Derivation of the Kalman filter equations

Joaquín Rapela

May 22, 2025

Theorem 1. *Given the linear dynamical systems model*

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

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

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

with

$$
\begin{aligned}
\mathbf{x}_{t|t-1} &= A_{t-1}\mathbf{x}_{t-1|t-1} \\
P_{t|t-1} &= A_{t-1}P_{t-1|t-1}A_{t-1}^{\intercal} + Q_{t-1} \\
\hat{\mathbf{y}}_{t|t-1} &= B_t\mathbf{x}_{t|t-1} \\
\mathbf{z}_t &= \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} \\
S_t &= Cov(\{\mathbf{z}_t\}|\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}) = B_tP_{t|t-1}B_t^{\intercal} + R_t \\
\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + K_t\mathbf{z}_t \\
\mathbf{P}_{t|t} &= (I - K_tB_t)P_{t|t-1} \\
\mathbf{K}_t &= P_{t|t-1}B_t^{\intercal}S_t^{-1} \\
\mathbf{x}_{0|0} &= \mathbf{m}_0
\end{aligned}
$$

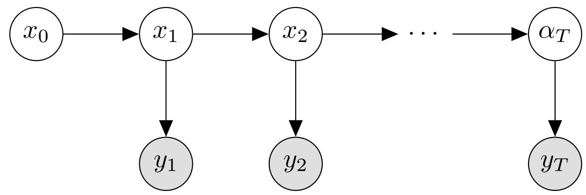

Figure 1: Graphical models for our linear dynamical system in Theorem 1.

$$P_{0|0} = V_0$$

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

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

$$\begin{aligned}\mathbf{x}_{t|t-1} &= E\{\mathbf{x}_t|Y_{t-1}\} = E\{A_{t-1}\mathbf{x}_{t-1} + \mathbf{w}_{t-1}|Y_{t-1}\} \\ &= A_{t-1}E\{\mathbf{x}_{t-1}|Y_{t-1}\} + E\{\mathbf{w}_{t-1}|Y_{t-1}\} \\ &= A_{t-1}\mathbf{x}_{t-1|t-1} + E\{\mathbf{w}_{t-1}\} = A_{t-1}\mathbf{x}_{t-1|t-1} \qquad (1)\end{aligned}$$

□

Notes:

1. the first equality in Eq. 1 holds because $\mathbf{w}_{t-1}$ is independent of Y_{t-1}.

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

$$E\left\{\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}\right\} = \begin{pmatrix}\boldsymbol{\mu}_x\\ \boldsymbol{\mu}_y\end{pmatrix}$$
$$Cov\left\{\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}\right\} = \begin{pmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{pmatrix}$$

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

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

Proof. Let

$$\mathbf{z} = \mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \tag{2}$$

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

$$\begin{aligned}
\mathrm{E}\{\mathbf{z}\} &= E\{\mathbf{x}\} = \boldsymbol{\mu}_x \\
\mathbf{z} - \boldsymbol{\mu}_z &= \big(\mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)\big) - \boldsymbol{\mu}_x = (\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
\mathrm{Cov}\{\mathbf{z}\} &= \mathrm{E}\{(\mathbf{z} - \boldsymbol{\mu}_z)(\mathbf{z} - \boldsymbol{\mu}_z)^\intercal\} \\
&= \mathrm{E}\{\big[(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)\big]\,\big[(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)\big]^\intercal\} \\
&= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} + \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} \\
&= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \\
\mathrm{Cov}\{\mathbf{y}, \mathbf{z}\} &= \mathrm{E}\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{z} - \boldsymbol{\mu}_z)^\intercal\} \\
&= \mathbf{E}\{(\mathbf{y} - \boldsymbol{\mu}_y)((\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y))^\intercal\} \\
&= \Sigma_{yx} - \Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} = 0
\end{aligned} \tag{3}$$

Because $(\mathbf{y}, \mathbf{z})$ are uncorrelated (Eq. 3) and jointly Gaussian, they are independent. Thus, $\mathrm{E}\{\mathbf{z}|\mathbf{y}\} = \mathrm{E}\{\mathbf{z}\}$ and $\mathrm{Cov}\{\mathbf{z}|\mathbf{y}\} = \mathrm{Cov}\{\mathbf{z}\}$.

From Eq. 2, $\mathbf{x} = \mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)$. Then

$$\begin{aligned}
\mathrm{E}\{\mathbf{x}|\mathbf{y}\} &= \mathrm{E}\{\mathbf{z}|\mathbf{y}\} + \mathrm{E}\{\Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)|\mathbf{y}\} \\
&= \mathrm{E}\{\mathbf{z}\} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
&= \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
\mathrm{Cov}\{\mathbf{x}|\mathbf{y}\} &= \mathrm{Cov}\{\mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)|\mathbf{y}\} = \mathrm{Cov}\{\mathbf{z}|\mathbf{y}\} \\
&= \mathrm{Cov}\{\mathbf{z}\} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}
\end{aligned} \tag{4}$$

Notes:

1. The last equality in Eq. 4 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.

$\square$

References

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