Lecture 10

Duality and Sensitivity

October 6, 2008

Outline

- Perturbed Primal Problem
- Primal Value Function
- Role in Sensitivity
- Role in Zero-Gap
- Duality and Problem Reformulations

Convex Primal and its Dual

$$\begin{array}{llll} \text{minimize} & f(x) & \qquad \text{maximize} & q(\mu,\lambda) \\ \text{subject to} & g(x) \le 0 & \qquad \text{subject to} & \mu \ge 0 \\ & Ax = b & & \\ & x \in X & & \end{array}$$

- x is a primal variable, $x \in \mathbb{R}^n$ [there are n decision variables]
- there are m inequalities, i.e., $g = [g_1, \ldots, g_m]^T$
- there are r linear equalities, i.e., $A \in \Re^{r \times n}$

- We assume that the problem is **convex** and its optimal value f^* **is finite**

Assumption 1.

(a) The set X is convex.

(b) The functions f and g_j are convex ($g = [g_1, \ldots, g_m]$).

(c) The optimal value f^* is finite.

Perturbed Problem and its Dual

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \le u \\ & Ax = b + v, \ x \in X \end{aligned}$$

- u is a parameter, $u \in \mathbb{R}^m$ [there are m inequalities]
- v is a parameter, $v \in \mathbb{R}^r$ [there are r linear equalities]

Derivation of the dual of the perturbed problem

$$\begin{aligned} \tilde{q}(\mu, \lambda) &= \inf_{x \in X} \left\{ f(x) + \mu^T(g(x) - u) + \lambda^T(Ax - b - v) \right\} \\ &= \inf_{x \in X} \left\{ f(x) + \mu^T g(x) + \lambda^T(Ax - b) \right\} - \mu^T u - \lambda^T v \\ &= q(\mu, \lambda) - \mu^T u - \lambda^T v \end{aligned}$$

The dual of the perturbed problem

$$\begin{aligned} \text{maximize} \quad & q(\mu, \lambda) - u^T \mu - v^T \lambda \\ \text{subject to} \quad & \mu \ge 0 \end{aligned}$$

Primal Value Function

Perturbed problem and its dual

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \le u \\ & Ax = b + v \\ & x \in X \end{array} \qquad \begin{array}{ll} \text{maximize} & q(\mu,\lambda) - u^T\mu - v^T\lambda \\ \text{subject to} & \mu \ge 0 \end{array}$$

- Let $p(u,v) : \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}$ be the function that, to each (u,v), assigns the optimal value of the perturbed problem:

$$p(u,v) = \inf_{\substack{g(x)\le u \\ Ax=b+v \\ x\in X}} f(x) \qquad \text{for all } (u,v) \in \mathbb{R}^m \times \mathbb{R}^r$$

- It is referred to as **primal value function** or **primal function**
- Note that $p(0,0) = f^*$
- Primal function $p(u,v)$ **is convex** when the primal problem is convex
- In some cases, the primal optimal value $p(0,0)$ provides some information about $p(u,v)$

Examples of Primal Value Functions

Example 1

$$\begin{array}{ll} \text{minimize} & -\sqrt{x} \\ \text{subject to} & x \le 0 \end{array} \qquad \mathrm{dom} f = \{x \in \mathbb{R} \mid x \ge 0\}$$

$$p(u) = \inf_{x \le u} \left(-\sqrt{x}\right) = \begin{cases} -\sqrt{u} & \text{for } u \ge 0 \\ +\infty & \text{for } u < 0 \end{cases}$$

Example 2

$$\begin{array}{ll} \text{minimize} & e^{-\sqrt{x_1 x_2}} \\ \text{subject to} & x_1 \le 0 \end{array} \qquad \mathrm{dom} f = \{(x_1, x_2) \mid x_1 \ge 0,\ x_2 \ge 0\}$$

$$p(u) = \inf_{x_1 \le u} e^{-\sqrt{x_1 x_2}} = \begin{cases} 1 & \text{for } u = 0 \\ 0 & \text{for } u > 0 \\ +\infty & \text{for } u < 0 \end{cases}$$

Convexity of the Primal Function

Theorem Let Assumption 1 hold. Then, the primal function $p(u, v)$ is convex.

Proof Suppose that $p(u, v)$ is finite everywhere. Let $(u_1, v_1), (u_2, v_2) \in \Re^{m\times r}$ and $\alpha \in [0, 1]$ be arbitrary. Let $\epsilon > 0$ be also arbitrary, and $x_i \in X, i = 1, 2$ be such that

$$f(x_i) \le p(u_i, v_i) + \epsilon \qquad \text{for } i = 1, 2. \tag{1}$$

Then, by convexity of g_j's and X, we have

$$g(\alpha x_1 + (1-\alpha)x_2) \le \alpha u_1 + (1-\alpha)u_2$$
$$A(\alpha x_1 + (1-\alpha)x_2) = b + \alpha v_1 + (1-\alpha)v_2$$
$$\alpha x_1 + (1-\alpha)x_2 \in X$$

Therefore $\alpha x_1 + (1-\alpha)x_2$ is feasible for the perturbed problem, and hence, by convexity of f,

$$\begin{aligned} p(\alpha(u_1, v_1) + (1-\alpha)(u_2, v_2)) &\leq f(\alpha x_1 + (1-\alpha)x_2) \\ &\leq \alpha f(x_1) + (1-\alpha)f(x_2) \qquad (2) \end{aligned}$$

By combining Eqs. (1) and (2), we obtain

$$p(\alpha(u_1, v_1) + (1-\alpha)(u_2, v_2)) \leq \alpha p(u_1, v_1) + (1-\alpha)p(u_2, v_2) + \epsilon$$

By letting $\epsilon \to 0$, we see that

$$p(\alpha(u_1, v_1) + (1-\alpha)(u_2, v_2)) \leq \alpha p(u_1, v_1) + (1-\alpha)p(u_2, v_2)$$

implying that p is convex.

When p is not finite everywhere, the proof is based on showing that the epigraph of p is a convex set. $\square$.

Global Sensitivity Result

- Consider the perturbed primal problem and its dual

- Lower-bound property of the dual to the perturbed problem yields:

$$p(u,v) \ \geq \ \sup_{\mu \geq 0}\{q(\mu,\lambda) - u^T\mu - v^T\lambda\} \qquad \text{for all } u, v$$

Theorem Assume that the primal problem has finite optimal value f^*, there is no duality gap $[q^* = f^*]$, and its dual has an optimal solution (μ^*, λ^*). Then

$$\begin{aligned} p(u,v) \ &\geq \ q(\mu^*,\lambda^*) - u^T\mu^* - v^T\lambda^* \\ &= p(0,0) - u^T\mu^* - v^T\lambda^* \qquad \text{for all } u, v \end{aligned}$$

Sensitivity Interpretation

- μ_j^* large: p increases greatly if we tighten constraint j $(u_j < 0)$
- μ_j^* small: p does not decrease much if we loosen constraint j $(u_j > 0)$
- λ_i^* large and positive: p increases greatly if we take $v_i < 0$
- λ_i^* large and negative: p increases greatly if we take $v_i > 0$
- λ_i^* small and positive: p does not decrease much if we take $v_i > 0$
- λ_i^* small and negative: p does not decrease much if we take $v_i < 0$

Local Sensitivity

Theorem Assume that the primal problem has finite optimal value f^*, there is no duality gap $[q^* = f^*]$, its dual has an optimal solution (μ^*, λ^*), and the primal value function p is differentiable at $(0, 0)$. Then:

$$\mu_j^* = -\frac{\partial p(0,0)}{\partial u_j}, \qquad \lambda_i^* = -\frac{\partial p(0,0)}{\partial v_i} \qquad \text{for all } i, j$$

<u>Proof</u>: We use global sensitivity relation

$$p(u, v) \geq p(0, 0) - u^T \mu^* - v^T \lambda^* \qquad \text{for all } u, v$$

For μ_j^*, this relation yields

$$\frac{\partial p(0,0)}{\partial u_j} = \lim_{t \searrow 0} \frac{p(te_j, 0) - p(0, 0)}{t} \geq -\mu_j^*$$

where e_j is the vector in $\mathbb{R}^m$ with i-th entry 1 and other entries 0. Similarly,

$$\frac{\partial p(0,0)}{\partial u_j} = \lim_{t \nearrow 0} \frac{p(te_j, 0) - p(0, 0)}{t} \leq -\mu_j^*$$

Hence, the equality for μ_j^* follows. The proof for λ_i is identical. $\square$

Interpretation of Local Sensitivity

$$\mu_j^* = -\frac{\partial p(0,0)}{\partial u_j}, \qquad \lambda_i^* = -\frac{\partial p(0,0)}{\partial v_i} \qquad \text{for all } i, j$$

The local sensitivity result provides information locally at a primal optimal solution x^* (when such exists). Specifically, it gives us a measure of sensitivity of a constraint at the optimal x^* in the following sense:

- If $g_j(x^*) < 0$ (**inactive constraint**), then $\mu_j^* = 0$ (why?). Hence, the constraint g_j can be tightened or loosened by a small amount without affecting the optimal value f^* at all
- If $g_j(x^*) = 0$ (**active constraint**) and μ_j^* is small, then tightening or loosening the constraint by a small amount has a small effect on the optimal value f^*
- If $g_j(x^*) = 0$ (**active constraint**) and μ_j^* is large, then tightening or loosening the constraint by a small amount has a large effect on the optimal value f^*

Example: Maximizing Profit

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & a^T x = b \end{array}$$

- The problem may represent maximizing the profit $-f(x)$ of a firm subject to the given level b at which the firm operates [the firm resources are utilized at level b]

- The firm is interested in determining how will the optimal profit change when the operative level changes from b to $b + u$ [corresponding to utilizing more resources when $u > 0$ and utilizing less when $u < 0$]

Formal assumptions: there is no gap, and an optimal primal solution x^* and a dual solution $\mu^* \in \mathbb{R}$ exist, and f is continuously differentiable [globally over entire $\mathbb{R}^n$, or locally within some neighborhood of x^*]

We are interested in **the change in the optimal value $f(x^*)$ resulting from a small change in the level b**

Example: Profit Sensitivity Analysis

$$\begin{array}{ll}\text{minimize} & f(x)\\ \text{subject to} & a^T x = b + u\end{array}$$

- Let $x^*(u)$ be an optimal solution to the perturbed problem
- Let Δx^* be the change in the optimal solution: $\Delta x^* = x^*(u) - x^*$
- Let Δf^* be the change in the optimal value: $\Delta f^* = f(x^*(u)) - f(x^*)$
- By KKT conditions (Lagrangian optimality in x), we have

$$\nabla f(x^*) = -\lambda^* a$$

- By using Taylor's first order expansion at x^*, we obtain

$$\Delta f^* = \nabla f(x^*)^T \Delta x^* + o(\|\Delta x^*\|) \approx -\lambda^* a^T \Delta x^*$$

- Using $a^T x^* = b$ and $a^T x^*(u) = b + u$, we have $a^T \Delta x^* = u$. Hence: $\Delta f^* \approx -\lambda^* u$. Insignificant approximation error implies that

$$\frac{\text{change in opt. profit}}{\text{change in oper. level}} = \frac{\Delta f^*}{u} = -\lambda^*$$

- The multiplier λ^* provides information on the rate of change in optimal profit per unit of change in operation level

Comments

- The dual problem may have multiple solutions
- Which of the multipliers to use in the sensitivity analysis?
  - Ideally: the multiplier with the smallest norm
  - The one that is available

Primal Value Function and Zero Duality Gap

Primal Problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \preceq 0 \\ & Ax = b, \quad x \in X \end{array}$$

Primal Value Function

$$p(u,v) = \inf_{\substack{g(x) \preceq u \\ Ax = b + v \\ x \in X}} f(x) \qquad \text{for all } (u,v) \in \mathbb{R}^m \times \mathbb{R}^r$$

The primal value function is closely related to the duality gap:

there is no gap in the primal problem if and only if p is closed at $(0,0)$

- Consider the set V given by

$$V = \{(u,v,w) \mid g(x) \preceq u, \quad Ax - b = v, \quad f(x) \le w, \quad x \in X\}$$

- There holds: $V \subset \mathrm{epi}\, p \subset \mathrm{cl}(V)$
- Duality gap can be investigated by exploring:

the closedness of V at $(0,0,f^*)$

Examples Illustrating Epigraph of p

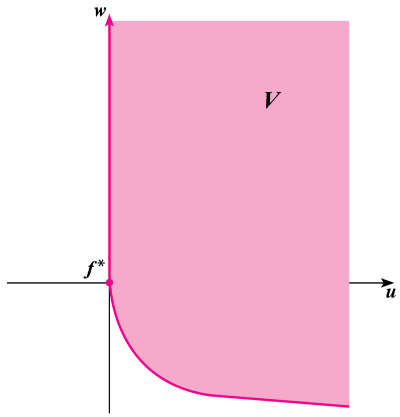
Example 1

minimize $-\sqrt{x}$
subject to $x \leq 0$

Optimal value $f^* = 0$
$\mathrm{dom} f = \{x \in \mathbb{R} \mid x \geq 0\}$

$p(u) = -\sqrt{u}$ for $u \geq 0$

$p(u) = +\infty$ otherwise



- Here: $\mathrm{epi}\, p = \{(u, w) \in \mathbb{R}^2 \mid -\sqrt{u} \leq w,\ u \geq 0\}$, hence $\mathrm{epi}\, p$ is closed
- Thus, p is closed at 0 [lower semi-continuous at 0], so zero duality holds

An Example with a Duality Gap

Example 2

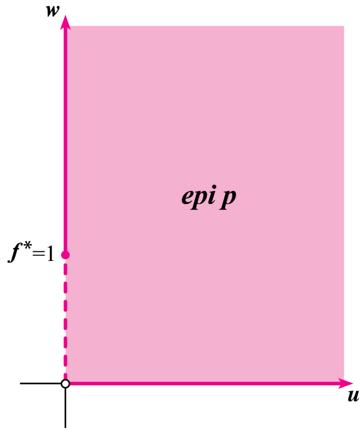
minimize $e^{-\sqrt{x_1 x_2}}$
subject to $x_1 \leq 0$

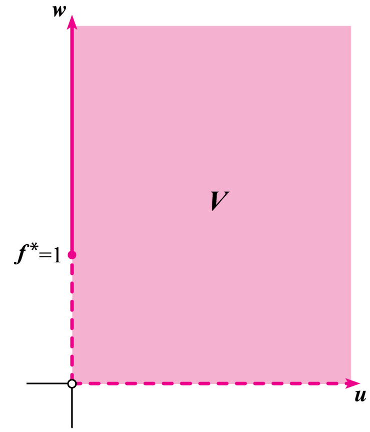
Optimal value $f^* = 1$
$\text{dom} f = \{(x_1, x_2) \mid x_1 \geq 0,\ x_2 \geq 0\}$

$p(u) = 1$ for $u = 0$

$p(u) = 0$ for $u > 0$

$p(u) = +\infty$ otherwise





- Here: $\text{epi}\, p = \{(0, w) \in \mathbb{R}^2 \mid 1 \leq w\} \cup \{(u, w) \in \mathbb{R}^2 \mid 0 \leq w,\ u > 0\}$
- The epigraph $\text{epi}\, p$ is not closed at 0, hence a duality gap exists
- Also: $V = \{(0, w) \in \mathbb{R}^2 \mid 1 \leq w\} \cup \{(u, w) \in \mathbb{R}^2 \mid 0 < w,\ u > 0\}$, showing that here, the set V is a strict subset of $\text{epi}\, p$

Duality and Problem Reformulations

Recall the Example:

$$\begin{array}{ll} \text{minimize} & -x_2 \\ \text{subject to} & \|x\| \leq x_1 \\ & x_2 \geq 0 \end{array} \qquad \begin{array}{ll} \text{minimize} & -x_2 \\ \text{subject to} & x_1 \geq 0 \\ & x_2 = 0 \end{array}$$

Formulation to the left: an **infinite duality gap**

Formulation to the right: a **zero duality gap**

Even when equivalent formulations of a problem lead to zero duality gap:

- Equivalent formulations can lead to very different dual problems
- Reformulating the primal problem can be useful when, for example:
  - The dual is difficult to derive
  - The dual is uninteresting (not very informative)

Some common reformulations include

- Introducing new variables and equality constraints
- Transforming the objective or constraint functions eg., replacing the objective $f(x)$ by $\phi(f(x))$ with ϕ convex and increasing

Introducing New Variables and Equality Constraints

Primal problem

$$\text{minimize} \;\; f(Ax+b)$$

Reformulated problem and its dual

$$\begin{array}{ll} \text{minimize} & f(y) \\ \text{subject to} & Ax+b-y=0 \end{array} \qquad \begin{array}{ll} \text{maximize} & \inf_y \{f(y)-\lambda^T y\} + b^T\lambda \\ \text{subject to} & A^T\lambda = 0 \end{array}$$

- The dual function follows from

$$\begin{aligned} q(\lambda) &= \inf_{x,y} \{f(y) - \lambda^T y + \lambda^T Ax + \lambda^T b\} \\ &= \inf_y \{f(y) - \lambda^T y\} + \inf_x (Ax)^T\lambda + b^T\lambda \\ &= \begin{cases} \inf_y \{f(y) - \lambda^T y\} + b^T\lambda & \text{for } A^T\lambda = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Example

Norm Approximation Problem: minimize $\|Ax - b\|$

Reformulated Problem:

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & y = Ax - b \end{array}$$

The dual function is:

$$q(\lambda) = \begin{cases} b^T\lambda + \inf_y \{\|y\| + \lambda^T y\} & \text{for } A^T\lambda = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Note that: $\inf_y \{\|y\| + \lambda^T y\} = \inf_y \{\|y\| - \lambda^T y\}$, hence

$$q(\lambda) = \begin{cases} b^T\lambda & \text{for } A^T\lambda = 0, \quad \|\lambda\| \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

Dual of Norm Approximation Problem

$$\begin{array}{ll} \text{maximize} & b^T\lambda \\ \text{subject to} & A^T\lambda = 0, \quad \|\lambda\| \le 1 \end{array}$$

Implicit Constraints

LP with box constraints: primal and dual problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{array} \qquad \begin{array}{ll} \text{maximize} & -b^T\lambda - \mathbf{1}^T\mu - \mathbf{1}^T\nu \\ \text{subject to} & c + A^T\lambda + \mu - \nu = 0 \\ & \mu \succeq 0, \ \nu \succeq 0 \end{array}$$

The dual is complicated so by making the box **constraints implicit** [equivalent to keeping the box constraints in the set X]

$$\begin{array}{ll} \text{minimize} & f(x) = \begin{cases} c^T x & \text{for } -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} & Ax = b \end{array}$$

The dual function now is:

$$q(\lambda) = \inf_{-\mathbf{1} \preceq x \preceq \mathbf{1}} \left\{ c^T x + \lambda^T (Ax - b) \right\} = -b^T\lambda - \|A^T\lambda + c\|_1$$

The dual problem: maximize $-b^T\lambda - \|A^T\lambda + c\|_1$

Feasibility Problems

Feasibility problem A in variables $x \in \mathbb{R}^n$:

$$g_j(x) < 0, \quad j = 1, \ldots, m, \qquad h_i(x) = 0, \quad i = 1, \ldots, r$$

Feasibility problem B in variables $\mu \in \mathbb{R}^m$, $\lambda \in \mathbb{R}^r$:

$$\mu \succeq 0, \qquad \mu \neq 0, \qquad q(\mu, \lambda) \geq 0$$

where $q(\mu, \lambda) = \inf_{x \in \mathbb{R}^n} \left\{ \sum_{j=1}^m \mu_j g_j(x) + \sum_{i=1}^r \lambda_i h_i(x) \right\}$

- Problem B is convex (q is concave), even when problem A is not
- Problems A and B are always **weak alternatives**: at most one is feasible

Proof: Assume $\tilde{x}$ is feasible in problem A and (μ, λ) is feasible in problem B

$$0 \leq q(\mu, \lambda) \leq \sum_{j=1}^m \mu_j g_j(\tilde{x}) + \sum_{i=1}^r \lambda_i h_i(\tilde{x}) < 0$$