Dual Estimates in Multiextremal Problems*

N. Z. SHOR
Institute of Cybernetics, Academy of Sciences of the Ukrainian SSR, Kiev 252207, Ukraine

(Received: 14 December 1990; accepted: 27 November 1991)

Abstract. We propose a technique of improving the dual estimates in nonconvex multiextremal problems of mathematical programming, by adding some additional constraints which are the consequences of the original constraints. This technique is used for the problems of finding the global minimum of polynomial functions, and extremal quadratic and boolean quadratic problems. In the article one ecological multiextremal problem and an algorithm for finding the dual estimate for it also considered. This algorithm is based upon a scheme of decomposition and nonsmooth optimization methods.

Key words. Dual estimates, polynomial functions, non-convex quadratic problems, nondifferentiable optimization.

In this report the general mathematical programming problem find:

$$f^* = \inf_{x \in X} f_0(x), \quad X \subseteq E^n \tag{1}$$

subject to the constraints:

$$f_i(x) \leqslant 0\,, \quad i \in I_1\,; \quad f_1(x) = 0\,, \quad i \in I_2 \tag{2}$$

is considered.

Let u be a vector of Lagrange multipliers and

$$L(x, u) = f_0(x) + \sum_{i \in I_1 \cup I_2} u_i f_i(x)$$

be the Lagrange function.

On the set $U = \{u : u_i \geqslant 0, i \in I_1\}$ let us consider the function:

$$\Psi(u) = \inf_{x \in X} L(x, u)\,.$$

The value $\Psi^* = \sup_{u \in U} \Psi(u)$ is called the dual estimate for f^*. It is clear that for $u \in U$ $\Psi(u) \leqslant f^*$ and consequently $\Psi^* \leqslant f^*$; therefore, this dual estimate is a lower estimate for f^*. In the nonconvex case the so-called "estimation gap" may occur:

$$\Delta := f^* - \Psi^* > 0\,.$$

*This paper was presented at the II. IIASA Workshop on Global Optimization, Sopron (Hungary), December 9–14, 1990.

Journal of Global Optimization **2**: 411–418, 1992.

One of the ways to diminish this gap consists in adding to the constraints in (2) formally new constraints which are consequences of the constraints (2). So, the set of feasible points of the problem (1)–(2) is not changed, but the set of Lagrange variables is extended. In some cases the gap can be reduced to zero by this way. In this article some examples are considered to illustrate such an approach.

1. The Global Minimization Problem for a Polynom

Let a (bounded-from-below) polynomial $P(x_1, x_2, \ldots, x_n)$ be given and let P^* be the value of the polynomial at the global minimum point.

By introducing new variables and making use of quadratic substitutions of the form: $x_i^2 = y_i$; $x_j x_k = z_{jk}$ and so forth we can reduce the minimization problem for the polynomial $P(x_1, \ldots, x_n)$ to a quadratic extremal problem with constraints in form of equalities. The direct application of the dual estimate technique to this quadratic problem results in nontrivial estimates only in rare cases. But if we modify the quadratic problem by generating simple quadratic equalities of the quadratic problem variables and by adding these equalities to the constraints some interesting results can be obtained for the modified problem.

THEOREM 1. *The dual estimate for the modified quadratic problem, which is equivalent to the problem of minimization of the polynomial* $P(x) = P(x_1, x_2, \ldots, x_n)$ *equals* P^* *iff the nonnegative polynomial* $P(x_1, x_2, \ldots, x_n) := P(x_1, x_2, \ldots, x_n) - P^*$ *can be represented as the sum of squares of other polynomials. Particularly, for* $n = 1$ *the dual estimate is exact.*

First, we give a more precise definition of the "modified quadratic problem" in Theorem 1 and an illustrative example.

Let $P^* > -\infty$. Then the highest degree S_i of each of the variables x_i must be even. Let $S_i = 2l_i$, $i = 1, \ldots, n$. Consider integer vectors $a = \{a_1, \ldots, a_n\}$ with non-negative elements and monomials of the type:

$$R[a] = x_1^{a_1} \ldots x_1^{a_n}, \quad 0 \leq a_i \leq l_i, \quad i = 1, \ldots n. \tag{3}$$

Then one obtains a system of identity relations:

$$R[a^{(1)}]R[a^{(2)}] - R[a^{(3)}]R[a^{(4)}] = 0 \tag{4}$$

for all

$$\{a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}\} \text{ whenever}$$
$$0 \leq a^{(1)} + a^{(2)} = a^{(3)} + a^{(4)} \leq S = \{S_i\}_{i=1}^n .$$

Any $P(x)$ with the vector $\{S_i\}_{i=1}^n = \{2l_i\}_{i=1}^n$ of highest degrees can be written as a quadratic function of the variables $R[a^{(i)}]$ which has the form

$$P(x) = L(R, \lambda) = \sum_{i,j} c_{ij} R[a^{(i)}]R[a^{(j)}]$$
$$+ \sum_{(k,l,m,n)} \lambda_{klmn} (R[a^{(k)}]R[a^{(1)}] - R[a^{(m)}]R[a^{(n)}] ,$$

where $R[a^{(i)}]R[a^{(j)}]$ are some representatives of monomials of type (3) the c_{ij} are appropriate coefficients, the λ_{klmn} are arbitrary multipliers at the left-hand side of (4). Alternatively, $L(R, \lambda)$ can be considered as a Lagrange function of the modified quadratic optimization problem

$$\text{minimize } K[R] = \sum_{i,j} c_{ij} R[a^{(i)}]R[a^{(j)}]$$

subject to (4). This problem is used to obtain the dual estimate in theorem 1. The set of constraints (4) is redundant as a rule.

EXAMPLE. Let

$$P(x_1) = x_1^6 + ax_1^5 + bx_1^4 + cx_1^3 + dx_1^2 + ex_1 ;$$

$S_1 = 6$; $l_1 = 3$. The problem is to find the global minima of $P(x_1)$. Enumerate the variables of the equivalent quadratic problem:

$$R[1] = x_1 ; \quad R[2] = x_1^2 = x_2 ; \quad R[3] = x_1^3 = x_3 .$$

The modified quadratic extremal problem has the form:

$$\text{minimize } K(x_1, x_2, x_3) = x_3^2 + ax_2x_3 + bx_2^2 + cx_3 + dx_2 + ex_1 \tag{5}$$

subject to the constraints

$$x_1^2 - x_2 = 0 , \tag{6}$$
$$x_1x_2 - x_3 = 0 , \tag{7}$$
$$x_2^2 - x_1x_3 = 0 . \tag{8}$$

Note that the last equation (8) is unnecessary. The Lagrange function $L(x, \lambda)$ for problem (5)–(8) has the form

$$L(x, \lambda) = K(x_1, x_2, x_3) + \lambda_1(x_1^2 - x_2) + \lambda_2(x_1x_2 - x_3) + \lambda_3(x_2^2 - x_1x_3) .$$

Let

$$\Psi(\lambda) = \inf_x L(x, \lambda) ; \quad \Omega = \operatorname{dom} \Psi ; \quad \Psi^* = \sup_{\lambda \in \operatorname{dom}\Psi} \Psi(\lambda) ,$$

and denote by λ^* the optimal vector λ if it exists. Note that if $b < 0$ then dom Ψ_{12}, $\Psi_{12}(\lambda_1, \lambda_2) = \Psi(\lambda_1, \lambda_2, 0)$ is empty so that the "unnecessary" equation (8) is necessary for finding the nontrivial dual lower bound Ψ^*. By using Theorem 1 it follows that $\Psi^* = \min_{x_1} P(x_1)$. Let $\Omega_0 \subset \Omega$ denote the interior of Ω. If Ω_0 is nonempty, then it is a convex set, since $\Omega_0 = \{\lambda_0 : L(x, \lambda_0)$ is positive definite on $x\}$. Problem (5)–(8) is equivalent to three "simplified" problems (I), (II), (III), where (I): (5), (6), (7); (II): (5), (6), (8); (III): (5), (7), (8). Suppose each of the problems (I), (II), (III) has a single optimal vector of Lagrange multipliers: $\lambda_I = \{\lambda_1^1, \lambda_2^1\}$, $\lambda_{II} = \{\lambda_1^2, \lambda_3^2\}$, $\lambda_{III} = \{\lambda_2^3, \lambda_3^3\}$. Appropriate vectors $\lambda(1) = \{\lambda_1^{(1)}, \lambda_2^{(1)}, 0\}$, $\lambda(2) = \{\lambda_1^{(2)}, 0, \lambda_3^{(2)}\}$, $\lambda(3) = \{0, \lambda_2^{(3)}, \lambda_3^{(3)}\}$ can be arranged on a line L. The statement "$L \cap \Omega \neq \{\varnothing\}$" is then equivalent to the statement "$\Omega^* = \min_{x_1} P(x_1)$". If $L \cap \Omega^0$ is nonempty then the global minimum is unique.

Proof of Theorem 1. Let $\Psi(\lambda) = \inf_R L(R, \lambda)$ $\Omega = \operatorname{dom} \Psi$ and denote by Ω_0 the interior of Ω. If $L(R, \lambda_0)$ is a positive definite quadratic function of the variables R then $\lambda \in \Omega_0$. If Ω is nonempty, then $\Psi(\lambda)$ is concave function on the convex set Ω.

Let $\sup_{\lambda \in \operatorname{dom} \Psi} \Psi(\lambda) = \Psi^*$. The proof of Theorem 1 is based on the following lemma.

LEMMA 1 (*about translation*). *Let the polynomial $P(z)$, $z \in E^n$, have the property that its quadratic dual estimate Ψ_P^* equals $\min_z P(z)$. Then for arbitrary $h \in E^n$ the polynom $P_h(z) = P(z + h)$ possesses the same property.*

The proof of this lemma is published in [8]. It essentially uses sufficient completeness of the set of Equations (4). We shall use also the following simple lemma.

LEMMA 2. *Any linear form $L(y) = \Sigma_{i=1}^k c_i y_i$ under the condition $\Sigma_{i=1}^k c_i = 0$ can be expressed as $\Sigma_{i,j=1}^k c_{ij}(y_i - y_j)$ for some coefficents $\{c_{ij}\}_{i,j=1}^n$.*

Using Lemma 1 and Lemma 2, Theorem 1 will follow from the following result.

THEOREM 2. *Suppose that we have a polynomial $P(z) = P(z_1, \ldots, z_n)$ with global minimum at the point $z = 0$, and $P(0) = 0$. Then the dual quadratic estimate satisfies $\Psi^* = 0$ iff the polynom $P(z)$ can be represented as the sum of squares of other polynoms.*

Proof. (Necessity). Let $\Psi^* = 0$. Then one can find $\lambda^* \in \Omega$ such that $L(R, \lambda^*)$ is positive semidefinite and $L(0, \lambda^*) = 0$. But arbitrary positive semidefinite quadratic functions can be represented as sum of squares of linear functions. Substitute in the formula of the function $L(R, \lambda^*)$ represented in form of sum of squares instead of variables $R[\alpha]$ equivalent monomials of the variables $z_1, \ldots, z_n$ to obtain a representation of $P(z)$ in the form of sum of squares of polynomials.

(Sufficiency). Let $P(z)$ be represented as a sum of squares of polynomials $P_1, \ldots, P_m$: $P(z) = \Sigma_{i=1}^m P_i^2(z)$. Polynomials $P_i(z)$, $i = 1, \ldots, m$, do not contain variable-free coefficients (if it were not the case, then $P(0) > 0$). Substituting instead of each monomial contained in the polynomials $P_i(z)$, $i = 1, \ldots, m$, the corresponding $c_\alpha R[\alpha]$ one obtains a homogeneous positive semidefinite quadratic form $\bar{K}(R)$ of the vector R satisfying $\min \bar{K}(R) = \bar{K}(0) = 0$. Consider the difference

$$\Delta(R) = \bar{K}(R) - L(R, 0); \quad (\lambda = 0).$$

The sum of the coefficients of $\Delta(R)$ must be equal to zero, so by Lemma 2 $\Delta(R)$ is represented as linear combination of the left sides of the constraints (4). The coefficients of this linear form can be interpreted as Lagrange multipliers at corresponding constraints. So we obtain a vector of Lagrange multipliers $\bar{\lambda}$ such that $L(R, \bar{\lambda}) = \bar{K}(R)$, $\Psi(\bar{\lambda}) = 0$. Theorem 2 is proved.

Let now $\min P(z) = P(z^*) = P^*$. Consider polynomials $P_0(z) = P(z - z^*) - P^*$; $\min P_0(z) = P_0(0) = 0$. Using Lemma 1 and Theorem 2 we obtain the statement of Theorem 1.

Unfortunately not every non-negative polynomial of several variables can be represented as the sum of squares of polynomials. This problem was investigated by Hilbert in 1888 [2]. He studied the homogeneous polynomial forms of even degree m with n variables. Hilbert showed that for $n = 3$, $m \geqslant 6$ and $n \geqslant 4$, $m \geqslant 4$ there exist non-negative forms which cannot be represented as the sum of squares of other forms.

Only for some general classes of forms this question is decided positively:

- $m = 2$, n is arbitrary (quadratic forms);
- $n = 2$, m is arbitrary even number (two-variable forms or corresponding polynomials of one variable);
- $m = 4$, $n = 3$ (biquadratic forms with three variables).

E. Artin in 1927 (cf. [1]) gave a positive answer on the 17th problem of Hilbert: each non-negative rational function can be represented as a sum of squares of rational functions. Using this result one can show that if $P(z)$ is a non-negative polynomial, then there exists a positive polynomial $P_0(z)$ such that $P_1(z) = P(z) * P_0(z)$ can be represented as a sum of squares of polynomials. But we do not know how to find $P_0(z)$ for $P(z)$.

2. The Problem of Minimizing a Quadratic Function on the Non-Negative Orthant

We consider the problem

$$\min_{x \geqslant 0} [(Kx, x) + (c, x)]\,; \quad x, c \in E^n\,, \tag{9}$$

where K is a symmetric matrix of order n. This problem is NP-hard if matrix K is not positive semidefinite (the problem is nonconvex). The dual approach results in the trivial estimate $-\infty$. Let us modify the problem by adding constraints of the form $x_i x_j \geqslant 0$. We showed in [9] that if $c = 0$ then the dual quadratic bound is exact for $n \leqslant 4$.

The problem (9) can be reduced by substitutions $x_i = y_i^2$, $i = 1, \dots, n$, to the problem of global minimization a 4th degree polynomial. Polya [4] proved the following theorem:

If the form $F(x_1, ,x_n)$ is positive for $x \geqslant 0$, $\Sigma_{i=1}^n x_i > 0$, then it can be represented as $F \equiv G/H$, where G and H are forms with positive coefficients. In particular, H can be taken in the form: $H = (x_1 + \cdots + x_n)^P$ for the appropriate P. This result illustrates a consequence of Artin's theorem: $G(x) = F(x)H(x)$. After substitutions $x_i = y_i^2$, $i = 1, \dots, n$, $G(\{y_i^2\}_{i=1}^n)$ is the sum of squares of polynomials.

3. Nonconvex Quadratic Programming

Let us consider a problem: to find the minimum of $K_0(x)$ with respect to linear constraints $l_i(x) \leqslant 0$, $i = 1, \dots, m$. If $K_0(x)$ is nonconvex, then we have the trivial dual bound $\Psi^* = -\infty$. In order to obtain better bounds it is possible to generate the quadratic constraints by multiplying the pairs of linear constraints: from $l_i(x) \leqslant 0$, $l_j(x) \leqslant 0$ we get $l_i(x) l_j(x) \geqslant 0$. Additional quadratic constraints allow us to affect the quadratic part of the Lagrange function. It is proved in [9] that the corresponding dual estimate for the modified quadratic problem is not worse than those which we can obtain by the linearization of the concave part of $K_0(x)$ performed by Pardalos–Rosen [3].

4. The Problem of Finding the Maximum Weight Independent Subset of Vertices in a Graph

Many boolean problems can be reduced to quadratic extremal problems. The dual estimates for quadratic boolean problems was investigated in [6]. Let us consider an interesting example of quadratic dual bounds application to Boolean optimization problem. Let $G(V, E)$ be a nondirected graph with vertex set $V = \{1, \dots, n\}$ and edge set $E = \{(i, j)\}$. A stable (or independent) set in the graph $G(V, E)$ is a set $S \subseteq V$ of vertices, any two of which are not adjacent. Let $w_i > 0$ be the weight of vertex i. The problem can be written as find

$$f^* = \max \sum_{i=1}^{n} w_i x_i \tag{10}$$

subject to

$$x_i x_j \leqslant 1\,, \quad (i, j) \in E\,, \tag{11}$$

$$x_i \in \{0, 1\}\,, \quad i = 1, \dots, n\,. \tag{12}$$

This problem is NP-complete. Usually the linear bounds are replaced by $0 \leqslant x_i \leqslant 1$, $i = 1, \dots, n$, when branch and bound algorithms are applied. By reducing (10)–(12) to a quadratic problem

$$\max \sum_{i=1}^{n} w_i x_i \quad \text{s.t.} \quad x_i x_j = 0\,, \quad (i, j) \in E\,; \quad x_i^2 - x_i = 0\,, \quad i = 1, \dots, n,$$

we can use dual quadratic bounds. Let $\Psi^* = \sup_{u \in \Omega} \Psi(u)$, where

$$\Psi(u) = \inf_x L(x, u) ,$$

$$L(x, u) = -\sum_{i=1}^{n} w_i x_i + \sum_{(i,j)\in E} u_{ij} x_i x_j + \sum_{j=1}^{n} u_j (x_j^2 - x_j) .$$

Lovasz [10], by using specific methods of the theory of coding, established such an upper bound for f^*:

$$\vartheta(G, w) = \max \sum_{(i,j)\in E} \sqrt{w_i w_j} b_{ij} ,$$

where $\{b_{ij}\}_{i,j=1}^{n}$ are symmetric positive semidefinite matrices, $\Sigma_{i=1}^{n} b_{ii} \leqslant 1$ and $b_{ij} = 0$ for $(i, j) \in E$.

This estimate is not worse than the linear one (see [9]). It has been shown in [9] that our estimate satisfies $\Psi^* = -\vartheta(G, w)$. Several numerical experiments have been performed by solving maximum weighted stable set problems. In those experiments the average value of $(|\Psi|^* - f^*)/f^*$ was about 3% (the linear relative error was about 25%). Due to this fact, the number of branches necessary to find the optimum was reduced drastically [9].

5. A Practical Ecological Problem

The practical significance of the problem of refining industrial water sinks is undoubted. One of the important goals in this area is an optimal distribution of costs for water refining among the enterprises whose sinks come to the same river. The mathematical model of the corresponding nonlinear distribution problem is the following: find

$$\mathbf{F}^* = \min \sum_{k=1}^{S} \sum_{i=1}^{N} A_{ki} \mathbf{x}_{ki}^{\lambda_{ki}} ; \quad A_{ki} \geqslant 0 ; \quad 0 \leqslant \lambda_{ki} < 1 , \quad \forall_{ki} \tag{13}$$

$$\mathbf{P}_k(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{x}_{ki} , \quad \text{for } \forall\, k = 1, S \tag{14}$$

$$\mathbf{T}_j(\mathbf{x}) = \sum_{k=1}^{S} \sum_{i=1}^{N} d_{kij} \mathbf{x}_{ki} \leqslant \mathbf{b}_j , \quad \text{for } \forall\; j = 1, M \tag{15}$$

$$0 \leqslant \mathbf{x}_{ki} \leqslant \mathbf{x}_{ki} \leqslant \mathbf{x}_{ki}^{+} \text{ or } \mathbf{x}_{ki} = 0 \quad \text{for } \forall (k, i) \tag{16}$$

(here the index k denotes a sink, **i** is an index of an active technological purification scheme, **j** is an index of admixture, $\mathbf{x}_{ki}$ means the part of water volume of sink **k** that is refining by scheme **i**, and the constraints (15) give an upper limit for the amount of admixtures is control points).

For solving this multiextremal problem we used the method of dual estimates for the constraints (15). The inner problem of determining $\psi(\lambda)$ was solved by the dynamic programming using the constraints (14), (16). The outer problem of $\max \psi(\lambda)$ on $\lambda \geqslant 0$ is solved by a subgradient-type method with dilatation of space in the direction of the difference of two sequential subgradients [5]. The approximate solution vector $\{\mathbf{x}_{ki}\}$ is computed by a special heuristic procedure, after getting λ^*. The corresponding algorithm was implemented by a collaborator of our Institute, Tukalevski S.L. on a personal computer PC-AT 286. His program solves practical-size problems ($k = \overline{1, 44}$; $i = \overline{1, 7}$; $j = \overline{1, 9}$) for about 60 minutes.

6. Final Remarks

The technique of dual quadratic estimates gives us a tool to obtain precise dual estimates for some classes of nonconvex and boolean problems. In other cases we hope to get better lower bounds than the bounds obtained by using the linear programming models. For solving the outer problems maximize $\Psi(\lambda)$, $\lambda \in \Omega$, we used in practical calculations with success the subgradient methods of nondifferentiable optimization with space dilatation [5]. Some details of such experiments are published in [9].

References

1. Artin, E. (1927), Ueber die Zerlegung definiter Funktionen in Quadrate, *Hamburg Abh.* **N5**, 100–15.
2. Hilbert, D. (1888), Ueber die Darstellung definiter Formen als Summen von Formen quadraten, *Mathematical Annals* **22**, 342–350.
3. Pardalos, P. M. and J. B. Rosen (1987), Constrained Global Optimization: Algorithms and Applications, *Lecture Notes in Computer Science* **268**, Springer-Verlag, New York.
4. Polya, G. (1928), Ueber positive Darstellung von Polynomen, *Vierteljahrsschrift d. naturforschenden Gesellsch. Zürich* **73**, 141–145.
5. Shor, N. Z. (1985), *Minimization Methods for Non-Differentiable Functions*, Springer-Verlag, Berlin.
6. Shor, N. Z. (1987), Quadratic Optimization Problems, Izvestia of Academy of Sciences of USSR, *Technical Cybernetics Moscow* **N1**, 128–139 (in Russian).
7. Shor, N. Z. (1987), One Idea of Getting Global Extremum in Polynomial Problems of Mathematical Programming, *Kibernetica* **26** (5) 102–106, (in Russian).
8. Shor, N. Z. (1987), One Class of Estimates for Global Minimum of Polynomial Functions, *Kibernetika* **26** (6), 9–11 (in Russian).
9. Shor, N. Z. and S.I. Stetsenko (1989), *Quadratic Extremal Problems and Nondifferentiable Optimization*, Naukova Dumka, Kiev, 200 pp. (in Russian).
10. Lovasz, L. (1979), On the Shannon Capacity of a Graph, *IEEE Transactions on Information Theory* **25**, 1–7.