

Note on LQG

Jean Walrand¹

Abstract

This note explains the LQG control problem. The key ideas are that the value function is quadratic, the optimal control is linear, and certainty equivalence.

Keywords

SDPE, Certainty Equivalence

¹walrand@berkeley.edu

Contents

| | | |
|---|----------------|---|
| 1 | Summary | 1 |
| 2 | Simple Model | 1 |
| 3 | Multiple Steps | 2 |

1. Summary

This note explains the LQG control results.

- We consider a one step control problem where the cost of the state after the step is quadratic. We show that the optimal control is linear and that the minimum cost is again quadratic.
- We show certainty equivalence in one step. It follows from the fact that the multiplier of the initial state does not depend on the noise variance. Thus, this noise could be an estimation error.
- We then look at a multi-step problem. If the minimum cost-to-go with n steps is quadratic, the optimal control is linear and the minimum cost-to-go with $n + 1$ steps is again quadratic. Thus, the minimum cost-to-go is always quadratic and the optimal control is linear at each step.
- Certainty equivalence follows automatically.
- The result is cool, but remember that certainty equivalence is the exception, not the rule!

2. Simple Model

Consider the following system:

$$X_1 = aX_0 + V + U.$$

In this model, X_0 is known and $V = \mathcal{N}(0, \sigma^2)$ is independent of X_0 and not observed. Also, U is a value that we will choose to minimize

$$C(X_0, U) = E[c + dX_1^2 + \alpha U^2]$$

where $\alpha > 0$ is given.

Interpretation

The interpretation is that X_1 is the state of some system that we want to get close to zero by exercising some control U . However, there is a tradeoff between how close we get the state to zero and the cost of the control. This is a cartoon version of the problem of getting a system to follow a desired trajectory while having to pay for the effort of controlling it.

The *myopic* strategy of minimizing the cost of the first step by choosing $U = 0$ is quite obviously not optimal because it results in a costly value of X_1^2 . In life, we tend to sacrifice the current pleasure in the hope of a better future. That's probably why you are in 126.

We find that

$$C(X_0, U) = E[c + d(aX_0 + V + U)^2 + \alpha U^2].$$

Taking the partial derivative with respect to U we get

$$\begin{aligned} \frac{\partial}{\partial U} C(X_0, U) &= \frac{\partial}{\partial U} E[\dots] = E\left[\frac{\partial}{\partial U} \dots\right] \\ &= 2E[adX_0 + dV + (d + \alpha)U] \\ &= 2[adX_0 + (d + \alpha)U]. \end{aligned}$$

To minimize over U we should set this derivative to zero. We find that the minimizing value U^* of U is given by

$$U^* = -\frac{ad}{d + \alpha} X_0.$$

Note that this optimal control is linear in X_0 . Moreover, the coefficient of X_0 does not depend on σ^2 . This last fact will imply *certainty equivalence* a bit later.

Another observation is that if $\alpha \approx 0$, then $U^* = -aX_0$, which should be intuitively obvious. Also, if $\alpha \gg 1$, then $U^* \approx 0$, which is also clear. In general, U^* is in-between.

Also, the resulting minimum cost is

$$\begin{aligned} C^*(X_0) &= C(X_0, U^*) \\ &= E\left[c + d(aX_0 + V - \frac{ad}{d+\alpha}X_0)^2 + \alpha \frac{a^2d^2}{(d+\alpha)^2}X_0^2\right] \\ &= c + d\sigma^2 + \frac{a^2\alpha d}{\alpha+d}X_0^2 \end{aligned}$$

Thus, if the cost $c + dX_1^2$ of being in X_1 is quadratic in X_1 , then the optimal control starting from X_0 is also quadratic in X_0 . This fact will imply that the result extends to multiple steps, as we will see shortly.

Certainty Equivalence

Assume that, instead of knowing X_0 , we have an estimate \hat{X}_0 with an estimation error that is $\mathcal{N}(0, w^2)$. That is,

$$X_0 = \hat{X}_0 + W$$

where $W = \mathcal{N}(0, w^2)$ is independent of V . In this case, we can write

$$X_1 = \hat{X}_0 + W + V + U = \hat{X}_0 + Z + U$$

where $Z = \mathcal{N}(0, \sigma^2 + w^2)$ is independent of \hat{X}_0 . To minimize over U

$$E[a + dX_1^2 + \alpha U^2],$$

we note that

$$U^* = -\frac{ad}{d+\alpha}\hat{X}_0.$$

Indeed, the problem is the same as before, except that now we know \hat{X}_0 and that σ^2 has been replaced by $\sigma^2 + w^2$. Thus, the optimal control is the same as when we know X_0 , except that X_0 is now replaced by its estimate \hat{X}_0 . This result is called *certainty equivalence*: do as if \hat{X}_0 were X_0 , i.e., as if you had no uncertainty about X_0 .

As expected, the uncertainty increases the optimal cost. Since σ^2 is replaced by $\sigma^2 + w^2$, we see that the optimal cost becomes

$$C^*(\hat{X}_0) = c + d(\sigma^2 + w^2) + \frac{a^2\alpha d}{\alpha+d}\hat{X}_0^2.$$

Thus, for a given value of \hat{X}_0 , we see that the cost increases linearly in the variance w^2 of the estimation error. The **moral of the story** is clear. Try to get as good an estimate of X_0 as you can. However, if you pay to improve the variance of the estimate, keep in mind the tradeoff between the cost dw^2 of uncertainty and the cost of reducing w .

A word of caution. In many situations, the optimal action depends on the degree of uncertainty. For instance, say that

the average value of Apple's stock in one month will be \$185 instead of \$171 today. If you are sure of that value, buy as much as you can afford. Now, if I tell you that it is equally likely to be \$220 or \$150, that should probably change your decision!

However, remarkably, for our simple Linear Quadratic Gaussian control problem, we saw that certainty equivalence holds.

3. Multiple Steps

We extend our simple example to one with multiple steps. Our detailed analysis of the simple model will pay off nicely here.

Assume that

$$X_{n+1} = aX_n + V_n + U_n.$$

Here, the V_n are i.i.d. $\mathcal{N}(0, \sigma^2)$ and independent of X_0 . We are allowed to choose U_n to be a function of X_n . That is, we observe the state at time n and then choose the control value U_n .

The goal is to minimize

$$E\left[\sum_{n=0}^N (X_n^2 + \alpha U_n^2)\right].$$

Stochastic Dynamic Programming Equations

We know X_0 and the minimum cost will be some function of X_0 . Let $V_N(X_0)$ be that minimum cost. The expected cost of $N+1$ steps is the cost of the initial step 0 plus the expected cost of the remaining N steps. Therefore, the minimum expected cost of $N+1$ steps is the minimum of the cost of step 0 plus the minimum expected cost of the remaining N steps.

Therefore,

$$\begin{aligned} V_N(X_0) &= \min_{U_0} \{X_0^2 + \alpha U_0^2 + E[V_{N-1}(X_1)]\} \\ &= \min_{U_0} \{X_0^2 + \alpha U_0^2 + E[V_{N-1}(aX_0 + V_0 + U_0)]\} \end{aligned}$$

where $V = \mathcal{N}(0, \sigma^2)$.

More generally, we have

$$V_{n+1}(X) = \min_U \{X^2 + \alpha U^2 + E[V_n(aX + V + U)]\}$$

Note that $V_0(X) = \min_{U_0} (X_0^2 + \alpha U_0^2) = X_0^2$.

Assume that $V_n(X) = c_n + d_n X^2$. Then, observe that finding the minimizing value of U in the expression above and the corresponding minimum cost V_{n+1} is precisely the simple problem that we studied in the previous section. We know that

$$U^* = -\frac{ad_n}{d_n + \alpha}X \tag{1}$$

and also that

$$V_{n+1}(X) = c_n + d_n \sigma^2 + \left(1 + \frac{a^2\alpha d_n}{\alpha + d_n}\right)X^2.$$

Thus, $V_{n+1}(X) = c_{n+1} + d_{n+1}X^2$ with

$$c_{n+1} = c_n + d_n \sigma^2 \quad (2)$$

$$d_{n+1} = 1 + \frac{a^2 \alpha d_n}{\alpha + d_n}. \quad (3)$$

Consequently, we have solved the multi-step problem. Indeed, at step 0, we are minimizing $V_N(X_0)$ and the optimal control is (1) with $n = N$. At step 1, this is the same but with $n = N - 1$ since there are $N - 1$ steps after step 1. Thus, at time m , there are $N - m$ steps to go and

$$U_m^* = -\frac{ad_{N-m}}{d_{N-m} + \alpha} X_m$$

where we find the constants d_k by iterating (3).

Certainty Equivalence

Observe that in the expressions (3), the variance σ^2 of the noise V_n does not enter. Thus, as in our simple example, if we observe \hat{X}_n which differs from X_n by some $\mathcal{N}(0, \sigma_n^2)$ error, then the optimal control is

$$U_m^* = -\frac{ad_{N-m}}{d_{N-m} + \alpha} \hat{X}_m$$

where the factor of \hat{X}_m is exactly the same as when if we had observed X_m . Thus, certainty equivalence prevails!

Of course, the cost increases with the estimation error variance, as in the simple case.

How do we get \hat{X}_m ? Funny you should ask. Assume that we observe

$$Y_{n+1} = cX_{n+1} + W_n$$

where the W_n are $\mathcal{N}(0, w^2)$ and independent of X_0 and the V_n . Then, we get \hat{X}_n as the output of the Kalman filter.

Figure 1 illustrates the result.

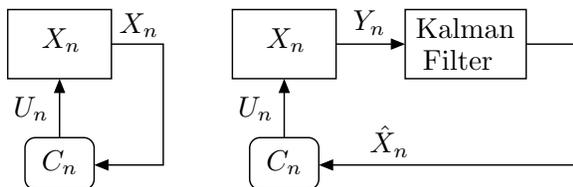


Figure 1. Certainty Equivalence: Complete observations (left) and noisy observations (right).