A Class of Problems where Dual Bounds Beat Underestimation Bounds

Dür, Mirjam

A Class of Problems where Dual Bounds Beat Underestimation Bounds

Mirjam Dür

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 79
October 2000

http://statmath.wu-wien.ac.at/

A Class of Problems where Dual Bounds Beat Underestimation Bounds *

Mirjam Dür
Department of Statistics, Vienna University of Economics and Business Administration, A-1090 Vienna, Austria

October 24, 2000

Abstract. We investigate the problem of minimizing a nonconvex function with respect to convex constraints, and we study different techniques to compute a lower bound on the optimal value: The method of using convex envelope functions on one hand, and the method of exploiting nonconvex duality on the other hand. We investigate which technique gives the better bound and develop conditions under which the dual bound is strictly better than the convex envelope bound. As a byproduct, we derive some interesting results on nonconvex duality.

Keywords: Nonconvex duality, Dual bounds, Convex underestimation.

1. Introduction

In this paper, we consider the global optimization problem of minimizing a nonconvex function subject to convex constraints:

$$(P) \qquad \begin{array}{c} \min f(x) \\ \text{s.t. } h_i(x) \leq 0, \quad i = 1, \dots, m, \\ x \in X, \end{array}$$

where $f : X \to \mathbb{R}$ is a lower semicontinuous function, $h_i : X \to \mathbb{R}$ $(i = 1, \dots, m)$ are convex functions, and $X \in \mathbb{R}^m$ is a convex compact set. Our aim is to study different methods to obtain lower bounds for the optimal value of (P).

The first technique which has been used since many years is to replace the objective function with some easier (i.e. convex or linear) subfunctional and solve the resulting problem. Obviously, the quality of a bound obtained by this means depends on the quality of the underestimating function. The best possible result is achieved when the so called convex envelope function is used.

DEFINITION 1. *Let $X \subset \mathbb{R}^n$ be convex and compact, and let $f : X \to \mathbb{R}$ be lower semicontinuous on X. A function $\varphi_f : X \to \mathbb{R}$ is called the convex envelope of f on X if it satisfies*

* Dedicated to Reiner Horst on the occasion of his 60th birthday.

EXAMPLE 2. *Consider the one dimensional problem*

$$\min_{x\in[-2,3]} \{-x^2 : x^2 - x - 2 \le 0\}.$$

The optimal value is $\min(P) = -4$, *attained at* $x = 2$. *The convexified problem* $(\bar{P})$ *takes the form*

$$\min_{x\in[-2,3]} \{-x - 6 : x^2 - x - 2 \le 0\}.$$

Its optimal value is $\min(\bar{P}) = -8$, *also attained at* $x = 2$. *The dual* (D) *of* (P) *is*

$$\sup_{\lambda\in\mathbb{R}_+} \min_{x\in[-2,3]} \{(\lambda - 1)x^2 - \lambda x - 2\lambda\},$$

which takes the optimal value $\sup(D) = -4.2$ *at* $\lambda = 6/5$ *and* $x = 3$. *The poor lower bound provided by* $(\bar{P})$ *is therefore improved considerably.*

In the remainder of the paper we develop conditions which guarantee that the dual bound is strictly better than the convex envelope bound.

2. Some Results on Nonconvex Duality

It is well known that in convex programming, Slater's constraint qualification ensures strong duality for (P) and (D), see, e.g., Geoffrion (1971). Since in nonconvex programming this condition turns out to be very useful as well, recall that problem (P) is said to fulfill Slater's condition if there exists a point $\hat{x} \in X$ such that $h_i(\hat{x}) < 0$ for all $i = 1, \ldots, m$.

In the sequel, we will use the notation

$$\mathcal{S} := \{\hat{x} \in X : h_i(\hat{x}) < 0 \text{ for all } i = 1, \ldots, m\}$$

to denote the set of all Slater points,

$$\bar{h}(x) := \max_{i=1,\ldots,m} h_i(x)$$

to denote the pointwise maximum of the constraint functions,

$$L(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$$

to denote the Lagrangean function of (P), and

$$\Theta(\lambda) := \min_{x\in X} L(x, \lambda)$$

Finally, using

$$\sup(D) = \sup_{\lambda \in \mathbb{R}^m_+} \Theta(\lambda) \geq \Theta(0) = \min_{x \in X} f(x),$$

we obtain

$$\rho(\hat{x}) = \frac{\sup(D) - f(\hat{x})}{\bar{h}(\hat{x})} \leq \frac{\min_{x \in X} f(x) - f(\hat{x})}{\bar{h}(\hat{x})}.$$

Obviously, $\|\bar{\lambda}\|_1 \leq \rho(\hat{x})$ for any Slater point $\hat{x} \in \mathcal{S}$. Hence we get the desired upper bound for $\|\bar{\lambda}\|_1$:

$$\|\bar{\lambda}\|_1 \leq \inf_{\hat{x} \in \mathcal{S}} \frac{\min_{x \in X} f(x) - f(\hat{x})}{\bar{h}(\hat{x})}.$$

□

This result seems interesting in its own right, but it may also prove useful in a numerical context: Note that any Slater point $\hat{x}$ gives the a priory bound

$$\|\bar{\lambda}\|_1 \leq \frac{\min_{x \in X} f(x) - f(\hat{x})}{\bar{h}(\hat{x})},$$

which may be helpful when solving dual problems with bundle–type methods. Of course, finding a Slater point is a difficult task in general, but may be easy when the constraints are simple, e.g. box constraints. This reasoning also applies to the so called standard quadratic problem of maximizing an indefinite quadratic form on the standard simplex, see Bomze (1998).

3. When are dual bounds better?

In this section we return to the question which of the bounds $\min(\bar{P})$ and $\sup(D)$ is better. Theorem 4 states under which assumptions on objective and constraint functions the dual bound beats the convex envelope bound.

But first observe that it may happen that $\min(P) = \min(\bar{P})$. In this case, it follows from (1) and weak duality that

$$\sup(D) = \min(\bar{P}) = \min(P),$$

in other words, the duality gap is zero and both bounds are equal. For this reason, the mentioned case is excluded in the theorem.

$$< f(x) + \sum_{i \in I(\bar{x})} \lambda_i h_i(x) + \sum_{i \notin I(\bar{x})} \lambda_i h_i(x) \tag{3}$$
$$= L(x, \lambda).$$

Inequality (3) holds because for $x \neq \bar{x}$ we have $h_i(x) < \langle x - \bar{x}, \nabla h_i(\bar{x}) \rangle$, as all h_i are strictly convex.

Since $\varphi_f(\bar{x}) < f(\bar{x})$ (recall that $\min(\bar{P}) < \min(P)$ by assumption), we obviously have $\tilde{L}(\bar{x}, \lambda) < L(\bar{x}, \lambda)$ for every $\lambda \in \mathbb{R}^m_+$, and hence

$$\tilde{L}(x, \lambda) < L(x, \lambda) \qquad \forall\, x \in X, \quad \forall \lambda \in \Lambda(\bar{x}).$$

Therefore, we get for the dual objective function $\tilde{\Theta}(\lambda)$ of $(\tilde{P})$

$$\tilde{\Theta}(\lambda) = \min_{x \in X} \tilde{L}(x, \lambda) < \min_{x \in X} L(x, \lambda) = \Theta(\lambda) \qquad \forall \lambda \in \Lambda(\bar{x}).$$

Next we show that $\max_{\lambda \in \mathbb{R}^m_+} \tilde{\Theta}(\lambda)$ is attained at some $\tilde{\lambda} \in \Lambda(\bar{x})$:

Let $\tilde{\lambda}$ denote a solution of $(\tilde{D})$, i.e. $\sup(\tilde{D}) = \tilde{\Theta}(\tilde{\lambda})$, and recall that $\bar{x}$ is the optimal solution of $(\tilde{P})$. Since $(\tilde{P})$ is a convex problem, the optimal primal dual pair $(\bar{x}, \tilde{\lambda})$ fulfills the complementary slackness condition (see Geoffrion (1971))

$$\sum_{i \in I(\bar{x})} \tilde{\lambda}_i \langle \bar{x} - \bar{x}, \nabla h_i(\bar{x}) \rangle + \sum_{i \notin I(\bar{x})} \tilde{\lambda}_i h_i(\bar{x}) = 0.$$

It follows that $\tilde{\lambda}_i = 0$ for all $i \notin I(\bar{x})$.

But from the assumptions of the theorem it follows that $\tilde{\lambda} \neq 0$: Assume that $\tilde{\lambda} = 0$. Then, since $(\bar{x}, \tilde{\lambda})$ is a saddle point of $\tilde{L}(x, \lambda)$, it follows that $\tilde{L}(\bar{x}, 0) \leq \tilde{L}(x, 0)$ for all $x \in X$, in other words, $\varphi_f(\bar{x}) \leq \varphi_f(x)$ for all $x \in X$, and therefore

$$\varphi_f(\bar{x}) = \min_{x \in X} \varphi_f(x).$$

As $\min_{x \in X} \varphi_f(x) = \min_{x \in X} f(x)$, we have $\varphi_f(\bar{x}) = \min_{x \in X} f(x)$. Since f is a concave function, the minimum of f over X is attained at some extremal point of X. Therefore, either $\bar{x}$ is an extreme point. Then $\varphi_f(\bar{x}) = f(\bar{x})$ and $\bar{x}$ would solve both $(\bar{P})$ and (P), a contradiction.

Or there exist $k \leq n + 1$ extremal points $v^1, \ldots, v^k$ of X, such that $\bar{x}$ is a convex combination of these extremal points and $f(v^j) = \varphi_f(\bar{x})$. But then φ_f would be constant on the convex hull of $\{v^1, \ldots, v^k\}$, which contradicts the assumptions as well.

Therefore, we conclude that $\tilde{\lambda} \in \Lambda(\bar{x})$.

To sum up, let $\bar{\rho}$ denote the maximum of the $\|\lambda\|_1$–bounds obtained via Theorem 3 for problems (P) and $(\tilde{P})$, respectively, and get

$$\sup(\tilde{D}) \;=\; \max \left\{ \tilde{\Theta}(\lambda) : \lambda \in \mathbb{R}^m_+, \|\lambda\|_1 \leq \bar{\rho}, \lambda \in \Lambda(\bar{x}) \right\}$$

Proof. Let $\tilde{x}$ denote the minimizer of problem (5). Clearly, the feasible set of problem (4) is contained in that of problem (5), hence

$$\varphi(\bar{x}) \geq \varphi(\tilde{x}).$$

Now assume that $\varphi(\bar{x}) > \varphi(\tilde{x})$. Since $\bar{x}$ is optimal for problem (4), there does not exist a feasible descent direction of φ at $\bar{x}$, i.e. there does not exist a direction d with

$$\langle d, \nabla h_i(\bar{x}) \rangle < 0 \quad \text{and} \quad \varphi'_d(\bar{x}) < 0, \tag{6}$$

where φ'_d denotes the directional derivative of φ in direction d.

Because of the strict convexity of all constraint functions and because of Slater's condition, we can assume that there exists a point $\breve{x}$ feasible for (5) such that

$$\langle \breve{x} - \bar{x}, \nabla h_i(\bar{x}) \rangle < 0 \quad \text{for all} \quad i \in I(\bar{x}),$$

and $\varphi(\breve{x}) < \varphi(\bar{x})$. But then $d := \breve{x} - \bar{x}$ is a feasible descent direction of φ at $\bar{x}$, since it fulfills conditions (6). This contradicts the optimality assumption on $\bar{x}$. □

References

Bomze, I.M. On Standard Quadratic Optimization Problems. *Journal of Global Optimization*, 13:369–387, 1998.

Dür, M. Dual Bounding Procedures Lead to Convergent Branch-and-Bound Algorithms. *Submitted*, 2000.

Dür, M. and R. Horst. Lagrange-Duality and Partitioning Techniques in Nonconvex Global Optimization. *Journal of Optimization Theory and Applications*, 95:347–369, 1997.

Falk, J.E. Lagrange Multipliers and Nonconvex Programs. *SIAM Journal on Control*, 7:534–545, 1969.

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

Horst, R. and H. Tuy. *Global Optimization*. Berlin (Springer), 1996.

Nowak, I. Dual Bounds and Optimality Cuts for All-Quadratic Programs with Convex Constraints. *Submitted*, 2000.

Thoai, N.V. On the Convergence and Implementation of a Decomposition Method Using Duality Bounds for Nonconvex Global Optimization. *Submitted*, 2000.