Properties of the Recourse Function

Prof. Jeff Linderoth

February 10, 2003
Outline

- Two-stage stochastic LP
  - Convexity
  - Continuity
  - Differentiability
  - Optimality Conditions
- L-Shaped Method!
minimize
\[ c^T x + \mathbb{E}_\omega [q^T y] \]
subject to
\[ Ax = b \]
\[ T(\omega)x + Wy(\omega) = h(\omega) \quad \forall \omega \in \Omega \]
\[ x \in \mathbb{R}_+^n \]
\[ y(\omega) \in \mathbb{R}_+^p \]

- \( Q(x, \omega) = \min_{y \in \mathbb{R}_+^p} \{ q^T y : Wy = h(\omega) - T(\omega)x \} \)
\[
\min_{x \in \mathbb{R}^n_+ : Ax = b} \left\{ c^T x + \mathbb{E}_\omega \left[ \min_{y \in \mathbb{R}^p_+} \{ q^T y : W y = h(\omega) - T(\omega)x \} \right] \right\}
\]

\[
\min_{x \in \mathbb{R}^n_+ : Ax = b} \left\{ c^T x + \mathbb{E}_\omega v(h(\omega) - T(\omega)x) \right\}
\]

\[
\min_{x \in \mathbb{R}^n_+ : Ax = b} \left\{ c^T x + \mathbb{E}_\omega Q(x, \omega) \right\}
\]

\[
\min_{x \in \mathbb{R}^n_+} \left\{ c^T x + Q(x) : Ax = b \right\}
\]
Proofs

• If LP duality holds...

\[ v(z) = \min_{y \in \mathbb{R}^p_+} \left\{ q^T y : Wy = z \right\} = \max_{t \in \mathbb{R}^m} \left\{ z^T t : W^T t \leq q \right\} \]

• Let \( \Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_{|\Lambda|}\} \) be the set of extreme points of \( \{t \in \mathbb{R}^m | W^T t \leq q\} \).

  ◦ Each of those extreme points \( \lambda_k \) is potentially an optimal solution to the LP.

  ◦ In fact, we are sure that there is no optimal solution better than one that occurs at an extreme point, so we can write...

\[ v(z) = \max_{k=1,\ldots,|\Lambda|} \left\{ z^T \lambda_k \right\}, z \in \mathbb{R}^m. \]
\[
\begin{align*}
\alpha v(z_1) + (1 - \alpha)v(z_2) &= \max_{k=1,2,\ldots,|\Lambda|} z_1^T \lambda_k + (1 - \alpha) \max_{k=1,2,\ldots,|\Lambda|} z_2^T \lambda_k \\
&\leq \alpha z_1^T \lambda^*_k + (1 - \alpha) z_2^T \lambda^*_k \\
&= (\alpha z_1 + (1 - \alpha)z_2)^T \lambda^*_k \\
&\leq \max_{k=1,2,\ldots,|\Lambda|} [(\alpha z_1 + (1 - \alpha)z_2)^T \lambda_k] \\
&= v((\alpha z_1 + (1 - \alpha)z_2))
\end{align*}
\]

What’s All This?

? What did I just prove?

Quite Enough Done???
• The above proof is hopelessly wrong

• Take $z_1, z_2 \in \text{dom}(v)$

\[
v((\alpha z_1 + (1 - \alpha)z_2)) = \max_{k=1,\ldots,|\Lambda|} \{(\alpha z_1 + (1 - \alpha)z_2)^T \lambda_k\} \\
= (\alpha z_1 + (1 - \alpha)z_2)^T \lambda_{k^*} \\
= (\alpha z_1^T \lambda_{k^*} + (1 - \alpha)z_2^T \lambda_{k^*} \\
\leq \alpha \max_{k=1,\ldots,|\Lambda|} z_1^T \lambda_k + (1 - \alpha) \max_{k=1,\ldots,|\Lambda|} z_2^T \lambda_k \\
= \alpha v(z_1) + (1 - \alpha)v(z_2)
\]
What if LP duality doesn’t hold. We Make It Hold!

- \( K_1 = \{ x \in \mathbb{R}^n_+ : Ax = b \} \)
- \( K_2 = \{ x | Q(x) < \infty \} \)

So problem is

\[
\min\{ c^T x + Q(x) : x \in K_1 \cap K_2 \}
\]

- A problem is said to have relatively complete recourse if \( K_1 \subseteq K_2 \).

? Why is this good?

⭐ Because we never have to worry about the case \( Q(x, \omega) = \infty \).
• $K_2(\omega) = \{x|Q(x, \omega) < \infty\}$
  ◦ The set of all feasible points for a given realization $\omega$

• $K_2 = \cap_{\omega \in \Omega} K_2(\omega)$

• A problem is said to have **complete recourse** if $\forall z \in \mathbb{R}^m$, $v(z) < \infty$. That is $\forall z \in \mathbb{R}^m, \exists y \in \mathbb{R}_+: Wy = z$.

• This implies that $\forall x, T(\omega), h(\omega), Q(x, \omega) < \infty$, since $z = h - Tx$.
  ◦ Complete recourse is a property of $W$.

  ◦ Namely if the columns of $W$ span $\mathbb{R}^m$, then $\forall z \in \mathbb{R}^m, \exists y \in \mathbb{R}^p : Wy = z$, and we have complete recourse.
Suppose $Q(x, \omega) = \min_{y \in \mathbb{R}^p_+} \{q^T y : Wy = h(\omega) - T(\omega)x\}$ is infeasible for some $x$. (i.e. LP duality doesn’t hold).

In practice, we don’t allow this.

Add additional slack (deviation) variables so that the columns of $W$ span $\mathbb{R}^m$.

- Adding $[I, -I]$ will do the trick.

Simple Example
What About $-\infty$?

- Suppose $Q(x, \omega) = \min_{y \in \mathbb{R}^p_+} \{ q^T y : W y = h(\omega) - T(\omega)x \}$ is unbounded.
- $Q(x, \omega) = -\infty$.
- We just don’t allow this!!
- $q \geq 0$ is sufficient to ensure it.
Other Highlights from Last Time

- **Thm:** If \( f_1(x), f_2(x), \ldots f_q(x) \) is an arbitrary collection of convex functions, then \( M(x) = \max\{f_1(x), f_2(x), \ldots f_q(x)\} \) is also a convex function.

- \( Q(x, \omega) \equiv v(h(\omega) - T(\omega)x) \) is convex.

- \( Q(x) \equiv \mathbb{E}_\omega Q(x, \omega) \) is convex
  - We only showed this for discrete \( \omega \), but the arguments based on sums also carry over to integrals. In fact...

- If \( g(x, y) \) is convex in \( x \), then \( \int g(x, y) dy \) is convex.
  - \( Q(x) = \int_\Omega Q(x, t)dF(t) \)
  - \( \Rightarrow Q(x) \) is convex
Other Properties—Continuity

- $Q(x)$ is Lipschitz-continuous.
  - In fact, all convex functions on the interior of their domain.
  - Some of you proved this on the homework.
- With some care to the technical details, you can also show continuity holds on exterior points as well.
Differentiability

Thm: Suppose LP duality holds, and the dual problem

\[ v(z) = \max_{t \in \mathbb{R}^m} \{ z^T t : W^T t \leq q \} \]

has a unique optimal solution \( \lambda^* \). Then \( \nabla v(z) = \lambda^* \)

Proof:

\[ v(z) = \max_{k=1, \ldots, |\Lambda|} \{ z^T \lambda_k \}, \, z \in \mathbb{R}^m. \]

Suppose that \( \lambda_{k^*} \) is the unique optimal solution to the problem. Then \( \lambda_{k^*} > \lambda_k \, \forall k \in \Lambda \setminus k^* \). Consider

\[ \lim_{h \to 0} \frac{v(z + he_j) - v(z)}{h} \]
Proof, Cont...

- By uniqueness of $\lambda_{k^*}$ and properties of LP,

$$
\lim_{h \to 0} \frac{v(z + he_j) - v(z)}{h} = \lim_{h \to 0} \frac{\lambda^{T}_{k^*} (z + he_j) - \lambda^{T}_{k^*} z}{h}
$$

By L’Hôpital’s rule\textsuperscript{a}, this is $\lambda_{k^*} e_j$. Do this for all directions and you get $\nabla v(z) = \lambda^*$

\textsuperscript{a}(Yikes – what the heck is that?!?!?)

Quite enough done.
• Just a refresher on L’Hôpital’s rule...

• Under some conditions on $f$ and $g$
  ◦ Both differentiable
  ◦ Derivative of $g$ nonzero
  ◦ Both limits go to zero

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

• Anyone remember the Chain rule?
  ◦ Here’s the 1-D version...

$$D(g(f(x))) = g'(f(x))f'(x)$$
What about $Q$?

$$D(g(f(x))) = g'(f(x))f'(x)$$

- Not necessarily so interested in $\nabla v(x)$
- We’re really interested in $\nabla Q(x, \omega)$ and the chain rule gives it to is...
- Apply chain rule with $f(x) = h - Tx$

$$\nabla v(h(\omega) - T(\omega)x) = \nabla Q(x, \omega) = -T\lambda^*$$

More Justification (if time permits)
Let $D(z)$ be the “dual problem”: $\max_{t \in \mathbb{R}^m} \{ z^T t : W^T t \leq q \}$ whose optimal value is $v(z)$.

**Thm:** Suppose $v(z)$ is finite $\forall z \in \mathbb{R}^m$. (LP duality holds). Then

$$\partial v(z) = \Lambda^*(z) \quad \forall z \in \mathbb{R}^m,$$

where $\Lambda^*$ is the set of all optimal solutions to the dual problem $D(z)$.

**Proof:** (You’ll probably need lots of space)
What we really care about are $\nabla Q(x)$ if it exists or $\partial Q(x)$ if it doesn’t.

What is $\partial Q(x) = \partial \mathbb{E}_\omega Q(x, \omega)$?

With much fancy convex analysis, we can show in our case that we can exchange $\mathbb{E}$ and $\partial$.

- Yeah! This means that we can compute $\partial Q(x)$ by decomposing it into subgradients for each $\partial Q(x, \omega)$.

$$\partial Q(x) = \mathbb{E}_\omega \partial Q(x, \omega)$$
• In particular, if $\omega$ comes from a discrete distribution,

$$\partial Q(x) = \sum_{s \in S} p_s Q(x, \omega_s)$$

If $\eta_s = -T(\omega_s)\lambda_s^* \in \partial Q(x, \omega_s)$, then

$$\eta = \sum_{s \in S} p_s \eta_s \in \partial Q(x)$$
Summary

- If $\omega$ comes from a finite distribution
  - $K_2$ is polyhedral. ($K_2 = \cap_{\omega \in \Omega} K_2(\omega)$)
  - $Q(x)$ is piecewise linear and convex on $K_2$
  - (We are going to focus on this case for a while)

- If $\omega$ comes from a continuous distribution with finite second moments.
  - (i.e. it has a bounded variance – Strange things can happen if you don’t – I’ll try to find a little example to give you on the homework).
  - $Q(x)$ is differentiable and convex
Discussion

• Computing $Q(x) = \int_{\Omega} Q(x, t) dF(t)$ in general requires numerical integration for a given value of $x$

• Computing $\nabla Q(x)$ also would require numerical integration.

★ This is only possible when $\omega$ is a vector of very small dimensionality.

• Typically people (and we will too) discretize the continuous distribution.

◊ We’ll talk about this...
Here, again for your convenience are the KKT conditions (in their non-differentiable extension).

- **Thm:** For a convex function \( f : \mathbb{R}^n \mapsto \mathbb{R} \), and convex functions \( g_i : \mathbb{R}^n \mapsto \mathbb{R}, i = 1, 2, \ldots m \), if we have some nice “regularity conditions” (which we have in this case), \( \hat{x} \) is an optimal solution to \( \min_{x \in \mathbb{R}_+^n} \{ f(x) : g_i(x) = 0 \ \forall i = 1, 2, \ldots m \} \) if and only if the following conditions hold:
  
  1. \( g_i(x) = 0 \ \forall i = 1, 2, \ldots m \)
  2. \( \exists \lambda_1, \lambda_2, \ldots \lambda_m \in \mathbb{R}, \mu_1, \mu_2, \ldots \mu_n \in \mathbb{R}_+ \) such that
     - \( 0 \in \partial f(\hat{x}) + \sum_{i=1}^{m} \lambda_i \partial g_i(\hat{x}) - \sum_{j=1}^{n} \mu_j \).
     - \( \mu_j \geq 0 \ \forall j = 1, 2, \ldots n \)
     - \( \mu_j \hat{x}_j = 0 \ \forall j = 1, 2, \ldots n \)
Apply to Our Problem

$$\min_{x \in \mathbb{R}^n_+} \{ c^T x + Q(x) : Ax = b \}$$

**Thm:** $\hat{x} \in K_1$ is optimal if and only if

- $\exists \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n_+$ such that
  - $0 \in c + \partial Q(\hat{x}) + A^T \lambda - \mu$
  - $\mu^T \hat{x} = 0$

Or

$$-c - A^T \lambda + \mu \in \partial Q(\hat{x})$$
Next time

- Algorithms!
  - The lshaped method.
  - Examples and (maybe) some of its variants...

- If I don’t know what you’re doing for a project, please come speak to me.

- Homework #2. :-(

February 5, 2003
Stochastic Programming – Lecture 9