S26 AerE 5740: Optimal Control

Practice Set 10: Continuous Time Stochastic Infinite Horizon DP

Instructor: Abhishek Halder

Problem 1.

Consider the infinite horizon discounted stochastic OCP in continuous time, given by

$$\min_{\boldsymbol{\gamma}\in\Gamma} \quad \mathbb{E}\int_0^\infty e^{-\beta t}\left(q(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{u}^\top \boldsymbol{R}(\boldsymbol{x})\boldsymbol{u}\right)\,\mathrm{d}t$$

$$\mathrm{d}\boldsymbol{x} = (\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u})\,\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{w}, \quad \beta > 0, \quad \boldsymbol{R}(\boldsymbol{x}) \succ \boldsymbol{0}, \quad \boldsymbol{x}\in\mathbb{R}^n, \quad \boldsymbol{u}\in\mathbb{R}^m, \quad \boldsymbol{w}\in\mathbb{R}^p,$$

where Γ denotes the set of all history-dependent randomized policies. Our objective is to compute the optimal state feedback control $\boldsymbol{u}_{\mathrm{opt}} = \boldsymbol{\gamma}_{\mathrm{opt}}(\boldsymbol{x})$, which from DP theory we learnt in class, must be a deterministic Markovian stationary state feedback.

(a) Deriving Stationary HJB PDE.

Derive the DP equation, i.e., the stationary HJB PDE for the value function $V_\infty(\boldsymbol{x})$ associated with the above OCP.

Solution for part (a):

From Lec. 26, the DP equation is

$$\beta V_\infty = \underset{\boldsymbol{u}\in\mathbb{R}^m}{\text{minimum}}\left\{q + \frac{1}{2}\boldsymbol{u}^\top\boldsymbol{R}\boldsymbol{u} + \langle\nabla V_\infty, \boldsymbol{f}+\boldsymbol{g}\boldsymbol{u}\rangle + \frac{1}{2}\text{trace}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^\top\text{Hess}V_\infty\right)\right\},$$

wherein performing the minimization, we find the optimal control

$$\boldsymbol{u}_{\mathrm{opt}} = -\boldsymbol{R}^{-1}\boldsymbol{g}^\top\nabla V_\infty.$$

Substituting this optimal control back in the DP equation gives the stationary HJB PDE:

$$\beta V_\infty = q + \langle\nabla V_\infty, \boldsymbol{f}\rangle - \frac{1}{2}\left(\nabla V_\infty\right)^\top\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^\top\nabla V_\infty + \frac{1}{2}\text{trace}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^\top\text{Hess}V_\infty\right).$$

This is a 2nd order nonlinear stationary PDE.

(b) Solution for the DP equation.

Show that Fleming's logarithmic transformation $V_\infty = -\log\varphi$ from Lec. 25 p. 7, does NOT make the PDE from part (a) linearly solvable even if the channel matching condition $\boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^\top = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ holds.

Solution for part (b):

From Lec. 25 p. 7-8, we have

$$\nabla V_\infty = -\frac{\nabla\varphi}{\varphi}, \quad \text{Hess}\, V_\infty = \frac{(\nabla\varphi)(\nabla\varphi)^\top}{\varphi^2} - \frac{\text{Hess}\,\varphi}{\varphi}.$$

Plugging these into the HJB PDE derived in part (a), we obtain

$$-\beta\log\varphi = q - \langle\nabla\log\varphi, \boldsymbol{f}\rangle + \frac{1}{2}\left(\nabla\log\varphi\right)^\top\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^\top - \boldsymbol{g}\boldsymbol{R}^{-1}\boldsymbol{g}^\top\right)\left(\nabla\log\varphi\right) - \frac{1}{2}\text{trace}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^\top\frac{\text{Hess}\,\varphi}{\varphi}\right),$$

which in the channel matching case simplifies to

$$-\beta\varphi\log\varphi - q\varphi + \langle\nabla\varphi, \boldsymbol{f}\rangle + \frac{1}{2}\text{trace}\left(\boldsymbol{\sigma}\boldsymbol{\sigma}^\top\text{Hess}\,\varphi\right) = 0.$$

This second order stationary PDE has a linear advection and linear diffusion term, but a nonlinear reaction term involving $\varphi\log\varphi$. This is why the PDE is NOT linearly solvable. For example, the special case $\boldsymbol{f} = \boldsymbol{0}, \boldsymbol{\sigma} = \boldsymbol{I}$ gives a nonlinear Poisson equation:

$$\frac{1}{2}\Delta\varphi = \beta\varphi\log\varphi + q\varphi.$$