E9 211: Adaptive Signal Processing

Kalman Filter
The discrete Kalman filter

Consider the following nonstationary state space model:

\[
\begin{align*}
    x(n) &= A(n-1)x(n-1) + w(n) \\
    y(n) &= C(n)x(n) + v(n)
\end{align*}
\]

where \( x(n) \) is the \( p \times 1 \) state vector, \( A(n-1) \) is the \( p \times p \) state transition matrix, \( w(n) \) is the state noise with 
\[
E\{w(n)w^H(n)\} = Q_w(n)\delta(n-k),
\]
\( y(n) \) is the \( q \times 1 \) observation vector, \( C(n) \) is the \( q \times p \) observation matrix, \( v(n) \) is the observation noise with 
\[
E\{v(n)v^H(k)\} = Q_v(n)\delta(n-k),
\]
and the state noise is independent of the observation noise.

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

Consider the linear measurement model

\[ y = Cx + v \]

where \( C \) is the \( q \times p \) observation matrix (\( q \geq p \)), \( x \) is the \( p \times 1 \) unknown vector, \( v \) is colored noise with \( E\{vv^H\} = Q_v \).

To determine the optimal estimator of \( x \), i.e., a minimum variance unbiased (MVU) estimator, we use a whitening approach and transform the above model to

\[ Q_{v}^{-1/2}y = Q_{v}^{-1/2}Cx + Q_{v}^{-1/2}v \]

such that the noise will be whitened as \( E\{Q_{v}^{-1/2}vv^HQ_{v}^{-1/2}\} = I \).

The MVU estimator of \( x \) is then the usual least squares estimator (based on the transformed model), i.e.,

\[ \hat{x} = \left(Q_{v}^{-1/2}C\right)^{\dagger}Q_{v}^{-1/2}y = \left(C^HQ_{v}^{-1/2}Q_{v}^{-1/2}C\right)^{-1}C^HQ_{v}^{-1/2}Q_{v}^{-1/2}y \]

so that

\[ \hat{x} = \left(C^HQ_{v}^{-1}C\right)^{-1}C^HQ_{v}^{-1}y. \]
Let us define \( \hat{x}(n|n-1) \) and \( \hat{x}(n|n) \) as the best linear estimate of \( x(n) \) given the observations \( y(n) \) up to time \( n-1 \) and \( n \), respectively. Let us denote the corresponding errors as

\[
e(n|n-1) = x(n) - \hat{x}(n|n-1)
\]

\[
e(n|n) = x(n) - \hat{x}(n|n)
\]

with covariance matrices

\[
P(n|n-1) = E\{e(n|n-1)e^H(n|n-1)\}
\]

\[
P(n|n) = E\{e(n|n)e^H(n|n)\}
\]

We can now derive the discrete Kalman filter.
Prediction stage

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

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

The prediction error will then be

$$e(n|n-1) = x(n) - \hat{x}(n|n-1) = A(n-1)x(n-1) + w(n) - A(n-1)\hat{x}(n-1|n-1)$$

$$= A(n-1)[x(n-1) - \hat{x}(n-1|n-1)] + w(n)$$

$$= A(n-1)e(n-1|n-1) + w(n)$$

with the covariance matrix

$$P(n|n-1) = A(n-1)P(n-1|n-1)A^H(n-1) + Q_w(n).$$
We can rewrite the estimate $\hat{x}(n|n-1)$ as follows

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

Augmenting the above system with $y(n)$ we have

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

with the covariance matrix of the augmented noise vector being

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

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

$$
\hat{x}(n|n) = \left( \left[ \begin{array}{c|c} I & CH(n) \end{array} \right] \left[ \begin{array}{c|c} P^{-1}(n|n-1) & Q_v^{-1}(n) \end{array} \right] \left[ \begin{array}{c} I \\ C(n) \end{array} \right] \right)^{-1}
$$

$$
\times \left[ \begin{array}{c|c} I & CH(n) \end{array} \right] \left[ \begin{array}{c|c} P^{-1}(n|n-1) & Q_v^{-1}(n) \end{array} \right] \left[ \begin{array}{c} \hat{x}(n|n-1) \\ y(n) \end{array} \right]
$$

$$
= \left[ P^{-1}(n|n-1) + CH(n)Q_v^{-1}(n)C(n) \right]^{-1}
\times \left[ P^{-1}(n|n-1)\hat{x}(n|n-1) + CH(n)Q_v^{-1}(n)y(n) \right]
$$
Correction stage

Using the matrix inversion lemma

\[(A + B^H C^{-1} B)^{-1} = A^{-1} - A^{-1} B^H (C + B A^{-1} B^H)^{-1} B A^{-1}\]

we have

\[
[P^{-1}(n|n-1) + C(n)^H Q_v^{-1}(n) C(n)]^{-1} = P(n|n-1) - P(n|n-1) C^H(n) \\
\times (Q_v(n) + C(n) P(n|n-1) C^H(n))^{-1} C P(n)
\]

and introducing the *Kalman gain* matrix, \(K(n)\), as

\[
K(n) = P(n|n-1) C^H(n) [Q_v(n) + C(n) P(n|n-1) C^H(n)]^{-1}
\]

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

\[
\hat{x}(n|n) = \hat{x}(n|n-1) + K(n) [y(n) - C(n) \hat{x}(n|n-1)] \\
= [I - K(n) C(n)] \hat{x}(n|n-1) + K(n) y(n)
\]
The error at the correction stage will then be

\[ e(n|n) = x(n) - \hat{x}(n|n) \]

\[ = x(n) - [I - K(n)C(n)] \hat{x}(n|n - 1) - K(n)y(n). \]

Substituting the expression for \( y(n) \) we get

\[ e(n|n) = x(n) - [I - K(n)C(n)] \hat{x}(n|n - 1) - K(n) [C(n)x(n) + v(n)] \]

\[ = [I - K(n)C(n)] (x(n) - \hat{x}(n|n - 1)) - K(n)v(n) \]

\[ = [I - K(n)C(n)] e(n|n - 1) - K(n)v(n). \]
Correction stage

The error covariance matrix \( P(n|n) \) can be computed as

\[
P(n|n) = [I - K(n)C(n)] P(n|n - 1) [I - K(n)C(n)]^H + K(n)Q_v(n)K^H(n)
\]

\[
= [I - K(n)C(n)] P(n|n - 1) - [I - K(n)C(n)] P(n|n - 1)C^H(n)
\]

\[
+ K(n)Q_v(n)K^H(n).
\]

By multiplying both sides of the Kalman gain matrix

\[
K(n) = P(n|n - 1)C^H(n)[Q_v(n) + C(n)P(n|n - 1)C^H(n)]^{-1}
\]

with \([Q_v(n) + C(n)P(n|n - 1)C^H(n)K^H(n)]\), it is easy to see that

\[
[[I - K(n)C(n)] P(n|n - 1)C^H(n) + K(n)Q_v(n)] K^H(n) = 0.
\]

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

\[
P(n|n) = [I - K(n)C(n)] P(n|n - 1).
\]
The discrete Kalman filter

Initializing the recursion with

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

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

\[
\begin{align*}
\hat{x}(n|n-1) &= A(n-1)\hat{x}(n-1|n-1) \\
\hat{P}(n|n-1) &= A(n-1)\hat{P}(n-1|n-1)A^H(n-1) + Q_w(n)
\end{align*}
\]

Correction stage:

\[
\begin{align*}
K(n) &= \hat{P}(n|n-1)C^H(n)[Q_v(n) + C(n)\hat{P}(n|n-1)C^H(n)]^{-1} \\
\hat{x}(n|n) &= \hat{x}(n|n-1) + K(n) [y(n) - C(n)\hat{x}(n|n-1)] \\
\hat{P}(n|n) &= [I - K(n)C(n)] \hat{P}(n|n-1)
\end{align*}
\]

Note that \( \hat{P}(n|n-1) \), \( K(n) \), and \( \hat{P}(n|n) \) are independent of the observations \( y(n) \) 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 \( A(n) = 0.8, C(n) = 1, Q_w = 0.36, \) and \( Q_v = 0.1 \), 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

\[ P(n|n - 1) = (0.8)^2P(n - 1|n - 1) + 0.36 \]
\[ K(n) = P(n|n - 1)[P(n|n - 1) + 0.1]^{-1} \]
\[ P(n|n) = [1 - K(n)]P(n|n - 1). \]
Filtering example

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

○ 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.