$$L_{022}(B) = (Y - \hat{Y})^{T}(Y - \hat{Y}) \qquad \hat{Y} = \hat{g}^{T} X$$

$$= (Y - \hat{g}^{T} X)^{T}(Y - \hat{g}^{T} X)$$

$$= (Y - \hat{g}^{T} X)^{T}(Y - \hat{g}^{T} X)$$

$$= Y^{T} Y + X^{T} \hat{g} \hat{g}^{T} X - 2X^{T} \hat{g}^{T} Y$$

$$\frac{\partial L_{022}(B)}{\partial B} = 2X^{T} X \hat{g} - 2X^{T} Y = 0$$

$$\Rightarrow X^{T} X \hat{g} = X^{T} Y$$

$$\Rightarrow \hat{g} = (X^{T} X)^{T} X^{T} Y$$

```
import numpy as np
data=np.genfromtxt("RegressionExample.txt", delimiter=' ',dtype =float)
x=data[:,1:]
y=data[:,0]

def closed_form_solution(y,x):
    y = y.reshape(y.shape[0],1)
    xtxi = np.linalg.inv((x.T@x))
    return xtxi@(x.T)@y

B_head = closed_form_solution(y,x)
print(B_head)
```

 $\frac{3}{c_{MP}} = \underset{\text{arg max}}{c_{MP}} \left\{ P(c|X) \right\} = \underset{\text{arg max}}{arg max} \left\{ P(c) + \underset{\text{ln}}{ln} \left(|\Sigma_c| \right) - (x_{-}u_c)^T \Sigma^{-1} (x_{-}u_c)^T \right\} \\
= \underset{\text{arg max}}{arg max} \left\{ \left(-\underset{\text{ln}}{ln} |P(c)| + \left| -\underset{\text{ln}}{ln} |P(x|c) \right| \right\} = c \\
= \underset{\text{arg min}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \left| -\underset{\text{ln}}{ln} |P(x|c) \right| \right\} = c \\
= \underset{\text{arg min}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \left| -\underset{\text{ln}}{ln} |P(x|c) \right| \right\} = c \\
= \underset{\text{arg min}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \left| -\underset{\text{ln}}{ln} |P(x|c) \right| \right\} = c \\
= \underset{\text{arg min}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \left| -\underset{\text{ln}}{ln} |P(x|c) \right| \right\} = c \\
= \underset{\text{arg min}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{arg min} \left\{ -\underset{\text{ln}}{ln} (P(c)) + \underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
= \underset{\text{ln}}{ln} \left\{ -\underset{\text{ln}}{ln} |P(c)| \right\} = c \\
=$

= avgmin[-ln(P(u)) + 0.5 ln ≥c[+0.5 (x-luc)] = lx2 = avgmin[-2ln(P(u)) + ln(|≥c|) + (x-luc)] = (x-luc)]

 $J = \sum_{k=1}^{n} w_{1}w - \sum_{k=1}^{n} x_{k} \left(y_{1}(w_{1}x_{1}) + \left(\sum_{k=1}^{n} x_{k} - \sum_{k=1}^{n} \beta_{1} y_{k} \right) \right)$ $\frac{\partial J}{\partial w} = W - \sum_{k=1}^{n} \alpha_{1}y_{1} x_{1} = 0 \quad \sum_{k=1}^{n} \alpha_{1}y_{1} = 0$ $W = \sum_{k=1}^{n} \alpha_{1}y_{1} x_{1} \qquad J = \underbrace{\sum_{k=1}^{n} \sum_{j=1}^{n} \alpha_{j}\alpha_{1}y_{1}y_{j}}_{A_{2}} x_{j}^{T}x_{1} - \sum_{k=1}^{n} \alpha_{1}y_{1} \sum_{j=1}^{n} \alpha_{1}y_{1} x_{j}^{T}x_{2} - \sum_{k=1}^{n} \alpha_{1}y_{1} x_{j}^{T}x_{2} + \sum_{k=1}^{n} \alpha_{1}y_{1}^{T}x_{2} + \sum_{$