Homework 2 - Group 076

Aprendizagem 2021/2022

1 Pen and Paper

1) Applying the linear basis function $\phi(\mathbf{x}) = (1, \|\mathbf{x}\|_2, \|\mathbf{x}\|_2^2, \|\mathbf{x}\|_2^3)$ (with $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2}$) to each instance $\mathbf{x}^{(i)}$ $(i = 1, ..., 8)$ in the training set, we get a new design matrix:

$$\mathbf{\Phi} = \begin{bmatrix} — & (\phi(\mathbf{x}^{(1)}))^T & — \\ — & (\phi(\mathbf{x}^{(2)}))^T & — \\ & \vdots & \\ — & (\phi(\mathbf{x}^{(8)}))^T & — \end{bmatrix} = \begin{bmatrix} 1.0 & 1.4142 & 2.0 & 2.8284 \\ 1.0 & 5.1962 & 27.0 & 140.2961 \\ 1.0 & 4.4721 & 20.0 & 89.4427 \\ 1.0 & 3.7417 & 14.0 & 52.3832 \\ 1.0 & 7.2801 & 53.0 & 385.8458 \\ 1.0 & 1.7321 & 3.0 & 5.1962 \\ 1.0 & 2.8284 & 8.0 & 22.6274 \\ 1.0 & 9.2195 & 85.0 & 783.6613 \end{bmatrix}$$

To learn the regression model, we must compute the weight vector $\mathbf{w}$ that minimizes the Sum of Squares error between the outputs $\mathbf{z} = \begin{bmatrix} 1 & 3 & 2 & 0 & 6 & 4 & 5 & 7 \end{bmatrix}^T$ and predictions $\hat{\mathbf{z}} = \mathbf{\Phi}\mathbf{w}$ (i.e., $\mathbf{w} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{z}$):

$$\mathbf{\Phi}^T\mathbf{\Phi} = \begin{bmatrix} 8.0 & 35.8843 & 212.0 & 1482.2811 \\ 35.8843 & 212.0 & 1482.2811 & 11436.0 \\ 212.0 & 1482.2811 & 11436.0 & 93573.5164 \\ 1482.2811 & 11436.0 & 93573.5164 & 793976.0 \end{bmatrix}$$

$$(\mathbf{\Phi}^T\mathbf{\Phi})^{-1} = \begin{bmatrix} 8.1955 & -6.2313 & 1.3049 & -0.0793 \\ -6.2313 & 5.0781 & -1.1044 & 0.0686 \\ 1.3049 & -1.1044 & 0.2472 & -0.0157 \\ -0.0793 & 0.0686 & -0.0157 & 0.001 \end{bmatrix}$$

$$(\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T = \begin{bmatrix} 1.7686 & -0.0811 & -0.6694 & -1.0069 & 1.3794 & 0.9051 & -0.785 & -0.5107 \\ -1.0644 & -0.0319 & 0.5312 & 0.904 & -1.307 & -0.3922 & 0.8501 & 0.5101 \\ 0.1933 & 0.0436 & -0.0907 & -0.1868 & 0.3232 & 0.0524 & -0.1954 & -0.1395 \\ -0.0107 & -0.0043 & 0.0044 & 0.0109 & -0.0214 & -0.0022 & 0.0123 & 0.011 \end{bmatrix}$$

$\mathbf{w} = (\mathbf{\Phi}^T\mathbf{\Phi})^{-1}\mathbf{\Phi}^T\mathbf{z} = \begin{bmatrix} 4.5835 & -1.6872 & 0.3377 & -0.0133 \end{bmatrix}^T$

2) Similarly to the previous question, we compute the image of each instance $\mathbf{x}^{(i)}$ $(i \in \{1, 2\})$ from the testing set and place the image $\phi(\mathbf{x}^{(i)})$ in each row of the matrix $\mathbf{\Phi}$. Using the obtained weight vector $\mathbf{w}$, we have the following estimates vector:

$$\hat{\mathbf{z}} = \mathbf{\Phi}\mathbf{w} = \begin{bmatrix} 1.0 & 2.0 & 4.0 & 8.0 \\ 1.0 & 2.4495 & 6.0 & 14.6969 \end{bmatrix} \begin{bmatrix} 4.5835 \\ -1.6872 \\ 0.3377 \\ -0.0133 \end{bmatrix} = \begin{bmatrix} 2.4536 \\ 2.2816 \end{bmatrix}$$

Computing the root mean square error, we have:

$$\text{RMSE}(\hat{\mathbf{z}}, \mathbf{z}) = \sum_{i=1}^{8} (\hat{z}_i - z_i)^2 = 1.2567$$