Universität Stuttgart

Prof.Dr.
Thomas Hobiger

Dynamic System Estimation

7

The discrete Kalman filter

The discrete Kalman filter

So far we have developed the framework that allows us to express a random process and the covariance of its state (with usual notation $\boldsymbol{P}$) in the form

$$\begin{aligned} \boldsymbol{x}_n &= \boldsymbol{\Phi}_{n-1}\boldsymbol{x}_{n-1} + \boldsymbol{u}_n \\ \boldsymbol{P}_n &= \boldsymbol{\Phi}_{n-1}\boldsymbol{P}_{n-1}\boldsymbol{\Phi}_{n-1}^T + \boldsymbol{Q} \end{aligned} \tag{7.1}$$

When we predict from t_{n-1} to t_n we do not have access to the actual contribution of $\boldsymbol{u}_n$, but we are sure that (unless it is zero) it increases our state covariance. Assuming now that our process is not only driven by the model expressed in (7.1) but we take observations $\boldsymbol{z}_n$ that relate to the state vector linearly

$$\boldsymbol{z}_n = \boldsymbol{H}_n\boldsymbol{x}_n + \boldsymbol{v}_n \tag{7.2}$$

where the matrix $\boldsymbol{H}_n$ provides the (noiseless) connection between the state and the observations and $\boldsymbol{v}_n$ represents the measurement errors, which are thought to be of white Gaussian noise nature and uncorrelated with $\boldsymbol{u}_n$.

We can now try to find a scheme that allows us to predict from t_{n-1} to t_n and then correct/update the state vector based on measurements collected at that epoch.

The discrete Kalman filter - cont'd

If we think of a method in which we first predict from t_{n-1} to t_n and then update the state and its covariance based on the measurements taken at t_n it is obvious that we we need to distinguish between a predicted state and an updated state. In order to make this distinction very clearly we use the subscript $n|n-1$ to indicate that the state (or its covariance) at t_n was predicted from a previous epoch t_{n-1}.

In addition we write $\hat{\boldsymbol{x}}$ when we deal with an estimated/predicted state instead of $\boldsymbol{x}$ for the true state. Thus, (7.1) becomes

$$\begin{aligned} \hat{\boldsymbol{x}}_{n|n-1} &= \boldsymbol{\Phi}_{n-1|n-1} \cdot \hat{\boldsymbol{x}}_{n-1|n-1} \\ \boldsymbol{P}_{n|n-1} &= \boldsymbol{\Phi}_{n-1|n-1} \cdot \boldsymbol{P}_{n-1|n-1} \cdot \boldsymbol{\Phi}^T_{n-1|n-1} + \boldsymbol{Q} \end{aligned} \tag{7.3}$$

Now we can ask ourselves how we can correct the predicted state $\boldsymbol{x}_{n|n-1}$ based on the measurements taken at t_n. Assuming that the difference $\boldsymbol{z}_n - \boldsymbol{H}_n\hat{\boldsymbol{x}}_{n|n-1}$ gives us an indication on how we should correct $\hat{\boldsymbol{x}}_{n|n-1}$ we can write

$$\hat{\boldsymbol{x}}_{n|n} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \left(\boldsymbol{z}_n - \boldsymbol{H}_n\hat{\boldsymbol{x}}_{n|n-1} \right) \tag{7.4}$$

where $\boldsymbol{K}_n$ is a weighting factor which we call Kalman gain and need to determine in the following

The discrete Kalman filter - cont'd

The error and covariance of the predicted state is

$$\begin{aligned} \boldsymbol{e}_{n|n-1} &= \boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1} \\ \boldsymbol{P}_{n|n-1} &= E\left(\boldsymbol{e}_{n|n-1}\boldsymbol{e}^T_{n|n-1}\right) \\ &= E\left(\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}\right)\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}\right)^T\right) \end{aligned} \tag{7.5}$$

In a similar way we can define the covariance of the updated state as

$$\boldsymbol{P}_{n|n} = E\left(\boldsymbol{e}_{n|n} \cdot \boldsymbol{e}^T_{n|n}\right) = E\left(\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}\right)\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}\right)^T\right) \tag{7.6}$$

Inserting (7.1) in (7.4) allows us to re-write (7.6) as

$$\begin{aligned} \boldsymbol{P}_{n|n} = \quad & E\left\{\left[\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}\right) - \boldsymbol{K}_n\left(\boldsymbol{H}_n\boldsymbol{x}_n + \boldsymbol{v}_n - \boldsymbol{H}_n\hat{\boldsymbol{x}}_{n|n-1}\right)\right] \cdot \right. \\ & \left.\left[\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}\right) - \boldsymbol{K}_n\left(\boldsymbol{H}_n\boldsymbol{x}_n + \boldsymbol{v}_n - \boldsymbol{H}_n\hat{\boldsymbol{x}}_{n|n-1}\right)\right]^T\right\} \end{aligned} \tag{7.7}$$

The discrete Kalman filter - cont'd

Since the prediction error $\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1}$ is uncorrelated with the measurement error $\boldsymbol{v}_n$ we can write (7.7) as

$$\boxed{\boldsymbol{P}_{n|n} = (\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H}_n)\,\boldsymbol{P}_{n|n-1}\,(\boldsymbol{I} - \boldsymbol{K}_n\boldsymbol{H}_n)^T + \boldsymbol{K}_n\boldsymbol{R}_n\boldsymbol{K}_n^T} \qquad (7.8)$$

where $\boldsymbol{R}_n$ is the covariance of the measurement noise, i.e. $\boldsymbol{R}_n = E(\boldsymbol{v}_n\boldsymbol{v}_n^T)$.

The basic idea of the Kalman filter is to minimize the mean-square error of the a posteriori state estimation, i.e. finding an optimal state $\hat{\boldsymbol{x}}_{n|n}$ that minimizes $E\left(\left(\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n}\right)^2\right)$. This is equivalent to minimizing the trace of the a posteriori estimate covariance matrix $\boldsymbol{P}_{n|n}$, i.e.

$$\boldsymbol{J}_n = \mathrm{tr}\left(\boldsymbol{P}_{n|n}\right) = min. \Rightarrow \frac{\partial \boldsymbol{J}_n}{\partial \boldsymbol{K}_n} = \boldsymbol{0} \qquad (7.9)$$

A (useful) relation for the following computations is

$$\frac{\partial \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}\boldsymbol{A}^T)}{\partial \boldsymbol{A}} = 2\boldsymbol{A}\boldsymbol{C}$$

if $\boldsymbol{C}$ is symmetric.

The discrete Kalman filter - cont'd

Thus we get from the two terms on the RHS in (7.9)

$$\frac{\partial \mathrm{tr}\left(\left(\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}_n\right)\boldsymbol{P}_{n|n-1}\left(\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}_n\right)^T\right)}{\partial \boldsymbol{K}_n} = -2\left(\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}_n\right)\boldsymbol{P}_{n|n-1}\boldsymbol{H}_n^T \quad (7.10)$$

and

$$\frac{\partial \mathrm{tr}\left(\boldsymbol{K}_n\boldsymbol{R}_n\boldsymbol{K}_n^T\right)}{\partial \boldsymbol{K}_n} = 2\boldsymbol{K}_n\boldsymbol{R}_n \quad (7.11)$$

Furthermore

$$\frac{\partial \boldsymbol{J}_n}{\partial \boldsymbol{K}_n} = \boldsymbol{0} = -2\left(\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}_n\right)\boldsymbol{P}_{n|n-1}\boldsymbol{H}_n^T + 2\boldsymbol{K}_n\boldsymbol{R}_n \quad (7.12)$$

which allows to solve for $\boldsymbol{K}_n$

$$\boxed{\boldsymbol{K}_n = \boldsymbol{P}_{n|n-1}\boldsymbol{H}_n^T\left(\boldsymbol{H}_n\boldsymbol{P}_{n|n-1}\boldsymbol{H}_n^T + \boldsymbol{R}_n\right)^{-1}} \quad (7.13)$$

and yields an analytic expression for the a posteriori covariance of the state

$$\boxed{\boldsymbol{P}_{n|n} = \left(\boldsymbol{I}-\boldsymbol{K}_n\boldsymbol{H}_n\right)\boldsymbol{P}_{n|n-1}} \quad (7.14)$$

The discrete Kalman filter - cont'd

In summary the Kalman filter can be sketched out as follows:

Project ahead:

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{\Phi}_{n-1|n-1} \cdot \hat{\boldsymbol{x}}_{n-1|n-1}$$

$$\boldsymbol{P}_{n|n-1} = \boldsymbol{\Phi}_{n-1|n-1} \boldsymbol{P}_{n-1|n-1} \boldsymbol{\Phi}^T_{n-1|n-1} + \boldsymbol{Q}$$

Compute Kalman gain:

$$\boldsymbol{K}_n = \boldsymbol{P}_{n|n-1} \boldsymbol{H}_n^T \left(\boldsymbol{H}_n \boldsymbol{P}_{n|n-1} \boldsymbol{H}_n^T + \boldsymbol{R}_n \right)^{-1}$$

Update with measurement:

$$\hat{\boldsymbol{x}}_{n|n} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \left(\boldsymbol{z}_n - \boldsymbol{H}_n \hat{\boldsymbol{x}}_{n|n-1} \right)$$

Compute error a posteriori covariance:

$$\boldsymbol{P}_{n|n} = \left(\boldsymbol{I} - \boldsymbol{K}_n \boldsymbol{H}_n \right) \boldsymbol{P}_{n|n-1}$$

The discrete Kalman filter - cont'd

You will find the examples discussed in this lecture as Jupyter notebook under

`https://github.com/spacegeodesy/ParameterEstimationDynamicSystems/blob/master/example07.ipynb`