References

Alfeld, G., & Herzberger, J. (1983). *Introduction to interval computations*. New York: Academic Press.

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. Englewood Cliffs: Prentice-Hall.

Andrews, A. (1981). Parallel processing of the Kalman filter. In *IEEE Proceedings of the International Conference on Parallel Processing* (pp. 216–220).

Aoki, M. (1989). *Optimization of stochastic systems: Topics in discrete-time dynamics*. New York: Academic Press.

Aström, K. J., & Eykhoff, P. (1971). System identification – a survey. *Automatica*, *7*, 123–162.

Balakrishnan, A. V. (1984, 1987). *Kalman filtering theory*. New York: Optimization Software, Inc.

Bierman, G. J. (1973). A comparison of discrete linear filtering algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, *9*, 28–37.

Bierman, G. J. (1977). *Factorization methods for discrete sequential estimation*. New York: Academic Press.

Blahut, R. E. (1985). *Fast algorithms for digital signal processing*. Reading: Addison-Wesley.

Bozic, S. M. (1979). *Digital and Kalman filtering*. New York: Wiley.

Brammer, K., & Sifflin, G. (1989). *Kalman-Bucy filters*. Boston: Artech House Inc.

Brown, R. G., & Hwang, P. Y. C. (1992, 1997). *Introduction to random signals and applied Kalman filtering*. New York: Wiley.

Bucy, R. S., & Joseph, P. D. (1968). *Filtering for stochastic processes with applications to guidance*. New York: Wiley.

Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). *Introduction to wavelets and wavelet transfroms: A primer*. Upper Saddle River: Prentice-Hall.

Carlson, N. A. (1973). Fast triangular formulation of the square root filter. *Journal of AIAA*, *11*, 1259–1263.

Catlin, D. E. (1989). *Estimation, control, and the discrete Kalman filter*. New York: Springer.

Cattivelli, F. S., & Sayed, A. H. (2010). Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Transactions on Automatic Control*, *55*(9), 2069–2084.

Chen, G. (1992). Convergence analysis for inexact mechanization of Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, *28*, 612–621.

Chen, G. (1993). *Approximate Kalman filtering*. Singapore: World Scientific.

C.K. Chui and G. Chen, *Kalman Filtering*, DOI 10.1007/978-3-319-47612-4

Chen, G., & Chui, C. K. (1986). Design of near-optimal linear digital tracking filters with colored input. *Journal of Computational and Applied Mathematics*, *15*, 353–370.

Chen, G., Chen, G., & Hsu, S. H. (1995). *Linear stochastic control systems*. Boca Raton: CRC Press.

Chen, G., Wang, J., & Shieh, L. S. (1997). Interval Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, *33*, 250–259.

Chen, H. F. (1985). *Recursive estimation and control for stochastic systems*. New York: Wiley.

Chui, C. K. (1984). Design and analysis of linear prediction-correction digital filters. *Linear and Multilinear Algebra*, *15*, 47–69.

Chui, C. K. (1997). *Wavelets: A mathematical tool for signal analysis*. Philadelphia: SIAM.

Chui, C. K., & Chen, G. (1989). *Linear systems and optimal control*. New York: Springer.

Chui, C. K., & Chen, G. (1992, 1997). *Signal processing and systems theory: Selected topics*. New York: Springer.

Chui, C. K., Chen, G., & Chui, H. C. (1990). Modified extended Kalman filtering and a real-time parallel algorithm for system parameter identification. *IEEE Transactions on Automatic Control*, *35*, 100–104.

Davis, M. H. A. (1977). *Linear estimation and stochastic control*. New York: Wiley.

Davis, M. H. A., & Vinter, R. B. (1985). *Stochastic modeling and control*. New York: Chapman and Hall.

Fleming, W. H., & Rishel, R. W. (1975). *Deterministic and stochastic optimal control*. New York: Springer.

Gaston, F. M. F., & Irwin, G. W. (1990). Systolic Kalman filtering: An overview. *IEE Proceedings-D*, *137*, 235–244.

Goodwin, G. C., & Sin, K. S. (1984). *Adaptive filtering prediction and control*. Englewood Cliffs: Prentice-Hall.

Haykin, S. (1986). *Adaptive filter theory*. Englewood Cliffs: Prentice-Hall.

Hong, L., Chen, G., & Chui, C. K. (1998). A filter-bank-based Kalman filtering technique for wavelet estimation and decomposition of random signals. *IEEE Transactions on Circuits and Systems (II)* (in press).

Hong, L., Chen, G., & Chui, C. K. (1998). Real-time simultaneous estimation and decomposition of random signals. *Multidimensional Systems and Signal Processing* (in press).

Jazwinski, A. H. (1969). Adaptive filtering. *Automatica*, *5*, 475–485.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. New York: Academic Press.

Jover, J. M., & Kailath, T. (1986). A parallel architecture for Kalman filter measurement update and parameter estimation. *Automatica*, *22*, 43–57.

Kailath, T. (1968). An innovations approach to least-squares estimation, part I: Linear filtering in additive white noise. *IEEE Transactions on Automatic Control*, *13*, 646–655.

Kailath, T. (1982). *Course notes on linear estimation*. Stanford: Stanford University.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, *82*, 35–45.

Kalman, R. E. (1963). New method in Wiener filtering theory. In *Proceedings of the Symposium on Engineering Applications of Random Function Theory and Probability*. New York: Wiley.

Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Transactions of the ASME – Journal of Basic Engineering*, *83*, 95–108.

Kumar, P. R., & Varaiya, P. (1986). *Stochastic systems: Estimation, identification, and adaptive control*. Englewood Cliffs: Prentice-Hall.

Kung, H. T. (1982). Why systolic architectures? *Computer*, *15*, 37–46.

Kung, S. Y. (1985). VLSI arrays processors. *IEEE ASSP Magazine*, *2*, 4–22.

Kushner, H. (1971). *Introduction to stochastic control*. New York: Holt Rinehart and Winston Inc.

Lewis, F. L. (1986). *Optimal estimation*. New York: Wiley.

Lu, M., Qiao, X., & Chen, G. (1992). A parallel square-root algorithm for the modified extended Kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, *28*, 153–163.
Lu, M., Qiao, X., & Chen, G. (1993). Parallel computation of the modified extended Kalman filter. *International Journal of Computer Mathematics*, *45*, 69–87.
Maybeck, P. S. (1982). *Stochastic models, estimation, and control* (Vol. 1,2,3). New York: Academic.
Mead, C., & Conway, L. (1980). *Introduction to VLSI systems*. Reading: Addison-Wesley.
Mehra, R. K. (1970). On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, *15*, 175–184.
Mehra, R. K. (1972). Approaches to adaptive filtering. *IEEE Transactions on Automatic Control*, *17*, 693–698.
Mendel, J. M. (1987). *Lessons in digital estimation theory*. Englewood Cliffs: Prentice-Hall.
Potter, J. E. (1963). New statistical formulas. Instrumentation Lab., MIT (Space Guidance Analysis Memo).
Probability Group (1975). Institute of Mathematics, Academia Sinica, China, ed.: Mathematical Methods of Filtering for Discrete-Time Systems (in Chinese) (Beijing)
Ruymgaart, P. A., & Soong, T. T. (1985, 1988). *Mathematics of Kalman-Bucy filtering*. New York: Springer.
Shiryayev, A. N. (1984). *Probability*. New York: Springer.
Siouris, G., Chen, G., & Wang, J. (1997). Tracking an incoming ballistic missile. *IEEE Transactions on Aerospace and Electronic Systems*, *33*, 232–240.
Sorenson, H. W. (Ed.). (1985). *Kalman filtering: Theory and application*. New York: IEEE Press.
Stengel, R. F. (1986). *Stochastic optimal control: Theory and application*. New York: Wiley.
Strobach, P. (1990). *Linear prediction theory: A mathematical basis for adaptive systems*. New York: Springer.
Wang, E. P. (1972). Optimal linear recursive filtering methods. *Journal of Mathematics in Practice and Theory (in Chinese)*, *6*, 40–50.
Wonham, W. M. (1968). On the separation theorem of stochastic control. *SIAM Journal on Control*, *6*, 312–326.
Xu, J. H., Bian, G. R., Ni, C. K., & Tang, G. X. (1981). *State estimation and system identification* (in Chinese) (Beijing).
Yang, W., Chen, G., Wang, X. F., & Shi, L. (2014). Stochastic sensor activation for distributed state estimation over a sensor network. *Automatics*, *50*, 2070–2076.
Young, P. (1984). *Recursive estimation and time-series analysis*. New York: Springer.
Yu, W. W., Chen, G., Wang, Z. D., & Yang, W. (2009). Distributed consensus filtering in sensor networks. *IEEE Transactions on Systems, Man and Cybernetics–Part B: Cybernetics*, *39*(6), 1568–1577.
Zhang, H. S., Song, X. M., & Shi, L. (2012). Convergence and mean square stability of suboptimal estimator for systems with measurements packet dropping. *IEEE Transactions on Automatic Control*, *57*(5), 1248–1253.

Answers and Hints to Exercises

Chapter 1

1.1. Since most of the properties can be verified directly by using the definition of the trace, we only consider $\mathrm{tr}AB = \mathrm{tr}BA$. Indeed,

$$\mathrm{tr}AB = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij}b_{ji}\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} b_{ji}a_{ij}\right) = \mathrm{tr}BA\,.$$

1.2.

$$(\mathrm{tr}A)^2 = \left(\sum_{i=1}^{n} a_{ii}\right)^2 \leq n\sum_{i=1}^{n} a_{ii}^2 \leq n\left(\mathrm{tr}AA^\top\right).$$

1.3.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

1.4. There exist unitary matrices P and Q such that

$$A = P\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^\top, \quad B = Q\begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{bmatrix} Q^\top,$$

and

$$\sum_{k=1}^{n} \lambda_k^2 \geq \sum_{k=1}^{n} \mu_k^2\,.$$

Let $P = [p_{ij}]_{n\times n}$ and $Q = [q_{ij}]_{n\times n}$. Then

$$p_{11}^2 + p_{21}^2 + \cdots + p_{n1}^2 = 1\,,\; p_{12}^2 + p_{22}^2 + \cdots + p_{n2}^2 = 1\,, \cdots,$$

C.K. Chui and G. Chen, *Kalman Filtering*, DOI 10.1007/978-3-319-47612-4

$$p_{1n}^2 + p_{2n}^2 + \cdots + p_{nn}^2 = 1\,, \quad q_{11}^2 + q_{21}^2 + \cdots + q_{n1}^2 = 1\,,$$

$$q_{12}^2 + q_{22}^2 + \cdots + q_{n2}^2 = 1\,, \cdots, q_{1n}^2 + q_{2n}^2 + \cdots + q_{nn}^2 = 1\,,$$

and

$$\begin{aligned}
\mathrm{tr} A A^\top &= \mathrm{tr}\left\{ P \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix} P^\top \right\} \\
&= \mathrm{tr} \begin{bmatrix} p_{11}^2\lambda_1^2 + p_{12}^2\lambda_2^2 + \cdots + p_{1n}^2\lambda_n^2 & & * \\ & p_{21}^2\lambda_1^2 + p_{22}^2\lambda_2^2 + \cdots + p_{2n}^2\lambda_n^2 & \\ * & & p_{n1}^2\lambda_1^2 + p_{n2}^2\lambda_2^2 + \cdots + p_{nn}^2\lambda_n^2 \end{bmatrix} \\
&= (p_{11}^2 + p_{21}^2 + \cdots + p_{n1}^2)\lambda_1^2 + \cdots + (p_{1n}^2 + p_{2n}^2 + \cdots + p_{nn}^2)\lambda_n^2 \\
&= \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2\,.
\end{aligned}$$

Similarly, $\mathrm{tr} B B^\top = \mu_1^2 + \mu_2^2 + \cdots + \mu_n^2$. Hence, $\mathrm{tr} A A^\top \geq \mathrm{tr} B B^\top$.

1.5. Denote

$$I = \int_{-\infty}^{\infty} e^{-y^2} dy\,.$$

Then, using polar coordinates, we have

$$\begin{aligned}
I^2 &= \left(\int_{-\infty}^{\infty} e^{-y^2} dy\right)\left(\int_{-\infty}^{\infty} e^{-x^2} dx\right) \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \\
&= \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r dr d\theta = \pi\,.
\end{aligned}$$

1.6. Denote

$$I(x) = \int_{-\infty}^{\infty} e^{-xy^2} dy\,.$$

Then, by Exercise 1.5,

$$I(x) = \frac{1}{\sqrt{x}}\int_{-\infty}^{\infty} e^{-(\sqrt{x}y)^2} d(\sqrt{x}y) = \sqrt{\pi/x}\,.$$

Hence,

$$\int_{-\infty}^{\infty} y^2 e^{-y^2} dy = -\frac{d}{dx} I(x)\Big|_{x=1}$$
$$= -\frac{1}{dx}\left(\sqrt{\pi/x}\right)\Big|_{x=1} = \frac{1}{2}\sqrt{\pi}\,.$$

1.7. (a) Let P be a unitary matrix so that

$$R = P^{\top} diag[\lambda_1, \cdots, \lambda_n] P\,,$$

and define

$$\mathbf{y} = \frac{1}{\sqrt{2}}\, diag[\sqrt{\lambda}_1, \cdots, \sqrt{\lambda}_n] P(\mathbf{x} - \underline{\mu})\,.$$

Then

$$E(X) = \int_{-\infty}^{\infty} \mathbf{x} f(\mathbf{x}) d\mathbf{x}$$
$$= \int_{-\infty}^{\infty} \left(\underline{\mu} + \sqrt{2} P^{-1}\, diag[\,1/\sqrt{\lambda}_1, \cdots, 1/\sqrt{\lambda}_n\,]\mathbf{y}\right) f(\mathbf{x}) d\mathbf{x}$$
$$= \underline{\mu} \int_{-\infty}^{\infty} f(\mathbf{x}) d\mathbf{x}$$
$$+\, Const. \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} e^{-y_1^2} \cdots e^{-y_n^2} dy_1 \cdots dy_n$$
$$= \underline{\mu} \cdot 1 + 0 = \underline{\mu}\,.$$

(b) Using the same substitution, we have

$$Var(X)$$
$$= \int_{-\infty}^{\infty} (\mathbf{x} - \underline{\mu})(\mathbf{x} - \underline{\mu})^{\top} f(\mathbf{x}) d\mathbf{x}$$
$$= \int_{-\infty}^{\infty} 2 R^{1/2} \mathbf{y}\mathbf{y}^{\top} R^{1/2} f(\mathbf{x}) d\mathbf{x}$$
$$= \frac{2}{(\pi)^{n/2}} R^{1/2} \Bigg\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} y_1^2 & \cdots & y_1 y_n \\ \vdots & & \vdots \\ y_n y_1 & \cdots & y_n^2 \end{bmatrix}$$
$$\cdot\ e^{-y_1^2} \cdots e^{-y_n^2} dy_1 \cdots dy_n \Bigg\} R^{1/2}$$
$$= R^{1/2} I R^{1/2} = R\,.$$

1.8. All the properties can be easily verified from the definitions.
1.9. We have already proved that if X_1 and X_2 are independent then $Cov(X_1, X_2) = 0$. Suppose now that $R_{12} = Cov(X_1, X_2) = 0$. Then $R_{21} = Cov(X_2, X_1) = 0$ so that

$$\begin{aligned} f(X_1, X_2) &= \frac{1}{(2\pi)^{n/2} det\, R_{11} det\, R_{22}} \\ &\quad \cdot\, e^{-\frac{1}{2}(X_1-\underline{\mu}_1)^\top R_{11}(X_1-\underline{\mu}_1)} e^{-\frac{1}{2}(X_2-\underline{\mu}_2)^\top R_{22}(X_2-\underline{\mu}_2)} \\ &= f_1(X_1)\cdot f_2(X_2)\,. \end{aligned}$$

Hence, X_1 and X_2 are independent.
1.10. Equation (1.35) can be verified by a direct computation. First, the following formula may be easily obtained:

$$\begin{aligned} &\begin{bmatrix} I & -R_{xy}R_{yy}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \begin{bmatrix} I & 0 \\ -R_{yy}^{-1}R_{xy}^\top & I \end{bmatrix} \\ &= \begin{bmatrix} R_{xx} - R_{xy}R_{yy}^{-1}R_{yx} & 0 \\ 0 & R_{yy} \end{bmatrix}. \end{aligned}$$

This yields, by taking determinants,

$$det \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = det\big[R_{xx} - R_{xy}R_{yy}^{-1}R_{yx}\big] \cdot det R_{yy}$$

and

$$\begin{aligned} &\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix}\right)^\top \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}^{-1} \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} - \begin{bmatrix} \underline{\mu}_x \\ \underline{\mu}_y \end{bmatrix}\right) \\ &= (\mathbf{x} - \tilde{\underline{\mu}})^\top \big[R_{xx} - R_{xy}R_{yy}^{-1}R_{yx}\big]^{-1} (\mathbf{x} - \tilde{\underline{\mu}}) + (\mathbf{y} - \underline{\mu}_y)^\top R_{yy}^{-1} (\mathbf{y} - \underline{\mu}_y)\,, \end{aligned}$$

where

$$\tilde{\underline{\mu}} = \underline{\mu}_x + R_{xy}R_{yy}^{-1}(\mathbf{y} - \underline{\mu}_y)\,.$$

The remaining computational steps are straightforward.
1.11. Let $\mathbf{p}_k = C_k^\top W_k \mathbf{z}_k$ and $\sigma^2 = E\big[\mathbf{p}_k^\top (C_k^\top W_k C_k)^{-1} \mathbf{p}_k\big]$. Then it can be easily verified that

$$F(\mathbf{y}_k) = \mathbf{y}_k^\top (C_k^\top W_k C_k)\mathbf{y}_k - \mathbf{p}_k^\top \mathbf{y}_k - \mathbf{y}_k^\top \mathbf{p}_k + \sigma^2\,.$$

From

$$\frac{dF(\mathbf{y}_k)}{d\mathbf{y}_k} = 2(C_k^\top W_k C_k)\mathbf{y}_k - 2\mathbf{p}_k = 0\,,$$

and the assumption that the matrix $(C_k^\top W_k C_k)$ is nonsingular, we have

$$\hat{\mathbf{y}}_k = (C_k^\top W_k C_k)^{-1}\mathbf{p}_k = (C_k^\top W_k C_k)^{-1} C_k^\top W_k \mathbf{z}_k \,.$$

1.12.

$$\begin{aligned} E\hat{\mathbf{x}}_k &= (C_k^\top R_k^{-1} C_k)^{-1} C_k^\top R_k^{-1} E(\mathbf{v}_k - D_k \mathbf{u}_k) \\ &= (C_k^\top R_k^{-1} C_k)^{-1} C_k^\top R_k^{-1} E(C_k \mathbf{x}_k + \underline{\eta}_k) \\ &= E\mathbf{x}_k \,. \end{aligned}$$

Chapter 2

2.1.

$$\begin{aligned} W_{k,k-1}^{-1} &= Var(\overline{\underline{\epsilon}}_{k,k-1}) = E(\overline{\underline{\epsilon}}_{k,k-1} \overline{\underline{\epsilon}}_{k,k-1}^\top) \\ &= E(\overline{\mathbf{v}}_{k-1} - H_{k,k-1}\mathbf{x}_k)(\overline{\mathbf{v}}_{k-1} - H_{k,k-1}\mathbf{x}_k)^\top \\ &= \begin{bmatrix} R_0 & & \\ & \ddots & \\ & & R_{k-1} \end{bmatrix} + Var \begin{bmatrix} C_0 \sum_{i=1}^k \Phi_{0i} \Gamma_{i-1} \underline{\xi}_{i-1} \\ \vdots \\ C_{k-1} \Phi_{k-1,k} \Gamma_{k-1} \underline{\xi}_{k-1} \end{bmatrix} . \end{aligned}$$

2.2. For any nonzero vector $\mathbf{x}$, we have $\mathbf{x}^\top A\mathbf{x} > 0$ and $\mathbf{x}^\top B\mathbf{x} \geq 0$ so that

$$\mathbf{x}^\top (A+B)\mathbf{x} = \mathbf{x}^\top A\mathbf{x} + \mathbf{x}^\top B\mathbf{x} > 0 \,.$$

Hence, $A + B$ is positive definite.
2.3.

$$\begin{aligned} & W_{k,k-1}^{-1} \\ =& E(\overline{\underline{\epsilon}}_{k,k-1} \overline{\underline{\epsilon}}_{k,k-1}^\top) \\ =& E(\overline{\underline{\epsilon}}_{k-1,k-1} - H_{k,k-1}\Gamma_{k-1}\underline{\xi}_{k-1})(\overline{\underline{\epsilon}}_{k-1,k-1} - H_{k,k-1}\Gamma_{k-1}\underline{\xi}_{k-1})^\top \\ =& E(\overline{\underline{\epsilon}}_{k-1,k-1} \overline{\underline{\epsilon}}_{k-1,k-1}^\top) + H_{k,k-1}\Gamma_{k-1} E(\underline{\xi}_{k-1}\underline{\xi}_{k-1}^\top)\Gamma_{k-1}^\top H_{k,k-1}^\top \\ =& W_{k-1,k-1}^{-1} + H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \Phi_{k-1,k}^\top H_{k-1,k-1}^\top \,. \end{aligned}$$

2.4. Apply Lemma 1.2 to $A_{11} = W_{k-1,k-1}^{-1}$, $A_{22} = Q_{k-1}^{-1}$ and

$$A_{12} = A_{21}^\top = H_{k-1,k-1}\Phi_{k-1,k}\Gamma_{k-1} \,.$$

2.5. Using Exercise 2.4 (or (2.9)), we have

$$
\begin{aligned}
& H_{k,k-1}^{\top} W_{k,k-1} \\
=& \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k-1,k-1} \\
& - \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k,k-1} H_{k,k-1} \Phi_{k-1,k} \Gamma_{k-1} \\
& \cdot (Q_{k-1}^{-1} + \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
& \cdot \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k-1,k-1} \\
=& \Phi_{k-1,k}^{\top} \{ I - H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} \\
& \cdot (Q_{k-1}^{-1} + \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
& \cdot \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} \} H_{k-1,k-1}^{\top} W_{k-1,k-1} \,.
\end{aligned}
$$

2.6. Using Exercise 2.5 (or (2.10)) and the identity $H_{k,k-1} = H_{k-1,k-1} \Phi_{k-1,k}$, we have

$$
\begin{aligned}
& (H_{k,k-1}^{\top} W_{k,k-1} H_{k,k-1}) \Phi_{k,k-1} \\
& \cdot (H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1})^{-1} H_{k-1,k-1}^{\top} W_{k-1,k-1} \\
=& \Phi_{k-1,k}^{\top} \{ I - H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} \\
& \cdot (Q_{k-1}^{-1} + \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} H_{k-1,k-1}^{\top} W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
& \cdot \Gamma_{k-1}^{\top} \Phi_{k-1,k}^{\top} \} H_{k-1,k-1}^{\top} W_{k-1,k-1} \\
=& H_{k,k-1}^{\top} W_{k,k-1} \,.
\end{aligned}
$$

2.7.

$$
\begin{aligned}
& P_{k,k-1} C_k^{\top} (C_k P_{k,k-1} C_k^{\top} + R_k)^{-1} \\
=& P_{k,k-1} C_k^{\top} (R_k^{-1} - R_k^{-1} C_k (P_{k,k-1}^{-1} + C_k^{\top} R_k^{-1} C_k)^{-1} C_k^{\top} R_k^{-1}) \\
=& (P_{k,k-1} - P_{k,k-1} C_k^{\top} R_k^{-1} C_k (P_{k,k-1}^{-1} + C_k^{\top} R_k^{-1} C_k)^{-1}) C_k^{\top} R_k^{-1} \\
=& (P_{k,k-1} - P_{k,k-1} C_k^{\top} (C_k P_{k,k-1} C_k^{\top} + R_k)^{-1} \\
& \cdot (C_k P_{k,k-1} C_k^{\top} + R_k) R_k^{-1} C_k (P_{k,k-1}^{-1} + C_k^{\top} R_k^{-1} C_k)^{-1}) C_k^{\top} R_k^{-1} \\
=& (P_{k,k-1} - P_{k,k-1} C_k^{\top} (C_k P_{k,k-1} C_k^{\top} + R_k)^{-1} \\
& \cdot (C_k P_{k,k-1} C_k^{\top} R_k^{-1} C_k + C_k)(P_{k,k-1}^{-1} + C_k^{\top} R_k^{-1} C_k)^{-1}) C_k^{\top} R_k^{-1} \\
=& (P_{k,k-1} - P_{k,k-1} C_k^{\top} (C_k P_{k,k-1} C_k^{\top} + R_k)^{-1} C_k P_{k,k-1} \\
& \cdot (C_k^{\top} R_k^{-1} C_k + P_{k,k-1}^{-1})(P_{k,k-1}^{-1} + C_k^{\top} R_k^{-1} C_k)^{-1}) C_k^{\top} R_k^{-1} \\
=& (P_{k,k-1} - P_{k,k-1} C_k^{\top} (C_k P_{k,k-1} C_k^{\top} + R_k)^{-1} C_k P_{k,k-1}) C_k^{\top} R_k^{-1}
\end{aligned}
$$

$$
\begin{aligned}
&= P_{k,k} C_k^\top R_k^{-1} \\
&= G_k \,.
\end{aligned}
$$

2.8.

$$
\begin{aligned}
& P_{k,k-1} \\
&= (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} \\
&= (\Phi_{k-1,k}^\top (H_{k-1,k-1}^\top W_{k-1,k-1} H_{k-1,k-1} \\
&\quad - H_{k-1,k-1}^\top W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1} \\
&\quad \cdot (Q_{k-1}^{-1} + \Gamma_{k-1}^\top \Phi_{k-1,k}^\top H_{k-1,k-1}^\top W_{k-1,k-1} H_{k-1,k-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
&\quad \cdot \Gamma_{k-1}^\top \Phi_{k-1,k}^\top H_{k-1,k-1}^\top W_{k-1,k-1} H_{k-1,k-1}) \Phi_{k-1,k})^{-1} \\
&= (\Phi_{k-1,k}^\top P_{k-1,k-1}^{-1} \Phi_{k-1,k} - \Phi_{k-1,k}^\top P_{k-1,k-1}^{-1} \Phi_{k-1,k} \Gamma_{k-1} \\
&\quad \cdot (Q_{k-1}^{-1} + \Gamma_{k-1}^\top \Phi_{k-1,k}^\top P_{k-1,k-1}^{-1} \Phi_{k-1,k} \Gamma_{k-1})^{-1} \\
&\quad \cdot \Gamma_{k-1}^\top \Phi_{k-1,k}^\top P_{k-1,k-1}^{-1} \Phi_{k-1,k})^{-1} \\
&= (\Phi_{k-1,k}^\top P_{k-1,k-1}^{-1} \Phi_{k-1,k})^{-1} + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \\
&= A_{k-1} P_{k-1,k-1} A_{k-1}^\top + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \,.
\end{aligned}
$$

2.9.

$$
\begin{aligned}
& E(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^\top \\
&= E(\mathbf{x}_k - (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} H_{k,k-1}^\top W_{k,k-1} \overline{\mathbf{v}}_{k-1}) \\
&\quad \cdot (\mathbf{x}_k - (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} H_{k,k-1}^\top W_{k,k-1} \overline{\mathbf{v}}_{k-1})^\top \\
&= E(\mathbf{x}_k - (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} H_{k,k-1}^\top W_{k,k-1} \\
&\quad \cdot (H_{k,k-1} \mathbf{x}_k + \underline{\overline{\epsilon}}_{k,k-1}))(\mathbf{x}_k - (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} \\
&\quad \cdot H_{k,k-1}^\top W_{k,k-1} (H_{k,k-1} \mathbf{x}_k + \underline{\overline{\epsilon}}_{k,k-1}))^\top \\
&= (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} H_{k,k-1}^\top W_{k,k-1} E(\underline{\overline{\epsilon}}_{k,k-1} \underline{\overline{\epsilon}}_{k,k-1}^\top) W_{k,k-1} \\
&\quad \cdot H_{k,k-1} (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} \\
&= (H_{k,k-1}^\top W_{k,k-1} H_{k,k-1})^{-1} \\
&= P_{k,k-1} \,.
\end{aligned}
$$

The derivation of the second identity is similar.

2.10. Since

$$\begin{aligned}\sigma^2 &= Var(x_k) = E(ax_{k-1} + \xi_{k-1})^2 \\ &= a^2 Var(x_{k-1}) + 2aE(x_{k-1}\xi_{k-1}) + E(\xi_{k-1}^2) \\ &= a^2\sigma^2 + \mu^2 ,\end{aligned}$$

we have

$$\sigma^2 = \mu^2/(1-a^2) .$$

For $j = 1$, we have

$$\begin{aligned}E(x_k x_{k+1}) &= E(x_k(ax_k + \xi_k)) \\ &= aVar(x_k) + E(x_k\xi_k) \\ &= a\sigma^2 .\end{aligned}$$

For $j = 2$, we have

$$\begin{aligned}E(x_k x_{k+2}) &= E(x_k(ax_{k+1} + \xi_{k+1})) \\ &= aE(x_k x_{k+1}) + E(x_k + \xi_{k+1}) \\ &= aE(x_k x_{k+1}) \\ &= a^2\sigma^2 ,\end{aligned}$$

etc. If j is negative, then a similar result can be obtained. By induction, we may conclude that $E(x_k x_{k+j}) = a^{|j|}\sigma^2$ for all integers j.
2.11. Using the Kalman filtering equations (2.17), we have

$$\begin{aligned}&P_{0,0} = Var(x_0) = \mu^2 , \\ &P_{k,k-1} = P_{k-1,k-1} , \\ &G_k = P_{k,k-1}(P_{k,k-1} + R_k)^{-1} = \frac{P_{k-1,k-1}}{P_{k-1,k-1} + \sigma^2} ,\end{aligned}$$

and

$$P_{k,k} = (1 - G_k)P_{k,k-1} = \frac{\sigma^2 P_{k-1,k-1}}{\sigma^2 + P_{k-1,k-1}} .$$

Observe that

$$P_{1,1} = \frac{\sigma^2\mu^2}{\mu^2+\sigma^2}\,,$$
$$P_{2,2} = \frac{\sigma^2 P_{1,1}}{P_{1,1}+\sigma^2} = \frac{\sigma^2\mu^2}{2\mu^2+\sigma^2}\,.$$
$$\cdots$$
$$P_{k,k} = \frac{\sigma^2\mu^2}{k\mu^2+\sigma^2}\,.$$

Hence,

$$G_k = \frac{P_{k-1,k-1}}{P_{k-1,k-1}+\sigma^2} = \frac{\mu^2}{k\mu^2+\sigma^2}$$

so that

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k(v_k - \hat{x}_{k|k-1})\\ &= \hat{x}_{k-1|k-1} + \frac{\mu^2}{\sigma^2+k\mu^2}(v_k - \hat{x}_{k-1|k-1})\end{aligned}$$

with $\hat{x}_{0|0} = E(x_0) = 0$. It follows that

$$\hat{x}_{k|k} = \hat{x}_{k-1|k-1}$$

for large values of k.
2.12.

$$\begin{aligned}\hat{Q}_N &= \frac{1}{N}\sum_{k=1}^{N}(\mathbf{v}_k\mathbf{v}_k^\top)\\ &= \frac{1}{N}(\mathbf{v}_N\mathbf{v}_N^\top) + \frac{1}{N}\sum_{k=1}^{N-1}(\mathbf{v}_k\mathbf{v}_k^\top)\\ &= \frac{1}{N}(\mathbf{v}_N\mathbf{v}_N^\top) + \frac{N-1}{N}\hat{Q}_{N-1}\\ &= \hat{Q}_{N-1} + \frac{1}{N}[(\mathbf{v}_N\mathbf{v}_N^\top) - \hat{Q}_{N-1}]\end{aligned}$$

with the initial estimation $\hat{Q}_1 = \mathbf{v}_1\mathbf{v}_1^\top$.
2.13. Use superimposition.

2.14. Set $\mathbf{x}_k = [(\mathbf{x}_k^1)^\top \cdots (\mathbf{x}_k^N)^\top]^\top$ for each k, $k = 0, 1, \cdots$, with $\mathbf{x}_j = 0$ (and $\mathbf{u}_j = 0$) for $j < 0$, and define

$$\begin{aligned}
\mathbf{x}_k^1 &= B_1\mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^2 + (A_1 + B_1 A_0)\mathbf{u}_{k-1}\,, \\
&\cdots\cdots \\
\mathbf{x}_k^M &= B_M\mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^{M+1} + (A_M + B_M A_0)\mathbf{u}_{k-1}\,, \\
\mathbf{x}_k^{M+1} &= B_{M+1}\mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^{M+2} + B_{M+1}A_0\mathbf{u}_{k-1}\,, \\
&\cdots\cdots \\
\mathbf{x}_k^{N-1} &= B_{N-1}\mathbf{x}_{k-1}^1 + \mathbf{x}_{k-1}^{N} + B_{N-1}A_0\mathbf{u}_{k-1}\,, \\
\mathbf{x}_k^{N} &= B_N\mathbf{x}_{k-1}^1 + B_N A_0\mathbf{u}_{k-1}\,.
\end{aligned}$$

Then, substituting these equations into

$$\mathbf{v}_k = C\mathbf{x}_k + D\mathbf{u}_k = \mathbf{x}_k^1 + A_0\mathbf{u}_k$$

yields the required result. Since $\mathbf{x}_j = 0$ and $\mathbf{u}_j = 0$ for $j < 0$, it is also clear that $\mathbf{x}_0 = 0$.

Chapter 3

3.1. Let $A = BB^\top$ where $B = [b_{ij}] \neq 0$. Then $\mathrm{tr}A = \mathrm{tr}BB^\top = \sum_{i,j} b_{ij}^2 > 0$.
3.2. By Assumption 2.1, $\underline{\eta}_\ell$ is independent of $\mathbf{x}_0$, $\underline{\xi}_0, \cdots, \underline{\xi}_{j-1}, \underline{\eta}_0, \cdots, \underline{\eta}_{j-1}$, since $\ell \geq j$. On the other hand,

$$\begin{aligned}
\hat{\mathbf{e}}_j &= C_j(\mathbf{x}_j - \hat{\mathbf{y}}_{j-1}) \\
&= C_j\Bigg(A_{j-1}\mathbf{x}_{j-1} + \Gamma_{j-1}\underline{\xi}_{j-1} - \sum_{i=0}^{j-1}\hat{P}_{j-1,i}(C_i\mathbf{x}_i + \underline{\eta}_i)\Bigg) \\
&\cdots\cdots \\
&= B_0\mathbf{x}_0 + \sum_{i=0}^{j-1} B_{1i}\underline{\xi}_i + \sum_{i=0}^{j-1} B_{2i}\underline{\eta}_i
\end{aligned}$$

for some constant matrices B_0, B_{1i} and B_{2i}. Hence, $\langle \underline{\eta}_\ell, \hat{\mathbf{e}}_j\rangle = O_{q\times q}$ for all $\ell \geq j$.
3.3. Combining (3.8) and (3.4), we have

$$\mathbf{e}_j = \|\mathbf{z}_j\|_q^{-1}\mathbf{z}_j = \|\mathbf{z}_j\|_q^{-1}\mathbf{v}_j - \sum_{i=0}^{j-1}\Big(\|\mathbf{z}_j\|_q^{-1}C_j\hat{P}_{j-1,i}\Big)\mathbf{v}_i\,;$$

that is, $\mathbf{e}_j$ can be expressed in terms of $\mathbf{v}_0, \mathbf{v}_1, \cdots, \mathbf{v}_j$. Conversely, we have

$$
\begin{aligned}
\mathbf{v}_0 &= \mathbf{z}_0 = \|\mathbf{z}_0\|_q \mathbf{e}_0 \,, \\
\mathbf{v}_1 &= \mathbf{z}_1 + C_1 \hat{\mathbf{y}}_0 = \mathbf{z}_1 + C_1 \hat{P}_{0,0} \mathbf{v}_0 \\
&= \|\mathbf{z}_1\|_q \mathbf{e}_1 + C_1 \hat{P}_{0,0} \|\mathbf{z}_0\|_q \mathbf{e}_0 \,, \\
&\cdots\cdots
\end{aligned}
$$

that is, $\mathbf{v}_j$ can also be expressed in terms of $\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_j$. Hence, we have

$$
Y(\mathbf{e}_0, \ \cdots, \ \mathbf{e}_k) = Y(\mathbf{v}_0, \ \cdots, \ \mathbf{v}_k) \,.
$$

3.4. By Exercise 3.3, we have

$$
\mathbf{v}_i = \sum_{\ell=0}^{i} L_\ell \mathbf{e}_\ell
$$

for some $q \times q$ constant matrices L_ℓ, $\ell = 0, 1, \cdots, i$, so that

$$
\langle \mathbf{v}_i, \ \mathbf{z}_k \rangle = \sum_{\ell=0}^{i} L_\ell \langle \mathbf{e}_\ell, \ \mathbf{e}_k \rangle \|\mathbf{z}_k\|_q^\top = O_{q\times q} \,,
$$

$i = 0, 1, \cdots, k-1$. Hence, for $j = 0, 1, \cdots, k-1$,

$$
\begin{aligned}
\langle \hat{\mathbf{y}}_j, \ \mathbf{z}_k \rangle &= \left\langle \sum_{i=0}^{j} \hat{P}_{j,i} \mathbf{v}_i, \ \mathbf{z}_k \right\rangle \\
&= \sum_{i=0}^{j} \hat{P}_{j,i} \langle \mathbf{v}_i, \ \mathbf{z}_k \rangle \\
&= O_{n\times q} \,.
\end{aligned}
$$

3.5. Since

$$
\begin{aligned}
\mathbf{x}_k &= A_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1} \\
&= A_{k-1}(A_{k-2}\mathbf{x}_{k-2} + \Gamma_{k-2}\underline{\xi}_{k-2}) + \Gamma_{k-1}\underline{\xi}_{k-1} \\
&= \cdots\cdots \\
&= B_0\mathbf{x}_0 + \sum_{i=0}^{k-1} B_{1i}\underline{\xi}_i
\end{aligned}
$$

for some constant matrices B_0 and B_{1i} and $\underline{\xi}_k$ is independent of $\mathbf{x}_0$ and $\underline{\xi}_i$ $(0 \le i \le k-1)$, we have $\langle \mathbf{x}_k, \ \underline{\xi}_k \rangle = 0$. The rest can be shown in a similar manner.

3.6. Use superimposition.
3.7. Using the formula obtained in Exercise 3.6, we have

$$\begin{cases} \hat{d}_{k|k} = \hat{d}_{k-1|k-1} + hw_{k-1} + G_k(v_k - \Delta d_k - \hat{d}_{k-1|k-1} - hw_{k-1}) \\ \hat{d}_{0|0} = E(d_0)\,, \end{cases}$$

where G_k is obtained by using the standard algorithm (3.25) with $A_k = C_k = \Gamma_k = 1$.
3.8. Let

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{x}_k^1 \\ \mathbf{x}_k^2 \\ \mathbf{x}_k^3 \end{bmatrix}, \quad \mathbf{x}_k^1 = \begin{bmatrix} \Sigma_k \\ \dot{\Sigma}_k \\ \ddot{\Sigma}_k \end{bmatrix}, \quad \mathbf{x}_k^2 = \begin{bmatrix} \Delta A_k \\ \Delta \dot{A}_k \\ \Delta \ddot{A}_k \end{bmatrix}, \quad \mathbf{x}_k^3 = \begin{bmatrix} \Delta E_k \\ \Delta \dot{E}_k \\ \Delta \ddot{E}_k \end{bmatrix},$$

$$\underline{\xi}_k = \begin{bmatrix} \underline{\xi}_k^1 \\ \underline{\xi}_k^2 \\ \underline{\xi}_k^3 \end{bmatrix}, \qquad \underline{\eta}_k = \begin{bmatrix} \eta_k^1 \\ \eta_k^2 \\ \eta_k^3 \end{bmatrix}, \qquad \mathbf{v}_k = \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \end{bmatrix},$$

$$A = \begin{bmatrix} 1 & h & h^2/2 \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}, \qquad \textit{and} \qquad C = [\,1\;0\;0\,]\,.$$

Then the system described in Exercise 3.8 can be decomposed into three subsystems:

$$\begin{cases} \mathbf{x}_{k+1}^i = A\mathbf{x}_k^i + \Gamma_k^i \underline{\xi}_k^i \\ v_k^i = C\mathbf{x}_k^i + \eta_k^i\,, \end{cases}$$

$i = 1, 2, 3$, where for each k, $\mathbf{x}_k$ and $\underline{\xi}_k$ are 3-vectors, v_k and η_k are scalars, Q_k a 3×3 non-negative definite symmetric matrix, and $R_k > 0$ a scalar.

Chapter 4

4.1. Using (4.6), we have

$$\begin{aligned} & L(A\mathbf{x} + B\mathbf{y},\ \mathbf{v}) \\ &= E(A\mathbf{x} + B\mathbf{y}) + \langle A\mathbf{x} + B\mathbf{y},\ \mathbf{v}\rangle \big[Var(\mathbf{v})\big]^{-1}(\mathbf{v} - E(\mathbf{v})) \\ &= A\big\{E(\mathbf{x}) + \langle \mathbf{x},\ \mathbf{v}\rangle \big[Var(\mathbf{v})\big]^{-1}(\mathbf{v} - E(\mathbf{v}))\big\} \\ &\quad + B\big\{E(\mathbf{y}) + \langle \mathbf{y},\ \mathbf{v}\rangle \big[Var(\mathbf{v})\big]^{-1}(\mathbf{v} - E(\mathbf{v}))\big\} \\ &= AL(\mathbf{x},\ \mathbf{v}) + BL(\mathbf{y},\ \mathbf{v})\,. \end{aligned}$$

4.2. Using (4.6) and the fact that $E(\mathbf{a}) = \mathbf{a}$ so that

$$\langle \mathbf{a},\ \mathbf{v}\rangle = E(\mathbf{a} - E(\mathbf{a}))\,(\mathbf{v} - E(\mathbf{v})) = 0\,,$$

we have

$$L(\mathbf{a},\ \mathbf{v}) = E(\mathbf{a}) + \langle \mathbf{a},\ \mathbf{v}\rangle\big[Var(\mathbf{v})\big]^{-1}(\mathbf{v} - E(\mathbf{v})) = \mathbf{a}\,.$$

4.3. By definition, for a real-valued function f and a matrix $A = [a_{ij}]$, $df/dA = [\partial f/\partial a_{ji}]$. Hence,

$$\begin{aligned}
0 &= \frac{\partial}{\partial H}\Big(\mathrm{tr}\|\mathbf{x} - \mathbf{y}\|_n^2\Big)\\
&= \frac{\partial}{\partial H}E((\mathbf{x} - E(\mathbf{x})) - H(\mathbf{v} - E(\mathbf{v})))^\top((\mathbf{x} - E(\mathbf{x})) - H(\mathbf{v} - E(\mathbf{v})))\\
&= E\frac{\partial}{\partial H}((\mathbf{x} - E(\mathbf{x})) - H(\mathbf{v} - E(\mathbf{v})))^\top((\mathbf{x} - E(\mathbf{x})) - H(\mathbf{v} - E(\mathbf{v})))\\
&= E(-2(\mathbf{x} - E(\mathbf{x})) - H(\mathbf{v} - E(\mathbf{v})))\,(\mathbf{v} - E(\mathbf{v}))^\top\\
&= 2\big(H\ E(\mathbf{v} - E(\mathbf{v}))\,(\mathbf{v} - E(\mathbf{v}))^\top - E(\mathbf{x} - E(\mathbf{x}))\,(\mathbf{v} - E(\mathbf{v}))^\top\big)\\
&= 2\big(H\|\mathbf{v}\|_q^2 - \langle \mathbf{x},\ \mathbf{v}\rangle\big)\,.
\end{aligned}$$

This gives

$$H^* = \langle \mathbf{x},\ \mathbf{v}\rangle\Big[\|\mathbf{v}\|_q^2\Big]^{-1}$$

so that

$$\mathbf{x}^* = E(\mathbf{x}) - \langle \mathbf{x},\ \mathbf{v}\rangle\Big[\|\mathbf{v}\|_q^2\Big]^{-1}(E(\mathbf{v}) - \mathbf{v})\,.$$

4.4. Since $\mathbf{v}^{k-2}$ is a linear combination (with constant matrix coefficients) of

$$\mathbf{x}_0,\ \underline{\xi}_0,\ \cdots,\ \underline{\xi}_{k-3},\ \underline{\eta}_0,\ \cdots,\ \underline{\eta}_{k-2}$$

which are all uncorrelated with $\underline{\xi}_{k-1}$ and $\underline{\eta}_{k-1}$, we have

$$\langle \underline{\xi}_{k-1},\ \mathbf{v}^{k-2}\rangle = 0 \qquad \textit{and} \qquad \langle \underline{\eta}_{k-1},\ \mathbf{v}^{k-2}\rangle = 0\,.$$

Similarly, we can verify the other formulas (where (4.6) may be used).
4.5. The first identity follows from the Kalman gain equation (cf. Theorem 4.1(c) or (4.19)), namely:

$$G_k(C_k P_{k,k-1} C_k^\top + R_k) = P_{k,k-1} C_k^\top\,,$$

so that

$$\begin{aligned} G_k R_k &= P_{k,k-1} C_k^\top - G_k C_k P_{k,k-1} C_k^\top \\ &= (I - G_k C_k) P_{k,k-1} C_k^\top . \end{aligned}$$

To prove the second equality, we apply (4.18) and (4.17) to obtain

$$\begin{aligned} &\langle \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1},\ \Gamma_{k-1}\underline{\xi}_{k-1} - K_{k-1}\underline{\eta}_{k-1}\rangle \\ &= \langle \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-2} - \langle \mathbf{x}^{\#}{}_{k-1},\ \mathbf{v}^{\#}{}_{k-1}\rangle \Big[\|\mathbf{v}^{\#}{}_{k-1}\|^2\Big]^{-1} \mathbf{v}^{\#}{}_{k-1}, \\ &\qquad \Gamma_{k-1}\underline{\xi}_{k-1} - K_{k-1}\underline{\eta}_{k-1}\rangle \\ &= \langle \mathbf{x}^{\#}{}_{k-1} - \langle \mathbf{x}^{\#}{}_{k-1},\ \mathbf{v}^{\#}{}_{k-1}\rangle \Big[\|\mathbf{v}^{\#}{}_{k-1}\|^2\Big]^{-1} (C_{k-1}\mathbf{x}^{\#}{}_{k-1} + \underline{\eta}_{k-1}), \\ &\qquad \Gamma_{k-1}\underline{\xi}_{k-1} - K_{k-1}\underline{\eta}_{k-1}\rangle \\ &= -\langle \mathbf{x}^{\#}{}_{k-1},\ \mathbf{v}^{\#}{}_{k-1}\rangle \Big[\|\mathbf{v}^{\#}{}_{k-1}\|^2\Big]^{-1} (S_{k-1}^\top \Gamma_{k-1}^\top - R_{k-1} K_{k-1}^\top) \\ &= O_{n\times n} , \end{aligned}$$

in which since $K_{k-1} = \Gamma_{k-1} S_{k-1} R_{k-1}^{-1}$, we have

$$S_{k-1}^\top \Gamma_{k-1}^\top - R_{k-1} K_{k-1}^\top = O_{n\times n} .$$

4.6. Follow the same procedure in the derivation of Theorem 4.1 with the term $\mathbf{v}_k$ replaced by $\mathbf{v}_k - D_k \mathbf{u}_k$, and with

$$\hat{\mathbf{x}}_{k|k-1} = L(A_{k-1}\mathbf{x}_{k-1} + B_{k-1}\mathbf{u}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1},\ \mathbf{v}^{k-1})$$

instead of

$$\hat{\mathbf{x}}_{k|k-1} = L(\mathbf{x}_k,\ \mathbf{v}^{k-1}) = L(A_{k-1}\mathbf{x}_{k-1} + \Gamma_{k-1}\underline{\xi}_{k-1},\ \mathbf{v}^{k-1}) .$$

4.7. Let

$$\begin{aligned} w_k &= -a_1 v_{k-1} + b_1 u_{k-1} + c_1 e_{k-1} + w_{k-1} , \\ w_{k-1} &= -a_2 v_{k-2} + b_2 u_{k-2} + w_{k-2} , \\ w_{k-2} &= -a_3 v_{k-3} , \end{aligned}$$

and define $\mathbf{x}_k = [\ w_k\ \ w_{k-1}\ \ w_{k-2}\ ]^\top$. Then,

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k + B u_k + \Gamma e_k \\ \quad v_k = C\mathbf{x}_k + D u_k + \Delta e_k , \end{cases}$$

where

$$A = \begin{bmatrix} -a_1 & 1 & 0 \\ -a_2 & 0 & 1 \\ -a_3 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 - a_1 b_0 \\ b_2 - a_2 b_0 \\ -a_3 b_0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} c_1 - a_1 c_0 \\ -a_2 c_0 \\ -a_3 b_0 \end{bmatrix},$$

$$C = [\ 1\ \ 0\ \ 0\ ], \qquad D = [b_0] \qquad \textit{and} \qquad \Delta = [c_0]\,.$$

4.8. Let

$$\begin{aligned} w_k &= -a_1 v_{k-1} + b_1 u_{k-1} + c_1 e_{k-1} + w_{k-1}\,, \\ w_{k-1} &= -a_2 v_{k-2} + b_2 u_{k-2} + c_2 e_{k-2} + w_{k-2}\,, \\ &\cdots\cdots \\ w_{k-n+1} &= -a_n v_{k-n} + b_n u_{k-n} + c_n e_{k-n}\,, \end{aligned}$$

where $b_j = 0$ for $j > m$ and $c_j = 0$ for $j > \ell$, and define

$$\mathbf{x}_k = [\ w_k \ \ w_{k-1} \ \ \cdots \ \ w_{k-n+1}\ ]^\top .$$

Then

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k + Bu_k + \Gamma e_k \\ v_k = C\mathbf{x}_k + Du_k + \Delta e_k\,, \end{cases}$$

where

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} b_1 - a_1 b_0 \\ \vdots \\ b_m - a_m b_0 \\ -a_{m+1} b_0 \\ \vdots \\ -a_n b_0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} c_1 - a_1 c_0 \\ \vdots \\ c_\ell - a_\ell c_0 \\ -a_{\ell+1} \\ \vdots \\ -a_n c_0 \end{bmatrix},$$

$$C = [\ 1\ \ 0\ \ \cdots\cdots\ \ 0\ ], \qquad D = [b_0]\,, \qquad \textit{and} \qquad \Delta = [c_0]\,.$$

Chapter 5

5.1. Since $\mathbf{v}^k$ is a linear combination (with constant matrices as coefficients) of

$$\mathbf{x}_0,\ \underline{\eta}_0,\ \underline{\gamma}_0,\ \cdots,\ \underline{\gamma}_k,\ \underline{\xi}_0,\ \underline{\beta}_0,\ \cdots,\ \underline{\beta}_{k-1}$$

which are all independent of $\underline{\beta}_k$, we have

$$\langle \underline{\beta}_k,\ \mathbf{v}^k \rangle = 0\,.$$

On the other hand, $\underline{\beta}_k$ has zero-mean, so that by (4.6) we have

$$L(\tilde{\underline{\beta}}_k,\ \mathbf{v}^k) = E(\tilde{\underline{\beta}}_k) - \langle \tilde{\underline{\beta}}_k,\ \mathbf{v}^k \rangle \Big[\|\mathbf{v}^k\|^2 \Big]^{-1} \big(E(\mathbf{v}^k) - \mathbf{v}^k\big) = 0\,.$$

5.2. Using Lemma 4.2 with $\mathbf{v} = \mathbf{v}^{k-1},\ \mathbf{v}^1 = \mathbf{v}^{k-2},\ \mathbf{v}^2 = \mathbf{v}_{k-1}$ and

$$\mathbf{v}^{\#}{}_{k-1} = \mathbf{v}_{k-1} - L(\mathbf{v}_{k-1},\ \mathbf{v}^{k-2})\,,$$

we have, for $\mathbf{x} = \mathbf{v}_{k-1}$,

$$\begin{aligned}
&L(\mathbf{v}_{k-1},\ \mathbf{v}^{k-1})\\
&= L(\mathbf{v}_{k-1},\ \mathbf{v}^{k-2}) + \langle \mathbf{v}^{\#}{}_{k-1},\ \mathbf{v}^{\#}{}_{k-1} \rangle \Big[\|\mathbf{v}^{\#}{}_{k-1}\|^2 \Big]^{-1} \mathbf{v}^{\#}{}_{k-1}\\
&= L(\mathbf{v}_{k-1},\ \mathbf{v}^{k-2}) + \mathbf{v}_{k-1} - L(\mathbf{v}_{k-1},\ \mathbf{v}^{k-2})\\
&= \mathbf{v}_{k-1}\,.
\end{aligned}$$

The equality $L(\underline{\gamma}_k,\ \mathbf{v}^{k-1}) = 0$ can be shown by imitating the proof in Exercise 5.1.

5.3. It follows from Lemma 4.2 that

$$\begin{aligned}
&\mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1}\\
&= \mathbf{z}_{k-1} - L(\mathbf{z}_{k-1},\ \mathbf{v}^{k-1})\\
&= \mathbf{z}_{k-1} - E(\mathbf{z}_{k-1}) + \langle \mathbf{z}_{k-1},\ \mathbf{v}^{k-1} \rangle \Big[\|\mathbf{v}^{k-1}\|^2 \Big]^{-1} \big(E(\mathbf{v}^{k-1}) - \mathbf{v}^{k-1}\big)\\
&= \begin{bmatrix} \mathbf{x}_{k-1} \\ \underline{\xi}_{k-1} \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_{k-1}) \\ E(\underline{\xi}_{k-1}) \end{bmatrix}\\
&\quad + \begin{bmatrix} \langle \mathbf{x}_{k-1}, \mathbf{v}^{k-1} \rangle \\ \langle \underline{\xi}_{k-1}, \mathbf{v}^{k-1} \rangle \end{bmatrix} \Big[\|\mathbf{v}^{k-1}\|^2 \Big]^{-1} \big(E(\mathbf{v}^{k-1}) - \mathbf{v}^{k-1}\big)
\end{aligned}$$

whose first n-subvector and last p-subvector are, respectively, linear combinations (with constant matrices as coefficients) of

$$\mathbf{x}_0,\ \underline{\xi}_0,\ \underline{\beta}_0,\ \cdots,\ \underline{\beta}_{k-2},\ \underline{\eta}_0,\ \underline{\gamma}_0,\ \cdots,\ \underline{\gamma}_{k-1}\,,$$

which are all independent of $\underline{\gamma}_k$. Hence, we have

$$B\langle \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1},\ \underline{\gamma}_k \rangle = 0\,.$$

5.4. The proof is similar to that of Exercise 5.3.
5.5. For simplicity, denote

$$B = [C_0 Var(\mathbf{x}_0) C_0^\top + R_0]^{-1}\,.$$

It follows from (5.16) that

$$\begin{aligned}
& Var(\mathbf{x}_0 - \hat{\mathbf{x}}_0) \\
&= Var(\mathbf{x}_0 - E(\mathbf{x}_0) \\
&\quad - [Var(\mathbf{x}_0)]C_0^\top [C_0 Var(\mathbf{x}_0) C_0^\top + R_0]^{-1}(\mathbf{v}_0 - C_0 E(\mathbf{x}_0))) \\
&= Var(\mathbf{x}_0 - E(\mathbf{x}_0) - [Var(\mathbf{x}_0)]C_0^\top B(C_0(\mathbf{x}_0 - E(\mathbf{x}_0)) + \underline{\eta}_0)) \\
&= Var((I - [Var(\mathbf{x}_0)]C_0^\top BC_0)(\mathbf{x}_0 - E(\mathbf{x}_0)) - [Var(\mathbf{x}_0)]C_0^\top B\underline{\eta}_0) \\
&= (I - [Var(\mathbf{x}_0)]C_0^\top BC_0) Var(\mathbf{x}_0)\ (I - C_0^\top BC_0[Var(\mathbf{x}_0)]) \\
&\quad + [Var(\mathbf{x}_0)]C_0^\top BR_0BC_0[Var(\mathbf{x}_0)] \\
&= Var(\mathbf{x}_0) - [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] \\
&\quad - [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] \\
&\quad + [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] \\
&\quad + [Var(\mathbf{x}_0)]C_0^\top BR_0BC_0[Var(\mathbf{x}_0)] \\
&= Var(\mathbf{x}_0) - [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] \\
&\quad - [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] + [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)] \\
&= Var(\mathbf{x}_0) - [Var(\mathbf{x}_0)]C_0^\top BC_0[Var(\mathbf{x}_0)]\,.
\end{aligned}$$

5.6. From $\hat{\underline{\xi}}_0 = 0$, we have

$$\hat{\mathbf{x}}_1 = A_0\hat{\mathbf{x}}_0 + G_1(\mathbf{v}_1 - C_1 A_0 \hat{\mathbf{x}}_0)$$

and $\hat{\underline{\xi}}_1 = 0$, so that

$$\hat{\mathbf{x}}_2 = A_1\hat{\mathbf{x}}_1 + G_2(\mathbf{v}_2 - C_2 A_1 \hat{\mathbf{x}}_1)\,,$$

etc. In general, we have

$$\begin{aligned}
\hat{\mathbf{x}}_k &= A_{k-1}\hat{\mathbf{x}}_{k-1} + G_k(\mathbf{v}_k - C_k A_{k-1}\hat{\mathbf{x}}_{k-1}) \\
&= \hat{\mathbf{x}}_{k|k-1} + G_k(\mathbf{v}_k - C_k \hat{\mathbf{x}}_{k|k-1})\,.
\end{aligned}$$

Denote

$$P_{0,0} = \left[\left[Var(\mathbf{x}_0)\right]^{-1} + C_0^\top R_0^{-1} C_0\right]^{-1}$$

and

$$P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^\top + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top .$$

Then

$$\begin{aligned} G_1 &= \begin{bmatrix} A_0 & \Gamma_0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_{0,0} & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} A_0^\top & C_1^\top \\ \Gamma_0^\top & C_1^\top \end{bmatrix} \\ &\quad \cdot \left([\, C_1 A_0 \;\; C_1 \Gamma_0 \,] \begin{bmatrix} P_{0,0} & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} A_0^\top & C_1^\top \\ \Gamma_0^\top & C_1^\top \end{bmatrix} + R_1 \right)^{-1} \\ &= \begin{bmatrix} P_{1,0} C_1^\top (C_1 P_{1,0} C_1^\top + R_1)^{-1} \\ 0 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} P_1 &= \left(\begin{bmatrix} A_0 & \Gamma_0 \\ 0 & 0 \end{bmatrix} - G_1 [\, C_1 A_0 \;\; C_1 \Gamma_0 \,] \right) \begin{bmatrix} P_{0,0} & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} A_0^\top & 0 \\ \Gamma_0^\top & 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & 0 \\ 0 & Q_1 \end{bmatrix} \\ &= \begin{bmatrix} [\, I_n - P_{1,0} C_1^\top (C_1 P_{1,0} C_1^\top + R_1)^{-1} C_1 \,] P_{1,0} & 0 \\ 0 & Q_1 \end{bmatrix}, \end{aligned}$$

and, in general,

$$G_k = \begin{bmatrix} P_{k,k-1} C_k^\top (C_k P_{k,k-1} C_k^\top + R_k)^{-1} \\ 0 \end{bmatrix},$$

$$P_k = \begin{bmatrix} [\, I_n - P_{k,k-1} C_k^\top (C_k P_{k,k-1} C_k^\top + R_k)^{-1} C_k] P_{k,k-1} & 0 \\ 0 & Q_k \end{bmatrix}.$$

Finally, if we use the unbiased estimate $\hat{\mathbf{x}}_0 = E(\mathbf{x}_0)$ of $\mathbf{x}_0$ instead of the somewhat more superior initial state estimate

$$\hat{\mathbf{x}}_0 = E(\mathbf{x}_0) - [Var(\mathbf{x}_0)] C_0^\top [C_0 Var(\mathbf{x}_0) C_0^\top + R_0]^{-1} [C_0 E(\mathbf{x}_0) - \mathbf{v}_0],$$

and consequently set

$$\begin{aligned} P_0 &= E\left(\begin{bmatrix} \mathbf{x}_0 \\ \underline{\xi}_0 \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_0) \\ E(\underline{\xi}_0) \end{bmatrix} \right) \left(\begin{bmatrix} \mathbf{x}_0 \\ \underline{\xi}_0 \end{bmatrix} - \begin{bmatrix} E(\mathbf{x}_0) \\ E(\underline{\xi}_0) \end{bmatrix} \right)^\top \\ &= \begin{bmatrix} Var(\mathbf{x}_0) & 0 \\ 0 & Q_0 \end{bmatrix}, \end{aligned}$$

then we obtain the Kalman filtering algorithm derived in Chaps. 2 and 3.

5.7. Let

$$\overline{P}_0 = [\,[Var(\mathbf{x}_0)]^{-1} + C_0^\top R_0^{-1} C_0]^{-1}$$

and

$$\overline{H}_{k-1} = [\, C_k A_{k-1} - N_{k-1} C_{k-1}\,].$$

Starting with (5.17b), namely:

$$P_0 = \begin{bmatrix} (\,[Var(\mathbf{x}_0)]^{-1} + C_0 R_0^{-1} C_0)^{-1} & 0 \\ 0 & Q_0 \end{bmatrix} = \begin{bmatrix} \overline{P}_0 & 0 \\ 0 & Q_0 \end{bmatrix},$$

we have

$$\begin{aligned} G_1 &= \begin{bmatrix} A_0 & \Gamma_0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{P}_0 & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} \overline{H}_0^\top \\ \Gamma_0^\top C_1^\top \end{bmatrix} \\ &\quad \cdot \left([\, H_0 \;\; C_1\Gamma_0\,] \begin{bmatrix} \overline{P}_0 & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} \overline{H}_0^\top \\ \Gamma_0^\top C_1^\top \end{bmatrix} + R_1 \right)^{-1} \\ &= \begin{bmatrix} \left(A_0 \overline{P}_0 \overline{H}_0^\top + \Gamma_0 Q_0 \Gamma_0^\top C_1^\top\right)\left(\overline{H}_0 \overline{P}_0 \overline{H}_0^\top + C_1 \Gamma_0 Q_0 \Gamma_0^\top C_1^\top + R_1\right)^{-1} \\ 0 \end{bmatrix} \\ &:= \begin{bmatrix} \overline{G}_1 \\ 0 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} P_1 &= \left(\begin{bmatrix} A_0 & \Gamma_0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \overline{G}_1 \\ 0 \end{bmatrix} [\, \overline{H}_0 \;\; C_1\Gamma_0\,] \right) \begin{bmatrix} \overline{P}_0 & 0 \\ 0 & Q_0 \end{bmatrix} \begin{bmatrix} A_0^\top & 0 \\ \Gamma_0^\top & 0 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & 0 \\ 0 & Q_1 \end{bmatrix} \\ &= \begin{bmatrix} (A_0 - \overline{G}_1 \overline{H}_0)\overline{P}_0 A_0^\top + (I - \overline{G}_1 C_1)\Gamma_0 Q_0 \Gamma_0^\top & 0 \\ 0 & Q_1 \end{bmatrix} \\ &:= \begin{bmatrix} \overline{P}_1 & 0 \\ 0 & Q_1 \end{bmatrix}. \end{aligned}$$

In general, we obtain

$$\begin{cases} \hat{\mathbf{x}}_k = A_{k-1}\hat{\mathbf{x}}_{k-1} + \bar{G}_k(\mathbf{v}_k - N_{k-1}\mathbf{v}_{k-1} - \bar{H}_{k-1}\hat{\mathbf{x}}_{k-1}) \\ \hat{\mathbf{x}}_0 = E(\mathbf{x}_0) - [Var(\mathbf{x}_0)]C_0^\top [C_0 Var(\mathbf{x}_0) C_0^\top + R_0]^{-1}[C_0 E(\mathbf{x}_0) - \mathbf{v}_0] \\ \bar{H}_{k-1} = [\, C_k A_{k-1} - N_{k-1} C_{k-1}\,] \\ \bar{P}_k = (A_{k-1} - \bar{G}_k \bar{H}_{k-1})\bar{P}_{k-1} A_{k-1}^\top + (I - \bar{G}_k C_k)\Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top \\ \bar{G}_k = (A_{k-1}\bar{P}_{k-1}\bar{H}_{k-1}^\top + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top C_k^\top) \cdot \\ \qquad (\bar{H}_{k-1}\bar{P}_{k-1}\bar{H}_{k-1}^\top + C_k \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^\top C_k^\top + R_{k-1})^{-1} \\ \bar{P}_0 = \left[\, [Var(\mathbf{x}_0)]^{-1} + C_0^\top R_0^{-1} C_0\right]^{-1} \\ k = 1, 2, \cdots . \end{cases}$$

By omitting the "bar" on $\bar{H}_k$, $\bar{G}_k$, and $\bar{P}_k$, we have (5.21).

5.8. (a)

$$\begin{cases} \underline{X}_{k+1} = A_c \underline{X}_k + \underline{\zeta}_k \\ v_k = C_c \underline{X}_k \,. \end{cases}$$

(b)

$$P_{0,0} = \begin{bmatrix} Var(\mathbf{x}_0) & 0 & 0 \\ 0 & Var(\underline{\xi}_0) & 0 \\ 0 & 0 & Var(\eta_0) \end{bmatrix},$$

$$P_{k,k-1} = A_c P_{k-1,k-1} A_c^\top + \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & Q_{k-1} & \\ & & & & r_{k-1} \end{bmatrix},$$

$$G_k = P_{k,k-1} C_c^\top \left(C_c^\top P_{k,k-1} C_c\right)^{-1},$$

$$P_{k,k} = (I - G_k C_c) P_{k,k-1},$$

$$\hat{\underline{X}}_0 = \begin{bmatrix} E(\mathbf{x}_0) \\ 0 \\ 0 \end{bmatrix},$$

$$\hat{\underline{X}}_k = A_c \hat{\underline{X}}_{k-1} + G_k (v_k - C_c A_c \hat{\underline{X}}_{k-1}) \,.$$

(c) The matrix $C_c^\top P_{k,k-1} C_c$ may not be invertible, and the extra estimates $\hat{\underline{\xi}}_k$ and $\hat{\eta}_k$ in $\hat{\underline{X}}_k$ are needed.

Chapter 6

6.1. Since

$$\mathbf{x}_{k-1} = A\mathbf{x}_{k-2} + \Gamma \underline{\xi}_{k-2} = \cdots = A^n \mathbf{x}_{k-n-1} + noise$$

and

$$\begin{aligned} \tilde{\mathbf{x}}_{k-1} &= A^n [N_{CA}^\top N_{CA}]^{-1} \big(C^\top \mathbf{v}_{k-n-1} + A^\top C^\top \mathbf{v}_{k-n} \\ &\quad + \cdots + (A^\top)^{n-1} C^\top \mathbf{v}_{k-2}\big) \\ &= A^n [N_{CA}^\top N_{CA}]^{-1} \big(C^\top C \mathbf{x}_{k-n-1} + A^\top C^\top C A \mathbf{x}_{k-n-1} \\ &\quad + \cdots + (A^\top)^{n-1} C^\top C A^{n-1} \mathbf{x}_{k-n-1} + noise\big) \\ &= A^n [N_{CA}^\top N_{CA}]^{-1} [N_{CA}^\top N_{CA}] \mathbf{x}_{k-n-1} + noise \\ &= A^n \mathbf{x}_{k-n-1} + noise \,, \end{aligned}$$

we have $E(\tilde{\mathbf{x}}_{k-1}) = E(A^n \mathbf{x}_{k-n-1}) = E(\mathbf{x}_{k-1})$.

6.2. Since

$$\frac{d}{ds}\left[A^{-1}(s)A(s)\right] = \frac{d}{ds}\mathrm{I} = 0\,,$$

we have

$$A^{-1}(s)\left[\frac{d}{ds}A(s)\right] + \left[\frac{d}{ds}A^{-1}(s)\right]A(s) = 0\,.$$

Hence,

$$\frac{d}{ds}A^{-1}(s) = -A^{-1}(s)\left[\frac{d}{ds}A(s)\right]A^{-1}(s)\,.$$

6.3. Let $P = U diag[\ \lambda_1, \cdots, \lambda_n\ ]U^{-1}$. Then

$$P - \lambda_{min}\mathrm{I} = U diag[\ \lambda_1 - \lambda_{min}, \cdots, \lambda_n - \lambda_{min}\ ]U^{-1} \geq 0\,.$$

6.4. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of F and J be its Jordan canonical form. Then there exists a nonsingular matrix U such that

$$U^{-1}FU = J = \begin{bmatrix} \lambda_1 & * & & & \\ & \lambda_2 & * & & \\ & & \ddots & \ddots & \\ & & & \ddots & * \\ & & & & \lambda_n \end{bmatrix}$$

with each $*$ being 1 or 0. Hence,

$$F^k = UJ^kU^{-1} = U \begin{bmatrix} \lambda_1^k & * & \cdots\cdots & * \\ & \lambda_2^k & * \ \cdots & * \\ & & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n^k \end{bmatrix},$$

where each $*$ denotes a term whose magnitude is bounded by

$$p(k)|\lambda_{max}|^k$$

with $p(k)$ being a polynomial of k and $|\lambda_{max}| = max(\ |\lambda_1|, \cdots, |\lambda_n|\ )$. Since $|\lambda_{max}| < 1$, $F^k \to 0$ as $k \to \infty$.

6.5. Since

$$0 \le (A - B)\,(A - B)^\top = AA^\top - AB^\top - BA^\top + BB^\top\,,$$

we have

$$AB^\top + BA^\top \le AA^\top + BB^\top\,.$$

Hence,

$$\begin{aligned}(A + B)\,(A + B)^\top &= AA^\top + AB^\top + BA^\top + BB^\top \\ &\le 2(AA^\top + BB^\top)\,.\end{aligned}$$

6.6. Since $\mathbf{x}_{k-1} = A\mathbf{x}_{k-2} + \Gamma\underline{\xi}_{k-2}$ is a linear combination (with constant matrices as coefficients) of $\mathbf{x}_0, \underline{\xi}_0, \cdots, \underline{\xi}_{k-2}$ and

$$\begin{aligned}\vec{\mathbf{x}}_{k-1} &= A\vec{\mathbf{x}}_{k-2} + G(\mathbf{v}_{k-1} - CA\vec{\mathbf{x}}_{k-2}) \\ &= A\vec{\mathbf{x}}_{k-2} + G(CA\mathbf{x}_{k-2} + C\Gamma\underline{\xi}_{k-2} + \underline{\eta}_{k-1}) - GCA\vec{\mathbf{x}}_{k-2}\end{aligned}$$

is an analogous linear combination of $\mathbf{x}_0, \underline{\xi}_0, \cdots, \underline{\xi}_{k-2}$ and $\underline{\eta}_{k-1}$, which are uncorrelated with $\underline{\xi}_{k-1}$ and $\underline{\eta}_k$, the two identities follow immediately.

6.7. Since

$$\begin{aligned}&P_{k,k-1}C_k^\top G_k^\top - G_kC_kP_{k,k-1}C_k^\top G_k^\top \\ =&G_kC_kP_{k,k-1}C_k^\top G_k^\top + G_kR_kG_k^\top - G_kC_kP_{k,k-1}C_k^\top G_k^\top \\ =&G_kR_kG_k^\top\,,\end{aligned}$$

we have

$$-(I - G_kC)P_{k,k-1}C^\top G_k^\top + G_kRG_k^\top = 0\,.$$

Hence,

$$\begin{aligned}P_{k,k} &= (I - G_kC)P_{k,k-1} \\ &= (I - G_kC)P_{k,k-1}(I - G_kC)^\top + G_kRG_k^\top \\ &= (I - G_kC)\,(AP_{k-1,k-1}A^\top + \Gamma Q\Gamma^\top)\,(I - G_kC)^\top + G_kRG_k^\top \\ &= (I - G_kC)AP_{k-1,k-1}A^\top(I - G_kC)^\top \\ &\quad + (I - G_kC)\Gamma Q\Gamma^\top(I - G_kC)^\top + G_kRG_k^\top\,.\end{aligned}$$

6.8. Imitating the proof of Lemma 6.8 and assuming that $|\lambda| \ge 1$, where λ is an eigenvalue of $(I - GC)A$, we arrive at a contradiction to the controllability condition.

6.9. The proof is similar to that of Exercise 6.6.

6.10. From

$$\begin{aligned} 0 \ &\le \ \langle \underline{\epsilon}_j - \underline{\delta}_j, \underline{\epsilon}_j - \underline{\delta}_j \rangle \\ &= \ \langle \underline{\epsilon}_j, \underline{\epsilon}_j \rangle - \langle \underline{\epsilon}_j, \underline{\delta}_j \rangle - \langle \underline{\delta}_j, \underline{\epsilon}_j \rangle + \langle \underline{\delta}_j, \underline{\delta}_j \rangle \end{aligned}$$

and Theorem 6.2, we have

$$\begin{aligned} &\langle \underline{\epsilon}_j, \underline{\delta}_j \rangle + \langle \underline{\delta}_j, \underline{\epsilon}_j \rangle \\ \le\ &\langle \underline{\epsilon}_j, \underline{\epsilon}_j \rangle + \langle \underline{\delta}_j, \underline{\delta}_j \rangle \\ =\ &\langle \hat{\mathbf{x}}_j - \mathbf{x}_j + \mathbf{x}_j - \vec{\mathbf{x}}_j, \hat{\mathbf{x}}_j - \mathbf{x}_j + \mathbf{x}_j - \vec{\mathbf{x}}_j \rangle + \|\mathbf{x}_j - \vec{\mathbf{x}}_j\|_n^2 \\ =\ &\|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 + \langle \mathbf{x}_j - \vec{\mathbf{x}}_j, \hat{\mathbf{x}}_j - \mathbf{x}_j \rangle \\ &+ \langle \hat{\mathbf{x}}_j - \mathbf{x}_j, \mathbf{x}_j - \vec{\mathbf{x}}_j \rangle + 2\|\mathbf{x}_j - \vec{\mathbf{x}}_j\|_n^2 \\ \le\ &2\|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 + 3\|\mathbf{x}_j - \hat{\mathbf{x}}_j\|_n^2 \\ \to\ &5(P^{-1} + C^\top R^{-1} C)^{-1} \end{aligned}$$

as $j \to \infty$. Hence, $B_j = \langle \underline{\epsilon}_j, \underline{\delta}_j \rangle A^\top C^\top$ are componentwise uniformly bounded.
6.11. Using Lemmas 1.4, 1.6, 1.7 and 1.10 and Theorem 6.1, and applying Exercise 6.10, we have

$$\begin{aligned} &\mathrm{tr}[F B_{k-1-i}(G_{k-i} - G)^\top + (G_{k-i} - G) B_{k-1-i}^\top F^\top] \\ \le&(n\ \mathrm{tr} F B_{k-1-i}(G_{k-i} - G)^\top (G_{k-i} - G) B_{k-1-i}^\top F^\top)^{1/2} \\ &+ (n\ \mathrm{tr}(G_{k-i} - G) B_{k-1-i}^\top F^\top F B_{k-1-i}(G_{k-i} - G)^\top)^{1/2} \\ \le&(n\ \mathrm{tr} F F^\top \cdot \mathrm{tr}\ B_{k-1-i} B_{k-1-i}^\top \cdot \mathrm{tr}(G_{k-i} - G)^\top (G_{k-i} - G))^{1/2} \\ &+ (n\ \mathrm{tr}(G_{k-i} - G)\ (G_{k-i} - G)^\top \cdot \mathrm{tr} B_{k-1-i}^\top B_{k-1-i} \cdot \mathrm{tr} F^\top F)^{1/2} \\ =&2(n\ \mathrm{tr}(G_{k-i} - G)\ (G_{k-i} - G)^\top \cdot \mathrm{tr} B_{k-1-i}^\top B_{k-1-i} \cdot \mathrm{tr} F^\top F)^{1/2} \\ \le& C_1 r_1^{k+1-i} \end{aligned}$$

for some real number r_1, $0 < r_1 < 1$, and some positive constant C independent of i and k.
6.12. First, solving the Riccati equation (6.6); that is,

$$c^2 p^2 + [(1 - a^2) r - c^2 \gamma^2 q] p - r q \gamma^2 = 0\,,$$

we obtain

$$p = \frac{1}{2c^2}\{c^2\gamma^2 q + (a^2 - 1) r + \sqrt{[(1 - a^2) r - c^2\gamma^2 q]^2 + 4c^2\gamma^2 q r}\,\}\,.$$

Then, the Kalman gain is given by

$$g = pc/(c^2 p + r)\,.$$

Chapter 7

7.1. The proof of Lemma 7.1 is constructive. Let $A = [a_{ij}]_{n\times n}$ and $A^c = [\ell_{ij}]_{n\times n}$. It follows from $A = A^c(A^c)^\top$ that

$$a_{ii} = \sum_{k=1}^{i} \ell_{ik}^2\,, \qquad i = 1, 2, \cdots, n,$$

and

$$a_{ij} = \sum_{k=1}^{j} \ell_{ik}\ell_{jk}\,, \quad j \neq i\,; \qquad i, j = 1, 2, \cdots, n.$$

Hence, it can be easily verified that

$$\ell_{ii} = \Big(a_{ii} - \sum_{k=1}^{i-1} \ell_{ik}^2\Big)^{1/2}, \qquad i = 1, 2, \cdots, n,$$
$$\ell_{ij} = \Big(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk}\Big)/\ell_{jj}\,, \qquad j = 1, 2, \cdots, i-1; \quad i = 2, 3, \cdots, n,$$

and

$$\ell_{ij} = 0\,, \qquad j = i+1, i+2, \cdots, n; \quad i = 1, 2, \cdots, n.$$

This gives the lower triangular matrix A^c. This algorithm is called the *Cholesky decomposition*. For the general case, we can use a (standard) singular value decomposition (SVD) algorithm to find an orthogonal matrix U such that

$$U\, diag[s_1, \cdots, s_r, 0, \cdots, 0]U^\top = AA^\top\,,$$

where $1 \le r \le n$, $s_1, \cdots, s_r$ are singular values (which are positive numbers) of the non-negative definite and symmetric matrix $AA^\top$, and then set

$$\tilde{A} = U\, diag[\sqrt{s_1}, \cdots, \sqrt{s_r}, 0, \cdots, 0]\,.$$

7.2.

$$(a) \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & -2 & 1 \end{bmatrix}. \qquad (b) \quad L = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ \sqrt{2}/2 & \sqrt{2.5} & 0 \\ \sqrt{2}/2 & 1.5/\sqrt{2.5} & \sqrt{2.6} \end{bmatrix}.$$

7.3.
(a)

$$L^{-1} = \begin{bmatrix} 1/\ell_{11} & 0 & 0 \\ -\ell_{21}/\ell_{11}\ell_{22} & 1/\ell_{22} & 0 \\ -\ell_{31}/\ell_{11}\ell_{33} + \ell_{32}\ell_{21}/\ell_{11}\ell_{22}\ell_{33} & -\ell_{32}/\ell_{22}\ell_{33} & 1/\ell_{33} \end{bmatrix}.$$

(b)

$$L^{-1} = \begin{bmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & 0 \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix},$$

where

$$\begin{cases} b_{ii} = \ell_{ii}^{-1}, \qquad i = 1, 2, \cdots, n; \\ b_{ij} = -\ell_{jj}^{-1} \sum_{k=j+1}^{i} b_{ik}\ell_{kj}, \\ \quad j = i-1, i-2, \cdots, 1; \quad i = 2, 3, \cdots, n. \end{cases}$$

7.4. In the standard Kalman filtering process,

$$P_{k,k} \simeq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

which is a singular matrix. However, its "square-root" is

$$P_{k,k}^{1/2} = \begin{bmatrix} \epsilon/\sqrt{1-\epsilon^2} & 0 \\ 0 & 1 \end{bmatrix} \simeq \begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix}$$

which is a nonsingular matrix.

7.5. Analogous to Exercise 7.1, let $A = [a_{ij}]_{n \times n}$ and $A^u = [\ell_{ij}]_{n \times n}$. It follows from $A = A^u(A^u)^\top$ that

$$a_{ii} = \sum_{k=i}^{n} \ell_{ik}^2, \qquad i = 1, 2. \cdots, n,$$

and

$$a_{ij} = \sum_{k=j}^{n} \ell_{ik}\ell_{jk}, \qquad j \neq i; \quad i, j = 1, 2, \cdots, n.$$

Hence, it can be easily verified that

$$\ell_{ii} = \Big(a_{ii} - \sum_{k=i+1}^{n} \ell_{ik}^2\Big)^{1/2}, \qquad i = 1, 2, \cdots, n,$$

$$\ell_{ij} = \Big(a_{ij} - \sum_{k=j+1}^{n} \ell_{ik}\ell_{jk}\Big)/\ell_{jj},$$

$$j = i+1, \cdots, n; \quad i = 1, 2, \cdots, n.$$

and

$$\ell_{ij} = 0, \qquad j = 1, 2, \cdots, i-1; \quad i = 2, 3, \cdots, n.$$

This gives the upper-triangular matrix A^u.

7.6. The new formulation is the same as that studied in this chapter except that every lower triangular matrix with superscript c must be replaced by the corresponding upper triangular matrix with superscript u.
7.7. The new formulation is the same as that given in Sect. 7.3 except that all lower triangular matrix with superscript c must be replaced by the corresponding upper triangular matrix with superscript u.

Chapter 8

8.1. (a) Since $r^2 = x^2 + y^2$, we have

$$\dot{r} = \frac{x}{r}\dot{x} + \frac{y}{r}\dot{y}\,,$$

so that $\dot{r} = v\ sin\theta$ and

$$\ddot{r} = \dot{v}\ sin\theta + v\dot{\theta}\ cos\theta\,.$$

On the other hand, since $tan\theta = y/x$, we have $\dot{\theta}sec^2\theta = (x\dot{y} - \dot{x}y)/x^2$ or

$$\dot{\theta} = \frac{x\dot{y} - \dot{x}y}{x^2\ sec^2\theta} = \frac{x\dot{y} - \dot{x}y}{r^2} = \frac{v}{r}cos\theta\,,$$

so that

$$\ddot{r} = a\ sin\theta + \frac{v^2}{r}cos^2\theta$$

and

$$\begin{aligned}\ddot{\theta} &= \left(\frac{\dot{v}r - v\dot{r}}{r^2}\right)cos\theta - \frac{v}{r}\dot{\theta}sin\theta \\ &= \left(\frac{ar - v^2 sin\theta}{r^2}\right)cos\theta - \frac{v^2}{r^2}sin\theta cos\theta\,.\end{aligned}$$

(b)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) := \begin{bmatrix} v\ sin\theta \\ a\ sin\theta + \frac{v^2}{r}cos^2\theta \\ (ar - v^2 sin\theta)cos\theta/r^2 - v^2 sin\theta\ cos\theta/r^2 \end{bmatrix}.$$

(c)

$$\mathbf{x}_{k+1} = \begin{bmatrix} x_k[1] + hv\ sin(x_k[3]) \\ \\ x_k[2] + ha\ sin(x_k[3]) + v^2 cos^2(x_k[3])/x_k[1] \\ \\ v cos(x_k[3])/x_k[1] \\ \\ (a x_k[1] - v^2 sin(x_k[3]))cos(x_k[3])/x_k[1]^2 \\ -v^2 sin(x_k[3])cos(x_k[3])/x_k[1]^2 \end{bmatrix} + \underline{\xi}_k$$

and

$$v_k = [\ 1\ \ 0\ \ 0\ \ 0\ ]\mathbf{x}_k + \eta_k\,,$$

where $\mathbf{x}_k := [\ x_k[1]\ \ x_k[2]\ \ x_k[3]\ \ x_k[4]\ ]^\top$.
(d) Use the formulas in (8.8).
8.2. The proof is straightforward.
8.3. The proof is straightforward. It can be verified that

$$\hat{\mathbf{x}}_{k|k-1} = A_{k-1}\hat{\mathbf{x}}_{k-1} + B_{k-1}\mathbf{u}_{k-1} = \mathbf{f}_{k-1}(\hat{\mathbf{x}}_{k-1})\,.$$

8.4. Taking the variances of both sides of the modified "observation equation"

$$\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0) = C_0(\theta)\mathbf{x}_0 - C_0(\theta)E(\mathbf{x}_0) + \underline{\eta}_0\,,$$

and using the estimate $(\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0))(\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0))^\top$ for $Var(\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0))$ on the left-hand side, we have

$$\begin{aligned}&(\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0))(\mathbf{v}_0 - C_0(\theta)E(\mathbf{x}_0))^\top\\ =&C_0(\theta)Var(\mathbf{x}_0)C_0(\theta)^\top + R_0\,.\end{aligned}$$

Hence, (8.13) follows immediately.
8.5. Since

$$E(\mathbf{v}_1) = C_1(\theta)A_0(\theta)E(\mathbf{x}_0)\,,$$

taking the variances of both sides of the modified "observation equation"

$$\begin{aligned}&\mathbf{v}_1 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0)\\ =&C_1(\theta)(A_0(\theta)\mathbf{x}_0 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0) + \Gamma(\theta)\underline{\xi}_0) + \underline{\eta}_1\,,\end{aligned}$$

and using the estimate $(\mathbf{v}_1 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0))(\mathbf{v}_1 - C_1(\theta)A_0(\theta)\cdot E(\mathbf{x}_0))^\top$ for the variance $Var(\mathbf{v}_1 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0))$ on the left-hand side, we have

$$\begin{aligned}&(\mathbf{v}_1 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0))(\mathbf{v}_1 - C_1(\theta)A_0(\theta)E(\mathbf{x}_0))^\top\\ =&C_1(\theta)A_0(\theta)Var(\mathbf{x}_0)A_0^\top(\theta)C_1^\top(\theta) + C_1(\theta)\Gamma_0(\theta)Q_0\Gamma_0^\top(\theta)C_1^\top(\theta) + R_1\,.\end{aligned}$$

Then (8.14) follows immediately.
8.6. Use the formulas in (8.8) directly.
8.7. Since $\underline{\theta}$ is a constant vector, we have $S_k := Var(\underline{\theta}) = 0$, so that

$$P_{0,0} = \mathrm{Var}\begin{pmatrix}\mathbf{x}\\ \underline{\theta}\end{pmatrix} = \begin{bmatrix}\mathrm{Var}(\mathbf{x}_0) & 0\\ 0 & 0\end{bmatrix}.$$

It follows from simple algebra that

$$P_{k,k-1} = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} \quad \textit{and} \quad G_k = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

where $*$ indicates a constant block in the matrix. Hence, the last equation of (8.15) yields $\hat{\underline{\theta}}_{k|k} \equiv \hat{\underline{\theta}}_{k-1|k-1}$.

8.8.

$$\begin{cases} \begin{bmatrix} \hat{x}_0 \\ \hat{c}_0 \end{bmatrix} = \begin{bmatrix} x^0 \\ c^0 \end{bmatrix}, \qquad P_{0,0} = \begin{bmatrix} p_0 & 0 \\ 0 & s_0 \end{bmatrix} \\ \textit{For } k = 1, 2, \cdots, \\ P_{k,k-1} = P_{k-1,k-1} + \begin{bmatrix} q_{k-1} & 0 \\ 0 & s_{k-1} \end{bmatrix} \\ G_k = P_{k,k-1} \begin{bmatrix} \hat{c}_{k-1} \\ 0 \end{bmatrix} \left[[\hat{c}_{k-1} \; 0] P_{k,k-1} \begin{bmatrix} \hat{c}_{k-1} \\ 0 \end{bmatrix} + r_k \right]^{-1} \\ P_{k,k} = [I - G_k[\hat{c}_{k-1} \; 0]] P_{k,k-1} \\ \begin{bmatrix} \hat{x}_k \\ \hat{c}_k \end{bmatrix} = \begin{bmatrix} \hat{x}_{k-1} \\ \hat{c}_{k-1} \end{bmatrix} + G_k (v_k - \hat{c}_{k-1} \hat{x}_{k-1}), \end{cases}$$

where c^0 is an estimate of $\hat{c}_0$ given by (8.13); that is,

$$v_0^2 - 2v_0 x^0 c^0 + [(x^0)^2 - p_0](c^0)^2 - r_0 = 0 .$$

Chapter 9

9.1. (a) Let $\vec{\mathbf{x}}_k = [x_k \;\; \dot{x}_k]^\top$. Then

$$\begin{cases} x_k = -(\alpha + \beta - 2)x_{k-1} - (1 - \alpha)x_{k-2} + \alpha v_k + (-\alpha + \beta)v_{k-1} \\ \dot{x}_k = -(\alpha + \beta - 2)\dot{x}_{k-1} - (1 - \alpha)\dot{x}_{k-2} + \frac{\beta}{h} v_k - \frac{\beta}{h} v_{k-1} . \end{cases}$$

(b) $0 < \alpha < 1$ and $0 < \beta < \frac{\alpha^2}{1-\alpha}$.

9.2. System (9.11) follows from direct algebraic manipulation.

9.3. (a)

$$\Phi = \begin{bmatrix} 1 - \alpha & (1 - \alpha)h & (1 - \alpha)h^2/2 & -s\alpha \\ -\beta/h & 1 - \beta & h - \beta h/2 & -s\beta/h \\ -\gamma/h^2 & 1 - \gamma/h & 1 - \gamma/2 & -s\gamma/h^2 \\ -\theta & -\theta/h & -\theta h^2/2 & s(1 - \theta) \end{bmatrix}$$

(b)

$$\begin{aligned}
&det[zI - \Phi] = \\
&z^4 + [(\alpha - 3) + \beta + \gamma/2 - (\theta - 1)s]z^3 \\
&+ [(3 - 2\alpha) - \beta + \gamma/2 + (3 - \alpha - \beta - \gamma/2 - 3\theta)s]z^2 \\
&+ [(\alpha - 1) - (3 - 2\alpha - \beta + \gamma/2 - 3\theta)s]z + (1 - \alpha - \theta)s\,.
\end{aligned}$$

$$\begin{aligned}
\tilde{X}_1 &= \frac{zV(z-s)}{det[zI - \Phi]}\{\alpha z^2 + (\gamma/2 + \beta - 2\alpha)z + (\gamma/2 - \beta + \alpha)\}\,, \\
\tilde{X}_2 &= \frac{zV(z-1)(z-s)}{det[zI - \Phi]}\{\beta z - \beta + \gamma\}/h\,, \\
\tilde{X}_3 &= \frac{zV(z-1)^2(z-s)}{\det[zI - \Phi]}\gamma/h^2\,,
\end{aligned}$$

and

$$W = \frac{zV(z-1)^3}{det[zI - \Phi]}\theta\,.$$

(c) Let $\check{X}_k = [\ x_k \ \ \dot{x}_k \ \ \ddot{x}_k \ \ w_k\ ]^\top$. Then

$$\begin{aligned}
x_k &= a_1 x_{k-1} + a_2 x_{k-2} + a_3 x_{k-3} + a_4 x_{k-4} + \alpha v_k \\
&\quad + (-2\alpha - s\alpha + \beta + \gamma/2)v_{k-1} + [\alpha - \beta + \gamma/2 \\
&\quad + (2\alpha - \beta - \gamma/2)s]v_{k-2} - (\alpha - \beta + \gamma/2)s v_{k-3}\,, \\
\dot{x}_k &= a_1\dot{x}_{k-1} + a_2\dot{x}_{k-2} + a_3\dot{x}_{k-3} a_4\dot{x}_{k-4} + (\beta/h)v_k \\
&\quad - [(2+s)\beta/h - \gamma/h]v_{k-1} + [\beta/h - \gamma/h \\
&\quad + (2\beta - \gamma)s/h]v_{k-2} - [(\beta - \gamma)s/h]v_{k-3}\,, \\
\ddot{x}_k &= a_1\ddot{x}_{k-1} + a_2\ddot{x}_{k-2} + a_3\ddot{x}_{k-3} + a_4\ddot{x}_{k-4} + (\gamma/h)v_k \\
&\quad - [(2+\gamma)\gamma/h^2]v_{k-1} + (1+2s)v_{k-2} - s v_{k-3}\,, \\
w_k &= a_1 w_{k-1} + a_2 w_{k-2} + a_3 w_{k-3} + a_4 w_{k-4} \\
&\quad + (\gamma/h^2)(v_k - 3v_{k-1} + 3v_{k-2} - v_{k-3})\,,
\end{aligned}$$

with the initial conditions $x_{-1} = \dot{x}_{-1} = \ddot{x}_{-1} = w_0 = 0$, where

$$\begin{aligned}
a_1 &= -\alpha - \beta - \gamma/2 + (\theta - 1)s + 3\,, \\
a_2 &= 2\alpha + \beta - \gamma/2 + (\alpha + \beta h + \gamma/2 + 3\theta - 3)s - 3\,, \\
a_3 &= -\alpha + (-2\alpha - \beta + \gamma/2 - 3\theta + 3)s + 1\,,
\end{aligned}$$

and

$$a_4 = (\alpha + \theta - 1)s\,.$$

(d) The verification is straightforward.
9.4. The verifications are tedious but elementary.
9.5. Study (9.19) and (9.20). We must have $\sigma_p, \sigma_v, \sigma_a \geq 0,\ \sigma_m > 0$, and $P > 0$.
9.6. The equations can be obtained by elementary algebraic manipulation.
9.7. Only algebraic manipulation is required.

Chapter 10

10.1. For (1) and (4), let $* \in \{+, -, \cdot, /\}$. Then

$$\begin{aligned} X * Y &= \left\{x * y | x \in X, y \in Y\right\} \\ &= \left\{y * x | y \in Y, x \in X\right\} \\ &= Y * X\,. \end{aligned}$$

The others can be verified in a similar manner. As to part (c) of (7), without loss of generality, we may only consider the situation where both $x \geq 0$ and $y \geq 0$ in $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$, and then discuss different cases of $\underline{z} \geq 0$, $\overline{z} \leq 0$, and $\underline{z}\overline{z} < 0$.
10.2. It is straightforward to verify all the formulas by definition. For instance, for part (j.1), we have

$$\begin{aligned} A^I(BC) &= \left[\sum_{j=1}^{n} A^I(i,j)\left[\sum_{\ell=1}^{n} B_{j\ell}C_{\ell k}\right]\right] \\ &\subseteq \left[\sum_{j=1}^{n}\sum_{\ell=1}^{n} A^I(i,j)B_{j\ell}C_{\ell k}\right] \\ &= \left[\sum_{\ell=1}^{n}\left[\sum_{j=1}^{n} A^I(i,j)B_{j\ell}\right]C_{\ell k}\right] \\ &= (A^I B)C\,. \end{aligned}$$

10.3. See: Alefeld, G. and Herzberger, J. (1983).
10.4. Similar to Exercise 1.10.
10.5. Observe that the filtering results for a boundary system and any of its neighboring system will be inter-crossing from time to time.
10.6. See: Siouris, G., Chen, G. and Wang, J. (1997).

Chapter 11

11.1.

$$\phi_2(t) = \begin{cases} \frac{1}{2}t^2 & 0 \le t < 1 \\ -t^2 + 3t - \frac{3}{2} & 1 \le t < 2 \\ \frac{1}{2}t^2 - 3t + \frac{9}{2} & 2 \le t < 3 \\ 0 & \textit{otherwise.} \end{cases}$$

$$\phi_3(t) = \begin{cases} \frac{1}{6}t^3 & 0 \le t < 1 \\ -\frac{1}{2}t^3 + 2t^2 - 2t + \frac{2}{3} & 1 \le t < 2 \\ \frac{1}{2}t^3 - 4t^2 + 10t - \frac{22}{3} & 2 \le t < 3 \\ -\frac{1}{6}t^3 + 2t^2 - 8t + \frac{32}{3} & 3 \le t < 4 \\ 0 & \textit{otherwise.} \end{cases}$$

11.2.

$$\widehat{\phi_n}(\omega) = \left(\frac{1 - e^{-i\omega}}{i\omega}\right)^n = e^{-in\omega/2}\left(\frac{\sin(\omega/2)}{\omega/2}\right)^n.$$

11.3. Simple graphs.
11.4. Straightforward algebraic operations.
11.5. Straightforward algebraic operations.

Chapter 12

12.1 Apply matrix analysis (see Cattivelli and Sayed 2010).
12.2 Straightforward algebraic operations.

Index

C.K. Chui and G. Chen, *Kalman Filtering*, DOI 10.1007/978-3-319-47612-4