ECE421 Problem Set 2

Micol Altomare

1 Linear Regression

Given data:

$x^{(i)})$	$t^{(i)})$
1	6
2	4
3	2
4	1
5	3
6	6
7	10

Table 1: Given data

1.1

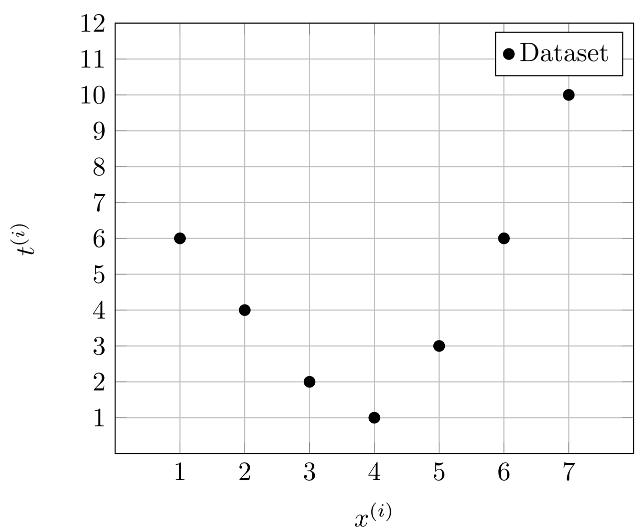

Figure 1: Scatter plot of $x^{(i)}$ vs. $t^{(i)}$

1.2

$$
\begin{aligned}
\mathcal{E}(w,b) &= \frac{1}{2N}\sum_{i=1}^{N}(y^{(i)}-t^{(i)})^2 \\
&= \frac{1}{2N}\sum_{i=1}^{N}(wx^{(i)}-t^{(i)})^2 \\
&= \frac{1}{2N}\sum_{i=1}^{N}\left[(wx^{(i)})^2 + wx^{(i)}b - wx^{(i)}t^{(i)} + wx^{(i)}b + b^2 - bt^{(i)} - wx^{(i)}t^{(i)} - t^{(i)}b + (t^{(i)})^2\right] \\
&= \frac{1}{2N}\sum_{i=1}^{N}\left[(w^2x^{(i)})^2 + 2wx^{(i)}b - 2wx^{(i)}t^{(i)} - 2bt^{(i)} + b^2 + (t^{(i)})^2\right] \\
&= \boxed{\frac{1}{2N}\sum_{i=1}^{N}\left[(x^{(i)})^2w^2 + 1b^2 + 2x^{(i)}wb - 2t^{(i)}x^{(i)}w - 2t^{(i)}b + (t^{(i)})^2\right]} \\
\implies A_i &= (x^{(i)})^2, B_i = 1, C_i = 2x^{(i)}, D_i = -2t^{(i)}x^{(i)}, E_i = -2t^{(i)}, F_i = (t^{(i)})^2
\end{aligned}
$$

in the form: $\mathcal{E}(w,b) = \frac{1}{2N}\sum_{i=1}^{N} A_iw^2 + B_ib^2 + C_iwb + D_iw + E_ib + F_i$

1.3

The loss function is minimized when $\frac{\partial\mathcal{E}}{\partial w} = 0$ and $\frac{\partial\mathcal{E}}{\partial b} = 0$. Where $A = \sum_i A_i$, $B = \sum_i B_i, C = \sum_i C_i, D = \sum_i D_i, E = \sum_i E_i$:

$$
\begin{aligned}
\frac{\partial\mathcal{E}}{\partial w} &= \frac{1}{2N}\sum_{i=1}^{N} 2wA_i + C_ib + D_i && (1)\\
&= 2wA + Cb + D = 0 && (2)\\
\implies w &= \frac{-Cb - D}{2A} && (3)\\
\frac{\partial\mathcal{E}}{\partial b} &= \frac{1}{2N}\sum_{i=1}^{N} 2B_ib + C_iw + E_i && (4)\\
&= 2Bb + Cw + E = 0 && (5)\\
\implies b &= \frac{-Cw - E}{2B} && (6)\\
\implies w &= \frac{-2Bb - E}{C} && (7)\\
\implies & \boxed{b = \frac{2AE - CD}{C^2 - 4AB}; w = \frac{2BD - CE}{C^2 - 4AB}} && (8)
\end{aligned}
$$

1.4

By plugging in numerical values from the dataset D (Table 1) and using the results in **1.2** and **1.3**, the values are found to be approximately: $b = 2.1429$ and $b = 0.6071$.

1.5

Using Excel's linear regression tool, it is found that $b = 2.1429$ and $w = 0.6071$.

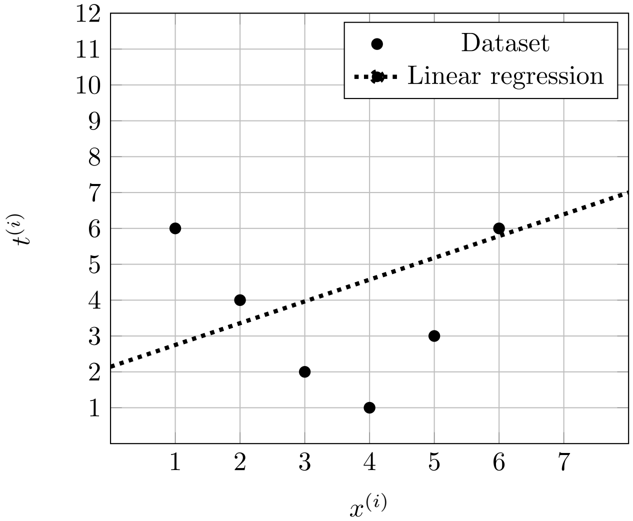

Figure 2: Scatter plot of $x^{(i)}$ vs. $t^{(i)}$

2 Least Squares

2.1

$$\begin{aligned}
g_w(\vec{x}) &= \vec{x}\vec{w} \\
&= \begin{bmatrix} x^{(i)} & 1 \end{bmatrix} \begin{bmatrix} w \\ b \end{bmatrix} \\
&= wx + (1)b \\
&= g_{w,b}(x) \\
\implies & \boxed{\vec{w} = \begin{bmatrix} w \\ b \end{bmatrix}}
\end{aligned}$$

2.2

$$\begin{aligned}
X\vec{w} - \vec{t} &= \begin{bmatrix} x^{(1)} & 1 \\ x^{(2)} & 1 \\ \vdots & \vdots \\ x^{(N)} & 1 \end{bmatrix} \begin{bmatrix} w \\ b \end{bmatrix} - \begin{bmatrix} t^{(1)} \\ t^{(2)} \\ \vdots \\ t^{(N)} \end{bmatrix} \\
&= \sum_{i=1}^{N} \begin{bmatrix} x^{(i)} & 1 \end{bmatrix} \vec{w} - t^{(i)} \\
&= \sum_{i=1}^{N} x^{(i)} w + b - t^{(i)} \\
\implies ||X\vec{w} - \vec{t}||^2 &= \sum_{i=1}^{N} (x^{(i)} w + b - t^{(i)})^2
\end{aligned}$$

$$\nabla_w ||X\vec{w} - \vec{t}||^2 = \frac{\partial ||X\vec{w} - \vec{t}||^2}{\partial w}$$

2.3

2.4

2.5

3 Problem 3

3.1

3.2

3.3

3.4

3.5

3.6