Case Study of Policy Iteration: LQR

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

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

Recall:

- $V(x_k) = x_k^T P x_k$
- $u_k = K \cdot x_k$

Let's apply the policy eval & policy improvement steps.

① Policy Evaluation: " $V^{j+1} = c(\cdot,\cdot) + \gamma V^j$ "

$$x_k^T P^{j+1} x_k = x_k^T Q x_k + \underbrace{u_k^T}_{u=Kx} R u_k + \underbrace{x_{k+1}^T}_{=Ax+Bu} P^j x_{k+1} \quad (V = x^T P x)$$

$$x_k^T P^{j+1} x_k = x_k^T \left[Q + K^T R K + (A+BK)^T P^j (A+BK) \right] x_k \quad \forall x_k$$

$$\Rightarrow P^{j+1} = Q + K^T R K + (A+BK)^T P^j (A+BK) \quad \text{for some } K$$

which provides an iterative algo to the Lyapunov Eqn.

② Policy Improvement: Recall

from step ①

$$w^* = \arg\min_{w} \left\{ x_k^T Q x_k + w^T R w + (A x_k + B w)^T P^{OLD} (A x_k + B w) \right\}$$

Differentiate w.r.t. w, set to zero, & re-arrangement

$$w^* = \underbrace{-\left[R + B^T P^{OLD} B\right]^{-1} B^T P^{OLD} A}_{\overset{\text{def}}{=} K^{NEW}} \cdot x_k$$

Summary of Iterative Method to Solve Infinite-time LQR

$$u_k = K^{j+1} \cdot x_k \quad \leftarrow \text{control}$$

$$K^{j+1} = -\left[R + B^T P^{j+1} B\right]^{-1} B^T P^{j+1} A \quad \leftarrow \text{policy improve}$$

$$P^{j+1} = Q + (K^j)^T R K^j + (A + BK^j)^T P^j (A + BK^j) \quad \leftarrow \text{policy evaluation}$$

These eqns are known as Hewer's method [Hewer 1971]

VALUE ITER (VI)