Discrete Time Optimal Control

Given a dynamical system:

$$x_{k+1} = f_k(x_k, u_k) \qquad\qquad x_0 = x_{init}$$

Find the control sequence $u_k = u(t_k)$, $k=0,\ldots,N-1$, which minimizes :

$$J = \sum_{k=0}^{N-1} l_k(x_k, u_k) + V_T(x_N) \qquad N \ total \ "stages"$$

Solution: introduce Lagrange multipliers $\lambda_0, \lambda_1, \ldots, \lambda_N$ for all constraints

$$J = \sum_{k=0}^{N-1} l_k(x_k, u_k) + \lambda_{k+1}^T(f_k(x_k, u_k) - x_{k+1}) + V_T(x_N) + \lambda_0^T(x_0 - x_{init})$$
$$= \sum_{k=0}^{N-1} H_k(x_k, u_k, \lambda_{k+1}) - \lambda_{k+1}^T x_{k+1} + V_T(x_N) + \lambda_0^T(x_0 - x_{init})$$

$$H_i(x_i, u_i, \lambda_{i+1}) \equiv l_i(x_i, u_i, \lambda_{i+1}) + \lambda_{i+1}^T f_i(x_i, u_i)$$

Extremize the augmented cost w.r.t.

$$u_k \ (k = 0, \dots, N-1); \qquad x_k \ (k = 0, \dots, N); \qquad \lambda_k \ (k = 0, \dots, N);$$

Since there are a discrete number of "states", we can use the classical finite dimensional extremal criteria:

$$\frac{\partial J}{\partial x_k} = 0; \quad \Rightarrow \quad \lambda_k = \frac{\partial H_k}{\partial x_k}(x_k, u_k, \lambda_{k+1}) \qquad k = 1, \dots, N-1$$

$$\frac{\partial J}{\partial x_N} = 0; \quad \Rightarrow \quad \lambda_N = \left.\frac{\partial V_T}{\partial x_N}\right|_{x_N}$$

$$\frac{\partial J}{\partial x_0} = 0; \quad \Rightarrow \quad x_0 = x_{init}$$

$$\frac{\partial J}{\partial u_i} = 0; \quad \Rightarrow \quad \frac{\partial H_i}{\partial u_i}(x_i, u_i, \lambda_{i+1}) = 0 \qquad i = 0, \dots, N-1$$

$$\frac{\partial J}{\partial \lambda_m} = 0; \quad \Rightarrow \quad x_{m+1} = f(x_m, u_m) \qquad m = 0, \dots, N-1$$

Discrete TIME LQR (a different derivation)

A linear dynamical system: $x_{k+1} = A_k x_k + B_k u_k \qquad x_0 = x_{init}$

Find the control sequence u_k which minimizes :

$$J = \frac{1}{2} x_N^T P_T x_N + \frac{1}{2} \sum_{k=0}^{N-1} x_k^T Q_k x_k + u_k^T R_k u_k$$

Solution: *Assuming* $u_k = -K_k x_k$, then $x_{k+1} = (A_k - B_k K_k) x_k \equiv \boxed{A_k^C} x_k$

$$J = \frac{1}{2} x_N^T P_T x_N + \frac{1}{2} \sum_{k=0}^{N-1} x_k^T (Q_k + K_k^T R_k K_k) x_k$$

"Cost-to-go" from t_p

Define:

$$J_p = \frac{1}{2} x_N^T P_T x_N + \frac{1}{2} \sum_{k=p}^{N-1} x_k^T (Q_k + K_k^T R_k K_k) x_k$$

Then note that :

$$\underbrace{J_{p+1} - J_p}_{\text{"Incremental" Cost}} = -\frac{1}{2} x_p^T (Q_p + K_p^T R_p K_p) x_p \qquad (\dagger\dagger)$$

Discrete TIME LQR (continued)

The controlled dynamical system will propagate as:

$$x_m = \underbrace{\Phi(m,k)}_{\text{Transition matrix}}\, x_k = \left(\prod_{i=k}^{m-1} A_i^C\right) x_k$$

The "cost-to-go" can be rewritten using the Transition Matrix:

$$J_p = \frac{1}{2} x_p^T \left(\Phi^T(N,p) P_T \Phi(N,p) + \sum_{k=p}^{N-1} \Phi^T(k,p)(Q_k + K_k^T R_k K_k)\Phi(k,p) \right) x_p$$

$$\equiv \frac{1}{2} x_p^T P_p x_p$$

P_p = "cost-to-go" (by abuse of language)

Then:

$$J_{p+1} - J_p = \frac{1}{2}(x_{p+1}^T P_{p+1} x_{p+1} - x_p^T P_p x_p) = \frac{1}{2} x_p^T \left((A_p^C)^T P_{p+1} A_p^C - P_p \right) x_p \quad (**)$$

Now equate (**) and (††), and then rearrange:

$$P_p = (A_p - B_p K_p)^T P_{p+1} \left(A_p - B_p K_p\right) + K_p^T R_p K_p + Q_p$$

Discrete TIME LQR (continued)

To derive the *optimal* feedback gain, we want to minimize the cost-to-go with respect to the gain:

$$\frac{\partial P_p}{\partial K_p} = 0; \quad \Rightarrow \quad \boxed{K_p \; = \; (R_p + B_p^T P_{p+1} B_p)^{-1} B_p^T P_{p+1} A_p}$$

Substituting the expression for K_p back into the expression for P_p yields:

$$P_p = Q_p + A_p^T P_{p+1} A_p - A_p^T P_{p+1} B_p (R_p + B_p^T \, P_{p+1} B_p)^{-1} B_p^T P_{p+1} A_p$$

$$\boxed{= Q_p + A_p^T \left[P_{p+1} - P_{p+1} B_p (R_p + B_p^T P_{p+1} B_p)^{-1} B_p^T P_{p+1} \right] A_p}$$

This is the discrete time Riccati equation (DRE).

– Because $(x_p{}^T P_p x_p)/2$ is the "cost-to-go", and at stage N the cost to go is $(x_N{}^T P_N x_N)/2$, the terminal value of P_N is given by: $P_N = P_T$