11 Certainty-equivalent control

We show that the addition of noise to a linear system with quadratic costs does not change the optimal control, as a function of state.

Consider the realised stochastic controllable dynamical system $(G, (\varepsilon_n)_{n\geqslant 1})$, where

$$G(x, a, \varepsilon) = Ax + Ba + \varepsilon, \quad x \in \mathbb{R}^d, \quad a \in \mathbb{R}^m,$$

and where $(\varepsilon_n)_{n\geqslant 1}$ are independent $\mathbb{R}^d$-valued random variables, with mean $\mathbb{E}(\varepsilon) = 0$ and variance $\mathbb{E}(\varepsilon\varepsilon^T) = N$. Thus the controlled process, for a given starting point $X_0 = x$, is given by

$$X_{n+1} = AX_n + BU_n + \varepsilon_{n+1}$$

where $U_n = u_n(X_0, \dots, X_n)$ is the control. We study the n-horizon problem with non-negative quadratic instantaneous costs $c(x, a)$ and final cost $c(x)$, as in the preceding section. Thus

$$c(x, a) = x^T Rx + x^T S^T a + a^T Sx + a^T Qa, \quad c(x) = x^T \Pi_0 x.$$

Set

$$V_n^u(x) = \mathbb{E}_x^u \left(\sum_{k=0}^{n-1} c(X_k, U_k) + c(X_n) \right), \quad V_n(x) = \inf_u V_n^u(x).$$

Suppose inductively that

$$V_n(x) = x^T \Pi_n x + \gamma_n.$$

This is true for $n = 0$ if we take $\gamma_0 = 0$. By a straightforward generalization[25] of Proposition 2.1, V_{n+1} is given by the optimality equation

$$V_{n+1}(x) = \inf_a \{c(x, a) + \mathbb{E}(V_n(Ax + Ba + \varepsilon))\}.$$

We have

$$\begin{aligned} \mathbb{E}(V_n(Ax + Ba + \varepsilon)) &= \mathbb{E}((Ax + Ba + \varepsilon)^T \Pi_n (Ax + Ba + \varepsilon)) + \gamma_n \\ &= (Ax + Ba)^T \Pi_n (Ax + Ba) + \mathbb{E}(\varepsilon^T \Pi_n \varepsilon) + \gamma_n \end{aligned}$$

and we showed in the preceding section that

$$\inf_a \{c(x, a) + (Ax + Ba)^T \Pi_n (Ax + Ba)\} = x^T r(\Pi_n) x,$$

with minimizing action $a = K(\Pi_n)$. Also

$$\mathbb{E}(\varepsilon^T \Pi_n \varepsilon) = \sum_{i,j} \mathbb{E}(\varepsilon_i (\Pi_n)_{ij} \varepsilon_j) = \sum_{i,j} \mathbb{E}(N_{ij} (\Pi_n)_{ij}) = \operatorname{trace}(N\Pi_n).$$

So $V_{n+1}(x) = x^T \Pi_{n+1} x + \gamma_{n+1}$, where $\Pi_{n+1} = r(\Pi_n)$ and $\gamma_{n+1} = \gamma_n + \operatorname{trace}(N\Pi_n)$. By induction, we have proved the following result.

[25] We have moved out of the setting of a countable state space used in Section 2. For a function F on $S \times A$, instead of writing PF as a sum, we can use the formula $PF(x, a) = \mathbb{E}(F(G(x, a, \varepsilon)))$.

Proposition 11.1. *For the linear system*

$$X_{n+1} = AX_n + BU_n + \varepsilon_{n+1},$$

with independent perturbations $(\varepsilon_n)_{n\geqslant 1}$*, having mean* 0 *and variance* N*, and with non-negative quadratic costs as above, the infimal cost function is given by*

$$V_n(x) = x^T \Pi_n x + \gamma_n$$

and the n*-horizon optimal control is* $U_k = K(\Pi_{n-1-k})X_k$*.*

This is *certainty-equivalent control* as the optimal control is the same as for $\varepsilon = 0$.