ML Assignment 2

Approach #1 → derivative = 0

$$RSS(w) = \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right)^2$$

$$\frac{\partial RSS(w)}{\partial w_0} = \frac{\partial}{\partial w_0} \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right)^2$$

$$= \sum_{i=1}^{N} \frac{\partial}{\partial w_0} \left(y_i - [w_0 + w_1 x_i]\right)^2$$

$$= \sum_{i=1}^{N} 2\left(y_i - [w_0 + w_1 x_i]\right)(-1)$$

$$= -2 \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right)$$

similarly.

$$\frac{\partial RSS(w)}{\partial w_1} = -2 \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right) x_i$$

$$\nabla RSS(w_0, w_1) = \begin{bmatrix} -2 \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right) \\ -2 \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right) x_i \end{bmatrix}$$

Putting them equal to 0 we get.

$$-2 \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i]\right) = 0$$

$$\sum y_i - w_0 N - w_1 \sum x_i = 0$$

$$\text{so } \hat{w}_0 = \frac{\sum y_i}{N} - \hat{w}_1 \frac{\sum x_i}{N} = \text{Mean}(y_i) - \hat{w}_1 \text{Mean}(x_i)$$

Eq(I)

$$-2\sum_{i=1}^{N}\left(y_i - [w_0 + w_1 x_i]\right)x_i = 0$$

$$\sum y_i x_i - w_0 \sum x_i - w_1 \sum x_i^2 = 0$$

Substituting value of w_0 from eq (1) we get

$$\sum y_i x_i - \left(\frac{\sum y_i}{N} - \hat{w}_1 \frac{\sum x_i}{N}\right)\sum x_i - \hat{w}_1 \sum x_i^2 = 0$$

$$\sum y_i x_i - \frac{\sum y_i \sum x_i}{N} \cancel{-\hat{w}_1}\left(\sum x_i^2 - \frac{(\sum x_i)^2}{N}\right) = 0$$

$$\Rightarrow \hat{w}_1 = \frac{\sum y_i \cancel{x_i} - \frac{\sum y_i \sum x_i}{N}}{\sum x_i^2 - \frac{(\sum x_i)^2}{N}}$$

dividing both numerator & denominator by N

$$\text{so } \hat{w}_1 = \frac{\text{Mean}(y_i x_i) - \text{Mean}(y_i)\,\text{Mean}(x_i)}{\text{Mean}(x_i^2) - (\text{Mean}(x_i))^2} \quad \text{— Eq (2)}$$

with Eq (2) we get 'slope' and with slope we get intercept with Eq (1)

Now moving to ques 1, 2 and 3.

Ques 1

year = [2004, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017]

revenue (in B$) = [61.2, 58.3, 67.1, 69.2, 68.9, 83.5, 89.1, 80, 92.3, 93, 97] input = years , output = revenue

Mean(x_i) = 2011.73 , Mean(y_i) = 78.145
Mean(x_i^2) = 4047060.1 , Mean($y_i x_i$) = 157251.53

$$\hat{w}_1 = \frac{\text{Mean}(y_i x_i) - \text{Mean}(y_i)\,\text{Mean}(x_i)}{\text{Mean}(x_i^2) - (\text{Mean}(x_i))^2} \quad \text{(slope)}$$

Eq(2)

$\hat{w}_1$ (slope) = 3.28

$\hat{w}_0$ = Mean(y_i) − $\hat{w}_1$ Mean(x_i) (intercept)
$\hat{w}_0$ (intercept) = −6520.32 Eq (1)

b) $\hat{y}_i$ (x_i = 2019) = $\hat{w}_0 + \hat{w}_1$(2019)
= 102.0 billion Rupees

c) error = output − predicted output

error = [8.4, −7.62, −2.1, −3.28, −6.86, 4.46, 6.78, −5.6, 3.42, 0.84, 1.56]

error keeps on decreasing and then increasing in magnitude.

RSS (w) = 302.90

Ques 2

ML = [75, 80, 93, 65, 87, 71, 98, 68, 84, 77]
HUR = [82, 78, 86, 72, 91, 80, 95, 72, 89, 74]

(a) Taking ML as independent variable

$x = ML$, $y = HUR$

$Mean(x_i) = 79.8$, $Mean(y_i) = 81.9$

$Mean(x_i^2) = 6472.2$, $Mean(y_i x_i) = 6604.5$

$$\hat{w}_1 = \frac{Mean(y_i x_i) - Mean(y_i)\,Mean(x_i)}{Mean(x_i^2) - (Mean(x_i))^2}$$

$\hat{w}_1 = 0.66$ slope

$\hat{w}_0 = Mean(y_i) - \hat{w}_1\,Mean(x_i)$

$\hat{w}_0 = 29.13$ intercept

$RSS(\hat{w}) = 143.40$

(c) $\hat{y}(x=96) = \hat{w}_0 + \hat{w}_1(96)$

$= 92.6$ marks in HUR

(b) Taking HUR as independent variable

$x = HUR$, $y = ML$

$Mean(x_i) = 81.9$, $Mean(y_i) = 79.8$

$Mean(x_i^2) = 6767.5$, $Mean(y_i x_i) = 6604.5$

Using the formula for $\hat{w}_1$ we get

$\hat{w}_1 = 1.15$ slope

and for $\hat{w}_0$ we get

$\hat{w}_0 = -14.39$ intercept

$RSS(\hat{w}) = 249.40$

(d) $\hat{y}(x=95) = \hat{w}_0 + \hat{w}_1(95)$

$= 94.86$ in ML

(c) Both linear regression models are different as they have different values of 'slope' and 'intercept' for example :-

in (d) expected value of marks in ML for 95 marks in HUR is $= \underline{94.86}$ (in 2nd model)

while acc to the equation in 1st model

$$\hat{x}(y=95) = \frac{95 - \hat{w}_0}{\hat{w}_1} = \underline{99.80} \text{ in ML}$$

which is different from 94.86 in ML, the value we got from 2nd model

So the two models are different.

Ques 3

$V = [54.3,\ 61.8,\ 72.4,\ 88.7,\ 118.6,\ 194]$

$P = [61.2,\ 49.5,\ 37.5,\ 28.4,\ 19.2,\ 10.1]$

taking input $x = \frac{1}{V}$ & $y = P$ output and applying simple linear regression model

Mean$(x_i) = 0.0122$, Mean$(y_i) = 34.32$

Mean$(x_i^2) = 0.000169$, Mean$(x_i y_i) = 0.496$

$$\hat{w}_1 = \frac{\text{Mean}(y_i x_i) - \text{Mean}(y_i)\,\text{Mean}(x_i)}{\text{Mean}(x_i^2) - (\text{Mean}(x_i))^2}$$

Slope $\hat{w}_1 = 3824.39$

$\hat{w}_0 = \text{Mean}(y_i) - \hat{w}_1 \,\text{Mean}(x_i)$

Intercept $\hat{w}_0 = -12.38$

$RSS(\hat{w}) = 32.13$

$\hat{y} = \hat{w_0} + \hat{w_1} x$

$\Rightarrow \hat{P} = \hat{w_0} + \hat{w_1} \frac{1}{V}$

$$\Rightarrow \boxed{\hat{P} = -12.38 + \frac{3824.39}{V}}$$

$\Rightarrow (\underbrace{\hat{P} + 12.38}_{P}) V' = \underline{3824.39}_{c}$

a) so for $PV^n = C$ (const)

$n = 1$, $C = 3824.39$

b) eq^n connecting P and V is

$$(P + 12.38) V = 3824.39$$

c) $\hat{P}(V=100) = -12.38 + \frac{3824.39}{100}$

$\hat{P} = \underline{25.85}$

Ques 4

$y = w_0 + w_1 x + w_2 x^2$

$RSS(w) = \sum_{i=1}^{N} \left(y_i - [w_0 + w_1 x_i + w_2 x_i^2] \right)^2$

$\frac{\partial RSS(w)}{\partial w_0} = \sum 2 \left(y_i - [w_0 + w_1 x_i + w_2 x_i^2] \right)(-1)$ —(1)

$\frac{\partial RSS(w)}{\partial w_1} = \sum 2 \left(y_i - [w_0 + w_1 x_i + w_2 x_i^2] \right)(-x_i)$ —(2)

$\frac{\partial RSS(w)}{\partial w_0} = \sum 2 \left(y_i - [w_0 + w_1 x_i + w_2 x_i^2] \right)(-x_i^2)$ —(3)

$$\nabla RSS(w) = \begin{bmatrix} \frac{\partial RSS(w)}{\partial w_0} \\ \frac{\partial RSS(w)}{\partial w_1} \\ \frac{\partial RSS(w)}{\partial w_2} \end{bmatrix}$$

now putting

$\nabla RSS(w) = 0$

$$\frac{\partial RSS(w)}{\partial w_0} = 0, \quad \frac{\partial RSS(w)}{\partial w_1} = 0, \quad \frac{\partial RSS(w)}{\partial w_2} = 0$$

3 equations, 3 variables :- w_0, w_1, w_2

let's call $A(x)$ = average of x_i's and similarly for y

$$(1) - \; -2\sum_{i=1}^{N} (y_i - [w_0 + w_1 x_i + w_2 x_i^2]) = 0$$

$$\sum y_i - w_0 N - w_1 \sum x_i - w_2 \sum x_i^2 = 0$$

$$\Rightarrow \quad w_0 = A(y) - w_1 A(x) - w_2 A(x^2) \qquad Eq(1)$$

$$(2) - \; -2\sum_{i=1}^{N} (y_i - [w_0 + w_1 x_i + w_2 x_i^2]) x_i = 0$$

$$\sum y_i x_i - w_0 \sum x_i - w_1 \sum x_i^2 - w_2 \sum x_i^3 = 0$$

(dividing by N)

$$A(yx) - w_0 A(x) - w_1 A(x^2) - w_2 A(x^3) = 0$$

Eq(2)

$$(3) - \; -2\sum_{i=1}^{N} (y_i - [w_0 + w_1 x_i + w_2 x_i^2]) x_i^2 = 0$$

$$\sum y_i x_i^2 - w_0 \sum x_i^2 - w_1 \sum x_i^3 - w_2 \sum x_i^4 = 0$$

(dividing by N)

$$A(yx^2) - w_0 A(x^2) - w_1 A(x^3) - w_2 A(x^4) = 0$$

Eq (3)

so the three equations are

$$w_0 = A(y) - w_1 A(x) - w_2 A(x^2) \quad - Eq(1)$$

$$A(yx) - w_0 A(x) - w_1 A(x^2) - w_2 A(x^3) = 0 \quad - Eq(2)$$

$$A(yx^2) - w_0 A(x^2) - w_1 A(x^3) - w_2 A(x^4) = 0 \quad - Eq(3)$$

substituting the value of w_0 from Eq (1) to Eq(2) we get

$$A(yx) - \left(A(y) - w_1 A(x) - w_2 A(x^2)\right) A(x) - w_1 A(x^2) - w_2 A(x^3) = 0$$

$$A(yx) - A(y)A(x) - w_1\left(A(x^2) - (A(x))^2\right) - w_2\left(A(x^3) - A(x^2)A(x)\right) = 0$$

$$\Rightarrow w_1 = \frac{A(yx) - A(y)A(x)}{A(x^2) - (A(x))^2} - w_2 * \frac{A(x^3) - A(x^2)A(x)}{A(x^2) - (A(x))^2}$$

Eq (4)

Substituting ~~the value~~ of w_0 from Eq(1) & w_1 from Eq(4) to Eq(3) we get

$$A(yx^2) - \left(A(y) - w_1 A(x) - w_2 A(x^2)\right) A(x^2) - w_1 A(x^3) - w_2 A(x^4) = 0$$

$$A(yx^2) - A(y)A(x^2) - w_1\left(A(x^3) - A(x)A(x^2)\right) - w_2\left(A(x^4) - (A(x^2))^2\right) = 0$$

$$\Rightarrow w_1 = \frac{A(yx^2) - A(y)A(x^2)}{A(x^3) - A(x)A(x^2)} - w_2 * \frac{A(x^4) - (A(x^2))^2}{A(x^3) - A(x)A(x^2)}$$

Eq (5)

equating Eq(4) & Eq(5) we get

$$\frac{A(yx)-A(y)A(x)}{A(x^2)-(A(x))^2} - w_2 * \frac{A(x^3)-A(x^2)A(x)}{A(x^2)-(A(x))^2}$$

$$= \frac{A(yx^2)-A(y)A(x^2)}{A(x^3)-A(x)A(x^2)} - w_2 * \frac{A(x^4)-(A(x^2))^2}{A(x^3)-A(x)A(x^2)}$$

let

$p = A(yx) - A(y)A(x)$	$s = A(yx^2) - A(y)A(x^2)$
$q = A(x^2) - (A(x))^2$	$t = A(x^4) - (A(x^2))^2$
$r = A(x^3) - A(x^2)A(x)$	

now equation become

$$\frac{p}{q} - w_2\frac{r}{q} = \frac{s}{r} - w_2\frac{t}{r}$$

$$\Rightarrow \frac{p - w_2 r}{q} = \frac{s - w_2 t}{r}$$

$$\Rightarrow r(p - w_2 r) = q(s - w_2 t)$$

$$\Rightarrow rp - w_2 r^2 = qs - w_2 qt$$

$$\Rightarrow \boxed{\hat{w}_2 = \frac{qs - rp}{qt - r^2}} \quad \text{Eq (6)}$$

and $\boxed{\hat{w}_1 = \frac{p - \hat{w}_2 r}{q}}$

and

$$\boxed{\hat{w}_0 = A(y) - \hat{w}_1 A(x) - \hat{w}_2 A(x^2)}$$

$X = [0, 1, 2, 3, 4, 5, 6]$

$Y = [2.4, 2.1, 3.2, 5.6, 9.3, 14.6, 21.9]$

So the values are

$A(y) = 8.44$, $A(x) = 3.0$, $A(x^2) = 13.0$

$p = 12.8$, $q = 4.0$, $r = 24.0$, $s = 85.6$, $t = 156.0$

$$\hat{w}_2 = \frac{qs - rp}{qt - r^2} = 0.73 \quad -(iii)$$

$$\hat{w}_1 = \frac{p - \hat{w}_2 r}{q} = -1.20 \quad -(ii)$$

$$\hat{w}_0 = A(y) - \hat{w}_1 A(x) - \hat{w}_2 A(x^2)$$

$$\hat{w}_0 = 2.51 \quad -(i)$$

$$\hat{y} = \hat{w}_0 + \hat{w}_1 x + \hat{w}_2 x^2$$

$RSS(w) = 0.164$

almost fit!