Instituto Tecnológico de Aeronáutica
Divisão de Engenharia Mecânica-Aeronáutica

MOQ-12
Cadeias de Markov

Professora:
Denise Beatriz T. P. do Areal Ferrari
denise@ita.br

Roteiro

- Introdução
- Processos Estocásticos
- Motivação
- Cadeias de Markov
  - Definição
  - Propriedade Markoviana
  - Propriedade de Estacionariedade
  - Matriz de Transição
  - Equação de Chapman-Kolmogorov
  - Vetor de Probabilidades Iniciais
  - Classificação de Estados

Introdução

- Hipótese de Independência

- Cadeias de Markov:
  - podem ser considerados como a generalização mais simples do esquema de experimentos independentes.

- Esses processos são o ponto inicial para o desenvolvimento de um novo e importante ramo da Teoria das Probabilidades:

  Teoria dos Processos Estocásticos.

Processos Estocásticos

- Funções Aleatórias
  - Intervalo de tempo: *série temporal*

    Exemplos:

    - flutuações de câmbio;
    - sinais (fala, áudio e vídeo);
    - dados médicos (eletrocardiograma, pressão sangüínea e temperatura);
    - movimentos aleatórios (Movimento Browniano e Passeio Aleatório)
  - Região do espaço: *campo aleatório*

    Exemplos:

    - imagens estáticas;
    - topografias aleatórias (satélite);
    - variações de composição em um material não homogêneo.

Motivação (1)

- Problema em estudo:

Estamos medindo a temperatura do ambiente externo a cada hora, durante um dia.

Temperatura $\Rightarrow$ fenômeno aleatório.

(resulta de uma conjuminação de fatores meteorológicos que não podem ser determinados com exatidão)

Motivação (2)

- Para medir esta temperatura, usamos um termômetro que mede desde 0°C até 50°C, com divisões de 10°C.
  - temos um conjunto finito de resultados ou estados possíveis: 0°C, 10°C, 20°C, 30°C, 40°C e 50°C.

Sejam:

X_1 = temperatura medida no início do período (primeira observação)

X_2 = temperatura medida na segunda observação (1 hora depois de X_1)

X_n = temperatura medida no n-ésimo período observado, n = 1,2,...

A seqüência X_1, X_2,..., X_{24} é um exemplo de Processo Estocástico.

Motivação (3)

A seqüência X_1, X_2,..., X_{24} é um exemplo de Processo Estocástico.

- Processo estocástico de ***parâmetro discreto****:* instantes pontuais de tempo (intervalos a cada hora)

X_1 = estado inicial do processo

X_n = estado do processo no instante n.

- Processo estocástico de ***parâmetro contínuo:*** Se estivéssemos monitorando a temperatura continuamente.

Motivação (4)

- *Espaço de estados* (E): $X_n \in E$

  - **Discreto**:

    $E = \{x \in \mathbb{Z} | x \geq 0°C\} = \{0°C, 10°C, 20°C, 30°C, ... \}$

    O termômetro utilizado pode medir apenas um número definido de temperaturas (variações de 10°C).

  - **Contínuo**:

    $E = \{x \in \mathbb{R} | x \geq 0°C\}$

    O termômetro utilizado pode medir qualquer temperatura.

Cadeias de Markov

- É um tipo especial de processo estocástico, que satisfaz as seguintes condições:

  - o parâmetro n é discreto (ex: tempo)
  - o espaço de estados **E** é discreto (coleção de estados possíveis)

    E pode ser finito ou infinito e enumerável. Vamos considerar **E** finito. As cadeias de Markov deste tipo são chamadas ***Cadeias de Markov Finitas***.
  - o estado inicial do processo ou o espaço de estados é conhecido.
  - vale a ***propriedade markoviana***
  - vale a ***propriedade de estacionariedade***

Propriedade Markoviana

Para n = 1,2,..., e qualquer seqüência de estados possíveis $s_1, s_2,\ldots,s_{n+1}$, com X_n, X_{n-1},..., X_1 conhecidos:

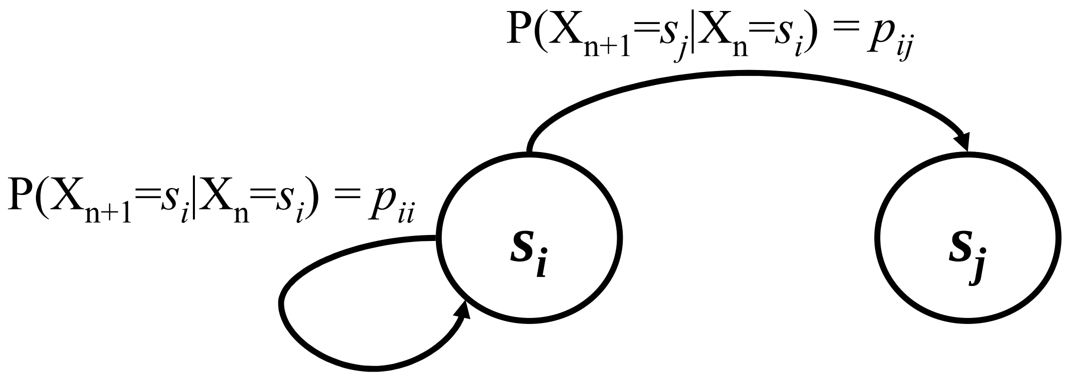
$$P(X_{n+1} = s_{n+1} \mid X_1 = s_1, X_2 = s_2, \ldots, X_n = s_n) =$$
$$= P(X_{n+1} = s_{n+1} \mid X_n = s_n)$$

Em palavras:

- As probabilidades de todos os estados futuros X_j (j>n) dependem somente do estado atual X_n, mas não dependem dos estados anteriores X_1,..., X_{n-1}.

O estado "futuro" do sistema depende do "presente", mas não depende do "passado".

Propriedade de Estacionariedade



- *Probabilidades de transição:* $P(X_{n+1}=s_j|X_n=s_i)$
- *Probabilidades de transição estacionárias:*

$$P(X_{n+1} = s_j \mid X_n = s_i) = cte. = p_{ij} \qquad n = 1,2,...$$

A cadeia de Markov é dita *homogênea* ou *estacionária*.

Matriz de Transição

- Cadeia de Markov Finita e Estacionária
- k possíveis estados: $s_1, s_2,\ldots,s_k$

$$P = \begin{array}{c c} & \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \end{array}$$

$$pij \geq 0$$

$$\sum_{j=1}^{k} pij = 1, \qquad i = 1,\ldots k$$

Exemplo

- Suponha que na Terra de Oz, nunca há dois dias ensolarados consecutivos.
  - Se em um dia qualquer faz sol, no dia seguinte pode tanto nevar quanto chover.
  - Se chover, metade das vezes continua chovendo no dia seguinte e nas outras ocasiões pode tanto fazer sol ou nevar.
  - Se nevar, apenas metade das vezes o dia seguinte também neva.

Queremos:

- Representar graficamente a Cadeia de Markov
- Construir sua matriz de transição
- Determinar a probabilidade de nevar daqui a dois dias?

Equação de Chapman-Kolmogorov

Sejam:

- P a matriz de transição de uma cadeia de Markov.
- O elemento $p_{ij}^{(n)}$ da matriz $P^{(n)}$ representa a probabilidade de que o processo, iniciado no estado s_i, estará no estado s_j, depois de n passos.
- u um instante qualquer entre 0 e n.

$$p_{ij}^{(n)} = \sum_{r=1}^{k} p_{ir}^{(u)} . p_{rj}^{(n-u)}, \quad 0 < u < n$$

ou, alternativamente:

$$p_{ij}^{(n)} = P(X_n = s_j \mid X_1 = s_i) = \sum_{r=1}^{k} p_{ir}^{(n-1)} p_{rj}, \quad n = 1,2,...$$

Vetor de Probabilidades Iniciais (*a priori*)

Em muitas situações, não conhecemos o estado da Cadeia de Markov no instante inicial.

- Cadeia de Markov Finita e Estacionária
- k possíveis estados: $s_1, s_2, \dots, s_k$

- Para i=1,...,k:

v_i = probabilidade de que o processo esteja no estado s_i no instante inicial

$$v_i \geq 0 \qquad \text{e} \qquad \sum_{i=1}^{k} v_i = 1$$

- A qualquer vetor **w** = $(w_1,..,w_k)$, tal que $w_i \geq 0$ e $w_1+w_2+\dots+w_k=1$ chamamos

vetor de probabilidades.

- O vetor **v** =$(v_1,\dots, v_k)$ é chamado ***vetor de probabilidades iniciais***, pois representa as probabilidades dos vários estados da cadeia no instante de início do processo.

Vetor de Probabilidades Iniciais (*a priori*)

- Teorema:
  - Seja P a matriz de transição de uma cadeia de Markov
  - **v** o vetor de probabilidades iniciais.

  Então, a probabilidade de que o processo esteja no estado s_j depois de n passos é a j-ésima componente do vetor:

$$\mathbf{v}^{(n)} = \mathbf{v}.P^n$$

onde:

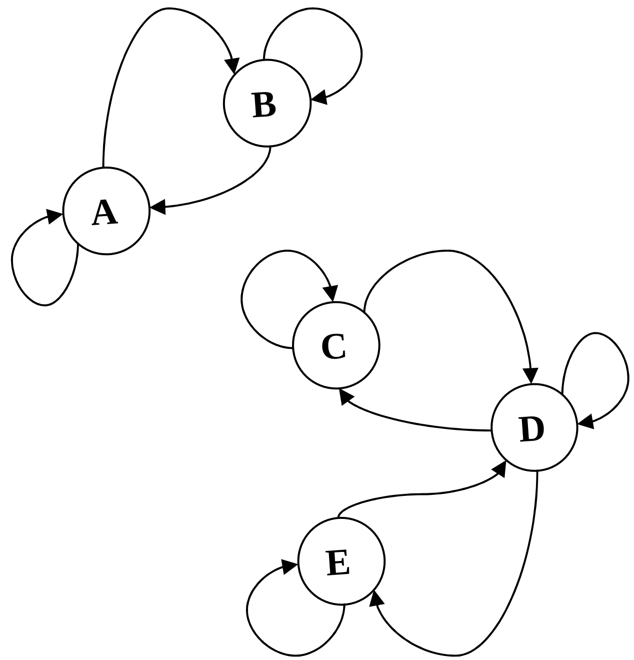
$\mathbf{v}^{(n)} = [v_1^{(n)}\ v_2^{(n)} \ldots v_k^{(n)}]$

$v_i^{(n)} = P(Xn = s_i)$

Classificação de Estados

- A fim de ilustrar as próximas definições, vamos utilizar a Cadeia de Markov representada a seguir:

$$P = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0.4 & 0.6 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 & 0 \\ 0 & 0 & 0.5 & 0.4 & 0.1 \\ 0 & 0 & 0 & 0.8 & 0.2 \end{bmatrix} \end{array}$$



- Definição: ***Caminho***

  Dados dois estados s_i e s_j, um *caminho* entre s_i e s_j é uma seqüência de transições que começam em s_i e terminam em s_j, tais que cada transição tem uma probabilidade positiva de ocorrência.

1. Um estado s_j é ***acessível*** a partir do estado s_i se existe um caminho que liga s_i a s_j.

2. Dois estados s_i e s_j são ***comunicáveis*** se s_j é acessível a partir de s_i e s_i é acessível a partir de s_j.

3. Um conjunto de estados **E** em uma cadeia de Markov é uma ***classe*** se nenhum estado fora de **E** é acessível por qualquer estado em **E**.

   Se a cadeia inteira é formada por uma única classe, isto é, todos os estados são comunicáveis, a cadeia é dita ***irredutível.***

4. Um estado si é absorvente se pii = 1.

5. Um estado si é transiente se existe um estado sj que é acessível a partir de si, mas o estado si não é acessível a partir de sj.

6. Se um estado não é transiente ele é recorrente.

7. Um estado s_i é ***periódico*** com período T>1 se T é o menor número tal que todos os caminhos que levam do estado s_i de volta a s_i tem comprimento múltiplo de T. Se um estado recorrente não é periódico ele é ***aperiódico***.

8. Se todos os estados em uma Cadeia de Markov são recorrentes, aperiódicos e comunicáveis entre si, então a cadeia é dita ***ergódica***.

Exemplo

- Suponha que só existem dois refrigerantes: guaraná e soda.
  - Se uma pessoa escolheu guaraná, existe 90% de chance de peça novamente guaraná.
  - Se a pessoa tiver escolhido soda, a chance de que peça este refrigerante outra vez é de 80%.

1. Se uma pessoa é atualmente consumidora de soda, qual a probabilidade de que escolha guaraná no segundo pedido futuro?

2. Se a pessoa é atualmente consumidora de guaraná, qual é a probabilidade de que escolha guaraná no terceiro pedido futuro?

Solução

- Os pedidos de cada consumidor podem ser interpretados como uma Cadeia de Markov de dois estados, descritos como:

Estado 1 = a pessoa escolheu guaraná da última vez (G)

Estado 2 = a pessoa escolheu soda da última vez (S)

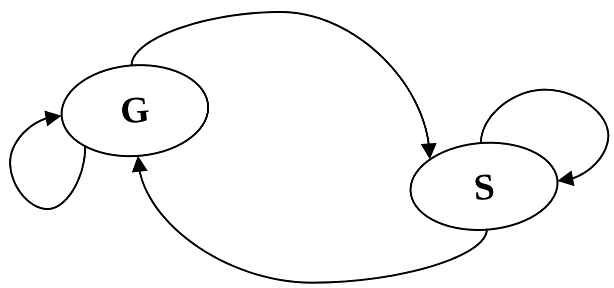
Vamos definir:

X_0= refrigerante escolhido no presente

X_n= refrigerante escolhido no n-ésimo pedido futuro

A seqüencia X_0, X_1,... pode ser descrita como a seguinte Cadeia de Markov:

$$P = \begin{array}{c c} & \begin{array}{cc} G & S \end{array} \\ \begin{array}{c} G \\ S \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \end{array}$$



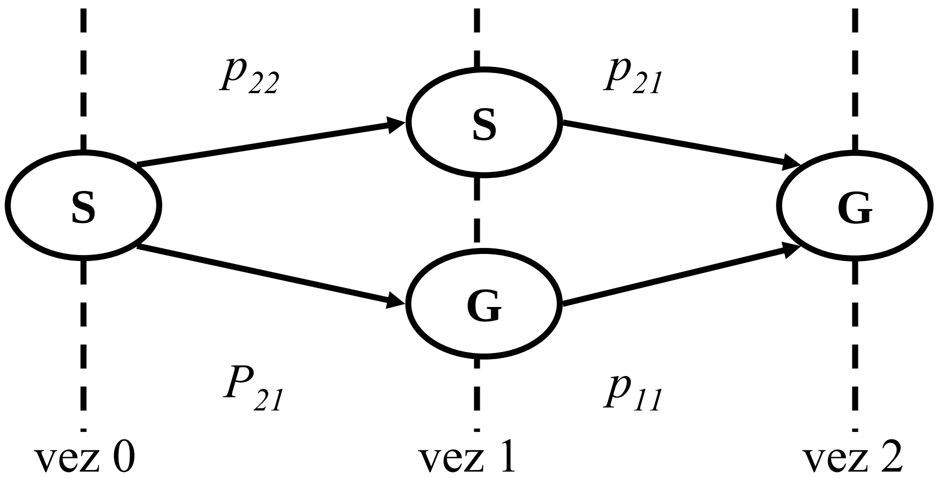
Agora sim, podemos responder às perguntas...

1. Se uma pessoa é atualmente consumidora de soda, qual a probabilidade de que escolha guaraná no segundo pedido futuro?

Queremos: $$P(X_2 = G \mid X_0 = S) = p_{21}^{(2)}$$

$$P^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ \boxed{0.34} & 0.66 \end{bmatrix}$$

De outra maneira:



$p_{21}^{(2)} =$

(probabilidade de que o próximo pedido seja soda e o segundo pedido seja soda)

+

(probabilidade de que o próximo pedido seja guaraná e o segundo pedido seja soda)

$= p_{21}p_{11} + p_{22}p_{21}$
$=(0.2)(0.9)+(0.8)(0.2)=0.34$

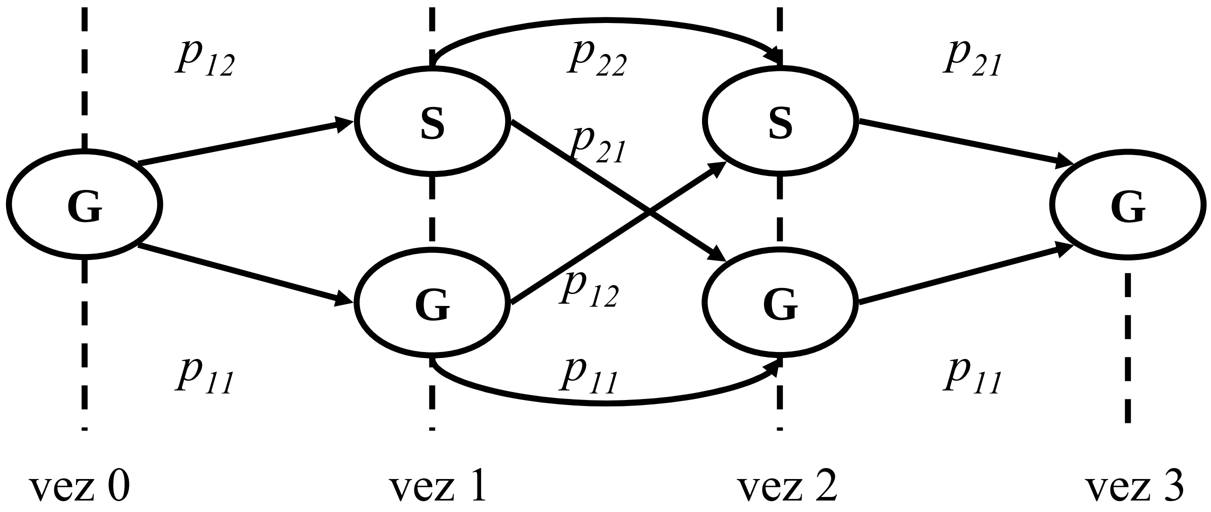
2. Se a pessoa é atualmente consumidora de guaraná, qual é a probabilidade de que escolha guaraná no terceiro pedido futuro?

Queremos: $P(X_3 = G \mid X_0 = G) = p_{11}^{(3)}$

$$P^3 = P.P^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} \boxed{0.781} & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

2. Se a pessoa é atualmente consumidora de guaraná, qual é a probabilidade de que escolha guaraná no terceiro pedido futuro?

De outra maneira:



$$p_{11}^{(3)} = p_{12}p_{22}p_{21} + p_{12}p_{21}p_{11} + p_{11}p_{12}p_{21} + p_{11}p_{11}p_{11} =$$

$$=(0.1)(0.8)(0.2)+(0.1)(0.2)(0.9)+(0.9)(0.1)(0.2)+(0.9)(0.9)(0.9) = 0.781$$

Suponha, que 60% das pessoas bebem guaraná e 40% bebem soda agora. Daqui a três pedidos, que fração das pessoas beberá guaraná?

Queremos: $\mathbf{v}^{(3)} = \mathbf{v}.\ P^3$

Temos: $\mathbf{v} = [0.6 \;\; 0.4]$

Portanto:

$$\mathbf{v}^{(3)} = \mathbf{v}.P^3 = [0.6 \quad 0.4]\begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix} = [\boxed{0.6438} \quad 0.3562]$$