12

Output Feedback Optimal Control with Constraints

Contributed by Tristan Perez and Hernan Haimovich

12.1 Overview

In Chapters 1 through 9 of the book (with the exception of a brief discussion on observers and integral action in Section 5.5 of Chapter 5) we considered constrained optimal control problems for systems without uncertainty, that is, with no unmodelled dynamics or disturbances, and where the full state was available for measurement. More realistically, however, it is necessary to consider control problems for systems with uncertainty. This chapter addresses some of the issues that arise in this situation. As in Chapter 9, we adopt a stochastic description of uncertainty, which associates probability distributions to the uncertain elements, that is, disturbances and initial conditions. (See Section 12.6 for references to alternative approaches to model uncertainty.)

When incomplete state information exists, a popular observer-based control strategy in the presence of stochastic disturbances is to use the certainty equivalence [CE] principle, introduced in Section 5.5 of Chapter 5 for deterministic systems. In the stochastic framework, CE consists of *estimating* the state and then using these estimates as if they were the *true* state in the control law that results if the problem were formulated as a deterministic problem (that is, without uncertainty). This strategy is motivated by the unconstrained problem with a quadratic objective function, for which CE is indeed the optimal solution (Åström 1970, Bertsekas 1976).

One of the aims of this chapter is to explore the issues that arise from the use of CE in RHC in the presence of constraints. We then turn to the obvious question about the optimality of the CE principle. We show that CE is, indeed, not optimal in general.

We also analyse the possibility of obtaining truly optimal solutions for single input linear systems with input constraints and uncertainty related to output feedback and stochastic disturbances. We first find the optimal solution for the case of horizon $N = 1$, and then we indicate the complications that arise in the case of horizon $N = 2$. Our conclusion is that, for the case of linear constrained systems, the extra effort involved in the optimal feedback policy is probably not justified in practice. Indeed, we show by example that CE can give near optimal performance. We thus advocate this approach in real applications.

12.2 Problem Statement

We consider the following time-invariant, discrete time linear system with disturbances

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{12.1}$$

where $x_k, w_k \in \mathbb{R}^n$ and $u_k, y_k, v_k \in \mathbb{R}$. The control u_k is constrained to take values in the set

$$\mathbb{U} = \{u \in \mathbb{R} : -\Delta \leq u \leq \Delta\},$$

for a given constant value $\Delta > 0$. The disturbances w_k and v_k are i.i.d. random vectors, with probability density functions (pdf) $p_w(\,\cdot\,)$ and $p_v(\,\cdot\,)$, respectively. The initial state, x_0, is characterised by a pdf $p_{x_0}(\,\cdot\,)$. We assume that the pair (A, B) is reachable and that the pair (A, C) is observable.

We further assume that, at time k, the value of the state x_k is not available to the controller. Instead, the following sets of past inputs and outputs, grouped as the *information vector* I^k, represent all the information available to the controller at the time instant k:

$$I^k = \begin{cases} \{y_0\} & \text{if} \quad k = 0, \\ \{y_0, y_1, u_0\} & \text{if} \quad k = 1, \\ \{y_0, y_1, y_2, u_0, u_1\} & \text{if} \quad k = 2, \\ \qquad \vdots & \qquad \vdots \\ \{y_0, y_1, \ldots, y_{N-1}, u_0, u_1, \ldots\, u_{N-2}\} & \text{if} \quad k = N-1. \end{cases}$$

Then, $I^k \in \mathbb{R}^{2k+1}$, and also $I^{k+1} = \{I^k, y_{k+1}, u_k\}$, where $I^k \subset I^{k+1}$.

For system (12.1), under the assumptions made, we formulate the optimisation problem:

$$\text{minimise } \mathbf{E}\Bigg\{F(x_N) + \sum_{k=0}^{N-1} L(x_k, u_k)\Bigg\}, \tag{12.2}$$

where

$$F(x_N) = x_N^{\mathrm{T}} P x_N,$$
$$L(x_k, u_k) = x_k^{\mathrm{T}} Q x_k + R u_k^2,$$

subject to the system equations (12.1) and the input constraint $u_k \in \mathbb{U}$, for $k = 0, \ldots, N-1$. Note that, under the stochastic assumptions, the expression $F(x_N) + \sum_{k=0}^{N-1} L(x_k, u_k)$ is a random variable. Hence, it is only meaningful to formulate the minimisation problem in terms of its statistics. A problem of practical interest is to minimise the expected value of this expression, which motivates the choice of the objective function in (12.2).

The result of the above minimisation problem will be a sequence of functions $\{\pi_0^{\text{OPT}}(\cdot), \pi_1^{\text{OPT}}(\cdot), \ldots, \pi_{N-1}^{\text{OPT}}(\cdot)\}$ that enable the controller to calculate the desired optimal control action depending on the information available to the controller at each time instant k, that is, $u_k^{\text{OPT}} = \pi_k^{\text{OPT}}(I^k)$. These functions also must ensure that the constraints be always satisfied. We thus make the following definition.

Definition 12.2.1 (Admissible Policies for Incomplete State Information) *A policy Π_N is a finite sequence of functions $\pi_k(\cdot) : \mathbb{R}^{2k+1} \to \mathbb{R}$ for $k = 0, 1, \ldots, N-1$, that is,*

$$\Pi_N = \big\{\pi_0(\cdot), \pi_1(\cdot), \cdots, \pi_{N-1}(\cdot)\big\}.$$

A policy Π_N is called an admissible control policy *if and only if*

$$\pi_k(I^k) \in \mathbb{U} \quad \textit{for all } I^k \in \mathbb{R}^{2k+1}, \quad \textit{for } k = 0, \ldots, N-1.$$

Further, the class of all admissible control policies will be denoted by

$$\bar{\Pi}_N = \big\{\Pi_N : \Pi_N \text{ is admissible}\big\}.$$

◦

Using the above definition, we can then state the optimal control problem of interest as follows.

Definition 12.2.2 (Stochastic Finite Horizon Optimal Control Problem) *Given the pdfs $p_{x_0}(\cdot)$, $p_w(\cdot)$ and $p_v(\cdot)$ of the initial state x_0 and the disturbances w_k and v_k, respectively, the problem considered is that of finding the control policy Π_N^{OPT}, called the* optimal control policy*, belonging to the class of all admissible control policies $\bar{\Pi}_N$, which minimises the objective function*

$$V_N(\Pi_N) = \mathop{\mathbf{E}}_{\substack{x_0, w_k, v_k \\ k=0,\ldots,N-1}} \left\{F(x_N) + \sum_{k=0}^{N-1} L(x_k, \pi_k(I^k))\right\}, \tag{12.3}$$

subject to the constraints

$$\begin{aligned} x_{k+1} &= Ax_k + B\ \pi_k(I^k) + w_k, \\ y_k &= Cx_k + v_k, \\ I^{k+1} &= \{I^k, y_{k+1}, u_k\}, \end{aligned}$$

for $k = 0, \ldots, N-1$. *In* (12.3) *the terminal state weighting* $F(\,\cdot\,)$ *and the per-stage weighting* $L(\,\cdot\,,\,\cdot\,)$ *are given by*

$$\begin{aligned} F(x_N) &= x_N^{\mathrm{T}} P x_N, \\ L(x_k, \pi_k(I^k)) &= x_k^{\mathrm{T}} Q x_k + R\pi_k^2(I^k), \end{aligned} \tag{12.4}$$

with $P > 0$, $R > 0$ *and* $Q \geq 0$.

The optimal control policy is then

$$\Pi_N^{\mathrm{OPT}} = \arg \inf_{\Pi_N \in \bar{\Pi}_N} V_N(\Pi_N),$$

with the following resulting optimal objective function value

$$V_N^{\mathrm{OPT}} = \inf_{\Pi_N \in \bar{\Pi}_N} V_N(\Pi_N). \tag{12.5}$$

◦

It is important to recognise that the optimisation problem of Definition 12.2.2 takes into account the fact that new information will be available to the controller at future time instants. This is called *closed loop optimisation*, as opposed to *open loop optimisation* where the control values $\{u_0, u_1, \ldots, u_{N-1}\}$ are selected all at once, at stage zero (Bertsekas 1976). For deterministic systems, in which there is no uncertainty, the distinction between open loop and closed loop optimisation is irrelevant, and the minimisation of the objective function over all sequences of controls or over all control policies yields the same result.

In what follows, and as in previous chapters, the matrix P in (12.4) will be taken to be the solution to the algebraic Riccati equation,

$$P = A^{\mathrm{T}} P A + Q - K^{\mathrm{T}} \bar{R} K, \tag{12.6}$$

where

$$K \triangleq \bar{R}^{-1} B^{\mathrm{T}} P A, \qquad \bar{R} \triangleq R + B^{\mathrm{T}} P B. \tag{12.7}$$

12.3 Optimal Solutions

The problem described in the previous section belongs to the class of the so-called sequential decision problems under uncertainty (Bertsekas 1976, Bertsekas 2000). A key feature of these problems is that an action taken at a particular stage affects all future stages. Thus, the control action has to be computed taking into account the future consequences of the current decision. The only general approach known to address sequential decision problems is dynamic programming.

The dynamic programming algorithm was introduced in Section 3.4 of Chapter 3 and was used in Chapters 6 and 7 to derive a closed form solution of a deterministic finite horizon optimal control problem. We next briefly show how this algorithm is used to solve the stochastic optimal control problem of Definition 12.2.2.

We define the functions

$$\begin{aligned}\tilde{L}_{N-1}(I^{N-1}, \pi_{N-1}(I^{N-1})) &= \mathbf{E}\left\{F(x_N) + L(x_{N-1}, \pi_{N-1}(I^{N-1}))\right. \\ &\qquad\qquad \left. | I^{N-1}, \pi_{N-1}(I^{N-1})\right\}, \\ \tilde{L}_k(I^k, \pi_k(I^k)) &= \mathbf{E}\left\{L(x_k, \pi_k(I^k)) | I^k, \pi_k(I^k)\right\} \quad \text{for } k = 0, \ldots, N-2.\end{aligned}$$

Then, the dynamic programming algorithm for the case of incomplete state information can be expressed via the following sequential optimisation (sub-) problems [$\mathcal{SOP}$]:

$$\begin{aligned}\mathcal{SOP}_{N-1}: \quad & J_{N-1}(I^{N-1}) = \inf_{u_{N-1} \in \mathbb{U}} \tilde{L}_{N-1}(I^{N-1}, u_{N-1}), \\ & \text{subject to:} \\ & x_N = Ax_{N-1} + Bu_{N-1} + w_{N-1},\end{aligned} \tag{12.8}$$

and, for $k = 0, \ldots, N-2$,

$$\begin{aligned}\mathcal{SOP}_k: \quad & J_k(I^k) = \inf_{u_k \in \mathbb{U}} \left[\tilde{L}_k(I^k, u_k) + \mathbf{E}\left\{J_{k+1}(I^{k+1}) | I^k, u_k\right\}\right], \\ & \text{subject to:} \\ & x_{k+1} = Ax_k + Bu_k + w_k, \\ & I^{k+1} = \{I^k, y_{k+1}, u_k\}, \\ & y_{k+1} = Cx_{k+1} + v_{k+1}.\end{aligned}$$

The dynamic programming algorithm starts at stage $N-1$ by solving $\mathcal{SOP}_{N-1}$ for all possible values of I^{N-1}. In this way, the law $\pi^{\text{OPT}}_{N-1}(\cdot)$ is obtained, in the sense that given the value of I^{N-1}, the corresponding optimal control is the value $u_{N-1} = \pi^{\text{OPT}}_{N-1}(I^{N-1})$, the minimiser of $\mathcal{SOP}_{N-1}$. The procedure then continues to solve the sub-problems $\mathcal{SOP}_{N-2}, \ldots, \mathcal{SOP}_0$ to obtain the laws $\pi^{\text{OPT}}_{N-2}(\cdot), \ldots, \pi^{\text{OPT}}_0(\cdot)$. After the last optimisation sub-problem is solved, the optimal control policy Π^{OPT}_N is obtained and the optimal objective function (see (12.5)) is

$$V^{\text{OPT}}_N = V_N(\Pi^{\text{OPT}}_N) = \mathbf{E}\{J_0(I^0)\} = \mathbf{E}\{J_0(y_0)\}.$$

12.3.1 Optimal Solution for N = 1

In the following proposition, we apply the dynamic programming algorithm to obtain the optimal solution of the problem in Definition 12.2.2 for the case $N = 1$.

Proposition 12.3.1 *For $N = 1$, the solution to the optimal control problem stated in Definition 12.2.2 is of the form $\Pi_1^{\text{OPT}} = \{\pi_0^{\text{OPT}}(\cdot)\}$, with*

$$u_0^{\text{OPT}} = \pi_0^{\text{OPT}}(I^0) = -\text{sat}_\Delta(K\,\mathbf{E}\{x_0|I^0\}) \quad \text{for all } I^0 \in \mathbb{R}, \tag{12.9}$$

where K is given in (12.7) *and* $\text{sat}_\Delta : \mathbb{R} \to \mathbb{R}$ *is the saturation function defined as*

$$\text{sat}_\Delta(z) = \begin{cases} \Delta & \text{if } z > \Delta, \\ z & \text{if } |z| \le \Delta, \\ -\Delta & \text{if } z < -\Delta. \end{cases}$$

Moreover, the last step in the dynamic programming algorithm has the value

$$\begin{aligned} J_0(I^0) = \mathbf{E}\left\{x_0^{\mathrm{T}} P x_0 | I^0\right\} + \bar{R}\Phi_\Delta(K\,\mathbf{E}\{x_0|I^0\}) \\ + \operatorname{tr}(K^{\mathrm{T}} K \operatorname{cov}\{x_0|I^0\}) + \mathbf{E}\{w_0^{\mathrm{T}} P w_0\}, \end{aligned} \tag{12.10}$$

where P and $\bar{R}$ are defined in (12.6) *and* (12.7), *respectively, and where $\Phi_\Delta : \mathbb{R} \to \mathbb{R}$ is given by*

$$\Phi_\Delta(z) = [z - \text{sat}_\Delta(z)]^2. \tag{12.11}$$

Proof. For $N = 1$, the only optimisation sub-problem to solve is $\mathcal{SOP}_0$ (see (12.8)).

$$\begin{aligned} J_0(I^0) &= \inf_{u_0 \in \mathbb{U}} \mathbf{E}\left\{F(x_1) + L(x_0, u_0) | I^0, u_0\right\} \\ &= \inf_{u_0 \in \mathbb{U}} \mathbf{E}\left\{(Ax_0 + Bu_0 + w_0)^{\mathrm{T}} P (Ax_0 + Bu_0 + w_0)\right. \\ &\quad \left. + x_0^{\mathrm{T}} Q x_0 + R u_0^2 | I^0, u_0\right\}. \end{aligned} \tag{12.12}$$

Using the fact that $\mathbf{E}\{w_0|I^0, u_0\} = E\{w_0\} = 0$ and that w_0 is neither correlated with the state x_0 nor correlated with the control u_0, (12.12) can be expressed, after distributing and grouping terms, as

$$\begin{aligned} J_0(I^0) = \mathbf{E}\{w_0^{\mathrm{T}} P w_0\} + \inf_{u_0 \in \mathbb{U}} \mathbf{E}\left\{x_0^{\mathrm{T}}(A^{\mathrm{T}} P A + Q) x_0\right. \\ \left. + 2u_0 B^{\mathrm{T}} P A x_0 (B^{\mathrm{T}} P B + R) u_0^2 | I^0, u_0\right\}. \end{aligned}$$

Further, using (12.6) and (12.7), the above becomes

$$
\begin{aligned}
J_0(I^0) = \mathbf{E}\{w_0^{\mathrm{T}} P w_0\} + \inf_{u_0 \in \mathbb{U}} \mathbf{E}\Big\{ & x_0^{\mathrm{T}} P x_0 + \bar{R}(x_0^{\mathrm{T}} K^{\mathrm{T}} K x_0 \\
& + 2u_0 K x_0 + u_0^2) | I^0, u_0 \Big\} \\
= \mathbf{E}\{w_0^{\mathrm{T}} P w_0\} + & \mathbf{E}\{x_0^{\mathrm{T}} P x_0 | I^0\} + \bar{R} \inf_{u_0 \in \mathbb{U}} \mathbf{E}\Big\{ (u_0 + K x_0)^2 | I^0, u_0 \Big\},
\end{aligned}
$$

where we have used the fact that the conditional pdf of x_0 given I^0 and u_0 is equal to the pdf where only I^0 is given. Finally, using properties of the expected value of quadratic forms (see, for example, Åström 1970) the optimisation problem to solve becomes

$$
\begin{aligned}
J_0(I^0) = \mathbf{E}\{w_0^{\mathrm{T}} P w_0\} + \mathbf{E}\{x_0^{\mathrm{T}} P x_0 | I^0\} + \mathrm{tr}(K^{\mathrm{T}} K \operatorname{cov}\{x_0 | I^0\}) \\
+ \bar{R} \inf_{u_0 \in \mathbb{U}} \left[(u_0 + K \, \mathbf{E}\{x_0 | I^0\})^2 \right].
\end{aligned}
\tag{12.13}
$$

It is clear from (12.13) that the unconstrained minimum is attained at $u_0 = -K\,\mathbf{E}\{x_0|I^0\}$. In the constrained case, equation (12.9) follows from the convexity of the quadratic function. The final value (12.10) is obtained by substituting (12.9) into (12.13). The result is then proved. □

Note that when $N = 1$ the optimal control law π_0^{OPT} depends on the information I^0 only through the *conditional expectation* $\mathbf{E}\{x_0|I^0\}$. Therefore, this conditional expectation is a sufficient statistic in this case, that is, it provides all the necessary information to implement the control.

We observe that the control law given in (12.9) is also the optimal control law for the cases in which:

- the state is measured (complete state information) and the disturbance w_k is still acting on the system;
- the state is measured and w_k is set equal to a fixed value or to its mean value (see (6.17) in Chapter 6 for the case $w_k = 0$).

Therefore, CE is optimal for horizon $N = 1$, that is, the optimal control law is the same law that would result from an associated deterministic optimal control problem in which some or all uncertain quantities were set to a fixed value.

12.3.2 Optimal Solution for N = 2

We now consider the case where the optimisation horizon is $N = 2$.

Proposition 12.3.2 *For $N = 2$, the solution to the optimal control problem stated in Definition 12.2.2 is of the form $\Pi_2^{\mathrm{OPT}} = \{\pi_0^{\mathrm{OPT}}(\,\cdot\,), \pi_1^{\mathrm{OPT}}(\,\cdot\,)\}$, with*

$$
u_1^{\mathrm{OPT}} = \pi_1^{\mathrm{OPT}}(I^1) = -\mathrm{sat}_\Delta(K\,\mathbf{E}\{x_1|I^1\}) \quad \textit{for all } I^1 \in \mathbb{R}^3,
$$

$$u_0^{\text{OPT}} = \pi_0^{\text{OPT}}(I^0) = \arg\inf_{u_0\in\mathbb{U}} \Big[\bar{R}(u_0 + K\,\mathbf{E}\{x_0|I^0\})^2 + \bar{R}\,\mathbf{E}\Big\{\Phi_\Delta(K\,\mathbf{E}\{x_1|I^1\})|I^0, u_0\Big\}\Big] \quad \textit{for all } I^0 \in \mathbb{R}, \tag{12.14}$$

where $\Phi_\Delta(\,\cdot\,)$ *is given in* (12.11).

Proof. The first step of the dynamic programming algorithm (see (12.8)) gives

$$J_1(I^1) = \inf_{u_1\in\mathbb{U}} \mathbf{E}\Big\{F(Ax_1 + B\ u_1 + w_1) + L(x_1, u_1)|I^1, u_1\Big\}.$$

We see that $J_1(I^1)$ is similar to $J_0(I^0)$ in (12.12). By comparison, we readily obtain

$$\begin{aligned}
\pi_1^{\text{OPT}}(I^1) &= -\text{sat}_\Delta(K\,\mathbf{E}\{x_1|I^1\}),\\
J_1(I^1) &= \mathbf{E}\{w_1^{\text{T}}Pw_1\} + \mathbf{E}\{x_1^{\text{T}}Px_1|I^1\} + \text{tr}(K^{\text{T}}K\,\text{cov}\{x_1|I^1\})\\
&\quad + \bar{R}\Phi_\Delta(K\,\mathbf{E}\{x_1|I^1\}).
\end{aligned}$$

The second step of the dynamic programming algorithm proceeds as follows:

$$\begin{aligned}
&J_0(I^0) = \inf_{u_0\in\mathbb{U}} \Big[\mathbf{E}\{L(x_0, u_0)|I^0, u_0\} + \mathbf{E}\left\{J_1(I^1)|I^0, u_0\right\}\Big],\\
&\text{subject to:}\\
&x_1 = Ax_0 + Bu_0 + w_0,\\
&I^1 = \{I^0, y_1, u_0\},\\
&y_1 = Cx_1 + v_1.
\end{aligned}$$

The objective function above can be written as

$$\begin{aligned}
J_0(I^0) = \inf_{u_0\in\mathbb{U}} \Big[&\mathbf{E}\{x_0^{\text{T}}Qx_0 + Ru_0^2|I^0, u_0\} + \mathbf{E}\{w_1^{\text{T}}Pw_1\}\\
&+ \mathbf{E}\{\mathbf{E}\{x_1^{\text{T}}Px_1|I^1\}|I^0, u_0\} + \text{tr}(K^{\text{T}}K\,\text{cov}\{x_1|I^1\})\\
&+ \bar{R}\,\mathbf{E}\left\{\Phi_\Delta(K\,\mathbf{E}\{x_1|I^1\})|I^0, u_0\right\}\Big].
\end{aligned} \tag{12.15}$$

Since $\{I^0, u_0\} \subset I^1$, using the properties of successive conditioning (Ash and Doléans-Dade 2000), we can express the third term inside the inf in (12.15) as

$$\begin{aligned}
\mathbf{E}\{\mathbf{E}\{x_1^{\text{T}}Px_1|I^1\}|I^0, u_0\} &= \mathbf{E}\{x_1^{\text{T}}Px_1|I^0, u_0\}\\
&= \mathbf{E}\Big\{(Ax_0 + Bu_0 + w_0)^{\text{T}}P(Ax_0\\
&\qquad + Bu_0 + w_0)|I^0, u_0\Big\}.
\end{aligned}$$

Using this, expression (12.15) becomes

$$
\begin{aligned}
J_0(I^0) = \inf_{u_0 \in \mathbb{U}} \Big[& \mathbf{E}\{x_0^{\mathrm{T}} Q x_0 + R u_0^2 + (Ax_0 + Bu_0 + w_0)^{\mathrm{T}} P (Ax_0 \\
& + Bu_0 + w_0) | I^0, u_0\} + \mathbf{E}\{w_1^{\mathrm{T}} P w_1\} + \mathrm{tr}(K^{\mathrm{T}} K \,\mathrm{cov}\{x_1 | I^1\}) \\
& + \bar{R}\, \mathbf{E}\left\{ \Phi_\Delta (K\, \mathbf{E}\{x_1 | I^1\}) | I^0, u_0 \right\} \Big]. \qquad (12.16)
\end{aligned}
$$

Note that the first part of (12.16) is identical to (12.12). Therefore, using (12.13), expression (12.16) can be written as

$$
\begin{aligned}
J_0(I^0) = \mathbf{E}\{x_0^{\mathrm{T}} P x_0 | I^0\} & + \sum_{j=0}^{1} \left[\mathrm{tr}(K^{\mathrm{T}} K \,\mathrm{cov}\{x_j | I^j\}) + \mathbf{E}\{w_j^{\mathrm{T}} P w_j\} \right] \\
& + \inf_{u_0 \in \mathbb{U}} \Big[\bar{R}(u_0^2 + K\, \mathbf{E}\{x_0 | I^0\})^2 + \bar{R}\, \mathbf{E}\left\{ \Phi_\Delta (K\, \mathbf{E}\{x_1 | I^1\}) | I^0, u_0 \right\} \Big], \\
& \qquad (12.17)
\end{aligned}
$$

where $\mathrm{tr}(K^{\mathrm{T}} K \,\mathrm{cov}\{x_1 | I^1\})$ has been left out of the minimisation because it is not affected by u_0 due to the linearity of the system equations (Bertsekas 1987, Bertsekas 2000). By considering only the terms that are affected by u_0, we find the result given in (12.14). □

To obtain an explicit form for π_0^{OPT}, we would need to express $\mathbf{E}\{x_1 | I^1\} = \mathbf{E}\{x_1 | I^0, y_1, u_0\}$ explicitly as a function of I^0, u_0 and y_1. The optimal law $\pi_0^{\mathrm{OPT}}(\,\cdot\,)$ depends on I^0 not only through $\mathbf{E}\{x_0 | I^0\}$, as was the case for $N = 1$. Indeed, Haimovich, Perez and Goodwin (2003) have shown that, even for Gaussian disturbances, when input constraints are present, the optimal control law $\pi_0^{\mathrm{OPT}}(\,\cdot\,)$ depends also on $\mathrm{cov}\{x_0 | I^0\}$.

To calculate $\mathbf{E}\{x_1 | I^1\}$, we need to find the conditional pdf $p_{x_1|I^1}(\,\cdot\,|I^1)$. At any time instant k, the conditional pdfs $p_{x_k|I^k}(\,\cdot\,|I^k)$ satisfy the Chapman–Kolmogorov equation and the observation update equation (see Section 9.8 in Chapter 9):

Time update

$$
\begin{aligned}
p_{x_k|I^{k-1},u_{k-1}}(x_k | I^{k-1}, u_{k-1}) = & \int_{\mathbb{R}^n} p_{x_k|x_{k-1},u_{k-1}}(x_k | x_{k-1}, u_{k-1}) \\
& \times p_{x_{k-1}|I^{k-1},u_{k-1}}(x_{k-1} | I^{k-1}, u_{k-1}) dx_{k-1}, \quad (12.18)
\end{aligned}
$$

Observation update

$$
\begin{aligned}
p_{x_k|I^k}(x_k | I^k) & = p_{x_k|I^{k-1},y_k,u_{k-1}}(x_k | I^{k-1}, y_k, u_{k-1}) \\
& = \frac{p_{y_k|x_k}(y_k | x_k) p_{x_k|I^{k-1},u_{k-1}}(x_k | I^{k-1}, u_{k-1})}{p_{y_k|I^{k-1},u_{k-1}}(y_k | I^{k-1}, u_{k-1})}, \qquad (12.19)
\end{aligned}
$$

where

$$p_{y_k|I^{k-1},u_{k-1}}(y_k|I^{k-1},u_{k-1}) = \int_{\mathbb{R}^n} p_{y_k|x_k}(y_k|x_k) \times p_{x_k|I^{k-1},u_{k-1}}(x_k|I^{k-1},u_{k-1})dx_k.$$

Remark 12.3.1. In general, depending on the pdfs of the initial state and the disturbances, it may be very difficult or even impossible to obtain an explicit form for the conditional pdfs that satisfy the recursion given by (12.18) and (12.19). If the pdfs of the initial state and the disturbances are Gaussian, however, all the conditional densities that satisfy (12.18) and (12.19) are also Gaussian. In this particular case, (12.18) and (12.19) lead to the well-known Kalman filter algorithm (see Section 9.6 in Chapter 9). The latter is a recursive algorithm in terms of the (conditional) expectation and covariance, which completely define any Gaussian pdf. ◦

Due to the way the information enters the conditional pdfs, it is, in general, very difficult to obtain an explicit form for the optimal control. On the other hand, even if the recursion given by (12.18) and (12.19) can be found explicitly, the implementation of such optimal control may also be complicated and computationally demanding. We illustrate this point by suggesting a way of implementing the optimal controller in such a case.

Let us first discretise the set $\mathbb{U}$ of admissible control values, and suppose that the discretised set $\mathbb{U}_d = \{u_{0i} \in \mathbb{U} : i = 1, 2, \ldots, r\}$ contains a finite number r of elements. We then approximate the optimal control as

$$u_0^{\text{OPT}} \approx \arg\inf_{u_0 \in \mathbb{U}_d} \Big[\bar{R}(u_0 + K\,\mathbf{E}\{x_0|I^0\})^2 + \bar{R}\,\mathbf{E}\{\Phi_\Delta(K\,\mathbf{E}\{x_1|I^1\})|I^0, u_0\}\Big]. \tag{12.20}$$

Hence, to solve the above minimisation, only a finite number of values of u_0 have to be considered for a given $I^0 = y_0$. To do this, given the measurement $I^0 = y_0$ and for every value u_{0i}, $i = 1, \ldots, r$, of u_0, we can evaluate the expression between square brackets in (12.20) in the following way:

(i) Given y_0, we evaluate the function $W_2(\cdot)$ given by

$$W_2(y_0) = \mathbf{E}\{x_0|I^0\} = \mathbf{E}\{x_0|y_0\} = \int_{\mathbb{R}^n} x_0\ p_{x_0|y_0}(x_0|y_0)dx_0. \tag{12.21}$$

(ii) Given y_0, $p_v(\cdot)$, $p_w(\cdot)$ and $p_{x_0}(\cdot)$, we obtain $p_{y_0|x_0}(\cdot|x_0)$ and $p_{x_1|y_0,y_1,u_0}(\cdot|y_0,y_1,u_0)$ in explicit form (as assumed) using the recursion (12.18) and (12.19) together with the system and measurement equations (12.1).

(iii) Using $p_{x_1|y_0,y_1,u_0}(\cdot|y_0,y_1,u_0)$, the expectation $\mathbf{E}\{x_1|y_0,y_1,u_0\}$ can be written as

$$\begin{aligned} h(y_0,y_1,u_0) &= \mathbf{E}\{x_1|y_0,y_1,u_0\} \\ &= \int_{\mathbb{R}^n} x_1 p_{x_1|y_0,y_1,u_0}(x_1|y_0,y_1,u_0)dx_1, \end{aligned} \tag{12.22}$$

which may not be expressible in explicit form even if (12.18) and (12.19) are so expressed. However, it may be evaluated numerically if y_0, y_1 and u_0 are given.

(iv) Using $p_{x_0|y_0}(\cdot|y_0)$, $p_v(\cdot)$, together with the measurement equation, we can obtain $p_{y_1|y_0,u_0}(\cdot|y_0,u_0)$. We can now express the expectation $\mathbf{E}\left\{\Phi_\Delta\left(Kh(y_0,y_1,u_0)\right)|I^0,u_0\right\}$ as

$$\begin{aligned} W_1(y_0,u_0) &= \mathbf{E}\left\{\Phi_\Delta\left(Kh(y_0,y_1,u_0)\right)|I^0,u_0\right\} \\ &= \int_{\mathbb{R}} \Phi_\Delta\left(Kh(y_0,y_1,u_0)\right) p_{y_1|y_0,u_0}(y_1|y_0,u_0)dy_1. \end{aligned} \tag{12.23}$$

Note that, in order to calculate this integral numerically, the function $h(y_0,y_1,u_0)$ has to be evaluated for different values of y_1, even if u_0 and y_0 are given.

To find the value of the objective function achieved by *one* of the r values u_{0i}, expressions (12.21), (12.22) and (12.23) may need further discretisations (for x_0, x_1 and y_1, respectively).

From the previous comments, it is evident that the approximation of the optimal solution can be very computationally demanding depending on the discretisations performed. Note that in the above steps all the pdfs are assumed known in explicit form so that the integrals can be evaluated. As already mentioned in Remark 12.3.1, this may not always be possible.

As an alternative approach to brute force discretisations, we could use Markov chain Monte Carlo [MCMC] methods (Robert and Casella 1999). These methods approximate continuous pdfs by discrete ones by drawing samples from the pdfs in question or from other approximations. However, save for some very particular cases, the exponential growth in the number of computations as the optimisation horizon is increased seems to be unavoidable. We observe, in passing, that the application of MCMC methods to the recursion given by (12.18) and (12.19) gives rise to a special case of the, so-called, particle filters (Doucet et al. 2001).

The above discussion suggests that, not only does it seem impossible to analytically proceed with the optimisation for horizons greater than two but also the implementation of the optimal law (even for $N = 2$) appears to be quite intricate and computationally burdensome. This leads us to consider suboptimal solutions. In the next section, we analyse two alternative suboptimal strategies.

12.4 Suboptimal Strategies

12.4.1 Certainty Equivalent Control

As mentioned before, certainty equivalent control [CEC] uses the control law obtained as the solution of an associated *deterministic* control problem derived from the original problem by removing all uncertainty. Specifically, the associated problem is derived by setting the disturbance w_k to a fixed typical value (for example, $\bar{w} = \mathbf{E}\{w_k\}$) and by also assuming perfect state information. The resulting control law is a function of the true state. Then, the control is implemented using some estimate of the state $\hat{x}(I^k)$ in place of the true state.

For our problem, we first obtain the optimal policy for the deterministic problem

$$\Pi_N^{\text{DET}} = \left\{\pi_0^{\text{DET}}(\,\cdot\,), \ldots, \pi_{N-1}^{\text{DET}}(\,\cdot\,)\right\}, \tag{12.24}$$

where $\pi_k^{\text{DET}} : \mathbb{R}^n \to \mathbb{R}$ for $k = 0, 1, \ldots, N-1$. Then, the CEC evaluates the deterministic laws at the estimate of the state, that is,

$$u_k^{\text{CE}} = \pi_k^{\text{DET}}\left(\hat{x}(I^k)\right). \tag{12.25}$$

As we saw in Section 6.2 of Chapter 6, the associated deterministic problem for linear systems with a quadratic objective function is an example of a case where the control policy can be obtained explicitly for any finite optimisation horizon. The following example illustrates this for an optimisation horizon $N = 2$.

Example 12.4.1 (Closed Loop CEC). For $N = 2$, the deterministic policy $\Pi_2^{\text{DET}} = \{\pi_0^{\text{DET}}(\,\cdot\,), \pi_1^{\text{DET}}(\,\cdot\,)\}$ is given by (see Theorem 6.2.1 in Chapter 6):

$$\pi_1^{\text{DET}}(x) = -\text{sat}_\Delta(Kx) \quad \text{for all } x \in \mathbb{R}^n$$

$$\pi_0^{\text{DET}}(x) = \begin{cases} -\text{sat}_\Delta(Gx + h) & \text{if } x \in \mathbb{Z}^-, \\ -\text{sat}_\Delta(Kx) & \text{if } x \in \mathbb{Z}, \\ -\text{sat}_\Delta(Gx - h) & \text{if } x \in \mathbb{Z}^+. \end{cases}$$

K is given by (12.7) and

$$G = \frac{K + KBKA}{1 + (KB)^2}, \qquad h = \frac{KB}{1 + (KB)^2}\Delta.$$

The sets $\mathbb{Z}^-, \mathbb{Z}, \mathbb{Z}^+$ form a partition of $\mathbb{R}^n$, and are given by

$$\begin{aligned} \mathbb{Z}^- &= \{x : K(A - BK)x < -\Delta\}, \\ \mathbb{Z} &= \{x : |K(A - BK)x| \leq \Delta\}, \\ \mathbb{Z}^+ &= \{x : K(A - BK)x > \Delta\}. \end{aligned}$$

Therefore, a closed loop CEC applies the controls

$$\begin{aligned} u_0^{\text{CE}} &= \pi_0^{\text{DET}}\left(\hat{x}(I^0)\right), \\ u_1^{\text{CE}} &= \pi_1^{\text{DET}}\left(\hat{x}(I^1)\right), \end{aligned}$$

where the estimate $\hat{x}(I^k)$ can be provided, for example, by the Kalman filter.

◦

12.4.2 Partially Stochastic CEC

This variant of CEC uses the control law obtained as the solution to an associated problem that assumes perfect state information but takes stochastic disturbances into account. To actually implement the controller, the value of the state is replaced by its estimate $\hat{x}_k(I^k)$.

In our case, given a partially stochastic CEC [PS–CEC] admissible policy

$$\Lambda_N = \big\{\lambda_0(\,\cdot\,), \ldots, \lambda_{N-1}(\,\cdot\,)\big\}, \tag{12.26}$$

that is, a sequence of admissible control laws $\lambda_k(\,\cdot\,) : \mathbb{R}^n \to \mathbb{U}$ that map the (estimates of the) states into admissible control actions, the PS–CEC solves the following perfect state information problem.

Definition 12.4.1 (PS–CEC Optimal Control Problem) *Assuming that the state $\hat{x}_k$ will be available to the controller at time instant k to calculate the control, and given the pdf $p_w(\,\cdot\,)$ of the disturbances w_k, find the admissible control policy $\Lambda_N^{\text{OPT}} = \big\{\lambda_0^{\text{OPT}}(\,\cdot\,), \ldots, \lambda_{N-1}^{\text{OPT}}(\,\cdot\,)\big\}$ that minimises the objective function*

$$\hat{V}_N(\Lambda_N) = \underset{\substack{w_k \\ k=0,\ldots,N-1}}{\mathbf{E}} \left\{ F(\hat{x}_N) + \sum_{k=0}^{N-1} L(\hat{x}_k, \lambda_k(\hat{x}_k)) \right\},$$

subject to $\hat{x}_{k+1} = A\hat{x}_k + B\lambda_k(\hat{x}_k) + w_k$ for $k = 0, \ldots, N-1$. ◦

The optimal control policy for perfect state information thus found will be used, as in CEC, to calculate the control action based on the estimate $\hat{x}_k$ provided by the estimator; that is,

$$u_k = \lambda_k^{\text{OPT}}(\hat{x}_k).$$

Next, we apply this suboptimal strategy to the problem of interest for horizons 1 and 2.

PS–CEC for N = 1.

Using the dynamic programming algorithm, we have

$$\hat{J}(\hat{x}_0) = \inf_{u_0 \in \mathbb{U}} \mathbf{E} \left\{ \hat{x}_1^{\text{T}} P \hat{x}_1 + \hat{x}_0^{\text{T}} Q \hat{x}_0 + R u_0^2 | \hat{x}_0, u_0 \right\}.$$

As with the true optimal solution for $N = 1$, the PS–CEC optimal control has the form

$$\begin{aligned} \hat{u}_0^{\text{OPT}} &= \lambda_0^{\text{OPT}}(\hat{x}_0) = -\text{sat}_\Delta(K\hat{x}_0), \\ \hat{J}_0(\hat{x}_0) &= \hat{x}_0^{\text{T}} P \hat{x}_0 + \bar{R}\Phi_\Delta(K\hat{x}_0) + \mathbf{E}\{w_0^{\text{T}} P w_0\}. \end{aligned}$$

We can see that if $\hat{x}_0 = \mathbf{E}\{x_0 | I^0\}$ then the PS–CEC for $N = 1$ coincides with the optimal solution.

PS–CEC for N = 2.

The first step of the dynamic programming algorithm yields

$$\begin{aligned}\hat{u}_1^{\text{OPT}} &= \lambda_1^{\text{OPT}}(\hat{x}_1) = -\text{sat}_\Delta(K\hat{x}_1),\\ \hat{J}_1(\hat{x}_1) &= \hat{x}_1^{\text{T}} P \hat{x}_1 + \bar{R}\Phi_\Delta(K\hat{x}_1) + \mathbf{E}\{w_1^{\text{T}} P w_1\}.\end{aligned}$$

For the second step, we have, after some algebra, that

$$\begin{aligned}&\hat{J}_0(\hat{x}_0) = \inf_{u_0\in\mathbb{U}} \left[\mathbf{E}\{L(\hat{x}_0,u_0) + \hat{J}_1(\hat{x}_1)|\hat{x}_0,u_0\}\right],\\ &\text{subject to:}\\ &\hat{x}_1 = A\hat{x}_0 + Bu_0 + w_0,\end{aligned}$$

$$\hat{u}_0^{\text{OPT}} = \arg\inf_{u_0\in\mathbb{U}} \left[\bar{R}(u_0+K\hat{x}_0)^2 + \bar{R}\,\mathbf{E}\{\Phi_\Delta[K(A\hat{x}_0+Bu_0+w_0)]|\hat{x}_0,u_0\}\right]. \quad (12.27)$$

Comparing $\hat{u}_0^{\text{OPT}}$ with expression (12.14) for the optimal control, we can appreciate that, given $\hat{x}_0$, even if $\mathbf{E}\left\{\Phi_\Delta[K(A\hat{x}_0 + Bu_0 + w_0)]|\hat{x}_0, u_0\right\}$ cannot be found in explicit form as a function of u_0, the numerical implementation of this suboptimal control action is much less computationally demanding than its optimal counterpart.

12.5 Simulation Examples

In this section we compare the performance of the suboptimal strategies CEC and PS–CEC by means of simulation examples. The performance is assessed by computing the achieved value of the objective function. The objective function is defined as the expected value of a random variable, which is a quadratic function of the states and controls in our case. Hence, a comparison between the values of the objective function incurred by using different policies is only meaningful in terms of these expected values. To numerically compute values of the objective function for a given control policy, different realisations of the initial state plus process and measurement disturbances have to be obtained and a corresponding realisation of the objective function evaluated. Then, the expected value can be approximated by averaging over the different realisations.

The following examples are simulated for the system:

$$A = \begin{bmatrix} 0.9713 & 0.2189 \\ -0.2189 & 0.7524 \end{bmatrix}, \qquad B = \begin{bmatrix} 0.0287 \\ 0.2189 \end{bmatrix}, \qquad C = [0.3700 \quad 0.0600]. \quad (12.28)$$

The disturbances w_k are assumed to have a uniform distribution with support on $[-0.5, 0.5] \times [-1, 1]$ and likewise for v_k with support on $[-0.1, 0.1]$. The initial state x_0 is assumed to have a Gaussian distribution with zero mean and covariance diag$\{300^{-1}, 300^{-1}\}$. A Kalman filter was implemented to provide the state estimates needed. Although this estimator is not the optimal one in this case because the disturbances are not Gaussian, it yields the best linear unbiased estimator for the state. The parameters for the Kalman filter were chosen as the true mean and covariance of the corresponding variables in the system. The saturation limit of the control was taken as $\Delta = 1$. The optimisation horizon is in both cases $N = 2$.

For PS–CEC, we discretise the set $\mathbb{U}$ so that only 500 values are considered, and the expected value in (12.27) is approximated by taking 300 samples of the pdf $p_w(\cdot)$ for every possible value of u_0 in the discretised set. For CEC, we implement the policy given in Example 12.4.1.

We simulated the closed loop system over two time instants and repeated the simulation a large number of times (between 2000 and 8000). For each simulation, a different realisation of the disturbances and the initial state was used. A realisation of the objective function was calculated for every simulation run for each one of the control policies applied (PS–CEC and CEC). The sample average of the objective function values achieved by each policy was computed, and the difference between them was always found to be less than 0.1%.

Although the examples are based on a simple simulated model, the comparison between the objective function values for the two control policies seems to indicate that the trade-off between better performance and computational complexity favours the CEC implementation over the PS–CEC.

It would be of interest, from a practical standpoint, to extend the optimisation horizon beyond $N = 2$. However, as we explained in a previous section and observed in the examples, due to computational issues this becomes very difficult. In order to achieve this extension, one is led to conclude that CEC may be, at this point, the only way forward.

Of course, the ultimate test for the suboptimal strategies would be to contrast them with the optimal one. It would be expected that, in this case, an appreciable difference in the objective function values may be obtained due to the fact that the optimal strategy takes into account the process and measurement disturbances in a unified manner, as opposed to the above mentioned suboptimal strategies, which use estimates provided by an estimator as if they were the true state.

12.6 Further Reading

For complete list of references cited, see References section at the end of book.

General

For more details on general sequential decision problems under uncertainty and the use of dynamic programming, the reader is referred to Bellman (1957), Bertsekas (1976) and Bertsekas (2000).

Section 12.3

The use of CE in RHC, due to its simplicity, has been advocated in the literature (Muske and Rawlings 1993) and reported in a number of applications (see, for example, Angeli, Mosca and Casavola 2000, Marquis and Broustail 1988, Perez, Goodwin and Tzeng 2000).

For RHC literature for uncertain systems using a stochastic uncertainty description, see Haimovich, Perez and Goodwin (2003) and Perez, Haimovich and Goodwin (2004) (on which this chapter is based). Also, in Filatov and Unbehauen (1995) output-feedback predictive control of nonlinear systems with uncertain parameters is addressed. The control is assumed unconstrained and only suboptimal solutions are considered. Batina, Stoorvogel and Weiland (2001) consider the RHC problem for the case of state feedback, input constraints and scalar disturbances. The optimal solution is approximated via a randomised algorithm (Monte Carlo sampling). Examples for an optimisation horizon of length 1 are presented. In Batina, Stoorvogel and Weiland (2002), the authors extend their previous result to the state constrained case.

An alternative approach to model uncertainty is via a set-membership description, which only gives information regarding the sets in which the uncertain elements take values. When addressing uncertain systems, the RHC literature has somewhat favoured the set-membership description; see, for example, Shamma and Tu (1998) and Lee and Kouvaritakis (2001). For example, Shamma and Tu (1998) propose an observer-based strategy that assumes unknown but bounded disturbances, and generates a set of possible states based on past input and output information. Then, to each estimated state the strategy associates a set of control values that meet the constraint requirements. The actual control applied to the system is selected to belong to the intersection of all the control value sets. As another example, Lee and Kouvaritakis (2001) present an extension of the dual-mode paradigm of Mayne and Michalska (1993), in which invariant sets of estimation errors are used for the case of unknown-but-bounded measurement noise and disturbances.