Jitka Dupačová

On maximizing a concave function subject to linear constraints by Newton's method

Aplikace matematiky, Vol. 13 (1968), No. 4, 339–355

Persistent URL: http://dml.cz/dmlcz/103178

ON MAXIMIZING A CONCAVE FUNCTION SUBJECT TO LINEAR CONSTRAINTS BY NEWTON'S METHOD

JITKA ŽÁČKOVÁ

(Received August 8, 1967)

1. Newton's method for finding the point at which a function $f(\mathbf{x})$ of several variables attains its maximum (minimum) is defined by the approximation scheme

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n[\mathbf{F}(\mathbf{x}_n)]^{-1} \nabla f(\mathbf{x}_n) . \tag{1}$$

Here $-\mathbf{F}(\mathbf{x})$ denotes the matrix of second-order derivatives of the function $f(\mathbf{x})$ at point $\mathbf{x}$. The α_n's are eligible; they can be chosen all equal to 1 (the classical case) or they can be chosen according to the principle of small steps or according to the principle of steepest ascent.

Newton's method can be adapted for solving nonlinear programming problems. We shall confine ourselves to the problem of maximizing a concave function subject to linear constraints. We shall derive such an adaptation by replacing the gradient direction by Newton's direction in J. B. Rosen's gradient projection method [3] and we shall discuss its properties both from the theoretical (convergence problems) as well as practical point of view (computational improvements, a numerical example). This is the contents of Sect. 2, 3, 4, 5 and 7 of the present paper. At the same time, a program in ALGOL is to appear in the respective part of this journal.

Another adaptation of Newton's method (for maximizing concave functions of a special type constrained to a simplex) has been suggested by Hájek [1, Sect. 5] in connection with a problem in statistical sampling techniques. In Section 5 of the present paper, the convergence of Hájek's method is proved in one-dimensional case, whilst a counter-example is given in the two-dimensional case.

2. (A) Let the convex polyhedral set

$$\mathscr{X} = \{\mathbf{x} \in E_m : \mathbf{a}_i^T\mathbf{x} - b_i = \lambda_i(\mathbf{x}) \geqq 0,\ i = 1, \ldots, k\}$$

be bounded and nonempty.

(B) Let the objective function $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ have continuous second-order derivatives on some open set containing $\mathscr{X}$; let the matrix

$$\mathbf{F}(\mathbf{x}) = -\left(\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right) \quad i, j = 1, \ldots, m$$

be positive definite at all points $\mathbf{x} \in \mathscr{X}$.

The problem is to find the point at which $f(\mathbf{x})$ attains its maximal value on $\mathscr{X}$.

Let $\mathbf{a}_i$, $i \in \{i_1, \ldots, i_q\} \subset \{1, \ldots, k\}$ be linearly independent vectors; set $\mathscr{Q} = \{\mathbf{x} \in E_m : \mathbf{a}_i^T\mathbf{x} - b_i = \lambda_i(\mathbf{x}) = 0,\ i = i_1, \ldots, i_q\}$, let $\mathbf{x}_0$ be any point of $\mathscr{Q}$. (Consequently, $\mathbf{x}_0$ lies in the intersection of q independent hyperplanes.) Suppose, without loss of generality, that $\{i_1, \ldots, i_q\} = \{1, \ldots, q\}$. Denote $\mathbf{F}_0 = \mathbf{F}(\mathbf{x}_0)$. In addition to the usual inner product and norm, denoted by $(.,.)$ and $\|.\|$ respectively, define the inner product

$$(\mathbf{x}, \mathbf{y})_0 = \mathbf{x}^T\mathbf{F}_0\mathbf{y}\,.$$

Denote as $\perp_0$ and $\|.\|_0$ the corresponding to it relation of orthogonality and norm, respectively. Let $\mathbf{n}_i$ be the inward pointing normal of the hyperplane $\mathbf{a}_i^T\mathbf{x} = 0$, i.e., the vector satisfying $\mathbf{n}_i \perp_0 \mathbf{x}$ for all $\mathbf{x}$ in this hyperplane and such that $\|\mathbf{n}_i\|_0 = 1$. Then $\mathbf{F}_0\mathbf{n}_i = \mu_i\mathbf{a}_i$ follows, i.e.,

$$\mathbf{n}_i = \mu_i\mathbf{F}_0^{-1}\mathbf{a}_i\,,$$

where $\mu_i = (\mathbf{a}_i^T\mathbf{F}_0^{-1}\mathbf{a}_i)^{-\frac{1}{2}}$. Because of the possibility of multiplying the inequalities $\mathbf{a}_i^T\mathbf{x} \geqq b_i$ by any positive number we can assume that $\mu_i = 1$, $i = 1, \ldots, q$.

Let $\mathscr{Q}^{(1)}$ be the q-dimensional subspace of E_m spanned by (independent) vectors $\mathbf{n}_1, \ldots, \mathbf{n}_q$; then $\mathscr{Q}^{(2)} = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = 0,\ i = 1, \ldots, q\}$ is obviously a $(m - q)$-dimensional subspace of E_m and it holds $\mathscr{Q}^{(1)} \perp_0 \mathscr{Q}^{(2)}$, $E_m = \mathscr{Q}^{(1)} \oplus \mathscr{Q}^{(2)}$.

Define the matrices $\mathbf{N}_q = [\mathbf{n}_1, \ldots, \mathbf{n}_q]$ and $\mathbf{A}_q = [\mathbf{a}_1, \ldots, \mathbf{a}_q]$. Similarly as in [3], the following lemma holds.

Lemma 1. *The matrix* $\mathbf{P}_q^{(1)} = \mathbf{N}_q(\mathbf{N}_q^T\mathbf{F}_0\mathbf{N}_q)^{-1}\,\mathbf{N}_q^T\mathbf{F}_0$ *is a projection matrix which takes any vector in* E_m *into* $\mathscr{Q}^{(1)}$ *and the matrix*

$$\mathbf{P}_q^{(2)} = \mathbf{E} - \mathbf{P}_q^{(1)} = \mathbf{E} - \mathbf{N}_q(\mathbf{N}_q^T\mathbf{F}_0\mathbf{N}_q)^{-1}\,\mathbf{N}_q^T\mathbf{F}_0 \tag{3}$$

is a projection matrix which takes any vector in E_m *into* $\mathscr{Q}^{(2)}$.

Further, let $\mathbf{N}_{q-1} = [\mathbf{n}_1, \ldots, \mathbf{n}_{q-1}]$ and denote $\mathbf{P}_{q-1}^{(1)}$, $\mathbf{P}_{q-1}^{(2)}$ the corresponding projection matrices. Then

$$\mathbf{P}_q^{(2)} = \mathbf{P}_{q-1}^{(2)} - \frac{\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|^2}{\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|_0^2}\,\mathbf{F}_0 \tag{4}$$

which is easy to verify by multiplying of partitioned matrices.

Some lemmas, which are similar to those in [3] as to the assertions as well as to the proofs, will be introduced now (without proof).

Lemma 2. *If* $\|(\mathbf{N}_q^T\mathbf{F}_0\mathbf{N}_q)^{-1}\| \leqq \eta$, *then* $\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|_0 \geqq \eta^{-\frac{1}{2}}$.

Lemma 3. *If* $\mathbf{x}_0 \in \mathscr{Q}$ *then* $\mathbf{x}_0 + \mathbf{P}_q^{(2)}\mathbf{y} \in \mathscr{Q}$ *for arbitrary* $\mathbf{y} \in E_m$. *If* $\mathbf{x}_0 \in \mathscr{Q}$ *and* $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{z} \in \mathscr{X}$ *then* $\mathbf{A}_q^T\mathbf{z} \geqq 0$.

Let $\gamma(\mathbf{x})$ be the maximal eigenvalue of the matrix

$$\mathbf{F}(\mathbf{x}) = -\left(\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right), \quad \text{let} \quad \gamma \geqq \max_{x \in \mathscr{X}} \gamma(\mathbf{x}) .$$

Further, denote the gradient of $f(\mathbf{x})$ by $\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$.

Lemma 4. *For arbitrary $\mathbf{x}$, $\mathbf{x}_0 \in \mathscr{X}$, the inequality*

$$(\mathbf{x} - \mathbf{x}_0)^T \mathbf{g}(\mathbf{x}_0) - \tfrac{1}{2}\gamma \|\mathbf{x} - \mathbf{x}_0\|^2 \leqq f(\mathbf{x}) - f(\mathbf{x}_0) \leqq (\mathbf{x} - \mathbf{x}_0)^T \mathbf{g}(\mathbf{x}_0)$$

holds.

Theorem 1. *Under assumptions* (A), (B), *the following assertions are valid*:

(i) *If $\mathbf{x}_0 \in \mathscr{X}$ lies on exactly q $(1 \leqq q \leqq m)$ hyperplanes which are linearly independent, say*

$$\mathbf{x}_0 \in \mathscr{Q} = \{\mathbf{x} : \mathbf{a}_i^T \mathbf{x} = b_i,\ i = 1, \ldots, q\} ,$$

then the function $f(\mathbf{x})$ attains its global maximum on $\mathscr{X}$ at the point $\mathbf{x}_0$ if and only if

(5) $$\mathbf{P}_q^{(2)} \mathbf{F}_0^{-1}\, \mathbf{g}(\mathbf{x}_0) = 0 \quad \text{and} \quad (\mathbf{N}_q^T \mathbf{F}_0 \mathbf{N}_q)^{-1}\, \mathbf{N}_q^T\, \mathbf{g}(\mathbf{x}_0) \leqq 0 .$$

(ii) *If $\mathbf{x}_0$ is interior to $\mathscr{X}$, then the function $f(\mathbf{x})$ attains its global maximum on $\mathscr{X}$ at the point $\mathbf{x}_0$ if and only if*

(6) $$\mathbf{F}_0^{-1}\, \mathbf{g}(\mathbf{x}_0) = 0 .$$

Proof. (i) The function $f(\mathbf{x})$ attains its maximum at the point $\mathbf{x}_0$ if and only if

$$-(\mathbf{x} - \mathbf{x}_0)^T\, \mathbf{g}(\mathbf{x}_0) \geqq 0$$

holds for all vectors $\mathbf{x}$ satisfying the inequalities

$$\mathbf{a}_i^T(\mathbf{x} - \mathbf{x}_0) \geqq 0 , \quad i = 1, \ldots, q .$$

But this holds true (see Karlin [2, Theor. B. 3.4]) if and only if there exists a q-dimensional vector $\mathbf{r} \leqq 0$ such that

$$\mathbf{g}(\mathbf{x}_0) = \sum_{i=1}^{q} r_i \mathbf{a}_i ,$$

or, equivalently,

(7) $$\mathbf{F}_0^{-1}\, \mathbf{g}(\mathbf{x}_0) = \mathbf{N}_q \mathbf{r} .$$

Multiplying both sides of (7) by $(\mathbf{N}_q^T \mathbf{F}_0 \mathbf{N}_q)^{-1}\, \mathbf{N}_q^T \mathbf{F}_0$, we get

$$\mathbf{r} = (\mathbf{N}_q^T \mathbf{F}_0 \mathbf{N}_q)^{-1}\, \mathbf{N}_q^T\, \mathbf{g}(\mathbf{x}_0) \leqq 0 ,$$

which completes the proof of (i). The proof of (ii) is obvious.

3. Let us suppose, for the sake of simplicity, that our problem is non-degenerated, i.e., in the polyhedral set $\mathscr{X}$, each $(m - e)$-dimensional face lies in the intersection of exactly e (independent) hyperplanes, $0 < e \leqq m$.

Let us define a finite or infinite sequence $\{\mathbf{x}_n\}$ according to the following rule. Let $\mathbf{x}_0 \in \mathscr{X}$. Denote

$$\mathbf{F}_n = \mathbf{F}(\mathbf{x}_n)\,, \quad \Delta(\mathbf{x}_n) = \mathbf{F}_n^{-1}\,\mathbf{g}(\mathbf{x}_n)\,, \quad \|\mathbf{y}\|_n^2 = \mathbf{y}^T\mathbf{F}_n\mathbf{y}\,, \quad \text{etc.}$$

Let $\mathbf{x}_n$ belong to the intersection $\mathscr{Q}$ of exactly q independent hyperplanes, $1 \leqq q \leqq m$; let us set

$$\mathbf{r} = \mathbf{r}(\mathbf{x}_n) = (\mathbf{N}_q^T\mathbf{F}_n\mathbf{N}_q)^{-1}\,\mathbf{N}_q^T\,\mathbf{g}(\mathbf{x}_n)\,, \tag{8}$$

$$\varrho_{\mathscr{Q}} = \max_{1 \leqq i \leqq q} r_i\,,$$

$$\mathbf{P}_q^{(1)}\,\Delta(\mathbf{x}_n) = \mathbf{N}_q\,\mathbf{r}(\mathbf{x}_n)\,, \quad \mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n) = \Delta(\mathbf{x}_n) - \mathbf{P}_q^{(1)}\,\Delta(\mathbf{x}_n)\,.$$

(Remark that matrix $\mathbf{N}_q$ depends on $\mathbf{x}_n$ and that formulas (8) are in accordance with (3).)

Choose $\eta > 0$ such that $\|(\mathbf{N}^T\,\mathbf{F}(\mathbf{x})\,\mathbf{N})^{-1}\| \leqq \eta$ for all points $\mathbf{x}$ from the boundary of $\mathscr{X}$. (Columns of $\mathbf{N}$ are again the normals of all the hyperplanes which $\mathbf{x}$ belongs to.)

a) Let either $\mathbf{x}_n$ be interior to $\mathscr{X}$ and (6) holds, or $\mathbf{x}_n \in \mathscr{Q}$ and (5) holds. Then the function $f(\mathbf{x})$ attains its maximum on $\mathscr{X}$ at the point $\mathbf{x}_n$ and the sequence terminates.

b) Let either $\mathbf{x}_n$ be interior to $\mathscr{X}$ and $\Delta(\mathbf{x}_n) \neq 0$, or $\mathbf{x}_n \in \mathscr{Q}$ and $\|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n > \max\{0, \frac{1}{2}\varrho_{\mathscr{Q}}\eta^{-\frac{1}{2}}\}$. The algorithm will be defined for the latter case only; the corresponding formulas for the former one follow by replacing $\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)$ by $\Delta(\mathbf{x}_n)$.

Define

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n\mathbf{z}_n \tag{9}$$

where

$$\mathbf{z}_n = \frac{\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)}{\|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n}\,,$$

and where τ_n is chosen in the way described below which ensures both $\mathbf{x}_{n+1} \in \mathscr{X}$ and $f(\mathbf{x}_{n+1}) > f(\mathbf{x}_n)$. According to Lemma 3, $\mathbf{x}_{n+1} \in \mathscr{Q}$ for arbitrary τ. In addition, $\lambda_i(\mathbf{x}_{n+1}) \geqq 0$ for $i = q + 1, \ldots, k$, is required. Denote

$$\tau_i = \frac{\lambda_i(\mathbf{x}_n)}{-\mathbf{a}_i^T\mathbf{z}_n} \quad \text{for} \quad \mathbf{a}_i^T\mathbf{z}_n < 0\,,$$

$$= \infty \qquad \text{for} \quad \mathbf{a}_i^T\mathbf{z}_n \geqq 0\,,$$

$$\tau_n^{(\mathrm{M})} = \min_{q+1 \leqq i \leqq k}\{\tau_i > 0\}\,. \tag{10}$$

Considering that $\mathbf{z}_n \neq 0$ and $\mathscr{X}$ is bounded, we have $\tau_n^{(M)} < \infty$. For $0 < \tau \leqq \tau_n^{(M)}$, we have $\mathbf{x}_{n+1} \in \mathscr{X}$. Now, if we choose,

$$\tau_n = \min\left(\tau_n^{(M)}, \frac{1}{\gamma}\frac{\|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n^3}{\|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|^2}\right) \tag{11}$$

we can easily establish (by means of Lemma 4) that

$$f(\mathbf{x}_n + \tau_n \mathbf{z}_n) - f(\mathbf{x}_n) \geqq \tfrac{1}{2}\tau_n \|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n\,. \tag{12}$$

c) Let $\varrho_2 > 0$ and $\|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n \leqq \frac{1}{2}\varrho_2\eta^{-\frac{1}{2}}$. Suppose for simplicity that $\varrho_2 = r_q$ and define

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n \mathbf{z}_n\,,$$

where

$$\mathbf{z}_n = \frac{\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)}{\|\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)\|_n}$$

and τ_n is again chosen in such a way that $\mathbf{x}_{n+1} \in \mathscr{X}$ and $f(\mathbf{x}_{n+1}) > f(\mathbf{x}_n)$. We have

$$\Delta(\mathbf{x}_n) = \mathbf{P}_q^{(1)}\,\Delta(\mathbf{x}_n) + \mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n) = \sum_{i=1}^{q} r_i \mathbf{n}_i + \mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\,,$$

hence

$$\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n) = r_q \mathbf{P}_{q-1}^{(2)}\mathbf{n}_q + \mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\,.$$

Comparing with (4), we conclude that

$$r_q = \frac{(\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q, \Delta(\mathbf{x}_n))_n}{\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|_n^2} = \frac{\mathbf{a}_q^{\mathrm{T}}\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)}{\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|_n^2}$$

and thus $\mathbf{a}_q^{\mathrm{T}}\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n) > 0$. We have again $\mathbf{x}_{n+1} \in \mathscr{X}$ for any $\tau \in (0, \tau_n^{(M)})$, where $\tau_n^{(M)}$ has the same meaning as above. Further, $\mathbf{z}_n \neq 0$, in view of

$$\|\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)\|_n \geqq r_q\|\mathbf{P}_{q-1}^{(2)}\mathbf{n}_q\|_n - \|\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}_n)\|_n \geqq \tfrac{1}{2}\varrho_2\eta^{-\frac{1}{2}}\,.$$

Choosing

$$\tau_n = \min\left(\tau_n^{(M)}, \frac{1}{\gamma}\frac{\|\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)\|_n^3}{\|\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)\|^2}\right) \tag{13}$$

we can establish by means of Lemma 4 that

$$f(\mathbf{x}_n + \tau_n \mathbf{z}_n) - f(\mathbf{x}_n) \geqq \tfrac{1}{2}\tau_n\|\mathbf{P}_{q-1}^{(2)}\,\Delta(\mathbf{x}_n)\|_n \geqq \tfrac{1}{4}\tau_n\varrho_2\eta^{-\frac{1}{2}}\,. \tag{14}$$

4. Now, let us go into the convergence problem of the suggested algorithm. Logically, two cases can occur. Either (i) there is an infinite subsequence $\{\mathbf{x}_{n_k}\}$ of the sequence $\{\mathbf{x}_n\}$ such that $\tau_{n_k} < \tau_{n_k}^{(M)}$ holds for each its term, or (ii) in the sequence $\{\mathbf{x}_n\}$, $\tau_n = \tau_n^{(M)}$ holds for all n starting from some n_1. First, let us follow the case (i).

(i) Suppose, without loss of generality, that the mentioned subsequence $\{\mathbf{x}_{n_k}\}$ is convergent, say $\mathbf{x}_{n_k} \to \mathbf{x}^*$, and that the points $\mathbf{x}_{n_k}$ are all relative interior points of the same face $\mathscr{S} \cap \mathscr{X}$ of the polyhedral set $\mathscr{X}$, where $\mathscr{S} = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = b_i, i = i_1, \ldots, i_s\}$. Denote $\mathbf{P}_s^{(2)}$ resp. $\mathbf{P}_s^{(1)} = \mathbf{E} - \mathbf{P}_s^{(2)}$ the projection matrices onto subspaces $\mathscr{S}^{(2)} = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = 0, i = i_1, \ldots, i_s\}$ and $\mathscr{S}^{(1)}$ (which is $\perp_0$ – orthogonal to $\mathscr{S}^{(2)}$), respectively. Further, let $\varrho_{\mathscr{S}}(\mathbf{x}) = \max\{0, \max_{1 \leq i \leq s} r_{\mathscr{S},i}(\mathbf{x})\}$, where the vector $\mathbf{r}_{\mathscr{S}}(\mathbf{x})$ satisfies the relation $\mathbf{P}_s^{(1)} \Delta(\mathbf{x}) = \mathbf{N}_s \mathbf{r}_{\mathscr{S}}(\mathbf{x})$, $\mathbf{N}_s = [\mathbf{n}_{i_1}, \ldots, \mathbf{n}_{i_s}]$. Then (12) and (14) give

$$f(\mathbf{x}_{n_{k+1}}) - f(\mathbf{x}_{n_k}) > f(\mathbf{x}_{n_k+1}) - f(\mathbf{x}_{n_k}) > \psi_{n_k}$$

where

$$\psi_j = \tfrac{1}{2}\tau_j \|\mathbf{P}_s^{(2)} \Delta(\mathbf{x}_j)\|_j \quad \text{for} \quad \|\mathbf{P}_s^{(2)} \Delta(\mathbf{x}_j)\|_j > \tfrac{1}{2}\eta^{-\frac{1}{2}} \varrho_{\mathscr{S}}(\mathbf{x}_j)$$

$$= \tfrac{1}{4}\tau_j\, \varrho_{\mathscr{S}}(\mathbf{x}_j)\, \eta^{-\frac{1}{2}} \quad \text{for} \quad \|\mathbf{P}_s^{(2)} \Delta(\mathbf{x}_j)\|_j \leqq \tfrac{1}{2}\eta^{-\frac{1}{2}} \varrho_{\mathscr{S}}(\mathbf{x}_j)\,.$$

(If $\mathscr{S} \cap \mathscr{X} = \mathscr{X}$ then the points $\mathbf{x}_{n_k}$ are interior to $\mathscr{X}$ and $\psi_j = \frac{1}{2}\tau_j\|\Delta(\mathbf{x}_j)\|_j$.)

Let the limit point $\mathbf{x}^*$ belong to $\mathscr{Q} = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = b_i,\ i = i_1, \ldots, i_q\} \subset \mathscr{S}$, let $\mathbf{P}_q^{(1)}$, $\mathbf{P}_q^{(2)}$ be the corresponding projection matrices and $\mathbf{r}$ be the vector of coordinates of $\mathbf{P}_q^{(1)} \Delta(\mathbf{x}^*)$ with respect to the basis $\{\mathbf{n}_{i_1}, \ldots, \mathbf{n}_{i_q}\}$. The sequence $f(\mathbf{x}_{n_k}) \nearrow f(\mathbf{x}^*)$, and because of continuity, $\mathbf{P}_s^{(2)} \Delta(\mathbf{x}_{n_k}) \to \mathbf{P}_s^{(2)} \Delta(\mathbf{x}^*)$ and $\varrho_{\mathscr{S}}(\mathbf{x}_{n_k}) \to \varrho_{\mathscr{S}}(\mathbf{x}^*)$. Moreover, $f(\mathbf{x}^*) - f(\mathbf{x}_{n_0}) > \sum_{k=0}^{\infty} \psi_{n_k}$, hence $\psi_{n_k} \to 0$ and $\mathbf{P}_s^{(2)} \Delta(\mathbf{x}^*) = 0$, $\varrho_{\mathscr{S}}(\mathbf{x}^*) = 0$. Evidently, $\mathbf{P}_q^{(2)} \Delta(\mathbf{x}^*) = 0$, too, and $\Delta(\mathbf{x}^*)$ can be written both as

$$\Delta(\mathbf{x}^*) = \mathbf{P}_q^{(1)} \Delta(\mathbf{x}^*) = \sum_{j=1}^{q} r_j(\mathbf{x}^*)\, \mathbf{n}_{i_j}\,,$$

as well as

$$\Delta(\mathbf{x}^*) = \mathbf{P}_s^{(1)} \Delta(\mathbf{x}^*) = \sum_{j=1}^{s} r_{\mathscr{S},j}(\mathbf{x}^*)\, \mathbf{n}_{i_j}$$

where $r_{\mathscr{S},j}(\mathbf{x}^*) \leqq \varrho_{\mathscr{S}}(\mathbf{x}^*) = 0$. Here $\{i_1, \ldots, i_s\} \subset \{i_1, \ldots, i_q\}$ and from the unique expression of the vector $\Delta(\mathbf{x}^*)$ in the basis $\{\mathbf{n}_{i_1}, \ldots, \mathbf{n}_{i_q}\}$ it follows $r_j(\mathbf{x}^*) \leqq 0$, $j = 1, \ldots, q$. Thus the necessary and sufficient condition for the point $\mathbf{x}^*$ be the maximum point of f on $\mathscr{X}$ is satisfied.

(ii) Let us follow the second case. Let the sequence $\{\mathbf{x}_n\}$ be infinite and let an integer n_1 exist, such that for every $n > n_1$, $\tau_n = \tau_n^{(M)}$ holds. In the original Rosen's paper [3], this alternative was omitted. Though we don't know any example, the possibility of its occurence is not excluded from the logical point of view and can be the cause of the "zigzagging" effect.

Now, let $\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n^{(M)}\mathbf{z}_n$ for $n > n_1$ and either

$$\|\mathbf{P}^{(2)} \Delta(\mathbf{x}_n)\|_n > c\,\varrho(\mathbf{x}_n) \quad \text{or} \quad \|\mathbf{P}^{(2)} \Delta(\mathbf{x}_n)\|_n \leqq c\,\varrho(\mathbf{x}_n)$$

where $c = \frac{1}{2}\eta^{-\frac{1}{2}}$ and the symbols $\mathbf{N}$, $\mathbf{P}^{(2)}$, ϱ refer to the intersection of all hyperplanes that $\mathbf{x}_n$ belongs to.

In the former case, a face of smaller dimension is reached for $\tau_n = \tau_n^{(M)}$. Hence this case cannot occur in an infinite number of steps, as a vertex of the polyhedral set $\mathscr{X}$ would be necessarily reached after a finite number of occurrences.

Hence, it remains to handle the following case: There is $n_0 \geqq n_1$ such that

$$\|\mathbf{P}^{(2)} \Delta(\mathbf{x}_n)\|_n \leqq c\,\varrho(\mathbf{x}_n)\,, \quad \varrho(\mathbf{x}_n) > 0$$

and $\tau_n = \tau_n^{(M)}$ hold for all $n \geqq n_0$ and, moreover, the points of the sequence $\{\mathbf{x}_n\}_{n_0}$ lie in faces of constant dimension $m - s$, $s > 0$.

If the sequence $\{\mathbf{x}_n\}$ possesses more than one point of accumulation then there exists a convergent subsequence $\{\mathbf{x}_{n'}\}$ to each of them and $\liminf_{n'} \tau_{n'}^{(M)} > 0$. From the monotone convergence of the sequences $\{f(\mathbf{x}_{n'})\}$, from (14) and from continuity, we can prove, similarly as in the case (i), that all the points of accumulation of the sequence $\{\mathbf{x}_n\}$ are the maximum points of f on $\mathscr{X}$ – which contradicts the strict concavity of f.

Hence, there is a limit $\mathbf{x}^*$ of the sequence $\{\mathbf{x}_n\}$; let $\mathbf{x}^* \in \mathscr{Q} = \{\mathbf{x} : \mathbf{a}_i^{\mathrm{T}}\mathbf{x} = b_i,\ i = i_1, \ldots, i_q\}$ and let all the points of $\{\mathbf{x}_n\}_{n_0}^{\infty}$ are relative interior to faces $\mathscr{S}_1 \cap \mathscr{X}, \ldots, \mathscr{S}_h \cap \mathscr{X}$ altogether of dimension $m - s$. Denote $\mathbf{P}_{s,e}^{(2)}$ the projection matrix into the subspace $\mathscr{S}_e^{(2)}$, $e = 1, \ldots, h$. (The index $^{(2)}$ has the same meaning as in the definition of $\mathscr{Q}^{(2)}$.) For given e, let $\{\mathbf{x}_{n'}\}$ be the subsequence of $\{\mathbf{x}_n\}$ containing all points $\mathbf{x}_n \in \mathscr{S}_e$; then

$$0 < \|\mathbf{P}_{s,e}^{(2)} \Delta(\mathbf{x}_{n'})\|_{n'} \leqq c\,\varrho_{\mathscr{S}_e}(\mathbf{x}_{n'})\,.$$

Lemma 5. *Suppose $\varrho_{\mathscr{S}_e}(\mathbf{x}^*) = 0$ for some e. Then $\mathbf{x}^*$ is the point at which f attains its maximum value on $\mathscr{X}$.*

Proof. If $\varrho_{\mathscr{S}_e}(\mathbf{x}^*) = 0$ for some $e \in \{1, \ldots, h\}$ then $\mathbf{P}_{s,e}^{(2)} \Delta(\mathbf{x}^*) = 0$ and $\mathbf{P}_q^{(2)} \Delta(\mathbf{x}^*) = 0$ follow. The vector $\Delta(\mathbf{x}^*)$ can be written both as

$$\Delta(\mathbf{x}^*) = \mathbf{P}_q^{(1)} \Delta(\mathbf{x}^*) = \sum_{j=1}^{q} r_j \mathbf{n}_{i_j}\,,$$

as well as

$$\Delta(\mathbf{x}^*) = \mathbf{P}_{s,e}^{(1)} \Delta(\mathbf{x}^*) = \sum_{j=1}^{s} r_{\mathscr{S}_e,j}(\mathbf{x}^*)\,\mathbf{n}_{i_j}$$

where $\mathbf{n}_{i_j}$ are normals to the hyperplanes whose intersection is $\mathscr{S}_e^{(2)}$, $r_{\mathscr{S}_e,j}(\mathbf{x}^*)$ are components of the vector $\mathbf{r}_{\mathscr{S}_e}$ and $r_{\mathscr{S}_e,j}(\mathbf{x}^*) \leqq \varrho_{\mathscr{S}_e}(\mathbf{x}^*) = 0$ holds. Because of the unique expression of $\Delta(\mathbf{x}^*)$ in the basis $\{\mathbf{n}_{i_1}, \ldots, \mathbf{n}_{i_q}\}$, $r_j \leqq 0$, $j = 1, \ldots, q$, follows.

The existence of such e will be proved for $m \leqq 3$.

Theorem 2. *For $m \leqq 3$, the sequence $\{\mathbf{x}_n\}$ defined in Section 3 converges to the point at which f attains its maximum value on $\mathscr{X}$.*

The proof will be given for $m = 3$; the case $m = 2$ can be handled in the same manner as sub 1. below. Moreover, we can consider the case (ii) only. For the non-degenerated problem we have to discuss the following five eventualities:

1. Let $\{\mathbf{x}^*\} = \mathscr{Q}$ be a vertex which is the intersection of edges $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_3$ each of which contains an infinite subsequence of $\{\mathbf{x}_n\}_{n0}^{\infty}$. Now, $\mathbf{P}_q^{(2)}\,\Delta(\mathbf{x}^*) = 0$ and $r_i \leqq 0$, $i = 1, 2, 3$, the return to any edge $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_3$ beeing impossible, because $\mathbf{x}^*$ is the maximum point of f on each set $\mathscr{S}_j \cap \mathscr{X}$, $j = 1, 2, 3$. The point $\mathbf{x}^*$ is the sought solution.

2. Let in the case 1., the infinite subsequences of $\{\mathbf{x}_n\}_{n0}^{\infty}$ can be drawn from the edges $\mathscr{S}_1$, $\mathscr{S}_2$ only. Numerate the respective points in such a way that for $k \geqq k_0 = [(n_0 + 1)/2]$, $\mathbf{x}_{2k} \in \mathscr{S}_1$, $\mathbf{x}_{2k+1} \in \mathscr{S}_2$ and $\mathbf{x}_{2k} \to \mathbf{x}^*$, $\mathbf{x}_{2k+1} \to \mathbf{x}^*$. Let

$$\mathscr{Q} = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = b_i,\ i = 1, 2, 3\}$$
$$\mathscr{S}_1 = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = b_i,\ i = 1, 3\}$$
$$\mathscr{S}_2 = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} = b_i,\ i = 2, 3\}$$
$$\mathscr{S}_0 = \{\mathbf{x} : \mathbf{a}_3^T\mathbf{x} = b_3\}$$

and $\mathbf{P}_1^{(2)}$ be the projection matrix onto the set $\mathscr{S}_0^{(2)} = \{\mathbf{x} : \mathbf{a}_3^T\mathbf{x} = 0\}$. Then we have for $n \geqq n_0$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \tau_n^{(M)} \frac{\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}_n)}{\|\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}_n)\|_n},$$

$$\mathbf{a}_1^T\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}_{2k}) > 0\,, \quad \mathbf{a}_2^T\mathbf{P}_1^{(2)}\,\Delta(x_{2k}) < 0\,,$$

$$\mathbf{a}_1^T\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}_{2k+1}) < 0\,, \quad \mathbf{a}_2^T\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}_{2k+1}) > 0$$

which gives $\mathbf{a}_1^T\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}^*) = \mathbf{a}_2^T\mathbf{P}_1^{(2)}\,\Delta(\mathbf{x}^*) = 0$ for $k \to \infty$, i.e. $\varrho_{\mathscr{S}_1}(\mathbf{x}^*) = \varrho_{\mathscr{S}_2}(\mathbf{x}^*) = 0$ and according to Lemma 5, $\mathbf{x}^*$ is the sought solution.

Similarly, it is possible to prove the convergence in the remaining three cases.

3. $\{\mathbf{x}^*\} = \mathscr{Q}$ is a vertex which is the intersection of two-dimensional faces $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_3$ each of which contains an infinite subsequence of $\{\mathbf{x}_n\}_{n0}^{\infty}$.

4. As in the case 3., but only the faces $\mathscr{S}_1$, $\mathscr{S}_2$ contain infinitely many points of $\{\mathbf{x}_n\}_{n_0}^{\infty}$.

5. $\mathbf{x}^* \in \mathscr{Q}$, where the edge $\mathscr{Q}$ is the intersection of two-dimensional faces $\mathscr{S}_1$, $\mathscr{S}_2$ each of which contains an infinite subsequence of $\{\mathbf{x}_n\}_{n_0}^{\infty}$.

For more-dimensional cases, the discussion is very complicated and it seems impossible to generalize this proof.

5. Up to this point, $\mathbf{x}_n$ has been supposed to lie on exactly q hyperplanes, supposed to be linearly independent. If $\mathbf{x}_n$ lies in addition on further hyperplanes which are linearly dependent on the original set, the degeneracy occurs. It is possible to remove it in a similar manner as in linear programming, i.e., by means of small perturbations of the position of the dependent hyperplanes. The algorithm itself can be adapted in the way suggested in [3].

Two modifications of the suggested algorithm, which reduce the amount of computations required per step and don't affect the convergence, will be mentioned now.

First, the definition of the inner product need not be changed and the inverse $\mathbf{F}_n^{-1}$ need not be computed at every step of the algorithm. For instance, let $\nu > 0$ be an integer given in advance. Suppose that at the n-th step, inner product $(\mathbf{x}, \mathbf{y})_n$ was defined and the matrix $\mathbf{F}_n^{-1}$ established. We can keep this definition of the inner product even in the following $(\nu - 1)$ steps, using $\mathbf{F}_n^{-1}\, \mathbf{g}(\mathbf{x}_{n+j})$ instead of Newton's direction $\Delta(\mathbf{x}_{n+j})$ for $1 \leqq j < \nu$. The projection matrices for $1 \leqq j < \nu$ can then be computed by means of recursion relations of the type (4) (cf. [3]).

Secondly, the use of formulas (11) and (13) for determining τ_n is not very advantageous. Instead, the length τ of each step can be chosen by the method of steepest ascent which means to solve (at least approximately) one-dimensional maximization at each step. If the function $f(\mathbf{x})$ possesses concave second derivatives in all directions, a result stated below as Theorem 3 can be used for solving this problem, i.e., for maximizing $f(\mathbf{x}_0 + \tau \mathbf{z}_0)$ on $0 \leqq \tau \leqq \tau^{(M)}$.

In spite of these modifications, the amount of computations per step is still quite large in the suggested method. It seems that it would be proper to use it for acceleration of computations or for refinement of solution yielded by some of the gradient methods in those cases when the gradient method converges slowly or even does not converge to the correct solution because of rounding errors.

6. Now, we shall investigate an adaptation of the classical Newton's method (i.e., of the scheme (1) with $\alpha_n = 1$) to the constrained problems. We shall start with the following quite simple one-dimensional result, which is nevertheless interesting, as it does not require the initial approximation to lie in the contractivity domain of the respective mapping.

Theorem 3. *Let $f(t)$ be a function of one real variable, let $f''(t)$ exist and be continuous, negative and concave on $\langle a, b \rangle$. Denote t^* the point at which $f(t)$ attains* *its maximal value on* $\langle a, b\rangle$. *Let* $t_0 \in \langle a, b\rangle$. *For* $n \geqq 1$, *define*

$$(15) \qquad t_{n+1} = t_n - \frac{f'(t_n)}{f''(t_n)} \quad \textit{for} \quad t_n - \frac{f'(t_n)}{f''(t_n)} \in \langle a, b\rangle\,,$$

$$= a \qquad \textit{for} \quad t_n - \frac{f'(t_n)}{f''(t_n)} < a\,,$$

$$= b \qquad \textit{for} \quad t_n - \frac{f'(t_n)}{f''(t_n)} > b\,.$$

Then one of the following cases occurs:

1. $\{t_n\}_0^\infty \searrow t^*$.
2. $\{t_n\}_0^\infty \nearrow t^*$.
3. *There is an integer* $n_0 = n_0(t_0)$ *such that*

$$t_0 < t_1 < \ldots < t_{n_0 - 1} < t^* < t_{n_0} \quad \textit{and} \quad \{t_n\}_{n_0}^\infty \searrow t^*\,.$$

4. *There is an integer* $n_0 = n_0(t_0)$ *such that*

$$t_0 > t_1 > \ldots > t_{n_0 - 1} > t^* > t_{n_0} \quad \textit{and} \quad \{t_n\}_{n_0}^\infty \nearrow t^*\,.$$

The proof will be carried out in two steps.

(i) First, we shall prove this auxiliary assertion:

(*) *If the sequence* $\{t_n\}$ *converges then* $\lim_{n\to\infty} t_n = t^*$.

By means of relations (15), a continuous mapping T of $\langle a, b\rangle$ into itself is defined:

$$Tt = t + \max\left\{a - t;\ \min\left[-\frac{f'(t)}{f''(t)}\,;\ b - t\right]\right\}.$$

If $\bar{t} = \lim_{n\to\infty} t_n$ then $\bar{t}$ is the fixed point of T, i.e.,

$$\max\left\{a - \bar{t};\ \min\left[-\frac{f'(\bar{t})}{f''(\bar{t})}\,;\ b - \bar{t}\right]\right\} = 0\,,$$

which can occur exactly in one of the following three cases:

a) $\bar{t} = a$, $f'(a) \leqq 0$,
b) $a < \bar{t} < b$, $f'(\bar{t}) = 0$,
c) $\bar{t} = b$, $f'(b) \geqq 0$.

This proves $\bar{t} = t^*$.

(ii) Let $t^* = b$; then $f'(t) \geqq 0$ for all $t \in \langle a, b\rangle$ and $t_n \leqq t_{n+1} \leqq b$ for arbitrary $t_n \in \langle a, b\rangle$. The sequence $\{t_n\}_0^\infty$ is nondecreasing, bounded from above and $\lim_{n\to\infty} t_n = b$

because of $(*)$. Similarly, if $t = a$ then the sequence $\{t_n\}_0^\infty$ is nonincreasing and $\lim_{n\to\infty} t_n = a$.

If $a < t^* < b$ then $f'(t^*) = 0$ and

$$-\frac{f'(t_n)}{f''(t_n)} = (t_n - t^*)\left[-\frac{f''(t^* + \theta(t_n - t^*))}{f''(t_n)}\right],$$

where $0 \leqq \theta \leqq 1$, thus

$$t_n - \frac{f'(t_n)}{f''(t_n)} - t^* = (t_n - t^*)\left[1 - \frac{f''(t^* + \theta(t_n - t^*))}{f''(t_n)}\right].$$

Suppose for the moment that $t_n - (f'(t_n)/f''(t_n)) \in \langle a, b\rangle$. Then

$$t_{n+1} - t^* = (t_n - t^*)\left[1 - \frac{f''(t^* + \theta(t_n - t^*))}{f''(t_n)}\right].$$

As a consequence of concavity, the function $f''(t)$ is either nondecreasing for $t < t^*$ or nonincreasing for $t > t^*$. Suppose $f''(t)$ is nonincreasing for $t > t^*$; the second case is similar. Then $f''(t^* + \theta(t - t^*))/f''(t) \leqq 1$ for any $t > t^*$. As soon as $t_{n_0} > t^*$, then $t_{n_0} > t_n \geqq t^*$ for all $n > n_0$, the sequence $\{t_n\}_{n_0}^\infty$ is nonincreasing, bounded from below and according to $(*)$ we have $\{t_n\}_{n_0}^\infty \searrow t^*$. Especially, for $t_0 > t^*$ case 1. occurs. For $t_0 < t^*$, $t_0 < t_1$ holds. Suppose $t_0 < t_1 < \ldots t_k < t^*$. Then either $t_k < t_{k+1} \leqq t^*$ or $t_{k+1} > t^*$ and the case 3. occurs with $n_0 = k + 1$. If such an integer n_0 does not exist then

$$1 > 1 - \frac{f''(t^* + \theta(t_n - t^*))}{f''(t_n)} \geqq 0$$

holds for all n and $\{t_n\} \nearrow t^*$.

If $t_n - (f'(t_n)/f''(t_n)) \notin \langle a, b\rangle$, then obviously $n = n_0 - 1$ and the steady convergence starts from the point $t_{n_0} = b$ resp. a.

In [1], a problem concerning statistical sampling techniques is studied and reduced to maximization of the concave function

$$f(\mathbf{p}) = \sum_{h=1}^{H} \left[\sum_{j=1}^{J-1} (a_{jh} - a_{Jh})\, p_j + a_{Jh}\right]^{\frac{1}{2}} \tag{16}$$

on the set

$$\mathscr{P} = \{\mathbf{p} \in E_{J-1} : p_j \geqq 0,\ j = 1, \ldots, J - 1,\ \sum_1^{J-1} p_j \leqq 1\}.$$

The a_{jh}'s are non-negative constants satisfying following conditions:

a) There are no numbers $\lambda_i \geqq 0$, $\sum_{i \neq k} \lambda_i = 1$ such that

$$a_{kh} \leqq \sum_{i \neq h} \lambda_i a_{ih}$$

holds for all $h = 1, \ldots, H$.

b) There is no decomposition of the set $\{1, \ldots, J\}$ in to sets $\mathscr{A}$, $\mathscr{B}$ such that

$$\sum_{h=1}^{H} a_{ih}^{1/2} \geqq \sum_{h=1}^{H} a_{jh} a_{ih}^{-1/2}$$

holds for all $i \in \mathscr{A}$, $j \in \mathscr{B}$.

This special problem is suggested to be solved by means of the following adaptation of the classical Newton's method. Denote

$$\mathbf{F}(\mathbf{p}) = -\left(\frac{\partial^2 f}{\partial p_i \, \partial p_j}\right)_{i,j=1,\ldots,J-1},$$

$$\Delta(\mathbf{p}) = \mathbf{F}(\mathbf{p})^{-1} \nabla f(\mathbf{p})$$

and

$$p_J = 1 - \sum_{1}^{J-1} p_j, \quad \Delta_J(\mathbf{p}) = -\sum_{1}^{J-1} \Delta_j(\mathbf{p}),$$

where $\Delta_j(\mathbf{p})$, $j = 1, \ldots, J-1$, are components of the vector $\Delta(\mathbf{p})$. Let $\mathbf{p}_n \in \mathscr{P}$. If $\mathbf{p}_n + \Delta(\mathbf{p}_n) \in \mathscr{P}$ define $\mathbf{p}_{n+1} = \mathbf{p}_n + \Delta(\mathbf{p}_n)$. Let $\mathbf{p}_n + \Delta(\mathbf{p}_n) \notin \mathscr{P}$ and

$$\Delta_j(\mathbf{p}_n) < 0 \quad \text{for} \quad j \in \mathscr{Z} \subset \{1, \ldots, J\}$$

$$\Delta_j(\mathbf{p}_n) > 0 \quad \text{for} \quad j \in \mathscr{K} \subset \{1, \ldots, J\}.$$

Denote

$$\varkappa_n = \frac{\sum_{j \in \mathscr{Z}} \min\left(p_{jn}; -\Delta_j(\mathbf{p}_n)\right)}{\sum_{j \in \mathscr{K}} \Delta_j(\mathbf{p}_n)},$$

where p_{jn}, $j = 1, \ldots, J-1$, are components of the vector $\mathbf{p}_n$ and $p_{Jn} = 1 - \sum_{j=1}^{J-1} p_{jn}$. For $j = 1, \ldots, J$ define

$$
\begin{aligned}
(17) \qquad p_{jn+1} &= p_{jn} + \Delta_j(\mathbf{p}_n) && \text{if} \quad \Delta_j(\mathbf{p}_n) \leqq 0 \quad \text{and} \quad p_{jn} + \Delta_j(\mathbf{p}_n) \geqq 0, \\
p_{jn+1} &= p_{jn} + \varkappa_n \Delta_j(\mathbf{p}_n) && \text{if} \quad \Delta_j(\mathbf{p}_n) > 0, \\
p_{jn+1} &= 0 && \text{if} \quad p_{jn} + \Delta_j(\mathbf{p}_n) < 0.
\end{aligned}
$$

Then obviously $\mathbf{p}_{n+1} = (p_{1n+1}, \ldots, p_{J-1n+1})^T \in \mathscr{P}$.

The question 1. about the convergence of this algorithm and 2. about its modification in case some derivatives of f fail to exist in some point of $\mathscr{P}$ (which occurs if and only if some $a_{jh} = 0$) remained open. Let us follow the case $J = 2$.

Theorem 4. *Let $J = 2$ in (16) and $a_{jh} > 0$ for all $j = 1, 2$ and $h = 1, \ldots, H$. Then the sequence $\{\mathbf{p}_n\}$ defined by formulas (17) converges to the point $\mathbf{p}^*$ at which the function $f(\mathbf{p})$ attains its maximal value on $\mathscr{P}$.*

Proof. Now, $\mathscr{P} = \langle 0, 1 \rangle$ and formulas (17) coincide with (15). The function f is concave and its second derivative exists at all points in $\langle 0, 1 \rangle$, is continuous and concave. The assumptions of Theorem 3 are satisfied.

If some $a_{jh} = 0$, we can indicate an interval $\mathrm{I} \subset \langle 0, 1 \rangle$ which contains the sought solution p^* and on which the algorithm converges.

Theorem 5. *Let $a_{1h} = 0$, $h \in \mathscr{H}_1$, $a_{1h} > 0$, $h \in \mathscr{H}_2$, $a_{2h} = 0$, $h \in \mathscr{H}_2$, $a_{2h} > 0$, $h \in \mathscr{H}_1$, where $\mathscr{H}_1 \cap \mathscr{H}_2 = \emptyset$, $\mathscr{H}_1, \mathscr{H}_2 \subset \{1, \ldots, H\}$.*

Then the point at which the function

$$f(p) = \sum_{h=1}^{H} [(a_{1h} - a_{2h})\, p + a_{2h}]^{\frac{1}{2}}$$

attains its maximum on $\langle 0, 1 \rangle$ belongs to the interval $\langle \varepsilon_1, 1 - \varepsilon_2 \rangle$ where

$$\varepsilon_1 = \tfrac{1}{2} \quad \text{for} \quad f'(\tfrac{1}{2}) \geqq 0\,, \tag{18}$$

$$\sqrt{\varepsilon_1} = \frac{\sum_{h \in \mathscr{H}_2} \sqrt{a_{1h}}}{\sum_{h \in \mathscr{H}_1} \sqrt{(2a_{2h})} - \sum_{h \notin \mathscr{H}_1 \cup \mathscr{H}_2} \frac{a_{1h} - a_{2h}}{\sqrt{(\frac{1}{2}(a_{1h} + a_{2h}))}}} < \frac{1}{\sqrt{2}} \quad \text{for} \quad f'(\tfrac{1}{2}) < 0\,,$$

$$\varepsilon_2 = \tfrac{1}{2} \quad \text{for} \quad f'(\tfrac{1}{2}) \leqq 0\,, \tag{19}$$

$$\sqrt{\varepsilon_2} = \frac{\sum_{h \in \mathscr{H}_1} \sqrt{a_{2h}}}{\sum_{h \in \mathscr{H}_2} \sqrt{(2a_{1h})} + \sum_{h \notin \mathscr{H}_1 \cup \mathscr{H}_2} \frac{a_{1h} - a_{2h}}{\sqrt{(\frac{1}{2}(a_{1h} + a_{2h}))}}} < \frac{1}{\sqrt{2}} \quad \text{for} \quad f'(\tfrac{1}{2}) < 0\,.$$

Proof. It is desired to find $\varepsilon_1, \varepsilon_2 \leqq \frac{1}{2}$ such that $f'(1 - \varepsilon_2) \leqq 0$ and $f'(\varepsilon_1) \geqq 0$ hold.

a) If $f'(\frac{1}{2}) \leqq 0$ then $\varepsilon_2 = \frac{1}{2}$ can be taken. Let $f'(\frac{1}{2}) > 0$, i.e.,

$$\sum_{h \in \mathscr{H}_2} \sqrt{(2a_{1h})} - \sum_{h \in \mathscr{H}_1} \sqrt{(2a_{2h})} + \sum_{h \notin \mathscr{H}_1 \cup \mathscr{H}_2} \frac{a_{1h} - a_{2h}}{\sqrt{(\frac{1}{2}(a_{1h} + a_{2h}))}} > 0$$

and

$$\sum_{h \in \mathscr{H}_2} \sqrt{(2a_{1h})} + \sum_{h \notin \mathscr{H}_1 \cup \mathscr{H}_2} \frac{a_{1h} - a_{2h}}{\sqrt{(\frac{1}{2}(a_{1h} - a_{2h}))}} > 0 \, .$$

If $\varepsilon \leqq \frac{1}{2}$, then

$$2f'(1 - \varepsilon) = - \frac{1}{\sqrt{\varepsilon}} \sum_{h \in \mathscr{H}_1} \sqrt{a_{2h}} + \sum_{h \notin \mathscr{H}_1} \frac{a_{1h} - a_{2h}}{\sqrt{(a_{1h} + \varepsilon(a_{2h} - a_{1h}))}} \leqq$$

$$\leqq - \frac{1}{\sqrt{\varepsilon}} \sum_{h \in \mathscr{H}_1} \sqrt{a_{2h}} + \sum_{h \in \mathscr{H}_2} \sqrt{(2a_{1h})} + \sum_{h \notin \mathscr{H}_1 \cup \mathscr{H}_2} \frac{a_{1h} - a_{2h}}{\sqrt{(\frac{1}{2}(a_{1h} + a_{2h}))}}$$

and for $\varepsilon = \varepsilon_2$ defined by (19), $f'(1 - \varepsilon_2) \leqq 0$ and $\varepsilon_2 \leqq \frac{1}{2}$ hold.

b) If $f'(\frac{1}{2}) \geqq 0$ then $\varepsilon_1 = \frac{1}{2}$ can be taken. For $f'(\frac{1}{2}) < 0$, the conditions $f'(\varepsilon_1) \geqq 0$ and $\varepsilon_1 \leqq \frac{1}{2}$ can be verified for ε_1 defined by (18) in quite similar manner as in the previous case.

A similar result holds even for $J > 2$. Before stating it, let us recall that the problem (16) and that of maximizing the function

$$f_1(\mathbf{p}) = \sum_{h=1}^{H} \left(\sum_{j=1}^{J} a_{jh} p_j\right)^{\frac{1}{2}} \tag{20}$$

on the set

$$\mathscr{P}_1 = \{\mathbf{p} \in E_j : p_j \geqq 0, j = 1, \ldots, J, \sum_{j=1}^{J} p_j = 1\} \, ,$$

are equivalent.

Theorem 6. *Let*

$$a_{ih_0} = 0 \, , \quad i \in \mathscr{A} \subset \{1, \ldots, J\} \, ,$$

$$a_{jh_0} \neq 0 \, , \quad j \in \mathscr{B} = \{1, \ldots, J\} - \mathscr{A} \, .$$

Then the function $f_1(\mathbf{p})$ does not attain its maximal value on $\mathscr{P}_1$ at such a point $\mathbf{p}^*$ for which $p_j^* = 0$ for some $j \in \mathscr{B}$. Moreover, to any relative interior point $\mathbf{p} \in \mathscr{P}_1$, there is $\varepsilon_0 > 0$ such that

$$f_1(\mathbf{p}^* + \varepsilon(\mathbf{p} - \mathbf{p}^*)) > f(\mathbf{p}^*)$$

holds for $0 < \varepsilon \leqq \varepsilon_0$.

Proof see in [4].

The problem of convergence for $J > 2$ is more complicated. If $a_{jh} > 0$ for all $j = 1, \dots, J$, $h = 1, \dots, H$, then the mapping T of the set $\mathscr{P}$ into itself given by formulas (17) is continuous and according to the Brouwer's fixed point theorem, T possesses at least one fixed point. However, the fixed point of the mapping T need not be the point at which the function $f(\boldsymbol{p})$ attains its maximum on $\mathscr{P}$ and the sequence $\{\boldsymbol{p}_n\}$ given by (17) need not converge to the solution of the problem, as demonstrated by the following example.

Example. Find the maximum of

$$f(p_1, p_2) = \sqrt{(\tfrac{1}{2}p_1 - \tfrac{1}{4}p_2 + \tfrac{1}{2})} + \sqrt{(\tfrac{1}{4}p_2 + \tfrac{1}{4})} + \sqrt{(-\tfrac{1}{9}p_1 + \tfrac{2}{9})}$$

on the set $\mathscr{P} = \{p_1, p_2 : p_1 \geqq 0,\ p_2 \geqq 0,\ p_1 + p_2 \leqq 1\}$. Here is $J = 3$, $H = 3$ and

$$(a_{jh}) = \begin{pmatrix} 1 & \frac{1}{4} & \frac{1}{9} \\ \frac{1}{4} & \frac{1}{2} & \frac{2}{9} \\ \frac{1}{2} & \frac{1}{4} & \frac{2}{9} \end{pmatrix}.$$

Let $p_{10} = 1$, $p_{20} = 0$. Then

$$\nabla f(\boldsymbol{p}_0) = \begin{pmatrix} \frac{1}{12} \\ \frac{1}{8} \end{pmatrix}, \quad \boldsymbol{F}_0 = \begin{pmatrix} \frac{7}{48} & -\frac{1}{32} \\ -\frac{1}{32} & \frac{9}{64} \end{pmatrix},$$

$$\boldsymbol{F}_0^{-1} = \frac{16}{5}\begin{pmatrix} \frac{9}{4} & \frac{1}{2} \\ \frac{1}{2} & \frac{7}{3} \end{pmatrix}, \quad \Delta(\boldsymbol{p}_0) = \begin{pmatrix} \frac{4}{5} \\ \frac{16}{5} \end{pmatrix}, \quad \Delta_3(\boldsymbol{p}_0) = -\frac{4}{5} - \frac{16}{5} = -4 < 0$$

and $\varkappa = 0$. The algorithm terminates in the point $p_{10} = 1$, $p_{20} = 0$. Nevertheless, there is a direction in which $f(\boldsymbol{p})$ does increase on $\mathscr{P}$. Let us have, e.g., $p_1 + p_2 = 1$; denoting $p = p_1$ we get the function $f_2(p) = \frac{1}{2}\sqrt{(3p + 1)} + \frac{5}{6}\sqrt{(2 - p)}$ which does not attain its maximum on the interval $\langle 0, 1 \rangle$ at the point $p = 1$ but at the point $p = \frac{137}{156}$. Consequently, $f(0, 1) < f(\frac{137}{156}, \frac{19}{156})$.

7. The above-mentioned example will be solved by means of the projected Newton's direction method, explained in Sections 2–4. Starting from the point $p_{10} = 1$, $p_{20} = 0$ and determining the length τ according to the principle of steepest ascent, the projected Newton's direction method yields the exact solution already after the first step, as can be easily seen.

Now, for illustration, let us start from an interior point, e.g., from $p_{10} = p_{20} = \frac{1}{3}$. Then

$$f(\boldsymbol{p}_0) = \sqrt{\tfrac{7}{12}} + \sqrt{\tfrac{1}{3}} + \sqrt{\tfrac{5}{27}} = 1.768\,,$$

$$\boldsymbol{g}(\boldsymbol{p}_0) = \begin{pmatrix} 0.1975 \\ 0.0522 \end{pmatrix}, \quad \boldsymbol{F}_1^{-0} = k\begin{pmatrix} 0.4652 & 0.2818 \\ 0.2818 & 1.3385 \end{pmatrix}$$

and

$$\Delta(\boldsymbol{p}_0) = k\begin{pmatrix} 0.1067 \\ 0.1256 \end{pmatrix},$$

where k is a positive constant. The condition $\boldsymbol{p}_0 + \tau\, \Delta(\boldsymbol{p}_0) \in \mathscr{P}$ gives $\tau \leqq \tau^{(M)} = (1/3 \cdot 0.2323) = 1.4349$. The function $f(\boldsymbol{p}_0 + \tau\, \Delta(\boldsymbol{p}_0))$ is increasing in the point $\tau = \tau^{(M)}$ thus we have $\tau_0 = \tau^{(M)}$ and $p_{11} = \frac{1}{3} + 1.4349 \cdot 0.1067 = 0.4864$, $p_{21} = \frac{1}{3} + 1.4349 \cdot 0.1256 = 0.5136$. Now,

$$f(\boldsymbol{p}_1) = 0.6148^{\frac{1}{2}} + 0.3784^{\frac{1}{2}} + 0.1682^{\frac{1}{2}} = 1.8093\,,$$

$$\boldsymbol{g}(\boldsymbol{p}_1) = \begin{pmatrix} 0.18336 \\ 0.04380 \end{pmatrix}, \quad \boldsymbol{F}_1^{-1} = 19.0051 \begin{pmatrix} 0.39818 & 0.25930 \\ 0.25930 & 0.69746 \end{pmatrix},$$

$$\Delta(\boldsymbol{p}_1) = \begin{pmatrix} 1.60346 \\ 1.48411 \end{pmatrix}.$$

The point $\boldsymbol{p}_1 \in \mathscr{Q} = \{\boldsymbol{p} \in E_2 : p_1 + p_2 = 1\}$, the corresponding normal is

$$\boldsymbol{n}_1 = k\, \boldsymbol{F}_1^{-1} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -K \begin{pmatrix} 0.65748 \\ 0.95676 \end{pmatrix}$$

where the value $K = 3.43132$ follows from the condition $\boldsymbol{n}_1^T \boldsymbol{F}_1 \boldsymbol{n}_1 = 1$. The projection

$$\boldsymbol{P}_q^{(2)}\, \Delta(\boldsymbol{p}_1) = \Delta(\boldsymbol{p}_1) - \boldsymbol{n}_1 \boldsymbol{n}_1^T\, \boldsymbol{g}(\boldsymbol{p}_1) = \begin{pmatrix} 0.346 \\ -0.346 \end{pmatrix}$$

and the one-dimensional maximization in the direction $\boldsymbol{P}_q^{(2)}\, \Delta(\boldsymbol{p}_1)$ yields $p_{12} = \frac{137}{156} \doteq 0.8782$, $p_{22} = 0.1218$, $\boldsymbol{p}_2 \in \mathscr{Q}$, $f(\boldsymbol{p}_2) \doteq 1.836$. The gradient

$$\boldsymbol{g}(\boldsymbol{p}_2) = 0.1049 \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

hence $\Delta(\boldsymbol{p}_2) = -0.1049\boldsymbol{n}_1$ and $\boldsymbol{P}_q^{(2)}\, \Delta(\boldsymbol{p}_2) = 0$, $\varrho_{\mathscr{Q}} = -0.1049 < 0$. According to Theorem 1, $\boldsymbol{p}_2$ is the desired solution.

References

[1] *J. Hájek:* Minimalisace nákladů při dosažení předepsané přesnosti současně u několika odhadů, Apl. mat., *7* (1962), p. 405–425.

[2] *S. Karlin:* Mathematical Methods and Theory in Games, Programming and Economics, Vol. I, London 1959.

[3] *J. B. Rosen:* The Gradient Projection Method for Nonlinear Programming. Part I. Linear Constraints, J. Soc. Industr. Appl. Math., *8* (1960), p. 181–217.

[4] *J. Žáčková:* Dva příspěvky k matematickému programování. Kandidátská disertační práce, matematicko-fyzikální fakulta Karlovy university, Praha 1966.

Výtah

POUŽITÍ NEWTONOVY METODY PRO VYHLEDÁNÍ MAXIMA KONKÁVNÍ FUNKCE PŘI LINEÁRNÍCH OMEZENÍCH

JITKA ŽÁČKOVÁ

Newtonovu metodu lze upravit pro řešení úloh nelineárního programování — pro vyhledání maxima konkávní funkce na omezeném konvexním polyedru. Navrhovaná modifikace spočívá v tom, že se používá Newtonova směru namísto gradientu v Rosenově [3] metodě projekce gradientu. V práci je odvozen příslušný algoritmus (odst. 2 a 3) a studuje se jeho konvergence (odst. 4). V odstavci 5 jsou navrženy některé úpravy vhodné pro numerické výpočty; postup výpočtu je ilustrován na příkladě (odst. 7).

Jinou úpravu Newtonovy metody pro vyhledání maxima konkávní funkce speciálního typu na jednotkovém simplexu navrhl Hájek [1, odst. 5] v souvislosti s řešením jedné úlohy pravděpodobnostního výběru. Tato metoda konverguje v jednorozměrném případě, jak je dokázáno v odstavci 5, avšak v odstavci 6 je uveden příklad, kdy ve dvourozměrném případě metoda nekonverguje.

Резюме

ПРИМЕНЕНИЕ МЕТОДА НЬЮТОНА К МАКСИМИЗАЦИИ ВОГНУТОЙ ФУНКЦИИ ПРИ ЛИНЕЙНЫХ ОГРАНИЧЕНИЯХ

ИТКА ЖАЧКОВА

В настоящей статье изучается модификация метода Ньютона для решения задач нелинейного программирования — для максимизации вогнутой функции при линейных ограничениях. Предложенная модификация заключается в замещении градиента направлением Ньютона в градиентном методу Розена [3]. Изучается сходимость соответствующего алгорифма (отдел 4) и в отделе 5 предложены некоторые изменения удобные для практических вычислений. Применение метода показано на примере (отдел 7).

Другой модификацией метода Ньютона для максимизации вогнутой функции частного типа на единичном симплексе занимался Гаек [1, отдел 5] в связи с решением одной задачи вероятностной выборки. В отделе 5 настоящей статьи доказана сходимость этого метода в одномерном случае и в отделе 6 показан двухмерный пример, в котором метод не сходится.

Author's address: Dr. *Jitka Žáčková* C.Sc., Matematicko-fyzikální fakulta KU, Praha 8, Sokolovská 83.