universität freiburg

Machine Learning for Stochastics **Summer semester 2025**

Lecturer: Prof. Dr. Thorsten Schmidt
Assistance: Simone Pavarana, MSc.

Tutorial 1

Exercise 1 (3 Points)**.**
Suppose a stationary Markov Decision Model is given with action space $A = \{0,1\}$. We denote:

$$r(x,1) = r_1(x), \quad r(x,0) = r_0(x),$$

for all $x \in E$, where r_1 and r_0 are bounded measurable reward functions. For a bounded measurable function $v : E \to \mathbb{R}$ and $a \in A$, define:

$$(Q_a v)(x) := \int v(y) Q(dy \mid x,a).$$

(a) Show that the assumptions (A_N) and (SA_N) are satisfied.

(b) Prove that the value function satisfies the recursion:

$$J_n = \max\{r_0 + \beta Q_0 J_{n-1}, r_1 + \beta Q_1 J_{n-1}\} =: \max\{L_0 J_{n-1}, L_1 J_{n-1}\}.$$

(c) Define $d_n(x) := L_1 J_{n-1}(x) - L_0 J_{n-1}(x)$ and show that:

$$d_{n+1} = L_1 L_0 J_{n-1} - L_0 L_1 J_{n-1} + \beta Q_1 d_n^+ - \beta Q_0 d_n^-.$$

Hint: For part (c), use the identity

$$\max\{x,y\} = x + (y-x)^+ = y + (y-x)^-, \quad \text{for all } x,y \in \mathbb{R}.$$

Exercise 2 (3 Points)**.**
A machine is in use over several periods. The state of the machine is randomly deteriorating and the reward which is obtained depends on the state of the machine. When should the machine be replaced by a new one? The new machine costs a fixed amount $K \geq 0$. We assume that the evolution of the state of the machine is a Markov process with state space $E = \mathbb{R}_+$ and transition kernel Q, where

$$Q([x,\infty) \mid x) = 1,$$

meaning the machine cannot improve spontaneously. A large state x refers to a worse condition/quality of the machine. The reward is $r(x)$ if the state of the machine is x. We assume that the measurable function $r : E \to \mathbb{R}$ is bounded. The terminal reward is $g = r$. We define:

$$(Qv)(x) := \int v(x')Q(dx' \mid x), \quad (Q_0 v)(x) := \int v(x')Q(dx' \mid 0),$$

where note that $(Q_0 v)(x)$ is independent of x.

(a) Show that the assumptions (A_N) and (SA_N) are satisfied.

(b) Show that the Bellman operator is given by:

$$(\mathcal{T}v)(x) = r(x) + \max\left\{\beta(Qv)(x), -K + \beta(Q_0 v)(x)\right\}.$$

(c) Define $d_n(x) := -K + \beta(Q_0 J_{n-1})(x) - \beta(QJ_{n-1})(x)$. Show that:

$$d_{n+1} = -(1-\beta)K - \beta Qr - \beta Q d_n^- + c_n,$$

where $c_n := \beta Q_0 J_n - \beta^2 Q_0 J_{n-1}$ is independent of x.

(d) Assume that r is decreasing and that Q is stochastically monotone. Prove that an optimal decision rule f_n^* is of threshold type, i.e., there exists $x_n^* \in \mathbb{R}_+$ such that:

$$f_n^*(x) = \begin{cases} \text{replace}, & \text{if } x \geq x_n^*, \\ \text{do not replace}, & \text{if } x < x_n^*. \end{cases}$$

Exercise 3 (3+1 Points)**.**
Consider the following special LQ-problem. The transition function is given by

$$T_n(x,a,z) := A_{n+1}x + B_{n+1}a + z,$$

where $x \in \mathbb{R}^m$, $a \in \mathbb{R}^d$, and A_n, B_n are deterministic matrices of appropriate dimension. The disturbances $Z_1, Z_2, \dots$ are independent and identically distributed with expectation and covariance matrix:

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}[ZZ^\top] = \Sigma.$$

The cost to be minimized is

$$\mathbb{E}_x^\pi \left[\sum_{k=0}^N X_k^\top Q_k X_k\right],$$

where Q_k are positive definite matrices.

(a) Show that assumptions (A_N) and (SA_N) are satisfied.

(b) Show that the minimal cost-to-go function is given by

$$V_0(x) = x^\top \widetilde{Q}_0 x + \sum_{k=1}^N \operatorname{tr}(\widetilde{Q}_k \Sigma), \quad x \in \mathbb{R}^m,$$

where the matrices $\widetilde{Q}_n$ are recursively defined by

$$\widetilde{Q}_N := Q_N,$$
$$\widetilde{Q}_n := Q_n + A_{n+1}^\top \widetilde{Q}_{n+1} A_{n+1} - A_{n+1}^\top \widetilde{Q}_{n+1} B_{n+1} \left(B_{n+1}^\top \widetilde{Q}_{n+1} B_{n+1}\right)^{-1} B_{n+1}^\top \widetilde{Q}_{n+1} A_{n+1}.$$

The optimal policy $(f_0^*, \dots, f_{N-1}^*)$ is given by

$$f_n^*(x) = -\left(B_{n+1}^\top \widetilde{Q}_{n+1} B_{n+1}\right)^{-1} B_{n+1}^\top \widetilde{Q}_{n+1} A_{n+1} x.$$

(c) **(Bonus)** Assume now that $A_k = A$, $B_k = B$, and $Q_k = Q$ for all k. Consider the discrete Riccati equation

$$\begin{aligned} \widetilde{Q}_N &:= Q, \\ \widetilde{Q}_n &:= Q + A^\top \widetilde{Q}_{n+1} A - A^\top \widetilde{Q}_{n+1} B \left(B^\top \widetilde{Q}_{n+1} B \right)^{-1} B^\top \widetilde{Q}_{n+1} A. \end{aligned}$$

Moreover, assume that the matrix

$$[B, AB, A^2B, \ldots, A^{N-1}B]$$

has full rank. Show that there exists a positive definite matrix $\widetilde{Q}$ such that

$$\lim_{n\to\infty} \widetilde{Q}_n = \widetilde{Q}.$$

Moreover, $\widetilde{Q}$ is the unique solution of the algebraic Riccati equation

$$\widetilde{Q} = Q + A^\top \widetilde{Q} A - A^\top \widetilde{Q} B \left(B^\top \widetilde{Q} B \right)^{-1} B^\top \widetilde{Q} A,$$

within the class of positive semidefinite matrices.

Hint: For part (c), consider the operator

$$\mathcal{R}(\widetilde{Q}) := Q + A^\top \widetilde{Q} A - A^\top \widetilde{Q} B \left(B^\top \widetilde{Q} B \right)^{-1} B^\top \widetilde{Q} A.$$

First, show that $\mathcal{R}$ is monotone increasing with respect to the Löwner order on symmetric matrices, i.e., if $\widetilde{Q}_1 \leq \widetilde{Q}_2$, then $\mathcal{R}(\widetilde{Q}_1) \leq \mathcal{R}(\widetilde{Q}_2)$.
Next, verify that $Q \leq \mathcal{R}(Q)$ for any Q, and use these facts to conclude that the sequence $(\widetilde{Q}_n)$ is monotone non-decreasing and bounded from above and below. This allows you to deduce convergence.