

Advanced Time Series Analysis

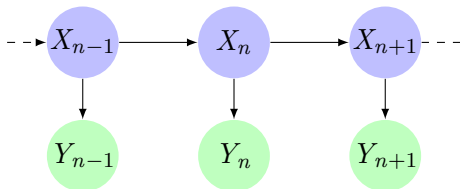
State Space Models and the Kalman Filter

Henrik Madsen

Renewable Energy Forecasting - Theory and Practice
Forecasting Summer School
International Institute of Forecasters (IIF)
`hmad@dtu.dk`

Summer/Fall 2020

State space models



- System model; A full description of the dynamical system (i.e. including the parameters):

$$X_t = f(X_{t-1}) + g(u_{t-1}) + e_{1,t}$$

- Observations; Noisy measurements of some parts (states) of the system:

$$Y_t = h(X_t) + e_{2,t}$$

- Goal; reconstruct and predict the state of the system

The linear stochastic state space model

System equation: $\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \mathbf{e}_{1,t}$

Observation equation: $\mathbf{Y}_t = \mathbf{C}\mathbf{X}_t + \mathbf{e}_{2,t}$

- $\mathbf{X}$: State vector
- $\mathbf{Y}$: Observation vector
- $\mathbf{u}$: Input vector
- $\mathbf{e}_1$: System noise
- $\mathbf{e}_2$: Observation noise
- $\dim(\mathbf{X}_t) = m$ is called the order of the system
- $\{\mathbf{e}_{1,t}\}$ and $\{\mathbf{e}_{2,t}\}$ mutually independent white noise
- $V[\mathbf{e}_1] = \Sigma_1$, $V[\mathbf{e}_2] = \Sigma_2$
- $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, Σ_1 , and Σ_2 are **known** matrices
- The state vector contains all information available for future evaluation; the process is a *Markov process*.
- Important: The formulation makes it possible to handle time-varying systems as well.

The Kalman filter

Initialization:

$$\begin{aligned}\hat{\mathbf{X}}_{1|0} &= E[\mathbf{X}_1] = \boldsymbol{\mu}_0 \\ \boldsymbol{\Sigma}_{1|0}^{xx} &= V[\mathbf{X}_1] = \mathbf{V}_0 \\ \boldsymbol{\Sigma}_{1|0}^{yy} &= \mathbf{C}\boldsymbol{\Sigma}_{1|0}^{xx}\mathbf{C}^T + \boldsymbol{\Sigma}_2\end{aligned}$$

For: $t = 1, 2, 3, \dots$

Reconstruction:

$$\begin{aligned}\mathbf{K}_t &= \boldsymbol{\Sigma}_{t|t-1}^{xx} \mathbf{C}^T \left(\boldsymbol{\Sigma}_{t|t-1}^{yy} \right)^{-1} \\ \hat{\mathbf{X}}_{t|t} &= \hat{\mathbf{X}}_{t|t-1} + \mathbf{K}_t \left(\mathbf{Y}_t - \mathbf{C}\hat{\mathbf{X}}_{t|t-1} \right) \\ \boldsymbol{\Sigma}_{t|t}^{xx} &= \boldsymbol{\Sigma}_{t|t-1}^{xx} - \mathbf{K}_t \boldsymbol{\Sigma}_{t|t-1}^{yy} \mathbf{K}_t^T\end{aligned}$$

Prediction:

$$\begin{aligned}\hat{\mathbf{X}}_{t+1|t} &= \mathbf{A}\hat{\mathbf{X}}_{t|t} + \mathbf{B}\mathbf{u}_t \\ \boldsymbol{\Sigma}_{t+1|t}^{xx} &= \mathbf{A}\boldsymbol{\Sigma}_{t|t}^{xx}\mathbf{A}^T + \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_{t+1|t}^{yy} &= \mathbf{C}\boldsymbol{\Sigma}_{t+1|t}^{xx}\mathbf{C}^T + \boldsymbol{\Sigma}_2\end{aligned}$$

Multi-step predictions

- Not part of the Kalman filter as stated above
- Can be calculated recursively for a given t starting with $k = 1$ for which $\hat{\mathbf{X}}_{t+k|t}$ and $\Sigma_{t+k|t}$ are calculated in the Kalman prediction step.

$$\hat{\mathbf{X}}_{t+k+1|t} = \mathbf{A}\hat{\mathbf{X}}_{t+k|t} + \mathbf{B}\mathbf{u}_{t+k}$$

$$\Sigma_{t+k+1|t}^{xx} = \mathbf{A}\Sigma_{t+k|t}^{xx}\mathbf{A}^T + \Sigma_1$$

- The future input must be known/assumed.

The Foundation of the Kalman filter

- Theorem 2.6 (Linear projection)
- The theorem is concerned with the random vectors $\mathbf{X}$ and $\mathbf{Y}$ for which the means, variances and covariances are used
- The state is called $\mathbf{X}_t$ and the observation is called $\mathbf{Y}_t$ and we could write down the theorem for these
- We have additional information; $\mathcal{Y}_{t-1}^T = (\mathbf{Y}_1^T, \dots, \mathbf{Y}_{t-1}^T)$
- We include this information by considering the random vectors $\mathbf{X}_t|\mathcal{Y}_{t-1}$ and $\mathbf{Y}_t|\mathcal{Y}_{t-1}$ instead

$$\begin{aligned} E[(\mathbf{X}_t|\mathcal{Y}_{t-1}) | (\mathbf{Y}_t|\mathcal{Y}_{t-1})] &= E[\mathbf{X}_t | \mathbf{Y}_t, \mathcal{Y}_{t-1}] = \\ &E[\mathbf{X}_t | \mathcal{Y}_{t-1}] + \text{Cov}[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] V^{-1}[\mathbf{Y}_t | \mathcal{Y}_{t-1}] (\mathbf{Y}_t - E[\mathbf{Y}_t | \mathcal{Y}_{t-1}]) \end{aligned}$$

$$\begin{aligned} V[(\mathbf{X}_t|\mathcal{Y}_{t-1}) | (\mathbf{Y}_t|\mathcal{Y}_{t-1})] &= V[\mathbf{X}_t | \mathbf{Y}_t, \mathcal{Y}_{t-1}] = \\ &V[\mathbf{X}_t | \mathcal{Y}_{t-1}] - \text{Cov}[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] V^{-1}[\mathbf{Y}_t | \mathcal{Y}_{t-1}] \text{Cov}^T[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] \end{aligned}$$

The Foundation of the Kalman filter II

$$\begin{aligned} E[\mathbf{X}_t | \mathbf{Y}_t, \mathcal{Y}_{t-1}] &= \\ E[\mathbf{X}_t | \mathcal{Y}_{t-1}] &+ \text{Cov}[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] V^{-1} [\mathbf{Y}_t | \mathcal{Y}_{t-1}] (\mathbf{Y}_t - E[\mathbf{Y}_t | \mathcal{Y}_{t-1}]) \\ V[\mathbf{X}_t | \mathbf{Y}_t, \mathcal{Y}_{t-1}] &= \\ V[\mathbf{X}_t | \mathcal{Y}_{t-1}] &- \text{Cov}[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] V^{-1} [\mathbf{Y}_t | \mathcal{Y}_{t-1}] \text{Cov}^T[\mathbf{X}_t, \mathbf{Y}_t | \mathcal{Y}_{t-1}] \end{aligned}$$

Using definitions from previous slide the update equations are:

$$\begin{aligned} \hat{\mathbf{X}}_{t|t} &= \hat{\mathbf{X}}_{t|t-1} + \Sigma_{t|t-1}^{xy} \left(\Sigma_{t|t-1}^{yy} \right)^{-1} \left(\mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1} \right) \\ \Sigma_{t|t}^{xx} &= \Sigma_{t|t-1}^{xx} - \Sigma_{t|t-1}^{xy} \left(\Sigma_{t|t-1}^{yy} \right)^{-1} \left(\Sigma_{t|t-1}^{xy} \right)^T \\ \mathbf{K}_t &= \Sigma_{t|t-1}^{xy} \left(\Sigma_{t|t-1}^{yy} \right)^{-1} \end{aligned}$$

$\mathbf{K}_t$ is called the *Kalman gain*, because it determines how much the 1-step prediction error influence the update of the state estimate

The Foundation of the Kalman filter III

The 1-step predictions are obtained directly from the state space model:

$$\begin{aligned}\widehat{\mathbf{X}}_{t+1|t} &= \mathbf{A}\widehat{\mathbf{X}}_{t|t} + \mathbf{B}\mathbf{u}_t \\ \widehat{\mathbf{Y}}_{t+1|t} &= \mathbf{C}\widehat{\mathbf{X}}_{t+1|t}\end{aligned}$$

Which results in the prediction errors:

$$\begin{aligned}\widetilde{\mathbf{X}}_{t+1|t} &= \mathbf{X}_{t+1} - \widehat{\mathbf{X}}_{t+1|t} = \mathbf{A}\widetilde{\mathbf{X}}_{t|t} + \mathbf{e}_{1,t+1} \\ \widetilde{\mathbf{Y}}_{t+1|t} &= \mathbf{Y}_{t+1} - \widehat{\mathbf{Y}}_{t+1|t} = \mathbf{C}\widetilde{\mathbf{X}}_{t+1|t} + \mathbf{e}_{2,t+1}\end{aligned}$$

And therefore:

$$\begin{aligned}\Sigma_{t+1|t}^{xx} &= \mathbf{A}\Sigma_{t|t}^{xx}\mathbf{A}^T + \Sigma_1 \\ \Sigma_{t+1|t}^{yy} &= \mathbf{C}\Sigma_{t+1|t}^{xx}\mathbf{C}^T + \Sigma_2 \\ \Sigma_{t+1|t}^{xy} &= \Sigma_{t+1|t}^{xx}\mathbf{C}^T\end{aligned}$$