Stage: Second

Subject: Linear Algebra Application

2022-2023

Lecturer Kawther Saeed Taha

Chapter five

Some application of Linear Algebra بعض التطبيقات لجبر الخطي:

Using matrices for the Linear Regression model(Regression Analysis): Matrix Formulation of Linear Regression:

Where X is the input data and each column is a data feature, b is a vector of coefficients and y is a vector of output variables for each row in X. Reformulated, the problem becomes a system of linear equations where the b vector values are unknown.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + + \beta_p X_{pi} + \varepsilon_i$$

Where:

$i = 1, 2, 3, \ldots, n$

$\underline{Y}$: is the dependent variable(the response vector).

X's : are the independent variables (Information matrix) (design matrix).

$\underline{\beta}$: is the vector of parameter.

ε_i: error vector

The Design Matrix:

$$\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad , \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \quad , \quad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Then: $Y = X\beta + \varepsilon_i$

Parameter Estimation of this model by using matrix is:

$$\underline{\hat{\beta}} = (X'X)^{-1} X'Y$$

Simple Linear Regression model نموذج الانحدا الخطي البسيط :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$$

Then $\underline{\hat{\beta}} = (X'X)^{-1} X'Y$

Where:

$$\underline{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad , \quad X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix}$$

$$X'X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \end{bmatrix} \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{1n} \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^{\ 2} \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{12} & \cdots & x_{1n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i \, y_i \end{bmatrix}$$

$then:$

$$\underline{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = (X'X)^{-1} X'Y$$

Ex1: **From the following data:**

X	-2	0.6	1.4	1.3	0	-1.6	-1.7	0.7
Y	-6.1	-0.5	7.2	6.9	-0.2	-2.1	-3.9	3.8

Estimate the Linear Regression Model.

Solution:

									Sum
X	-2	0.6	1.4	1.3	0	-1.6	-1.7	0.7	-1.3
Y	-6.1	-0.5	7.2	6.9	-0.2	-2.1	-3.9	3.8	5.1
XY	12.2	-0.3	10.08	8.97	0	3.36	6.63	2.66	43.6
X^2	4	0.36	1.96	1.69	0	2.56	2.89	0.49	13.95

$$X'X = \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^{\,2} \end{bmatrix} = \begin{bmatrix} 8 & -1.3 \\ -1.3 & 13.95 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{adj(X'X)}{|X'X|}$$

$$|X'X| = \begin{vmatrix} 8 & -1.3 \\ -1.3 & 13.95 \end{vmatrix} = 111.6 - 1.69 = 109.91$$

$$adj(X'X) = \begin{bmatrix} 13.95 & 1.3 \\ 1.3 & 8 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{109.91} \begin{bmatrix} 13.95 & 1.3 \\ 1.3 & 8 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix} = \begin{bmatrix} 5.1 \\ 43.6 \end{bmatrix}$$

$$\underline{\hat{\beta}} = (X'X)^{-1} X'Y$$

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \frac{1}{109.91} \begin{bmatrix} 13.95 & 1.3 \\ 1.3 & 8 \end{bmatrix} \begin{bmatrix} 5.1 \\ 43.6 \end{bmatrix}$$

$$= \frac{1}{109.91} \begin{bmatrix} 71.145 + 56.68 \\ 6.63 + 348.8 \end{bmatrix} = \frac{1}{109.91} \begin{bmatrix} 127.825 \\ 355.43 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1.16 \\ 3.23 \end{bmatrix}$$

$$\therefore \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

$$= 1.16 + 3.23 X_i$$

Ex2:

If we have this information in 6 observations:

$$\sum_{i=1}^{n} x_i = 153.8 \qquad \sum_{i=1}^{n} x_i^{\,2} = 5859.26 \quad , \sum_{i=1}^{n} y_i = 18.7 \quad , \sum_{i=1}^{n} x_i\, y_i = 682.77$$

Find the estimated equation of the regression line.

Ex3:

From the following data

X_i	4	2	3	6	7	5	9	8	4	11
Y_i	3	3	2	5	7	6	6	4	2	7

Explain the relation between (X_i) and (Y_i).

Solution: