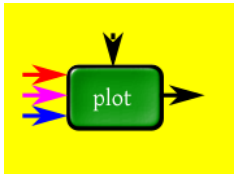
Anova Table:
Orthogonality of columns

Bhaskara Rahul Sanku

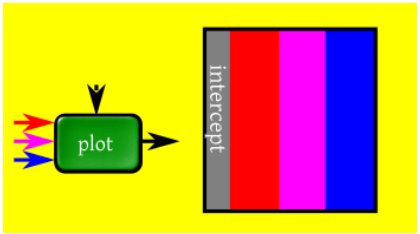
October 2022

Linear models in modern age includes situations of various types of inputs(only covariates or only factors or both) based on which we have a particular model.When the statisticians of the past were thinking about linear models, they were exclusively working on anova model(i.e only factor inputs) which is simpler in setup.So regression was initially considered to be a seperate field of study.But even then it wasn't an easy task to come up with an algebraic identity(i.e total variation of output in terms of components of variations of various inputs)

When we have only one factor input, then they could come up with the simple algebraic identity which made them look for identities in more general setup.But later did they realise that obtaining such an identity is a formidable task.Consider an example as follows,



Suppose we have 3 factor inputs which are colour coded as red purple and blue.Now we want to obtain such an algebraic identity(not always possible).The condition under which it is possible could be expressed in linear algebraic terms and for that we have to look at it's design matrix.



Every input contributes a number of columns to our design matrix.Suppose for 3 factor inputs, the corresponding parameters are $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3, \gamma_4$

then we have 3-α columns, 2-β columns, 4-γ columns respectively.Typically we have intercept column(the very first column of design matrix denoted by grey colour in above picture).

Now we have shown that each input provides it's own batch of columns and soon found that the condition under which such a splitting is possible is when the space spanned by the red columns is orthogonal to space spanned by the purple columns and the space spanned by blue columns is orthogonal to both column spaces of red and purple.Note that column spaces may have some elements in common but are removed as span of intercept from them.To better understand it consider the following example.

Suppose in R^3 we have $x-y$ and $x-z$ planes and we can say that they are perpendicular to each other.But they are not orthogonal to each other as not every vector in $x-y$ plane is perpendicular to every vector in $x-z$ plane.To make the two planes orthogonal we have to remove their common component from every vector in two planes which is x component in this case.So here the left out y component of any vector in $x-y$ plane is always orthogonal to left out out z component of any vector in $x-z$ plane.

So we are saying that the spans of red purple and blue inputs may have something in common that is span of gray (intercept).Once you remove that common thing if their spans are mutually orthogonal(which in general may not be true) then it is possible to split up the **total variation in the output in terms of component variations that are ascribable to different inputs including the random error**.