

Derivation of the Kalman filter equations

Joaquín Rapela

May 28, 2025

Theorem 1. *Given the linear dynamical systems model*

$$\begin{aligned}\mathbf{x}_{t+1} &= A_t \mathbf{x}_t + \mathbf{w}_t && \text{with } \mathbf{w}_t \sim N(0, Q_t) \\ \mathbf{y}_t &= B_t \mathbf{x}_t + \mathbf{v}_t && \text{with } \mathbf{v}_t \sim N(0, R_t) \\ \mathbf{x}_0 &\sim N(\mathbf{m}_0, V_0)\end{aligned}$$

(represented in Fig. 1), then the predictive distribution, $p(\mathbf{x}_t | \mathbf{y}_1, \dots, \mathbf{y}_{t-1})$, and the filtering distribution, $p(\mathbf{x}_t | \mathbf{y}_1, \dots, \mathbf{y}_t)$, are

$$\begin{aligned}p(\mathbf{x}_t | \mathbf{y}_1, \dots, \mathbf{y}_{t-1}) &= N(\mathbf{x}_t | \mathbf{x}_{t|t-1}, P_{t|t-1}) \\ p(\mathbf{x}_t | \mathbf{y}_1, \dots, \mathbf{y}_t) &= N(\mathbf{x}_t | \mathbf{x}_{t|t}, P_{t|t})\end{aligned}$$

with

$$\mathbf{x}_{t|t-1} = A_{t-1} \mathbf{x}_{t-1|t-1} \tag{1}$$

$$P_{t|t-1} = A_{t-1} P_{t-1|t-1} A_{t-1}^\top + Q_{t-1} \tag{2}$$

$$\hat{\mathbf{y}}_{t|t-1} \triangleq E\{\mathbf{y}_t | \mathbf{y}_1, \dots, \mathbf{y}_{t-1}\} = B_t \mathbf{x}_{t|t-1} \tag{3}$$

$$\mathbf{z}_t \triangleq \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$$

$$S_t \triangleq \text{Cov}\{\mathbf{z}_t | \mathbf{y}_1, \dots, \mathbf{y}_{t-1}\} = B_t P_{t|t-1} B_t^\top + R_t \tag{4}$$

$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t-1} + K_t \mathbf{z}_t \tag{5}$$

$$\mathbf{P}_{t|t} = (I - K_t B_t) P_{t|t-1} \tag{6}$$

$$\mathbf{K}_t = P_{t|t-1} B_t^\top S_t^{-1} \tag{7}$$

$$\mathbf{x}_{0|0} = \mathbf{m}_0 \tag{8}$$

$$P_{0|0} = V_0 \tag{9}$$

The following proof adds a few details to that given in Section 4.3.1 of [Durbin and Koopman \(2012\)](#).

Proof. Call $Y_t = \{\mathbf{y}_1, \dots, \mathbf{y}_t\}$, then

$$\mathbf{x}_{t|t-1} = E\{\mathbf{x}_t | Y_{t-1}\} = E\{A_{t-1} \mathbf{x}_{t-1} + \mathbf{w}_{t-1} | Y_{t-1}\}$$

$$\begin{aligned}
&= A_{t-1}E\{\mathbf{x}_{t-1}|Y_{t-1}\} + E\{\mathbf{w}_{t-1}|Y_{t-1}\} \\
&= A_{t-1}\mathbf{x}_{t-1|t-1} + E\{\mathbf{w}_{t-1}\} = A_{t-1}\mathbf{x}_{t-1|t-1}
\end{aligned} \tag{10}$$

This proves Eq. 1.

$$\begin{aligned}
\mathbf{P}_{t|t-1} &= \text{Cov}\{\mathbf{x}_t|Y_{t-1}\} = E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top|Y_{t-1}\} \\
&= E\{(A_{t-1}\mathbf{x}_{t-1} + w_{t-1} - A_{t-1}\mathbf{x}_{t-1|t-1})(A_{t-1}\mathbf{x}_{t-1} + w_{t-1} - A_{t-1}\mathbf{x}_{t-1|t-1})^\top|Y_{t-1}\} \\
&= E\{(A_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1}) + w_{t-1})(A_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1}) + w_{t-1})^\top|Y_{t-1}\} \\
&= A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top|Y_{t-1}\}A_{t-1}^\top + \\
&\quad E\{w_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top|Y_{t-1}\}A_{t-1}^\top + \\
&\quad A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})w_{t-1}^\top|Y_{t-1}\} + \\
&\quad E\{w_{t-1}w_{t-1}^\top|Y_{t-1}\} \\
&= A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top|Y_{t-1}\}A_{t-1}^\top + \\
&\quad E\{w_{t-1}|Y_{t-1}\}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top|Y_{t-1}\}A_{t-1}^\top + \\
&\quad A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})|Y_{t-1}\}E\{w_{t-1}^\top|Y_{t-1}\} + \\
&\quad E\{w_{t-1}w_{t-1}^\top\} \\
&= A_{t-1}P_{t-1|t-1}A_{t-1}^\top + Q_{t-1}
\end{aligned} \tag{11}$$

$$\tag{12}$$

This proves Eq. 2.

$$\begin{aligned}
\hat{\mathbf{y}}_{t|t-1} &= E\{\mathbf{y}_t|Y_{t-1}\} = E\{B_t\mathbf{x}_t + \mathbf{v}_t|Y_{t-1}\} = B_tE\{\mathbf{x}_t|Y_{t-1}\} + E\{\mathbf{v}_t|Y_{t-1}\} \\
&= B_t\mathbf{x}_{t|t-1} + E\{\mathbf{v}_t\} = B_t\mathbf{x}_{t|t-1}
\end{aligned}$$

This proves Eq. 3.

Because

$$\mathbf{z}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} = B_t\mathbf{x}_t + \mathbf{v}_t - B_t\mathbf{x}_{t|t-1} = B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t \tag{13}$$

Y_{t-1} and \mathbf{z}_t are fixed if and only if Y_t is fixed¹. Then

$$\begin{aligned}
\mathbf{x}_{t|t} &= E\{\mathbf{x}_t|Y_t\} = E\{\mathbf{x}_t|Y_{t-1}, \mathbf{z}_t\} \\
&= E\{\mathbf{x}_t|Y_{t-1}\} + \text{Cov}(\mathbf{x}_t, \mathbf{z}_t|Y_{t-1}) \text{Cov}(\mathbf{z}_t|Y_{t-1})^{-1} \mathbf{z}_t
\end{aligned} \tag{14}$$

$$\begin{aligned}
\text{Cov}(\mathbf{x}_t, \mathbf{z}_t|Y_{t-1}) &= \text{Cov}(\mathbf{x}_t, B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t|Y_{t-1}) \\
&= E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t)^\top|Y_{t-1}\} \\
&= E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top|Y_{t-1}\}B_t^\top \\
&\quad + E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})\mathbf{v}_t^\top|Y_{t-1}\} \\
&= P_{t|t-1}B_t^\top
\end{aligned} \tag{15}$$

$$S_t = \text{Cov}(\mathbf{z}_t|Y_{t-1}) = E\{\mathbf{z}_t\mathbf{z}_t^\top|Y_{t-1}\}$$

¹If we now Y_{t-1} and \mathbf{z}_t , then we know $\hat{\mathbf{y}}_{t|t-1}$ and \mathbf{z}_t , then (by the first equality in Eq. 13) we know \mathbf{y}_t , thus we know Y_t . Also, if we know Y_t , we know $\hat{\mathbf{y}}_{t|t-1}$ and \mathbf{y}_t and (by the first equality in Eq. 13) we know \mathbf{z}_t .

$$\begin{aligned}
&= E\{(B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t) (B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t)^\top | Y_{t-1}\} \\
&= B_t E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1}) (\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top | Y_{t-1}\} B_t^\top \\
&\quad + E\{\mathbf{v}_t \mathbf{v}_t^\top | Y_{t-1}\} \\
&= B_t P_{t|t-1} B_t^\top + R_t
\end{aligned} \tag{16}$$

This proves Eq. 4.

Combining Eqs. 14, 15 and 16 we obtain

$$\begin{aligned}
\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + P_{t|t-1} B_t^\top S_t^{-1} \mathbf{z}_t \\
&= \mathbf{x}_{t|t-1} + K_t \mathbf{z}_t \quad \text{with } K_t = P_{t|t-1} B_t^\top S_t^{-1} \\
P_{t|t} &= \text{Cov}(\mathbf{x}_t | Y_t) = \text{Cov}(\mathbf{x}_t | Y_{t-1}, \mathbf{z}_t) = P_{t|t-1} - P_{t|t-1} B_t^\top S_t^{-1} B_t P_{t|t-1} \\
&= (I - P_{t|t-1} B_t^\top S_t^{-1} B_t) P_{t|t-1} = (I - K_t B_t) P_{t|t-1}
\end{aligned} \tag{17}$$

This proves Eqs. 5, 6 and 7.

Using Eqs. 8 and 9 in Eqs. 1 and 2 we obtain

$$\begin{aligned}
\mathbf{x}_{1|0} &= A_0 \mathbf{x}_{0|0} = A_0 \mathbf{m}_0 \\
\mathbf{P}_{1|0} &= A_0 P_{0|0} A_0^\top + Q_0 = A_0 V_0 A_0^\top + Q_0
\end{aligned}$$

If Eqs. 8 and 9 are correct, then the density of \mathbf{x}_1 should be $p(\mathbf{x}_1) = \mathcal{N}(\mathbf{x}_1 | \mathbf{x}_{1|0}, P_{1|0})$. We now calculate this density using the linear dynamical system model in Theorem 1.

$$\begin{aligned}
p(\mathbf{x}_1) &= \int p(\mathbf{x}_1, \mathbf{x}_0) d\mathbf{x}_0 = \int p(\mathbf{x}_1 | \mathbf{x}_0) p(\mathbf{x}_0) d\mathbf{x}_0 = \int \mathcal{N}(\mathbf{x}_1 | A_0 \mathbf{x}_0, Q_0) \mathcal{N}(\mathbf{x}_0 | \mathbf{m}_0, V_0) d\mathbf{x}_0 \\
&= \mathcal{N}(\mathbf{x}_1 | A_0 \mathbf{m}_0, A_0 V_0 A_0^\top + Q_0) = \mathcal{N}(\mathbf{x}_1 | \mathbf{x}_{1|0}, \mathbf{P}_{1|0})
\end{aligned} \tag{18}$$

This proves Eqs. 8 and 9.

□

Notes:

1. the first equality in Eq. 10 holds because \mathbf{w}_{t-1} is independent of Y_{t-1} .
2. the second and third terms in Eq. 11 hold because w_{t-1} is independent of x_{t-1} given Y_{t-1} .
3. Eq. 12 holds because $E\{\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1} | Y_{t-1}\} = 0$.
4. the last equality in Eq. 14 follows from Eq. ?? in Lemma 1
5. the last equality in Eq. 15 holds because \mathbf{x}_t is independent (and therefore uncorrelated) of \mathbf{v}_t .
6. the second equality of Eq. 16 uses Eq. 13.
7. the third equality of Eq. 16 holds because \mathbf{x}_t is independent of \mathbf{v}_t given Y_{t-1} .
8. the last equality of Eq. 16 holds because \mathbf{v}_t is independent of Y_{t-1} .

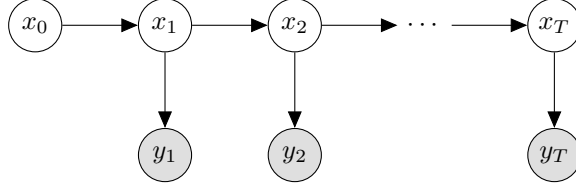


Figure 1: Graphical models for our linear dynamical system in Theorem 1.

9. in the third equality of Eq. 17 we used Eq. 20 in Lemma 1 with $\mathbf{x} = \mathbf{x}_t|Y_{t-1}$ and $\mathbf{y} = \mathbf{z}_t|Y_{t-1}$ giving

$$\begin{aligned}
\Sigma_{x|y} &= \Sigma_{\mathbf{x}_t|\mathbf{z}_t, Y_{t-1}} = \Sigma_{\mathbf{x}_t|Y_t} = P_{t|t} \\
\Sigma_{xx} &= \Sigma_{\mathbf{x}_t|Y_{t-1}} = P_{t|t-1} \\
\Sigma_{xy} &= \Sigma_{\mathbf{x}_t\mathbf{z}_t|Y_{t-1}} = \text{Cov}(\mathbf{x}_t, \mathbf{z}_t|Y_{t-1}) = P_{t|t-1}B_t^\top \\
\Sigma_{yy} &= \Sigma_{\mathbf{z}_t\mathbf{z}_t|Y_{t-1}} = \text{Cov}(\mathbf{z}_t|Y_{t-1}) = S_t \\
&\text{thus} \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}B_t^\top S_t^{-1}B_t P_{t|t-1}
\end{aligned}$$

10. in the fourth equality of Eq. 18 we used Lemma ??.

Lemma 1. Let \mathbf{x} and \mathbf{y} be jointly Gaussian distributed random vectors with

$$E\left\{\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}\right\} = \begin{pmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{pmatrix} \quad (19)$$

$$\text{Cov}\left\{\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}\right\} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \quad (20)$$

where Σ_{yy} is assumed to be non-singular. Then the conditional distribution of \mathbf{x} given \mathbf{y} is Gaussian with mean vector

$$E\{\mathbf{x}|\mathbf{y}\} = \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)$$

and covariance matrix

$$\text{Cov}\{\mathbf{x}|\mathbf{y}\} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$$

Proof. Let

$$\mathbf{z} = \mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \quad (21)$$

Since (\mathbf{x}, \mathbf{y}) are jointly Gaussian, and (\mathbf{z}, \mathbf{y}) is an affine transformation of (\mathbf{x}, \mathbf{y}) , then (\mathbf{z}, \mathbf{y}) are jointly Gaussian.

We have

$$\begin{aligned}
E\{\mathbf{z}\} &= E\{\mathbf{x}\} = \boldsymbol{\mu}_x \\
\mathbf{z} - \boldsymbol{\mu}_z &= (\mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)) - \boldsymbol{\mu}_x = (\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
\text{Cov}\{\mathbf{z}\} &= E\{(\mathbf{z} - \boldsymbol{\mu}_z)(\mathbf{z} - \boldsymbol{\mu}_z)^\top\} \\
&= E\{[(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)] [(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)]^\top\} \\
&= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} + \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} \\
&= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \\
\text{Cov}\{\mathbf{y}, \mathbf{z}\} &= E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{z} - \boldsymbol{\mu}_z)^\top\} \\
&= E\{(\mathbf{y} - \boldsymbol{\mu}_y)((\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y))^\top\} \\
&= \Sigma_{yx} - \Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} = 0
\end{aligned} \tag{22}$$

Because (\mathbf{y}, \mathbf{z}) are uncorrelated (Eq. 22) and jointly Gaussian, they are independent. Thus, $E\{\mathbf{z}|\mathbf{y}\} = E\{\mathbf{z}\}$ and $\text{Cov}\{\mathbf{z}|\mathbf{y}\} = \text{Cov}\{\mathbf{z}\}$.

From Eq. 21, $\mathbf{x} = \mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)$. Then

$$\begin{aligned}
E\{\mathbf{x}|\mathbf{y}\} &= E\{\mathbf{z}|\mathbf{y}\} + E\{\Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)|\mathbf{y}\} \\
&= E\{\mathbf{z}\} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
&= \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
\text{Cov}\{\mathbf{x}|\mathbf{y}\} &= \text{Cov}\{\mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)|\mathbf{y}\} = \text{Cov}\{\mathbf{z}|\mathbf{y}\} \\
&= \text{Cov}\{\mathbf{z}\} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}
\end{aligned} \tag{23}$$

Notes:

1. The last equality in Eq. 23 holds because, when conditioning on \mathbf{y} , the term $\Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)$ is a constant, and constants are irrelevant when computing covariances.

□

References

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*, volume 38. OUP Oxford.