The Discrete Kalman Filter

These notes are taken from *Statistical Digital Signal Processing and Modeling* by Monson Hayes, John Wiley, New York, NY, 1996. ISBN 0-47159431-8

Step 1

Since no new measurements are used to estimate **x**(*n*), all that is known is that **x**(*n*) evolves according to the state equation

$$\mathbf{x}(n) = \mathbf{A}(\nu-1)\boldsymbol{\xi}(\nu-1) + \boldsymbol{\omega}(\nu)\ .$$

Since **w**(*n*) is a zero mean white noise process (and the values of **w**(*n*) are unknown), we predict **x**(*n*) as follows

$$\boxed{\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(\nu-1)\hat{\boldsymbol{\xi}}(\nu-1|\nu-1)}$$

which has an estimation error given by

$$\begin{aligned}\mathbf{e}(n|n-1) &= \boldsymbol{\xi}(\nu) - \hat{\boldsymbol{\xi}}(\nu|\nu-1)\\ &= \mathbf{A}(\nu-1)\boldsymbol{\xi}(\nu-1) + \boldsymbol{\omega}(\nu) - \mathbf{A}(\nu-1)\hat{\boldsymbol{\xi}}(\nu-1|\nu-1)\\ &= \mathbf{A}(\nu-1)\boldsymbol{\varepsilon}(\nu-1|\nu-1) + \boldsymbol{\omega}(\nu)\end{aligned}$$

Note that since **w**(*n*) has zero mean, if $\hat{\mathbf{x}}(n-1|\nu-1)$ is an unbiased estimate of **x**(*n* – 1), i.e.

$$E[\mathbf{e}(n-1|\nu-1)] = 0$$

then $\hat{\mathbf{x}}(n|n-1)$ will be an unbiased estimate of **x**(*n*),

$$E[\mathbf{e}(n|n-1)] = 0\ .$$

Finally since the estimation error $\mathbf{e}(n-1|\nu-1)$ (past) is uncorrelated with **w**(*n*) (present), we have

$$\begin{aligned}\mathbf{P}(n|n-1) &= E[\mathbf{e}(n|n-1)\mathbf{e}^H(n|n-1)]\\ &= E[(\mathbf{A}(n-1)\mathbf{e}(n-1|n-1) + \mathbf{w}(n))(\mathbf{e}^H(n-1|n-1)\mathbf{A}^H(n-1) + \mathbf{w}^H(n))]\end{aligned}$$

or

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

Step 2

In the second step we incorporate the new measurement **y**(*n*) into the estimate $\hat{\mathbf{x}}(n|n-1)$. A linear estimate of **x**(*n*) that is based on $\hat{\mathbf{x}}(n|n-1)$ and **y**(*n*) is of the form

$$\hat{\mathbf{x}}(n|n) = \mathbf{K}'(\nu)\hat{\boldsymbol{\xi}}(\nu|\nu-1) + \mathbf{K}(\nu)\boldsymbol{\psi}(\nu)$$

where **K**(*n*) and **K**’(*n*) are matrices, yet to be specified.

The requirement that is imposed on $\hat{\mathbf{x}}(n|n)$ is that it be unbiased, $E[\mathbf{e}(n|n)] = 0$ and that it minimize the MSE $E[\|\mathbf{e}(n|n)\|^2]$.

We express $E[\mathbf{e}(n|n)]$ in terms of $E[\mathbf{e}(n|n-1)]$ as follows.

$$\begin{aligned}\mathbf{e}(n|n) &= \boldsymbol{\xi}(\nu) - \mathbf{K}'(\nu)\hat{\boldsymbol{\xi}}(\nu|\nu-1) - \mathbf{K}(\nu)\boldsymbol{\psi}(\nu) \\ &= \boldsymbol{\xi}(\nu) - \mathbf{K}'(\nu)[\boldsymbol{\xi}(\nu) - \boldsymbol{\varepsilon}(\nu|\nu-1)] - \mathbf{K}(\nu)[\mathbf{X}(\nu)\boldsymbol{\xi}(\nu) + \boldsymbol{\varpi}(\nu)] \\ &= [\mathbf{I} - \mathbf{K}'(\nu) - \mathbf{K}(\nu)\mathbf{X}(\nu)]\boldsymbol{\xi}(\nu) + \mathbf{K}'(\nu)\boldsymbol{\varepsilon}(\nu|\nu-1) - \mathbf{K}(\nu)\boldsymbol{\varpi}(\nu)\end{aligned}$$

Since $E[\mathbf{v}(n)] = 0$ and $E[e(n|n-1)] = 0$, $\hat{\mathbf{x}}(n|n)$ will be unbiased for any **x**(*n*) only if the term in brackets is zero,

$$\mathbf{K}'(n) = \mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\ .$$

With this constraint it follows that $\hat{\mathbf{x}}(n|n)$ has the form

$$\boxed{\hat{\mathbf{x}}(n|n) = \hat{\boldsymbol{\xi}}(\nu|\nu-1) + \mathbf{K}(\nu)[\boldsymbol{\psi}(\nu) - \mathbf{X}(\nu)\hat{\boldsymbol{\xi}}(\nu|\nu-1)]}$$

and the error is

$$\begin{aligned}\mathbf{e}(n|n) &= \mathbf{K}'(\nu)\boldsymbol{\varepsilon}(\nu|\nu-1) - \mathbf{K}(\nu)\boldsymbol{\varpi}(\nu) \\ &= [\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]\boldsymbol{\varepsilon}(\nu|\nu-1) - \mathbf{K}(\nu)\boldsymbol{\varpi}(\nu)\end{aligned}\ .$$

Since **v**(*n*) is uncorrelated with **w**(*n*), then **v**(*n*) is uncorrelated with **x**(*n*) and therefore, it is uncorrelated with $\hat{\mathbf{x}}(n|n-1)$. In addition, since $\mathbf{e}(n|n-1) = \boldsymbol{\xi}(\nu) - \hat{\boldsymbol{\xi}}(\nu|\nu-1)$, then **v**(*n*) is uncorrelated with $\mathbf{e}(n|n-1)$,

$$E[\mathbf{e}(n|n-1)v(n)] = 0.$$

Thus the error covariance matrix for $\mathbf{e}(n|n)$ is

$$\begin{aligned}\mathbf{P}(n|n) &= E[\boldsymbol{\varepsilon}(\nu|\nu)\boldsymbol{\varepsilon}^H(\nu|\nu)] \\ &= [\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]\boldsymbol{\Pi}(\nu|\nu-1)[\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]^H + \mathbf{K}(\nu)\boldsymbol{\Theta}_{\varpi}(\nu)\mathbf{K}^H(\nu)\end{aligned}$$

Step 3

We now determine the value for the Kalman gain **K**(*n*), that minimizes the MSE

$$\xi(n) = \mathrm{tr}\{\mathbf{P}(n|n)\}\ .$$

We proceed by computing the gradient of ξ(*n*) w.r.t. **K**(*n*). Note the following matrix differention formulas

$$\begin{aligned}\frac{\partial}{\partial\mathbf{K}}\mathrm{tr}(\mathbf{KA}) &= \mathbf{A}^H \\ \frac{\partial}{\partial\mathbf{K}}\mathrm{tr}(\mathbf{KAK}^H) &= 2\mathbf{KA}\end{aligned}\ .$$

The covariance matrix for $\mathbf{e}(n|n)$ is given by

$$\mathbf{P}(n|n) = [\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]\boldsymbol{\Pi}(\nu|\nu-1)[\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]^H + \mathbf{K}(\nu)\boldsymbol{\Theta}_{\varpi}(\nu)\mathbf{K}^H(\nu)$$

and the gradient of the MSE by

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}(n)}\mathrm{tr}\{\mathbf{P}(n|n)\} &= 2[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)[-\mathbf{C}^H(n)] + 2\mathbf{K}(n)\mathbf{Q}_v(n) \\ &= -2[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)\mathbf{C}^H(n) + 2\mathbf{K}(n)\mathbf{Q}_v(n)\end{aligned}.$$

Solving for **K**(*n*) we have

$$\frac{\partial}{\partial \mathbf{K}(n)}\mathbf{P}(n|n) = 0$$

which yields

$$\begin{aligned}&-2[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)\mathbf{C}^H(n) + 2\mathbf{K}(n)\mathbf{Q}_v(n) = \mathbf{0} \\ &\Rightarrow [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)\mathbf{C}^H(n) = \mathbf{K}(n)\mathbf{Q}_v(n) \\ &\Rightarrow \mathbf{P}(n|n-1)\mathbf{C}^H(n) - \mathbf{K}(n)\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) = \mathbf{K}(n)\mathbf{Q}_v(n) \\ &\Rightarrow \mathbf{P}(n|n-1)\mathbf{C}^H(n) = \mathbf{K}(n)[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n) + \mathbf{Q}_v(n)]\end{aligned}$$

or

$$\boxed{\mathbf{K}(n) = \mathbf{\Pi}(\nu|\nu-1)\mathbf{X}^H(\nu)[\mathbf{X}(\nu)\mathbf{\Pi}(\nu|\nu-1)\mathbf{X}^H(\nu) + \mathbf{\Theta}_{\varpi}(\nu)]^{-1}}.$$

Given the Kalman gain, we can simplify the error covariance matrix, $\mathbf{P}(n|n)$ as follows. We first rearrange the matrix

$$\begin{aligned}\mathbf{P}(n|n) &= [\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1)[\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]^H + \mathbf{K}(\nu)\mathbf{\Theta}_{\varpi}(\nu)\mathbf{K}^H(\nu) \\ &= [\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1) - [\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1)\mathbf{X}^H(\nu)\mathbf{K}^H(\nu) + \mathbf{K}(\nu)\mathbf{\Theta}_{\varpi}(\nu)\mathbf{K}^H(\nu) \\ &= [\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1) - \{[\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1)\mathbf{X}^H(\nu) + \mathbf{K}(\nu)\mathbf{\Theta}_{\varpi}(\nu)\}\mathbf{K}^H(\nu)\end{aligned}.$$

We notice that the term in braces is equal to the gradient which we set to zero. Therefore

$$\boxed{\mathbf{P}(n|n) = [\mathbf{I}-\mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu|\nu-1)}.$$

We have thus derived the Kalman filtering equations for recursively estimating the state vector **x**(*n*). We now initialize the recursion. Since the value of the initial state is unknown, in the absence of any observed data at *n* = 0, the initial estimate is chosen to be

$$\hat{\mathbf{x}}(0|0) = E[\boldsymbol{\xi}(0)]$$

and the error covariance matrix is initialized as

$$\mathbf{P}(0|0) = E[\boldsymbol{\xi}(0)\boldsymbol{\xi}^H(0)].$$

Finally, we note the initial conditions make $\hat{\mathbf{x}}(0|0)$ an unbiased estimator of **x**(0) and ensures $\hat{\mathbf{x}}(n|n)$ will be unbiased for all *n* (as required in our derivation).