problem 2 : ISL book (chapter 2)

2) a] Inference → which factors affect CEO sallary.

n = 500 (the number of observations)

p = profit, number of employees, industry (features)

regression.

b] prediction → wish to know wether a new product will be a success or f.

n → 20 similar products

p → price charged, marketing budget, competition pric + 10 other var.

classification → success or failaur

c] prediction → predicting the % change in the USD/Euro exchange rate.

n → weekly data for all of 2012

p → % change in the US market.

→ % change in the British market.

→ % change in the German market

output → % chang in the USD/Euro exchange rate

⇓

regression

Classification examples

example 1

4) a] we are considering a sequence of DNA and wish to know wether it will be a promoter or not. => prediction

⇓

(discrete) ← response (classification)

n = we collect 1000 known promoter sequence in human genome

→ find the % similarity between these sequences. => predictors (feature)

example 2 → our goal is investigating which factors effect lung cancer

⇒ Inference.

n = 100 lung cancer patients

p = smoking, aging

→ find a factor has effect on creating lung cancer or not

⇓

classification

example 3 → digit classification

⇓

we have a digit number & our goal is detecting this number

↳ we can pixel the picture

⇓

then count black & white houses.

/ /

regression

b) example 1 → evaluation the PSA level in prostate cancer

input features: age, prostate weight, ...

example 2 → prediction the % of death cause to Breast cancer in iran

input features → Nutrition, age

example 3 → the % of different factors can affect diabetes

input features → inheritance, amount of sugar consumption

clustering

ex.1

c) we have a gene expression data but the data is completely unlable.

↳ we can cluster data based on similar gene expression ranges.

⇓

by unsupervised learning models.

ex.2 → in ex.1 we also can cluster each gene cluster to –

two groups based on high or low gene expression

⇓ this groups can show for example healthy & cancer groups.

ex.3 → find different cancer signatures mutations from different

sources can also be a kind of clustering

↳ for example → in lung cancer → features: age, smoking

(we can solve these type of problems by blind source separation)

Ecleadian distance

7) a)

obs.	x_1	$(x-x_1)^2$, $x=0$	x_2	$(x-x_2)^2$, $x=0$	x_3	$(x-x_3)^2$, $x=0$	$\sqrt{\sum_{j=1}^{3}(x_j-x_j[i])^2}$	Y
1	0	0	3	9	0	0	$\sqrt{9}=3$	R
2	2	4	0	0	0	0	$\sqrt{4}=2$	R
3	0	0	1	1	3	9	$\sqrt{10}$	R
4	0	0	1	1	2	4	$\sqrt{5}$	G
5	-1	1	0	0	1	1	$\sqrt{2}$	G
6	1	1	1	1	1	1	$\sqrt{3}$	R

b) $k=1$, $x_1=x_2=x_3=0$ (test point) → min of Eucleadian distance $=\sqrt{2}$

→ Since the output is classified, majority voting among neighbors is useful.

$\Rightarrow \hat{f}(x)=$ Red

c) $k=3$, $x_1=x_2=x_3=0$ → three min of Eucledian dis $= x[5], x[6], x[2]$

$\Rightarrow$ Like previous question, we can use majority voting → $\hat{f}(x)=\max(\text{Red}, \text{Green}, \text{Red}) =$ Red

HW 1

problem 3 : linear Algebra

a) $A = \begin{bmatrix} 1 & 2 \\ 5 & 3 \end{bmatrix}_{2\times2}$ $B = \begin{bmatrix} 2 & 5 \\ 7 & 3 \end{bmatrix}_{2\times2}$ $v = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

$C = A \times B \rightarrow C = AB \in R^{m\times p}$, $c_{ij} = \sum_{i=1}^{n} a_{ik} b_{kj}$

$\Rightarrow C = A \times B = \begin{bmatrix} 1 & 2 \\ 5 & 3 \end{bmatrix} \times \begin{bmatrix} 2 & 5 \\ 7 & 3 \end{bmatrix} = \begin{bmatrix} 2+14 & 5+6 \\ 10+21 & 25+9 \end{bmatrix} = \begin{bmatrix} 16 & 11 \\ 31 & 34 \end{bmatrix}$

$A^T \Rightarrow (A^T)_{ij} = A_{ji} \rightarrow A^T = \begin{bmatrix} 1 & 5 \\ 2 & 3 \end{bmatrix}$

$B^{-1} \rightarrow \begin{bmatrix} 2 & 5 \\ 7 & 3 \end{bmatrix}^{-1} = \frac{1}{2\times3 - 5\times7} \times \begin{bmatrix} 3 & -5 \\ -7 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{3}{29} & \frac{-5}{29} \\ -\frac{7}{29} & -\frac{2}{29} \end{bmatrix}$

$6 \quad 35$

-29

$|B| = 2\times3 - 5\times7 = -29$

$tr(B) = \sum_{i=1}^{n} B_{ii} = 2+3 = 5$

$v^T B v = \begin{bmatrix} -5 & 3 \end{bmatrix}_{1\times2} \begin{bmatrix} 2 & 5 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} -5 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 & -16 \end{bmatrix}_{1\times2} \begin{bmatrix} -5 \\ 3 \end{bmatrix}_{2\times1} = -55 + (-48) = -103$

$\|v\|_1 = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = \sum_{i=1}^{n} |x_i| = |-5| + |3| = 8$

$\|v\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{|-5|^2 + |3|^2} = \sqrt{34}$

b) singular matrix: a square matrix that does not have a matrix inverse $\Rightarrow$ determinant is zero.

(i) $A=\begin{bmatrix} K & 6 \\ 4 & 3 \end{bmatrix} \Rightarrow$ det(A) = 0 $\quad 3K - 24 = 0 \rightarrow \boxed{K=8}$

(ii) $B=\begin{bmatrix} 1 & 2 & -7 \\ -3 & 4 & K \\ -4 & 2 & 6 \end{bmatrix} \rightarrow |B| = 0 \rightarrow 1\times\begin{bmatrix} 4 & K \\ 2 & 6 \end{bmatrix} - 2\begin{bmatrix} -3 & K \\ -4 & 6 \end{bmatrix} - 7\begin{bmatrix} -3 & 4 \\ -4 & 2 \end{bmatrix} = 0$

$\rightarrow 24 - 2K - 2(-18 + 4K) - (-6 + 16) = 0$

$24 - 2K + 36 - 6K - 10 = 0$

$12 = 8K \rightarrow \boxed{K = \frac{12}{8}}$

c) the rank of a matrix

$A=\begin{bmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 1 & 5 & 11 \end{bmatrix} \rightarrow a\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + b\begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix} + c\begin{bmatrix} 3 \\ 9 \\ 11 \end{bmatrix} = 0$

3x3

$a + 2b + 3c = 0$

$3a + 6b + 9c = 0$

$a + 5b + 11c = 0$

the maximum rank of this matrix $\rightarrow$ 3

$\boxed{\text{rank}(A) = 2}$

because row 1 & 2 $\rightarrow$ is linealy dependent.

Scanned by CamScanner

d) $\beta \in R^p$, $C \in R^{p\times p}$ (a symmetric matrix)

$f(\beta)=\beta^T C\beta \rightarrow$ Calculate gradiant & Hessian matrix (respect to β)

$$f(\beta)=\sum_{i=1}^{p}\sum_{j=1}^{p} \beta_i C_{ij}\beta_j = \sum_{i=1}^{p}\left(C_{ii}\beta_i^2 + \sum_{j\neq i} \beta_i C_{ij}\beta_j\right)$$

$$\nabla f(\beta)=\begin{bmatrix} \frac{\partial}{\partial \beta_1} f(\beta) \\ \frac{\partial}{\partial \beta_2} f(\beta) \\ \vdots \\ \frac{\partial}{\partial \beta_p} f(\beta) \end{bmatrix} = \begin{bmatrix} 2\sum_{j=1}^{p} \beta_j C_{j1} + \sum_{j=1}^{p} \beta_j C_{1j} \\ \vdots \\ 2\sum_{i=1}^{p} \beta_j C_{jp} + \sum_{j=1}^{p} \beta_j C_{pj} \end{bmatrix} = \underbrace{\begin{bmatrix} 2\sum_{j=1}^{p} \beta_j C_{j1} \\ \vdots \\ 2\sum_{i=1}^{p} \beta_j C_{jp} \end{bmatrix}}_{\beta C^T} + \underbrace{\begin{bmatrix} \sum_{j=1}^{p} \beta_j C_{1j} \\ \vdots \\ \sum_{j=1}^{p} \beta_j C_{pj} \end{bmatrix}}_{\beta C} =$$

$$= (C^T + C)\beta \underset{C=\text{symmetric}}{=\!=} 2C\beta$$

e) positive definite matrix: $A \in R^{m\times m}$ is PD matrix if $x^T A x > 0$ $(x \neq 0)$
$(x \in R^m)$

$$A=\begin{bmatrix} 7 & 2 \\ 2 & 1 \end{bmatrix} \quad x=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 7 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \overset{M}{7x_1^2 + x_2^2 + 4x_1x_2}$$

$$\Downarrow$$

$$3x_1^2 + \underbrace{4x_1^2 + x_2^2 + 4x_1x_2}$$

$$\Downarrow$$

پس ماتریس A، PD است. $\Longleftarrow$ $3x_1^2 + \underbrace{(2x_1+x_2)^2}_{\text{این عبارت همیشه مثبت است.}} > 0$