Integer Dualities

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, let us consider the binary programming problem

$$z_{IP} = \max\{c^T x \mid Ax \le b,\ x \in \{0,1\}^n\}, \tag{1}$$

and let z_{LP} denote the optimum of its continuous relaxation, i.e.,

$$z_{LP} = \max\{c^T x \mid Ax \le b,\ x \in [0,1]^n\}.$$

In this section we shall consider a series of upper bounds to the maximum of (1).

Lagrangean Duality

Given a subset $I \subseteq [m] = \{1, 2, ..., m\}$ of the row indices, let $J = [m] \setminus I$ and for simplicity let us denote by $A_I = A[I;]$ and $A_J = A[J;]$ the submatrices of A formed by the rows with indices in I and J, respectively. Similarly, let b_I and b_J denote the subvectors of $b \in \mathbb{R}^m$ formed by the components with indices belonging to I and J, respectively.

For an arbitrary $\lambda \in \mathbb{R}^I_+$ let us consider the following *Lagrangean relaxation* of (1):

$$z_I(\lambda) = \lambda^T b_I + \max\left\{(c^T - \lambda^T A_I)x \;\middle|\; A_J x \le b_J,\ x \in \{0,1\}^n\right\}. \tag{2}$$

Lemma 1 *For every $I \subseteq [m]$ the function $z_I(\lambda)$ is piecewise linear and convex. Furthermore, for every $\lambda \in \mathbb{R}^I_+$ we have*

$$z_I(\lambda) \ge z_{IP}.$$

Proof. There are just finitely many binary vectors, so piecewise linearity is obvious.

For $\lambda, \lambda' \in \mathbb{R}^I_+$ let us consider the optimal solution $x^* \in \{0,1\}^n$ attaining $z_I(\frac{1}{2}\lambda + \frac{1}{2}\lambda')$. This is also feasible solution in the problems computing $z_I(\lambda)$ and $z_I(\lambda')$. Thus, convexity follows.

Finally, let us consider the optimal solution $\bar{x}$ attaining z_{IP}. This is also feasible in the problem computing $z_I(\lambda)$, and thus the inequality $z_I(\lambda) \ge z_{IP}$ follows. $\square$

Thus, to get the best upper bound to z_{IP} of the type (2), we need to solve the so called *Lagrangean dual problem*:

$$z_{LD}(I) = \min\left\{z_I(\lambda) \;\middle|\; \lambda \in \mathbb{R}^I_+\right\} \tag{3}$$

Due to Lemma 1, problem (3) is a convex minimization over a convex domain, and hence it can be solved efficiently – as long as the function values $z_I(\lambda)$ can be computed efficiently.

Theorem 1 *If the maximization problem in* (2) *can be solved in polynomial time, then* $z_{LD}(I)$ *can also be computed in polynomial time.*

Proof. Consider the set

$$X = \{x \in \{0,1\}^n \mid A_J x \leq b_J\}$$

and the polyhedron

$$P = \{(\lambda, \nu) \mid \nu \geq \lambda^T b_I + (c^T - \lambda^T A_I)x \;\; \forall x \in X\}.$$

Then we have[1]

$$z_{LD}(I) = \min\{\nu \mid (\lambda, \nu) \in P\}.$$

Since the separation problem for P can be solved by solving the maximization problem[2] in (2), $z_{LD}(I)$ can indeed be computed in polynomial time by using the ellipsoid method, whenever the maximization problem in (2) can be solved efficiently. □

Lemma 2 *We have* $z_{LP} = z_{LD}([m])$, $z_{IP} = z_{LD}(\emptyset)$, *and* $z_{LD}(I) \geq z_{LD}(I')$ *whenever* $I \supseteq I'$.

Proof. Homework ... □

For the general case, Held and Karp (1970,1971) proposed a column generation technique to compute $z_{LD}(I)$, which is similar to the decomposition technique proposed by Benders (1960, 1962) (cf. Dantzig and Wolf (1960,1961)). Note that if for some set I the continuous relaxation of the

[1]WHY?
[2]WHY??

maximization problem in (2) has integral optimum, then by LP-duality we have $z_{LD}(I) = z_{LP}$. This observation helped to apply this technique to solve some traveling salesman problems (see Held and Karp (1970, 1971)).

Lagrangean duality was introduced and studied by Lorie and Savage (1955), Everett (1963), Nemhauser and Ullman (1968), and Geoffrion (1974).

Surrogate Duality

Given $u \in \mathbb{R}^m_+$, define

$$z_S(u) = \max\left\{c^T x \;\middle|\; u^T A x \leq u^T b,\ x \in \{0,1\}^n\right\}, \tag{$S(u)$}$$

and let

$$z_{SD} = \min\{z_S(u) \mid u \in \mathbb{R}^m_+\}. \tag{SD}$$

Lemma 3 *For all $u \in \mathbb{R}^m_+$ we have $z_S(u) \geq z_{IP}$, and*

$$z_{LP} \geq z_{SD} \geq z_{IP}.$$

Proof. See in class ... □

Given an optimization problem (P), let us denote by $\Omega(P)$ the set of its optimal solutions, and by $\Phi(P)$ the set of its feasible solutions. Let further

$$U(Z) = \{u \in \mathbb{R}^m_+ \mid z_S(u) \leq Z\}.$$

Lemma 4 *For all $Z \in \mathbb{R}$ the set $U(Z)$ is convex.*

Proof. Assume that $u, v \in U(Z)$ and $w = \frac{1}{2}u + \frac{1}{2}v \notin U(Z)$ for some $Z \in \mathbb{R}$, i.e., that $z_S(w) > Z$, $z_S(u) \leq Z$, and $z_S(v) \leq Z$. Thus, $\Phi(S(w)) \neq \emptyset$, and hence there exists $x^w \in \Omega(S(w))$. Then, $x_w \notin \Phi(S(u)) \cup \Phi(S(v))$, and hence we must have $u^T A x^w > u^T b$ and $v^T A x^w > v^T b$, implying $w^T A x^w > w^T b$, i.e., that $x^w \notin \Phi(S(w))$, contradicting the selection of $x^w \in \Omega(S(w)) \subseteq \Phi(S(w))$. □

Lemma 5 *If $z_S(v) < z_S(u)$ for some $u, v \in \mathbb{R}^m_+$, and $x^u \in \Omega(S(u))$, then we must have*

$$v^T(b - Ax^u) < 0.$$

Proof. $\Omega(S(u)) \cap \Phi(S(v)) = \emptyset$. $\square$

Let us assume now that A and b are integral, and let

$$L = \max_{x \in \{0,1\}^n} \|Ax - b\|.$$

Note that $\log L = poly(size(A), size(b))$.

Lemma 6 *Assuming $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, we have for all $u \in \mathbb{Z}_+^m$ and all $v \in \mathbb{R}_+^m$ with $\|u - v\| < \frac{1}{L}$ that $v \in U(z_S(u))$.*

Proof. We claim that

$$\Phi(S(v)) \subseteq \Phi(S(u)).$$

To see this, let us assume indirectly that $x \in \Phi(S(v)) \setminus \Phi(S(u))$, i.e., that

$$v^T Ax \leq b \quad \text{and} \quad u^T Ax \geq u^T b + 1$$

(remember, A, b, x and u are all integers). From this we get

$$(v^T - u^T)(b - Ax) \geq 1.$$

Since we assumed $\|u - v\| < \frac{1}{L}$, and since $L \geq \|b - Ax\|$ by definition, we get a contradiction by the Cauchy-Schwartz inequality. $\square$

Lemma 7 *There exists an integer vector $u \in U(z_{SD}) \cap \mathbb{Z}_+^m$ for which $u \leq L^m$.* $\square$

Lemma (5) proves that the separation problem for the convex set $U(z_{SD})$ can be solved by solving a knapsack problem in n binary variables. Lemmas (6) and (7) show that the ellipsoid algorithm will stop in polynomially many steps (in the size of A and b, hence proving that the surrogate dual z_{SD} can be computed by solving polynomially many knapsack problems (see Boros(1985)).

Surrogate duality was introduced by Glover (1968, 1975) and Greenberg and Pierskalla (1970). First solutions (see e.g., Karwan and Rardin (1979), Dyer (1980)) are based on the simplex algorithm without guaranteeing polynomiality of the number of knapsack problems that have to be solved.

Superadditive Duality

A function $F : \mathbb{R}^m \to \mathbb{R}$ is called *superadditive* if

(i) $F(x) + F(y) \leq F(x+y)$ for all $x, y \in \mathbb{R}^m$,

(ii) $F(0) \leq 0$, and

(iii) $F(x) \leq F(y)$ whenever $x \leq y$.

Fact 1 *If F, G are superadditive $\mathbb{R}^m \to \mathbb{R}$ functions, then so are λF for $\lambda \in \mathbb{R}_+$, $\lfloor F \rfloor$, $F + G$ and* $\min\{F, G\}$. $\square$

Fact 2 *If $F_j : \mathbb{R}^m \to \mathbb{R}$ for $j = 1, ..., k$ and $H : \mathbb{R}^k \to \mathbb{R}$ are all superadditive, then so is their composition $H(F_1, \ldots, F_k)$.* $\square$

Given $A \in \mathbb{Q}^{m \times n}$ and $b \in \mathbb{R}^m$, let $P(b) = \{x \in \mathbb{R}^n_+ \mid Ax \leq b\}$ and let $P_I(b) = \operatorname{conv}(P(b) \cap \mathbb{Z}^n_+)$. We shall assume in the sequel that $P_I(b)$ **is bounded for all** $b \in \mathbb{R}^m$, or equivalently[3] that $\{x \in \mathbb{R}^n_+ \mid Ax \leq 0\} = \{0\}$.

We say that an inequality $a^T x \leq d$ is *valid for $P_I(b)$* if all vectors $x \in P_I(b)$ satisfy it. Furthermore, let us denote by a_j, $j = 1, ..., n$ the column vectors of matrix A, as before,.

Lemma 8 *If F is superadditive, then*

$$\sum_{j=1}^{n} F(a_j) x_j \leq F(b)$$

is valid for $P_I(b)$.

Proof. For an arbitrary $x \in P_I(b)$ we have

$$\sum_{j=1}^{n} F(a_j) x_j \leq F(\sum_{j=1}^{n} a_j x_j) = F(Ax) \leq F(b)$$

[3]WHY is this equivalent? Hint: if x^1, x^2,... is an infinite sequence of nonnegative integer vectors then we must have indices $i \neq j$ such that $x^i \leq x^j$.

by (i), (iii) and by the nonnegativity and integrality of x. $\square$

For $c \in \mathbb{R}^n$ let us define

$$F_c(b) = \begin{cases} \max\left\{c^T x \mid x \in P_I(b)\right\} & \text{if } P_I(b) \neq \emptyset, \\ -\infty & \text{otherwise.} \end{cases} \tag{4}$$

Let us note that since we assume $P_I(b)$ to be bounded, the maximum in (4) is always finite. The function F_c is called the value function of the integer programming problem $\max\left\{c^T x \mid x \in P_I(b)\right\}$ (see e.g., Everett (1963), Johnson (1973), Geoffrion (1974), and for some special cases Gilmore and Gomory (1966) and Gomory (1965,1967)).

Lemma 9 *$F_c(b)$ is superadditive.*

Proof. Clearly, we have $P_I(b) \subseteq P_I(b')$ whenever $b \leq b'$, thus $F_c(b) \leq F_c(b')$ follows. Furthermore, $\{0\} = P_I(0)$ by our assumption of boundedness of the $P_I(b)$ polyhedra, and hence we have $F_c(0) = 0$. To complete the proof we need to show that

$$F_c(b) + F_c(b') \leq F_c(b + b') \tag{5}$$

holds for all $b, b' \in \mathbb{R}^m$. Clearly, this inequality holds if $P_I(b) = \emptyset$ or if $P_I(b') = \emptyset$. Otherwise we have

$$P_I(b) \oplus P_I(b') = \{x + x' \mid x \in P_I(b) \text{ and } x' \in P_I(b')\} \subseteq P_I(b + b').$$

In particular, if x and x' are optimal solutions to the respective maximization problems in (4), i.e., if $c^T x = F_c(b)$ and $c^T x' = F_c(b')$, then we have $x + x' \in P_I(b + b')$, implying (5). $\square$

Lemma 10 *We have*

$$F_c(a_j) \geq c_j$$

for all $j = 1, ..., n$.

Proof. Denoting by e_j the j-th unit vector, we have $e_j \in P_I(a_j)$ for $j = 1, ..., n$. $\square$

Lemma 11 *If $\gamma^T x \leq \beta$ is a valid inequality for $P_I(b)$ then we have*

$$\gamma^T x \leq \sum_{j=1}^{n} F_\gamma(a_j)x_j \leq F_\gamma(b) \leq \beta$$

for all $x \in P_I(b)$.

Proof. The first inequality is implied by Lemma 10, while the second one follows by Lemma 8. To see the last inequality, note that for all $x \in P_I(b)$ we have $\gamma^T x \leq \beta$ by the validity if this inequality. $\square$

Theorem 2 *All tightest valid inequalities arise from superadditive functions as in Lemma 8.*

Proof. According to lemma 11, if $\gamma^T x \leq \beta$ is a valid inequality, then $\sum_{j=1}^{n} F_\gamma(a_j)x_j \leq F_\gamma(b)$ is another valid inequality, which is at least as tight as $\gamma^T x \leq \beta$. $\square$

Given $A \in \mathbb{Q}^{m\times n}$ and $c \in \mathbb{R}^n$, let us introduce

$$\mathcal{F}_c \;=\; \{F \mid \text{ superadditive, and } F(a_j) \geq c_j \text{ for } j = 1, ..., n\}\,.$$

Theorem 3 (Weak superadditive duality) *For every $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $x \in P_I(b)$ and for every superadditive function $F \in \mathcal{F}_c$ we have*

$$c^T x \leq F(b).$$

Proof. Since x is nonnegative and F is superadditive, we can write

$$c^T x \leq \sum_{j=1}^{n} F(a_j)x_j \leq F(Ax) \leq F(b).$$

$\square$

Theorem 4 (Strong superadditive duality) *For every $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ we have*

$$\max_{x\in P_I(b)} c^T x \quad = \quad \min_{F\in\mathcal{F}_c} F(b).$$

Proof. Note that Theorem 3 and by the boundedness of $P_I(b)$ we have

$$\max_{x \in P_I(b)} c^T x \leq \min_{F \in \mathcal{F}_c} F(b).$$

To see the reverse inequality, let us observe that by Lemmas 9 and 10 we have $F_c \in \mathcal{F}_c$, and thus applying Lemma 11 to the valid inequality $c^T x \leq \beta = \max_{y \in P_I(b)} c^T y$ we get

$$\min_{F \in \mathcal{F}_c} F(b) \leq F_c(b) \leq \beta = \max_{x \in P_I(b)} c^T x.$$

$\square$

Theorem 5 (Superadditive complementary slackness) *Let $x^* \in P_I(b)$ be an optimal solution for the optimization problem $\max\{c^T x \mid x \in P_I(b)\}$, and let $F^* \in \mathcal{F}_c$ be an optimal solution for the dual optimization problem $\min\{F(b) \mid F \in \mathcal{F}_c\}$. Then, for every $0 \leq x \leq x^*$ we have*

$$F^*(Ax) = c^T x \quad \textit{and} \quad F^*(Ax) + F^*(b - Ax) = F^*(b).$$

Proof. Let us note that for $b = Ax$ and $b' = A(x^* - x)$ we have $x \in P_I(b)$ and $x^* - x \in P_I(b')$, and thus applying Theorem 3 we get

$$c^T x \leq F^*(Ax) \quad \text{and} \quad c^T(x^* - x) \leq F^*(A(x^* - x)).$$

Adding these inequalities, and using the superadditivity of F^* and the fact that $Ax^* \leq b$ we obtain

$$c^T x^* \leq F^*(Ax) + F^*(A(x^* - x)) \leq F^*(Ax) + F^*(b - Ax) \leq F^*(b) = c^T x^*$$

implying the statement. $\square$

The theory of superadditive duality appears in Johnson (1973), Gomory and Johnson (1973), and later in Blair and Jeroslow (1977), Jeroslow (1978), etc. It is strongly related to the so called *group problem*, introduced by Gomory (1965).

Chvátal-Gomory Cuts

Let us note that for $u \in \mathbb{R}^m_+$ the funtion $F_u(y) = \lfloor u^T y \rfloor$ is a superadditive function over $y \in \mathbb{R}^m$.

Let us then consider the integer programming problem

$$P_I \; = \; \{x \in \mathbb{Z}^n_+ \mid Ax \leq b\}$$

for some $A = [a^1, ..., a^n] \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then by the above theorems we have

$$\sum_{j=1}^{n} F_u(a^j) x_j \leq F_u(b)$$

is a valid inequality for P_I for an arbitrary $u \in \mathbb{R}^m_+$. These inequalities are named Chvátal-Gomory cuts (or in short CG-cuts) after the works of R. Gomory (1965) and V. Chvátal (1973).