Optimal Control and LQR Example

In stochastic optimal control problems, the state evolves as

$$x_{t+1} = f(x_t, u_t, w_t),$$

where x_t is the state, u_t is the control, and w_t is a random disturbance. Because of the randomness, we cannot pre-compute a fixed sequence of actions $u_0, u_1, \ldots, u_T$. Instead, we need a *policy (a rule)*, i.e. a function mapping the current state to a control:

$$u_t = U_t^\pi(x_t).$$

Linear Quadratic Regulator (LQR)

A classical special case is the Linear Quadratic Regulator (LQR), where:

- **Linear dynamics:**
$$x_{t+1} = Ax_t + Bu_t,$$
- **Quadratic cost:**
$$L_t(x_t, u_t) = x_t^\top Q x_t + u_t^\top R u_t,$$
with $Q, R \succeq 0$ (positive semidefinite matrices).

Dynamic programming shows that the value function has quadratic form

$$V_t(x) = x^\top P_t x,$$

and the optimal control is a linear feedback policy:

$$u_t^* = -K_t x_t,$$

with feedback matrix

$$K_t = (R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A.$$

Numerical Example (1D case)

Consider a one-dimensional system:

$$x_{t+1} = x_t + u_t, \quad A = 1,\ B = 1,$$

with cost

$$L(x, u) = x^2 + u^2, \quad Q = 1,\ R = 1.$$

Finite-horizon solution (T=1). At the last step ($t = 1$), the cost is

$$V_1(x) = x^2.$$

At $t = 0$, Bellman's equation gives

$$V_0(x) = \min_u \big(x^2 + u^2 + (x+u)^2\big).$$

Expand:

$$V_0(x) = \min_u \big(2x^2 + 2u^2 + 2xu\big).$$

Differentiate w.r.t. u:

$$\frac{d}{du}(2x^2 + 2u^2 + 2xu) = 4u + 2x.$$

Set equal to 0:

$$4u + 2x = 0 \quad \Rightarrow \quad u^* = -\tfrac{1}{2}x.$$

Interpretation. The optimal feedback policy is proportional:

$$u^* = -Kx, \quad K = \tfrac{1}{2}.$$

If the state (speed error, inventory deviation, etc.) is positive, apply a negative control to reduce it; if the state is negative, apply a positive control. The further the state is from zero, the stronger the correction.