# Part - I Written Exercises

1.

Ans- Given :

% of units of carbon in compound = $\theta$ (actual parameter)

Literature is the prior

belief that $\theta$ is from Gaussian Distribution,

with mean = 25, std. deviation = 3

i.e. $\theta \sim (25, 3)$

Measuring Device is the posterior

sample y from Gaussian Distribution,

with mean = $\theta$, std. deviation = 5

i.e. $y = 28 \sim (\theta, 5)$

(a)

We know that :

Baye's Rule is $\boxed{p(\theta|y) = \frac{p(y|\theta) \cdot p(\theta)}{p(y)}}$

~~The~~ We also know that

a density function $\boxed{p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}}$ where $\mu$ = sample mean, $\sigma$ = sample std. deviation prior $p(\theta)$ given from $\theta \sim (25, 3)$

$$\therefore p(\theta) = \frac{1}{\sqrt{2\pi}(3)} e^{-\left(\frac{(\theta-25)^2}{2(3)^2}\right)}$$

$$= \frac{1}{3\sqrt{2\pi}} e^{-\left(\frac{(\theta-25)^2}{18}\right)}$$

likelihood $p(y|\theta)$ given from $y = 28 \sim (\theta, 5)$

$$\therefore p(y|\theta) = \frac{1}{\sqrt{2\pi}(5)} e^{-\left(\frac{(\theta-28)^2}{2(5)^2}\right)}$$

$$= \frac{1}{5\sqrt{2\pi}} e^{-\left(\frac{(\theta-28)^2}{50}\right)}$$

Also,

$$p(y) = \int_{-\infty}^{\infty} p(y|\theta) p(\theta)$$

$\therefore$ Baye's Rule:

$$p(y|\theta) = \frac{\frac{1}{3\sqrt{2\pi}} e^{-\left(\frac{(\theta-25)^2}{18}\right)} \cdot \frac{1}{5\sqrt{2\pi}} e^{-\left(\frac{(28-\theta)^2}{50}\right)}}{\int_{-\infty}^{\infty} \frac{1}{3\sqrt{2\pi}} e^{-\left(\frac{(\theta-25)^2}{18}\right)} \cdot \frac{1}{5\sqrt{2\pi}} e^{-\left(\frac{(28-\theta)^2}{50}\right)} d\theta}$$

Since evaluating integrals is difficult, we will consider the MAP estimate

MAP estimate assumes that $p(\theta|y)$ has a narrow peak around its mode, thus reducing it to a single point.

For MAP estimation, we don't have to consider the denominator $p(y)$, i.e, it is not necessary to calculate it

$$\therefore \hat{\theta}_{MAP} = \underset{\theta}{\operatorname{argmax}}\ p(\theta|y)$$

Let us calculate $p(\theta|y)$ first

$$p(\theta|y) = p(y|\theta) \cdot p(\theta) \quad (\because \text{MAP estimate})$$

$$= \frac{1}{5\sqrt{2\pi}} e^{-\left(\frac{(28-\theta)^2}{50}\right)} \cdot \frac{1}{3\sqrt{2\pi}} e^{-\left(\frac{(25-\theta)^2}{18}\right)}$$

$$= \frac{1}{15(\sqrt{2\pi})^2} e^{-\frac{(784+\theta^2-56\theta)}{50} - \frac{(625+\theta^2-50\theta)}{18}}$$

To make our computation easier, we will apply log
The $\theta$ that makes $p(\theta|y)$ max, also makes $\log p(\theta|y)$ max.

$$\therefore \log p(\theta|y) = \log(1) - \log(15) - \log(2\pi) - \frac{(784+\theta^2-56\theta)}{50} - \frac{(625+\theta^2-50\theta)}{18}$$

$$\theta_{MAP} = \underset{\theta}{\operatorname{argmax}} \log p(\theta|y)$$

$$= \underset{\theta}{\operatorname{argmax}}\left[0 - \log(15) - \log(2\pi) - \frac{(784+\theta^2-56\theta)}{50} - \frac{(625+\theta^2-50\theta)}{18}\right]$$

Since it is $\underset{\theta}{\operatorname{argmax}}$, we take the partial derivative of above and set it to 0.

$\theta_{MAP}$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(0 - \log(15) - \log(2\pi) - \frac{(784+\theta^2-56\theta)}{50} - \frac{(625+\theta^2-50\theta)}{18}\right) = 0$$

$$\Rightarrow 0 - 0 - 0 - \frac{1}{50}(0+2\theta-56) - \frac{1}{18}(0+2\theta-50) = 0$$

$$\Rightarrow -\frac{1}{50}(2\theta-56) - \frac{1}{18}(2\theta-50) = 0$$

$$\Rightarrow \frac{1}{50}(56-2\theta) + \frac{1}{18}(50-2\theta) = 0$$

$$\Rightarrow \frac{9(56-2\theta)+25(50-2\theta)}{450} = 0$$

$$\Rightarrow 9(56-2\theta)+25(50-2\theta) = 0$$

$$\Rightarrow 504 - 18\theta + 1250 - 50\theta = 0$$

$$\Rightarrow 1754 - 68\theta = 0$$

$$\Rightarrow -68\theta = -1754$$

$$\Rightarrow \theta = \frac{\cancel{-}1754}{\cancel{-}68}$$

$$= 25.7941176471$$

$$\therefore \theta_{MAP} = 25.7941176471$$

(b)

Given that:

prior : 28 ± 5 , i.e,

$$\theta \sim N(28, 5)$$

data : 25, 3 , ie,

$$y = 25 \sim N(\theta, 3)$$

Calculating $\theta_{MAP}$ in this case, we get

$\theta_{MAP} = \underset{\theta}{\operatorname{argmax}}\, p(\theta | y)$ where:

$$p(\theta|y) = p(y|\theta) \cdot p(y)$$

$$= \frac{1}{3\sqrt{2\pi}} e^{-\left(\frac{(25-\theta)^2}{18}\right)} \cdot \frac{1}{5\sqrt{2\pi}} e^{-\left(\frac{(\theta-28)^2}{50}\right)}$$

$$= \frac{1}{15(2\pi)} e^{-\left(\frac{625+\theta^2-50\theta}{18}\right) - \left(\frac{\theta^2+784-56\theta}{50}\right)}$$

Applying log, we get

$$\log p(\theta|y) = 0 - \log(15) - \log(2\pi) - \left(\frac{625+\theta^2-50\theta}{18}\right) - \left(\frac{\theta^2+784-56\theta}{50}\right)$$

Applying partial derivation to above and setting to 0, we get:

$$\theta_{MAP} \Rightarrow \frac{\partial}{\partial\theta} \log p(\theta | y) = 0$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(-\log(15)-\log(2\pi)\right)-\left(\frac{625+\theta^2-50\theta}{18}\right)-\left(\frac{\theta^2+784-56\theta}{50}\right)=0$$

$$\Rightarrow 0-0-\frac{1}{18}(0+2\theta-50)-\frac{1}{50}(2\theta-56)=0$$

$$\Rightarrow \frac{50-2\theta}{18}+\frac{56-2\theta}{50}=0$$

$$\Rightarrow \frac{25(50-2\theta)+9(56-2\theta)}{450}=0$$

$$\Rightarrow 1250-50\theta+504-18\theta=0$$

$$\Rightarrow -68\theta=-1754$$

$$\Rightarrow \theta=\frac{-1754}{-68}$$

$$=25.7941176471$$

$\therefore \theta_{MAP}=25.7941176471$ even in this case

$\therefore$ Both the $\theta_{MAP}$ are <u>same</u>.

This happens ~~happen~~ because we are only considering one sample. $\therefore$ ~~Bo~~ prior and data produces the same resulting as contribution is same.

(c)

Given that:

prior (literature):

$\theta \sim N(\mu_1, \sigma_1)$

data (experiment)

$y \sim N(\theta, \sigma_2)$

W.K.T Baye's Rule:

$$p(y|\theta) = \frac{p(\theta|y) \cdot p(\theta)}{p(y)}$$

For $\theta_{MAP} = \underset{\theta}{\operatorname{argmax}}\ p(y|\theta)$ (Not considering the denominator)

$$= \underset{\theta}{\operatorname{argmax}}\ p(\theta) \cdot p(\theta|y)$$

$$= \underset{\theta}{\operatorname{argmax}} \left[ \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\left(\frac{(\theta-\mu_1)^2}{2\sigma_1^2}\right)} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\left(\frac{(y-\theta)^2}{2\sigma_2^2}\right)} \right]$$

$$= \underset{\theta}{\operatorname{argmax}} \left[ \frac{1}{(\sqrt{2\pi})^2 \sigma_1 \sigma_2} e^{-\left(\frac{\theta^2+\mu_1^2-2\theta\mu_1}{2\sigma_1^2}\right) - \left(\frac{y^2+\theta^2-2y\theta}{2\sigma_2^2}\right)} \right]$$

$$= \underset{\theta}{\operatorname{argmax}} \left[ \frac{1}{2\pi\sigma_1\sigma_2} e^{\left(\frac{2\theta\mu_1-\theta^2-\mu_1^2}{2\sigma_1^2}\right) + \left(\frac{2y\theta-y^2-\theta^2}{2\sigma_2^2}\right)} \right]$$

Applying log, we get

$$= \underset{\theta}{\operatorname{argmax}} \log \left[ \frac{1}{2\pi\sigma_1\sigma_2} e^{\left(\frac{2\theta\mu_1 - \theta^2 - \mu_1^2}{2\sigma_1^2}\right) + \left(\frac{2y\theta - y^2 - \theta^2}{2\sigma_2^2}\right)} \right]$$

$$= \underset{\theta}{\operatorname{argmax}} \left[ \log(1) - \log(2\pi\sigma_1\sigma_2) + \frac{2\theta\mu_1 - \theta^2 - \mu_1^2}{2\sigma_1^2} + \frac{2y\theta - y^2 - \theta^2}{2\sigma_2^2} \right]$$

Since $\underset{\theta}{\operatorname{argmax}}$, we apply partial derivative w.r.t $\theta$ and set to 0

$\therefore \theta_{MAP}$

$$\Rightarrow \frac{\partial}{\partial\theta} \left[ -\log(2\pi\sigma_1\sigma_2) + \frac{2\theta\mu_1 - \theta^2 - \mu_1^2}{2\sigma_1^2} + \frac{2y\theta - y^2 - \theta^2}{2\sigma_2^2} \right] = 0$$

$$\Rightarrow 0 + \frac{1}{2\sigma_1^2}(2\mu_1 - 2\theta) + \frac{1}{2\sigma_2^2}(2y - 2\theta) = 0$$

$$\Rightarrow \frac{\cancel{2}\mu_1}{\cancel{2}\sigma_1^2} - \frac{\cancel{2}\theta}{\cancel{2}\sigma_1^2} + \frac{\cancel{2}y}{\cancel{2}\sigma_2^2} - \frac{\cancel{2}\theta}{\cancel{2}\sigma_2^2} = 0$$

$$\Rightarrow \frac{\mu_1}{\sigma_1^2} - \frac{\theta}{\sigma_1^2} + \frac{y}{\sigma_2^2} - \frac{\theta}{\sigma_2^2} = 0$$

$$\Rightarrow \frac{-\theta}{\sigma_1^2} - \frac{\theta}{\sigma_2^2} = \frac{-\mu_1}{\sigma_1^2} - \frac{y}{\sigma_2^2}$$

$$\Rightarrow \frac{\theta}{\sigma_1^2} + \frac{\theta}{\sigma_2^2} = \frac{\mu_1}{\sigma_1^2} + \frac{y}{\sigma_2^2}$$

$$\Rightarrow \theta\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right) = \left(\frac{1}{\sigma_1^2}\right)\mu_1 + \left(\frac{1}{\sigma_2^2}\right)y$$

$$\Rightarrow \theta = \frac{\left(\frac{1}{\sigma_1^2}\right)}{\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)}\mu_1 + \frac{\left(\frac{1}{\sigma_2^2}\right)}{\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)} y$$

$$\Rightarrow \theta_{MAP} = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}\, y + \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2}\, \mu_1$$

Thus, this is the generalized equation
It corresponds to the special formula from the text and the class

$$\theta_{BAYES} = E[\theta|y]$$

$$= \frac{N/\sigma_2^2}{1/\sigma_1^2 + N/\sigma_2^2}\, y + \frac{1/\sigma_1^2}{1/\sigma_1^2 + N/\sigma_2^2}\, \mu_1$$

where $N = 1$

2.

Ans- Given that:

$$P_v = P(X = v)$$

$$|V| = 3 \, , \, m = 0.1$$

$$\hat{p}_v = \frac{N_v + m}{N + |V| m} = \frac{N_v + 0.1}{N + 0.3}$$

W.K.T

Bias = estimate - actual

$$= E[\hat{p}_v] - p_v = E\left[\frac{N_v + 0.1}{N + 0.3}\right] - p_v = \frac{E[N_v] + E[0.1]}{E[N] + E[0.3]} - p_v$$

$$= \frac{E[N_v] + 0.1}{N + 0.3} - p_v$$

As it is given that $N_v$ is a random variable and $y_i$ is Bernoulli Random Variable.

$$\therefore E[N_v] = E[y_1 + y_2 + \dots + y_N] = N \cdot p_v$$

$$\therefore \text{Bias} = \frac{N p_v + 0.1}{N + 0.3} - p_v = \frac{\cancel{N p_v} + 0.1 - \cancel{N p_v} - 0.3 p_v}{N + 0.3}$$

$$\boxed{\Rightarrow \text{Bias} = \frac{0.1 - 0.3 p_v}{N + 0.3}}$$

3.

Ans- Given that:

purchase approved but stolen card = F.P
purchase not approved but not stolen card = F.N
purchase not approved, stolen card = T.N
purchase approved, card not stolen = T.P

Let:

Action $\alpha$ - {Approved, Not Approved}

Cost associated with Risk - $\lambda$

Class C - {Stolen, Not Stolen}

$\lambda$ for F.P = \$ $x$
$\lambda$ for T.P = \$ - (2% of \$ $x$)
$\lambda$ for T.N = \$ 0
$\lambda$ for F.N = \$ $5x$

| Action | Class | |
|---|---|---|
| $\alpha$ | Stolen | Not Stolen |
| Approved | F.P | T.P |
| Not Approved | T.N | F.N |

Also, mentioned:

$\$ x = \$350$

$\$(2\% \text{ of } \$x) = -\$7$

$\$5x = \$1750$

In this example, therefore:

| | Class | |
|---|---|---|
| Action $\alpha$ | Stolen $C_1$ | Not Stolen $C_2$ |
| Approved $\alpha_1$ | \$ 350 | -\$7 |
| Not Approved $\alpha_2$ | \$0 | \$ 1750 |

Also, $P(C_1 = \text{Stolen}) = 0.26$

(a) Cost to the company if purchase is approved and card is stolen
From the above table, it is clear that

$$\boxed{\lambda_{11} = \$350}$$

(b) Similarly,
Cost to the company if purchase is approved and card is not stole

$$\boxed{\lambda_{12} = -\$7}$$

(c) if P(Stolen) = 0.26, expected cost to company if purchased is approved.

W.K.T

$$R(\alpha_i | x) = \sum_{k=1}^{K} \lambda_{ik} P(C_k | x)$$

in this case,

$R(\alpha_1 = \text{Approved} \mid 350)$

$= 350 \cdot P(\text{Stolen} \mid 350) + (-7) \cdot P(\text{Not Stolen} \mid 350)$

$= 350\,(0.26) - 7\,(0.74)$ ($\because$ P(Not stolen) = 1 - P(stolen) = 1 - 0.26 = 0.74)

$= 91 - 5.18$

$= \$\,85.82$

$$\boxed{\therefore R(\text{Approved} \mid 350) = \$\,85.82}$$

(d) Similarly,

$R(\alpha_2 = \text{Not Approved} \mid 350)$

$= 0 \cdot P(C_1 | 350) + 1750\, P(C_2 | 350)$

$= 0 + 1750(0.74)$

$= \$1295$

$\therefore$ $\boxed{R(\text{Not Approved} \mid 350) = \$1295}$

(e)

Ans- As $R(\text{Approved} \mid 350) < R(\text{Not Approved} \mid 350)$,

i.e., $\$85.92 < \$1295$,

the decision of approving minimizes the company's risk

4.

Ans- Given :

| Attributes | | Output (dependent) |
|---|---|---|
| $x_1$ | $x_2$ | $r$ |
| 1 | 3 | 3 |
| 4 | 2 | 7 |
| 5 | 4 | 2 |
| 8 | 7 | 1 |

(a) $g(x) = w_k x_k + w_{k-1} x_{k-1} + \dots + w_1 x_1 + w_0$

$$D = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \dots & x_k^1 \\ 1 & x_1^2 & x_2^2 & \dots & x_k^2 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_1^N & x_2^N & \dots & x_k^N \end{bmatrix} \qquad r = \begin{bmatrix} r^1 \\ r^2 \\ \dots \\ r^N \end{bmatrix}$$

(a)

$$\therefore D = \begin{array}{c} \begin{matrix} x_1 & x_2 & y \end{matrix} \\ \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \\ 1 & 5 & 4 \\ 1 & 8 & 7 \end{bmatrix}_{4\times3} \end{array}$$

It's transpose is :-

$$D^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix}_{3\times4}$$

$$\therefore (D^T \cdot D) = D^T_{3\times4} * D_{4*3}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix}_{3\times4} * \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \\ 1 & 5 & 4 \\ 1 & 8 & 7 \end{bmatrix}_{4\times3}$$

$$= \begin{bmatrix} 4 & 18 & 16 \\ 18 & 106 & 87 \\ 16 & 87 & 78 \end{bmatrix}$$

$$\boxed{\therefore (D^T \cdot D) = \begin{bmatrix} 4 & 18 & 16 \\ 18 & 106 & 87 \\ 16 & 87 & 78 \end{bmatrix}_{3\times3}}$$

Now, $(D^T \cdot D)^{-1}$, ie taking the inverse, we get:

$$\boxed{(D^T D)^{-1} = \begin{bmatrix} \frac{699}{500} & -\frac{3}{125} & -\frac{13}{50} \\ -\frac{3}{125} & \frac{14}{125} & -\frac{3}{25} \\ -\frac{13}{50} & -\frac{3}{25} & \frac{1}{5} \end{bmatrix}_{3\times 3}}$$

$$\underset{}{w} = \underset{3\times3}{(D^T D)^{-1}} \cdot \underset{3\times4}{D^T} \cdot \underset{4\times1}{\gamma}$$

$$= \begin{bmatrix} 699/500 & -3/125 & -13/50 \\ -3/125 & 14/125 & -3/25 \\ -13/50 & -3/25 & 1/5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \\ 3 & 2 & 4 & 7 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 7 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3559/500 \\ 77/125 \\ -83/50 \end{bmatrix} \begin{matrix} - w_0 \\ - w_1 \\ - w_2 \end{matrix}$$

$$\boxed{\therefore\ w = \begin{bmatrix} 3559/500 \\ 77/125 \\ -83/50 \end{bmatrix}}$$

$\therefore$ Resulting linear equation

$$g(x) = w_2 x_2 + w_1 x_1 + w_0$$

$$\Rightarrow g(x) = \left(-\frac{83}{50}\right)x_2 + \left(\frac{77}{125}\right)x_1 + \left(\frac{3559}{500}\right)$$

$$\boxed{\therefore g(x) = -1.66x_2 + 0.616x_1 + 7.118}$$

(b)

Taking ~~t~~ only the above 2 examples,

$$D = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \end{bmatrix}_{2\times 3} \qquad D^T = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 2 \end{bmatrix}_{3\times 2}$$

$$D^T \cdot D = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 3 & 2 \end{bmatrix}_{3\times 2} \cdot \begin{bmatrix} 1 & 1 & 3 \\ 1 & 4 & 2 \end{bmatrix}_{2\times 3}$$

$$\boxed{D^T \cdot D = \begin{bmatrix} 2 & 5 & 5 \\ 5 & 17 & 11 \\ 5 & 11 & 13 \end{bmatrix}}$$

$(D^T \cdot D)^{-1}$ does not exist as $\det(D^T \cdot D) = 0$

$\therefore (D^T \cdot D)$ is not invertible

Also we can say that two examples cannot solve an equation with three variables.

(c) Regularized error function,

$$E_{2,r} := \frac{1}{2}\left(\sum_t \frac{1}{2}\left(g(x^t) - r^t\right)^2 + \lambda \sum_{i=1}^{2} w_i^2\right)$$

$$\Rightarrow \frac{1}{2}\left[\left(\frac{[(3w_2+w_1+w_0)-(3)]^2}{2} + \frac{[(2w_2+4w_1+w_0)-(7)]^2}{2} + \frac{[(4w_2+5w_1+w_0)-(2)]^2}{2} + \frac{[(7w_2+8w_1+w_0)-(1)]^2}{2}\right) + \lambda(w_1^2+w_2^2)\right]$$

$$\Rightarrow \frac{1}{2}\left[\frac{(9w_2^2+w_1^2+w_0^2+6w_1w_2+2w_0w_1+6w_0w_2+9-18w_2-6w_1-6w_0)}{2}\right.$$

$$+\frac{(4w_2^2+16w_1^2+w_0^2+16w_1w_2+8w_0w_1+4w_0w_2+49-28w_2-56w_1-14w_0)}{2}$$

$$+\frac{(16w_2^2+25w_1^2+w_0^2+40w_1w_2+10w_0w_1+8w_0w_2+4-16w_2-20w_1-4w_0)}{2}$$

$$\left.+\frac{(49w_2^2+64w_1^2+w_0^2+112w_1w_2+16w_0w_1+14w_0w_2+1-14w_2-16w_1-2w_0)}{2} + \lambda w_1^2 + \lambda w_2^2\right]$$

$$= \frac{1}{2}\left[\frac{78\omega_2^2+106\omega_1^2+4\omega_0^2+174\omega_1\omega_2+36\omega_0\omega_1+32\omega_0\omega_2+63 - 76\omega_2-98\omega_1-26\omega_0+2\lambda\omega_1^2+2\lambda\omega_2^2}{2}\right]$$

$$= \frac{1}{2}\left[\frac{(78+2\lambda)\omega_2^2+(106+2\lambda)\omega_1^2+4\omega_0^2+174\omega_1\omega_2 +36\omega_0\omega_1+32\omega_0\omega_2+63-76\omega_2-98\omega_1-26\omega_0)}{2}\right]$$

$$= \frac{1}{2}\left[(39+\lambda)\omega_2^2+(53+\lambda)\omega_1^2+2\omega_0^2+87\omega_1\omega_2+18\omega_0\omega_1 +16\omega_0\omega_2+\frac{63}{2}-38\omega_2-49\omega_1-13\omega_0\right]$$

$$\therefore E_{2,\gamma} := \frac{(39+\lambda)}{2}\omega_2^2+\frac{(53+\lambda)}{2}\omega_1^2+\omega_0^2+\frac{87}{2}\omega_1\omega_2+9\omega_0\omega_1 +8\omega_0\omega_2+\frac{63}{4}-19\omega_2-\frac{49}{2}\omega_1-\frac{13}{2}\omega_0$$

(d) Setting $\lambda = 2$, we get:

$$E_{2,\gamma} := \frac{41}{2}\omega_2^2 + \frac{55}{2}\omega_1^2 + \omega_0^2 + \frac{87}{2}\omega_1\omega_2 + 9\omega_0\omega_1 + 8\omega_0\omega_2$$

$$+ \frac{63}{4} - 19\omega_2 - \frac{49}{2}\omega_1 - \frac{13}{2}\omega_0$$

Taking partial derivatives w.r.t $\omega_0, \omega_1, \omega_2$, we get:

$$\frac{\partial E}{\partial \omega_0} = 0 + 0 + 2\omega_0 + 0 + 9\omega_1 + 8\omega_2 + 0 - 0 - 0 - \frac{13}{2}$$

$$\boxed{\therefore \frac{\partial E}{\partial \omega_0} = 2\omega_0 + 9\omega_1 + 8\omega_2 - \frac{13}{2}}$$

$$\frac{\partial E}{\partial \omega_1} = 0 + 55\omega_1 + 0 + \frac{87}{2}\omega_2 + 9\omega_0 + 0 + 0 - 0 - \frac{49}{2}$$

$$\boxed{\therefore \frac{\partial E}{\partial \omega_0} = 9\omega_0 + 55\omega_1 + \frac{87}{2}\omega_2 - \frac{49}{2}}$$

$$\frac{\partial E}{\partial \omega_2} = 41\omega_2 + 0 + 0 + \frac{87}{2}\omega_1 + 0 + 8\omega_0 + 0 - 19 - 0 - 0$$

$$\boxed{\frac{\partial E}{\partial \omega_2} = 8\omega_0 + \frac{87}{2}\omega_1 + 41\omega_2 - 19}$$

(e) Solving

$$\left.\begin{aligned} 2w_0 + 9w_1 + 8w_2 - \frac{13}{2} &= 0 \\ 9w_0 + 55w_1 + \frac{87}{2}w_2 - \frac{49}{2} &= 0 \\ 8w_0 + \frac{87}{2}w_1 + 41w_2 - 19 &= 0 \end{aligned}\right\}$$

by using np.linalg.solve, we get:

$w_0 = 6.20791246$

$w_1 = 0.13131313$

$w_2 = -0.88720539$

$\therefore$ The equation of the linear function that minimizes the regularized error function $E_{2,\gamma}$ error on this data set is:

$$\boxed{g(x) = -0.8872x_2 + 0.1313x_1 + 6.2079}$$