Problem 1. [50 points] Discrete Time Finite Horizon DP

Before start solving this problem, please review Lec. 16, p. 4-8.

Given a controlled dynamical system

$$x_{k+1} = x_k + u_k, \quad u_k \in \{-1, 0\}, \quad k = 0, 1,$$

with the initial condition $x_0 \equiv 1$. Consider the problem

$$\underset{\gamma \in \Gamma}{\text{minimize}} \quad \alpha \left(x_1^2 + x_2^2\right) + (1 - \alpha) \left(u_0^2 + u_1^2\right)$$

subject to the above dynamics, where Γ is the collection of all history dependent randomized policies, and α is a given constant satisfying $0 \leq \alpha \leq 1$.

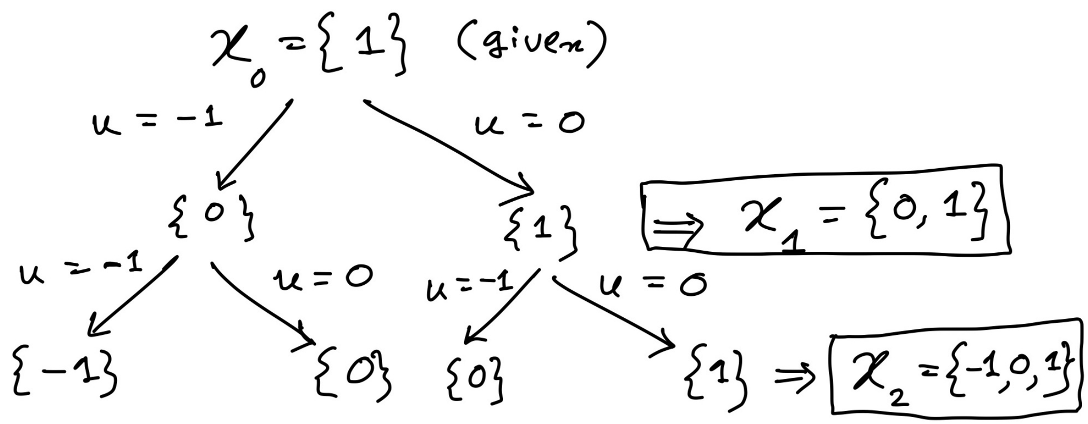
(a) [(2+2+2) + (3+3+3) = 15 points] State space and costs

(i) Clearly write down the state space $\mathcal{X}_k$ for each $k = 0, 1, 2$.

(ii) Looking at the objective, clearly write down the costs c_k for each $k = 0, 1, 2$.

Solution for Problem 1(a):

(i)



To derive the state spaces $\mathcal{X}_k$, we draw a tree rooted at $\mathcal{X}_0 = \{1\}$ (given) as shown in the figure above. This gives $\mathcal{X}_0 = \{1\}$, $\mathcal{X}_1 = \{0, 1\}$, $\mathcal{X}_2 = \{-1, 0, 1\}$.

(ii) We rewrite the given objective in the standard form:

$$\alpha\left(x_1^2 + x_2^2\right) + (1-\alpha)\left(u_0^2 + u_1^2\right) = c_2(x_2) + \sum_{k=0}^{2-1} c_k(x_k, u_k).$$

From the above, we identify $c_0(x_0, u_0) = (1-\alpha)u_0^2$, $c_1(x_1, u_1) = \alpha x_1^2 + (1-\alpha)u_1^2$, $c_2(x_2) = \alpha x_2^2$.

(b) [35 points] Optimal cost-to-go

Apply DP recursions to compute the optimal cost-to-go (a.k.a. value function) V_0. Your final answer should be in the form:

$$V_0 = \begin{cases} f(\alpha) & \text{if} \quad \alpha \leq \alpha_0, \\ g(\alpha) & \text{if} \quad \alpha > \alpha_0, \end{cases}$$

where you need to explicitly compute $f(\cdot), g(\cdot), \alpha_0$.

Solution for Problem 1(b):

From DP recursion, $V_2(x_2) = c_2(x_2) = \alpha x_2^2$, and

$$\begin{aligned} V_1(x_1) &= \min_{u \in \{-1,0\}} \quad c_1(x_1, u) + V_2\left(x_1 + u\right) \\ &= \min_{u \in \{-1,0\}} \quad \alpha x_1^2 + (1-\alpha)u^2 + \alpha(x_1 + u)^2 \\ &= \min\{2\alpha x_1^2, 2\alpha x_1^2 - 2\alpha x_1 + 1\}. \end{aligned}$$

Since $x_1 \in \mathcal{X}_1 = \{0, 1\}$, we get

$$V_1(0) = \min\{0, 1\} = 0, \quad \text{and} \quad V_1(1) = \min\{2\alpha, 1\} = \begin{cases} 2\alpha & \text{if} \quad \alpha \leq 1/2, \\ 1 & \text{if} \quad \alpha > 1/2. \end{cases}$$

Next,

$$\begin{aligned} V_0(x_0) &= \min_{u \in \{-1,0\}} \quad c_0(x_0, u) + V_1\left(x_0 + u\right) \\ &= \min_{u \in \{-1,0\}} \quad (1-\alpha)u^2 + V_1(x_0 + u). \end{aligned}$$

Since $x_0 \in \mathcal{X}_0 = \{1\}$, therefore,

$$
\begin{aligned}
V_0(1) &= \min_{u\in\{-1,0\}} \quad (1-\alpha)u^2 + V_1(1+u)\\
&= \min\{1-\alpha+\underbrace{V_1(0)}_{=0},\ \underbrace{V_1(1)}_{=\min\{2\alpha,1\}}\ \}\\
&= \min\{1-\alpha, \min\{2\alpha,1\}\}\\
&= \begin{cases} \min\{1-\alpha, 2\alpha\} & \text{if} \quad \alpha \leq 1/2,\\ \underbrace{\min\{1-\alpha, 1\}}_{=1-\alpha} & \text{if} \quad \alpha > 1/2,\end{cases}\\
&= \min\{\min\{1-\alpha, 2\alpha\}, 1-\alpha\}\\
&= \min\{1-\alpha, 2\alpha\}\\
&= \begin{cases} 2\alpha & \text{if} \quad \alpha \leq 1/3,\\ 1-\alpha & \text{if} \quad \alpha > 1/3.\end{cases}
\end{aligned}
$$

Hence $f(\alpha) = 2\alpha$, $g(\alpha) = 1-\alpha$, $\alpha_0 = 1/3$.

Problem 2. [50 points] Continuous Time DP

(a) [25 points]

Suppose we have a continuous time DP problem given by the HJB PDE initial value problem

$$\frac{\partial V}{\partial \tau} + H_{\text{opt}}\left(P\frac{\partial V}{\partial x} + q\right) = 0, \qquad V(\tau=0,x) = \phi(x), \quad x \in \mathbb{R}^n,$$

where $H_{\text{opt}}(z)$ is a convex and 1-coercive function of $z \in \mathbb{R}^n$. The matrix $P \succ 0$, and the vector $q \in \mathbb{R}^n$ are given.

Derive a Hopf-Lax formula for the solution $V(\tau, x)$.

Solution for Part 2(a):

Since all conditions stated in Lec. 19, p. 18-19, are satisfied in this case, we can use the Hopf-Lax formula in Lec. 20, p. 1, applied to the affine composition $H_{\text{opt}}(P\xi + q)$.

Suppose that the Legendre-Fenchel conjugate $\left(H_{\text{opt}}(\xi)\right)^* = H^*_{\text{opt}}(\eta)$. Then, from Lec. 20, p. 2, it follows that

$$\left(H_{\text{opt}}(P\xi+q)\right)^* = H^*_{\text{opt}}(P^{-\top}\eta) - \langle q, P^{-\top}\eta\rangle = H^*_{\text{opt}}(P^{-1}\eta) - \langle q, P^{-1}\eta\rangle \quad \text{since } P \succ 0.$$

See also p. 2 of the posted supplementary notes "CalculusOfConvexConjugates.pdf".

Therefore, the Hopf-Lax formula in this case becomes

$$\begin{aligned} V(\tau, x) &= \min_{y \in \mathbb{R}^n} \quad \phi(y) + \tau H^*_{\text{opt}}\left(P^{-1}\frac{x-y}{\tau}\right) - \tau\left\langle q, P^{-1}\frac{x-y}{\tau}\right\rangle \\ &= \min_{y \in \mathbb{R}^n} \quad \phi(y) + \tau H^*_{\text{opt}}\left(P^{-1}\frac{x-y}{\tau}\right) - \left\langle q, P^{-1}(x-y)\right\rangle. \end{aligned}$$

(b) [25 points]

Consider the continuous time OCP

$$\underset{\gamma \in \Gamma}{\text{minimize}} \quad \|x(T)\|_2 + \int_0^T \frac{1}{2}\|u\|_2^2 \, \mathrm{d}t$$

subject to $\dot{x} = u$, where $x \in \mathbb{R}^n$. Prove that the value function is of the form

$$V(t, x) = \begin{cases} a(t, x) & \text{if} \quad \|x\|_2 \leq T - t, \\ b(t, x) & \text{if} \quad \|x\|_2 > T - t. \end{cases}$$

Explicitly determine the functions $a(t, x)$ and $b(t, x)$.

Solution for part 2(b):

Proceeding as in Lec. 20, p. 3-5, we arrive at the HJB PDE IVP:

$$\frac{\partial V}{\partial \tau} + \frac{1}{2}\left\|\frac{\partial V}{\partial x}\right\|_2^2 = 0, \qquad V(\tau = 0, x) = \|x\|_2,$$

where $\tau := T - t > 0$, and the Hopf-Lax formula becomes

$$V(\tau, x) = \min_{y \in \mathbb{R}^n} \quad \|y\|_2 + \frac{1}{2\tau}\|x - y\|_2^2 = M^{\tau}_{\|\cdot\|_2}(x) \quad \text{(the Moreau-Yosida envelope of } \|\cdot\|_2\text{)}.$$

To explcitly compute the argmin $y_{\text{opt}} := \text{prox}_{\tau\|\cdot\|_2}(x)$ of the above unconstrained convex minimization problem, we observe from the definition of prox operator that $y_{\text{opt}} = 0$ if $x = 0$. To determine y_{opt} when $x \neq 0$, we take the derivative of the objective and set it equal to zero to get

$$\left(\frac{\tau}{\|y_{\text{opt}}\|_2} + 1\right) y_{\text{opt}} = x, \quad x \neq 0. \qquad (*)$$

Applying $\|\cdot\|_2$ to both sides of the above, we obtain $\|y_{\text{opt}}\|_2 = \|x\|_2 - \tau$ provided $\|x\|_2 \geq \tau$ since $\|y_{\text{opt}}\|_2$ must be nonnegative. Therefore, $(*)$ gives

$$y_{\text{opt}} = \left(1 - \frac{\tau}{\|x\|_2}\right) x, \quad x \neq 0, \quad \|x\|_2 \geq \tau.$$

Combining the $x = 0$ and $x \neq 0$ cases, we get

$$y_{\text{opt}} = \text{prox}_{\tau\|\cdot\|_2}(x) = \begin{cases} (\|x\|_2 - \tau)_+ \dfrac{x}{\|x\|_2}, & \text{if} \quad x \neq 0, \\ 0, & \text{if} \quad x = 0, \end{cases}$$
$$= \left(1 - \frac{\tau}{\max\{\|x\|_2, \tau\}}\right) x, \qquad (**)$$

wherein $(z)_+ := \max\{z, 0\}$.

Substituting the minimizer $(**)$ back in the objective (as in Lec. 20, p. 7), we obtain the minimum value

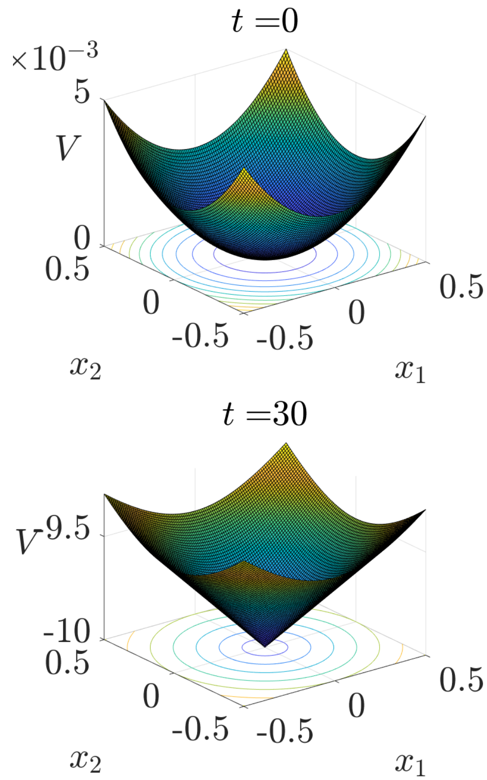
$$M^{\tau}_{\|\cdot\|_2}(x) = \|\text{prox}_{\tau\|\cdot\|_2}(x)\|_2 + \frac{1}{2\tau}\|x - \text{prox}_{\tau\|\cdot\|_2}(x)\|_2^2$$
$$= \begin{cases} \dfrac{1}{2\tau}\|x\|_2^2 & \text{if} \quad \|x\|_2 \leq \tau, \\ \|x\|_2 - \dfrac{\tau}{2} & \text{if} \quad \|x\|_2 > \tau, \end{cases}$$

which is the Huber function well-known in machine learning.

Recalling that $\tau := T - t$, the statement follows with

$$a(t, x) = \frac{1}{2(T - t)}\|x\|_2^2, \qquad b(t, x) = \|x\|_2 - \frac{T - t}{2}.$$

A plot of the corresponding value functions is shown below for $T = 50$ and state space $[-0.5, 0.5]^2$; notice the time-dependent switching of the shape of V as pedicted above.


$t = 0$
$\times 10^{-3}$
5
V
0
0.5
0
-0.5
x_2
-0.5
0
0.5
x_1
$t = 10$
$t = 20$
$t = 30$
-9.5
-10
$t = 40$
-4.5
-5
$t = 50$
0.5