

---

## Aprendizagem Automática

FICHA LABORATORIAL N. 3

ENUNCIADO

Nome: Ana Sofia Preto Oliveira

Número: A39275

---

NOTA: Fixa de respostas múltiplas. Só uma única resposta em cada alínea está correta. Cada alínea vale 2 valores. Respostas erradas descontam 0.5 valores.

---

ATENÇÃO: Considere para esta ficha o conjunto de dígitos manuscritos do ficheiro `MNISTsmall.p`. Não modifique os dados (ex: não remova a média ao conjunto de dados).

No texto, quando são referidos os exemplos do conjunto (colunas da matriz  $X$ ) bem como as dimensões dos dígitos (linhas da matriz  $X$ ), estes estão indexados a começar por um e não por zero. Por exemplo, ao mencionar a 3ª dimensão do 20º dígito, o valor corresponde à entrada da matriz de dados  $X[2, 19]$ .

---

1. Considere os 381 primeiros exemplos do conjunto de teste do dígito “0”.

- (a) Para esta alínea arredonde os valores pedidos a 0 casas decimais. Considere o 173º dígito da matriz de dados. Pretende-se projetar e reconstruir este dígito usando um número mínimo de componentes principais tal que os dados projetados tenham 75% da variância total dos dados originais. O erro absoluto médio entre o dígito original e o reconstruído é:

i. 14

ii. 7

iii. 1

iv. 6

- (b) Para esta alínea, arredonde os valores pedidos a 2 casas decimais.

- i. Existem 299 dimensões dos dados que não variam.  
ii. O coeficiente de correlação entre a dimensão 218 e 412 é 0.21.  
iii. Todas as respostas anteriores.

iv. Nenhuma das respostas anteriores.

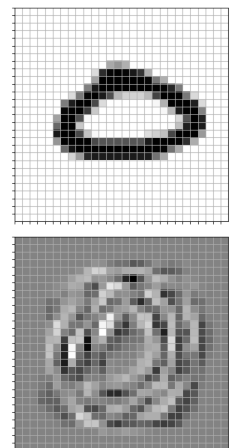
- (c) . Pretende-se visualizar os dígitos deste conjunto e os vetores próprios da matriz de covariância dos dados.

- i. A figura de cima é a imagem invertida do 3º dígito do conjunto.

- ii. A figura de baixo é a imagem (ou a imagem invertida) do 2º vetor próprio da matriz de covariância dos dados.

- iii. Todas as respostas anteriores.

iv. Nenhuma das respostas anteriores.



- (d) Considere o 178º dígito da matriz de dados. Considere ainda que se pretende projetar e reconstruir este dígito usando as 13 primeiras componentes principais. O erro quadrático médio entre o dígito original e o reconstruído é (arredondando a 0 casas decimais):
- i. 1243      ii. 1103      iii. 1573      iv. 995
- (e) Considere que se estima a matriz de covariância deste conjunto.
- i. A dimensão da matriz é  $784 \times 784$ .  
ii. O determinante da matriz de covariância não é igual a 0.  
iii. Todas as respostas anteriores.  
iv. Nenhuma das respostas anteriores.
- (f) Pretende-se projetar os dígitos nas suas componentes principais. O número mínimo de componentes principais para que os dados projetados tenham acima de 90% da variância total dos dados originais, é:
- i. 46      ii. 49      iii. 59      iv. 35
- (g) O número de pixels com valor igual a zero para todos os exemplos deste conjunto é:
- i. 387      ii. 310      iii. 437      iv. 320
- (h) Para esta alínea, arredonde os valores pedidos a 0 casas decimais.
- i. A norma do vetor de média dos dados é 1917.  
ii. O desvio padrão da dimensão 158 dos dados é 3.  
iii. Todas as respostas anteriores.  
iv. Nenhuma das respostas anteriores.
- (i) O número de valores próprios superiores a  $10^{-10}$  da matriz de covariância dos dados é:
- i. 330      ii. 285      iii. 406      iv. 380
- (j) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Considere que os 225º, 226º, 274º e 292º dígitos são representados pelos vetores  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ ,  $\mathbf{x}_3$ , e  $\mathbf{x}_4$ .
- i. O produto interno entre os vetores  $\mathbf{x}_1$  e  $\mathbf{x}_2$  é 3591347.  
ii. O produto interno entre os vetores  $\mathbf{x}_3$  e  $\mathbf{x}_4$  é 2975543.  
iii. Todas as respostas anteriores.  
iv. Nenhuma das respostas anteriores.