Ordinary Least Square Method (Geometric Interpretation)

Suppose, We are given data points and corresponding target as $(x_1,x_2) \rightarrow y$

- $P_1(1,2) \rightarrow 3$
- $P_2(0,1) \rightarrow 1$
- $P_3(0,0) \rightarrow 2$

This represented as features / basis vectors is :

$\oint_1$	$\oint_2$	y
1	2	3
0	1	1
0	0	2

Here, **($\oint_1$, $\oint_2$)** are the basis vectors commonly known as features. The **y** is the target variable / value to be predicted. With a general representation as :

$$\begin{array}{cccc} \phi_1 & \phi_2 & & \phi_d \end{array}$$
$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix}$$

Solution using OLS

The solution is the Linear combination of basis vectors represented as:

$$\hat{y} = w\oint = w_1\oint_1 + w_2\oint_2 + w_3\oint_3 + \ldots\ldots\ldots + w_d\oint_d$$

ŷ is the linear combination of basis vectors, this will always span the column space created by (∮ 1 , ∮ 2….. ∮ d).

The optimal solution can be obtained by minimizing the squared errors between **y** & **ŷ.**

$$\textbf{i.e. min}(\sum_{0}^{M} (y_i - \hat{y}_i)^2)$$

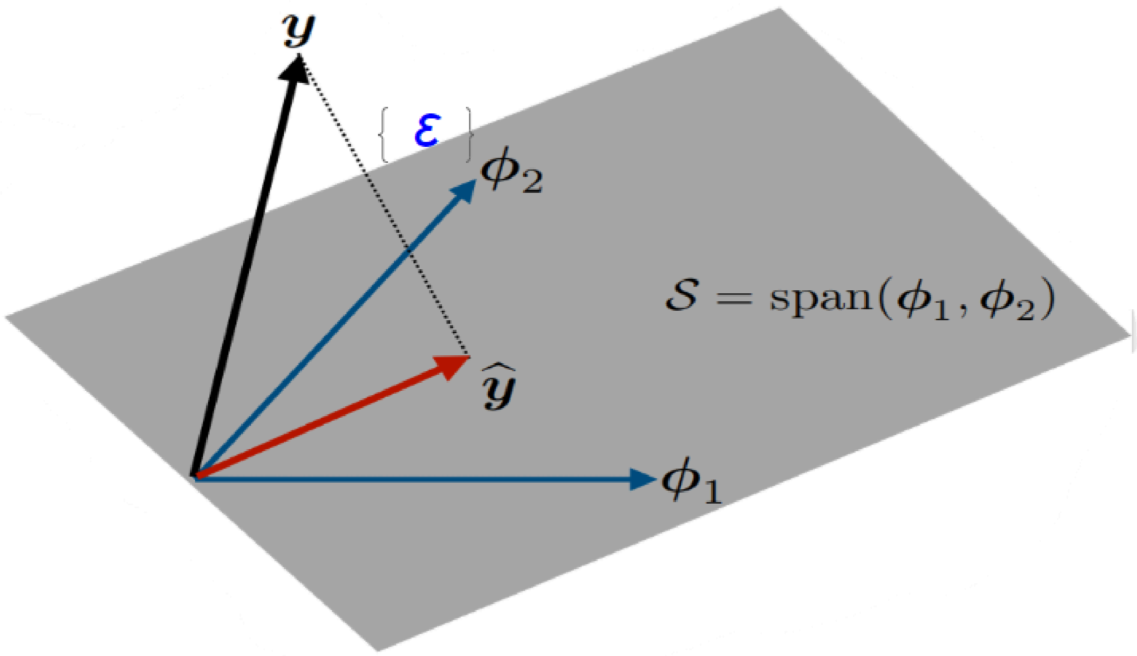
This can be done by solving for weights(***w***) as:

$$w = (\oint^T \oint)^{-1} \oint y$$

Geometric Intuition : Geometric representation of basis vectors

When the basis vectors are represented in $\mathbf{R^n}$ space , we get vector **ŷ in the space columns of** ($\oint_1$, $\oint_2$….. $\oint_d$). The Least square method tries to optimize the solution such that it is orthogonal to the span of data points X i.e. ($\oint_1$, $\oint_2$….. $\oint_d$). The difference between predicted (**ŷ**) and actual values(**y**) is the error term given by:

$$\varepsilon = y - \hat{y} = y - (w_1\oint_1 + w_2\oint_2 + w_3\oint_3 + \ldots\ldots\ldots + w_d\oint_d)$$



References

1. Geometric Interpretation of OLS by Ben Lambert. (Link)
2. Least Squares and OLS by Penn State University. (Link)
3. Pattern Recognition & Machine Learning <Page 142-143> (Link)