Estimate vs Estimator

Welcome to week four of the course on observation theory. In the previous week we learned about the least squares estimation.

We saw how to find a “solution” for inconsistent functional models, based on the least squares or weighted least squares principle.

But, so far, when we discussed the least squares estimation, we treated it in a deterministic way. We defined the vector of errors or residuals, and said let’s minimise this vector… and we derive the least squares solution.

During our discussion, we did not consider the randomness in the observations, …or… in the errors and their statistical properties.

This week, we want to bring statistical properties of observations into the estimation problem, and in this video.

I would like start by introducing the concept of an “estimator” as the stochastic counterpart of an “estimate”.

To start, let’s consider the general case that we have an observation vector y, and the functional model y equals Ax.

When we estimate x, the solution, or the “estimate” xhat is a function of y: let’s say that xhat equals G(y).

For example, in the case of least squares estimation, xhat is equal to a linear function of y.

Now recall the concept of “observable” versus “observation”. The observation vector y is a realization of a random vector underline{y} which we call it the observable vector.

The values in the observable vector should be interpreted as random variables. So if we now use the observable vector in our function G, we get a random vector which we write it as underline{xhat}.

Note that, the output of the function of a random variable is itself random too. That is if we put a random variable or random vector in a function, the output is always random.

In a same way, if we apply G to the observable vector y, we get the random vector xhat.

It is important here to stress the distinction between the underline{xhat} and the xhat. The random vector xhat is called the estimator of x, and xhat is called an estimate of x.

We can say the estimate xhat is simply a single realisation of the random estimator underline{xhat}.

What I've just said, in other words, is that the estimate is a deterministic value, but the estimator is a stochastic variable or stochastic vector.

This distinction between estimates and estimators is helpful when we want to think about the quality of the results of an estimation. How close is an estimate xhat to the true but unknown value of x?

This question cannot be answered in a deterministic sense. In other words, it is not possible to evaluate the individual difference between xhat and the true value which is unknown.

However, it is possible to answer the question in a probabilistic way, and for example discuss the statistical properties of the closeness of an estimator to the true value.

Let's consider now the least squares case. The weighted least squares "estimator" is written as the following function of an observable vector y.

If the estimator is stochastic, it means it has statistical characteristics. It should have a probability distribution function. It should have an expectation or dispersion in the form of the covariance matrix of xhat.

Some important questions that we can ask are:

If the observables are normally distributed with the mean equal to Ax, then what will be the distribution of the weighted least squares estimator? Can we say that it is also normally distributed?

Or what can we say about the expectation of the estimator?

Or even about its covariance matrix?

Let's look at some exercises and then come back and continue with discussing some properties of estimators in general, and of the least squares estimator in particular.