Mathematical Programming 8 (1975) 189–206.
North-Holland Publishing Company

A FINITE ALGORITHM TO MAXIMIZE CERTAIN PSEUDOCONCAVE FUNCTIONS ON POLYTOPES *

Balder von HOHENBALKEN
University of Alberta, Edmonton, Alberta, Canada

Received 5 April 1974
Revised manuscript received 15 January 1975

This paper develops and proves an algorithm that finds the exact maximum of certain nonlinear functions on polytopes by performing a finite number of logical and arithmetic operations. Permissible objective functions need to be pseudoconcave and allow the closed-form solution of sets of equations $\partial f(Dy + \hat{x}^k)/\partial y = 0$, which are first order conditions associated with the unconstrained, but affinely transformed objective function. Examples are pseudoconcave quadratics and especially the homogeneous function $cx + m(xVx)^{1/2}, m < 0$, V positive definite, for which sofar no finite algorithm existed.

In distinction to most available methods, this algorithm uses the internal representation [6] of the feasible set to selectively decompose it into simplices of varying dimensions; linear programming and a gradient criterion are used to select a sequence of these simplices, which contain a corresponding sequence of strictly increasing, relative and relatively interior maxima, the greatest of which is shown to be the global maximum on the feasible set. To find the interior maxima on these simplices in a finite way, calculus maximizations on the affine hulls of subsets of their vertices are necessary; thus the above requirement that $\partial f(Dy + \hat{x}^k)/\partial y = 0$ be explicitly solvable.

The paper presents a flow structure of the algorithm, its supporting theory, its decision-theoretic use, and an example, computed by an APL-version of the method.

1. The algorithm in this paper is the product of a search for a finite method to solve programs of the type

$$\max_{x \in X} \{f(x) \equiv cx + m\,(xVx)^{1/2}\}$$

$$X \equiv \{x \in \mathbf{R}^n : Ax \leqslant b,\ x \geqslant 0\}, \qquad m < 0, \quad V \text{ positive definite.}$$

The search was motivated by Van Moeseke's work [9] who developed a criterion for decision-making under risk (see Section 5), whose empirical implementation needed a solution method for the above problem.

* Research supported in part by The Canada Council, Ottawa.

Earlier versions of the paper were presented at the European meeting of the Econometric Society in Budapest, September 1972 and at the meeting of the Econometric Society in San Francisco, December 1974.

Van Moeseke himself devised an infinite algorithm that depended essentially on the homogeneity of f, and some of the fundamental ideas of the method discussed here are due to him; Dréze and Van Moeseke [3] later proposed a finite method for homogeneous programs on unit simplices.

The algorithm to be discussed is a primal method, in the sense that it attacks the given primal problem, and the algorithmic sequence starts and remains in the primal feasible set. On polytopes it finds, after a finite number of logical and arithmetic operations, the exact maximum of objective functions, that are pseudoconcave and allow the closed-form solution of sets of equations $\partial f(Dy + \hat{x}^k)/\partial y = 0$ (these are first-order conditions for maxima on internally represented manifolds, to be discussed below). Examples that exhibit these properties are the homogeneous programs stated above, and pseudoconcave quadratic programs [5].

The well-known primal methods of Rosen [7], Wolfe [13], Zangwill [14] and others approach a maximizer on the feasible polytope X by searching on (feasible pieces of) tangent manifolds, that are directly defined by the constraints active at a particular instance. In Rockafellar's terminology, the polytope and these manifolds are said to be externally represented, i.e., as the intersection of halfspaces and hyperplanes, respectively. Our approach uses the *internal* representation of the feasible polytope, based on Carathéodory's theorem:

Theorem 1.1. *Be $X \subset \mathbf{R}^n$ a nonempty convex polytope; then every $x \in X$ lies in the relative interior of one of a finite number of simplices whose vertices are extreme points of X. The union of this collection of simplices equals X.*

Proof. See Rockafellar [6, Theorem 17.1, p. 155; Corollary 18.5.1, p. 167; Corollary 19.1.1, p. 172].

The choice of suitable extreme points of X is made by gradient-guided linear programs, which furnish the bridge between external and internal representation. At each step a new affine basis (see definitions below) is proposed, which generates a manifold and an equidimensional embedded simplex. Via calculus maximizations on this manifold and/or suitably chosen submanifolds (see Section 3) members of the basis with nonpositive barycentric weights are dropped, until a relative maximizer with all positive barycentric coordinates is found, which then lies in the relative interior of the simplex spanned by the residual basis. In this way

the algorithm generates a sequence of simplices with strictly increasing relative maxima, the largest of which must appear after finitely many steps and is the maximum of f on X.

The simplicial decomposition principle based on Carathéodory's theorem[1] can be profitably applied in more general algorithms, which locate maxima in simplices by line searches of more or less sophisticated nature [11]. Algorithms of this kind are capable of finding local maxima of arbitrary functions on polytopes, but finiteness is lost, of course. A further generalization to nonlinear constraints is not possible, because these methods depend essentially on linear programs to search among the finite number of extreme points of polyhedral sets.

Notation. Let $x \in \mathbf{R}^n$, with elements $x_i \in R$;
$x_i > 0$ and $x_i \geqslant 0$ have the usual meaning;
$x > 0$ means $x_i > 0$ for all i,
$x \geqq 0$ means $x_i \geqslant 0$ for all i,
$x \geqslant 0$ means $x \geqq 0$ and $x \neq 0$.

Vectors with special properties (e.g., being a maximizer, an extreme point, etc.) are superscripted: x^*, $\hat{x}^k$, etc.

The gradient of f is denoted by $\partial f/\partial x$ or f_x; evaluated at x^*, we write it as $\partial f/\partial x|_{x=x^*}$ or f_{x^*}. Inner products are stated by mere juxtaposition; e.g., $f_{x^*}(x-x^*)$ is written for $\Sigma_{i=1}^n \partial f/\partial x_i|_{x_i=x_i^*}(x_i-x_i^*)$; the transposition sign "T" is used only to avoid ambiguities in matrix products.

Familiarity with the elementary theory of convex sets is assumed. The following definitions cover only some special notions that are frequently employed in the paper (see also [6]).

The set of points $B \equiv \{\hat{x}^1, \ldots, \hat{x}^k\}$ is *affinely independent* if the convex hull S of B is $(k-1)$-dimensional; B is then called an *affine basis,* and S a *simplex.*

An *affine combination* is a linear combination whose weights sum to unity; the set M of all affine combinations of a given set B is a (linear) *manifold; M* is also the *affine hull* of B; if B is an affine basis, the unique weights expressing any $x \in M$ are called *barycentric coordinates.*

[1] This procedure has little in common with the decomposition principle of Dantzig and Wolfe [2], and it is not related to the simplex interpretation of Dantzig's Simplex method [1, 1963, ch. 23].

$T : \mathbf{R}^k \to \mathbf{R}^n$ is an affine transformation if $Ty = Dy + \hat{x}$, where D is a linear transformation and $\hat{x} \in \mathbf{R}^n$.

The interior of a *convex set* $S \subset \mathbf{R}^n$ relative to $\mathbf{R}^n$ is denoted by int S; the interior of S relative to the affine hull of S is denoted by ri S. Clearly, $S \neq \emptyset$ implies ri $S \neq \emptyset$; if the dimensionality of $S \subset \mathbf{R}^n$ equals n, then int S = ri S.

A vertex $\hat{x}^i$ of a simplex S is a *carrier* of $x \in S$ if the i^{th} barycentric coordinate of x is positive; $x \in$ ri S if all vertices of S carry x.

$f : \mathbf{R}^n \to \mathbf{R}$ is a *pseudoconcave* function if $f_x(y - x) \leqslant 0$ implies $f(y) \leqslant f(x)$ for all $x, y \in \mathbf{R}^n$ [4, ch. 9].

2. This section presents a flow structure of the algorithm, which gives computational instructions and bivalued decision criteria for every step. Given are a feasible polytope

$$X \equiv \{x \in \mathbf{R}^n : \; Ax \leqq b, \; x \geqq 0\},$$

and a pseudoconcave objective function $f : \mathbf{R}^n \to \mathbf{R}$, that permits an explicit solution of $\partial f(Dy + \hat{x}^k)/\partial y = 0$.

The next two sections contain the supporting theorems and a commentary on this flow structure.

The algorithm

Initial step. Solve the linear program

$$\max_{x \in X} f_0 x$$

where $f_0 = \partial f/\partial x|_{x=0}$. This locates an extreme point of X, say $\hat{x}^1$; for $t = 0$ set $x^{t+1} = x^1 = \hat{x}^1$, $B^{t+1} = B^1 = \{\hat{x}^1\}$, where B^1 is the affine basis of the zero-dimensional simplex S^1, with maximizer $x^1 \in$ ri S^1; go to basic step 1.

Basic step 1. Set $x^t = x^{t+1}$, $B^t = B^{t+1}$ and let S^t and M^t be the simplex and the manifold generated by B^t; use linear programming to locate the extreme point $\hat{x}^k$ that solves

$$\max_{x \in X} f_{x^t} x.$$

(a) If $f_{x^t}(\hat{x}^k - x^t) = 0$, terminate, with $x^t \in$ ri S^t being a maximizer of f on S^t and on X;

(b) if $f_{x^t}(\hat{x}^k - x^t) > 0$, go to 2.

Basic step 2. (a) If $f(\hat{x}^k) > f(x^t)$, set $x^{t+1} = \hat{x}^k$, $B^{t+1} = \{\hat{x}^k\}$ and go to 1;

(b) if $f(\hat{x}^k) \leqslant f(x^t)$, augment the basis B^t by $\hat{x}^k$ to form the new affine basis $B \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}, \hat{x}^k\}$; B generates a simplex $S \subset X$ and a manifold $M \supset S$, on which x^t is not a maximizer; go to 3.

Basic step 3. Attempt to find a maximizer of f on the manifold M, generated by B, using an affine transformation of the argument x.

(a) If f possesses a maximizer x^* on M, it satisfies $f(x^*) > f(x^t)$; go to 4;

(b) if f does not possess a maximizer on M, find its maximizer x' on M', where M' is the manifold through $\hat{x}^k$ and parallel to $M \cap M^t$, again using an affine transformation; x' exists and satisfies $f(x') > f(x^t)$.

The barycentric representation of x' w.r.t. the basis B is $x' = Bw' = x'w'_1 + ... + \hat{x}^k w'_k$, with $w'_k > 0$ and at least one $w'_{i \neq k} \leqslant 0$, i.e., $x' \notin \text{ri } S$; go to 5.

Basic step 4. The barycentric representation of the maximizer x^* of f on M is $x^* = Bw^* = \hat{x}^1 w_1^* + ... + \hat{x}^k w_k^*$, with $w_k^* > 0$ (Section 4).

(a) If $w_i^* > 0$ for all $i = 1, ..., k$, $x^* \in \text{ri } S$ is a maximizer of f on S; set $x^{t+1} = x^*$, $B^{t+1} = B$ and go to 1,

(b) if at least one $w_{i \neq k}^* \leqslant 0$, i.e., $x^* \notin \text{ri } S$, go to 5.

Basic step 5. Intersect the segment $\overline{x^t x'}$, or the segment $\overline{x^t x^*}$ with the boundary of S; the intersection point $x^r = Bw^r = x^1 w_1^r + ... + x^k w_k^r$ satisfies $f(x^r) > f(x^t)$ and will have $w_i^r \geqslant 0$, all i, with $w_k^r > 0$ and at least one $w_{i \neq k}^r = 0$. Drop the vertices $\hat{x}^i$ that are not carriers of x^r from B to get the reduced affine basis B^r; B^r generates a sub-simplex $S^r \subset S$ such that $x^r \in \text{ri } S^r$.

(a) If S^r is zero-dimensional, x^r is its only point and thus a relatively interior maximizer of f on S^r; set $x^{t+1} = x^r$, $B^{t+1} = B^r$ and go to 1;

(b) if S^r has dimensionality larger than zero, set $x^t = x^r$, $B = B^r$, $M = M^r$, $S = S^r$ and go to 3.

3. For simplicity of exposition we have assumed that the linear programming algorithm used in the initial step and in basic step 1 delivers *some* extreme point of X if f_0 or f_{x^t} happen to be the zero vector.

Basic step 1 is entered either from the initial step or from basic step 4. In both cases, $x^t \in \text{ri } S^t$ is a maximizer on the simplex S^t and on the manifold M^t, which are both spanned by the affine basis B^t. Given x^t, the linear program guided by the gradient $f_{x^t} \equiv \partial f / \partial x|_{x = x^t}$ picks another extreme point $\hat{x}^k \in X$. With its help it can be decided whether x^t is also a maximizer of f on X.

Theorem 3.1. *Let f be pseudoconcave, and let $\hat{x}^k$ solve* $\max_{x \in X} f_{x^t} x$; *then x^t maximizes f on X if and only if* $f_{tx}(\hat{x}^k - x^t) = 0$.

Proof. It is easy to show that the condition is necessary for a maximum of any differentiable function (e.g., see [9, Lemma 3.1.5, p. 236]). Sufficiency follows directly from pseudoconcavity of f.

Equivalently, $f_{x^t}(\hat{x}^k - x^t) > 0$ means that x^t is *not* a maximizer on X.

The first step of basic step 2 permissibly short-circuits the algorithmic procedure in case $f(\hat{x}^k) > f(x^t)$ by returning to a zero-dimensional simplex, as after the initial step. This prevents a certain matrix used in a calculus maximization below from ever becoming empty (see Remark 4.3).

With $f(\hat{x}^k) \leqslant f(x^t)$ one still has $f_{x^t}(\hat{x}^k - x^t) > 0$; this means $\hat{x}^k \notin M^t$, because x^t being a maximizer on M^t implies $f_{x^t}(x - x^t) = 0$ for all $x \in M^t$. Thus the augmented set $B \equiv \{\hat{x}^1, \ldots, \hat{x}^{k-1}, \hat{x}^k\}$ is an affine basis, that spans the simplex S; replacing X by S in Theorem 3.1 shows that x^t is not a maximizer of f on S, and a fortiori not on $M \supset S$.

Basic step 3 is the centerpiece of the algorithm; the ingredients available at this point are the bases B^t and B, that span the simplices and manifolds S^t and M^t, S and M, respectively. From previous discussion it is clear that $B^t \subset B$, $S^t \subset S \subset M$, $S^t \subset M^t \subset M$. We also have $x^t \in \text{ri } S^t$, that maximizes f on M^t, but not on M. It is therefore obvious that M must own points whose f-value is strictly larger than $f(x^t)$, and a maximizer x^* on M, if it exists, will satisfy $f(x^*) > f(x^t)$. (See Figs. 1a and 1b).

For certain classes of functions it can be decided by calculus criteria whether $x^* \in M$ exists, and if this is the case, x^* can be found exactly by calculus (see Section 5). If x^* happens to lie in ri S, an iteration has been completed successfully with $x^{t+1} = x^*$. If $x^* \notin \text{ri } S$, the location of x^* is used to suitably reduce B and thus the dimensionality of S and M (see Theorem 3.8).

In case the maximizer on M recedes to infinity, it cannot possibly lie in ri S, and does not even furnish a reference location, which would indicate which member(s) of B ought to be dropped. As a surrogate location, any point outside ri S whose f-value is larger than $f(x^t)$ would do conceptually (Theorem 3.8(c) below). In order to keep the algorithm finite, however, one needs some submanifold on which such a point can be found by an exact, one-step calculus procedure, analogous to the one used to find x^* on M.

The submanifold $M' \subset M$ which fits the bill cuts across the path of the receding maximizer (see Fig. 1c): it is the manifold *parallel* to M^t,

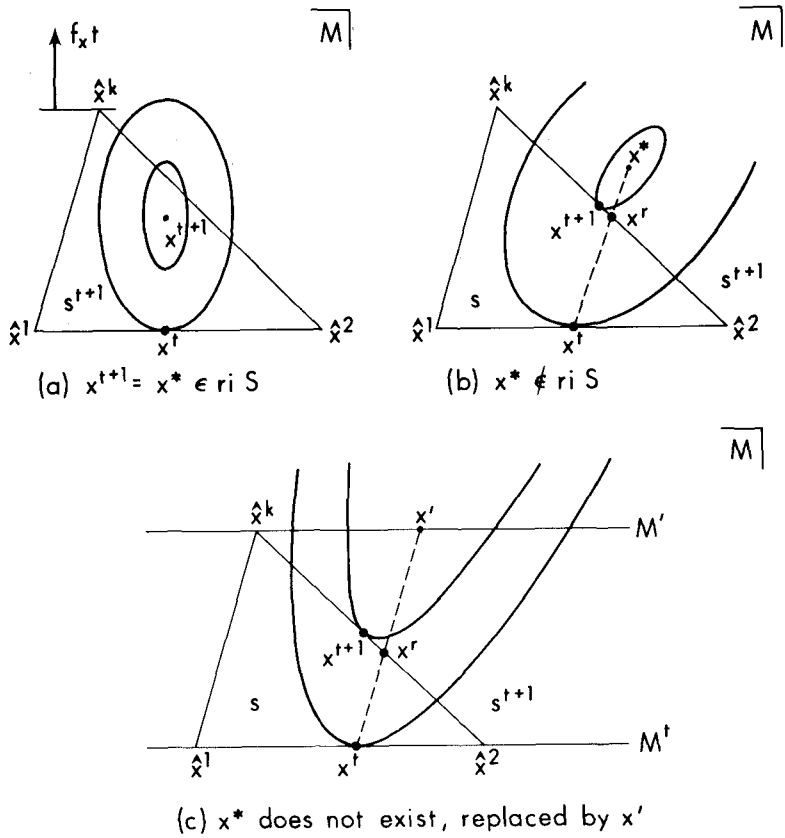

Fig. 1. The three varieties of basic step 3 of the algorithm.

that passes through $\hat{x}^k \notin M^t$, i.e., $M' = M^t + \hat{x}^k$. The proof that M' owns a *maximizer* x' that satisfies $f(x') > f(x^t)$ when no maximizer of f on M exists is given in Theorem 3.6; the proof rests mainly on the following result, which Rockafellar proves using asymptotic cones:

Lemma 3.2. *Let U be a closed convex set, and let M^t be a manifold such that $M^t \cap U$ is nonempty and bounded; then $M' \cap U$ is bounded for every manifold M' parallel to M^t* [6, *p.* 64, *Corollary* 8.4.1].

Lemma 3.3 (corollary to Lemma 3.2). *Let U be an unbounded, closed convex set with a nonempty interior, and let M^t be a supporting hyperplane to U, with $U \cap M^t$ bounded. Further, let $\hat{x}^k \notin M^t$ be a point in the same halfspace as U, say above M^t. Then $M' \cap$ int U is nonempty, where M' is the hyperplane parallel to M^t through $\hat{x}^k$.*

Proof. Assume, contrary to assertion, that $M' \cap \text{int } U$ is empty; since $\text{int } U \neq \emptyset$, this means that M' bounds U from above, while M^t bounds U from below. Further, our assumptions imply the premisses of Lemma 3.2, and therefore the intersection of U with every plane parallel to M^t is bounded. Thus it is impossible for U to recede to infinity anywhere on or between the two bounding planes M^t and M'; but this contradicts unboundedness of U, and therefore $M' \cap \text{int } U$ cannot be empty.

Lemma 3.4. *Let f be pseudoconcave, and let $f(x') > f(x^t)$; then any*

$$x^r = \lambda x' + (1 - \lambda)x^t, \qquad 0 < \lambda \leqslant 1$$

satisfies

$$f(x^r) > f(x^t).$$

Proof. By pseudoconcavity, $f(x^r) > f(x^t)$ implies $f_{x^t}(x^r - x^t) > 0$. Rearranging the expression for x^r, one gets

$$x' - x^t = \frac{1}{\lambda}(x^r - x^t), \qquad 1 \leqslant \frac{1}{\lambda} < \infty;$$

premultiplication by f_{x^t} yields

$$f_{x^t}(x' - x^t) = \frac{1}{\lambda} f_{x^t}(x^r - x^t) > 0.$$

Now assume $f_{x^t}(x' - x^t) \leqslant 0$; by pseudoconcavity of f, this implies $f(x') \leqslant f(x^t)$, which is a contradiction.

Lemma 3.5. *Let f be pseudoconcave, and let*

$$U \equiv \{x \in M: f(x) \geqslant f(x^t)\}$$

own a point x^r such that $f(x^r) > f(x^t)$. Then any $x' \in \text{int } U$ satisfies $f(x') > f(x^t)$.

Proof. Because $x' \in \text{int } U$, one can always find a point $y \in U$ such that

$$x' = \lambda x^r + (1 - \lambda)y, \qquad 0 < \lambda \leqslant 1.$$

Since $f(y) \geqslant f(x^t)$ by definition of U, $f(x') > f(x^t)$ follows by pseudoconcavity of f and by Lemma 3.4.

Theorem 3.6. *Let x^t be a member of a bounded set of maximizers of the pseudoconcave function f on the manifold M^t,[2] with affine basis $B^t \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}\}$; assume the point $\hat{x}^k$ satisfies $f_{x^t}(\hat{x}^k - x^t) > 0$, and let $B \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}, \hat{x}^k\}$ span the manifold $M \supset M^t$, on which f does not possess a maximizer.*

Then f does have a maximizer x' on $M' \equiv (M^t + \hat{x}^k) \subset M$, that satisfies $f(x') > f(x^t)$.

Proof. Let $U \equiv \{x \in M: f(x) \geqslant f(x^t)\}$ be the upper contour set of f w.r.t. x^t. By pseudoconcavity of f, U is a closed, convex set; $f_{x^t}(\hat{x}^k - x^t) > 0$ implies by Theorem 3.1, that x^t is not a maximizer of f on M, which means that int U is nonempty; nonexistence of a maximizer of f on M means, on the other hand, that U is unbounded.

Relative to the manifold M, the submanifolds M^t and M' are hyperplanes; since x^t maximizes the differentiable function f on $M^t \subset M$, M^t is a tangent plane to U at x^t; the convex set U lies entirely "above" M^t, i.e., on the same side as $\hat{x}^k$. $U \cap M^t$ is obviously nonempty and bounded by assumption. Now, by Lemma 3.3, $M' \cap U$ is bounded and by Lemma 3.4, $M' \cap$ int U is nonempty, which implies that there exists a maximizer $x' \in M'$; x' satisfies $f(x') > f(x^t)$ by Lemma 3.5.

Corollary 3.7. *Let M^r be a manifold spanned by any subbasis of B, that contains at least $\hat{x}^k$; assume f has no maximizer on M^r, and that M^r contains a point x^r on or below M' such that $f(x^r) > f(x^t)$. Then $M^r \cap M'$ owns a maximizer x' that satisfies $f(x') > f(x^t)$.*

Proof. By assumption $M^r \cap$ int U owns points above and below or on M'; therefore by convexity of U, $M' \cap M^r \cap$ int U is nonempty; because $M' \cap M^r \cap U$ is also bounded, there exists a maximizer $x' \in M' \cap M^r$ such that $f(x') > f(x^t)$.

Corollary 3.7 applies to the unlikely but possible case, that after one or several reductions of the dimensionality of M still no maximizer exists on the submanifold M^r spanned by the remaining basis B^r; we showed that then again a maximizer $x' \in M^r \cap M'$ exists, whose properties allow a further dimensional reduction.

[2] Due to the fact that the maximizer $x^t \in M^t$ is found by an exact calculus procedure (see Section 4), the x^t appearing in the algorithm is actually *unique*.

The following theorem summarizes and proves the supporting results for basic steps 4 and 5 of the algorithm. These steps use explicitly the internal or barycentric representation in terms of the basis $B \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}, \hat{x}^k\}$ to decide whether $x^* \in \text{ri } S$ or not; which members of B to drop with the help of $x^* \notin \text{ri } S$ or of $x' \notin \text{ri } S$; the proof of part (c) of Theorem 3.8 shows how this can be done with the help of a trivial linear program.

Theorem 3.8. *Be $x^t \in M^t \subset M$, $x^* \in M$, $x' \in M' \equiv M^t + \hat{x}^k \subset M$, such that $x^t \in \text{ri } S^t$ and maximizes f on M^t, $f_{x^t}(\hat{x}^k - x^t) > 0$, $f(x^*) > f(x^t)$ and $f(x') > f(x^t)$.*

Let f be pseudoconcave, and let $B^t \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}\}$ be the affine basis of M^t and S^t, $B \equiv \{\hat{x}^1, ..., \hat{x}^{k-1}, \hat{x}^k\}$ the basis of M and S. Then

(a) *$x^* = Bw^*$ has barycentric coordinates $w_k^* > 0$, w_i^* arbitrary for $i = 1, ..., k-1$,*

(b) *$x' = Bw'$ has barycentric coordinates $w_k' = 1$, $w_i' \leqslant 0$ for at least one $i = 1, ..., k-1$, i.e., $x' \notin \text{ri } S$,*

(c) *the segment between x^t and any point x^* (or x') with $w_k^* > 0$ and $w_i^* \leqslant 0$ for at least one $i = 1, ..., k-1$ contains a point x^r, with $w_i^r \geqslant 0$ for all i, $w_k^r > 0$, and at least one $w_{i \neq k}^r = 0$.*

Proof. (a) M has dimensionality $k-1$, while $M^t \subset M$ has dimensionality $k-2$; M^t therefore is a separating plane on M. By Theorem 1.1, $f_{x^t}(\hat{x}^k - x^t) > 0$ implies $\hat{x}^k \notin M^t$; as a member of the basis B, $\hat{x}^k$ has $w_k = 1$; by pseudoconcavity, $f(x^*) > f(x^t)$ implies $f_{x^t}(x^* - x^t) > 0$; thus both $x^*, \hat{x}^k \notin M^t$ but on M on the same side of M^t, which implies $w_k^* > 0$.

(b) Since $x' \in M' \equiv M^t + \hat{x}^k$, *all* points on M' have $w_k = 1$; because also $\Sigma_{i=1}^k w_i' = 1$, one must have $w_i' \leqslant 0$ for at least one $i = 1, ..., k-1$; this means $x' \notin \text{ri } S$.

(c) Since $x^t \in \text{ri } S^t \subset M^t \subset M$ has coordinates $w_k^t = 0$, $w_{i \neq k}^t > 0$, one finds the barycentric coordinates of the desired point x^r by solving

$$\begin{aligned} &\text{maximize } \lambda \\ &\text{subject to } \lambda w_i^* + (1-\lambda) w_i^t \geqslant 0, \qquad i = 1, ..., k-1 \end{aligned}$$

(the k^{th} constraint is never binding); the solution can directly be written as

$$\lambda^r = \left(\max_i \frac{w_i^t - w_i^*}{w_i^t} \right)^{-1}, \qquad i = 1, ..., k-1,$$

$$w_i^r = \lambda^r w_i^* + (1-\lambda^r) w_i^t \geqslant 0, \qquad i = 1, ..., k,$$

with $w_k^r > 0$ and the binding component(s) l having $w_l^r = 0$.

It was mentioned earlier that if $x^* \in \text{ri } S$, one has achieved the objective of stepping from a maximizer $x^t \in \text{ri } S^t \subset X$ to a maximizer $x^{t+1} \in \text{ri } S^{t+1} \subset X$ such that $f(x^{t+1}) > f(x^t)$ ($x^{t+1} = x^*$, $S^{t+1} = S$). Otherwise, one either has the point $x^* \notin \text{ri } S$ or, in case it is infinite, one can construct the point $x' \notin \text{ri } S$ with $f(x') > f(x^t)$. By Theorem 3.8, the segment $\overline{x'x^t}$ intersects the boundary of $S \subset M$ at the point x^r. By dropping from B to the noncarrying vertices j (i.e., those with $w_j^r = 0$), one changes only the barycentric reference frame of x^r, *not* its location in X. x^r lies then in $\text{ri } S^r \subset S$, $S^r \neq S$, i.e., in a subsimplex spanned by the reduced basis $B^r \subset B$.

By Lemma 3.4, x^r satisfies $f(x^r) > f(x^t)$, and therefore a new maximizer x^* on the associated submanifold M^r will satisfy $f(x^*) \geqslant f(x^r) > f(x^t)$. When x^* exists and $x^* \in \text{ri } S^r$, the iteration again has been completed successfully; otherwise, the procedure repeats itself, with $S = S^r$, $B = B^r$, $M = M^r$; a possibly necessary new $x' \in M' \cap M^r$ has the required properties by Corollary 3.7.

4. Basic step 3 of the algorithm requires the maximization of f on a manifold M (or M'). The usual external representation of a $(k-1)$-dimensional manifold $M \subset \mathbf{R}^n$ is the intersection of $n - k + 1$ hyperplanes, or algebraically the solution set of $n - k + 1$ simultaneous linear equations; those equations represent the set of locally active constraints of a program, and a calculus maximization of f on M would proceed by forming the appropriate Lagrangean and solving the associated $2n - k + 1$ first order conditions (if this is possible). The internal representation employed here sees a $(k-1)$-dimensional manifold M as the affine hull of k affinely independent points, which constitute an affine basis B. Given in this way,

$$M = \left\{ x \in \mathbf{R}^n : x = Bw, \sum_{i=1}^{k} w_i = 1 \right\},$$

and the problem $\max_{x \in M} f(x)$ could be linearly transformed into

$$\begin{aligned} &\underset{w}{\text{maximize}} && f(Bw), \\ &\text{subject to} && \sum_{i=1}^{k} w_i = 1. \end{aligned}$$

This last problem would require the solution of $k + 1$ first order conditions, a considerably smaller number than $2n - k + 1$, particularly if M is of low dimensionality.

A possible difficulty in the external approach is the linear dependence of the active constraints, which renders singular certain matrices required in the maximization procedure. In the internal approach used here it can happen, that the manifold M is a subspace (i.e., M owns the origin); the associated affine basis B (of order $n \times k$) will then have only rank $k-1$, which again causes the breakdown of the maximization by calculus. Lemma 4.1 opens an elegant route to surmount this difficulty.

Lemma 4.1. *Each nonempty manifold M is parallel to a unique subspace L; in particular, if M has the affine basis $B \equiv \{\hat{x}^1, \ldots, \hat{x}^{k-1}, \hat{x}^k\}$, then L has the linear basis $B \equiv \{\hat{x}^1 - \hat{x}^k, \ldots, \hat{x}^{k-1} - \hat{x}^k\}$, and $M = L + \hat{x}^k$ [6, Theorem 1.2, pp. 4, 7].*

Using Lemma 4.1,

$$M = L + \hat{x}^k \equiv \{x \in \mathbf{R}^n : x = Dy + \hat{x}^k,\ y \in \mathbf{R}^{k-1}\},$$

which allows the original maximization problem $\max_{x \in M} f(x)$ to be affinely transformed into

$$\max_y f(Dy + \hat{x}^k).$$

This is a constraintless problem with only $k-1$ first order conditions to be solved simultaneously; furthermore, by Lemma 4.1, D is guaranteed to be of full rank, since B is an affine basis.

Theorem 4.2. *If f is pseudoconcave and permits a closed-form solution of*

$$\frac{\partial}{\partial y} f(Dy + \hat{x}^k) = 0,$$

then the algorithm described in Section 2 is finite.

Proof. The algorithm basically generates two sequences: There is the main one of simplices in X, such that each simplex S^{t+1} possesses a relatively interior maximum $f(x^{t+1})$, that is strictly greater than $f(x^t)$, the maximum on the predecessor S^t. This implies that no simplex can recur, and thus the members of the sequence are a subset of the set of simplices in X, which is finite by Theorem 1.1.

An auxiliary sequence comes into play at the passage from S^t to S^{t+1}, if S^{t+1} does not own a relatively interior maximizer; in this case the dimension of S^{t+1} is reduced step by step until a relatively interior maxim-

izer is obtained. This happens at the latest, when S^{t+1} has shrunk to a point. That these reductions (obviously finite in number) can be done by a *finite procedure* is ensured by Theorems 3.6 and 3.8, given closed-form solvability of $\partial f(Dy + \hat{x}^k)/\partial y = 0$ (no inherently infinite sequence of line searches is necessary).

In the actual computations on the manifold M spanned by the affine basis $B \equiv \{\hat{x}^1, ..., \hat{x}^k\}$ we therefore represent points on M by

$$x = Dy + \hat{x}^k, \qquad y \in \mathbf{R}^{k-1},$$

where $D = \{\hat{x}^1 - \hat{x}^k, ..., \hat{x}^{k-1} - \hat{x}^k\}$ is the linear basis of the $(k-1)$-dimensional subspace parallel to M.

Points on the mutually parallel submanifolds $M' \subset M$, $M^t \subset M$ are given by

$$x = D^t y + \hat{x}^k, \qquad y \in \mathbf{R}^{k-2},$$

$$x = D^t y + \hat{x}^{k-1}, \qquad y \in \mathbf{R}^{k-2},$$

where

$$D^t = [\hat{x}^1 - \hat{x}^{k-1}, ..., \hat{x}^{k-2} - \hat{x}^{k-1}]$$

is the linear basis of the $(k-2)$-dimensional subspace parallel to both M' and M^t.

Remark 4.3. Notice that if the original affine basis B consists of k extreme points of X, D^t contains $k-2$ points; therefore, if B has (or was reduced to) only two members, and step 3b were taken, D^t would be empty; it is easy to show that step 2a prevents this occurrence.

5. On polytopes, the algorithm above is finite for any pseudoconcave objective function that admits the closed-form solution of $\partial f(Dy + \hat{x}^k)/\partial y = 0$.

Relatively simple examples of this class are pseudoconcave quadratics [5], which, of course, include semi-definite and definite quadratic functions. Our computational trials suggest that our method is efficient enough to be competitive.

Especially interesting is the algorithm's ability to finitely solve problems involving the linear homogeneous concave function

$$f(x) = cx + m(xVx)^{1/2}, \qquad m < 0, \quad V \text{ positive definite.}$$

It is this functional for which the method was originally conceived, because it puts the computational coping stone on a theory of decision making under risk developed by Van Moeseke [9].

Consider the stochastic linear programming model

$$\max_{x \in X} \gamma x,$$

where γ is a random vector with a distribution characterized by the vector of means c and the variance-covariance matrix V; the feasible set is

$$X = \{x \in \mathbf{R}^n : Ax \leqq b,\ x \geqq 0\}$$

with A and b deterministic.[3]

In a stochastic problem, each possible decision is associated with a multiplicity of outcomes corresponding to the various states of nature. One method of attack is to employ a decision criterion that maps the space of consequences into the reals, thus achieving a preference ranking of decisions. The best one (under the chosen criterion) is then found by maximization.

With Van Moeseke's *truncated minimax criterion* [8,9] the ranking is achieved via the risk preference functional

$$f(x) = cx + m(xVx)^{1/2}$$

which represents a dimension-preserving linear weighting of expectation and standard deviation. $(xVx)^{1/2}$ is considered as a scalar risk measure, and $m \leqslant 0$ expresses the degree of risk aversion. For a given m, a corresponding best decision is then found by maximizing f over the set X of feasible decisions.

There are other ways to view the operation of the criterion: It compares the lower ends of the distributions of γx for every decision x after truncating their tails, whose size is determined by m; an interpretation in terms of confidence limits is therefore possible [9, ch. 2]. The truncated minimax criterion converts a stochastic linear program as above into a deterministic linear homogeneous, but non-linear, problem. Thus the theory of homogeneous programming is applicable, which in the special case of portfolio selection yields some interesting and sharp results [8, 10].

The algorithm in this paper allows the computational implementation of Van Moeseke's criterion in the case where the decision set is poly-

[3] This model fits well only those interpretations where the constraints are known with near-certainty (e.g., production constraints in crop farming, the budget constraints in portfolio problems), but the objective function contains randomness (market prices for agricultural produce, security prices).

hedral. We now adapt the method for the objective functional

$$f(x) = cx + m(xVx)^{1/2}\ , \quad m < 0, \quad V \text{ positive definite};$$

its gradient is $f_x = c + (m/\sigma)Vx$, where $\sigma = (xVx)^{1/2}$. It is apparent that f_x is not defined if $\sigma = 0$, i.e., if $xVx = 0$. Assuming V to be positive definite reduces the set of $x \in \mathbf{R}^n$ with undefined gradients to $\{0\}$; the following theorem accommodates this singular case in the framework of the algorithm (which otherwise requires differentiability everywhere).

Theorem 5.1. *Let $f(x)$ be a linear homogeneous concave function to be maximized on a convex set X, and let $\hat{x} \neq 0$ be any point in X. Then, if the origin is the solution of* $\max_{x \in X} f_{\hat{x}} x$, *the origin solves* $\max_{x \in X} f(x)$.

Proof. By concavity, $f_{\hat{x}}(x - \hat{x}) \geqslant f(x) - f(\hat{x})$ for all $x \in X$; since $f_{\hat{x}}\hat{x} = f(\hat{x})$ for any $\hat{x} \neq 0$ by linear homogeneity, this simplifies to $f_{\hat{x}} x \geqslant f(x)$ for all $x \in X$; by assumption $f_{\hat{x}} 0 \geqslant f_{\hat{x}} x$ for all $x \in X$, and because $f_{\hat{x}} 0 = 0 = f(0)$ by homogeneity, $f(0) \geqslant f(x)$ for all $x \in X$ follows.

Accordingly, the algorithm is adapted as follows: The linear program in the initial step is solved using the vector c rather than $f_0 \equiv \partial f/\partial x|_{x=0}$. Whenever the origin appears as the solution of a linear program (in the initial step or in basic step 1), one terminates with 0 being the maximizer.

To maximize $cx + m(xVx)^{1/2}$ on a manifold by calculus, one applies the affine transformation described in Section 4; differentiation yields

$$\frac{\partial}{\partial y} f(Dy + \hat{x}^k) = cD + \frac{m}{\sigma}(Dy + \hat{x}^k)^{\mathrm{T}} VD = 0, \tag{1}$$

where

$$\sigma^2 = (Dy + \hat{x}^k)^{\mathrm{T}} V(Dy + \hat{x}^k). \tag{2}$$

The solution for y^t proceeds by first expressing y in terms of σ; from (1) one gets

$$y = -\left(\frac{\sigma}{m} cD + \hat{x}^k VD\right)(D^{\mathrm{T}} V D)^{-1} \tag{3}$$

which is substituted into (2). The resulting quadratic equation in σ simplifies to

$$\sigma^2 = m^2[\hat{x}^k V \hat{x}^k - \hat{x}^k VD(D^{\mathrm{T}} V D)^{-1} D^{\mathrm{T}} V \hat{x}^k] \times$$

$$\times [m^2 - cD(D^{\mathrm{T}} V D)^{-1} D^{\mathrm{T}} c]^{-1}. \tag{4}$$

Table 1

```
     ∇ HOMOGENEOUS M;B;D;E;F;G;I;J;O;P;Q;R;S;T;U;W;X;Y
[1]    B←((ρC),2+ρC)ρE←1E¯9
[2]    Y←0 SIMPLEX G←C
[3]    →6×ι∧/0=B[;S←W←,1]←X←Y
[4]    G←C+V+.×X×M÷(X+.×V+.×X←B[;S]+.×W)*÷2
[5]    →3×ι∧/0=Y←1 SIMPLEX G
[6]    ((G+.×X),X),W,[1] B[;S]
[7]    →0×ιE>|G+.×X-B[;T←ρU←W,0]←Y
[8]    J←S∘.=S←ιρO←T=ιT
[9]    →3×ι0=1+ρJ←J[;ιT]-J[;Tρ1+T←T-1]
[10]   Q←(P←D+.×⌹(⍉D)+.×V+.×D)+.×⍉D←B[;S]+.×J
[11]   →15×ιE≥R←Y+.×(V-V+.×Q+.×V)+.×Y÷(M*2)-C+.×Q+.×C
[12]   →4×ι∧/E<W←O+J+.×((C×R*÷2)-Y+.×V)+.×P
[13]   W←U+(W-U)÷⌈/(U-W)÷U+E×U=0
[14]   →8,ρB[;ιT←ρU←W[S]]←B[;S←(E<W)/S]
[15]   →9,ρ⎕←'COMPLEXITY'
     ∇
```

```
      C           A           M
1 1 1       1 1 1       ¯5
      V     1 1 1             SLK
 1 0 1            B      1 ¯1
 0 1 1     2 1
 1 1 3
```

HOMOGENEOUS M

¯8	1
2	2
0	0
0	0

¯5.07107	0.5	0.5
1	2	0
1	0	2
0	0	0

COMPLEXITY

¯3.25421	0.235921	0.764079
0.471842	2	0
0.764079	0	1
0	0	0

COMPLEXITY

¯2.53553	0.5	0.5
0.5	0	1
0.5	1	0
0	0	0

Since V is positive definite it is clear from (2) that σ^2 must be nonnegative.

Now the decision required in step 3 of the algorithm is made: If $\sigma^2 = 0$ from (4), $\sigma = 0$ and (1) becomes meaningless, i.e., no maximizer of f on M exists; route 3b is then taken. If $\sigma^2 > 0$ from (4), one substitutes $\sigma > 0$ back into (3) to get y^t, whereupon w^t and x^t are found by the appropriate transformations.

Table 1 exhibits the algorithm for the linear homogeneous case written in APL,[4] the data for an example and its solution sequence. Each iteration is represented by a matrix containing

$$\left[\begin{array}{c|c} f(x^t) & w^t \\ \hline x^t & B^t \end{array}\right] = \left[\begin{array}{c|c} \text{maximum} & \text{barycentric coordinates} \\ \hline \text{maximizer} & \text{affine basis} \end{array}\right].$$

The last matrix gives the exact solution. The example is selected such that step 3b is taken; its occurrence is signalled by the inserted word "COMPLEXITY".

References

[1] G.B. Dantzig, *Linear programming and extensions* (Princeton University Press, Princeton, N.J., 1963).

[2] G.B. Dantzig and P. Wolfe, "Decomposition principle for linear programs", *Operations Research* 8(1) (1960).

[3] J. Dréze and P. Van Moeseke, "An algorithm for homogeneous programming", CORE Discussion Paper No. 2023 (1971).

[4] O.L. Mangasarian, *Nonlinear programming* (McGraw-Hill, New York, 1969).

[5] W.C. Mylander, "Finite algorithms for solving quasiconvex quadratic programs", *Operations Research* 20 (1) (1972) 167–173.

[6] R.T. Rockafellar, *Convex analysis* (Princeton University Press, Princeton, N.J., 1970).

[7] J. Rosen, "The gradient projection method for nonlinear programming, I. Linear constraints", *Journal of the Society for Industrial and Applied Mathematics* 8 (1960) 181–217.

[8] P. Van Moeseke, "Towards a theory of efficiency", in: J. Quirk and A. Zabel, eds., *Papers in quantitative economics* (University of Kansas Press, St. Louis, 1968).

[9] P. Van Moeseke, "Stochastic linear programming: A study in resource allocation under risk", *Yale Economic Essays* 5 (1965) 196–254.

[10] P. Van Moeseke and B. von Hohenbalken, "Efficient and optimal portfolios by homogeneous programming", *Zeitschrift für Operations Research* 18 (1974) 205–214.

[4] For the sake of efficiency this latest APL-version of the algorithm omits step 2a (which ensured that the basis D^t could never become empty; see Remark 4.3). Instead the (rarely occurring) emptiness of D^t itself is used to return to $B = \{\hat{x}^k\}$; see line 9 of the APL-function.

The Simplex subroutine which we employ to solve the consecutive linear programs is adjusted to start at the extreme point of the feasible set that was optimal in the previous round. This measure reduces the total number of basis changes in the linear programs to about half.

[11] B. von Hohenbalken, "Differentiable programming on polytopes", Research Paper Series No. 14, Department of Economics, University of Alberta (1972).
[12] B. von Hohenbalken, "Simplicial decomposition in nonlinear programming algorithms", Research Paper Series No. 21, Department of Economics, University of Alberta (1974).
[13] P. Wolfe, "Methods of nonlinear programming", in: J. Abadie, ed., *Nonlinear programming* (Wiley, New York, 1967), pp. 97-131.
[14] W.I. Zangwill, *Nonlinear programming: A unified approach* (Prentice-Hall, Englewood Cliffs, N.J., 1969).