S26 AerE 5740: Optimal Control

Practice Set 09: Continuous Time Deterministic Finite Horizon DP

Instructor: Abhishek Halder

Problem 1.

Consider the finite horizon (fixed terminal time T) deterministic OCP in continuous time, given by

$$\begin{aligned}&\min_{\gamma\in\Gamma} \quad e^{-\beta T} q x(T) \;+\; \int_0^T e^{-\beta t}\left(x + cu^2\right)\,\mathrm{d}t\\ &\dot{x} = a - u\sqrt{x}, \quad x(0) = x_0 \in \mathbb{R}_{>0}, \quad x \in \mathcal{X} \equiv \mathbb{R}_{>0}, \quad u \in \mathcal{U} \equiv \mathbb{R}_{>0},\end{aligned}$$

where Γ denotes the set of all history-dependent randomized policies. The parameters $a, \beta, c, q > 0$. Our objective is to compute the optimal state feedback control $u_{\text{opt}} = \gamma_{\text{opt}}(t, x)$.

(a) Deriving HJB PDE IVP.

Derive the Dynamic Programming (DP) equation (i.e., the HJB PDE IVP) for the value function $V(t, x)$ associated with the above OCP.

Solution for part (a):

Here, the Lagrangian $L(t, x, u) = e^{-\beta t}\left(x + cu^2\right)$, and the controlled vector field $f(t, x, u) = a - u\sqrt{x}$, which result in the HJB PDE

$$\begin{aligned}&\frac{\partial V}{\partial t} + \min_{u\in\mathcal{U}}\left\{L(t,x,u) + \left\langle \frac{\partial V}{\partial x}, f(t,x,u)\right\rangle\right\} = 0,\\ \Rightarrow &\frac{\partial V}{\partial t} + \min_{u\in\mathcal{U}}\left\{e^{-\beta t}\left(x + cu^2\right) + \frac{\partial V}{\partial x}\left(a - u\sqrt{x}\right)\right\} = 0.\end{aligned}$$

Performing the above minimization, we get $u_{\text{opt}} = \dfrac{e^{\beta t}}{2c}\dfrac{\partial V}{\partial x}\sqrt{x}$, which upon substituting back in the DP equation, results in the 1st order nonlinear PDE IVP:

$$\frac{\partial V}{\partial t} + e^{-\beta t}x + a\frac{\partial V}{\partial x} - \frac{e^{\beta t}}{4c}\left(\frac{\partial V}{\partial x}\right)^2 x = 0, \quad V(T, x(T)) = e^{-\beta T} q x(T).$$

(b) Solution for the DP equation.

It is not possible to use the Hopf-Lax variaional formula to solve the PDE IVP in part (a) because the optimized Hamiltonian has explicit dependence on x and t, in addition to dependence on $\frac{\partial V}{\partial x}$. To solve this, make the structural guess

$$V(t,x) = e^{-\beta t}\left(A(t)x + B(t)\right),$$

where $A(t), B(t)$ are to-be-determined functions of time t. Derive the optimal state-feedback controller $u_{\text{opt}} = \gamma_{\text{opt}}(t,x)$.

Use this solution to numerically generate the following 3 figures: (i) trajectory plots: A, B as functions of t, (ii) surface plot: u_{opt} as function of t, x, (iii) surface plot: V as function of t, x. Use the data: $a = 1, \beta = 2, c = 3, q = 5, T = 2, x = 0.1 : 0.1 : 10$.

Solution for part (b):

Substituting the supplied anstaz for the value function $V(t,x)$ in the HJB PDE IVP derived in part (a), we get:

$$\left(\dot{A} - \beta A + 1 - \frac{A^2}{4c}\right)x \;+\; \left(\dot{B} - \beta B + aA\right) \;=\; 0, \quad e^{-\beta T}\left(A(T)x + B(T)\right) = qxe^{-\beta T}.$$

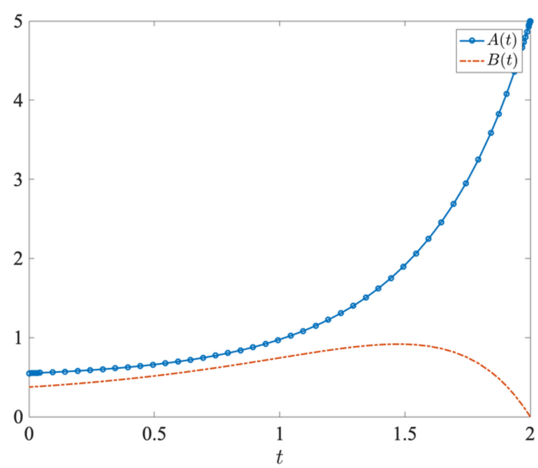
For the above to hold for all t, x, we must have

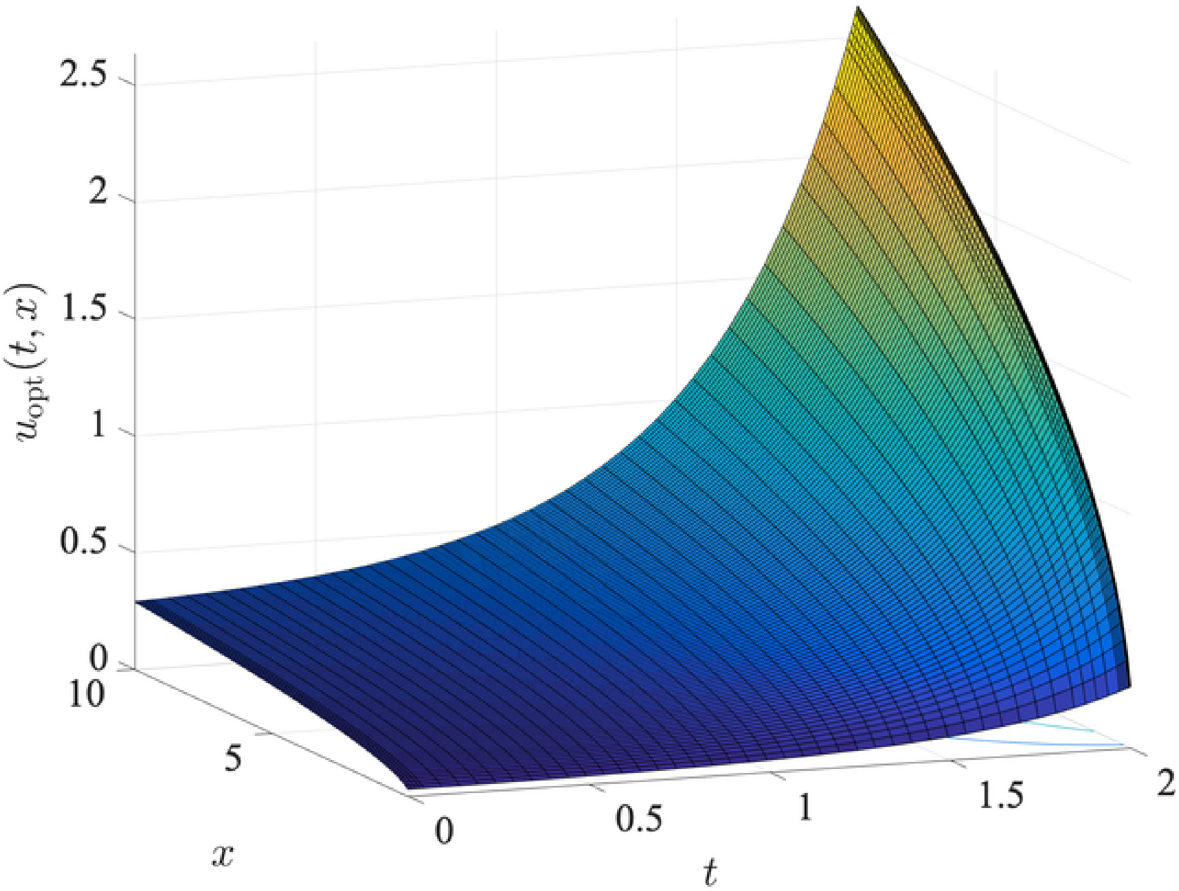
$$\begin{aligned} &\dot{A} - \beta A + 1 - \frac{A^2}{4c} = 0, && A(T) = q,\\ &\dot{B} - \beta B + aA = 0, && B(T) = 0. \end{aligned}$$

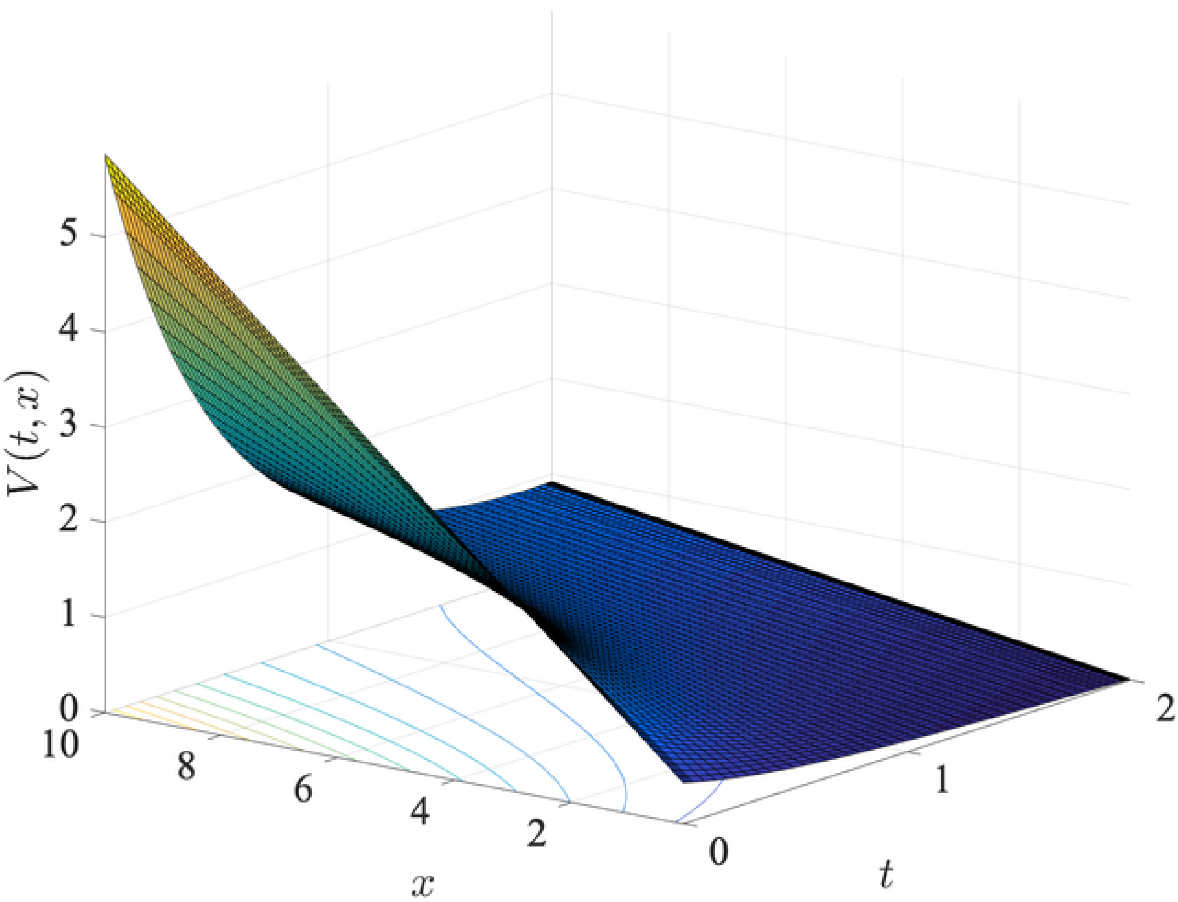
The optimal feedback control and the value function, are obtained in terms of A, B as

$$u_{\text{opt}} = \gamma_{\text{opt}}(t,x) = \frac{e^{\beta t}}{2c}\frac{\partial V}{\partial x}\sqrt{x} = \frac{\sqrt{x}}{2c}A(t), \qquad V(t,x) = e^{-\beta t}\left(A(t)x + B(t)\right).$$

The desired plots obtained by numerically integrating the A, B ODEs are as follows.




2.5
2
1.5
1
0.5
0
$u_{\rm opt}(t,x)$
10
5
x
0
0.5
1
1.5
2
t


5
4
3
2
1
0
$V(t,x)$
10
8
6
4
2
x
0
1
2
t