Universidade Federal do Esprito Santo - UFES

Centro de Ciências Exatas

Departamento de Matemática

Mestrado Profissional em Matemática em Rede Nacional - PROFMAT

TEORIA BÁSICA DAS CADEIAS DE MARKOV

Autor: Mateus Mendes Magela

Orientador: Dr. Florêncio Ferreira Guimarães Filho

13 de Outubro de 2015

Mateus Mendes Magela

TEORIA BÁSICA DAS CADEIAS DE MARKOV

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional da Sociedade Brasileira de Matemática (PROFMAT) em parceria com o Centro de Ciências Exatas da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do grau de Mestre em Matemática, sob orientação do Professor Doutor Florêncio Ferreira Guimarães Filho.

13 de Outubro de 2015

Agradecimentos

Aos meus pais, pela bravura diante de todas as dificuldades, pela dedicação que sempre tiveram na minha educação, por todo carinho e amor dedicados ao Henrique, pelos exemplos de dignidade, honestidade e sobretudo trabalho.

A minha irmã pela amizade e carinho.

A minha esposa Flávia, minha heroína, que sempre me apoiou e incentivou nas horas difíceis. Por estar sempre ao meu lado, pelo seu carinho, por sua atenção, pelo seu amor. Te amo.

Ao meu filho Henrique, pelo seu amor incondicional.

A minha nova família, Garcia, Inês e Amanda, pela confiança, por todo apoio recebido e pelo carinho e amor dedicados ao Henrique.

Ao meu orientador Prof. Dr. Florêncio Guimarães Filho por tudo que me ensinou, devo a ele a oportunidade que tive de chegar até aqui, sua dedicação e competência como professor são exemplos que levarei por toda minha vida.

A Escola São Domingos pela confiança e pelo carinho que sempre tiveram comigo, por me proporcionar conhecimento não apenas profissional, sobretudo a formação do meu caráter. Pelos grandes amigos aqui aí encontrei. Muito obrigado por tudo.

Resumo

As cadeias de Markov desempenham papel cresecente e importante na resolução de problemas em diversas áreas do conhecimento como exemplo: Administração, Biologia, Genética, Sociologia, Metereologia, Teoria de Jogos. A finalidade desse trabalho é apresentar a importância das aplicações da distribuição estacionária das cadeias de Markov. Uma revisão sobre os pré-requisitos necessários para compreensão da teoria é apresentada nos primeiros capítulos. Em seguida, é apresentada a teoria geral das cadeias de Markov introduzindo algumas de suas aplicações. O último capítulo trata com destaque o algoritmo Page Rank, uma importante aplicação das cadeias de Markov utilizada pelo Google para apresentar no topo as páginas da web mais interessantes sobre o assunto pesquisado.

Palavras chave: Probabilidade.Matriz Estocástica.Cadeias de Markov.Page Rank.

Abstract

Markov chains play a rising and important role in problems solving in several knowledge areas such as: Administration, Biology, Genetics, Meteorology and Game theory. This paper aims show the use of Markov chains stationary distribution. A review about prerequisites necessary for comprehending the theory will be presented in the first chapters. Then, we will introduce the general theory of Markov chains focusing on some of their applications. The last chapter will highlight the Page Rank algorithm, an important feature of Markov chain used by Google to place on top the most interesting web pages related to the researched topic.

Key-words: Probabilility.Stochastic Matrix.Markov chains.Page Rank.

Conteúdo

1 Introdução

Esta dissertação, de caráter elementar, pretende introduzir os principais elementos teóricos sobre as cadeias de Markov.

De maneira simples podemos descrever um sistema em cadeia de Markov discreto, como sendo um processo estocástico discreto em que o estado do sistema no instante futuro não se altera pelo seu histórico, mas apenas é afetado pelo seu estado presente. Esta caracterśtica confere as cadeias de Markov o apelido de processo desmemoriado.

A finalidade das cadeias de Markov é prever um estado futuro apenas com relação ao estado atual do sistema. Como exemplo: um matemático vai lançar uma moeda honesta quatro vezes. Qual a probabilidade de que o resultado seja cara, cara, cara, cara? A resposta é 6,25%. Agora, suponhamos que o jogador já tivesse jogado a moeda uma vez, e o resultado tenha sido cara, como pode a partir daí prever o futuro? Esse é um típico exemplo de uma cadeia de Markov.

Através das cadeias de Markov, foi possível compreender melhor fenômenos científicos e sociais em diversas áreas do conhecimento.Como exemplos podemos destacar o engenheiro da computação, que necessita construir um algoritmo de busca que classifica páginas da internet levando em consideração sua relevância na web. O sociólogo, que necessita entender como memes, ou de que forma características culturais espalham-se pela sociedade. O genteticista, que necessita construir o mapa genético de uma espécie. O meteorologista, que precisa prever cenários climáticos para uma determinada região. Ou até mesmo um biólogo, que necessecita estudar o comportamento futuro da população de uma determinada espécie marinha submetida a pesca industrial.

Essas aplicações decorrem em virtude da existência de uma distribuição de probabilidades estacionária das cadeias de Markov regulares, cujo valor é determinado por meio de uma artmética elementar envolvendo matrizes, tornando sua aplicação uma tarefa simples para os programas algébricos de computação já existentes.

As cadeias de Markov são assim denominadas em homenagem ao matemático russo Andrei A. Markov (1856 - 1922). Ele analisou a sequência de vogais e consoantes no poema Eugene Onegin (1883). Ele verificou empiricamente que uma vogal era seguida de uma consoante em 87% das vezes e que um consoante era seguida por uma vogal 66% das vezes. Este trabalho ficou conhecido como cadeias de Markov. Na época ele não imaginava que seu trabalho tivesse alguma outra aplicação.

Dentre os principais matemáticos que contribuíram para o seu avanço dessa teoria podemos destacar o matemático russo Andrei Kolmorogov.

2 Teoria das Probabilidades

2.1 Introdução

Quando o assunto é condomínio, a garagem é sempre um ponto delicado, em especial quando se trata de distribuição das vagas entre as unidades. Para minimizar as reclamações, um edifício decidiu fazer sorteio das vagas de garagem de ano em ano. O edifício possui 48 vagas sendo 36 vagas presas (ruins),ou seja, aquela que depende de seu vizinho de garagem para que se possa acessá-la e 12 vagas soltas (boas), em outras palavras vagas isoladas, aquelas em que o morador tem livre acesso. O sorteio acontece por ordem de chegada da seguinte forma: o morador que chegar primeiro, é também o primeiro a sortear uma vaga, o morador que chegar em segundo é o segundo a sortear e assim por diante. Entretanto, essa medida não foi suficiente para amenizar as reclamações, pois alguns moradores passaram a protestar que o morador que tinha condições de ser o primeiro a chegar levava grande vantagem sobre os demais . Sendo assim, surgem alguns questionamentos. O primiero morador a sortear a vaga tem maior probabilidade de tirar uma vaga boa do que o segundo no sorteio?

Existe alguma posição no sorteio que privilegia o resultado de uma vaga boa?

A Teoria das Probabilidades é o ramo da Matemática que se dedica no desenvolvimento e na busca por modelos que podem ser usados para resolver esse e demais problemas que envolvam experimentos aleatórios.

2.2 Antecedentes Históricos

O desenvolvimento dessa teoria começou a ter destaque no século XVII atráves de debates sobre problemas de jogos de azar entre Blaise Pascal, Pierre de Fermat e Antonie Gombaud, também conhecido como Chevalier de Méré. Embora tenha uma longa história de aplicação prática, apenas a partir das décadas de 20 e 30 do século XX que a teoria das probabilidades foi sistematizada com sólida fundamentação teórica, passanod a assumir com grande destaque estudos em diversas áreas como: Física, Química, Biologia, Medicina, Psicologia, Sociologia, Ciências Políticas, Educação, Economia, Engenharia dentre outras.

2.3 Conceitos Básicos

2.3.1 Experimento Aleatório

Experimentos que submetidos as mesmas condições, implicam em resultados arbitrariamente distintos, são chamados de experimentos aleatórios. Podemos dizer em outras palavras, é todo aquele experimento cujos resultados não podem ser previstos com certeza antes de sua execução como exemplo: lançamento de um dado, sorteio da Mega-Sena, lançamento de uma moeda, pesquisa de satisfação, risco econômico de crétido, temperatura máxima no verão.

2.3.2 Espaço Amostral

O conjunto de todos os resultados possíveis de um experimento é denominado espaço amostral. Cada elemento do espaço amostral é chamado de ponto amostral ou simplismente amostra.

Exemplo 2.3.2.1: O experimento aleatório dado pela obsevarção da face voltada para cima no lançamento de um dado tem como espaço amostral o conjunto A dado por:

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

2.3.3 Evento

Um evento é um conjunto de resultados possíveis de um experimento. Quando o espaço amostral é um conjunto finito, então todo subconjunto é um evento. O mesmo não acontece quando o espaço amostral é um conjunto infinito.

Exemplo 2.3.3.1: No lançamento de um dado, podemos considerar como evento E a observação de um número par na face voltada para cima.

$$E = \{2, 4, 6\}$$

É importante observarmos que podemos combinar eventos para obter novos eventos usando as operações elementares dos conjuntos.

Exemplos:

a) $A \cup B$ é o evento que ocorre se, e somente se, ocorre apenas A, ocorre apenas B ou ocorre ambos;

b) $A \cap B$ é o evento que ocorre se, e somente se, ocorrem A e B simultâneamente.

Definição: Dois eventos A e B são chamados de mutuamente exclusivos se eles são disjuntos, ou seja, se $A \cap B = \oslash$. Podemos dizer em outras palavras, A e B são multuamente exclusivos se, e somente se, não podem ocorrer simultâneamente.

Exemplo 2.3.3.2: Seja A o evento em que o resultado no lançamento de um dado seja um número par e B o evento em que o resultado no lançamento de um dado seja um número ímpar. Então temos:

$$A = \{2, 4, 6\}\,; B = \{1, 3, 5\}$$

Os eventos A e B são multuamente exclusivos uma vez que não é possível que um número seja par e ímpar simultâneamente.

2.3.4 Definição Cássica de Probabilidades

Consideremos o evento E $= \{2, 4, 6\}$. Se lançarmos um dado honesto um grande número de vezes, é intuitivo esperar que um número par ocorrerá em aproximadamente metade das observações. Isso decorre dos seguintes fatos:

1) os eventos elementares equiprováveis (elementos favoráveis) são aqueles igualmente prováveis;

2) o número de elementos de E é justamente a metade dos elementos do espaço amostral(todos os elementos possíveis).

Estas condições provocam a seguinte definição de probabilidade.

$$\boxed{\textbf{Probabilidade} = \frac{\textbf{número de resultados favoráveis}}{\textbf{número de resultados possíveis}}}$$

Dessa forma se um evento E ocorre de a maneiras dentre um total de b maneiras igualmente prováveis, então a probabilidade (chance) P(E) do evento E ocorrer é dada por

$$\boxed{\mathrm{P(E)} = \frac{a}{b}}$$

2.4 Espaço Equiprovável e Probabilidade

Seja S um espaço amostral finito tal que $S = \{s_1, s_2, \cdots, s_n\}$. Um espaço finito de probabilidades é obtido associando-se a cada elemento $s_j \in S$ um número real não negativo p_j, denominado a probabilidade de s_j, tal que $p_1 + p_2 + \cdots + p_n = 1$:
Indicamos por P(E) a probabilidade de um evento E, definida como sendo a soma das probabilidades dos elementos de E.

Exemplo 2.4.1. Consideremos o lançamento simultâneo de 3 moedas e observemos o número de caras. O espaço amostral desse experimento aleatório é dado pelo conjunto

$$A = \{0, 1, 2, 3\}$$

Obtemos um espaço de probabilidades pela seguinte associação:

$$P(0) = \tfrac{1}{8},\ P(1) = \tfrac{3}{8},\ P(2) = \tfrac{3}{8},\ P(3) = \tfrac{1}{8}$$

em que P(0) é a probabilidade de sair nenhuma cara, P(1) é a probabilidade de sair 1 cara, P(2) é a probabilidade de sair 2 caras e P(3) é a probabilidade de sair 3 caras. Note que cada probabilidade é não negativa e a soma é 1.
Agora seja A o evento em que ocorre pelo menos 1 cara e seja B o evento em que ocorrem apenas caras ou apenas coroas.

Então pela definição,

$$P(A) = P(1) + P(2) + P(3) = \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8}$$

$$P(B) = P(0) + P(3) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

Definição: Um espaço finito de probabilidades S, em que cada ponto amostral tem a mesma probabilidade, é denominado um espaço equiprovável. Em particular, se S contém n elementos, então a probabilidade de cada ponto amostral é $\frac{1}{n}$. Além disso, se um evento E contém m elementos, então sua probabilidade é dada por P(E) $= \frac{m}{n}$. Em outras palavras,

$$\boxed{\mathrm{P(E)} = \frac{\textbf{número de elemento de A}}{\textbf{número de elementos de S}}}$$

É importante ressaltar que, a equiprobabilidade é uma característica do modelo de probabilidade e não do espaço amostral.

Exemplo 2.4.2: A Mega Sena é uma das loterias mais populares do Brasil. Organizada pela Caixa Econômica Federal, é a loteria que mais premia no país, principalmente quando acumula o prêmio por várias semanas. Qual é a chance real de alguém acertar os seis números?

Primeiro passo é verificar o total de agrupamentos possíveis, 6 a 6, dos 60 números. Eles podem ser calculados pelo princípio multiplicativo da contagem por meio do produto 60 x 59 x 58 x 57 x 56 x 55 = 36.045.979.200.

Mas, um cartão é contado 720 vezes, o total de possibilidades de trocas na ordem desses 6 números e, portanto, para achar o número de cartões distintos é preciso dividir o resultado por 720. Ao fazermos a conta encontramos 50.063.860. Logo, a probabilidade de um apostador ganhar na Mega-Sena com apenas um cartão de seis números é

$$\frac{1}{50.063.860} \cong 0,000001997\%$$

É mais provável jogar uma moeda 25 vezes seguidas e conseguir sempre o mesmo resultado do que ganhar o prêmio da loteria.

$$\textbf{Total de combinações dos lançamentos} = \underbrace{2 \cdot 2 \cdot \cdots \cdot 2}_{25 vezes} = 2^{25} = 33.554.432.$$

Logo, a probabilidade de que ocorre um mesmo resultado em todos os 25 lançamentos consecutivos de uma moeda é

$$\frac{1}{33.554.432} \cong 0,000002980\%$$

2.5 Probabilidade Condicional

Seja E um evento arbitrário num espaço amostral S com $P(E) > 0$. A probabilidade de que um evento A ocorra na certeza da ocorrência de um evento E inidcada por $P(A|E)$ é definida por:

$$\boxed{\mathrm{P}(A|E) = \frac{P(A \cap E)}{P(E)}}$$

Note-se que $P(A|E)$ somente está definida quando $P(E)$ >0.

Se $P(E) = 0$ a probabilidade condicional $P(A|E)$ pode ser definida arbitrariamente. Convenientemente quando $P(E) = 0$ definimos $P(A|E) = P(A)$. Sendo assim, sempre que $P(E)$ >0 podemos escrever

$$\boxed{\mathrm{P}(\mathrm{A} \cap E) = P(E) \cdot P(A|E)}$$

Podemos observar no diagrama a seguir que $P(A\,|E)$ mede, num certo sentido, a probabilidade relativa de A com relação ao espaço reduzido E.

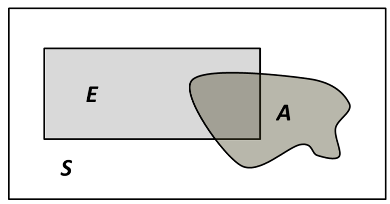

Figura 2.1: Diagrama

Exemplo 2.5.1: Um experimento foi conduzido com o objetivo de avaliar o poder germinativo de duas culturas de cebola, conforme mostra a tabela abaixo.

Desejando-se fazer uma avaliação do poder germinativo de uma das culturas de cebola, uma amostra foi retirada ao acaso. Sabendo-se que a amostra escolhida germinou, qual a probabilidade de essa amostra pertencer à Cultura A?

Germinação de sementes de duas culturas de cebola

Culturas	Germinação		TOTAL
	Germinaram	Não Germinaram	
A	392	8	400
B	381	19	400
TOTAL	773	27	800

Figura 2.2: BUSSAB, W. O; MORETIN, L. G. Estatística para as ciências agrárias e biológicas (adaptado).

Queremos encontrar a probabilidade da amostra retirada pertencer a cultura A considerando como espaço amostral reduzido somente as sementes que germinaram. O experimento foi realizado com um total de 800 sementes dentre as quais 773 germinaram. Das 773 sementes que germinaram, 392 pertencem a cultura A. Sendo assim, a probabilidade condicional de retirar uma semente da cultura A sabendo que a semente germinou é dada por

$$\frac{392}{773} \cong 50\%$$

Exemplo 2.5.2: Flávia quer enviar uma carta a Amanda. A probabilidade de Flávia escreva a carta é de $\frac{8}{10}$. A probabilidade de que o correio não a perca é de $\frac{9}{10}$. A proabilidade de que o carteiro a entregue é de $\frac{9}{10}$. Dado que Amanda não recebeu a carta, qual a probabilidade condicional de que Flávia não tenha escrito a carta?

$$P(\textbf{não escreve/não recebe}) = \frac{P(\textbf{não escreve})}{P(\textbf{não recebe})} = \frac{2/10}{\frac{2}{10} + \frac{8}{10}.\frac{1}{10} + \frac{8}{10}.\frac{9}{10}.\frac{1}{10}} = \frac{25}{44}.$$

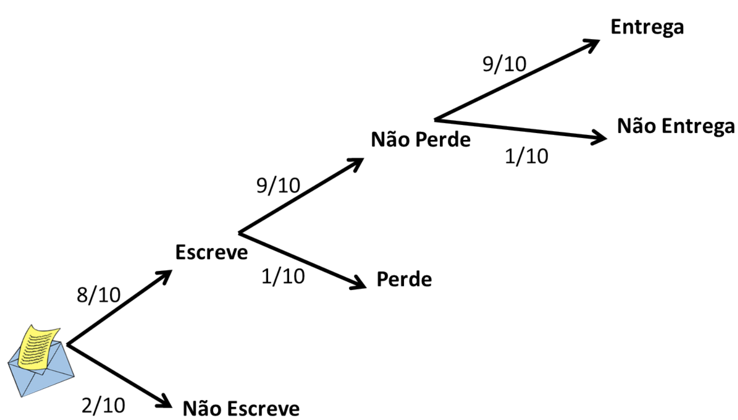

Figura 2.3: Diagrama de árvore - problema da carta

Teorema da Probabilidade Composta: Sejam A e B dois conjuntos quaisquer tais que $A \cap B$ é não vazio. Então temos:

1) $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$;

2) Se $P(A_1 \cap A_2 \cap \cdots \cap A_n) \neq 0$, então $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$.

Demonstração: 1) Sabemos que

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

Daí segue que,

$$P(A \cap B) = P(B)P(A|B);$$
$$P(B \cap A) = P(A)P(B|A);$$

Como $P(A \cap A) = P(B \cap A)$ temos que

$$P(B)P(A|B) = P(A)P(B|A)$$

Como queríamos mostrar.

2)Para dois conjuntos A_1 e A_2 a fórmula é verdadeira, pois já vimos que $P(A \cap B) = P(A)P(B/A)$. Agora, vejamos a sentença para três conjuntos A_1, A_2 e A_3.

$$P(A_1 \cap A_2 \cap A_3) = P(A_3|(A_1 \cap A_2))P(A_1 \cap A_2) = P(A_3|(A_1 \cap A_2))P(A_2|A_1)P(A_1),$$

como queríamos mostrar. No caso geral a demonstração é análoga e segue o Princípio da Indução Completa.

Exemplo 2.5.3: Em um lote de 12 peças, das quais 4 são defeituosas, três peças são escolhidas ao acaso, uma após a outra. Qual a probabilidade de que todas as três peças retiradas sejam não defeituosas.

Solução: A probabilidade de que a primeira peça seja não defeituosa é $\frac{8}{12}$, uma vez que 8 das 12 peças são não defeituosas. Se a primeira peça é não defeituosa, então a probabilidade de que a segunda peça escolhida seja não defeituosa é $\frac{7}{11}$, pois 7 das 11 peças restantes são não defeituosas. Se as duas primeiras peças escolhidas são não defeituosas, então a probabilidade da terceira peça escolhida ser não defeituosa é $\frac{6}{10}$. Assim, pelo teorema do produdo segue que a probabilidade de se retirar três peças não defeituosas é dada por

$$\frac{8}{12} \cdot \frac{7}{11} \cdot \frac{6}{10} = \frac{14}{55}$$

2.6 Processos Estocásticos Finitos

2.6.1 Introdução

Uma sequência finita de experimentos, na qual cada experimento tem um número finito de ocorrências com probabilidades conhecidas, é chamado de um processo estocástico finito. Um método conveniente para descrever um processo estocástico finito e calcular a probabilidade de qualquer evento é através de um diagrama de árvore.

Exemplo 2.6.1.1: Escolhemos aleatoriamente uma dentre duas moedas disponíveis A e B. Sabemos que a moeda A é honesta e que a moeda B possui cara (C) em ambos os lados. A moeda escolhida é lançada. Se o resultado for coroa (K), joga-se um dado honesto, caso contrário joga-se a moeda novamente. A primeira etapa do experimento é escolher uma moeda. A segunda etapa é lançar a moeda escolhida e a terceira etapa é lançar uma moeda ou um dado, dependendo dos resultados obtidos nas duas primeiras etapas. Indicamos todos os possíveis resultados do experimento pela árvore de probabilidade abaixo:

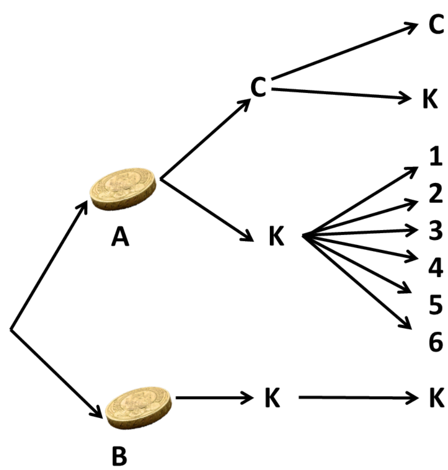

Figura 2.4: Diagrama de árvore

Os possíveis resultados para o experimento são:

$$
\begin{gathered}
(A, C, C)\,, (A, C, K)\,, (A, K, 1)\,, \\
(A, K, 2)\,, (A, K, 3)\,, (A, K, 4)\,, \\
(A, K, 5)\,, (A, K, 6)\,, (B, K, K)\,.
\end{gathered}
$$

Cada resultado pode ser identificado como um caminho através da árvore. Cada caminho é composto por segmentos chamados de ramos. Nesta árvore, há nove caminhos com três ramos cada. o processo descrito acima, pode ser realizado para qualquer experiência que ocorre em etapas. Exigimos apenas que haja num número finito de resultados em cada etapa e que saibamos as probabilidades para qualquer resultado na enésima etapa dado o conhecimento do resultado da $(n-1)$-ésima etapa. Para cada n obtemos uma árvore T_n. A probabilidade de que um caminho particular da árvore ocorra é dado pelo teorema da multiplicação como sendo o produto das probabilidades de cada ramo do caminho. Por exemplo, a probabilidade de se escolher a moeda A e a partir daí obter o número 3 no lançamento de dado honesto é dado pelo produto:

$$\frac{1}{2}.\frac{1}{2}.\frac{1}{6} = \frac{1}{24}$$

2.6.2 Independência

Os eventos aleatórios A e B são estocasticamente independentes se

$$\boxed{\mathrm{P(A}\cap B) = P(A) \cdot P(B)}$$

Os eventos de probabilidade 0 ou 1 são independentes de qualquer outro.

Se $P(A) = 0$, então $P(A \cap B) = 0$ e A e B são independentes.

Se $P(B) = 1$, então $P(A \cap B) = P(A) - P(A \cap B^c) = 0$ e $P(A \cap B) = P(A)P(B)$

Logo A e B são independentes.

Exemplo 2.6.2.1: Uma moeda é lançada duas vezes. Obtemos nesse caso o espaço amostral equiprovável S dado por

$$S = \{CC, CK, KC, KK\}$$

em que C representa cara e K representa coroa.

Consideremos o evento A em que cara ocorra no primeiro lançamento e B o evento em que cara ocorra no segundo lançamento.

$$P(A) = \frac{2}{4} = \frac{1}{2} \qquad\qquad P(B) = \frac{2}{4} = \frac{1}{2}$$

Por outro lado, a probabilidade de que cara ocorre no primeiro e no segundo lançamento é dada por

$$P(A \cap B) = \frac{1}{4}$$

Portanto, P(A$\cap B) = P(A) \cdot P(B)$

Consequentemente, A e B são eventos independentes.

2.6.3 Problema Motivador

Na introdução deste capítulo iniciamos o texto com um problema motivador. O problema em questão é denominado Problema das Garagens. Esse problema foi vivenciado pelo professor Florêncio Guimarães Filho em um antigo prédio em que morava, e desde que ele apresentou esse problema em uma aula do PIC (Programa de Iniciação Científica da OBMEP) na qual eu participava como seu monitor, sempre o apresento em minhas aulas sobre probabilidade como um problema motivador, afim de despertar o interesse e a partición dos alunos. Fato interessante é que, alguns alunos acabam se identificam com o problema, relatando que já passaram por uma situação semelhante. Segue a seguir uma solução para esse problema.

Solução: A probabilidade de um morador retirar um vaga boa (B) sendo o primeiro a sortear é de $\frac{12}{48} = \frac{1}{4} = 25\%$.

Agora, sendo o segundo podem ocorrer dois casos: BB(boa,boa) ou RB(ruim,boa).

$$P(BB) = \frac{12}{48} \cdot \frac{11}{47} \qquad\qquad P(RB) = \frac{36}{48} \cdot \frac{12}{47}$$

Sendo assim, a probabilidade do morador sortear uma vaga sendo o segundo na ordem é de

$$P(BB) + P(RB) = \frac{12}{48} \cdot \frac{11}{47} + \frac{36}{48} \cdot \frac{12}{47} = \frac{12}{48}\left(\frac{11}{47} + \frac{36}{47}\right) = \frac{12}{48} \cdot \frac{47}{47} = \frac{12}{48} = \frac{1}{4}$$

Portanto a probabilidade do morador sortear uma vaga boa sendo o primeiro ou segundo da lista é a mesma.

Suponhamos agora que o morador seja o terceiro da lista, então podem ocorrer quatro casos favoráveis: BBB,BRB,RRB,RBB.

Logo a probabilidade de uma vaga boa ser sorteada nessas condições é dada por

$$P(BBB)+P(BRB)+P(RRB)+P(RBB) = \frac{12}{48}\cdot\frac{11}{47}\cdot\frac{10}{46}+\frac{12}{48}\cdot\frac{36}{47}\cdot\frac{11}{46}+\frac{36}{48}\cdot\frac{35}{47}\cdot\frac{12}{46}+\frac{36}{48}\cdot\frac{12}{47}\cdot\frac{11}{46}$$

$$= \frac{12}{48}\left(\frac{11}{47} \cdot \frac{10}{46} + \frac{36}{47} \cdot \frac{11}{46} + \frac{35}{46} \cdot \frac{36}{47} + \frac{36}{46} \cdot \frac{11}{47}\right)$$

$$= \frac{12}{48}\left(\frac{11.10 + 36.11 + 35.36 + 36.11}{46.47}\right)$$

$$= \frac{12}{48} \cdot \frac{2162}{2162} = \frac{1}{4}$$

Então, a probabilidade de uma vaga boa sair em primeiro, segundo, terceiro.

Agora supomos que em um determinado momento do sorteiro temos b vagas boas e r vagas ruins. A probabilidade de um morador sortear uma vaga boa nessa situação é dada por

$$\frac{b}{b+r}.$$

E a probabilidade do próximo morador sortear uma vaga boa é dada por

$$\frac{b}{b+r} \cdot \frac{b-1}{b+r-1} + \frac{r}{b+r} \cdot \frac{b}{b+r-1} = \frac{b(b+r-1)}{(b+r)(b+r-1)} = \frac{b}{b+r}$$

Portanto, não existem motivos para que os protestos dos moradores quanto a ordem do sorteio.

Uma outra forma de esclarecer esse problema é pensar no resultado do sorteio como sendo uma lista em que cada sorteio resume-se a uma lista de vagas boas (B) ou ruins (R). Dessa forma queremos saber se existe uma posição mais privilegiada na lista final.

Note que, o total de listas distintas é dada por

$$\frac{48!}{12! \cdot 36!}$$

Fixando uma vaga boa arbitrariamente em qualquer posição na lista, o número de listas distintas dessa forma é dado por

$$\frac{47!}{11! \cdot 36!}$$

Sendo assim, a probabilidade de que uma vaga boa aparecer em uma posição arbitraria na lista é

$$\frac{\frac{47!}{11! \cdot 36!}}{\frac{48!}{12! \cdot 36!}} = \frac{12}{48} = \frac{1}{4}$$

Portanto, a probabilidade de uma vaga boa aparecer em qualquer posição na lista é mesma, não havendo dessa forma uma posição privilegiada na ordem do sorteio.

3 Matrizes

3.1 Antecedentes Históricos

Por muitos séculos, matemáticos trabalharam na elaboração e no entendimento de problemas envolvendo blocos numéricos como exemplo os quadrados mágicos. No quadrado mágico mais simples, deve-se arranjar os números 1,2,3,4,5,6,7,8 e 9 usando-os uma única vez, numa tabela 3 x 3, de modo que, os números em cada linha, coluna ou diagonal somem 15.

2	9	4
7	5	3
6	1	8

Figura 3.1: Quadrado Mágico 3x3

Segundo uma lenda, o primerio quadrado mágico surgiu há 4000 anos, inscrito sobre um casco de tartaruga que se arrastava saindo do rio Lo, na China. O rio havia provocado uma grande enchente, e o imperador Yu ordenou que se fizesse sacrifícios para agradar o deus do rio. Em contrapartida, o deus enviou uma tartaruga, cujo padrão numérico inscrito no seu casco, destinava-se a ajudar o imperador a controlar o rio. Uma vez descoberto o arranjo numérico, os matemáticos chineses começaram a construir quadrados maiores que funcionassem da mesma forma. Esses quadrados eram reconhecidos como detentores de grandes propriedades mágicas. Há evidências de que os quadrados mágicos foram levados para Índia por mercadores chineses que lidavam não somente com especiarias, mas também com ideias matemáticas. Os quadrados mágicos também foram populares na cultura islâmica medieval. Uma das primeiras exibições de quadrados mágicos na Europa é o quadrado que aparece na gravura *Melancholia*, de Albrecht Durer.

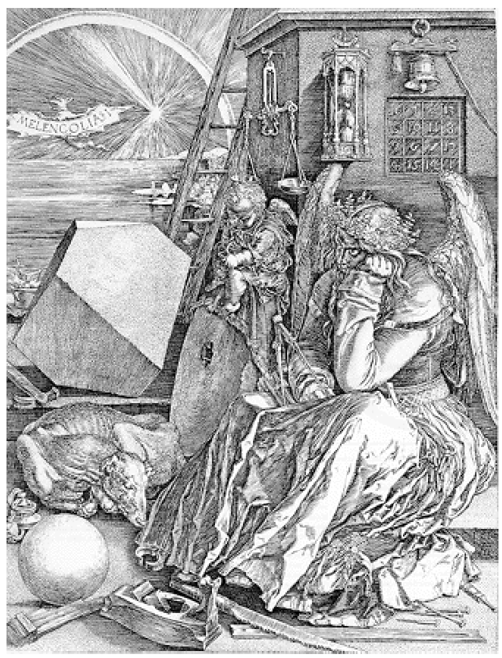

Figura 3.2: Melancholia, de Albercht Durer

Nesse quadrado, os números de 1 a 16 estão dispostos em linhas, colunas e diagonais que somam 34. Durer ainda dispôs os dois números centrais da última linha para registrar o ano da gravura: 1514.

Quadrados mágicos de diferentes tamanhos eram tradicionalmente associados a planetas do sistema solar. O clássico 3 x 3 era associado a Saturno, o quadrado 4 x

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

Figura 3.3: Quadrado Mágico de Durer

4, a Júpter, enquanto o 9 x 9, era atribuído à Lua. Uma explicação para o uso que Durer fez do quadrado é que este refletia a crença mística de que o espírito alegre de Júpter podia se contrapor ao senso de melancolia presente na gravura. Entretanto, a ideia de tratar blocos numéricos como um único número decolou apenas há 150 anos com um pequeno grupo de matemáticos que reconheceram seu potencial. Doravante, o progresso de uma álgebra unidimensional para uma álgebra de múltipla dimensões demonstrou ser incrivelvente potente para aplicações sofisticadas.

Augustin-Louis Cauchy (1789 - 1857), matemático francês, foi o primeiro a nomear a essas configurações numéricas de tableau, que em francês, significa tabela em 1826, mas o nome matriz surgiu em 1850 com James Joseph Sylvester. Entretanto, foi Arthur Cayley em 1858, na sua publicação de *Memoir on the Theory of Matrices*, quem formulou a primeira definição abstrata de matriz. Cayley também forneceu uma álgebra matricial, definindo adição, multiplicação de matrizes, multiplicação por escalar e matriz invertível. Foi a partir dos trabalhos de Cayley que as matrizes passaram a ter relevância e gradativamente foram superando os determinantes em grau de importância.

As matrizes estão envolvidas em diversas áreas do conhecimento e em diversas atividades humanas, como por exemplo: nos jogos eletrônicos, os programas de informática, a aplicação de bancos de dados e internet, nos sistemas de rede elétrica e de transportes, são apenas alguns exemplos em que a aplicação de matrizes é fundamental.

Neste capítulo apresentaremos as definições básicas da teoria de matrizes necessárias no desenvolvimento e na ánalise do estudos das cadeias de Markov.

3.2 Introdução a Teoria das Matrizes

Frequentemente encontramos em revistas, jornais ou documentos, informações nu-méricas organizadas na forma de tabelas retangulares dispostas em linhas e colunas. Estas tabelas de dados numéricos são simplificadas, utilizando uma forma denominada matriz. As matrizes estão presentes em várias atividades científicas dentre as quais podemos destacar: análise de redes elétricas, programação linear geométricas, cadeias de Markov, jogos de estratégia, modelos econômicos, administração de florestas, computação gráfica, criptografia dentre.

Exemplo 3.2.1. A tabela a seguir é parte da análise das condições de vida da população brasileira em 2014 realizada pelo IBGE.

Os dados são referentes ao número de pessoas de 16 anos ou mais ocupadas em alguma atividade economica nos estados de Alagoas e Espírito Santo.

	HOMENS	MULHERES
AL	757000	494000
ES	1076000	783000

Tabela 3.1: IBGE, Pesquisa Nacional por Amostra de Domicílios 2013

Podemos representar a tabela acima na forma de uma matriz M dada por

$$\mathbf{M} = \begin{pmatrix} 757000 & 494000 \\ 1076000 & 783000 \end{pmatrix}$$

Dessa forma defimos matriz como sendo qualquer arranjo retangular de informações numéricas organizadas em linhas e colunas.

As linhas são contadas de cima para baixo. A primeira e a segunda linha da matriz M são respectivamente

$$(757000 \quad 494000) \quad ; \quad (1076000 \quad 783000).$$

As colunas são contadas da esquerda para direita. A primeira e a segunda coluna da matriz M são respectivamente

$$\begin{pmatrix} 757000 \\ 1076000 \end{pmatrix} ; \begin{pmatrix} 494000 \\ 783000 \end{pmatrix}.$$

O elemento da matriz localizado na linha i e na coluna j é denotado por a_{ij}. Então, na matriz M temos: $a_{11} = 757000$, $a_{12} = 494000$, $a_{21} = 1076000$ e $a_{22} = 783000$.

Exemplo 3.2.2. Um grupo de alunos dos cursos A, B e C solicita transferência para outro curso, escolhido entre os mesmos A, B e C. A matriz abaixo, representa o resultado obtido após o processo de transferência, admitindo que cada aluno pode se matricular em apenas um curso

$$\mathbf{T} = \begin{pmatrix} 137 & 7 & 8 \\ 12 & 115 & 13 \\ 14 & 15 & 119 \end{pmatrix}$$

em que a linha 1 representa a quantidade de alunos oriundos do curso A dentre os quais 137 pernaneceram, 7 se transferiram para o curso B e 8 se transferiram para o curso C. A linha 2 representa a quantidade de alunos oriundos do curso B dentre os quais 12 alunos se transferiram para o curso A, 115 permaneceram no curso B e 13 se transferiram para o curso C. A linha 3 representa a quantidade de alunos procedentes do curso C dentre os quais 14 se transferiram para o curso A, 15 se transferiram para o curso B e 119 permaneceram matriculados no curso C.

Podemos dizer também por exemplo que a coluna 1 representa a quantidade de alunos matriculados no curso A após encerrado o período de transferências dentre os quais, 137 já frequentavam o curso A, 12 estudavam no curso B e 14 estudavam no curso C. A segunda coluna representa a quantidade de alunos matriculados no curso B após encerrado o período de transferências dentre os quais 7 estudavam no curso A, 115 permaneceram no curso B e 15 são oriundos do curso C. A terceira linha representa a quantidade de alunos matriculados no curso C após encerrado o período de transferências dentre os quais 8 estudavam no curso A, 13 estudavam no curso B e 119 permaneceram matriculados no curso C.

Analisando a matriz T, qual a quantidade total do número de alunos transferidos?

$$a_{12} + a_{13} + a_{21} + a_{23} + a_{31} + a_{32} = 69.$$

Portanto, 69 alunos transferiram de curso

Qual a quantidade de alunos que estavam matriculados no curso A antes do processo de transferência?

Para isso, basta somarmos o número de alunos que permaneceram no curso A com aqueles que se transferiram do curso A para o curso B ou C. Dessa forma fazemos,

$$a_{11} + a_{12} + a_{13} = 147.$$

Portando, haviam 147 estudantes matriculados no curso A.

Exemplo 3.2.3. Cada elemento a_{ij} da matriz M indica o tempo, em minuto, que um semáforo fica aberto num período de 2 minutos, para que haja o fluxo de veículos da rua i para a rua j, considerando que cada rua tenha mão dupla.

$$\mathbf{A} = \begin{pmatrix} 0 & 1.5 & 0.5 \\ 1.5 & 0 & 1 \\ 0.5 & 1 & 0 \end{pmatrix}$$

Com base nos dados, e admitindo que é possível até 20 carros passarem por minuto cada vez que o semáforo se abre, em um período de 2 horas, qual o número máximo de veículos podem passar da rua 3 para a rua 1?

Solução: Segue da matriz M que o semáforo que permite aos veículos passarem da rua 3 para a rua 1 fica aberto por 0.5 minutos a cada 2 minutos, isto é, a cada uma hora este semáforo fica aberto por 30 x 0.5 = 15 minutos. Agora considerando o fluxo máximo de carros, temos 20 x 15 = 300 carros passando da rua 3 para a rua 1 em uma hora. Portanto, em duas horas, podem passar da rua 3 para a rua 1, 600 carros.

Definição: Uma matriz é uma tabela retangular com m linhas e n colunas cujos os elementos a_{ij} representam elementos com dupla entrada, ou seja, com índices duplos em que 1 $\leq$ i$\leq$ m e 1 $\leq$ j$\leq$ n sendo m,n inteiros positivos.

Representamos uma matriz $A = (a_{ij})$ de ordem m x n, na qual o elemento a_{ij} encontra-se na interseção da i-ésima linha com a j-ésima coluna da seguinte forma:

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

Então, podemos escrever que $A = (a_{ij})_{mxn}$

Exemplo 3.2.4: Considere a matriz $A = (a_{ij})_{mxn}$, tal que os elemntos a_{ij} são dados por $a_{ij} = 3i - 2j$ para resolver os itens abaixo:

a) represente a matriz A tal que m = 3 e n = 2;

b) determine a soma de todos os elemntos da segunda coluna da matriz A.

Solução: a) Obedecendo a lei de formação dos elementos de A temos que:

$$\begin{array}{ll} \mathrm{a}_{11} = 3 \cdot 1 - 2 \cdot 1 = 3 - 2 = 1 & \mathrm{a}_{12} = 3 \cdot 1 - 2 \cdot 2 = 3 - 4 = -1 \\ \mathrm{a}_{21} = 3 \cdot 2 - 2 \cdot 1 = 6 - 2 = 4 & \mathrm{a}_{22} = 3 \cdot 2 - 2 \cdot 2 = 6 - 4 = 2 \\ \mathrm{a}_{31} = 3 \cdot 3 - 2 \cdot 1 = 9 - 2 = 7 & \mathrm{a}_{32} = 3 \cdot 3 - 2 \cdot 2 = 9 - 4 = 5 \end{array}$$

Logo, a matriz A é representada por:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 4 & 2 \\ 7 & 5 \end{pmatrix}$$

Solução: b) A soma dos elementos da segunda coluna é -1 + 2 + 5 = 6.

Exemplo 3.2.5. Uma fábrica produz três modelos de carros utilizando diferentes peças para a montagem do motor. Considerando a matriz abaixo

$$\mathbf{A} = \begin{pmatrix} 15 & 10 & 12 \\ 10 & 11 & 13 \\ 14 & 12 & 11 \end{pmatrix}$$

em que cada elemento a_{ij} representa a quantidade de peças do tipo j utilizada na fabricação de um carro do modelo i.

Quantas peças serão utilizadas para fabricar o modelo 2 de carro?

Solução: Basta somarmos os elementos da segunda linha, $10 + 11 + 13 = 34$. Logo, serão utilizadas 34 peças.

3.3 Álgebra Matricial

3.3.1 Adição de Matrizes

Uma vantagem fundamental da álgebra de matricial é que podemos considerar uma grande quantidade de informações numéricas, como sendo, apenas um objeto.

Exemplo 3.3.1.1: Observe a tabela abaixo:

	produto 1	**produto 2**	**produto 3**	**produto 4**
fábrica 1	7	5	0	1
fábrica 2	0	4	3	7
fábrica 3	3	2	0	2

Tabela 3.2: Tabela de produção de Janeiro 2015

Suponhamos que essa tabela represente a produção da empresa Alfa no mês de janeiro de 2015, e que essa empresa, tenha três fábricas situadas em distintas parte do país e a produção de cada um de seus quatro produtos, é medida em milhões de unidades. Assim, podemos representar sua produção através de uma matriz A, de modo que, cada elemento a_{ij} representa a produção em milhões de unidades da fábrica i e do produto j.

$$\mathbf{A} = \begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix}$$

Agora, suponhamos que no mês seguinte, a produção dessa empresa seja demonstrada pela tabela a seguir.

	produto 1	**produto 2**	**produto 3**	**produto 4**
fábrica1	9	4	1	0
fábrica2	0	5	1	8
fábrica3	4	1	1	0

Tabela 3.3: Tabela de produção de Fevereiro 2015

Sendo assim, podemos representar os dados de produção do mês de fevereiro de 2015 através da matriz B representada a seguir:

$$\mathbf{B} = \begin{pmatrix} 9 & 4 & 1 & 0 \\ 0 & 5 & 1 & 8 \\ 4 & 1 & 1 & 0 \end{pmatrix}$$

Para apresentar os resultados da produção total do primeiro bimestre de 2015 da empresa Alfa, é necessário somar as produções dos meses de janeiro e fevereiro de

2015, fazemos isso somando os elementos correspondentes nos dois quadros.

Essa situação representa uma adição de matrizes, conforme demonstrada a seguir:

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix}+\begin{pmatrix} 9 & 4 & 1 & 0 \\ 0 & 5 & 1 & 8 \\ 4 & 1 & 1 & 0 \end{pmatrix}=\begin{pmatrix} 7+9 & 5+4 & 0+1 & 1+0 \\ 0+0 & 4+5 & 3+1 & 7+8 \\ 3+4 & 2+1 & 0+1 & 2+0 \end{pmatrix}$$

Portanto, a soma das matrizes A e B é dada por

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix} 16 & 9 & 1 & 1 \\ 0 & 9 & 4 & 15 \\ 7 & 3 & 1 & 2 \end{pmatrix}$$

Exemplo 3.3.1.2. Uma editora pretende publicar uma coleção de livros de Álgebra e Geometria em duas versões: volumes 1,2 e 3 e uma outra volume único. A tabela I mostra a quantidade de cada volume a ser lançado no primeiro semestre do ano e a tabela II mostra a quantidade de cada volume a ser lançado no segundo semestre do ano.

Volume	**Álgebra**	**Geometria**
1	200	250
2	220	230
3	260	240
Único	300	310

Tabela 3.4: Tabela I: Quantidades de exemplares no primeiro semestre (em milhares de unidades)

Volume	**Álgebra**	**Geometria**
1	170	150
2	120	140
3	120	140
Único	350	410

Tabela 3.5: Tabela II: Quantidades de exemplares no segundo semestre (em milhares de unidades)

Quantos livros de Álgebra e Geometria de cada volume serão lançados por essa editora nesse ano?

Solução: Para determinarmos o total de livros lançados é preciso somar os elementos correspondentes nas duas tabelas.

Volume	**Álgebra**	**Geometria**
1	370	400
2	340	370
3	380	380
Único	650	720

Tabela 3.6: Tabela III: Quantidades de exemplares no ano (em milhares de unidades)

Esse problema pode ser resolvido através da soma de duas matrizes, conforme mostrado a seguir:

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix} 200 & 250 \\ 220 & 230 \\ 260 & 240 \\ 300 & 310 \end{pmatrix}+\begin{pmatrix} 170 & 150 \\ 120 & 140 \\ 120 & 140 \\ 350 & 410 \end{pmatrix}=\begin{pmatrix} 370 & 400 \\ 340 & 370 \\ 380 & 380 \\ 650 & 720 \end{pmatrix}$$

Note que a soma de matrizes segue uma soma algébrica convencional entre os elementos correspondentes (de mesmo índice duplo). Portanto, podemos somar duas matrizes somente quando elas forem de mesma ordem, isto é, se ambas tiverem o mesmo número de linhas e colunas.

Definição: Definimos a soma de duas matrizes, como sendo a soma entre os seus elementos correspondentes, isto é, de mesma posição.

De modo geral, dadas as matrizes $A = (a_{ij})$ e $B = (b_{ij})$, definimos a soma $A + B = (c_{ij})$ como sendo:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}+\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}$$

tal que, $c_{ij} = a_{ij} + b_{ij}, \forall\ i, j$

Sendo assim, à adição de matrizes segue as mesmas propriedades da adição de números reais: comutativa, associativa, elemento neutro da adição e elemento oposto.

3.3.2 Propriedades da Adição de Matrizes

Sejam A, B e C matrizes de mesma ordem, então valem as seguintes propriedades:

(1) **Comutativa:** $\boxed{\text{A + B = B + A}}$

Demonstração: Sejam X e Y matrizes tais que X = A + B e Y = B + A. Observando que $x_{ij} = a_{ij} + b_{ij}$ e que $y_{ij} = b_{ij} + a_{ij}$. Como a_{ij} e $b_{ij} \in \mathbb{R}$, então vale a propriedade comutativa dos números reais, $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ implicando que $x_{ij} = y_{ij}$, logo X = Y. Portanto, A + B = B + A, como queríamos mostrar.

(2) **Associativa:** $\boxed{\text{A + (B + C) = (A + B) + C}}$

Demonstração: Sejam X e Y matrizes tais que X = A + (B + C) e Y = (A + B) + C. Observando que $x_{ij} = a_{ij} + (b_{ij} + c_{ij})$ e $y_{ij} = (a_{ij} + b_{ij}) + c_{ij}$. Como a_{ij}, b_{ij} e c_{ij} são números reais, vale a propriedade associativa dos números reais, então $x_{ij} = y_{ij}$ logo, $X = Y$. Portanto, A + (B + C) = (A + B) + C, como queríamos mostrar.

(3) **Existência do elemento simétrico:** $\boxed{\text{A + (-A) = 0}}$

Demonstração: Seja X uma matriz tal que X + A = 0, em que 0 representa uma matriz nula (denotada por 0, uma matriz nula todos os elementos são iguais a zero). Observamos que $x_{ij} + a_{ij} = 0$, $\forall\ i, j$. Como a_{ij} é um número real, segue que $x_{ij} + a_{ij} = 0$ implica, $x_{ij} + a_{ij} + (-a_{ij}) = 0 + (-a_{ij})$. Logo, $x_{ij} = (-a_{ij})\ \forall\ i, j$. Portanto, todo elemento da matriz X é o oposto do elemento correspondente em A. Então, existe uma matriz (- A) = -(a_{ij}) tal que A + (- A) = 0

(4) **Existência do elemento neutro:** $\boxed{\text{A + 0 = A}}$

Demonstração: Seja X uma matriz tal que A + X = A. Então, $a_{ij} + x_{ij} = a_{ij}$, $\forall$ i, j, isto é, $x_{ij} = 0\ \forall\ i, j$. Portanto, a única matriz que satisfaz a equação $a_{ij} + x_{ij} = a_{ij}$ é a matriz nula. Portanto, existe uma matriz 0 tal que A + 0 = A.

3.3.3 Multiplicação de Matrizes

Voltando ao exemplo da empresa Alfa, suponhamos que a direção da empresa, afim de atender à demanda de seus sócios, precisa apresentar a receita gerada no primeiro bimestre do ano de 2015. Sabe-se que os preços de venda, em reais, de cada unidade, de seus produtos 1,2,3 e 4 são respectivamente, 3,9,8,2.
Então a receita total gerada nesse período pela fábrica 1 é dada por:

$$7 \text{ x } 3 + 5 \text{ x } 9 + 0 \text{ x } 8 + 1 \text{ x } 2 = 68 \text{ (milhões de reais)}$$

Entretanto, ao invês de calcularmos apenas a receita gerada por uma fábrica, podemos calcular simultâneamente com a mesma simplicidade a receita gerada por todas as fábricas. A resposta será uma matriz de ordem 1 x 3, em que o elemento da primeira linha corresponde à receita gerada pela fábrica 1, o elemento da segunda linha corresponde à fábrica 2 e por fim, o elemento da terceira linha corresponde à receita gerada pela fábrica 3.

$$\mathbf{T} = \begin{pmatrix} 7 & 5 & 0 & 1 \\ 0 & 4 & 3 & 7 \\ 3 & 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 9 \\ 8 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \cdot 3 + 5 \cdot 9 + 0 \cdot 8 + 1 \cdot 2 \\ 0 \cdot 3 + 4 \cdot 9 + 3 \cdot 8 + 7 \cdot 2 \\ 3 \cdot 3 + 2 \cdot 9 + 0 \cdot 8 + 2 \cdot 2 \end{pmatrix} = \begin{pmatrix} 68 \\ 74 \\ 31 \end{pmatrix}$$

Podemos observar que, a multiplicação entre os elementos de uma linha pelos elementos de uma coluna é uma ação fundamental para a multiplicação de matrizes. Este exemplo sugere a definição geral para a multiplicação de matrizes.
Definição: Sejam $A = (a_{ij})$ e $B = (b_{ij})$ matrizes de ordem m x k e k x n respectivamente. O produto AB é a matriz $C = (c_{ij})$, de ordem m x n, em que

cada elemento c_{ij} é dado por:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}$$

$$\begin{pmatrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & \vdots & \vdots \\ * & * & \cdots & * \end{pmatrix} \cdot \begin{pmatrix} * & \cdots & b_{1j} & \cdots & * \\ * & \cdots & b_{2j} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \\ * & \cdots & b_{kj} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \\ * & \cdots & b_{kj} & \cdots & * \end{pmatrix} = \begin{pmatrix} * & \cdots & \cdots & \cdots & * \\ & & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & \cdots & c_{ij} & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & \cdots & \cdots & * \end{pmatrix}$$

Exemplo 3.3.3.1 Dadas as matrizes

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & -1 & 2 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \\ 2 & 4 & 1 \end{pmatrix}$$

Vamos determinar produto $AB = C = (c_{ij})$.

Note que o elemnto c_{22} é dado pela soma dos produtos entre os elementos correspondentes da segunda linha da matriz A pelos da segunda coluna da matriz B conforme segue demonstrado abaixo.

$$\begin{array}{c} \downarrow \qquad\qquad\qquad\qquad \downarrow \end{array}$$

$$\longrightarrow \begin{pmatrix} * & * & * \\ 4 & 5 & 6 \\ * & * & * \end{pmatrix} \cdot \begin{pmatrix} * & 2 & * \\ * & 0 & * \\ * & 4 & * \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & \boxed{4 \,.\, 2 + 5 \,.\, 0 + 6 \,.\, 4} & * \\ * & * & * \end{pmatrix}$$

Logo,

$$\downarrow \qquad\qquad\qquad \downarrow$$

$$\longrightarrow \begin{pmatrix} * & * & * \\ 4 & 5 & 6 \\ * & * & * \end{pmatrix} \cdot \begin{pmatrix} * & 2 & * \\ * & 0 & * \\ * & 4 & * \end{pmatrix} = \begin{pmatrix} * & * & * \\ * & 32 & * \\ * & * & * \end{pmatrix}$$

Calculando $c_{ij}\ \forall\ i, i$ obtemos:

$$\begin{aligned}
c_{11} &= 1 \cdot 1 + 2 \cdot (-1) + 3 \cdot 2 = 5 \\
c_{12} &= 1 \cdot 2 + 2 \cdot 0 + 3 \cdot 2 = 14 \\
c_{13} &= 1 \cdot 3 + 2 \cdot 3 + 3 \cdot 1 = 12 \\
c_{21} &= 4 \cdot 1 + 5 \cdot (-1) + 6 \cdot 2 = 11 \\
c_{22} &= 4 \cdot 2 + 5 \cdot 0 + 6 \cdot 0 = 32 \\
c_{23} &= 4 \cdot 3 + 5 \cdot 3 + 6 \cdot 1 = 33 \\
c_{31} &= 1 \cdot 1 + (-1) \cdot (-1) + 2 \cdot 2 = 6 \\
c_{32} &= 1 \cdot 2 + (-1) \cdot 0 + 2 \cdot 4 = 10 \\
c_{33} &= 1 \cdot 3 + (-1) \cdot 3 + 2 \cdot 1 = 2
\end{aligned}$$

Portanto, a matriz $C = c_{ij}$ é dada por

$$C = \begin{pmatrix} 5 & 14 & 12 \\ 11 & 32 & 33 \\ 2 & 10 & 2 \end{pmatrix}$$

Exemplo 3.3.3.2: Dados os aeroportos de Londrina-PR (**L**) e Porto Alegre - RS (**P**), e os aéroportos de Esmeraladas - MG (**E**), Brasília DF (**B**) e Três Lagoas - MS (**T**). A figura a seguir, mostra as possíveis rotas de voos diretos entre essas cidades.

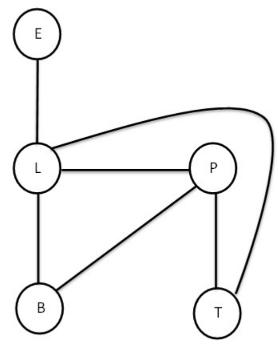

Figura 3.4: Rotas de voos diretos

A fim de analizarmos as rotas de voo, primeiro codificamos a rede em uma matriz. Se há um voo direto entre as cidades, colocamos 1, caso contrário 0.

Sendo assim o quadro que indica as rotas de voos pode ser representado por:

	L	P	E	B	T
L	0	1	1	1	1
P	1	0	0	1	1
E	1	0	0	0	0
B	1	1	0	0	0
T	1	1	0	0	0

Tabela 3.7: Matriz de voos diretos

Dessa forma podemos observar que não existem voos diretos entre as cidades de

Esmeralda, Brasília e Três Lagoas.

Podemos interpretar a matriz $AA = A^2$ como sendo a tabela das possíveis rotas de voos entre as cidades com exatamente uma escala.

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 0 & 1 & 1 \\ 2 & 3 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 \end{pmatrix}$$

Note que, há três possibilidades de viagens de ida e volta a Porto Alegre passando por outras cidades, mas não há rotas entre as cidades e Londrina e Esmeralda com uma escala.

3.3.4 Associatividade do Produto de Matrizes

Uma propriedade importante do produto matricial, é a associatividade, isto é, dadas as matrizes A_{mxn},B$_{nxp}$ e C_{pxq}. Então,

$$\boxed{\text{(AB)C = A(BC)}}$$

Demonstração: Sejam X e Y matrizes tais que $X = (AB)C$ e $Y = A(BC)$, e sejam x_{ij} e y_{ij} elementos quaisquer de X e Y respectivamente, queremos mostrar que $X = Y$.

Tomando de X um elemento x_{ij} arbitrário temos que

$$
\begin{aligned}
x_{ij} &= \textstyle\sum_{i=1}^{p} (ab)_{ik} c_{kj} \\
&= \textstyle\sum_{k=1}^{p} (\sum_{s=1}^{n} a_{is} b_{sk}) c_{kj} \\
&= \textstyle\sum_{k=1}^{p} (\sum_{s=1}^{n} a_{is} b_{sk} c_{kj}) \\
&= \textstyle\sum_{s=1}^{n} (\sum_{k=1}^{p} a_{is} b_{sk} c_{kj}) \\
&= \textstyle\sum_{s=1}^{n} a_{is} (\sum_{k=1}^{p} b_{sk} c_{kj}) \\
&= \textstyle\sum_{s=1}^{n} a_{is} (bc)_{sj} \\
&= (a\,(bc))_{ij} = y_{ij}
\end{aligned}
$$

Portanto, $X = Y$, como queríamos mostrar.

3.3.5 Distributividade do Produto em relação a Soma

Outra propriedade importante é a distribuitividade do produto em relação a soma, isto é, das as matrizes A, B e C, compatíveis com as condições de existência da soma e produto temos:

$$\boxed{\text{A(B + C) = AB + AC}}$$

Demonstração: Sejam X e Y matrizes tais que $X = A(B + C)$ e $Y = AB + AC$, queremos mostrar que $X = Y$. Para isso, tomamos x_{ij} um elemento qualquer de X. Logo, podemos escrever que $x_{ij} = (a(b + c))_{ij}$. Então $x_{ij} = (a(b + c))_{ij} \Longrightarrow \mathrm{x}_{ij} = a_{ik}(b + c)_{kj} \Longrightarrow \mathrm{x}_{ij} = a_{ik}(b_{kj} + c_{kj}) \Longrightarrow \mathrm{x}_{ij} = a_{ik}b_{kj} + a_{ik}c_{kj} \Longrightarrow \mathrm{x}_{ij} = (ab)_{ij} + (ac)_{ij} \Longrightarrow \mathrm{x}_{ij} = (ab + ac)_{ij} \Longrightarrow \mathrm{x}_{ij} = y_{ij}$.

Por outro lado, tomando y_{ij} um elemento qualquer de Y, podemos escrever que, $y_{ij} = (ab + ac)_{ij}$. Logo, sabemos que, $y_{ij} = (ab)_{ij} + (ac)_{ij} \Longrightarrow \mathrm{y}_{ij} = a_{ik}b_{kj} + a_{ik}c_{kj} \Longrightarrow \mathrm{y}_{ij} = a_{ik}(bkj + c_{kj}) \Longrightarrow \mathrm{y}_{ij} = a_{ik}(b+c)_{kj} \Longrightarrow \mathrm{y}_{ij} = (a(b+c))_{ij}$. Portanto, $y_{ij} = x_{ij}$. Então, podemos concluir que $X = Y$, como queríamos mostrar.

3.3.6 Álgebra Numérica versus Álgebra Matricial

Podemos estabelecer muitas semelhanças entre a álgebra númerica e a álgebra matricial, pórem existem quatro diferenąs fundamentais.

O produto AB não está definido para quaisquer matrizes A e B, pois o produto apenas faz sentido quando o número de linhas da matriz A é igual ao número de colunas da matriz B.

O produto AB não é comutativo. Mesmo que os produtos AB e BA estejam definidos, não se tem necessáriamente $AB = BA$. Por exemplo:

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

Enquanto que

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

O produto de duas matrizes não nulas, pode gerar a matriz nula. Como exemplo:

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

Para todo $x \in \mathbb{R}$ diferente de zero, existe $x^{-1} \in \mathbb{R}$, tal que, $x.x^{-1} = 1$.

Esse fato não ocorre na álgebra matricial. Como exemplo:

Sejam A e B duas matrizes não nulas tais que $AB = 0$. Então nem A nem B possuem inversa. Pois,

$$A^{-1}(AB) = (A^{-1}A)B = B = 0$$

4 Matrizes Estocásticas

4.1 Introdução

As matrizes estocásticas exercem papel fundamental na teoria das cadeias de Markov. Veja a seguir um exemplo típico de uma matriz estocástica.

$$A = \begin{pmatrix} 0,1 & 0,01 & 0,3 \\ 0,2 & 0,99 & 0,3 \\ 0,7 & 0 & 0,4 \end{pmatrix}$$

Toda matriz estocástica possui duas propriedades:

(1) todos os seus elementos são maiores ou iguais a zero;

(2) a soma dos elementos em cada coluna é igual a 1.

Para o desenvolvimento da teoria das cadeias de Markov, estaremos interessados nas potências dessas matrizes. As matrizes estocásticas estão fortemente conectadas com a teoria das Probabilidades.

Exemplo 4.1.1: Um homem ou dirige seu carro ou toma um ônibus para ir ao trabalho todos os dias. Suponhamos que ele nunca tome o ônibus dois dias seguidos, mas se ele dirige até o trabalho, então no dia seguinte, a probabilidade dele dirigir novamente é a mesma dele tomar um ônibus.

A matriz estocástica que representa as probabilidades de escolhas entre ir ao trabalho de carro ou de ônibus é dada por:

$$\mathbf{A} = \begin{pmatrix} 0,5 & 1 \\ 0,5 & 0 \end{pmatrix}$$

em que a primeira coluna corresponde a hipótese de que no dia seguinte em que ele dirigiu até o trabalho, ele dirige ou toma um ônibus no dia seguinte com igual probabilidade. A segunda coluna corresponde a hipótese de que ele nunca vai ao trabalho de ônibus dois dias seguidos.

Definição: Uma matriz quadrada é denominada matriz estocástica quando todos os seus elementos são números reais não negativos e a soma dos elementos em cada coluna é sempre igual a 1.

Exemplo 4.1.2: A matriz P é estocástica.

$$\mathbf{P} = \begin{pmatrix} 3/4 & 1/2 & 0 & 0 \\ 0 & 1/4 & 3/5 & 1/3 \\ 1/8 & 1/4 & 1/5 & 1/3 \\ 1/8 & 0 & 1/5 & 1/3 \end{pmatrix}$$

Pois, todos os seus elementos são maiores ou igual zero e a soma dos elementos de cada coluna é igual a 1.

Teorema 4.1.1: Sejam A $= (a_{ij})\ e\ B = (b_{ij})$ matrizes estocásticas de ordem m. Então o produto AB é uma matriz estocástica. Em particular, todas as potênicas A^k com k inteiro positivo, são matrizes estocásticas.

Demonstração: Considere as matrizes quadradas A $= a_{ij}$ e B= b_{ij} de mesma ordem. Sabemos que o produto AB é dado por AB= c_{ij}, em que

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj}$$

Afim de mostrar que o produto AB é estocástica, tomamos arbitrariamente uma coluna k de AB. Somando seus elementos obtemos

$$\begin{aligned}
&c_{1k} + c_{2k} + \cdots + c_{mk} = \\
&= (a_{11}b_{1k} + \cdots + a_{1m}b_{mk}) + (a_{21}b_{1k} + \cdots + a_{2m}b_{mk}) + \cdots + (a_{m1}b_{1k} + \cdots + a_{mm}b_{mk}) \\
&= b_{1k}\,(a_{11} + \cdots + a_{m1}) + b_{2k}\,(a_{12} + \cdots + a_{m2}) + \cdots + b_{mk}(a_{1m} + \cdots + a_{mm})
\end{aligned}$$

Como a matriz A=(a_{ij}) é uma matriz estocástica segue que

$$c_{1k} + c_{2k} + \cdots + c_{mk} = b_{1k}.1 + b_{2k}.1 + \cdots + b_{mk}.1 = b_{1k} + \cdots + b_{mk}$$

Entretanto, como B=(b_{ij}) também é estocástica temos que $b_{1k} + \cdots + b_{mk}$.

Logo, $c_{1k} + c_{2k} + \cdots + c_{mk} = 1$

Portanto, o produto entre duas matrizes estocásticas é uma matriz estocástica.

4.2 Matriz Estocástica Regular

Uma matriz estocástica P é dita uma matriz estocástica regular, se todos os elementos de alguma potência P^n são estritamente positivos, sendo n um inteiro positivo.

Exemplo 4.2.1. A migração de animais é o deslocamento deles em grupos de um lugar para outro. Muitos deles fazem esse movimento na procura de alimento ou na busca de um local para se reproduzir, exemplo disso é a tartaruga-de-couro (*Dermochelys coriacea*). Sua imigração é uma das mais impressionantes. As fêmeas ficam circulando entre a América e a Europa e escolhem o litoral brasileiro para desovar [13].

Agora suponha que um biólogo especialista em comportamento migratório, realizou um estudo sobre a migração da tartaruga-de-couro após o período da desova afim de estabelecer um padrão migratório para espécie. Ele constatou empiricamente que dentre as fêmeas monitoradas que partiram dos EUA, 50% delas migraram para o México após a desova e que 50% migraram para a Europa. Dentre as fêmeas que partiram do México, 20% migraram para os EUA, 50% retornaram ao México e que 30% migraram para a Europa. Dentre as fêmeas que partiram da Europa, todas migraram para o México após o período de desova.

Considere:

$$\begin{cases} \textbf{x = número de fêmeas que partiram dos EUA} \\ \textbf{y = número de fêmeas que partiram do México} \\ \textbf{z = número de fêmeas que partiram da Europa} \end{cases}$$

$$\begin{cases} \textbf{x' = número de fêmeas que migraram para os EUA} \\ \textbf{y' = número de fêmeas que migraram para o México} \\ \textbf{z' = número de fêmeas que migraram para Europa} \end{cases}$$

Então de acordo com os dados obtemos:

$$\begin{cases} \textbf{x' = 20\% de y} \\ \textbf{y' = 50\% de x + 50\% de y + z} \\ \textbf{z' = 50\% de x + 30\% de y} \end{cases}$$

Dessa forma obtemos a seguinte equação:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 0,2 & 0 \\ 0,5 & 0,5 & 1 \\ 0,5 & 0,3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Indicando por M a matriz

$$M = \begin{pmatrix} 0 & 0,2 & 0 \\ 0,5 & 0,5 & 1 \\ 0,5 & 0,3 & 0 \end{pmatrix}$$

e denotando

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad\qquad v' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

obtemos a seguinte equação

$$v' = Mv$$

Vamos admitir que as taxas de migração sejam mantidas a cada ano. Além disso vamos supor que a população monitorada permanece constante.

Para o próximo período de desova a população de tartarugas-de-couro que migrou para, EUA, México e Europa será dada por

$$v' = M(Mv)$$

Como o produto de matrizes é associativo, tem - se

$$v' = M^2v.$$

Observe que

$$M^2 = \begin{pmatrix} 0 & 0,2 & 0 \\ 0,5 & 0,5 & 1 \\ 0,5 & 0,3 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0,2 & 0 \\ 0,5 & 0,5 & 1 \\ 0,5 & 0,3 & 0 \end{pmatrix} = \begin{pmatrix} 0,1 & 0,1 & 0,2 \\ 0,75 & 0,65 & 0,5 \\ 0,15 & 0,25 & 0,3 \end{pmatrix}$$

Portanto, a cada ano a população que migra é obtida pela multiplicação da população anterior matriz M. Daí se conclui que decorridos k anos, a população que migra é dada por

$$M^k v.$$

Por este motivo diz - se que o fluxo migratório das tartarugas-de-couro monitoradas é representada pela matriz M. Podemos observar através da matriz M, que após a desova, apenas migraram para os EUA, animais que partiram do México. Mas no segundo período do ciclo migratório de desova, todos os locais de observação EUA, México e Europa receberão animais vindo de todos os três locais de partida. Isto pode ser visto do fato de que todas as entradas de M^2 são positivas.
Observe que o produto de uma matriz com entradas positivas por uma matriz estocástica resulta sempre numa matriz com entradas positivas.

4.3 Ponto fixo de uma Matriz Estocástica Regular

Seja A uma matriz quadrada. Um vetor não nulo $w \in \mathbb{R}$ é chamado de autovetor da matriz A, se existir $\lambda \in \mathbb{R}$ tal que

$$Aw = \lambda w$$

O escalar λ é chamado de autovalor da matriz A associado ao autovetor w.
Quando o autovetor w está associado ao autovalor $\lambda = 1$, dizemos que w é um ponto fixo da matriz A. Assim um ponto fixo de A é um vetor w, não nulo, w tal que

$$Aw = w$$

Exemplo 4.3.1. Observe que o vetor

$$\mathbf{w} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

é um autovetor da matriz

$$\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$$

associado ao autovalor $\lambda = 3$, pois Aw = 3w. Vejamos:

$$\begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Teorema 4.3.1. Seja A uma matriz quadrada e k um inteiro positivo. Se w é um autovetor de A associado ao autovalor λ então w é um autovetor de A^k associado ao autovalor λ^k.

Demonstração: Vamos mostrar por indução que a sentença

$$A^k w = \lambda^k w$$

é verdadeira para todo $k \in \mathbb{N}$.

Observe que para $k = 1$ a sentença é verdadeira, pois por hipótese λ é um autovetor da matriz A associado ao autovetor w, ou seja,

$$Aw = \lambda w.$$

Agora segue da equação acima que

$$A^2w = A(Aw) = A(\lambda w) = \lambda(Aw) = \lambda^2 w$$

Agora, supondo que a sentença seja verdadeira para algum $k = n$ com $n \in \mathbb{N}$, vamos mostrar que a ela é verdadeira para $k = n + 1$. De fato, suponha que

$$A^n w = \lambda^n w, n \in \mathbb{N}.$$

Multiplicando ambos os membros da igualdade pela matriz A, obtemos

$$A(A^n w) = A(\lambda^n w).$$

Pela propriedade associativa do produto de matrizes segue que

$$(AA^n)w = \lambda^n(Aw) = \lambda^n(\lambda w)$$

$$A^{n+1}w = \lambda^{n+1}w.$$

Assim, mostramos que a sentença é verdadeira para $k = n + 1$.
Portanto, pelo Princípio da Indução Finita, $\mathrm{A}^k w = \lambda^k w$, é verdadeira para todo k inteiro positivo, como queríamos mostrar.

Em particular, quando $\lambda = 1$, obtemos o seguinte.

Corolário 4.3.1. Se w é um ponto fixo de A então w é ponto fixo de A^k, $\forall k \in \mathbb{N}$ com $k \geq 1$.

Teorema 4.3.2. Toda matriz quadrada não nula $A = (a_{ij})$ tem pelo menos um ponto fixo se, somente se, o determinante da matriz A - I é nulo, em que I é a matriz identidade.

Demonstração: A matriz $A = (a_{ij})$ tem ponto fixo se, somente se, existe um vetor w, não nulo, dado por

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}$$

tal que Aw = w, isto é,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{pmatrix}$$

Então, segue que

$$\begin{cases} a_{11}w_1 + a_{12}w_2 + \cdots + a_{1m}w_m = w_1 \\ a_{21}w_1 + a_{22}w_2 + \cdots + a_{2m}w_m = w_2 \\ \vdots \\ a_{m1}w_1 + a_{m2}w_2 + \cdots + a_{mm}w_m = w_m \end{cases}$$

Em outras palavras, podemos dizer que w é um ponto fixo de $A = (a_{ij})$, se existir solução não trivial para o sistema homogêneo abaixo.

$$\begin{cases} (a_{11} - 1)\, w_1 + a_{12}w_2 + \cdots + a_{1m}w_m = 0 \\ a_{21}w_1 + (a_{22} - 1)\, w_2 + \cdots + a_{2m}w_m = 0 \\ \vdots \\ a_{m1}w_1 + a_{m2}w_2 + \cdots + (a_{mm} - 1)\, w_m = 0 \end{cases}$$

Como um sistema homogêneo possui solução não trivial se, somente se, o determinante da matriz dos coeficientes é nulo, concluímos que

$$\begin{vmatrix} (a_{11}-1) & a_{12} & \cdots & a_{1m} \\ a_{21} & (a_{22}-1) & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & (a_{mm}-1) \end{vmatrix} = 0$$

Portanto podemos concluir que uma matriz quadrada possui ponto fixo não nulo se, e somente se, det(A - I) = 0.

Teorema 4.3.3. Toda matriz estocástica A tem pelo menos um ponto fixo.

Demonstração: Pelo fato da matriz A ser estocástica a soma dos elementos de cada coluna da matriz A - I é zero. Logo a soma de todas as linhas da matriz $A - I$ é a linha nula. O determinante de uma matriz não se altera quando se troca uma linha pela soma dela com a soma das outras linhas. Substituindo a última linha da matriz A - I pela soma de todas as linhas obtemos a linha nula. Logo det (A - I) = 0. Portanto, pelo teorema 4.3.2 toda matriz estocástica possui pelo menos um ponto fixo.

Exemplo 4.3.2. Consideremos a matriz

$$\mathbf{A} = \begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}$$

Então, o vetor

$$\mathbf{w} = \begin{pmatrix} 2/5 \\ 3/5 \end{pmatrix}$$

é um ponto fixo de A. De fato, pois,

$$\begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix} \begin{pmatrix} 2/5 \\ 3/5 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}.\frac{2}{5} + \frac{1}{3}.\frac{3}{5} \\ \\ \frac{1}{2}.\frac{2}{5} + \frac{2}{3}.\frac{3}{5} \end{pmatrix} = \begin{pmatrix} 2/5 \\ 3/5 \end{pmatrix}$$

Portanto,

$$Aw = w.$$

Teorema 4.3.4. Seja w um ponto fixo da matriz A e seja $x \in \mathbb{R}$ com $x \neq 0$, então o vetor xw também é um ponto fixo de A.

Demonstração: Sabemos que w é um ponto fixo de A então podemos escrever

$$Aw = w \Rightarrow x(Aw) = xw \Rightarrow A(xw) = xw.$$

O que demonstra que o vetor xw também é um ponto fixo de A.

Observação: Vamos mostrar que todos os pontos fixos da matriz A do exemplo 4.3.2 são colineares.

De fato seja (x, y) um ponto fixo de A. Então

$$A(x, y) = (x, y)$$

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{3} \\ \\ \frac{1}{2} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x \\ \\ y \end{pmatrix}$$

Logo

$$\begin{cases} \frac{1}{2}x + \frac{1}{3}y = x \\ \\ \frac{1}{2}x + \frac{2}{3}y = y \end{cases}$$

O que equivale a $y = \frac{3}{2}x$.

Então

$$(x, y) = \left(x, \frac{3}{2}x\right) = x\left(1, \frac{3}{2}\right) = \frac{5}{2}x\left(\frac{2}{5}, \frac{3}{5}\right).$$

Definição: Se o vetor w é um ponto fixo de A e além disso $w_1 + w_2 + \cdots + w_m = 1$ com $w_i \geq 0$ com $i = 1, \cdots, m$. Dizemos que w é um vetor fixo de probabilidade.

Teorema Fundamental da Matriz Estocástica Regular: Seja P uma matriz estocástica regular. Então, P possui um único vetor de probabilidade fixo w e todas as coordenadas de w são positivas.

Demonstração: Como P é uma matriz estocástica, pelo teorema 4.3.3 P possui pelo menos um ponto fixo w, logo

$$Pw = w.$$

Pelo corolário 4.3.1, w é um ponto fixo de P^n, para todo $n \geq 1$. Como P é uma matriz estocástica, seja $m \in \mathbb{N}$ tal que todas as entradas de P^m são positivas, logo

$$P^m w = w.$$

Seja $P^m = (a_{ij})$, $a_{ij} > 0$, $1 \leq i, j \leq n$.

Então,

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}.$$

Trocando w por $-w$, se necessário, podemos supor que w possui alguma coordenada positiva. Vamos mostrar que todas as coordenadas de w são positivas.

Suponha por absurdo que w possui k coordenadas não positivas e consequentemente w possui $n - k$ coordenadas positivas. Sem perda de generalidade podemos supor que $w_j \leq 0$ com $1 \leq j \leq k$ e que $w_j > 0$ com $k + 1 \leq j \leq n$.

Então podemos escrever o sistema

$$\begin{cases} a_{11}w_1 + a_{12}w_2 + \cdots + a_{1n}w_n = w_1 \\ a_{21}w_1 + a_{22}w_2 + \cdots + a_{2n}w_n = w_2 \\ \vdots \\ a_{k1}w_1 + a_{k2}w_2 + \cdots + a_{kn}w_n = w_k \\ \vdots \\ a_{n1}w_1 + a_{n2}w_2 + \cdots + a_{nn}w_n = w_n. \end{cases}$$

Considerando apenas as k primeiras equações obtemos

$$\begin{cases} a_{11}w_1 + a_{12}w_2 + \cdots + a_{1k}w_k + a_{1k+1}w_{k+1} + \cdots + a_{1n}w_n = w_1 \\ a_{21}w_1 + a_{22}w_2 + \cdots + a_{2k}w_k + a_{2k+1}w_{k+1} + \cdots + a_{2n}w_n = w_2 \\ \vdots \\ a_{k1}w_1 + a_{k2}w_2 + \cdots + a_{kk}w_k + a_{kk+1}w_{k+1} + \cdots + a_{kn}w_n = w_k \end{cases}$$

Isolando as $(n-k)$ últimas variáveis obtemos

$$\begin{cases} a_{1k+1}w_{k+1} \cdots + a_{1n}w_n = w_1 - (a_{11}w_1 + a_{12}w_2 + \cdots + a_{1k}w_k) \\ a_{2k+1}w_{k+1} \cdots + a_{2n}w_n = w_2 - (a_{21}w_1 + a_{22}w_2 + \cdots + a_{2k}w_k) \\ \vdots \\ a_{kk+1}w_{k+1} \cdots + a_{kn}w_n = w_k - (a_{k1}w_1 + a_{k2}w_2 + \cdots + a_{kk}w_k)\,. \end{cases}$$

Somando as k equações temos:

$$w_{k+1}\left(a_{1k+1} + \cdots + a_{kk+1}\right) + \cdots + w_n\left(a_{1n} + \cdots + a_{kn}\right) = w_1\left(1 - \sum_{i=1}^{k} a_{i1}\right) + \cdots + w_k\left(1 - \sum_{i=1}^{k} a_{ik}\right)$$

Lembrando que $a_{ij} > 0$, para todo $1 \leq i,\, j \leq n$, logo

$$0 < \underbrace{w_{k+1}}_{>0}\underbrace{\left(a_{1k+1} + \cdots + a_{kk+1}\right)}_{>0} + \cdots + \underbrace{w_n}_{>0}\underbrace{\left(a_{1n} + \cdots + a_{kn}\right)}_{>0} = \underbrace{\sum_{j=1}^{k} w_j\left(1 - \sum_{i=1}^{k} a_{ij}\right)}_{\leq 0} \leq 0$$

Isto caracteriza um flagrante absurdo. Portanto w não possui coordenadas menores ou igual a zero.

Logo todas as coordenadas de w são positivas, isto é, $w = (w_1, \cdots, w_n)$ com $w_i > 0$ $\forall\ 1 \leq i \leq n$.

Tomando $w^* = \dfrac{1}{w_1 + \cdots + w_n} w$, concluímos que w é um vetor de probabilidade fixo, com todas as coordenadas positivas.

Agora vamos mostrar que o vetor de probabilidade de P é único. Suponhamos que existam dois vetores de probabilidade u e v, $u \neq v$, tais que u e v são pontos fixos de P. Então

$$Pu = u \quad e \quad Pv = v.$$

Além disso

$$\sum_{i=1}^{n} (u - v)_i = \sum_{i=1}^{n} u_i - \sum_{i=1}^{n} v_i = 1 - 1 = 0.$$

Como $u - v \neq 0$ então o vetor $u - v$ possui pelo menos uma coordenada maior que zero e pelo menos uma coordenada menor que zero. Mas isso contraria ao que demonstrado anteriormente.

Logo, P possui um único vetor de probabilidade fixo, como queríamos demonstrar.

5 Cadeias de Markov

5.1 Introdução

Observando a natureza, pode-se perceber uma equilibrada dicotomia. Pois não existem duas coisas exatamente iguais, mas todas elas tendem a seguir algum padrão não manifestado claramente. Platão acreditava que a verdadeira forma do universo estava oculta. Observando a natureza, apenas conseguiríamos atingir um conhecimento aproximado sobre a verdade pura em outras palavras, a verdade estava escondida em planos impercepitíveis aos nossos sentidos. Sendo assim, a verdadeira forma pura das coisas, somente era acessível por meio do raciocínio abstrato da filosofia e da matemática, como exemplo um círculo, cuja a distância da borda até o centro é sempre a mesma. Entretanto, nós nunca encontraremos uma manifestação material de um círculo perfeito ou de uma linha reta perfeita. Embora interessado, Platão especulava que, passados uma quantidade arbitrariamente grande de anos, o universo encontraria um estado ideal, retornando a sua forma perfeita. Esse foco platônico em formas puramente abstratas, perdurou até o século XVI, momento em que as pessoas passaram a interpretar o obscuro, as variações do mundo real e a aplicar a matemática como ferramenta para elucidar os padrões subjacentes da natureza.

Essa corrente de ideias foi amplamente extendida, notando que, não apenas as coisas convergem para uma razão esperada, mas a probabilidade da variação em relação as medidas, também seguem uma familiar forma subjacente ou uma distribuição.

Um excelente exemplo dessa ideia, é a máquina de feijões de Francis Galton.

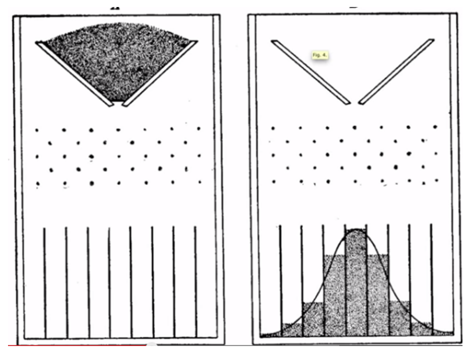

Figura 5.1: Máquina de feijões de Francis Galton

Imagine cada colisão como um evento independente, como uma virada em uma moeda honesta. Após 10 colisões, o feijão cai em um recipiente que representa a razão entre a deflaxão entre esquerda e direita ou ainda como sendo cara ou coroa. Essa curvatura é conhecida como distribuição binomial. Aparenta que a média do destino desses eventos é de alguma forma predeterminada, conhecida hoje como Teorema Central do Limite.

Isso representou uma ideia filosófica perigosa para alguns. Parul Nekrasov, um teólogo por formação que mais tarde acabou seguindo os passos da matemática, era um influente proponente da doutrina religiosa do livre arbítrio. Ele não admitia a ideia de termos esse destino estatístico predeterminado. Ele afirmava que a independência era uma condição necessária para a lei dos grandes números desde que a independência descrevesse esses exemplos simples como os que utilizam feijões, dados ou moedas, os quais o resultado de um evento anterior não afeta a probabilidade da ocorrência do evento futuro. No entanto, como podemos observar, a grande parte dos fenômenos da física, são claramente dependentes de um resultado anterior.

Quando a probabilidade de algum evento depende ou é condicional a um evento anterior, dizemos que eles são eventos dependentes.

Essa afirmação irritou um matemático russo, Andrey A. Markov, o qual manteve publicamente um desafeto com Nekrasov. Markov estendeu os trabalhos realizados por Bernoulli sobre variáveis dependentes utilizando uma construção engenhosa. Imagine uma moeda a qual não seja independente, mas dependente apenas do resultado anterior, então esse processo teria uma memória de curto prazo. Esse processo pode ser visualizado utilizando uma máquina hipotética composta por dois potes, os quais chamamos de estados. Em um estado temos uma mistura de 50 bolas pretas e 50 bolas brancas indistinguíveis. Enquanto no outro, temos uma mistura com mais bolas pretas do que brancas. Um pote podemos chamar de estado 0 (zero), ele representa a ocorrência anterior a uma bola preta, e o outro, podemos chamar de 1 (um), ele representa a bola branca tendo ocorrido anteriormente.

Para iniciarmos o uso da máquina, simplismente começamos em um dos estados retirando uma bola. A partir daí, iremos para o estado 0 ou 1, dependendo do resultado daquele evento. Como esse processo possui dois estados, podemos identificar 4 possíveis transições de estados. Se estamos no estado 0 e ocorre uma bola preta, permanecemos no estado 0, e retiramos uma outra bola. Se uma bola branca é retirada, o processo passa para o estado 1, o qual também pode permanecer em si mesmo, ou passar para o estado 0 se uma bola preta for retirada.

A probabilidade de uma bola branca ser retirada em relação uma bola preta é claramente não independente

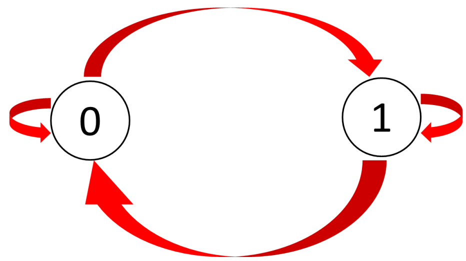

Figura 5.2: Máquina de dois estados

Entretanto, Markov demonstrou que quando todos os estados da máquina são acessíveis, ao ligar a máquina em uma sequência, ela atingirá o equilíbrio, ou seja, não importa em que estado comece, uma vez que, iniciada a sequência, o número de vezes que o processo visita cada estado, converge para um equilíbrio, isto é, para uma distribuição de probabilidades.

Esse simples experimento, não dialóga com a afirmação de Nekrasov de que apenas eventos independentes podem convergir para distribuições previsíveis. No entanto, a prática de modelar as sequências de eventos aleatórios utilizando estados e suas transições ficou conhecido como cadeias de Markov.

As cadeias de Markov são assim denominadas em homenagem ao matemático russo Andrei Andreyevich Markov (1856 -1922) que as utilizou para analisar alterações de vogais e de consoantes no poema Eugene Onegin, de Pushkin. Ele não enxergava outra aplicação para sua teoria na época, certamente ficaria muito satisfeito se soubesse do sucesso e da ampla utilização de sua descoberta. A cadeia de Markov é um processo estocástico desmemoriado, isto é, um processo estocástico no qual o estado futuro depende apenas do estado presente, mas não dos estados passados.

5.2 Andrei Andreyevich Markov

Andrei Andreyevich Markov nasceu em 14 de junho de 1856 em Ryazan, Rússia, e morreu em 20 de julho de 1922 em São Petersburgo foi o primeiro filho do casal Andrei Grigorievich Markov e Nadezhda Petrovna.

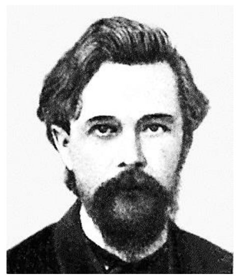

Figura 5.3: Andrei Andreyevich Markov

Markov conviveu grande parte de sua infância com uma saúde muito frágil, apresentando dificuldades até mesmo para andar. Na adolescência demonstrou seu talento excepcional para a matemática, mas não teve o mesmo sucesso nas outras disciplinas. Publicou seu primeiro artigo quando ainda cursava o equivalente ao Ensino Médio cujo tema abordado foi Equações Diferenciais Lineares e, mesmo apresentando resultados já conhecidos, o artigo despertou a atenção de dois importantes professores da Universidade de São Petersburg, Aleksandr Korkin e Yegor Ivanovich Zolotarev.

A partir daí, ficou evidente que a Matemática era o caminho natural para Markov e, em 1874, ingressou na Faculdade de Física e Matemática de São Petersburgo. Por lá, matriculou-se no seminário dirigido pelos professores Korkin e Zolotarev, mas também assistiu a muitas palestras do professor Pafnuty Chebyshev, na época chefe do departamento de matemática.

Markov se formou em 1878 ganhando nesse ano o prêmio de melhor artigo publicado envolvendo o tema Integração de Equações Diferenciais por meio de Funções Contínuas. Querendo tornar-se professor, trabalhou em seu mestrado ao longo de dois anos obtendo o título em 1880 com um trabalho sobre Formas Quadráticas Binárias com Determinantes Positivos. Esse trabalho foi muito elogiado por Chebyshev e representou uma dos melhores trabalhos realizados sobre Teorias dos Nmeros de toda a Matemática Russa. Embora sua dissertação tenha sido publicada simultaneamente em francês, ela não foi absorvida imediatamente por matemáticos da Europa Ocidental, dando a medida do quanto Markov havia se aprofundado no assunto. Terminado o mestrado, Markov começou a lecionar na Universidade de São Petersburgo, enquanto trabalhava em seu doutoramento, concluindo-o em 1884 com a tese sobre Aplicações de Frações Contínuas.

Agora professor, Markov havia ganhado notabilidade social suficiente para assumir sua antiga paixão por Maria Ivanova Valvatyeva, pedindo-a em casamento. Os dois já se conheciam desde crianças, pois ela era a filha do proprietário da fazenda em que seu pai gerenciava. No entanto, a mãe de Ivanova não aceitava a ideia de que sua filha casasse com o filho do gerente da fazenda. Entretanto, os esfor{ccos para evitar a união não foram suficientes e, em 1883, a mãe de Ivanova concedeu a anuncia e o casamento aconteceu naquele mesmo ano.

Markov construiu uma sólida carreira como professor na Universidade de São Petersburgo e, além disso, ingressou na Academia Russa de Ciências indicado por Chebyshev, em 1883. Aposentou-se formalmente em 1905, mas continuou a ensinar por alguns anos.

Markov foi o porta-voz mais elegante das ideias e direções de pesquisa em Teoria de Probabilidades de Chebyshev. Ele é lembrado particularmente pelo estudo desenvolvido sobre Sequências de Variáveis Aleatórias em que a próxima variável é determinada no máximo pelo estado presente, mas é independente da forma pelo qual o estado atual surgiu de seus antecessores. Markov analisou a sequência de vogais e consoantes no poema Eugene Onegin, escrito por Alexander Pushkin em 1883. Ele verificou empiricamente que uma vogal era seguida por uma consoante em 87% das vezes e que uma consoante era seguida por uma vogal 66% das vezes. Este trabalho é conhecido como Cadeias de Markov e representa um marco na Teoria das Probabilidades. Em 1923, Nobert Winer tornou-se o primeiro a tratar rigorosamente os Processos Contínuos Markovianos. As bases de uma teoria geral foi fornecida durante a década de 1930 por Andrei Kolmogorov.

Markov viveu um período de grande atividade política na Rússia e, por ter opiniões firmes, tornou-se um militante político. Em 1913, a dinastia Romanov, que estava no poder na Rússia desde 1613, comemorou seus 300 anos de poder, mas Markov fez questão de demonstrar toda sua desaprovação pela celebração. Com o início da Revolução Russa em 1917, Markov foi enviado para Zaraisk, uma pequena cidade do interior da Rússia, onde ensinou matemática na escola secundária sem receber qualquer remuneração. Por fim, retornou a São Petersburgo, apresentando uma saúde debilitada, vindo a falecer em julho de 1922 depois de meses de sofrimento.

5.3 Conceitos Básicos sobre Cadeias de Markov

5.3.1 Introdução

Essa seção propõem-se a apresentar os principais conceitos e definições sobre as cadeias de Markov, fixando as ideias por meio da apresentação de alguns exemplos.

Em particular, os conceitos e ferramentas necessários para a análise da convergência de uma cadeia de Markov regular para sua distribuição estacionária.
Um conceito muito importante que será apresentado é sobre a matriz de transição de uma cadeia de Markov, elemento esse fundamental na teoria. A partir da matriz de transição, podemos caracterizá - la e doravante estudarmos sua distribuição de probabilidade estacionária utilizando operações matriciais básicas facilmente programadas. Isso é uma particularidade importante na aplicação da teoria.

5.3.2 Definição de uma cadeia de Markov

Considere um processo estocástico finito $\{X_n\}_{n>0}$ em que cada variável aleatória X_i assume, com certa probabilidade, um estado $s_i \in \{s_1, s_2, \cdots, s_n\}$ com $i = 1, 2, \cdots, n$.
Definição: Uma cadeia de Markov finita é um processo estocástico finito tal que:

$$P\left(X_n = s_n | X_1 = s_1, X_2 = s_2, ..., X_{n-1} = s_{n-1}\right) = P\left(X_n = s_n | X_{n-1} = s_{n-1}\right)$$

Podemos dizer outras palavras, que uma cadeia de Markov é um processo estocástico desmemoriado, isto é, o resultado de qualquer tentativa depende apenas do resultado da tentativa imediatamente anterior, não dependendo das demais tentativas.
A condição acima é conhecida como propriedade de Markov.
Seja $\{X_n\}_{n=0}^{\infty}$ uma cadeia de Markov, se $X_n = s_j$, dizemos que a cadeia de Markov, está no estado s_j no parâmetro n ou que a cadeia de Markov visita o estado s_j no parâmetro n.
A probabilidade do sistema passar do estado s_i para para o estado s_j no tempo n é

denotada por:

$$p_{ij}(n) = P(X_n = s_i | X_{n-1} = s_j)$$

As probabilidades de transição $p_{ij}(n)$ não dependem de n. Sendo assim, denotamos apenas por p_{ij}.

Se uma cadeia de Markov possui n estados possíveis, então as probabilidades de transição são convenientemente representadas em uma matriz quadrada de ordem n, $P = (p_{ij})$ denotada matriz de transição sendo representada por:

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

Assim, a cada estado s_j correspondente a j-ésima coluna da matriz P. Então o vetor coluna representa a distribuição da probabilidade condicional de X_n dado que $X_{n-1} = j$, sendo portanto, um vetor de probabilidade.

O vetor de probabilidade inicial é o vetor $w_0 = \left\{ p_j^{(0)} \right\} = \{P(X_0 = s_j)\}$.

Observe que P tem as seguintes propriedades:

$$p_{ij} \geq 0$$

$$\sum_{j=1}^{n} p_{ij} = 1$$

Exemplo 5.3.2.1. O que é maior: a probabilidade de que neve depois de uma de sol quente ou depois de um dia nublado e frio?
Situações como essa podem serem estudadas e modeladas pelas cadeias de Markov como segue na situação a seguir.
Na terra de Nárnia nunca ocorrem dois dias ensolarados seguidos. Se em um dia faz sol, então para o dia seguinte fica igualmente propenso a nevar ou chover. Se chover, no dia seguinte a probabilidade de chover novamente é de $\frac{1}{2}$ e fica igualmente propenso nevar ou fazer sol. Caso tenhamos um dia com preciptação de neve, para o dia seguinte temos a probabilidade de nevar novamente de $\frac{1}{2}$ e a de chover ou fazer sol igualmente propensos. Dessa forma a matriz de transição é dada por

$$P = \begin{pmatrix} 0,5 & 0,5 & 0,25 \\ 0,25 & 0 & 0,25 \\ 0,25 & 0,5 & 0,5 \end{pmatrix}$$

Ao usarmos uma cadeia de Markov para explicar um fenômeno da natureza, é fundamental conhecermos alguns detalhes. Em função da independência temporal, que está presente no conceito das cadeias de Markov. Então podemos ignorar o tempo de 10 dias atrás para prever o tempo amanhã. Entretanto, isso não significa dizer que o futuro não dependa do passado, ou ainda, que o passado não influência no futuro. Ao contrário, uma série de eventos contém mais informações que dois eventos isolados. Contudo, para prever o tempo amanhã, e em virtude das cadeias de Markov, sabemos que podemos desprezar o tempo de 10 dias atrás.
Na verdade, com a probabilidade, apenas medimos a chance de que algo venha acontecer, isso não quer dizer que sabemos com certeza o que vai acontecer.
Exemplo 5.3.2.2. Um homem joga em duas máquinas distintas de caça níquel. A

primeira paga o prêmio com probabilidade c, e a segunda com probabilidade d. Se ele perde, ele joga novamente na mesma máquina, se vence, ele muda para a outra máquina. Seja s_i o estado em que ele ele joga na máquina i. Então a matriz de transição é dada por

$$\mathbf{P} = \begin{pmatrix} c & 1-d \\ 1-c & d \end{pmatrix}$$

Exemplo 5.3.2.3. Um psicólogo faz as seguintes hipóteses, quanto ao comportamento de ratos sujeitos a um horário especial de alimentação. Para qualquer tentativa particular, 80% dos ratos que foram para direita no experimento anterior irão para a direita na próxima tenativa, e 60% dos que foram para a esquerda no experimento anterior irão para a direita na próxima tentativa. Os estados do sistema são D (direita) e E (esquerda). A matriz de tarnsição é dada por

$$\mathbf{P} = \begin{pmatrix} 0,8 & 0,6 \\ 0,2 & 0,4 \end{pmatrix}$$

Exemplo 5.3.2.4. Uma empresa de aluguel de automóveis tem três lojas. Um cliente pode alugar um carro em qualquer uma das lojas e devolvê-lo também em qualquer uma das três lojas. Por meio de um estudo realizado por essa empresa, estima-se que os clientes devolvem os automóveis às diferentes lojas, de acordo com as seguintes probabilidades representadas na matriz de transição a seguir, que dependem apenas da loja onde o automóvel foi alugado

$$\mathbf{P} = \begin{pmatrix} 0,8 & 0,3 & 0,2 \\ 0,1 & 0,2 & 0,6 \\ 0,1 & 0,5 & 0,2 \end{pmatrix}$$

Nesta matriz o valor $p_{23} = 0.6$ corresponde à probabilidade de um carro ser alugado na loja 2 ser devolvido à loja 3.

Exemplo 5.3.2.5. Em sociologia é conveniente classificar as pessoas por faixa de renda: classe baixa, classe média ou classe alta. Sociólogos descubriram que o fator predominante que influência a classe de renda de uma pessoa, é a classe de renda de seus pais. A matriz de transição P mostra as probabilidades de um indivíduo mudar de classe social dependendo apenas da classe de renda de seus pais.

$$\mathbf{P} = \begin{array}{c|ccc} & \textbf{classe baixa} & \textbf{classe média} & \textbf{classe alta} \\ \textbf{classe baixa} & 0,65 & 0,15 & 0,12 \\ \textbf{classe média} & 0,28 & 0,67 & 0,36 \\ \textbf{classe alta} & 0,07 & 0,18 & 0,52 \end{array}$$

A matriz de transição P mostra a probabilidade de mudança de classe de renda de uma geração para a próxima. Agora desejamos investigar as probabilidades de mudanças de classe de renda para duas gerações. Por exemplo, para pais de classe alta, qual a probabilidade de que tenham um neto de classe média?

A probabilidade de que pais de classe alta tenham netos de classe média é dada por

$$0.52 \cdot 0.36 + 0.36 \cdot 0.67 + 0.12 \cdot 0.28 = 0.462 = 46,2\%$$

5.4 Passeios Aleatórios Simples

Nesta seção iremos apresentar vários exemplos simples de cadeias de Markov. Esses exemplos dizem a respeito ao que chamamos de passeio aleatório.

Sejam $X_1, X_2, \cdots, X_n, \cdots$ variáveis aleatórias independentes e igualmente distribuídas. Sejam, $s_0 = c$ e $s_n = s_0 + \sum_{i=1}^{n} X_i$, com $n \geq 1$. O processo estocástico $\{s_n, n \geq 0\}$ é chamado de passeio aleatório simples. Em outras palavras, uma passeio aleatório simples consiste de uma partícula inicialmente num ponto $x \in Z$, que se move em Z e a cada instante pode pular d eum ponto x para um dos pontos vizinhos, $x+1$ ou $x-1$, com probabilidade p de saltar para direita e probabilidade $q = 1-p$ de saltar para esquerda. Se $p = q = \frac{1}{2}$, então dizemos que é um passeio aleatório simétrico em que s_n denota a posição da partícula no instante n.

Imagine um homem que se move em linha reta em passos com comprimento 1, cada passo corresponde a se deslocar uma unidade para a direita com probabilidade p ou uma unidade para a esquerda, com probabilidade q. Move-se até atingir um dos dois pontos extremos que são chamados de pontos de fronteira. As possibilidades de seu comportamento nestes pontos determinam vários tipos diferentes de cadeias de Markov. Os estados são as posições possíveis. Nos exemplos a seguir, estaremos levando em consideração o caso com 5 estados. Os estados 0 e 4 são os estados de fronteiras.

Exemplo 5.4.1. Um homem está num ponto de coordenada inteira sobre o eixo x entre a origem e o ponto 4. Ele dá um passo de uma unidade para a direita com probabilidade p ou para a esquerda com probabilidade $q = 1-p$, exceto quando ele estiver na origem, caso em que dá um passo para a direita chegando ao ponto 1, ou quando estiver sobre o ponto 4, caso em que dá um passo para esquerda chegando ao ponto 3. Indiquemos por X_n a sua posição, após n passos. Esta é uma (C.M.), cujo espaço de estados é $S = \{0, 1, 2, 3, 4\}$, em que cada estado $s_i = i$ representa que o homem está sobre o eixo x no ponto de abscissa i, com $i = 0, 1, 2, 3, 4$. A matriz de transição é dada por

$$\mathbf{P} = \begin{pmatrix} 0 & q & 0 & 0 & 0 \\ 1 & 0 & q & 0 & 0 \\ 0 & p & 0 & q & 0 \\ 0 & 0 & p & 0 & 1 \\ 0 & 0 & 0 & p & 0 \end{pmatrix}$$

Cada coluna da matriz, exceto a primeira e a última, correspondem ao fato de que o homem se move do estado s_i para o estado s_{i+1}, com probabilidade p ou para o estado s_{i-1} com probabilidade $q = 1 - p$. A primeira coluna corresponde ao fato de que o homem passa do estado s_0 sempre para o estado s_1 e a última coluna corresponde ao fato de que o homem sempre passa do estado s_4 para o estado s_3.

Exemplo 5.4.2. Supomos agora que sempre que o homem atinge um dos estados de fronteira vai diretamente para o estado do centro s_3. A matriz de transição é dada por

$$\mathbf{P} = \begin{pmatrix} 0 & q & 1 & 0 & 0 \\ 0 & 0 & q & 0 & 0 \\ 1 & p & 0 & q & 0 \\ 0 & 0 & p & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Exemplo 5.4.3. Suponha agora que quando o homem atinge um estado de fronteira ele permanece nesse estado, com probabilidade $\frac{1}{2}$ e se move para o outro estado de fronteira também com probabilidade $\frac{1}{2}$. Nesse caso a matriz de transição é dada por

$$\mathbf{P} = \begin{pmatrix} 0,5 & q & 0 & 0 & 0,5 \\ 0 & 0 & q & 0 & 0 \\ 0 & p & 0 & q & 0 \\ 0 & 0 & p & 0 & p \\ 0,5 & 0 & 0 & p & 0,5 \end{pmatrix}$$

Exemplo 5.4.4. A seguir, consideramos uma versão modificada do passeio aleatório. Se o processo está em um dos três estados interiores,então tem igual probabilidade de se deslocar para a direita, para a esquerda, ou permanecer em seu estado atual. Se for na fronteira, não pode permanecer, entretanto tem igual probabilidade de se deslocar para qualquer um dos outros quatro estados. A matriz de transição é dada por

$$\mathbf{P} = \begin{pmatrix} 0 & 1/3 & 0 & 0 & 1/4 \\ 1/4 & 1/3 & 1/3 & 0 & 1/4 \\ 1/4 & 1/3 & 1/3 & 1/3 & 1/4 \\ 1/4 & 0 & 1/3 & 1/3 & 1/4 \\ 1/4 & 0 & 0 & 1/3 & 0 \end{pmatrix}$$

5.5 Probabilidades de Transição Superiores

O elemento p_{ij} da matriz de transição P de uma cadeia de Markov representa a probabilidade de que o sistema passe do estado s_j para o estado s_i em apenas um estágio: $s_j \to$s$_i$. Agora queremos determinar, a probabilidade $p_{ij}^{(n)}$ de que o sistema mude do estado s_j para o estado s_i em exatamente n estágios.

$$s_j \rightarrow \mathrm{s}_{k_1} \rightarrow \mathrm{s}_{k_2} \rightarrow \cdots \rightarrow \mathrm{s}_{n-1} \rightarrow \mathrm{s}_i$$

Teorema 5.5.1. Seja P a matriz de transição de um processo em cadeia de Markov. Se $p = (p_i)$ é a distribuição de probabilidades do sistema num instante arbitrário, então Pp_i é a distribuição de probabilidades do sistema após um estágio e $P^n p_i$ é a distribuição de probabilidades do sistema após n estágios. Em particular, temos que

$$p^{(1)} = Pp^{(0)}; p^{(2)} = Pp^{(1)}; p^{(3)} = Pp^{(2)}; \cdots ; p^{(n)} = Pp^{(n-1)}$$

Logo, podemos escrever que $p^{(n)} = Pp^{(0)}$.

Demonstração: Suponhamos que o espaço de estados seja $S = \{s_1, s_2, \cdots, s_n\}$. A probabilidade de que o sistema esteja no estado s_j no instante k e passe para o estado s_i no instante $k+1$ é o produto $p_{ij}p_i$, isto é,

$$P\left(X_{k+1} = s_i | X_k = s_j\right) = p_{ij}p_j$$

Dessa forma, a probabilidade de que o sistema esteja no estado s_i no instante $k+1$ é dada por

$$P\left(X_{k+1} = s_i | X_k = s_1\right) + P\left(X_{k+1} = s_i | X_k = s_2\right) + \cdots + P\left(X_{k+1} = s_i | X_k = s_n\right).$$

Ou seja,

$$p_{i1}p_1 + p_{i2}p_2 + \cdots + p_{in}p_n$$

Logo, a distribuição de probabilidades no instante $k+1$ é dada por:

$$p^* = \begin{pmatrix} p_{11}p_1 + p_{12}p_2 + \cdots + p_{1n}p_n \\ p_{21}p_1 + p_{22}p_2 + \cdots + p_{2n}p_n \\ \vdots \\ p_{n1}p_1 + p_{n2}p_2 + \cdots + p_{nn}p_n \end{pmatrix}$$

Por outro lado, isso mostra que o vetor p^* é exatamente o produto da matriz $P = (p_{ij})$ pelo vetor p_i.

Portanto, $p^* = Pp$. Sendo assim tem - se

$$p^{(1)} = Pp^{(0)}; p^{(2)} = Pp^{(1)}; \cdots ; p^{(n)} = Pp^{(n-1)}$$

Agora note que

$$\begin{gathered} p^{(2)} = Pp^{(1)} = P(Pp^{(0)}) = P^2 p^{(0)} \\ p^{(3)} = Pp^{(2)} = P(P^2 p^{(0)}) = P^3 p^{(0)} \\ \vdots \\ p^{(n)} = Pp^{(n-1)} = P(P^{(n-1)} p^{(0)}) = P^n p^{(0)} \end{gathered}$$

como queríamos demonstrar.

Teorema 5.5.2. Teorema das Potências da Matriz de Transição: Seja P a matriz de transição de um processo em cadeia de Markov. Então, a matriz de transição para n estágios é igual a enésima potência de P, em que $p^{(n)} = P^n$.

Demonstração: Suponhamos que o sistema esteja no estado s_j no instante k.

A distribuição de probabilidade do sistema no instante k, uma vez que o sistema está no estado s_j, é o vetor e_j, tal que e_j é o vetor coluna com valor 1 na j-ésima coordenada e zeros nos demais índices. Pelo teorema anterior a distribuição de probabilidades no instante $k + n$ é dada pelo vetor

$$p^{(n+k)} = P^n p^k = P^n e_j.$$

Entretanto, $P^n e_j$ é a j-ésima coluna da matriz P^n. Dessa forma, $p_{ij}^{(n)}$ é a i-ésima componente da j-ésima coluna de P^n. Portanto, $p^{(n)} = P^n$, como queríamos demonstrar.

Suponhamos agora, que em um instante arbitrário, a probabilidade de que o sistema esteja no estado s_i seja p_{ij}. Indicamos estas probabilidades pelo vetor

$$\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$$

o qual determina a distribuição de probabilidades do sistema naquele instante. Em particular, indicamos por

$$\mathbf{p^{(0)}} = \begin{pmatrix} p_1^{(0)} \\ p_2^{(0)} \\ \vdots \\ p_n^{(0)} \end{pmatrix}$$

como sendo a distribuição inicial de probabilidades, isto é, a distribuição de probabilidades que existe quando o processo se inicia, e indicamos por

$$\mathbf{p^{(n)}} = \begin{pmatrix} p_1^{(n)} \\ p_2^{(n)} \\ \vdots \\ p_n^{(n)} \end{pmatrix}$$

como sendo a distribuição de probabilidades após n estágios.

Exemplo 5.5.1. Um homem ou dirige ou toma um ônibus para ir trabalhar todos os dias. Suponhamos que ele nunca tome o ônibus dois dias consecutivos, mas se ele dirige até o trabalho, então no dia seguinte, a probabilidade dele dirigir novamente é a mesma dele tomar o ônibus.

A matriz de transição da cadeia de Markov associado a este processo é:

$$\begin{pmatrix} 0 & 0,5 \\ 1 & 0,5 \end{pmatrix}$$

Nesse caso, o espaço de estados é definido tal que $S = \{s_1 = \hat{o}nibus, s_2 = carro\}$

Então, queremos conhecer, qual a probabilidade de que o sistema passe do estado s_1 para o estado s_2, em exatamente 4 estágios.

Primeiramente, calculamos P^4

$$P^4 = \begin{pmatrix} 0 & 0,5 \\ 1 & 0,5 \end{pmatrix} \begin{pmatrix} 0 & 0,5 \\ 1 & 0,5 \end{pmatrix} \begin{pmatrix} 0 & 0,5 \\ 1 & 0,5 \end{pmatrix} \begin{pmatrix} 0 & 0,5 \\ 1 & 0,5 \end{pmatrix} = \begin{pmatrix} 0,375 & 0,3125 \\ 0,625 & 0,6875 \end{pmatrix}$$

Dessa forma, podemos dizer que a probabilidade do sistema passar do estado s_1 para o estado s_2, em exatamente 4 estágios é dada por $p_{12}^{(4)} = 0,625$.

Suponhamos agora que no primeiro dia o homem tenha lançado um dado honesto

para decidir se iria ao trabalho de carro ou ônibus. Ele definiu que iria dirindo para o trabalho se, e somente se, o resultado do lançamento do dado fosse 6, caso contrário iria de ônibus. Com isso podemos assumir que a distribuição inicial de probabilidades do processo é, $p^{(0)} = (5/6, 1/6)$ é a distribuição inicial de probabilidades do sistema.
Logo, a distribuição de probabilidades após 4 dias é dada por:

$$p^{(4)} = P^4 p^{(0)} = \begin{pmatrix} 0,375 & 0,3125 \\ 0,625 & 0,6875 \end{pmatrix} \begin{pmatrix} 5/6 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 35/96 \\ 61/96 \end{pmatrix}$$

Portanto, a distribuição de probabilidades após 4 dias é dada por:

$$\mathbf{p^{(4)}} = \begin{pmatrix} 35/96 \\ 61/96 \end{pmatrix}$$

Em outras palavras, após espera-se que o homem tenha ido ao trabalho de ônibus em 36,4% das vezes e tenha ido de carro em 63,6% das vezes.

5.6 Cadeias de Markov Regulares

5.6.1 Introdução

Suponhamos que em uma determinada região, observa-se que se chover bastante durante o ano, a probabilidade de que chova bastante no ano seguinte é de 25%, e que a probabilidade de que se tenha uma escassez de chuva é de 75%. Ainda sabemos que, se houver uma escassez de chuvas em um ano, no ano seguinte a probabilidade de haver bastante chuva ou uma seca, será a mesma. Suponhamos também, para simplificar o modelo, que estas probabilidades permaneçam inalteradas com o decorrer dos anos, o que não ocorre na prática, embora possamos usar essa simplificação, como recurso para termos um indicador da situação.

A longo prazo, o que podemos esperar com maior probabilidade. Que essa região sofra com períodos de muita chuva? Ou com períodos com escassez de chuvas?

As cadeias de Markov mostram-se eficazes na resolução desse tipo de problema. Sendo assim, muito importantes na avaliação de propostas de intervenção na realidade utilizando conhecimentos algébricos. Esse capítulo dedica-se na análise do comportamento estacionário da distribuição de probabilidade de uma cadeia de Markov regular, essencial na aplicação.

As cadeias de Markov regulares, exercem papel fundamental nas aplicações, em virtude de sua distribuição estacionária de probabilidades, sobretudo, pois sua convergência para sua distribuição é obtida por meio de operações matriciais elementares aplicadas na matriz de transição as quais são facilmente programadas e esta é uma importante particularidade nas aplicações das cadeias regulares, o que propcia apresentar o tema em turmas de Ensino Médio. A contemporalidade do tema e suas diversas aplicações corroboram para o desenvolvimento do pensamento crítico e investigativo dos alunos, potencializando a aprendizagem de temas como probabilidades e matrizes.

5.6.2 Distribuição Estacionária de uma cadeia de Markov Regular

Definição: Uma cadeia de Markov é dita regular se, e somente se, a partir de um estado inicial arbitrário é possível acessar qualquer estado após um certo número de passos, em outras palavras, se sua matriz de transição P for uma matriz estocática regular.

Lema 5.6.2: Seja P uma matriz de transição de ordem r com todas as entradas são positivas. Seja ϵ a menor entrada da matriz P. Seja x um vetor coluna com r coordenadas, sendo M_0 sua maior coordenada e m_0 sua menor coordenada, e sejam M_1 e m_1 respectivamente a maior e a menor coordenada do vetor Px. Então

$$M_1 \leq M_0, m_0 \leq m_1 \qquad\qquad M_1 - m_1 \leq (1 - 2\epsilon)(M_0 - m_0).$$

Demonstração: Seja x^* um vetor obtido substituindo todas as componentes do vetor x por M_0 exceto a componente m_0, então, $x_i \leq x_i^*$. Cada componente de Px^* é da forma

$$Px_i^* = a \cdot m_0 + (1 - a) \cdot M_0 = M_0 - a(M_0 - m_0)$$

em que $a \geq \epsilon$. Assim, $Px_i^* \leq M_0 - \epsilon(M_0 - m_0)$. Entretanto, como $x_i \leq x_i^*$, temos

$$M_1 \leq M_0 - \epsilon(M_0 - m_0).$$

Se aplicarmos esse resultado no vetor $-x$, obtemos

$$-m_1 \leq -m_0 - \epsilon(-m_0 + M_0)$$

Segue dos resultados anteriores que

$$M_1 - m_1 \leq M_0 - m_0 - 2\epsilon(M_0 - m_0) = (1-2\epsilon)(M_0 - m_0) \Longrightarrow M_1 - m_1 \leq (1-2\epsilon)(M_0 - m_0)$$

Teorema Fundamental da Convergência das Cadeias de Markov Regulares: Seja P uma matriz de transição regular. Então, as potências P^n se aproximam da matriz estocástica A, tal que todas as linhas são iguais ao ponto fixo w da matriz P. Além disso, cada coluna de A é o mesmo vetor de probabilidades w, isto é, $A = w\xi$, em que ξ é o vetor linha com todas as componentes iguais a 1.

Demonstração: Primeiro vamos assumir que P não tem entrada zero e seja ϵ sua menor entrada. Seja ρ_j um vetor coluna com 1 na j-ésima coordenada e 0 nas demais. Sejam M_n e m_n o maior e o menor valores dentre as coordenadas do vetor $P^n \rho_j$. Desde que $P^n \rho_j = PP^{n-1}\rho_j$ segue do lema anterior. que $M_1 \geq M_2 \geq M_3 \geq \cdots$ e $m_1 \leq m_2 \leq m_3 \leq \cdots$ então

$$M_n - m_n \leq (1-2\epsilon)(M_{n-1} - m_{n-1})$$

$\forall n \geq 1$.Se tomarmos $d_n = M_n - m_n$, isso nos diz que

$$d_n \leq (1-2\epsilon)^n d_0 = (1-2\epsilon)^n$$

Assim, quando n tende ao infinito d_n tende a zero, M_n e m_n aproximam de um limite comum, e portanto $P^n p_j$ tende para o vetor com todas as componentes iguais. Seja w_j esse valor comum. Fica claro que, $\forall n$, $m_n \leq w_j \leq M_n$. Em particular, desde que $0 < m_1$ e $M_1 < 1$, temos que $0 < w_j < 1$. Agora $P^n \rho_j$ é a j-ésima coluna de P^n. Assim, a j-ésima coluna de P^n tende para o vetor com todas componentes iguais ao valor w_j, isto é, P^n tende para a matriz A em que todas as linhas são iguais ao vetor $w = (w_1, \cdots, w_r)$. Desde que as somas das componentes dos vetores linhas da matriz P^n sejam sempre iguais a 1, o mesmo deve ser verdade para o limite. Isso completa a demonstração para o caso me que todas as entradas da matriz são positivas.

Consideramos agora que P seja regular. Seja N tal que, cada potência P^N não tenha entradas zero. Tome ϵ^* a menor das entradas de P^N. Aplicando a primeira parte da demonstração para a matriz P^N, temos

$$d_{kN} \leq (1 - 2\epsilon^*)^k$$

Portanto, a sequência d_n, que não é crescente, possui uma subsquência tendendo a zero. Assim d_n tende a zero e o resto da prova é analoga a prova para a matriz com todas as entradas positivas.

Portanto, a sequência P, P^2, P^3,$\cdots$,P^n, $\cdots$ das potências de P tende à matriz cujas as colunas são todas iguais ao vetor de probabilidade fixo w, como queríamos demosntrar.

Esse teorema assegura a existência de um método simples para determinar o vetor de probabilidade fixo de uma matriz de transição de uma cadeia de Markov.

Para obtermos o vetor de probabilidade fixo basta tomarmos uma potência P^n com n suficientemente grande.

O vetor de probabilidade fixo w será denominado vetor da distribuição estacionária da cadeia de Markov regular determinada pela matriz de transição P.

Suponhamos que uma cadeia de Markov seja regular, isto é, que sua matriz de transição P seja regular. Pelo teorema da convergência, a sequência das matrizes de transição para n estágios, P^n, tende a matriz T, cujas linhas são todas iguais ao unico vetor de probabilidade fixo w de P. Logo, a probabilidade $w_{ij}^{(n)}$ de que o estado j ocorra, para n suficientemente grande, é independente do estado inicial i e tende a componente w_i de w. Em outras palavras, a longo prazo, a probabilidade de que qualquer estado j ocorra é aproximadamente igual a componente w_i do único vetor de probabilidade fixo w de P.

Dessa forma, observamos que a influência do estado inicial ou da distribuição inicial de probabilidades do processo diminui a medida em que o número de estágios do processo aumenta. Além disso, toda sequência de distribuição de probabilidades tende ao vetor de probabilidade fixo de P, o qual representa a distribuição estacionária de uma cadeia de Markov.

Essa característica permite com que as cadeia de Markov Regulares, sejam um ótimo modelo de modelagem matemática para a análise de previsões a longo prazo.

Exemplo 5.6.1. Suponhamos que em uma determinada região, observa-se que se chover bastante durante o ano, a probabilidade de que chova bastante no ano seguinte é de $\frac{1}{4}$, e que a probabilidade de que se tenha uma escassez de chuva é de $\frac{3}{4}$. Ainda sabemos que, se houver uma escassez de chuvas em um ano, no ano seguinte a probabilidade de haver bastante chuva ou uma seca, será a mesma. Suponhamos também, para simplificar o modelo, que estas probabilidades permaneçam inalteradas com o decorrer dos anos, o que não ocorre na prática, embora possamos usar essa simplificação, como recurso para termos um indicador da situação. Sendo assim, a matriz de transição desse processo em cadeia de Markov é dada por

$$M = \begin{pmatrix} 0,25 & 0,5 \\ 0,75 & 0,5 \end{pmatrix}$$

Note que M é uma matriz estocástica regular, uam vez que todos os seus elementos são positivos. Sendo assim, podemos concluir que, quaisquer que sejam as probabilidades iniciais, a distribuição de probabilidades a longo prazo é dada pelo vetor de probabilidades fixo da matriz M.

$$Mw = w$$

$$\begin{pmatrix} 0,25 & 0,5 \\ 0,75 & 0,5 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} w_{chuva} \\ w_{seca} \end{pmatrix}$$

Da equação acima segue o sistema

$$\begin{cases} 0,25x + 0,5 - 0,5x = x \\ \\ 0,75x + 0,5 - 0,5x = 1 - x \end{cases}$$

Então, devemos ter

$$0,25x + 0,5 - 0,5x = x$$
$$x + 2 - 2x = 4x$$

Da equação acima segue que, x = $\frac{2}{5} = 0,4$ e 1 - x = $\frac{3}{5} = 0,6$.

$$\begin{pmatrix} w_{chuva} \\ w_{seca} \end{pmatrix} = \begin{pmatrix} 0,4 \\ 0,6 \end{pmatrix}$$

Portanto, a longo prazo, a probabilidade de termos nessa região um ano ocm bastante chuva é de 40%, enquanto que a probabilidade de termos um ano com escassez de chuvas é de 60%, dentro das hipóteses simplificadoras. Com base nesses a dados podemos concluir, que a região tenderá a uma ligeira aridez.

Exemplo 5.6.2 Suponhamos que em um município, a cada ano 3% da população da zona rural, migra para a zona urbana, enquanto 1% da população da zona urbana migra para a zona rural. Se todas essas porcentagens não mudarem, qual deve ser a relação entre as populações urbana e rural desse município a longo prazo?

A Matriz de transição desse processo markoviano regular é dada por

$$\begin{pmatrix} 0.99 & 0.03 \\ 0.01 & 0.97 \end{pmatrix}$$

Como a a matriz é regular, então, a longo prazo as probabilidades $w_{rural} = 1 - x$ e $w_{urbana} = x$, devem satisfazer a equação

$$\begin{pmatrix} 0,99 & 0,03 \\ 0,01 & 0,97 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix}$$

Da equação acima segue que

$$\begin{cases} 0,99x + 0,03 - 0,03x = x \\ \\ 0,01x + 0,97 - 0,97x = 1 - x \end{cases}$$

Daí segue que

$$\begin{aligned} &0,99x + 0,03 - 0,03x = x \\ &0,04x = 0,03 \end{aligned}$$

Então temos que x = 0,75. Logo, w_{rural} = 25% e w_{urbana} = 75%. Portanto, se não houver modificações nas politicas públicas de migração daquela região, teremos a longo prazo, 25% da população vivendo na zona rural do município e 75% da população vivendo na zona urbana.

Exemplo 5.6.3 Observa-se experiemntalmente que, em condições naturais e sem ser submetida à pesca industrial, a quantidade de uma certa espécie de peixe varia do seguinte modo: se em um ano a população diminui, a probabilidade de que diminua ainda mais no ano seguinte é de 60% e, se em um determinado ano a população aumenta, a probabilidade de que diminua no ano seguinte é de 30%. Entretanto, observa-se que, sendo essa mesma espécie de peixe submetida à pesca industrial, quando a população de peixes aumenta num determinado ano, a probabilidade de que dimiua no ano seguinte se altera para 50%, enquanto que se a população diminua no ano seguinte continua sendo 60%. Deseja-se conhecer, como a longo prazo, a pesca industrial estará afetando a população de peixes dessa espeécie. Desse modo, seria possível determinar, se a pesca industrial deve diminuir para preservar a espécie ou até mesmo para determinar que essa atividade comercial tem potencial para expansão. Os estado desse processo são: diminuição da população de peixes (D) e aumento da população de peixes (A). Então, vamos analisar as probabilidades de evolução populacional dessa espécie de peixe, sem haver pesca industrial, a matriz de probabilidades de transição desse processo markoviano é dada por

$$\begin{pmatrix} 0,7 & 0,4 \\ 0,3 & 0,6 \end{pmatrix}$$

Como a matriz é regular, as probabilidades $w_D = 1 - x$ da população diminuir e $w_A = x$ da população aumentar a longo prazo são tais que

$$\begin{pmatrix} 0,7 & 0,4 \\ 0,3 & 0,6 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix}$$

Então,

$$\begin{cases} 0,7x + 0,4 - 0,4x = x \\ 0,3x - 0,6 - 0,6x = 1 - x \end{cases}$$

Daí segue que

$$\begin{aligned} &0,7x + 0,4 - 0,4x = x \\ &0,7x = 0,4 \\ &x = \tfrac{4}{7} \Longrightarrow 1 - x = \tfrac{3}{7} \end{aligned}$$

Portanto, a probabilidade da população dessa espécie de peixes aumentar é $w_A = \frac{4}{7}$ e a probabilidade da população de peixes diminuir é $w_D = \frac{3}{7}$. Podemos concluir que, em condições naturais, a espécie tem a sobrevivência razoavelmente garantida.

Exemplo 5.6.3. Agora vamos analisar como a longo prazo a pesca industrial efeta a população de peixes. Com esse novo cenário, a matris de transição se altera para

$$\begin{pmatrix} 0,5 & 0,4 \\ 0,5 & 0,6 \end{pmatrix}$$

Como a matriz é regular, a longo prazo $w_A =$ x e $w_D =$ 1 - x, são tais que

$$\begin{pmatrix} 0,5 & 0,4 \\ 0,5 & 0,6 \end{pmatrix} \begin{pmatrix} x \\ 1-x \end{pmatrix} = \begin{pmatrix} x \\ 1-x \end{pmatrix}$$

Da equação acima segue que

$$\begin{cases} 0,5x + 0,4 - 0,4x = x \\ 0,5x + 0,6 - 0,6x = 1 - x \end{cases}$$

Do sistema acima segue que

$$\begin{aligned} &0,5x + 0,4 - 0,4x = x \\ &0,9x = 0,4 \\ &x = \tfrac{4}{9} \Longrightarrow 1 - x = \tfrac{5}{9} \end{aligned}$$

Desse modo temos, $w_A = \frac{4}{9}$ e $w_D = \frac{5}{9}$. Como a probabilidade da população diminuir é maior, se a espécie for submetida à pesca industrial, sua sobrevivência será ameaçada e, portanto a pesca deve ser diminuída ou até mesmo proibida.
Podemos observar que os processos markovianos, podem ser modelados e estudados, afim de estabelecermos cenários a longo prazo, os quais podem oferecer informações úteis para tomadas de decisão como por exemplo, na gestão de recursos naturais e na elaboração de políticas públicas.

5.7 Cadeias de Markov Absorventes

Um estado i de uma cadeia de Markov é denominado absorvente se

$$p_{ii} = 1.$$

Uma cadeia de Markov é denominada absorvente se existe pelo menos um estado absorvente e se for possível, a partir de qualquer estado arbitrário, existir uma sequência de transições até um estado absorvente. Um estado que não é absorvente é demoninado estado de transição.

Exemplo 5.7.1. Suponhamos uma cadeia de Markov em que sua matriz de transição é dada por

$$\begin{array}{cc} \begin{matrix} \mathtt{A} & B & C & D & E \end{matrix} & \\ \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} \end{array}$$

Note que os estados $s_2 = B$ e $s_5 = E$ são ambos absorventes, pois tanto a segunda coluna como a quinta coluna possuem 1 na diagonal principal e 0 nas demais posições. Assim, um estado i é absorvente se, e somente se, a i-ésima linha da matriz de transição P possui a componente na diagonal principal igual a 1, e zeros nas demais posições.

Exemplo 5.7.2. Um homem lança uma moeda honesta até que ocorram 3 caras consecutivas. Seja $X_n = k$ se, na enésima tentativa, a última coroa tenha ocorrido na $(n-k)$ ésima tentativa, isto é, X_n indica a maior sequência de caras que termina na enésima tentativa. Esse é um processo em cadeia de Markov, cujo espaço de estados é $S = \{s_0, s_1, s_2, s_3\}$, em que s_i significa que a sequência de caras tem comprimento i. A matriz de transição é

$$
\begin{array}{cc}
\begin{array}{cccc} \mathtt{A} & B & C & D \end{array} & \\
\begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 1
\end{pmatrix} & \begin{array}{c} A \\ B \\ C \\ D \end{array}
\end{array}
$$

Qualquer coluna exceto a última, corresponde ao fato de que uma sequência de caras é interrompida, se ocorrer coroa ou então, é ampliada se ocorrer cara. A última coluna corresponde ao fato de que o jogo termina se ocorrerem três caras em sequência. Note que D é um estado absorvente.

Teorema 5.7.1. Se uma matriz estocástica de ordem $n \geq 2$ tem 1 na diagonal principal, então P não é regular. Podemos dizer em outras palavras, que se uma cadeia de Markov possui pelo menos um estado absorvente, então a cadeia não é regular.

Demonstração: Seja s_j um estado absorvente de uma cadeia de Markov com matriz de transição P. Então para i, a probabilidade de transição após n estágios é $p_{ij}^{(n)} = 0$, $\forall n$. Portanto, qualquer potência de P tem um elemento nulo. Logo, P não é regular.

6 Como o Google googla?

6.1 Introdução

O Google tornou-se um dos principais buscadores utilizados pelos usuários de internet em todo o mundo. Grande parte desse sucesso deve-se a matemática aplicada em seu algoritmo de busca. O Google foi criado com a missão de organizar a informação mundial e torná-la universalmente acessível.

Neste capítulo vamos mostrar como um simples gesto de pesquisar um assunto no Google tem por trás muita matemática, mais precisamente, cadeias de Markov.

Numa época em que a internet se transformou em um gigantesco acervo de informações e conhecimento, motores de busca como o Google se transformaram em uma espécie de sonar que leva o pescador a encontrar um bom cardume. Quando pesquisamos um tema, o Google aplica um algoritmo que, em microssegundos, apresenta uma excelente ordenação para encontrarmos o que procuramos. Resumidamende, ao perquisarmos um assunto no Google, ele envia para cerca de 1 milhão de servidores ao redor do mundo e estes por sua vez processam mais de 1 bilhão de pesquisas gerando aproximadamente 20 petabytes de dados o equivalente a 22.517.998.136.852.500 bytes.

Suponhamos que uma pessoa deseja preparar uma deliciosa carne assada para o jantar e vai a procura de uma receita. Nessa altura a pessoa percebe que existem 451000 páginas relacionadas com o tema. Certamente muitas dessas páginas não são interessantes e mesmo assim seria impossível ler todas antes do jantar, mas como num passe de "mágica"normalmente encontramos o que queremos nas primeiras sugestões do Google. Mas como o Google sabe quais são as páginas mais interessantes sobre carne assada a ponto de apresentá - las em primeiro lugar? Mágica ou Matemática?

Os primeiros buscadores desenvolvidos limitavam-se a contar o número de vezes que a expressão carne assada ocorria em cada página apresentando a página em que a expressão ocorria com maior frenquência em primeiro lugar. Entretanto este não é o melhor critério, pois um site que repete a expressão carne assada não significa que tem a melhor receita até mesmo porque pode ser uma página em que o autor repete diversas vezes que destesta carne assada. Para se esquivar desse problema, no final da década de 90, o Google apresentou um novo método de classificar as páginas na internet com base na interação existente entre elas na Web.

Se existem muitos links para uma determinada página que contém uma receita para carne assada então essa deve ser uma página com uma receita saborosa. Segundo o algoritmo utilizado pelo Google, a importância de uma página depende essencialmente da importância dos sites que possuem link para ela.

O algoritmo aplicado pelo Google utiliza a estrutura de links da Web para produzir um ranking de importância para as páginas da Web. Esse algoritmo é chamado de PageRank, ele ajuda ao Google a qualificar a grande heterogeneidade das páginas da internet.

Desenvolvedores de websites se dedicam intensamente para conseguir uma classificação no topo das pesquisas em motores de busca como o Google. Com o uso da criptografia como critério para classificação das páginas, muitos sites tornarão suas páginas mais seguras para os usuários. Entretanto, a segurança de um site terá um peso menor na classificação em relação a outros fatores, mas sua importância poderá aumentar com o tempo.

Atualmente, há outros fatores adicionais que influênciam na ordenação dos sites. Entretanto, a essência do ranqueamento para qualquer usuário da ferramenta de busca é a mesma, o PageRank.

6.2 A Origem do Google

Em 1995, dois jovens estudantes de ciências da computação, Larry Page com 22 anos e Sergey Brin com 21, criaram um mecanismo de pesquisa inicialmente chamado de BackRub, que chegou a ser utlizado em servidores da universidade de Stanford durante mais de um ano, e acabacou em desuso por usar uma largura de banda excessiva para os padrões que a universidade tinha disponível naquela época.
A partir daí, eles decidiram que o mecanismo de busca por eles criado, precisava de um outro nome. Após diversas sugestões, optaram por escolher Google, um brincadeira com a palavra "googol", termo matemático para designar o número representado pelo dígito 1 seguido de cem dígitos 0.

$$10^{100}$$

Andy Bechtolsheim, doou um cheque no valor 100 mil dólares para uma entidade que ainda inexistente: uma empresa chamada Google Inc. Improvisado em uma garagem no endereço 232 Santa margarida, Menlo Park, o Google inicia seus trabalhos. A formalização da criação da empresa na Califórnia ocorreu em 4 de setembro de 1998. Em seguida, Larry e Sergey abrem uma conta em um banco em nome da nova empresa e depositam o cheque doado por Andy Bechtolsheim. Com o crescimento da empresa foi necessário abandonar o pequeno escritório na garagem, e transferir-se

para um novo endereço, mais amplo e que atendesse a nova demanda da empresa e por isso em fevereiro do ano de 1999, o Google inaugura sua nova sede no endereço 165 University Avenue em Palo Alto, com apenas oito funcionários.
A partir daí, o Google teve uma ascensão meteórica e hoje é o site mais popular de pesquisa na internet e uma das maiores empresas de todo o mundo.

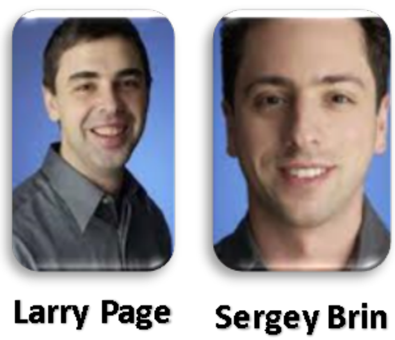

Figura 6.1: Larry Page e Sergey Brin

6.3 A Web e as cadeias de Markov

6.3.1 Introdução

Entender o mecanismo de funcionamento do algoritmo PageRank, é essencial para qualquer desenvolvedor de sites que deseja ter sua página acessada com frequência, uma vez que, aparecer listado no topo de uma pesquisa no Google provoca muitas vizualizações. Na verdade, em função da proeminência do Google como um motor de busca, o seu método de classificação provocou uma forte influência sobre o

desenvolvimento e estrutura da internet, e principalmente, alterou o formato das informações e serviços oferecidos via a rede mundial de computadores.
Ferramentas de pesquisa como o Google se propoem a realizar três ações básicas:

1. rastrear a web e localizar todas as páginas com acesso público;
2. indexar os dados a partir da primeira ação, de modo a serem pesquisados de forma eficiente por meio de palavra-chave ou experssões relevantes;
3. calcular a taxa de importância de cada página do banco de dados, de modo que, ao realizarmos uma busca dentro do subconjunto de páginas na base de dados com a informação desejada, as páginas mais importantes sejam apresentadas no topo da lista,isto é, com maior prioridade.

O valor da importância das páginas da web não é o único fator levado em consideração, mas é o critério mais significativo.

6.3.2 Web Fortemente Conectada

Definição: Uma rede é dita admissível se, e somente se, todos os sites possuem pelos menos uma link de saída.
Definição: Uma rede é dita fortemente conectada se é possível passar de uma página para outra qualquer apenas clicando nos links.
Exemplo 6.3.2.1. A princípio consideremos um caso simples para ilustrar a importância dos sites de uma determinada rede da internet, representada pelo grafo

a seguir:

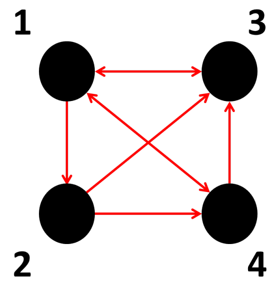

Figura 6.2: Web Admissivel

No exemplo acima, a rede é composta por 4 sites. Cada flecha orientada representa um link existente de um site para o outro. Trata-se de uma Web admissível, isto é, cada página possui pelo menos um link de saída.

Considere que um usuário, estando no site 1 tem igual probabilidade de acessar os demais sites com os quais ele possui link de saída, ou seja x_i o índice de importância do site i, com $x_i \geq 0$ para qualquer página i.

Inicialmente poderiamos encontrar a importância do site 4, somando as importâncias dos sites 1 e 2, ou seja, $x_4 = x_1 + x_2$. Entretanto, podemos observar que o site 1 possui link de saída para os sites 2,3 e 4.

Desse modo, a importância da página 1 deve ser dividida igualmente em três.

Da mesma forma, como a página 2 possui link de saída para as páginas 3 e 4 sua importância deve ser dividida em duas partes iguais.

Portanto, $x_4 = \frac{1}{3}x_1 + \frac{1}{2}x_2$.

As importâncias dos demais sites seguem equações análogas. Dessa forma, obtemos a seguinte matriz de transição.

$$P = \begin{pmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{pmatrix}$$

Observe que P não representa uma cadeia de Markov absorvente, e

$$P^2 = \begin{pmatrix} 1/2 & 3/4 & 0 & 1/2 \\ 0 & 0 & 1/3 & 1/6 \\ 1/3 & 1/4 & 1/3 & 1/6 \\ 1/6 & 0 & 1/3 & 1/6 \end{pmatrix}$$

$$P^{16} = \begin{pmatrix} 0,3872645672 & 0,3872851923 & 0,3871504167 & 0,3872007577 \\ 0,1277596375 & 0,1277831474 & 0,1277928625 & 0,1278153643 \\ 0,28944822217 & 0,28941951440 & 0,28947241393 & 0,28943110624 \\ 0,19552757300 & 0,19551214576 & 0,19558430667 & 0,19555277158 \end{pmatrix}$$

$$P^{32} = \begin{pmatrix} 0,3872216851 & 0,3872216899 & 0,3872216784 & 0,3872216866 \\ 0,12778315422 & 0,12778315296 & 0,12778315897 & 0,12778315637 \\ 0,2894482101 & 0,2894482090 & 0,2894482087 & 0,2894482076 \\ 0,19554695050 & 0,19554694812 & 0,19554695279 & 0,19554694930 \end{pmatrix}$$

$$P^{1024} = \begin{pmatrix} 0,3872216844 & 0,3872216844 & 0,3872216844 & 0,3872216844 \\ 0,12778315585 & 0,12778315585 & 0,12778315585 & 0,12778315585 \\ 0,2894482090 & 0,2894482090 & 0,2894482090 & 0,2894482090 \\ 0,19554695062 & 0,19554695062 & 0,19554695062 & 0,19554695062 \end{pmatrix}$$

Sabemos que P é uma matriz de transição de uma cadeia de Markov regular. Note que as componentes dos vetores linhas de probabilidades da matriz P^n para valores de n suficientemente grandes tornam-se constantes, ou seja, essa cadeia de Markov possui uma distribuição estacionária.

Dessa forma podemos concluir que a probabilidade de que um usuário acesse a página 1 após 1024 entradas ou sáidas entre os links dessa rede é igual a 38,72%, que é também a probabilidade de que se acesse o site 1 depois de 1024 cliques de links saindo da página 2,3 ou 4. Portanto, a probabilidade de um usuário acessar a página 1 após 1024 entradas ou saídas entre esses links é 38,72%, enquanto que para as páginas 2,3 e 4 as probabilidades são respectivamente 12,77%,28,94% e 19,55%, isto é, o site 2 é o menos acessado nessa rede de links. Por outro lado, a página 1 é a mais popular, isto é, aquela que possui uma maior frequência de vizualições e acessos.

Para análisar a convergência de uma cadeia de Markov regular para uma distribuição estacionária de modo prático basta encontrar um autovetor da matriz de transição P associado ao autovalor 1 e nesse caso temos:

$$Pw = w$$

$$\begin{pmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

Desse modo a equação anterior envolvendo autovetor e autovalor, pode ser conduzida a um problema de sistema linear homogêneo dado por:

$$\begin{cases} w_1 - w_3 - \frac{w_4}{2} = 0 \\ \\ \frac{w_1}{3} - w_2 = 0 \\ \\ \frac{w_1}{3} + \frac{w_2}{2} - w_3 + \frac{w_4}{2} = 0 \\ \\ \frac{w_1}{3} + \frac{w_2}{2} - w_4 = 0 \\ \end{cases}$$

Resolvendo o sistema, sua solução será

$$\mathbf{w} = t \begin{pmatrix} 12 \\ 4 \\ 9 \\ 6 \end{pmatrix}$$

tal que $t \in \mathbb{R}$.

Sendo assim, o sistema possui infinitas soluções, mas o que é relevante é o ranqueamento das importâncias, isto é, sua ordenação. E para valores positivos de t a ordenação será sempre a mesma, por isso, tomamos t de modo que $w_1 + w_2 + w_3 + w_4 = 1$. Nesse caso temos:

$$w_1 = \frac{12}{31} \quad w_2 = \frac{4}{31} \quad w_3 = \frac{9}{31} \quad w_4 = \frac{6}{31}$$

Portanto, o vetor de probabilidade

$$\mathbf{w} = \begin{pmatrix} \frac{12}{31} \\ \frac{4}{31} \\ \frac{9}{31} \\ \frac{6}{31} \end{pmatrix}$$

é a distribuição estacionária de probabilidades da cadeia de Markov representada pela matriz de transição P e também um ponto fixo de P.

Exemplo 6.3.2.2. Considere uma web com 3 sites, mas agora de forma que nesses sites existam autolinks, isto é, links que mantém o usuário no mesmo site, e em alguns casos a probabilidade de permanecer no site é maior que a de saída. Dizemos que essa rede é fortemente conectada, pois podemos passar de um site qualquer para outro site arbitário apenas clicando nos links.

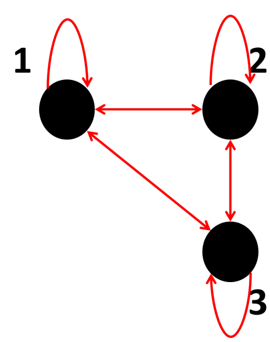

Figura 6.3: Web

Dessa forma a matriz de transição P é dada por

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

Desse modo para analisarmos a distribuição estacionária dessa cadeia de Markov descrita pela matriz de transição P devemos encontrar o w de P, isto é, devemos resolver a equação

$$Pw = w$$

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Portanto,

$$\mathbf{w} = \begin{pmatrix} x \\ \frac{3}{4}x \\ \frac{5}{4}x \end{pmatrix}.$$

Entretanto devemos ter $x_1 = \frac{1}{3}$, daí segue que

$$\mathbf{w} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{4} \\ \frac{5}{12} \end{pmatrix}.$$

6.3.3 Web não Fortemente Conectada

Na internet muitas páginas são criadas a todo momento e algumas dessas páginas podem conter apenas um autolink, isto é, um único link para si mesmo. Essas páginas trazem problemas para a aplicação do algoritmo PangRank, pois nesse caso, a cadeia é absorvente e não regular. Sendo assim, o teorema da convergência não se aplica. Para esses casos é possível fazer uma modificação na matriz original de links P substituindo-a, por uma matriz M, em que M é uma soma afim de P e A, tal que

$$M = (1 - m)P + mA$$

em que $0 < m < 1$ e $A = a_{ij} = \frac{1}{n}$ $\forall i, j$, sendo A uma matriz quadrada de ordem n. O valor utilizado pelo Google é $m = 0,15$. Observe que, quanto menor o valor de m, mais peso se atribui a matriz original de links P e menos peso é atribuido a matriz A. Nesse sentido, dizemos que a matriz A é uma matriz neutra, pois ela distribui sua importância igualmente entre todos os links da web.

Exemplo 6.3.3.1. Considere uma processo em cadeia de Markov representado pela seguinte matriz de transição

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 1 \end{pmatrix}$$

em que a primeira coluna representa as probabilidades de transição do site A para os demais, a segunda coluna representa as probabilidades de transição do site B para os demais sites, a terceira coluna representa as probabilidades de transição do site C para os demais sites e assim por diante até a sétima coluna representando as probabilidades de transição do site G para os demais sites.

Note que P é absorvente, mas não é regular sendo assim, vamos aplicar a substituição

$$M = (1 - m)P + mA$$

em que $m = 0,15$ e

$$A = \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{pmatrix}$$

Então temos:

$$M = (1 - 0,15)P + 0,15A$$

$$M = 0,85 \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 1 \end{pmatrix} + 0,15 \begin{pmatrix} \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{pmatrix}$$

Essa substituição distribui igualmente a probabilidade de $\frac{1}{7}$ para um site qualquer passar para outro site arbitrário.

$$M = \begin{pmatrix} 0,021 & 0,446 & 0,021 & 0,021 & 0,021 & 0,021 & 0,021 \\ 0,021 & 0,021 & 0,304 & 0,021 & 0,446 & 0,021 & 0,021 \\ 0,871 & 0,021 & 0,021 & 0,021 & 0,021 & 0,304 & 0,021 \\ 0,021 & 0,021 & 0,304 & 0,871 & 0,021 & 0,021 & 0,021 \\ 0,021 & 0,446 & 0,021 & 0,021 & 0,021 & 0,304 & 0,021 \\ 0,021 & 0,021 & 0,304 & 0,021 & 0,446 & 0,021 & 0,021 \\ 0,021 & 0,021 & 0,021 & 0,021 & 0,021 & 0,304 & 0,871 \end{pmatrix}$$

Calculando a potência M^{1000} temos:

$$P^{1000} = \begin{pmatrix} 0,055 & 0,055 & 0,055 & 0,055 & 0,055 & 0,055 & 0,055 \\ 0,080 & 0,080 & 0,080 & 0,080 & 0,080 & 0,080 & 0,080 \\ 0,091 & 0,091 & 0,091 & 0,091 & 0,091 & 0,091 & 0,091 \\ 0,316 & 0,316 & 0,316 & 0,316 & 0,316 & 0,316 & 0,316 \\ 0,078 & 0,078 & 0,078 & 0,078 & 0,078 & 0,078 & 0,078 \\ 0,080 & 0,080 & 0,080 & 0,080 & 0,080 & 0,080 & 0,080 \\ 0,295 & 0,295 & 0,295 & 0,295 & 0,295 & 0,295 & 0,295 \end{pmatrix}$$

Sendo assim, a distribuição estacionária da cadeia de Markov representada pela matriz de transição P é dada pelo vetor

$$w = \begin{pmatrix} 0,055 \\ 0,080 \\ 0,091 \\ 0,316 \\ 0,078 \\ 0,080 \\ 0,295 \end{pmatrix}$$

Portanto, o site acessado é o C com popularidade de 31,6% entre os usuários dessa pequena web.

Bibliografia

[1] KEMENY, G.J.; SNELL, L.J. ***Finite Markov Chains***. New York: Springer - Verlag, 1976.

[2] LISPCHUTZ, S. ***Matemática Finita***. São Paulo: McGraw - Hill do Brasil, 1966.

[3] MORGADO, A.C.; CARVALHO, J.B.P.; CARVALHO, P.C.P.; FERNANDEZ, P. ***Análise Combinatória e Probabilidade***. Sétima Edição. Rio de Janeiro: SBM, 1991.

[4] BOLDRINI, J. L.; COSTA, S.I.R.; RIBEIRO, V.L.F.F.; WELTZER, H.G. ***Álgebra Linear***. Segunda Edição. São Paulo: Harbra, 1980.

[5] SÁ, C.C; ROCHA, J. ***Treze Viagens pelo Mundo da Matemática***. Segunda Edição. Rio de Janeiro: SBM, 2012.

[6] CRILLY, T. ***50 cosas que hay que saber sobre Matemáticas***. Quinta Edição. Buenos Aires: Ariel, 2011.

[7] BRYAN, K; lEISE, T. ***The $ 25,000,000,000 eigenvector: the Linear Algebra behind Google.*** SIAM Review, 2006. Vol. 48, n° 3,569. Disponível em: http:// www.rose-hulman.edu/bryan/googleFinalVersionFixed.pdf.

[8] RPM, Rio de Janeiro, n° 80, p. 42 - 45, 1° quadrimestre de 2013.

[9] NORRIS, J.R. ***Markov Chains***. First Edition. New York: Cambridge University Press, 1997.

[10] MELO, M.P. ***Ordenação das páginas do Google - Page Rank***. 2009. Dissertação. (Mestrado em Ciências) - Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, 2009.

[11] ISTO é Matemática - Como é que o Google googla. Portugual: Sigma 8, 22 de outubro de 2012. Disponível em: clubes.obmep.org.br/blog/video-4/.

[12] CONNOR, J.J.; ROBERTSON, E.F. **MacTutor History of Mathematics**. Scotland: 2006. Dispanível em:http:www-history.mcs.st-andrews.ac.uk/Biographies/Markov.html. Acesso em: 08 ago. 2015.

[13] Tartaruga - de - Couro Gigante, Disponível em http: www.tartarugas.avph.com.br. Acesso em 10 de outubro de 2015.