Minimizing Convex Functions Over a Simplex

Sam C. Saunders
Gordon B. Crawford

Mathematics Research

April 1965
MINIMIZING CONVEX FUNCTIONS OVER A SIMPLEX

by

Sam C. Saunders
Gordon B. Crawford

Mathematical Note No. 404
Mathematics Research Laboratory
BOEING SCIENTIFIC RESEARCH LABORATORIES
April 1965
SUMMARY

We present an iteration procedure to locate the minimum of a continuously differentiable strictly convex function over the unbounded simplex in Euclidean n-space, and we prove that the procedure converges to the unique minimum. This procedure is constructed to facilitate its adaptation to machine programming. Applications of this procedure to maximum likelihood estimation in certain non-parametric cases are mentioned.
1. INTRODUCTION

Recently, there have occurred some applications of the problem of minimizing a convex function, say $f$, defined on Euclidean $n$-space $\mathbb{R}^n$, over the unbounded simplex

$$\mathbf{S} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 \leq x_2 \leq \cdots \leq x_n\}.$$ 

Special cases of this problem arise in the maximum likelihood estimation of parameters subject to known constraints. These problems have been treated by Brunk et al in a series of papers [1], [2], [3]. They also arise in the maximum likelihood estimation in certain non-parametric situations as treated by Marshall and Proschan [5].

The method used in the cases cited requires that the function $f$ be of the form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i)$$

where each $f_i$ is convex over $\mathbb{R}_i$.

Let $u(j-r, j+s)$ be the value of $x$ which minimizes $\sum_{i=j-r+1}^{j+p-1} f_i(x)$. Then the minimizing point of $f$ over the
simplex $S$, call it $(a_1, \ldots, a_n)$, is known to be given by

$$a_j = \max_{r > 1} \min_{s > 1} u(j-r, j+s).$$

This straightforward method works for non-parametric estimates in the case of densities with increasing failure rates.

The general problem of minimizing a function over compact subsets of $R^n$ in the case the function is strictly convex in each coordinate and continuously differentiable and assumes the minimum in the interior has been treated by Warga [6]. He proposes the use of an iteration procedure of minimizing successively one coordinate after another beginning at any point.

If $f$ is convex and continuously differentiable in $R^n$ and is to be minimized over the bounded simplex

$$S_M = \{(x_1, \ldots, x_n) \in R^n : -M \leq x_1 \leq \cdots \leq x_n \leq M\}$$

for some $M > 0$, then one may utilize the Warga iteration procedure to find the minimum of $f$ over the $n$-fold Cartesian product of the interval $[-M, M]$. Let it be $(a_1^*, \ldots, a_n^*)$. Then knowing the result (proved in [3]) that if $a_j^* > a_{j+1}^*$, then the point $(a_1^*, \ldots, a_n^*)$ which minimizes $f$ over $S_n$ must have $a_j = a_{j+1}^*$. We can, in at most $n$ applications of the Warga procedure, obtain the minimum over $S_M$.

The problem which arises in the determination of the nonparametric maximum likelihood estimator in the case the density has
a convex failure rate is that of minimizing a convex function (which does not have property (1.1)) over the unbounded simplex. The belief which prompted this effort was that something could be found which was alternate to the method above for some $M$ sufficiently large.

2. THE ITERATION OPERATOR

Let $f$ be a function which is to be minimized over the convex set $S$. We assume that $f$ satisfies the following:

1° For any $\alpha, \beta \in S$ and $t \in (0,1)$
$$f(t\alpha + (1-t)\beta) < tf(\alpha) + (1-t)f(\beta),$$
and letting $\delta_{j}$ be the vector which is zero in every coordinate except the $j^{th}$ which is unity, we have
$$D_{j}f(\alpha) = \lim_{t \to 0} \frac{f(\alpha + t\delta_{j}) - f(\alpha)}{t},$$
the partial derivative with respect to the $j^{th}$ coordinate, exists and is continuous for $j = 1, \ldots, n$.

2° If $\{\alpha^{i}\}$ is a sequence of points in $S$ such that at least one coordinate, say the $j^{th}$, has the property that
$$\lim \sup_{i \to \infty} |\alpha_{j}^{i}| = \infty$$
then
$$\lim \sup_{i \to \infty} f(\alpha^{i}) = \infty.$$
S exists which by the strict convexity must be unique.

For any vector \( \alpha \in S \) and any two integers \( j, k \) such that \( j \geq 1, k \leq n \) and \( k \geq j \) we make the notational convention for \( x \) a real number

\[
(2.1) \quad (x; v, k, \alpha) = \alpha + x(\delta_j + \cdots + \delta_k)
\]

where \( \delta_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{in}) \) with \( \delta_{ij} \) the Kronecker delta.

Now, if we use \( \Delta \) as the difference operator \( \Delta \alpha_j = \alpha_j - \alpha_{j-1} \), then \( -\Delta \alpha_j \leq x \leq \Delta \alpha_{k+1} \) implies that \( (x; j, k, \alpha) \in S \). Let the value of \( x \in [-\Delta \alpha_j, \Delta \alpha_{k+1}] \) which minimizes \( f(x; j, k, \alpha) \) be denoted by \( p(j, k, \alpha) \).

**Lemma 1:** For given integers such that \( j \geq 1, k \leq n, k \geq j \) we have

\[ p(j, k, \alpha) \text{ is a continuous function of } \alpha \text{ on } S. \]

**Proof:** Let \( j, k \) as required and \( \alpha \in S \) be fixed arbitrarily. To show continuity, let \( \{\alpha^n\} \) be a sequence of points in \( S \) such that \( \alpha^n \to \alpha \in S \). If \( p_n = p(j, k, \alpha^n) \) does not converge to \( p(j, k, \alpha) \), then by compactness there is a subsequent such that

\[ p_{n_1} + p_0 \neq p(j, k, \alpha) \]

**Case I** \(-\Delta \alpha_j < \Delta \alpha_{k+1}\). Take \( x \in (-\Delta \alpha_j, \Delta \alpha_{k+1}) \). Then for that \( x \) there is an \( N \) sufficiently large that
\[ x \in \cap_{i>N} \left[ -\Delta a_i^j, \Delta a_i^{k+1} \right] \]

and then for \( i > N \), we have

\[ f(p_n^i : j, k, a^i) \leq f(x : j, k, a^i). \]

Now letting \( i \to \infty \), we have by continuity of \( f \) that

\[ f(p_0^i : j, k, a^i) \leq f(x : j, k, a). \]

This inequality holds for arbitrary \( x \in (-\Delta a_j, \Delta a_{k+1}) \) and by continuity of \( f \) must hold for all \( x \) in the closed interval. Hence, by strict convexity, it must follow that \( p_0 \) is the minimizing value and by definition \( p_0 = p(j, k, a) \) which is a contradiction.

**CASE II** \(-\Delta a_j = \Delta a_{k+1}\). Since \( \Delta a_1 \geq 0 \), we must have

\[ \Delta a_j = \Delta a_{k+1} = 0 \] and since \( a^n \to a \) it must follow that

\[ |\Delta a_j^n| + |\Delta a_{k+1}^n| \to 0 \] as \( n \to \infty \). But \(-\Delta a_j^n \leq p_n \leq \Delta a_{k+1}^n\) and hence \( p_n \to 0 = p(j, k, a) \). This completes the proof.

We now define the transformations

\[ A_{jk}(\alpha) = \alpha + p(j, k, a)(\delta_j + \cdots + \delta_k) \] for \( 1 \leq j \leq k \leq n \).

Following immediately from Lemma 1 we make the obvious

**REMARK:** For each \( j, k \) as prescribed, the transformation \( A_{jk} \) is a continuous map from \( S \) into \( S \).
in turn, we set

\[
A_1 = A_{nn} \cdots A_{jj} \cdots A_{22} A_{11}
\]

\[
\vdots
\]

\[
A_r = A_{n+1-r,n} \cdots A_{2r+1-r,1r}
\]

\[
\vdots
\]

\[
A_n = A_{1n}
\]

(2.3)

where juxtaposition indicates composition of the transformations.

Finally, we set

\[
B = A_n \cdots A_2 A_1
\]

Since the composition of continuous functions is continuous, we have

THEOREM 1: The transformation B is a continuous map of S into itself.

We now prove

LEMMA 2: If \( A_{jk}(a) \neq a \), then \( f(A_{jk}(a)) < f(a) \).

PROOF: By definition

\[
f(A_{jk}(a)) \leq f(a + x_1 \delta_1 + \cdots + x_k \delta_k)
\]

for all \( x \in [-\Delta_1, \Delta_{k+1}] \)

and by strict convexity we obtain equality iff \( x = p(j, k, a) \).

Since always \( 0 \in [-\Delta_1, \Delta_{k+1}] \) for \( a \in S \) we have \( f(A_{jk}(a)) \leq f(a) \) with equality iff \( p(j, k, a) = 0 \). Clearly then \( p(j, k, a) \neq 0 \)
Implies $f(A_{j,k}(a)) < f(a)$ and by equation 2.3 it follows that $A_{j,k}(a) \neq a$ iff $p(j,k,a) \neq 0$. This completes the proof.

3. PROOF OF CONVERGENCE

Definition: A point $a \in S$ is a fixed point of $B$ iff $B(a) = a$.

One checks easily that the property of the $A_{j,k}$'s expressed in Lemma 2 is preserved under composition. Thus, we have

THEOREM 2: $B$ is a transformation defined from $S$ into $S$ such that if $a$ is not fixed, then $f(Ba) < f(a)$.

There follows immediately

THEOREM 3: If $\mu$ is the unique minimum of $f$ over $S$, then $\mu$ is a fixed point of $B$.

PROOF: Otherwise, by Theorem 2 $B(\mu) \neq \mu$ which would contradict the fact that $\mu$ was the minimum.

We now derive some properties of this minimum which shall be needed subsequently. We define

\begin{equation}
\delta_f(a;\beta-a) = \lim_{t \to 0} \frac{1}{t}[f(a + t\beta - ta) - f(a)]
\end{equation}

and as we know we have

\begin{equation}
= \sum_{i=1}^{n} (\delta_{i,a_i}) D_i f(a).
\end{equation}

In the above, we have followed the notation of [4].
LEMMA 3: Suppose the minimum \( \mu \) of \( f \) is such that

\[
\mu_{j-1} < \mu_j = \cdots = \mu_k < \mu_{k+1}
\]

for \( j < k \)

then it follows that for \( j \leq r \leq k \)

\[
(3.4.1) \quad \delta f(u; \sum_{i=1}^{k} \delta_i) = \sum_{r} D_i f(\mu) \geq 0,
\]

\[
(3.4.2) \quad \delta f(u; \sum_{i=1}^{r} \delta_i) = \sum_{j} D_i f(\mu) \leq 0
\]

and

\[
(3.4.3) \quad \delta f(u; \sum_{i=1}^{k} \delta_i) = \sum_{j} D_i f(\mu) = 0
\]

PROOF: If \( \gamma \) is any point in \( S \), then

\[
(3.5) \quad \delta f(u; \gamma - \mu) = \sum_{i=1}^{n} (\gamma_i - \mu_i) \delta_i f(u) \geq 0,
\]

otherwise \( \mu \) would not be the minimum since there would exist a \( t > 0 \) such that

\[
\frac{1}{t} [f(\mu + t(\gamma - \mu))] - f(\mu) < 0
\]

which would be a contradiction to \( \mu \) being the minimum. Now take

\[
\gamma_i = \begin{cases} \mu_j - x & \text{if } i = j, \ldots, r \\ \mu_i & \text{otherwise.} \end{cases}
\]
For some \( x > 0 \) such that \( \gamma \in S \) it follows

\[
\sum_{i=1}^{n} (\gamma_i - \mu_i) D_i f(\mu) = -x \delta f(\mu; \sum_{j} \delta_i) \geq 0.
\]

This proves (3.4.2). The other case is done similarly. The proof of (3.4.3) follows immediately as a corollary since we must have the l.h.s. both \( \geq 0 \) and \( \leq 0 \).

We can now state the crucial

**THEOREM 4:** If \( \alpha \) is not the minimum of \( f \) over \( S \), then

is not a fixed point of \( B \).

By (3.3) we have

\[
\delta f(\alpha; u - \alpha) = \sum_{j=1}^{r} (u_j - \alpha_j) D_j f(\alpha) < 0.
\]

If we consider the intervals of indices across which \( \alpha_i \) is constant, say,

\[
l_0 < i < l_1, \ l_1 < i < l_2, \ldots, l_N < i < l_{N+1} = n + 1
\]

we may rewrite (3.6) as

\[
\delta f(\alpha; u - \alpha) = \sum_{i=1}^{N} \sum_{j=l_i}^{l_{i+1}} (i_j - \alpha_j) D_j f(\alpha) < 0.
\]

Since \( \alpha \) is a fixed point of \( B \) we must have over each interval in which \( \alpha_i \) is constant,
\[ (3.7.1) \quad \partial f(a; \delta_{j} + \cdots + \delta_{j+1}) = 0, \quad j = 0, \ldots, m \]

otherwise \( a \) would not be a fixed point of \( A_{j} \).

Moreover, for every \( j = 0, \ldots, m \), we must have for each \( k \),

\[ (3.7.2) \quad \partial f(a; \delta_{j} + \cdots + \delta_{k}) \leq 0 \]

\[ (3.7.3) \quad \partial f(a; \delta_{k} + \cdots + \delta_{j+1}) \geq 0 \]

otherwise, \( a \) would not be fixed for the appropriate \( A_{j,k} \)’s.

Since the sum in \( (3.7) \) is negative, it follows that at least one of the summands must be negative. Suppose without loss of generality that it is the first. If we also consider the subintervals of \( \ell_{1} \leq j < \ell_{2} \) across which \( \mu_{j} \) is constant, say for some \( N \geq 1 \),

\[ \ell_{1} = k_{1} \leq j < k_{2}, \quad k_{2} \leq j < k_{3}, \quad \ldots, \quad k_{m} \leq j < k_{m+1} = \ell_{2} \]  

We have from \( (3.7) \) by using \( (3.2) \) that

\[ (3.7.4) \quad \sum_{i=0}^{m} (\mu_{k_{i}} - a_{i}) \partial f(a; \delta_{k_{i}} + \cdots + \delta_{k_{i}+1}) < 0. \]

If \( \mu_{j} \) is constant across \( \ell_{1} \leq j < \ell_{2} \) (i.e. \( m = 1 \) in \( (3.7.4) \)), then we have a contradiction with \( (3.7.1) \). Thus, we
assume \( m > 1 \). Because \( \mu_{k_1} \) is an increasing function of \( i \), it must cross \( \alpha_{k_1} \) at most once. Thus, we can write (3.7.4) as

\[
0 > \left( \sum_{i=1}^{s} + \sum_{i=s+1}^{m} \right) (\mu_{k_1} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_1+1}).
\]

On the right-hand side of (3.7.5) we label the first sum \( I \) and the second \( II \) where

\[
(3.7.6) \quad \mu_{k_1} \leq \alpha_{k_1} \quad \text{for} \quad i = 1, \ldots, s
\]

\[
\mu_{k_1} > \alpha_{k_1} \quad \text{for} \quad i = s+1, \ldots, m
\]

and perhaps one of the summations is vacuous.

Consider the first term of \( I \) keeping (3.7.2) and (3.7.6) in mind. Now

\[
(\mu_{k_1} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_2}) > 0.
\]

Since \( \mu_{k_1} \) is an increasing function of \( i \), we have also

\[
0 \leq (\mu_{k_2} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_2}) - (\mu_{k_2} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_2}).
\]

By combining the terms \( i = 1, 2 \), we have

\[
I > \sum_{i=3}^{s} (\mu_{k_2} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_1+1})
\]

\[
+ (\mu_{k_2} - \alpha_{k_1}) \delta f(a; \delta_{k_1} + \ldots + \delta_{-1+k_3}).
\]
Repeating the argument \( s \) times we see that eventually we show

\[
\| A_k - A_{k+1} \| \geq 2^s \delta^{1+s} + \cdots + 2^{s+1} \delta^{1+s} + \cdots \geq 0.
\]

A similar argument holds for showing \( \| f(0) \| \geq 0 \) and thus we have a contradiction to (3.7.5).

We now have

**THEOREM 5:** For any \( \alpha \in S \) the sequence \( \{B^n(\alpha)\} \) converges to a fixed point of \( f \).

**PROOF:** Let \( a_n = B^n(\alpha) \) for \( n = 1, 2, \ldots \) then the sequence \( \{a^n\} \) has a convergent subsequence since it is from the set \( \{\beta \in S : f(\beta) \leq f(\alpha)\} \) which is closed and bounded by virtue of \( 2^s \) and being a subset of Euclidean \( n \)-space is compact.

Set \( a_n = f(a^n) \). The sequence \( \{a_n\} \) is a decreasing sequence of real numbers bounded below and therefore converges to \( a_0 \), say. Clearly all limit points of \( \{a^n\} \) have the same \( f \) value, namely \( a_0 \).

Let \( a_n \to \gamma \), say, then

\[
b_0 = f(\gamma) \leq fB(a_n) \leq f(a_n).
\]

and letting \( k \to \infty \) we have
\[ \lim_{k \to \infty} f^B(a_k^n) = f(y). \]

But by Theorem 1, \( B \) is continuous and by assumption \( f \) is also thus we have

\[ \lim_{k \to \infty} f^B(a_k^n) = f^B(y) \]

and \( f^B(y) = f(y) \) and \( y \) is a fixed point which is the unique minimum.
REFERENCES


