Jaroslav Hrouda

Two minimax-type methods for solving systems of nonlinear equations

Aplikace matematiky, Vol. 14 (1969), No. 1, 29–53

Persistent URL: http://dml.cz/dmlcz/103206

TWO MINIMAX-TYPE METHODS FOR SOLVING SYSTEMS OF NONLINEAR EQUATIONS

JAROSLAV HROUDA

(Received April 20, 1967)

1. ASSIGNMENT OF EXTREMUM PROBLEMS

Let us have a system of equations

$$f_i(x) = 0\,,\ i \in K\,, \qquad p_i(x) = 0\,,\ i \in L \tag{1.1}$$

where f_i are continuously differentiable (nonlinear) real functions,[1]) p_i linear non-constant functions, x is a point of E_n, K and L are disjunctive sets of indices.[2]) If r stands for the number of equations of the system, then $r \lesseqgtr n$. We shall use a common notation h_i, $i \in K \cup L = I$ for both types of the functions in (1.1).

The system (1.1) is assigned functions

$$\alpha(x) = \max_{i \in I}\ h_i(x)\,, \tag{1.2}$$

$$\beta(x) = \max_{i \in I} |h_i(x)| \tag{1.3}$$

and a set of E_n

$$\Omega = \{x \mid h_i(x) \geqq 0,\, i \in I\} \tag{1.4}$$

which is a closed and, in general, disconnected set. Then the solving of the system can be formulated as

A. a constrained extremum problem

$$\min_x \{\alpha(x) \mid x \in \Omega\}\,, \tag{1.5}$$

[1]) Defined where necessary.

[2]) One of the sets K and L may be empty.

B. an unconstrained extremum problem

$$(1.6) \qquad \min_x \beta(x) .$$

The methods we shall apply to solving these minimization problems will be based upon the principle of successive relaxation of the values $\alpha(x)$ or $\beta(x)$. Iterative processes will generally converge to a so-called A-stationary point in the case (1.5) and B-stationary point in (1.6). Of course, all roots (absolute minima) are included in both classes of these stationarities.

To develop these methods we shall make use of the ideas of *Zoutendijk*'s *method of feasible directions* [1] [2] known from mathematical programming.[3]) A similar idea is used in [3] to obtain the Chebyshev solution of an inconsistent linear system.

The system (1.1) can be assigned extremum problems still in another way: Let us denote by Ω' and $\tilde{\Omega}'$ the sets of E_{n+1}

$$(1.7) \qquad \Omega' = \{(x, x_{n+1}) \mid 0 \leqq h_i(x) \leqq x_{n+1}, \ i \in I\} ,^4)$$

$$(1.8) \qquad \tilde{\Omega}' = \{(x, x_{n+1}) \mid |h_i(x)| \leqq x_{n+1}, \ i \in I\} .$$

Then, clearly, each root of the system (1.1) is also a solution of both the problems

$$(1.9) \qquad \min_{(x, x_{n+1})} \{x_{n+1} \mid (x, x_{n+1}) \in \Omega'\} ,$$

$$(1.10) \qquad \min_{(x, x_{n+1})} \{x_{n+1} \mid (x, x_{n+1}) \in \tilde{\Omega}'\} .$$

In contrast to objective functions of (1.5) and (1.6) those corresponding to (1.9) and (1.10) are differentiable. Therefore it is possible to solve (1.9), (1.10) by applying the general methods of mathematical programming [1] [6] [7]. It should be noticed, however, that there are no methods which would always – without more assumptions – provide absolute extrema.[5])

---

[3]) Accurately speaking, one of its version called by Zoutendijk "algorithm P1 with precaution AZ1". Knowledge of this method is not necessary for understanding the article. We shall use only a few basic concepts from the theory of mathematical programming.

[4]) (x, x_{n+1}) is a vector of E_{n+1}; it is obtained by adding the component x_{n+1} to the vector $x \in E_n$.

[5]) In this connection Mr. J. Nedoma suggested another quite natural possibility:

$$\min_x \{\sum_i h_i(x) \mid x \in \Omega\} .$$

Of course, this approach is not a minimax-type one.

2.1. Definitions and properties of concepts. We resume the notation introduced in the foregoing section. All the following definitions and assertions will apply to some point $x \in \Omega$. Let us denote by $I_+(x)$, $I_0(x)$, and $I(x)$ the sets of indices

$$\begin{aligned} I_+(x) &= \{i \in I \mid h_i(x) = \alpha(x)\}\,, \\ I_0(x) &= \{i \in I \mid h_i(x) = 0\}\,, \\ I(x) &= I_+(x) \cup I_0(x)\,. \end{aligned} \tag{2.1}$$

Then the meaning of $K_+(x)$, $L_+(x)$, etc. will be obvious. Further, we denote by $A(x)$ the set of vectors

$$\begin{aligned} A(x) = \{s \neq 0 \mid \nabla h_i(x)^{\mathrm{T}}\, s &< 0\,, \quad i \in I_+(x)\,, \\ \nabla f_i(x)^{\mathrm{T}}\, s &> 0\,, \quad i \in K_0(x)\,, \\ \bar{p}_i(s) &\geqq 0\,, \quad i \in L_0(x)\}\,.^{6)} \end{aligned} \tag{2.2}$$

Here $\bar{p}_i(s) = \nabla p_i(x)^{\mathrm{T}}\, s$; so the bar denotes breaking away of the absolute term of a linear function.

Definition 1. *Vector s is called an A-direction of the point x if $s \in A(x)$.*

Definition 2. *x is called an A-stationary point of the system (of level ϱ) if $A(x) = \emptyset$ (and $\alpha(x) = \varrho$).*[7]

In this way each $x \in \Omega$ either is assigned a set of vectors (A-directions) or its A-stationarity is stated (no A-direction of x exists).

Remark 1. All roots of the system (1.1) are just all its A-stationary points of level 0. For if x is a root, then $x \in \Omega$, $\alpha(x) = 0, I_+(x) = I_0(x) = I$, and thus according to (2.2) it must be $A(x) = \emptyset$. The converse assertion follows from the relations $0 \leqq h_i(x) \leqq \alpha(x) = 0,\ i \in I$.

Lemma 1. *If $s \in A(x)$, then there is a number $\tilde{\lambda} > 0$ such that $\alpha(x + \lambda s) < \alpha(x)$, $x + \lambda s \in \Omega$ for all $0 < \lambda \leqq \tilde{\lambda}$ (more precisely: $f_i(x + \lambda s) > 0$, $i \in K$; $p_i(x + \lambda s) > 0$, $i \in L - L_0(x)$).*

Proof. First notice that in virtue of Remark 1 $\alpha(x) > 0$. From (2.1), (2.2), and the continuity assumption of partial derivatives of h_i it follows that there exist num-

[6]) $\nabla h_i(x) = \operatorname{grad} h_i(x)$, the symbol $^{\mathrm{T}}$ means scalar product. The notation $A(x)$ does not relate with that in [4] and [5].

[7]) The terms from both definitions will be often used in an abbreviated form: A-direction, A-stationary point.

bers $\tau_i > 0$ $(i \in I)$ such that for all $0 < \tau \leqq \tau_i$ the following inequalities hold:

$$0 < h_i(x + \tau s) = \alpha(x) + \tau \nabla h_i(x + \Theta_i \tau s)^T s < \alpha(x) \quad \text{for} \quad i \in I_+(x)\,,^{8})$$
$$\alpha(x) > f_i(x + \tau s) = \tau \nabla f_i(x + \Theta_i \tau s)^T s > 0 \quad \text{for} \quad i \in K_0(x)\,,$$
$$\alpha(x) > p_i(x + \tau s) = \tau \bar{p}_i(s) \geqq 0 \quad \text{for} \quad i \in L_0(x)\,,$$
$$0 < h_i(x + \tau s) < \alpha(x) \quad \text{for} \quad i \in I - I(x)\,.$$

Putting down $\tilde{\lambda} = \min_{i \in I} \tau_i$ finishes the proof.

Lemma 1 shows that when we locally move from the point x along its A-direction we can decrease the value $\alpha(x)$ remaining within the set Ω. The following assertion provides a criterion for deciding about the A-stationarity of a point.

Lemma 2. *The point x is an A-stationary point if and only if*

$$\begin{aligned} \max_{(s,\sigma)} \{\sigma \mid \nabla h_i(x)^T s + \theta_i \sigma &\leqq 0\,, \quad i \in I_+(x)\,, \\ -\nabla f_i(x)^T s + \theta_i \sigma &\leqq 0\,, \quad i \in K_0(x)\,, \\ -\bar{p}_i(s) \qquad &\leqq 0\,, \quad i \in L_0(x)\} = 0 \end{aligned} \tag{2.3}$$

where θ_i are some positive numbers.

Proof. Let us denote by $\bar{\sigma}$ the left-hand side of the equality (2.3). Let $\bar{\sigma} = 0$. If there exists an $\tilde{s} \in A(x)$, the vector $(\tilde{s}, \tilde{\sigma})$ where

$$\tilde{\sigma} = \min_i \left\{ -\frac{\nabla h_i(x)^T \tilde{s}}{\theta_i}\,,\ i \in I_+(x), \frac{\nabla f_i(x)^T \tilde{s}}{\theta_i}\,,\ i \in K_G(x) \right\} > 0$$

would satisfy both the conditions in (2.3) and $\tilde{\sigma} > \bar{\sigma}$. This is impossible, however, thus $A(x) = \emptyset$. On the other hand, let $A(x) = \emptyset$. If there were a vector $(\tilde{s}, \tilde{\sigma})$ satisfying the conditions in (2.3) and $\tilde{\sigma} > 0$, it would mean that an A-direction of the point x exists. The vector $(s, \sigma) = (0, 0)$ fulfils the conditions in (2.3), hence $\bar{\sigma} = 0$.

Remark 2. If $x \in \Omega$ is such that $\nabla h_i(x) = 0$ for some $i \in K(x)$, then x is A-stationary. Indeed, in this case a condition $\theta_i \sigma \leqq 0$ occurs in (2.3), so $\bar{\sigma} = 0$.

By words, an A-stationary point x can be characterized as follows: Out of this point there passes no direction along which it would be possible locally to decrease the value $\alpha(x)$, and at the same time to increase all the zero residua of the equations $f_i(x) = 0$, and not to decrease none of the zero residua of the equations $p_i(x) = 0$. For if such a direction exists, it could be proved by consideration similar to that in Lemma 1 that it is an A-direction. As for its sense, the conception of A-stationary

---

[8]) $0 < \Theta_i < 1$.

point is therefore equivalent to that of Ω-quasistationary point of the function $\alpha(x)$ [5]. (If $\nabla\alpha(x)$ exists, both conceptions are identical even by definition.) If the set Ω fulfils the regularity condition, the corresponding notion will be the Ω-stationary point of the function $\alpha(x)$.

Concerning the mathematical programming problem (1.9) the following statement is valid: x is an A-stationary point of the system if and only if $(x, \alpha(x))$ is an Ω'-quasistationary point of the function x_{n+1}.[9] Under the regularity condition for the set Ω', $(x, \alpha(x))$ is an Ω'-stationary point of the function x_{n+1}.

The following interpretation will make the meaning of A-direction still clearer: Let us denote by $\Omega_{\alpha(x)}$ the set

$$\Omega_{\alpha(x)} = \{y \mid 0 \leqq h_i(y) \leqq \alpha(x),\ i \in I\}\,, \tag{2.4}$$

The A-direction of the point x is a vector with x as origin and pointing to the interior of the set (2.4) (i.e., excluding tangent position to those boundaries of the set on which the point x lies unless they are "low" linear boundaries).

In order to make possible to compare directional qualities of A-directions, it is necessary to eliminate somehow the influence of their lengths. This will be also useful for working out methods of calculating these vectors.

Definition 3. *We shall call a set of vectors $N(x)$ "normalization set relating to $A(x)$" if it has the following properties:*

a) *If $A(x) \neq \emptyset$, then*

aa) *for each $s \in A(x)$ there is a positive number $\bar{\gamma}(s)$ such that*

$$\gamma s \in N(x) \quad \textit{for all} \quad 0 \leqq \gamma \leqq \bar{\gamma}(s)\,, \quad \bar{\gamma}(s)\,\|s\| \geqq \omega_1 \tag{2.5}$$

where $\omega_1 > 0$ is a constant (independent on x);

ab) *for each $s \in N(x)$ it holds*

$$\|s\| \leqq \omega_2 \tag{2.6}$$

where ω_2 is a constant ($\omega_2 \geqq \omega_1$);

ac) *$N(x)$ is closed.*

b) *If $A(x) = \emptyset$, then $N(x)$ can be an arbitrary set containing the zero element.*

Now, we shall execute the required normalization of the set $A(x)$ by setting

$$\bar{A}(x) = A(x) \cap N(x)\,. \tag{2.7}$$

[9]) It can be obtained by Zoutendijk's method of feasible directions (without the need for the regularity condition).

Obviously, all A-directions of the point x remain in the set (2.7) if not regarding their lengths. Thus the conception of the A-stationary point will not be influenced if defined by (2.7); i.e. it holds

Lemma 3. *$A(x) = \emptyset$ if and only if $\bar{A}(x) = \emptyset$.*

Remark 3. Lemma 2 will remain valid if the normalization condition $s \in N(x)$ is inserted into (2.3). This can be proved by adapting slightly the proof of Lemma 2 and applying Lemma 3.

2.2. Algorithm. In this section an iterative procedure for obtaining one of the A-stationary points of the system (1.1) will be given. First, we shall describe the algorithmic scheme of one iteration (we will justify it afterwards).

Let us have a point $x \in \Omega$ and a number $0 < \delta < \frac{1}{2}$. Let us denote by $I_+(x, \delta)$, $I_0(x, \delta)$, and $I(x, \delta)$ the sets of indices

$$\begin{aligned} I_+(x, \delta) &= \{i \in I \mid h_i(x) \geqq (1 - \delta)\, \alpha(x)\}\,, \\ I_0(x, \delta) &= \{i \in I \mid h_i(x) \leqq \delta\alpha(x)\}\,, \\ I(x, \delta) &= I_+(x, \delta) \cup I_0(x, \delta)\,.^{10)} \end{aligned} \tag{2.8}$$

Replacing in (2.2) $I(x)$ by $I(x, \delta)$ we get a set $A(x, \delta)$.[11] By inserting $A(x, \delta)$ into Definition 3 instead of $A(x)$ a normalization set $N(x, \delta)$ will be defined.

For given x and δ the *auxiliary problem* for finding an A-direction of the point x is formulated as follows:

$$\begin{aligned} \max_{(s,\sigma)} \{\sigma \mid \nabla h_i(x)^{\mathrm{T}} s + \|\nabla h_i(x)\|\, \sigma &\leqq 0\,, \quad i \in I_+(x, \delta)\,, \\ -\nabla f_i(x)^{\mathrm{T}} s + \|\nabla f_i(x)\|\, \sigma &\leqq 0\,, \quad i \in K_0(x, \delta)\,, \\ -\bar{p}_i(s) &\leqq 0\,, \quad i \in L_0(x, \delta)\,, \\ s \in N(x, \delta) & \qquad\qquad \}\,.^{12)} \end{aligned} \tag{2.9}$$

The extremum problem (2.9) is a mathematical programming one. Let the symbol $\Pi(x, \delta)$ stand for it. Let us suppose that all norms of the gradients in (2.9) are non-zero. The auxiliary problem is always solvable (if $A(x, \delta) \neq \emptyset$, it follows from the definition of the set $N(x, \delta)$; if $A(x, \delta) = \emptyset$, the problem has a zero optimal solution, which can be proved like Lemma 2).

[10]) Analogically for the letters K and L.

[11]) It is $A(x, \delta) \subset A(x)$.

[12]) $\| \; \|$ means the Euclidean norm of a vector.

Let $(s_\delta^*, \sigma_\delta^*)$ be an optimal solution of the auxiliary problem (2.9). If $\sigma_\delta^* > 0$, then $\|s_\delta^*\| \neq 0$;[13]) put down

$$(2.10) \qquad (s', \sigma') = (s_\delta^*, \sigma_\delta^*)\,,$$

$$(2.11) \qquad \delta' = \begin{cases} \delta & \text{if} \quad \sigma'/\|s'\| > \delta\,, \\ \frac{1}{2}\delta & \text{if} \quad \sigma'/\|s'\| \leqq \delta\,. \end{cases}$$

If $\sigma_\delta^* = 0$, then solve the problem $\Pi(x, 0)$. If also $\sigma_0^* = 0$, then x is an A-stationary point (see Lemma 2 and Remark 3). Otherwise put

$$(2.12) \qquad (s', \sigma') = (s_0^*, \sigma_0^*)\,,$$

$$(2.13) \qquad \delta' = \tfrac{1}{2}\delta\,.$$

The vector s' is an (optimal) A-direction of the point x.

If there is $\|\nabla h_i(x)\| = 0$ for some $i \in K(x, \delta)$, turn immediately to the problem $\Pi(x, 0)$. If there is a zero gradient even there, then x is an A-stationary point (according to Remark 2).

The A-direction s' being known, Lemma 1 guarantees the existence of a number $\lambda' > 0$ such that the point

$$(2.14) \qquad x' = x + \lambda' s'$$

will satisfy the relations

$$(2.15) \qquad \alpha(x') < \alpha(x)\,, \quad x' \in \Omega\,.$$

Choose the number $\lambda' > 0$ so that the inequality in (2.15) may hold more strongly, namely as $\alpha(x) - \alpha(x') \geqq \varepsilon$ where $\varepsilon > 0$ is a given constant. If no such a number exists, take $\lambda' = \lambda^*$ where λ^* is a solution of the one-dimensional constrained minimization problem

$$(2.16) \qquad \min_\lambda \{\alpha(x + \lambda s') \mid x + \lambda s' \in \Omega, \lambda \geqq 0\}\,.$$

The latter way of getting λ' is an optimal one: it makes $\alpha(x)$ decrease as much as possible in the direction s' remaining in Ω.

Remark 4. The norm of the vector s' satisfies the inequalities

$$(2.17) \qquad \omega_1 \leqq \|s'\| \leqq \omega_2\,.$$

This is a consequence of the extremum nature of s', as well as of (2.5) and (2.6).

[13]) It follows from the inequalities in (2.9) since $I_+(x, \delta) \neq \emptyset$.

Remark 5. Since (s', σ') is a solution of (2.9), it holds

$$(2.18) \qquad 0 < \frac{\sigma'}{\|s'\|} = \min_i \{|\cos(\nabla h_i(x), s')|, \; i \in I_+(x, \delta) \cup K_0(x, \delta)\} \leqq 1 .^{14})$$

Now, if an initial point $x^0 \in \Omega$ and an initial value $0 < \delta_0 < \frac{1}{2}$ are given, we can construct the sequences

$$(2.19) \qquad \{x^k\}, \{s^k\}, \{\delta_k\}, \{\lambda_k\}, \{\sigma_k\}^{15})$$

by recursive application of the procedure described above (writing x^k instead of x, x^{k+1} instead of x', etc.).

Before formulating the convergence assertion we will attempt to clear up some parts of the algorithm:

In the $(k + 1)$th iteration we can either state that an A-stationary point has been reached or construct — on the basis of x^k — a new point x^{k+1} satisfying $\alpha(x^{k+1}) < \alpha(x^k)$ and $x^{k+1} \in \Omega$. Such a point can be found along an A-direction of x^k. The solving of the problem (2.9) provides a vector (an optimal one) from the set of A-directions. In order to make it more illustrative, let us write down the problem (2.9) in the form

$$(2.20) \qquad \min_s \{\|s\| \max_i \{\cos(\nabla h_i(x^k), s), \; i \in I_+(x^k, \delta_k), \\ \cos(-\nabla f_i(x^k), s), \; i \in K_0(x^k, \delta_k)\}| \\ \bar{p}_i(s) \geqq 0, \; i \in L_0(x^k, \delta_k), \; s \in N(x^k, \delta_k)\}$$

and remind the interpretation of an A-direction by (2.4). The solution of (2.20) yields a vector which directs towards the interior of the set $\Omega_{\alpha(x^k)}$ departing as much as possible — in Chebyshevian sense and with respect to the normalization applied — from those boundaries of the set the point x^k is near (except for the "low" linear boundaries where also parallel position is admitted). The parameter δ_k serves as a measure of this nearness of boundaries. It appears necessary to guarantee the convergence of the process (it avoids small steps towards close boundaries). The value of δ_k is reduced during the computation whenever there appears an optimal A-direction only slightly deviated from the boundaries of $\Omega_{\alpha(x^k)}$. So, the influence of the parameter δ_k is gradually getting weaker if it would prevent proceeding towards an A-stationary point. In the extremal situation when no non-zero solution of the problem (2.9) exists, it is necessary to search for an A-direction without respect to δ_k, i.e. in the whole set $A(x^k)$. From the setting of (2.20) it is also apparent why $\|\nabla h_i(x^k)\|$

[14]) Symbol $\cos(a, b)$ means $a^T b/(\|a\| \, \|b\|)$.

[15]) Here the index number is written as superscript if the sequence is a vectorial one and as subscript if scalar.

occur in the problem (2.9) — in this way the undesirable influence of the lengths of gradients of h_i upon choice of optimal A-directions is avoided.[16])

As concerns the number λ_{k+1}, undoubtedly, it is of advantage to take it (optimally) such as to bring about the deepest decrease of the value $\alpha(x)$ in the $(k + 1)$th iteration. In general, this is a difficult task. But in case there exists an interval $\mu_{k+1} < \lambda < \nu_{k+1}$ where it holds $\alpha(x^k + \lambda s^{k+1}) + \varepsilon \leqq \alpha(x^k)$, $x^k + \lambda s^{k+1} \in \Omega$, the algorithm enables us to choose λ_{k+1} as an arbitrary number from this interval.

And now, we proceed to the convergence questions. Let us suppose that none of the points x^k is A-stationary, so that the sequences (2.19) are infinite.

Theorem 1. *If for some $\varrho > 0$ the set Ω_ϱ* [17]) *is bounded and $x^k \in \Omega_\varrho$ for some $k = k_0$, then there exists a cluster point of the sequence $\{x^k\}$ which is an A-stationary point of the system* (1.1).

Proof. Since $x^k \in \Omega_\varrho$ for all $k \geqq k_0$, the sequence $\{x^k\}$ is bounded. Let us denote $\bar{\delta} = \lim \delta_k$. Two possibilities are to be distinguished:

1. $\bar{\delta} = 0$. Then there must be an infinite subsequence of "halving" $\{\delta_{k_l+1} = \frac{1}{2}\delta_{k_l}\}$. We can take, without loss of generality, $k_l = l$. Then we select a convergent subsequence $\{x^{l_q}\}$ with a limit $\bar{x}$. Obviously, $\bar{x} \in \Omega$. Again, we take for the sake of simplicity $l_q = q$. Let us suppose that $\bar{x}$ is not an A-stationary point. Then $\alpha(\bar{x}) > 0$, $\|\nabla h_i(\bar{x})\| > 0$ for $i \in I(\bar{x})$ according to Remarks 1 and 2. It holds

$$I(x^q, \delta_q) \subset I(\bar{x}) \quad \text{for asl. } q\,. \tag{2.21}$$ [18])

Indeed, $i \notin I(\bar{x})$ indicates $0 < h_i(\bar{x}) < \alpha(\bar{x})$. But then (with regard to the convergences and continuities) it will be also for asl. q

$$\delta_q\, \alpha(x^q) < h_i(x^q) < (1 - \delta_q)\, \alpha(x^q)\,,$$

i.e. $i \notin I(x^q, \delta_q)$.

Considering Lemma 2, we can postulate a vector $(\tilde{s}, \tilde{\sigma})$ such that $\tilde{s} \neq 0$, $\tilde{\sigma} > 0$, and

$$\begin{aligned}
\nabla h_i(\bar{x})^{\mathrm{T}}\, \tilde{s} + \|\nabla h_i(\bar{x})\|\, \tilde{\sigma} &\leqq 0\,, \quad i \in I_+(\bar{x})\,,\\
-\nabla f_i(\bar{x})^{\mathrm{T}}\, \tilde{s} + \|\nabla f_i(\bar{x})\|\, \tilde{\sigma} &\leqq 0\,, \quad i \in K_0(\bar{x})\,,\\
-\bar{p}_i(\tilde{s}) \qquad\qquad &\leqq 0\,, \quad i \in L_0(\bar{x})\,.
\end{aligned}$$

[16]) Naturally, such a procedure will be of value only if the solving of the auxiliary problems is much less laborious than that of the original system of equations.

[17]) See (2.4).

[18]) asl. means *all sufficiently large.*

Because of continuity it holds

$$\begin{aligned}
\nabla h_i(x^q)^{\mathrm{T}}\,\tilde{s} + \tfrac{1}{2}\|\nabla h_i(x^q)\|\,\tilde{\sigma} &\leqq 0\,, \quad i \in I_+(\bar{x})\,,\\
-\nabla f_i(x^q)^{\mathrm{T}}\,\tilde{s} + \tfrac{1}{2}\|\nabla f_i(x^q)\|\,\tilde{\sigma} &\leqq 0\,, \quad i \in K_0(\bar{x})\,,\\
-\bar{p}_i(\tilde{s}) \qquad\qquad\qquad\qquad &\leqq 0\,, \quad i \in L_0(\bar{x})
\end{aligned}$$

for asl. q, hence according to (2.21) $\tilde{s} \in A(x^q, \delta_q)$ for asl. q – say $q > Q$. For each of these q there exists a number $\bar{\gamma}_q$ such that $\gamma\tilde{s} \in N(x^q, \delta_q)$ for $0 \leqq \gamma \leqq \bar{\gamma}_q$ and $\bar{\gamma}_q \geqq \omega_1/\|\tilde{s}\|$.[19]) Setting

$$\bar{\gamma} = \inf_{q>Q} \bar{\gamma}_q$$

gives $\bar{\gamma} > 0$ and $\bar{\gamma}\tilde{s} \in N(x^q, \delta_q)$, $q > Q$. Then the vector $\bar{\gamma}(\tilde{s}, \tfrac{1}{2}\tilde{\sigma})$ is a feasible solution of all the auxiliary problems $\Pi(x^q, \delta_q)$, $q > Q$, thus

$$\sigma^*_{\delta_q} = \sigma_{q+1} \geqq \tfrac{1}{2}\bar{\gamma}\tilde{\sigma} > 0 \quad \text{for asl. } q\,. \tag{2.22}$$

But at the same time we get for these q according to (2.11) and (2.17)

$$\sigma_{q+1} \leqq \delta_q\|s^{q+1}\| \leqq \delta_q\omega_2\,.$$

Because of the assumption $\bar{\delta} = 0$ this implies $\sigma_{q+1} \to 0$, which is in contradiction with (2.22).

2. $\bar{\delta} > 0$. Again, we can have convergent subsequences $x^{k_l} \to \bar{x} \in \Omega$, $s^{k_l+1} \to \bar{s}$, $\sigma_{k_l+1} \to \bar{\sigma}$[20]) and can simplify indexing: $k_l = l$. Clearly, $\delta_l = \bar{\delta}$ for asl. l, thus they are the problems $\Pi(x^l, \delta_l)$ (not $\Pi(x^l, 0)$!) to be solved, and according to (2.11), (2.17) it holds $\sigma_{l+1} > \delta_l\|s^{l+1}\| \geqq \bar{\delta}\omega_1$, which implies $\bar{\sigma} > 0$. Let us suppose that $\bar{x}$ is not A-stationary. Now, it can be easily seen that the inclusion

$$I(\bar{x}) \subset I(x^l, \delta_l) \tag{2.23}$$

is valid for asl. l. Really, if $i \in I_+(\bar{x})$, then

$$\frac{h_i(x^l)}{\alpha(x^l)} \geqq 1 - \bar{\delta} = 1 - \delta_l$$

for asl. l; if $i \in I_0(\bar{x})$, then $h_i(x^l) \leqq \bar{\delta}\alpha(\bar{x}) < \delta_l\alpha(x^l)$ for asl. l.

---

[19]) See the property aa) of the normalization set. Here the normalization set from the auxiliary problem $\Pi(x^q, \delta_q)$ is considered.

[20]) The inequalities (2.17) and (2.18) indicate boundedness of the sequences $\{s^k\}$ and $\{\sigma_k\}$.

It follows from (2.23) that the inequalities

$$(2.24)\qquad \begin{aligned} &\nabla h_i(x^l)^{\mathrm{T}} s^{l+1} + \|\nabla h_i(x^l)\|\, \sigma_{l+1} \leqq 0\,, && i \in I_+(\bar{x})\,,\\ -&\nabla f_i(x^l)^{\mathrm{T}} s^{l+1} + \|\nabla f_i(x^l)\|\, \sigma_{l+1} \leqq 0\,, && i \in K_0(\bar{x})\,,\\ -&\bar{p}_i(s^{l+1}) \qquad\qquad\qquad\qquad \leqq 0\,, && i \in L_0(\bar{x}) \end{aligned}$$

are valid for asl. l. By limiting them we get

$$\begin{aligned} &\nabla h_i(\bar{x})^{\mathrm{T}} \bar{s} + \|\nabla h_i(\bar{x})\|\, \bar{\sigma} \leqq 0\,, && i \in I_+(\bar{x})\,,\\ -&\nabla f_i(\bar{x})^{\mathrm{T}} \bar{s} + \|\nabla f_i(\bar{x})\|\, \bar{\sigma} \leqq 0\,, && i \in K_0(\bar{x})\,,\\ -&\bar{p}_i(\bar{s}) \qquad\qquad\qquad \leqq 0\,, && i \in L_0(\bar{x})\,. \end{aligned}$$

Since $\|\nabla h_i(\bar{x})\|\, \bar{\sigma} > 0$ for $i \in I(\bar{x})$, the vector $\bar{s} \neq 0$ is an A-direction of the point $\bar{x}$. By virtue of Lemma 1 there is a number $\tilde{\lambda} > 0$ such that

$$\begin{aligned} 0 < f_i(\bar{x} + \tilde{\lambda}\bar{s}) < \alpha(\bar{x})\,, \quad i \in K\,,\\ 0 \leqq p_i(\bar{x} + \tilde{\lambda}\bar{s}) < \alpha(\bar{x})\,, \quad i \in L\,. \end{aligned}$$

But then for asl. l it will be

$$0 \leqq h_i(x^l + \tilde{\lambda} s^{l+1}) < \alpha(x^{l+1})\,, \quad i \in I\,,\ ^{21})$$

which is in contradiction with the optimal determination of the numbers λ_{l+1} that occurs for asl. l owing to monotony as well as boundedness of the sequence $\{\alpha(x^k)\}$. This completes the proof of Theorem 1.

Remark 6. Theorem 1 is true also for a rather simplified algorithm which contains arbitrary constants $\theta_i > 0$ instead of $\|\nabla h_i(x)\|$ in (2.9) and an arbitrary constant $\varkappa > 0$ instead of $\|s^l\|$ in (2.11).

Finally, we shall formulate a *sufficient condition* to assure the method of A-directions will tend to a root. Let us denote by $J(x)$ the Jacobian matrix of the system (1.1)

$$J(x) = \left(\frac{\partial h_i(x)}{\partial x_j}\right), \quad \begin{aligned} i &= 1, \ldots, r\,,\\ j &= 1, \ldots, n\,. \end{aligned}$$

Theorem 2. *Let $r \leqq n$. If for some $\varrho > 0$ the set Ω_ϱ is non-empty,[22]) bounded and the rank of $J(x)$ is r for all $x \in \Omega_\varrho - \Omega_0$, then the system (1.1) has solutions. The method of A-directions determines one of them starting from any initial point $x^0 \in \Omega_\varrho$.*

[21]) $\tilde{\lambda}$ can be chosen so as to imply $p_i(\bar{x} + \tilde{\lambda}\bar{s}) = 0$ at most for $i \in L_0(\bar{x})$, but in this case $\bar{p}_i(s^{l+1}) \geqq 0$ because of (2.24).

[22]) This assumption is no restrictive one (see Section 4.3).

Proof. According to Theorem 1 the method of A-directions provides an A-stationary point (in the set Ω_ϱ). But there are no A-stationary points in Ω_ϱ, except roots. Indeed, for each $x \in \Omega_\varrho - \Omega_0$ we could obtain an A-direction by solving the algebraic system

$$\nabla h_i(x)^{\mathrm{T}} s = \varkappa_i \,, \quad i \in I(x)$$

where $\varkappa_i$ are some arbitrary numbers – negative for $i \in I_+(x)$ and positive for $i \in I_0(x)$.

Remark 7. In order that $\bar{x}$ may be an A-stationary point of the system ($r \leqq n$) of level $\varrho > 0$, ·it is necessary for the matrix $J(\bar{x})$ to be of the rank less than r. For otherwise an A-direction of $\bar{x}$ could be obtained as in the proof of Theorem 2.

3. METHOD OF B-DIRECTIONS

Now, we shall deal – more concisely already – with solving the problem (1.6). Let x be an arbitrary point. Let us denote by $I_+(x)$ and $I_-(x)$ the sets of indices

$$(3.1) \qquad \begin{aligned} I_+(x) &= \{i \in I \mid h_i(x) = \beta(x)\} \,, \\ I_-(x) &= \{i \in I \mid h_i(x) = -\beta(x)\} \end{aligned}$$ [23])

and by $B(x)$ the set of vectors

$$(3.2) \qquad \begin{aligned} B(x) = \{s \neq 0 \mid \nabla h_i(x)^{\mathrm{T}} s &< 0 \,, \quad i \in I_+(x) \,, \\ \nabla h_i(x)^{\mathrm{T}} s &> 0 \,, \quad i \in I_-(x)\} \,. \end{aligned}$$

The following terms can be defined by analogy to Section 2.1: B-direction of the point x, B-stationary point of the system, normalization set relating to $B(x)$ (the notation $N(x)$ will be used also here). The reader himself will easily formulate and prove the assertions corresponding to all lemmas and remarks of the Section 2.1.

B-stationary point x can be described as follows: No direction leaves it along which it could be possible locally to decrease the value $\beta(x)$. So, the sense of this concept is equivalent to that of stationary point of a function. (When $\nabla\beta(x)$ exists, both concepts are equivalent even by definition: $\nabla\beta(x) = 0$.)

Regarding the problem (1.10) the following is true: x is a B-stationary point of the system if and only if $(x, \beta(x))$ is a $\tilde{\Omega}'$-stationary point of the function x_{n+1} [5]. (The regularity condition is always fulfilled for the set $\tilde{\Omega}'$.)

Concerning the correspondence between the classes of A-stationary and B-stationary points if follows immediately: If $x \in \Omega$ and $B(x) = \emptyset$, then also $A(x) = \emptyset$. (If $x \notin \Omega$, the set $A(x)$ is not defined.)

---

[23]) Likewise for K and L. The coincidence of this notation with that in the method of A-directions will not cause misunderstanding.

Analogically to (2.4), let us denote by $\tilde{\Omega}_{\beta(x)}$ the set

$$\tilde{\Omega}_{\beta(x)} = \{y \mid |h_i(y)| \leqq \beta(x),\ i \in I\} . \tag{3.3}$$

Then the B-direction is a vector directing, in the environment of the point x, "sharply" into the interior of this set.

Now, we describe an *algorithm* for determining a B-stationary point. The scheme of one iteration will be similar to that of Section 2.2.

Let us have a point x and a number $0 < \delta < \frac{1}{2}$. Denote by $I_+(x, \delta)$ and $I_-(x, \delta)$ the sets of indices

$$\begin{aligned} I_+(x, \delta) &= \{i \in I \mid h_i(x) \geqq (1 - \delta)\, \beta(x)\} , \\ I_-(x, \delta) &= \{i \in I \mid h_i(x) \leqq -(1 - \delta)\, \beta(x)\} . \end{aligned} \tag{3.4}$$

By means of them, the sets of vectors $B(x, \delta)$, $N(x, \delta)$ are defined. Solve the *auxiliary problem*

$$\begin{aligned} \max_{(s,\sigma)} \{\sigma \mid \nabla h_i(x)^{\mathrm{T}} s + \|\nabla h_i(x)\|\, \sigma \leqq 0\,, \quad & i \in I_+(x, \delta)\,, \\ -\nabla h_i(x)^{\mathrm{T}} s + \|\nabla h_i(x)\|\, \sigma \leqq 0\,, \quad & i \in I_-(x, \delta)\,, \\ s \in N(x, \delta) \qquad & \} . \end{aligned} \tag{3.5}$$

Further proceed the same way as in Section 2.2 (remembering appropriate changes in the notation) until a B-direction s' of the point x is obtained. Construct a point

$$x' = x + \lambda' s' \tag{3.6}$$

taking λ' so as to get $\beta(x) - \beta(x') \geqq \varepsilon$ or, if impossible, $\lambda' = \lambda^*$ where λ^* is a solution of the problem

$$\min_{\lambda} \{\beta(x + \lambda s') \mid \lambda \geqq 0\} . \tag{3.7}$$

Starting from an initial point x^0 and a number $0 < \delta_0 < \frac{1}{2}$ and applying recursively the procedure described above, we either reach a B-stationary point or obtain an infinite sequence $\{x^k\}$. In the latter case the following assertion is true:

Theorem 3. *If for some $\varrho > 0$ the set $\tilde{\Omega}_\varrho$ is bounded and $x^k \in \tilde{\Omega}_\varrho$ for some $k = k_0$, then there exists a cluster point of the sequence $\{x^k\}$ which is a B-stationary point of the system* (1.1).

The proof is an analogy to that of Theorem 1. We still ask the reader for the final kindness – to formulate and prove by himself "B-pendants" of Remark 6, Theorem 2, and Remark 7.

If there exists a (finite) solution of the minimization problem (1.6), it is obviously a B-stationary point, and this represents the *Chebyshev solution* of the system (1.1).

In one special case — if the system is linear — each B-stationary point of the system is guaranteed to be its Chebyshev solution. This is the subject of the following

Theorem 4. *Each B-stationary point of the system* (1.1) *with* $K = \emptyset$ *is an absolute minimum of the function* $\beta(x)$.

Proof. Let $\bar{x}$ is a B-stationary point. If there were x such that $\beta(x) < \beta(\bar{x})$, the vector $x - \bar{x}$ would be a B-direction of $\bar{x}$. Indeed, because it is

$$\bar{p}_i(x - \bar{x}) = p_i(x) - p_i(\bar{x}) = p_i(x) - \beta(\bar{x}) \leqq \beta(x) - \beta(\bar{x}) < 0$$

for $i \in I_+(\bar{x})$ and

$$\bar{p}_i(x - \bar{x}) = p_i(x) - p_i(\bar{x}) = p_i(x) + \beta(\bar{x}) \geqq -\beta(x) + \beta(\bar{x}) > 0$$

for $i \in I_-(\bar{x})$.

4. REALIZATION OF THE METHODS

4.1. Some normalizations. As we already know, the normalization enables us to make use of the metric properties of the scalar product to obtain the optimal A/B-directions[24]). It is the choice of a normalization set on which the proceeding of the computations essentially depends. Doing this we fall into usual controversy: quality of the direction versus laboriousness of its computation. We will describe four types of normalization sets that we believe to be interesting from the view of practice. First for the method of A-directions:

1) $$N_1(x, \delta) = \{s \mid \|s\| \leqq 1\}\,.$$

This normalization is theoretically the best — the optimality of the A-direction is considered on the basis of angular deviations. In this case the appropriate auxiliary problem (2.9) is a non-linear programming problem (with one nonlinear constraint $\|s\| \leqq 1$). Zoutendijk suggested a special method to solve such a problem [1, section 8.2] [6, section 9-5]. However, his approach seems to be enormously laborious. (Of course, except for the trivial case when $I(x, \delta)$ contains only one index.) Here the approximate methods appear more suitable for practical purposes:

1a) The auxiliary problem is solved by cutting-plane method [9], which is an iterative procedure starting from the approximation $|s_j| \leqq 1$, $j = 1, \ldots, n$ for the constraint $\|s\| \leqq 1$. In every iteration one linear constraint is added to improve the current approximation. It is of advantage that a direction more or less close to the optimal one is available in every iteration of the cutting-plane method. The initial iteration gives the direction optimal in the normalization N_2 (see bellow).

[24]) Read: A-directions or B-directions.

1b) The auxiliary problem is converted into an equivalent quadratic programming problem. Its objective function is replaced by a piecewise linear function and the linear programming problem so obtained (with a great number of variables, of course) is then solved using decomposition principle. The reader is referred for more details to [6, section 9-5].

2) $$N_2(x, \delta) = \{s \mid |s_j| \leqq 1, j = 1, \ldots, n\}.$$

This normalization set leads to linear auxiliary problems with variables subjected to both lower and upper bounds.[25] Such problems can be conveniently solved by dual simplex method. (This is described in many books on linear programming, see e.g. [8].)

3) $$N_3(x, \delta) = \{s \mid \sum_{j=1}^{n} |s_j| \leqq 1\}.$$

The auxiliary problem can be solved as a linear programming problem, of course with enlarged number of variables [6, section 9-5].

4) The fourth normalization set will be defined only for the auxiliary problem (2.9) when all $\|\nabla h_i(x)\| \neq 0$:[26]

$$N_4(x, \delta) = \{s \mid \mu(s) \leqq 1, \|s\| \leqq M\} \tag{4.1}$$

where

$$\mu(s) = \min_i \left\{ \frac{|\nabla h_i(x)^T s|}{\|\nabla h_i(x)\|}, i \in I_+(x, \delta) \cup K_0(x, \delta) \right\}, \tag{4.2}$$

$M \geqq 1$ is a given constant. This normalization set, in contrast to the foregoing ones, depends on x and δ. Let us make sure that it has the properties required by Definition 3:

aa) In virtue of (4.2) and Cauchy-Schwarz inequality it is

$$\mu(s) \leqq \|s\|. \tag{4.3}$$

Taking $\omega_1 \leqq 1$, we can for $s \in A(x, \delta)$ put down

$$\bar{\gamma}(s) = \frac{1}{\|s\|}.$$

Really, we have for $0 \leqq \gamma \leqq \bar{\gamma}(s)$ according to (4.2) and (4.3)

$$\mu(\gamma s) = \gamma \mu(s) \leqq \bar{\gamma}(s)\, \mu(s) = \frac{\mu(s)}{\|s\|} \leqq 1, \quad \|\gamma s\| \leqq \bar{\gamma}(s)\, \|s\| = 1 \leqq M.$$

---

[25]) The variable σ fulfils the inequalities $0 \leqq \sigma \leqq \sqrt{n}$ owing to (2.18).

[26]) Only such auxiliary problems are to be solved.

ab) Take $\omega_2 \geqq M$.

ac) Evident from the setting

$$N_4(x, \delta) = (E_n - \{s \mid \mu(s) > 1\}) \cap \{s \mid \|s\| \leqq M\} .$$

b) The set (4.1) contains the zero element.

Remark 8. If $M = 1$, then $N_4(x, \delta) \equiv N_1(x, \delta)$. It can be readily verified using (4.3).

For practical application of the normalization N_4 the following assertion will be useful:

Lemma 4. *Let A' stands for a set of vectors (s, σ) satisfying the inequalities in (2.9). The sets of optimal solutions of the problems*

$$\Pi_1: \max_{(s,\sigma)} \{\sigma \mid (s, \sigma) \in A',\ s \in N_4(x, \delta)\} ,$$

$$\Pi_2: \max_{(s,\sigma)} \{\sigma \mid (s, \sigma) \in A',\ \sigma \leqq 1,\ \mu(s) = \sigma,\ \|s\| \leqq M\}$$

are identical.

Proof. Each optimal solution (s^*, σ^*) of the problem Π_1 satisfies

$$\sigma^* = \mu(s^*) \leqq 1 , \quad \|s^*\| \leqq M .$$

Therefore it is also a feasible solution of Π_2, i.e. each optimal solution (s^{**}, σ^{**}) of the problem Π_2 fulfils the inequality $\sigma^* \leqq \sigma^{**}$. Conversely, each optimal solution of Π_2 is a feasible solution of Π_1, thus $\sigma^* \geqq \sigma^{**}$.

The contents of Lemma 4 can be expressed briefly as follows: The normalization N_4 in the auxiliary problems (2.9) is *induced* by the constraints

$$\sigma \leqq 1 , \quad \mu(s) = \sigma , \quad \|s\| \leqq M . \tag{4.4}$$

If M is sufficiently large, the constraint $\|s\| \leqq M$ can be omitted in practical computations (when number of iterations is finite). That means, we calculate with a value $M > \max \|s^k\|$ which, of course, need not be known in advance.[27])

An auxiliary problem

$$\max_{(s,\sigma)} \{\sigma \mid (s, \sigma) \in A', \sigma \leqq 1, \mu(s) = \sigma\} \tag{4.5}$$

---

[27]) The sophistry that we suspect here consists in substituting the infinite iterative process by a finite one which can only be said about that it has relaxation property. Every use of this method in practice is then "uncertain" from this point of view. As a rule, a practician is compliant to undergo such a risk, especially if he can, like here, easily recognize reaching or approaching the result (according to the level of $\alpha(x)$).

can conveniently be solved by a special simplex-type method [1, section 8.5]. (Satisfying the condition $\mu(s) = \sigma$ is naturally supplied by the simplex mechanism of the method.)

Remark 9. The normalization N_4 lets considerable freedom in the choice of the optimal A-directions. Exactly speaking: If $M \geqq 1/\eta$, $0 < \eta \leqq 1$, then any direction $\tilde{s} \in A(x, \delta)$ of the form

$$\tilde{s} = \frac{1}{\mu(\bar{s})} \bar{s},$$

where $\|\bar{s}\| = 1$, $\mu(\bar{s}) \geqq \eta$, can be selected as an optimal one. Indeed, the vector $(\tilde{s}, 1)$ is a solution of the auxiliary problem Π_2 because of $\mu(\tilde{s}) = 1$, $(\tilde{s}, 1) \in A'$, and

$$\|\tilde{s}\| = \|\bar{s}\|/\mu(\bar{s}) \leqq 1/\eta \leqq M.$$

In the method of B-directions the first three normalization sets can be used without any change. The fourth one is to be modified in the following way:

$$\tilde{N}_4(x, \delta) = \{s \mid \tilde{\mu}(s) \leqq 1, \|s\| \leqq M\}$$

where

$$\tilde{\mu}(s) = \min_i \left\{ \frac{|\nabla h_i(x)^T s|}{\|\nabla h_i(x)\|}, i \in I_+(x, \delta) \cup I_-(x, \delta) \right\}.$$

This normalization will be induced in the auxiliary problems (3.5) again by the constraints (4.4) – with the same practical consequences as in the former case.

Remark 10. Naturally, the normalizations described can be interchanged during computation. It seems reasonable to apply the normalization N_1 in those iterations where the sets $I(x, \delta)$ or $I(x, 0)$ contain one index only. Otherwise, some other less laborious normalizations should be used.

4.2. Evaluation of λ'. We shall continue using the notation of Sections 2.2 and 3. If ε is sufficiently small, then in practical computations (i.e. with a finite number of iterations) it will be not necessary to determine λ' by the optimal way. All we need is to decrease the value of $\alpha(x)$ or $\beta(x)$. Here holds the same as for the set N_4. Nevertheless, it will be desirable to approach the optimal value of λ as much as possible. At the same time, however, the economy achieved in less number of iterations should not be depreciated by necessarily larger volume of computations. We shall describe one rather simple and general procedure to compute λ' which has proved to be of practical value. First for the method of A-directions:

Take a number $t > 0$ (called *basic step*). Proceed through the points

$$x^{(0)} \equiv x, \tag{4.6}$$

$$x^{(l)} \equiv x + 2^{l-1} \frac{t}{\|s'\|} s', \quad l = 1, 2, \ldots$$

as long as the inequalities

a) $\alpha(x^{(l-1)}) > \alpha(x^{(l)})$,

b) $h_i(x^{(l)}) \geqq 0, \quad i \in I$,

c) $l < \bar{l}$

are true ($\bar{l}$ is a given constant). If for some $l = l'$ this is no more the case, then

1) if $l' > 1$ and a) or b) fails, set

$$\lambda' = 2^{l'-2} \frac{t}{\|s'\|}; \tag{4.7}$$

2) if c) fails, set

$$\lambda' = 2^{l'-1} \frac{t}{\|s'\|}; \tag{4.8}$$

3) if $l' = 1$, then proceed "in the opposite direction" through the points

$$x^{(-m)} \equiv x + 2^{-m} \frac{t}{\|s'\|} s', \quad m = 1, 2, \ldots \tag{4.9}$$

testing the inequalities

a) $\alpha(x^{(-m)}) < \alpha(x)$,

b) $h_i(x^{(-m)}) \geqq 0, \quad i \in I$.

As soon as for some $m = m'$ both the inequalities are satisfied, set

$$\lambda' = 2^{-m'} \frac{t}{\|s'\|}. \tag{4.10}$$ [28]

Usually, we are not interested in the knowledge of the explicit value of λ'. What we need is the point x'. We can obtain it – in the three cases (4.7), (4.8), and (4.10) – according to the formulae respectively

$$x' = \begin{cases} x^{(l'-1)}, \\ x^{(l')}, \\ x^{(-m')}. \end{cases} \tag{4.11}$$

[28]) Theoretically, such m' exists. However, at the computer computation the appropriate λ' could sometimes be reduced to the machinary zero (owing to various reasons each of which, after all, is due to the finite machinary precision). Therefore for programming we recommend to "ensure" this point like this: If no $m' \leqq \bar{m}$ will occur, start the computation newly with an initial point chosen from an environment of the point x; for instance, it could be $x + \tau s'$. Here $\bar{m} > 1$, $\tau > 0$ are given constants.

It is convenient to take in the $(k+1)$th iteration the basic step $t = t_k$ equal to the length of motion made in the kth iteration, i.e.

$$(4.12) \qquad t_k = \|x^k - x^{k-1}\| = \lambda_k \|s^k\| .$$

The reason of this is to achieve certain selfadaptability of the length of the basic step assuring that the numbers l_{k+1}, m_{k+1} [29]) will not become very large. The initial basic step t_0 may be taken as an arbitrary positive number. In virtue of (4.12), (4.7), (4.8), and (4.10) the values of basic steps can be obtained recursively according to the rules

$$(4.13) \qquad t' = \lambda' \|s'\| = \begin{cases} 2^{l'-2} t , \\ 2^{l'-1} t , \\ 2^{-m'} t . \end{cases}$$

The procedure here described can be readily adapted for the method of B-directions: it is sufficient to introduce the function $\beta(x)$ in the inequalities a) and leave out the inequalities b).

4.3. Preparation of the initial point. The calculation according to the method of A-directions can be started from an arbitrary point x^0 if we shall solve – instead of (1.1) – the equivalent system

$$(4.14) \qquad H_i(x) \equiv \zeta_i(x^0)\, h_i(x) , \quad i \in I$$

where

$$\zeta_i(x^0) = \begin{cases} +1 & \text{if} \quad h_i(x^0) \geqq 0 , \\ -1 & \text{if} \quad h_i(x^0) < 0 . \end{cases}$$

In this connection, let us notice that A-stationary points of non-zero level are not invariant relating to "sign" transformations of the system.

As concerns the method of B-directions, every point x^0 is immediately available as a starting point. The notion of B-stationary point is independent on the "signs" of the equations.

4.4. Calculation of function values and derivatives. In the programs of both algorithms the values as well as gradients of the functions h_i are to be calculated. In practice the following three approaches proved satisfactory:

1) Annex subroutines for the function values and partial derivatives.

2) If h_i are polynomials, input them into the computer through a special code (arranging coefficients, indices of variables, and exponents); use this code to compute the values of the polynomials and to obtain the partial derivatives by "machine derivation".

---

[29]) l_{k+1}, m_{k+1} stand for the values l', m' in the $(k+1)$th iteration.

3) Approximate

$$\frac{\partial h_i(x)}{\partial x_j} \approx \frac{\Delta_{x_j} h_i(x)}{\Delta x_j}, \quad 1 \leqq j \leqq n \tag{4.15}$$

using the function values obtained by one of the preceding ways.

4.5. Linear equations in the method of Δ-directions. A priori we can imagine that the direction obtained by solving the auxiliary problem (2.9) with inequalities $\bar{p}_i(s) \geqq 0$ could leave the region Ω. It could be caused by numerical inaccuracy. This danger can be faced by inserting artificial small systematic errors into the auxiliary problem:

$$\bar{p}_i(s) \geqq \eta_i \quad (\eta_i > 0)\,.$$

4.6. Numerical illustrations. In all computations that we shall mention here the value $\delta_0 = 0{\cdot}125$ has been used. The numbers λ_k have been obtained by means of the procedure described in Section 4.2 with the following values of the parameters: $t_0 = 0{\cdot}01$; $l = 10$; $\bar{m} = 10$. Partial derivatives – unless it is said otherwise – were calculated according to 2) in Section 4.4. We will illustrate the behaviour of the algorithms using these examples:

I.[30]

$$\begin{aligned} x_1^2x_2^2 - 2x_1^3 - 5x_2^3 + 10 &= 0 \\ x_1^4 - 8x_2 + 1 &= 0 \end{aligned}$$

II.

$$\begin{aligned} 2x_1^2 - x_2^2 + x_3^2 + 3x_1x_3 + x_1 + 1 &= 0 \\ x_2^2 - 2x_3^2 + x_1x_2 - x_1 + x_2 - x_3 + 2 &= 0 \\ x_1^2 + x_3^2 - 3x_1x_2 + x_2x_3 + x_1 + x_2 - 1 &= 0 \end{aligned}$$

III. (system of 4 linear equations with 4 unknowns [11, section 16, table II.1a])

IV.

$$\begin{aligned} x_3^4 + x_4^3 - 2x_1x_3 + 3x_2 - 11 &= 0 \\ x_2^3 - 3x_1x_4 + x_3x_4 - 2x_1 + 4x_2 - x_4 - 8 &= 0 \\ x_1^2 - 2x_3^2 + x_2x_4 + 3x_1 - x_4 + 6 &= 0 \\ 3x_1^2 + x_2^2 - 2x_4^2 + x_1x_2 - 4x_2x_3 + 5 &= 0 \end{aligned}$$

V.

$$\begin{aligned} x_2^2 + x_3x_7 + x_5 - 3 &= 0 \\ x_3^2 + x_2x_6 + x_1 + x_4 - 4 &= 0 \\ x_1^2 + x_4^2 + x_1x_5 + x_2 + x_3 + x_7 - 6 &= 0 \\ x_2^2 + x_5^2 + x_4x_7 + x_2 + x_3 + x_6 - 6 &= 0 \\ x_3^2 + x_6^2 + x_3x_6 + x_1 + x_4 + x_5 - 6 &= 0 \\ x_4^2 + x_7^2 + x_2x_5 + x_4 + x_5 - 5 &= 0 \\ x_5^2 + x_1x_4 + x_3 + x_6 - 4 &= 0 \end{aligned}$$

[30]) Taken from [10].

In Table 1 some sequences obtained in computing Example II by the method of A-directions with normalization N_2 are presented. The process has converged from the initial point $(10, -10, 15)$ to the root $(0{\cdot}535\,777,\ -2{\cdot}122\,983,\ 0{\cdot}940\,767)$. The

Table 1.

k	$\alpha(x^k)$	$I_+(x^k, \delta_k)$	$I_0(x^k, \delta_k)$	$\sigma_k/\lVert s^k \rVert$	l_k	m_k	t_k
0	$0{\cdot}7860_{10} + 3$ [a]	1* [b]		—	—	—	$0{\cdot}1000_{10} - 1$
1	$0{\cdot}3390_{10} + 3$	1*, 3		0·8978	10		$0{\cdot}5120_{10} + 1$
2	$0{\cdot}9987_{10} + 2$	3*		0·8377	2		$0{\cdot}5120_{10} + 1$
3	$0{\cdot}6535_{10} + 2$	3*		0·6566	1	2	$0{\cdot}1280_{10} + 1$
4	$0{\cdot}5711_{10} + 2$	3*		0·6455	1	2	0·3200
5	$0{\cdot}5298_{10} + 2$	3*	1	0·6434	1	1	0·1600
6	$0{\cdot}4891_{10} + 2$	3*	1	0·3322	3		0·3200
7	$0{\cdot}4695_{10} + 2$	2, 3*		0·3182	1	1	0·1600
8	$0{\cdot}4278_{10} + 2$	2, 3*	1	0·6694	2		0·1600
9	$0{\cdot}1340_{10} + 2$	1*		$0{\cdot}8132_{10} - 1$	8		$0{\cdot}1024_{10} + 2$
10	$0{\cdot}7324_{10} + 1$	3*		0·9425	1	4	0·6400
11	$0{\cdot}2819_{10} + 1$	3*		0·6287	2		0·6400
12	$0{\cdot}1832_{10} + 1$	2*		0·6082	1	1	0·3200
13	$0{\cdot}1217_{10} + 1$	2*		0·9482	1	2	$0{\cdot}8000_{10} - 1$
14	$0{\cdot}1063_{10} + 1$	2*	1	0·9565	1	2	$0{\cdot}2000_{10} - 1$
15	0·9237	2*, 3		0·2270	4		$0{\cdot}8000_{10} - 1$
16	0·8485	2*, 3		0·5958	1	3	$0{\cdot}1000_{10} - 1$
17	0·7919	2*, 3	1	0·7239	2		$0{\cdot}1000_{10} - 1$
18	0·5438	2*		$0{\cdot}9774_{10} - 1$	7		0·3200
19	0·4098	2*		0·9636	1	4	$0{\cdot}2000_{10} - 1$
20	0·3428	2*		0·9657	1	1	$0{\cdot}1000_{10} - 1$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
50	$0{\cdot}1458_{10} - 3$	2*		0·1982	4		$0{\cdot}4883_{10} - 5$
51	$0{\cdot}1384_{10} - 3$	2*	1	0·9699	1	2	$0{\cdot}1221_{10} - 5$
52	$0{\cdot}1368_{10} - 3$	2*, 3		0·1982	2		$0{\cdot}1221_{10} - 5$
53	$0{\cdot}1340_{10} - 3$	2*, 3	1	0·7274	1	1	$0{\cdot}6104_{10} - 6$
54	$0{\cdot}8219_{10} - 4$	2*		0·1052	9		$0{\cdot}7812_{10} - 4$
55	$0{\cdot}5231_{10} - 4$	2*		0·9699	1	4	$0{\cdot}4883_{10} - 5$
56	$0{\cdot}3739_{10} - 4$	2*		0·9699	1	1	$0{\cdot}2441_{10} - 5$
57	$0{\cdot}2993_{10} - 4$	2*		0·9699	1	1	$0{\cdot}1221_{10} - 5$
58	$0{\cdot}2805_{10} - 4$	2*	1	0·9699	1	2	$0{\cdot}3052_{10} - 6$
59	$0{\cdot}2652_{10} - 4$	2*, 3		0·1982	4		$0{\cdot}1221_{10} - 5$
60	$0{\cdot}2513_{10} - 4$	2*, 3	1	0·7274	1	2	$0{\cdot}3052_{10} - 6$
61	$0{\cdot}1720_{10} - 4$	1*		0·1052	8		$0{\cdot}1953_{10} - 4$
62	$0{\cdot}7575_{10} - 5$	1*		0·9752	1	4	$0{\cdot}1221_{10} - 5$

[a]) $q_{10} \pm p = q \,.\, 10^{\pm p}$.

[b]) The asterisk means that the index belongs to the set $I_+(x^k)$.

Table 2.

No	x^0	$\beta(x^0)$	MN	K	$\beta(x^K)$
I	$(-4, 4)$	$0{\cdot}2250_{10}+3$	AN_2	45	$0{\cdot}7063_{10}-5$
			AN_2 [a])	46	$0{\cdot}6795_{10}-5$
			AN_4	30	$0{\cdot}8583_{10}-5$
			BN_2	23	$0{\cdot}1669_{10}-5$
	$(-0{\cdot}1, 0{\cdot}1)$	$0{\cdot}9997_{10}+1$	AN_2	96	$0{\cdot}9924_{10}-5$
			AN_2 [a])	62	$0{\cdot}7838_{10}-5$
			AN_4	64	$0{\cdot}4560_{10}-5$
			BN_2	22	$0{\cdot}5841_{10}-5$
	$(20, -20)$	$0{\cdot}1840_{10}+6$	AN_2	37	$0{\cdot}7987_{10}-5$
			AN_2 [a])	35	$0{\cdot}8106_{10}-5$
			AN_4	74	$0{\cdot}6050_{10}-5$
			BN_2	21	$0{\cdot}6914_{10}-5$
	$(3, -3)$	$0{\cdot}1720_{10}+3$	AN_2	35	$0{\cdot}7093_{10}-5$
			AN_2 [a])	42	$0{\cdot}5633_{10}-5$
			AN_4	32	$0{\cdot}4053_{10}-5$
			BN_2	17	$0{\cdot}6261_{10}+1$ [b])
II	$(-4, 3, 4)$	$0{\cdot}7800_{10}+2$	AN_2	43	$0{\cdot}9719_{10}-5$
			AN_4	35	$0{\cdot}8501_{10}-5$
			BN_2	7	$0{\cdot}4508_{10}-6$
	$(10, -10, 15)$	$0{\cdot}7860_{10}+3$	AN_2	62	$0{\cdot}7575_{10}-5$
			AN_4	51	$0{\cdot}8879_{10}-5$
			BN_2	49	$0{\cdot}9989_{10}-5$
	$(-0{\cdot}45016554,$	$0{\cdot}7646$	AN_2	110	$0{\cdot}1732_{10}-2$ [c])
	$0{\cdot}027210277,$		AN_4	110	$0{\cdot}9521_{10}-3$
	$0{\cdot}70557485)$		BN_2	110	$0{\cdot}3790_{10}-4$
	$(4, 3, -4)$	$0{\cdot}1000_{10}+2$	AN_2	44	$0{\cdot}5700$ [b])
			AN_4	32	$0{\cdot}5700$ [b])
			BN_2	41	$0{\cdot}3019$ [b])
III	$(0, 0, 0, 0)$	$0{\cdot}9000$	AN_2	26	$0{\cdot}6147_{10}-5$
			AN_4	21	$0{\cdot}8244_{10}-5$
			BN_2	46	$0{\cdot}2064_{10}-5$
IV	$(-6, -5, 6, 7)$	$0{\cdot}1685_{10}+4$	AN_2	174	$0{\cdot}6649_{10}-4$
	$(1, 1, -1, -2)$	$0{\cdot}1300_{10}+2$	AN_2	75	$0{\cdot}1259_{10}-1$ [c])
			AN_4	75	$0{\cdot}4731_{10}-1$ [c])
			BN_2	75	$0{\cdot}1262_{10}-4$
V	$x_j^0 = -1$	$0{\cdot}6000_{10}+1$	AN_2	71	$0{\cdot}3653$ [c])
	$x_j^0 = 7$	$0{\cdot}1620_{10}+3$	AN_2	100	$0{\cdot}1667_{10}-4$
			AN_4	40	$0{\cdot}1153_{10}+1$ [c])
			BN_2	100	$0{\cdot}1407_{10}-3$

[a]) Partial derivatives approximated according to (4.15).
[b]) A/B-stationary point of non-zero level.
[c]) Slow convergence.

parameter δ has changed its value only once – in the 9th iteration. The computation of 62 iterations has taken 11 minutes (prints included).[31])

Table 2 contains some information on the solving of Examples I–V. In the column MN the method applied is indicated as $M = A$ or B and the normalization as $N = N_2$ or N_4. Column K contains numbers of iterations carried out. Let us notice that for the method of A-directions the system was always transformed so that it might be $x^0 \in \Omega$ (see Section 4.3), and thus $\alpha(x^0) = \beta(x^0)$. The auxiliary problems were solved by the methods recommended in Section 4.1. The computations took from 80 sec. to 45 min.[31]) Numbers of iterations given in Table 2 are not quite reliable characteristics of laboriousness. In a few cases there have been rather significant differences in average time of one iteration in calculating an example (from the same initial point) by various methods.

4.7. Conclusion. Finally, we will express some not very exact jedgements concerning practical aspects of the methods of A/B-directions which we base either upon the results described in Section 4.6, either upon some other experience.

The normalization N_2 represents a good compromise in the contradiction of demands mentioned in Section 4.1. The normalization N_4 (with sufficiently large M) sometimes gives surprisingly good results, but sometimes creates the optimal directions little deviated from the boundaries of $\Omega_{\alpha(x)}$, which could unfavourably influence the speed of convergence.[32])

The difference approximation of gradients is practically as good for application as the exact gradients (sometimes it is even better).

None of the methods of A/B-directions can be said to be systematically better than the other.

Two properties characterize the convergence behaviour of the methods:

a) the convergence tends to be slow;

b) the process can converge to an A/B-stationary point of non-zero level.

Similar behaviour was observed, e.g., in the method of gradient minimization of the sum of squared residua. Of course, this is algorithmically much simpler. Nevertheless, even such complicated methods as those of A/B-directions might have their justification: Setting aside various intuitive imaginations,[33]) it is sure that the region of convergence to a root (see Theorem 2) is, in general, different from that of other methods (compare, e.g., with [12], [13]).

---

[31]) On the small-size computer National ELLIOTT 803B.

[32]) The normalizations N_1, N_3, and $\tilde{N}_4$ have not been examined.

[33]) For instance: The functions $\alpha(x)$, $\beta(x)$ are "less nonlinear" than $q(x) = \sum[h_i(x)]^2$. There is less number of A/B-stationary points than of stationary points of the function $q(x)$.

The property b) can be — to an extent — neutralized by means of a strategy often recommended in such cases:[34]) to do more computations starting from different (randomly chosen) initial points. By this means more roots, if any, may be obtained.

The programming of the methods of A/B-directions is not difficult if a subroutine for solving the auxiliary problems is available. The programs of both methods are nearly identical.

The author thanks Mr. JOSEF NEDOMA for his valuable comments and Mrs. MILENA DRTINOVÁ for help in programming.

References

[1] *Zoutendijk, G.:* Methods of feasible directions; Elsevier, Amsterdam 1960.

[2] *Altman, M.:* A feasible direction method for solving the non-linear programming problem; Bull. Acad. Polon. Sci., math., astr., phys. *12* (1964), No 1, 43–50.

[3] *Зуховицкий, С. И., Поляк, Р. А., Примак, М. Е.:* Алгорифм для решения задачи выпуклого чебышевского приближения; ДАН СССР *151* (1963), № 1, 27—30.

[4] *Altman, M.:* Stationary points in non-linear programming; Bull. Acad. Polon. Sci., math., astr., phys. *12* (1964), No 1, 29–35.

[5] *Hrouda, J.:* On a classification of stationary points in nonlinear programming; this issue.

[6] *Hadley, G.:* Nonlinear and dynamic programming; Addison-Wesley, Reading 1964.

[7] *Зуховицкий, С. И., Авдеева, Л. И.:* Линейное и выпуклое программирование; Наука, Москва 1964.

[8] *Юдин, Д. Б., Гольштейн, Е. Г.:* Линейное программирование; Физматгиз, Москва 1963.

[9] *Kelley, J. E.:* The cutting-plane method for solving convex programs; J. SIAM *8* (1960), No 4, 703–712.

[10] *Загускин, В. Л.:* Справочник по численным методам решения алгебраических и трансцендентных уравнений; Физматгиз, Москва 1960.

[11] *Фаддеев, Д. К., Фаддеева, В. Н.:* Вычислительные методы линейной алгебры; Физматгиз, Москва 1963.

[12] *Goldstein, A. A.:* Cauchy's method of minimization; Numer. Math. *4* (1962), No 2, 146–150.

[13] *Яковлев, М. Н.:* О некоторых методах решения нелинейных уравнений; Труды матем. инст. В. А. Стеклова *84*, Наука, Москва 1965, 8—40.

[14] *Hrouda, J.:* Řešení soustav nelineárních rovnic; Závěrečná zpráva o úkolu R7.3/66, VÚTECHP, Praha 1966, 6–15.

[34]) For lack of some better one.

Souhrn

DVĚ MINIMAXOVÉ METODY ŘEŠENÍ SOUSTAV NELINEÁRNÍCH ROVNIC

JAROSLAV HROUDA

K řešení soustavy (konečných) rovnic

$$h_i(x) = 0\,, \quad i = 1, \ldots, r\,, \quad x \in E_n$$

se v článku používá variačního principu: minimalizovat funkce

A. $\alpha(x) = \max\limits_i h_i(x)$ za podmínky $h_i(x) \geqq 0$ $(i = 1, \ldots, r)$,

B. $\beta(x) = \max\limits_i |h_i(x)|$.

Tyto extremální úlohy se řeší iteračními metodami založenými na relaxaci hodnot $\alpha(x)$ (s podmínkou $h_i(x) \geqq 0$) nebo $\beta(x)$. Obecně je zaručena konvergence jen k jistým stacionárním bodům (A-stacionární, B-stacionární bod); všechny kořeny soustavy jsou však mezi nimi zahrnuty. Obě metody jsou vypracovány na basi Zoutendijkovy metody přípustných směrů, známé z teorie nelineárního programování. Lze je tedy v podstatě chápat jako metody přípustných směrů řešení minimalizačních úloh A., B. s *nediferencovatelnými* účelovými funkcemi $\alpha(x)$, $\beta(x)$. Stranou hlavní teorie je ukázáno, jak je možno k řešení soustavy rovnic použít nelineárního programování přímo.

Je udána jedna postačující podmínka pro konvergenci ke kořenu. Je-li soustava lineární, je každý její B-stacionární bod jejím čebyševovským přiblížením.

Každá iterace sestává po numerické stránce ze dvou částí:

1. určení směru relaxace (A-směr, B-směr) řešením pomocné extremální úlohy (zpravidla úlohy lineárního programování speciálního typu);

2. určení délky postupu v daném směru (k tomu je navržen jeden jednoduchý algoritmus).

Praktická účinnost nových metod je ilustrována na řešení 5 soustav polynomických rovnic ($2 \leqq n, r \leqq 7$). Numerická pracnost obou metod je nemalá. Přesto mohou mít svůj význam; doplňují totiž — dosti chudé — konvergenční možnosti jiných metod.

Author's address: Jaroslav Hrouda, Výzkumný ústav technicko-ekonomický chemického průmyslu, Štěpánská 15, Praha 2.