

# Derivation of the Kalman filter equations

Joaquín Rapela

November 13, 2025

**Theorem 1.** *Given the linear dynamical systems model*

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

(represented graphically 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

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

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

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

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

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

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

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

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

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

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

The following proof adds 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

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

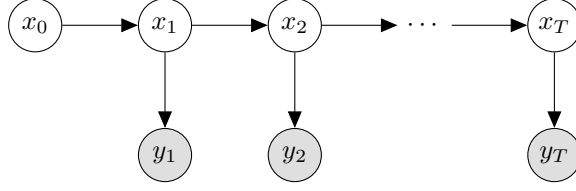


Figure 1: Graphical models for our linear dynamical system in Theorem 1.

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

This proves Eq. 1.

$$\begin{aligned} \mathbf{P}_{t|t-1} &= \text{Cov}\{\mathbf{x}_t|Y_{t-1}\} = E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top | Y_{t-1}\} \\ &= E\{(\mathbf{u}_t + A_{t-1}\mathbf{x}_{t-1} + w_{t-1} - \mathbf{u}_t - A_{t-1}\mathbf{x}_{t-1|t-1}) (\mathbf{u}_t + A_{t-1}\mathbf{x}_{t-1} + w_{t-1} - \mathbf{u}_t - A_{t-1}\mathbf{x}_{t-1|t-1})^\top | Y_{t-1}\} \\ &= E\{(A_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1}) + w_{t-1}) (A_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1}) + w_{t-1})^\top | Y_{t-1}\} \\ &= A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top | Y_{t-1}\}A_{t-1}^\top + \\ &\quad E\{w_{t-1}(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top | Y_{t-1}\}A_{t-1}^\top + \\ &\quad A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})w_{t-1}^\top | Y_{t-1}\} + \\ &\quad E\{w_{t-1}w_{t-1}^\top | Y_{t-1}\} \\ &= A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top | Y_{t-1}\}A_{t-1}^\top + \\ &\quad E\{w_{t-1}|Y_{t-1}\}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})^\top | Y_{t-1}\}A_{t-1}^\top{}^2 + \\ &\quad A_{t-1}E\{(\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1})|Y_{t-1}\}E\{w_{t-1}^\top | Y_{t-1}\} + \\ &\quad E\{w_{t-1}w_{t-1}^\top\} \\ &= A_{t-1}P_{t-1|t-1}A_{t-1}^\top + Q_{t-1}{}^3 \end{aligned}$$

This proves Eq. 2.

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

This proves Eq. 3.

Because

$$\mathbf{z}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} = \mathbf{a}_t + B_t\mathbf{x}_t + \mathbf{v}_t - \mathbf{a}_t - B_t\mathbf{x}_{t|t-1} = B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t \quad (10)$$

$Y_{t-1}$  and  $\mathbf{z}_t$  are fixed if and only if  $Y_t$  is fixed<sup>4</sup>. Then

<sup>1</sup> $\mathbf{w}_{t-1}$  is independent of  $Y_{t-1}$ .

<sup>2</sup> $\mathbf{w}_{t-1}$  is independent of  $x_{t-1}$  given  $Y_{t-1}$ .

<sup>3</sup> $E\{\mathbf{x}_{t-1} - \mathbf{x}_{t-1|t-1}|Y_{t-1}\} = E\{\mathbf{x}_{t-1}|Y_{t-1}\} - \mathbf{x}_{t-1|t-1} = \mathbf{x}_{t-1|t-1} - \mathbf{x}_{t-1|t-1} = 0$ .

<sup>4</sup>If we know  $Y_{t-1}$  and  $\mathbf{z}_t$ , then we know  $\hat{\mathbf{y}}_{t|t-1}$  and  $\mathbf{y}_t$ , then (by the first equality in Eq. 10) 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. 10) we know  $\mathbf{z}_t$ .

$$\begin{aligned}\mathbf{x}_{t|t} &= E\{\mathbf{x}_t|Y_t\} = E\{\mathbf{x}_t|Y_{t-1}, \mathbf{z}_t\}^5 \\ &= E\{\mathbf{x}_t|Y_{t-1}\} + \text{Cov}(\mathbf{x}_t, \mathbf{z}_t|Y_{t-1}) \text{Cov}(\mathbf{z}_t|Y_{t-1})^{-1} \mathbf{z}_t^6\end{aligned}\quad (11)$$

$$\begin{aligned}\text{Cov}(\mathbf{x}_t, \mathbf{z}_t|Y_{t-1}) &= \text{Cov}(\mathbf{x}_t, B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t|Y_{t-1}) \\ &= E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t)^\top|Y_{t-1}\} \\ &= E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top|Y_{t-1}\}B_t^\top \\ &\quad + E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})\mathbf{v}_t^\top|Y_{t-1}\} \\ &= P_{t|t-1}B_t^\top\end{aligned}\quad (12)$$

$$\begin{aligned}S_t = \text{Cov}(\mathbf{z}_t|Y_{t-1}) &= E\{\mathbf{z}_t\mathbf{z}_t^\top|Y_{t-1}\} \\ &= E\{(B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t)(B_t(\mathbf{x}_t - \mathbf{x}_{t|t-1}) + \mathbf{v}_t)^\top|Y_{t-1}\}^8 \\ &= B_tE\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top|Y_{t-1}\}B_t^\top + B_tE\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})\mathbf{v}_t^\top|Y_{t-1}\} \\ &\quad + E\{\mathbf{v}_t(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top\}B_t^\top + E\{\mathbf{v}_t\mathbf{v}_t^\top|Y_{t-1}\} \\ &= B_tE\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})^\top|Y_{t-1}\}B_t^\top + \\ &\quad + E\{\mathbf{v}_t\mathbf{v}_t^\top\}^{79} \\ &= B_tP_{t|t-1}B_t^\top + R_t\end{aligned}\quad (13)$$

This proves Eq. 4.

Combining Eqs. 11, 12 and 13 we obtain

$$\begin{aligned}\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + P_{t|t-1}B_t^\top S_t^{-1}\mathbf{z}_t \\ &= \mathbf{x}_{t|t-1} + K_t\mathbf{z}_t \quad \text{with } K_t = P_{t|t-1}B_t^\top S_t^{-1} \\ P_{t|t} &= \text{Cov}(\mathbf{x}_t|Y_t) = \text{Cov}(\mathbf{x}_t|Y_{t-1}, \mathbf{z}_t) = P_{t|t-1} - P_{t|t-1}B_t^\top S_t^{-1}B_tP_{t|t-1}^{10} \\ &= (I - P_{t|t-1}B_t^\top S_t^{-1}B_t)P_{t|t-1} = (I - K_tB_t)P_{t|t-1}\end{aligned}$$

This proves Eqs. 5, 6 and 7.

---

<sup>5</sup>The validity of the last equality follow from measure theory arguments (that I don't know).

<sup>6</sup>Refer to Eq. 14 in Lemma 1.

<sup>7</sup> $E\{(\mathbf{x}_t - \mathbf{x}_{t|t-1})\mathbf{v}_t^\top|Y_{t-1}\} = E\{\mathbf{x}_t - \mathbf{x}_{t|t-1}|Y_{t-1}\}E\{\mathbf{v}_t^\top|Y_{t-1}\} = (E\{\mathbf{x}_t|Y_{t-1}\} - \mathbf{x}_{t|t-1})E\{\mathbf{v}_t^\top|Y_{t-1}\} = (\mathbf{x}_{t|t-1} - \mathbf{x}_{t|t-1})E\{\mathbf{v}_t^\top|Y_{t-1}\} = 0$  because  $\mathbf{x}_t$  is independent of  $\mathbf{v}_t$  given  $Y_{t-1}$ .

<sup>8</sup>Eq. 10.

<sup>9</sup> $\mathbf{v}_t$  is independent from  $Y_{t-1}$ .

<sup>10</sup>Refer to Eq. 15 in Lemma 1 with  $\mathbf{x} = \mathbf{x}_t|Y_{t-1}$  and  $\mathbf{y} = \mathbf{z}_t|Y_{t-1}$  giving

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

$$\begin{aligned}
p(\mathbf{x}_1) &= \mathcal{N}(\mathbf{x}_1|\mathbf{x}_{1|0}, \mathbf{P}_{1|0}) = \int p(\mathbf{x}_1, \mathbf{x}_0) d\mathbf{x}_0 = \int p(\mathbf{x}_1|\mathbf{x}_0) p(\mathbf{x}_0) d\mathbf{x}_0 \\
&= \int \mathcal{N}(\mathbf{x}_1|\mathbf{u}_0 + A_0 \mathbf{x}_0, Q_0) \mathcal{N}(\mathbf{x}_0|\mathbf{m}_0, V_0) d\mathbf{x}_0 = \mathcal{N}(\mathbf{x}_1|\mathbf{u}_0 + A_0 \mathbf{m}_0, A_0 V_0 A_0^\top + Q_0)^{11}
\end{aligned}$$

then

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

and by Eqs. 1 and 2

$$\begin{aligned}
\mathbf{x}_{0|0} &= \mathbf{m}_0 \\
\mathbf{P}_{1|0} &= V_0
\end{aligned}$$

This proves Eqs. 8 and 9.

□

---

<sup>11</sup>Lemma 2.

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

$$E \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\} = \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}$$

$$Cov \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\} = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$$

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

$$E\{\mathbf{x}|\mathbf{y}\} = \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \quad (14)$$

and covariance matrix

$$Cov\{\mathbf{x}|\mathbf{y}\} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \quad (15)$$

*Proof.* Let

$$\mathbf{z} = \mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \quad (16)$$

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

$$\begin{aligned} E\{\mathbf{z}\} &= E\{\mathbf{x}\} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) = \boldsymbol{\mu}_x \\ \mathbf{z} - \boldsymbol{\mu}_z &= (\mathbf{x} - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)) - \boldsymbol{\mu}_x = (\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\ Cov\{\mathbf{z}\} &= E\{(\mathbf{z} - \boldsymbol{\mu}_z)(\mathbf{z} - \boldsymbol{\mu}_z)^\top\} \\ &= E\{[(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)] [(\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)]^\top\} \\ &= E\{(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{x} - \boldsymbol{\mu}_x)^\top\} - E\{(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{y} - \boldsymbol{\mu}_y)^\top\}\Sigma_{yy}^{-1}\Sigma_{xy} - \Sigma_{xy}\Sigma_{yy}^{-1}E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{x} - \boldsymbol{\mu}_x)^\top\} + \\ &\quad \Sigma_{xy}\Sigma_{yy}^{-1}E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^\top\}\Sigma_{yy}^{-1}\Sigma_{xy} \\ &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} + \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} \\ &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \\ Cov\{\mathbf{y}, \mathbf{z}\} &= E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{z} - \boldsymbol{\mu}_z)^\top\} \\ &= E\{(\mathbf{y} - \boldsymbol{\mu}_y)((\mathbf{x} - \boldsymbol{\mu}_x) - \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y))^\top\} \\ &= \Sigma_{yx} - \Sigma_{yy}\Sigma_{yy}^{-1}\Sigma_{yx} = 0 \end{aligned} \quad (17)$$

Because  $(\mathbf{y}, \mathbf{z})$  are uncorrelated (Eq. 17) and jointly Gaussian, they are independent. Thus,  $E\{\mathbf{z}|\mathbf{y}\} = E\{\mathbf{z}\}$  and  $Cov\{\mathbf{z}|\mathbf{y}\} = Cov\{\mathbf{z}\}$ .

From Eq. 16,  $\mathbf{x} = \mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)$ . Then

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

$$\begin{aligned}
&= \mathbb{E}\{\mathbf{z}\} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
&= \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\
\text{Cov}\{\mathbf{x}|\mathbf{y}\} &= \text{Cov}\{\mathbf{z} + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y)|\mathbf{y}\} = \text{Cov}\{\mathbf{z}|\mathbf{y}\}^{12} \\
&= \text{Cov}\{\mathbf{z}\} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}
\end{aligned} \tag{18}$$

□

---

<sup>12</sup>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.

**Lemma 2.** *Let*

$$\begin{aligned} p(\mathbf{y}|\mathbf{x}) &= \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x} + \mathbf{b}, \Sigma) \\ p(\mathbf{x}) &= \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Lambda) \end{aligned}$$

*then*

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\Lambda\mathbf{A}^\top + \Sigma)$$

*Proof.*

$$\begin{aligned} \ln p(\mathbf{x}, \mathbf{y}) &= \ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x}) \\ &= -\frac{1}{2} (\mathbf{y} - (\mathbf{A}\mathbf{x} + \mathbf{b}))^\top \Sigma^{-1} (\mathbf{y} - (\mathbf{A}\mathbf{x} + \mathbf{b})) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \Lambda^{-1} (\mathbf{x} - \boldsymbol{\mu}) + K_1 \\ &= -\frac{1}{2} \mathbf{y}^\top \Sigma^{-1} \mathbf{y} + \frac{1}{2} \mathbf{y}^\top \Sigma^{-1} \mathbf{A} \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{A}^\top \Sigma^{-1} \mathbf{y} - \frac{1}{2} \mathbf{x}^\top (\mathbf{A}^\top \Sigma^{-1} \mathbf{A} + \Lambda^{-1}) \mathbf{x} \\ &\quad + \frac{1}{2} \mathbf{y}^\top \Sigma^{-1} \mathbf{b} + \frac{1}{2} \mathbf{x}^\top (-\mathbf{A}^\top \Sigma^{-1} \mathbf{b} + \Lambda \boldsymbol{\mu}) + \frac{1}{2} \mathbf{b}^\top \Sigma^{-1} \mathbf{y} + \frac{1}{2} (-\mathbf{b}^\top \Sigma^{-1} \mathbf{A} + \boldsymbol{\mu}^\top \Lambda) \mathbf{x} + K_2 \\ &= -\frac{1}{2} [\mathbf{x}^\top, \mathbf{y}^\top] \begin{bmatrix} \mathbf{A}^\top \Sigma^{-1} \mathbf{A} + \Lambda^{-1} & -\mathbf{A}^\top \Sigma^{-1} \\ -\Sigma^{-1} \mathbf{A} & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \\ &\quad + \frac{1}{2} [\mathbf{x}^\top, \mathbf{y}^\top] \begin{bmatrix} -\mathbf{A}^\top \Sigma^{-1} \mathbf{b} + \Lambda^{-1} \boldsymbol{\mu} \\ \Sigma^{-1} \mathbf{b} \end{bmatrix} + \frac{1}{2} [-\mathbf{b}^\top \Sigma^{-1} \mathbf{A} + \boldsymbol{\mu}^\top \Lambda^{-1}, \mathbf{b}^\top \Sigma^{-1}] \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + K_2 \quad (19) \end{aligned}$$

where  $K_1$  and  $K_2$  are constants that do not depend on  $\mathbf{x}$  or  $\mathbf{y}$ .

Because  $\ln p(\mathbf{x}, \mathbf{y})$  is a quadratic form, then  $p(\mathbf{x}, \mathbf{y})$  is a normal probability density function (pdf), thus its marginal  $p(\mathbf{y})$  is also a normal pdf. Our aim is to derive the mean and covariance of  $\mathbf{y}$ ,  $\boldsymbol{\mu}_y$  and  $\Gamma_{yy}$ , respectively.

Call

$$p(\mathbf{x}, \mathbf{y}) = \mathcal{N} \left( \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \middle| \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}, \Gamma \right)$$

with

$$\Phi^{-1} = \Gamma = \begin{bmatrix} \Gamma_{xx} & \Gamma_{xy} \\ \Gamma_{yx} & \Gamma_{yy} \end{bmatrix}$$

Next,

$$\begin{aligned} \ln p(\mathbf{x}, \mathbf{y}) &= -\frac{1}{2} [(\mathbf{x} - \boldsymbol{\mu}_x)^\top, (\mathbf{y} - \boldsymbol{\mu}_y)^\top] \Phi [(\mathbf{x} - \boldsymbol{\mu}_x), (\mathbf{y} - \boldsymbol{\mu}_y)] + K_1 \\ &= -\frac{1}{2} [\mathbf{x}^\top, \mathbf{y}^\top] \Phi \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + \frac{1}{2} [\mathbf{x}^\top, \mathbf{y}^\top] \Phi \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix} + \frac{1}{2} [\boldsymbol{\mu}_x^\top, \boldsymbol{\mu}_y^\top] \Phi \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} + K_2 \quad (20) \end{aligned}$$

where  $K_1$  and  $K_2$  are constants that do not depend on  $\mathbf{x}$  or  $\mathbf{y}$ .

From Eqs. 19 and 20 it follows that

$$\Phi = \begin{bmatrix} A^\top \Sigma^{-1} A + \Lambda^{-1} & -A^\top \Sigma^{-1} \\ -\Sigma^{-1} A & \Sigma^{-1} \end{bmatrix}$$

and

$$\Phi \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix} = \begin{bmatrix} -A^\top \Sigma^{-1} \mathbf{b} + \Lambda^{-1} \boldsymbol{\mu} \\ \Sigma^{-1} \mathbf{b} \end{bmatrix}$$

Then

$$\begin{aligned} \Gamma &= \begin{bmatrix} \Gamma_{xx} & \Gamma_{xy} \\ \Gamma_{yx} & \Gamma_{yy} \end{bmatrix} = \Phi^{-1} = \begin{bmatrix} A^\top \Sigma^{-1} A + \Lambda^{-1} & -A^\top \Sigma^{-1} \\ -\Sigma^{-1} A & \Sigma^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \Lambda & \Lambda A \\ A \Lambda & \Sigma + A \Lambda A^\top \end{bmatrix}^{13} \quad (21) \\ \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix} &= \Phi^{-1} \begin{bmatrix} -A^\top \Sigma^{-1} \mathbf{b} + \Lambda^{-1} \boldsymbol{\mu} \\ \Sigma^{-1} \mathbf{b} \end{bmatrix} = \Gamma \begin{bmatrix} -A^\top \Sigma^{-1} \mathbf{b} + \Lambda^{-1} \boldsymbol{\mu} \\ \Sigma^{-1} \mathbf{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \\ A \boldsymbol{\mu} + \mathbf{b} \end{bmatrix} \end{aligned}$$

Thus,

$$\begin{aligned} \Gamma_{yy} &= \Sigma + A \Lambda A^\top \\ \boldsymbol{\mu}_y &= A \boldsymbol{\mu} + \mathbf{b} \end{aligned}$$

□

---

<sup>13</sup>Lemma 3.



**Lemma 3.**

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{bmatrix}$$

with

$$M = (A - BD^{-1}C)^{-1}$$

*Proof.*

$$\begin{aligned} & \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{bmatrix} = \\ & \begin{bmatrix} AM - BD^{-1}CM & -AMB D^{-1} + BD^{-1} + BD^{-1}CMBD^{-1} \\ CM - CM & -CMBD^{-1} + I + CMBD^{-1} \end{bmatrix} = \\ & \begin{bmatrix} (A - BD^{-1}C)M & (-A + M^{-1} + BD^{-1}C)MBD^{-1} \\ 0 & I \end{bmatrix} = \\ & \begin{bmatrix} M^{-1}M & (-M^{-1} + M^{-1})MBD^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \\ & \begin{bmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \\ & \begin{bmatrix} MA - MBD^{-1}C & MB - MB \\ -D^{-1}CMA + D^{-1}C + D^{-1}CMBD^{-1}C & -D^{-1}CMB + I - D^{-1}CMB \end{bmatrix} = \\ & \begin{bmatrix} M(A - BD^{-1}C) & 0 \\ -D^{-1}CM(A - BD^{-1}C) + D^{-1}C & I \end{bmatrix} = \\ & \begin{bmatrix} MM^{-1} & 0 \\ -D^{-1}CMM^{-1} + D^{-1}C & I \end{bmatrix} = \\ & \begin{bmatrix} I & 0 \\ -D^{-1}C + D^{-1}C & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned}$$

□

## References

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