10 Observability

We present observability and the LQG model with process and observation noise.

10.1 Observability

The previous lecture introduced the $[A, B, C]$ system,

$$x_t = Ax_{t-1} + Bu_{t-1}, \tag{10.1}$$

$$y_t = Cx_{t-1}. \tag{10.2}$$

$y \in \mathbb{R}^k$ is observed, but x is not. C is $k \times n$. The **observability** question is: can we infer x at a prescribed time by subsequent y values?

Definition 10.1 *A system is said to be* **r-observable** *if x_0 can be inferred from knowledge of the subsequent observations $y_1, \dots, y_r$ and relevant control values $u_0, \dots, u_{r-2}$ for any initial x_0. A system of dimension n is said to be* **observable** *if it is r-observable for some r.*

From (10.1) and (10.2) we can determine y_t in terms of x_0 and subsequent controls:

$$x_t = A^t x_0 + \sum_{s=0}^{t-1} A^s B u_{t-s-1},$$

$$y_t = Cx_{t-1} = C\left[A^{t-1}x_0 + \sum_{s=0}^{t-2} A^s B u_{t-s-2}\right].$$

Thus, if we define the 'reduced observation'

$$\tilde{y}_t = y_t - C\left[\sum_{s=0}^{t-2} A^s B u_{t-s-2}\right],$$

then x_0 is to be determined from the system of equations

$$\tilde{y}_t = CA^{t-1}x_0, \quad 1 \le t \le r. \tag{10.3}$$

By hypothesis, these equations are mutually consistent, and so have a solution; the question is whether this solution is unique. This is the reverse of the situation for controllability, when the question was whether the equation for u had a solution at all, unique or not. Note that an implication of the system definition is that the property of observability depends only on the matrices A and C; not upon B at all.

Theorem 10.2 *(i) The system $[A, \cdot, C]$ is r-observable if and only if the matrix*

$$N_r = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{r-1} \end{bmatrix}$$

has rank n, or (ii) equivalently, if and only if the $n \times n$ matrix

$$N_r^\top N_r = \sum_{j=0}^{r-1} (A^\top)^j C^\top C A^j$$

is nonsingular. (iii) If the system is r-observable then it is s-observable for $s \ge \min(n, r)$, and (iv) the determination of x_0 can be expressed

$$x_0 = (N_r^\top N_r)^{-1} \sum_{j=1}^{r} (A^\top)^{j-1} C^\top \tilde{y}_j. \tag{10.4}$$

Proof. If the system has a solution for x_0 (which is so by hypothesis) then this solution must is unique if and only if the matrix N_r has rank n, whence assertion (i). Assertion (iii) follows from (i). The equivalence of conditions (i) and (ii) can be verified directly as in the case of controllability.

If we define the deviation $\eta_t = \tilde{y}_t - CA^{t-1}x_0$ then the equation amounts to $\eta_t = 0$, $1 \le t \le r$. If these equations were not consistent we could still define a 'least-squares' solution to them by minimizing any positive-definite quadratic form in these deviations with respect to x_0. In particular, we could minimize $\sum_{t=0}^{r-1} \eta_t^\top \eta_t$. This minimization gives (10.4). If equations (10.3) indeed have a solution (i.e., are mutually consistent, as we suppose) and this is unique then expression (10.4) must equal this solution; the actual value of x_0. The criterion for uniqueness of the least-squares solution is that $N_r^\top N_r$ should be nonsingular, which is also condition (ii). ∎

Note that we have again found it helpful to bring in an optimization criterion in proving (iv); this time, not so much to construct one definite solution out of many, but rather to construct a 'best-fit' solution where an exact solution might not have existed. This approach lies close to the statistical approach necessary when observations are corrupted by noise.

10.2 Observability in continuous-time

Theorem 10.3 *(i) The n-dimensional continuous-time system $[A, \cdot, C]$ is observable if and only if the matrix N_n has rank n, or (ii) equivalently, if and only if*

$$H(t) = \int_0^t e^{A^\top s} C^\top C e^{As}\, ds$$

is positive definite for all $t > 0$. (iii) If the system is observable then the determination of $x(0)$ can be written

$$x(0) = H(t)^{-1} \int_0^t e^{A^\top s} C^\top \tilde{y}(s)\, ds,$$

where

$$\tilde{y}(t) = y(t) - \int_0^t CA^{t-s} Bu(s)\, ds.$$

10.3 Examples

Example. Observation of population
Consider two populations whose sizes are changing according to the equations

$$\dot{x}_1 = \lambda_1 x_1, \qquad \dot{x}_2 = \lambda_2 x_2.$$

Suppose we observe $x_1 + x_2$, so

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}.$$

and so the individual populations are observable if $\lambda_1 \neq \lambda_2$.

Example. Radioactive decay
Suppose $j = 1, 2, 3$ and decay is from 1 to 2 at rate α, from 1 to 3 at rate β and from 2 to 3 at rate γ. We observe only the accumulation in state 3. Then $\dot{x} = Ax$, where

$$A = \begin{pmatrix} -\alpha-\beta & 0 & 0 \\ \alpha & -\gamma & 0 \\ \beta & \gamma & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix},$$

$$N_3 = \begin{pmatrix} C \\ CA \\ CA^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \beta & \gamma & 0 \\ \alpha\gamma - \beta(\alpha+\beta) & -\gamma^2 & 0 \end{pmatrix}.$$

The determinant equals $\gamma(\alpha+\beta)(\gamma-\beta)$, so we require $\gamma > 0$, $\alpha+\beta > 0$, $\beta \neq \gamma$ for a nonzero determinant and observability.

Example. Satellite
Recall the linearised equation $\dot{x} = Ax$, for perturbations of the orbit of a satellite, (here taking $\rho = 1$), where

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} r-\rho \\ \dot{r} \\ \theta - \omega t \\ \dot{\theta} - \omega \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{pmatrix}.$$

By taking $C = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ we see that we can see that system is observable on the basis of angle measurements alone but, not observable for $\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$, i.e., on the basis of radius movements alone.

$$N_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \\ -6\omega^3 & 0 & 0 & -4\omega^2 \end{bmatrix} \quad \tilde{N}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & -\omega^2 & 0 & 0 \end{bmatrix}$$

10.4 Imperfect state observation with noise

The full LQG model, whose description has been deferred until now, assumes linear dynamics, quadratic costs and Gaussian noise. Imperfect observation is the most important point. The model is

$$x_t = Ax_{t-1} + Bu_{t-1} + \epsilon_t, \tag{10.5}$$
$$y_t = Cx_{t-1} + \eta_t, \tag{10.6}$$

where ϵ_t is process noise, y_t is the observation at time t and η_t is the observation noise. The state observations are degraded in that we observe only Cx_{t-1}. Assume

$$\operatorname{cov}\begin{pmatrix} \epsilon \\ \eta \end{pmatrix} = E\begin{pmatrix} \epsilon \\ \eta \end{pmatrix}\begin{pmatrix} \epsilon \\ \eta \end{pmatrix}^\top = \begin{pmatrix} N & L \\ L^\top & M \end{pmatrix}$$

and that $x_0 \sim N(\hat{x}_0, V_0)$. Let $W_t = (Y_t, U_{t-1}) = (y_1, \ldots, y_t; u_0, \ldots, u_{t-1})$ denote the observed history up to time t. Of course we assume that t, A, B, C, N, L, M, $\hat{x}_0$ and V_0 are also known; W_t denotes what might be different if the process were rerun. In the next lecture we turn to the question of estimating x from y.

We consider the issues of state estimation and optimal control and shall show:

(i) $\hat{x}_t$ can be calculated recursively from the Kalman filter (a linear operator):

$$\hat{x}_t = A\hat{x}_{t-1} + Bu_{t-1} + H_t(y_t - C\hat{x}_{t-1}),$$

which is like the plant equation except that the noise is given by an **innovation process**, $\tilde{y}_t = y_t - C\hat{x}_y$, rather than the white noise.

(ii) If with full information (i.e., $y_t = x_t$) the optimal control is $u_t = K_t x_t$, then without full information the optimal control is $u_t = K_t \hat{x}_t$, where $\hat{x}_t$ is the best linear least squares estimate of x_t based on the information (Y_t, U_{t-1}) at time t.

Many of the ideas we encounter in this analysis are unrelated to the special state structure and are therefore worth noting as general observations about control with imperfect information.