Video 4 Summary

Reetanjan Kali Roychowdhury

11 August 2022

1 Overview

Firstly , note that here we wish to express the **3*1** vector **Y**= $\begin{bmatrix} 9.8 \\ 9.1 \\ 7.0 \end{bmatrix}$ approximately as a linear combination of the two **3*1** column vectors $\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$, i.e. we want to find **A** and **B** such that the following holds approximately :-

$$\begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} * \mathbf{A} + \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} * \mathbf{B} \approx \begin{bmatrix} 9.8 \\ 9.1 \\ 7.0 \end{bmatrix}$$

This is of the form$\mathbf{Y} \approx \mathbf{X}\beta$, where $\mathbf{Y}= \begin{bmatrix} 9.8 \\ 9.1 \\ 7.0 \end{bmatrix}$,$\mathbf{X} = \begin{bmatrix} 3 & 4 \\ 4 & 1 \\ 2 & 3 \end{bmatrix}$ and $\beta = \begin{bmatrix} A \\ B \end{bmatrix}$.

Now , we can also write this approximate model as the following:-
$\bar{\mathbf{Y}} = \mathbf{X}\bar{\beta} + \varepsilon$, where Y is a **n*1** vector , X is a **n*p**, β is a **p*1** and ε is a **n*1** vector .

The above form $\bar{\mathbf{Y}} = \mathbf{X}\bar{\beta} + \varepsilon$ is called the general form of a **Linear Model**, which is nothing but an approximate system of Linear Equations ! [Note :- We are going to put some statistical assumptions over ε later on , so that the Model shall then be called a **Linear Statistical Model**]

2 Solving for β

Now , we try to find β so that $\mathbf{Y} \approx \mathbf{X}\beta$ holds approximately . Using some Linear Algebra , we shall see that to solve for the required $\hat{\beta}$ we shall need the orthogonal projection of vector $\mathbf{Y}$ onto the column space of matrix $\mathbf{X}$!

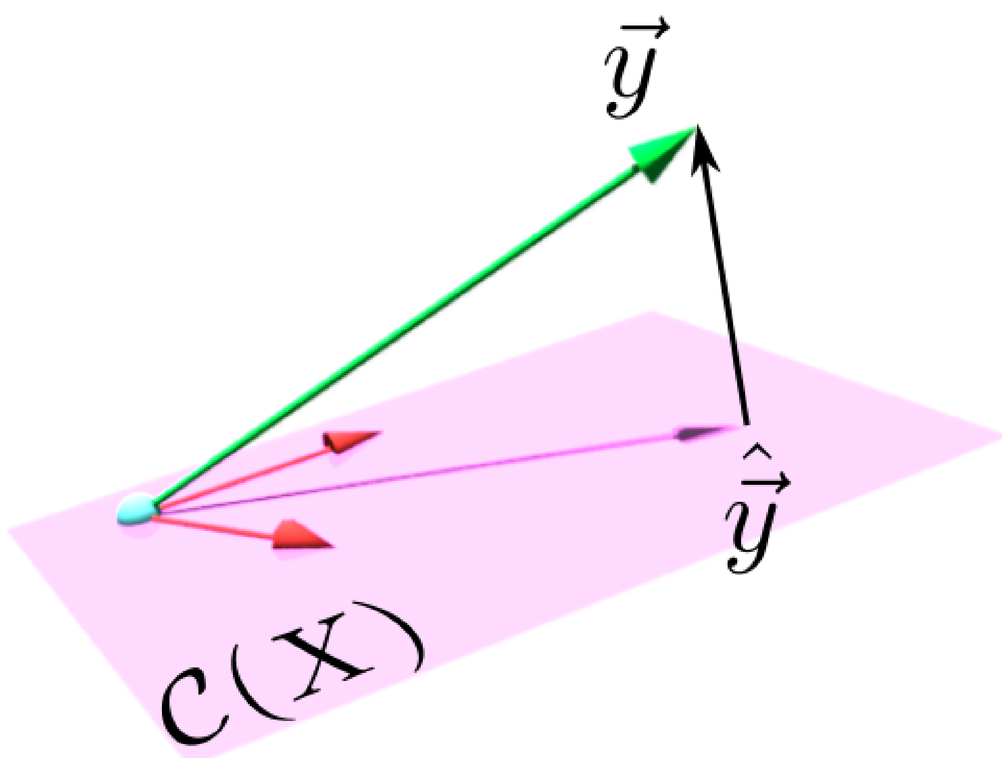

Figure 1: **Projection onto Column Space**

There is an easier technique using Linear Algebra to solve for this :-

1 . Given Equation

$$Y \approx X * \beta$$

2. Multiplying both sides by X transpose and considering equality , we have

$$X^T Y = X^T X \beta \text{ (.... } i)$$

Solving the above equation shall give us our required β and hence the above equation is also the **NORMAL EQUATION** ! Also , this equation is always consistent thus guaranteeing atleast one solution for *beta* .

3. Now if $X^T X$ is non-singular , then we have a unique solution for β else we shall infinitely many solutions . So , considering $X^T X$ to be non-singular i.e. invertible , we multiply $X^T X^{-1}$ on both sides of equation (..... i) to get:-

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

The $\hat{\beta}$ thus obtained through the process by solving the **NORMAL EQUATION** is called the Least Squares Estimator !