Ritam Dey (BS1913)

So why do people use the interaction model instead of the cell means model more frequently? The reason lies in the way the user perceives the interaction between factors. Naturally we dont want interaction to be there because it is bad news. When there is no interaction we can talk about the two input separately. While trying to find a model what we actually do is we assume that the respose is linearly related to the feature variables and we try to interpret the coefficients or in statistical terminology the weights on those features. The weight for that particular feature will give us an understanding of the qualitative and the quantitative effect of the particular feature on a certain response .

$$Y = 10x_1 + 2x_2 + \epsilon$$

In the above model we see that the factor (we replace "feature" with "factor") x_1 has more effect than the factor x_2 in the output Y. For the construction of this model the assumption of non interaction have been undertaking so it becomes more easy to interpret the individual contribution of each factor in the model. But having interactions spoils the this nice interpretation. Assuming there is interaction between the two feature in the previous model we get something like:

$$Y = 10x_1 + 2x_2 + 3x_{1,2} + \epsilon$$

earlier we could have easily said that suppose i keep increasing the factor x_1 , my natural intution will be the model response will also increase. But the case for the interaction model in slightly different . If we keep x_2 fixed and keep increasing the desired change in the response Y might not be same as the previous model. It might so happen that upto a certain level the response Y increased as x_1 increases, but after come crossing some threshold the interaction $x_{1,2}$ becomes negetive and the response does not increase linearly as we would expect or decrease significantly. all because of $x_{1,2}$ is the interaction term between x_1 and x_2 and the x_1 is changing so does the interaction effect . Which motivates us to include this ineraction term in the model seperately to caption the interrelation between the factors x_1 and x_2.

consider an example. In glucose water solution if we mix more glucose the sweetness increases if we add more water the sweetness decreases. so we might model the situation in such a way:

$$sweetness = A.glucose - B.water + \epsilon,$$

where A and B are positive numbers. Here we see that if we increase water the sweetness will decrease . This model might not be mathematically correct but this a good understanding of hte interaction concept. Now water is a product of interaction between hydrogen and oxygen . So what we want to look at is the how the sweetness of ht mixture is affected when we add hydrogen or oxygen to the mixture. Consider the model :

$$sweetness = C.oxygen + D.hydrogen + A.glucose + B.water + \epsilon$$

if we wanted to measure the effect of adding oxygen in the sweetness we would find that individually the sweetness is not affected by the gas. But if the environment is such that electricity and hydrogen is present then their interaction will create water and the sweetness will decrease. So here we see that individually the gases might have no effect but their interaction plays a major role in the sweetness of the water.

1 Working with interaction

Previously we saw that we have two ways of expressing the interaction models:

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}............(A)$$
$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}.............(B)$$

We prefer the model B over the model A. The first model somehow tells us that the two factors are entirely mixed up and cannot be analyzed seperately. We study the μ_{ij} to study the effect of each pair individually, calculate their mean and are done. But in the model B we see there are two parts . the first $\mu + \alpha_i + \beta_j$ part is the "good additive part". And the γ_{ij} is the interaction between them which might not have the additive effect. The reason we put there is because we dont want to influence the model by that interaction term. Rather it is to test whether their is interaction or not.

Such an interaction based exression allows us to test , whether the inteaction terms are zero or not. If we do a hypothesis test with out null being "$gamma_{ij} = 0$". If we accept the null then we conclude that there is no interaction and we are happy about our model. Whereas if we assume the formulation of the model A then we can not test for the interaction term "only" . Although we can use complex linear expression among the μ_{ij}'s to extract out the total effect of a particular factor but that is tiresome and not at all computation friendly.

In case of interraction is there , we are not interested aout how the interaction plays a role in the response , rather we are interested in knowing that which set or sets of factors have an interaction between them so that we stay catious whie dealing with those particular factors simultaneously. In that case we just work with those particular factors separately and mention them.

So summarizing, we are performing some experiments and we are trying to find the effect of certain features . First, we fit the model B and look for whether there is any interaction , that is whether γ_{ij} is zero or not for all i and j and then move to model A if necessary to find the combined effect of the i,j pair of factors. If they are zero we work with the additive model and report that. If one of the $g\gamma_{ij}$ is zero then we treat the case seperately and report that but dont include the pairs in the general model. Like for a particular j we find the response for all the i's and for that particular i we find the response for all hte j's. and report that.