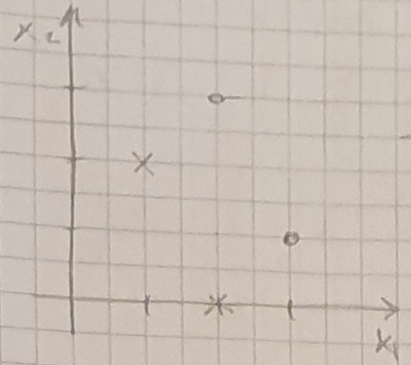
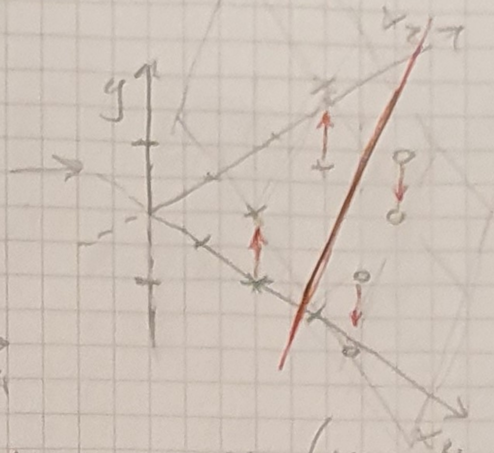
ELE 520 Exercise 3

Daniel Fylling

P1

a)





b) $g(x) = \theta^T x = y$, $x = (x_1, x_2, 1)^T$

LS-estimate for the above:

$$\hat{\theta} = (X^T X)^{-1} X^T y$$

$$X = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

Using symbolab.com to solve

$$\Rightarrow \theta = \begin{bmatrix} -\frac{4}{3} \\ -\frac{2}{3} \\ \frac{11}{3} \end{bmatrix} \Rightarrow y = -\frac{4}{3}x_1 - \frac{2}{3}x_2 + \frac{11}{3}$$

Decision boundary at $y = 0$

$$\frac{2}{3}x_2 = \frac{11}{3} - \frac{4}{3}x_1 \Rightarrow x_2 = 5.5 - 2x_1$$

c) 1. No longer possible for $g(x)$ to pass through all points

2. $g(x)$ will be less declining to accommodate for the "shallower" point.

ELE520 Exercise 3

P1 d) LMS-method

$\theta^{(1)} = [1\ 1\ 1]$, $\theta = 1$, $\mu = 0.5$, $\mu_i = \frac{\mu}{i}$

$\| \mu_i \cdot (y_i - \theta^T x_i) x_i \|$

$X = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \\ 2 & 3 & 1 \end{bmatrix}$ $y = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$ $i = 1, 5, 9$; $i = 2, 6, 10$; $i = 3, 7, 11$; $i = 4, 8, 12$

$i = 1$ $\mu_i = \frac{0.5}{1} = 0.5$

$\theta_{\frac{1}{2}} = 0.5\left(1 - \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1\\2\\1\end{bmatrix}\right)\begin{bmatrix}1\\2\\1\end{bmatrix} = 0.5 \cdot (-3)\begin{bmatrix}1\\2\\1\end{bmatrix} = \begin{bmatrix}-1.5\\-3\\-1.5\end{bmatrix}$

$\|\theta_{\frac{1}{2}}\| = 3.67 > \theta \Rightarrow$ Not satisfied - update θ_2

$\theta_2 = \theta_1 + \theta_{\frac{1}{2}} = \begin{bmatrix}-0.5\\-2\\-0.5\end{bmatrix}$

$i = 2$ $\mu_i = \frac{0.5}{2} = 0.25$

$\theta_{\frac{1}{2}} = 0.25\left(1 - \begin{bmatrix}-0.5\\-2\\-0.5\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}\right)\begin{bmatrix}2\\0\\1\end{bmatrix} = 0.25(2.5)\begin{bmatrix}2\\0\\1\end{bmatrix} = \begin{bmatrix}1.25\\0\\0.625\end{bmatrix}$

$\|\theta_{\frac{1}{2}}\| = 1.40 > \theta \rightarrow$ Not satisfied $\rightarrow$ update

$\theta_3 = \theta_2 + \theta_{\frac{1}{2}} = \begin{bmatrix}0.75\\-2\\0.125\end{bmatrix}$

$i = 3$ $\mu_i = \frac{0.5}{3}$

$\theta_{\frac{1}{2}} = \frac{0.5}{3}\left(-1 - \begin{bmatrix}0.75\\-2\\0.125\end{bmatrix}\begin{bmatrix}3\\1\\1\end{bmatrix}\right)\begin{bmatrix}3\\1\\1\end{bmatrix} = \begin{bmatrix}-0.69\\-0.23\\-0.23\end{bmatrix}$

$\|\theta_{\frac{1}{2}}\| = 0.76 > \theta \rightarrow$ Satisfied $\rightarrow$ no update

$\theta_4 = \theta_3 = \begin{bmatrix}0.75\\-2\\0.125\end{bmatrix}$

ELE520 Exercise 3

P1(d)
cont) $i = 4$ (Made python script to calculate)

$$\mu = \frac{0.5}{4} \quad \theta_{\frac{1}{2}} = \begin{bmatrix} 0,84 \\ 1,27 \\ 0,42 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 1,58 \rightarrow \theta_4 = \begin{bmatrix} 1,59 \\ -0,73 \\ 0,55 \end{bmatrix}$$

$i = 5$

$$\mu = \frac{0.5}{5} \quad \theta_{\frac{1}{2}} = \begin{bmatrix} 0,03 \\ 0,07 \\ 0,03 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 0,08 \rightarrow \theta_5 = \theta_4 \quad \text{Ok!}$$

$i = 6$

$$\mu = \frac{0.5}{6} \quad \theta_{\frac{1}{2}} = \begin{bmatrix} -0,45 \\ 0 \\ -0,23 \end{bmatrix}, \quad \|\theta_{\frac{1}{2}}\| = 0,51 \rightarrow \theta_6 = \theta_5 \quad \text{Ok!}$$

$i = 7$

$$\mu = \frac{0.5}{7} \quad \theta_{\frac{1}{2}} \begin{bmatrix} -1,20 \\ -0,40 \\ -0,40 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 1,33 \rightarrow \theta_7 = \begin{bmatrix} 0,40 \\ -1,13 \\ 0,15 \end{bmatrix}$$

$i = 8$

$$\mu = \frac{0.5}{8} \quad \theta_{\frac{1}{2}} \begin{bmatrix} 0.18 \\ 0.27 \\ 0.09 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 0,34 \rightarrow \theta_8 = \theta_7 \quad \text{Ok!}$$

$i = 9$

$$\mu = \frac{0.5}{9} \quad \theta_{\frac{1}{2}} \begin{bmatrix} 0.15 \\ 0.30 \\ 0.15 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 0.37 \Rightarrow \theta_9 = \theta_7 \quad \text{Ok!}$$

$i = 10$

$$\mu = \frac{0.5}{10} \quad \theta_{\frac{1}{2}} = \begin{bmatrix} 0.006 \\ 0 \\ 0,003 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 0,007 \rightarrow \theta_{10} = \theta_9 \quad \text{Ok!}$$

$i = 11$

$$\mu = \frac{0.5}{11} \quad \theta_{\frac{1}{2}} = \begin{bmatrix} -0.16 \\ -0,05 \\ -0,05 \end{bmatrix} \quad \|\theta_{\frac{1}{2}}\| = 0,18 \quad \theta_{11} = \theta_{10} \quad \text{Ok!}$$

All 4 datapoints are satisfied with the proposed solution of $\theta = \begin{bmatrix} 0.4 \\ 1.13 \\ 0.15 \end{bmatrix}$

ELE520 Exercise 3

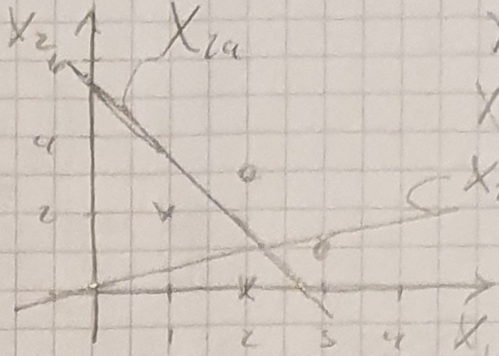
P1 d)

cont... $y = 0.4x_1 - 1.13x_2 + 0.15 = 0$

$$\Rightarrow x_2 = \frac{0.4}{1.13}x_1 + \frac{0.15}{1.13} = 0.35x_1 + 0.13$$

$x_{2a} = 5.5 - 2x_1$

$x_{2d} = 0.13 + 0.35x_1$



This does not look too good... Suspecting that threshold $\theta = 1$ is too lenient and produced a solution too easily.

P2 a)

$$\begin{bmatrix} x_{11} & x_{21} & \cdots & x_{L1} & 1 \\ x_{12} & & & x_{L2} & 1 \\ \vdots & & & \vdots & \vdots \\ x_{1N} & x_{2N} & \cdots & x_{LN} & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{L+1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

$$X\theta = y \Rightarrow \theta = X^{-1}y$$

X would have to be rectangular to compute inverse, so since we basically never have the same number of features as samples the equation is not generally solvable.

ELE520 Exercise 3

P2

b) Will find minimum point by differentiating error and setting to zero.

$$\|X\theta - y\|^2$$

$$(X\theta - y)^T(X\theta - y)$$

$$X^T\theta^T X\theta - 2X^T\theta^T y + y^T y$$

$$2\theta X^T X - 2X^T y$$

$$\theta X^T X = X^T y \quad \Rightarrow \quad \theta = (X^T X)^{-1} X^T y$$

* here the inverse of $X^T X$ will allways exist.

c) $$\|X\theta - y\|^2$$

$$(X\theta - y)^T(X\theta - y)$$

$$\theta^T X^T X\theta - y^T X\theta - \theta^T X^T y + y^T y$$

$$[(X^T X)^{-1} X^T y]^T X^T X[(X^T X)^{-1} X^T y] - 2[(X^T X)^{-1} X^T y]^T X^T y + y^T y$$

$$y^T X(X^T X)^{-1} X^T X(X^T X)^{-1} X^T y - 2y^T X(X^T X)^{-1} X^T y + y^T y$$

$$y^T X(X^T X)^{-1} I X^T y - 2y^T X(X^T X)^{-1} X^T y + y^T y$$

$$y^T(I - X(X^T X)^{-1} X^T)y$$

d) –

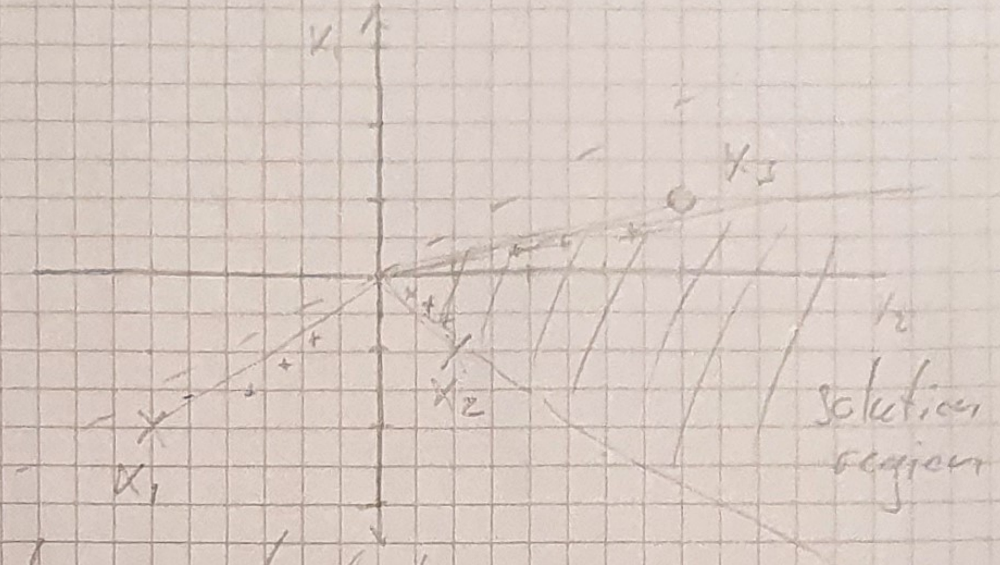
ELE520 Exercise 3

P3

a) $y = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \Rightarrow x = \begin{bmatrix} -3 & -2 \\ 1 & -1 \\ 4 & 1 \end{bmatrix} \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix}$

b)

c)



*Points normalised by y_n

d) $\theta_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\mu_1 = 1$ criterion $\theta = 1$

$\theta_2 = \begin{matrix} 0 \\ 0 \end{matrix} + \begin{matrix} -3 \\ -2 \end{matrix} + \begin{matrix} 1 \\ -1 \end{matrix} + \begin{matrix} 4 \\ 1 \end{matrix} = \begin{matrix} 2 \\ -2 \end{matrix}$

$\theta_3 = \begin{matrix} 2 \\ -2 \end{matrix} + \begin{matrix} 1 \\ -1 \end{matrix} = \begin{matrix} 3 \\ -3 \end{matrix}$

$\begin{matrix} 4 \\ -4 \end{matrix}$

$\vdots$ this went nowhere fast.

x_1 ✓
x_2 ✗
x_3 ✓
x_1 ✓
x_2 ✗
x_3 ✓

e) $\theta_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\mu_1 = \frac{1}{1}$ criterion $\theta = 1$

$\theta_2 = \begin{matrix} 2 \\ -2 \end{matrix}$

$\mu_2 = \frac{1}{2} \Rightarrow \theta_3 = \begin{matrix} 2 \\ -2 \end{matrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -2.5 \end{bmatrix}$

Hmm... for this example I could not find a convincing convergence ... maybe something is wrong.