Kalman Filter using MATLAB
5: Kalman filter derivation

by
Midhun T. Augustine

Overview

1. Kalman filter: basic idea
2. Kalman filter derivation
   - Optimization based approach
   - Statistical approach
3. Kalman filter: Algorithm

Kalman filter: basic idea

Kalman filter: basic idea

- In Kalman filter, the measurements upto k^{th} time instant are used for computing the estimate $\hat{\mathbf{x}}_k$, i.e., $l = k$.
- Consider the stochastic linear system

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{d}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \end{aligned} \tag{1}$$

- Assume that the disturbance, noise and initial state vector as Gaussian with mean $\mathbf{E}(\mathbf{d}_k) = \mathbf{0}, \mathbf{E}(\mathbf{v}_k) = \mathbf{0}$ and $\mathbf{E}(\mathbf{x}_0)$ is known.
- Also, the vectors $\mathbf{x}_k, \mathbf{d}_k, \mathbf{v}_k$ are assumed to be independent:

$$\mathbf{V}(\mathbf{x}_k, \mathbf{d}_k) = \mathbf{0}, \quad \mathbf{V}(\mathbf{x}_k, \mathbf{v}_k) = \mathbf{0}, \quad \mathbf{V}(\mathbf{d}_k, \mathbf{v}_k) = \mathbf{0} \tag{2}$$

- Kalman filter computes the estimate of the state using the following difference equation:

$$\hat{\mathbf{x}}_k = \mathbf{A}_{k-1} \hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1} \mathbf{u}_{k-1} + \mathbf{L}_k [\mathbf{y}_k - \hat{\mathbf{y}}_k]. \tag{3}$$

Kalman filter derivation

Kalman filter derivation

- Let $\mathbf{x}$ and $\mathbf{d}$ are two independent random vectors such that $\mathbf{E}(\mathbf{x}) = \mathbf{0}$, $\mathbf{E}(\mathbf{d}) = \mathbf{0}$, then
$$\begin{aligned} \mathbf{E}(\mathbf{A}\mathbf{x} + \mathbf{d}) &= \mathbf{A}\mathbf{E}(\mathbf{x}) + \mathbf{E}(\mathbf{d}) = \mathbf{0} \\ \mathbf{V}(\mathbf{A}\mathbf{x} + \mathbf{d}) &= \mathbf{A}\mathbf{V}(\mathbf{x})\mathbf{A}^T + \mathbf{V}(\mathbf{d}) \end{aligned} \tag{4}$$
- Similarly, let $\mathbf{x}$, $\mathbf{d}$, $\mathbf{v}$ are three independent random vectors such that $\mathbf{E}(\mathbf{x}) = \mathbf{0}$, $\mathbf{E}(\mathbf{d}) = \mathbf{0}$, $\mathbf{E}(\mathbf{d}) = \mathbf{0}$, then
$$\begin{aligned} \mathbf{E}(\mathbf{A}\mathbf{x} + \mathbf{L}\mathbf{v} + \mathbf{d}) &= \mathbf{A}\mathbf{E}(\mathbf{x}) + \mathbf{L}\mathbf{E}(\mathbf{v}) + \mathbf{E}(\mathbf{d}) = \mathbf{0} \\ \mathbf{V}(\mathbf{A}\mathbf{x} + \mathbf{L}\mathbf{v} + \mathbf{d}) &= \mathbf{A}\mathbf{V}(\mathbf{x})\mathbf{A}^T + \mathbf{L}\mathbf{V}(\mathbf{d})\mathbf{L}^T + \mathbf{V}(\mathbf{d}) \\ \mathbf{V}(\mathbf{A}\mathbf{x}, \mathbf{C}\mathbf{x}) &= \mathbf{A}\mathbf{V}(\mathbf{x})\mathbf{C}^T \end{aligned} \tag{5}$$
- In Kalman estimators we define $\mathbf{x}_{e_k} = \mathbf{x}_k - \hat{\mathbf{x}}_k$, and consider $\mathbf{E}(\mathbf{x}_k) = \hat{\mathbf{x}}_k$. This results in
$$\begin{aligned} \mathbf{E}(\mathbf{x}_{e_k}) &= \mathbf{E}(\mathbf{x}_k) - \hat{\mathbf{x}}_k = \mathbf{0} \\ \mathbf{V}(\mathbf{x}_{e_k}) &= \mathbf{E}\big([\mathbf{x}_{e_k} - \mathbf{E}(\mathbf{x}_{e_k})][\mathbf{x}_{e_k} - \mathbf{E}(\mathbf{x}_{e_k})]^T\big) \\ &= \mathbf{E}(\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T) = \mathbf{E}\big([\mathbf{x}_k - \hat{\mathbf{x}}_k][\mathbf{x}_k - \hat{\mathbf{x}}_k]^T\big) \\ &= \mathbf{V}(\mathbf{x}_k). \end{aligned} \tag{6}$$

Optimization based approach: derivation 1

- For the estimation error vector $\mathbf{x}_{e_k}$, the error dynamics becomes

$$\begin{aligned}
\mathbf{x}_{e_k} =& \mathbf{x}_k - \hat{\mathbf{x}}_k \\
=& \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{d}_{k-1} \\
& - \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} - \mathbf{B}_{k-1}\mathbf{u}_{k-1} - \mathbf{L}_k[\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k - \hat{\mathbf{y}}_k] \\
=& \mathbf{A}_{k-1}\mathbf{x}_{e_{k-1}} + \mathbf{d}_{k-1} - \mathbf{L}_k\big[\mathbf{C}_k[\mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{d}_{k-1}] \\
& + \mathbf{v}_k - \mathbf{C}_k[\mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} - \mathbf{B}_{k-1}\mathbf{u}_{k-1}]\big] \\
=& [\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]\big[\mathbf{A}_{k-1}\mathbf{x}_{e_{k-1}} + \mathbf{d}_{k-1}\big] - \mathbf{L}_k\mathbf{v}_k
\end{aligned} \tag{7}$$

- For which the variance matrix $\mathbf{P}_k = \mathbf{V}(\mathbf{x}_{e_k})$ is obtained as

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]\big[\mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}\big][\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{L}_k\mathbf{R}_k\mathbf{L}_k^T \tag{8}$$

which is the DRE for the Kalman filter.

Optimization based approach: derivation 1

- The cost function for the Kalman filter is chosen as

$$J = E(\mathbf{x}_{e_k}^T \mathbf{x}_{e_k}) = Trace(\mathbf{P}_k) \tag{9}$$

- We have the Kalman filter DRE as

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{L}_k \mathbf{C}_k]\big[\mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}\big][\mathbf{I} - \mathbf{L}_k \mathbf{C}_k]^T + \mathbf{L}_k \mathbf{R}_k \mathbf{L}_k^T \tag{10}$$

- The estimator gain $\mathbf{L}_k$ is chosen to minimize the cost J, which leads to $\frac{\partial Trace(\mathbf{P}_k)}{\partial \mathbf{L}_k} = \mathbf{0}$. This results in

$$\begin{aligned} &-2[\mathbf{I} - \mathbf{L}_k \mathbf{C}_k]\big[\mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}\big]\mathbf{C}_k^T + 2\mathbf{L}_k \mathbf{R}_k = \mathbf{0} \\ &\implies \mathbf{L}_k = \Big[\big[\mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}\big]\mathbf{C}_k^T\Big]\Big[\mathbf{C}_k\big[\mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}\big]\mathbf{C}_k^T + \mathbf{R}_k\Big] \end{aligned} \tag{11}$$

Optimization based approach: derivation 2

- Denote $\hat{\mathbf{x}}_{k|i}$ as the estimate of $\mathbf{x}_k$ computed using the information at i^{th} time instant.
- Using this the Kalman filter eqn can be rewritten as a two-stage process:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k[\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}].\end{aligned} \tag{12}$$

- Define the prediction/forecast error vector as $\mathbf{x}_{e_{k|k-1}} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$ and the error dynamics becomes

$$\begin{aligned}\mathbf{x}_{e_{k|k-1}} &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \\ &= \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{d}_{k-1} - \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{B}_{k-1}\mathbf{u}_{k-1} \\ &= \mathbf{A}_{k-1}\mathbf{x}_{e_{k-1|k-1}} + \mathbf{d}_{k-1}\end{aligned} \tag{13}$$

- For which the variance matrix $\mathbf{P}_{k|k-1} = \mathbf{V}(\mathbf{x}_{e_{k|k-1}})$ is obtained as

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}. \tag{14}$$

Optimization based approach: derivation 2

- Similarly, the estimation error is defined as $\mathbf{x}_{e_{k|k}} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$ and the error dynamics becomes

$$\begin{aligned}\mathbf{x}_{e_{k|k}} &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k} \\ &= \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{L}_k[\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}] \\ &= [\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]\mathbf{x}_{e_{k|k-1}} - \mathbf{L}_k\mathbf{v}_k\end{aligned} \tag{15}$$

- For which the variance matrix $\mathbf{P}_{k|k} = \mathbf{V}(\mathbf{x}_{e_{k|k}})$ is obtained as

$$\begin{aligned}\mathbf{P}_{k|k} &= [\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]\mathbf{P}_{k|k-1}[\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{L}_k\mathbf{R}_k\mathbf{L}_k^T \\ &= \mathbf{P}_{k|k-1} - 2\mathbf{P}_{k|k-1}\mathbf{C}_k^T\mathbf{L}_k^T + \mathbf{L}_k\left[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k\right]\mathbf{L}_k^T\end{aligned} \tag{16}$$

- Then the Kalman gain is obtained from $\frac{\partial Trace(\mathbf{P}_{k|k})}{\partial \mathbf{L}_k} = \mathbf{0}$, which gives

$$\begin{aligned}&-2\mathbf{P}_{k|k-1}\mathbf{C}_k^T + 2\mathbf{L}_k\left[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k\right] = \mathbf{0} \\ &\implies \mathbf{L}_k = \mathbf{P}_{k|k-1}\mathbf{C}_k^T\left[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k\right]^{-1}.\end{aligned} \tag{17}$$

- The Kalman filter equations can also obtained from the conditional expectation equation by defining $\hat{\mathbf{x}}_{k|k} = \mathbf{E}(\mathbf{x}_k|\mathbf{y}_k)$.
- The conditional expectation formula gives

$$
\begin{aligned}
\hat{\mathbf{x}}_{k|k} &= \mathbf{E}(\mathbf{x}_k|\mathbf{y}_k) = \mathbf{E}(\mathbf{x}_k) + \mathbf{L}_k[\mathbf{y}_k - \mathbf{E}(\mathbf{y}_k)] \\
&= \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k[\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}].
\end{aligned} \tag{18}
$$

- In which the estimator gain $\mathbf{L}_k$ is obtained as

$$
\begin{aligned}
\mathbf{L}_k &= \mathbf{V}(\mathbf{x}_k, \mathbf{y}_k)\mathbf{V}(\mathbf{y}_k)^{-1} \\
&= \mathbf{V}(\mathbf{x}_k, \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k)\mathbf{V}(\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k)^{-1} \\
&= \mathbf{P}_{k|k-1}\mathbf{C}_k^T\left[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k\right]^{-1}.
\end{aligned} \tag{19}
$$

Statistical approach

- Similarly, the conditional variance equation gives

$$\begin{aligned}\mathbf{P}_{k|k} &= \mathbf{V}(\mathbf{x}_k|\mathbf{y}_k) = \mathbf{V}(\mathbf{x}_k) - \mathbf{L}_k\mathbf{V}(\mathbf{y}_k, \mathbf{x}_k)\\ &= \mathbf{P}_{k|k-1} - \mathbf{L}_k\mathbf{V}(\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k, \mathbf{x}_k)\\ &= \mathbf{P}_{k|k-1} - \mathbf{L}_k\mathbf{C}_k\mathbf{V}(\mathbf{x}_k) = \mathbf{P}_{k|k-1} - \mathbf{L}_k\mathbf{C}_k\mathbf{P}_{k|k-1} \quad sub.\ \mathbf{L}_k\\ &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]^{-1}\mathbf{C}_k\mathbf{P}_{k|k-1}.\end{aligned} \tag{20}$$

- The DRE obtained from the optimization approach can be rewritten as

$$\begin{aligned}\mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - 2\mathbf{P}_{k|k-1}\mathbf{C}_k^T\mathbf{L}_k^T + \mathbf{L}_k[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]\mathbf{L}_k^T, \quad sub.\ \mathbf{L}_k\\ &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]^{-1}\mathbf{C}_k\mathbf{P}_{k|k-1}.\end{aligned} \tag{21}$$

- For scalar systems, this becomes

$$P_{k|k} = P_{k|k-1} - \frac{P_{k|k-1}^2 C_k^2}{C_k^2 P_{k|k-1} + R_k} \tag{22}$$

which gives $P_{k|k} \leq P_{k|k-1}$.

Statistical approach

- In general

$$
\begin{aligned}
\mathbf{V}(\mathbf{x}_{e_{k|k-1}}) &= \mathbf{P}_{k|k-1} \\
\mathbf{V}(\mathbf{y}_k) &= \mathbf{V}(\mathbf{C}\mathbf{x}_k + \mathbf{v}_k) = \mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k \\
\mathbf{V}(\mathbf{x}_{e_{k|k}}) &= \mathbf{P}_{k|k} \\
&= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]^{-1}\mathbf{C}_k\mathbf{P}_{k|k-1}
\end{aligned}
\tag{23}
$$

- We have $\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k > 0$, which results in

$$
\mathbf{P}_{k|k} \leq \mathbf{P}_{k|k-1} \quad or \quad \mathbf{V}(\mathbf{x}_{e_{k|k}}) \leq \mathbf{V}(\mathbf{x}_{e_{k|k-1}}) \tag{24}
$$

- Therefore, by using the correction term (measurement information) the variance of the state estimate can be reduced.

Kalman filter: Algorithm

Kalman filter: Algorithm

- Define the sets $\mathbb{A} = \{\mathbf{A}_0, \mathbf{A}_1, ..., \mathbf{A}_{N-1}\}, \mathbb{B} = \{\mathbf{B}_0, \mathbf{B}_1, ..., \mathbf{B}_{N-1}\}, \mathbb{C} = \{\mathbf{C}_1, \mathbf{C}_2, ..., \mathbf{C}_N\}, \mathbb{Q} = \{\mathbf{Q}_0, \mathbf{Q}_1, ..., \mathbf{Q}_{N-1}\}, \mathbb{R} = \{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_N\}$.

Algorithm 1 (Kalman filter)

Require $\mathbb{A}.\mathbb{B}, \mathbb{C}, \mathbb{Q}, \mathbb{R}$
Initialize $\hat{\mathbf{x}}_{0|0}$ and $\mathbf{P}_{0|0}$
for $k = 1$ *to* N **do**
  $\mathbf{A}_{k-1} = [\mathbb{A}]_k$, $\mathbf{B}_{k-1} = [\mathbb{B}]_k, \mathbf{C}_k = [\mathbb{C}]_k$, $\mathbf{Q}_{k-1} = [\mathbb{Q}]_k$, $\mathbf{R}_k = [\mathbb{R}]_k$
  $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1}$
  $\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{A}_{k-1}^T + \mathbf{Q}_{k-1}$
  $\mathbf{L}_k = \mathbf{P}_{k|k-1}\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]^{-1}$
  Obtain $\mathbf{y}_k, \hat{\mathbf{y}}_{k|k-1}$ from sensor measurements and the output equation
  $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k[\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}]$
  $\mathbf{P}_{k|k} = [\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]\mathbf{P}_{k|k-1}[\mathbf{I} - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{L}_k\mathbf{R}_k\mathbf{L}_k^T$
end for

Thank you