Journal of Global Optimization **22:** 263–270, 2002.

Convergence of duality bound method in partly convex programming

NGUYEN VAN THOAI
Department of Mathematics, University of Trier, 54286 Trier, Germany (E-mail: thoai@uni-trier.de)

Abstract. We discuss the convergence of a decomposition branch-and-bound algorithm using Lagrangian duality for partly convex programs in the general form. It is shown that this decomposition algorithm has all convergence properties as any known branch-and-bound algorithm in global optimization under usual assumptions. Thus, some strict assumptions discussed in the literature are avoidable.

Key words: Partly convex programming problems, Decomposition branch and bound algorithms, Lagrangian duality.

To Reiner Horst on his 60th birthday

1. Introduction

For the implementation of branch and bound algorithms in global optimization, the Lagrangian duality can serve as an efficient tool. Partly convex programming problems belong to the class of nonconvex optimization problems, for which the Lagrangian duality bound method can be successfully applied. In Ben-Tal et al (1994), Duer and Horst (1997), Duer (1999), Duer et al. (2000), Thoai (1997, 2000), branch and bound algorithms for some special problems of this class are presented, in which Lagrangian dual problems can be formulated equivalently as ordinary linear programs, and the resulting bounds are shown to be at least as good as the bounds computed by using the classical convexification techniques for nonconvex problems. Different kinds of assumptions are made for proving convergence properties of these algorithms.

In this article, we discuss the convergence of a decomposition algorithm using Lagrangian duality for partly convex programs in the general form. It is shown that this decomposition algorithm has all convergence properties as any known branch and bound algorithm in global optimization under the usual assumptions. Thus, some strict assumptions discussed in the above-mentioned papers are avoidable.

Preliminaries on partly convex programming problems and a decomposition branch and bound algorithm are given in the next section. The main results on the convergence of the algorithm is disccused in Section 3. The last section contains some conclusions.

2. Preliminaries

Let $\overline{X}$ and $\overline{Y}$ be convex subsets of $I\!R^n$ and $I\!R^p$, respectively. A function $f : \overline{X} \times \overline{Y} \to I\!R$ is called *partly convex* if the function $f(x, \cdot)$ is convex on $\overline{Y}$ for each $x \in \overline{X}$. A partly convex program can be formulated as follows:

$$\begin{array}{ll} \inf & F(x,y) \\ \text{s.t.} & G_i(x,y) \leqslant 0 \ (i = 1, \cdots, m) \\ & x \in X, \ y \in Y, \end{array} \tag{P}$$

where X and Y are closed convex subsets of $\overline{X}$ and $\overline{Y}$, respectively, and F and G_i $(i = 1, \cdots, m)$ are partly convex functions defined on $\overline{X} \times \overline{Y}$. We denote by Z the feasible set of Problem (P), i.e.,

$$Z = \{(x, y) : G_i(x, y) \leqslant 0 \ (i = 1, \cdots, m), x \in X, y \in Y\}. \tag{1}$$

Throughout this paper we assume that in Problem (P), the sets X and Y are compact, $F(x, y) > -\infty$ for $x \in X, y \in Y$, and that there exists an optimal solution whenever the feasible set Z defined in (1) is nonempty.

To apply a decomposition branch and bound algorithm for solving Problem (P), we assume that one can construct a simple compact convex set S^0 such that $\overline{X} \supset S^0 \supset X$, (e.g., S^0 is a simplex or a rectangle). The algorithm can be briefly described as follow.

ALGORITHM. Start with S^0. Compute a lowerbound $\mu_0 = \mu(S^0)$ and an upper bound γ_0 for the optimal value of the problem

$$\inf\{F(x, y) : G_i(x, y) \leqslant 0 \ (i = 1, \cdots, m), x \in S^0, y \in Y\}. \tag{2}$$

($\gamma_0 = F(x^0, y^0)$ if some feasible Solution $(x^0, y^0) \in Z$ is found, otherwise, $\gamma_0 = +\infty$). At Iteration $k \geqslant 0$, if $+\infty > \mu_k \geqslant \gamma_k$ or $\mu_k = +\infty$, then stop, (in the first case, (x^k, y^k) with $F(x^k, y^k) = \gamma_k$ is an optimal solution, in the second case, Problem (P) has no feasible solution). Otherwise, divide S^k into finitely many convex sets $S_1^k, \ldots, S_r^k$ satisfying $\bigcup_{i=1}^r S_i^k = S^k$ and $S^i \cap S^j = \emptyset$ for $i \neq j$, (the sets S^k and S_i^k are called 'partition sets'). Compute for each partition set a lower bound and an upper bound. Update the lower bound μ_k by choosing the minimum of lower bounds of all existing partition sets, and update the upper bound γ_k by using feasible points found so far. Delete all partition sets S such that $\mu(S) \geqslant \gamma_k$. If not all partition sets are deleted, let S^{k+1} be a partition set with the minimum lower bound. Go to Iteration $k + 1$.

For each partition set S generated throughout the algorithm, a lower bound $\mu(S)$ is computed by solving the Lagrangian dual problem of Problem (2) according to S, i.e.,

$$\mu(S) = \sup_{\lambda \geqslant 0} \inf\{F(x, y) + \sum_{i=1}^{m} G_i(x, y)\lambda_i : x \in S, y \in Y\}. \tag{3}$$

In the case that the algorithm does not terminate after a finite number of iterations, it generates at least one infinite 'decreasing' sequence $\{S^q\}$ of convex sets, i.e., $S^{q+1} \subset S^q \ \forall\, q$. For this case, the convergence of the algorithm is discussed in the next section.

3. Convergence of the algorithm

We begin the establishment of the convergence of the algorithm by showing some useful properties of the Lagrangian duality bound computed in (3).

LEMMA 1. *Let S, R be two partition sets satisfying $S \supseteq R$. Then*
(i) $\mu(S) \leqslant \mu(R)$;
(ii) If there exists $\lambda^0 \geqslant 0$ such that $\sum_{i=1}^m G_i(x, y)\lambda_i^0 > 0$ for all $x \in S, y \in Y$, then $\mu(S) = +\infty$.

Proof. (i) follows immediately from the definition of $\mu(S)$. To show (ii), let $\Lambda = \{\lambda \in I\!R^m : \lambda = k\lambda^0, k \geqslant 0\}$. Then we have

$$\begin{aligned}\mu(S) &= \sup_{\lambda \geqslant 0} \inf\{F(x, y) + \sum_{i=1}^m G_i(x, y)\lambda_i : x \in S, y \in Y\}\\ &\geqslant \sup_{\lambda \in \Lambda} \inf\{F(x, y) + \sum_{i=1}^m G_i(x, y)\lambda_i : x \in S, y \in Y\}\\ &= \sup_{\lambda \in \Lambda} \inf\{F(x, y) + k\sum_{i=1}^m G_i(x, y)\lambda_i^0 : x \in S, y \in Y\} = +\infty.\end{aligned}$$

The last equation follows from the assumption that $F(x, y) > -\infty$ for $x \in S, y \in Y$, and the assumption in (ii) by letting $k \to +\infty$. □

REMARK 1. The condition $\sum_{i=1}^m G_i(x, y)\lambda_i^0 > 0$ for $x \in S, y \in Y$, is fulfilled if, e.g., there is an index j such that

$$\{(x, y) : G_j(x, y) \leqslant 0, x \in S, y \in Y\} = \emptyset.$$

In this case, one can choose $\lambda^0 = (0, \cdots, 0, \lambda_j^0, 0, \cdots, 0)$ with $\lambda_j^0 = 1$.

Next, let us recall the concept of 'upper semicontinuity' of a point-to-set mapping (see e.g., Bank et al., 1983).

DEFINITION 1. Let $A \subset I\!R^n$. A point-to-set mapping $M : A \to I\!R^m$ is called 'upper semicontinuous according to Berg (u.s.c.B.)' at a point $x^* \in A$, if for each open set Ω containing $M(x^*)$ there exists a $\delta = \delta(\Omega) > 0$ such that $M(x) \subset \Omega$ $\forall x \in U(x^*, \delta) \cap A$, where $U(x^*, \delta)$ is an open ball with radius δ around x^*.

LEMMA 2. *Assume that, in Problem (P), the functions G_i $(i = 1, \cdots, m)$ are lower semicontinuous on $S^0 \times Y$ (recall that S^0 is a compact convex set containing X). Then the point-to-set mapping $M : S^0 \to I\!R^m$ defined by*

$$M(x) = \{\lambda \in I\!R^m : G_i(x, y) \leqslant \lambda_i \leqslant t_i \ (i = 1, \cdots, m) \text{ for some } y \in Y\}, \tag{4}$$

where, for each i,

$$t_i = \sup\{G_i(x, y) : x \in S^0, y \in Y\},$$

is u.s.c.B. at each $x \in S^0$.

Proof. First, it is easy to verify that the set $M(x)$ is compact for each $x \in S^0$. Next, for each $x \in S^0$, let $\{x^k\} \subset S^0$ be an arbitrary sequence such that $x^k \to x$, and let $\{\lambda^k\} \subset I\!R^m$ such that $\lambda^k \in M(x^k) \setminus M(x)$. From compactness of $M(x^k)$, by passing to a subsequence if necessary, we assume that and $\lambda^k \to \lambda$. We show that

$$\lambda \in M(x). \tag{5}$$

To do this, for each k, let $y^k \in Y$ be a point such that $G_i(x^k, y^k) \leqslant \lambda_i \leqslant t_i$ (see definition of M in (4)). Again by passing to a subsequence if necessary, assume that and $y^k \to y$. Since the functions G_i $(i = 1, \cdots, m)$ are lower semicontinuous on $S^0 \times Y$, it follows that, for each $(i = 1, \cdots, m)$,

$$G_i(x, y) \leqslant \lim_{k\to\infty} G_i(x^k, y^k) \leqslant \lim_{k\to\infty} \lambda_i^k = \lambda_i \leqslant t_i,$$

which implies by definition that $\lambda \in M(x)$.

Suppose that M is not u.s.c.B at x. Then, by definition, there exist an open set Ω containing $M(x)$, a sequence $x^k \to x$ and a sequence λ^k such that $\lambda^k \in M(x^k) \setminus \Omega$. Thus, let $\lambda^k \to \lambda$, then $\lambda \notin \Omega$, i.e., $\lambda \notin M(x)$, which is a contradiction to (5). □

LEMMA 3. *Assume that the algorithm generates an infinite subsequence of partition sets, $\{S^q\}$, such that*

$$S^{q+1} \subset S^q \text{ for all } q, \text{ and } \lim_{q\to\infty} S^q = \bigcap_{q=1}^{\infty} S^q = \{x^*\}. \tag{6}$$

Then

$$(\{x^*\} \times Y) \cap Z \neq \emptyset. \tag{7}$$

Proof. Suppose that (7) does not hold, i.e.,

$$\{y \in I\!R^p : G_i(x^*, y) \leqslant 0 \ (i = 1, \cdots, m), y \in Y\} = \emptyset. \tag{8}$$

From (8) it follows that

$$\{\lambda \in \mathbb{R}^m : \lambda_i \leqslant 0,\ G_i(x^*, y) \leqslant \lambda_i\ (i = 1, \cdots, m) \text{ for some } y \in Y\} = \emptyset. \tag{9}$$

Equation (9) implies that the two closed sets

$$T_1 = \{\lambda \in \mathbb{R}^m : \lambda_i \leqslant 0\ (i = 1, \cdots, m)\} \tag{10}$$

and

$$T_2 = \{\lambda \in \mathbb{R}^m : G_i(x^*, y) \leqslant \lambda_i\ (i = 1, \cdots, m) \text{ for some } y \in Y\} \tag{11}$$

are disjointed. Note that T_1 is convex and T_2 is exactly the projection of the convex set

$$\{(y, \lambda) \in \mathbb{R}^{p+m} : G_i(x^*, y) - \lambda_i \leqslant 0\ (i = 1, \cdots, m),\ \ y \in Y\}$$

on $\mathbb{R}^m$, and hence is also convex. Thus, T_1 and T_2 can be separated by a hyperplane of the form

$$\{\lambda \in \mathbb{R}^m : \lambda^0 \lambda = 0\} \text{ with some } \lambda^0 \in \mathbb{R}^m \tag{12}$$

such that

$$\lambda^0 \lambda \leqslant 0 \text{ for } \lambda \in T_1 \text{ and } \lambda^0 \lambda > 0 \text{ for } \lambda \in T_2. \tag{13}$$

Moreover, it must hold that

$$\lambda^0 \geqslant 0, \tag{14}$$

because otherwise, letting $\lambda_j^0 < 0$ and choosing

$$\overline{\lambda} = (0, \cdots, 0, \overline{\lambda}_j = -1, 0, \cdots, 0) \in T_1,$$

we have $\lambda^0 \overline{\lambda} = -\lambda_j^0 > 0$, which contradicts the property $\lambda^0 \lambda \leqslant 0$ for $\lambda \in T_1$.

Next, let Ω be an open convex set containing the closed convex set T_2 such that T_1 and Ω can still be separated by the hyperplane (12), i.e., we also have

$$\lambda^0 \lambda > 0 \text{ for } \lambda \in \Omega. \tag{15}$$

Let M be the point-to-set mapping defined in (4). Then

$$M(x^*) = T_2 \subset \Omega.$$

Since M is u.s.c.B. at x^* (by Lemma 2) and

$$\lim_{q \to \infty} S^q = \bigcap_{q=1}^{\infty} S^q = \{x^*\},$$

it follows that there exists an index q^* such that $M(x) \subset \Omega$ for all $x \in S^{q^*}$. Therefore, if for each $x \in S^{q^*}$ and each $y \in Y$, let $\lambda(x, y)$ be a vector with $\lambda_i(x, y) = G_i(x, y)$ $(i = 1, \cdots, m)$, then $\lambda(x, y) \in M(x) \subset \Omega$. From (14) and (15), it follows that

$$\sum_{i=1}^{m} G_i(x, y)\lambda_i^0 = \sum_{i=1}^{m} \lambda_i(x, y)\lambda_i^0 = \lambda^0\lambda(x, y) > 0 \text{ for } x \in S^{q^*},\ y \in Y, \quad (16)$$

which implies by Lemma 1 that $\mu(S^{q^*}) = +\infty$, i.e., the partition set S^{q^*} has to be removed at Iteration q^* of the algorithm. By this contradiction, the proof of (7) is completed. □

Using the previous results we can now prove the convergence properties of the algorithm.

THEOREM 1. *Assume that the algorithm generates an infinite subsequence of partition sets, $\{S^q\}$, such that*

(i) $S^{q+1} \subset S^q$ *for all* q *and* $\lim_{q\to\infty} S^q = \bigcap_{q=1}^{\infty} S^q = \{x^*\}$,
(ii) there is zero duality gap at x^*, *i.e.,*

$$\begin{aligned} &\inf\{F(x^*, y) : G_i(x^*, y) \leqslant 0\ (i = 1, \cdots, m), y \in Y\} = \\ &= \sup_{\lambda \geqslant 0}\ \inf\{F(x^*, y) + \sum_{i=1}^{m} G_i(x^*, y)\lambda_i : y \in Y\}. \end{aligned} \quad (17)$$

Then (x^*, y^*) *is an optimal solution of Problem (P), where* y^* *is an optimal solution of the convex program*

$$\inf\{F(x^*, y) : G_i(x^*, y) \leqslant 0\ (i = 1, \cdots, m), y \in Y\}. \quad (18)$$

Proof. For each q, let

$$w_q(\lambda) = \inf\{F(x, y) + \sum_{i=1}^{m} G_i(x, y)\lambda_i : x \in S^q, y \in Y\}, \quad (19)$$

and let λ^q be an optimal solution of the problem $\sup\{w_q(\lambda) : \lambda \geqslant 0\}$, i.e., $\mu(S^q) = w_q(\lambda^q)$. Moreover, let

$$w^*(\lambda) = \inf\{F(x^*, y) + \sum_{i=1}^{m} G_i(x^*, y)\lambda_i : y \in Y\}. \quad (20)$$

First, we show that $w^*(\lambda) = \sup_q w_q(\lambda)$ for each λ. By definition, it is obvious that $w^*(\lambda) \geqslant \sup_q w_q(\lambda)$. On the other hand, for each q, let $x^q \in S^q$, $y^q \in Y$ such that

$$w_q(\lambda) = F(x^q, y^q) + \sum_{i=1}^{m} G_i(x^q, y^q)\lambda_i. \quad (21)$$

Then, from Assumption (i) and Lemma 3, it follows, by passing to subsequences if necessary, that $\lim_{q\to\infty}(x^q, y^q) = (x^*, y^*) \in Z$, where $y^* \in Y$. This implies that

$$\sup_q w_q(\lambda) = \lim_{q\to\infty} w_q(\lambda) = F(x^*, y^*) + \sum_{i=1}^{m} G_i(x^*, y^*)\lambda_i \geqslant w^*(\lambda). \qquad (22)$$

Thus, we have $w^*(\lambda) = \sup_q w_q(\lambda)$.

Since the sequence $\{\mu(S^q)\}$ of lower bounds is nondecreasing (Lemma 1) and bounded by the optimal value of Problem (P), its limit, μ^*, exists, and we have

$$\begin{aligned} \mu^* &= \lim_{q\to\infty} \mu(S^q) = \lim_{q\to\infty} w_q(\lambda^q) = \lim_{q\to\infty} \sup_{\lambda\geqslant 0} w_q(\lambda) \\ &= \sup_q \ \sup_{\lambda\geqslant 0} w_q(\lambda) = \sup_{\lambda\geqslant 0} \ \sup_q w_q(\lambda) = \sup_{\lambda\geqslant 0} w^*(\lambda). \end{aligned} \qquad (23)$$

From Assumption (ii), it follows then

$$\mu^* = \inf\{F(x^*, y) : G_i(x^*, y) \leqslant 0 \ (i = 1, \cdots, m), y \in Y\}. \qquad (24)$$

Let y^* be an optimal solution of this problem. Then, since μ^* is a lower bound of the optimal value of (P), it follows that (x^*, y^*) is an optimal solution of (P). □

REMARK 2. Since the problem

$$\inf\{F(x^*, y) : G_i(x^*, y) \leqslant 0 \ (i = 1, \cdots, m), y \in Y\}$$

is an ordinary convex program, Condition (ii) of Theorem 1 is fulfilled under well known *constraint qualifications* (cf., e.g., Geoffrion, 1971; Mangasarian, 1969).

DEFINITION 2. A subsequence of partition sets $\{S^q\}$ generated by the algorithm is called to be 'exhaustive' if $S^{q+1} \subset S^q$ for all q and

$$\lim_{q\to\infty} S^q = \bigcap_{q=1}^{\infty} S^q = \{x^*\},$$

where x^* is a point in $\mathbb{R}^n$.

A partition process used within the algorithm is called exhaustive if every subsequence of partition sets generated by the algorithm is exhaustive.

Using this concept of an exhaustive partition process, we obtain immediately from Lemma 3 the following.

THEOREM 2. *Assume that within the algorithm an exhaustive partition process is used. Then the algorithm terminates after finitely many iterations if the feasible set Z of Problem (P) is empty.*

4. Conclusions

In this article we have shown that for partly convex programs, the decomposition branch and bound algorithm using duality bounds has all convergence properties as any known branch and bound algorithm in global optimization. These properties are stated in the theorems 1 and 2 under the usual assumptions that the partition process is exhaustive in the space $I\!R^n$ and there is some constraint qualification for each convex program obtained from Problem (P) by fixing x. Thus, we see that all other assumptions discussed in Ben-Tal et al. (1993), Duer and Horst (1997) and Thoai (1997) are avoidable. For the case $p = 0$, Problem (P) is a nonconvex programming problem of the form

$$\begin{aligned} \inf \quad & F(x) \\ \text{s.t.} \quad & G_i(x) \leqslant 0 \ (i = 1, \cdots, m) \\ & x \in X. \end{aligned} \tag{25}$$

The results in the Theorems 1 and 2 are stated by the following.

THEOREM 3. *(i) If the algorithm generates an infinite subsequence of partition sets, $\{S^q\}$, such that $S^{q+1} \subset S^q$ for all q and*

$$\lim_{q\to\infty} S^q = \bigcap_{q=1}^{\infty} S^q = \{x^*\},$$

then x^* is an optimal solution of Problem (25).

(ii) If Problem (25) has no feasible solution, and an exhaustive partition process is used, then the algorithm terminates after finitely many iterations indicating this fact.

References

Bank, B., Guddat, J., Klatte, D., Kummer, B. and Tammer, K. (1983), *Nonlinear Parametric Optimization*. Birkhaeuser, Basel.

Ben-Tal, A., Eiger, G. and Gershovitz, V. (1994), Global Minimization by Reducing the Duality Gap. *Mathematical Programming* 63: 193–212.

Duer, M. (1999), Duality in Global Optimization: Optimality Conditions and Algorithmical Aspects Ph.D. Thesis, University of Trier.

Duer, M. and Horst, R. (1997), Lagrange Duality and Partioning Techniques in Nonconvex Global Optimization. *Journal of Optimization Theory and Applications* 95: 347–369.

Duer, M., Horst, R. and Thoai, N.V. (2001), Solving Sum-of-Ratios Fractional Programs Using Efficient Points *Optimization* 49: 447–466.

Geoffrion, A.M. (1971), Duality in Nonlinear Programming: A Simplified Application-Oriented Development. *SIAM Review* 13, 1-37.

Mangasaian, O.L. (1979), *Nonlinear Programming*. Robert E. Krieger, Huntington, New York.

Thoai, N.V. (1997), On Convergence and Application of a Decomposition Method Using Duality Bounds for Nonconvex Global Optimization. Research Report Nr. 97–24, University of Trier, Germany. Forthcoming in *Journal of Optimization Theory and Applications*.

Thoai, N.V. (2000), Duality Bound Methods for the General Quadratic Programming Problem with Quadratic Constraints *Journal of Optimization Theory and Applications* 107, 331–354.