UCSD MAE288A
Optimal Control

CLASS WILL START AT 5PM

Spring 2020
Lecture 9
Section 5:
Discounted-Cost Discrete-Time Stochastic Systems
Recall Discounted-Cost Problem Class

- **Dynamics:**
  \[
  \xi_{t+1} = f(\xi_t, \mu_t(\xi), w_t), \quad (D)
  \]
  \[
  \xi_0 = x \in \mathbb{R}^n. \quad (IC)
  \]

- \(\{w_t\} = \{w_t\}_{t=0}^{T-1}\) IID; \(w_t(\omega) \in \mathbb{R}^k; U \subseteq \mathbb{R}^m\).

- We’ve “seen” that feedbacks are sufficient because of the Markov nature of the noise. Let
  \[
  \mathcal{M}_s^f \doteq \{ \{\mu_t\}_{t=s}^{\infty} \mid \mu_t : \mathbb{R}^n \rightarrow U \ \forall \ t \geq s \}. \]

- Discounted-cost infinite time-horizon payoff (with discount factor \(\alpha \in (0, 1)\)) and value function:
  \[
  J(x, \mu.) \doteq E|_{\xi_0=x} \left\{ \sum_{t=0}^{\infty} \alpha^t L(\xi_t, \mu_t(\xi_t)) \right\} = E\{\{w_t\}_{t=0}^{\infty} \left\{ \sum_{t=0}^{\infty} \alpha^t L(\xi_t, \mu_t(\xi_t)) \right\}, \quad (P)
  \]
  \[
  V(x) = \inf_{\mu. \in \mathcal{M}_0^f} J(x, \mu.). \quad (V)
  \]
Admissible Feedbacks

- A minor concern is that the payoff could be infinite. Example with $\alpha = 0.75$:

$$
\xi_{t+1} = (3 + \mu_t(\xi_t))\xi_t, \quad \xi_0 = 0.4,
$$

$$
L(x, v) \doteq |x|,
$$

$$
\mu_t(x) \doteq -1 \quad \forall \ t, x.
$$

- The set of admissible feedback policies is

$$
\tilde{M}_0 = \{ \{\mu_t\}_{t=0}^{\infty} \mid \mu_t : \mathbb{R}^n \to U \text{ s.t. } J(x, \mu_\cdot) \in \mathbb{R} \ \forall \ t \geq 0 \}.
$$

- Value:

$$
V(x) = \inf_{\mu_\cdot \in \tilde{M}_0} J(x, \mu_\cdot).
$$

- Assume

$$
\exists D < \infty \text{ s.t. } |L(x, v)| \leq D \quad \forall x \in \mathbb{R}^n, \ v \in U.
$$

(A.1)
The Dynamic Programming Equation

Theorem (1)

Under Assumption (A.1), the value function is the unique solution of the DPE given by

$$V(x) = \inf_{v \in U} \{L(x, v) + \alpha E[V(f(x, v, w))]\}.$$  (DPE)

Rough proof sketch:
The right-hand side of \((DPE)\) takes in a function, \(V : R^n \rightarrow R\) and produces a function, say \(\tilde{V} : R^n \rightarrow R\).

Given \(V(\cdot)\), let \(\tilde{V}(x) = G[V](x)\), where
\[
G[V](x) = \inf_{v \in U} \{ L(x, v) + \alpha E[V(f(x, v, w))] \} \quad \forall x \in R^n.
\]

We want to solve \(V = G[V]\).

We want the fixed point of operator \(G\).

For all \(t\), the optimal control will be (if \(\min\) exists)
\[
\mu_t^*(x) = \bar{u}(x) \in \arg\min_{v \in U} \{ L(x, v) + \alpha E[V(f(x, v, w))] \}.
\]

Two common approaches:

1. Value Iteration (a.k.a. BFPT, CMT).
2. Policy Iteration.
Value Iteration

1. Guess some $V_0$.
2. Given $V_n$, $V_{n+1} = \mathcal{G}[V_n]$. Repeat...
3. Stop at some convergence criterion, yielding approximate solution, $V_N \simeq \bar{V}$ (where $\bar{V} = \mathcal{G}[\bar{V}]$) and approximate optimal control $\bar{u}(x) \simeq u_N(x) \in \text{argmin}_{\nu \in U} \{L(x, \nu) + \alpha \mathbb{E}[V_N(f(x, \nu, w))]\}$.

$n$ is NOT time!
Let $X$ be a space. $X$ is a vector space if $ax + by \in X$ for all $x, y \in X$ and all $a, b \in \mathbb{R}$ [plus other conditions. . . ]

Examples:
Let $\mathcal{X}$ be a vector space. $\| \cdot \|$ is a norm on $\mathcal{X}$ and $(\mathcal{X}, \| \cdot \|)$ is a normed vector space if:

1. $\|x\| \geq 0 \quad \forall x \in \mathcal{X}$; $\|x\| = 0$ iff $x = 0$.
2. $\|ax\| = |a|\|x\| \quad \forall a \in \mathbb{R}, \ x \in \mathcal{X}$.
3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathcal{X}$.

Examples:
Let $y_n \in X$ for all $n$. $\{y_n\}_{n=1}^\infty$ is Cauchy if given $\varepsilon > 0$, there exists $N < \infty$ s.t. $\|y_n - y_m\| \leq \varepsilon$ for all $n, m \geq N$.

Examples:
Let \((X, \| \cdot \|)\) be a normed vector space. \(X\) is complete if every Cauchy sequence has a limit in \(X\). In that case, \((X, \| \cdot \|)\) is a Banach space.

Examples:
Let \((\mathcal{Y}, \| \cdot \|)\) be a Banach space (or just an NVSp). \(F : \mathcal{Y} \to \mathcal{Y}\) is a contraction if there exists \(K < 1\) s.t. \(\|F(y) - F(z)\| \leq K\|y - z\|\) for all \(y, z \in \mathcal{Y}\).

Examples:
The Banach Fixed-Point Theorem

- A.k.a. the contraction mapping principle.
- Will yield Value Iteration.

**Theorem (Banach Fixed-Point Theorem)**

Let \((\mathcal{Y}, \| \cdot \|)\) be a Banach Space. Suppose \(F : \mathcal{Y} \rightarrow \mathcal{Y}\) is a contraction. Then, there exists unique \(\bar{y} \in \mathcal{Y}\) such that \(\bar{y} = F(\bar{y})\). Further, if \(y_0 \in \mathcal{Y}\) and \(y_{n+1} = F(y_n)\) for all \(n \geq 0\), then \(y_n \rightarrow \bar{y}\).