MPEI 2015-2016

Aula 17 – Cadeias de Markov (continuação)

11 Novembro 2015

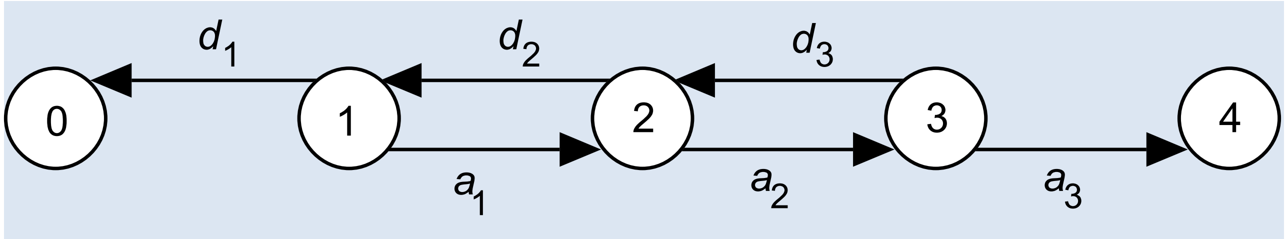
Assuntos principais da aula anterior

- Vector estado
- Múltiplas transições
  - Equações de Chapman-Kolmogorov
  - Potências da matriz de transição
- Estado estacionário
- Tipos de estados: recorrentes, transientes …

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

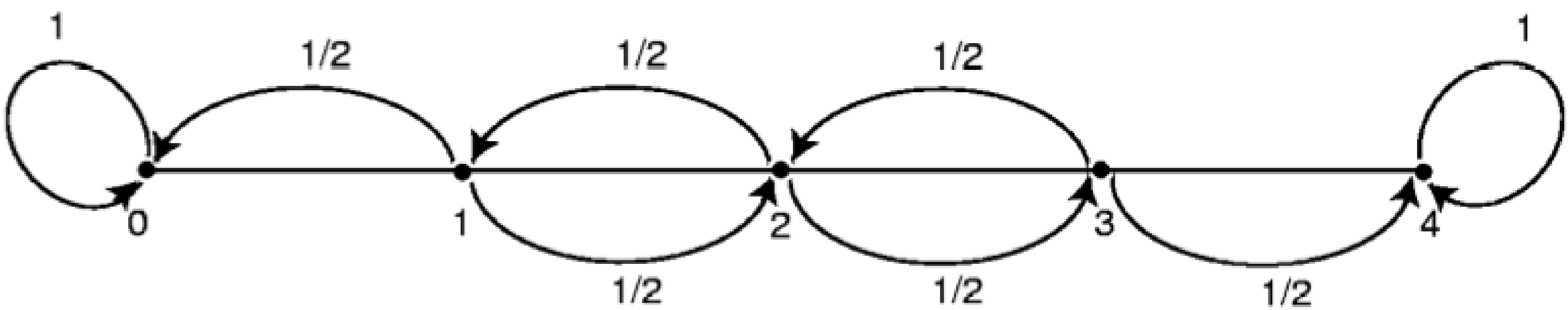
Cadeias absorventes

- 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

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

t: # estados transientes

a:# estados absorventes

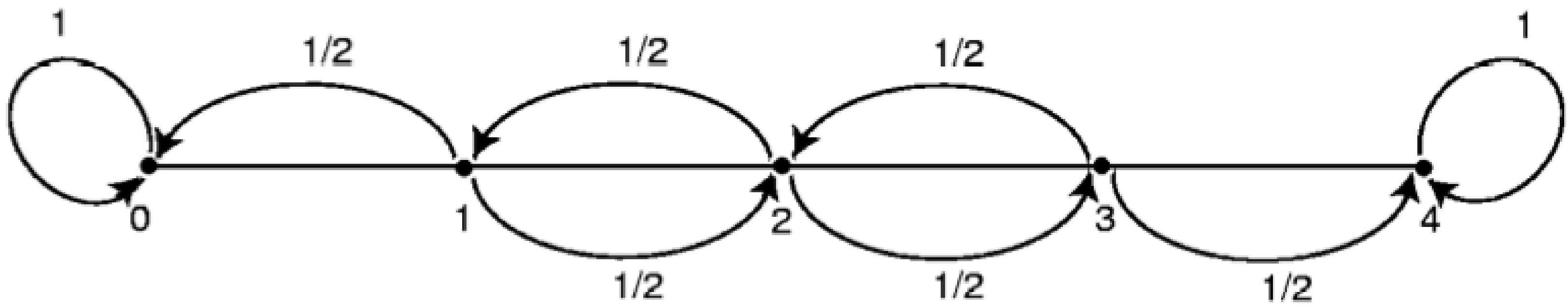
a x a matriz identidade

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



-

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

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

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

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 as matrizes 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 páginas 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 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 número de esperado 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

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

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

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

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

$$\sum_j F_{ji}$$

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

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

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

…

- 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

- $t = F' \ 1$

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

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

Probabilidades de absorção

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

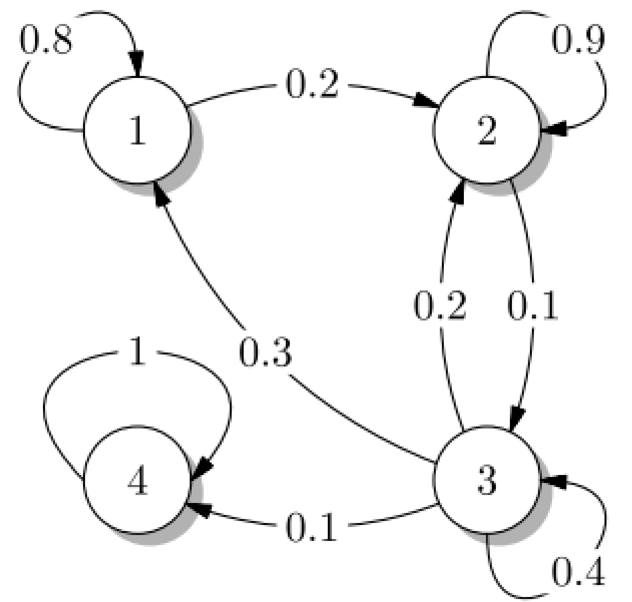
$$T = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \end{bmatrix} \text{ e } F = \begin{bmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{bmatrix}$$

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

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

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

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
```

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

Exercício de revisão

Cadeias de Markov

Problema 1

- Implemente uma função Matlab chamada *markov_estadoestacionario.m* que utilize o método das potências para calcular as probabilidades em estado estacionário para uma cadeia de Markov com N estados.
- Um dos parâmetros de entrada será a matriz de transição
  - Assumida como irredutível e aperiódica
- Outros parâmetros:
  - Vector estado inicial
  - Limiar para terminar o processo (máximo da diferença entre os vectores em duas iterações deverá ser inferior a esse valor)

- Inicialize o processo com um vector uniforme, isto é $x = \frac{[1,1,1,\ldots 1]}{N}$, e limiar=1e-5 caso apenas seja fornecida a matriz de transição

Problema 2

Considere a matriz T seguinte:

T =

0	0.5000	0	0
1.0000	0	0.6000	0.6000
0	0	0	0.4000
0	0.5000	0.4000	0

- Use a função que criou no problema 1 para calcular o vector estado estacionário usando o vector inicial $\mathbf{x} = [0.25, 0.25, 0.25, 0.25]$
- Confirme o resultado calculando o vector estacionário por outro método
- Adicione à sua função a capacidade de mostrar num gráfico o valor do segundo elemento de x em função da iteração (que deverá ser o eixo do xx) e repita o ponto anterior

Problema 3

- Repita o problema anterior para outros vectores iniciais
  - Exemplos: [1,0,0,0] e [0,1,0,0] etc

- Continua a convergir para o mesmo vector ?

Até à próxima aula