Time Series Analysis

Week 10 – State space models, 1st part

Peder Bacher

Department of Applied Mathematics and Computer Science

Technical University of Denmark

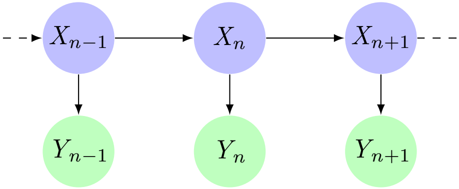
April 12, 2024

Week 10: Outline of the lecture

State space models, 1st part:

- The advantages
- The linear state space model
- Determining model structure
- Example
- An example on application of the Kalman filter.

State space models



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

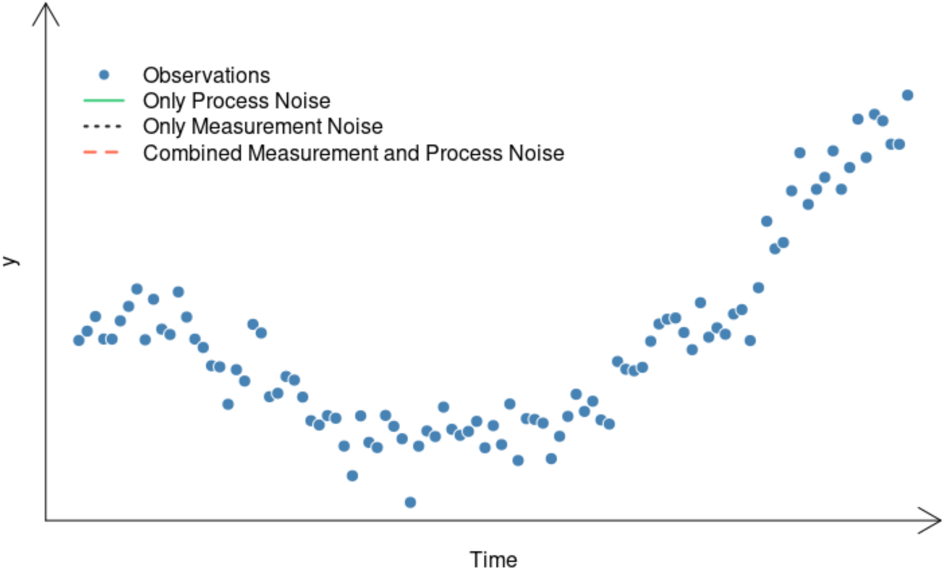
$$X_t = f(X_{t-1}) + g(u_{t-1}) + e_{1,t}$$

- 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

State space models



The linear stochastic state space model

$$\text{System equation:} \quad \boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{B}\boldsymbol{u}_{t-1} + \boldsymbol{e}_{1,t}$$
$$\text{Observation equation:} \quad \boldsymbol{Y}_t = \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{e}_{2,t}$$

- $\boldsymbol{X}$: State vector
- $\boldsymbol{Y}$: Observation vector
- $\boldsymbol{u}$: Input vector
- $\boldsymbol{e}_1$: System noise
- $\boldsymbol{e}_2$: Observation noise

- $\dim(X_t) = m$ is called the order of the system
- $\{\boldsymbol{e}_{1,t}\}$ and $\{\boldsymbol{e}_{2,t}\}$ mutually independent white noise
- $V[\boldsymbol{e}_1] = \boldsymbol{\Sigma}_1$, $V[\boldsymbol{e}_2] = \boldsymbol{\Sigma}_2$
- $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_2$ are **known** matrices

- The state vector contains all information available for future evaluation; the process is a *Markov process*.

Examples

- Find examples of systems where state-space models are preferred compared to models introduced previously in this course.
- What is the typical kind of system for which state-space models are needed?

Determining the model structure

1. The system model is often based on physical considerations; start by formulating the model using differential equations
2. Rewrite m'th order differential equation as m 1st order differential equations.
3. Find the discrete-time model for a particular time step by formulating the 1-step predictions.
4. Pray that the resulting model is linear or enrol in Advanced Time Series Analysis (02427).
5. Add noise to appropriate states.
6. Formulate observation equation.

Example – a falling body

An object is dropped at some height with some initial velocity. We want to estimate the position over time given noisy observations of it.

1. Formulate physical equations: $\frac{d^2z}{dt^2} = -g$
2. Rewrite second-order ODE as two first-order ODEs:

$$\boldsymbol{x}'(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ -1 \end{bmatrix} g$$

.

3. Find discrete time model for time-step: $(T = 1)$

$$\boldsymbol{x}_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_{k-1} + \begin{bmatrix} -1/2 \\ -1 \end{bmatrix} g.$$

4. It's linear, thanks god!
5. Add noise: $\boldsymbol{x}_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_{k-1} + \begin{bmatrix} -1/2 \\ -1 \end{bmatrix} g + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{e}_{1,k}$.
6. Formulate observation equation: $\boldsymbol{y}_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}_k + e_{2,k}$.

How to use the model – a falling body

Given measurements of the position at time points $1, 2, \ldots, k$ we could:

- **Predict** the future position and velocity $\boldsymbol{x}_{k+n|k}$ $(n > 0)$.
- **Reconstruct** the current position and velocity from noisy measurements $\boldsymbol{x}_{k|k}$.
- **Smooth** to find the best estimate of the position and velocity at a previous time point $\boldsymbol{x}_{k+n|k}$ $(n < 0)$ (estimate the path in the state space).

Requirement – observability

In order to predict, reconstruct or smooth, the system needs to be observable

$$\boldsymbol{X}_t = \boldsymbol{A}\boldsymbol{X}_{t-1} + \boldsymbol{B}\boldsymbol{u}_{t-1} + \boldsymbol{e}_{1,t}$$
$$\boldsymbol{Y}_t = \boldsymbol{C}\boldsymbol{X}_t + \boldsymbol{e}_{2,t}$$

- Observability is the ability to measure the internal states of a system by examining its outputs.
- Provide examples, either mathematically or intuitively where this is not the case.
- In general the linear state space model is observable if and only if:

$$\text{rank}\left[\boldsymbol{C}^T \vdots (\boldsymbol{C}\boldsymbol{A})^T \vdots \cdots \vdots \left(\boldsymbol{C}\boldsymbol{A}^{m-1}\right)^T\right] = m.$$

- For the falling body (from the discrete-time description of the system):

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$\left[C^T \vdots (CA)^T\right] = \left[\begin{matrix} 1 \\ 0 \end{matrix} \vdots \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right)^T\right] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

```
> qr( cbind(t(C), t(C %*% A)) )$rank
[1] 2
```

The Kalman filter

Initialization:

$$\widehat{\boldsymbol{X}}_{1|0} = E[\boldsymbol{X}_1] = \boldsymbol{\mu}_0$$
$$\boldsymbol{\Sigma}^{xx}_{1|0} = V[\boldsymbol{X}_1] = \boldsymbol{V}_0$$
$$\boldsymbol{\Sigma}^{yy}_{1|0} = \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{1|0}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2$$

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

Reconstruction:

$$\boldsymbol{K}_t = \boldsymbol{\Sigma}^{xx}_{t|t-1}\boldsymbol{C}^T\left(\boldsymbol{\Sigma}^{yy}_{t|t-1}\right)^{-1}$$
$$\widehat{\boldsymbol{X}}_{t|t} = \widehat{\boldsymbol{X}}_{t|t-1} + \boldsymbol{K}_t\left(\boldsymbol{Y}_t - \boldsymbol{C}\widehat{\boldsymbol{X}}_{t|t-1}\right)$$
$$\boldsymbol{\Sigma}^{xx}_{t|t} = \boldsymbol{\Sigma}^{xx}_{t|t-1} - \boldsymbol{K}_t\boldsymbol{\Sigma}^{yy}_{t|t-1}\boldsymbol{K}_t^T$$

Prediction:

$$\widehat{\boldsymbol{X}}_{t+1|t} = \boldsymbol{A}\widehat{\boldsymbol{X}}_{t|t} + \boldsymbol{B}\boldsymbol{u}_t$$
$$\boldsymbol{\Sigma}^{xx}_{t+1|t} = \boldsymbol{A}\boldsymbol{\Sigma}^{xx}_{t|t}\boldsymbol{A}^T + \boldsymbol{\Sigma}_1$$
$$\boldsymbol{\Sigma}^{yy}_{t+1|t} = \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{t+1|t}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2$$

Example: The falling body revised

Description of the system:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \boldsymbol{B} = \begin{bmatrix} -1/2 \\ -1 \end{bmatrix} \qquad \boldsymbol{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 2.0 & 0.8 \\ 0.8 & 1.0 \end{bmatrix} \qquad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 10000 \end{bmatrix}$$

Initialization: Released 10000 m above ground at 0 m/s

$$\widehat{\boldsymbol{X}}_{1|0} = \begin{bmatrix} 10000 \\ 0 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{1|0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{yy}_{1|0} = \begin{bmatrix} 10000 \end{bmatrix}$$

Simulation of a falling body – initialization

```r
z0 <- 10000
A <- matrix(c(1,0,1,1),nrow=2)
B <- matrix(c(-.5,-1),nrow=2)
C <- matrix(c(1,0),nrow=1)
Sigma1 <- matrix(c(2,.8,.8,1),nrow=2)
Sigma2 <- matrix(10000)
g <- 9.82; N <- 300
X <- matrix(nrow=2,ncol=N) ## Allocating space
X[,1] <- c(z0,0)
Y <- numeric(N)
Y[1] <- C%*%X[,1]+sqrt(Sigma2) %*% rnorm(1)
```

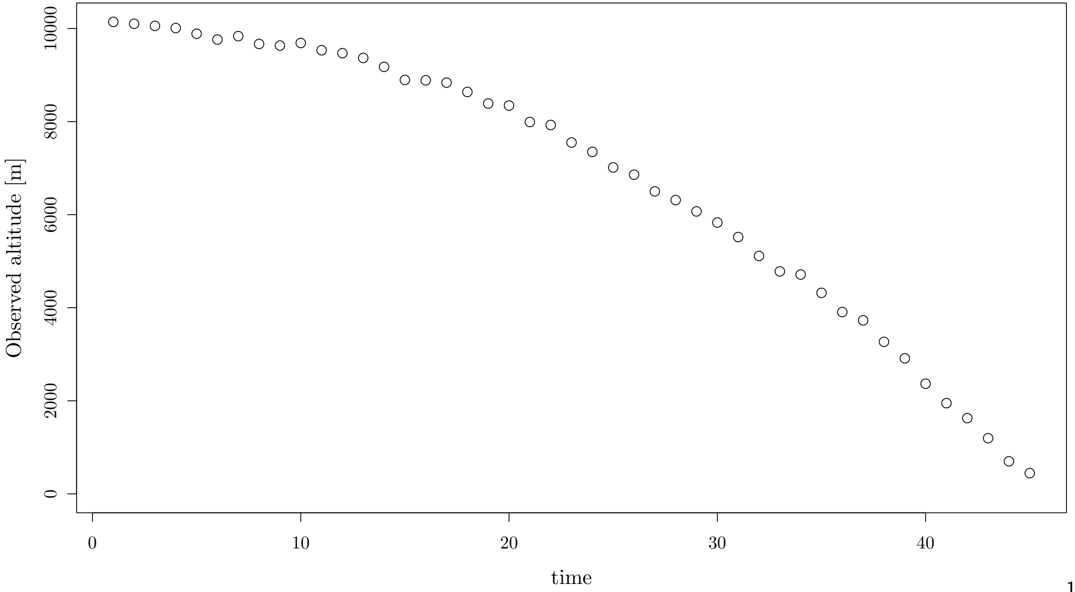
Simulation of a falling body - simulation

```
for (I in 2:N){
    X[,I] <- A %*% X[,I-1,drop=FALSE] + B%*%g +
        chol(Sigma1) %*% matrix(rnorm(2),ncol=1)
    Y[I] <- C %*% X[,I] + sqrt(Sigma2) %*% rnorm(1)
}
Nhit <- min(which(X[1,]<0))-1
X <- X[,1:Nhit]
Y <- Y[1:Nhit]
```

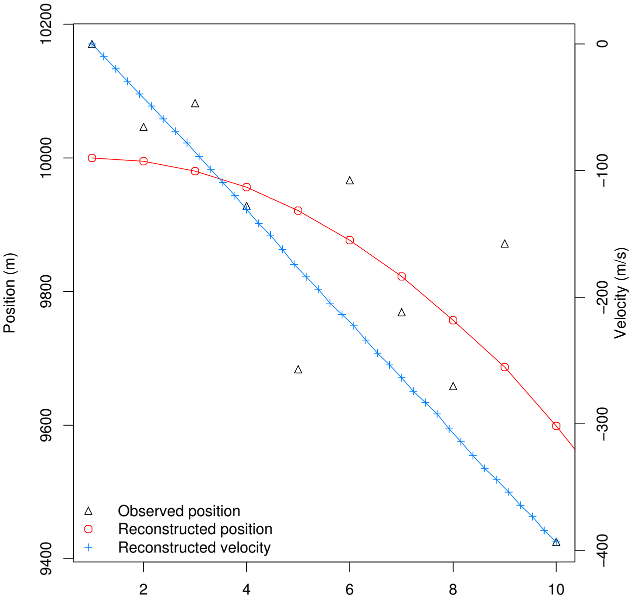
- Remember that if $Z \sim N(0, I)$, then $Y = QZ \sim N(0, QQ^T)$.
- The Cholesky factorization is one way to solve $QQ^T = \Sigma$ for Q.

The falling body – observations

```
plot(Y,xlab="time",ylab="Observed altitude [m]", ylim=c(0,max(Y)))
```



Falling body – the 10 first time points



Falling body – wrong initial state