9 Infinite Horizon Limits

We define stabilizability and discuss the LQ regulation problem over an infinite horizon.

9.1 Linearization of nonlinear models

Linear models are important because they arise naturally via the linearization of nonlinear models. Consider the state-structured nonlinear model:

$$\dot{x} = a(x, u).$$

Suppose x, u are perturbed from an equilibrium $(\bar{x}, \bar{u})$ where $a(\bar{x}, \bar{u}) = 0$. Let $x' = x - \bar{x}$ and $u' = u - \bar{u}$ and immediately drop the primes. The linearized version is

$$\dot{x} = Ax + Bu$$

where

$$A = \left.\frac{\partial a}{\partial x}\right|_{(\bar{x},\bar{u})}, \quad B = \left.\frac{\partial a}{\partial u}\right|_{(\bar{x},\bar{u})}.$$

If $\bar{x}, \bar{u}$ is to be a stable equilibrium point then we must be able to choose a control that can stabilise the system in the neighbourhood of $(\bar{x}, \bar{u})$.

9.2 Stabilizability

Suppose we apply the stationary control $u = Kx$ so that $\dot{x} = Ax + Bu = (A + BK)x$. So with $\Gamma = A + BK$, we have

$$\dot{x} = \Gamma x, \quad x_t = e^{\Gamma t} x_0, \quad \text{where } e^{\Gamma t} = \sum_{j=0}^{\infty} (\Gamma t)^j / j!$$

Similarly, in discrete-time, we have can take the stationary control, $u_t = Kx_t$, so that $x_t = Ax_{t-1} + Bu_{t-1} = (A + BK)x_{t-1}$. Now $x_t = \Gamma^t x_0$.

We are interested in choosing Γ so that $x_t \to 0$ and $t \to \infty$.

Definition 9.1

Γ *is a* **stability matrix** *in the continuous-time sense if all its eigenvalues have negative real part, and hence* $x_t \to 0$ *as* $t \to \infty$.

Γ *is a* **stability matrix** *in the discrete-time sense if all its eigenvalues of lie strictly inside the unit disc in the complex plane,* $|z| = 1$*, and hence* $x_t \to 0$ *as* $t \to \infty$.

The $[A, B]$ *system is said to* **stabilizable** *if there exists a* K *such that* $A + BK$ *is a stability matrix.*

Note that $u_t = Kx_t$ is linear and Markov. In seeking controls such that $x_t \to 0$ it is sufficient to consider only controls of this type since, as we see below, such controls arise as optimal controls for the infinite-horizon LQ regulation problem.

9.3 Example: pendulum

Consider a pendulum of length L, unit mass bob and angle θ to the vertical. Suppose we wish to stabilise θ to zero by application of a force u. Then

$$\ddot{\theta} = -(g/L)\sin\theta + u.$$

We change the state variable to $x = (\theta, \dot{\theta})$ and write

$$\frac{d}{dt}\begin{pmatrix} \theta \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} \dot{\theta} \\ -(g/L)\sin\theta + u \end{pmatrix}$$
$$\sim \begin{pmatrix} \dot{\theta} \\ -(g/L)\theta \end{pmatrix} + \begin{pmatrix} 0 \\ u \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 \\ -(g/L) & 0 \end{pmatrix}\begin{pmatrix} \theta \\ \dot{\theta} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u.$$

Suppose we try to stabilise with a control $u = -K\theta = -Kx_1$. Then

$$A + BK = \begin{pmatrix} 0 & 1 \\ -(g/L) - K & 0 \end{pmatrix}$$

and this has eigenvalues $\pm\sqrt{-(g/L) - K}$. So either $-(g/L) - K > 0$ and one eigenvalue has a positive real part, in which case there is in fact instability, or $-(g/L) - K < 0$ and eigenvalues are purely imaginary, which means we will in general have oscillations. So successful stabilization must be a function of $\dot{\theta}$ as well, (and this would come out of solution to the LQ regulation problem.)

9.4 Infinite-horizon LQ regulation

Consider the time-homogeneous case and write the finite-horizon cost in terms of time to go s. The terminal cost, when $s = 0$, is denoted $F_0(x) = x^\top \Pi_0 x$. In all that follows we take $S = 0$, without loss of generality.

Lemma 9.2 *Suppose* $\Pi_0 = 0$*,* $R \geq 0$*,* $Q \geq 0$ *and* $[A, B, \ \cdot\ ]$ *is controllable or stabilizable. Then* $\{\Pi_s\}$ *has a finite limit* Π.

Proof. Costs are non-negative, so $F_s(x)$ is non-decreasing in s. Now $F_s(x) = x^\top \Pi_s x$. Thus $x^\top \Pi_s x$ is non-decreasing in s for every x. To show that $x^\top \Pi_s x$ is bounded we use one of two arguments.

If the system is controllable then $x^\top \Pi_s x$ is bounded because there is a policy which, for any $x_0 = x$, will bring the state to zero in at most n steps and at finite cost and can then hold it at zero with zero cost thereafter.

If the system is stabilizable then there is a K such that $\Gamma = A + BK$ is a stability matrix and using $u_t = Kx_t$, we have

$$F_s(x) \le x^\top \left[\sum_{t=0}^{\infty} (\Gamma^\top)^t (R + K^\top QK)\Gamma^t \right] x < \infty.$$

Hence in either case we have an upper bound and so $x^\top \Pi_s x$ tends to a limit for every x. By considering $x = e_j$, the vector with a unit in the jth place and zeros elsewhere, we conclude that the jth element on the diagonal of Π_s converges. Then taking $x = e_j + e_k$ it follows that the off diagonal elements of Π_s also converge. ■

Both value iteration and policy improvement are effective ways to compute the solution to an infinite-horizon LQ regulation problem. Policy improvement goes along the lines developed in Lecture 6.

The following theorem establishes the efficacy of value iteration. It is similar to Theorem 4.2 which established the same fact for D, N and P programming. The LQ regulation problem is a negative programming problem, however we cannot apply Theorem 4.2, because in general the terminal cost of $x^\top \Pi_0 x$ is not zero.

Theorem 9.3 *Suppose that $R > 0$, $Q > 0$ and the system $[A, B, \,\cdot\,]$ is controllable. Then (i) The equilibrium Riccati equation*

$$\Pi = f\Pi \tag{9.1}$$

has a unique non-negative definite solution Π. (ii) For any finite non-negative definite Π_0 the sequence $\{\Pi_s\}$ converges to Π. (iii) The gain matrix Γ corresponding to Π is a stability matrix.

Proof. (*starred*) Define Π as the limit of the sequence $f^{(s)}0$. By the previous lemma we know that this limit exists and that it satisfies (9.1).

Consider $u_t = Kx_t$ and $x_{t+1} = (A + BK)x_t = \Gamma x_t = \Gamma^t x_0$, for arbitrary x_0, where $K = -(Q + B^\top \Pi B)^{-1} B^\top \Pi A$ and $\Gamma = A + BK$. We can write (9.1) as

$$\Pi = R + K^\top QK + \Gamma^\top \Pi \Gamma. \tag{9.2}$$

and hence

$$x_t^\top \Pi x_t = x_t^\top (R + K^\top QK)x_t + x_{t+1}^\top \Pi x_{t+1} \ge x_{t+1}^\top \Pi x_{t+1}.$$

Thus $x_t^\top \Pi x_t$ decreases and, being bounded below by zero, it tends to a limit. Thus $x_t^\top (R+K^\top QK)x_t$ tends to 0. Since $R+K^\top QK$ is positive definite this implies $x_t \to 0$, which implies (iii). Hence for arbitrary finite non-negative definite Π_0,

$$\Pi_s = f^{(s)}\Pi_0 \ge f^{(s)}0 \to \Pi. \tag{9.3}$$

However, if we choose the fixed policy $u_t = Kx_t$ then it follows that

$$\Pi_s \le f^{(s)}0 + (\Gamma^\top)^s \Pi_0 \Gamma^s \to \Pi. \tag{9.4}$$

Thus (9.3) and (9.4) imply (ii).

Finally, if non-negative definite $\tilde{\Pi}$ also satisfies (9.1) then $\tilde{\Pi} = f^{(s)}\tilde{\Pi} \to \Pi$, whence (i) follows. ■

9.5 The $[A, B, C]$ system

The notion of controllability rested on the assumption that the initial value of the state was known. If, however, one must rely upon imperfect observations, then the question arises whether the value of state (either in the past or in the present) can be determined from these observations. The discrete-time system $[A, B, C]$ is defined by the plant equation and observation relation

$$x_t = Ax_{t-1} + Bu_{t-1},$$
$$y_t = Cx_{t-1}.$$

Here $y \in \mathbb{R}^r$ is observed, but x is not. We suppose C is $r \times n$. The observability question is whether or not we can infer x_0 from the observations $y_1, y_2, \ldots$. The notion of observability stands in dual relation to that of controllability; a duality that indeed persists throughout the subject.