1)

$$\frac{\partial}{\partial b} \underbrace{\sum_{i=1}^{N} (w^T x_i + b - y_i)^2} = 0$$

$$= \frac{\partial}{\partial b} \left[\cancel{\sum_{i=1}^{N} (w^T x_i)^2} + \sum b^2 + \cancel{\sum (y_i)^2} + 2 \sum w^T x_i b - 2 \sum b y_i \right]$$

$$= 2 \cdot b \cdot N + 2 \sum w^T x_i - 2 \sum y_i = 2bN + 2 \sum w^T x_i - y_i = 0$$

$$bN = \underbrace{\sum y_i}_{0 \text{ since } \#y_1 = \#y_{-1}} - w^T \sum x_i \Rightarrow \quad \mu_{-1} = \frac{1}{N_{-1}} \sum_{i: y_i = -1} x_i \Rightarrow (N_{-1}) \mu_{-1} = \sum_{i: y_i = -1} x_i$$

$$\mu_{+1} = \frac{1}{N_{+1}} \sum_{i: y_i = 1} x_i$$

$$b \cancel{N} = - w^T \frac{\cancel{N}}{2} (\mu_{-1} + \mu_{+1}) \Rightarrow \boxed{\hat{b} = -\frac{1}{2} w^T (\mu_{-1} + \mu_{+1})}$$

2) $$\frac{\partial}{\partial w} \sum (w^T x_i + b - y_i)^2 = 0$$

$$= \frac{\partial}{\partial w} \left[\sum (w^T x_i)^2 + 2 \sum w^T x_i b - 2 \sum w^T x_i y \right] = 0 \qquad \left((b^T \cdot c) \cdot a = (a \cdot c^T) \cdot b \right)$$

$$= \cancel{2} \sum w^T x_i x_i^T + \cancel{2} \sum x_i b = \cancel{2} \sum x_i y_i = \sum \left[w^T x_i^2 - \frac{1}{2} w^T (\mu_{-1} + \mu_{+1}) x_i \right] = \sum x_i y_i$$

$$= \sum w^T \left(x_i - \frac{1}{2} (\mu_{-1} + \mu_{+1}) \right) x_i = \sum x_i \left(x_i - \frac{1}{2} (\mu_{-1} + \mu_{+1}) \right)^T w = \sum x_i y_i$$

$$\sum x_i y_i = \sum_{i: y_i = 1} x_i - \sum_{i: y_i = -1} x_i = \frac{N}{2} (\mu_{+1} - \mu_{-1})$$

$$\Rightarrow \left[\sum x_i x_i^T - \frac{1}{2} (\mu_{-1} + \mu_{+1})^T \underbrace{\sum x_i}_{\frac{N}{2}(\mu_{-1} + \mu_{+1})} \right] \cdot w = \frac{N}{2} (\mu_{+1} - \mu_{-1})$$

$$= w \left[\sum x_i x_i^T - \frac{N}{4} (\mu_{-1} + \mu_{+1}) (\mu_{-1} + \mu_{+1})^T \right] = \frac{N}{4} \left[\mu_{-1} \mu_{-1}^T + \mu_{-1} \mu_{+1}^T + \mu_{+1} \mu_{-1}^T + \mu_{+1} \mu_{+1}^T \right]$$

$$= w\left[\sum x_i x_i^T - \frac{N}{4}\left(\mu_{-1}\mu_{-1}^T + \mu_{-1}\mu_{+1}^T + \mu_{+1}\mu_{-1}^T + \mu_{+1}\mu_{+1}^T\right)\right] = \frac{N}{2}(\mu_{+1} - \mu_{-1})$$

$$S_B = \mu_{+1}\mu_{+1}^T - \mu_{+1}\mu_{-1}^T - \mu_{-1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T \qquad S_w = \frac{1}{N}\sum^{N}(x_i - \mu_{y_i})(x_i - \mu_{y_i})^T$$

$$S_w = \frac{1}{N}\sum\Big[x_i x_i^T - \underbrace{\mu_{y_i} x_i^T}_{\frac{N}{2}(\mu_{-1}+\mu_{+1})} - x_i \mu_{y_i}^T + \mu_{y_i}\mu_{y_i}^T\Big] = \frac{1}{N}\sum x_i x_i^T - \frac{N}{2}\left(\mu_1\mu_{y_i}^T + \mu_1\mu_{y_i}^T + \mu_{y_i}\mu_{-1}^T + \mu_{y_i}\mu_{+1}^T\right) + \mu_{y_i}\mu_{y_i}^T$$

$$= \frac{1}{N}\Big[\sum x_i x_i^T + \frac{1}{4}\cdot S_B - \frac{1}{2}\left(\mu_{+1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T\right)\Big] w = \frac{\mu_{+1} - \mu_{-1}}{2}$$

$$\frac{1}{N}\Big[\frac{N}{2}\left(\mu_{+1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T\right) - \underbrace{\mu_{+1}\frac{N}{2}\mu_{+1}^T - \mu_{-1}\frac{N}{2}\mu_{-1}^T}\ \underbrace{- \mu_{+1}\frac{N}{2}\mu_{+1}^T - \mu_{-1}\, N/2\, \mu_{-1}^T}\Big]$$

$$= \frac{1}{N}\sum \mu_{y_i}\mu_{y_i}^T - \frac{1}{N}\sum \mu_{y_i} x_i^T - \frac{1}{N}\sum x_i \mu_i^T$$

$$= \frac{1}{N}\Big[\sum x_i x_i^T - \mu_{y_i} x_i^T - x_i \mu_{y_i}^T + \mu_{y_i}\mu_{y_i}^T + \frac{1}{4} S_B\Big]\hat{w} = \frac{(\mu_{+1} - \mu_{-1})}{2} = \left(S_w + \frac{S_B}{4}\right)\hat{w} = \frac{\mu_1 - \mu_{-1}}{2}$$

$$= \frac{1}{N}\Big[$$

$$\cancel{x_i x_i^T} - \frac{1}{2}\mu_{+1}\mu_{+1}^T - \frac{1}{2}\mu_{-1}\mu_{-1}^T = \cancel{x_i x_i^T} - \mu_{y_i} x_i^T - x_i\mu_{y_i}^T + \mu_{y_i}\mu_{y_i}^T$$

$$\mu_{+1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T - \mu_{+1}\mu_{+1}^T - \mu_{-1}\mu_{-1}^T - \frac{1}{2}\left(\mu_{+1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T\right)$$

$$\underbrace{\frac{N}{2}\left(\mu_{+1}\mu_{+1}^T + \mu_{-1}\mu_{-1}^T\right)}_{\sum \mu_{y_i}\mu_{y_i}^T} \underbrace{- \mu_{+1}\frac{N}{2}\mu_{+1}^T - \mu_{-1}\frac{N}{2}\mu_{-1}^T - \mu_{+1}\frac{N}{2}\mu_{+1}^T - \mu_{-1}\frac{N}{2}\mu_{-1}^T}_{-\sum \mu_{y_i} x_i^T + x_i\mu_{y_i}^T}$$

" END "

1.3)

$$\Big[\underset{\downarrow \atop D\times D}{S_w} + \frac{1}{4}\underset{\downarrow \atop D\times 1}{(\mu_1-\mu_{-1})}\underset{\downarrow \atop 1\times D}{(\mu_1-\mu_{-1})^T}\Big]\underset{\downarrow \atop D\times 1}{\hat{w}} = \frac{1}{2}(\mu_1-\mu_{-1})$$

$$(S_w)\hat{w} = (\mu_1-\mu_{-1})\left[\frac{1}{2} - \frac{1}{4}\underbrace{\underbrace{(\mu_1-\mu_{-1})^T\hat{w}}_{1\times D\ \cdot\ D\times 1}}_{\text{scalar}}\right]$$

$$\hat{w} = S_w^{-1}\cdot(\mu_1-\mu_{-1})\cdot J$$