Gauss-Markov model:Departures from Gauss-Markov set up

Srikar Babu Pilli

August 2022

We have learned one of the most important theorems of linear models i.e. the Gauss-Markov theorem.It is very appealing that we just put the barest possible restriction on the linear parametric function that we want to estimate and we got the best linear unbiased estimate,in a very constructive way.But,we have not reached our goal yet as there are problems,in order to arrive at this result,we have put a number of assumptions.If those assumptions are violated or nearly violated,then this optimality will be of no use,which is often the case in real life.

For example,we made an assumption that $var(\epsilon_i) = \sigma^2$(i.e. they are all homoscedastic),but in many practical situations,we might find that they are either correlated or they are not homoscedastic,as we have seen while measuring the length of something,if we make some of the measurements using a very precise instrument and some using a less precise instrument,in that case,we will not naturally be comfortable in taking the simple average,so if errors violate this assumption(i.e. we have a situation where the variances are not equal)then,we need to generalize this concept,which is called generalized least squares method.

$V(\vec{\epsilon}) \neq \sigma^2 I \implies$
Generalised Least Squares (GLS)

We shall go into the generalized linear squares method later,remember that the Gauss-Markov theorem is not something where you just put everything blindly and believe whatever that comes blindly.Let us now see another way by which things can go wrong.Let us assume that we are in a situation where there are no outliers.Every statistical method when it is originally created,assumes that there are no bad observations,but in practice,it is quite possible that there may be some outliers.Generally,the Least Square Estimators that we get are not robust.For example,even in the case of measuring just the length,the least squares estimator is the mean,and we know that mean is not robust.

So if there are one or two outliers,some measurements that are quite far away from the rest of the measurements will pull the mean towards them and completely ruin it.So if you suspect that there will be outliers,then you should use robust techniques and robust estimators in general are not linear.So,the fact that something is BLUE(Best Linear Unbiased Estimator) is not going to help us because it is the best among linear unbiased estimators.So if there are non-linear estimators,they might be better.It is not necessary that we should always focus our attention on linear estimators,especially when there are outliers,non-linear estimators might outperform the linear estimators.

Outliers $\Longrightarrow$
Robust (nonlinear) methods

A similar point is about unbiasedness,there are situations where we have what is called multicollinearity.Multicollinearity means that the columns of X have got some approximate dependence.For example,suppose we measure height of a person in inches and then measure the height of the same person in centimeters,so because of some amount of numerical error and some amount of measurement error,these two will not be exact multiples of each other,but very nearly 2.54 times one will be the other and so they are independent(because they are not exact multiples of each other),but almost dependent.In that case,we will find that $X^T X$ is nearly singular(because if those two columns are exact multiples of each other,then the rank of X will be one less,here it is not one less but on the verge of being that).

In such a case,we will find that it is much better idea to allow a little bit of bias than working with unbiased estimators.Unbiased estimators like least square estimators will have tremendously high variance.If we allow a little bit of bias,we can reduce the variance a lot.So the mean square error will be better.We shall see one such example,the ridge regression,later.

Multicollinearity $\Longrightarrow$
Allowing bias is good.

So,do not trust the Gauss-Markov theorem blindly.It is correct,it is mathematically proven theorem,but the assumptions that go inside them may be violated and in that case,the conclusion is also a great suspect.