Markov Decision Processes with Applications
Day 2

Nicole Bäuerle

Accra, February 2020
Overview

▶ Semicontinuous MDP
▶ Stochastic LQ Problems
▶ Terminal Wealth Problems
Semicontinuous MDP
A function $v : M \rightarrow \bar{\mathbb{R}}$ is called **upper semicontinuous** (usc) if for all sequences $(x_n) \subset M$ with $\lim_{n \rightarrow \infty} x_n = x \in M$ it holds

$$\limsup_{n \rightarrow \infty} v(x_n) \leq v(x).$$

A function $v : M \rightarrow \bar{\mathbb{R}}$ is called **lower semicontinuous** if $-v$ is upper semicontinuous.
Example of an Upper Semicontinuous Function

Figure: Graph of an upper semicontinuous function.
Theorem

Let $M$ be compact. If $v : M \to \overline{\mathbb{R}}$ is upper semicontinuous then the function $v$ attains its supremum.
A set-valued mapping (also multifunction or correspondence) $D(\cdot)$ from $E$ to $A$ is a function s.t. $D(x) \subset A$ is non-empty for all $x \in E$.

Here we consider only compact-valued mappings $x \mapsto D(x)$, i.e. $D(x)$ is compact for $x \in E$.

**Definition**

The set-valued mapping $x \mapsto D(x)$ is *upper semicontinuous* (usc) if it has the following property for all $x \in E$:

If $x_n \to x$ and $a_n \in D(x_n)$ for all $n \in \mathbb{N}$, then $(a_n)$ has an accumulation point in $D(x)$. 
Assume the MDP has an upper bounding function $b$ and for all $n = 0, 1, \ldots, N - 1$:

(i) $D_n(x)$ is compact for all $x \in E$ and $x \mapsto D_n(x)$ is usc,

(ii) $(x, a) \mapsto \int v(x')Q_n(dx'|x, a)$ is usc for all usc $v$,

(iii) $(x, a) \mapsto r_n(x, a)$ is usc,

(iv) $x \mapsto g_N(x)$ is usc.

Then the sets $\mathbb{M}_n := \{ v \mid v \text{ is usc} \}$ and $\Delta_n := F_n$ satisfy (SA$_N$). In particular, $V_n \in \mathbb{M}_n$ and there exists a maximizer $f_n^* \in F_n$ of $V_{n+1}$. The policy $(f_0^*, \ldots, f_{N-1}^*)$ is optimal.
Stochastic LQ-Problems
Stochastic Linear-Quadratic Problems

- \( E := \mathbb{R}^m \) where \( x \in E \) denotes the system state,
- \( A := \mathbb{R}^d = D_n(x) \) where \( a \in A \) denotes the action,
- \( Z := \mathbb{R}^{(m,m)} \times \mathbb{R}^{(m,d)} \) where \( Z = (A, B) \) denotes the random transition coefficients of the linear system,
- \( T_n(x, a, A, B) := Ax + Ba, \)
- \( Q^Z(\cdot | x, a) := \text{distribution of } Z_{n+1} := (A_{n+1}, B_{n+1}) \ (= Q^Z(\cdot)) \) independent of \((x, a)\),
- \( r_n(x, a) := -x^\top Q_n x, \ Q_n \) positive definite,
- \( g_N(x, a) := -x^\top Q_N x, \ Q_N \) positive definite,
- \( \beta := 1. \)
The aim is to minimize

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

over all \( N \)-stage policies \( \pi \).
Stochastic Linear-Quadratic Problems

- We have $r \leq 0$. Thus, $(A_N)$ is satisfied.
- We treat this problem as a cost minimization problem, i.e. we suppose that $V_n$ is the minimal cost in the period $[n, N]$.
- For the calculation we assume that all expectations exist.
- The minimal cost operator is given by

$$T_n v(x) = \inf_{a \in \mathbb{R}^d} \left\{ x^\top Q_n x + \mathbb{E} v(A_{n+1} x + B_{n+1} a) \right\}. $$
Assume \( \mathbb{E} \left[ B_{n+1}^\top Q B_{n+1} \right] \) is positive definite for all symmetric positive definite \( Q \).

Consider \( \mathcal{M}_n \) given by

\[
\mathcal{M}_n := \{ \nu : \mathbb{R}^m \to \mathbb{R}_+ \mid \nu(x) = x^\top Q x \text{ with } Q \text{ sym., pos. definite} \}
\]

and \( \Delta_n := \Delta \cap F_n \) with

\[
\Delta := \{ f : E \to A \mid f(x) = C x \text{ for some } C \in \mathbb{R}^{(d,m)} \}.
\]
Stochastic Linear-Quadratic Problems

- Obviously $V_N(x) = x^\top Q_N x \in \mathbb{M}_N$.
- Now let $v(x) = x^\top Qx \in \mathbb{M}_{n+1}$. We try to solve

$$T_n v(x) = \inf_{a \in \mathbb{R}^d} \left\{ x^\top Q_n x + \mathbb{E} v(A_{n+1} x + B_{n+1} a) \right\}$$

$$= \inf_{a \in \mathbb{R}^d} \left\{ x^\top Q_n x + x^\top \mathbb{E} \left[ A_{n+1}^\top Q A_{n+1} \right] x + 2x^\top \mathbb{E} \left[ A_{n+1}^\top Q B_{n+1} \right] a + a^\top \mathbb{E} \left[ B_{n+1}^\top Q B_{n+1} \right] a \right\}.$$
Stochastic Linear-Quadratic Problems

The unique minimum point is given by

\[ f_n^*(x) = -\left( \mathbb{E} \left[ B_{n+1}^\top Q B_{n+1} \right] \right)^{-1} \mathbb{E} \left[ B_{n+1}^\top Q A_{n+1} \right] x. \]

Inserting the minimum point yields

\[ T_n v(x) = x^\top \left( Q_n + \mathbb{E} \left[ A_{n+1}^\top Q A_{n+1} \right] - \mathbb{E} \left[ A_{n+1}^\top Q B_{n+1} \right] \left( \mathbb{E} \left[ B_{n+1}^\top Q B_{n+1} \right] \right)^{-1} \mathbb{E} \left[ B_{n+1}^\top Q A_{n+1} \right] \right) x =: x^\top \tilde{Q} x \]

\[ \tilde{Q} \] is symmetric and since \( x' \tilde{Q} x = T_n v(x) \geq x^\top Q_n x \), it is also positive definite. Thus \( T_n v \in \mathbb{M}_n \)
Stochastic Linear-Quadratic Problems

Theorem

a) Let the matrices $\tilde{Q}_n$ be recursively defined by

\[
\tilde{Q}_N := Q_N \\
\tilde{Q}_n := Q_n + \mathbb{E} \left[ A_{n+1}^\top \tilde{Q}_{n+1} A_{n+1} \right] \\
- \mathbb{E} \left[ A_{n+1}^\top \tilde{Q}_{n+1} B_{n+1} \right] \left( \mathbb{E} \left[ B_{n+1}^\top \tilde{Q}_{n+1} B_{n+1} \right] \right)^{-1} \mathbb{E} \left[ B_{n+1}^\top \tilde{Q}_{n+1} A_{n+1} \right].
\]

Then $\tilde{Q}_n$ are sym., pos. semidefinite and $V_n(x) = x^\top \tilde{Q}_n x$.

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

\[
f_n^*(x) := - \left( \mathbb{E} \left[ B_{n+1}^\top \tilde{Q}_{n+1} B_{n+1} \right] \right)^{-1} \mathbb{E} \left[ B_{n+1}^\top \tilde{Q}_{n+1} A_{n+1} \right] x.
\]
Terminal Wealth Problems
Terminal Wealth Problems

► We have an investor with utility function $U : \text{dom } U \rightarrow \mathbb{R}$ with $\text{dom } U = [0, \infty)$ or $\text{dom } U = (0, \infty)$ and initial wealth $x > 0$.

► A riskless bond with $S_0^0 \equiv 1$ and

$$S_{n+1}^0 := S_n^0 (1 + i_{n+1}), \quad n = 0, 1, \ldots, N-1$$

$i_{n+1}$ denotes the deterministic interest rate in $[n, n+1)$.

► There are $d$ risky assets with $S_0^k = s_0^k, k = 1, \ldots, d$ and

$$S_{n+1}^k = S_n^k \tilde{R}_{n+1}^k, \quad n = 0, 1, \ldots, N-1.$$  

► Assume that $\tilde{R}_1, \ldots, \tilde{R}_N$ are independent,
Terminal Wealth Problems

**Definition**

A *portfolio* or a *trading strategy* is given by \( \pi = (f^n_0, f^n) \) where \( f^n_0 \in \mathbb{R} \) and \( f^n = (f^n_1, \ldots, f^n_d) \in \mathbb{R}^d \). \( f^n_k \) denotes the amount of money invested in asset \( k \) at time \( n \).

**Definition**

A portfolio strategy \( \pi \) is called *self-financing* if

\[
f^n_0 + f^n \cdot e = f^{n-1}_0 (1 + i^n) + f^{n-1} \cdot \tilde{R}_n. \quad \mathbb{P} \text{-a.s.}
\]

for all \( n \), i.e. the current wealth is just reassigned to the assets.
Wealth Equation

\[ X_{n+1} = X_0 + \sum_{t=1}^{n+1} (X_t - X_{t-1}) \]

\[ = X_0 + \sum_{t=1}^{n+1} \left( f_t^0 - f_{t-1}^0 + f_t \cdot e - f_{t-1} \cdot e \right) \]

\[ = X_0 + \sum_{t=1}^{n+1} \left( f_{t-1}^0 i_t + f_{t-1} \cdot (\tilde{R}_t - e) \right) \]

\[ = X_n + (f_n^0 i_{n+1} + f_n \cdot (\tilde{R}_{n+1} - e)) \]

\[ = X_n (1 + i_{n+1}) + \sum_{k=1}^{d} f_n^k \left( \tilde{R}_{n+1}^k - 1 - i_{n+1} \right). \]
Wealth Equation

\[ X_{n+1} = X_n (1 + i_{n+1}) + \sum_{k=1}^{d} f_n^k \left( \tilde{R}_{n+1}^k - 1 - i_{n+1} \right). \]

When we introduce the so-called relative risk process \((R_n)\) defined by \(R_n := (R_n^1, \ldots, R_n^d)\) and

\[ R_n^k := \frac{\tilde{R}_n^k}{1 + i_n} - 1, \quad k = 1, \ldots, d, \]

we obtain

\[ X_{n+1} = (1 + i_{n+1}) \left( X_n + f_n \cdot R_{n+1} \right). \]
Terminal Wealth Problems

- The wealth process \((X_n)\) evolves as follows

\[
X_{n+1} = (1 + i_{n+1}) \left( X_n + f_n \cdot R_{n+1} \right)
\]

where \(\pi = (f_n)\) is a portfolio strategy.

- The optimization problem is then

\[
\begin{aligned}
\max \quad & \mathbb{E}_X^\pi U(X_N) \\
\text{subject to} \quad & \pi = (f_n) \text{ is a portfolio strategy and } X_N \in \text{dom } U \text{ } \mathbb{P}\text{-a.s.}
\end{aligned}
\]
Terminal Wealth Problems

- $E := \text{dom } U$ where $x$ denotes the wealth,
- $A := \mathbb{R}^d$ where $a$ is the amount of money invested in the risky assets,
- $D_n(x) := \left\{ a \in \mathbb{R}^d \mid (1 + i_{n+1})(x + a \cdot R_{n+1}) \in \text{dom } U \; \mathbb{P}\text{-a.s.} \right\}$,
- $Z := [-1, \infty)^d$ where $z$ denotes the relative risk,
- $T_n(x, a, z) = (1 + i_{n+1})(x + a \cdot z)$,
- $Q_n^Z(\cdot | x, a) := \text{distribution of } R_{n+1}$ (independent of $(x, a)$),
- $r_n \equiv 0$, and $g_N(x) := U(x)$. 
Terminal Wealth Problems

Theorem

For the multiperiod terminal wealth problem it holds:

a) The value functions $V_n$ are strictly increasing, strictly concave and continuous.

b) The value functions can be computed recursively by the Bellman equation

$$V_N(x) = U(x),$$
$$V_n(x) = \sup_{a \in D_n(x)} \mathbb{E} V_{n+1} \left( (1 + i_{n+1})(x + a \cdot R_{n+1}) \right), x \in E.$$

(c) There exist maximizers (under no-arbitrage) $f_n^*$ of $V_{n+1}$, and $(f_0^*, f_1^*, \ldots, f_{N-1}^*)$ is optimal for the $N$-stage terminal wealth problem.
Power Utility

Suppose that the utility function is given by

$$U(x) = \frac{1}{\gamma} x^\gamma, \quad x \in [0, \infty)$$

with $0 < \gamma < 1$. Here we have $E = [0, \infty)$. Since it will be more convenient to work with fractions of invested money instead of amounts we define the set of admissible fractions by

$$A_n := \{ \alpha \in \mathbb{R}^d \mid 1 + \alpha \cdot R_{n+1} \geq 0 \ \mathbb{P}\text{-a.s.} \}$$

and the generic one-period optimization problem by

$$v_n := \sup_{\alpha \in A_n} \mathbb{E} \left( 1 + \alpha \cdot R_{n+1} \right)^\gamma.$$
Theorem

Let $U$ be the power utility with $0 < \gamma < 1$. Then:

a) The value functions are given by

$$V_n(x) = d_n x^\gamma, \quad x \geq 0$$

with

$$d_N = \frac{1}{\gamma} \quad \text{and} \quad d_n = \frac{1}{\gamma} \prod_{k=n}^{N-1} (1 + i_{k+1})^\gamma v_k.$$ 

b) The optimal fractions are given by $f_n^*(x) = \alpha_n^*$ where $\alpha_n^*$ is the optimal solution of the generic one-period problem. The optimal portfolio strategy is given by $(f_0^*, f_1^*, \ldots, f_{N-1}^*)$. 

Power Utility
Proof

We will show that

\[ \mathbb{M}_n := \{ v : E \to \mathbb{R}_+ \mid v(x) = bx^\gamma \text{ for } b > 0 \}, \]

\[ \Delta_n := \{ f \in F_n \mid f(x) = \alpha x \text{ for } \alpha \in \mathbb{R}^d \}. \]

Obviously \( g_N = U \in \mathbb{M}_N \).

Let \( v \in \mathbb{M}_{n+1} \). Then we obtain

\[ T_n v(x) = \sup_{a \in D_n(x)} \mathbb{E} v \left( (1 + i_{n+1})(x + a \cdot R_{n+1}) \right) \]

\[ = b(1 + i_{n+1})^\gamma \sup_{a \in D_n(x)} \mathbb{E} (x + a \cdot R_{n+1})^\gamma. \]
Proof

- If $x = 0$ then only $a = 0$ is admissible. Hence suppose $x > 0$.
- We use the transformation $a = \alpha x$ to obtain

$$T_n v(x) = b(1 + i_{n+1})^\gamma x^\gamma \sup_{\alpha \in A_n} E(1 + \alpha \cdot R_{n+1})^\gamma$$

$$= b(1 + i_{n+1})^\gamma x^\gamma v_n.$$

- Thus $T_n v(x) = \tilde{b} x^\gamma \in \mathbb{M}_n$ with $\tilde{b} := b(1 + i_{n+1})^\gamma v_n > 0$.
- For all $v \in \mathbb{M}_{n+1}$ there exists a maximizer $f_n^*$ of $v$ with $f_n^* \in \Delta_n$. 
Proof

The statement now follows by induction. We obtain

\[ V_n(x) = \sup_{\alpha \in A_n} d_{n+1}(1 + i_{n+1})^\gamma x^\gamma \mathbb{E} \left( 1 + \alpha \cdot R_{n+1} \right)^\gamma \]

\[ = d_{n+1}(1 + i_{n+1})^\gamma v_n x^\gamma. \]

If we define \( d_n := d_{n+1}(1 + i_{n+1})^\gamma v_n > 0 \) then the statements in part a) follow.