

Lecture 11

The time-varying Kalman filter

Giancarlo Ferrari Trecate¹

¹Dependable Control and Decision Group
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
giancarlo.ferraritrecate@epfl.ch

Recap from the last lecture

Proposition (Linear combination of Gaussian RVs)

If $x = [x_1, \dots, x_n]^T \sim N(\mu_x, C_x)$ and

$$y = Ax + b, \quad b \in \mathbb{R}^m, \quad A \in \mathbb{R}^{m \times n}$$

(a) $y \in \mathbb{R}^m$ is Gaussian with

$$E[y] = A\mu_x + b, \quad \text{Var}[y] = AC_xA^T$$

(b) $z = \begin{bmatrix} x \\ y \end{bmatrix}$ is Gaussian with

$$E[z] = \begin{bmatrix} \mu_x \\ A\mu_x + b \end{bmatrix} \quad \text{Var}[z] = \begin{bmatrix} C_x & C_xA^T \\ AC_x & AC_xA^T \end{bmatrix}$$

Recap from the last lecture

Proposition (Conditioning for Gaussian RVs)

Let $X \in \mathbb{R}^n$, $Y \in \mathbb{R}^m$ and

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \begin{bmatrix} C_{XX} & C_{XY} \\ C_{YX} & C_{YY} \end{bmatrix} \right)$$

Then, $X|Y$ is Gaussian with

$$\begin{aligned} E[X|Y] &= \overbrace{\mu_X + C_{XY} C_{YY}^{-1} (y - \mu_Y)}^{(a)} \quad \text{"a posteriori" mean} \\ \text{Var}[X|Y] &= \underbrace{C_{XX} - C_{XY} C_{YY}^{-1} C_{YX}}_{(b)} \quad \text{"a posteriori" variance} \end{aligned}$$

(a): Shift in the mean

(b): "reduction" of the original uncertainty C_{XX}

State estimation: linear Gaussian setting

$$x_{k+1} = Ax_k + w_k$$

$$y_k = Cx_k + v_k$$

$$x_0 \sim N(\bar{x}_0, \Sigma_0)$$

Assumptions

- (a) x_0, w_i, v_j independent $\forall i, j$
- (b) $w_k \sim WGN(0, W) \quad W \geq 0$
- (c) $v_k \sim WGN(0, V) \quad V > 0$

State estimation: conditioning on the measured outputs

Let $Y_k = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix}$

- $\begin{bmatrix} x_k \\ Y_k \end{bmatrix}$ and $\begin{bmatrix} x_{k+1} \\ Y_k \end{bmatrix}$ are Gaussian
- Let $x_{k|k} = x_k | Y_k$ and define

$$\hat{x}_{k|k} = E[x_k | Y_k]$$

$$\Sigma_{k|k} = \text{Var}[x_k | Y_k]$$

Similar notation for $x_{k+1|k} = x_{k+1} | Y_k$

$\hat{x}_{k|k}$: filtered state

$\hat{x}_{k+1|k}$: state prediction

Problem: How to compute them along with $\Sigma_{k|k}, \Sigma_{k+1|k}$?

Naive method

$$\begin{bmatrix} x_k \\ Y_k \end{bmatrix} \sim N \left(\begin{bmatrix} \bar{x}_k \\ \bar{Y}_k \end{bmatrix}, \begin{bmatrix} \Sigma_{x_k x_k} & \Sigma_{x_k Y_k} \\ \Sigma_{Y_k x_k} & \Sigma_{Y_k Y_k} \end{bmatrix} \right)$$

Therefore

$$\hat{x}_{k|k} = \bar{x}_k + \Sigma_{x_k Y_k} \Sigma_{Y_k Y_k}^{-1} (Y_k - \bar{Y}_k)$$

(similar formula for $\hat{x}_{k+1|k}$)

Problem

The dimension of Y_k and $\Sigma_{Y_k Y_k}$ grows with time!

↪ impractical

Kalman Filtering (KF)

A recursive way for computing all desired quantities

Kalman filter and predictor

At $k = 0$, $x_0 \sim N(\bar{x}_0, \Sigma_0)$. Rename $\bar{x}_0 \rightarrow \hat{x}_{0|-1}$, $\Sigma_0 \rightarrow \Sigma_{0|-1}$

$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ is a linear transformation of a Gaussian random vector because

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} x_0 \\ v_0 \end{bmatrix}$$

then

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \sim N \left(\begin{bmatrix} \hat{x}_{0|-1} \\ C\hat{x}_{0|-1} \end{bmatrix}, \begin{bmatrix} \Sigma_{0|-1} & \Sigma_{0|-1}C^T \\ C\Sigma_{0|-1} & C\Sigma_{0|-1}C^T + V \end{bmatrix} \right) \quad (*)$$

1) Filtering step (measurement update)

$x_0|y_0$ is Gaussian with mean and variance

$$\hat{x}_{0|0} = \hat{x}_{0|-1} + \Sigma_{0|-1}C^T(C\Sigma_{0|-1}C^T + V)^{-1}(y_0 - C\hat{x}_{0|-1})$$

$$\Sigma_{0|0} = \Sigma_{0|-1} - \Sigma_{0|-1}C^T(C\Sigma_{0|-1}C^T + V)^{-1}C\Sigma_{0|-1}$$

Remark: the estimate $\hat{x}_{0|0}$ is based on the new measurement y_0

2) Prediction step (time update)

· $x_1|y_0$ is given by $A(x_0|y_0) + w_0$

· $y_1|y_0$ is given by $C(x_1|y_0) + v_1$

(■)

$$\hookrightarrow x_1|y_0 \sim N(\underbrace{A\hat{x}_{0|0}}_{\hat{x}_{1|0}}, \underbrace{A\Sigma_{0|0}A^T + W}_{\Sigma_{1|0}})$$

Then, using (■)

$$y_1|y_0 \sim N(C\hat{x}_{1|0}, C\Sigma_{1|0}C^T + V)$$

Moreover, $x_1|y_0$ and $y_1|y_0$ are linear combinations of Gaussian vectors $\rightarrow$ they are jointly Gaussian and, using the formulae for linear combinations of Gaussian vectors

$$\begin{bmatrix} x_1|y_0 \\ y_1|y_0 \end{bmatrix} \sim N \left(\begin{bmatrix} \hat{x}_{1|0} \\ C\hat{x}_{1|0} \end{bmatrix}, \begin{bmatrix} \Sigma_{1|0} & \Sigma_{1|0}C^T \\ C\Sigma_{1|0} & C\Sigma_{1|0}C^T + V \end{bmatrix} \right) \quad (**)$$

Key observation enabling iterations

$(*)$ is identical to $(*)$ up to the replacements

$$\hat{x}_{0|-1} \rightarrow \hat{x}_{1|0}$$

$$\Sigma_{0|-1} \rightarrow \Sigma_{1|0}$$

Idea: iterate the procedure for $k = 1, 2, \dots$

Kalman predictor and filter

Init $\hat{x}_{0|-1} \in \mathbb{R}^n, \Sigma_{0|-1} \in \mathbb{R}^{n \times n}$ (statistics of x_0)

Filtering step:

$$\hat{x}_{k|k} \stackrel{\text{def}}{=} \hat{x}_{k|k-1} + \Sigma_{k|k-1} C^T (C \Sigma_{k|k-1} C^T + V)^{-1} (y_k - C \hat{x}_{k|k-1})$$

$$\Sigma_{k|k} \stackrel{\text{def}}{=} \Sigma_{k|k-1} - \Sigma_{k|k-1} C^T (C \Sigma_{k|k-1} C^T + V)^{-1} C \Sigma_{k|k-1}$$

Prediction step:

$$\hat{x}_{k+1|k} = A \hat{x}_{k|k}$$

$$\Sigma_{k+1|k} = A \Sigma_{k|k} A^T + W$$

Remarks

- Sometimes, in the literature, “Kalman filter” denotes the predictor...
- Does $\hat{x}_{k|k}$ coincide with the quantity $x_{k|k} = E[x_k | Y_k]$ defined before?

Lemma (conditioning on the whole past)

$$\begin{aligned}x_k | Y_k &\sim N(\hat{x}_{k|k}, \Sigma_{k|k}) \\ x_{k+1} | Y_k &\sim N(\hat{x}_{k+1|k}, \Sigma_{k+1|k})\end{aligned}$$

(■)

Remarks

- Lemma $\Rightarrow$ Optimality of the predictor and the filter in a statistical sense
- $\Sigma_{k|k}$ and $\Sigma_{k+1|k}$ can be computed at all $k \geq 0$ before having any measurement
 $\hookrightarrow$ Update = Difference Riccati Equation (DRE) for filtering
- Observer form: from the algorithm we obtain

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + \underbrace{A\Sigma_{k|k-1}C^T(C\Sigma_{k|k-1}C^T + V)^{-1}}_{L_k}(y_k - \underbrace{C\hat{x}_{k|k-1}}_{\text{estimated output}})$$

L_k is the (time-varying) Kalman gain

Generalisation to systems with inputs

System:

$$x_{k+1} = Ax_k + Bu_k + w_k$$

$$y_k = Cx_k + v_k$$

$$x_0 \sim N(\bar{x}_0, \Sigma_0)$$

- Add the effect of the input in the prediction step, which becomes

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k$$

$\hookrightarrow Bu_k$ just changes the mean of $x_{k+1|k}$, not the variance

KF as a minimum variance estimator

System:

$$x_{k+1} = Ax_k + Bu_k + w_k$$

$$y_k = Cx_k + v_k$$

$$x_0 \sim N(\bar{x}_0, \Sigma_0)$$

Assume a linear predictor dynamics

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + \underbrace{L_k}_{\substack{\downarrow \\ \text{How to tune } L_k?}} [y_k - C\hat{x}_{k|k-1}]$$

$$\hat{x}_0 = \bar{x}_0$$

Dynamics of the error $e_{k|k-1} \stackrel{\text{def}}{=} x_k - \hat{x}_{k|k-1}$:

$$e_{k+1|k} = \underbrace{(A - L_k C)}_{A_{c,k}} e_{k|k-1} + \underbrace{w_k - L_k v_k}_{\text{new, compared to Luenberger}}$$

$$= A_{c,k} e_{k|k-1} + B_{c,k} \xi_k$$

$$B_{c,k} = \begin{bmatrix} I & -L_k \end{bmatrix} \quad \xi_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix}$$

- $e_{k|k-1}$ is a Gaussian random vector (linear combination of $e_{0|-1} = x_0 - \bar{x}_0$ and $v_0, \dots, v_{k-1}, w_0, \dots, w_{k-1}$ which are jointly Gaussian)

Statistics of $e_{k|k-1}$

- $E[e_{k+1|k}] = A_{c,k} E[e_{k|k-1}]$
Setting $\hat{x}_0 = \hat{x}_{0|-1} = \bar{x}_0$, one has $E[e_{0|-1}] = 0$ and $E[e_{k|k-1}] = 0, \forall k \geq 0$
- Definitions
 - ▶ Variance of the error: $E[\|e_{k|k-1}\|^2] = E[e_{k|k-1}^T e_{k|k-1}] \in \mathbb{R}$
 - ▶ Covariance of the error: $E[e_{k|k-1} e_{k|k-1}^T] \in \mathbb{R}^{n \times n}$
- Abuse of terminology: so far $E[e_{k|k-1} e_{k|k-1}^T]$ was termed “variance”

Problem (variance minimization)

Compute L_k so as to solve

$$\min_{L_k} E[\|e_{k|k-1}\|^2]$$

Theorem

The matrix L_k minimizing $E \left[\|e_{k|k-1}\|^2 \right]$ is

$$L_k = A \Sigma_{k|k-1} C^T [C \Sigma_{k|k-1} C^T + V]^{-1}$$

where $\Sigma_{k|k-1} = E \left[e_{k|k-1} e_{k|k-1}^T \right]$ is given by the DRE

$$\Sigma_{k+1|k} = A \Sigma_{k|k-1} A^T + W - A \Sigma_{k|k-1} C^T [C \Sigma_{k|k-1} C^T + V]^{-1} C \Sigma_{k|k-1} A^T$$

with initial condition

$$\Sigma_{0|-1} = \Sigma_0 \quad (\blacksquare)$$

Proof by direct minimization (omitted)

Remarks

- Same formulae for $\Sigma_{k|k-1}$ and L_k obtained before
- One can show that the filtered error $e_{k|k} = x_k - \hat{x}_{k|k}$ has zero mean and variance $\text{Var}[e_{k|k}] = \Sigma_{k|k}$ computed as in KF
- Riccati equation for the FH-LQ regulator:

$$P_k = Q + A^T P_{k+1} A - A^T P_{k+1} B [B^T P_{k+1} B + R]^{-1} B^T P_{k+1} A$$

... for the Kalman predictor:

$$\Sigma_{k+1|k} = W + A \Sigma_{k|k-1} A^T - A \Sigma_{k|k-1} C^T [C \Sigma_{k|k-1} C^T + V]^{-1} C \Sigma_{k|k-1} A^T$$

They are related through the substitutions:

FH-LQ	Kalman Predictor
k	$-k$ (reversed!)
A	A^T
B	C^T
Q	W
R	V
P_k	$\Sigma_{k k-1}$

Called “duality relations”!

Generalization

Straightforward extension to:

- linear time-varying systems
 $\hookrightarrow$ Substitute $A \rightarrow A_k$, $C \rightarrow C_k$, $B \rightarrow B_k$ in the formulae for the Kalman predictor and filter
- noise with time-varying variance
 $w_k \sim N(0, W_k)$, $W_k \geq 0$
 $v_k \sim N(0, V_k)$, $V_k > 0$

Innovation sequence

Consider the KF update

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \Sigma_{k|k-1} C^T \left(C \Sigma_{k|k-1} C^T + V \right)^{-1} (y_k - C \hat{x}_{k|k-1})$$

Definition

The innovation is $\nu_k = y_k - C \hat{x}_{k|k-1}$

Remarks

ν_k quantifies the additional “information content” brought by the measurement y_k , compared to the prediction $C \hat{x}_{k|k-1}$ which is based on measurements $y_{k-1}, y_{k-2}, \dots$

Statistics of the innovation

- ν_k are jointly Gaussian random vectors
- Zero mean, conditioned to the past measurements

$$E[\nu_{k+1}|Y_k] = E[y_{k+1} - \underbrace{C\hat{x}_{k+1|k}}_{\text{deterministic}} | Y_k] = E[y_{k+1}|Y_k] - C\hat{x}_{k+1|k} = 0$$

- Variance given by

$$\begin{aligned} S_{k+1} &= E[\nu_{k+1} \nu_{k+1}^T] = E[(y_{k+1} - C\hat{x}_{k+1|k})(y_{k+1} - C\hat{x}_{k+1|k})^T] \\ &= V + C\Sigma_{k+1|k}C^T \end{aligned}$$

- ν_k is NOT correlated with ν_j , $j \neq k$ (proof omitted), i.e. ν_k is white Gaussian noise
 - ▶ This can be used for checking if the KF works well and for tuning KF parameters (see later)
 - ▶ Note that y_k and y_j , $k \neq j$, are correlated
- ν_k is also uncorrelated with past measurements y_j , $j < k$ (proof omitted)
 - ▶ expected, if ν_k captures all and only the additional information brought by y_k

Example 1: estimation of position and velocity of a ground vehicle¹

Motivation: develop a navigation and tracking system for vehicle

Problem: How can the vehicle know its position and velocity?

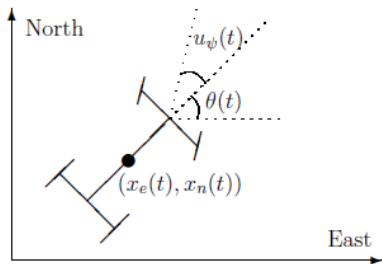
- **Solution 1:** use encoders attached to the vehicle wheels
 $\hookrightarrow$ imprecise $\rightarrow$ the estimated and real positions can be far away from the real ones
- **Solution 2:** use GPS measurements
 $\hookrightarrow$ can give much better results, but measurements are noisy and velocities are not measured

Idea: Use GPS with a KF!

¹ mathworks.com/help/control/getstart/estimating-states-of-time-varying-systems-using-kalman-filters.html

Example 1 (cont.)

The vehicle moves in 2-D space: position and velocity in the north and east directions.



- For simulation: non-linear non-holonomic model where the input is steering angle of the vehicle (u_ψ)
- For the KF: discrete-time linear model describing the evolution of the position and velocity over time in response to model initial conditions as well as position measurements obtained from GPS

Example 1 (cont.)

Variables of interest:

- $\hat{x}_1(k)$: east position estimate
- $\hat{x}_2(k)$: north position estimate
- $\hat{v}_1(k)$: east velocity estimate
- $\hat{v}_2(k)$: north velocity estimate

$$\hat{x}(k) = \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \\ \hat{v}_1(k) \\ \hat{v}_2(k) \end{bmatrix}$$

Linear model:

$$x(k+1) = Ax(k) + Gw(k)$$

$$y(k) = Cx(k) + v(k)$$

where x is the state vector, $y \in \mathbb{R}^2$ is the vector of measured positions, $w(k) = \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix}$ is the process noise and $v(k) = \begin{bmatrix} v_1(k) \\ v_2(k) \end{bmatrix}$ is the measurement noise

Example 1 (cont.)

$$x(k+1) = Ax(k) + Gw(k)$$

$$y(k) = Cx(k) + v(k)$$

$$A = \begin{bmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} \frac{T_s}{2} & 0 \\ 0 & \frac{T_s}{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Considerations:

- velocities are modelled as a random walk:

$$v_i(k+1) = v_i(k) + w_i(k), \quad i \in \{1, 2\}$$

- positions are represented by the following discretization of $\frac{d}{dt}x = v$:

$$\frac{x_i(k+1) - x_i(k)}{T_s} = \frac{v_i(k+1) + v_i(k)}{2}, \quad i \in \{1, 2\}$$

- $T_s = 1s$

Example 1 (cont.)

$$w \sim N(0, W_k) \text{ and } v \sim N(0, V)$$

- variance of the measurement noise: $V = \text{diag} \{50, 50\}$
- variance of the process noise: it should describe how much the vehicle velocity can change over one sampling interval $\rightarrow$ time-varying W

$$W_k = GQ(k)G^T \quad Q(k) = \begin{bmatrix} 1 + \frac{250}{f_{sat}(\hat{v}_1^2(k))} & 0 \\ 0 & 1 + \frac{250}{f_{sat}(\hat{v}_2^2(k))} \end{bmatrix}$$

where $f_{sat}(z) = \min(\max(z, 25), 625)$ (values obtained experimentally) and $\hat{v}_1, \hat{v}_2$ are the estimated velocities

$\hookrightarrow$ Captures the intuition that typical values of w are smaller when velocity is large.

$\hookrightarrow$ A diagonal W_k represents the naive assumption that the velocity changes in the north and east directions are uncorrelated

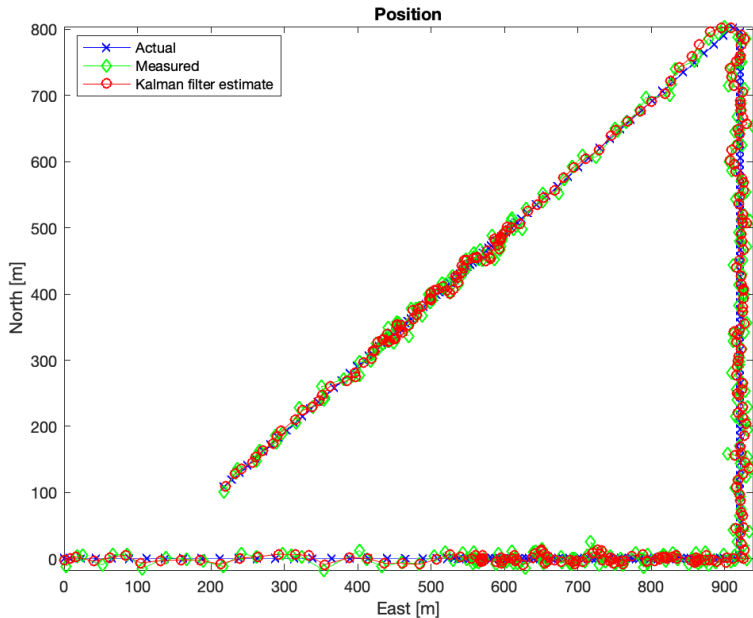
- Initial state: $x_0 \sim N(0, 10I)$

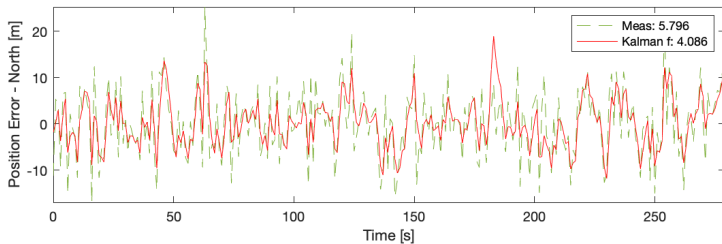
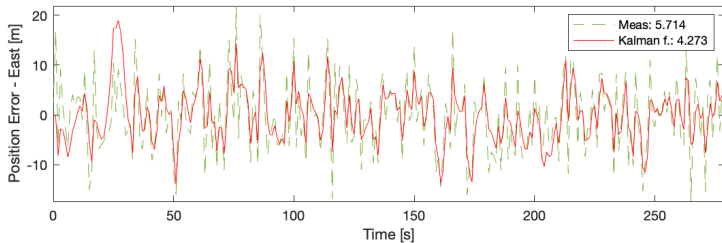
Example 1 (cont.)

Simulation scenario: the vehicle makes the following maneuvers:

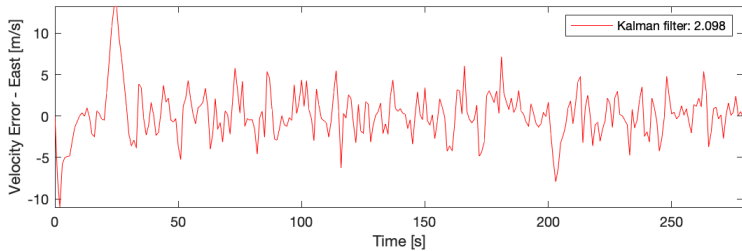
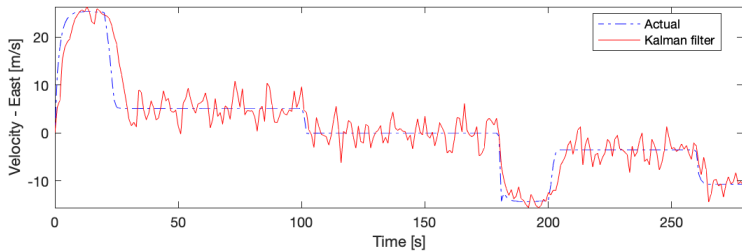
- 1 At $t = 0$ the vehicle is at $x_e(0) = 0$, $x_n(0) = 0$ and is stationary
- 2 Heading east, it accelerates to 25m/s. It decelerates to 5m/s at $t=20$ s.
- 3 At $t = 100$ s, it turns toward north and accelerates to 10m/s.
- 4 At $t = 180$ s, it accelerates to 20m/s with southwest direction.
- 5 At $t = 200$ s, it decelerates to 5m/s.
- 6 At $t = 260$ s, it accelerates to 15m/s.

Example 1 (cont.)

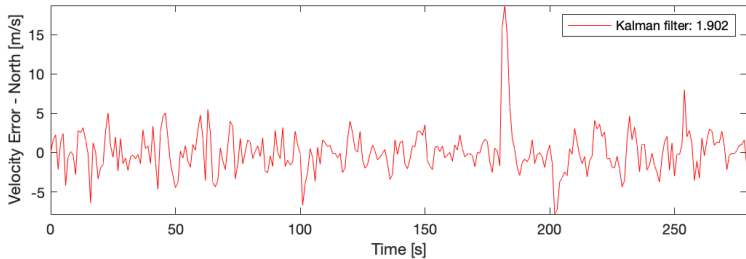
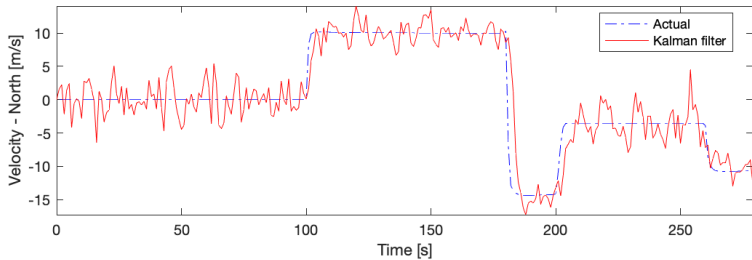




$$\text{Meas: } \frac{1}{N_s} \sum_{k=1}^{N_s} |x_i(k) - y_i(k)| \quad \text{Kalman f: } \frac{1}{N_s} \sum_{k=1}^{N_s} |x_i(k) - \hat{x}_i(k)| \quad i = \{1, 2\}$$



$$\text{Kalman filter: } \frac{1}{N_s} \sum_{k=1}^{N_s} |v_1(k) - \hat{v}_1(k)|$$



$$\text{Kalman filter: } \frac{1}{N_s} \sum_{k=1}^{N_s} |v_2(k) - \hat{v}_2(k)|$$

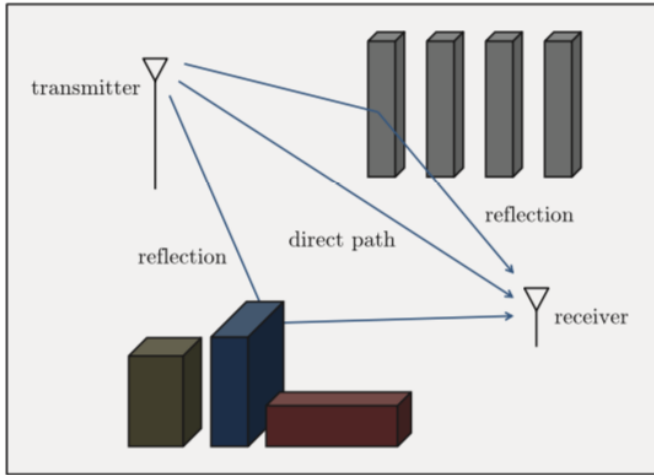
Comments:

- The Kalman filter gives better position estimates than the raw measurements
 - ▶ error reduction: $\sim 25\%$ on the east position and $\sim 30\%$ on the north position
 - The peaks in the velocity match with the sharp turn and sudden acceleration of the car, e.g.:
 - ▶ at $t = 20s$ and $t = 200s$ in the east velocity
 - ▶ at $t = 1800s$ and $t = 200s$ in the north velocity
- ↪ After a few time steps, the filter estimates catch up with the actual velocity.

Additional example (check at home!)

Example 2: Channel estimation in communication systems²

Wireless communications systems → signals from the transmitter may not reach the receiver directly due to scattering → **DELAYS!**



²N. Kovvali, M. Banavar, A. Spanias. *An Introduction to Kalman Filtering with MATLAB Examples*.

Example 2 (cont.)

- The received signal at time k is a superposition of:
 - ▶ scaled version of emitted signal at time k
 - ▶ scaled and shifted versions of emitted signals at time $k - 1, k - 2, \dots$
 - ▶ noise
- The propagation channel changes over time due to:
 - ▶ movements of the transmitter/receiver
 - ▶ changes of the environment

} Multipath propagation

A time-varying Finite Impulse Response (FIR) filter let us model the multipath channel:

→ Consider the sequence $\{c(k)\}$ as the sent signal and

$$y(k) = \sum_{d=1}^3 x_d(k) c(k - d + 1) + v(k)$$

where $y(k)$ is the received signal, $x_i(k)$ for $i = 1, 2, 3, \dots$ are the time varying FIR coefficients and $v(k)$ represents additive noise

Example 2 (cont.)

FIR coefficients can be estimated by transmitting and receiving a (known) test signal through the channel.

We can use KF to estimate the FIR coefficients of the channel!

- One possible model: “random walk”

$$x_i(k+1) = x_i(k) + w_i(k) \quad w(k) \sim N(0, W)$$

- Other option:

→ Take into account the correlation (α_i) between two sequent values

$$x_i(k+1) = \alpha_i x_i(k) + w_i(k) \quad w(k) \sim N(0, W)$$

- ▶ Model the channel as a time-varying FIR filter of order $D = 3$
- ▶ Let $x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$ be the channel (FIR) coefficients at time step k
- ▶ Consider α_1 , α_2 and α_3 as given constants (they depend on the sampling time and the frequency of the received signal)
- ▶ Consider iid noise $w(k) = \begin{bmatrix} w_1(k) \\ w_2(k) \\ w_3(k) \end{bmatrix} \sim N(0, W)$

Example 2 (cont.)

State-space model of the FIR coefficients:

$$x(k+1) = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} x(k) + w(k) \quad (*)$$

Measurements: obtained as the filtered output for a test signal $\{c(k)\}$ propagated through the channel (modelled as a FIR)

$$y(k) = \sum_{d=1}^3 x_d(k) c(k-d+1) + v(k)$$

where $v(k) \sim N(0, V)$

In matrix notation,

$$y(k) = \begin{bmatrix} c(k) & c(k-1) & c(k-2) \end{bmatrix} x(k) + v(k) \quad (**)$$

Remark: $(*)$ and $(**)$ provide a LTV system

Example 2 (cont.)

Simulations:

Set

- $\alpha_1 = 0.85$, $\alpha_2 = 1$ and $\alpha_3 = -0.95$
- $W = \text{diag}\{0.1, 0.1, 0.1\}$
- $V = 0.1$

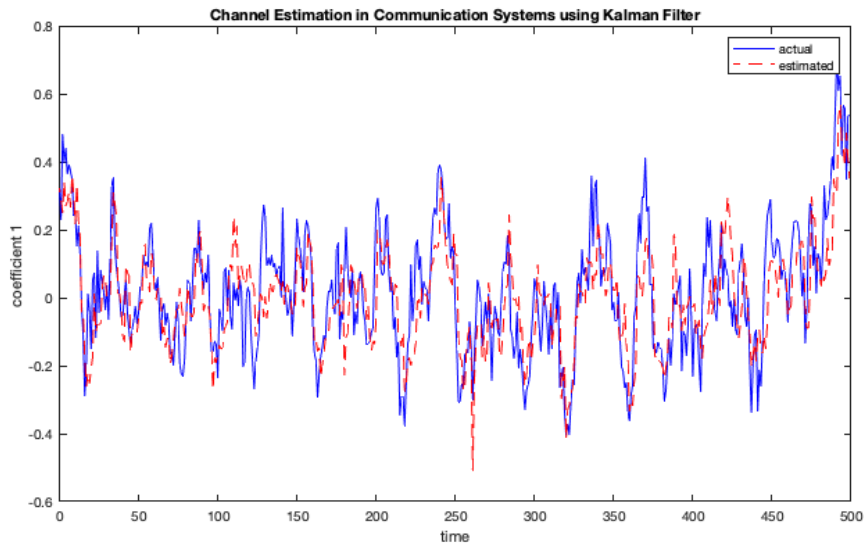
and generate a test signal sequence $c(k)$ (for example a random *but known* signal).

Also consider:

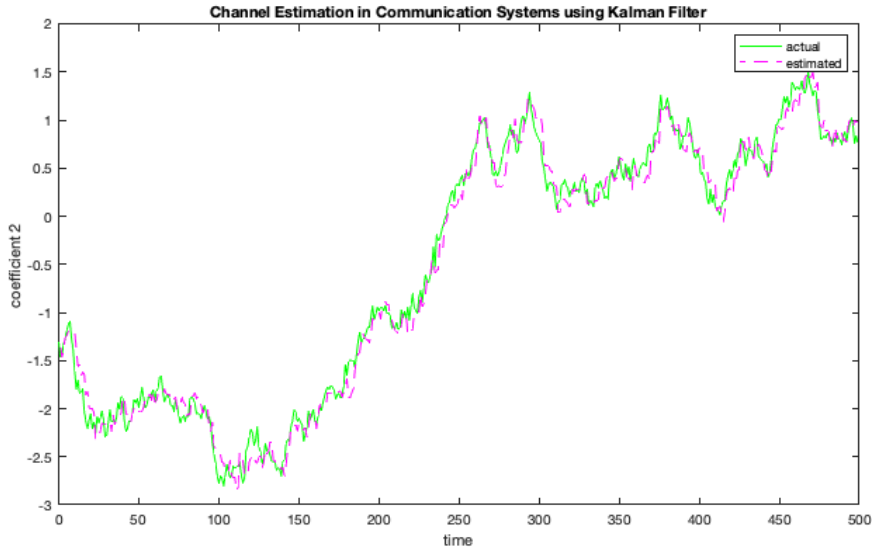
- initial channel coefficients $x(0) \sim N(0, I)$
- KF initial variance $\Sigma_{0|-1} = I$

→ After performing $N = 500$ time steps of Kalman filtering, the actual and estimated channel coefficients are shown in the plots.

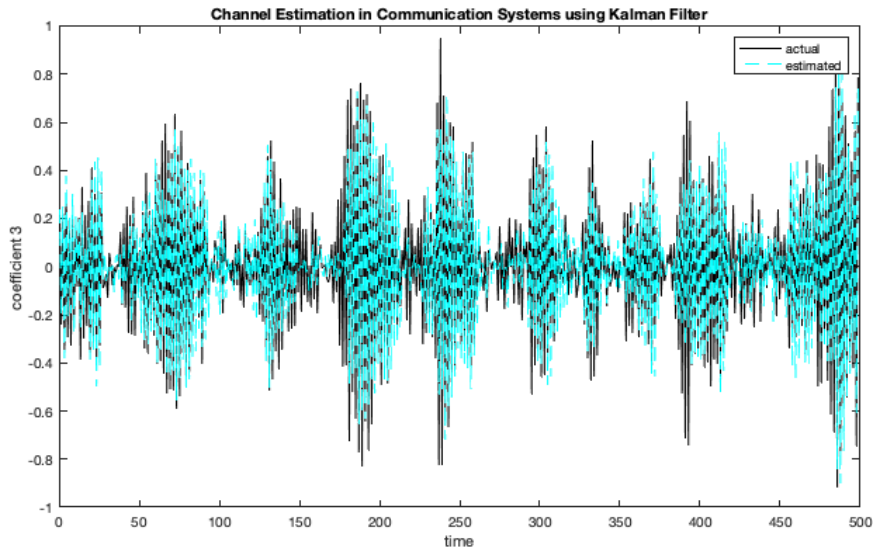
Example 2 (cont.)



Example 2 (cont.)



Example 2 (cont.)



Example 2 (cont.)

Remarks:

- The vertical scale of coefficient 2 is larger than the others
- The sign of the values of coefficient 3 usually changes since $\alpha_3 < 0$

The Kalman filter is able to estimate the time-varying channel coefficients with good accuracy.