Derivation of the Kalman Filter II

Kai Borre

Danish GPS Center, Denmark

State Updates $\hat{\boldsymbol{x}}_{k|k-1}$ and $\hat{\boldsymbol{x}}_{k|k}$

The left side of the big least-squares problem is now dealt with. We know the last row of $(\mathcal{A}_k^{\mathrm{T}} \Sigma_k^{-1} \mathcal{A}_k)^{-1}$. Now we multiply by the right side $\mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \boldsymbol{\mathscr{b}}_k$, when $\boldsymbol{\mathscr{b}}_k$ includes the newest observation $\boldsymbol{b}_k$.

The predicted value of $\boldsymbol{x}_k$ (before that observation) is simple to understand. Only the state equation has been added to the system. We can solve it exactly by

$$\hat{\boldsymbol{x}}_{k|k-1} = F_{k-1} \hat{\boldsymbol{x}}_{k-1|k-1}. \tag{1}$$

This is our best estimate of $\boldsymbol{x}_k$, based on the state equation and the old observations. It solves the new state equation exactly, and it keeps the best solution to the earlier equations.

So it maintains the correct least-squares solution, when the new row and column are added to the system.

Now we include the new observation. This changes everything. The earlier estimates $\hat{\boldsymbol{x}}_{i|k-1}$ are "smoothed" in the new $\hat{\boldsymbol{x}}_{i|k}$. We leave those smoothing formulas (for $i < k$) until later. The predicted $\hat{\boldsymbol{x}}_{k|k-1}$ in (1) changes to a corrected value $\hat{\boldsymbol{x}}_{k|k}$. This is what we compute now. It is the last component of the weighted least-squares solution to the complete system $\mathcal{A}_k x_k \approx \mathscr{b}_k$:

$$\begin{bmatrix} A_0 & & & & \\ -F_0 & I & & & \\ & A_1 & & & \\ & & \ddots & & \\ & & A_{k-1} & & \\ & & -F_{k-1} & I & \\ & & & A_k \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_0 \\ \vdots \\ \boldsymbol{x}_k \end{bmatrix} \approx \begin{bmatrix} \boldsymbol{b}_0 \\ \boldsymbol{0} \\ \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_{k-1} \\ \boldsymbol{0} \\ \boldsymbol{b}_k \end{bmatrix} = \mathscr{b}_k. \tag{2}$$

The least-squares solution is always $\hat{\boldsymbol{x}} = (\mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \mathcal{A}_k)^{-1} \mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \mathscr{b}_k$. We want the last block $\hat{\boldsymbol{x}}_{k|k}$ in this least-squares solution.

So we use a result from last lecture:

$$\hat{\boldsymbol{x}}_{k|k} = \big(\text{last row of } (\mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \mathcal{A}_k)^{-1}\big) \mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \boldsymbol{\mathcal{b}}_k$$

$$= (I - K_k A_k)\big(\text{last row of } (\mathcal{S}_k^{\mathrm{T}} \Sigma^{-1} \mathcal{S}_k)^{-1}\big) \mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \boldsymbol{\mathcal{b}}_k. \qquad (3)$$

We start with $\boldsymbol{\mathcal{b}}_k$ on the right side, and carry out each multiplication in this equation. Separate the old observations in $\boldsymbol{\mathcal{b}}_{k-1}$ from the new $\boldsymbol{b}_k$, and multiply by Σ^{-1}:

$$\Sigma^{-1} \boldsymbol{\mathcal{b}}_k = \begin{bmatrix} \Sigma_{k-1}^{-1} \boldsymbol{\mathcal{b}}_{k-1} \\ \Sigma_{e,k}^{-1} \boldsymbol{b}_k \end{bmatrix}.$$

Multiply next by $\mathcal{A}_k^{\mathrm{T}} = [\, \mathcal{S}_k^{\mathrm{T}} \quad W^{\mathrm{T}} \,]$ and recall that $W = [\, 0 \quad \ldots \quad 0 \quad A_k \,]$:

$$\mathcal{A}_k^{\mathrm{T}} \Sigma^{-1} \mathscr{b}_k = \mathcal{S}_k^{\mathrm{T}} \Sigma_{k-1}^{-1} \mathscr{b}_{k-1} + A_k^{\mathrm{T}} \Sigma_{e,k}^{-1} \boldsymbol{b}_k. \tag{4}$$

Now multiply by the last row of $(\mathcal{S}_k^{\mathrm{T}} \Sigma^{-1} \mathcal{S}_k)^{-1}$. This produces the least-squares solution $\hat{\boldsymbol{x}}_{k|k-1}$ in the old $k-1$ part. Watch what it produces in the new part:

$$\begin{aligned}
&\left(\text{last row of } (\mathcal{S}_k^{\mathrm{T}} \Sigma^{-1} \mathcal{S}_k)^{-1}\right)\left(\mathcal{S}_k^{\mathrm{T}} \Sigma_{k-1}^{-1} \mathscr{b}_{k-1} + A_k^{\mathrm{T}} \Sigma_{e,k}^{-1} \boldsymbol{b}_k\right) \\
&\qquad = \hat{\boldsymbol{x}}_{k|k-1} + P_{k|k-1} A_k^{\mathrm{T}} \Sigma_{e,k}^{-1} \boldsymbol{b}_k.
\end{aligned} \tag{5}$$

Finally equation (3) multiplies this by $(I - K_k A_k)$ to yield $\hat{\boldsymbol{x}}_{k|k}$:

$$\hat{\boldsymbol{x}}_{k|k} = (I - K_k A_k) \hat{\boldsymbol{x}}_{k|k-1} + K_k \boldsymbol{b}_k. \tag{6}$$

That final term used the identity

$$(I - K_k A_k) P_{k|k-1} A_k^{\mathrm{T}} = K_k \Sigma_{e,k}$$

to replace $(I - K_k A_k) P_{k|k-1} A_k^{\mathrm{T}}$ by $K_k \Sigma_{e,k}$.

This completes the sequence of Kalman filter update equations. As we hoped, formula (6) for $\hat{\boldsymbol{x}}_{k|k}$ can be expressed as the prediction $\hat{\boldsymbol{x}}_{k|k-1}$ plus a correction:

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + K_k (\boldsymbol{b}_k - A_k \hat{\boldsymbol{x}}_{k|k-1}). \tag{7}$$

The correction is Kalman's gain matrix times the ***innovation***.