Kalman Filter Derivation

Overview

1. Discuss several useful matrix identities.
2. Derive Kalman filter algorithms.
3. Discuss alternate form (Alternate Gain Expression) of the Kalman filter.

References

1. Applied Optimal Estimation. Edited by Arthur Gelb. M.I.T. Press 1986.
2. Introduction to Random Signals and Applied Kalman Filtering. 2nd Edition. R.G. Brown and P.Y.C. Hwang. John Wiley and Sons, Inc. New York. 1992.

Kalman Filter Equations

In this section, we will derive the five Kalman filter equations

1. State extrapolation

$$\hat{\mathbf{x}}_{k+1} = \Phi_{k+1}\,\hat{\mathbf{x}}_k$$

2. Covariance Extrapolation

$$\mathbf{P}^{-}_{k+1} = \Phi_{k+1}\,\mathbf{P}_k\,\Phi^{T}_{k+1} + \mathbf{Q}_k$$

3. Kalman Gain Computation

$$\mathbf{K}_{k+1} = \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\left[\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right]^{-1}$$

4. State Update

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} + \mathbf{K}_{k+1}\left[\mathbf{z}_{k+1} - \mathbf{H}_{k+1}\hat{\mathbf{x}}^{-}_{k+1}\right]$$

5. Covariance Update

$$\mathbf{P}_{k+1} = \mathbf{P}^{-}_{k+1} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}$$

Definitions and Identities

Vector and Matrix Differentiation

Let z be a scalar and $\mathbf{x}$ be a column vector

$$\frac{\delta_z}{\delta_{\mathbf{x}}} = \begin{bmatrix} \dfrac{\delta_z}{\delta_{x_1}} \\ \vdots \\ \dfrac{\delta_z}{\delta_{x_n}} \end{bmatrix}_{nx1}$$

Likewise, for the matrix $\mathbf{A}_{mxn}$

$$\frac{\delta_z}{\delta_{\mathbf{A}}} = \begin{bmatrix} \dfrac{\delta_z}{\delta_{a_{11}}} & \cdots & \dfrac{\delta_z}{\delta_{a_{1n}}} \\ \vdots & \ddots & \vdots \\ \dfrac{\delta_z}{\delta_{a_{m1}}} & \cdots & \dfrac{\delta_z}{\delta_{a_{mn}}} \end{bmatrix}$$

Definitions and Identities

(1) $$\frac{\delta \mathbf{x}^T \mathbf{y}}{\delta \mathbf{x}} = \mathbf{y} = \frac{\delta \mathbf{y}^T \mathbf{x}}{\delta \mathbf{x}}$$

(2) $$\frac{\delta \mathbf{x}^T \mathbf{N} \mathbf{x}}{\delta \mathbf{x}} = 2\, \mathbf{N}\, \mathbf{x}$$ (where $\mathbf{N}$ is symmetric)

(3) $$\frac{\delta (\mathbf{A}\, \mathbf{x} + \mathbf{b})^T\, \mathbf{M} (\mathbf{A}\, \mathbf{x} + \mathbf{b})}{\delta \mathbf{x}} = 2\, \mathbf{A}^T \mathbf{M}\, \mathbf{A}\, \mathbf{x} + 2\, \mathbf{A}^T\, \mathbf{M}\, \mathbf{b} = 2\, \mathbf{A}^T\, \mathbf{M}\, (\mathbf{A}\, \mathbf{x} + \mathbf{b})$$

(4) $$\frac{\delta}{\delta \mathbf{A}} \operatorname{trace}(\mathbf{AC}) = \mathbf{C}^T$$

Note for $\mathbf{A}\ \mathbf{C}$ to be square, $\dim \mathbf{A} = \dim \mathbf{C}^T$

(5) $$\frac{\delta}{\delta \mathbf{A}} \operatorname{trace}\left(\mathbf{ABA}^T\right) = 2\, \mathbf{A}\, \mathbf{B}$$ (where $\mathbf{B}$ is symmetric)

Kalman Filter Derivation

Definitions and Identities

(6) Gradient Expression

$$\frac{\delta}{\delta \mathbf{x}}\left[(\mathbf{Hx}-\mathbf{z})^{T}\mathbf{W}(\mathbf{Hx}-\mathbf{z})\right] = 2\mathbf{H}^{T}\mathbf{W}(\mathbf{Hx}-\mathbf{z}) \qquad \textbf{(}\text{for } \mathbf{W} \text{ symmetric}\textbf{)}$$

(7) Gain Expression Proof

$$\mathbf{H}^{T} = \mathbf{H}^{T}$$

$$\mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HPH}^{T} + \mathbf{H}^{T} = \mathbf{H}^{T} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HPH}^{T}$$

$$\mathbf{H}^{T}\mathbf{R}^{-1}\left[\mathbf{HPH}^{T} + \mathbf{R}\right] = \left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]\mathbf{PH}^{T}$$

$$\left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\mathbf{H}^{T}\mathbf{R}^{-1} = \mathbf{PH}^{T} + \left[\mathbf{HPH}^{T} + \mathbf{R}\right]^{-1}$$

(8) Inversion Lemma Proof

$$\left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1} = \left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\left[\mathbf{I} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HP} - \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HP}\right]$$

$$= \left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]\mathbf{P} - \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HP}$$

$$= \mathbf{P} - \left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{HP}$$

Use identity above to obtain **"**Inversion Lemma**"**

$$\left[\mathbf{P}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}\right]^{-1} = \mathbf{P} - \mathbf{PH}^{T}\left[\mathbf{HPH}^{T} + \mathbf{R}\right]^{-1}\mathbf{H}$$

Assumptions

Assume the following form of the estimator

- linear
- recursive

Goal is to show that the Kalman Filter Equations provide the minimum variance estimator over all unbiased estimators which have this form

No assumptions are made concerning the particular distribution of the process or measurement noise

Kalman Filter Derivation

Model

Process: $\mathbf{x}_{k+1} = \Phi_{k+1}\mathbf{x}_k + \mathbf{w}_k$

Measurement: $\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + v_k$

Assumptions:

$$E[\mathbf{x}_0] = \mu_0^{\mathbf{x}}$$

$$E[\mathbf{w}_k] = 0\forall k$$

$$E[\mathbf{v}_k] = 0\forall k$$

$$\operatorname{cov}\{\mathbf{w}_k, \mathbf{w}_j\} = \mathbf{Q}_k\delta_{kj}$$

$$\operatorname{cov}\{\mathbf{v}_k, \mathbf{v}_j\} = \mathbf{R}_k\delta_{kj}$$

$$\operatorname{cov}\{\mathbf{x}_0, \mathbf{x}_0\} = \mathbf{P}_0$$

$$\operatorname{cov}\{\mathbf{w}_k, \mathbf{v}_j\} = 0, \forall k,$$

$$\operatorname{cov}\{\mathbf{x}_0, \mathbf{w}_k\} = 0, \forall k,$$

$$\operatorname{cov}\{\mathbf{x}_0, \mathbf{v}_j\} = 0, \forall j$$

Kalman Filter Derivation

Assumptions

1. Assume that at time $k+1$ we have available an unbiased estimate of the state at time k

   The error term $\tilde{\mathbf{x}}_k = \hat{\mathbf{x}}_k - \mathbf{x}_k$ in addition to having zero mean, has cov ariance $\mathbf{P}_k$

2. At time k, we have a measurement $\mathbf{z}_k$ available, where

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$$

Kalman Filter Derivation

Goal

The goal is to find unbiased minimum variance estimator of the state at time k+1, of the form

$$\hat{\mathbf{x}}_{k+1} = \mathbf{K}'_{k+1}\hat{\mathbf{x}}_k + \mathbf{K}_{k+1}\mathbf{z}_{k+1}$$

Note that this estimator is both

linear in $\hat{\mathbf{x}}_k$ and $\mathbf{z}_{k+1}$

recursive - only processes the current measurement $\mathbf{z}_{k+1}$

Kalman Filter Derivation	Derivation Steps

Step 1. For unbiased $\hat{\mathbf{x}}_k$, develop expression for $\mathbf{K}'_{k+1}$. Substitute for $\mathbf{K}'_{k+1}$ to obtain state update expression (equation 4).

Step 2. Develop expression for $\mathbf{K}'_{k+1}$ that minimizes variance.

2a. Define $\hat{\mathbf{x}}_{\overline{k}+1}$ (equation 1)

Define $\mathbf{P}_{\overline{k}+1}$ (equation 2)

2b. Define $\mathbf{P}_{k+1}$ (equation 5 – Joseph form)

2c. Define $\mathbf{K}_{k+1}$ (equation 3)

Define $\mathbf{P}_{k+1}$ (equation 5 – short form)

Kalman Filter Derivation

Step 1

The unbiased criteria forces a certain relationship between $\mathbf{K}'_{k+1}$ and $\mathbf{K}_{k+1}$. What we will see, over the next few pages, is that this criteria requires that $\mathbf{K}'_{k+1} = \Phi_{k+1} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\,\Phi_{k+1}$.

For $\hat{\mathbf{x}}_{k+1}$ to be unbiased, $E\,\underbrace{\left[\hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1}\right]}_{\tilde{\mathbf{x}}_{k+1}} = 0.$

Kalman Filter Derivation | Step 1

Substituting for the terms in the brackets gives

$$\hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1} = \mathbf{K}'_{k+1}\,\hat{\mathbf{x}}_k + \mathbf{K}_{k+1}\,\mathbf{z}_{k+1} - \mathbf{x}_{k+1}$$

Adding and subtracting two terms, and further substitution gives

$$= \mathbf{K}'_{k+1}\hat{\mathbf{x}}_{k+1} + \mathbf{K}_{k+1}\left(\mathbf{H}_{k+1}\,\mathbf{x}_{k+1} + \mathbf{v}_{k+1}\right) - \mathbf{x}_{k+1} \underbrace{- \mathbf{K}'_{k+1}\,\mathbf{x}_k + \mathbf{K}'_{k+1}\mathbf{x}_k}_{\text{Add and Subtract}}$$

Rearranging terms gives

$$= \mathbf{K}'_{k+1}\left[\hat{\mathbf{x}}_k + \mathbf{x}_k\right] + \mathbf{K}_{k+1}\left(\mathbf{H}_{\mathbf{k+1}}\left[\Phi_{k+1}\,\mathbf{x}_k + \mathbf{w}_{k+1}\right] + \mathbf{v}_{\mathbf{k+1}}\right) - \left(\Phi_{k+1}\,\mathbf{x}_k + \mathbf{w}_{k+1}\right) + \mathbf{K}'_{k+1}$$

$$= \mathbf{K}'_{k+1}\left[\hat{\mathbf{x}}_k + \mathbf{x}_k\right] + \left[\mathbf{K}_{k+1}\,\mathbf{H}_{\mathbf{k+1}}\,\Phi_{k+1} - \Phi_{k+1} + \mathbf{K}'_{k+1}\right]\mathbf{x}_{\mathbf{k}} + \left(\mathbf{K}_{k+1}\,\mathbf{H}_{k+1} - \mathbf{I}\right)\mathbf{w}_{\mathbf{k+1}} + \mathbf{K}_{\mathbf{k+1}}\mathbf{v}_{\mathbf{k}}$$

Kalman Filter Derivation | **Step 1**

The final step is to take the expectation of this expression and set it equal to zero. For the right hard side to be equal to zero, the following must be true

$$E\left[\hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1}\right] = \left[\mathbf{K}_{k+1}\,\mathbf{H}_{\mathbf{k+1}}\,\Phi_{K+1} - \Phi_{\mathbf{k+1}} + \mathbf{K}'_{k+1}\right] E\left[\boldsymbol{x}_k\right] = 0$$

which implies

$$\mathbf{K}_{\mathbf{k+1}}\,\mathbf{H}_{\mathbf{k+1}}\,\Phi_{\mathbf{k+1}} - \Phi_{\mathbf{k+1}} + \mathbf{K}'_{k+1} = 0$$

or

$$\mathbf{K}'_{k+1} = \left(\mathbf{I} - \mathbf{K}_{\mathbf{k+1}}\,\mathbf{H}_{\mathbf{k+1}}\right)\Phi_{\mathbf{k+1}}$$

Kalman Filter Derivation

Step 1

Thus, to satisfy the unbiased criteria:

$$\hat{\mathbf{x}}_{\mathbf{k+1}} = \left(\mathbf{I} - \mathbf{K}_{\mathbf{k+1}}\,\mathbf{H}_{\mathbf{k+1}}\right)\Phi_{\mathbf{k+1}}\,\hat{\mathbf{x}}_{\mathbf{k}} + \mathbf{K}_{\mathbf{k+1}}\,\mathbf{z}_{\mathbf{k+1}}$$

or equivalently

$$\hat{\mathbf{x}}_{\mathbf{k+1}} = \underbrace{\Phi_{\mathbf{k+1}}\,\hat{\mathbf{x}}_{\mathbf{k}}}_{\substack{\text{extrapolated}\\ \text{state}}} + \underbrace{\mathbf{K}_{\mathbf{k+1}}\left(\mathbf{z}_{\mathbf{k+1}} - \mathbf{H}_{k+1}\,\Phi_{k+1}\,\hat{\mathbf{x}}_{\mathbf{k}}\right)}_{\substack{\text{residual of measurement}\\ \text{and prediction of}\\ \text{measurement}}}$$

which is the state update equation (equation 4)

It remains to find the value of $\mathbf{K}_{k+1}$ which minimizes the covariance of the estimation error

Kalman Filter Derivation

Step 1

The estimation error is

$$\tilde{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1}$$

The covariance of this error

$$\mathbf{P}_{k+1} \text{ looks like } \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}$$

The goal will be to find $\mathbf{K}_{k+1}$ such that

$$\sigma_1^2 + \cdots + \sigma_n^2 = \text{Trace } \mathbf{P}_{k+1}$$

is minimized (i.e., minimum variance).

Kalman Filter Derivation

Step 2

Step 2 is to find $\mathbf{K}_{k+1}$ which minimizes Trace $\mathbf{P}_{k+1}$, where

$$\mathbf{P}_{k+1} = E\left[\tilde{\mathbf{x}}_{k+1}\, \tilde{\mathbf{x}}_{k+1}^{T}\right]$$

A. First, find the covariance of the extrapolated estimate error,

$\mathbf{P}^{-}_{k+1}$ (equation 2).

The extrapolated estimate is defined as (equation 1)

$$\hat{\mathbf{X}}^{-}_{k+1} = \Phi_{k+1} \hat{\mathbf{X}}_{k}$$

Kalman Filter Derivation | Step 2

The extrapolated estimate error is then

$$\tilde{\mathbf{x}}^{-}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} - \mathbf{x}_{k+1} = \Phi_{k+1} + \hat{\mathbf{x}}_k - \Phi_{k+1}\,\mathbf{x}_k + \mathbf{w}_{k+1}$$

we can see $\underbrace{\hat{\mathbf{x}}^{-}_{k+1} - \mathbf{x}_{k+1}}_{\text{has cov } \mathbf{P}^{-}_{k+1}} = \Phi_{k+1}\underbrace{\left[\hat{\mathbf{x}}_{k+1} - \mathbf{x}_k\right]}_{\text{has Cov } \mathbf{P}_k} + \underbrace{\mathbf{w}_{k+1}}_{\text{has Cov } \mathbf{Q}_{k+1}}$

To obtain equation 2, take the expected value of both sides

$$\begin{aligned}
\mathbf{P}_{k+1} &= E\left\{\tilde{\mathbf{x}}_{k+1}\,\tilde{\mathbf{x}}^{-T}_{k+1}\right\} \\
&= \Phi_{k+1} E\left\{\left(\tilde{\mathbf{x}}_k - \mathbf{x}_k\right)\left(\tilde{\mathbf{x}}_k - \mathbf{x}_k\right)^T\right\}\Phi^T_{k+1} + E\left\{\mathbf{w}_{k+1}\,\mathbf{w}^T_{k+1}\right\} \\
&= \Phi_{k+1}\,\mathbf{P}_k\,\Phi^T_{k+1} + \mathbf{Q}_{k+1}
\end{aligned}$$

Thus, equation 2 is $\mathbf{P}^{-}_{k+1} = \Phi_{k+1}\,\mathbf{P}_k\,\Phi^T_{k+1} + \mathbf{Q}_{k+1}$

Kalman Filter Derivation — Step 2

B. Find the covariance $\mathbf{P}_{k+1}$ (the covariance of the final estimation error, equation 5). It will be a function of $\mathbf{K}_{k+1}$ and $\mathbf{P}^{-}_{k+1}$.

Using the identity $\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} - \mathbf{K}_{k+1}\left(\mathbf{z}_{k+1} - \mathbf{H}_{k+1}\hat{\mathbf{x}}^{-}_{k+1}\right)$ in

$$\tilde{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1}$$

Gives

$$\tilde{\mathbf{x}}_{k+1} = [\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}]\hat{\mathbf{x}}^{-}_{k+1} + \mathbf{K}_{k+1}\,\mathbf{z}_{k+1} - \mathbf{x}_{k+1}$$

$$= [\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}]\hat{\mathbf{x}}^{-}_{k+1} + \mathbf{K}_{k+1}\,[\mathbf{H}_{k+1}\,\mathbf{x}_{k+1} + \mathbf{v}_{k+1}] - \mathbf{x}_{k+1}$$

$$= \hat{\mathbf{x}}^{-}_{k+1} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\,\mathbf{x}^{-}_{k+1} + \mathbf{K}_{k+1}\,\mathbf{H}^{T}_{k+1}\,\mathbf{x}_{k+1} + \mathbf{K}_{k+1}\mathbf{v}_{k+1} - \mathbf{x}_{k+1}$$

$$= \left[\hat{\mathbf{x}}^{-}_{k+1} - \mathbf{x}_{k+1}\right] - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\left[\hat{\mathbf{x}}^{-}_{k+1} - \mathbf{x}_{k+1}\right] + \mathbf{K}_{k+1}\mathbf{v}_{k+1}$$

$$= [\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}]\left[\hat{\mathbf{x}}^{-}_{k+1} - \mathbf{x}_{k+1}\right] + \mathbf{K}_{k+1}\,\mathbf{v}_{k+1}$$

$$= [\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}]\tilde{\mathbf{x}}^{-}_{k+1} - \mathbf{K}_{k+1}\,\mathbf{v}_{k+1}$$

Kalman Filter Derivation — Step 2

Thus

$$\tilde{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1} = \left[\mathbf{I} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\right]\tilde{\mathbf{x}}^{-}_{k+1} + \mathbf{K}_{k+1}\,\mathbf{x}_{k+1}$$

Taking the expected value of both sides will provide us with an expression for $\mathbf{P}_{k+1}$

$$\mathbf{P}_{k+1} = \mathrm{Cov}\,\tilde{\mathbf{x}}_{k+1} = E\left[\tilde{\mathbf{x}}_{k+1}\,\tilde{\mathbf{x}}^{\mathbf{T}}{}_{k+1}\right]$$

$$= \left[\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]\mathbf{P}^{-}_{k+1}\left[\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]^{T} + \mathbf{K}_{k+1}\mathbf{R}_{k+1}\,\mathbf{K}^{T}_{k+1}$$

Now we have found the expression for the covariance update, equation 5. Note that $\mathbf{P}_{k+1}$ is a function of $\mathbf{K}_{k+1}$, $\mathbf{P}^{-}_{k+1}$, and $\mathbf{R}_{k+1}$.

Kalman Filter Derivation | Step 2

c. The final step is to find an expression for $\mathbf{K}_{k+1}$ (equation 3) which minimizes the trace of $\mathbf{P}_{k+1}$.

Using $\mathbf{P}$ as shorthand notation for $\mathbf{P}^-_{k+1}$, $\mathbf{K}$ for $\mathbf{K}_{k+1}$, $\mathbf{R}$ for $\mathbf{R}_{k+1}$, and $\mathbf{H}$ for $\mathbf{H}_{k+1}$

$$
\begin{aligned}
\mathbf{P}_{k+1} &= (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{P}\,(\mathbf{I} - \mathbf{K}\,\mathbf{H})^T + \mathbf{K}\,\mathbf{R}\,\mathbf{K}^T \\
&= (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{P}\left(\mathbf{I} - \mathbf{H}^T\,\mathbf{K}^T\right) + \mathbf{K}\,\mathbf{R}\,\mathbf{K}^T \\
&= \mathbf{P} - \mathbf{K}\,\mathbf{H}\,\mathbf{P} - \mathbf{P}\,\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\,\mathbf{H}\,\mathbf{P}\,\mathbf{H}^T\,\mathbf{K}^T + \mathbf{K}\,\mathbf{R}\,\mathbf{K}^T
\end{aligned}
$$

$$
\mathrm{Tr}\,\mathbf{P}_{k+1} = \mathrm{Tr}\,\mathbf{P} - 2\,\mathrm{Tr}\,\mathbf{K}\,\mathbf{H}\,\mathbf{P} + \mathrm{Tr}\,\mathbf{K}\left(\mathbf{H}\,\mathbf{P}\,\mathbf{H}^T\right)\mathbf{K}^T + \mathrm{Tr}\,\mathbf{K}\,\mathbf{R}\,\mathbf{K}^T
$$

Kalman Filter Derivation

Step 2

Using the identities $\left(\mathbf{P}\,\mathbf{H}^T\mathbf{K}^T\right)^T = \mathbf{K}\,\mathbf{H}\,\mathbf{P}$

$$\mathrm{Tr}\left(\mathbf{P}\,\mathbf{H}^T\mathbf{K}^T\right) = \mathrm{Tr}\left(\mathbf{K}\,\mathbf{H}\,\mathbf{P}\right)$$

$$\frac{\delta\ \mathrm{Tr}\ \mathbf{A}\,\mathbf{B}\,\mathbf{A}^T}{\delta\mathbf{A}} = 2\ \mathbf{A}\,\mathbf{B} \text{ where } \mathbf{B} \text{ is symmetric}$$

$$\frac{\delta\ \mathrm{Tr}\ \mathbf{A}\,\mathbf{C}}{\delta\mathbf{A}} = \mathbf{C}^T$$

We obtain the partial of Tr (P_{k+1})with request to K

$$\frac{\delta\ \mathrm{Tr}\ \mathbf{P}_{k+1}}{\delta\mathbf{K}} = -2\ \mathbf{P}\,\mathbf{H}^T + 2\,\mathbf{K}\,\mathbf{H}\,\mathbf{P}\,\mathbf{H}^T + 2\,\mathbf{K}$$

Kalman Filter Derivation | Step 2

Taking the partial with respect to **K**,

$$\frac{\delta \operatorname{Tr} \mathbf{P}_{k+1}}{\delta \mathbf{K}} = -2\,\mathbf{P}\,\mathbf{H}^T + 2\,\mathbf{K}\,\mathbf{H}\,\mathbf{P}\,\mathbf{H}^T + 2\,\mathbf{K}$$

And setting this equal to zero and solving for K gives

$$\mathbf{K} = \mathbf{P}\,\mathbf{H}^T\left(\mathbf{H}\,\mathbf{P}\,\mathbf{H}^T + \mathbf{R}\right)^{-1}$$

which is the Kalman gain (equation 3)!

It can be verified that this is indeed a minimum (reference Gelb, pg. 109).

Kalman Filter Derivation

Step 2

Now we have an expression for $\mathbf{K}_{k+1}$ that optimizes the estimate

$$\mathbf{K}_{k+1} = \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\left[\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right]^{-1}$$

We can substitute $\mathbf{K}_{k+1}$ in our $\mathbf{P}_{k+1}$ expression

$$\mathbf{P}_{k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\right)\mathbf{P}^{-}_{k+1}\left[\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]^{T} + \mathbf{K}_{k+1}\,\mathbf{R}_{k+1}\,\mathbf{K}^{T}_{k+1}$$

To get the $\mathbf{P}_{k+1}$ expression

$$\mathbf{P}_{k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\right)\mathbf{P}^{-}_{k+1}$$

Kalman Filter Derivation

Summary

Thus we have the recursive algorithm

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} + \mathbf{K}_{k+1}\left[\mathbf{z}_{k+1} - \mathbf{H}_{k+1}\,\hat{\mathbf{x}}^{-}_{k+1}\right]$$

where

$$\hat{\mathbf{x}}^{-}_{k+1} = \Phi_{k+1}\,\hat{\mathbf{x}}_{\mathbf{k}}$$

$$\mathbf{P}^{-}_{k+1} = \Phi_{k+1}\mathbf{P}_{k}\,\Phi^{T}_{k+1} + \mathbf{Q}_{k+1}$$

$$\mathbf{K}_{k+1} = \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\left[\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right]^{-1}$$

$$\mathbf{P}_{k+1} = \left[\mathbf{I} - \mathbf{K}_{k+1}\,\mathbf{K}_{k+1}\right]\mathbf{P}^{-}_{k+1}$$

Kalman Filter Derivation

Alternate Gain Expression

The standard Kalman Filter algorithm computes the gain $\mathbf{K}_{k+1}$, then computes the updated covariance $\mathbf{P}_{k+1}$ as a function of the gain.

$$\mathbf{K}_{k+1} = \mathbf{P}^{-}_{k+1}\mathbf{H}^{T}_{k+1}\left[\mathbf{H}_{k+1}\mathbf{P}^{-}_{k+1}\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right]^{-1}$$

$$\mathbf{P}_{k+1} = \left[\mathbf{I} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]\mathbf{P}^{-}_{k+1}$$

This computation involves taking the inverse of a *m* x *m* matrix, where *m* = dim **z**

Usually dim **z** < dim **x** (size measurement vector smaller than number of states), so this formulation is desirable.

Alternate Gain Expression

Another formulation exists which involves reversing this, i.e. computing the updated covariance $\mathbf{P}_{k+1}$ first, and finding $\mathbf{K}_{k+1}$ as a function of $\mathbf{P}_{k+1}$.

This form is

$$\mathbf{P}_{k+1} = \left[\left(\mathbf{P}^{-}_{k+1} \right)^{-1} + \mathbf{H}^{T}_{k+1} \mathbf{R}^{-1}_{k+1} \mathbf{H}_{k+1} \right]^{-1}$$

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1} \mathbf{H}^{T}_{k+1} \mathbf{R}^{-1}_{k+1}$$

Note that this involves computing the inverse of n dimensional matrices, where n = dim $\mathbf{x}$, the state (in addition to $\mathbf{R}^{-1}_{k+1}$)

There are situations where this would be computationally preferable - if dim $\mathbf{x}$ < dim $\mathbf{z}$ and $\mathbf{R}_{k+1}$ is of a simple form ($\mathbf{I}$, diagonal, etc.)

Alternate Gain Expression

Derivation of the Alternate formulation follows directly from the Matrix Inversion Lemma (MIL) and Gain Expression (GE) Identity given earlier.

$$\mathbf{P}_{k+1} = \left(\mathbf{I} - \mathbf{K}_{k+1}\,\mathbf{H}_{k+1}\right)\mathbf{P}^{-}_{k+1} = \mathbf{P}^{-}_{k+1} - \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\left(\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right)^{-1}\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}$$

$$\overset{\text{MIL}}{=} \left[\left(\mathbf{P}^{-}_{k+1}\right)^{-1} + \mathbf{H}^{T}_{k+1}\,\mathbf{R}^{-1}_{k+1}\,\mathbf{H}_{k+1}\right]^{-1}$$

$$\mathbf{K}_{k+1} = \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\left(\mathbf{H}_{k+1}\,\mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1} + \mathbf{R}_{k+1}\right)^{-1} \overset{\text{GE}}{=} \left[\left(\mathbf{P}^{-}_{k+1}\right)^{-1} + \mathbf{H}^{T}_{k+1}\,\mathbf{R}^{-1}_{k+1}\,\mathbf{H}_{k+1}\right]^{-1}\mathbf{H}^{T}_{k+1}\,\mathbf{R}^{-1}_{k+1}$$

$$= \mathbf{P}^{-}_{k+1}\,\mathbf{H}^{T}_{k+1}\,\mathbf{R}^{-1}_{k+1}$$