6.231 DYNAMIC PROGRAMMING

LECTURE 15

LECTURE OUTLINE

- Review of basic theory of discounted problems
- Monotonicity and contraction properties
- Contraction mappings in DP
- Discounted problems: Countable state space with unbounded costs
- Generalized discounted DP
- An introduction to abstract DP

DISCOUNTED PROBLEMS/BOUNDED COST

- Stationary system with arbitrary state space

$$x_{k+1} = f(x_k, u_k, w_k), \qquad k = 0, 1, \ldots$$

- Cost of a policy $\pi = \{\mu_0, \mu_1, \ldots\}$

$$J_\pi(x_0) = \lim_{N \to \infty} \underset{\substack{w_k \\ k=0,1,\ldots}}{E} \left\{ \sum_{k=0}^{N-1} \alpha^k g\big(x_k, \mu_k(x_k), w_k\big) \right\}$$

with $\alpha < 1$, and for some M, we have $|g(x, u, w)| \leq M$ for all (x, u, w)

- Shorthand notation for DP mappings (operate on functions of state to produce other functions)

$$(TJ)(x) = \min_{u \in U(x)} \underset{w}{E} \left\{ g(x, u, w) + \alpha J\big(f(x, u, w)\big) \right\}, \ \forall\, x$$

TJ is the optimal cost function for the one-stage problem with stage cost g and terminal cost αJ.

- For any stationary policy μ

$$(T_\mu J)(x) = \underset{w}{E} \left\{ g\big(x, \mu(x), w\big) + \alpha J\big(f(x, \mu(x), w)\big) \right\}, \ \forall\, x$$

“SHORTHAND” THEORY – A SUMMARY

- Cost function expressions [with $J_0(x) \equiv 0$]

$$J_\pi(x) = \lim_{k\to\infty}(T_{\mu_0}T_{\mu_1}\cdots T_{\mu_k}J_0)(x), \quad J_\mu(x) = \lim_{k\to\infty}(T_\mu^k J_0)(x)$$

- Bellman's equation: $J^* = TJ^*, \ \ J_\mu = T_\mu J_\mu$
- Optimality condition:

$$\mu\text{: optimal} \quad <==> \quad T_\mu J^* = TJ^*$$

- Value iteration: For any (bounded) J and all x:

$$J^*(x) = \lim_{k\to\infty}(T^k J)(x)$$

- Policy iteration: Given μ^k,
  - Policy evaluation: Find J_{μ^k} by solving

$$J_{\mu^k} = T_{\mu^k}J_{\mu^k}$$

  - Policy improvement: Find μ^{k+1} such that

$$T_{\mu^{k+1}}J_{\mu^k} = TJ_{\mu^k}$$

MAJOR PROPERTIES

- Monotonicity property: For any functions J and J' on the state space X such that $J(x) \leq J'(x)$ for all $x \in X$, and any μ

$$(TJ)(x) \leq (TJ')(x), \qquad \forall\, x \in X,$$

$$(T_\mu J)(x) \leq (T_\mu J')(x), \qquad \forall\, x \in X.$$

- Contraction property: For any bounded functions J and J', and any μ,

$$\max_x \big|(TJ)(x) - (TJ')(x)\big| \leq \alpha \max_x \big|J(x) - J'(x)\big|,$$

$$\max_x \big|(T_\mu J)(x) - (T_\mu J')(x)\big| \leq \alpha \max_x \big|J(x) - J'(x)\big|.$$

- Shorthand writing of the contraction property

$$\|TJ - TJ'\| \leq \alpha \|J - J'\|, \;\; \|T_\mu J - T_\mu J'\| \leq \alpha \|J - J'\|,$$

where for any bounded function J, we denote by $\|J\|$ the sup-norm

$$\|J\| = \max_{x \in X} \big|J(x)\big|.$$

CONTRACTION MAPPINGS

- Given a real vector space Y with a norm $\|\cdot\|$ (see text for definitions).

- A function $F : Y \mapsto Y$ is said to be a *contraction mapping* if for some $\rho \in (0,1)$, we have

$$\|Fy - Fz\| \leq \rho\|y - z\|, \qquad \text{for all } y, z \in Y.$$

ρ is called the *modulus of contraction* of F.

- Linear case, $Y = \Re^n$: $Fy = Ay + b$ is a contraction (for some norm $\|\cdot\|$) if and only if all eigenvalues of A are strictly within the unit circle.

- For $m > 1$, we say that F is an *m-stage contraction* if F^m is a contraction.

- Important example: Let X be a set (e.g., state space in DP), $v : X \mapsto \Re$ be a positive-valued function. Let $B(X)$ be the set of all functions $J : X \mapsto \Re$ such that $J(s)/v(s)$ is bounded over s.

- The *weighted sup-norm* on $B(X)$:

$$\|J\| = \max_{s \in X} \frac{|J(s)|}{v(s)}.$$

- Important special case: The discounted problem mappings T and T_μ [for $v(s) \equiv 1$, $\rho = \alpha$].

A DP-LIKE CONTRACTION MAPPING

- Let $X = \{1, 2, \ldots\}$, and let $F : B(X) \mapsto B(X)$ be a linear mapping of the form

$$(FJ)(i) = b(i) + \sum_{j \in X} a(i,j)\, J(j), \qquad \forall\ i$$

where $b(i)$ and $a(i,j)$ are some scalars. Then F is a contraction with modulus ρ if

$$\frac{\sum_{j \in X} |a(i,j)|\, v(j)}{v(i)} \leq \rho, \qquad \forall\ i$$

[Think of the special case where $a(i,j)$ are the transition probs. of a policy].

- Let $F : B(X) \mapsto B(X)$ be the mapping

$$(FJ)(i) = \min_{\mu \in M} (F_\mu J)(i), \qquad \forall\ i$$

where M is parameter set, and for each $\mu \in M$, F_μ is a contraction from $B(X)$ to $B(X)$ with modulus ρ. Then F is a contraction with modulus ρ.

CONTRACTION MAPPING FIXED-POINT TH.

- Contraction Mapping Fixed-Point Theorem: If $F : B(X) \mapsto B(X)$ is a contraction with modulus $\rho \in (0, 1)$, then there exists a unique $J^* \in B(X)$ such that

$$J^* = FJ^*.$$

Furthermore, if J is any function in $B(X)$, then $\{F^k J\}$ converges to J^* and we have

$$\|F^k J - J^*\| \leq \rho^k \|J - J^*\|, \qquad k = 1, 2, \ldots.$$

- Similar result if F is an m-stage contraction mapping.

- This is a special case of a general result for contraction mappings $F : Y \mapsto Y$ over normed vector spaces Y that are complete: every sequence $\{y_k\}$ that is Cauchy (satisfies $\|y_m - y_n\| \to 0$ as $m, n \to \infty$) converges.

- The space $B(X)$ is complete [see the text (Section 1.5) for a proof].

GENERAL FORMS OF DISCOUNTED DP

- Monotonicity assumption: If $J, J' \in R(X)$ and $J \leq J'$, then

$$H(x, u, J) \leq H(x, u, J'), \qquad \forall\, x \in X,\ u \in U(x)$$

- Contraction assumption:
  - For every $J \in B(X)$, the functions $T_\mu J$ and TJ belong to $B(X)$.
  - For some $\alpha \in (0, 1)$ and all $J, J' \in B(X)$, H satisfies

$$\left| H(x, u, J) - H(x, u, J') \right| \leq \alpha \max_{y \in X} \left| J(y) - J'(y) \right|$$

for all $x \in X$ and $u \in U(x)$.

- We can show all the standard analytical and computational results of discounted DP based on these two assumptions (with identical proofs!)

- With just the monotonicity assumption (as in shortest path problem) we can still show various forms of the basic results under appropriate assumptions (like in the SSP problem)

EXAMPLES

- Discounted problems

$$H(x,u,J) = E\{g(x,u,w) + \alpha J\big(f(x,u,w)\big)\}$$

- Discounted Semi-Markov Problems

$$H(x,u,J) = G(x,u) + \sum_{y=1}^{n} m_{xy}(u)J(y)$$

where m_{xy} are "discounted" transition probabilities, defined by the transition distributions

- Deterministic Shortest Path Problems

$$H(x,u,J) = \begin{cases} a_{xu} + J(u) & \text{if } u \neq t, \\ a_{xt} & \text{if } u = t \end{cases}$$

where t is the destination

- Minimax Problems

$$H(x,u,J) = \max_{w \in W(x,u)} \big[g(x,u,w) + \alpha J\big(f(x,u,w)\big)\big]$$

RESULTS USING CONTRACTION

- The mappings T_μ and T are sup-norm contraction mappings with modulus α over $B(X)$, and have unique fixed points in $B(X)$, denoted J_μ and J^*, respectively (cf. Bellman's equation). *Proof*: From contraction assumption and fixed point Th.

- For any $J \in B(X)$ and $\mu \in \mathcal{M}$,

$$\lim_{k\to\infty} T_\mu^k J = J_\mu, \qquad \lim_{k\to\infty} T^k J = J^*$$

(cf. convergence of value iteration). *Proof*: From contraction property of T_μ and T.

- We have $T_\mu J^* = TJ^*$ if and only if $J_\mu = J^*$ (cf. optimality condition). *Proof*: $T_\mu J^* = TJ^*$, then $T_\mu J^* = J^*$, implying $J^* = J_\mu$. Conversely, if $J_\mu = J^*$, then $T_\mu J^* = T_\mu J_\mu = J_\mu = J^* = TJ^*$.

- Useful bound for J_μ: For all $J \in B(X)$, $\mu \in \mathcal{M}$

$$\|J_\mu - J\| \leq \frac{\|T_\mu J - J\|}{1-\alpha}$$

Proof: Take limit as $k \to \infty$ in the relation

$$\|T_\mu^k J - J\| \leq \sum_{\ell=1}^{k} \|T_\mu^\ell J - T_\mu^{\ell-1} J\| \leq \|T_\mu J - J\| \sum_{\ell=1}^{k} \alpha^{\ell-1}$$

RESULTS USING MON. AND CONTRACTION I

- Existence of a nearly optimal policy: For every $\epsilon > 0$, there exists $\mu_\epsilon \in \mathcal{M}$ such that

$$J^*(x) \le J_{\mu_\epsilon}(x) \le J^*(x) + \epsilon v(x), \qquad \forall\, x \in X$$

Proof: For all $\mu \in \mathcal{M}$, we have $J^* = TJ^* \le T_\mu J^*$. By monotonicity, $J^* \le T_\mu^{k+1} J^* \le T_\mu^k J^*$ for all k. Taking limit as $k \to \infty$, we obtain $J^* \le J_\mu$.

Also, choose $\mu_\epsilon \in \mathcal{M}$ such that for all $x \in X$,

$$\|T_{\mu_\epsilon} J^* - J^*\| = \big\|(T_{\mu_\epsilon} J^*)(x) - (TJ^*)(x)\big\| \le \epsilon(1-\alpha)$$

From the earlier error bound, we have

$$\|J_\mu - J^*\| \le \frac{\|T_\mu J^* - J^*\|}{1-\alpha}, \qquad \forall\, \mu \in \mathcal{M}$$

Combining the preceding two relations,

$$\frac{\big|J_{\mu_\epsilon}(x) - J^*(x)\big|}{v(x)} \le \frac{\epsilon(1-\alpha)}{1-\alpha} = \epsilon, \qquad \forall\, x \in X$$

- Optimality of J^* over stationary policies:

$$J^*(x) = \min_{\mu \in \mathcal{M}} J_\mu(x), \qquad \forall\, x \in X$$

Proof: Take $\epsilon \downarrow 0$ in the preceding result.

RESULTS USING MON. AND CONTRACTION II

- Nonstationary policies: Consider the set Π of all sequences $\pi = \{\mu_0, \mu_1, \ldots\}$ with $\mu_k \in \mathcal{M}$ for all k, and define for any $J \in B(X)$

$$J_\pi(x) = \limsup_{k \to \infty} (T_{\mu_0} T_{\mu_1} \cdots T_{\mu_k} J)(x), \qquad \forall\, x \in X,$$

(the choice of J does not matter because of the contraction property).

- Optimality of J^* over nonstationary policies:

$$J^*(x) = \min_{\pi \in \Pi} J_\pi(x), \qquad \forall\, x \in X$$

Proof: Use our earlier existence result to show that for any $\epsilon > 0$, there is μ_ϵ such that $\|J_{\mu_\epsilon} - J^*\| \leq \epsilon(1 - \alpha)$. We have

$$J^*(x) = \min_{\mu \in \mathcal{M}} J_\mu(x) \geq \min_{\pi \in \Pi} J_\pi(x)$$

Also

$$T^k J \leq T_{\mu_0} \cdots T_{\mu_{k-1}} J$$

Take limit as $k \to \infty$ to obtain $J \leq J_\pi$ for all $\pi \in \Pi$.

6.231 Dynamic Programming and Stochastic Control
Fall 2015