HOME-WORK-2
M.I.S

Problem 2.8 =

$$P(x) = c_1 + c_2 x + \text{-----} \; c_n x^{n-1}$$

(a) = find an n-vector a for which, (where $\alpha < \beta$)

$$a^T c = \int_{\alpha}^{\beta} p(x)\, dx \quad \text{, always holds.}$$

$$a^T c = \int_{\alpha}^{\beta} \left(c_1 + c_2 x + \text{----} \; c_n x^{(n-1)} \right) dx$$

$$a^T c = \left[c_1 x + c_2 \frac{x^2}{2} + c_3 \frac{x^3}{3} \; \text{-----} \; \frac{c_n x^{(n-1)+1}}{(n-1)+1} \right]_{\alpha}^{\beta}$$

$$a^T c = \left[c_1 x + \frac{c_2 x^2}{2} + \frac{c_3 x^3}{3} \; \text{----} \; \frac{c_n x^n}{n} \right]_{\alpha}^{\beta}$$

$$a^T c = \left(c_1 \beta + \frac{c_2 \beta^2}{2} + \frac{c_3 \beta^3}{3} \; \text{-----} + \frac{c_n \beta^n}{n} \right)$$

$$- \left(c_1 \alpha + \frac{c_2 \alpha^2}{2} + \frac{c_3 \alpha^3}{3} \; \text{----} + \frac{c_n \alpha^n}{n} \right)$$

$$a^T c = c_1 (\beta - \alpha) + c_2 \left(\frac{\beta^2}{2} - \frac{\alpha^2}{2} \right) + c_3 \left(\frac{\beta^3}{3} - \frac{\alpha^3}{3} \right) \text{---} + c_n \left(\frac{\beta^n}{n} - \frac{\alpha^n}{n} \right)$$

$$a^T c = (\beta - \alpha) c_1 + \left(\frac{\beta^2}{2} - \frac{\alpha^2}{2} \right) c_2 + \left(\frac{\beta^3}{3} - \frac{\alpha^3}{3} \right) c_3 \text{----} + \left(\frac{\beta^n}{n} - \frac{\alpha^n}{n} \right) c_n$$

$$a^T c = \begin{bmatrix} (\beta - \alpha) \\ \left(\frac{\beta^2}{2} - \frac{\alpha^2}{2} \right) \\ \vdots \\ \left(\frac{\beta^n}{n} - \frac{\alpha^n}{n} \right) \end{bmatrix}^T \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

Hence n vector $a = \begin{bmatrix} (\beta-\alpha) \\ \left(\frac{\beta^2}{2}-\frac{\alpha^2}{2}\right) \\ \left(\frac{\beta^3}{3}-\frac{\alpha^3}{3}\right) \\ \vdots \\ \left(\frac{\beta^n}{n}-\frac{\alpha^n}{n}\right) \end{bmatrix}$

(b)= Let α be a number, find n-vector b for which

$$b^T c = p'(\alpha)$$

$$p(x) = c_1 + c_2 x + c_3 x^2 - - - - + c_n x^{n-1}$$

Then $p'(x) = \frac{d\,P(x)}{dx} = \frac{d}{dx}\left(c_1 + c_2 x + - - - - - + c_n x^{n-1}\right)$

$$p'(x) = \frac{d}{dx} c_1 + \frac{d}{dx} c_2 x + \frac{d}{dx} c_3 x^2 - - - + \frac{d}{dx} c_n x^{n-1}$$

$$p'(x) = 0 + c_2\left(\frac{d}{dx} x\right) + c_3\left(\frac{d}{dx} x^2\right) - - - + c_n\left(\frac{d}{dx} x^{n-1}\right)$$

$$p'(x) = 0 + c_2 + 2c_3 x - - - - + (n-1)\, c_n x^{(n-1)-1}$$

$$p'(x) = 0 + c_2 + 2c_3 x - - - - - - + (n-1)\, c_n x^{n-2}$$

$$p'(\alpha) = 0 + c_2 + 2c_3(\alpha) - - - - - + (n-1)\, c_n \alpha^{n-2}$$

$$p'(\alpha) = 0.c_1 + 1.c_2 + 2\alpha.c_3 - - - - + (n-1)\alpha^{n-2}.c_n$$

$$b^T c = p'(\alpha) = 0.c_1 + 1.c_2 + 2\alpha.c_3 - - - + (n-1)\alpha^{n-2}.c_n$$

Hence,

$$b^T C = \begin{bmatrix} 0 \\ 1 \\ 2\alpha \\ \vdots \\ (n-1)\alpha^{n-2} \end{bmatrix}^T \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix}$$

Hence n-vector $b = \begin{bmatrix} 0 \\ 1 \\ 2\alpha \\ \vdots \\ (n-1)\alpha^{n-2} \end{bmatrix}$

Problem 2.12 =

Regression Model to predict the profit total. when the product prices are changed. is given by -

$\hat{P} = \beta^T x + P$ where $\hat{P}$ = Predicted profit

P = profit with current Prices

$x_i = (P_i^{new} - P_i) / P_i$ (fractional change in Price of Product x_i)

we can write,

$$\hat{P} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_n \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + P$$

$$\hat{P} = (\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \ldots \beta_n x_n) + P$$

(a) = What does it mean if $(\beta_3 < 0)$

$$\hat{P} = (\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \; --- \; \beta_n x_n) + P$$

if $(\beta_3 < 0)$, That means value of regressional coefficient for Product x_3 is negative. So if the Product Price of x_3 is increased That will reduce the overall profit, and if the product price of x_3 is decreased from previous value then overall predicted profit will increase.

lets consider x_3 is the product for which prices are changed while other products have same prices as before. Then in that case fractional change for all products where prices are not changed becomes 0.

$$\hat{P} = (\beta_1 \cdot 0 + \beta_2 \cdot 0 + \beta_3 x_3 \; ---- \; + \beta_n \cdot 0) + P$$

$$\hat{P} = \beta_3 x_3 + P \quad \text{(Price of Product -3 is changed.}$$

Case.1 = $(\beta_3 < 0)$ and if x_3 increases i.e. the product price in increased from previous value – $(P_3^{new} > P_3)$. Then Value of $\hat{P}$ will decrease

Case.2 = $(\beta_3 < 0)$ and if x_3 decreases i.e. the product price is decreased from previous value $(P_3^{new} < P_3)$ Then value of $\hat{P}$ will increase.

Case.3 = if x_3 remains 0 (i.e. Product price is not changed, i.e. $(P_3^{new} = P_3)$ in that case it does not impact predicted profit.

we can say if β_3 is negative,

Then predicted profit $(\hat{P})$ is inversely proportional to fractional change (x_i) in Product-3 $\left(x_3 = \frac{P_3^{new} - P_3}{P_3} \right)$.

(b)= if I am given permission to change the price of one product by up to 1%, to increase total profit.

Then I will choose ith product for which the absolute value of regressional coefficient is maximum.

i.e. $|\beta_i|$ is maximum among all $(\beta_1, \beta_2, \beta_3 --- \beta_n)$ coefficients.

case 1 = if (β_i is positive) and ($|\beta_i|$ = maximum) Then I will Increase the price of product i by 1%. i.e P_i^{new} will be 1% more than P_i. In that case.

$$x_i = \left(\frac{P_i^{new} - P_i}{P_i} \right) \Rightarrow \text{Positive value}$$

Hence $\hat{P} = \beta_i x_i + P$

$(+ve)$ $(+ve)$ → will have +ve product and add maximum value to predicted profit.

case 2 = if (β_i is negative) and ($|\beta_i|$ = maximum) Then I will decrease the price of product i by 1%. i.e. P_i^{new} will be 1% less than P_i. In that case.

$$x_i = \left(\frac{P_i^{new} - P_i}{P_i} \right) \Rightarrow -\text{ve value}$$

Hence $\hat{P} = \beta_i x_i + P$

$(-ve)$ $(-ve)$ → will have +ve product and add maximum value to predicted profit.

(C) = If I am allowed to change the prices of two products each by upto 1% to maximize the profit

Then I will choose the product i and k such that the absolute values of regressional coefficients for product i and k are maximum values among all other coefficients.

i.e. $|\beta_i|$ and $|\beta_K|$ = are two largest values in the set of $(\beta_1, \beta_2 ---- \beta_n)$

the relation could be described as —

$$\underbrace{|\beta_1| \leq |\beta_2| \leq |\beta_3| ---- \leq |\beta_K| \leq |\beta_i|}_{\text{Two largest values.}}$$

Case 1 = if (β_i = positive, β_K = positive) Then I will increase the price of both products i and k by 1%.

[P_i^{new} = 1% more than P_i] [P_K^{new} = 1% more than P_K]

In that case ($\hat{P} = \beta_i x_i + \beta_K x_K + P$) will be maximized.

$\underbrace{+ve \quad +ve}_{+ve \text{ product}}$ $\underbrace{+ve \quad +ve}_{+ve \text{ product}}$

Case 2 = if (β_i = positive, β_K = negative) Then I will increase the price of product i by 1% and decrease the price of product k by 1%.

In that case $\hat{P} = \beta_i x_i + \beta_K x_K + P$ will maximize profit.

$\underbrace{+ve \quad +ve}_{+ve \text{ product}}$ $\underbrace{-ve \quad -ve}_{+ve \text{ product}}$

case.3 = if (β_i = negative) (β_K = positive) Then I will decrease the price of product i by 1%.

and Increase the price of product k by 1%.

In that case = $\hat{P} = \beta_i x_i + \beta_K x_K + P$ will maximize profit.

$$\underbrace{\underset{-ve}{\beta_i}\ \underset{-ve}{x_i}}_{+ve\ \text{product}} + \underbrace{\underset{+ve}{\beta_K}\ \underset{+ve}{x_K}}_{+ve\ \text{product}}$$

case.4 = if (β_i = negative) (β_K = negative) Then I will decrease the price of product i by 1%.

and decrease the price of product k by 1%.

In that case = $\hat{P} = \beta_i x_i + \beta_K x_K + P$ will maximize the profit.

$$\underbrace{\underset{-ve}{\beta_i}\ \underset{-ve}{x_i}}_{+ve\ \text{product}} + \underbrace{\underset{-ve}{\beta_K}\ \underset{-ve}{x_K}}_{+ve\ \text{product}}$$

Problem 3.5 ⇒

for an n-vector x. vector norm 1-norm $\|x\|_1$ and ∞-norm $\|x\|_\infty$ are defined as –

$$\|x\|_1 = |x_1| + |x_2| \cdots + |x_n|$$

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \cdots, |x_n|\}$$

show that 1-norm $\|x\|_1$ and ∞-norm $\|x\|_\infty$ satisfy the four norm properties.

verification for 1-norm ⇒

a) :- Non-negative Homogenity ⇒ $\boxed{\|\beta x\|_1 = |\beta| \, \|x\|_1}$

lets consider n-vector x ⇒ $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Then $\|x\|_1 = |x_1| + |x_2| \cdots |x_n| \longrightarrow ①$

now lets multiply vector x with scalar β ⇒

$$\beta x = \begin{bmatrix} \beta x_1 \\ \beta x_2 \\ \vdots \\ \beta x_n \end{bmatrix}$$

$$\|\beta x\|_1 = |\beta x_1| + |\beta x_2| + \cdots + |\beta x_n|$$

$$\|\beta x\|_1 = |\beta||x_1| + |\beta| |x_2| + \cdots + |\beta| |x_n|$$

$$\|\beta x\|_1 = |\beta| \left[|x_1| + |x_2| + \cdots + |x_n| \right]$$

$$\|\beta x\|_1 = |\beta| \, \|x\|_1 \quad \text{[By equation ①]}$$

(b) **Triangle inequality** $\boxed{\|x+y\|_1 \leq \|x\|_1 + \|y\|_2}$

lets consider two n-vectors, x and y.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

Then $X+Y = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \\ \vdots \\ x_n+y_n \end{bmatrix}$

now $\|x+y\|_1 = |x_1+y_1| + |x_2+y_2| - - - - + |x_n+y_n|$ → equation (1)

$\|x\|_1 = |x_1| + |x_2| + - - - - |x_n|$ → equation (2)

$\|y\|_1 = |y_1| + |y_2| + - - - |y_n|$ → equation (3)

By $\|x\|_1 + \|y\|_1$ → equation (2) + equation (3)

we get,

$\|x\|_1 + \|y\|_1 = [|x_1| + |x_2| + - - + |x_n|] + [|y_1| + |y_2| + - - |y_n|]$

$= (|x_1| + |y_1|) + (|x_2| + |y_2|) + - - - - + (|x_n| + |y_n|)$

→ equation (4)

By Properties of modulus operator we know that ⇒

$$\boxed{|a+b| \leq |a| + |b|}$$

for all values of real number a, b.

Hence we can say that ⇒

$$|x_1 + y_1| \leq |x_1| + |y_1|$$
$$|x_2 + y_2| \leq |x_2| + |y_2|$$
$$\vdots \qquad\qquad \vdots$$
$$|x_n + y_n| \leq |x_n| + |y_n|$$

(L.H.S.) (R.H.S.)

Summing the equations on left and right we get ⇒

$$|x_1+y_1| + |x_2+y_2| + \dots |x_n+y_n| \leq (|x_1| + |y_1|) + (|x_2| + |y_2|) + \dots + (|x_n| + |y_n|)$$

using previously derived equations ① and ④ we can write.

$$\boxed{\|x+y\|_1 \leq \|x\|_1 + \|y\|_1}$$ proved

(C) ⇒ Non-negtivity. $\boxed{\|x\|_1 \geq 0}$

n vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ Then $\|x\|_1 = |x_1| + |x_2| \dots + |x_n|$

↳ equation ①

we know that $\boxed{|a| \geq 0}$ always for all real values of a.

Then we can say that

$$|x_1| \geq 0$$
$$|x_2| \geq 0$$
$$|x_3| \geq 0$$
$$\vdots \qquad \vdots$$
$$|x_n| \geq 0$$

Summing Right and left side we get ⇒

$$|x_1| + |x_2| + |x_3| - - - - + |x_n| \geq 0$$

By using equation ① we can write =

$$\boxed{\|x\|_1 \geq 0}$$ Proved

(d) = Definiteness ⇒ $\boxed{\|x\|_1 = 0 \text{ only if } x = 0}$

n-vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ Then $\|x\|_1 = |x_1| + |x_2| + - - - + |x_n|$

↳ equation ①

we know that modulus of any real number is zero only when the number itself is zero.

$$\boxed{|a| = 0 \text{ when } a = 0}$$

Hence we can say,

$$|x_1| = 0 \quad \text{only if } x_1 = 0$$
$$|x_2| = 0 \quad \text{only if } x_2 = 0$$
$$\vdots \qquad \vdots$$
$$|x_n| = 0 \quad \text{only if } x_n = 0$$

Summing Right & left sides we get =

$$|x_1| + |x_2| - - + |x_n| = 0$$ [only if $x_1, x_2 - - - x_n$ all are zero]

$$= \|x\|_1 = 0$$

Hence $\|x\|_1 = 0$

only when $x_1, x_2 \text{ ---- } x_n$ all entries are zero.

if $x_1 = x_2 = x_3 \text{ --- } = x_n = 0$ Then x becomes a zero vector

i.e. $\boxed{x = 0}$

we can say that $\boxed{\|x\|_1 = 0}$ only if vector $x = 0$

Verification for ∞-norm ⇒

a) = Non-negative Homogeneity ⇒ $\boxed{\|\beta x\|_\infty = |\beta| \, \|x\|_\infty}$

lets consider n-vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

Then $\|x\|_\infty = \text{Max.}\{|x_1|, |x_2| \text{ --- } |x_n|\}$ ⟶ equation①

Multiplying vector x with scalar β. we get,

$$\beta x = \begin{bmatrix} \beta x_1 \\ \beta x_2 \\ \vdots \\ \beta x_n \end{bmatrix}$$

Then $\|\beta x\|_\infty = \text{Max.}\{|\beta x_1|, |\beta x_2|, \text{ --- } |\beta x_n|\}$

we know that $\boxed{|ab| = |a|\,|b|}$

Hence we can write =

$$\|\beta x\|_\infty = \text{Max.}\{|\beta|\,|x_1|, |\beta|\,|x_2| \text{ ---- } |\beta|\,|x_n|\}$$

since β is constant and multiplied to all the terms we can say that.

$$\|\beta x\|_\infty = \max\{|\beta||x_1|, |\beta||x_2| ---- |\beta||x_n|\}$$

$$\|\beta x\|_\infty = |\beta| \cdot \max\{|x_1|, |x_2| --- |x_n|\}$$

$$\boxed{\|\beta x\|_\infty = |\beta| \, \|x\|_\infty}$$ (from equation 1)

(b)= <u>Triangle Inequality</u> $\boxed{\|x+y\|_\infty \le \|x\|_\infty + \|y\|_\infty}$

lets consider two n-vectors x and y.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Then
$$x+y = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

$$\|x\|_\infty = \max\{|x_1|, |x_2| --- |x_n|\}$$

$$\|y\|_\infty = \max\{|y_1|, |y_2| --- |y_n|\}$$

$$\|x+y\|_\infty = \max\{|x_1+y_1|, |x_2+y_2| --- |x_n+y_n|\}$$

lets assume that $\boxed{\|x+y\|_\infty = |x_i+y_i|}$ (ith term is maximum)

in that case

$$\boxed{|x_i+y_i| \le |x_i| + |y_i|} \longrightarrow 1$$

[b/c we know that $|a+b| \le |a|+|b|$]

we can say that,

$|x_i|$ could be the largest value in the set $(|x_1|, |x_2| --- |x_n|)$ or it could be lesser than the largest value in this set so we can write

$$\boxed{|x_i| \leq \max \{|x_1|, |x_2| --- |x_n|\}}$$

②

Similarly,

$|y_i|$ could be the largest value in the set $(|y_1|, |y_2| --- |y_n|)$ or it could be lesser than the largest value, so we can write.

$$\boxed{|y_i| \leq \max \{|y_1|, |y_2| --- |y_n|\}}$$

③

Summing Equation ② and ③ we get -

$$|x_i| + |y_i| \leq \max \{|x_1|, |x_2| --- |x_n|\} + \max \{|y_1|, |y_2| --- |y_n|\}$$

Hence we know that, It will be true to write ⇒

$$|x_i| + |y_i| \leq \|x\|_\infty + \|y\|_\infty$$

But we also know that.

$$|x_i + y_i| \leq |x_i| + |y_i|$$

so we can say that.

$$\boxed{|x_i + y_i| \leq |x_i| + |y_i| \leq \|x\|_\infty + \|y\|_\infty}$$

But $|x_i + y_i| = \|x+y\|_\infty$

so we can write the equation as =

$$\boxed{\|x+y\|_\infty \le \|x\|_\infty + \|y\|_\infty}$$ proved

(c)= Non-negativity $\boxed{\|x\|_\infty \ge 0}$

n vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ Then $\|x\|_\infty = \max\{|x_1|, |x_2| \dots |x_n|\}$

$\|x\|_\infty$ = maximum value among absolute values of entries of x-vector.

we know that $\boxed{|a| \ge 0}$ for all real numbers.

Hence all the absolute values of entries of x →

$|x_1| \ge 0$, $|x_2| \ge 0$ --- $|x_n| \ge 0$

Hence $\max\{|x_1|, |x_2| \dots |x_n|\} \ge 0$

Hence, we can say $\boxed{\|x\|_\infty \ge 0}$ proved

(d)= Definiteness → $\boxed{\|x\|_\infty = 0 \text{ only if } x = 0}$

n-vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ Then $\|x\|_\infty = \max\{|x_1|, |x_2| \dots |x_n|\}$

$\max\{|x_1|, |x_2| \dots |x_n|\}$ could be zero if all elements in the set $(|x_1|, |x_2| \dots |x_n|)$ are zero i.e.

$|x_1| = 0$, $|x_2| = 0$, --- $|x_n| = 0$

i.e. $x_1 = 0$, $x_2 = 0$ ---- $x_n = 0$

(5) = ~~Definiteness~~

Hence we can say that.

$\|x\|_\infty = 0$, only if all the absolute values of vector x entries are zero.

we know that $|a| = 0$ Then $a = 0$.

so if $|x_1| = 0$, Then $x_1 = 0$

$|x_2| = 0$ Then $x_2 = 0$

$\vdots$

$|x_n| = 0$ Then $x_n = 0$

This means. if $\|x\|_\infty = 0$

Then x = zero vector with all entries $= 0$.

$$x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

NOTE = if any element in the entry set is lesser than zero or greater than zero.

ie $x_i < 0$ or $x_i > 0$

in that case $|x_i|$ = Positive non zero value.

in that case $\|x\|_\infty$ = can never be zero.

Problem 3.24 ⇒ $z_1, z_2 --- z_n$ is collection of n-vector.

x is another n-vector.

vector z_j is the (distance) nearest neigbour of x

if $\|x - z_j\| \leq \|x - z_i\|$, $i = 1, 2 --- n$

vector z_j is the angle nearest neighbour of x if

$$\angle(x, z_j) \leq \angle(x, z_i) , \quad i = 1, 2, --- n$$

(a) = Give a simple specific numerical Example where the distance nearest neighbour is not the same as the angle nearest neighbours.

lets consider n-vector x, (where $n=5$).

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

z_1 and z_2 is the collection of two vectors. (5-vectors)

Such that $z_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 10 \end{bmatrix}$

$$z_2 = 3x = \begin{bmatrix} 3 \\ 6 \\ 9 \\ 12 \\ 15 \end{bmatrix}$$

Then distances of vectors z_1 and z_2 from x are –

① = $\|x - z_1\|$ = distance of vector x from z_1

$$x - z_1 = \begin{bmatrix} 1-0 \\ 2-0 \\ 3-0 \\ 4-0 \\ 5-10 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ -5 \end{bmatrix}$$

$$\|x - z_1\| = \sqrt{(1)^2 + (2)^2 + (3)^2 + (4)^2 + (-5)^2}$$

$$= \sqrt{1 + 4 + 9 + 16 + 25}$$

$$= \sqrt{55}$$

$$= \sqrt{55}$$

$$\cong 7.41$$

② = $\|x - z_2\|$ = distance of x from z_2

$$x - z_2 = \begin{bmatrix} 1-3 \\ 2-6 \\ 3-9 \\ 4-12 \\ 5-15 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ -6 \\ -8 \\ -10 \end{bmatrix}$$

$$\|x - z_2\| = \sqrt{(-2)^2 + (-4)^2 + (-6)^2 + (-8)^2 + (-10)^2}$$

$$= \sqrt{4 + 16 + 36 + 64 + 100}$$

$$= \sqrt{4 + 16 + 36 + 64 + 100}$$

$$= \sqrt{220}$$

$$\cong 14.83$$

$\|x - z_2\| > \|x - z_1\|$

$\boxed{\text{Hence distance nearest neighbour of } x \text{ is } z_1.}$

now the angles of vectors z_1 and z_2 with respect to x are

① $\angle(x, z_1) \Rightarrow$ angle b/w x and z_1 vector.

$$\theta = \arccos\left(\frac{x^T z_1}{\|x\| \|z_1\|}\right)$$

$$\theta = \arccos\left(\frac{\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}^T \begin{bmatrix}0\\0\\0\\0\\10\end{bmatrix}}{\sqrt{1^2+2^2+3^2+4^2+5^2}\,\sqrt{(0)^2+(0)^2+0^2+0^2+10^2}}\right)$$

$$\theta = \arccos\left(\frac{1.0+2.0+3.0+4.0+5.10}{\sqrt{55}\,\sqrt{100}}\right)$$

$$\theta = \arccos\left(\frac{50}{10\sqrt{55}}\right)$$

$$\theta = \arccos\left(\frac{5}{\sqrt{55}}\right)$$

$$\theta = \arccos\left(\frac{5}{7.41}\right)$$

$$\theta = \arccos(0.6747)$$

$$\theta \cong 47.56$$

hence $\boxed{\angle(x, z_1) = \theta \cong 47.56}$

(2) = $\angle(x, z_2)$ = angle b/w x and z_2 vector

$$\cos\theta = \frac{x^T z_2}{\|x\| \, \|z_2\|}$$

$$\theta = \arccos\left[\frac{\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}^T \begin{bmatrix}3\\6\\9\\12\\15\end{bmatrix}}{\sqrt{1^2+2^2+3^2+4^2+5^2}\,\sqrt{3^2+6^2+9^2+12^2+15^2}}\right]$$

$$\theta = \arccos\left[\frac{3+12+27+48+75}{\sqrt{55}\,\sqrt{9+36+81+144+225}}\right]$$

$$\theta = \arccos\left[\frac{165}{\sqrt{55}\,\sqrt{495}}\right]$$

$$\theta = \arccos\left(\frac{165}{\sqrt{27225}}\right)$$

$$\theta = \arccos\left[\frac{165}{165}\right]$$

$$\theta = \arccos\,[1]$$

$$\boxed{\theta = 0^\circ}$$

Hence angle b/w vectors x and z_2 is 0.

we can see that

$$\boxed{\begin{array}{l} \angle(x, z_1) > \angle(x, z_2) \\ (47.56 > 0^\circ) \\ \text{Hence we can say that angle nearest neighbour of } x \\ \text{is } z_2. \end{array}}$$

we saw in that Example that.

distance nearest neighbour of vector x is z_1

while angle nearest neighbour of vector x is z_2.

where $x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$, $z_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 10 \end{bmatrix}$, $z_2 = \begin{bmatrix} 3 \\ 6 \\ 9 \\ 12 \\ 15 \end{bmatrix}$

(b) = vectors $z_1, z_2, \ldots z_m$ are normalized.
i.e. $\|z_i\| = 1$. , $i = 1, 2 \ldots M$

show that in this case distance nearest neighbour and angle nearest neighbour is always same.

lets consider n-vector $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$

lets consider z_i is the vector which is angle nearest neighbour of vector x.

if z_i is the vector which is nearest angle neighbour of x then we can say -

$$\angle(x, z_i) < \angle(x, z_j)$$ [where $j = 1,2,3 \ldots$ Excluding i]

we can write =

$$\arccos\left(\frac{x^T z_i}{\|x\| \|z_i\|}\right) < \arccos\left(\frac{x^T z_j}{\|x\| \|z_j\|}\right)$$

It is given that if $\boxed{\arccos(u) > \arccos(v)}$

Then $\boxed{-1 \leq u < v \leq 1}$

so we can write that =

$$\frac{x^T z_i}{\|x\| \|z_i\|} > \cancel{\arc} \frac{x^T z_j}{\|x\| \|z_j\|} \longrightarrow (1)$$

we know that $[\|z_1\| = 1 \; \|z_2\| = 1 \ldots \|z_m\| = 1]$

so we can rewrite equation (1) as =

$$\frac{x^T z_i}{\|x\| \cdot 1} > \frac{x^T z_j}{\|x\| \cdot 1}$$

$$\Rightarrow \boxed{x^T z_i > x^T z_j}$$ [$\|x\|$ = positive, multiply both side by $\|x\|$]

we Have equation =

$$x^T z_i > x^T z_j$$

Multiply by Scalar -2 on both sides →

$$-2(x^T z_i) < -2(x^T z_j)$$

add $[1 + \|x\|^2]$ on both sides, its a positive value →

$$1 + \|x\|^2 - 2(x^T z_i) < 1 + \|x\|^2 - 2(x^T z_j)$$

we know that $[\|z_i\| = \|z_j\| = 1]$ so we can write =

$$\|x\|^2 + \|z_i\|^2 - 2(x^T z_i) < \|x\|^2 + \|z_j\|^2 - 2(x^T z_j)$$

$$\Rightarrow (\|x - z_i\|)^2 < (\|x - z_j\|)^2$$

Taking square root both sides we get =

$$\sqrt{(\|x - z_i\|)^2} < \sqrt{(\|x - z_j\|)^2}$$

$$\boxed{\|x - z_i\| < \|x - z_j\|}$$ Proved

$\|x - z_i\| \Rightarrow$ is the distance of vector x from z_i

$\|x - z_j\| =$ is the distance of vector x from all other vectors z_j where $(J = 1, 2, 3 --- M$ Excluding $i)$

Here we can see that, when $\|z\| = 1$

$$\angle(x, z_i) < \angle(x, z_j)$$

Then

$$\|x - z_i\| < \|x - z_j\|$$

i.e. The angle nearest neighbour is same as distance nearest neighbour.

Problem 3.26 =

T-vector x is a non constant time series.

$$\boxed{\mu = \frac{(1^T x)}{T}}$$ = Mean Value.

Auto Correlation $R(\tau)$ = Defined as correlation coeffici-ent of two vectors for $\tau = 0, 1, 2 \ldots$ $(x, \mu 1_\tau)$ and $(\mu 1_\tau, x)$

vector $(x, \mu 1_\tau)$ and $(\mu 1_\tau, x)$ have Mean μ.

(a) = Explain why $R(0) = 1$ and $R(\tau) = 0$ for $\tau \geq T$.

when $\boxed{\tau = 0}$ Then $\rightarrow$

if T-vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}$

Then vector $(x, \mu 1_0)$ and $(\mu 1_0, x)$ will also be same as x.

vector $(x, \mu 1_0) = (\mu 1_0, x)$ for $(\tau = 0)$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}$$

~~correlation coefficient~~ (ρ).

lets denote vector $(x, \mu 1_0)$ By a.
and vector $(\mu 1_0, x)$ By b.

Then correlation coefficient $\rho = \dfrac{\hat{a}^T \hat{b}}{\|\hat{a}\| \, \|\hat{b}\|}$

where $\hat{a} = a - \text{avg}(a)\,1$
$\hat{b} = b - \text{avg}(b)\,1$

$$\hat{a} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} - \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_T = \begin{bmatrix} x_1 - \mu \\ x_2 - \mu \\ \vdots \\ x_T - \mu \end{bmatrix}$$

Similarly

$$\hat{b} = \begin{bmatrix} x_1 - \mu \\ x_2 - \mu \\ \vdots \\ x_T - \mu \end{bmatrix}$$

(b/c $\hat{a}$ and $\hat{b}$ will be same b/c a and b are same for $\tau = 0$).

$$\|\hat{a}\| = \sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 - \cdots (x_T-\mu)^2}$$
$$\|\hat{b}\| = \sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 - \cdots (x_T-\mu)^2}$$

lets calculate the value of $R(\tau)$ = correlation coefficient where $(\tau = 0)$.

$$R(0) = \mathcal{S} = \frac{\begin{bmatrix} x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \end{bmatrix}^T \begin{bmatrix} x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \end{bmatrix}}{\sqrt{(x_1-\mu)^2+(x_2-\mu)^2 - - - +(x_T-\mu)^2}\sqrt{(x_1-\mu)^2+(x_2-\mu)^2 + - +(x_T-\mu)^2}}$$

$$R(0) = \frac{(x_1-\mu)^2+(x_2-\mu)^2 + - - - - - + (x_T-\mu)^2}{(x_1-\mu)^2+(x_2-\mu)^2 + - - - - - (x_T-\mu)^2}$$

$\boxed{R(0) = 1}$ $\left(b/c \ \frac{\text{numerator}}{\text{denominator}} = 1, \text{ since both are same} \right)$

(2) = $\boxed{R(\tau) = 0, \text{ for } \tau \geq T}$

a = vector $(x, \mu 1_\tau)$ = $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix}$ $\begin{matrix} \updownarrow T \\ \\ \updownarrow \tau \end{matrix}$

a in $T+\tau$ vector

(τ time)

b = vector $(\mu 1_\tau, x)$ = $\begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}$ $\begin{matrix} \updownarrow (\tau \text{ times}) \\ \\ \updownarrow T \text{ times.} \end{matrix}$

b is $(T+\tau)$ vector

mean of $a = \mu$, Mean of $b = \mu$

$\hat{a} = a - \text{avg}(a)\mathbf{1}$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ \tau \\ \tau \\ \vdots \\ \tau \end{bmatrix}_{T+\tau} - \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix}_{T+\tau} = \begin{array}{c} \begin{bmatrix} x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \updownarrow T \\ \\ \updownarrow \tau \end{array}$$

$\hat{b} = b - \text{avg}(b)\mathbf{1}$

$$= \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}_{T+\tau} - \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix}_{T+\tau} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \end{bmatrix} \begin{array}{l} \updownarrow \tau \\ \\ \updownarrow T \end{array}$$

lets calculate $\hat{a}^T\hat{b} \Rightarrow$ when $(\tau \geq T)$ $\left(\begin{array}{c}\tau = T \\ \text{and when } \tau > T\end{array}\right)$

in that case

$$\begin{bmatrix} x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ x_1-\mu \\ x_2-\mu \\ \vdots \\ x_T-\mu \end{bmatrix}$$

(when $\tau = T$)

$$\Rightarrow (x_1-\mu).0 + (x_2-\mu).0 \ldots\ldots + (x_T-\mu).0$$
$$+ 0\cdot(x_1-\mu) + 0\cdot(x_2-\mu) \ldots + 0\cdot(x_T-\mu)$$

$= 0$

when $(\tau > T)$ In that case

$$\hat{a}^T \hat{b} = (x_1-\mu).0 + (x_2-\mu).0 \;\cdots\; (x_T-\mu).0$$
$$+ 0.0 + 0.0 \;\cdots$$
$$+ 0.(x_1-\mu) + 0.(x_2-\mu) \;\cdots\; + 0.(x_T-\mu)$$

$$\boxed{\hat{a}^T \hat{b} = 0}$$

Here $[0.0 + 0.0 \cdots$ term will come for $(\tau - T)$ times$]$

we now that when $\boxed{\tau \geq T}$ Then $\boxed{\hat{a}^T \hat{b} = 0}$

Hence $R(\tau) = \rho = \dfrac{\hat{a}^T \hat{b}}{\|\hat{a}\| \cdot \|\hat{b}\|}$

$$R(\tau) = \rho = \frac{0}{\|\hat{a}\| \, \|\hat{b}\|} = 0$$

Hence $\boxed{R(\tau) = 0}$ when $\boxed{\tau \geq T}$

(b) = if z denote standardized or z scored version of x !
show that for $\tau = 0, 1 \ldots\ldots T-1$

$$R(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} z_t z_{t+\tau}$$

$$\text{vector } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}$$

$$\text{vector } (x, \mu 1_\tau) = a = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} T \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \tau \end{matrix}, \quad \text{avg}(a) = \mu$$

$$\text{vector } (\mu 1_\tau, x) = b = \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \tau \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} T \end{matrix}, \quad \text{avg}(b) = \mu$$

$$\hat{a} = a - \text{avg}(a) 1$$
$$= a - \mu 1$$
$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} - \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 - \mu \\ x_2 - \mu \\ \vdots \\ x_T - \mu \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} T \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \tau \end{matrix}$$

$$\hat{b} = b - \text{avg}(b) 1$$
$$= b - \mu I$$
$$= \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} - \mu \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_1 - \mu \\ x_2 - \mu \\ \vdots \\ x_T - \mu \end{bmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} \tau \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} T \end{matrix}$$

lets calculate ⇒

$$\hat{a}^T \hat{b} = (x_1 - \mu).0 + (x_2 - \mu).0 \text{ ----- } + (x_\tau - \mu).0$$
$$+ (x_{\tau+1} - \mu)(x_1 - \mu) + (x_{\tau+2} - \mu)(x_2 - \mu) \text{---} + (x_T - \mu)(x_{T-\tau} - \mu)$$
$$+ 0.(x_{T-\tau+1} - \mu) \text{ ----- } 0.(x_T - \mu)$$

$$\hat{a}^T \hat{b} = (x_{\tau+1} - \mu)(x_1 - \mu) + (x_{\tau+2} - \mu)(x_2 - \mu) \text{ ------}$$
$$+ (x_T - \mu)(x_{T-\tau} - \mu)$$

$$\Rightarrow x_1 x_{\tau+1} - \mu x_{\tau+1} - \mu x_1 + \mu^2 + x_2 x_{\tau+2}$$
$$- \mu x_{\tau+2} - \mu x_2 + \mu^2 \text{ ---------}$$
$$\text{----} + x_{T-\tau} x_T - \mu x_T - \mu x_{T-n} + \mu^2$$

lets devide both sides by $(\|\hat{a}\| . \|\hat{b}\|)$. ⇒

$$\frac{\hat{a}^T \hat{b}}{\|\hat{a}\| \|\hat{b}\|} = \frac{(x_{\tau+1} - \mu)(x_1 - \mu) + (x_{\tau+2} - \mu)(x_2 - \mu) \text{----} + (x_T - \mu)(x_{T-\tau} - \mu)}{\|\hat{a}\| \|\hat{b}\|}$$

⇒ lets multiply both sides by T.

$$T. \left[\frac{\hat{a}^T \hat{b}}{\|\hat{a}\| \|\hat{b}\|}\right] = T. \frac{\left[(x_{\tau+1} - \mu)(x_1 - \mu) + (x_{\tau+2} - \mu)(x_2 - \mu) \text{---} + (x_T - \mu)(x_{T-\tau} - \mu)\right]}{\|\hat{a}\| \|\hat{b}\|}$$

By spreading Expression we can write -

$$T \cdot (R(\tau)) = \frac{\sqrt{T}(x_1 - \mu)}{\|\hat{a}\| \, \|\hat{b}\|} \cdot \frac{\sqrt{T}(x_{\tau+1} - \mu)}{\cancel{\|a}}$$

$$+ \frac{\sqrt{T}(x_2 - \mu) \cdot \sqrt{T}(x_{\tau+2} - \mu)}{\|\hat{a}\| \, \|\hat{b}\|}$$

$$\vdots$$

$$+ \frac{\sqrt{T}(x_{T-\tau} - \mu) \cdot \sqrt{T}(x_T - \mu)}{\|\hat{a}\| \, \|\hat{b}\|}$$

Put value of $\|\hat{a}\|$ and $\|\hat{b}\|$ in above equation -

$$\boxed{\begin{aligned} T \cdot (R(\tau)) = {} & \frac{\sqrt{T}(x_1 - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 - \cdots (x_T-\mu)^2}} \cdot \frac{\sqrt{T}(x_{\tau+1} - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots (x_T-\mu)^2}} \\ & + \frac{\sqrt{T}(x_2 - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots (x_T-\mu)^2}} \cdot \frac{\sqrt{T}(x_{\tau+2} - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots (x_T-\mu)^2}} \\ & \vdots \\ & + \frac{\sqrt{T}(x_{T-\tau} - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 + \cdots (x_T-\mu)^2}} \cdot \frac{\sqrt{T}(x_T - \mu)}{\sqrt{(x_1-\mu)^2 - \cdots + (x_T-\mu)^2}} \end{aligned}}$$

Note - $\|\hat{a}\| = \|\hat{b}\| = \sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 - \cdots - (x_T-\mu)^2}$

Equation (1)

now lets calculate standized version of x.

which is given by $z = \frac{1}{std(x)} (x - avg(x) 1)$

$$std(x) = \frac{\left\| x - \left(\frac{1^T . x}{T}\right) . 1 \right\|}{\sqrt{T}}, \quad (avg(x) = u)$$

$$\Rightarrow \frac{\| x - avg(x) . 1 \|}{\sqrt{T}} = \frac{\left\| \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} - u \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\|}{\sqrt{T}}$$

$$= \frac{\left\| \begin{bmatrix} x_1 - u \\ x_2 - u \\ \vdots \\ x_T - u \end{bmatrix} \right\|}{\sqrt{T}}$$

$$\Rightarrow \frac{\sqrt{(x_1-u)^2 + (x_2-u)^2 - - - + (x_T-u)^2}}{\sqrt{T}}$$

Then $$Z_1 = \frac{\sqrt{T} . (x_1 - u)}{\sqrt{(x_1-u)^2 + (x_2-u)^2 - - - (x_T-u)^2}}$$

$$Z_2 = \frac{\sqrt{T} (x_2 - u)}{\sqrt{(x_1-u)^2 + (x_2-u)^2 - - - (x_T-u)^2}}$$

$\vdots$

$$Z_T = \frac{\sqrt{T}\,(x_T - \mu)}{\sqrt{(x_1-\mu)^2 + (x_2-\mu)^2 \cdots (x_T-\mu)^2}}$$

let's compare equation ① with values of $Z_1, Z_2 \cdots Z_T$ then we find that -

$$T.\left(R(\tau)\right) = \sum_{t=1}^{T-\tau} Z_t \cdot Z_{t+\tau}$$

Hence.

$$\boxed{R(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} Z_t \cdot Z_{t+\tau}}$$

(C)= find the auto correlation for time series

$x = (+1, -1, +1, -1 - \cdots +1, -1)$

where T is even.

if T is even then mean of x can be calculated as.

$$\frac{x^T 1}{T} = \frac{\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ \vdots \\ 1 \\ -1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}{T}$$

$$= \frac{1.1 + (-1).1 - \cdots + 1(1) + (-1).1}{T}$$

$$= \frac{1 + (-1) + 1 + (-1) - \cdots + 1 + (-1)}{T}$$

$$\text{mean}(x) = \frac{0}{T} = 0$$ [sum of series 1, -1 ---- till even number of times is zero]

$$\boxed{\text{mean}(x) \text{ or } \text{avg}(x) = 0}$$

$a =$ vector $(x, \mu 1_\tau)$ where $\mu = 0$

$$a = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \updownarrow T \\ \\ \updownarrow \tau \end{matrix}$$

Similarly $b = \text{vector}(u 1_\tau, t)$ where $u = 0$

$$b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \tau \\ \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} T \end{array}$$

avg(a) and avg(b) is also $= u = 0$.

$$\hat{a} = a - \text{avg}(a) \cdot 1$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - 0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Similarly

$$\hat{b} = b - \text{avg}(b) \cdot 1$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} - 0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}$$

Lets calculate $(\hat{a}^T \hat{b}) \Rightarrow$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} \Rightarrow \begin{array}{l} x_1 \cdot 0 + x_2 \cdot 0 \text{ ----- } + (x_\tau) \cdot 0 \\ + (x_{\tau+1} \cdot x_1) + (x_{\tau+2} \cdot x_2) + \cdots (x_T \cdot x_{T-\tau}) \\ + 0 \cdot (x_{T-\tau+1}) \text{ ---- } + 0 \cdot (x_T) \end{array}$$

$$\Rightarrow (x_{\tau+1} \cdot x_1) + (x_{\tau+2} \cdot x_2) \text{ ---- } + (x_T \cdot x_{T-\tau})$$

$$\|\hat{a}\| = \sqrt{x_1^2 + x_2^2 \text{----} x_T^2} \doteq \sqrt{(1)^2 + (-1)^2 \text{---} (-1)^2}$$

$$= \sqrt{T}$$

$$\|\hat{b}\| = \sqrt{(1)^2 + (-1)^2 \text{-----} (-1)^2}$$

$$= \sqrt{1+1+1+1 \text{----} 1}$$

$$= \sqrt{T}$$

Hence $\|\hat{a}\| = \|\hat{b}\| = \sqrt{T}$

Then

$$R(\tau) = \rho = \frac{\hat{a}^T \hat{b}}{\|\hat{a}\| \, \|\hat{b}\|}$$

$$R(\tau) = \frac{(x_{\tau+1} \cdot x_1) \pm (x_{\tau+2} \cdot x_2) \text{---} \pm (x_T \cdot x_{T-\tau})}{\sqrt{T} \cdot \sqrt{T}}$$

$$\boxed{R(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} x_t \, x_{t+\tau}}$$

Answer

(d)= x denotes the number of meals served by a restaurant on day τ.

It is observed that $R(7)$ = fairly High

$R(14)$ = also High but not as high.

<u>Explanation</u> ⇒

with previous Examples we understood that.

$R(0) = 1$

i.e. correlation b/w data points is highest when there is no time lag $(\tau=0)$, data points are same.

but as we increase time lag. Then $R(\tau)$ start decreasing b/c correlation b/w data points will decrease over time.

ultimately when $(\tau \geq T)$ $R(\tau) = 0$.

In above question $R(7)$ = fairly high b/c autocorrelation function is telling us that number of meals served from day 1 to 7 form patterns which are highly correlated

but $R(14)$ is is not as high b/c since there are more data points for over 14days the correlation b/w the data pattern decreases.

Hence $R(14)$ is less.

Problem 4.2 ⇒

(a)= Repeting k-means algorithm on given data set twice for k=2, k=5, k=10 values and observing results

k=2 observations =

1st Repeatition = In first repeatition final convergence was achieved in 8 iterations. and data was grouped into two clusters. Below are the values of J_{clust} in each iteration.

$$\left[J_1 = 0.914, J_2 = 0.897, J_3 = 0.896, J_4 = 0.887, J_5 = 0.874, J_6 = 0.873, J_7 = 0.872, J_8 = 0.872 \right]$$

convergence was achieved in 7th and 8th iteration. and data was grouped in two clusters.

2nd Repeatition = In second Repeatition results changed dramatically. data was grouped into clusters in 6 iterations. J_{clust} values ⇒

$$\left[J_1 = 0.93, J_2 = 0.917, J_3 = 0.902, J_4 = 0.90, J_5 = 0.899, J_6 = 0.899 \right]$$

convergence was achieved in 5th and 6th iterations.

Conclusion = for (k=2) the first run was slower which took more iterations to cluster the data but it minimized J effectively Hence it did clustering more effectively ($J_8 = 0.872$) On the other hand 2nd run was fast but clustering was not effective b/c value of J is High. $J_6 = 0.899$ compared to 1st run.

$K = 5$ Observations =

1st repeatition = In first repeatition clustering process was little faster, data was clustered into 5 clusters in 7 iterations. J values are given as —

$$\left[J_1 = 0.875,\ J_2 = 0.825,\ J_3 = 0.810,\ J_4 = 0.775,\ J_5 = 0.770,\ J_6 = 0.770 \right]$$

convergence was achieved in 7th iteration.

2nd repeatition = In second repeatition clustering process was little slower than the first repeatition, data was clustered into 5 clusters in 9 iterations. J values are given as follow →

$$\cancel{\left[J_1 = 0.864,\ J_2 = 0.790,\ J_3 = 0.787,\ J_4 = 0.785,\ J_5 = 0.775,\ J_6 = 0.775,\ J_7 = \right.}$$

$$\left[J_1 = 0.864,\ J_2 = 0.796,\ J_3 = 0.794,\ J_4 = 0.790,\ J_5 = 0.787,\ J_6 = 0.786,\ J_7 = 0.785,\ J_8 = 0.780,\ J_9 = 0.780 \right]$$

Conclusion = for $K = 5$, 1st repeatition was more effective and efficient, it completed only in 6 iteration compared to 2nd repeatition which completed in 9 iterations. also in 1 repeatition value of J was minimized effectively. we can see final J value which is $J_6 = 0.770$.

$$\boxed{\underset{\text{In first Run}}{J_{min} = 0.770} < \underset{\text{In Second Run}}{J_{min} = 0.780}}$$

K=10 obnervations =

1st repeatition ⇒ for k=10 first repeatition took Total 8 iterations to do the data clustering, to group data into 10 clusters. value of J are as below –

$$\begin{bmatrix} J_1 = 0.825, & J_2 = 0.765, & J_3 = 0.745, & J_4 = 0.740 \\ J_5 = 0.739, & J_6 = 0.738, & J_7 = 0.738, & J_8 = 0.738 \end{bmatrix}$$

convergenie was achieved in 6th, 7th & 8th iterations

2nd Repeatition = for k=10, second Run also took 8 iterations to cluster the data into 10 groups. value of J are as follow –

$$\begin{pmatrix} J_1 = 0.820, & J_2 = 0.736, & J_3 = 0.734, & J_4 = 0.731, \\ J_5 = 0.730, & J_6 = 0.729, & J_7 = 0.727, & J_8 = 0.727 \end{pmatrix}$$

convergence was achieved in 8th iteration.

conclusion = Both the ~~iter~~ Runs took same number of iterations to cluster the data into 10 groups but the second iteration Run was more effective than the first Run because In the second Run value of J was minimized in better way.

$$\boxed{\text{In first run } J_{Min} = 0.738 \;>\; \text{In Second Run } J_{min} = 0.727}$$

Hence 2nd Run was more effective

(b) = Topic discovery | Titles | words →

I chose K=5. and below are the results =
K=5 means There would be 5 clusters of documents articles. My analysis gives following Information.

(Titles)

Centroid 1 ⇒

Top articles associated →
- Brock
- Bulbasaur
- Deoxy
- Eeve
- Gameplay of Pokemon
- mew
- Pikachu
- Pokemon anime
- ⋮
- Total 33 Titles

Top words associated →
- Pokemon
- game
- games
- player
- Nintendo
- Pikachu

Description - This cluster indicates that the most articles are about a specific game for Pokemons which players can play on nintendo.

Centroid 2 = (Titles)

article titles associated → Amplitude modulation
→ Amplitude shift keying
→ Analog signal
→ Antenna
→ Asynchronous Transfer mode
→ Broadcasting
⋮
(55 article titles)

words associated → radio
→ signal
→ frequency
→ telephone
→ digital
⋮
(1000 more)

Description = The articles in cluster 2 are about "Transmission of signals" and Signal Properties.

Centroid 3 =

Article titles associated

- Acid rain
- Albedo
- Anemometer
- Atmosphere
- Atmosphere of Earth
- Barograph
- Barometer
- ⋮
- Erosion
- famine
- flood

(100 titles more)

associated words
- weather
- wind
- pressure
- air
- temperature
- Ice.

Description = articles associated to this group are mainly about, earth climate, Atmosphere & weather.

Centroid 4 =

article Titles Associated
- conventions on the Rights of Persons with disabilities
- Headquarters of the united Nations
- International Bank of Reconstruction and development
- world Bank group
- United nations
- UN women
- UNESCO
- UNICEF

associated words
- nations
- International
- members
- council
- general
- convention
- assembly.
- rights.

Description = The articles in this Cluster are mainly about International organization which work for people around the world globally like UN, UNESCO, etc.

centroid 5 =

associated article Titles → A bar at the folies Bergere
→ Alfred Sisley
→ Armand guillaumin
→ Art institute of chicago
→ Paul signac
→ still life
→ The childs Bath
→ woman with a parasol

associated words → painting
→ art
→ artists
→ monet
→ paris
→ works
→ salon
→ group.

Description = The articles in this group will mainly be about different artforms like acting, painting and different famous artworks of famous artist in their respective field.