

We consider the Euclidean space E^n with vectors $z = \{z_1, z_2, \dots, z_n\}$. We shall understand a polynomial mathematical programming problem to be a problem of the following kind:
Find:

$$\inf P_0(z) \quad (1)$$

under the constraints

$$P_i(z) = 0, \quad i = 1, \dots, m, \quad (2)$$

where $P_0(z)$, $P_i(z)$, $i = 1, \dots, m$ are polynomial functions of z .

The fact that the constraints (2) are given in the form of equalities does not substantially limit the generality of the problem since any polynomial inequality of the form $R(z) = 0$ can be reduced to the polynomial equality $R(z) + t^2 = 0$, where t is an additional variable.

Introducing new variables and using quadratic substitutions of the form $z_i^2 = y_i$, $z_{jk} = z_j z_k$, $y_i^2 = v_i$, etc., the degree of the polynomials in (1) and (2) can be reduced to quadratics by considering them as functions of an expanded set of variables; new quadratic equalities corresponding to the above-mentioned substitutions appear here. Any problem of the form (1)-(2) can therefore be reduced to a quadratic extremal problem: Find

$$\inf_x K_0(x), \quad x \in E^{\bar{n}}, \quad \bar{n} \geq n, \quad (3)$$

under the constraints

$$K_i(x) = 0; \quad i = 1, \dots, \bar{m}; \quad \bar{m} \geq m. \quad (4)$$

where $K_v(x)$, $v = 0, 1, \dots, \bar{m}$ are quadratic functions.

A method is proposed in [1] to obtain the lower bound for the optimal value of the target function and which we shall designate dual since it uses Lagrange multipliers.

Let us consider the Lagrange function: $L(x, u) = K_0(x) + \sum_{i=1}^{\bar{m}} u_i K_i(x) = (A(u)x, x) +$

$(\ell(u), x) + c(u)$. Here $A(u)$ are symmetric $\bar{n} \times \bar{n}$ matrices, $\ell(u)$ are vectors of dimensionality \bar{n} , and $c(u)$ are constants dependent on the vector of Lagrange multipliers $u = \{u_1, \dots, u_{\bar{m}}\}$. Let $\Omega(\bar{\Omega})$ be the set of such values of $u = \{u_1, \dots, u_{\bar{m}}\}$ for which the matrix $A(u)$ is positive definite (non-negative definite). For $u \in \Omega$ $\min_x L(x, u) = \psi(u)$ is achieved at a certain point $x(u)$ that is the solution of the linear system of equations

$$2A(u)x + \ell(u) = 0. \quad (5)$$

Let us determine $\psi(u) = L(x(u), u)$ for $u \in \Omega$. For any allowable \bar{x} , $L(\bar{x}, u) = K_0(\bar{x})$, consequently $\psi(u) \leq K_0(\bar{x})$ for arbitrary u . Hence, $\psi^* = \sup_{u \in \Omega} \psi(u) \leq K_0(\bar{x})$ for arbitrary allowable \bar{x} , i.e., $\psi^* \leq f^*$, where f^* is the optimal value of the target function in the problem (3)-(4) (if the system of equations (4) is incompatible, we will consider that $f^* = +\infty$), $\psi(u)$ is a concave function on $\bar{\Omega}$, and $\bar{\Omega}$ is a convex or empty set. In this latter case we consider $\psi^* = -\infty$.

The problem of finding $\psi^* = \sup_{u \in \Omega} \psi(u)$ refers to convex programming. Sufficiently effective algorithms have been developed for it, in particular good practical results are given by an r-algorithm with special regulation of the step multiplier [2]. Problems of the form

Translated from *Kibernetika*, No. 5, pp. 102-106, September-October, 1987. Original article submitted July 14, 1986.

(1)-(2) and their corresponding (3)-(4) can be non-convex and multiextremal. Consequently, it is impossible to guarantee the equality $\psi^* = f^*$ in the general case. One approach to obtaining a global optimum can be the method of branches and boundaries with utilization of the above-mentioned dual estimates. The algorithm to find the maximal weighted independent set of vertices of a graph, described in [3], can be an example of such an approach. However, the method of branches and boundaries is not sufficiently effective for all problems. It is interesting to investigate those classes of polynomial non-convex problems for which $\psi^* = f^*$. Let us note that when going over from polynomial to quadratic problems different quadratic substitutions can be utilized and ψ^* will vary depending on this. Moreover, new constraints can be generated that are algebraic consequences of the previous ones by keeping the domain of allowable solutions unchanged. The dual estimates do not decrease here but grow in certain cases and increase the chances of agreement between ψ^* and f^* .

It is easy to prove the following theorem.

THEOREM 1. If $\psi^* = \sup_{u \in \Omega} \psi(u)$ is achieved on the set Ω , then $\psi^* = f^*$.

Proof. The gradient of the function $\psi(u)$ for $u \in \Omega$ agrees with the vector of the residual

$$g_\psi(u) = \{K_i(x(u))\}_{i=1}^m.$$

if the maximum $\psi(u)$ is achieved at a certain point $u^* \in \Omega$, then $K_i(x(u^*)) = 0$, $i = 1, \dots, m$, i.e., $x(u^*)$ is an allowable point, here $f^* \leq K_0(x(u^*)) = \psi(u^*) = \psi^*$. But ψ^* is the lower bound for f^* . It thus follows that $f^* = \psi^*$. Q.E.D.

Therefore, if $\psi^* < f^*$, then the supremum of $\psi(u)$ is achieved on the boundary of the domain Ω . Let u^* be a point on $\bar{\Omega} \setminus \Omega$ in any neighborhood of which there are, for an arbitrarily given $\varepsilon > 0$, points $u \in \Omega$ for which $\psi^* - \psi(u) < \varepsilon$ is valid. The rank of the matrix $A(u^*)$ should agree with the rank of the adjoint matrix $(A(u^*) | \varrho(u^*))$, i.e., the system (5) should have a solution for $u = u^*$.

Let us clarify the above by examples.

I. Find $\min [(Ax, x) + (c, x)]$ under the constraints $(x, x) - 1 = 0$, $x \in E^n$.

We reduce the matrix A to diagonal form by an orthogonal transformation. In the new variables $y = \{y_1, \dots, y_n\}$ we obtain the following problem: Find $\min K_0(y)$ where

$$K_0(y) = \sum_{i=1}^n \lambda_i y_i^2 + \sum_{i=1}^n a_i y_i, \quad (6)$$

under the constraints

$$K_1(y) = \sum_{i=1}^n y_i^2 - 1 = 0. \quad (7)$$

Let $\lambda_1, \dots, \lambda_n$ be the eigennumbers of the matrix A written in non-decreasing order with their multiplicity taken into account, and the minimal eigennumber has the multiplicity 1. Let us examine the Lagrange function of the problem (6)-(7)

$$L(y, u) = K_0(y) + u K_1(y).$$

For $u > -\lambda_1$ $L(y, u)$ is positive definite and can be written in the following form

$$L(y, u) = \sum_{i=1}^n (\lambda_i + u) \left(y_i + \frac{a_i}{2(\lambda_i + u)} \right)^2 - \sum_{i=1}^n \frac{a_i^2}{4(\lambda_i + u)} - u.$$

For $u + \lambda_1 > 0$

$$\psi(u) = \min_x L(x, u) = -\frac{1}{4} \sum_{i=1}^n \frac{a_i^2}{\lambda_i + u} - u.$$

If $a_1 \neq 0$, then for $u > -\lambda_1$, $u \rightarrow -\lambda_1$, $\psi(u) \rightarrow -\infty$. Therefore, $\max \psi(u)$ is achieved in the domain of positive-definiteness $-\lambda_1 < u^* < +\infty$ and the global extremum of the problem (6)-(7) corresponds to it. If $a_1 = 0$, then $\max_u \psi(u)$ can be achieved on both the boundary of the

positive-definiteness domain ($u + \lambda_1 = 0$), hence $\psi^* = f^*$. Analogous reasoning holds even in the case of multiple eigenvalues of the matrix A.

II. We consider the problem of minimizing a fourth degree polynomial of one variable

$$\min(x^4 + ax^2 + bx).$$

Let us make the substitution: $x^2 - y = 0$. We obtain the following quadratic problem

$$\min(y^2 + ay + bx)$$

under the constraint $x^2 - y = 0$. The Lagrange function has the following form (for $u > 0$)

$$L(x, y, u) = y^2 + ay + bx + u(x^2 - y) = \left(y + \frac{a-u}{2}\right)^2 + u\left(x + \frac{b}{2u}\right)^2 - \left(\frac{a-u}{2}\right)^2 - \frac{b^2}{4u}.$$

For $u > 0$ (in the positive-definiteness domain in y, x of the Lagrange function) $\psi(u) = -\left(\frac{a-u}{2}\right)^2 - \frac{b^2}{4u}$. For $b \neq 0$, $u \rightarrow +0$, $\psi(u) \rightarrow -\infty$. For $u \rightarrow +\infty$, $\psi(u) \rightarrow -\infty$. This means

$\max_{u>0} \psi(u)$ is achieved at the point u^* , $0 < u^* < +\infty$ of the domain $\Omega: u > 0$ and the unique global optimum $x = -\frac{b}{2u^*}$ will correspond to it. For $b = 0$, $a = 0$, we have $u^* = a$, $\psi(u^*) = 0$, $x^* = 0$. For $b = 0$, $a < 0$, $\sup_{u>0} \psi(u)$ is achieved at a point on the boundary of Ω , $u^* = 0$, here $y^* =$

$x^{*2} = -a/2$, $\psi^* = -a^2/4$, which corresponds to two distinct global minimums $x^* = \pm\sqrt{-a/2}$. Therefore, in all cases $\sup_{u>0} \psi(u) = \psi^* = f^*$.

It turns out that the result obtained for polynomials of fourth degree can be extended to polynomials of arbitrary even power in one variable. Let us consider a polynomial of the even power $2n$, $n \geq 1$, of the variable x_1 for which the coefficient of the highest power equals 1:

$$P_{2n}(x_1) = x_1^{2n} + \sum_{k=1}^{2n} a_{2n-k} x_1^{2n-k}. \quad (8)$$

Let us introduce the notation $x_k = x_1^k$, $k = 0, 1, \dots, n$. In this notation (we note that $x_0 = 1$), the problem of minimizing $P_{2n}(x_1)$ is transformed into a quadratic extremal problem of the following kind:

Find the minimum of

$$K_{2n}(x) = x_n^2 + \sum_{k=1}^{n-1} a_{2n-k} x_n x_{n-k} + \sum_{i=0}^n a_{n-i} x_{n-i} \quad (9)$$

under the constraints

$$\begin{aligned} x_p x_q - x_r x_s &= 0, \quad p + q = r + s \leq 2n - 2, \\ p \geq q, \quad p > r \geq s, \end{aligned} \quad (10)$$

where p, q, r, s are non-negative integers. The expression in the left side of the equality of the form (10) is denoted by $R(p, q; r, s)$ while Q_{2n} denotes the set of different allowable tetraeders.

Remark. Among the equalities of the form (10) part is evidently algebraically redundant; certain equalities are even linear consequences of others, precisely these can be discarded without damage to the further discussion. We knowingly conserve the whole set of expressions $R(p, q; r, s)$ to simplify the notation and the proof.

We compare the Lagrange multiplier $\lambda(p, q; r, s)$ to each equality of the form $R(p, q; r, s) = 0$. The set of Lagrange multipliers forms the vector λ .

Let $x = (x_1, \dots, x_n)$, $x_0 = 1$.

We form the Lagrange function of the quadratic problem (9)-(10):

$$L(x, \lambda) = K_{2n}(x) + \sum_{(p,q;r,s) \in Q_{2n}} \lambda(p, q; r, s) R(p, q; r, s).$$

Let $\Omega(P_{2n})(\bar{\Omega}(P_{2n}))$ be the set of such vectors λ for which the quadratic function $L(x, \lambda)$ in x is positive (non-negative) definite. The function $\psi(\lambda) = \min_x L(x, \lambda)$ is defined for $\lambda \in \Omega(P_{2n})$. If the supremum $\psi(\lambda)$ agrees with $f^* = \min_{x_1} P(x_1)$ for $\lambda \in \bar{\Omega}(P_{2n})$, then we will say that P_{2n} possesses the ω -property. If this supremum is achieved on $\lambda^* \in \Omega(P_{2n})$ then we will say that the polynomial P_{2n} possesses the strong ω -property, here (see Theorem 1), $\psi(\lambda^*)$ agrees with f^* , i.e., P_{2n} possesses the ω -property.

Let us show that the ω -property is conserved under a shift of the origin. Thus the following theorem is valid.

THEOREM 2 (on shift). If $P(x_1)$ possesses the ω -property, then the polynomial $P(x_1) = P(x_1 + a)$ also possesses the ω -property for arbitrary a .

Proof. We introduce the variable $\bar{x}_1 = x_1 - a$. Then $\bar{P}(\bar{x}_1) = P(x_1)$. Therefore, the passage from the polynomial P to \bar{P} is reduced to the substitution: $x_1 = \bar{x}_1 + a$ in $P(x_1)$. If we define $\bar{x}_k = x_1^{-k}$, $k = 1, \dots, n$, then we obtain the following expressions for x_k

$$\bar{x}_k = (\bar{x}_1 + a)^k = \sum_{i=0}^k C_k^i a^i \bar{x}_{k-i}. \quad (11)$$

We need two lemmas to continue the proof.

LEMMA 1. Let there be given a system of equalities of the form

$$y_i - y_j = 0; \quad i, j = 1, \dots, k; \quad i \neq j. \quad (12)$$

Then any linear form $\ell(y) = \sum_{i=1}^n c_i y_i$ under the condition $\sum_{i=1}^n c_i = 0$ can be written in the form of a linear combination of the left sides of (12).

Proof. For $k = 1$ the lemma is trivial. For $k > 1$ it can easily be proved by induction from n to $n + 1$ by using the simple identity

$$\sum_{i=1}^{n+1} c_i y_i = \sum_{i=1}^n c_i y_i + (c_n + c_{n+1}) y_n + c_{n+1} (y_{n+1} - y_n).$$

LEMMA 2. Any expression of the form $x_s x_q - x_p x_r$; $\bar{s} + \bar{q} = \bar{p} + \bar{r}$, where $\bar{s}, \bar{q}, \bar{p}, \bar{r}$ are non-negative integers, go over into expressions of the variables $\bar{x}_0, \bar{x}_1, \dots, \bar{x}_n$ under substitutions of the form (11), which are linear combinations of expressions of the form $x_s x_q - x_p x_r$, $s + q = p + r$; s, q, p, r are non-negative integers.

Proof. Let us use the equalities (11)

$$\begin{aligned} x_s x_q - x_p x_r &= \left(\sum_{i=0}^{\bar{s}} a^i C_s^i \bar{x}_{s-i} \right) \left(\sum_{j=0}^{\bar{q}} a^j C_q^j \bar{x}_{q-j} \right) - \left(\sum_{i=0}^{\bar{p}} a^i C_p^i \bar{x}_{p-i} \right) \left(\sum_{j=0}^{\bar{r}} a^j C_r^j \bar{x}_{r-j} \right) = \\ &= \sum_{t=0}^{\bar{s}+\bar{q}} a^t \left(\sum_{i,j:i+j=t} C_s^i C_q^j \bar{x}_{s-i} \bar{x}_{q-j} \right) - \left(\sum_{i,j:i+j=t} C_p^i C_r^j \bar{x}_{p-i} \bar{x}_{r-j} \right). \end{aligned}$$

As follows from the binomial identities and equalities $\bar{s} + \bar{q} = \bar{p} + \bar{r}$

$$\sum_{i,j:i+j=t} C_s^i C_q^j - \sum_{i,j:i+j=t} C_p^i C_r^j = C_{\bar{s}+\bar{q}}^t - C_{\bar{p}+\bar{r}}^t = 0.$$

Using Lemma 1 repeatedly for different t , we find that the expression $x_s x_q - x_p x_r$ is representable in the form of a linear combination of expressions of the form $x_s x_q - x_p x_r$, $s + q = p + r$; s, q, p, r are non-negative integers. The lemma is proved.

We continue the proof of the theorem.

The polynomial $P(x_1)$ possesses the ω -property. This means that there is a vector of the Lagrange multiplier λ^* such that the Lagrange function $L(x, \lambda)$ of the quadratic problem (9)-(10) for $\lambda = \lambda^*$ is non-negative-definite in x and $\psi(\lambda^*) = f^*$, where $f^* = \min_{x_1} P(x_1)$. In place of components of the vector x let us substitute expressions from (11) into $L(x, \lambda^*)$. (The

formulas (11) can be considered as the passage from one coordinate system to another). We obtain a quadratic function $\bar{L}(\bar{x}, \lambda^*)$ in \bar{x} and since a positive (non-negative) definite quadratic function goes over into a positive (non-negative) definite function under nonsingular coordinate transformations, then $\bar{L}(\bar{x}, \lambda^*)$ will be non-negative definite. Let us note that the domains of the values $L(x, \lambda^*)$ and $\bar{L}(\bar{x}, \lambda^*)$ are in agreement, from which $\bar{\psi}^* = \min_{\bar{x}} \bar{L}(\bar{x}, \lambda^*) = f^*$. On the other hand, by using Lemma 2, we arrive at the deduction that the expression $\bar{L}(\bar{x}, \lambda^*)$ can be written in the form of the Lagrange function of the quadratic problem corresponding to finding $\min \bar{P}(\bar{x}_1)$;

$$\min \bar{K}(\bar{x}) \quad (13)$$

under the constraints

$$\bar{x}_p \bar{x}_q - \bar{x}_r \bar{x}_s = 0, \quad (p, q; r, s) \in Q_{2n}, \quad (14)$$

where $\bar{K}(\bar{x})$ is the result of substituting the expressions (11) into $K_{2n}(x)$ from (9) in place of components of the vector x ; the appropriate Lagrange multipliers are here expressed linearly in terms of components of λ^* . Therefore, for the problem (13) and (14) there exists a vector of the Lagrange multiplier $\bar{\lambda}^*$ for which the appropriate quadratic function of \bar{x} is non-negatively definite and its minimum agrees with f^* by the value of the minimum of the polynomials $P(x_1)$ and $\bar{P}(\bar{x}_1)$. This means that the polynomial \bar{P} also possesses the ω -property. Theorem 2 is proved.

Let us turn to the proof of the main result.

THEOREM 3. Any polynomial $P_{2n}(x_1)$ of the form (8) of even degree possesses the ω -property.

Proof. The ω -property is proved for $n = 2$ (see example 2). We prove the theorem by induction over n . Let the ω -property be valid for polynomials of the form (8) of degree $2n$. We prove its validity for polynomials of degree $2(n+1)$. According to Theorem 2 (on the shift), the ω -property is conserved for a shift in the argument, consequently, it can be assumed without limiting the generality that the global minimum is found at 0 and the value of the polynomial is 0 at 0. It will hence follow that the coefficients of the lowest powers are $a_0 = a_1 = 0$ and $a_2 = 0$ for the polynomial $P_{2n+2}^0(x_1)$ under consideration, i.e., it is representable in the form $P_{2n+2}^0(x_1) = x_1^2 \cdot P_{2n}(x_1)$, where the polynomial $P_{2n}(x_1)$ takes on non-negative values on the whole axis. By the assumption of induction the ω -property is satisfied for $P_{2n}(x_1)$, i.e., for the corresponding quadratic problem a vector is found for the Lagrange multiplier λ^* for which the Lagrange function $L_{2n}(x, \lambda^*)$ will take on non-negative values for any x , i.e., can be written in the form $L_{2n}(x) = \sum_{i=1}^n l_i(x) + r^2$, where $l_i(x)$ are linear functions of (x_1, \dots, x_n) , and r^2 is a non-negative constant.

Let us consider the expression $x_1^2 L_{2n}^*(x)$. The quadratic form

$$L_{2n+2}^*(\bar{x}) = \sum_{i=1}^n \bar{l}_i^2(\bar{x}) + r^2 x_1^2,$$

corresponds to it, where $\bar{l}_i(\bar{x})$ is a linear function of the variables x_1, \dots, x_n, x_{n+1} that is obtained from l_i as follows: if $l_i(x) = \sum_{j=1}^n l_{ij} x_j + l_{i0}$, then $\bar{l}_i(\bar{x}) = \sum_{j=0}^n l_{ij} x_{j+1}$. It is easy

to see that $L_{2n+2}^*(\bar{x})$ is obtained from the Lagrange function of the quadratic problem corresponding to minimization of the polynomial $P_{2n+2}^0(x_1)$ for the same values of the Lagrange multipliers that form the vector λ^* , however these values refer to "transformed" constraints: to constraints of the form $x_p x_q - x_r x_s = 0$ that refer to the problem of minimizing $P_{2n}(x_1)$, will correspond to the constraints $x_{p+1} x_{q+1} - x_{r+1} x_{s+1} = 0$ in the problem of minimizing $P_{2n+2}^0(x_1)$. The minimal value of $L_{2n+2}^*(\bar{x})$ is achieved for $\bar{x} = 0$ and equals 0. Therefore, for definite values of the Lagrange multiplier of the quadratic problem corresponding to minimization of the polynomial $P_{2n+2}^0(x_1)$ an exact estimate is obtained, i.e., the polynomial $P_{2n+2}^0(x_1)$ possesses the ω -property. By the theorem on the shift the arbitrary polynomial $P_{2n+2}(x_1)$ of degree $2n+2$ with coefficient 1 in the highest term will possess this same property. Theorem 3 is proved.

We conclude with some remarks.

1. Reduction of the problem of minimizing $P_{2n}(x_1)$ to a quadratic problem is not unique. This is related to the non-uniqueness of the representation of x_1^k , e.g., $x_1^5 = x_1x_4 = x_2x_3 = x_5$, etc. Depending on the representation taken, the optimal Lagrange multipliers change, however, the ω -property is independent of a specific representation.

2. The minimal number of constraints that must be taken into account in formulating the equivalent quadratic problem for the minimization of $P_{2n}(x_1)$ is $(n - 1)$ since the variables x_2, \dots, x_n must be determined. The remaining constraints are redundant, i.e., their addition does not narrow the domain of allowable values. The role of the redundant constraints is to broaden the number of dual variables of the Lagrange function which generally results in more exact estimates. Addition of constraints, which is a linear combination of the available ones, is not reflected in the accuracy of the dual estimates since the "contribution" of this constraint to the Lagrange function is equivalent to a definite change in the Lagrange multiplier for the available constraints.

3. As a rule, redundancy in the number of constraints results in non-uniqueness of the vector of the optimal Lagrange multiplier λ^* . According to the geometric meaning of the optimal Lagrange multipliers, they are coefficients of the expansion in gradients of the constraints orthogonal to the manifold cutout by constraints antigradient to the target function at the optimal point. Under non-uniqueness of the expansion, the set of allowable vectors of the coefficients form a linear manifold. The intersection of this manifold with $\bar{\Omega}(P)$ indeed yields the set of optimal Lagrange vectors. The assumption that under minimization of the polynomial function of many variables $P(x_1, \dots, x_n)$ by the construction of an equivalent quadratic problem when using a fixed set of redundant constraints, an exact dual estimate is obtained for the value of the global minimum (analog of Theorem 3), is very probable. If this assumption can be proved successfully, then a method based on dual estimates appears in the arsenal of methods to find the global minimum of polynomial functions of several variables. This method can be combined in different forms with the method of branches and boundaries and the method of relaxation of constraints. Of still greater interest is the extension of the theory noted in this paper to the general class of polynomial problems with constraints of the form of (1) and (2). In the general case, can a method be given to generate a redundant system of constraints in order to assure that the dual estimate will turn out to be exact? An experimental investigation of the efficiency of the proposed method of dual estimates goes on at present in a number of classes of extremal problems of graph theory and nonconvex quadratic problems with linear constraints [3].

LITERATURE CITED

1. N. Z. Shor and A. S. Davydov, "On a method to obtain estimates in quadratic extremal problems with Boolean variables," *Kibernetika*, No. 2, 48-50 (1985).
2. N. Z. Shor, *Methods of Minimizing Nondifferentiable Functions and Their Application* [in Russian], Naukova Dumka, Kiev (1979).
3. S. I. Stetsenko, "On algorithms of the solution of certain sampling problems," *Investigations of Methods of Solving Extremal Problems* [in Russian], Inst. Kibern. Akad. Nauk UkrSSR, 12-16 Kiev (1986).