Kalman Filter using MATLAB
3: Kalman predictor derivation

by
Midhun T. Augustine

Overview

1. Kalman estimators: revisited
2. Kalman predictor derivation
   - Optimization based approach
   - Statistical approach
3. Kalman predictor: Algorithm

Kalman estimators: revisited

Kalman estimators: revisited

- The Kalman estimator gives an estimate of the state vector $\mathbf{x}_k$ in terms of it's expectation and variance, using the available information.
- Let the state estimate $\hat{\mathbf{x}}_k$ at time instant k is computed using the measurement data up to time instant l, i.e., $\{\mathbf{y}_0, \mathbf{y}_1, ..., \mathbf{y}_l\}$.
- Kalman estimator: versions
  1. **Kalman predictor**: for which $l = k - 1$.
  2. **Kalman filter**: for which $l = k$.
  3. **Kalman smoother**: for which $l = N$.
- Approaches used for derivation
  1. **Optimization based approach**: the estimator gain is computed by minimizing a quadratic function of the estimation error.
  2. **Statistical approach**: in which the estimator gain and variance is computed from the conditional expectation and variance equations.

Kalman predictor derivation

Kalman predictor

- In Kalman predictor the measurements upto $k-1^{th}$ time instant are used for computing the estimate $\hat{\mathbf{x}}_k$, i.e., $l = k-1$.
- 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:

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

- The Kalman predictor is defined as

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

in which, $\mathbf{A}_k \hat{\mathbf{x}}_k + \mathbf{B}_k \mathbf{u}_k$ is known as the prediction/forecast term, and $\mathbf{y}_k - \hat{\mathbf{y}}_k$ is known as the correction/innovation term.

Kalman predictor 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

$$\begin{aligned} &\mathbf{E}(\mathbf{Ax} + \mathbf{d}) = \mathbf{AE}(\mathbf{x}) + \mathbf{E}(\mathbf{d}) = \mathbf{0}, \\ &\mathbf{E}(\mathbf{x}\mathbf{d}^T) = \mathbf{0}, \qquad \mathbf{E}(\mathbf{d}\mathbf{x}^T) = \mathbf{0} \end{aligned} \tag{4}$$

- Using which $\mathbf{V}(\mathbf{Ax} + \mathbf{d})$ is obtained as

$$\begin{aligned} \mathbf{V}(\mathbf{Ax} + \mathbf{d}) &= \mathbf{E}\big([\mathbf{Ax} + \mathbf{d} - \mathbf{E}(\mathbf{Ax} + \mathbf{d})][\mathbf{Ax} + \mathbf{d} - \mathbf{E}(\mathbf{Ax} + \mathbf{d})]^T\big) \\ &= \mathbf{E}([\mathbf{Ax} + \mathbf{d}][\mathbf{Ax} + \mathbf{d}]^T) \\ &= \mathbf{E}\big(\mathbf{A}\mathbf{x}\mathbf{x}^T\mathbf{A}^T + \mathbf{A}\mathbf{x}\mathbf{d}^T + \mathbf{d}\mathbf{x}^T\mathbf{A}^T + \mathbf{d}\mathbf{d}^T\big) \\ &= \mathbf{AE}(\mathbf{x}\mathbf{x}^T)\mathbf{A}^T + \mathbf{AE}(\mathbf{x}\mathbf{d}^T) + \mathbf{E}(\mathbf{d}\mathbf{x}^T)\mathbf{A}^T + \mathbf{E}(\mathbf{d}\mathbf{d}^T) \\ &= \mathbf{AV}(\mathbf{x})\mathbf{A}^T + \mathbf{V}(\mathbf{d}) \end{aligned} \tag{5}$$

Optimization based approach

- Define the error vector $\mathbf{x}_{e_k} = \mathbf{x}_k - \hat{\mathbf{x}}_k$, and the error dynamics becomes

$$\begin{aligned}\mathbf{x}_{e_{k+1}} &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1} = \mathbf{A}_k\mathbf{x}_k + \mathbf{B}_k\mathbf{u}_k + \mathbf{d}_k - \mathbf{A}_k\hat{\mathbf{x}}_k - \mathbf{B}_k\mathbf{u}_k - \mathbf{L}_k[\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}_k\hat{\mathbf{x}}_k] \\ &= [\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{x}_{e_k} + \mathbf{d}_k - \mathbf{L}_k\mathbf{v}_k.\end{aligned} \tag{6}$$

- From which the variance of the estimation error is obtained as

$$\begin{aligned}\mathbf{V}(\mathbf{x}_{e_{k+1}}) &= \mathbf{E}\big([\mathbf{x}_{e_{k+1}} - \mathbf{E}(\mathbf{x}_{e_{k+1}})][\mathbf{x}_{e_{k+1}} - \mathbf{E}(\mathbf{x}_{e_{k+1}})]^T\big) = \mathbf{E}(\mathbf{x}_{e_{k+1}}\mathbf{x}_{e_{k+1}}^T) \\ &= \mathbf{E}\big([[\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{x}_{e_k} + \mathbf{d}_k - \mathbf{L}_k\mathbf{v}_k]\,[[\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{x}_{e_k} + \mathbf{d}_k - \mathbf{L}_k\mathbf{v}_k]^T\big) \\ &= \mathbf{E}\big([\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T[\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{d}_k\mathbf{d}_k^T + \mathbf{L}_k\mathbf{v}_k\mathbf{v}_k^T\mathbf{L}_k^T\big) \\ &\qquad\qquad \because \mathbf{E}(\mathbf{x}_{e_k}\mathbf{d}_k^T) = \mathbf{E}(\mathbf{x}_{e_k}\mathbf{v}_k^T) = \mathbf{E}(\mathbf{d}_k\mathbf{v}_k^T) = \mathbf{0} \\ &= [\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{V}(\mathbf{x}_{e_k})[\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{V}(\mathbf{d}_k) + \mathbf{L}_k\mathbf{V}(\mathbf{v}_k)\mathbf{L}_k^T.\end{aligned} \tag{7}$$

- Define $\mathbf{P}_k = \mathbf{V}(\mathbf{x}_{e_k})$, $\mathbf{Q}_k = \mathbf{V}(\mathbf{d}_k)$ and $\mathbf{R}_k = \mathbf{V}(\mathbf{v}_k)$ which gives

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

which is the Difference Riccati Equation (DRE) for the Kalman predictor.

Optimization based approach

- In Kalman predictor, the cost is chosen as a quadratic function of the estimation error (also known as the mean square error cost function):

$$
\begin{aligned}
J &= E(\mathbf{x}_{e_{k+1}}^T \mathbf{x}_{e_{k+1}}) = E(Trace(\mathbf{x}_{e_{k+1}} \mathbf{x}_{e_{k+1}}^T)) \\
&= Trace(\mathbf{E}(\mathbf{x}_{e_{k+1}} \mathbf{x}_{e_{k+1}}^T)) = Trace(\mathbf{P}_{k+1}).
\end{aligned} \tag{9}
$$

- The Kalman gain $\mathbf{L}_k$ is cosen to minimize the cost J, and this leads to $\frac{\partial Trace(\mathbf{P}_{k+1})}{\partial \mathbf{L}_k} = \mathbf{0}$, which results in

$$
\begin{aligned}
&-2[\mathbf{A}_k - \mathbf{L}_k \mathbf{C}_k]\mathbf{P}_k \mathbf{C}_k^T + 2\mathbf{L}_k \mathbf{R}_k = \mathbf{0} \\
&\implies \mathbf{L}_k = \mathbf{A}_k \mathbf{P}_k \mathbf{C}_k^T [\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T + \mathbf{R}_k]^{-1}.
\end{aligned} \tag{10}
$$

Statistical approach

- This approach uses the conditional expectation and variance equations for computing the Kalman predictor gain and DRE.
- Firstly, the conditional expectation equation is used to compute $\hat{\mathbf{x}}_k = \mathbf{E}(\mathbf{x}_k|\mathbf{y}_{k-1})$, which reults in

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \mathbf{E}(\mathbf{x}_{k+1}|\mathbf{y}_k) = \mathbf{E}(\mathbf{x}_{k+1}) + \mathbf{L}_k[\mathbf{y}_k - \mathbf{E}(\mathbf{y}_k)] \\ &= \mathbf{A}_k\hat{\mathbf{x}}_k + \mathbf{B}_k\mathbf{u}_k + \mathbf{L}_k[\mathbf{y}_k - \hat{\mathbf{y}}_k].\end{aligned} \tag{11}$$

- In which $\mathbf{L}_k$ is given by

$$\begin{aligned}\mathbf{L}_k &= \mathbf{V}(\mathbf{x}_{k+1}, \mathbf{y}_k)\mathbf{V}(\mathbf{y}_k)^{-1} \\ &= \mathbf{V}(\mathbf{A}_k\mathbf{x}_k + \mathbf{B}_k\mathbf{u}_k + \mathbf{d}_k, \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k)\mathbf{V}(\mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k)^{-1} \\ &= \mathbf{E}(\mathbf{A}_k\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T\mathbf{C}_k^T)\mathbf{E}(\mathbf{C}_k\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T\mathbf{C}_k^T + \mathbf{R}_k)^{-1} \\ &= \mathbf{A}_k\mathbf{P}_k\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_k\mathbf{C}_k^T + \mathbf{R}_k]^{-1}.\end{aligned} \tag{12}$$

Statistical approach

- Similarly, denote $\mathbf{P}_{k+1} = \mathbf{V}(\mathbf{x}_{k+1}|\mathbf{y}_k)$ which can be derived from the conditional variance equation:

$$\begin{aligned}
\mathbf{P}_{k+1} &= \mathbf{V}(\mathbf{x}_{k+1}|\mathbf{y}_k) = \mathbf{V}(\mathbf{x}_{k+1}) - \mathbf{L}_k\mathbf{V}(\mathbf{y}_k, \mathbf{x}_{k+1}) \\
&= \mathbf{E}(\mathbf{A}_k\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T\mathbf{A}_k^T + \mathbf{Q}_k) - \mathbf{L}_k\mathbf{E}(\mathbf{C}_k\mathbf{x}_{e_k}\mathbf{x}_{e_k}^T\mathbf{A}_k^T) \\
&= \mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T + \mathbf{Q}_k - \mathbf{A}_k\mathbf{P}_k\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_k\mathbf{C}_k^T + \mathbf{R}_k]^{-1}\mathbf{C}_k\mathbf{P}_k\mathbf{A}_k^T.
\end{aligned} \tag{13}$$

- This matches with the DRE obtained from the optimization based approach, since rewriting (8) results in

$$\begin{aligned}
\mathbf{P}_{k+1} &= [\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]\mathbf{P}_k[\mathbf{A}_k - \mathbf{L}_k\mathbf{C}_k]^T + \mathbf{Q}_k + \mathbf{L}_k\mathbf{R}_k\mathbf{L}_k^T \\
&= \mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T + \mathbf{Q}_k - 2\mathbf{L}_k\mathbf{C}_k\mathbf{P}_k\mathbf{A}_k^T + \mathbf{L}_k[\mathbf{C}_k\mathbf{P}_k\mathbf{C}_k^T + \mathbf{R}_k]\mathbf{L}_k^T \qquad sub.\ \mathbf{L}_k \\
&= \mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T + \mathbf{Q}_k - \mathbf{A}_k\mathbf{P}_k\mathbf{C}_k^T[\mathbf{C}_k\mathbf{P}_k\mathbf{C}_k^T + \mathbf{R}_k]^{-1}\mathbf{C}_k\mathbf{P}_k\mathbf{A}_k^T
\end{aligned} \tag{14}$$

Kalman predictor: Algorithm

Kalman predictor: Algorithm

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

Algorithm 1 (Kalman predictor)

Require $\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{Q}, \mathbb{R}$
Initialize $\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$
for $k = 0$ *to* $N - 1$ **do**
    $\mathbf{A}_k = [\mathbb{A}]_{k+1}$, $\mathbf{B}_k = [\mathbb{B}]_{k+1}, \mathbf{C}_k = [\mathbb{C}]_{k+1}$, $\mathbf{Q}_k = [\mathbb{Q}]_{k+1}$, $\mathbf{R}_k = [\mathbb{R}]_{k+1}$
    $\mathbf{L}_k = \mathbf{A}_k \mathbf{P}_k \mathbf{C}_k^T [\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T + \mathbf{R}_k]^{-1}$
    Obtain $\mathbf{y}_k, \hat{\mathbf{y}}_k$ from sensor measurements and the output equation
    Obtain $\mathbf{u}_k$ from the control law (i.e. $\mathbf{u}_k = -\mathbf{K}\hat{\mathbf{x}}_k$)
    $\hat{\mathbf{x}}_{k+1} = \mathbf{A}_k \hat{\mathbf{x}}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{L}_k [\mathbf{y}_k - \hat{\mathbf{y}}_k]$
    $\mathbf{P}_{k+1} = [\mathbf{A}_k - \mathbf{L}_k \mathbf{C}_k] \mathbf{P}_k [\mathbf{A}_k - \mathbf{L}_k \mathbf{C}_k]^T + \mathbf{Q}_k + \mathbf{L}_k \mathbf{R}_k \mathbf{L}_k^T$
end for

Thank you