Aplikace matematiky

Ladislav Lukšan

Dual method for solving a special problem of quadratic programming as a subproblem at nonlinear minimax approximation

Aplikace matematiky, Vol. 31 (1986), No. 5, 379–395

Persistent URL: http://dml.cz/dmlcz/104215

31 (1986) APLIKACE MATEMATIKY No. 5, 379—395

DUAL METHOD FOR SOLVING A SPECIAL PROBLEM OF QUADRATIC PROGRAMMING AS A SUBPROBLEM AT NONLINEAR MINIMAX APPROXIMATION

LADISLAV LUKŠAN

(Received June 26, 1985)

Summary. The paper describes the dual method for solving a special problem of quadratic programming as a subproblem at nonlinear minimax approximation. Two cases are analyzed in detail, differring in linear dependence of gradients of the active functions. The complete algorithm of the dual method is presented and its finite step convergence is proved.

Keywords: Nonlinear minimax approximation, quadratic programming, dual method, algorithm.

AMS Classification: 65K05.

1. INTRODUCTION

This papers concerns a special problem of quadratic programming which occurs as a subproblem at nonlinear minimax approximation, where a point $x^* \in R_n$ is sought such that

$$F(x^*) = \min_{x \in R_n} (\max_{i \in M} f_i(x)) \tag{1.1}$$

where $f_i(x)$, $i \in M$ are real-valued functions defined in the n-dimensional vector space R_n, with continuous second-order derivatives, and $M = \{1, \ldots, m\}$. Recently the problem (1.1) has been attracting considerable attention. To solve this problem, several approaches have been developed, especially the least p-th approximation methods [2], [3], steepest descent method [4], variable metric methods making use of properties of the generalized differential [9], recursive linear programming methods [10], [7], recursive quadratic programming methods [8].

The methods of recursive quadratic programming for solving the problems of minimax approximation were developed by analogy with their original application in the field of nonlinear programming. The problem (1.1) can be transformed into the equivalent problem of nonlinear programming, where we seek a pair $(x^*, z^*) \in$ $\in N_{n+1}$ such that

(1.2) $$z^* = \min_{(x,z)\in N_{n+1}} z ,$$

where

$$N_{n+1} = \{(x, z) \in R_{n+1}: f_i(x) \leqq z,\ i \in M\} .$$

Applying the method of recursive quadratic programming [11] to the problem (1.2) we obtain one of the methods described in [8]. This method can be described roughly in the following way.

Step 1. We choose an initial point $x \in R_n$ and an initial symmetric positive definite matrix G. We compute $f_i = f_i(x)$, $i \in M$, $a_i = g_i(x)$, $i \in M$ and $F = F(x) = \max_{i\in M} f_i(x)$ ($g_i(x)$ is the value of the gradient of $f_i(x)$ at the point $x \in R_n$).

Step 2. We find the pair $(s, z) \in R_{n+1}$ which is the solution of the quadratic programming problem (1.3).

Step 3. If $\|s\| \leqq \varepsilon$ where ε is a small enough positive number, the computation is terminated; else we find a steplength α satisfying

$$F(x + \alpha s) \leqq F(x) - \eta\alpha s^{\mathrm{T}} G s ,$$

where $0 < 2\eta < 1$. Taking $x^+ = x + \alpha s$, we compute $f_i^+ = f_i(x^+)$, $i \in M$, $a_i^+ =$ $= g_i(x^+)$, $i \in M$, and $F^+ = F(x^+) = \max_{i\in M} f_i(x^+)$.

Step 4. We transform the matrix G to make it positive definite and an as good approximation as possible to the Hessian matrix of the Lagrangian function for (1.2). Usually this goal is attained by means of quasi-Newton updates determined by using the differences $x^+ - x$, $a_i^+ - a_i$, $i \in M$, and Lagrange multipliers for the quadratic programming problem (1.3).

Step 5. We set $x = x^+$, $f_i = f_i^+$, $i \in M$, $a_i = a_i^+$, $i \in M$, $F = F^+$ and go to Step 2.

The most important step of this method is the solution of the quadratic programming subproblem, in which we seek a pair $(s^*, z^*) \in L_{n+1}$ such that

(1.3) $$\varphi(s^*, z^*) = \min_{(s,z)\in L_{n+1}} \varphi(s, z) ,$$

where

$$\varphi(s, z) = \tfrac{1}{2} s^{\mathrm{T}} G s + z$$

and

$$L_{n+1} = \{(s, z) \in R_{n+1}: f_i + a_i^{\mathrm{T}} s \leqq z,\ i \in M\} .$$

The function $\varphi(s, z)$ can be rewritten in the form

$$\varphi(s, z) = \tfrac{1}{2}[s^{\mathrm{T}}, z] \begin{bmatrix} G, & 0 \\ 0, & 0 \end{bmatrix} \begin{bmatrix} s \\ z \end{bmatrix} + [0, 1] \begin{bmatrix} s \\ z \end{bmatrix}$$

where G is a symmetric positive definite matrix. Now it is apparent that (1.3) is a quadratic programming problem with a singular positive semidefinite matrix. Usual methods solving this problem require nonsingularity. However, owing to the very specific form of the matrix in our case, it is possible to develop methods overcoming the singularity drawback. In this paper, we are going to present one of these methods, which consists in the solution of the dual quadratic programming problem.

The fact that the matrix G is positive definite implies that the problem (1.3) is convex and therefore we may apply the duality theory to it [1]. Thus we obtain a dual quadratic programming problem, where we seek a vector $u^* \in R_m$ such that

$$\psi(u^*) = \min_{u \in L_m} \psi(u)\,, \tag{1.4}$$

where

$$\psi(u) = \tfrac{1}{2}u^{\mathrm{T}}A^{\mathrm{T}}HAu - f^{\mathrm{T}}u$$

and

$$L_m = \{u \in R_m: e^{\mathrm{T}}u = 1,\ u \geqq 0\}\,.$$

Here $A = [a_1, \ldots, a_m]$ is a matrix the columns of which are vectors a_i, $i \in M$, $f = [f_1, \ldots, f_m]^{\mathrm{T}}$, $e = [1, \ldots, 1]^{\mathrm{T}}$, $H = G^{-1}$. The solution of the problem (1.3) can be obtained from the solution of the problem (1.4) by means of

$$\begin{aligned} s^* &= -HAu^*\,, \\ z^* &= f^{\mathrm{T}}u^* - (u^*)^{\mathrm{T}}\,A^{\mathrm{T}}HAu^*\,, \end{aligned} \tag{1.5}$$

which follows from the theory of duality. The vector u^* which is the solution of (1.4), is also the optimal vector of the Lagrange multipliers for (1.3).

The problem (1.4) is a convex one. Hence the vector $u^* \in R_m$ is the solution of (1.4) if and only if the Kuhn-Tucker conditions are valid [1], i.e. if and only if

$$\begin{aligned} e^{\mathrm{T}}u^* &= 1\,, \\ u^* &\geqq 0\,, \end{aligned} \tag{1.6}$$

and there exists a number z^* such that

$$\begin{aligned} v^* = A^{\mathrm{T}}HAu^* - f + z^*e &\geqq 0\,, \\ (v^*)^{\mathrm{T}}\,u^* &= 0\,. \end{aligned} \tag{1.7}$$

The vector v^* is the vector of Lagrange multipliers for (1.4). (1.6) and (1.7) together imply that z^* in (1.7) coincides with the z^* in (1.5). This in turn implies that v^* is, at the same time, the vector of the values of constraints of (1.3).

The dual method for solving (1.3) that is under examination in this paper is essentially the method of active constraints applied to (1.4). In each iteration of this method we start with a feasible point of the problem (1.4), for which $u_i > 0$, $i \in I \subset M$ and $u_i = 0$, $i \in M \setminus I$, and we try to find the optimal point of the problem (1.4) in the subspace defined by the constraints $u_i = 0$, $i \in M \setminus I$. If we meet the boundary of the feasible set, we delete a convenient index from the set $I \subset M$ and the whole process is repeated. If we find the optimal point of the problem (1.4) in the subspace defined by the constraints $u_i = 0$, $i \in M \setminus I$, we test whether the conditions (1.7) are fulfilled.

When this is the case, we terminate the computation, else we add a convenient index to the set $I \subset M$ and the whole process is repeated.

The convergence of the dual method can be guaranteed by choosing individual feasible points of the problem (1.4) so as to make the function $\varphi(s, z)$ monotone increasing. However, one problem appears, namely that of linear dependence of the vectors a_i, $i \in I$. We must discern two cases. In Section 2 we analyse the nonsingular case with linearly independent vectors a_i, $i \in I$ and we generalize these considerations in Section 3. Section 4 contains the complete algorithm of the dual method as well as the proof of its convergence.

This paper was motivated by [6], which provides a description of the dual method for solving the standard quadratic programming problem with a positive definite matrix in the quadratic term.

2. ANALYSIS OF THE NONSINGULAR CASE

Let $I \subset M$. Let $D(I)$ denote the problem which results if we substitute I for M in the problem (1.4). Let $P(I)$ have the similar meaning with respect to the problem (1.3). The problem $D(I)$ is dual to the problem $P(I)$.

In this section, we are dealing with nonsingular case only, where the vectors a_i, $i \in I$ are linearly independent. Let the Lagrange multipliers u_i, $i \in I$ be the solution of the problem $D(I)$ and let the pair $(s, z) \in R_{n+1}$ be the solution of the problem $P(I)$. In order to simplify the notation, we introduce symbols u denoting the vector containing all u_i, $i \in I$, f denoting the vector of f_i, $i \in I$, A denoting the matrix containing a_i, $i \in I$ as its columns, and e denoting the vector containing only units, having the same dimension as u and f. Furthermore, we write

$$
\begin{aligned}
C &= (A^T H A)^{-1}, \\
Q &= H - HACA^T H, \\
p &= Ce,
\end{aligned}
\tag{2.1}
$$

where $H = G^{-1}$. Clearly $QA = 0$ and $QGQ = Q$.

Definition 1.1. *We say that the Lagrange multipliers u_i, $i \in I$ are a basic solution of the problem $D(I)$ and that the pair $(s, z) \in R_{n+1}$ is a basic solution of the problem $P(I)$ if $v_i = a_i^T HAu - f_i + z = 0$ for all indices $i \in I$.*

Lemma 2.1. *Let the Lagrange multipliers u_i, $i \in I$ be a basic solution of the problem $D(I)$ and let the pair $(s, z) \in R_{n+1}$ be a basic solution of the problem $P(I)$. Then*

$$
\begin{aligned}
z &= \frac{p^T f - 1}{p^T e}, \\
u &= C(f - ze), \\
s &= -HAu.
\end{aligned}
\tag{2.2}
$$

Proof. Since $v_i = 0$ for all indices $i \in I$, we have $u = C(f - ze)$ by (1.7). Applying (1.6) we obtain $1 = p^T f - z p^T e$, which means $z = (p^T f - 1)/p^T e$. Mutual duality of $D(I)$ and $P(I)$ brings $s = -HAu$ (see (1.5)). □

The formulae (2.2) may be formally applied to an arbitrary subset $I \subset M$. However, a situation can arise in which $u_i \geqq 0$ does not hold for all indices $i \in I$. In this case there is no basic solution of the problem $D(I)$. Each problem $D(\{k\})$, where $\{k\} \in M$ is a single-element subset of M, has a basic solution, for $u_k = 1$ holds of necessity by (1.6).

Suppose the Lagrange multipliers u_i, $i \in I$ are a basic solution of the problem $D(I)$ and the pair $(s, z) \in R_{n+1}$ is a basic solution of the problem $P(I)$. If $v_i = a_i^T HAu - f_i + z \geqq 0$ for all indices $i \in M \setminus I$, then the vector $u^* = [u^T, 0]^T$ is the solution of the problem (1.4) (assuming a suitable ordering of indices) and the pair $(s^*, z^*) = (s, z)$ is the solution of the problem (1.3). In the other case there exists an index $k \in M \setminus I$ such that $v_k = a_k^T HAu - f_k + z < 0$, which suggests that the index k has to be added to I.

Let us set $I^+ = I \cup \{k\}$. Since the problem $D(I^+)$ need not have any basic solution, we want to find a subset $\bar{I} \subset I^+$, $k \in \bar{I}$, such that the problem $D(\bar{I})$ may have a basic solution and, at the same time, $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$, where $(\bar{s}, \bar{z})$ is a basic solution of the problem $P(\bar{I})$ and $\varphi(s, z)$ is defined by (1.3).

Let $D_\lambda(I^+)$ be the problem we obtain from $D(I^+)$ after substituting $f_i(\lambda) = f_i + (1 - \lambda) v_i$, $i \in I^+$ for f_i, $i \in I^+$ and let $P_\lambda(I^+)$ have the analogous meaning with respect to $P(I^+)$. Let us suppose $0 \leqq \lambda \leqq 1$. Let the Lagrange multipliers $u_i(\lambda)$, $i \in I^+$ be the solution of the problem $D_\lambda(I^+)$ and let $(s(\lambda), z(\lambda)) \in R_{n+1}$ be the solution of the problem $P_\lambda(I^+)$. Let $u(\lambda)$ denote the vector containing $u_i(\lambda)$, $i \in I$.

The Lagrange multipliers $u_i(0) = u_i$, $i \in I$ and $u_k(0) = 0$ are a basic solution of the problem $D_0(I^+)$. We want to find the maximum value of the parameter λ such that the Lagrange multipliers $u_i(\lambda)$, $i \in I^+$ are a basic solution of the problem $D_\lambda(I^+)$.

Lemma 2.2. *Suppose the Lagrange multipliers $u_i(0)$, $u_i(\lambda)$, $i \in I^+$ and the pairs $(s(0), z(0)) \in R_{n+1}$, $(s(\lambda), z(\lambda)) \in R_{n+1}$ are basic solutions of the problems $D_0(I^+)$, $D_\lambda(I^+)$, $P_0(I^+)$, $P_\lambda(I^+)$, respectively. Let us introduce*

(2.3)
$$\begin{aligned} q_k &= CA^T H a_k\,, \\ \beta_k &= 1 - e^T q_k\,, \\ \gamma_k &= \beta_k / p^T e\,, \\ \delta_k &= a_k^T Q a_k = a_k^T H(a_k - A q_k) \end{aligned}$$

assuming $\beta_k \gamma_k + \delta_k \neq 0$. Then

(2.4)
$$\begin{aligned} u(\lambda) &= u(0) - \alpha(q_k + \gamma_k p)\,, \\ u_k(\lambda) &= u_k(0) + \alpha\,, \end{aligned}$$

$$z(\lambda) = z(0) + \alpha\gamma_k ,$$

where

$$\alpha = -\lambda \frac{v_k}{\beta_k\gamma_k + \delta_k} . \tag{2.5}$$

Proof. Using (1.7) we obtain

$$\begin{bmatrix} A^T H A, & A^T H a_k \\ a_k^T H A, & a_k^T H a_k \end{bmatrix} \begin{bmatrix} u(\lambda) - u(0) \\ u_k(\lambda) - u_k(0) \end{bmatrix} = - \begin{bmatrix} (z(\lambda) - z(0))\, e \\ \lambda v_k + (z(\lambda) - z(0)) \end{bmatrix} .$$

Hence, by (2.1) and (2.3)

$$u(\lambda) - u(0) = -(z(\lambda) - z(0))\, p - (u_k(\lambda) - u_k(0))\, q_k \tag{2.6}$$

and

$$a_k^T H A(u(\lambda) - u(0)) + a_k^T H a_k(u_k(\lambda) - u_k(0)) = -\lambda v_k - (z(\lambda) - z(0)) .$$

This equality together with (2.6) gives

$$-q_k^T e(z(\lambda) - z(0)) + a_k^T Q a_k(u_k(\lambda) - u_k(0)) = -\lambda v_k - (z(\lambda) - z(0)) ,$$

and consequently,

$$u_k(\lambda) - u_k(0) = -\frac{\lambda v_k + \beta_k(z(\lambda) - z(0))}{\delta_k} . \tag{2.7}$$

Considering (1.6) we get

$$e^T(u(\lambda) - u(0)) + (u_k(\lambda) - u_k(0)) = 0 ,$$

which, by virtue of (2.6), yields

$$-(z(\lambda) - z(0))\, e^T p + (1 - e^T q_k)(u_k(\lambda) - u_k(0)) = 0 .$$

Hence we obtain

$$z(\lambda) - z(0) = \gamma_k(u_k(\lambda) - u_k(0)) . \tag{2.8}$$

Finally, substituting (2.8) into (2.6) and (2.7), we get (2.4) and (2.5). □

Lemma 2.3. *Let the assumptions of Lemma 2.2 hold. Then*

$$\varphi(s(\lambda), z(\lambda)) = \varphi(s(0), z(0)) + \tfrac{1}{2}\alpha(\beta_k\gamma_k + \delta_k)(u_k(\lambda) + u_k(0)) . \tag{2.9}$$

Proof. Using (1.5) and (2.4) we get

$$\begin{aligned} s(\lambda) - s(0) &= -HA(u(\lambda) - u(0)) - Ha_k(u_k(\lambda) - u_k(0)) = \\ &= \alpha HA(q_k + \gamma_k p) - \alpha H a_k = \alpha(\gamma_k HAp - Qa_k) , \end{aligned}$$

so that

$$\begin{aligned} (s(\lambda) - s(0))^T G\, s(0) &= -\alpha(\gamma_k HAp - Qa_k)^T G(HA\, u(0) + Ha_k\, u_k(0)) = \\ &= -\alpha\gamma_k(1 - u_k(0)) - \alpha\gamma_k e^T q_k\, u_k(0) + \alpha\delta_k\, u_k(0) = \\ &= -\alpha\gamma_k + \alpha(\beta_k\gamma_k + \delta_k)\, u_k(0) \end{aligned}$$

and

$$(s(\lambda) - s(0))^T G(s(\lambda) - s(0)) = \alpha^2(\gamma_k HAp - Qa_k)^T G(\gamma_k HAp - Qa_k) =$$

$$= \alpha^2(\gamma_k^2 e^{\mathrm{T}} p + a_k^{\mathrm{T}} Q a_k) = \alpha^2(\beta_k \gamma_k + \delta_k)$$

since $QA = 0$ and $QGQ = Q$. Thus

$$\begin{aligned}\varphi(s(\lambda), z(\lambda)) &= \varphi(s(0), z(0)) + (s(\lambda) - s(0))^{\mathrm{T}} G\, s(0) + \\ &+ \tfrac{1}{2}(s(\lambda) - s(0))^{\mathrm{T}} G(s(\lambda) - s(0)) + (z(\lambda) - z(0)) = \\ &= \varphi(s(0), z(0)) - \alpha\gamma_k + \alpha(\beta_k\gamma_k + \delta_k)\, u_k(0) + \\ &+ \tfrac{1}{2}\alpha(\beta_k\gamma_k + \delta_k)(u_k(\lambda) - u_k(0)) + \alpha\gamma_k = \\ &= \varphi(s(0), z(0)) + \tfrac{1}{2}\alpha(\beta_k\gamma_k + \delta_k)(u_k(\lambda) + u_k(0))\end{aligned}$$

and the proof is completed. □

The maximum value of the parameter λ for which the problem $D_\lambda(I^+)$ has a basic solution, is determined by the condition $u(\lambda) \geqq 0$. Let us write

$$\alpha_1 = -\frac{v_k}{\beta_k\gamma_k + \delta_k}, \tag{2.10}$$

$$\alpha_2 = \frac{u_j(0)}{q_{kj} + \gamma_k p_j} = \min_{i \in \hat{I}} \frac{u_i(0)}{q_{ki} + \gamma_k p_i}$$

where $\hat{I} = \{i \in I: q_{ki} + \gamma_k p_i > 0\}$, q_{ki} is the i-th component of the vector q_k and p_i is the i-th component of the vector p. Let us set $\alpha = \min(\alpha_1, \alpha_2)$. Then the maximum value λ_0 of the parameter λ is defined as $\lambda_0 = \alpha/\alpha_1$.

When $\alpha = \alpha_1$ (i.e. $\lambda_0 = 1$), the Lagrange multipliers $u_i(\lambda_0)$, $i \in I^+$ are a basic solution of the problem $D(I^+)$ and we can set $\bar{I} = I^+$. If the vectors a_i, $i \in I^+$ are linearly independent (i.e. if $\delta_k \neq 0$) we can construct the matrices $\bar{A} = [A, a_k]$ and $\bar{C} = (\bar{A}^{\mathrm{T}} H \bar{A})^{-1}$. Then

$$\bar{C} = \begin{bmatrix} C + \dfrac{q_k q_k^{\mathrm{T}}}{\delta_k}, & -\dfrac{q_k}{\delta_k} \\ -\dfrac{q_k^{\mathrm{T}}}{\delta_k}, & \dfrac{1}{\delta_k} \end{bmatrix} \tag{2.11}$$

(for the derivation of this formula see for instance [5]). If the vectors a_i, $i \in I^+$ are linearly dependent (i.e. $\delta_k = 0$), the matrix $\bar{C}$ is not defined and we must proceed in the manner described in Section 3.

When $\alpha \neq \alpha_1$ (i.e. $\lambda_0 < 1$), we have $u_j(\lambda_0) = 0$ by (2.10). Let us set $I_1 = I \setminus \{j\}$, $I_1^+ = I^+ \setminus \{j\}$, and $v_i^{(1)} = (1 - \lambda_0)\, v_i$, $i \in I_1^+$. Let $D_\lambda(I_1^+)$ denote the problem resulting from the problem $D(I_1^+)$ after substituting $f_i(\lambda) = f_i + (1 - \lambda)\, v_i^{(1)}$, $i \in I_1^+$ (where $0 \leqq \lambda \leqq 1$) for f_i, $i \in I_1^+$. The Lagrange multipliers $u_i^{(1)}(0) = u_i(\lambda_0)$, $i \in I_1^+$ are a basic solution of the problem $D_0(I_1^+)$. Again we want to find the maximum value of the parameter λ, for which the Lagrange multipliers $u_i^{(1)}(\lambda)$, $i \in I_1^+$ are a basic solution of the problem $D_\lambda(I_1^+)$. For this purpose we can apply the preceding process (Lemma 2.2 and Lemma 2.3). except that instead of the values referring to the problem $D(I^+)$ we use the values referring to the problem $D(I_1^+)$. Especially, the matrices A,

C are to be replaced by the matrices A_1, C_1, respectively, such that $A_1 = A^{(j)}$, which is the matrix A with the j-th column removed, and

$$C_1 = C^{(jj)} - \frac{C_j^{(j)}(C_j^{(j)})^{\mathrm{T}}}{C_{jj}}, \tag{2.12}$$

where $C^{(jj)}$ results from C by removing the j-th row and the j-th column, $C_j^{(j)}$ results from the j-th column of C by removing the element C_{jj}. (For the derivation of the formula (2.12) see e.g. [5]).

Suppose $\lambda_0^{(1)} = \alpha^{(1)}/\alpha_1^{(1)}$ is the maximum value of the parameter λ, for which the problem $D_\lambda(I_1^+)$ has a basic solution. If $\lambda_0^{(1)} = 1$, we set $\bar{I} = I_1^+$, else we repeat the whole process. In this manner we obtain a sequence $I_1^+, \ldots, I_p^+$ of subsets of the set I^+. The cardinality of each of these subsets is by one element less than the cardinality of its precursor. But the set I^+ is finite and the problem $D(\{k\})$ has a basic solution, therefore we obtain, after a finite number of steps, a subset $I_p^+ \subset I^+$, $k \in I_p^+$ such that the problem $D(I_p^+)$ has a basic solution. Thus we can set $\bar{I} = I_p^+$.

So far we have been treating the case $\beta_k\gamma_k + \delta_k \neq 0$. Now let us suppose $\beta_k\gamma_k + \delta_k = 0$. In this case, there exists no nonzero value of the parameter λ such that the problem $D_\lambda(I^+)$ has a basic solution. On the other hand, the problem $D_0(I^+)$ has more basic solutions that are defined by the equations.

$$\begin{aligned} u(\alpha) &= u(0) - \alpha(q_k + \gamma_k p), \\ u_k(\alpha) &= u_k(0) + \alpha. \end{aligned} \tag{2.13}$$

The condition $\beta_k\gamma_k + \delta_k = 0$ is valid only if $\beta_k = 0$, $\gamma_k = 0$ and $\delta_k = 0$ (this is implied by the facts that C is positive definite, Q is positive semidefinite, and by (2.3)). We have immediately $z(\alpha) = z(0)$ and $s(\alpha) = s(0)$. The problem $P_0(I^+)$ has a unique solution $(s(\alpha), z(\alpha)) = (s(0), z(0)) \in R_{n+1}$ so that $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0))$ for an arbitrary value of the parameter α.

The Lagrange multipliers $u_i(\alpha)$, $i \in I^+$ are a basic solution of the problem $D_0(I^+)$ only if $u(\alpha) \geqq 0$.

Lemma 2.4. *There exists a finite maximum value of the parameter α in (2.13) for which the Lagrange multipliers $u_i(\alpha)$, $i \in I^+$ are a basic solution of the problem $D_0(I^+)$.*

Proof. Using (2.3) we get

$$e^{\mathrm{T}}(q_k + \gamma_k p) = e^{\mathrm{T}}q_k + \frac{1 - e^{\mathrm{T}}q_k}{e^{\mathrm{T}}p}\, e^{\mathrm{T}}p = e^{\mathrm{T}}q_k + 1 - e^{\mathrm{T}}q_k = 1.$$

Therefore there exists at least one index $i \in I$ such that $q_{ki} + \gamma_k p_i > 0$. Hence necessarily $\alpha \leqq \alpha_2$, where α_2 is a finite value determined by (2.10). □

When we choose $\alpha = \alpha_2$, then $u_j(\alpha) = 0$ holds for some index $j \in I$. Let us set $I_1 = I \setminus \{j\}$, $I_1^+ = I^+ \setminus \{j\}$, and $v_i^{(1)} = v_i$, $i \in I_1^+$. Let $D_\lambda(I_1^+)$ denote the problem $D(I_1^+)$ after substituting $f_i(\lambda) = f_i + (1 - \lambda)\, v_i^{(1)}$, $i \in I_1^+$ (where $0 \leqq \lambda \leqq 1$) for

f_i, $i \in I_1^+$. Then the Lagrange multipliers $u_i^{(1)}(0) = u_i(\alpha)$, $i \in I_1^+$ are a basic solution of the problem $D_0(I_1^+)$. Thus we can proceed in the same way as we did in the case where $\beta_k\gamma_k + \delta_k \neq 0$ and $\alpha \neq \alpha_1$, only for $\beta_k\gamma_k + \delta_k = 0$ we formally set $\alpha_1 = \infty$ in (2.10).

Let $\bar{I}$ be a set we have obtained by the process described in this section. It remains to prove that $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$. where $(\bar{s}, \bar{z}) \in R_{n+1}$ is the solution of the problem $P(\bar{I})$.

Theorem 2.1. *Suppose $(s, z) \in R_{n+1}$ is the solution of the problem $P(I)$ and $(\bar{s}, \bar{z}) \in \in R_{n+1}$ is the solution of the problem $P(\bar{I})$. Then $\varphi(\bar{s},\bar{z}) > \varphi(s, z)$.*

Proof. The set $\bar{I}$ is obtained after a finite number of steps, in which we construct subsets $\bar{I} = I_p^+ \subset \ldots \subset I_1^+ \subset I^+$. Since all steps are formally equivalent, it suffices to analyse the first step. Let $(s(0), z(0)) \in R_{n+1}$ be the solution of the problem $P_0(I^+)$ and $(s(\alpha), z(\alpha)) \in R_{n+1}$ the solution of the problem $P_0(I_1^+)$. Two cases are possible. If $\beta_k\gamma_k + \delta_k = 0$, then $s(\alpha) = s(0)$ and $z(\alpha) = z(0)$, so that $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0))$. If $\beta_k\gamma_k + \delta_k \neq 0$, we get, by (2.4) and (2.9),

$$\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0)) + \tfrac{1}{2}\alpha(\beta_k\gamma_k + \delta_k)(2u_k(0) + \alpha) .$$

But $\beta_k\gamma_k + \delta_k > 0$ (because $\beta_k\gamma_k + \delta_k \geqq 0$ and $\beta_k\gamma_k + \delta_k \neq 0$), $u_k(0) \geqq 0$ and $\alpha \geqq 0$, so that $\varphi(s(\alpha), z(\alpha)) \geqq \varphi(s(0), z(0))$ and $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0))$ holds if and only if $\alpha = 0$. Combining both cases we obtain

$$\varphi(\alpha) \geqq \varphi(0) , \tag{2.14a}$$

$$\varphi(\alpha) = \varphi(0) \Leftrightarrow s(\alpha) = s(0) , \quad z(\alpha) = z(0) , \tag{2.14b}$$

where $\varphi(\alpha) = \varphi(s(\alpha), z(\alpha))$ and $\varphi(0) = \varphi(s(0), z(0))$. Now we will prove that $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$. The validity of (2.14a) in each step yields $\varphi(\bar{s}, \bar{z}) \geqq \varphi(s, z)$. Now let us suppose $\varphi(\bar{s}, \bar{z}) = \varphi(s, z)$. Since (2.14b) is valid in each step, we have $\bar{s} = s$ and $\bar{z} = z$. Therefore

$$\bar{v}_k = \bar{z} - a_k^{\mathrm{T}}\bar{s} - f_k = z - a_k^{\mathrm{T}}s - f_k = v_k < 0 ,$$

which is a contradiction, for $k \in \bar{I}$ and $(\bar{s}, \bar{z}) \in R_{n+1}$ is a basic solution of the problem $P(\bar{I})$, and consequently, $\bar{v}_k = 0$. □

3. ANALYSIS OF THE SINGULAR CASE

Supposing $I \subset M$, let $J \subset I$ be the maximum subset of the set I such that the vectors a_i, $i \in J$ are linearly independent. In the singular case we have $J \neq I$. Let the Lagrange multipliers u_i, $i \in I$ be a basic solution of the problem $D(I)$ and the pair $(s, z) \in R_{n+1}$ be the basic solution of the problem $P(I)$. Let u denote the vector containing u_i, $i \in J$, let f denote the vector containing elements f_i, $i \in J$ and let A denote the matrix containing vectors a_i, $i \in J$ as its columns. Let e be the vector containing only units and let C, Q be the matrices defined by (2.1)

In the singular case, we proceed in the similar way as in the nonsingular one. Again we seek a subset $\bar{I} \subset I^+$, $k \in \bar{I}$ such that the problem $P(\bar{I})$ has a basic solution $(\bar{s}, \bar{z}) \in R_{n+1}$ together with $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$ $(I^+ = I \cup \{k\}$, where $k \in M \setminus I$ and $v_k = a_k^{\mathrm{T}}HA - f_k + z < 0)$. However, different formulae are needed for the computation of the Lagrange multipliers.

Let us suppose that $I = J \cup \{l\}$ (we are dealing with only those cases where the subset $I \setminus J$ has at most one element). Let $D_\lambda(I^+)$, $P_\lambda(I^+)$ be the problems defined in Section 2. The Lagrange multipliers $u_i(0) = u_i$, $i \in I$ and $u_k(0)$ are a basic solution of the problem $D_0(I^+)$. We are seeking the maximum value of the parameter λ for which the Lagrange multipliers $u_i(\lambda)$, $i \in I^+$ are a basic solution of the problem $D_\lambda(I^+)$.

Lemma 3.1. *Suppose* $u_i(0)$, $i \in I^+$, $(s(0), z(0)) \in R_{n+1}$, $u_i(\lambda)$, $i \in I^+$, $(s(\lambda), z(\lambda)) \in R_{n+1}$ *are basic solutions of the problems* $D_0(I^+)$, $P_0(I^+)$, $D_\lambda(I^+)$, $P_\lambda(I^+)$, *respectively. Let us introduce*

$$(3.1)\qquad \begin{aligned} q_l &= CA^{\mathrm{T}}Ha_l\,, \\ q_k &= CA^{\mathrm{T}}Ha_k\,, \\ \beta_l &= 1 - e^{\mathrm{T}}q_l\,, \\ \beta_k &= 1 - e^{\mathrm{T}}q_k\,, \\ \delta_k &= a_k^{\mathrm{T}}Qa_k = a_k^{\mathrm{T}}H(a_k - Aq_k) \end{aligned}$$

and suppose $\beta_l \neq 0$, $\delta_k \neq 0$. Then

$$(3.2)\qquad \begin{aligned} u(\lambda) &= u(0) - \alpha\left(q_k - \frac{\beta_k}{\beta_l}\, q_l\right), \\ u_l(\lambda) &= u_l(0) - \alpha\frac{\beta_k}{\beta_l}\,, \\ u_k(\lambda) &= u_k(0) + \alpha\,, \\ z(\lambda) &= z(0)\,, \end{aligned}$$

where

$$(3.3)\qquad \alpha = -\lambda\frac{v_k}{\delta_k}\,.$$

Proof. Using (1.7) we get

$$\begin{bmatrix} A^{\mathrm{T}}HA, & A^{\mathrm{T}}Ha_l, & A^{\mathrm{T}}Ha_k \\ a_l^{\mathrm{T}}HA, & a_l^{\mathrm{T}}Ha_l, & a_l^{\mathrm{T}}Ha_k \\ a_k^{\mathrm{T}}HA, & a_k^{\mathrm{T}}Ha_l, & a_k^{\mathrm{T}}Ha_k \end{bmatrix} \begin{bmatrix} u(\lambda) - u(0) \\ u_l(\lambda) - u_l(0) \\ u_k(\lambda) - u_k(0) \end{bmatrix} = - \begin{bmatrix} (z(\lambda) - z(0))\, e \\ z(\lambda) - z(0) \\ z(\lambda) - z(0) + \lambda v_k \end{bmatrix}$$

so that, by (2.1) and (3.1),

$$u(\lambda) - u(0) = -(z(\lambda) - z(0))\, p - (u_l(\lambda) - u_l(0))\, q_l - (u_k(\lambda) - u_k(0))\, q_k$$

and

$$\beta_l(z(\lambda) - z(0)) + a_l^{\mathrm{T}} Q a_l (u_l(\lambda) - u_l(0)) + a_l^{\mathrm{T}} Q a_k (u_k(\lambda) - u_k(0)) = 0\,,$$
$$\beta_k(z(\lambda) - z(0)) + a_k^{\mathrm{T}} Q a_l (u_l(\lambda) - u_l(0)) + a_k^{\mathrm{T}} Q a_k (u_k(\lambda) - u_k(0)) = -\lambda v_k\,.$$

Since the vector a_l is a linear combination of the vectors a_i, $i \in J$, we get $Qa_l = 0$. Thus

$$z(\lambda) - z(0) = 0\,,$$

$$u_k(\lambda) - u_k(0) = -\lambda \frac{v_k}{a_k^{\mathrm{T}} Q a_k} = \alpha$$

and

(3.4) $$u(\lambda) - u(0) = -(u_l(\lambda) - u_l(0))\, q_l - (u_k(\lambda) - u_k(0))\, q_k\,.$$

Using (1.6) we get

$$e^{\mathrm{T}}(u(\lambda) - u(0)) = -(u_l(\lambda) - u_l(0)) - (u_k(\lambda) - u_k(0))\,,$$

which together with (3.4) yields

$$u_l(\lambda) - u_l(0) = -\frac{\beta_k}{\beta_l}(u_k(\lambda) - u_k(0)) = -\frac{\beta_k}{\beta_l}\alpha\,.$$

When substituting the last formula into (3.4), we finally obtain $u(\lambda) - u(0) = -\left(q_k - \frac{\beta_k}{\beta_l} q_l\right)\alpha.$ □

Lemma 3.2. *Let us suppose that all suppositions of lemma* 3.1 *hold. Then*

(3.5) $$\varphi(s(\lambda), z(\lambda)) = \varphi(s(0), z(0)) + \tfrac{1}{2}\alpha\, \delta_k (u_k(\lambda) - u_k(0))\,.$$

Proof. Using (1.5) and (3.2) we get

$$s(\lambda) - s(0) = -HA(u(\lambda) - u(0)) - Ha_l(u_l(\lambda) - u_l(0)) - Ha_k(u_k(\lambda) - u_k(0)) =$$
$$= \alpha HA\left(q_k - \frac{\beta_k}{\beta_l} q_l\right) - \alpha H\left(a_k - \frac{\beta_k}{\beta_l} a_l\right) = -\alpha\left(Qa_k - \frac{\beta_k}{\beta_l} Qa_l\right) = -\alpha Q a_k\,,$$

so that

$$(s(\lambda) - s(0))^{\mathrm{T}}\, G\, s(0) = \alpha a_k^{\mathrm{T}} Q G(HA\, u(0) + Ha_l\, u_l(0) + Ha_k\, u_k(0)) =$$
$$= \alpha a_k^{\mathrm{T}} Q a_k\, u_k(0) = \alpha \delta_k\, u_k(0)$$

and

$$(s(\lambda) - s(0))^{\mathrm{T}}\, G(s(\lambda) - s(0)) = \alpha^2 a_k^{\mathrm{T}} Q G Q a_k = \alpha^2 \delta_k$$

for $QA = 0$, $Qa_l = 0$, and $QGQ = Q$. Thus we obtain

$$\varphi(s(\lambda) - z(\lambda)) = \varphi(s(0), z(0)) + (s(\lambda) - s(0))^{\mathrm{T}}\, G\, s(0) +$$
$$+ \tfrac{1}{2}(s(\lambda) - s(0))^{\mathrm{T}}\, G(s(\lambda) - s(0)) + (z(\lambda) - z(0)) =$$
$$= \varphi(s(0), z(0)) + \alpha\delta_k\, u_k(0) + \tfrac{1}{2}\alpha\delta_k(u_k(\lambda) - u_k(0)) =$$
$$= \varphi(s(0), z(0)) + \tfrac{1}{2}\alpha\delta_k(u_k(\lambda) + u_k(0))\,.$$ □

The maximum value of the parameter λ for which the problem $D_\lambda(I^+)$ has a basic solution is defined by the condition $u(\lambda) \geqq 0$, $u_l(\lambda) \geqq 0$. Let us write $w_i = q_{ki} - (\beta_k/\beta_l)\, q_{li}$, $i \in J$ and $w_l = (\beta_k/\beta_l)$ (q_{ki} is the i-th component of the vector q_k and q_{li} is the i-th component of the vector q_l). Furthermore, let us write

(3.6)
$$\alpha_1 = -\frac{v_k}{\delta_k},$$

$$\alpha_2 = \frac{u_j(0)}{w_j} = \min_{i \in \hat{I}} \left(\frac{u_i(0)}{w_i} \right),$$

where $\hat{I} = \{i \in I : w_i > 0\}$. Let us set $\alpha = \min(\alpha_1, \alpha_2)$. Then the maximum value λ_0 of the parameter λ is defined by $\lambda_0 = \alpha/\alpha_1$.

If $\alpha = \alpha_1$ (i.e. if $\lambda_0 = 1$), then the Lagrange multipliers $u_i(\lambda_0)$, $i \in I^+$ are a basic solution, of the problem $D(I^+)$ and we can set $\bar{I} = I^+$. Since by assumption $\delta_k \neq 0$, the vector a_k is not a linear combination of the vectors a_i, $i \in J$, so that we can set $\bar{J} = J \cup \{k\}$ and construct the matrices $\bar{A} = [A, a_k]$, $\bar{C} = (\bar{A}^T H \bar{A})^{-1}$ (the matrix $\bar{C}$ is computed according to (2.11)).

If $\alpha \neq \alpha_1$ (i.e. $\lambda_0 < 1$), then, by (3.6), $u_j(\lambda_0) = 0$. If we set $J_1 = J \setminus \{j\}$, $I_1 = I \setminus \{j\}$, $I_1^+ = I^+ \setminus \{j\}$, and $v_i^{(1)} = (1 - \lambda_0)\, v_i$, $i \in I_1^+$, then the Lagrange multipliers $u_i^{(1)}(0) = u_i(\lambda_0)$, $i \in I_1^+$ are a basic solution of the problem $D_0(I_1^+)$ and we can repeat the whole process (the matrices A_1, C_1 are defined in the same way as in the nonsingular case). However, two cases are possible.

Lemma 3.3. *The elements q_{li}, $i \in J$ are the uniquely determined coefficients of linear dependence of the vector a_l on the vectors a_i, $i \in J$.*

Proof. $Qa_l = H(a_l - Aq_l) = 0$ implies

(3.7)
$$a_l = Aq_l .$$

Since the vectors a_i, $i \in J$ are linearly independent, the expression (3.7) is uniquely determined. □

If $q_{lj} = 0$, the vector a_l is a linear combination of the vectors a_i, $i \in J_1$. In this case we have the set J_1 as well as the matrices A_1, C_1 unchanged. If $q_{lj} \neq 0$, the vector a_l is not a linear combination of the vectors a_i, $i \in J_1$. In this case we add the index l to the set J_1 and reconstruct the matrices A_1 and C_1 (thus obtaining a nonsingular case).

So far we have been treating the case with $\delta_k \neq 0$. Now let us suppose $\delta_k = 0$. In this case there exists no nonzero value of the parameter λ such that the problem $D_\lambda(I^+)$ has a basic solution. On the other hand, the problem $D_0(I^+)$ has more basic solutions that are defined by the equations

(3.8)
$$u(\alpha) = u(0) - \alpha\left(q_k - \frac{\beta_k}{\beta_l}\, q_l\right),$$

$$u_l(\alpha) = u_l(0) - \alpha \frac{\beta_k}{\beta_l},$$

$$u_k(\alpha) = u_k(0) + \alpha\,,$$

and $z(\alpha) = z(0)$, $s(\alpha) = s(0)$ hold (for $s(\alpha) = s(0) - \alpha Q a_k$, where $Q a_k = 0$ because Q is positive semidefinite and $a_k^T Q a_k = \delta_k = 0$). Hence the problem $P_0(I^+)$ has only one solution $(s(\alpha), z(\alpha)) = (s(0), z(0)) \in R_{n+1}$, so that $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0))$ for an arbitrary value of α.

The Lagrange multipliers $u_i(\alpha)$, $i \in I^+$ are a basic solution of the problem $D_0(I^+)$ only if $u(\alpha) \geqq 0$ and $u_l(\alpha) \geqq 0$.

Lemma 3.4. *There exists a finite maximum value of the parameter α in (3.8) for which the Lagrange multipliers $u_i(\alpha)$, $i \in I^+$ are a basic solution of the problem $D_0(I^+)$.*

Proof. Denoting

$$w = q_k - \frac{\beta_k}{\beta_l} q_l \quad \text{and} \quad w_l = \frac{\beta_k}{\beta_l}$$

and using (3.1) we obtain

$$e^T w + w_l = e^T q_k - \frac{\beta_k}{\beta_l} e^T q_l + \frac{\beta_k}{\beta_l} = (1 - \beta_k) - \frac{\beta_k}{\beta_l}(1 - \beta_l) + \frac{\beta_k}{\beta_l} = 1\,.$$

Since $e^T w + w_l = 1$, there exists at least one index $i \in I$ such that $w_i > 0$. Hence of necessity $\alpha \leqq \alpha_2$, where α_2 is a finite value defined by (3.6). □

If we choose $\alpha = \alpha_2$, then $u_j(\alpha) = 0$ is valid for some $j \in I$. Setting $J_1 = J \setminus \{j\}$, $I_1 = I \setminus \{j\}$, $I_1^+ = I^+ \setminus \{j\}$, and $v_i^{(1)} = v_i$, $i \in I_1^+$, the Lagrange multipliers $u_i^{(1)}(0) = u_i(\alpha)$, $i \in I_1^+$ are a basic solution of the problem $D_0(I_1^+)$. Therefore we can proceed in the same manner as if $\delta_k \neq 0$, $\alpha \neq \alpha_1$, except that for $\delta_k = 0$ we formally set $\alpha_1 = \infty$ in (3.6).

We have shown that $\alpha \neq \alpha_1$ provided $\delta_k = 0$, and thereby we have demonstrated the correctness of our assumption that the set $I \setminus J$ has at most one element.

It remains to prove that $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$, where $(\bar{s}, \bar{z}) \in R_{n+1}$ is a solution of the problem $P(\bar{I})$.

Theorem 3.1. *Let $(s, z) \in R_{n+1}$, $(\bar{s}, \bar{z}) \in R_{n+1}$ be the solutions of the problems $P(I)$, $P(\bar{I})$, respectively. Then $\varphi(\bar{s}, \bar{z}) > \varphi(s, z)$.*

Proof. The set $\bar{I}$ results after a finite number of steps in which we construct the subsets $\bar{I} = I_p^+ \subset \ldots \subset I_1^+ \subset I^+$. We will prove that

$$\varphi(\alpha) \geqq \varphi(0)\,, \tag{3.9a}$$

$$\varphi(\alpha) = \varphi(0) \Leftrightarrow s(\alpha) = s(0)\,, \quad z(\alpha) = z(0)\,, \tag{3.9b}$$

where $\varphi(0) = \varphi(s(0), z(0))$ is the value of the function (1.3) at the beginning of a current step and $\varphi(\alpha) = \varphi(s(\alpha), z(\alpha))$ is the value of the function (1.3) at the end of a current step. (3.9) has been already proved (Section 2, (2.14)) for those steps in which the nonsingular case occurs. Since all steps that involve the singular case are formally identical, it suffices to analyze the first one. Two cases are possible. If $\delta_k = 0$, then $s(\alpha) = s(0)$ and $z(\alpha) = z(0)$ so that $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0))$. If $\delta_k \neq 0$, we have, by (3.2) and (3.5), $\varphi(s(\alpha), z(\alpha)) = \varphi(s(0), z(0)) + \frac{1}{2}\alpha\delta_k(2u_k(0) + \alpha)$. But $\delta_k > 0$ (since $\delta_k \geqq 0$ and $\delta_k \neq 0$), $u_k(0) \geqq 0$, $\alpha \geqq 0$. Thus $\varphi(s(\alpha), z(\alpha)) \geqq \varphi(s(0), z(0))$ and the equality is valid if and only if $\alpha = 0$. Combining the both cases we obtain (3.9) and proceed in the same manner as in the proof of Theorem 2.1. □

4. ALGORITHM OF THE DUAL METHOD

In Section 2 and Section 3, we have described the construction of the principal step of the dual method for solving the problem (1.3). Now we will describe the algorithm that contains these major steps.

Algorithm 4.1.

Step 1. Choose arbitrarily an index $k \in M$.

Step 2. Set $I := \{k\}$, $J := \{k\}$, $u := [1]$, $e := [1]$, $A := [a_k]$, $C := [1/a_k^{\mathrm{T}} H a_k]$ and calculate $z := f_k - a_k^{\mathrm{T}} H a_k$.

Step 3. Calculate

$$s := -H \sum_{i \in I} a_i u_i$$

and

$$v_k := z - f_k - a_k^{\mathrm{T}} s = \min_{i \in M \setminus I} (z - f_i - a_i^{\mathrm{T}} s) .$$

If $v_k \geqq 0$, terminate ($(s, z) \in R_{n+1}$ is the solution of the problem (1.3)). If $v_k < 0$, set $u_k := 0$ and go to Step 4 provided $J = I$, else go to Step 8 provided $J \neq I$.

Step 4. (Nonsingular case.) Calculate $p := Ce$, $q_k := CA^{\mathrm{T}} H a_k$, $\beta_k := 1 - e^{\mathrm{T}} q_k$, $\gamma_k := \beta_k / e^{\mathrm{T}} p$, and $\delta_k := a_k^{\mathrm{T}} H(a_k - A q_k)$. If $\beta_k \gamma_k + \delta_k = 0$, set $\alpha_1 := \infty$, else set

$$\alpha_1 := -\frac{v_k}{\beta_k \gamma_k + \delta_k} .$$

Calculate

$$\alpha_2 := \frac{u_j}{q_{kj} + \gamma_k p_j} = \min_{i \in \hat{I}} \left(\frac{u_i}{q_{ki} + \gamma_k p_i} \right),$$

where $\hat{I} = \{ i \in I : q_{ki} + \gamma_k p_i > 0 \}$. Set $\alpha := \min(\alpha_1, \alpha_2)$ and calculate $u := u - \alpha(q_k + \gamma_k p)$, $u_k := u_k + \alpha$, $z := z + \alpha \gamma_k$, $v_k := (1 - \alpha/\alpha_1) v_k$. If $\alpha = \alpha_1$, go to Step 5, else go to Step 6.

Step 5. Set $I := I \cup \{k\}$. If $\delta_k = 0$, set $l := k$ and go to Step 3. If $\delta_k \neq 0$ set $J := J \cup \{k\}$, $u := [u^{\mathrm{T}}, u_k]^{\mathrm{T}}$, $e := [e^{\mathrm{T}}, 1]^{\mathrm{T}}$, $A := [A, a_k]$ and

$$C := \begin{bmatrix} C + \dfrac{q_k q_k^{\mathrm{T}}}{\delta_k}, & -\dfrac{q_k}{\delta_k} \\ -\dfrac{q_k^{\mathrm{T}}}{\delta_k}, & \dfrac{1}{\delta_k} \end{bmatrix}$$

and go to Step 3.

Step 6. If the set J contains one element and $J = I$, then go to Step 2. If J contains one element and $J \neq I$, set $I := \{l\}$, $J := \{l\}$, $u := [u_l]$, $e := [1]$, $A := [a_l]$, $C := [1/a_l^{\mathrm{T}} H a_l]$ and go to Step 4. If J does not contain one element, set $I := I \setminus \{j\}$, $J := J \setminus \{j\}$, $u := u^{(j)}$, $e := e^{(j)}$ and $A := A^{(j)}$, where $u^{(j)}$, $e^{(j)}$ result from u, e by deleting the elements u_j, e_j, respectively, and $A^{(j)}$ results from A by deleting the column a_j. Then set

$$C := C^{(jj)} - \frac{C_j^{(j)} (C_j^{(j)})^{\mathrm{T}}}{C_{jj}},$$

where $C^{(jj)}$ results from C by deleting both the j-th row and the j-th column, $C_j^{(j)}$ results from the j-th column of C by deleting the component C_{jj}. If $J = I$ go to Step 4, else go to Step 7.

Step 7. If $q_{lj} = 0$, set $q_l := q_l^{(j)}$, where $q_l^{(j)}$ results from q_l by deleting the element q_{lj}, and go to Step 8. If $q_{lj} \neq 0$, calculate $q_l := CA^{\mathrm{T}} H a_l$ and $\delta_l := a_l^{\mathrm{T}} H(a_l - Aq_l)$. Set $J := J \cup \{l\}$, $u := [u^{\mathrm{T}}, u_l]^{\mathrm{T}}$, $e := [e^{\mathrm{T}}, 1]^{\mathrm{T}}$, $A := [A, a_l]$ and

$$C := \begin{bmatrix} C + \dfrac{q_l q_l^{\mathrm{T}}}{\delta_l}, & -\dfrac{q_l}{\delta_l} \\ -\dfrac{q_l^{\mathrm{T}}}{\delta_l}, & \dfrac{1}{\delta_l} \end{bmatrix},$$

and go to Step 4.

Step 8. (Singular case.) Calculate $q_k := CA^{\mathrm{T}} H a_k$, $\beta_k := 1 - e^{\mathrm{T}} q_k$, and $\delta_k := a_k^{\mathrm{T}} H(a_k - Aq_k)$. If $\delta_k = 0$, set $\alpha_1 := \infty$. If $\delta_k \neq 0$, set

$$\alpha_1 := -\frac{v_k}{\delta_k}.$$

Calculate

$$\alpha_2 := \frac{u_j}{w_j} = \min_{i \in \hat{I}} \left(\frac{u_i}{w_i} \right),$$

where $w_i = q_{ki} - (\beta_k/\beta_l)\, q_{li}$, $i \in J$, $w_l = \beta_k/\beta_l$, $\hat{I} = \{i \in I : w_i > 0\}$. Set $\alpha := \min(\alpha_1, \alpha_2)$. Calculate $u := u - \alpha(q_k - (\beta_k/\beta_l)\, q_l)$, $u_l := u_l - \alpha(\beta_k/\beta_l)$, $u_k := u_k + \alpha$, and $v_k := (1 - \alpha/\alpha_1)\, v_k$. If $\alpha = \alpha_1$, set $q_l := [q_l^{\mathrm{T}}, 0]^{\mathrm{T}}$ and go to Step 5. If $\alpha \neq \alpha_1$ and $j = l$, set $I := I \setminus \{l\}$ and go to Step 4. If $\alpha \neq \alpha_1$ and $j \neq l$, go to Step 6.

Algorithm 4.1 has a considerably complex logical structure, but its numerical calculations are not more expensive than those of a usual algorithm for solving the standard quadratic programming problems. Similar operations take place in both the singular and nonsingular cases, the only exception being Step 7 where, when performing the transition from a singular to a nonsingular case, it is necessary to calculate repeatedly the vector q_l in order to make the reconstruction of the matrix C possible.

Let us show that, in the singular case, $\beta_l \neq 0$ holds (assumption of Lemma 3.1). This inequality is valid in each transition from a nonsingular case to a singular case ($\beta_k \gamma_k + \delta_k \neq 0$ and $\delta_k = 0$). Inspecting Step 7 and Step 8 we can see that the value β_l remains unchanged in the singular case (the vector q_l is changed only by adding or by deleting a zero element). Therefore we have always $\beta_l \neq 0$ in the singular case, so that the assumption of Lemma 3.1 is valid.

Now we will prove the convergence of the dual method for solving the problem (1.3).

Theorem 4.1. *Algorithm* 4.1 *finds the solution of the problem* (1.3) *after a finite number of steps.*

Proof. Let $(s^*, z^*) \in R_{n+1}$ be the solution of the problem (1.3) and let $I^* \subset M$ be a set of indices such that $(s^*, z^*) \in R_{n+1}$ is a basic solution of the problem $P(I^*)$. During the execution of Algorithm 4.1 we construct a sequence of subsets $\bar{I}_j \subset M$, $j \geqq 0$. Theorems 2.1 and 3.1 guarantee the validity of $\varphi(\bar{s}_j, \bar{z}_j) > \varphi(\bar{s}_{j-1}, \bar{z}_{j-1})$, $j \geqq 1$ (the pair $(\bar{s}_j, \bar{z}_j) \in R_{n+1}$ is a basic solution of the problem $P(\bar{I}_j)$). Hence the sets $\bar{I}_j \subset M$, $j \geqq 0$ must be distinct. Since M is finite, the sequence of mutually different subsets $\bar{I}_j \subset M$, $j \geqq 0$ is also finite. The last element of this sequence must be the set I^*, for, if it were not so, it could be possible to continue in constructing the next subset according to Algorithm 4.1. □

Algorithm 4.1 uses the matrices $H = G^{-1}$, $C = (A^{\mathrm{T}}HA)^{-1}$. If we desire to obtain a numerically more stable version of the algorithm, we can replace the matrix $H = G^{-1}$ with a triangular decomposition $G = LL^{\mathrm{T}}$ (L is a lower triangular matrix). Similarly, instead of the matrix C, we can make use of the orthogonal decomposition

$$L^{-1}A = Q\begin{bmatrix} R \\ 0 \end{bmatrix},$$

where Q is an orthogonal matrix and R is an upper triangular matrix. More details relative to these decompositions are presented in [6].

References

[1] *M. S. Bazaraa, C. M. Shetty*: Nonlinear programming. Theory and algorithms. New York: Wiley 1979.

[2] *C. Charalambous, J. W. Bandler*: Nonlinear minimax optimization as a sequence of least p-th optimization with finite values of p. Faculty Engn., McMaster University, Hamilton, Ontario, Canada, Rept. SOC-3, 1973.

[3] *C. Charalambous:* Acceleration of the least$p$-th algorithm for minimax optimization with engineering applications. Math. Programming 17, 270—297, (1979).

[4] *V. F. Demyanov, V. N. Malozemov:* Introduction to minimax, Chap. 3, § 5. New York: Wiley 1974.

[5] *D. Golfarb:* Extension of Davidon's variable metric method to maximization under linear inequality and equality constraints. SIAM J. Appl. Math. 17, 739—764, (1969).

[6] *D. Goldfarb, A. U. Idnani:* A numerically stable dual method for solving strictly convex quadratic programs. The City College of New York, Dept. of Computer Sci., Rept. 81—102, (1981).

[7] *J. Hald, K. Madsen:* Combined LP and Quasi-Newton methods for minimax optimization. Math. Programming 20, 49—62, (1981).

[8] *S. P. Han:* Variable metric methods for minimizing a class of nondifferentiable functions. Math. Programming 20, 1—13, (1981).

[9] *L. Lukšan:* Variable metric methods for linearly constrained nonlinear minimax approximation. Computing 30, 315—334, (1983).

[10] *K. Madsen:* An algorithm for minimax solution of overdetermined systems of nonlinear equations. J. Inst. Math. Appl. 16, 321—328, (1975).

[11] *M. J. D. Powell:* A fast algorithm for nonlinearly constrained optimization claculations. In "Numerical analysis, Dundes 1977", (G. A. Watson, ed.), Lecture Notes in Mathematics 630, Berlin: Springer-Verlag 1978.

Souhrn

DUÁLNÍ METODA PRO ŘEŠENÍ SPECIÁLNÍ ÚLOHY KVADRATICKÉHO PROGRAMOVÁNÍ, KTERÁ SE VYSKYTUJE PŘI NELINEÁRNÍ MINIMAXOVÉ APROXIMACI

LADISLAV LUKŠAN

V článku je popsána duální metoda pro řešení speciální úlohy kvadratického programování, která se vyskytuje jako podúloha při nelineární minimaxové aproximaci. Podrobně jsou analyzovány dva případy, které se liší lineární závislostí gradientů funkcí aktivních v daném bodě. Závěrem je uveden podrobný algoritmus duální metody a je dokázána jeho konvergence po konečném počtu kroků.

Резюме

ДУАЛЬНЫЙ МЕТОД РЕШЕНИЯ СПЕЦИАЛЬНОЙ ЗАДАЧИ КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ КОТОРАЯ ЯВЛЯЕТСЯ ПОДЗАДАЧЕЙ НЕЛИНЕЙНОЙ МИНИМАКСНОЙ АППРОКСИМАСИИ

LADISLAV LUKŠAN

В статье описывается дуальный метод решения задачи квадратичного программирования, которая является подзадачей нелинейной минимаксной аппроксимации. Подробнее анализируются два случая, которые отличаются друг от друга линейной зависимостью градиентов функций, активных в данной точке. В заключение приводится подробный алгоритм дуального метода, и доказывается его сходимость после конечного числа шагов.

Author's address: Ing. *Ladislav Lukšan,* CSc., Středisko výpočetní techniky ČSAV, Pod vodárenskou věží 2, 182 07 Praha 8.