Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica

Mestrado Profissional em Matemática

Cadeias de Markov Homogêneas discretas

Dissertação de Mestrado

Francisco Zuilton Gonçalves Vieira

Prof. Dr. Simão Nicolau Stelmastchuk

Orientador

Campinas-SP, 2011

Cadeias de Markov Homogêneas Discretas

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por **Francisco Zuilton Gonçalves Vieira** e aprovada pela Comissão Julgadora.

Campinas, 28 de Fevereiro de 2011.

Simão Nicolau Stelmastchuk

Prof. Dr. Simão Nicolau Stelmastchuk
Orientador

Banca Examinadora:
Prof. Dr. Simão Nicolau Stelmastchuk (FAFIUV)
Prof. Dr. Cristiano Torezzan (FCA-UNICAMP)
Prof. Dr. Marcelo Oliveira Veloso (UFSJ)

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do título de **Mestre em Matemática**.

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Fabiana Bezerra Müller – CRB8 / 6162

Vieira, Francisco Zuilton Gonçalves

V673c Cadeias de Markov homogêneas discretas/Francisco Zuilton Gonçalves Vieira-- Campinas, [S.P. : s.n.], 2011.

Orientador : Simão Nicolau Stelmastchuk

Dissertação (mestrado profissional) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1.Probabilidades. 2.Markov, Cadeias de. I. Stelmastchuk, Simão Nicolau. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: Discrete homogeneous Markov chains

Palavras-chave em inglês (Keywords): 1. Probabilities. 2. Markov chains.

Titulação: Mestre em Matemática

Banca examinadora: Prof. Dr. Simão Nicolau Stelmastchuk (FAFIUV)
Prof. Dr. Cristiano Torezzan (FCA - UNICAMP)
Prof. Dr. Marcelo Oliveira Veloso (UFSJ)

Data da defesa: 28/02/2011

Programa de Pós-Graduação: Mestrado Profissional em Matemática

Dissertação de Mestrado Profissional defendida em 28 de fevereiro de 2011 e aprovada pela Banca Examinadora composta pelos Profs. Drs.

Simão N. Stelmastchuk

Prof. (a). Dr (a). SIMÃO NICOLAU STELMASTCHUK

Prof. (a). Dr (a). CRISTIANO TOREZZAN

Prof. (a). Dr (a). MARCELO OLIVEIRA VELOSO

Aos meus familiares e amigos que
partiram deixando saudades.
Dedico

Agradecimentos

À Deus, em quem acredito.
À minha família, pela paciência.

Ao Prof. Dr. Simão Nicolau Stelmastchuk que, com sua competência e dedicação, tornou possível a elaboração deste trabalho.

Ao corpo de professores do Mestrado Profissional em Matemática, de modo especial, à Profª. Drª. Sueli Irene Rodrigues Costa, por sua atençao durante o curso.

A todos os colegas do mestrado, em especial, André, Waléria, Portela, Coimbra e Luis Faustino, pela amizade e companheirismo.

Ao amigo Prof. Carlos Hentique Sales, pelo incentivo.

Aos Professores e funcionários do Centro de Estudos Superiores de Caxias/UEMA, na pessoa da Profª. Isabel Dolores Leão Brito (*in memoria*)

Instruir-te-ei, ensinar-te-ei o caminho que deves seguir; guiar-te-ei com os meus olhos.
Salmo 32:8.

Resumo

Está dissertação tem como tema o estudo das cadeias de Markov discretas com valores em um espaço de estados enumerável. Cadeias de Markov são processos estocásticos no seguinte sentido: dado o momento presente, o futuro não depende do passado, mas somente do momento presente.

Nosso estudo é realizado sobre cadeias de Markov homogêneas (CMH) discretas. Inicialmente, introduzimos a definição e conceitos básicos das CMH discretas. Tais estudos nos conduzem ao conceito de topologia das matrizes de Transição associada as CMH. A topologia de tais cadeias é a ferramenta necessária para o estudo dos conjuntos recorrentes e transcientes, os quais são de grande importância nesta teoria. O estudo de estados estacionários e a propriedade forte de Markov também são abordados. Esta última propriedade serve para construção do conceito de estado recorrente. A partir deste último conceito trabalhamos com os conceitos de positivo e nulo recorrente. Por fim, estudamos o importante conceito de tempo absorção, o qual é entendido como o tempo que algum estado é absorvido a um conjunto recorrente.

Palavras chaves: Probabilidades, cadeias de Markov

Abstract

This dissertation deals with the study of discrete Markov chains with values in a countable state space. Markov chains are processes stochastic in the following sense: given the present moment, the future does not depend on the past, but only in the present moment. Our study is conducted on homogeneous Markov chains (HMC) discrete. Initially, we introduced the definition and the basic concepts of discrete HMC.

Such studies lead us to understand the concept of topology Transition matrices associated to HMC. The topology of these chains is a necessary tool for the study of the recurrent and transient sets, which are of great importance in this theory. The study of steady states and the strong Markov properties are also addressed. This latter property serves to build the concept of recurrent state. From this latter concept we work with the concepts of positive and null recurrent. Finally, we studied the important concept of absorption time, which is understood as the time that some state is absorbed to a set recurrent.

Keywords: Probabilities, Markov chains.

Sumário

Introdução

O estudo de processos estocásticos remonta aos tempos mais antigos, muitos estudiosos, devotaram grande parte de seu tempo na análise e desenvolvimento de teorias com vistas a dotar este processos de uma estrutura teórica sólida. Processos aleatórios cuja propriedade caracteristica principal é a de que os resultados obtidos em um tempo passado não tem influência sobre o que poderá ocorrer em um tempo futuro, ou seja o conhecimento do que acontece no presente é o que importa para se prever o que poderá vir acontecer no futuro, estes processos são denominados de processos markovianos, e quando o tempo é discreto de Cadeias de Markov.

Neste trabalho, tratamos da análise das cadeias de Markov que evoluem em tempo discreto. Cumprindo o propósito do Mestrado Profissinal o nosso trabalho procura auxiliar quem deseja iniciar nesta teoria. Com está visão desenvolvemos um trabalho minucioso nas demonstrações e cálculos que se encontram em nosso estudo. Acreditamos cumprir com uma necessidade na literatura de Cadeias de Markov quanto ao detalhamento de suas contas e demonstrações, principalmente para o iníco desta teoria. Além disso, trazemos um texto em português sobre cadeias de Markov homogêneas discretas infinitas enumeráveis, o qual até o momento de nosso estudo não havíamos encontrado.

Processos Markovianos são estudados em três casos distintos quanto ao espaço de estados: discretos finitos, discretos enumeráveis e contínuos. Nosso foco é no caso discreto enumerável, ou seja, o espaço de estado é um conjunto enumerável. Este estudo se diferencia profundamente do caso discreto finito, como alguém poderia pensar. Um exemplo disto é o Teorema 3.9.

No primero capítulo, introduzimos uma pequena introdução dos conceitos, definições e propriedades de Teoria de Probabilidade as quais serão necessárias em nossos propósitos posteriores.

No segundo capítulo, a definição de cadeia de Markov homogênea discreta (CMHD) é apresentada. Após a definição, apresentamos alguns exemplos de CMHD. Continuando definimos a matriz de transição associada a uma CMHD e mostramos como descrever os termos no m-ésimo passo desta matriz. Apresentamos, também, uma forma de construir cadeias de Markov e mostramos que alguns exemplos conhecidos de CMHD são criados a partir desta construção. O conceito de Análise de primeiro passo é desenvolvido de maneira extensiva através de exemplos. Este é o germe dos conceito de absorção apresentado no capítulo 3. Com estes conceitos básicos entendidos introduzimos um conceito sofisticado: topologia da matriz de transição. O conceito de distribuição estacionária encontrado no caso finito é abordado em nosso estudo, com as devidas considerações. Além disso, apresentamos a propriedade forte de Markov, o qual é uma propriedade necessária para o estudo do número de visitas a um estado e dos ciclos regenerativos, ambos precurssores dos conceitos de recorrência e transciência.

No capítulo três, introduzimos os conceitos de recorrência e transciência, mais ainda, os conceitos de recorrência positiva e recorrência nula. Sendo o caso de recorrência positiva o mais interessante, pois ele diz que saindo de um estado i a cadeia de Markov com está propriedade retorna a i infintas vezes em um tempo médio finito. Apresentamos ferramentas que visam a caracaterizar condições para que a CMHD possua está propriedade. Uma outra características das CMHD é a possível existência de estados transcientes, estados em que a CMHD pode ou não retornar a eles. Aqui o interesse da teoria é distinguir se estes estados serão absorvidos por um conjunto recorrente ou não. Assim, estudamos uma série de resultados que auxiliam na elucidação destes casos.

Por fim, deixamos claro que não esgotamos o estudo de CMHD, pois ainda existem muitos tópicos para serem estudados nesta Teoria, como por exemplo a questão de acoplamento, ergodicidade e cadeias de Markov Discretas não homogêneas, entre outros. Em resumo, nosso trabalho se caracteriza como um material introdutório para o estudo de cadeias de Markov homgêneas discretas.

Capítulo 1

Conceitos básicos

1.1 Probabilidades

Muitos estudiosos têm contribuído para desenvolvimento da Teoria da Probabilidade, que teve início em épocas remotas, passando pelos trabalhos de Cardano, Paccioli, Tartaglia e Galileu. O primeiro grande tratado de probabilidade foi *Ars Conjectandi* - A Arte da Conjectura escrito pelo matemático suíço Jaques Bernoulli(1654-1705). Este trabalho não estava concluído quando de seu falecimento, sendo então concluído por seu sobrinho Nicolaus(I) Bernoulli (1687-1759), que também deixou contribuições importantes teoria de probabilidades, onde se destaca o Paradoxo de São Petersburgo, descrito a seguir: dois jogadores lançam uma moeda honesta, um deles A paga ao outro jogador B, 2^k unidades monetárias se o acontecimento *cara* aparece pela primeira vez no k-ésimo lançamento da moeda. O problema é determinar o quanto B deve pagar a A para entrar no jogo de modo que os valores esperados dos ganhos de cada um sejam iguais e, sendo assim o jogo seria honesto. Observa-se que o valor esperado tende ao infinito, o que obrigaria a B a pagar uma quantia infinita para entrar no jogo.

O estudo de problemas como este envolvem conceitos de probabilidade, que trataremos a seguir.

Neste capítulo seguiremos como referência [1], [2], [3].

1.2 Elementos de Probabilidade

Modelos matemáticos podem *ser determinísticos* quando as condições sob as quais o experimento é realizado determinam o resultado do experimento e *não determinísticos* quando não é possível prever de antemão seus resultados. O objeto da probabilidade é o estudo dos experimentos *não determinísticos.* Existem quatro interpretações relacionadas à probabilidade:

a) Clássica: se em um experimento em que o número de resultados possíveis é n e os resultados são mutuamente exclusivos e igualmente prováveis e, sendo n_A, o número de ocorrências do evento A, e n o número de resultados possíveis, então a probabilidade de ocorrer o evento A é dada pela razão entre o número de ocorrências do evento A (n_A) e o número total de possíveis resultados do experimento (n) sendo denotado por $\mathbb{P}(A) = \dfrac{n_A}{n}$.

b) Frequência relativa: Se um experimento é repetido n vezes em condições idênticas e o número de resultados de um evento B é n_B, o limite da razão de n_B por n com n suficientemente grande é definido como a probabilidade de ocorrência de B.

c) Subjetiva. Quando o observador estima a probabilidade de ocorrência do evento.

d) Axiomática. A que usaremos neste trabalho e será definida adiante.

No que segue, daremos algumas definições básicas com respeito a teoria de probabilidade.

Definição 1.0.1. *O conjunto de pontos que representam os possíveis resultados de um experimento é chamado espaço amostral do experimento.*

Definição 1.0.2. *Um espaço amostral é denominado discreto se os resultados podem ser postos em correspondência biunívoca com o conjunto dos números inteiros positivos.*

Definição 1.0.3. *Seja Ω o espaço amostral do experimento. Todo subconjunto $A \subset \Omega$ será denominado de evento. Ω é o evento certo e $\emptyset$ o evento impossível. Se $\omega \in \Omega$, o evento $\{\omega\}$ é chamado de evento elementar.*

Definição 1.0.4. *Um evento A ao qual atribuímos uma probabilidade será denominado de evento aleatório.*

Nem sempre é possível atribuir uma probabilidade a um evento, este fato não constitui dificuldade, pois podemos restringir o espaço amostral. Maiores informações encontramos em [1].

Uma probabilidade é uma lei que associa a cada evento de Ω um número. Por razões técnicas não adotaremos o conjunto das partes de Ω, ou seja, $P(\Omega)$ para desenvolver a teoria de probabilidades. Aqui iremos nos restringir às σ-álgebras.

Definição 1.0.5. *Seja Ω um conjunto não vazio. Uma classe $\mathbb{A}$ de subconjuntos de Ω satisfazendo às seguintes condições:*

i) $\Omega \in \mathbb{A}$.

ii) $A \in \mathbb{A}$ *e* $A^c \in \mathbb{A}$

iii) *Se* $A_n \in \mathbb{A}$ *para* $n= 1\ ,\ 2\ ,\ 3\ ,\ \ldots,$ *então* $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathbb{A}$.

é denominada uma σ-álgebra de subconjuntos de Ω.

Definição 1.0.6. *Dois eventos A e B são mutuamente exclusivos se não ocorrem simultaneamente, ou seja, $A \cap B = \emptyset$.*

Proposição 1.0.1. *Seja $\mathbb{A}$ uma σ-álgebra de subconjuntos de Ω, vale então, as seguintes propriedades:*

i) $\emptyset \in \mathbb{A}$;

ii) $\forall n, A_1, \ldots, A_n$, *temos* $\displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathbb{A}$.

A demonstração é imediata da definição 1.2.5.

Usando o conceito de σ-álgebra definimos probabilidade do seguinte modo

Definição 1.0.7. Axiomática de Probabilidade *uma medida de probabilidade definida em uma σ-álgebra $\mathbb{A}$ de Ω é uma função $P : \mathbb{A} \longrightarrow [0, 1]$ que satisfaz as seguintes condições:*
a) $0 \leq \mathbb{P}(A) \leq 1$, *para todo A pertencente ao espaço amostral Ω;*

b) $P(\Omega) = 1$

c) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, *com os* $A_i's$ *disjuntos.*

Definição 1.0.8. *Uma variável aleatória é uma função* $X : \Omega \longrightarrow E$. *E é finito ou infinito enumerável tal que para todo* $a \in \mathbb{R}$, *ao evento* $\{X = a\} = \{\omega; X(\omega) = a\}$ *pode ser atribuida uma probabilidade, ou seja,*

$$\{X = a\} \in \mathbb{A}.$$

A atribuição de uma probabilidade a um evento pode ser efetuada com base em informações de características teóricas que o fenômeno aleatório apresenta ou ainda através das freqüências relativas observadas na ocorrência da variável que está sendo observada, isto é, para um número grande de experimentos a frequência relativa pode ser utilizada como a probabilidade de ocorrer o evento.

A probabilidade da união de eventos é calculada através da *regra da adição de probabilidades*, ou seja, se A e B são eventos do espaço amostral Ω, então

$$\mathbb{P}(A \cup B) = P(A) + P(B) - P(A \cap B).$$

No caso em que A e B são disjuntos, a expressão acima reduz-se à soma das probabilidades de A e B. Uma consequência da regra da adição é a de que para qualquer evento $A \subset \Omega$, temos

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c).$$

1.3 Probabilidade condicional e Independência estocástica

Os fenômenos aleatórios muitas vezes podem ser separados em etapas e, a informação ocorrida em uma determinada etapa, pode influenciar nas probabilidades de ocorrências das etapas que se sucedem. A informação que um fenômeno aleatório ocorreu permite obter um novo cálculo das probabilidades. A nova probabilidade obtida é denominada *probabilidade condicional.*

Definição 1.0.9. *Dados dois eventos A e B, a probabilidade condicional de A dado que já ocorreu B é definida por*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \bigcap B)}{\mathbb{P}(B)}, \ \ \mathbb{P}(B) > 0, \tag{1.1}$$

também chamamos $\mathbb{P}(A|B)$ de probabilidade condicional de A dado que B ocorreu. Se $P(B) = 0$, $P(A|B)$ pode ser definida de modo arbitrário, e, nesse caso $P(A|B) = P(A)$.

Supondo que A e B são eventos do espaço amostral Ω. Então

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)P(B),$$

com $\mathbb{P}(B) > 0$. É claro que se A e B são independentes

$$\mathbb{P}(A|B) = \mathbb{P}(A), \ \ \text{e} \ \ \mathbb{P}(B|A) = \mathbb{P}(B). \tag{1.2}$$

Da forma simétrica de (1.1)

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)P(B) = \mathbb{P}(B|A)\mathbb{P}(A) \tag{1.3}$$

definida com $\mathbb{P}(B) > 0$.

Regras de Bayes [2]

Teorema 1.1 (Regra da inversão)**.** *Se $\mathbb{P}(A) > 0$, temos que*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}. \tag{1.4}$$

Demonstração: Prova direta de (1.3) □

Teorema 1.2 (Regra de exclusivas e exaustivas causas)**.** *Dados $B_1, B_2, \ldots$ tal que*

$$\sum_{i=1}^{\infty} B_i = \Omega,$$

então para todo $A \in \mathbb{A}$

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i) = \mathbb{P}(A|B_i)\mathbb{P}(B_i). \tag{1.5}$$

Demonstração: Decompondo o evento A em relação aos B_i's temos

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}\left(A \cap \left(\sum_{i=1}^{\infty} B_i\right)\right) = \mathbb{P}\left(\sum_{i=1}^{\infty}(A \cap B_i)\right) \\ &= \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i).\end{aligned}$$

$\square$

Teorema 1.3 (fórmula sequencial)**.** *Para qualquer sequência de eventos $A_1, \ldots, A_n$,*

$$\mathbb{P}(\cap_{i=1}^{k} A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2 \ldots)\mathbb{P}(A_k| \cap_{i=1}^{k-1} A_i).$$

Demonstração:

Procedemos a demonstração por indução. Supondo que 1.3 é verdadeira para k. Escrevemos

$$\mathbb{P}(\cap_{i=1}^{k+1} A_i) = \mathbb{P}((\cap_{i=1}^{k} A_i) \cap A_{k+1}) = \mathbb{P}(A_{k+1}| \cap_{i=1}^{k} A_i)\mathbb{P}((\cap_{i=1}^{k} A_i)$$

e substituímos $\mathbb{P}(\cap_{i=1}^{k} A_i)$ pela igualdade assumida 1.3. $\square$

Definição 1.3.1 (Independência de eventos)**.** *Dados os eventos A e B, se a ocorrência ou não de B não altera a probabilidade da ocorrência de A, ou seja, $\mathbb{P}(A|B) = P(A)$, $\mathbb{P}(B) > 0$, ou ainda, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, então A e B são ditos eventos independentes.*

Definição 1.3.2 (Independência de variávais aleatórias)**.** *As variáveis aleatórias X e Y são independentes se para todo a e b pertencentes a $\mathbb{R}$,*

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b),$$

onde $\{X = a\} = \{\omega; X(\omega) = a\}$ e $\{Y = b\} = \{\omega; Y(\omega) = b\}$.

Definição 1.3.3. *Dizemos que os eventos A e B são condicionalmente independentes dado o evento C se*

$$\mathbb{P}(A \bigcap B|C) = \mathbb{P}(A|C)\mathbb{P}(B|C).$$

Sejam X, Y, Z variáveis aleatórias tomando valores em conjuntos enumeráveis E, F, G, respectivamente. Dizemos que X e Y são condicionalmente independente dado Z se para todo x, y, z em E, F, G, respectivamente, os eventos $\{X = x\}$ e $\{Y = y\}$ são condicionalmente independentes dado $\{Z = z\}$.

Teorema 1.4 (Propriedade de Markov para eventos)**.** *Dados os eventos A_1, A_2 e A_3 com probabilidade positiva. Os eventos A_1 e A_3 serão condicionalmente independentes dado A_2 se, e somente se,*

$$\mathbb{P}(A_3|A_1 \cap A_2) = \mathbb{P}(A_3|A_2).$$

Demonstração: Assumamos a independência condicional. Assim, pela definição de probabilidade condicional e fórmula sequencial de Bayes (ver Teorema 1.3),

$$\begin{aligned}\mathbb{P}(A_3|A_1 \cap A_2) &= \frac{\mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1 \cap A_2)} \\ &= \frac{\mathbb{P}(A_1 \cap A_3|A_2)\mathbb{P}(A_2)}{\mathbb{P}(A_1 \cap A_2)}.\end{aligned}$$

Sendo A_1 e A_3 condicionalmente independente dado A_2, segue que

$$\mathbb{P}(A_3|A_1 \cap A_2) = \frac{\mathbb{P}(A_1|A_2)\mathbb{P}(A_3|A_2)\mathbb{P}(A_2)}{\mathbb{P}(A_1 \cap A_2)}.$$

Observando que $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1|A_2)\mathbb{P}(A_2)$ temos

$$\mathbb{P}(A_3|A_1 \cap A_2) = \mathbb{P}(A_3|A_2).$$

Inversamente, basta seguir o caminho inverso do argumento acima. □

Exemplo 1.4.1 (Lançamento de moedas)**.**

Dada uma sequência de infinitos lançamentos de uma moeda. O espaço amostral Ω será a coleção de todas as sequências

$$\omega = \{X_n\}_{n\geq 1},$$

onde $X_n = 1$ ou $X_n = 0$ dependendo de que o n-ésimo lançamento seja cara ou coroa. Uma amostra do exemplo pode ser:

$$\{1, 1, 0, 1, 0, 0, 0, 1, ...\}$$

Distribuições de probabilidade

A distribuição de probabilidade de uma variável aleatória X em que o contradomínio é $S_X = \{x_1, x_2, \ldots\}$, para o caso em que a variável é discreta é dada pelo conjunto de pares (x_i, p_i), $i = 1, 2, \ldots$, onde os p_i são equivalentes a $P(X = x_i)$. As probabilidades p_i devem satisfazer:

i) $p_i \geq 0$, propriedade da não negatividade e;

ii) $\sum_{i\geq 1} p_i = 1$, propriedade da normalização.

Exemplo 1.4.2 (Distribuição de Poisson). *A distribuição de Poisson é uma distribuição discreta utilizada em situações probabilísticas onde a área de oportunidade de ocorrência de um evento é grande, mas a oportunidade de ocorrência em um intervalo particular é pequena.*

Em termos matemáticos a probabilidade de exatamente s ocorrências de um evento é dada por

$$P(s) = \frac{(\lambda)^s e^{-\lambda}}{s!}$$

em que λ é a média da distribuição e a variância de $P(s)$ é também λ.

1.4 Esperança

Seja X uma variável aleatória discreta com função de probabilidade $p(x_i)$. A espe-rança de X para o caso discreto, é definida por

$$E[X] = \sum_i x_i p(x_i) = \sum_i x_i \mathbb{P}(X = x_i).$$

A esperança de X também recebe a denominação de média de X ou valor esperado de X. A esperança de X é uma média ponderada de acordo com a distribuição de X.

1.5 Esperança Condicional de Variáveis Discretas

Definição 1.4.1. *Sejam X e Y variáveis aleatórias discretas tomando valores nos conjuntos enumeráveis F e G, respectivamente e uma função não negativa $g : F \times G \longrightarrow \mathbb{R}$. Definimos a função $\psi : G \longrightarrow \mathbb{R}$ por:*

$$\psi(y) = \sum_{x \in F} g(x, y) P(X = x | Y = y).$$

Para cada $y \in G$, $\psi(y)$ é denominada de esperança condicional de $g(X, Y)$ dado $Y = y$ sendo denodada por

$$E^{Y=y}[g(X, Y)], \quad ou \quad E[g(X, Y)|Y = y] = \psi(y).$$

Concluímos este capítulo com um exemplo envolvendo os conceitos abordados até aqui.

Exemplo 1.4.3 (Lançamento de moedas)**.**

Considere o experimento: Uma sequência infinita de lançamentos de uma moeda. O espaço amostral Ω será o conjunto de todas as sequências $\omega = \{x_n\}_{n \geq 1}$, onde $x_n = 1$ se no n-ésimo lançamento ter ocorrido cara e $x_n = 0$ se no n-ésimo lançamento for coroa. Como exemplo de uma sequência, podemos ter o evento:

$$A = \{1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, \ldots\}$$

onde $1 \leftrightarrow cara$ e $0 \leftrightarrow coroa$.

Uma variável aleatória X_n será definida como o número obtido no n-ésimo lançamento, isto é,

$$X_n(\omega) = x_n.$$

A probabilidade de um evento $A = \{X_1 = a_1, X_2 = a_2, \ldots, X_k = a_k\}$, onde $a_1, a_2, \ldots, a_k$ são arbitrários no conjunto $\{0, 1\}$, é dada por:

$$\mathbb{P}(A) = \frac{1}{2^k}.$$

Como as variáveis aleatórias X_n são independentes. O evento $\{X_k = a_k\}$ é a soma direta dos eventos $\{X_1 = a_1, X_2 = a_2, \ldots, X_{k-1} = a_{k-1}, X_k = a_k\}$ para todos os valores possíveis de $(a_1, a_2, \ldots, a_{k-1})$. Desde que 2^{k-1} sejam valores e cada um tenha probabilidade 2^{-k}, temos assim,

$$\begin{aligned} \mathbb{P}(X_k = a_k) &= 2^{k-1}2^{-k} \\ \mathbb{P}(X_k = 0) &= \mathbb{P}(X_k = 1) = \frac{1}{2}. \end{aligned}$$

Portanto, $\mathbb{P}(X_1 = a_1, X_2 = a_2, \ldots, X_k = a_k) = \mathbb{P}(X_1 = a_1) \ldots \mathbb{P}(X_k = a_k)$ para todos os $a_1, a_2, \ldots, a_k \in \{0, 1\}$, de onde, por definição de variável independente $X_1, X_2, \ldots, X_k$ são variáveis aleatórias independentes. Agora iremos calcular a esperança ou valor esperado do número de caras ocorridas em n lançamentos, isto é,

$$\begin{aligned} S_n &= X_1 + X_2 + \ldots + X_n \\ s_n &= x_1 + x_2 + \ldots + x_n. \end{aligned}$$

Esta é uma variável aleatória tomando valores inteiros de 0 a n. O evento $\{S_n = k\}$ significa que k variáveis aleatórias $X_1, X_2, \ldots, X_k$ são iguais a 1.

Existem $\begin{pmatrix} n \\ k \end{pmatrix}$ distintas combinações assumindo k valores 1 e as demais $n - k$ combinações tomam o valor 0 nas variáveis aleatórias $X_1, X_2, \ldots, X_n$, com a mesma probabilidade $p = 2^{-n}$. Assim $P(S_n = k) = \begin{pmatrix} n \\ k \end{pmatrix} \frac{1}{2^n}$. A esperança de S_n é dada por

$$
\begin{aligned}
E[S_n] &= \sum_{k=0}^{n} kP(S_n = k) \\
&= \sum_{k=0}^{n} k \left(\begin{array}{c} n \\ k \end{array} \right) \frac{1}{2^n} \\
&= \frac{1}{2^n} \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} \\
&= \frac{1}{2^n} \sum_{k=0}^{n} k \frac{n(n-1)!}{k(k-1)!(n-k)!} \\
&= \frac{n}{2^n} \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} \\
&= \frac{n}{2^n} 2^{n-1} \\
&= \frac{n}{2}.
\end{aligned}
$$

Portanto a esperança é

$$E[S_n] = \frac{n}{2}.$$

Capítulo 2

Cadeias de Markov

O estudo dos processos estocásticos, tendo em vista uma teoria bastante desenvolvida, tem especial importância quando da aplicação às ciências físicas, biológicas, econômicas e sociais. Neste capítulo estudamos os fundamentos das Cadeias de Markov em Tempo Discreto, definições e comportamento, bem como, são apresentados exemplos canônicos que ilustram a teoria das cadeias de Markov. Andrei Andreyevich Markov nasceu no dia 14 de junho de 1856 em Ryazan, na Rússia. Morreu no dia 20 de julho de 1922 em Petrograd (agora St Petersburg), Rússia. Se formou na universidade de St Petersburg (1878), onde se tornou professor em 1886. Os primeiros trabalhos de Markov foram principalmente em teoria dos números, análise, frações contínuas, limites de integrais, teoria da aproximação e a convergência de séries.

Após 1900, Markov aplicou o método das frações contínuas, inicialmente desenvolvido por Pafnuty Chebyshev, na teoria da probabilidade. Estudou as seqüências de variáveis mutuamente independentes, esperando estabelecer as leis da probabilidade de forma mais geral, também provou o teorema central do limite.

Markov é particularmente lembrado pelo estudo de cadeias de Markov. Cadeias de Markov são um formalismo de modelagem de sistemas que descrevem o sistema como um processo estocástico. Deste ponto de vista o sistema modelado é caracterizado pelos seus estados e a forma pela qual eles se alternam.

Em 1923 Norbert Winter se tornou o primeiro a tratar rigorosamente um processo contínuo de Markov. A fundação da teoria geral ocorreu em 1930 por Andrei Kolmogorov.

As principais referências para este capítulo são [2], [7], [10].

2.1 Propriedade Básicas das Cadeias de Markov

A descrição do procedimento de um sistema operando sobre algum período de tempo, são de interesse dos processos estocásticos. Seja uma sequência $\{X_n\}_{n\geq 0}$ de variáveis aleatórias e E um conjunto de estados enumeráveis.

Definição 2.0.2 (Processo Estocástico)**.** *Denominamos processo estocástico ao conjunto de variáveis aleatórias, denotadas por $\{X_n\}_{n\geq 0}$, indexadas por um parâmetro n pertencente a um dado conjunto T, que representa uma característica de interesse, mensurável no tempo n.*

Exemplificando, consideremos o experimento de lançar um dado e observar o número que ocorre na face voltada para para cima, X_n pode representar esta ocorrência. Outro processo estocástico é dado pelo exemplo 1.3.1 do capítulo 1.

Definição 2.0.3 (Cadeia de Markov)**.** *Seja a sequência $\{X_n\}_{n\geq 0}$ um processo estocástico com espaço estado E enumerável. Se para todo inteiro $n \geq 0$ e todos os estados $i_0, i_1, \ldots, i_{n-1}, i, j \in E$ temos*

$$\mathbb{P}(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j|X_n = i), \qquad (2.1)$$

quando ambos os lados da expressão acima estão bem definidos, o processo estocástico é denominado de Cadeia de Markov, e se o lado direito independe de n a cadeia é dita Homogênea.

A expressão (2.1) pode ser interpretada como a probabilidade condicional de qualquer evento futuro ser independente de qualquer evento passado, e depender apenas do estado presente $X_n = i$.

As probabilidades condicionais,

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

são denominadas probabilidades de transição da Cadeia. Quando a probabilidade de transição é independente de n a denominamos de *Probabilidade Estacionária.*

Definição 2.0.4 (Função de transição)**.** *Seja a sequência* $\{X_n\}_{n\geq 0}$ *uma cadeia de Markov homogênea com espaço de estado* E*. A função* $\mathbb{P}(i,j)$*, com* $i, j \in E$*, definida por*

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

denomina-se **função de transição** *e, como os valores da função são probabilidades, devem satisfazer:*

$$P(i,j) \geq 0 \quad e \quad \sum_j \mathbb{P}(i,j) = 1$$

para todo $i, j \in E$*.*

Exemplo 2.0.4 (Ruína do jogador)**.** *Consideremos um jogo em que, se o jogador vence, ganha \$ 1,00 unidade monetária, com probabilidade* $p = 0,4$ *e se não vence, perde \$ 1,00 unidade monetária com probabilidade* $q = 1 - p = 0,6$*. Supondo que o jogador estabeleceu uma regra em que quando atingir a quantia de \$* **N** *unidades monetária, para de jogar e que, quando ficar sem recursos é obrigado a parar. Seja* X_n *a quantia de que dispõe depois de n jogos. Intuitivamente o processo possui a propriedade de Markov. Em outras palavras, que dado o estado corrente, qualquer informação sobre o passado é irrelevante para predizer o próximo estado* X_{n+1}*. Para confirmação, observamos que se o jogador permanece jogando no momento n, seus recursos* $X_n = i$*, estão restritos a* $0 < i < N$*, segue que, para alguma possível história de seu ganho,* $i_{n-1}, \ \ i_{n-2}, \ldots, i_0$ *a probabilidade para a próxima jogada é:*

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = 0,4,$$

de modo que para aumentar seu ganho, terá que vencer a próxima aposta e o resultado da aposta anterior não fornece nenhuma informação que possa ser útil para predizer o próximo resultado.

Exemplo 2.0.5 (Urnas de Ehrenfest)**.** . *Sejam duas caixas identificadas por* **A** *e* **B***, contendo k cartões numerados com 1, 2, ... , k. Parte dos cartões são colocados na caixa* **A** *e o restante na caixa* **B***. É então escolhido aleatoriamente um número compreendido entre* 1 *e* k *e o cartão correspondente a esse número é transferido para a outra caixa. O processo é repetido indefinidamente com os sorteios independentes. Seja* X_n *o número de cartões na caixa* **A** *no n-ésimo sorteio. A sequência* $\{X_n\}_{n\geq 0}$ *é uma cadeia de Markov com espaço de estado* $E = 0, 1, 2, \ldots, k$*. Supondo que existam i cartões na caixa* **A** *no tempo n. A escolha do cartão da caixa* **A** *e transferência para a caixa* **B** *na* $(n + 1)$*-ésima retirada terá probabilidade* $p = \frac{i}{k}$*. Procedendo do mesmo modo, escolhemos um cartão da caixa* **B** *com a probabilidade* $q = \frac{k-i}{k}$ *e transferimos para a caixa* **A***. Então a função de transição é dada por:*

$$P(i,j) = \begin{cases} \frac{i}{k} & se \;\; j = i-1 \\ 1 - \frac{i}{k} & se \;\; j = i+1 \\ 0 & de \; outro \; modo. \end{cases}$$

Os exemplos descritos acima descrevem uma Cadeia de Markov, pois de modo intuitivo, tanto no primeiro como no segundo exemplo, o que acontecerá no próximo passo não depende do que ocorreu anteriormente e somente do estado presente.

Mostraremos abaixo um exemplo mais elaborado que descreve uma cadeia de Markov homogênea, extraido de [2]. Neste exemplo, em favor do leitor, detalhamos todas as passagens.

Exemplo 2.0.6 (Substituição de máquina)**.** *Seja* $\{U_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias independentes e identicamente distribuídas, tomando valores no conjunto* $\{1, 2, \ldots, +\infty\}$*. A variável aleatória* U_n *pode ser interpretada como a vida útil de alguma máquina, no n-ésimo tempo, sendo substituída no* $(n+1)$*-ésimo tempo, após uma falha. Desse modo no tempo 0, a máquina 1 é colocada para funcionar até sua quebra no tempo* U_1*, no que é imediatamente substituída pela máquina 2, que quebra no tempo* $U_1 + U_2$*, e assim por diante. O tempo decorrido de uma máquina em serviço no tempo n é denotado por* X_n*. Assim, o processo* $\{X_n\}_{n\geq 0}$ *toma valores em* $E = \mathbb{N}$ *e cresce linearmente de 0 até o tempo* $R_k = \sum_{i=1}^{k} U_i$ *para* $U_{k+1} - 1$ *no* *tempo* $R_{k+1} - 1$.

A sequência $\{R_k\}_{k\geq 0}$ *definida dessa maneira, com* $R_0 = 0$, *é chamada uma sequência de substituição e* X_n *é chamada de tempo recorrente de volta em* n *(ver figura 2.1).*

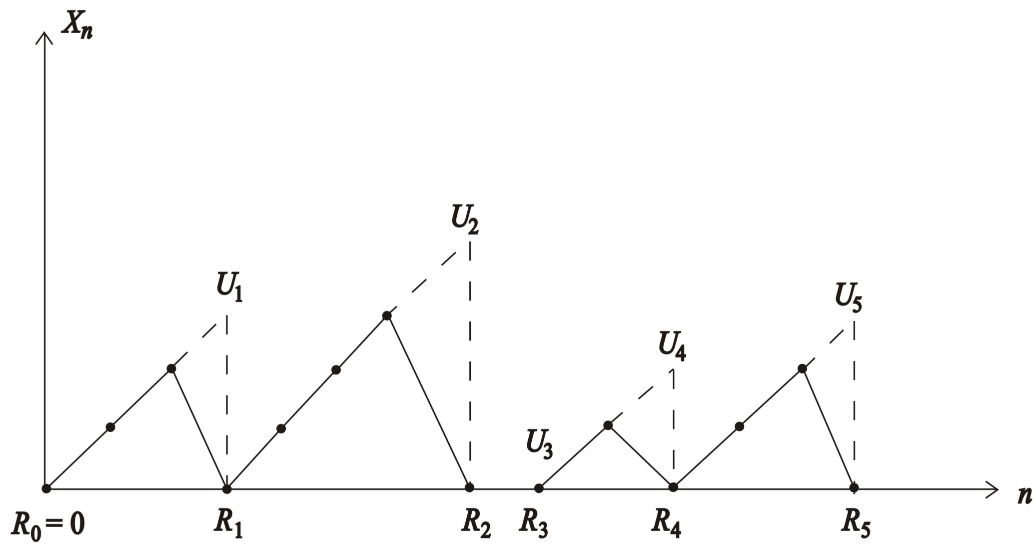

Figura 2.1: Tempo recorrente de volta

O processo $\{X_n\}_{n\geq 0}$ *é uma Cadeia de Markov Homogênea com espaço estado* $E = \mathbb{N}$, *e as entradas não nulas da matriz de transição são da forma* $p_{i,i+1}$ *e* $p_{i,0} = 1 - p_{i,i+1}$, *onde*

$$p_{i,i+1} = \frac{\mathbb{P}(U_1 > i+1)}{\mathbb{P}(U_1 > i)}.$$

No que segue mostraremos que é verdadeira expressão $p_{i,0} = 1 - p_{i,i+1}$. *Por definição*

$$p_{i,i+1} = \mathbb{P}(X_{n+1} = i+1 | X_n = i).$$

Usando a definição de X_n temos que

$$\begin{aligned} 1 &= \mathbb{P}(X_{n+1} = 0 \,\dot{\cup}\, X_{n+1} = i+1 | X_n = i) \\ &= \mathbb{P}(X_{n+1} = 0 | X_n = i) + \mathbb{P}(X_{n+1} = i+1 | X_n = i) \\ &= p_{i,0} + p_{i,i+1}. \end{aligned}$$

Portanto,

$$p_{i,0} = 1 - p_{i,i+1}.$$

Agora, vamos mostrar que

$$p_{i,i+1} = \frac{\mathbb{P}(U_1 > i+1)}{\mathbb{P}(U_1 = i)}. \tag{2.2}$$

Por definição,

$$p_{i,i+1} = \mathbb{P}(X_{n+1} = i+1 | X_n = i) = \frac{\mathbb{P}(X_{n+1} = i+1, X_n = i)}{\mathbb{P}(X_n = i)}.$$

Afirmamos que $\{X_n = i\} = \{U_{l+1} > i\}$. De fato para $n \in \mathbb{N}$, temos que $n \in [R_l, R_{l+1}]$, para algum $l \in \mathbb{N}$. Em $[R_l, R_{l+1}]$ temos que $X_n \in \{0, 1, \ldots, U_{l+1} - 1\}$. Seja $X_n = i \in \{0, 1, \ldots, U_{l+1} - 1\}$. Como U_{l+1} é maior que qualquer dos elementos do conjunto $X_n = i \in \{0, 1, \ldots, U_{l+1} - 1\}$, segue que $U_{l+1} > i$. Portanto,

$$\{X_n = i\} = \{U_{l+1} > i\}.$$

Disto, como as probabilidades são identicamente distribuídas, temos:

$$\mathbb{P}(X_n = i) = \mathbb{P}(U_{l+1} > i) = \mathbb{P}(U_1 > i).$$

De modo análogo, é verdade que

$$\{X_{n+1} = i+1\} = \{U_{l+1} > i+1\}.$$

Logo

$$\{X_{n+1} = i+1, X_n = i\} = \{U_{l+1} > i+1, U_l > i\} = \{U_{l+1} > i+1\},$$

pois, $\{U_{l+1} > i+1\} \subset \{U_l > i\}$*. Isto segue que*

$$\mathbb{P}(X_{n+1} = i+1, X_n = i) = \mathbb{P}(U_{l+1} > i+1).$$

Pelo fato que as U_l*'s são identicamente distribuídas, concluímos que*

$$\mathbb{P}(X_{n+1} = i+1, X_n = i) = \mathbb{P}(U_1 > i+1).$$

Portanto

$$p_{i,i+1} = \frac{\mathbb{P}(U_1 > i+1)}{\mathbb{P}(U_1 > i)}.$$

Em seguida vamos verificar a equação,

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = 0) = \mathbb{P}(X_{n+1} = j | X_n = i_n),$$

ou seja, que $\{X_n\}$ *é uma Cadeia de Markov Homogênea. Fazendo* $B = \{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}$*, escrevemos as probabilidades condicionais acima como*

$$\frac{\mathbb{P}(X_{n+1} = j, X_n = i, B)}{\mathbb{P}(X_n = i, B)} = \frac{\mathbb{P}(X_{n+1} = j, X_n = i)}{\mathbb{P}(X_n = i)}$$

para sequências $i_0, \ldots, i_{n-1}, i, j$*, tal que* $\mathbb{P}(X_n = i, B) > 0$ *e* $\mathbb{P}(X_n = i) > 0$*. Em particular* $j = i+1$ *ou* 0*, e* $i_{n-1} = i-1, \ldots, i_{n-i} = 0$*, pela construção de* X_n*.*

Seja $v(n)$ *o tempo de substituições* R_k *no intervalo* $[1, n]$*. Tomando por exemplo,* $j = i+1$ *e escrevendo* $D = \{X_{n-i-1} = i_{n-i-1}, \ldots, X_0 = 0\}$*, temos*

$$\begin{aligned} & \mathbb{P}(X_{n+1} = j, X_n = i, B) \\ = \; & \mathbb{P}(X_{n+1} = j, X_n = i, X_{n-1} = i-1, \ldots, X_{n-i} = 0, D, \Omega) \\ = \; & \sum_{k=0}^{\infty} \mathbb{P}(X_{n+1} = j, X_n = i, X_{n-1} = i-1, \ldots, X_{n-1} = 0, D, v(n) = k), \quad (2.3) \end{aligned}$$

onde na última igualdade usamos o fato de $\Omega = \bigcup_{k=0}^{\infty} \{v(n) = k\}$*, e que esta união é disjunta.*

Analisemos o termo geral da série acima. Já sabemos que

$$\{X_{n+1} = i+1\} = \{U_{k+1} > i+1\}$$

e

$$\{X_n = i\} = \{U_{k+1} > i\}.$$

Logo

$$\{X_{n+1} = i+1, X_n = i, \ldots, X_{n-i} = 0\} = \{U_{k+1} > i+1\}.$$

Como $v(n) = k$ *temos* $R_k = n - i$, *porque no tempo* $n - i$ *foi onde houve a última substituição. Disto,*

$$\begin{aligned} & \mathbb{P}(X_{n+1} = i+1, X_n = i, X_{n-1} = i-1, \ldots, X_{n-i} = 0, D, v(n) = k) \\ = \; & \mathbb{P}(U_{k+1} > i+1, R_k = n-i, D). \end{aligned}$$

Já que $R_k = \sum_{i=1}^{k} U_i$ *e as* $\{U_k\}$*'s são independentes, obtemos*

$$\begin{aligned} \mathbb{P}(U_{k+1} > i+1, R_k = n-i, D) &= \mathbb{P}(U_{k+1} > i+1)\mathbb{P}(R_k = n-i, D) \\ &= \mathbb{P}(U_1 > i+1)\mathbb{P}(R_k = n-i, D), \end{aligned}$$

onde também utilizamos que U_{k+1} *independe de* D *e o fato das* $\{U_k\}'s$ *serem identicamente distribuídas. Retornando a (2.3) temos*

$$\mathbb{P}(X_{n=1} = i+1, X_n = i, B) = \mathbb{P}(U_1 > i+1)\left(\sum_{k=0}^{\infty} \mathbb{P}(R_k = n-i, D)\right). \qquad (2.4)$$

Cálculo semelhante produz

$$\mathbb{P}(X_n = i, B) = \mathbb{P}(U_1 > i)\left(\sum_{k=0}^{\infty} \mathbb{P}(R_k = n-i, D)\right). \qquad (2.5)$$

Agora de (2.4) e (2.5) concluímos que

$$\mathbb{P}(X_{n+1} = i+1 | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \frac{\mathbb{P}(U_1 > i+1)}{\mathbb{P}(U_1 > i)}.$$

Os cálculos anteriores conduzem a mesma avaliação para $\mathbb{P}(X_{n+1} = i+1 | X_n = i)$. *Completando assim a demonstração de que* X_n *é cadeia de Markov homogênea.*

2.2 Distribuição de uma Cadeia de Markov Homogênea

A variável aleatória X_0 é por convenção o estado inicial da cadeia e sua distribuição é denotada por v_0.

Definição 2.0.5 (Distribuição inicial)**.** *A distribuição inicial é a função definida por*

$$v_0(i) = \mathbb{P}(X_0 = i).$$

A distribuição inicial satisfaz as condições:

$$v_0(i) \geq 0 \quad \text{e} \quad \sum_i v_0(i) = 1,$$

para todo $i \in E$.

Definição 2.0.6 (Matriz de Transição)**.** *A matriz* $\mathbf{P} = \{p_{ij}\}_{i,j \in E}$*, onde*

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) \tag{2.6}$$

é a matriz de transição da Cadeia de Markov Homogênea.

As linhas da matriz de transição são probabilidades, ou seja, satisfazem as condições:

$$p_{ij} \geq 0, \quad \sum_{k \in E} p_{ik} = 1.$$

Exemplo 2.0.7. *Em uma cidade existem dois supermercados que denotaremos por A e B. O supermercado A encomendou uma pesquisa de mercado com referência à fidelidade dos clientes, tendo recebido o seguinte resuldado, 80% dos clientes de A continuariam comprando em A e 20% migrariam para o supermercado B. Entretanto, foi observado que somente 45% dos clientes de B continuaria comprando em B enquanto 55% passariam a comprar em A. Com relação ao número de consumidores do mercado observou-se que 60% compram em A e os restantes 45% compram em B. A direção do supermercado A deseja saber qual o percentual de consumidores passarão a comprar em A na próxima compra.*

Vamos efetuar estes cálculos. O supermercado A detem 60% dos consumidores e 80% comprarão em A, então:

1. *O percentual de consumidores que comprarão em A será* $0,60 \times 0,80 = 0,48$.

2. *O percentual de consumidores que compravam em B e comprarão em A será* $0,40 \times 0,55 = 0,22$.

3. *Total dos consumidores que comprarão em A será igual a* $(1) + (2) = 0,70$.

Portanto o supermercado A terá 70% dos consumidores. Aproveitamos para calcular o percentual de consumidores que irão comprar em B.

1. *O percentual de consumidores que comprarão em B será* $0,40 \times 0,45 = 0,18$.

2. *O percentual de consumidores que compravam em A e comprarão em B será* $0,60 \times 0,20 = 0,12$.

3. *Total dos consumidores que comprarão em B é dado por* $(1) + (2) = 0,30$.

Assim 30% dos consumidores comprarão em B o que já era esperado.

Faremos agora os mesmos cálculos utilizando uma outra abordagem. As infomações obtidas permitem construir as tabelas seguintes:

1. *Preferencia dos consumidores*

Tabela 2.1: Preferencia dos consumidores

supermercado	comprarão em A	comprarão em B
compram em A	0,80	0,20
compram em B	0,55	0,45

2. *Consumo atual*

Tabela 2.2: Consumo atual

Compram em A	Compram em B
0,60	0,40

Com os dados das tabelas 1 e 2, podemos construir as matrizes

$$\mathbf{P} = \begin{pmatrix} 0.80 & 0.20 \\ 0.55 & 0.45 \end{pmatrix}$$

e

$$\mathbf{c} = \begin{pmatrix} 0.60 & 0.40 \end{pmatrix}.$$

Multiplicando $\mathbf{c}$ *por* $\mathbf{P}$ *obtemos*

$$\mathbf{cP} = \begin{pmatrix} 0.70 & 0.30 \end{pmatrix},$$

que é o resultado obtido no primeiro cálculo. A matriz $\mathbf{P}$ *é no caso a matriz de transição.*

Teorema 2.1. *A distribuição de uma Cadeia de Markov Homogênea em tempo discreto é determinada por sua distribuição inicial e sua matriz de transição.*

Demonstração: Seja X_0 uma variável aleatória que denota o *estado inicial* de uma cadeia de Markov em tempo discreto. Tomemos

$$v_0(i) = \mathbb{P}(X_0 = i),$$

como sendo a distribuição inicial. Usando a fórmula sequencial de Bayes (Teorema 1.3), segue que

$$\begin{aligned} & \mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_k = i_k) \\ = \quad & \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1 | X_0 = i_0) \dots \mathbb{P}(X_k = i_k | X_{k-1} = i_{k-1}, \dots, X_0 = i_0). \end{aligned}$$

A expressão acima é independente de k, assim, pela propriedade da Cadeia de Markov Homogênea e a definição de matriz de transição, obtemos

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_k = i_k) = v(i_0) p_{i_0 i_1} \dots p_{i_{k-1} - i_k}.$$

$\square$

Uma generalização importante da distribuição inicial é a distribuição no tempo n. Tal conceito é usado frequentemente em nosso trabalho. Para maiores referências consultar [2].

Definição 2.1.1. *A distribuição no tempo n da cadeia é o vetor v_n, o qual é definido por*

$$v_n(i) = \mathbb{P}(X_n = i). \tag{2.7}$$

Como consequência da regra de Bayes de exclusivas e exaustivas causas (Teorema 1.2),

$$v_{n+1}(j) = \sum_{i \in E} v_n(i) p_{ij}.$$

De fato,

$$\begin{aligned} v_{n+1}(j) &= \mathbb{P}(X_{n+1} = j) \\ &= \mathbb{P}(X_{n+1} = j, \Omega) \\ &= \mathbb{P}(X_{n+1} = j, \dot{\bigcup_{i \in E}} \{X_n = i\}) \end{aligned}$$

Agora, pela definição de probabilidade,

$$\begin{aligned} v_{n+1}(j) &= \sum_{i \in E} \mathbb{P}(X_{n+1} = j, X_n = i) \\ &= \sum_{i \in E} \frac{\mathbb{P}(X_{n+1} = j | X_n = i)\mathbb{P}(X_n = i)}{\mathbb{P}(X_n = i)} \\ &= \sum_{i \in E} v_n(i) \mathbb{P}(X_{n+1} = j | X_n = i) \\ &= \sum_{i \in E} v_n(i) p_{ji} \end{aligned}$$

que é, na forma matricial, $v_{n+1}^T = v_n^T \mathbf{P}$. Da iteração desta igualdade resulta:

$$v_{n+1}^T = v_0^T \mathbf{P}^n. \tag{2.8}$$

A matriz $\mathbf{P}^n$ é chamada de matriz de transição de n-passos, devido ao termo geral que é:

$$p_{ij}(m) = \mathbb{P}(X_{n+m} = j | X_n = i). \tag{2.9}$$

De fato, usando a regra sequêncial de Bayes (Teorema 1.3) e a propriedade de Markov encontramos para o lado direito da igualdade (2.9)

$$p_{ij}(m) = \sum_{i_1, \ldots, i_m \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{m-1} j}. \tag{2.10}$$

e este é o termo geral da *m-ésima* potência de $\mathbf{P}$.

No que segue mostraremos tal afirmação. Usando a propriedade de distribuição condicional segue

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}(X_{n+m} = j, X_n = i)}{\mathbb{P}(X_n = i)}.$$

Sendo $X_{m+n-1} : \Omega \longrightarrow E$ uma variável aleatória discreta,

$$\Omega = \dot{\bigcup_{i_{m-1} \in E}} \{X_{m+n-1} = i_{m-1}\} = \dot{\bigcup_{i_{m-1} \in E}} X^{-1}_{m+n-1}(i_{m-1}).$$

Disto, podemos escrever

$$\begin{aligned} \mathbb{P}(X_{m+n} = j, X_n = i) &= \mathbb{P}(X_{n+m} = j, \Omega, X_n = i) \\ &= \mathbb{P}(X_{m+n} = j, \dot{\bigcup_{i_{m-1} \in E}} X_{m+n-1} = i_{m-1}, X_n = i). \end{aligned}$$

Segue que,

$$\begin{aligned} \mathbb{P}(X_{n+m} = j | X_n = i) &= \frac{\mathbb{P}(X_{m+n} = j, \dot{\bigcup_{i_{m-1} \in E}} X^{-1}_{m+n-1}(i_{m-1}), X_n = i)}{\mathbb{P}(X_n = i)} \\ &= \frac{\mathbb{P}\left(\dot{\bigcup_{i_{m-1} \in E}} \{X_{m+n} = j, X^{-1}_{m+n-1}(i_{m-1}), X_n = i\} \right)}{\mathbb{P}(X_n = i)} \\ &= \sum_{i_{m-1} \in E} \left(\frac{\mathbb{P}(X_{m+n} = j, X_{m+n-1} = i_{m-1}, X_n = i)}{\mathbb{P}(X_n = i)} \right), \end{aligned} \tag{2.11}$$

onde na última igualdade usamos a definição axiomática de probabilidade.

Multiplicando e dividindo (2.11) por $\mathbb{P}(X_{n+m-1} = i_{m-1}, X_n = i)$, vemos que

$$\begin{aligned} &= \sum_{i_{m-1} \in E} \left(\frac{\mathbb{P}(X_{m+n} = j, X_{m+n-1} = i_{m-1}, X_n = i)}{\mathbb{P}(X_{n+m-1} = i_{m-1}, X_n = i)} \right) . \left(\frac{\mathbb{P}(X_{n+m-1} = i_{m-1}, X_n = i)}{\mathbb{P}(X_n = i)} \right) \\ &= \sum_{i_{m-1} \in E} \left(\mathbb{P}(X_{m+n} = j | X_{m+n-1} = i_{m-1}, X_n = i) \frac{\mathbb{P}(X_{m+n-1} = i_{m-1}, X_n = i)}{\mathbb{P}(X_n = i)} \right). \end{aligned}$$

Usando a propriedade de Markov, obtemos

$$\mathbb{P}(X_{m+n} = j|X_n = i) = \sum_{i_{m-1} \in E} p_{i_{m-1}j} \frac{\mathbb{P}(X_{n+m-1} = i_{m-1}, X_n = i)}{\mathbb{P}(X_n = i)}.$$

De maneira análoga,

$$\mathbb{P}(X_{m+n} = j|X_n = i) = \sum_{i_{m-1} \in E} \sum_{i_{m-2} \in E} p_{i_{m-1}j} p_{i_{m-2}i_{m-1}} \frac{P(X_{n+m-2} = i_{m-1}, X_n = i)}{P(X_n = i)}.$$

Por iteração, pois m é finito, chegamos a

$$\mathbb{P}(X_{m+n} = j|X_n = 1) = \sum_{i_1, \ldots, i_m \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{m-1}j}.$$

Assim mostramos a afirmação feita anteriormente.

Outro fato importante é que a propriedade de Markov (2.1) pode ser ampliada para

$$\begin{aligned} & \mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+k} = j_k | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ = \ & \mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+k} = j_k | X_n = i) \end{aligned}$$

Mostraremos tal propriedade para o caso

$$\begin{aligned} & \mathbb{P}(X_{n+2} = j_2, X_{n+1} = j_1 | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ = \ & \mathbb{P}(X_{n+2} = j_2, X_{n+1} = j_1 | X_n = i). \end{aligned}$$

O caso geral se dá por indução. Denotemos por

$A_4 = \{X_{n+2} = j_2\}, A_3 = \ \{X_{n+1} = j_1\}, \ A_2 = \{X_n = i\}$ e $A_1 = \{X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\}$

Por definição,

$$\mathbb{P}(A_4, A_3 | A_2, A_1) = \frac{\mathbb{P}(A_4, A_3, A_2, A_1)}{\mathbb{P}(A_2, A_1)}$$

Pela fórmula sequencial de Bayes (Teorema 1.3),

$$\begin{aligned} \mathbb{P}(A_4, A_3 | A_2, A_1) &= \frac{\mathbb{P}(A_1)\mathbb{P}(A_2|A_1).\mathbb{P}(A_3|A_2, A_1)\mathbb{P}(A_4|A_3, A_2, A_1)}{\mathbb{P}(A_2, A_1)} \\ &= \mathbb{P}(A_3|A_2, A_1)\mathbb{P}(A_4|A_3, A_2, A_1). \end{aligned}$$

Utilizando a propriedade de Markov obtemos

$$\mathbb{P}(A_4, A_3|A_2, A_1) = \mathbb{P}(A_3|A_2)\mathbb{P}(A_4|A_3) = \frac{\mathbb{P}(A_2)\mathbb{P}(A_3|A_2)\mathbb{P}(A_4|A_3)}{\mathbb{P}(A_2)}$$

Como os eventos A_4 e A_2 são condicionalmente independentes dado A_3 (ver Teorema 1.4) segue que

$$\mathbb{P}(A_4, A_3|A_2, A_1) = \frac{\mathbb{P}(A_2)\mathbb{P}(A_3|A_2)\mathbb{P}(A_4|A_3, A_2)}{\mathbb{P}(A_2)}.$$

Novamente pela regra sequencial de Bayes (Teorema 1.3), concluímos que

$$\mathbb{P}(A_4, A_3|A_2, A_1) = \frac{\mathbb{P}(A_4, A_3, A_2)}{\mathbb{P}(A_2)} = \mathbb{P}(A_4, A_3|A_2).$$

Assim,

$$\mathbb{P}(X_{n+2} = j_2, X_{n+1} = j_1|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+2} = j_2, X_{n+1} = j_1|X_n = i)$$

para todos $i_0, \ldots, i_{n-1}, i, j_1, j_2$ tal que ambos os lados da igualdade estão definidos. Escrevendo

$$A = \{X_{n+1} = j_1, \ldots, X_{n+k} = j_k\} \quad e \quad B = \{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}$$

vemos que $\mathbb{P}(A|X_n = i, B) = \mathbb{P}(A|X_n = i)$, o que por sua vez produz

$$\mathbb{P}(A \cap B|X_n = i) = P(A|X_n = i)P(B|X_n = i).$$

De fato,

$$\mathbb{P}(A \cap B|X_n = i) = \frac{\mathbb{P}(A \cap B, X_n = i)}{\mathbb{P}(X_n = i)}$$

Pela regra sequencial de Bayes (Teorema 1.3),

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

Portanto

$$\mathbb{P}(A \cap B|X_n = i) = \mathbb{P}(B|X_n = i)\mathbb{P}(A|X_n = i).$$

Em palavras, o futuro no tempo n e o passado no tempo n são condicionalmente independentes do estado presente dado $X_n = i$. Isto mostra em particular que a propriedade de Markov é independente da direção do tempo.

Exemplo 2.1.1 (Cadeia de Markov de dois Estados). *Em uma estação meteorológica são realizadas apenas duas observações, se chove ou não chove. Supondo que esteja chovendo no dia n, existe uma probabilidade p de que esteja chovendo no dia seguinte e, ainda, existe uma probabilidade q de que não esteja chovendo no dia n e chova no dia seguinte. Seja $v_0(0)$ a probabilidade de que esteja chovendo no 0-ésimo dia. Vamos convencionar e representar por 0 o estado em que está chovendo e por 1 o estado em que não chove. Designemos por X_n a variável aleatória que representa o estado de estar chovendo ou não, no tempo n. Como calcular as probabilidades de estar ou não chovendo?*

A probabilidade de transição de um estado para outro é dada por

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

Como consequência temos:
i) probabilidade de chover ($i = 0$) no dia n e não chover ($j = 1$) no dia $n + 1$ (dia seguinte) será dada por

$$p_{01} = \mathbb{P}(X_{n+1} = 1|X_n = 0) = p;$$

ii) probabilidade de não chover ($i = 1$) no dia n e de chover ($j = 0$) no dia $n + 1$ (dia seguinte) será dada por

$$p_{10} = \mathbb{P}(X_{n+1} = 0|X_n = 1) = q;$$

iii) a probabilidade de estar chovendo e continuar a chover é dada por

$$p_{00} = \mathbb{P}(X_{n+1} = 0|X_n = 0) = 1 - p;$$

iv) a probabilidade não estar chovendo e sem chuva no dia seguinte, dada por

$$p_{11} = \mathbb{P}(X_{n+1} = 1|X_n = 1) = 1 - q;$$

v) a probabilidade de no tempo 0 não estar chovendo é

$$v_0(1) = \mathbb{P}(X_0 = 1) = 1 - v_0(0);$$

vi) e por último, a probabilidade de está chovendo no tempo 0 *é igual a*

$$v_0(0) = \mathbb{P}(X_0 = 0).$$

Tendo em vista o exposto, podemos agora calcular as probabilidades de estar chovendo ou não. A probabilidade de estar chovendo no dia seguinte a um dia n *qualquer é dada por:*

$$\begin{aligned}
\mathbb{P}(X_{n+1} = 0) &= \mathbb{P}(X_{n+1} = 0 \cap X_n = 0) + \mathbb{P}(X_{n+1} = 0 \cap X_n = 1) \\
&= \mathbb{P}(X_n = 0)\mathbb{P}(X_{n+1} = 0 | X_n = 0) + \mathbb{P}(X_n = 1)P(X_{n+1} = 0 | X_n = 1) \\
&= \mathbb{P}(X_n = 0)(1 - p) + \mathbb{P}(X_n = 1)q \\
&= (1 - p - q)\mathbb{P}(X_n = 0) + q.
\end{aligned}$$

onde usamos na última igualdade que $\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0)$. *Como* $v_0(0) = \mathbb{P}(X_0 = 0)$, *segue que*

$$\mathbb{P}(X_1 = 0) = (1 - p - q)v_0(0) + q$$

e

$$\begin{aligned}
\mathbb{P}(X_2 = 0) &= (1 - p - q)\mathbb{P}(X_1 = 0) + q \\
&= (1 - p - q)^2 v_0(0) + q(1 + (1 - p - q)).
\end{aligned}$$

Por iteração, vemos que

$$\mathbb{P}(X_n = 0) = (1 - p - q)^n v_0(0) + q \sum_{j=0}^{n-1} (1 - p - q)^j.$$

Como

$$\sum_{j=0}^{n-1} (1 - p - q)^j = \frac{1 - (1 - p - q)^n}{p + q}$$

temos

$$\mathbb{P}(X_n = 0) = \frac{q}{p + q} + (1 - p - q)^n \left(v_0(0) - \frac{q}{p + q} \right)$$

e

$$\mathbb{P}(X_n = 1) = \frac{p}{p+q} + (1-p-q)^n \left(v_0(1) - \frac{p}{p+q} \right).$$

Sendo $0 < p < 1$, *e* $0 < q < 1$, *adicionando-se (-1) aos termos da desigualdade, obtemos -1<-(1-p-q)<1, que é equivalente a*

$$|1-p-q| < 1.$$

Neste caso quando n tende ao ∞ *temos que*

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0) = \frac{q}{p+q}$$

e

$$\lim_{n\to\infty} \mathbb{P}(X_n = 1) = \frac{p}{p+q}.$$

2.3 Cadeias de Markov por recorrência

Nesta seção apresentamos uma técnica (Ver [2]) para descrever algumas cadeias de Markov usando equações de recorrência. O teorema que segue descreve tal técnica.

Teorema 2.2 (Cadeia de Markov Homogênea dirigida por Ruído Branco)**.** *Seja* $\{Z_n\}_{n\geq 1}$ *uma sequência independente e identicamente distribuída de variáveis aleatórias com valores em um espaço arbitrário* F. *Seja* E *um espaço contável e,* $f : E \times F \longrightarrow E$ *uma função qualquer. Seja* X_0 *uma variável aleatória tomando valores em* E, *independente de* $\{Z_n\}_{n\geq 1}$. *A equação recorrente*

$$X_{n+1} = f(X_n, Z_{n+1}) \tag{2.12}$$

define uma Cadeia de Markov Homogênea.

Demonstração: Por hipótese as variáveis aleatórias $X_0, Z_1, \ldots, Z_n$ são independentes de Z_{n+1}. Por iteração da equação recorrente $X_{n+1} = f(X_n, Z_{n+1})$, obtemos, para

$n = 0, \quad X_1 = f(X_0, Z_1)$
$n = 1, \quad X_2 = f(X_1, Z_2) = f(f(X_0, Z_1), Z_2)$
$n = 2, \quad X_3 = f(X_2, Z_3) = f(f(f(X_0, Z_1), Z_2), Z_3)$
Continuando, vemos que existe uma função g_n, com $n \geq 1$ tal que

$$X_n = g_n(X_0, Z_1, Z_2, ..., Z_n).$$

Por hipótese

$$X_{n+1} = f(X_n, Z_{n+1}). \tag{2.13}$$

Substituindo na equação acima o valor de X_n por i, isto é, $X_n = i$, resulta que

$$X_{n+1} = f(i, Z_{n+1}). \tag{2.14}$$

Afirmamos que $f(i, Z_{n+1})$ é independente de $g_n(X_0, Z_1, \ldots, Z_{n-1})$. De fato,

$$\begin{aligned}
& \mathbb{P}(f(i, Z_{n+1}) = j, g_n(X_0, Z_1, \ldots, Z_{n-1}) = l) \\
= \; & \mathbb{P}(f_i(Z_{n+1}) = j, g_n(X_0, Z_1, \ldots, Z_{n-1}) = l) \\
= \; & \mathbb{P}(Z_{n+1} \in f_i^{-1}(j), (X_0, Z_1, \ldots, Z_{n-1}) \in g_n^{-1}(l)) \\
= \; & \mathbb{P}(Z_{n+1} \in f_i^{-1}(j))\mathbb{P}((X_0, Z_1, \ldots, Z_{n-1}) \in g_n^{-1}(l)) \\
= \; & \mathbb{P}(f(i, Z_{n+1}) = j)\mathbb{P}(g_n(X_0, Z_1, \ldots, Z_{n-1}) = l),
\end{aligned}$$

onde na igualdade 4 usamos a independência de Z_{n+1} e $(X_0, Z_1, \ldots, Z_{n-1})$. Portanto, $f(i, Z_{n+1})$ e $g_n(X_0, Z_1, \ldots, Z_{n-1})$ são independentes.

Substituindo o valor de X_{n+1} encontrado no lado esquerdo da propriedade de Markov, obtemos a igualdade

$$\begin{aligned}
& \mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\
= \; & \mathbb{P}(f(i, Z_{n+1}) = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\
= \; & \mathbb{P}(f(i, Z_{n+1}) = j)
\end{aligned}$$

porque o evento $\{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i\}$ é independente de Z_{n+1}, como provamos acima.

Mostraremos agora que $\mathbb{P}(X_{n+1} = j|X_n = i) = \mathbb{P}(f(i, Z_{n+1}) = j)$. Vimos anteriormente que $X_n = g_n(X_0, Z_1, \ldots, Z_n)$ e, usando o lado direito da propriedade de Markov (2.1), com as devidas substituições como efetuamos anteriormente, obtemos

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j|X_n = i) &= \mathbb{P}(f(i, Z_{n+1}) = j|X_n = i,) \\
&= \mathbb{P}(f(i, Z_{n+1}) = j|g_n(X_0, Z_1, Z_2, ..., Z_n) = i) \\
&= \mathbb{P}(f(i, Z_{n+1}) = j).
\end{aligned}$$

Portanto X_n é uma Cadeia de Markov, a qual é homogênea, já que $\mathbb{P}(f(i, Z_{n+1}) = j)$ é independente de n. □

A probabilidade de transição p_{ij} da CMH do Teorema acima é dada por

$$\mathbb{P}(X_{n+1} = j|X_n = i) = \mathbb{P}(f(i, Z_{n+1}) = j).$$

Assim, em um tempo zero

$$p_{ij} = \mathbb{P}(f(i, Z_1) = j). \tag{2.15}$$

A seguir mostraremos exemplos canônicos que ilustram esta teoria.

Exemplo 2.2.1 (Passeio aleatório). *Seja uma variável aleatória com valores em $\mathbb{Z}$. Consideremos a sequência $\{Z_n\}_{n\geq 1}$ de variáveis aleatórias, independentes de X_0 e tomando valores no conjunto* +1 *ou* -1, *com distribuição de probabilidade igual a:*

$$\mathbb{P}(Z_n = +1) = p,$$

onde $p \in (0, 1)$. O processo descrito por

$$X_{n+1} = X_n + Z_{n+1}$$

é uma CMH, denominada de passeio aleatório sobre $\mathbb{Z}$. De fato, segue do Teorema 2.2 e da observação que

$$X_{n+1} = f(X_n, Z_{n+1})$$

com

$$f : \mathbb{Z} \times \{+1, -1\} \longrightarrow \mathbb{Z}$$

definida por

$$f(X_n, Z_{n+1}) = X_n + Z_{n+1}.$$

Exemplo 2.2.2 (Reparo de máquinas)**.** *Durante o dia n, Z_{n+1} máquinas quebram e vão para a loja de reparos no dia $n+1$. Cada dia uma máquina entre aquelas à espera de serviço é reparada. Portanto, denotando por X_n o número de máquinas na loja no dia n,*

$$X_{n+1} = (X_n - 1)^+ + Z_{n+1} \tag{2.16}$$

onde a^+=max$(a, 0)$. Particularmente, se $\{Z_n\}_{n\geq 1}$ é uma sequência independente e identicamente distribuída, independente do estado inicial X_0, então $\{X_n\}_{n\geq 0}$ é uma Cadeia Homogênea de Markov. Em termos de distribuição de probabilidade:

$$\mathbb{P}(Z_1 = k) = a_k, k \geq 0. \tag{2.17}$$

Se existem X_n máquinas para reparo no dia n. No dia seguinte uma máquina é reparada, por isso o número de máquinas na oficina será $X_n - 1$, no entanto no tempo $n+1$ quebram Z_{n+1} máquinas, por esta razão, o número de máquinas para reparo no tempo $n+1$, será $X_{n+1} = (X_n - 1)^+ + Z_{n+1}$. O espaço de estado é $E = \mathbb{N}$ e a função é definida como

$$f : E \times \mathbb{N} \longrightarrow E$$
$$f(n, m) = (n-1)^+ + m. \tag{2.18}$$

Pelo Teorema 2.2, (2.16) e (2.18) concluímos que $\{X_n\}$ é uma cadeia de Markov homogênea. Sua matriz de transição é

$$\mathbf{P} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}. \tag{2.19}$$

De fato, por (2.15) e (2.18) vemos que

$$\begin{aligned} p_{ij} &= \mathbb{P}(f(i, Z_1) = j) \\ f(i, z_1) &= (i-1)^+ + Z_1. \end{aligned}$$

Agora, de (2.17) e (2.18) obtemos

$$\begin{aligned} p_{ij} &= \mathbb{P}((i-1)^+ + Z_1) = j) \\ &= \mathbb{P}(Z_1 = j - (i-1)^+) \\ &= a^+_{j-(i-1)}. \end{aligned}$$

2.4 Probabilidade e tempo médio de absorção

Existe uma técnica importante para avaliar a probabilidades e o tempo médio de absorção em um conjunto fechado de uma CMH. Esta técnica é chamada de *Análise do primeiro passo* [2]. Os exemplos a seguir ilustram esta técnica.

Exemplo 2.2.3. *Neste exemplo apresentaremos duas situações que serão utilizadas para ilustrar o exemplo que segue. Consideremos uma primeira situação. Um jogador dispõe de duas unidades monetárias e entra em um jogo com a seguinte regra: quando ganha recebe uma unidade com probabilidade p e quando perde paga uma unidade com probabilidade q. Vamos então calcular a probabilidade de que o jogador venha a vencer, estabelecendo que: quando atingir 4 unidades monetárias o jogo termina, ou quando estiver sem recursos. Este é um típico exemplo de passeio aleatório, exemplo 2.2.1, com espaço de estados $E = \{0, 1, 2, 3, 4\}$.*

Denotemos por $u(i)$, $i = 0, 1, 2, 3, 4$, as probabilidades do jogador ganhar quando ele possui \$ i unidades monetárias. As condições limitantes são $u(4) = 1$ e $u(0) = 0$. Como o jogador parte com \$ 2, devemos calcular $u(2)$. Mas, para encontrarmos $u(2)$ precisamos saber quais são as probabilidades de ele vencer quando ele possuir \$ 1 ou \$ 3, pois $u(2)$ depende de $u(1)$ e $u(3)$. De fato,

$$u(2) = pu(3) + qu(1).$$

Da mesma maneira, temos que $u(3) = pu(4) + qu(2)$ e $u(1) = pu(2) + qu(0)$. Um cálculo direto dá

$$u(2) = \frac{p^2}{1 - 2pq}.$$

Fazendo $p = q = \dfrac{1}{2}$ *obtemos*

$$u(2) = \frac{\frac{1}{4}}{1 - 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

Para a segunda situação as regras são as mesmas da anterior, diferindo no seguinte aspecto, o jogador agora dispõe de \$ 3 e vence quando atinge a quantia de \$ 5. As condições limitantes são agora $u(5) = 1$ *e* $u(0) = 0$. *Calculemos* $u(3)$. *Inicialmente temos que*

$$u(3) \quad = \quad pu(4) + qu(2).$$

Vamos calcular em separado $u(2)$. *Do mesmo modo,* $u(2)$ *depende de* $u(1)$ *e* $u(3)$, *isto é,*

$$u(2) = pu(3) + qu(1).$$

Também, usando a condição limtante inferior temos que

$$u(1) = pu(2) + qu(0) = pu(2).$$

Usando as igualdades acima é facil ver que

$$u(2) = \frac{p}{1 - qp} u(3).$$

Sendo $u(4) = pu(5) + qu(3)$, *por cálculos semelhantes aos acima concluímos que*

$$u(3) = \frac{(1 - qp)p^2}{(1 - qp)^2 - qp}.$$

Em particular, se $p = q = \dfrac{1}{2}$, *então* $u(3)$ *será igual a*

$$u(3) \quad = \quad \frac{\left(1 - \frac{1}{4}\right)\frac{1}{4}}{\left(1 - \frac{1}{4}\right)^2 - \frac{1}{4}} = \frac{\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)}{\frac{9}{16} - \frac{1}{4}} = \frac{\frac{3}{16}}{\frac{9-4}{16}} = \frac{3}{5}.$$

O exemplo acima é um caso particular de uma situação geral que trataremos no próximo exemplo.

Exemplo 2.2.4 (A Ruína do jogador). *Dois jogadores A e B jogam Cara ou Coroa, onde a probabilidade de ocorrer Cara é $p \in (0,1)$, e os resultados sucessivos são uma sequência independente e identicamente distribuída (i.i.d). Seja X_n o ganho em unidades monetárias do jogador A no tempo n, $X_{n+1} = X_n + Z_{n+1}$ onde $Z_{n+1} = +1$ ou respectivamente -1, com probabilidade p e $q = 1-p$, respectivamente. É claro que $\{Z_n\}_{n\geq 1}$ é i.i.d. De modo mais claro, A aposta \$ 1 no resultado Cara e B aposta* 1 *no resultado Coroa em cada lançamento da moeda. As quantias que cada jogador dispõem inicialmente são a e b, respectivamente para os jogadores A e B. O jogo termina quando um dos jogadores perde toda a quantia de que dispõe. O processo $\{X_n\}_{n\geq 1}$ é um caminho aleatório, como foi descrito no exemplo 2.1.1, com excessão de que é restrito ao espaço $E = \{0, \ldots, a, a+1, \ldots, a+b = c\}$. O jogo dura um tempo T, o primeiro tempo n em que $X_n = 0$ ou c, e a probabilidade de A vencer é*

$$u(a) = P(X_T = c|X_0 = a).$$

Observemos que T é o tempo de absorção, ou seja, uma vez que o processo atinge 0 *ou c, em T ele não se modificará. Em vez de calcular $u(a)$ isoladamente, a Análise do Primeiro Passo calcula:*

$$u(i) = \mathbb{P}(X_T = c|X_0 = i)$$

para todos os estados $i \in [0,c]$, e para isto, primeiro é gerado uma equação recorrente para os $u(i)'s$ pela quebra do evento "A vence" de acordo com o que pode acontecer depois do primeiro passo, ou seja, o primeiro lançamento da moeda. Se $X_0 = i \in [1, c-1]$ então $X_1 = i+1$ com probabilidade p ou $X_1 = i-1$ com probabilidade q. A probabilidade do fracasso de B começa a partir dos recursos iniciais de que A dispõe: se for $i+1$ é $u(i+1)$ ou no caso de $i-1$ será $u(i-1)$. Portanto para $i \in [1, c-1]$

$$u(i) = pu(i+1) + qu(i-1) \tag{2.19}$$

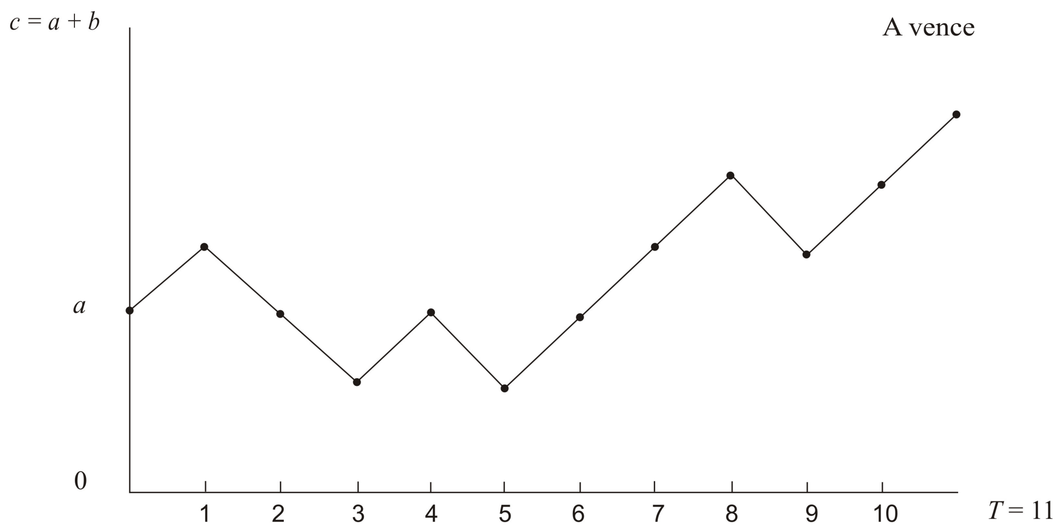

Figura 2.2: A ruína do jogador

tendo como limite a condição:

$$u(0) = 0, \quad u(c) = 1.$$

Demonstraremos a seguir a equação (2.19). Seja $Y = (Y_n)_{n\geq 0}$ *a cadeia de Markov homogênea obtida pelo retardo de* $X = (X_n)_{n\geq 0}$ *por uma unidade de tempo* $Y_n = X_{n-1}$.

Definimos

$$u(i) = \mathbb{P}(X_T = c|X_0 = i).$$

Podemos reescrever $u(i)$ *como*

$$u(i) = \mathbb{P}(X \text{ \textit{é absorvido por} } c|X_0 = i)$$

Pela definição de probabilidade condicional,

$$u(i) = \frac{\mathbb{P}(X \text{ \textit{é absorvido por} } 0, X_0 = i)}{\mathbb{P}(X_0 = i)}. \tag{2.20}$$

O númerador desta igualdade pode ser escrita como

$$\begin{aligned}
& \mathbb{P}(X \ \textit{é absorvido por } c, X_0 = i) \\
= & \ \mathbb{P}(X \ \textit{é absorvido por } c, \Omega, X_0 = i) \\
= & \ \mathbb{P}(X \ \textit{é absorvido por } c, \{X_1 = i+1\} \cup \{X_i = i-1\}, X_0 = i) \\
= & \ \mathbb{P}(X \ \textit{é absorvido por } c, X_1 = i+1, X_0 = i) \\
+ & \ \mathbb{P}(X \ \textit{é absorvido por } c, X_1 = i-1, X_0 = i)
\end{aligned}$$

Aqui observamos que as seguintes probabilidades são equivalentes,

$$\mathbb{P}(X \quad \textit{é absorvido por } c, X_1 = i \pm 1, X_0 = i)$$

e

$$\mathbb{P}(Y \quad \textit{é absorvido por } c, X_1 = i \pm 1, X_0 = i).$$

Agora, já que $(Y_n)_{n \geq 0}$ *e* X_0 *são independentes dado* X_1*, temos que* $\mathbb{P}(Y \quad \textit{é absorvido por } c, X_1 = i \pm 1, X_0 = i)$ *é igual a*

$$\mathbb{P}(X_0 = i, X_1 = i \pm 1)\mathbb{P}(Y \quad \textit{é absorvido por } c | Y_0 = i \pm 1),$$

onde usamos o fato de $Y_0 = X_1$*. Observando que as duas cadeias tem a mesma matriz de transição, e quando elas possuem o mesmo estado inicial, elas tem a mesma distribuição. Disto,*

$$\mathbb{P}(Y \quad \textit{é absorvido por } c | Y_0 = i \pm 1) = u(i \pm 1).$$

Usando as conclusões acima em (2.20) obtemos

$$\begin{aligned}
u(i) &= u(i+1)\frac{\mathbb{P}(X_0 = i, X_1 = i+1)}{\mathbb{P}(X_0 = i)} \\
&+ u(1-1)\frac{\mathbb{P}(X_0 = i, X_1 = i-1)}{\mathbb{P}(X_0 = i)} \\
&= pu(i+1) + qu(i-1).
\end{aligned}$$

A equação característica associada com esta equação linear recorrente é

$$pr^2 - r + q = 0. \tag{2.21}$$

Encontramos esta equação do seguinte modo. Primeiro, dividimos a equação (2.19) por $u(i-1)$, *a saber,*

$$\frac{u(i)}{u(i-1)} = p\frac{u(i+1)}{u(i-1)} + q.$$

Assim, é suficiente mostrar que $r = \dfrac{u(i)}{u(i-1)}$ *e* $r^2 = \dfrac{u(i+1)}{u(i-1)}$. *Iniciemos o cálculo para* $\dfrac{u(i)}{u(i-1)}$. *Calculando,*

$$\begin{aligned}\frac{u(i)}{u(i-1)} &= \frac{\mathbb{P}(X_T = c|X_0 = i)}{\mathbb{P}(X_T = c|X_0 = i-1)} = \frac{\dfrac{\mathbb{P}(X_T = c, X_0 = i)}{\mathbb{P}(X_0 = i)}}{\dfrac{\mathbb{P}(X_T = c, X_0 = i-1)}{\mathbb{P}(X_0 = i-1)}} \\ &= \frac{\mathbb{P}(X_T = c, X_0 = i)}{\mathbb{P}(X_T = c, X_0 = i-1)}\frac{\mathbb{P}(X_0 = i-1)}{\mathbb{P}(X_0 = i)}.\end{aligned}$$

No caso de $\dfrac{u(i+1)}{u(i-1)}$ *obtemos*

$$\begin{aligned}\frac{u(i+1)}{u(i-1)} &= \frac{\mathbb{P}(X_T = c|X_0 = i+1)}{\mathbb{P}(X_T = c|X_0 = i-1)} = \frac{\dfrac{\mathbb{P}(X_T = c, X_0 = i+1)}{\mathbb{P}(X_0 = i+1)}}{\dfrac{\mathbb{P}(X_T = c, X_0 = i-1)}{\mathbb{P}(X_0 = i-1)}} \\ &= \frac{\mathbb{P}(X_T = c, X_0 = i+1)}{\mathbb{P}(X_T = c, X_0 = i-1)}.\frac{\mathbb{P}(X_0 = i-1)}{\mathbb{P}(X_0 = i)}.\end{aligned}$$

Vamos agora multiplicar o último termo da igualdade por

$$\frac{\mathbb{P}(X_T = c|X_0 = i)}{\mathbb{P}(X_T = c|X_0 = i)}.\frac{\mathbb{P}(X_0 = i)}{\mathbb{P}(X_0 = i)}$$

para obter

$$\frac{\mathbb{P}(X_T = c|X_0 = i+1)}{\mathbb{P}(X_T = c|X_0 = i-1)}.\frac{\mathbb{P}(X_0 = i-1)}{\mathbb{P}(X_0 = i+1)}.\frac{\mathbb{P}(X_T = c|X_0 = i)}{\mathbb{P}(X_T = c|X_0 = i)}.\frac{\mathbb{P}(X_0 = i)}{\mathbb{P}(X_0 = i)}.$$

Reordenando os termos podemos reescrever $\dfrac{u(i+1)}{u(i-1)}$ *como*

$$\left(\frac{\mathbb{P}(X_T = c|X_0 = i+1)}{\mathbb{P}(X_T = c|X_0 = i)}.\frac{\mathbb{P}(X_0 = i)}{\mathbb{P}(X_0 = i+1)}\right)\left(\frac{\mathbb{P}(X_T = c|X_0 = i)}{\mathbb{P}(X_T = c|X_0 = i-1)}.\frac{\mathbb{P}(X_0 = i-1)}{\mathbb{P}(X_0 = i)}\right).$$

Os dois termos do produto acima são iguais a

$$\frac{u(i+1)}{u(i)} \quad e \quad \frac{u(i)}{u(i-1)},$$

os quais pela equação de recorrência (2.19) são iguais a r. *Disto, segue a equação (2.21).*

A equação (2.21) têm duas raízes distintas $r_1 = 1$ *e* $r_2 = \frac{p}{q}$, *se* $p \neq q$ *e raíz dupla,* $r_1 = 1$ *se* $p = q = \frac{1}{2}$. *Portanto, a solução geral será :*

$$u(i) = \lambda r_1^i + \mu r_2^i = \lambda + \mu(\frac{q}{p})^i,$$

quando $p \neq q$, *e*

$$u(i) = \lambda r_1^i + \mu i r_1^i = \lambda + \mu i,$$

quando $p = q = \frac{1}{2}$. *Levando em conta as condições limites se pode determinar os valores de* λ *e* μ. *O resultado será para* $p \neq q$,

$$u(i) = \frac{1 - (\frac{q}{p})^i}{1 - (\frac{q}{p})^c}, \tag{2.22}$$

e para $p = q = \frac{1}{2}$

$$u(i) = \frac{i}{c}. \tag{2.23}$$

No caso em que $p = q = \frac{1}{2}$, *a probabilidade* $v(i)$ *em que* B *vence, quando dispõem de recursos* $c - i$ *é obtida por substituição de* i *por* $c - i$ *na expressão (2.26).*

$$v(i) = \frac{c-i}{c} = 1 - \frac{i}{c}.$$

Podemos comprovar que $u(i) + v(i) = 1$. *O que significa que a probabilidade do jogo não terminar é nula.*

Exemplo 2.2.5 (O lobo e o cordeiro)**.** *Um cordeiro entra em um labirinto. Se no tempo* n *ele está em um compartimento com* k *compartimentos adjacentes, estará no tempo* $n+1$ *em um dos* k *compartimentos adjacentes, escolhido um ao acaso, cada um com probabilidade* $\frac{1}{k}$. *Um lobo faminto permanece o tempo todo em um determinado* *compartimento, e uma porção de ração está disponível para o cordeiro em outro compartimento (ver figura 2.3). O lobo não está desatento, se o cordeiro entra no compartimento onde está o o lobo, o lobo devora o cordeiro. Qual a probabilidade do cordeiro comer a ração, quando ele parte do compartimento 1 e o lobo e a ração estão respectivamente nos compartimentos 4 e 6?*

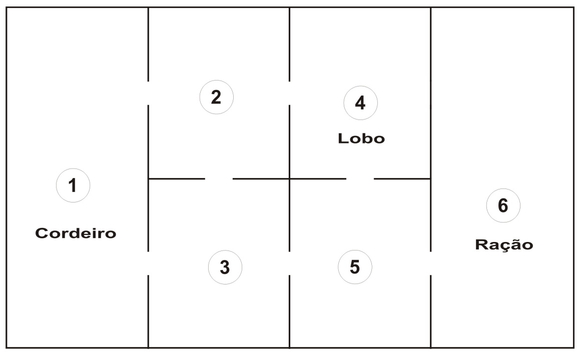

Figura 2.3: O lobo e o cordeiro

Aplicando a analise do primeiro passo, seja $u(i)$ a probabilidade que tem o cordeiro inicialmente no quarto i conseguir a ração sem ser devorada pelo lobo. As condições limites são

$$u(4) = 0 \; e \, u(6) = 1.$$

Se a ovelha está no compartimento 1, ela pode se mover para os compartimento 2 ou 3 com probabilidade de $\frac{1}{2}$ e a probabilidade de comer a ração será, então, $u(2)$ ou $u(3)$. Em consequência,

$$u(1) = \frac{1}{2}u(2) + \frac{1}{2}u(3);$$

Igualmente,

$$u(2) = \frac{1}{3}u(1) + \frac{1}{3}u(3) + \frac{1}{3}u(4);$$

$$u(3) = \frac{1}{3}u(1) + \frac{1}{3}u(2) + \frac{1}{3}u(5);$$

$$u(4) = 0;$$

$$u(5) = \frac{1}{3}u(3) + \frac{1}{3}u(4) + \frac{1}{3}u(6);$$

$$u(6) = 1.$$

Nosso objetivo é determinar o valor de u(1). Temos definidos os valores de $u(4) = 0$ *e* $u(6) = 1$. *Vamos colocar os valores de* $u(2)$ *e* $u(5)$ *em função de* $u(1)$ *e* $u(3)$. *Como*

$$u(2) = \frac{1}{3}u(1) + \frac{1}{3}u(3) + \frac{1}{3}u(4)$$

e $u(4) = 0$ *temos*

$$u(2) = \frac{1}{3}u(1) + \frac{1}{3}u(3).$$

Calculando $u(5)$. *Os valores de* $u(4)$ *e* $u(6)$ *são respectivamente* 0 *e* 1, *logo*

$$\begin{aligned} u(5) &= \frac{1}{3}u(3) + \frac{1}{3}(0) + \frac{1}{3}(1) \\ u(5) &= \frac{1}{3}u(3) + \frac{1}{3}. \end{aligned}$$

Calculemos $u(3)$. *Sabemos que*

$$u(3) = \frac{1}{3}u(1) + \frac{1}{3}u(2) + \frac{1}{3}u(5).$$

Vamos substituir na expressão acima os valores encontrados de $u(2)$ *e* $u(5)$. *Então,*

$$u(3) = \frac{1}{3}u(1) + \frac{1}{3}\left(\frac{1}{3}u(1) + \frac{1}{3}u(3)\right) + \frac{1}{3}\left(\frac{1}{3}u(3) + \frac{1}{3}\right).$$

Seguindo,

$$u(3) = \frac{1}{3}u(1) + \left(\frac{1}{9}u(1) + \frac{1}{9}u(3)\right) + \left(\frac{1}{9}u(3) + \frac{1}{9}\right)$$

$$u(3) = \frac{4}{9}u(1) + \frac{2}{9}u(3) + \frac{1}{9}$$

$$u(3) - \frac{2}{9}u(3) = \frac{4}{9}u(1) + \frac{1}{9}$$

$$\frac{7}{9}u(3) = \frac{4}{9}u(1) + \frac{1}{9}.$$

De onde,

$$u(3) = \frac{\frac{4}{9}u(1) + \frac{1}{9}}{\frac{7}{9}} = \left(\frac{4}{9}u(1) + \frac{1}{9}\right)\frac{9}{7} = \frac{4}{7}u(1) + \frac{1}{7}$$

Na continuação vamos colocar $u(2)$ em função de $u(1)$. Substituindo o valor encontrado de $u(3)$ na expressão $u(2)$ e efetuando os cálculos devidos temos

$$\begin{aligned} u(2) &= \frac{1}{3}u(1) + \frac{1}{3}\left(\frac{4}{7}u(1) + \frac{1}{7}\right) \\ &= \frac{1}{3}u(1) + \frac{4}{21}u(1) + \frac{1}{21} \\ &= \frac{11}{21}u(1) + \frac{1}{21}. \end{aligned}$$

Substituindo os valores de $u(2)$ e $u(3)$ em $u(1)$ vamos encontrar

$$\begin{aligned} u(1) &= \frac{1}{2}u(2) + \frac{1}{2}u(3) \\ &= \frac{1}{2}\left(\frac{11}{21}u(1) + \frac{1}{21}\right) + \frac{1}{2}\left(\frac{4}{7}u(1) + \frac{1}{7}\right) \\ &= \frac{11}{42}u(1) + \frac{1}{42} + \frac{4}{14}u(1) + \frac{1}{14} \\ &= \frac{23}{42}u(1) + \frac{4}{42}. \end{aligned}$$

De onde

$$\begin{aligned} u(1) - \frac{23}{42}u(1) &= \frac{4}{42} \\ \frac{19}{42}u(1) &= \frac{4}{42}. \end{aligned}$$

Por fim

$$u(1) = \frac{\frac{4}{42}}{\frac{19}{42}} = \left(\frac{4}{42}\right)\left(\frac{42}{19}\right) = \frac{4}{19}.$$

Portanto a probabilidade do cordeiro comer a ração quando saindo do compartimento 1 é de $\frac{4}{19}$.

A seguir apresentamos algumas situações quando modificamos a posição do cordeiro, lobo e ração com suas respectivas probabilidades

Tabela 2.3: Probabilidades em diferentes situações

Situação	Codeiro	Lobo	Ração	$u(i)$
S_1	1	2	6	$\frac{2}{21}$
S_2	1	3	6	$\frac{1}{19}$
S_3	2	1	6	$\frac{5}{24}$
S_4	5	1	6	$\frac{13}{24}$
S_5	5	2	6	$\frac{10}{21}$
S_6	4	2	6	$\frac{5}{19}$

Calcularemos, agora, o tempo médio antes da absorção do exemplo 2.3.1, utilizando a técnica da análise do primeiro passo.

Exemplo 2.2.6 (A ruína do jogador)**.** *A duração média* $m(i) = E[T|X_0 = i]$ *do jogo quando os recursos no início de que dispõe o jogador A é i satisfaz a equação recorrente:*

$$m(i) = 1 + pm(i+1) + qm(i-1) \tag{2.24}$$

para $i \in [1, c-1]$*. Basta olhar a equação recorrente (2.19). De fato a moeda será lançada, no mínimo uma vez, com probabilidade* p *ou* q *e os recursos do jogador A será* $i+1$ *ou* $i-1$*, respectivamente. Portanto,* $m(1+i)$ *ou* $m(i-1)$ *lançamentos a mais serão necessários em média antes que um dos jogadores fique sem nenhum recurso. As condições limitantes são:*

$$m(0) = 0, \quad m(c) = 0 \tag{2.25}$$

Para resolver a equação (2.24) com as condições limitantes (2.25), escreve-se (2.24) na forma

$$-1 = p(m(i+1) - m(i)) - q(m(i) - m(i-1)).$$

onde usamos que $p + q = 1$*. Definindo*

$$y_i = m(i) - m(i-1),$$

tem-se, para $i \in [1, c-1]$,

$$-1 = py_{i+1} - qy_i$$

e

$$m(i) = y_1 + y_2 + \ldots + y_i. \tag{2.26}$$

Resolvendo para $p = q = \dfrac{1}{2}$. *tem-se:*

$$\begin{aligned}
-1 &= \frac{1}{2}y_2 - \frac{1}{2}y_1,\\
-1 &= \frac{1}{2}y_3 - \frac{1}{2}y_2,\\
&\vdots\\
-1 &= \frac{1}{2}y_i - \frac{1}{2}y_{i-1}.
\end{aligned}$$

Efetuando a soma das expressões acima, obtém-se

$$-(i-1) = \frac{1}{2}y_i - \frac{1}{2}y_1,$$

ou seja, para $i \in [1, c]$,

$$y_i = y_1 - 2(i-1).$$

Substituindo esta expressão em (2.26) e observando que $y_1 = m(1)$, *obtém-se*

$$m(i) = im(1) - 2[1 + 2 + \ldots + (i-1)] = im(1) - i(i-1).$$

Agora a condição limitante $m(c) = 0$, *dá* $cm(1) = c(c-1)$, *consequentemente,* $m(1) = \dfrac{c(c-1)}{c}$ *e, portanto*

$$m(i) = i(c-i). \tag{2.27}$$

2.5 Topologia da Matriz de Transição

As propriedades topológicas tem especial importância no estudo das cadeias de Markov no que tange á transciência e recorrência, já que serão utilizadas nas demonstrações dos teoremas que seguem. Além de serem úteis para verificação da irredutibilidade das cadeias. Nesta seção apresentaremos as definições e estudaremos as propriedades topológicas.

Definição 2.2.1 (Estados acessíveis). *Um estado j é acessível a partir de um estado i, se existe um inteiro $M \geq 0$ tal que $p_{ij}(M) > 0$.*

Proposição 2.2.1. *Um estado i é sempre acessível a partir de si mesmo.*

Demonstração: Por definição,

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

Fazendo $M = 0$ e $i = j$, obtemos

$$p_{ii}(0) = \mathbb{P}(X_{0+n} = i | X_n = i) = 1.$$

□

Definição 2.2.2 (Estados Comunicantes). *Dados dois estados i e j, dizemos que os estados i e j são comunicantes se j é acessível a partir de i e vice versa. Indicamos isto por*

$$i \longleftrightarrow j.$$

Proposição 2.2.2. *Sendo $M \geq 1$, por (2.10) vemos que*

$$p_{ij}(M) = \sum_{i_1,\ldots,i_{M-1}j} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M-1}j}.$$

Portanto $p_{ij}(M) > 0$ se, e somente se, existe pelo menos um caminho $i, i_1, \ldots, i_{m-1}, j$ de i para j tal que $p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M-1}j} > 0$.

O interesse nos estados comunicantes é que eles formam uma partição da Cadeia de Markov Homogênea. Isto se dá pelo fato de *ser comunicante* ser uma classe de equivalência. Mostremos isto.
$i)$ Reflexiva:

$$i \longleftrightarrow i$$

Pela proposição 2.2.1 um estado i é sempre acessível a si mesmo.
$ii)$ Simétrica:

Se $i \neq j$

$$i \longleftrightarrow j \Longrightarrow j \longleftrightarrow i$$

Se i e j são comunicantes implica, por definição, que j é acessível a partir de i e vice versa, isto é, existe M_1 e M_2 tal que $p_{ij}(M_1) > 0$ e $p_{ji}(M_2) > 0$. Como as probabilidades são ambas positivas existe um caminho mínimo de j para i e de i para j.

$iii)$Transitiva:

$$i \longleftrightarrow j, j \longleftrightarrow k \Longrightarrow i \longleftrightarrow k,$$

Supondo que j é acessível a partir de i e k é acessível a partir de j, vamos mostrar que k é acessível a partir de i. Com efeito, seja M_1 e M_2 inteiros positivos, tais que:

$$p_{ij}(M_1) = \sum_{i_1,\ldots,i_{M-1} \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M_1-1} j} > 0$$

e

$$p_{jk}(M_2) = \sum_{j_1,\ldots,j_{M-2} \in E} p_{jj_1} p_{j_1 j_2} \ldots p_{j_{M_2-1} k} > 0.$$

Segue que para k temos

$$p_{ik}(M_1 + M_2) = \sum_{i_1,\ldots,i_{M_1+M_2} \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M_1+M_2-1} k}.$$

Observamos que $i_{M_1} \in E$ em algum momento j, desse modo

$$p_{ik}(M_1 + M_2) = \sum_{i_1,\ldots,i_{M_1},i_{M_1+1},\ldots,i_{M_1+M_2-1 \in E}} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M_1-1} i_{M_1}} p_{i_{M_1} i_{M_1+1}} \ldots p_{i_{M_1+M_2-1} k}$$

$$= \sum_{i_{M_1} \in E} \left(\sum_{i_1,\ldots,i_{M_1},i_{M_1+1},\ldots,i_{M_1+M_2-1} \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M_1-1} i_{M_1}} p_{i_{M_1} i_{M_1+1}} \ldots p_{i_{M_1+M_2-1} k} \right)$$

$$\geq \left(\sum_{i_1,\ldots,i_{M_1},i_{M_1+1},\ldots,i_{M_1+M_2-1} \in E} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{M_1-1} i_{M_1}} p_{i_{M_1} i_{M_1+1}} \ldots p_{i_{M_1+M_2-1} k} \right)$$

$$= \sum_{i_1,\ldots,i_{M_1-1}} \cdot \sum_{j_1,\ldots,j_{M_2-1}} p_{ii_1}\ldots p_{i_{M_1-1}j}.p_{jj_1}\ldots p_{j_2}\ldots p_{j_{M_2-1}k}$$

onde $j_1 = i_{M_1}, j_2 = i_{M_1+1}, \ldots, j_{M_2-1} = i_{M_1+M_2-1}$. Continuando,

$$\begin{aligned}
&= \sum_{i_1,\ldots,i_{M_1-1}} p_{ii_1}\ldots p_{i_{M_1-1}} \left(\sum_{j_1,\ldots,j_{M_2-1}} p_{jj_1}p_{j_1j_2}\ldots p_{j_{M_2-1}k} \right) \\
&= \sum_{i_1,\ldots,i_{M_1-1}} p_{ii_1}\ldots p_{i_{M_1-1}}p_{jk(M_2)} \\
&= p_{jk}(M_2) \sum_{i_1,\ldots,i_{M_1-1}} p_{ii_1}\ldots p_{i_{M_1-1}j} \\
&= p_{jk}(M_2)p_{ij}(M_1) > 0.
\end{aligned}$$

Consequentemente, pela definição 2.2.2, k é acessível a partir de i. Portanto os estados comunicantes são uma relação de equivalência e produzem uma partição do Espaço Estado E em classes de equivalência chamadas *classes comunicantes.*

Definição 2.2.3 (Conjunto Fechado)**.** *Um estado i tal que $p_{ii} = 1$ é definido como fechado. Geralmente, um conjunto C de estados tal que para todo $i \in C$, $\sum_{j\in C} p_{ij} = 1$ é dito fechado.*

Exemplo 2.2.7. *O diagrama de transição da figura 2.4 tem três classes comunicantes: $\{1, 2, 3, 4\}$, $\{5, 7, 8\}$ e $\{6\}$. O estado 6 é fechado. A classe comunicante $\{5, 7, 8\}$ não é fechada, no entanto a classe $\{5, 6, 7, 8\}$ é fechada.*

Definição 2.2.4. *Irredutibilidade*
Se existe somente uma classe comunicante, então a cadeia, é chamada de irredutível.

2.6 Período

Consideremos o passeio aleatório sobre Z descrito no exemplo 2.2.1. Desde que $p \in (0, 1)$, a cadeia é irredutível. Observa-se que $E = C_0 + C_1$, onde C_0 e C_1, são os conjuntos de pares e ímpares relativos aos inteiros, tendo a seguinte propriedade: Se partindo de $i \in C_0$ ($C_0 = \{0, \pm2, \pm4, \ldots\}$) em um único passo somente pode

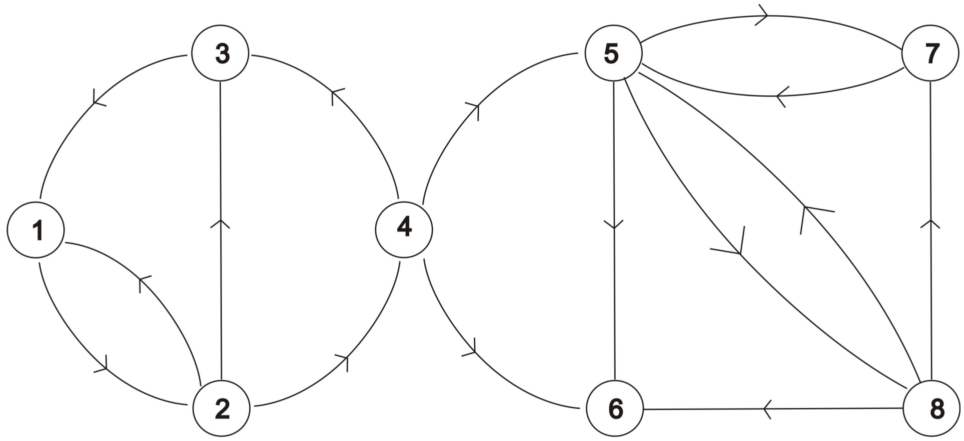

Figura 2.4: Diagrama com 3 classes comunicantes

alcançar o estado $j \in C_1$ ($C_1 = \{\pm 1, \pm 3, \pm 5, \ldots\}$) e, quando parte de um estado $i \in C_1$ só chega a um estado $j \in C_0$. A cadeia $\{X_n\}$ passa alternadamente de uma classe para outra. Neste sentido, a cadeia tem um comportamento periódico, e que corresponde ao período 2.

Iniciamos com o conceito geral de período para estados simples.

Definição 2.2.5 (Definição Aritmética de Período)**.** *O período d_i do estado $i \in E$, é por definição,*

$$d_i = mdc\{n \geq 1; p_{ii}(n) > 0\},$$

convencionando-se que $d_i = \infty$, se não existir $n \geq 1$ com $p_{ij}(n) > 0$. Se $d_i = 1$ o estado i é denominado de aperiódico.

Observamos que para o estado i o período d_i é o menor número natural tal que $p_{ii}(d_i) > 0$. Também, qualquer outro número natural tal que $p_{ii}(n) > 0$ deve ser da forma $n = Nd_i$, para algum $N > 0$. Em outras palavras i é acessível a i em caminhos com quantidade de passos múltiplos de d_i. Além disso, temos uma estrutura cíclica de acessibilidade de i para i com um espaço de tempo d_i.

Mostraremos a seguir que período é uma propriedade de classe comunicante.

Teorema 2.3. *Se os estados i e j se comunicam, então têm o mesmo período.*

Demonstração: Como i e j se comunicam, existe inteiros N e M tais que $p_{ij}(M) > 0$ e $p_{ji}(N) > 0$. Para qualquer $k \geq 1$,

$$p_{ij}(M + nk + N) \geq p_{ij}(M)(p_{jj}(k))^n p_{ji}(N).$$

De fato, o caminho $X_0 = i, \ldots, X_M = j, \ldots, X_{M+nk+N} = i$ é o caminho para ir de i para i em $M+nk+n$ passos. Portanto, para algum $k \geq 1$ tal que $p_{jj}(k) > 0$ tem-se que $p_{ii}(M+nk+N) > 0$, para todo $n \geq 1$. Assim, d_i divide $M+nk+N$ para todo $n \geq 1$. Em particular, d_i divide k. Desse modo d_i divide todo k tal que $p_{jj}(k) > 0$. Ainda, d_i divide d_j e, finalmente, $d_i = d_j$. □

A seguir demonstraremos um Teorema que nos auxiliará no estudo das CMH irredutíveis.

Teorema 2.4 (Teorema do entrelaçamento)**.** *Seja P uma matriz estocástica irredutível de período d. Então para todos os estados i, j, existem $M \geq 0$ e $n_0 \geq 0$, com M e n possivelmente dependentes de i, j tal que*

$$p_{ij}(M + nd) > 0, \ \forall n \geq n_0. \tag{2.28}$$

Demonstração: Basta provar para isso que $i = j$. Na verdade existe M tal que $p_{ij}(M) > 0$, porque j é acessível a partir de i e, portanto, se para algum $n_0 \geq 0$, tem-se $p_{jj}(nd) > 0$ para todo $n \geq n_0$, então $P_{ij}(M + nd) \geq p_{ij}(M)p_{jj}(nd) > 0$ para todo $n \geq n_0$. O mdc do conjunto $A = \{k \geq 1; p_{jj}(k) > 0\}$ é d, e A é fechado para a adição. O conjunto A portanto, contém todos números finitos de múltiplos positivos de d (ver Teorema 1.1 apêndice). Desse modo, existe n_0 tal que $n > n_0$ implica em $p_{jj}(nd) > 0$. □

Com o Teorema acima conseguimos generalizar o exemplo dado no início desta subseção. Provamos este resultado de maneira própria.

Teorema 2.5 (Estrutura Cíclica)**.** *Para toda cadeia irredutível de Markov encontra-se uma única partição de E em d classes $C_0, C_1, \ldots, C_{d-1}$, tal que para todo*

$k \in \{0, \ldots, d-1\}$ *e* $i \in C_k$,

$$\sum_{j \in C_{k+1}} p_{ij} = 1.$$

Por convenção, $C_d = C_0$*, e onde* d *é o período da CMH.*

Demonstração: Pelo Teorema 2.4, sabemos que para todo i, j existem $m > 0$ e $n \geq n_0$.

Vamos utilizar este fato na construção da partição. Fixemos um estado $i \in E$ arbitrário e a partir de i construímos os seguintes conjuntos

$$C_0 = \{j \in E : \exists \ \ n_0 \geq 0 \ \text{ tal que } \ p_{ij}(nd) > 0, \forall n \geq n_0\}$$

$$C_1 = \{j \in E : \exists \ \ n_1 \geq 0 \ \text{ tal que } \ p_{ij}(1 + nd) > 0, \forall n \geq n_1\}$$

$$C_2 = \{j \in E : \exists \ \ n_2 \geq 0 \ \text{ tal que } \ p_{ij}(2 + nd) > 0, \forall n \geq n_2\}$$

$$\vdots$$

$$C_{d-1} = \{j \in E : \exists \ \ n_{d+1} \geq 0 \ \text{ tal que } \ p_{ij}((d-1) + nd) > 0, \forall n \geq n_{d-1}\}.$$

É claro que $E = C_0 \cup C_1 \cup C_2 \ldots C_{d-1}$. Além disso, tal construção independe dos i's. De fato, dado $i \in C_0$ um elemento qualquer podemos gerar as outras classes $C_1, \ldots, C_{d-1}$.

Para concluirmos que $C_0, C_1, \ldots, C_{d-1}$ é uma partição, nos resta mostrar que estes conjuntos são mutuamente disjuntos. Sejam $k < l < d$. Suponhamos que j pertença a C_k e a C_l. Por construção existem $n_k \geq 0$ e $n_d \geq 0$ tais que

$$p_{ij}(k + nd) \geq 0, \forall n \geq n_k \ \text{ e } \ p_{ij}(l + nd) > 0, \quad \forall n \geq n_l.$$

Tome $n \geq max\{n_k, n_l\}$. De maneira explicita

$$\mathbb{P}(X_{1+k+nd} = j | X_0 = i) > 0.$$

Destas duas conclusões vemos que a cadeia visitou j no tempo $1 + (k + nd)$ e no tempo $1 + (l + nd)$. Consequentemente, existe um probabilidade positiva da CMH visitar j num tempo $k - l \geq 1$ quando partiu de j. Em outras palavras

$$p_{jj}(k - l) > 0.$$

Porém, isto é uma contradição com a definição de período. Portanto, $C_k \bigcap C_l = \emptyset$ para todo $k, l < d$ De acordo com o que foi mostrado acima,

$$E = C_0 \cup C_1 \cup \ldots \cup C_{d-1}$$

é uma partição de E.

Para a outra afirmação, primeiro observamos que $\sum_{j \in C_k} p_{kj} = 1$, para todo $0 \leq k \leq d$. Mostraremos a afirmação para C_0 e C_1, os outros casos são análogos. Calculando

$$\begin{aligned} p_{ij}(1+nd) &= \sum_{k \in E} p_{ik}(nd) p_{kj} \\ &= \sum_{k \in C_0} p_{ik}(nd) p_{kj} \end{aligned}$$

onde usamos a definição de C_0 na última igualdade. Somando sobre os termos de C_1 vemos que

$$\sum_{j \in C_1} p_{ij}(1+nd) = \sum_{k \in C_0} p_{ii}(nd) \sum_{j \in C_1} p_{kj}.$$

Suponhamos que $\sum_{k \in C_1} p_{kj} < 1$. Então

$$1 = \sum_{k \in C_0} p_{ik}(nd) \sum_{j \in C_1} p_{kj} < \sum_{k \in C_0} p_{ik}(nd) = 1.$$

Isto é um absurdo. Portanto $\sum_{j \in C_1} p_{kj} = 1$ para qualquer $K \in C_0$. □

O número $d \geq 1$ é denominado *período* da matriz ou do diagrama de transição. As classes $C_0, C_1, \ldots, C_{d-1}$, são ditas *classes cíclicas*.

A cadeia se move de uma classe para outra em cada transição ciclicamente, ver figura 2.5.

Exemplo 2.5.1. *A cadeia com diagrama de transição 2.6 é irredutível e tem período $d = 3$, com as classes cíclicas $C_0 = \{1, 2\}, C_1 = \{4, 7\}, C_3 = \{3, 5, 6\}$.*

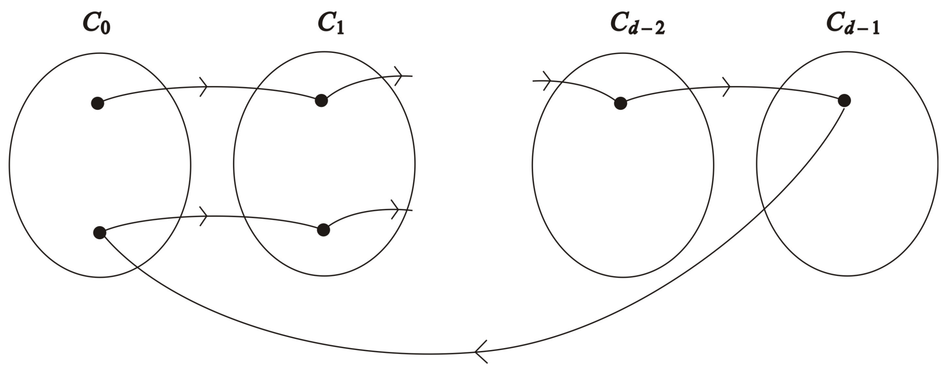

Figura 2.5: Ciclos

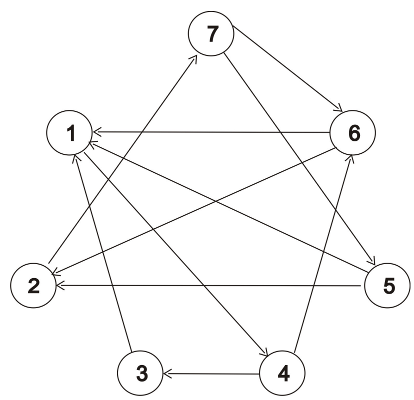

Figura 2.6: Diagrama de transição com 3 períodos

Seja uma cadeia irredutível de período d com as classes cíclicas $C_0, C_1, \ldots, C_d$. Renumerando os estados de E, caso necessário, a matriz de transição tem a estrutura de bloco abaixo, onde $d = 4$,

$$\mathbf{P} = \begin{array}{c} \\ C_0 \\ C_1 \\ C_2 \\ C_3 \end{array} \begin{array}{c} \begin{array}{cccc} C_0 & C_1 & C_2 & C_3 \end{array} \\ \begin{pmatrix} 0 & A_0 & 0 & 0 \\ 0 & 0 & A_1 & 0 \\ 0 & 0 & 0 & A_2 \\ A_3 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

e portanto P^2, P^3 e P^4 também tem estrutura em bloco correspondendo a C_0, C_1, C_2

e C_3. Em termo de matrizes

$$\mathbf{P}^2 = \begin{pmatrix} 0 & 0 & B_0 & 0 \\ 0 & 0 & 0 & B_1 \\ B_2 & 0 & 0 & 0 \\ 0 & B_3 & 0 & 0 \end{pmatrix} \mathbf{P}^3 = \begin{pmatrix} 0 & 0 & 0 & D_0 \\ D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 \\ 0 & 0 & D_3 & 0 \end{pmatrix} \mathbf{P}^4 = \begin{pmatrix} E_0 & 0 & 0 & 0 \\ 0 & E_1 & 0 & 0 \\ 0 & 0 & E_2 & 0 \\ 0 & 0 & 0 & E_3 \end{pmatrix}$$

São observados dois fenômenos: transição de blocos e forma diagonal em blocos em $\mathbf{P}^4$. É claro que a matriz de transição $\mathbf{P}^d$ tem a forma diagonal de bloco correspondente à classe cíclica $C_0, C_1, \ldots, C_{d-1}$:

$$\mathbf{P}^d = \begin{array}{c} \\ C_0 \\ C_1 \\ \vdots \\ C_{d-1} \end{array} \begin{array}{c} \begin{array}{cccc} C_0 & C_1 & \cdots & C_{d-1} \end{array} \\ \begin{pmatrix} E_0 & 0 & \cdots & 0 \\ 0 & E_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & E_{d-1} \end{pmatrix} \end{array} \tag{2.29}$$

O d-passo da matriz de transição P^d é também uma matriz estocástica e que a P-classe cíclica $C_0, C_1, C_2, \ldots C_{d-1}$ são diferentes P^d classes comunicantes, como a estrutura em bloco diagonal apresentada acima. Pode-se questionar: Existe na classe C_0 mais que uma P^d classe comunicante? Verifica-se que não e, portanto a matriz E_0 em (2.29) é uma matriz estocástica irredutível. Para verificar toma-se dois diferentes estados $i, j \in C_0$. Sendo estes P-comunicantes, assumindo que a cadeia é irredutível, existe $m > 0$ e $n > 0$ tal que $p_{ij}(m) > 0$ e $p_{ji}(n) > 0$ e que P tem período d. Então $m = Md$, $n = Nd$ para algum $M > 0$ e $N > 0$. Assim, $p_{ij}(Md) > 0$ e $P_{ji}(Nd) > 0$. Mas, $p_{ij}(Md)$ é o (i, j) termo de $(P^d)^M$, e igualmente $p_{ji}(Nd)$ é o (j, i) termo de $(P^d)^N$. Provou-se, portanto que i e j são P^d-comunicantes.

2.7 Estados Estáveis

Definição 2.5.1 (Distribuição estacionária). *A distribuição de probabilidade π que satisfaz*

$$\pi^T = \pi^T \mathbf{P} \tag{2.30}$$

é denominada de distribuição estacionária da matriz de transição $\mathbf{P}$ *ou da correspondente CMH.*

A equação de equilíbrio global (2.30) diz-se que para todos os estados i,

$$\pi(i) = \sum_{j \in E} \pi(j) p_{ji}. \tag{2.31}$$

Iterando (2.30) vemos que $\pi^T = \pi^T \mathbf{P}^n$ para todo $n \geq 0$. Se a distribuição inicial for $v = \pi$, por (2.8), obtemos $\mathbf{v}_n^T = \mathbf{v}\mathbf{P}^n = \pi^T$. Assim, $\mathbf{v}_n = \pi$. De maneira análoga ao Teorema 2.1 é possível mostrar que

$$\mathbb{P}(X_n = i_0, X_{n+1}. = i_1, \ldots, X_{n+k} = i_k) = v_n(i_0) p_{i_0} i_1 \ldots p_{i_{k-1} i_k}.$$

Sendo $\mathbf{v} = \pi$ a distribuição estacionária, segue que

$$\mathbb{P}(X_n = i_0, X_{n+1} = i_1, \ldots, X_{n+k} = i_k) = \pi(i_0) p_{i_0} i_1 \ldots p_{i_{k-1} i_k}.$$

É claro que a distribuição acima não depende de n. Em resumo, pode se dizer que a cadeia está em *regime estacionário*, em *equilíbrio* ou em *estado estacionário.*

Exemplo 2.5.2 (Cadeias de Markov de Dois Estados)**.** *Tomando* $E = \{1, 2\}$*, definimos a matriz de transição*

$$P = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

onde $\alpha, \beta, \in (0, 1)$*. As equações de equilibrio global são:*

$$\begin{aligned} \pi(1) &= \pi(1)(1-\alpha) + \pi(2)\beta \\ \pi(2) &= \pi(1)\alpha + \pi(2)(1-\beta). \end{aligned}$$

Este é um sistema dependente onde é reduzido a uma única equação $\pi(1)\alpha = \pi(2)\beta$*. Usando* $\pi(1) + \pi(2) = 1$ *concluímos que*

$$\pi(1) = \frac{\beta}{\alpha + \beta}, \quad \pi(2) = \frac{\alpha}{\alpha + \beta}.$$

Exemplo 2.5.3 (Caminho Aleatório Simétrico)**.** *Um caminho aleatório simétrico em $\mathbb{Z}$ não pode ter uma distribuição estacionária. Pelo Exemplo 2.2.1 vemos que este processo é uma CMH. Realmente, a solução da equação de equilíbrio*

$$\pi(i) = \frac{1}{2}\pi(i-1) + \frac{1}{2}\pi(j+1)$$

para $i \geq 0$, com os dados iniciais $\pi(0)$ e $\pi(1)$, é

$$\begin{aligned}
\pi(i) &= \lambda + \mu i \\
\pi(0) &= \lambda \\
\pi(1) &= \lambda + \mu \\
\mu &= \pi(1) - \lambda \\
\mu &= \pi(1) - \pi(0).
\end{aligned}$$

$$\pi(i) = \pi(0) + (\pi(1) - \pi(0))i,$$

ver exemplo 2.2.4. Desde que $\pi(i) \in [0,1]$, necessariamente $\pi(1) - \pi(0) = 0$. Portanto, π é uma constante. Agora, sabemos que: $\sum_{i \in E} \pi(i) < \infty$. Mas,

$$\sum_{i \in E} \pi(i) = \sum_{i \in E} \pi(0) = \pi(0) \sum_{i \in E} i,$$

Por estes dois fatos, $\pi(0) = 0$. Concluímos que $\pi(i) = 0$. Como isto é válido para todo i, segue que $\pi = 0$. O que contradiz o fato de π ser uma distribuição de probabilidades.

Seja $\{X_n\}_{n \geq 0}$ uma CHM com matriz de transição $\mathbf{P}$ e admitindo uma distribuição estacionária π tal que

$$\pi(i) > 0$$

para todos os estados i. Definamos uma matriz $\mathbf{Q}$, indexada por E, tal que

$$\pi(i)q_{ij} = \pi(j)p_{ji}. \tag{2.32}$$

$\mathbf{Q}$ é uma matriz estocástica. De fato,

$$\sum_{j\in E} q_{ij} = \sum_{j\in E} \frac{\pi(j)}{\pi(i)} p_{ji} = \frac{1}{\pi(i)} \sum_{j\in E} \pi(j) p_{ji} = \frac{\pi(i)}{\pi(i)} = 1,$$

onde na terceira igualdade usamos a definição de distribuição estacionária. Temos a seguinte interpretação. Suponhamos que a distribuição inicial de $\{X_n\}$ seja π, ou seja, para todo $n \geq 0$

$$\mathbb{P}(X_n = i) = \pi(i). \tag{2.33}$$

Pela regra da inversão de Bayes (Teorema 1.1),

$$\mathbb{P}(X_n = j | X_{n+1} = i) = \frac{\mathbb{P}(X_{n+1} = i | X_n = j)\mathbb{P}(X_n = j)}{\mathbb{P}(X_{n+1} = i)}$$

Por (2.32) e (2.33) é visto que

$$\mathbb{P}(X_n = j | X_{n+1} = i) = q_{ji}. \tag{2.34}$$

Tem-se , então que $\mathbf{Q}$ é uma matriz de transição de uma cadeia de Markov reversa quando o tempo é reverso.

Exemplo 2.5.4. *Seja a matriz de transição* $\mathbf{P}$ *da cadeia de Markov de dois estados, ver exemplo 2.1.1,*

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

Sabemos pelo exemplo 2.5.2 que sua distribuição estacionária é dada por

$$\pi(1) = \frac{\beta}{\alpha+\beta}, \quad \pi(2) = \frac{\alpha}{\alpha+\beta}.$$

Queremos encontrar a matriz $\mathbf{Q}$ *para este exemplo. Para tal observamos que* $E = \{1, 2\}$ *e tomamos a equação (2.32).*

$$\pi(i) q_{ij} = \pi(j) p_{ji}.$$

Desta equação encontramos as quatro entradas de $\mathbf{Q}$ *dadas por*

$$\begin{aligned}\pi(1)q_{11} &= \pi(1)p_{11}\\ \pi(1)q_{12} &= \pi(2)p_{21}\\ \pi(2)q_{21} &= \pi(1)p_{12}\\ \pi(2)q_{22} &= \pi(2)p_{22}.\end{aligned}$$

Cãlculos simples mostra que a matriz $\mathbf{Q}$ *é dada por*

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

Teorema 2.6 (Teste de reversão)**.** *Sejam* $\mathbf{P}$ *e* $\mathbf{Q}$ *matrizes estocásticas indexadas por um conjunto enumerável* E. *Seja* π *uma distribuição de probabilidade tal que para todo* $i, j \in E$, *têm-se*

$$\pi(i)q_{ij} = \pi(j)p_{ji}. \tag{2.35}$$

Então π *é uma distribuição estacionária de* $\mathbf{P}$.

Demonstração: Para um estado $i \in E$ fixo, somamos ambos os membros da igualdade (2.35) em relação a $j \in E$ obtemos

$$\sum_{j\in E} \pi(i)q_{ij} = \sum_{j\in E} \pi(j)p_{ji}.$$

Temos que o lado esquerdo da expressão é igual a $\pi(i)\sum_{j\in E} q_{ij} = \pi(i)$, e, portanto, para todo $i \in E$,

$$\pi(i) = \sum_{j\in E} \pi(j)p_{ji}.$$

□

2.8 Regeneração

Determinar o tempo em que um estado é alcançado é uma tarefa importante no estudo das cadeias de Markov. Para determinação de recorrência e transciência dos estados, o tempo de retorno informa se um estado é recorrente ou transciente, pois conhecendo esta informação será possível identificar sua natureza

2.8.1 Propriedade Forte de Markov

Definição 2.6.1 (Tempo de parada). *Seja um processo estocástico* $\{X_n\}_{n\geq 0}$. *Uma variável aleatória* $\tau : \Omega \longrightarrow \mathbb{N}\cup\{+\infty\}$ *tal que, para todos os inteiros* $m \geq 0$, *o evento* $\{\tau = m\} = \{\omega : \tau(\omega) = m\}$ *possa ser expresso em termos de* $X_0, X_1, \ldots, X_m$, *é dito tempo de parada.*

Exemplo 2.6.1 (Tempo de Retorno). *Na teoria de Cadeia de Markov, o mais importante tempo de parada é o tempo de retorno ao estado* $i \in E$ **[2]**, *isto é*

$$T_i = inf\{n \geq 1; X_n = i\},$$

onde $T_i = \infty$ *se* $X_n \neq i$ *para todo* $n \geq 1$. *Em verdade* T_i *é um tempo de parada. De fato*

$$\{T_i = m\} = \{X_m = i, \sum_{n=1}^{m} 1_{\{X_n=1\}} = 1\},$$

o qual é escrito em termos de $X_0, X_1, \ldots, X_m$.

Seja $\tau : \Omega \to \mathbb{N}$ um tempo de parada. Olhando a cadeia de Markov X do seguinte modo

$$X : \Omega \times \mathbb{N} \longrightarrow E$$

$$(\omega, m) \longrightarrow X(\omega, m)$$

podemos fazer a composição de X com τ, isto é, $X_{\tau(\omega)}(\omega) = X(\omega, \tau(\omega))$. No caso de $\tau \in \mathbb{N}$ não há problemas. Porém, quando $\tau = \infty$ tomamos um elemento arbitrário Δ que não esteja em E para definirmos

$$X_\infty = \Delta.$$

Sendo possível compor uma cadeia de Markov X com os tempos de parada, definimos o processo "$\{X_n\}$ anterior a τ" como

$$\{X_{n\wedge\tau}\}_{n\geq 0},$$

onde $n \wedge \tau = \min\{n, \tau\}$. O processo "$\{X_n\}$ posterior a τ" é definido por

$$\{X_{n+\tau}\}_{n\geq 0}.$$

Teorema 2.7. *Propriedade Forte de Markov*

Seja $\{X_n\}_{n\geq 0}$ uma CMH em um espaço enumerável E e uma matriz de transição $\mathbf{P}$. Seja τ, um tempo de parada com respeito a esta cadeia. Então para qualquer estado $i \in E$, dado que $X_\tau = i$ (em particular $\tau < \infty$, desde que $i \neq \Delta = \infty$) a seguintes afirmações são válidas

(a) O processo posterior a τ e o processo anterior a τ são independentes;

(b) O processo posterior a τ é uma CMH com matriz de transição $\mathbf{P}$.

Demonstração:

(a) Vamos mostrar que para todos os tempos $k \geq 1, n \geq 0$, e todos os estados $i_0, \ldots, i_n, i, j_1, \ldots, j_k$, vale

$$\begin{aligned}\mathbb{P}(X_{\tau+1} = j_1, \ldots, X_{\tau+k} = j_k | X_\tau = i, X_{\tau\wedge n} = i_0, \ldots, X_{\tau\wedge n} = i_n)\\ = \mathbb{P}(X_{\tau+1} = j_1, \ldots, X_{\tau+k} = j_k | X\tau = i).\end{aligned} \tag{2.36}$$

Mostraremos a seguir uma versão simplificada da igualdade (2.36),

$$\mathbb{P}(X_{\tau+k} = j | X_\tau = i, X_{\tau\wedge n} = i_n) = \mathbb{P}(X_{\tau+k} = j | X_\tau = i). \tag{2.37}$$

A generalização pode ser obtida por argumentos semelhantes. O lado esquerdo da igualdade acima é igual a

$$\frac{\mathbb{P}(X_{\tau+k} = j, X_\tau = i, X_{\tau\wedge n} = i_n)}{\mathbb{P}(X_\tau = i, X_{\tau\wedge n} = i_n)} \tag{2.38}$$

Observando que $\tau : \Omega \longrightarrow \mathbb{N} \cup \{+\infty\}$ e $\Omega = \bigcup_{r\geq 0}\{\tau = r\}$ o numerador acima pode ser reescrito como

$$\begin{aligned}
& \mathbb{P}(X_{\tau+k}=j, X_\tau=i, X_{\tau\wedge n}=i_n) \\
= \; & \mathbb{P}(\Omega, X_{\tau+k}=j, X_\tau=i, X_{\tau\wedge n}=i_n) \\
= \; & \mathbb{P}(\bigcup_{r\geq 0}\{\tau=r\}, x_{\tau+k}=j, X_\tau=i, X_{\tau\wedge n}=i_n) \\
= \; & \sum_{r\geq 0}\mathbb{P}(\{\tau=r\}, X_{\tau+k}=j, X_\tau=i, X_{\tau\wedge n}=i_n) \\
= \; & \sum_{r\geq 0}\mathbb{P}(\tau=r, X_{r+k}=j, X_r=i, X_{r\wedge n}=i_n). \qquad (2.39)
\end{aligned}$$

Desenvolvamos o termo geral da série acima.

$$\begin{aligned}
& \mathbb{P}(\tau=r, X_{r+k}=j, X_r=i, X_{r\wedge n}=i_n) \\
= \; & \frac{\mathbb{P}(\tau=r, X_{r+k}=j, X_r=i, X_{r\wedge n}=i_n)}{\mathbb{P}(\tau=r, X_r=i, X_{r\wedge n}=i_n)}.\mathbb{P}(\tau=r, X_{r\wedge n}=i_n, X_r=i) \\
= \; & \mathbb{P}(X_{r+k}=j|X_r=i, X_{r\wedge n}=i_n, \tau=r).\mathbb{P}(\tau=r, X_{r\wedge n}=i_n, X_r=i).
\end{aligned}$$

Já que $r\wedge n\leq r$ e $\{\tau=r\}\in X_0^r$, o evento $B=\{X_{r\wedge n}=i_n, \tau=r\}$ está en X_0^r. Portanto pela propriedade de Markov

$$\mathbb{P}(X_{r+k}=j|X_r=i, X_{r\wedge n}=i_n, \tau=r)=\mathbb{P}(X_{r+k}=j|X_r=i)=p_{ij}(k).$$

Assim expressão (2.39) fica reduzida a

$$\sum_{r\geq 0}p_{ij}(k)\mathbb{P}(\tau=r, X_{r\wedge n}=i_n, X_r=i)=p_{ij}(k)\mathbb{P}(X_{\tau=i}, X_{\tau\wedge n}=i_n).$$

Portanto o lado esquerdo da expressão (2.37) é $p_{ij}(k)$. Cálculos semelhantes mostram que o lado direito de (2.37) também é $p_{ij}(k)$, e desse modo fica demonstrado o item (a).

(b) Devemos mostrar que para todos os estados $i, j, k, i_{n-1}, \ldots, i_1$

$$\begin{aligned}
& \mathbb{P}(X_{\tau+n+1}=k|X_{\tau+n}=j, X_{\tau+n-1}=i_{n-1}, \ldots, , X_\tau=i) \\
& \qquad = \mathbb{P}(X_{\tau+n+1}=k|X_{\tau+n}=j)=p_{jk}.
\end{aligned}$$

A primeira igualdade segue do que foi demonstrado no item (a) para o tempo de parada $\tau' = \tau + n$, o processo anterior e o posterior a τ' são independentes dado $X_{\tau'} = j$. A segunda igualdade é obtida através de cálculos semelhantes ao demonstrado no item (a). $\square$

2.8.2 Ciclos regenerativos

O objetivo desta seção é determinar o número de visitas a um estado. A expressão abaixo, conta o número de visitas ao estado i,

$$N_i = \sum_{n \geq 1} \mathbf{1}_{(X_n = i)}. \tag{2.40}$$

A seguir mostraremos alguns fatos que nos ajudarão na demonstração do Teorema à frente. Observamos que $\{X_n = i\} = \{\omega : X_n(\omega) = i\}$ e que

$$\mathbf{1}_{\{X_n(\omega)=i\}} = \begin{cases} 1, & \text{se} \quad \omega \in \{X_n = i\} \\ 0, & \text{se} \quad \omega \notin \{X_n = i\}. \end{cases}$$

De (2.40) podemos escrever

$$N_i = N_i(\omega) = 1_{\{X_0=i\}}(\omega) + 1_{\{X_1=i\}}(\omega) + 1_{\{X_2=i\}}(\omega) + \ldots$$

Supondo $X_0 = j$ o estado inicial e r o número de visitas ao estado i, ou seja, $N_i = r$, temos

$$\mathbb{P}(N_i = r | X_0 = j) = \mathbb{P}_j(N_i = r),$$

o que é a probabilidade do número de visitas ao estado i ser r quando partindo do estado inicial j. Após estas considerações enunciamos o seguinte teorema.

Teorema 2.8 (Número visitas a um estado). *A distribuição de N_i dado $X_0 = j$ é*

$$\mathbb{P}_j(N_i = r) = \begin{cases} f_{ji} f_{ii}^{r-1}(1 - f_{ii}) & \textit{para} \quad r \geq 1 \\ 1 - f_{ji} & \textit{para} \quad r = 0, \end{cases} \tag{2.41}$$

onde

$$f_{ji} = \mathbb{P}_j(T_i < \infty)$$

e T_i *é o tempo de retorno ao estado* i*.*

Demonstração: Para $r = 0$, queremos mostrar que $\mathbb{P}_j(N_i = 0) = 1 - f_{ji}$, onde $f_{ji} = \mathbb{P}(T_i < \infty | X_0 = j)$. É claro que $\{T_i < \infty\} \cup \{N_i = 0\} = \Omega$.

Supondo que $\omega \in \{N_i = 0\}$, ou seja, $N_i(\omega) = 0$. Logo $\omega \notin \{T_i < \infty\}$. Portanto, $\{T_i < \infty\} \cap \{N_i = 0\} = \emptyset$, o que significa que os eventos são complementares, isto é,

$$\{T_i < \infty\} \dot{\cup} \{N_i = 0\} = \Omega.$$

Segue então que

$$\mathbb{P}(\{T_i < \infty\} \dot{\cup} \{N_i = 0\} | X_0 = j) = \mathbb{P}(\Omega | X_0 = j) = 1.$$

Continuando

$$\mathbb{P}(\{T_i < \infty\} | X_0 = j) + \mathbb{P}(\{N_i = 0\} | X_0 = j) = 1$$

Portanto, quando $r = 0$, a probabilidade será

$$\mathbb{P}_j(N_i = 0) = 1 - f_{ji}.$$

Seja $r \geq 1$ e assumamos (2.41) verdadeira para todo $k \in [1, r]$. Em particular,

$$\mathbb{P}_j(N_i > r) \overset{(1)}{=} 1 - \sum_{k=0}^{r} \mathbb{P}_j(N_i = k) \overset{(2)}{=} f_{ji} f_{ii}^r.$$

Verifiquemos a igualdade (1). Vemos que

$$\{N_i > r\} = \dot{\bigcup_{k=r+1}} \{N_i = k\}.$$

É claro que

$$\mathbb{P}_j(\bigcup_{k=r+1}^{\infty} \{N_i = k\} \dot{\cup} \bigcup_{k=0}^{r} \{N_i = k\}) = \mathbb{P}_j(\Omega) = 1.$$

Disto

$$\mathbb{P}_j(N_i > r) + \mathbb{P}_j(\cup_{k=0}^{r}\{N_i = k\}) = 1.$$

Sendo $\dot{\bigcup\limits_{k=0}}\{N_i = k\}$ disjunta,

$$\mathbb{P}_j(N_i > r) + \sum_{k=0}^{r} \mathbb{P}_j(N_i = k) = 1.$$

Portanto,

$$\mathbb{P}_j(N_i > r) = 1 - \sum_{k=0}^{r} \mathbb{P}_j(N_i = k).$$

Agora, verificaremos a igualdade (2). Da igualdade (1) deduzimos que

$$\begin{aligned}
1 - \sum_{k=0}^{r} \mathbb{P}_j(N_i = k) &= 1 - (\mathbb{P}_j(N_i = 0) + \mathbb{P}_j(N_i = 1) + \mathbb{P}_j(N_i = 2) + \ldots + \mathbb{P}_j(N_i = r)) \\
&= 1 - [(1 - f_{ji}) + f_{ji}f_{ii}^0(1 - f_{ii}) + f_{ji}f_{ii}^1(1 - f_{ii}) + f_{ji}f_{ii}^1(1 - f_{ii}^2) \\
&+ f_{ji}f_{ii}^3(1 - f_{ii}) + \ldots + f_{ji}f_{ii}^{r-1}(1 - f_{ii})] \\
&= 1 - (1 - f_{ji} + f_{ji}f_{ii}^0 - f_{ji}f_{ii} + f_{ji}f_{ii} - f_{ji}f_{ii}^2 + f_{ji}f_{ii}^2 - f_{ji}f_{ii}^3 \\
&+ f_{ji}f_{ii}^3 \ldots + f_{ji}f_{ii}^{r-1} - f_{ji}f_{ii}^r) \\
&= 1 - [1 - f_{ji}f_{ii}^r] \\
&= f_{ji}f_{ii}^r,
\end{aligned}$$

onde usamos a hipótese indutiva na segunda igualdade.

Supondo que τ_r é o r-ésimo tempo de retorno para o estado i, temos a igualdade

$$\{N_i = r + 1\} = \{X_{\tau_{r+1} = i}\}.$$

Deste fato deduzimos que

$$\mathbb{P}_j(N_i = r + 1) = \mathbb{P}_j(N_i = r + 1, X_{\tau_{r+1}} = i).$$

Como $\{N_i = r + 1\}$ significa que o estado i foi visitado $r + 1$ vezes e somente esta quantidade, temos que o $(r + 2)$-ésimo tempo de retorno é infinito. Em outras palavras $\{N_i = r+1\} = \{\tau_{r+2} - \tau_{r+1} = \infty\}$. Segue que

$$\begin{aligned}\mathbb{P}_j(N_i = r+1) &= \mathbb{P}_j(\tau_{r+2} - \tau_{r+1} = \infty, X_{\tau_{r+1}} = i) \\ &= \mathbb{P}_j(\tau_{r+2} - \tau_{r+1} = \infty | X_{\tau_{r+1}} = i)\mathbb{P}_j(X_{\tau_{r+1}} = i).\end{aligned}$$

Entretanto,

$$\begin{aligned}\mathbb{P}_j(\tau_{r+2} - \tau_{r+1} = \infty | X_{r+1} = i) &= \mathbb{P}(\tau_{r+2} - \tau_{r+1} = \infty | X_{\tau_{r+1}} = i, X_0 = j) \\ &= \mathbb{P}(\tau_{r+2} - \tau_{r+1} = \infty | X_{\tau_{r+1}} = i),\end{aligned}$$

pela propriedade forte de Markov. Desde que $\tau_{r+2} - \tau_{r+1}$ é o tempo de retorno para i do processo depois de τ_{r+1} a Propriedade forte de Markov dá

$$\mathbb{P}(\tau_{r+2} - \tau_{r+1} = \infty | X_{\tau_{r+1}} = i) = \mathbb{P}(T_i = \infty | X_0 = i).$$

Também

$$\mathbb{P}_j(X_{\tau_{r+1}} = i) = \mathbb{P}_j(N_i > r)$$

(se $N_i \leq r$, então $X_{\tau_{r+1}} = X_\infty = \Delta \notin E$). Portanto

$$\mathbb{P}_j(N_i = r+1) = \mathbb{P}_i(T_i = \infty)P_j(N_i > r) = (1 - f_{ii})f_{ji}f_{ii}^r,$$

o que completa a demonstração. $\square$

A distribuição de N_i, dado $X_0 = j$ e $N_i \geq 1$ é geométrica [9]. Isto acarreta duas consequências principais. Em primeiro lugar

$$\mathbb{P}_i(T_i < \infty) = 1 \Leftrightarrow \mathbb{P}_i(N_i = \infty) = 1.$$

Demonstremos a afirmação acima. Sejam $p \neq q$ quaisquer e $\omega \in \{N_i = p\} \cap \{N_i = q\}$. Então $\omega \in \{N_i = p\}$ e $\omega \in \{N_i = q\}$. Isto implica que $N_i(\omega) = p$ e $N_i(\omega) = q$. Disto obtemos $p = q$ uma contradição.

Pelo Teorema (2.8)(número de visitas a um estado),

$$\mathbb{P}_i(N_i = r) = \begin{cases} f_{ii}f_{ii}^{\tau-1}(1 - f_{ii}) & \text{para} \quad r \geq 1 \\ 1 - f_{ii} & \text{para} \quad r = 0. \end{cases}$$

Suponhamos que $\mathbb{P}_i(T_i < \infty) = 1$, ou seja, $f_{ii} = 1$. Então $\mathbb{P}(N_i = r) = 0$ para todo $r \geq 0$. Como $\Omega = \cup_{n=0}^{\infty}\{N_i = r\}$ temos

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{n=0}^{\infty}\{N_i = r\}) = \sum_{n=0}^{\infty} \mathbb{P}(N_i = r) = \mathbb{P}_i(N_i = \infty).$$

Por outro lado, sabemos que $1 = \sum_{n=0}^{\infty} \mathbb{P}_i(N_i = r)$ e suponhamos que $\mathbb{P}_i(N_i = \infty) = 1$. Pelo fato de que $\mathbb{P}_i(N_i = r) \geq 0$ para todo $r \geq 0$ é possível concluir que $\mathbb{P}_i(N_i = r) \geq 0$, para todo $r \geq 0$. Em particular $\mathbb{P}_i(N_i = 0) = 0$. Obtemos desta maneira $fii = 1$ e, consequentemente, $\mathbb{P}_i(T_i < \infty) = 1$. Em outras palavras, quando partindo de i, se quase com certeza se a retorna a i, então i será visitado infinitamente.

Em segundo lugar, tem-se

$$E_i[N_i] = \sum_{r=1}^{\infty} r\mathbb{P}_i(N_i = r) = \sum_{r=1}^{\infty} r f_{ii}^r(1 - f_{ii}) = \frac{f_{ii}}{1 - f_{ii}}.$$

Em particular

$$\mathbb{P}_i(T_i < \infty) < 1 \Leftrightarrow E_i[N_i] < \infty.$$

Temos $f_{ii} = \mathbb{P}_i(T_i < \infty) < 1$ e

$$\begin{aligned}
\sum_{r \geq 0}^{\infty} r(f_{ii}^r - f_{ii}^{r+1}) &= (f_{ii} - f_{ii}^2) + 2(f_{ii}^2 - f_{ii}^3) + 3(f_{ii}^3 - f_{ii}^4) + 4(f_{ii}^4 - f_{ii}^5) + \dots \\
&+ r(f_{ii}^r - f_{ii}^{r+1}) + \dots \\
&= f_{ii} + f_{ii}^2 + f_{ii}^3 + f_{ii}^4 + \dots + f_{ii}^r + \dots \\
&= \sum_{r \geq 0}^{\infty} (f_{ii})^r = \frac{f_{ii}}{1 - f_{ii}}.
\end{aligned}$$

Teorema 2.9 (Recorrência)**.** *Para algum estado $i \in E$,*

$$\mathbb{P}_i(T_i < \infty) = 1 \Leftrightarrow \mathbb{P}_i(N_i = \infty) = 1,$$

e

$$\mathbb{P}_i(T_i < \infty) < 1 \Leftrightarrow \mathbb{P}_i(N_i = \infty) \Leftrightarrow \mathbb{P}_i(N_i = \infty) = 0 \Leftrightarrow E_i[N_i] < \infty.$$

Particularmente, o evento $\{N_i = \infty\}$ tem probabilidade $\mathbb{P}_i$ igual a 0 ou 1.

Consideremos uma Cadeia de Markov com estado inicial convencionado ser 0, tal que $\mathbb{P}_0(T_0 < \infty) = 1$. Levando em conta o último Teorema, a cadeia partindo do estado 0 retornará infinita e frequentemente muitas vezes a este estado. Seja $\tau_1 = T_0, \tau_2, \ldots$ os sucessivos tempos de retorno a 0, e tome $\tau_0 \equiv 0$. Pela Propriedade Forte de Markov, para algum $k \geq 1$, o processo posterior a τ_k é independente do processo anterior a τ_k. Observa-se que a condição $X_{\tau_k} = 0$ é sempre satisfeita, e o processo posterior a τ_k é uma Cadeia de Markov como a mesma matrix de transição que a cadeia original e com estado inicial 0, por construção.

Teorema 2.10 (Teorema do ciclo regenerativo)**.** *Seja $\{X_n\}_{n\geq 0}$ uma CMH com estado inicial 0 que é sempre certamente frequentemente visitado infinitamente. Denotando por $\tau_0 = 0, \tau_1, \tau_2, \ldots$ os sucessivos tempos de visita a 0, as partes da trajetória*

$$\{X_{\tau_k}, X_{\tau_k+1}, \ldots, X_{\tau_{k+1}-1}\}, k \geq 0,$$

são independentes e identicamente distribuídas.

Tais partes são denominadas *ciclos regenerativos* da cadeia entre as visitas ao estado 0. Cada tempo aleatório τ_k é um *tempo de regeneração*, nesse sentido que $\{X_{\tau_k+n}\}_{n\geq 0}$ é independente do passado $X_0, \ldots, X_{\tau_k-1}$ e tem a mesma distribuição que $\{X_n\}_{n\geq 0}$. Em particular, a sequência $\{\tau_k - \tau_{k-1}\}_{k\geq 1}$ é independente e identicamente distribuída.

Capítulo 3

Recorrência, Transciência e Absorção

Neste capítulo estudaremos os critérios para determinação de recorrência e transciência, assim como da estabilidade de uma cadeia de Markov. O critério da matriz potência e a estrutura da matriz de transição são apresentados e, a partir destes mostramos que a recorrência é uma propriedade de classe. Serão abordados ainda os conceitos de absorção e tempo de absorção.

3.1 Critério da Potência da Matriz

3.1.1 Estados Recorrentes e Transcientes

No capítulo dois, o Teorema 2.9 trata de um primeiro conceito de recorrência. Nesta seção usaremos tal resultado para definir o que são estados recorrentes e transcientes.

Consideremos uma cadeia de Markov tomando valores em um espaço de estados $E = \mathbb{N}$. Seja $i \in \mathbb{N}$ o estado inicial da cadeia. Existe a possibilidade de que a cadeia nunca visite este estado, novamente, depois de decorrido um certo tempo aleatório finito. O que é uma característica indesejável. Como por exemplo, seja uma cadeia que conta o número de pessoas que esperam atendimento em uma fila de um certo banco.

A recorrência positiva é utilizada para avaliar a estabilidade de uma cadeia de Markov irredutível, quando o estado é visitado um número infinito de vezes e o tempo médio entre duas visitas sucessivas é finito. Para determinação da estabilidade de uma cadeia estudaremos as condições necessárias e suficientes, ou seja, o critério da matriz potência e o critério da distribuição estacionária. Iniciaremos com a definição fundamental para esta seção

Definição 3.0.1 (Recorrência e Transciência). *Seja $i \in E$, i será recorrente se*

$$\mathbb{P}_i(T_i < \infty) = 1, \tag{3.1}$$

de outro modo é denominado transciente. Um estado $i \in E$ que satisfaz

$$E_i[T_i < \infty] \tag{3.2}$$

é denominado recorrente positivo, se não, é dito recorrente nulo.

Exemplo 3.0.1 (Sucesso na corrida). *Seja um jogo que obedece a seguinte regra: uma moeda é lançada repetidamente. Seja p a probabilidade de ocorrer cara e q a probabilidade de ocorrer coroa. Considerando uma escada, se o resultado for coroa, sua probabilidade será $q = 1 - p$, e subiremos um degrau, se o resultado for cara desceremos um degrau. Seja X_n nossa posição no tempo n, pelo Exemplo 2.2.1, o processo $\{X_n\}_{n \geq 0}$ será uma CMH com espaço de estados $E = \mathbb{N}$.*

No diagrama de transição da cadeia do exemplo em questão (Figura 3.1), observamos que o espaço de estados é $E = \mathbb{N}$.

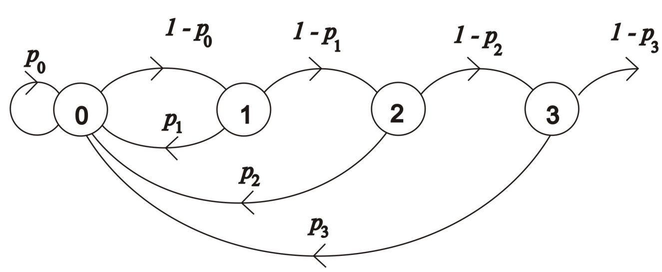

Figura 3.1: Diagrama de transição sucesso na corrida

Para garantir a irredutibilidade vamos impor a condição

$$p_i \in (0,1) \quad \textit{para todo} \quad i \in E.$$

A seguir calculemos a probabilidade de retorno ao estado inicial 0 e o tempo médio de retorno ao mesmo estado. Partindo do estado 0, existe apenas um caminho para retornar ao estado 0 em exatamente n passos. O caminho é $0, 1, \ldots, n-1, 0$. Portando, $\mathbb{P}_0(T_0 = 1) = p_0$, para $n \geq 1$,

$$\mathbb{P}_0(T_0 = n) = (1 - p_0) \ldots (1 - p_{n-2}) p_{n-1}.$$

Fato demonstrado em (2.10). Definamos $u_0 = 1$ e para $n \geq 1$

$$u_n = (1 - p_0) \ldots (1 - p_{n-1}).$$

Vemos que

$$\mathbb{P}(T_0 = n) = u_{n-1} - u_n.$$

Calculando

$$\mathbb{P}_0(T_0 < \infty) = \sum_{n=1}^{\infty} \mathbb{P}_0(T_0 = n) = \lim_{m \uparrow \infty} \sum_{n=1}^{m} \mathbb{P}(T_0 = n) = \lim_{m \uparrow \infty} (1 - u_m).$$

Tomando o limite e substituindo u_m temos

$$\mathbb{P}(T_0 < \infty) = 1 - \lim_{m \uparrow \infty} \prod_{i=0}^{m-1} (1 - p_i).$$

Tendo em vista o resultado de infinitos produtos (Ver Teorema 3.1 apêndice)

$$\mathbb{P}(T_0 < \infty) = 1 \Longleftrightarrow \lim_{m \uparrow \infty} \prod_{i=0}^{m-1} (1 - p_i) = 0 \Longleftrightarrow \sum_{i=0}^{\infty} p_i = \infty.$$

De um modo geral não é fácil determinar se um estado é recorrente ou transciente. A teoria das cadeias de Markov objetiva mostrar critérios para determinação da recorrência de uma cadeia. Em alguns casos o critério é satisfeito por condições necessárias e em outros por uma condição suficiente.

3.1.2 Matriz Potência

No capítulo 2 mostramos que $\mathbf{P}^n$ é a matriz de transição de uma cadeia de Markov em n passos. Uma matriz $\mathbf{G}$ associada com a matriz de transição $\mathbf{P}$ definida por

$$\mathbf{G} = \sum_{n\geq 0} \mathbf{P}^n$$

é denominada de matriz potência. Seu termo geral

$$g_{ij} = \sum_{n=0}^{\infty} p_{ij}(n) = \sum_{n=0}^{\infty} P_i(X_n = j) = \sum_{n=0}^{\infty} E_i[1_{\{X_n=j\}}] = E_i\left[\sum_{n=0}^{\infty} 1_{\{X_n=j\}}\right]$$

é a média de visitas ao estado j, quando a cadeia inicia no estado i.

Teorema 3.1 (Critério da matriz potência)**.** *O estado $i \in E$ é recorrente se, e somente se,*

$$\sum_{n=0}^{\infty} p_{ii}(n) = \infty. \tag{3.3}$$

Demonstração: Este teorema apenas reformula o Teorema 2.9 do capítulo 2. □

Observamos que o Exemplo 3.0.1 é uma aplicação deste Teorema.

Exemplo 3.1.1 (Passeio aleatório)**.** *A cadeia de Markov correspondente foi descrita no Exemplo 2.2.1. Os termos não nulos da matriz de transição são:*

$$p_{i,i+1} = p, \quad p_{i,i-1} = 1 - p,$$

onde $p \in (0,1)$. Nós podemos estudar a natureza de qualquer estado, pois este processo é invariante por translações para cima ou para baixo. Verifiquemos o caso do estado 0. Observamos que é impossível em uma quantidade impar de passos retornar a 0 *dado que se partiu de* 0*. Portanto, temos que $p_{00}(2n+1) = 0$. Em uma quantidade par de passos, ou seja,* $2n$ *passos, os únicos caminhos que retornan para* 0 *são aqueles que sobem n vezes e descem n vezes. Assim, temos uma combinação* *de* $2n$ *em* n. *Sendo que para cada vez que o caminho sobe temos a probabilidade* p *cada vez que desce a probabilidade* $1-p$, *segue que*

$$p_{00}(2n) = \frac{(2n)!}{n!n!} p^n (1-p)^n.$$

Pela fórmula equivalente de Stirling, $n! \approx n/e^n \sqrt{2\pi n}$ **[2]**, *a expressão acima é equivalente a*

$$\frac{[4p(1-p)]^n}{\sqrt{\pi n}}. \tag{3.4}$$

A natureza da série $\sum_{n=1}^{\infty} p_{00}(n)$ *(converge ou diverge) é o da série com termo geral (3.4). Se* $p \neq \frac{1}{2}$, *resulta* $[4p(1-p)]^n \leq 1$, *a série converge e se* $p = \frac{1}{2}$, *caso em que* $[4p(1-p)]^n = 1$, *a série diverge. Resumindo: os estados do passeio aleatório são transcientes se* $p \neq \frac{1}{2}$ *e recorrente se* $p = \frac{1}{2}$.

Consideremos os estados $i, j \in E$ tal que i é recorrente e acessível a partir de $j \in E$. Como i é acessível a partir de j existe uma probabilidade positiva de i ser visitado pelo menos uma vez e, partindo de i, por ser recorrente, o número médio de visitas é infinito. Portanto, partindo de j o número médio de visitas a i é infinito, e calculado por

$$E_j[N_i] = \sum_{n \geq 1} p_{ji}(n) = \infty.$$

Para o caso que i é transciente, de igual modo, o número médio de visitas para qualquer estado j é dado por

$$E_j[N_i] = \sum_{n \geq 1} p_{ji}(n) < \infty.$$

3.2 Estrutura da Matriz de Transição

Anteriormente, definimos os conceitos de recorrência e transciêcia, recorrência positiva e nula para estados simples. Porém, como veremos agora tais estados são propriedades de classe.

Teorema 3.2 (Recorrência é uma propriedade de classe). *Se* i *e* j *são comunicantes, ambos são recorrentes ou ambos são transcientes.*

Demonstração: Por definição, os estados i e j são comunicantes se existem inteiros M e N, tais que $p_{ij}(M) > 0$ e $p_{ji}(N) > 0$. Indo para j a partir de i em M passos, de j a partir de j em n passos e, por fim, de j para i em N passos, temos que $p_{ii}(M + n + N) \geq p_{ij}(M)p_{jj}(n)p_{ji}(N)$. De modo análogo $p_{ji}(N + n + M) \geq p_{ji}(N)p_{ii}(n)p_{ij}(M)$. Seja $\alpha = p_{ij}(M)p_{ji}(N) > 0$. Isto nos leva a $p_{ii}(M + n + N) \geq \alpha p_{jj}(n)$ e $p_{jj}(M + N + n) \geq \alpha p_{ii}(n)$. Estes fatos implicam que as séries

$$\sum_{n=0}^{\infty} p_{ii}(n) \quad \text{e} \quad \sum_{n=0}^{\infty} p_{jj}(n)$$

ambas convergem ou ambas divergem. O Teorema 3.1 conclui a demonstração. $\square$

Dados i positivo ou nulo recorrente, se j se comunica com i então j é positivo ou nulo recorrente, respectivamente. Deste modo, estas são propriedades de classe.

Os estados de uma cadeia de Markov irredutível são todos da mesma natureza, ou seja, transcientes, recorrentes positivos ou recorrentes nulos, é possível determinar se a cadeia é transciente, recorrente positiva ou recorrente nula, examinando apenas um dos estados e determinar a qual classe pertence, para tanto escolhe-se aquele que propicie maior facilidade no cálculo.

Pelo exposto, existem duas classes: as classes recorrentes e as classes transcientes. Denotaremos por T a classe transcientes, a classe T contém todos os estados transcientes e por R as classes recorrentes, cujos estados são todos recorrentes. O conjunto de classes recorrentes R pode ser composto por vários subconjuntos disjuntos de classes comunicantes $\{R_1, R_2, \ldots\}$. Os conjuntos $\{R_i\}$ são fechados. De fato se a cadeia vai de $i \in R_i$ para algum $j \in E$ com certeza voltará para i, pois i e j são comunicantes, portanto $j \in R_i$. A figura 3.2, mostra a estrutura de comunicação da matriz de transição $\mathbf{P}$.

3.3 Recorrência e Medidas Invariantes

A medida invariante é uma ferramenta utilizada para determinação da natureza dos estados de uma cadeia de Markov, pois é uma condição necessária. Sua noção se estende à distribuição estacionária.

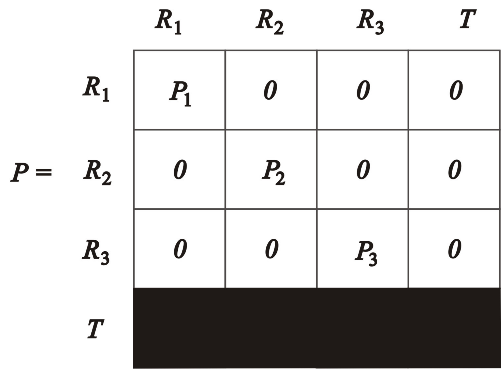

Figura 3.2:

Definição 3.2.1 (Medida Invariante)**.** *Um vetor não nulo* $\mathbf{x} = \{x_i\}_{i \in E}$ *é denominado uma medida invariante da matriz estocástica* $\mathbf{P} = \{p_{ij}\}_{i,j \in E}$*, se para todo* $i \in E$*,*

$$x_i \in [0, \infty) \tag{3.5}$$

e

$$x_i = \sum_{j \in E} x_j p_{ij}. \tag{3.6}$$

Em notação abreviada, $0 \leq \mathbf{x} < \infty$ e $\mathbf{x}^T \mathbf{P} = \mathbf{x}^T$.

Definamos para todo $i \in E$ o seguinte

$$x_i = E_0 \left[\sum_{n \geq 1} \mathbf{1}_{\{X_n = i\}} \mathbf{1}_{\{n \leq T_0\}} \right], \tag{3.7}$$

onde T_0 é o tempo de retorno ao estado 0.

Para $n \in [1, T_0]$, onde T_0 é o tempo de retorno ao estado 0, $X_n = 0$ se, e somente se, $n = T_0$. Portanto,

$$\begin{aligned} x_0 &= E_0 \left[\sum_{n \geq 1} \mathbf{1}_{\{X_n = 0\}} \mathbf{1}_{\{n \leq T_0\}} \right] \\ &= E_0[1] = 1 \qquad (3.8) \\ & \qquad\qquad\qquad (3.9) \end{aligned}$$

Calculemos $\sum_{i \in E} x_i$. Pela linearidade de E_0, basta calcular

$$\begin{aligned}\sum_{i \in E} \sum_{n \geq 1} \mathbf{1}_{\{X_n = i\}} \mathbf{1}_{\{n \leq T_0\}} &= \sum_{n \geq 1} \left\{ \sum_{i \in E} \mathbf{1}_{\{X_n = i\}} \right\} \mathbf{1}_{\{n \leq T_0\}} \\ &= \sum_{n \geq 1} \mathbf{1}_{\{n \leq T_0\}} \\ &= T_0.\end{aligned}$$

Consequentemente,

$$\sum_{i \in E} x_i = E_0[T_0]. \tag{3.10}$$

Para simplificar a demonstração do Teorema 3.4 abaixo, adotaremos a seguinte notação

$${}_0p_{0i}(n) \stackrel{\text{def}}{=} E_0[\mathbf{1}_{\{X_n = i\}} \mathbf{1}_{\{n \leq T_0\}}] = P_0(X_1 \neq 0, \ldots, X_{n-1} \neq 0, X_n = i). \tag{3.11}$$

É claro que

$$x_i = \sum_{n \geq 1} {}_0p_{0i}(n). \tag{3.12}$$

Lema 3.3. *Para $n \geq 2$ temos que*

$${}_0p_{0i}(n) = \sum_{j \neq 0} {}_0p_{0j}(n-1) p_{ji}.$$

Demonstração: Seja a união disjunta dos eventos $\{X_{n-1} \neq 0\} = \dot{\bigcup}_{j \neq 0} \{X_{n-1} = j\}$. Por definição

$$\begin{aligned}{}_0p_{0i}(n) &= \mathbb{P}_0(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-1} \neq 0, X_n = i) \\ &= \mathbb{P}_0(X_1 \neq 0, X_2 \neq 0, \ldots, \dot{\cup} \{X_{n-1} = j\}, X_n = i).\end{aligned}$$

Disto

$${}_0p_{0i}(n) = \sum_{j \neq 0} \mathbb{P}_0(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_n = i). \tag{3.13}$$

Analisaremos agora o termo geral da série acima. Por definição,

$$\begin{aligned}
& \mathbb{P}_0(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_n = i) \\
= & \ \mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_n = i | X_0 = 0) \\
= & \ \frac{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_n = i, X_0 = 0)}{\mathbb{P}(X_0 = 0)}.
\end{aligned}$$

Multiplicando a expressão acima por

$$\frac{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_0 = 0)}{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_0 = 0)},$$

e procedendo a reordenação dos termos do numerador e do denominador obtemos

$$\begin{aligned}
= & \ \frac{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_n = i, X_0 = 0)}{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_0 = 0)} \times \\
& \ \frac{\mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j, X_0 = 0)}{\mathbb{P}(X_0 = 0)}.
\end{aligned}$$

Logo, por definição de probabilidade condicional,

$$\begin{aligned}
= & \ \mathbb{P}(X_n = i | X_{n-1} = j, X_2 \neq 0, \ldots, X_2 \neq 0, X_1 \neq 0, X_0 = 0) \\
\times & \ \mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j | X_0 = 0).
\end{aligned}$$

De acordo com a propriedade de Markov

$$\begin{aligned}
& \mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-1} = j, X_n = i) \\
= & \ \mathbb{P}(X_n = i | X_{n-1} = j) \mathbb{P}(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-2} \neq 0, X_{n-1} = j | X_0 = 0) \\
= & \ {}_0p_{0j}(n-1) p_{ji}.
\end{aligned}$$

Assim, da série (3.13) resulta que

$${}_0p_{0i}(n) = \sum_{j \neq 0} {}_0p_{ij}(n-1) p_{ji}.$$

$\square$

Teorema 3.4 (Forma Regenerativa da Medida Invariante)**.** *Seja* $\mathbf{P}$ *uma matriz de transição de uma Cadeia de Markov Homogênea recorrente e irredutível* $\{X_n\}_{n\geq 0}$. *Seja 0 um estado arbitrário e* T_0 *o tempo de retorno para o estado 0. Para todo* $i \in E$, *o vetor*

$$x_i \in (0, \infty), \tag{3.14}$$

e $\mathbf{x}$ *é uma medida invariante de* $\mathbf{P}$.

Demonstração: Inicialmente observamos que

$$\begin{aligned} {}_0p_{0i}(1) &= E_0\left[\sum_{n\geq 1} \mathbf{1}_{\{X_1=i\}} \mathbf{1}_{\{1\leq T_0\}}\right] \\ &= \mathbb{P}_0(X_1 = i) = \mathbb{P}(X_1 = i | X_0 = 0) \\ &= p_{0i}. \end{aligned}$$

Agora, para todo $n \geq 2$, o Lema 3.3 mostra que

$$ {}_0p_{0i}(n) = \sum_{j\neq 0} {}_0p_{0j}(n-1)p_{ji}. \tag{3.15}$$

Somando as igualdades acima e tendo em conta (3.12), obtém-se

$$x_i = p_{0i} + \sum_{j\neq 0} x_j p_{ji},$$

que é (3.6), já que $x_0 = 1$. De forma explicita

$$
\begin{aligned}
x_i &= \sum_{n\geq 1} {}_0p_{0i}(n) \\
&= {}_0p_{0i}(1) + \sum_{n\geq 2} {}_0p_{0i}(n) \\
&= p_{0i} + \sum_{n\geq 2}\left(\sum_{j\neq 0} {}_0p_{0j}(n-1)p_{ji}\right) \\
&= p_{0i} + \sum_{j\neq 0}\left(\sum_{n\geq 2} {}_0p_{0j}(n-1)p_{ji}\right) \\
&= p_{0i} + \sum_{j\neq 0} x_j p_{ji} \\
&= x_0 p_{0i} + \sum_{j\neq 0} x_j p_{ji} \\
&= \sum_{j\in E} x_j p_{ji}.
\end{aligned}
$$

Em seguida mostraremos que $x_i > 0$ para todo $i \in E$. Realmente, por iteração de (3.6), encontra-se $\mathbf{x}^T = \mathbf{x}^T\mathbf{P}^n$. Já que $x_0 = 1$,

$$x_i = \sum_{j\in E} x_j p_{ji}(n) = p_{0i}(n) + \sum_{j\neq 0} x_j p_{ji}(n).$$

Se x_i fosse nulo para algum $i \in E$, $i \neq 0$, a última igualdade implicaria que $p_{0i}(n) = 0$ para todo $n \geq 0$. Significando que 0 e i não são comunicantes, em contradição a irredutibilidade pressuposta.

Continuando, mostraremos que $x_i < \infty$ para todo $i \in E$. Como antes, encontra-se que

$$1 = x_0 = \sum_{j\in E} x_j p_{j0}(n)$$

para todo $n \geq 1$ e, portanto, se $x_i = \infty$ para algum i, necessariamente $p_{i0}(n) = 0$ para todo $n \geq 1$, o que também contradiz a irredutibilidade. $\square$

Teorema 3.5 (Unicidade da Medida Invariante)**.** *A medida invariante de uma matriz estocástica recorrente irredutível é única a menos de um fator multiplicativo.*

Demonstração: Na demonstração do Teorema 3.4 foi mostrado que para uma medida invariante y de uma cadeia irredutível, $y_i > 0$ para todo $i \in E$. Portanto, definiremos, para todo $i, j \in E$, a matriz $\mathbf{Q}$ por

$$q_{ji} = \frac{y_i}{y_j} p_{ij}. \tag{3.16}$$

Esta é uma matriz de transição, pois

$$\sum_{i \in E} q_{ji} = \sum_{i \in E} \frac{y_i}{y_j} p_{ij} = \frac{1}{y_j} \sum_{i \in E} y_i p_{ij} = \frac{y_j}{y_j} = 1.$$

O termo geral de $\mathbf{Q}^n$ é

$$q_{ji}(n) = \frac{y_i}{y_j} p_{ij}(n). \tag{3.17}$$

Supondo (3.17) verdadeira para n, vemos de modo análogo ao Lema 3.3 que

$$\begin{aligned} q_{ji}(n+1) &= \sum_{k \in E} q_{jk} q_{ki}(n) = \sum_{k \in E} \frac{y_k}{y_j} p_{kj} \frac{y_i}{y_k} p_{ik}(n) \\ &= \frac{y_i}{y_j} \sum_{k \in E} p_{ik}(n) p_{kj} = \frac{y_i}{y_j} p_{ij}(n+1), \end{aligned}$$

e (3.17) segue por indução, para todo $n \geq 1$.

Claramente, $\mathbf{Q}$ é irredutível, desde que $\mathbf{P}$ é irredutível, pois $q_{ji}(n) > 0$ se, e somente se, $p_{ij}(n) > 0$, tendo em vista (3.17). Tambem é verdade que $p_{ii}(n) = q_{ii}(n)$ e, portanto, $\sum_{n \geq 0} q_{ii}(n) = \sum_{n \geq 0} p_{ii}(n)$. Isto garante que $\mathbf{Q}$ é recorrente, pelo critério da matriz potência. Denominemos por $g_{ji}(n)$ a probabilidade relativa à cadeia determinada pela matriz de transição $\mathbf{Q}$, de retorno ao estado i pela primeira vez no passo n, quando partindo de j. De maneira explicita

$$g_{ji}(n) = \mathbf{Q}_j(X_1 \neq 0, X_2 \neq 0, \ldots, X_{n-1} \neq 0, X_n = i).$$

A análise do primeiro passo dá

$$g_{i0}(n+1) = \sum_{j \neq 0} q_{ij} g_{j0}(n),$$

o que nada mais é que a aplicação do Lema 3.3. Usando (3.17) obtemos

$$y_i g_{i0}(n+1) = \sum_{j \neq 0} (y_j g_{j0}(n)) p_{ji}.$$

Voltemos a ${}_0p_{0i}(n+1) = \sum_{j\neq 0} {}_0p_{0j}(n)p_{ji}$ ou, equivalentemente,

$$y_{00}p_{0i}(n+1) = \sum_{j\neq 0}(y_{00}p_{0j}(n))p_{ji}.$$

Vê-se, portanto, que as sequências $\{y_{00}p_{0i}(n)\}$ e $\{y_i g_{0i}(n)\}$ satisfazem a mesma equação de recorrência. Seus primeiros termos ($n = 1$), são respectivamente, $y_{00}p_{0i}(1) = y_0 p_{0i}$ e $y_i y_{i0}(1) = y_i q_{i0}$, que são iguais em vista de (3.17). Portanto, para todo $n \geq 1$,

$$_0p_{0i}(n) = \frac{y_i}{y_0} g_{i0}(n).$$

Somando com respeito a $n \geq 1$ e usando $\sum_{n\geq 1} g_{i0}(n) = 1$ ($\mathbf{Q}$ é recorrente), obtemos o resultado anunciado $x_i = \dfrac{y_i}{y_0}$. □

A igualdade (3.10) e a definição de recorrência positiva são necessárias no seguinte Teorema.

Teorema 3.6 (Positiva $\times$ nula recorrência)**.** *Uma Cadeia de Markov Homogênea recorrente irredutível é positiva recorrente se e somente se sua medida invariante* $\mathbf{x}$ *satisfaz*

$$\sum_{i\in E} x_i < \infty. \tag{3.18}$$

Demonstração: Basta observar que $\sum_{i\in E} x_i = E_0[T_0]$ e que recorrência positiva é uma propriedade de classe. □

Observação 3.6.1. *Uma CMH pode ser irredutível e possuir uma medida invariante, e ainda não ser recorrente. O exemplo mais simples é o passeio aleatório não simétrico, Exemplo 3.1.1, que foi mostrado ser transciente, e que admite* $x_i \equiv 1$ *para medida invariante.*

3.4 Recorrência Positiva

3.4.1 Critério da Distribuição Estacionária

Na seção anterior assumimos que a CMH irredutível era recorrente. Com esta hipótese encontramos o critério para recorrência positiva, ver Teorema 3.6. Porém o contrário não foi demonstrado. Nesta seção usando distribuição estacionária encontramos um critério mais forte para recorrência positiva, sem a necessidade de recorrência

Teorema 3.7 (Critério da distribuição estacionária)**.** *Uma cadeia de Markov homogênea irredutível é recorrente positiva se, e somente se, existe uma distribuição estacionária. Além disso, a distribuição estacionária* π *é, quando existe, única e* $\pi > 0$.

Demonstração: A parte direta decorre dos Teoremas 3.4 e 3.6. Para o lado oposto, assumimos a existência de uma distribuição estacionária π. Iterando $\pi^T = \pi^T \mathbf{P}^n$, ou seja, para todo $i \in E$,

$$\pi(i) = \lim_{n\uparrow\infty} \sum_{j\in E} \pi(j) p_{ji}(n).$$

Se a cadeia fosse transciente, então tendo em conta o critério da matriz potência e as observações seguintes do Exemplo 3.1.1, para todos os estados i, j,

$$\lim_{n\uparrow\infty} p_{ji}(n) = 0.$$

Dado que $p_{ji}(n)$ é uniformemente limitado por 1 em j e n e $\sum_{i\in E} \pi(i) = 1$, pelo teorema das séries convergentes (ver Teorema 3.2 apêndice),

$$\pi(i) = \lim_{n\uparrow\infty} \sum_{j\in E} \pi(j) p_{ji}(n) = \sum_{j\in E} \pi(j) \left(\lim_{n\uparrow\infty} p_{ji}(n) \right) = 0.$$

Isto contradiz a hipótese que π é uma distribuição estacionária. Portanto a cadeia é recorrente e pelo Teorema 3.6 é recorrente positiva.

A distribuição estacionária π de uma cadeia irredutível recorrente positiva é única, pelo Teorema 3.5, e o fato de não existir escolha para o fator multiplicativo, além de 1. Lembramos também que $\pi(i) > 0$ para todo $i \in E$ (ver Teorema 3.4). $\square$

Teorema 3.8 (Significado do Tempo de Retorno)**.** *Seja π a única distribuição estacionária de uma CMH irredutível recorrente positiva e T_i o tempo de retorno para o estado i. Então,*

$$\pi(i)E_i[T_i] = 1. \tag{3.19}$$

Demonstração: Esta igualdade é consequência direta da expressão (3.7) para a medida invariante. Pois a distribuição π é obtida é pela normalização de x, isto é, para todo $i \in E$,

$$\pi(i) = \frac{x_i}{\sum_{j\in E} x_j}.$$

Em particular, para $i = 0$, usamos (3.8) e (3.10) para obter

$$\pi(0) = \frac{x_0}{\sum_{j\in E} x_j} = \frac{1}{E_0[T_0]}.$$

Desde que o estado 0 não desempenha papel especial na análise, (3.19) é verdadeira para todo $i \in E$. $\square$

Para o caso em que o espaço estado é finito, esta situação é simplificada.

Teorema 3.9 (Espaço estado finito e Recorrência Positiva)**.** *Uma Cadeia de Markov irredutível com espaço estado finito é recorrente positiva.*

Demonstração: Devemos demonstrar em primeiro lugar a recorrência. Se a cadeia fosse transciente então, pelo critério da potência da matriz e das observações seguintes, para todo $i,\ j \in E$,

$$\sum_{n\in E} p_{ij}(n) < \infty.$$

Como o espaço estado E é finito

$$\sum_{j\in E}\sum_{n\geq 0} p_{ij}(n) < \infty.$$

O último somatório é igual a

$$\sum_{n\geq 0}\sum_{j\in E} p_{ij}(n) = \sum_{n\geq 0} 1 = \infty.$$

daí, a contradição. Portanto, a cadeia é recorrente. Pelo Teorema 3.4 a cadeia tem uma medida invariante $\mathbf{x}$. Desde que E seja finito, $\sum_{i\in R} x_i < \infty$. Consequentemente a cadeia é recorrente positiva, pelo Teorema 3.7. $\square$

Exemplo 3.9.1 (Corrida do sucesso)**.** *Vamos utilizar o Exemplo 3.1.1 para aplicação do Critério da distribuição estacionária. Escrevemos* $q_i = 1 - p_i$*. A expressão* $\mathbf{x}^T = \mathbf{x}^T$ *será*

$$x_0 = p_0 x_0 + p_1 x_1 + p_2 x_2 + \ldots, \tag{3.20}$$

e, para $i \geq 1$*,*

$$x_i = q_{i-1} x_{i-1}.$$

Olhando agora para a última expressão, quando $i \geq 1$*, vemos que*

$$x_i = (q_0 q_1 \ldots q_{n1}) x_0. \tag{3.21}$$

Excluamos o caso em que $x_0 \leq 0$ *pois isto implicaria que* x *negativo ou nulo. Substituindo (3.21) em (3.20), reescrevemos (3.20) como*

$$x_0 = p_0 x_0 + p_1 q_0 x_0 + p_2 q_0 q_1 x_0 + \ldots\ .$$

Desde modo

$$1 = p_0 + q_0 p_1 + q_0 q_1 p_2 + \ldots\ .$$

Já que

$$q_0 q_1 \ldots q_{n-1} p_n = q_0 q_1 \ldots q_{n-1} - q_0 q_1 \ldots q_n,$$

obtemos

$$1 = 1q_0 + (q_0 - q_0 q_1) + (q_0 q_1 - q_0 q_1 q_2) + (q_0 q_1 q_2 - q_0 q_1 q_2 q_3) + \ldots\ .$$

Fazendo a soma telescópica resulta que

$$1 = 1 + \prod_{i=0}^{\infty} q_i = 0$$

e, portanto,

$$\prod_{i=0}^{\infty} q_i = 0. \tag{3.22}$$

Isto significa que $\mathbf{x}$ *satisfaz (3.20) se, e somente, se* $\prod_{i=0}^{\infty} q_i = 0$. *Sendo que* $q_i = 1 - p_i$ *e* $p_i \in (0,1)$, *o critério de convergência para infinitos produtos (ver teorema 3 do apêndice) assegura que (3.22) é equivalente a*

$$\sum_{i=0}^{\infty} p_i = \infty.$$

Do que foi mostrado acima, a condição necessária e suficiente para a existência de uma medida invariante é a divergência da série $\sum_{i=0}^{\infty} p_i = \infty$.

Relembramos que a existência de uma medida invariante não implica em recorrência (ver observação 3.6.1). No entanto a existência de uma distribuição estacionária implica em recorrência pelo critério da distribuição estacionária.

Pela condição (3.22), existe uma medida invariante, e esta medida tem massa finita ($\sum_{i=0}^{\infty} x_i < \infty$ *(ver Teorema 3.4)) se, e somente se,*

$$1 + \sum_{n=0}^{\infty} \left(\prod_{i=0}^{n-1} q_i \right) < \infty.$$

Usando o fato que

$$\sum_{n=0}^{\infty} x_i = \left(1 + \sum_{n=0}^{\infty} \left(\prod_{i=0}^{n-1} q_i \right) \right) x_0,$$

por (3.21). A distribuição estacionária definida a partir de $\mathbf{x}$ *é*

$$\pi(i) = \frac{x_i}{\sum_{i=0}^{\infty} x_i}.$$

Desta forma concluímos que a distribuição estacionária é dada por

$$\pi(0) = \left(1 + \sum_{i=0}^{\infty} \left(\prod_{i=0}^{n-1} q_i \right) \right)^{-1}$$

e

$$\pi(i) = \left(\prod_{i=0}^{n-1} q_i\right)\pi(0).$$

3.5 Período antes da Absorção

3.5.1 Permanência passageira Infinita

Na seção em que tratamos de estrutura da matriz de transição, temos visto pelo Teorema 3.2, que os estados de uma CMH podem ser classificados em recorrentes e transcientes, que a comunicação entre os estados é uma propriedade de classe, e ainda, que o espaço de estados pode ser decomposto como segue

$$E = T + \sum_j R_j,$$

onde os R_i's representam as classes recorrentes disjuntas e T é o conjunto de estados transcientes.

Como antes obtivemos resultados que diziam respeito às CMH irredutíveis e recorrentes, no que segue, estudaremos as cadeias de Markov que apresentam várias classes de comunicação.

A matriz pode ser particionada em blocos de estados com as mesmas características, queremos dizer, estados recorrentes e estados transcientes, como mostramos a seguir.

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cccc} R_1 & R_2 & \cdots & T \end{array} \\ \begin{array}{c} R_1 \\ R_2 \\ \vdots \\ T \end{array} & \begin{pmatrix} \mathbf{P}_1 & 0 & 0 & 0 \\ 0 & \mathbf{P}_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ \mathbf{B}(1) & \mathbf{B}(2) & \cdots & \mathbf{Q} \end{pmatrix} \end{array}, \tag{3.23}$$

ou em notação compacta,

$$\mathbf{P} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{B} & \mathbf{Q} \end{pmatrix}. \tag{3.24}$$

A matriz $\mathbf{P}$ mostra que um estado pertencente a uma determinada classe recorrente não pode ir a qualquer outro estado que não seja da mesma classe, ou seja, uma classe recorrente é fechada.

Uma pergunta clássica é: Qual a probabilidade de um estado ser absorvido por uma classe recorrente quando parte de uma classe transciente? No capítulo 2 quando tratamos da análise do primeiro passo, o exemplo o lobo e o cordeiro, trata desta questão, lá são atribuídas condições limitantes, o que permite encontrar uma solução única, dado que o espaço de estado é finito. Nosso objetivo agora é encontrar uma matriz algébrica que seja solução do problema.

Uma outra questão é a de que um conjunto de estados transciente nunca seja absorvido por um conjunto de estados recorrentes. Consideremos a obtenção da distribuição do tempo e a probabilidade de absorção por um estado recorrente.

Seja A um subconjunto do estado de espaço E (tipicamente o conjunto de estados transcientes, mas não necessariamente). O objetivo agora é calcular a probabilidade de qualquer estado inicial $i \in A$ de permanecer para sempre em A. Em notação

$$v(i) \quad = \quad \mathbb{P}_i(X_r \in A; \ r \geq 0).$$

Definamos

$$v_n(i) = \mathbb{P}_i(X_1 \in A, \dots, X_n \in A).$$

Pela continuidade sequencial monótona, temos

$$\lim_{n \longrightarrow \infty} v_n(i) = v(i).$$

De fato, usando que

$$\{X_0 \in A, \dots, X_n \in A\} \supset \{X_0 \in A, \dots, X_n \in A, X_{n+1} \in A\},$$

fazemos os seguintes cálculos

$$
\begin{aligned}
v(i) &= \mathbb{P}_i(X_n \in A : n \geq 0) \\
&= \mathbb{P}(X_n \in A : n \geq 0 | X_0 = i) \\
&= \mathbb{P}(X_n \in A : n \geq 0 | X_0 = i) \\
&= \frac{\mathbb{P}(X_n \in A : n \geq 0, X_0 = i)}{P(X_0 = i)} \\
&= \frac{\mathbb{P}(\bigcap_{n=0}^{\infty} \{X_1 \in A, ..., X_n \in A\}, X_0 = i)}{P(X_0 = i)} \\
&= \lim_{n \longrightarrow \infty} \frac{\mathbb{P}(\{X_1 \in A, X_2 \in A, \dots, X_n \in A\}, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \lim_{n \longrightarrow \infty} \mathbb{P}_i(\{X_1 \in A, X_2 \in A, \dots, X_n \in A\}) \\
&= \lim_{n \longrightarrow \infty} v_n(i).
\end{aligned}
$$

Todavia, para $j \in A$,

$$\mathbb{P}_i(X_1 \in A, \dots, X_{n-1} \in A, X_n = j) = \sum_{i_1 \in A} \dots \sum_{i_{n-1} \in A} p_{ii_1} \dots p_{i_{n-1}j}$$

que é o termo geral $q_{ij}(n)$ da n-ésima iteração de $\mathbf{Q}$, restrição de $\mathbf{P}$ ao conjunto A. Portanto,

$$v_n(i) = \sum_{j \in A} q_{ij}(n).$$

A última igualdade resulta em

$$v_n = \mathbf{Q}^n \mathbf{1}_A, \tag{3.25}$$

onde $\mathbf{1}_A$ é um vetor coluna indexado por A com todas as entradas igual a 1. Desta igualdade obtemos

$$v_{n+1} = \mathbf{Q}^{n+1} \mathbf{1}_A.$$

Daí

$$v_{n+1} = \mathbf{Q} v_n.$$

Pela convergência dominada $\mathbf{v} = \mathbf{Q}\mathbf{v}$. Além disso, $\mathbf{0}_A \leq \mathbf{v} \leq \mathbf{1}_A$, onde $\mathbf{0}_A$ é um vetor coluna indexado por com todas as entradas iguais a 0.

Teorema 3.10 (Permanência Passageira Infinita)**.** *O vetor* $\mathbf{v}$ *é solução maximal de*

$$\mathbf{v} = \mathbf{Q}\mathbf{v} \tag{3.26}$$

$$\mathbf{0}_A \leq \mathbf{v} \leq \mathbf{1}_A \tag{3.27}$$

Além disso, ou $\mathbf{v} = \mathbf{0}_A$ *ou* $sup_{i\in A} v(i) = 1$.

Demonstração: Para provar a maximalidade consideremos o vetor $\mathbf{u}$ indexado por A tal que $\mathbf{u} = \mathbf{Q}\mathbf{u}$ e $\mathbf{0}_A \leq \mathbf{u} \leq \mathbf{1}_A$. Iterando $\mathbf{u} = \mathbf{Q}\mathbf{u}$, obtemos $\mathbf{u} = \mathbf{Q}^n\mathbf{u}$, e $\mathbf{u} \leq \mathbf{1}_A$ implica que $\mathbf{Q}^n\mathbf{u} \leq \mathbf{Q}^n\mathbf{1}_A = \mathbf{v}_n$. Portanto $\mathbf{u} \leq \mathbf{v}_n$, que dá $\mathbf{u} \leq \mathbf{v}$ quando passamos o limite. Para a última afirmação, seja $c = \sup_{i\in A} v(i)$. De $\mathbf{v} \leq c\mathbf{1}_A$ obtemos $\mathbf{v} \leq c\mathbf{v}_n$ como acima, passando o limite $\mathbf{v} \leq c\mathbf{v}$. Isto implica em $\mathbf{v} = \mathbf{0}_A$ ou $c = 1$. □

A equação $\mathbf{v} = \mathbf{Q}\mathbf{v}$, com a restrição $\mathbf{0}_A \leq \mathbf{v} \leq \mathbf{1}_A$, pode ser expressa por

$$v(i) = \sum_{j ß A} p_{ii} v(j)(i \in A). \tag{3.28}$$

A análise do primeiro passo mostra que (3.28) é uma condição necessária para que $i \in A$ para sempre.

A seguir utilizaremos o estudo acima para mostrar uma CMH transciente.

Exemplo 3.10.1 (Caminho Aleatório Refletido em 0)**.**

A matriz de transição de um caminho aleatório sobre $\mathbb{N}$ com um obstáculo refletido em 0, é

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & & & & \\ q & 0 & p & & & \\ & q & 0 & p & & \\ & & q & 0 & p & \\ & & & . & . & . \end{pmatrix},$$

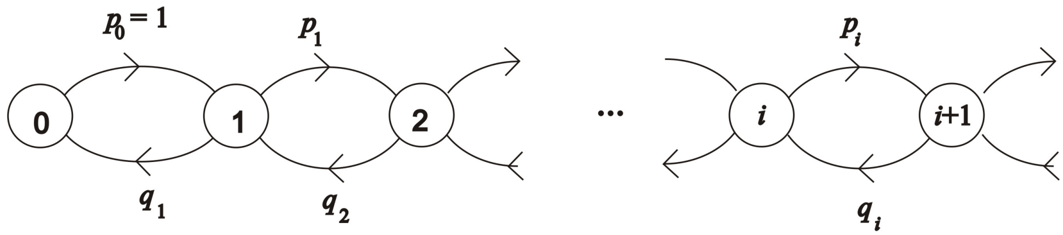

Figura 3.3: Caminho aleatório refletido

onde $p \in (0,1)$, tal cadeia é claramente irredutível (ver figura 3.3).

Intuitivamente, se $p > q$, existe um pulo para a direita, e espera-se que a cadeia seja transciente. Isto será provado formalmente quando demonstrarmos que a probabilidade $v(i)$ de nunca visitar o estado 0 quando partir do estado $i \geq 1$ é estritamente positiva. Na sequência aplicaremos o Teorema 3.10 com $A = \mathbb{N} - \{0\}$. Vamos encontrar uma solução geral para $\mathbf{u} = \mathbf{Q}\mathbf{u}$. Por (2.19) sabemos que

$$u(i) = qu(i-1) + pu(i+1).$$

Logo,

$$\begin{aligned} u(1) &= \sum_{j \in A} p_{1j} u(j) = pu(2) \\ u(2) &= \sum_{j \in A} p_{2j} u(j) = qu(1) + pu(3) \\ u(3) &= \sum_{j \in A} p_{3j} u(j) = qu(2) + pu(3) \\ &\vdots \end{aligned}$$

e sua solução geral é

$$u(i) = u(1) \sum_{j=0}^{i-1} (\frac{q}{p})^j.$$

O maior valor de $u(1)$ em relação ao limite de $u(i) \in [0,1]$ é $u(1) = 1 - (\frac{q}{p})$. A solução $v(i)$ é obtida fixando este valor, e portanto

$$v(i) = 1 - (\frac{q}{p})^i. \tag{3.29}$$

Observação 3.10.1 (Conjunto Transiente Finito)**.** *Se um conjunto de estados transcientes T é finito, a probabilidade de infinita permanência em T é nula. Isto é intuitivamente óbvio, fato que pode ser deduzido da aplicação da equação (3.28) com $A = T$. Realmente, com a notação de (3.24),*

$$\mathbf{P}^n = \begin{pmatrix} \mathbf{D}^n & \mathbf{0} \\ \mathbf{B}_n & \mathbf{Q}^n \end{pmatrix}. \tag{3.30}$$

Portanto o termo geral de $\mathbf{Q}^n$ é $p_{ij}(n)$, mas, sabe-se que para qualquer estado transiente j, $\lim_{n\uparrow\infty} p_{ij}(n) = 0$ para todo $i \in E$. Portanto se T é finito,

$$v(i) = \lim_{n\uparrow\infty} \sum_{j\in T} p_{ij}(n) = 0.$$

3.5.2 Tempo para absorção

Dada uma CMH e seja τ tempo de saída do conjunto transciente T. Nesta seção vamos calcular a distribuição de τ. O Teorema 3.10 afirma que $\mathbf{v} = \{v(i)\}_{i\in T}$, onde

$$v(i) = \mathbb{P}_i(\tau = \infty),$$

é a maior solução de $\mathbf{v} = \mathbf{Q}\mathbf{v}$ sujeita à restrição $\mathbf{0}_T \leq \mathbf{v} \leq \mathbf{1}_T$, onde $\mathbf{Q}$ é a restrição de $\mathbf{P}$ para o conjunto transciente T. De forma explicita,

$$v(i) = \mathbb{P}_i(\tau = \infty) = \mathbb{P}(X_r \in T; r \geq 0 | X_0 = i).$$

O cálculo da distribuição de τ é facilitado com a utilização da identidade abaixo

$$\mathbb{P}_i(\tau = n) = \mathbb{P}_i(\tau \geq n) - \mathbb{P}_i(\tau \geq n+1),$$

a qual se demonstra da seguinte maneira. Sendo

$$\{\tau \geq n\} = \{\tau = n\} \dot{\cup} \{\tau > n\}$$

segue que

$$\begin{aligned} \mathbb{P}_i(\tau \geq n) &= \mathbb{P}_i(\tau = n) + \mathbb{P}_i(\tau > n) \\ &= \mathbb{P}_i(\tau = n) + \mathbb{P}_i(\tau \geq n+1). \end{aligned}$$

De onde

$$\mathbb{P}_i(\tau = n) = \mathbb{P}_i(\tau \geq n) - \mathbb{P}_i(\tau \geq n+1).$$

Observamos que

$$\{\tau \geq n\} = \{X_{n-1} \in T\} \quad \text{e} \quad \{\tau \geq n+1\} = \{X_n \in T\}.$$

Logo

$$\begin{aligned} \mathbb{P}_i(\tau = n) &= \mathbb{P}_i(X_{n-1} \in T) - \mathbb{P}_i(X_n \in T) \\ &= \mathbb{P}(X_{n-1} \in T | X_0 = i) - \mathbb{P}(X_n \in T | X_0 = i). \end{aligned}$$

Escrevendo $\{X_{n-1} \in T\} = \dot{\bigcup_{j \in T}} \{X_{n-1} = j\}$ e $\{X_n \in T\} = \dot{\bigcup_{j \in T}} \{X_n = j\}$, deduzimos o seguinte

$$\begin{aligned} \mathbb{P}(\tau = n) &= \mathbb{P}(\dot{\bigcup_{j \in T}} \{X_{n-1} = j\} | X_0 = i) - \mathbb{P}(\dot{\bigcup_{j \in T}} \{X_n = j\} | X_0 = i) \\ &= \sum_{j \in T} \mathbb{P}(X_{n-1} = j | X_0 = i) - \sum_{j \in T} \mathbb{P}(X_n = j | X_0 = i) \\ &= \sum_{j \in T} p_{ij}(n-1) - \sum_{j \in T} p_{ij}(n) \\ &= \sum_{j \in T} (p_{ij}(n-1) - p_{ij}(n)). \end{aligned} \tag{3.31}$$

Sendo $p_{ij}(n)$, para $i, j \in T$, o termo geral de $\mathbf{Q}^n$, tem-se o resultado a seguir.

Teorema 3.11 (Distribuição do Tempo Para Absorção)**.**

$$\mathbb{P}_i(\tau = n) = \{(\mathbf{Q}^{n-1} - \mathbf{Q}^n)\mathbf{1}_T\}_i.$$

Em particular, se $\mathbb{P}_i(\tau = \infty) = 0$

$$\mathbb{P}_i(\tau > n) = \{\mathbf{Q}^n \mathbf{1}_\tau\}_i. \tag{3.32}$$

Demonstração: A primeira igualdade vem da equação (3.31)

$$\mathbb{P}_i(\tau = n) = \{(\mathbf{Q}^{n-1} - \mathbf{Q}^n)1_T\}_i.$$

Agora, de $\{\tau > n\} = \{\tau \geq n+1\} = \{X_n \in T\}$ conseguimos

$$\mathbb{P}_i(\tau > n) = \mathbb{P}(X_n \in T) = \sum_{j \in T} p_{ij}(n) = \{\mathbf{Q}^n \mathbf{1}_T\}_i.$$

$\square$

3.6 Absorção

3.6.1 Matriz Fundamental

Um estado transciente pode permanecer para sempre em um conjunto transciente ou ser absorvido por um conjunto recorrente. Para calcular a probabilidade de um estado transciente ser absorvido por uma dada classe recorrente R_j, quando parte de um estado inicial i usaremos, o conceito de matriz fundamental que está vinculado à matriz potência.

Analisaremos apenas o caso em que os estados $i, j \in T$, e verificaremos o número de visitas ao estado $j \in T$ antes de ser absorvido por $R = R_1 \cup R_2 \cup \ldots$, quando parte do estado $i \in T$.

Quando a cadeia tem início em algum estado i, o número médio de visitas a um estado arbitrário j é dado por

$$E_i\left[\sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j\}}\right] = \sum_{n=0}^{\infty} p_{ij}(n). \tag{3.33}$$

De forma clara,

$$E_i\left[\sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j\}}\right] = \sum_{n=0}^{\infty} E_i\left[\mathbf{1}_{\{X_n=j\}}\right] = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n=j) = \sum_{n=0}^{\infty} p_{ij}(n).$$

Este é o (i, j) elemento da matriz potência

$$\mathbf{G} = \sum_{n=0}^{\infty} \mathbf{P}^n. \tag{3.34}$$

Os estados transcientes i e j, requerem uma melhor análise, pois nas outras situações, a expressão (3.33) é nula ou infinita. Exemplificando, se i e j pertence a mesma classe recorrente, o número de visitas a j a partir de i é infinito, será zero se i e j pertencem á classes recorrentes diferentes , ou se i é um estado recorrente e j é um estado transciente. Se i é transciente e j é recorrente, o termo correspondente é infinito se j é acessível de i, e zero se j não é acessível de j.

Para avaliarmos (3.33) para $i, t \in T$. Usamos a representação (3.24) da matriz de transição para obter

$$\mathbf{G} = \begin{pmatrix} \mathbf{E} & \mathbf{O} \\ \mathbf{F} & \mathbf{S} \end{pmatrix}, \tag{3.35}$$

onde

$$\mathbf{S} = \sum_{n=0}^{\infty} \mathbf{Q}^n. \tag{3.36}$$

Definição 3.11.1 (Cadeias Absorventes e Matriz Fundamental)**.** *Uma CMH com pelo menos um estado transciente e um estado recorrente é denominada absorvente e a matriz* $\mathbf{S}$ *é sua matriz fundamental.*

Como $\mathbf{Q}$ representa a probabilidade de um estado transciente i ser acessível de um estado j, por definição, o número de visitas em j dado que partimos de i é finita. Desta observação, concluímos que a matriz fundamental possui termos finitos. Em outras palavras $\sum \mathbf{Q}^n$ converge para $\mathbf{S}$. Vê-se, então, que deixando N ir para ∞, $\mathbf{S}$ é a solução de

$$\mathbf{S}(\mathbf{I} - \mathbf{Q}) = \mathbf{I} \tag{3.37}$$

Se T é finito, $\mathbf{S}$ é portanto a inversa de $\mathbf{I} - \mathbf{Q}$. Quando T é infinito, podem existir várias soluções de (3.37).

Teorema 3.12 (Caracterização Maximal da Matriz Fundamental)**.** *A matriz* $\mathbf{S}$ *dada por (3.37) é a menor solução de* $\mathbf{S}(\mathbf{I} - \mathbf{Q}) = \mathbf{I}$ *sujeita à restrição* $\mathbf{S} \geq \mathbf{0}$.

Demonstração: Seja $\mathbf{S}$ uma solução qualquer. Escrevendo $\mathbf{S}(\mathbf{I} - \mathbf{Q}) = \mathbf{I}$, temos $\mathbf{S} = \mathbf{I} + \mathbf{S}\mathbf{Q}$. Por iteração da última igualdade obtemos

$$\mathbf{S} = \mathbf{I} + \mathbf{Q} + \ldots + \mathbf{Q}^n + \mathbf{S}\mathbf{Q}^{n+1}.$$

Mas $\mathbf{S} \geq \mathbf{0}$ implica $\mathbf{S}\mathbf{Q}^{n+1} \geq \mathbf{0}$, e portanto

$$\mathbf{S} \geq \mathbf{I} + \mathbf{Q} + \ldots + \mathbf{Q}^{n-1}.$$

Passando ao limite obtém-se $\mathbf{S} \geq \sum_{n=0}^{\infty} \mathbf{Q}^n$. □

Exemplo 3.12.1. *Seja uma matriz de transição dada por*

$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.3 & 0.3 & & & & & \\ 0 & 0.6 & 0.4 & & & & & \\ 0.5 & 0.5 & 0 & & & & & \\ & & & 0 & 1 & & & \\ & & & 0.8 & 0.2 & & & \\ 0 & 0 & 0 & & & 04. & 0.6 & 0 \\ 0.4 & 0.4 & 0 & & & 0 & 0 & 0.2 \\ 0.1 & 0 & 0.3 & & & 0.6 & 0 & 0 \end{bmatrix}$$

Com espaço de estado $E = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Neste caso temos duas classes recorrentes $R_1 = \{1, 2, 3\}$ e $R_2 = \{4, 5\}$ e uma classe transciente $T = \{6, 7, 8\}$. Aqui

$$\mathbf{Q} = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0 & 0 & 0.2 \\ 0.6 & 0 & 0 \end{bmatrix}$$

e

$$\mathbf{S} = (\mathbf{I} - \mathbf{Q})^{-\mathbf{1}} = \frac{1}{66} \begin{bmatrix} 125 & 75 & 15 \\ 15 & 75 & 15 \\ 75 & 45 & 75 \end{bmatrix}.$$

Também, nenhum estado T leva a R_2, e portanto $g_{ij} = 0$ para $i \in T, j \in R_2$. Todos *os estados de T levam para R_1, e portanto $g_{ij} = +\infty$ para $i \in T, j \in R_1$. Resumindo,*

$$\mathbf{P} = \begin{bmatrix} +\infty & +\infty & +\infty & 0 & 0 & 0 & 0 & 0 \\ +\infty & +\infty & +\infty & 0 & 0 & 0 & 0 & 0 \\ +\infty & +\infty & +\infty & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +\infty & +\infty & 0 & 0 & 0 \\ 0 & 0 & 0 & +\infty & +\infty & 0 & 0 & 0 \\ +\infty & +\infty & +\infty & 0 & 0 & & & \\ +\infty & +\infty & +\infty & 0 & 0 & & & \\ +\infty & +\infty & +\infty & 0 & 0 & & & \end{bmatrix}$$

onde a parte que falta é $\mathbf{S}$.

3.6.2 Matriz Absorvente

Por fim procuraremos calcular a probabilidade de absorção por uma dada classe recorrente quanto partindo de um dado estado transciente. Como será visto mais adiante, é suficiente para a teoria tratar o caso onde as classes recorrentes possuem um único elemento. Portanto, supomos que a matriz de transição tem a forma

$$\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{B} & \mathbf{Q} \end{pmatrix}. \tag{3.38}$$

Seja $f_{ij} = \mathbb{P}(X_n = j|X_0 = i)$ a probabilidade de absorção pela classe recorrente $R_j = \{j\}$, quando tem início no estado transciente i. De (3.38),

$$\mathbf{P}^n = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{L}^n & \mathbf{Q}^n \end{pmatrix},$$

onde $\mathbf{L}_n = (\mathbf{I} + \mathbf{Q} + \ldots + \mathbf{Q}^n) \cdot \mathbf{B}$. Portanto, $\lim_{n\uparrow\infty} \mathbf{L}_n = \mathbf{SB}$, onde $\mathbf{S}$ é a matriz fundamental (Definição 3.11.1). Para $i \in T$, o termo (i,j) de $\mathbf{L}_n$ é

$$\mathbf{L}_n(i,j) = \mathbb{P}(X_n = j|X_0 = i).$$

Agora, se T_j é o primeiro tempo que visita a $R_j = \{j\}$ após o tempo 0, então

$$\mathbf{L}_n(i,j) = \mathbb{P}_i(T_j \leq n),$$

já que $R_j = \{j\}$ é um estado fechado. Deixando n ir ao ∞ obtém-se o Teorema.

Teorema 3.13 (Probabilidade de Absorção)**.** *Considere uma CMH com a matriz de transição* $\mathbf{P}$ *como em (3.38). A probabilidade de absorção pela classe recorrente* $R_j = \{j\}$ *a partir do estado i transciente é*

$$\mathbb{P}_i(T_{R_j} < \infty) = (\mathbf{SB})_{ij}.$$

O caso geral, onde as classes recorrentes não necessariamente possuem um único estado, pode ser reduzido a este como segue. Seja $\mathbf{P}^*$ uma matriz obtida da matriz de transição $\mathbf{P}$ dada por (3.23), por agrupamento das classes recorrentes R_j a um único estado $\hat{j}$, isto é,

$$\mathbf{P}^* = \begin{array}{c} \\ \hat{1} \\ \hat{2} \\ \vdots \\ T \end{array} \begin{array}{c} \begin{array}{cccc} \hat{1} & \hat{2} & \cdots & T \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ \hat{\mathbf{b}}_1 & \hat{\mathbf{b}}_2 & \cdots & \mathbf{Q} \end{pmatrix} \end{array}$$

(3.39)

onde $\mathbf{b}_{\hat{j}} = B(j)\mathbf{1}_T$ é obtido pelo somatório das colunas de $B(j)$, a matriz consistindo da coluna $j \in R$ de $\mathbf{B}$.

Exemplo 3.13.1. *Considere a Cadeia de Markov com o espaço de estados* $E = \{1, 2, 3, 4, 5, 6, 7\}$ *e a matriz de transição*

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & & & & & \\ 0.8 & 0.2 & & & & & \\ & & 0 & 0.4 & 0.6 & & \\ & & 1 & 0 & 0 & & \\ & & 1 & 0 & 0 & & \\ 0.1 & 0 & 0.2 & 0.1 & 0.2 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0 & 0.1 & 0.2 & 0.4 \end{bmatrix}.$$

Existem duas classes recorrentes $R_1 = \{1, 2\}$, $R_1 = \{3, 4, 5\}$ *e uma classe transciente* $T = \{6, 7\}$. *Em notação antes referida,*

$$\mathbf{B}_1 = \begin{bmatrix} 0.1 & 0 \\ 0.1 & 0.1 \end{bmatrix}, \ \mathbf{B}_2 = \begin{bmatrix} 0.2 & 0.1 & 0.2 \\ 0.1 & 0 & 0.1 \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} 0.3 & 0.1 \\ 0.2 & 0.4 \end{bmatrix}.$$

Multiplicando $\mathbf{B}_1$ *pelo vetor* $(1 \ \ 1)^T$ *e* $\mathbf{B}_2$ *pelo vetor* $(1 \ \ 1 \ \ 1)^T$ *obtemos*

$$\mathbf{b}_1 = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 0.5 \\ 0.2 \end{bmatrix}, \ \hat{\mathbf{B}} = \begin{bmatrix} 0.1 & 0.5 \\ 0.2 & 0.2 \end{bmatrix}.$$

Com isto criamos a matriz $\mathbf{P}^*$,

$$\mathbf{P}^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.1 & 0.5 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0.2 & 0.4 \end{bmatrix}.$$

Da matriz $\hat{\mathbf{B}}$ *concluímos que*

$$\mathbf{S} = (\mathbf{I} - \mathbf{Q})^{-1} = \begin{bmatrix} 1.5 & 0.25 \\ 0.5 & 1.75 \end{bmatrix}, \ \mathbf{S}\hat{\mathbf{B}} = \begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix}.$$

Assim, verificamos as seguintes probabilidades de absorção:
a) do estado transciente 6 para a classe $\{3, 4, 5\}$ *é 0.8 e para a classe* $\{1, 2\}$ *é 0.2,*
b) do estado transciente 7 para a classe $\{3, 4, 5\}$ *é 0.6 e para a classe* $\{1, 2\}$ *é 0.4.*

Apêndice

Teorema 3.1. *Lema de Cesaro Seja* $\{b_n\}_{n\geq 0}$ *uma sequência de números reais tal que*

$$\lim_{n\uparrow\infty} b_n = 0.$$

Então

$$\lim_{n\uparrow\infty} \frac{b_1 + \ldots + b_n}{n} = 0.$$

Demonstração: A sequência $\{b_n\}_{n\geq 0}$ é limitada em valor absoluto denominado de K. Para um $\epsilon > 0$ fixado arbitrariamente, existe n_0 tal que $n > n_0$ implica em que $|b_n| \leq \epsilon$, e portanto

$$\begin{aligned} \left|\frac{b_1 + \ldots + b_n}{n}\right| &\leq \left|\frac{b_1 + \ldots + b_{n_0}}{n}\right| + \left|\frac{b_{n_0} + \ldots + b_n}{n}\right| \\ &\leq \frac{n_0 K}{n} + \frac{n - n_0}{n}\epsilon \leq 2\epsilon, \end{aligned}$$

se n é suficientemente grande.

$\square$

Teorema 3.2. *Convergência dominada Seja* $\{a_{nk}\}_{n\geq 1, k\geq 1}$ *uma matriz de números reais tal que, para alguma sequência* $\{b_k\}_{k\geq 1}$ *de números não negativos satisfazendo*

$$\sum_{k=1}^{\infty} b_k < \infty, \tag{3.40}$$

então para todo $n \geq 1, \quad k \geq 1$,

$$|a_{nk}| \leq b_k \tag{3.41}$$

Se além disso para $k \geq 1$,

$$\lim_{n\uparrow\infty} a_{nk} = a_k \tag{3.42}$$

então

$$\lim_{n\uparrow\infty} \sum_{k=1}^{\infty} a_{nk} = \sum_{k=1}^{\infty} a_k. \tag{3.43}$$

Demonstração: Seja $\epsilon > 0$ fixo. Já que $\sum_{k=1}^{\infty} b_k$ é uma série convergente, podemos encontrar $M = M(\epsilon)$ tal que $\sum_{k=M+1}^{\infty} b_k < \frac{\epsilon}{3}$. Já que $|a_{nk}| \leq b_k$ e portanto $|a_k| \leq b_k$, temos

$$\sum_{k=M+1}^{\infty} |a_{nk}| + \sum_{k=M+1}^{\infty} |a_k| \leq \frac{2\epsilon}{3}.$$

Para n suficientemente grande,

$$\sum_{k=M+1}^{\infty} |a_{nk} - a_k| \leq \frac{2\epsilon}{3}.$$

Portanto, para n suficientemente grande,

$$\left| \sum_{k=1}^{\infty} a_{nk} - \sum_{k=1}^{\infty} |a_k \right| \leq \sum_{k=1}^{M} |a_{nk} - a_k| + \sum_{k=M+1}^{\infty} |a_{nk}| + \sum_{k=M+1}^{\infty} |a_k| \leq \frac{\epsilon}{3} + \frac{2\epsilon}{3} = \epsilon.$$

$\square$

Teorema 3.3. *Produtos infinitos Seja $\{a_n\}_{n\leq 1}$ uma sequência de números no intervalo [0,1).*

(a) Se $\sum_{n=1}^{\infty} a_n < \infty$, então

$$\lim_{n\uparrow\infty} \prod_{k=1}^{n} (1 - a_k) > 0$$

(b) Se $\sum_{n=1}^{\infty} a_n = \infty$, então

$$\lim_{n\uparrow\infty} \prod_{k=1}^{n} (1 - a_k) = 0$$

Demonstração:

Item (a)

Para quaisquer números $c_1, \ldots, c_n$ no intervalo $[0.1)$, e tal que $(1-c_1)(1-c_2)\ldots(1-c_n) \geq 1 - c_1 - c_2 - \ldots - c_n$(a demonstração se dá por indução). E já que $\sum_{n=1}^{\infty} a_n$ converge, existe N tal que para todo $n \geq N$,

$$a_N + \ldots + a_n < \frac{1}{2}.$$

Portanto, definindo $\pi(n) = \prod_{k=1}^{n}(1 - a_k)$, teremos para todo $n \geq N$

$$\frac{\pi(n)}{\pi(N-1)} = (1 - a_N) \ldots (1 - a_n) \geq 1 - (a_N + \ldots + a_n) \geq \frac{1}{2}.$$

Portanto a sequência $\{\pi(n)\}_{n \geq N}$ é uma sequência não crescente limitada inferiormente por$\frac{1}{2}\pi(N-1) > 0$, de modo que $\lim_{n\uparrow\infty} \pi(n) > 0$.

Item(b)

Usando a desigualdade $1-a \leq e^{-a}$ quando $a \in [0, 1)$, temos que $\pi(n) \leq e^{-a_1-a_2-\ldots-a_n}$ e portanto, se $\sum_{n=1}^{\infty} a_n = \infty$, $\lim_{n\uparrow\infty} \pi(n) = 0$ $\square$

Considerações Finais

O objetivo deste trabalho foi o de detalhar as demonstrações dos principais teoremas que dão suporte a teoria das cadeias de Markov homogêneas discretas. Apresentamos exemplos canônicos que ilustram a teoria ao longo do trabalho. Neste sentido, consideramos termos alcançado o nosso propósito, que foi o de apresentar um material didático, bem como, adquirimos conhecimentos necessários para melhor desempenhar nossa prática pedagógica, principalmente no que se refere aos processos estocásticos.

É nossa pretensão continuar o estudo das cadeias de Markov, tanto pela utilização desta teoria em aplicações de natureza prática, como também na orientação de trabalhos de conclusão de curso.

Referências Bibliográficas

[1] BARRY, R. James. **Probabilidade: um curso de nível intermediário**. 3. ed. IMPA. Rio de Janeiro - RJ, 2008.

[2] BRÉMAUD, Pierre. **Markov Chains - Gibbs fields, Monte Carlo Simulation and queues**. Spring-Verlag. New York - NY - USA. 1999.

[3] CANAVOS, George C. **Applied probability and Statistical methods**. Litle, Brown Company (Canada) Limited. Toronto. 1984.

[4] CHING, WAY-KI & K. NG.**Markov Chains: Models, Algorithms and Applications**. Spring Science, New York, NY: 2006.

[5] FREEDMAN, D. **Markov Chains**. Holden-Day, Inc. San Francisco-CA. USA. 1971.

[6] HOEL, G. Paul. **Introducion to Mathematical Statistics**, 5^a ed. John Wiley Sons, Inc. USA: 1984.

[7] HOEL, G. Paul, Port, Sidney C, Stone,Charles J. **Introduction to Stochastic Processes**. Waveland Press.Inc, Boston-USA,1987.

[8] MAGALHÃES, M. N. de Lima, A.C.P. **Noções de Probabilidade e Estatistica**, 6. ed. EDUSP, São Paulo-SP, 2004.

[9] MEYER, P.L. **Probabilidade - Aplicacações à Estatistica**. 2a. ed. LTC. Rio de Janeiro - RJ: 1983.

[10] NORRIS, J.R. **Markov chains**. Cambridge University Press. New York - NY, 1997.

[11] REVUZ, Z. **Markov Chains**. 2a. ed. Elsevier Science Puplisher B. V. Amsterdan - Netherlands. 1991.

[12] ROSS, Sheldon M. **Introduction to Probability Models** 9a. ed. Elsevier Inc, San Diego, California - USA. 2007.

[13] SHIRYAEU. A. N. **Probability**. 2nd ed. Springer-Verlag. New York Inc. New York-NY. USA. 1996.