Kalman Filtering

The goal

Estimate the state of a system (presumably its model is known), given knowledge of the statistics of past inputs and outputs

Method: least squares estimation (innovation method!)

System model: $$\begin{cases} x_{i+1} = A_i x_i + B_i u_i \\ y_i = C_i x_i + v_i \end{cases}$$

The goal: find estimate $\hat{x}_i$ for x_i given observations $y_0, y_1, \ldots, y_{i-1}$
and $\Pi_0 = \mathrm{E} x_0 x_0'$

Second order statistics:

- u_i, v_i noise vectors with positive def. covariances $\mathrm{E} u_i u_i' = Q_i, \mathrm{E} v_i v_i' = R_i$
- all are uncorrelated over all time: $\mathrm{E} u_i u_j' = 0, \mathrm{E} v_i v_j' = 0 \text{ for } i \neq j, \mathrm{E} u_i v_j' = 0 \text{ all } i, j$

Note: matrix correlations:

$$\mathrm{E} x y' = \mathrm{E}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \right) = \left[\mathrm{E} x_i y_j \right]$$

The Kalman solution

$$\hat{x}_{i+1} = A_i \hat{x}_i + K_{p,i}\left(y_i - C_i \hat{x}_i\right)$$

where the "Kalman Gain" $K_{p,i}$

is found by the following recurrence:

$$K_{p,i} = K_i R_{e,i}^{-1}, R_{e,i} = R_i + C_i P_i C_i', K_i = A_i P_i C_i'$$

$$P_{i+1} = A_i P_i A_i' + B_i Q_i B_i' - K_{p,i} R_{e,i} K'_{p,i}$$

herein:

$$e_{x,i} = x_i - \hat{x}_i$$

the state innovation

$$P_i = \mathrm{E} e_{x,i} e'_{x,i} \; , R_{e,i} = \mathrm{E} e_{y,i} e'_{y,i} = R_i + C_i P_i C_i'$$

state and innovation covariances

Initial conditions: $\hat{x}_0 = 0, \quad P_0 = \Pi_0$

Criticism: quadratic recursion, numerically unstable!

Working with innovations and Proof

The “Wiener Principle”: the innovations should be orthogonal on the observed data.

Step from *i* to *i+1*:

Simplification: $e_{x,i} = x_i - \hat{x}_i$ is already orthogonal on $y_0, \cdots, y_{i-1}$

Strategy: let’s try to use only the most recent known data (linear combination):

$$\hat{x}_{i+1} = X_i \hat{x}_i + Y_i y_i$$

for some X_i and Y_i , and let us request that the new innovation is orthogonal just on $\hat{x}_i$ (a linear combination of past data) and the new data (we’ll check that this is sufficient):

$$\mathrm{E} e_{x,i+1} \hat{x}_i' = 0 \quad \Rightarrow \quad X_i = A_i - Y_i C_i$$

and

$$\mathrm{E} e_{x,i+1} y_i' = 0 \quad \Rightarrow \quad Y_i (R_i + C_i P_i C_i') = A_i P_i C_i'$$

now identify $K_{p,i} := Y_i$ and $R_{e,i} := R_i + C_i P_i C_i'$ and observe moreover *(with* $e_{y,i} := y_i - \hat{y}_i$ *)*

$$e_{x,i+1} + K_{p,i} e_{y,i} = A_i e_{x,i} + B_i u_i \Rightarrow P_{i+1} + K_{p,i} R_{e,i} K_{p.i}' = A_i P_i A_i' + B_i Q_i B_i'$$

to reproduce the Kalman formulas

Orthogonality check

Filling out the expressions for x_i and $\hat{x}_i$ we obtain:

$$e_{x,i+1} = (A_i - K_{p,i}C_i)e_{x,i} + B_i u_i - K_{p,i}\nu_i$$

(a useful expression for the innovation!)

Let now *j<i*. We should check that also $\mathrm{E}e_{x,i+1}y'_j = 0$!

Given the hypothesis on the noise, this will be the case if already $\mathrm{E}e_{x,i}y'_j = 0$.
For *i=j+1* this is true by construction. It will hence also be true for
i=j+2 etc... and hence for all *i>j* recursively.

The square root algorithm

Instead of working with the covariances, we work with their "square roots":

write $P_i = R_i^{1/2} R_i^{'/2}$ with $R_i^{1/2}$ lower triangular, and similarly $P_{e,i} = R_{e,i}^{1/2} R_{e.i}^{'/2}$

$$(R_i^{'/2} = (R_i^{1/2})')!$$

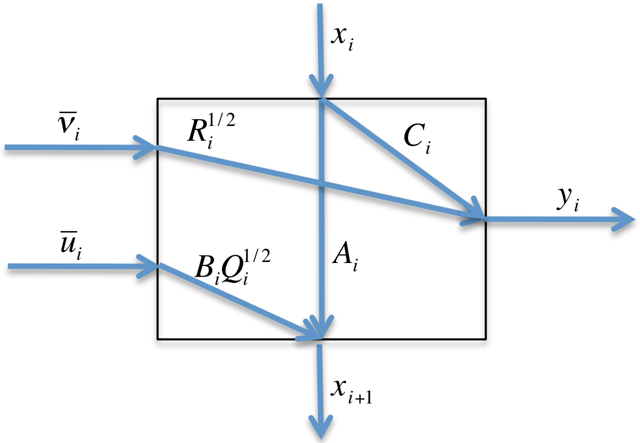
The Kalman formulas then produce the LQ square root algorithm (starting with the first row – and with $\overline{K}_{p,i} = K_{p,i} R_{e,i}^{1/2}$):

$$\begin{bmatrix} C_i P_i^{1/2} & R_i^{1/2} & 0 \\ A_i P_i^{1/2} & 0 & B_i Q_i^{1/2} \end{bmatrix} U_i = \begin{bmatrix} R_{e,i}^{1/2} & 0 & 0 \\ \overline{K}_{p,i} & P_{i+1}^{1/2} & 0 \end{bmatrix}$$

This formula is due to Kailath. Linear and numerically stable!

Connection to Inner-outer

The square root recursion is nothing else than an outer-inner factorization of the original filter:

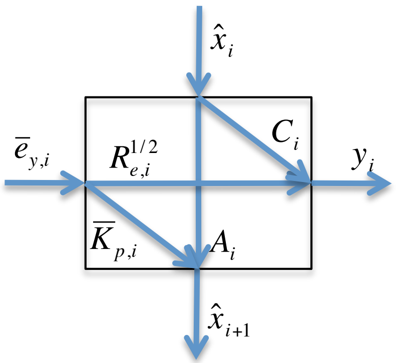


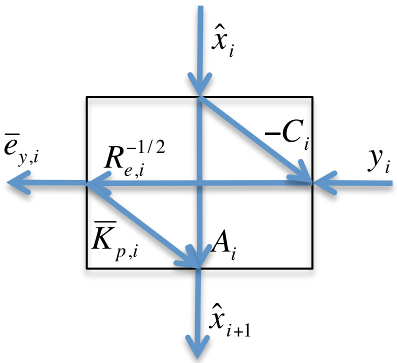
$$\begin{bmatrix} C_iY_i & R_i^{1/2} & 0 \\ A_iY_i & 0 & B_iQ_i^{1/2} \end{bmatrix} U_i = \begin{bmatrix} R_{e,i}^{1/2} & 0 & 0 \\ \overline{K}_{p,i} & Y_{i+1} & 0 \end{bmatrix}$$

(inputs are normalized white noise)

Inverse (estimation and innovation) filter:

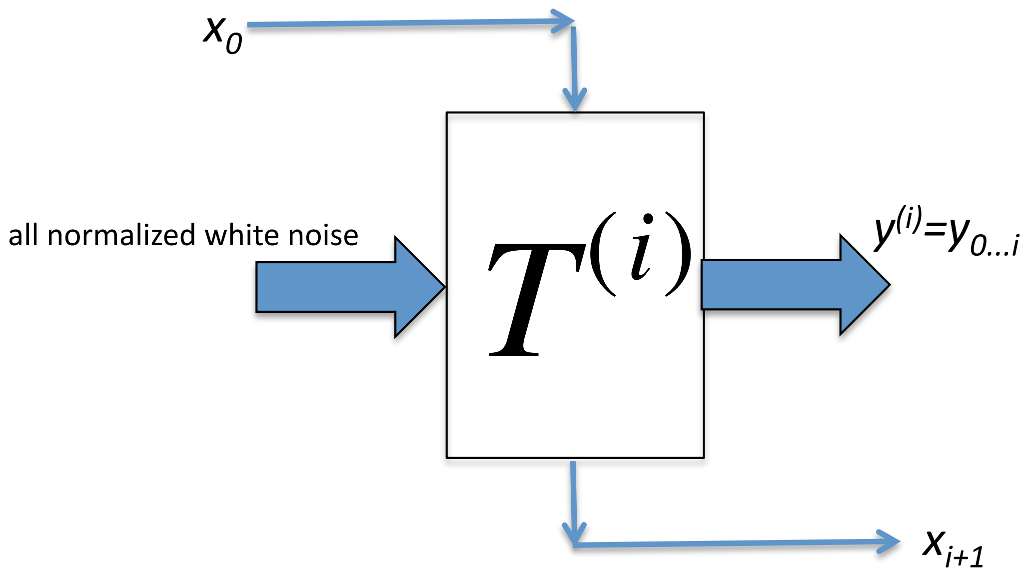
Outer filter:





Outer-inner interpretation

We can view the first *i* stages of the original system as an operator $T^{(i)}$:



The outer-inner factorization gives globally:

$$T^{(i)} = T_o^{(i)} U^{(i)\prime}$$

with $T_o^{(i)}$ causally invertible. Hence the overall covariance:

$$\mathrm{E}y^{(i)}y^{(i)\prime} = T^{(i)}T^{(i)\prime} = T_o^{(i)}T_o^{(i)\prime}$$

and $\left[T_o^{(i)}\right]^{-1} y^{(i)}$ is the overall innovation (a causal filter computing the innovation from the given data).

Extensions

1. Smoothing

    Estimate the state at time point *i* using data up to *i+k*
    (happens often in telecommunications, e.g. GSM)

2. Non-linear systems: use differentials
3. Is it possible to estimate the system?

    Not a well-conditioned problem in general, two solutions:
    1. parametrized model estimation, using a learning sequence
    2. model reduction on a high order model
    If that does not work, estimate some characteristics, and use appropriate statistics (spectral estimation, principal components, independent components). This is more an art than a science!

Main application of Kalman filtering: model based control.