11. Certainty equivalent control

Recall from §3 the notion of a realized stochastic controllable dynamical system $(G, (\varepsilon_n)_{n\geq 1})$:

- $(\varepsilon_n)_{n\geq 1}$ sequence of i.i.d. r.v's in E.
- $G : S \times A \times E \longrightarrow S$.

The controlled process, for a given starting point $X_0 = x$ and a given control $u : S_0^* \longrightarrow A$, is given by

$$X_{n+1} = G(X_n, U_n, \varepsilon_{n+1}), \qquad U_n = u_n(X_0, \dots, X_n), \quad n \geq 0.$$

Consider the n-horizon problem with non-negative instantaneous costs $c(x,a)$ and final costs $V_0(x)$.

Thus we seek to minimize over all controls

$$V_n^u(x) = \mathbb{E}_x^u\left(\sum_{k=0}^{n-1} c(X_k, U_k) + V_0(X_n) \right).$$

We write $V_n(x) = \inf_u V_n^u(x)$ for the infimal cost function as usual.

Given an initial state x and a control u, write $a = u_0(x)$ and define a new control $\tilde{u}$ by

$$\tilde{u}_n(x_0, \dots, x_n) = u_{n+1}(x, x_0, \dots, x_n), \quad n \geq 0.$$

Then

$$V^u_{n+1}(x) = c(x,a) + \mathbb{E}\left(V^{\tilde{u}}_n (G(x, a, \varepsilon_1))\right).$$

To see this consider the controlled process $(X_n)_{n \geq 0}$ starting from x with control u, and condition on $\varepsilon_1 = \varepsilon$. Set $\tilde{X}_n = X_{n+1}$, $\tilde{\varepsilon}_n = \varepsilon_{n+1}$, $\tilde{U}_n = \tilde{u}_n(\tilde{X}_0, \dots, \tilde{X}_n)$, then $\tilde{X}_0 = G(x, a, \varepsilon)$ and $\tilde{X}_{n+1} = G(\tilde{X}_n, \tilde{U}_n, \tilde{\varepsilon}_n)$.

Hence

$$\mathbb{E}_x^u \left(\sum_{k=1}^{n} c(X_k, U_k) + V_0(X_{n+1}) \,\middle|\, \varepsilon_1 = \varepsilon \right)$$

$$= \mathbb{E}_x^u \left(\sum_{k=0}^{n-1} c(\tilde{X}_k, \tilde{U}_k) + V_0(\tilde{X}_n) \,\middle|\, \varepsilon_1 = \varepsilon \right)$$

$$= V_n^{\tilde{u}}(G(x, a, \varepsilon))$$

and hence

$$V_{n+1}^u(x) = c(x, a) + \mathbb{E}_x^u \left(\sum_{k=1}^{n} c(X_k, U_k) + V_0(X_{n+1}) \right)$$

$$= c(x, a) + \mathbb{E}\left(V_n^{\tilde{u}} (G(x, a, \varepsilon_1)) \right)$$

Now specialize to the case

- $G(x, a, \varepsilon) = Ax + Ba + \varepsilon, \quad x \in \mathbb{R}^d, \ a \in \mathbb{R}^m, \ \varepsilon \in \mathbb{R}^d,$
- $(\varepsilon_n)_{n \geq 1}$ sequence of zero-mean independent r.v's in $\mathbb{R}^d$ with $\operatorname{var}(\varepsilon) = N$ (ie $\operatorname{cov}(\varepsilon_i, \varepsilon_j) = \mathbb{E}(\varepsilon_i \varepsilon_j) = N_{ij}$),
- $c(x, a) = x^T R x + x^T S^T x + a^T S x + a^T Q a,$
- $V_0(x) = x^T \Pi_0 x.$

The dynamics are thus

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

Proposition 11.1

For the linear system $X_{n+1} = AX_n + BU_n + \varepsilon_{n+1}$, with independent zero-mean perturbations ε_n of variance N, for the n-horizon problem with non-negative instantaneous and final costs

$$c(x,a) = x^T R x + x^T S^T a + a^T S x + a^T Q a, \qquad V_0(x) = x^T \Pi_0 x,$$

the infimal cost function and optimal control are given by

$$V_n(x) = x^T \Pi_n x + \gamma_n, \qquad U_k = K(\Pi_{n-k-1}) X_k, \quad k = 0, 1, \ldots, n-1,$$

where $\Pi_{n+1} = r(\Pi_n)$ and $\gamma_{n+1} = \gamma_n + \operatorname{trace}(N\Pi_n)$ for $n \geq 0$, and $\gamma_0 = 0$.

<u>Proof</u> The case $n = 0$ is trivial. Suppose inductively that the claim holds for n. Then

$$V_n^u(x) \geq x^T \Pi_n x \quad \text{with equality when } u_k(x_0, \dots, x_k) = K(\Pi_{n-k-1}) x_k .$$

Given a starting point x and a control u, set $a = u_0(x)$ and $\tilde{u}_k(x_0, \dots, x_k) = u_{k+1}(x, x_0, \dots, x_k)$. Then

$$\begin{aligned} V_{n+1}^u(x) &= c(x,a) + \mathbb{E}\left(V_n^{\tilde{u}}(Ax + Ba + \varepsilon_1)\right) \\ &\geq c(x,a) + \mathbb{E}\left((Ax + Ba + \varepsilon)^T \Pi_n (Ax + Ba + \varepsilon_1)\right) + \gamma_n \\ &= \left\{ c(x,a) + (Ax+Ba)^T \Pi_n (Ax + Ba) \right\} + \mathbb{E}\left(\varepsilon_1^T \Pi_n \varepsilon_1\right) + \gamma_n \end{aligned}$$

with equality when $u_{k+1}(x, x_0, \dots, x_k) = K(\Pi_{n-k-1}) x_k$, for $k = 0, 1, \dots, n-1$.

Now

$$\mathbb{E}(\varepsilon_i \Pi_{ij} \varepsilon_j) = \Pi_{ij}\, \mathbb{E}(\varepsilon_i \varepsilon_j) = \Pi_{ij} N_{ij}$$

so

$$\mathbb{E}(\varepsilon_1^T \Pi_n \varepsilon_1) = \operatorname{trace}(N \Pi_n).$$

By Step 2 in §10,

$$c(x,a) + (Ax+Ba)^T \Pi_n (Ax+Ba) \geq x^T r(\Pi_n) x,$$

with equality when $u_0(x) = a = K(\Pi_n)x$.

Hence the claim holds for $n+1$ and the induction proceeds.

$\square$

- In particular, V_n satisfies the optimality equation

$$V_{n+1}(x) = \inf_a \{ C(x,a) + \mathbb{E}(V_n(Ax + Ba + \varepsilon_1)) \}.$$

- The control values are a function of the present state only and are the same such function for all N, in particular for the case $N = 0$, which we dealt with in §10. This is called <u>certainty equivalence control</u>.

- In the case where $\mathrm{var}(\varepsilon_n) = N_n$, a similar argument shows that control is still certainty equivalent, with

$$V_n(x) = x^T \Pi_n x + \sum_{k=1}^{n} \mathrm{trace}(N_k \Pi_{n-k}).$$

12. LQG systems and the Kalman filter

We consider the same dynamics as in §11:

$$X_{n+1} = AX_n + BU_n + \varepsilon_{n+1}, \quad n \geq 0$$

but now the controls must be chosen as a function $U_n = u_n(Y_1, \ldots, Y_n)$ of a sequence of observations in $\mathbb{R}^p$:

$$Y_{n+1} = CX_n + \eta_{n+1}, \quad n \geq 0.$$

Here C is a given $p \times d$-matrix, and $(\eta_n)_{n\geq 1}$ is a sequence of r.v.'s in $\mathbb{R}^p$. We assume that

$$X_0, \begin{pmatrix} \varepsilon_1 \\ \eta_1 \end{pmatrix}, \begin{pmatrix} \varepsilon_2 \\ \eta_2 \end{pmatrix}, \ldots$$

are independent Gaussian (normal) r.v.'s, with

$$\mathbb{E}(X_0) = x, \ \mathrm{var}(X_0) = \Sigma_0, \ \mathrm{var}(\varepsilon_n) = N, \ \mathrm{cov}(\varepsilon_n, \eta_n) = L, \ \mathrm{var}(\eta_n) = M.$$

There are non-negative quadratic cost functions, as in §§10, 11:

$$c(x,a) = x^T R x + x^T S^T a + a^T S x + a^T Q a, \quad c(x) = x^T \Pi_0 x.$$

In the n-horizon problem, we seek to minimize

$$V_n^u(x, \Sigma_0) = \mathbb{E}^u_{(x,\Sigma_0)}\left(\sum_{k=0}^{n-1} c(X_k, U_k) + c(X_n) \right)$$

over all controls $u_n : (\mathbb{R}^p)^n \to \mathbb{R}^m$, $n \geq 0$, writing as usual V_n for the infimal cost function

$$V_n(x, \Sigma_0) = \inf_u V_n^u(x, \Sigma_0).$$

- There is no observation at $n=0$, so U_0 has to be chosen without any knowledge of the process.
- Observations of the form $Y_n = CX_n + \eta_n$ are often more natural. This can be transformed to the given problem, which has the advantage of neater formulas.
- Note that V_n^u now has a dependence on Σ_0, reflecting our uncertainty about the state of the system.

Lemma 12.1

Let (X, Y) be a Gaussian random variable in $\mathbb{R}^d \times \mathbb{R}^p$, having zero mean and

$$\operatorname{var}(X) = U, \quad \operatorname{cov}(X, Y) = W, \quad \operatorname{var}(Y) = V,$$

with V invertible. Then

$$X = \hat{X} + Z$$

where $\hat{X} = WV^{-1}Y$, and where Z is independent of Y, with

$$\operatorname{var}(Z) = U - WV^{-1}W^T.$$

Proof

Since (Y, Z) is Gaussian, we can show Y and Z are independent by showing their covariance is zero. We have

$$\operatorname{cov}(Z, Y) = \mathbb{E}(ZY^T) = \mathbb{E}(XY^T) - \mathbb{E}(WV^{-1}YY^T)$$
$$= W - WV^{-1}V = 0$$

and

$$\operatorname{var}(Z) = \mathbb{E}(ZZ^T) = \mathbb{E}(ZX^T)$$
$$= \mathbb{E}(XX^T) - \mathbb{E}(WV^{-1}YX^T) = U - WV^{-1}W^T.$$

$\square$

Example

$$X_{n+1} = X_n + U_n, \qquad X_0 \sim N(x, v)$$

$$Y_{n+1} = X_n + \eta_{n+1}, \qquad \eta_1, \eta_2, \ldots \text{ independent } N(0,1)$$

$$V_n^u(x, v) = \mathbb{E}^u_{(x,v)} \left(\sum_{k=0}^{n-1} U_k^2 + D X_n^2 \right)$$

Find the optimal control and evaluate the infimal cost function.

We show how to obtain from the partially observed system

$$X_{n+1} = AX_n + BU_n + \varepsilon_{n+1}, \qquad X_0 \sim N(x, v),$$
$$Y_{n+1} = CX_n + \eta_{n+1},$$

a related fully observed system

$$\hat{X}_{n+1} = A\hat{X}_n + BU_n + \hat{\varepsilon}_{n+1}$$

which will allow us to determine the in final cost function and optimal control for the partially observed system with quadratic costs.

We seek a decomposition

$$X_n = \hat{X}_n + \Delta_n$$

Where

- $\hat{X}_n$ is a function of $Y_1, \dots, Y_n$

  (in particular $\hat{X}_0$ is a constant),

- Δ_n is a zero mean Gaussian, independent of $Y_1, \dots, Y_n$.

For $n=0$, we can take

$$\hat{X}_0 = x, \qquad \Delta_0 = X_0 - x \sim N(0, \Sigma_0), \qquad \Sigma_0 = v.$$

Suppose a decomposition has been found for n:

$$X_n = \hat{X}_n + \Delta_n, \quad \Delta_n \sim N(0, \Sigma_n).$$

Then

$$\begin{pmatrix} X_{n+1} \\ Y_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} A\hat{X}_n + BU_n \\ C\hat{X}_n \end{pmatrix}}_{\substack{\text{a function of} \\ Y_1, \dots, Y_n}} + \underbrace{\begin{pmatrix} \xi_{n+1} \\ \zeta_{n+1} \end{pmatrix}}_{\substack{\text{zero-mean Gaussian} \\ \text{independent of } Y_1, \dots, Y_n}}, \quad \begin{pmatrix} \xi_{n+1} \\ \zeta_{n+1} \end{pmatrix} = \begin{pmatrix} \varepsilon_{n+1} + A\Delta_n \\ \eta_{n+1} + C\Delta_n \end{pmatrix}.$$

By Lemma 12.1, there are matrices H_{n+1}, Σ_{n+1} such that

$$\xi_{n+1} = \hat{\xi}_{n+1} + \Delta_{n+1}, \quad \hat{\xi}_{n+1} = H_{n+1}\zeta_{n+1}, \quad \Delta_{n+1} \sim N(0, \Sigma_{n+1})$$

with Δ_{n+1} independent of ζ_{n+1}.

We compute H_{n+1}, Σ_{n+1} using Lemma 12.1. We have

$$\operatorname{var}(\xi_{n+1}) = \tilde{N} = N + A\Sigma_n A^T,$$

$$\operatorname{var}(\zeta_{n+1}) = \tilde{M} = M + C\Sigma_n C^T,$$

$$\operatorname{cov}(\xi_{n+1}, \zeta_{n+1}) = \tilde{L} = L + A\Sigma_n C^T.$$

So

$$H_{n+1} = H(\Sigma_n) := \tilde{L}\tilde{M}^{-1} = (L + A\Sigma_n C^T)(M + C\Sigma_n C^T)^{-1},$$

$$\Sigma_{n+1} = G(\Sigma_n) := \tilde{N} - \tilde{L}\tilde{M}^{-1}\tilde{L}^T$$

$$= (N + A\Sigma_n A^T) - (L + A\Sigma_n C^T)(M + C\Sigma_n C^T)^{-1}(L + A\Sigma_n C^T)^T.$$

Then

- $\hat{\varepsilon}_{n+1} = H_{n+1}(Y_{n+1} - C\hat{X}_n)$ is a function of $Y_1, \dots, Y_{n+1}$
- Δ_{n+1} is independent of $Y_1, \dots, Y_{n+1}$.

Set

$$\hat{X}_{n+1} = A\hat{X}_n + BU_n + \hat{\varepsilon}_{n+1}$$

then we have the desired decomposition for $n+1$:

$$X_{n+1} = \hat{X}_{n+1} + \Delta_{n+1}.$$

We have computed the conditional distribution of the state X_n given the observations $Y_1, \dots, Y_n$:

$$X_n \sim N(\hat{X}_n, \Sigma_n)$$

where

$$\hat{X}_{n+1} = A\hat{X}_n + BU_n + H(\Sigma_n)(Y_{n+1} - C\hat{X}_n), \qquad \hat{X}_0 = x$$

$$\Sigma_{n+1} = G(\Sigma_n), \quad \Sigma_0 = v.$$

This is called the <u>Kalman filter</u>.

We can consider

$$\hat{X}_{n+1} = A\hat{X}_n + BU_n + \hat{\varepsilon}_{n+1}, \quad \Sigma_{n+1} = \sigma(\Sigma_n)$$

as a fully observed controllable linear system.

The random variables $(\hat{\varepsilon}_n)_{n\geq 1}$ are independent, with

$$\hat{\varepsilon}_{n+1} \sim N(0, \hat{N}(\Sigma_n))$$

where

$$\hat{N}(\Sigma_n) = \mathrm{var}(\xi_{n+1}) - \mathrm{var}(\Delta_{n+1}) = \tilde{N} - \Sigma_{n+1} = \tilde{L}\tilde{M}^{-1}\tilde{L}^T$$

$$= (L + A\Sigma_n C^T)(M + C\Sigma_n C^T)^{-1}(L + A\Sigma_n C^T)^T.$$

Note that

$$\mathbb{E}(c(X_k, U_k)) = \mathbb{E}(c(\hat{X}_k + \Delta_k, U_k))$$
$$= \underbrace{\mathbb{E}(\Delta_k^T R \Delta_k)}_{= \operatorname{trace}(R\Sigma_k)} + \mathbb{E}(c(\hat{X}_k, U_k))$$

and, similarly

$$\mathbb{E}(c(X_n)) = \operatorname{trace}(\Pi_0 \Sigma_n) + \mathbb{E}(c(\hat{X}_n)).$$

So

$$V_n^u(x, \Sigma_0) = \underbrace{\hat{V}_n^u(x, \Sigma_0)}_{\substack{\text{total expected cost for } n\text{-horizon problem with same cost} \\ \text{functions on the fully observed system}}} + \sum_{k=0}^{n-1} \operatorname{trace}(R\Sigma_k) + \operatorname{trace}(\Pi_0 \Sigma_n).$$

By certainty equivalence, the optimal control is given by

$$U_k = K(\Pi_{n-k-1})\hat{X}_k, \quad k=0,1,\dots,n-1,$$

where $\Pi_{n+1} = r(\Pi_n)$ for $n \geq 0$. Also

$$\hat{V}_n(x, \Sigma_0) = x^T \Pi_n x + \sum_{k=1}^{n} \operatorname{trace}(\hat{N}_k \Pi_{n-k})$$

where $\hat{N}_k = \operatorname{var}(\hat{\varepsilon}_k) = \hat{N}(\Sigma_{k-1})$, $k \geq 1$. So the infimal cost function for the partially observed system is given by

$$V_n(x, \Sigma_0) = x^T \Pi_n x + \sum_{k=1}^{n} \operatorname{trace}(\hat{N}_k \Pi_{n-k}) + \sum_{k=0}^{n-1} \operatorname{trace}(R\Sigma_k) + \operatorname{trace}(\Pi_0 \Sigma_n).$$

The product form of the optimal control is interesting

$$U_k = K(\Pi_{n-k-1})\,\hat{X}_k$$

$K(\Pi_{n-k-1})$ determined from $A, B, \underbrace{R, S, Q}_{\text{costs}}, \Pi_0, n$

$\hat{X}_k$ from the Kalman filter, using $A, B, \underbrace{C, L, M, N, \Sigma_0}_{\text{observations \& noise}}$

This splitting is called the <u>separation principle</u>.

13. Observability

Consider the system

$$x_{n+1} = Ax_n \qquad \text{state in } \mathbb{R}^d$$
$$y_{n+1} = Cx_n \qquad \text{observation in } \mathbb{R}^p$$

Say the system is <u>observable in</u> n <u>steps</u> if, for all $x_0 \in \mathbb{R}^d$, $y_1, \ldots, y_n$ determine uniquely the initial state x_0.

Also, it is <u>observable</u> if it is observable in n step for some $n \geq 1$.

We have

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = N_n x_0, \qquad N_n = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

So the system is observable in n steps if and only if

$$\operatorname{rank}(N_n) = d.$$

By the Cayley-Hamilton theorem, the system is observable if and only if

$$\operatorname{rank}(N_d) = d.$$

Consider now the continuous-time system

$$\dot{x}_t = A x_t \qquad \text{state in } \mathbb{R}^d$$

$$\dot{y}_t = C x_t, \quad y_0 = 0 \qquad \text{observation in } \mathbb{R}^p$$

The system is <u>observable</u> in time t if, for all $x_0 \in \mathbb{R}^d$, $(y_s)_{0 \le s \le t}$ determines uniquely the initial state x_0.

Note that

$$x_t = e^{At} x_0, \qquad y_t = \int_0^t C e^{As} x_0 \, ds.$$

By induction on n, we have $\left(\frac{d}{dt}\right)^n y_t = CA^{n-1}x_t$, $n \geq 1$

For any $t > 0$, $(y_s)_{0 \leq s \leq t}$ determines the values of these derivatives at 0, so the condition $\operatorname{rank}(N_d) = d$ is sufficient for observability in time t.

On the other hand, if $\operatorname{rank}(N_d) \leq d-1$, then there exists $x_0 \in \mathbb{R}^d \setminus \{0\}$ such that $CA^n x_0 = 0$ for $n = 0, 1, \dots, d-1$, and hence for all n by Cayley-Hamilton. Then

$$y_t = \int_0^t C \sum_{n=0}^{\infty} \frac{(As)^n}{n!} x_0 = 0 \quad , \quad t \geq 0,$$

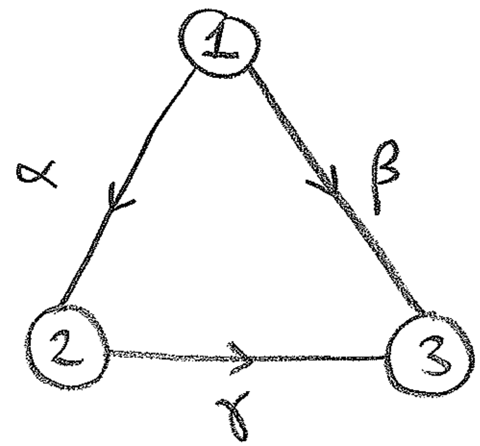
so we cannot distinguish x_0 from 0. The condition $\operatorname{rank}(N_d) = d$ is thus necessary for observability (in any time $t > 0$)

13.5 .

Example – the sum of two populations

Suppose $\dot{x}_t = \lambda x_t$, $\dot{z}_t = \mu z_t$ and we observe $y_t = x_t + z_t$.
Can we determine x_0 and z_0 ?

Example – radioactive decay



Suppose

$$\dot{x}_t^1 = -(\alpha+\beta)x_t^1$$
$$\dot{x}_t^2 = \alpha x_t^1 - \gamma x_t^2$$
$$\dot{x}_t^3 = \beta x_t^1 + \gamma x_t^2$$

We observe $(x_t^3)_{t\geq 0}$. Can we determine x_0^1, x_0^2 ?