MATH: Operations Research **2014-15 First Term**

Handout 5: Linear programming duality theory – Strong duality

Instructor: Junfeng Yang *September 28, 2014*

Contents

5.1 Preliminaries

Definition 5.1 (线性类) *A set L in R^n is said to be a* linear variety, *if given any $x, y \in L$, we have $\lambda x + (1-\lambda)y \in L$ for all $\lambda \in R$.*

Let L be a linear variety in R^n. Clearly, for any $x_0 \in L$, the set $L - x_0 := \{x - x_0 : x \in L\}$ is a subspace of R^n. The dimension of L is then defined as that of $L - x_0$.

Definition 5.2 (超平面) *A hyperplane in R^n is an $(n-1)$-dimensional linear variety.*

A hyperplane is a largest linear variety in a space, other than the entire space itself.

Theorem 5.3 *Let $0 \neq a \in R^n$ and $c \in R$. The set $H := \{x \in R^n : a^T x = c\}$ is a hyperplane in R^n. On the other hand, any hyperplane in R^n can be represented in this form.*

Proof: Let H be a hyperplane. Take $x_1 \in H$ and define $M := H - x_1$. Then, M is a subspace. Since H is a hyperplane, the dimension of M is $n-1$. Thus, $M^\perp$ is one-dimensional subspace. Take $0 \neq a \in M^\perp$ and let $c = a^T x_1$. It can be verified that $H = \{x \in R^n : a^T x = c\}$. ∎

Definition 5.4 *Let $0 \neq a \in R^n$ and $c \in R$. Corresponding to the hyperplane $H := \{x \in R^n : a^T x = c\}$ are the positive and negative closed half spaces*

$$H_+ = \{x \in R^n : a^T x \geq c\}, \quad H_- = \{x \in R^n : a^T x \leq c\}$$

and the positive and negative open half spaces

$$H_+^\circ = \{x \in R^n : a^T x > c\}, \quad H_-^\circ = \{x \in R^n : a^T x < c\}.$$

Definition 5.5 (凸多胞形) *A set which can be expressed as the intersection of a finite number of closed half spaces is called a convex polytope.*

Corollary 5.6 *The feasible set $\mathcal{F} = \{x \in R^n : Ax = b, x \geq 0\}$ of standard form LP is a convex polytope since $\mathcal{F}$ can be reduced to*

$$\{x = (x_B, x_D) \in R^m \times R^{n-m} : x_B = B^{-1}b - B^{-1}Dx_D \geq 0, x_D \geq 0\},$$

where $A = [B, D]$ and B is a basis matrix.

Definition 5.7 (凸多面体) *A nonempty bounded polytope is called a polyhedron.*

Theorem 5.8 (分离超平面定理) *Let C be a convex set in R^n and let y be a point exterior to the closure of C. Then there is a nonzero vector $a \in R^n$ such that*

$$a^T y < \inf_{x \in C} a^T x.$$

Proof: Define $\delta = \inf_{x \in C} \|x - y\| > 0$. There is an x_0 on the boundary of C such that $\|x_0 - y\| = \delta$. This follows because the continuous function $f(x) = \|x - y\|$ achieves its minimum over any closed and bounded set and it is clearly only necessary to consider x in the intersection of the closure of C and the sphere of radius 2δ centered at y. We shall show that $a = x_0 - y \neq 0$ satisfies the conclusion of the theorem.

Let $x \in C$. For any $\alpha \in [0,1]$, the point $x_0 + \alpha(x - x_0) \in \text{cl}(C)$ (the closure of C) and thus

$$\|x_0 + \alpha(x - x_0) - y\|^2 \geq \|x_0 - y\|^2.$$

By expanding the left hand side and letting $\alpha \to 0+$, we obtain $(x_0 - y)^T(x - x_0) \geq 0$, which implies

$$(x_0 - y)^T x \geq (x_0 - y)^T x_0 = (x_0 - y)^T y + (x_0 - y)^T(x_0 - y) = (x_0 - y)^T y + \delta^2.$$

Thus, $\inf_{x \in C}(x_0 - y)^T x \geq (x_0 - y)^T y + \delta^2 > (x_0 - y)^T y$. Setting $a = x_0 - y$ completes the proof.

Corollary 5.9 *The hyperplane $H_1 := \{x \in R^n : a^T x = a^T y\}$ contains y and contains C in its open half space $(H_1)^\circ_+$.*

Corollary 5.10 *The hyperplane*

$$H_2 := \left\{x \in R^n : a^T x = a^T \frac{x_0 + y}{2}\right\}$$

separates y and C because

$$a^T y < a^T \frac{x_0 + y}{2} < a^T x$$

for any $x \in C$. Clearly, separating hyperplanes are not unique.

Theorem 5.11 (Projection onto closed convex set) *Let C be a closed convex set in R^n. For any $y \in R^n$, there is a* **unique** *point $x_0 \in C$ such that $\|y - x_0\| = \inf_{x \in C}\|y - x\|$. The point x_0, denoted by $Proj_C(y)$, is called the projection of y onto C and satisfies*

$$(Proj_C(y) - y)^T(x - Proj_C(y)) \geq 0, \ \ \forall\, x \in C.$$

Proof: From the proof of the separating hyperplane theorem, there is a $x_0 \in C$ such that $\|y - x_0\| = \inf_{x \in C}\|y - x\|$ and, for any $x \in C$, it hods that $(x_0 - y)^T(x - x_0) \geq 0$. If there is $x_0' \in C$ different from x_0 and servers the same role as x_0, then $(x_0 - y)^T(x_0' - x_0) \geq 0$ and $(x_0' - y)^T(x_0 - x_0') \geq 0$ lead to a contradiction. ∎

Theorem 5.12 (支撑超平面定理) *Let C be a convex set in R^n and let y be a boundary point of C. Then there is a nonzero vector $a \in R^n$ such that*

$$a^T y \leq \inf_{x \in C} a^T x,$$

i.e., the hyperplane $H = \{x \in R^n : a^T x = a^T y\}$ contains y and contains C in its closed half space H_+. H is thus called the supporting hyperplane of C at y.

Proof: Let $\{y_k\}$ be a sequence of vectors, exterior to the closure of C, converging to y. Let $\{a_k\}$ be the sequence of corresponding vectors constructed in the separating hyperplane theorem, normalized so that $\|a_k\| = 1$, such that

$$a_k^T y_k < \inf_{x \in C} a_k^T x.$$

Taking subsequence if necessary, we assume that $\{a_k\}$ converges to a. Thus, for any $x \in C$, it holds that

$$a^T y = \lim_{k \to \infty} a_k^T y_k \leq \lim_{k \to \infty} a_k^T x = a^T x.$$

The theorem follows by taking infimum on the right hand side with respect to x in C. ∎

Definition 5.13 (Affine hull, 仿射包) *Let S be a subset of R^n. The affine hull of S, denoted by* $\mathrm{aff}(S)$*, contains all the affine combinations of points in S, i.e.,*

$$\mathrm{aff}(S) := \left\{ x = \sum_{i=1}^{k} \alpha_i x_i \in R^n \;\middle|\; \begin{array}{c} x_i \in S, \alpha_i \in R, i = 1, 2, \ldots, k; \\ \sum_{i=1}^{k} \alpha_i = 1; \\ k \text{ is any positive integer.} \end{array} \right\}$$

Equivalently, $\mathrm{aff}(S)$ *is the smallest affine set containing S.*

Definition 5.14 (Relative interior, 相对内部) *Let $B := \{x \in R^n : \|x\| \le 1\}$ be the unit ball in R^n. The relative interior of a convex set C in R^n is defined as*

$$ri(C) := \{x \in \mathrm{aff}(C) : \exists \varepsilon > 0, (x + \varepsilon B) \cap \mathrm{aff}(C) \subset C\}.$$

Definition 5.15 (Cone, 锥) *A set C is called a cone if $x \in C$ implies $\alpha x \in C$ for all $\alpha \ge 0$. A cone that is also convex is called a convex cone.*

Definition 5.16 (Generated cone, 生成锥) *Let $\{x_1, x_2, \ldots, x_m\}$ be a set of m points in R^n. Then*

$$C := \left\{ x = \sum_{i=1}^{m} \alpha_i x_i \in R^n : \alpha_i \ge 0, i = 1, 2, \ldots, m \right\}$$

is a **closed convex cone***, which is called the cone generated by the set of points $\{x_1, x_2, \ldots, x_m\}$.*

Theorem 5.17 (分离超平面定理) *Let C_1 and C_2 be convex sets in R^n with no common relative interior points, i.e., the only common points of C_1 and C_2, if any, are boundary points. Then there is a hyperplane separating C_1 and C_2. In particular, there exist a nonzero vector $a \in R^n$ such that*

$$\sup_{x_1 \in C_1} a^T x_1 \le \inf_{x_2 \in C_2} a^T x_2.$$

Proof: Consider the set $G = C_2 - C_1$. It is easy to see that G is convex and that 0 is not a relative interior point of G. Hence, the separating hyperplane theorem or the supporting hyperplane theorem applies and gives the appropriate hyperplane. ∎

5.2 Farkas lemma

Lemma 5.18 (Farkas Lemma - 1st form) *Let $A \in R^{m \times n}$ and $b \in R^m$. Define*

$$\mathcal{X} := \{x \in R^n : Ax = b, x \ge 0\} \quad \textit{and} \quad \mathcal{Y} := \{y \in R^m : A^T y \le 0, b^T y > 0\}.$$

Then, one and only one of the sets $\mathcal{X}$ and $\mathcal{Y}$ is nonempty (empty).

Proof: "$\mathcal{X} \ne \emptyset \Rightarrow \mathcal{Y} = \emptyset$": Assume $\mathcal{Y} \ne \emptyset$. Take $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Then, $0 \ge (A^T y)^T x = y^T A x = y^T b > 0$ gives a contradiction.

"$\mathcal{X} = \emptyset \Rightarrow \mathcal{Y} \ne \emptyset$": Define $S := \{Ax \in R^m : x \ge 0\}$, which is a nonempty polyhedral set and hence it is **closed** and convex. $\mathcal{X} = \emptyset$ implies that $b \notin S$. According to separating hyperplane theorem, there is a $y \in R^m$ such that

$$b^T y < \inf_{z \in S} y^T z = \inf_{x \ge 0} y^T A x.$$

It follows from $\inf_{x \geq 0} y^T Ax \leq 0$ that $b^T y < 0$. Furthermore, $A^T y \geq 0$ must hold because otherwise

$$b^T y < \inf_{x \geq 0} y^T Ax = -\infty,$$

which is impossible because $b^T y$ is a fixed constant. Thus, $\mathcal{Y} \neq \emptyset$ since $-y \in \mathcal{Y}$. ∎

Remark 5.2.1 *$\mathcal{Y} \neq \emptyset \implies \mathcal{X} = \emptyset$. Thus, a point $y \in \mathcal{Y}$ is called an infeasibility certificate for Primal-LP1. Furthermore, $\mathcal{Y} \neq \emptyset$ implies that Dual-LP1 is either infeasible ($d_1^* = -\infty$) or feasible but unbounded ($d_1^* = +\infty$).*

Lemma 5.19 (Farkas Lemma - 2nd form) *Let $A \in R^{m \times n}$ and $c \in R^n$. Define*

$$\mathcal{X} := \{x \in R^n : Ax = 0, x \geq 0, c^T x < 0\} \quad \textit{and} \quad \mathcal{Y} := \{y \in R^m : A^T y \leq c\}.$$

Then, one and only one of the sets $\mathcal{X}$ and $\mathcal{Y}$ is nonempty (empty).

Proof: Proof the 2nd form of Farkas Lemma is left to yourself. ∎

Remark 5.2.2 *$\mathcal{X} \neq \emptyset \implies \mathcal{Y} = \emptyset$. Thus, a point $x \in \mathcal{X}$ is called an infeasibility certificate for Dual-LP1. Furthermore, $\mathcal{X} \neq \emptyset$ implies that Primal-LP1 is either infeasible ($p_1^* = +\infty$) or feasible but unbounded ($p_1^* = -\infty$).*

The two forms of Farkas Lemma discussed above correspond to Primal-LP1 and Dual-LP1. Farkas Lemma has several other forms corresponds to nonstandard form LPs and their dual problems. The implications of these variants and their corresponding proofs are similar to what we have presented here.

5.3 Strong duality

Theorem 5.20 (Strong Duality for the pair (Primal-LP1, Dual-LP1)) *If either of the problems Primal-LP1 or Dual-LP1 has a finite optimal solution, so does the other, and the corresponding values of the objective functions are equal.*

Proof: Because either problem can be converted to standard form and the roles of primal and dual are reversible (dual to each other), it is sufficient to assume that Primal-LP1 has a finite optimal solution and show that Dual-LP1 has a solution with the same value.

Assume Primal-LP1 has a finite optimal value p_1^*. Define

$$C = \{(r, w) \in R^{m+1} : r = tp_1^* - c^T x, w = tb - Ax, x \geq 0, t \geq 0\}.$$

It can be shown that C is a closed convex cone. First, we claim that $(1, 0) \notin C$ (note that here $(1, 0) = (1, \mathbf{0})$). Otherwise, there exist $x_0 \geq 0$, $t_0 \geq 0$ such that $1 = t_0 p_1^* - c^T x_0$ and $0 = t_0 b - Ax_0$. If $t_0 > 0$, then x_0/t_0 is feasible for Primal-LP1. Then, $1/t_0 = p_1^* - c^T(x_0/t_0) \leq 0$ is a contradiction. If $t_0 = 0$, then $c^T x_0 = -1$, $Ax_0 = 0$ and $x_0 \geq 0$. Let x be any feasible solution of Primal-LP1. Then, for any $\alpha \geq 0$, $x + \alpha x_0$ is also a feasible solution of Primal-LP1. The corresponding function value is

$$c^T(x + \alpha x_0) = c^T x + \alpha c^T x_0,$$

which goes to $-\infty$ as $\alpha \to +\infty$. This contradicts to the fact that p_1^* is finite. In summary, $(1, 0) \notin C$.

Since C is a closed convex set and $(1, 0) \notin C$, by the separating hyperplane theorem, there is a nonzero vector $(s, \lambda) \in R \times R^m$ and a constant **const** such that

$$s = (s, \lambda)^T (1, 0) < \mathbf{const} = \inf\{sr + \lambda^T w : (r, w) \in C\}.$$

Since C is a cone, $sr + \lambda^T w \geq 0$ must hold for any $(r, w) \in C$. Otherwise, if there were $(r, w) \in C$ such that $sr + \lambda^T w < 0$, then $\alpha(r, w) \in C$ for any $\alpha \geq 0$. It thus follows that

$$s < \inf\{\alpha(sr + \lambda^T w) : (r, w) \in C\} = -\infty,$$

which is clearly a contradiction since s is a fixed constant. Therefore, $\mathbf{const} \geq 0$. On the other hand, $\mathbf{const} \leq 0$ since $(0, 0) \in C$. As a result, $\mathbf{const} = 0$ and $s < 0$. Without loss of generality, we assume that $s = -1$.

From the above, there exists $\lambda \in R^m$ such that $-r + \lambda^T w \geq 0$ for all $(r, w) \in C$. Equivalently, using the definition of C,

$$(c^T - \lambda^T A)x - tp_1^* + t\lambda^T b \geq 0$$

for all $x \geq 0$ and $t \geq 0$. Setting $t = 0$ yields $A^T\lambda \leq c$, which implies that λ is feasible for Dual-LP1. Setting $x = 0$ and $t = 1$ yields $p_1^* \leq b^T\lambda$, which in view of the weak duality and its corollary shows that λ is optimal for Dual-LP1. Therefore, $p_1^* = d_1^*$ and no duality gap. ■

Remark 5.3.1 *Alternatively, after proving $(1, \mathbf{0}) \notin C$, we can use* **Farkas Lemma** *(2nd form). In fact, $(1, \mathbf{0}) \notin C$ implies that the system*

$$(A, -b)\begin{bmatrix} x \\ t \end{bmatrix} = 0, \quad \begin{bmatrix} x \\ t \end{bmatrix} \geq 0, \quad \begin{bmatrix} c \\ -p_1^* \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} = -1$$

has no feasible solution. According to **Farkas Lemma** *(2nd form), there exists $\lambda \in R^m$ such that*

$$(A, -b)^T\lambda = \begin{bmatrix} A^T \\ -b^T \end{bmatrix}\lambda \leq \begin{bmatrix} c \\ -p_1^* \end{bmatrix},$$

or equivalently, $A^T\lambda \leq c$ and $b^T\lambda \geq p_1^*$. $A^T\lambda \leq c$ implies that λ is feasible for Dual-LP1, while $b^T\lambda \geq p_1^*$, together with weak duality, implies that $d_1^* = p_1^*$.

5.4 The optimality system

The following result is a corollary of strong duality.

Theorem 5.21 *Consider the pair of linear programs Primal-LP1 and Dual-LP1. Assume that both are feasible. Then, $x \in R^n$ and $(\lambda, s) \in R^m \times R^n$ are, respectively optimal for Primal-LP1 and Dual-LP1 if and only if the following conditions are all satisfied:*

$$\begin{aligned} \textbf{Primal feasibility}: &\quad Ax = b, \quad && x \geq 0 \\ \textbf{Dual feasibility}: &\quad A^T\lambda + s = c, \quad && s \geq 0 \\ \textbf{Complementarity}: &\quad c^T x - b^T\lambda = 0 \quad && (or\ \ s^T x = 0). \end{aligned}$$

5.5 Summary

The primal and dual LPs can be infeasible simultaneously. A trivial example: $\min_x\{x : \ s.t.\ 0 \cdot x \geq 1\}$, whose dual problem is $\max_\mu\{\mu : \ s.t.\ 0 \cdot \mu = 1, \mu \geq 0\}$.

$$\begin{array}{rl} \min & -4x_1 + 2x_2 \\ \text{s.t.} & -x_1 + x_2 \geq 2 \\ & x_1 - x_2 \geq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

(a) Primal LP

$$\begin{array}{rl} \max & 2\lambda_1 + \lambda_2 \\ \text{s.t.} & -\lambda_1 + \lambda_2 \leq -4 \\ & \lambda_1 - \lambda_2 \leq 2 \\ & \lambda_1, \lambda_2 \geq 0 \end{array}$$

(b) Dual LP

Figure 5.1: It can be verified that the primal and the dual LPs are infeasible simultaneously.

Table 5.1: All possible cases for Primal-LP1 and Dual-LP1. “unbounded" means feasible and unbounded below for minimization and feasible and unbounded above for maximization.

Dual-LP1	Primal-LP1	$d_1^* \leq p_1^*$
feasible	feasible	$d_1^* = p_1^*$
infeasible	infeasible	$-\infty = d_1^* < p_1^* = +\infty$
unbounded	infeasible	$d_1^* = p_1^* = +\infty$
infeasible	unbounded	$-\infty = d_1^* = p_1^*$

Table 5.2: Primal-LP1 and Dual-LP1: the rest 5 cases cannot happen!

	p_1^* finite	$p_1^* = -\infty$	$p_1^* = +\infty$
d_1^* finite	$\surd$	$\times$	$\times$
$d_1^* = +\infty$	$\times$	$\times$	$\surd$
$d_1^* = -\infty$	$\times$	$\surd$	$\surd$

“$\surd$" means possible, “$\times$" means impossible.

5.6 Relations to the simplex procedure

Theorem 5.22 (Strong duality) *Suppose Primal-LP1 has an optimal basic feasible solution corresponding to the basis B. Then the vector λ satisfying $\lambda^T = c_B^T B^{-1}$ is an optimal solution to Dual-LP1. The optimal values of both problems are equal.*

Proof: Let $A = [B, D]$. Since x is optimal, we have

$$r_D^T = c_D^T - c_B^T B^{-1} D \geq 0.$$

It is easy to verify that λ is dual feasible and $\lambda^T b = c^T x$. Therefore, λ is dual optimal. ∎

Where to find the dual optimal solution λ in the final simplex tableau? Suppose we have found an initial basic feasible solution (by phase I of two-phase method, or by introducing slack variables; In all, there is an identity matrix in the reduced form of $[A, b]$). The first and the last simplex tableaus are

$$\begin{bmatrix} (\cdots) & I & (\vdots) \\ (\cdots) & c_I^T & 0 \end{bmatrix} \text{ and } \begin{bmatrix} B^{-1}(\cdots) & B^{-1} & B^{-1}(\vdots) \\ (\cdots) - c_B^T B^{-1}(\cdots) & c_I^T - c_B^T B^{-1} & -c_B^T B^{-1}(\vdots) \end{bmatrix},$$

respectively, where B is the optimal basis. $c_B^T B^{-1}$ is in the final tableau!

$$\begin{array}{rl} \min & -x_1 - 4x_2 - 3x_3 \\ \text{s.t.} & 2x_1 + 2x_2 + x_3 + x_4 = 4 \\ & x_1 + 2x_2 + 2x_3 + x_5 = 6 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{array}$$

(a) Primal-LP1

$$\begin{array}{rl} \max & 4\lambda_1 + 6\lambda_2 \\ \text{s.t.} & 2\lambda_1 + \lambda_2 \leq -1 \\ & 2\lambda_1 + 2\lambda_2 \leq -4 \\ & \lambda_1 + 2\lambda_2 \leq -3 \\ & \lambda_1, \lambda_2 \leq 0. \end{array}$$

(b) Dual-LP1

Figure 5.2: A pair of primal and dual LPs.

Example 5.6.1 *Consider the pair of primal and dual LPs in Figure 5.2. The first and last simplex tableaus are, respectively,*

$$\begin{array}{cccccc} 2 & 2 & 1 & 1 & 0 & 4 \\ 1 & 2 & 2 & 0 & 1 & 6 \\ -1 & -4 & -3 & 0 & 0 & 0 \end{array} \quad \textit{and} \quad \begin{array}{cccccc} 3/2 & 1 & 0 & 1 & -1/2 & 1 \\ -1 & 0 & 1 & -1 & 1 & 2 \\ 2 & 0 & 0 & 1 & 1 & 10 \end{array}$$

The dual optimal solution is $\lambda^T = (-1, -1)$!

5.7 Explanations of dual variables

5.7.1 Simplex multipliers

In the iteration of the simplex method, the vector $\lambda^T = c_B^T B^{-1}$ changes with B from iteration to iteration, and it is not optimal for the dual problem until an optimal basis B is found for the primal problem. Since λ is used to compute the relative cost vector by

$$r_D^T = c_D^T - \lambda^T D,$$

λ is often called the vector of *simplex multipliers*. Take the diet problem (with equality constraints) for example. Note that each column vector of A corresponds to a certain food. Suppose B is the current basis (consisting certain m foods). All other foods can be synthetically constructed from these m foods. In particular, one unit of the ith nutrient (corresponds to the ith column vector of the identity matrix of order m, denote it by e_i) can be synthetically constructed as

$$e_i = By,$$

where $y = B^{-1}e_i$. The synthetical price of one unit of the ith nutrient is

$$c_B^T y = c_B^T B^{-1} e_i = \lambda^T e_i = \lambda_i.$$

Thus, at each iteration, λ_i represents the *synthetic price* of the ith nutrient under the current basis, and $\lambda^T a_j$ is the *synthetic price* of the jth food. When optimality for primal problem is attained, synthetic price of any food a_j must be no more expensive than the market price c_j, i.e., $\lambda^T A \leq c^T$ (dual feasibility). Simplex method (for primal problem) maintains primal feasibility, 0-duality gap, because

$$c^T x - \lambda^T b = c_B^T x_B - c_B^T B^{-1} b = c_B^T x_B - c_B^T x_B = 0,$$

and keeps improving the primal objective function value until it cannot be improved further, that is, when dual feasibility is met by λ.

5.7.2 Marginal price

Consider $p^*(b) := \min\{c^T x : s.t.\ Ax = b, x \geq 0\}$, where the optimal value is viewed as a function of b. Suppose the optimal basis is B. Primal and dual optimal solutions are $x = (x_B, 0) = (B^{-1}b, 0)$ and $\lambda^T = c_B^T B^{-1}$, respectively. Assuming nondegeneracy (which implies $B^{-1}b > 0$), then small changes in b will not cause the optimal basis to change. This is because

$$B^{-1}b > 0 \quad \text{and} \quad r_D = c_D^T - c_B^T B^{-1} D \geq 0$$

remain unaffected if the change in b is small. Suppose Δb is small enough such that $B(b + \Delta b) > 0$. Then,

$$(x_B + \Delta x_B, 0) = (B^{-1}b + B^{-1}\Delta b, 0)$$

is an optimal solution corresponding to $b + \Delta b$. The increment in the cost function is

$$\Delta p^* = p^*(b + \Delta b) - p^*(b) = c_B^T \Delta x_B = \lambda^T \Delta b,$$

which implies that

$$\lambda = \frac{\partial p^*}{\partial b} = \nabla p^*(b).$$

Therefore, λ gives the sensitivity of the optimal cost with respect to small changes in b. Thus, λ_i is also known as the *marginal price* of the component b_i, since if b_i is changed to $b_i + \Delta b_i$ the optimal value changes by $\lambda_i \Delta b_i$.

5.7.3 Shadow price

Consider the manufacturing problemand its dual given in Figure 5.3. Optimal solutions for primal and dual are, respectively,

$$x = (4, 2)^T \ \text{ and } \ \lambda = (3/2, 1/8, 0)^T.$$

Note that $p^* = b^T \lambda$ and $\partial p^* / \partial b = \lambda$. Therefore, λ represents the sensitivity of p^* with respect to b, given all other conditions (A and c).

From Figure 5.4, it can be seen that

$$
\begin{aligned}
p^* := \max \quad & 2x_1 + 3x_2 \\
\text{s.t.} \quad & x_1 + 2x_2 \leq 8 \\
& 4x_1 \leq 16 \\
& 4x_2 \leq 12 \\
& x_1, x_2 \geq 0.
\end{aligned}
$$

(a) Primal-LP1

$$
\begin{aligned}
d^* := \min \quad & 8\lambda_1 + 16\lambda_2 + 12\lambda_3 \\
\text{s.t.} \quad & \lambda_1 + 4\lambda_2 \geq 2 \\
& 2\lambda_1 + 4\lambda_3 \geq 3 \\
& \lambda_1, \lambda_2, \lambda_3 \geq 0.
\end{aligned}
$$

(b) Dual-LP1

Figure 5.3: Primal and dual of the manufacturing problem.

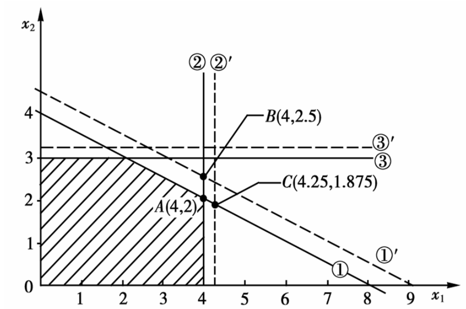

Figure 5.4: Illustration of the manufacturing problem.

1. If available machine hours are increased by 1 (replace "$x_1 + 2x_2 \leq 8$" by "$x_1 + 2x_2 \leq 9$"), then the neat profit will increase by $\lambda_1 = 3/2$.
2. If resource I is increased by 1 (replace "$4x_1 \leq 16$" by "$4x_1 \leq 17$"), then the neat profit will increase by $\lambda_2 = 1/8$.
3. If resource II is increased by 1 (replace "$4x_2 \leq 12$" by "$4x_2 \leq 13$"), then the neat profit remains unchanged ($\lambda_3 = 0$).

If sell out all resource, the selling price should be prime cost plus shadow price. If the market price of b_i is less than its shadow price, then purchase b_i and expand production; If the otherwise, then sell b_i and shrink production; λ is called the *shadow price* of resource b (closely related to current production plan, varies from situation to situation).

Theorem 5.23 *Consider the following pair of LPs:*

$$\min\{c^T x : s.t.\ Ax \geq b, x \geq 0\} \;\; \textit{and} \;\; \max\{b^T \lambda : s.t.\ A^T \lambda \leq c, \lambda \geq 0\}.$$

Let x and λ be, respectively, primal and dual feasible. Denote the jth row of A by a^j. A necessary and sufficient condition that they both be optimal is that for all i and j

- $x_i > 0 \Rightarrow \lambda^T a_i = c_i$
- $x_i = 0 \Leftarrow \lambda^T a_i < c_i$
- $\lambda_j > 0 \Rightarrow a^j x = b_j$
- $\lambda_j = 0 \Leftarrow a^j x > b_j$.

Proof: Left to yourself. ■

Think about the primal LP as the diet problem and give explanations of these four relations.

5.8 Dual simplex method

Dual simplex method maintains dual feasibility, 0-duality gap and keeps improving the dual objective function value until it cannot be improved further (i.e., when primal feasibility is satisfied):

1. starting from an initial dual feasible basic solution, i.e., $x = (B^{-1}b, 0)$, where $B^{-1}b \geq 0$ may not be true but $\lambda^T A = c_B^T B^{-1} A \leq c^T$.
2. at each iteration first determine one variable to leave basis (so that dual objective can be increased)
3. then choose one variable appropriately to enter basis (so that dual feasibility can be maintained)
4. do pivoting and repeat until optimal (i.e., primal feasibility is satisfied).

Given the current dual feasible basis B, i.e., the corresponding dual variable $\lambda^T = c_B^T B^{-1}$ is feasible for the dual problem (i.e., $\lambda^T A \leq c^T$). If $x_B = B^{-1}b \geq 0$, then B is also primal feasible and thus optimal. By assuming nondegeneracy, we see that if B is not primal feasible, then there must be i such that $(x_B)_i < 0$. This i determines which variable is going to leave basis. For simplicity, we assume the current basis contains the first m columns of A. It is easy to verify that

$$\lambda^T a_j \begin{cases} = c_j, & j = 1, 2, \ldots, m; \\ < c_j, & j = m+1, \ldots, n. \end{cases}$$

To develop one cycle of the dual simplex method, we find a new vector $\bar{\lambda}$ such that one of the equalities becomes an inequality and one of the inequality becomes equality, while at the same time increasing the value of the dual objective function. The m equalities in the new solution then determine a new basis. Denote the ith row of B^{-1} by u^i. Then for

$$\bar{\lambda}^T = \lambda^T - \varepsilon u^i,$$

we have

$$\bar{\lambda}^T a_j = \lambda^T a_j - \varepsilon u^i a_j = \lambda^T a_j - \varepsilon y_{ij}.$$

Here y_{ij} is the ijth element of the tableau. Therefore,

$$\begin{cases} \bar{\lambda}^T a_j = c_j, & j \in \{1, \ldots, m\} \setminus \{i\}; \\ \bar{\lambda}^T a_j = c_i - \varepsilon, & j = i; \\ \bar{\lambda}^T a_j = \lambda^T a_j - \varepsilon y_{ij}, & j \in \{m+1, \ldots, n\}. \end{cases}$$

The dual objective value at $\bar{\lambda}$:

$$\bar{\lambda}^T b = \lambda^T b - \varepsilon u^i b = \lambda^T b - \varepsilon (x_B)_i.$$

The above arguments lead to the dual simplex method:

1. Given a dual basic feasible solution x_B. If $x_B \geq 0$, then already optimal. Otherwise, find i such that $(x_B)_i < 0$.
2. If $y_{ij} \geq 0$ for all $j = 1, 2, \ldots, n$, then the dual problem is unbounded above.
3. If $y_{ij} < 0$ for some j, then let

$$\varepsilon_0 = \frac{(A^T\lambda)_k - c_k}{y_{ik}} = \min_j \Big\{ \frac{(A^T\lambda)_j - c_j}{y_{ij}} \;:\; y_{ij} < 0 \Big\}.$$

4. Compute $\bar{\lambda}^T = \lambda^T - \varepsilon_0 u^i$. Form a new basis B by replacing a_i by a_k. Use this basis to determine the new x_B and repeat.

Example 5.8.1 (Dual simplex method)

$$\begin{aligned} \min \quad & 3x_1 + 4x_2 + 5x_3 \\ s.t. \quad & x_1 + 2x_2 + 3x_3 \geq 5 \\ & 2x_1 + 2x_2 + x_3 \geq 6 \\ & x_1, x_2, x_3 \geq 0. \end{aligned}$$

The initial tableau of dual simplex method:

x_1	x_2	x_3	x_4	x_5	b
-1	-2	-3	1	0	-5
(-2)	-2	-1	0	1	-6
3	4	5	0	0	0

($x = (0,0,0,-5,-6)^T, c_B^T B^{-1} A = 0 \leq c$.) First choose x_5 to leave basis because -6 is the most negative one among $\{-5,-6\}$; then choose (-2) as pivot element because $3/(-2)$ is the maximum negative ratio among all negative ratios

$$\{3/(-2), 4/(-2), 5/(-1)\}.$$

The second tableau of dual simplex method:

x_1	x_2	x_3	x_4	x_5	b
0	(-1)	-5/2	1	-1/2	-2
1	1	1/2	0	-1/2	3
0	1	7/2	0	3/2	9

Choose (-1) as pivot element because $-2 < 0$ and

$$1/(-1) = \max\{1/(-1), (7/2)/(-5/2), (3/2)/(-1/2)\}.$$

The final tableau of dual simplex method:

x_1	x_2	x_3	x_4	x_5	b
0	1	5/2	-1	1/2	2
1	0	-2	1	-1	1
0	0	1	1	1	11

Optimal solutions for primal and dual LPs are, respectively, $x = (1,2,0)^T$ and $\lambda = (-1,-1)^T$.

The primal simplex method maintains primal feasibility, 0-duality gap and keeps decreasing primal objective function value until dual feasibility. On the contrary, the dual simplex method maintains dual feasibility, 0-duality gap and keeps increasing dual objective function value until primal feasibility.

References

[Luenberger-Ye] David G. Luenberger and Yinyu Ye Linear and nonlinear programming.