Lecture 11

The time-varying Kalman filter

Giancarlo Ferrari Trecate[1]

[1]Dependable 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

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

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

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

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

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

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, $\begin{bmatrix} X \\ Y \end{bmatrix}$ is Gaussian with

$$E[X|Y] = \overbrace{\mu_X + C_{XY} C_{YY}^{-1}(y - \mu_Y)}^{(a)} \quad \text{“a posteriori” mean}$$

$$Var[X|Y] = \underbrace{C_{XX} - C_{XY} C_{YY}^{-1} C_{YX}}_{(b)} \quad \text{“a posteriori” variance}$$

(a): Shift in the mean

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

State estimation: linear Gaussian setting

$$x_{k+1} = Ax_k + w_k \qquad x_k \in \mathbb{R}^n \quad w_k \in \mathbb{R}^n$$
$$y_k = Cx_k + v_k \qquad y_k \in \mathbb{R}^p \quad v_k \in \mathbb{R}^p$$
$$x_0 \sim N(\bar{x}_0, \Sigma_0)$$

Assumptions

(a) x_0, w_i , v_j independent $\forall i, j$

(b) $w_k \sim WGN(0, W) \quad W \geq 0$

(c) $v_k \sim WGN(0, V) \quad V > 0$

Example

$$\xi = \begin{bmatrix} x_0 \\ w_1 \\ v_4 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{x}_0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_0 & 0 & 0 \\ 0 & W & 0 \\ 0 & 0 & V \end{bmatrix} \right)$$

State estimation: conditioning on the measured outputs

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

$x_1 = Ax_0 + w_0$

$y_0 = Cx_0 + v_0$

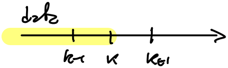
$x_2 = A^2 x_0 + A w_0 + w_1$

$\dots$

- $\begin{bmatrix} x_k \\ Y_k \end{bmatrix}$ and $\begin{bmatrix} x_{k+1} \\ Y_k \end{bmatrix}$ are Gaussian $\rightarrow$ Because they are linear combinations of $x_0, w_0, \dots w_k, v_0, \dots, v_k$
- Let $x_{k|k} = x_k | Y_k$ and define

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

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!
$\hookrightarrow$ 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 $\bar{x}_0 \to \hat{x}_{0|-1}$, $\Sigma_0 \to \Sigma_{0|-1}$

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

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

then

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \sim N\left(\begin{bmatrix} \hat{x}_{0|-1} \\ C\hat{x}_{0|-1} \end{bmatrix}, \overbrace{\begin{bmatrix} \Sigma_{0|-1} & \Sigma_{0|-1} C^T \\ C\Sigma_{0|-1} & C\Sigma_{0|-1} C^T + V \end{bmatrix}}^{\tau \begin{bmatrix} \Sigma_0 & 0 \\ 0 & V \end{bmatrix} \tau^T} \right) \qquad (*)$$

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)

$\cdot\ x_1|y_0$ is given by $A(x_0|y_0) + w_0$
$\cdot\ y_1|y_0$ is given by $C(x_1|y_0) + v_1$ } "open loop" predictions $\quad (\blacksquare)$

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

since $\begin{bmatrix} x_1|y_0 \\ y_1|y_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} x_1|y_0 \\ v_1 \end{bmatrix}$

Then, using ($\blacksquare$)

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

$\begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} \Sigma_{1|0} & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} I & C^T \\ 0 & I \end{bmatrix}$

since

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

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

Key observation enabling iterations

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

$$\hat{x}_{0|-1} \to \hat{x}_{1|0}$$
$$\Sigma_{0|-1} \to \Sigma_{1|0}$$

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

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:

statistics of $x_k | y_k$

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

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)

$$x_k | Y_k \sim N(\hat{x}_{k|k}, \Sigma_{k|k})$$
$$x_{k+1} | Y_k \sim N(\hat{x}_{k+1|k}, \Sigma_{k+1|k}) \qquad (\blacksquare)$$

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:

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + w_k \\
y_k &= Cx_k + v_k \\
x_0 &\sim N(\bar{x}_0, \Sigma_0)
\end{aligned}
$$

- 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

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

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

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 $v_0, \ldots, v_{k-1}, w_0, \ldots, w_{k-1}$ 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: $E\left[\left\|e_{k|k-1}\right\|^2\right] = E\left[e_{k|k-1}^T e_{k|k-1}\right] \in \mathbb{R}$
  - Covariance of the error: $E\left[e_{k|k-1} e_{k|k-1}^T\right] \in \mathbb{R}^{n \times n}$
- Abuse of terminology: so far $E\left[e_{k|k-1} e_{k|k-1}^T\right]$ was termed "variance"

Problem (variance minimization)

Compute L_k so as to solve

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

Theorem

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

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

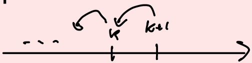
with initial condition

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

Proof by direct minimization (omitted)

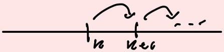
Remarks

- 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 $Var[e_{k|k}] = \Sigma_{k|k}$ 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 $A \to A_k$, $C \to C_k$, $B \to B_k$ 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 $y_{k-1}, y_{k-2}, \ldots$

Statistics of the innovation

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

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

- Variance given by

$$S_{k+1} = E\left[\nu_{k+1}\,\nu_{k+1}^T\right] = E\left[\left(y_{k+1} - C\hat{x}_{k+1|k}\right)\left(y_{k+1} - C\hat{x}_{k+1|k}\right)^T\right]$$
$$= V + C\Sigma_{k+1|k}C^T$$

uncorrelated

$$\nu_{k+1} = C x_{k+1} + v_{k+1} - C\hat{x}_{k+1|k} = C e_{k+1|k} + v_{k+1}$$

$$= \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} e_{k+1|k} \\ v_{k+1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} C & I \end{bmatrix} \begin{bmatrix} \Sigma_{k+1|k} & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} C^T \\ I \end{bmatrix}\right)$$

i.e V_γ is WGN

- 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 tuining 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

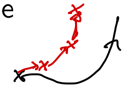
E.g. $y_0 = Cx_0 + v_0$, $y_1 = CAx_0 + Cw_0 + v_1$

$$\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} C & 0 & I & 0 \\ CA & C & 0 & I \end{bmatrix} \begin{bmatrix} x_0 \\ w_0 \\ v_0 \\ v_1 \end{bmatrix} \rightarrow V_{ar}\begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} C & 0 & I & 0 \\ CA & C & 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_0 & & & 0 \\ & W_0 & & \\ & & V_0 & \\ 0 & & & V_1 \end{bmatrix} \begin{bmatrix} C^T & A^TC^T \\ 0 & C^T \\ I & 0 \\ 0 & I \end{bmatrix}$$

example: the bloc 1,2 is $\begin{bmatrix} C\Sigma_0 & 0 & V_0 & 0 \end{bmatrix} \begin{bmatrix} A^TC^T \\ C^T \\ 0 \\ I \end{bmatrix} = C\Sigma_0 A^T C^T \neq 0$

Example 1: estimation of position and velocity of a ground vehicle[1]

Motivation: develop a navigation and tracking system for vehicle



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

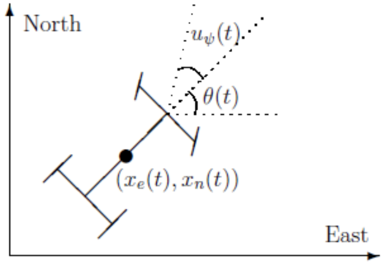
- Solution 1: use encoders attached to the vehicle wheels
  $\hookrightarrow$ imprecise $\rightarrow$ the estimated and real positions can be far away from the real ones
- Solution 2: use GPS measurements
  $\hookrightarrow$ can give much better results, but measurements are noisy and velocities are 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:

$$x(k+1) = Ax(k) + Gw(k)$$
$$y(k) = Cx(k) + \mathrm{v}(k)$$

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 $\mathrm{v}(k) = \begin{bmatrix} \mathrm{v}_1(k) \\ \mathrm{v}_2(k) \end{bmatrix}$ is the measurement noise

Example 1 (cont.)

$$x(k+1) = Ax(k) + Gw(k)$$
$$y(k) = Cx(k) + \mathrm{v}(k)$$

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

Considerations:

- velocites are modelled as a random walk:

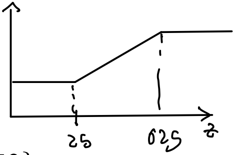
$$v_i(k+1) = v_i(k) + w_i(k), \qquad i \in \{1, 2\}$$

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

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

- $T_s = 1s$

Example 1 (cont.)



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

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

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

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

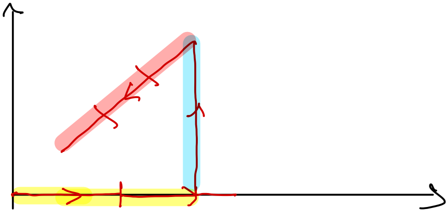
$\hookrightarrow$ Captures the intuition that typical values of w are smaller when velocity is large.
$\hookrightarrow$ A diagonal W_k represents the naive assumption that the velocity changes in the north and east directions are uncorrelated

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

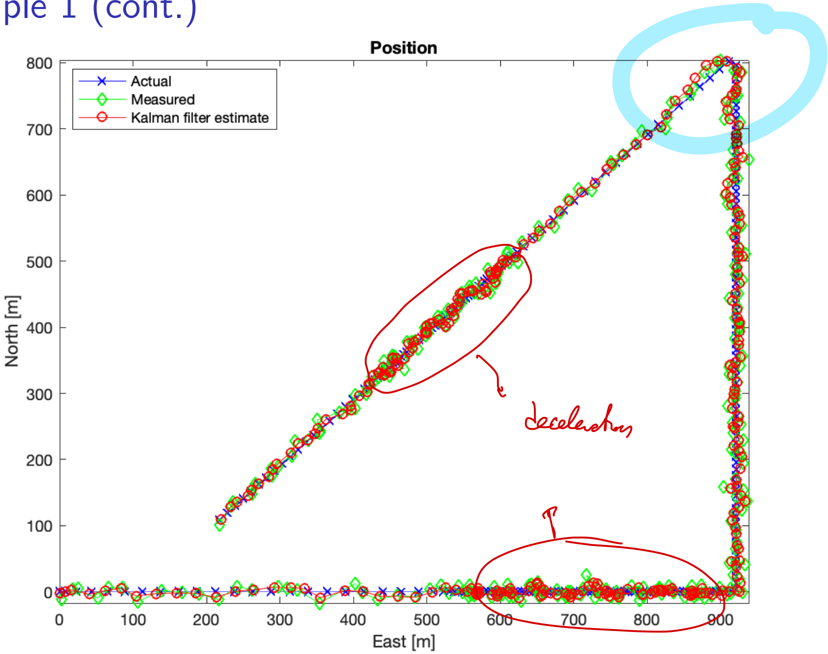
Example 1 (cont.)

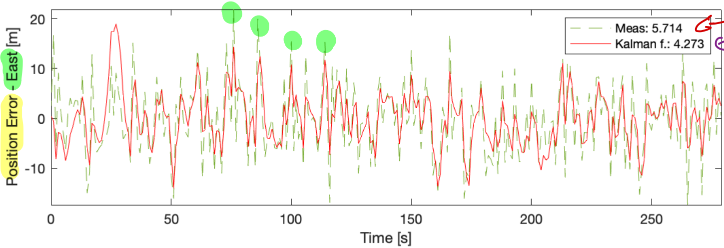
Simulation scenario: the vehicle makes the following maneuvers:

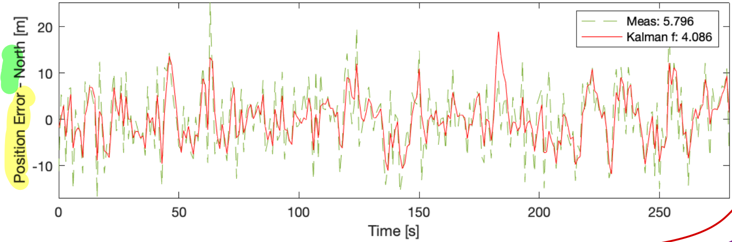
1. At t $= 0$ the vehicle is at $x_e(0) = 0$, $x_n(0) = 0$ and is stationary
2. Heading east, it accelerates to 25m/s. It decelerates to 5m/s at t=20s.
3. At t $=$ 100s, it turns toward north and accelerates to 10m/s.
4. At t $=$ 180s, it accelerates to 20m/s with southwest direction.
5. At t $=$ 200s,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



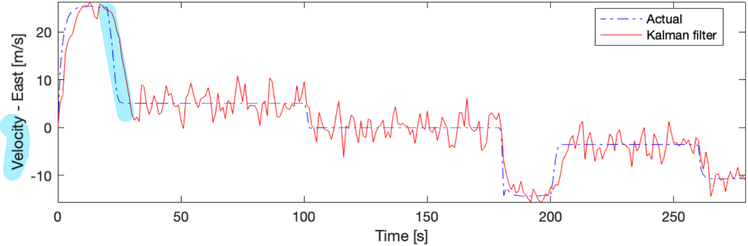
Example 1 (cont.)

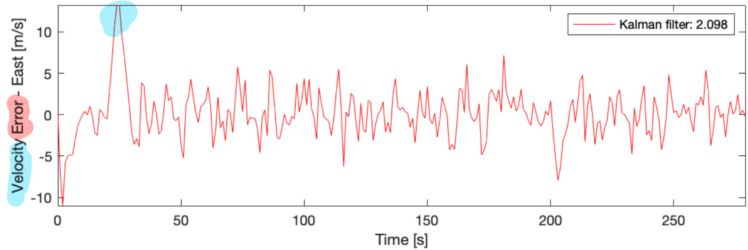




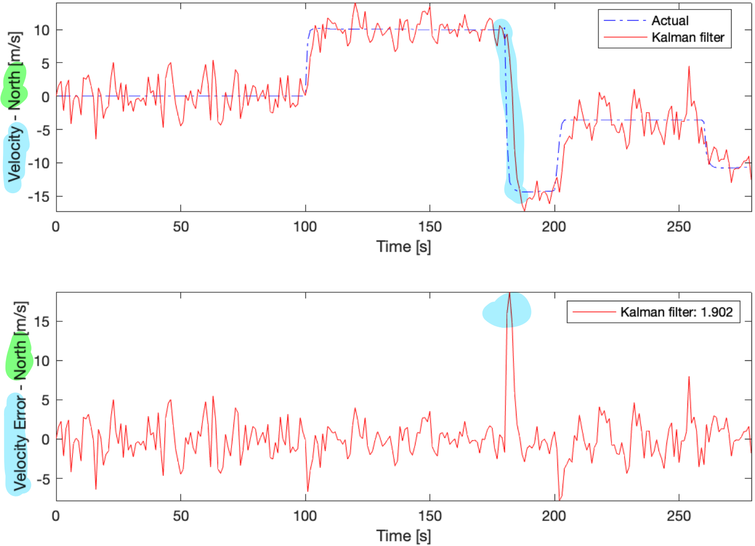


$$\text{Meas:} \frac{1}{N_s} \sum_{k=1}^{N_s} |x_i(k) - y_i(k)| \qquad \text{Kalman f:} \frac{1}{N_s} \sum_{k=1}^{N_s} |x_i(k) - \hat{x}_i(k)| \qquad i = \{1, 2\}$$





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



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

Comments:

- The Kalman filters 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

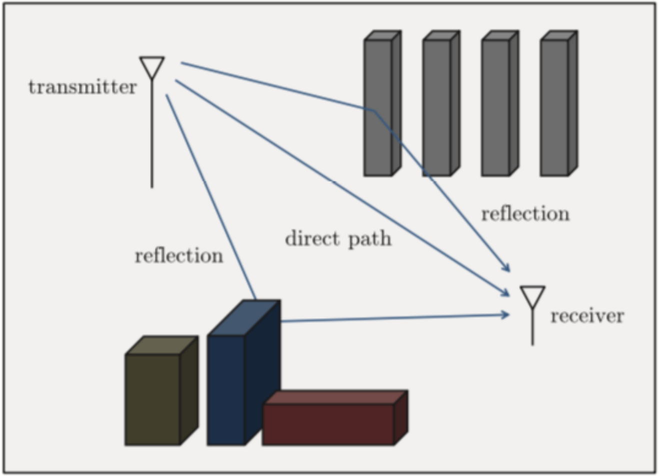
$\hookrightarrow$ After a few time steps, the filter estimates catch up with the actual velocity.

1h 12

Additional example (check at home!)

Example 2: Channel estimation in communication systems[2]

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



[2] N. 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 $k-1, k-2, \ldots$
  - noise

  Multipath propagation

- The propagation channel changes over time due to:
  - movements of the transmitter/receiver
  - changes of the environment

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

$\rightarrow$ 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 $i = 1, 2, 3, \ldots$ 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*"
$$x_i(k+1) = x_i(k) + w_i(k) \qquad w(k) \sim N(0, W)$$
- Other option:
$\rightarrow$ Take into account the correlation (α_i) between two sequent values
$$x_i(k+1) = \alpha_i x_i(k) + w_i(k) \qquad w(k) \sim N(0, W)$$

- 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:

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

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,

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

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

Example 2 (cont.)

Simulations:

Set

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

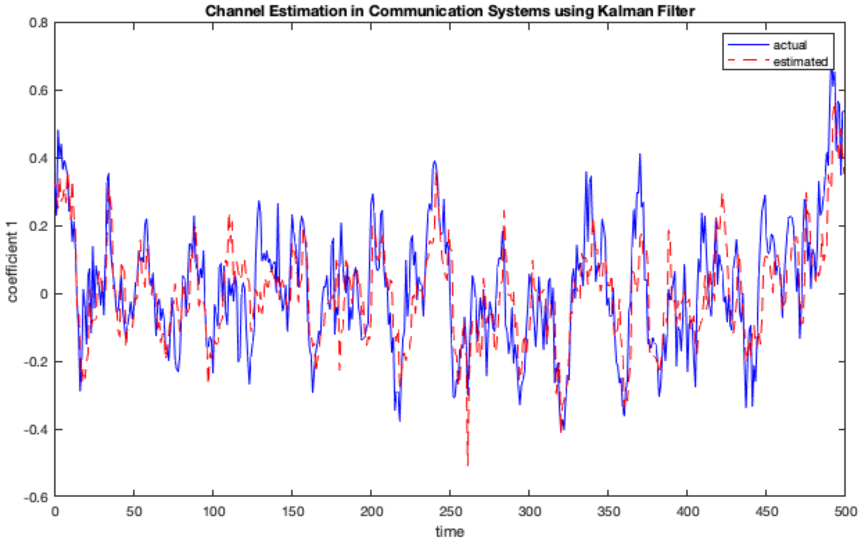
and generate a test signal sequence $c(k)$ (for example a random *but known* signal).

Also consider:

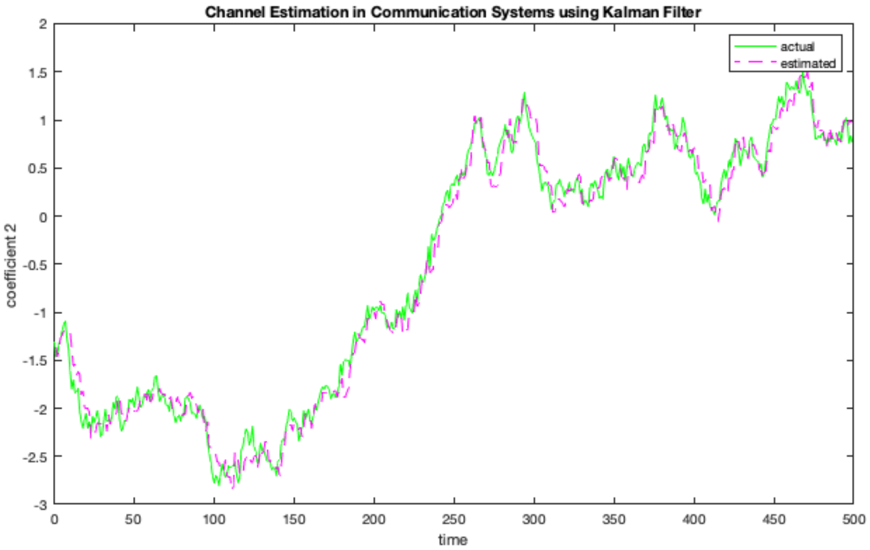
- initial channel coefficients $x(0) \sim N(0, I)$
- KF initial variance $\Sigma_{0|-1} = I$

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

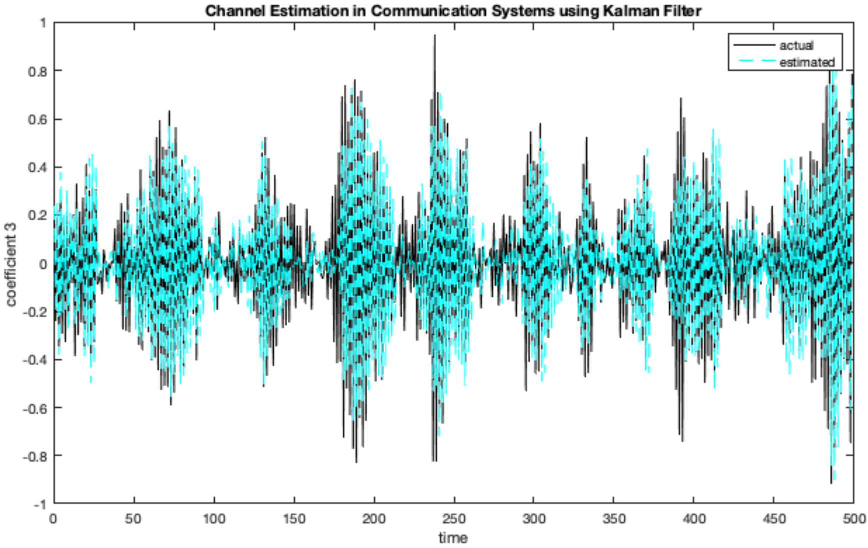
Example 2 (cont.)



Example 2 (cont.)



Example 2 (cont.)



Example 2 (cont.)

Remarks:

- The vertical scale of coefficient 2 is larger that 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.