Potter Square Root Filter

ASEN 5070
October 30, 2002
George H. Born

Introduction

In general, square root filters are more numerically stable than the conventional Kalman filter. Note that the condition number for the square root of a covariance matrix is the square root of the condition number of the covariance matrix. Hence, the square root filter will be less affected by numerical problems

The first square root filter development is due to Potter who developed an algorithm for the limited case of uncorrelated scalar observations with no process noise. This filter was used in the Lunar Excursion Module (LEM) for the Apollo Program.

Potter Algorithm Derivation

Begin with the time update equation for the estimation error covariance (we will assume that the time update is from t_{k-1} to t_k and drop the indices)

$$\bar{P} = \Phi P \Phi^T . \tag{1}$$

Define the square root of P

$$W W^T = P . \tag{2}$$

From Eq. (1) and (2)

$$\bar{P} = \Phi W W^T \Phi^T \equiv \bar{W} \bar{W}^T \tag{3}$$

where

$$\bar{W} = \Phi W . \tag{4}$$

Next write the expression for the Kalman gain and the measurement update for P in terms of W. (Note that observations are processed one at a time and are assumed to have uncorrelated errors.)

$$\begin{aligned} K &= \bar{P}\tilde{H}^T \left(\tilde{H}\bar{P}\tilde{H}^T + \sigma^2\right)^{-1} \\ &= \bar{W}\bar{W}^T \tilde{H}^T \left(\tilde{H}\bar{W}\bar{W}^T \tilde{H}^T + \sigma^2\right)^{-1} \end{aligned} \tag{5}$$

where σ^2 is the variance of the observation error.

Let

$$\alpha \equiv \left(\tilde{H}\bar{W}\bar{W}^T\tilde{H}^T + \sigma^2\right)^{-1} \tag{6}$$

where α is a scalar. Define

$$\tilde{F} = W^T\tilde{H}^T . \tag{7}$$

then

$$\alpha = \left(\tilde{F}^T\tilde{F} + \sigma^2\right)^{-1} \tag{8}$$

and

$$K = \alpha\bar{W}\tilde{F} . \tag{9}$$

The measurement update for P is

$$\begin{aligned} P = WW^T &= \left(I - K\tilde{H}\right)\bar{P} \\ &= \left(I - \alpha\bar{W}\tilde{F}\tilde{H}\right)\bar{W}\bar{W}^T \\ &= \bar{W}\left(I - \alpha\tilde{F}\tilde{H}\bar{W}\right)\bar{W}^T \\ &= \bar{W}\left(I - \alpha\tilde{F}\tilde{F}^T\right)\bar{W}^T \quad . \end{aligned} \tag{10}$$

Potter observed that if a matrix $\bar{A}$ could be found such that

$$\bar{A}\bar{A}^T = \left(I - \alpha\tilde{F}\tilde{F}^T\right) \tag{11}$$

then

$$P = \bar{W}\bar{A}\bar{A}^T\bar{W}^T = WW^T . \tag{12}$$

To find $\bar{A}$ introduce the scalar γ so that

$$\begin{aligned} \bar{A}\bar{A}^T &= \left(I - \gamma\alpha\tilde{F}\tilde{F}^T\right)\left(I - \gamma\alpha\tilde{F}\tilde{F}^T\right) \\ &= \left(I - \alpha\tilde{F}\tilde{F}^T\right) \quad . \end{aligned} \tag{13}$$

Solving for γ ,

$$I - \alpha\tilde{F}\tilde{F}^T = I - 2\gamma\alpha\tilde{F}\tilde{F}^T + \gamma^2\alpha^2\tilde{F}\tilde{F}^T\tilde{F}\tilde{F}^T . \tag{14}$$

Define

$$\beta \equiv \tilde{F}_{1\times n}^{T}\tilde{F}_{n\times 1}\,, \tag{15}$$

where β is a scalar. Then

$$I - \alpha\tilde{F}\tilde{F}^{T} = I - 2\gamma\alpha\tilde{F}\tilde{F}^{T} + \gamma^{2}\alpha^{2}\beta\tilde{F}\tilde{F}^{T}$$

or

$$\begin{aligned}&\left(\gamma^{2}\alpha^{2}\beta - 2\gamma\alpha + \alpha\right)\tilde{F}\tilde{F}^{T} = 0\\ &= \left(\alpha\beta\gamma^{2} - 2\gamma + 1\right)\alpha\tilde{F}\tilde{F}^{T} = 0\ .\end{aligned}$$

$\alpha\tilde{F}\tilde{F}^{T} = 0$ is a trivial solution.

Thus

$$\alpha\beta\gamma^{2} - 2\gamma + 1 = 0\,. \tag{16}$$

Using the solution for a quadratic equation (after some algebra)

$$\gamma = \frac{1}{1 \mp \sqrt{\alpha\sigma^{2}}}\,, \tag{17}$$

where the + sign is chosen to prevent the possibility that $\gamma = \infty$ when $\alpha\sigma^{2} = 1$.

Recall from Eq. (12) that

$$\begin{aligned}W &= \overline{W}\overline{A}\\ &= \overline{W}\left(I - \gamma\alpha\tilde{F}\tilde{F}^{T}\right)\end{aligned} \tag{18}$$

and

$$K = \alpha\overline{W}\tilde{F}\,. \tag{19}$$

Hence,

$$W = \overline{W} - \gamma K\tilde{F}^{T} \tag{20}$$

which is the measurement update for *W*.

The Potter Computational Algorithm:

Given $\bar{W}_k$, $\bar{x}_k$, y_k, $\tilde{H}_k$ where $\bar{P}_k = \bar{W}_k \bar{W}_k^T$.

Compute

1. $\tilde{F}_k = \bar{W}_k^T \tilde{H}_k^T$

2. $\alpha_k = \left(\tilde{F}_k^T \tilde{F}_k + \sigma^2 \right)^{-1}$

3. $K_k = \alpha_k \bar{W}_k \tilde{F}_k$

4. $\hat{x}_k = \bar{x}_k + K_k \left(y_k - \tilde{H}_k \bar{x}_k \right)$

5. $\gamma_k = \dfrac{1}{1 + \sqrt{\alpha_k \sigma^2}}$

6. $W_k = \bar{W}_k - \gamma_k K_k \tilde{F}_k^T$, $P_k = W_k W_k^T$

7. Integrate the reference orbit and $\dot{\Phi} = A\Phi$ forward to $k+1$

8. Time update W_k and $\hat{x}_k$ to $k+1$

   $$\bar{W}_{k+1} = \Phi\left(t_k, t_{k+1}\right) W_k$$
   $$\bar{x}_{k+1} = \Phi\left(t_k, t_{k+1}\right) \hat{x}_k$$

9. Return to step 1 with $k = k+1$

We have assumed here that the observation vector, y_k, contains a single measurement. If y_k contains more than one data type at each observation time, we would return to step 1 after completing step 6 for each element of y_k. After we have processed all elements of y_k, we would proceed to the time update of step 8.

Note that unlike P, W contains n^2 distinct elements (i.e., W is not symmetric). The Carlson algorithm reduces W to an upper (or lower) triangular matrix (see Bierman, 1977).