Aula 12

Aula passada

- Cadeias de Markov
- Definição e exemplos
- Modelo On-Off
- Sem memória
- Distribuição no tempo
- Irredutibilidade
- Aperiodicidade

Aula de hoje

- Distribuição estacionária
- Tempo de chegada
- Distância de variação total
- Convergência
- Reversibilidade
- Passeios aleatórios

Cadeia de Markov

- Seja S o espaço de estados da CM, e P a matriz de transição de estados
- Seja X_t uma v.a. que determina o valor do estado da cadeia no instante de tempo t, para t = 0,1,2,…
- Estamos interessados em
  - $P[\, X_t = s \,] = \pi_s(t)$
- Sabemos que para um estado incial, temos

  $\pi(t) = \pi(t-1)P = \pi(0)P^t$

Mas para onde vai π(t)?

- será que depende em $\pi(0)$?

Possível Convergência

- Observação 1: X_t não converge
  - X_t passeia pela cadeia para sempre, trocando de estado
- Observação 2: $P[X_t = s_i]$ pode convergir
  - Prob. de encontrar X_t no estado s_i, ou fração de vezes que X_t visita o estado s_i (intuitivamente)
- Estamos interessados em valores grandes de t

$$\lim_{t \to \infty} \pi(t) = \lim_{t \to \infty} \pi(0) P^t$$

  - mas temos que formalizar este limite

Distribuição Estacionária

- Seja π um vetor de distribuição em uma CM com matriz de transição *P*
- Dizemos que π (vetor linha) é uma distribuição estacionária sse

$$\pi_s \geq 0 \qquad \sum_{s \in S} \pi_s = 1$$

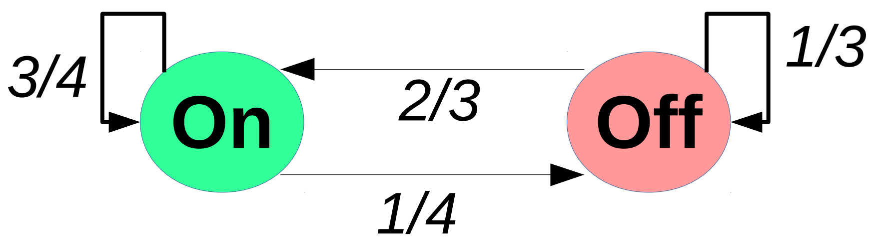
Vetor de probabilidade

$$\boxed{\pi P = \pi} \longrightarrow \pi_i = \sum_j \pi_j P_{ji}$$

, para todo *i*

- Ou seja, ao multiplicar π por *P* temos π de volta
  - "estacionou", não temos mais dinâmica!
- Repare que se $\pi(0)$ é uma distribuição estacionária, então $\pi(1) = \pi(0)P = \pi(0)$

Exemplo Modelo On-Off



$$P = \begin{pmatrix} .75 & .25 \\ .67 & .33 \end{pmatrix}$$

$$\pi = \begin{pmatrix} 8/11 & 3/11 \end{pmatrix}$$ ← Distribuição estacionária

- Verificando: πP = ?

$$\pi_1 = 8/11 * 3/4 + 3/11 * 2/3 = 8/11$$
$$\pi_2 = 8/11 * 1/4 + 3/11 * 1/3 = 3/11$$

- Se $\pi(0) = \pi$ então temos que $\pi(1) = \pi(0)P = \pi$
  - vetor de probabilidade está estacionado!

Tempo de Chegada

- T_{ij}: Tempo necessário para sair de um estado s_i e chegar a outro estado s_j (*hitting time*)
  - número de transições, T_{ij} é aleatório, $\tau_{ij} = E[\ T_{ij}\ ]$

$$T_{ij} = min\{t \mid X_t = s_j \wedge X_0 = s_i\}$$

- Se $i = j$, temos τ_{ii}, chamado de tempo médio de retorno ao estado s_i
  - número médio de transições para sair e voltar ao estado s_i
- T_{ij} não depende de $\pi(t)$ pois estamos condicionando em estar em s_i

Todos os estados aperiódicos

Tempo de Chegada

grafo fortemente conexo

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer dois estados s_i e s_j, temos o seguinte:

$$P[T_{ij} < \infty] = 1$$

$$P[T_{ij} = \infty] = 0$$

- Probabilidade de T_{ij} ser infinito é zero
- Não há chances de sair de s_i ficar circulando pela cadeia e nunca chegar a s_j

Em algum momento vou chegar em todos os estados

$$E[T_{ij}] = \tau_{ij} < \infty$$

- Valor esperado do tempo de retorno para qualquer estado s_i é finito
- Boa notícia!

Distribuição Estacionária e Tempo de Retorno

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer estado s_i, temos a seguinte relação

$$\boxed{\pi_i = \frac{1}{\tau_{ii}}}$$

- Relação entre tempo médio de retorno e distribuição estacionária
- Conhecer um determina o outro!
- **Intuição:** na média o processo visita s_i 1 vez a cada τ_{ii} passos

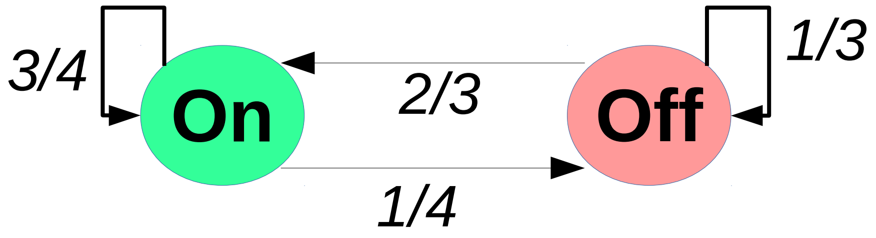
$\tau_{ii} = 1/\pi_i$

↳ probabilidade

Parece o valor esperado de uma geométrica

Figueiredo 2018

Exemplo



$$P = \begin{pmatrix} .75 & .25 \\ .67 & .33 \end{pmatrix}$$

$$\pi = \begin{pmatrix} 8/11 & 3/11 \end{pmatrix}$$

- Tempo médio de retorno do estado On (1) e Off (2) ?
- τ_{11} = $1/\pi_1$ = 11/8 = 1.375
- τ_{22} = $1/\pi_2$ = 11/3 = 3.666

Variação Total

- Como medir a distância entre dois vetores?
  - Precisamos disto para medir aproximação da distribuição estacionária
  - há muitas maneiras, uma delas é variação total
- Seja α e β dois vetores de probabilidade em *S*, a distância de variação total entre eles é dada por

$$d_{TV}(\alpha,\beta)=\frac{1}{2}\sum_k |\alpha_k-\beta_k|$$

d_{TV} tem valor entre 0 e 1

↳ normalização

- Sequência de vetores α_t converge para β em variação total sse

$$\lim_{t\to\infty} d_{TV}(\alpha_t,\beta)=0$$

A maior diferença acontece quando colocamos toda a massa de prob. de α em α_i e a de β em β_j

$\alpha = [0 \ 1]$

$\beta = [1 \ 0]$

$\rightarrow \sum |\alpha_k - \beta_k| = |1-0| + |0-1| = 2$

Convergência de CM

Teorema Fundamental das Cadeias de Markov

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer condição inicial $\pi(0)$, temos que

↳ sequência de vetores no tempo

$$\lim_{t \to \infty} d_{TV}(\pi(t), \pi) = 0$$

π é uma distribuição estacionária da CM

- Além disso, π é única! Corolário
  - CM sempre converge para a mesma distribuição estacionária, independente da condição inicial
  - Também chamado de estado estacionário, ou equilíbrio da CM

Figueiredo 2018

Encontrando π

- Ótima notícia, mas como encontrar π ?

1) Método iterativo

$$\pi(t)=\pi(t-1)P=\pi(0)P^t$$

- fazer iteração até critério de convergência

2) Método direto

$$\pi=\pi P$$

- resolver sistema de equações, adicionando a equação

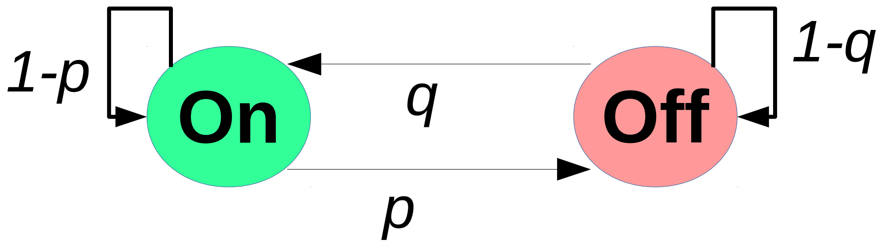
$$\sum_i \pi_i=1$$

3) Monte Carlo

- Usar própria cadeia para gerar amostras para estimar π_i ou estimar τ_{ii} para todo s_i

Modelo On-Off

- Distribuição estacionária?
- Sistemas de equações



$$\pi_i = \sum_j \pi_j P_{ji} \qquad \sum_i \pi_i = 1$$

$$\pi_1 = \pi_1 P_{11} + \pi_2 P_{21} = (1-p)\pi_1 + q\pi_2$$
$$\pi_2 = \pi_1 P_{12} + \pi_2 P_{22} = p\pi_1 + (1-q)\pi_2$$
$$\pi_1 + \pi_2 = 1$$

- Substituindo $\pi_2 = 1 - \pi_1$ na primeira equação, temos

$$\pi_1 = (1-p)\pi_1 + q(1-\pi_1) \rightarrow \pi_1 = \frac{q}{p+q}$$

$$e \quad \pi_2 = \frac{p}{p+q}$$

Reversibilidade

- Uma classe bem especial de CM
  - usada por muitos algoritmos, incluindo MCMC
- Reversível no tempo: evolução de X_t é o mesmo se t vai para frente ou para trás
- Uma CM é dita reversível para a distribuição de probabilidade π sse

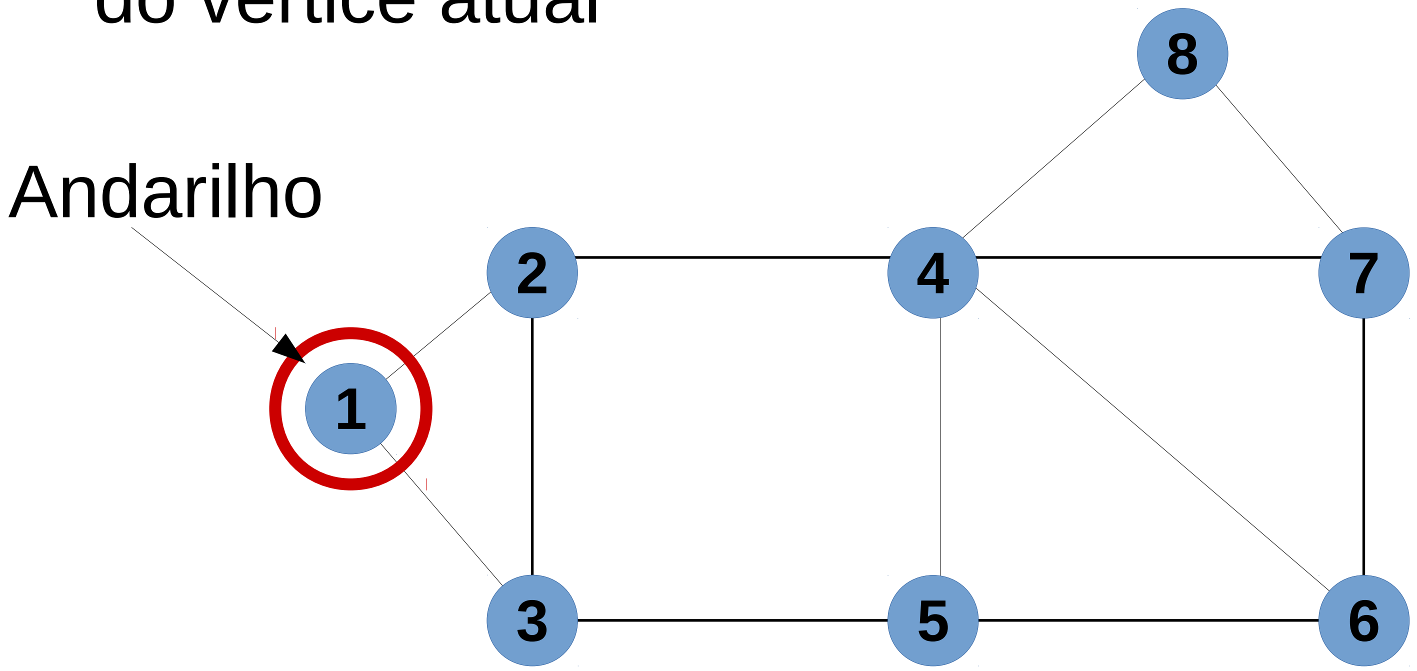
$$\pi_i P_{ij} = \pi_j P_{ji}$$

Massa de prob fluindo de s_i para s_j = Massa de prob fluindo de s_j para s_i

- Se π existe então π é também distribuição estacionária
- Uma forma ainda mais forte de equilíbrio!

Passeios Aleatórios

- Seja G=(V, E) um grafo não direcionado
- Considere um andarilho que passeia pelo grafo de forma aleatória, sem preferência e sem memória
  - escolhe uniformemente próximo vértice entre vizinhos do vértice atual



- p_{12} = ½, p_{13} = ½ , p_{21} = 1/3, p_{23} = 1/3, p_{24} = 1/3, ...

Passeios Aleatórios

- Nosso andarilho induz uma cadeia de Markov sobre o grafo G
  - forma de colocar o grafo em “movimento”
- X_t : vértice onde andarilho se encontra no tempo *t*
- Matriz de transição de probabilidade

$$P = D^{-1} A$$

  - D^{-1} é a matriz diagonal com o inverso do grau de cada vértice
  - *A* é a matriz de adjacência de G

- Outra forma: $P_{ij} = 1/d_i$ se (i, j) são vizinhos, 0 c.c.
  - onde d_i é o grau do vértice *i*

Passeios Aleatórios

- Qual a distribuição estacionária do andarilho ?

$$\pi_i = \frac{d_i}{K} \qquad K = \sum_i d_i = 2m$$

m é o número de arestas em G

- **Incrível:** só depende do grau do vértice *i*
  - *K* é constante de normalização
  - Todos os vértices com mesmo grau em igual prob
- Além disso, esta CM é reversível

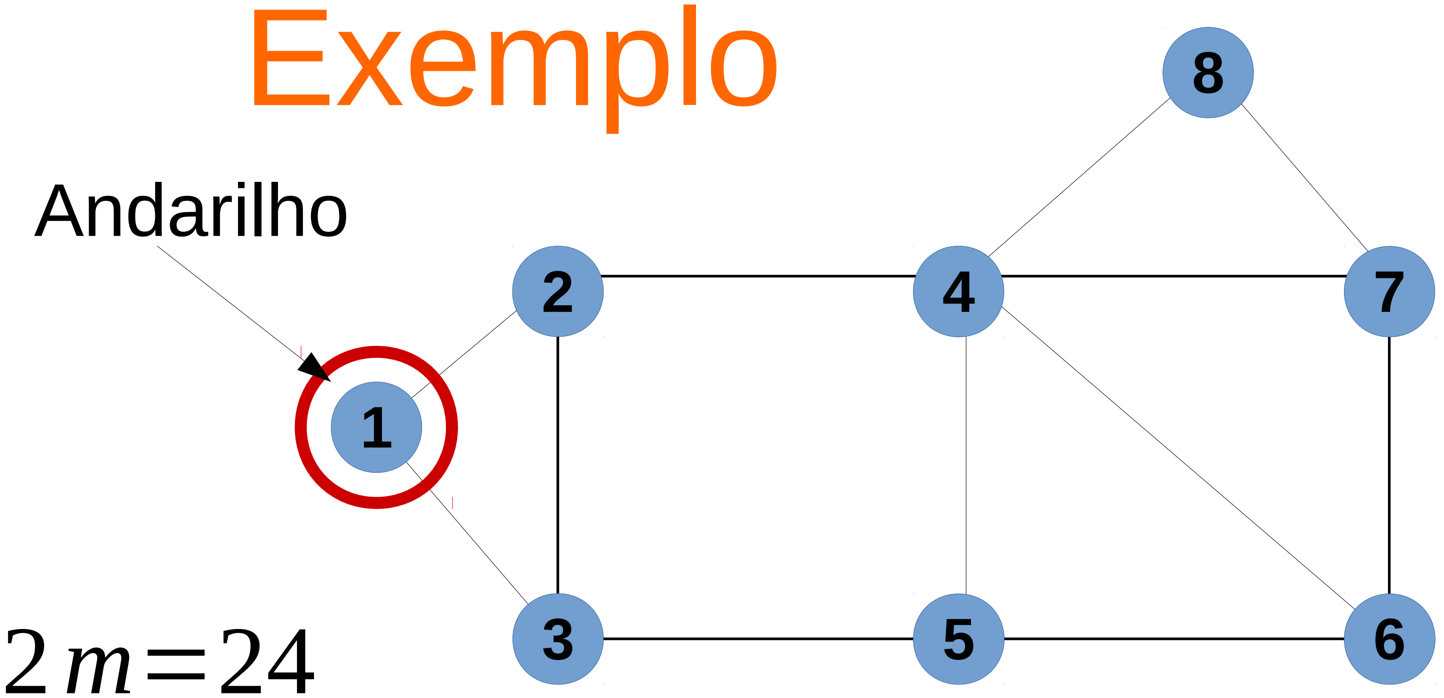
$$\pi_i P_{ij} = \pi_j P_{ji}$$

Podemos verificar

  - tanto faz andarilho andar para frente ou para trás
  - equilíbrio mais forte para esta CM

Figueiredo 2018

Exemplo

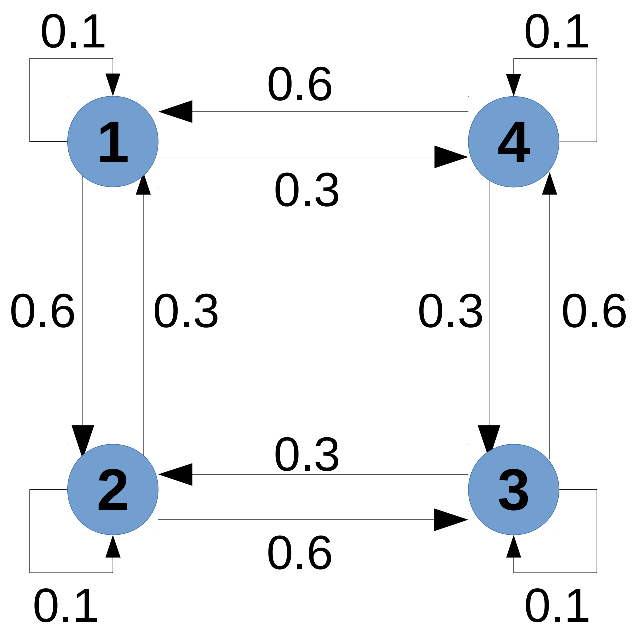


$$K = \sum_i d_i = 2m = 24$$

$$\pi = \left(\frac{2}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{5}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{2}{24} \right)$$

- Independente de onde andarilho inicia, distribuição de sua posição converge para π
  - prob de encontrar o andarilho no vértice *i*

Outro Exemplo



- CM é irredutível? **Sim!**
- CM é aperiódica? **Sim!** → self - loops
- Qual é sua distribuição estacionária?
  - por inspeção, sem fazer conta

  π_i = ¼ para i = 1, 2, 3, 4
- Podemos verificar facilmente!

- CM é reversível? **Não!**

$$\pi_1 p_{12} = \frac{1}{4}\frac{6}{10} = \frac{3}{20} \neq \pi_2 p_{21} = \frac{1}{4}\frac{3}{10} = \frac{3}{40}$$

- **Intuição:** processo gira mais em uma direção (anti-horário)
  - direção do tempo (para frente ou para trás) determina direção mais girada, não sendo reversível

Figueiredo 2018

Distribuição Estacionária: π é uma distribuição estacionária da MC se:

$$\boxed{\pi P = \pi} \quad \text{ou} \quad \boxed{\pi_i = \sum_j \pi_j \cdot P_{ji}}$$

Tempo de Chegada: T_{ij} é uma variável aleatória que indica o tempo mínimo para ir de $i \to j$.

$$\tau_{ij} = \mathbb{E}[T_{ij}]$$

Se a MC for irredutível e aperiódica:

$$\boxed{P[T_{ij} < \infty] = 1 \quad (\tau_{ij} < \infty)} \quad \text{e} \quad \boxed{\pi_i = 1/\tau_{ii}}$$

Reversibilidade: Não consigo diferenciar se $i \to j$ ou se $j \to i$

$$\boxed{\pi_j \cdot P_{ji} = \pi_i \cdot P_{ij}}$$

Convergência: Se a MC é irredutível e aperiódica, então a convergência para π é única e independe de $\pi(0)$