Numerische Mathematik 1, 253—268 (1959)

Newton's Method for Convex Programming and Tchebycheff Approximation

By

E. W. CHENEY and A. A. GOLDSTEIN

§ 1. Introduction. The rationale of Newton's method is exploited here in order to develop effective algorithms for solving the following general problem: given a convex continuous function F defined on a closed convex subset K of E_n, obtain (if such exists) a point x of K such that $F(x) \leqq F(y)$ for all y in K. The manifestation of Newton's method occurs when, in the course of computation, convex hypersurfaces are replaced by their support planes.

The problems of infinite systems of linear inequalities and of infinite linear programming are subsumed by the above problem, as are certain Tchebycheff approximation problems for continuous functions on a metric compactum. In regard to the latter, special attention is devoted in §§ 27—30 to the feasibility of replacing a continuum by a finite subset in such a way that a discrete approximation becomes an accurate substitute for the continuous approximation.

It is to be pointed out that the basic idea of the algorithms below is not new, having been first used by REMEZ [*1, 2, 3, 4*]. Other authors who have put it to use in one form or another are NOVORDVORSKII-PINSKER [*5*], BEALE [*6, 7*], BRATTON [*8*], STIEFEL [*9, 10*], WOLFE [*11*], STONE [*12*], and KELLEY [*13*].

The general problem described above is put aside until § 21 in order that the main ideas may be developed in a simpler setting. Consider, then, a matrix having a finite number, $n+1$, of columns and at least $n+1$ rows (the number of rows may be non-denumerable),

$$\begin{pmatrix} A_1^1 \dots A_n^1 & b_1 \\ A_1^2 \dots A_n^2 & b_2 \\ \vdots \quad\; \vdots & \vdots \end{pmatrix}.$$

A point $x=(x_1, \dots, x_n) \in E_n$ is sought which will minimize the function

$$F(x) = \sup_i \sum_{j=1}^{n} A_j^i x_j - b_i .$$

The linear Tchebycheff approximation problem is already included by this problem. Note that any continuous convex function F defined throughout E_n may be represented as above via its support planes. These planes assume the form

$$R^i(x) = \sum_{j=1}^{n} A_j^i x_j - b_i .$$

§ 2. Nomenclature. Consider the three matrices

$$A=\begin{pmatrix} A_1^1\dots A_n^1 \\ A_1^2\dots A_n^2 \\ \vdots \quad \vdots \end{pmatrix} \qquad B=\begin{pmatrix} A_1^1\dots A_n^1 & 1 \\ A_1^2\dots A_n^2 & 1 \\ \vdots \quad \vdots & \vdots \end{pmatrix} \qquad C=\begin{pmatrix} A_1^1\dots A_n^1 & 1 & b_1 \\ A_1^2\dots A_n^2 & 1 & b_2 \\ \vdots \quad \vdots & \vdots & \vdots \end{pmatrix}.$$

The *Haar condition* on A is the requirement that every $n\times n$ submatrix of A be non-singular. The *solvency condition* on A is the requirement that every $(n+1)\times(n+1)$ submatrix of B be non-singular. The *normality condition* on C is the requirement that every $(n+2)\times(n+2)$ submatrix of C be non-singular. The functions R^i defined above are known as *residual functions*, and the hyperplanes in E_{n+1} whose equations are

$$z = R^i(x)$$

are *residual planes*. An *edge* is the intersection of any n residual planes. When the Haar condition is fulfilled, each edge is a 1-dimensional manifold in E_{n+1} along which z is not constant. When the solvency condition is met, each set of $n+1$ residual planes has in common a single point called a *vertex*. Assuming the normality condition, no more than $n+1$ residual planes can pass thru a vertex.

The solvency condition is equivalent to the following: given a point $x\in E_n$ and a set of rows $\{A^{i_0},\dots,A^{i_n}\}$ from A, there corresponds a unique vector $u=(u_0,\dots,u_n)$ satisfying $\Sigma u_j=1$ and $x=\Sigma u_j A^{i_j}$. Stated otherwise, each set of $n+1$ rows from A has an n-simplex for its convex hull.

The notation $o(I)$ denotes the number of elements in the set I. The notation $H\{A^i: i\in I\}$ denotes the *convex hull* of the set of points $\{A^i: i\in I\}$; i.e., the set of all linear combinations $\Sigma u_i A^i$ where $u_i\geqq 0$, $o\{i: u_i>0\}<\infty$, and $\Sigma u_i=1$. The notation $[u, v]$ will be used for $\Sigma u_i v_i$. The notation $C\{A^i: i\in I\}$ denotes the *conical hull* of the set of points $\{A^i: i\in I\}$; i.e., the set of all linear combinations $\Sigma u_i A^i$ where $u_i\geqq 0$ and $o\{i: u_i>0\}<\infty$. A *half-space* in E_n is a set of the form $\{x: [a, x]\geqq k\}$. A *polytope* in E_n is the intersection of a finite number of half-spaces.

§ 3. Lemmas. *A*. Let K be a closed convex set in Hilbert space and u a point not in K. There exists an unique $v\in K$ closest to u. Furthermore, for $x\in K$,

$$[x, u-v]\leqq[v, u-v]<[u, u-v].$$

This theorem is well-known. See for example [*14*].

B. Let Ω denote a subset of E_n and x a point of $H(\Omega)$. There exists $\Omega'\subset\Omega$ such that $o(\Omega')\leqq n+1$ and $x\in H(\Omega')$. If Ω is connected then there exists $\Omega'\subset\Omega$ such that $o(\Omega')\leqq n$ and $x\in H(\Omega')$. See [*15*, p. 9] and the references given there.

C. The distance between two polytopes in E_n is attained. Thus, if bounded below, the function $F(x)=\max\limits_{1\leqq i\leqq m}\sum\limits_{j=1}^{n} A_j^i x_j - b_i$ attains its infimum in E_n. See [*14*].

D. Let a matrix B result from an $n\times n$ non-singular matrix A by replacement of its r^{th} row by a vector b. Let the columns of A^{-1} be designated by $C_1,\dots,C_n$. If $\lambda\equiv[b, C_r]\neq 0$, then B is non-singular and the columns $D_1,\dots,D_n$ of its inverse are given by $D_r=\lambda^{-1}C_r$ and $D_j=C_j-\lambda^{-1}[b, C_j]C_r$ $(j\neq r)$.

E. The convex hull of a compactum in E_n is itself compact.

Proof. Let Ω be a compactum in E_n, K its convex hull. By § 3 B to each $x \in K$ there corresponds a representation $x = \sum_{i=0}^{n} t_i(x) A^i(x)$ with $t_i(x) \geqq 0$, $\Sigma_i t_i(x) = 1$, and $A^i(x) \in \Omega$. If $\{x_j\}$ is a sequence in K then there exists by the compactness of Ω and of $Q = \{(t_0, \ldots, t_n) : t_i \geqq 0, \Sigma t_i = 1\}$ a sequence j_k such that $\lim_{k\to\infty} A^i(x_{j_k}) \equiv A$ exists in Ω and $\lim_k \left(t_0(x_{j_k}), \ldots, t_n(x_{j_k})\right) \equiv (t_0, \ldots, t_n)$ exists in Q. Clearly then $\Sigma t_i A^i \in K$, proving the compactness of K.

In a general Banach space, the convex hull of a compactum is totally bounded, by a theorem of Mazur, but not necessarily closed.

§ 4. Lemma. Let E denote an arbitrary linear space, Ω a set of linear functionals on E, and K the convex hull of Ω. The system of linear inequalities

$$\text{(S)} \qquad f(x) < 0 \qquad (f \in \Omega)$$

possesses a finite inconsistent subsystem if and only if $0 \in K$.

Proof. (i) Assume $0 \in K$. Then an equation of the form $0 = \sum_{i=1}^{m} c_i f_i$ holds where $f_i \in \Omega$, $c_i > 0$ and $\Sigma c_i = 1$. Thus for any $x \in E$, $0 = \Sigma c_i f_i(x)$, showing that the system

$$\text{(S')} \qquad f_i(x) < 0 \qquad (1 \leqq i \leqq m)$$

is inconsistent.

(ii) Assume that system (S) has system (S′) as an inconsistent subsystem. Denote by N the set of solutions of

$$f_i(z) = 0 \qquad (1 \leqq i \leqq m)$$

and select $x_1, \ldots, x_n \in E$ $(n \leqq m)$ so that $N \oplus x_1 \oplus \cdots \oplus x_n = E$. Since every $x \in E$ has a representation $x = x_0 + \sum_{j=1}^{n} c_j x_j$ with $x_0 \in N$, $f_i(x) = \Sigma c_j f_i(x_j)$. The system (S′) may therefore be written

$$\text{(S'')} \qquad \sum_{j=1}^{n} A_j^i c_j < 0 \qquad (1 \leqq i \leqq m)$$

where $A_j^i = f_i(x_j)$, and this system too is inconsistent in E_n. We shall show that $0 \in H\{A^1, \ldots, A^m\} \equiv K'$ where $A^i = (A_1^i, \ldots, A_n^i)$. Indeed if this is not the case, then by § 3 A, there exists a halfspace $\{x : [x, c] < 0\}$ in E_n containing K'. This would make c a solution of (S″). Thus $0 \in K'$, and an equation of the form $0 = \sum_{i=1}^{m} e_i A^i$ must obtain with $e_i \geqq 0$ and $\Sigma e_i = 1$. From this we obtain easily $\Sigma e_i f_i = 0$, which completes the proof. For infinite systems this lemma generalizes corollary 5 of [*16*].

§ 5. Lemma. Let Ω denote a compact subset of E_n and b a continuous real-valued function on Ω. For $x \in E_n$ define $f(x) = \sup_{A \in \Omega} [A, x]$ and $F(x) = \sup_{A \in \Omega} [A, x] - b(A)$. Consider also two systems of linear inequalities

$$\text{(s)} \qquad [A, z] < 0 \qquad\qquad \text{(S)} \qquad [A, z] < b(A) + M \qquad (A \in \Omega).$$

The following statements are equivalent

(i) F is bounded below.

(ii) f is bounded below.

(iii) System (s) is inconsistent.

(iv) $0 \in H(\Omega)$.

(v) System (s) has an inconsistent subsystem comprising at most $n+1$ inequalities.

(vi) For some M, system (S) has an inconsistent subsystem comprising at most $n+1$ inequalities.

(vii) For some M, system (S) is inconsistent.

Proof. (i)→(ii).

$$F(x) = \sup_A \{[A, x] - b(A)\} \leqq \sup_A [A, x] - \inf_A b(A) = f(x) - \inf_A b(A).$$

(ii)→(iii). If (s) is consistent and is satisfied by z^0, then $q \equiv \sup_A [A, z^0] < 0$, for the supremum is necessarily attained, Ω being closed and bounded. But now $f(t z^0) = t q \to -\infty$ as $t \to +\infty$.

(iii)→(iv). $H(\Omega)$ is compact by § 3 E. If $0 \notin H(\Omega)$ then by § 3 A, there is a halfspace $\{x : [x, z] < 0\}$ containing $H(\Omega)$. Thus z satisfies system (s).

(iv)→(v). If $0 \in H(\Omega)$ then by § 3 B, there exist $A^0, \ldots, A^n$ such that $0 \in H\{A^0, \ldots, A^n\}$. By § 4, then, the system

$$\text{(s')} \qquad [A^i, z] < 0 \qquad (0 \leqq i \leqq n)$$

is inconsistent.

(v)→(vi). Assume system (s') inconsistent. Put $M = -\max_{0 \leqq i \leqq n} b(A^i)$ then for all z, $\max_{0 \leqq i \leqq n} \{[A^i, z] - b(A^i)\} \geqq \max_i [A^i, z] - \max_i b(A^i) \geqq M$. Thus the system

$$\text{(S')} \qquad [A^i, z] < b(A^i) + M \qquad (0 \leqq i \leqq n)$$

is inconsistent.

(vi)→(vii). Trivial.

(vii)→(i). If (S) is inconsistent, then clearly $F(x) \geqq M$ for all x.

§ 6. Lemma. Let $S_1 \supset S_2 \supset \ldots$ be a nested sequence of compact sets in E_n. Then $H(\cap S_i) = \cap H(S_i)$.

Proof. Clearly $\cap S_i \subset \cap H(S_i)$. Hence by the convexity of $\cap H(S_i)$, $H(\cap S_i) \subset \cap H(S_i)$. For the converse, let $x \equiv (x_1, \ldots, x_n)$ be a point of $\cap H(S_i)$. Then for each i there is, by § 3 B a representation

$$x = \sum_{j=0}^{n} t_j^{(i)} y_j^{(i)}$$

where the $(n+1)$-tuple $t^{(i)} \equiv (t_0^{(i)}, \ldots, t_n^{(i)})$ lies in the set

$$Q = \{(t_0, \ldots, t_n) : t_j \geqq 0, \ \Sigma_j t_j = 1\}$$

and where $y_j^{(i)} \in S_i$. By the compactness of Q and of S_1 there exists a sequence of integers $i_1, i_2, \ldots$ such that $\lim_{k\to\infty} t^{(i_k)} \equiv t \equiv (t_0, \ldots, t_n)$ exists in Q and $\lim_{k\to\infty} y^{(i_k)} \equiv y_j$ exists in S_1. For each k, all but a finite number of $y_j^{(i_1)}, y_j^{(i_2)}, \ldots$ lie in S_k because the S_k's are nested. Since each S_k is closed, $y_j \in S_k$; thus $y_j \in \cap S_i$. Hence $x = \sum_{j=0}^{n} t_j y_j$, showing $x \in H(\cap S_i)$. Q.E.D.

§ 7. Theorem. Let Ω be a compact subset of E_n and b a continuous real valued function on Ω. Define $R(A, x) = [A, x] - b(A)$ and $F(x) = \sup_{A \in \Omega} R(A, x)$. If there exists an $x^0 \in E_n$ such that $F(x^0) \leqq F(x)$ for all x, then there exists a set $\{A^0, \ldots, A^k\} \subset \Omega$ with $k \leqq n$ such that $\inf_x \max_{0 \leqq i \leqq k} R(A^i, x) = F(x^0) \equiv M = R(A^i, x^0)$ $(0 \leqq i \leqq k)$.

Proof. Define for each $i = 1, 2, \ldots$ the set $\Omega_i = \left\{A \in \Omega : R(A, x^0) \geqq F(x^0) - \frac{1}{i}\right\}$. Clearly the sets Ω_i are compact and nested. We shall prove that for each i the following system is inconsistent:

(1) $$[A, z] < 0 \qquad (A \in \Omega_i).$$

Indeed, if z^0 satisfies (1), then define $q = \sup_{A \in \Omega_i} [A, z^0]$. Since Ω_i is compact and $[A, z^0]$ is a continuous function of A, there is an $A^0 \in \Omega_i$ for which $[A^0, z^0] = q$. Thus $q < 0$. Let $c = \|z^0\| \sup_{A \in \Omega} \|A\|$. Since $R\left(A, x^0 + \frac{1}{2ic} z^0\right) = R(A, x^0) + \frac{1}{2ic} [A, z^0]$, it is clear that for $A \in \Omega_i$ we have $R\left(A, x^0 + \frac{1}{2ic} z^0\right) \leqq F(x^0) + \frac{q}{2ic}$, while for $A \notin \Omega_i$,

$$R\left(A, x^0 + \frac{1}{2ic} z^0\right) < F(x^0) - \frac{1}{i} + \frac{1}{2ic} \|A\| \cdot \|z^0\| \leqq F(x^0) - \frac{1}{2i}.$$

Thus $F\left(x^0 + \frac{1}{2ic} z^0\right) < M$, a contradiction. This shows that for all $i = 1, 2, \ldots$, system (1) is inconsistent. By § 5, $0 \in H(\Omega_i)$. By § 6, $0 \in H\left(\bigcap_{i=1}^{\infty} \Omega_i\right)$. Furthermore $\cap \Omega_i = \{A \in \Omega : R(A, x^0) = F(x^0)\}$. Now by § 5 there exist $A^0, \ldots, A^k$ in $\cap \Omega_i$ such that the system

$$[A^i, z] < 0 \qquad (0 \leqq i \leqq k)$$

is inconsistent, whence the theorem. A related result may be found in [*17*].

§ 8. Remark. Let Ω denote a closed, bounded, connected subset of E_n and b a continuous function on Ω. If $\inf_x \sup_{A \in \Omega} \sum_{j=1}^{n} A_j x_j - b(A) > -\infty$, then Ω contains vanishing $n \times n$ determinants; in other words, the Haar condition is violated.

Proof. By § 5, $0 \in H(\Omega)$. Since Ω is connected, by § 3 B a there exist points $A^1, \ldots, A^n$ in Ω and non-negative coefficients $c_1, \ldots, c_n$ fulfilling $\Sigma c_i = 1$ and $\Sigma c_i A^i = 0$. This latter equation exhibits a linear dependence among $A^1, \ldots, A^n$.

§ 9. Lemma. If A is an $(n+1) \times n$ matrix of rank n and if the function $f(x) = \max_i [A^i, x]$ is bounded below, then the solvency condition follows.

Proof. If the solvency condition fails, there exists a vector $(u_0, \ldots, u_n) \neq 0$ such that

$$\begin{pmatrix} 1 & A_1^0 \ldots A_n^0 \\ \vdots & \vdots \qquad \vdots \\ 1 & A_1^n \ldots A_n^n \end{pmatrix} \begin{pmatrix} u_0 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

thus $[A^i, u] = -u_0$, where $0 \leqq i \leqq n$ and where $u = (u_1, \ldots, u_n)$. If $u_0 = 0$, the rank hypothesis is violated. If $u_0 \neq 0$, then either $+u$ or $-u$ is a solution of the inequalities $[A^i, z] < 0$. Thus $f(z) < 0$, whence $\lim_{t \to +\infty} f(tz) = -\infty$.

§ 10. Lemma. Assume that the set $\{A^i: 0 \leqq i \leqq n+1\}$ satisfies the Haar condition and that $0 \in H\{A^i: 0 \leqq i \leqq n\}$. Then there exist unique indices, i_0 and i_1 among $\{0, \ldots, n\}$ such that $0 \in H\{A^i: 0 \leqq i \leqq n+1,\ i \neq i_0\}$ and $-A^{n+1} \in C\{A^i: 0 \leqq i \leqq n,\ i \neq i_1\}$. Furthermore, $i_0 = i_1$. See Satz 5, p. 4, of [*10*].

Proof. Since $0 \in H\{A^i: 0 \leqq i \leqq n\}$, and since $\{A^i: 0 \leqq i \leqq n\}$ satisfies the solvency condition, there exist unique coefficients $u_0, \ldots, u_n$ such that $u_i \geqq 0$, $\Sigma u_i = 1$, and $\Sigma u_i A^i = 0$. Because of the Haar condition $u_i > 0$. By § 9, there exist unique coefficients $v_0, \ldots, v_n$ such that $\Sigma v_i = 1$ and $\Sigma v_i A^i = A^{n+1}$. For any $j \leqq n$ we have $0 = A^{n+1} - \Sigma v_i A^i = A^{n+1} - v_j A^j - \Sigma' v_i A^i = A^{n+1} - v_j \Sigma' - \frac{u_i}{u_j} A^i - \Sigma' v_i A^i = A^{n+1} + \Sigma' \left(\frac{v_j u_i}{u_j} - v_i \right) A^i$ throughout which, Σ' abbreviates $\Sigma^n_{i=0, i \neq j}$. If $\frac{v_j u_i}{u_j} \geqq v_i$ for all i, then this equation will furnish a barycentric representation of 0 in terms of $\{A^i: 0 \leqq i \leqq n+1,\ i \neq j\}$. This will indeed be the case if j is chosen to fulfill $\frac{v_j}{u_j} \geqq \frac{v_i}{u_i}$ for all i. If this value of j be denoted by i_0, we have in fact $\frac{v_{i_0}}{u_{i_0}} > \frac{v_i}{u_i}$ for all $i \neq i_0$, due again to the Haar condition. The uniqueness of i_0 may be seen at once from the fact that any other choice will lead to a negative coefficient in the above representation of zero. For each j this representation is unique up to scalar multiplication due to the Haar condition. Similar arguments apply to i_1.

§ 11. Remark. Assume that the set $\{A^1, \ldots, A^{n+2}\} \subset E_n$ satisfies the Haar condition. The system of inequalities

$$(1) \qquad [A^i, z] < 0 \qquad (1 \leqq i \leqq n+2)$$

is inconsistent if and only if it possesses precisely two inconsistent proper subsystems.

Proof. The "if" part being trivial, we proceed at once to the "only if" part. If the system (1) is inconsistent, then by § 4, $0 \in H\{A^i: 1 \leqq i \leqq n+2\}$, and thus, by § 3B, there exists an index i_0 such that $0 \in H\{A^i: 1 \leqq i \leqq n+2,\ i \neq i_0\}$. By §§ 5 and 9, the solvency condition holds for the set $\{A^i: 1 \leqq i \leqq n+2,\ i \neq i_0\}$. By § 10, there exists a unique index $i_1 \neq i_0$ such that

$$0 \in H\{A^i: 1 \leqq i \leqq n+2,\ i \neq i_1\}.$$

By § 4, the inconsistent subsystems are obtained.

§ 12. Remark. Consider two related matrices

$$A = \begin{pmatrix} A_1^1 \ldots A_n^1 \\ A_1^2 \ldots A_n^2 \\ \vdots \quad \vdots \end{pmatrix} \qquad A^* = \begin{pmatrix} A_0^1 & A_1^1 \ldots A_n^1 \\ A_0^2 & A_1^2 \ldots A_n^2 \\ \vdots & \vdots \quad \vdots \end{pmatrix}$$

in which the number of rows is finite or denumerable. If every $(n+1) \times n$ submatrix of A is of full rank then $(A_0^1, A_0^2, \ldots)^T$ may be chosen so that in A^* every $(n+1) \times (n+1)$ submatrix is of full rank.

Proof. Set $A_0^i = t^i$, and expand a typical $(n+1) \times (n+1)$ determinant of A^* by the elements of its first column, obtaining thereby a polynomial in t whose coefficients are $n \times n$ determinants from A which are not all zero by hypothesis.

The set S of all the roots of all the polynomials obtained in this way is an at most denumerable set. One may therefore select any $t \notin S$ to obtain the desired conclusion.

§ 13. Theorem. Assume that the function $f(x) = \max_{1 \leq i \leq m} [A^i, x]$ is bounded below, that $m \geqq n+2$, and that the Haar conditions prevails. Then there exist at least $m-n$ sets $I \subset \{1, \ldots, m\}$ such that $o(I) = n+1$ and $0 \in H\{A^i : i \in I\}$. This bound is best possible.

Proof. By § 5 and the boundedness of f there exists a set $I_0 \subset \{1, \ldots, m\}$ such that $o(I_0) = n+1$ and $0 \in H\{A^i : i \in I_0\}$. By § 5, the function $\max_{i \in I_0} [A^i, x]$ is bounded below. By the Haar condition, the set $\{A^i : i \in I_0\}$ has rank n. By § 9, then, the solvency condition is satisfied by this set. By § 10, to each index $j \notin I_0$ there corresponds uniquely an index $i_j \in I_0$ in such a way that $0 \in H\{A^i : i \in I + j - i_j\}$.

Since j may be selected in $m-n-1$ ways, the number $m-n$ is established. That this bound is best possible is shown in the next paragraph. Observe that it has been shown that to each i there corresponds an $I \subset \{1, \ldots, m\}$ such that $i \in I$, $o(I) = n+1$ and $0 \in H\{A^i : i \in I\}$.

§ 14. Example. Let positive numbers $t_1, \ldots, t_{m-n}$ be selected. Define

$$A = \begin{pmatrix} -1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & -1 \\ t_1 & t_1^2 & t_1^3 & \ldots & t_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ t_{m-n} & t_{m-n}^2 & t_{m-n}^3 & \ldots & t_{m-n}^n \end{pmatrix}.$$

It turns out that (i) A satisfies the Haar condition; (ii) $\inf_x \max_i [A^i, x] > -\infty$; and that (iii) there are precisely $m-n$ sets $I \subset \{1, \ldots, m\}$ such that $o(I) = n+1$ and $0 \in H\{A^i : i \in I\}$.

Proof. (i) Suppose that A contains a singular $n \times n$ submatrix B. Then a dependence $\Sigma c_j B_j = 0$ exists among the columns of B. If exactly k of the first n rows of A are present in B, then this equation implies the vanishing of k of the coefficients c_j as well as the vanishing of the polynomial $\Sigma c_j t^j$ at exactly $n-k$ positive points. Since such a polynomial has at most $n-k-1$ changes of sign, it can have by DESCARTES' rule at most $n-k-1$ positive roots. Thus $c \equiv 0$.

(ii) This follows at once, using § 5, from the observation that

$$0 \in H\{A^1, \ldots, A^{n+1}\}.$$

(iii) If $0 \in H\{A^i : i \in I\}$ and $o(I) = n+1$, then one may write $0 = \sum_{i \in I} k_i A^i$ where $k_i \geqq 0$ and $\Sigma k_i = 1$. It may be seen at once from these conditions that all the first n rows of A must be among $\{A^i : i \in I\}$. The number of ways of obtaining a set of $n+1$ rows from A including the first n is $m-n$.

§ 15. Theorem. Define $F(x) = \max_{1 \leq i \leq m} R^i(x)$ and $f(x) = \min_{1 \leq i \leq m} R^i(x)$. If either of the numbers $F_0 = \inf_x F(x)$, $f_0 = \sup_x f(x)$ is finite, then the other is also, and these values are achieved at appropriate points; furthermore $f_0 \leqq F_0$; i.e.

$$\max_x \min_i R^i(x) \leqq \min_x \max_i R^i(x).$$

Equality occurs here if and only if there exists a point y and a number M satisfying $R^i(y)=M$ $(1 \leqq i \leqq m)$.

Proof. By § 3C, if $F_0 > -\infty$, an $x^0 \in E_n$ exists for which $F_0 = F(x^0)$. Define then $I = \{i: R^i(x^0) = F_0\}$ and observe that the system of inequalities $[A^i, z] < 0$, $(i \in I)$ is inconsistent. This implies that x^0 maximizes the concave function $\min_{i \in I} R^i(x)$. Thus $F_0 = \sup_x \min_{i \in I} R^i(x) \geqq \sup_x \min_i R^i(x) = f_0$. Here we obtain strict inequality unless $R^i(x^0) = F_0$ for all i. The arguments are the same if one begins with the assumption $f_0 < \infty$.

§ 16. Algorithm I. We are given a bounded subset Ω of E_n and a bounded real-valued function on Ω which we write A_0 for $A \in \Omega$. For each $A \in \Omega$ define $R(A, x) = [A, x] - A_0$. Also define $F(x) = \sup_{A \in \Omega} R(A, x)$. It is desired to obtain an $x \in E_n$ for which $F(x) \leqq F(y)$ for all y if such an x exists. Assume that in getting started a subset $\{A^0, \ldots, A^l\} \subset \Omega$ is known which spans E_n and satisfies $0 \in H\{A^0, \ldots, A^l\}$. In this connection, see § 18. At step k $(k \geqq l)$ in the algorithm there is given $\{A^0, \ldots, A^k\} \subset \Omega$. Select x^k to minimize the function $F^k(x) \equiv \max\{R(A^i, x): 0 \leqq i \leqq k\}$. (This may be accomplished by the algorithm of § 17, by the methods of [*18*], by linear programming, etc., etc.) Select $A = A^{k+1} \in \Omega$ to maximize $R(A, x^k)$, or to come within $1/k$ of this maximum. Repeat this cycle, obtaining thereby a sequence $x^l, x^{l+1}, \ldots$. The validity of this algorithm is established in § 22.

§ 17. Algorithm II. We are given a subset $\{A^1, \ldots, A^m\}$ of E_n satisfying the Haar condition and an m-tuple $(b_1, \ldots, b_m)$. It is desired to obtain a minimum for the function $F(x) = \max_{1 \leqq i \leqq m} R^i(x)$ where $R^i(x) = \sum_{j=1}^{n} A^i_j x_j - b_i$. It is necessary to assume that $\inf_x F(x) > -\infty$, or equivalently, $0 \in H\{A^1, \ldots, A^m\}$. At each stage there is given a set $I \subset \{1, \ldots, m\}$ of $n+1$ elements and a point $y \in E_n$. Select $j \in \{1, \ldots, m\}$ to maximize $R^j(y)$. Select y' to minimize $\max_{i \in I+j} R^i(y')$. Select $h \in I$ to minimize $R^h(y')$. Define $I' = I + j - h$, and start anew with I' and y'. A starting procedure is given in § 18 and a formulary in § 19. This algorithm is now in use on the IBM 704 computer, having been programmed by Mr. NORMAN LEVINE.

The following remarks will assist in interpreting § 19. In each cycle there is a set of indices $I = \{i_0, i_1, \ldots, i_n\}$, a point $x = (x_1, \ldots, x_n)$ and a number x_0 such that $R^i(x) = -x_0$ for $i \in I$. This equation may be written $x^* = D b^*$, where $x^* = (x_0, x_1, \ldots, x_n)$, $b^* = (b_{i_0}, \ldots, b_{i_n})$ and where D is the inverse of

$$\begin{pmatrix} 1 & A^{i_0}_1 \ldots A^{i_0}_n \\ \vdots & \vdots \qquad \vdots \\ 1 & A^{i_n}_1 \ldots A^{i_n}_n \end{pmatrix}.$$

In the formulary, $l=0$ signifies the problem of minimizing $F(x) = \max_i R^i(x)$; $l=1$ signifies the problem of obtaining a solution of $F(x) < 0$; and $l=2$ signifies the problem of minimizing $F^*(x) = \max_i |R^i(x)|$.

§ 18. Starting Procedure for Algorithms. Assume the Haar condition and that the function $F(x) = \max_{1 \leq i \leq m} R^i(x)$ possesses a greatest lower bound M. By § 3C there exists an x^0 such that $F(x^0) = M$. Algorithm II requires, for starting, a set of rows $\{A^{i_0}, \ldots, A^{i_n}\}$ from the matrix having the property that the system of inequalities

$$[A^{i_j}, z] < 0 \qquad (0 \leq j \leq n)$$

be inconsistent. The existence of such sets (indeed, $m - n$ of them) is guaranteed by § 13. To obviate the search for such a set, a new row A^0 is adjoined to the matrix, and a number b_0 is selected in such a way that

(1) $$R^0(x^0) \equiv [A^0, x^0] - b_0 < M;$$

that is, x^0 remains a solution of the augmented problem. Toward this end, define $A^0 = -\sum_{i=1}^{n} A^i$, and suppose b_0 is large enough to validate equation (1) above. It is to be verified that the set $\{A^0, \ldots, A^n\}$ satisfies the Haar condition and the condition that the system

(2) $$[A^i, z] < 0 \qquad (0 \leq i \leq n)$$

be inconsistent. Indeed, supposing a non-trivial linear equation $0 = \sum_{\substack{i=0 \\ i \neq i_0}}^{n} u_i A^i$, the Haar condition on $\{A^1, \ldots, A^n\}$ implies that $i_0 \neq 0$ and that $u_0 \neq 0$. The replacement of A^0 in this equation by $-\sum_{i=1}^{n} A^i$ will then yield an equation which exhibits linear dependence among $A^1, \ldots, A^n$.

Finally, we obviously have $0 \in H\{A^0, \ldots, A^n\}$, showing (§ 4) that system (2) is inconsistent.

If the presence of A^0 vitiates the Haar or solvency condition at a subsequent juncture in the algorithm, the above technique may be repeated. Specifically, suppose that at a certain juncture the set I contains 0, and that the set $\{A^i : i \in I\}$ fails either of the two desired conditions. Then A^0 could be replaced by $-\sum_{\substack{i \in I \\ i \neq 0}} A^i$, and the computations may proceed.

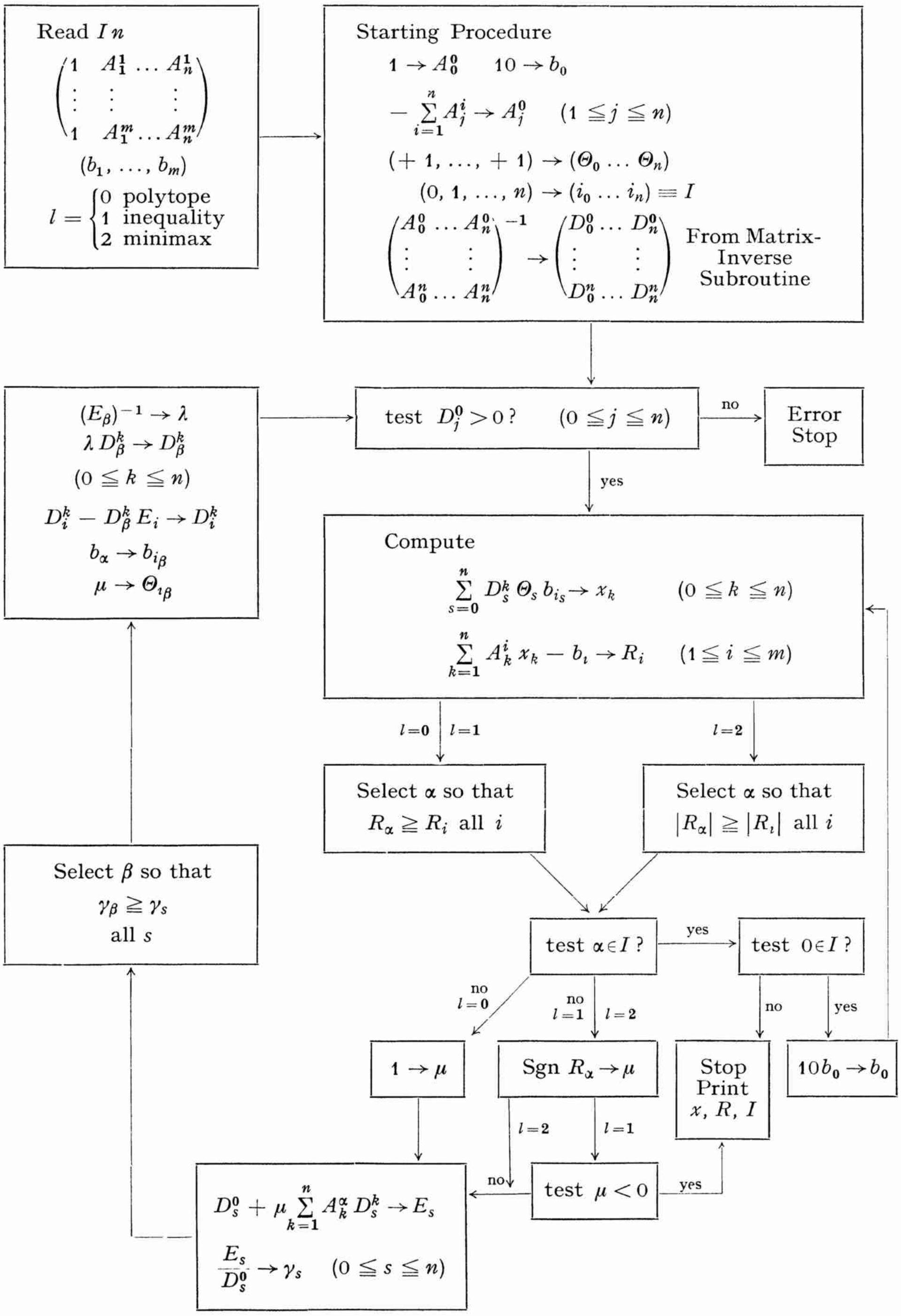
Since there is no a priori knowledge of the number $[A^0, x^0] - M$, b_0 is chosen in practice by trial, and the number

$$M^* = \inf_x \max_{0 \leq i \leq m} R^i(x)$$

is calculated by means of the algorithm. If, in the last cycle of the algorithm, $0 \in I$, then condition (1) above is not satisfied by b_0, and b_0 must be increased. As b_0 increases, M^* decreases, but if $M^* < \min_{1 \leq i \leq m} (-b_i)$, then by § 15, the (original) function F is not bounded below.

Assuming the Haar and Normality conditions, the Tchebycheff problem of minimizing the function $F(x) = \max_{1 \leq i \leq m} |R^i(x)|$ has an alternate starting procedure. Let J be any set of $n+1$ indices, and let x^* be a point which minimizes $\max_{i \in J} |R^i(x)|$. Such a point may be determined by methods of [18], for example. Define $R^{m+i} = -R^i(x)$. The set $I \equiv \{i : 1 \leq i \leq 2m, \ R^i(x^*) = F(x^*)\}$ is a satisfactory starting set.

§ 19. Formulary for Algorithm II



§ 20. Algorithm III. Consider the problem of obtaining the minimizing point for the function $F(x) = \sup_{A \in \Omega} [A, x] - b(A)$, Ω being a countable compact set in E_n and b a continuous function on Ω. Let the elements of Ω be enumerated: $A^1, A^2, \ldots$. For each m let $x^{(m)}$ be chosen to minimize the auxiliary function

$$F^{(m)}(x) = \max_{1 \leq i \leq m} R(A^i, x).$$

By § 7, there is a subset $\{A^{i_0}, \ldots, A^{i_n}\}$ of Ω having the property that the minimum of

$$F^*(x) = \max_{0 \leq j \leq n} R(A^{i_j}, x)$$

equals the minimum of F. Thus in the presence of the Haar condition, the sequence $x^{(1)}, x^{(2)}, \ldots$ is eventually stationary and gives the minimum point for F. The algorithm therefore converges in a finite number of steps.

§ 21. Algorithm IV. Let Ω and Λ denote bounded subsets of E_n and b and c bounded real-valued functions on Ω and Λ respectively. Define the closed convex set $K = \{x \in E_n : [B, x] \leq c(B)$, all $B \in \Lambda\}$. Define the continuous convex function $F(x) = \sup\{R(A, x) : A \in \Omega\}$, where $R(A, x) = [A, x] - b(A)$. It is desired to obtain, if such exists, a point y of K inducing a minimum value of F. Assume, in getting started, that finite sets $\Omega^0 \subset \Omega$ and $\Lambda^0 \subset \Lambda$ are available for which $0 \in \operatorname{int} H(\Omega^0 \cup \Lambda^0)$. In this connection, see § 18. At the m^{th} step there are given two finite sets $\Omega^m \subset \Omega$ and $\Lambda^m \subset \Lambda$. Define $F^m(x) = \sup\{R(A, x) : A \in \Omega^m\}$ and $K^m = \{x : [B, x] \leq c(B)$, all $B \in \Lambda^m\}$. Select x^m to minimize F^m on K^m. In this connection, see § 23. Select $A' \in \Omega$ to maximize $R(A, x^m)$ within a tolerance of $1/m$. Select $B' \in \Lambda$ to maximize $[B, x^m] - c(B)$ within a tolerance of $1/m$. Begin anew with $\Omega^{m+1} = \Omega^m \cup \{A'\}$ and $\Lambda^{m+1} = \Lambda^m \cup \{B'\}$.

§ 22. Theorem. If K is non-empty then algorithm IV is effective in the sense that

(i) $F^m(x^m) \nearrow p \equiv \inf\{F(x) : x \in K\}$;

(ii) the sequence $\{x^m : m = 0, 1, 2, \ldots\}$ possesses cluster points, each of which lies in K and minimizes F thereon.

Proof. We have finite sets Ω^0 and Λ^0 fulfilling $0 \in \operatorname{int} H(W)$ where $W = \Omega^0 \cup \Lambda^0$. Define $\vartheta = \inf_{\|x\|=1} \max_{w \in W} [w, x]$. If $\vartheta \leq 0$, then there exists an $x^0 \neq 0$ such that $[w, x^0] \leq 0$ for all $w \in W$. Hence $[w, x^0] \leq 0$ for all $w \in H(W)$. Since this is incompatible with the fact that $0 \in \operatorname{int} H(W)$, $\vartheta > 0$.

Now assume $K \neq \varphi$. Select $v \in K$ and define $M = \vartheta^{-1} \max[\sup_{B \in \Lambda} c(B), F(v) + \sup_{A \in \Omega} b(A)]$. Let x be an arbitrary point satisfying $\|x\| > M$. Select $w^0 \in W$ such that $[w^0, x] = \max_{w \in W} [w, x]$. If $w^0 \in \Omega^0$, then $F^0(x) \geq [w^0, x] - b(w^0) \geq \|x\| \vartheta - \sup_{A \in \Omega} b(A) > F(v) \geq F^0(v)$. On the other hand, if $w^0 \in \Lambda^0$, then $[w^0, x] - c(w^0) \geq \|x\| \vartheta - \sup_{B \in \Lambda} c(B) > 0$, so that $x \notin K^0$. This argument establishes that $\inf_{x \in K^0} F^0(x) = \inf_{x \in K^0, \|x\| \leq M} F^0(x)$, so that, due to the continuity of F^0, the infimum on the left is attained. Thus x^0 exists. By the same argument, $\inf_{x \in K} F(x) = \inf_{x \in K, \|x\| \leq M} F(x) \equiv p < \infty$. The following inequality is obvious: $F^m(x^m) = \inf_{x \in K^m} F^m(x) \leq \inf_{x \in K^{m+1}} F^{m+1}(x) = F^{m+1}(x^{m+1}) \leq p$. Since $F^0(x^m) \leq F^m(x^m) \leq p \leq F(v)$, and since $x^m \in K^0$, we know by virtue of the preceding paragraph that $\|x^m\| \leq M$. Let y be any cluster point of the sequence $x^0, x^1, x^2, \ldots$. If for some m, $y \notin K^m$, then select $\delta > 0$ such that $[B^m, y] - c(B^m) > \delta$. Select $i \leqq m$ so that $\|x^i - y\| \leqq \delta/r$ where $r = \sup_{B \in \Lambda} \|B\|$. Then $[B^m, x^i] - c(B^m) = [B^m, y] - c(B^m) + [B^m, x^i - y] > \delta - \|B^m\| \cdot \|x^i - y\| \geqq 0$, contradicting $x^i \in K^i \subset K^m$. Hence $y \in \bigcap_{m=0}^{\infty} K^m$. Now define $G(x) = \sup_{B \in \Lambda} [B, x] - c(B)$. If $y \notin K$, then put $\delta = \frac{1}{3} G(y)$. Take $m \geqq 1/\delta$ so that $\|x^m - y\| < \delta/r$. Then $G(y) < G(x^m) + \delta \leqq [B', x^m] - c(B') + \frac{1}{m} + \delta < [B', y] - c(B') + 3\delta \leqq 3\delta = G(y)$, a contradiction. Note that we use here the fact that $y \in K^{m+1}$. Thus $y \in K$.

It was shown earlier that $F^m(x^m)$ is non-decreasing and bounded by p. Thus $F^m(x^m) \nearrow p - 3\varepsilon$ for some $\varepsilon \geqq 0$. If $\varepsilon > 0$, select $m \geqq 1/\varepsilon$ so that $\|y - x^m\| < \varepsilon/2q$, where $q = \sup_{A \in \Omega} \|A\|$. Then $F(x^m) > F(y) - \varepsilon \geqq p - \varepsilon$. Take $i > m$ so that $\|y - x^i\| < \varepsilon/2q$. Using $R^{m+1}(x) = [A', x] - b(A')$, we have $R^{m+1}(x^m) - R^{m+1}(x^i) \geqq F(x^m) - \frac{1}{m} - F^i(x^i) \geqq F(x^m) - \varepsilon - p + 3\varepsilon \geqq \varepsilon$. On the other hand, $R^{m+1}(x^m) - R^{m+1}(x^i) \leqq \|A'\| \cdot \|x^m - x^i\| < \varepsilon$, a contradiction. Hence $\varepsilon = 0$, establishing (i). As for (ii), observe that every cluster point y of $\{x^m\}$ lies in K. Then as above, $p \leqq F(y) \leqq F(x^m) + \delta \leqq R^{m+1}(x^m) + 2\delta < R^{m+1}(x^i) + 3\delta \leqq F^i(x^i) + 3\delta < p + 3\delta$, Q.E.D.

§ 23. Algorithm V. Define $F(x) = \max_{1 \leqq i \leqq k} R^i(x)$ and $G(x) = \max_{k < i \leqq m} R^i(x)$, where $R^i(x) = [A^i, x] - b_i$ and $1 \leqq k \leqq m$. It is desired to obtain the minimum of F on the domain $K = \{x \in E_n : G(x) \leqq 0\}$, if such exists. It is assumed that $\{A^1, \ldots, A^m\}$ satisfies the Haar condition.

In each cycle of the algorithm, a set $I \subset \{1, \ldots, m\}$ is given such that

(i) $o(I) = n + 1$
(ii) $0 \in H\{A^i : i \in I\}$
(iii) $I \cap \{1, \ldots, k\}$ non-empty.
(See § 18 for starting procedure.)
Obtain a point x and a number μ such that
(iv) $i \in I \cap \{1, \ldots, k\} \Rightarrow R^i(x) = \mu$
(v) $i \in I \cap \{k+1, \ldots, m\} \Rightarrow R^i(x) = 0$.
If $G(x) \leqq 0$ and $F(x) = \mu$, then x is a solution.
If $G(x) > 0$, select $p \in \{k+1, \ldots, m\}$ so that $R^p(x) > 0$.

If $G(x) \leqq 0$ and $F(x) > \mu$, select $p \in \{1, \ldots, k\}$ so that $R^p(x) > \mu$. In both the latter cases, select $q \in I$ so that $0 \in H\{A^i : i \in I'\}$ where $I' = I \cup \{p\} - \{q\}$. See § 10 in this connection. Begin the next cycle with I' in place of I.

§ 24. Effectiveness of Algorithm V. Consider a set I satisfying (i), (ii), and (iii). Put $I = \{i_0, \ldots, i_n\}$, with $\{i_0, \ldots, i_j\} \subset \{1, \ldots, k\}$ and $\{i_{j+1}, \ldots, i_n\} \subset \{k+1, \ldots, m\}$. We show first that the following matrix is non-singular.

$$B = \begin{pmatrix} A_1^{i_0} & \ldots & A_n^{i_0} & 1 \\ \vdots & & \vdots & \vdots \\ A_1^{i_j} & \ldots & A_n^{i_j} & 1 \\ A_1^{i_{j+1}} & \ldots & A_n^{i_{j+1}} & 0 \\ \vdots & & \vdots & \vdots \\ A_1^{i_n} & \ldots & A_n^{i_n} & 0 \end{pmatrix}.$$

Suppose, on the contrary, that there exists a non-zero vector $u = (u_1, \ldots, u_{n+1})$ for which $Bu = 0$. Then $[A^{i_p}, u^*] = -u_{n+1}$ for $0 \leqq p \leqq j$ and $[A^{i_p}, u^*] = 0$ for $j+1 \leqq p \leqq n$, where $u^* = (u_1, \ldots, u_n)$. If $u_{n+1} = 0$ then the Haar condition is violated. If $u_{n+1} \neq 0$, then write, in accordance with (ii): $0 = \sum \lambda_p A^{i_p}$ with $\sum \lambda_p = 1$ and $\lambda_p \geqq 0$. By the Haar condition, $\lambda_p > 0$. Thus $0 = [0, u] = \sum \lambda_p [A^{i_p}, u]$, a contradiction. Thus there is no difficulty in obtaining x and μ satisfying (iv) and (v).

We show now that if $G(x) \leqq 0$ and $F(x) = \mu$ then x is a solution. If not then there exists y such that $F(y) < \mu$ and $G(y) \leqq 0$. Then the vector $z = y - x$ satisfies the inequalities

$$\begin{aligned} (A^i, z) &< 0 \qquad i \in I \cap \{1, \ldots, k\} \\ (A^i, z) &\leqq 0 \qquad i \in I \cap \{k+1, \ldots, m\}, \end{aligned}$$

and one obtains a contradiction as above. Hence x is a solution.

If $G(x) > 0$ or $F(x) > \mu$ then the algorithm specifies how to obtain a set I' satisfying conditions (i) and (ii). We now show that I' satisfies (iii) as well. If not, then $o(I \cap \{1, \ldots, k\}) = 1$ and $k < p \leqq m$. Select $y \in K$. Then $[A^i, y] \leqq b_i$ for $k < i \leqq m$. Thus $[A^i, y - x] \leqq 0$ for $i \in I \cap \{k+1, \ldots, m\}$ and $[A^p, y - x] < 0$. This contradicts property (ii) for I'. Thus there is no difficulty in obtaining an x' and μ' satisfying

$$\begin{aligned} [A^i, x'] - b_i &= \mu' \qquad i \in I' \cap \{1, \ldots, k\} \\ [A^i, x'] - b_i &= 0 \qquad i \in I' \cap \{k+1, \ldots, m\}. \end{aligned}$$

By subtracting, we find

$$\begin{aligned} [A^i, x' - x] &= \mu' - \mu \quad i \in I \cap I' \cap \{1, \ldots, k\} \\ [A^i, x' - x] &= 0 \quad i \in I \cap I' \cap \{k+1, \ldots, m\} \\ [A^p, x' - x] &< \mu' - \mu \quad \text{when} \quad p \in \{1, \ldots, k\} \\ [A^p, x' - x] &< 0 \quad \text{when} \quad p \in \{k+1, \ldots, m\}. \end{aligned}$$

Since $0 \in H\{A^i : i \in I'\}$, $\mu' - \mu > 0$. Thus in proceeding from one cycle of the algorithm to the next, the value of μ increases. Since there are but a finite number of sets I satisfying (i), the algorithm will terminate with a point x satisfying $F(x) = \mu$ and $G(x) \leqq 0$. This completes the proof. It may be observed that if K is empty the algorithm will indicate this fact by the impossibility of computing x and μ in some cycle.

§ 25. Algorithm VI. It is desired to obtain $x^* \in E_n$ minimizing the function $F(x) = \max_{1 \leqq i \leqq m} R^i(x)$. This algorithm has the feature that at the k-th step, an upper bound r_k and a lower bound s_k are provided for the unknown number $p \equiv \inf_x F(x)$; furthermore, $r_k \searrow p$ and $s_k \nearrow p$. The desirability of such a feature was pointed out in [9].

Assume now that the Haar and normality conditions are fulfilled and that F is bounded below. For any $I \subset \{1, \ldots, m\}$ define $F_I(x) = \max_{i \in I} R^i(x)$. In each computing cycle there will be given a point $x \in E_n$ and a set $I \subset \{1, \ldots, m\}$, the latter satisfying $o(I) = n+1$ and $0 \in H\{A^i : i \in I\}$. Define then $J = \{j : R^j(x) = F(x)\}$.

Select z to minimize $F_{I\cup J}$. If $z=x$, then z is a solution. Otherwise proceed to select an x' which will minimize F on the ray $\{x+t(z-x): t\geqq 0\}$. Define $I'=\{i\in I\cup J: R^i(z)=F_{I\cup J}(z)\}$. Begin anew with x' and I'. See § 18 for starting procedure.

§ 26. Effectiveness of Algorithm VI. The proof will be given in six parts.

(1) The point x' is well-defined because on the indicated ray, F is a polygonal function which is bounded below and thus attains its minimum (§ 3 C).

(2) If $x\neq z$ then $F(x')<F(x)$. To prove this, observe that for each $j\in J$, $R^j(z)\leqq F_{I\cup J}(z)<F_{I\cup J}(x)\leqq F(x)$, the strict inequality being due to the uniqueness of z (a consequence of the Haar condition). Thus for small $t>0$, $F(x+t(z-x))<F(x)$. Note that $F(x)$ has then the properties claimed above for r_k.

(3) I' satisfies the conditions laid down for I. To prove this, observe that the system of inequalities $[A^i, d]<0$, $(i\in I')$ is inconsistent. Thus $0\in H\{A^i: i\in I'\}$. By the Haar condition, then, $o(I')>n$. By the normality condition, $o(I')\leqq n+1$.

(4) If $x=z$ then z is a solution. Indeed if $x=z$, then x minimizes $F_{I\cup J}$. By the uniqueness of z, $F_{I\cup J}$ is increasing in a neighborhood of z. Now $F(x)=F_{I\cup J}(x)$ and $F_{I\cup J}\leqq F$ always. Thus F is increasing in a neighborhood of x, and x must be a solution.

(5) If $x'\neq z'$ then $F_{I''}(z')>F_{I'}(z)$. To establish this, note first that $\min F_{I'}\leqq \min F_{I'\cup J'}=\min F_{I''}$. If equality occurs here, then $z=z'$ due to uniqueness of z. In this event, $x'=x''$ because x, x', z are colinear, as are x', x'', z', so that the minimum of F on the ray xz occurs with the minimum of F on the ray $x'z'$. Hence $x'=z'$ as well. Note that $F_{I'}(z)$ has therefore the properties claimed above for s_k.

(6) There are but a finite number of sets I in $\{1, \ldots, m\}$, and in each cycle of the algorithm a new I occurs because of (5). Thus the algorithm terminates at some cycle in which $x=z$, such a point being a solution, by (4). This concludes the proof.

§ 27. Approximation in $C(T)$. Let T denote a compact metric space and $C(T)$ the linear space of all continuous real-valued functions defined on T. For any subset S of T define $|S|=\sup_{t\in T}\inf_{s\in S} d(s,t)$, and define a semi-norm in $C(T)$ by writing $\|f\|_S=\sup_{s\in S}|f(s)|$. Let M denote a given subset of $C(T)$ and g a fixed element of $C(T)$. The problem of approximating g by elements of M is to be investigated. Specifically, given $\varepsilon>0$, it is desired to obtain by a simple algorithm an $f_0\in M$ fulfilling the condition

$$\|f_0-g\|_T-\varepsilon<p\equiv\inf_{f\in M}\|f-g\|_T.$$

Since in practice it is easier to compute the semi-norms $\|\ \|_S$ instead of $\|\ \|_T$, it is advantageous that the following principle be valid.

(P) To each $\varepsilon>0$ there corresponds a $\delta>0$ such that $p-\varepsilon<\|f_0-g\|_S\leqq\|f_0-g\|_T<p+\varepsilon$ whenever $|S|<\delta$ and $\|f_0-g\|_S<\delta+\inf_{f\in M}\|f-g\|_S$.

§ 28. Lemma. Let M designate a finite dimensional subspace of $C(T)$. There exist two constants $q>0$ and $Q>0$ such that $\|f\|_T\leqq Q$ whenever $|S|\leqq q$ and $\|f-g\|_S\leqq p+1$.

Proof. Let $\{f_1, \ldots, f_n\}$ be a basis for M, and define for each $t \in T$ an n-tuple $A(t) = (f_1(t), \ldots, f_n(t))$. Due to the independence of the f_i's, the set $\{A(t): t \in T\}$ has a zero orthogonal complement in E_n and thus contains a basis $\{A(t_1), \ldots, A(t_n)\}$. Hence the $n \times n$ determinant whose entries are $A_{ij} = f_i(t_j)$ is non-zero, and by the continuity of the determinant function, there exist positive numbers δ and r such that $|\det B_{ij}| \geqq r$ whenever $\max_{ij} |B_{ij} - A_{ij}| \leqq \delta$. Select $q > 0$ so that $|f_i(s) - f_i(t)| \leqq \delta$ whenever $d(s, t) \leqq q$. Assume $|S| \leqq q$, $\|f - g\|_S \leqq p + 1$, and $f = \sum_{i=1}^{n} x_i f_i$. For each j select $s_j \in S$ satisfying $d(s_j, t_j) \leqq q$. Clearly $|f(s_j) - g(s_j)| \leqq p + 1$. Thus $|\sum x_i f_i(s_j)| \leqq p + 1 + \|g\|_T \equiv c$. Since $d(s_j, t_j) \leqq q$, $|f_i(s_j) - f_i(t_j)| \leqq \delta$, and $|\det f_i(s_j)| \geqq r$. By CRAMER's Rule, each $|x_i|$ has an upper bound $d \equiv c n!\, r^{-1} (\delta + \max |A_{ij}|)^{n-1}$, whence $\|f\|_T \leqq \sum |x_i| \cdot \|f_i\|_T \leqq d \sum \|f_i\|_T \equiv Q$.

§ 29. Theorem. Principle (P) is valid under either of the two following conditions:

(i) M is an equicontinuous subset of $C(T)$;

(ii) M is a subset of a finite dimensional subspace of $C(T)$.

Proofs. (i) Given $\varepsilon > 0$, take $\delta < \varepsilon/3$ such that $|f(s) - f(t)| < \varepsilon/3$ and $|g(s) - g(t)| < \varepsilon/3$ whenever $f \in M$ and $d(s,t) < \delta$. Suppose $|S| < \delta$ and $\|f_0 - g\|_S < \delta + \inf_f \|f - g\|_S$. Select $t \in T$ so that $|f_0(t) - g(t)| = \|f_0 - g\|_T$. Select $s \in S$ so that $d(s, t) < \delta$. Then $p \leqq \|f_0 - g\|_T = |f_0(t) - g(t)| \leqq |f_0(t) - f_0(s)| + |f_0(s) - g(s)| + |g(s) - g(t)| \leqq \varepsilon/3 + \|f_0 - g\|_S + \varepsilon/3 < \varepsilon + \inf_f \|f - g\|_S \leqq \varepsilon + p$. Thus $p - \varepsilon < p - 2\varepsilon/3 \leqq \|f_0 - g\|_S \leqq \|f_0 - g\|_T < \varepsilon + p$.

(ii) By the Lemma, if $\delta < \min(1, q)$, $|S| \leqq \delta$, and $\|f_0 - g\|_S \leqq \delta + \inf_{f \in M} \|f - g\|_S$, then $\|f_0\|_T \leqq Q$. Thus the approximating functions are taken from a bounded subset of a finite-dimensional subspace of $C(T)$, which is therefore equicontinuous. Hence (ii) reduces to (i).

§ 30. Examples. *A.* Let $T = [0, 1]$ and let M consist of all functions on T having a first-derivative bounded in modulus by a constant k. Then M is equicontinuous since $|f(s) - f(t)| = |f'(v)| \cdot |s - t| \leqq k\,|s - t|$. This M is infinite dimensional, containing the independent set $\{e^{at}: 0 \leqq a \leqq \frac{1}{2} \log k\}$.

B. Now let U denote an arbitrary (non-topological) set and $B(U)$ the Banach space of all bounded real-valued functions on U, normed by $\|\varphi\| = \sup_{u \in U} |\varphi(u)|$. Let N designate a finite dimensional subspace of $B(U)$ spanned by $\{\varphi_1, \ldots, \varphi_n\}$. Let ϑ denote any fixed element of $B(U)$. We seek $\varphi^* \in N$ such that $\|\varphi^* - \vartheta\| \leqq \|\varphi - \vartheta\|$ for all $\varphi \in N$. This problem may be treated by Algorithm I after recasting the problem as follows. For each $u \in U$, define an n-tuple $A^u = (\varphi_1(u), \ldots, \varphi_n(u))$. We then seek $x^* \in E_n$ which minimizes the function $F(x) = \sup_{u \in U} |[A^u, x] - \vartheta(u)|$.

It is also possible to recast the problem into the form of § 27. Define $B^u = (\varphi_1(u), \ldots, \varphi_n(u), \vartheta(u))$ and denote by T the closure of the set $\{B^u: u \in U\}$ in E_{n+1}. On E_{n+1} define the functions $f_i(y) = i$-th component of y, $1 \leqq i \leqq n+1$. Clearly T is compact and $f_i \in C(T)$. It turns out that for each $x \in E_n$, $\|\sum x_i \varphi_i - \vartheta\|_U = \|\sum x_i f_i - f_{n+1}\|_T$. Thus by § 29, Principle (P) is valid.

References

[1] Remez, E.: Sur un Procédé Convergent d'Approximations Successives pour Determiner les Polynomes d'Appproximation. C. R. Acad. Sci. Paris **198**, 2063—2065 (1934).

[2] Remez, E.: Sur le Calcul Effectif des Polynomes d'Approximation de Tschebyscheff. C. R. Acad. Sci. Paris **199**, 337—340 (1934).

[3] Remez, E. Ya.: On the Method of Best, in the Sense of Tchebycheff, Approximate Representation of Functions, (Ukrainian), Kiev, 1935. See also Reference 4 below.

[4] Remez, E. Ya.: General Computation Methods for Chebyshev Approximation. Problems with Real Parameters Entering Linearly. Izdat. Akad. Nauk Ukrainsk. SSR. Kiev, 1957. 454 pp. See also MR 19-580 (Russian).

[5] Novodvorskii, E. N., and I. Sh. Pinsker: On a Process of Equalization of Maxima, Uspehi Matem. Nauk, N.S. **6**, 174—181 (1951). See also Shenitzer, A.: Chebyshev Approximations. J. Assoc. Comput. Mach. **4**, 30—35 (1957), MR 13-728.

[6] Beale, E. M. L.: An Alternative Method for Linear Programming. Proc. Cambridge Philos. Soc. **50**, 512—523 (1954). MR 16-155.

[7] Beale, E. M. L.: On Minimizing a Convex Function Subject to Linear Inequalities. J. Roy. Stat. Soc., Ser. B **17**, 173—177 (1955).

[8] Bratton, Donald: New Results in the Theory and Techniques of Chebyshev Fitting. Abstract 546-34. Notices Am. Math. Soc. **5**, 248 (1958).

[9] Stiefel, Eduard L.: Numerical Methods of Tchebycheff Approximation, pp. 217—232 in R. E. Langer (ed.), On Numerical Approximation. Madison 1959. 480 pp.

[10] Stiefel, E.: Über diskrete und lineare Tschebyscheff-Approximationen. Numerische Mathematik **1**, 1—28 (1959).

[11] Wolfe, Philip: Programming with Nonlinear Constraints. Preliminary Report, Abstract 548-102. Notices Am. Math. Soc. **5**, 508 (1958).

[12] Stone, Jeremy J.: The Cross Section Method, presented orally at Symposium for Mathematical Programming, RAND Corporation, March 19, 1959.

[13] Kelley, James E.: A General Technique for Convex Programming, presented orally at Symposium on Mathematical Programming, RAND Corporation, March 19, 1959.

[14] Cheney, E. W., and A. A. Goldstein: Proximity Maps for Convex Sets. Proc. Amer. Math. Soc. **10**, 448—450 (1959).

[15] Bonnesen, T., u. W. Fenchel: Theorie der konvexen Körper. Berlin 1934.

[16] Fan, K.: On Systems of Linear Inequalities. In: Linear Inequalities and Related Systems, ed. by H. W. Kuhn and A. W. Tucker, pp. 99—156. Princeton 1956.

[17] Bram, Joseph: Chebychev Approximation in Locally Compact Spaces. Proc. Am. Math. Soc. **9**, 133—136 (1958).

[18] Goldstein, A. A., and E. W. Cheney: A Finite Algorithm for the Solution of Consistent Linear Equations and Inequalities and for the Tchebycheff Approximation of Inconsistent Linear Equations. Pac. J. Math. **8**, 415—427 (1958).

Convair Astronautics
Department 591.10
San Diego 12, California

(Received July 3, 1959)