Introdução aos Processos Estocásticos

Luiz Renato Fontes

www.ime.usp.br/∼lrenato

Livro: James Norris, *Markov Chains*, Cambridge U Press

Processo Estocástico

- Família de variáveis aleatórias: $(X_t)_{t\in\mathcal{T}}$; $X_t \in \mathcal{S}$, $t \in \mathcal{T}$;
- $\mathcal{T}$ um cj de índices/tempos; a priori, arbitrário; neste curso:
  - $\mathcal{T} = \mathbb{N} = \{0, 1, \dots\}$ (ou $\mathbb{Z}$): tempo discreto; ou
  - $\mathcal{T} = \mathbb{R}^+ = [0, \infty)$; tempo contínuo;
  - Notação:
    - 1^o caso: $(X_n)_{n\in\mathbb{N}} = (X_n)_{n\geq 0} = \{X_0, X_1, \dots\}$;
    - 2^o caso: $(X_t)_{t\in\mathbb{R}^+} = (X_t)_{t\in[0,\infty)} = (X_t)_{t\geq 0}$;
- $\mathcal{S}$ é o espaço de estados, a priori arbitrário; neste curso:
  - $\mathcal{S}$ enumerável; muitas vezes $\mathcal{S} \subset \mathbb{N}$;
- $\mathbf{X} := (X_t)_{t\in\mathcal{T}}$ é uma fç aleatória; $X : \mathcal{T} \to \mathcal{S}$;
  - qdo $\mathcal{T} = \mathbb{R}^+$: X contínua à direita, com limites à esquerda;
    - ex: processo de salto.

Distribuição de probabilidades de $\mathbf{X}$

Caracterizada por suas distribuições *finito-dimensionais*:

$$\mathbb{P}(X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_n} = x_n),$$

$x_i \in \mathcal{S},\ t_i \in \mathcal{T},\ i = 1, \ldots, n,\ n \geq 0.$

Cadeias de Markov

$(X_t)_{t \in \mathcal{T}}$ é uma *Cadeia de Markov* se valer a *proprie// de Markov*:

$$\mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1}, \ldots, X_0 = x_0)$$

$$= \mathbb{P}(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n),$$

$x_1, x_2, \ldots \in \mathcal{S};\ 0 \leq t_1 < t_2 < \ldots \in \mathcal{T};\ n \geq 1.$

Homogeneidade temporal

$\mathbb{P}(X_t = y | X_s = x) = \mathbb{P}(X_{t-s} = y | X_0 = x),\ x, y \in \mathcal{S};\ 0 \leq s < t \in \mathcal{T}.$

Cadeia de Markov em tempo discreto ($\mathcal{T} = \mathbb{N}$)

X cadeia de Markov (CM) em $\mathcal{S}$ em tempo discreto.

Distribuição inicial

Seja μ uma distribuição de probabilidade em $\mathcal{S}$, i.e.,

$$\mu : \mathcal{S} \to [0,1]; \quad \sum_{x \in \mathcal{S}} \mu(x) = 1.$$

Dizemos que μ é a distribuição inicial de X se

$$\mathbb{P}(X_0 = x) = \mu(x),\ x \in \mathcal{S}.$$

Matriz de transição

$$P(x,y) := \mathbb{P}(X_1 = y | X_0 = x),\ \ x, y \in \mathcal{S}$$

Cadeia de Markov (em tempo discreto)

Obs. $\mathbf{P} = (P(x,y),\ x,y \in \mathcal{S})$ é uma *matriz estocástica*, i.e.

$$P(x,y) \in [0,1],\ x,y \in \mathcal{S};\ \sum_{y \in \mathcal{S}} P(x,y) = 1,\ x \in \mathcal{S}.$$

Proposição 1. A distribuição inicial μ e a matriz de transição $\mathbf{P}$ determinam a distribuição de probabilidades de $\mathbf{X}$.

Dem. Basta mostrar que as distribuições finito-dimensionais $\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n)$, $x_0, \ldots, x_n \in \mathcal{S}$, $n \geq 1$ são determinadas por μ e $\mathbf{P}$. De fato, condicionando sucessivamente, temos que a última probabilidade vale

$$\begin{aligned}
&\mathbb{P}(X_0 = x_0) P(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_2 = x_2 | X_1 = x_1, X_0 = x_0) \times \\
&\quad \cdots \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) \\
&\overset{\text{Markov}}{=} \mathbb{P}(X_0 = x_0) \mathbb{P}(X_1 = x_1 | X_0 = x_0) \cdots \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}) \\
&= \mu(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n).
\end{aligned}$$

□

Cadeia de Markov (cont.)

Proposição 1′. Reciprocamente, dadas uma distribuição de probabilidade μ e uma matriz estocástica $\mathbf{P}$ em $\mathcal{S}$, existe uma Cadeia de Markov $\mathbf{X}$ em $\mathcal{S}$ com distribuição inicial μ e matriz de transição $\mathbf{P}$.

Dem. Basta notar que $\mu(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n)$, $x_0, \ldots, x_n \in \mathcal{S}$, $n \geq 0$, formam uma família *consistente* de distribuições finito-dimensionais, e invocar o Teorema de Kolmogorov. $\square$

Notação. $X \sim \mathrm{CM}(\mu, \mathbf{P})$: X é uma/a Cadeia de Markov em $\mathcal{S}$ com dist inicial μ e matriz de transição $\mathbf{P}$.

Transição em n passos

Def.

$P^{(n)}(x,y) = \mathbb{P}(X_n = y | X_0 = x)$ ... prob de transição em n passos de x a y, $x, y \in \mathcal{S}$, $n \geq 1$

$\mathbf{P}^{(n)} = (P^{(n)}(x,y),\ x, y \in \mathcal{S})$ ... matriz de transição em n passos

Obs.

1)Verifique que $P^{(n)}(x,y) = \mathbb{P}(X_{n+m} = y | X_m = x)$, $m, n \geq 1$;

2) $\mathbf{P}^{(1)} = \mathbf{P}$; $\mathbf{P}^{(0)} := I$, a matriz identidade;

3) Tb usaremos a notação $P^{(n)}_{xy} = P^{(n)}(x,y)$;

4) Tb usaremos a notação $\mathbb{P}_\mu = \mathbb{P}$ para indicar a distr inicial μ.

Transição em n passos (cont.)

Proposição 2. $\mathbf{P}^{(n)} = \mathbf{P}^n$, $n \geq 0$

Dem. Casos $n = 0$ e 1 são claros. Suponha a tese válida para $n \geq 1$; então, introduzindo a notação $\mathbf{P}^n =: (P^n_{xy})_{x,y\in\mathcal{S}}$,

$$
\begin{aligned}
P^{(n+1)}_{xy} &= \mathbb{P}(X_{n+1} = y | X_0 = x) = \sum_{z\in\mathcal{S}} \mathbb{P}(X_{n+1} = y, X_n = z | X_0 = x) \\
&\overset{\text{cond.}}{=} \sum_{z\in\mathcal{S}} \mathbb{P}(X_{n+1} = y | X_n = z, X_0 = x)\mathbb{P}(X_n = z | X_0 = x) \\
&\overset{\text{Markov}}{=} \sum_{z\in\mathcal{S}} \mathbb{P}(X_{n+1} = y | X_n = z) P^{(n)}_{xz} \\
&\overset{\text{hip. ind.}}{=} \sum_{z\in\mathcal{S}} P_{zy} P^n_{xz} = P^{n+1}_{xy}.
\end{aligned}
$$

Transição em n passos (cont.)

Obs.

1) $\mathbf{P}^n$ é uma matriz estocástica, $n \geq 0$.

2) Valem as *Equações de Chapman-Kolmogorov*

$$\mathbb{P}(X_{m+n} = y | X_0 = x) = P^{(m+n)}(x, y) \overset{\text{Prop 2}}{=} P^{m+n}(x, y)$$
$$= (\mathbf{P}^m \mathbf{P}^n)(x, y) = \sum_{z \in \mathcal{S}} P^m_{xz} P^n_{zy} = \sum_{z \in \mathcal{S}} P^{(m)}_{xz} P^{(n)}_{zy}$$
$$= \sum_{z \in \mathcal{S}} \mathbb{P}(X_m = z | X_0 = x) \mathbb{P}(X_n = y | X_0 = z),$$

$x, y \in \mathcal{S}$, $n, m \geq 1$.

3) Distribuição de X_n

$$\mathbb{P}(X_n = y) = \sum_{x \in \mathcal{S}} \mathbb{P}(X_0 = x) \mathbb{P}(X_n = y | X_0 = x)$$
$$= \sum_{x \in \mathcal{S}} \mu(x) P^{(n)}_{xy} = (\mu \mathbf{P}^{(n)})(y) = (\mu \mathbf{P}^n)(y), \; y \in \mathcal{S}$$

Em outras palavras, $\mathbb{P}(X_n = \cdot) = \mu \mathbf{P}^n$, $n \geq 0$.

Exemplos

Há exemplos (mais ou menos simples) em que se pode obter as probs de trans em n passos explicitamente.

1) CM com dois estados: $\mathcal{S} = \{1, 2\}$; $\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$,

ie, $P_{11} = 1 - \alpha$, $P_{12} = \alpha$, $P_{21} = \beta$, $P_{22} = 1 - \beta$; $\alpha, \beta \in [0, 1]$.
Suponha que $\alpha + \beta > 0$, se não $\mathbf{P} = I$, e temos um caso trivial.
De $\mathbf{P}^{(n+1)} = \mathbf{P}^{(n)}\mathbf{P}$, segue

$$\begin{aligned} P_{11}^{(n+1)} &= P_{11}^{(n)} P_{11} + P_{12}^{(n)} P_{21} = P_{11}^{(n)} (1-\alpha) + P_{12}^{(n)} \beta \\ &= (1-\alpha) P_{11}^{(n)} + \beta \left(1 - P_{11}^{(n)}\right) = (1 - \alpha - \beta) P_{11}^{(n)} + \beta, \end{aligned}$$

$n \geq 0$.

Exemplo 1 (cont.)

A equação de diferença

$$\begin{cases} P_{11}^{(n+1)} = (1-\alpha-\beta)P_{11}^{(n)} + \beta, \ n \geq 0; \\ P_{11}^{(0)} = 1, \end{cases}$$

tem como solução $P_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n, \ n \geq 0.$

Exemplos (cont.)

$1'$) $N \geq 2$ variedades de um vírus. A cada geração o vírus pode se manter na mesma variedade, ou sofrer uma mutação, o que acontece com prob $\alpha \in (0, 1]$, e, neste caso, uma das outras $N - 1$ variedades ocorre com distribuição uniforme. Qual a probabilidade de que a variedade da n-ésima geração seja a mesma do que a inicial?
Temos que as sucessivas variedades do vírus ao longo das gerações pode ser descrita por uma CM em $\mathcal{S} = \{1, \ldots, N\}$, com matriz de transição $\mathbf{P}$ tal que

$$P_{xx} = 1 - \alpha, \; P_{xy} = \frac{\alpha}{N - 1}, \; x \neq y \in \mathcal{S}, \; \alpha > 0.$$

Pela simetria do modelo, temos que $P_{xx}^{(n)}$ não depende de $x \in \mathcal{S}$, e logo $P_{11}^{(n)}$ é probabilidade que procuramos.

Exemplo 1′ (cont.)

Podemos ainda simplificar este problema da seguinte forma. Seja

$$\tilde{X}_n = \begin{cases} 1, & \text{se } X_n = 1; \\ 2, & \text{se } X_n \neq 1. \end{cases}$$

Verifica-se (faça-o) que $(\tilde{X}_n)$ é uma CM com matriz de trans

$$\tilde{\mathbf{P}} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \ \beta = \frac{\alpha}{N-1}.$$

Com isto, temos finalmente que

$$P_{11}^{(n)} = \tilde{P}_{11}^{(n)} = \frac{1}{N} + \left(1 - \frac{1}{N}\right)\left(1 - \alpha\frac{N}{N-1}\right)^n, \ n \geq 0.$$

Exemplos (cont.)

2) $\mathcal{S} = \{1, 2, 3\}$ (CM com 3 estados)

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

Vamos diagonalizar $\mathbf{P}$. Eq característica:

$$\det(xI - \mathbf{P}) = x\left(x - \frac{1}{2}\right)^2 - \frac{1}{4} = (x-1)\left(x^2 + \frac{1}{4}\right) = 0$$

Autovalores: $\lambda_1 = 1$, $\lambda_2 = i/2$, $\lambda_3 = -i/2$, e logo

$$\mathbf{P} = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & i/2 & 0 \\ 0 & 0 & -i/2 \end{pmatrix} U^{-1},$$

para certa matriz inversível U.

Exemplo 2 (**Obs.**)

1) U pode ser escrita como (v_1^t, v_2^t, v_3^t), onde $v_i = (v_{i1}, v_{i2}, v_{i3})$ é um autovalor de $\mathbf{P}$ associado ao autovalor λ_i, $i = 1, 2, 3$.
2) $\lambda = 1$ é sempre um autovalor de $\mathbf{P}$, associado ao autovetor $(1, 1, 1)$, pelo fato de $\mathbf{P}$ ser estocástica.
3) De posse de U, temos que

$$\mathbf{P}^{(n)} = \mathbf{P}^n = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & (i/2)^n & 0 \\ 0 & 0 & (-i/2)^n \end{pmatrix} U^{-1}.$$

Note que $\left(\pm\frac{i}{2}\right)^n = \left(\frac{1}{2}e^{\pm i\frac{\pi}{2}}\right)^n = \left(\frac{1}{2}\right)^n \left\{\cos n\frac{\pi}{2} \pm i \sin n\frac{\pi}{2}\right\}$, e, p.ex.

$$P_{11}^{(n)} = \alpha + \left(\frac{1}{2}\right)^n \left\{\beta \cos n\frac{\pi}{2} + \gamma \sin n\frac{\pi}{2}\right\}, \ n \geq 0,$$

para certas constantes α, β, γ, que podem ser obtidas de U, mas...

Exemplo 2 (cont)

... vamos, alternativamente, simplesmente comparar a expressão acima à forma original de **P**, dois slides acima.

$P_{11}^{(0)} = 1 = \alpha + \beta, \quad P_{11}^{(1)} = 0 = \alpha + \gamma/2, \quad P_{11}^{(2)} = 0 = \alpha - \beta/4$

Segue que $\alpha = 1/5, \quad \beta = 4/5, \quad \gamma = -2/5$. Logo,

$$P_{11}^{(n)} = \frac{1}{5} + \left(\frac{1}{2}\right)^n \left\{\frac{4}{5}\cos n\frac{\pi}{2} - \frac{2}{5}\sin n\frac{\pi}{2}\right\}, \; n \geq 0,$$

e podemos similarmente obter expressões para $P_{xy}^{(n)}$, $x, y \in \mathcal{S}$.

Exemplo 2 (obs.)

Obs.

1) Caso geral ($\mathcal{S}$ finito). De posse dos autovalores de $\mathbf{P}$, $\lambda_1, \ldots, \lambda_m$, $m = \dim \mathcal{S}$, *se os autovals forem todos distintos*, podemos usar a abordagem acima, e obter

$$(*) \qquad P_{11}^{(n)} = \alpha_1 \lambda_1^n + \ldots + \alpha_m \lambda_m^n, \quad n \geq 0,$$

para certos coeficientes constantes $\alpha_1, \ldots, \alpha_m$, que podem ser obtidos de maneira similar ao que fizemos acima.

1′) No caso de autovals com multiplicidade ≥ 2, a representação $(*)$ pode se complicar, com o surgimento de coeficientes polinomiais em n (qdo $\mathbf{P}$ não for diagonalizável).

2) No Exemplo 1 pode ser resolvido desta forma — matrizes estocásticas 2×2 podem ser sempre diagonalizadas.

Perda de memória

Pode ocorrer que

$$\mathbf{P}^n \to \Pi = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}, \text{ qdo } n \to \infty,$$

onde π é uma dist de prob em $\mathcal{S}$. A conv acima significa que

$$P^{(n)}(x,y) = \mathbb{P}(X_n = y | X_0 = x) \to \pi(y),\ x, y \in \mathcal{S},$$

e, como o lado direito *não depende* de x, dizemos que *a memória* sobre a condição inicial da CM em questão *se perde* qdo $n \to \infty$.

Obs. Note que neste caso $\mathbb{P}(X_n = \cdot) = \mu\mathbf{P}^n \to \mu\Pi = \pi$ qdo $n \to \infty$ para qualquer ditribuição inicial μ.

Perda de memória (cont.)

No Exemplo 1 acima, se $0 < \alpha + \beta < 2$, então $\mathbf{P}^n \to \begin{pmatrix} \pi \\ \pi \end{pmatrix}$, onde $\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$ — isto é claro para $P_{11}^{(n)}$, mas pode ser verificado tb para as demais probs de trans.

No Ex. $1'$, temos perda de memória se $\alpha < 1$ ou $N \geq 3$.

No Ex. 2, é imediato da expressão obtida acima que $P_{11}^{(n)} \to \frac{1}{5}$; obtemos de forma análoga que $P_{xy}^{(n)} \to \pi_y$, $\forall x, y$ (com $\pi_1 = \frac{1}{5}$). Tb podemos fazer

$$\mathbf{P}^n \to U \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} U^{-1} =: \Pi.$$

Usando o fato que $U = (v_1^t, v_2^t, v_3^t)$, com $v_1 = (1, 1, 1)$, concluímos prontamente que Π_{xy} não depende de x.

Convergência

Um dos objetivos iniciais deste curso é estabelecer conds para convergência de $\mathbf{P}^n$ com perda de memória, ie

$$\mathbf{P}^n \to \Pi, \quad \text{qdo } n \to \infty, \text{ com } \Pi_{xy} = \pi_y, \forall x, y \in \mathcal{S},$$

e de tal forma que $\sum_{y \in \mathcal{S}} \pi_y = 1$.

P.ex., no Ex. 1, se $\alpha + \beta = 0$, então temos convergência, mas sem perda de memória; se $\alpha + \beta = 2$, então não temos convergência.

As abordagens diretas dos Exs. 1 e 2 não funcionam bem em geral, mesmo para $|\mathcal{S}| < \infty$, porém, neste caso, da teoria de matrizes estocásticas/não negativas, temos que todos os avals de $\mathbf{P}$ são ≤ 1 em módulo. Lembrando que 1 é sempre aval associado a avet $(1, \dots, 1)$, se ocorrer de todos os demais $m - 1$ avals serem < 1 em módulo, então

$$\mathbf{P}^n = UJ^nU^{-1} \to \begin{pmatrix} 1 & \cdots \\ \vdots & \vdots \\ 1 & \cdots \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} U^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} U^{-1} =: \Pi$$

Convergência (cont.)

Note que $\Pi = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}$, onde π é a primeira linha de U^{-1};

Π tem que ser estocástica, logo π é uma prob em $\mathcal{S}$; vê-se então temos conv com perda de memória.

Obs.

1) A conv acima é exponencialmente rápida, como nos exs de perda de memória vistos antes.

2) Uma cond suficiente para que todos os avals de $\mathbf{P}$ sejam < 1 em módulo (a menos de um aval) é que exista $n_0 \geq 1$ tal que se $n \geq n_0$, então $P^n_{xy} > 0 \ \forall x, y \in \mathcal{S}$.

Se $P_{xy} > 0 \ \forall x, y \in \mathcal{S}$, então a cond vale com $n_0 = 1$.

Estrutura de classes

$\mathbf{X} \sim \mathrm{CM}(\cdot, \mathbf{P})$ em $\mathcal{S}$. *Not:* $\mathbb{P}(\mathbf{X} \in \cdot | X_0 = x) =: \mathbb{P}_x(\mathbf{X} \in \cdot)$.

Dados $x, y \in \mathcal{S}$, dizemos que x *atinge* y, *not:* $x \to y$, se

$$\mathbb{P}_x(X_n = y \text{ para algum } n \geq 0) > 0.$$

E dizemos que x e y *se comunicam* (ou x *se comunica com* y), *not:* $x \leftrightarrow y$, se $x \to y$ e $y \to x$.

Obs. Note que $x \to x$, e logo $x \leftrightarrow x$, $x \in \mathcal{S}$.

Teorema. Para $x, y \in \mathcal{S}$, $x \neq y$, são equivalentes:
(i) $x \to y$;
(ii) $\exists\ n \geq 1$ e $x_0, x_1, \ldots, x_n \in \mathcal{S}$, com $x_0 = x$ e $x_n = y$, tq

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

(iii) $P^{(n)}_{xy} > 0$ para algum $n \geq 1$.

Estrutura de classes (cont.)

Dem. (i $\Rightarrow$ iii) Por subaditividade:

$$\sum_{n\geq 1}\mathbb{P}_x(X_n = y) \geq \mathbb{P}_x(\cup_{n\geq 1}\{X_n = y\}) \overset{(i)}{>} 0 \Rightarrow \text{(iii)}$$

(iii $\Rightarrow$ ii)

$$\begin{aligned}
&\mathbb{P}_x(X_n = y) \\
&= \sum_{x_i,\dots,x_{n-1}\in\mathcal{S}} \mathbb{P}_x(X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n = x_n) \\
&\overset{(*)}{=} \sum_{x_i,\dots,x_{n-1}\in\mathcal{S}} P_{xx_1}\cdots P_{x_{n-1}x_n} \overset{(iii)}{>} 0 \Rightarrow \text{(ii)}
\end{aligned}$$

(ii $\Rightarrow$ i) De (ii) e $(*)$ segue que $\mathbb{P}_x(X_n = y) > 0$; logo

$$\mathbb{P}_x(\cup_{\ell\geq 1}\{X_\ell = y\}) \geq \mathbb{P}_x(X_n = y) > 0. \qquad \square$$

Estrutura de classes (cont.)

Proposição. $\leftrightarrow$ é uma relação de equivalência em $\mathcal{S}$.

Dem.

1) $x \leftrightarrow x$, como já tínhamos observado. (*Reflexividade*)

2) Se $x \leftrightarrow y$ e $y \leftrightarrow z$, então $\exists\, m, n \geq 0$ tq $P^m_{xy} P^n_{yz} > 0$.

Logo, $P^{m+n}_{xz} = \sum_{y \in \mathcal{S}} P^m_{xw} P^n_{wz} \geq P^m_{xy} P^n_{yz} > 0$, e $x \to z$.

Similarmente, $z \to x$, e $x \leftrightarrow z$. (*Transitividade*)

3) É óbvio que $x \leftrightarrow y \Rightarrow y \leftrightarrow x$. (*Simetria*) $\square$

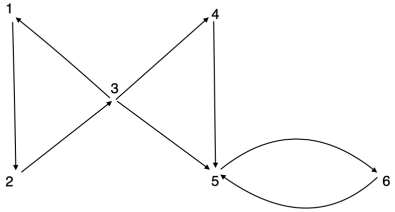
Logo, $\leftrightarrow$ particiona $\mathcal{S}$ em *classes de comunicação*.

Diremos que uma classe $\mathcal{C}$ de $\mathcal{S}$ é *fechada* se, dados $x \in \mathcal{C}$ e $y \in \mathcal{S}$, se $x \to y$, então $y \in \mathcal{C}$.

Exemplo

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

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



Classes (de comunicação): $\{1, 2, 3\}$, $\{4\}$ e $\{5, 6\}$.

$\{5, 6\}$ é a única classe fechada.

Dizemos que $x \in \mathcal{S}$ é *absorvente* se $\{x\}$ for uma classe fechada ($\Leftrightarrow P_{xx} = 1$).

Uma CM para a qual $\mathcal{S}$ for uma classe é dita *irredutível* ($\Leftrightarrow x \leftrightarrow y \ \forall x, y \in \mathcal{S}$).