) + d(-1,1).$$

Obtemos o sistema linear

$$\begin{cases} c & - & d & = & 0 \\ c & + & d & = & 1 \end{cases}$$

Assim, $c = \frac{1}{2}$ *e* $d = \frac{1}{2}$.

Logo,

$$[I]_{\beta}^{can} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}.$$

Uma outra maneira de encontrar $[I]_{\beta}^{can}$ *é através do cálculo da inversa de* $[I]_{can}^{\beta}$.

Assim, se $v^0 = (v_1^0, v_2^0)$ *é a distribuição inicial, então*

$$v^n = T^n v^0 = ([f]_{can}^{can})^n v^0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{3} \end{bmatrix}^n \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} v^0.$$

Como nosso objetivo é estudar o que irá acontecer em um experimento a longo prazo, vamos discutir a existência do limite $\lim\limits_{n\to\infty}([f]_{can}^{can})^n$.

Note que

$$\lim_{n\to\infty}([f]_{can}^{can})^n = \lim_{n\to\infty}\{[I]_{can}^{\beta} \cdot ([f]_{\beta}^{\beta})^n \cdot [I]_{\beta}^{can}\} = [I]_{can}^{\beta} \cdot \{\lim_{n\to\infty}([f]_{\beta}^{\beta})^n\} \cdot [I]_{\beta}^{can} =$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \lim_{n\to\infty} \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{3} \end{bmatrix}^n \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \lim_{n\to\infty} \begin{bmatrix} 1^n & 0 \\ 0 & (\frac{2}{3})^n \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Desse modo, podemos concluir que:

$$\begin{aligned}
\lim_{n\to\infty} v^n = \lim_{n\to\infty} T^n v^0 &= \Big(\lim_{n\to\infty}([f]_{can}^{can})^n\Big)v^0 \\
&= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} v_1^0 \\ v_2^0 \end{bmatrix} \\
&= \Big(\frac{1}{2}v_1^0 + \frac{1}{2}v_2^0, \frac{1}{2}v_1^0 + \frac{1}{2}v_2^0\Big) \\
&= \Big(\frac{1}{2}(v_1^0 + v_2^0), \frac{1}{2}(v_1^0 + v_2^0)\Big) \\
&= \Big(\frac{1}{2}, \frac{1}{2}\Big).
\end{aligned}$$

A última igualdade se deve ao fato de $v_1^0 + v_2^0 = 1$.

Ou seja, a medida que n *aumenta, a distribuição* v^n *se aproxima de* $(\frac{1}{2}, \frac{1}{2})$*, independente da distribuição inicial* v^0*. Logo, a probabilidade do interruptor estar ligado ou desligado, a longo prazo, será a mesma:* $\frac{1}{2}$*, independente do estado anterior.*

4.1.2 Convergência de matrizes de transição regulares

Definição 4.1.4. *Uma matriz das probabilidades de transição é regular se alguma de suas potências tem todos os elementos não nulos.*

Note, por exemplo, que a matriz das probabilidades de transição dada por

$$T = \begin{pmatrix} 0 & 0,4 \\ 1 & 0,6 \end{pmatrix},$$

é regular, visto que

$$T^2 = \begin{pmatrix} 0,4 & 0,24 \\ 0,6 & 0,76 \end{pmatrix}.$$

Vale ressaltar que nem toda matriz de transição é regular. As matrizes diagonais são exemplos imediatos. A matriz de transição:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

significa que há probabilidade nula de mudança de estado ou seja, a probabilidade de transição do j-ésimo estado para o i-ésimo estado é igual a 1 (100%), se $i = j$ e é igual a 0 (0%), se $i \neq j$. Neste caso, dizemos que a matriz das probabilidades de transição é absorvente, ou seja, se em determinado passo, caindo em um estado, seja "impossível"sair dele.

Agora iremos apresentar o Teorema de Perron-Frobenius, que nos fornece uma ferramenta para o cálculo de previsões a longo prazo, para os casos onde a matriz de transição é regular.

Teorema 4.1.5. *(Perron - Frobenius) Seja T uma matriz estocástica com a seguinte propriedade: existe um inteiro $m \geq 1$, tal que $(T^m)_{i,j} > 0$, para todo i, j, ou seja, T é regular. Então:*

1. *T possui o autovalor $\lambda = 1$ e qualquer outro autovalor λ' de T satisfaz $|\lambda'| < \lambda$. Além disso, λ é não-degenerado: existe um único autovetor w associado, $Tw = w$, que pode ser escolhido, tal que $\displaystyle\sum_i w_i = 1$, com $w_i \geq 0$, para todo i.*
2. *Para qualquer distribuição inicial v, $T^n v \longrightarrow w$.*

Exemplo 15. *Podemos resolver o Exemplo 14 através do Teorema de Perron - Frobenius. Como a matriz*

$$T = [f]_{can}^{can} = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \end{bmatrix}$$

é regular, pelo Teorema de Perron - Frobenius, $[f]_{can}^{can}$ possui $\lambda = 1$ como autovalor e existe um único autovetor $w = (w_1, w_2)$ associado, $Tw = w$, que pode ser escolhido tal que $w_1 + w_2 = 1$.

Como vimos acima, $v_1 = (y_1, y_1)$ é um autovetor para $\lambda = 1$. Assim, temos que $w = (w_1, w_1)$ é tal que $w_1 + w_1 = 1$. Logo, $w = (\frac{1}{2}, \frac{1}{2})$.

Podemos concluir então que

$$([f]_{can}^{can})^n v^0 \longrightarrow \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix},$$

quando $n \longrightarrow \infty$.

Reescreveremos o Teorema 4.1.5 (Perron - Frobenius), numa versão que não utiliza conhecimentos avançados de álgebra linear. A aplicação deste teorema, nesta versão, nos possibilitará encontrar tais previsões fazendo apenas o uso de conceitos estudados no

Ensino Médio: matrizes e resolução de sistemas lineares. As ferramentas que aparecem na demonstração deste resultado, que apresentaremos em seguida, não serão utilizados em nossos problemas de aplicações voltados para o Ensino Médio. Porém a prova é interessante para estudantes de álgebra linear.

Teorema 4.1.6. *Se T é uma matriz estocástica regular $r \times r$ então:*

i) *T^n se aproxima de uma matriz M, no sentido de que cada entrada da matriz T^n aproxima-se da entrada correspondente em M;*

ii) *Todas as colunas de M são iguais, sendo dadas por um vetor coluna*

$$w = \begin{bmatrix} w_1 \\ \vdots \\ w_r \end{bmatrix},$$

com $w_i > 0$, para $i = 1, \dots r$.

iii) *Para qualquer vetor de probabilidades inicial*

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_r \end{bmatrix},$$

o vetor de probabilidades $T^n v$ aproxima-se de w, quando $n \to \infty$;

iv) *O vetor w é o único que satisfaz $Tw = w$ e é chamado vetor estacionário.*

Demonstração. Faremos aqui a prova para matrizes 2×2. A mesma ideia pode ser aplicada para matrizes $r \times r$.

a. Vamos supor inicialmente que T é uma matriz com entradas todas não nulas e que $\varepsilon > 0$ seja uma entrada da matriz, cujo valor é menor ou igual que outras entradas. Assim, podemos supor que:

$$T = \begin{bmatrix} \alpha & \varepsilon \\ 1-\alpha & 1-\varepsilon \end{bmatrix},$$

com $\alpha > 0$.

Seja x um vetor linha tendo mínima componente m_0 e máxima componente M_0, ou seja, $m_0 = min(x)$ e $M_0 = max(x)$. E seja m_1 e M_1 a mínima e máxima componente da xT.

Supondo que $x = [m_0 \;\; M_0]$, temos que $xT = [m_0\alpha + M_0(1-\alpha) \quad m_0\varepsilon + M_0(1-\varepsilon)] = [M_0 - \alpha(M_0 - m_0) \quad M_0 - \varepsilon(M_0 - m_0)]$, e desta forma, temos $M_1 = M_0 - \varepsilon(M_0 - m_0)$. O mesmo ocorre se $x = [M_0 \;\; m_0]$. A expressão xT também pode ser escrita como:

$$xT = [m_0 + (1-\alpha)(M_0 - m_0) \quad m_0 + (1-\varepsilon)(M_0 - m_0)].$$

Como, por hipótese, $1 - \alpha \geq \varepsilon$, segue que $m_1 = m_0 + (1-\alpha)(M_0 - m_0) \geq m_0 + \varepsilon(M_0 - m_0)$ e, portanto,

$$-m_1 \leq -m_0 - \varepsilon(M_0 - m_0).$$

Desta forma, como $M_1 = M_0 - \varepsilon(M_0 - m_0)$, somando ambos os membros com a expressão anterior, temos

$$M_1 - m_1 \leq M_0 - m_0 - 2\varepsilon(M_0 - m_0)$$

$$M_1 - m_1 \leq (1 - 2\varepsilon)(M_0 - m_0).$$

Seja e_j o vetor linha com o número 1 na j-ésima entrada (no nosso caso $1 \leq j \leq 2$). E sejam M_n e m_n os valores máximo e mínimo das componentes do vetor e_jT^n, ou seja, $M_n = max(e_jT^n)$ e $m_n = min(e_jT^n)$, para n = 1, 2, 3, ... Note que, $M_1 \geq M_2 \geq M_3 \geq \ldots$ e $m_1 \leq m_2 \leq m_3 \leq \ldots$.

Vamos considerar aqui $T^0 = I$ e, portanto, $M_0 = max(e_jT^0) = max(e_j) = 1$ e $m_0 = min(e_jT^0) = min(e_j) = 0$. Como $M_1 = max(e_jT)$ e $m_1 = min(e_jT)$, pelo resultado obtido acima segue que:

$$M_1 - m_1 \leq (1 - 2\varepsilon)(M_0 - m_0) = (1 - 2\varepsilon)(1 - 0) = (1 - 2\varepsilon).$$

Analogamente, podemos escrever $M_2 = max(e_jT^2) = max((e_jT)T)$ e $m_2 = min(e_jT^2) = min((e_jT)T)$, e aplicando novamente o resultado acima, chegamos a

$$M_2 - m_2 \leq (1 - 2\varepsilon)(M_1 - m_1) \leq (1 - 2\varepsilon)(1 - 2\varepsilon),$$

ou seja,

$$M_2 - m_2 \leq (1 - 2\varepsilon)^2.$$

Para um n qualquer podemos escrever $M_n = max(e_jT^n) = max((e_jT^{n-1})T)$ e $m_n = min(e_jT^n) = min((e_jT^{n-1})T)$, aplicando o procedimento descrito anteriormente sucessivas vezes obtemos que

$$M_n - m_n \leq (1 - 2\varepsilon)^n,$$

para $n \geq 1$.

Assim, $M_n - m_n \to 0$, quando $n \to \infty$ e M_n e m_n se aproximam para um limite comum, digamos w_j. É claro que $m_n \leq w_j \leq M_n$. Em particular, como $0 < m_1$ e $M_1 < 1$, temos que $0 < w_j < 1$. Portanto, e_jT^n tende a um vetor em que a maior e a menor componente se aproximam, ou seja, um vetor onde todas as componentes tendem a w_j. Logo, a *j-ésima* linha de M é dada por um vetor de entradas w_j. E portanto, as colunas de M são iguais a um vetor

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Com $w_1 + w_2 = 1$ já que T^n é uma matriz estocástica pra cada n, e portanto, o mesmo deve valer para o limite M.

b. Vamos supor que T é regular e que não necessariamente todas as entradas de T sejam diferentes de zero. Seja N, tal que T^N é a matriz cujas entradas são não nulas. Seja ε' o menor valor das entradas de T^N. Aplicando o mesmo raciocínio do item **a.** para T^N, temos $M_{NK} = max(e_j(T^N)^K)$ e $m_{NK} = min(e_j(T^N)^K)$. Logo,

$$M_{NK} - m_{NK} \leq (1 - 2\varepsilon')^K.$$

Portanto, a sequência não-crescente $(M_n - m_n)$ tem uma subsequência $(M_{NK} - m_{NK})$ tendendo a zero. Logo, $(M_n - m_n)$ tende a zero e o resto da prova é análogo ao item **a.**.

Isto prova os itens **(i)** e **(ii)**.

Prova do item (iii): Temos que $T^n v$ se aproxima de Mv, quando $n \to \infty$. Além disso, uma vez que $v_1 + v_2 = 1$, segue que:

$$Mv = \begin{bmatrix} w_1 & w_1 \\ w_2 & w_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} w_1(v_1+v_2) \\ w_2(v_1+v_2) \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = w.$$

Portanto,

$$T^n v \to w.$$

Prova do item (iv): Temos que $T^n T \to MT$. Por outro lado, $T^{n+1} \to M$. Como $T^{n+1} = T^n T$, pela unicidade do limite, $MT = M$. Analogamente, $TM = M$. E assim temos:

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} w_1 & w_1 \\ w_2 & w_2 \end{bmatrix} = \begin{bmatrix} w_1 & w_1 \\ w_2 & w_2 \end{bmatrix}.$$

E desta equação matricial extraímos

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Ou seja,

$$Tw = w.$$

Vamos agora mostrar a unicidade de w. Suponha que $\tilde{w}$ seja outro vetor de probabilidade com $T\tilde{w} = \tilde{w}$. Logo, $T^n\tilde{w} = \tilde{w}$, para todo n.

E assim, $T^n\tilde{w} \to \tilde{w}$. Mas por **(iii)** sabemos que $T^n\tilde{w} \to w$. Logo, pela unicidade do limite, segue que $\tilde{w} = w$.

$\square$

A demonstração completa do Teorema 4.1.6 para T uma matriz $r \times r$ se encontra em [9].

Observação 4.1.7. *Com esse teorema percebe-se que a previsão a longo prazo não dependerá do vetor de probabilidade inicial.*

Observação 4.1.8. *Note que o Teorema acima, é de fato o próprio Teorema de Perron - Frobenius, visto que, se T é uma matriz estocástica regular, então $\lambda = 1$ é um autovalor de T e assim, w é tal que $Tw = 1 \cdot w$.*

Assim, para encontrarmos uma previsão w a longo prazo, basta conhecermos a matriz das probabilidades de transição T e resolvermos o sistema linear $Tw = w$, o que torna possível explorarmos este assunto com alunos do Ensino Médio, visto que o conteúdo de matrizes, probabilidade e sistemas lineares é trabalhado com os alunos nesta etapa.

Capítulo 5

Aplicação das cadeias de Markov no Ensino Médio

Para ilustrarmos como o Teorema 4.1.6 pode ser aplicado no contexto de cadeias de Markov no Ensino Médio, veremos a seguir exemplos, onde alguns foram retirados de [8] e [10].

Exemplo 16. *Duas substâncias distintas estão em contato e trocam íons de sódio entre si. Sabe-se (por dedução teórica, ou experimentação) que um íon de sódio do meio (1) tem probabilidade 0,7 de passar ao meio (2), enquanto que um íon de sódio que esteja no meio (2) tem probabilidade 0,1 de passar ao meio (1).*

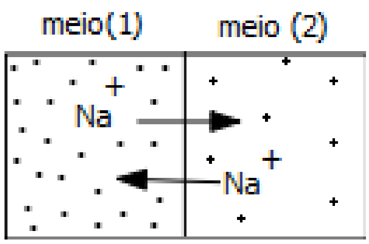

Figura 5.1: Trocas de íons de sódio.

Colocando-se dois moles de sódio no meio (1), quais serão as concentrações de sódio em cada um dos meios, após um longo período de tempo?

Os estados deste processo são: o íon está no meio (1) e o íon está no meio (2). Considerando que um íon de sódio do meio (1) tem probabilidade 0,7 de passar ao meio

(2), segue que ele tem probabilidade 0,3 de permanecer onde está e se um íon de sódio do meio (2) tem probabilidade 0,1 de passar ao meio (1), ele terá probabilidade 0,9 de permanecer onde está.

Podemos representar esta situação através do seguinte diagrama de transição:

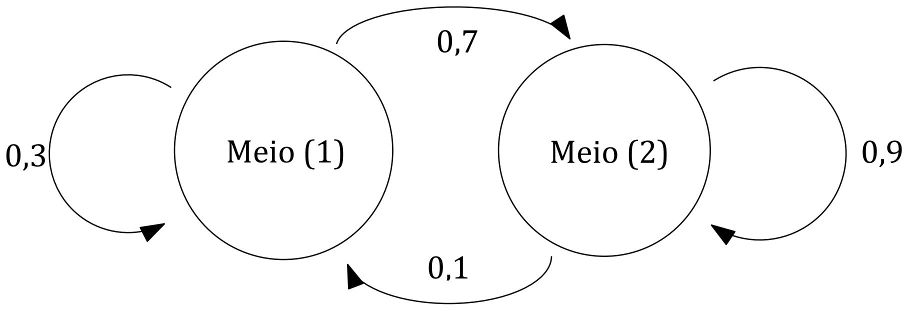

Figura 5.2: Probabilidades de transição dos íon de sódio.

A matriz de probabilidades de transição será dada por:

	meio (1)	*meio (2)*
meio (1)	$0,3$	$0,1$
meio (2)	$0,7$	$0,9$

$$\Rightarrow T = \begin{bmatrix} 0,3 & 0,1 \\ 0,7 & 0,9 \end{bmatrix}.$$

Sejam p_1 e p_2 as probabilidades de estar no meio (1) e (2), respectivamente. Como T é regular, a longo prazo as probabilidades não dependem das probabilidades iniciais e podemos aplicar o Teorema 4.1.6. Logo,

$$\begin{bmatrix} 0,3 & 0,1 \\ 0,7 & 0,9 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}.$$

O que nos leva a resolver o sistema linear:

$$\begin{cases} 0,3p_1 & + & 0,1p_2 & = & p_1 \\ 0,7p_1 & + & 0,9p_2 & = & p_2 \end{cases} \Rightarrow \begin{cases} -0,7p_1 & + & 0,1p_2 & = & 0 \\ 0,7p_1 & + & -0,1p_2 & = & 0 \end{cases}$$

Concluímos então que $0,7p_1 = 0,1p_2 \Rightarrow p_2 = 7p_1$. Como $p_1 + p_2 = 1$ segue que:

$$p_1 + 7p_1 = 1 \Rightarrow p_1 = \frac{1}{8} \;\; e \;\; p_2 = \frac{7}{8}.$$

Logo, as concentrações finais, depois de um longo período, em cada meio são: $\frac{1}{8} \cdot 2 = 0,25$ moles no meio (1) e $\frac{7}{8} \cdot 2 = 0,75$ moles no meio (2).

Exemplo 17. *Num polo industrial de uma cidade, os dados sobre a qualidade do ar são classificados como satisfatório (S) e insatisfatório (I). Assuma que, se num dia é registrado (S), a probabilidade de se ter (S) no dia seguinte é de* $\frac{2}{5}$ *e que, uma vez registrado (I), tem-se* $\frac{1}{5}$ *de probabilidade de ocorrer (S) no dia seguinte.*

a) Qual é a probabilidade do quarto dia ser satisfatório (S), se o primeiro dia foi insatisfatório (I)?

b) O que se pode dizer a longo prazo sobre a probabilidade de termos dias (S) ou (I)?

Os estados deste processo são: a qualidade do ar está satisfatória (S) e a qualidade do ar está insatisfatória (I). Considerando que se num dia é registrado (S), a probabilidade de se ter (S) no dia seguinte é de $\frac{2}{5} = 0,4$*, segue que ele tem probabilidade* $\frac{3}{5} = 0,6$ *de ser (I) no dia seguinte e se registrado (I), tem-se* $\frac{1}{5} = 0,2$ *de probabilidade de ocorrer (S) no dia seguinte, ele terá* $\frac{4}{5} = 0,8$ *de probabilidade de ocorrer (I) no dia seguinte.*

A matriz de probabilidades de transição será dada por:

	S	I
S	$0,4$	$0,2$
I	$0,6$	$0,8$

$$\Rightarrow T = \begin{bmatrix} 0,4 & 0,2 \\ 0,6 & 0,8 \end{bmatrix}.$$

Veja que T representa a matriz de probabilidades para o segundo dia. Podemos representar esta situação através do seguinte diagrama de transição:

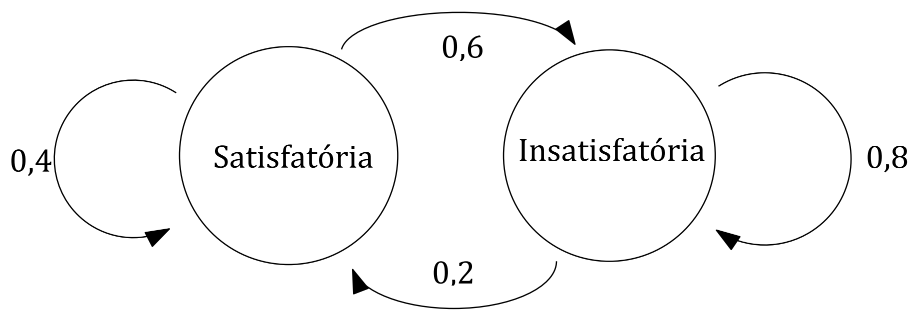

Figura 5.3: Probabilidades de transição da qualidade do ar.

Para o item a), a probabilidade de ocorrer (S) no quarto dia tendo ocorrido (I) no primeiro dia pode ser obtida calculando o elemento a_{12} da matriz T^3. Temos

$$T^3 = \begin{bmatrix} 0,4 & 0,2 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 0,4 & 0,2 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 0,4 & 0,2 \\ 0,6 & 0,8 \end{bmatrix} = \begin{bmatrix} 0,256 & 0,248 \\ 0,744 & 0,752 \end{bmatrix}.$$

Assim, a probabilidade do quarto dia ser satisfatório (S), se o primeiro dia foi insatisfatório (I) é $0,248 = 24,8\%$.

Para o item b), sejam p_1 *e* p_2 *as probabilidades de o ar ser classificado como satisfatório (S) e insatisfatório (I), respectivamente. Como* T *é regular, a longo prazo as probabilidades não dependem das probabilidades iniciais e podemos aplicar o Teorema 4.1.6. Logo,*

$$\begin{bmatrix} 0,4 & 0,2 \\ 0,6 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}.$$

O que nos leva a resolver o sistema linear:

$$\begin{cases} 0,4p_1 & + & 0,2p_2 & = & p_1 \\ 0,6p_1 & + & 0,8p_2 & = & p_2 \end{cases} \Rightarrow \begin{cases} -0,6p_1 & + & 0,2p_2 & = & 0 \\ 0,6p_1 & + & -0,2p_2 & = & 0 \end{cases}$$

Concluímos então que $0,6p_1 = 0,2p_2 \Rightarrow p_2 = 3p_1$. *Como* $p_1 + p_2 = 1$ *segue que:*

$$p_1 + 3p_1 = 1 \Rightarrow p_1 = \frac{1}{4} = 0,25 \quad e \quad p_2 = 0,75.$$

Logo, a longo prazo, a probabilidade de termos dias satisfatórios é $0,25 = 25\%$ *e insatisfatórios* $0,75 = 75\%$.

Exemplo 18. *É observado que as probabilidades de um time de futebol ganhar (G), perder (P) e empatar (E) uma partida depois de conseguir uma vitória são* $\frac{1}{2}$, $\frac{1}{5}$, *e* $\frac{3}{10}$ *respectivamente; e depois de ser derrotado são* $\frac{3}{10}$, $\frac{3}{10}$, *e* $\frac{2}{5}$, *respectivamente; e depois de empatar são* $\frac{1}{5}$, $\frac{2}{5}$, *e* $\frac{2}{5}$, *respectivamente. Se o time não melhorar e nem piorar, conseguirá mais vitórias que derrotas a longo prazo?*

A partir das informações fornecidas no enunciado podemos construir sua matriz de transição:

	G	P	E
G	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{5}$
P	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$
E	$\frac{3}{10}$	$\frac{2}{5}$	$\frac{2}{5}$

$$\Rightarrow T = \begin{bmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,3 & 0,4 \\ 0,3 & 0,4 & 0,4 \end{bmatrix}.$$

Sejam p_G, p_P *e* p_E *as probabilidades de um time de futebol ganhar (G), perder (P) e empatar (E), respectivamente. Como* T *é regular, a longo prazo as probabilidades não* *dependem das probabilidades iniciais e podemos aplicar o Teorema 4.1.6. Logo,*

$$\begin{bmatrix} 0,5 & 0,3 & 0,2 \\ 0,2 & 0,3 & 0,4 \\ 0,3 & 0,4 & 0,4 \end{bmatrix} \cdot \begin{bmatrix} p_G \\ p_P \\ p_E \end{bmatrix} = \begin{bmatrix} p_G \\ p_P \\ p_E \end{bmatrix}.$$

O que nos leva a resolver o sistema linear:

$$\left\{ \begin{array}{ccccccc} 0,5p_G & + & 0,3p_P & + & 0,2p_E & = & p_G \\ 0,2p_G & + & 0,3p_P & + & 0,4p_E & = & p_P \\ 0,3p_G & + & 0,4p_P & + & 0,4p_E & = & p_E \end{array} \right. \Rightarrow \left\{ \begin{array}{ccccccc} -0,5p_G & + & 0,3p_P & + & 0,2p_E & = & 0 \\ 0,2p_G & + & -0,7p_P & + & 0,4p_E & = & 0 \\ 0,3p_G & + & 0,4p_P & + & -0,6p_E & = & 0 \end{array} \right.$$

Escalonando este sistema linear, obtemos:

$$\left\{ \begin{array}{ccccccc} -0,5p_G & + & 0,3p_P & + & 0,2p_E & = & 0 \\ & & -0,29p_P & + & 0,24p_E & = & 0 \end{array} \right.$$

Concluímos então que $0,29p_P = 0,24p_E \Rightarrow p_P = \frac{24}{29}p_E$.

Assim,

$$\begin{gathered} -0,5p_G + 0,3p_P + 0,2p_E = 0 \\ -\frac{5}{10}p_G + \frac{3}{10} \cdot \frac{24}{29}p_E + \frac{2}{10}p_E = 0 \\ -\frac{5}{10}p_G + \frac{72}{290}p_E + \frac{2}{10}p_E = 0 \\ -\frac{5}{10}p_G + \frac{130}{290}p_E = 0 \\ -\frac{5}{10}p_G + \frac{13}{29}p_E = 0 \\ \frac{13}{29}p_E = \frac{5}{10}p_G = \frac{1}{2}p_G \\ p_G = 2 \cdot \frac{13}{29}p_E \\ p_G = \frac{26}{29}p_E. \end{gathered}$$

Como $p_G + p_P + p_E = 1$ *segue que:*

$$\begin{gathered} \frac{26}{29}p_E + \frac{24}{29}p_E + p_E = 1 \\ \frac{26 + 24 + 29}{29}p_E = 1 \end{gathered}$$

$$\frac{79}{29}p_E = 1$$

$$p_E = \frac{29}{79}.$$

Desse modo,

$$p_G = \frac{26}{29}p_E = \frac{26}{29} \cdot \frac{29}{79} = \frac{26}{79}$$

e

$$p_P = \frac{24}{29}p_E = \frac{24}{29} \cdot \frac{29}{79} = \frac{24}{79}.$$

Assim, a longo prazo, a probabilidade de um time de futebol ganhar uma partida será $\frac{26}{79}$*, perder* $\frac{24}{79}$ *e empatar* $\frac{29}{79}$ *Ou seja, a probabilidade será de aproximadamente* $32,9\%$ *de ganhar,* $30,4\%$ *de perder e* $36,7\%$ *de empatar. Logo, o time conseguirá mais vitórias que derrotas a longo prazo.*

Exemplo 19. *Insetos da ordem blattodea, quando submetidos a altas temperaturas, se deslocam de maneira aleatória. Considere o seguinte modelo: uma barata (periplaneta americana) se desloca em um tabuleiro* 3×3 *de maneira aleatória, e de acordo com a seguinte regra: a probabilidade do inseto se deslocar, a cada passo, para uma das casas vizinhas (inclusive na diagonal), é idêntica. Considere que o inseto não permanece na casa em que estava no instante anterior.*

Construa a matriz de transição correspondente e mostre que depois de um tempo suficientemente grande, a probabilidade do inseto se encontrar na casa do meio é de 1/5.

Na figura a seguir temos um modelo do tabuleiro, onde as casas foram numeradas de 1 a 9. Note que, por exemplo, a probabilidade do inseto sair da casa 4 e ir para uma de suas casas vizinhas (1, 2, 5, 7 e 8) será de $\frac{1}{5}$ *e para as demais casas (3, 6 e 9) temos probabilidade 0.*

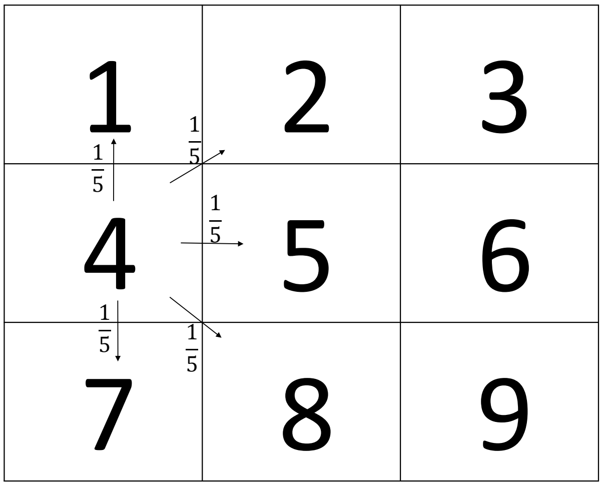

Figura 5.4: Probabilidades de transição da casa 4 para as casas vizinhas.

Temos que a matriz de transição será dada por:

$$T = \begin{bmatrix} 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{8} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} & \frac{1}{8} & \frac{1}{5} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & \frac{1}{8} & \frac{1}{5} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{5} & 0 & 0 & \frac{1}{8} & 0 & \frac{1}{3} & \frac{1}{5} & 0 \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{3} & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{3} & \frac{1}{5} & \frac{1}{3} \\ 0 & \frac{1}{5} & \frac{1}{3} & 0 & \frac{1}{8} & 0 & 0 & \frac{1}{5} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{8} & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{8} & \frac{1}{5} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{1}{5} & 0 & \frac{1}{5} & 0 \end{bmatrix}.$$

Para os cálculos deste problema, utilizamos o software "caculadora de matrizes"[7], no qual é uma ferramenta para cálculos envolvendo matrizes e resolução de sistemas lineares.

Assim, T^2 é dada por:

$$T^2 = \begin{bmatrix} \frac{7}{40} & \frac{13}{200} & \frac{13}{120} & \frac{13}{200} & \frac{1}{20} & \frac{13}{200} & \frac{13}{120} & \frac{13}{200} & \frac{1}{24} \\ \frac{13}{120} & \frac{143}{600} & \frac{13}{120} & \frac{11}{120} & \frac{2}{15} & \frac{11}{120} & \frac{13}{120} & \frac{21}{200} & \frac{13}{120} \\ \frac{13}{120} & \frac{13}{200} & \frac{7}{40} & \frac{13}{200} & \frac{1}{20} & \frac{13}{200} & \frac{1}{24} & \frac{13}{200} & \frac{13}{120} \\ \frac{13}{120} & \frac{11}{120} & \frac{13}{120} & \frac{143}{600} & \frac{2}{15} & \frac{21}{200} & \frac{13}{120} & \frac{11}{120} & \frac{13}{120} \\ \frac{2}{15} & \frac{16}{75} & \frac{2}{15} & \frac{16}{75} & \frac{4}{15} & \frac{16}{75} & \frac{2}{15} & \frac{16}{75} & \frac{2}{15} \\ \frac{13}{120} & \frac{11}{120} & \frac{13}{120} & \frac{21}{200} & \frac{2}{15} & \frac{143}{600} & \frac{13}{120} & \frac{11}{120} & \frac{13}{120} \\ \frac{13}{120} & \frac{13}{200} & \frac{1}{24} & \frac{13}{200} & \frac{1}{20} & \frac{13}{200} & \frac{7}{40} & \frac{13}{200} & \frac{13}{120} \\ \frac{13}{120} & \frac{21}{200} & \frac{13}{120} & \frac{11}{120} & \frac{2}{15} & \frac{11}{120} & \frac{13}{120} & \frac{143}{600} & \frac{13}{120} \\ \frac{1}{24} & \frac{13}{200} & \frac{13}{120} & \frac{13}{200} & \frac{1}{20} & \frac{13}{200} & \frac{13}{120} & \frac{13}{200} & \frac{7}{40} \end{bmatrix}.$$

Logo, T é regular e podemos aplicar o Teorema 4.1.6.

Sejam p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8 e p_9 as probabilidades do inseto estar, depois de um longo período de tempo ($n \longrightarrow \infty$), nas casas 1, 2, 3, 4, 5, 6, 7, 8 e 9, respectivamente.

Logo, aplicando o Teorema, item (iv), temos

$$\begin{bmatrix} 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{8} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} & \frac{1}{8} & \frac{1}{5} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & \frac{1}{8} & \frac{1}{5} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{5} & 0 & 0 & \frac{1}{8} & 0 & \frac{1}{3} & \frac{1}{5} & 0 \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{3} & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{3} & \frac{1}{5} & \frac{1}{3} \\ 0 & \frac{1}{5} & \frac{1}{3} & 0 & \frac{1}{8} & 0 & 0 & \frac{1}{5} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{8} & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{8} & \frac{1}{5} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{1}{5} & 0 & \frac{1}{5} & 0 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \\ p_8 \\ p_9 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \\ p_8 \\ p_9 \end{bmatrix}.$$

O que nos leva a resolver o sistema linear:

$$\left\{\begin{array}{ccccccccccccccccccc}
0p_1 & + & \frac{1}{5}p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & p_1\\
\frac{1}{3}p_1 & + & 0p_2 & + & \frac{1}{3}p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & p_2\\
0p_1 & + & \frac{1}{5}p_2 & + & 0p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & p_3\\
\frac{1}{3}p_1 & + & \frac{1}{5}p_2 & + & 0p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & \frac{1}{3}p_7 & + & \frac{1}{5}p_8 & + & 0p_9 & = & p_4\\
\frac{1}{3}p_1 & + & \frac{1}{5}p_2 & + & \frac{1}{3}p_3 & + & \frac{1}{5}p_4 & + & 0p_5 & + & \frac{1}{5}p_6 & + & \frac{1}{3}p_7 & + & \frac{1}{5}p_8 & + & \frac{1}{3}p_9 & = & p_5\\
0p_1 & + & \frac{1}{5}p_2 & + & \frac{1}{3}p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & 0p_7 & + & \frac{1}{5}p_8 & + & \frac{1}{3}p_9 & = & p_6\\
0p_1 & + & 0p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & 0p_7 & + & \frac{1}{5}p_8 & + & 0p_9 & = & p_7\\
0p_1 & + & 0p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & \frac{1}{3}p_7 & + & 0p_8 & + & \frac{1}{3}p_9 & = & p_8\\
0p_1 & + & 0p_2 & + & 0p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & \frac{1}{5}p_8 & + & 0p_9 & = & p_9
\end{array}\right. \Rightarrow$$

$$\left\{\begin{array}{ccccccccccccccccccc}
-1p_1 & + & \frac{1}{5}p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & 0\\
\frac{1}{3}p_1 & - & 1p_2 & + & \frac{1}{3}p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & 0\\
0p_1 & + & \frac{1}{5}p_2 & - & 1p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & 0p_8 & + & 0p_9 & = & 0\\
\frac{1}{3}p_1 & + & \frac{1}{5}p_2 & + & 0p_3 & - & 1p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & + & \frac{1}{3}p_7 & + & \frac{1}{5}p_8 & + & 0p_9 & = & 0\\
\frac{1}{3}p_1 & + & \frac{1}{5}p_2 & + & \frac{1}{3}p_3 & + & \frac{1}{5}p_4 & - & 1p_5 & + & \frac{1}{5}p_6 & + & \frac{1}{3}p_7 & + & \frac{1}{5}p_8 & + & \frac{1}{3}p_9 & = & 0\\
0p_1 & + & \frac{1}{5}p_2 & + & \frac{1}{3}p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & - & 1p_6 & + & 0p_7 & + & \frac{1}{5}p_8 & + & \frac{1}{3}p_9 & = & 0\\
0p_1 & + & 0p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & 0p_6 & - & 1p_7 & + & \frac{1}{5}p_8 & + & 0p_9 & = & 0\\
0p_1 & + & 0p_2 & + & 0p_3 & + & \frac{1}{5}p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & \frac{1}{3}p_7 & - & 1p_8 & + & \frac{1}{3}p_9 & = & 0\\
0p_1 & + & 0p_2 & + & 0p_3 & + & 0p_4 & + & \frac{1}{8}p_5 & + & \frac{1}{5}p_6 & + & 0p_7 & + & \frac{1}{5}p_8 & - & 1p_9 & = & 0
\end{array}\right.$$

Resolvendo este sistema obtemos: $p_1 = p_9$, $p_2 = \frac{5}{3}p_9$, $p_3 = p_9$, $p_4 = \frac{5}{3}p_9$, $p_5 = \frac{8}{3}p_9$, $p_6 = \frac{5}{3}p_9$, $p_7 = p_9$, $p_8 = \frac{5}{3}p_9$ *e* $p_9 = p_9$.

Como $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 + p_9 = 1$. *Segue que:*

$$p_9 + \frac{5}{3}p_9 + p_9 + \frac{5}{3}p_9 + \frac{8}{3}p_9 + \frac{5}{3}p_9 + p_9 + \frac{5}{3}p_9 + p_9 = 1 \Longrightarrow p_9 = \frac{3}{40}.$$

Como estamos interessados em encontrar a probabilidade do inseto se encontrar na casa do meio, estamos procurando o valor de p_5. *Logo,*

$$p_5 = \frac{8}{3}p_9 = \frac{8}{3} \cdot \frac{3}{40} = \frac{1}{5}.$$

Note ainda que: $p_1 = 0,075$, $p_2 = 0,125$, $p_3 = 0,075$, $p_4 = 0,125$, $p_5 = 0,2$, $p_6 = 0,125$, $p_7 = 0,075$, $p_8 = 0,125$ *e* $p_9 = 0,075$. *Assim, podemos concluir também que a casa do meio tem maior chance da barata ser encontrada.*

Exemplo 20. *Podemos resolver o item b do Exemplo10 utilizando matriz de transição e o Teorema 4.1.6. Para isto, denotemos por **1** o estado livre e por **2** o estado ocupado. Os elementos da matriz de transição são dados por*

$a_{11} = P(A_{n+1}|A_n) = 1 - p;$ $\qquad a_{12} = P(A_{n+1}|B_n) = q;$

$a_{21} = P(B_{n+1}|A_n) = p;$ $\qquad a_{22} = P(B_{n+1}|B_n) = 1 - q.$

Logo,

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

Pelo Teorema 4.1.6, para encontrarmos $\lim\limits_{n\to\infty} x_n$*, basta resolvermos o sistema*

$$\begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix} \cdot \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix},$$

onde $x = \lim\limits_{n\to\infty} x_n$ *é a probabilidade do telefone estar livre e* $1-x$ *é a probabilidade dele estar ocupado a longo prazo.*

Temos

$$\left\{\begin{array}{ccccc} (1-p)x & + & q(1-x) & = & x \\ px & + & (1-q)(1-x) & = & 1-x \end{array}\right. \Longrightarrow \left\{\begin{array}{ccccc} (-p-q)x & + & q & = & 0 \\ (p+q)x & - & q & = & 0 \end{array}\right.$$

Obtemos então

$$x = \frac{q}{p+q}.$$

5.1 Plano de aula

O conteúdo de matrizes, sistemas lineares e probabilidade é trabalhado nas turmas do 2º ano do Ensino Médio de acordo com o Currículo do Estado de São Paulo [11]. Apresentamos a seguir uma sugestão de plano de aula para aplicação dos problemas propostos anteriormente sobre cadeias de Markov para alunos do Ensino Médio.

Plano de Aula:

Tema: Uma abordagem matricial sobre cadeias de Markov.

Objetivos Gerais:

- Utilizar elementos de matrizes para organizar e justificar a resolução de situações-problema envolvendo cadeias de Markov.

Objetivos Específicos:

- Compreender o significado das cadeias de Markov;

- Adquirir conhecimentos básicos sobre probabilidade;

- Saber expressar, por meio de matrizes, situações relativas a cadeias de Markov;

- Saber resolver sistemas de equações lineares.

Conteúdo:

- Matrizes: Representação de uma matriz, operações com matrizes, representação matricial de sistemas lineares, matrizes estocásticas;

- Noções de probabilidade;

- Cadeias de Markov: processos estocásticos, processos de Markov, matriz das probabilidades de transição, matrizes regulares, vetor de probabilidade a longo prazo.

Metodologia:

- Aula expositiva do conteúdo proposto;

- Apresentação das situações-problema propostas neste trabalho para a resolução em sala de aula;

- Observar os recursos utilizados pelos alunos no desenvolvimento dos problemas sugeridos utilizando a "calculadora de matrizes";

- Apresentar a resolução dos problemas envolvendo cadeias de Markov utilizando o Teorema 4.1.6.

Recursos Didáticos:

- Quadro, giz e sala de informática.

Avaliação:

- Observação da participação do aluno na resolução dos problemas sugeridos;

- Entrega da atividade escrita desenvolvida em sala de aula.

5.2 Aplicação dos problemas propostos nesta dissertação

Relatos da experiência em sala de aula

Para a aplicação dos exemplos deste trabalho foi selecionada uma escola estadual da cidade de Bauru e foi trabalhado o assunto com os alunos de uma turma do 2º ano do

Ensino Médio. Como leciono nesta escola, já tinha a informação sobre os conhecimentos prévios de meus alunos. Inicialmente, foram trabalhados os seguintes conteúdos que compõem o Currículo do Estado de São Paulo para alunos nesta série:

- Matrizes: definição, operações com matrizes, matriz inversa e determinantes.
- Sistemas lineares: equações lineares, resolução de sistemas lineares: método de Cramer e escalonamento.
- Análise combinatória e probabilidade: métodos de contagem, probabilidade, probabilidade condicional.

O conteúdo que envolve matrizes e sistemas lineares foi trabalhado durante o 2º bimestre e o conteúdo de análise combinatória e probabilidade foi trabalhado no 3º bimestre. Em seguida, foi apresentado o contexto de cadeias de Markov: processos estocásticos, processos de Markov, matriz das probabilidades de transição, matrizes regulares, vetor de probabilidade a longo prazo. Ao tomar conhecimento sobre o assunto, foi proposto a resolução do Exemplo 17 apresentado neste trabalho através das aproximações do vetor $v^n = T^n v^0$ para valores de n cada vez maiores. Para isto, os alunos foram levados para uma sala de informática e foi utilizado o software "calculatora de matrizes"a fim de facilitar os cálculos das potências da matriz de transição encontrada. Assim, os alunos puderam perceber que havia uma possível convergência da matriz de transição a medida que os valores de n aumentavam.

Em seguida, foi apresentado aos alunos o Teorema 4.1.6 a fim de mostrar uma ferramenta que facilite os cálculos e nos dê uma previsão concisa a longo prazo para matrizes estocásticas regulares. Com isso, foi sugerido a resolução do problema anterior, utilizando agora, o Teorema apresentado.

A aplicação da atividade envolvendo cadeias de Markov levou uma semana, ocupando 5 aulas de 50 minutos.

Embora sabemos que em uma sala de aula há muitos alunos desinteressados e com grande defasagem de conteúdos envolvendo os conhecimentos prévios sobre os assuntos trabalhados no Ensino Fundamental, alguns alunos conseguiram desenvolver toda a atividade proposta corretamente e relataram gostar da atividade diferenciada em sala de aula.

O tema cadeias de Markov, permite ao professor, desenvolver assuntos que, muitas vezes, é de difícil compreensão por parte dos alunos e assim, despertar a curiosidade dos mesmos. Concluímos assim que, nós professores, devemos estimular nossos alunos a buscar novos conhecimentos e mostrar como a matemática pode ser aplicada na resolução de alguns problemas cotidianos.

Referências Bibliográficas

[1] JAMES, B. R. **Probabilidade:** um curso em nível intermediário. 2. ed. Rio de Janeiro: IMPA, 1996.

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

[3] HAIRER, M. **Ergodic properties of Markov processes.** Warwick: 2006. Lecture notes. Disponível em: <http://www.hairer.org/notes/Markov.pdf>. Acesso em: 9 dez. 2017.

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

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

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

[7] MATRIX CALCULATOR. Disponível em: <https://matrixcalc.org/pt/>. Acesso em: 5 dez. 2017.

[8] BOLDRINI, J. L. et al. **Álgebra linear.** 3. ed. Campinas: Universidade Estadual de Campinas, Departamento de Matemática, 1980.

[9] SILVA, F. B.; ROTA, I. S. Convergência de matrizes estocásticas regulares. **C.Q.D:** revista eletrônica paulista de matemática, Bauru, v. 8, p. 4-14, dez. 2016.

[10] MALAJOVICH, G. **Álgebra linear**. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Departamento de Matemática Aplicada, 2010. Disponível em: <http://www.labma.ufrj.br/ gregorio/livro/al2.pdf>. Acesso em: 9 dez. 2017.

[11] São Paulo (Estado). Secretaria da Educação. **Currículo do Estado de São Paulo:** matemática e suas tecnologias. São Paulo, 2011.