Journal of Global Optimization (2006) 35: 255–282
DOI 10.1007/s10898-005-3838-0

Towards Strong Duality in Integer Programming*

D. LI[1] and X.L. SUN[2]
[1]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong* (*e-mail: dli@se.cuhk.edu.hk*)
[2]*Department of Mathematics, Shanghai University, Baoshan, Shanghai 200444, P. R. China* (*e-mail: xlsun@staff.shu.edu.cn*)

(Accepted 6 October 2005)

Abstract. We consider in this paper the Lagrangian dual method for solving general integer programming. New properties of Lagrangian duality are derived by a means of perturbation analysis. In particular, a necessary and sufficient condition for a primal optimal solution to be generated by the Lagrangian relaxation is obtained. The solution properties of Lagrangian relaxation problem are studied systematically. To overcome the difficulties caused by duality gap between the primal problem and the dual problem, we introduce an equivalent reformulation for the primal problem via applying a pth power to the constraints. We prove that this reformulation possesses an asymptotic strong duality property. Primal feasibility and primal optimality of the Lagrangian relaxation problems can be achieved in this reformulation when the parameter p is larger than a threshold value, thus ensuring the existence of an optimal primal-dual pair. We further show that duality gap for this partial pth power reformulation is a strictly decreasing function of p in the case of a single constraint.

Key words: asymptotic strong duality, integer programming, Lagrangian duality, pth power reformulation

1. Introduction

The general integer programming problem we address in this paper is of the following form:

$$(P)\quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \leqslant b_i, \;\; i=1,\ldots,m, \\ & x \in X, \end{array} \tag{1}$$

where f and g_i's are continuous functions and X is a finite integer set in $\mathbb{R}^n$.

*Dedicated to Professor Alex Rubinov on the occasion of his 65th birthday. Research supported by the Research Grants Council of Hong Kong under Grant CUHK 4214/01E, and the National Natural Science Foundation of China under Grants 79970107 and 10571116.

By associating with the ith constraint in (P) a nonnegative $\lambda_i, i = 1, 2, \ldots, m$, the Lagrangian function of (P) is defined as:

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i), \tag{2}$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T$. The Lagrangian relaxation of (P) is given by

$$(L_\lambda) \quad d(\lambda) = \min_{x \in X} L(x,\lambda). \tag{3}$$

Let

$$S = \{x \in X \mid g_i(x) \leqslant b_i, \ \ i = 1, \ldots, m\},$$
$$f^* = \min_{x \in S} f(x).$$

Since $d(\lambda) \leqslant f(x), \forall x \in S, \forall \lambda \geqslant 0$, weak duality relation always holds:

$$f^* \geqslant d(\lambda), \quad \forall \lambda \geqslant 0. \tag{4}$$

The Lagrangian dual problem of (P) is then to search for a multiplier vector $\lambda^* \geqslant 0$ which maximizes $d(\lambda)$ for all $\lambda \geqslant 0$:

$$(D) \quad d(\lambda^*) = \max_{\lambda \geqslant 0} d(\lambda). \tag{5}$$

Lagrangian methods for linear integer programming have been extensively studied in the literature (see e.g., Bell and Shapiro (1977), Fisher and Shapiro (1974), Fisher (1981), Geoffrion (1974) and Nemhauser and Wolsey (1988)). A survey of the use of Lagrangian techniques in integer programming can be found in Shapiro (1979). Lagrangian relaxation and decomposition methods have been also investigated in nonlinear integer programming (see e.g. Guignard and Kim (1987), Michelon and Maculan (1991, 1993)). In most situations, the Lagrangian dual problem (D) is unable to provide an optimal solution, or even a feasible solution, to the primal problem (P) due to the presence of a duality gap. Holmberg (1994), Williams (1996) and Wolsey (1981) studied the duality theory in linear integer programming. A geometric study of duality gaps in general integer programming was given in Lemaréchal and Renaud (2001). Dentcheva and Romisch (2004) and Sen et al. (2000) investigated duality gaps in integer stochastic programming. Duality theory was also discussed in the context of discrete convex analysis (Murota (1998)).

If an optimal solution x^* to (P) can be produced by solving (L_λ) with $\lambda=\lambda^*$, then we say that λ^* is an *optimal generating multiplier* (OGM) vector of (P) for x^* (Li and White (2000)). Let $g(x)=(g_1(x),\dots,g_m(x))^T$. Denote by E the map of set X under the mapping $x\mapsto(g(x),f(x))$, i.e.,

$$E=\{(y,z)=(g(x),f(x))\in\mathbb{R}^{m+1}\mid x\in X\}. \tag{6}$$

Geometrically, the existence of an OGM vector for x^* is equivalent to the existence of a supporting plane to set E at $(g(x^*),f(x^*))$. The existence of the supporting plane, however, could not be guaranteed in general in the Lagrangian formulation (1)–(5) as demonstrated by the following example.

EXAMPLE 1.

$$\begin{aligned}\min\ & f(x)=4+x_1x_2x_3x_4-x_1+3x_2+x_3-2x_4\\ \text{s.t. } & g_1(x)=x_1-2x_2+x_3+3\leqslant b,\\ & x\in X=\{0,1\}^4.\end{aligned}$$

Take $b=2.5$. Figure 1 illustrates set E of Example 1. We can see from Figure 1 that point $(2,4)$, the map of $(1,1,0,1)^T$, has the lowest value of $f(x)$ among the points in E located on the left of line $y=2.5$. Thus the optimal solution of this example is $x^*=(1,1,0,1)^T$ with $f(x^*)=4$. However, there does not exist an OGM vector for x^* as no line with slope $-\lambda(\lambda\geqslant 0)$ can support the set E at $(2,4)$.

Another fundamental question is whether an OGM vector is necessarily an optimal solution to the dual problem (D). It turns out that the answer is also negative in general as will be shown by Example 2 in Section 3.

If the dual optimal solution λ^* is an OGM vector for an optimal solution x^* to (P), then (x^*,λ^*) is said to be an *optimal primal-dual pair* of (P) (Li and White (2000)). It is clear that the existence of an optimal primal-dual pair provides a platform for the success of dual search. A well-known sufficient condition (see Shapiro (1979)) for strong duality (zero duality gap) can be stated as follows. If there exists a pair (x^*,λ^*) with $x^*\in X$ and $\lambda^*\geqslant 0$ such that

$$d(\lambda^*)=L(x^*,\lambda^*), \tag{7}$$

$$\lambda_i^*(g_i(x^*)-b_i)=0,\quad i=1,\dots,m, \tag{8}$$

$$g_i(x^*)\leqslant b_i,\quad i=1,\dots,m, \tag{9}$$

then (x^*,λ^*) is an optimal primal-dual pair of (P) and strong duality holds: $f(x^*)=d(\lambda^*)$. In contrast to its counterpart in continuous

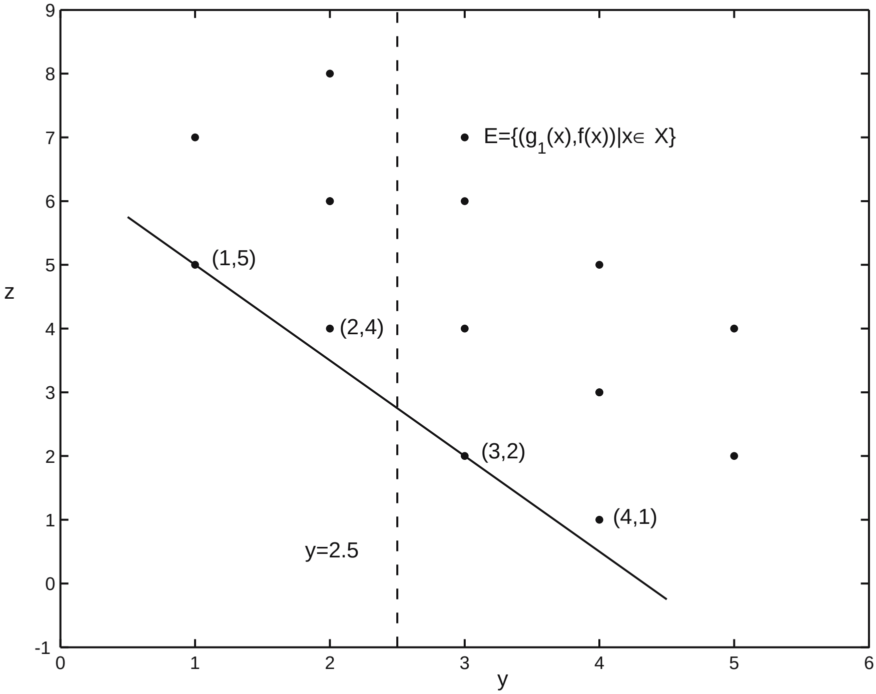

Figure 1. Illustration of set E in Example 1 with $b = 2.5$.

optimization, the complementary slack condition (8) rarely holds true for an optimal solution of integer programming since the constraint $g_i(x) \leqslant b_i$ is often inactive at x^* for index i with $\lambda_i^* > 0$.

The goal of this paper is twofold. First, we establish some new properties of Lagrangian duality theory for general integer programming by a means of perturbation analysis. Based on the relationship between duality and the perturbation function, a necessary and sufficient condition for the existence of an OGM vector of (P) is derived. Solution properties of the Lagrangian relaxation problem are investigated systematically. Second, we propose a new approach to ensure the existence of an optimal primal-dual pair via a partial pth power equivalent reformulation of the original problem. This reformulation is formed by applying a pth power transformation to the constraints. It is proved that the existence of an optimal primal-dual pair and an asymptotic strong duality can be assured in this partial pth power reformulation. The idea of using a pth power transformation was first introduced in Li (1995) to achieve a zero duality gap for nonconvex continuous optimization problems. The pth power Lagrangian was proposed for general nonlinear integer programming in Li and White (2000) and Li and Sun (2000). Although the existence of an OGM can be guaranteed in the pth power formulation of

Li and White (2000) and Li and Sun (2000), primal feasibility and existence of the primal-dual pair can be assured only for singly constrained cases resulted from applying a nonlinear surrogate constraint method in Li (1999). An alternative Lagrangian formulation of a logarithmic-exponential type for integer programming was investigated in Sun and Li (2000). Similarly, primal feasibility and the existence of the primal-dual pair can only be proved in the logarithmic-exponential-type nonlinear Lagrangian formulation for singly constrained cases. The most prominent contribution of this paper in nonlinear Lagrangian theory is an assurance of primal feasibility and the existence of the primal-dual pair in the proposed partial pth power Lagrangian formulation for general multiply constrained situations. The results obtained in this paper should better our understanding of the primal-dual methods for general integer programming.

The organization of the paper is as follows. In Section 2 we analyze the relationship between the Lagrangian duality and the perturbation function for integer programming. New results are derived to characterize the OGM vector and primal-dual pair. In Section 3 we establish new solution properties for the Lagrangian problem (L_{λ}^{*}), where λ^* is the dual optimum. In particular, primal feasibility and primal optimality of optimal solutions to (L_{λ^*}) are addressed. To achieve an asymptotic strong duality we introduce in Section 4 a partial pth power reformulation of (P). This partial pth power Lagrangian formulation ensures the existence of a primal-dual pair when p is sufficiently large. The decreasing monotonicity of duality gap in the partial pth power formulation is proved in Section 5 for the single-constraint case of (P). Finally, we give concluding remarks in Section 6.

2. Duality and Perturbation Function

In this section we establish some basic properties of the perturbation function of (P) and investigate the relationship between the perturbation function and the Lagrangian duality.

We make the following assumption on problem (P):

ASSUMPTION 1. *$S \neq \emptyset$ and there is at least one $x \in X \setminus S$ such that $f(x) < f^*$.*

Assumption 1 ensures that the problem (P) is feasible and cannot be trivially reduced to an unconstrained integer programming problem.

For any vectors x and $y \in \mathbb{R}^m, x \leqslant y$ iff $x_i \leqslant y_i, i = 1, \ldots, m$. A function $h(x)$ defined on $\mathbb{R}^m$ is said to be nonincreasing if for any x and $y \in \mathbb{R}^m, x \leqslant y$ implies $h(x) \geqslant h(y)$.

Let $b=(b_1,\dots,b_m)^T$. The perturbation function of (P) is defined as

$$w(y)=\min\{f(x)\mid g(x)\leqslant y,\ \ x\in X\}. \tag{10}$$

The domain of w is

$$Y=\{y\in\mathbb{R}^m\mid \text{there exists } x\in X \text{ such that } g(x)\leqslant y\}.$$

Note that Y is not always a convex set. The perturbation function w can be extended to the convex hull of Y by defining $w(y)=+\infty$ for $y\in \mathrm{conv}(Y)\setminus Y$.

By definition (10), $w(g(x))\leqslant f(x)$ for any $x\in X$ and $w(b)=f^*$. Furthermore, w is a nonincreasing and piecewise constant $(+\infty)$ function of y on $\mathrm{conv}(Y)$. In a process of increasing y, if there is a new point $\tilde{x}\in X$ such that $f(\tilde{x})<w(y)$ for any $y\in\{z\in Y\mid z\leqslant g(\tilde{x}),\ \ z\neq g(\tilde{x})\}$, the perturbation function w has a downward jump at $y=g(\tilde{x})$. The point $g(\tilde{x})$ corresponding to this new point $\tilde{x}$ is called a *corner point* of the perturbation function w in the y space. Since f and g_i's are continuous functions and X is a finite integer set, there are only finite number of corner points, say K corner points, $c_1,c_2,\dots,c_K$. Let $f_i=w(c_i), i=1,\dots,K$. Define the sets of corner points in y space and $\{y,w(y)\}$ space by

$$C=\{c_i=(c_{i1},c_{i2},\dots,c_{im})^T\mid i=1,\dots,K\},$$
$$\Phi_c=\{(c_i,f_i)\mid i=1,\dots,K\},$$

respectively. From the definition of the corner point, $(y,w(y))\in\Phi_c$ iff for any $z\in Y$ satisfying $z\leqslant y$ and $z\neq y$, it holds $w(z)>w(y)$.

By the definition of w, $\prod_{i=1}^m[y_i,+\infty)\subseteq Y$ if $y\in Y$. Let e^i denote the ith unit vector in $\mathbb{R}^m$. Then, e^i's are the extreme directions of $\mathrm{conv}(Y)$. Also, the set of extreme points of $\mathrm{conv}(Y)$ is a subset of C. Denote

$$\Lambda=\left\{\mu\in\mathbb{R}^K\,\middle|\,\sum_{i=1}^K\mu_i=1,\ \ \mu_i\geqslant 0,\ \ i=1,\dots,K\right\}.$$

The convex hull of Y can be expressed as

$$\begin{aligned}\mathrm{conv}(Y)&=\left\{\sum_{i=1}^K\mu_ic_i+\sum_{i=1}^m\alpha_ie^i\,\middle|\,\mu\in\Lambda,\ \ \alpha_i\geqslant 0,\ \ i=1,\dots,m\right\}\\ &=\{y\mid y\geqslant\sum_{i=1}^K\mu_ic_i,\ \ \mu\in\Lambda\}.\end{aligned}\tag{11}$$

From the definition of the corner point, the domain Y can be decomposed into K subsets with each c_i as the lower end of each subset Y_i. More specifically, we have $Y=\cup_{i=1}^{K} Y_i$ with $c_{ij}=\min\{y_j \mid y\in Y_i\}$, $j=1,\dots,m$, and w takes a constant f_i over Y_i:

$$w(y)=f_i, \quad \forall y\in Y_i, \ i=1,\dots,K. \tag{12}$$

Note that some Y_i may not be a single rectangular strip and there may exist different Y_i's on which $w(y)$ takes the same value. Define

$$\Phi=\{(y,w(y)) \mid y\in Y\}.$$

By the definition of Y_i, $c_i\in Y_i$ and $w(c_i)=f_i$ for each i. Thus $\Phi_c\subset\Phi$. Also, by the definition of E (cf. (6)), we have $\Phi_c\subset E$.

Consider an example of (P) with $f(x)=3-x_1-x_2-x_1x_2$, $g_1(x)=x_1$, $g_2(x)=x_2$, and $X=\{(0,0)^T,(0,1)^T,(1,1)^T,(0,2)^T,(2,0)^T,(2,2)^T\}$. By definition, we have

$$\begin{aligned}
&Y_1=[0,2)\times[0,1), \ c_1=(0,0)^T, \ f_1=3,\\
&Y_2=[0,1)\times[1,2), \ c_2=(0,1)^T, \ f_2=2,\\
&Y_3=[0,1)\times[2,+\infty), \ c_3=(0,2)^T, \ f_3=1,\\
&Y_4=[2,+\infty)\times[0,1), \ c_4=(2,0)^T, \ f_4=1,\\
&Y_5=[1,2)\times[1,+\infty)\cup[1,+\infty)\times[1,2), \ c_5=(1,1)^T, \ f_5=0,\\
&Y_6=[2,+\infty)\times[2,+\infty), \ c_6=(2,2)^T, \ f_6=-5.
\end{aligned}$$

We see that Y_5 is not a single rectangular strip and w takes the same value of 1 over Y_3 and Y_4. Figure 2 illustrates the perturbation function of this example.

A point $x\in X$ is said to be *noninferior* if there is no $\bar{x}\in X$ with $w(g(\bar{x}))=w(g(x))$ such that $g(\bar{x})\leqslant g(x)$ and $g(\bar{x})\neq g(x)$. The following lemma shows some useful properties of the perturbation function. Most importantly, the lemma proves that any noninferior optimal solution of (P) is corresponding to a corner point.

LEMMA 1. (i) *For any $y\in Y$, if x_y solves the perturbation problem*

$$w(y)=\min\{f(x) \mid g(x)\leqslant y, \ x\in X\},$$

then $(g(x_y), f(x_y))\in\Phi$.

(ii) *For any $c_i\in C$, there exists $\bar{x}\in X$ such that $(c_i,f_i)=(g(\bar{x}),f(\bar{x}))\in\Phi_c$.*

(iii) *For any noninferior optimal solution x^* to (P), $(g(x^*),f(x^*))\in\Phi_c$.*

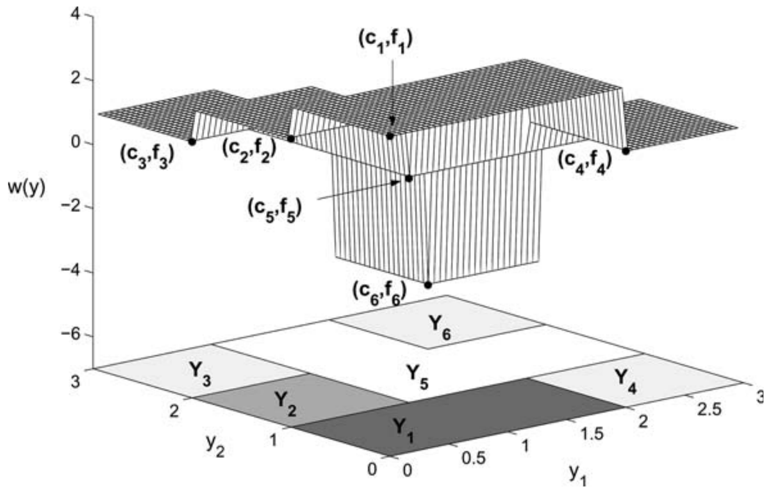

Figure 2. Illustration of perturbation function w and decomposition of Y.

(iv) *If $b \in Y_k$ for some $k \in \{1, \dots, K\}$, then $f^* = f_k$ and any optimal solution to the perturbation problem*

$$w(c_k) = \min\{f(x) \mid g(x) \leqslant c_k,\ \ x \in X\}$$

is a noninferior optimal solution to (P).

(v) *For any $\lambda \geqslant 0$, there exists $x_\lambda \in X$ that solves (L_λ) and satisfies $(g(x_\lambda), f(x_\lambda)) \in \Phi_c$.*

Proof. (i) Since $g(x_y) \leqslant y$ and w is a nonincreasing function, we have $f(x_y) = w(y) \leqslant w(g(x_y))$. On the other hand, since x_y is feasible in the perturbation problem

$$w(g(x_y)) = \min\{f(x) \mid g(x) \leqslant g(x_y),\ \ x \in X\},$$

we have $w(g(x_y)) \leqslant f(x_y)$. Thus, $w(g(x_y)) = f(x_y)$, i.e., $(g(x_y), f(x_y)) \in \Phi$.

(ii) Suppose that $\bar{x}$ solves the perturbation problem $w(c_i) = \min\{f(x) \mid g(x) \leqslant c_i,\ x \in X\}$, then $f(\bar{x}) = w(c_i) = f_i$ and $g(\bar{x}) \leqslant c_i$. By part (i), we have $(g(\bar{x}), f(\bar{x})) \in \Phi$. It then follows from the definition of c_i that $g(\bar{x}) = c_i$ and so $(g(\bar{x}), f(\bar{x})) = (c_i, f_i)$.

(iii) By part (i), we have $(g(x^*), f(x^*)) \in \Phi$. Let $z \in Y$ be such that $z \leqslant g(x^*)$ and $z \neq g(x^*)$. Suppose that $\bar{x}$ solves the perturbation problem

$$w(z) = \min\{f(x) \mid g(x) \leqslant z,\ \ x \in X\}.$$

Then $w(z) = f(\bar{x})$ and $g(\bar{x}) \leqslant z \leqslant g(x^*)$ with $g(\bar{x}) \neq g(x^*)$. Since x^* is a noninferior optimal solution, we must have $w(z) = f(\bar{x}) > f(x^*) = w(g(x^*))$. Thus $(g(x^*), f(x^*)) \in \Phi_c$.

(iv) Suppose that x^* solves the problem $w(c_k) = \min\{f(x) \mid g(x) \leqslant c_k,\ x \in X\}$. Then, by (12), $f^* = w(b) = w(c_k) = f_k = f(x^*)$. So x^* is an optimal solution to (P). If there exists another optimal solution $\bar{x}$ to (P) such that $g(\bar{x}) \leqslant g(x^*)$ and $g(\bar{x}) \neq g(x^*)$, then $g(\bar{x}) \leqslant g(x^*) \leqslant c_k \leqslant b$ and $g(\bar{x}) \neq c_k$. Since (c_k, f_k) is a corner point and $w(y)$ is a nonincreasing function, we have

$$f(\bar{x}) \geqslant w(g(\bar{x})) > w(c_k) = f_k = f(x^*),$$

which contradicts to the optimality of $\bar{x}$. Therefore, x^* is a noninferior optimal solution of (P).

(v) Let $\bar{x} \in X$ be an optimal solution to (L_λ). We claim that $f(\bar{x}) = w(g(\bar{x}))$. Otherwise, $f(\bar{x}) > w(g(\bar{x}))$. Let $\tilde{x} \in X$ solve $\min\{f(x) \mid g(x) \leqslant g(\bar{x}),\ x \in X\}$. Then $g(\tilde{x}) \leqslant g(\bar{x})$ and $f(\bar{x}) > w(g(\bar{x})) = f(\tilde{x})$. We have

$$L(\tilde{x}, \lambda) = f(\tilde{x}) + \lambda^T(g(\tilde{x}) - b) < f(\bar{x}) + \lambda^T(g(\bar{x}) - b) = L(\bar{x}, \lambda),$$

which contradicts the optimality of $\bar{x}$ to (L_λ). Now, let $g(\bar{x}) \in Y_k$ for some $k \in \{1, \ldots, K\}$. Then $c_k \leqslant g(\bar{x})$. By part (ii), there exists $x_\lambda \in X$ such that $(c_k, f_k) = (g(x_\lambda), f(x_\lambda))$. By (12), $f(x_\lambda) = f_k = w(c_k) = w(g(\bar{x})) = f(\bar{x})$. We have $L(x_\lambda, \lambda) \leqslant L(\bar{x}, \lambda)$ and hence x_λ is also an optimal solution to (L_λ) and $(g(x_\lambda), f(x_\lambda)) \in \Phi_c$. □

Let epi(w) denote the epigraph of w:

$$\text{epi}(w) = \{(y, z) \mid z \geqslant w(y),\ \ y \in \text{conv}(Y)\}. \tag{13}$$

Note that $f_i = w(c_i), i = 1, \ldots, K$. By (11) and (13), the convex hull of epi(w), conv(epi(w)), can be expressed as

$$\text{conv}(\text{epi}(w)) = \left\{(y, z) \mid (y, z) \geqslant \left(\sum_{i=1}^{K} \mu_i c_i, \sum_{i=1}^{K} \mu_i f_i\right), \mu \in \Lambda\right\}. \tag{14}$$

Define the *convex envelope function* of w on conv(Y):

$$\psi(y) = \min\{z \mid (y, z) \in \text{conv}(\text{epi}(w))\}. \tag{15}$$

By expression (14), (15) is equivalent to

$$\begin{aligned}\psi(y)=\min\ & \sum_{i=1}^{K}\mu_i f_i\\ \text{s.t.}\ & \sum_{i=1}^{K}\mu_i c_i\leqslant y,\\ & \sum_{i=1}^{K}\mu_i=1,\ \ \mu_i\geqslant 0,\quad i=1,\dots,K.\end{aligned}\tag{16}$$

The dual problem of (16) is

$$\begin{aligned}\psi(y)=\max\ & -\lambda^T y+r\\ \text{s.t.}\ & -\lambda^T c_i+r\leqslant f_i,\quad i=1,\dots,K,\\ & \lambda\geqslant 0,\ \ r\in\mathbb{R}.\end{aligned}\tag{17}$$

We see from (17) that ψ is a nonincreasing piecewise linear convex function on $\mathrm{conv}(Y)$. By definitions (13) and (15), it holds

$$w(y)\geqslant\psi(y),\quad y\in\mathrm{conv}(Y).\tag{18}$$

The dual expression of ψ in (17) also indicates that ψ is the greatest convex function majored by w.

THEOREM 1. *Let* μ^* *and* $(-\lambda^*, r^*)$ *be optimal solutions to* (16) *and* (17) *with* $y=b$*, respectively. Then*

(i) λ^* *is an optimal solution to the dual problem* (D) *and*

$$\psi(b)=\max_{\lambda\geqslant 0} d(\lambda)=d(\lambda^*).$$

(ii) *For each* i *with* $\mu_i^*>0$*, any* $\bar{x}\in X$ *satisfying* $(g(\bar{x}), f(\bar{x}))=(c_i, f_i)$ *is an optimal solution to the Lagrangian problem* (L_{λ^*}).

Proof. (i) For any $\lambda\geqslant 0$, by Lemma 1 (v), there exists $j\in\{1,\dots,K\}$ such that

$$\min_{x\in X} L(x,\lambda)=\min\{f_i+\lambda^T(c_i-b)\mid i=1,\dots,K\}=f_j+\lambda^T(c_j-b).$$

Let $r_\lambda = f_j + \lambda^T c_j$, then $f_i + \lambda^T c_i \geqslant r_\lambda$, $i = 1, \ldots, K$. Thus

$$\begin{aligned}\max_{\lambda\geqslant 0} d(\lambda) &= \max_{\lambda\geqslant 0}\min_{x\in X} L(x,\lambda)\\ &= \max_{\lambda\geqslant 0}(-\lambda^T b + r_\lambda)\\ &= \max_{\lambda\geqslant 0,\ r\in\mathbb{R}}\{-\lambda^T b + r \mid f_i + \lambda^T c_i \geqslant r,\ i = 1,\ldots,K\}.\\ &= \max_{\lambda\geqslant 0,\ r\in\mathbb{R}}\{-\lambda^T b + r \mid -\lambda^T c_i + r \leqslant f_i,\ i = 1,\ldots,K\}. \end{aligned} \tag{19}$$

On the other hand, by (17), we have

$$\begin{aligned}\psi(b) = \max\ & -\lambda^T b + r\\ \text{s.t. } & -\lambda^T c_i + r \leqslant f_i, \quad i = 1,\ldots,K,\\ & \lambda \geqslant 0,\ r\in\mathbb{R}.\end{aligned} \tag{20}$$

Combining (19) with (20) leads to

$$\psi(b) = \max_{\lambda\geqslant 0} d(\lambda) = d(\lambda^*).$$

Thus λ^* is a dual optimal solution.

(ii) By the complementary slackness condition of linear program (20), we have

$$\mu_i^*[(-\lambda^*)^T c_i + r^* - f_i] = 0, \quad i = 1,\ldots,K.$$

So for each $\mu_i^* > 0$, it holds $r^* = f_i + (\lambda^*)^T c_i$. Hence

$$d(\lambda^*) = \psi(b) = (-\lambda^*)^T b + r^* = f_i + (\lambda^*)^T (c_i - b). \tag{21}$$

By Lemma 1 there exists $\bar{x} \in X$ such that $(g(\bar{x}), f(\bar{x})) = (c_i, f_i)$. It then follows from (21) that $d(\lambda^*) = L(\bar{x}, \lambda^*)$, which means $\bar{x}$ is an optimal solution to (L_{λ^*}). □

We point out that part (i) of Theorem 1 was obtained before in Li and White (2000) and Lemaréchal and Renaud (2001), using different proofs. The following theorem characterizes a necessary and sufficient condition for the existence of an OGM vector of (P).

THEOREM 2. *Let x^* be an optimal solution to (P). Then there exists an OGM vector for x^* if and only if $w(g(x^*)) = \psi(g(x^*))$.*

Proof. Let $-\lambda^*\leqslant 0$ be a subgradient of ψ at $g(x^*)\in Y$. We have

$$\psi(y)\geqslant\psi(g(x^*))+(-\lambda^*)^T(y-g(x^*)),\quad\forall y\in Y. \tag{22}$$

For any $x\in X$, setting $y=g(x)\in Y$ in (22) and using (18), we get

$$f(x)\geqslant w(g(x))\geqslant\psi(g(x))\geqslant\psi(g(x^*))+(-\lambda^*)^T(g(x)-g(x^*)). \tag{23}$$

Since x^* is an optimal solution to (P), from Lemma 1 (i), we have $f(x^*)=w(g(x^*))$. If the condition $w(g(x^*))=\psi(g(x^*))$ holds, then we deduce from (23) that

$$f(x)+(\lambda^*)^T(g(x)-b)\geqslant f(x^*)+(\lambda^*)^T(g(x^*)-b),\quad\forall x\in X, \tag{24}$$

which means x^* is an optimal solution to (L_{λ^*}) and hence λ^* is an OGM vector for x^*.

Conversely, if there exists an OGM vector $\lambda^*\geqslant 0$ for x^*, then (24) holds. For any $y\in Y$, there exists $x\in X$ satisfying $f(x)=w(y)$ and $g(x)\leqslant y$. From (24), we have

$$\begin{aligned}w(y)&=f(x)\\&\geqslant f(x^*)-(\lambda^*)^T(g(x)-g(x^*))\\&\geqslant w(g(x^*))-(\lambda^*)^T(y-g(x^*))\end{aligned} \tag{25}$$

for all $y\in Y$. Recall that ψ is the greatest convex function majorized by w. We therefore deduce from (25) that

$$\psi(y)\geqslant w(g(x^*))-(\lambda^*)^T(y-g(x^*)),\quad\forall y\in Y.$$

Letting $y=g(x^*)$ in the above inequality yields $\psi(g(x^*))\geqslant w(g(x^*))$. Together with (18), this implies $w(g(x^*))=\psi(g(x^*))$. □

COROLLARY 1. *Let x^* be a noninferior optimal solution to (P). If all corner points are on the convex envelope function, i.e.,*

$$\psi(c_i)=f_i,\quad i=1,\dots,K, \tag{26}$$

then there exists an OGM vector for x^*.

Proof. From Lemma 1 (iii), $(g(x^*),f(x^*))\in\Phi_c$. By the assumption, $\psi(g(x^*))=f(x^*)=w(g(x^*))$. The conclusion then follows from Theorem 2. □

The following example, however, shows that condition (26) is not enough to guarantee the existence of an optimal primal-dual pair of (P).

EXAMPLE 2.

$$\begin{aligned}
\min\ & -3\sqrt{x_1}-2x_2\\
\text{s.t. } & x_1\leqslant 5,\\
& x_2\leqslant 5,\\
& x\in X=\{(1,4)^T,(2,2)^T,(5,7)^T,(8,8)^T,(9,7)^T\}.
\end{aligned}$$

The optimal solution of this problem is $x^*=(1,4)^T$ with $f(x^*)=-11$. The corner points are $(c_i,f_i), i=1,\dots,5$, with $c_1=(1,4)^T$, $f_1=-11$, $c_2=(2,2)^T$, $f_2=-8.2426$, $c_3=(5,7)^T$, $f_3=-20.7082$, $c_4=(8,8)^T$, $f_4=-24.4853$, $c_5=(9,7)^T$, $f_5=-23$. The optimal dual solution to (D) of this problem is $\lambda^*=(0.57287,2.14946)^T$ with $d(\lambda^*)=-16.4095$. There are three optimal solutions to the Lagrangian problem (L_{λ^*}): $(2,2)^T$, $(5,7)^T$ and $(9,7)^T$, among which only $(2,2)^T$ is feasible. However, $(2,2)^T$ with $f((2,2)^T)=-8.2426$ is not an optimal solution to the primal problem. Hence there is no optimal primal-dual pair in this problem. We can verify, however, condition (26) is satisfied and $\lambda=(1.01311,1.88524)^T$ is an OGM vector for $x^*=(1,4)^T$. This example also shows that an OGM vector is not necessarily an optimal solution to the dual problem (D).

In the following, we consider the existence of an optimal primal-dual pair in single-constraint cases of (P). Notice that the corner point set $\Phi_c=\{(c_i,f_i)\mid i=1,\dots,K\}$ now is a set in $\mathbb{R}^2$ and by the monotonicity of w we can assume without loss of generality that $c_1<c_2<\cdots<c_K$ and $f_1>f_2>\cdots>f_K$. The domain of w is $Y=[c_1,+\infty)$.

Define the envelope function of w in singly constrained cases as

$$\phi(y)=\begin{cases} f_1+\xi_1(y-c_1), & c_1\leqslant y<c_2\\ f_2+\xi_2(y-c_2), & c_2\leqslant y<c_3\\ \dots & \dots\\ f_{K-1}+\xi_{K-1}(y-c_{K-1}), & c_{K-1}\leqslant y<c_K\\ f_K, & c_K\leqslant y<\infty\end{cases} \tag{27}$$

where

$$\xi_i=\frac{f_{i+1}-f_i}{c_{i+1}-c_i}<0,\quad i=1,\dots,K-1. \tag{28}$$

It is clear that ϕ is a convex function if and only if $\xi_1\leqslant\xi_2\leqslant\cdots\leqslant\xi_{K-1}$. We have the following theorem.

THEOREM 3. *Suppose that* $m=1$ *and* ϕ *is convex on* $Y=[c_1,+\infty)$. *If* x^* *is a noninferior optimal solution to* (P), *then there exists* $\lambda^*\geqslant 0$ *such that* (x^*,λ^*) *is an optimal primal-dual pair of* (P).

Proof. By Assumption 1 and Lemma 1 (ii), there exists $k\in\{1,\dots,K-1\}$ satisfying $b\in[c_k,c_{k+1})$ and $(g(x^*),f(x^*))=(c_k,f_k)\in\Phi_c$. Let $\lambda^*=-\xi_k$. We first prove that x^* solves problem (L_{λ^*}). Since ξ_k is a subgradient of ϕ at $y=g(x^*)=c_k$, we have

$$\begin{aligned} w(y)&\geqslant\phi(y)\geqslant\phi(g(x^*))+\xi_k(y-g(x^*))\\ &=f(x^*)+\xi_k(y-g(x^*)),\quad \forall y\in Y. \end{aligned}\tag{29}$$

For any $x\in X$, let $y=g(x)$. It follows from (29) that

$$\begin{aligned} f(x)\geqslant w(g(x))=w(y)&\geqslant f(x^*)+\xi_k(y-g(x^*))\\ &=f(x^*)+\xi_k(g(x)-g(x^*)), \end{aligned}$$

which in turn yields

$$\begin{aligned} L(x,\lambda^*)&=f(x)+\lambda^*(g(x)-b)\\ &\geqslant f(x^*)+\lambda^*(g(x^*)-b)=L(x^*,\lambda^*). \end{aligned}\tag{30}$$

Thus x^* solves (L_{λ^*}). Next, we prove that λ^* solves the dual problem (D). For any fixed $\lambda\geqslant 0$, suppose that x_λ solves (L_λ). Thus, $L(x_\lambda,\lambda)\leqslant L(x,\lambda)$, for any $x\in X$. Then, for $(c_i,f_i), i=k,k+1$, we have

$$f_i\geqslant f(x_\lambda)-\lambda(c_i-g(x_\lambda)),\quad i=k,k+1.\tag{31}$$

Also, since $b\in[c_k,c_{k+1})$, there exists a $\mu\in(0,1]$ such that $b=\mu c_k+(1-\mu)c_{k+1}$. We thus obtain from (27), (28) and (31) that

$$\begin{aligned} d(\lambda^*)&=\min_{x\in X}\ L(x,\lambda^*)\\ &=f(x^*)-\xi_k(g(x^*)-b)\\ &=f_k-\frac{f_{k+1}-f_k}{c_{k+1}-c_k}[c_k-(\mu c_k+(1-\mu)c_{k+1})]\\ &=\mu f_k+(1-\mu)f_{k+1}\\ &\geqslant\mu[f(x_\lambda)-\lambda(c_k-g(x_\lambda))]+(1-\mu)[f(x_\lambda)-\lambda(c_{k+1}-g(x_\lambda))]\\ &=f(x_\lambda)+\lambda(g(x_\lambda)-b)\\ &=\min_{x\in X}L(x,\lambda)\\ &=d(\lambda). \end{aligned}$$

Hence λ^* solves (D). Therefore, (x^*,λ^*) is an optimal primal-dual pair of (P). □

3. Solution Properties of Lagrangian Problems

In this section, we focus on the solution properties of Lagrangian relaxation problem (L_{λ^*}):

$$d(\lambda^*) = \min_{x\in X} L(x, \lambda^*), \tag{32}$$

where λ^* is an optimal solution to the dual problem (D).

A key question arises from the problem (L_{λ^*}): Is there always an optimal solution to (L_{λ^*}) which is feasible in the primal problem? The answer is negative in general situations as shown in the following example.

EXAMPLE 3.

$$\begin{aligned} \min\ & f(x) = 3x_1 + 2x_2 - 1.5x_1^2 \\ \text{s.t.}\ & g_1(x) = \sqrt{15 - 7x_1 + 2x_2} \leqslant 2\sqrt{3}, \\ & g_2(x) = \sqrt{15 + 2x_1^2 - 7x_2} \leqslant 2\sqrt{3}, \\ & x \in X = \{(0,1)^T, (0,2)^T, (1,0)^T, (1,1)^T, (2,0)^T, (2,2)^T\}. \end{aligned}$$

The optimal solution of the problem is $x^* = (1,1)^T$ with $f(x^*) = 3.5$. The optimal solution to the dual problem (D) is $\lambda^* = (0, 1.0166)^T$ with $d(\lambda^*) = 1.3538$. The Lagrangian relaxation problem (L_λ) with $\lambda = \lambda^*$ has two optimal solutions: $(0,1)^T, (2,0)^T$, none of which is feasible. Notice that the problem has only two feasible solutions $(1,1)^T$ and $(2,2)^T$.

Nevertheless, we will show that the primal feasibility of problem (L_{λ^*}) can be always achieved in single-constraint cases.

THEOREM 4. *If $m = 1$, then there exists at least one optimal solution to the Lagrangian problem (L_{λ^*}) which is feasible in the primal problem.*

Proof. Suppose on the contrary there is no feasible optimal solution to (L_{λ^*}). Then

$$L(x, \lambda^*) > L(x^*, \lambda^*), \quad \forall x \in S, \tag{33}$$

where $x^* \in X \backslash S$ is an optimal solution to (L_{λ^*}) which is infeasible in the primal problem. Let

$$\bar{\lambda} = \min_{x\in S} \frac{f(x) - f(x^*)}{g(x^*) - g(x)}. \tag{34}$$

Then by (33), we have $\bar{\lambda} > \lambda^*$. Let $\bar{x} \in S$ be such that

$$L(\bar{x}, \bar{\lambda}) = \min_{x \in S} L(x, \bar{\lambda}). \tag{35}$$

Now for any $x \in X \backslash S$, since $g(x) - b > 0$, we have

$$\begin{aligned} L(x, \bar{\lambda}) &= f(x) + \bar{\lambda}(g(x) - b) > f(x) + \lambda^*(g(x) - b) \\ &= L(x, \lambda^*) \geqslant L(x^*, \lambda^*). \end{aligned} \tag{36}$$

On the other hand, for any $x \in S$, by (34) and (35), we have

$$\begin{aligned} L(x, \bar{\lambda}) &\geqslant L(\bar{x}, \bar{\lambda}) \\ &= f(\bar{x}) + \bar{\lambda}(g(\bar{x}) - b) \\ &= f(\bar{x}) + \bar{\lambda}(g(\bar{x}) - g(x^*)) + \bar{\lambda}(g(x^*) - b) \\ &\geqslant f(\bar{x}) + \frac{f(\bar{x}) - f(x^*)}{g(x^*) - g(\bar{x})}(g(\bar{x}) - g(x^*)) + \bar{\lambda}(g(x^*) - b) \\ &> f(x^*) + \lambda^*(g(x^*) - b) \\ &= L(x^*, \lambda^*). \end{aligned} \tag{37}$$

Combining (36) with (37), we infer that

$$d(\bar{\lambda}) = \min_{x \in X} L(x, \bar{\lambda}) > L(x^*, \lambda^*) = d(\lambda^*),$$

which contradicts the optimality of λ^*. □

Interestingly, the following theorem and corollary reveal that the primal infeasibility is assured for at least one optimal solution to (L_{λ^*}) in general situations, including both singly constrained and multiply constrained cases, where there exists a nonzero duality gap.

THEOREM 5. *Assume that $\psi(y) < w(y)$ for some $y \in conv(Y)$. Let μ^* be an optimal solution to* (16). *Then there is at least an $i \in \{1, \dots, K\}$ such that $\mu_i^* > 0$ and $c_i \in C$ with $c_i \nleqslant y$.*

Proof. For a $y \in Y$, by (16), there exists $\mu^* \in \Lambda$ that solves the following problem:

$$\begin{aligned} &\psi(y) = \min \textstyle\sum_{i=1}^K \mu_i f_i, \\ &\text{s.t. } \textstyle\sum_{i=1}^K \mu_i c_i \leqslant y, \ \ \mu \in \Lambda. \end{aligned} \tag{38}$$

Let $I = \{i \mid \mu_i^* > 0\}$. Suppose that $c_i \leqslant y$ for all $i \in I$. We claim that $f_k = f_l$ for any $k, l \in I$. Otherwise, suppose that $f_k > f_l$ for some $k, l \in I$. Define

$\tilde{\mu}=(\tilde{\mu}_1,\dots,\tilde{\mu}_K)$ as follows: $\tilde{\mu}_i=\mu_i^*$, if $i\neq k$ and $i\neq l$; and $\tilde{\mu}_k=\mu_k^*-\epsilon$, $\tilde{\mu}_l=\mu_l^*+\epsilon$, with $\epsilon>0$ being small enough that $\tilde{\mu}_k>0$.

Note that $\tilde{\mu}\in\Lambda$ and $\tilde{\mu}_j=0$ iff $\mu_j^*=0$. Since by assumption $c_i\leqslant y$ for all $i\in I$, it follows that $\sum_{i=1}^K\tilde{\mu}_ic_i=\sum_{i\in I}\tilde{\mu}_ic_i\leqslant\sum_{i\in I}\tilde{\mu}_iy=y$. Thus, $\tilde{\mu}$ is feasible to problem (38). Moreover,

$$\sum_{i=1}^K(\tilde{\mu}_if_i-\mu_i^*f_i)=\epsilon(f_l-f_k)<0,$$

which contradicts that μ^* is an optimal solution to (38). Therefore, $f_k=f_l$ for any $k,l\in I$. It then follows that $\psi(y)=f_i$ for any $i\in I$. Since $c_i\leqslant y$ for all $i\in I$, $w(c_i)\geqslant w(y)$. Thus, $\psi(y)=f_i=w(c_i)\geqslant w(y)$, contradicting the assumption that $\psi(y)<w(y)$. □

COROLLARY 2. *Assume that the duality gap between (P) and (D) is nonzero, i.e., $d(\lambda^*)<f^*$. Then there is at least one optimal solution to the Lagrangian problem (L_{λ^*}) which is infeasible in the primal problem.*

Proof. Notice from Theorem 1 (i) that $\psi(b)=d(\lambda^*)$. Thus, $\psi(b)<f^*=w(b)$. Applying Theorem 5 with $y=b$, we conclude that there exists an $i\in I$ such that $c_i\nleqslant b$. Let $\bar{x}$ be such that $(g(\bar{x}),f(\bar{x}))=(c_i,f_i)$. Then $\bar{x}$ is infeasible and by Theorem 1 (ii), $\bar{x}$ solves (L_{λ^*}). □

4. *p*th Power Reformulation

In this section we introduce a new partial pth power reformulation of (P). Based on the results obtained in the previous sections, we show that this partial pth power reformulation possesses an asymptotic strong duality property and can guarantee the existence of an optimal primal-dual pair, thus providing a platform for the success of dual search for the reformulated problem. Without loss of generality, we can assume the following.

ASSUMPTION 2. *The functions g_i's are strictly positive over X and $b_i>0$ for all i.*

Notice that Assumption 2 can be always satisfied via some suitable equivalent transformations for (P).

For any $y\in\mathbb{R}_+^m$ and $p>0$, denote by y^p the vector $(y_1^p,\dots,y_m^p)^T$. Consider the following equivalent reformulation of (P):

$$\begin{aligned}&\min f(x)\\&\text{s.t. } [g(x)]^p\leqslant b^p,\quad x\in X.\end{aligned}\tag{39}$$

The Lagrangian relaxation of (39) is

$$d_p(\lambda) = \min_{x \in X} L_p(x, \lambda), \tag{40}$$

where $L_p(x, \lambda) = f(x) + \lambda^T \{[g(x)]^p - b^p\}$. The dual problem of (39) is

$$\theta(p) = \max_{\lambda \geqslant 0} \ d_p(\lambda). \tag{41}$$

The pth power dual formulation was introduced previously in Li and White (2000) and Li and Sun (2000) where pth power is imposed on both objective and constraint functions. Formulation (39) can be viewed as a partial pth power reformulation. In the following, the results obtained in the previous sections will be applied to (39)–(41). Denote by w_p the perturbation function of (39):

$$w_p(y) = \min\{f(x) \mid [g(x)]^p \leqslant y, \ \ x \in X\}.$$

It is clear that the domain of w_p is $Y_p = \{y^p \mid y \in Y\}$ and $w_p(y^p) = w(y)$ for any $y \in Y$. Moreover, the corner point set of w_p is $\{(c_i^p, f_i) \mid i = 1, \ldots, K\}$.

Let ψ_p denote the convex envelope function of w_p. Then, by (16), $\psi_p(y)$ can be expressed as

$$\begin{aligned} \psi_p(y) = \min \ \ & \textstyle\sum_{i=1}^K \mu_i(p) f_i \\ \text{s.t.} \ \ & \textstyle\sum_{i=1}^K \mu_i(p) c_i^p \leqslant y, \ \ \mu(p) \in \Lambda. \end{aligned} \tag{42}$$

Let $v \in Y$ and $\mu(p) \in \Lambda$ for any $p > 0$. Define the following index sets:

$$I(p) = \{i \in \{1, \ldots, K\} \mid \mu_i(p) > 0\}, \tag{43}$$

$$I_1(p) = \{i \in \{1, \ldots, K\} \mid \mu_i(p) > 0, \ \ c_i \leqslant v\}, \tag{44}$$

$$I_2(p) = I(p) \backslash I_1(p). \tag{45}$$

LEMMA 2. *Let $v \in Y$ and $p \to \infty$. For each $p > 0$, let $\mu(p) \in \Lambda$ be such that $\sum_{i=1}^K \mu_i(p) c_i^p \leqslant v^p$. Then*

$$\lim_{p \to \infty} \sum_{i \in I_2(p)} \mu_i(p) = 0, \tag{46}$$

$$\lim_{p \to \infty} \sum_{i \in I_1(p)} \mu_i(p) = 1. \tag{47}$$

Proof. Suppose on the contrary that (46) does not hold. Then there exist a subsequence $\{p_k\}$ with $p_k \to \infty$ and a constant $\epsilon > 0$ such that $\sum_{i \in I_2(p_k)} \mu_i(p_k) \geqslant \epsilon$ for all k. Since $|I_2(p_k)| \leqslant K$, there must exist an $s \in I_2(p_k)$ and $\mathcal{K} \subseteq \{1, 2, \dots\}$ such that for each $k \in \mathcal{K}, \mu_s(p_k) \geqslant \epsilon / K$ holds. Moreover, since $s \in I_2(p_k)$, we have from (45) that $c_s \nleqslant v$ and there exists $t \in \{1, 2, \dots, m\}$ such that $c_{st} > v_t$. Thus, by assumption, we have

$$v_t^{p_k} \geqslant \sum_{i \in I(p_k)} \mu_i(p_k) c_{it}^{p_k} \geqslant \sum_{i \in I_2(p_k)} \mu_i(p_k) c_{it}^{p_k} \geqslant \mu_s(p_k) c_{st}^{p_k} \geqslant (\epsilon/K) c_{st}^{p_k}, \quad \forall k \in \mathcal{K}. \tag{48}$$

Let

$$\tau = \max \left\{ \frac{v_j}{c_{ij}} \mid c_{ij} > v_j, \;\; j \in \{1, \dots, m\}, \;\; i \in \{1, \dots, K\} \right\}.$$

It is clear that $\tau < 1$. Then by (48),

$$\tau^{p_k} \geqslant \left(\frac{v_t}{c_{st}} \right)^{p_k} \geqslant \epsilon / K > 0, \quad \forall k \in \mathcal{K}.$$

This contradicts $\tau < 1$ and $p_k \to \infty$. Thus (46) holds. Since $\mu(p) \in \Lambda$, (46) implies (47). □

THEOREM 6. *There exists $p_0 > 1$ such that*

$$\psi_p(c_i^p) = f_i, \quad i = 1, \dots, K, \tag{49}$$

when $p \geqslant p_0$.

Proof. From (18), we have

$$f_i = w_p(c_i^p) \geqslant \psi_p(c_i^p), \quad \forall p > 0. \tag{50}$$

We prove the theorem by contradiction. Suppose that the conclusion of the theorem does not hold. Then there exists $l \in \{1, \dots, K\}$ and a sequence $\{p_k\}$ with $p_k \to \infty$ such that

$$f_l > \psi_{p_k}(c_l^{p_k}), \quad \forall k. \tag{51}$$

Let $\mu(p_k)$ be an optimal solution to (42) with $y=c_l^{p_k}$. Then $\mu(p_k)\in\Lambda$ and

$$\psi_{p_k}(c_l^{p_k})=\sum_{i=1}^{K}\mu_i(p_k)f_i, \tag{52}$$

$$\sum_{i=1}^{K}\mu_i(p_k)c_i^{p_k}\leqslant c_l^{p_k}. \tag{53}$$

We claim that $\mu_l(p_k)=0$ for any k, i.e., $l\notin I(p_k)$ for any k, where $I(p_k)$ is defined by (43). We note first that $\mu_l(p_k)\neq 1$, since by (52) $\mu_l(p_k)=1$ implies $\psi_{p_k}(c_l^{p_k})=f_l$, contradicting (51). If $0<\mu_l(p_k)<1$, then we can rewrite (53) as $\sum_{i=1}^{K}\hat{\mu}_i(p_k)c_i^{p_k}\leqslant c_l^{p_k}$, where $\hat{\mu}_i(p_k)=\mu_i(p_k)/(1-\mu_l(p_k))$ for $i\neq l$ and $\hat{\mu}_l(p_k)=0$. Thus, $\hat{\mu}(p_k)\in\Lambda$ and $\hat{\mu}(p_k)$ is feasible to (42) with $y=c_l^{p_k}$. Moreover, we have

$$\begin{aligned}\sum_{i=1}^{K}\hat{\mu}_i(p_k)f_i&=\frac{\psi_{p_k}(c_l^{p_k})-\mu_l(p_k)f_l}{1-\mu_l(p_k)}\\&=\psi_{p_k}(c_l^{p_k})+\frac{\mu_l(p_k)}{1-\mu_l(p_k)}(\psi_{p_k}(c_l^{p_k})-f_l).\end{aligned} \tag{54}$$

Since, by (51), $f_l>\psi_{p_k}(c_l^{p_k})$, (54) implies $\sum_{i=1}^{K}\hat{\mu}_i(p_k)f_i<\psi_{p_k}(c_l^{p_k})$, contradicting the optimality of $\mu(p_k)$.

Now, let $I_1(p_k)$ and $I_2(p_k)$ be defined in (44) and (45) with $v=c_l$, respectively. Let

$$\delta=\min\{f_i-f_l\mid c_i\leqslant c_l,\ i\neq l,\ i\in\{1,\dots,K\}\}.$$

It follows from the definition of a corner point that $c_i\leqslant c_l$ and $c_i\neq c_l$ implies $f_i>f_l$, and hence $\delta>0$. Since $l\notin I_1(p_k)$ for all k, we have

$$f_i\geqslant f_l+\delta,\quad \forall i\in I_1(p_k),\ \forall k.$$

Since $w_p(c_l^{p_k})=f_l>\psi_{p_k}(c_l^{p_k})$, we have $I_2(p_k)\neq\emptyset$ for all k by Theorem 5. Applying Lemma 2 with $v=c_l$, we conclude that there exists k_0 such that when $k\geqslant k_0$, $I_1(p_k)\neq\emptyset$ holds and

$$\sum_{i\in I_1(p_k)}\mu_i(p_k)f_i\geqslant\sum_{i\in I_1(p_k)}\mu_i(p_k)(f_l+\delta)\geqslant f_l+\frac{1}{2}\delta, \tag{55}$$

$$\sum_{i\in I_2(p_k)}\mu_i(p_k)f_i\geqslant-\frac{1}{4}\delta. \tag{56}$$

Combining (52) with (55) and (56) yields

$$\begin{aligned}\psi_{p_k}(c_l^{p_k}) &= \sum_{i\in I_1(p_k)} \mu_i(p_k) f_i + \sum_{i\in I_2(p_k)} \mu_i(p_k) f_i \\ &\geqslant f_l + \frac{1}{2}\delta - \frac{1}{4}\delta \\ &= f_l + \frac{1}{4}\delta \end{aligned} \tag{57}$$

for $k \geqslant k_0$. Inequality (57) contradicts (51). The proof is completed. □

Let $\lambda(p)$ be a dual optimal solution to (41). Denote the Lagrangian problem with $\lambda = \lambda(p)$ as:

$$d_p(\lambda(p)) = \min_{x\in X} L_p(x, \lambda(p)). \tag{58}$$

Let us consider again Example 1 with $b = 2.5$ in Section 1 as an illustration of Theorem 6. Figure 3 depicts the functions w_p and ψ_p for $p = 3$. We can see from Figure 3 that condition (49) is satisfied when $p = 3$. It can be verified that $\lambda(3) = -(4-2)/(8-27) = 2/19$ is an optimal solution to the dual problem (41) in this example and $x^* = (1, 1, 0, 1)^T$ can be generated by the Lagrangian problem (58) when $p = 3$.

The following theorem further shows that the primal feasibility of (58) and the existence of an optimal primal-dual pair of (39) can be also ensured when p is larger than a threshold value. Moreover, the partial pth power reformulation possesses an asymptotic strong duality.

THEOREM 7. (i) *There exists $p_1 \geqslant 1$ such that there exists at least an optimal solution to* (58) *that is feasible to* (P) *when $p \geqslant p_1$.*

(ii) $\lim_{p\to\infty} \theta(p) = f^*$.

(iii) *For any noninferior optimal solution x^* of (P), there is $p_2 \geqslant p_1$ such that there exists one optimal primal-dual pair $(x^*, \lambda(p))$ of (39) when $p \geqslant p_2$.*

Proof. We first notice from Theorem 1 (i) that $\theta(p) = d_p(\lambda(p)) = \psi_p(b^p)$. Let $\mu(p) \in \Lambda$ be an optimal solution to (42) with $y = b^p$. Then

$$\psi_p(b^p) = \sum_{i=1}^{K} \mu_i(p) f_i, \tag{59}$$

$$\sum_{i=1}^{K} \mu_i(p) c_i^p \leqslant b^p. \tag{60}$$

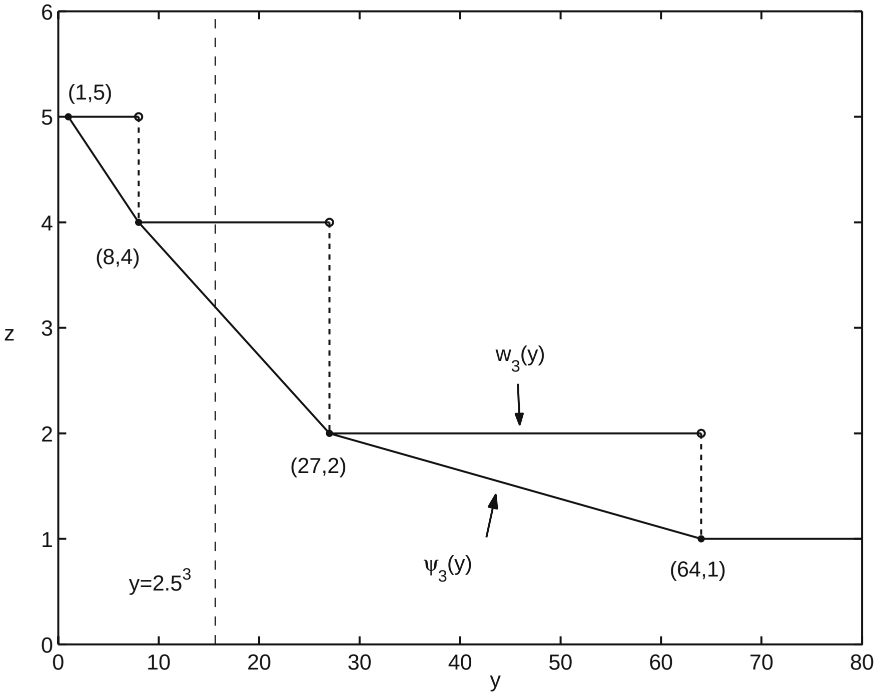

Figure 3. Illustration of $w_3(y)$ and $\psi_3(y)$ for Example 1 with $b = 2.5$.

Let $I(p)$ be defined in (43) and $I_1(p)$ and $I_2(p)$ be defined in (44) and (45) with $v = b$, respectively. By Lemma 2, we have

$$\lim_{p\to\infty} \sum_{i\in I_2(p)} \mu_i(p) = 0, \tag{61}$$

$$\lim_{p\to\infty} \sum_{i\in I_1(p)} \mu_i(p) = 1. \tag{62}$$

(i) Note that if $I_1(p) \neq \emptyset$, then for any $i \in I_1(p)$, by Lemma 1 (ii), there is $\bar{x} \in X$ satisfying $g(\bar{x}) = c_i \leqslant b$. Moreover, by Theorem 1 (ii), $\bar{x}$ is an optimal solution to (58). Thus, it suffices to show that there exists $p_1 > 0$ such that $I_1(p) \neq \emptyset$ when $p \geqslant p_1$. Since $I_2(p) = \emptyset$ implies $I_1(p) = I(p) \neq \emptyset$, we assume $I_2(p) \neq \emptyset$. It follows from (62) that $I_1(p) \neq \emptyset$ for sufficiently large p.

(ii) By part (i), $I_1(p) \neq \emptyset$ for $p \geqslant p_1$. For any $i \in I_1(p)$, $f_i = w(c_i) \geqslant w(b) = f^*$ by the monotonicity of the perturbation function w. We obtain from (59) and (62) that

$$\begin{aligned}\psi_p(b^p) &= \sum_{i\in I_1(p)} \mu_i(p) f_i + \sum_{i\in I_2(p)} \mu_i(p) f_i \\ &\geqslant \sum_{i\in I_1(p)} \mu_i(p) f^* \to f^*, \qquad p\to\infty.\end{aligned} \tag{63}$$

On the other hand, the weak duality relation (4) and Theorem 1 (i) give

$$f^* \geqslant d_p(\lambda(p)) = \psi_p(b^p). \tag{64}$$

Combining (63) with (64) yields part (ii).

(iii) Notice first that if $f_i = f^*$ for some $i \in I_1(p)$, then there exists $x^* \in S$ such that $(g(x^*), f(x^*)) = (c_i, f_i) = (c_i, f^*)$. Hence x^* is an optimal solution to (39). By Theorem 1 (ii), x^* solves problem (58) and thus $(x^*, \lambda(p))$ is an optimal primal-dual pair of (39). We now prove that there exists $i \in I_1(p)$ satisfying $f_i = f^*$ when p ($\geqslant p_1$) is sufficiently large. Suppose on the contrary there exists a sequence $\{p_k\}$ with $p_k \to \infty$ and for each k, $f_i > f^*$ for all $i \in I_1(p_k)$. Let

$$\delta^* = \min\{f_i - f^* \mid c_i \leqslant b, \;\; f_i \neq f^*, \;\; i \in \{1, \ldots, K\}\} > 0.$$

Using the similar arguments as in the proof of (57), we can deduce from (61) and (62) that

$$\theta(p_k) = \psi_{p^k}(b^{p_k}) > f^* + \frac{1}{4}\delta^*$$

when k is sufficiently large. This, however, contradicts part (ii). □

Next, we study the relationship among the parameters p_0, p_1 and p_2, which are defined in Theorems 6 and 7, respectively. By the definition of the optimal primal-dual pair, it always holds $p_1 \leqslant p_2$. When $m = 1$, we also know from Theorems 3 and 4 that $p_1 = 1$ and $p_2 \leqslant p_0$. Thus, for singly constrained problems, we have

$$1 = p_1 \leqslant p_2 \leqslant p_0. \tag{65}$$

The strict inequality $p_2 < p_0$ in (65) could hold when condition (49) is satisfied for c_i around $y = b$ and thus there exists an optimal primal-dual pair, while condition (49) is not satisfied for c_i far away from $y = b$. Consider Example 1 with $b = 3.5$. The perturbation function $w(y)$ and the convex envelope function $\psi(y)$ of this problem are illustrated in Figure 4. It can be verified that the optimal solution of this problem is $x^* = (0, 0, 0, 1)^T$ which corresponds to point $(3, 2)^T$ in Figure 4. Also, $\lambda^* = 1$ is the optimal solution to (D) and (x^*, λ^*) is an optimal primal-dual pair. The corner points are $(c_i, f_i), i = 1, \ldots, 4$, with $c_1 = 1, f_1 = 5, c_2 = 2, f_2 = 4, c_3 = 3, f_3 = 2, c_4 = 4, f_4 = 1$. Yet $\psi(c_2) = 3.5 < 4 = f_2$. Hence $1 = p_1 = p_2 < p_0$. It is noticed from Figure 3 that $p_0 \leqslant 3$.

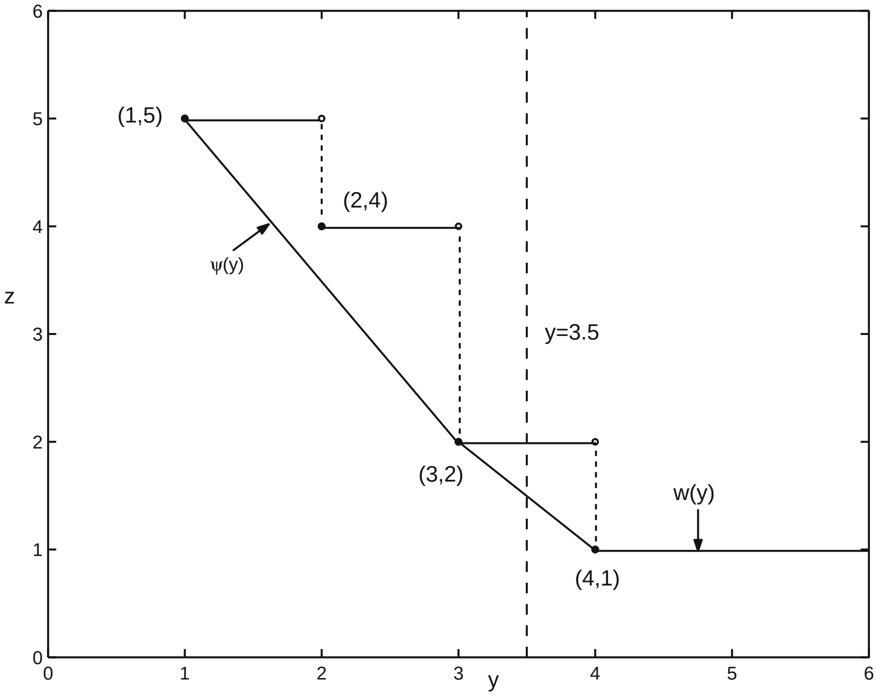

Figure 4. Illustration of $w(y)$ and $\psi(y)$ for Example 1 with $b=3.5$.

For multiply constrained cases, either of the following two cases may happen:

$$1 \leqslant p_1 \leqslant p_2 \leqslant p_0, \tag{66}$$

$$1 \leqslant p_0 \leqslant p_1 \leqslant p_2. \tag{67}$$

Let us consider Example 3 in Section 3 again. The corner points of Example 3 are (c_i, f_i), $i = 1, \ldots, 6$, with $c_1 = (4.1231, 2.8284)^T$, $f_1 = 2$, $c_2 = (4.3589, 1)^T$, $f_2 = 4$, $c_3 = (2.8284, 4.1231)^T$, $f_3 = 1.5$, $c_4 = (3.1623, 3.1623)^T$, $f_4 = 3.5$, $c_5 = (1, 4.7958)^T$, $f_5 = 0$, $c_6 = (2.2361, 3)^T$, $f_6 = 4$. It can be verified that

$$\psi(c_1) = 2 = f_1, \quad \psi(c_2) = 4 = f_2, \quad \psi(c_3) = 0.6839 < 1.5 = f_3,$$
$$\psi(c_4) = 1.6834 < 3.5 = f_4, \quad \psi(c_5) = 0 = f_5, \quad \psi(c_6) = 4 = f_6.$$

Note that (c_4, f_4) corresponds to the optimal solution $x^* = (1, 1)^T$.

Applying the partial pth power reformulation to Example 3, it can be verified that the primal feasibility of the pth power Lagrangian relaxation problem (58) can be achieved when $p \geqslant 2$. However, this is not enough to guarantee the existence of the optimal primal-dual pair. For instance, take $p = 2$, we have $\lambda(2) = (0.1951, 0.3414)^T$ and the optimal solutions to (58) are $(0,1)^T$, $(2,0)^T$, $(0,2)^T$ and $(2,2)^T$. Thus, $(x^*, \lambda(2))$ is not an optimal primal-dual pair. We can verify that for $p = 2$,

$$\psi_p(c_1^p)=2=f_1,\ \ \psi_p(c_2^p)=4=f_2,\ \ \psi_p(c_3^p)=0.8<1.5=f_3,$$
$$\psi_p(c_4^p)=2.6829<3.5=f_4,\ \ \psi_p(c_5^p)=0=f_5,\ \ \psi_p(c_6^p)=4=f_6.$$

So, condition (49) is not satisfied. Since $\psi_p(c_4^p) < f_4$ and (c_4^p, f_4) corresponds to the optimal solution x^*, there is no OGM vector for x^* when $p=2$. We can further increase the value of p. When $p \geqslant 6.3$, condition (49) is satisfied and $(x^*, \lambda(p))$ becomes an optimal primal-dual pair. For instance, take $p=6.3$, we have $\lambda(6.3)=(0.2874\times 10^{-3}, 0.3609\times 10^{-3})^T$ and the optimal solution to (58) are $(0,1)^T$, $(1,0)^T$, $(1,1)^T$ and $(2,2)^T$. Therefore, we have $1<p_1<p_2=p_0$ and hence (66) holds in this example.

To show that (67) may happen, let us consider Example 2 in Section 2. Although condition (26) (or (49) with $p_0=1$) is satisfied for this example, there does not exist an optimal primal-dual pair in the original problem setting. Applying the partial pth power reformulation to Example 2 with $p=3$, we make the generation of an optimal primal-dual pair with $\lambda(3)=(0.0038, 0.0331)^T$ and $x^*=(1,4)^T$. Thus, $1=p_0=p_1<p_2$.

5. Monotonicity of Duality Gap

Let $\kappa(p)$ denote duality gap between the problem (39) (or (1)) and the dual problem (41): $\kappa(p)=f^*-\theta(p)$. The main result of this section is to show that $\kappa(p)$ is a strictly decreasing function of p ($p>0$) for single-constraint cases of (P).

Let $m=1$. Assume that points in C, the set of corner point of the perturbation function $w(y)$, are in an increasing order: $0<c_1<c_2<\cdots<c_K$. The convex envelope function ψ_p can be expressed as

$$\psi_p(y)=\min\{f_i+\xi_{ij}(p)(y^p-c_i^p)\mid c_i\leqslant y<c_j,\ \ 1\leqslant i<j\leqslant K\}, \tag{68}$$

where

$$\xi_{ij}(p)=\frac{f_j-f_i}{c_j^p-c_i^p}<0,\quad 1\leqslant i<j\leqslant K. \tag{69}$$

We need the following lemma.

LEMMA 3. *Let*

$$\varphi(p)=\frac{\gamma^p-\delta^p}{\eta^p-\delta^p},\quad p>0,\quad 0<\delta\leqslant\gamma<\eta. \tag{70}$$

Then φ is a strictly decreasing function on $(0,+\infty)$.

Proof. Let $u=\gamma/\delta$ and $v=\eta/\delta$. By the assumption, we have $1\leqslant u<v$. The function φ can be rewritten as

$$\varphi(p)=\frac{u^p-1}{v^p-1}.$$

For any $p>0$, we have

$$\begin{aligned}\varphi'(p)&=\frac{(v^p-1)u^p\ln(u)-(u^p-1)v^p\ln(v)}{(v^p-1)^2}\\&=\frac{u^pv^p(\ln(u)-\ln(v))+v^p\ln(v)-u^p\ln(u)}{(v^p-1)^2}\\&=\frac{u^pv^p(\ln(u^p)-\ln(v^p))+v^p\ln(v^p)-u^p\ln(u^p)}{p(v^p-1)^2}.\end{aligned}\tag{71}$$

Let $\alpha=u^p$ and $\beta=v^p$. Then $1\leqslant\alpha<\beta$. By (71), $\varphi'(p)<0$ if and only if

$$\alpha\beta(\ln(\alpha)-\ln(\beta))+\beta\ln(\beta)-\alpha\ln(\alpha)<0,$$

which is in turn equivalent to

$$\frac{\alpha}{\alpha-1}\ln(\alpha)<\frac{\beta}{\beta-1}\ln(\beta).\tag{72}$$

Now, consider the function

$$s(t)=\frac{t}{t-1}\ln(t),\quad t>1.$$

We have

$$s'(t)=\frac{(t-1-\ln(t))}{(t-1)^2}.\tag{73}$$

Note that $\ln(t)<t-1$ for all $t>1$. It follows from (73) that $s'(t)>0$ for all $t>1$ and hence s is a strictly increasing function on $(1,+\infty)$. Therefore (72) holds. □

THEOREM 8. *Duality gap $\kappa(p)$ is a strictly decreasing function of p for $p>0$.*

Proof. By the definition, it suffices to show that $\psi_p(b^p)$ is a strictly increasing function of p. Let $1\leqslant k\leqslant K$ be such that $c_k\leqslant b<c_{k+1}$. From (68), we have

$$\psi_p(b^p) = \min\{f_i + \xi_{ij}(p)(b^p - c_i^p) \mid 1 \leqslant i \leqslant k,\ k+1 \leqslant j \leqslant K\}$$
$$= \min\left\{ f_i + \frac{f_j - f_i}{c_j^p - c_i^p}(b^p - c_i^p) \mid 1 \leqslant i \leqslant k,\ k+1 \leqslant j \leqslant K \right\}. \quad (74)$$

Since $c_i \leqslant c_k \leqslant b < c_{k+1} \leqslant c_j$, by Lemma 3, $(b^p - c_i^p)/(c_j^p - c_i^p)$ is a strictly decreasing function on $(0, +\infty)$. Moreover, $f_j - f_i < 0$ for $i < j$. Thus, we deduce from (74) that $\psi_p(b^p)$ is a strictly increasing function of p. □

6. Concluding Remarks

We have presented in this paper new results on the Lagrangian duality theory for general integer programming problems. Fresh insights into some fundamental properties of Lagrangian duality have been obtained by virtue of perturbation analysis for integer programming. In particular, we have derived a necessary and sufficient condition for the existence of an OGM vector. New solution properties of Lagrangian relaxation problems have been identified. To ensure the existence of an optimal primal-dual pair, we have proposed a partial pth power reformulation of the primal problem. This reformulation ensures primal feasibility of the optimal solution produced by the dual search and the existence of the optimal primal-dual pair and possesses an asymptotic strong duality. Monotonicity of duality gap in the partial pth power reformulation has been proved for single-constraint cases. Although the monotonicity of $\kappa(p)$ in multiple-constraint cases is witnessed in numerical experiments, it is still an open question to prove the monotonicity of duality gap for general multiply constrained cases.

Acknowledgments

The authors are grateful to an anonymous referee for his extremely valuable comments and suggestions, which have contributed to significant improvements of the paper. Especially, the current version of the description of the decomposition of domain Y in Section 2 and the expression of the convex envelope function of w are based on the corrections from the referee. The first author appreciates greatly the stimulating suggestions from Prof. Douglas J. White, especially on a conjecture that duality gap is strictly decreasing with respect to p in the pth power Lagrangian formulation.

References

1. Bell, D.E. and Shapiro, J.F. (1977), A convergent duality theory for integer programming, *Operations Research*, 25, 419–434.
2. Dentcheva, D. and Romisch, W. (2004), Duality gaps in nonconvex stochastic optimization, *Mathematical Programming*, 101, 515–535.

3. Fisher, M.L. (1981), The Lagrangian relaxation method for solving integer programming problems, *Management Science*, 27, 1–18.
4. Fisher, M.L. and Shapiro, J.F. (1974), Constructive duality in integer Programming, *SIAM Journal on Applied Mathematics*, 27, 31–52.
5. Geoffrion, A.M. (1974), Lagrangian relaxation for integer programming, *Mathematical Programming Study*, 2, 82–114.
6. Guignard, M. and Kim, S. (1987), Lagrangian decomposition: A model yielding stronger Lagrangian relaxation bounds, *Mathematical Programming*, 33, 262–273.
7. Holmberg, K. (1994), Cross decomposition applied to integer programming problems: Duality gaps and convexification in parts, *Operations Research*, 42, 657–668.
8. Lemaréchal, C. and Renaud, A. (2001), A geometric study of duality gaps, with applications, *Mathematical Programming*, 90, 399–427.
9. Li, D. (1995), Zero duality gap for a class of nonconvex optimization problems, *Journal of Optimization Theory and Applications*, 85, 309–324.
10. Li, D. (1999), Zero duality gap in integer programming: P-norm surrogate constraint method, *Operations Research Letters*, 25, 89–96.
11. Li, D. and Sun, X.L. (2000), Success guarantee of dual search in nonlinear integer programming: P-th Power Lagrangian Method, *Journal of Global Optimization*, 18, 235–254.
12. Li, D. and White, D.J. (2000), P-th power Lagrangian method for integer programming, *Annals of Operations Research*, 98, 151–170.
13. Michelon, P. and Maculan, N. (1991), Lagrangian decomposition for integer nonlinear programming with linear constraints, *Mathematical Programming*, 52, 303–313.
14. Michelon, P. and Maculan, N. (1993), Lagrangian methods for 0-1 quadratic programming, *Discrete Applied Mathematics*, 42, 257–269.
15. Murota, K. (1998), Discrete convex analysis, *Mathematical Programming*, 83, 313–371.
16. Nemhauser, G.L. and Wolsey, L.A. (1988), *Integer and Combinatorial Optimization*. John Wiley & Sons, New York.
17. Sen, S., Higle, J.L. and Birge, J.R. (2000), Duality gaps in stochastic integer programming, *Journal of Global Optimization*, 18, 189–194.
18. Shapiro, J.F. (1979), A survey of Lagrangian techniques for discrete optimization, *Annals of Discrete Mathematics*, 5, 113–138.
19. Sun, X.L. and Li, D. (2000), Asymptotic strong duality for bounded integer programming: A logarithmic-exponential dual formulation, *Mathematics of Operations Research*, 25, 625–644.
20. Williams, H.P. (1996), Duality in mathematics and linear and integer programming, *Journal of Optimization Theory and Applications*, 90, 257–278.
21. Wolsey, L.A. (1981), Integer programming duality: Price functions and sensitivity analysis, *Mathematical Programming*, 20, 173–195.