Discrete-Time Optimal Filtering

The linear system:

$$\underline{x}(n+1) = \underline{\Phi}(n)\,\underline{x}(n) + \underline{G}(n)\,\underline{w}(n)$$

$$\underline{z}(n) = \underline{H}(n)\,\underline{x}(n) + \underline{v}(n)$$

$$E\{\underline{x}(0)\} = E\{\underline{w}(n)\} = E\{\underline{v}(n)\} = \underline{0}$$

$$E\{\underline{x}(0)\underline{x}^T(0)\} = \underline{\Psi} \qquad E\{\underline{w}(n)\underline{w}^T(m)\} = \underline{Q}(n)\,\delta(n-m)$$

$$E\{\underline{v}(n)\underline{v}^T(m)\} = \underline{R}(n)\,\delta(n-m)$$

$\underline{x}(0)$, $\underline{w}(n)$, $\underline{v}(m)$ all independent Gaussian r.v.

Notation:

Let $\hat{\underline{x}}(n|m)$ = "best" estimate of $\underline{x}(n)$ given $\underline{z}(1), \ldots, \underline{z}(m)$ + a priori information

and

$$\underline{\Sigma}(n|m) = E\{(\underline{x}(n) - \hat{\underline{x}}(n|m))(\underline{x}(n) - \hat{\underline{x}}(n|m))^T\}$$

Using the optimal (static) linear estimator, we can compute $\hat{\underline{x}}(n|m)$ from $\mathcal{Y}_m = \{\underline{z}(k);\ k \le m\}$.

However, it is easier to implement a <u>recursive</u> estimator and compute $\hat{\underline{x}}(n|n)$ from $\hat{\underline{x}}(n-1|n-1)$ and $\underline{z}(n)$.

Assume an <u>a priori</u> estimate (before $z(n)$ is used) of $\underline{x}(n)$, $\hat{\underline{x}}(n|n-1)$, is known & is based upon all measurements through $n-1$, and the covariance of the estimation error is $\underline{\Sigma}(n|n-1)$.

<u>Desired</u>: A map from $\hat{\underline{x}}(n|n-1)$ and $\underline{z}(n) \longrightarrow \hat{\underline{x}}(n|n)$

" " " $\Sigma(n|n-1) \longrightarrow$ & $\underline{\Sigma}(n|n)$.

<u>Assume</u> the update map is linear:

$$\hat{\underline{x}}(n|n) = \underline{K}'(n)\,\hat{\underline{x}}(n|n-1) + \underline{K}(n)\,\underline{z}(n)$$

(Note: This assumption is not necessary.)

Define estimation errors before and after update:

$$\hat{\underline{x}}(n|n) = \underline{x}(n) + \tilde{\underline{x}}(n|n)$$

$$\hat{\underline{x}}(n|n-1) = \underline{x}(n) + \tilde{\underline{x}}(n|n-1)$$

and substitute these expressions into the update map:

$$\hat{\underline{x}}(n|n) = \left[\underline{K}'(n) + \underline{K}(n)\underline{H}(n) - \underline{I}\right]\underline{x}(n) + \underline{K}'(n)\,\tilde{\underline{x}}(n|n-1) + \underline{K}(n)\,\underline{v}(n)$$

Require that the estimate be unbiased:

$$E\left[\tilde{\underline{x}}(n|n)\right] = 0 = \left[\underline{K}'(n) + \underline{K}(n)\underline{H}(n) - \underline{I}\right]E\left[\underline{x}(n)\right] + \underline{K}'(n)\,\cancelto{0}{E\left[\tilde{\underline{x}}(n|n-1)\right]} + \underline{K}(n)\,\cancelto{0}{E\left[\underline{v}(n)\right]}$$

So $\underline{K}'(n) = \left[\underline{I} - \underline{K}(n)\underline{H}(n)\right]$ is required for an unbiased update or,

$$\underline{\hat{x}}(n|n) = [\underline{I} - \underline{K}(n)\underline{H}(n)]\,\underline{\hat{x}}(n|n-1) + \underline{K}(n)\underline{z}(n)$$

$$= \underline{\hat{x}}(n|n-1) + \underline{K}(n)\underbrace{[\underline{z}(n) - \underline{H}(n)\underline{\hat{x}}(n|n-1)]}$$

The error between the actual & predicted measurement given a priori information.

The error in the updated estimate is:

$$\underline{\tilde{x}}(n|n) = (\underline{I} - \underline{K}(n)\underline{H}(n))\,\underline{\tilde{x}}(n|n-1) + \underline{K}(n)\underline{v}(n)$$

and the covariance of this error is

$$\underline{\Sigma}(n|n) = E[\underline{\tilde{x}}(n|n)\underline{\tilde{x}}^T(n|n)]$$

$$= (\underline{I} - \underline{K}(n)\underline{H}(n))\,E[\underline{\tilde{x}}(n|n-1)\underline{\tilde{x}}^T(n|n-1)](\underline{I} - \underline{K}(n)\underline{H}(n))^T$$

$$+ (\underline{I} - \underline{K}(n)\underline{H}(n))\,E[\underline{\tilde{x}}(n|n-1)\underline{v}^T(n)]\underline{K}^T(n)$$

$$+ \underline{K}(n)E[\underline{v}(n)\underline{\tilde{x}}^T(n|n-1)](\underline{I} - \underline{K}(n)\underline{H}(n))^T$$

$$+ \underline{K}(n)E[\underline{v}(n)\underline{v}^T(n)]\underline{K}^T(n)$$

Since

$$E[\underline{\tilde{x}}(n|n-1)\underline{\tilde{x}}^T(n|n-1)] = \underline{\Sigma}(n|n-1)$$

$$E[\underline{v}(n)\underline{v}^T(n)] = \underline{R}(n)$$

$$E[\underline{\tilde{x}}(n|n-1)\underline{v}^T(n)] = \underline{0}$$

Then

$$\underline{\Sigma}(n|n) = (\underline{I} - \underline{K}(n)\underline{H}(n))\,\underline{\Sigma}(n|n-1)(\underline{I} - \underline{K}(n)\underline{H}(n))^T + \underline{K}(n)\underline{R}(n)\underline{K}^T(n)$$

Now, choose $\underline{K}(n)$ to minimize $J(n) = \mathrm{tr}[\underline{\Sigma}(n|n)]$

Fact: $\dfrac{\partial}{\partial \underline{A}} \mathrm{tr}[\underline{A}\underline{B}\underline{A}^T] = 2\underline{A}\underline{B}$ for symmetric $\underline{B}$.

$$\therefore \quad -2\,(\underline{I} - \underline{K}(n)\underline{H}(n))\,\underline{\Sigma}(n|n-1)\,\underline{H}^T(n) + 2\,\underline{K}(n)\underline{R}(n) = \underline{0}$$

or

$$\boxed{\underline{K}(n) = \underline{\Sigma}(n|n-1)\,\underline{H}^T(n)\left[\underline{H}(n)\,\underline{\Sigma}(n|n-1)\,\underline{H}^T(n) + \underline{R}(n)\right]^{-1}}$$

This is the discrete time Kalman filter gain. (optimal linear estimator)

and

$$\boxed{\begin{aligned} \underline{\Sigma}(n|n) &= \underline{\Sigma}(n|n-1) - \underline{\Sigma}(n|n-1)\,\underline{H}^T(n)\left[\underline{H}(n)\,\underline{\Sigma}(n|n-1)\,\underline{H}^T(n) + \underline{R}(n)\right]^{-1} \\ &\qquad \times \underline{H}(n)\,\underline{\Sigma}(n|n-1) \\ &= \left[\underline{I} - \underline{K}(n)\,\underline{H}(n)\right]\underline{\Sigma}(n|n-1) \end{aligned}}$$

We have already done the prediction step:

$$\underline{\hat{x}}(n+1|n) = \underline{\Phi}(n)\,\underline{\hat{x}}(n|n)$$

$$\underline{\Sigma}(n+1|n) = \underline{\Phi}(n)\,\underline{\Sigma}(n|n)\,\underline{\Phi}^T(n) + \underline{G}(n)\,\underline{Q}(n)\,\underline{G}^T(n)$$

Summary : The Discrete-Time Kalman Filter

System Model:

$$\underline{x}(n+1) = \underline{\Phi}(n)\underline{x}(n) + \underline{G}(n)\underline{w}(n)$$

$$\underline{w}(n) \sim N(\underline{0}, \underline{Q}(n))$$

Measurement Model:

$$\underline{z}(n) = \underline{H}(n)\underline{x}(n) + \underline{v}(n)$$

$$\underline{v}(n) \sim N(\underline{0}, \underline{R}(n))$$

Initial Conditions: $\hat{\underline{x}}(0) = E[\underline{x}(0)]$, $\underline{\Sigma}(0|0) = E[(x(0)-\hat{x}(0))(x^T(0)-\hat{x}^T(0)$

Other Assumptions: $\underline{x}(0)$, $\underline{w}(n)$, $\underline{v}(m)$ indep. Gaussian r.v. $\forall\, n, m$

Prediction Step:

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

$$\underline{\Sigma}(n|n-1) = \underline{\Phi}(n-1)\underline{\Sigma}(n-1|n-1)\underline{\Phi}^T(n-1) + \underline{G}(n-1)\underline{Q}(n-1)\underline{G}^T(n-1)$$

Update Step:

$$\hat{\underline{x}}(n|n) = \hat{\underline{x}}(n|n-1) + \underline{K}(n)\left[\underline{z}(n) - \underline{H}(n)\hat{\underline{x}}(n|n-1)\right]$$

$$\underline{\Sigma}(n|n) = \left[\underline{I} - \underline{K}(n)\underline{H}(n)\right]\underline{\Sigma}(n|n-1)$$

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

A Simpler Form:

$$\underline{\Sigma}^{-1}(n|n) = \underline{\Sigma}^{-1}(n|n-1) + \underline{H}^T(n)\underline{R}^{-1}(n)\underline{H}(n)$$

$$\underline{K}(n) = \underline{\Sigma}(n|n)\underline{H}^T(n)\underline{R}^{-1}(n)$$

A trivial example:

$$x(n+1) = 0.95\,x(n) + w(n)$$

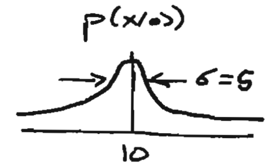
$$E[w^2(n)] = 1$$

$$E[w(n)] = 0$$

$$E[x(0)] = 10$$

$$E[\tilde{x}^2(0)] = 25$$

$$\Longrightarrow$$



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

$$E[v^2(n)] = \alpha$$

$$E[v(n)] = 0$$

$$\hat{x}(0) = 10, \quad \Sigma(0|0) = 25$$

Prediction Step:

$$\hat{x}(n|n-1) = 0.95\,\hat{x}(n-1|n-1)$$

$$\Sigma(n|n-1) = 0.95^2\,\Sigma(n-1|n-1) + 1$$

Update Step:

$$\hat{x}(n|n) = \hat{x}(n|n-1) + K(n)\left[z(n) - \hat{x}(n|n-1)\right]$$

$$\Sigma(n|n) = \left[I - K(n)\right]\Sigma(n|n-1)$$

$$K(n) = \frac{\Sigma(n|n-1)}{\Sigma(n|n-1) + \alpha}$$

$$\Rightarrow \quad \Sigma(n|n) = \frac{\alpha\,\Sigma(n|n-1)}{\Sigma(n|n-1) + \alpha}$$

A second way to derive the update equation:

Let $\underline{x}, y$ be zero mean r.v., and partition y:

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Suppose $\hat{\underline{x}}(n|y_1)$ is known, and y_2 is then provided (a second measurement).

We could compute $\hat{\underline{x}}(n|y_1, y_2) = \hat{\underline{x}}(n|y_1) = \underline{\Gamma}_{xy}\underline{\Gamma}_{yy}^{-1} y$
(the best linear estimate, or MMSE estimate).

But, we already have $\hat{\underline{x}}(n|y_1)$.

If $\underline{\Gamma}_{y_1 y_2} = \underline{0}$, then

$$\hat{\underline{x}}(n|y) = \underline{\Gamma}_{xy}\underline{\Gamma}_{yy}^{-1} y = [\underline{\Gamma}_{xy_1} \; \underline{\Gamma}_{xy_2}] \begin{bmatrix} \underline{\Gamma}_{y_1y_1} & \underline{0} \\ \underline{0} & \underline{\Gamma}_{y_2y_2} \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \underline{\Gamma}_{xy_1}\underline{\Gamma}_{y_1y_1}^{-1} y_1 + \underline{\Gamma}_{xy_2}\underline{\Gamma}_{y_2y_2}^{-1} y_2$$

$$= \hat{\underline{x}}(n|y_1) + \hat{\underline{x}}(n|y_2)$$

$$\Rightarrow \quad \hat{\underline{x}}(n|y) = \hat{\underline{x}}(n|y_1) + \hat{\underline{x}}(n|y_2) \text{ when } E[y_1 y_2^T] = \underline{0}$$

But, given y_1, we know how to remove the portion from y_2 that is correlated with y_1

Suppose $\nu_2 = y_2 -$ component of y_2 correlated w/ y_1

$\nu_2 =$ innovations process — contains only <u>new</u> information.

If $\nu_2 = y_2 - A y_1$, then

$$\Gamma_{\nu_2 y_1} = E[\nu_2 y_1^T] = \Gamma_{y_2 y_1} - A \Gamma_{y_1 y_1} = 0$$

$$\text{or} \quad A = \Gamma_{y_2 y_1} \Gamma_{y_1 y_1}^{-1}$$

and

$$\begin{bmatrix} y_1 \\ \nu_2 \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ -\Gamma_{y_2 y_1}\Gamma_{y_1 y_1}^{-1} & I \end{bmatrix}}_{\text{invertible, so no information has been lost.}} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

→ We can find $\hat{x}(n|y_1, y_2)$ in the form of $\hat{x}(n|y_1, \nu_2)$

Since $\Gamma_{\nu_2 y_1} = 0$, $\hat{x}(n|y_1, \nu_2) = \hat{x}(n|y_1) + \hat{x}(n|\nu_2)$

$$= \hat{x}(n|y_1) + \Gamma_{x\nu_2}\Gamma_{\nu_2\nu_2}^{-1}(y_2 - \hat{y}_2|y_1)$$

$$= \hat{x}(n|y_1) + \Gamma_{x\nu_2}\Gamma_{\nu_2\nu_2}^{-1}(y_2 - \Gamma_{y_2 y_1}\Gamma_{y_1 y_1}^{-1} y_1)$$

You can also prove that

$$\Sigma(n|y_1, y_2) = \Sigma(n|y_1) - \Gamma_{x\nu_2}\Gamma_{\nu_2\nu_2}^{-1}\Gamma_{\nu_2 x}$$

For the Kalman filter,

associate $\underline{\hat{x}}(n|\underline{y}_1)$ with $\underline{\hat{x}}(n|n-1)$

" $\underline{y}_1$ " $\underline{z}(1), \underline{z}(2), \ldots, \underline{z}(n-1)$

" $\underline{y}_2$ " $\underline{z}(n)$

Given $\underline{\hat{x}}(n|n-1)$,

$$\underline{\hat{x}}(n|n) = \underline{\hat{x}}(n|n-1) + \underline{K}(n)\,\underline{\nu}(n)$$

where $\underline{K}(n) = E\{\underline{x}(n)\underline{\nu}^T(n)\}\,E\{\underline{\nu}(n)\underline{\nu}^T(n)\}^{-1}$

$$\begin{aligned}
\underline{\nu}(n) &= \underline{z}(n) - \underline{\hat{z}}(n|n-1) \\
&= \underline{z}(n) - \underline{H}(n)\,\underline{\hat{x}}(n|n-1) \\
&= \underline{H}(n)\left(\underline{x}(n) - \underline{\hat{x}}(n|n-1)\right) + \underline{v}(n)
\end{aligned}$$

Exercise: Find $E\{\underline{x}(n)\underline{\nu}^T(n)\}$ and $E\{\underline{\nu}(n)\underline{\nu}^T(n)\}$ and continue, to derive the Kalman update equations.

Suggested HW Problems

1. Let $\underline{v} \in \mathbb{R}^2$, $\underline{v} \sim N(\underline{0}, \underline{I})$

   Define $\underline{x} = \underline{H}\,\underline{v}$ and $\underline{H} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$, $\underline{U} = \underline{V}$

   and
$$\underline{U} = \begin{bmatrix} 0.7071 & -0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}, \quad \underline{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2)$$

   and $\sigma_1, \sigma_2 \in \{0.1, 0.2, 0.5, 10\}$ (16 cases).

   (a) Generate a 3D plot of $p(\underline{v})$ for $\|\underline{v}\|_\infty \le 3$.

   (b) Using 3D plots, show how $p(\underline{x})$ behaves on $\|\underline{v}\|_\infty \le 3$. Explain the roles of σ_i and the columns of $\underline{U}$.

2. Let $\underline{x} \sim N(\underline{0}, \underline{I})$, $\underline{x} \in \mathbb{R}^5$ and $\underline{v} \sim N(\underline{0}, \underline{G})$, $\underline{v} \in \mathbb{R}^3$. Assume that $\underline{y} = \underline{H}\,\underline{x} + \underline{v}$ is observed where

$$\underline{H} = \begin{bmatrix} -0.1170 & 0.3164 & 0.3806 & 0.2085 & 0.0213 \\ 0.0933 & 0.2491 & 0.2646 & 0.2921 & 0.1446 \\ 0.2398 & 0.3678 & 0.3714 & 0.5017 & 0.2839 \end{bmatrix}$$

$$\underline{G} = \begin{bmatrix} 0.7954 & -0.2128 & -0.3427 \\ -0.2128 & 0.7787 & -0.3565 \\ -0.3427 & -0.3565 & 0.4260 \end{bmatrix}$$

   (a) On what subspace of $\mathbb{R}^5$ can the projection of $\underline{x}$ be determined exactly given the observation $\underline{y}$?

   (b) On what subspace of $\mathbb{R}^5$ does the observation $\underline{y}$ provide no information about the projection of $\underline{x}$?

(c) On what maximal subspace of $\mathbb{R}^3$ does the measurement noise vector $\underline{v}$ have an effect on $\underline{y}$?

(d) Defining the estimate $\hat{\underline{x}} = \underline{E}\,\underline{y}$, $\hat{\underline{x}} \in \mathbb{R}^5$, find the value of $\underline{E}$ that minimizes $tr(cov(\underline{x} - \hat{\underline{x}}))$. What is the optimal estimate $\hat{\underline{x}}$ of $\underline{x}$ projected onto the subspace found in (b) for any $\underline{y}$?

The remaining problems are from Gelb, Chapters 3 & 4.

3-9, 3-11, 4-4, 4-8, 4-11, 4-13, 4-16, 4-22, 4-26, 4-28