Universidade Federal da Bahia - UFBA
Instituto de Matemática e Estatística - IME
Sociedade Brasileira de Matematica - SBM
Mestrado Profissional em Matemática em Rede Nacional - PROFMAT
Dissertação de Mestrado

Cadeias de Markov Absorventes: Uma aplicação para o Ensino Médio

Thiago Emmanoel Amaral Nascimento

Salvador - Bahia
Abril de 2019

Cadeias de Markov Absorventes: Uma aplicação para o Ensino Médio

Thiago Emmanoel Amaral Nascimento

Dissertação de Mestrado apresentada à Comissão Acadêmica Institucional do PROFMAT-UFBA como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Joseph Nee Anyah Yartey

Salvador - Bahia

Abril de 2019

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

Nascimento, Thiago Emmanoel Amaral
Cadeias de Markov Absorventes: Uma aplicação para o ensino / Thiago Emmanoel Amaral Nascimento. -- Salvador, 2019.
91 f. : il

Orientador: Prof. Dr. Joseph Nee Anyah Yartey.
Dissertação (Mestrado - Mestrado Profissional em Matemática em Rede Nacional - PROFMAT) -- Universidade Federal da Bahia, Instituto de Matemática e Estatística - IME, 2019.

1. Cadeias de Markov. 2. Cadeias de Markov Absorventes. 3. Ensino Médio. I. Nee Anyah Yartey, Prof. Dr. Joseph. II. Título.

Cadeias de Markov Absorventes.

Thiago Emmanoel Amaral Nascimento

Dissertação de Mestrado apresentada à comissão Acadêmica Institucional do PROFMAT-UFBA como requisito parcial para obtenção do título de Mestre em Matemática, aprovada em 25/04/2019.

Banca Examinadora:

Prof. Dr. Joseph Nee Anyah Yartey (orientador)
UFBA

Mariana Cassol

Profª. Drª. Mariana Cassol
UFBA

Celso Costa

Prof. Dr. Celso José da Costa
UFF

Cláudia Ribeiro Santana

Profª. Drª. Cláudia Ribeiro Santana
UESC

À minha família

Agradecimentos

Agradeço a Deus pela minha vida e por me assegurar força neste caminho que apenas eu e ele sabemos o quanto foi duro e árduo caminhar. Agradeço a minha esposa e filha por compreender os momentos ausentes e pelo amor a mim dedicado que me fazem superar as horas de medo e angústia. Agradeço aos meus pais o ensinamento de que a educação é um bem valioso e por todo o sacrifício que fizeram para me garantir a mesma.

Agradeço a todos os colegas de curso pelas horas de estudos e pela união que fez de nossa turma um grupo de agradável e alegre convivência. Agradeço a todos os professores que nos cederam um pouco de sua sabedoria e paciência, que com certeza nos elevaram a um outro patamar de conhecimento. Em especial, agradeço ao professor Joseph, meu orientador, um grande mestre, que impressiona pela sua humildade e inteligência.

Agradeço a todos os meus colegas das escolas onde trabalho por acreditar e persistir na difícil missão de educar nossos jovens. Agradeço ainda ao Instituto de Matemática e Estatística da UFBA por adotar e manter o Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT como local de aprimoramento dos professores de matemática do estado da Bahia. Por fim, gostaria de agradecer a Sociedade Brasileira de Matemática - SBM e ao Instituto de Matemática Pura e Aplicada - IMPA pela criação e empenho na manutenção do mestrado profissional que com certeza ajuda, e muito, a melhorar o ensino de matemática em nosso país.

"Ensinar a resolver problemas é
educar a vontade".
George Polya

Resumo

No início do século XX, Andrei Markov começou um importante estudo de um novo tipo de processo onde em um experimento, o resultado atual influencia o resultado que ocorrerá naquele imediatamente posterior a ele. A este tipo de processo chamamos de **Cadeias de Markov**. Estas cadeias são melhores estudadas quando são consideradas alguns tipos especiais da mesma.

Neste trabalho vamos estudar as Cadeias de Markov Absorventes, fazendo uma fundamentação teórica e dando alguns exemplos que mostram o funcionamento matemático destas cadeias. Iremos ainda criar algumas atividades e materiais pedagógicos que auxiliem as aulas do professor que deseje ensinar as Cadeias de Markov Absorventes em turmas de matemática do ensino médio.

Palavra-Chave: Cadeias de Markov; Cadeias de Markov. Absorventes

Abstract

In the early twentieth century, Andrei Markov began an important study of a new kind of process where in an experiment, the current result influences the outcome that will occur in that immediately after it. To this type of process we call **Markov chains**. These chains are best studied when considering some special types of them.

In this work we will study the Absorbent Markov Chains making a theoretical basis and giving some examples that show the mathematical functioning of these chains. We will also create some activities and teaching materials that will support the teacher's classes that wish to teach the Absorbent Markov Chains in high school math classes.

Keyword: Markov Chains; Absorbing Markov Chains.

Lista de Figuras

Lista de Tabelas

Sumário

Introdução

Um processo estocástico é um modelo matemático que evolui ao longo do tempo de forma probabilística. Quando a probabilidade que governa esta evolução não depende dos resultados anteriores ao atual, ou seja, o processo não possui memória, dizemos que este processo possui a propriedade markoviana.

Um processo estocástico que possui a propriedade markoviana e tem o seu espaço de estados discreto é denominado cadeia de Markov. Assim para que um dado processo estocástico seja considerado uma cadeia de Markov deve ter as seguintes características: ter a propriedade markoviana; o conjunto de estados possíveis ao processo é discreto e que ele irá ocupar algum desses estados a cada etapa do processo; o estado inicial em que ele se encontra seja conhecido e as probabilidades de transição entre os estados também estejam definidas.

Indexando os possíveis estados de uma cadeia de Markov podemos associar a ela uma matriz de transição quadrada de ordem i, onde i é a quantidade de estados da cadeia. Assim, cada componente da matriz, representa a probabilidade de que a cadeia estará no estado associado ao número da coluna dado que na etapa, imediatamente anterior, encontrava-se no estado associado ao número da linha.

Com esta matriz de transição e a imposição de algumas restrições podemos estudar tipos característicos de cadeias de Markov que possuem aplicações nas mais diversas situações. Particularmente neste trabalho iremos focar nas cadeias de Markov com estados absorventes. Dizemos que um estado é absorvente se ao adentrar nele é impossível sair. Um estado absorvente está para o processo, assim como, um buraco negro está para a vida de uma estrela ou a morte está para um ser vivo.

As cadeias de Markov, em geral, podem ser aplicadas a uma ampla quantidade de problemas, como podemos ver facilmente ao fazer uma breve pesquisa na internet. São aplicações na área da termodinâmica, Química, reconhecimento de fala, teoria das filas, ciências da informação, aplicações de internet, estatística, economia e finanças, ciências sociais, Biologia, estratégias militares, genética e até mesmo na música.

Iremos estruturar este trabalho da seguinte maneira. No capítulo 1 faremos uma breve biografia de Andrei Andreyevich Markov (1856-1922) um jovem matemático russo que foi o criador das cadeias que levam o seu nome. Markov tem uma história de vida e acadêmica muito peculiar. Acreditamos que seja válido e relevante conhecer o homem por trás da teoria, para melhor entender sua obra. Em seguida, iremos tratar de alguns conceitos básicos associados as cadeias de Markov, como o conjunto de estados de uma cadeia e as probabilidades de transição da mesma. Ainda neste capítulo introduzimos as definições e teoremas que nos permitem explorar as propriedades mais genéricas das cadeias de Markov. Finalizando este capítulo damos alguns exemplos que nos parecem pertinentes e que esclarecem os conceitos e definições dadas neste capítulo.

Começamos o capítulo 2 com a classificação dos estados de uma cadeia de Markov. Um estado de uma cadeia de Markov será chamado de absorvente quando ao entrar naquele estado o processo não irá mais sair dele. Daí temos que uma cadeia de Markov é absorvente se possui um estado absorvente e, além disso, este estado é acessível a qualquer outro estado não absorvente. Em seguida definimos a forma canônica de uma cadeia absorvente, bem como a matriz fundamental associada a ela.

Com essa matriz fundamental podemos extrair várias informações relevantes sobre a cadeia que nos permite prever muitos resultados interessantes sobre as mesmas. Dentre eles estão o número de vezes que o processo irá ter até que seja absorvido e ainda a probabilidade de que seja absorvido dado que iniciou o processo num estado não absorvente. No final do capítulo também damos alguns exemplos que servem para consolidar e ilustrar os conceitos associados as cadeias de Markov absorventes.

O último capítulo é dedicado a comentar os exercícios e atividades propostas no site. São diretrizes que irão ajudar o professor na condução das aulas e os estudantes na resolução das questões. Acreditamos que o público ideal para a aplicação exitosa do material produzido sejam alunos do 2º ou 3º ano do ensino médio. Porque nestes casos eles já estudaram matrizes, determinantes e sistemas lineares, assim como os conceitos básicos de probabilidade.

Nos apêndices estão as listas de exercícios, suas respectivas soluções e alguns cálculos efetuados, nos exemplos dos capítulos anteriores, que achamos necessário deixar explícitos. Nos anexos temos um breve resumo com conceitos básicos sobre matrizes, determinantes e probabilidade, que caso o leitor precise estará disponível para consulta. Esperamos que aproveite a leitura.

Capítulo 1

Cadeias de Markov

Em 1907, Andrei Markov começou um importante estudo de um novo tipo de processo onde o resultado de um experimento atual é afetado pelo resultado que ocorreu no resultado imediatamente anterior a ele. A este tipo de processo chamamos de **Cadeias de Markov**. No entanto antes de definir o que é uma Cadeia de Markov, vejamos uma breve biografia de Andrei Markov que tem como base as informações encontradas nos artigos de J J O'Connor and E F Robertson na url www-groups.dcs.st-and.ac.uk/history/Biographies/Markov.html acessado em 24 de fevereiro de 2019 e o artigo de Brian Hayes acessado na url www.americanscientist.org/article /first-links-in-the-markov-chain em 06 de março de 2019.

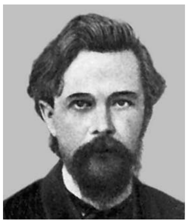


Figura 1.1: *Andrei Andreyevich Markov*

1.1 Andrei Andreyevich Markov

Andrei Andreyevich Markov filho de Nadezhda Petrovna e Andrei Grigorievich Markov, foi um grande matemático russo nascido na cidade de Riazan em 14 de junho de 1856. Teve muitas irmãs e um irmão mais novo chamado Vladimir, que também era um matemático conhecido na época que morreu muito jovem aos 25 anos, de tuberculose. Andrei Markov foi um jovem também de saúde frágil que até os dez anos andava com ajuda de muletas.

Desde o ginásio, Andrei Markov, já revelara seu talento para matemática e nesta mesma época escreveu seu primeiro artigo matemático, cujo tema era a integração de equações diferenciais. Apesar de não ser um artigo original, lhe rendeu um encontro com os professores universitários Aleksandr Korkin(1837-1908) e Yegor Ivanovich Zolotarev(1847-1878) da universidade de São Petersburgo, onde ele ingressou, em 1874, na Faculdade de Física e Matemática.

Na universidade ele se matriculou nos seminários de Korkin e Zolotarev, mas se encantou mesmo com as aulas do grande matemático Pafnuti Chebyshev(1821-1894), que eram estimulantes para Markov, pois nelas os alunos eram imersos em uma atmosfera de pesquisa, tendo sempre novas questões e problemas para que os estudantes investigassem.

Markov graduou-se em 1878, quando ganhou uma medalha de ouro por submeter o melhor artigo de sua faculdade naquele ano, cujo título era - *Sobre a integração de equações diferenciais por meio de frações contínuas.* Após sua graduação ele começou seu mestrado com o desejo de se tornar professor universitário. Conseguiu concluir seu mestrado em apenas dois anos e em 1880 apresentou sua dissertação cujo título foi - *Sobre as formas quadráticas binárias com determinante positivo.* Ainda neste mesmo ano Markov começou a ensinar na Universidade de São Petersburgo e iniciou seu doutoramento que concluiu em 1884 com sua tese *Sobre algumas aplicações das frações contínuas.*

Maria Ivanova Valvatyeva, sua futura esposa, era uma amiga de infância, mas apesar disso, a mãe de Ivanova não era a favor do casamento entre eles. Visto que, Markov era filho do gerente de sua propriedade e portanto, para ela, não possuía o status social necessário para casar com sua filha. Porém em 1883 a mãe de Ivanova concorda com o casamento, talvez pela trajetória de sucesso que Markov vinha traçando e eles se casam neste mesmo ano.

Andrei Markov tornou-se professor titular da Universidade de São Petersburgo em 1893. Chebyshev propôs que Markov fosse membro da Academia Russa de Ciências em 1886. Ele foi eleito como um membro extraordinário em 1890 e acadêmico em 1896. Aposentou-se formalmente em 1905, mas continuou a ensinar durante à maior parte de sua vida.

O trabalho inicial de Markov foi em análise e teoria dos números, frações algébricas contínuas, limites de integrais, teoria de aproximação e convergência de séries. Depois de 1900 ele aplicou o método das frações contínuas, de Chebyshev, a teoria da probablidade.

Entretanto Markov é particularmente lembrado por seu estudo das cadeias de Markov, este trabalho fundou um ramo completamente novo da teoria da probabilidade e lançou a teoria dos processos estocásticos. Em 1923, Norbert Wiener(1894-1964) tornou-se o primeiro a tratar rigorosamente um processo contínuo de Markov. A fundação de uma teoria geral foi fornecida a partir dos anos de 1930 por Andrei Kolmogorov(1903-1987) e Sergei Bernstein(1880-1968).

O curso clássico de Markov sobre o cálculo de probabilidades e suas notas originais, são lembrados por sua exatidão e clareza de exposição. Contribuíram muito para a transformação da teoria da probabilidade em uma das áreas mais estudadas da matemática, e para a ampla disseminação dos métodos e das pesquisas de Chebyshev. Sua análise profunda das dependências entre os fenômenos aleatórios observados, permitiu a Markov estender a teoria da probabilidade de uma maneira essencial através da introdução e investigação de quantidades aleatórias dependentes.

Suas palestras eram marcadas por um rigor irrepreensível de argumentação e buscava desenvolver em seus alunos um pensamento matemático crítico, que não toma nada como garantido. Ele incluiu em seus cursos muitos resultados recentes de investigações e freqüentemente omitia questões tradicionais.

Conforme Hayes, segundo a maioria dos relatos, Markov era um personagem irritante, abrasivo mesmo com amigos, ferozmente combativo com rivais, muitas vezes envolvido em protestos e brigas públicas. Um fato que nos dá um vislumbre de sua forte personalidade são suas correspondências com o estatístico Alexander Chuprov (1874-1926). Suas cartas a Chuprov são repletas de comentários desdenhosos denegrindo o trabalho de outros - inclusive o de Chuprov.

A combatividade de Markov se estendeu além da matemática, passando pela política e pela vida pública. Quando a igreja russa excomungou Leo Tolstoy, Markov pediu que ele fosse expulso também, o que foi feito. Em 1902, o escritor de esquerda Maxim Gorky foi eleito para a Academia, mas a eleição foi vetada pelo czar Nicolau II. Em protesto, Markov anunciou que recusaria todas as honras futuras do czar. Em 1913, quando o czar convocou celebrações de 300 anos de governo Romanov, Markov respondeu organizando um simpósio comemorando um aniversário diferente: a publicação de Ars Conjectandi 200 anos antes, livro sobre combinação e probabilidade escrito por Jacob Bernoulli.

Porém o maior conflito de Markov foi reservado para outro matemático, Pavel Nekrasov (1853–1924), cujo trabalho Markov descreveu como “um abuso da matemática”. Nekrasov estava na faculdade da Universidade de Moscou, que era então um reduto da Igreja Ortodoxa Russa. Nekrasov havia começado a estudar em um seminário teológico antes de se voltar para a matemática e aparentemente acreditava que as duas vocações poderiam se apoiar mutuamente.

Em um artigo publicado em 1902, Nekrasov injetou a lei de grandes números no debate teológico centenário sobre o livre-arbítrio versus a predestinação. Seu argumento foi mais ou menos assim: Atos voluntários - expressões de livre arbítrio - são como os eventos independentes da teoria da probabilidade, sem ligações causais entre eles. A lei dos grandes números aplica-se apenas a tais eventos independentes.

Markov e Nekrasov encontravam-se em pólos opostos: um republicano secular de Petersburgo confrontava um monarquista eclesiástico de Moscou. Mas quando Markov lançou seu ataque a Nekrasov, ele não o fez com bases em suas diferenças faccionais ou ideológicas. Ele se concentrou em um erro matemático. Nekrasov assumiu que a lei dos grandes números requer o princípio da independência. Embora essa noção tenha sido um lugar comum da teoria da probabilidade desde o tempo de Jacob Bernoulli, Markov começou a mostrar que a suposição é desnecessária. A lei dos grandes números aplica-se perfeitamente aos sistemas de variáveis dependentes se satisfizerem certos critérios e assim nasceram as cadeias de Markov.

A Revolução Russa começou no início de 1917, quando os suprimentos de comida se esgotaram. Em setembro daquele ano, Markov solicitou que a Academia o mandasse para uma cidade desfavorecida no interior da Rússia. Ele foi enviado para Zaraisk, uma pequena cidade do interior, onde lecionava matemática na escola secundária sem receber nenhuma remuneração.

Ele retornou a São Petersburgo, mas sua saúde já estava se deteriorando. Embora em 1921 ele estivesse com um estado de saúde muito ruim, de forma que mal conseguia se manter em pé, ainda assim continuou a dar palestra sobre probabilidade na universidade. Sua morte em julho de 1922 veio depois de meses de sofrimento muito grave. Markov teve um filho (de mesmo nome) que nasceu em 9 de setembro de 1903 e seguiu seu pai também tornando-se um renomado matemático.

1.2 Noções Básicas

A maioria dos estudos de probabilidade utiliza o *Princípio de Independência de Eventos*, ou seja, que em um experimento aleatório o resultado atual ou anterior não tem influência sobre um resultado futuro. Para ilustrar tal princípio, basta lembrar que no lançamento de uma moeda honesta, a probabilidade de que a moeda dê cara ou coroa em um lançamento não se altera devido a ocorrência de cara ou corôa em qualquer um dos lançamentos anteriores. Porém, existem estudos em que a probabilidade de um determinado resultado é influenciada pela ocorrência do resultado atual ou anteriores.

As Cadeias de Markov modelam processos onde o próximo resultado é influenciado pelo resultado atual obtido. Antes de formalizar uma definição matemática mais rigorosa para uma *Cadeia de Markov*, vamos descrever quais os elementos que a formam e como a mesma funciona. Considere um conjunto de *estados* possíveis $S = \{s_1, s_2, \ldots, s_r\}$ de uma variável aleatória. Uma Cadeia de Markov tem início em um desses estados e a probabilidade de que haja mudança neste estado depende do estado em que o processo se encontra. Chamaremos cada uma dessas mudanças de estado de *passo ou etapa.*

Se a cadeia encontra-se atualmente no estado s_i e se move para o estado s_j no próximo passo, a probabilidade, indicada por p_{ij}, não depende dos resultados que ocorreram em passos anteriores ao estado atual da cadeia. Por isso dizemos que uma Cadeia de Markov é um processo que não possui memória, ou seja, o que ocorrerá no próximo passo depende apenas do estado atual da cadeia.

As probabilidades p_{ij} são chamadas de *probabilidades de transição.* O processo pode permanecer no mesmo estado em um passo qualquer e neste caso a probabilidade de transição será indicada por p_{ii}. Uma distribuição inicial de probabilidade, definida sobre S, e um estado inicial define uma cadeia de Markov, veja no diagrama abaixo uma cadeia de Markov com quatro estados.

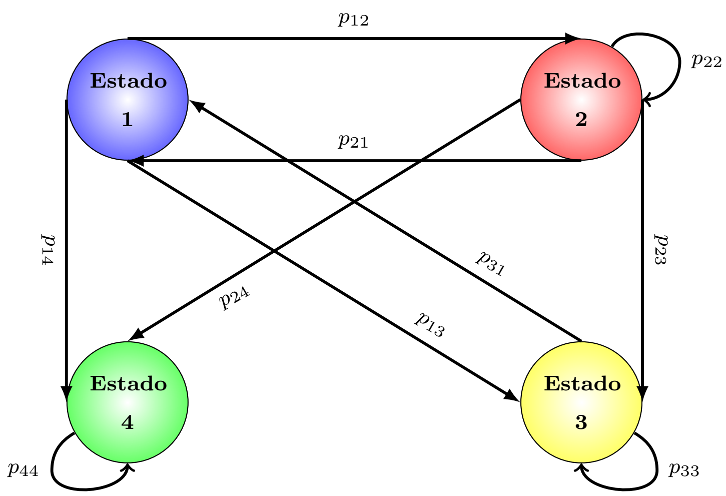


Figura 1.2: *Diagrama de estados*

Na tabela 1.1 a seguir, estão registradas as respectivas probabilidades de transição para cada mudança de estado.

Tabela 1.1: Distribuição de Probabilidades

Probabilidade de transição (p_{ij})	**Valor**	**Probabilidade de transição** (p_{ij})	**Valor**
Permanecer no Estado 1 (p_{11})	0	Ir do Estado 3 para o 1 (p_{31})	0,25
Ir do Estado 1 para o 2 (p_{12})	0,5	Ir do Estado 3 para o 2 (p_{32})	0,25
Ir do Estado 1 para o 3 (p_{13})	0,30	Permanecer no Estado 3 (p_{33})	0,5
Ir do Estado 1 para o 4 (p_{14})	0,20	Ir do Estado 3 para o 4 (p_{34})	0
Ir do Estado 2 para o 1 (p_{21})	0,40	Ir do Estado 4 para o 1 (p_{41})	0
Permanecer no Estado 2 (p_{22})	0,20	Ir do Estado 4 para o 2 (p_{42})	0
Ir do Estado 2 para o 3 (p_{23})	0,30	Ir do Estado 4 para o 3 (p_{43})	0
Ir do Estado 2 para o 4 (p_{24})	0,10	Permanecer no Estado 4 (p_{44})	1

Vejamos agora como interpretar a Cadeia da figura 1.2 associada as probabilidades de transição dadas na tabela 1.1. Por exemplo, observe que se o processo encontra-se no estado 2 ele pode ir no próximo passo para o estado 3 (p_{23}) e isso ocorre com uma probabilidade de 30%. Temos ainda que se o processo encontra-se no estado 3 a probabilidade dele permanecer neste estado (p_{33}) é de 50%. Note também que do estado 1 podemos ir para qualquer um dos outros estados, mas nunca permaneceremos neste mesmo estado no próximo passo, pois a probabilidade de permanecer no estado 1 (p_{11}) é zero. Porém, observe que ao chegar no estado 4 o processo se encerra, pois irá permanecer neste estado para sempre já que a probabilidade de permanecer em 4 (p_{44}) é de 100%, no capítulo 2 chamaremos este tipo de estado de absorvente. Dado esses esclarecimentos iniciais vejamos como definir formalmente uma Cadeia de Markov.

1.3 Definições

Nesta seção vamos definir formalmente alguns conceitos que são necessários para caracterizar uma Cadeia de Markov.

Definição 1.3.1. *Um **processo estocástico** $\{X(t) \in E \mid t \in U\}$ é uma sequência de variáveis aleatórias $X(t)$ que descreve a evolução de alguma característica X de um processo sob análise ao longo do tempo $t \in U$, tal que $U \subset \mathbb{R}$.*

Para as definições a seguir, seja U o conjunto dos números naturais e a variável aleatória X, representa o *estado do sistema* no instante t. O conjunto E, chamado de *espaço de estados*, é discreto e finito, ou seja, possui M elementos, com $M \in \mathbb{N}$. Desta forma o processo estocásticos $\{X(t)\}$ fornece uma representação matemática de como o processo evolui ao longo do tempo.

Definição 1.3.2. *Um processo estocástico é denominado **processo markoviano** de parâmetro discreto, se dada uma sequência de tempos $t_0 < t_1 < t_2 < \cdots < t_n < t_{n+1}$, a distribuição de probabilidade condicional de $X(t_{n+1})$ para dados valores de $X(t_0), X(t_1), \ldots, X(t_n)$ depende unicamente de $X(t_n)$, ou seja,*

$$P\{X(t_{n+1}) = j \mid X(t_0) = k_0, X(t_1) = k_1, \ldots, X(t_{n-1}) = k_{n-1}, X(t_n) = i\}$$

$$=$$

$$P\{X(t_{n+1}) = j \mid X(t_n) = i\}$$

Um processo markoviano que possui espaço de estados E finito é denominado **cadeia de Markov**. As probabilidades condicionais $P\{X(t_{n+1}) = j \mid X(t_n) = i\}$ para uma cadeia de Markov são chamadas **probabilidades de transição** para um *passo* ou etapa.

Para simplificar a escrita vamos utilizar as seguintes notações:

$$p_{ij} = P\{X(t+1) = j \mid X(t) = i\}$$

$$p_{ij}^{(n)} = P\{X(t+n) = j \mid X(t) = i\}$$

Note que as $p_{ij}^{(n)}$ são as probabilidades condicionais para que se esteja em um estado i e se passe para um estado j em n passos. Além disso, como essas quantidades representam probabilidades condicionais são números **não** negativos. Observe ainda que o processo deve realizar sempre uma transição para algum estado em cada passo, logo elas devem satisfazer as seguintes propriedades:

$$p_{ij}^{(n)} \geq 0 \text{ para todo i e j; } n \in \mathbb{N}$$

e

$$\sum_{j=0}^{M} p_{ij}^{(n)} = 1 \text{ para todo i e j; } n \in \mathbb{N}$$

Outro fato importante é que $p_{ij}^{(1)} = p_{ij}$, e que $p_{ij}^{(0)} = \begin{cases} 1, \text{ se } & i = j \\ 0, \text{ se } & i \neq j \end{cases}$.

O exemplo a seguir é uma adaptação daquele que se encontra em Poole (2004, p. 206) e achamos importante citá-lo, pois esclarece e ilustra as definições dadas até o presente momento.

Exemplo 1.3.1. *Um psicológo coloca um rato em uma gaiola de 5 compartimentos, como mostra a figura abaixo. O rato foi treinado para selecionar uma porta aleatoriamente sempre que tocarem um sinal, e dirigir-se através dela para o próximo compartimento.*

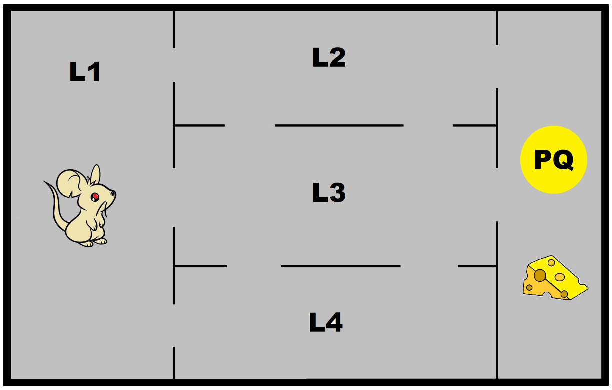


Figura 1.3: *Labirinto do Rato*

Esta situação pode ser modelada por uma Cadeia de Markov, pois o compartimento ao qual o rato irá acessar após cada sinal depende, apenas, do compartimento onde ele se encontra atualmente(propriedade markoviana). Ademais esta é uma cadeia com 5 estados, onde o compartimento PQ(paraíso do queijo) é um estado absorvente. As probabilidades de transição entre os compartimentos estão representadas no diagrama 1.4.

Note que as setas em cada nó do diagrama, representam a probabilidade de transição que o rato possui para que estando naquele compartimento acesse aquele que a seta está indicando. Além disso, o somatório das probabilidades das setas que saem de cada nó são sempre iguais a 1, ou seja, $\sum_{j=1}^{5} p_{ij} = 1$ para todo $i \in 1, 2, 3, 4, 5$.

Uma maneira mais conveniente e eficaz de representar todas as probabilidades de transição em n etapas de uma cadeia de Markov é utilizar o formato matricial, que definimos abaixo.

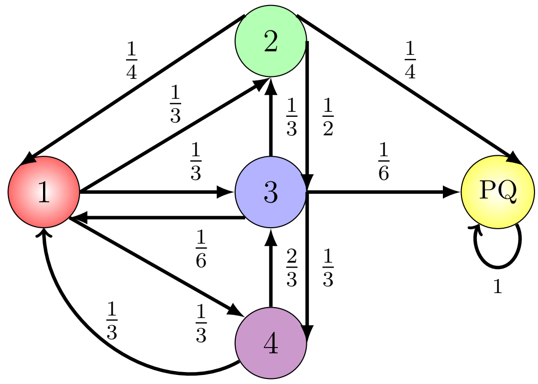


Figura 1.4: *Diagrama de estados - Labirinto do Rato*

Definição 1.3.3. *A* ***matriz de transição em n-etapas*** *de uma cadeia de Markov é a matriz $P^{(n)}$ com entradas $p_{ij}^{(n)}$.*

$$
\mathbf{P^{(n)}} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ M \end{array}
\begin{array}{c}
\begin{array}{cccc} \quad 1 \quad & \quad 2 \quad & \cdots & \quad M \quad \end{array} \\
\left[\begin{array}{cccc}
p_{11}^{(n)} & p_{12}^{(n)} & \cdots & p_{1M}^{(n)} \\
p_{21}^{(n)} & p_{22}^{(n)} & \cdots & p_{2M}^{(n)} \\
\vdots & \vdots & \ddots & \vdots \\
p_{M1}^{(n)} & p_{M2}^{(n)} & \cdots & p_{MM}^{(n)}
\end{array}\right]
\end{array}
$$

(Column header row labelled *Estado*: $1 \quad 2 \quad \cdots \quad M$.)

A probabilidade de transição se dá do estado de linha para o estado de coluna. Além disso, quando $n = 1$, eliminamos o superescrito (n) e neste caso temos a *matriz de transição* da cadeia de Markov.

Para o exemplo 1.3.1 podemos utilizar o diagrama 1.4 para construir a matriz de transição do experimento, que fica da seguinte maneira.

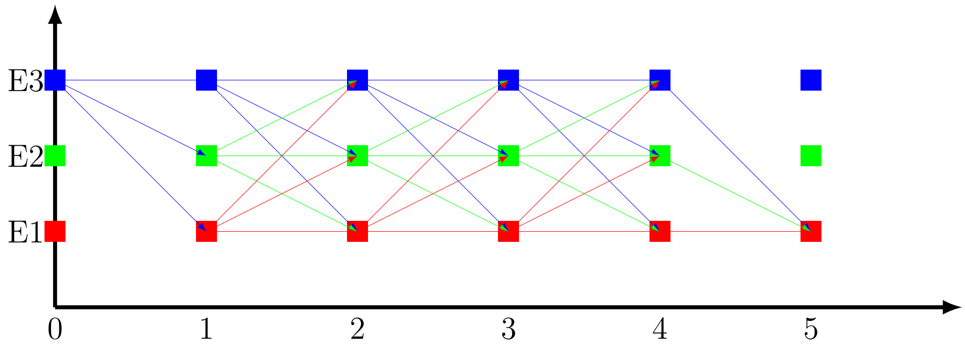
$$
\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ PQ \end{array}
\begin{array}{c}
\begin{array}{ccccc} 1 & 2 & 3 & 4 & PQ \end{array} \\
\left[\begin{array}{ccccc}
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\
\frac{1}{4} & 0 & \frac{2}{4} & 0 & \frac{1}{4} \\
\frac{1}{6} & \frac{2}{6} & 0 & \frac{2}{6} & \frac{1}{6} \\
\frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{array}\right]
\end{array}
$$

Vamos agora enunciar e provar o primeiro teorema relacionado as cadeias de Markov.

1.4 Equação de Chapman-Kolmogorov

Um importante teorema que nos garante uma relação entre as probabilidades de transição entre estados em n-etapas e as potências da matriz de transição de uma cadeia de Markov é a equação de **Chapman-Kolmogorov**.

A figura abaixo nos ajuda a compreender melhor o enunciado do teorema. Tomemos como exemplo uma cadeia de Markov com 3 estados possíveis e observemos como se comporta a probabilidade de transição de passar do estado 3 para o estado 1 em 5 etapas, vejamos:



A equação de Chapman-Kolmogrov afirma que a probabilidade de que o processo se encontre, por exemplo, no estado 1 dado que ele iniciou no estado 3 em 5 etapas é dada pela soma do produto entre as probabilidades de ir do estado 3 para cada um dos estados em 3 etapas e as probabilidades de ir de cada um dos estados até o estado 1 em duas etapas. A partição do exemplo foi de 3 para 2 etapas, mas poderia ser qualquer outra, cujo resultado fosse igual a cinco.

Para ter uma boa intuição da veracidade dessa afirmação basta perceber que a figura acima é uma forma condensada do diagrama de árvores das probabilidades condicionais do problema. Com isso é fácil perceber que de fato a afirmação parece verdadeira, mas como em matemática apenas a intuição não basta, vejamos agora o enunciado e a demonstração deste teorema.

Teorema 1.4.1. *Seja $\{X(t)\}$ uma cadeia de Markov, com probabilidades de transição dada por p_{ij}. Para qualquer par de tempos r e n, tal que $0 \leq r \leq n$ e para todo $k \in E$ temos que:*

$$p_{ij}^{(n)} = \sum_{k=1}^{M} p_{ik}^{(r)} \cdot p_{kj}^{(n-r)}$$

Demonstração. Por definição sabemos que para passar do estado atual i para o estado j em n etapas, a probabilidade de transição será dada por:

$$p_{ij}^{(n)} = P\{X(n) = j | X(0) = i\}$$

Usando a definição de probabilidade condicional, temos que:

$$P\{X(n)=j|X(0)=i\}=\frac{P\{X(n)=j\cap X(0)=i\}}{P\{X(0)=i\}}$$

Pelo teorema da probabilidade total, podemos afirmar que:

$$\frac{P\{X(n)=j\cap X(0)=i\}}{P\{X(0)=i\}}=\sum_{k=1}^{M}\frac{P\{X(n)=j\cap X(r)=k\cap X(0)=i\}}{P\{X(0)=i\}}$$

Novamente pela definição de probabilidade condicional podemos escrever que:

$$\sum_{k=1}^{M}\frac{P\{X(n)=j\cap X(r)=k\cap X(0)=i\}}{P\{X(0)=i\}}=\sum_{k=1}^{M}\frac{P\{X(n)=j|X(r)=k\cap X(0)=i\}\cdot P\{X(r)=k\cap X(0)=i\}}{P\{X(0)=i\}}$$

Daí, como $r\geq 0$ e aplicando a propriedade markoviana podemos afirmar que:

$$P\{X(n)=j|X(r)=k\cap X(0)=i\}=P\{X(n)=j|X(r)=k\}$$

Portanto:

$$\sum_{k=1}^{M}\frac{P\{X(n)=j|X(r)=k\cap X(0)=i\}\cdot P\{X(r)=k\cap X(0)=i\}}{P\{X(0)=i\}}=\sum_{k=1}^{M}P\{X(n)=j|X(r)=k\}\cdot P\{X(r)=k|X(0)=i\}$$

Para finalizar basta aplicar a definição da probabilidade de transição as duas probabilidades condicionais encontradas para concluirmos que: $p_{ij}^{(n)}=\sum_{k=1}^{M}p_{ik}^{(r)}\cdot p_{kj}^{(n-r)}$ ■

Uma importante consequência da equação de Chapman-Kolmogorov é o seguinte corolário:

Corolário 1.4.1. *Se P é a matriz de probabilidades de transição do estado i para o estado j em um passo, então a matriz de probabilidades de transição do estado i para o estado j em n passos é dada pela potência de P elevada a n, ou seja, $P^{(n)}=P^n$.*

Demonstração. Usaremos a indução finita sobre n para provar a sentença acima. Para $n=1$ temos que $P^{(1)}=P^1=P$ que é verdade por definição. Suponha que a sentença seja verdadeira para $n=m$, com $m\in\mathbb{N}$ e $m>1$, isto é, $P^{(m)}=P^m$, devemos mostrar que a sentença é verdadeira quando $n=m+1$. De fato, usando a equação de Chapman-Komogorov, para $r=1$, temos que:

$$\begin{aligned} P^{(m+1)} = p_{ij}^{(m+1)} &= \sum_{k\in E}^{M} p_{ik}^{(1)} \cdot p_{kj}^{(m)} \\ &= \sum_{k\in E}^{M} p_{ik}^{1} \cdot p_{kj}^{m} \\ &= P.P^{m} = P^{m+1} \end{aligned}$$

Logo pelo princípio da indução finita a sentença é verdadeira para todo $n \in \mathbb{N}$. ■

Assim, a equação de Chapman-Kolmogorov nos mostra que para saber qual a probabilidade de se passar de um estado i para um estado j em n-etapas, basta que calculemos as potência da matriz de transição elevada a n para obter a probabilidade desejada.

Isso nem sempre é simples de se fazer, pois envolve multiplicação de matrizes quadradas, de ordem M que é a quantidade de estados que a variável pode assumir. Quando a matriz de transição for diagonalizável o trabalho é razoavelmente facilitado, quando não, podemos utilizar algum programa de computador para realizar essa tarefa e para o caso de alunos do ensino médio essa segunda opção é a única possível.

Vejamos agora uma última definição que facilita o tratamento de uma cadeia de Markov, o vetor de probabilidade.

1.5 O vetor de probabilidade

Definição 1.5.1. *O* ***vetor de probabilidade****, indicado por $\pi^{(n)}$ é um vetor linha, cujo i-ésimo componente x_i é a probabilidade do sistema estar, naquela observação, no i-ésimo estado.*

Definição 1.5.2. *Chamamos de* ***vetor de probabilidade inicial*** *o vetor de probabilidade cujo n é igual a zero, ou seja, $\pi^{(0)}$.*

Vale salientar que sendo um vetor de probabilidades suas entradas serão todas não negativas e de soma igual a 1. Observe ainda que, cada uma das componentes x_i do vetor de probabilidade inicial $\pi^{(0)}$ representa a probabilidade da cadeia encontrar-se no estado s_i, no início do processo.

Agora note que com a definição do vetor de probabilidade e a matriz de transição P de uma cadeia de Markov, podemos escrever a equação de Chapman-Kolmogorov para $n \geq 1$ na forma matricial da seguinte maneira:

$$\pi^{(n)} = \pi^{(n-1)} \cdot P$$

O teorema que iremos enunciar e demonstrar abaixo, nos mostra que o vetor de distribuição de probabilidade após n-etapas, pode também ser calculado através do produto entre o vetor de probabilidade inicial $\pi^{(0)}$ e a n-ésima potência da matriz de transição P, vejamos:

Teorema 1.5.1. *Seja P a matriz de transição da cadeia de Markov e $\pi^{(0)}$ o vetor de probabilidade que representa sua distribuição inicial. Então a probabilidade de que a cadeia esteja no estado s_i depois de n passos, é a componente x_i do vetor $\pi^{(n)} = \pi^{(0)} \cdot P^n$*

Demonstração. Sabemos que pela forma matricial da equação de Chapman-Kolmogorov:

$$\pi^{(n)} = \pi^{(n-1)} \cdot P$$

Aplicando a equação de forma recursiva chegamos a:

$$\pi^{(n)} = \pi^{(n-2)} \cdot P \cdot P$$

$$\vdots$$

$$\pi^{(n)} = \pi^{(0)} \cdot \underbrace{P \cdot P \cdot P \cdots P}_{\text{n vezes}}$$

Donde podemos concluir que: $\pi^{(n)} = \pi^{(0)} \cdot P^n$. ∎

Em particular, se quisermos observar o comportamento da cadeia para um estado s_i específico, basta fazer com que o vetor de probabilidade inicial tenha componentes todas iguais a zero exceto à componente x_i, que terá valor igual a 1.

1.6 Exemplos

Podemos encontrar muitos exemplos de aplicações das Cadeias de Markov na literatura. Dentre eles podemos citar: problemas da área de Administração que envolvem a *Lealdade do Consumidor* em relação a determinados produtos ou ações do mercado financeiro. Encontramos ainda aplicações na área da Genética onde a matriz de transição da cadeia representa a probabilidade de numa reprodução a prole ter um gene do tipo dominante, recessivo ou híbrido. Ainda temos fenômenos físicos que podem ser modelados através de uma Cadeia de Markov, como é o caso do modelo de Ehrenfest que é usado para explicar como a difusão de gases ocorre num determinado ambiente.

No entanto escolhemos inicialmente um exemplo que envolve *o estado do tempo em uma cidade*, pois este é mais eficiente para nos ajudar a compreender alguns dos conceitos básicos apresentados.

1.6.1 Tempo

Considere uma sequência de variáveis aleatórias $X(t)$ cujos elementos pertencem ao espaço $E = \{Sem\ nuvens, Nublado, Chovendo\}$. A tabela 1.2 abaixo, nos mostra como esteve o tempo a cada dia em um certo mês em Salvador, iremos considerar essa nossa observação inicial.

Tabela 1.2: Registro do Tempo ao longo de um mês

Dia 1	**Dia 2**	**Dia 3**	**Dia 4**	**Dia 5**
Nublado	Sem nuvens	Nublado	Nublado	Chovendo
Dia 6	**Dia 7**	**Dia 8**	**Dia 9**	**Dia 10**
Sem nuvens	Sem nuvens	Nublado	Sem nuvens	Nublado
Dia 11	**Dia 12**	**Dia 13**	**Dia 14**	**Dia 15**
Nublado	Chovendo	Chovendo	Nublado	Chovendo
Dia 16	**Dia 17**	**Dia 18**	**Dia 19**	**Dia 20**
Chovendo	Sem nuvens	Sem nuvens	Sem nuvens	Nublado
Dia 21	**Dia 22**	**Dia 23**	**Dia 24**	**Dia 25**
Sem nuvens	Sem nuvens	Nublado	Nublado	Chovendo
Dia 26	**Dia 27**	**Dia 28**	**Dia 29**	**Dia 30**
Sem nuvens	Sem nuvens	Nublado	Sem nuvens	Nublado

Pela tabela 1.2, segue que $P_{Sem\ nuvens} = \frac{12}{30}$, $P_{Nublado} = \frac{12}{30}$ e $P_{Chovendo} = \frac{6}{30}$. Fazendo corresponder o estado Sem nuvens a 1, o estado Nublado a 2 e o estado Chovendo a 3 temos que o vetor de probabilidade inicial para este fenômeno seria dado por $\pi^{(0)} = (\frac{12}{30}, \frac{12}{30}, \frac{6}{30})$ ou $\pi^{(0)} = (0,40; 0,40; 0,20)$.

Suponha que o Instituto Nacional de Pesquisas Espaciais - INPE, tenha ao longo de 20 anos de observação gerado uma matriz de transição do tempo para a cidade de Salvador com as seguintes entradas:

$$\begin{array}{cc} & \begin{array}{ccc} \quad 1 & \quad 2 & \quad 3 \end{array} \\ \mathbf{P} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0,40 & 0,55 & 0,05 \\ 0,20 & 0,60 & 0,20 \\ 0,10 & 0,60 & 0,30 \end{bmatrix} \end{array}$$

(com rótulo *Estado* sobre as colunas)

Observando a matriz acima podemos ver que se hoje temos um tempo sem nuvens (*estado* 1) no céu há uma probabilidade de 55% de que amanhã teremos um dia nublado (*estado* 2), ou ainda, que se hoje está chovendo (*estado* 3) temos uma probabilidade de 30% que amanhã continue chovendo.

Como nossa variável é mensal ao multiplicar o vetor de probabilidade inicial $\pi^{(0)}$ pela matriz de transição P teremos o vetor de probabilidade $\pi^{(1)}$ que representa a distribuição de probabilidade para o próximo mês e assim sucessivamente.

Este exemplo é importante para o aprendizado das cadeias de Markov por dois motivos. O tempo em uma cidade é um fenômeno que de fato independe do seu estado há um ano atrás mas é influenciado pelo estado em que se encontra no dia anterior, ou seja, este é um fenômeno que possui a propriedade markoviana. O outro motivo é que a matriz de transição neste caso é simples de ser interpretada e tem um tamanho reduzido pois a quantidade de estados possíveis da variável é realmente pequeno.

A seguir iremos estudar o exemplo que envolve a *Lealdade do consumidor* sobre as compras que fazem durante o mês.

1.6.2 Lealdade do consumidor

O enunciado abaixo é uma versão adaptada do problema que encontra-se em Sullivan e Mizrahi (2006, p. 502), vejamos:

> *Em um determinado bairro de uma cidade existem 4 supermercados em que a população do bairro faz suas compras. Através de uma pesquisa feita em 1 de janeiro foi determinado que* 25% *da população compra no supermercado 1,* 32% *no supermercado 2,* 23% *no supermercado 3* e 20% *no supermercado 4. Além disso, sabe-se ainda que a cada mês o supermercado 1 mantém* 75% *dos seus clientes e perde* 10% *para o supermercado 2,* 5% *para o supermercado 3,* 10% *para o supermercado 4. O supermercado 2 mantém* 60% *dos seus clientes e perde* 8% *para o supermercado 1,* 15% *para o supermercado 3,* 17% *para o supermercado 4. O supermercado 3 mantém* 50% *dos seus clientes e perde* 10% *para o supermercado 1,* 20% *para o supermercado 2,* 20% *para o supermercado 4. Por fim o supermercado 4 mantém* 80% *dos seus clientes e perde* 10% *para o supermercado 2,* 10% *para o supermercado 3.*

O vetor de probabilidade inicial neste caso é dado por $\pi^{(0)} = (0,25; 0,32; 0,23; 0,20)$ e para facilitar a montagem da matriz de transição P, façamos inicialmente um diagrama de estados com os dados do problema e em seguida sua matriz de transição.

$$\mathbf{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0,75 & 0,10 & 0,05 & 0,10 \\ 0,08 & 0,60 & 0,15 & 0,17 \\ 0,10 & 0,20 & 0,50 & 0,20 \\ 0 & 0,10 & 0,10 & 0,80 \end{bmatrix} \end{array}$$

O processo descrito no problema é uma cadeia de Markov. Assim podemos saber qual o vetor seguinte de distribuição de probabilidades de clientes em cada mercado ao longo do tempo usando a equação $\pi^{(n+1)} = \pi^{(n)} \cdot P$. Na tabela abaixo podemos visualizar esses cálculos para os próximos 31 meses, com aproximação de 4 casas decimais:

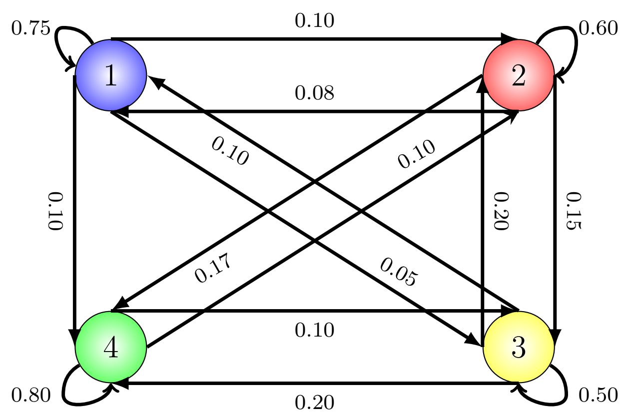


Figura 1.5: *Diagrama de estados - Lealdade do consumidor*

Tabela 1.3: Vetor de probabilidade - Lealdade do Consumidor

Vetor	**Valor**	**Vetor**	**Valor**
$\pi^{(0)}$	(0,2500; 0,3200; 0,2300; 0,2000)	$\pi^{(16)}$	(0,1460; 0,2348; 0,1740; 0,4453)
$\pi^{(1)}$	(0,2361; 0,2830; 0,1955; 0,2854)	$\pi^{(17)}$	(0,1457; 0,2348; 0,1740; 0,4455)
$\pi^{(2)}$	(0,2193; 0,2611; 0,1805; 0,3391)	$\pi^{(18)}$	(0,1454; 0,2348; 0,1741; 0,4457)
$\pi^{(3)}$	(0,2034; 0,2486; 0,1743; 0,3737)	$\pi^{(19)}$	(0,1453; 0,2348; 0,1741; 0,4458)
$\pi^{(4)}$	(0,1899; 0,2417; 0,1720; 0,3964)	$\pi^{(20)}$	(0,1451; 0,2348; 0,1741; 0,4459)
$\pi^{(5)}$	(0,1789; 0,2381; 0,1714; 0,4116)	$\pi^{(21)}$	(0,1451; 0,2348; 0,1741; 0,4460)
$\pi^{(6)}$	(0,1704; 0,2362; 0,1715; 0,4219)	$\pi^{(22)}$	(0,1450; 0,2348; 0,1741; 0,4461)
$\pi^{(7)}$	(0,1638; 0,2352; 0,1719; 0,4290)	$\pi^{(23)}$	(0,1449; 0,2348; 0,1741; 0,4461)
$\pi^{(8)}$	(0,1589; 0,2348; 0,1723; 0,4340)	$\pi^{(24)}$	(0,1449; 0,2348; 0,1742; 0,4461)
$\pi^{(9)}$	(0,1552; 0,2346; 0,1727; 0,4375)	$\pi^{(25)}$	(0,1449; 0,2348; 0,1742; 0,4461)
$\pi^{(10)}$	(0,1524; 0,2346; 0,1731; 0,4399)	$\pi^{(26)}$	(0,1449; 0,2348; 0,1742; 0,4461)
$\pi^{(11)}$	(0,1504; 0,2346; 0,1733; 0,4417)	$\pi^{(27)}$	**(0,1448; 0,2348; 0,1742; 0,4462)**
$\pi^{(12)}$	(0,1489; 0,2346; 0,1735; 0,4429)	$\pi^{(28)}$	(0,1448; 0,2348; 0,1742; 0,4462)
$\pi^{(13)}$	(0,1478; 0,2347; 0,1737; 0,4438)	$\pi^{(29)}$	(0,1448; 0,2348; 0,1742; 0,4462)
$\pi^{(14)}$	(0,1470; 0,2347; 0,1738; 0,4445)	$\pi^{(30)}$	(0,1448; 0,2348; 0,1742; 0,4462)
$\pi^{(15)}$	(0,1464; 0,2347; 0,1739; 0,4449)	$\pi^{(31)}$	(0,1448; 0,2348; 0,1742; 0,4462)

Na tabela 1.3 destacamos em negrito o vetor $\pi^{(27)}$, pois a partir dele, para uma aproximação de 4 casas decimais, os vetores de distribuição de probabilidades continuam iguais, ou seja, todo vetor $\pi^{(n)}$ tal que $n > 27$ terá a mesma distribuição de probabilidade que ele.

Observe que podemos interpretar este fato da seguinte maneira: se as pessoas deste bairro continuarem a frequentar os mercados segundo a matriz de transição dada, após vinte e sete meses teremos: $14,48\%$ dos moradores comprando no supermercado 1, $23,48\%$ dos moradores comprando no supermercado 2, $17,42\%$ dos moradores comprando no supermercado 3 e $44,62\%$ dos moradores comprando no supermercado 4. Este vetor é chamado de *vetor de probabilidade estacionário*.

Façamos agora o mesmo cálculo utilizando o teorema 1.5.1, ou seja, encontraremos

$\pi^{(n)}$, multiplicando o vetor de probabilidade inicial $\pi^{(0)}$ pelas enésimas potências da matriz de transição P. Os cálculos completos encontram-se no apêndice 1.

No entanto, o que nos interessa neste momento é perceber que a partir da trigésima primeira potência de P, com uma aproximação de 4 casas decimais o vetor estacionário aparece em cada linha da matriz de transição, ou seja, podemos calcular o vetor de distribuição estacionário apenas fazendo as potências de P.

Essa é uma propriedade muito interessante de certos tipos de Cadeias de Markov denominadas *regulares*. Essas cadeias possuem matriz de transição que se caracterizam por ter um único vetor de probabilidade estacionário e, além disso, este independe do vetor de distribuição inicial, porque como vimos podem ser calculados apenas utilizando a matriz de transição da cadeia. Formalmente definimos uma matriz de transição como regular da seguinte maneira.

Definição 1.6.1 (Cadeias de Markov Regulares)**.** *Seja P uma matriz de transição de uma cadeia de Markov. A matriz P é chamada regular se alguma potência positiva de P tem todas as suas entradas positivas.*

Para maiores informações sobre Cadeias de Markov regulares recomendamos a leitura de Poole (2004).

1.6.3 Nível econômico

O exemplo abaixo vem da sociologia. Segundo esta área do conhecimento o nível econômico de uma pessoa é fortemente influenciado pela classe econômica dos seus pais. Façamos a classificação do nível econômico de um indivíduo em três categorias: rico (R), classe média (M), e pobre (P).

Supondo que cada homem tem um filho e que observando sucessivas gerações das famílias desses homens descobrimos que dos filhos de um homem rico 95% continuam ricos e 5% se tornam classe média. No caso de um indivíduo de classe média, 10% ficam ricos, 70% se mantém na classe média e 20% ficam pobres. E, finalmente, no caso de um homem pobre, 30% vão para classe média e 70% permanecem pobres. Então a matriz de transição neste caso seria dada por:

$$\mathbf{P} = \begin{array}{c} \\ R \\ M \\ P \end{array} \begin{array}{c} \begin{array}{ccc} R & M & P \end{array} \\ \begin{bmatrix} 0,95 & 0,05 & 0 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{bmatrix} \end{array}$$

Porém na realidade nem todo homem, necessariamente, tem um filho. Vamos supor que a probabilidade de que um homem venha a ter um filho seja de 90%. Podemos então formar uma nova cadeia de Markov com quatro estados, onde o estado delta representa o estado de não ter filho.

$$\mathbf{K} = \begin{array}{c} \\ R \\ M \\ P \\ \Delta \end{array} \begin{array}{c} \begin{array}{cccc} R & M & P & \Delta \end{array} \\ \begin{bmatrix} 0,855 & 0,045 & 0 & 0,10 \\ 0,09 & 0,63 & 0,18 & 0,10 \\ 0 & 0,27 & 0,63 & 0,10 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Note que o estado sem filhos é *absorvente*, ou seja, se um homem de uma geração para outra não tem filho, isso significa que não existe a possibilidade que chegando a este estado ele saia do mesmo. Por isso a probabilidade de permanecer nesse estado ao chegar nele é de 100%. São Cadeias de Markov que possuem essa característica que iremos estudar no próximo capítulo.

Capítulo 2

Cadeias de Markov Absorventes

De forma geral as Cadeias de Markov são estudadas a partir de características específicas que elas possuem em seus estados ou em sua matriz de transição. Como se pode ver em Poole (2004, p.294) as Cadeias de Markov *regulares* possuem um vetor de estado estacionário, pois sua matriz de transição é regular. No caso das cadeias de Markov Absorventes iremos nos ocupar em responder para um dado estado inicial s_i, não absorvente, as seguintes questões:

- Em média, quantas vezes o processo visitará cada um dos estados não absorventes?
- Qual o número de etapas esperado antes da absorção?
- Qual é a probabilidade de que um estado absorvente seja alcançado?

Para responder tais questões, iremos inicialmente definir formalmente as possíveis características que um estado $s_i \in E$ pode ter. As definições na próxima seção são baseadas naquelas encontradas no livro de Hillier e Lieberman (2013).

2.1 Classificação de Estados

Para todas as definições abaixo considere que s_i e $s_j \in E$, onde E é o espaço de estados de uma cadeia de Markov.

Definição 2.1.1. *Dizemos que um estado s_j é **acessível** a partir do estado s_i se $p_{ij}^{(n)} > 0$ para algum $n \geq 1$ e escrevemos $s_i \to s_j$.*

Ou seja, o estado s_j ser acessível a partir do estado s_i significa que é possível ao sistema entrar eventualmente no estado s_j quando ele se encontra no estado s_i.

No exemplo do tempo na subseção 1.6.1 do capítulo anterior, temos que todos os estados são acessíveis a partir de qualquer um dos outros, pois para todo s_i e s_j, temos $p_{ij} > 0$. No exemplo que envolve a Lealdade do Consumidor, para $n \geq 2$, todos os estados também são acessíveis a partir de todos os outros visto que para todo s_i e s_j, temos que $p_{ij}^{(2)} > 0$. Entretanto, no exemplo do nível econômico com 4 estados, nenhum dos outros estados, diferentes de Δ, são acessíveis a partir do estado *Delta*.

Definição 2.1.2. *Se o estado s_j for acessível a partir do estado s_i e o estado s_i for acessível a partir do estado s_j, então dizemos que os estados s_i e s_j se* ***comunicam*** *e escrevemos $s_i \leftrightarrow s_j$.*

Observe que a relação $\leftrightarrow$ para estados que se comunicam, assim definida é uma relação de equivalência, pois:

i) Qualquer estado se comunica consigo mesmo; *(Reflexiva)*

ii) Se o estado s_i se comunica com o estado s_j, então o estado s_j se comunica com o estado s_i;*(Simétrica)*

iii) Se o estado s_i se comunica com o estado s_j e o estado s_j se comunica com o estado s_k, então o estado s_i se comunica com o estado s_k.*(Transitiva)*

O item 1 se justifica pois $p_{ii}^{(0)} = P\{X_0 = i | X_0 = i\} = 1$. O item 2 é verificado pela própria definição 2.1.2. O item 3 é demonstrado da seguinte maneira.

Demonstração. Suponhamos que $s_i \to s_j$ e $s_j \to s_k$, isto significa que $\exists n \geq 1$ e $m \geq 1$, tais que $p_{ij}^{(n)} > 0$ e $p_{jk}^{(m)} > 0$. Sabemos que pela equação de Chapman-Kolmogorov que existe um estado s, tal que:

$$p_{ik}^{(n+m)} = \sum_{s=1}^{M} p_{is}^{(n)} \cdot p_{sk}^{(m)}$$

Mas,

$$\sum_{s=1}^{M} p_{is}^{(n)} \cdot p_{sk}^{(m)} \geq p_{ij}^{(n)} \cdot p_{jk}^{(m)} > 0$$

Logo $s_i \leftrightarrow s_k$. ∎

Por ser uma relação de equivalência os estados de uma cadeia de Markov podem ser subdivididos em **classes** distintas, tais que aqueles estados que comunicam-se entre si estão na mesma classe. Uma classe pode ser formada por um único estado. Se uma cadeia de Markov posuir uma única classe, isto é, todos os estados comunicarem-se, não necessariamente em um único passo, a cadeia de Markov é chamada **irredutível** ou **ergôdica**.

Um fato curioso de se notar é que por definição uma cadeia de Markov regular é irredutível, mas nem toda matriz irredutível é regular. Um exemplo clássico desta ocorrência é a cadeia cuja matriz de transição é dada por:

$$
\begin{array}{cc}
 & \begin{array}{cc} 1 & 2 \end{array} \\
\mathbf{P} = \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
\end{array}
$$

Observe que podemos acessar o estado 2 a partir do estado 1 e acessar o estado 1 a partir do estado 2 portanto a matriz é irredutível. No entanto as entradas das potencias de P nunca serão todas positivas. Se n for par a matriz será igual a I_2 e se n for ímpar a matriz será igual a P.

Definição 2.1.3. *Um estado é dito **transiente** se, após entrar nesse estado, existe a possibilidade do processo jamais retornar a esse estado novamente.*

Portanto o estado s_i é transiente se, e somente se, existir um estado $s_j (s_j \neq s_i)$ que seja acessível do estado s_i, mas não o contrário, isto é o estado s_i não é acessível a partir do estado s_j. Note que se um estado é transiente isto significa que ele será visitado apenas um número finito de vezes, pois existe a possibilidade de que ele passe para o estado s_j e não retorne mais para o estado s_i.

É o caso, por exemplo, dos compartimentos 1, 2, 3 e 4 do exemplo 1.3.1 do labirinto do rato visto que é possível que estes estados nunca mais sejam visitados pelo rato. Pois ao adentrar no Paraíso do Queijo ele não mais sairá de lá. Uma outra possibilidade para um estado de uma cadeia de Markov é dado na próxima definição.

Definição 2.1.4. *Um estado é dito ser **recorrente** se, após adentrar aquele estado, o processo <u>com certeza retornar</u> a esse estado novamente. Portanto um estado é recorrente se, e somente se, ele não for transiente.*

Já que um estado recorrente será revisitado após cada visita, ele retornará com frequência infinita para esse estado caso o processo se mantenha eternamente. Os estados no exemplo do clima e da lealdade do consumidor são recorrentes visto que eles sempre ocorrerão caso o processo seja infinito.

Definição 2.1.5. *Diz-se que um estado é **absorvente** caso, após adentrar esse estado, o processo jamais deixará esse estado novamente. Portanto o estado s_i é um estado absorvente se, e somente se, $p_{ii} = 1$.*

No exemplo do nível econômico como já tínhamos antecipado, o estado Δ(não ter filhos) é um estado absorvente. Visto que ao adentrar neste estado o processo jamais o deixará. Como podemos ver em Hillier e Lieberman (2013, p. 706), temos que:

A recorrência é uma propriedade de classe, ou seja, todos os estados em uma classe são recorrentes ou então transientes. Além disso, em uma cadeia de Markov de estados finitos, nem todos os estados podem ser transientes. Assim, todos os estados em uma cadeia de Markov de estados finitos irredutíveis são recorrentes.

2.2 A Forma Canônica

Como já definimos o que são estados transientes e absorventes podemos definir o que é uma cadeia de Markov Absorvente, vejamos:

Definição 2.2.1. *Uma cadeia de Markov é chamada de* ***absorvente*** *se, e somente se, ela contiver pelo menos um estado absorvente e se for possível ir de qualquer estado não absorvente para um estado absorvente em uma ou mais etapas.*

Em outras palavras dizemos que uma cadeia de Markov é absorvente quando algum dos seus estados é capaz de capturar o processo e não permitir que nenhum outro estado seja acessado. Vale salientar que a exigência de que os estados não absorventes acessem os estados absorventes pode ocorrer em mais de uma etapa. Vejamos um exemplo para esclarecer a definição acima.

Considere duas cadeias de Markov cujas matrizes de transição são P_1 e P_2.

$$\mathbf{P_1} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{bmatrix} 0,3 & 0,6 & 0,1 & 0 \\ 0 & 1 & 0 & 0 \\ 0,2 & 0 & 0,4 & 0,4 \\ 0,3 & 0 & 0,3 & 0,4 \end{bmatrix} \end{array} \qquad \mathbf{P_2} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \end{array} \begin{array}{c} \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0,4 & 0,6 \\ 0 & 0,2 & 0,8 \end{bmatrix} \end{array}$$

A cadeia cuja matriz de transição é P_1 é absorvente, pois o estado E_2 é absorvente e é possível através de E_1 acessar E_2. Observe ainda que E_3 e E_4 acessam E_1 diretamente, logo esses estados também acessam indiretamente E_2. Podemos observar claramente que isso ocorre com os cálculos efetuados no apêndice 1 seção 2.

No entanto, a cadeia cuja matriz de transição é dada por P_2 não é absorvente, porque apesar do estado E_1 ser absorvente não é possível acessá-lo nem a partir de E_2 nem a partir de E_3, ou seja estes estados serão recorrentes. Assim se o processo tiver início em algum desses estados nunca ocorrerá a absorção.

Considere agora uma cadeia de Markov absorvente qualquer, lembre-se que a mesma será composta de r estados absorventes e t estados transientes. Dessa forma podemos reorganizar a matriz de transição da cadeia de modo que os estados transientes apareçam primeiro e depois os estados absorventes, fazendo isso teremos a seguinte configuração.

$$\mathbf{P} = \begin{array}{c c} & \begin{array}{cc} Trans & Absor \end{array} \\ \begin{array}{c} Trans \\ Absor \end{array} & \left[\begin{array}{c|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbf{I_r} \end{array}\right] \end{array}$$

Note que $\mathbf{I_r}$ é uma matriz identidade de ordem r, a matriz $\mathbf{0}$ é uma matriz nula com r linhas e t colunas, $\mathbf{R}$ é uma matriz não nula com t linhas e r colunas e $\mathbf{Q}$ é uma matriz quadrada de ordem t. Os primeiros t estados serão transientes e os r estados restantes serão absorventes.

É importante perceber o papel de cada uma dessas matrizes para a dinâmica do processo. A matriz Q nos diz quais as probabilidades de transição entre estados transientes. A matriz R é responsável pelo acesso de estados transientes aos estados absorventes. A matriz nula nos informa que é impossível sair de um estado absorvente para um estado transiente e a matriz identidade nos indica que ao chegar num estado absorventes lá iremos permanecer.

Agora vejamos o que ocorre com a probabilidade de transição entre estados transientes a medida que à cadeia passa para etapas seguintes. Sabemos que para isso basta que elevemos a matriz P a potência n, onde n representa a quantidade de etapas que queremos avançar. Assim, temos que:

$$\mathbf{P^{(n)}} = \begin{array}{c c} & \begin{array}{cc} Trans & Absor \end{array} \\ \begin{array}{c} Trans \\ Absor \end{array} & \left[\begin{array}{c|c} \mathbf{Q^n} & * \\ \hline \mathbf{0} & \mathbf{I_r} \end{array}\right] \end{array}$$

Lembre que Q^n nos diz qual a probabilidade de se iniciar em um dos r estados transientes e continuar em um desses estados transientes após n etapas. Nosso primeiro importante teorema da próxima seção prova que se a cadeia inicia em um estado transiente a probabilidade de permanecer em um estado transiente quando n tende ao infinito é zero. Assim todas as entradas de Q^n se aproximam de zero quando n tende ao infinito, o que provaremos abaixo:

Teorema 2.2.1. *Em uma cadeia de Markov absorvente, a probabilidade de que o processo seja absorvido é igual a 1.* (*i.e.* $n \to \infty \Rightarrow Q^n \to 0$)

Demonstração. Por definição se a cadeia de Markov é absorvente então é possível a cada estado transiente s_j acessar um estado absorvente. Seja m_j o número mínimo de etapas necessárias para que se chegue a um estado absorvente partindo de s_j. Seja p_j a probabilidade de que, partindo de s_j, o processo não alcance um estado absorvente em m_j etapas, ou seja, p_j é a probabilidade de que o processo permaneça num estado transiente após m_j etapas. Note que p_j é uma probabilidade logo $p_j < 1$. Seja m o maior dos m_j e seja p o maior dos p_j, isto é, m é a maior quantidade de etapas que são necessárias para que todos os estados acessem um estado absorvente e p é a maior probabilidade de que um desses estágios permaneçam num estado transiente. Portanto a probabilidade de não ser absorvido em m etapas é menor ou igual a p logo em $2m$ etapas será menor ou igual a p^2, e assim sucessivamente. Como $p < 1$ esta probabilidade tende a 0. Portanto a probabilidade de não ser absorvido em n etapas é monotonicamente decrescente, logo as probabilidades de todos os p_j também tendem a 0, consequentemente $\lim_{n \to \infty} \mathbf{Q}^n = 0$. ■

Já temos agora os elementos necessários para começar a responder as questões propostas no início do capítulo. Na próxima seção iremos definir a matriz fundamental de uma cadeia de Markov absorvente que é a chave para descobrir as respostas às questões feitas.

2.3 A matriz fundamental

Uma das questões que queremos responder é: *Em média, quantas vezes a cadeia de Markov absorvente visitará cada um dos estados não absorventes?* Para responder a esta pergunta vamos definir o que é a matriz fundamental de uma cadeia de Markov absorvente, vejamos:

Definição 2.3.1. *Para uma cadeia de Markov absorvente chamamos de* ***matriz fundamental F*** *a matriz inversa de* $(I_t - Q)$.

O índice t da matriz identidade corresponde a quantidade de estados transientes do processo. As entradas dessa matriz **F** nos diz qual o número esperado de vezes que a cadeia estará num estado s_j, dado que a cadeia partiu de um estado s_i. Para provar este teorema iremos primeiro demonstrar o lema abaixo.

Lema 2.3.1. *A matriz fundamental* $\boldsymbol{F}$ *existe e* $F = I+Q+Q^2+Q^3+\cdots$, *quando* $n \to \infty$.

Demonstração. Considere a igualdade:

$$(I - Q) \cdot (I + Q + Q^2 + Q^3 + \cdots + Q^{n-1}) = I - Q^n$$

Sabemos que pelo teorema 2.2.1 se $n \to \infty$ então $Q^n \to 0$ então quando $n \to \infty$ o segundo membro da igualdade acima tende para a matriz identidade I. Logo se aplicarmos o determinante a ambos os membros vemos que quando $n \to \infty$ o determinante do primeiro membro da igualdade tende a 1, ou seja,

$$|(I - Q) \cdot (I + Q + Q^2 + Q^3 + \cdots + Q^{n-1})| \to 1$$

Então para um n suficientemente grande temos que o determinante de $I - Q^n$ é não-nulo. Além disso, sabemos que o determinante do produto de duas matrizes é igual ao produto dos determinantes da mesma, logo para n suficientemente grande temos:

$$|(I - Q)| \cdot |(I + Q + Q^2 + Q^3 + \cdots + Q^{n-1})| \neq 0$$

Portanto o determinante da matriz $(I - Q)$ é diferente de zero, consequentemente ela possui uma matriz inversa desde que n seja suficientemente grande. Implicando que a matriz $F = (I - Q)^{-1}$ existe para um n suficientemente grande.

Como $F = (I - Q)^{-1}$ existe, podemos multiplicar ambos os lados da igualdade por F, façamos isso:

$$F \cdot (I - Q) \cdot (I + Q + Q^2 + Q^3 + \cdots + Q^{n-1}) = F \cdot (I - Q^n)$$

Mas $F \cdot (I - Q) = I$, portanto no primeiro membro da igualdade temos apenas $(I + Q + Q^2 + Q^3 + \cdots + Q^{n-1})$. Quando $n \to \infty$ temos $Q^n \to 0$, assim o segundo membro da igualdade fica $F \cdot I = F$ e portanto $F = I + Q + Q^2 + Q^3 + \cdots$, quando $n \to \infty$. ■

Teorema 2.3.1. *A f_{ij} entrada da matriz F é igual ao número de vezes que a cadeia visitará o estado s_j, dado que começa no estado s_i.*

Demonstração. Sejam s_i e s_j dois estados transientes de tal modo que os valores i e j sejam fixos. Seja $X^{(k)}$ uma variável aleatória que é igual a 1 se à cadeia está no estado s_j após k etapas e igual a zero nos outros casos. Temos que:

$$P(X^{(k)} = 1) = q_{ij}^{(k)} \quad e \quad P(X^{(k)} = 0) = 1 - q_{ij}^{(k)}$$

Onde $q_{ij}^{(k)}$ é a ij-ésima entrada de Q^k. Para manter a coerência nessas equações basta definir que para $k = 0$ temos $Q^0 = I$. Assim sendo $X^{(k)}$ é uma variável aleatória composta por zeros e uns de modo que o valor esperado $E(X^{(k)}) = q_{ij}^{(k)}$.

O número esperado de vezes que a cadeia visitará o estado s_j em n etapas, dado que iniciou no estado s_i é:

$$E(X^{(0)} + X^{(1)} + \cdots + X^{(n)}) = q_{ij}^{(0)} + q_{ij}^{(1)} + \cdots + q_{ij}^{(n)}$$

Fazendo $n \to \infty$ temos que

$$E(X^{(0)} + X^{(1)} + \cdots) = q_{ij}^{(0)} + q_{ij}^{(1)} + \cdots = f_{ij}$$

. ■

Vamos considerar agora a segunda questão proposta no início do capítulo: *Dado* *que a cadeia iniciou no estado s_i quantas etapas são necessárias para que entre em um estado absorvente?* A resposta será dada no teorema seguinte.

Teorema 2.3.2. *Seja t_i o número de etapas antes que a cadeia seja absorvida, dado que a cadeia iniciou no estado s_i e seja t um vetor coluna cujas entradas sejam t_i então $t = F.c$ onde c é um vetor coluna com todas as entradas iguais a 1.*

Demonstração. Se adicionarmos todas as entradas da i-ésima linha de **F**, nós teremos o número esperado de vezes que o processo irá visitar qualquer um dos estados transientes, visto que iniciou em s_i, isto é, o número de vezes que o processo visitará aquele estado antes da absorção. Assim t_i é a soma das entradas da i-ésima linha de **F**. Se escrevermos esta afirmação na forma matricial segue o teorema, ou seja, $t = F.c$. ■

Nosso último teorema responde a última questão que fizemos no início do capítulo, ou seja, nos diz qual a probabilidade de uma cadeia de Markov absorvente ser absorvida por um dos seus estados absorventes s_j dado que ela iniciou em um estado transiente s_i. Segue abaixo o enunciado e demonstração deste teorema.

Teorema 2.3.3. *Seja b_{ij} a probabilidade de que uma cadeia de Markov absorvente seja absorvida por um estado s_j dado que ela iniciou num estado transiente s_i. Seja $\boldsymbol{B}$ a matriz com entradas b_{ij}, então $\boldsymbol{B}$ tem ordem t por r e $B = F \cdot R$, onde t é a quantidade de estados transientes, r é a quantidade de estados absorventes, F é a matriz fundamental e R é matriz de ligação entre estados transientes e absorventes na forma canônica.*

Demonstração. Note que começando o processo num estado transiente s_i ele pode ser absorvido por s_j em uma ou mais etapas. A probabilidade de ser absorvido em um único passo é dado r_{ij}. Se isso não acontece, o processo pode se mover para um outro estado absorvente (e neste caso é impossível a ele acessar o estado s_j) ou se mover para um estado transiente s_k. Neste último caso há uma probabilidade b_{kj} de ser absorvido do estado s_k para o estado s_j. Portanto temos que:

$$b_{ij} = r_{ij} + \sum_k q_{ik} \cdot b_{kj}$$

A equação acima pode ser escrita na forma matricial da seguinte maneira:

$$\mathbf{B} = \mathbf{R} + \mathbf{QB}$$

Daí,

$$\mathbf{B} - \mathbf{QB} = \mathbf{R} \Rightarrow$$

$$\mathbf{B} \cdot (\mathbf{I} - \mathbf{Q}) = \mathbf{R} \Rightarrow$$

$$\mathbf{B} = (\mathbf{I} - \mathbf{Q})^{-1} \cdot \mathbf{R} \Rightarrow$$

$$\mathbf{B} = \mathbf{F} \cdot \mathbf{R}$$

■

Estas definições e teoremas nos fornecem as ferramentas necessárias para trabalhar com cadeias de Markov absorventes, na próxima seção daremos alguns exemplos do funcionamento deste tipo de cadeia de Markov.

2.4 Exemplos

Usaremos como exemplos de cadeias de Markov absorventes duas situações que nos parecem esclarecer muito bem as definições e teoremas vistos na seção anterior, são eles: *A ruína do jogador e a gestão médica em caso de pedras na vesícula.*

2.4.1 A ruína do jogador

Considere a seguinte situação que descreve um jogo de apostas entre duas pessoas:

Maria tem R$ 3,00 e João tem R$ 4,00 eles irão jogar uma moeda honesta e se der cara Maria perde R$ 1,00 para João e se der coroa João perde R$ 1,00 para Maria. O jogo termina quando um deles não tiver mais dinheiro para apostar.

Façamos um diagrama para entender melhor como funciona o jogo entre estas duas pessoas. Observe que nesta situação o jogador pode ter de zero a sete reais e que cada uma dessas possibilidades irá representar um possível estado em que o jogador se encontra. Além disso, a probabilidade de transição entre estados será de 50%, visto que a aposta será definida através do lançamento de uma moeda honesta. Assim nosso diagrama terá 8 nós e vejamos como ficarão as ligações entre eles na figura abaixo:

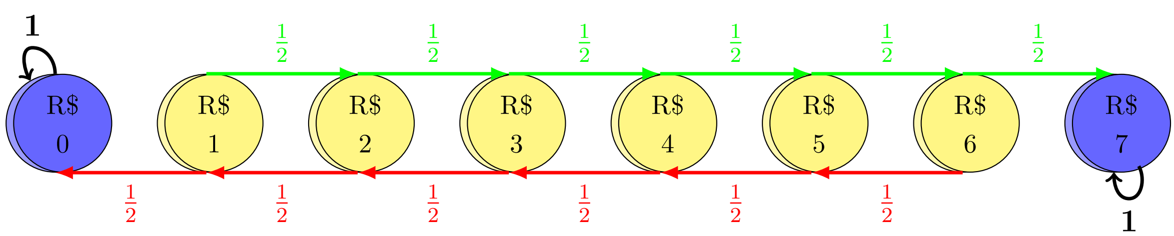


Figura 2.1: *Diagrama de estados - Ruína do Jogador*

Olhando o diagrama fica claro que os estados 0 e 7 são absorventes, visto que o jogo termina quando um dos jogadores perde todo o seu dinheiro e obviamente se ele perdeu todo o seu dinheiro a outra o ganhou e chegará ao estado 7 e daí não sairá. Note ainda que os estados 1, 2, 3, 4, 5, 6 são transientes, pois em algum momento nunca voltaremos a eles. Também queremos deixar em evidência, no diagrama, a ideia de classe a partir das cores dos nós. Façamos agora a matriz de transição para esta cadeia.

$$\mathbf{P} = \begin{array}{c} \\ \mathbf{0} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{7} \end{array} \begin{array}{c} \begin{array}{cccccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} \end{array} \\ \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,5 & 0 & 0,5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,5 & 0 & 0,5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \end{array}$$

Podemos verificar com a matriz de transição P que os estados transientes adjacentes se comunicam e que é possível acessar os estados absorventes através dos estados 1 e 6. Logo a cadeia em questão é absorvente, pois possui estados absorventes e esses são acessíveis a todos os estados transientes. Escrevamos agora P em sua forma canônica, isto é, vamos reorganizar a matriz de modo que os estados transientes apareçam a esquerda e em cima.

$$\mathbf{P} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \\ \mathbf{0} \\ \mathbf{7} \end{array} \begin{array}{c} \begin{array}{cccccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{0} & \mathbf{7} \end{array} \\ \left[\begin{array}{cccccc|cc} 0 & 0,5 & 0 & 0 & 0 & 0 & 0,5 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,5 & 0 & 0,5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,5 & 0 & 0,5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 0 & 0 & 0,5 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \end{array}$$

Com a forma canônica, podemos calcular a matriz fundamental $F = (I_6 - Q)^{-1}$, onde I_6 é a matriz identidade de ordem 6 e Q é a matriz quadrada cujas entradas são as probabilidades de que um estado transiente permaneça em um estado transiente, ou seja, a matriz quadrada de ordem 6 posicionada no canto esquerdo superior.

É importante perceber que neste caso já não temos uma tarefa simples ao tentar determinar a matriz F. Porque precisamos determinar a matriz inversa de uma matriz quadrada de ordem 6. Contudo, com o uso do Geogebra isto pode ser feito sem maiores dificuldades, ensinamos todo o processo de forma detalhada no vídeo dedicado a esta seção no site. Fazendo isso achamos então que a matriz fundamental F tem as seguintes entradas.

$$
\mathbf{F} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \end{array}
\begin{array}{c}
\begin{array}{cccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} \end{array} \\
\begin{bmatrix}
1,714 & 1,429 & 1,143 & 0,857 & 0,571 & 0,286 \\
1,429 & 2,857 & 2,286 & 1,714 & 1,143 & 0,571 \\
1,143 & 2,286 & 3,429 & 2,571 & 1,714 & 0,857 \\
0,857 & 1,714 & 2,571 & 3,429 & 2,286 & 1,143 \\
0,571 & 1,143 & 1,714 & 2,286 & 2,857 & 1,429 \\
0,286 & 0,571 & 0,857 & 1,143 & 1,429 & 1,714
\end{bmatrix}
\end{array}
$$

Como vimos as entradas f_{ij} da matriz F representa a quantidade de vezes, em média, que iniciando no estado s_i a cadeia irá passar pelo estado s_j antes de ser absorvido. Na matriz F acima podemos ver, por exemplo, que se um jogador iniciar a partida com R$ 2,00, ele deve ficar com R$ 3,00 em 2,286 etapas antes de perder todo seu dinheiro ou ganhar todo o dinheiro de seu adversário. Ainda podemos através da matriz fundamental F saber quanto o jogo duraria a partir da quantia inicial de um jogador, para isto basta somar as entradas em cada linha da matriz que obteremos esta quantidade, assim temos:

$$
\mathbf{t} = \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \end{array}
\begin{bmatrix} 6 \\ 10 \\ 12 \\ 12 \\ 10 \\ 6 \end{bmatrix}
$$

A matriz t nos informa que, em média, se uma pessoa iniciar o jogo com R$ 5,00 deve jogar 10 rodadas até que o jogo se encerre. Ou ainda, no caso do jogo entre João e Maria a partida deve ser encerrada após 12 jogadas.

A última informação que podemos obter sobre esta situação, utilizando os teoremas da seção anterior, é a probabilidade de perder ou ganhar o jogo quando iniciamos a mesma com uma determinada quantia. Como vimos, podemos obter essa informação fazendo o produto entre a matriz fundamental F e a matriz R que é a matriz que na forma canônica ocupa o canto superior direito, fazendo isso obtemos a matriz B igual a:

$$\mathbf{B} = \begin{array}{c} \\ \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \\ \mathbf{4} \\ \mathbf{5} \\ \mathbf{6} \end{array} \begin{array}{c} \begin{array}{cc} \mathbf{0} & \mathbf{7} \end{array} \\ \begin{bmatrix} 0,857 & 0,143 \\ 0,714 & 0,286 \\ 0,571 & 0,429 \\ 0,429 & 0,571 \\ 0,286 & 0,714 \\ 0,143 & 0.857 \end{bmatrix} \end{array}$$

Observe que b_{27} vale 28,6 %, ou seja, essa é a probabilidade de se iniciar uma partida com a quantia de R$ 2,00 e ser o vencedor da partida. Também podemos observar que uma pessoa que inicie o jogo com R$ 6,00 tem uma probabilidade de apenas 14,3% de sair perdedor.

Algumas observações precisam ser feitas sobre os resultados encontrados. A primeira é sobre a simetria desses resultados, isso só acontece devido a probabilidade de transição em todos os casos serem iguais a $\frac{1}{2}$. Ademais o tamanho da matriz de transição da cadeia cresce muito a medida que os valores das apostas aumentam. Vamos agora a outro exemplo interessante que pode ser modelado por uma cadeia de Markov absorvente.

2.4.2 Pedra na vesícula

Este exemplo encontra-se em Lay (1999) e aborda um dilema de médicos que tratam de pacientes com pedras na vesícula. Os médicos que diagnosticam precocemente esta doença vivem o dilema entre retirar imediatamente, através de uma cirurgia, as pedras para prevenir possíveis complicações que ameacem a vida do paciente ou de postergar esta cirurgia até que as complicações ocorram para daí realizá-la. Na ausência de um estudo clínico mais rigoroso, um médico poderia utilizar as cadeias de Markov para auxiliar a avaliação de riscos e benefícios para o paciente, se o cenário apresentado fosse como o descrito no texto abaixo:

> *Suponha que o paciente continuará assintomático (estado A) em 95% dos casos, durante um período de 4 meses. Pode ocorrer grandes complicações (estado C) no estado de saúde do paciente como uma colecistite, o que leva a uma intervenção cirúrgica em 4% dos casos, ou a morte (estado M), em pacientes de idades específicas, em 1% dos casos. Se a doença progride e se torna sintomática o paciente tem um risco de morte, causado por complicações cirúrgicas em 0,5% dos casos. Se a cirurgia é bem sucedida, o paciente entra no estado recuperação(estado R). Após 1 ano, noventa por cento dos pacientes que se submeteram a cirurgia ficam sãos(estado S), 9% permanecem no estado de recuperação e 1% morrem por causas naturais. Quando o paciente fica são ele permanece nesse estado em 99% dos casos.*

Observe que o estudo clínico pode ser modelado por uma cadeia de Markov absorvente, pois temos um estado absorvente (estado M) e este é acessível por todos os outros estados. Observe na figura 2.2 como fica o diagrama de estados desta cadeia e a partir dele sua correspondente matriz de transição.

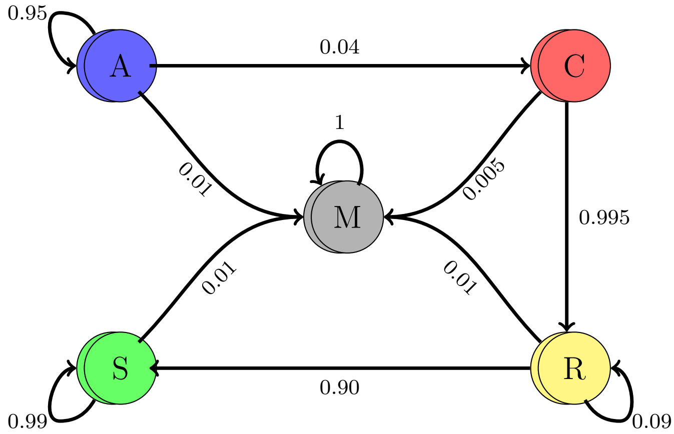


Figura 2.2: *Diagrama de estados - Pedra na Vesícula*

$$\mathbf{P} = \begin{array}{c} \\ \mathbf{A} \\ \mathbf{C} \\ \mathbf{R} \\ \mathbf{S} \\ \mathbf{M} \end{array} \begin{array}{c} \begin{array}{ccccc} \mathbf{A} & \mathbf{C} & \mathbf{R} & \mathbf{S} & \mathbf{M} \end{array} \\ \left[\begin{array}{cccc|c} 0,95 & 0,04 & 0 & 0 & 0,01 \\ 0 & 0 & 0,995 & 0 & 0,005 \\ 0 & 0 & 0,09 & 0,90 & 0,01 \\ 0 & 0 & 0 & 0,99 & 0,01 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

Observe que neste caso a matriz de transição P acima já está em sua forma canônica, portanto, podemos calcular a matriz fundamental $F = (I_4 - Q)^{-1}$, onde Q é a matriz quadrada de ordem 4 que contém as probabilidades de transição entre os estados transientes e I_4 é a matriz identidade de ordem 4. Calculando F chegamos a seguinte matriz:

$$\mathbf{F} = \begin{array}{c} \\ \mathbf{A} \\ \mathbf{C} \\ \mathbf{R} \\ \mathbf{S} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{A} & \mathbf{C} & \mathbf{R} & \mathbf{S} \end{array} \\ \begin{bmatrix} 20 & 0,8 & 0,87 & 78,73 \\ 0 & 1 & 1,09 & 98,41 \\ 0 & 0 & 1,10 & 98,90 \\ 0 & 0 & 0 & 100 \end{bmatrix} \end{array}$$

Como sabemos as entradas f_{ij} da matriz fundamental nos informam qual o número de vezes que o processo irá visitar o estado s_j visto que iniciou o processo no estado s_i. Assim, neste caso, a interpretação que podemos dar a matriz fundamental é a seguinte: cada uma das entradas nos indica quantas vezes o paciente visitará algum dos estados que o mantém vivo (assintomático, cirurgia, recuperação ou são) antes de vir a óbito. Por exemplo, se um paciente está inicialmente assintomático ele continuará neste estado, em média, por 20 períodos, antes de morrer. Como o período é de 4 meses, isso significa que um paciente assintomático pode continuar neste estado durante aproximadamente 6,7 anos.

Somando as linhas da matriz fundamental F chegamos aos seguintes valores para a matriz t, que nos diz quantos períodos o paciente permanecerá vivo, se iniciar o processo em qualquer um dos estados transientes.

$$\mathbf{t} = \begin{array}{c} \mathbf{A} \\ \mathbf{C} \\ \mathbf{R} \\ \mathbf{S} \end{array} \begin{bmatrix} 100,4 \\ 100,5 \\ 100 \\ 100 \end{bmatrix} \quad ou \quad \mathbf{t} = \begin{array}{c} \mathbf{A} \\ \mathbf{C} \\ \mathbf{R} \\ \mathbf{S} \end{array} \begin{bmatrix} 33,4 \\ 33,5 \\ 33,3 \\ 33,3 \end{bmatrix} em \quad anos.$$

Por fim podemos calcular qual a probabilidade de que um paciente venha à óbito dado que ele inicia o processo em algum estado transiente. Para isso calculamos o valor da matriz $B = Q \cdot R$, vejamos:

$$\mathbf{B} = \begin{bmatrix} 20 & 0,8 & 0,87 & 78,73 \\ 0 & 1 & 1,09 & 98,41 \\ 0 & 0 & 1,10 & 98,90 \\ 0 & 0 & 0 & 100 \end{bmatrix} \cdot \begin{bmatrix} 0,01 \\ 0.005 \\ 0.01 \\ 0.01 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Note que a probabilidade de que um paciente vá morrer é de 100% para qualquer um dos estados iniciais que ele possa ter. O que infelizmente, é lamentável, mas completamente compatível com a realidade de qualquer ser humano, tendo ele pedras na vesícula ou não.

Aqui encerramos nossas considerações teóricas sobre as cadeias de Markov absorventes. No próximo capítulo, iremos fazer algumas observações de como devem ser aplicadas as atividades propostas no apêndice e no site https://sites.google.com/view/cadeiasdemarkovabsorventes/.

Capítulo 3

Materiais Pedagógicos

Neste capítulo iremos comentar e esclarecer algumas questões propostas nas atividades escritas e virtuais. Como já dito na introdução, achamos que o ideal seja aplicar estas atividades no 2º ou 3º ano do ensino médio, visto que desta forma os estudantes já estudaram o conteúdo de matrizes e de probabilidade condicional.

O principal objetivo destas atividades é mostrar como funcionam as cadeias de Markov Absorventes e em quais situações do cotidiano elas podem ser usadas. É uma tentativa de responder aos alunos a pergunta que sempre fazem nas aulas de matemática: *Pra que serve isso?*

É um conteúdo que não faz parte da grade curricular do ensino médio. O professor que desejar aplicar o material, deve avaliar se o mesmo dispõe de tempo hábil em suas aulas para fazê-lo. Acreditamos que uma abordagem tradicional com aulas expositivas seja suficiente para dar conta de um nível de aprendizagem satisfatório.

Dessa forma, pode-se usar uma aula para as noções básicas sobre as cadeias de Markov e outra com os conceitos relativos as cadeias Absorventes. A partir daí pode-se pedir que os estudantes visitem o site e façam as listas de exercícios.

É importante destacar que este é um patamar mínimo de planejamento. Esta ideia inicial pode ser expandida, até mesmo para um projeto interdisciplinar. Seja com a área de ciências humanas, explorando exercícios como o do nível econômico, do nível de escolaridade ou da produção de uma safra de alimentos. Como podemos também utilizar os exercícios sobre genética, saúde e trânsito para trabalhar com os professores de ciências da natureza.

Enfim, achamos que a dosagem, à ênfase e a metodologia que será usada para trabalhar o assunto depende da intenção e do planejamento que o docente fará para tratar do tema. A seguir faremos alguns comentários que avaliamos pertinentes em relação aos materiais pedagógicos produzidos.

3.1 Listas de Exercícios

Produzimos três listas de exercícios que seguem a mesma estrutura desta dissertação. Na primeira exploramos questões relacionadas a conceitos básicos das cadeias de Markov. Nessa lista, as duas primeiras questões exploram a construção de diagrama de estados e a definição de matriz de transição de uma cadeia de Markov.

Os alunos devem perceber que nem toda matriz formada por decimais ou frações é uma matriz de transição. É necessário enfatizar que as linhas de uma matriz de transição devem ser iguais a 1. Ademais o aluno deverá ter a habilidade de intercambiar as representações pictórica(diagramas) e algébrica(matrizes) para descrever uma cadeia de Markov.

As questões seguintes até a sétima estão relacionadas ao reconhecimento do significado de cada componente de uma matriz de transição e ao que ocorre com o vetor de probabilidade de uma cadeia após cada etapa do processo. Nas questões 3 e 4, por exemplo, o aluno terá de perceber o que significa cada entrada da matriz de transição, em sua forma decimal ou de fração.

Além disso, deverá determinar a distribuição de probabilidades para uma etapa dado que o processo iniciou em determinado estado. Um fato a ser destacado, novamente, é que se o processo tem início em um determinado estado temos que o vetor de distribuição inicial será igual a 1 no estado em que o processo inicia e igual a zero nos demais.

Ademais o aluno deverá ser capaz de efetuar a multiplicação entre o vetor de probabilidade inicial e a matriz de transição para determinar a próxima distribuição de probabilidade do processo, ou ainda, que isso pode ser feito através do produto entre uma potência da matriz de transição e o vetor de probabilidade inicial utilizando o teorema 1.5.1.

Nas questões de 8 a 10 temos situações problemas que podem ser modeladas por uma cadeia de Markov. Os alunos deverão ser capazes de extrair do texto o diagrama de estados, a matriz de transição e o vetor de probabilidade inicial e depois fazer algum tipo de multiplicação para determinar o que se pede no problema. Apesar de serem basicamente as mesmas habilidades requeridas nas questões anteriores, a leitura e interpretação de texto serão exigidas nestes casos, o que as elevam a um nível maior de dificuldade.

Vale lembrar que esta é a única das três listas que recomendamos que seja feita sem nenhum auxílio computacional(Geogebra, calculadora...), pois de fato não é necessário o uso do mesmo. Gostaríamos de enfatizar que apesar da simplicidade destas questões é essencial que sejam feitas pelos alunos.

A segunda lista trata do comportamento a longo prazo do vetor de probabilidade em cadeias de Markov que sejam *regulares*. O uso de um recurso computacional, como o Geogebra, nos habilita a tratar as questões propostas sem maiores dificuldades.

A primeira questão apenas serve para que o aluno aplique a definição de matriz regular, ou seja, ele deverá ser capaz de avaliar se a matriz analisada ou alguma de suas potências tem todas as entradas positivas.

Na segunda questão o uso do Geogebra é indispensável visto que o aluno terá que elevar as matrizes dadas a potências cujo expoentes tem um valor muito grande para ser feito a mão.

Recomendamos o uso da janela planilhas no Geogebra, caso o computador seja a ferramenta disponível. Ou, o que nos parece ser mais viável, principalmente no caso das escolas públicas estaduais, o uso do Geogebra na versão Android, na medida em que uma grande quantidade dos alunos tem acesso ao dispositivo. Em ambos os casos temos vídeo-aulas de como utilizar o Geogebra para efetuar estes cálculos na página Cadeia de Markov disponível no site.

A terceira questão mais uma vez exige que o estudante seja capaz de escrever o diagrama de estados para daí determinar a matriz de transição e, nesta questão especificamente, terá apenas que elevar a matriz de transição encontrada a um n suficientemente grande para que a distribuição estacionária apareça.

As questões 4 e 5, apesar de matematicamente não acrescentarem muito pois continuam sendo resolvidas como as anteriores, mostra ao aluno como a ideia estacionária tem sua relevância no mundo físico e social. Na questão 4 temos uma situação onde a genética determina como uma população a longo prazo será distribuída se levarmos em consideração apenas esse aspecto.

Na questão 5 temos uma matriz de transição que nos mostra as tendências de qual o nível de instrução que a geração seguinte terá levando em consideração o nível de instrução da geração atual. Podemos propor uma investigação do que ocorre com o vetor de probabilidade estacionário quando existem mudanças na matriz de transição desta situação. Ainda é possível discutir com os alunos o impacto de políticas públicas educacionais para gerações futuras com base nas informações encontradas e isto pode ser feito de forma interdisciplinar com a ajuda de professores de história, sociologia e geografia.

A questão 6 é uma analogia da situação do labirinto do rato proposta na apresentação de slides. Pois nestas situações as probabilidades de transição entre os estados não são dadas no texto da questão, é necessário que o aluno seja capaz de descobri-las a partir da figura dada. Outro fato interessante nesta questão é que pela primeira vez o aluno terá que quantificar a distribuição de probabilidade encontrada para determinar quantos robôs ficarão em cada posição.

Para finalizar os comentários sobre a lista 2 é importante que seja destacado um fato essencial sobre as cadeias de Markov com matriz de transição regular. Este tipo de cadeia nos mostra que o estado em que algo se encontra inicialmente(vetor de probabilidade inicial) é irrelevante a longo prazo. O que verdadeiramente importa são as escolhas (matriz de transição) que iremos fazer ao longo do processo.

A terceira lista é dedicada a questões relacionadas à cadeias de Markov absorventes. Esta lista deve ser precedida da resolução das anteriores, pois é necessário que os alunos já tenham adquirido certas habilidades que são indispensáveis para a resolução dessa terceira lista.

Uma dessas habilidades é saber escrever o diagrama de estados a partir de um texto que descreva uma situação problema que possa ser modelada por uma cadeia de Markov. Outra habilidade necessária é a de determinar à matriz de transição associada a um diagrama de estados. Ainda é essencial que nesta fase os discentes já consigam usar sem maiores dificuldades o Geogebra para manipular matrizes.

Na primeira questão o aluno deverá ser capaz de identificar os estados absorventes de uma matriz de transição. É muito importante que ele compreenda que não basta que a probabilidade de transição seja de 100%, mas sim que a probabilidade de permanecer naquele estado seja de 100%. Além disso, na questão dois eles devem identificar quais das matrizes da questão 1 representam uma matriz de transição de uma cadeia de Markov absorvente, ou seja, devem verificar se além de possuir estados absorventes, se estes estados são acessíveis a todos os outros não absorventes.

A forma canônica e a matriz fundamental são exploradas na terceira questão. A manipulação das matrizes neste caso deve ser feita com muito cuidado pois ao deslocar linhas e colunas há o risco de que se altere a matriz de transição, o que altera completamente as informações. Na questão 4 a ideia é que os alunos consigam interpretar corretamente a matriz de transição F de uma cadeia de Markov absorvente, o problema da ruína do jogador é um bom exemplo de como isto acontece, por isso retornamos a ele nesta questão.

As questões 5 e 6 são situações problema na área de educação e saúde onde são dadas a matriz de transição em ambos os casos. Os itens propostos em cada questão são todos eles facilmente respondidos contanto que os estudantes saibam interpretar corretamente a matriz fundamental F em cada situação.

Nas questões 7 e 8 às matrizes de transição associadas a cada problema não são dadas e o aluno deverá construí-las a partir da leitura e interpretação do texto. Recomendamos que o diagrama de estados preceda a escrita da matriz de transição porque geralmente é mais dinâmico o processo de transformar a informação textual em diagrama para a partir daí construir a matriz de transição, como feito nas listas anteriores.

A questão 9 é um problema sobre genética e possui alguns aspectos inéditos no que diz respeito a sua resolução. O primeiro deles é que os estados do processo são formados a partir da combinação dos possíveis estados dos indivíduos estas combinações são indicadas pela tabela 5.2.

Situações semelhantes a essa, são muito bem exploradas no livro do Hillier e Lieberman (2013), em problemas onde máquinas são usadas para fazer algum produto ou serviço. Os estados de transição neste caso também são obtidos a partir da combinação entre estados individuais(funcionando, em manutenção) das máquinas que realizam o processo.

Outro aspecto inédito da solução da questão 9 é que as probabilidades de transição entre estados apesar de serem dadas no item a, é pedido que o aluno verifique se as mesmas estão corretas. Portanto ele deverá calcular cada uma das probabilidades de transição e para isso ele precisará obedecer as Leis de Mendell, que são:

1. Cada característica é determinada por um par de fatores herdados um de cada progenitor.

2. Os fatores para duas ou mais características distribuem-se de maneira independente para os descendentes e combinam-se ao acaso.

3. Cada fator puro para cada característica é transmitida à geração seguinte de forma independente uma da outra seguindo as duas leis anteriores.

Aplicando estas leis temos que a matriz do item (a) é de fato a matriz de transição para o problema. A solução dos outros itens seguem o mesmo padrão que os das questões anteriores.

Por fim a questão 10 também não nos indica diretamente quais as probabilidades de transição entre estados. Elas também precisam ser calculadas. Neste caso existem duas condições que precisam ser destacadas. Primeiro os tiros acontecem de forma simultânea a cada etapa e segundo é que cada um sempre atira no seu mais forte oponente vivo. Observe ainda que o conjunto dos estados possíveis é composto por $E = \{ABC,\ AC,\ BC,\ A,\ B,\ C,\ Nenhum\}$, o estado AB nunca pode ocorrer por causa da segunda condição. Sendo assim a matriz de transição para os carros de combate que sobrevivem a cada etapa é dada por:

$$
\mathbf{P} = \begin{array}{c} \\ \mathbf{ABC} \\ \mathbf{AC} \\ \mathbf{BC} \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{N} \end{array}
\begin{array}{c}
\begin{array}{ccccccc} \mathbf{ABC} & \mathbf{AC} & \mathbf{BC} & \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{N} \end{array} \\
\begin{bmatrix}
\frac{5}{18} & \frac{5}{18} & \frac{4}{18} & 0 & 0 & \frac{4}{18} & 0 \\
0 & \frac{5}{12} & 0 & \frac{5}{12} & 0 & \frac{1}{12} & \frac{1}{12} \\
0 & 0 & \frac{10}{18} & 0 & \frac{5}{18} & \frac{2}{18} & \frac{1}{18} \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\end{array}
$$

É muito provável que os alunos tenham dificuldade em construir esta matriz. Por isso recomendamos que a mesma seja feita sob orientação do professor seguindo os passos que constam no apêndice 3. Os demais itens da questão podem ser facilmente respondidos através do cálculo da matriz fundamental F e da matriz B.

Finalizamos destacando a tentativa de colocar em todas as listas de exercícios um grau de dificuldade crescente nas questões. Existem sempre aquelas que avaliam apenas entendimento de teoremas e definições e outras que buscam aplicar em alguma situação problema estes conhecimentos. Ademais as listas ora propostas, podem ser complementadas ou ajustadas conforme o nível da turma ou o desejo do professor, visto que podemos disponibilizar os arquivos .tex das listas pelo site.

3.2 O Beamer

O Beamer é uma classe de documentos LaTeX para criar slides para apresentações. É uma ferramenta muito útil principalmente quando as notas de aula que fazemos forem criados no LaTeX, a produção dos slides se torna basicamente um copiar e colar com muito estilo.

Nas apresentações que fizemos, disponíveis no site, usamos como base a fundamentação teórica feita nesta dissertação. São duas apresentações, a primeira com as seções: Cadeias de Markov; Vetor de probabilidade e a segunda com Cadeias de Markov Absorventes. Sugerimos que estes slides sejam usados pelo professor para auxiliá-lo em suas explicações acerca das cadeias de Markov.

Na primeira apresentação iniciamos com uma foto de Andrei Markov e um parágrafo introdutório sobre as suas cadeias. A seguir temos um slide onde é mostrado um diagrama de estados para que os conceitos de estados e probabilidades de transição sejam introduzidos.

Como primeiro exemplo para a apresentação, escolhemos à situação problema que envolve o labirinto do rato, pois nos parece muito adequada para um primeiro contato dos estudantes com as cadeias de Markov.

Podemos com este problema mostrar aos estudantes como se constrói um diagrama de estados avançar para o slide sobre a matriz de transição e daí transformar o diagrama feito em uma matriz de transição. O professor pode usar o quadro branco e piloto para fazer isto juntamente com os estudantes.

A seguir temos alguns slides dedicados as equações de Chapman-Kolmogorov que devem ser mais ou menos aprofundado conforme a avaliação do docente. A segunda parte da apresentação introduz o conceito do vetor de probabilidade. O primeiro exemplo que usa este conceito é o da previsão do tempo na cidade de Salvador, no qual está ilustrado o que é um vetor de probabilidade inicial e como podemos saber qual a previsão do tempo no mês seguinte a partir do produto entre o vetor inicial e sua matriz de transição.

Com o exemplo da lealdade do consumidor é importante que o professor peça aos alunos para construírem o diagrama de estados e a matriz de transição antes de mostrá-los no projetor. De modo que consiga avaliar como está a compreensão dos conceitos vistos anteriormente. Com este exemplo introduzimos o conceito de vetor de probabilidade estacionário e de matriz de transição regular.

Na segunda apresentação utilizamos o exemplo do nível econômico para introduzir o conceito de estado absorvente. Em seguida temos vários slides definindo todos os elementos necessários para plena caracterização das cadeias de Markov absorventes.

Esta é uma parte crítica da apresentação, porque são feitas muitas definições e os estudantes podem se dispersar comprometendo o entendimento necessário para resolução de problemas futuros. Exemplificar estas definições e tentar ser o mais breve possível pode ajudar neste processo. Finalizando esta parte, os resultados obtidos são aplicados no exemplo do labirinto do rato.

Concluímos a apresentação com o problema da ruína do jogador, que apesar de ser uma situação simples e de fácil interpretação nos permite ilustrar bem todos os conceitos necessários para resolução dos problemas que envolvem cadeias de Markov absorventes.

3.3 O site

Construímos o site (`https://sites.google.com/view/cadeiasdemarkovabsorventes/`) que é basicamente um local público para que os professores e alunos possam ter contato com o material elaborado nesta dissertação. Usamos o Google sites como plataforma de criação, optamos por ela pois já tínhamos algum know-how de como usá-la.

O site contém quatro páginas: *Biografia, Cadeias de Markov, Cadeias de Markov Absorventes e Curiosidades.* A página que contém a biografia é uma cópia fiel da seção 1.1 desta dissertação, por que além de não ser um texto muito longo, não existe, pelo menos não encontramos nas pesquisas realizadas, um artigo ou mesmo referências, em português, sobre o assunto que possuam mais do que dois parágrafos.

Em seguida temos a página sobre as Cadeias de Markov onde temos nove link's que nos levam a alguns dos materiais que produzimos. No link Noções Básicas temos uma pequena apostila onde introduzimos o conceito das cadeias de Markov da mesma forma que fizemos aqui.

Contudo omitimos as demonstrações dos teoremas pois avaliamos que o texto ficaria um pouco rebuscado para um primeiro contato com o assunto numa apostila. No entanto para o professor ou aluno mais curioso as demonstrações estarão disponíveis no link dissertação nesta mesma página.

No link Exemplos temos a discussão dos problemas sobre as cadeias de Markov que aqui foram feitas onde também omitimos alguns trechos com comentários que são pertinentes apenas para o contexto desta dissertação.

Logo depois estão os links para as duas listas que possuem exercícios referentes as cadeias de Markov e ainda um link para uma Apresentação em Beamer que pode ser usada por professores para guiar sua aula sobre o assunto seguindo o mesmo caminho que fizemos aqui.

Na segunda coluna de links, iniciamos com aquele que nos direciona a um vídeo com um tutorial passo a passo de como podemos utilizar a versão do Geogebra para PC para efetuar os cálculos matriciais necessários para resolver problemas que envolvem cadeias de Markov.

O link seguinte nos leva a um pequeno vídeo mostrando como podemos efetuar esses mesmos cálculos na versão para celular do Geogebra. Isto se faz necessário, porque nestes dispositivos o programa Geogebra não disponibiliza o modo planilha utilizado no vídeo anterior para a construção das matrizes. Também é bastante provável que muitos alunos não irão dispor de um PC para esta tarefa, mas com certeza terão acesso a algum celular que os permitam efetuar estes cálculos.

O penúltimo link utiliza uma ferramenta muito interessante do Google que são os Formulários em seu modo teste. Criamos um pequeno quiz com alguns conceitos ligados as Cadeias de Markov, onde o aluno após responder as questões pode visualizar seu despenho automaticamente.

Além disso, se o professor desejar poderá adaptar este quiz de maneira que ele possa ser armazenado em uma planilha e assim sirva como instrumento de avaliação no processo de aprendizagem dos estudantes.

O processo de criação dos formulários requer um certo esforço mas as possibilidades que são geradas com os mesmos compensam esse gasto inicial de energia. Ao atribuir valores as questões é possível quantificar e armazenar uma nota para cada estudante, além do que são geradas estatísticas de erros e acertos de cada questão. Tudo isso é feito de forma automática, sem que seja necessária nenhuma intervenção, após a criação do formulário e preenchimento das respostas.

Chegando a página cujo título é Cadeias de Markov Absorventes temos os mesmos links gerados na página anterior. Uma parte teórica com Definições e Teoremas pertinentes ao tema, onde desta vez mantivemos todos devidamente demonstrados, no primeiro link. Em seguida temos os Exemplos, Lista de Exercícios, Apresentação, Quiz e Tutorial Geogebra todos com a mesma funcionalidade da página anterior.

Na página Curiosidades temos alguns links interessantes que encontramos na world wide web na fase de pesquisa desta dissertação que achamos importante de serem compartilhados. O primeiro link está direcionado para a página da wikipédia sobre o assunto, nessa página podemos encontrar diversas aplicações das cadeias de Markov em áreas distintas do conhecimento.

O segundo link é de um vídeo da Khan Academy, em inglês, sobre as cadeias de Markov onde podemos ver que elas foram o resultado da busca do seu criador para garantir que a lei dos grandes números também era válida para eventos dependentes.

O link Processos Estocástico nos direciona para o canal do You Tube do professor

Luís Rincon do departamento de Matemática da Faculdade de Ciências da Universidade Nacional Autônoma do México. Um curso que apesar de estar em espanhol é completamente compreendido e de uma riqueza de detalhes que vale a pena conferir.

No link Simulação temos uma página da internet que nos mostra diagramas de estado de uma cadeia de Markov com dois ou três estados. É possível manipular as probabilidades de transição e perceber como isso modifica o processo através de uma animação.

Seguindo temos um vídeo, em inglês, onde as noções básicas das cadeias de Markov são apresentadas de uma forma simples e rápida que com certeza os estudantes irão gostar. E pode até servir de ponte para um trabalho interdisciplinar com a disciplina de língua estrangeira.

Por fim temos os dois links finais, um para o excelente livro de probabilidade de domínio público, em inglês, de Grimstead e Snell (1997), que inclusive faz parte da referência bibliográfica dessa dissertação. E o último é o artigo de história da matemática, aqui também usado, de Brian Hayes que fala sobre a vida e obra de Andrei Andreyevich Markov.

Capítulo 4

Considerações finais

Ao final desta dissertação ainda estamos fascinados pelas cadeias de Markov. Apesar das cadeias absorventes terem grande relevância teórica, com a pesquisa desenvolvida, vimos que são apenas a ponta do iceberg. Por isso desejamos continuar produzindo material com este tema, para atualizar o site com aquilo que encontrarmos e puder ser adaptado para alunos do ensino médio.

Além disso, constatamos que há uma escassez de material sobre as cadeias de Markov em língua portuguesa, o que nos impulsiona a continuar pesquisando sobre as mesmas. Até existem apostilas e monografias mas sempre com enfoque no ensino superior. As dissertações do PROFMAT sobre o tema foi o único lugar que encontramos materiais em português com o foco no ensino médio.

Outro desejo que possuímos é dar acesso aos estudantes do ensino médio a temas e problemas que envolvam matemática aplicada com os assuntos que eles vêm nesta fase de ensino. Queremos criar oficinas itinerantes que possam ser realizadas em escolas públicas de ensino médio no nosso estado, para que os estudantes tenham aí um outro estímulo para fomentar o desejo de aprender ou até mesmo seguir carreira na área de exatas.

É grande a satisfação de chegar ao final desse texto. Mais uma vez gostaria de agradecer a existência do Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, que possibilita ao professor do ensino básico trabalhar com pesquisa e perceber que é capaz de fazê-lo. O professor-pesquisador, no nível básico, apesar de teoricamente ser um objetivo do nosso sistema educacional, tem poucas iniciativas que verdadeiramente fomentam esta prática.

A criação do site e o contato com novas tecnologias como o Geogebra, o LaTeX e os aplicativos Google, fez com que o interesse por metodologias de ensino mais ativas crescesse e que buscássemos pelo menos conhecer os meios necessários para implementação das mesmas em nossas aulas, isso também se deu graças a escrita desse texto e a este programa.

O LaTeX, em particular, foi uma ferramenta de grande ajuda para vislumbrar novos horizontes para planejamento e execução de aulas. Muitas horas foram necessárias para conseguir escrever o texto que agora se lê. Mas a qualidade do texto e mudança na lógica do digitar para programar aquilo que se escreve é indescritível. Ademais o Beamer e o Tik'Z foram ambientes que nos ajudaram bastante e nos garantem infinitas possibilidades na criação de apresentações. Ainda em relação ao $\LaTeX$, gostaríamos de destacar o canal do You Tube do português Antero Neves que possui várias vídeo-aulas que muito nos ajudaram.

Concluindo, acreditamos que o caráter interdisciplinar do assunto aqui tratado possibilita ao discente uma relação de aproximação entre o vivido e o estudado que tanto buscamos atingir no dia a dia da sala de aula. É importante observar que trabalhos com matemática aplicada dão ao estudante uma visão não separatista e pragmática da matemática, melhorando sua relação com a mesma, que como sabemos nem sempre é harmoniosa.

Por fim deixamos uma citação do livro *A fórmula preferida do professor* (OGAWA, 2017), que tem como personagens uma cuidadora, seu filho e um matemático, que devido a um acidente, tem um problema de memória recente e só se lembra dos últimos 80 minutos de sua vida.

O trecho destacado nos traz uma visão platônica muito bela do que é a matemática e como ela contribui na jornada humana sobre a Terra. A parte citada abaixo acontece logo após o professor pedir para a sua cuidadora desenhar numa folha de papel uma reta, o que ela o faz com o auxílio de um hashi e em seguida é indagada por ele da seguinte maneira:

> - Muito bem, isso é uma reta. Você compreendeu bem a definição de uma reta. Mas pense um pouco. A reta que você traçou tem começo e fim, certo? Então se trata de um segmento de reta ... Ou seja, é impossível desenhar uma reta verdadeira em uma folha de papel.
> Observei atentamente a ponta do lápis.
> - Então, onde é possível encontrar uma reta verdadeira. A resposta é: só aqui!
> O professor colocou a mão sobre o peito. O mesmo gesto que ele fizera ao me ensinar sobre números imaginários.
> - A verdade eterna, que não se deixa influenciar pela matéria ou pelos fenômenos naturais, é invisível aos nossos olhos. A matemática desvenda e revela a aparência dessa verdade. Nada pode detê-la.
> A existência dessa verdade eterna, à qual o Professor se referia, era necessária para mim, que esfregava o assoalho do escritório com o estômago vazio e sempre preocupada com o meu filho. Precisava sentir que, de fato, esse mundo invisível sustentava o mundo visível. A reta verdadeira sem espessura ou superfície, aquela que atravessava majestosamente as trevas estendendo-se além de qualquer limite - só ela poderia me trazer um pouco de tranquilidade.

Capítulo 5

Apêndices

5.1 Apêndice 1 - Cálculos

5.1.1 Lealdade do Consumidor

Abaixo temos algumas potências da matriz P do exemplo da Lealdade do Consumidor:

$$\mathbf{P} = \begin{bmatrix} 0,7500 & 0,1000 & 0,0500 & 0,1000 \\ 0,0800 & 0,6000 & 0,1500 & 0,1700 \\ 0,1000 & 0,2000 & 0,5000 & 0,2000 \\ 0 & 0,1000 & 0,1000 & 0,8000 \end{bmatrix} \quad \mathbf{P^5} = \begin{bmatrix} 0,3040 & 0,2155 & 0,1448 & 0,3357 \\ 0,1558 & 0,2594 & 0,1803 & 0,4044 \\ 0,1599 & 0,2448 & 0,1829 & 0,4124 \\ 0.0815 & 0,2243 & 0,1770 & 0,5172 \end{bmatrix}$$

$$\mathbf{P^{10}} = \begin{bmatrix} 0,1765 & 0,2322 & 0,1688 & 0,4225 \\ 0,1496 & 0,2357 & 0,1739 & 0,4408 \\ 0,1496 & 0,2352 & 0,1737 & 0,4414 \\ 0,1302 & 0,2351 & 0,1762 & 0,4586 \end{bmatrix} \quad \mathbf{P^{15}} = \begin{bmatrix} 0,1512 & 0,2344 & 0,1731 & 0,4413 \\ 0,1459 & 0,2348 & 0,1740 & 0,4452 \\ 0,1459 & 0,2348 & 0,1740 & 0,4453 \\ 0.1417 & 0,2350 & 0,1747 & 0,4486 \end{bmatrix}$$

$$\mathbf{P^{20}} = \begin{bmatrix} 0,1461 & 0,2347 & 0,1740 & 0,4452 \\ 0,1450 & 0,2348 & 0,1741 & 0,4460 \\ 0,1450 & 0,2348 & 0,1741 & 0,4460 \\ 0,1442 & 0,2349 & 0,1743 & 0,4467 \end{bmatrix} \quad \mathbf{P^{25}} = \begin{bmatrix} 0,1451 & 0,2348 & 0,1741 & 0,4460 \\ 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0.1447 & 0,2348 & 0,1742 & 0,4463 \end{bmatrix}$$

$$\mathbf{P^{30}} = \begin{bmatrix} 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \end{bmatrix} \quad \mathbf{P^{35}} = \begin{bmatrix} \boxed{\begin{matrix} 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0.1448 & 0,2348 & 0,1742 & 0,4462 \end{matrix}} \end{bmatrix}$$

$$\mathbf{P^{40}} = \begin{bmatrix} 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0.1448 & 0,2348 & 0,1742 & 0,4462 \end{bmatrix} \qquad \mathbf{P^{45}} = \begin{bmatrix} 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0.1448 & 0,2348 & 0,1742 & 0,4462 \end{bmatrix}$$

Então pelo teorema 2.5.1 temos que:

$$\pi^{(5)} = \pi^{(0)} \cdot P^5 = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,3040 & 0,2155 & 0,1448 & 0,3357 \\ 0,1558 & 0,2594 & 0,1803 & 0,4044 \\ 0,1599 & 0,2448 & 0,1829 & 0,4124 \\ 0.0815 & 0,2243 & 0,1770 & 0,5172 \end{bmatrix}$$

$$\pi^{(5)} = \begin{bmatrix} 0,1789 & 0,2381 & 0,1714 & 0,4116 \end{bmatrix}$$

$$\pi^{(10)} = \pi^{(0)} \cdot P^{10} = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,1765 & 0,2322 & 0,1688 & 0,4225 \\ 0,1496 & 0,2357 & 0,1739 & 0,4408 \\ 0,1496 & 0,2352 & 0,1737 & 0,4414 \\ 0,1302 & 0,2351 & 0,1762 & 0,4586 \end{bmatrix}$$

$$\pi^{(10)} = \begin{bmatrix} 0,1524 & 0,2346 & 0,1731 & 0,4399 \end{bmatrix}$$

$$\pi^{(15)} = \pi^{(0)} \cdot P^{15} = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,1512 & 0,2344 & 0,1731 & 0,4413 \\ 0,1459 & 0,2348 & 0,1740 & 0,4452 \\ 0,1459 & 0,2348 & 0,1740 & 0,4453 \\ 0.1417 & 0,2350 & 0,1747 & 0,4486 \end{bmatrix}$$

$$\pi^{(15)} = \begin{bmatrix} 0,1464 & 0,2347 & 0,1739 & 0,4449 \end{bmatrix}$$

$$\pi^{(20)} = \pi^{(0)} \cdot P^{20} = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,1461 & 0,2347 & 0,1740 & 0,4452 \\ 0,1450 & 0,2348 & 0,1741 & 0,4460 \\ 0,1450 & 0,2348 & 0,1741 & 0,4460 \\ 0,1442 & 0,2349 & 0,1743 & 0,4467 \end{bmatrix}$$

$$\pi^{(20)} = \begin{bmatrix} 0,1451 & 0,2348 & 0,1741 & 0,4459 \end{bmatrix}$$

$$\pi^{(25)} = \pi^{(0)} \cdot P^{25} = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,1451 & 0,2348 & 0,1741 & 0,4460 \\ 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0.1447 & 0,2348 & 0,1742 & 0,4463 \end{bmatrix}$$

$$\pi^{(25)} = \begin{bmatrix} 0,1449 & 0,2348 & 0,1742 & 0,4461 \end{bmatrix}$$

$$\pi^{(30)} = \pi^{(0)} \cdot P^{30} = \begin{bmatrix} 0,2500 & 0,3200 & 0,2300 & 0,2000 \end{bmatrix} \cdot \begin{bmatrix} 0,1449 & 0,2348 & 0,1742 & 0,4461 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \\ 0,1448 & 0,2348 & 0,1742 & 0,4462 \end{bmatrix}$$

$$\pi^{(30)} = \begin{bmatrix} 0,1448 & 0,2348 & 0,1742 & 0,4462 \end{bmatrix}$$

Observe que os valores encontrados correspondem aos achados na tabela 1.3.

5.1.2 Forma Canônica

Segue abaixo as potências de P_1 do exemplo da seção 2.2.

$$\mathbf{P_1^2} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0,11 & 0,78 & 0,07 & 0,04 \\ 0 & 1 & 0 & 0 \\ 0,26 & 0,12 & 0,3 & 0,32 \\ 0,27 & 0,18 & 0,27 & 0,28 \end{array}\right] \end{array} \quad \mathbf{P_1^5} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0,03 & 0,91 & 0,03 & 0,03 \\ 0 & 1 & 0 & 0 \\ 0,16 & 0,53 & 0,16 & 0,16 \\ 0,14 & 0,58 & 0,14 & 0,14 \end{array}\right] \end{array}$$

$$\mathbf{P_1^{10}} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0,01 & 0,97 & 0,01 & 0,01 \\ 0 & 1 & 0 & 0 \\ 0,05 & 0,85 & 0,05 & 0,05 \\ 0,05 & 0,86 & 0,05 & 0,05 \end{array}\right] \end{array} \quad \mathbf{P_1^{15}} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,02 & 0,94 & 0,02 & 0,02 \\ 0,02 & 0,94 & 0,02 & 0,02 \end{array}\right] \end{array}$$

$$\mathbf{P_1^{20}} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0 & 1 & 0 & 10 \\ 0 & 1 & 0 & 0 \\ 0,01 & 0,97 & 0,01 & 0,01 \\ 0 & 1 & 0 & 0 \end{array}\right] \end{array} \quad \mathbf{P_1^{25}} = \begin{array}{c} \\ E_1 \\ E_2 \\ E_3 \\ E_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \left[\begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \end{array}$$

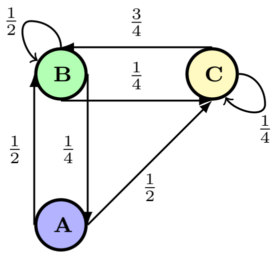
5.2 Apêndice 2 - Listas de Exercícios

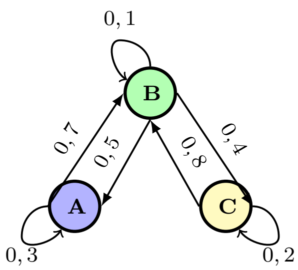
5.2.1 1º Bloco - Noções Básicas

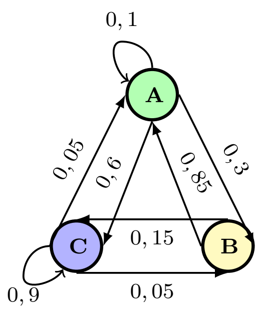
1. Decida, utilizando a definição, quais das matrizes abaixo poderia ser uma matriz de transição para uma cadeia de Markov. Esboce um diagrama para estas matrizes.

$$\text{a)} \begin{bmatrix} 0,5 & 0 \\ 0 & 0,5 \end{bmatrix} \qquad \text{b)} \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ 1 & 0 \end{bmatrix} \qquad \text{c)} \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \text{d)} \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ 2 & 0 & 1 \\ 1 & \frac{2}{3} & 0 \end{bmatrix}$$

2. Escreva uma matriz de transição associada a cada um dos diagramas abaixo.

a) 

b) 

c) 

3. Considere uma cadeia de Markov com matriz de transição:

$$\begin{array}{c c} & \begin{array}{cc} \text{Estado 1} & \text{Estado 2} \end{array} \\ \begin{array}{c} \text{Estado 1} \\ \text{Estado 2} \end{array} & \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \end{array}$$

a) O que representa a entrada $\frac{1}{4}$ nesta matriz?

b) Supondo que o sistema esteja inicialmente no estado 1, determine a distribuição de probabilidade após uma observação. Qual será após duas observações?

c) Supondo que o sistema esteja inicialmente no estado 2, determine a distribuição de probabilidade após uma observação. Qual será após duas observações?

4. Considere uma cadeia de Markov com matriz de transição:

$$\begin{array}{c c} & \begin{array}{cc} \text{Estado 1} & \text{Estado 2} \end{array} \\ \begin{array}{c} \text{Estado 1} \\ \text{Estado 2} \end{array} & \begin{bmatrix} 0,1 & 0,9 \\ 0,5 & 0,5 \end{bmatrix} \end{array}$$

a) O que representa a entrada $0,7$ nesta matriz?

b) Supondo que o sistema esteja inicialmente no estado 1, determine a distribuição de probabilidade após uma observação. Qual será após duas observações?

c) Se a distribuição de probabilidade inicial for $[\frac{1}{4}, \frac{3}{4}]$, qual será a probabilidade após duas observações?

5. Considere uma cadeia de Markov com matriz de transição:

$$\mathbf{P} = \begin{bmatrix} 0,2 & 0,7 & 0,1 \\ 0,6 & 0,2 & 0,2 \\ 0,4 & 0,1 & 0,5 \end{bmatrix}$$

   Se a distribuição inicial for $[0,25; 0,25; 0,50]$, qual será a distribuição de probabilidade na próxima observação?

6. Determine os valores de a, b e c que farão da matriz a seguir uma matriz de transição de uma cadeia de Markov.

$$\mathbf{P} = \begin{bmatrix} 0,2 & a & 0,4 \\ b & 0,6 & 0,3 \\ 0 & c & 0 \end{bmatrix}$$

7. Ache as três primeiras potências de cada uma das matrizes abaixo. Para cada uma das matrizes de transição, determine a probabilidade de que iniciando o processo no estado 1 ele esteja no estado 2 após três etapas.

   a) $\begin{bmatrix} 1 & 0 \\ 0,7 & 0,3 \end{bmatrix}$ b) $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ 1 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 0,5 & 0,5 \\ 0,64 & 0,36 \end{bmatrix}$

   d) $\begin{bmatrix} 0,3 & 0,5 & 0,2 \\ 0 & 0 & 1 \\ 0,6 & 0,3 & 0,1 \end{bmatrix}$ e) $\begin{bmatrix} 0,1 & 0,1 & 0,8 \\ 0,2 & 0,3 & 0,5 \\ 0,6 & 0,3 & 0,1 \end{bmatrix}$

8. Uma pequena e remota vila recebe transmissões de duas estações de rádio, uma de notícias e outra de músicas. Dos ouvintes que estão sintonizados na estação de notícias, 70% permanecerão ouvindo noticiários após o intervalo que ocorre a cada meia hora, enquanto 30% trocarão de estação durante o intervalo. Dos ouvintes que estão sintonizados na estação de música, 60% trocarão de estação durante o intervalo, enquanto 40% permanecerão ouvindo música. Suponha que as 08:15 todos estão ouvindo o noticiário.

   a) Desenhe um diagrama que represente o comportamento dos moradores dessa cidade em relação as estações de rádio.
   b) Escreva a matriz de transição associada ao diagrama do item anterior.
   c) Determine o vetor de probabilidade inicial.
   d) Que percentual dos ouvintes estarão sintonizados na estação de música às 09:25 (após os intervalos de programação das 08:30 e das 09:00).

9. Um estudo das safras de nozes-de-pinha no sudoeste americano, de 1940 a 1947, levantou à hipótese de a produção de nozes seguir uma cadeia de Markov. Os dados sugeriram que, se a safra de um ano fosse boa, as probabilidades de a safra do ano seguinte ser boa, regular ou péssima seria de. respectivamente, 0,08; 0,07 e 0,85; se a safra do ano fosse regular as probabilidades de a safra do ano seguinte ser boa, regular ou péssima seria de respectivamente, 0,09; 0,11 e 0,80; se a safra do ano fosse péssima as probabilidades de a safra do ano seguinte ser boa, regular ou péssima seria de. respectivamente, 0,11; 0,05 e 0,84.

   a) Desenhe um diagrama que represente esta cadeia.

   b) Escreva a matriz de transição associada ao diagrama do item anterior.

   c) Se a safra de 1940 foi boa, ache as probabilidades de uma boa safra ocorrer em 1943.

10. Um estudo sobre a resposta do sistema imunológico de coelhos os classificou em 4 grupos, de acordo com a força da resposta. De uma semana para outra, os coelhos trocam de classificação de um grupo para outro de acordo com a seguinte matriz de transição:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} \frac{5}{7} & \frac{2}{7} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \end{bmatrix} \end{array}$$

   a) Qual a proporção de coelhos do grupo 1 que permanecem no grupo 1 após 4 semanas?

   b) Na primeira semana haviam 9 coelhos no grupo 1, 4 no grupo 2 e nenhum coelho nos demais. Quantos coelhos são esperados em cada grupo após 4 semanas?

Bons Estudos!!!

5.2.2 2º Bloco - Vetor de probabilidade

1. Quais das matrizes de transição abaixo são regulares?

a) $\begin{bmatrix} 0,3 & 0,7 \\ 0,1 & 0,9 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 0 \\ 0,2 & 0,8 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

d) $\begin{bmatrix} 0,3 & 0,2 & 0,5 \\ 1 & 0 & 0 \\ 0,5 & 0,1 & 0,4 \end{bmatrix}$

2. Ache o vetor de equilíbrio para cada uma das matrizes de transição abaixo.

a) $\begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

b) $\begin{bmatrix} 0,3 & 0,7 \\ 0,6 & 0,4 \end{bmatrix}$

c) $\begin{bmatrix} 0,1 & 0,1 & 0,8 \\ 0,4 & 0,4 & 0,2 \\ 0,1 & 0,2 & 0,7 \end{bmatrix}$

d) $\begin{bmatrix} 0,5 & 0,2 & 0,3 \\ 0,1 & 0,3 & 0,6 \\ 0,2 & 0,2 & 0,6 \end{bmatrix}$

3. Uma comerciante abastece a sua loja com três tipos de detergente das marcas A, B e C. Quando a marca A é vendida, a probabilidade de que ela reabasteça com a marca A é de 0,7 e de 0,15 para cada uma das marcas B e C. Quando ela vende a marca B, a probabilidade de reabastecimento com a marca B é de 0,8 e de 0,1 para cada uma das marcas A e C. Por fim, quando ela vende a marca C, a probabilidade de que ela reabasteça com a marca C é de 0,6 e de 0,2 para cada uma das marcas A e B. Determine a matriz de transição. A longo prazo, como ficará distribuído o estoque de detergentes.

4. As probabilidades de que uma mãe loira venha a ter uma filha loira, ou uma morena, ou uma ruiva são de 0,6, 0,2 e 0,2, respectivamente. As probabilidades de que uma mãe morena venha a ter uma filha loira, ou uma morena, ou uma ruiva são de 0,1, 0,7 e 0,2, respectivamente. E as probabilidades de que uma mãe ruiva venha a ter uma filha loira, ou uma morena, ou uma ruiva são de 0,4, 0,2 e 0,4, respectivamente.

   a) Qual é a probabilidade de que uma mulher loira seja avó de uma morena?

   b) Se a população atual de mulheres é de 50% de morenas, 30% de loiras e o restante de ruivas, qual será a distribuição:

      i. Após duas gerações?

      ii. A longo prazo?

5. Use os dados da tabela abaixo e suponha que as tendências indicadas se mantenham, de modo a responder as questões a seguir:

Tabela 5.1: Tendências Educacionais

	Faculdade	Médio	Fundamental
Faculdade	80%	18%	2%
Médio	40%	50%	10%
Fundamental	20%	60%	20%

a) Qual é a matriz de transição?

b) Qual a probabilidade de que um neto de um concluinte do ensino fundamental complete a faculdade?

c) Qual a probabilidade de que um neto de um concluinte do ensino médio complete o ensino médio?

d) O nível mais elevado atingido por uma população em um determinado ano foi de 13% para a faculdade, 62% para o ensino médio e 25% para o ensino fundamental. Qual será a distribuição do nível de educação para os seus bisnetos?

e) Qual será a distribuição de longo prazo?

6. Robôs foram programados para atravessar o labirinto mostrado na figura abaixo e em cada junção, escolher aleatoriamente um caminho para seguir.

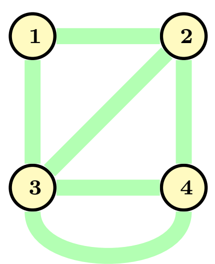


Figura 5.1: *Labirinto*

a) Construa a matriz de transição para a Cadeia de Markov que modela a situação.

b) Suponha que comecemos com 15 robôs em cada junção. Ache a distribuição estacionária dos robôs.(Considere que os robôs levam sempre o mesmo tempo para ir de uma junção a outra.)

Bons Estudos!!!

5.2.3 3º Bloco - Cadeias de Markov Absorventes

1. Identifique todos os estados absorventes das matrizes de transição abaixo.(Considere que o número de cada linha ou de cada coluna corresponde ao índice de um dos estados possíveis do processo)

a) $\begin{bmatrix} 0,15 & 0,05 & 0,8 \\ 0 & 1 & 0 \\ 0,4 & 0,6 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 0,4 & 0 & 0,6 \\ 0 & 1 & 0 \\ 0,9 & 0 & 0,1 \end{bmatrix}$

c) $\begin{bmatrix} 0,32 & 0,41 & 0,16 & 0,11 \\ 0,42 & 0,30 & 0 & 0,28 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 0,2 & 0,5 & 0,1 & 0,2 \\ 0 & 1 & 0 & 0 \\ 0,9 & 0,02 & 0,04 & 0,04 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

2. Quais das matrizes de transição do item anterior representam uma cadeia de Markov absorvente?

3. Escreva as matrizes na forma canônica e determine a matriz fundamental F das cadeias de Markov absorventes abaixo.

a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0,2 & 0,2 & 0,6 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$ d) $\begin{bmatrix} \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$

e) $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0,1 & 0,2 & 0,3 & 0,2 & 0,2 \\ 0,3 & 0,5 & 0,1 & 0 & 0,1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ f) $\begin{bmatrix} 0,4 & 0,2 & 0,3 & 0 & 0,1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0,1 & 0,5 & 0,1 & 0,1 & 0,2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

4. Suponha que para uma certa cadeia de Markov absorvente, a matriz fundamental F seja dada por:

$$\begin{array}{c} \\ R\$\ 1 \\ R\$\ 2 \\ R\$\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} R\$\ 1 & R\$\ 2 & R\$\ 3 \end{array} \\ \begin{bmatrix} 2,5 & 1 & 0,5 \\ 2 & 3 & 1 \\ 1,5 & 2 & 1,5 \end{bmatrix} \end{array}$$

(a) Qual é o número esperado de vezes que uma pessoa terá R\$ 3, dado que ela iniciou com R\$ 1? E com R\$ 2?

(b) Se uma jogadora iniciar com R\$ 3, de quantas jogadas ela pode esperar participar antes da absorção?E se iniciar com R\$ 1?

5. Dados foram coletados sobre a probabilidade de que um professor ou um estudante com interesse declarado em ensino, continue nessa carreira no ano seguinte. Para simplificar a análise, os dados originais foram classificados em quatro categorias: estudantes de ensino médio ou faculdade (E), professores recém formados(RF), professores veteranos (P) e aqueles que deixam a profissão (S). As probabilidades de transição estão na matriz abaixo.

$$\begin{array}{c c} & \begin{array}{cccc} E & RF & P & S \end{array} \\ \begin{array}{c} E \\ RF \\ P \\ S \end{array} & \left[\begin{array}{cccc} 0,70 & 0,11 & 0 & 0,19 \\ 0 & 0 & 0,86 & 0,14 \\ 0 & 0 & 0,88 & 0,12 \\ 0 & 0 & 0 & 1 \end{array}\right] \end{array}$$

   (a) Encontre o número de anos esperados para que um estudante com interesse em ensinar levará sendo professor veterano;

   (b) Determine o número de anos esperados para que um professor recém-formado ficará como professor veterano;

   (c) Encontre o número adicional de anos que um professor veterano irá continuar ensinando;

   (d) Note que o número de anos da letra b é muito maior do que o da letra a e ainda porque o da letra c é ainda maior. Explique porque isso já era esperado;

6. Um estudo de pacientes do Hospital Universitário da Carolina do Norte usa um modelo de cadeia de Markov com três categorias de pacientes: 0 (morte), 1 (estado desfavorável) e 2 (estado favorável). A matriz de transição para um ciclo de 72 horas é a seguinte:

$$\begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0,085 & 0,779 & 0,136 \\ 0,017 & 0,017 & 0,966 \end{array}\right] \end{array}$$

   (a) Ache a matriz fundamental.

   (b) Para um paciente com estado favorável, determine a expectativa do número de ciclos que o paciente vai continuar neste estado antes de morrer.

   (c) Para um paciente com estado desfavorável, determine a expectativa do número de ciclos que o paciente vai ficar em um estado favorável antes de morrer.

7. Os dados a seguir foram obtidos do setor de admissão para uma faculdade com um curso de dois anos de duração. Da classe do primeiro ano (P), 75% tornam-se no ano seguinte estudantes do segundo ano (S) e 25% abandonam o curso (A). Dentre aqueles que foram para o segundo, durante um certo ano, 90% se graduam (G) no ano seguinte, e 10% abandonam o curso.

   (a) Quantos e quais estados são absorventes?

   (b) Escreva a matriz de transição P na forma canônica.

   (c) Determine a matriz F.

   (d) Determine a probabilidade de que um estudante, ingressando no primeiro ano, irá se formar no final.

8. A cidade de Sacramento concluiu recentemente um novo sistema de trilhos para veículos leves (VLT) para levar passageiros e compradores ao centro da cidade e aliviar o congestionamento da rodovia. Os planejadores da cidade estimam que a cada ano, 15% daqueles que dirigem ou andam de automóvel vão mudar para o VLT; 80% irão continuar utilizando automóveis; e o restante não irá mais frequentar o centro da cidade. Entre os usuários do VLT estima-se que 5% irão voltar a utilizar automóveis e 80% irão continuar utilizando o VLT e o restante deixará de frequentar o centro da cidade. Considere que as pessoas que deixam de ir ao centro nunca volta a frequentar esta área da cidade.

   (a) Escreva a matriz de transição desta situação.

   (b) Qual a probabilidade de que uma pessoa que vai de carro para o centro deixe de frequentá-lo?

   (c) Determine a matriz fundamental F e o produto FR.

   (d) Encontre o número esperado de anos que uma pessoa que vá de automóvel para o centro irá levar para não mais frequentá-lo.

9. Realizou-se uma pesquisa onde dois animais de sexo opostos acasalavam, da sua prole pegavam-se aleatoriamente descendentes de sexo opostos e os colocavam para acasalar e assim por diante. Por uma questão de simplicidade admitiremos que a característica em questão independe do sexo. Seja **A** um gene dominante para alguma característica e **a** o gene recessivo. Os pais podem inicialmente carregar os genes AA, Aa ou aa. Há seis maneiras possíveis para a combinação de pais, elas estão listados na tabela abaixo.

   (a) Suponha que o acasalamento entre descendentes seja aleatório. Verifique se a matriz de transição desta situação corresponde a matriz abaixo:

Estado	1	2	3	4	5	6
Combinação	AA e AA	AA e Aa	AA e aa	Aa e Aa	Aa e aa	aa e aa

Tabela 5.2: Genética

$$\begin{array}{c c}
 & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{16} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
\end{array}$$

(b) Identifique os estados absorventes.

(c) Escreva a matriz na forma canônica.

(d) Determine a matriz fundamental e o produto FR.

(e) Se dois pais com os genes Aa acasalarem, encontre o número de pares de descendentes com esses genes que podem ser esperados antes que o gene dominante ou o gene recessivo não apareça mais.

(f) Se dois pais com genes Aa acasalarem, encontre a probabilidade de que o gene recessivo eventualmente desapareça.

10. Três carros de combate A, B e C vão batalhar. O carro de combate A tem probabilidade de $\frac{1}{2}$ de destruir o seu alvo, o B tem probabilidade de $\frac{1}{3}$ de destruir o seu alvo, e o C tem probabilidade de $\frac{1}{6}$ de destruir seu alvo. Os carros de combate atiram ao mesmo tempo e cada um atira sempre no seu mais forte oponente ainda não destruído. Usando como estados os carros sobreviventes em qualquer assalto (etapa), estabeleça uma cadeia de Markov e responda as seguintes questões:

(a) Quantos estados tem essa cadeia?

(b) Quantos estados são absorventes?

(c) Determine o número esperado de assaltos.

(d) Determine a probabilidade de que A sobreviva.

(e) Determine a probabilidade de que não haja sobreviventes.

5.3 Apêndice 3 - Probabilidades Carros de Combate

A seguir mostraremos como construir as probabilidades de transição do exercício dos carros de combate.

Nomearemos os carros de A, B e C e o estado onde não há nenhum dos carros sobrevivendo de N.

Na simbologia abaixo a letra "v"entre as letras que representam os carros de combate significa que o carro anterior sobrevive ao tiro de carro posterior e a letra "m"entre as letras que representam os carros de combate significa que o carro anterior morre com o tiro do carro posterior.

Lembre que os tiros são simultâneos e o tiro acontece sempre no mais forte oponente vivo.

Assim para calcular as probabilidades de transição, tendo como estados os carros que sobrevivem a cada etapa, temos:

- $ABC \to ABC = \underbrace{\frac{2}{3} \cdot \frac{5}{6}}_{\text{Av(B e C)}} \cdot \underbrace{\frac{1}{2}}_{\text{BvA}} \cdot \underbrace{1}_{\text{C vive}} = \frac{5}{18}$
- $ABC \to AC = \underbrace{\frac{2}{3} \cdot \frac{5}{6}}_{\text{Av(B e C)}} \cdot \underbrace{\frac{1}{2}}_{\text{BmA}} \cdot \underbrace{1}_{\text{C vive}} = \frac{5}{18}$
- $ABC \to BC = \underbrace{\frac{2}{3}}_{\text{AvB}} \cdot \underbrace{\frac{1}{6}}_{\text{AmC}} \cdot \underbrace{\frac{1}{2}}_{\text{BvA}} \cdot \underbrace{1}_{\text{C vive}} + \underbrace{\frac{1}{3}}_{\text{AmB}} \cdot \underbrace{\frac{1}{2}}_{\text{BvA}} \cdot \underbrace{1}_{\text{C vive}} = \frac{4}{18}$
- $ABC \to C = \underbrace{\frac{2}{3}}_{\text{AvB}} \cdot \underbrace{\frac{1}{6}}_{\text{AmC}} \cdot \underbrace{\frac{1}{2}}_{\text{BmA}} \cdot \underbrace{1}_{\text{C vive}} + \underbrace{\frac{1}{3}}_{\text{AmB}} \cdot \underbrace{\frac{1}{2}}_{\text{BmA}} \cdot \underbrace{1}_{\text{C vive}} = \frac{4}{18}$
- $AC \to AC = \underbrace{\frac{5}{6}}_{\text{AvC}} \cdot \underbrace{\frac{1}{2}}_{\text{CvA}} = \frac{5}{12}$
- $AC \to A = \underbrace{\frac{5}{6}}_{\text{AvC}} \cdot \underbrace{\frac{1}{2}}_{\text{CmA}} = \frac{5}{12}$
- $AC \to C = \underbrace{\frac{1}{6}}_{\text{AmC}} \cdot \underbrace{\frac{1}{2}}_{\text{CvA}} = \frac{1}{12}$
- $AC \to N = \underbrace{\frac{5}{6}}_{\text{AmC}} \cdot \underbrace{\frac{1}{2}}_{\text{CmA}} = \frac{5}{12}$
- $BC \to BC = \underbrace{\frac{5}{6}}_{\text{BvC}} \cdot \underbrace{\frac{2}{3}}_{\text{CvB}} = \frac{10}{18}$

- BC $\rightarrow$ C $= \underbrace{\frac{1}{6}}_{\text{BmC}} \cdot \underbrace{\frac{2}{3}}_{\text{CvB}} = \frac{2}{18}$
- BC $\rightarrow$ B $= \underbrace{\frac{5}{6}}_{\text{BvC}} \cdot \underbrace{\frac{1}{3}}_{\text{CmB}} = \frac{5}{18}$
- BC $\rightarrow$ N $= \underbrace{\frac{1}{6}}_{\text{BmC}} \cdot \underbrace{\frac{1}{3}}_{\text{CmB}} = \frac{1}{18}$

  Os estados onde apenas um dos carros sobrevive ou onde nenhum dos carros de combate sobrevivem são absorventes.

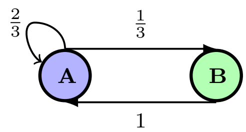
5.4 Apêndice 4 - Soluções

5.4.1 1º Bloco - Noções Básicas

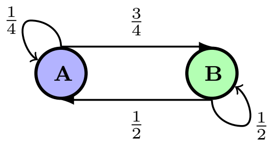
Usaremos $(\cdot)$ como separador decimal nas soluções abaixo.

1. a) Não representa uma cadeia de Markov, pois apesar das entradas da matriz serem maiores ou iguais a zero e menores do que 1, o somatório das entradas das linhas da matriz não é igual a 1.

   b) Sim.



   c) Sim.



   d) Não representa uma cadeia de Markov, pois existe entrada maior do que 1.

2. Fazendo A = estado1, B=estado2 e C=estado3

   a) $\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}$ b) $\begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.5 & 0.1 & 0.4 \\ 0 & 0.8 & 0.2 \end{bmatrix}$ c) $\begin{bmatrix} 0.1 & 0.3 & 0.6 \\ 0.85 & 0 & 0.15 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}$

3. a) A probabilidade de transição p_{21}.
   b) 1ª observação $[\frac{1}{3}, \frac{2}{3}]$
   2ª observação $[0.28, 0.72]$
   c) 1ª observação $[\frac{1}{4}, \frac{3}{4}]$
   2ª observação $[0.27, 0.73]$

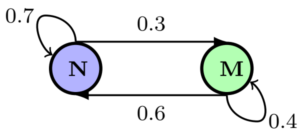
4. a) A probabilidade de transição p_{12}.
   b) 1ª observação $[0.1, 0.9]$
   2ª observação $[0.46, 0.54]$
   c) 1ª observação $[0.5, 0.5]$
   2ª observação $[0.3, 0.7]$

5. $[0.4, 0.275, 0.625]$

6. $a = 0.4$, $b = 0.1$ e $c = 1$

7. a) 0 b) 0.5185 c) 0.4398 d) 0.279 e) 0.198

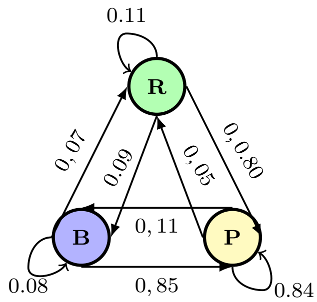
8. a)



   b) $P = \begin{bmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{bmatrix}$ c) $\pi^{(0)} = [1, 0]$

d) Basta determinar o valor de $\pi^{(2)} = \pi^{(0)} \cdot P^2 = \frac{1}{3}$

9\.

a) 

b) $P = \begin{bmatrix} 0.08 & 0.07 & 0.85 \\ 0.09 & 0.11 & 0.80 \\ 0.11 & 0.05 & 0.84 \end{bmatrix}$

c) $\pi^{(3)} = [0.1057, 0.0554, 0.8388]$, logo a probabilidade de se ter uma safra boa em 1943, dado que a safra foi boa em 1940 é de 10,57%.

10\.

a) Elevando a matriz de transição dada a quarta potência podemos ver que 26,03% dos coelhos permanecem no estado 1 após 4 semanas.

b) Tomando o vetor $\pi^{(0)} = [9, 4, 0, 0]$ e multiplicando pela matriz encontrada no item a temos que após 4 semanas teremos a seguinte distribuição dos coelhos $[2.34, 2.62, 3.07, 4, 96]$.

5.4.2 2º Bloco - Vetor de probabilidade

1. Apenas a matriz do item c) não é regular.

2\.

a) $[0.4, 0.6]$

b) $[0.4615, 0.5385]$

c) $[0.1687, 0.2289, 0.6024]$

d) $[0.2540, 0.2222, 0.5238]$

3. Basta elevar a matriz de transição até $n \geq 24$ numa aproximação de 4 casas decimais para que se chegue a distribuição estacionária de 30,77% da marca A, 46,15% marca B e 23,08% da marca C.

4\.

a) O parentesco de avó acontece na segunda geração, portanto basta elevar a matriz de transição ao quadrado. Assim temos que a probabilidade de que uma avó loira tenha uma neta morena é de 30%

b) i) $[0.377, 0.375, 0.248]$

ii) $[0.35, 0.40, 0.25]$

5.

a) $P = \begin{bmatrix} 0.8 & 0.18 & 0.02 \\ 0.4 & 0.5 & 0.1 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}$

b) Observando P^2 vemos que a probabilidade de que um neto de um concluinte do ensino fundamental terminar a faculdade é de 44%.

c) Novamente observando P^2 temos que a probabilidade pedida é de 38,2%.

d) $[0.5993, 0.3358, 0.0649]$

e) $[0.6489, 0.2877, 0.0534]$

6.

a) $P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix}$

b) Como o vetor estacionário da matriz P é igual a $[\frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{4}]$, temos que a distribuição final dos robôs será da por $[10, 15, 20, 15]$

5.4.3 3º Bloco - Cadeias de Markov Absorventes

1.

a) Estado 2

b) Estado 2

c) Nenhum

d) Estado 2 e Estado 4

2. Apenas as matrizes dos itens a) e d) representam uma cadeia de Markov Absorvente

3.

a) $P = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} 2.5 \end{bmatrix}$

b) $P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} 1.5 \end{bmatrix}$

c) $P = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} \frac{4}{3} & \frac{1}{3} \\ 0 & 1 \end{bmatrix}$

d) $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$

e) $P = \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0 & 0.3 & 0.5 & 0.1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} 1.47059 & 0.29412 \\ 0.14706 & 1.02941 \end{bmatrix}$

f) $P = \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0.2 & 0.2 \\ 0.1 & 0 & 0.3 & 0.5 & 0.1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$; $F = \begin{bmatrix} \frac{10}{9} & \frac{185}{999} \\ 0 & \frac{5}{3} \end{bmatrix}$

4. a) $f_{13} = 0.5$ e $f_{23} = 1$

   b) $f_{31} + f_{32} + f_{33} = 5$ e $f_{11} + f_{12} + f_{13} = 4$

5. A matriz fundamental neste caso será dada por $F = \begin{bmatrix} \frac{10}{3} & \frac{33}{90} & \frac{2365}{900} \\ 0 & 1 & \frac{645}{90} \\ 0 & 0 & \frac{75}{9} \end{bmatrix}$

   a) Aproximadamente 2.63 anos

   b) Aproximadamente 7.17 anos

   c) Aproximadamente 8.33 anos

   d) Um estudante, por sua inexperiência, não sabe direito, ou ainda pode modificar sua opção em sua carreira. Enquanto que um recém formado apenas irá mudar de profissão se, de fato, não se identifica de jeito nenhum com sua escolha ou se arranjar melhor remuneração em outra, o mesmo vale para o professor veterano.

6. a) $F = \begin{bmatrix} 6.536 & 26.144 \\ 3.268 & 42.484 \end{bmatrix}$

   b) 42.484 ciclos ou 127.45 dias.

   c) 26.144 ciclos ou 78.43 dias

7. a) 2 estados. Graduar-se e Abandonam

   b) $P = \begin{array}{c} P \\ S \\ G \\ A \end{array} \begin{bmatrix} 0 & 0.75 & 0 & 0.25 \\ 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

   c) $F = \begin{bmatrix} 1 & 0.75 \\ 0 & 1 \end{bmatrix}$

   d) Neste caso temos que $B = \begin{bmatrix} 0.675 & 0.375 \\ 0.9 & 0.1 \end{bmatrix}$ portanto a probabilidade pedida é 67,5%.

8. A - automóvel; DC - Deixa de ir ao centro

a) $$P = \begin{array}{c} A \\ VLT \\ DC \end{array} \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.05 & 0.8 & 0.15 \\ 0 & 0 & 1 \end{bmatrix}$$

b) 5% .

c) $$F = \begin{bmatrix} 6.154 & 4.615 \\ 1.538 & 6.154 \end{bmatrix}; \; B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

d) 10.769 anos.

9. a) Sim. Basta levar em consideração as leis de Mendell e lembrar de considerar os casos inversos em cada casal.

b) Estado 1 e estado 6.

c)

$$\begin{array}{c|cccc|cc} & 2 & 3 & 4 & 5 & 1 & 6 \\ \hline 2 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 3 & 0 & 0 & 1 & 0 & 0 & 0 \\ 4 & \frac{1}{4} & \frac{1}{16} & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} \\ 5 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ \hline 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

d) $$F = \begin{bmatrix} \frac{8}{3} & \frac{1}{6} & \frac{4}{3} & \frac{2}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} & \frac{4}{3} \\ \frac{4}{3} & \frac{1}{3} & \frac{8}{3} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{6} & \frac{4}{3} & \frac{8}{3} \end{bmatrix};$$

$$B = \begin{array}{c} \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 6 \end{array} \\ \begin{bmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \end{array}$$

e) $\frac{17}{3}$ ou 5.6 gerações.

f) 50%

10. a) 7 estados.

b) 4 estados.

c) $$t = \begin{array}{c} ABC \\ AB \\ BC \end{array} \begin{bmatrix} 2.7363 \\ 1.7143 \\ 2.25 \end{bmatrix}$$

d) Soma de todas as possíveis absorções efetuadas pelo estado A 81.73%.

e) Soma de todas as possíveis absorções efetuadas pelo estado (N) nenhum sobrevivente 36.13%.

Capítulo 6

Anexos

Nestes anexos temos algumas definições e teoremas que foram usados ao longo do texto e que são necessárias para definir e operar os objetos associados as cadeias de Markov. É composto por duas seções, uma dedicada a definições e teoremas associados a álgebra, mais especificamente a álgebra linear, e um outro, a conceitos associados à teoria da probabilidade.

6.1 Álgebra Linear

Usaremos como referência para as definições e teoremas abaixo (HEFEZ e FERNANDEZ, 2016), livro texto do PROFMAT para álgebra linear. Vamos começar definindo o que é uma matriz.

Definição 6.1.1. *Dados m e n em $\mathbb{N}$, definimos uma* ***matriz de ordem m por n****, ou simplesmente uma matriz m por n(escreve-se $m \times n$), como uma tabela formada por elementos de $\mathbb{R}$ distribuídos em m linhas e n colunas. Estes elementos de $\mathbb{R}$ são chamados de* ***entradas*** *da matriz.*

É usual indicarmos as entradas de uma matriz arbitrária A por A_{ij} ou por a_{ij}, onde os índices indicam, nessa ordem, a linha e a coluna onde o elemento se encontra. Assim uma matriz A de ordem $m \times n$ é usualmente representada por

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

A depender dos valores de m e n, uma matriz $m \times n$ recebe um nome especial. De fato, toda matriz $1 \times n$ é denominada uma *matriz linha* e toda matriz $m \times 1$ é denominada *matriz coluna*. Uma matriz $n \times n$ é chamada de *matriz quadrada de ordem n*. Se A é uma matriz quadrada de ordem n, as entradas a_{ii}, com $1 \leq i \leq n$,formam a *diagonal principal* de A. Uma matriz cujas entradas são todas iguais a zero é chamada *matriz nula*.

Definição 6.1.2. *Uma **matriz diagonal de ordem n** é uma matriz quadrada de ordem n em que os elementos que não pertencem a diagonal principal são iguais a zero.*

Definição 6.1.3. *A matriz diagonal de ordem n cujas entradas da diagonal principal são iguais ao número 1, é chamada **matriz identidade de ordem n** e usualmente denotada por I_n.*

Seja $M(m,n)$ o símbolo que denota o conjunto das matrizes de ordem $m \times n$.

Definição 6.1.4. *Dizemos que duas matrizes $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$, de mesma ordem, são **iguais** escrevendo $A = B$, quando $a_{ij} = b_{ij}$ para todo $1 \leq i \leq m$ e para todo $1 \leq j \leq n$.*

Definição 6.1.5. *Se $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$ são duas matrizes de mesma ordem $m \times n$, **a soma de A e B**, denotada por $A + B$, é a matriz $C = [c_{ij}]$ de ordem $m \times n$ tal que $c_{ij} = a_{ij} + b_{ij}$, para todo $1 \leq i \leq m$ e para todo $1 \leq j \leq n$.*

Definição 6.1.6. *Dada uma matriz $A = [a_{ij}]$, define-se a **matriz oposta de A**, como a matriz $-A = [-a_{ij}]$.*

A adição de matrizes também tem propriedades semelhantes à adição nos números reais, ou seja, associatividade, comutatividade, elemento neutro e elemento oposto.

Proposição 6.1.1. *Se A, B e C são matrizes de mesma ordem $m \times n$, então:*

i) $A + (B + C) = (A + B) + C$ *(associatividade da adição);*

ii) $A + B = B + A$ *(comutatividade da adição);*

iii) $A + 0 = A$, *onde* 0 *é a matriz nula* $m \times n$;

iv) *A + (-A) = 0.*

Todas as proposições acima decorrem das definições de igualdade e de adição de matrizes feitas anteriormente, combinada com as respectivas propriedades que são válidas no corpo dos números reais para cada uma das entradas.

Definição 6.1.7. *Dada a matriz $A = [a_{ij}]_{m\times n}$, definimos **o produto de A pelo número real k**, como sendo a matriz $kA = [ka_{ij}]_{m\times n}$.*

Tendo definido a operação de adição sobre o conjunto $M(m,n)$, podemos definir a operação de **subtração** da maneira usual, isto é: dada duas matrizes A e B em $M(m,n)$,

$$A - B = A + (-B)$$

Proposição 6.1.2. *Para quaisquer A e $B \in M(m,n)$, e a e b em $\mathbb{R}$, temos que:*

i) $a \cdot (A + B) = aA + aB$;

ii) $(a + b) \cdot A = aA + bA$;

iii) $a \cdot (bA) = (ab) \cdot A$;

iv) $1A = A$.

As proposições acima podem ser demonstradas diretamente através da definição 6.1.7 e das propriedades da multiplicação de números reais.

A seguir temos à definição clássica da *multiplicação de matrizes* conforme Artur Cayley(1821-1895) o fez em seu trabalho "A Memoir on the Theory of Matrices"publicada em 1858 na revista *Philosophical Transactions of the Royal Society of London.* Cayley neste trabalho já notara, que assim definida, a multiplicação de matrizes era uma operação que não obedecia a lei do corte, nem tinha a propriedade comutativa, e ainda que uma matriz não nula nem sempre era invertível, vejamos a definição:

Definição 6.1.8. *Sejam $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{n\times p}$ duas matrizes. O produto AB, de A por B, nesta ordem, é definido como a matriz $C = [c_{ij}]_{m\times p}$ tal que*

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1} + b_{1j} + \cdots + a_{in}b_{nj}$$

Para todo $1 \leq i \leq m$ e para todo $1 \leq j \leq p$.

Note que, para o produto entre as matrizes A e B estar definido é necessário que o número de colunas de A seja igual ao número de linhas de B.

Proposição 6.1.3. *Sejam A, B e C matrizes, desde que as operações sejam possíveis, temos que:*

i) $A \cdot (B + C) = AB + AC$;

ii) $(A + B) \cdot C = AC + BC$;

iii) $(AB) \cdot C = A \cdot (BC)$;

iv) $IA = AI = A$, *onde I é a matriz identidade.*

Definida a multiplicação de matrizes, definimos a *potenciação* de matrizes da seguinte maneira. Sejam $A \in M(n,n)$ e $k \in \mathbb{N}$, com $k \geq 2$ temos que:

$$A^0 = I_n, \qquad A^1 = A \qquad e \qquad A^k = \underbrace{A \cdot A \cdots A}_{k fatores}$$

Definição 6.1.9. *Dada uma matriz $A = [a_{ij}]_{m\times n}$ chamamos de **transposta** de A, e denotamos por A^t, matriz $[b_{ij}]_{n\times m}$, onde $b_{ij} = a_{ji}$, $\forall i$ tal que $1 \leq i \leq n$ e $\forall j$ tal que $1 \leq j \leq n$.*

Uma matriz quadrada A é chamada *simétrica* se $A^t = A$ e antissimétrica se $A^t = -A$

Uma matriz A é chamada de uma *matriz em blocos* se A está subdividida em matrizes menores, chamadas **blocos**. Esta subdivisão geralmente é feita por linhas horizontais e verticais, vejamos um exemplo:

$$\mathbf{P} = \left[\begin{array}{ccc|cc} 2 & -1 & 3 & 2 & 0 \\ 1 & -2 & 9 & 1 & 1 \\ 0 & 0 & 1 & 6 & 7 \\ \hline 1 & -1 & -1 & 9 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Uma propriedade interessante da partição de uma matriz em blocos é que o resultado das operações de adição e multiplicação com matrizes em blocos, podem ser obtidos efetuando o cálculo com os blocos, como se os mesmos fossem elementos da matriz.

Definição 6.1.10. *Dada uma matriz quadrada A de ordem n, chamamos de **inversa** de A a uma matriz quadrada B de ordem n tal que $AB = BA = I_n$.*

Note que nem toda matriz quadrada possui inversa. Basta observa que a matriz quadrada nula de ordem n multiplicada por uma matriz quadrada X qualquer de ordem n será sempre diferente de I_n.

Mesmo matrizes não nulas podem não ter inversa, como por exemplo a matriz $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ não possui inversa, pois não existe uma matriz quadrada B de ordem 2, tal que $AB = I_2$.

Uma matriz quadrada A é chamada *invertível* se A admite uma matriz inversa.

Teorema 6.1.1. *Se uma matriz A possui uma inversa, então essa inversa é única.*

Demonstração. De fato, suponhamos que B e C são inversas de uma matriz A de ordem n. Então $AB = I_n$ e $AC = I_n$ pela proposição 1.1.3, temos que:

$$C = C \cdot I_n = C(AB) = (CA)B = I_n \cdot B = B$$

■

Como, se existir, a inversa de uma matriz A é única, escrevemos A^{-1} para denotar a inversa de A.

Definição 6.1.11. ***Equação Linear*** *nas incógnitas* $x_1, x_2, \ldots, x_n$ *é toda equação do tipo:* $\mathbf{a_1x_1 + a_2x_2 + \cdots + a_nx_n = b}$ *em que* $a_1, a_2, \ldots, a_n$ *e* b *são coeficientes reais. Chamamos* b *de coeficiente(ou termo) independente da equação.*

Uma equação linear é chamada de *homogênea* quando seu termo independente é igual a zero.

Definição 6.1.12. ***Sistema Linear*** *é o conjunto* $S_{m\times n}$ *formado por* m *equações lineares e* n *incógnitas, que pode ser indicado da seguinte maneira:*

$$S = \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

onde os a_{ij}*'s e* b_i*'s, para* $1 \leq i \leq m$ *e* $1 \leq j \leq n$*, são números reais dados.*

Definição 6.1.13. *A* ***solução*** *de um sistema linear* $m\times n$ *é toda ênupla* $(\alpha_1, \alpha_2, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ *que é solução de cada uma das* m *equações desse sistema.*

Vale lembrar que um sistema linear pode ter conjunto solução vazio e neste caso o sistema linear é chamado de *impossível.* No caso em que seu conjunto solução é unitário chama-se de *possível e determinado.* Ou ainda, se o conjunto solução tiver infinitos elementos é chamado de *possível e indeterminado.*

Determinar ou descrever o mais explicitamente possível a solução de um sistema linear é a tarefa fundamental que temos quando nos deparamos com os mesmos. Desde a antiguidade as mais diversas áreas do conhecimento modelam problemas do cotidiano através de sistemas lineares.

Para resolver um sistema linear utilizamos uma série de *transformações elementares* para simplificar o sistema de modo que possamos explicitar sua solução. Chamamos de transformações elementares as seguintes operações:

i) Trocar a posição relativa de duas equações do sistema;

ii) Trocar uma equação pela soma membro a membro da própria equação com um múltiplo de outra;

iii) Trocar uma equação dada por um de seus múltiplos, ou seja, multiplicar ambos os membros da equação por um número real não-nulo.

Dizemos que dois sistemas lineares são *sistemas equivalentes*, se pudermos obter um sistema do outro a partir de uma sequência finita de transformações lineares. Observe que esta relação entre dois sistemas é uma relação de equivalência(reflexiva, transitiva e simétrica), portanto sistemas lineares equivalentes possuem mesmo conjunto solução.

É importante salientar que estas transformações elementares nos permitem determinar, se existir, a solução de um sistema linear qualquer, porém a medida que à quantidade de equações e incógnitas aumentam, este processo se torna muito trabalhoso.

O que há de essencial em um sistema de equações lineares são os seus coeficientes. Sejam os vetores $(a_{m1}, a_{m2}, \ldots, a_{mn}, b_m) \in \mathbb{R}^{n+1}$ que representam os coeficientes de um sistema S, acrescido do seu termo independente. Se os organizarmos em linhas, teremos a *matriz ampliada* associada ao sistema S:

$$\left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & b_m \end{array}\right]$$

Quando o sistema de equações lineares é homogêneo, a matriz associada à ele contém apenas os seus coeficientes, eliminando-se a coluna de zeros da direita da matriz.

Sendo assim podemos representar um sistema de equações lineares, de modo perfeito, através da equação matricial $AX = B$, onde:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \text{ e } B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}$$

As matrizes A, X e B são chamadas, respectivamente, de matriz dos *coeficientes do sistema*, *matriz das incógnitas*, *matriz dos termos independentes*.

Note que com a forma matricial de um sistema linear, ganhamos mais uma maneira de determinar sua solução caso a matriz inversa da matriz dos coeficientes exista, pois pela definição de matriz inversa e pela proposição 6.1.3, temos:

$$AX = B \Rightarrow X = (A^{-1}A)X = A^{-1}(AX) = A^{-1}B$$

Um fato importante sobre a forma matricial de um sistema linear é que se tivermos uma forma de caracterizar quando uma matriz é invertível, temos uma forma eficiente de determinar quando um sistema linear possui ou não uma solução, o que já é um avanço na tentativa de solucionar o mesmo. Para isso precisamos do conceito do determinante de uma matriz, que iremos definir a seguir.

Foi Colin Maclaurin(1698-1746) em seu trabalho de título *Um Tratado sobre Álgebra em Três Partes*, publicado em 1748 quem pela primeira vez utilizou o que chamamos hoje de *Regra de Cramer*, para resolver sistemas com até 4 incógnitas e 4 equações. Entretanto foi Gabriel Cramer(1704-1752) quem generalizou seu método, por isso a regra é chamada de Cramer até os dias de hoje. Apenas a título de ilustração e motivação do conceito de determinante, vejamos como ele extraiu seu teorema para os casos onde os sistemas são 2×2 e 3×3. Sejam

$$S_1 = \begin{cases} ax & + & by & = & e \\ cx & + & dy & = & f \end{cases} \quad e \quad S_2 = \begin{cases} ax & + & by & + & cz & = & m \\ dx & + & ey & + & fz & = & n \\ gx & + & hy & + & kz & = & p \end{cases}$$

Aplicando as transformações lineares adequadas Maclaurin, descobriu que a solução de sistema S_1, se $ad - bc \neq 0$ era igual a:

$$x = \frac{ed - fb}{ad - bc}, \quad y = \frac{af - ce}{ad - bc}$$

E que se $aek - ahf + dhc - dbk + gbf - gec \neq 0$ a solução de S_2 seria dada por:

$$x = \frac{mek - mfh + bfp - bnk + cnh - cep}{aek - ahf + dhc - dbk + gbf - gec},$$

$$y = \frac{nak - ncg + mfg - mdk + pcd - paf}{aek - ahf + dhc - dbk + gbf - gec} \ e$$

$$z = \frac{aep - ahn + dhm - dbp + gbn - gem}{aek - ahf + dhc - dbk + gbf - gec}$$

.

Maclaurin notou que, tanto para S_1 como para S_2 o numerador que aparece nas soluções são os mesmos e que consiste na soma alternada de vários produtos dos coeficientes das incógnitas do sistema. Outro fato que ele observou é que os numeradores também são somas alternadas de produtos dos coeficientes das demais incógnitas e dos termos independentes do sistema. Os numeradores e denominadores nas soluções de Maclaurin foi o que motivou o conceito de *determinantes* que temos hoje, termo introduzido por Gauss em 1801. A definição formal para o determinante, ou melhor, para a *função determinante* de uma matriz quadrada A de ordem n, é um pouco mais elaborada e depende de outras definições que daremos a seguir seguindo os passos que podem ser encontrados em (ANTON, 2001).

Definição 6.1.14. *Dados n objetos distintos $a_1, \ldots, a_n$ uma* ***permutação*** *destes objetos consiste em dispô-los em uma determinada ordem.*

A *quantidade de permutações* de n objetos distintos é dada por $n!$, que é lido como n fatorial e $n! = n \cdot (n-1) \cdot (n-2) \ldots 3 \cdot 2 \cdot 1$ se $n > 0$ e define-se ainda que $0! = 1$.

Definição 6.1.15. *Dada uma permutação dos inteiros $1, 2, \ldots, n$ existe uma* ***inversão*** *quando um inteiro precede outro menor do que ele.*

Uma permutação de números inteiros será classificado como *par* se o seu número de inversões for par e será classificada como *ímpar* se o seu número de inversões for ímpar.

Para ilustrar a definição acima observe a tabela abaixo que nos mostra o número de inversões e sua classificação para cada permutação dos números 1,2 e 3.

Tabela 6.1: Número de inversões

Permutação	**Nºde inversões**	**Classificação**
(123)	0	Par
(132)	1	Ímpar
(213)	1	Ímpar
(231)	2	Par
(312)	2	Par
(321)	3	Ímpar

Definição 6.1.16. *Seja a matriz quadrada $A = [a_{ij}]$ de ordem n, dizemos que um produto de n entradas de A, tais que não há duas entradas de mesma linha ou de mesma coluna de A, de um* ***produto elementar*** *da matriz A.*

Observe que a quantidade de produtos elementares de uma matriz quadrada de ordem n será dado por $n!$, visto que pela definição acima cada uma das escolhas que fizermos para um produto elementar estará associada uma permutação das n colunas (ou linhas) da matriz. Assim esses produtos são da forma $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$, onde $(j_1, j_2, j_3, \ldots, j_n)$ é uma permutação do conjunto $\{1, 2, \ldots, n\}$.

Um produto elementar $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ multiplicado por $+1$ ou -1 é chamado de *produto elementar com sinal.* Usamos o sinal $(+)$ se $(j_1, j_2, j_3, \ldots, j_n)$ é uma permutação par e o sinal de $(-)$ no caso contrário.

Definição 6.1.17. *Seja A uma matriz quadrada de ordem n. A* ***função determinante*** *que associa a cada matriz A um número chamado determinante de A, é denotada por $det(A)$ ou $|A|$ e é definida como a soma de todos os produtos elementares com sinal de A.*

Note que o cálculo de determinantes para matrizes de ordem maiores que 3 não é uma tarefa muito simples de computar dado que, por exemplo, para calcular o determinante de uma matriz de ordem 6, pela definição, precisaríamos fazer 6!=720 produtos e depois efetuar a soma dessas 720 parcelas. A seguir temos algumas propriedades sobre o determinante de uma matriz que irá nos ajudar a provar alguns resultados sobre as Cadeias de Markov.

Proposição 6.1.4. *Seja A uma matriz quadrada de ordem n, se todos os elementos de uma linha (ou coluna) de A forem iguais à zero então* $det(A) = 0$.

Demonstração. Basta observar que se todos os elementos de uma linha (ou coluna) são iguais a zero então todos os produtos elementares sempre terão um fator igual a zero, logo todos os produtos elementares serão iguais a zero e portanto pela definição do determinante temos $det(A) = 0$. ■

Proposição 6.1.5. *O determinante da matriz identidade* I_n *é igual a 1 para todo n.*

Demonstração. Lembre que a matriz identidade I_n tem todas as suas entradas iguais a zero exceto à sua diagonal principal, que tem entradas iguais a 1. Assim todos os produtos elementares de I_n serão iguais a zero pois terão um fator igual a zero, exceto o produto elementar formado pelos elementos de sua diagonal principal que neste caso é igual a 1. Como na diagonal principal à permutação formada pelos números das suas colunas, tem zero inversões, isto é, é uma permutação par, temos $det(I_n) = 1$ ■

Teorema 6.1.2. *Sejam A e B matrizes de ordem* $n \times n$ *quaisquer, então o determinante do produto de A e B é igual ao produto entre os determinantes de A e de B, ou seja,* $det(A \cdot B) = det(A) \cdot det(B)$.

A demonstração desse teorema requer uma série de resultados sobre *matrizes elementares*, além de resultados associados ao *escalonamento* de matrizes, é uma prova muito bonita, porém muito extensa e foge ao escopo desse trabalho. O leitor que tiver interesse pode consultar (BOLDRINI, 1980, p. 84).

Teorema 6.1.3. *Seja A uma matrizes de ordem* $n \times n$ *quaisquer, dizemos que A é uma matriz invertível se* $det(A) \neq 0$.

Demonstração. Sabemos que por definição $A.A^{-1} = I_n$, aplicando o determinante a ambos os membros dessa igualdade temos $det(A.A^{-1}) = det(I_n)$, como vimos na proposição 6.1.5 $det(I_n) = 1$, para todo n. Pelo teorema 6.1.2 temos que $det(A.A^{-1}) = det(A).det(A^{-1})$, logo $det(A^{-1}) = \frac{1}{det(A)}$. Assim $det(A^{-1})$, só existe, se $det(A) \neq 0$. Portanto A é invertível quando $det(A) \neq 0$ ■

Com o teorema acima encerramos os resultados relacionados a álgebra linear, que são necessários para desenvolver os conceitos e resultados associados as Cadeias de Markov.

Vale salientar, que a resolução de um sistema linear de equações, ou a determinação de uma matriz inversa, ou o cálculo de determinantes são tarefas que para matrizes de ordem muito grande, e muito grande nesse caso seria algo em torno de 6 ou 7, o uso de alguma ferramenta computacional é indispensável para efetuar tais tarefas com sucesso.

No caso do cálculo de determinantes a *expansão de Laplace* é um método interessante para trabalhar com matrizes de ordem 4, mas se torna inviável para matrizes de ordem maiores que 4, pois seu custo computacional cresce muito. O mesmo vale para a *regra de Cramer* na resolução de sistemas, que apesar de ser uma fórmula eficaz para resolver sistemas lineares tem um custo computacional também muito grande até mesmo para computadores atuais. Vale o mesmo comentário para a determinação da matriz inversa através do seu determinante e sua matriz adjunta, o custo computacional para efetuar as contas é muito alto.

Por isso é necessário que ao trabalhar com as Cadeias de Markov, se tenha acesso ao Geogebra para auxiliar os cálculos. Pois a quantidade de contas necessárias pode acabar entediando os estudantes e prejudicando o foco nos resultados associados as cadeias. Na próxima seção definiremos os conceitos e resultados básicos associados a probabilidade que são essenciais para o desenvolvimento das Cadeias de Markov.

6.2 Probabilidade

Usaremos (DANTAS, 2013) e (MORGADO e CARVALHO, 2015) como fontes bibliográficas para as definições e teoremas desta seção. Além disso, como já fizemos no caso da álgebra linear, vamos apenas definir e mostrar aquilo que nos parece imprescindível a compreensão das cadeias de Markov. Comecemos então com a definição de experimento aleatório.

Definição 6.2.1. *Experimentos que ao serem repetidos nas mesmas condições, não produzem o mesmo resultado são denominados* ***experimentos aleatórios****.*

É importante ao falar sobre experimentos aleatórios nas aulas de matemática no ensino médio, fazer a contraposição dessa definição com a de experimento *determinístico* muito utilizada nas ciências naturais e na maioria das vezes utilizando apenas o termo experimento para designá-lo, o que pode confundir os estudantes e já criar um obstáculo cognitivo na aprendizagem do assunto.

Definição 6.2.2. *Chamamos de* ***espaço amostral*** *associado a um experimento o conjunto de seus resultados possíveis. Denotaremos com a letra grega Ω o espaço amostral.*

Definição 6.2.3. *Chamamos de* ***evento*** *a todo resultado ou subconjunto de resultados de um experimento.*

Vale lembrar que um evento é representado por um subconjunto do conjunto que forma o espaço amostral. Os eventos representados por um conjunto unitário serão chamados de *eventos simples*. Um evento que nunca ocorre é chamado de *evento impossível*, um evento que sempre ocorre é chamado de *evento certo*.

Sendo o espaço amostral um conjunto e os eventos associados a ele subconjuntos, podemos fazer operações entre os eventos.

Definição 6.2.4. *A* ***reunião*** *de dois eventos A e B, denotada $A \cup B$, é o evento que ocorre se pelo menos um deles ocorre.*

Definição 6.2.5. *A* ***intersecção*** *de dois eventos A e B, denotada $A \cap B$, é o evento que ocorre se ambos ocorrem.*

Definição 6.2.6. *O* ***complementar*** *do evento A, denotado $\overline{A}$, é o evento que ocorre quando A não ocorre.*

Definição 6.2.7. *A* ***diferença*** *de dois eventos A e B, denotado $A \backslash B$, é o evento em que ocorre A e* ***não*** *ocorre B.*

Definição 6.2.8. *Dizemos que o evento A* ***implica*** *o evento B, que denotamos $A \subset B$, se para todo $a \in A$ tivermos que $a \in B$.*

Dizer que o evento A implica o evento B significa dizer que a ocorrência de A garante à ocorrência de B.

Os eventos A e B são *iguais* se $A \subset B$ e $B \subset A$.

Definição 6.2.9. *Os eventos A e B são ditos* ***mutuamente excludentes****, se eles não podem ocorrer simultaneamente, ou seja, se $A \cap B = \emptyset$.*

As operações acima definidas, estão relacionadas as análogas operações com conjuntos e gozam das mesmas propriedades. Podemos agora definir o que é a probabilidade de evento num espaço amostral Ω.

Definição 6.2.10. ***Probabilidade*** *é uma função definida numa classe $\mathcal{F}$ de eventos de Ω que satisfaz as seguintes condições:*

a*)* $P(A) \geq 0$ *para todo* $A \in \mathcal{F}$*;*

b*) Se $(A_n)_{n\geq 1}$ é uma sequência de eventos de (F), que são mutuamente excludentes, então: $P(\cup_{n=1}^{k} A_n) = \displaystyle\sum_{n=1}^{k} P(A_n)$, com $k \in \mathbb{N}$;*

c*)* $P(\Omega) = 1$

É relevante dizer que a definição de probabilidade busca uma forma eficaz de associar números aos eventos(conjuntos). E as propriedades abaixo nos mostram como ocorre a relação entre as operações dos eventos e a aritmética dos números associados a eles, ou seja, suas probabilidades. Vejamos como isso acontece.

Proposição 6.2.1. *Denotemos por ϕ o evento impossível. Temos $P(\phi) = 0$.*

Demonstração. Note que $P(\Omega) = P(\Omega \cup \phi)$, como o evento certo e o evento impossível são mutuamente excludentes, por definição temos que $P(\Omega \cup \phi) = P(\Omega) + P(\phi)$. Assim por transitividade $P(\Omega) = P(\Omega) + P(\phi)$, logo $P(\phi) = 0$. ■

Proposição 6.2.2. *Seja $A \subset \Omega$ um evento e $\overline{A}$ seu evento complementar. Temos que $P(\overline{A}) = 1 - P(A)$.*

Demonstração. Pelo item c) da definição 6.2.10 $P(\Omega) = 1$, sabemos ainda que $A \cup \overline{A} = \Omega$, portanto $P(\Omega) = P(A \cup \overline{A}) = 1$, como A e $\overline{A}$ são mutuamente excludentes, pelo item b) da definição de probabilidade temos $P(A) + P(\overline{A}) = 1$, logo $P(\overline{A}) = 1 - P(A)$. ■

Proposição 6.2.3. *Sejam A e B dois eventos do espaço amostral Ω. Temos que $P(A \backslash B) = P(A) - P(A \cap B)$.*

Demonstração. Note que $(A \backslash B) \cup (A \cap B) = A$. Assim $P[(A \backslash B) \cup (A \cap B)] = P(A)$. Mas $(A \backslash B)$ e $(A \cap B)$ são eventos mutuamente excludentes, portanto $P(A \backslash B) + P(A \cap B) = P(A)$ o que equivale a $P(A \backslash B) = P(A) - P(A \cap B)$. ■

Proposição 6.2.4. *Se A e $B \subset \Omega$ então $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.*

Demonstração. Temos que $P(A \cup B) = P[(A \backslash B) \cup B] = P(A \backslash B) + P(B)$, pois $A \backslash B$ e B, são mutuamente excludentes. Pela Proposição 6.2.3 sabemos que $P(A \backslash B) = P(A) - P(A \cap B)$, daí concluímos que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ■

Proposição 6.2.5. *Se $B \subset A$ então $P(B) \leq P(A)$.*

Demonstração. Como $P(A \backslash B) = P(A) - P(A \cap B)$, se $B \subset A$ resulta que $P(A \backslash B) = P(A) - P(B)$. Como $P(A \backslash B) \geq 0$ temos $P(B) \leq P(A)$. ■

A seguir vamos definir o que é probabilidade condicional e enunciar dois teoremas importantes usados nas Cadeias de Markov.

Definição 6.2.11. *Sejam A e B dois eventos de Ω, supondo $P(A) > 0$, a* ***probabilidade condicional*** *de B dado que a A acontece é definida por:*

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

. Observe que podemos reescrever a equação da definição acima da seguinte maneira: $P(A \cap B) = P(B|A) \cdot P(A)$.

Teorema 6.2.1. *Sejam $A_1, A_2, \ldots, A_n$ eventos do espaço amostral Ω, onde está definida a probabilidade P, tem-se:*

$$P(A_1 \cap A_2, \ldots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \ldots A_{n-1}) \quad (6.1)$$

Demonstração. Vamos demonstrar por indução finita. Para $n = 2$ o teorema se reduz a definicão de probabilidade condicional que é verdadeira. Suponha que a sentença seja válida para $n - 1$ eventos, ou seja, $P(A_1 \cap A_2 \ldots A_{n-1}) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \ldots P(A_{n-1}|A_1 \cap A_2 \ldots A_{n-2})$ devemos mostrar que ela vale para n eventos:
De fato, aplicando a definição 6.2.11 aos eventos $A_1 \cap A_2, \ldots A_{n-1}$ e A_n, temos: $P(A_1 \cap A_2, \ldots A_n) = P(A_1 \cap A_2 \ldots A_{n-1}) \cdot P(A_n|A_1 \cap A_2 \ldots A_{n-1})$, por fim aplicamos a hipótese de indução e chegamos a equação 6.1. ■

Teorema 6.2.2 (Teorema das Probabilidades Totais)**.** *Sejam $B_1, B_2, \ldots, B_n$ uma partição do espaço amostral Ω, isto é, esses eventos são mutuamente excludentes e sua reunião é Ω, seja A um evento e P uma probabilidade definida nos eventos de Ω, temos:*

$$P(A) = \sum_{k=1}^{n} P(A|B_k)P(B_k).$$

Demonstração. Como $\Omega = \bigcup_{k=1}^{n} B_k$ e $A = A \cap \Omega = A \cap (\bigcup_{k=1}^{n} B_k) = \bigcup_{k=1}^{n} A \cap B_k$. Calculando-se a probabilidade A obtemos: $P(A) = \sum\limits_{k=1}^{n} P(A \cap B_k)$, para finalizar bastar aplicar a definição 6.2.11 em $P(A \cap B_k)$ e chegamos a igualdade desejada. ■

Teorema 6.2.3 (Teorema de Bayes)**.** *Seja B um evento e A_1 e A_2 uma partição do espaço amostral Ω. Seja P uma probabilidade definida nos eventos de Ω. Temos para $i = 1, 2$:*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}.$$

Demonstração. Pela definição de probabilidade condicional temos:

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

Daí, basta escrever $P(A_i \cap B) = P(A_i)P(B|A_i)$ e escrever $P(B)$ utilizando o teorema das probabilidades totais 6.2.2 para concluir a demonstração. ■

Vale lembrar que o teorema de Bayes continua válido quando se considera uma partição finita do espaço amostral Ω, isto é, se $A_1, A_2, \ldots, A_n$ é uma partição de Ω e B é um evento de Ω, então, para todo i, tal que $1 \leq i \leq n$ e $n \in \mathbb{N}$, tem-se:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}$$

Para ver a demonstração dessa generalização o leitor pode consultar (DANTAS, 2013, p. 48). A seguir temos à definição de quando dois eventos de uma espaço amostral são independentes.

Definição 6.2.12. *Sejam A e B dois eventos de espaço amostral Ω e suponha $P(A) > 0$. O evento B é dito* ***independente*** *do evento A se $P(B|A) = P(B)$.*

Note que esta definição vem da ideia de que se os eventos A e B são independentes, isto significa que a probabilidade de B não se altera com a ocorrência de A. Temos ainda que se os eventos A e B são independentes então $P(A \cap B) = P(A) \cdot P(B)$. Outro fato importante sobre a independência de eventos é sua simetria, ou seja, se B não depende de A então A não depende de B, para mostrar isso basta utilizar a definição 6.2.11, vejamos:

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

Observe que podemos afirmar pelos resultados anteriores que A é independente de B se, e somente se, $P(A \cap B) = P(A) \cdot P(B)$. Para finalizar iremos definir quando n eventos são independentes, para n igual a um inteiro positivo e o que são variáveis aleatórias discretas um dos conceitos necessários para a definição das Cadeias de Markov.

Definição 6.2.13. *Os eventos $A_1, A_2, \ldots, A_n$ são* ***independentes*** *se: $P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \ldots P(A_{i_k})$ para todo $k = 2, 3, \ldots n$ e todo $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$ tal que $i_1 \leq i_2 \leq \ldots \leq i_k$.*

Definição 6.2.14. *Uma* ***variável aleatória*** *é uma função definida num espaço amostral, que assume valores reais.*

Definição 6.2.15. *As variáveis aleatórias que assumem valores em um conjunto enumerável serão denominadas* ***discretas*** *e aquelas que assumem valores num intervalo da reta real serão denominadas* ***contínuas****.*

Referências Bibliográficas

[1] ANTON, H., E RORRES, C. *Álgebra Linear com aplicações*, 8 ed. Porto Alegre, Bookman, 2001.

[2] BOLDRINI, J. L., ET AL. *Álgebra Linear*, 3 ed. São Paulo, Harper e Row do Brasil, 1980.

[3] DANTAS, C. A. B. *Probabilidade: Um curso introdutório*, 3 ed. São Paulo, Edusp, 2013.

[4] GRIMSTEAD, C. M., E SNELL, J. L. *Introduction to Probability*, 2 ed. Rhode Island, AMS-American Matemathical Society, 1997.

[5] HAYES, B. *First link in Markov Chain.* www.americanscientist.org/article/first-links-in-the-markov-chain, Acessado em 06-03-2019.

[6] HEFEZ, A., E FERNANDEZ, C. S. *Introdução à álgebra linear*, 2 ed. Rio de Janeiro, SBM - Sociedade Brasileira de Matemática, 2016.

[7] HILLIER, F. S., E LIEBERMAN, G. J. *Introdução à pesquisa operacional*, 9 ed. Porto Alegre, AMGH Editora, 2013.

[8] LAY, D. C. *Álgebra Linear e sua aplicações*, 2 ed. Rio de Janeiro, LTC, 1999.

[9] MORGADO, A. C., E CARVALHO, P. C. P. *Matemática Discreta*, 2 ed. Rio de Janeiro, SBM - Sociedade Brasileira de Matemática, 2015.

[10] O'CONNOR, J. J., E ROBERTSON, E. F. *Andrei Andreyevich Markov.* http://www-history.mcs.st-andrews.ac.uk/Biographies/Markov.html, Acessado em 24-02-2019.

[11] OGAWA, Y. *A fórmula preferida do professsor*, 1 ed. São Paulo, Estação Liberdade, 2017.

[12] POOLE, D. *Álgebra Linear*, 1 ed. São Paulo, Pioneira Thomson Learning, 2004.

[13] SULLIVAN, M., E MIZRAHI, A. *Matemática Finita*, 9 ed. Rio de Janeiro, LTC, 2006.