Mohamad Lakkis

February, 2024

1 Proving that the OLS estimates are equivariant under scaling transformation and under rotations

Before jumping into trying to actually prove the claim, let's try to understand more about what is meant by it !
Equivariance means that if we transform (according to **Scaling** or **Rotation**) the variable axes (i.e. X) to a new variable axes(X'), then carry out the OLS in the new transformed system X' and then transform back the solution, we will obtain the same results as in the original system X ! Note that we are scaling/rotating the predictors and not the design matrix, i.e. the column of 1s remains unchanged because it is not a variable that can be scaled(as mentioned in the claim). So our **GOAL:** show $\hat{y} = \hat{y}'$ for both cases !

Scaling:
We will work with a diagonal matrix A with values 1 and a_i on the diagonal, we will X rather than $X(p)$ to be able to write $\hat{\beta} = (X^T X)^{-1} X^T y$, and so,

$$\hat{\beta}' = (X'^T X')^{-1} X'^T y$$

Substituting $X' = XA$; and applying some basic matrix rules, plus that the transpose of a product of two matrices is equal to the product of the transposes in reverse order(same for the inverse),and so we can get, ,and $A = A^T$, and that $(A^T)^{-1} A^T = I$

$$\hat{\beta}' = A^{-1}(X^T X)^{-1} X^T y = A^{-1}\hat{\beta}$$

So to transform back we can just multiply the following with A !
Now,

$$\hat{y}' = X'\hat{\beta}'$$

Substituting $\hat{\beta}'$, and $X' = XA$we get,

$$\hat{y}' = XAA^{-1}\hat{\beta}$$

And so,

$$\hat{y}' = X\hat{\beta} = \hat{y}$$

Rotation:
Here we will follow a very similar analysis as the previous proof, rather than $X^{'} = XA$, we will have $X^{'} = XO^T$, and here we will use the property of an orthogonal matrix $O^T = O^{-1}$, plus that the transpose of a product of two matrices is equal to the product of the transposes in reverse order(same for the inverse), and so we can derive,

$$\hat{\beta}^{'} = (X^{'T}X')^{-1}X^{'T}y$$

Substituting $\hat{\beta}^{'}$, and $X' = XO^T$we get,

$$\hat{\beta}^{'} = ((XO^T)^T(XO^T))^{-1}(XO^T)^Ty$$

using the fact that, the inverse of a product of two or more matrices is equal to the product of the inverse in reverse order

$$\hat{\beta}^{'} = (O^T)^T(X^TX)^{-1}O^TOX^Ty$$

Since $O^TO = I_p$, and $(O^T)^T = O$,

$$\hat{\beta}^{'} = O(X^TX)^{-1}X^Ty = O\hat{\beta}$$

Now,

$$\hat{y}^{'} = X'\hat{\beta}' = XO^TO\hat{\beta} = \hat{y}$$