E9 211: Adaptive Signal Processing

Kalman Filter

The discrete Kalman filter

Consider the following nonstationary *state space model*:

$$\begin{aligned} \mathbf{x}(n) &= \mathbf{A}(n-1)\mathbf{x}(n-1) + \mathbf{w}(n) \\ \mathbf{y}(n) &= \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n) \end{aligned}$$

where $\mathbf{x}(n)$ is the $p \times 1$ state vector, $\mathbf{A}(n-1)$ is the $p \times p$ state transistion matrix, $\mathbf{w}(n)$ is the state noise with $E\{\mathbf{w}(n)\mathbf{w}^H(n)\} = \mathbf{Q}_w(n)\delta(n-k)$, $\mathbf{y}(n)$ is the $q \times 1$ observation vector, $\mathbf{C}(n)$ is the $q \times p$ observation matrix, $\mathbf{v}(n)$ is the observation noise with $E\{\mathbf{v}(n)\mathbf{v}^H(k)\} = \mathbf{Q}_v(n)\delta(n-k)$, and the state noise is independent of the observation noise.

We will derive the best linear estimate of $\mathbf{x}(n)$ for observations $\mathbf{y}(n)$ up to n using a weighted least squares formulation.

Weighted least squares

Consider the linear measurement model

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{v}$$

where $\mathbf{C}$ is the $q \times p$ observation matrix ($q \geq p$), $\mathbf{x}$ is the $p \times 1$ unknown vector, $\mathbf{v}$ is colored noise with $E\{\mathbf{v}\mathbf{v}^H\} = \mathbf{Q}_v$.
To determine the optimal estimator of $\mathbf{x}$, i.e., a *minimum variance unbiased* (MVU) estimator, we use a *whitening* approach and transform the above model to

$$\mathbf{Q}_v^{-1/2}\mathbf{y} = \mathbf{Q}_v^{-1/2}\mathbf{C}\mathbf{x} + \mathbf{Q}_v^{-1/2}\mathbf{v}$$

such that the noise will be whitened as $E\{\mathbf{Q}_v^{-1/2}\mathbf{v}\mathbf{v}^H\mathbf{Q}_v^{-1/2}\} = \mathbf{I}$.
The MVU estimator of $\mathbf{x}$ is then the usual least squares estimator (based on the transformed model), i.e.,

$$\hat{\mathbf{x}} = \left(\mathbf{Q}_v^{-1/2}\mathbf{C}\right)^{\dagger}\mathbf{Q}_v^{-1/2}\mathbf{y} = \left(\mathbf{C}^H\mathbf{Q}_v^{-1/2}\mathbf{Q}_v^{-1/2}\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{Q}_v^{-1/2}\mathbf{Q}_v^{-1/2}\mathbf{y}$$

so that

$$\hat{\mathbf{x}} = \left(\mathbf{C}^H\mathbf{Q}_v^{-1}\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{Q}_v^{-1}\mathbf{y}.$$

The discrete Kalman filter

Let us define $\hat{\mathbf{x}}(n|n-1)$ and $\hat{\mathbf{x}}(n|n)$ as the best linear estimate of $\mathbf{x}(n)$ given the observations $\mathbf{y}(n)$ up to time $n-1$ and n, respectively.
Let us denote the corresponding errors as

$$\mathbf{e}(n|n-1) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1)$$
$$\mathbf{e}(n|n) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|n)$$

with covariance matrices

$$\mathbf{P}(n|n-1) = E\{\mathbf{e}(n|n-1)\mathbf{e}^H(n|n-1)\}$$
$$\mathbf{P}(n|n) = E\{\mathbf{e}(n|n)\mathbf{e}^H(n|n)\}$$

We can now derive the discrete Kalman filter.

Prediction stage

Given the state estimate $\hat{\mathbf{x}}(n-1|n-1)$ at time $n-1$, we can compute the prediction as

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n-1)\hat{\mathbf{x}}(n-1|n-1)$$

The prediction error will then be

$$\begin{aligned}
\mathbf{e}(n|n-1) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1) &= \mathbf{A}(n-1)\mathbf{x}(n-1) + \mathbf{w}(n) - \mathbf{A}(n-1)\hat{\mathbf{x}}(n - \\
&= \mathbf{A}(n-1)\left[\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1|n-1)\right] + \mathbf{w}(\\
&= \mathbf{A}(n-1)\mathbf{e}(n-1|n-1) + \mathbf{w}(n)
\end{aligned}$$

with the covariance matrix

$$\mathbf{P}(n|n-1) = \mathbf{A}(n-1)\mathbf{P}(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n).$$

Correction stage

We can rewrite the estimate $\hat{\mathbf{x}}(n|n-1)$ as follows

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{x}(n) + \mathbf{e}(n|n-1).$$

Augmenting the above system with $\mathbf{y}(n)$ we have

$$\begin{bmatrix} \hat{\mathbf{x}}(n|n-1) \\ \mathbf{y}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{C}(n) \end{bmatrix} \mathbf{x}(n) + \begin{bmatrix} \mathbf{e}(n|n-1) \\ \mathbf{v}(n) \end{bmatrix},$$

with the covariance matrix of the augmented noise vector being

$$E\left\{ \begin{bmatrix} \mathbf{e}(n|n-1) \\ \mathbf{v}(n) \end{bmatrix} \begin{bmatrix} \mathbf{e}(n|n-1) \\ \mathbf{v}(n) \end{bmatrix}^H \right\} = \begin{bmatrix} \mathbf{P}(n|n-1) & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_v(n) \end{bmatrix}.$$

Correction stage

The estimate $\hat{\mathbf{x}}(n|n)$ can be obtained by solving the augmented system of equations using the weighted least squares approach:

$$\hat{\mathbf{x}}(n|n) = \left(\left[\begin{array}{c|c} \mathbf{I} & \mathbf{C}^H(n) \end{array}\right] \left[\begin{array}{c|c} \mathbf{P}^{-1}(n|n-1) & \\ \hline & \mathbf{Q}_v^{-1}(n) \end{array}\right] \left[\begin{array}{c} \mathbf{I} \\ \hline \mathbf{C}(n) \end{array}\right] \right)^{-1}$$

$$\times \left[\begin{array}{c|c} \mathbf{I} & \mathbf{C}^H(n) \end{array}\right] \left[\begin{array}{c|c} \mathbf{P}^{-1}(n|n-1) & \\ \hline & \mathbf{Q}_v^{-1}(n) \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{x}}(n|n-1) \\ \hline \mathbf{y}(n) \end{array}\right]$$

$$= \left[\mathbf{P}^{-1}(n|n-1) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)\right]^{-1}$$
$$\times \left[\mathbf{P}^{-1}(n|n-1)\hat{\mathbf{x}}(n|n-1) + \mathbf{C}^H(n)\mathbf{Q}_v^{-1}(n)\mathbf{y}(n)\right]$$

Correction stage

Using the matrix inversion lemma

$$(\mathbf{A}+\mathbf{B}^H\mathbf{C}^{-1}\mathbf{B})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}^H(\mathbf{C}+\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^H)^{-1}\mathbf{B}\mathbf{A}^{-1}$$

we have

$$\begin{aligned}\left[\mathbf{P}^{-1}(n|n-1)+\mathbf{C}(n)^H\mathbf{Q}_v^{-1}(n)\mathbf{C}(n)\right]^{-1} &= \mathbf{P}(n|n-1) - \mathbf{P}(n|n-1)\mathbf{C}^H(n)\\ &\times(\mathbf{Q}_v(n)+\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n))^{-1}\mathbf{C}\mathbf{P}(n\end{aligned}$$

and introducing the *Kalman gain* matrix, $\mathbf{K}(n)$, as

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{Q}_v(n)+\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)]^{-1}$$

we can obtain the recursion for computing $\hat{\mathbf{x}}(n|n)$ as

$$\begin{aligned}\hat{\mathbf{x}}(n|n) &= \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)\left[\mathbf{y}(n)-\mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)\right]\\ &= \left[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)\right]\hat{\mathbf{x}}(n|n-1)+\mathbf{K}(n)\mathbf{y}(n)\end{aligned}$$

Correction stage

The error at the correction stage will then be

$$\begin{aligned}\mathbf{e}(n|n) &= \mathbf{x}(n) - \hat{\mathbf{x}}(n|n) \\ &= \mathbf{x}(n) - [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\hat{\mathbf{x}}(n|n-1) - \mathbf{K}(n)\mathbf{y}(n).\end{aligned}$$

Substituting the expression for $\mathbf{y}(n)$ we get

$$\begin{aligned}\mathbf{e}(n|n) &= \mathbf{x}(n) - [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\hat{\mathbf{x}}(n|n-1) - \mathbf{K}(n)\,[\mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n)] \\ &= [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1)) - \mathbf{K}(n)\mathbf{v}(n) \\ &= [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{e}(n|n-1) - \mathbf{K}(n)\mathbf{v}(n).\end{aligned}$$

Correction stage

The error covariance matrix $\mathbf{P}(n|n)$ can be computed as

$$\begin{aligned}\mathbf{P}(n|n) &= [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\,[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]^H + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}^H(n) \\ &= [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1) - \big[[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}^H(n) \\ &\qquad +\mathbf{K}(n)\mathbf{Q}_v(n)]\,\mathbf{K}^H(n).\end{aligned}$$

By multiplying both sides of the Kalman gain matrix

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{Q}_v(n)+\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)]^{-1}$$

with $[\mathbf{Q}_v(n)+\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n]\mathbf{K}^H(n)$, it is easy to see that

$$\big[[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}^H(n)+\mathbf{K}(n)\mathbf{Q}_v(n)\big]\,\mathbf{K}^H(n) = \mathbf{0}.$$

So the error covariance matrix $\mathbf{P}(n|n)$ can be recursively updated using

$$\mathbf{P}(n|n) = [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1).$$

The discrete Kalman filter

Initializing the recursion with

$$\hat{\mathbf{x}}(0|0) = E\{\mathbf{x}(0)\} \text{ and } \hat{\mathbf{P}}(0|0) = E\{\mathbf{x}(0)\mathbf{x}^H(0)\}$$

we obtain the following optimal recursion for $n = 1, 2, \ldots$ at the Prediction stage:

$$\begin{aligned}\hat{\mathbf{x}}(n|n-1) &= \mathbf{A}(n-1)\hat{\mathbf{x}}(n-1|n-1) \\ \mathbf{P}(n|n-1) &= \mathbf{A}(n-1)\mathbf{P}(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n)\end{aligned}$$

Correction stage:

$$\begin{aligned}\mathbf{K}(n) &= \mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{Q}_v(n) + \mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)]^{-1} \\ \hat{\mathbf{x}}(n|n) &= \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)\left[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)\right] \\ \mathbf{P}(n|n) &= \left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)\end{aligned}$$

Note that $\mathbf{P}(n|n-1)$, $\mathbf{K}(n)$, and $\mathbf{P}(n|n)$ are independent of the observations $\mathbf{y}(n)$ and thus can be computed off-line prior to any filtering.

Filtering example

We consider the following noisy measurement model

$$y(n) = x(n) + v(n)$$

where $v(n)$ is white noise with variance $\sigma_v^2 = 0.1$. Suppose $x(n)$ is an AR(1) process given by

$$x(n) = 0.8x(n-1) + w(n)$$

where $w(n)$ is white noise with variance $\sigma_w^2 = 0.36$. Thus, with $\mathbf{A}(n) = 0.8$, $\mathbf{C}(n) = 1$, $\mathbf{Q}_w = 0.36$, and $\mathbf{Q}_v = 0.1$, the Kalman filter state estimation equation is

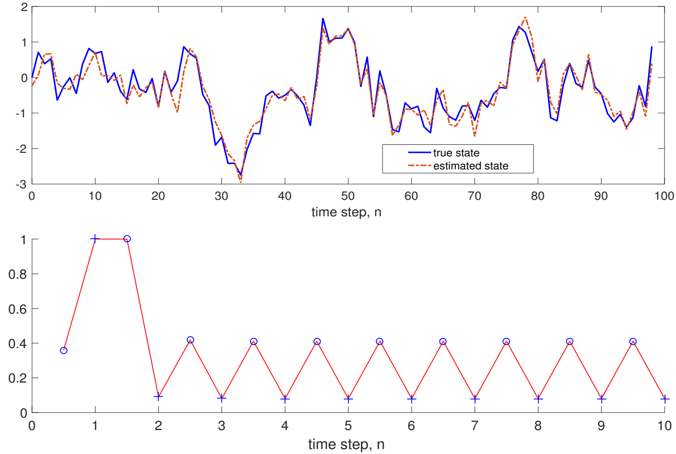
$$\hat{x}(n|n) = 0.8\hat{x}(n-1|n-1) + K(n)[y(n) - 0.8\hat{x}(n-1|n-1)].$$

For the scalar state, the equations for updating the error covariance matrices are

$$\begin{aligned} P(n|n-1) &= (0.8)^2 P(n-1|n-1) + 0.36 \\ K(n) &= P(n|n-1)[P(n|n-1) + 0.1]^{-1} \\ P(n|n) &= [1 - K(n)]P(n|n-1). \end{aligned}$$

Filtering example

With $\hat{x}(0|0) = 0$ and $P(0|0) = 1$



$\circ$ denotes the prediction error and $+$ denotes the correction error. The prediction stage increases the error, while the correction stage decreases it. After a few iterations the error settles down into its steady state.