II. Dynamic Programming

Case Study: Linear Quadratic Regulator (LQR)

Note $\gamma = 1$

Consider minimize $J = \sum_{k=0}^{\infty} \left[x_k^T Q x_k + u_k^T R u_k \right]$

subject to: $x_{k+1} = A x_k + B u_k \quad k = 0, 1, \ldots$

$x_0 = x_{init}$

where $Q = Q^T \underbrace{\geq 0}_{\text{positive semi-definite (psd)}}$, $R = R^T \underbrace{> 0}_{\text{p.d.}}$

We will solve the LQR problem w/ DP, to arrive at the Discrete-Time Algebraic Riccati Eqns

We will discover

- $V(x_k) = x_k^T P x_k$ (quadratic) $P = P^T \geq 0$ $P \in \mathbb{R}^{n \times n}$
- $\pi^*(x_k) = K \cdot x_k$ (linear) $K \in \mathbb{R}^{p \times n}$

Bellman's Optimality Eqn for $V(x_k) = x_k^T P x_k$

is ... $V(x_k) = c(x_k, u_k) + \gamma \cdot V(x_{k+1})$

apply to LQR $x_k^T P x_k = (x_k^T Q x_k + u_k^T R u_k) + 1 \cdot x_{k+1}^T P x_{k+1}$

substitute $u_k = K x_k$ ...

$$x_k^T P x_k = x_k^T \left[Q + K^T R K + (A + BK)^T P (A + BK) \right] x_k$$

must hold for all x_k

We have the "Lyapunov" matrix eqn:

$$(A+BK)^T P (A+BK) - P + Q + K^T R K = 0$$ linear in P

If K is fixed, then we can use the Lyap eqn to find P.

This yields $P = P^T \geq 0$ such that $V(x_k) = x_k^T P x_k$

$$V(x_k) = \sum_{\tau=k}^{\infty} x_\tau^T Q x_\tau + u_\tau^T R u_\tau \;\Big|\; u_\tau = K x_\tau$$

$$= \sum_{\tau=k}^{\infty} x_\tau^T \left[Q + K^T R K\right] x_\tau = x_k^T P x_k$$

To find an expression for K, write Bellman's optimality eqn

$$x_k^T P x_k = \min_{w} \{ x_k^T Q x_k + w^T R w + (A x_k + B w)^T P (A x_k + B w) \}$$

differentiate w.r.t. w and set to zero.

$$2Rw + B^T P (A x_k + B w) = 0$$

$$\Rightarrow w^* = \underbrace{-(R + B^T P B)^{-1} B^T P A}_{= K} \cdot x_k$$

Thus we have $u_k^* = K \cdot x_k$ where K

Substitute back into the Bellman eqn

... and simplify to yield

$$A^T P A - P + Q - A^T P B (R + B^T P B)^{-1} B^T P A = 0$$

which is quadratic in P.

This is the discrete-time algebraic Riccati Eqn (DARE)

Summary of Infinite-Time LQR

$u_k^{lqr} = K \cdot x_k$ where $K = -(R + B^T P B)^{-1} B^T P A$

and P solve DARE:

$$A^T P A - P + Q - A^T P B (R + B^T P B)^{-1} B^T P A = 0$$

Value Fcn: $V(x_k) = x_k^T P x_k$