Hat Matrix

The Hat Matrix

$$H = X(X^TX)^{-1}X^T$$

Its called the Hat Matrix cause :

$$\hat{Y} = X\hat{\beta}$$

$$\hat{Y} = X(X^TX)^{-1}X^TY$$

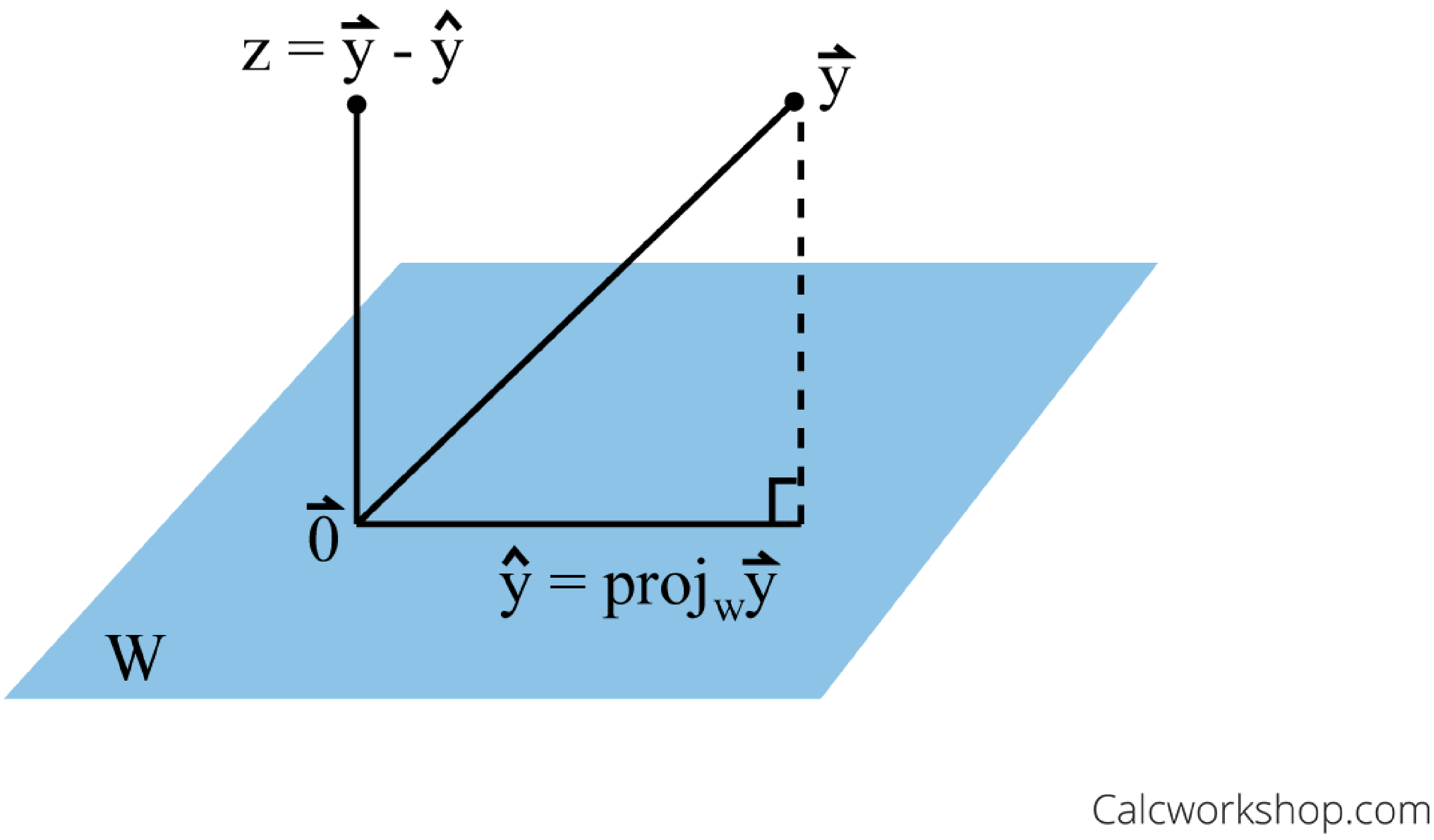
$$\hat{Y} = HY$$

- Its transforms $Y \to \hat{Y}$
- So the Hat Matrix is a Transformation Matrix

$$e_i = Y - \hat{Y} = Y - HY = (I - H)Y$$

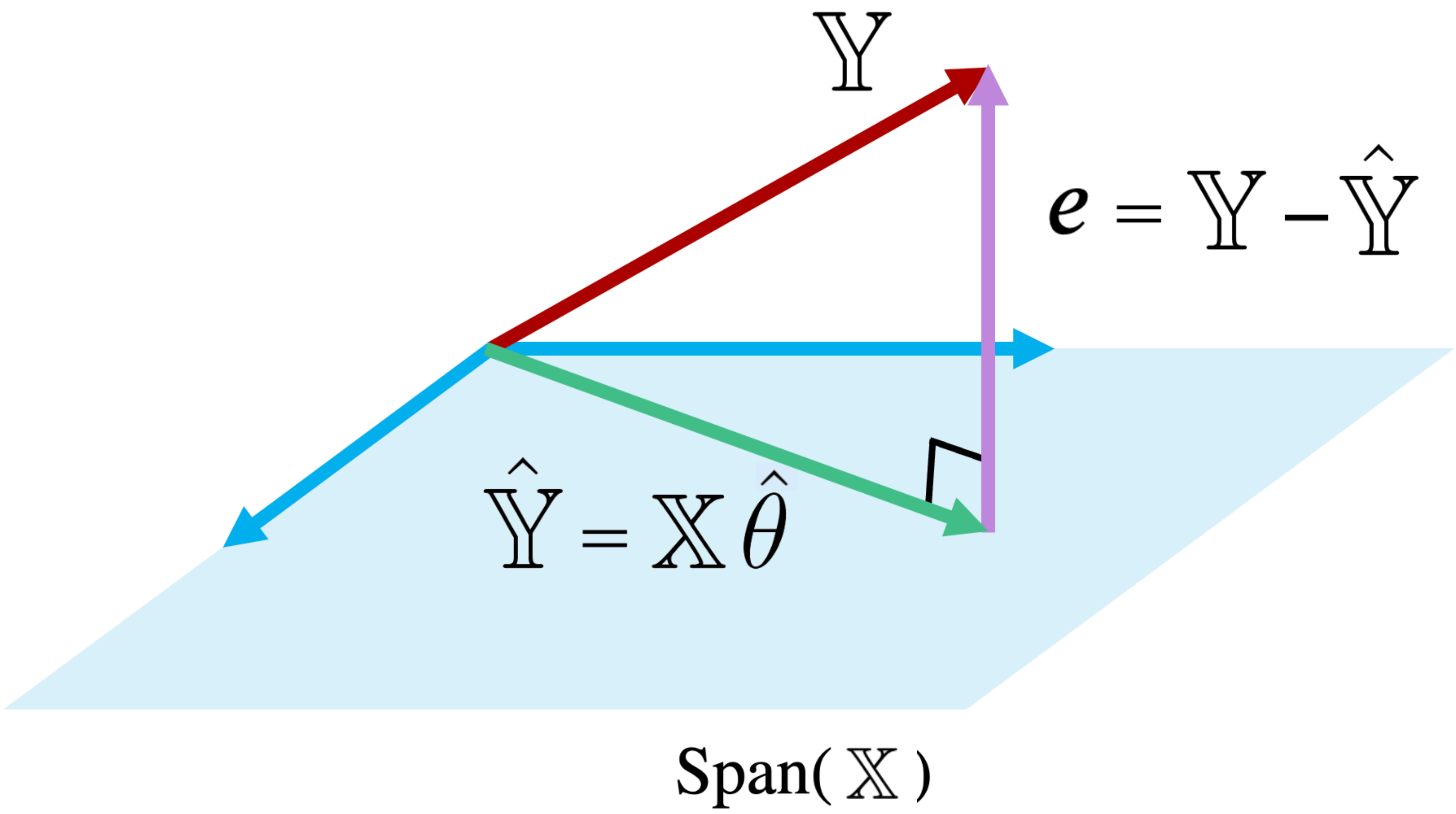
Orthogonal Projection

- Its a decomposition of a vector into sum of two orthogonal vectors



- $\hat{Y}$ is the projection of Y in an Orthogonal projection in the W subspace
- Where $Z + \hat{Y} = Y$
- We decomposed Y into a sum of two orthogonal vectors $Z, \hat{Y}$

Multiple Linear Regression interpretation :



- Think of rectangle as the column space on X (They are vectors of form $X\hat{\beta}$), they span a subspace and we call it **The column space** $c(x)$
- $\hat{Y}$ is an orthogonal projection of Y into the **Column space** $c(x)$
- Now Y is the sum of two **perpendicular vectors** $Y = e + \hat{Y}$
- e is the least squares method explained in a graphical way
  Now the Hat Matrix comes in

$$\text{Least Squares } = e = Y - \hat{Y} = (I - H)Y \in c(x)^{\perp}$$

$$\text{Proj}_{c(x)}Y = \hat{Y} = X\hat{\beta} = HY$$

Using Linear mapping logic : (Hat matrix Transform Y Orthogonaly into $c(x)$ no matter where Y is)

$$\hat{Y} = HY$$

Conclusion :

- The Response vector Y is decomposed into 2 orthogonal components

$$Y = \hat{Y} + e$$
$$\hat{Y} = x\hat{\beta} \in c(x) \text{ Estimation Space}$$
$$e = Y - \hat{Y} \text{ Error space}$$