11. Dynamic Programming

Daisuke Oyama

Mathematics II

May 27, 2025

Sequential Problem

Let X be a nonempty set.

We consider the following problem:

$$\max_{(x_t)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \tag{$*$}$$
$$\text{s.t. } x_{t+1} \in \Gamma(x_t), \quad t = 0, 1, \ldots$$
$$x_0 \in X : \text{given},$$

where

- $\Gamma\colon X \to X$ is the nonempty-valued correspondence describing the feasibility constraints,
- $F\colon A \to \mathbb{R}$ is the one-period return function, where $A = \{(x, y) \in X \times X \mid y \in \Gamma(x)\}$, and
- $\beta \in (0, 1)$ is the discount factor.

- For each $x_0 \in X$, we denote

$$\Pi(x_0) = \{(x_t)_{t=0}^{\infty} \mid x_{t+1} \in \Gamma(x_t),\ t = 0, 1, \ldots\}.$$

- We call elements of $\Pi(x_0)$ *feasible paths from* x_0.
- We denote a typical element of $\Pi(x_0)$ by

$$\underline{x} = (x_0, x_1, x_2, \ldots).$$

Example: "Cake Eating"

Fix any $\bar{x} > 0$, and let $X = [0, \bar{x}]$.

Consider

$$\begin{aligned} \max_{(c_t)_{t=0}^{\infty}} \ & \sum_{t=0}^{\infty} \beta^t u(c_t) \\ \text{s.t. } \ & c_t \in [0, x_t] \\ & x_{t+1} = x_t - c_t, \quad t = 0, 1, \ldots \\ & x_0 \in X : \text{given.} \end{aligned}$$

where we assume $u(c) = \frac{1}{\alpha} c^{\alpha}$ ($\alpha < 1, \alpha \neq 0$).

- $\Gamma(x) = \{y \in X \mid y = x - c \text{ for some } c \in [0, x]\} = [0, x]$.
- $F(x, y) = u(x - y)$.

Assumptions

Assumption 1

X is a subset of $\mathbb{R}^N$.
The feasibility correspondence Γ is compact-valued and upper and lower semi-continuous.

Assumption 2

The return function F is continuous.

Assumption 3

The return function F is bounded.

- Many typical examples from economics do not satisfy Assumption 3 *without modification*.
- Sometimes (but not always) one can restrict the state space X to be a compact set.

For $\underline{x} = (x_0, x_1, \ldots) \in \Pi(x_0)$, we write

$$U(\underline{x}) = \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}).$$

Observation 1

For any $x_0 \in X$ and any $\underline{x} = (x_0, x_1, \ldots) \in \Pi(x_0)$,

$$U(\underline{x}) = \sum_{\tau=0}^{t-1} \beta^\tau F(x_\tau, x_{\tau+1}) + \beta^t U(\underline{x}^t),$$

where $\underline{x}^t = (x_t, x_{t+1}, \ldots) \in \Pi(x_t)$.

Optimal Value Function

- $v^*\colon X \to \mathbb{R}$: optimal value function:

$$v^*(x_0) = \sup_{\underline{x} \in \Pi(x_0)} U(\underline{x}) \qquad (x_0 \in X).$$

- v^* is well defined and is a bounded function by Assumption 3 (bounded returns):

  If $|F(x,y)| \leq M$ for all $(x,y) \in A$, then $|v^*(x)| \leq M/(1-\beta)$ for all $x \in X$.

Policy Functions

- A *feasible policy function* (or simply *policy function*, or *policy*) is a function $g\colon X \to X$ such that $g(x) \in \Gamma(x)$ for all $x \in X$.

  Denote by $\mathcal{G}$ the set of all feasible policy functions.

- For each $x_0 \in X$, a policy function g generates a feasible path from x_0,

  $$\underline{x}^g = (x_0, g(x_0), g^2(x_0), \ldots) \in \Pi(x_0),$$

  where $g^{t+1}(x_0) = g(g^t(x_0))$, $t = 1, 2, \ldots$.

Policy Functions

- Define the *policy value function* v_g for g by

$$v_g(x_0) = U(\underline{x}^g) \quad (x_0 \in X).$$

  v_g is a bounded function by Assumption 3 (bounded returns).

- $g \in \mathcal{G}$ is an *optimal policy function* if $v_g(x_0) = v^*(x_0)$ for all $x_0 \in X$.

- We will show that an optimal policy function exists.

Bellman Operator

- Denote by $\mathcal{B}(X)$ the set of bounded functions from $X \to \mathbb{R}$, and by $\mathcal{C}_b(X)$ the set of bounded and continuous functions from $X \to \mathbb{R}$.
- Given a function $v \in \mathcal{B}(X)$, define the function $w \colon X \to \mathbb{R}$ by

$$w(x) = \sup_{y \in \Gamma(x)} F(x, y) + \beta v(y).$$

- w is bounded since F and v are bounded, i.e., $w \in \mathcal{B}(X)$.
- Denote this mapping $v \mapsto w$ by T, i.e.,

$$(Tv)(x) = \sup_{y \in \Gamma(x)} F(x, y) + \beta v(y).$$

  This is a mapping from $\mathcal{B}(X)$ to $\mathcal{B}(X)$.
- This is called the *Bellman Operator*.

Bellman Operator

- If $v \in \mathcal{B}(X)$ is continuous, i.e., $v \in \mathcal{C}_b(X)$, then the "sup" in the definition of T is attained by the Extreme Value Theorem, so that

$$(Tv)(x) = \max_{y \in \Gamma(x)} F(x, y) + \beta v(y).$$

- In this case, Tv is a (bounded and) continuous function, i.e., $Tv \in \mathcal{C}_b(X)$, by the Theorem of Maximum.
- Thus, $T(\mathcal{C}_b(X)) \subset \mathcal{C}_b(X)$.

One-Period Return Operator

- Given a policy function $g \in \mathcal{G}$, define the operator $T_g \colon \mathcal{B}(X) \to \mathcal{B}(X)$ by

$$(T_g v)(x) = F(x, g(x)) + \beta v(g(x)).$$

($T_g v \in \mathcal{B}(X)$ whenever $v \in \mathcal{B}(X)$ by the boundedness of F.)

The Operators

- Bellman operator:

  $T\colon \mathcal{B}(X) \to \mathcal{B}(X)$ defined by:

  $$(Tv)(x) = \sup_{y \in \Gamma(x)} F(x, y) + \beta v(y).$$

- For a policy $g \in \mathcal{G}$,

  $T_g\colon \mathcal{B}(X) \to \mathcal{B}(X)$ defined by

  $$(T_g v)(x) = F(x, g(x)) + \beta v(g(x)).$$

- By definition, $T_g v \leq Tv$.
- g is said to be v-*greedy* if $T_g v = Tv$.

Monotonicity

Observation 2

T and T_g are monotone,
i.e., if $v \leq w$, then

$$
\begin{aligned}
&Tv \leq Tw, \\
&T_g v \leq T_g w.
\end{aligned}
$$

Contraction

Observation 3

T and T_g are contraction mappings with coefficient $\beta \in (0,1)$ for $d(v,w) = \sup_{x \in X}|v(x) - w(x)|$.

Proof

- Fix any $x \in X$.
- Then we have

$$
\begin{aligned}
(Tv)(x) &= \sup_{y \in \Gamma(x)} F(x,y) + \beta[w(y) + (v(y) - w(y))] \\
&\leq \sup_{y \in \Gamma(x)} F(x,y) + \beta w(y) + \beta d(v,w) \\
&= (Tw)(x) + \beta d(v,w),
\end{aligned}
$$

  or, $(Tv)(x) - (Tw)(x) \leq \beta d(v,w)$.

- Similarly, we have $(Tw)(x) - (Tv)(x) \leq \beta d(v,w)$.

Proposition 11.1

For any policy $g \in \mathcal{G}$:

1. v_g *is a fixed point of* T_g*; and*
2. *it is a unique fixed point of* T_g*.*

Proof

1. For each $x_0 \in X$,

$$\begin{aligned} v_g(x_0) &= F(x_0, g(x_0)) \\ &\qquad + \beta F(g(x_0), g^2(x_0)) + \beta^2 F(g^2(x_0), g^3(x_0)) + \cdots \\ &= F(x_0, g(x_0)) + \beta v_g(g(x_0)). \end{aligned}$$

2. $\because T_g$ is a contraction mapping.

Principle of Optimality

Proposition 11.2

1. *T has a unique fixed point in $\mathcal{C}_b(X)$.*
2. *The unique fixed point of T equals v^*.*
3. *g^* is an optimal policy if and only if $T_{g^*}v^* = Tv^*$.*
4. *g^* is an optimal policy if and only if $T_{g^*}v_{g^*} = Tv_{g^*}$.*
5. *An optimal policy exists.*

Principle of Optimality

That is:

- The value function v^* is a unique solution to the *Bellman equation*

$$v(x) = \max_{y \in \Gamma(x)} F(x, y) + \beta v(y)$$

on $\mathcal{C}_b(X)$, where the variable is a (continuous) function v.

- g^* is an optimal policy if and only if g^* is v^*-greedy (also said to be *conserving*), i.e.,

$$g^*(x) \in \arg\max_{y \in \Gamma(x)} F(x, y) + \beta v^*(y)$$

for all $x \in X$.

- In another equivalent expression:

  g^* is an optimal policy if and only if g^* is v_{g^*}-greedy (also said to be *unimprovable*), i.e.,

  $$g^*(x) \in \arg\max_{y \in \Gamma(x)} F(x,y) + \beta v_{g^*}(y)$$

  for all $x \in X$.

  (This is sometimes called the *one-shot deviation principle*.)

- Since v^* is continuous, such a g^* exists, which generates an optimal path.

- Hence, we in fact have $v^*(x_0) = \max_{\underline{x} \in \Pi(x_0)} U(\underline{x})$.

Key Lemma

Lemma 11.3
For $v \in \mathcal{B}(X)$:

1. *If $v \geq Tv$, then $v \geq v^*$.*
2. *If $v \leq Tv$, then $v \leq v^*$.*
3. *If v is a fixed point of T, then it is a unique fixed point of T and equals v^*.*

Proof of Lemma 11.3

- Let $v \in \mathcal{B}(X)$, and suppose that $v \geq Tv$.
- Take any $\underline{x} = (x_0, x_1, \ldots) \in \Pi(x_0)$.
- For all t, we have

$$v(x_t) \geq \sup_{y \in \Gamma(x_t)} F(x_t, y) + \beta v(y) \geq F(x_t, x_{t+1}) + \beta v(x_{t+1}).$$

- Therefore we have

$$\begin{aligned} v(x_0) &\geq F(x_0, x_1) + \beta v(x_1) \\ &\geq F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 v(x_2) \\ &\vdots \\ &\geq \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) + \beta^T v(x_T). \end{aligned}$$

- Let $T \to \infty$.

  Since v is bounded and $\beta \in (0, 1)$, we have $\beta^T v(x_T) \to 0$, so that $v(x_0) \geq U(\underline{x})$.

- Since $\underline{x}$ has been taken arbitrarily, this implies that $v \geq v^*$.

- Let $v \in \mathcal{B}(X)$, and suppose that $v \leq Tv$.
- Fix any $\varepsilon > 0$ and any $x_0 \in X$.
- Define $\underline{x} \in \Pi(x_0)$ as follows:

  For each t, let $x_{t+1} \in \Gamma(x_t)$ be such that

  $$(Tv)(x_t) \leq F(x_t, x_{t+1}) + \beta v(x_{t+1}) + (1-\beta)\varepsilon,$$

  so that

  $$v(x_t) \leq F(x_t, x_{t+1}) + \beta v(x_{t+1}) + (1-\beta)\varepsilon.$$

- Then we have

  $$\begin{aligned} v(x_0) &\leq F(x_0, x_1) + \beta v(x_1) + (1-\beta)\varepsilon \\ &\leq F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 v(x_2) + (1+\beta)(1-\beta)\varepsilon \\ &\vdots \\ &\leq \sum_{t=0}^{T-1} \beta^t F(x_t, x_{t+1}) + \beta^T v(x_T) + \sum_{t=0}^{T-1} \beta^t (1-\beta)\varepsilon. \end{aligned}$$

- Let $T \to \infty$.

  Since v is bounded and $\beta \in (0,1)$, we have $\beta^T v(x_T) \to 0$, so that $v(x_0) \leq U(\underline{x}) + \varepsilon$.

- This implies that $v(x_0) \leq v^*(x_0) + \varepsilon$ for all $x_0 \in X$.

- Since $\varepsilon > 0$ has been taken arbitrarily, this implies that $v \leq v^*$.

Proof of Proposition 11.2

1. [T has a unique fixed point in $\mathcal{C}_b(X)$.]

   By the Contraction Mapping Fixed Point Theorem (Proposition 10.9; see also the Remark there).

2. [The unique fixed point of T equals v^*.]

   By Lemma 11.3.

3. [Principle of Optimality]

$$
\begin{aligned}
& v^* = v_{g^*} \\
& \iff T_{g^*} v^* = v^* \ (\because v_{g^*} \text{ is the unique fixed point of } T_{g^*}) \\
& \iff T_{g^*} v^* = T v^* \ (\because v^* \text{ is a fixed point of } T)
\end{aligned}
$$

4. [One-Shot Deviation Principle]

$$
\begin{aligned}
& v^* = v_{g^*} \\
& \iff T v_{g^*} = v_{g^*} \ (\because v^* \text{ is the unique fixed point of } T) \\
& \iff T v_{g^*} = T_{g^*} v_{g^*} \ (\because v_{g^*} \text{ is a fixed point of } T_{g^*})
\end{aligned}
$$

5. [Existence of optimal policy]

For each $x \in X$, since $F(x, \cdot)$ and v^* are continuous and $\Gamma(x)$ is compact, $g^*(x) \in \Gamma(x)$ such that

$$
F(x, g^*(x)) + \beta v^*(g^*(x)) = \max_{y \in \Gamma(x)} F(x, y) + \beta v^*(y)
$$

exists by the Extreme Value Theorem.

Value Iteration

Proposition 11.4

For any $v_0 \in \mathcal{B}(X)$,

$$d(T^n v_0, v^*) \leq \beta^n d(v_0, v^*) \quad (n = 0, 1, 2, \ldots),$$

so that $T^n v_0$ converges uniformly to v^* as $n \to \infty$.

Proof

- Since v^* is a fixed point of T,

$$\begin{aligned} d(T^n v_0, v^*) &= d(T^n v_0, T^n v^*) \\ &\leq \beta d(T^{n-1} v_0, T^{n-1} v^*) \\ &\vdots \\ &\leq \beta^n d(v_0, v^*). \end{aligned}$$

Solution Algorithms

- Value iteration
- Policy iteration
- Modified policy iteration
- Linear programming

ε-Optimality

Let v^* be the value function.

- v is a δ-*approximation* of v^* if $d(v, v^*) < \delta$.
- g is an ε-*optimal policy* if v_g is an ε-approximation of v^*.

Error Bounds

Lemma 11.5

For any v,

$$d(v^*, Tv) \leq \frac{\beta}{1-\beta} d(Tv, v).$$

Proof

- $d(v^*, Tv) \leq d(v^*, T^m v) + d(T^m v, Tv)$, where

$$\begin{aligned} \text{Second term} &\leq \sum_{k=1}^{m-1} d(T^{k+1} v, T^k v) \\ &\leq \sum_{k=1}^{m-1} \beta^k d(Tv, v) = \frac{\beta - \beta^m}{1-\beta} d(Tv, v). \end{aligned}$$

Let $m \to \infty$.

Lemma 11.6

For any v and any Tv-greedy policy g,

$$d(v_g, Tv) \leq \frac{\beta}{1-\beta} d(Tv, v).$$

Proof

- Denote $u = Tv$.

  Recall that $v_g = T_g v_g$ and $T_g u = Tu$.

- Then,

$$\begin{aligned} d(v_g, u) &= d(T_g v_g, u) \\ &\leq d(T_g v_g, Tu) + d(Tu, u) \\ &= d(T_g v_g, T_g u) + d(Tu, Tv) \\ &\leq \beta d(v_g, u) + \beta d(u, v). \end{aligned}$$

  Rearranging terms yields the desired inequality.

Proposition 11.7

For any v and any Tv-greedy policy g,

$$d(v_g, v^*) \leq \frac{2\beta}{1-\beta} d(Tv, v).$$

Proof

- By the previous two lemmas,

$$\begin{aligned} d(v_g, v^*) &\leq d(v_g, Tv) + d(Tv, v^*) \\ &\leq \frac{\beta}{1-\beta} d(Tv, v) + \frac{\beta}{1-\beta} d(Tv, v). \end{aligned}$$

Value Iteration with a Termination Condition

Specify $\varepsilon > 0$.

1. Set $n = 0$.

   Choose any v^0.

2. Let $v^{n+1} = Tv^n$.

3. If $d(v^{n+1}, v^n) < \frac{1-\beta}{2\beta}\varepsilon$, then return $\hat{v} = v^{n+1}$ and a $\hat{v}$-greedy policy $\hat{g}$.

   Otherwise, let $n = n + 1$ and go to Step 2.

Proposition 11.8

Given an $\varepsilon > 0$, the value iteration algorithm as described terminates in a finite number of iterations, and

- *$\hat{g}$ is an ε-optimal policy and*
- *$\hat{v}$ is an $\frac{\varepsilon}{2}$-approximation of v^*.*

Proof

- By $d(v^{n+1}, v^n) \to 0$, Proposition 11.7, and Lemma 11.5.

Policy Iteration

1. Set $n = 0$.

   Choose any g^0.

2. [Policy evaluation]

   Compute the value v_{g^n}, i.e., the function v_{g^n} such that $v_{g^n} = T_{g^n} v_{g^n}$.

3. [Policy improvement]

   Compute a v_{g^n}-greedy policy g^{n+1}, i.e., a g^{n+1} such that $T_{g^{n+1}} v_{g^n} = T v_{g^n}$.

4. If $g^{n+1} = g^n$, then return $\hat{g} = g^n$ and $\hat{v} = v_{g^n}$.

   Otherwise, let $n = n + 1$ and go to Step 2.

Proposition 11.9

Let $\{g^n\}$ be a sequence obtained by policy iteration.

1. $v_{g^n} \leq T v_{g^n} \leq v_{g^{n+1}}$.
2. $T^n v_{g^0} \leq v_{g^n}$ *($\leq v^*$).*
3. $v_{g^n} \nearrow v^*$ *as* $n \to \infty$.
4. *If* $v_{g^n} = v_{g^{n+1}}$, *then* $v_{g^n} = T v_{g^n}$ *and hence* g^n *is optimal;*
   if $v_{g^n} \neq v_{g^{n+1}}$, *then* $v_{g^n} \neq T v_{g^n}$ *and hence* g^n *is not optimal.*

Proof

1. Show $v_{g^n} \leq T v_{g^n} \leq v_{g^{n+1}}$:
   - We have
   $$\begin{aligned} v_{g^n} &= T_{g^n} v_{g^n} && (v_{g^n} \text{ is a fixed point of } T_{g^n}) \\ &\leq T v_{g^n} && (T_g v \leq T v \text{ for any } v \text{ and } g) \\ &= T_{g^{n+1}} v_{g^n} && (\text{definition of } g^{n+1}). \end{aligned}$$
   - By the monotonicity of $T_{g^{n+1}}$, it follows that
   $$T_{g^{n+1}} v_{g^n} \leq T^2_{g^{n+1}} v_{g^n} \leq T^3_{g^{n+1}} v_{g^n} \leq \cdots \leq T^k_{g^{n+1}} v_{g^n}.$$
   - Let $k \to \infty$.
   Since for any v, $T^k_{g^{n+1}} v \to v_{g^{n+1}}$, we have $T^k_{g^{n+1}} v_{g^n} \to v_{g^{n+1}}$.

2. Show $T^n v_{g^0} \le v_{g^n}$ by induction:
    - Trivial for $n = 0$.
    - Assume that $T^n v_{g^0} \le v_{g^n}$ holds.
    - By the monotonicity of T, we have

    $$T^{n+1} v_{g^0} \le T v_{g^n}.$$

    - By 1, $T v_{g^n} \le v_{g^{n+1}}$.
3. Since $T^n v_{g^0} \to v^*$, we have $v_{g^n} \to v^*$ by 2.
4. By 1.

Proposition 11.10

Suppose that X is a finite set.
The policy iteration algorithm terminates in a finite number of iterations, and $\hat{g}$ is an optimal policy and $\hat{v}$ is the optimal value.

Proof

- Because the values are nondecreasing by Proposition 11.9, and there are finitely many possible policies when there are finitely many states.

Policy Evaluation

- The value v_{g^n} of a policy g^n is a unique solution to

$$v = T_{g^n} v.$$

- If X has N elements, this is a system of N linear equations with N variables.

  $\cdots$ Solvable (by a linear equation solver software).

- If N is huge, solving the equation can take much time.
- Since $T_{g^{n+1}}^k v_{g^n} \to v_{g^{n+1}}$,
  we have $T_{g^{n+1}}^k v_{g^n} \approx v_{g^{n+1}}$ for large k.
- The version of policy iteration with this approximation is called "modified" policy iteration.
- If $k = 1$, then $T_{g^{n+1}}^k v_{g^n} = T v_{g^n}$ $\cdots$ value iteration.

  If $k \to \infty$, then $T_{g^{n+1}}^k v_{g^n} \to v_{g^{n+1}}$ $\cdots$ "exact" policy iteration.

Modified Policy Iteration

Specify $\varepsilon > 0$ and $k \geq 1$.

1. Set $n = 0$.

   Choose any v^0 such that $Tv^0 \geq v^0$.

2. [Policy improvement]

   Compute a v^n-greedy policy g^{n+1}, i.e., a g^{n+1} such that $T_{g^{n+1}} v^n = Tv^n$.

3. Compute $u = Tv^n$ $(= T_{g^{n+1}} v^n)$.

   If $d(u, v^n) < \frac{1-\beta}{2\beta}\varepsilon$, then return u and g^{n+1}.

   Otherwise, go to Step 4.

4. [Partial policy evaluation]

   Compute $v^{n+1} = T^k_{g^{n+1}} u$ $(= T^{k+1}_{g^{n+1}} v^n)$.

   Set $n = n + 1$, and go to Step 2.

Proposition 11.11

Let $\{(g^n, v^n)\}$ be a sequence obtained by modified policy iteration.

1. $v^n \leq Tv^n$ *implies* $Tv^n \leq v^{n+1} \leq Tv^{n+1}$.
2. $v^n \leq v^{n+1}$ *and* $T^n v^0 \leq v^n \leq v^*$.
3. $v^n \nearrow v^*$ *as* $n \to \infty$.

Proof

1. Suppose that $v^n \le Tv^n$. Show that $Tv^n \le v^{n+1} \le Tv^{n+1}$.
   - We have
$$\begin{aligned} v^n &\le Tv^n && \text{(assumption)} \\ &= T_{g^{n+1}} v^n && \text{(definition of } g^{n+1}\text{)}. \end{aligned}$$
   - By the monotonicity of $T_{g^{n+1}}$, it follows that
$$Tv^n = T_{g^{n+1}} v^n \le T^2_{g^{n+1}} v^n \le \cdots \le T^k_{g^{n+1}} v^n \le T^{k+1}_{g^{n+1}} v^n = v^{n+1}.$$
   - By $T^k_{g^{n+1}} v^n \le v^{n+1}$, we also have
$$\begin{aligned} v^{n+1} = T^{k+1}_{g^{n+1}} v^n &\le T_{g^{n+1}} v^{n+1} && \text{(monotonicity of } T_{g^{n+1}}\text{)} \\ &\le Tv^{n+1} && (T_g v \le Tv \text{ for any } v \text{ and } g). \end{aligned}$$

2. Since $v^0 \leq Tv^0$ by assumption, by 1 we have $v^0 \leq Tv^0 \leq v^1 \leq Tv^2 \leq \cdots$.

   In particular, we have $v^n \leq v^{n+1}$ and $Tv^n \leq v^{n+1}$.

   From the latter, we have $T^n v^0 \leq v^n$ by induction.

3. Since $T^n v^0 \to v^*$, we have $v^n \to v^*$ by 2.

Linear Programming

- From Proposition 11.2 and Lemma 11.3,
  - $v^* = Tv^*$, and
  - if $v \geq Tv$, then $v \geq v^*$.
- I.e., v^* is the smallest function that satisfies $v \geq Tv$.

Linear Programming

Suppose that X is a finite set.

- v^* is a (unique) solution to the optimization problem:

$$\begin{aligned} \min_{v \in \mathbb{R}^{|X|}} \quad & \sum_{x \in X} v(x) \\ \text{s.t.} \quad & v(x) \geq (Tv)(x) \quad \text{(for all } x \in X\text{)}, \end{aligned}$$

  where $(Tv)(x) = \max_{y \in \Gamma(x)} F(x, y) + \beta v(y)$;

- or equivalently, to the linear program:

$$\begin{aligned} \min_{v \in \mathbb{R}^{|X|}} \quad & \sum_{x \in X} v(x) \\ \text{s.t.} \quad & v(x) \geq F(x, y) + \beta v(y) \quad \text{(for all } (x, y) \in A\text{)}, \end{aligned}$$

  where $A = \{(x, y) \in X \times X \mid y \in \Gamma(x)\}$.

Envelope Theorem

- We want to consider the derivatives of the value function v^*, $\dfrac{\partial v^*}{\partial x_i}$.
- Recall v^* satisfies

$$v^*(x) = \max_{y \in \Gamma(x)} F(x, y) + \beta v^*(y).$$

- Write

$$G(x) = \arg\max_{y \in \Gamma(x)} F(x, y) + \beta v^*(y).$$

- If the differentiability of v^* is assumed
  and if $F(\cdot, y)$ is differentiable,
  then it is a routine work to derive the envelope formula:

$$\frac{\partial v^*}{\partial x_i}(x) = \frac{\partial F}{\partial x_i}(x, y^*)$$

  for any $y^* \in G(x)$.

Envelope Theorem

- The issue is the differentiability of v^*.
- A sufficient condition is that $F(x, y)$ is concave in (x, y), with some additional conditions.
- We suppose that
  - $X \subset \mathbb{R}^N$ is convex; and
  - $A = \{(x, y) \in X \times X \mid y \in \Gamma(x)\}$ is convex.

Concavity of v^*

Proposition 11.12

If F is concave, then v^* is concave.

Proof

- Suppose that $v^*(x_0) > a$ and $v^*(x_0') > a'$.

  We want to show that $v^*((1-\lambda)x_0 + \lambda x_0') > (1-\lambda)a + \lambda a'$. (Recall Problem **??** in Homework 3.)

- By definition, there exist $\underline{x} \in \Pi(x_0)$ and $\underline{x}' \in \Pi(x_0')$ such that $U(\underline{x}) > a$ and $U(\underline{x}') > a'$.

- Since F is concave, we have

$$\sum_{t=0}^{T} \beta^t F((1-\lambda)x_t + \lambda x_t', (1-\lambda)x_{t+1} + \lambda x_{t+1}')$$

$$\geq (1-\lambda)\sum_{t=0}^{T} \beta^t F(x_t, x_{t+1}) + \lambda \sum_{t=0}^{T} \beta^t F(x_t', x_{t+1}')$$

  for all T.

- It follows that

$$
\begin{aligned}
U((1-\lambda)\underline{x} + \lambda\underline{x}') &\geq (1-\lambda)U(\underline{x}) + \lambda U(\underline{x}') \\
&> (1-\lambda)a + \lambda a'.
\end{aligned}
$$

- Hence,

$$
v^*((1-\lambda)x_0 + \lambda x_0') > (1-\lambda)a + \lambda a'.
$$

Differentiability of v^*

Proposition 11.13

Suppose that

- F *is concave;*
- $x_0 \in \operatorname{Int} X$, $v^*(x_0) < \infty$, $y^* \in G(x_0)$*;*
- *for some neighborhood* $D \subset X$ *of* x_0*,* $y^* \in \Gamma(x)$ *for all* $x \in D$*; and*
- $\dfrac{\partial F}{\partial x_i}(x_0, y^*)$ *exists.*

Then $\dfrac{\partial v^*}{\partial x_i}(x_0)$ *exists, and* $\dfrac{\partial v^*}{\partial x_i}(x_0) = \dfrac{\partial F}{\partial x_i}(x_0, y^*)$*.*

- Sometimes this proposition is called
  the Benveniste-Scheinkman theorem.
- Its main content is the differentiability of v^*,
  and not deriving the envelope formula, which is just a routine once we have the differentiability
  (as you have done in Problem **??** in Homework 7)!
- An elementary proof is given by Milgrom and Segal (2002).

  ($\rightarrow$ Problem **??** in Homework 7.)
- See "On the Differentiability of the Value Function"
  on the course webpage.