Time Series Analysis

Henrik Madsen

hm@imm.dtu.dk

Informatics and Mathematical Modelling
Technical University of Denmark
DK-2800 Kgs. Lyngby

Outline of the lecture

State space models, 1st part:

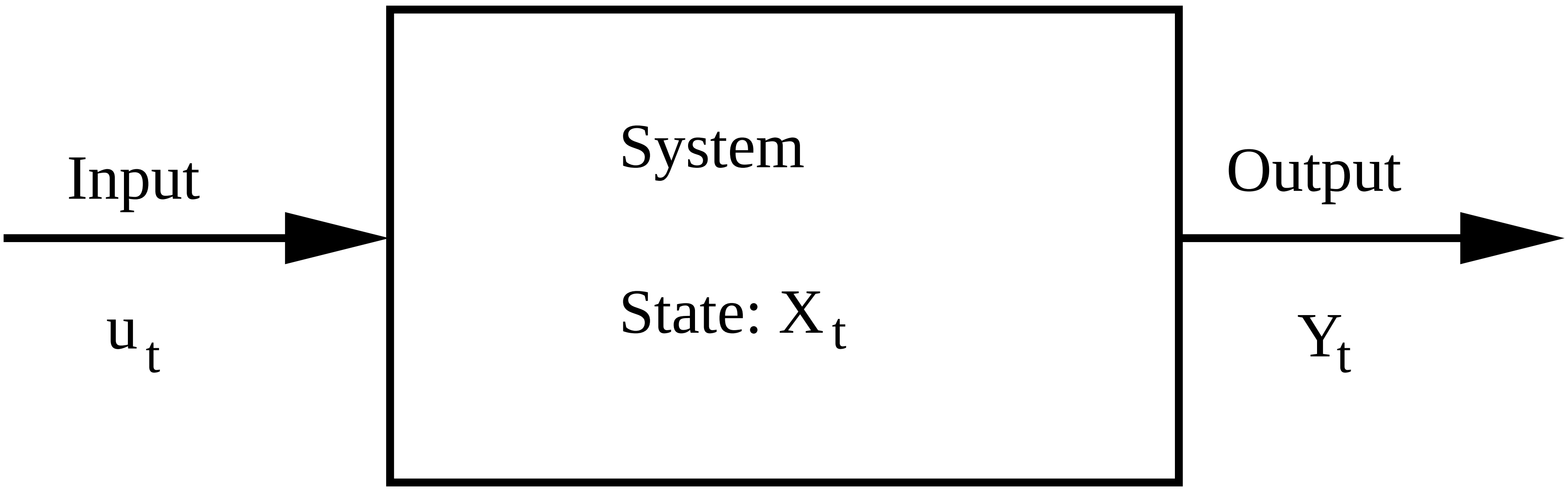
- Model: Sec. 10.1
- The Kalman filter: Sec. 10.3

Cursory material:

- Sec. 10.3.2 (Empirical-Bayesian description)

State space models

- System model; A full description of the dynamical system (i.e. including the parameters)
- Observations; Noisy measurements on some parts (states) of the system
- Goal; reconstruct and predict the state of the system



State space models; examples

- 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 current position
- A model of a car engine: Input; fuel. State; Fuel and temperature in various parts. Observarions: Sensor output
- PK/PD-modeling: State: Amount of drug in blood, liver, muscules, . . . Observations: Amount in blood (with noise), Input: Drug.

Determining the model structure

- The system model is often based on physical considerations; this often leads to dynamical models consisting of differential equations
- An m'th order differential equation can be formulated as m 1st order differential equations
- Sampling such a system leads to a linear state space model and there exist a way of coming from the coefficients in continuous time to the coefficients in discrete time

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

Example – a falling body

- Height above ground: $z(t)$
- Initial conditions: Position $z(t_0)$ and velocity $z'(t_0)$
- Physical considerations: $\dfrac{d^2z}{dt^2} = -g$
- States: Position $x_1(t) = z(t)$ and velocity $x_2(t) = z'(t)$
- Only the position is measured $y(t) = x_1(t)$
- Continuous time description $\boldsymbol{x}(t) = [x_1(t)\ x_2(t)]^T$:

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

Example – a falling body (cont'nd)

- Solving the equations:

$$
\begin{aligned}
x_2(t) &= -g(t-t_0) + x_2(t_0) \\
x_1(t) &= -\frac{g}{2}(t-t_0)^2 + (t-t_0)x_2(t_0) + x_1(t_0)
\end{aligned}
$$

- Sampling: $t = sT$, $t_0 = (s-1)T$, and $T = 1$

$$
\begin{aligned}
\boldsymbol{x}_s &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_{s-1} + \begin{bmatrix} -1/2 \\ -1 \end{bmatrix} g \\
\boldsymbol{y}_s &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}_s
\end{aligned}
$$

- Adding disturbances and measurement noise:

$$
\begin{aligned}
\boldsymbol{x}_s &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_{s-1} + \begin{bmatrix} -1/2 \\ -1 \end{bmatrix} g + \boldsymbol{e}_{1,s} \\
\boldsymbol{y}_s &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}_s + e_{2,s}
\end{aligned}
$$

Example – a falling body (cont'nd)

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

- **Predict** the future position and velocity $\boldsymbol{x}_{s+k|s}$ ($k > 0$)
- **Reconstruct** the current position and velocity from noisy measurements $\boldsymbol{x}_{s|s}$
- **Interpolate** to find the best estimate of the position and velocity at a previous time point $\boldsymbol{x}_{s+k|s}$ ($k < 0$) (estimate the path in the state space; vary k so that $s + k$ varied from 1 to s)

we will focus on reconstruction and prediction

Requirement

In order to predict, reconstruct or interpolate the m-dimensional state in the system

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

the system must be observable, i.e.

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

For the falling body (S-PLUS):

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

Where A and C is taken from the discrete time description of the system.

The Kalman filter

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

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

Reconstruction:

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

Prediction:

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

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 $\widehat{\boldsymbol{X}}_{t+k|t}$ and $\boldsymbol{\Sigma}_{t+k|t}$ are calculated as part of the Kalman filter

$$\begin{aligned}\widehat{\boldsymbol{X}}_{t+k+1|t} &= \boldsymbol{A}\widehat{\boldsymbol{X}}_{t+k|t} + \boldsymbol{B}\boldsymbol{u}_{t+k} \\ \boldsymbol{\Sigma}^{xx}_{t+k+1|t} &= \boldsymbol{A}\boldsymbol{\Sigma}^{xx}_{t+k|t}\boldsymbol{A}^T + \boldsymbol{\Sigma}_1\end{aligned}$$

- The future input must be *decided*

Naming and history

- The filter is named after Rudolf E. Kalman, though Thorvald Nicolai Thiele and Peter Swerling actually developed a similar algorithm earlier.
- It was during a visit of Kalman to the NASA Ames Research Center that he saw the applicability of his ideas to the problem of trajectory estimation for the Apollo program, leading to its incorporation in the Apollo navigation computer.

From `http://en.wikipedia.org/wiki/Kalman_filter`

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

$$E[(\boldsymbol{X}_t|\mathcal{Y}_{t-1}) \,|\, (\boldsymbol{Y}_t|\mathcal{Y}_{t-1})] = E[\boldsymbol{X}_t|\boldsymbol{Y}_t, \mathcal{Y}_{t-1}] =$$

$$E[\boldsymbol{X}_t|\mathcal{Y}_{t-1}] + C[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]V^{-1}[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}](\boldsymbol{Y}_t - E[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}])$$

$$V[(\boldsymbol{X}_t|\mathcal{Y}_{t-1}) \,|\, (\boldsymbol{Y}_t|\mathcal{Y}_{t-1})] = V[\boldsymbol{X}_t|\boldsymbol{Y}_t, \mathcal{Y}_{t-1}] =$$

$$V[\boldsymbol{X}_t|\mathcal{Y}_{t-1}] - C[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]V^{-1}[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}]C^T[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]$$

The Foundation of the Kalman filter (cont'nd)

$$E[\boldsymbol{X}_t|\boldsymbol{Y}_t, \mathcal{Y}_{t-1}] =$$
$$E[\boldsymbol{X}_t|\mathcal{Y}_{t-1}] + C[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]V^{-1}[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}](\boldsymbol{Y}_t - E[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}])$$
$$V[\boldsymbol{X}_t|\boldsymbol{Y}_t, \mathcal{Y}_{t-1}] =$$
$$V[\boldsymbol{X}_t|\mathcal{Y}_{t-1}] - C[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]V^{-1}[\boldsymbol{Y}_t|\mathcal{Y}_{t-1}]C^T[\boldsymbol{X}_t, \boldsymbol{Y}_t|\mathcal{Y}_{t-1}]$$

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

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

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

The Foundation of the Kalman filter (cont'nd)

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

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

Which results in the prediction errors:

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

And in therefore:

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

$\boldsymbol{\Sigma}^{xy}_{t+1|t}$ can also be calculated

Kalman filter applied to a falling body

Description of the system:

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

$$\Sigma_1 = \begin{bmatrix} 2.0 & 0.8 \\ 0.8 & 1.0 \end{bmatrix} \qquad \Sigma_2 = \begin{bmatrix} 10000 \end{bmatrix}$$

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

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

Kalman filter applied to a falling body (cont'nd)

1st observation ($t = 1$): $y_1 = 10171$

Reconstruction: $K_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$

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

Prediction:

$$\widehat{\boldsymbol{X}}_{2|1} = \begin{bmatrix} 9995.09 \\ -9.82 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{2|1} = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 1 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{yy}_{2|1} = \begin{bmatrix} 10002 \end{bmatrix}$$

Kalman filter applied to a falling body (cont'nd)

2nd observation ($t = 2$): $y_2 = 10046$

Reconstruction: $K_2 = \begin{bmatrix} 0.00020 & 0.00008 \end{bmatrix}^T$

$$\widehat{\boldsymbol{X}}_{2|2} = \begin{bmatrix} 9995.1 \\ -9.81 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{2|2} = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

Prediction:

$$\widehat{\boldsymbol{X}}_{3|2} = \begin{bmatrix} 9980.38 \\ -19.63 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{3|2} = \begin{bmatrix} 6.6 & 2.6 \\ 2.6 & 2 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{yy}_{3|2} = \begin{bmatrix} 10006.6 \end{bmatrix}$$

Kalman filter applied to a falling body (cont'nd)

3rd observation ($t = 3$): $y_3 = 10082$

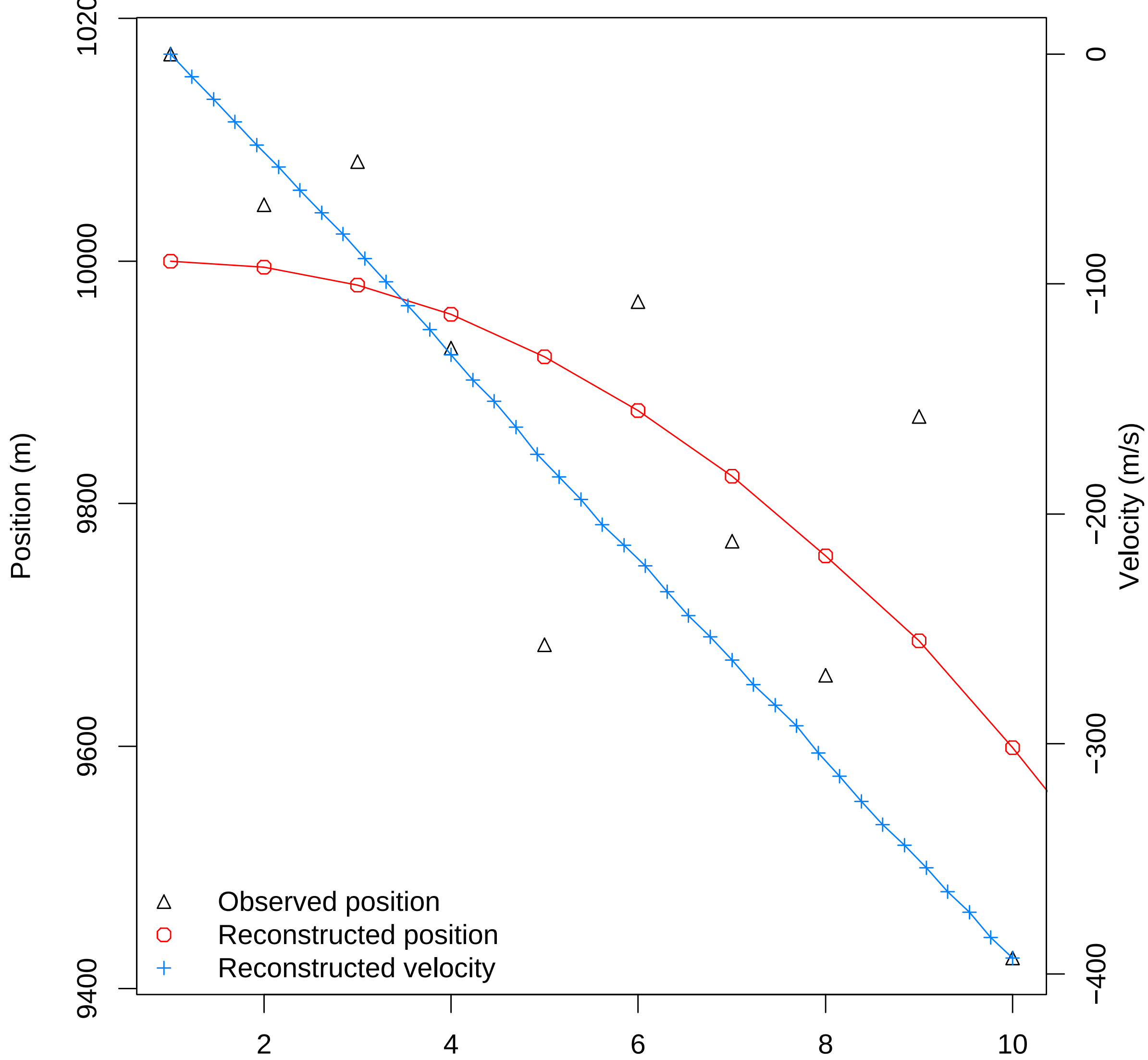
Reconstruction: $K_3 = \left[\begin{array}{cc} 0.00066 & 0.00026 \end{array}\right]^T$

$$\widehat{\boldsymbol{X}}_{3|3} = \begin{bmatrix} 9980.45 \\ -19.6 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{3|3} = \begin{bmatrix} 6.59 & 2.6 \\ 2.6 & 2 \end{bmatrix}$$

Prediction:

$$\widehat{\boldsymbol{X}}_{4|3} = \begin{bmatrix} 9955.94 \\ -29.41 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{xx}_{4|3} = \begin{bmatrix} 15.79 & 5.4 \\ 5.4 & 3 \end{bmatrix} \qquad \boldsymbol{\Sigma}^{yy}_{4|3} = \left[\begin{array}{c} 10015.79 \end{array}\right]$$

Falling body – the 10 first time points



Falling body – wrong initial state