# ORF 418 : OPTIMAL LEARNING

## LECTURE 3 : September 10, 2025

### LINEAR QUADRATIC REGULATOR

1) General Formulation
2) An example
3) Solution to the example

# I. GENERAL FORMULATION.

This is a deterministic problem with known functions. So there is no learning nor randomness.

The state $x_t \in \mathbb{R}^d$ is a d-vector.

The control $u_t \in \mathbb{R}^\ell$ is a $\ell$-vector

The admissible controls $\mathcal{U}$ are any processes.

## a) Dynamics

Given a control process $u_0, u_1, \ldots$ and an initial condition $x_0$, the state process solves

$$x_{k+1} = Ax_k + Bu_k, \qquad k=0,1,\ldots$$

where $A$ is a $(d \times d)$ matrix and $B$ is $(d \times \ell)$ matrix both known.

Note that there is no issue of solvability or uniqueness. Indeed

$$x_1 = Ax_0 + Bu_0$$

$$x_2 = Ax_1 + Bu_1 = A^2 x_0 + ABu_0 + Bu_1$$

$$x_3 = Ax_2 + Bu_2 = A^3 x_0 + A^2 Bu_0 + ABu_1 + Bu_2$$

$$\ldots$$

b) Cost Functional: infinite horizon

$$J_\infty(x_0,u) = \sum_{k=1}^{\infty} \rho^k \left(x_k^T M x_k + u_k^T N u_k\right)$$

where $\rho > 0$ is the discount factor, $M$ is a $(d \times d)$ matrix, and $N$ is a $(\ell \times \ell)$ matrix. All are known.

and we assume that $M$ and $N$ are symmetric and positive definite. Equivalently all eigenvalues of $M$ and $N$ are strictly positive. In particular, there is a constant $c_*$ such that

$$x^T M x \geq c_* |x|^2 , \quad u^T N u \geq c_* |u|^2 .$$

c) Cost Functional: finite horizon $n$

$$J_n(x_0,u) := \sum_{k=1}^{n-1} \left(x_k^T M x_k + u_k^T N u_k\right) + x_n^T \hat{M} x_n$$

where $M, N$ are as before and $\hat{M}$ is another $(d \times d)$, symmetric, positive definite matrix.

---

# II. AN EXAMPLE

Just to fix the ideas, we consider a one-dimensional example with $d=\ell=1$, $A=4$, $B=M=N=1$, $\rho=1/2$, $x_0=1$

Then, the problem is to

$$\text{minimize} \sum_{k=0}^{\infty} \left(\tfrac{1}{2}\right)^k \left(x_k^2+u_k^2\right)$$

subject to

$$x_{k+1}=4x_k+u_k, \qquad k=0,1,\dots$$

## III SOLUTION TO THE EXAMPLE

We can choose $u_k$ as a constant multiple of $x_k$ itself : $u_k=-f\,x_k$ for some constant $f$.

Then,

$$x_{k+1}=(4-f)x_k \qquad k=0,1,\dots$$

Since $x_0=1$, we have

$$x_k=(4-f)^k \qquad k=0,1,\dots$$

This implies that

$$J=\sum_{k=0}^{\infty}\left(\tfrac{1}{2}\right)^k\left(x_k^2+(-f x_k)^2\right)$$

$$=\sum_{k=0}^{\infty}\left(\tfrac{1}{2}\right)^k(1+f^2)x_k^2$$

$$=(1+f^2)\sum_{k=0}^{\infty}\left(\tfrac{1}{2}\right)^k(4-f)^{2k} \qquad \text{geometric sum}$$

$$=(1+f^2)\sum_{k=0}^{\infty}\left(\frac{(4-f)^2}{2}\right)^k$$

$$= (1+f^2)\,\frac{1}{1-\frac{(4-f)^2}{2}}$$

$$= \frac{2(1+f^2)}{8f-f^2-14}.$$

This sum is finite if $(4-f)^2<2$. To minimize, we differentiate:

$$0 = \frac{4f}{8f-f^2-14} - \frac{2(1+f^2)(8-2f)}{(8f-f^2-14)^2}$$

$$\Rightarrow \quad f(8f-f^2-14) = (1+f^2)(4-f)$$

$$\Rightarrow \quad 8f^2-f^3-14f = 4-f+4f^2-f^3$$

$$\Rightarrow \quad 4f^2-13f-4=0$$

We numerically compute that

$$f^* \approx 3.53305.$$

This is indeed the optimal control as we prove generally in the next lecture.

general initial data.

Instead of $x_0=1$, we now consider a general $x_0$.

By the same reasoning

$$x_k = (4-f)^k x_0.$$

and
$$J = \frac{2(1+f^2)}{8f - f^2 - 14} x_0^2 .$$

Hence the optimal $f$ does not depend on $x_0$.

## finite horizon

Consider the problem
$$\text{minimize} \quad J_n = \sum_{k=0}^{n-1} (x_k^2 + u_k^2)$$
subject to
$$x_{k+1} = 4x_k + u_k .$$

Since the problem is no longer time-homogenous there is no reason for optimal gains constant $f$ to be independent of time. But we can solve for this constant recursively in time.

$n=1$ Problem is simple:
$$\text{minimize} \quad x_0^2 + u_0^2 \quad \text{over } u_0 .$$
Then, $u_0^* = 0$ and the minimum value is
$$v_1(x_0) := \inf_{u_0} x_0^2 + u_0^2 = x_0^2 .$$

$n=2$ The problem is

$$V_2(x_0) := \underset{u_0, u_1}{\text{minimum}}\{(x_0^2+u_0^2)+(x_1^2+u_1^2)\}$$

subject to $x_1 = 4x_0 + u_0$. Hence,

$$V_2(x_0) = \min_{u_0}\{x_0^2+u_0^2+\underbrace{\min_{u_1}(x_1^2+u_1^2)}_{V_1(x_1)=(4x_0+u_0)^2}\}$$

$$= \min_{u_0}\{x_0^2+u_0^2+(4x_0+u_0)^2\}$$

By calculus $u_0^* = -2x_0$ (and we also have $u_1^*=0$).

Hence,

$$V_2(x_0) = x_0^2+(-2x_0)^2+(4x_0-2x_0)^2 = 9x_0^2.$$

$n=3$

$$V_3(x_0) := \underset{u_0,u_1,u_2}{\text{minimum}}\{x_0^2+u_0^2+x_1^2+u_1^2+x_2^2\}.$$

$$= \underset{u_0}{\text{minim}}\{x_0^2+u_0^2+\underbrace{\underset{u_1,u_2}{\text{minimize}}\{x_1^2+u_1^2+(4x_1+u_1)^2\}}_{=9x_1^2 \text{ from case } n=2 = 9(4x_0+u_0)^2}\}$$

and $u_1^* = -2x_1$, $u_0^*=0$.

$$= \underset{u_0}{\text{minimize}}\{x_0^2+u_0^2+9(4x_0+u_0)^2\}$$

$$u_0^* + 9(4x_0+u_0^*) = 0 \Rightarrow u_0^* = -3.6x_0$$

This is dynammic programming.