where $\overline{\Sigma}$ is given by

$$\overline{\Sigma} = \Sigma - \Sigma C'(C\Sigma C' + N)^{-1}C\Sigma,$$

and Σ is the unique positive semidefinite symmetric solution of the Riccati equation

$$\Sigma = A\big(\Sigma - \Sigma C'(C\Sigma C' + N)^{-1}C\Sigma\big)A' + M.$$

The assumptions required for this are that *the pair (A, C) is observable and that the matrix M can be written as $M = DD'$, where D is a matrix such that the pair (A, D) is controllable*. The steady-state controller of Eqs. (5.10)-(5.12) is particularly attractive for practical implementation. Furthermore, as shown in Appendix E, it results in a stable closed-loop system, under the preceding controllability and observability assumptions.

5.3 MINIMUM VARIANCE CONTROL OF LINEAR SYSTEMS

We have considered so far the control of linear systems in state variable form as in the previous section. However, linear systems are often modeled by means of an input-output equation, which is more economical in the number of parameters needed to describe the system dynamics. In this section we consider single-input, single-output, linear, time-invariant systems, and a special type of quadratic cost function. The resulting optimal policy is particularly simple and has found wide application. We first introduce some of the basic facts regarding linear systems in input-output form. Detailed discussions may be found in the books by Aström and Wittenmark [AsW84], [AsW90], Goodwin and Sin [GoS84], and Whittle [Whi63].

We consider a single-input single-output discrete-time linear system, which is specified by an equation of the form

$$y_k + a_1 y_{k-1} + \cdots + a_m y_{k-m} = b_0 u_k + b_1 u_{k-1} + \cdots + b_m u_{k-m}, \quad (5.13)$$

where a_i, b_i are given scalars. The scalar sequences $\{u_k \mid k = 0, \pm 1, \pm 2, \ldots\}$, $\{y_k \mid k = 0, \pm 1, \pm 2, \ldots\}$ are viewed as the input and output of the system, respectively. Note that we allow time to extend to $-\infty$ as well as $+\infty$; this will be useful for describing generic properties of the system relating to stability. We will later revert to our usual convention of starting at time 0 and proceeding forward.

It is convenient to describe this type of system by means of the *backward shift operator*, denoted s, which when operating on a sequence $\{x_k \mid k = 0, \pm 1, \pm 2, \ldots\}$ shifts its index by one unit; that is,

$$s(x_k) = x_{k-1}, \qquad k = 0, \pm 1, \pm 2, \ldots$$

We denote by s^r the operator resulting from r successive applications of s:

$$s^r(x_k) = x_{k-r}, \qquad k = 0, \pm 1, \pm 2, \ldots \tag{5.14}$$

We also write for simplicity $s^r x_k = x_{k-r}$. The *forward shift operator*, denoted s^{-1}, is the inverse of s and is defined by

$$s^{-1}(x_k) = x_{k+1}, \qquad k = 0, \pm 1, \pm 2, \ldots$$

Thus the notation (5.14) holds for all integers r. We can form linear combinations of operators of the form s^r. Thus, for example, the operator $(s + 2s^2)$ is defined by

$$(s + 2s^2)(x_k) = x_{k-1} + 2x_{k-2}, \qquad k = 0, \pm 1, \pm 2, \ldots$$

With this notation, Eq. (5.13) can be written as

$$A(s)y_k = B(s)u_k,$$

where $A(s)$, $B(s)$ are the operators

$$A(s) = 1 + a_1 s + \cdots + a_m s^m,$$

$$B(s) = b_0 + b_1 s + \cdots + b_m s^m.$$

Sometimes it is convenient to write the equation $A(s)y_k = B(s)u_k$ as

$$y_k = \frac{B(s)}{A(s)} u_k$$

or

$$\frac{A(s)}{B(s)} y_k = u_k.$$

The meaning of both equations is that the sequences $\{y_k\}$ and $\{u_k\}$ are related via $A(s)y_k = B(s)u_k$. There is a certain ambiguity here in that, for a fixed $\{u_k\}$, the equation $A(s)y_k = B(s)u_k$ has an infinite number of solutions in $\{y_k\}$. For example, the equation

$$y_k + a y_{k-1} = u_k$$

for $u_k \equiv 0$ has as solutions all sequences of the form $y_k = \beta(-a)^k$, where β is any scalar; the solution becomes unique only after some boundary condition for the sequence $\{y_k\}$ is specified. As will be discussed shortly, however, for stable systems and for a *bounded* sequence $\{u_k\}$ there is a unique solution $\{y_k\}$ that is *bounded*. It is this solution that will be denoted by $\big(B(s)/A(s)\big)u_k$ in what follows. The reader who is familiar with linear

dynamic system theory will note that $B(s)/A(s)$ can be viewed as a *transfer function* involving z-transforms.

We now introduce some terminology. When the sequences $\{y_k\}$ and $\{u_k\}$ satisfy $A(s)y_k = B(s)u_k$, we say that y_k is *obtained by passing* u_k *through the filter* $B(s)/A(s)$. This comes from engineering terminology, where linear time-invariant systems are commonly referred to as filters. We also refer to the equation $A(s)y_k = B(s)u_k$ as the filter $B(s)/A(s)$.

A filter $B(s)/A(s)$ is said to be *stable* if the polynomial $A(s)$ has all its (complex) roots strictly outside the unit circle of the complex plane; that is, $|\rho| > 1$ for all complex ρ satisfying $A(\rho) = 0$. A stable filter $B(s)/A(s)$ has the following two properties:

(a) Every solution $\{y_k\}$ of
$$A(s)y_k = 0$$
satisfies $\lim_{k\to\infty} y_k = 0$; that is, the output y_k tends to zero if the input sequence $\{u_k\}$ is identically zero.

(b) For every bounded sequence $\{\overline{u}_k\}$, the equation
$$A(s)y_k = B(s)\overline{u}_k$$
has a *unique* solution $\{\overline{y}_k\}$ within the class of bounded sequences. Furthermore, every solution $\{y_k\}$ (possibly unbounded) of the equation satisfies
$$\lim_{k\to\infty} (y_k - \overline{y}_k) = 0.$$
For example, consider the system
$$y_k - 0.5y_{k-1} = u_k.$$
Given the bounded input sequence $\overline{u}_k = \{\ldots, 1, 1, 1, \ldots\}$, the set of all solutions is given by
$$y_k = 2 + \frac{\beta}{2^k},$$
where β is a scalar, but of these the only bounded solution is $\overline{y}_k = \{\ldots, 2, 2, 2, \ldots\}$. The solution $\{\overline{y}_k\}$ can thus be naturally associated with the input sequence $\{u_k\}$; it is also known as the *forced response* of the system due to the input $\{u_k\}$.

ARMAX Models – Reduction to State Space Form

We now consider a linear system with output y_k, which is driven by two inputs: a random noise input ϵ_k, and a control input u_k. It has the form

$$\begin{aligned} y_k + a_1 y_{k-1} + \cdots + a_m y_{k-m} = {} & b_1 u_{k-1} + \cdots + b_m u_{k-m} \\ & + \epsilon_k + c_1 \epsilon_{k-1} + \cdots + c_m \epsilon_{k-m}, \end{aligned} \tag{5.15}$$

and it is known as an ARMAX model (AutoRegressive, Moving Average, with eXogenous input). We assume throughout that the random variables ϵ_k are mutually independent. We can write the model in the shorthand form

$$A(s)y_k = B(s)u_k + C(s)\epsilon_k,$$

where the polynomials $A(s)$, $B(s)$, and $C(s)$ are given by

$$\begin{aligned} A(s) &= 1 + a_1 s + \cdots + a_m s^m, \\ B(s) &= b_1 s + \cdots + b_m s^m, \\ C(s) &= 1 + c_1 s + \cdots + c_m s^m. \end{aligned}$$

The ARMAX model is very common and its derivation is outlined in Appendix F, where it is shown that without loss of generality we can assume that $C(s)$ *has no roots strictly inside the unit circle*. For much of the analysis in subsequent sections, it will be necessary to exclude the critical case where $C(s)$ has roots *on* the unit circle and assume that $C(s)$ has all its roots strictly outside the unit circle. This assumption is usually satisfied in practice.

In several situations, analysis and algorithms relating to the ARMAX model are greatly simplified if $C(s) = 1$ so that the noise terms $C(s)\epsilon_k = \epsilon_k$ are independent. However, this is typically an unrealistic assumption. To emphasize this point and see how easily the noise can be correlated, suppose that we have a first-order system

$$x_{k+1} = ax_k + w_k,$$

where we observe

$$y_k = x_k + v_k.$$

Then

$$\begin{aligned} y_{k+1} &= x_{k+1} + v_{k+1} \\ &= ax_k + w_k + v_{k+1} \\ &= a(y_k - v_k) + w_k + v_{k+1}, \end{aligned}$$

so finally

$$y_{k+1} = ay_k + v_{k+1} - av_k + w_k.$$

However, the noise sequence $\{v_{k+1} - av_k + w_k\}$ is correlated even if $\{v_k\}$ and $\{w_k\}$ are individually and mutually independent.

The ARMAX model (5.15) can be put into state space form. The process is based on state augmentation and can perhaps be best understood in terms of an example. Consider the system

$$y_k + a_1 y_{k-1} + a_2 y_{k-2} = b_1 u_{k-1} + b_2 u_{k-2} + \epsilon_k + c_1 \epsilon_{k-1}. \tag{5.16}$$

We have

$$\begin{pmatrix} y_{k+1} \\ y_k \\ u_k \\ \epsilon_{k+1} \end{pmatrix} = \begin{pmatrix} -a_1 & -a_2 & b_2 & c_1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_k \\ y_{k-1} \\ u_{k-1} \\ \epsilon_k \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 1 \\ 0 \end{pmatrix} u_k + \begin{pmatrix} \epsilon_{k+1} \\ 0 \\ 0 \\ \epsilon_{k+1} \end{pmatrix}. \tag{5.17}$$

By setting

$$x_k = \begin{pmatrix} y_k \\ y_{k-1} \\ u_{k-1} \\ \epsilon_k \end{pmatrix}, \qquad w_k = \begin{pmatrix} \epsilon_{k+1} \\ 0 \\ 0 \\ \epsilon_{k+1} \end{pmatrix},$$

$$A = \begin{pmatrix} -a_1 & -a_2 & b_2 & c_1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

we can write Eq. (5.17) as

$$x_{k+1} = Ax_k + Bu_k + w_k,$$

where $\{w_k\}$ is a stationary independent process. We have arrived at this state space model through state augmentation. Notice that the state x_k includes ϵ_k. Thus if the controller is assumed to know at time k only the present and past outputs $y_k, y_{k-1}, \ldots$, and past controls $u_{k-1}, u_{k-2}, \ldots$ (but not $\epsilon_k, \epsilon_{k-1}, \ldots$), we are faced with a model of imperfect state information. If $c_1 = 0$ in Eq. (5.16) then the state space model can be simplified so that

$$x_k = \begin{pmatrix} y_k \\ y_{k-1} \\ u_{k-1} \end{pmatrix},$$

in which case we have perfect state information. More generally, we have perfect state information in the ARMAX model (5.15) if $b_1 \neq 0$ and $c_1 = c_2 = \cdots = c_m = 0$.

Minimum Variance Control: Perfect State Information Case

Consider the perfect state information case of the ARMAX model (5.15):

$$y_k + a_1 y_{k-1} + \cdots + a_m y_{k-m} = b_1 u_{k-1} + \cdots + b_m u_{k-m} + \epsilon_k,$$

where $b_1 \neq 0$. A problem of interest, known as the *minimum variance control problem*, is to select u_k as a function of the present and past outputs $y_k, y_{k-1}, \ldots$, as well as the past controls $u_{k-1}, u_{k-2}, \ldots$, so as to minimize the cost

$$E\left\{\sum_{k=1}^{N} (y_k)^2\right\}.$$

There are no constraints on u_k. By transforming the system to state space form, we see that this problem can be reduced to a perfect state information linear-quadratic problem where the state x_k is

$$\big(y_k, y_{k-1}, \ldots, y_{k-m+1}, u_{k-1}, \ldots, u_{k-m+1}\big)'.$$

The problem is of the same nature as the linear-quadratic problem of Section 4.1 except that the corresponding matrices R_k in the quadratic cost function are zero here. Nonetheless, in Section 4.1 we used the invertibility of R_k only to ensure that various matrices in the optimal policy and the Riccati equation are invertible. If invertibility of these matrices can be guaranteed by other means, the same analysis applies even if R_k is positive semidefinite. This is indeed the case here. An analysis analogous to the one of Section 4.1 shows that the optimal control u_k^* at time k (given $y_k, y_{k-1}, \ldots, y_{k-m+1}$ and $u_{k-1}, \ldots, u_{k-m+1}$) is the same as the one that would be applied if all future disturbances $\epsilon_{k+1}, \ldots, \epsilon_N$ were set equal to zero, their expected value (certainty equivalence). It follows that

$$\mu_k^*(y_k, \ldots, y_{k-m+1}, u_{k-1}, \ldots, u_{k-m+1}) = \frac{1}{b_1}(a_1 y_k + \cdots + a_m y_{k-m+1} - b_2 u_{k-1} - \cdots - b_m u_{k-m+1}),$$

and $\{u_k^*\}$ is generated via the equation

$$b_1 u_k^* + b_2 u_{k-1}^* + \cdots + b_m u_{k-m+1}^* = a_1 y_k + a_2 y_{k-1} + \cdots + a_m y_{k-m+1}.$$

In other words, $\{u_k^*\}$ is generated by passing $\{y_k\}$ through the linear filter $\overline{A}(s)/\overline{B}(s)$, where

$$\overline{A}(s) = a_1 + a_2 s + \cdots + a_m s^{m-1} = s^{-1}\big(A(s) - 1\big),$$

$$\overline{B}(s) = b_1 + b_2 s + \cdots + b_m s^{m-1} = s^{-1} B(s),$$

as shown in Fig. 5.3.1. The resulting closed-loop system is

$$y_k = \epsilon_k \tag{5.18}$$

and the associated cost is

$$N\, E\{(\epsilon_k)^2\}.$$

Notice that the optimal policy, called *minimum variance* control law, is time invariant and does not depend on the horizon N.

Whereas the optimal closed-loop system as given by Eq. (5.18) is clearly stable, we can anticipate serious difficulties if the filter $\overline{A}(s)/\overline{B}(s)$ in the feedback loop is unstable. For if $\overline{B}(s)$ has some roots inside the unit circle, then the sequence $\{u_k\}$ will tend to be unbounded. This is illustrated by the following example.

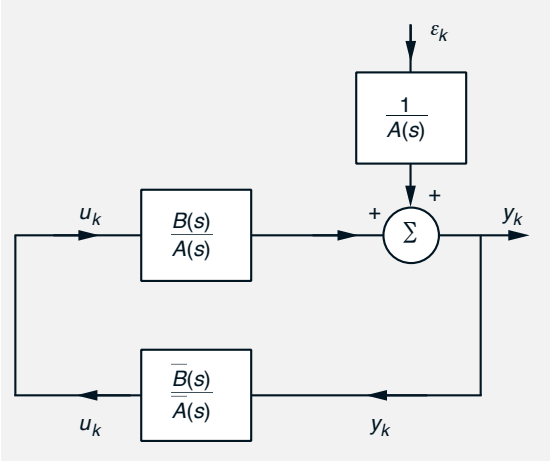

Figure 5.3.1 Minimum variance control with perfect state information. Structure of the optimal closed-loop system, where $A(s) = 1 + a_1 s + \cdots + a_m s^m$, $B(s) = b_1 s + \cdots + b_m s^m$, $\overline{A}(s) = s^{-1}\big(A(s) - 1\big)$, and $\overline{B}(s) = s^{-1}B(s)$.

Example 5.3.1 (An Optimal but Unstable Controller)

Consider the system

$$y_k + y_{k-1} = u_{k-1} - 2u_{k-2} + \epsilon_k.$$

The optimal policy is

$$u_k = 2u_{k-1} + y_k$$

and the optimal closed-loop system is

$$y_k = \epsilon_k,$$

which is a stable system. On the other hand, the last two equations yield

$$u_k = 2u_{k-1} + \epsilon_k.$$

Thus, u_k is generated by an *unstable* system, and in fact it is given by

$$u_k = \sum_{n=0}^{k} 2^n \epsilon_{k-n}.$$

Therefore, even though the output y_k stays bounded, the control u_k typically becomes unbounded.

For another view of the same difficulty, suppose that the coefficients $b_1, \ldots, b_m$ of $\overline{B}(s)$ are slightly different from the ones of the true system.

We will show that if the feedback filter $\overline{A}(s)/\overline{B}(s)$ is unstable, then the closed-loop system is also unstable in the sense that both u_k and y_k become unbounded with probability one.

Assume that the system is governed by

$$A^0(s)y_k = B^0(s)u_k + \epsilon_k, \tag{5.19}$$

while the policy is calculated under the assumption that the system model is

$$A(s)y_k = B(s)u_k + \epsilon_k,$$

where the coefficients of $A(s)$ and $B(s)$ differ slightly from those of $A^0(s)$, $B^0(s)$. Define $\overline{A}^0(s)$, $\overline{B}^0(s)$ by

$$1 + s\overline{A}^0(s) = A^0(s),$$

$$s\overline{B}^0(s) = B^0(s).$$

Note that $\overline{A}^0(s) = \overline{A}(s)$ and $\overline{B}^0(s) = \overline{B}(s)$ if $A^0(s) = A(s), B^0(s) = B(s)$. By multiplying Eq. (5.19) with $\overline{B}(s)$ and by using the relation defining the optimal policy

$$\overline{B}(s)u_k = \overline{A}(s)y_k,$$

we obtain

$$\overline{B}(s)A^0(s)y_k = B^0(s)\overline{A}(s)y_k + \overline{B}(s)\epsilon_k.$$

If the coefficients of $\overline{A}^0(s)$ and $\overline{B}^0(s)$ are close to those of $\overline{A}(s)$, $\overline{B}(s)$, then the roots of the polynomial

$$\overline{B}(s) + s\big(\overline{B}(s)\overline{A}^0(s) - \overline{B}^0(s)\overline{A}(s)\big)$$

are close to the roots of $\overline{B}(s)$. Thus *the closed-loop system is stable only if the roots of* $\overline{B}(s)$ *are outside the unit circle*, or equivalently, if and only if the filter $\overline{A}(s)/\overline{B}(s)$ is stable. If our model is exact, the closed-loop system will be stable due to what is commonly referred to as a *pole-zero cancellation*. However, the slightest modeling discrepancy will induce instability.

The conclusion from the preceding analysis is that the minimum variance control law is advisable only if it can be realized through a stable filter [$\overline{B}(s)$ has roots outside the unit circle]. Even if $\overline{B}(s)$ has its roots outside the unit circle, but some of these roots are near the unit circle, the performance of the minimum variance policy can be very sensitive to variations in the parameters of the polynomials $A(s)$ and $B(s)$. One way to overcome this sensitivity is to change the cost to

$$E\left\{\sum_{k=1}^{N}\big((y_k)^2 + R(u_{k-1})^2\big)\right\},$$

where R is some positive parameter. This requires solution via the Riccati equation as in Section 4.1. For a detailed derivation, see Aström [Ast83].

In some problems, the system equation includes an additional external input sequence $\{v_k\}$, the values of which can be measured by the controller as they occur. In particular, consider the scalar system

$$y_k + a_1 y_{k-1} + \cdots + a_m y_{k-m} = b_1 u_{k-1} + \cdots + b_m u_{k-m} + d_1 v_{k-1} + \cdots + d_m v_{k-m} + \epsilon_k,$$

where each value v_k becomes known to the controller without error at time k. The minimum variance controller then takes the form

$$\mu_k^*(y_k, \ldots, y_{k-m+1}, u_{k-1}, \ldots, u_{k-m+1}, v_k, \ldots, v_{k-m+1}) = \frac{1}{b_1}\big(a_1 y_k + \cdots + a_m y_{k-m+1} - d_1 v_k - \cdots - d_m v_{k-m+1} - b_2 u_{k-1} \cdots - b_m u_{k-m+1}\big),$$

and the optimal controls $\{u_k^*\}$ are generated by

$$\overline{B}(s) u_k^* = \overline{A}(s) y_k - \overline{D}(s) v_k,$$

where

$$\overline{A}(s) = a_1 + a_2 s + \cdots + a_m s^{m-1},$$
$$\overline{B}(s) = b_1 + b_2 s + \cdots + b_m s^{m-1},$$
$$\overline{D}(s) = d_1 + d_2 s + \cdots + d_m s^{m-1}.$$

The closed-loop system is again $y_k = \epsilon_k$, but for practical purposes it is stable only if $\overline{B}(s)$ has its roots outside the unit circle. The process whereby external inputs are measured and used for control is commonly referred to as *feedforward control*.

Imperfect State Information Case

Consider now the general ARMAX model

$$y_k + a_1 y_{k-1} + \cdots + a_m y_{k-m} = b_M u_{k-M} + \cdots + b_m u_{k-m} + \epsilon_k + c_1 \epsilon_{k-1} + \cdots + c_m \epsilon_{k-m}$$

or, equivalently,

$$A(s) y_k = B(s) u_k + C(s) \epsilon_k,$$

where

$$A(s) = 1 + a_1 s + \cdots + a_m s^m,$$
$$B(s) = b_M s^M + \cdots + b_m s^m,$$
$$C(s) = 1 + c_1 s + \cdots + c_m s^m.$$

We assume the following:

(1) $b_M \neq 0$ and $1 \leq M \leq m$.

(2) $\{\epsilon_k\}$ is a zero mean, independent, stationary process.

(3) The polynomial $C(s)$ has all its roots outside the unit circle. (As explained in Appendix F, this assumption is not overly restrictive.)

The controller knows at each time k the past inputs and outputs. Thus the information vector at time k is

$$I_k = \big(y_k, y_{k-1}, \ldots, y_{-m+1}, u_{k-1}, u_{k-2}, \ldots, u_{-m+M}\big).$$

(We include in the information vector the control inputs $u_{-1}, \ldots, u_{-m+M}$. If control starts at time 0, these inputs will be zero.) There are no constraints on u_k. The problem is to find a policy $\big\{\mu_0(I_0), \ldots, \mu_{N-1}(I_{N-1})\big\}$ that minimizes

$$E\left\{\sum_{k=1}^{N}(y_k)^2\right\}.$$

By using state augmentation, we can cast this problem into the framework of the linear-quadratic problem of Section 5.2. The corresponding linear system in state space format involves a state x_k given by

$$x_k = (y_{k+M-1}, \ldots, y_{k+M-m}, u_{k-1}, \ldots, u_{k+M-m}, \epsilon_{k+M-1}, \ldots, \epsilon_{k+M-m}).$$

Because $y_{k+M-1}, \ldots, y_{k+1}$ and $\epsilon_{k+M-1}, \ldots, \epsilon_{k+M-m}$ are unknown to the controller, we are faced with a problem of imperfect state information.

An analysis analogous to the one of Section 5.2 shows that certainty equivalence holds; that is, the *optimal control u_k^* at time k given I_k is the same as the one that would be applied in the deterministic problem where the current state*

$$x_k = (y_{k+M-1}, \ldots, y_{k+M-m}, u_{k-1}, \ldots, u_{k+M-m}, \epsilon_{k+M-1}, \ldots, \epsilon_{k+M-m})$$

is set equal to its conditional expected value given I_k, and the future disturbances $\epsilon_{k+M}, \ldots, \epsilon_N$ are set equal to zero (their expected value).

Thus the optimal control $u_k^* = \mu_k^*(I_k)$ is obtained by solving for u_k the equation

$$\begin{aligned} E\{y_{k+M} \mid u_k, I_k\} &= E\big\{y_{k+M} \mid y_k, y_{k-1}, \ldots, y_{-m+1}, u_k, u_{k-1}, \ldots, u_{-m+M}\big\} \\ &= 0. \end{aligned}$$

This leads to the problem of calculating $E\{y_{k+M} \mid I_k, u_k\}$, known as the *forecasting* or *prediction* problem, which is important in its own right. We first treat the easier case where there is no delay ($M = 1$) and then discuss the more general case where the delay can be positive.

Forecasting for ARMAX Models – No Delay ($M = 1$)

Assume that $M = 1$. We would like to generate an equation for the forecast $E\{y_{k+1} \mid I_k, u_k\}$, and then determine the optimal control $u_k^* = \mu_k^*(I_k)$ by setting this forecast to zero. Let us introduce an auxiliary sequence $\{z_k\}$ via the equation

$$z_k = y_k - \epsilon_k.$$

A key fact is that, since $\{\epsilon_k\}$ is an independent, zero-mean sequence, we have

$$E\{z_{k+1} \mid I_k, u_k\} = E\{y_{k+1} \mid I_k, u_k\}.$$

We can thus obtain the desired forecast of y_{k+1} by forecasting z_{k+1} instead. We can then obtain the optimal control u_k^* by setting $E\{z_{k+1} \mid I_k, u_k^*\} = 0$.

By using the definition $z_k = y_k - \epsilon_k$ to express y_k in terms of z_k in the ARMAX model equation for $M = 1$, we obtain

$$z_{k+1} + c_1 z_k + \cdots + c_m z_{k-m+1} = b_1 u_k + \cdots + b_m u_{k-m+1} + w_k, \quad (5.20)$$

where

$$w_k = (c_1 - a_1)y_k + \cdots + (c_m - a_m)y_{k-m+1}.$$

We note that w_k is perfectly observable by the controller; however, the scalars $z_k, \ldots, z_{k-m+1}$ are not known to the controller because the initial conditions $z_0, \ldots, z_{1-m}$ of the system (5.20) are unknown. Nonetheless, the system (5.20) is stable, since the roots of the polynomial $C(s)$ have been assumed to be outside the unit circle. As a result, the initial conditions do not matter asymptotically. In other words, if we generate a sequence $\{\hat{y}_k\}$ using the system (5.20) and zero initial conditions, i.e.,

$$\hat{y}_{k+1} + c_1 \hat{y}_k + \cdots + c_m \hat{y}_{k-m+1} = b_1 u_k + \cdots + b_m u_{k-m+1} + w_k,$$

with

$$\hat{y}_0 = 0, \qquad \hat{y}_{-1} = 0, \qquad \ldots \qquad \hat{y}_{1-m} = 0,$$

then we will have

$$\lim_{k \to \infty} \big(\hat{y}_k - z_k\big) = 0.$$

Thus, $\hat{y}_{k+1}$ is an asymptotically accurate approximation to the optimal forecast $E\{y_{k+1} \mid I_k, u_k\}$.

Minimum Variance Control: Imperfect State Information and No Delay

Based on the earlier discussion, an asymptotically accurate approximation to the minimum variance policy is obtained by setting u_k to the value that makes $\hat{y}_{k+1} = 0$; that is, by solving for u_k the equation

$$\hat{y}_{k+1} + c_1 \hat{y}_k + \cdots + c_m \hat{y}_{k-m+1} = b_1 u_k + \cdots + b_m u_{k-m+1} + w_k.$$

If this policy is followed, however, the earlier forecasts $\hat{y}_k, \ldots, \hat{y}_{k-m+1}$ will be equal to zero. Thus the (approximate) minimum variance policy is given by

$$\begin{aligned} u_k &= \frac{1}{b_1}\big(w_k - b_2 u_{k-1} - \cdots - b_m u_{k-m+1}\big) \\ &= \frac{1}{b_1}\big((a_1 - c_1)y_k + \cdots + (a_m - c_m)y_{k-m-1} \\ &\qquad - b_2 u_{k-1} - \cdots - b_m u_{k-m+1}\big). \end{aligned}$$

By substituting this policy in the ARMAX model

$$\begin{aligned} y_{k+1} + a_1 y_k + \cdots + a_m y_{k-m+1} = b_1 u_k + \cdots + b_m u_{k-m+1} \\ + \epsilon_{k+1} + c_1 \epsilon_k + \cdots + c_m \epsilon_{k-m+1}, \end{aligned}$$

we see that the closed-loop system becomes

$$y_{k+1} - \epsilon_{k+1} + c_1(y_k - \epsilon_k) + \cdots + c_m(y_{k-m+1} - \epsilon_{k-m+1}) = 0,$$

or equivalently $C(s)(y_k - \epsilon_k) = 0$. Since $C(s)$ has its roots outside the unit circle, this is a stable system, and we have

$$y_k = \epsilon_k + \gamma(k),$$

where $\gamma(k) \to 0$ as $k \to \infty$.

Forecasting: The General Case

Consider now the general case where the delay M can be greater than 1. The forecasting problem can still be nicely solved by using a certain trick to transform the ARMAX equation into a more convenient form. To this end, we first obtain polynomials $F(s)$ and $G(s)$ of the form

$$F(s) = 1 + f_1 s + \cdots + f_{M-1} s^{M-1},$$

$$G(s) = g_0 + g_1 s + \cdots + g_{m-1} s^{m-1},$$

which satisfy

$$C(s) = A(s)F(s) + s^M G(s). \tag{5.21}$$

The coefficients of $F(s)$ and $G(s)$ are uniquely determined from those of $C(s)$ and $A(s)$ by equating coefficients of both sides of the relation

$$\begin{aligned} 1 + c_1 s + \cdots + c_m s^m = (1 + a_1 s + \cdots + a_m s^m)(1 + f_1 s + \cdots + f_{M-1} s^{M-1}) \\ + s^M (g_0 + g_1 s + \cdots + g_{m-1} s^{m-1}). \end{aligned}$$

Example 5.3.2

Let $m = 3$ and $M = 2$. Then the preceding equation takes the form

$$1 + c_1 s + c_2 s^2 + c_3 s^3 = (1 + a_1 s + a_2 s^2 + a_3 s^3)(1 + f_1 s) + s^2(g_0 + g_1 s + g_2 s^2),$$

and by equating coefficients we have

$$c_1 = a_1 + f_1, \quad c_2 = a_2 + a_1 f_1 + g_0, \quad c_3 = a_3 + a_2 f_1 + g_1, \quad a_3 f_1 + g_2 = 0,$$

from which f_1, g_0, g_1, and g_2 are uniquely determined.

The ARMAX model can be written as

$$A(s)y_{k+M} = \overline{B}(s)u_k + C(s)\epsilon_{k+M}, \tag{5.22}$$

where

$$\overline{B}(s) = s^{-M}B(s) = b_M + b_{M+1}s + \cdots + b_m s^{m-M}.$$

Multiplying both sides of Eq. (5.22) with $F(s)$, we have

$$F(s)A(s)y_{k+M} = F(s)\overline{B}(s)u_k + F(s)C(s)\epsilon_{k+M},$$

and using Eq. (5.21) to express $F(s)A(s)$ as $C(s) - s^M G(s)$, we obtain

$$\big(C(s) - s^M G(s)\big)y_{k+M} = F(s)\overline{B}(s)u_k + F(s)C(s)\epsilon_{k+M},$$

or equivalently

$$C(s)\big(y_{k+M} - F(s)\epsilon_{k+M}\big) = F(s)\overline{B}(s)u_k + G(s)y_k. \tag{5.23}$$

Let us now introduce the auxiliary sequence $\{z_k\}$ via the equation

$$z_{k+M} = y_{k+M} - F(s)\epsilon_{k+M} = y_{k+M} - \epsilon_{k+M} - f_1\epsilon_{k+M-1} - \cdots - f_{M-1}\epsilon_{k+1}.$$

Note that when $M = 1$, we have $F(s) = 1$ and $z_k = y_k - \epsilon_k$, so $\{z_k\}$ is the same sequence as the one introduced earlier for the case of no delay. Again, since $\{\epsilon_k\}$ is an independent, zero-mean sequence, by taking expectations in the definition $z_{k+M} = y_{k+M} - F(s)\epsilon_{k+M}$, we obtain

$$E\{z_{k+M} \mid I_k, u_k\} = E\{y_{k+M} \mid I_k, u_k\},$$

and we can obtain the desired forecast of y_{k+M} by forecasting z_{k+M} in its place. Furthermore, by Eq. (5.23), z_{k+M} is written as

$$C(s)z_{k+M} = w_k$$

or

$$z_{k+M} + c_1 z_{k+M-1} + \cdots + c_m z_{k+M-m} = w_k, \tag{5.24}$$

where

$$w_k = F(s)\overline{B}(s)u_k + G(s)y_k. \tag{5.25}$$

Since the scalar w_k of Eq. (5.25) is available at time k (i.e., it is determined from I_k and u_k), the system (5.24) can serve as a basis for forecasting z_{k+M}. We would be able to predict exactly z_{k+M} and use it as a forecast of y_{k+M} if we knew appropriate initial conditions with which to start the equation (5.24) that generates it. We don't know such initial conditions, but because this equation represents a stable system, the choice of initial conditions does not matter asymptotically, as we proceed to explain more formally.

We consider the sequence $\hat{y}_{k+M}$ generated by

$$\hat{y}_{k+M} + c_1 \hat{y}_{k+M-1} + \cdots + c_m \hat{y}_{k+M-m} = w_k$$

with initial condition

$$\hat{y}_{M-1} = \hat{y}_{M-2} = \cdots = \hat{y}_{M-m} = 0, \tag{5.26}$$

and we claim that the forecast $E\{z_{k+M} \mid I_k\}$ can be approximated by $\hat{y}_{k+M}$. To see this, note that from Eqs. (5.24) to (5.26) we have

$$z_{k+M} = \hat{y}_{k+M} + \big(\gamma_1(k) z_{M-1} + \cdots + \gamma_m(k) z_{M-m}\big)$$

and

$$E\{z_{k+M} \mid I_k, u_k\} = \hat{y}_{k+M} + \sum_{i=1}^{m} \gamma_i(k) E\{z_{M-i} \mid I_k, u_k\},$$

where $\gamma_1(k), \ldots, \gamma_m(k)$ are appropriate scalars depending on k. Since $C(s)$ has all its roots outside the unit circle, we have (compare with the discussion on stability earlier in this section)

$$\lim_{k\to\infty} \gamma_1(k) = \lim_{k\to\infty} \gamma_2(k) = \cdots = \lim_{k\to\infty} \gamma_m(k) = 0.$$

It follows that, for large values of k,

$$\hat{y}_{k+M} \simeq E\{z_{k+M} \mid I_k, u_k\} = E\{y_{k+M} \mid I_k, u_k\}.$$

(More precisely, we have $|\hat{y}_{k+M} - E\{y_{k+M} \mid I_k, u_k\}| \to 0$ as $k \to \infty$, where the convergence is in the mean-square sense.)

In conclusion, an asymptotically accurate approximation to the optimal forecast $E\{y_{k+M} \mid I_k, u_k\}$ is given by $\hat{y}_{k+M}$ and is generated by the equation

$$\hat{y}_{k+M} + c_1 \hat{y}_{k+M-1} + \cdots + c_m \hat{y}_{k+M-m} = F(s)\overline{B}(s)u_k + G(s)y_k \tag{5.27}$$

with the initial condition

$$\hat{y}_{M-1} = \hat{y}_{M-2} = \cdots = \hat{y}_{M-m} = 0. \tag{5.28}$$

Minimum Variance Control: The General Case

Based on the earlier discussion, the minimum variance policy is obtained by solving for u_k the equation $E\{y_{k+M} \mid I_k, u_k\} = 0$. Thus an asymptotically accurate approximation is obtained by setting u_k to the value that makes $\hat{y}_{k+M} = 0$, that is, by solving for u_k the equation [cf. Eqs. (5.27) and (5.28)]

$$F(s)\overline{B}(s)u_k + G(s)y_k = c_1\hat{y}_{k+M-1} + \cdots + c_m\hat{y}_{k+M-m}.$$

If this policy is followed, however, the earlier forecasts $\hat{y}_{k+M-1}, \ldots, \hat{y}_{k+M-m}$ will be equal to zero. Thus the (approximate) minimum variance policy is given by

$$F(s)\overline{B}(s)u_k + G(s)y_k = 0; \tag{5.29}$$

that is, u_k^* is generated by passing y_k through the linear filter

$$-G(s)/F(s)\overline{B}(s),$$

as shown in Fig. 5.3.2.

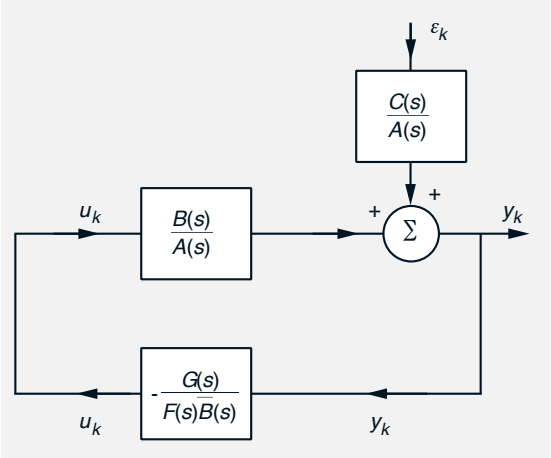

Figure 5.3.2 Minimum variance control with imperfect state information. Structure of the optimal closed-loop system.

From Eqs. (5.23) and (5.29), we obtain the equation for the closed-loop system

$$C(s)\big(y_{k+M} - F(s)\epsilon_{k+M}\big) = 0.$$

Since $C(s)$ has its roots outside the unit circle, we obtain

$$y_{k+M} = F(s)\epsilon_{k+M} + \gamma(k),$$

where $\gamma(k) \to 0$ as $k \to \infty$. So asymptotically, the closed-loop system takes the form

$$y_k = \epsilon_k + f_1\epsilon_{k-1} + \cdots + f_{M-1}\epsilon_{k-M+1}.$$

Let us consider now the stability properties of the closed-loop system when the true system parameters differ slightly from those of the assumed model. Let the true system be described by

$$A^0(s)y_k = s^M\overline{B}^0(s)u_k + C^0(s)\epsilon_k, \tag{5.30}$$

while u_k is given by the minimum variance policy

$$F(s)\overline{B}(s)u_k + G(s)y_k = 0, \tag{5.31}$$

where

$$C(s) = A(s)F(s) + s^M G(s).$$

Operating on Eq. (5.30) with $F(s)\overline{B}(s)$ and using Eq. (5.31), we obtain

$$F(s)\overline{B}(s)A^0(s)y_k = -s^M\overline{B}^0(s)G(s)y_k + F(s)\overline{B}(s)C^0(s)\epsilon_k.$$

Combining the last two equations and collecting terms, we have

$$\Big\{F(s)\overline{B}(s)A^0(s) + \big(C(s) - A(s)F(s)\big)\overline{B}^0(s)\Big\}y_k = F(s)\overline{B}(s)C^0(s)\epsilon_k$$

or

$$\Big\{\overline{B}^0(s)C(s) + F(s)\big(\overline{B}(s)A^0(s) - A(s)\overline{B}^0(s)\big)\Big\}y_k = F(s)\overline{B}(s)C^0(s)\epsilon_k.$$

If the coefficients of $A^0(s)$, $\overline{B}^0(s)$, and $C^0(s)$ are near those of $A(s)$, $\overline{B}(s)$, and $C(s)$, then the poles of the closed-loop system will be near the roots of $\overline{B}(s)C(s)$. Thus *the closed-loop system will be in effect stable only if the roots of* $\overline{B}(s)$ *are strictly outside the unit circle*, similar to the perfect state information case examined earlier.

5.4 SUFFICIENT STATISTICS

The main difficulty with the DP algorithm for imperfect state information problems is that it is carried out over a state space of expanding dimension. As a new measurement is added at each stage k, the dimension of the state (the information vector I_k) increases accordingly. This motivates an effort to reduce the data that are truly necessary for control purposes. In other words, it is of interest to look for quantities known as *sufficient statistics*,

APPENDIX F: Modeling of Stochastic Linear Systems

In this appendix we show how controlled linear time-invariant systems with stochastic inputs can be represented by the ARMAX model used in Section 5.3.

F.1 LINEAR SYSTEMS WITH STOCHASTIC INPUTS

Consider a linear system with output $\{y_k\}$, control input $\{u_k\}$, and an additional zero-mean random input $\{w_k\}$. We assume that $\{w_k\}$ is a stationary (up to second order) stochastic process. That is, $\{w_k\}$ is a sequence of random variables satisfying, for all $i, k = 0, \pm 1, \pm 2, \ldots,$

$$E\{w_k\} = 0, \qquad E\{w_0 w_i\} = E\{w_k w_{k+i}\} < \infty.$$

(All references to stationary processes in this section are meant in the limited sense just described.) By linearity, y_k is the sum of one sequence $\{y_k^1\}$ due to the presence of $\{u_k\}$ and another sequence $\{y_k^2\}$ due to the presence of $\{w_k\}$:

$$y_k = y_k^1 + y_k^2. \tag{F.1}$$

We assume that y_k^1 and y_k^2 are generated by some filters $B_1(s)/A_1(s)$ and $B_2(s)/A_2(s)$, respectively:

$$A_1(s) y_k^1 = B_1(s) u_k, \tag{F.2a}$$

$$A_2(s)y_k^2 = B_2(s)w_k. \tag{F.2b}$$

Operating on Eqs. (F.2a) and (F.2b) with $A_2(s)$ and $A_1(s)$, respectively, adding, and using Eq. (F.1), we obtain

$$\overline{A}(s)y_k = \overline{B}(s)u_k + v_k, \tag{F.3}$$

where $\overline{A}(s) = A_1(s)A_2(s)$, $\overline{B}(s) = A_2(s)B_1(s)$, and $\{v_k\}$, given by

$$v_k = A_1(s)B_2(s)w_k, \tag{F.4}$$

is a zero-mean, generally correlated, stationary stochastic process.

We are interested in the case where u_k is a control input applied after y_k has occurred and has been observed, so that in Eq. (F.2a) we have $B_1(0) = 0$. Then, we may assume that the polynomials $\overline{A}(s)$ and $\overline{B}(s)$ have the form

$$\overline{A}(s) = 1 + \overline{a}_1 s + \cdots + \overline{a}_{m_0} s^{m_0}, \qquad \overline{B}(s) = \overline{b}_1 s + \cdots + \overline{b}_{m_0} s^{m_0}$$

for some scalars $\overline{a}_i$ and $\overline{b}_i$, and some positive integer m_0.

To summarize, we have constructed a model of the form

$$\overline{A}(s)y_k = \overline{B}(s)u_k + v_k,$$

where $\overline{A}(s)$ and $\overline{B}(s)$ are polynomials of the preceding form and $\{v_k\}$ is some zero-mean, correlated, stationary stochastic process. We now need to model further the sequence $\{v_k\}$.

F.2 PROCESSES WITH RATIONAL SPECTRUM

Given a zero-mean, stationary scalar process $\{v_k\}$, denote by $V(k)$ the autocorrelation function

$$V(k) = E\{v_i v_{i+k}\}, \qquad k = 0, \pm 1, \pm 2, \ldots$$

We say that $\{v_k\}$ has *rational spectrum* if the transform of $\{V(k)\}$ defined by

$$S_v(\lambda) = \sum_{k=-\infty}^{\infty} V(k)e^{-jk\lambda}$$

exists for $\lambda \in [-\pi, \pi]$ and can be expressed as

$$S_v(\lambda) = \sigma^2 \frac{|C(e^{i\lambda})|^2}{|D(e^{j\lambda})|^2}, \qquad \lambda \in [-\pi, \pi], \tag{F.5}$$

where σ is a scalar, $C(z)$ and $D(z)$ are some polynomials with real coefficients

$$C(z) = 1 + c_1 z + \cdots + c_m z^m, \tag{F.6a}$$

$$D(z) = 1 + d_1 z + \cdots + d_m z^m, \tag{F.6b}$$

and $D(z)$ has no roots on the unit circle $\{z \mid |z| = 1\}$.

The following facts are of interest:

(a) If $\{v_k\}$ is an uncorrelated process with $V(0) = \sigma^2$, $V(k) = 0$ for $k \neq 0$, then

$$S_v(\lambda) = \sigma^2, \qquad \lambda \in [-\pi, \pi],$$

and clearly $\{v_k\}$ has rational spectrum.

(b) If $\{v_k\}$ has rational spectrum S_v given by Eq. (F.5), then S_v can be written as

$$S_v(\lambda) = \tilde{\sigma}^2 \frac{|\tilde{C}(e^{j\lambda})|^2}{|\tilde{D}(e^{j\lambda})|^2}, \qquad \lambda \in [-\pi, \pi],$$

where $\tilde{\sigma}$ is a scalar and $\tilde{C}(z)$, $\tilde{D}(z)$ are unique real polynomials of the form

$$\tilde{C}(z) = 1 + \tilde{c}_1 z + \cdots + \tilde{c}_m z^m,$$

$$\tilde{D}(z) = 1 + \tilde{d}_1 z + \cdots + \tilde{d}_m z^m,$$

such that:

(1) $\tilde{C}(z)$ has all its roots outside or on the unit circle, and if $C(z)$ has no roots on the unit circle, then the same is true for $\tilde{C}(z)$.

(2) $\tilde{D}(z)$ has all roots strictly outside the unit circle.

These facts are seen by noting that if $\rho \neq 0$ is a root of $D(z)$, then $|D(e^{j\lambda})|^2 = D(e^{j\lambda})D(e^{-j\lambda})$ contains a factor

$$(1 - \rho^{-1}e^{j\lambda})(1 - \rho^{-1}e^{-j\lambda}) = \rho^{-2}(\rho - e^{j\lambda})(\rho - e^{-j\lambda}).$$

A little reflection shows that the roots of $\tilde{D}(z)$ should be ρ or ρ^{-1} depending on whether ρ is outside or inside the unit circle. Similarly, the roots of $\tilde{C}(z)$ are obtained from the roots of $C(z)$. Thus the polynomials $\tilde{C}(z)$ and $\tilde{D}(z)$ as well as $\tilde{\sigma}^2$ can be uniquely determined. We may thus assume without loss of generality that $C(z)$ and $D(z)$ in Eq. (F.5) have no roots inside the unit circle.

There is a fundamental result here that relates to the realization of processes with rational spectrum. The proof is hard; see for example, Ash and Gardner [AsG75, pp. 75-76].

Proposition F.1: If $\{v_k\}$ is a zero-mean, stationary stochastic process with rational spectrum

$$S_v(\lambda) = \sigma^2 \frac{|C(e^{j\lambda})|^2}{|D(e^{j\lambda})|^2}, \qquad \lambda \in [-\pi, \pi],$$

where the polynomials $C(s)$ and $D(s)$ are given by

$$C(s) = 1 + c_1 s + \cdots + c_m s^m, \qquad D(s) = 1 + d_1 s + \cdots + d_m s^m,$$

and are assumed (without loss of generality) to have no roots inside the unit circle, then there exists a zero-mean, uncorrelated stationary process $\{\epsilon_k\}$ with $E\{\epsilon_k^2\} = \sigma^2$ such that for all k

$$v_k + d_1 v_{k-1} + \cdots + d_m v_{k-m} = \epsilon_k + c_1 \epsilon_{k-1} + \cdots + c_m \epsilon_{k-m}.$$

F.3 THE ARMAX MODEL

Let us now return to the problem of representation of a linear system with stochastic inputs. We had arrived at the model

$$\overline{A}(s) y_k = \overline{B}(s) u_k + v_k. \tag{F.7}$$

If the zero-mean stationary process $\{v_k\}$ has rational spectrum, the preceding analysis and proposition show that there exists a zero-mean, uncorrelated stationary process $\{\epsilon_k\}$ satisfying

$$D(s) v_k = C(s) \epsilon_k,$$

where $C(s)$ and $D(s)$ are polynomials, and $C(s)$ has no roots inside the unit circle. Operating on both sides of Eq. (F.7) with $D(s)$ and using the relation $D(s)v_k = C(s)\epsilon_k$, we obtain

$$A(s) y_k = B(s) u_k + C(s) \epsilon_k, \tag{F.8}$$

where $A(s) = D(s)\overline{A}(s)$ and $B(s) = D(s)\overline{B}(s)$. Since $\overline{A}(0) = 1$, $\overline{B}(0) = 0$, we can write Eq. (F.8) as

$$y_k + \sum_{i=1}^{m} a_i y_{k-i} = \sum_{i=1}^{m} b_i u_{k-i} + \epsilon_k + \sum_{i=1}^{m} c_i \epsilon_{k-i},$$

for some integer m and scalars a_i, b_i, c_i, $i = 1, \ldots, m$. This is the ARMAX model that we have used in Section 5.3.

may not help appreciably in solving a different instance.

(b) *The problem data changes as the system is being controlled.* As an example, consider the route planning example in case (a) above, and assume that new service points to be visited arise as the vehicle is on its way. It is possible in principle to model these data changes in terms of stochastic disturbances, but then we may end up with a problem that is too complicated for analysis or solution by DP. A frequently employed alternative is to use *on-line replanning*, whereby the problem is resolved on-line with the new data, as soon as these data become available, and control continues with a policy that corresponds to the new data.

A common feature of the above situations, which can seriously impact the solution, is that there may be stringent time constraints for the computation of the controls. This may substantially exacerbate the "curse of dimensionality" problem mentioned above.

As indicated by the above discussion, in practice one often has to settle for a suboptimal control scheme that strikes a reasonable balance between convenient implementation and adequate performance. In this chapter we discuss some general approaches for suboptimal control, which are based on approximations to the DP algorithm. We begin with two general schemes to simplify the DP computation, certainty equivalent control (Section 6.1), which replaces the stochastic quantities of the problem by deterministic nominal values, and open-loop-feedback control (Section 6.2), which ignores in part the availability of information in the future. These two schemes set the stage for limited lookahead control, which together with its many variations (Sections 6.3-6.5), is one of the principal approaches for suboptimal control. We also discuss adaptive control in the context of certainty equivalent control. This discussion is not used in subsequent developments, so the reader may skip Sections 6.1.1-6.1.4 if desired.

6.1 CERTAINTY EQUIVALENT AND ADAPTIVE CONTROL

The *certainty equivalent controller* (CEC) is a suboptimal control scheme that is inspired by linear-quadratic control theory. It applies at each stage the control that would be optimal if the uncertain quantities were fixed at some "typical" values; that is, it acts as if a form of the certainty equivalence principle were holding.

The advantage of the CEC is that it replaces the DP algorithm with what is often a much less demanding computation: the solution of a *deterministic* optimal control problem at each stage. This problem yields an optimal control sequence, the first component of which is used at the current stage, while the remaining components are discarded. The main attractive characteristic of the CEC is its ability to deal with stochastic

and even imperfect information problems by using the mature and effective methodology of deterministic optimal control.

We describe the CEC for the general problem with imperfect state information of Section 5.1. As can be expected, the implementation is considerably simpler if the controller has perfect state information. Suppose that we have an "estimator" that uses the information vector I_k to produce a "typical" value $\overline{x}_k(I_k)$ of the state. Assume also that for every state-control pair (x_k, u_k) we have selected a "typical" value of the disturbance, which we denote by $\overline{w}_k(x_k, u_k)$. For example, if the state spaces and disturbance spaces are convex subsets of Euclidean spaces, the expected values

$$\overline{x}_k(I_k) = E\{x_k \mid I_k\}, \qquad \overline{w}_k(x_k, u_k) = E\{w_k \mid x_k, u_k\},$$

can serve as typical values.

The control input $\overline{\mu}_k(I_k)$ applied by the CEC at each time k is determined by the following rule:

(1) Given the information vector I_k, compute the state estimate $\overline{x}_k(I_k)$.

(2) Find a control sequence $\{\overline{u}_k, \overline{u}_{k+1}, \ldots, \overline{u}_{N-1}\}$ that solves the deterministic problem obtained by fixing the uncertain quantities x_k and $w_k, \ldots, w_{N-1}$ at their typical values:

$$\text{minimize} \quad g_N(x_N) + \sum_{i=k}^{N-1} g_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big)$$

subject to the initial condition $x_k = \overline{x}_k(I_k)$ and the constraints

$$u_i \in U_i, \qquad x_{i+1} = f_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big), \qquad i = k, k+1, \ldots, N-1.$$

(3) Use as control the first element in the control sequence found:

$$\overline{\mu}_k(I_k) = \overline{u}_k.$$

Note that step (1) is unnecessary if we have perfect state information; in this case we simply use the known value of the x_k. The deterministic optimization problem in step (2) must be solved at each time k, once the initial state $\overline{x}_k(I_k)$ becomes known by means of an estimation (or perfect observation) procedure. A total of N such problems must be solved by the CEC at every system run. In many cases of interest, these deterministic problems can be solved by powerful numerical methods such as conjugate gradient, Newton's method, augmented Lagrangian, and sequential quadratic programming methods; see e.g. Luenberger [Lue84] or Bertsekas [Ber99]. Furthermore, the implementation of the CEC requires no storage of the type required for the optimal feedback controller.

An alternative to solving N optimal control problems in an "on-line" fashion is to solve these problems a priori. This is accomplished by computing an optimal feedback controller for the deterministic optimal control problem obtained from the original problem by replacing all uncertain quantities by their typical values. It is easy to verify, based on the equivalence of open-loop and feedback implementation of optimal controllers for deterministic problems, that the implementation of the CEC given earlier is equivalent to the following.

Let $\{\mu_0^d(x_0), \ldots, \mu_{N-1}^d(x_{N-1})\}$ be an optimal controller obtained from the DP algorithm for the deterministic problem

$$\text{minimize} \quad g_N(x_N) + \sum_{k=0}^{N-1} g_k\big(x_k, \mu_k(x_k), \overline{w}_k(x_k, u_k)\big)$$

$$\text{subject to} \quad x_{k+1} = f_k\big(x_k, \mu_k(x_k), \overline{w}_k(x_k, u_k)\big), \quad \mu_k(x_k) \in U_k, \quad k \geq 0.$$

Then the control input $\overline{\mu}_k(I_k)$ applied by the CEC at time k is given by

$$\overline{\mu}_k(I_k) = \mu_k^d\big(\overline{x}_k(I_k)\big)$$

as shown in Fig. 6.1.1.

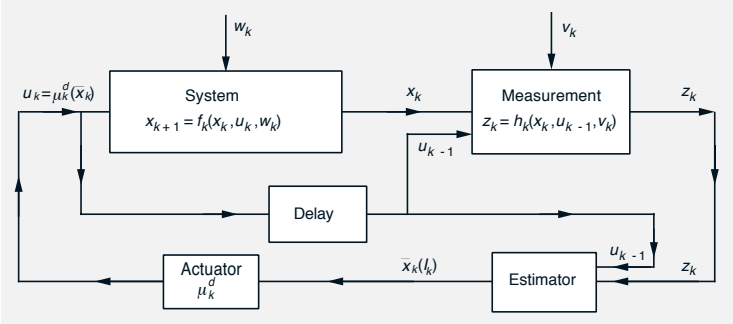

Figure 6.1.1 Structure of the certainty equivalent controller when implemented in feedback form.

In other words, an equivalent alternative implementation of the CEC consists of finding a feedback controller $\{\mu_0^d, \mu_1^d, \ldots, \mu_{N-1}^d\}$ that is optimal for a corresponding deterministic problem, and subsequently using this controller for control of the uncertain system [modulo substitution of the state x_k by its estimate $\overline{x}_k(I_k)$]. Either one of the definitions given for the CEC can serve as a basis for its implementation. Depending on the nature of the problem, one method may be preferable to the other.

The CEC approach often performs well in practice and yields near-optimal policies. In fact, for the linear-quadratic problems of Sections 4.1

and 5.2, the CEC is identical to the optimal controller (certainty equivalence principle). It is possible, however, that a CEC performs strictly worse than the optimal open-loop controller (see Exercise 6.2).

In what follows in this section, we will discuss a few variants of the CEC, and we will then focus on one particular type of methodology, adaptive control of systems with unknown parameters.

Certainty Equivalent Control with Heuristics

Even though the CEC approach simplifies a great deal the computations, it still requires the solution of a deterministic optimal control problem at each stage. This problem may be difficult, and a more convenient approach may be to solve it suboptimally using a heuristic algorithm. To simplify notation, let us assume perfect state information [the ideas to be discussed can also be applied to imperfect state information problems, by substituting x_k with its estimate $\overline{x}_k(I_k)$]. Then, in this approach, given x_k, we use some (easily implementable) heuristic to find a suboptimal control sequence $\{\overline{u}_k, \overline{u}_{k+1}, \ldots, \overline{u}_{N-1}\}$ for the problem

$$\text{minimize} \quad g_N(x_N) + \sum_{i=k}^{N-1} g_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big)$$

subject to

$$u_i \in U_i(x_i), \qquad x_{i+1} = f_i\big(x_i, u_i, \overline{w}_i(x_i, u_i)\big), \qquad i = k, k+1, \ldots, N-1.$$

We then use $\overline{u}_k$ as the control for stage k.

An important enhancement of this idea is to use minimization over the first control u_k and to use the heuristic only for the remaining stages $k+1, \ldots, N-1$. To implement this variant of the CEC, we must apply at time k a control $\overline{u}_k$ that minimizes over $u_k \in U_k(x_k)$ the expression

$$g_k\big(x_k, u_k, \overline{w}_k(x_k, u_k)\big) + H_{k+1}\Big(f_k\big(x_k, u_k, \overline{w}_k(x_k, u_k)\big)\Big), \tag{6.1}$$

where H_{k+1} is the cost-to-go function corresponding to the heuristic, i.e., $H_{k+1}(x_{k+1})$ is the cost incurred over the remaining stages $k+1, \ldots, N-1$ starting from a state x_{k+1}, using the heuristic, and assuming that the future disturbances will be equal to their typical values $\overline{w}_i(x_i, u_i)$. Note that for any next-stage state x_{k+1}, it is not necessary to have a closed-form expression for the heuristic cost-to-go $H_{k+1}(x_{k+1})$. Instead we can generate this cost by running the system forward from x_{k+1} and accumulating the corresponding single-stage costs. Since the heuristic must be run for each possible value of the control u_k to calculate the costs $H_{k+1}\big(f_k\big(x_k, u_k, \overline{w}_k(x_k, u_k)\big)\big)$ needed in the minimization, it is necessary to discretize the control constraint set if it is not already finite.

Note that the general structure of the preceding variant of the CEC is similar to the one of standard DP. It involves minimization of the expression (6.1), which is the sum of a current stage cost and a cost-to-go starting from the next state. The difference with DP is that the optimal cost-to-go $J^*_{k+1}(x_{k+1})$ is replaced by the heuristic cost $H_{k+1}(x_{k+1})$, and the disturbance w_k is replaced by its typical value $\overline{w}_k(x_k, u_k)$ (so that there is no need to take expectation over w_k). We thus encounter for the first time an important suboptimal control idea, based on an approximation to the DP algorithm: *minimizing at each stage k the sum of approximations to the current stage cost and the optimal cost-to-go*. This idea is central in other types of suboptimal control such as the limited lookahead, rollout, and model predictive control approaches, which will be discussed in Sections 6.3-6.5.

Partially Stochastic Certainty Equivalent Control

In the preceding descriptions of the CEC all future disturbances are fixed at their typical values. A useful variation for some imperfect state information problems is to take into account the stochastic nature of these disturbances, and to treat the problem as one of perfect state information, using an estimate $\overline{x}_k(I_k)$ of x_k as if it were exact. Thus, if $\{\mu^p_0(x_0), \ldots, \mu^p_{N-1}(x_{N-1})\}$ is an optimal policy obtained from the DP algorithm for the stochastic *perfect state information* problem

$$\text{minimize} \quad E\left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k\big(x_k, \mu_k(x_k), w_k\big) \right\}$$

$$\text{subject to} \quad x_{k+1} = f_k\big(x_k, \mu_k(x_k), w_k\big), \quad \mu_k(x_k) \in U_k, \quad k = 0, \ldots, N-1,$$

then the control input $\overline{\mu}_k(I_k)$ applied by this variant of CEC at time k is given by

$$\overline{\mu}_k(I_k) = \mu^p_k\big(\overline{x}_k(I_k)\big).$$

Generally, there are several variants of the CEC, where the stochastic uncertainty about some of the unknown quantities is explicitly dealt with, while all other unknown quantities are replaced by estimates obtained in a variety of ways. Let us provide some examples.

Example 6.1.1 (Multiaccess Communication)

Consider the slotted Aloha system described in Example 5.1.1. It is very difficult to obtain an optimal policy for this problem, primarily because there is no simple characterization of the conditional distribution of the state (the system backlog), given the channel transmission history. We therefore resort to a suboptimal policy. As discussed in Section 5.1, the perfect state information version of the problem admits a simple optimal policy:

$$\mu_k(x_k) = \frac{1}{x_k}, \qquad \text{for all } x_k \geq 1.$$

As a result, there is a natural partially stochastic CEC,

$$\overline{\mu}_k(I_k) = \min\left[1, \frac{1}{\overline{x}_k(I_k)}\right],$$

where $\overline{x}_k(I_k)$ is an estimate of the current packet backlog based on the entire past channel history of successes, idles, and collisions (which is I_k). Recursive estimators for generating $\overline{x}_k(I_k)$ are discussed by Mikhailov [Mik79], Hajek and van Loon [HaL82], Tsitsiklis [Tsi87], and Bertsekas and Gallager [BeG92].

Example 6.1.2 (Finite-State Systems with Imperfect State Information)

Consider the case where the system is a finite-state Markov chain under imperfect state information. The partially stochastic CEC approach is to solve the corresponding problem of perfect state information, and then use the controller thus obtained for control of the imperfectly observed system, modulo substitution of the exact state by an estimate obtained via the Viterbi algorithm described in Section 2.2.2. In particular, suppose that $\{\mu_0^p, \ldots, \mu_{N-1}^p\}$ is an optimal policy for the corresponding problem where the state is perfectly observed. Then the partially stochastic CEC, given the information vector I_k, uses the Viterbi algorithm to obtain (in real time) an estimate $\overline{x}(I_k)$ of the current state x_k, and applies the control

$$\overline{\mu}_k(I_k) = \mu_k^p\big(\overline{x}_k(I_k)\big).$$

Example 6.1.3 (Systems with Unknown Parameters)

We have been dealing so far with systems having a known system equation. In practice, however, there are many cases where the system parameters are not known exactly or change over time. One possible approach is to estimate the unknown parameters from input-output records of the system by using system identification techniques. This is a broad and important methodology, for which we refer to textbooks such as Kumar and Varaiya [KuV86], Ljung and Soderstrom [LjS83], and Ljung [Lju86]. However, system identification can be time consuming, and thus difficult to apply in an on-line control context. Furthermore, the estimation must be repeated if the parameters change.

The alternative is to formulate the stochastic control problem so that unknown parameters are dealt with directly. It can be shown that problems involving unknown system parameters can be embedded within the framework of our basic problem with imperfect state information by using state augmentation. Indeed, let the system equation be of the form

$$x_{k+1} = f_k(x_k, \theta, u_k, w_k),$$

where θ is a vector of unknown parameters with a given a priori probability distribution. We introduce an additional state variable $y_k = \theta$ and obtain a system equation of the form

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} f_k(x_k, y_k, u_k, w_k) \\ y_k \end{pmatrix}.$$

This equation can be written compactly as

$$\tilde{x}_{k+1} = \tilde{f}_k(\tilde{x}_k, u_k, w_k),$$

where $\tilde{x}_k = (x_k, y_k)$ is the new state, and $\tilde{f}_k$ is an appropriate function. The initial state is

$$\tilde{x}_0 = (x_0, \theta).$$

With a suitable reformulation of the cost function, the resulting problem becomes one that fits our usual framework.

Unfortunately, however, since y_k (i.e., θ) is unobservable, we are faced with a problem of imperfect state information even if the controller knows the state x_k exactly. Thus, typically an optimal solution cannot be found. Nonetheless, the partially stochastic CEC approach is often convenient. In particular, suppose that for a fixed parameter vector θ, we can compute the corresponding optimal policy

$$\big\{\mu_0^*(I_0, \theta), \ldots, \mu_{N-1}^*(I_{N-1}, \theta)\big\};$$

this is true for example if for a fixed θ, the problem is linear-quadratic of the type considered in Sections 4.1 and 5.2. Then a partially stochastic CEC takes the form

$$\overline{\mu}_k(I_k) = \mu_k^*(I_k, \hat{\theta}_k),$$

where $\hat{\theta}_k$ is some estimate of θ based on the information vector I_k. Thus, in this approach, the system is identified while it is being controlled. However, the estimates of the unknown parameters are used as if they were exact.

The approach of the preceding example is one of the principal methods of *adaptive control*, that is, control that adapts itself to changing values of system parameters. In the remainder of this section, we discuss some of the associated issues. Because adaptive control is somewhat disjoint from other material in the chapter, the reader may skip directly to Section 6.2.

6.1.1 Caution, Probing, and Dual Control

Suboptimal control is often guided by the qualitative nature of optimal control. It is therefore important to try to understand some of the characteristic features of the latter in the case where some of the system parameters are unknown. One of these is the need for balance between "caution" (the need for conservatism in applying control, since the system is not fully known), and "probing" (the need for aggressiveness in applying control, in order to excite the system enough to be able to identify it). These notions cannot be easily quantified, but often manifest themselves in specific control schemes. The following example provides some orientation; see also Bar-Shalom [Bar81].

Example 6.1.4 [Kum83]

Consider the linear scalar system

$$x_{k+1} = x_k + bu_k + w_k, \qquad k = 0, 1, \ldots, N-1,$$

and the quadratic terminal cost $E\{(x_N)^2\}$. Here everything is as in Section 4.1 (perfect state information) except that the control coefficient b is unknown. Instead, it is known that the a priori probability distribution of b is Gaussian with mean and variance

$$\overline{b} = E\{b\} > 0, \qquad \sigma_b^2 = E\{(b - \overline{b})^2\}.$$

Furthermore, w_k is zero mean Gaussian with variance σ_w^2 for each k.

Consider first the case where $N = 1$, so the cost is calculated to be

$$E\{(x_1)^2\} = E\{(x_0 + bu_0 + w_0)^2\} = x_0^2 + 2\overline{b}x_0u_0 + \left(\overline{b}^2 + \sigma_b^2\right)u_0^2 + \sigma_w^2.$$

The minimum over u_0 is attained at

$$u_0 = -\frac{\overline{b}}{\overline{b}^2 + \sigma_b^2}x_0,$$

and the optimal cost is verified by straightforward calculation to be

$$\frac{\sigma_b^2}{\overline{b}^2 + \sigma_b^2}x_0^2 + \sigma_w^2.$$

Therefore, the optimal control here is *cautious* in that the optimum $|u_0|$ decreases as the uncertainty in b (i.e., σ_b^2) increases.

Consider next the case where $N = 2$. The optimal cost-to-go at stage 1 is obtained by the preceding calculation:

$$J_1(I_1) = \frac{\sigma_b^2(1)}{\left(\overline{b}(1)\right)^2 + \sigma_b^2(1)}x_1^2 + \sigma_w^2, \tag{6.2}$$

where $I_1 = (x_0, u_0, x_1)$ is the information vector and

$$\overline{b}(1) = E\{b \mid I_1\}, \qquad \sigma_b^2(1) = E\left\{\left(b - \overline{b}(1)\right)^2 \mid I_1\right\}.$$

Let us focus on the term $\sigma_b^2(1)$ in the expression (6.2) for $J_1(I_1)$. We can obtain $\sigma_b^2(1)$ from the equation $x_1 = x_0 + bu_0 + w_0$ (which we view as a noise-corrupted measurement of b) and least-squares estimation theory (see Appendix E). The formula for $\sigma_b^2(1)$ will be of no further use to us, so we just state it without going into the calculation:

$$\sigma_b^2(1) = \frac{\sigma_b^2\sigma_w^2}{u_0^2\sigma_b^2 + \sigma_w^2}.$$

The salient feature of this equation is that $\sigma_b^2(1)$ is affected by the control u_0. Basically, if $|u_0|$ is small, the measurement $x_1 = x_0 + bu_0 + w_0$ is dominated by w_0 and the "signal-to-noise ratio" is small. Thus to achieve small error variance $\sigma_b^2(1)$ [which is desirable in view of Eq. (6.2)], we must apply a control u_0 that is large in absolute value. A choice of large control to enhance parameter identification is often referred to as *probing*. On the other hand, if $|u_0|$ is large, $|x_1|$ will also be large, and this is not desirable in view of Eq. (6.2). Therefore, in choosing u_0 we must strike a balance between caution (choosing a small value to keep x_1 reasonably small) and probing (choosing a large value to improve the signal-to-noise ratio and enhance estimation of b).

The tradeoff between the control objective and the parameter estimation objective is commonly referred to as *dual control*.

6.1.2 Two-Phase Control and Identifiability

An apparently reasonable form of suboptimal control in the presence of unknown parameters (cf. Example 6.1.3) is to separate the control process into two phases, a *parameter identification phase* and a *control phase*. In the first phase the unknown parameters are identified, while the control takes no account of the interim results of identification. The final parameter estimates from the first phase are then used to implement an optimal control law in the second phase. This alternation of identification and control phases may be repeated several times during any system run in order to take into account subsequent changes of the parameters.

One drawback of this approach is that information gathered during the identification phase is not used to adjust the control law until the beginning of the second phase. Furthermore, it is not always easy to determine when to terminate one phase and start the other.

A second difficulty, of a more fundamental nature, is due to the fact that the control process may make some of the unknown parameters invisible to the identification process. This is the problem of parameter *identifiability*, discussed by Ljung [Lju86], which is best explained by means of an example.

Example 6.1.5

Consider the scalar system

$$x_{k+1} = ax_k + bu_k + w_k, \qquad k = 0, 1, \ldots, N-1,$$

with the quadratic cost

$$E\left\{\sum_{k=1}^{N}(x_k)^2\right\}.$$

We assume perfect state information, so if the parameters a and b are known, this is a minimum variance control problem (cf. Section 5.3), and the optimal

control law is

$$\mu_k^*(x_k) = -\frac{a}{b}x_k.$$

Assume now that the parameters a and b are unknown, and consider the two-phase method. During the first phase the control law

$$\tilde{\mu}_k(x_k) = \gamma x_k \tag{6.3}$$

is used (γ is some scalar; for example, $\gamma = -\overline{a}/\overline{b}$, where $\overline{a}$ and $\overline{b}$ are a priori estimates of a and b, respectively). At the end of the first phase, the control law is changed to

$$\overline{\mu}_k(x_k) = -\frac{\hat{a}}{\hat{b}}x_k,$$

where $\hat{a}$ and $\hat{b}$ are the estimates obtained from the identification process. However, with the control law (6.3), the closed-loop system is

$$x_{k+1} = (a + b\gamma)x_k + w_k,$$

so the identification process can at best identify the value of $(a + b\gamma)$ but not the values of both a and b. In other words, the identification process cannot discriminate between pairs of values (a_1, b_1) and (a_2, b_2) such that $a_1 + b_1\gamma = a_2 + b_2\gamma$. Therefore, a and b are not identifiable when feedback control of the form (6.3) is applied.

One way to correct the difficulty is to add an additional known input δ_k to the control law (6.3); that is, use

$$\tilde{\mu}_k(x_k) = \gamma x_k + \delta_k.$$

Then the closed-loop system becomes

$$x_{k+1} = (a + b\gamma)x_k + b\delta_k + w_k,$$

and the knowledge of $\{x_k\}$ and $\{\delta_k\}$ makes it possible to identify $(a+b\gamma)$ and b. Given γ, one can then obtain estimates of a and b. Actually, to guarantee this in a more general context where the system is of higher dimension, the sequence $\{\delta_k\}$ must satisfy certain conditions: it must be "persistently exciting" (see for example Ljung and Soderstrom [LjS83] for further explanation of this concept).

A second possibility to bypass the identifiability problem is to change the structure of the system by artificially introducing a one-unit delay in the control feedback. Thus, instead of considering control laws of the form $\tilde{\mu}_k(x_k) = \gamma x_k$, as in Eq. (6.3), we consider controls of the form

$$u_k = \hat{\mu}_k(x_{k-1}) = \gamma x_{k-1}.$$

The closed-loop system then becomes

$$x_{k+1} = ax_k + b\gamma x_{k-1} + w_k,$$

and given γ, it is possible to identify both parameters a and b. This technique can be generalized for systems of arbitrary order, but artificially introducing a control delay makes the system less responsive to control.

6.1.3 Certainty Equivalent Control and Identifiability

At the opposite extreme of the two-phase method we have the certainty equivalent control approach, where the parameter estimates are incorporated into the control law as they are generated, and they are treated as if they were true values. In terms of the system

$$x_{k+1} = f_k(x_k, \theta, u_k, w_k)$$

considered in Example 6.1.3, suppose that, for each possible value of θ, the control law $\pi^*(\theta) = \{\mu_0^*(\cdot, \theta), \ldots, \mu_{N-1}^*(\cdot, \theta)\}$ is optimal with respect to a certain cost $J_\pi(x_0, \theta)$. Then the (suboptimal) control used at time k is

$$\hat{\mu}_k(I_k) = \mu_k^*(x_k, \hat{\theta}_k),$$

where $\hat{\theta}_k$ is an estimate of θ based on the information

$$I_k = \{x_0, x_1, \ldots, x_k, u_0, u_1, \ldots, u_{k-1}\}$$

available at time k; for example,

$$\hat{\theta}_k = E\{\theta \mid I_k\}$$

or, more likely in practice, an estimate obtained via an on-line system identification method (see [KuV86], [LjS83], [Lju86]).

One would hope that when the horizon is very long, the parameter estimates $\hat{\theta}_k$ will converge to the true value θ, so the certainty equivalent controller will become asymptotically optimal. Unfortunately, we will see that difficulties related to identifiability arise here as well.

Suppose for simplicity that the system is stationary with a priori known transition probabilities $P\{x_{k+1} \mid x_k, u_k, \theta\}$ and that the control law used is also stationary:

$$\hat{\mu}_k(I_k) = \mu^*(x_k, \hat{\theta}_k), \qquad k = 0, 1, \ldots$$

There are three systems of interest here (cf. Fig. 6.1.2):

(a) The system (perhaps falsely) believed by the controller to be true, which evolves probabilistically according to

$$P\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}_k), \hat{\theta}_k\}.$$

(b) The true closed-loop system, which evolves probabilistically according to

$$P\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}_k), \theta\}.$$

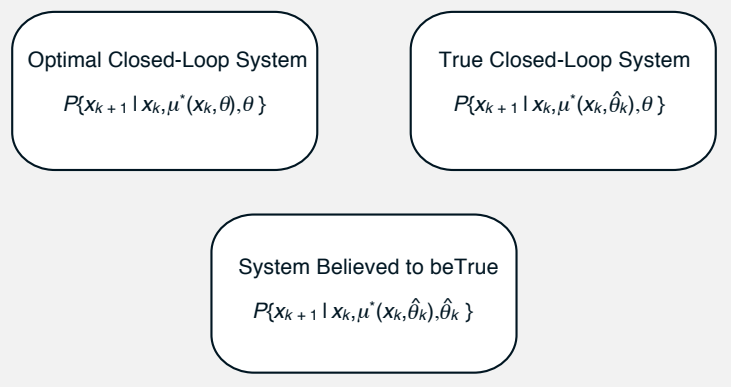

Figure 6.1.2 The three systems involved in certainty equivalent control, where θ is the true parameter and $\hat{\theta}_k$ is the parameter estimate at time k. Loss of optimality occurs when the true system differs asymptotically from the optimal closed-loop system. If the parameter estimates converge to some value $\hat{\theta}$, the true system typically becomes asymptotically equal to the system believed to be true. However, the parameter estimates need not converge, and even if they do, both systems may be different asymptotically from the optimal.

(c) The optimal closed-loop system that corresponds to the true value of the parameter, which evolves probabilistically according to

$$P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \theta), \theta\big\}.$$

For asymptotic optimality, we would like the last two systems to be equal asymptotically. This will certainly be true if $\hat{\theta}_k \to \theta$. However, it is quite possible that either

(1) $\hat{\theta}_k$ does not converge to anything, or that

(2) $\hat{\theta}_k$ converges to a parameter $\hat{\theta} \neq \theta$.

There is not much we can say about the first case, so we concentrate on the second. To see how the parameter estimates can converge to a wrong value, assume that for some $\hat{\theta} \neq \theta$ and all x_{k+1}, x_k, we have

$$P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}), \hat{\theta}\big\} = P\big\{x_{k+1} \mid x_k, \mu^*(x_k, \hat{\theta}), \theta\big\}. \tag{6.4}$$

In words, *there is a false value of parameter for which the system under closed-loop control looks exactly as if the false value were true*. Then, if the controller estimates at some time the parameter to be $\hat{\theta}$, subsequent data will tend to reinforce this erroneous estimate. As a result, a situation may develop where the identification procedure locks onto a wrong parameter value, regardless of how long information is collected. This is a difficulty with identifiability of the type discussed earlier in connection with two-phase control.

On the other hand, if the parameter estimates converge to some (possibly wrong) value, we can argue intuitively that the first two systems (believed and true) typically become equal in the limit as $k \to \infty$, since, generally, parameter estimate convergence in identification methods implies that the data obtained are asymptotically consistent with the view of the system one has based on the current estimates. However, the believed and true systems may or may not become asymptotically equal to the optimal closed-loop system. We first present two examples that illustrate how, even when the parameter estimates converge, the true closed-loop system can differ asymptotically from the optimal, thereby resulting in a certainty equivalent controller that is strictly suboptimal. We then discuss the special case of the self-tuning regulator for ARMAX models with unknown parameters, where, remarkably, it turns out that all three of the above systems are typically equal in the limit, even though the parameter estimates typically converge to false values.

Example 6.1.6 [BoV79]

Consider a two-state system with two controls u^1 and u^2. The transition probabilities depend on the control applied as well as a parameter θ, which is known to take one of two values θ^* and $\hat{\theta}$. They are as shown in Fig. 6.1.3. There is zero cost for a transition from state 1 to itself and a unit cost for all other transitions. Therefore, the optimal control at state 1 is the one that maximizes the probability of the state remaining at 1. Assume that the true parameter is θ^* and that

$$p_{11}(u^1, \hat{\theta}) > p_{11}(u^2, \hat{\theta}), \qquad p_{11}(u^1, \theta^*) < p_{11}(u^2, \theta^*).$$

Then the optimal control is u^2, but if the controller *thinks* that the true parameter is $\hat{\theta}$, it will apply u^1. Suppose also that

$$p_{11}(u^1, \hat{\theta}) = p_{11}(u^1, \theta^*).$$

Then, under u^1 the system looks identical for both values of the parameter, so if the controller estimates the parameter to be $\hat{\theta}$ and applies u^1, subsequent data will tend to reinforce the controller's belief that the true parameter is indeed $\hat{\theta}$.

More precisely, suppose that we estimate θ by selecting at each time k the value that maximizes

$$P\{\theta \mid I_k\} = \frac{P\{I_k \mid \theta\} P(\theta)}{P(I_k)},$$

where $P(\theta)$ is the a priori probability that the true parameter is θ (this is a popular estimation method). Then if $P(\hat{\theta}) > P(\theta^*)$, it can be seen, by using induction, that at each time k, the controller will estimate falsely θ to be $\hat{\theta}$ and apply the incorrect control u^1. To avoid the difficulty illustrated in this example, it has been suggested to occasionally deviate from the certainty

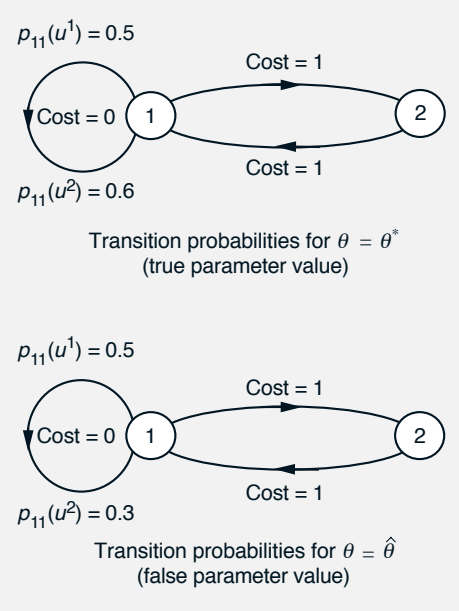

Figure 6.1.3 Transition probabilities for the two-state system of Example 6.1.6. Under the nonoptimal control u^1, the system looks identical under the true and the false values of the parameter θ.

equivalent control, applying other controls that enhance the identification of the unknown parameter (see Doshi and Shreve [DoS80], and Kumar and Lin [KuL82]). For example, by making sure that the control u^2 is used infrequently but infinitely often, we can guarantee that the correct parameter value will be identified by the preceding estimation scheme.

Example 6.1.7 [Kum83]

Consider the linear scalar system

$$x_{k+1} = ax_k + bu_k + w_k,$$

where we know that the parameters are either $(a, b) = (1, 1)$ or $(a, b) = (0, -1)$. The sequence $\{w_k\}$ is independent, stationary, zero mean, and Gaussian. The cost is quadratic of the form

$$\sum_{k=0}^{N-1} \big((x_k)^2 + 2(u_k)^2\big),$$

where N is very large, so the stationary form of the optimal control law is used (see Section 4.1). This control law can be calculated via the Riccati

equation to be

$$\mu^*(x_k) = \begin{cases} -\frac{x_k}{2} & \text{if } (a,b) = (1,1), \\ 0 & \text{if } (a,b) = (0,-1). \end{cases}$$

To estimate (a, b), we use a least-squares identification method. The value of the least-squares criterion at time k is given by

$$V_k(1,1) = \sum_{i=0}^{k-1} (x_{i+1} - x_i - u_i)^2, \qquad \text{for } (a,b) = (1,1), \tag{6.5}$$

$$V_k(0,-1) = \sum_{i=0}^{k-1} (x_{i+1} + u_i)^2, \qquad \text{for } (a,b) = (0,-1). \tag{6.6}$$

The control applied at time k is

$$u_k = \tilde{\mu}_k(I_k) = \begin{cases} -\frac{x_k}{2} & \text{if } V_k(1,1) < V_k(0,-1), \\ 0 & \text{if } V_k(1,1) > V_k(0,-1). \end{cases}$$

Suppose the true parameters are $\theta = (0, -1)$. Then the true system evolves according to

$$x_{k+1} = -u_k + w_k. \tag{6.7}$$

If at time k the controller estimates incorrectly the parameters to be $\hat{\theta} = (1, 1)$, because $V_k(\hat{\theta}) < V_k(\theta)$, the control applied will be $u_k = -x_k/2$ and the *true* closed-loop system will evolve according to

$$x_{k+1} = \frac{x_k}{2} + w_k. \tag{6.8}$$

On the other hand, the controller *thinks* (given the estimate $\hat{\theta}$) that the closed-loop system will evolve according to

$$x_{k+1} = x_k + u_k + w_k = x_k - \frac{x_k}{2} + w_k = \frac{x_k}{2} + w_k, \tag{6.9}$$

so from Eqs. (6.7) and (6.8) we see that *under the control law $u_k = -x_k/2$, the closed-loop system evolves identically for both the true and the false values of the parameters* [cf. Eq. (6.4)].

To see what can go wrong, note that if $V_k(\hat{\theta}) < V_k(\theta)$ for some k we will have, from Eqs. (6.5)-(6.9),

$$x_{k+1} + u_k = x_{k+1} - x_k - u_k,$$

so from Eqs. (6.5) and (6.6) we obtain

$$V_{k+1}(\hat{\theta}) < V_{k+1}(\theta).$$

Therefore, if $V_1(\hat{\theta}) < V_1(\theta)$, the least-squares identification method will yield the wrong estimate $\hat{\theta}$ *for every* k. To see that this can happen with positive probability, note that, since the true system is $x_{k+1} = -u_k + w_k$, we have

$$V_1(\hat{\theta}) = (x_1 - x_0 - u_0)^2 = (w_0 - x_0 - 2u_0)^2,$$
$$V_1(\theta) = (x_1 + u_0)^2 = w_0^2.$$

Therefore, the inequality $V_1(\hat{\theta}) < V_1(\theta)$ is equivalent to

$$(x_0 + 2u_0)^2 < 2w_0(x_0 + 2u_0),$$

which will occur with positive probability since w_0 is Gaussian.

The preceding examples illustrate that loss of identifiability is a serious problem that frequently arises in the context of certainty equivalent control.

6.1.4 Self-Tuning Regulators

We described earlier the nature of the identifiability issue in certainty equivalent control: under closed-loop control, incorrect parameter estimates can make the system behave as if these estimates were correct [cf. Eq. (6.4)]. As a result, the identification scheme may lock onto false parameter values. This is not necessarily bad, however, since it may happen that the control law implemented on the basis of the false parameter values is near optimal. Indeed, through a fortuitous coincidence, it turns out that *in the practically important minimum variance control formulation (Section 5.3), when the parameter estimates converge, they typically converge to false values, but the resulting control law typically converges to the optimal*. We can get an idea about this phenomenon by means of an example.

Example 6.1.8

Consider the simplest ARMAX model:

$$y_{k+1} + ay_k = bu_k + \epsilon_{k+1}.$$

The minimum variance control law when a and b are known is

$$u_k = \mu_k(I_k) = \frac{a}{b} y_k.$$

Suppose now that a and b are not known but are identified on-line by means of some scheme. The control applied is

$$u_k = \frac{\hat{a}_k}{\hat{b}_k} y_k, \tag{6.10}$$

where $\hat{a}_k$ and $\hat{b}_k$ are the estimates obtained at time k. Then the difficulty with identifiability occurs when

$$\hat{a}_k \to \hat{a}, \qquad \hat{b}_k \to \hat{b},$$

where $\hat{a}$ and $\hat{b}$ are such that the true closed-loop system given by

$$y_{k+1} + ay_k = \frac{b\hat{a}}{\hat{b}} y_k + \epsilon_{k+1}$$

coincides with the closed-loop system that the controller thinks is true on the basis of the estimates $\hat{a}$ and $\hat{b}$. This latter system is

$$y_{k+1} = \epsilon_{k+1}.$$

For these two systems to be identical, we must have

$$\frac{a}{b} = \frac{\hat{a}}{\hat{b}},$$

which means that the control law (6.10) asymptotically becomes optimal despite the fact that the asymptotic estimates $\hat{a}$ and $\hat{b}$ may be incorrect.

Example 6.1.8 can be extended to the general ARMAX model of Section 5.3 with no delay:

$$y_k + \sum_{i=1}^{m} a_i y_{k-i} = \sum_{i=1}^{m} b_i u_{k-i} + \epsilon_k + \sum_{i=1}^{m} c_i \epsilon_{k-i}.$$

If the parameter estimates converge (regardless of the identification method used and regardless of whether the limit values are correct), then a minimum variance controller *thinks* that the closed-loop system is asymptotically

$$y_k = \epsilon_k.$$

Furthermore, parameter estimate convergence intuitively means that the true closed-loop system is also asymptotically $y_k = \epsilon_k$, and this is clearly the optimal closed-loop system. Results of this type have been proved in the literature in connection with several popular methods for parameter estimation. In fact, surprisingly, in some of these results, the model adopted by the controller is allowed to be incorrect to some extent.

One issue that we have not discussed is whether the parameter estimates indeed converge. A complete analysis of this issue is quite difficult. We refer to the survey paper by Kumar [Kum85], and the textbooks by Goodwin and Sin [GoS84], Kumar and Varaiya [KuV86], and Aström and Wittenmark [AsW90] for a discussion and sources on this subject. However, extensive simulations have shown that with proper implementation, these estimates typically converge for the type of systems likely to arise in many applications.