Observation Theory

Script V42A – BLUE

After introducing the concepts of "least squares estimation", followed by "weighted least squares estimation", we now introduce a third estimation procedure: BLUE, or best linear unbiased estimation

Let's have a look again at our results for least squares estimation.

We found the expression for the unweighted, (or 'ordinary') least squares estimate x-hat.

It is only dependent on the design matrix A.

If we would insert a unit-matrix 'I' in this equation, we find that this UNWEIGHTED estimate is, in fact, a special case of the WEIGHTED least-squares estimate where a weight matrix W is introduced that gives some observations more importance (or weight) than others.

With these equations we now have the basic elements to obtain an estimate for x, given a specific set of observations y.

Two remarks can be made based on these equations.

First of all, we can see that the unweighted and weighted least squares estimates are basically treated in a deterministic way.

The observation vector y is not underlined, which means that we are dealing with a particular set of observations.

But also the estimate x-hat is not underlined, which implies that this vector _also_ represents a specific set of values.

In other words, we did not introduce any information on the actual quality of the observations, which could in principle be known beforehand.

The second remark is that this set of equations does not give us a clue on “how” to find the optimal weight matrix W.

It seems rather subjective to insert particular values in this matrix.

Both of these issues we are going to discuss in the context of BLUE.

But first lets recap the ‘weighted’ least squares estimation solution.

We formulated our y=Ax problem, and came to the least-squares estimate x-hat.

Essentially, what we did was to minimize the differences between our actual measurements (the red crosses) and the linear model that we assumed to be applicable (which is the green dashed line).

But we did this by giving some observations more weight than others, as expressed in the weight matrix W.

In the plot, the observations with a bigger cross represent observations with more weight.

The differences (or ‘residuals’) are indicated by the little vertical blue lines, and we call these the “least squares estimates of the residuals”, or e-hat.

Mathematically, we obtained this solution by minimizing the “weighted sum of the squared residuals”, which then yields the “weighted least squares estimate”.

Best Linear Unbiased Estimation is another method for parameter estimation, with a different historical background, and based on different concepts.

But it is also remarkably consistent with the other forms of estimation we just discussed.

Now, in this video, we could derive the BLUE estimators step by step, going through tedious derivations, and arriving at the corresponding equations at the end of a very long video.

However, I would rather spoil the story and tell you directly how it ends.

Here you see the final appearance of the best linear unbiased estimator x, (at the bottom) in comparison to the weighted least squares estimate (at the top).

Do you see the similarity and the differences?

The form of the equation is identical, but there are two differences.

First, in stead of an 'arbitrary' weight matrix W, BLUE uses one specific weight matrix which is the inverse of the covariance matrix Qy.

Remember that the covariance matrix holds information on the quality of the observations.

This makes sense, since we want to make sure that high quality observations have a greater weight to the final estimate.

And similarly, low-quality observations should have a lower weight.

The other difference between the two observations is the underline, indicating that for BLUE, we really consider the observations y and the estimator x to be stochastic variables.

(Obviously, once we insert real numbers in this equation, the underline disappears.)

To be complete: the similarity in the equations also holds for the yhat which can be elaborated further and the ehat which is elaborated in this way.

Now, there are a lot of equations on this slide, which may seem difficult if we present them this way.

But remember that a large part of this just follows from filling in the expressions.

We could take this part away again and have a final look at the result.

Thus, the main thing to remember from this introduction is that BLUE uses the quality of the observations as input to optimize the estimation.

This quality is stored in the covariance matrix of the observations.

To summarize, here we see the three different forms for parameter estimation.

The (ordinary, or unweighted) least-squares method,

The weighted least-squares method

And "Best Linear Unbiased estimation."

If you take some time to grasp the essence of the equations you could classify them in different ways.

For example, you can regard BLUE and unweighted least-squares as special cases of weighted least squares

Have a look.

The unweighted least-squares solution is the same as the weighted least-squares solution with equal weights.

And BLUE is the situation in which the weight matrix is the inverse of the covariance matrix of the observations

But we could also classify them in a deterministic class and a probabilistic class.

In this video, we introduced Best Linear Unbiased Estimation, not via some kind of mathematical derivation, but by simply stating the final set of equations.

I would like to encourage you to memorize this expression for BLUE: in my view, it is one of the most valuable equations in engineering, with a very wide range of applications.