Teoria da Informação — Exercícios

Miguel Barão (mjsb@uevora.pt)

November 24, 2017

1 Álgebra linear (revisão)

1. Realize os seguintes produtos matriciais:

$$\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}, \qquad \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}5 & 6\\7 & 8\end{bmatrix}, \qquad \begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}3\\4\end{bmatrix}, \qquad \begin{bmatrix}1\\2\end{bmatrix}\begin{bmatrix}3 & 4\end{bmatrix},$$

2. Calcule os seguintes determinantes:

$$\det\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}, \qquad \det\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}, \qquad \det\begin{bmatrix}0 & 2 & 0\\1 & 0 & 2\\1 & -1 & 3\end{bmatrix}$$

3. Calcule os valores e vectores próprios das seguintes matrizes:

$$\mathbf{A} = \begin{bmatrix}11 & -4\\8 & -1\end{bmatrix}, \qquad \mathbf{B} = \frac{1}{3}\begin{bmatrix}5 & -4\\4 & -5\end{bmatrix}$$

4. Calcule as soluções dos seguintes problemas

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}5\\6\end{bmatrix}, \qquad \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}5\\10\end{bmatrix}, \qquad \begin{bmatrix}1 & 2\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}5\\5\end{bmatrix}$$

5. Confira os resultados das alíneas anteriores com o interpretador ipython (requer python>=3.5 e numpy):

```
from numpy import array, vstack, hstack, eye, zeros, ones
from numpy.linalg import det, eig, solve

A = array([[1,2],[3,4]])  # array bidimensional: matriz
B = array([5,6])          # array unidimensional: vector linha ou coluna
A[i,j]                    # elemento na linha i, coluna j
A[i,:]                    # array unidimensional com a linha i
A[:,j]                    # array unidimensional com a coluna j
A.T                       # matriz transposta
A + A                     # soma matricial
A @ B                     # produto matricial (B interpretado como coluna)
A * A                     # multiplicação elemento-a-elemento (não é o produto matricial)
det(A)                    # determinante
val, vec = eig(A)         # valores e vectores proprios
val[0]                    # primeiro valor proprio
vec[:,0]                  # primeiro vector proprio (coluna 0)
x = solve(A, B)           # solucao do sistema de equacoes lineares Ax = B
vstack((A,B))             # junta A e B na vertical, fica matriz 3x2
hstack((A,B))             # junta A e B na horizontal, fica matriz 2x3
sum(A, 0)                 # soma ao longo do indice 0
I = eye(3)                # matriz identidade 3x3
Z = zeros((2,3))          # matriz de zeros 2x3
O = ones((2,3))           # matriz de uns 2x3
```

2 Probabilidades (revisão)

1. Considere duas variáveis aleatórias X e Y que representam o estado do tempo, chuva e Sol respectivamente. A função de probabilidade conjunta $p(x, y)$ tem as probabilidades indicadas na seguinte tabela:

$X \backslash Y$	Sol	Encoberto
Chove	0.05	0.3
Não chove	0.5	0.15

   (a) Indique que símbolos compõem os alfabetos $\mathcal{X}$ e $\mathcal{Y}$.

   (b) Calcule as probabilidades marginais $p(x)$ e $p(y)$.

   (c) Calcule as probabilidades condicionais $p(x|y)$ e $p(y|x)$.

   (d) Verifique que

$$\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p(x, y) = 1, \qquad \sum_{x \in \mathcal{X}} p(x) = 1, \qquad \sum_{y \in \mathcal{Y}} p(y) = 1, \qquad \sum_{x \in \mathcal{X}} p(x|y) = 1, \qquad \sum_{y \in \mathcal{Y}} p(y|x) = 1.$$

   (e) Repita os cálculos das perguntas anteriores usando ipython.

```
from numpy import array, sum

# probabilidades conjuntas. x no indice 0, y no indice 1
Pxy = array([[0.05, 0.3], [0.5, 0.15]])

# marginais
Py = sum(Pxy, 0)                # soma sobre todos os valores de x
Px = sum(Pxy, 1)                # soma sobre todos os valores de y

# condicionais. cada coluna é uma distribuicao de probabilidade
Px_dado_y = Pxy / Py            # usa "broadcasting"
Py_dado_x = Pxy.T / Pxy         # usa "broadcasting"
```

3 Análise (revisão)

1. Prove que

   (a) $\sum_{n=0}^{+\infty} r^n = \frac{1}{1-r}$, para $0 < r < 1$.

   (b) $\sum_{n=1}^{+\infty} n r^n = r/(1-r)^2$, para $0 < r < 1$.

2. Calcule o resultado das seguintes expressões:

   (a) $\displaystyle\sum_{n=1}^{+\infty} r^n$, com $0 < r < 1$

   (b) $\displaystyle\sum_{n=1}^{N} r^n$, com $0 < r < 1$

   (c) $\displaystyle\frac{\mathrm{d}}{\mathrm{d}t}\left(\sin(2t) + 2^{-t} + \log_2(t) + t^2\right)$

   (d) $\displaystyle\int_0^1 t^2 \mathrm{d}t$

   (e) $\displaystyle\int_0^{+\infty} e^{-2t} \mathrm{d}t$

   (f) $\displaystyle\lim_{x \to +\infty} \frac{2x+1}{x}$

   (g) $\displaystyle\lim_{x \to 0} x \log x$

3. Confira os resultados anteriores usando ipython (requer sympy).

```
from sympy import symbols, Sum, diff, integrate, limit

t, x, r = symbols('t, x, r')          # define variaveis simbolicas

Sum(0.5**t, (t, 0, +oo)).doit()       # serie geometrica de razao 0.5
diff(log(t, 2) + t**2, t)             # derivada em ordem a t
integrate(exp(-2*x), (x, 0, +oo))     # integral de 0 a +infinito
limit(x*log(x,2), x, 0)               # limite quando x->0
```

4 Entropia

1. Considere as variáveis aleatórias X e Y com probabilidades conjuntas indicada na tabela seguinte:

$X\backslash Y$	S	N
S	0.05	0.3
N	0.5	0.15

   (a) Calcule as entropias $H(X)$ e $H(Y)$.
   (b) Calcule a entropia conjunta $H(X,Y)$.
   (c) Confirme os resultados usando ipython.
   (d) Considere uma nova variável $Z = X$. Qual a distribuição de probabilidade conjunta $p(x,z)$?
   (e) Calcule a entropia conjunta $H(X,Z)$. Compare com $H(X)$ e $H(Z)$.
   (f) Calcule a entropia condicional $H(X|Z)$. Interprete o resultado.
   (g) Considere uma nova variável W, independente de X, e com probabilidades $(0.1,\ 0.9)$. Qual a distribuição de probabilidade conjunta $p(x,w)$?
   (h) Calcule a entropia conjunta $H(X,W)$. Compare com $H(X) + H(W)$.
   (i) Calcule a entropia condicional $H(X|W)$. Interprete o resultado.
   (j) Calcule $I(X;Y)$, $I(X;Z)$ e $I(X;W)$. Interprete o resultado.

2. Considere duas variáveis aleatórias independentes X e Y com funções de probabilidade

$$p(x) = \begin{bmatrix} 0.1 \\ 0.3 \\ 0.2 \\ 0.05 \\ 0.35 \end{bmatrix}, \qquad p(y) = \begin{bmatrix} 0.25 \\ 0.3 \\ 0.05 \\ 0.4 \end{bmatrix}. \tag{1}$$

   (a) Calcule a distribuição de probabilidade conjunta $p(x,y)$.
   (b) Calcule as probabilidades condicionais $p(x|y)$ e $p(y|x)$.
   (c) Calcule as entropias $H(X)$, $H(Y)$ e $H(X,Y)$.

3. Considere uma fonte em que, para gerar um símbolo, é realizada uma sequência de lançamentos de uma moeda. O número de vezes em que sai cara consecutivamente desde o primeiro lançamento é o símbolo gerado.
   (a) Qual o alfabeto desta fonte?
   (b) Determine a distribuição de probabilidade $p(x)$.
   (c) Verifique que a soma das probabilidades é 1.
   (d) Calcule a entropia $H(X)$ da fonte.

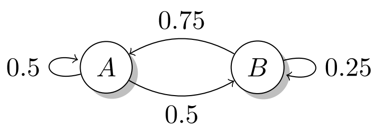
4. Considere um tabuleiro de xadrez de dimensão infinita onde apenas existe um rei que se move aleatoriamente (o rei pode mover-se uma casa em qualquer direcção). Pretende-se determinar a entropia no movimento do rei. Indique qual o alfabeto que deve ser considerado, as probabilidades $p(x)$ e a entropia $H(X)$.

5. Considere um jogo de sorte jogado com dois dados. Em cada jogada é somada a pontuação obtida pelo lançamento dos dois dados. Calcule a entropia $H(Z)$ correspondente à soma $Z = X + Y$, onde X e Y são as pontuações dos dados.

6. Sabe-se que uma certa doença afecta 1% da população. Para diagnosticar essa doença pode ser efectuado um teste que não é completamente fiável. Sabe-se que o teste dá 5% de falsos positivos e 10% de falsos negativos.
   (a) Calcule a probabilidade de estar doente sabendo que um teste deu positivo (use o teorema de Bayes).
   (b) Calcule a probabilidade de estar doente sabendo que dois testes deram positivos.
   (c) Calcule a entropia da variável X (estar ou não estar doente) antes dos testes, depois do primeiro teste, e finalmente depois dos dois testes.
   (d) Calcule a informação mútua entre X e o resultado do teste Y.

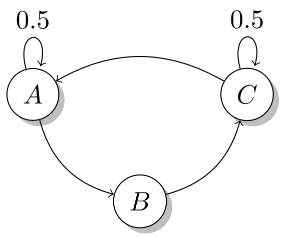
5 Cadeias de Markov

7. Considere o processo estocástico descrito pela figura seguinte.



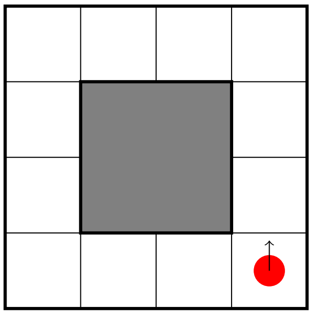
   (a) Este processo é uma cadeia de Markov? (justifique)
   (b) Escreva a matriz das probabilidades de transição correspondente.
   (c) Sabendo que o estado inicial é $X_0 = A$, calcule $\Pr\{X_t = A\}$ para $t = 1, \ldots, 10$. (pode usar o computador).
   (d) Determine a distribuição estacionária μ.
   (e) Assumindo regime estacionário, calcule a entropia condicional $H(X_t|X_{t-1})$.
   (f) Partindo da condição inicial $x_0 = A$, determine a evolução da distribuição de probabilidade dos estados ao longo do tempo. *I.e.*, calcule $p(x_0)$, $p(x_1)$, $p(x_2)$, $p(x_n)$.

8. Considere o processo estocástico descrito pela cadeia de Markov da figura seguinte:



   (a) Verifique se a cadeia de Markov é irredutível e aperiódica.
   (b) Calcule o ritmo de entropia $H'(\mathcal{X})$.

9. Considere um robot móvel que percorre um corredor fechado (ver figura abaixo). Em cada segundo o robot pode avançar ou recuar com probabilidades 0.9 e 0.1, respectivamente. Sabendo que inicialmente o robot está localizado no canto inferior direito, escreva uma função em octave/python que actulize o estado de conhecimento (distribuição de probabilidade) acerca da localização do robot ao longo de um intervalo de 10 minutos.



6 Códigos

1. Considere uma fonte com alfabeto e probabilidades indicadas na tabela seguinte:

x	$p(x)$	$C_1(x)$	$C_2(x)$	$C_3(x)$	$C_4(x)$	$C_5(x)$
A	0.3	0	00	110	110	00
B	0.1	1	11	11	0	111
C	0.2	10	00	001	111	110
D	0.4	11	10	00	10	010

   (a) Classifique cada um dos códigos em códigos singulares, não-singulares, univocamente descodificáveis e instantâneos.

   (b) Calcule o comprimento médio de cada código.

2. Considere uma fonte descrita pelas seguintes probabilidades:

x	$p(x)$
A	0.1
B	0.4
C	0.5

   (a) Construa um código de Shannon para esta fonte.

   (b) Calcule o comprimento médio $L(C)$. Compare com a entropia da fonte $H(X)$.

   (c) Agrupe os símbolos aos pares (x_i, x_{i+1}), e repita as duas alíneas anteriores.

   (d) Qual o comprimento médio por símbolo original? A compressão obtida é melhor ou pior que a anterior?

3. Mostre que o comprimento médio de um código de Shannon satisfaz a desigualdade

$$L(C) < H(X) + 1.$$

4. Considere uma fonte com alfabeto $\mathcal{X} = \{0, 1\}$, e probabilidades $[0.1, 0.9]$, respectivamente.

   (a) Desenhe códigos de Huffman para grupos de 1, 2 e 3 símbolos.

   (b) Calcule a entropia $H(X)$ da fonte e os comprimentos médios $L(C^1)$, $L(C^2)/2$ e $L(C^3)/3$ por símbolo da fonte. Comente.

5. Pretende-se comprimir a string ASCII "AABACCABCCC".

   (a) Codifique a string usando o código de Huffman adaptativo.

   (b) Escreva o código obtido em binário, incluindo os símbolos ASCII. (`A=0x41, B=0x42, C=0x43`).

   (c) Descodifique a string binária anterior e compare com a mensagem emitida pela fonte.

   (d) Compare o comprimento (medido em bits) da string original com a obtida depois da compressão.

6. Considere um ficheiro com comprimento de N bytes. Cada byte é considerado como um símbolo do alfabeto. Suponha que um ficheiro de texto ASCII é comprimido usando o código de Huffman Adaptativo, obtendo-se a seguinte sequência de bits:

   010000100011000010001101110111010

   (a) Quantos símbolos tem o alfabeto?

   (b) Descomprima e descodifique a sequência de bits acima de modo a obter o texto original.

7. Será que, usando o código de Huffman adaptativo, se obtém sempre um ficheiro mais pequeno que o original?

8. Será que existe algum esquema de compressão para o qual se consegue sempre comprimir os dados? (*i.e.*, dado um ficheiro original, a aplicação sucessiva de compressão sem perdas resulta sucessivamente em ficheiros mais pequenos.)

9. Considere uma fonte com alfabeto $\mathcal{X} = \{0, 1\}$ e probabilidades $[0.2,\ 0.8]$.

   (a) Desenhe um código Shannon-Fano-Elias para grupos de dois símbolos. Calcule o comprimento médio por símbolo e compare com a entropia.

   (b) Suponha que foi gerada a sequência de símbolos "1101110111111110". Comprima esta mensagem usando:

      i. Shannon-Fano-Elias (obtido anteriormente)

      ii. Codificação aritmética

10. Considere a string "AABAABBBAAA". Codifique a string usando:

    (a) LZ77, com buffers de comprimento 8.

    (b) LZ78

    (c) LZW

11. Considere uma fonte binária com alfabeto $\mathcal{X} = \{0, 1\}$, descrita por uma cadeia de Markov em que a probabilidade de dois símbolos iguais se sucederem é de 0.75, i.e. $p(X_{t+1} = a | X_t = a) = 0.75$, $a \in \mathcal{X}$.

    (a) Desenhe a cadeia de Markov, indicando todas as probabilidades de transição. Escreva a matriz de transição P respectiva.

    (b) Calcule a distribuição estacionária μ para este processo.

    (c) Calcule a entropia condicional $H(X_{t+1}|X_t)$.

    (d) Supondo que o estado no instante $t = 0$ é $x_0 = 1$, qual a probabilidade de no instante $t = 3$ o estado ser $x_3 = 1$?

    (e) Desenhe um código de Huffman para blocos de 3 bits desta fonte. Calcule o comprimento médio $L(C)$ e compare com o ritmo de entropia $H'(\mathcal{X})$.

7 Capacidade do canal

1. Considere um canal binário simétrico onde a probabilidade de um bit ser corrompido é $p_e = 0.1$.

   (a) Desenho o grafo correspondente indicando todas as probabilidades de transição.

   (b) Escreva a matriz de transição correspondente.

   (c) Calcule a capacidade do canal.

2. Um canal binário simétrico funciona de acordo com a equação booleana $y = x \oplus 1$, onde $\oplus$ representa a operação XOR (ou exclusivo). Qual é a capacidade deste canal?

3. Um canal binário simétrico funciona de acordo com a equação booleana $y = x \oplus e$, onde $\Pr\{E = 1\} = 0.9$. Qual é a capacidade deste canal?

4. Considere um *erasure channel* onde a probabilidade de um bit ser eliminado é $p_e = 0.2$. Calcule a capacidade deste canal.

5. Considere um canal binário simétrico com $p_e = 0.1$. Suponha que os símbolos x à entrada do canal têm as probabilidades indicadas na tabela seguinte:

x	$p(x)$
0	0.3
1	0.7

Pretende-se calcular a distribuição de probabilidade do símbolo x enviado, dado o símbolo recebido y. Isto é, pretende-se calcular as probabilidades condicionais $p(x|y)$. *(Sugestão: lei de Bayes)*

6. Uma cadeia de Markov com probabilidades de transição

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \tag{2}$$

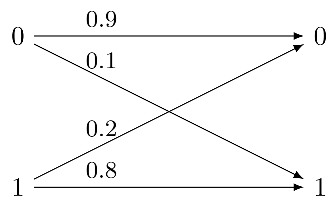
é observada por intermédio de um canal binário simétrico com $p_e = 0.8$. Pretende-se estimar qual o estado da cadeia de Markov ao longo do tempo à medida que são obtidas observações y. Admita conhecido o estado inicial $x_0 = 0$. Calcule a distribuição de probabilidade do estado nos instantes $t = 1, 2, 3$, sabendo que foi observada a sequência $y_{1:3} = [0, 1, 1]$. Assuma para cada instante de tempo t não é conhecido o futuro (i.e., em $t = 2$ não se conhece y_3).

Este problema é conhecido como cadeias de Markov escondidas, ou HMM (Hidden Markov Models).

7. Supondo que é usado o código de Hamming (7,4), faça a correcção das seguintes palávras recebidas:
   (a) 0001100
   (b) 1010101

8. Suponha que o código de Hamming $(7, 4)$ é usado para transmitir por um canal binário simétrico com $p_e = 0.1$. Calcule a probabilidade de erro na descodificação de uma palavra de código. A probabilidade de erro é igual para todas as palavras de código?

9. Suponha que o código de Hamming $(7, 4)$ é usado para transmitir por um canal binário



Calcule a probabilidade de erro na descodificação de cada palavra de código.

8 Variáveis Contínuas

1. Calcule a entropia diferencial $h(X)$ para as funções densidade de probabilidade seguintes:
   (a) Uniforme: $X \sim \text{unif}(1 - \varepsilon, 1 + \varepsilon)$
   (b) Mistura de uniformes: $X \sim a\ \text{unif}(1 - \varepsilon, 1 + \varepsilon) + (1 - a)\ \text{unif}(2 - \varepsilon, 2 + \varepsilon)$
   (c) Gaussiana: $X \sim \mathcal{N}(\mu, \sigma^2)$, onde

$$\mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

2. Considere um canal Gaussiano com entrada $X \in [0, 5]$, saída Y, e ruído aditivo Gaussiano $Z \sim \mathcal{N}(0, 1)$. O canal Gaussiano vai ser usado para transmitir símbolos de um alfabeto binário (valores lógicos True/False).
   (a) Como deve ser feita a codificação dos valores lógicos?
   (b) Como deve ser feita a descodificação da saída Y?

(c) Calcule a probabilidade de erro na descodificação e calcule o canal binário equivalente.

(d) Calcule a capacidade do canal.

(e) Generalize a capacidade do canal para $X \in [-X_{\max}, X_{\max}]$ e $Z \sim \mathcal{N}(0, \sigma^2)$.

3. Considere um canal Gaussiano com entrada X, saída Y e ruído Gaussiano $Z \sim \mathcal{N}(0, 1)$. A entrada X tem "potência" limitada a $P = x^2 \leq 4$.

   (a) O canal Gaussiano vai ser usado para transmitir símbolos de um alfabeto trenário $X^\Delta \in \{0, 1, 2\}$, pretendendo-se obter à saídas símbolos $Y^\Delta \in \{0, 1, 2\}$. Determine de que modo os símbolos X^Δ podem ser codificados e os símbolos Y^Δ descodificados.

   (b) Calcule as probabilidades de erro na descodificação para cada símbolo discreto transmitido.