1^{st} Homework

Chalkiopoulos Georgios | $p3352124$

January 12, 2022

Exercise 1.

(i) $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2 + x_1^2$

a. There are 4 parameters $(\theta_0, \theta_1, \theta_2, \theta_3)$

b. The dependence is Linear.

(ii) $y = sign(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2 + \theta_4 x_1^2)$

a. There are 5 parameters $(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$

b. The dependence is non-Linear.

(iii) $y = 2x_1 + sign(3-7)x_2 + ReLU(3)x_1 x_2$

a. There are no parameters, since we only have constants.

b. The dependence in Non-Linear.

(iv) $y = \theta + \theta x_1 + \theta x_2 + \theta x_1 x_2$

a. The parameter involved is one: θ

b. The dependence is linear.

Exercise 2.

1. $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2$

The model is parametric.

2. $y = min(x_1, x_2)$

The model is non-parametric, since there are no parameters

3. $y = ReLU(\theta_0 + \theta_1 x_1)$

The model is parametric.

4. $y = \sum_{i=1}^{N} \theta_i (x_{i1}x_1 - x_{i2}x_2)$

The model is non-parametric since the output depends on the size of $\boldsymbol{x}$

Exercise 3.

a) Define the parametric set of the **quadratic** functions $f_\theta : R \rightarrow R$ give two instances of it. What is the dimensionality of $\boldsymbol{\theta}$?

For a given $\boldsymbol{x} = [x_1, x_2]^T$ it is:

$$f_\theta(\boldsymbol{x}) = \theta_0 + \theta_1 x_1 + \theta_2 x_1^2, \quad \boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2]^T$$

and

$$F_{lin} := f_{\boldsymbol{\theta}}(\cdot) : \boldsymbol{\theta} \in R^3$$

i) If we choose e.g., $\boldsymbol{\theta} = [1, 2, 3]^T$ have the following instance of F_{lin}

$$f_\theta(\boldsymbol{x}) = 1 + 2x_1 + 3x_1^2$$

ii) If we choose e.g., $\boldsymbol{\theta} = [42, 42, 42]^T$ have the following instance of F_{lin}

$$f_\theta(\boldsymbol{x}) = 42 + 42x_1 + 42x_1^2$$

b) Define the parametric set of the **3$^{\text{rd}}$ degree polynomials** $f_\theta : R^2 \rightarrow R$ give two instances of it. What is the dimensionality of $\boldsymbol{\theta}$?

For a given $\boldsymbol{x} = [x_1]^T$ it is:

$$f_\theta(\boldsymbol{x}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2 + \theta_5 x_1 x_2 + \theta_6 x_1^2 x_2 + \theta_7 x_1 x_2^2 + \theta_8 x_1^3 + \theta_9 x_2^3$$

thus,

$$\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9]^T$$

and

$$F_{lin} := f_{\boldsymbol{\theta}}(\cdot) : \boldsymbol{\theta} \in R^{10}$$

i) If we choose e.g., $\boldsymbol{\theta} = [10, 1, 2, 3, 4, 5, 6, 7, 8, 9]^T$ have the following instance of F_{lin}

$$f_\theta(\boldsymbol{x}) = 10 + x_1 + 2x_2 + 3x_1^2 + 4x_2^2 + 5x_1 x_2 + 6x_1^2 x_2 + 7x_1 x_2^2 + 8x_1^3 + 9x_2^3$$

ii) If we choose e.g., $\boldsymbol{\theta} = [42, 42, 42, 42, 42, 42, 42, 42, 42, 42]^T$ have the following instance of F_{lin}

$$f_\theta(\boldsymbol{x}) = 42 + 42x_1 + 42x_2 + 42x_1^2 + 42x_2^2 + 42x_1x_2 + 42x_1^2x_2 + 42x_1x_2^2 + 42x_1^3 + 42x_2^3$$

c) Define the parametric set of the **3rd degree polynomials** $f_\theta : R^3 \to R$ give two instances of it. What is the dimensionality of $\boldsymbol{\theta}$?

For a given $\boldsymbol{x} = [x_1, x_2, x_3]^T$ it is:

$$\begin{aligned} f_\theta(\boldsymbol{x}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2 + \\ \theta_5 x_1 x_2 + \theta_6 x_1^2 x_2 + \theta_7 x_1 x_2^2 + \theta_8 x_1^3 + \theta_9 x_2^3 + \\ \theta_{10} x_3 + \theta_{11} x_3^2 + \theta_{12} x_3^3 + \theta_{13} x_1 x_3 + \theta_{14} x_1 x_3^2 + \\ \theta_{15} x_1^2 x_3 + \theta_{16} x_2 x_3 + \theta_{17} x_2 x_3^2 + \theta_{18} x_2^2 x_3 + \theta_{19} x_1 x_2 x_3 \end{aligned}$$

thus,

$$\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}, \theta_{16}, \theta_{17}, \theta_{18}, \theta_{19}]^T$$

and

$$F_{lin} := f_{\boldsymbol{\theta}}(\cdot) : \boldsymbol{\theta} \in R^{20}$$

i) If we choose e.g., $\boldsymbol{\theta} = [10, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]^T$ have the following instance of F_{lin}

$$\begin{aligned} f_\theta(\boldsymbol{x}) = 10 + 1x_1 + 2x_2 + 3x_1^2 + 4x_2^2 + \\ 5x_1x_2 + 6x_1^2x_2 + 7x_1x_2^2 + 8x_1^3 + 9x_2^3 + \\ 10x_3 + 11x_3^2 + 12x_3^3 + 13x_1x_3 + 14x_1x_3^2 + \\ 15x_1^2x_3 + 16x_2x_3 + 17x_2x_3^2 + 18x_2^2x_3 + 19x_1x_2x_3 \end{aligned}$$

ii) If we choose e.g., $\boldsymbol{\theta} = [42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42]^T$ have the following instance of F_{lin}

$$\begin{aligned} f_\theta(\boldsymbol{x}) = 42 + 42x_1 + 42x_2 + 42x_1^2 + 42x_2^2 + \\ 42x_1x_2 + 42x_1^2x_2 + 42x_1x_2^2 + 42x_1^3 + 42x_2^3 + \\ 42x_3 + 42x_3^2 + 42x_3^3 + 42x_1x_3 + 42x_1x_3^2 + \\ 42x_1^2x_3 + 42x_2x_3 + 42x_2x_3^2 + 42x_2^2x_3 + 42x_1x_2x_3 \end{aligned}$$

d) Cosider the function $f_{\boldsymbol{\theta}}(\boldsymbol{x}) : R^5 \rightarrow R$, $f_\theta(\boldsymbol{x}) = \frac{1}{1+exp(-\boldsymbol{\theta}^T\boldsymbol{x})}$ Define the associated parametric set and give two instances of it. What is the dimensionality of $\boldsymbol{\theta}$?

For a given $\boldsymbol{x} = [x_1, x_2, x_3, x_4, x_5]^T$ it is:

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + exp(-(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4 + \theta_5 x_5))}$$

thus,

$$\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T$$

and

$$F_{lin} := f_{\boldsymbol{\theta}}(\cdot) : \boldsymbol{\theta} \in R^6$$

i) If we choose e.g., $\boldsymbol{\theta} = [0, 1, 2, 3, 4, 5]^T$ have the following instance of F_{lin}

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + exp(-(1x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5))}$$

ii) If we choose e.g., $\boldsymbol{\theta} = [42, 42, 42, 42, 42, 42]^T$ have the following instance of F_{lin}

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \frac{1}{1 + exp(-(42 + 42x_1 + 42x_2 + 42x_3 + 42x_4 + 42x_5))}$$

e) In which of the above cases $\boldsymbol{f_\theta}$ is linear with respect to $\boldsymbol{\theta}$?

$\boldsymbol{f_\theta}$ is linear with respects to $\boldsymbol{\theta}$ in cases a, b, c, while non linear in case d.

Exercise 4.

For the left part of the equation we have:

$$\begin{aligned}(\boldsymbol{\theta}^T\boldsymbol{x})\boldsymbol{x} &= \left(\begin{bmatrix}\theta_1, \ldots, \theta_l\end{bmatrix} \begin{bmatrix} x1 \\ \vdots \\ x_l \end{bmatrix} \right) \boldsymbol{x} \\ &= \left(\sum_{i=1}^{l} \theta_i x_i \right) \begin{bmatrix} x1 \\ \vdots \\ x_l \end{bmatrix} \end{aligned} \tag{3.1}$$

For the right part of the equation we have:

$$(\boldsymbol{x}\boldsymbol{x}^T)\boldsymbol{\theta} = \left(\begin{bmatrix} x1 \\ \vdots \\ x_l \end{bmatrix} \begin{bmatrix} x_1, \dots, x_l \end{bmatrix} \right) \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_l \end{bmatrix} = \begin{bmatrix} x1 \\ \vdots \\ x_l \end{bmatrix} \left(\begin{bmatrix} x_1, \dots, x_l \end{bmatrix} \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_l \end{bmatrix} \right)$$
$$= \begin{bmatrix} x1 \\ \vdots \\ x_l \end{bmatrix} \left(\sum_{i=1}^{l} \theta_i x_i \right) \tag{3.2}$$

The last expression (3.2) can be written exactly as (3.1) since $\sum_{i=1}^{l} \theta_i x_i$ is a number (vector 1x1) and we know that for a matrix $A : cA = Ac$.

Exercise 5.

a) Verify the identities.

i) We have:

$$X^T X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1l} & x_{2l} & \iddots & x_{Nl} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1l} \\ x_{21} & x_{22} & \dots & x_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \iddots & x_{Nl} \end{bmatrix}$$
$$= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1}^T \\ \boldsymbol{x_2}^T \\ \vdots \\ \boldsymbol{x_N}^T \end{bmatrix}$$
$$= \sum_{n=1}^{N} \boldsymbol{x_n} \boldsymbol{x_n}^T$$

ii) Similarly:

$$
\begin{aligned}
X^T\boldsymbol{y} &= \begin{bmatrix} x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1l} & x_{2l} & \ddots & x_{Nl} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \\
&= \sum_{n=1}^{N} \boldsymbol{x_n} y_n = \sum_{n=1}^{N} y_n \boldsymbol{x_n} \qquad (5.1)
\end{aligned}
$$

We are able to interchange y_n in (5.1) since when creating the inner product, each y_i is a single number and again using $A : cA = Ac$ we come to this form.

b) What are the size of the matrices?

1. $X_{N\times l}$
2. $\boldsymbol{y}_{N\times 1}$
3. For X^TX we have: $X^T_{l\times N} \cdot X_{N\times l} \rightarrow (X^TX)_{l\times l}$
4. For $X^T\boldsymbol{y}$ we have: $X^T_{l\times N} \cdot \boldsymbol{y}_{N\times 1} \rightarrow (X^T\boldsymbol{y})_{l\times 1}$

c) Assume that a column vector of 1's is added in front of the 1st column of X.

i) What will be the changes in the dimensionality of the quantities in (b)?

1. $X_{N\times(l+1)}$
2. $\boldsymbol{y}_{N\times 1}$
3. For X^TX we have: $X^T_{(l+1)\times N} \cdot X_{N\times(l+1)} \rightarrow (X^TX)_{(l+1)\times(l+1)}$
4. For $X^T\boldsymbol{y}$ we have: $X^T_{(l+1)\times N} \cdot \boldsymbol{y}_{N\times 1} \rightarrow (X^T\boldsymbol{y})_{(l+1)\times 1}$

ii) Do the identities given in (a) still hold? i) We now have:

$$X^T X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_{11} & x_{21} & \dots & x_{N1} \\ x_{12} & x_{22} & \dots & x_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1l} & x_{2l} & \ddots & x_{Nl} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1l} \\ 1 & x_{21} & x_{22} & \dots & x_{2l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \ddots & x_{Nl} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & \dots & 1 \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{x_1}^T \\ 1 & \boldsymbol{x_2}^T \\ \vdots & \vdots \\ 1 & \boldsymbol{x_N}^T \end{bmatrix} \tag{5.2}$$

We now consider the vectors $\boldsymbol{x}'_n = [1, x_{n1}, \dots, x_{nl}]^T$. By substituting in (5.2) we have:

$$X^T X = \begin{bmatrix} \boldsymbol{x}'_1 & \boldsymbol{x}'_2 & \dots & \boldsymbol{x}'_N \end{bmatrix} \begin{bmatrix} \boldsymbol{x'_1}^T \\ \boldsymbol{x'_2}^T \\ \vdots \\ \boldsymbol{x'_N}^T \end{bmatrix}$$

$$= \sum_{n=1}^{N} \boldsymbol{x'_n} \boldsymbol{x'_n}^T$$

$$\overset{\mathrm{x'_n = x_n}}{=} \sum_{n=1}^{N} \boldsymbol{x_n} \boldsymbol{x_n}^T$$

In a similar way we can show that the second identity still holds.

Exercise 6. (No grade)

Exercise 6 was written in paper for practicing purposes.

Exercise 7. A body moves on a straight line and performs a smoothly accelerating motion (we begin to study its motion at the time instance $t = 0$). In the following table is given the velocity at certain time instances

t(sec)	1	2	3	4	5
ν(m/sec)	5.1	6.8	9.2	10.9	13.1

(a) Estimate the initial velocity and the acceleration of the body, based on the above measurements, utilizing the least squares error criterion.
We define:

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \quad \textit{and} \quad v = \begin{bmatrix} 5.1 \\ 6.8 \\ 9.2 \\ 10.9 \\ 13.1 \end{bmatrix}$$

It is: $T^T T = \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}$ and $(T^T T)^{-1} = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix}$. Also $T^T v = \begin{bmatrix} 45.1 \\ 155.4 \end{bmatrix}$.
Thus $\boldsymbol{\theta} = (T^T T)^{-1} T^T v = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 45.1 \\ 155.4 \end{bmatrix} = \begin{bmatrix} 2.99 \\ 2.01 \end{bmatrix}$
Thus, the Least Squares line is:

$$v = 2.99 + 2.01 \cdot t$$

By setting $t = 0$ we get that the initial velocity is $v_0 = 2.99$.

(b) Write down the equation that expresses the velocity of the body as a function of time t. The function is:

$$v = 2.99 + 2.01 \cdot t \tag{7.1}$$

(c) Estimate the velocity of the body at t=2.3. By substituting $t = 2.3$ in (**??**) we get that:

$$v = 2.99 + 2.01 \cdot 2.3 = 7.613$$