Jaromír Abrham

The multiplex method and its application to concave programming

Czechoslovak Mathematical Journal, Vol. 12 (1962), No. 3, 325–345

Persistent URL: http://dml.cz/dmlcz/100521

ЧЕХОСЛОВАЦКИЙ МАТЕМАТИЧЕСКИЙ ЖУРНАЛ
Математический институт Чехословацкой Академии наук
Т. *12* (87) ПРАГА 20. IX. 1962 г., No 3

THE MULTIPLEX METHOD AND ITS APPLICATION TO CONCAVE PROGRAMMING

JAROMÍR ABRHAM, Praha

(Received July 7, 1960)

The paper consists of two parts. In the first part, Frisch' multiplex method for linear programming is derived and the finiteness of this method is proved. In the second part, an iterative provess based upon this method is given for finding the maximum of a concave function subject to linear inequality constraints. The convergence of this process is proved and an estimate of the error is constructed.

I. THE MULTIPLEX METHOD FOR LINEAR PROGRAMMING

1. Formulation of the Problem, Outline of the Method. We shall consider the following programming problem: To maximize a linear preference function $f(X) = \sum_{j=1}^{N} c_j x_j$ on the set $\mathfrak{M}$ of all finite groups $X = (x_1, \ldots, x_N)$ of real numbers (to be considered as points in the Euclidean N-space $\mathscr{E}_N$) subject to the constraints that

(1) $$\sum_{j=1}^{N} a_{ij} x_j = a_{i0}\,, \quad i = 1, \ldots, m < N$$

(2) $$\bar{x}_j \leqq x_j \leqq \bar{\bar{x}}_j\,, \quad j = 1, \ldots, N$$

where a_{ij}, $i = 1, \ldots, m$, $j = 0, 1, \ldots, N$ are given real constants and $\bar{x}_j < \bar{\bar{x}}_j$, $j = 1, \ldots, N$ are either real constants or symbols $-\infty$ or $+\infty$ (in this case the equality sign in (2) is to be excluded). The rank of the matrix of Eq. (1) will be supposed to equal m and the set $\mathfrak{M}$ will be supposed to possess at least two different elements. When explaining the method for solving the above-stated problem a point $X_1 \in \mathfrak{M}$ will be assumed to be given.

It is well-known that the set $\mathfrak{M}$ is a convex polyhedron contained in $\mathscr{E}_N$, $\dim \mathfrak{M} \leqq N - m$. When using some classical methods (e. g. the simplex method) for solving the problem in question we should have to pass from a given vertex of $\mathfrak{M}$ to another one so as to increase the value of the preference function. Using the multiplex method we shall however take a different algorithm. From a given point of $\mathfrak{M}$ we shall construct another one passing on a line giving in some sense the largest increase of the preference function.

From this point of view it would seem natural to find a vector $\boldsymbol{u} = (u_1, \ldots, u_N)$, i. e. a solution of the linear homogeneous equations

$$(3) \qquad \sum_{j=1}^{N} a_{ij} u_j = 0, \quad i = 1, \ldots, m$$

maximizing $\sum_{j=1}^{N} c_j u_j$ subject to $\sum_{j=1}^{N} u_j^2 = \text{const.}$ (i. e. the length of $\boldsymbol{u}$ being fixed). However the computations necessary to find such a vector are rather difficult, especially when a boundary point of $\mathfrak{M}$ is given. The multiplex method to be explained below seems to be more convenient.

Eqs. (1) being linearly independent we can without any loss of generality suppose that

$$\begin{vmatrix} a_{1,n+1}, \ldots, a_{1N} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m,n+1}, \ldots, a_{mN} \end{vmatrix} \neq 0,$$

where $n = N - m$. Then the variables $x_{n+1}, \ldots, x_N$ can be uniquely expressed in terms of $x_1, \ldots, x_n$ in the form

$$(4) \qquad x_j = b_{j0} + \sum_{k=1}^{N} b_{jk} x_k .$$

Eqs. (4) are obviously valid for $j = n + 1, \ldots, N$. In the following however, it will be useful to generalize Eqs. (4) so that they also hold for $j = 1, \ldots, n$. For, it suffices to put

$$b_{j0} = 0, \quad j = 1, \ldots, n; \quad b_{jk} = \delta_{jk}, \quad j, k = 1, \ldots, n,$$

where δ_{jk} is well-known Kronencker's symbol. Linear equations describing the polyhedron $\mathfrak{M}$ will therefore in the following always be supposed to be given in the form (4) for $j = 1, \ldots, N$ which corresponds to a parametric description of the linear space determined by Eqs. (1). By using Eqs. (4) the considered preference function can easily be transformed into

$$(5) \qquad f(X) = b_{00} + \sum_{k=1}^{n} b_{0k} x_k$$

where $b_{0k} = \sum_{j=1}^{N} c_j b_{jk}$, $k = 0, 1, \ldots, n$.

It would obviously be possible to describe the polyhedron $\mathfrak{M}$ by linear inequalities involving only the "basic variables" $x_1, \ldots, x_n$ in the form

$$\bar{x}_k \leqq x_k \leqq \bar{\bar{x}}_k, \quad k = 1, \ldots, n;$$

$$\bar{x}_j \leqq b_{j0} + \sum_{k=1}^{n} b_{jk} x_k \leqq \bar{\bar{x}}_j, \quad j = n + 1, \ldots, N$$

i. e. it would be possible to take only the space of variables $x_1, \ldots, x_n$ into consideration. In fact, for computational purposes it is convenient to make use of these relations and to construct the vector of the maximal increase of the preference function (5) in this space with some additional linear constraints to be seen from the considerations below.

First, we have to find the above-mentioned vector giving the maximal increase of the function (5). In the following, this vector will be referred to, regardless of its length, as a *gradient*. However, we must remember that some variables may acquire the boundary values and that some of the corresponding components of the gradient can be such that it will not be possible to proceed from the given point in the gradient direction in $\mathfrak{M}$; this difficulty can be avoided by seeking the gradient in some face of the polyhedron $\mathfrak{M}$.

For further considerations some basic notions will be introduced.

A variable x_j as coordinate of a given point $X \in \mathfrak{M}$ will be termed a *bound-attained variable* (*coordinate*) if either $x_j = \bar{x}_j$ or $x_j = \bar{\bar{x}}_j$.

The linear dependency and independency of variables will now be introduced. To each variable x_j we can let there uniquely correspond a vector $\mathbf{b}_j$ with components $b_{j1}, \ldots, b_{jn}$ (vector of coefficients of $x_1, \ldots, x_n$ in Eqs. (4)). Then, *the variables* $x_{i_1}, \ldots, x_{i_\nu}$ *will be said to be linearly dependent* (*independent*) *according to whether the corresponding vectors* $\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_\nu}$ *are linearly dependent* (*independent*). A necessary and sufficient condition that the variables $x_{i_1}, \ldots, x_{i_\nu}$ be linearly dependent (independent) is that the Gram's determinant

$$\begin{vmatrix} M_{i_1,i_1}, \ldots, M_{i_1,i_\nu} \\ \ldots\ldots\ldots\ldots \\ M_{i_\nu,i_1}, \ldots, M_{i_\nu,i_\nu} \end{vmatrix}$$

where $M_{ij} = \sum_{k=1}^{n} b_{ik}b_{jk}$, vanish (be positive) (see e. g. [3], p. 203 and following).

By an *r-dimensional face of* $\mathfrak{M}$ the set of all points $X = (x_1, \ldots, x_N) \in \mathfrak{M}$ will be meant, exactly $n - r$ linearly independent coordinates of which are bound-attained variables lying on a fixed boundary of the corresponding interval in (2).

The operation set of a point $X \in \mathfrak{M}$ is an arbitrary set of linearly independent bound-attained coordinates of X (it need not contain the maximal number of its linearly independent bound-attained coordinates).

Now, let $\mathbf{u} = (u_1, \ldots, u_N)$ be an arbitrary vector of $\mathfrak{M}$, i. e. a vector, for which the relations (3) or equivalent relations

$$u_j = \sum_{k=1}^{n} b_{jk}u_k\,, \quad j = 1, \ldots, N$$

are valid. Then *a component* u_j *of* $\mathbf{u}$ *is said to be admissible with respect to a point* $X = (x_1, \ldots, x_N) \in \mathfrak{M}$ *if*

$$u_j \geqq 0 \text{ when } x_j = \bar{x}_j\,, \quad u_j \leqq 0 \text{ when } x_j = \bar{\bar{x}}_j\,,$$
$$u_j \text{ is arbitrary when } \bar{x}_j < x_j < \bar{\bar{x}}_j\,.$$

A vector $\boldsymbol{u}$ of $\mathfrak{M}$ is admissible with respect to a point $X \in \mathfrak{M}$ if all its components are admissible.

Let X be a point contained in an r-dimensional face of $\mathfrak{M}$ ane let an operation set $\{x_{i_1}, \ldots, x_{i_v}\}$ $(0 \leqq v \leqq n - r)$ be given. We wish to find the gradient of the function (5) in the face of $\mathfrak{M}$ described by the given position of the bound-attained variables $x_{i_1}, \ldots, x_{i_v}$. We have therefore to find real numbers $u_1, \ldots, u_n$ maximizing $\sum_{k=1}^{n} b_{0k}u_k$ and subject to the constraints that

(6) $$\sum_{k=1}^{n} b_{i_s,k}u_k = 0\,, \quad s = 1, \ldots, v\,;$$

(7) $$\sum_{k=1}^{n} u_k^2 = L^2$$

where L is a positive constant. For this purpose we construct the Lagrangian function

$$\psi = \sum_{k=1}^{n} b_{0k}u_k + \tfrac{1}{2}A \sum_{k=1}^{n} u_k^2 + \sum_{s=1}^{v} B_{i_s} \sum_{k=1}^{n} b_{i_s,k}u_k\,,$$

where $A, B_{i_1}, \ldots, B_{i_v}$ are Lagrangian multipliers, and we consider the system of equations $\partial\psi/\partial u_k = 0$, $k = 1, \ldots, n$ together with Eqs. (6) and (7). We get first

(8) $$b_{0k} + Au_k + \sum_{s=1}^{v} B_{i_s}b_{i_s,k} = 0\,, \quad k = 1, \ldots, n\,.$$

Multiplying the k-th equation in (8) $b_{i_t,k}$ and adding over k from 1 to n we get with regard to Eq. (6) the system of equations

(9) $$M_{i_t,0} + \sum_{s=1}^{v} B_{i_s}M_{i_s,i_t} = 0\,, \quad t = 1, \ldots, v$$

for determining the parameters B_{i_s}, $s = 1, \ldots, v$, where $M_{ij} = \sum_{k=1}^{n} b_{ik}b_{jk}$, $i, j = 0, 1, \ldots$ $\ldots, N$. The variables $x_{i_1}, \ldots, x_{i_v}$ being linearly independent the determinant of Eq. (9) is positive and the B_{i_s}'s can therefore be computed uniquely. According to R. Frisch [1], [2] the numbers B_{i_s}, $s = 1, \ldots, v$ will be termed *regression coefficients* of the operation set $\{x_{i_1}, \ldots, x_{i_v}\}$. From Eq. (8) we obtain then $Au_k = -d_k$ where $d_k = b_{0k} + \sum_{s=1}^{v} B_{i_s}b_{i_s,k}$, $k = 1, \ldots, n$ and using Eq. (7) we have

(10) $$A^2L^2 = \sum_{k=1}^{n} d_k^2\,.$$

Let us now put

$$A = -\frac{1}{L}\sqrt{\sum_{k=1}^{n} d_k^2}\,.$$

If $A = 0$ we are not able to find the gradient in question; this case will be discussed in the following section. Suppose therefore that $A < 0$. We have $(\partial^2\psi/\partial u_k \partial u_l) = A\delta_{kl}$, $k, l = 1, \ldots, n$. Let us consider the quadratic form

$$\frac{1}{|A|} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\partial^2 \psi}{\partial u_k \partial u_l} \, \mathrm{d}u_k \mathrm{d}u_l = - \sum_{k=1}^{n} (\mathrm{d}u_k)^2 .$$

Supposing that from the relations $\sum_{k=1}^{n} b_{i_s,k} \, \mathrm{d}u_k = 0$, $s = 1, \ldots, \nu$ e. g. the differentials $\mathrm{d}u_1, \ldots, \mathrm{d}u_\nu$ can be expressed in the form $\mathrm{d}u_s = \sum_{\mu=\nu+1}^{n} p_{s\mu} \, \mathrm{d}u_\mu$, $s = 1, \ldots, \nu$ and substituting this into the quadratic form in question we get the quadratic form

$$- \sum_{k=\nu+1}^{n} (\mathrm{d}u_k)^2 - \sum_{k=1}^{\nu} \Big[\sum_{\mu=\nu+1}^{n} p_{k\mu} \, \mathrm{d}u_\mu \Big]^2$$

which is obviously negative definite. The vector found is therefore the gradient of the preference function (5) in the space of variables $x_1, \ldots, x_n$ with the given operation set. The numbers u_k being proportional to the numbers d_k, $k = 1, \ldots, n$ with a positive proportionality coefficient $|A|$, the numbers $d_1, \ldots, d_n$ can be taken as components of the gradient. In agreement with Frisch' terminology the quantities $d_1, \ldots, d_n$ will be termed *basic direction numbers*. The *general direction numbers* d_j, $j = 1, \ldots, N$ are then defined by

$$d_j = \sum_{k=1}^{n} b_{jk} d_k , \quad j = 1, \ldots, N .$$

For $j = 1, \ldots, N$, it follows that

$$d_j = \sum_{k=1}^{n} b_{jk} \Big(b_{0k} + \sum_{s=1}^{\nu} B_{i_s} b_{i_s,k}\Big) = M_{j0} + \sum_{s=1}^{\nu} B_{i_s} M_{i_s,j} .$$

By this relation and Eq. (9) we obtain again: *If x_j belongs to the operation set then* $d_j = 0$.

The admissibility of direction numbers with respect to a given point $X \in \mathfrak{M}$ is defined in the same way as the admissibility of components of a vector of $\mathfrak{M}$.

2. An Optimality Criterion. *An optimal solution* (*or an optimum*) is a point $X_0 \in \mathfrak{M}$ for which the considered preference function acquires its maximum on $\mathfrak{M}$.

We shall now consider the case when $A = 0$ in Eq. (8). From Eq. (10) we can see that $A = 0$ if and only if $\mathrm{d}_k = 0$, $k = 1, \ldots, n$. The natural question arises whether from the relations that $d_k = 0$, $k = 1, \ldots, n$ we can conclude that an optimal solution has been found. The example below shows that this is not the case.

Example. Let the preference function be $3x_1 + 2x_2 + 5x_3$ and $\mathfrak{M}$ be the set of all points $X = (x_1, \ldots, x_5)$ satisfying

$$\begin{aligned}
x_1 &= x_1 , & 0 &\leqq x_1 \leqq 5 ; \\
x_2 &= x_2 , & 0 &\leqq x_2 \leqq 3 ; \\
x_3 &= x_3 , & 0 &\leqq x_3 ; \\
x_4 &= -1 + 2x_1 - x_2 + x_3 , & -\infty &< x_4 < +\infty ; \\
x_5 &= -5 + 3x_1 + 3x_3 , & x_5 &\leqq 38 .
\end{aligned}$$

The point $Y = (5, 3, \frac{28}{3}, \frac{46}{3}, 38)$ belongs obviously to $\mathfrak{M}$ and $f(Y) = 67\frac{2}{3}$. Choosing the operation set $\{x_1, x_2, x_5\}$ the regression coefficients fulfil

$$\begin{aligned} 3 + \;\; B_1 \qquad\quad + \;\; 3B_5 &= 0\,, \\ 2 \qquad\quad + B_2 \qquad\quad &= 0\,, \\ 24 + 3B_1 \qquad\quad + 18B_5 &= 0 \end{aligned}$$

and therefore $B_1 = 2$, $B_2 = -2$, $B_2 = -\frac{5}{3}$. It follows that $d_1 = d_2 = d_3 = 0$, hence also $d_4 = d_5 = 0$. But $\mathfrak{M}$ contains also the point $Z = (0, 3, \frac{43}{3}, \frac{31}{3}, 38)$ for which $f(Z) = 78\frac{2}{3} > f(Y)$.

In order to be able to establish an optimality criterion the following definition will be useful:

Let $X = (x_1, \ldots, x_N) \in \mathfrak{M}$ and let x_i belong to the operation set. Then the regression coefficient B_i is sign correct if $B_i \geqq 0$ then $x_i = \bar{x}_i$, but $B_i \leqq 0$ when $x_i = \bar{\bar{x}}_i$.

Ler us now denote $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{in})$, $i = 1, \ldots, N$. Then the following theorem is valid:

Let $\hat{X} = (\hat{x}_1, \ldots, \hat{x}_N) \in \mathfrak{M}$ and let $\{x_{i_1}, \ldots, x_{i_v}\}$ be such an operation set of $\hat{X}$ that all regression coefficients $B_{i_1}, \ldots, B_{i_v}$ are sign correct and that the relation

(11)
$$\boldsymbol{b}_0 = -\sum_{s=1}^{v} B_{i_s} \boldsymbol{b}_{i_s}$$

holds. Then $\hat{X}$ in an optimal solution.

Proof. Forming scalar products of Eq. (11) with the vectors b_{i_t}, $t = 1, \ldots, v$ we get Eq. (9). If follows that, if $\boldsymbol{b}_0$ can be expressed as a linear combination of $\boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_v}$, the coefficients in this linear combination can be chosen equal to regression coefficients with opposite signs.

Let now $Y = (y_1, \ldots, y_N)$ be an arbitrary point of $\mathfrak{M}$. The obviously

$$f(Y) = f(\hat{X}) + \sum_{k=1}^{n} b_{0k}(y_k - \hat{x}_k)\,;$$

according to Eq. (11), then

$$f(Y) - f(\hat{X}) = \sum_{k=1}^{n} b_{0k}(y_k - \hat{x}_k) = -\sum_{k=1}^{n} (y_k - \hat{x}_k) \sum_{s=1}^{v} B_{i_s} b_{i_s,k} =$$
$$= -\sum_{s=1}^{v} B_{i_s} \sum_{k=1}^{n} b_{i_s,k}(y_k - \hat{x}_k)\,.$$

Since

$$\sum_{k=1}^{n} b_{i_s,k} y_k = y_{v_s} - b_{i_s,0}, \quad \sum_{k=1}^{n} b_{i_s,k} \hat{x}_k = x_{i_s} - b_{i_s,0}\,, \quad s = 1, \ldots, v\,,$$

we have

$$f(Y) - f(\hat{X}) = -\sum_{s=1}^{v} B_{i_s}(y_{i_s} - \hat{x}_{i_s})\,.$$

As $Y \in \mathfrak{M}$, $\hat{X} \in \mathfrak{M}$, the vector $Y - \hat{X}$ is admissible with respect to $\hat{X}$. All regression coefficients being sign correct, if follows that $B_{i_s}(y_{i_s} - \hat{x}_{i_s}) \geqq 0$, $s = 1, \ldots, v$, hence $f(Y) - f(\hat{X}) \leqq 0$.

From this, the following assertion is easy to establish: *If all regression coefficients are sign correct and all basic direction numbers vanish, then an optimal solution has been found.*

3. The Iterative Process by the Multiplex Method. First, we have to introduce the preference direction number d_0. Putting $j = 0$ in the definition of general direction numbers we get $d_0 = M_{00} + \sum_{s=1}^{v} B_{i_s} M_{i_s,0}$. Expressing B_{i_s}, $s = 1, \ldots, v$ from Eq. (9) by Cramer's rule and substituting into the relation $d_0 = M_{00} + \sum_{s=1}^{v} B_{i_s} M_{i_s,0}$ we find

$$d_0 = \begin{vmatrix} M_{00}, & M_{0,i_1}, & \ldots, & M_{0,i_v} \\ M_{i_1,0}, & M_{i_1,i_1}, & \ldots, & M_{i_1,i_v} \\ \cdots & \cdots & \cdots & \cdots \\ M_{i_v,0}, & M_{i_v,i_1}, & \ldots, & M_{i_v,i_v} \end{vmatrix} : \begin{vmatrix} M_{i_1,i_1}, & \ldots, & M_{i_1,i_v} \\ \cdots & \cdots & \cdots \\ M_{i_v,i_1}, & \ldots, & M_{i_v,i_v} \end{vmatrix} .$$

The preference direction number d_0 is therefore the ratio between two Gram's determinants and hence $d_0 \geqq 0$. Moreover, from the assertions of Sec. 2 it follows that, *if $d_0 = 0$ and if all regression coefficients are sign correct, then an optimal solution has been found.*

For explaining the multiplex process, four logically possible and mutually exclusive cases for the situation in the operation set are to be distinguished. According to the situation described by these cases we shall either change the operation set, or construct a new point in $\mathfrak{M}$ so as to increase the value of the preference function, or we shall see that an optimum has already been found.

The four cases in question are:

I. At least one regression coefficient is sign incorrect.

In cases II–IV all regression coefficients are sign correct.

II. At least one direction number is inadmissible.
III. All direction numbers are admissible, $d_0 > 0$.
IV. $d_0 = 0$. Then an optimal solution has been found.

We shall now discuss the cases I–III:

Case I. Some regression coefficient being sign incorrect we are e. g. even not able to say whether an optimum has been found, especially when a situation as in the example of Sec. 2 arises. From the reasons to be seen below the following instruction is the most convenient: *Drop from the operation set one or more variables having sign incorrect regression coefficient.*

The elements of the new operation set will clearly be linearly independent and we are again in a situation corresponding to some of the four cases in question. When dropping two or more variables we are unable to say anything more about the new operation set. If we drop only one variable, e. g. x_α, we can say that in the operation set thus obtained, the variable x_α (which is bound-attained but does not belong to the new operation set) will have an admissible direction number.

For the proof of this assertion we shall make use of the relation for matrix inversion proved in Sec. 5. If (a_{ik}) is a ν-rowed symmetric regular matrix, $(\bar{a}_{ik})$ a $\nu + 1$-rowed symmetric regular matrix obtained from (a_{ik}) by adding one (e. g. the $\nu + 1$-th) row and column, then

$$(12) \qquad A_{ik} = \bar{A}_{ik} - \frac{\bar{A}_{i,\nu+1}\bar{A}_{h,\nu+1}}{\bar{A}_{\nu+1,\nu+1}}, \quad i, h = 1, \ldots, \nu$$

where A_{ik}, $\bar{A}_{ik}$ are elements of the inverse to (a_{ik}), $(\bar{a}_{ik})$ respectively.

Let $\bar{B}_r$, $r = i_1, \ldots, i_\nu$ be regression coefficients in the original operation set, B_r, $r = i_1, \ldots, i_\nu$, $r \neq \alpha$ the regression coefficients in the new (reduced) operation set, d_j, $j = 1, \ldots, N$ direction numbers computed from the point of view of the reduced operation set. $\bar{M}_{rs}^{-1}$, $r, s = i_1, \ldots, i_\nu$ and M_{rs}^{-1}, $r, s = i_1, \ldots, i_\nu$, $r \neq \alpha \neq s$ will denote the elements of the inverse to the matrix of Gram's determinant of the original and reduced operation set respectively. According to Eq. (12) we have

$$(12') \qquad M_{rs}^{-1} = \bar{M}_{rs}^{-1} - \frac{\bar{M}_{r\alpha}^{-1}\,\bar{M}_{s\alpha}^{-1}}{\bar{M}_{\alpha\alpha}^{-1}}, \quad r, s = i_1, \ldots, i_\nu, \; r \neq \alpha \neq s\,.$$

From Eq. (9) it follows that

$$B_r = -\sum_{s \neq \alpha} M_{s0} M_{rs}^{-1}, \quad r = i_1, \ldots, i_\nu, \; r \neq \alpha,$$

where the index s runs over $i_1, \ldots, i_\nu$. By (12′) we get

$$B_r = -\sum_{s \neq \alpha} M_{s0}\bar{M}_{rs}^{-1} + \frac{\bar{M}_{r\alpha}^{-1}}{\bar{M}_{\alpha\alpha}^{-1}} \sum_{s \neq \alpha} M_{s0}\bar{M}_{s\alpha}^{-1} = -\sum_{s \neq \alpha} M_{s0}\bar{M}_{rs}^{-1} +$$
$$+ \frac{\bar{M}_{r\alpha}^{-1}}{\bar{M}_{\alpha\alpha}^{-1}} \sum_{s} M_{s0}\bar{M}_{s\alpha}^{-1} - M_{\alpha 0}\bar{M}_{r\alpha}^{-1} = \bar{B}_r - \frac{\bar{M}_{r\alpha}^{-1}}{\bar{M}_{\alpha\alpha}^{-1}}\bar{B}_\alpha, \quad r = i_1, \ldots, i_\nu, \; r \neq \alpha\,.$$

Further, it follows that

$$d_\alpha = M_{\alpha 0} + \sum_{s \neq \alpha} M_{\alpha s} B_s = M_{\alpha 0} + \sum_{s \neq \alpha} M_{\alpha s}\bar{B}_s -$$
$$- \frac{\bar{B}_\alpha}{\bar{M}_{\alpha\alpha}^{-1}} \sum_{s \neq \alpha} M_{\alpha s}\bar{M}_{\alpha s}^{-1} = M_{\alpha 0} + \sum_s M_{\alpha s}\bar{B}_s - M_{\alpha\alpha}\bar{B}_\alpha - \frac{\bar{B}_\alpha}{\bar{M}_{\alpha\alpha}^{-1}}(1 - M_{\alpha\alpha}\bar{M}_{\alpha\alpha}^{-1}) = \frac{\bar{B}_\alpha}{\bar{M}_{\alpha\alpha}^{-1}}$$

as $M_{\alpha 0} + \sum_s M_{\alpha s}\bar{B}_s = 0$; therefore

$$(13) \qquad \bar{B}_\alpha = -\bar{M}_{\alpha\alpha}^{-1}\, d_\alpha\,.$$

Then obviously $\bar{M}_{\alpha\alpha}^{-1} > 0$, hence sign $\bar{B}_\alpha = -$ sign d_α.

Of course, it could occasionally happen that, step by step, all elements of the original operation set would be dropped. If the operation set is empty we clearly take

$d_k = b_{0k}$, $k = 1, \ldots, n$ (basic direction numbers) and have $d_j = \sum_{k=1}^{n} b_{0k} b_{jk} = M_{j0}$, $j = 1, \ldots, N$ (general direction numbers). If no operation set containing exactly one element has a sign correct regression coefficient, then all these direction numbers are admissible, as follows from Eq. (13).

Case II. The vector of direction numbers being inadmissible, we cannot pass from the given point in its direction to another point in $\mathfrak{M}$. Then it seems natural to seek the gradient with a larger number of constraints and we come to the following instruction: *Add to the operation set one or more variables having inadmissible direction numbers.*

When adding two or more bound-attained variables to the operation set we are unable to prove any of its properties. Let us now suppose the operation set has been enlarged by adding exactly one bound-attained variable x_α. Then the elements of the new operation set will be linearly independent. If this were not the case, real numbers $c_{\alpha s}$, $s = i_1, \ldots, i_\nu$ would exist such that $b_{\alpha k} = \sum_s c_{\alpha s} b_{sk}$, $k = 1, \ldots, n$. It further follows that

$$d_\alpha = \sum_{k=1}^{n} b_{\alpha k} d_k = \sum_s c_{\alpha s} \sum_{k=1}^{n} b_{sk} d_k = \sum_n c_{\alpha s} d_s \,.$$

But we have $d_s = 0$, $s = i_1, \ldots, i_\nu$, hence $d_\alpha = 0$, i. e. d_α is admissible. Moreover, using Eq. (13) we conclude that, in the enlarged operation set, the regression coefficient corresponding to x_α will be sign correct.

Case III. Let $X_1 = (x_{11}, \ldots, x_{1N})$ be the considered point with the operation set $\{x_{i_1}, \ldots, x_{i_\nu}\}$ and let us put $X_2(\lambda) = X_1 + \lambda \boldsymbol{d}$, where $\boldsymbol{d} = (d_1, \ldots, d_N)$. Then $X_2(\lambda) \in \mathfrak{M}$ for all sufficiently small $\lambda > 0$ and we have

$$f(X_2(\lambda)) - f(X_1) = \lambda \sum_{k=1}^{n} b_{0k} d_k = \lambda \sum_{k=1}^{n} b_{0k}(b_{0k} + \sum_{s=1}^{\nu} B_{i_s} b_{i_s,k}) =$$

$$= \lambda(M_{00} + \sum_{s=1}^{\nu} B_{i_s} M_{0,i_s}) = \lambda d_0 > 0 \,.$$

If $X_2(\lambda) \in \mathfrak{M}$ for all $\lambda > 0$ then obviously $\sup_{X \in \mathfrak{M}} f(X) = +\infty$. In the contrary case we choose λ as large as possible. For, we define the parameters λ_j by the relations

$$\lambda_j = \frac{\bar{\bar{x}}_j - x_{1j}}{d_j} \quad \text{if} \quad d_j > 0 \,, \qquad \lambda_j = \frac{x_{1j} - \bar{x}_j}{-d_j} \quad \text{if} \quad d_j < 0$$

and put $\lambda^* = \min_{\{j : d_j \neq 0\}} \lambda_j$. Then $X_2(\lambda^*) \in \mathfrak{M}$ but $X_2(\lambda)$ non $\in \mathfrak{M}$ for $\lambda > \lambda^*$ and there is at least one index ϱ such that $\lambda^* = \lambda_\rho$. Now, we put $X_2 = X_2(\lambda^*)$ and repeat the whole consideration for X_2. As for the operation set of X_1, we are obviously in the situation described by Case II. The values of $x_{i_1}, \ldots, x_{i_\nu}$ have not changed, hence all regression coefficients are again sign correct, but the direction numbers d_ρ are inadmissible with respect to X_2. The instruction of Case II is therefore to be used first.

Remark. If, in Case I, all direction numbers are admissible and $\sum_{k=1}^{n} |d_k| > 0$, the instruction of Case III may also be applied, but it would probably be possible to find in $\mathfrak{M}$ a direction giving a larger increase of the preference function. If $\sum_{k=1}^{n} |d_k| = 0$ we must clearly proceed according to the instruction for Case I.

A short expository of the multiplex method can also be found in [4].

4. Proof of the Finiteness of the Multiplex Method. In this section we shall prove that by using the multiplex process an optimal solution will be obtained after a finite number of steps. First, we shall introduce the following definition:

Let $\mathbf{v} = (v_1, \ldots, v_N)$ *be a vector of* $\mathfrak{M}$. *Then the length of* $\mathbf{v}$ *is a number* $|\mathbf{v}|$ *satisfying*

$$|\mathbf{v}| = \sqrt{\sum_{k=1}^{n} v_k^2}\,.$$

The length of $\mathbf{v}$ is therefore defined only in terms of its "basic" components $v_1, \ldots, v_n$. We can easily verify that the length thus defined has all the usual properties.

A unit vector of $\mathfrak{M}$ is a vector $\mathbf{u}$ for which $|\mathbf{u}| = 1$.

Let us now suppose that, using the above-described multiplex process, we can construct an infinite sequence of points X_r, $r = 1, 2, \ldots$. As there is only a finite number of faces of $\mathfrak{M}$, there exists at least one face containing infinitely many points of the sequence $\{X_r\}$. Furthermore, the points X_r, X_{r+1} are not contained in the same face for any r. Let us denote

$$\mathbf{u}_r = \frac{X_{r+1} - X_r}{|X_{r+1} - X_r|}\,;$$

$\mathbf{u}_r$ is then a unit vector parallel to and oriented similarly to the vector of direction numbers used to construct X_{r+1} from X_r. Then we can find a finite group $X_p, X_{p+1}, \ldots$ $\ldots, X_{p+q}$ of points of $\{X_r\}$ such that:

1. The points X_p and X_{p+q} are contained in the same face $\hat{\mathfrak{M}}$ of the polyhedron $\mathfrak{M}$.
2. $\mathbf{b}_0\mathbf{u}_p \leqq \mathbf{b}_0\mathbf{u}_{p+j}$, $j = 1, \ldots, q - 1$ where $\mathbf{b}_0\mathbf{u} = \sum_{k=1}^{n} b_{0k}u_k$.

Further, let $X_{p+j+1} = X_{p+j} + \lambda_{p+j}\mathbf{u}_{p+j}$, $j = 0, 1, \ldots, q - 1$; then $\lambda_{p+j} > 0$ for $j = 0, 1, \ldots, q - 1$. We construct now two auxiliary vectors

$$\mathbf{v} = \Big|\sum_{j=0}^{q-1} \lambda_{p+j}\mathbf{u}_{p+j}\Big|^{-1} \cdot \sum_{j=0}^{q-1} \lambda_{p+j}\mathbf{u}_{p+j}\,,$$

$$\mathbf{t} = \Big(\sum_{j=0}^{q-1} \lambda_{p+j}\Big)^{-1} \cdot \sum_{j=0}^{q-1} \lambda_{p+j}\mathbf{u}_{p+j}\,.$$

It follows clearly that $|\mathbf{t}| \leqq |\mathbf{v}| = 1$ and $\mathbf{v}, \mathbf{t}$ are parallel and similarly oriented. Let $\mathbf{w}$ be a unit gradient of the preference function in question in the face $\hat{\mathfrak{M}}$. Then obviously

$\mathbf{b}_0\mathbf{w} \geqq \mathbf{b}_0\mathbf{v} \geqq \mathbf{b}_0\mathbf{t}$. We shall now prove that $\mathbf{b}_0\mathbf{t} \geqq \mathbf{b}_0\mathbf{u}_p$. Let us suppose the opposite inequality $\mathbf{b}_0\mathbf{t} < \mathbf{b}_0\mathbf{u}_p$, i. e.

$$\sum_{j=0}^{q-1} \lambda_{p+j} \cdot \mathbf{b}_0\mathbf{u}_{p+j} < \mathbf{b}_0\mathbf{u}_p \sum_{j=0}^{q-1} \lambda_{p+j}$$

is valid. Then we get by the assumption 2 that

$$\sum_{j=1}^{q-1} \lambda_{p+j}\mathbf{b}_0\mathbf{u}_{p+j} < \sum_{j=1}^{q-1} \lambda_{p+j}\mathbf{b}_0\mathbf{u}_{p+j}$$

which is a contradiction. On the other hand, the relation $\mathbf{b}_0\mathbf{u}_p \geqq \mathbf{b}_0\mathbf{w}$ must hold as $\mathbf{u}_p$ is a unit gradient computed with a less or equal number of constraints as $\mathbf{w}$. It follows that $\mathbf{b}_0\mathbf{w} = \mathbf{b}_0\mathbf{v} = \mathbf{b}_0\mathbf{t} = \mathbf{b}_0\mathbf{u}_p$; but $\mathbf{t} = \alpha\mathbf{v}$, where

$$\alpha = \left| \sum_{j=0}^{g-1} \lambda_{p+j}\mu_{p+j} \right| : \sum_{j=0}^{g-q} \lambda_{p+j} \, .$$

As all vectors $\mathbf{u}_p, \ldots, \mathbf{u}_{p+q-1}$ cannot be parallel it follows that $0 < \alpha < 1$. From the equalities just proved it follows that $\mathbf{b}_0\mathbf{v} = \mathbf{b}_0\mathbf{t} = \alpha\mathbf{b}_0\mathbf{v}$ which implies $\mathbf{b}_0\mathbf{v} = 0$, i. e. $f(X_{p+q}) = f(X_p)$, and this is impossible.

5. An Auxiliary Relation for Matrix Inversion. Let (a_{hi}) be a symmetric regular ν-rowed matrix and A_{ik} be the elements of the inverse of (a_{hi}). Let $(\bar{a}_{hi})$ be a symmetric regular $\nu + 1$-rowed matrix obtained from (a_{hi}) by adding one (e. g. the $\nu + 1$-th) row and column. The elements of the inverse of $(\bar{a}_{hi})$ will be denoted by $\bar{A}_{ih}$. Let D, $\bar{D}$ be determinants of (a_{hi}), $(\bar{a}_{hi})$ respectively. M_{ij} will denote the subdeterminant of D obtained by dropping from D the i-th row and the j-th column. Finally let $D_{ij} = (-1)^{i+j} M_{ij}$, $\bar{D}_{ij}$ be minors of D, $\bar{D}$ respectively. Then, for $i = 1, \ldots, \nu$

$$\bar{D}_{i,\nu+1} = (-1)^{i+\nu+1} \sum_{h=1}^{\nu} a_{h,\nu+1}(-1)^{h+\nu} M_{ih} = -\sum_{h=1}^{\nu} a_{h,\nu+1} D_{ih}$$

and hence

$$\bar{A}_{i,\nu+1} = -\frac{D}{\bar{D}} \sum_{h=1}^{\nu} a_{h,\nu+1} A_{ih} \, .$$

Further clearly $\bar{A}_{\nu+1,\nu+1} = D/\bar{D}$. For $i, h = 1, \ldots, \nu$ the relations

$$\bar{A}_{ih} = A_{ih} + \frac{\bar{A}_{i,\nu+1}\bar{A}_{h,\nu+1}}{\bar{A}_{\nu+1,\nu+1}} \tag{12}$$

are valid.

We shall now prove that the numbers $\bar{A}_{ih}$ satisfying Eq. (12) are elements of the inverse of $(\bar{a}_{hi})$. Multiplying Eq. (12) by a_{ri} $(1 \leqq r \leqq \nu)$ and adding over i from 1 to ν we get

$$\sum_{i=1}^{\nu} a_{ri}\bar{A}_{ih} = \sum_{i=1}^{\nu} a_{ri}A_{ih} + \frac{\bar{A}_{h,\nu+1}}{\bar{A}_{\nu+1,\nu+1}} \sum_{i=1}^{\nu} a_{ri}\bar{A}_{i,\nu+1} \, .$$

Since $\sum_{i=1}^{\nu} a_{ri}\bar{A}_{i,\nu+1} = \sum_{i=1}^{\nu+1} a_{ri}\bar{A}_{i,\nu+1} - a_{r,\nu+1}\bar{A}_{\nu+1,\nu+1} = - a_{r,\nu+1}\bar{A}_{\nu+1,\nu+1}$, we have

$$\sum_{i=1}^{\nu} a_{ri}A_{ih} = \delta_{rh} - \frac{\bar{A}_{h,\nu+1}}{\bar{A}_{\nu+1,\nu+1}} a_{r,\nu+1}\bar{A}_{\nu+1,\nu+1} = \delta_{rh} - a_{r,\nu+1}\bar{A}_{h,\nu+1}$$

and hence

$$\sum_{i=1}^{\nu+1} a_{ri}\bar{A}_{ih} = \delta_{rh} .$$

6. A Numerical Example. We shall apply the multiplex method to the following simple programming problem: Maximize the linear preference function $2x_1 + x_2 + 3x_3 - x_4 + x_5$ on the set $\mathfrak{M}$ described by the conditions

$$(14) \qquad x_1 + x_2 + 2x_3 + x_4 - x_5 = 4 , \quad 2x_1 - x_2 + x_3 - x_4 = 1$$

$$(15) \qquad 0 \leqq x_1 \leqq 5 , \quad 0 \leqq x_2 \leqq 3 , \quad 0 \leqq x_3 , \quad - \infty < x_4 < + \infty , \quad x_5 \leqq 38 .$$

Transforming Eq. (14) into

$$\begin{aligned} x_1 &= x_1 , \\ x_2 &= x_2 , \\ x_3 &= x_3 , \\ x_4 &= - 1 + 2x_1 - x_2 + x_3 , \\ x_5 &= - 5 + 3x_1 + 3x_3 , \end{aligned}$$

we obtain the preference function in the form

$$f(X) = - 4 + 3x_1 + 2x_2 + 5x_3 .$$

We construct now the matrix $\boldsymbol{M} = (M_{ij})$, $i, j = 0, 1, \ldots, 5$ where $M_{ij} = \sum_{k=1}^{3} b_{ik}b_{jk}$. We get

$$\boldsymbol{M} = \begin{Vmatrix} 38 & 3 & 2 & 5 & 9 & 24 \\ 3 & 1 & 0 & 0 & 2 & 3 \\ 2 & 0 & 1 & 0 & -1 & 0 \\ 5 & 0 & 0 & 1 & 1 & 3 \\ 9 & 2 & -1 & 1 & 6 & 9 \\ 24 & 3 & 0 & 3 & 9 & 18 \end{Vmatrix} \quad \begin{matrix} i = 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix}$$

$$j = 0 \; 1 \; 2 \; 3 \; 4 \; 5 .$$

Let the initial point be $X_1 = (1, 1, 1, 1, 1)$; $f(X_1) = 6$. As X_1 has no boundary attained coordinate the operation set is empty and we have the basic direction numbers $d_1 = 3$, $d_2 = 2$, $d_3 = 5$. From the relations

$$d_4 = 2d_1 - d_2 + d_3 , \quad d_5 = 3d_1 + 3d_3$$

we obtain $d_4 = 9$, $d_5 = 24$. We now put $X_2(\lambda) = X_1 + \lambda\boldsymbol{d}$, where $\boldsymbol{d} = (d_1, \ldots, d_5)$, $\lambda > 0$. Hence $X_2(\lambda) = (1 + 3\lambda, 1 + 2\lambda, 1 + 5\lambda, 1 + 9\lambda, 1 + 24\lambda)$ and from (15) we get the following constraints for $\lambda : \lambda \leqq \frac{4}{3}, \lambda \leqq 1, \lambda \leqq \frac{37}{24}$. Hence we choose $\lambda = 1$ and have $X_2 = (4, 3, 6, 10, 25)$, $f(X_2) = 44$. The point X_2 possess one bound-attained coordinate $x_2 = 3$ and the above-computed direction number d_2 is not admissible with respect to X_2. The operation set $\{x_2\}$ will therefore be chosen. The regression coefficient B_2 fulfils $M_{20} + B_2 M_{22} = 0$ hence $B_2 = -2$ which is sign correct. Computing the basic direction numbers $d_k = d_{0k} + B_2 b_{2k}$, $k = 1, 2, 3$ we find $d_1 = 3$, $d_2 = 0$, $d_3 = 5$, hence $d_4 = 11$, $d_5 = 24$. All direction numbers being admissible we put $X_3(\lambda) = (4 + 3\lambda, 3, 6 + 5\lambda, 10 + 11\lambda, 25 + 24\lambda)$. Using (15) we conclude that $\lambda \leqq \frac{1}{3}$, therefore $\lambda = \frac{1}{3}$ is taken and $X_3 = (5, 3, 23/3, 41/3, 33)$, $f(X_3) = 55\frac{1}{3}$. At X_3, the operation set $\{x_1, x_2\}$ is to be taken. Computing B_1, B_2 from equations $M_{10} + B_1 M_{11} + B_2 M_{12} = 0$, $M_{20} + B_1 M_{12} + B_2 M_{22} = 0$, we obtain $B_1 = -3$, $B_2 = -2$. Both regression coefficients being sign correct we find $d_1 = 0$, $d_2 = 0$, $d_3 = 5$, $d_4 = 5$, $d_5 = 15$. As all d_j's are admissible we put $X_4(\lambda) = (5, 3, 23/3 + 5\lambda, 41/3 + 5\lambda, 33 + 15\lambda)$ and by choosing $\lambda = \frac{1}{3}$ the point $X_4 = (5, 3, 28/3, 46/3, 38)$ is obtained and $f(X_4) = 63\frac{2}{3}$. Adding the new bound-attained variable x_5 to the operation set we find the regression coefficients $B_1 = 2$, $B_2 = -2$, $B_5 = -5/3$. As B_1 is sign incorrect we drop x_1 from the operation set. For the reduced operation set $\{x_2, x_5\}$ we have $B_2 = -2$, $B_5 = -4/3$. As both B_2 and B_2 are sign correct we compute $d_1 = -1$, $d_2 = 0$, $d_3 = 1$, $d_4 = -1$, $d_5 = 0$. Then we have $X_5(\lambda) = (5 - \lambda, 3, 28/3 + \lambda, 46/3 - \lambda, 38)$. From the inequalities (15) the condition $\lambda \leqq 5$ follows. Taking $\lambda = 5$ the point $X_5 = (0, 3, 43/3, 31/3, 38)$ is obtained, $f(X_5) = 73\frac{2}{3}$. At X_5, x_1 is a bound-attained variable again (x_1 has moved from its upper to its lower boundary). Taking the operation set $\{x_1, x_2, x_5\}$ we find the sign correct regression coefficients $B_1 = 2$, $B_2 = -2$, $B_5 = -5/3$ and hence $d_j = 0$, $j = 1, \ldots, 5$, i. e. X_5 is an optimum.

II. APPLICATION OF THE MULTIPLEX METHOD TO CONCAVE PROGRAMMING

1. Formulation of the Problem and the Iterative Process. In this chapter we shall construct a modification of the multiplex process for solving the following programming problem: To find the maximum of a concave preference function $\varphi(x_1, \ldots, x_N)$ on the set $\mathfrak{M}$ described by conditions (2) and (4). Substituting into $\varphi(x_1, \ldots, x_N)$ for $x_{n+1}, \ldots, x_N$ from Eq. (4) the preference function will become a function of $x_1, \ldots, x_n$ only. This function will be denoted by $f(x_1, \ldots, x_n)$ or by $f(X)$. The preference function $f(X)$ will obviously be concave, too.

Let us remember that $f(X)$ is concave on $\mathfrak{M}$ if, for any two different points $X, Y \in \mathfrak{M}$ and any two positive numbers α, β such that $\alpha + \beta = 1$ the inequality $f(\alpha X + \beta Y) \geqq \alpha f(X) + \beta f(Y)$ is valid. If the relation $f(\alpha X + \beta Y) > \alpha f(X) + \beta f(Y)$ always holds then $f(X)$ is strictly concave on $\mathfrak{M}$. As an example of a concave (strictly concave) function we can take the function $\mathbf{a}X' - X\mathbf{C}X'$ where $\mathbf{a}X'$ is any linear form and $X\mathbf{C}X'$ any positive semidefinite (positive definite) quadratic form.

As for $f(X)$, the two following assumptions will still be made:

1. We shall suppose that the set $\mathfrak{N}(X^*) = \{X \in \mathfrak{M} : f(X) \geqq f(X^*)\}$ is bounded for every $X^* \in \mathfrak{M}$.

2. The function $f(X)$ will be supposed to be defined and concave and to possess continuous partial derivatives of the first order with respect to all variables on some set $\mathfrak{R}$ described by Eq. (4) and by the inequalities

$$\bar{x}_j - \eta < x_j < \bar{\bar{x}}_j + \eta\,, \quad j = 1, \ldots, N\,,$$

where η is an arbitrarily small positive number.

The set $\mathfrak{M}$ will again be supposed to possess at least two different elements.

The symbol $\nabla f(X)$ will denote the n-vector with components $\partial f/\partial x_1, \ldots, \partial f/\partial x_n$, the derivatives being computed at the point X.

We are now in a position to explain the iterative process for finding the maximum of $f(X)$ on $\mathfrak{M}$. To this purpose we shall construct a sequence $\{X_r\}$ of points of $\mathfrak{M}$ and then prove that

$$\lim_{r \to \infty} f(X_r) = \max_{X \in \mathfrak{M}} f(X)\,.$$

The sequence $\{X_r\}$ will be constructed by induction. Let X_1 be any point of $\mathfrak{M}$ and suppose the point $X_r = (x_{r1}, \ldots, x_{rN})$ $(r \geqq 1)$ has already been found. Then the auxiliary linear function

$$g_r(X) = \nabla f(X_r)(X - X_r) \equiv \sum_{k=1}^{n} \frac{\partial f(X_r)}{\partial x_k}(x_k - x_{rk})$$

will be taken and we apply the multiplex process for determining the maximum of $g_r(X)$ on $\mathfrak{M}$ in the following way: We find an operation set of X_r and classify it according to four fundamental cases described in Chapter I, Sec. 4. If we are in a situation corresponding to Case I or II we proceed according to the corresponding instruction. When in a situation corresponding to Case III (this situation and the situation of Case IV will be in the following expressed by saying that *the operation set is admissible*) we compute all direction number. Let $\mathbf{d}(X_r) = (d_1(X_r), \ldots, d_N(X_r))$ be the vector of these direction numbers. Then we put

$$X_{r+1} = X_r + \lambda_r \mathbf{d}(X_r)\,,$$

where λ_r maximizes the function $f(X_r + \lambda \mathbf{d}(X_r))$ of one real variable λ subject to $X_r + \lambda \mathbf{d}(X_r) \in \mathfrak{M}$.

We will now discuss the situation corresponding to Case IV, i. e. the situation when all regression coefficients are sign correct an all direction numbers vanish. Let us suppose that this situation occurs at some point $X_0 \in \mathfrak{M}$. Then the linear function $g_0(X) = \nabla f(X_0)(X - X_0)$ acquires obviously its maximum at X_0, i. e.

$$\max_{X \in \mathfrak{M}} g_0(X) = 0$$

which is equivalent to the condition that $g_0(X) \leqq 0$ for all $X \in \mathfrak{M}$. We shall now prove the following statement.

A necessary and sufficient condition for $f(X)$ to acquire its maximum on $\mathfrak{M}$ at the point X_0 is that $g_0(X) \leqq 0$ for all $X \in \mathfrak{M}$.

As for the sufficiency of this condition, let us remember that the linear function $z = f(X_0) + \nabla f(X_0)(X - X_0)$ represents the tangent superplane of the supersurface $z = f(X)$ at the point X_0. From the concavity property of $f(X)$ we conclude that the inequality $f(X) \leqq f(X_0) + \nabla f(X_0)(X - X_0) = f(X_0) + g_0(X)$ holds for all $X \in \mathfrak{M}$. If $g_0(X) \leqq 0$ for all $X \in \mathfrak{M}$ we thus obtain that, for every $X \in \mathfrak{M}$, $f(X) \leqq \leqq f(X_0)$.

Let us now suppose that

$$f(X_0) = \max_{X \in \mathfrak{M}} f(X) .$$

Then the relation

$$0 \geqq \frac{\mathrm{d}}{\mathrm{d}t} f(X_0 + t(X - X_0))|_{t=0} = \nabla f(X_0)(X - X_0) = g_0(X)$$

is valid for every $X \in \mathfrak{M}$.

2. Proof of the Convergence of the Iterative Process. From the assumptions made in Sec. 1 follows the existence of at least one point $Y \in \mathfrak{M}$ such that $f(Y) = \max_{X \in \mathfrak{M}} f(X)$. We shall now prove that

$$\lim_{r \to \infty} f(X_r) = \max_{X \in \mathfrak{M}} f(X) .$$

Let $\{X_{r_s}\}$ be a subsequence of $\{X_r\}$. Then the sequence $\{X_{r_s}\}$ is *homogeneous* if the following conditions are fulfilled:

1. The sequence $\{X_{r_s}\}$ is convergent.
2. All points X_{r_s} are contained in the same face of $\mathfrak{M}$.
3. For $s = 1, 2, \ldots$, the point X_{r_s+1} is constructed from X_{r_s} by using the same operation set $\{x_{i_1}, \ldots, x_{1_v}\}$.

A homogeneous sequence $\{X_{r_s}\}$ is *regular* if the operation set $\{x_{i_1}, \ldots, x_{i_v}\}$ is also admissible for the point $X_0 = \lim_{s \to \infty} X_{r_s}$.

Let $\{X_{r_s}\}$ be any homogeneous sequence and $\lim_{s \to \infty} X_{r_s} = X_0$. The regression coefficients being obviously continuous functions of $X \in \mathfrak{M}$, we have

$$\lim_{s \to \infty} B_{i_j}(X_{r_s}) = B_{i_j}(X_0) , \quad j = 1, \ldots, v .$$

As the operation set in question is admissible at X_{r_s}, all regression coefficients $B_{i_j}(X_{r_s})$, and therefore also all $B_{i_j}(X_0)$'s, are sign correct. Hence we conclude that the homogeneous sequence $\{X_{r_s}\}$ is regular if and only if all direction numbers $d_j(X_0)$, $j = = 1, \ldots, N$ are admissible with respect to X_0. Of course, it also follows that

$$d_j(X_0) = \lim_{s \to \infty} d_j(X_{r_s})$$

because all direction numbers are continuous functions on $\mathfrak{M}$, too.

We can easily prove that the sequence $\{X_r\}$ contains at least one homogeneous sequence $\{X_{r_s}\}$. Let now $\{X_{r_s}\}$ be an arbitrary homogeneous sequence. We shall distinguish two special cases according to whether the sequence $\{X_{r_s}\}$ is regulat or not.

I. *Let the sequence $\{X_{r_s}\}$ be regular and denote*

$$X_0 = \lim_{s\to\infty} X_{r_s}\,.$$

Then

$$f(X_0) = \max_{X\in\mathfrak{M}} f(X)\,.$$

Proof. According to the definition of regular sequences the same operation set $\{x_{i_1}, \ldots, x_{i_v}\}$ is admissible both for all points X_{r_s}, $s = 1, 2, \ldots$ and for X_0. Let us now suppose that $\mathbf{d}(X_0) \neq \mathbf{0}$. Then the function $g_0(X) = \nabla f(X_0)(X - X_0)$ does not acquire its maximum on $\mathfrak{M}$ at X_0 and therefore there exists a positive number λ_0 such that $X_0 + \lambda_0\mathbf{d}(X_0) \in \mathfrak{M}$, $f(X_0 + \lambda_0\mathbf{d}(X_0)) > f(X_0)$. We shall now construct a number λ^* $(0 < \lambda^* \leqq \lambda_0)$ such that $X_{r_s} + \lambda^*\mathbf{d}(X_{r_s}) \in \mathfrak{M}$ for every sufficiently large s. As the vector $\mathbf{d}(X_{r_s})$ is admissible with respect to X_{r_s} there is either a real number $t_{r_s} > 0$ such that $X_{r_s} + t_{r_s}\mathbf{d}(X_{r_s}) \in \mathfrak{M}$ but $X_{r_s} + t\mathbf{d}(X_{r_s})$ non $\in \mathfrak{M}$ for $t > t_{r_s}$ or $X_{r_s} + t\mathbf{d}(X_{r_s}) \in \mathfrak{M}$ for all $t > 0$. Then we put $t_{r_s} = +\infty$. Let us put $t_0 = \liminf_{s\to\infty} t_{r_s}$; then obviously $t_0 > 0$ and it suffices to choose e. g. $\lambda^* = \min(\lambda_0, \frac{1}{2}t_0)$. Then $\lambda^* > 0$ and we have

$$f(X_0 + \lambda^*\mathbf{d}(X_0)) = f\left[\frac{\lambda^*}{\lambda_0}(X_0 + \lambda_0\mathbf{d}(X_0)) + \left(1 - \frac{\lambda^*}{\lambda_0}\right)X_0\right] \geqq$$

$$\geqq \frac{\lambda^*}{\lambda_0} f(X_0 + \lambda_0\mathbf{d}(X_0)) + \left(1 - \frac{\lambda^*}{\lambda_0}\right)f(X_0) > f(X_0)\,,$$

$$\lim_{s\to\infty} f(X_{r_s} + \lambda^*\mathbf{d}(X_{r_s})) = f(X_0 + \lambda^*\mathbf{d}(X_0))$$

and hence, for all sufficiently large s,

$$f(X_{r_s} + \lambda^*\mathbf{d}(X_{r_s})) > f(X_0) \geqq f(X_{r_s} + \lambda_{r_s}\mathbf{d}(X_{r_s}))$$

which contradicts the construction of λ_{r_s}. It must therefore follow that $\mathbf{d}(X_0) = \mathbf{0}$ which implies

$$f(X_0) = \max_{X\in\mathfrak{M}} f(X)\,.$$

II. Let the homogeneous sequence $\{X_{s_s}\}$ not be regular. In order to treat this case we shall first prove two lemmas:

1. *Let $\{X_{r_s}\}$ be a homogeneous sequence which is not regular. Then* $\lim_{s\to\infty} \lambda_{r_s} = 0$.

Proof. Let $\lim_{s\to\infty} X_{r_s} = X_0$. The operation set of the points X_{r_s} being inadmissible at X_0 the vector $\mathbf{d}(X_0) = \lim_{s\to\infty} \mathbf{d}(X_{r_s})$ is inadmissible. If the relation $\lim_{s\to\infty} \lambda_{r_s} = 0$ did not hold then there would be a subsequence $\{\lambda_{r_{s_p}}\}$ such that $\lim_{p\to\infty} \lambda_{r_{s_p}} = \bar{\lambda} > 0$. But then

$$\lim_{p\to\infty} X_{r_{s_p}+1} = \lim_{p\to\infty} X_{r_{s_p}} + \lambda_{r_{s_p}} \mathbf{d}(X_{r_{s_p}}) = X_0 + \bar{\lambda}\mathbf{d}(X_0) \in \mathfrak{M}\,,$$

i. e. the vector $\mathbf{d}(X_0)$ would be admissible.

2. *Let $\{X_{r_s}\}$ be a homogeneous sequence which is not regular and let* $\lim_{s\to\infty} X_{r_s} = X_0 = (x_{01}, \ldots, x_{0N})$. *Let α be such an index that the direction number $d_\alpha(X_0)$ is inadmissible and that all $d_j(X_0)$'s, for $j \neq \alpha$, are admissible. Then $x_{r_s+1,\alpha} = x_{0\alpha}$ for all sufficiently large s.*

Proof. Let, for example, $x_{0\alpha} = \bar{x}_\alpha$. Then $d_\alpha(X_0) < 0$ and therefore there is an s_0 such that also $d_\alpha(X_{r_s}) < 0$ for $s > s_0$. The considered operation set being admissible at X_{r_s} we have obviously $x_{r_s,\alpha} > \bar{x}_\alpha$ for $s > s_0$. If the equality $x_{r_s+1,\alpha} = \bar{x}_\alpha$ did not hold there would be infinitely many indices s_p, $p = 1, 2, \ldots$ such that $x_{r_{s_p}+1,\alpha} > \bar{x}_\alpha$. Let us now consider our problem modified so that the variable x_α has to acquire values from the closed interval $\langle \bar{x}_\alpha - \varepsilon, \bar{\bar{x}}_\alpha \rangle$ where ε is an arbitrarily small positive number. Then the quantities $x_{r_{s_p}+1,\alpha}$ would not change but $d_\alpha(X_0)$ would become admissible and similarly as in the previous case we conclude that $\mathbf{d}(X_0) = \mathbf{0}$.

In the following, two special cases are to be distinguished again:

IIa. Let the vector $\mathbf{d}(X_0) = \lim_{s\to\infty} \mathbf{d}(X_{r_s})$ *have exactly one inadmissible component* $d_\alpha(X_0)$. Then we can without any loss of generality suppose that x_α is the only bound-attained coordinate of X_0 not included in the operation set. For, if x_β is another boundary attained coordinate of X_0 not included in the operation set, e. g. $x_{0\beta} = \bar{x}_\beta$, then the direction number $d_\beta(X_0)$ is admissible with respect to X_0 (hence $d_\beta(X_0) \geqq 0$) and it suffices to replace the condition $\bar{x}_\beta \leqq x_\beta \leqq \bar{\bar{x}}_\beta$ in (2) by $\bar{x}_\beta - \varepsilon \leqq x_\beta \leqq \bar{\bar{x}}_\beta$, where ε is an arbitrarily small positive number and x_β stops being a boundary attained variable. Then, as follows from Lemma 2, the situation in the considered operation set $\{x_{i_1}, \ldots, x_{i_\nu}\}$ will be the same both for all points X_{r_s+1} and for X_0 and the new operation set formed according to the corresponding instructions (Chapter I, Sec. 4) will therefore be admissible also for X_0 (by Lemma 1 we conclude that $\lim_{s\to\infty} X_{r_s+1} = X_0$), i. e. the sequence $\{X_{r_s+1}\}$ is regular and hence

$$f(X_0) = \max_{X \in \mathfrak{M}} f(X)\,.$$

IIb. Let the vector $\mathbf{d}(X_0) = \lim_{s\to\infty} \mathbf{d}(X_{r_s})$ *have inadmissible components* $d_{\alpha_1}(X_0), \ldots$ $\ldots, d_{\alpha_\mu}(X_0)$, $\mu \geqq 2$. Analogically as in Case IIa, the set of all bound-attained coordinates of X_0 can be supposed to be the union of the operation set $\{x_{i_1}, \ldots, x_{i_\nu}\}$ with the set $\{x_{\alpha_1}, \ldots, x_{\alpha_\mu}\}$. To fix the ideas we shall, for example, suppose that $x_{0,\alpha_j} = \bar{x}_{\alpha_j}$, $j = 1, \ldots, \mu$ (other possible cases being treated similarly). Then we shall define the auxiliary parameters $\lambda_{\alpha_j}(r_s)$ as follows:

$$\lambda_{\alpha_j}(r_s) = \frac{x_{r_s,\alpha_j} - \bar{x}_{\alpha_j}}{-d_{\alpha_j}(X_{r_s})}\,, \quad s = 1, 2, \ldots,\ j = 1, \ldots, \mu\,.$$

There exist certainly an integer q $(1 \leqq q \leqq \mu)$ and a subsequence $\{X_{r_{s_p}}\}$ of $\{X_{r_s}\}$ such that, for all p,

$$\lambda_{\alpha_q}(r_{s_p}) = \max\,[\lambda_{\alpha_1}(r_{s_p}), \ldots, \lambda_{\alpha_\mu}(r_{s_p})]\,.$$

Let us now put

$$Y_p = X_{r_{s_p}} + \lambda_{\alpha_q}(r_{s_p})\,\mathbf{d}(X_{r_{s_p}})\,.$$

Then clearly Y_p non $\in \mathfrak{M}$ but to every $\varepsilon > 0$ there exists a p_0 such that $\varrho(X_{r_{s_p}+1}, Y_p) < \varepsilon$ for $p > p_0$ ($\varrho(X, Y)$ denotes the distance between the points X, Y in $\mathscr{E}_N$), hence $Y_p \in \mathfrak{R}$ for all sufficiently large p (the set $\mathfrak{R}$ was defined in Sec. 1).

We shall now define the set $\mathfrak{A}$ as follows: $\mathfrak{A}$ is the set of all points $X \in \mathscr{E}_N$ satisfying Eq. (4) and the inequalities

$$\bar{x}_i \leqq x_i \leqq \bar{\bar{x}}_i\,, \quad 1 \leqq i \leqq N\,, \quad i \neq \alpha_1, \alpha_2, \ldots, \alpha_{q-1}, \alpha_{q+1}, \ldots, \alpha_\mu\,;$$

$$\bar{x}_i - \varepsilon \leqq x_i \leqq \bar{\bar{x}}_i\,, \quad i = \alpha_1, \alpha_2, \ldots, \alpha_{q-1}, \alpha_{q+1}, \ldots, \alpha_\mu$$

where ε is a positive number such that $\mathfrak{A} \subset \mathfrak{R}$. Clearly $\mathfrak{M} \subset \mathfrak{A}$ and $Y_p \in \mathfrak{A}$ for all sufficiently large p as $\lim_{p\to\infty} Y_p = X_0$. The points Y_p, $p = 1, 2, \ldots$ and X_0 have the same properties as the points X_{r_s+1}, $s = 1, 2, \ldots$ and X_0 in Case IIa. The sequence $\{Y_p\}$ is therefore regular in $\mathfrak{A}$ and we conclude that

$$f(X_0) = \max_{X\in\mathfrak{A}} f(X) \quad \text{and a fortiori} \quad f(X_0) = \max_{X\in\mathfrak{M}} f(X)\,.$$

Remark. From what we have just proved it follows that for each limit point X_0 of the sequence $\{X_r\}$ the equality

$$f(X_0) = \max_{X\in\mathfrak{M}} f(X)$$

holds. If the preference function $f(X)$ is even strictly concave on $\mathfrak{M}$ then the point for which it acquires its maximum on $\mathfrak{M}$ is determined uniquely, as is well-known. The sequence $\{X_r\}$ is therefore in that case convergent and the relation

$$f(\lim_{r\to\infty} X_r) = \max_{X\in\mathfrak{M}} f(X)$$

is valid.

3. An Estimate of the Error. Throughout this section, the set $\mathfrak{M}$ will be supposed to be bounded. Let us denote

$$\sigma_r = \max_{X\in\mathfrak{M}} g_r(X)\,, \quad r = 1, 2, \ldots, \quad c = \max_{X\in\mathfrak{M}} f(X)\,.$$

Then obviously $\sigma_r \geqq 0$, $\sigma_r = 0$ if and only if $f(X_r) = c$. We shall now prove that the number c fulfils the inequalities

$$f(X_r) \leqq c \leqq f(X_r) + \sigma_r\,, \quad r = 1, 2, \ldots\,.$$

The left-hand side part of this inequality being trivial we shall prove its right-hand side part. Under our assumptions, there exist points X_0, Y_1, Y_2, ... such that $f(X_0) = c$,

$g_r(Y_r) = \sigma_r$, $r = 1, 2, \ldots$. If the considered inequalities were not valid then, for some r, the opposite inequality

$$f(X_r) < c - \sigma_r = f(X_0) - \nabla f(X_r)(Y_r - X_r)$$

would hold. But from the concavity property of $f(X)$ follows

$$f(X_0) \leqq f(X_r) + \nabla f(X_r)(X_0 - X_r)$$

and substituting from this relation into the previous inequality for $f(X_0)$ we get

$$f(X_r) < f(X_r) + \nabla f(X_r)(X_0 - X_r) + \nabla f(X_r)(X_r - Y_r)$$

and hence $\nabla f(X_r)(X_0 - Y_r) > 0$. It follows that

$$\sigma_r = \nabla f(X_r)(Y_r - X_r) = \nabla f(X_r)(Y_r - X_0) + \\ + \nabla f(X_r)(X_0 - X_r) < \nabla f(X_r)(X_0 - X_r) = g_r(X_0)$$

which contradicts the assumptions that $\sigma_r = \max_{X \in \mathfrak{M}} g_r(X)$.

We are now in a position to prove that $\lim_{r \to \infty} \sigma_r = 0$. For the proof, the following auxiliary assertion will be useful:

Let $\{Z_r\}$ be any convergent sequence of points contained in $\mathfrak{M}$, $\lim_{s \to \infty} Z_r = Z_0$. Let $\{X_{r_s}\}$ be a convergent subsequence of $\{X_r\}$, $\lim X_{r_s} = X_0$. Then $\lim_{s \to \infty} g_{r_s}(Z_{r_s}) = g_0(Z_0)$.

The proof follows immediately from the continuity property of all derivatives of the 1st order.

Let us now suppose that the relation $\lim_{r \to \infty} \sigma_r = 0$ does not hold. Then there exists a subsequence $\{\sigma_{r_p}\}$ such that

$$\lim_{p \to \infty} \sigma_{r_p} = \sigma_0 > 0 .$$

The sequences $\{X_{r_p}\}$, $\{Y_{r_p}\}$ being bounded there exist convergent subsequences $\{X_{r_{p_q}}\}$, $\{Y_{r_{p_q}}\}$,

$$\lim_{q \to \infty} X_{r_{p_q}} = X_0 , \quad \lim_{q \to \infty} Y_{r_{p_q}} = Y_0 .$$

Then it follows from the above stated auxiliary assertion that

$$\lim_{q \to \infty} \sigma_{r_{p_q}} = \lim_{q \to \infty} g_{r_{p_q}}(Y_{r_{p_q}}) = g_0(Y_0) = \sigma_0 > 0$$

which contradicts the fact that $g_0(X) \leqq 0$ for all $X \in \mathfrak{M}$.

Having found the point X_r and the value $f(X_r)$ we can therefore find an upper estimate σ_r of the difference $c - f(X_r)$. The problem of finding σ_r is then a usual linear programming problem consisting in maximizing the linear function $g_r(X)$ on the set $\mathfrak{M}$. The determination of σ_r being rather laborious, it is convenient to compute it only when fairly small values of the differences $f(X_{r+1}) - f(X_r)$ occur. The estimate in question is also rather rough at the beginning of the iterative process.

Remark. Under our assumptions, the set

$$\mathfrak{N}(X^*) = \{X \in \mathfrak{M} : f(X) \geqq f(X^*)\}$$

is bounded for every $X^* \in \mathfrak{M}$. If we are able to compute the quantity

$$\tau_r = \max_{X \in \mathfrak{N}(X_r)} g_r(X)$$

we can obviously use also τ_r as an upper estimate of the difference $c - f(X_r)$, even when the set $\mathfrak{M}$ is not bounded. Of course, the computation of τ_r is not a linear programming problem any more.

The analogical convex programming problem has also been studied e. g. by G. B. Dantzig [5] and J. E. Kelley [6]. For a special case, the quadratic programming, various finite methods have already been found (see e. g. [7], [8]).

References

[1] *Ragnar Frisch:* The Multiplex Method for Linear Programming. Memorandum fra Socialøkonomisk Institutt, Universitetet i Oslo, 12. September 1958.

[2] *Ragnar Frisch:* The Multiplex Method for Linear Programming. Sankhya Vol. *18*, Parts 3 & 4, 329 — 362.

[3] *Ф. Р. Гантмахер:* Теория матриц. Москва 1954.

[4] *C. J. van Eijk, J. Sandee:* Quantitative Determination of an Optimum Economic Policy. Econometrica Vol. *27* (1959), 1 — 13.

[5] *G. B. Dantzig:* On the Significance of Solving Linear Programming Problems with Some Integer Variables. Econometrica Vol. *28* (1960), 30 — 44 (see especially pp. 38 — 39).

[6] *J. E. Kelley:* A Computational Approach to Convex Programming (an abstract). Econometrica Vol. *27* (1959), p. 276.

[7] *P. Wolfe:* The Simplex Method for Quadratic Programming. Econometrica Vol. *27* (1959), 382 — 398.

[8] *H. S. Houthakker:* The Capacity Method for Quadratic Programming. Econometrica Vol. *28* (1960), 62 — 87.

Резюме

МЕТОД МУЛЬТИПЛЕКСА И ЕГО ИСПОЛЬЗОВАНИЕ В НЕЛИНЕЙНОМ ПРОГРАММИРОВАНИИ

ЯРОМИР АБРГАМ (Jaromír Abrham), Прага

Работа разделяется на две части. В первой части наново формулируется метод мультиплекса Фриша (см. [1]) для определения максимума линейной

целевой функции $b_{00} + \sum_{k=1}^{n} b_{jk}x_k$ на множестве $\mathfrak{M}$, описанном системой линейных уравнений

$$x_j = b_{j_0} + \sum_{k=1}^{n} b_{jk}x_k\,, \quad j = 1, \dots, N > n\,,$$

где b_{jk}, $j = 0, 1, \dots, N$, $k = 0, 1, \dots, N$ — данные вещественные числа, для которых $b_{j_0} = 0$, $b_{jk} = \delta_{jk}$, $j = 1, \dots, n$, и системой линейных неравенств

$$\bar{x}_j \leqq x_j \leqq \bar{\bar{x}}_j\,, \quad j = 1, \dots, N\,,$$

где $\bar{x}_j$, соотв. $\bar{\bar{x}}_j$ — или вещественные числа или символы $-\infty$, $+\infty$. Показано, что вектор $(d_j, \dots, d_N)$, где d_j — направляющие числа, является в известном смысле вектором наибольшего возрастания рассматриваемой целевой функции. В п. 4 доказывается финитность метода мультиплекса для линейного программирования.

Во второй части формулируется способ итераций для определения максимума выпуклой функции $f(X) = f(x_1, \dots, x_n)$ на множестве $\mathfrak{M}$. Относительно функции $f(X)$ предполагается, что она определена на некотором открытом множестве $\mathfrak{R} \supset \mathfrak{M}$ и обладает на этом множестве непрерывными частными производными первого порядка по всем переменным. Построена последовательность $\{X_r\}_{r=1}^{\infty}$ точек множества $\mathfrak{M}$, о которой доказано, что

$$\lim_{r \to \infty} f(X_r) = \max_{X \in \mathfrak{M}} f(X)\,.$$

В заключение, при условии, что множество $\mathfrak{M}$ ограничено, построена последовательность $\{\sigma_r\}_{r=1}^{\infty}$ неотрицательных чисел таких, что для всех r имеет место

$$c - f(X_r) \leqq \sigma_r\,, \quad \text{где} \quad c = \max_{X \in \mathfrak{M}} f(X)\,.$$

Каждое из чисел σ_r вычисляется как максимум некоторой специальной линейной функции на множестве $\mathfrak{M}$, т. е. вычисление каждого из этих чисел является задачей линейного программирования. Далее доказывается, что $\lim_{r \to \infty} \sigma_r = 0$.