02417: Time Series Analysis

Week 10 – State space models, 1st part

Christian Ankerstjerne Thilker
DTU Compute

Based on previous material from the course

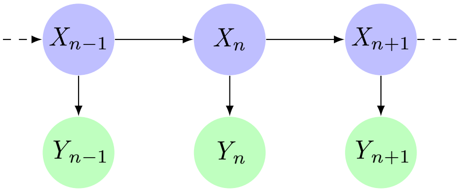
April 8, 2026

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.

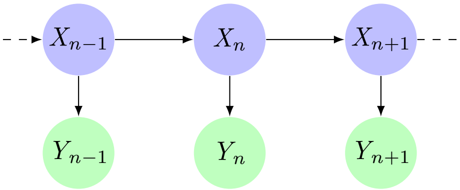
State space models



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

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

State space models



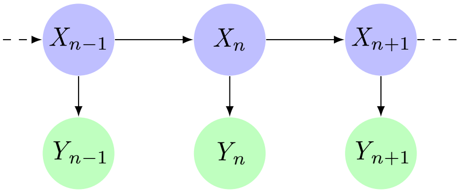
- System model; A full description of the dynamical system (i.e. including the parameters):

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

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

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

State space models



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

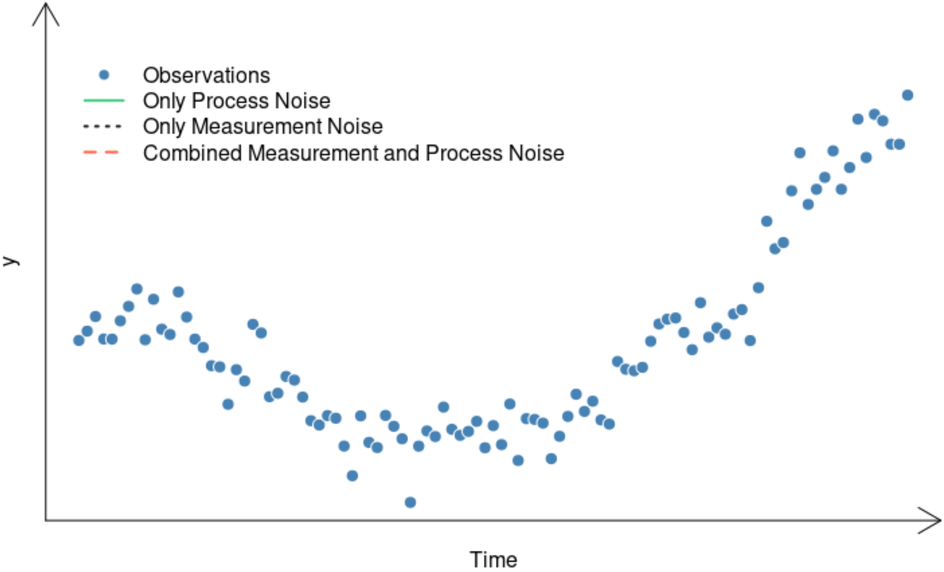
$$X_t = f(X_{t-1}) + g(u_{t-1}) + e_{1,t}$$

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

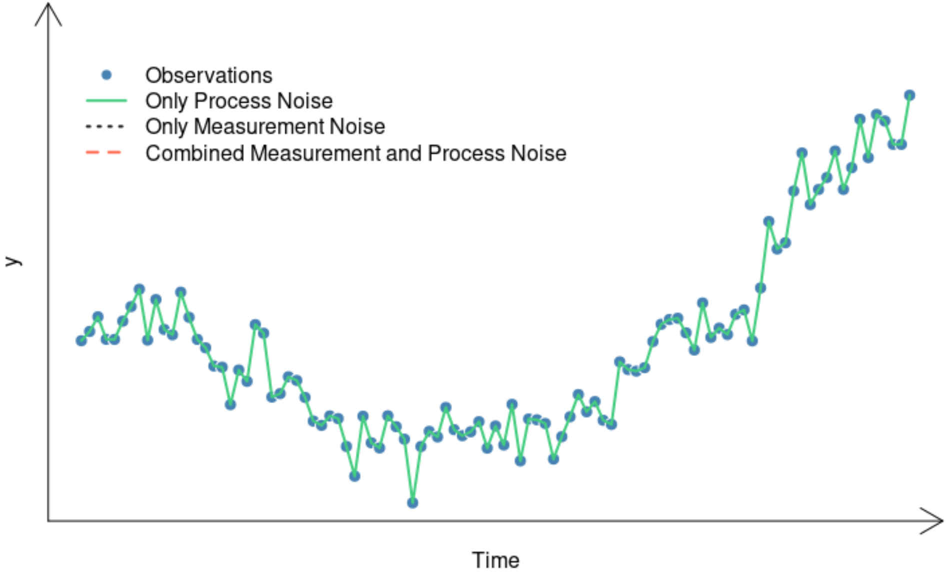
$$Y_t = h(X_t) + e_{2,t}$$

- Goal; reconstruct and predict the state of the system

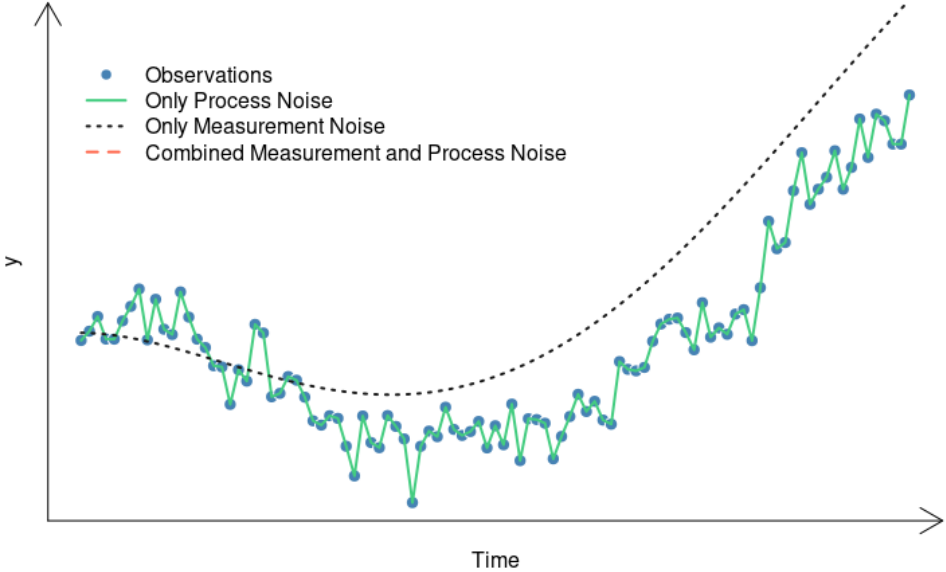
State space models



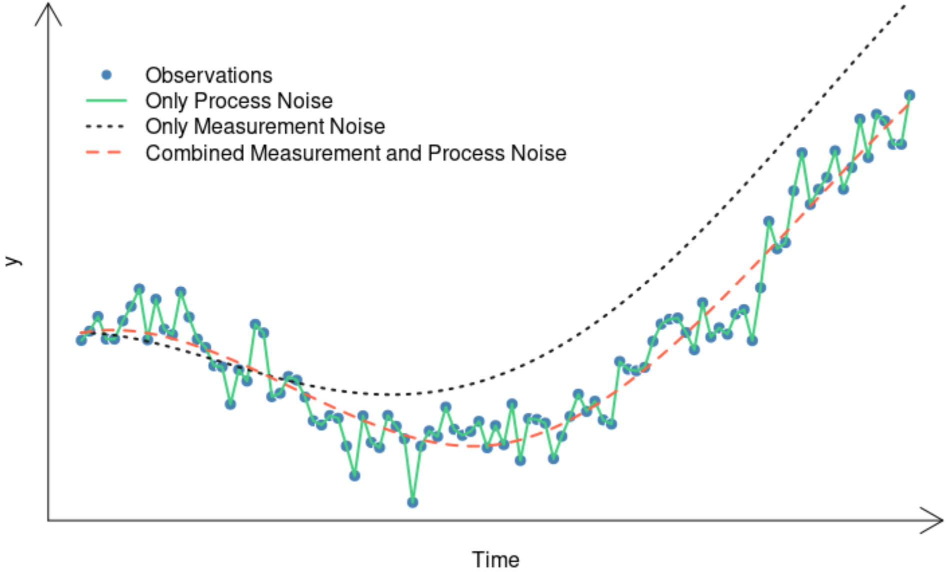
State space models



State space models



State space models



The linear stochastic state space model

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

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

The linear stochastic state space model

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

- $\boldsymbol{X}$: 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 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

- Typically physical systems where where state-space models are needed:

Examples

- Typically physical systems where where state-space models are needed: The state and input can be separated (which is useful for control).

Examples

- Typically physical systems where where state-space models are needed: The state and input can be separated (which is useful for control).
- Estimate the temperature inside a solid block of material when we measure the temperature on the surface (with noise)

Examples

- Typically physical systems where where state-space models are needed: The state and input can be separated (which is useful for control).
- Estimate the temperature inside a solid block of material when we measure the temperature on the surface (with noise)
- Noisy measurements of the position of a ship; give a better estimate of the actual current position

Examples

- Typically physical systems where where state-space models are needed: The state and input can be separated (which is useful for control).
- Estimate the temperature inside a solid block of material when we measure the temperature on the surface (with noise)
- Noisy measurements of the position of a ship; give a better estimate of the actual current position
- PK/PD-modelling: State: Amount of drug in blood, liver, muscles, ... Observations: Amount in blood (with noise), Input: Drug.

Examples

- Typically physical systems where where state-space models are needed: The state and input can be separated (which is useful for control).
- Estimate the temperature inside a solid block of material when we measure the temperature on the surface (with noise)
- Noisy measurements of the position of a ship; give a better estimate of the actual current position
- PK/PD-modelling: State: Amount of drug in blood, liver, muscles, ... Observations: Amount in blood (with noise), Input: Drug.
- Amount of persons in room: State CO2 concentration, amount of persons in room Observation: CO2 concentration (with noise)

Determining the model structure

- The system model is often based on physical considerations; start by formulating the model using differential equations

Determining the model structure

- The system model is often based on physical considerations; start by formulating the model using differential equations
- Rewrite m'th order differential equation as m 1st order differential equations.

Determining the model structure

- The system model is often based on physical considerations; start by formulating the model using differential equations
- Rewrite m'th order differential equation as m 1st order differential equations.
- Find the discrete-time model for a particular time step by formulating the 1-step predictions.

Determining the model structure

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

Determining the model structure

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

Determining the model structure

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

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.

- Formulate physical equations: $\frac{d^2z}{dt^2} = -g$

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.

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

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

.

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.

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

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

.

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

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

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.

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

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

.

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

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

- It's linear, thanks god!

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.

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

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

.

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

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

- It's linear, thanks god!
- Add noise:

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.

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

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

.

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

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

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

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.

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

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

.

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

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

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

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

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

.

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

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

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

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 $\boldsymbol{x}_{k+n|k}$ $(n > 0)$.

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 $\boldsymbol{x}_{k|k}$.

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).

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).

But how do we do this?

How to use the model – a falling body

Given measurements of the position at time points $1, 2, \dots, k$ 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).

But how do we do this? Answer: *The Kalman filter*

Requirement – observability

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

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

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.

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.

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:

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

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):

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

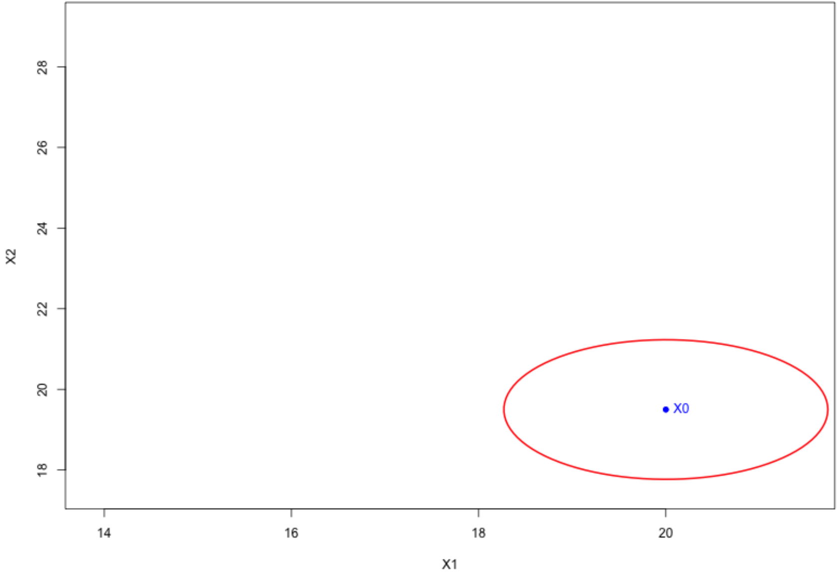
$$\left[C^T \vdots (CA)^T\right] = \left[\begin{matrix} 1 \\ 0 \end{matrix} \vdots \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right)^T\right] = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

The Kalman filter

What is the Kalman filter?

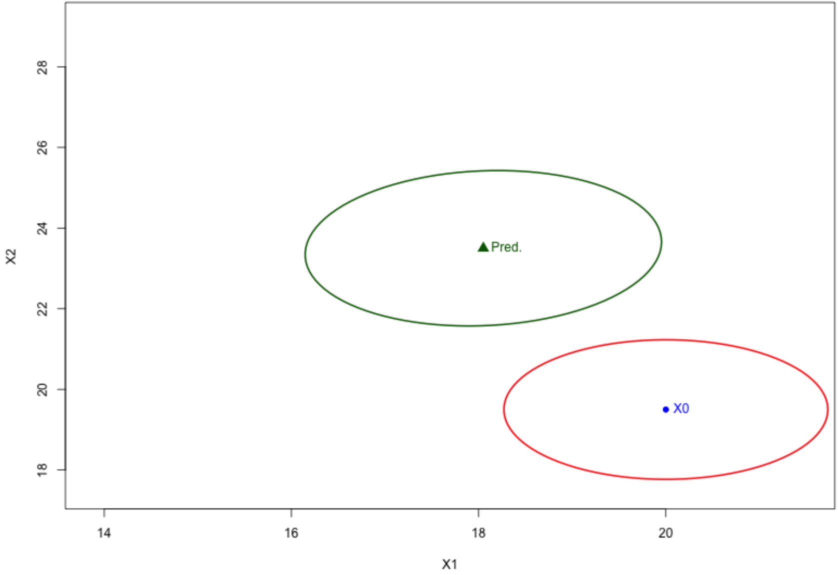
The Kalman filter

What is the Kalman filter?



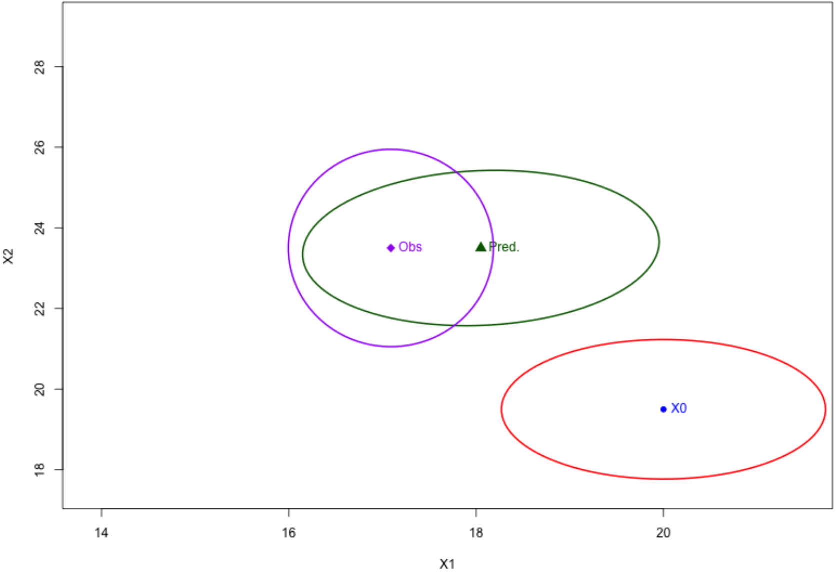
The Kalman filter

What is the Kalman filter?



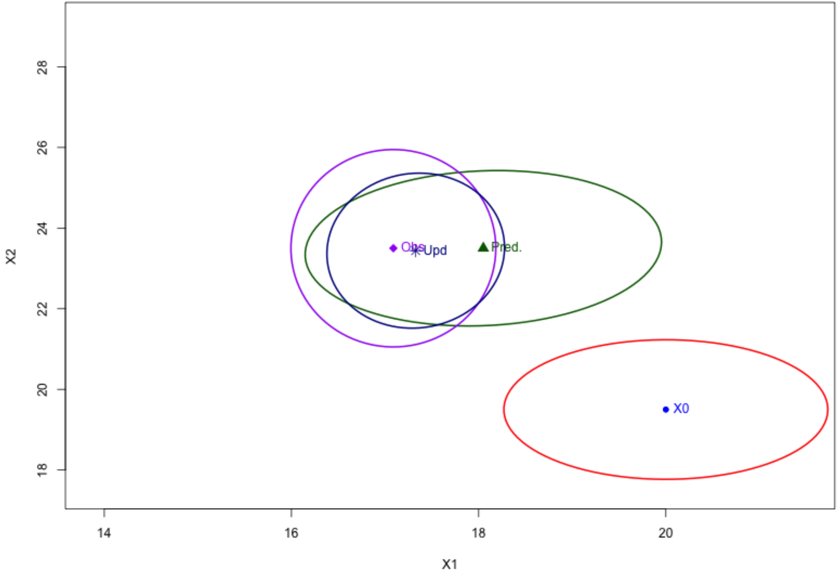
The Kalman filter

What is the Kalman filter?



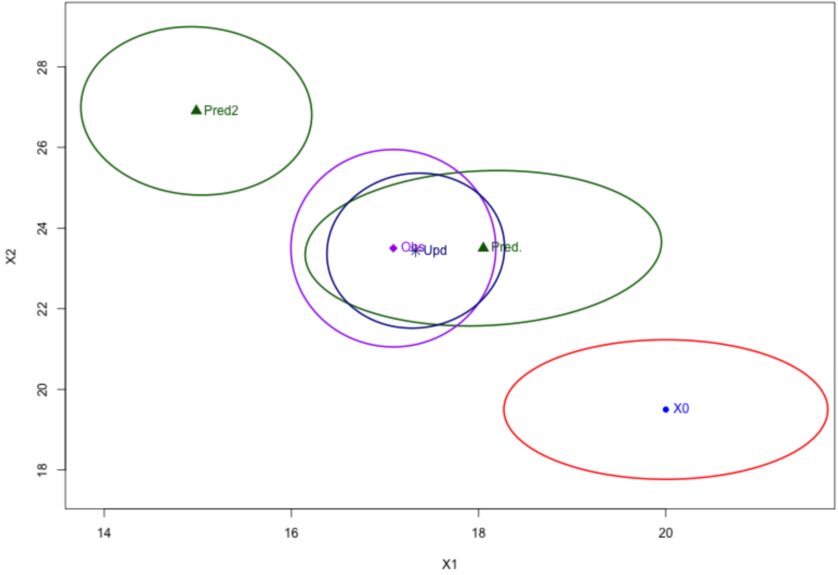
The Kalman filter

What is the Kalman filter?



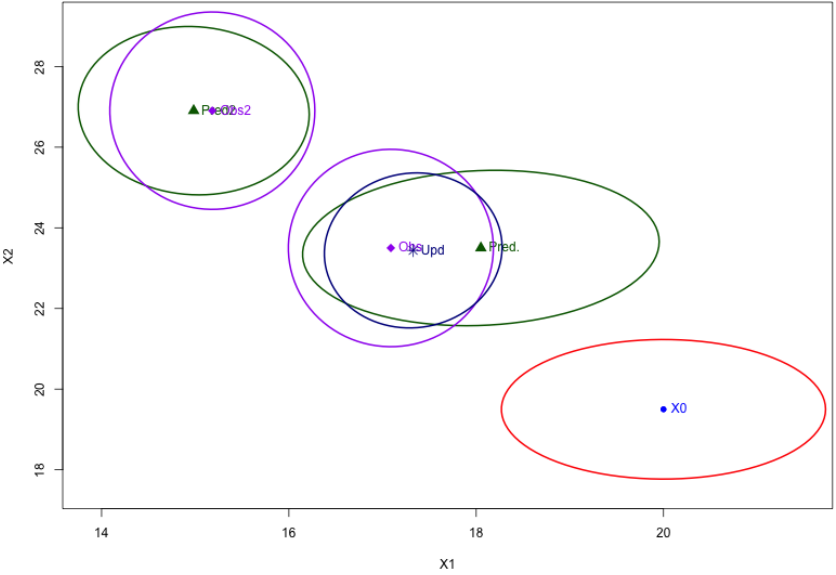
The Kalman filter

What is the Kalman filter?



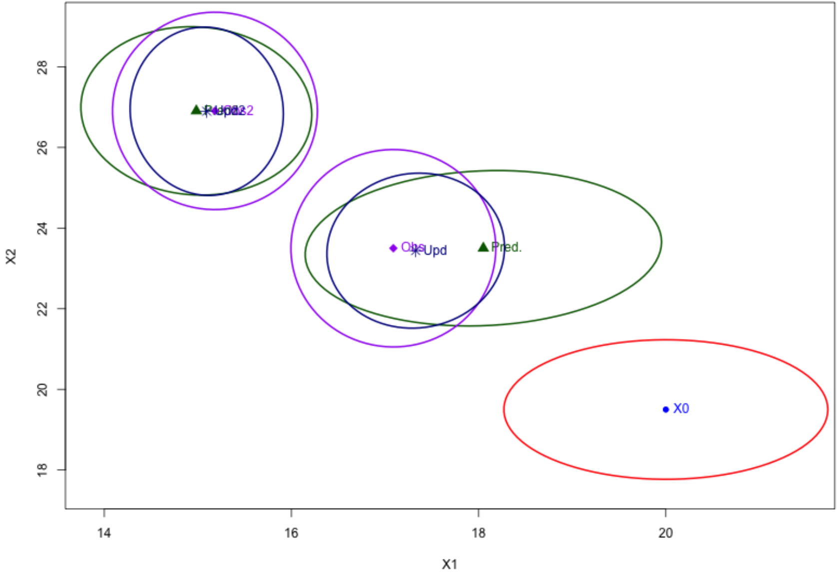
The Kalman filter

What is the Kalman filter?



The Kalman filter

What is the Kalman filter?



The Kalman filter

- The Kalman filter works (in theory) for models on the given form:

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

The Kalman filter

- The Kalman filter works (in theory) for models on the given form:

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

- Basic assumptions: everything is linear and Gaussian.

The Kalman filter

- The Kalman filter works (in theory) for models on the given form:

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

- Basic assumptions: everything is linear and Gaussian.
- If noise is non-Gaussian or if the model is non-linear, the prediction/ observation densities are not Gaussian (we shall work briefly with this in the next assignment).

The Kalman filter

- The Kalman filter works (in theory) for models on the given form:

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

- Basic assumptions: everything is linear and Gaussian.
- If noise is non-Gaussian or if the model is non-linear, the prediction/ observation densities are not Gaussian (we shall work briefly with this in the next assignment).
- If that is the case, we can use what is called *the extended* Kalman filter (enroll to Advanced Time Series Analysis (02427) to learn about this).

The Foundation of the Kalman filter

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

The Foundation of the Kalman filter

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

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

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

The Foundation of the Kalman filter II

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

The Foundation of the Kalman filter II

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

Using definitions from previous slide the update equations are:

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

$$
\boldsymbol{K}_t = \boldsymbol{\Sigma}^{xy}_{t|t-1} \left(\boldsymbol{\Sigma}^{yy}_{t|t-1} \right)^{-1}
$$

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

The Kalman filter

Initialization:

$$\widehat{\boldsymbol{X}}_{1|0} = E\left[\boldsymbol{X}_1\right] = \boldsymbol{\mu}_0$$

$$\boldsymbol{\Sigma}^{xx}_{1|0} = V\left[\boldsymbol{X}_1\right] = \boldsymbol{V}_0$$

$$\boldsymbol{\Sigma}^{yy}_{1|0} = \boldsymbol{C}\boldsymbol{\Sigma}^{xx}_{1|0}\boldsymbol{C}^T + \boldsymbol{\Sigma}_2$$

The Kalman filter

Initialization:

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

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

Reconstruction:

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

The Kalman filter

Initialization:

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

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)
```

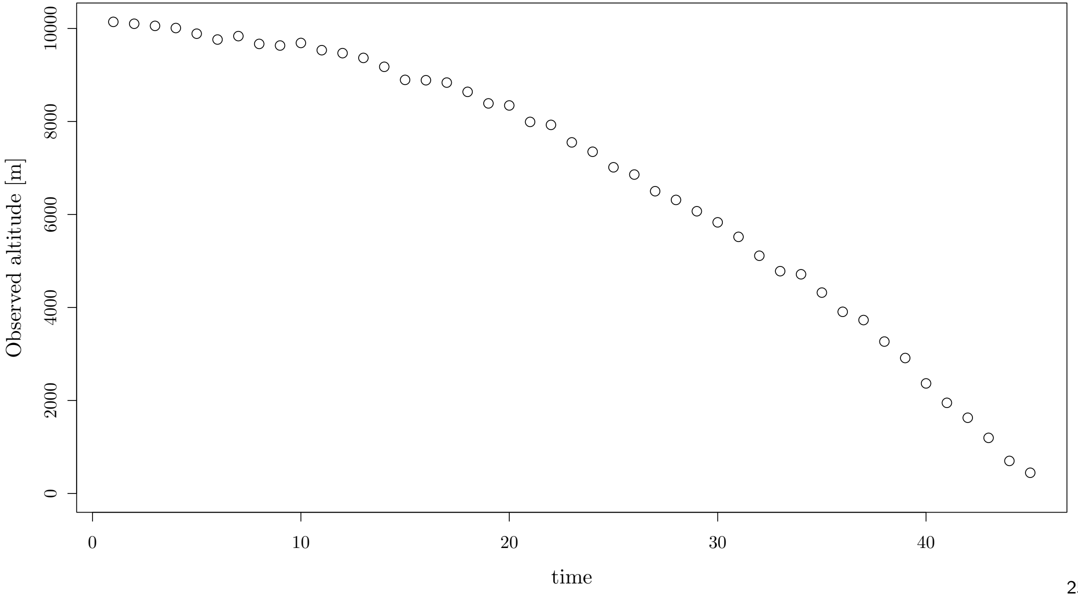
Simulation of a falling body - simulation

```
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]
```

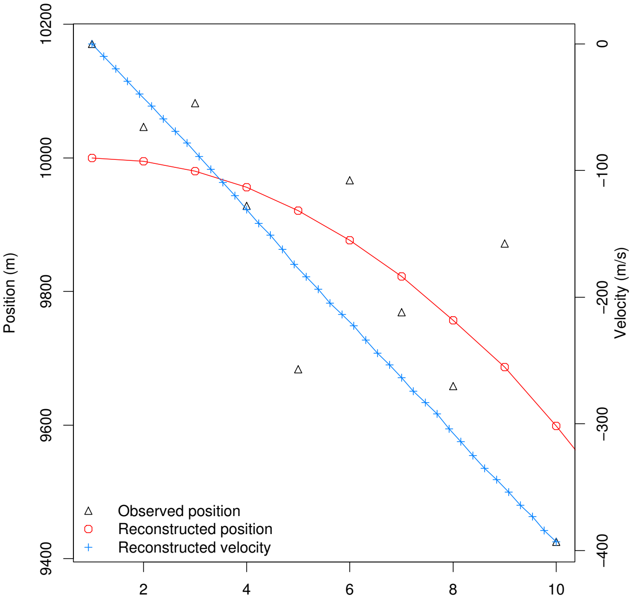
- Remember that if $Z \sim N(0, I)$, then $Y = QZ \sim N(0, QQ^T)$.
- The Cholesky factorization is one way to solve $QQ^T = \Sigma$ for Q.

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