Lecture 10

Kalman Filtering

In this lecture we introduce the optimal linear state estimator known as the Kalman filter. The filter is optimal under the assumption that the plant is a linear system driven by white noise and that we want to minimize the mean square of the estimation error. Combined with an optimal linear state feedback (the LQ controller, developed in the previous lecture), they together form a linear-quadratic Gaussian (LQG) controller, see Figure 10.1.

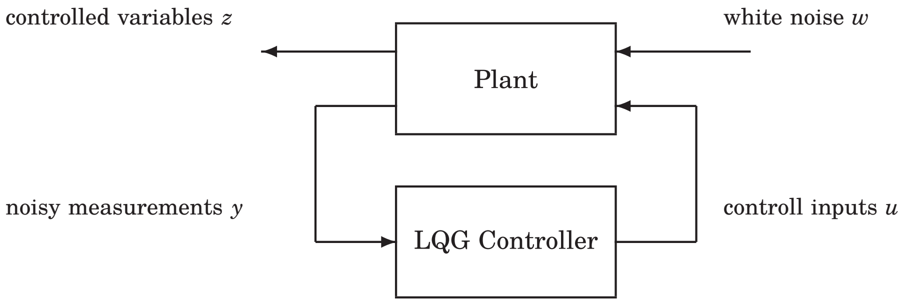

Figure 10.1 The LQG controller is optimal for a linear plant driven by white noise with a quadratic performance index.

The LQG controller minimizes a performance index of the form

$$J = \mathrm{E}\left\{x^T Q_1 x + 2x^T Q_{12} u + u^T Q_2 u\right\},$$

which can be seen as a weighted sum of the stationary variance of the signals in the control loop. Under the assumption of white process and measurement noise, all signals in the loop will be Gaussian.

10.1 Output feedback

A controller that is composed of a state observer and a feedback from the estimated state is called an output feedback controller, see Figure 10.2. In this lecture we will assume that the plant is subject to process disturbances w_1 as well as measurement noise w_2. In state-space

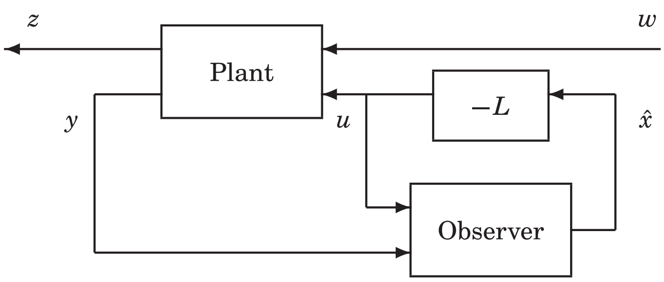

Figure 10.2 An output feedback controller.

form, a plant description is then

$$\begin{aligned}\frac{dx(t)}{dt} &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) \qquad\qquad\;\; + w_2(t)\end{aligned}$$

The controller consists of an observer that produces a state estimate $\hat{x}$ and a linear feedback from that estimate:

$$\begin{aligned}\frac{d\hat{x}(t)}{dt} &= A\hat{x}(t) + Bu(t) + K[y(t) - C\hat{x}(t)] \\ u(t) &= -L\hat{x}(t)\end{aligned}$$

Combining the two system descriptions above, we can formulate the closed-loop system and eliminate u and y as follows:

$$\begin{aligned}\frac{dx(t)}{dt} &= Ax(t) - BL\hat{x}(t) + w_1(t) \\ \frac{d\hat{x}(t)}{dt} &= A\hat{x}(t) - BL\hat{x}(t) + K[Cx(t) - C\hat{x}(t)] + Kw_2(t)\end{aligned}$$

Introducing the observer error $\tilde{x} = x - \hat{x}$, the state of the closed loop can equivalently be represented by $[x \;\; \tilde{x}]^T$, which has the dynamics

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ \tilde{x}(t)\end{bmatrix} = \begin{bmatrix} A - BL & BL \\ 0 & A - KC\end{bmatrix}\begin{bmatrix} x(t) \\ \tilde{x}(t)\end{bmatrix} + \begin{bmatrix} w_1(t) \\ w_1(t) - Kw_2(t)\end{bmatrix} \tag{10.1}$$

Note the block triangular structure of the system matrix. The eigenvalues of the state feedback (called the *control poles*) and of the Kalman filter (called the *observer poles*) are separate and are given by $\det(sI - A + BL) = 0$ and $\det(sI - A + KC) = 0$, respectively. This implies that the state feedback and the observer can be designed independently of each other. A formal statement regarding this *separation principle* will be given in the next lecture.

The state observer has in general dual goals: to estimate those state variables that cannot be directly measured and to filter out measurement noise. These goals are in general in conflict with each other. From the equation above we can see that we can make the observer error converge faster by increasing the filter gain K. At the same time, this will amplify the measurement noise w_2 more. In the next section we will find the optimal balance between the speed of estimation and the measurement noise attenuation in the Kalman filter.

10.2 The Kalman filter

The Kalman filter was first developed by Rudolf Kalman for discrete-time linear dynamical systems in 1960. A year later the results were extended to continuous-time systems by Kalman and Bucy, and this formulation is what we will use in this lecture. Kalman's great innovation was to give a state-space formulation of the optimal filtering problem, which allowed a recursive update of the estimate based on the most recent measurement value. As a bonus, the filter can also handle time-varying system and noise parameters. Previous optimal linear filtering results by Norbert Wiener (1949) were only applicable to stationary input-output formulations.

In feedback control applications we are most often interested in the optimal state estimate given measurements up to and including the current time. This is called the *filter case*. When treating prerecorded measurement data, a better estimate can be obtained by also considering measurements ahead of the current point—this is called *smoothing*. Another variant is to try to predict future values based on old measurements—this is called the *prediction* problem. Mathematically, the general problem is to estimate $x(k+m)$ given $\{y(i), u(i) \,|\, i \le k\}$, see Figure 10.3. In this lecture we will only discuss the filter problem, but the solution to the general problem is a simple extension.

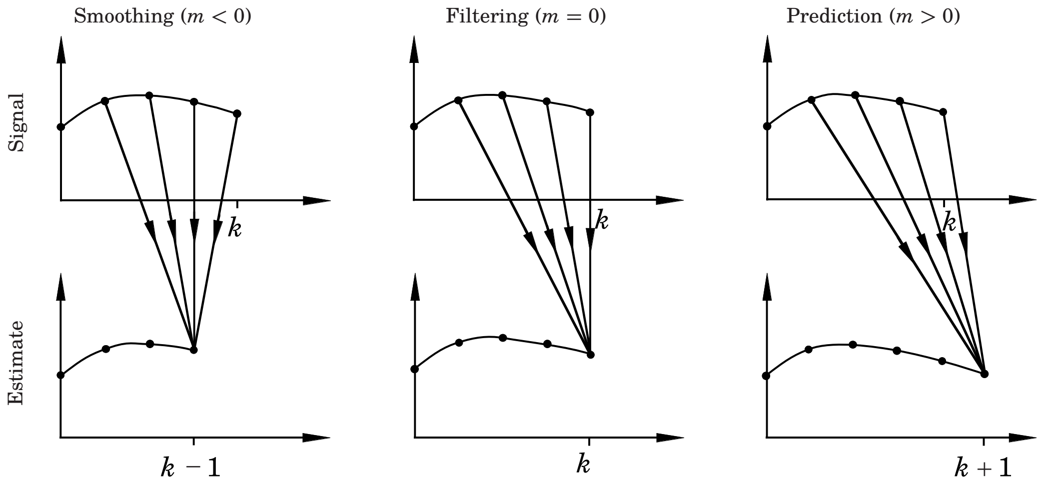

Figure 10.3 Smoothing, filtering, and prediction are variants of the same general problem: to estimate $x(k+m)$ given $\{y(i),\, u(i) \mid i \le k\}$.

Derivation of the Kalman filter

The Kalman filter can be derived in a number of different ways, under various assumptions. Here we will assume that the process is a continuous-time linear system driven by white noise, and that the filter has the structure of the standard linear observer discussed in the previous section. Further, we assume that the plant is *detectable*, meaning that any unstable modes must be observable. Starting from (10.1), the observer error dynamics are given by

$$\frac{d\tilde{x}}{dt} = (A - KC)\tilde{x} + \begin{pmatrix} I & -K \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

The disturbance process $w = \begin{pmatrix} w_1 & w_2 \end{pmatrix}^T$ is assumed to be white with constant spectral density

$$\Phi_w(\omega) = \begin{pmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{pmatrix} > 0$$

To find the optimal observer, we are seeking the observer gain K that minimizes the stationary observer error variance

$$P = \mathrm{E}\,\tilde{x}\,\tilde{x}^T.$$

Assuming $(A - KC)$ to be stable, the stationary variance P can be calculated from the Lyapunov equation

$$(A-KC)P + P(A-KC)^T + \begin{pmatrix} I & -K \end{pmatrix} \begin{pmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{pmatrix} \begin{pmatrix} I \\ -K^T \end{pmatrix} = 0$$

Completing the square,

$$\begin{aligned} AP+PA^T+R_1+(K-(PC^T+R_{12})R_2^{-1})R_2(K-(PC^T+R_{12})R_2^{-1})^T \\ -(PC^T+R_{12})R_2^{-1}(PC^T+R_{12})^T = 0 \end{aligned}$$

we find that the minimium variance will be attained for the filter gain

$$K = (PC^T + R_{12})R_2^{-1}$$

What remains is an algebraic Riccati equation,

$$AP + PA^T + R_1 - (PC^T + R_{12})R_2^{-1}(PC^T + R_{12})^T = 0$$

We summarize the result in the following theorem:

Table 10.1 Conversion of parameters between the LQ control problem and the Kalman filtering problem

LQ control	Kalman filter
A	A^T
B	C^T
Q_1	R_1
Q_2	R_2
Q_{12}	R_{12}
S	P
L	K^T

THEOREM 10.1—THE KALMAN FILTER
Given a detectable linear plant disturbed by white noise,

$$\begin{cases} \dot{x} = Ax + Bu + w_1 \\ y = Cx + w_2 \end{cases} \qquad \Phi_w = \begin{pmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{pmatrix} > 0$$

the optimal observer (in the mean-square error sense) is given by the Kalman filter

$$\frac{d\hat{x}}{dt} = A\hat{x} + Bu + K(y - C\hat{x})$$

where the Kalman gain K is given by

$$K = (PC^T + R_{12})R_2^{-1}$$

where $P = \mathrm{E}(x - \hat{x})(x - \hat{x})^T$ is the positive definite solution to

$$AP + PA^T + R_1 - (PC^T + R_{12})R_2^{-1}(PC^T + R_{12})^T = 0$$

□

Note that the solution does not depend on what states we are interested in. The Kalman filter produces the optimal estimate of all states at the same time. Further, the optimal observer gain K is static since we are solving a steady-state problem. The Kalman filter can also be derived for finite-horizon problems and problems with time-varying system matrices. We then obtain a Riccati differential equation for $P(t)$ and a time-varying filter gain $K(t)$.

Duality between LQ control and Kalman filtering

The optimal state feedback problem and the optimal filtering problem display many similarities, and they are in fact dual problems. The algebraic Riccati equations associated with both problems are very similar, and one problem can be symbolically translated into the other using Table 10.1.

10.3 Examples

Here we give two different examples of Kalman filtering.

EXAMPLE 1. Consider an integrator process with both process and measurement noise:

$$\begin{aligned} \dot{x}(t) &= w_1(t) \\ y(t) &= x(t) + w_2(t) \end{aligned} \qquad \Phi_w = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix}$$

The Kalman filter for the process is given by

$$\frac{d\hat{x}}{dt} = A\hat{x}(t) + Bu(t) + K[y(t) - C\hat{x}(t)]$$

where K is obtained by first solving the algebraic Riccati equation,

$$R_1 - P^2/R_2 \quad \Rightarrow \quad P = \sqrt{R_1 R_2},$$

which then yields the filter gain

$$K = P/R_2 = \sqrt{R_1/R_2}$$

Note that the optimal gain only depends on the ratio of R_1 and R_2. When R_2 is large compared to R_1, the filter attenuates measurement noise by reducing the gain K. At the same time the filter pole $A - KC = -K$ becomes slower, which gives slow error convergence. In the Laplace domain, the filter equation becomes

$$\hat{X}(s) = \frac{K}{s+K} Y(s)$$

which shows that the Kalman filter in this case is a first-order low-pass filter with a break frequency that depends on $\sqrt{R_1/R_2}$. □

EXAMPLE 2. Consider the problem of tracking a moving object in two dimensions, relying on very noisy GPS position readings, see Figure 10.4(a). Not knowing anything about the object we are tracking, we can model it as a double integrator in each dimension, driven by white noise. Let (p_1, p_2) denote the position coordinates. We can then set up the dynamical model

$$\begin{aligned} \ddot{p}_1 &= w_1 \\ \ddot{p}_2 &= w_2 \end{aligned}$$

where the noise processes w_1 and w_2 are assumed independent for simplicity. Introducing the state vector

$$x = (p_1 \;\; \dot{p}_1 \;\; p_2 \;\; \dot{p}_2)^T$$

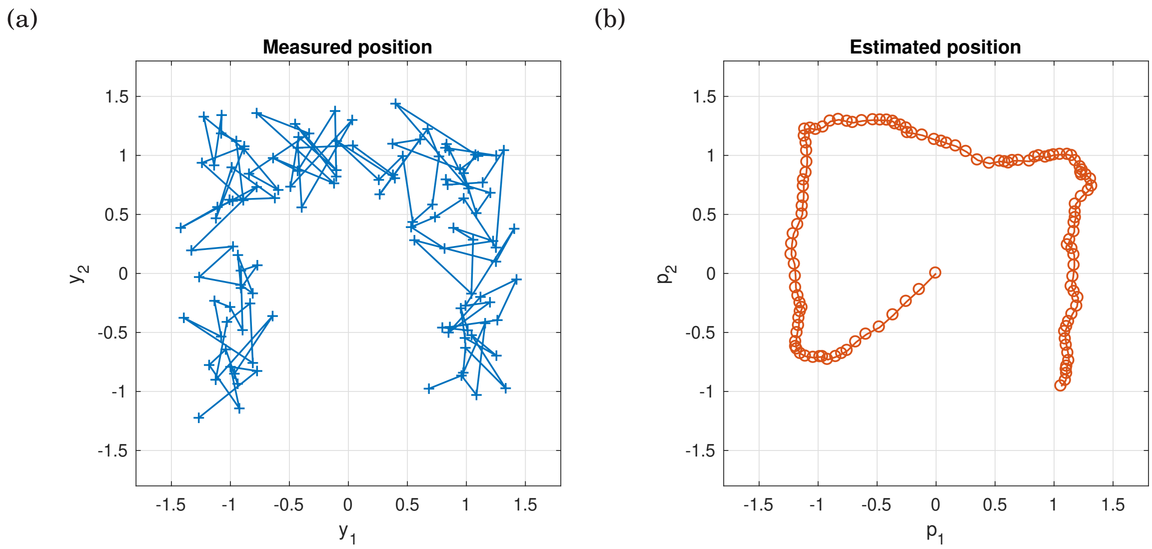

Figure 10.4 Tracking of a moving object in two dimensions: (a) Position measurements with noise. (b) Position estimates from a Kalman filter.

we can write the model in state-space form as

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} w_{11}$$

$$y = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} x + w_{12}$$

To design a Kalman filter we need to specify the noise intensities, which are unknown to us. Fixing the process noise intensity to $R_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we can tune the speed and noise attenuation of the filter by selecting different values of R_2. A larger value of R_2 gives better noise rejection but slower tracking. A simulation of the output of the Kalman filter starting with the initial condition $\hat{x}(0) = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^T$ is shown in Figure 10.4(b). □

10.4 Frequency-domain noise shaping

So far we have assumed that the process disturbance w_1 and the measurement noise w_2 have constant spectral densities. Recall from Lecture 3 that, in many cases, we have some knowledge about the characteristics of the disturbances. For instance, load disturbances are often dominated by low frequencies, while measurement noise has more high-frequency components. The Kalman filter can be tuned in the frequency domain by extending the plant model with filters that shape the process disturbance and measurement noise spectra accordingly. The idea is illustrated in Figure 10.5(a). The spectrum of the process disturbance is shaped via the filter H_1, while the spectrum of the measurement noise is shaped via the filter H_2. The relative size of $|H_1(i\omega)|$ and $|H_2(i\omega)|$ at each frequency will then influence the gain of the Kalman filter for that particular frequency. Modeling a large input disturbance will increase the Kalman gain, while a large measurement disturbance will decrease the gain. It can be shown that the Kalman filter will have zeros at the locations of the poles of the output noise filter $H_2(s)$.

A common example of this design approach is to model an integral disturbance acting on the process input, see Figure 10.5(b). Since the Kalman filter contains a model of the plant, this will introduce an integrator in the Kalman filter. The disturbance state x_i will be observable but not controllable. The effect of the input disturbance can however be canceled out by the controller by extending the feedback law according to

$$u(t) = -L\hat{x}(t) - \hat{x}_i(t)$$

A different way to introduce integral action in the LQG controller will be discussed in the next lecture.

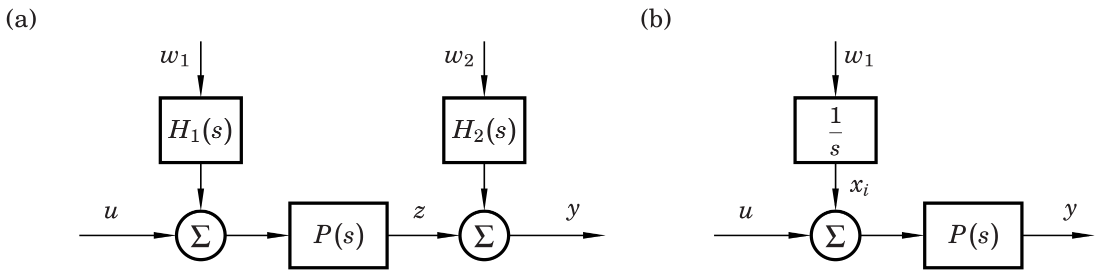

Figure 10.5 Extended plant models with shaped noise. (a) General input and output noise filters. (b) Integrator input disturbance.