1, and 10 s, and simulated vertical gust velocity, $x(t)$, of the filtered white noise with $T = 0.1$, 1, and 10 s. Note that the maximum value of $S(\omega)$ is $1.125a^2$ which occurs near $\omega T = 1$. Note from the Bode plots of the filter with various T values, that the white noise filter for Dryden turbulence acts as a *low-pass filter* with a cut-off frequency, $1/T$. Consequently, as T increases the filtered output (representing turbulence) becomes smoother, because more high-frequency noise is blocked by the filter.

7.4 The Kalman Filter

In the previous section we saw how we can represent stochastic systems by passing white noise through linear systems. Such a representation of stochastic systems is useful for dealing with a plant which we cannot model accurately using only a deterministic model, because of the presence of modeling uncertainties (called *process noise*) and *measurement noise*. A noisy plant is thus a stochastic system, which is modeled by passing white noise through an appropriate linear system. Consider such a plant with the following linear, time-varying state-space representation:

$$\mathbf{x}^{(1)}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{F}(t)\mathbf{v}(t) \tag{7.40}$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) + \mathbf{z}(t) \tag{7.41}$$

where $\mathbf{v}(t)$ is the *process noise vector* which may arise due to modeling errors such as neglecting nonlinear or higher-frequency dynamics, and $\mathbf{z}(t)$ is the *measurement noise vector*. By assuming $\mathbf{v}(t)$ and $\mathbf{z}(t)$ to be *white noises*, we will only be extending the methodology of the previous section for a description of the stochastic plant. However, since now we are dealing with a time-varying system as the plant, our definition of white noise has to be modified. For a time-varying stochastic system, the output is a *non-stationary* random signal. Hence, the random noises, $\mathbf{v}(t)$ and $\mathbf{z}(t)$, could in general be *non-stationary white noises*. A *non-stationary white noise* can be obtained by passing the stationary white noise through an *amplifier* with a time-varying gain. The *correlation matrices* of non-stationary white noises, $\mathbf{v}(t)$ and $\mathbf{z}(t)$, can be expressed as follows:

$$\mathbf{R}_{v}(t, \tau) = \mathbf{V}(t)\delta(t-\tau) \tag{7.42}$$

$$\mathbf{R}_{z}(t, \tau) = \mathbf{Z}(t)\delta(t-\tau) \tag{7.43}$$

where $\mathbf{V}(t)$ and $\mathbf{Z}(t)$ are the time-varying power spectral density matrices of $\mathbf{v}(t)$ and $\mathbf{z}(t)$, respectively. Note that Eqs. (7.42) and (7.43) yield *infinite* covariance matrices, $\mathbf{R}_{v}(t, t)$ and $\mathbf{R}_{z}(t, t)$, respectively, which can be regarded as a characteristic of white noise – stationary or non-stationary.

For designing a control system based on the stochastic plant, we cannot rely on full-state feedback, because we cannot predict the state-vector, $\mathbf{x}(t)$, of the stochastic plant. Therefore, an *observer* is required for estimating the state-vector, based upon a measurement of the output, $\mathbf{y}(t)$, given by Eq. (7.41) and a known input, $\mathbf{u}(t)$. Using the pole-placement methods of Chapter 5 we can come up with an observer that has poles at desired locations. However, such an observer would not take into account the power spectra of the process and measurement noise. Also, note the difficulty encountered in Chapter 5 for designing observers for multi-input, multi-output plants, which limits the pole-placement approach of observer design largely to single-output plants. To take into account the fact that the measured output, $\mathbf{y}(t)$, and state-vector of the plant, $\mathbf{x}(t)$, are random vectors, we need an observer that estimates the state-vector based upon *statistical* (rather than deterministic) description of the *vector* output and plant state. Such an observer is the *Kalman Filter*. Rather than being an ordinary observer of Chapter 5, the *Kalman filter* is an *optimal* observer, which *minimizes* a statistical measure of the *estimation error*, $\mathbf{e_o}(t) = \mathbf{x}(t) - \mathbf{x_o}(t)$, where $\mathbf{x_o}(t)$ is the *estimated state-vector*. The state-equation of the Kalman filter is that of a time-varying observer (similar to Eq. (5.103)), and can be written as follows:

$$\mathbf{x_o}^{(1)}(t) = \mathbf{A}(t)\mathbf{x_o}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{L}(t)[\mathbf{y}(t) - \mathbf{C}(t)\mathbf{x_o}(t) - \mathbf{D}(t)\mathbf{u}(t)] \tag{7.44}$$

where $\mathbf{L}(t)$ is the *gain matrix* of the Kalman filter (also called the *optimal observer gain matrix*). Being an *optimal* observer, the Kalman filter is a counterpart of the optimal regulator of Chapter 6. However, while the optimal regulator minimizes an objective function based on transient and steady-state response and control effort, the Kalman filter minimizes the *covariance of the estimation error*, $\mathbf{R_e}(t,t) = E[\mathbf{e_o}(t)\mathbf{e_o}^{\mathbf{T}}(t)]$. Why is it useful to minimize the covariance of estimation error? Recall that the state-vector, $\mathbf{x}(t)$, is a random vector. The estimated state, $\mathbf{x_o}(t)$, is based on the measurement of the output, $\mathbf{y}(t)$, for a *finite* time, say T, where $t \geq T$. However, a true statistical average (or *mean*) of $\mathbf{x}(t)$ would require measuring the output for an infinite time (i.e. taking *infinite* number of samples), and then finding the expected value of $\mathbf{x}(t)$. Hence, the *best estimate* that the Kalman filter could obtain for $\mathbf{x}(t)$ is not the *true mean*, but a *conditional mean*, $\mathbf{x_m}(t)$, based on only a *finite* time record of the output, $\mathbf{y}(t)$, for $T \leq t$, written as follows:

$$\mathbf{x_m}(t) = E[\mathbf{x}(t) : \mathbf{y}(T), T \leq t] \tag{7.45}$$

There may be a *deviation* of the estimated state-vector, $\mathbf{x_o}(t)$, from the conditional mean, $\mathbf{x_m}(t)$, and we can write the estimated state-vector as follows:

$$\mathbf{x_o}(t) = \mathbf{x_m}(t) + \Delta\mathbf{x}(t) \tag{7.46}$$

where $\Delta\mathbf{x}(t)$ is the deviation from the conditional mean. The *conditional covariance matrix* (i.e. the covariance matrix based on a finite record of the output) of the estimation error is given by

$$\mathbf{R_e}(t,t) = E[\mathbf{e_o}(t)\mathbf{e_o}^{\mathbf{T}}(t) : \mathbf{y}(T), T \leq t] = E[\{\mathbf{x}(t) - \mathbf{x_o}(t)\}\{\mathbf{x}^{\mathbf{T}}(t) - \mathbf{x_o}^{\mathbf{T}}(t)\} : \mathbf{y}(T), T \leq t] \tag{7.47}$$

Equation (7.47) can be simplified using Eq. (7.45) as follows:

$$\mathbf{R_e}(t,t) = E[\mathbf{x}(t)\mathbf{x}^{\mathbf{T}}(t)] - \mathbf{x_o}(t)\mathbf{x_m}^{\mathbf{T}}(t) - \mathbf{x_o}^{\mathbf{T}}(t)\mathbf{x_m}(t) + \mathbf{x_o}(t)\mathbf{x_o}^{\mathbf{T}}(t) \tag{7.48}$$

Finally, substituting Eq. (7.46) into Eq. (7.48) and simplifying, we get

$$\mathbf{R_e}(t,t) = E[\mathbf{x}(t)\mathbf{x}^{\mathbf{T}}(t)] - \mathbf{x_m}(t)\mathbf{x_m}^{\mathbf{T}}(t) + \Delta\mathbf{x}(t)\Delta\mathbf{x}^{\mathbf{T}}(t) \tag{7.49}$$

From Eq. (7.49) it is clear that the best estimate of state-vector, implying $\Delta\mathbf{x}(t) = 0$ (i.e. $\mathbf{x_o}(t) = \mathbf{x_m}(t)$), would result in a *minimization* of the conditional covariance matrix, $\mathbf{R_e}(t, t)$. In other words, minimization of $\mathbf{R_e}(t, t)$ yields the optimal (i.e. the best) observer, which is the Kalman filter.

Let us derive the expression for the gain matrix of the Kalman filter, $\mathbf{L}(t)$, which minimizes $\mathbf{R_e}(t, t)$, i.e. which makes the estimated state-vector equal to the conditional mean vector. The *optimal* estimation error is thus $\mathbf{e_o}(t) = \mathbf{x}(t) - \mathbf{x_m}(t)$. Subtracting Eq. (7.44) from Eq. (7.40) and substituting Eq. (7.41), we can write the following state-equation for the optimal estimation error:

$$\mathbf{e_o}^{(1)}(t) = [\mathbf{A}(t) - \mathbf{L}(t)\mathbf{C}(t)]\mathbf{e_o}(t) + \mathbf{F}(t)\mathbf{v}(t) - \mathbf{L}(t)\mathbf{z}(t) \tag{7.50}$$

Note that since $\mathbf{v}(t)$ and $\mathbf{z}(t)$ are white noises, the following vector is also a (non-stationary) white noise

$$\mathbf{w}(t) = \mathbf{F}(t)\mathbf{v}(t) - \mathbf{L}(t)\mathbf{z}(t) \tag{7.51}$$

To find the covariance of the estimation error, we must somehow find an expression for the solution of Eq. (7.50), which is a linear, time-varying system excited by a non-stationary white noise, $\mathbf{w}(t)$. Let us write Eqs. (7.50) and (7.51) as follows:

$$\mathbf{e_o}^{(1)}(t) = \mathbf{A_o}(t)\mathbf{e_o}(t) + \mathbf{w}(t) \tag{7.52}$$

where $\mathbf{A_o}(t) = [\mathbf{A}(t) - \mathbf{L}(t)\mathbf{C}(t)]$. The solution to Eq. (7.52) for a given initial condition, $\mathbf{e_o}(t_0)$, can be expressed as follows:

$$\mathbf{e_o}(t) = \Phi(t, t_0)\mathbf{e_o}(t_0) + \int_{t_0}^{t} \Phi(t, \lambda)\mathbf{w}(\lambda)\, d\lambda \tag{7.53}$$

where $\Phi(t, t_0)$ is the *state-transition matrix* of the time-varying estimation error state-equation, Eq. (7.52). Then the conditional covariance of estimation error can be written as follows (dropping the notation $\mathbf{y}(T)$, $T \le t$, from the expected value for convenience):

$$\begin{aligned}\mathbf{R_e}(t, t) = E[\mathbf{e_o}(t)\mathbf{e_o^T}(t)] = E\Big[&\Phi(t, t_0)\mathbf{e_o}(t_0)\mathbf{e_o^T}(t_0)\Phi^{\mathbf{T}}(t, t_0) \\ &+\mathbf{e_o}(t_0)\int_{t_0}^{t} \Phi(t, \lambda)\mathbf{w}(\lambda)\, d\lambda + \left\{\int_{t_0}^{t} \Phi(t, \lambda)\mathbf{w}(\lambda)\, d\lambda\right\}\mathbf{e_o}(t_0) \\ &+ \int_{t_0}^{t}\int_{t_0}^{t} \Phi(t, \lambda)\mathbf{w}(\lambda)\mathbf{w^T}(\xi)\Phi^{\mathbf{T}}(t, \xi)\, d\lambda d\xi\Big]\end{aligned} \tag{7.54}$$

or, using the properties of the expected value operator, we can write

$$\begin{aligned}\mathbf{R_e}(t, t) = &\Phi(t, t_0)E[\mathbf{e_o}(t_0)\mathbf{e_o^T}(t_0)]\Phi^{\mathbf{T}}(t, t_0) + \mathbf{e_o}(t_0)\int_{t_0}^{t} \Phi(t, \lambda)E[\mathbf{w}(\lambda)]\, d\lambda \\ &+ \left\{\int_{t_0}^{t} \Phi(t, \lambda)E[\mathbf{w}(\lambda)]\, d\lambda\right\}\mathbf{e_o}(t_0) \\ &+ \int_{t_0}^{t}\int_{t_0}^{t} \Phi(t, \lambda)E[\mathbf{w}(\lambda)\mathbf{w^T}(\xi)]\Phi^{\mathbf{T}}(t, \xi)\, d\lambda\, d\xi\end{aligned} \tag{7.55}$$

Since the expected value of white noise is zero, i.e. $E[\mathbf{w}(t)] = \mathbf{0}$, and the *correlation matrix* of white noise is given by

$$E[\mathbf{w}(\lambda)\mathbf{w}^{\mathbf{T}}(\xi)] = \mathbf{W}(\lambda)\delta(\lambda - \xi) \tag{7.56}$$

we can simplify Eq. (7.55) as follows:

$$\mathbf{R_e}(t, t) = \boldsymbol{\Phi}(t, t_0)E[\mathbf{e_o}(t_0)\mathbf{e_o^T}(t_0)]\boldsymbol{\Phi}^{\mathbf{T}}(t, t_0) + \int_{t_0}^{t} \boldsymbol{\Phi}(t, \lambda)\mathbf{W}(\lambda)\boldsymbol{\Phi}^{\mathbf{T}}(t, \lambda)\, d\lambda \tag{7.57}$$

If the initial estimation error, $\mathbf{e_o}(t_0)$, is also a random vector, we can write the initial conditional covariance matrix as follows:

$$\mathbf{R_e}(t_0, t_0) = E[\mathbf{e_o}(t_0)\mathbf{e_o^T}(t_0)] \tag{7.58}$$

Substituting Eq. (7.58) into Eq. (7.57), we can write

$$\mathbf{R_e}(t, t) = \boldsymbol{\Phi}(t, t_0)\mathbf{R_e}(t_0, t_0)\boldsymbol{\Phi}^{\mathbf{T}}(t, t_0) + \int_{t_0}^{t} \boldsymbol{\Phi}(t, \lambda)\mathbf{W}(\lambda)\boldsymbol{\Phi}^{\mathbf{T}}(t, \lambda)\, d\lambda \tag{7.59}$$

Equation (7.59) describes how the optimal estimation error covariance evolves with time. However, the state-transition matrix for the time varying system, $\boldsymbol{\Phi}(t, t_0)$, is an *unknown quantity*. Fortunately, we have already encountered an integral similar to that in Eq. (7.59) while deriving the *optimal regulator gain* in Chapter 6. An equivalent integral for the *optimal control problem* is given in Eq. (6.8) for $\mathbf{M}(t, t_f)$. Comparing Eqs. (7.59) and (6.8) we find that $\boldsymbol{\Phi}_{\mathbf{CL}}^{\mathbf{T}}(\tau, t)$ in Eq. (6.8) is replaced by $\boldsymbol{\Phi}(t, \lambda)$ in Eq. (7.59), where τ and λ are the variables of integration in Eqs. (6.8) and (7.59), respectively. Furthermore, the matrix $[\mathbf{Q}(\tau) + \mathbf{K}^{\mathbf{T}}(\tau)\mathbf{R}(\tau)\mathbf{K}(\tau)]$ in Eq. (6.8) is replaced by $\mathbf{W}(\lambda)$ in Eq. (7.59). Also, the direction of integration in time is *opposite* in Eqs. (6.8) ($t \to t_f$) and (7.59) ($t_0 \to t$). Thus, taking a cue from the similarity (and differences) between Eqs. (6.8) and (7.59), we can write a *differential equation* for $\mathbf{R_e}(t, t)$ similar to that for $\mathbf{M}(t, t_f)$, Eq. (6.14), as follows:

$$d\mathbf{R_e}(t, t)/dt = \mathbf{A_o}(t)\mathbf{R_e}(t, t) + \mathbf{R_e}(t, t)\mathbf{A_o^T}(t) + \mathbf{W}(t) \tag{7.60}$$

Note that Eq. (7.60) is an *ordinary differential equation*, rather than a *partial differential equation* (Eq. (6.14)). Also, due to the fact that time progresses in a *forward direction* ($t_0 \to t$) in Eq. (7.60), rather than in a *backward direction* ($t_f \to t$) in Eq. (6.14), the negative sign on the left-hand side of Eq. (6.14) is replaced by a positive sign on the left-hand side of Eq. (7.60). Equation (7.60) is called the *covariance equation* for the Kalman filter, and must be solved with the initial condition given by Eq. (7.58). Note that we *do not* need to know the state-transition matrix, $\boldsymbol{\Phi}(t, t_0)$, for solving for the optimal covariance matrix. Equation (7.60) is the counterpart of the matrix Riccati equation for the Kalman filter.

Substituting Eqs. (7.51), (7.42) and (7.43) into Eq. (7.56), and *assuming* that the two white noise signals, $\mathbf{v}(t)$ and $\mathbf{z}(t)$, are *uncorrelated with each other*, i.e. $E[\mathbf{v}(t)\mathbf{z}^{\mathbf{T}}(\tau)] = E[\mathbf{z}(t)\mathbf{v}^{\mathbf{T}}(\tau)] = \mathbf{0}$, we can write the following expression relating $\mathbf{W}(t)$ to the spectral densities of the two white noise signals, $\mathbf{V}(t)$ and $\mathbf{Z}(t)$, as follows:

$$\mathbf{W}(t) = \mathbf{F}(t)\mathbf{V}(t)\mathbf{F}^{\mathbf{T}}(t) + \mathbf{L}(t)\mathbf{Z}(t)\mathbf{L}^{\mathbf{T}}(t) \tag{7.61}$$

Substituting Eq. (7.61) into Eq. (7.60) and substituting $\mathbf{A_o}(t) = [\mathbf{A}(t) - \mathbf{L}(t)\mathbf{C}(t)]$, we can express the covariance equation as follows:

$$\begin{aligned} d\mathbf{R_e}(t,t)/dt = {} & [\mathbf{A}(t) - \mathbf{L}(t)\mathbf{C}(t)]\mathbf{R_e}(t,t) + \mathbf{R_e}(t,t))[\mathbf{A}(t) - \mathbf{L}(t)\mathbf{C}(t)]^{\mathbf{T}} \\ & + \mathbf{F}(t)\mathbf{V}(t)\mathbf{F}^{\mathbf{T}}(t) + \mathbf{L}(t)\mathbf{Z}(t)\mathbf{L}^{\mathbf{T}}(t) \end{aligned} \tag{7.62}$$

Comparing Eq. (7.62) with Eq. (6.21) and using the steps similar to those of Section 6.1.2, we can write the *optimal Kalman filter gain*, $\mathbf{L^o}(t)$, as follows:

$$\mathbf{L^o}(t) = \mathbf{R_e^o}(t,t)\mathbf{C}^{\mathbf{T}}(t)\mathbf{Z}^{-1}(t) \tag{7.63}$$

where $\mathbf{R_e^o}(t,t)$ is the *optimal covariance matrix* satisfying the following *matrix Riccati equation*:

$$\begin{aligned} d\mathbf{R_e^o}(t,t)/dt = {} & \mathbf{A}(t)\mathbf{R_e^o}(t,t) + \mathbf{R_e^o}(t,t)\mathbf{A}^{\mathbf{T}}(t) \\ & - \mathbf{R_e^o}(t,t)\mathbf{C}^{\mathbf{T}}(t)\mathbf{Z}^{-1}(t)\mathbf{C}(t)\mathbf{R_e^o}(t,t) + \mathbf{F}(t)\mathbf{V}(t)\mathbf{F}^{\mathbf{T}}(t) \end{aligned} \tag{7.64}$$

We can derive a more general matrix Riccati equation for the Kalman filter when the two noise signals are correlated with each other with the following *cross-correlation matrix*:

$$E[\mathbf{v}(t)\mathbf{z}^{\mathbf{T}}(\tau)] = \Psi(t)\delta(t-\tau) \tag{7.65}$$

where $\Psi(t)$ is the *cross-spectral density matrix* between $\mathbf{v}(t)$ and $\mathbf{z}(t)$. Then the optimal Kalman filter gain can be shown to be given by

$$\mathbf{L^o}(t) = [\mathbf{R_e^o}(t,t)\mathbf{C}^{\mathbf{T}}(t) + \mathbf{F}(t)\Psi(t)]\mathbf{Z}^{-1}(t) \tag{7.66}$$

where $\mathbf{R_e^o}(t,t)$ is the *optimal covariance matrix* satisfying the following general *matrix Riccati equation*:

$$\begin{aligned} d\mathbf{R_e^o}(t,t)/dt = {} & \mathbf{A_G}(t)\mathbf{R_e^o}(t,t) + \mathbf{R_e^o}(t,t)\mathbf{A_G^T}(t) \\ & - \mathbf{R_e^o}(t,t)\mathbf{C}^{\mathbf{T}}(t)\mathbf{Z}^{-1}(t)\mathbf{C}(t)\mathbf{R_e^o}(t,t) + \mathbf{F}(t)\mathbf{V_G}(t)\mathbf{F}^{\mathbf{T}}(t) \end{aligned} \tag{7.67}$$

with

$$\mathbf{A_G}(t) = \mathbf{A}(t) - \mathbf{F}(t)\Psi(t)\mathbf{Z}^{-1}(t)\mathbf{C}(t) \tag{7.68}$$

$$\mathbf{V_G}(t) = \mathbf{V}(t) - \Psi(t)\mathbf{Z}^{-1}(t)\Psi^{\mathbf{T}}(t) \tag{7.69}$$

For simplicity of notation, we will use $\mathbf{L}$ to denote the optimal gain matrix of the Kalman filter, rather than $\mathbf{L^o}(t)$.

The appearance of matrix Riccati equation for the Kalman filter problem is not surprising, since the Kalman filter is an optimal observer. Hence, Kalman filter problem is solved quite similarly to the optimal control problem. Usually, we are interested in a *steady Kalman filter*, i.e. the Kalman filter for which the covariance matrix converges to a constant in the limit $t \to \infty$. Such a Kalman filter results naturally when the plant is *time-invariant* and the noise signals are *stationary white noises*. In such a case, the estimation error would also be a stationary white noise with a *constant optimal covariance matrix*, $\mathbf{R}_e^o$. For the steady-state time-varying problem with non-stationary white noise signals, or the time-invariant problem with stationary white noise signals, the following *algebraic Riccati equation* results for the optimal covariance matrix, $\mathbf{R}_e^o$:

$$\mathbf{0} = \mathbf{A}_G\mathbf{R}_e^o + \mathbf{R}_e^o\mathbf{A}_G^T - \mathbf{R}_e^o\mathbf{C}^T\mathbf{Z}^{-1}\mathbf{C}\mathbf{R}_e^o + \mathbf{F}\mathbf{V}_G\mathbf{F}^T \tag{7.70}$$

where the matrices on the right-hand side are either constant (time-invariant), or steady-state values for the time-varying plant. The sufficient conditions for the existence of a *unique, positive definite* solution to the algebraic Riccati equation (Eq. (7.70)) are the same as those stated in Chapter 6: the system with state-dynamics matrix, **A**, and observation matrix, **C**, is *detectable*, and the system with state-dynamics matrix, **A**, and controls coefficient matrix, $\mathbf{B} = \mathbf{F}\mathbf{V}^{1/2}$, is *stabilizable* ($\mathbf{V}^{1/2}$ denotes the *matrix square-root* of **V**, which satisfies $\mathbf{V}^{1/2}(\mathbf{V}^{1/2})^T = \mathbf{V}$). These sufficient conditions will be met if the system with state-dynamics matrix, **A**, and observation matrix, **C**, is *observable*, **V** is a *positive semi-definite* matrix, and **Z** is a *positive definite matrix*. We can solve the algebraic Riccati equation for steady-state Kalman filter using either the MATLAB command *are*, or more specifically, the specialized Kalman filter commands *lqe* or *lqe2*, which are called as follows:

```
>>[L,P,E] = lqe(A,F,C,V,Z,Psi) <enter>
```

where **A**, **F**, **C**, are the plant's state coefficient matrices, **V** is the process noise spectral density matrix, **Z** is the measurement noise spectral density matrix, and **Psi** $= \Psi$, the cross-spectral density matrix of process and measurement noises. If **Psi** is not specified (by having only the first four input arguments in *lqe*), it is assumed that $\Psi = \mathbf{0}$. **L** is the returned Kalman filter optimal gain, $\mathbf{P} = \mathbf{R}_e^o$, the returned optimal (conditional) covariance matrix of the estimation error, and **E** is a vector containing the eigenvalues of the Kalman filter (i.e. the eigenvalues of **A-LC**). The command *lqe2*, is used in a manner similar to *lqe*, but utilizes a more numerically robust algorithm for solving the algebraic Riccati equation than *lqe*. A third MATLAB command, *lqew*, is also available, which solves a special Kalman filter problem in which the output equation is $\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) + \mathbf{G}(t)\mathbf{v}(t) + \mathbf{z}(t)$, where $\mathbf{v}(t)$ is the process noise that affects the output through the coefficient matrix, $\mathbf{G}(t)$, and is *uncorrelated* with the measurement noise, $\mathbf{z}(t)$.

Example 7.6

Let us design a Kalman filter for the fighter aircraft of Example 5.13. It is assumed that only the first two of the three state variables are measured. The state coefficient matrices for the linear, time-invariant model are as follows:

$$\mathbf{A} = \begin{bmatrix} -1.7 & 50 & 260 \\ 0.22 & -1.4 & -32 \\ 0 & 0 & -12 \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} -272 \\ 0 \\ 14 \end{bmatrix} 0; \quad \mathbf{F} = \begin{bmatrix} 0.02 & 0.1 \\ -0.0035 & 0.004 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}; \quad \mathbf{D} = 0 \tag{7.71}$$

The third-order plant has a single input, two outputs and two process noise variables. The process noise spectral density matrix for the bomber is assumed to be $\mathbf{V} = \mathbf{F}^\mathbf{T}\mathbf{F}$, while that for the measurement noise is, $\mathbf{Z} = 10\mathbf{C}\mathbf{C}^\mathbf{T}$. Also, assume that the cross-spectral density of process and measurement noise is zero, i.e. $\Psi = \mathbf{0}$. Then the Kalman filter gain is calculated using the MATLAB command *lqe* as follows:

```
>>A=[-1.7 50 260; 0.22 -1.4 -32;0 0 -12];F=[0.02 0.1;-0.0035 0.004;0 0];
  C=[1 0 0; 0 1 0]; <enter>

>>[L,P,E] = lqe(A,F,C,F'*F,10*C*C') <enter>
L =
        3.5231          0.2445
        0.2445          0.0170
        0               0

P =
        35.2306         2.4450          0
        2.4450          0.1697          0
        0               0               0

E =
 -12.0000
 -4.8700
 -1.7700
```

The Kalman filter is thus *stable* with eigenvalues at $\lambda = -12$, $\lambda = -4.87$, and $\lambda = -1.77$. Let us simulate the Kalman filter estimation error with $\mathbf{v}(t)$ and $\mathbf{z}(t)$ generated using *randn* as follows:

```
>>randn('seed',0); t=0:0.01:10; v = randn(size(t,2),2); z = randn(size
  (t,2),2); w = F*v'-L*z'; <enter>

>>[e,X] = lsim(A-L*C,eye(3),eye(3),zeros(3,3),w',t); <enter>
```

The simulated white noise, representing an element of $\mathbf{v}(t)$ or $\mathbf{z}(t)$, and the simulated estimation error, $\mathbf{e_o}(t) = [e_{01}(t)e_{02}(t)e_{03}(t)]^T$, are plotted in Figure 7.7.

Note that the third element of the estimation error vector, $\mathbf{e_o}(t)$, is identically zero (because the last row of $\mathbf{F}$ is zero), while the first two elements are much smoother and with smaller magnitudes than the noise. The first two elements of the estimation error are random variables with zero mean values. How accurate is our simulated estimation error? Not very accurate, because the noise vector we have simulated does not have *exactly* the same power spectral density that we have assumed. In fact, the *simulated* white noise is far from being a perfect white noise, which we can verify by calculating the covariance matrices and mean values of $\mathbf{v}(t)$ and $\mathbf{z}(t)$, as follows:

```
>>cov(v) <enter>
ans =
        1.0621 -0.0307
        -0.0307         1.0145
```

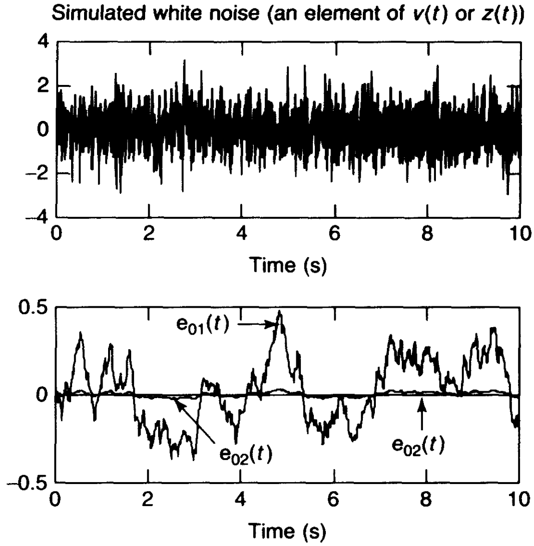

Figure 7.7 Simulated white noise and estimation errors of the Kalman filter for a fighter aircraft

```
>>cov(z) <enter>
ans =
      1.0229        0.0527 -0.0133
      0.0527        0.8992         0.0131
      -0.0133       0.0131         0.9467

>>mean(v) <enter>
ans =
 -0.0139 -0.0047

>>mean(z) <enter>
ans =
 -0.0125 0.0025 -0.0421
```

Recall that a white noise vector must have *infinite* covariance matrix and a *zero* mean vector. The covariance of the *simulated* estimation error is the following:

```
>>cov(e) <enter>
ans =
      0.0347        0.0024         0
      0.0024        0.0002         0
      0             0              0
```

Note that the simulated estimation error's covariance is different from the optimal covariance matrix, **P**, obtained as the solution to the algebraic Riccati equation of the Kalman filter. However, the ratios between the elements of **P** are the same as those between the elements of *cov(e)*. Hence, the matrix P must be scaled by a scalar constant to represent the covariance of estimation error. To reduce the difference between the covariance matrices of the simulated and optimal estimation error, we should modify our assumptions of the power spectral densities, or simulate the white noise more accurately using the random number generator, *randn*. Since the actual noise will almost *never* be a white noise, there is no point in spending the time to accurately model white noise better on a computer. Instead, we should *fine tune* the Kalman filter gain by appropriately selecting the spectral densities $\mathbf{V}$, $\mathbf{Z}$, and Ψ. After some trial and error, by selecting $\mathbf{V} = \mathbf{F}^T\mathbf{F}$, $\mathbf{Z} = 0.01\mathbf{C}\mathbf{C}^T$, and $\Psi = \mathbf{0}$, we get the Kalman filter gain and optimal estimation error covariance as follows:

```
>>[L,P,E] = lqe(A,F,C,F'*F,0.01*C*C') <enter>

L =
        3.5247          0.2446
        0.2446          0.0170
        0               0

P =
        0.0352          0.0024          0
        0.0024          0.0002          0
        0               0               0

E =
        -12.0000
        -4.8708
        -1.7709
```

and the simulated estimation error is re-calculated as

```
>>w = F*v'-L*z';sysob=ss(A-L*C,eye(3),eye(3),zeros(3,3));[e,t,X] = lsim
  (sysob,w',t); <enter>
```

with the simulated estimation error's covariance matrix given by

```
>>cov(e) <enter>
ans =
          0.0347          0.0024          0
          0.0024          0.0002          0
          0               0               0
```

which is the same as calculated previously, and quite close to the new optimal covariance, $\mathbf{P}$. The new Kalman filter gain and eigenvalues are not changed by much (indicating little change in the estimation error time response), but the scaling of the optimal covariance of the estimation error is now greatly improved. The *mean* estimation error vector, $\mathbf{e}_{om}$, is calculated as follows:

```
>> mean(e) <enter>

ans =
          0.0305          0.0021          0
```

which is quite close to zero vector, as desired. The accuracy of the mean value will, of course, improve by taking more time points in the simulation, and in the limit of *infinite* number of time points, the mean would become *exactly* zero.

The Kalman filter approach provides us with a procedure for designing observers for multivariable plants. Such an observer is guaranteed to be optimal in the presence of white noise signals. However, since white noise is rarely encountered, the power spectral densities used for designing the Kalman filter can be treated as *tuning* parameters to arrive at an observer for multivariable plants that has desirable properties, such as performance and robustness. The linear Kalman filter can also be used to design observers for *nonlinear* plants, by treating nonlinearities as process noise with appropriate power spectral density matrix.

Example 7.7

Let us design a Kalman filter to estimate the states of a double-pendulum (Example 4.12, Figure 1.5). A choice of the state variables for this fourth order system is $x_1(t) = \theta_1(t)$, $x_2(t) = \theta_2(t)$, $x_3(t) = \theta_1^{(1)}(t)$; $x_4(t) = \theta_2^{(1)}(t)$, which results in *nonlinear* state-equations given by Eq. (4.93). The function M-file for evaluating the time derivative of the state-vector, $\mathbf{x}^{(1)}(t)$, is called *doub.m* and is tabulated in Table 4.7. Thus, M-file assumes a known input torque acting on the pendulum given by $u(t) = 0.01\sin(5t)N - m$. It is desired to design a *linear* Kalman filter based on the known input, $u(t)$, and measurement of the angular position and angular velocity of the mass, m_1, i.e. $\mathbf{y}(t) = [\theta_1(t); \theta_1^{(1)}(t)]^T$, with the following *linearized* plant model:

$$\begin{aligned}
x_1^{(1)}(t) &= x_3(t) \\
x_2^{(1)}(t) &= x_4(t) \\
x_3^{(1)}(t) &= [m_2 g x_2(t) - (m_1 + m_2) g x_1(t)]/(m_1 L_1) \qquad (7.72) \\
x_4^{(1)}(t) &= -[g x_2(t) + L_1 dx_3(t)/dt]/L_2 + u(t)/(m_2 L_2^2)
\end{aligned}$$

which results in the following linear state coefficient matrices for the plant:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -(m_1+m_2)g/(m_1L_1) & m_2 g/(m_1 L_1) & 0 & 0 \\ (m_1+m_2)g/(m_1L_2) & -g(m_2/m_1+1)/L_2 & 0 & 0 \end{bmatrix};$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ -1/(m_2 L_2^2) \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad \mathbf{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (7.73)$$

The solution of the exact nonlinear state-equations of the plant with zero initial condition, i.e. $\mathbf{x}(0) = \mathbf{0}$, is obtained using the MATLAB Runge–Kutta solver *ode45* as follows:

```
>>[t,X] = ode45(@doub,[0 20],zeros(4,1)); <enter>
```

However, it is assumed that the state-vector, $\mathbf{x}(t)$, solved above, is unavailable for measurement, and only the output, $\mathbf{y}(t)$, can be measured, which is calculated by

```
>>C = [1 0 0 0; 0 0 1 0]; y = C*X'; <enter>
```

The state coefficient matrices, **A** and **B**, of the linearized plant are calculated with $m_1 = 1$ kg, $m_2 = 2$ kg, $L_1 = 1$ m, $L_2 = 2$ m, and $g = 9.8$ m/s^2 (same values as those used in *doub.m* for the nonlinear plant) as follows:

```
>>m1=1;m2=2;L1=1;L2=2;g=9.8; A = [0 0 1 0; 0 0 0 1; -(m1+m2)*g/(m1*L1)
  m2*g/(m1*L1) 0 0; (m1+m2)*g/(m1*L2) -g*(m2/m1+1)/L2 0 0] <enter>

A =
 0              0              1.0000        0
 0              0              0             1.0000
 -29.4000       19.6000        0             0
 14.7000        -14.7000       0             0

>>B = [0 0 0 1/(m2*L2*L2)]' <enter>


B =
     0
     0
     0
  0.1250
```

Since nonlinearities appear as *process noise* for all the state variables of the linear plant model, the process noise coefficient matrix is assumed to be an identity matrix, i.e. $\mathbf{F} = \mathbf{I}$. The linear Kalman filter is to be designed using the spectral densities of process and measurement noise, **V**, **Z**, and $\boldsymbol{\Psi}$ such that the exact state-vector, $\mathbf{x}(t)$, is accurately estimated. After some trial and error, we select $\mathbf{V} = 10^6\mathbf{I}$, $\mathbf{Z} = \mathbf{CC}^{\mathrm{T}}$, and $\Psi = 0$, and calculate the Kalman filter gain as follows:

```
>>[L,P,E]=lqe(A,eye(4),C,1e6*eye(4),C*C') <enter>

L =
  9.9989e+002 -1.3899e+001
  1.5303e+001  1.0256e+003
 -1.3899e+001  1.0202e+003
  2.1698e+001  4.8473e+002
```

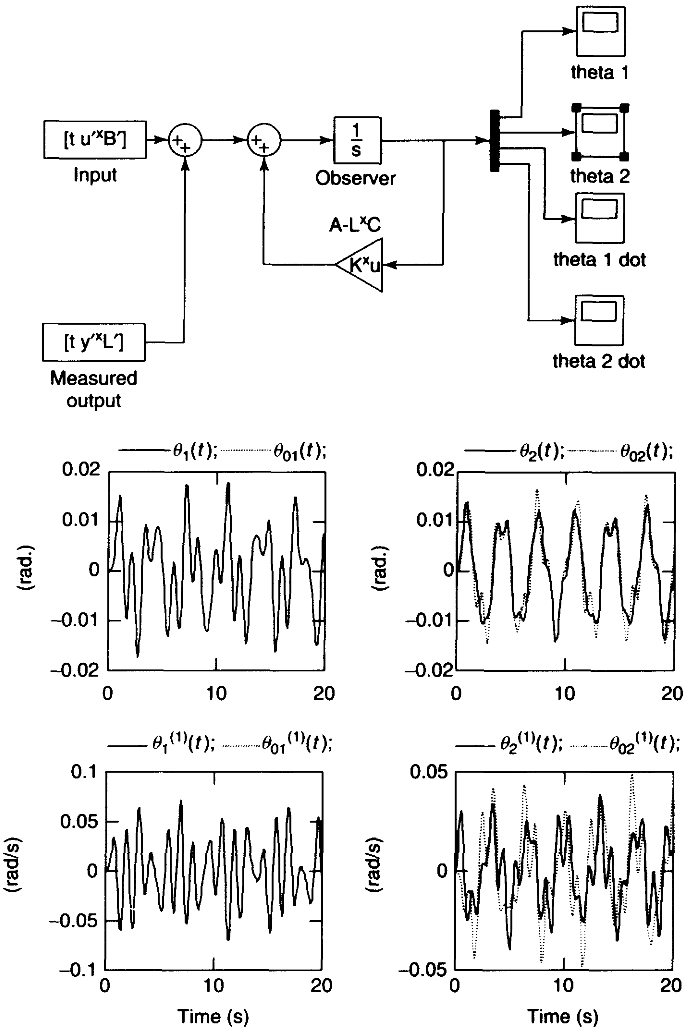

Figure 7.8 Actual (simulated) state-vector, $\mathbf{x}(t) = [\theta_1(t); \theta_2(t); \theta_1^{(1)}(t); \theta_2^{(1)}(t)]^T$, and estimated state-vector, $\mathbf{x}_o(t) = [\theta_{o1}(t); \theta_{o2}(t); \theta_{o1}^{(1)}(t); \theta_{o2}^{(1)}(t)]^T$, with a linear Kalman filter for the nonlinear dynamics of a double pendulum

```
P =
  9.9989e+002 1.5303e+001 -1.3899e+001 2.1698e+001
  1.5303e+001 5.3371e+004  1.0256e+003 2.6027e+004
 -1.3899e+001 1.0256e+003  1.0202e+003 4.8473e+002
  2.1698e+001 2.6027e+004  4.8473e+002 1.2818e+006

E =
 -1.0000e+003+ 1.5202e+001i
 -1.0000e+003- 1.5202e+001i
 -1.8785e+001
 -1.3039e+000
```