Introduction to Mathematical Foundations and Principles of Autonomy - Homework 1

Matthieu J. Capuano

Fall 2019

1 Problem 1

(a) We want to show that the open-loop control law that minimizes

$$J(x_0, u(\cdot)) = \frac{1}{2}\left[x^T(t_f)Sx(t_f) + \int_0^{t_f} [x^T(t)Qx(t) + u^T(t)Ru(t)]dt\right]$$

has form:

$$u(t) = -R^{-1}B^T\lambda(t)$$

We can do this by starting from the Hamilton-Jacobi-Bellman equation directly, however, we choose to begin from the more general nonlinear time-varying problem, and *then* derive the HJB equation to maximize learning (i.e. this author wants to do it to make sure he gets it, feel free to skip directly to "(Yayyyy!)" if this is irrelevant to you). This derivation is also paraphrased from Dr. Tsiotras' notes on deriving the HJB equation, so it will be a useful reference for problem 2. Thus, we begin with the more general nonlinear time-varying performance measure:

$$V = \int_t^T l(x, u, \tau)d\tau + m[x(T)]$$

and search for a formulation of some control law $u = \phi(x, t)$ that minimizes V subject to the nonlinear system dynamics:

$$\dot{x} = f(x, u, t)$$

Here $l(x, u, \tau)$ is our loss function, and $m(x)$ is the terminal cost. Define the minimal value of our total performance V to be $V^*(x, t)$ if we start at time t and state $x(t)$. Then we minimize our performance measure to obtain it's optimal value:

$$V^*(x, t) = \min_{t \leq u \leq T} \left\{ \int_t^T l(x, u, \tau)d\tau + m[x(T)] \right\}$$

(Equation 7 in Dr. Tsiotras' notes) At this point in our (Dr. Tsiotras') derivation, there are a couple of steps using the optimality principle and properties of integral such that we can instead focus on the time interval $[t, t + \Delta t]$ and get:

$$V^*(x,t) = \min_{t \le u \le t+\Delta t} \left\{ \int_t^{t+\Delta t} l(x,u,\tau) d\tau + V^*[x(t+\Delta t), t+\Delta t] \right\}$$

(Equation 11 in Dr. Tsiotras' notes) Effectively, since our original V^* was an optimal solution starting at time t down to time T, by the optimality principle, V^* must also be optimal over time interval $[t, t + \Delta t]$. We then expand the term $V^*[x(t + \Delta t), t + \Delta t]$ through a multivariate Taylor series expansion, and approximate the integral term with $l(x, u, \tau)\Delta t$, both of which are valid for very small Δt values. This gives us:

$$V^*(x,t) = \min_{u(t)} \left\{ l(x,u,t)\Delta t + V^*(x,t) + \frac{\partial V^*}{\partial t}\Delta t + \frac{\partial V^*}{\partial x} f(x,u,t)\Delta t + O(\Delta t) \right\}$$

$O(\Delta t)$ refers to second and higher order terms in Δt. We can pull $V^*(x,t)$ and $\frac{\partial V^*}{\partial t}\Delta t$ out of the minimization since they don't depend on u:

$$0 = \frac{\partial V^*}{\partial t}\Delta t + \min_{u(t)} \left\{ l(x,u,t)\Delta t + \frac{\partial V^*}{\partial x} f(x,u,t)\Delta t + O(\Delta t) \right\}$$

Finally, we can divide by Δt and take the limit as $\Delta t \longrightarrow 0$ to get the Hamilton-Jacobi-Bellman equation:

$$-\frac{\partial V^*}{\partial t} = \min_{u(t)} \left\{ l(x,u,t) + \frac{\partial V^*}{\partial x} f(x,u,t) \right\}$$

(Yayyyy!)

With this, we can solve the linear-quadratic problem, which we will show produces the desired matrix Riccati equation $u(t) = -R^{-1}B^T\lambda(t)$.For our LQ problem $l(x,u,t) = x^T(t)Qx(t) + u^T(t)Ru(t)$, $f(x,u,t) = Ax + Bu$, and $m(x) = x^T(t_f)Sx(t_f)$, so we are trying to find:

$$-\frac{\partial V^*}{\partial t} = \min_{u(t)} \left\{ x^T(t)Qx(t) + u^T(t)Ru(t) + \frac{\partial V^*}{\partial x}(Ax + Bu) \right\}$$

We can do this by taking the gradient of the term inside the minimization with respect to u and setting it equal to zero. We simplify notation with $u = u(t)$:

$$\begin{aligned} 0 &= \frac{\partial}{\partial u}\left(x^T(t)Qx(t) + u^T Ru + \frac{\partial V^*}{\partial x}(Ax + Bu) \right) && (1) \\ &= 0 + 2Ru + 0 + \frac{\partial V^*}{\partial x}B && (2) \\ -2Ru &= \frac{\partial V^*}{\partial x}B && (3) \\ \therefore u^* &= -\frac{1}{2}R^{-1}B^T\frac{\partial V^*}{\partial x} && (4) \end{aligned}$$

In the last step, we left multiply by the inverse of R, which we know exists since R is given to be positive definite. We also move B to the left of the gradient term by taking it's transpose, which we can do since B is a vector, so the operation is commutative. if we choose

$$\lambda(t) = \frac{1}{2}\frac{\partial V^*}{\partial x}$$

Then we have shown that the optimal open-loop control for the given problem is

$$u^*(t) = -R^{-1}B^T\lambda(T)$$

(b) Assuming that the solution to the HJB equation is given by

$$V^*(x,t) = x'P(t)x$$

for some matrix $P(t)$, we also have that:

$$\frac{\partial V^*(x,t)}{\partial x} = 2P(t)x$$

We are additionally given that $\dot{x} = Ax + Bu$ and our cost function is $l(x,u,t) = x^T(t)Qx(t) + u^T Ru$. Going back to our general definition of u from equation (4), we can substitute the $\frac{\partial V^*(x,t)}{\partial x}$ we just computed:

$$\begin{aligned} u^* &= -\frac{1}{2}R^{-1}B^T\frac{\partial V^*}{\partial x} && (5)\\ &= -\frac{1}{2}R^{-1}B^T 2P(t)x && (6)\\ u^* &= -R^{-1}B^T P(t)x && (7) \end{aligned}$$

We can start from our HJB equation with the minimization carried out (since we are substituting an equation for u^*, which is optimal) as shown below. Note that, again, we leave the time-variant nature of the states and controls implied: $x = x(t)$ and $u = u(t)$:

$$\begin{aligned} -\frac{\partial V^*}{\partial t} &= l(x,u^*,t) + \frac{\partial V^*}{\partial x}f(x,u^*,t) && (8)\\ &= \left(x^TQx + u^{*T}Ru^*\right) + \frac{\partial V^*}{\partial x}\left(Ax + Bu^*\right) && (9) \end{aligned}$$

We expand and simplify each of these terms individually.
The second term:

$$\begin{aligned} u^{*T}Ru^* &= \big(-R^{-1}B^TPx)^TR(-R^{-1}B^TPx) && (10)\\ &= \big(x^TPBR^{-1})R(R^{-1}B^TPx) && (11)\\ &= x^TPBR^{-1}B^TPx && (12) \end{aligned}$$

The third term:

$$\begin{aligned} \frac{\partial V^*}{\partial x}Ax &= 2PxAx && (13)\\ &= 2x^TPAx && (14)\\ &= x^T(A^TP + PA)x && (15) \end{aligned}$$

Where this last identity comes from the Matrix cookbook (https://www.math.uwaterloo.ca/ hwolkowi/matrixcookbook.pdf). The fourth term:

$$\begin{aligned} \frac{\partial V^*}{\partial x}Bu^* &= 2PxB(-R^{-1}B^TPx) && (16)\\ &= -2x^TPBR^{-1}B^TPx && (17) \end{aligned}$$

Substituting (13), (16), and (18) back into (9):

$$-\frac{\partial x^T P x}{\partial t} = x^T Q x + x^T P B R^{-1} B^T P x + x^T A^T P + P A x + -2 x^T P B R^{-1} B^T P x \tag{18}$$

$$-x^T \dot{P} x = x^T (Q + P B R^{-1} B^T P + A^T P + P A + -2 P B R^{-1} B^T P) x \tag{19}$$

$$-x^T \dot{P} x = x^T (Q + A^T P + P A + -P B R^{-1} B^T P) x \tag{20}$$

The boundary condition is:

$$V^*(x, t_f) = x^T(t_f) S x(t_f) = x^T S x$$

Equation (20) shows that the differential equation for $P(t)$ is:

$$\dot{P} = Q + A^T P + P A + -P B R^{-1} B^T P$$

With: $P(t_f) = S$

(c) We aim to confirm that the open-loop and closed-loop solutions from (a) and (b) produce the same optimal trajectories. We do so by plotting both control histories on the same figure.

2 Problem 2

Principle of Optimality: If $u^*(\tau)$ is optimal over the interval [t, T], starting at state $x(t)$, then $u^*(\tau)$ must be optimal over the sub-interval $[t + \Delta t, T]$ for any Δt such that $T - t \geq \Delta t > 0$.[1]

In our derivation of the HJB equation, which has been re-conducted in the first part of our solution to question 1, we use the optimality principle in two places. Looking at the lecture notes titled "Principal of Optimality and Derivation of the Hamilton-Jacobi-Bellman Equation", we first use it when generalizing the problem formulated by equation (1) (of the lecture notes), which optimizes over the interval $[t_0, t_f]$, to the more general case, which optimizes over the interval $[t, t_f]$, without loss of generality. Second, we use it in our derivation we going from optimizing:

$$V^*(x,t) = \min_{t \leq u \leq T} \left\{ \int_t^T l(x,u,\tau)d\tau + m[x(T)] \right\}$$

To optimizing:

$$V^*(x,t) = \min_{t \leq u \leq t+\Delta t} \left\{ \int_t^{t+\Delta t} l(x,u,\tau)d\tau + V^*[x(t+\Delta t), t+\Delta t] \right\}$$

Whereby we have shifted our focus from interval $[t, T]$ to interval $[t, t + \Delta t]$, a sub-interval of the original, through the Principle of Optimality.

Similarly, in the discrete-time case, we applied the Principle of Optimality when shifting our minimization from:

$$V(x,k) = \min_{\{u_k,\ldots,u_{N-1}\}} \left\{ \phi(x_N) + \sum_{j=k}^{N-1} L_j(x_j, u_j) \right\}$$

To the minimization of:

$$V(x,k) = \min_{u_k} \left\{ \min_{\{u_{k+1},\ldots,u_{N-1}\}} \left\{ \phi(x_N) + L_k(x_k, u_k) + \sum_{j=k+1}^{N-1} L_j(x_j, u_j) \right\} \right\}$$

Where k represents a discrete timestep. So we effectively shifted from minimizing over the interval $[k, N-1]$ to the sub-interval $[k+1, N-1]$. Note that these two equations are effectively the equivalents of equations (7) and (9) (if the first integral is moved inside the inner minimization) from the lecture notes respectively, for the continuous time case.

[1] This wording is slightly paraphrased from "Linear Quadratic Control, An Introduction" by Peter Dorato, Chaouki T. Abdallah, and Vito Cerone.

3 Problem 3

We complete the indicated portions of the code by adding the logic in the last page of this submission to the backpropagation loop in the DDP_iteration function within file DDP_discrete.m. Note that the longer lines have been split over multiple lines for readability.

We then plot our individual state variables, cost values, and controller gains over time (check next page). It is clear from our plots that the cart is trying to optimize, however, our cost plots indicate that it is failing to converge.