MPEI 2018-2019

21/22 – Cadeias de Markov (parte 2)

Assuntos principais da aula anterior

- Noção de processo estocástico
- Cadeias de Markov
- Propriedade de Markov
- Matriz de transição T

- Representação gráfica

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

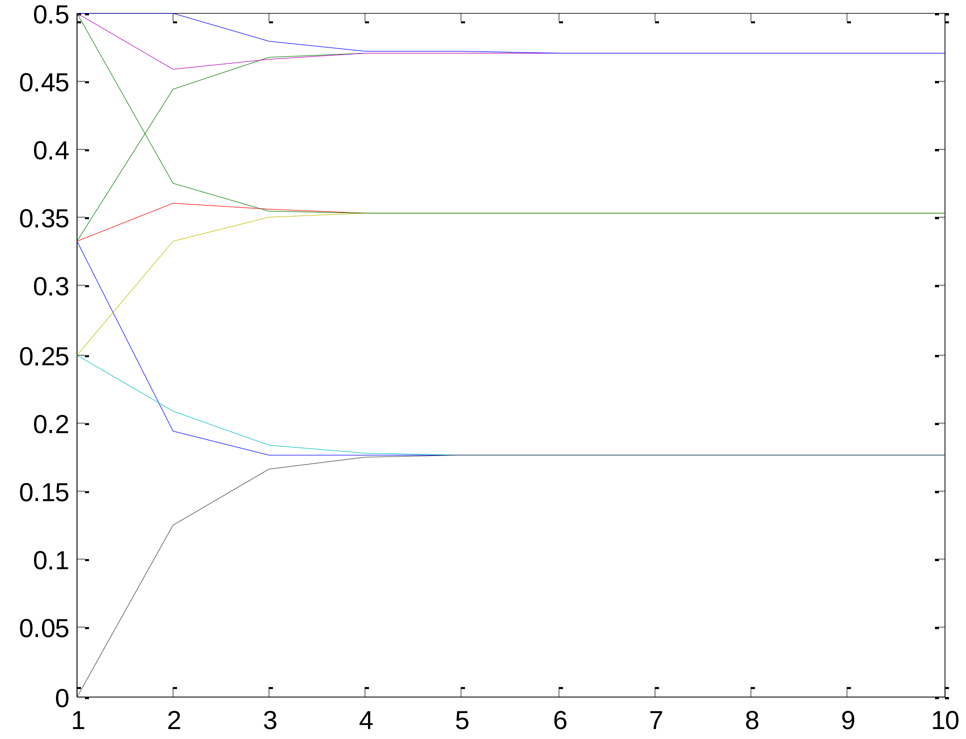
Tn =

0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	0.3529

% n=100

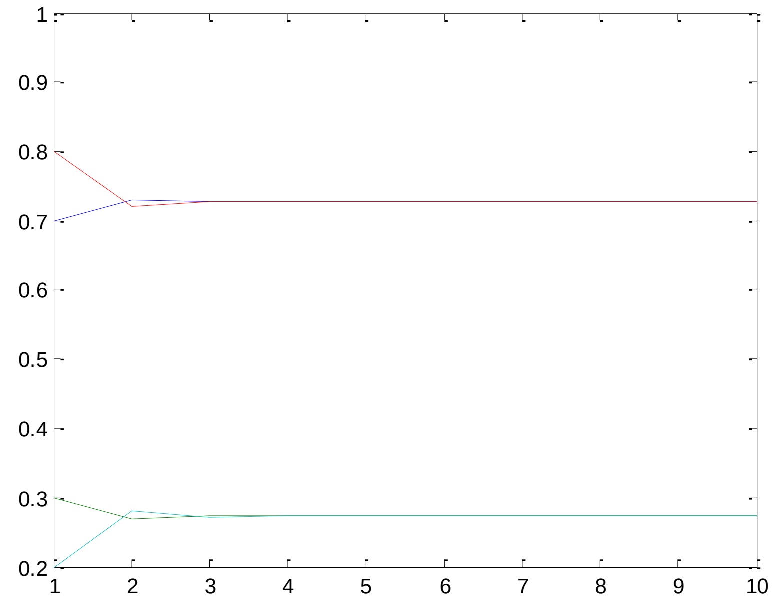
Tn =

0.1765	0.1765	0.1765
0.4706	0.4706	0.4706
0.3529	0.3529	0.3529



Exemplo 1 (faltar/não faltar)

```
Tn=[0.7 0.2
      0.3 0.8]
T= [1 0]’
pij=[];
for n=1:10
   Tn=Tn*T;
   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:
- $T = \begin{bmatrix} 0 & \mathrm{X} & 0 \\ 0 & 0 & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & 0 \end{bmatrix}, T^2 = \begin{bmatrix} 0 & 0 & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & 0 \\ 0 & \mathrm{X} & \mathrm{X} \end{bmatrix} \ldots, T^8 = \begin{bmatrix} \mathrm{X} & \mathrm{X} & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & \mathrm{X} \\ \mathrm{X} & \mathrm{X} & \mathrm{X} \end{bmatrix}$

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:

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

Com todas as colunas idênticas

- Cada coluna **u** é um vector probabilidade

Vector estado estacionário (steady-stade vector)

- Sendo T uma matriz de transição regular e A e $\mathbf{u}$ o resultado 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 $\mathbf{Tu} = \mathbf{u}$

Cálculo do vector estado estacionário

- Sabemos que o vector correspondente ao estado estacionário é única.
- Usamos a equação que ele satisfaz para o calcular: $\mathbf{Tu} = \mathbf{u}$

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

Exemplo 1 (aulas)

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

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

...

- 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_S02/LABS/LAB2/lab2_w01/node9.html

Exemplo 2

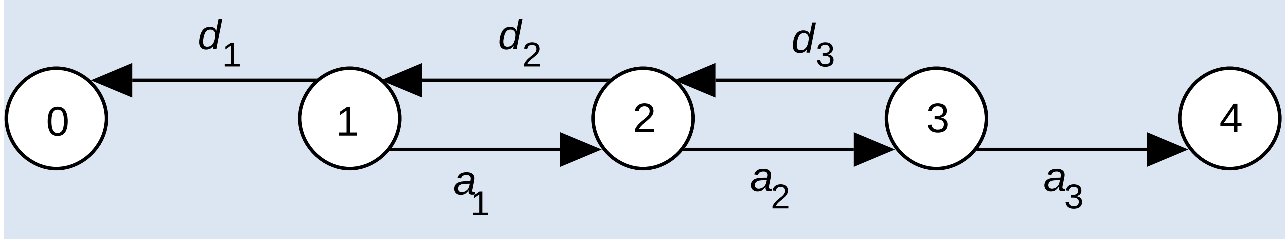
- Aplicando a última técnica ao nosso exemplo 2 (grupos) teremos

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

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

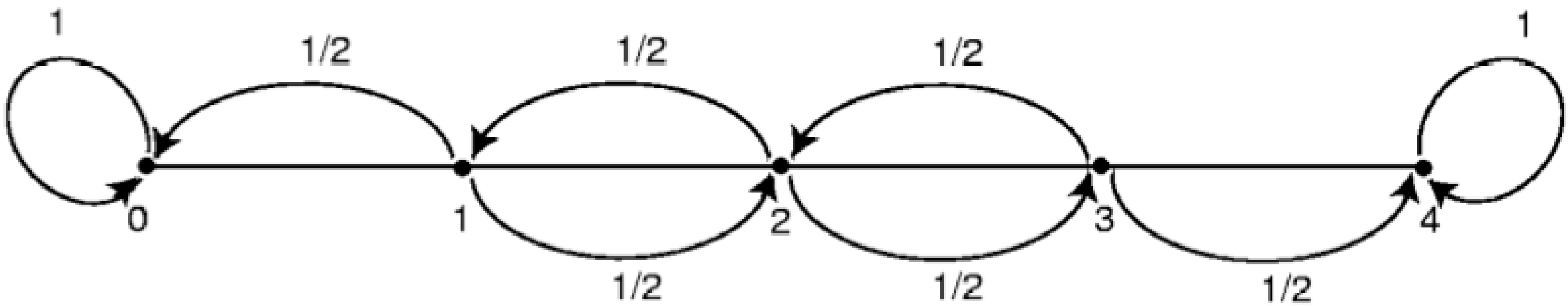
Cadeias absorventes

- 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

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

a x a matriz identidade

t: # estados transientes

a:# estados absorventes

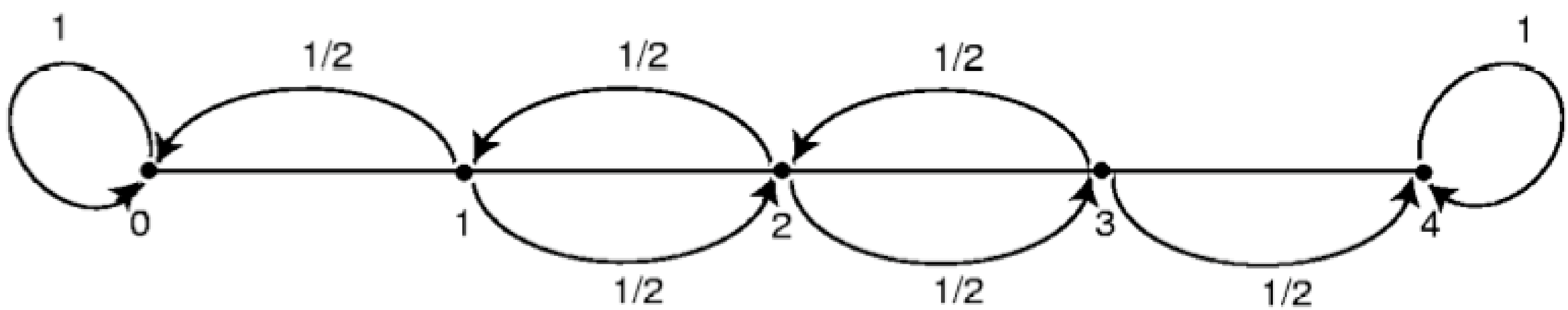
Aplicação a um exemplo

- 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



-

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

Forma canónica

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

Q

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

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 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
  - Que 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:
- $X_k(ji) = \begin{cases} 1, & se\ estiver\ em\ j\ após\ k\ passos, \\ & partindo\ de\ i \\ 0, & caso\ contrário \end{cases}$
- 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 $n \rightarrow \infty$ - **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 de visitas a todos os estados transientes até à absorção

- Ou seja a **soma de uma coluna de F**

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

- 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 até absorção

- Soma de uma coluna de F:
  - Valor esperado do **número de vezes num estado transiente para um dado estado inicial** s_i
  - Valor esperado do tempo necessário até absorção
  - É isto que qualquer valor do vector **t** é

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 somatório:
- $B_{ji} = \sum_k \sum_n r_{jk} {q^{(n)}}_{ki}$
- 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:

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

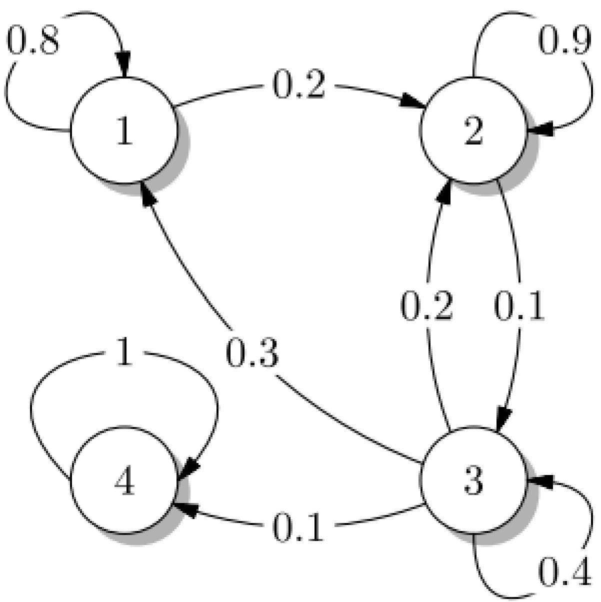
- E portanto $R = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$

- 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

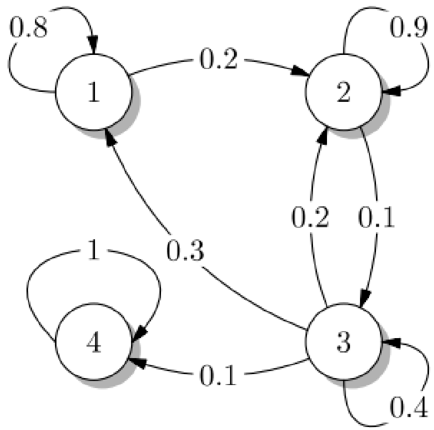
```
estados=[1 2 3 4];
% matriz T
Tcan=zeros(4);
Tcan(1,1)=0.8;  Tcan(2,1)=0.2;
Tcan(2,2)=0.9; Tcan(3,2)=0.1;
Tcan(1,3)=0.3; Tcan(2,3)=0.2; Tcan(3,3)=0.4; Tcan(4,3)=0.1;
Tcan(4,4)=1;

%%  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