Exercise for Optimal control – Week 6

Choose **1.5** problems to solve.

Disclaimer
This is not a complete solution manual. For some of the exercises, we provide only partial answers, especially those involving numerical problems. If one is willing to use the solution manual, one should judge whether the solutions are correct or wrong by him/herself.

Exercise 1. Derive the policy iteration scheme for the LQR problem

$$\min_{u(\cdot)} \sum_{k=1}^{\infty} x_k^\top Q x_k + u_k^\top R u_k$$

with $Q = Q^\top \geq 0$ and $R = R^\top > 0$ subject to:

$$x_{k+1} = Ax_k + Bu_k.$$

Assume the system is stabilizable. Start the iteration with a stabilizing policy. Run the policy iteration and value iteration on a computer for the following matrices:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R = I$$

Compare the convergence rates of the two iterations scheme for the policies and value functions.

Solution. Choose $u_1(x) = K_1 x$ such that $A + BK_1$ is Schur. Then assume $J_k(x) = x^\top P_k x$. The policy iteration takes the form

$$x^\top P_1 x = x^\top Q x + x^\top K_1^\top R K_1 x + x^\top (A + BK_1)^\top P_1 (A + BK_1) x$$

or

$$P_1 = Q + K_1^\top R K_1 + (A + BK_1)^\top P_1 (A + BK_1)$$

This is a LME with unknown P_1. To get u_2, solve

$$\begin{aligned} u_2(x) &= \arg\min_u \{x^\top Q x + u^\top R u + (Ax + Bu)^\top P_1 (Ax + Bu)\} \\ &= -(B^\top P_1 B + R)^{-1} B^\top P_1 A x. \end{aligned}$$

Thus

$$K_2 = (B^\top P_1 B + R)^{-1} B^\top P_1 A.$$

Repeat this, one would get $u_k, u_{k+1}, \cdots$ and $J_k, J_{k+1}, \cdots$. The policy iteration – the controller gain K – has quadratic convergence rate, see [1], whereas value iteration has only linear convergence rate.

Exercise 2 (LQR for LTV systems)**.** Consider a controllable LTV system

$$\dot{x} = A(t)x + B(t)u$$

with $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and cost function

$$J = x(T)^\top Q_f x(T) + \int_{t_0}^{T} x(t)^\top Q(t) x(t) + u(t)^\top R(t) u(t) \mathrm{d}t$$

where Q_f, $Q(t) \geq 0$ and $R(t) > 0$ for all $t \geq 0$. In addition, we assume $A(\cdot)$, $B(\cdot)$, $Q(\cdot)$ and $R(\cdot)$ are continuous. The objective is to find an optimal control u^* such that J is minimized.

1) The dynamic programming works also for time varying systems. Write down the Hamiltonian $H(t, x, u, p)$ for this problem and derive the optimal controller using the verification rule. *Hint:* consider value function of the form $J^*(t, x) = x^\top P(t)x$.

2) Show that the HJB equation reduces to an ODE:

$$-\dot{P}(t) = Q(t) + P(t)A(t) + A(t)^\top P(t) - P(t)B(t)R(t)^{-1}B(t)^\top P(t). \tag{1}$$

with boundary condition

$$P(T) = Q_f.$$

3) Prove that the equation (1) has a unique symmetric semi-positive definite solution on interval $[0, T]$ for any $T > 0$. In particular, there is no finite escape time.

Solution. The Hamiltonian function is

$$H(x, u, p, t) = p^\top (A(t)x + B(t)u) - x^\top Q(t)x - u^\top R(t)u$$

We implement the first two steps of the verification rule:

Step 1: solve the maximization $\max_u H(x, u, p, t)$, resulting in

$$u^* = \arg\max_u H(x, u, p, t) = \frac{1}{2}R(t)^{-1}B(t)^\top p$$

and

$$H(x, u^*, p, t) = p^\top A(t)x - x^\top Q(t)x + \frac{1}{4}p^\top B(t)R(t)^{-1}B(t)^\top p$$

Step 2: Replace p by $-\frac{\partial V}{\partial x}$ to get the HJB

$$-V_t - V_x A(t)x - x^\top Q(t)x + \frac{1}{4}V_x B(t)R(t)^{-1}B(t)^\top V_x^\top = 0.$$

Consider a candidate $V(t, x) = x^\top P(t)x$. Then

$$u^* = \frac{1}{2}R(t)^{-1}B(t)^\top P(t)x$$

and the HJB equation reduces to

$$-\dot{P}(t) = Q(t) + P(t)A(t) + A(t)^\top P(t) - P(t)B(t)R(t)^{-1}B(t)^\top P(t). \tag{2}$$

with boundary condition

$$P(T) = Q_f.$$

The first order ODE (2) is called *differential Riccati equation* (DRE). Thus the continuous LQR problem on finite horizon reduces to solving the DRE (2).

Notice that the right hand side of (2) is quadratic in P (thus locally Lipschitz!) and that $A(\cdot)$, $B(\cdot)$, $Q(\cdot)$ and $R(\cdot)$ are continuous, therefore local existence and uniqueness of solutions are guaranteed. This also implies that the solution to (2) is symmetric: if $P(t)$ is a solution, so is $P(t)^\top$, while both have the same terminal condition, thus $P(t) = P(t)^\top$.

We show next that there is no finite escape time. Suppose that the maximum interval of existence of solutions to the DRE is $(t_1, T]$ for some finite $t_1 \in \mathbb{R}$. Then by construction, for any $t_2 \in (t_1, T]$, and $x(t_2) \in \mathbb{R}^n$, we have

$$-x(t)^\top \dot{P}(t)x(t) + H(x(t), u(t), -2P(t)x(t)) \leq 0, \tag{3}$$

for any $u(\cdot)$ since $-V_t + \sup H(x, u, p) = 0$. One can check that the equation (3) is

$$-\frac{d}{dt}x^\top(t)P(t)x(t) \leq x^\top(t)Q(t)x(t) + u^\top(t)R(t)u(t)$$

from which it follows that

$$x(t_2)^\top P(t_2)x(t_2) \leq x(T)^\top Q_f x(T) + \int_{t_2}^{T} x(t)^\top Q(t)x(t) + u(t)^\top R(t)u(t)\mathrm{d}t, \quad \forall u(\cdot). \tag{4}$$

(The inequality becomes equality for $u = u^*$, thus we also get $P(t_2) \geq 0$.) In particular, this is true for $u \equiv 0$. In this case,

$$x(T) = \Phi(T, t_2)x(t_2)$$

where $\Phi(T, t)$ is the state transition matrix of the system $\dot{x} = A(t)x$. Recall that A is continuous, thus $A(\cdot)$ as well as $\Phi(T, t)$ is bounded on $[t_1, T]$. Therefore, we may conclude from (4) that

$$x(t_2)^\top P(t_2)x(t_2) \leq c|x(t_2)|^2$$

for some c, depending on t_1. Therefore $P(t_2) < cI$ for all $t_2 \in (t_1, T]$ (no blow-up!). To show that P can be extended outside $(t_1, T]$, let t_2 be sufficiently close to t_1. Since $P(t)$ is uniformly bounded on $(t_1, T]$, then near t_2, the DRE is Lipschitz on $(t_1, t_2]$. Thus the solution can be extended to t_1. This is a contradiction since we assumed $(t_1, T]$ is the maximum interval of existence.

Exercise 3. 1) Derive the HJB equation for the time optimal control problem of the double integrator

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= u\end{aligned}$$

with initial condition $(1, 1)$ and terminal condition $(0, 0)$ under the constraint $|u| \leq 1$.

2) Solve the HJB equation using method of characteristics.

Solution. We already know the HJB equation for this problem is

$$-V_t - V_{x_1}x_2 + |V_{x_2}| - 1 = 0$$

with boundary condition

$$V(t_f, 0) = 0.$$

Remember $\tilde{H}(x, V_x) = H(x, u^*(x, V_x), -V_x) = -V_{x_1}x_2 + |V_{x_2}| - 1$, thus by the method of characteristics, $\dot{p} = -\tilde{H}_x$, or

$$\begin{aligned}\dot{p}_1 &= 0\\ \dot{p}_2 &= p_1\end{aligned}$$

which is exactly the costate equation from in maximum principle. Then one solve a BVP for the system (x_1, x_2, p_1, p_2) and then the HJB equation can be solved along the characteristic curve by integrating

$$\frac{d}{dt}V = p^\top F_p = -p_1 x_2 + |p_2|$$

References

[1] Gary Hewer. An iterative technique for the computation of the steady state gains for the discrete optimal regulator. *IEEE Transactions on Automatic Control*, 16(4):382–384, 1971.