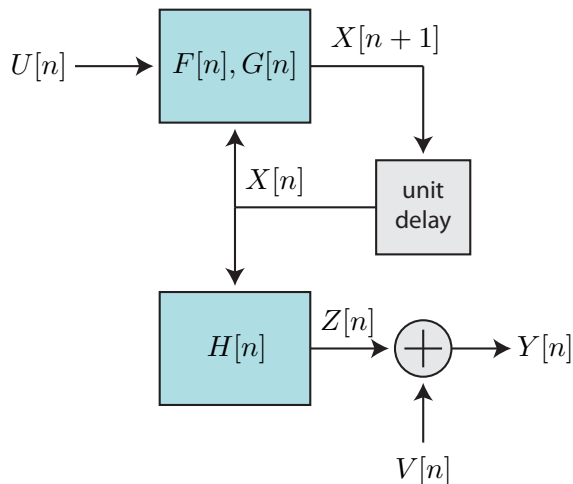


ECE531 Screencast 7.3: The Kalman Filter

D. Richard Brown III

Worcester Polytechnic Institute

Dynamic MMSE State Estimation (Filtering)



We will focus on the **filtering** problem, i.e. estimating $X[\ell]$ given the observations $Y[0], \dots, Y[\ell]$. We know that the MMSE state estimator is the conditional mean

$$\hat{X}[\ell | \ell] := E \{ X[\ell] | \mathcal{Y}_0^\ell \}$$

where

$$\mathcal{Y}_0^\ell := [Y^\top[0], \dots, Y^\top[\ell]]^\top$$

is the "super vector" of all observations up to and including time ℓ .

MMSE State Prediction $\hat{X}[0]$ given no observations

Kalman's big idea was to find a **computationally efficient recursion** for MMSE state estimation. To see how this works, we need to approach the MMSE estimation problem from the beginning.

THE BEGINNING: We want to estimate the initial state $X[0]$ before we have any observations. What is the best estimator of the initial state?

Recall that we have assumed a Gaussian distributed initial state $X[0] \sim \mathcal{N}(m[0], \Sigma[0])$ where $m[0]$ and $\Sigma[0]$ are known. Prior to receiving the first observation, the MMSE estimate (prediction) of the initial state $X[0]$ is simply

$$\hat{X}[0 | -1] := \mathbb{E} \{X[0] \mid \text{no observations}\} = m[0].$$

The error covariance matrix of this MMSE estimator (predictor) is then

$$\begin{aligned} \Sigma[0 | -1] &:= \mathbb{E} \left\{ \left(\hat{X}[0 | -1] - X[0] \right) \left(\hat{X}[0 | -1] - X[0] \right)^\top \right\} \\ &= \mathbb{E} \left\{ \left(m[0] - X[0] \right) \left(m[0] - X[0] \right)^\top \right\} = \Sigma[0] \end{aligned}$$

MMSE State Estimate $\hat{X}[0]$ given $Y[0]$ (part 1)

At time $n = 0$, we receive the observation

$$Y[0] = H[0]X[0] + V[0]$$

where $X[0] \sim \mathcal{N}(m[0], \Sigma[0])$, $V[0] \sim \mathcal{N}(0, R[0])$, and $H[0]$ is known.

We use the standard result for jointly Gaussian random vectors with

$$\mathbb{E}[X] = m[0] \qquad \mathbb{E}[Y] = H[0]m[0]$$

$$\text{cov}(X, Y) = \Sigma[0]H^{\top}[0] \qquad \text{cov}(Y, Y) = H[0]\Sigma[0]H^{\top}[0] + R[0]$$

to write an expression for the MMSE estimate of $X[0]$ given $Y[0]$ as

$$\begin{aligned} \hat{X}[0|0] &:= \mathbb{E}\{X[0] | Y[0]\} \\ &= m[0] + \Sigma[0]H^{\top}[0] \left(H[0]\Sigma[0]H^{\top}[0] + R[0] \right)^{-1} (Y[0] - H[0]m[0]) \\ &= \hat{X}[0|-1] + \\ &\quad \Sigma[0|-1]H^{\top}[0] \left(H[0]\Sigma[0|-1]H^{\top}[0] + R[0] \right)^{-1} (Y[0] - H[0]\hat{X}[0|-1]) \end{aligned}$$

MMSE State Estimate $\hat{X}[0]$ given $Y[0]$ (part 2)

If we define the **Kalman Gain** as

$$\begin{aligned} K[0] &:= \text{cov}(X[0], Y[0]) [\text{cov}(Y[0], Y[0])]^{-1} \\ &= \Sigma[0 | -1] H^\top[0] (H[0] \Sigma[0 | -1] H^\top[0] + R[0])^{-1} \end{aligned}$$

we can write

$$\hat{X}[0 | 0] = \hat{X}[0 | -1] + K[0] (Y[0] - H[0] \hat{X}[0 | -1])$$

Remarks:

- ▶ The term $\tilde{Y}[0 | -1] := Y[0] - H[0] \hat{X}[0 | -1]$ is sometimes called the “innovation”. It is the error between the MMSE **prediction** of $Y[0]$ and the actual observation $Y[0]$.
- ▶ Substituting for $Y[0]$, note that the innovation can be written as

$$\begin{aligned} \tilde{Y}[0 | -1] &= (H[0]X[0] + V[0] - H[0]\hat{X}[0 | -1]) \\ &= H[0] (X[0] - \hat{X}[0 | -1]) + V[0] \end{aligned}$$

MMSE State Estimate $\hat{X}[0]$ given $Y[0]$ (part 3)

The error covariance matrix of the MMSE estimator $\hat{X}[0|0]$ can be computed as

$$\begin{aligned}\Sigma[0|0] &:= \mathbb{E} \left\{ \left(\hat{X}[0|0] - X[0] \right) \left(\hat{X}[0|0] - X[0] \right)^\top \mid Y[0] \right\} \\ &= \text{cov} \{ X[0] \mid Y[0] \}\end{aligned}$$

We can use the standard result for jointly Gaussian random vectors to write

$$\begin{aligned}\Sigma[0|0] &= \Sigma[0] - \Sigma[0]H^\top[0] \left(H[0]\Sigma[0]H^\top[0] + R[0] \right)^{-1} H[0]\Sigma[0] \\ &= \Sigma[0| -1] - K[0]H[0]\Sigma[0| -1]\end{aligned}$$

where we used our definitions for $\Sigma[0| -1]$ and $K[0]$ in the last equality. Note that, given $K[0]$ and $H[0]$, the error covariance matrix of the MMSE state estimator after the observation $Y[0]$ is only a function of the error covariance matrix of the prediction $\Sigma[0| -1]$.

Remarks

What we have done so far:

1. Predicted the first state with no observations: $\hat{X}[0 | -1]$.
2. Computed the error covariance matrix of this prediction: $\Sigma[0 | -1]$.
3. Received the observation $Y[0]$.
4. Estimated the first state given the observation: $\hat{X}[0 | 0]$.
5. Computed the error covariance matrix of this estimate: $\Sigma[0 | 0]$.

Interesting observations:

- ▶ The **estimate** $\hat{X}[0 | 0]$ is expressed in terms of the **prediction** $\hat{X}[0 | -1]$ and the observation $Y[0]$.
- ▶ The **estimate error covariance matrix** $\Sigma[0 | 0]$ is expressed in terms of the **prediction error covariance matrix** $\Sigma[0 | -1]$.

Along similar lines, we can show the **prediction** $\hat{X}[1 | 0]$ can be expressed in terms of the **estimate** $\hat{X}[0 | 0]$ and the **prediction error covariance** $\Sigma[1 | 0]$ can be expressed in terms of the **estimation error covariance matrix** $\Sigma[0 | 0]$. These steps can be repeated to develop a general recursion.

The Discrete-Time Kalman-Bucy Filter (1961)

Theorem

Under the squared error cost assignment, the linear system model, and the white Gaussian input, noise, and initial state assumptions discussed previously, the optimal estimates for the current state (filtering) and the next state (prediction) are given recursively as

$$\hat{X}[\ell | \ell] = \hat{X}[\ell | \ell - 1] + K[\ell] \left(Y[\ell] - H[\ell] \hat{X}[\ell | \ell - 1] \right) \text{ for } \ell = 0, 1, \dots$$

$$\hat{X}[\ell + 1 | \ell] = F[\ell] \hat{X}[\ell | \ell] \text{ for } \ell = 0, 1, \dots$$

with the initialization $\hat{X}[0 | -1] = m[0]$ and where the matrix

$$K[\ell] = \Sigma[\ell | \ell - 1] H^{\top}[\ell] \left(H[\ell] \Sigma[\ell | \ell - 1] H^{\top}[\ell] + R[\ell] \right)^{-1}$$

with $\Sigma[\ell | \ell - 1] := \text{cov} \{ X[\ell] | Y[0], \dots, Y[\ell - 1] \}$ and $R[\ell] := \text{cov} \{ V[\ell] \}$.

Kalman Filter: Summary of General Recursion

Initialization (predictions):

$$\begin{aligned}\hat{X}[0 | -1] &= m[0] \\ \Sigma[0 | -1] &= \Sigma[0]\end{aligned}$$

Recursion, beginning with $\ell = 0$:

$$\begin{aligned}K[\ell] &= \Sigma[\ell | \ell - 1]H^T[\ell] \left(H[\ell]\Sigma[\ell | \ell - 1]H^T[\ell] + R[\ell] \right)^{-1} \\ \hat{X}[\ell | \ell] &= \hat{X}[\ell | \ell - 1] + K[\ell] \left(Y[\ell] - H[\ell]\hat{X}[\ell | \ell - 1] \right) \\ \Sigma[\ell | \ell] &= \Sigma[\ell | \ell - 1] - K[\ell]H[\ell]\Sigma[\ell | \ell - 1] \\ \hat{X}[\ell + 1 | \ell] &= F[\ell]\hat{X}[\ell | \ell] \\ \Sigma[\ell + 1 | \ell] &= F[\ell]\Sigma[\ell | \ell]F[\ell]^T + G[\ell]Q[\ell]G^T[\ell]\end{aligned}$$