Stochastic Model, part 2

In the previous lecture we introduced the stochastic model and looked at an example based on the tide gauge observations. In this lecture we will look at another example, where the covariance matrix takes a different form.

For most of the estimation problems that we will be looking at in this course, it will be assumed that the observations are independent and thus uncorrelated, resulting in a diagonal covariance matrix. Where the individual variances on the diagonal may or may not differ from each other.

But there are situations where the observations are correlated.

For instance, observations may be correlated in time, such that the covariances are no longer equal to zero, and the covariance matrix might be fully populated.

This may happen for instance with GPS receivers taking observations at a very high rate. A new observation is then still partially based on the preceding ones, introducing the time correlation.

The example that we will be looking at is a time series of range or distance observations to an object which is moving away from the measurement device at a constant but unknown speed.

If we assume that at time t_0 the distance is zero, the observation equation for a single observable becomes this one.

We have distance observations y_i in meters, the known times of observation t_i, and the unknown velocity v.

We will now show the 30 observations that we got with two different devices.

The green line is the true distance as function of time. At first sight you may conclude that device B which provided the observations on the right-hand side are more precise, since the random errors show a much smaller spread.

Furthermore, there is an offset from the truth, from which you might conclude that a systematic bias was present reducing the accuracy.

However, if we make a closer inspection of the corresponding random errors with both devices, something strange can be seen.

On the left-hand side we see that subsequent errors with device A can take arbitrary sign and size within a certain range, the preceding value of the error does not tell you anything about what can be expected in the next epoch.

With device B on the other hand, subsequent errors tend to be very close to each other, which is emphasized by connecting the errors with the blue line here. The reason is that these observations are strongly correlated in time.

The correlation as function of the time lag for both devices is shown in the following figures.

Recall that a correlation coefficient close to 1 indicates a strong positive correlation.

For device A, time correlation is absent – since for all time lags larger than zero, the correlation coefficient is zero.

For device B, however, subsequent observations may have a correlation coefficient very close to 1, which explains that the random errors are very similar.

And even after 30 epochs there is still a high correlation of almost 0.8, which means that in our example the last observation is still strongly correlated to the first one.

Let's now finish by showing the covariance matrices for the two devices.

Assuming the same standard deviation for all observations, for device A it is simply a scaled identity matrix.

Here the values are shown by colors. The colorbar shows that the variances on the diagonal are equal to 0.1, and zero elsewhere.

For device B, the covariance matrix is shown on the right-hand side.

The variances on the diagonal are again equal to 0.1 but the covariances are slowly decreasing away from the main diagonal.

Note that the color scaling on the right-hand side is different, and blue in this case corresponds to a covariance of .075.

This example concludes the lectures of this module.

You now know how to construct the mathematical model for a given estimation problem, and you learned about important properties of the functional model.

With this, you are ready to enter the following module where we will introduce the least squares estimation principle.