ME 7247: Advanced Control Systems	**Fall 2022–23**

Lecture 16: LQG control

Tuesday November 1, 2022

Lecturer: Laurent Lessard *Scribe: Youbang Sun*

In the last lecture we studied the controllability and observability of a system in deterministic systems, and using these tools we studied system identification and model reduction. This lecture we go back to the stochastic system with noise, and consider Linear–Quadratic–Gaussian (LQG) control. We want to find the optimal control policy for a stochastic system where the observation is also affected by Gaussian noise.

1 LQG Problem Formulation

The stochastic system of question is described by the following equations,

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= Cx_t + v_t, \\ w_t &\sim \mathcal{N}(0, W), v_t \sim \mathcal{N}(0, V). \end{aligned} \tag{1}$$

The system is a standard linear stochastic system, x_t is the system states at time t, u_t is the system input at time t, y_t is the system observation at time t, lastly w_t and v_t are the system noise and observation noise at time t, both noise follow a Gaussian distribution.

Previously we looked at when we only have noisy observations of the system and we try to find the optimal estimation of the system states from the observations, the method is called Kalman Filtering. We also looked at the task where we know the state of a stochastic system and we want to find the optimal control input minimizing a given quadratic cost, this method is called LQR. Here we want to combine the two together. Suppose we only have the noisy observations of the stochastic system, and we want to find the optimal control based on the observation, such that the expectation of the quadratic cost is minimized.

Similar to previous study on stochastic systems, since the noise in the system is not deterministic, we aim to minimize the expectation of the quadratic cost. The objective is defined as follows,

$$J = \underset{u_0,\dots,u_{N-1}}{\text{minimize}} \quad \mathbf{E}\left[\sum_{t=0}^{N-1} (x_t^\mathsf{T} Q x_t + u_t^\mathsf{T} Q u_t) + x_N^\mathsf{T} Q_f x_N \;\middle|\; y_0, \dots, y_{N-1} \right] \tag{2}$$

Note. The notation above isn't quite correct accurate. What we mean is the the decision u_t should depend on $y_0, \dots, y_{t-1}$. So each decision depends on a different subset of the information.

1.1 Separation Principle

We want to prove the **Separation Principle**. In this case, Separation Principle means that the optimal control can be generated by first use Kalman Filter to find optimal system estimation, then we use LQR to find the optimal output based on the optimal estimation, and the end result is indeed the optimal control feedback output for the LQG task.

Another note is that although we only provide analysis of Separation Principle in the LQG setting, **Separation Principle can be applied to more general systems, even when the system is not Linear, cost is not Quadratic, and noise is not Gaussian.** However, the nice formulas we will derive for the estimation and control gains only work when we make the LQG assumptions.

The separation principle is different from the notion that standard LQR has the same optimal policy as stochastic LQR. This fact is known as **certainty equivalence**. Unlike the separation principle, certainty equivalence requires the LQG assumptions. When we do not have linear dynamics, for example, both LQR and stochastic LQR can be solved using dynamic programming, but they will not necessarily have the same optimal policy.

2 Optimal Controller

2.1 Value Function Definition

First we define the information available at time t as i_t,

$$i_t := \{y_0, \dots, y_{t-1}, u_0, \dots, u_{t-1}\}. \tag{3}$$

At time t we define the following value function.

$$V_t(i_t) := \underset{u_t,\dots,u_{N-1}}{\text{minimize}}\, \mathbf{E}\left[\sum_{k=t}^{N-1}(x_k^\top Q x_k + u_k^\top Q u_k) + x_N^\top Q_f x_N \;\middle|\; i_t, u_t\right] \tag{4}$$

To reduce the confusion between a random variable and its realization, we denote random variables with capital letter and an instance of random variable with lower case letter.

2.2 Optimal Solution

Using the new notations, we can write the general version of the principle of optimality. See the supplementary notes for a proof of this result.

$$\begin{aligned} V_t(i_t) &= \min_u \ \mathbf{E}\Big[X_t^\top Q X_t + U_t^\top Q U_t + V_{t+1}(I_{t+1}) \;\Big|\; I_t = i_t, U_t = u\Big], \\ I_{t+1} &= \{I_t, Y_t, U_t\}. \end{aligned} \tag{5}$$

We know that at the terminal time step N, $V_N = \mathbf{E}[X_N^\top Q_f X_N \mid I_N = i_N] = \hat{x}_N^\top Q_f \hat{x}_N + \mathbf{tr}(Q\Sigma_N)$. Here we denote the conditional distribution of X_N given i_N as $X_N|i_N \sim \mathcal{N}(\hat{x}_N, \Sigma_N)$.

Next we want to prove that for all time t, $V_t(i_t)$ can be written as $V_t(i_t) = \hat{x}_t^\top P_t \hat{x}_t + r_t$. We will prove this by induction. We assume it is true at time $t+1$, then at time t, we have

$$\begin{aligned} V_t(i_t) &= \min_u \ \mathbf{E}\Big[X_t^\top Q X_t + U_t^\top Q U_t + V_{t+1}(I_{t+1}) \;\Big|\; I_t = i_t, U_t = u\Big] \\ &= \min_u \Big\{\hat{x}_t^\top Q \hat{x}_t + \mathbf{tr}(Q\Sigma_t) + u^\top Q u + \mathbf{E}\Big[\hat{X}_{t+1}^\top P_{t+1} \hat{X}_{t+1} \;\Big|\; I_t = i_t, U_t = u\Big] + r_{t+1}\Big\}, \end{aligned} \tag{6}$$

where, according to Kalman Filtering, we have

$$
\begin{aligned}
\hat{x}_{t+1} &= (A + L_t C)\hat{x}_t + Bu - L_t y_t \\
&= (A + L_t C)\hat{x}_t + Bu - L_t(Cx_t + v_t) \\
&= A\hat{x}_t + Bu - L_t(C(x_t - \hat{x}_t) + v_t) \\
&\sim \mathcal{N}(A\hat{x}_t + Bu, L_t(C\Sigma_t C^\top + V)L_t^\top)
\end{aligned} \tag{7}
$$

Plugging (7) into (6), we have

$$
\begin{aligned}
V_t(i_t) &= \min_u \left\{ \hat{x}_t^\top Q\hat{x}_t + \mathbf{tr}(Q\Sigma_t) + u^\top Qu + \mathbf{E}[\hat{X}_{t+1}^\top P_{t+1}\hat{X}_{t+1} | I_t = i_t, U_t = u] + r_{t+1} \right\} \\
&= \min_u \Big\{ \hat{x}_t^\top Q\hat{x}_t + u^\top Qu + (A\hat{x}_t + Bu)^\top P_{t+1}(A\hat{x}_t + Bu) \\
&\qquad\qquad + \mathbf{tr}(Q\Sigma_t) + \mathbf{tr}(P_{t+1}L_t(C\Sigma_t C^\top + V)L_t^\top) + r_{t+1} \Big\} \\
&= \min_u \left\{ \hat{x}_t^\top Q\hat{x}_t + u^\top Qu + (A\hat{x}_t + Bu)^\top P_{t+1}(A\hat{x}_t + Bu) \right\} \\
&\qquad\qquad + \mathbf{tr}(Q\Sigma_t) + \mathbf{tr}(P_{t+1}L_t(C\Sigma_t C^\top + V)L_t^\top) + r_{t+1}
\end{aligned}
$$

The last three terms do not depend on u, and the first three terms are **exactly the same as in the LQR problem**, hence the optimal u is the same as in LQR but we replace x_t with $\hat{x}_t$. This proves the separation principle.

Therefore, we compute P_t and K_t exactly as in LQR, and we can write:

$$
\begin{aligned}
V_t(i_t) &= \hat{x}_t^\top P_t \hat{x}_t + r_t \\
u_t &= K_t \hat{x}_t
\end{aligned} \tag{8}
$$

The optimal LQG controller is a *dynamical system* with an internal state $\hat{x}_t$ that estimates the internal state of the plant, x_t. We can draw it as a block diagram:

$$
u_t \longleftarrow \boxed{\begin{aligned} \hat{x}_{t+1} &= (A + L_t C)\hat{x}_t + Bu_t - L_t y_t \\ u_t &= K_t \hat{x}_t \end{aligned}} \longleftarrow y_t \tag{9}
$$

We can write the entire system diagram as Fig. 1, the optimal controller is in the red block.

3 Optimal cost of LQG controller

To consider the cost of the system, we need to go back to the definition of value function

$$
\begin{aligned}
J &= \hat{x}_0^\top P_0 \hat{x}_0 + r_0 \\
&= \hat{x}_0^\top P_0 \hat{x}_0 + \sum_{t=0}^{N-1} \Big(\mathbf{tr}(Q\Sigma_t) + \mathbf{tr}(P_{t+1}L_t(C\Sigma_t C^\top + V)L_t^\top) \Big) + \mathbf{tr}(Q_f \Sigma_N)
\end{aligned} \tag{10}
$$

This cost expression does not appear to be symmetrical. For example, it involves L_t but not K_t. It involves Q but not W. This is an illusion. It turns out we can manipulate this expression for the cost and show that it is actually symmetrical.

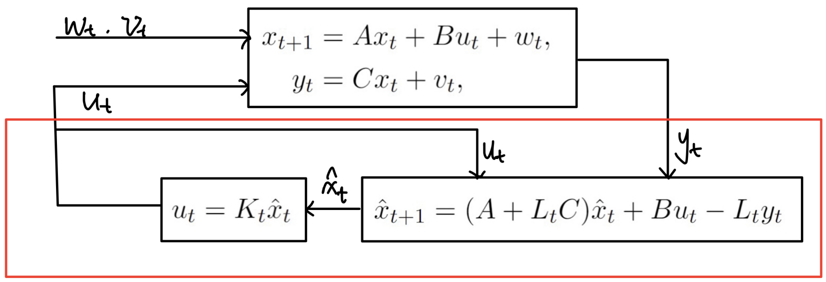

Figure 1: Optimal LQG controller.

Recall the following relationships.

$$
\begin{aligned}
L_t &= A\Sigma_t C^\top (C\Sigma_t C^\top + V)^{-1} \\
\Sigma_{t+1} &= A\Sigma_t A^\top + W - A\Sigma_t C^\top (C\Sigma_t C^\top + V)^{-1} C\Sigma_t A^\top \\
&= A\Sigma_t A^\top + W - L_t (C\Sigma_t C^\top + V) L_t^\top
\end{aligned}
$$

If $X_0 \sim \mathcal{N}(0, \Sigma_x)$, we replace $\hat{x}_0^\top P_0 \hat{x}_0$ by $\mathbf{tr}(P_0\Sigma_x)$ and obtain Therefore (10) can be written as

$$
\begin{aligned}
J &= \mathbf{tr}(P_0\Sigma_x) + \sum_{t=0}^{N-1} \Big(\mathbf{tr}(Q\Sigma_t) + \mathbf{tr}(P_{t+1}(A\Sigma_t A^\top + W - \Sigma_{t+1})) \Big) + \mathbf{tr}(Q_f \Sigma_N) \\
&= \mathbf{tr}(P_0\Sigma_x) + \sum_{t=0}^{N-1} \Big(\mathbf{tr}(Q\Sigma_t) + \mathbf{tr}(P_{t+1}A\Sigma_t A^\top) + \mathbf{tr}(P_{t+1}W) - \mathbf{tr}(P_{t+1}\Sigma_{t+1}) \Big) + \mathbf{tr}(Q_f \Sigma_N)
\end{aligned}
$$

This expression is now symmetric in P and Σ. We can also use the LQR Riccati equation to obtain an expression involving W, Σ_t, B, and K_t analogous to (10).

4 Stability of closed-loop system

We use the notion of steady-state LQG controller to denote the system where we replace P_t, K_t, Σ_t, L_t with their steady-state versions. The dynamics of the steady-state system are

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + w_t \\
y_t &= Cx_t + v_t \\
\hat{x}_{t+1} &= (A + LC)\hat{x}_t + Bu_t - Ly_t \\
u_t &= K\hat{x}_t
\end{aligned}
\tag{11}
$$

We can eliminate two internal signals y_t, u_t and get,

$$
\begin{aligned}
x_{t+1} &= Ax_t + BK\hat{x}_t + w_t \\
\hat{x}_{t+1} &= (A + LC)\hat{x}_t + BK\hat{x}_t - LCx_t + v_t
\end{aligned}
\tag{12}
$$

Combining the two states as one vector, we get,

$$\begin{bmatrix} x_{t+1} \\ \hat{x}_{t+1} \end{bmatrix} = \begin{bmatrix} A & BK \\ -LC & A+BK+LC \end{bmatrix} \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & -L \end{bmatrix} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \tag{13}$$

To examine the stability of the system, we perform a change of coordinates.

$$\begin{bmatrix} x_t \\ e_t \end{bmatrix} = \underbrace{\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}}_{T} \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix} \tag{14}$$

The system dynamics can be transformed using T to obtain

$$\begin{aligned} \begin{bmatrix} x_{t+1} \\ e_{t+1} \end{bmatrix} &= \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} A & BK \\ -LC & A+BK+LC \end{bmatrix} \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}^{-1} \begin{bmatrix} x_t \\ e_t \end{bmatrix} + \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -L \end{bmatrix} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \\ &= \begin{bmatrix} A+BK & -BK \\ 0 & A+LC \end{bmatrix} \begin{bmatrix} x_t \\ e_t \end{bmatrix} + \begin{bmatrix} I & 0 \\ I & L \end{bmatrix} \begin{bmatrix} w_t \\ v_t \end{bmatrix} \end{aligned} \tag{15}$$

State transformations do not change the eigenvalues, so the eigenvalues of the closed-loop map are precisely those of $A+BK$ and $A+LC$, which we know are stable by design since they come from the LQR controller and Kalman filter, respectively.

In this section, we made no assumption that the L and K in (11) were chosen optimally. So as long as K and L are chosen such that $A+BK$ and A_LC are stable, the controller will have an observer-regulator structure as in Fig. 1 and the closed-loop map will be stable.