Reinforcement Learning - Lecture 5

Utilities, Distributional Planning and Von Neumann-Morgenstern Theorem

ENS M2

October 19, 2025

Contents

1 Distributional Statistics Functions

1.1 General Functional

Let $\psi : \mathcal{M}_1(\mathbb{R}) \to \mathbb{R}^K$ be defined such that:

$$\psi(\nu^\pi(s,t)) = \phi(\psi(\nu^\pi(s', t+1)))$$

We saw that this does not work directly with $\psi(\nu) = \mathrm{Var}(\nu)$ but it works with:

$$\psi(\nu) = (\mathbb{E}[\nu], \mathbb{E}[\nu^2])$$

Warning

Exercise
Show that it works with $\psi(\nu) = (\mathbb{E}[\nu], \mathbb{E}[\nu^2], \dots, \mathbb{E}[\nu^K])$ for any $K \geq 1$.
This requires at least:

$$\forall x \in \mathbb{R}, \nu, \nu' \in \mathcal{M}_1(\mathbb{R}), \psi(\nu) = \psi(\nu') \Rightarrow \psi(\tau_x \nu) = \psi(\tau_x \nu')$$

where τ_x is the translation operator.

2 Quantiles and Limitations

2.1 Problem with Quantiles

For quantiles, we also need:

$$\psi(\nu_1) = \psi(\nu_1') \text{ and } \psi(\nu_2) = \psi(\nu_2') \Rightarrow \psi\left(\frac{1}{2}(\nu_1 + \nu_2)\right) = \psi\left(\frac{1}{2}(\nu_1' + \nu_2')\right)$$

However, this is not true in general.

3 Utilities and Utility Functions

3.1 Definition of Utilities

Definition

Utility
We consider a utility: $\psi(\nu) = \mathbb{E}_{X\sim\nu}[f(X)]$. For a utility: $\psi = \psi_f$
We have:

$$\psi\left(\sum_k q_k \nu_k\right) = \mathbb{E}_{X_k \sim \nu_k, I \sim q}[f(X_I)] = \sum_k q_k \mathbb{E}_{X_k\sim\nu_k}[f(X_k)]$$

For utilities, this condition is sufficient.

We need:

$$\forall x \in \mathbb{R}, \nu, \nu' \in \mathcal{M}_1(\mathbb{R}), X \sim \nu, Y \sim \nu', \mathbb{E}[f(X)] = \mathbb{E}[f(Y)] \Rightarrow \mathbb{E}[f(X+x)] = \mathbb{E}[f(Y+x)]$$

3.2 Examples of Utilities

Example

Exponential Utility
$f(x) = e^{\beta x}$:
$$\mathbb{E}[e^{\beta(X+x)}] = e^{\beta x}\mathbb{E}[e^{\beta X}] = e^{\beta x}\mathbb{E}[e^{\beta Y}] = \mathbb{E}[e^{\beta(Y+x)}]$$

This means:
$$\mathbb{E}\left[e^{\beta\left(r(S_t,\pi(S_t),S_{t+1})+\sum_{k=t+1}^{T} R_k\right)}\middle|S_t = s\right]$$
$$= \sum_{s'\in\mathcal{S}} k(s,\pi(s),s')e^{\beta r(s,\pi(s),s')} \cdot \mathbb{E}\left[e^{\beta\sum_{k=t+1}^{T} R_k}\middle|S_{t+1} = s'\right]$$
$$W(s,t) = \sum_{s'\in\mathcal{S}} k(s,\pi(s),s')e^{\beta r(s,\pi(s),s')} \cdot W(s',t+1)$$

4 Risk Measures and β-Means

4.1 Definition of β-Means

Definition

β-Mean
If $\nu \in \mathcal{M}_1(\mathbb{R})$ (bounded) and $\beta \in \mathbb{R}$, we define:
$$U_\beta(\nu) = \frac{1}{\beta}\ln(\mathbb{E}_{X\sim\nu}[e^{\beta X}])$$
for $\beta \neq 0$ and $U_0(\nu) = \mathbb{E}(\nu)$ for $\beta = 0$.

Tip

Properties
The function $\beta \to U_\beta(\nu)$ is increasing on $\mathbb{R}$.
If $\mathbb{P}(a \leq X \leq b) = 1$ then $a \leq U_\beta(\nu) \leq b$ and $\lim_{\beta\to 0} U_\beta(X) = \mathbb{E}(X) = U_0(X)$.

4.2 Special Cases

Example

Normal Distribution
If $X \sim \mathcal{N}(\mu, \sigma^2)$ then:
$$\mathbb{E}[e^{\beta X}] = e^{\beta\mu}\mathbb{E}[e^{\beta\sigma Z}]$$
with $Z \sim \mathcal{N}(0,1)$, which gives:
$$\mathbb{E}[e^{\beta X}] = e^{\beta\mu+\frac{\beta^2\sigma^2}{2}}$$

Asymptotic behavior:

- When $\beta \to +\infty$, $U_\beta(X) \to \sup(X)$
- When $\beta \to -\infty$, $U_\beta(X) \to \inf(X)$

5 Characterization of Translation-Invariant Utilities

5.1 General Question

Are there other utilities f such that:

$$\forall x \in \mathbb{R}, \mathbb{E}[f(X)] = \mathbb{E}[f(Y)] \Rightarrow \mathbb{E}[f(X+x)] = \mathbb{E}[f(Y+x)]?$$

5.2 Special Cases

Case $K = 1$:

- $f =$ constant works
- $f(x) = \alpha + \beta x$ works
- $f(x) = Ce^{\beta x}$ works

5.3 Complete Characterization

Theorem

Characterization Theorem
The only utilities satisfying the translation invariance condition are of the form:

$$f(x) = x^k e^{\beta x}$$

for $0 \leq k \leq K - 1$ and $\beta \in \mathbb{R}$.

This property is called **Bellman closure**.

6 Beyond Utilities: Von Neumann-Morgenstern Theorem

6.1 Stochastic Order

Assume that ν has support in $[\underline{x}, \bar{x}]$ and that if $\nu_1 \leq_{st} \nu_2$, then $\psi(\nu_1) \leq \psi(\nu_2)$ with $\leq_{st}$ a partial order on CDFs defined as:

$$\nu_1 \leq_{st} \nu_2 \Leftrightarrow \forall x \in \mathbb{R}, F_{\nu_1}(x) \geq F_{\nu_2}(x)$$

6.2 Main Theorem

Theorem

Von Neumann-Morgenstern Theorem
Under the above assumptions, if:

$$\psi(\nu_1) = \psi(\nu_2) \Rightarrow \psi(p\nu_1 + (1-p)\nu_2) = \psi(p\nu_1 + (1-p)\nu_2)$$

then there exists a utility ψ_f such that:

$$\forall \nu_1, \nu_2, \psi(\nu_1) \leq \psi(\nu_2) \Leftrightarrow \psi_f(\nu_1) \leq \psi_f(\nu_2)$$

with $\psi_f(\nu_1) = \int f(x)d\nu_1(x)$ and $\psi_f(\nu_2) = \int f(x)d\nu_2(x)$.
This condition is natural for MDPs because of the mixture law.

6.3 Proof Sketch

Proof:

By induction: if $\forall i, \psi(\nu_i) = \psi(\nu_i')$ then $\psi(\sum_{i=1}^k p_i\nu_i) = \psi(\sum_{i=1}^k p_i\nu_i')$

Take $\nu = \sum p_i \delta_{x_i}$. Note that: $\delta_{\underline{x}} \leq_{st} \nu \leq_{st} \delta_{\bar{x}}$

In particular: $\delta_{\underline{x}} \leq_{st} \delta_x \leq_{st} \delta_{\bar{x}} \Rightarrow \psi(\delta_{\underline{x}}) \leq \psi(\delta_x) \leq \psi(\delta_{\bar{x}})$

There exists $f(x) \in [0,1]$ such that:

$$\psi(\delta_x) = f(x)\psi(\delta_{\bar{x}}) + (1-f(x))\psi(\delta_{\underline{x}})$$

Define: $U(\nu) = \int f(x)d\nu(x)$

$$U\left(\sum p_i \delta_{x_i}\right) = \sum f(x_i)p_i$$

7 Distributional Planning

7.1 General Problem

Planning Recap:

$$Q(s,a,t) = \sum_{s' \in \mathcal{S}} k(s,a,s')[r(s,a,s') + V^*_{s',t+1}]$$

$$V^*_{s,t} = \max_{a \in \mathcal{A}} Q(s,a,t)$$

And $\pi_t^*(s) = \arg\max_{a \in \mathcal{A}} Q(s,a,t)$ allows finding an optimal policy for $\mathbb{E}[\sum_{t=1}^T R_t | S_1 = s]$.

Question: What else can we optimize?

$$\max_{\pi} \mathbb{E}[\psi(\sum_{t=1}^T R_t)|S_1 = s]$$

7.2 Naive Distributional Planning Algorithm

$$Q_\nu(s,a,t) = \sum_{s' \in \mathcal{S}} k(s,a,s') \tau_{r(s,a,s')} \nu(s', t+1)$$

$$\psi^*(s,t) = \arg\max_{a \in \mathcal{A}} \psi(Q_\nu(s,a,t))$$

$$\nu^*(s,t) = Q_\nu(s, \psi^*(s,t), t)$$

This is a generalization of the backward induction algorithm (we replaced $\mathbb{E}$ by ψ).