

PL 3

Cadeias de Markov

Nota: Adapte a definição da matriz de transição (de estados) em que o elemento t_{ij} da matriz corresponde à probabilidade de transição do estado j para o estado i .

3.1 Exercícios base

1. Considere a seguinte situação e responda às alíneas abaixo:

Um aluno do primeiro ano de um curso de Engenharia tem todas as semanas 2 aulas Teórico-Práticas de uma Unidade Curricular X às 9:00, às quartas e sextas.

Todos os dias que tem aulas desta UC, o aluno decide se vai à aula ou não da seguinte forma: Se tiver estado presente na aula anterior a probabilidade de ir à aula é 70%; se faltou à anterior, a probabilidade de ir é 80%.

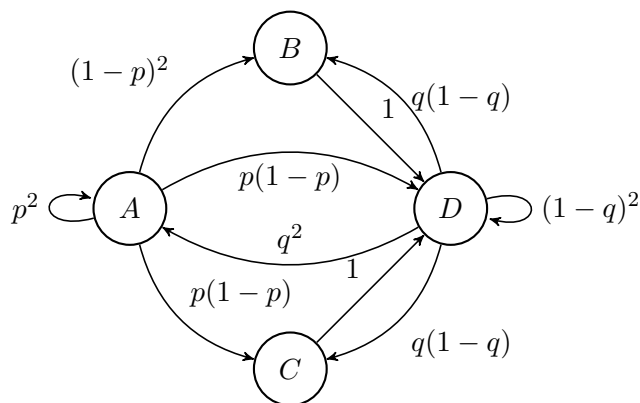
- (a) Se estiver presente na aula de quarta numa determinada semana, qual a probabilidade de estar presente na aula de quarta da semana seguinte?
Sugestão: Comece por definir a matriz de transição de estados e o vetor estado correspondentes.
- (b) Se não estiver presente na aula de quarta numa determinada semana, qual a probabilidade de estar presente na aula de quarta da semana seguinte?
- (c) Sabendo que esteve presente na primeira aula, qual a probabilidade de estar na última aula, assumindo que o semestre tem exactamente 15 semanas de aulas e não existem feriados?
- (d) Represente num gráfico a probabilidade de faltar a cada uma das 30 aulas, assumindo que a probabilidade de estar presente na primeira aula é de 85%.

2. Considere a seguinte “dança” de grupos: Divide-se uma turma em 3 grupos (A, B e C) no início do semestre e no final de cada aula efectua-se os seguintes movimentos:

- 1/3 do grupo A vai para o grupo B e outro 1/3 do grupo A vai para o grupo C;
- 1/4 do grupo B vai para A e 1/5 de B vai para C
- Metade do grupo C vai para o grupo B; a outra mantém-se no grupo C.

- (a) Crie em Matlab a matriz de transição de estados que representa as trocas entre grupos.
Confirme que se trata de uma matriz estocástica.
- (b) Crie o vector relativo ao estado inicial considerando que no total temos 90 alunos, o grupo A tem o dobro da soma dos outros dois e os grupos B e C têm o mesmo número de alunos.
- (c) Quantos elementos integrarão cada grupo no fim da aula 30 considerando como estado inicial o definido na alínea anterior?
- (d) Quantos elementos integrarão cada grupo no fim da aula 30 considerando que inicialmente se distribuíram os 90 alunos equitativamente pelos 3 grupos?

3. Gere aleatoriamente uma matriz de transição de estados para uma cadeia de 20 estados (identificados de 1 a 20) recorrendo à função do Matlab *rand*. Com base nessa matriz:
- Confirme que a matriz de transição de estados é estocástica.
 - Qual a probabilidade de o sistema, começando no estado 1, estar no estado 20 após 2 transições? E após 5? E após 10? E após 100? Apresente os resultados em percentagem e com 5 casas decimais. O que conclui?
4. Considere o seguinte diagrama representativo de uma Cadeia de Markov:

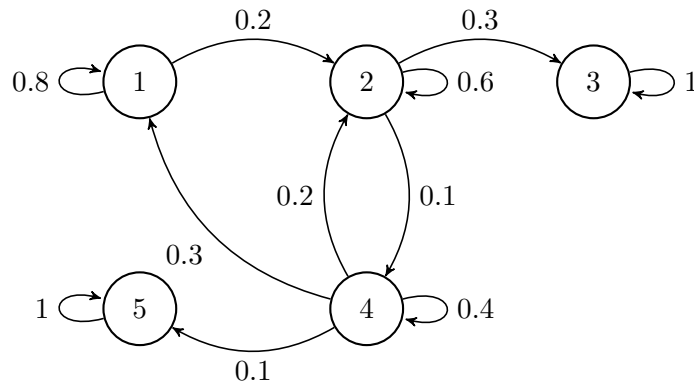


- Defina, em Matlab, a matriz de transição de estados T assumindo $p = 0,4$ e $q = 0,6$.
 - Assuma que o sistema se encontra inicialmente no estado A. Qual a probabilidade de estar em cada estado ao fim de 5 transições? E de 10 transições? E de 100 transições? E de 200 transições?
 - Determine as probabilidades limite de cada estado. Compare estes valores com os obtidos na alínea anterior. O que conclui?
5. Considere que o tempo em cada dia é genericamente classificado num de 3 estados – sol, nuvens e chuva – e que o tempo num determinado dia apenas depende do tempo no dia anterior. Assuma que estamos no primeiro dia de janeiro e que as probabilidades de transição de estados são as da tabela seguinte.

dia $n \setminus$ dia $n + 1 \rightarrow$	sol	nuvens	chuva
sol	0,7	0,2	0,1
nuvens	0,2	0,3	0,5
chuva	0,3	0,3	0,4

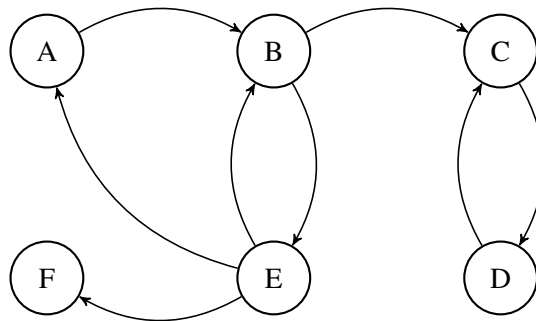
- Defina, em Matlab, a correspondente matriz de transição.
- Qual a probabilidade de estar sol no segundo dia e no terceiro dia de janeiro quando o primeiro dia é de sol?
- Qual a probabilidade de não chover nem no segundo dia nem no terceiro dia de janeiro quando o primeiro dia é de sol?
- Assumindo que o primeiro dia é de sol, determine o número médio de dias de sol, de nuvens e de chuva que se espera ter em todo o mês de janeiro.
- Assumindo que o primeiro dia é de chuva, determine o número médio de dias de sol, de nuvens e de chuva que se espera ter em todo o mês de janeiro. Compare estes resultados com os da alínea anterior. O que conclui?
- Considere uma pessoa com reumatismo crónico que tem dores reumáticas com probabilidades de 10%, 30% e 50% quando os dias são de sol, de nuvens ou de chuva, respetivamente. Qual o número esperado de dias que a pessoa vai sofrer de dores reumáticas em janeiro quando o primeiro dia é de sol? E quando o primeiro dia é de chuva?

6. Considere a cadeia de Markov com o diagrama de transição de estados seguinte:



- Defina em Matlab a matriz de transição de estados T , com T_{ij} sendo a probabilidade de ir do estado j para o estado i num único passo.
- Faça um gráfico com a probabilidade de, começando no estado 1, estar no estado 2 ao fim de n passos, com n a variar de 1 até 100. Justifique o que observa.
- Faça um gráfico com a probabilidade de, começando no estado 1, estar no estado 3 ao fim de n passos. Na mesma figura, faça um segundo gráfico com a probabilidade de, começando no estado 1, estar no estado 5 ao fim de n passos. Em ambos os casos, considere n a variar de 1 até 100. Justifique o que observa.
- Determine a matriz Q .
- Determine a matriz fundamental F .
- Qual a média (valor esperado) do número de passos antes da absorção começando no estado 1? E começando no estado 2? E se começando no estado 4?
- Começando no estado 1, qual é a probabilidade de absorção do estado 3? E do estado 5? Verifique a coerência destes valores com o que observou na alínea 6 (c).

7. Considere o conjunto de páginas Web e respetivas hiperligações entre si dado pelo diagrama seguinte:



- Usando a matriz H das hiperligações, obtenha a estimativa do pagerank de cada página ao fim de 10 iterações. Relembre que deve considerar (i) a mesma probabilidade de transição de cada página para todas as páginas seguintes possíveis e (ii) a probabilidade da página inicial deve ser igual para todas as páginas. Qual/quais a(s) página(s) com maior pagerank e qual o seu valor?
- Identifique a "spider trap" e o "dead-end" contidos neste conjunto de páginas.
- Altere a matriz H para resolver apenas o problema do "dead-end" e recalcule o pagerank de cada página novamente em 10 iterações.
- Resolva agora ambos os problemas e recalcule o pagerank de cada página novamente em 10 iterações (assuma $\beta = 0,8$).
- Calcule agora o pagerank de cada página considerando um número mínimo de iterações que garanta que nenhum valor muda mais do que 10^{-4} em 2 iterações consecutivas. Quantas iterações são necessárias? Compare os valores de pagerank obtidos com os da alínea anterior. O que conclui?

3.2 Exercícios suplementares¹

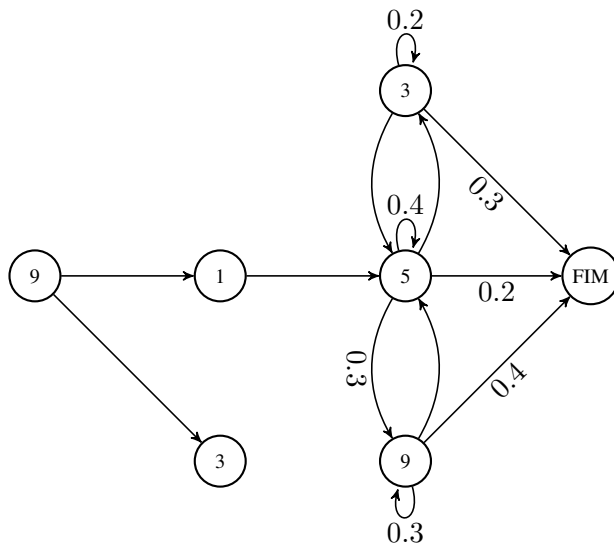
1. Relativamente ao problema 6 da secção anterior:

- Confirme os valores calculados nas alíneas (f) e (g) através de simulação (faça a média de várias simulações). Use o código Matlab na última página como base para criar a suas simulações.
- Guarde numa variável do tipo cell array do Matlab 10000 sequências que começam num dos estados não absorventes (escolhido com igual probabilidade). Calcule os comprimentos mínimo e máximo das sequências geradas.

2. Considere sequências de dígitos ímpares de 1 a 9 com as seguintes propriedades:

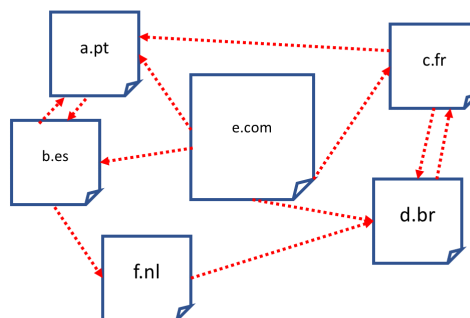
- todas as sequências começam por 9;
- na segunda posição da sequência podemos ter 1 ou 3, com igual probabilidade;
- na terceira posição e seguintes podem aparecer os dígitos 3, 5 e 9;
- os inícios de sequência 913, 915 e 919 são equiprováveis;
- da segunda para a terceira posições a transição de 3 para 5 tem probabilidade 0.6 e não é possível a transição de 3 para 3;
- da terceira posição em diante as probabilidades são as assinaladas na figura abaixo.

- Complete o diagrama de estados abaixo (incompleto!...) relativo à cadeia de Markov correspondente indicando todos os estados e transições possíveis segundo o enunciado. Crie em Matlab a matriz de transição de estados (T) correspondente.



- Qual a probabilidade de uma sequência de 7 dígitos iniciada por 91 e terminada em 9 ?
- Qual a probabilidade de sequências de 7 dígitos iniciadas por 91 ?

3. Considere as páginas web com as ligações apresentadas na figura:



¹Os Alguns dos exercícios desta secção foram utilizados em exames práticos de anos anteriores

- (a) Usando a matriz da Google A , com $\beta = 0.85$, obtenha a estimativa do pagerank de cada página usando um método iterativo, repetindo o processo iterativo de cálculo até que a diferença de pagerank entre duas iterações não exceda 0.01 em valor absoluto para todas as páginas. As colunas e linhas da matriz devem seguir a ordem alfabética dos nomes das páginas.
- Quais as páginas com o menor e maior pagerank e qual o seu valor ?
- (b) Confirme os valores de pagerank obtidos usando um processo não iterativo.

Anexo

```
% a state transition matrix example
H = [0.5 0.5 0;
      0.5 0.4 0;
      0   0.1 1];

% how to use crawl()
state = crawl(H, 1, 3)
% Note:
% if the Markov chain has more than 1 absorbing state the third argument
% must be a vector with the absorbing states
% for example in a chain with 5 states and considering that
% states 4 and 5 are absorbing states:
% state = crawl(H5, 1, [4 5])
%
% Random walk on the Markov chain
% Inputs:
% H - state transition matrix
% first - initial state
% last - terminal or absorbing state
function [state] = crawl(H, first, last)
    % the sequence of states will be saved in the vector "state"
    % initially, the vector contains only the initial state:
    state = [first];
    % keep moving from state to state until state "last" is reached:
    while (1)
        state(end+1) = nextState(H, state(end));
        if ismember(state(end), last) % verifies if state(end) is absorbing
            break;
        end
    end
end

% Returning the next state
% Inputs:
% H - state transition matrix
% currentState - current state
function state = nextState(H, currentState)
    % find the probabilities of reaching all states starting at the current one:
    probVector = H(:,currentState)'; % probVector is a row vector
    n = length(probVector); % n is the number of states
    % generate the next state randomly according to probabilities probVector:
    state = discrete_rnd(1:n, probVector);
end

% Generate randomly the next state.
% Inputs:
% states = vector with state values
% probVector = probability vector
function state = discrete_rnd(states, probVector)
    U=rand();
    i = 1 + sum(U > cumsum(probVector));
    state= states(i);
end
```