PROCESSOS ESTOCÁSTICOS

pedro j. fernandez

Esta monografia é uma versão eletrônica do minicurso ministrado em Julho de 1997 durante a realização do 10º Colóquio Brasileiro de Matemática em Poços de Caldas, MG, Brasil

Pedro J. Fernandez
I M P A
Outubro 2014

O propósito desta monografia é modesto: cobrir integralmente o material a ser exposto em oito aulas-conferências sobre processos estocásticos. Não é, nem pretende ser, um livro sobre processos estocásticos. Não contém uma exposição rigorosa dos temas tratados, porque essa requeriria da audiência pelo menos um curso avançado em Teoria das Probabilidades (que por sua vez requer cursos avançados de matemática como Teoria da Medida por exemplo) e muito mais.

A maioria das demonstrações apelam para argumentos intuitivos. No caso de que estes não possam ser fornecidos, referências adequadas são mencionadas. Várias demonstrações são simplesmente esboçadas. Em compensação uma grande quantidade de exemplos são introduzidos e são parcial ou totalmente estudados com os elementos provenientes da teoria exposta anteriormente.

É ao estudo dos exemplos que eu sugeriria que o leitor se concentrasse. Neles (como em outras áreas da matemática) pode se encontrar a motivação para o estudo mais avançado desta teoria e as ideias que podem gerar novos exemplos e novas aplicações.

Rio de Janeiro, maio de l975

Pedro J. Fernandez

ÍNDICE

1

Introdução aos Processos Estocásticos

1.1 Introdução

Suponhamos que para cada realização de um experimento aleatório o *resultado* fosse uma função (digamos no intervalo $[0, 1]$ para fixar as idéias). Dois possíveis resultados estão indicados na Figura 1.1.

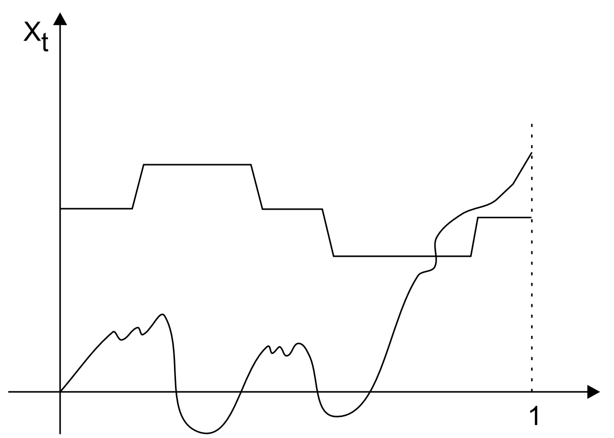


Fig. 1.1

t é usualmente interpretado como o tempo, mas poderia ser uma distância a um ponto fixo, um volume, etc.

Suponhamos para sermos mais concretos que ao tempo 0 observamos uma partícula (de um gás por exemplo) na origem de coordenadas de um sistema de coordenadas tridimendionais. Suponhamos que para cada instante de tempo $t \geq 0$ registramos o valor da primeira das coordenadas. Obteríamos um gráfico como o da Figura 1.2.

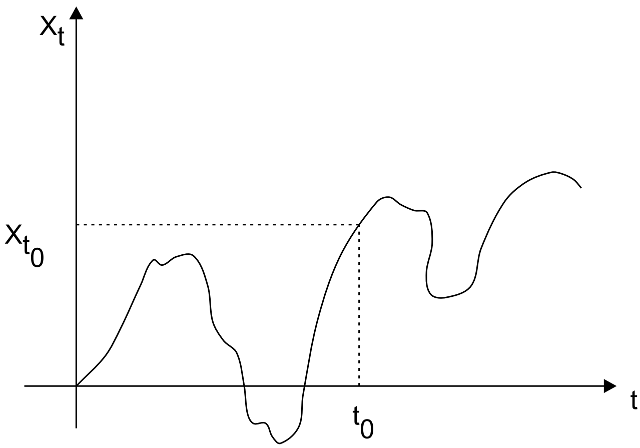


Fig. 1.2

X_{t_0} é então o valor da primeira coordenada no instante de tempo t_0. Este fenômeno foi observado pelo botânico inglês Robert Brown em 1827 e estudado posteriormente por A. Einstein, N. Wiener, P. Levy, etc. Estudaremos ele com cuidado no Capítulo 4.

Consideremos agora o fenômeno que consiste em contar o número de partículas emitidas por uma substância radioativa num intervalo de tempo $[0, t]$, $t \geq 0$. Uma realização "típica"deste experimento está representada na Figura 1.3.

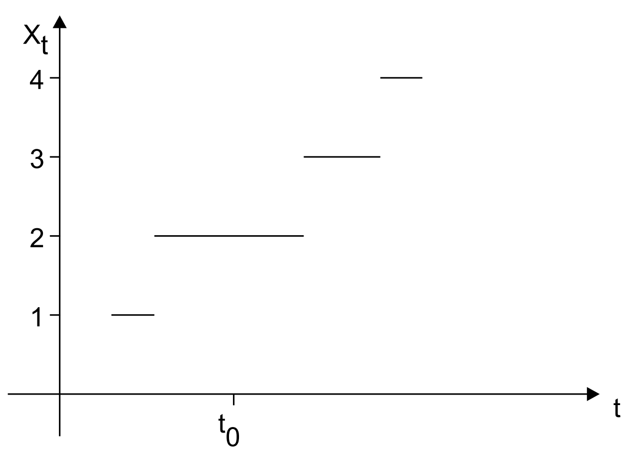


Fig. 1.3

X_{t_0} indica o número de partículas (2) emitidas pela substância até o tempo t_0. Um gráfico semelhante seria obtido se o fenômeno observado fosse o número de chamadas telefônicas que chegam a uma central telefônica, ou o número de pessoas atendidas num supermercado, etc.

Em forma imprecisa diremos que um *processo estocástico* é um modelo matemático utilizado para o estudo de fenômenos aleatórios que têm como resultados funções. Essas funções que chamaremos *trajetórias* estão definidas sobre um conjunto arbitrário T mas que nós tomaremos usualmente como sendo um intervalo na reta real ou o conjunto $[0, +\infty)$. O conjunto T será chamado *conjunto de parâmetros.*

Como para cada ponto ω no espaço amostral Ω temos uma trajetória, um processo estocástico também pode ser pensado como uma função de duas variáveis, $\omega \in \Omega$ e $t \in T$, a valores em um conjunto E que chamaremos *espaço de estados* que usualmente será $[0, +\infty)$, $\mathbb{R}^1$, ou um conjunto enumerável ou finito.

Se fixamos t, temos uma variável aleatória sobre Ω. Podemos pensar então todo processo estocástico como *uma família de variáveis aleatórias* (a valores em E) $\{X_t\}_{t \in T}$. Colocamos esta idéia em forma de definição.

Definição 1.1. Um *processo estocástico* consiste de um conjunto não vazio T que chamamos *espaço paramétrico* e na associação para cada $t \in T$ de uma variável aleatória $X_t \colon \Omega \to E$ todas elas definidas sobre o mesmo espaço de probabilidades.

Tomaremos usualmente como T o conjunto $[0, +\infty)$ ou $\mathbb{N} = \{0, 1, 2, \dots\}$. No primeiro caso, falaremos de um processo estocástico a *parâmetro contínuo* e no segundo de um processo estocástico a *parâmetro discreto.* Para uma definição completamente rigorosa de processo estocástico veja o Apêndice 5. A menos que seja mencionado explicitamente o contrário o conjunto E será $\mathbb{R}^1$.

1.2 Famílias consistentes

Seja T um conjunto de índices como os mencionados anteriormente. Suponhamos que para cada $n \geq 1$ e para cada subconjunto $t_1 < t_2 < \cdots < t_n$ de elementos de T associamos uma probabilidade $P_{t_1 \dots t_n}$ sobre $\mathbb{R}^n$. Dizemos que esta família de probabilidades é *consistente* se quando consideramos variáveis aleatórias $X_{t_1} \dots X_{t_n}$ com distribuição conjunta $P_{t_1 \dots t_n}$ a distribuição conjunta de cada subconjunto de $n-1$ variáveis $X_{t_1} \dots X_{t_{i-1}}, X_{t_{i+1}} \dots X_{t_n}$ é igual a $P_{t_1 \dots t_{i-1} t_{i+1} \dots t_n}$.

Um processo estocástico $\{X_t\}_{t\in T}$ gera uma família consistente de probabilidades: basta tomar $P_{t_1\ldots t_n} = \mathcal{L}(X_{t_1}\ldots X_{t_n})$. Reciprocamente dada uma família de distribuições consistente será que existe um processo estocástico que gera essa família? A resposta é afirmativa e está contida no seguinte teorema devido a Kolmogorov. (Veja também Apêndice 5.)

Teorema 1.2.1. *Seja T um conjunto de índices (intervalo finito ou infinito de números reais ou $\{0, 1, \ldots, \}$). Para toda família consistente de probabilidades com conjunto de parâmetros T existe um processo estocástico $\{X_t\}_{t\in T}$ que gera esta família.*

Em outras palavras um processo estocástico pode ser construido de forma à ajustar-se as seguintes condições. Damos primeiro para cada $t \in T$ a distribuição P_t sobre $\mathbb{R}^1$ que queremos que seja a distribuição de X_t.

Depois para cada t_1 e t_2 damos as distribuições $P_{t_1t_2}$ que queremos que sejam as distribuiçções de $\big(X_{t_1}, X_{t_2}\big)$. Logicamente as duas marginais de $P_{t_1t_2}$ devem ser P_{t_1} e P_{t_2}. Depois damos as distribuições $P_{t_1t_2t_3}$ que queremos seja a distribuição de $\big(X_{t_1}, X_{t_2}, X_{t_3}\big)$. Logicamente as três marginais de $P_{t_1t_2t_3}$ devem coincidir com $P_{t_1t_2}$, $P_{t_1t_3}$ e $P_{t_2t_3}$. Se continuamos com esse processo de especificação mantendo as condições de consistência teremos uma família consistente de probabilidades. O Teorema 1 garante a existência de um processo estocástico gerando essa família.

1.3 Alguns tipos gerais de processos estocásticos

Diz-se que um processo estocástico tem *incrementos estacionários* se $\forall\, t < x$, $\forall\, t' < s'$ com $s - t = s' - t'$ temos

$$\mathcal{L}(X_s - X_t) = \mathcal{L}(X_{s'} - X_{t'}).$$

De outra forma se a distribuição de $X_s - X_t$ depende de s e t só através de sua diferença $s - t$.

Um processo estocástico $\{X_t\}_{t\in T}$ é dito com *incrementos independentes* se $\forall\, n$, $\forall\, t_1 < t_2 < \cdots < t_n$ as variáveis aleatórias

$$X_{t_2} - X_{t_1}\,, X_{t_3} - X_{t_2}\,, \ldots\,, X_{t_n} - X_{t_{n-1}}$$

são independentes.

Um processo com *incrementos independentes estacionários* é, naturalmente um processo com incrementos simultaneamente independentes e estacionários.

Para um processo deste tipo consideremos $P_t = \mathcal{L}(X_t)$ para $t > 0$ ($T = [0, +\infty)$).
Temos que

$$\begin{aligned} P_{t+s} &= \mathcal{L}(X_{t+s}) = \mathcal{L}(X_{t+s} - X_t + X_t) \\ &= \mathcal{L}(X_{t+s} - X_t) * \mathcal{L}(X_t) = \\ &= \mathcal{L}(X_s) * \mathcal{L}(X_t) = \\ &= P_s + P_t\,. \end{aligned}$$

Uma família de probabilidades sobre $\mathbb{R}^1$ com esta propriedade é chamada um *semigrupo.* Reciprocamente suponhamos que temos dado sobre $\mathbb{R}^1$ um semigrupo de probabilidades. Será que existe um processo com incrementos independentes e estacionários gerando esse semigrupo? A resposta é afirmativa e a construção é realizada usando o Teorema de Consistência de Kolmogorov. Dado $0 < t_1 < t_2 < \cdots < t_n$ considere $P_{t_1}, P_{t_2-t_1}, \ldots, P_{t_n-t_{n-1}}$ r.a. independentes $z_1, \ldots, z_n$ com estas distribuições. Defina $P_{t_1 \ldots t_n}$ como a distribuição conjunta de

$$(z_1, z_1 + z_2, z_1 + z_2 + z_3, \ldots, z_1 + z_2 + \cdots + z_n).$$

A verificação de que estas distribuições são consistentes é simples. A existência do processo é dada então pelo Teorema da Consistência.

Finalmente vamos introduzir duas classes de processos muito importantes. Vamos tomar para fixar idéias $T = [0, +\infty)$.

Um processo $\{X_t\}_{t \geq 0}$ é chamado *estacionário* se $\forall\, s \geq 0$, $\forall\, 0 \leq t_1 < t_2 < \cdots < t_n$,

$$\mathcal{L}\big(X_{t_1}, \ldots, X_{t_n}\big) = \mathcal{L}\big(X_{t_1+s}, \ldots, X_{t_n+s}\big)$$

ou seja, a distribuição conjunta de $X_{t_1}, \ldots, X_{t_n}$ é igual a distribuição conjunta de $X_{t_1+s}, \ldots, X_{t_n+s}$. Intuitivamente podemos pensar que para um processo estacionário a probabilidade de uma trajetória e a probabilidade de uma translação dessa trajetória são iguais. Como a distribuição de (X_t, x_{t+s}) não depende de t o mesmo acontece com $X_{t+s} - X_t$ e portanto todo processo estacionário tem incrementos estacionários. A recíproca não é verdadeira para o qual basta ver que para um processo estacionário a distribuição de X_t não depende de t o que não necessariamente é válido para um processo com incrementos estacionários.

Um processo estocástico $\{X_t\}_{t\geq 0}$ é chamado um *processo de Markov* se $\forall\, n \geq 3$, $\forall\, t_1 < t_2 < \cdots < t_n$, $\forall\, x_1, x_2, \ldots, x_n$ reais

$$P\big[X_{t_n} < x_n \mid X_{t_1} = x_1, \ldots, X_{t_{n-1}} = x_{n-1}\big] = \\ = P\big[X_{t_n} < x_n \mid X_{t_{n-1}} = x_{n-1}\big].$$

Ou seja a distribuição condicional de uma variável dado um número finito de variáveis anteriores depende somente da última destas variáveis. De outra forma o futuro não depende da forma em que chegamos a $X_{t_{n-1}} = x_{n-1}$; depende só do fato que no tempo t_{n-1} estamos no estado x_{n-1}.

Se um processo tem incrementos independentes então é de Markov.

Nos Capítulos 2 e 3 estudaremos em detalhe um tipo muito importante de processo estocástico a parâmetro discreto e com espaço de estados E finito ou enumerável (Cadeias de Markov a parâmetro discreto).

No Capítulo 4 permitiremos que o parâmetro seja contínuo ($t = [0, +\infty)$).

No Capítulo 5 falaremos de alguns processos a parâmetro contínuo com espaços de estados $\mathbb{R}^1$ (Movimento Browniano, etc.).

Exercício 1. Seja $\{X_t\}_{t\geq 0}$ um processo com incrementos independentes e estacionários. Seja $c > 0$, e a, b e d números reais. Então o processo

$$Y_u = a + bu + dX_{c(u-u_0)} \quad u \geq u_0 \geq 0$$

é também um processo com incrementos independentes e estacionários.

Exercício 2. Seja $\{X_t\}_{t\geq 0}$ um processo com incrementos independentes e estacionários. Suponhamos que $\exists\, M \geq 0$ tal que $\sup\limits_{0\leq t\leq 1} EX_t^2 \leq M$ e $X_0 = 0$. Então

a) $EX_t - t\,EX_1$ e $\operatorname{var} X_t = t\ \operatorname{var}(X_1)$ (ver Apêndice 4);

b) Prove que $\forall\, t > 0$

$$P\big[|X_t - t\,EX_1| \geq \delta\sqrt{t\ \operatorname{var} X_1}\big] \geq 1 - \frac{1}{\delta^2}$$

e portanto o conjunto de trajetórias que no tempo t satisfaçam a condição

$$|X_t - t\,EX_1| \leq 2\sqrt{t\ \mathrm{var}\,X_1}$$

tem probabilidade maior a $0,75$.

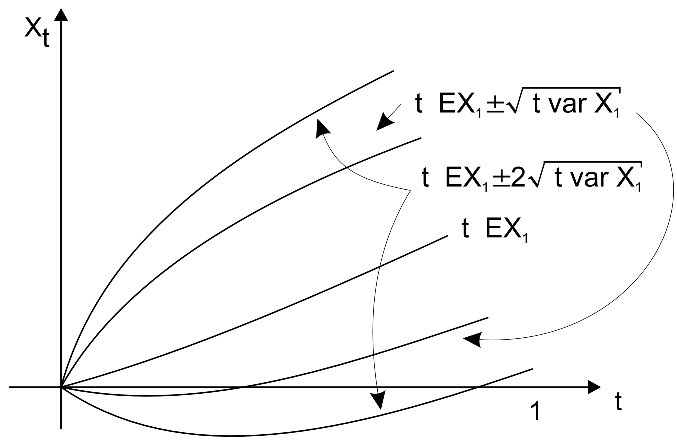


Fig. 1.4

Exercício 3. Seja $\{P_t\}_{0<t\leq\delta}$, $\delta > 0$, um semigrupo de probabilidades sobre $\mathbb{R}^1$. Provar que existe um único semigrupo sobre $(0, +\infty)$, $\{Q_t\}_{t>0}$, tal que $\forall\, t$ com $0 < t \leq \delta$, $Q_t = P_t$.

2

Cadeias de Markov

2.1 Definição e exemplos

Seja E um conjunto enumerável ou finito (que em geral vai ser um subconjunto dos inteiros), os elementos de E serão chamados *estados* e E o *espaço de estados*. Seja $\{p_i\}_{i\in I}$ uma sucessão de números tais que $p_i \geq 0$ e $\sum\limits_{i\in E} p_i = 1$.

Seja P uma matriz de números p_{ij}, $i, j \in E$, tais que $p_{ij} \geq 0$ e $\forall\, i \in E \quad \sum\limits_{j\in E} p_{ij} = 1$. Uma tal matriz é chamada *matriz markoviana*.

Definição 2.1.1. Uma sucessão de variáveis aleatórias $\{X_n\}_{n=0,1,2,\dots}$ é chamada uma *Cadeia de Markov* com distribuição inicial $\{p_i\}_{i\in E}$ e matriz de transição P se:

a) $\forall\, i \in E \quad P[X_0 = i] = p_i$.

b) $\forall\, n;\ \forall\, i_0 i_1 \dots i_n i_{n+1} \in E$ tal que

$$P[X_0 = i_0, \dots, X_n = i_n] > 0$$
$$P\big[X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n\big] = p_{i_n\, i_{n+1}}.$$

A propriedade a) diz que a distribuição de variável aleatória X_0 é a determinada pela sucessão $\{p_i\}_{i\in E}$. A propriedade b) diz que $\forall\, n$, a distribuição condicional de X_{n+1} dado $X_0 X_1 \dots X_n$, não depende nem da sucessão $i_0 \dots i_n$, nem do instante n, só depende do estado i_n ocupado no tempo n.

Outra definição está contida na seguinte proposição.

Proposição 2.1.1. *Dado um espaço de estados E, uma distribuição $\{p_i\}_{i\in E}$ é uma matriz markoviana P; uma sucessão de variáveis aleatórias $\{X_n\}_{n=0,1,\dots}$ é uma cadeia de Markov se, e somente se, $\forall\, n$, $\forall\, i_1 \dots i_n \in E$.*

$$p[X_0 = i_0, \dots, X_n = i_n] = p_{i_0} p_{i_0} i_1 \dots p_{i_{n-1}} i_n \,.$$

Demonstração: Que a condição é suficiente e facilmente verificável.

Vamos, agora, provar a necessidade por indução. Para $n = 0$, é evidente. Consideremos

$$P[X_0 = i_0, \dots, X_{n+1} = i_{n+1}].$$

Podemos supor $p_{i_0} > 0$. Se $P[X_0 = i_0, \dots, X_n = i_n] > 0$, temos que

$$\begin{aligned}
&P[X_0 = i_0, \dots, X_{n+1} = i_{n+1}] = \\
&= [X_0 = i_0, \dots, X_n = i_n] = \\
&= P[X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n]
\end{aligned}$$

(por indução)

$$= p_{i_0} p_{i_0 i_1}, \dots, p_{i_{n-1} i_n} p_{i_n i_{n+1}} \,.$$

Se $P[X_0 = i_0, \dots, X_n = i_n] = 0$, então $P[X_0 = i_0, \dots, X_{n+1} = i_{n+1}] = 0$ e existe j, $0 \le j \le n-1$, tal que $p_{i_j i_j + 1} = 0$.

Portanto, ambos os membros da identidade a ser provada são iguais a 0.

Nota: O espaço de probabilidades sobre o qual estão definidas as variáveis aleatórias que entram na definição de cadeia de Markov não foi descrito explicitamente porque isso ultrapassaria o nível desta monografia. Para os problemas que vamos discutir não é necessário. O problema da existência de uma cadeia de Markov, dada uma distribuição inicial $\{p_i\}_{i\in I}$, é uma matriz markoviana P, tampouco vai ser discutido pelas mesmas razões. Uma tal cadeia de Markov existe sempre. O leitor interessado e com algumas noções de Teoria da Medida pode consultar por exemplo [1]. (A construção é feita utilizando os resultados do Apêndice 5.)

Exemplo 2.1.1 - Vamos supor que uma certa mensagem é transmitida usando só os dígitos 0 e 1.

A probabilidade de que um dígito seja transmitido corretamente é p e que mude, $q = 1 - p$. Podemos representar este processo por uma cadeia de Markov com os estados 0 e 1, e com a matriz de transição

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} p & q \\ q & p \end{pmatrix} \end{array}$$

onde o elemento i, j da matriz indica a probabilidade de transmitir j quando o verdadeiro dígito é i.

Exemplo 2.1.2 - Consideremos um jogador que ganha ou perde uma unidade com probabilidades p e q respectivamente. Vamos supor que o seu capital inicial é x e que o seu adversário tem um capital de $a - x$ ($a \geq x$, e portanto o capital total é a). O jogo continua até que um dos jogadores fique arruinado, ou seja, até que o primeiro jogador aumente seu capital até a ou perde x.

A matriz de transição desta cadeia é a seguinte

$$P = \begin{array}{cc} & \begin{array}{ccccccc} 0 & 1 & 2 & \cdot & \cdot & \cdot & a \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ a-1 \\ a \end{array} & \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ q & 0 & p & \cdot & \cdot & \cdot & 0 \\ 0 & q & 0 & p & \cdot & \cdot & 0 \\ & & & & & & \\ \cdot & \cdot & \cdot & \cdot & q & 0 & q \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Exemplo 2.1.3 - Vamos supor que o nível econômico de um homem é classificado em três categorias: rico (R), classe média (M), e pobre (P). Vamos supor que dos filhos de um homem rico 95% são ricos e 5% de classe média. No caso de um indivíduo de classe média, 10% são ricos, 70% de classe média e 20% pobres, E, finalmente, no caso de um homem pobre, 30% são de classe média e 70% são pobres. Supondo que cada homem tem um filho, podemos formar uma cadeia de Markov observando uma família através de gerações sucessivas. A matriz de probabilidade de transição é

$$\begin{array}{cc} & \begin{array}{ccc} \mathrm{R} & \mathrm{M} & \mathrm{P} \end{array} \\ \begin{array}{c} \mathrm{R} \\ \mathrm{M} \\ \mathrm{P} \end{array} & \begin{pmatrix} 0,95 & 0,05 & 6 \\ 0,10 & 0,70 & 0,20 \\ 0 & 0,30 & 0,70 \end{pmatrix} \end{array}$$

Exemplo 2.1.4 - Vamos modificar o Exemplo 2.1.3 supondo que todo homem não tem, necessariamente, um filho. Vamos supor que a probabilidade de ter um filho é de 90%. Podemos então formar uma cadeia de Markov com quatro estados, agregados um, que vamos chamar Δ, aos três anteriores. Este vai ser um estado que, depois de alcançado, não pode ser abandonado (absorvente). A matriz de probabilidades de transição neste caso é:

$$\begin{array}{c} \\ R \\ M \\ P \\ \Delta \end{array} \begin{array}{c} \begin{array}{cccc} \;R\quad & \quad M\quad & \quad P\quad & \quad \Delta \end{array} \\ \begin{pmatrix} 0,855 & 0,045 & 0 & 0,10 \\ 0,09 & 0,63 & 0,18 & 0,10 \\ 0 & 0,27 & 0,63 & 0,10 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Exemplo 2.1.5 - Um jogo consiste no seguinte: um ponto inicial é escolhido no interior da rede da Figura 2.1.1 (vamos dizer o ponto 5). Um movimento ao acaso é começado nesse ponto passando a cada um dos pontos vizinhos com igual probabilidade (por exemplo, se o ponto inicial foi o ponto 5 no primeiro movimento é possível alcançar os pontos 2, 4, 6 e 7 com probabilidade 1/4 para cada um deles). O processo é terminado quando alguns dos pontos da fronteira (A, B, C, D, E, F, G, H, I, J) são alcançados.

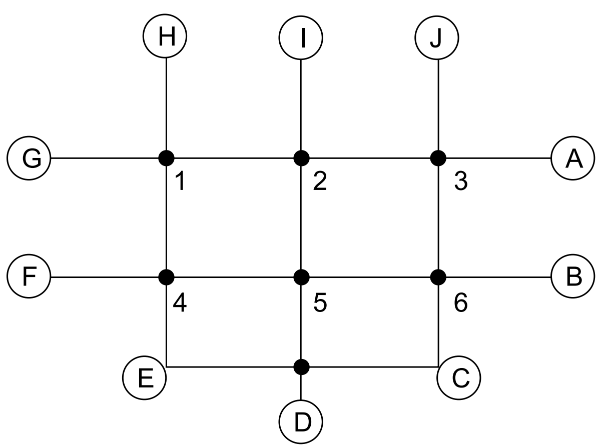


Fig. 2.1.1

Esta cadeia de Markov está descrita pela matriz de transição:

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ A \\ B \\ C \\ D \\ E \\ F \\ G \\ H \\ I \\ J
\end{array}
\begin{array}{c}
\begin{array}{ccccccc|cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & A & B & C & D & E & F & G & H & I & J
\end{array} \\
\left(\begin{array}{ccccccc|cccccccccc}
0 & \frac14 & 0 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac14 & \frac14 & 0 & 0 \\
\frac14 & 0 & \frac14 & 0 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac14 & 0 \\
0 & \frac14 & 0 & 0 & 0 & \frac14 & 0 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac14 \\
\frac14 & 0 & 0 & 0 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & \frac14 & \frac14 & 0 & 0 & 0 & 0 \\
0 & \frac14 & 0 & \frac14 & 0 & \frac14 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac14 & 0 & \frac14 & 0 & 0 & 0 & \frac14 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac14 & 0 & 0 & 0 & \frac14 & \frac14 & \frac14 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & & & & & & 1 & & & & & & & & & \\
 & & & & & & & & 1 & & & & & & & & \\
 & & & & & & & & & \cdot & & & 0 & & & & \\
 & & & & & & & & & & \cdot & & & & & & \\
 & & & 0 & & & & & & & & \cdot & & & & & \\
 & & & & & & & & & & & & \cdot & & & & \\
 & & & & & & & & & & & & & \cdot & & & \\
 & & & & & & & & 0 & & & & & & \cdot & & \\
 & & & & & & & & & & & & & & & \cdot & \\
 & & & & & & & & & & & & & & & & 1
\end{array}\right)
\end{array}
$$

Exemplo 2.1.6 (Passeio ao acaso) - Suponhamos uma partícula na origem da reta real. Por algum mecanismo aleatório selecionamos 1 e -1 com probabilidade p e q respectivamente. Se o resultado é 1 a partícula se move uma unidade para a direita (de 0 a 1); se o resultado é -1, uma unidade para a esquerda (0 a -1). Se depois de i movimentos (tempo i) a partícula se encontra na posição k e o resultado é 1, a partícula se move para a posição $k+1$. se é -1, para a posição $k-1$. O espaço de estados é os inteiros e a matriz de transição está dada por:

$$p_{i,i+1} = p, \quad p_{i,i-1} = q, \quad p \geq 0, \quad p + q = 1.$$

Uma variante deste exemplo se obtém tomando como espaço de estados os inteiros não negativos e definindo P da seguinte forma:

$$
P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \end{array}
\begin{array}{c}
\begin{array}{cccccc} 0 & 1 & 2 & \cdot & \cdot & \cdot \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & \cdot & \cdot & \cdot \\
q & 0 & p & \cdot & \cdot & \cdot \\
0 & q & 0 & p & \cdot & \cdot \\
\vdots & & & & &
\end{pmatrix}
\end{array}
\qquad p \neq 0, \quad q \neq 0
$$

Exemplo 2.1.7 - Uma seqüência de tarefas é realizada em uma ordem determinada cada uma com sua probabilidade de sucesso.

Sucesso em uma tarefa implica em passar a realizar a tarefa seguinte; fracasso implica em começar tudo de novo. Tomamos como espaço de estados para esta cadeia os inteiros não negativos e como P a matriz

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & \cdots \end{array} \\ \begin{pmatrix} q_1 & p_1 & 0 & 0 & \cdots \\ q_2 & 0 & p_2 & 0 & \cdots \\ q_3 & 0 & 0 & p_3 & \cdots \end{pmatrix} \end{array}.$$

O valor p_i é a probabilidade de transição do estado $i-1$ ao estado i, e q_i é a probabilidade de transição do estado $i-1$ ao estado 0. p_i é a probabilidade de sucesso ao realizar a tarefa i. Vamos supor que $\forall\, i$, $0 < p_i$ e que $p_i < 1$ para infinitos i.

Exemplo 2.1.8 - *Processo de ramificação.* O espaço de estados neste exemplo vai ser os inteiros não negativos. Seja Q uma probabilidade fixa sobre os inteiros não negativos (determinada por $\{q_i\}_{i=0,1,\ldots}$, $q_i \geq 0$, $\sum\limits_{i=0}^{\infty} q_i = 1$). Seja $\{Y_n\}_{n=1,2,\ldots}$ uma sucessão de variáveis aleatórias independentes identicamente distribuidas com lei comum $Q(\mathcal{L}(Y_n) = Q, \forall\, n)$. Definimos $p_{ij} = P\big[\sum\limits_{k=1}^{i} Y_k = j\big]$ para $i \geq 1$, $p_{oo} = 1$ e $p_{oj} = 0$ para $j \geq 1$.

A situação a que corresponde este modelo é a seguinte: a probabilidade de que uma bactéria dê lugar a n delas antes de morrer, é q_n . A v.a. $\sum\limits_{k=1}^{i} Y_k$ indica o número de bactérias originadas por i delas antes de morrer. A v.a. X_n , correspondente à cadeia de Markov associada a esta matriz de transição, indica o número de bactérias na colônia no instante de tempo n.

Exemplo 2.1.9 - Consideremos a carteira de um grande estabelecimento comercial. Cada conta nesse porta-documentos pode ser classificada em um dos seguintes "estados". Paga (P), antiguidade 0 (se todas as faturas na conta tem menos de um mês), antiguidade 1 (se a fatura com maior antiguidade tem data de emissão entre 1 e 2 meses atrás); antiguidade 2 e 3 são definidas em forma semelhante. Finalmente uma conta é colocada como perda (L) (e outros procedimentos são adaptados para sua cobrança) se sua antiguidade é superior a 4 meses.

Utilizando um procedimento que veremos posteriormente a matriz de probabilidade de transição é igual a

$$P = \begin{array}{c} \\ P \\ L \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccccc} P & L & 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0,3 & 0 & 0,6 & 0,1 & 0 & 0 \\ 0,2 & 0 & 0,2 & 0,5 & 0,1 & 0 \\ 0,1 & 0 & 0,2 & 0,5 & 0,4 & 0,1 \\ 0,1 & 0,1 & 0,1 & 0,1 & 0,2 & 0,4 \end{pmatrix} \end{array}.$$

Exemplo 2.1.10 *Cadeia de Ehrenfest* - Suponhamos ter d bolinhas enumeradas $1, 2, \ldots, d$. Inicialmente estas d bolinhas estão distribuidas em duas urnas. Um número entre 1 e d é escolhido aleatoriamente e a bolinha com esse número é transferida da urna em que se encontra para a outra. Seja $\{X_n\}_{n=0,1,2,\ldots}$ o número de bolinhas na urna 1 no instante de tempo n. $\{X_n\}_{n=0,1,\ldots}$ é uma cadeia de Markov com matriz de probabilidades de transição dada por

$$P(i,j) = \begin{cases} \dfrac{1}{d} & j = i-1 \\ 1 - \dfrac{i}{d} & j = i+1 \\ 0 & j \neq i = 1, \;\; j \neq i+1 \end{cases}$$

Exemplo 2.1.11 - *Cadeia fila.* Consideremos um lugar onde pessoas chegam em diferentes momentos e são eventualmente atendidas na ordem da sua chegada (central telefônica, caixa em um supermercado, etc.). Todas as pessoas que chegaram mas que ainda não foram atendidas uma fila.

Vamos considerar um modelo um pouco artificial para descrever este sistema. Dividamos o tempo em intervalos iguais (por exemplo minutos). Vamos supor que haja pessoas esperando, ao início de um período exatamente uma será atendida nesse período. Se não há pessoas esperando nenhuma será atendida durante esse período. Seja Y_n uma v.a. que indica o número de clientes que chegam durante o intervalo n. Vamos supor que estas v.a. são independentes e identicamente distribuidas com distribuição determinada pela função g, onde $\forall\, n$, $P[Y_n = k] = g(k)$, $k = 0, 1, 2, \ldots$. Seja X_n o número de pessoas na fila no final do período n. (X_0 indica o número de clientes presentes inicialmente). Se $X_n = 0$

então $X_{n+1} = Y_{n+1}$; se $X_n \geq 1$ então $X_{n+1} = X_n + Y_{n+1} - 1$. Resulta que X_n é uma cadeia de Markov com espaço de estados $E = \{0, 1, 2, \dots\}$ e com matriz de probabilidades de transição dada por

$$P(0, j) = g(j)$$
$$P(i, j) = g(j - i + 1),\ i \geq 1,\ j \geq i - 1$$

Para os restantes valores de i, j, P é definido como sendo 0.

Seja μ a esperança da distribuição g e φ a sua função geradora de probabilidades.

2.2 Probabilidades de transição de ordem superior

Vamos indicar com $P^{(n)}$ a potência n da matriz de transição P, e com $P^{(n)}(i, j)$ o termo i, j dessa matriz. A probabilidade condicional $P[\cdot \mid X_0 = i]$ e as esperanças com respeito a esta probabilidade serão demonstradas com P^i e E^i, respectivamente.

Vamos estabelecer agora uma relação muito útil para o cálculo de probabilidades de transição em n etapas.

Proposição 2.2.1

$$P[X_n = j \mid X_0 = i] = P^{(n)}(i, j) \quad n = 1, 2, \dots .$$

Demonstração: Vamos fazer o argumento para o caso $n = 2$. O caso geral é similar:

$$P[X_2 = j \mid X_0 = i] = \frac{P[X_0 = i, X_2 = j]}{p_i} =$$

$$\frac{\sum\limits_{k \in E} P[X_0 = i, X_1 = k, X_2 = j]}{p_i} = \frac{\sum\limits_{k \in E} p_i p_{ik} p_{kj}}{p_i} =$$

$$= \sum_{k \in E} p_{ik} \mp_{kj} = P^{(2)}(i, j).$$

Corolário 2.2.1. *Se π indica o vetor da distribuição inicial $\{p_i\}_{i \in E}$, então a distribuição de X_n está determinada pelo vetor $\pi\, P^{(n)}$.*

Demonstração:

$$
\begin{aligned}
P[X_n = j] &= \sum_{\{i:i\in E, p_i>0\}} P[X_n = j \mid X_0 = i] P[X_0 = i] \\
&= \sum_{\{i:i\in E, p_i>0\}} p_i \, P^{(n)}(i,j) = \sum_{i\in E} p_i \, P^{(n)}(i,j).
\end{aligned}
$$

Proposição 2.2.2. a) *Se* $P[X_0 = i_0, \ldots, X_m = i_m] > 0$, *então,*

$$P[X_{n+m} = j \mid X_0 = i_0, \ldots, X_m = i_m] = P^{(n)}(i_m, j).$$

b) $f\colon E - \mathbb{R}$ ($f \geq 0$ *ou limitada*), *então* $\forall\, i \in E$

$$E^i\, f(X_n) = \sum_{j\in E} P^{(n)}(i,j)\ f(j).$$

c) $E^i\, f(X_{n+m}) = \sum_{k\in E} P^{(n)}(i,k)\, E^k\, f(X_n)$.

Demonstração: Só vamos provar c).

$$
\begin{aligned}
E^i\, f(X_{n+m}) &= E^i \left\{ \sum_{i_1, i_2, \ldots, i_{m-1}, k, j} f(j)\, I_{[X_0=i, X_1=i_1, \ldots, X_m-k, X_{m+n}=j]} \right\} = \\
&= \sum_{\substack{i_1, i_2, \ldots, k, j \\ P[X_0=i, X_1=i_1, \ldots, X_m=k]>0}} f(j) P^i[X_0 = i, \ldots, X_{m+n} = j] = \\
&= \sum_{\substack{i_1, i_2, \ldots, k, j \\ P[X_0=i, X_1=i_1, \ldots, X_m=k]>0}} f(j) \frac{1}{p_i} P[X_0 = i, \ldots, X_m = k] \cdot \\
&\qquad \cdot P[X_{m+n} = j \mid X_0 = i, \ldots, X_m = k] = \\
&= \sum_{i_1, i_2, \ldots, k, j} f(j) \frac{1}{p_i}\, P[X_0 = i, \ldots, X_m = k] P^{(n)}(k,j) = \\
&= \sum_k \left(\underbrace{\sum_{i_1, i_2, \ldots, i_{m-1}} \frac{1}{p_i} P[X_0 = i, \ldots, X_m = k]}_{P^i[X_m=k]} \right) \left(\underbrace{\sum_j f(j) P^{(n)}(k,j)}_{=E^k f(X_n)} \right) = \\
&= \sum_k P^{(m)}(i,k) E^k f(X_n).
\end{aligned}
$$

2.3 Classificação de estados

Definição 2.3.1. Dizemos que $i \in E$ *conduz* a $j \in E$ e escrevemos $i \to j$ se $P^{(n)}(i,j) > 0$ para algum $n \geq 1$. Dizemos que $i \in E$ e $j \in E$ *comunicam-se*, e escrevemos $i \leftrightarrow j$ se e somente se $i \to j$ e $j \to i$.

Proposição 2.3.1. *A relação* $\leftrightarrow$ *é uma relação de equivalência.*

Demonstração: Verificaremos só a transitividade. Suponhamos $i \to j$ e $j \to k$. Ou seja $\exists\, n \geq 1$ e $m \geq 1$ tais que $P^{(n)}(i,j) > 0$ e $P^{(m)}(j,k) > 0$. Então

$$P^{(n+m)}(i,k) = \sum_{s \in E} P^{(n)}(i,s) P^{(m)}(s,k) \geq P^{(n)}(i,j) P^{(m)}(j,k) > 0.$$

Portanto $i \to k$ o que prova a transitividade.

Esta relação de equivalência divide o espaço de estados em classes de equivalência.

Definição 2.3.2. Uma cadeia é chamada *irredutível*, se tem só uma classe de equivalência.

Seja $f_{ii}^{(n)} = P^i[X_n = i, X_1 \neq i, X_2 \neq i, \ldots, X_{n-1} \neq i]$. $f_{ii}^{(n)}$ indica a probabilidade de retornar ao estado i depois de exatamente n período de tempo. (Na notação do Apêndice 2 $f_{ii}^{(n)} = P^i[\bar{\tau}_i = n]$). Seja $\overline{H}_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)}$ a probabilidade de retornar alguma vez ao estado i. Definimos também $\mu_i = \sum_{n=1}^{\infty} n\, f_{ii}^{(n)}$. μ_i indica o número médio de etapas necessárias para retornar ao estado inicial i. μ_i é chamado *tempo médio de recorrência.*

Definição 2.3.3. Um estado $i \in E$ é chamado *recorrente* se $\overline{H}_{ii} = 1$. É chamado *transitório* se $\overline{H}_{ii} < 1$.

Definição 2.3.4. Um estado recorrente é chamado *recorrente nulo* (*recorrente positivo*) se $\mu_i = \infty$, ($\mu_i < \infty$).

O seguinte teorema é importante mas a sua demonstração é mais complicada e foi portanto incluida para os interessados no Apêndice 2.

Teorema 2.3.1. *Um estado* $i \in E$ *é recorrente ou transitório de acordo com que*

$$\sum_{n=1}^{\infty} P^{(n)}(i,i) = \infty \quad ou \quad < \infty.$$

Uma conseqüência imediata deste teorema é a seguinte.

Proposição 2.3.2. *Se $i \leftrightarrow j$ e i é recorrente então j é recorrente. (Ou seja a propriedade de ser recorrente é uma propriedade de classe).*

Demonstração: Sejam n e m tais que $P^{(n)}(i,j) > 0$ e $P^{(m)}(j,i) > 0$. Então

$$P^{(n+m+s)}(j,j) \geq P^{(m)}(j,i)P^{(s)}(i,i)P^{(n)}(i,j).$$

Portanto

$$\sum_{s=0}^{\infty} P^{(n+m+s)}(j,j) \geq P^{(m)}(j,i)P^{(s)}(i,i)P^{(n)}(i,j)\sum_{s=0}^{\infty} P^{(s)}(i,i).$$

Se i é recorrente $\sum\limits_{s=0}^{\infty} P^{(s)}(i,i) = \infty$ e portanto esquerdo é também ∞ o que prova o resultado.

Definição 2.3.5. Um estado $i \in E$ é chamado *absorvente* se $P(i,i) = 1$. Portanto $\overline{H}_{ii} = 1$ e $\mu_i = 1$.

Definição 2.3.6. Se $i \to i$ então o máximo divisor comum (m.d.c.) de $\{n : n > 0, P^{(n)}(i,i) > 0\}$ é chamado *período do estado i*.

Proposição 2.3.3. *Se $i \to j$ então, período de i = período de j.*

Demonstração: Escolhamos a e b tais que $P^{(a)}(i,j) > 0$ e $P^{(b)}(j,i) > 0$. Temos também que

$$P^{(a+m+b)}(i,i) \geq P^{(a)}(i,j)P^{(m)}(j,j)P^{(b)}(j,i).$$

Se $P^{(m)}(j,j) > 0$ então $P^{(2m)}(j,j) > 0$ e $P^{(a+m+b)}(i,i) > 0$ e $P^{(a+2m+b)}(i,i) > 0$.

Portanto o período de i divide a $(a+m+b)$ e $(a+2m+b)$ e conseqüentemente a sua diferença ou seja m. Quer dizer o período de i divide qualquer número m tal que $P^{(m)}(j,j) > 0$. Portanto o período de i divide o período de j. Como i e j podem ser trocados neste argumento o resultado fica provado.

Uma conseqüência deste resultado é que o *período é uma propriedade de classe.* Podemos falar do *período de uma classe* como o período de qualquer um dos estados que compõem a classe.

A notação $m \equiv n(d)$ significa que $m - n$ é divisível por d. Seja $i \in E$ e seja I a classe a que i pertence; I com período d. Temos

Proposição 2.3.4. *Para cada $j \in I$ existe $r_j = 0, 1, \ldots, d-1$ tal que $P^{(n)}(i,j) > 0$ implica $n \equiv r_j(d)$.*

Demonstração: Seja tal que $P^{(a)}(j,i) > 0$. Se $P^{(m)}(i,j) > 0$ e $P^{(n)}(i,j) > 0$ então $P^{(m+a)}(i,i) > 0$ e $P^{(n+a)}(i,i) > 0$.

Portanto d divide $m+a$ e $n+a$. Então d divide a sua diferença ou seja $m-n$ ou seja $m \equiv n(d)$. Podemos definir então r_j como o resto de dividir n por d, para qualquer n com $P^{(n)}(i,j) > 0$.

Seja F um subconjunto dos inteiros contendo pelo menos um elemento diferente de 0. Denotaremos com grupo (F) (semigrupo (F)) o grupo (semigrupo) gerado por F. O semigrupo por F consiste de todas as somas $f_1 + f_2 + \cdots + f_m$ onde $m > 0$ e $f_i \in F$, $i = 1, 2, \ldots, m$.

É fácil verificar também que grupo $(F) = \{a - b;\ \in$ semigrupo (F), $b \in$ semigrupo $(F)\}$.

Para $a \in$ grupo (F) e k inteiro positivo $ka = \sum_{n=1}^{k} a \in$ grupo (F).

Lema 2.3.1. m.d.c. (F) = *menor elemento positivo de grupo* (F).

Demonstração: Seja a o menor elemento positivo de grupo (F). Pela forma que os elementos de grupo (F) são gerados resulta que m.d.c. $(F) \leq a$. Seja $f \in F$. $f = ka + r$ onde $0 \leq r < a$. Agora $r = f - ka \in$ grupo (F). Como é menor que a tem que ser $r = 0$. Portanto a divide f para todo $f \in F$, e conseqüentemente o m.d.c. (F) o que prova $a \leq$ m.d.c. (F).

Proposição 2.3.5. $F \subseteq \{1, 2, \ldots\}$. *Seja* $d =$ m.d.c. (F). *Então existe m_0 tal que $\forall\, m \geq m_0$, $md \in$ semigrupo (F).*

Demonstração: Pelo lema anterior temos que $d = a - b$ com a e b pertecendo ao semigrupo (F). Como semigrupo $(F) \supseteq$ semigrupo $\{a, b\}$ vamos provar a tese para semigrupo $\{a, b\}$. Escolhemos $m_0 = \left(\frac{b}{d}\right)^2$. Seja $m \geq m_0$, $m = k\,\frac{b}{d} + r$ com $0 \leq r < \frac{b}{d}$. Como $m \geq m_0$, $k \geq \frac{b}{d}$ e portanto $k - r > 0$. Multiplicando por d temos

$$md = kb + rd = kb + r(a - b) = (k - r)b + ra \in \text{ semigrupo } (F).$$

Proposição 2.3.6. *Seja k o período do estado i. Então existe m_0 tal que $\forall\, m \geq m_0$, $P^{(md)}(i,i) > 0$.*

Demonstração: $\{n : n \geq 1, P^{(n)}(i,i) > 0\}$ é um semigrupo. A proposição decorre agora da proposição anterior.

Definição 2.3.7. Uma classe de estados com período 1 é chamada *aperiódica.*

Proposição 2.3.7. *Seja P a matriz de probabilidades de transição de uma cadeia de Markov finita, irredutível e aperiódica. Então $\exists\, N$ tal que $\forall\, n \geq N$ todas as entradas de $P^{(n)}$ são positivas.*

Demonstração: Dado i e $j \in E$, $\exists\, k(i,j)$ tal que $P^{k(i,i)}(i,j) > 0$. Também existe, pela Proposição 2.3.6, $N(i,j)$ tal que $\forall\, n \geq N(i,j)$

$$P^{(k(i,j)+n)}(i,j) > 0.$$

Se definimos

$$N = \max_{i,j \in E} \{k(i,j) + N(i,j)\}$$

então $\forall\, n \geq N,\ \forall\, i \in E,\ \forall\, j \in E,\ P^{(n)}(i,j) > 0.$

Definição 2.3.8. Um estado $i \in E$ aperiódica e recorrente positivo é chamado *ergódico.*

2.4 Representação canônica da matriz de probabilidades de transição de uma Cadeia de Markov finita

Enumere os estados de forma tal que os estados que estejam na mesma classe de equivalência tenham números consecutivos.

Enumere as classes de equivalência de forma tal que o processo possa passar de um estado a outro na mesma classe ou em classe precedente, mas não em uma classe posterior na enumeração. Isto pode ser conseguido colocando as classes recorrentes primeiro e depois as classes transitórias, estas últimas ordenadas sempre com o mesmo critério. Sejam $C_1, C_2, \ldots, C_r, C_{r+1}, \ldots, C_n$ as classes de equivalência onde $C_1, C_2, \ldots, C_r$ são recorrentes e $C_{r+1}, \ldots, C_n$ são transitórias. A matriz de probabilidades de transição pode ser representada então da seguinte forma:

$$P = \begin{pmatrix} P_1 & & & & & & \\ & P_2 & & & & & \\ \bigcirc & & \ddots & & & \bigcirc & \\ A_{r+1,1} & & \ldots & A_{r+1,r} & B_{r+1} & & \\ A_{r+2,1} & & \ldots & A_{r+2,r} & A_{r+2,r+1} & B_{r+2} & \\ & & & & & & \ddots \\ A_{n,1} & & \ldots & A_{n,r} & A_{n,r+1} & \ldots & B_n \end{pmatrix}$$

$P_1, P_2, \dots, P_r, B_{r+1}, \dots, B_n$ são as matrizes de probabilidades de transição correspondentes às classes de equivalência e as $A_{i,j}$ são as matrizes que dão as probabilidades de transição de uma classe para as classes anteriores.

Proposição 2.4.1. *Numa cadeia de Markov finita a probabilidade de que o processo esteja num estado transitório converge a* 0*.*

Demonstração: Seja T o conjunto de estados transitórios. $T = \{i_1, i_2, \dots, i_s\}$. Enttão $\forall\, \ell$, $1 \le \ell \le s$, existe n_ℓ tal que $P^{i_\ell}[X_{n_\ell} \notin T] > 0$. Se $n = \sum_{\ell=1}^{s} n_\ell$ e $\delta = \min_{1 \le \ell \le s} P^{i_\ell}[X_{n_\ell} \notin T]$ então $\forall\, i \in T$, $P^i[X_n \notin T] \ge \delta > 0$ equivalentemente $\sum_{r \in T} P^{(n)}(i, r) \le 1 - \delta$, $\forall\, i \in T$.
Agora

$$P^{(nk)}(i, r) = \sum_{r_1} \cdots \sum_{r_{k-1}} P^{(n)}(i, r_1) P^{(n)}(r_1, r_2) \dots P^{(n)}(r_{k-1}, r)$$

e

$$\sum_{r \in T} P^{(nk)}(i, r) \le (1 - \delta)^k.$$

Ou seja $P^i[X_{nk} \in T] \le (1-\delta)^k$. Como com P^i-probabilidade 1 $[X_{nk+\ell} \in T] \subseteq [X_{nk} \in T]$, $0 \le \ell \le n$, resulta que $P^i[X_{nk+\ell} \in T] \le (1 - \delta)^k$, $0 \le \ell < n$. Quando $k \to \infty$ o membro direito converge a 0, o que prova a proposição.

Corolário 2.4.1. *Se i e j são estados transitórios então $P^{(n)}(i, j) \to 0$ geometricamente. É dizer que existem $c > 0$ e $0 < r < 1$ tal que*

$$P^{(n)}(i, j) \le c r^n.$$

Demonstração: Pela proposição anterior temos que

$$P^{(nk+\ell)}(i, j) \le (1 - \delta)^k \quad i \in T,\ j \in T.$$

Seja $m = nk + \ell$. Temos

$$P^{(m)}(i, j) \le \big(1 - \delta\big)^{\frac{m-\ell}{m}} \le \big(1 - \delta\big)^{\frac{m}{n} - 1} = \frac{1}{(1 - \delta)} \big((1 - \delta)^{1/n}\big)^m.$$

Definindo $c = \dfrac{1}{1 - \delta}$ e $r = (1 - \delta)^{1/n}$ temos o corolário.

Corolário 2.4.2. *Uma cadeia de Markov finita tem pelo menos um estado recorrente.*

2.5 Estudo de algumas cadeias especiais

2.5.1 Cadeia de nascimento e morte

Chamaremos assim, uma cadeia que tem como espaço de estados algum dos conjuntos $\{0,1,2,\ldots,r\}$ ou $\{0,1,\ldots\}$ e como matriz de transição a dada por

$$P(i,j) = \begin{cases} q_i & j = i-1 \\ r_i & j = i \\ p_i & j = i+1 \end{cases}$$

onde $p_i + r_i + q_i = 1$, $q_0 = 0$, e no caso de E ser finito $p_r = 0$. Supomos também que $p_i > 0$ e $q_i > 0$, $\forall\, 0 < i < r$.

Para a e $b \in E$, $a < b$, definimos

$$u(i) = P^i[\tau_a < \tau_b], \quad a < i < b.$$

Definimos $u(a) = 1$ e $u(b) = 0$.

A função u satisfaz a seguinte equação

$$u(j) = q_j\, u(j-1) + r_j\, u(j) + p_j\, u(j+1) \quad a < j < b.$$

Como $r_j = 1 - p_j - q_j$ temos

$$\begin{aligned} u(j+1) - u(j) &= \frac{q_j}{p_j}\,[u(j) - u(j-1)], \quad a < j < b \\ &= \frac{q_j q_{j-1}}{p_j p_{j-1}}\,[u(j-1) - u(j-2)] = \cdots = \\ &= \frac{q_j q_{j-1} \cdots q_{a+1}}{p_j p_{j-1} \cdots p_{q+1}}\,[u(a+1) - u(a)]. \end{aligned}$$

Chamando $\alpha_i = \dfrac{q_1 q_2 \cdots q_i}{p_1 p_2 \cdots p_i}$, $\alpha_0 = 1$, temos que

(I) $$u(j+1) - u(j) = \frac{\alpha_j}{\alpha_a}\,[u(a+1) - u(a)], \quad a \le j < b.$$

Somando sobre j entre a e $b-1$ temos

$$\underbrace{u(b) - u(a)}_{-1} = \frac{u(a+1) - u(a)}{\alpha_a} \left(\sum_{i=a}^{b-1} \alpha_i \right)$$

ou seja

$$\frac{u(a+1)-u(a)}{\alpha_a} = -\frac{1}{\sum\limits_{i=a}^{b-1}\alpha_i}.$$

Substituindo em (I) temos

$$u(j+1)-u(j) = -\frac{\alpha_j}{\sum\limits_{i=a}^{b-1}\alpha_i}, \quad a \le j < b.$$

Somando para j entre i e $b-1$ temos:

$$\underbrace{u(b)-u(i)}_{-u(i)} = -\frac{\sum\limits_{j=i}^{b-1}\alpha_j}{\sum\limits_{i=a}^{b-1}\alpha_i}.$$

Portanto

$$u(i) = \frac{\sum\limits_{j=i}^{b-1}\alpha_j}{\sum\limits_{j=a}^{b-1}\alpha_j} = P^i[\tau_a < \tau_b].$$

Resulta então que a probabilidade do complemento é igual a

$$P^i[\tau_b < \tau_a] = \frac{\sum\limits_{j=a}^{i-1}\alpha_j}{\sum\limits_{j=a}^{b-1}\alpha_j}, \quad a < i < b.$$

Vamos agora estudar quando uma cadeia de nascimento e morte, irredutível sobre $E = \{0,1,2,\dots\}$, é recorrente ou transitória.

Temos que (com relação a P^1)

$$1 \le \tau_2 < \tau_3 < \tau_4 < \cdots$$

e então $P^1[\tau_n \ge n-1] = 1$. Resulta que $P^1[\tau_n \uparrow \infty] = 1$. Temos então

$$P^1[\tau_0 < \infty] = \lim_n P^1[\tau_0 < \tau_n] = \lim_n \left(1 - \frac{1}{\sum\limits_{i=0}^{n-1}\alpha_i}\right) = 1 - \frac{1}{\sum\limits_{i=0}^{\infty}\alpha_i}.$$

Se a cadeia é recorrente $P^1[\tau_0 < \infty] = 1$ e portanto $\sum_{i=0}^{\infty} \alpha_i = \infty$. Reciprocamente, se $\sum_{i=0}^{\infty} \alpha_i = 1$, $P^1[\tau_0 < \infty] = 1$ e $P^0[\tau_0 < \infty] = P(0,0)+ P(0,1)$, $P^1[\tau_0 < \infty] = P(0,0) + P(0,1) = 1$.

Portanto 0 é um estado recorrente e então toda a cadeia é recorrente.

Resumindo, a cadeia é recorrente se e somente se

$$\sum_{i=1}^{\infty} \frac{q_1 q_2 \cdots q_i}{p_1 p_2 \cdots p_i} = \infty.$$

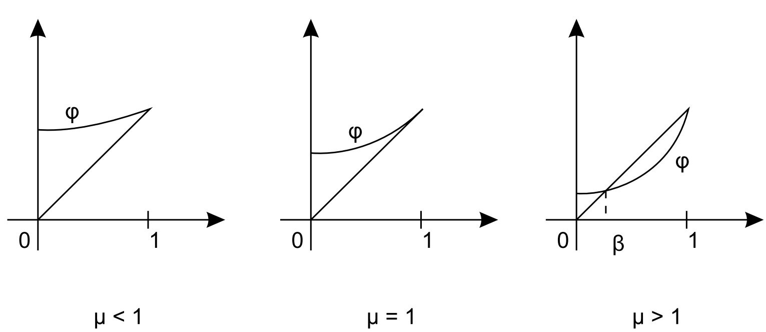
2.5.2 Cadeia de ramificação

Vamos calcular $P^1[\tau_0 < \infty] = \beta$ ou seja a probabilidade de extinção dos descendentes de um indivíduo. Se começarmos com i indivíduos, a probabilidade de extinção será β^i. Para calcular β e para exemplos futuros vamos utilizar o seguinte:

Lema. *Seja $\beta_0, \beta_1, \beta_2, \ldots$ uma distribuição de probabilidades sobre os inteiros não negativos e $\varphi(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots$ a função geradora de probabilidade. Seja μ a esperança da distribuição dada ($0 \leq \mu \leq +\infty$). Suponhamos que $\beta_1 < 1$. Então se $\mu \leq 1$ a equação $\varphi(t) = t$ não tem raízes no intervalo $[0, 1)$. Se $\mu > 1$ a equação $\varphi(t) = t$ tem uma única raiz no intervalo $[0, 1)$.*

Esquema de demonstração:

Os gráficos de φ para os distintos valores de μ são os indicados na figura.



Note que $\varphi(0) = \beta_0$, $\varphi(1) = 1$ e $\lim_{t \to 1} \varphi'(t) = \mu$. Análise de $\varphi'(t)$ em $[0, 1]$ dá a demonstração.

Passamos agora ao cálculo de β. Temos por exercício que

$$\begin{aligned}\beta = P^1[\tau_0 < \infty] &= P(1,0) + \sum_{k=1}^{\infty} P(1,k)P^k[\tau_0 < \infty] \\ &= P(1,0) + \sum_{k=1}^{\infty} P(1,k)\beta^k.\end{aligned}$$

Seja φ a função geradora de probabilidades da distribuição $[P(1,k)]_{k=0,1,\ldots}$. Vamos supor que $P(1,1) < 1$. Temos então que $\beta = \varphi(\beta)$. Ou seja β é raiz da equação $\varphi(t) = t$. Seja μ a esperança dessa distribuição. Então se $\mu \leq 1$, resulta $\beta = 1$ ou seja, temos extinção certa. Se $\mu > 1$, β é raiz da equação $\varphi(t) = t$ no intervalo $[0,1]$. Portanto é igual a 1 ou a um número no intervalo $[0,1)$. Vamos provar que $\beta < 1$. Como as partículas iniciais atuar independentes na formação dos seus descendentes, resulta que a probabilidade $P^i[\tau_0 \leq n]$ de que os descendentes de cada uma das i partículas iniciais tenham desaparecido antes do tempo n, é igual a $(P^1[\tau_0 \leq m])^i$. Portanto

$$\begin{aligned}P^1[\tau_0 \leq n+1] &= P(1,0) + \sum_{k=1}^{\infty} P(1,k)P^k(\tau_0 \leq n) = \\ &= P(1,0) + \sum_{k=1}^{\infty} P(1,k)(P^1(\tau_0 \leq n))^k = \\ &= \beta_0 + \sum_{k=1}^{\infty} \beta_k (P^1(\tau_0 \leq n))^k.\end{aligned}$$

Ou seja

$$P^1(\tau_0 \leq n+1) = \varphi(P^1(\tau_0 \leq n)), \quad n \geq 0.$$

Seja β_0 a raiz em $[0,1)$ de $\varphi(t) = t$.
Agora $P^1[\tau_0 \leq 0] = 0 \leq \beta_0$ e por indução

$$P^1(\tau_0 \leq n+1) = \varphi(P^1(\tau_0 \leq n)) \leq \varphi(\beta_0) = \beta_0\,.$$

Portanto $\forall\, n\, P^1[\tau_0 \leq n] \leq \beta_0$ e passando ao limite $P^1[\tau_0 \leq \infty] = \beta \leq \beta_0 < 1$. Como β tem que ser igual a β_0 ou a 1 resulta $\beta = \beta_0$.

2.5.3 Cadeia fila

Temos pelo Exercício 2.2,c)

$$P^1[\tau_0 < \infty] = P(1,0) + \sum_{k=1}^{\infty} P(1,k)P^k[\tau_0 < \infty]$$

e

$$P^1[\tau_0 < \infty] = P(0,0) + \sum_{k=1}^{\infty} P(0,k)P^k[\tau_0 < \infty].$$

Como para esta cadeia $P(0,k) = P(1,k) = g(k)\ \forall\, k$. Resulta que $P^1[\tau_0 < \infty] = P^0[\tau_0 < \infty]$.

Vamos provar agora que se $\mu \leq 1$ e a cadeia é irredutível, então, é recorrente (condição para que seja irredutível podem ser encontradas nos exercícios). Seja $\beta = P^0[\tau_0 < \infty] = P^1[\tau_0 < \infty]$. Se provarmos que $\varphi(\beta) = \beta$, então, pelo lema $\beta = 1$ o que provaria o resultado. (Note-se que $g(1) < 1$ porque a cadeia é irredutível.)
Temos que

$$P^0[\tau_0 < \infty] = g(0) + \sum_{k=1}^{\infty} g(k)P^k[\tau_0 < \infty].$$

Agora por exercício

$$P^k[\tau_0 < \infty] = P^k[\tau_{k-1} < \infty]P^{k-1}[\tau_{k-2} < \infty] \dots P^1[\tau_0 < \infty].$$

Por outra parte, para $i \geq 1$,

$$\begin{aligned} P^i[\tau_{i-1} = n] &= P\{\omega : \min_{m>0}\{i + (Y_1 - 1) + \cdots + (Y_m - 1) = i - 1\{= n\} \\ &= P\{\omega : \min_{m>0}\{Y_1 + Y_2 + \cdots + Y_m = m - 1\} = n\}. \end{aligned}$$

Ou seja, $P^i[\tau_{i-1} = n]$ é independente de i e portanto $P^i[\tau_{i-1} < \infty]$ é também independente de i. Temos então

$$P^k[\tau_0 < \infty] = \beta^k$$

e

$$\beta = P^0[\tau_0 < \infty] = g(0) + \sum_{k=1}^{\infty} g(k)\beta^k = \varphi(\beta).$$

Vamos estudar agora o caso $\mu > 1$.

Como $\beta = \varphi(\beta)$ temos para β duas possibilidades: $\beta = \beta_0 < 1$ ou $\beta = 1$ onde $\beta_0 = \varphi(\beta_0)$ é a única raiz de $\varphi(t) = t$ em $[0, 1)$. Vamos provar que $\beta = \beta_0$. Temos que (ver exercício)

$$P^1[\tau_0 \leq n+1] = P(1,0) + \sum_{k=1}^{\infty} P(1,k) P^k[\tau_0 \leq n]$$

ou seja

$$\text{(I)} \qquad P^1[\tau_0 \leq n+1] = g(0) = \sum_{k=1}^{\infty} g(k) P^k[\tau_0 \leq n], \quad n \geq 0.$$

Agora por exercício temos que,

$$P^k[\tau_0 \leq n] \leq P^k[\tau_{k-1} \leq n] P^{k-1}[\tau_{k-2} \leq n[\dots P^1[\tau_0 \leq n]$$

como $P^i[\tau_{i-1} \leq n] = P^1[\tau_0 \leq n]$ resulta que

$$P^k[\tau_0 \leq n] \leq (P^1[\tau_0 \leq n])^k.$$

Portanto a equação (I) transforma-se em

$$P^1[\tau_0 \leq n+1] \leq g(0) + \sum_{k=1}^{\infty} g(k) (P^1[\tau_0 \leq n])^k = \varphi(P^1[\tau_0 \leq n]).$$

Como $P^1[\tau_0 \leq 0] = 0 \leq \beta_0$, por indução usando a equação anterior resulta que

$$\forall\, n \geq 0 \quad P^1[\tau_0 \leq n] \leq \beta_0 .$$

Passando ao limite quando $n \to \infty$, temos que

$$\beta = P^1[\tau_0 < \infty] \leq \beta_0$$

o que prova que a cadeia é transitória ($\beta = P^0[\tau_0 < \infty] = \beta_0 < 1$).

2.6 Distribuibuição estacionária de uma cadeia de Markov

Seja P a matriz de probabilidades de transição de uma cadeia de Markov. Seja $\{\pi_i\{_{i \in E}$ uma probabilidade sobre E. Esta distribuição é chamada estacionária se

$$\forall\, j \in E, \quad \sum_{i \in E} \pi_i \, P_{ij} = \pi_j$$

ou em notação matricial
$$\pi' P = \pi.$$
É fácil ver que também $\pi' P^{(n)} = \pi$. (Ver Exercício 2.8.)

Se a distribuição inicial de uma cadeia é uma distribuição estacionária, então a cadeia é um processo estacionário (Exercício 2.9).

Se π é uma idstribuição estacionária e $\lim_n P^{(n)}(i,j) = \pi_j \quad \forall\, i$, então π é a única distribuição estacionária. (Exercício 2.10.)

Vamos agora estudar um exemplo. Seja $\{X_n\}_{n\geq 0}$ uma cadeia de nascimento e morte sobr4e $E = \{0, 1, 2,$
$dots\}$ tal que $p_i > 0$ se $i \geq 0$ e $q_i > 0$ para $i \geq 1$. Vamos resolver o sistema:

$$(\pi_0, \pi_1, \dots) \begin{pmatrix} r_0 & p_0 & 0 & 0 & & \dots & 0 \\ q_1 & r_1 & p_1 & & & \dots & 0 \\ 0 & q_2 & r_2 & p_2 & & \dots & 0 \\ 0 & 0 & q_3 & r_3 & p_3 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \end{pmatrix}$$

Resulta
$$\pi_0 r_0 + \pi_1 q_1 = \pi_0$$
$$p_{i-1}\pi_{i-1} + r_i \pi_i + q_{i+1}\pi_{i+1} = \pi_i\,, \quad i \geq 1$$
Como $q_i + r_i + p_i = 1$ temos
$$q_1 \pi_1 - p_0 \pi_0 = 0$$
$$q_{i-1}\pi_{i+1} - p_i \pi_i = q_i \pi_i - p_{i-1}\pi_{i-1}\,, \quad i \geq 1.$$
Resulta então que
$$q_{i+1}\pi_{i+1} - p_i \pi_i = 0, \quad i \geq 0$$
e portanto,
$$\pi_{i+1} = \frac{p_i}{q_{i+1}}\,\pi_i\,.$$
Desta relação resulta que
$$\pi_i = \underbrace{\frac{p_0 p_1 \cdots p_{i-1}}{q_1 q_2 \cdots q_i}}_{a_i}\,\pi_0\,, \quad i \geq 1.$$

Como $\{\pi_i\}_{i \in E}$ tem que ser uma distribuição sobre E a soma $\sum\limits_{i \in E} \pi_i$ deve ser igual a 1. Temos que

$$\sum_{i \geq 0} \pi_i = \pi_0 + \sum_{i=1}^{\infty} a_i \pi_0 = \pi_0 \big(1 + \sum_{i=1}^{\infty} a_i \big).$$

Uma condição necessária para a existência de distribuição estacionária é que

$$\sum_{i=1}^{\infty} a_i < \infty.$$

A condição é também suficiente porque se $\sum\limits_{i=1}^{\infty} a_i < \infty$, definimos $a_0 = 1$ e

$$\pi_0 = \frac{1}{\sum\limits_{i=0}^{\infty} a_i}$$

e

$$\pi_i = a_i \pi_0 = \frac{a_i}{\sum\limits_{j=0}^{\infty} a_j}, \quad i \geq 1.$$

Vamos agora estudar o comportamento assintótico do número de visitas a um estado j. A notação que usaremos será a do Apêndice 2. Seja $n_j^{(m)} = \sum\limits_{n=1}^{m} I_{\{j\}}(X_n)$, $\quad I_{\{j\}}(X_0) = n_j^{(m)} \underset{m \to}{\longrightarrow} n_j$. Seja $N_{ij}^{(m)} = E^i \big(n_j^{(m)} \big)$.

No Apêndice 2 é provada a seguinte fórmula

$$N_{ij} = H_{ij} \, N_{ij} \, .$$

Se j é transitória $N_{jj} < \infty$ e então

$$N_{ij} = E^i n_j = E^i \Big(\sum_{n=0}^{\infty} I_{\{j\}}(X_n) \Big) = \sum_{n=0}^{\infty} P^{(n)}(i,j) < \infty.$$

Daqui resulta que $P^{(n)}(i,j) \underset{m \to}{\longrightarrow} 0$. Também como n_j e P^i-finita resulta que $n_j^{(m)} \to a$ algo finito $(n_j - I_{\{j\}}(X_0))$ e portanto

$$\frac{n_j^{(m)}}{m} \to 0 \quad \text{(q.t.p. } P^i, \forall\, i).$$

Também
$$E^i\, I_{\{j\}}(X_0) + E^i\, n_j^{(m)} \to E^i\, n_j$$
ou seja
$$\delta_{ij} + N_{ij}^{(m)} \to N_{ij} < \infty.$$

Portanto $\dfrac{N_{ij}^{(m)}}{m} \to 0.$

Note que $\dfrac{N_{ij}^{(m)}}{m} = E^i\big(\frac{n_j^{(m)}}{m}\big).$

Seja agora j recorrente com tempo médio de recorrência $\mu_j = E^j(\bar{\tau}_j)$. Temos o seguinte

Teorema 2.6.1.

a) $\dfrac{n_j^{(m)}}{m} \to \dfrac{I_{[\tau_j \leq \infty]}}{\mu_j}$ (q.t.p. P^i, $\forall\, i$)

b) $\dfrac{N_{ij}^{(m)}}{m} \to \dfrac{H_{ij}}{\mu_j}\cdot$

(Se $\mu_j \to \infty$ o quociente é definido como sendo 0).
Não demonstraremos este Teorema; uma demonstração pode ser encontrada em [6] e em forma mais completa em [1] ou [8].
As duas partes a) e b) são intuitivamente muito razoáveis. Se ao longo de uma trajetória alcançamos o estado j ($\tau_j < \infty$) voltaremos a visitar j em intervalos de tempo de comprimento médio μ_j . Portanto a proporção de tempo que a cadeia está no estado j é $\dfrac{1}{\mu_j}\cdot$ A parte b) decorre de a) tomando esperanças.

Enunciamos agora uma série de proposições; sua demonstração é deixada como exercício.

Proposição 2.6.1. *Se $i \to j$ e i é recorrente positivo então j é recorrente positivo.*

Proposição 2.6.2. *Toda cadeia de Markov finita tem pelo menos um estado recorrente positivo.*

Proposição 2.6.3. *Toda cadeia de Markov finita e irredutível é recorrente positiva.*

Proposição 2.6.4. *Toda cadeia de Markov finita não tem estados recorrentes nulos.*

Proposição 2.6.5. *Se $\{\pi_i\}_{i\in E}$ é uma distribuiçlão estacionária e j é um estado transitório ou recorrente nulo, então $\pi_j = 0$.*

Demonstração: Temos

$$\pi_j = \sum_{i\in E} \pi_i \, P^{(n)}(i,j).$$

Somando para $n = 1, 2, \ldots, n$ resulta

$$m\,\pi_j = \sum_{i\in E} \pi_i \big(\sum_{n=1}^{m} P^{(n)}(i,j) \big).$$

Dividindo por m fica

$$\pi_j = \sum_{i\in E} \pi_i \, \frac{N_{i,j}^{(m)}}{m}.$$

Como j é transitório ou recorrente nulo, $\dfrac{N_{i,j}^{(m)}}{m} \to 0$. Pelo Teorema da Convergência dominada temos que

$$\pi_j \, \lim_m \sum_{i\in E} \pi_i \, \frac{N_{i,j}^{(m)}}{m} = \sum_{i\in E} \pi_i \, \lim_m \frac{N_{i,j}^{(m)}}{m} = 0.$$

A demonstração do seguinte teorema pode ser encontrada em [6].

Teorema 2.6.2. *Uma cadeia de Markov irredutível e recorrente positiva tem uma única distribuição estacionária dada por*

$$\pi_i \, \frac{1}{\mu_1}, \quad i \in E, \quad (\mu_i = E^i(\bar{\tau}_i)).$$

Proposição 2.6.6. *Uma cadeia de Markov irredutível é recorrente positiva se e somente se, tem uma distribuição estacionária.*

Demonstração: Decorre imediatamente do Teorema 2.6.2 e da Proposição 2.6.5.

Enunciaremos finalmente, um teorema que determina o comportamento assintótico de $P^{(n)}(i,j)$. Sua demonstração pode ser encontrada em [8].

Teorema 2.6.3. *Suponhamos j recorrente.*

a) *Se* $\mu_j \to +\infty$, $\lim\limits_{n\to\infty} P^{(n)}(i,j) = 0$.

b) *Se* $\mu_j < \infty$ *e período de* j *é igual a* 1

$$\lim_{n\to\infty} P^{(n)}(i,j) = \frac{H_{ij}}{\mu_j)}.$$

c) *Seja* $\phi_n(i,j) = P^i[\tau_j = n]$ $(\pi_0(i,j) = \delta_{ij}$ *e* $\phi_n(i,j) = 0$ *se* $n \geq 1)$. *Se* $\mu_i \leq +\infty$ *e período de* $j = d$ *então para cada* $r = 0, 1, \ldots, d-1$

$$\lim_{n\to\infty} P^{(nd+r)}(i,j) = \frac{d\big(\sum\limits_{m=0}^{\infty} \phi_{md+r}(i,j)\big)}{\mu_j}.$$

Exercício 2.1. Um jogador faz uma série de apostas de 1 cruzeiro. Ele tem probabilidade de ganhar $\frac{9}{10}$ e de perder $\frac{10}{19}$ em cada jogada. Decide jogar até ganhar 25 cruzeiros ou perder todo o seu capital de 10 cruzeiros. Provar que a probabilidade de ganhar é 0,047 e que sua perda esperada é de 8,36.

Exercício 2.2. Provar

a) $P^i[\tau_j = n+1] = \sum\limits_{k\neq j} P(i,k)P^k[\tau_j = n]$, $n \geq 1$.
(Sugestão: usar propriedade de Markov)

b) $P^i[\tau_j \leq n+1] = P(i,j) + \sum\limits_{k\neq j} P(i,k)P^k[\tau_j \geq n]$, $n \geq 0$

c) $P^i[\tau_j < \infty] = P(i,j) + \sum\limits_{k\neq j} P(i,k)P^k[\tau_j < \infty]$.

Exercício 2.3. Um homem tem exatamente três filhos que independentemente tem probabilidade $\frac{1}{2}$ de ser homem ou mulher. Suponhamos que o número de homens na geração n forme uma cadeia de ramificação. Provar que a probabilidade de extinção é igual a $\sqrt{5} - 2$. Provar que se todo homem tem dois filhos homens e uma mulher a probabilidade de extinção é $4(\sqrt{5} - 2)$.

Exercício 2.4. Na cadeia cola percebe-se que

1) Se $g(0) = 0$ ou $g(0) + g(1) = 1$ a cadeia não é irredutível.

2) Se $g(0) > 0$ e $g(0) + g(1) < 1$ a cadeia é irredutível.

3) Classifique os estados em transitórios recorrentes (e absorventes) quando a cadeia não é irredutível considerando separadamente os casos:

   a) $g(1) = 1$
   b) $g(0) = 1$
   c) $g(0) > 0$, $g(1) > 0$ e $g(0) + g(1) = 1$
   d) $g(0) = 0$ e $g(1) < 1$.

4) Prove que se $\mu > 1$ e a cadeia não é irredutível, estamos no caso d) de 3).

5) Prove que para $i \geq 2$ e $m \geq 1$

$$P^i[\tau_0 = m] = \sum_{k=1}^{m-1} P^i[\tau_{i-1} = k]\, P^{i-1}[\tau_0 = m - k].$$

(Sugestão: usar a propriedade forte de Markov)

6) Somando sobre m em 5) concluir que

$$P^i[\tau_0 < \infty] = P^i[\tau_{i-1} < \infty]\, P^{i-1}[\tau_0 < \infty], \;\; i \geq 2.$$

7) Somando sobre $m = 1, 2, \ldots, n$, provar que

$$P^i[\tau_0 \leq n] \leq P^i[\tau_{i-1} \leq n]\, P^{i-1}[\tau_0 \leq n], \;\; i \geq 2.$$

Exercício 2.5. Considere uma cadeia de nascimento e morte sobre $E = \{0, 1, 2, \ldots\}$

1) Se $\sum_{j=0}^{\infty} \alpha_j = \infty$ então $P^i[\tau_0 \leq \infty] = 1, \;\; \forall\, i \geq 1$.

2) Se $\sum_{j=0}^{\infty} \alpha_j < \infty$ então $P^i[\tau_0 \leq \infty] = \dfrac{\sum_{j=1}^{\infty} \alpha_j}{\sum_{j=0}^{\infty} \alpha_j}$

3) Considere o caso particular onde $q_i = q$, $i = 1, 2, \ldots,$ e $p_i = p$, $i = 1, 2, \ldots$ (cadeia de mina do jogador)

Prove que para $p \leq q$, $P^i[\tau_0 \leq \infty] = \left(\frac{q}{p}\right)^i$, $\; i \geq 1$.

Exercício 2.6. Considere uma cadeia irredutível sobre os inteiros não negativos, tal que $\dfrac{q_i}{p_i} = \left(\dfrac{1}{i+1}\right)^2$, $i \geq 1$. Prove que é transitória. Prove que

$$P^i[\tau_0 < \infty] = 1 - \frac{6}{\pi^2}\left(1 + \frac{1}{4} + \cdots + \frac{1}{i^2}\right).$$

(Sugestão: $\sum_{k=1}^{\infty} \dfrac{1}{k^2} = \dfrac{\pi^2}{6}$).

Exercício 2.7. Seja μ uma probabilidade concentrada nos inteiros e tal que

$$\mu(\{1,2,\dots\}) > 0, \quad \mu(\{-1,-2,\dots\}) > 0$$

e se $F = \{k : \mu(\{k\}) > 0\}$ então m.e.d. $(F) = 1$.
Seja $\{X_i\}_{i=1,2,\dots}$ uma seqüência de r.a.i.i.d. com distribuição μ, e $S_n = \sum_{i=1}^{n} X_i$, $S_0 = 0$, a cadeia de Markov das somas parciais. Provar que esta cadeia é irredutível. Provar que se $m = \sum_k \mu(\{k\})$ existe e $m \neq 0$, então a cadeia é transitória.
(Sugestão: usar a lei forte dos grandes números).
Provar que se $m = 0$ e $\{k : \mu(\{k\}) > 0\}$ é finito então a cadeia é transitória.

Exercício 2.8. Provar que se π é uma distribuição estacionária, então

$$\forall\, n \geq 1, \quad \pi'\, P^{(n)} = \pi.$$

Exercício 2.9. Provar que π é uma distribuição estacionária, e $\mathcal{L}(X_0) = \pi$, então a cadeia $\{X_n\}_{n\geq 0}$ é um processo estacionário.

Exercício 2.10. Se π é uma distribuição estacionária, e $\lim_n P^{(n)}(i,j) = \pi_j$, $\forall\, i$, então π é a única distribuição estacionária.

Exercício 2.11. Seja $\{X_n\}_{n\geq 0}$ uma cadeia de nascimento e morte sobre $E = \{0, 1, \dots, d\}$. Seja

$$a_i = \begin{cases} 1 & i = 0 \\[2ex] \dfrac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i} & i \geq 1 \end{cases}$$

Então a única distribuição estacionária desta cadeia está dada por

$$\pi_i = \frac{a_i}{\sum_{j=0}^{\alpha} a_j}, \quad 0 \leq i \leq d.$$

Exercício 2.12. Provar a Proposição 2.6.1.

Exercício 2.13. Provar a Proposição 2.6.2.

Exercício 2.14. Provar a Proposição 2.6.3.

Exercício 2.15. Provar a Proposição 2.6.4.

Exercício 2.16. No Exemplo 2.1.2 provar que a probabilidade de mina começando com um capital de x é igual a

$$\frac{\left(\frac{q}{p}\right)^x - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^a}.$$

Chamando $\dfrac{q}{p} = r$ temos que esta probabilidade é igual (dividindo por r^a)

$$\frac{1 - r^{-(a-x)}}{1 - r^{-1}} > 1 - r^{-(a-x)}.$$

Desta desigualdade resulta que se o capital do oponente $(a-x)$ é grande a probabilidade de ruina está perto de 1.

3

Cadeias de Markov Finitas

Consideremos uma Cadeia de Markov finita com r estados recorrentes e s estados transitórios. Os r estados recorrentes podem estar agrupados em séries de classes de equivalência. A matriz de probabilidades de transição P pode ser representada em forma conveniente da seguinte forma,

$$P = \begin{pmatrix} p_1 & 0 \\ A & B \end{pmatrix}$$

onde P_1 é $r \times r$, A é $s \times r$ e B é $s \times s$.

Estamos interessados nas seguintes perguntas. Dado que o processo começa no estado transitório i. Qual é o número esperado de visitas a outro estado transitório j antes de entrar em uma das classes recorrentes? Qual é a variânça desse número de visitas? Qual é a probabilidade de que começando no estado transitório i entrar eventualmente no estado recorrente j? etc.

A potência n da matriz P é fácil de se ver que é da forma

$$P^{(n)} = \begin{pmatrix} * & 0 \\ * & B^n \end{pmatrix}$$

onde $*$ indica matrizes nas quais não estamos interessados por enquanto.

3.1 A matriz fundamental

Proposição 3.1.1. *A inversa da matriz $(I - B)$ sempre existe e temos*

$$(I - B)^{-1} = I + B + B^2 + \cdots$$

Demonstração: Temos

$$(I-B)(I+B+B^2+\cdots+B^{n-1})=I-B^n.$$

Pela Proposição 2.4.1, B^n converge a matriz 0. Portanto o determinante $|I-B^n|\to 1$. Ou seja, para $n\geq n_0 \quad |I-B^n|\neq 0$ e portanto

$$|I-B|\,|I+B+B^2+\cdots+B^{n+1}|\neq 0$$

de onde $|I-B|\neq 0$, ou seja $I-B$ tem inversa. Multiplicando por $(I-B)^{-1}$ temos

$$I+B+B^2+\cdots+B^{n-1}=(I-B)^{-1}(I-B^n).$$

Tomando limite quando $n\to\infty$ temos

$$(I-B)^{-1}=I+B+B^2+\cdots$$

A matriz $C=(I+B)^{-1}$ é chamada *matriz fundamental*. Os elementos da matriz C têm uma interpretação probabilística muito importante. Seja T o conjunto dos estados transitórios e i e j elementos de T.

$$\begin{aligned}
C(i,j) &= S_{ij}+B(i,j)+B^2(i,j)+\cdots=\\
&= P^i[X_0=j]+P^i[X_1=j]+P^i[X^2=j]+\cdots\\
&= E^i\,I_{[X_0=j]}+E^i\,I_{[X_1=j]}+E^i\,I_{[X_2=j]}+\cdots\\
&= E^i\{[I_{[X_0=j]}+I_{[X_1=j]}+I_{[X_2=j]}+\cdots\}=\\
&= E^i(n_j)=N_{ij}\,.
\end{aligned}$$

Ou seja, $C(i,j)$ indica o *número médio de vezes que o processo visita o estado transitório j antes de entrar numa classe recorrente, começando no estado i.*

O número $\sum_{j\in T} C(i,j)$ indica para cada estado transitório i o *tempo médio requerido para deixar o conjunto de estados transitórios T.*

Vamos indicar com $\text{var}^i(n_j)$ a variânça de n_j com respeito a probabilidade P^i, $\overline{C}$ vai indicar a matriz que tem como entradas os quadrados dos elementos de C.

Proposição 3.1.2. *Sejam i e j elementos de T,*

$$||\,\text{var}^i(n_j)||=C(2\,C_D-I)-\overline{C}.$$

Demonstração: $\mathrm{var}^i(n_j) = E^i(n_j^2) - (E^i\, n_j)^2$. Como $i \in T$ e $j \in T$ $(E^i\, n_j)^2 = \overline{C}$. Vamos provar agora que $E^i(n_j^2) < \infty$. Como $n_j = \sum\limits_{n=0}^{\infty} I_{\{j\}}(X_n)$ temos que

$$\begin{aligned} E^i(n_j^2) &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} E^i\, I_{\{j\}}(X_{n_1}) I_{\{j\}}(X_{n_2}) = \\ &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} P^i\big[X_{n_1} = j, X_{n_2} = j\big]. \end{aligned}$$

Seja $u = n_1 \wedge n_2$ e $v = |n_1 - n_2|$. Por proposição anterior sabemos que existem $C > 0$ e $0 < r < 1$ tais que $\forall\, i, j \in T$, $P^{(n)}(i,j) \leq C r^n$. Temos então

$$\begin{aligned} P^i\big[X_{n_1} = j, X_{n_2} = j\big] &= P^i[X_u = j] P^i[X_{u+v} = j \mid X_u = j] = \\ &= P^{(n)}(i,j) P^{(v)}(j,j) \leq \\ &\leq C r^u\, C r^v = C^2\, r^{u+v} = C^2\, r^{n_1 \wedge n_2}. \end{aligned}$$

Portanto

$$\begin{aligned} \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} P^i\big[X_{n_1} = j, X_{n_2} = j\big] &\leq C^2 \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} r^{n_1 \wedge n_2} = \\ &= C^2 \sum_{n_1=0}^{\infty} \left[(n_1+1) r^{n_1} + \frac{r^{n_1+1}}{1-r}\right] = \\ &= C^2 \left\{ \sum_{n_1=0}^{\infty} n_1\, r^{n_1} + \sum_{n_1=0}^{\infty} r^{n_1} + \frac{1}{1-r} \sum_{n_1=0}^{\infty} r^{n_1+1} \right\} = \\ &= C^2 \left\{ \frac{r}{1-r} + \frac{r}{1-r} + \frac{1}{1-r}\,\frac{r}{1-r} \right\} = C^2 \left\{ \frac{2r}{1-r} + \frac{1}{1-r} \right\} < \\ &< \infty. \end{aligned}$$

Vamos calcular agora $E^i(n_j^2)$ para i e $j \in T$.

Seja $\bar{n}_j(\omega) = n_j(\theta_1(\omega))$. Temos

$$\begin{aligned} E^i(n_j^2) &= E^i\, n_j^2\, I_{[X_1 \in T]} + E^i\, n_j^2\, I_{[X_1 \in T^c]} = \\ &= E^i\, n_j^2\, I_{[X_1 \in T]} + S_{ij}\, P^i[X_i \in T^c]. \end{aligned}$$

Vamos calcular $E^i\, n_j^2\, I_{[X_1 \in T]}$.

$$\begin{aligned} E^i\, n_j^2\, I_{[X_1\in T]} &= E^i(\bar{n}_j + I_{\{j\}}(X_0))^2\, I_{[X_1\in T]} = \\ &= E^i\, \bar{n}_j^2\, I_{[X_1\in T]} + 2E^i\, \bar{n}_j\, I_{\{j\}}(X_0) I_{[X_1\in T]} + \\ &\quad + E^i\, I_{\{j\}}(X_0) I_{[X_1\in T]}. \end{aligned}$$

Primeiro termo

$$\begin{aligned} E^i\, \bar{n}_j^2\, I_{[X_1\in T]} &= E^i\big[E^i\, \bar{n}_j^2\, I_{[X_1\in T]} \mid X_1\big] = \\ &= E^i\, I_{[X_1\in T]} E^{X_1}\, n_j^2 = E^i \sum_{r\in T} I_{[X_1=r]} E^r\, n_j^2 = \\ &= \sum_{r\in T} P(i,r) E^r\, n_j^2\,. \end{aligned}$$

Segundo termo

$$\begin{aligned} &2E^i \bar{n}_j\, I_{\{j\}}(X_0) I_{[X_1\in T]} = \\ &\quad = 2E^i\, E^i[\bar{n}_j\, I_{\{j\}}(X_0) I_{[X_1\in T]} \mid X_0, X_1] = \\ &\quad = 2E^i\, I_{\{j\}}(X_0) I_{[X_1\in T]} E^i[\bar{n}_j \mid X_0, X_1] = \\ &\quad = 2S_{ij}\, E^i\, I_{[X_1\in T]} E^{X_1}\, n_j = 2S_{ij} \sum_{r\in T} P(i,r) N(r,j). \end{aligned}$$

Terceiro termo

$$E^i\, I_{\{j\}}(X_0) I_{[X_1\in T]} = S_{ij} \sum_{r\in T} P(i,r).$$

Temos então

$$\begin{aligned} E^i(n_j^2) &= \sum_{r\in T} P(i,r) E^r\, n_j^2 + 2S_{ij} \sum_{r\in T} P(i,r) N(r,j) + \\ &\quad + S_{ij} \underbrace{\sum_{r\in T} P(i,r)}_{P^i[X_1\in T]} + S_{ij}\, P^i[X_1 \in T^c] = \\ &= \sum_{r\in T} P(i,r) E^r(n_j^2) + 2S_{ij} \sum_{r\in T} P(i,r) N(r,j) + S_{ij}\,. \end{aligned}$$

Em notação matricial

$$||E^i(n_j^2)|| = B||E^i(n_j^2)|| + 2(BC)_D + I.$$

Portanto

$$\begin{aligned}||E^i(n_j^2)|| &= (I-B)^{-1}\,[2(BC)_D + I] \\ BC &= B[I+B+B^2+B^3+\cdots] = B + B^2 + \cdots = \\ &= C - I.\end{aligned}$$

Também temos que $(C-I)_D = C_D - I$. Portanto

$$||E^i(n_j^2)|| = C[2(C_D - I) + I]$$

e $\mathrm{var}^i(n_j) = C[2C_D - I] - \overline{C}$.

Seja $N = \sum_{j\in T} n_j$ e $N_i = E^i\,N$.

Proposição 3.1.3. $N_i = E^i\,N = E^i\big(\sum_{j\in T} n_j\big) = \sum_{j\in T} E^i(n_j) = \sum_{j\in T} c_{ij}\,.$

Proposição 3.1.4. $||\,\mathrm{var}^i(N)|| = (2C-I)||E^i\,N|| - ||(E^i\,N)^2|| = (2C-I)||N_i|| - ||N_i^2||$.

Demonstração: Começamos como anteriormente, calculando $E^i(N^2)$. Temos

$$\begin{aligned}E^i\,N^2 &= E^i\big(N^2\,I_{[X_1\in T^c]} + N^2\,I_{[X_1\in T]}\big) = \\ &= P^i(X_1\in T^c) + E^i\,I_{[X_1\in T]}(N\circ\theta_1 + 1)^2 = \\ &= P^i[X_1\in T^c] + E^i\,I_{X_1\in T]}E^i((N\circ\theta_1)^2 + 2(N\circ\theta_1) + \\ &\quad + 1 \mid X_0, X_1) = \\ &= P^i[X_1\in T^c] + P^i[X_1\in T] + E^i\,I_{[X_1\in T]}\,E^{X_1}\,N^2 + \\ &\quad + 2E^i\,I_{[X_1\in T]}\,E^{X_1}\,N = \\ &= 1 + \sum_{r\in T} P(i,r)E^r\,N^2 + 2\sum_{r\in T} P(i,r)E^r\,N.\end{aligned}$$

Em notação matricial temos

$$||E^i(N^2)|| = B||E^i(N^2)|| + 2B||E^i\,N|| + 1$$

ou seja

$$\begin{aligned}||E^i(N^2)|| &= (I-B)^{-1}\{2B||E^i\,N|| + 1\} = \\ &= 2CB||E^i\,N|| + ||E^i\,N|| = \\ &= 2(C-I)||E^i\,N|| + ||E^i\,N|| = \\ &= (2C-I)||E^i\,N||.\end{aligned}$$

Agora

$$|| \operatorname{var}^i N|| = (2C - I)||E^i N|| - ||(E^i N)|| = \\ = (2C - I)||N_i|| - ||N_i^2||.$$

Seja $i \in T$ e $j \in T^c$. SEja $V = \{\omega :$ primeira vez que ω visita T^c entre no estado $j\}$.

Seja $d_{ij} = P^i(V)$ e D a matriz dos d_{ij} .

Proposição 3.1.5. $D = CA$.

Demonstração:

$$\begin{aligned} d_{ij} = P^i(V) &= P^i[(V \cap [X_1 \in T]) + (V \cap [X_1 \in T^c]) = \\ &= P(i,j) + P^i(V \cap [X_1 \in T]) = \\ &= P(i,j) + P^i([X_1 \in T] \cap [\theta_1 \in V]) = \\ &= P(i,j) + E^i\, I_{[X_1 \in T]}\, P^i[\theta_1 \in V \mid X_0, X_1] = \\ &= P(i,j) + E^i\, I_{[X_1 \in T]}\, P^{X_1}(V) = \\ &= P(i,j) + \sum_{r \in T} P(i,r) P^r(V). \end{aligned}$$

Ou seja

$$D(i,j) = P(i,j) + (BD)(i,j)$$

ou em notgação matricial

$$D = A + BD.$$

Portanto

$$D - BD = (I - B)D = A$$

e

$$D = (I - B)^{-1}\, A = CA.$$

3.2 O limite de $P^{(n)}$

O objetivo desta seção é estudar o limite de $P^{(n)}$ e algumas conseqüências importantes. Começamos pelo

Lema 3.2.1. *Seja* $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$ *e* P *uma probabilidade sobre* Ω *determinada por* $\{p_i\}_{i=1,2,\ldots,n}$ *onde* $p_j \geq \delta > 0$. *Se* X *é uma variável aleatória e* $M = \max\limits_{1\leq i\leq n} X(\omega_i)$ *e* $m = \min\limits_{1\leq i\leq n} X(\omega_i)$ *então*

$$\forall\, j, \quad X(\omega_j)\delta + m(1-\delta) \leq EX \leq X(\omega_j)\delta + M(1-\delta).$$

Demonstração: Temos que

$$X(\omega_j)I_{\{\omega_j\}} + mI_{\{\omega_j\}^c} \leq X \leq X(\omega_j)I_{\{\omega_j\}^c}$$

portanto

$$X(\omega_j)p_j + m(1-p_j) \leq EX \leq X(\omega_j)p_j + M(1-p_j).$$

Como $p_j \geq \delta > 0$

$$X(\omega_j)p_j + m(1-p_j) \geq X(\omega_j)\delta + m(I-\delta)$$

e

$$X(\omega_j)p_j + M(1-p_j) \geq X(\omega_j)\delta + M(1-\delta).$$

Teorema 3.2.1. *Seja* $\{X_n\}_{n=0,1,\ldots}$ *uma cadeia de Markov com matriz de transição* P. *Se existe* d *inteiro,* $\delta > 0$, *e* $j \in E$ *tal que* $\min\limits_{i\in E} P^{(d)}(i,j) \geq \delta > 0$, *então para toda função* f *definida sobre* E, $E^i f(X_n)$ *converge a uma constante que é independente do estado inicial* i.

Demonstração: Consideremos primeiro o caso $d = 1$. Vamos definir $M_n = \max\limits_{i\in E} E^i f(X_n)$ e $m_n = \min\limits_{i\in E} E^i f(X_n)$ para $n \geq 1$. Resulta que m_n é uma sucessão crescente e M_n é uma sucessão decrescente. Vamos verificar só a afirmação ferente a m_n

$$m_{n+1} = E^i f(X_{n+1}) = \sum_{k\in E} p_{ik}\, E^k f(X_n) \geq \sum_{k\in E} p_{ik}\, m_n = m_n\,.$$

A segunda igualdade decorre da propriedade de Markov.

Vamos estimar agora a diferença $M_n - m_n$. O que vamos provar é a desigualdade

$$M_n - m_n \leq (1-\delta)^n (M-m),$$

onde $M = \max_{i\in E} f(i)$ e $m = \min_{i\in E} f(i)$.

Utilizando o Lema 3.2.1 temos que

$$\forall\, i, \quad f(j)\delta + m(1-\delta) \leq E^i\, f(X_1) \leq f(j)\delta + M(1-\delta).$$

Portanto

$$M_1 - m_1 \leq M(1-\delta) - m(1-\delta) = (M-m)(1-\delta).$$

Agora,

$$E^i\, f(X_2) = \sum_{k\in E} p_{ik}\, E^k\, f(X_1).$$

Usando novamente o lema sobre a variável $k \to E^k\, f(X_1)$, temos que

$$\begin{aligned} E^j\, f(X_1)\delta + \big[\min_{k\in E} E^k\, f(X_1)\big](1-\delta) &\leq E^i\, f(X_2) \leq \\ &\leq E^j\, f(X_1)\delta + \big[\max_{k\in E} E^k\, f(X_1)\big](1-\delta) \end{aligned}$$

e, portanto

$$\begin{aligned} M_2 - m_2 &\leq (1-\delta)\big\{\max_{k\in E} E^k\, f(X_1) - \min_{k\in E} E^k\, f(X_1)\big\} \leq \\ &\leq (1-\delta)(1-\delta)(N-m) = (1-\delta^2)(M-m). \end{aligned}$$

O resultado geral decorre facilmente por indução.

Como $\delta > 0$, temos que $M_n - m_n \underset{n\to\infty}{\longrightarrow} 0$ e, portanto, $E^i\, f(X_n)$ converge a uma constante que não depende de i.

Para o caso geral, $d > 0$, basta observar que $P^{(d)}$ é também uma matriz markoviana. Utilizando o resultado provado até agora podemos concluir que:

$$M_{nd} - m_{nd} \underset{n\to\infty}{\longrightarrow} 0.$$

Como M_n é crescente e m_n decrescente a convergência desta subsucessão implica a convergência de toda a sucessão.

Corolário 3.2.1. *Nas condições do Teorema* 3.2.1*, existem números* $\{\mu_j\}_{j\in E}$ *tal que* $\mu_n \geq 0$*,* $\sum_{j\in E} \mu_j = 1$ *tal que* $\forall\, i \in E$*,* $P^{(n)}(i,j) \to u_j$ *.*

Demonstração: Substituir f por $I_{\{j\}}$ para cada $j \in E$.

$$\mu_j = \lim_{n\to\infty} E^i\, I_{\{j\}}(X_n) = \lim_{n\to\infty} P^i[X_n = j].$$

Logicamente, $\mu_j \geq 0$ e

$$1 = \lim_n \sum_{j \in E} P^i[X_n = j] = \sum_{j \in E} \mu_j \, .$$

Se A é uma matriz $(1 \times n)$ com a propriedade de que $P[X_0 = i] = a_i$, então a distribuição de X_n está determinada pela matriz $AP^{(m)}$. Temos o seguinte:

Corolário 3.2.2. *Nas condições do Teorema* 3.2.1*,*

$$AP^{(m)} \underset{n \to \infty}{\longrightarrow} (\mu_1, \mu_2, \cdots, \mu_n) = \Lambda.$$

Demonstração:

$$\sum_{i=1}^{n} a_i \, P^{(n)}(i,j) \to \sum_{i=1}^{n} a_i \, \mu_j = \mu_j \, .$$

Corolário 3.2.3. *Nas condições do Teorema* 3.2.1*, a matriz* Λ *é a única solução da equação* $A = AP$*, onde os componentes de* A *são não-negativos e soma* 1.

Demonstração: Seja B uma matriz $(n \times n)$ tal que cada uma das filas de B é igual a Λ. Da equação

$$P^{(n+1)} = P^{(n)} P,$$

passando ao limite, temos

$$B = BP.$$

Portanto (tomando uma fila qualquer de B), $\Lambda = \Lambda P$. Se A é outra matriz tal que $A = AP$ pelo Corolário 10.4.2, temos

$$A = AP \overset{\forall n}{=} AP^{(n)} = \lim_n AP^{(n)} = \Lambda.$$

Corolário 3.2.4. *Seja* P *a matriz de probabilidade de transição de uma cadeia de Markov finita, aperiódica e irredutível. Então* $\lim\limits_{n \to \infty} P^{(n)}(i,j) = \pi_j$ *existe, não depende de* i*,* $\pi_j > 0$*, e* $\sum\limits_{j \in E} \pi_j = 1$*.*

Demonstração: Ver exercícios.

3.3 Tempos de primeira passagem

O objetivo desta seção é calculara $E^i(\bar{\tau}_i)$ para todo $i \in E$. Provaremos a seguinte

Proposição 3.3.1. *Seja $\pi = \{\pi_i\}_{i\in E}$ a distribuição limite de uma cadeia de Markov, irredutível e aperiódica. Então*

$$E^i(\bar{\tau}_i) = \frac{1}{\pi_i}\cdot$$

Demonstração:

$$\begin{aligned}
E^i(\bar{\tau}_j) &= E^i\big(\bar{\tau}_j\, I_{[\bar{\tau}_j I_{[\bar{\tau}_j=1]}]} + \bar{\tau}_j\, I_{[\bar{\tau}_j>1]}\big) = \\
&= P^i[\bar{\tau}_j = 1] + E^i\,\bar{\tau}_j\, I_{[\bar{\tau}_j>1]} = \\
&= P(i,j) + E^i\,\bar{\tau}_j\, I_{[\bar{\tau}_j>1]} = \\
&= P(i,j) + E^i\big\{[(\bar{\tau}_j\circ\theta_1)+1]I_{[\bar{\tau}_j>1]}\big\} = \\
&= P(i,j) + P^i[\bar{\tau}_j > 1] + E^i\, I_{[\bar{\tau}_j>1]}E^i(\bar{\tau}_j\circ\theta_i \mid X_0, X_1) = \\
&= 1 + E^i\, I_{[\bar{\tau}_j>1]}E^{X_1}(\bar{\tau}_j) = 1 + \sum_{\substack{r\in E\\ r\neq j}} P(i,r)E^r(\bar{\tau}_j).
\end{aligned}$$

Se chamarmos $h_{ij} = E^i(\bar{\tau}_j)$ temos que

$$h_{ij} = 1 + (Ph)(i,j) - P(i,j)h_{jj} = 1 + (ph)(i,j) - (Ph_D)(i,j).$$

Ou seja, se E é uma matriz composta de 1's.

$$h = E + P\,h - P\,h_D = P(h-h_D) + E$$

multiplicando pelo vetor linha π

$$\pi h = \pi P(h-h_D) + \pi E = \pi(h-h_D) + \pi E$$

ou seja

$$\pi h_D = \pi E$$

de onde resulta

$$\pi_i\, h_{ii} = 1$$

ou seja

$$h_{ii} = E^i(\bar{\tau}_i) = \frac{1}{\pi_i}$$

como queríamos provar.

3.4 Tempos de ocupação

Vamos denotar com $n_j^{(n)} = \sum_{k=1\{j\}} (X_k)$ o número de vezes que o processo visita o estado j no intervalo de tempo $[1,n]$, A fração de tempo que o processo gasta no estado j nesse intervalo está dado por

$$\frac{n_j^{(n)}}{n}.$$

Temos o seguinte resultado.

Proposição 3.4.1. *Dada uma cadeia de Markov finita irredutível e periódica, com distribuição estacionária $\{\pi_j\}_{j\in E}$ temos que*

$$E^i \left(\frac{n_j^{(n)}}{n} - \pi_j \right)^2 \underset{n\to\infty}{\longrightarrow} 0.$$

Em outras palavras $n_j^{(n)}$ converge em L_2 a π_j . Uma conseqüência desta propriedade é que converge em L_1 , e em probabilidade.

Demonstração:

$$\begin{aligned} E^i \left(\frac{n_j^{(n)}}{n} - \pi_j \right)^2 &= E^i \left(\frac{1}{n} \sum_{k=1}^{n} (I_{\{j\}}(X_k) - \pi_j) \right)^2 = \\ &= \frac{1}{n^2} E^i \left\{ \sum_{k=1}^{n} \sum_{\ell=1}^{n} (I_{\{j\}}(X_k) - \pi_j)(I_{\{j\}}(X_\ell) - \pi_j) \right\} = \\ &= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{n} E^i \big(I_{\{j\}}(X_k) - \pi_j\big)\big(I_{\{j\}}(X_\ell - \pi_j)\big). \end{aligned}$$

Agora

$$\begin{aligned} &E^i \big(I_{\{j\}}(X_k) - \pi_j\big)\big(I_{\{j\}}(X_\ell - \pi_j)\big) = \\ &= E^i\, I_{\{j\}}(X_k) I_{\{j\}}(X_\ell) - \pi_j\, P^{(k)}(i,j) - \pi_j\, P^{(\ell)}(i,j) + \pi_j^2 = \end{aligned}$$

Se $u = k \wedge \ell$ e $v = |k-\ell|$ temos

$$= P^{(u)}(i,j) P^{(v)}(i,j) - \pi_j\, P^{(k)}(i,j) - \pi_j\, P^{(\ell)}(i,j) + \pi_j^2\,.$$

Por proposição anterior nós sabemos que

$$P^{(n)}_{(i,j)} = \pi_j + \varepsilon^{(n)}_{ij} \quad \text{onde} \quad |\varepsilon^{(n)}_{ij}| \leq c\, r^n$$

com $c > 0$ e $0 \leq r < 1$.
Substituindo temos que a expressão anterior é igual a

$$\begin{aligned}
(\pi_j + \varepsilon^{(u)}_{ij})(\pi_j + \varepsilon^{(v)}_{ij}) - \pi_j(\pi_j + \varepsilon^{(k)}_{ij}) - \pi_j(\pi_j + \varepsilon^{(\ell)}_{ij}) + \pi_j^2 &= \\
= \pi_j\big\{\varepsilon^{(v)}_{ij} + \varepsilon^{(u)}_{ij} - \varepsilon^{(k)}_{ij} - \varepsilon^{(\ell)}_{ij}\big\} + \varepsilon^{(u)}_{ij}\, \varepsilon^{(v)}_{ij} &\leq \\
\leq c\, r^v + c\, r^u + c\, r^k + c\, r^\ell + c^2\, r^{u+v} &= \\
= c[r^{k\wedge\ell} + r^{|k-\ell|} + r^k + r^\ell - c\, r^{k\wedge\ell}]&.
\end{aligned}$$

Agora

$$\#\{s : s = k \wedge \ell, k = 1, 2, \ldots, n; \ell = 1, 2, \ldots, n\} = n - (k \wedge \ell) + 1 \leq n,$$

e

$$\#\{s : |k - \ell| = s; k = 1, 2, \ldots, n; \ell = 1, 2, \ldots, n\} \leq 2n.$$

Temos então que

$$\begin{aligned}
E^i\left(\frac{n^{(n)}_j}{n} - \pi_j\right)^2 &\leq \frac{1}{n^2}\sum_{k=1}^{n}\sum_{\ell=1}^{n} c\big[r^{k\wedge\ell} + r^{|k-\ell|} + r^k + r^\ell + c\, r^{k\wedge\ell}\big] \leq \\
&\leq \frac{1}{n^2}\, c\big[n\sum_{s=0}^{n} r^s + 2n\sum_{s=0}^{n} n\sum_{s=0}^{n} r^s + c\sum_{s=0}^{n} r^s\big] = \\
&= \frac{cn\big(\sum_{s=0}^{\infty} r^s\big)|1+2+1+1+c|}{n^2} \underset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

Exercício 3.1. Seja P a matriz de probabilidades de transição de uma cadeia de Markov com m estados, e seja $|p - \lambda I| = 0$ a equação característica da matriz P. Usando a propriedade que a matriz satisfaz sua própria equação característica, prove que

$$P^n = a_1\, P^{n-1} + a_2\, P^{n-2} + \cdots + a_m\, P^{n-m}$$

onde

$$a_i = -\frac{a'_{m-i}}{a'_m}\,,\ i = 1, 2, \ldots, m$$

e
$$|P - \lambda I| = a'_m\, P^m + a'_{m-1}\, P^{m-1} + \cdots + a'_0 = 0.$$

Exercício 3.2. Consideremos uma cadeia com 2 estados ($m = 2$) com matriz de transição
$$P = \begin{pmatrix} 1-a & a \\ 0 & 1-b \end{pmatrix}.$$
Então
$$P^n = (2-a-b)P^{n-1} - (1-a-b)P^{n-2}.$$

Exercício 3.3. Se $P^{(n)}(j,i) > 0$ existe m_0 tal que
$$\forall\, m \geq m_0\,, \quad P^{(n+md_i)}(j,i) > 0.$$

Exercício 3.4. Nas condições do Teorema 3.2.1, provar que se $P^{(n)}(i,j) = \pi_j + \varepsilon_{ij}^{(n)}$, existe $c > 0$ e $0 < r < 1$ tal que
$$|c_{ij}^{(n)}| \leq c\, r^n$$
se a matriz P não tem elementos nulos, então $c = 1$.

Solução: Na demonstração do Teorema 3.2.1, para o caso $d = 1$ provamos que
$$M_m - m_n \leq (1-\delta)^n (M - M).$$
Como $M - m \leq 1$ e $|c_{ij}^{(n)}| \leq M_n - m_n \leq (1-\delta)^n$ basta tomar $r = 1 - \delta$. Para o caso $d > 1$, a mesma demonstração de
$$M_{sd} - m_{sd} \leq (1-\delta)^s\, (M - m).$$
Se $0 \leq \ell \leq d$
$$M_{sd+\ell} - m_{sd+\ell} \leq (1-\delta)^s\, (M - m).$$
Definindo $n = sd + \ell$ temos $s = \dfrac{n-\ell}{d}$ e
$$M_n - m_n \leq (1-\delta)^{\frac{n}{d}}\, (1-\delta)^{-\frac{\ell}{d}}\, (M - m).$$
Como $\dfrac{\ell}{d} \leq 1$ temos
$$M_m - m_n \leq \big(\frac{1}{1-\delta}\big)(M - m)[(1-\delta)^{1/d}]^n.$$

Em nosso caso $M - m \leq 1$. Chamando $c = \dfrac{1}{1-\delta}$ e $r = (1-\delta)^{\frac{n}{d}}$ temos o resultado.

Exercício 3.5. Provar que toda cadeia de Markov finita tem pelo menos um estado recorrente positivo.

Exercício 3.6. Provar que a propriedade de um estado ser recorrente positivo, é uma propriedade de classe.

Exercício 3.7. Seja $\{X_n\}_{n=0,1,2,\dots}$ uma cadeia de Markov finita com matriz de probabilidades de transição P. Seja $R(i,j)$ uma certa recompensa associada à transição $i - j$. Seja $V^{(n)} = \sum_{k=0}^{n-1} R(X_k, X_{k+1})$ a recompensa total depois de n transições.

Se $V_i^{(n)} = E^i\, V^{(n)}$, provar que

$$V_i^{(n)} = \sum_{j \in E} P_{ij}(E_{ij} + V_j^{(n-1)}), \quad i \in E, \quad n = 1, 2, \dots$$

Exercício 3.8. Seja $\{X_n\}_{n=0,1,\dots}$ uma cadeia de Markov finita com distribuição estacionária com todos seus elementos positivos. Provar que esta cadeia é irredutível, aperiódica e recorrente positiva.

4

Cadeias de Markov e Parâmetro Contínuo

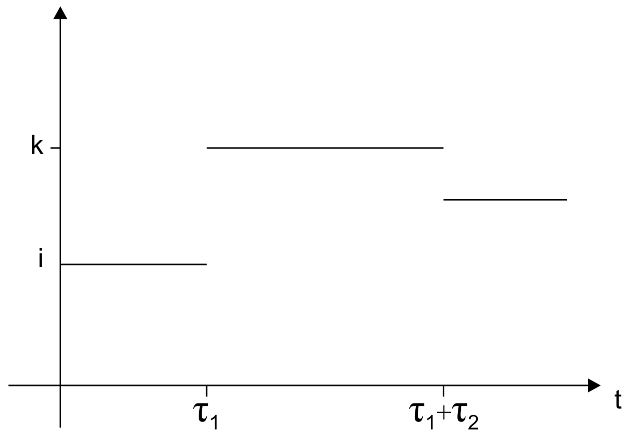
4.1 Esboço de sua construção

Seja E um conjunto enumerável (espaço de estados), $Q(i,j)$ uma matriz estocástica definida em $E \times E$ tal que $Q(i,i) = 0$, $\forall\, i \in E$ e $\{\lambda(i)\}_{i\in E}$ uma seqüência de números > 0. Uma distribuição exponencial com parâmetro $\lambda(i)$ e uma distribuição em $\mathbb{R}^1$ com densidade dada pela fórmula

$$f(t) = \lambda(i)e^{-\lambda(i)t}\, I_{[0,+\infty]}(t).$$

Uma trajetória "típica"do processo que queremos construir é obtida da seguinte forma. Suponhamos que começamos o processo no estado $i \in E$. Esperamos um tempo τ_1 com uma distribuição exponencial com parâmetro $\lambda(i)$. No instante de tempo τ_1 , pulamos para outro estado usando a distribuição $Q(i, \cdot\,)$. Suponhamos que passamos ao estado $k \in E$. Esperamos agora um tempo τ_2 exponencial com parâmetro $\lambda(k)$. Ao fim deste intervalo de tempo pulamos para outro estado usando a distribuição $Q(k, \cdot\,)$, etc.

O início desta trajetória está indicada na seguinte figura



$X_t(\omega)$ indicará como sempre o estado do processo no instante de tempo t.

Passamos a descrever rapidamente a construção formal. Seja $q(j,k) = \lambda(j)\, Q(j,k)$. Temos $q(j,k) \geq 0$, $q(i,j) = 0$, $\sum\limits_k q(j,k) = \lambda(j)$.

Seja $X_0, X_1, \ldots$ uma cadeia de Markov com espaço de estados E e matriz de probabilidades de transição $Q(j,k)$. Para cada $i \in E$ e para cada $n \geq 0$ seja $\tau_n(i)$ uma variável aleatória com distribuição exponencial com parâmetro $\lambda(i)$. A construção é feita sobre um espaço de probabilidades de maneira que $(X_0, X_1, \ldots)$ e $\{\tau_n(i)\}_{\substack{i\in E\\ n=0,s,\ldots}}$ sejam conjuntos de v.a. independentes.

Definamos $R_0 = 0$ e $R_n = \tau_0(X_0) + \cdots + \tau_{n-1}(X_{n-1})$.

Definamos $\tilde{n}(t,n) = n$ se $R_n(\omega) \leq t < R_{n+1}(\omega)$.

Finalmente, seja

$$X(t,\omega) = X_{\tilde{n}(t,\omega)}(\omega).$$

Temos o seguinte resultado:

Teorema 4.1.1. *Se $\tilde{n}$ está bem definida (ou seja $R_n \underset{n\to\infty}{\longrightarrow} \infty$ com P^i-probabilidade 1, $\forall\, i$) então $X(t,\omega)$ é uma cadeia de Markov com parâmetro contínuo que satisfaz as condições infinitesimais*

$$P^j(X_t) = k) = P_t(j,k) = q(j,k)t + 0(t), \quad k \neq j$$
$$P^j(X_t = j) = P_t(j,j) = 1 - \lambda(j)t + 0(t).$$

A continuação deixamos sem demonstração uma condição necessária e suficiente para que $\tilde{n}$ esteja bem definida.

Proposição 4.1.1. *$\tilde{n}$ está bem definida se e somente se $\forall\, i \in E$, $P^i\big(\sum_{n=0}^{\infty} \dfrac{1}{\lambda(X_n)} < \infty\big) = 0$.*

Um corolário imediato é o seguinte.

Corolário 4.1.1. *Se* $\sup \lambda(j) < \infty$ *então $\tilde{n}$ está bem definida.*

Vamos ver um pouco mais adiante que para um *processo de nascimento puro* $X_n = n$ e portanto, a condição necessária e suficiente é neste caso

$$\sum_{n=0}^{\infty} \frac{1}{\lambda(n)} = \infty.$$

Para o *processo de Poisson* ($\lambda(n) = \lambda, \forall\, n$) e para o *Processo de Yule* ($\lambda(n) = n\lambda$, $\lambda > 0$) esta condição é claramente satisfeita.

Da propriedade de Markov sai

$$P^i(X_t = k, X_{t+s} = j) = P^i(X_t = k)P^k(X_s = j) = P_t(i,k)P_s(k,j).$$

Portanto

$$P_{t+s}(i,j) = P^i(X_{t+s} = j) = \sum_{k \in E} P^i(X_t = k, X_{t+s} = j) =$$
$$= \sum_{k \in E} P_t(i,k)P_s(k,j).$$

Ou em notação matricial

$$P_{t+s} = P_t P_s\,.$$

Esta identidade é conhecida como *equação de Chapman-Kolmogorov.*

4.2 Equações para frente e para trás

A propriedade forte de Markov poderia ser utilizada para provar a seguinte identidade que intuitivamente deveria ser satisfeita

$$P^i(\tau_1 \le t, X_{\tau_1} = k, X_t = j) = \int_0^t \lambda(i)e^{-\lambda(i)s}\, Q_{ik}\, P_{t-s}(k,j)\, ds.$$

Somando sobre k temos

$$P^i(\tau_1 \le t, X_t = j) = \int_0^t \lambda(i)e^{-\lambda(i)s}\big(\sum_{k \ne i} Q_{ik}P_{t-s}(k,j)\big)\, ds$$

Agora

$$P^i(\tau_1 > t, X_t = j) = \delta_{ij}\, e^{-\lambda(i)t}.$$

Portanto $P_t(i,j)$ satisfaz a seguinte equação integral

$$P^i(X_t = j) = \delta_{ij}\, e^{-\lambda(i)t} +$$
$$+ \int_0^t \lambda(i) e^{-\lambda(i)s} \big(\sum_{k \neq i} Q_{ik}\, P_{t-s}(k,s) \big)\, ds.$$

Decorre desta expressão que $P_t(i,j)$ é uma função contínua de t. Mudando de variável $(t-s-s)$ temos

$$P_i(X_t = j) = \delta_{ij}\, e^{-\lambda(i)t} +$$
$$+ \lambda(i) e^{-\lambda(i)t} \int_0^t e^{\lambda(i)s} \big(\sum_{k \neq i} Q_{ik}\, P_s(k,j) \big)\, ds.$$

Derivando com respeito a t temos

$$\text{(I)} \qquad P_t'(i,j) = -\lambda(i)\delta_{ij}\, e^{-\lambda(i)t} +$$
$$+ \lambda(i) e^{-\lambda(i)t}\, e^{\lambda(i)t} \sum_{k \neq i} Q_{ik}\, P_t(k,j) -$$
$$- \lambda^2(i) e^{-\lambda(i)t} \int_0^t e^{\lambda(i)s} \big(\sum_{k \neq i} Q_{ik}\, P_s(k,j) \big)\, ds =$$
$$= -\lambda(i) P_t(i,j) + \lambda(i) \sum_{k \neq i} Q_{ik}\, P_t(k,j).$$

Em particular

$$P_0'(i,j) = -\lambda(i) P_0(i,j) + \lambda(i) \sum_{k \neq i} Q_{ik}\, P_0(k,j) =$$
$$= -\lambda(i)\delta_{ij} + \lambda(i) \sum_{k \neq i} Q_{ik}\, \delta_{kj} =$$
$$= -\lambda(i)\delta_{ij} + \lambda(i) Q_{ij}\,.$$

Chamando $q(i,j) = P_0'(i,j)$ temos

$$q(i,j) = \begin{cases} -\lambda(i) & i = j \\ \lambda(i) Q(i,j) & i \neq j \end{cases}$$

Portanto

$$\sum_{j\neq i} q(i,j) = \lambda(i) = -q(i,i).$$

Vamos chamar a matriz $A = ||q(i,j)||$ de *operador infinitesimal* e aos elementos de A *parâmetros infinitesimais.*

Usando estas quantidades, a equação (I) transforma-se em

$$P_t'(i,j) = \sum_{k\in E} q(i,k)P_t(k,j), \quad t \geq 0$$

ou em notação matricial

$$P_t' = A\,P_t\,, \quad \forall\, t \geq 0.$$

Esta equação é chamada, *equação para trás.* Consideremos agora a equação de Chapman-Kolmogorov

$$P_{t+s}(i,j) = \sum_{k\in E} P_t(i,k)P_s(k,j).$$

Se E for finito podemos derivar com relação a s obtendo

$$P_{t+s}'(i,j) = \sum_{k\in E} P_t(i,k)P_s'(k,j).$$

Para $t = 0$, temos

$$P_t'(i,j) = \sum_{k\in E} P_t(i,k)P_0'(k,j).$$

Em notação matricial

$$P_t' = P_t A, \quad \forall\, t \geq 0.$$

Esta equação é chamada *equação para frente.* Esta equação é válida também para o caso em que E tenha um número enumerável de estados se por exemplo, o processo não tem explosões (ou equivalentemente as trajetórias têm só um número finito de saltos em cada intervalo finito de tempo).

4.3 Exemplos

4.3.1 Processo de nascimentos e morte

Seja $E = \{0, 1, 2, \dots\}$. Um *processo de nascimento e morte* sobre E é uma cadeia de Markov e parâmetro contínuo e tal que $q(i, j) = 0$ se $|i - j| > 1$. Ou seja, um processo de nascimento e morte que começa em i pode ir em um salto a $i + 1$ ou a $i - 1$. Chamaremos $\lambda_i = q(i, i + 1)$ e $\mu_i = q(i, i - 1)$ com $\mu_0 = 0$. λ_i e μ_i são chamados *taxa de nascimento* e *taxa de mortalidade* respectivamente. Temos

$$\lambda(i) = -q(i, i) = q(i, i + 1) + q(i, i - 1) = \lambda_i + \mu_i\,.$$

Então i é absorvente se e somente se $\lambda_i = \mu_i = 0$. Se i é não absorvente

$$Q(i, j) = \begin{cases} \dfrac{\lambda_i}{\lambda_i + \mu_i} & j = i + 1 \\ \\ \dfrac{\mu_i}{\lambda_i + \mu_i} & j = i - 1 \end{cases}$$

Se $\lambda_i = 0$, $\forall\, i$, o processo é chamado de *morte pura*. Se $\mu_i = 0$, $\forall\, i$, o processo é chamado de *nascimento puro*. Outro espaço de estados de utilidade para considerar processos de nascimento e morte, é o espaço $E = \{0, 1, 2, \dots, r\}$. Condições suficientes para a existência de processos de nascimento e morte, podem ser encontradas nos exercícios. Uma condição necessária é evidentemente $\mu_0 = 0$ e para o caso E finito $\lambda\, r = 0$.

4.1.2 Processo de Poisson

O processo de Poisson é um processo de nascimento puro sobre $E = \{0, 1, \dots\}$ com taxa de nascimento $\lambda(i) = \lambda > 0$, $\forall\, i \in E$. Temos $P^i(X_t = i) = P_t(i, i) = e^{-\lambda t}$. Consideremos a equação para frente $(P' = PA)$

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & \dots \\ 0 & -\lambda & \lambda & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Se $j > 0$ temos

$$P'_{ij}(t) = -\lambda\, P_{ij}(t) + \lambda\, P_{ij-1}(t).$$

A solução desta equação (ver Apêndice ....) é

$$P_{ij}(t) = \underbrace{P_{ij}(0)}_{\delta_{ij}} e^{-\lambda t} + \lambda \int_0^t e^{-\lambda(t-s)} P_{ij-1}(s)\, dx.$$

Se $j > i$ temos

$$P_{ij}(t) = \lambda \int_0^t e^{-\lambda(t-s)} P_{ij-1}(s)\, ds.$$

Seja $j = i + 1$. Então

$$P_{ii+s}(t) = \lambda \int_0^t e^{-\lambda(t-s)}\, e^{-\lambda s}\, ds = \lambda \int_0^t e^{-\lambda t}\, ds = \lambda t\; e^{-\lambda t}.$$

Seja $j = i + 2$. Temos

$$\begin{aligned} P_{ii+2} &= \lambda \int_0^t e^{-\lambda(t-s)}\, (\lambda s) e^{-\lambda s}\, ds = \\ &= \lambda^2 \int_0^t e^{-\lambda t}\, s\, ds = \lambda^2\, e^{-\lambda t} \int_0^t s\, ds = \\ &= \frac{\lambda^2 t^2}{2}\, e^{-\lambda t} = e^{-\lambda t}\, \frac{(\lambda t)^2}{2}. \end{aligned}$$

Por indução se $j > i$

$$P_{ij}(t) = \frac{e^{-\lambda t}\, (\lambda t)^{j-i}}{(j-1)!}.$$

Note-se que $P_{ij}(t) = P_{0(j-i)}(t)$, $t \geq 0$; e que se $X_0 = i$ então $\mathcal{L}(X_t - 1) =$ Poisson (λt). Em geral se $0 \leq s < t$, $\mathcal{L}(X_t - S_s) =$ Poisson $\lambda(t - s)$. Com efeito,

$$\begin{aligned} P^i(X_t - X_s = k) &= \sum_{\ell=0}^{\infty} P^i(X_s = \ell, X_t = \ell + k) = \\ &= \sum_{\ell=0}^{\infty} P^i(X_s = \ell) P^{\ell}(X_{t-s} = \ell + k) = \\ &= \sum_{\ell=0}^{\infty} P^i(X_s = \ell) P_{t-s}(0, k) = P_{t-s}(0, k) = \\ &= P^0(X_{t-s} = k) = \frac{[\lambda(t-s)]^k\, e^{-\lambda(t-s)}}{k!}. \end{aligned}$$

O Processo de Poisson é um processo com incrementos independentes. Sejam $0 \leq t_1 < t_2 < \cdots < t_n$. Temos que provar que $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ são v.a. i. Temos

$$
\begin{aligned}
&P^i(X_{t_2} - X_{t_1} = j_1, \ldots, X_{t_n} - X_{t_{n-1}} = j_{n-1}) = \\
&= \sum_{\ell=0}^{\infty} P^i(X_{t_1} = \ell, X_{t_2} - X_{t_1} = j_1, \ldots, X_{t_n} - S_{t_{n-'}} = j_{n-1}) = \\
&= \sum_{\ell=0}^{\infty} P^i(X_{t_1} = \ell) P^{\ell}(X_{t_2 - t_1} = \ell + j_1) P^{j_1+\ell}(X_{t_3 - t_2} = j_2 + j_1 + \ell) = \\
&= \sum_{\ell=0}^{\infty} P^i(X_{t_1} = \ell) P_{t_2 - t_1}(0, j_1) P_{t_3 - t_2}(0, j_2) \ldots P_{t_n - t_{n-1}}(0, j_{n-1}) = \\
&= P_{t_2 - t_1}(0, j_1) P_{t_3 - t_2}(0, j_2) \ldots P_{t_n - t_{n-1}}(0, j_{n-1}) = \\
&= P^i(X_{t_2} - X_{t_1} = j_1) \ldots P^i(X_{t_n} - X_{t_{n-1}} = j_{n-1}).
\end{aligned}
$$

4.4 Algumas propriedades de cadeias de Markov a parâmetro contínuo. Distribuições estacionárias.

Seja $\{X_t\}_{t \geq 0}$ uma cadeia de Markov a parâmetro contínuo com espaço de estados E.

Seja τ_1 o tempo do primeiro pulo ($\tau_1 = +\infty$ se a trajetória é constante). Seja para $i \in E$

$$\tau_i(\omega) = \min\{t : t \geq \tau_1(\omega) \text{ e } X_t(\omega) = i\}.$$

Se $\tau_1(\omega) = +\infty$ ou $\forall\, t \geq \tau_1(\omega)$, $X_t(\omega) \neq i$, definimos $\tau_i(\omega) = +\infty$. $\tau_i(\omega)$ indica o tempo da primeira visita ao estado i depois do primeiro pulo. Suponhamso que i é não absorvente ($\lambda(i) > 0$). Dizemos então que i é *recorrente* (resp. *transitório*) se $P^i[\tau_i < \infty] = 1$ (resp. $P^i[\tau_i < \infty] < 1$). Uma cadeia será chamada recorrente (resp. transitória) se todos os seus estados são recorrentes (resp. transitórios).

Se para todo par $i, j \in E$

$$P^i[\tau_j < \infty] > 0$$

dizemos que a cadeia é *irredutível*.

A cadeia de Markov a tempo discreto determinada pela matriz estocástica Q é chamada *cadeia mergulhada.* Um estado da cadeia de Markov é recorrente, transitório ou a cadeia é irredutível, se e somente se, o estado é recorrente ou transitório na cadeia mergulhada é irredutível

$$(P^i[\tau_j < \infty] = P^i[\bar{\tau}_j < \infty]).$$

Uma seqüência $\{\pi_i\}_{i \in E}$ de números não negativos e tais que $\sum\limits_{i \in E} \pi_i = 1$ é uma distribuição estacionária para a cadeia se

$$\forall\, j \in E, \quad \forall\, t \geq 0, \quad \sum_{i \in E} \pi_i \, P_t(i,j) = \pi_j \, .$$

Em notação matricial

$$\text{(I)} \qquad \pi' \, P_t = \pi, \quad \forall\, t \geq 0 \quad (\pi' = \text{ transposta de } \pi).$$

Note-se que se fosse possível derivar, teríamos

$$\pi' \, P_t' = 0.$$

Fazendo $t = 0$ temos a relação

$$\text{(II)} \qquad \pi' \, A = 0$$

Pode ser provado que a relação (II) é válida se e somente se (I) é válida.

Um estado i é chamado *recorrente positivo* (resp. *recorrente nulo*) se $\mu_i = E^i$, $\tau_i < \infty$ (resp. $\mu_i = E^i$, $\tau_i = +\infty$). Como anteriormente um processo é chamado recorrente positivo (resp. recorrente nulo) se todos os seus estados são recorrentes positivos (resp. recorrente nulo). Para uma cadeia irredutível, todos os estados são do mesmo tipo. Também toda distribuição estacionária está concentrada nos estados recorrentes positivos e portanto, um processo que é transitório ou recorrente nulo não tem distribuição estacionária. Uma cadeia irredutível recorrente positiva tem uma única distribuição estacionária, dada por

$$\pi_i = \frac{1}{\lambda(i)\mu(i)}\, , \quad i \in E.$$

Esta idnetidade é bem razoável, porque em um intervalo de tempo t grande, o processo faz $\dfrac{t}{\mu_i}$ visitas ao estado i. Em cada visita ele demora em média $\frac{1}{\lambda(i)}\cdot$ Portanto o total de tempo que passa no estado i é $\frac{t}{\mu_i\lambda(i)}$ em proporção $\frac{1}{\mu_i\lambda(i)}\cdot$

Para toda cadeia irredutível e recorrente positiva com distribuição estacionária $\{\pi_i\}_{i\in E}$

$$\lim_{t\to\infty} P_t(i,j) = \pi_j\,.$$

4.5 Processo de nascimento e morte

Consideremos um processo de nascimento e morte irredutível. Este processo é transitório se e somente se a cadeia de nascimento e morte mergulhada é transitória. Como $p_i = Q(i,i+1) = \frac{\lambda_i}{\lambda_i+\mu_i}$, $Q(i,i-1) = \frac{\mu_i}{\lambda_i+\mu_i} = q_i$ esta cadeia é transitória, se e somente se

$$\sum_{i=1}^{\infty} \frac{q_1\dots q_i}{p_1\dots p_i} < \infty$$

ou seja, neste caso se e somente se

$$\sum_{i=1}^{\infty} \frac{\mu_1\dots\mu_i}{\lambda_1\dots\lambda_i} < \infty.$$

Para calcular a distribuição estacionária, resolveremos a equação $\pi A = 0$ onde

$$A = \begin{pmatrix} -\lambda(0) & \lambda_0 & & & \\ \mu_1 & -\lambda(1) & & & \bigcirc \\ & \mu_2 & -\lambda(2) & \lambda_2 & \\ \bigcirc & & & & \end{pmatrix}$$

Temos então

(I) $$\pi_1\,\lambda_1 - \lambda_0\,\pi_0$$

e

$$\pi_{i-1}\,\lambda_{i-1} + \pi_i(-\lambda_i - \mu_i) + \pi_{i+1}\,\mu_{i+1} = 0,\ \ i \ge 1$$

$$\pi_{i+1}\,\mu_{i+1} - \pi_i\,\lambda_i = \pi_i\,\mu_i - \pi_{i-1}\,\lambda_{i-1}\,.$$

Usando (I) e indução resulta que

$$\pi_{i+1}\,\mu_{i+1} = \pi_i\,\lambda_i\,, \;\; i \geq 1$$

e portanto

$$\pi_{i+1} = \left(\frac{\lambda_i}{\mu_{i+1}}\right)\pi_i\,, \quad i \geq 0.$$

Desta equação resulta que

(II) $$\pi_i = \frac{\lambda_0\lambda_1\ldots\lambda_{i-1}}{\mu_1\mu_2\ldots\mu_i}\,\pi_0\,, \quad i \geq 1.$$

Chamando

$$\tilde{\pi}_i = \begin{cases} 1 \\ \dfrac{\lambda_0\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i} & i \geq 1 \end{cases}$$

A equação (II) fica então

$$\pi_i = \tilde{\pi}_i\,\pi_0\,, \quad i \geq 0.$$

Se $\sum\limits_{i=0}^{\infty} \tilde{\pi}_i \leq \infty$, ou seja

(III) $$\sum_{i=1}^{\infty} \frac{\lambda_0\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i} < \infty$$

o processo tem uma única distribuição estacionária dada por

$$i \longrightarrow \frac{\tilde{\pi}_i}{\sum\limits_{j=0}^{\infty}\tilde{\pi}_j}\,, \quad i \geq 0.$$

Se (III) não se satisfaz o processo não tem distribuição estacionária.

Resumindo, se

$$\sum_{i=1}^{\infty} \frac{\mu_1\ldots\mu_i}{\lambda_1\ldots\lambda_i} < \infty$$

o processo é transitório.
Se

$$\sum_{i=1}^{\infty} \frac{\lambda_0\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i} < \infty$$

o processo é recorrente positivo.
Se

$$\sum_{i=1}^{\infty} \frac{\mu_1 \dots \mu_i}{\lambda_1 \dots \lambda_i} = \infty \quad \text{e} \quad \sum_{i=1}^{\infty} \frac{\lambda_0 \dots \lambda_{i-1}}{\mu_1 \dots \mu_i} = \infty$$

o processo é recorrente nulo.

Se o processo irredutível de nascimento e morte tivesse como espaço de estados $E = \{0, 1, \dots, r\}$ seria necessariamente recorrente positivo. Tem uma única distribuição estacionária, dada por

$$i \longrightarrow \frac{\tilde{\pi}_i}{\sum\limits_{j=0}^{r} \tilde{\pi}_j}, \quad 0 \leq i \leq r.$$

Exercício 4.1. Se $\lambda(i)$ e $\mu(i)$, $i \in [0, 1, 2, \dots)$, são $\mu(0) = 0$, são seqüências de números não negativos e $\lambda(i) \leq A + Bi \quad (A \geq 0, B \geq 0)$ então existe um processo de nascimento e morte que tem estas duas seqüências como parâmetros infinitesimais.

Exercício 4.2. Sejam $X_1, X_2, \dots, X_n$, n v.a.i.i.d., $\mathcal{L}(X_i) = \exp(\lambda(i))$

a) $\lambda\big(\min\limits_{1 \leq i \leq n}(X_1, \dots, X_n)\big) = \exp\big(\sum\limits_{i=1}^{n} \lambda(i)\big)$

b) $P(X_k = \min(X_1, \dots, X_n)) = \dfrac{\lambda(k)}{\sum\limits_{i=1}^{n} \lambda(i)}$

c) $P\big(\sum\limits_{i \neq j} [X_i = X_j]\big) = 0.$

Exercício 4.3. Seja $\{X_t\}_{t \geq 0}$ um processo de nascimento e morte sobre $E = \{0, 1, 2, \dots\}$ e tal que $\lambda_i > 0$ e $\mu_i > 0, \quad \forall\, i \geq 1$. Seja

$$\alpha_i = \frac{\mu_1 \dots \mu_i}{\lambda_1 \dots \lambda_i} \quad i \geq 1, \;\; \alpha_0 = 1.$$

Provar:

a) Se $\sum\limits_{i=1}^{\infty} \alpha_i = \infty \Rightarrow P^i[\tau_0 < \infty] = 1$, $i \geq 1$

b) Se $\sum\limits_{i=1}^{\infty} \alpha_i \leq \infty \Rightarrow P^i[\tau_0 < \infty] = \dfrac{\sum\limits_{j=i}^{\infty} \alpha_j}{\sum\limits_{j=0}^{\infty} \alpha_j}$, $i \geq 1$.

5

Outros Processos Estocásticos

Neste capítulo vamos introduzir informalmente o processo de Wiener (Movimento Browniano) e enunciar e provar alguma das suas propriedades. Também introduziremos a noção de processo gaussiano e daremos alguns exemplos.

5.1 Movimento Browniano

Suponhamos que ao tempo 0, você observa através de um microscópio uma partícula de um líquido ou gás que nesse instante se encontra na origem de um sistema de coordenadas em três dimensões. Se por exemplo a primeira coordenada é representada em função do tempo, obteríamos um gráfico como o da figura 5.1.1

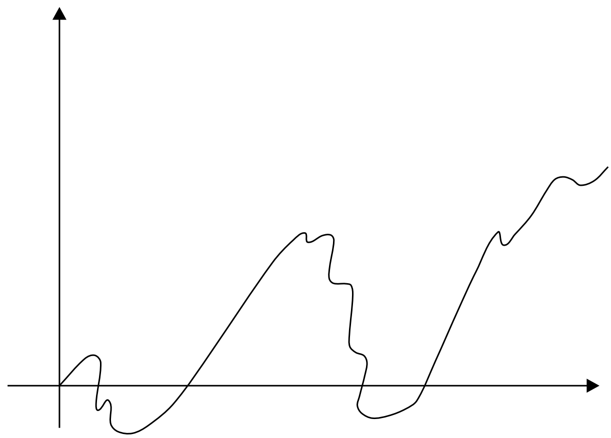


Fig. 5.1.1

Este fenômeno foi observado por um botânico inglês, Robert Brown em 1827, e estudado matematicamente por A. Einstein em 1905, e posteriormente por matemáticos como N. Wiener e P. Levy. Este experímento aleatório é chamado de Movimento Browniano ou Processo de Wiener.

Seja $\Omega = \{\omega : \omega \colon [0, +\infty) \to \mathbb{R}^1 \text{ a contínua}\}$.

Seja $X_t(\omega) = \omega(t)$, $t \geq 0$, e seja $\mathcal{A}$ a menor σ-álgebra, tal que todas as funções X_t resultem mensuráveis.

Enunciamos sem demonstração o seguinte teorema.

Teorema 5.1.1 (Existência do Movimento Browniano). *Existe uma probabilidade P sobre $\mathcal{A}$ tal que o processo estocástico $\{X_t\}_{t\geq 0}$ tem (com relação a P) as seguintes propriedades:*

1) $\forall\, t \geq 0,\quad \mathcal{L}(X_t) = N(0,t)$

2) $\forall\, n,\quad 0 \leq t_0 < t_1 < \cdots < t_n$ *as variáveis*

$$X_{t_0}, X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

são independentes.

Vamos agora enunciar e esboçar as demonstrações de várias importantes propriedades deste processo.

Começamos pelo chamado Princípio da reflexão, ferramenta utilíssima na demonstração dos outros resultados.

Princípio da Reflexão. Definimos para cada $t \geq 0$,

$$M_t(\omega) \max_{0\leq u\leq 1} X_u(\omega).$$

A seguinte igualdade é conhecida como Princípio da Reflexão.

Teorema 5.1.2. $P[\omega : M_t(\omega) \geq \lambda] = 2P[\omega : X_t(\omega) \geq \lambda]$ *onde* λ *é real não negativo.*

Vamos fazer uma justificação intuitiva desta identidade.

Seja $A = [M_t \geq \lambda]$. Então temos

(I) $$A = A \cap [X_t \geq \lambda] + A \cap [X_1 < \lambda].$$

A cada trajetória de $A \cap [X_t \geq \lambda]$ corresponde por "reflexão"a partir do primeiro instante de tempo u tal que $X_u(\omega) = \lambda$ uma trajetória que

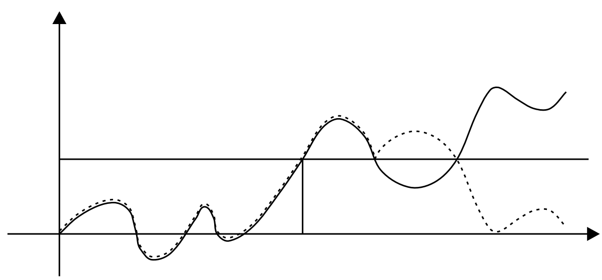


Fig. 5.1.2

pertence a $A \cap [X_t < \lambda]$ e reciprocamente. É portanto mais ou menos natural esperar que a probabilidade dos dois eventos que aparecem no membro direito da identidade (I) sejam iguais. Temos portanto

$$P(A) = 2P\{A \cap [X_t \geq \lambda]\} = 2P[X_t \geq 1];$$

Teorema 5.1.3. *Quase todas as trajetórias do Movimento Browniano não são deriváveis em ponto nenhum.*

Demonstração: Vamos nos limitar ao caso da origem. Suponhamos $h\colon [0, +\infty) \to \mathbb{R}'$ é tal que

a) $h(0) = 0$.

b) Existe $\{t_n\}_{n=1,2,\ldots}$ e $\{t'_n\}_{n=1,2,\ldots}$ duas seqüências tais que $t_n > 0$, $t'_n > 0$, $t_n \to 0$, $t'_n \to 0$, $h(t_n) \geq (t_n)^\delta$ e $h(y'_n) \leq -(t'_n)^\delta$ com $0 < \delta < 1$.

Então h não é derivável na origem porque

$$\frac{h(t_n)-h(0)}{t_n} = \frac{h(t_n)}{t_n} \geq \frac{(t_n)^\delta}{t_n} = \frac{1}{t_n^{(1-\delta)}} \to +\infty$$
$$\frac{h(t'_n)-h(0)}{t'_n} = \frac{h(t'_n)}{t'_n} \leq -\frac{(t'_n)^\delta}{t'_n} = -\frac{1}{t_n'^{(1-\delta)}} \to -\infty.$$

Seja $B_n = \{\omega : \exists\, u,\, 0 < u < \frac{1}{n}, X_u(\omega) \geq u^\delta\}$ com $/2 < \delta < 1$. Temos $\forall\, t$, $0 < t < \frac{1}{n}$, $B_n \supseteq |M_t \geq t^\delta]$ e portanto

$$P(B_n) \overset{\forall t}{\geq} P|M_t \geq t^\delta] \overset{\forall t}{=} 2P|X_t \geq t^\delta] \overset{\forall t}{=} 2P\Big[\frac{X_t}{\sqrt{t}} \geq t^{\delta-1/2}\Big] \overset{\forall t}{=} 2\int_{t^{\delta-1/2}}^{+\infty} f_1(x).$$

Como quando $t \to 0$ o último membro desta cadeia de igualdades converge a 1, temos que $\forall\, n$ $P(B_n) = 1$. Daqui resulta que $P\big(\bigcap_{n=1}^{\infty} B_n\big) = 1$. Portanto para que $\omega\, \exists$ uma seqüência t_n, $t_n > 0$, $t_n \to 0$ tal que $X_{t_n}(\omega) \geq t_n^\delta$. Com o mesmo procedimento podemos construir uma seqüência t'_n e portanto quase toda trajetória ω não é derivável na origem.

Teorema 5.1.4. *O conjunto de zeros de uma trajetória é um conjunto não enumerável, fechado, sem pontos isolados, de dimensão topológica* 0*, e de medida de Lebesgue* 0*.*

Esboço de parte da demonstração:

Que o conjunto de zeros é fechado decorre do fato das trajetórias serem funções contínuas. Vamos provar agora que ou não é um ponto isolado ou que é essencialmente suficiente.

Seja $t > 0$,

$$P[M_t > 0] = P\Big\{\bigcup_{n=1}^{\infty}\Big[M_t \geq \frac{1}{n}\Big]\Big\} = \lim_n P\Big[M_t \geq \frac{1}{n}\Big] =$$
$$= \lim_n 2P\Big[X_t \geq \frac{1}{n}\Big] = 2P[X_t > 0] = 1.$$

É dizer que qualquer trajetória toma um valor positivo em qualquer intervalo contendo a origem. Da mesma forma podemos provar que toma valores negativos em qualquer intervalo, e portanto, como as trajetórias são contínuas, anula-se em qualuer intervalo. É dizer-se

$$B_n = \{\omega : \exists t,\, 0 < t < \frac{1}{n}\,,\, X_t(\omega) = 0\}$$

$P(B_n) = 1$ e portanto $P\big(\bigcap_{n=1}^{\infty} B_n\big) = 1$ o que prova o resultado.

Seja agora $h\colon (0,1[\to \mathbb{R}$ uma função contínua, não-negativa. Consideremos o evento

$$\{\omega > \exists\, \varepsilon > 0 \text{ tal que } X_t(\omega) < h(t), \forall\, t\, 0 < t \leq \varepsilon\}.$$

Por aplicação da Lei 0-1 de Blumenthal este evento tem probabilidade 0 ou 1.

Definição 5.1.1. h é dita de classe *superior* (resp. *inferior*) se a probabilidade desse evento é 1 (resp. 0).

Teorema 5.1.5 (Lei do logaritmo iterado). *Se* $h(t) = (1+\varepsilon)\big[2t \log \log\big(\frac{1}{t}\big)\big]^{1/2}$ *então se* $\varepsilon > 0$, h *pertence à classe superior; se* $\varepsilon < 0$, h *pertence à classe inferior.*

Teorema 5.1.6 (Kolmogorov). *h crescente e* $\dfrac{h(t)}{\sqrt{t}}$ *é decrescente em algum intervalo da forma* $(0,\delta)$, *então h pertence à classe superior ou inferior de acordo com que* $\displaystyle\int_{(0,\delta)} t^{-3/2}\, e^{-h^2/2t}\, h\, dt$ *seja finito ou não.*

Seja agora h uma função contínua em $[1,+\infty)$, positiva, e tal que $\dfrac{h(t)}{t}$ é decrescente e $\dfrac{h(t)}{\sqrt{t}}$ é crescente em algum intervalo da forma $[a,+\infty)$. Seja

$$A = \{\omega : \exists\, a > 0 \text{ com } X_t(\omega) < h(t)\ \forall\, t \geq a\}.$$

Então

$$P(A) = \begin{cases} 1 & \text{se} \quad \displaystyle\int_1^{\infty} t^{-3/2}\, e^{-h^2/2t}\, h(t) dt < \infty \\ \\ 0 & \text{se} \quad \displaystyle\int_1^{\infty} t^{-3/2}\, e^{-h^2/2t}\, h(t) dt = \infty \end{cases}$$

Corolário 5.1.1. *Se* $h(t) = (1+\varepsilon)\sqrt{2t \log\log t}$ *então*

$$P(A) = \begin{cases} 1 & se \quad \varepsilon > 0 \\ 0 & se \quad \varepsilon < 0 \end{cases}$$

5.2 Processo gaussianos

Seja $\{X_t\}_{t\in T}$ um processo estocástico tal que $\forall\, t \in T$, $EX_t^2 < \infty$. Definimos então a função $\mu(t) = EX_t$, $t \in T$, e a função

$$K(s,t) = \text{cov}(X_s, X_t) = E(X_s - EX_s)(X_t - EX_t)$$

onde s e t pertencem a T.

μ é chamada *função valor médio* e K *função de covariança.*

Seja agora $\{X_t\}_{t\in T}$ um processo com incrementos independentes cov $X_0 = 0$.
Temos para $0 \leq s \leq t$

$$\begin{aligned} K(s,t) &= \text{cov}(X_s, X_t) = \text{cov}(X_s, X_s + X_t - X_s) = \\ &= \text{cov}(X_s, X_s) + \text{cov}(X_s, X_t - X_s) = \\ &= \text{var}(X_s). \end{aligned}$$

Em geral temos para um processo deste tipo

$$K(s,t) = \text{var}(X_{s\wedge t}).$$

Se o processo além de ter incrementos independentes, estes fossem estacionários teríamos para $0 \leq s \leq t$

$$K(s,t) = \text{var}(X_s) = s\,\text{var}(X_1)$$

e em geral

$$K(s,t) = (s \wedge t)\,\text{var}(X_1).$$

Sabemos também que para um processo deste tipo

$$\mu(t) = EX_t = tEX_1\,.$$

Em particular, se $\{X_t\}_{t\geq -}$ é o processo de Poisson temos que $K(s,t) = (s \wedge t)\lambda$.

Se $\{X_t\}_{t\geq 0}$ é o processo de Wiener $K(s,t) = (s \wedge t)$.

Um processo estocástico é com *covariança estacionária*, se $K(s,t)$ é uma função da diferença $t - s$. De outra, existe uma função $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ (necessariamente par) tal que

$$K(s,t) = f(s-t).$$

Os dois exemplos anteriores (i.e. Processo de Poisson e Processo de Wiener) são exemplos de processo que não tem covariança estacionária.

Todo processo estacionário tem covariança estacionária: se $m = EX_0$

$$\begin{aligned} K(s+u,t+u) &= E(X_{s+u}-m)(X_{t+u}-m) = \\ &= \int_{\mathbb{R}^2}(x-m)(y-m)\mathcal{L}(X_{s+u},X_{t+u})(dx,dy) = \\ &= \int_{\mathbb{R}^2}(x-m)(y-m)\mathcal{L}(X_s,X_t)(dx,dy) = \\ &= e(X_s-m)(X_t-m) = K(s,t). \end{aligned}$$

Um processo estocástico $\{X_t\}_{t\in T}$ é chamado *gaussiano* ou *normal* se $\forall\, n$, $\forall\, t_1,t_2,\ldots,t_n \in T$ a distribuição conjunta de $X_{t_1},\ldots,X_{t_n}$ e normal em $\mathbb{R}^n$.

Consideremos o caso de $T = \mathbb{R}^1$ ou $T = [0,+\infty)$.

Seja $t_1 < t_2 < \cdots < t_n$. Se m indica o vetor de médias das variáveis $X_{t_1},\ldots,X_{t_n}$ e Σ a matriz de covariança temos que a função característica de $\mathcal{L}(X_{t_1},\ldots,X_{t_n})$ é igual a

$$\text{(I)} \qquad \varphi(u) = e^{iu'm-1/2(i'\Sigma u)}, \quad u = (u_1,\ldots,u_n)'.$$

Suponhamos agora que o processo gaussiano $\{X_t\}_{t\in T}$ tenha μ (função de médias) constante e covariança estacionária. Se consideramos as variáveis $X_{t_1+h},\ldots,X_{t_n+h}$ resulta que a função característica da distribuição conjunta destas variáveis é igual a função característica correspondente às variáveis $X_{t_1},\ldots,X_{t_n}$. Basta para isso observar a expressão (I). Isto prova então que todo processo gaussiano com função de médias μ constante e covariança estacionária é um processo estacionário.

Um exemplo de processo gaussiano é o movimento Browniano. Seja

$\{X_t\}_{t\geq 0}$ movimento Browniano e seja $0 \leq t_1 < t_2 < \cdots < t_n$. Temos

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & & \dots & 0 \\ 1 & 1 & & \dots & 0 \\ 1 & 1 & 1 & \dots & \\ \vdots & & & & \\ 1 & & & \dots & 1 \end{pmatrix} \begin{pmatrix} X_{t_1} \\ X_{t_2} - X_{t_1} \\ \vdots \\ X_{t_n} - X_{t_{n-1}} \end{pmatrix}.$$

Como $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ são v.a. independentes e normais e $X_{t_1}, \dots, X_{t_n}$ é obtido como uma transformação linear dessas variáveis resulta que $\mathcal{L}(X_{t_1}, \dots, X_{t_n})$ é normal em $\mathbb{R}^n$.

Vamos construir agora outro importante exemplo de processo gaussiano.

Seja $\{X_t\}_{t\geq 0}$ movimento browniano e $g\colon \mathbb{R}^1 \to \mathbb{R}^1$ uma função positiva, estritamente crescente e tal que $g(0) = 1$.

Seja $Y_t = \dfrac{1}{\sqrt{g(t)}} X_{ag(t)}$, $a > 0$, $t \in \mathbb{R}^1$. Este novo processo estocástico com espaço paramétrico $T = \mathbb{R}^1$ tem as seguintes propriedades:

$$EY_t = 0, \quad \operatorname{var} Y_t = \frac{1}{g(t)} \operatorname{var} X_{ag(t)} = a.$$

Se $t_1 < t_2 < \cdots < t_n$

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{g(t_1)}} & & \bigcirc \\ & \ddots & \\ \bigcirc & & \frac{1}{\sqrt{g(t_n)}} \end{pmatrix} \begin{pmatrix} X_{ag(t_1)} \\ \vdots \\ X_{ag(t_n)} \end{pmatrix}$$

Como $(X_{ag(t_1)}, \dots, X_{ag(t_n)})$ é normal resulta que $(T_{t_1}, \dots, Y_{t_n})$ é também normal. Ou seja $\{Y_t\}_{t\in\mathbb{R}^1}$ é um processo gaussiano. Vamos agora escolher g da forma tal que $\{Y_t\}_{t\in\mathbb{R}^1}$ resulte estacionário. como é gaussiano com média 0 é suficiente que seja com covariança estacionária.

Seja $h > 0$ e consideremos

$$E(Y_{t+h}, Y_t) = \frac{E(X_{ag(t)} X_{ag(t+h)})}{\sqrt{g(t)g(t+h)}} =$$

$$= \frac{\operatorname{cov}(X_{ag(t)}, X_{ag(t+h)})}{\sqrt{g(t)g(t+h)}} = \frac{\operatorname{var}(X_{ag(t)})}{\sqrt{g(t)g(t+h)}} =$$

$$= \frac{ag(t)}{\sqrt{g(t)g(t+h)}} = a\sqrt{\frac{g(t)}{g(t+h)}}$$

queremos que esta última expressão seja independente de t. Ou seja

$$\frac{g(t)}{g(t+h)} = \text{ cte. } = \frac{g(0)}{g(h)} = \frac{1}{g(h)}.$$

Portanto devemos ter $\forall\, t \in \mathbb{R}^1$, $\forall\, h > 0$, como $\forall\, t \in (0,1]$, $g(t) < g(1)$ estamos em condições de utilizar os resultados do Apêndice 4. Resulta que

$$g(t) = [g(1)]^t.$$

Como $g(1) > g(0) = 1$, podemos escrever $g(1)$ como $g(1) = e^{2b}$ com $b > 0$.
Temos então

$$g(t) = e^{2bt}.$$

Representando esta expressão na fórmula da covariança temos que

$$\operatorname{cov}(Y_t, Y_{t+h}) = a \left[\frac{e^{2bt}}{e^{2b(t+h)}} \right]^{1/2} = a\, e^{-bh}.$$

Resumindo: o processo

$$Y_t = e^{-bt} X_{a\, e^{2bt}}\,, \quad t \in \mathbb{R}^1, \quad a > 0, \quad b > 0$$

é um processo gaussiano com $EY_t = 0$, $\operatorname{var}(Y_t) = 0$, estacionário, e com função de covariança

$$K(s,t) = a\, e^{-b|s-t|}.$$

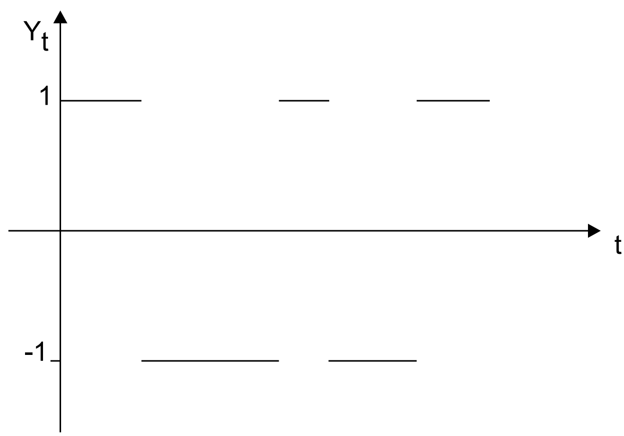
Como é uma função do movimento Brownianotem trajetórias contínuas, não diferenciáveis, sem raízes isoladas. O processo Y_t é chamado *processo de Ornstein-Uhlenbeck*.

Exercício 5.1. Seja $\{X_t\}_{t\geq 0}$ o processo de Poisson com parâmetro $\lambda > 0$. Seja Z uma v.a. independente deste processo e tal que

$$Z = \begin{cases} +1 & 1/2 \\ -1 & 1/2 \end{cases}$$

Seja $Y_t = Z(-1)^{X_t}$.

O processo Y_t é chamado freqüentemente *sinal de telégrafo aleatório*. Uma trajetória típica está representada na seguinte figura



Provar:

a) O processo Y_t é estacionário;

b) Provar que para este processo

$$K(s,t) = e^{-2\lambda|t-s|}.$$

Note que se tomamos um processo de Ornstein-Uhlenbeck restringido a valores de $t \geq 0$, $a = 1$ e $b = 2\lambda$ obtemos um processo com a mesma função de covariança. Assim dois processos tão diferentes como estes dois, podem ter a mesma função de covariança.

APÊNDICE 1

Propriedade Forte de Markov

Seja f uma função definida sobre o espaço de trajetórias, mensurável e não negativa ou limitada.

Seja θ o operador translação por α ou seja, $\theta_n(\omega)(k) = \omega(k+n)$. A propriedade de Markov pode ser escrita na forma:

$$E^i(f : \theta_n \mid X_0, X_1, \ldots, X_n) = E^{X_n}(f)$$

(a igualdade é válida em q.t.p. $[P^i]$, $\forall\, i \in E$).

Seja agora τ um tempo de parada e

$$\mathcal{J}_\tau = \{A : A \cap [\tau = n] \in \sigma[X_0, X_1, \ldots, X_n], \forall\, n\}.$$

Seja $\theta_\tau(\omega)(k) = \omega(\tau(\omega) + k)$.

A propriedade forte de Markov substitui o índice n (ou tempo n) pelo "tempo aleatório" τ da forma natural.

A propriedade forte de Markov é enunciada de forma precisa da seguinte forma

$$E^i(f \circ \theta_\tau \mid \mathcal{J}_\tau) = E^{X_\tau} f$$

sobre $[\tau \leq \infty]$ (a igualdade é válida q.t.p. $[P^i]$).
X_τ é a função $\omega \to X_{\tau(\omega)}(\omega)$.

APÊNDICE 2

Seja n_j a v.a. que indica o número de vezes que o processo visita o estado $j \in E$ e τ_j o tempo de parada que indica a primeira vez que o estado $j \in E$ é visitado.

Seja $A_j = \{\omega : X_n(\omega) = j, \text{ algum } n \geq 0\}$ e $\overline{A}_j = \{\omega : X_n(\omega) = j, \text{ algum } n \geq 1\}$.

Seja $\bar{\tau}_j$ o tempo de parada que indica a primeira visita ao estado j depois do tempo 0. Definimos $N_{ij} = E^i(n_j)$ e $\overline{H}_{ij} = P^i(\overline{A}_j)$. A matriz N é chamada *matriz de potencial* da cadeia.

Proposição 1. *A probabilidade de retornar ao estado i ao menos k vezes é igual a $(\overline{H}_{ii})^k$.*

Demonstração: Para $k = 0$ trivial. Procederemos por indução.

Seja $A(A')$ o evento que acontece se o processo retorna ao estado i ao menos $k(k-1)$ vezes. Seja θ o operador translação.

$$
\begin{aligned}
P^i[A] &= P^i\left[\theta_{\bar{\tau}_i \in A'}\,, \bar{\tau}_i < \infty\right] = \\
&= E^i\, E^i I_{[\bar{\tau}_i \in A']}\, I_{[\bar{\tau}_i < \infty]}\, [\delta_{\bar{\tau}_i}] =
\end{aligned}
$$

(prop. forte de Markov)

$$
\begin{aligned}
&= E^i\, I_{[\bar{\tau}_i < \infty]} E^{X_{\bar{\tau}_i}}\, (I_{A'}) = \\
&= P^i[A'] P^i[\bar{\tau}_i < \infty] \overset{\text{(indução)}}{=} \overline{H}_{ii}^{k-1}\, \overline{H}_{ii} = \\
&= \overline{H}_{ii}^{k}\,.
\end{aligned}
$$

Proposição 2. *O estado i é transitório se e somente se $N_{ii} < \infty$. Nesse caso $N_{ii} = \dfrac{1}{1 - \overline{H}_{ii}}$.*

Demonstração:

$$N_{ii} = \sum_{k=0}^{\infty} k\, P^i[n_i = k] = \sum_{k=1}^{\infty} k\, P^i[n_i = k] =$$
$$= \sum_{k=1}^{\infty} \sum_{m=k}^{\infty} P^i[n_i = m] = \sum_{k=1}^{\infty} P^i[n_i \geq k] =$$
$$= \sum_{k=1}^{\infty} \overline{H}_{ii}^{k-1}\,.$$

Esta última série é convergente se e somente se $\overline{H}_{ii} < 1$. Nesse caso sua soma é igual $\dfrac{1}{1-\overline{H}_{ii}}$ o que prova a Proposição 2. Note-se que $N_{ii} < \infty$ se e somente se $\sum\limits_{n=0}^{\infty} P^{(n)}(i,i) < \infty$ o que é equivalente a $\sum\limits_{n=1}^{\infty} P^{(n)}(i,k) < \infty$.

Algumas outras relações úteis

Seja $H_{ij} = P^i(A_j)$. Temos a seguinte:

Proposição 3. $\overline{H} = PH$.

Demonstração:

$$\overline{H}_{ij} = P^i(\overline{A}_j) = P^i(\theta_1 \in A_j) = \sum_{k \in E} P^i[X_1 = k] P^k(A_j) =$$
$$= \sum_{k \in E} P(i,k) H(k,j) = (PH)(i,j).$$

Proposição 4. $N_{ij} = H_{ij}\, N_{jj}$.

Demonstração: $N_{ij} = E^i(n_j)$. Sobre $[\tau_j < \infty]\, n_j = n_j \circ \theta_{\tau_j}$.
Portanto

$$E^i(n_j) = E^i((n_j \circ \theta_{\tau_j}) I_{[\tau_j < \infty]}) =$$
$$= E^i\, I_{[\tau_j<\infty]} E^{X_{\tau_j}}\, n_j = E^i\, I_{[\tau_j<\infty]} E^j(n_j)$$
$$= N_{ij}\, E^i[\tau_j < \infty] = N_{jj}\, P^i(A_j) = N_{jj}\, H_{ij}\,.$$

APÊNDICE 3

Martingalas

Em $(\Omega, \mathcal{A}, P)$ seja $\{X_n\}_{n=0,1,2,\ldots}$ uma seqüência de v.a. e $\{\mathcal{J}_n\}_{n=0,1,\ldots}$ uma seqüência crescente de sub-σ-álgebras de $\mathcal{A}$. Dizemos que $\{X_n, \mathcal{J}_n, n \geq 0\}$ é uma *martingala* se

1) $\forall\, n, \quad X_n$ é integrável;

2) X_n é $\mathcal{J}_n$-mensurável;

3) $E(X_{n+1} \mid \mathcal{J}_n) = X_n \quad$ q.c.

Se em 3) em lugar de $=$, tivéssemos $\geq$, temos uma *submartingala*; se fosse $\leq$ uma *supermatingala.*

Um exemplo básico de martingala consiste em considerar uma seqüência de v.a.i.i.d. $\{\xi_n\}_{n=0,1,\ldots}$ com $E(\xi_0) = 0$.

Então se $X_n = \sum\limits_{i=0}^{n} \xi_i$ e $\mathcal{J}_n = \sigma(\xi_0, \ldots, \xi_n)$ resulta que $\{X_n\,, \mathcal{J}_n\,, n \geq 0\}$ é uma martingala.

Outro exemplo importante é o seguinte. Seja $\{x_n\}_{n\geq 0}$ uma cadeia de Markov com espaço de estados E e matriz de probabilidades de transição P. Seja h uma função real definida em E e tal que $h(i) = \sum\limits_{j\in E} P(i,j)h(j)$. Então se $\forall\, n$, $h(X_n)$ é integrável resulta que $\{h(X_n), \sigma[X_0, \ldots, X_n]; n \geq 0\}$ é uma martingala.

Vamos enunciar agora alguns dos resultados básicos da Teoria das Martingalas. As demonstrações destes resultados podem ser encontradas em [1].

Teorema 1. *Se $\{X_n\,, \mathcal{J}_n\,; n \geq 0\}$ é uma submartingala tal que $E[X_n] \leq K \leq \infty \quad \forall\, n$, então $\lim\limits_{m\to\infty} X_n(\omega)$ existe e é finito para quase todo ω.*

Corolário 1. *Qualquer supermartingala não negativa, converge a uma função finita em quase todo ponto. Em particular, qualquer martingala não negativa, converge quase certamente.*

Definição 1. Dada uma seqüência não decrescente de σ-álgebras $\{\mathcal{J}_n\}_{n=0,1,\dots}$ um *tempo aleatório* τ é uma variável aleatória que toma os valores $\{0, 1, 2, \dots, +\infty\}$ e tal que $\forall\, n$ $[\tau \le n] \in \mathcal{J}_n$.

(Se em lugar de ter uma seqüência de σ-álgebra tivermos uma família crescente de σ-álgebra $\{\mathcal{J}_t\}_{t\ge 0}$ a condição seria que $[\tau \le t] \in \mathcal{J}_t$, $\forall\, t \ge 0$).

Se τ é um tempo aleatório finito (*tempo de parada*) e $\{X_n\}_{n\neq 0}$ é uma seqüência de variáveis aleatórias. Definimos

$$X_\tau(\omega) = X_{\tau(\omega)}(\omega).$$

Resulta que X_τ é uma v.a.

Teorema 2. *Seja $\{X_n, \mathcal{J}_n;\, n \ge 0\}$ uma martingala e τ um tempo de parada. Se*

1) $E|X_\tau| < \infty$;

2) $\lim_n \int_{[\tau > \infty]} X_n = 0$

então

$$E(X_\tau) = E(X_0).$$

Um resultado análogo é válido para submartingalas e supermartingalas representando a igualdade pela desigualdade correspondente.

Corolário 2. $\{X_n, \mathcal{J}_n;\, n \ge 0\}$ *uma martingala e τ um tempo de parada. Se $|X_n| \le K < \infty$ $\forall\, n$ então $EX_\tau = EX_0$.*

Algumas aplicações e exemplos.

Seja $\{\xi_n\}_{n=1,2,\dots}$ uma seqüência de v.a.i.i.d.

$$\xi_n = \begin{cases} +1 & 1/2 \\ -1 & 1/2 \end{cases}$$

Seja $X_n = \sum_{o=1}^{n} \xi_i$, $X_0 = 0$. $\{X_n, \sigma[\xi_1, \dots, \xi_n\}.n \ge 0\}$ é uma martingala. Seja M, N inteiros positivos e τ o tempo de parada que indica a primeira vez que o processo alcança M ou $-N$. Seja $p = P[X_\tau = M]$ e $q = 1 = p = P[X_\tau = -N]$. Temos que

$$0 = EX_0 = EX_\tau \, Mp - Nq - N(1-p) = (M+N)p - N.$$

Portanto $p = \dfrac{N}{M+N}$ e $q = \dfrac{M}{M+N}\cdot$

Este resultado pode ser aplicado ao Exemplo 2.1.2 ($p = q = 1/2$) dando como probabilidade de ruina do jogador com capital inicial x, $\dfrac{a-x}{a} = \dfrac{1-x}{a}\cdot$ Se $p \neq q$ então $\left(\dfrac{q}{p}\right)^{X_n}$ é uma martingala.

Exercício. Seja $\{X_n\}_{n\geq 0}$ a cadeia de Markov correspondente ao passeio casual simétrico em duas dimensões. Provar que $z_n = ||X_n||^2 - n$ é uma martingala. (Portanto $\forall\, n \geq 0 \;\; E||X_n||^2 = n$).

APÊNDICE 4

Lema 1. *$f\colon (0,+\infty) \to \mathbb{R}$, $|f(x)| \leq K$ para $0 \leq x \leq 1$, e $f(x+y) = f(x) + f(y)$ $\forall\, x > 0$ e $\forall\, y > 0$. Então $f(x) = f(1)x$ $\forall\, x > 0$.*

Demonstração: Seja $F(x) = f(x) - f(1)x$. Então F tem as seguintes propriedades:

$$|F(x)| \leq 2K \quad \forall\, 0 \leq x \leq 1; \quad F(1) = 0$$

e

$$F(x+y) = F(x) + F(y) \quad \forall\, x > 0 \text{ e } \forall\, y > 0.$$

Vamos provar que $F \equiv 0$. Temos se $x_1 > 0, x_2 > 0, \dots, x_n > 0$ que

$$\begin{aligned} F\Big(\sum_{i=1}^{n} x_i\Big) &= F\Big(\sum_{i=1}^{n-1} x_i\Big) + F(x_n) = \\ &= F\Big(\sum_{i=1}^{n-2} x_i\Big) + F(x_{n-1}) + F(x_n) = \cdots = \\ &= \sum_{i=1}^{n} F(x_i). \end{aligned}$$

Colocando $x_2 = x_3 = \cdots = x_n = 1$ temos que $\forall\, x_1 > 0$, $F(x_1 + n - 1) = F(x_1) + (n-1)F(1) = F(x_1)$. Portanto F é periódica. Resulta então que $|F| \leq 2K$. Se $F \not\equiv 0$, $\exists\, x_0 > 0$ tal que $F(x_0) \neq 0$. Então $|F(nx_0)| = n|F(x_0)|$. Como $F(x_0) \neq 0$ quando n cresce $n|F(x_0)|$ seria eventualmente maior que $2K$ o que é uma contradição.

Lema 2. *Seja $g\colon (0,+\infty) \to \mathbb{R}$ tal que $|g(x)| \leq K$, $\forall\, 0 < x \leq 1$ e*

$$g(x+y) = g(x)g(y) \qquad \forall\, x \geq 0, \ \forall\, y \geq 0.$$

Então $g(1) \geq 0$, $g(x) = [g(1)]^x$. Ou seja g é identicamente nula ou sempre positiva e da forma

$$g(x) = [g(1)]^2 = e^{\log(g(1))x}.$$

Demonstração: Temos que $g(2x) = g(x)g(x) = (g(x))^2$, $\forall\, x > 0$. Portanto $g \geq 0$. Se $g \equiv 0$ o lema está provado. Caso contrário, existe x_0

tal que $g(x_0) > 0$. Vamos provar que então $g > 0$. Temos primeiro que $\forall\, n \geq 1,\ g(nx_0) = [g(x_0)]^n > 0$; ou seja sobre a seqüência nx_0 , $g > 0$. Vamos provar agora que se $g(u) > 0$ então $g(x) > 0,\quad \forall\, 0 < x \leq u$. Isto decorre da identidade

$$0 < g(u) = g(x + (u - x)) = g(x)g(u - x).$$

Portanto $g > 0$. Seja então $f(x) = \log g(x)$. Se $0 < x < 1$ temos $0 < g(1) = g(x)g(1 - x) \leq g(x)K$.
Portanto $\dfrac{g(1)}{K} \leq g(x) \leq K,\quad \forall\, x,\ 0 < x < 1$. Desta desigualdade resulta que $f(x) = \log g(x)$ e limitada no intervalo $(0, 1]$. Portanto pelo Lema 1 $f(x) = f(1)x$ ou seja, $\log g(x) = [\log(g(1))]x$ e portanto $g(x) = e^{[\log(g(1))]x}$.

Notas: a) O fato de g ser limitada no intervalo $(0, 1]$ pode ser trocado pela sua limitação em algum intervalo da forma $(0, \varepsilon]$, $\varepsilon > 0$.

b) A hipótese de limitação pode ser trocada pela de *continuidade* em algum intervalo da forma $[0, \varepsilon]$, $\varepsilon > 0$.

APÊNDICE 5

Seja $\Omega, \mathcal{A}, P)$ um espaço de probabilidades e $(E, \mathcal{C})$ um espaço mensurável. Seja T um conjunto arbitrário.

Definição 1. Uma família $\{X_t : t \in T\}$ de aplicações $\mathcal{A}-\xi$ mensuráveis é chamada um *processo estocástico* com *espaço de estados* E e *conjunto de índices* (de parâmetros) T.

Seja $\mathcal{U}$ a família de todos os subconjuntos finitos de T ($\mathcal{U} = \mathbb{P}_f(T)$). Indicaremos com α, β, γ, δ, etc. os elementos de $\mathcal{U}$.

Seja $\alpha \in \mathcal{U}$ (resp. α e β dois elementos de $\mathcal{U}$ tais que $\alpha \subseteq \beta$). Denotaremos com π_α (resp. $\pi_{\alpha\beta}$) a projeção canônica de E^T sobre E^α (resp. de E^β sobre E^α). As aplicações π_α e $\pi_{\alpha\beta}$ são mensuráveis com respeito as correspondentes σ-álgebras produto.
Temos também que

$$\pi_{\alpha\beta}\, \pi_\beta = \pi_\alpha \, (\alpha, \beta \in \mathcal{U},\ \alpha \subseteq \beta)$$
$$\pi_{\alpha\beta}\, \pi_{\beta\gamma} = \pi_{\alpha\gamma} \, (\alpha, \beta \text{ e } \gamma \in \mathcal{U},\ \alpha \subseteq \beta \subseteq \gamma)$$

Suponhamos que temos uma probabilidade P_α sobre cada espaço mensurável (E^α, ξ^α), $\alpha \in \mathcal{U}$. Dizemos que a família $\{P_\alpha\}_{\alpha \in \mathcal{U}}$ é *consistente* (ou que constitui um *sistema projetivo de probabilidades*) se $\forall\, \alpha$, $\beta \in \mathcal{U}$, $\alpha \subseteq \beta$

$$\pi_{\alpha\beta}(P_\beta) = P_\beta\, \pi_{\alpha\beta}^{-1} = P_\alpha\,.$$

Uma pergunta natural é a seguinte: dada uma família consistente $\{P_\alpha\}_{\alpha \in \mathcal{U}}$ será que existe uma probabilidade P sobre (E^T, ξ^T) tal que $\forall\, \alpha \in \mathcal{U}$

$$P_\alpha = P\, \pi_\alpha^{-1} = \pi_\alpha(P).$$

Para dar uma resposta a esta pergunta introduzimos a seguinte

Definição 2. Uma classe de conjuntos $\mathcal{C}$ é chamada *compacta* se para toda seqüência de conjuntos com $C_n \in \mathcal{C}$ e tais que $\forall\, n$, $\bigcap_{i=1}^{n} C_i \neq \emptyset$ temos que $\bigcap_{i=1}^{\infty} C_i \neq \emptyset$.

Definição 3. Seja P uma probabilidade sobre o espaço mensurável $(\Omega, \mathcal{A})$. Seja $\mathcal{C}$ uma classe compacta contida em $\mathcal{A}$. Dizemos que P é *regular* com respeito se $\forall\, A \in \mathcal{A}$

$$P(A) = \sup\{P(C) : C \in \mathcal{C}, C \subseteq A\}.$$

Temos agora o seguinte

Teorema (de Consistência de Kolmogorov). *Seja* $\{P_\alpha\}_{\alpha \in \mathcal{U}}$ *uma família consistente de probabilidades. Suponhamos que* $\forall\, t \in T$*, existe* C_t *uma classe compacta contida em* ξ *tal que* $P_{\{t\}}$ *é regular com relação a* $\mathcal{C}_t$ *. Então existe uma probabilidade* P *sobre* (E^T, ξ^T) *tal que* $\forall\, \alpha \in \mathcal{U}$*,* $P\,\pi_\alpha^{-1} = P_\alpha \cdot P$ *é única.*

Uma demonstração deste resultado pode ser encontrada em [5].

REFERÊNCIAS

[1] Kemeny, J.G., Snell, J.L., Knapp, A.W. - (1966) - *Denumerable Markov Chains.* Van Nostrand.

[2] Fernandez, Pedro J. - (1973) - *Introdução à Teoria das Probabilidades.* Ao Livro Técnico e Científico Editora S.A., Rio de Janeiro, RJ.

[3] Feller, W. - (1950) - *An Introduction to Probability Theory and its Applications.* Vol. I e II, John Wiley & Sons, Inc.

[4] Thomasian, A.J. - (1969) - *The Structure of Probability Theory with Applications.* Mc Graw-Hill.

[5] Neveu, J. - (1965) - *Mathematical Foundations of the Calculus of Probability.* Holden-Day, Inc.

[6] Hoel, P.G., Port, S.C., Stone, G.J. - (1972) - *Introduction to Stochastic Processes.* Boston, Mifflin.

[7] Doob, J.L. - (1953) - *Stochastic Processes.* New York, J. Wiley.

[8] Freedman, D. - (1971) - *Markov Chains.* Holden-Day, Inc.