Basic Kalman Filtering

İlker Bayram
ibayram@ieee.org

This note presents a derivation of the classical Kalman filter. The discussion roughly follows that in Särkkä’s book [1].

Consider a model of the form

\begin{align}
    x_k &= A_{k-1} x_{k-1} + q_{k-1}, \\
    y_k &= H_k x_k + r_k,
\end{align}

where \( q_k \sim \mathcal{N}(0, Q_k) \), and \( r_k \sim \mathcal{N}(0, R_k) \).

Suppose we know that \( x_{k-1} | y_{1:k-1} \sim \mathcal{N}(m_{k-1}, P_{k-1}) \). We would like to find the distribution of \( x_k | y_{1:k} \). Thanks to linearity in the model, we know that this is also a Gaussian, i.e., \( x_k | y_{1:k} \sim \mathcal{N}(m_k, P_k) \). Therefore, given a new observation \( y_k \), we seek a relation between \( (m_k, P_k) \) and \( (m_{k-1}, P_{k-1}) \). The Kalman filter provides this relation.

The plan is to first find the joint distribution of \( (x_k, y_k) | y_{1:k-1} \), which is, again, Gaussian. Given this distribution, we obtain the distribution of \( x_k | y_{1:k} \) by conditioning on \( y_k | y_{1:k-1} \), where the latter distribution is given by marginalizing the joint Gaussian distribution \( (x_k, y_k) | y_{1:k-1} \).

The last conditioning step can be achieved using the following fact about Gaussian random variables.

**Lemma 1.** Suppose \( (x_k, y_k) | y_{1:k-1} \) is distributed as

\[
    \mathcal{N}( \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} A & C \\ C^T & B \end{bmatrix} ).
\]

Then it follows that \( x_k | y_{1:k} \sim \mathcal{N}(m_k, P_k) \) with

\begin{align}
    m_k &= a + C B^{-1} (y_k - b), \\
    P_k &= A - C B^{-1} C^T.
\end{align}

Instead of proving this, consider an equivalent statement with a slightly simpler notation.

**Lemma 2.** Suppose \( (x, y) \) is distributed as

\[
    \mathcal{N}( \begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} \Sigma_x & C \\ C^T & \Sigma_y \end{bmatrix} ).
\]

Then

\[
    x | y = y^0 \sim \mathcal{N}(m_x + C \Sigma_y^{-1} (y^0 - m_y), \Sigma_x - C \Sigma_y^{-1} C^T).
\]

**Proof.** To prove this statement, consider first a special case. Suppose \( x = y_1 + y_2 \), where \( y_1 \sim \mathcal{N}(m_1, \Sigma_1) \) and \( y_2 \sim \mathcal{N}(m_2, \Sigma_2) \) are independent. Then, given an observation \( y_1 = y^0 \), we have \( x | y_1 = y^0 \sim \mathcal{N}(y^0 + m_2, \Sigma_2) \).
Consider now the original distribution in (5). In general, we can find a matrix $W$, such that for $x$ expressed as

$$x = Wy + z,$$

(7)

$z$ is also Gaussian and independent of $y$. By the orthogonality principle, this is equivalent to

$$\mathbb{E}((x - m_x) - W(y - m_y)) (y - m_y)^T = 0.$$

(8)

Expanding and plugging in the terms from the joint the covariance matrix of $(x, y)$, this gives

$$C - W\Sigma_y = 0 \iff W = C\Sigma_y^{-1}.$$

(9)

Using this, we find the expression for $x$ as $x = c - C\Sigma_y^{-1}y$. First, note that

$$\mathbb{E}(z) = m_x - C\Sigma_y^{-1}m_y.$$

(10)

Since $z$ and $Wy$ are independent, we should have

$$\text{cov}(z) + \text{cov}(Wy) = \text{cov}(x) \iff \text{cov}(z) = \Sigma_x - W\Sigma_y WT = \Sigma_x - C\Sigma_y^{-1}CT.$$

(11)

Therefore, using the observation in the beginning of the proof ($y_1$ replaced by $Wy$), we find

$$x|y = y^0 \sim \mathcal{N}(m_x + C\Sigma_y^{-1}(y^0 - m_y), \Sigma_x - C\Sigma_y^{-1}CT).$$

(12)

Given Lemma 1, the problem boils down to determining the joint distribution $(x_k, y_k)|y_{1:k-1}$ as a function of $(m_{k-1}, P_{k-1})$.

We get to the joint distribution $(x_k, y_k)|y_{1:k-1}$ in two steps. First, we consider the distribution $x_k|y_{1:k-1}$. Note that, with the usual abuse of notation regarding pdfs of random variables\(^1\),

$$p(x_k|y_{1:k-1}) = \int p(x_k, x_{k-1}|y_{1:k-1}) dx_{k-1}$$

(13)

$$= \int p(x_k|x_{k-1}, y_{1:k-1}) p(x_{k-1}|y_{1:k-1}) dx_{k-1}$$

(14)

$$= \int p(x_k|x_{k-1}) p(x_{k-1}|y_{1:k-1}) dx_{k-1}.$$  

(15)

where in the last step, we used the fact that given $x_{k-1}$, $x_k$ is independent of previous observations. Note that the last integral is just the marginal of the joint distribution of $(x_k, x_{k-1})$, where $x_{k-1}$’s distribution is replaced by that of $x_{k-1}|y_{1:k-1}$, which is $\mathcal{N}(m_{k-1}, P_{k-1})$. Thus, using (1), we find

$$x_k|y_{1:k-1} \sim \mathcal{N}(\hat{m}_k, \hat{P}_k),$$

(16)

with

$$\hat{m}_k = A_{k-1} m_{k-1},$$

(17)

$$\hat{P}_k = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1}.$$  

(18)

Here, the auxiliary parameters with hats may be interpreted as encoding our prediction of the current state’s distribution given past observations.

---

\(^1\)By this, I mean the practice of representing the pdf of a random variable $X$ as $p(x)$, rather than something like $p_X(\cdot)$. 
We can now write

\[ p(x_k, y_k|y_{1:k-1}) = p(y_k|x_k, y_{1:k-1}) p(x_k|y_{1:k-1}) \]

\[ = p(y_k|x_k) p(x_k|y_{1:k-1}), \]  \hfill (19)

Therefore, the joint distribution \((y_k, x_k)|y_{1:k-1}\) can be written as in (3) with

\begin{align*}
    a &= \hat{m}_k, \quad \hfill (24) \\
    b &= H_k \hat{m}_k, \quad \hfill (25) \\
    A &= \hat{P}_k, \quad \hfill (26) \\
    B &= H_k \hat{P}_k H_k^T + R_k, \quad \hfill (27) \\
    C &= \hat{P}_k H_k^T. \quad \hfill (28)
\end{align*}

Using (4), we thus obtain

\begin{align*}
    m_k &= \hat{m}_k + \hat{P}_k H_k^T (H_k \hat{P}_k H_k^T + R_k)^{-1} (y_k - H_k \hat{m}_k), \quad \hfill (29) \\
    P_k &= \hat{P}_k - \hat{P}_k H_k^T (H_k \hat{P}_k H_k^T + R_k)^{-1} H_k \hat{P}_k. \quad \hfill (30)
\end{align*}

To summarize, the Kalman filter consists of the following operations:

**Algorithm 1 Kalman Update**

**Require:** Given a state update equation as in (1), and an observation equation as in (2), we have the matrices \(A_{k-1}, H_k\), and the state update covariance matrix \(Q_{k-1}\), and the observation noise covariance matrix \(R_k\). We also know that \(x_{k-1}|y_{1:k-1} \sim N(m_{k-1}, P_{k-1})\). The goal is to compute \((m_k, P_k)\).

1: \(\hat{m}_k \leftarrow A_{k-1} m_{k-1}\) \%compute \(\mathbb{E}(x_k|y_{1:k-1})\)
2: \(\hat{P}_k \leftarrow A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1}\) \%compute \(\text{var}(x_k|y_{1:k-1})\)
3: \(m_k \leftarrow \hat{m}_k + \hat{P}_k H_k^T (H_k \hat{P}_k H_k^T + R_k)^{-1} (y_k - H_k \hat{m}_k)\)
4: \(P_k \leftarrow \hat{P}_k - \hat{P}_k H_k^T (H_k \hat{P}_k H_k^T + R_k)^{-1} H_k \hat{P}_k\)

**References**