Exercise 1

Parametric f(.): Linear Regression; where f(x) = a +bx, x being the independent variable. The parameters are a and b and are calculated using the data of the problem.

Non-parametric f(.): Decision Trees; in the sense that the only parameter is the size of the tree which depends on the size of the dataset but not the actual values of the data.

Exercise 2

a. *Parametric set of the quadratic functions f_ϑ: R→R*

That is, for a given $x = [x_1]^T \in R$ it is:

$$f_\theta(x_1) = \theta_0 + \theta_1 x_1 + \theta_2 x_1^2, \qquad \theta = [\theta_0, \theta_1, \theta_2]^T$$

and $F := \{f_\theta(\cdot): \theta \in R^3\}$

eg.a.1 $\theta = [4,5,3]^T$ then the instance of F is $f_\theta(x_1) = 4 + 5x_1 + 3x_1^2$

eg.a.2 $\theta = [2,1,0]^T$ then the instance of F is $f_\theta(x_1) = 2 + x_1$

b. *Parametric set of the 3rd degree polynomials functions f_ϑ: R2→R*

That is, for a given $x = [x_1, x_2]^T \in R^2$ it is:

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

$$\theta = [\theta_0, \theta_1, \theta_2, \theta_3, \theta_y, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9]^T$$

and $F := \{f_\theta(\cdot): \theta \in R^{10}\}$

eg.b.1 $\theta = [1,4,3,2,5,6,7,1,2,3]^T$ then the instance of F is

$$f_\theta(x_1, x_2) = 1 + 4x_1 + 3x_2 + 2x_1 x_2 + 5x_1^2 + 6x_2^2 + 7x_1^2 x_2 + x_2^2 x_1 + 2x_1^3 + 3x_2^2$$

eg.b.2 $\theta = [0,1,34,0,7,0,1,2,0,1]^T$ then the instance of F is

$$f_\theta(x_1, x_2) = x_1 + 34x_2 + 7x_1^2 + x_1^2 x_2 + 2x_2^2 x_1 + x_2^2$$

c. *Parametric set of the 3rd degree polynomials functions f_ϑ: R3→R*

That is, for a given $x = [x_1, x_2, x_3]^T \in R^3$ it is:

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

$$\theta = [\theta_0, \theta_1, \theta_{2,\dots,} \theta_{19}]^T$$

and $F \coloneqq \{f_\theta(\cdot): \theta \in R^{20}\}$

<u>eg.c.1</u> $\theta = [1,4,3,2,5,6,7,1,2,3,0,0,0,1,12,9,54,0,1,0]^T$ then the instance of F is

$$f_\theta(x_1, x_2, x_3) = 1 + 4x_1 + 3x_2 + 2x_3 + 5x_1^2 + 6x_2^2 + 7x_3^2 + x_1x_2 + 2x_1x_3 + 3x_2x_3 + x_1^2x_2 + 12x_1^2x_3 + 9x_2^2x_1 + 54x_2^2x_3 + x_3^2x_2$$

<u>eg.c.2</u> $\theta = [0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0]^T$ then the instance of F is

$$f_\theta(x_1, x_2, x_3) = x_1^2 + x_2^2x_3$$

d. $f_\theta(x): R^5 \to R, f_\theta(x) = \frac{1}{1+exp(-\theta^T x)}$

That is, for a given x = [x_1, x_2, x_3, x_4, x_5]^T∈R^3 and $\theta = [\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T$ it is:

$$f_\theta(x_1, x_2, x_3, x_4, x_5) = \frac{1}{1 + exp(-\theta_0 - \theta_1x_1 - \theta_2x_2 - \theta_3x_3 - \theta_4x_4 - \theta_5x_5)}$$

and $F \coloneqq \{f_\theta(\cdot): \theta \in R^6\}$

<u>eg.d.1</u> $\theta = [-1,1,-1,1,1,-1]^T$ then the instance of F is

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

<u>eg.d.2</u> $\theta = [0,0,0,0,-1]^T$ then the instance of F is

$$f_\theta(x_1, x_2, x_3, x_4, x_5) = \frac{1}{1 + exp(x_5)}$$

e. f is linear with respect to θ in a, b, c but not in d

Exercise 3

$(\theta^T x)x = ([\theta_1\ \theta_2 \ldots \theta_l][x_1\ x_2 \ldots x_l]^T)[x_1\ x_2 \ldots x_l]^T = (x_1\theta_1 + \theta_2 x_2 + \ldots + \theta_l x_l)\ [x_1\ x_2 \ldots x_l]^T =$

$$\begin{bmatrix} \theta_1 x_1^2 + \theta_2 x_2 x_1 + & \cdots & +\theta_1 x_l x_1 \\ \theta_1 x_1 x_2 + \theta_2 x_2^2 + & \cdots & +\theta_1 x_l x_2 \\ \vdots & & \\ \theta_1 x_1 x_l + \theta_2 x_2 x_l + & \cdots & +\theta_l x_l^2 \end{bmatrix} = \begin{bmatrix} x_1^2 & x_2 x_1 & \cdots & x_l x_1 \\ x_1 & x_2 x_2^2 & \cdots & x_l x_2 \\ & \vdots & & \\ x_1 x_l & x_2 x_l & \cdots & x_l^2 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_l \end{bmatrix} = \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_l \end{bmatrix} [x_1\ x_2 \ldots x_l] \right)\theta^T = (xx^T)\theta$$

Thus, $(\theta^T x)x = (xx^T)\theta$

Exercise 4

First, we will show that $X^T = [x_1 \quad x_2 \quad \cdots \quad x_N]$

Since $X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1l} \\ x_{21} & x_{22} & \cdots & x_{2l} \\ \vdots & & & \\ x_{N1} & x_{N2} & \cdots & x_{Nl} \end{bmatrix}$, then $X^T = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{N1} \\ x_{12} & x_{22} & \cdots & x_{N2} \\ \vdots & & & \\ x_{1l} & x_{2l} & \cdots & x_{Nl} \end{bmatrix} = [x_1 \quad x_2 \quad \cdots \quad x_N]$

a. $X^T X = [x_1 \quad x_2 \quad \cdots \quad x_N] \begin{bmatrix} {x_1}^T \\ {x_2}^T \\ \vdots \\ {x_N}^T \end{bmatrix} = x_1 {x_1}^T + x_2 {x_2}^T + \cdots + x_N {x_N}^T = \sum_{n=1}^N x_n x_n^T$

b. $X^T y = [x_1 \quad x_2 \quad \cdots \quad x_N] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = x_1 y_1 + x_1 y_1 + \cdots + x_N y_N = \sum_{n=1}^N x_n y_n$

But $\sum_{n=1}^N x_n y_n = \sum_{n=1}^N y_n x_n$, since y_n is just a number.

Exercise 6

The velocity of a body performing smoothly accelerating motion is given by the equation:

$$v = v_0 + at$$

Where t is time, v_0 is the initial speed and α is the acceleration.

$$y = v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 5.1 \\ 6.8 \\ 9.2 \\ 10.9 \\ 13.1 \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \\ 1 & t_5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}, \qquad \theta^T = [v_0 \quad \alpha]$$

Taking advantage of the least squares error criterion $\hat{\theta} = (X^T X)^{-1} X^T y$, we will find the initial speed and the acceleration.

First, we need to prove that the inverse matrix of $X^T X$ exists.

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 15 \\ 15 & 55 \end{bmatrix}$$

Thus, det=275-225=50 and $(X^T X)^{-1} = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix}$

$$(X^T X)^{-1} X^T = \begin{bmatrix} 1.1 & -0.3 \\ -0.3 & 0.1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.5 & 0.2 & -0.1 & -0.4 \\ -0.2 & -0.1 & 0 & 0.1 & 0.2 \end{bmatrix}$$

$$(X^T X)^{-1} X^T y = \begin{bmatrix} 0.8 & 0.5 & 0.2 & -0.1 & -0.4 \\ -0.2 & -0.1 & 0 & 0.1 & 0.2 \end{bmatrix} \begin{bmatrix} 5.1 \\ 6.8 \\ 9.2 \\ 10.9 \\ 13.1 \end{bmatrix} = \begin{bmatrix} 2.99 \\ 2.01 \end{bmatrix}$$

In conclusion, $\hat{\theta} = \begin{bmatrix} 2.99 \\ 2.01 \end{bmatrix}$

Initial speed is 2.99, acceleration is 2.01 and velocity at any time point t>0 is given by the equation

$$v = 2.99 + 2.01t$$

At t = 2.3 velocity of the body is 7.613.