Chapter 4

Linear-Quadratic Optimal Control: Full-State Feedback

[1] Linear quadratic optimization is a basic method for designing controllers for linear (and often nonlinear) dynamical systems and is actually frequently used in practice, for example in aerospace applications. Moreover it also has interpretations in terms of "classical control" notions, such as disturbance rejection, phase and gain margin, etc. (topics we will not cover, but a reference is [AM07]). In estimation, it leads to the Kalman filter, which we will encounter in chapter 5. In this chapter however, we continue with our investigation of problems where the full state of the system is observable, and describe the solution of the Linear Quadratic Regulator (LQR) problem.

Some references: [Ber07, section 4.1], Slides from Stephen Boyd's EE363 class [Boya], [Ath71, AM07] (mostly continuous-time).

4.1 Model

We consider in this chapter a system with linear dynamics

$$x_{k+1} = A_k x_k + B_k u_k + w_k,$$

where x_k and u_k are real vectors of dimension n and m respectively, the states and controls u_k are *unconstrained* ($\mathsf{X}_k = \mathbb{R}^n$, $\mathsf{U}_k(x_k) = \mathbb{R}^m$ for all k), and the disturbances w_k are independent random vectors (independent of x_k and u_k), with a known probability distribution with zero mean and finite second moment matrix $W_k = E[w_k w_k^T]$. The matrices $A_k \in \mathbb{R}^{n \times n}$ and $B_k \in \mathbb{R}^{n \times m}$ are called the dynamics and input matrices respectively. Such a model sometimes comes from the discretization of a continuous-time system, as in example 1.1.2 of chapter 1, but sometimes the discrete-time nature of the model is more intrinsic, for example in production planning or inventory control problems.

[1] This version: September 19 2009

We wish to minimize the quadratic (finite-horizon) cost

$$E\left\{\sum_{k=0}^{N-1}(x_k^T Q_k x_k + u_k^T R_k u_k) + x_N^T Q_N x_N\right\}, \tag{4.1}$$

where the expectation is with respect to the disturbances $w_0, \ldots, w_{N-1}$. We assume that the matrices $Q_k, k = 0, \ldots, N$ (the stage cost and final stage cost matrices) are positive semidefinite (denoted $Q_k \succeq 0$) and the matrices $R_k, k = 0, \ldots, N-1$ (the input cost matrices) are positive definite ($R_k \succ 0$). The signification of this cost function is that we wish to bring the state close to the origin $x = 0$ (regulation problem) using the term $x_k^T Q_k x_k$. Linear systems theory tells us that in the absence of disturbances we can always bring the state to 0 in at most n steps (recall that n is the dimension of the state space), but this might require large control inputs. The additional terms $u_k^T Q_k u_k$ penalizes large inputs and thus seeks to obtain more realistic designs since systems are always subject to input constraints. This problem, often without the disturbances w_k, is called the Linear-Quadratic Regulator (LQR) problem. There are numerous variations and complications of this basic version of the problem (see e.g. [AM07]), such as adding cross-terms in the cost function, state-dependent noise statistics, random dynamics and input matrices ([Ber07, p. 159]; this is useful in recent work on networked control), etc. We will explore some of them in the problems.

Finally, it is important to note that the choice of the weighting matrices Q_k and R_k in the cost function is not trivial. Essentially, the LQ formulation translates the difficulty of the classical control problems, where specifications are typically given in terms of settling times, slew rates, stability and phase margins, and other specifications on input and output signals, into the choice of the coefficients of the cost matrices. Once these matrices are chosen, the design of the optimum controller is automatic (in the sense that you can call a MATLAB function to do it for you). In practice, an iterative procedure is typically followed where the properties of the synthesized controller are tested with respect to the given specifications, the cost matrix coefficients are readjusted depending on the observed performance, the new design is retested, and so on. There are also guidelines to understand the impact of the choice of cost coefficients on classical specifications [AM07]. Testing is necessary in any case to verify that differences between the real-world system and the mathematical model specified by the matrices A_k, B_k does not lead to an excessive drop in performance.

Remark. We consider the addition of hard state and control constraints in chapter 11 on model predictive control. The unconstrained linear quadratic problem, even if less realistic, is still widely used, if nothing else because it yields an analytical formula for the control law.

4.2 Solution of the Linear Quadratic Regulator

Intuition for the Solution

As usual, we write $J_0^*(x_0)$ for the value function (cost function of the optimal policy). Suppose for a moment that there are no disturbances, so that the problem is deterministic and we can solve it as an optimization problem to design an open-loop control policy, see chapter 2 and problem B.1.7. Then we can view $J_0^*(x_0)$ as the optimum value of a (linearly constrained) quadratic program for which x_0 is a parameter on the right-hand side of the linear constraints. Hence we know from standard convex optimization theory that J_0^* is a *convex* function of x_0. In fact, since $J_0^*(x_0)$ can be seen as the partial minimization of a quadratic function of x, u, over all variables except those in x_0, we have, essentially by using the Schur complement [BV04, appendix A.5.5], that $J_0^*(x_0)$ is a *convex quadratic* function of x_0. It turns out that in the presence of stochastic disturbances $\{w_k\}_{0\leq k\leq N-1}$, we simply need to add a constant term to the deterministic solution for $J_0^*(x_0)$.

Since we will use the Schur complement in a moment, let us recall the result here. Consider the minimization of the quadratic function

$$\min_u \ x^T Ax + u^T Cu + 2x^T Bu = \min_u \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

over some of its variables u, under the assumption $C \succ 0$. The solution is $u = -C^{-1}B^T x$ and the minimum value is equal to $x^T Sx$, where $S = A - BC^{-1}B^T$ is called the Schur complement of C in the matrix

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}.$$

If $X \succeq 0$, we know from convex analysis that the minimum value is a convex function of x, so $S \succeq 0$.

Exercise 11. Rederive the formula $S = A - BC^{-1}B^T$.

Remark. Under additional assumptions, the problem has sometimes a solution even if C is singular. For example, if $C \succeq 0$ and $Bx \in \text{Im}(C)$, then the same result holds with C^{-1} replaced by $C^\dagger$, the pseudo-inverse of C. If $Bx \notin \text{Im}(C)$ or $C \not\succeq 0$, the problem is unbounded. Note also that we have the following converse for the result above. If $C \succ 0$ and $S \succeq 0$, then $X \succeq 0$, because in this case the minimization over x and u (which can be performed by first minimizing over u and then over x) yields a finite value.

Solution using Backward Induction

In this section, we compute the optimal cost and optimal policy for the LQR problem. The DP algorithm gives

$$\begin{aligned} J_N^*(x_N) &= x_N^T Q_N x_N \\ J_k^*(x_k) &= \min_{u_k} E_{w_k} \left\{ x_k^T Q_k x_k + u_k^T R_k u_k + J_{k+1}^*(A_k x_k + B_k u_k + w_k) \right\}. \end{aligned}$$

Similarly to chapter 3, we show the key property of the value function using backward induction. In this case, we show that $J_k^*(x_k) = x_k^T P_k x_k + q_k$, for some matrix $P_k \succeq 0$ and constant $q_k \geq 0$. That is, the cost function is *convex quadratic* in the state (with no linear terms). The induction hypothesis is true for $k = N$, with $P_N = Q_N$ and $q_N = 0$. Assuming it is true for index $k+1$, we show that it is then true for index k. In the DP recursion, we obtain

$$\begin{aligned} J^*(x_k) = \min_{u_k} \{ & x_k^T Q_k x_k + u_k^T R_k u_k \\ & + E_{w_k}[(A_k x_k + B_k u_k + w_k)^T P_{k+1}(A_k x_k + B_k u_k + w_k)] \} + q_{k+1}. \end{aligned}$$

When evaluating the expectation, all terms that are linear in w_k vanish because w_k is zero mean. So we obtain

$$\begin{aligned} J^*(x_k) = \min_{u_k} \{ & x_k^T (A_k^T P_{k+1} A_k + Q_k) x_k + u_k^T (B_k^T P_{k+1} B_k + R_k) u_k \\ & + 2x^T (A_k^T P_{k+1} B_k) u_k \} + E_{w_k}[w_k^T P_{k+1} w_k]. \end{aligned}$$

We can rewrite the last term (it's useful to remember this trick for other optimization problems, although here it's not adding much)

$$E_{w_k}[w_k^T P_{k+1} w_k] = E_w[\mathrm{Tr}(P_{k+1} w_k w_k^T)] = \mathrm{Tr}(P_{k+1} W_k).$$

Now the minimization over u_k in the first term corresponds to our result in the previous paragraph on the partial minimization of a quadratic function, with matrix

$$X = \begin{bmatrix} A_k^T P_{k+1} A_k + Q_k & A_k^T P_{k+1} B_k \\ B_k^T P_{k+1} A_k & B_k^T P_{k+1} B_k + R_k \end{bmatrix}.$$

Because of our assumption that $R_k \succ 0$, the matrix $B_k^T P_{k+1} B_k + R_k$ is positive definite. The solution is then a control law that is *linear in the state*

$$u_k = K_k x_k, \text{with } K_k = -(B_k^T P_{k+1} B_k + R_k)^{-1} B_k^T P_{k+1} A_k, \tag{4.2}$$

and using the Schur complement result, we obtain directly that $J_k^*(x_k) = x_k^T P_k x_k + q_k$, with

$$P_k = A_k^T P_{k+1} A_k + Q_k - A_k^T P_{k+1} B_k (B_k^T P_{k+1} B_k + R_k)^{-1} B_k^T P_{k+1} A_k \tag{4.3}$$

and $q_k = q_{k+1} + \mathrm{Tr}(P_{k+1} W_k)$. Note that for all k we have $P_k \succeq 0$, which can be seen simply by observing that the cost $J_k^*(x_k)$ is by its initial definition lower bounded by 0, hence the minimization of $x_k^T P_k x_k$ yields a bounded value. Finally, we see that the optimal cost for the problem is

$$\begin{aligned} J_0^*(x_0) &= x_0^T P_0 x_0 + \sum_{k=0}^{N-1} \mathrm{Tr}(P_{k+1} W_k) \\ &= \mathrm{Tr}(P_0 X_0) + \sum_{k=0}^{N-1} \mathrm{Tr}(P_{k+1} W_k), \end{aligned} \tag{4.4}$$

where $X_0 = x_0 x_0^T$.

Remark. Without the assumption $R_k \succ 0$, we could use the more general Schur complement result and replace inverses by pseudo-inverses under appropriate assumptions, following the remark in the previous paragraph.

This solution has a number of attractive properties which help explain its popularity. First, we automatically obtain a closed-loop feedback law (as always with the DP approach) (4.2) which has the convenient property of being *linear*. Hence we can synthesize automatically a optimum linear controller once the cost matrices have been specified. This control law depends on the availability of the *gain matrices* K_k, which can be computed in advance however and stored in the memory of the controller. This computation requires the computation of the matrices P_k using the *backward* difference equation (4.3), called the (discrete-time) Riccati difference equation, initialized with $P_N = Q_N$.

Moreover, a remarkable property of the solution is that the gain matrices K_k and the Riccati equation do not depend on the actual characteristics of the disturbances w_k, which only enter in the total cost (4.4). The deterministic problem with no disturbance (variables $\{w_k\}$ deterministic and equal to their mean, 0 in this case) has the exact same solution, with total cost simply equal to $\mathrm{Tr}(P_0 X_0)$. This special property shared by most linear quadratic optimal control problems is called the *certainty equivalence principle*. We could have designed the *optimal* controller even if we had assumed that the values of the disturbances were certain and fixed to their mean. Note that although this might be an positive characteristic from a computational point of view, it is not necessarily so from a modeling point of view. Indeed, a consequence of the certainty equivalence principle is that the controller does not change with the disturbance variability as long as it is zero-mean, i.e., the LQR controller is *risk-neutral*. There is a more general class of problems (linear exponential quadratic [Jac73]), which allows us to take into account the second order properties of the process noise, while still admitting a nice linear solution for the control law. A brief introduction to this *risk sensitive control* problem will be given in the problems.

Finally, note that we could use this solution for the inventory control problem, for which the dynamics are linear, as long as we assume the cost to be quadratic. This was proposed in 1960 by Holt et al. [HMMS60]. However, according to Porteus [Por02, p. 112], "the weakness of this model lies in the difficulty of fitting its parameters to the cost functions observed in practice and in the fact that it must allow for redemption of excess stock as well as replenishment of low stock levels". Indeed, this does not allow for a fixed positive ordering cost. Moreover, reductions in the stock level are allowed.

4.3 A First Glimpse at the Infinite-Horizon Control Problem

Let us briefly consider the problem as the length N of the horizon increases. We assume now that the problem to be time-homogeneous (often called stationary), i.e., the matrices A, B, Q, R, W are independent of k, although we can still have a different terminal cost matrix Q_N. Let us consider the simple 2-dimensional system

$$x_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k + w_k, \;\; x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.5}$$

with $W = 10^{-4} I$, $Q = I$, $R = \rho I$, $Q_N = \lambda I$. Fig. 4.1 shows two examples of control and state trajectories, for $\rho = 0.1$ and $\rho = 10$. We see that increasing the control costs clearly has an effect on the amplitude of the resulting optimal controls. Here however we are more interested in fig. 4.2, which shows that the gain matrices K_k and in fact the matrices P_k in the Riccati equation quickly converge to constant values as the terminal stage becomes more distant. Only when the final period N of the problem approaches do the gain values change noticeably.

This phenomenon is general (in economics, it is related to the concept of "turnpike theory" [tur, Ace08]), and exhibited by solutions of the Riccati difference equation (4.3). Because this equation has a number of important applications, in control theory and dynamical systems in particular, it has been extensively studied, with entire books devoted to the subject, see e.g. [BLW91]. In particular, its asymptotic properties are well-understood. A introductory discussion of these properties can be found in [Ber07, p. 151][2]. Suffice to say for now that under certain assumptions on the problem matrices (e.g. (A, B) controllable, $(A, Q^{1/2})$ observable), the matrices P_k converge as $k \to -\infty$ to a steady-state matrix P, solution of the *algebraic discrete-time Riccati equation* (ARE)

$$P = A^T P A + Q - A^T P B (B^T P B + R)^{-1} B^T P A. \tag{4.6}$$

Equivalently, we start at $k = 0$ and consider an infinite-horizon problem. The optimal control for this problem is then a constant-gain linear feedback $u_k = K x_k$, with $K = -(B^T P B + R)^{-1} B^T P A$. Moreover, this stationary control law has the important property of *stabilizing the system.* In our stochastic framework, this means that this control law is guaranteed to bring the state close to 0. This is another major advantage of linear quadratic methods, namely we automatically get stabilizing controllers, whereas in classical control stabilization is treated separately.

[2] we might discuss this again when we cover dynamic programming for infinite-horizon problems, depending on the background of the class in linear systems theory

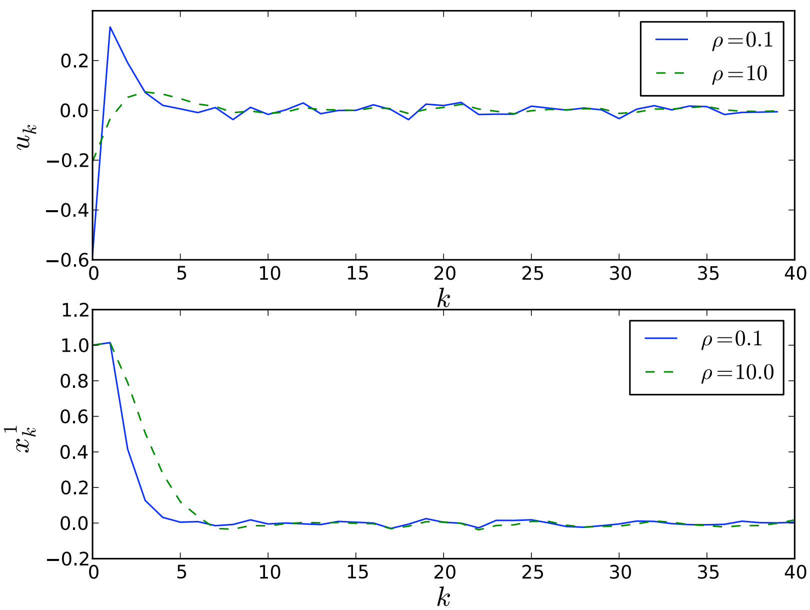

Figure 4.1: Sample state and control trajectories for the problem (4.5) with two different control costs. The realization of the noise trajectory is the same in both cases.

Note here already that there is an apparent difficulty in defining the optimal control problem as $N \to \infty$ because the total cost (4.4) diverges, at least in the presence of noise. The appropriately modified cost function in this case consists in studying the *average-cost*

$$J^* = \lim_{N \to \infty} \frac{1}{N} E \left\{ \sum_{k=0}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k) \right\}. \tag{4.7}$$

The optimal average-cost value, obtained by using the steady state controller K described above, is then $\mathrm{Tr}(PW)$, and in particular it is independent of the initial condition! For practical applications, the optimal steady-state controller is often used even for finite horizon problems because it is simpler to compute and much easier to store in memory than the time-varying controller. Its good practical performance can be explained by the rapid convergence observed as in Fig. 4.2.

4.4 Practice Problems

Problem 4.4.1. Do all the exercises found in the chapter.

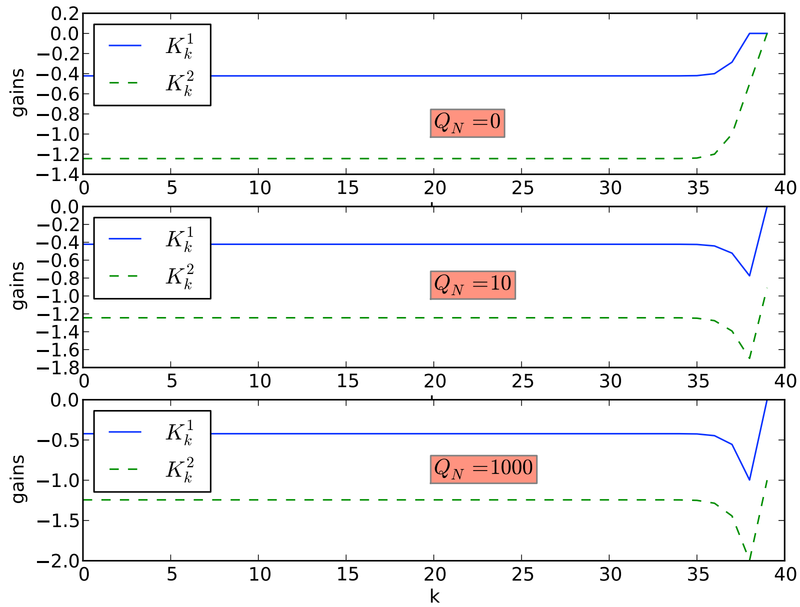

Figure 4.2: Controller gains K_k for different values of the final state cost matrix. Note in every case the rapid convergence as the horizon becomes more distant.