# SML Assignment 1

Dongju Ma
A1942340

June 20, 2025
Trimester 2

## Question 1a: Closed-Form Solution Derivation

### Derivation

$$\nabla_W(J(W)) = \frac{\partial}{\partial W}\frac{1}{2m}||\hat{Y} - Y||_2^2 = \frac{\partial}{\partial W}\frac{1}{2m}(XW - Y)^T(XW - Y).$$

Denote that $X \in \mathbb{R}^{m\times n}, Y \in \mathbb{R}^{m\times 1}$ and $W \in \mathbb{R}^{n\times 1}$,
hence $(W^T X^T Y)^T = Y^T XW = W^T X^T Y$ due to each term is scalar.
Using the identify $(AB)^T = B^T A^T$ and expanding the square

$$J(W) = \frac{1}{2m}(XW - Y)^T(XW - Y) = \frac{1}{2m}\left(W^T X^T XW - 2W^T X^T Y + Y^T Y\right).$$

Apply partial derivatives to $J(W)$

$$\nabla_W(J(W)) = \frac{1}{2m}(2X^T XW - 2X^T Y)$$

To minimize $J(W)$, we should compare $\nabla_W(J(W))$ to 0, therefore

$$\nabla_W(J(W)) = 0 \quad \Rightarrow \quad X^T XW = X^T Y$$

multiply both sides of the equation by $(X^T X)^{-1}$, thus

$$W = (X^T X)^{-1} X^T Y$$

### Matrix Alignment Verification

Given that $X \in \mathbb{R}^{5\times 3}$ and $Y \in \mathbb{R}^{5\times 1}$, we can compute the dimension of $W$ by the following steps, Since

$$W = (X^T X)^{-1} X^T Y,$$

$$\begin{aligned} X^T \in \mathbb{R}^{3\times 5} &\Rightarrow X^T X \in \mathbb{R}^{3\times 3} \Rightarrow (X^T X)^{-1} \in \mathbb{R}^{3\times 3} \\ &\Rightarrow (X^T X)^{-1} X^T \in \mathbb{R}^{3\times 5} \\ &\Rightarrow (X^T X)^{-1} X^T Y = W \in \mathbb{R}^{3\times 1} \end{aligned}$$

Thus $\hat{Y}$ has the dimension of that

$$\hat{Y} = XW \Rightarrow \hat{Y} \in \mathbb{R}^{5\times 1}$$

Which makes $\hat{Y}$ has the same dimension with $Y$, the alignment is correct.