Cadeias de Markov

2023-2024

Problema Exemplo 1

F: "falta à aula"

- Suponhamos que em cada dia que têm aulas de MPEI acordam e decidem se vêm ou não à aula
- Se vieram à aula anterior, a probabilidade de virem é 70% $= P(\bar{F} \mid \bar{F})$
- se faltaram à anterior, essa probabilidade é 80% $= P(\bar{F} \mid F)$

- Algumas questões:
  - Se vieram à aula esta segunda, qual a probabilidade de virem na aula de SEGUNDA da próxima semana ?
  - Assumindo que o semestre tem duração infinita (que horror!), qual a percentagem aproximada de aulas a que estariam presentes ?

Problema Exemplo 2

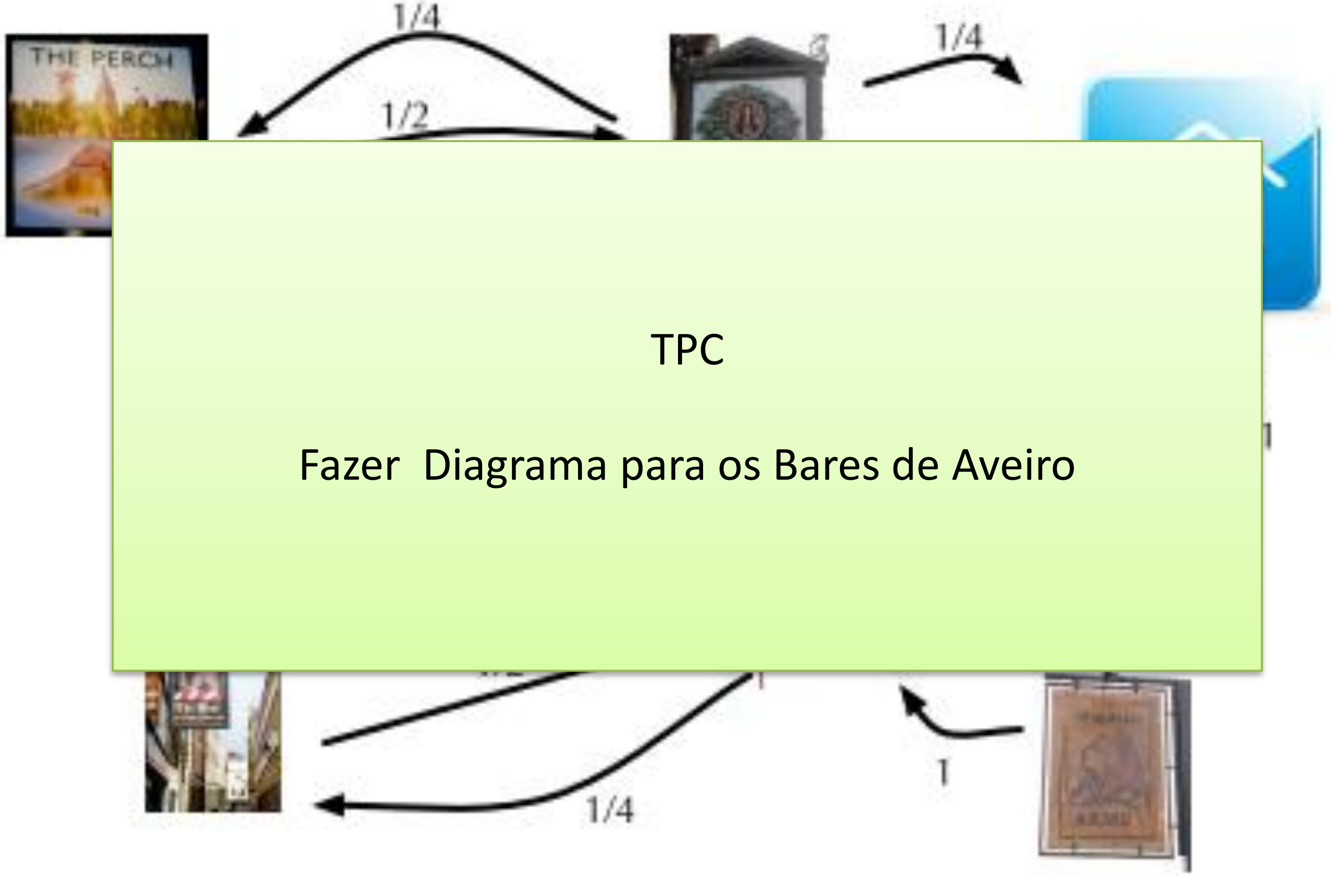
- Dividir a turma em 3 grupos A, B e C no início do semestre
- No final de cada aula:
  - 1/3 do grupo A vai para o B e outro 1/3 do grupo A vai para o grupo C
  - ¼ do grupo B vai para A e ¼ de B vai para C
  - ½ do grupo C vai para o grupo B

Podia ser a migração entre países

- Como ficarão os grupos ao fim de *n* aulas ?

Exemplo 3 – “Pub Crawl”

- Bares junto a uma conhecida Universidade:



TPC

Fazer Diagrama para os Bares de Aveiro

Muitas áreas de aplicação

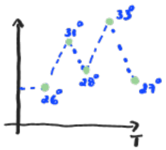
- Muitas vezes estamos interessados na transição de algo entre certos estados

- Exemplos:
  - Movimento de pessoas entre regiões
  - Estado do tempo
  - Movimento entre as posições num jogo de Monopólio
  - Pontuação ao longo de um jogo
  - Estado de Filas de atendimento

Em vez de $(t_1, t_2, t_3) \longrightarrow (x_1, x_2, x_3)$

$(T(t_1), T(t_2), T(t_3))$

Processos estocásticos



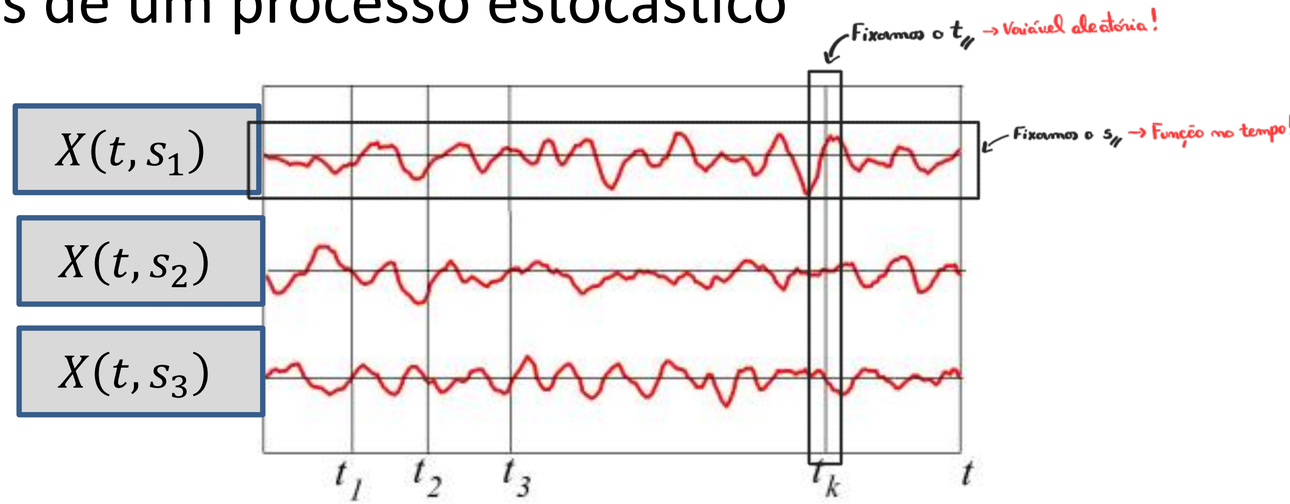
- Estendem o conceito de variável aleatória

- O processo estocástico mapeia o evento para números diferentes em tempos diferentes

- O que implica que em lugar de termos um número $X(s)$ temos $X(t, s)$
  - Sendo t $\in T$ geralmente um conjunto de tempos

Processos estocásticos

- 3 realizações de um processo estocástico



- Se fixamos s, $X(t)$ é uma função real do tempo
  - $X(t, s)$ pode então ser vista como uma colecção de funções no tempo
- Se fixarmos t temos uma função $X(s)$ que depende apenas de s, ou seja uma variável aleatória
- Um nome alternativo é processos aleatórios

Classificação de processos estocásticos

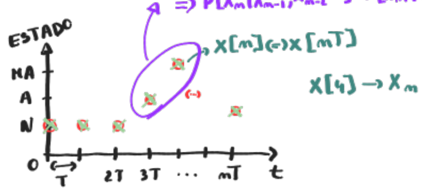
- Podem ser classificados segundo t e os valores que pode assumir (estados do processo)

- Quanto ao tempo
  - Tempo contínuo:
    - Se tempo é um intervalo contínuo → todos os valores
  - Tempo discreto: ←
    - Se o tempo é um conjunto contável
    - Também chamada sequência aleatória e representada por $X[n]$
- Quanto ao conjunto de estados (E):
  - Contínuo
  - Discreto ← Δt

Caso discreto

Nota: só existe uma dependência com o acontecimento anterior

$\Rightarrow P[X_n | X_{n-1}, X_{n-2} \cdots] = P[X_n | X_{n-1}]$



Mas, sabemos que:

$t \in \mathbb{R}$ ou $t \begin{cases} \text{M. Alta} \\ \text{Alta} \\ \text{Normal} \\ \varnothing \end{cases}$ (estados)

Definição

- Um processo de Markov (de 1ª ordem) é um processo estocástico em que a probabilidade de o sistema estar num estado específico num determinado período de observação depende apenas do seu estado no período de observação imediatamente precedente
  - O futuro apenas depende do presente e não do passado

Tipos de processos de Markov

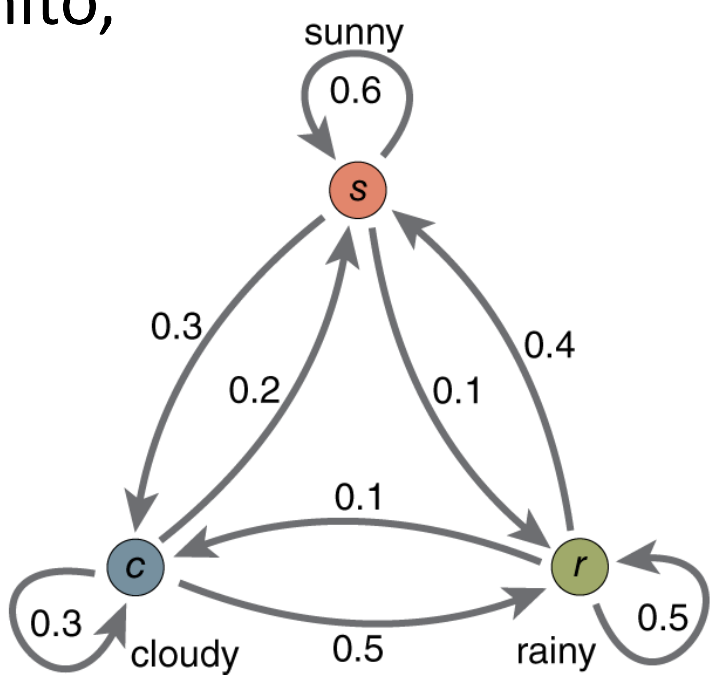
- Discretas/contínuas

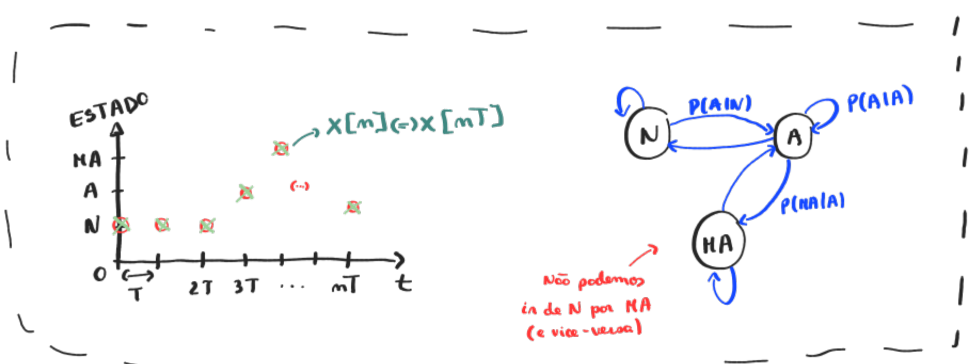
		Espaço de estados	
		Discreto	Contínuo
Tempo	Discreto	**Cadeia de Markov tempo discreto**	Processo de Markov em tempo discreto
	Contínuo	Cadeia de Markov tempo contínuo	Processo de Markov em tempo contínuo

- Focaremos a nossa atenção em cadeias de Markov de tempo discreto

Cadeias de Markov discretas

- X_n: estado após *n* transições
  - Pertence a um conjunto finito,
    - Em geral $\{1, 2, \dots, m\}$
  - X_0 é dado ou aleatório





Questões comuns relativas a cadeias de Markov

- Qual a probabilidade de transição entre dois estados em *n* observações ?

- Existe algum equilíbrio ?

- Existe uma estabilidade a longo prazo ?

Propriedade de Markov

- Probabilidade de transição do estado i para o estado j:

- $\boldsymbol{p_{ji}} = P(X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0)$

$$= \boldsymbol{P(X_{n+1} = j | X_n = i)}$$

$$\boxed{p_{ji} = P(X_{m+1} = j \mid X_m = i)} \Rightarrow \boxed{p_{ji} = p_{j|i}}$$

- Quando estas probabilidades $\boldsymbol{p_{ji}}$ não dependem de n a cadeia diz-se **homogénea**
  - Focaremos a nossa atenção neste tipo de cadeias de Markov

Especificação de uma cadeia

- Identificar os estados possíveis
- Identificar as transições possíveis
- Identificar as probabilidades de transição

Aplicando ao exemplo 1 – faltar ou não faltar à aula

<u>Sabemos que</u>:

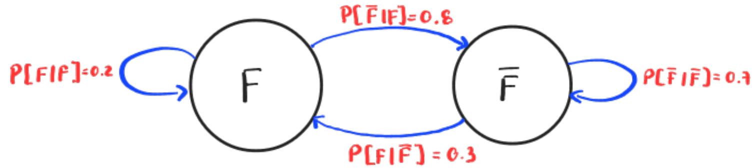
$P[\bar{F}|\bar{F}]=0.7$

$P[\bar{F}|F]=0.8$

$\Rightarrow$ $P[F|\bar{F}]=0.3$

$P[F|F]=0.2$

- Estados ?



- Transições ?

- Probabilidades de transição ?

<u>Matriz de transição</u>:

i - "sabendo que"

$$T_{ji} = \begin{array}{c|cc} & \bar{F} & F \\ \hline \bar{F} & 0.7 & \boxed{0.8} \\ F & 0.3 & 0.2 \end{array}$$

$\rightarrow P[\bar{F}|F]$

j - "próximo"

Primeiro a linha depois a coluna

Aplicando ao exemplo 1 – faltar ou não faltar à aula

- Estados ?
  - 2: {faltar, não faltar}

- Transições ?
  - Faltar-> não faltar
  - Não faltar -> faltar
  - Faltar -> faltar
  - Não faltar -> não faltar

- Probabilidades de transição ?
  - Faltar-> não faltar   : 0,8
  - Não faltar -> faltar   : 0,3
  - Faltar -> faltar   : 0,2
  - Não faltar -> não faltar: 0,7

Matriz de transição

- É usual representar as probabilidades de transição através de uma matriz, **chamada de matriz de transição**
- Tendo o sistema n estados possíveis, para cada par i, j fazemos t_{ji} igual à probabilidade de mudar **do estado *i* para o estado *j***.
- A matriz T cujo valor na posição $\text{linha} = j, coluna = i$ é t_{ji} é a matriz de transição

$$\begin{array}{cc} & \xrightarrow{\quad \text{Xn} \quad} \\ \text{Xn+1} \Big\downarrow & T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix} \end{array}$$

  - Nota: Alguns autores adoptam t_{ij} como a probabilidade de mudar do estado i para o estado j

Matriz T do Exemplo 1

- $T = \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix}$

- Considerando estado 1 “não faltar”, temos

- $T = \begin{array}{l} não\ faltar \rightarrow \\ faltar \rightarrow \end{array} \begin{pmatrix} 0{,}7 & 0{,}8 \\ 0{,}3 & 0{,}2 \end{pmatrix}$

Matriz T do Exemplo 2

$$T = \begin{array}{cccc} & A & B & C \\ A & ? & ? & ? \\ B & ? & ? & ? \\ C & ? & ? & ? \end{array}$$

Matriz T do Exemplo 2

- No final de cada aula:
  - 1/3 do grupo A vai para o B e outro 1/3 do grupo A vai para o grupo C
  - ¼ do grupo B vai para A e ¼ de B vai para C
  - ½ do grupo C vai para o grupo B

← "sabendo que" - i

$$T = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{ccc} A & B & C \\ \hline \frac{1}{3} & \boxed{\frac{1}{4}} & 0 \\ \boxed{\frac{1}{3}} & \frac{1}{2} & \boxed{\frac{1}{2}} \\ \boxed{\frac{1}{3}} & \boxed{\frac{1}{4}} & \frac{1}{2} \end{array}$$

futuro - j

Futuro Estado

Matriz T do Exemplo 2

$$T = \begin{array}{c|ccc|} & A & B & C \\ A & 1/3 & 1/4 & 0 \\ B & 1/3 & 1/2 & 1/2 \\ C & 1/3 & 1/4 & 1/2 \end{array}$$

Futuro Estado

Matriz T é estocástica

- A matriz de transição reflecte propriedades importantes das probabilidades:

  - Todas as entradas são não-negativas
  - Os valores em cada COLUNA somados dão sempre resultado 1

- Devido a estas propriedades a matriz é denominada de **matriz estocástica**

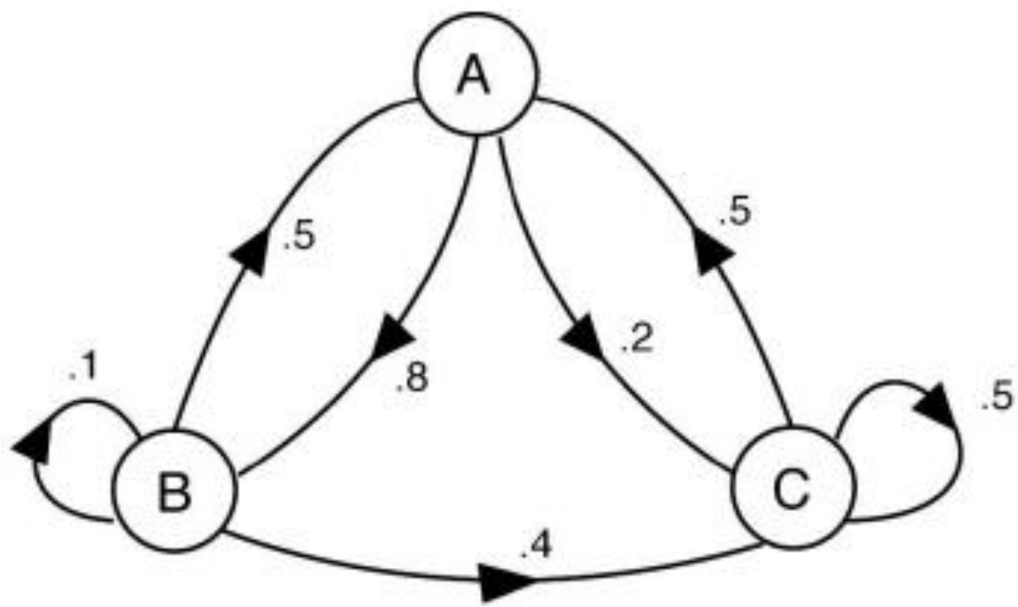
Representação gráfica da cadeia

- Apropriada e possível para número de estados pequeno

- Nós: representam todos os estados
- Setas: para todas as transições permitidas (one-step)
  - $\boxed{\text{Ou seja, seta entre i e j apenas de } p_{ji} > 0}$

Representação gráfica da cadeia

- Exemplo:



Simulação / Visualização dinâmica

- Estão disponíveis online formas de visualizar as transições entre estados ao longo do tempo ...
- Um desses exemplos é **Markov Chains - A visual explanation by [Victor Powell](http://setosa.io)**
  - http://setosa.io/blog/2014/07/26/markov-chains/index.html

  Que inclui:
  - http://setosa.io/markov/index.html#%7B%22tm%22%3A%5B%5B0.5%2C0.5%5D%2C%5B0.5%2C0.5%5D%5D%7D

- Para usar precisamos apenas de introduzir a matriz T
  - Que define o número de estados, quais as transições possíveis e as probabilidades associadas a essas transições

Estado da cadeia num determinado instante

- O **estado** de uma cadeia de Markov com n estados no tempo (time step) k é dado pelo vector estado

$$\mathbf{x}^{(k)} = \begin{pmatrix} \boldsymbol{p_1}^{(\boldsymbol{k})} \\ \boldsymbol{p_2}^{(\boldsymbol{k})} \\ \\ \boldsymbol{p_n}^{(\boldsymbol{k})} \end{pmatrix}$$

- Onde ${p_j}^{(k)}$ é a probabilidade de o sistema estar no estado j no instante de tempo k

Vector estado/probabilidade

- Considerando o exemplo 1:
  - Suponhamos que após 10 aulas a probabilidade de faltar e não faltar são iguais
  - Então o vector representativo do estado (*state vector*) seria:

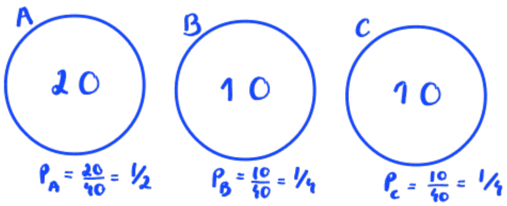
$$\mathbf{x}^{(10)} = \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix}$$

- Também se designa por **vector de probabilidade**
  - Todos elementos não-negativos
  - Soma dos elementos igual a um

Exemplo 2

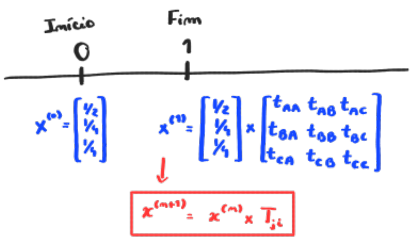
- Supondo que começávamos com:
  - 20 estudantes no grupo A e
  - 10 estudantes nos outros dois grupos

- O vector relativo ao estado inicial seria:



$$\mathbf{x}^{(0)} = \begin{pmatrix} 0{,}5 \\ 0{,}25 \\ 0{,}25 \end{pmatrix}$$

Probabilidade inicial



Vector estado após uma transição

- Como obter $\mathbf{x}^{(k+1)}$?

- O vector de estado $\mathbf{x}^{(k+1)}$ no período de observação $k+1$ pode ser determinado a partir do vector $\mathbf{x}^{(k)}$ através de:

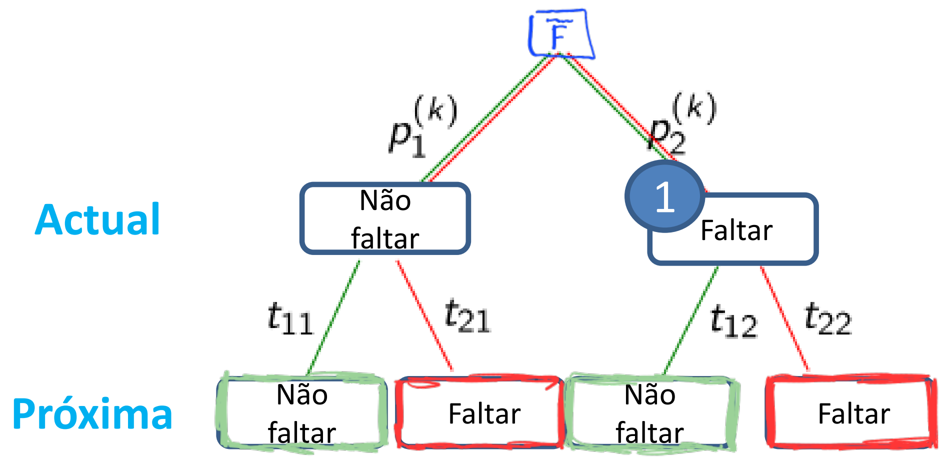
$$\boxed{\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)}}$$

- Que resulta da probabilidade condicional:
- $P(\;estado\; j\; em\; t = k+1)$

$$= \sum_{\mathrm{i}=1}^{n} P(transição\; do\; estado\; i\; para\; o\; j) P(estado\; \mathrm{i}\; em\; t = k)$$

Exemplo de aplicação – Exemplo 1

- De que forma depende a probabilidade de ir à aula seguinte da probabilidade de estar na aula actual ?



$$p_1^{(k+1)} = t_{11}p_1^{(k)} + t_{12}p_2^{(k)}$$
$$p_2^{(k+1)} = t_{21}p_1^{(k)} + t_{22}p_2^{(k)}$$

Multiplicação dos transições pelos probabilidades iniciais

Estado após múltiplas transições

- Ataquemos agora problemas do “tipo”:
  - Qual a probabilidade de transição entre dois estados em n observações/transições?

→ “não é no dia seguinte, mas $\underline{n}$ dias a seguir”

$$x^{(k+1)} = T\, x^{(k)}$$

$$x^{(k+2)} = T\, x^{(k+1)} = T(T\, x^{k}) = T^2 x^{k}$$

$$\boxed{x^{k+n} = T^n x^{k}}$$

- Exemplo 1:
  - Qual a probabilidade dos que estiveram na aula de uma segunda virem à aula na segunda seguinte
    - Assumindo as probabilidades do nosso exemplo !
    - Tendo em conta que temos aulas segunda e quarta (TP1).

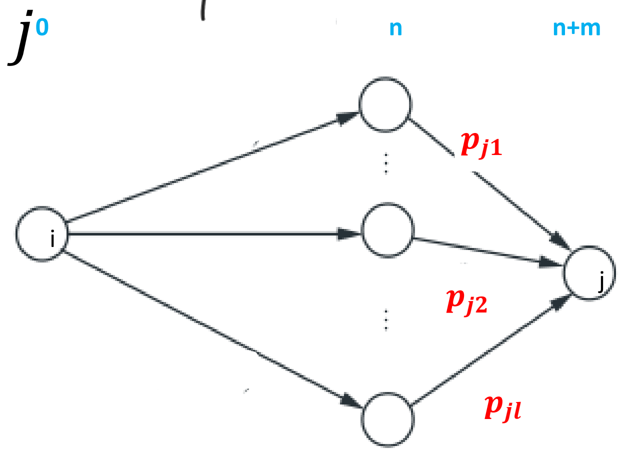
Probabilidade em n transições

- Definindo a transição em n passos ${p_{ji}}^n$ como a probabilidade de **um processo no estado $\boldsymbol{i}$ se encontrar no estado *j* após *n* transições** adicionais. Ou seja:
- ${p_{ji}}^n = P(X_{n+k} = \boldsymbol{j} | X_k = \boldsymbol{i}),\ n \geq 0, i, j \geq 0$
- Obviamente ${p_{ji}}^1 = p_{ji}$
- E para n > 1 ?

$p_{ji}^{\,n}$

- É fácil de compreender se tivermos em conta que $\boxed{p_{ki}^{\,n} p_{jk}^{\,m}}$ representa a probabilidade de:
  - O processo ir do estado i para o estado j em $n + m$ transições..
  - Através de um caminho que o leva ao estado k na transição n

Aveiro → Porto
(em n dias)



- Logo, somando para todos os estados intermédios k obtém-se a probabilidade de estar no estado j ao fim de $n + m$ transições

Equações de Chapman-Kolmogorov

- As **equações de Chapman-Kolmogorov** permitem calcular estas probabilidades

$$p_{ji}{}^{n+m} = \sum_k p_{ki}{}^{n}\ p_{jk}{}^{m} \quad \forall n, m \geq 0, \forall i, j$$

↳ posição intermédia (n passos)

Em termos de matrizes

- Se usarmos $\mathbf{T}^{(\mathbf{n})}$ para representar a matriz com as probabilidades de n transições, a equação anterior transforma-se em:

$$\boxed{\mathbf{T}^{(n+m)} = \mathbf{T}^{(n)} \, . \, \mathbf{T}^{(m)}}$$

Em que o "." significa multiplicação de matrizes

- Desta equação obtém-se facilmente:

$$\mathbf{T}^{(2)} = \mathbf{T}^{(1+1)} = \mathbf{T} \, . \, \mathbf{T} = \mathbf{T}^{2}$$

// Provado...

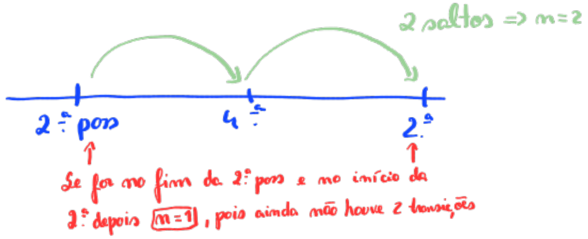
O difícil no enunciado é contar o valor de n //

- E por indução $\mathbf{T}^{(n)} = \mathbf{T}^{(n-1+1)} = \mathbf{T}^{n-1} \, . \, \mathbf{T} = \mathbf{T}^{n}$
  - Ou seja, a matriz de transição relativa a n transições pode ser obtida multiplicando $\mathbf{T}$ por si própria n vezes

Aplicação ao Exemplo 1

- Voltando a uma questão colocada no início da aula …

- *Se vieram à aula esta segunda, qual a probabilidade de virem na aula de SEGUNDA da próxima semana ?*



- Solução:
- Temos $\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, significando "não faltar"
- Pretendemos $\mathbf{x}^{(2)}$, 0= "esta segunda"

0.49+0.24=0.73
0.21+0.06=0.27

$$\boxed{X^{(k+m)} = x^{(k)} T^{m}} \Rightarrow x^{(2)} = x^{(0)} T^{2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{array}{c|cc} & \bar{F} & F \\ \hline \bar{F} & 0.7 & 0.8 \\ F & 0.3 & 0.2 \end{array}^{2} = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \begin{bmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.73 \\ 0.27 \end{bmatrix}$$

← Temos de saber fazer! //

...

- $$\boxed{\mathbf{x}^{(2)} = \mathrm{T}\mathbf{x}^{(1)} = T\left(T\mathbf{x}^{(0)}\right) = T^{2}\mathbf{x}^{(0)}}$$

$$= \begin{pmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{pmatrix}^{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} = \begin{array}{c} \bar{F} \\ F \end{array} \begin{pmatrix} \boxed{0.73} \\ 0.27 \end{pmatrix}$$

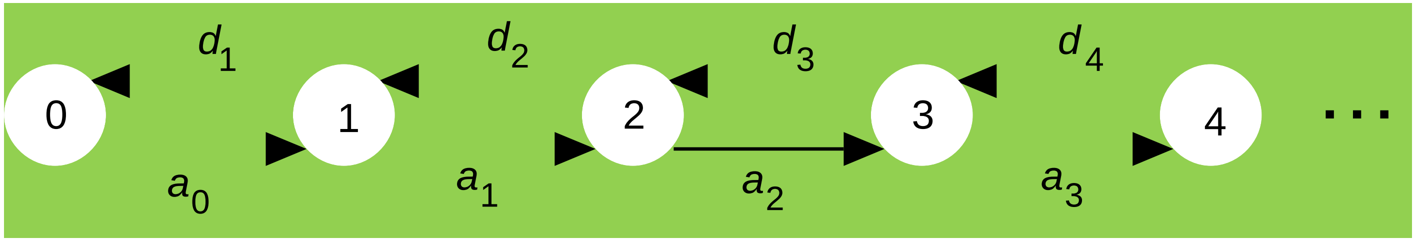
- Ou seja probabilidade igual a $\boxed{0.73}$ de virem na próxima Segunda

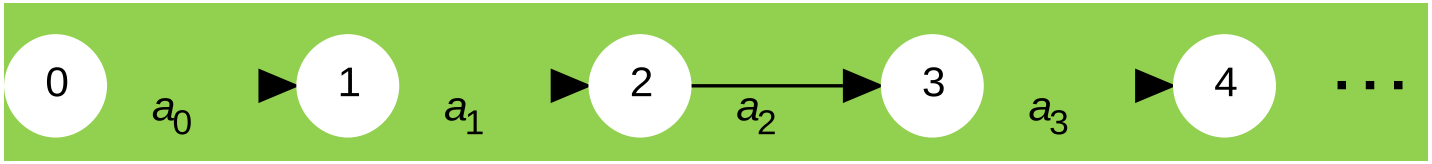
Terminologia

Tipos de estados

Acessibilidade de um estado

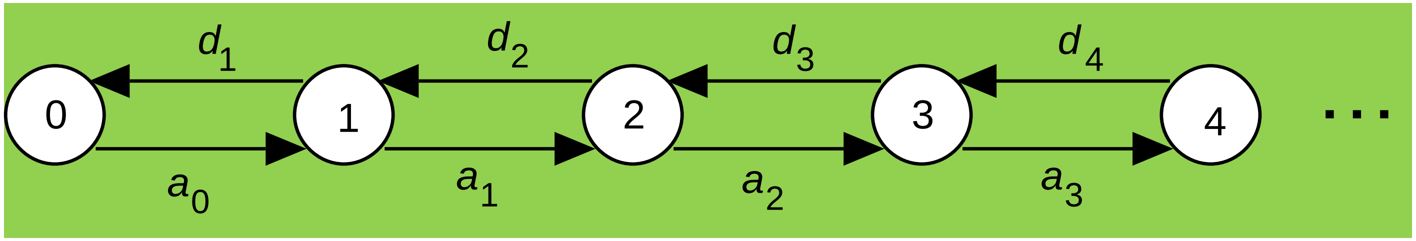
- Possibilidade de ir do estado i para o estado j (existe caminho na cadeia de i para j).





Estados comunicantes

- Dois estados comunicam se ambos são acessíveis a partir do outro.



- Um sistema é não redutível (irreducible) se todos os estados comunicam

- **Classe**: conjunto de estados que comunicam entre si

Estado recorrente

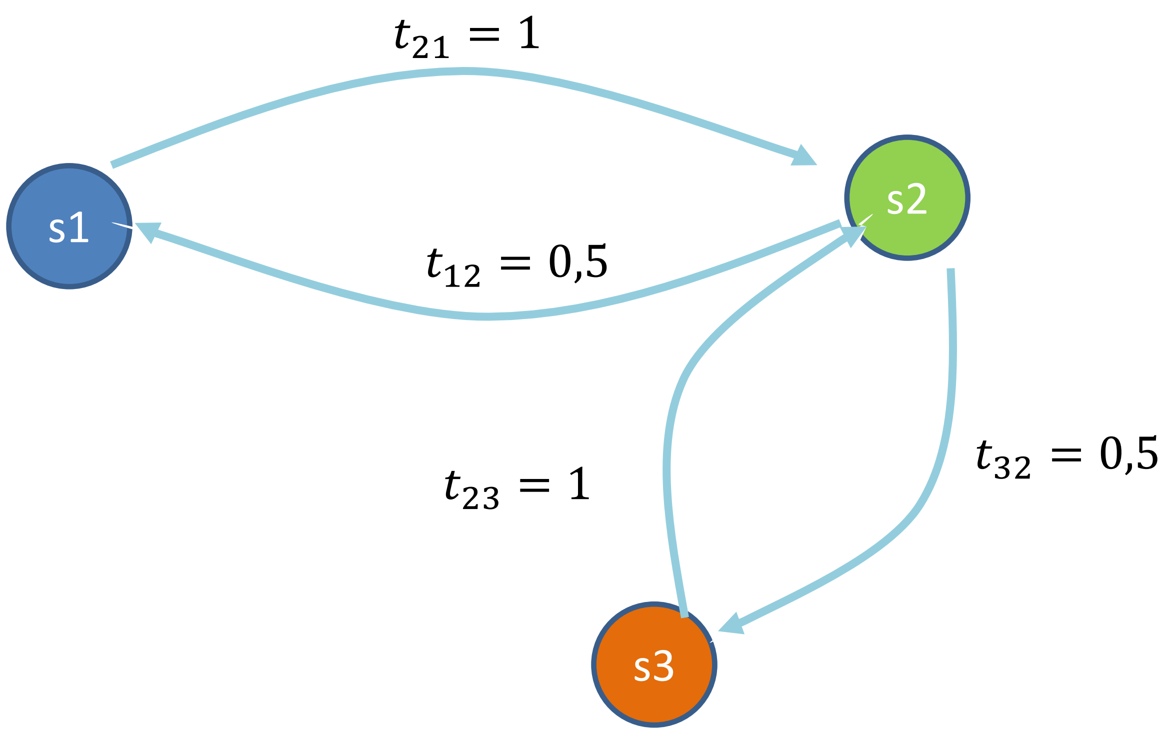
- Um estado s_i é um **estado recorrente** se o sistema puder sempre voltar a ele
  – depois de sair dele

- De uma forma mais formal:

  s_i é um **estado recorrente** se, para todos os estados s_j, a existência de um inteiro r_j tal que $p_{ji}^{(r_j)} > 0$ implica que existe um inteiro r_i tal que $p_{ij}^{(r_i)} > 0$

- Um estado não recorrente é transiente

Estados recorrentes ?



- Os 3 estados são recorrentes

Estado transiente

- Um estado é transiente se existe um outro estado qualquer para o qual o processo de Markov pode transitar, mas do qual o processo não pode retornar

- Ou seja, se existe um estado s_j e um inteiro l tal que ${p_{ji}}^{(l)} \neq 0$ e ${p_{ij}}^{(r)} = 0$ para $r = 0,1,2,\ldots$

- A probabilidade destes estados tende para zero quando *n* tende para infinito
  – Pois apenas são visitados um número finito de vezes

Estado periódico

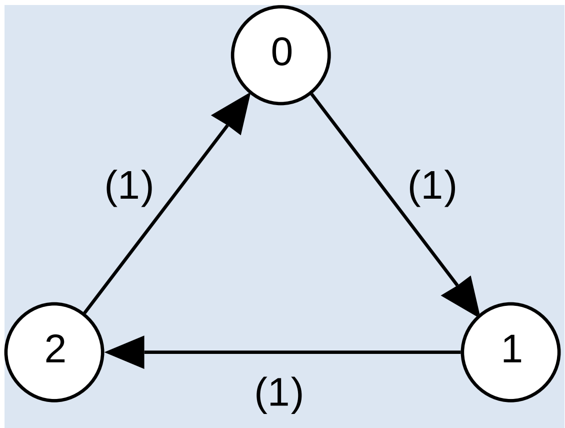
- Um estado é **periódico** se apenas se pode regressar a ele após um número fixo de transições superior a 1 (ou múltiplos desse número)

- Formalizando:
  - Um estado recorrente s_i diz-se **periódico** se existe um inteiro $c > 1$ tal que ${p_{ii}}^{(r)}$ é igual a zero para todos os valores de r excepto $r = c, 2c, 3c, \dots$

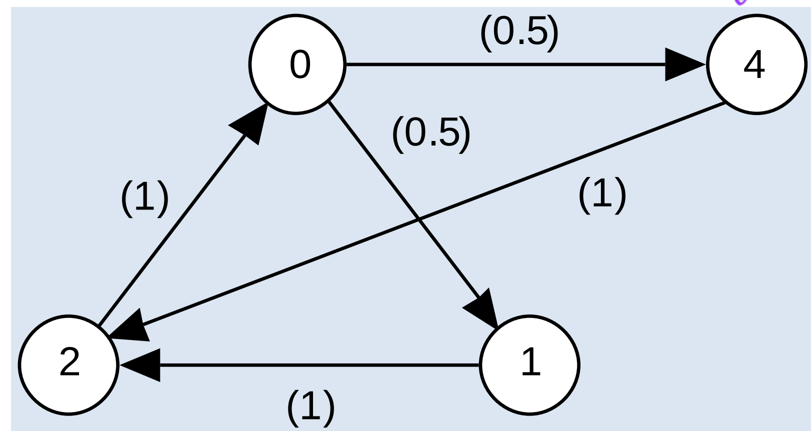
Estado periódico

•

ex: Multiplos



Todos estados visitados em múltiplos de 3 iterações



Todos estados visitados em múltiplos de 3 iterações

- Um estado não periódico é aperiódico
  - Como era de esperar!

Assuntos principais até ao momento

- Noção de processo estocástico
- Cadeias de Markov ← Processos estocásticos (discretos no tempo e no espaço) que preservem a <u>Propriedade de Markov</u>
- Propriedade de Markov ← (K + 1) só depende de (K)
- Matriz de transição T
- Representação gráfica → $\boxed{1}$ Fazer diagrama
  $\boxed{2}$ Fazer matriz T (identificar os estados! e as linhas e colunas)
- $\mathbf{T}^{(n)}$

Demos

- **Wolfram:**
  - **Finite-State, Discrete-Time Markov Chains**
    - **http://demonstrations.wolfram.com/ATwoStateDiscreteTimeMarkovChain/**
    - **http://demonstrations.wolfram.com/FiniteStateDiscreteTimeMarkovChains/**

O que acontece ao fim de muitas transições ?

Potências de T quando $n \to \infty$

- Exemplo 2 (3 grupos de alunos):
- Vejamos o comportamento de T^n ao aumentar *n*…

$$T = \begin{pmatrix} 0.333333 & 0.25 & 0. \\ 0.333333 & 0.5 & 0.5 \\ 0.333333 & 0.25 & 0.5 \end{pmatrix}$$

$$T^2 = \begin{pmatrix} 0.194444 & 0.208333 & 0.125 \\ 0.444444 & 0.458333 & 0.5 \\ 0.361111 & 0.333333 & 0.375 \end{pmatrix}$$

$$T^3 = \begin{pmatrix} 0.175926 & 0.184028 & 0.166667 \\ 0.467593 & 0.465278 & 0.479167 \\ 0.356481 & 0.350694 & 0.354167 \end{pmatrix}$$

$$T^4 = \begin{pmatrix} 0.17554 & 0.177662 & 0.175347 \\ 0.470679 & 0.469329 & 0.472222 \\ 0.353781 & 0.353009 & 0.352431 \end{pmatrix}$$

$$T^5 = \begin{pmatrix} 0.176183 & 0.176553 & 0.176505 \\ 0.470743 & 0.47039 & 0.470775 \\ 0.353074 & 0.353057 & 0.35272 \end{pmatrix}$$

$$T^6 = \begin{pmatrix} 0.176414 & 0.176448 & 0.176529 \\ 0.470636 & 0.470575 & 0.470583 \\ 0.35295 & 0.352977 & 0.352889 \end{pmatrix}$$

A prob de estares no estados vão ser constantes

Continuando... (em Matlab)

% n =10

Tn =

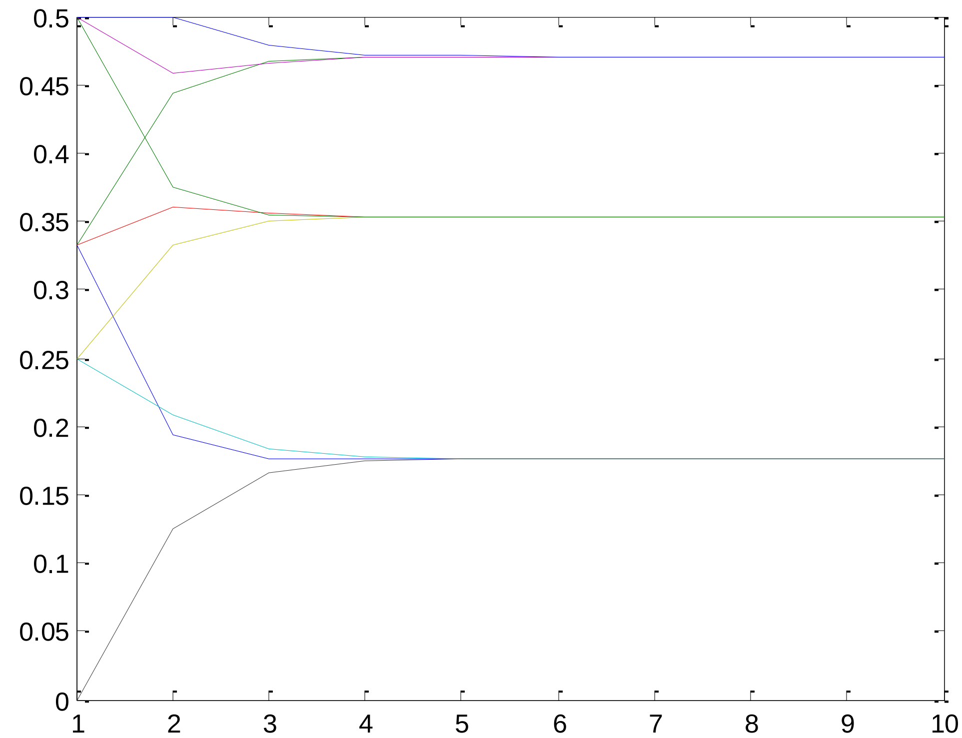
0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	0.3529

% n=100

Tn =

0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	+0.3529

1.0

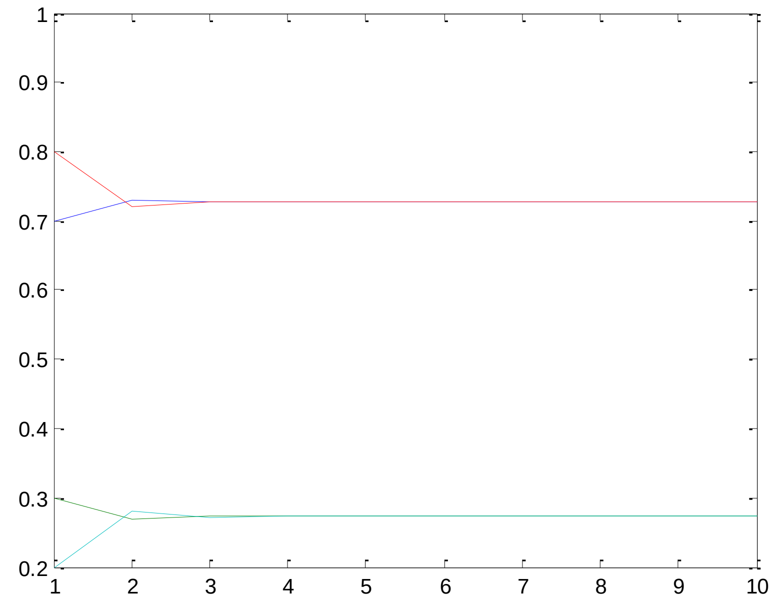


↳ Multiplicando pelo número de alunos, temos uma estimativa.

Exemplo 1 (faltar/não faltar)

```
T=[0.7 0.8
     0.3 0.2]
Tn= T;
pij= Tn(:);
for n=2:10
    Tn= T*Tn;
    pij=[ pij  Tn(:)];
    plot(pij')
    drawnow
end
Tn
```

$T_n = T \times T_n$




```
Tn =

    0.7273    0.7273
    0.2727    0.2727
```

Questões ?

- Converge ?
- Para quê ?

Equilíbrio

- As cadeias dos nossos exemplos atingem um equilíbrio.
- Quando isso acontece a probabilidade de qualquer estado torna-se constante independentemente do passo (step) e das condições iniciais
- Para analisar essa situação é necessário considerar um certo tipo de cadeias de Markov…

$$\lim_{n\to\infty} T^n$$

- Se T é a matriz de transição de um processo de Markov regular então:
- $\lim_{n\to\infty} T^n$ é a matriz:

$$A = \begin{bmatrix} u_1 \, u_1 \cdots u_1 \\ u_2 \, u_2 \cdots u_2 \\ \cdots \; \cdots \cdots \cdots \\ u_N u_N \ldots u_N \end{bmatrix}$$

$n \to \infty$

$\Rightarrow \begin{bmatrix} u_1 \\ u_2 \\ \vdots \end{bmatrix}_{n} \overset{\text{Não muda}}{\longleftrightarrow} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \end{bmatrix}_{n+1}$

$\Rightarrow x^{(n+1)} = T x^{(n)} \quad (x^{n+1} = x^n = u)$

$\Rightarrow \boxed{u = T u}$

$\vdots$

$\Rightarrow (T - I) u = 0$

$\hookrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Com todas as colunas idênticas

- Cada coluna **u** é um vetor probabilidade em que todas as componentes são positivas

Vector estado estacionário (steady-stade vector)

- Sendo T uma matriz de transição regular e A e $\mathbf{u}$ o resultado de $\lim_{n\to\infty} T^n$ (slide anterior), demonstra-se que:

a) Para qualquer vector de probabilidade $\mathbf{x}$, $\boldsymbol{T^n}\,\mathbf{x} \rightarrow \mathbf{u}$ quando $n \rightarrow \infty$

Sendo $\mathbf{u}$ o vector estado estacionário (**steady-state vector**)

b) $\mathbf{u}$ é o único vector de probabilidade que satisfaz a equação matricial $\mathbf{Tu} = \mathbf{u}$

c) $\lim_{\boldsymbol{n}\to\infty}\left(\frac{\boldsymbol{\eta(i,n)}}{\boldsymbol{n}}\right) = \boldsymbol{u}(\boldsymbol{i})$, em que $\eta(i,n)$ é o número de visitas ao estado ***i*** em ***n*** passos (transições)

Cálculo do vector estado estacionário

- Sabemos que o vector correspondente ao estado estacionário é único.
- Usamos a equação que ele satisfaz para o calcular: $\boxed{\underset{x^{(n)}}{\mathbf{Tu}} = \underset{x^{(n+1)}}{\mathbf{u}}}$

- Ou, na forma matricial, $(\mathbf{T} - \mathbf{I})\mathbf{u} = 0$

  ↳ em matlab $\boxed{\text{eye()}}$

Exemplo 1 (aulas)

- $\mathbf{Tu} = \mathbf{u}$
- $\begin{bmatrix} 0,7 & 0,8 \\ 0,3 & 0,2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$
- $\begin{cases} \frac{7}{10}u_1 + \frac{8}{10}u_2 = u_1 \\ \frac{3}{10}u_1 + \frac{2}{10}u_2 = u_2 \\ \boxed{u_1 + u_2 = 1} \end{cases} = \begin{cases} \frac{-3}{10}u_1 + \frac{8}{10}u_2 = 0 \\ \frac{3}{10}u_1 + \frac{-8}{10}u_2 = 0 \\ u_1 + u_2 = 1 \end{cases}$

Não esquecer!

- ...

Em Matlab

Uma possível solução:

```
% matriz de transição
T=[7 8; 3 2]/10
% (T-I)u aumentado com
u1+u2
M=[T-eye(2);
  ones(1,2)]
%
x=[0 0 1]'
% resolver para obter u
u=M\x
```

Resultado:

0.7273

0.2727

Ou seja aprox. 72 % de probabilidade de não faltarem

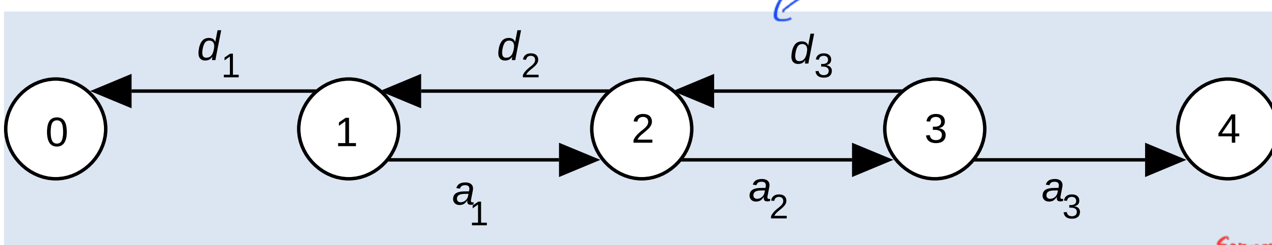
Cadeias com estados absorventes

• Existem matrizes que não colaboram…

Estado absorvente

- Um estado **absorvente** é um estado do qual não é possível sair
  - – ou seja transitar para outro estado

ex: existe um bebado no meio!



Estados Transientes

$$\begin{array}{c|ccc|cc} & 1 & 2 & 3 & 0 & 4 \\ \hline 1 & & & & 0 & 0 \\ 2 & & & & 0 & 0 \\ 3 & & & & 0 & 0 \\ \hline 0 & & & & 1 & 0 \\ 4 & & & & 0 & 1 \end{array}$$

Separar os estados absorventes

0

- Os estados 0 e 4 são absorventes

Resolvido mais à frente!

- Uma cadeia é absorvente se tiver pelo menos um estado absorvente

- Exemplo: http://demonstrations.wolfram.com/AbsorbingMarkovChain/

Forma canónica da Matriz de Transição

- Se numa matriz de transição **agruparmos todos os estados absorventes** obtemos a denominada forma canónica (standard form)

- O mais usual é colocar primeiro os não absorventes e depois os absorventes

- **A forma canónica** é muito útil para determinar as matrizes em situações limite de cadeias de Markov absorventes
  - Como veremos…

Forma canónica

- Rearranjar os estados da matriz T por forma a que os estados transientes apareçam primeiro

Matriz t x t

$$\begin{matrix} & \begin{matrix}\text{TR.} & \text{ABS.}\end{matrix} \\ \begin{matrix}\text{TR.} \\ \text{ABS.}\end{matrix} & \left(\begin{array}{c|c} \boxed{\mathbf{Q}} & \mathbf{0} \\ \hline \mathbf{R} & \mathbf{I}\end{array}\right)\end{matrix}$$

t: # estados transientes

a:# estados absorventes

a x a matriz identidade

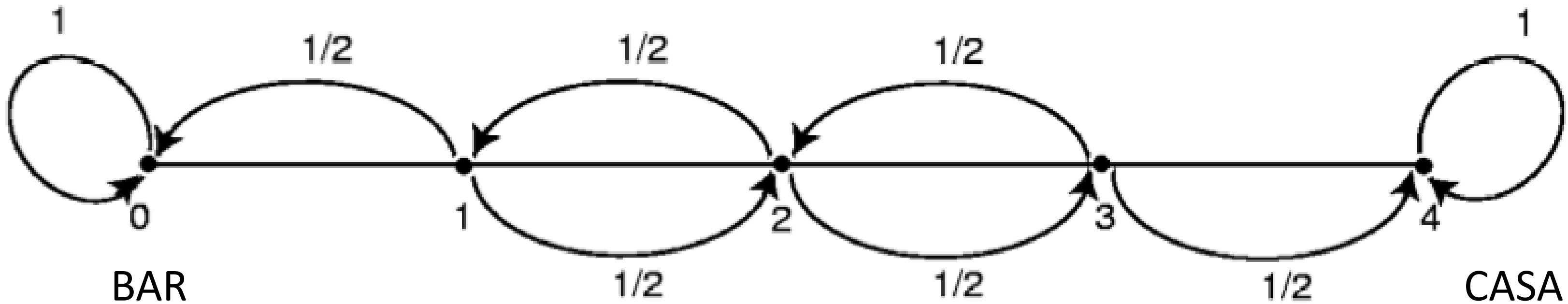
Q e R

- A sub-matriz Q descreve probabilidades de transição de estados não-absorventes para estados não-absorventes

- R contém as probabilidades de transição de não-absorventes para absorventes

Exemplo

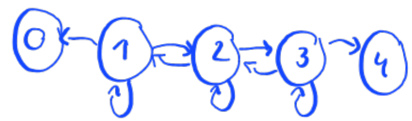
- Diagrama de transição



- Matriz de transição (não canónica):
  – Estados 0, 1, 2 …

- $T = \begin{bmatrix} 1 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1 \end{bmatrix}$

Forma canónica



- $$T = \begin{bmatrix} 1 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1 \end{bmatrix}$$

$$\begin{array}{l} \qquad\quad \overbrace{\begin{array}{ccc} 1 \quad & 2 \quad & 3 \end{array}}^{\text{Estados transientes}} \quad \overbrace{\begin{array}{cc} \mathbf{0} & \mathbf{4} \end{array}}^{\text{Estados Absorventes}} \\ T = \left[\begin{array}{ccc|cc} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \\ \hline 1/2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \end{array}\right] \end{array}$$

Q

0

R

I

Situação limite

Situação limite

- Como é óbvio a cadeia irá acabar por ficar indefinidamente num dos estados absorventes !

- Mas mesmo assim existem questões relevantes:
  - Qual o estado absovervente mais provável quando temos vários ?
  - Dado um estado inicial, qual o número esperado de transições até ocorrer absorção ?
  - Dado um estado inicial, qual a probabilidade de ser absorvido por um estado absorvente em particular ?

Potências de T

- Multiplicando repetidamente a matriz de transição na sua forma canónica vê-se que:

$$T^n = \begin{bmatrix} Q^n & 0 \\ X & I \end{bmatrix}$$

- A expressão exacta de X não tem interesse, mas Q e Q^n são importantes

Q^n

- A matriz Q^n representa a probabilidade de permanecer em estados não-absorventes após *n* passos

- $\boxed{Q^n \to 0}$ quando $n \to \infty$

$$\lim_{n \to \infty} Q^n = 0$$

Matriz fundamental

- Multiplicando verifica-se que $(I-Q)(I+Q+Q^2+\cdots+Q^n)=I-Q^{n+1}$
- Fazendo $n\to\infty$ temos
$$(I-Q)(I+Q+Q^2+\cdots)=I$$
  - porque $Q^n\to 0$

- Isto mostra que
$$(I-Q)^{-1}=I+Q+Q^2+Q^3+\cdots$$

Matriz Fundamental

$$\boxed{\boldsymbol{F} = (\boldsymbol{I} - \boldsymbol{Q})^{-1}}$$

$$= I + Q + Q^2 + \ldots$$

é a matriz fundamental do percurso aleatório

Interpretação de F

- Sejam $X_k(ji)$ as variáveis aleatórias definidas por:

- $$X_k(ji) = \begin{cases} 1, & se\ estiver\ em\ j\ após\ k\ passos, \\ & partindo\ de\ i \\ 0, & caso\ contrário \end{cases}$$

- A soma $\overset{k=0}{X_0}(ji) + X_1(ji) + \cdots + \overset{k=0}{X_n}(ji)$ = Soma dos probabilidades de passar naquele estado representa o número de visitas ao estado j, partindo do estado i, ao fim de n passos.

- O seu valor médio é dado por

← i

$$\left[\begin{array}{c} X_k(ji) \\ \\ \\ \end{array} \right]$$

↑ j

Somamos o valor em cada K transição ...

$$E[X_0(ji) + X_1(ji) + \cdots + X_n(ji)] = \sum_{k=0}^{n} E[X_k(ji)]$$

// Soma dos v.a.

Iremos abordar no futuro

- Lembrar média de soma de variáveis !

Interpretação de F (continuação)

Valor está em Q^k

- Mas $E[\overbrace{X_k(ji)}^{\text{Variavel de Bernoulli}}] = 1 \times p + 0 \times (1-p)$ $= p$ como em qualquer variável de Bernoulli
- E p designa a probabilidade de atingir o estado j após k passos, partindo de i
  - Ou seja exactamente o valor da coluna i e linha j de Q^k.

- Logo:

$$E[X_0(ji) + X_1(ji) + \cdots + X_n(ji)] = \sum_{k=0}^{n} Q^k(ji)$$

$$= (1 - Q)^{-1}$$

Interpretação de *F* (continuação)

- Os elementos de $I + Q + Q^2 + Q^3 + \cdots + Q^n$ exprimem portanto o número médio de visitas ao estado *j* partindo do estado *i* em *n* passos

$$\begin{array}{cc} & i \\ j & \begin{bmatrix} F_{ji} & \\ & \end{bmatrix} \end{array}$$

- Logo, **a matriz fundamental F** – que é o limite dessa quantidade quando $n \to \infty$ - **representa o número médio de visitas a cada estado antes da absorção**

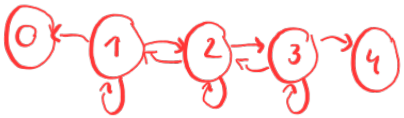
- F_{ji} dá-nos o valor esperado para o número de vezes que um processo se encontra no estado s_j se começou no estado s_i
  - Antes de ser absorvido

Aplicando ao nosso exemplo

- $Q = \begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{bmatrix}$

$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- $I - Q = \begin{bmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{bmatrix}$



Em média passamos $\frac{3}{2}$ das vezes começando no 1

- $F = (I - Q)^{-1} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} 3/2 & 1 & 1/2 \\ \boxed{1} & 2 & 1 \\ 1/2 & 1 & 3/2 \end{bmatrix}$

→ Se somarmos as colunas, sabemos o número de vezes que em média chegamos ao estado absorvente!

↓

"Tempo médio absorvente"

"Começando no 1 passamos em média 1 vez no 2"

Tem o seu sentido...

Tempo médio até à absorção

- O **tempo médio até à absorção** será a soma do número médio de visitas a todos os estados transientes até à absorção

- Ou seja a **soma da coluna *i* de *F***

$$t_i = \sum_j F_{ji}$$ // Somando a coluna!

- Na forma matricial pode obter-se o vector t usando

$$\boxed{t = F' \, \mathbf{1}}$$ Soma as colunas, bom para matlab

- Em que :
  - $\mathbf{1}$ é uma vector coluna com uns

Tempo médio até absorção

- A soma da coluna *i* de *F* representa:
  - O valor esperado do **número de vezes que a cadeia passa por um qualquer estado transiente partindo do estado inicial** *i* antes da absorção
  - Valor esperado do tempo necessário até absorção partindo do estado *i*
- O vetor *t* contém os tempos médios até à absorção partindo dos vários estados transientes

Aplicando ao Exemplo: Tempo até absorção

- $t = F' \; 1$

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

$$= \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} .$$

Probabilidades de absorção

- As **probabilidades de absorção** b_{ji} no estado s_j se se iniciar no estado s_i podem ser obtidas através de:

$$B = R\,F$$

$$\begin{bmatrix} Q & 0 \\ R & I \end{bmatrix}$$

$$= \begin{array}{c} \\ 0 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \left[\begin{array}{ccc} \; & \; & \; \\ \; & \; & \; \end{array} \right] \end{array}$$

- Em que B é uma matriz a x t com entradas b_{ji}

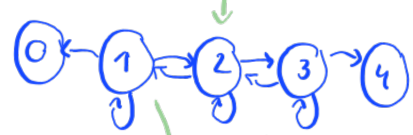
Aplicação ao nosso exemplo

$\boxed{B = RF}$

- Relembremos que temos:

$$T = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \end{bmatrix} \text{ e } F = \begin{bmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{bmatrix}$$

Aqui é ½ (0 ou 4)



Faz sentido!

- E portanto $R = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$

Começando no 1, a prob. de ser absorvido pelo 0 é ¾

- Multiplicando R e F obtemos $B = \begin{array}{c} \\ 0 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 3/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 3/4 \end{bmatrix} \end{array}$

Prob. de absorção!