Universität Stuttgart

Prof.Dr.
Thomas Hobiger

Dynamic System Estimation

Backward filter and smoothing

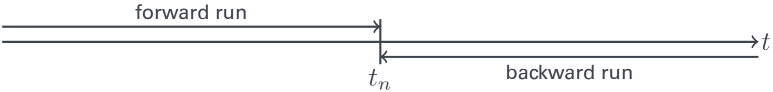
8

Backward filter and smoothing

The forward (real time) Kalman filter requires $t_{n-1} < t_n$ and updates the state vector and its variance by

$$\boxed{\begin{aligned} \hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \left(\boldsymbol{z}_n - \boldsymbol{H}_n \hat{\boldsymbol{x}}_{n|n-1} \right) \\ \boldsymbol{P}_{n|n} &= \left(\boldsymbol{I} - \boldsymbol{K}_n \boldsymbol{H}_n \right) \boldsymbol{P}_{n|n-1} \end{aligned}} \tag{8.1}$$

As long as we apply the correct state transition model one can run also the Kalman filter in time-backward direction.



This is of course only possible in post-processing, i.e. non-real time. We denote the estimated states from the backward run with $\hat{\boldsymbol{x}}^b_{n|n}$ and $\boldsymbol{P}^b_{n|n}$.

Question: How to combine the results from the forward run for an optimal estimate based on all data?

Backward filter and smoothing - cont'd

Linear combination (weighted average)

$$
\begin{aligned}
\hat{\boldsymbol{x}}_n &= \boldsymbol{A}\hat{\boldsymbol{x}}_{n|n} + (\boldsymbol{I}-\boldsymbol{A})\hat{\boldsymbol{x}}^b_{n|n} \\
\boldsymbol{P}_n &= \boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T + (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T
\end{aligned}
\tag{8.2}
$$

$\boldsymbol{A}$ and $(\boldsymbol{I}-\boldsymbol{A})$ are the weight matrices.

Question: How to choose $\boldsymbol{A}$?

Selection of an optimality criterion. Minimize the trace of covariance matrix of the result.

$$
\mathrm{tr}(\boldsymbol{P}_n) = \mathrm{tr}\left(\boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T + (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T\right) \rightarrow \min. \tag{8.3}
$$

Take the derivative of equ. (8.3) w.r.t. the matrix $\boldsymbol{A}$ and equate to zero.

Backward filter and smoothing - cont'd

A useful relation (see also previous lecture) is the following, given that the matrix $\boldsymbol{C}$ is symmetric.

$$\frac{\partial \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}\boldsymbol{A}^T)}{\partial \boldsymbol{A}} = 2\boldsymbol{A}\boldsymbol{C} \tag{8.4}$$

With that we start with

$$\begin{aligned}&\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T + (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T\right) = \\ &\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T\right) + \mathrm{tr}\left((\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T\right)\end{aligned} \tag{8.5}$$

and then compute the partial derivative w.r.t. $\boldsymbol{A}$

$$\begin{aligned}&\frac{\partial}{\partial \boldsymbol{A}}\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T + (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T\right) = \\ &\frac{\partial}{\partial \boldsymbol{A}}\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{P}_{n|n}\boldsymbol{A}^T\right) + \frac{\partial}{\partial \boldsymbol{A}}\mathrm{tr}\left((\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(\boldsymbol{I}-\boldsymbol{A})^T\right)\end{aligned} \tag{8.6}$$

which we then evaluate and equate to $\boldsymbol{0}$, i.e.

$$2\boldsymbol{A}\boldsymbol{P}_{n|n} + 2(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}^b_{n|n}(-\boldsymbol{I}) = \boldsymbol{0} \tag{8.7}$$

Backward filter and smoothing - cont'd

From equ. (8.7) we get

$$\boldsymbol{A} = \boldsymbol{P}^b_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1} \quad \text{and} \quad \boldsymbol{I} - \boldsymbol{A} = \boldsymbol{P}_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1} \tag{8.8}$$

Inserting equ. (8.8) in equ. (8.2) gives

$$\begin{aligned} \boldsymbol{P}_n = \quad & \boldsymbol{P}^b_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1} \boldsymbol{P}_{n|n}\left(\boldsymbol{P}^b_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1}\right)^T + \\ & \boldsymbol{P}_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1} \boldsymbol{P}^b_{n|n}\left(\boldsymbol{P}_{n|n}\left(\boldsymbol{P}^b_{n|n} + \boldsymbol{P}_{n|n}\right)^{-1}\right)^T \end{aligned} \tag{8.9}$$

If we use the matrix identity $\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B} = (\boldsymbol{A}^{-1} + \boldsymbol{B})^{-1})^{-1}$ we get

$$\boxed{\begin{aligned} \boldsymbol{P}_n = \quad & \left((\boldsymbol{P}^b_{n|n})^{-1} + (\boldsymbol{P}_{n|n})^{-1}\right)^{-1} \\ \hat{\boldsymbol{x}}_n = \quad & \boldsymbol{P}_n\left((\boldsymbol{P}_{n|n})^{-1}\hat{\boldsymbol{x}}_{n|n} + (\boldsymbol{P}^b_{n|n})^{-1}\hat{\boldsymbol{x}}^b_{n|n}\right) \end{aligned}} \tag{8.10}$$

You will find the examples discussed in this lecture as Jupyter notebook under
`https://github.com/spacegeodesy/ParameterEstimationDynamicSystems/blob/master/example08.ipynb`