Introdução aos Processos Estocásticos

Luiz Renato Fontes

Reversão no tempo

Propriedade de Markov: simétrica no tempo;
Convergência ao equilíbrio: assimétrica
Distribuição inicial de equilíbrio: restaura simetria temporal

Teorema 1. Seja $\mathbf{P}$ irredutível e com distribuição invariante π, e $\mathbf{X} \sim \mathrm{CM}(\pi, \mathbf{P})$. Dado $N \geq 0$, seja $Y_n = X_{N-n}$, $n = 0, \dots, N$.

Então $(Y_n)_{0 \leq n \leq N} \sim \mathrm{CM}(\pi, \hat{\mathbf{P}})$, onde $\hat{\mathbf{P}} = (\hat{P}_{xy})_{x,y \in \mathcal{S}}$ é dada por

$$\hat{P}_{xy} = \tfrac{\pi_y}{\pi_x} P_{yx} \qquad (\Leftrightarrow \pi_x \hat{P}_{xy} = \pi_y P_{yx}).$$

Além disso, $\hat{\mathbf{P}}$ é irredutível e tem distribuição invariante π.

Dem. Vamos verificar que

1) $\hat{\mathbf{P}}$ é estocástica: dado $x \in \mathcal{S}$,

$$\textstyle\sum_{y \in \mathcal{S}} \hat{P}_{xy} \overset{*}{=} \frac{1}{\pi_x} \sum_{y \in \mathcal{S}} \pi_y P_{yx} \overset{\text{inv}}{=} \frac{1}{\pi_x} \pi_x = 1.$$

---

*Lembre que $\pi_x > 0 \; \forall x \in \mathcal{S}$.

Dem. do Teorema 1 (cont)

2) π é invariante para $\hat{\mathbf{P}}$: dado $x \in \mathcal{S}$,

$$\sum_{x\in\mathcal{S}} \pi_x \hat{P}_{xy} = \sum_{x\in\mathcal{S}} \pi_y P_{yx} = \pi_y \overbrace{\sum_{x\in\mathcal{S}} P_{yx}}^{1} = \pi_y .$$

3) PM: Dados $x_0, \ldots, x_N \in \mathcal{S}$,

$$\begin{aligned}
& \mathbb{P}(Y_0 = x_0, Y_1 = x_1 \ldots, Y_N = x_N) \\
&= \mathbb{P}(X_0 = x_N, X_1 = x_{N-1} \ldots, X_N = x_0) \\
&= \pi_{x_N} P_{x_N x_{N-1}} \cdots P_{x_1 x_0} && (1) \\
&= \underbrace{\frac{\pi_{x_N}}{\pi_{x_{N-1}}} P_{x_N x_{N-1}}}_{\hat{P}_{x_{N-1} x_N}} \pi_{x_{N-1}} P_{x_{N-1} x_{N-2}} \cdots P_{x_1 x_0} = \ldots = \\
&= \hat{P}_{x_{N-1} x_N} \hat{P}_{x_{N-2} x_{N-1}} \cdots \hat{P}_{x_0 x_1} \pi_{x_0}, && (2)
\end{aligned}$$

e concluímos da Proposição 1, 1º álbum, que

$$(Y_n)_{0\leq n\leq N} \sim \mathrm{CM}(\pi, \hat{\mathbf{P}}).$$

Dem. do Teorema 1 (cont)

4) Irredutibilidade: Da irredutibilidade de $\mathbf{P}$, temos que, dados $x, y \in \mathcal{S}$, existem $n \geq 0$ e $x = x_0, \ldots, x_n = y$ tais que

$$P_{x_0 x_1} \cdots P_{x_{n-1} x_n} > 0.$$

Então $\hat{P}_{x_n x_{n-1}} \cdots \hat{P}_{x_1 x_0} \stackrel{(1,2)}{=} \frac{\pi_{x_0}}{\pi_{x_n}} P_{x_0 x_1} \cdots P_{x_{n-1} x_n} > 0.$ $\square$

Def. Dadas uma matriz estocástica $\mathbf{P}$ e uma medida μ, dizemos que μ e $\mathbf{P}$ estão *em equilíbrio detalhado* (ED), se

$$\mu_x P_{xy} = \mu_y P_{yx}, \ \forall x, y \in \mathcal{S} \tag{3}$$

Lema 1. Se μ e $\mathbf{P}$ estiverem em equilíbrio detalhado, então μ é invariante para $\mathbf{P}$.

Dem. Somando (3) em $y \in \mathcal{S}$:

$$\sum_{y \in \mathcal{S}} \mu_y P_{yx} = \sum_{y \in \mathcal{S}} \mu_x P_{xy} = \mu_x \overbrace{\sum_{y \in \mathcal{S}} P_{xy}}^{1} = \mu_x.$$ $\square$

Reversibilidade

Def. Dada $\mathbf{X} \sim \mathrm{CM}(\mu, \mathbf{P})$, dizemos que $\mathbf{X}$ é *reversível* se para todo $N \geq 0$, $(X_{N-n})_{0 \leq n \leq N} \sim \mathrm{CM}(\mu, \mathbf{P})$.

Teorema 2. Sejam $\mathbf{P}$ uma matriz estocástica irredutível e μ uma probabilidade em $\mathcal{S}$. Suponha que $\mathbf{X} \sim \mathrm{CM}(\mu, \mathbf{P})$. Então as duas afirmações a seguir são equivalentes.

(i) $\mathbf{X}$ é reversível;

(ii) μ e $\mathbf{P}$ estão em equilíbrio detalhado.

Dem. (ii $\Rightarrow$ i) Do Lema 1, temos que μ é invariante para $\mathbf{P}$; de (3) e da definição de $\hat{\mathbf{P}}$, temos que $\hat{\mathbf{P}} = \mathbf{P}$. Do Teorema 1, temos que $(Y_n)_{0 \leq n \leq N} \sim \mathrm{CM}(\mu, \mathbf{P})$.

(i $\Rightarrow$ ii) Da definição de reversibilidade com $N = 1$:

$$\underbrace{\mathbb{P}(X_0 = x, X_1 = y)}_{\mu_x P_{xy}} = \underbrace{\mathbb{P}(X_0 = y, X_1 = x)}_{\mu_y P_{yx}} \qquad \square$$

Reversibilidade (cont)

Obs. As identidades em (3) podem ser tratadas como um sistema de equações lineares para μ, cujas soluções, se houver, são medidas invariantes para $\mathbf{P}$. Se estas medidas invariantes forem finitas, então podem ser normalizadas para fornecer uma distribuição invariante.

Exemplos. 1) $\mathbf{P} = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}$

$\mathbf{P}$ é irredutível e duplamente estocástica[†], logo, $\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ é a distribuição invariante.

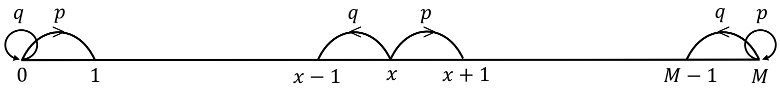
Mas π e $\mathbf{P}$ não estão em equilíbrio detalhado e logo a cadeia não é reversível.

[†] $\sum_{x \in \mathcal{S}} P_{xy} = 1 \ \forall \ y \in \mathcal{S}$

Exemplos (cont)

2) PAS com reflexão na fronteira

$$p = 1 - q \in (0,1)$$



As equações de ED são: $\mu_x \overbrace{P_{xx+1}}^{p} = \mu_{x+1} \overbrace{P_{x+1x}}^{q}$, $x = 0, \ldots, M-1$.

Equivalentemente: $\mu_{x+1} = \frac{p}{q}\mu_x$, $x = 0, \ldots, M-1$.

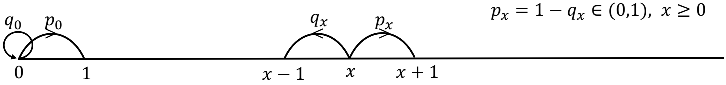
Uma solução: $\mu_x = \left(\frac{p}{q}\right)^x$, $x = 0, \ldots, M$.

Logo, μ normalizada, π, digamos, é a distribuição invariante,
e a cadeia começando de π é reversível ($\forall p \in (0,1)$).

2′) $M = \infty$: resultado vale se $0 < p < \frac{1}{2}$.

Exemplos (cont)

3) Processo de Nascimento e Morte (PNM)



Equações de ED: $\mu_x p_x = \mu_{x+1} q_{x+1} \Leftrightarrow \mu_{x+1} = \rho_x \mu_x$, $\rho_x = \frac{p_x}{q_{x+1}}$, $x \geq 0$.

Uma solução: $\mu_0 = 1$, $\mu_x = \prod_{y=0}^{x-1} \rho_y = \prod_{y=0}^{x-1} \frac{p_x}{q_{x+1}}$, $x \geq 1$.

Logo, se $\sum_{x\geq 1} \mu_x = \sum_{x\geq 0} \prod_{y=0}^{x} \frac{p_y}{q_{y+1}} < \infty$, então $\pi = \mu$ normalizada é a distribuição invariante, e a cadeia começando de π é reversível.

Em particular, se $p_x \equiv p < 1/2$, então π é a distribuição Geométrica com parâmetro de sucesso $1 - \rho := 1 - \frac{p}{q} > 0$.

Exemplo 3. **Obs.**

Pode-se mostrar que neste modelo a condição necessária e suficiente para transitoriedade é $\sum_{x\geq 0}\prod_{y=0}^{x}\frac{q_y}{p_y}<\infty$.

Segue a seguinte classificação do PNM:

i) Se $\sum_{x\geq 0}\prod_{y=0}^{x}\frac{q_y}{p_y}<\infty$, então o PNM é transitório;

ii) Se $\sum_{x\geq 0}\prod_{y=0}^{x}\frac{q_y}{p_y}=\sum_{x\geq 0}\prod_{y=0}^{x}\frac{p_y}{q_{y+1}}=\infty$, então o PNM é recorrente nulo;

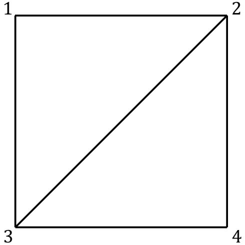
iii) Se $\sum_{x\geq 0}\prod_{y=0}^{x}\frac{p_y}{q_{y+1}}<\infty$, então o PNM é recorrente positivo.

O caso $p_x\equiv p<1/2$ foi discutido no slide anterior; se $p=1/2$, então segue do critério acima que o PNM é recorrente nulo; e se $p>1/2$, então temos transitoriedade.

Exemplos (cont)

4) PAS num grafo conexo

Seja $G = (\mathcal{S}, \mathcal{E})$ um grafo, em que $\mathcal{S}$ são os sítios[‡] e $\mathcal{E}$ os elos (conectando pares de sítios de $\mathcal{S}$). Por exemplo:



$$\mathcal{S} = \{1, 2, 3, 4\}$$

$$\mathcal{E} = \{(1,2), (1,3), (2,3), (2,4), (3,4)\}$$

Para $x \in \mathcal{S}$, seja $V_x = \{y \in \mathcal{S} : (x, y) \in \mathcal{E}\}$ o conjunto de vizinhos mais próximos de x, e v_x = *valência* de $x = \# V_x$, o número de tais vizinhos.

Como G é suposto conexo: $v_x > 0 \; \forall x \in \mathcal{S}$. Vamos supor adicionalmente que $v_x < \infty \; \forall x \in \mathcal{S}$.

[‡] $|\mathcal{S}| \geq 2$

Exemplo 4 (cont)

CM: A partir de $x \in \mathcal{S}$, a cadeia salta para um dos vizinhos de x uniformemente ao acaso, isto é,

$$P_{xy} = \frac{1}{v_x},\ y \in V_x.$$

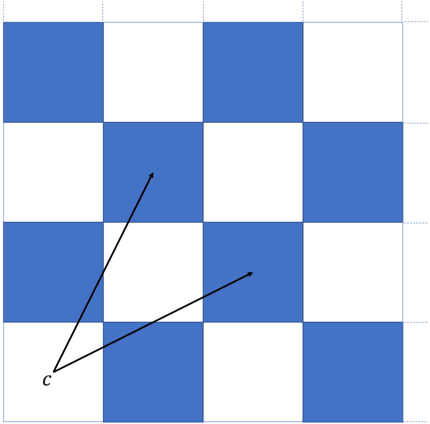
Então, a medida $\nu := \{v_x,\ x \in \mathcal{S}\}$ está em equilíbrio detalhado com $\mathbf{P}$ (claramente); logo, *se* $\mathcal{S}$ *for finito*, como $\mathbf{P}$ é irredutível,

$$\pi = \tfrac{1}{\sigma}\nu,\ \sigma = \textstyle\sum_{x \in \mathcal{S}} v_x, \text{ é a distribuição invariante de } \mathbf{P},$$

e $\mathbf{X} \sim \mathrm{CM}(\pi, \mathbf{P})$ é reversível.

Exemplos (cont)

4′) *Cavalo de xadrez aleatório*



A cada salto, cavalo faz todo movimento permitido com igual probabilidade.

Começando de um canto, quanto tempo em média leva para voltar?

Exemplo 4′ (cont)

$v_c = 2$, e verifique que a CM é irredutível, e que no tabuleiro (completo) de xadrez há

- 4 casas com valência 2,
- 8 casas com valência 3,
- 20 casas com valência 4,
- 16 casas com valência 6 e
- 16 casas com valência 8

$$\begin{aligned}
\therefore \quad \mathbb{E}_c(T_c) &= \frac{1}{\pi_c} = \frac{1}{\frac{v_c}{\sum_{x \in \mathcal{S}} v_x}} = \frac{\sum_{x \in \mathcal{S}} v_x}{v_c} \\
&= \frac{4 \times 2 + 8 \times 3 + 20 \times 4 + 16 \times 6 + 16 \times 8}{2} \\
&= 168
\end{aligned}$$

Teorema Ergódico

Seja $(X_n) \sim \text{CM}(\cdot, \mathbf{P})$ em $\mathcal{S}$. Para $x \in \mathcal{S}$ e $n \geq 1$, seja

$$V_x(n) = \sum_{i=0}^{n-1} \mathbb{1}\{X_i = x\}$$

o número de visitas de (X_n) a x até o instante $n-1$.

Teorema 2 (Teorema Ergódico)

Suponha $\mathbf{P}$ irredutível, e tomemos μ uma probabilidade qualquer em $\mathcal{S}$. Seja $(X_n) \sim \text{CM}(\mu, \mathbf{P})$.

a) Então, para todo $x \in \mathcal{S}$: $\dfrac{V_x(n)}{n} \underset{n\to\infty}{\longrightarrow} \dfrac{1}{m_x}$ qc.

b) Se além disto, $\mathbf{P}$ for rec pos, então $\forall\, f : \mathcal{S} \to \mathbb{R}$ limitada, temos:

$$\frac{1}{n}\sum_{i=0}^{n-1} f(X_i) \xrightarrow[n\to\infty]{\text{qc}} \bar{f} := \mathbb{E}_\pi\big(f(X_0)\big) = \sum_{x\in\mathcal{S}} f_x \pi_x,$$

onde π é a distribuição invariante para $\mathbf{P}$.

Dem. do Teorema 2

a) Se $\mathbf{P}$ for transitória, então $V_x(\infty) = \lim_{n\to\infty} V_x(n) < \infty$ qc.

$$\therefore \quad \limsup_{n\to\infty} \frac{V_x(n)}{n} \leq \limsup_{n\to\infty} \frac{V_x(\infty)}{n} = 0 \text{ qc}$$

Se $\mathbf{P}$ for recorrente, então, dado $x \in \mathcal{S}$, seja H^x o tempo de chegada a x. Pela PFM, $(X_{H_x+n})_{n\geq 0} \sim \mathrm{CM}(\delta_x, \mathbf{P})$ [§], e

$$\left| \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}\{X_i = x\} - \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}\{X_{H^x+i} = x\} \right|$$
$$= \frac{1}{n}\sum_{i=n}^{n-1+H^x} \mathbb{1}\{X_i = x\}^{\P} \leq \frac{H^x}{n} \underset{n\to\infty}{\longrightarrow} 0 \text{ qc},$$

já que $H^x < \infty$ qc.

Basta então considerar o caso em que $\mu = \delta_x$.

[§] δ_x é a distribuição em $\mathcal{S}$ que atribui probabilidade 1 a x.
[¶] $\sum_{i=n}^{n-1} \cdots = 0$ por convenção.

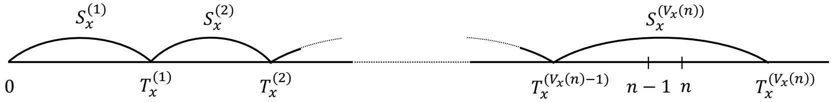
Dem. do Teorema 2 (cont)

Sejam $T_x^{(0)} = 0$, e para $r \geq 1$, $T_x^{(r)} = \inf\{n > T_x^{(r-1)} : X_n = x\}$,

e $S_x^{(r)} = T_x^{(r)} - T_x^{(r-1)}$. Pela PFM:

$S_x^{(r)}$, $r \geq 1$, é uma sequência iid com $\mathbb{E}(S_x^{(1)}) = \mathbb{E}_x(T_x) = m_x$.

Note que $T_x^{(r)} = \sum_{i=1}^r S_x^{(i)}$ e $T_x^{(V_x(n)-1)} < n \leq T_x^{(V_x(n))}$.



Logo, $$\frac{V_x(n)-1}{V_x(n)} \frac{T_x^{(V_x(n)-1)}}{V_x(n)-1} \leq \frac{n}{V_x(n)} \leq \frac{T_x^{(V_x(n))}}{V_x(n)}.$$

Dem. do Teorema 2 (cont)

Como $V_x(n) \underset{n\to\infty}{\longrightarrow} \infty$ qc (pela recorrência), temos, da Lei Forte dos Grandes Números, que o quociente do lado dir e o 2^o quociente do lado esq, ambos, $\underset{n\to\infty}{\longrightarrow} m_x$ qc [‖], e o 1^o quoc do lado esq $\underset{n\to\infty}{\overset{\text{qc}}{\longrightarrow}} 1$.

Logo, $\dfrac{n}{V_x(n)} \underset{n\to\infty}{\overset{\text{qc}}{\longrightarrow}} m_x$, e, como $m_x > 0$, $\dfrac{V_x(n)}{n} \underset{n\to\infty}{\overset{\text{qc}}{\longrightarrow}} \dfrac{1}{m_x}$. $\square_{a)}$

b) Podemos supor $max_{x\in\mathcal{S}}|f(x)| \leq 1$. Então

$$\left|\frac{1}{n}\sum_{i=0}^{n-1} f(X_i) - \bar{f}\right| \leq \sum_{x\in\mathcal{S}}\left|\left(\frac{V_x(n)}{n} - \pi_x\right) f_x\right| \leq \sum_{x\in\mathcal{S}}\left|\frac{V_x(n)}{n} - \pi_x\right|$$

$$= \sum_{x\in\mathcal{U}}\left|\frac{V_x(n)}{n} - \pi_x\right| + \sum_{x\in\mathcal{S}\setminus\mathcal{U}}\left|\frac{V_x(n)}{n} - \pi_x\right|, \tag{4}$$

onde $\mathcal{U} \subset \mathcal{S}$ a ser escolhido mais abaixo.

[‖] Mesmo quando $m_x = \infty$, como ocorre no caso recorrente nulo.

Dem. do Teorema 2 (cont)

A última soma em (4) pode ser cotada superiormente por

$$\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\left(\frac{V_x(n)}{n}+\pi_x\right)=\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\left(\frac{V_x(n)}{n}-\pi_x\right)+2\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\pi_x$$

$$=-\sum_{x\in\mathcal{U}}\left(\frac{V_x(n)}{n}-\pi_x\right)+2\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\pi_x$$

$$\leq\sum_{x\in\mathcal{U}}\left|\frac{V_x(n)}{n}-\pi_x\right|+2\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\pi_x$$

Logo, (4) pode ser cotada superiormente por

$$2\sum_{x\in\mathcal{U}}\left|\frac{V_x(n)}{n}-\pi_x\right|+2\sum_{x\in\mathcal{S}\setminus\mathcal{U}}\pi_x.$$

Dem. do Teorema 2 (cont)

Agora, dado $\varepsilon > 0$, escolhamos $\mathcal{U}$ finito tq $\sum_{x \in \mathcal{S} \setminus \mathcal{U}} \pi_x \leq \frac{\varepsilon}{2}$

(o que é possível por π ser probabilidade). Então

$$\limsup_{n\to\infty} \left| \frac{1}{n} \sum_{i=0}^{n-1} f(X_i) - \bar{f} \right| \leq 2 \sum_{x \in \mathcal{U}} \left| \overbrace{\lim_{n\to\infty} \frac{V_x(n)}{n} - \pi_x}^{=0,\ \text{por a)}} \right| + \varepsilon$$

$$= \varepsilon$$

e o resultado segue de ε ser arbitrário. $\square$

Aplicações

Votando ao exemplo do passeio aleatório num grafo conexo finito, podemos usar a *frequência empírica* $\frac{1}{n}V_n(x)$ para *estimar* a valência relativa do sítio x (já que, como segue do Teorema Ergódico,

$$\frac{v_x}{\sum_{y\in\mathcal{S}} v_y} \sim \frac{1}{n}V_n(x) \text{ quando } n \text{ é bastante grande).}$$

Ideias como esta são usadas por máquinas de busca na internet (para estimar quais são as páginas mais acessadas).

Outros aplicações dessas ideias se dão na simulação de distribuições de probabilidade, que podem ser difíceis de estimar mais diretamente (p.ex., em reconhecimento de imagens). Uma opção é achar uma Cadeia de Markov mais fácil de simular (digamos irredutível, recorrente positiva e aperiódica) cuja distribuição invariante é a distr de interesse (MCMC):

Do Teorema Ergódico, de novo, $\frac{1}{n}V_n(x) \sim \pi_x$, o peso em x da distribuição de interesse.

Métodos para achar tais cadeias existem em vários casos, normalmente envolvendo cadeias reversíveis (p.ex., o método de Metropolis-Hastings, mas há muitos outros).