TTK4115

Lecture 12

The Kalman filter part 2

Morten. O. Alver (based on material by Morten D. Pedersen)

This lecture

1. Discrete time modeling

2. Kalman filtering in discrete time

3. Time varying models

Topic

1. Discrete time modeling

2. Kalman filtering in discrete time

3. Time varying models

Discrete time Kalman filter

Measurements $\mathbf{y}$ are typically obtained through sampling at discrete intervals in time $t = kT,\ k = 0, 1, 2, \ldots$. Furthermore, estimates $\hat{\mathbf{x}}$ will typically be requested at discrete intervals. For these reasons (and others), the discrete Kalman filter is the version that sees most frequent use (by far).

Discrete time analysis

The passage from continuous to discrete time introduces a range of changes, some of which are quite subtle.

Continuous time random process

The continuous time plant model is given by the random process

$$\dot{\mathbb{x}} = \mathbf{A}\mathbb{x} + \mathbf{B}\mathbf{u} + \mathbf{G}\mathbb{w}, \quad \mathbb{y} = \mathbf{C}\mathbb{x} + \mathbb{v}$$

where the noise and disturbance are unbiased ($\mathbf{m}_\mathbf{v} = \mathbf{0},\ \mathbf{m}_\mathbf{w} = \mathbf{0}$) and white

$$\mathcal{A}_\mathbf{v}(t,\tau) = \mathrm{E}[\mathbb{v}(t)\mathbb{v}(\tau)^\mathsf{T}] = \delta(t-\tau)\mathbf{R}, \quad \mathcal{A}_\mathbf{w}(t,\tau) = \mathrm{E}[\mathbb{w}(\tau)\mathbb{w}(t)^\mathsf{T}] = \delta(t-\tau)\mathbf{Q}$$

Exact solution

Knowing the solution permits exact discretization. For the process model given above, an **exact** solution is furnished by

$$\mathbb{x}(t) = e^{\mathbf{A}t}\mathbb{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{G}\mathbb{w}(\tau)\,d\tau, \quad \mathbb{y}(t) = \mathbf{C}\mathbb{x}(t) + \mathbb{v}(t)$$

Exact solution

$$\mathbb{x}(t) = e^{\mathbf{A}t}\mathbb{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\, d\tau + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{G}\mathbb{w}(\tau)\, d\tau, \quad \mathbb{y}(t) = \mathbf{C}\mathbb{x}(t) + \mathbb{v}(t)$$

Discretization

Starting the solution at $t = kT$ and terminating it at $t = (k+1)T$ produces

$$\mathbb{x}[k+1] = e^{\mathbf{A}T}\mathbb{x}[k] + \int_0^T e^{\mathbf{A}\alpha}\mathbf{B}\mathbf{u}((k+1)T - \alpha)\, d\alpha + \int_0^T e^{\mathbf{A}\alpha}\mathbf{G}\mathbb{w}((k+1)T - \alpha)\, d\alpha$$

A calculation variable $\alpha = (k+1)T - \tau$ is here introduced to make life easier.

$$\bar{\mathbf{A}} \triangleq e^{\mathbf{A}T}, \quad \bar{\mathbf{B}} \triangleq \int_0^T e^{\mathbf{A}\alpha}\mathbf{B}\, d\alpha, \quad \bar{\mathbb{w}}[k] \triangleq \int_0^T e^{\mathbf{A}\alpha}\mathbf{G}\mathbb{w}((k+1)T - \alpha)\, d\alpha$$

Assuming that the deterministic input $\mathbf{u}(t)$ varies little over $(k+1)T \leq t \leq (k+1)T$ yields the discretized model

$$\mathbb{x}[k+1] = \bar{\mathbf{A}}\mathbb{x}[k] + \bar{\mathbf{B}}\mathbf{u}[k] + \bar{\mathbb{w}}[k]$$

Note that the discretized noise contribution is quite different from the continuous time variety, $\bar{\mathbb{w}}[k] \neq \mathbb{w}(kT)$.

Discrete time white **disturbances**

The discrete time white disturbance signal is now subjected to a closer examination.

$$\bar{\mathbb{w}}[k] \triangleq \int_0^T e^{\mathbf{A}\alpha}\mathbf{G}\mathbb{w}((k+1)T-\alpha)\,d\alpha$$

It is straightforward to verify that $\bar{\mathbb{w}}[k]$ inherits the unbiased nature of $\mathbb{w}(t)$. But, the autocovariance (incl. variance) changes in a subtle fashion. The discrete time autocovariance of $\bar{\mathbb{w}}[k]$ is given by

$$\begin{aligned}\bar{\mathcal{A}}_{\mathbf{w}}[k,l] &= \mathsf{E}[\bar{\mathbb{w}}[k]\bar{\mathbb{w}}[l]^{\mathsf{T}}] \\ &= \int_0^T\int_0^T e^{\mathbf{A}\alpha_1}\mathbf{G}\ \underbrace{\mathsf{E}[\mathbb{w}((k+1)T-\alpha_1)\mathbb{w}((l+1)T-\alpha_2)^{\mathsf{T}}]}_{\mathcal{A}_{\mathbf{w}}((k+1)T-\alpha_1,(l+1)T-\alpha_2)=\delta((l-k)T+\alpha_1-\alpha_2)\mathbf{Q}_{\mathbf{w}}}\ \mathbf{G}^{\mathsf{T}}e^{\mathbf{A}^{\mathsf{T}}\alpha_2}\,d\alpha_1\,d\alpha_2\end{aligned}$$

Kronecker's δ-function satisfies

$$\delta[k,l] = \begin{cases} 1 & k=l \\ 0 & k\neq l\end{cases}$$

Noting that $\delta((l-k)T+\alpha_1-\alpha_2) = \delta[k,l]\delta(\alpha_1-\alpha_2)$ the result follows

$$\bar{\mathcal{A}}_{\mathbf{w}}[k,l] = \delta[k,l]\bar{\mathbf{Q}}_{\mathbf{w}}, \quad \bar{\mathbf{Q}}_{\mathbf{w}} \triangleq \int_0^T e^{\mathbf{A}\alpha}\mathbf{G}\mathbf{Q}_{\mathbf{w}}\mathbf{G}^{\mathsf{T}}e^{\mathbf{A}^{\mathsf{T}}\alpha}\,d\alpha$$

Exact discretization has rendered the infinite variance of $\mathbb{w}(t)$ *finite* and equal to $\bar{\mathcal{A}}_{\mathbf{w}}[k,k] = \bar{\mathbf{Q}}_{\mathbf{w}}$ in discrete time (this in fact a consequence of the *central limit theorem*).

Discrete time white noise

The measurement model in continuous time is given by

$$\mathbb{y} = \mathbf{C}\mathbb{x} + \mathbb{v}, \quad \mathcal{A}_{\mathbf{v}}(t,\tau) = \mathsf{E}[\mathbb{v}(t)\mathbb{v}(\tau)^{\mathsf{T}}] = \delta(t-\tau)\mathbf{R}_{\mathbf{v}}$$

A naïve conversion to discrete time would suggest

$$\mathbb{y}[k] = \mathbf{C}\mathbb{x}[k] + \mathbb{v}[k], \quad \bar{\mathbf{R}}_{\mathbf{v}} = \mathsf{E}[\mathbb{v}[k]\mathbb{v}[k]^{\mathsf{T}}] = \delta(0)\mathbf{R}_{\mathbf{v}}$$

This interpretation leads to extreme exaggerations of noise in discrete time (but is suitable in continuous time).

Averaging convention

Rather than interpreting measurement noise as occuring at the instant of sampling, it can be interpreted in a *averaged* sense. This idea is captured in the convention

$$\bar{\mathbb{v}}[k] \triangleq \frac{1}{T}\int_0^T \mathbb{v}(kT-\alpha)\, d\alpha$$

The discrete time noise vector inherits the unbiased nature of the continuous time signal, whilst the autocovariance transforms to

$$\bar{\mathcal{A}}_{\mathbf{v}}[k,l] = \mathsf{E}[\bar{\mathbb{v}}[k]\bar{\mathbb{v}}[l]^{\mathsf{T}}] = \frac{1}{T^2}\int_0^T\int_0^T \mathsf{E}[\mathbb{v}(kT-\alpha_1)\mathbb{v}(lT-\alpha_2)^{\mathsf{T}}]\, d\alpha_1 d\alpha_2 = \delta[k,l]\bar{\mathbf{R}}_{\mathbf{v}}, \quad \bar{\mathbf{R}}_{\mathbf{v}} \triangleq \mathbf{R}_{\mathbf{v}}/T$$

Discrete time random process

The discrete time plant model is given by the random process

$$\mathbb{x}[k+1] = \bar{\mathbf{A}}\mathbb{x}[k] + \bar{\mathbf{B}}\mathbf{u}[k] + \bar{\mathbb{w}}[k], \quad \mathbb{y}[k] = \mathbf{C}\mathbb{x}[k] + \bar{\mathbb{v}}[k]$$

where the noise and disturbance are unbiased ($\mathbf{m_v} = \mathbf{0},\ \mathbf{m_w} = \mathbf{0}$) and white

$$\bar{\mathcal{A}}_{\mathbf{v}}[k,l] = \mathrm{E}[\bar{\mathbb{v}}[k]\bar{\mathbb{v}}[l]^{\mathsf{T}}] = \delta[k,l]\bar{\mathbf{R}}_{\mathbf{v}}, \quad \bar{\mathcal{A}}_{\mathbf{w}}[k,l] = \mathrm{E}[\bar{\mathbb{w}}[k]\bar{\mathbb{w}}[l]^{\mathsf{T}}] = \delta[k,l]\bar{\mathbf{Q}}_{\mathbf{w}}$$

It will be assumed that the noise and disturbance processes are uncorrelated $\mathrm{E}[\bar{\mathbb{v}}[k]\bar{\mathbb{w}}[l]^{\mathsf{T}}] = \mathbf{0}$.

Continuous to discrete conversion - sampling time T.

Transition matrix: Obtained from exact discretization.

$$\bar{\mathbf{A}} = e^{\mathbf{A}T}$$

Input matrix: Obtained from exact discretization & assumption of constant $\mathbf{u}$ over sampling period.

$$\bar{\mathbf{B}} = \int_0^T e^{\mathbf{A}\alpha}\mathbf{B}\, d\alpha$$

Disturbance covariance: Obtained from exact discretization.

$$\bar{\mathbf{Q}}_{\mathbf{w}} = \int_0^T e^{\mathbf{A}\alpha}\mathbf{G}\mathbf{Q}_{\mathbf{w}}\mathbf{G}^\top e^{\mathbf{A}^\top\alpha}\, d\alpha$$

Noise covariance: Obtained through an averaging convention.

$$\bar{\mathbf{R}}_{\mathbf{v}} = \mathbf{R}_{\mathbf{v}}/T$$

Van Loan's method[1]

The integrals in the preceding slide are often quite intractable. It is however possible to arrive at the correct matrices without integrating. This is done with *Van Loan's method*. The key result is

$$\exp\left(\begin{bmatrix} \mathbf{A} & \mathbf{G}\mathbf{Q}_\mathbf{w}\mathbf{G}^\mathsf{T} \\ \mathbf{0} & -\mathbf{A}^\mathsf{T} \end{bmatrix} T\right) = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{0} & \mathbf{M}_{22} \end{bmatrix}, \quad \bar{\mathbf{A}} = \mathbf{M}_{11}, \quad \bar{\mathbf{Q}}_\mathbf{w} = \mathbf{M}_{12}\mathbf{M}_{11}^\mathsf{T}$$

Matrix exponentials are readily computed numerically, obviating the need for integration. The input matrix can be computed from

$$\exp\left(\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} T\right) = \begin{bmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{0} & \mathbb{I} \end{bmatrix}, \quad \bar{\mathbf{A}} = \mathbf{N}_{11}, \quad \bar{\mathbf{B}} = \mathbf{N}_{12}$$

[1] Van Loan C.F. (1978), *Computing Integrals Involving the Matrix Exponential*, IEEE Transactions on Automatic Control, Vol. 23, No. 3. See also page 126 in the Brown & Hwang book.

Topic

Discrete observer

Discrete time requires a sligthly more explicit observer design. The estimate is generated in two distinct phases:

---

1 - A priori (denoted $\hat{\mathbb{x}}^-[k]$): The best guess for $\mathbb{x}[k]$ **prior** to incorporation of the measurement $\mathbb{y}[k]$. The deterministic model is to arrive at this estimate.

$$\hat{\mathbb{x}}^-[k] = \bar{\mathbf{A}}\hat{\mathbb{x}}[k-1] + \bar{\mathbf{B}}\mathbf{u}[k-1]$$

---

2 - A posteriori (denoted $\hat{\mathbb{x}}[k]$): The best guess for $\mathbb{x}[k]$ **after** incorporation of the measurement $\mathbb{y}[k]$. A linear blend of what the model suggests ($\hat{\mathbb{x}}^-[k]$) and the new measurement $\mathbb{y}[k]$ is used to arrive at this final estimate. The Kalman gain $\mathbf{L}[k]$ serves as the *blending factor*, viz.

$$\hat{\mathbb{x}}[k] = \hat{\mathbb{x}}^-[k] + \mathbf{L}[k](\mathbb{y}[k] - \mathbf{C}\hat{\mathbb{x}}^-[k])$$

Kalman gain

The Kalman gain is (as for the continuous time case) designed to minimize the mean-square error of the estimate at time k.

$$J[k] = \mathrm{tr}(\mathbf{P}[k]), \quad \mathbf{P}[k] \triangleq \mathrm{E}[(\mathbb{x}[k] - \hat{\mathbb{x}}[k])(\mathbb{x}[k] - \hat{\mathbb{x}}[k])^\mathsf{T}]$$

A **priori** error and covariance matrix

The a priori and a posteriori estimation errors and covariance matrices are given by

$$\mathbb{e}^-[k] \triangleq \mathbb{x}[k] - \hat{\mathbb{x}}^-[k], \quad \mathbf{P}^-[k] \triangleq \mathsf{E}[\mathbb{e}^-[k]\mathbb{e}^-[k]^\mathsf{T}]$$
$$\mathbb{e}[k] \triangleq \mathbb{x}[k] - \hat{\mathbb{x}}[k], \quad \mathbf{P}[k] \triangleq \mathsf{E}[\mathbb{e}_k\mathbb{e}_k^\mathsf{T}]$$

The process model produces the following state at k

$$\mathbb{x}[k] = \bar{\mathbf{A}}\mathbb{x}[k-1] + \bar{\mathbf{B}}\mathbf{u}[k-1] + \bar{\mathbb{w}}[k-1]$$

whereas the a priori estimate reads as

$$\hat{\mathbb{x}}^-[k] = \bar{\mathbf{A}}\hat{\mathbb{x}}[k-1] + \bar{\mathbf{B}}\mathbf{u}[k-1]$$

This permits the following expression for the **a priori** error

$$\mathbb{e}^-[k] = \bar{\mathbf{A}}\mathbb{e}[k-1] + \bar{\mathbb{w}}[k-1]$$

The **a priori** covariance matrix follows as

$$\mathbf{P}^-[k] = \mathsf{E}[(\bar{\mathbf{A}}\mathbb{e}[k-1] + \bar{\mathbb{w}}[k-1])(\bar{\mathbf{A}}\mathbb{e}[k-1] + \bar{\mathbb{w}}[k-1])^\mathsf{T}] = \bar{\mathbf{A}}\mathbf{P}[k-1]\bar{\mathbf{A}}^\mathsf{T} + \bar{\mathbf{Q}}_\mathbf{w}$$

The disturbance at k is uncorrelated to the a-posteriori estimate at k, hence $\mathsf{E}[\mathbb{e}[k]\bar{\mathbb{w}}[k]^\mathsf{T}] = \mathbf{0}$.

A **posteriori** error and covariance matrix

The a priori and a posteriori estimation errors and covariance matrices are given by

$$\mathbb{e}^-[k] \triangleq \mathbb{x}[k] - \hat{\mathbb{x}}^-[k], \quad \mathbf{P}^-[k] \triangleq \mathsf{E}[\mathbb{e}^-[k]\mathbb{e}^-[k]^\mathsf{T}]$$
$$\mathbb{e}[k] \triangleq \mathbb{x}[k] - \hat{\mathbb{x}}[k], \quad \mathbf{P}[k] \triangleq \mathsf{E}[\mathbb{e}_k\mathbb{e}_k^\mathsf{T}]$$

The a posteriori estimate can be expanded to read

$$\hat{\mathbb{x}}[k] = \hat{\mathbb{x}}^-[k] + \mathbf{L}[k](\mathbb{y}[k] - \mathbf{C}\hat{\mathbb{x}}^-[k]) = \hat{\mathbb{x}}^-[k] + \mathbf{L}[k]\mathbf{C}\mathbb{e}^-[k] + \mathbf{L}[k]\bar{\mathbb{v}}[k]$$

This permits the following expression for the **a posteriori** error

$$\mathbb{e}[k] = (\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbb{e}^-[k] - \mathbf{L}[k]\bar{\mathbb{v}}[k]$$

The **a posteriori** covariance matrix follows as

$$\begin{aligned}\mathbf{P}[k] = \mathsf{E}[((\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbb{e}^-[k] - \mathbf{L}[k]\bar{\mathbb{v}}[k])((\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbb{e}^-[k] - \mathbf{L}[k]\bar{\mathbb{v}}[k])^\mathsf{T}] \\ = (\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbf{P}^-[k](\mathbb{I} - \mathbf{L}[k]\mathbf{C})^\mathsf{T} + \mathbf{L}[k]\bar{\mathbf{R}}_\mathbf{v}[k]\mathbf{L}[k]^\mathsf{T}\end{aligned}$$

The noise at k is uncorrelated to the a-priori estimate at k, hence $\mathsf{E}[\mathbb{e}^-[k]\bar{\mathbb{v}}[k]^\mathsf{T}] = \mathbf{0}$.

Optimal estimation

The **a posteriori** covariance matrix describes the covariance of the final estimate error $\mathbb{e}[k] = \mathbb{x}[k] - \hat{\mathbb{x}}[k]$. We now seek to minimize the mean-square error

$$J[k] = \mathrm{tr}(\mathbf{P}[k])$$

We need to differentiate w.r.t. to the Kalman gain and solve for the extremum.

Matrix differentiation rules

These are the differentiation rules that we will use:

$$\frac{\partial}{\partial \mathbf{X}} \mathrm{tr}(\mathbf{X}\mathbf{A}) = \mathbf{A}^{\mathsf{T}}$$
$$\frac{\partial}{\partial \mathbf{X}} \mathrm{tr}(\mathbf{A}\mathbf{X}^{\mathsf{T}}) = \mathbf{A}$$
$$\frac{\partial}{\partial \mathbf{X}} \mathrm{tr}(\mathbf{X}\mathbf{B}\mathbf{X}^{\mathsf{T}}) = \mathbf{X}\mathbf{B}^{\mathsf{T}} + \mathbf{X}\mathbf{B}$$

Optimal estimation:

These are the differentiation rules that we will use:

$$\frac{\partial}{\partial \mathbf{X}}\text{tr}(\mathbf{XA}) = \mathbf{A}^\mathsf{T}, \quad \frac{\partial}{\partial \mathbf{X}}\text{tr}(\mathbf{AX}^\mathsf{T}) = \mathbf{A}, \quad \frac{\partial}{\partial \mathbf{X}}\text{tr}(\mathbf{XBX}^\mathsf{T}) = \mathbf{XB}^\mathsf{T} + \mathbf{XB}$$

The expression for $\mathbf{P}[k]$ is:

$$\mathbf{P}[k] = (\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbf{P}^-[k](\mathbb{I} - \mathbf{L}[k]\mathbf{C})^\mathsf{T} + \mathbf{L}[k]\bar{\mathbf{R}}_\mathbf{v}[k]\mathbf{L}[k]^\mathsf{T}$$

Optimal estimation

The **a posteriori** covariance matrix describes the covariance of the final estimate error $\mathbb{e}[k] = \mathbb{x}[k] - \hat{\mathbb{x}}[k]$. We now seek to minimize the mean-square error

$$J[k] = \mathrm{tr}(\mathbf{P}[k])$$

Differentiation w.r.t. to the Kalman gain and solving for the extremum yields:

$$\begin{aligned}\frac{\partial \mathrm{tr}(\mathbf{P}[k])}{\partial \mathbf{L}[k]} &= \frac{\partial}{\partial \mathbf{L}[k]} \mathrm{tr}\left((\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbf{P}^{-}[k](\mathbb{I} - \mathbf{L}[k]\mathbf{C})^{\mathsf{T}} + \mathbf{L}[k]\bar{\mathbf{R}}_{\mathbf{v}}[k]\mathbf{L}[k]^{\mathsf{T}} \right) \\ &= -2\mathbf{P}^{-}[k]\mathbf{C}^{\mathsf{T}} + 2\mathbf{L}[k](\mathbf{C}\mathbf{P}^{-}[k]\mathbf{C}^{\mathsf{T}} + \bar{\mathbf{R}}_{\mathbf{v}}) = \mathbf{0}\end{aligned}$$

The *Kalman gain* thus follows as

$$\mathbf{L}[k] = \mathbf{P}^{-}[k]\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{P}^{-}[k]\mathbf{C}^{\mathsf{T}} + \bar{\mathbf{R}}_{\mathbf{v}})^{-1}$$

Discrete Kalman filter algorithm

The filter is initialized at

$$\hat{\mathbf{x}}^-[0] = \mathrm{E}[\mathbb{x}(0)] = \mathbf{m}_{\mathbf{x}_0}$$
$$\mathbf{P}^-[0] = \mathrm{E}[\mathbb{e}^-[0]\mathbb{e}^-[0]^\mathsf{T}] = \mathrm{E}[(\mathbb{x}[0] - \mathbf{m}_{\mathbf{x}_0})(\mathbb{x}[0] - \mathbf{m}_{\mathbf{x}_0})^\mathsf{T}] = \mathcal{C}_{\mathbf{x}_0}$$

The recursive algorithm running over $k = 0 \ldots K$ is summarized by

1 - Compute Kalman gain

$$\mathbf{L}[k] = \mathbf{P}^-[k]\mathbf{C}^\mathsf{T}(\mathbf{C}\mathbf{P}^-[k]\mathbf{C}^\mathsf{T} + \bar{\mathbf{R}}_\mathbf{v})^{-1}$$

2 - Update estimate with measurement

$$\hat{\mathbb{x}}[k] = \hat{\mathbb{x}}^-[k] + \mathbf{L}[k](\mathbb{y}[k] - \mathbf{C}\hat{\mathbb{x}}^-[k])$$

3 - Update error covariance matrix

$$\mathbf{P}[k] = (\mathbb{I} - \mathbf{L}[k]\mathbf{C})\mathbf{P}^-[k](\mathbb{I} - \mathbf{L}[k]\mathbf{C})^\mathsf{T} + \mathbf{L}[k]\bar{\mathbf{R}}_\mathbf{v}[k]\mathbf{L}[k]^\mathsf{T}$$

4 - Project ahead

$$\hat{\mathbb{x}}^-[k+1] = \bar{\mathbf{A}}\hat{\mathbb{x}}[k] + \bar{\mathbf{B}}\mathbf{u}[k]$$
$$\mathbf{P}^-[k+1] = \bar{\mathbf{A}}\mathbf{P}[k]\bar{\mathbf{A}}^\mathsf{T} + \bar{\mathbf{Q}}_\mathbf{w}$$

Example: handheld GPS

Problem

GPS measurements are typically available at a sample time $T \sim 1[s]$. It is assumed that the horisontal measurements are approximately normally distributed around the true position $\mathbb{p} = [\mathbb{p}_1 \ \mathbb{p}_2]^T$ with a standard deviation $\sigma_v \sim 5[m]$. A measurement model is thus

$$\mathbb{y}[k] = \mathbb{p}[k] + \mathbb{v}[k], \quad \bar{\mathbf{R}} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}$$

---

1. How can one improve upon the direct measurement?
2. Is it possible to obtain velocity estimates?

Solution

The desired improvements can be had by incorporating system knowledge. The position of the handheld GPS unit will change in a manner that cannot be predicted exactly. We assume instead that the user moves in accordance with the random model

$$\tau \ddot{\mathbb{p}}_1 + \dot{\mathbb{p}}_1 = \mathbb{w}_1$$
$$\tau \ddot{\mathbb{p}}_2 + \dot{\mathbb{p}}_2 = \mathbb{w}_2$$

Physically, this model represents a mass-damper perturbed by an unknown force.

Note that the velocities $\dot{\mathbb{p}}$ enter as states of the model and can therefore be <u>estimated</u>.

Example: handheld GPS

Continuous time random process

The intensities of the disturbance signals and the time-constant τ should be tuned through practical experiments. A useful model structure can however be supplied as

$$\overbrace{\begin{bmatrix} \dot{\mathbb{p}}_1 \\ \dot{\mathbb{p}}_2 \\ \ddot{\mathbb{p}}_1 \\ \ddot{\mathbb{p}}_2 \end{bmatrix}}^{\dot{\mathbb{x}}} = \overbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\tau^{-1} & 0 \\ 0 & 0 & 0 & -\tau^{-1} \end{bmatrix}}^{\mathbf{A}} \overbrace{\begin{bmatrix} \mathbb{p}_1 \\ \mathbb{p}_2 \\ \dot{\mathbb{p}}_1 \\ \dot{\mathbb{p}}_2 \end{bmatrix}}^{\mathbb{x}} + \overbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \tau^{-1} & 0 \\ 0 & \tau^{-1} \end{bmatrix}}^{\mathbf{G}} \overbrace{\begin{bmatrix} \mathbb{w}_1 \\ \mathbb{w}_2 \end{bmatrix}}^{\mathbb{w}}$$

$$\underbrace{\begin{bmatrix} \mathbb{y}_1 \\ \mathbb{y}_2 \end{bmatrix}}_{\mathbb{y}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} \mathbb{p}_1 \\ \mathbb{p}_2 \\ \dot{\mathbb{p}}_1 \\ \dot{\mathbb{p}}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbb{v}_1 \\ \mathbb{v}_2 \end{bmatrix}}_{\mathbb{v}}$$

where

$$\mathbf{Q} = q \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{R} = \sigma_v^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} T$$

Discrete time random process

The Handheld GPS problem is solved using a discrete time Kalman filter. Tuning constants are chosen as $\tau = 200$ and $q = 25^2$.

Using Van Loan's method the discrete time system matrices $\bar{\mathbf{A}}$ and $\bar{\mathbf{Q}}$ can be found precisely:

$$\bar{\mathbf{A}} = \begin{bmatrix} 1 & 0 & 0.9975 & 0 \\ 0 & 1 & 0 & 0.9975 \\ 0 & 0 & 0.995 & 0 \\ 0 & 0 & 0 & 0.995 \end{bmatrix}, \quad \bar{\mathbf{Q}} = \begin{bmatrix} 0.0052 & 0 & 0.0078 & 0 \\ 0 & 0.0052 & 0 & 0.0078 \\ 0.0078 & 0 & 0.0155 & 0 \\ 0 & 0.0078 & 0 & 0.0155 \end{bmatrix}$$

Note: the discrete system, instead of 2 uncorrelated disturbance signals, has 4 signals with off-diagonal correlations. This is because the disturbances propagate through the system during the time period T.

The final model reads as

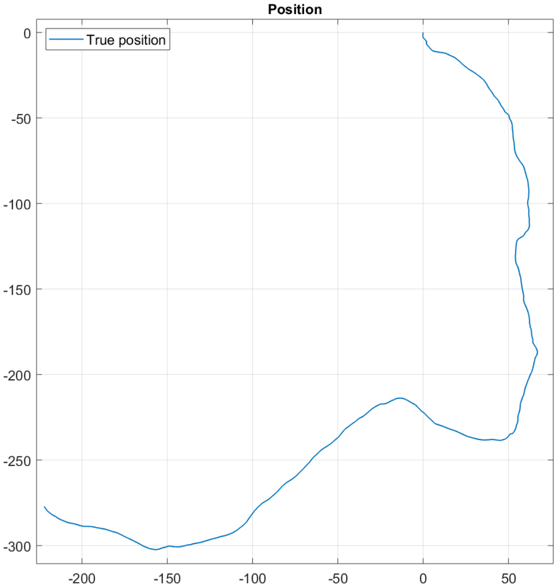
$$\mathbb{x}[k+1] = \bar{\mathbf{A}}\mathbb{x}[k] + \bar{\mathbb{w}}[k], \quad \mathbb{y}[k] = \mathbf{C}\mathbb{x}[k] + \bar{\mathbb{v}}[k]$$

where $\bar{\mathbf{Q}}$ and $\bar{\mathbf{R}}$ describe the respective covariances of the disturbance and noise signals.

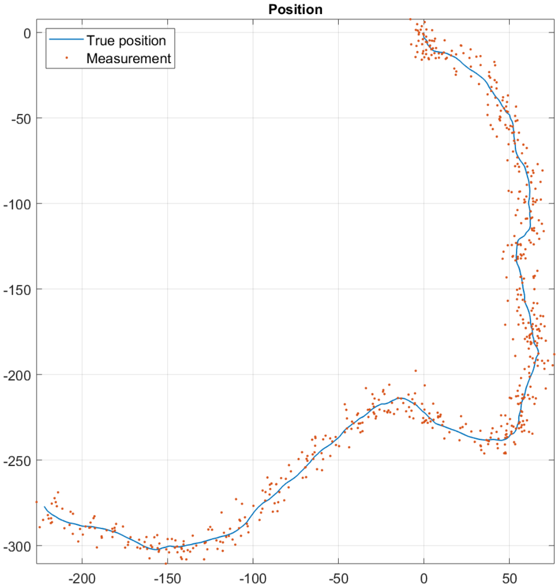
Matlab Demo

Model setup, discretization, simulation and plotting is done in the *handheld_GPS.m* Matlab script (available on Blackboard).

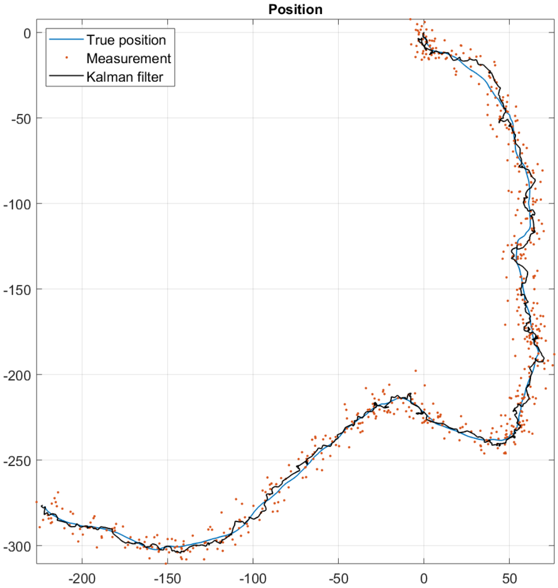
Example realization of random walk



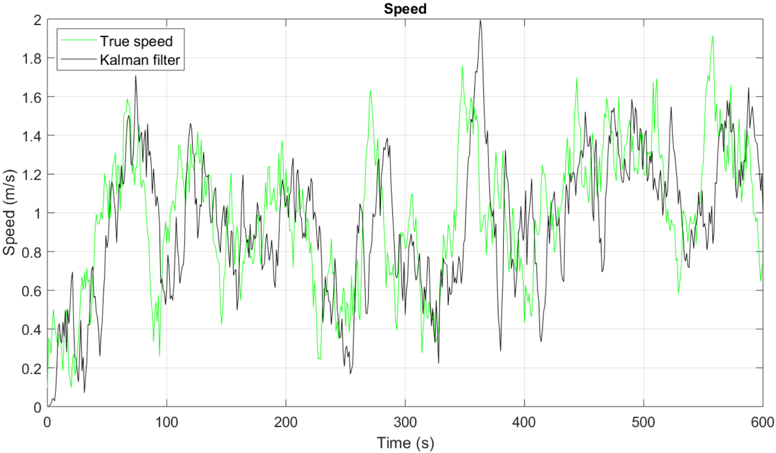
Noisy measurements of true position



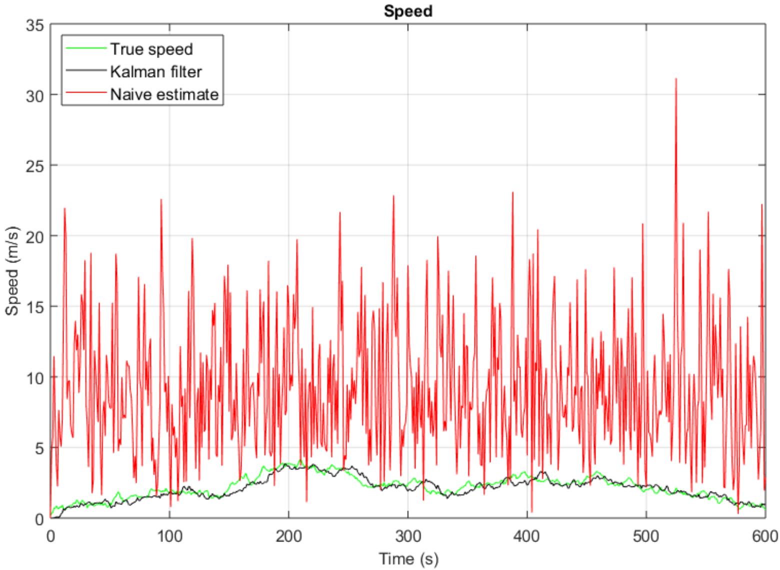
Kalman filter estimate of position



True speed vs. estimated velocity



Comparison to ”naive” velocity estimate



Topic

Time varying models

The continuous and discrete time Kalman filters are not limited to time-invariant plants. They can in fact serve as optimal estimators for time-varying systems!

LTV system

Let a linear time-varying random process be given by

$$\dot{\mathbb{x}}(t) = \mathbf{A}(t)\mathbb{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbb{w}(t), \quad \mathbb{y}(t) = \mathbf{C}(t)\mathbb{x}(t) + \mathbb{v}(t)$$

where the noise and disturbance are unbiased ($\mathbf{m_v} = \mathbf{0},\ \mathbf{m_w} = \mathbf{0}$) and white

$$\mathcal{A}_{\mathbf{v}}(t,\tau) = \mathsf{E}[\mathbb{v}(t)\mathbb{v}(\tau)^{\mathsf{T}}] = \delta(t-\tau)\mathbf{R}(t), \quad \mathcal{A}_{\mathbf{w}}(t,\tau) = \mathsf{E}[\mathbb{w}(\tau)\mathbb{w}(t)^{\mathsf{T}}] = \delta(t-\tau)\mathbf{Q}(t)$$

Optimal estimator[2]

An optimal estimator for the LTV process is given by

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}(t)\hat{\mathbf{x}}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{L}(t)(\mathbf{y}(t) - \mathbf{C}(t)\hat{\mathbf{x}}(t)), \quad \mathbf{L}(t) = \mathbf{P}(t)\mathbf{C}^{\mathsf{T}}(t)\mathbf{R}_{\mathbf{v}}^{-1}(t)$$

The covariance matrix is here computed by solving the Riccati Equation

$$\dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}(t)^{\mathsf{T}} + \mathbf{G}(t)\mathbf{Q}_{\mathbf{w}}(t)\mathbf{G}(t)^{\mathsf{T}} - \mathbf{P}(t)\mathbf{C}(t)^{\mathsf{T}}\mathbf{R}_{\mathbf{v}}(t)^{-1}\mathbf{C}(t)\mathbf{P}(t)$$

[2] Simulation of the continuous time Riccati equation can be challenging. This is one of the reasons that a discrete time formulation is preferred.

Time varying models

The continuous and discrete time Kalman filters are not limited to time-invariant plants. They can in fact serve as optimal estimators for time-varying systems!

DLTV system

Let a discrete time-varying random process plant model be given by

$$\mathbb{x}[k+1] = \bar{\mathbf{A}}[k]\mathbb{x}[k] + \bar{\mathbf{B}}[k]\mathbf{u}[k] + \bar{\mathbb{w}}[k], \quad \mathbb{y}[k] = \mathbf{C}[k]\mathbb{x}[k] + \bar{\mathbb{v}}[k]$$

where the noise and disturbance are unbiased ($\mathbf{m_v} = \mathbf{0},\ \mathbf{m_w} = \mathbf{0}$) and white

$$\bar{\mathcal{A}}_{\mathbf{v}}[k,l] = \mathrm{E}[\bar{\mathbb{v}}[k]\bar{\mathbb{v}}[l]^{\mathsf{T}}] = \delta[k,l]\bar{\mathbf{R}}_{\mathbf{v}}[k], \quad \bar{\mathcal{A}}_{\mathbf{w}}[k,l] = \mathrm{E}[\bar{\mathbb{w}}[k]\bar{\mathbb{w}}[l]^{\mathsf{T}}] = \delta[k,l]\bar{\mathbf{Q}}_{\mathbf{w}}[k]$$

Optimal estimator

The optimal estimator for the preceding system is furnished, quite simply, by letting the matrices in the Kalman filter algorithm be time-varying.

Kalman filter algorithm, general case

The filter is initialized at

$$\hat{\mathbf{x}}^-[0] = \mathrm{E}[\mathbb{x}(0)] = \mathbf{m}_{\mathbf{x}_0}$$
$$\mathbf{P}^-[0] = \mathrm{E}[\mathbb{e}^-[0]\mathbb{e}^-[0]^\mathsf{T}] = \mathrm{E}[(\mathbb{x}[0] - \mathbf{m}_{\mathbf{x}_0})(\mathbb{x}[0] - \mathbf{m}_{\mathbf{x}_0})^\mathsf{T}] = \mathcal{C}_{\mathbf{x}_0}$$

The recursive algorithm running over $k = 0 \dots K$ is summarized by

1 - Compute Kalman gain

$$\mathbf{L}[k] = \mathbf{P}^-[k]\mathbf{C}[k]^\mathsf{T}(\mathbf{C}[k]\mathbf{P}^-[k]\mathbf{C}[k]^\mathsf{T} + \bar{\mathbf{R}}_\mathbf{v}[k])^{-1}$$

---

2 - Update estimate with measurement

$$\hat{\mathbf{x}}[k] = \hat{\mathbf{x}}^-[k] + \mathbf{L}[k](\mathbf{y}[k] - \mathbf{C}[k]\hat{\mathbf{x}}^-[k])$$

---

3 - Update error covariance matrix

$$\mathbf{P}[k] = (\mathbb{I} - \mathbf{L}[k]\mathbf{C}[k])\mathbf{P}^-[k](\mathbb{I} - \mathbf{L}[k]\mathbf{C}[k])^\mathsf{T} + \mathbf{L}[k]\bar{\mathbf{R}}_\mathbf{v}[k]\mathbf{L}[k]^\mathsf{T}$$

---

4 - Project ahead

$$\hat{\mathbf{x}}^-[k+1] = \bar{\mathbf{A}}[k]\hat{\mathbf{x}}[k] + \bar{\mathbf{B}}[k]\mathbf{u}[k]$$
$$\mathbf{P}^-[k+1] = \bar{\mathbf{A}}[k]\mathbf{P}[k]\bar{\mathbf{A}}[k]^\mathsf{T} + \bar{\mathbf{Q}}_\mathbf{w}[k]$$

...repeat with $k = k + 1$...