Part 1 - Theoretical Exercises

1) Data matrix

x	bias	y
-1	1	-1
-1	1	1
1	1	2
2	1	3

using the definition $X^{+} = (X^{T}X)^{-1}X^{T}$

$$X^{T} = \begin{pmatrix} -1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \qquad (X^{T}X) = \begin{pmatrix} 7 & 1 \\ 1 & 4 \end{pmatrix}$$

$$(X^{T}X)^{-1} = \frac{1}{|X^{T}X|} \cdot \mathrm{Adj}(X^{T}X) = \begin{pmatrix} 4/27 & -1/27 \\ -1/27 & 7/27 \end{pmatrix}$$

$$X^{+} = \begin{pmatrix} 4/27 & -1/27 \\ -1/27 & 7/27 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -5/27 & -5/27 & 1/9 & 7/27 \\ 8/27 & 8/27 & 2/9 & 5/27 \end{pmatrix}$$

$$\theta^{*} = X^{+} \cdot y = \begin{pmatrix} -5/27 & -5/27 & 1/9 & 7/27 \\ 8/27 & 8/27 & 2/9 & 5/27 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\theta^{*} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Computing the min loss, using θ^* from our calculation

$$J(\theta^*) = (0+1)^2 + (0-1)^2 + 0^2 + 0^2 = \boxed{2}$$

2) Expressing A, B, C, D, E and f

$$J(\theta) = (\theta_0 - \theta_1 + 1)^2 + (\theta_0 - \theta_1 - 1)^2 + (\theta_0 + \theta_1 - 2)^2 + (\theta_0 + 2\theta_1 - 3)^2$$

$$= \theta_0^2 + \theta_1^2 + 1 + 2(\cancel{-\theta_0\theta_1} \cancel{-\theta_1} \cancel{+\theta_0}) + \theta_0^2 + \theta_1^2 + 1 + 2(\cancel{-\theta_0\theta_1} \cancel{+\theta_1} \cancel{-\theta_0})$$

$$+ \theta_0^2 + \theta_1^2 + 4 + 2(\theta_0\theta_1 - 2\theta_1 - 2\theta_0) + \theta_0^2 + 4\theta_1^2 + 9 + 2(\cancel{2\theta_0}\theta_1 - 6\theta_1 - 3\theta_0)$$

$$= 4\theta_0^2 + 7\theta_1^2 + 15 + 2\theta_0\theta_1 - 16\theta_1 - 10\theta_0$$

$$\frac{\partial}{\partial \theta_0} = 8\theta_0 + 2\theta_1 - 10, \quad \frac{\partial}{\partial \theta_1} = 14\theta_1 + 2\theta_0 - 16$$

$$\nabla J(\theta^*) = (8 + 2 - 10, 14 + 2 - 16) = (0, 0)$$

3)

3) $x = 1.5$, bias $= 1$

$\hat{y}$ using $\theta^* = 1.5 + 1 = 2.5$

$\hat{y}$ using k-nn$_{k=2}$ $= \frac{1}{2}(2+3) = 2.5$