Cadeias de Markov

?!

- O que esteve na origem do grande sucesso inicial da Google ?

- O que tem em comum esse sucesso com a capacidade de interagir por voz com computadores, robôs e smartphones ?
  - No reconhecimento de fala ?
  - Na síntese de fala ?

Exemplo 1

- 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%;
- se faltaram à anterior, essa probabilidade é 80%

- 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 ?

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

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

Exemplo 3 – "Pub Crawl"

- Bares junto a uma conhecida Universidade:

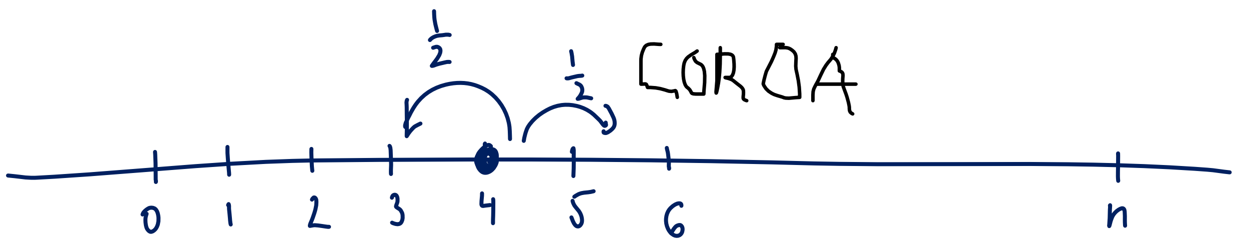
TPC

Fazer Diagrama para os Bares de Aveiro

Outro exemplo

- Passeio aleatório (random walk)

Lançar moeda



Cara Coroa Coroa … Cara …

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

Princípios básicos

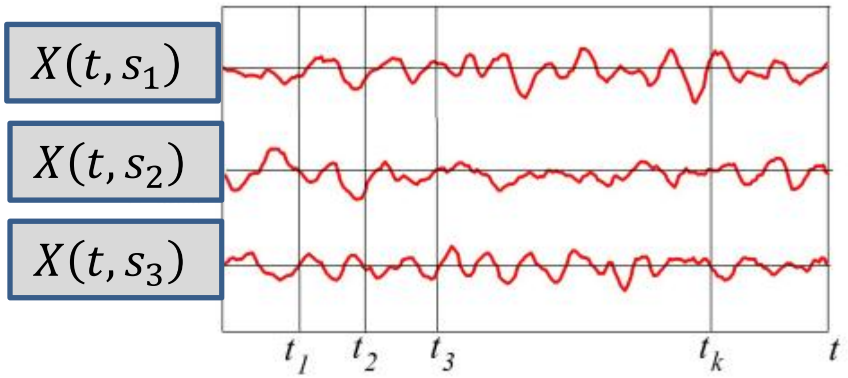
Processos estocásticos

- Estendem o conceito de variável aleatória
- Lidam com a dinâmica da teoria de probabilidades

- Uma v.a. X mapeia um acontecimento $s \in \Omega$ num número $X(s)$
- O processo 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



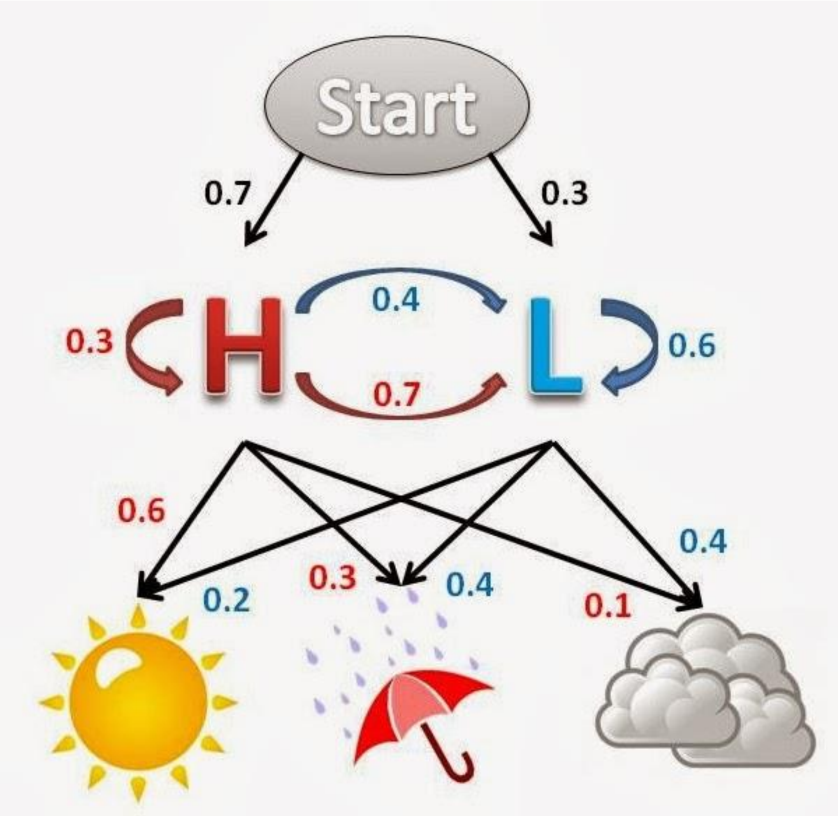
- Se fixarmos s, $X(t)$ é uma função real do tempo
- $X(t, s)$ pode então ser vista como uma colecção de funções no tempo
- Se fixarmos t temos uma função $X(s)$ que depende apenas de s, ou seja uma variável aleatória
- Um nome alternativo é processos aleatórios

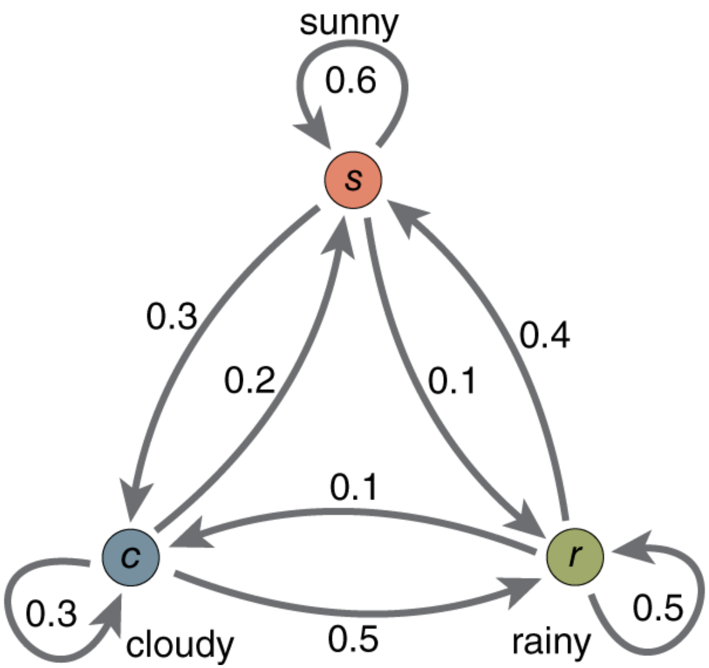
Classificação de processos estocásticos

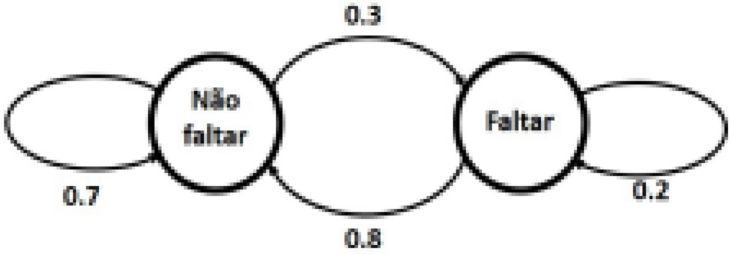
- Podem ser classificados segundo t e os valores que pode assumir (estados do processo)

- Quanto ao tempo :
  - Tempo contínuo: Se tempo é um intervalo contínuo
  - 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

Estados


Start
0.7
0.3
0.3
H
0.4
0.7
L
0.6
0.6
0.2
0.3
0.4
0.1
0.4


sunny
0.6
s
0.3
0.2
0.1
0.4
0.1
c
r
0.3
cloudy
0.5
rainy
0.5


0.3
Não faltar
Faltar
0.7
0.2
0.8

Definição

- Um processo de Markov 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:
- $\boldsymbol{p_{ji}} = P(X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, \dots, X_0 = x_0)$

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

- 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

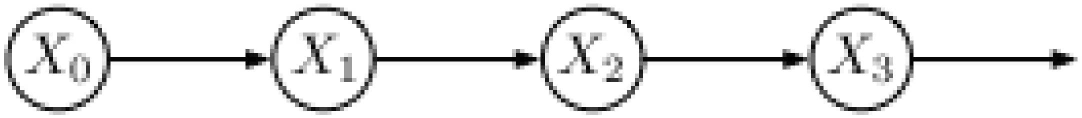
Propriedade de Markov

- $P(X_0 = x_0, X_1 = x_1, X_2 = x_2 \ldots) = ?$

$= P(X_0 = x_0)\, P(X_1 = x_1 | X_0 = x_0)\, P(X_2 = x_2 | X_0 = x_0, X_1 = x_1) \ldots$

$= P(X_0 = x_0)\, P(X_1 = x_1 | X_0 = x_0)\, P(\boldsymbol{X_2 = x_2 | X_1 = x_1}) \ldots$

- O processo "não tem memória"



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

- Estados ?

- Transições ?

- Probabilidades de transição ?

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

Xn

Xn+1

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix}$$

  - 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

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

Futuro Estado

Matriz T do Exemplo 2

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

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**

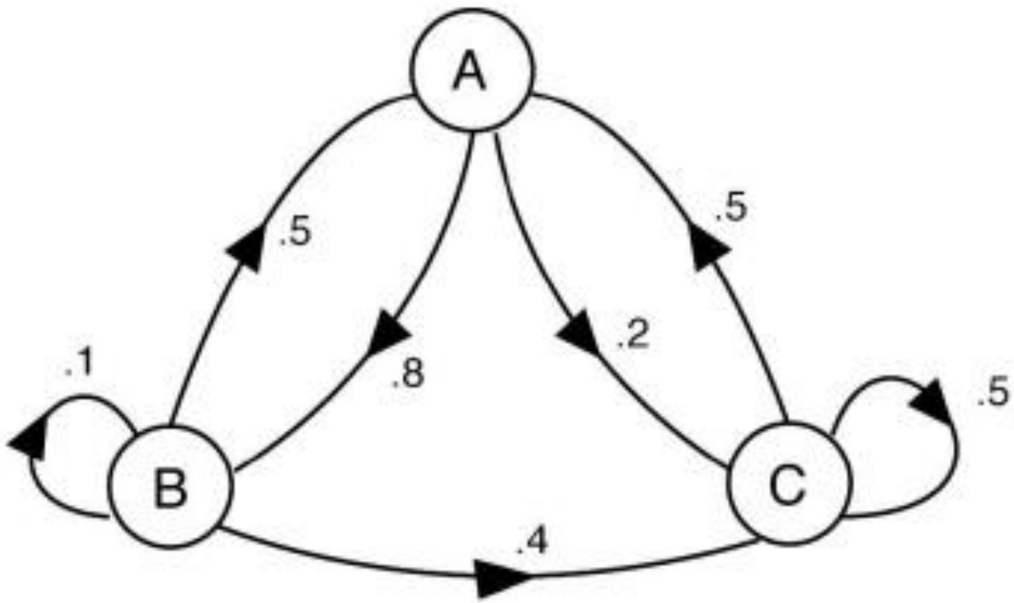
Representação gráfica da cadeia

- Apropriada e possível para número de estados pequeno

- Nós: representam todos os estados
- Setas: para todas as transições permitidas (one-step)
  - 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/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

Simulando os nossos exemplos

- Exemplo 1:
  - Matriz:
    [[0.7, 0.3],
    [0.8, 0.2]]
  - http://setosa.io/markov/index.html#%7B%22tm%22%3A%5B%5B0.7%2C0.3%5D%2C%5B0.8%2C0.2%5D%5D%7D

- Exemplo 2:
  - Matriz:

  [[0.33,0.33,0.34],
  [0.25,0.5,0.25],
  [0,0.5,0.5]]

- Outro exemplo

  [[0,1,0,0],
  [0,0,1,0],
  [0,0,0,1],
  [0.2,0.3,0.3,0.2]]

  - O que vamos ver ?
  - Acesso directo:

  http://setosa.io/markov/index.html#%7B%22tm%22%3A%5B%5B0%2C1%2C0%2C0%5D%2C%5B0%2C0%2C1%2C0%5D%2C%5B0%2C0%2C0%2C1%5D%2C%5B0.2%2C0.3%2C0.3%2C0.2%5D%5D%7D

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:

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

- Este vector 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

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

Vector estado após uma transição

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

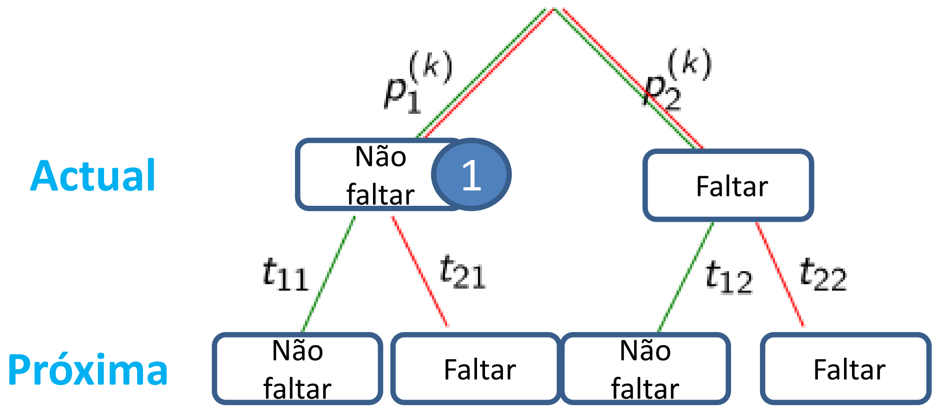
- O vector de estado $\mathbf{x}^{(k+1)}$ no período de observação $k+1$ pode ser determinado a partir do vector $\mathbf{x}^{(k)}$ através de:

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

- 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 ?



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

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

Estado após múltiplas transições

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

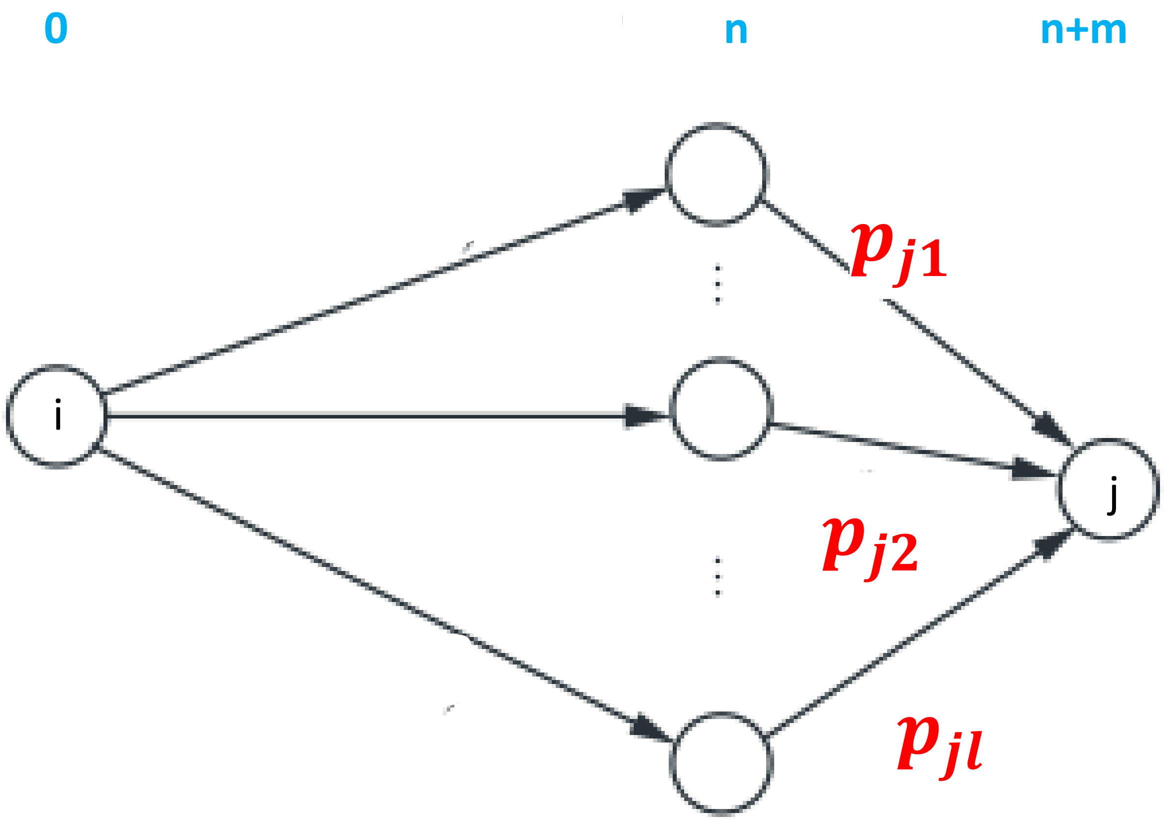
Equações de Chapman-Kolmogorov

- Definindo a transição em n passos ${p_{ji}}^n$ como a probabilidade de **um processo no estado 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 ${p_{ji}}^1 = p_{ji}$
- As **equações de Chapman-Kolmogorov** permitem calcular estas probabilidades

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

Interpretação



Interpretação

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

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

“Demonstração” Eqs. Chapman-Kolmogorov

- ${p_{ji}}^{n+m} = P(X_{n+m} = j | X_0 = i)$
- $= \sum_k P(X_{n+m} = j, X_n = k | X_0 = i)$
- $= \sum_k P(X_{n+m} = j | X_n = k)\, P(X_n = k | X_0 = i)$
- $\sum_k \quad {p_{jk}}^{m} {p_{ki}}^{n}$

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:

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

Em que o "." significa multiplicação de matrizes

- Desta equação obtém-se facilmente:

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

- 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= hoje

…

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

$$= \begin{pmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} = \begin{pmatrix} 0.73 \\ 0.27 \end{pmatrix}$$

- Ou seja probabilidade igual a 0.73 de virem na próxima Segunda

Terminologia

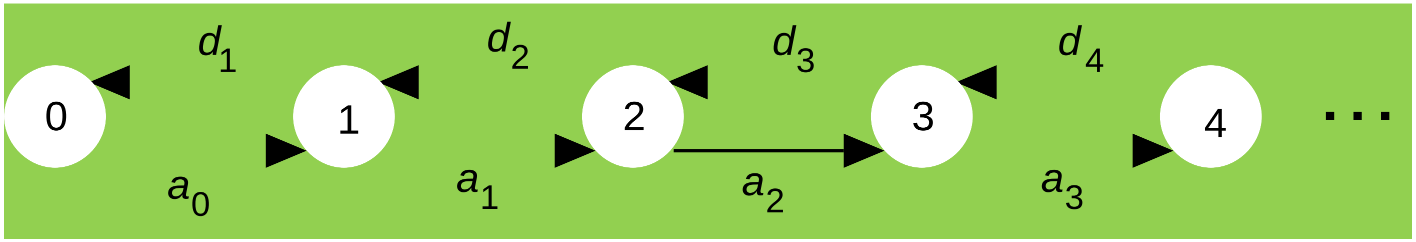
Tipos de estados

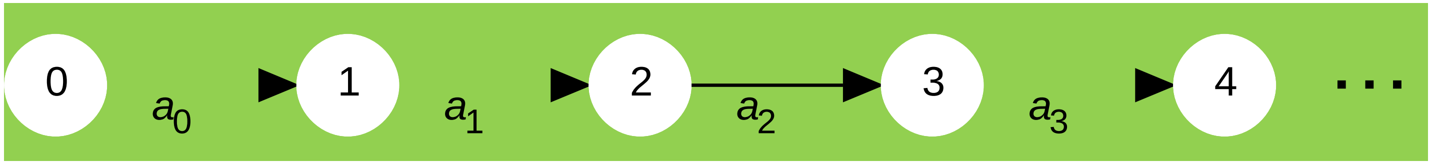
Tipos de matrizes de transição

...

Acessibilidade de um estado

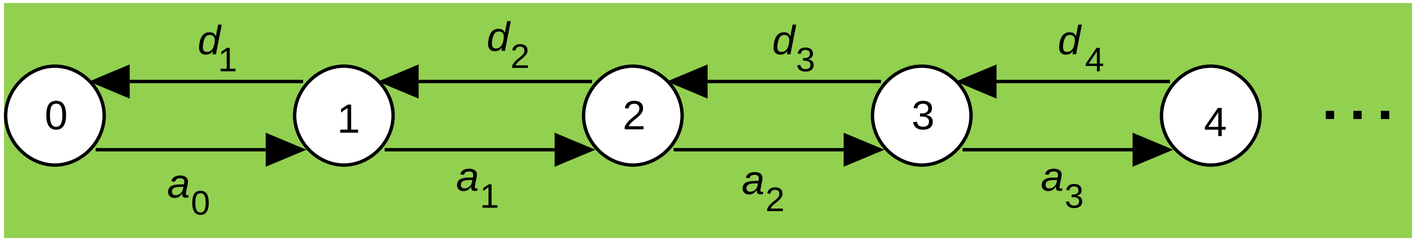
- Possibilidade de ir do estado i para o estado j (existe caminho na cadeia de i para j).





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

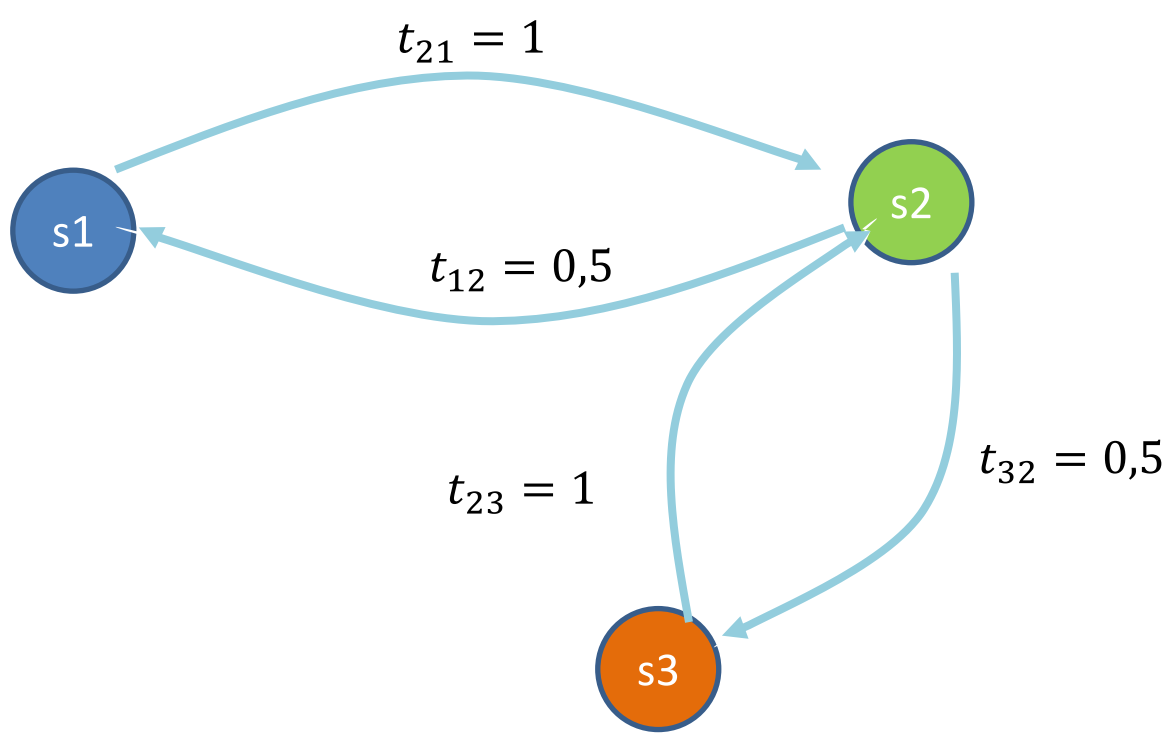
Estado recorrente

- 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

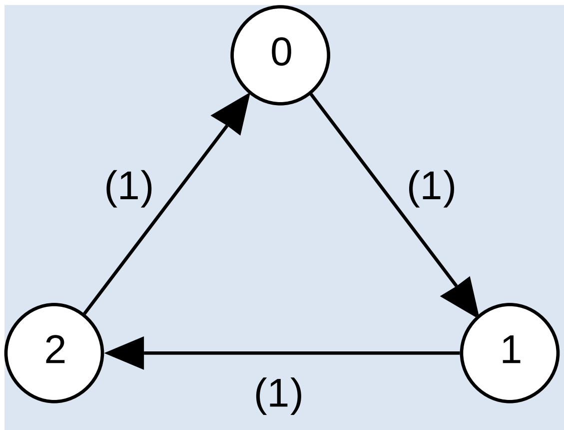
Estado periódico

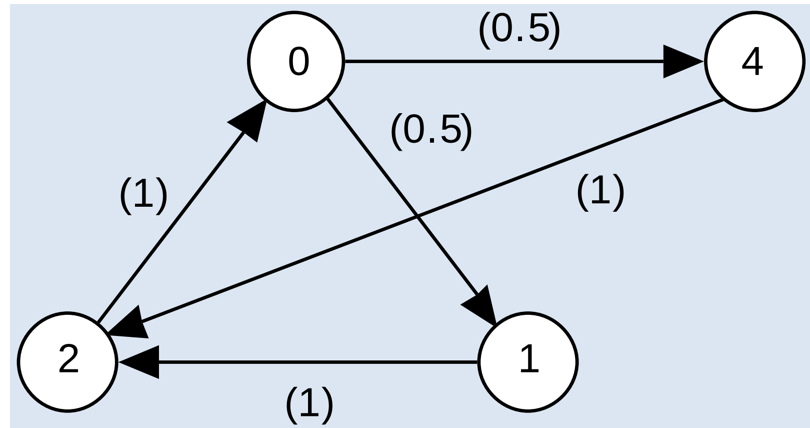
- Um estado é **periódico** se apenas se pode regressar a ele após um número fixo de transições superior a 1 (ou múltiplos desse número).

- 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 $r = c, 2c, 3c, \ldots$

Estado periódico

•





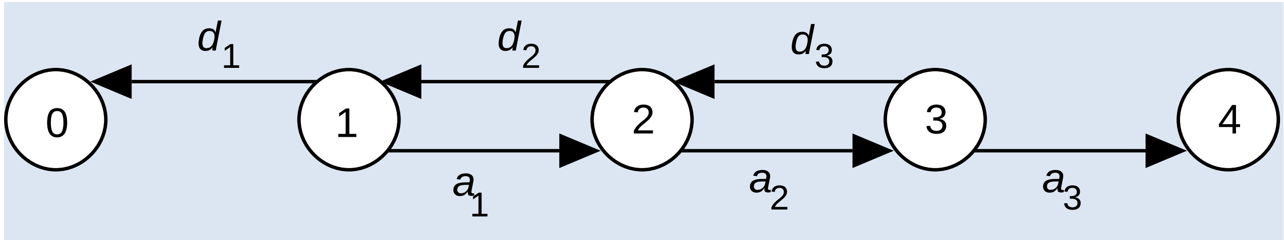
Todos estados visitados em múltiplos de 3 iterações

Todos estados visitados em múltiplos de 3 iterações

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

Estado absorvente

- Um estado **absorvente** é um estado do qual não é possível sair (ou seja transitar para outro estado)
- Uma cadeia é absorvente se tiver pelo menos um estado absorvente

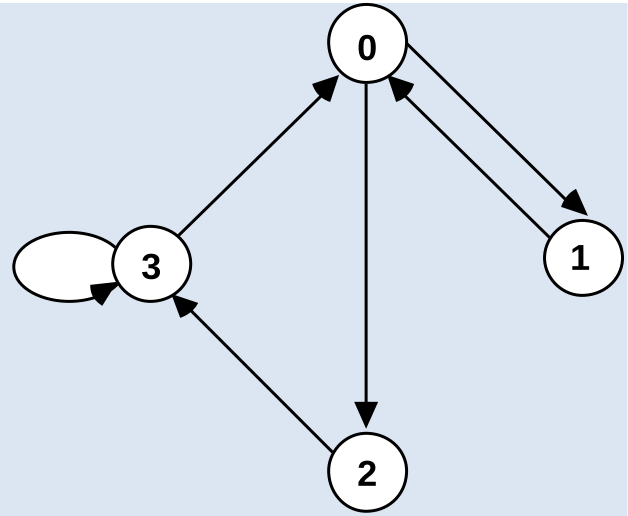


- Os estados 0 e 4 são absorventes

Aplicação dos conceitos

- Exemplo:

$$
\begin{array}{c|cccc}
\text{State} & 0 & 1 & 2 & 3 \\
\hline
0 & & & & \\
1 & & & & \\
2 & & & & \\
3 & & & &
\end{array}
\quad
\begin{array}{c}
\\ 0 \\ 1 \\ 2 \\ 3
\end{array}
\begin{bmatrix}
0 & X & X & 0 \\
X & 0 & 0 & 0 \\
0 & 0 & 0 & X \\
X & 0 & 0 & X
\end{bmatrix}
$$



- Todos os pares de estados comunicam, formando um única classe recorrente
  - Os estados são aperiódicos
- Em consequência o processo é aperiódico e irredutível

Assuntos principais dados anteriormente

- Noção de processo estocástico
- Cadeias de Markov
- Propriedade de Markov
- Matriz de transição T
- Representação gráfica
- $\mathbf{T}^{(n)}$

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...

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

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

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

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

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

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

Continuando… (em Matlab)

% n =10

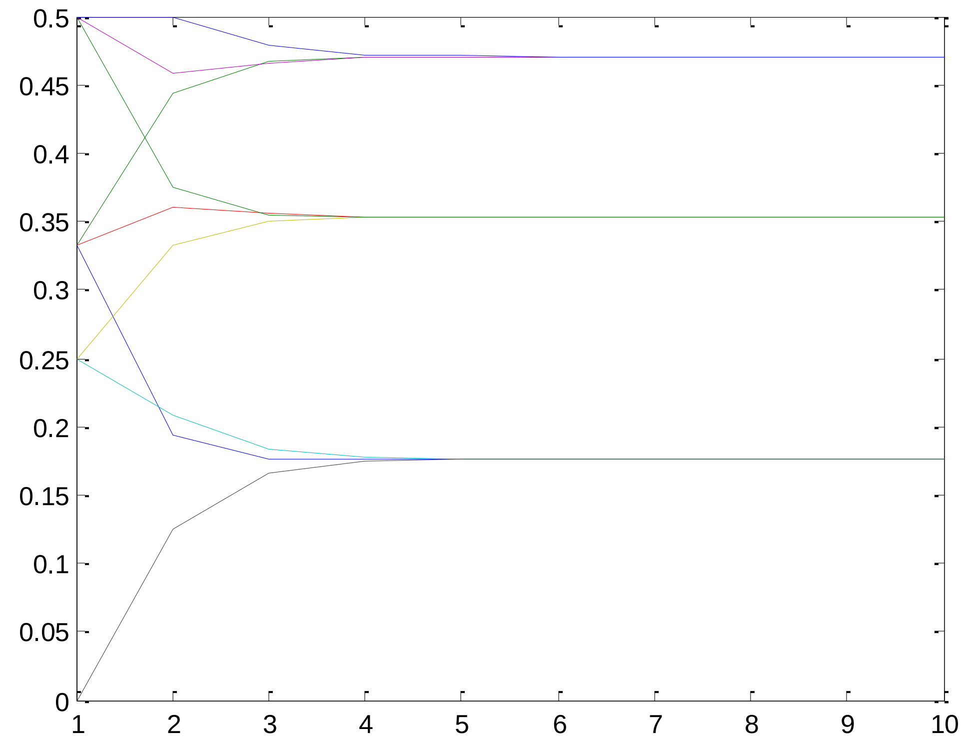
Tn =

0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	0.3529

% n=100

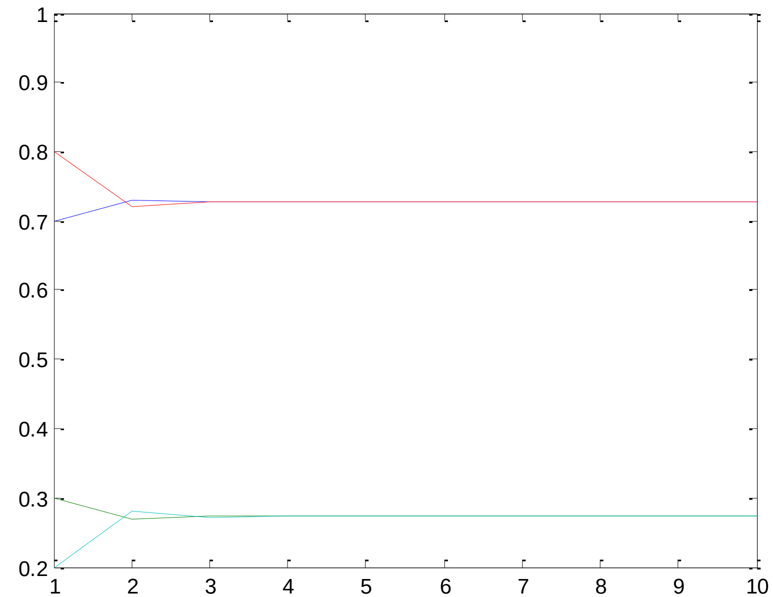
Tn =

0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	0.3529



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
```




```
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…

Matriz/ Processo regular

- A matriz de transição (ou o processo de Markov correspondente) é **regular** se alguma potência da matriz tem todos os valores não-nulos.
  - Existe uma probabilidade de mudar de qualquer estado para qualquer estado

- Qualquer matriz de transição sem elementos nulos é uma matriz regular.
- No entanto, uma matriz contendo elementos nulos pode ser regular.
  - Por exemplo: $T = \begin{bmatrix} 0 & 0 & 0.5 \\ 1 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix}$

...

- No caso de matrizes com elementos nulos, pode verificar-se se é regular substituindo os elementos não-nulos por "X" e calculando potências sucessivas

- No nosso exemplo:

- $T = \begin{bmatrix} 0 & \mathrm{X} & 0 \\ 0 & 0 & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & 0 \end{bmatrix}, T^2 = \begin{bmatrix} 0 & 0 & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & 0 \\ 0 & \mathrm{X} & \mathrm{X} \end{bmatrix} \dots, T^8 = \begin{bmatrix} \mathrm{X} & \mathrm{X} & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & \mathrm{X} \end{bmatrix}$

Cadeia **ergódica**

- Uma cadeia de Markov diz-se ergódica se é possível efectuar transições de qualquer estado para qualquer outro estado

- Em consequência, uma cadeia regular é também ergódica

Cadeia ergódica

- No entanto, nem todas as cadeias ergódicas são regulares
  - Exemplo: se de um determinado estado se pode transitar para alguns estados apenas num número par de transições e para outros num número ímpar de transições, então todas as potências da matriz de transição terão elementos nulos

$$T = \begin{bmatrix} 0 & 0.5 & 0 \\ 1 & 0 & 1 \\ 0 & 0.5 & 0 \end{bmatrix}$$

…

- Potências pares

```
>> T^30

ans =

   0.5000         0    0.5000
        0    1.0000         0
   0.5000         0    0.5000

>> T^100

ans =

   0.5000         0    0.5000
        0    1.0000         0
   0.5000         0    0.5000
```

- Potências ímpares

```
>> T^5
ans =

        0    0.5000         0
   1.0000         0    1.0000
        0    0.5000         0

>> T^51

ans =

        0    0.5000         0
   1.0000         0    1.0000
        0    0.5000         0
```

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

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

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

Com todas as colunas idênticas

- Cada coluna **u** é um vector 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 $\mathbf{x}$, $\boldsymbol{T^n}\,\mathbf{x} \to \mathbf{u}$ quando $n \to \infty$

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) $\lim_{n\to\infty} \left(\frac{\boldsymbol{\eta(i,n)}}{\boldsymbol{n}}\right) = \boldsymbol{u(i)}$, 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: $\mathbf{T}\mathbf{u} = \mathbf{u}$

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

Exemplo 1 (aulas)

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

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

Exemplo1.m

Outra possibilidade

```
%  matriz de transição
T=[7 8; 3 2]/10;

% resolver equação Au= b
aux= (T-eye(length(T)));

% usas as primeiras linhas de T + equação x1+x2= 1
A= [aux(1:end-1,:); 1 1];

b= [0 1]';

u= inv(A)*b
```

...

- Pode também ser resolvido usando uma matriz aumentada e a função **rref**()

```
%Matlab
C= [M   x]
rref(C)
```

$$\left(\begin{array}{cc|c} -3/10 & 8/10 & 0 \\ 3/10 & -8/10 & 0 \\ 1 & 1 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 0 & 8/11 \\ 0 & 1 & 3/11 \\ 0 & 0 & 0 \end{array}\right)$$

- Mais informação: https://www.math.ucdavis.edu/~daddel/MATH22AL/LABS/LAB2/lab2.pdf

Exemplo 2

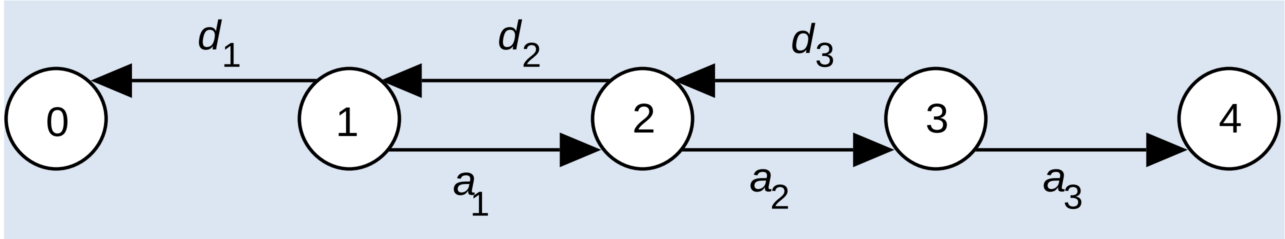
- Aplicando a última técnica ao nosso exemplo 2 (grupos) teremos

$$\left(\begin{array}{ccc|c} -2/3 & 1/4 & 0 & 0 \\ 1/3 & -1/2 & 1/2 & 0 \\ 1/3 & 1/4 & -1/2 & 0 \\ 1 & 1 & 1 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3/17 \\ 0 & 1 & 0 & 8/17 \\ 0 & 0 & 1 & 6/17 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Cadeias com estados absorventes

Estados absorventes

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



- Os estados 0 e 4 são absorventes

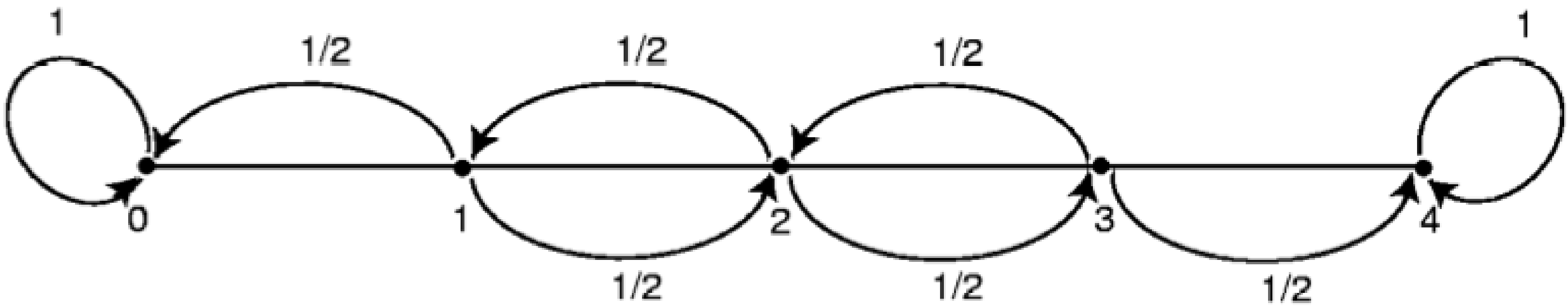
Cadeias absorventes

- Uma **cadeia** é **absorvente** se:

(1) tiver pelo menos um estado absorvente

(2) é possível ir de cada um dos estados não absorventes para pelo menos um dos estados absorventes num número finito de passos.

Exemplo simples



Demo

- Absorbing Markov Chain
  - http://demonstrations.wolfram.com/AbsorbingMarkovChain/

Forma canónica da matriz de transição

Forma canónica

- 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

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

t: # estados transientes

a:# estados absorventes

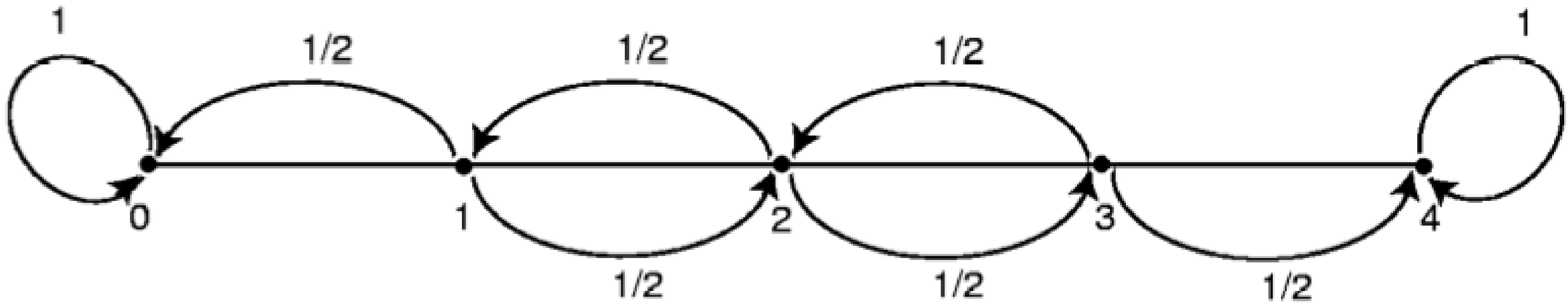
a x a matriz identidade

Aplicação a um exemplo

- Homem a caminhar para casa de um bar
  - 4 quarteirões entre o bar e a casa
  - 5 estados no total
- Estados absorventes:
  - Esquina 4 – Casa
  - Esquina 0 – Bar
- No limite de cada quarteirão existe igual probabilidade de seguir em frente ou retroceder

Diagrama e matriz de transição

- Diagrama de transição



- Matriz de transição :

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

Forma canónica

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

$$T = \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 0 & 4 \end{array} \\ \left[\begin{array}{ccc|cc} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \\ \hline 1/2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \end{array}\right] \end{array}$$

Q

0

R

I

Diferentes notações

- *Na nossa notação a estrutura é:*

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

- *Na notação alternativa:*

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

Q

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

Situação limite

Situação limite

- Situações limite de cadeias de Markov absorventes ?

- 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

  – Q^n tende para zero quando *n* aumenta

    - $Q^n \to 0$ quando $n \to \infty$

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

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

é a matriz fundamental do percurso aleatório

Interpretação de F

- Sejam $X_k(ji)$ as variáveis aleatórias definidas por:
- $X_k(ji) = \begin{cases} 1, \; se\ estiver\ em\ j\ após\ k\ passos, \\ \qquad partindo\ de\ i \\ 0, \qquad caso\ contrário \end{cases}$

- A soma $X_0(ji) + X_1(ji) + \cdots + X_n(ji)$ representa o número de visitas ao estado j, partindo do estado i, ao fim de n passos.
- O seu valor médio é dado por

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

- Lembrar média de soma de variáveis !

Interpretação de F (continuação)

- Mas $E[X_k(ji)] = 1 \times p + 0 \times (1-p)$ 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:

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

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

- Logo, **a matriz fundamental $\boldsymbol{F}$** – 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

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

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

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

Exemplo2.m

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***

$$t_i = \sum_j F_{ji}$$

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

$$t = F' \, \mathbf{1}$$

- 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

- $t = F' \; 1$

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

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

Exemplo2.m

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:

$$B = R\,F$$

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

Origem da expressão

- $B_{ji} = \sum_n \sum_k r_{jk} {q^{(n)}}_{ki}$
  - De *i* para *k* (transientes) e de *k* para *j* (absorvente)
  - Lembra-se de Chapman-Kolmogorov ?
- Trocando a ordem dos somatórioa:
- $B_{ji} = \sum_k \sum_n r_{jk} {q^{(n)}}_{ki}$
- Usando definição da matriz fundamental:
- $B_{ji} = \sum_k \ r_{jk} F_{ki}$
- De onde se obtém
- $B_{ji} = (R\ F)_{ji}$

Aplicação ao nosso exemplo

- Relembremos que temos:

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

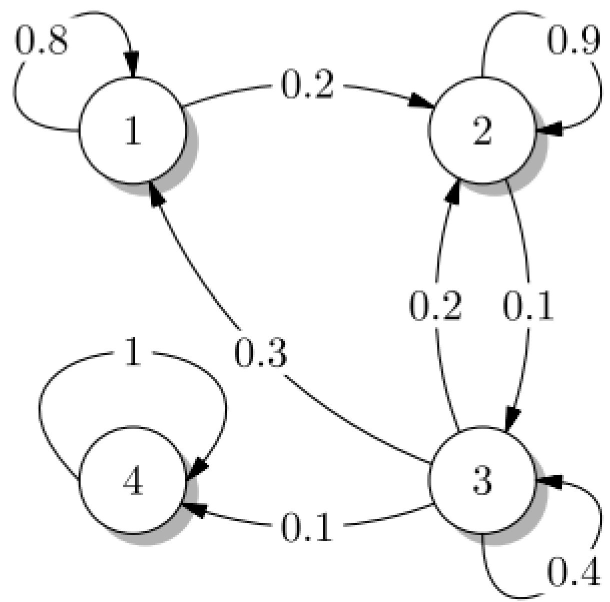
- E portanto $R = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$

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

Exemplo2.m

Aplicação a páginas web…

- Consideremos o conjunto de páginas web da figura:



- Qual o número médio de visitas às páginas 1,2 e 3 quando se parte da página 1?

- Quais os tempos médios até absorção ?

Número médio de visitas às páginas 1,2 e 3 quando se parte da página 1?

- São dados directamente pela matriz *F*
- Em Matlab …

% OBTER T na forma canónica

% Obter Q

submatriz de 3x3

% calcular F

Exemplo3.m

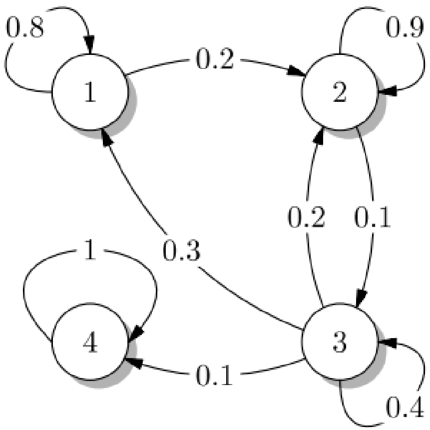
Matlab

```
estados=[1 2 3 4];
% matriz T
Tcan=zeros(4);
Tcan(1,1)=0.8;  Tcan(2,1)=0.2;
Tcan(2,2)=0.9; Tcan(3,2)=0.1;
Tcan(1,3)=0.3; Tcan(2,3)=0.2; Tcan(3,3)=0.4; Tcan(4,3)=0.1;
Tcan(4,4)=1;

%%  Q
Q=Tcan(1:3,1:3)

%%  F
aux= eye(size(Q)) - Q
F=inv(aux)
```

```
F =
20.0000   15.0000   15.0000
60.0000   60.0000   50.0000
10.0000   10.0000   10.0000
```

Resposta à questão

- Os valores que nos interessam são os da coluna 1 (correspondentes a começar na página 1)



- Quando se parte da página 1, o número médio de visitas aos estados 1, 2 e 3 antes de ocorrer absorção será 20, 60 e 10, respectivamente
  - A página 2 receberá mais visitas
  - A página 3, com ligação directa ao estado absorvente terá muito menos visitas

Tempos médios até absorção ?

- Basta obter o vector *t* correspondente à soma das colunas de *F*

%Em Matlab...

```
t=F' * ones(3,1)   % ou sum(F)
t =
  90.0000
  85.0000
  75.0000
```

Exemplo3.m

Matriz B ?

- Neste exemplo não faz sentido pedir B pois só temos um estado absorvente

- Mas se fizermos $B = R\ F$ obtemos um vector de 1x3 só com uns
  - Confirmando o esperado