Kalman Smoothing

John Kormylo

Kalman smoothing can obtain closed-form least-squares solutions to problems with a very large number of variables, provided they can be formulated using a state vector model. Examples include such things as cubic splines, refraction statics and image processing, as well as any combination of linear ordinary differential equations.

Orthogonality Principle

Consider a general linear estimator of the form

$$\hat{x} = \sum_{i=1}^{N} a_i z(i)$$

where x is a random quantity to estimated, the $z(i)$ are observations presumably related to x, and the a_i are chosen so as to minimize

$$J = \mathcal{E}\left\{(x - \hat{x})^2\right\} \quad .$$

Setting the partial of J with respect to each a_i we get the orthogonality condition

$$\mathcal{E}\left\{(x - \hat{x})z(i)\right\} = 0 \quad \forall \quad i = 1, 2, \ldots, N$$

which holds for each variable no matter how may variables are estimated simultaneously.

State Vector Models

The origin of state vector models comes from the realization that one can express any linear ordinary differential equation as a first order vector differential equation of the form

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) + B\mathbf{u}(t) \tag{1}$$

where $\mathbf{u}(t)$ contains the driving functions. The initial conditions problem is solved by specifying $\mathbf{x}(0)$. The poles of the Laplace transform correspond to the eigenvalues of matrix A.

Since the goal here is digital signal processing, let us move directly to the discrete-time state vector model

$$\mathbf{x}(k+1) = \Phi\mathbf{x}(k) + \Gamma\mathbf{u}(k) \tag{2}$$

and

$$\mathbf{z}(k) = H\mathbf{x}(k) + \mathbf{n}(k) \tag{3}$$

where $\mathbf{z}(k)$ represents the set of observable values at time k and $\mathbf{n}(k)$ represents observation error. We assume that $\mathbf{n}(k)$ is a zero-mean white noise process, so that

$$\mathcal{E}\left\{\mathbf{n}(k)\right\} = \mathbf{0} \tag{4}$$

and

$$\mathcal{E}\left\{\mathbf{n}(i)\mathbf{n}'(j)\right\} = \begin{cases} R & \text{when } i = j \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $'$ denotes transpose. Similarly, driving signal $\mathbf{u}(k)$ can be expressed as

$$\mathbf{u}(k) = \mathbf{b}(k) + \boldsymbol{\mu}(k) \tag{6}$$

where $\mathbf{b}(k)$ is a known bias function and system noise $\boldsymbol{\mu}(k)$ is a zero-mean white process, so that

$$\mathcal{E}\left\{\mathbf{u}(k)\right\} = \mathbf{b}(k) \tag{7}$$

and

$$\mathcal{E}\left\{[\mathbf{u}(i) - \mathbf{b}(i)][\mathbf{u}(j) - \mathbf{b}(j)]'\right\} = \mathcal{E}\left\{\boldsymbol{\mu}(i)\boldsymbol{\mu}'(j)\right\} = \begin{cases} Q & \text{when } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Finally, we assume that the system noise and observation noise are statistically independent from each other and from initial state $\mathbf{x}(0)$, so that

$$\begin{aligned} \mathcal{E}\left\{\mathbf{n}(i)\boldsymbol{\mu}'(j)\right\} &= 0 \quad \forall \quad i, j \\ \mathcal{E}\left\{\mathbf{x}(0)\mathbf{n}'(k)\right\} &= 0 \quad \forall \quad k \geq 0 \\ \mathcal{E}\left\{\mathbf{x}(0)\boldsymbol{\mu}'(k)\right\} &= 0 \quad \forall \quad k \geq 0 \end{aligned} \tag{9}$$

To handle corrolated noise processes, simply include an appropriate noise filter into the state vector model.

The (matrix) impulse response is given by

$$H(k) = H\Phi^k\Gamma \tag{10}$$

and the poles of the Z-transform correspond to the eigenvalues of transition matrix Φ.

Kalman Filter

We use the notation

$$\hat{\mathbf{x}}(k|j) = \mathcal{E}\left\{\mathbf{x}(k) \,|\, \mathbf{z}(0), \mathbf{z}(1), \ldots, \mathbf{z}(j)\right\} \tag{11}$$

and

$$\tilde{\mathbf{x}}(k|j) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|j) \tag{12}$$

to denote the estimate and estimation error. We also represent the covariance matrix of the estimation error using

$$P(k|j) = \mathcal{E}\left\{\tilde{\mathbf{x}}(k|j)\tilde{\mathbf{x}}'(k|j)\right\} \quad . \tag{13}$$

One can easily show that

$$\hat{\mathbf{x}}(k+1|j) = \Phi\hat{\mathbf{x}}(k|j) + \Gamma\mathbf{b}(k) \quad \forall j \leq k \tag{14}$$

satisfies the orthogonality condition

$$\mathcal{E}\left\{\tilde{\mathbf{x}}(k+1|j)\mathbf{z}'(i)\right\} = 0 \quad \forall i \leq j \tag{15}$$

assuming $\hat{\mathbf{x}}(k|j)$ also satisfies the orthogonality condition. Specifically, from (2) and (12) we see that

$$\tilde{\mathbf{x}}(k+1|j) = \Phi\tilde{\mathbf{x}}(k|j) + \Gamma\boldsymbol{\mu}(k) \tag{16}$$

since $\mathbf{x}(j)$ and therefore $\mathbf{z}(j)$ are independent of $\boldsymbol{\mu}(k)$ for $j \leq k$.

In particular we are interested in the case when $j = k$, namely

$$\hat{\mathbf{x}}(k+1|k) = \Phi\hat{\mathbf{x}}(k|k) + \Gamma\mathbf{b}(k) \quad . \tag{17}$$

Similarly, the covariance equation is given by

$$P(k+1|k) = \Phi P(k|k)\Phi' + \Gamma Q\Gamma' \quad . \tag{18}$$

This is called the prediction step of a Kalman filter.

For the correction step of a Kalman filter, assume a linear estimator of the form

$$\hat{\mathbf{x}}(k|k) = A\hat{\mathbf{x}}(k|k-1) + B\mathbf{z}(k) \tag{19}$$

and solve for matrices A and B to satisfy the orthogonality condition

$$\mathcal{E}\left\{\tilde{\mathbf{x}}(k|k)\mathbf{z}'(j)\right\} = 0 \quad \forall \quad j \leq k. \tag{20}$$

From equations (3) and (12) we see that

$$\tilde{\mathbf{x}}(k|k) = (I - BH)\mathbf{x}(k) - A\hat{\mathbf{x}}(k|k-1) - B\mathbf{n}(k)$$

or when $A = (I - BH)$

$$\tilde{\mathbf{x}}(k|k) = (I - BH)\tilde{\mathbf{x}}(k|k-1) - B\mathbf{n}(k)$$

which satisfies (20) for $j < k$. For $j = k$ from (20) we get

$$\mathcal{E}\left\{\tilde{\mathbf{x}}(k|k)\mathbf{z}'(k)\right\} = (I - BH)P(k|k)H' - BR = 0$$

which has the solution

$$B = P(k|k)H'\left[HP(k|k)H' + R\right]^{-1} \doteq K(k) \tag{21}$$

which is the Kalman gain function.

Substituting our solutions back into (19) we get

$$\begin{aligned}\hat{\mathbf{x}}(k|k) &= [I - K(k)H]\hat{\mathbf{x}}(k|k-1) + K(k)\mathbf{z}(k) \\ &= \hat{\mathbf{x}}(k|k-1) + K(k)\tilde{\mathbf{z}}(k|k-1)\end{aligned} \tag{22}$$

where the inovations process $\tilde{\mathbf{z}}(k|k-1)$ given by

$$\tilde{\mathbf{z}}(k|k-1) = \mathbf{z}(k) - H\hat{\mathbf{x}}(k|k-1) = \mathbf{n}(k) + H\tilde{\mathbf{x}}(k|k-1) \tag{23}$$

is an uncorrelated random process (white noise). The covariance for this estimate is given by

$$P(k|k) = [I - K(k)H]P(k|k-1)[I - K(k)H]' + R = [I - K(k)H]P(k|k-1) \quad . \tag{24}$$

Now all one needs is an initial estimate

$$\hat{\mathbf{x}}(0) = \mathcal{E}\left\{\mathbf{x}(0)\right\}$$

and covariance $P(0)$, which is usually chosen as steady state solution

$$P = \Phi P \Phi' + \Gamma R \Gamma'$$

to begin alternating the prediction and corrections steps of the Kalman filter.

Note, when covariance R is diagonal (independent observations), the matrix correction in (21-24)) can be replaced by a series of vector updates. In other words, one can repeat the correction step for each $z_i(k)$ independently and sequentially and avoid the matrix inversion in (21).

Kalman Smoother

We can express the Kalman Smoother as a general linear estimator of the form

$$\hat{\mathbf{x}}(k|N) = \sum_{i=1}^{N} A_i(k)\tilde{\mathbf{z}}(i|i-1) \tag{25}$$

and solve for the matrices $A_i(k)$ to satisfy the orthogonality condition

$$\mathcal{E}\left\{\tilde{\mathbf{x}}(k|N)\tilde{\mathbf{z}}'(i|i-1)\right\} = 0 \quad . \tag{26}$$

Because the inovations process is uncorrelated, one can solve for the $A_i(k)$ independently, and independent of number of observations used.

For example,

$$\hat{\mathbf{x}}(k|k-1) = \sum_{i=1}^{k-1} A_i(k)\tilde{\mathbf{z}}(i|i-1) \tag{27}$$

and therefore

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + A_k(k)\tilde{\mathbf{z}}(k|k-1) \tag{28}$$

where from (22) we see that $A_k(k) = K(k)$. We can also simplify the Kalman smoother to the form

$$\hat{\mathbf{x}}(k|N) = \hat{\mathbf{x}}(k|k-1) + K(k)\tilde{\mathbf{z}}(k|k-1) + \sum_{i=k+1}^{N} A_i(k)\tilde{\mathbf{z}}(i|i-1) \quad . \tag{29}$$

It should also be noted that

$$\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k) = \sum_{i=k+1}^{N} A_i(k+1)\tilde{\mathbf{z}}(i|i-1)$$

is a linear combination of the inovations process with the same range as the sum in (29) which satisfies the orthogonality condition. In fact, solving the respective constants reveals that

$$A_i(k) = P(k|k)\Phi' P^{-1}(k+1|k)A_i(k+1) \quad \forall \quad i > k \tag{30}$$

further simplifying the Kalman smoother to the form

$$\begin{aligned}\hat{\mathbf{x}}(k|N) &= \hat{\mathbf{x}}(k|k-1) + K(k)\tilde{\mathbf{z}}(k|k-1) \\ &\quad + P(k|k)\Phi' P^{-1}(k+1|k)[\hat{\mathbf{x}}(k+1|N) - \hat{\mathbf{x}}(k+1|k)] \quad .\end{aligned} \tag{31}$$

To reduce computations, define residual state vector $\mathbf{r}(k)$ as

$$\mathbf{r}(k) = P^{-1}(k|k-1)[\hat{\mathbf{x}}(k|N) - \hat{\mathbf{x}}(k|k-1)] \tag{32}$$

which can be computed recursively using

$$\mathbf{r}(k) = [I - K(k)H]'\Phi'\mathbf{r}(k+1) + H'[HP(k|k-1)H' + R]^{-1}\tilde{\mathbf{z}}(k|k-1) \tag{33}$$

for $k = N, N-1, \ldots, 0$ starting from $\mathbf{r}(N+1) = \mathbf{0}$.

Smoothed estimates of the input process $\boldsymbol{\mu}(k)$ are obtained by observing that

$$\hat{\mathbf{x}}(k+1|N) = \Phi\hat{\mathbf{x}}(k|N) + \Gamma[\mathbf{b}(k) + \hat{\boldsymbol{\mu}}(k|N)] \tag{34}$$

which produces

$$\hat{\boldsymbol{\mu}}(k|N) = Q\Gamma'\mathbf{r}(k+1) \tag{35}$$

after many substitutions.

It should also be noted that the covariance for the residual state vector

$$S(k) = \mathcal{E}\left\{\mathbf{r}(k)\mathbf{r}'(k)\right\} \tag{36}$$

can be computed recursively using

$$S(k) = [I - K(k)H]'\Phi' S(k+1)\Phi[I - K(k)H] + H'[HP(k|k-1)H' + R]^{-1}H \tag{37}$$

for $k = N, N-1, \ldots, 0$ starting from S(N+1) = 0. The error covariance for the smoothed state vector estimates is given by

$$P(k|N) = P(k|k-1) - P(k|k-1)S(k)P(k|k-1) \tag{38}$$

and the error covariance for smoothed input is given by

$$\mathcal{E}\left\{\tilde{\boldsymbol{\mu}}(k|N)\tilde{\boldsymbol{\mu}}'(k|N)\right\} = Q - Q\Gamma S(k)\Gamma' Q \quad . \tag{39}$$