Residuals : Extra Contribution of a Part

Samahriti Mukherjee
Roll No.: BS2003
B.Stat 3rd year
Indian Statistical Institute

1 Introduction

We will consider a more general form.
We have a model like this :

$$\vec{y} = X\vec{\beta} + \vec{\epsilon}$$

Now we consider X and split it into two parts , X_1 and X_2.
So

$$X\vec{\beta} = [X_1 X_2] \begin{bmatrix} \vec{\beta_1} \\ \vec{\beta_2} \end{bmatrix}$$

Where β_1 and β_2 are coefficient vectors corresponding to X_1 and X_2 respectively.

2 Goal

We are interested in how much does X_2 help to explain the behaviour of $\vec{y}$ when X_1 is already a part of the model.
So instead of fitting

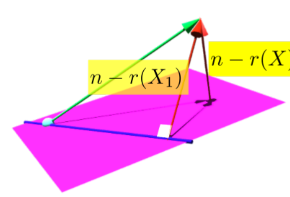
$$\vec{y} = X\vec{\beta} + \vec{\epsilon}$$

we will fit

$$\vec{y} = X_1\vec{\beta_1} + \vec{\epsilon}$$

So how much inclusion of X_2 in the model is justified?
This will lead us to an efficient way of scaling the residual sum of squares.
This is based on the diagram on the right.

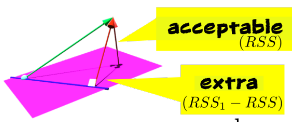
The blue line denotes the column space of X_1. Green arrow denotes the $\vec{y}$ vector. Now if we consider only the blue line, of course we will be able to explain less of the behaviour of the green arrow. But the question is, how much less?
Now considering only the column space of X_1 we try to fit the model, i.e. we drop a perpendicular from the tip of the $\vec{y}$



vector on $\mathscr{C}(X_1)$.
From the diagram, it is clear that when we restrict the X space, the length of the perpendicular goes up, i.e. the error goes up. Let us see how much increase is there in the perpendicular.

We can see that

$$RSS = ||\vec{y} - P_X\vec{y}||^2$$
$$RSS_1 = ||\vec{y} - P_{X_1}\vec{y}||^2$$



where $P_X\vec{y}$ and $P_{X_1}\vec{y}$ are the projection of $\vec{y}$ onto $\mathscr{C}(X)$ and $\mathscr{C}(X_1)$ respectively.
So by including the X_2 part, extra error=$RSS_1 - RSS$, which is positive anyways,
We are committing the RSS_1 part anyways. So we may consider this number

$$\frac{RSS_1 - RSS}{RSS_1},$$

which gives you a measure of what is the additional benefit of including X_2 when X_1 is already in the model. This measure always lie between 0 and 1 as

$$RSS_1 - RSS \geq 0$$
$$\Rightarrow \frac{RSS_1 - RSS}{RSS_1} \geq 0$$

Also

$$RSS \geq 0$$
$$\Rightarrow RSS_1 - RSS \leq RSS_1$$
$$\Rightarrow \frac{RSS_1 - RSS}{RSS_1} \leq 1$$