# AN ALGORITHM FOR NONCONVEX PROGRAMMING * 

G. GRAVES and A. WHINSTON<br>Krannert School<br>Purdue University<br>Lafayette Indiana

## 1. Introduction.

This paper presents an alogrithm to solve the most general mathematical programming problem

$$
\begin{array}{ll}
\text { s.t. } g^{\mathrm{i}}(y) \leqslant 0 & i=1,2, \ldots, m \\
\operatorname{Min} \cdot g(y) & y=\left(y_{1}, \ldots, y_{\mathrm{n}}\right)
\end{array}
$$

The only restriction required is that the functions $g^{i}, g$ be real valued. The general formulation allows for nonlinear or linear integer programming, mixed integer programming and general nonconvex continuous variable programming. The extant algorithms for this most general problem can usually be viewed as local search procedures. They suffer from two serious difficulties which can be characterized as the «dimensionality problem» and the problem of «trapping at local optima». These difficulties are illustrated by the «local corner search» where each of the $2^{n}$ adjacent

corners of a current point are evaluated and the best of these is used as the next current point. The number of function evaluations increases exponentially with the number of variables and the procedure is impossible except for problems with very few variables. As is well known, this procedure stabilizes at local optima. Traditionally, convexity is invoked by mathe-

[^0]maticians to eliminate this sort of unpleasantness. As a practical matter with real problems, convexity is never established. In fact, the essence of location of facilities problems is precisely the tradeoff between the economies of scale in production and the transportation cost. (Economies of scale imply minimization of concave functions).

In a recent paper [4], Falk and Soland have presented a method which is intended for nonconvex problems where the criterion function is separable. It employs the general ideas of branch and bound where branching is effected by repeated partitioning of the domain. Bounds are obtained for the various partitions replacing the criterion function by its convex support and optimizing the resulting problem. Falk and Soland have given a proof of convergence. However, a major drawback of their method is the fact that the subproblems formed to determine the bound are in general nonconvex optimization problems. This can be seen by noting that while the subproblem's criterion function is convex by construction, the constraints are determined by intersecting the sets of points defined by the original constraint set with the constraints defining the partition. If the original constraint set is nonconvex, which is the form they support to treat, the resulting feasible region cannot be guaranteed to be convex. Presumably, the effective scope of application of their method would be limited to convex constraint sets.

The classical approaches, then, have been essentially «local» or « neighborhood» techniques dependent on derivatives (or finite difference approximations to derivatives). Only unrealistic assumptions such as «convexity» or vague arm waving such as «try a representative sample of starting points» have been advocated to deal with the global problem. (Obtaining a <representative sample of starting points» is feasible with small generally artificial examples). We feel this sweeps the very quintessence of many economic problems under the rug. Our central aim here is to present a new framework for reaching global optimum. The procedure involves two interconnected mechanisms, a method for structuring the search and a decision rule for selecting the course of the search.

## 2. Structuring the Search.

Structuring the search consists of introducing a framework for reducing the general problem to that of «implicit enumeration» [1] suitable for machine implementation. In general, given a bounded domain $P$, it can be symmetrically partitioned into components $\mathrm{P}_{1}, \mathrm{P}_{2}, \ldots, \mathrm{P}_{2} n$. For example


Technically :

$$
\begin{array}{ll}
\text { given } & b(i) \leqslant y(i) \leqslant s(i) \\
\text { define } & r(i)=(s(i)-b(i)) / 2 \\
& \bar{y}(i)=b(i)+r(i)
\end{array}
$$

and introduce the class C of finite maps

$$
\omega:\{1, \ldots, \mathrm{n}\} \rightarrow\{0,1\}
$$

Now a $1-1$ correspondence can be setup between the components $P_{i}$ of the partition of P and the class of maps C by defining the upper and lower bounds of a component in terms of a map $\omega$

$$
\begin{aligned}
\mathrm{L}(i, \omega(i)) & =\bar{y}(i)-(1-\omega(i)) \cdot r(i) \\
\mathrm{U}(i, \omega(i)) & =\bar{y}(i)+\omega(i) \cdot r(i)
\end{aligned}
$$

To illustrate these formulas, we can apply them to the two dimensional unit square. In this event,

$$
0 \leqslant y(i) \leqslant 1 \quad i=1,2
$$

e.g.

$$
b(1)=0 \quad s(1)=1
$$

$$
b(2)=0
$$

$$
s(2)=1
$$

and

$$
\begin{array}{ll}
r(1)=1 / 2 & r(2)=1 / 2 \\
\bar{y}(1)=1 / 2 & \bar{y}(2)=1 / 2
\end{array}
$$

Using these quantities :

$$
\begin{aligned}
& \mathrm{L}(1, \omega(1))=1 / 2-(1-\omega(1)) \cdot 1 / 2 \\
& \mathrm{U}(1, \omega(1))=1 / 2+\omega(1) \cdot 1 / 2
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{L}(2, \omega(2)) & =1 / 2-(1-\omega(2)) \cdot 1 / 2 \\
\mathrm{U}(2, \omega(2)) & =1 / 2+\omega(2) \cdot 1 / 2
\end{aligned}
$$

The choice of any of the four different maps ( $\omega$ (1), $\omega$ (2)) specifies a particular rectangle.


For example, consider the map ( 0,1 ), e.g.

$$
\begin{aligned}
& \omega(1)=0 \\
& \omega(2)=1
\end{aligned}
$$

This map specifies rectangle 3

$$
\begin{array}{ll}
\mathrm{L}(1, \omega(1))=0 & \mathrm{U}(1, \omega(1))=1 / 2 \\
\mathrm{~L}(2, \omega(2))=1 / 2 & \mathrm{U}(2, \omega(2))=1
\end{array}
$$

The problem is now reduced to choosing a desirable map $\omega * \in \mathrm{C}$ and further refining the corresponding component until a point is specified to any predetermined accuracy.

Technically this can be setup recursively by taking

$$
\begin{aligned}
& r^{0}(i)=(s(i)-b(i)) / 2 \\
& \bar{y}^{0}(i)=b(i)+r^{0}(i) \\
& r^{t}(i)=r^{(t-1)}(i) / 2 \\
& \bar{y}^{t}(i)=\bar{y}^{(t-1)}-\left(1-\omega \cdot{ }^{(t-1)}(i)\right) \cdot r^{t}(i)+\omega *^{(t-1)}(i) \cdot r^{t}(i)
\end{aligned}
$$

and
and again for any may $\omega^{t}$ at the $t^{\text {th }}$ stage

$$
\begin{aligned}
& \mathrm{L}\left(i, \omega^{t}(i)\right)=\bar{y}^{\mathrm{t}}(\mathrm{i})-\left(1-\omega^{\mathrm{t}}(\mathrm{i})\right) \cdot r^{t}(i) \\
& \mathrm{U}\left(i, \omega^{t}(i)\right)=\bar{y}^{\mathrm{t}}(\mathrm{i})+\omega^{\mathrm{t}}(i) \cdot r^{\mathrm{t}}(\mathrm{i})
\end{aligned}
$$

Now specifying a sequence of maps

$$
\left(\omega{ }^{0}, \omega *^{1}, \omega *^{2}, \ldots\right)
$$

specifies a sequence of nested intervals for each $i$

$$
\left[\mathrm{L}\left(i, \omega *^{t}(i)\right) \mathrm{U}\left(i, \omega *^{t}(i)\right)\right]
$$

such that the limits of $\mathrm{L}\left(\mathrm{i}, \omega^{\top}(i)\right)$ monotonically increases with $t$ and the limite of $\mathrm{U}\left(i, \omega *^{t}(i)\right)$ monotonically decreases with $t$ and the difference $\left[\mathrm{U}\left(i, \omega *^{\mathrm{t}}(\mathrm{i})\right)-\mathrm{L}\left(i, \omega^{\mathrm{t}}(\mathrm{i})\right)\right]=(s(\mathrm{i})-b(i)) / 2^{t+1}$ approaches zero as $t$ increases.

Therefore, a sequence of maps $\left(\omega *^{0}, \omega{ }^{1}, \ldots\right)$ defines an $n$-tuple of real numbers or a point in $\mathrm{R}^{\mathrm{n}}$. (Recall the Weirstrauss-Heine development of the real numbers. Their definition is: «A real number is a nest of intervals $\left(x_{\mathrm{n}}, y_{\mathrm{n}}\right)$ such that $\left\{x_{\mathrm{n}}\right\}$ is monotonic decreasing, and $d_{\mathrm{n}}=\left(y_{\mathrm{n}}-x_{\mathrm{n}}\right) \rightarrow 0$ as $n \rightarrow \infty »$. See Knopp, [3], Chapter 1). Now for any stipulated accuracy of the solution $y^{*}(i) \pm \varepsilon$ take the first positive integer $T$ such that $\left(s^{\mathrm{T}}(i)-b^{\mathrm{T}}(i)\right) / 2^{\mathrm{T}+1} \leqslant \varepsilon$ for all $i$ or $2^{\mathrm{T}+1} \geqslant\left(s^{\mathrm{T}}(i)-b(i)\right) / \varepsilon$.

For a choice of $t / 0=\mathrm{T}$, introduce the class $\overline{\mathrm{c}}$ of meta-maps

$$
\begin{aligned}
& \dot{Q}=\left\{\omega^{1}, \omega^{2}, \ldots, \omega^{\mathrm{T}}\right\} \text { or } \\
& \Phi=\{1,2, \ldots, n \times \mathrm{T}\} \rightarrow\{0,1\}
\end{aligned}
$$

The choice of a $\Phi$ determines a «quantitized» point in the domain of interest. The problem is reduced to choosing the optimal meta-map $\Phi^{*}=\left(\omega *^{1}, \omega^{2}, \ldots, \omega *^{\mathrm{T}}\right)$. The algorithm we propose is to implicitly enumerate the class $\overline{\mathrm{c}}$ of meta-maps. There are of course many other ways of «quantitizing» the domain suitable for implicit enumeration. The employment of the present structure and, in particular, the T sub-maps $\left(\omega^{1}, \omega^{2}, \ldots, \omega^{T}\right)$ to specify $\Phi$ is to isolate for easy exploitation the nested components of the successive partitions identified by the $\omega^{t}$. It is these nested components that allow us to introduce set functionals for decision ma'ing and a global approach to calculating the optimum independent of such restrictions as convexity on the original functions.

## 3. Decision Rules for Directing the Search.

The most common set functional in mathematics is the ordinary integral. It is our contention that use of this functional instead of resorting to the derivative or its finite difference counterpart of the «local» procedures should enable us to utilize global information. Liberating ouf decision
process from the myopic local neighborhood processes should render us insensitive to trapping at local optima and enable us to dispense with inapplicable mathematical assumptions such as «convexity». The most elementary use of the integral would be to simply calculate for each component (defined by an element $\omega^{t}$ of the meta-map) the following quantities :

$$
\begin{aligned}
& \operatorname{AV}\left(\omega^{\mathrm{t}}, g\right)=\frac{1}{\prod_{\mathrm{i}=1}^{\mathrm{n}} r^{\mathrm{L}}(i)} \int_{\mathrm{L}\left(1, \omega^{\mathrm{t}}(1)\right)}^{\mathrm{U}\left(1, \omega^{\mathrm{t}}(1)\right)} \cdots \int_{\mathrm{L}\left(\mathrm{n}, \omega^{\mathrm{t}}(\mathrm{n})\right)}^{\mathrm{U}\left(\mathrm{n}, \omega^{t}(\mathrm{n})\right)} g(y) d y_{1} \ldots d y_{\mathrm{n}} \\
& \mathrm{SS}\left(\omega^{\mathrm{t}}, g\right)=\frac{1}{\prod_{\mathrm{i}=1}^{\mathrm{n}} r^{\mathrm{t}}(i)} \int_{\mathrm{L}\left(1, \omega^{\mathrm{t}}(1)\right)}^{\mathrm{U}\left(1, \omega^{\mathrm{t}}(1)\right)} \cdots \int_{\mathrm{L}\left(\mathrm{n}, \omega^{\mathrm{t}}(\mathrm{n})\right)}^{\mathrm{U}\left(\mathrm{n}, \omega^{\mathrm{t}}(\mathrm{n})\right)} g^{2}(y) d y_{1} \ldots d y_{\mathrm{n}} \\
& \\
& \operatorname{SGM}(\omega, g)=\left[\left(\mathrm{SS}\left(\omega^{\mathrm{L}}, g\right)-\operatorname{AV}^{2}\left(\omega^{\mathrm{t}}, g\right)\right]^{1 / 2}\right. \\
& d\left(\omega^{\mathrm{L}}, g\right)=\operatorname{AV}\left(\omega^{\mathrm{t}}, g\right)-v \cdot \operatorname{SGM}\left(\omega^{\mathrm{t}}, g\right)
\end{aligned}
$$

The element of the meta-map $\omega^{t}$ chosen would be such that

$$
d\left(\omega *^{\mathrm{t}}, g\right)=\min _{\omega^{\prime} \in \mathrm{C}} d\left(\omega^{\mathrm{t}}, g\right)
$$

The decision functional $d\left(\omega^{\mathrm{L}}, g\right)$ is a simple estimator of the minimum value of the function $g(y)$ on the associated component of the partition. If no knowledge of the underlying distribution is available, the parameter $v$ in the definition of $d\left(\omega^{\mathrm{t}}, g\right)$ would have to be determined empirically or several runs made using various values.

This simple procedure suffers from the same «dimensionality problem» as the local search procedures. The evaluation of the decision functional $d\left(\omega^{L}, g\right)$ for all possible $2^{n}$ maps $\omega^{\mathrm{L}}$ would impose an intolerable computational burden (except for artificial mathematical examples). This «dimensionality problem» can be eliminated, however, by resorting to an $n$-stage sequential decision process. The total map $\omega^{t}$ would be constructed in $n$-steps by sequentially fixing elements of the map. Suppose an arbitrary set of $k$ out of the possible $n$ elements of the domain are fixed. At the $(k+1)^{\text {st }}$ step an additional element of the domain, say $l_{\mathrm{k}+1}$ is chosen and

$$
l_{\mathrm{k}+1} \rightarrow 0 \quad \text { or } \quad l_{\mathrm{k}+1} \rightarrow 1
$$

Now if the order of fixing elements of the domain is completely arbitrary, there would be $2(n-k)$ possible choices of a couple $\left(l_{\mathrm{k}+1}, 0\right)$ or $\left(l_{\mathrm{k}+1}, 1\right)$ at each stage. The total number of functional evaluations would reduce to

$$
\sum_{\mathrm{k}=0}^{\mathrm{n}-1} 2(n-k)=2 \sum_{\mathrm{k}=1}^{\mathrm{n}} k=n(n+1)
$$

(This reduction is insignificant for 3 or 4 variables, but with as few as 20 variables we would achieve a reduction from
to

$$
\begin{aligned}
& 2^{20}=1,048,576 \\
& 20 \cdot 21=420 .)
\end{aligned}
$$

In the $n$-stage sequential process, it is necessary to use a slightly more sophisticated decision functional. Each choice is now determined by expected values over all completions of the $k$-partial map. Given a $k$-partial map,

$$
\left(\begin{array}{ccccc}
i_{1}, & i_{2}, \ldots, & i_{\mathrm{k}}, & i_{\mathrm{k}+1}, \ldots, & i_{\mathrm{n}} \\
\bar{\omega}^{\mathrm{t}}\left(i_{1}\right), & \bar{\omega}^{\mathrm{t}}\left(i_{2}\right), & \ldots, & \bar{\omega}^{\mathrm{t}}\left(i_{\mathrm{k}}\right), & \omega^{t}\left(i_{\mathrm{k}+1}\right), \\
& \ldots, & \omega^{\mathrm{t}}\left(i_{\mathrm{n}}\right)
\end{array}\right)
$$

where the $\left\{i_{p}\right\}$ are an arbitrary ordering of the elements of the domain and the barred elements are considered fixed, we need to consider its completion class $\mathrm{C}_{\mathrm{k}}{ }^{\mathrm{t}}$. This completion class consists of all possible completions leaving the first $k$ assignments fixed. It is the subset of maps which give the same assignment for a specific set of $k$ elements of the domain. We then employ the following expected values over the completion class $\mathrm{C}_{\mathrm{k}}{ }^{t}$

$$
\begin{array}{rlll}
\mathrm{E}_{\mathrm{C}_{\mathrm{k}}{ }^{\mathrm{t}}}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right)=\frac{1}{2^{\mathrm{n}-\mathrm{k}} \Pi r(j)} \int_{\mathrm{L}\left(1, \bar{\omega}^{t}(1)\right)}^{\mathrm{U}\left(1, \bar{\omega}^{t}(1)\right)} & \cdots & \int_{\mathrm{L}\left(\mathrm{k}, \bar{\omega}^{t}(\mathrm{k})\right)}^{\mathrm{U}\left(\mathrm{k}, \bar{\omega}^{t}(\mathbf{k})\right)} \\
\cdots & \int_{\mathrm{L}(\mathrm{k}+1,0)}^{\mathrm{U}(\mathrm{k}+1,1)} \cdots & \cdots & \int_{\mathrm{L}(\mathrm{n}, 0)}^{\mathrm{U}(\mathrm{n}, 1)} g^{2}(y) d y_{1} \ldots d y_{\mathrm{n}}{ }^{1} \\
\mathrm{E}_{\mathrm{C}_{\mathrm{k}}{ }^{\mathrm{t}}}\left(\mathrm{SS}\left(\omega^{\mathrm{t}}, g\right)\right)=\frac{1}{2^{\mathrm{n}-\mathrm{k}} \Pi r(j)} \int_{\mathrm{L}\left(1, \bar{\omega}^{t}(1)\right)}^{\mathrm{U}\left(1, \bar{\omega}^{t}(1)\right)} & \cdots & \int_{\mathrm{L}\left(\mathrm{k}, \bar{\omega}^{t}(\mathrm{k})\right)}^{\mathrm{U}\left(\mathrm{k}, \bar{\omega}^{t}(\mathrm{k})\right)} \\
& \cdots \int_{\mathrm{L}(\mathbf{k}+1,0)}^{\mathrm{U}(\mathrm{k}+1,1)} & \cdots & \int_{\mathrm{L}(\mathrm{n}, 0)}^{\mathrm{U}(\mathrm{n}, 1)} g^{2}(y) d y_{1} \ldots d y_{\mathrm{n}}{ }^{1}
\end{array}
$$

These results, of course, rely on the «additivity» of the limits of integration.
Using these more sophisticated quantities we proceed as before by calculating

$$
\operatorname{SGM}^{(\mathrm{k})}\left(\omega^{\mathrm{t}}, g\right)=\left[\mathrm{E}_{\mathrm{C}_{\mathrm{k}}}\left(\mathrm{SS}\left(\omega^{\mathrm{t}}, g\right)\right)-\mathrm{E}_{\mathrm{C}_{\mathrm{k}}}^{2}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right)\right]^{1 / 2}
$$

that is, the standard deviation of $g(y)$ on the components and

$$
d^{(\mathrm{k})}\left(\omega^{\mathrm{t}}, g\right)=\mathrm{E}_{\mathrm{C}_{\mathrm{k}}}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right)-v \cdot \mathrm{SGM}^{(\mathrm{k})}\left(\omega^{\mathrm{t}}, g\right)
$$

The decision functional $d^{(k)}\left(\omega^{\dagger}, g\right)$ is evaluated for the $2(n-k)$ possible couples, say $\left(l_{\mathrm{k}+1}\right) \rightarrow 1$ or $\left(l_{\mathrm{k}+1}, 0\right)$ and $l_{\mathrm{k}+1}$ any «free» element of the $k$-partial map. The minimum value of $d^{(\mathrm{k})}\left(\omega^{\mathrm{t}}, g\right)$ determines the next couple to be fixed.

This whole $n$-stage sequential decision process is then carried out $T$ times as indicated in Section 1 to yield a «point» in $\mathrm{R}^{\mathrm{n}}$ which is hopefully very close to the global minimum of $g(y)$. In any event, by continuing and employing a «confidence level implicit enumeration» (see [1] and [2]) of the whole class $\bar{c}$ of meta-maps, we should achieve a highly sophisticated search of the whole domain. The only point to note in employing the mechanism of the «confidence level enumeration» is that the recursive definition of the components would require $\omega^{i}$ to be entirely fixed before any element of $\omega^{\mathrm{i}+1}$.

## 4. Additional Observations.

(A) Limiting value.

When the function $g(y)$ is continuous, it might be worth noting that

$$
\begin{aligned}
& d^{(\mathrm{n})}\left(\omega^{\mathrm{t}}, g\right) \rightarrow g\left(y^{*}\right) \text { as } t \rightarrow \infty \\
& \text { where } y^{*} \in \mathrm{R}^{\mathrm{n}}
\end{aligned}
$$

is the point defined by the sequence of maps ( $\omega{ }^{0}, \omega *^{1}, \ldots$ ). This follows immediately from the Mean Value Theorem for Integrals which says:

$$
\frac{1}{\text { AREA }} \int_{\mathrm{D}} g d \mathrm{~A}=g \tilde{(y)}
$$

where $\tilde{y} \in \mathrm{D}$.
Applying this result to the terms of $d^{(n)}\left(\omega^{2}, g\right)$ yields

$$
\begin{array}{ll}
\mathrm{E}_{\mathrm{C}_{\mathrm{n}}}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right) \rightarrow g\left(y^{*}\right) & \text { as } t \rightarrow \infty \\
\mathrm{SGM}^{(\mathrm{n})}\left(\left(\omega^{\mathrm{t}}, g\right) \rightarrow 0\right. & \text { as } t \rightarrow \infty
\end{array}
$$

and hence

$$
d^{(n)}\left(\omega^{t}, g\right) \rightarrow g\left(y^{*}\right) \quad \text { for any } v \text { as } t \rightarrow \infty .
$$

(B) Indefinite Integral.

The evaluation of the integrals employed in the definition of the decision functional $d^{(k)}\left(\omega^{\mathrm{t}}, g\right)$ can be carried out in various ways. With continuous functions, the simplest procedure is to employ the closed form given by the indefinite integral; for example,

$$
\begin{gathered}
\frac{1}{\text { AREA }} \int_{\mathrm{D}} y_{1} y_{2} d y_{1} d y_{2}=\frac{\left(\mathrm{U}_{1}{ }^{2}-\mathrm{L}_{1}{ }^{2}\right) \cdot\left(\mathrm{U}_{2}{ }^{2}-\mathrm{L}_{2}{ }^{2}\right)}{4\left(\mathrm{U}_{1}-\mathrm{L}_{1}\right) \cdot\left(\mathrm{U}_{2}-\mathrm{L}_{2}\right)} \\
=\frac{\left(\mathrm{U}_{1}+\mathrm{L}_{1}\right) \cdot\left(\mathrm{U}_{2}+\mathrm{L}_{2}\right)}{4}
\end{gathered}
$$

## (C) Stratified Sampling.

When the function is not known in closed form or the indefinite integral is not available, it may become necessary to resort to stratified sampling of the various components of the domain defined by the limits of integration in the decision functional; for example,


We could determine an appropriate sample size $k_{\mathrm{i}}$ for each strata $S_{\mathrm{i}}$ and on the basis of this sample calculate estimates

$$
\hat{\mathrm{E}}_{\mathrm{C}_{\mathrm{k}}}\left(\operatorname{AV}\left(\omega^{\mathrm{t}}, g\right)\right) \quad \text { and } \quad \hat{\operatorname{SGM}}{ }^{(k)}\left(\omega^{\mathrm{t}}, g\right)
$$

and from these calculate $d^{(k)}\left(\omega^{\mathrm{t}}, g\right)$. At any decision point, we are stratifying a domain of the form :

$$
\begin{aligned}
& \mathrm{L}\left(1, \omega^{\mathrm{t}}(1)\right) \leqslant y_{1} \leqslant \mathrm{U}\left(1, \omega^{\mathrm{t}}(1)\right) \\
& \mathrm{L}\left(k, \omega^{\mathrm{L}}(k)\right) \leqslant \begin{array}{c}
y_{\mathrm{k}}
\end{array} \leqslant \mathrm{U}\left(k, \omega^{\mathrm{L}}(k)\right) \\
& \mathrm{L}(k+1,0) \leqslant y_{\mathrm{k}+1} \leqslant \mathrm{U}(k+1,1) \\
& \mathrm{L}(n, 0) \quad \leqslant \begin{array}{c}
\vdots \\
\vdots
\end{array} \leqslant \mathrm{U}(n, 1)
\end{aligned}
$$

(It is also possible to apply Analysis of Variance and other more sophisticated statistical techniques in refining the components).
(D) Discrete Variables.

It is, of course, not necessary that the variables be continuous. The Rieman-Stieltzes Integral is available to deal with discrete variables. Recall the usual Unit Step Function

$$
I(y)=\left\{\begin{array}{ll}
0 & (y \leqslant 0) \\
1 & (y>0)
\end{array}\right\}
$$

and Standard Counting Measure

$$
\alpha_{\mathrm{i}}\left(y_{\mathrm{i}}\right)=\mathrm{I}\left(y_{\mathrm{i}}\right)+\mathrm{I}\left(y_{\mathrm{i}}-1\right)
$$

that would be employed with zero-one discrete variables. (A slight generalization would eliminate the reduction to zero-one discrete variables). In this formulation,

$$
« \text { AREA }>=\int_{0}^{1+\epsilon} \ldots \int_{0}^{1+\epsilon} d \alpha_{1} \ldots d \alpha_{\mathrm{n}}=2^{\mathrm{n}}
$$

and for illustrative purposes, consider the simple linear case

$$
g(y)=\sum_{\mathrm{i}=1}^{\mathrm{n}} a_{\mathrm{i}} y_{\mathrm{i}}
$$

Take $S(i)=1+\varepsilon$ and $b(i)=0$

$$
\begin{aligned}
& \mathrm{I}_{2} \\
& P_{1} \quad \mathrm{E}_{\mathrm{C}_{\mathrm{k}}}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right)=\frac{1}{2} \int_{0}^{1+\epsilon} \int_{0}^{1 / 2}\left(a_{1} y_{1}+a_{2} y_{2}\right) d \alpha_{1} d \alpha_{2} \\
&=\frac{1}{2} \int_{0}^{1+\epsilon} a_{2} y_{2} d \alpha_{2}=a_{2} / 2 \\
& P_{2} \quad \mathrm{E}_{\mathrm{C}_{\mathrm{K}}}\left(\mathrm{AV}\left(\omega^{\mathrm{t}}, g\right)\right)=\frac{1}{2} \int_{0}^{1+\epsilon} \int_{1 / 2}^{1+\epsilon}\left(a_{1} y_{1}+a_{2} y_{2}\right) d \alpha_{1} d \alpha_{2} \\
&=\frac{1}{2} \int_{0}^{1+\epsilon}\left(a_{1}+a_{2} y_{2}\right) d \alpha_{2}=a_{1}+a_{2} / 2
\end{aligned}
$$

Hence, as expected, the decision of whether $y_{1} \rightarrow 0$ or $y_{1} \rightarrow 1$ is determined by whether $a_{1}<0$ or $a_{1}>0$. This general approach reduces to techniques expounded in great detail in the paper, «A New Approach to Discrete Mathematical Programming». It should be stressed that the Rieman-Stieltzes Integral Approach developed in this section is perfectly capable of handling
pure continuous variables, mixed continuous and integer variables, or pure integer variables.

## (E) Constraints.

The ideas developed in this paper can be extended to treat constraints of the form

$$
g_{\mathrm{i}}(y) \leqslant 0
$$

by introducing conditional expected values. The simplest way to achieve this is through the use of a Regression Equation. Instead of using AV $\left(\omega^{t}, g\right)$, this would require employment of :
$\operatorname{AV}\left(\omega^{\mathrm{t}},\left(g_{1} \mid g_{2}\right)\right)=\operatorname{AV}\left(\omega^{\mathrm{t}}, g_{2}\right)+\frac{\operatorname{COV}\left(\omega^{\mathrm{t}}, g_{1}, g_{2}\right)}{\operatorname{Var}\left(g_{1}\right)} \cdot\left(g_{1}-\operatorname{AV}\left(\omega^{\mathrm{t}}, g_{1}\right)\right.$
the conditional expected value of the function $g_{2}$ given a value of function $g_{1}$. In this procedure, it would be necessary to estimate the maximum or minimum of $\left(g_{1}-\operatorname{AV}\left(\omega^{\mathrm{t}}, g_{1}\right)\right)$ depending on the sign of the covariance on the components specified by the current $k$-partial map $\omega^{t}$. This could be done in turn in terms of the variance of $g_{1}$ and its mean. It would also be necessary to establish an appropriate confidence level that $g_{1}(y) \leqslant 0$ on the component. When the confidence drops to low it is necessary to «backtrack» in the construction of the meta-map. It should be observed that «normality» assumptions are not required for this procedure, but in the event of non-normality, the linear regression equation reduces to a firstorder approximation. Again, these ideas are developed at greater length in [1].

## 5. Examples.

The following simple examples will illustrate some of the ideas developed in this paper.

Example 1 (single variable two minima).

$$
\begin{array}{cll}
g(y)=1 / 4 y^{4}-7 / 6 y^{3}+7 / 4 y^{2}-y \\
y=1 / 2 & f(y)=-37 / 192 & \\
\text { (local minimum) } \\
y=1 & f(y)=-1 / 6 & \\
\text { (local maximum) } \\
y=2 & f(y)=-1 / 3 & \\
\text { (global minimum) }
\end{array}
$$



## Subroutine.

Sum.
$\mathrm{S}=1 / 20\left(\mathrm{U}^{5} \cdot \mathrm{~L}^{5}\right)+7 / 24\left(\mathrm{U}^{4}-\mathrm{L}^{4}\right)_{1} 7 / 12\left(\mathrm{U}^{3}-\mathrm{L}^{3}\right)-1 / 2\left(\mathrm{U}^{2}-\mathrm{L}^{2}\right)$
Sum of Squares.

$$
\begin{aligned}
\mathrm{SS} & =1 / 144\left(\mathrm{U}^{9}-\mathrm{L}^{9}\right)-7 / 96\left(\mathrm{U}^{8}-\mathrm{L}^{8}\right)+161 / 504\left(\mathrm{U}^{7}-\mathrm{L}^{7}\right) \\
& -55 / 72\left(\mathrm{U}^{6}-\mathrm{L}^{6}\right)+259 / 240\left(\mathrm{U}^{5}-\mathrm{L}^{5}\right)-7 / 8\left(\mathrm{U}^{4}-\mathrm{L}^{4}\right) \\
& +1 / 3\left(\mathrm{U}^{3}-\mathrm{L}^{3}\right)
\end{aligned}
$$

Results.

$$
\begin{gathered}
\text { Run } 1 \quad b(1)=0 \quad s(1)=3 \quad t=1.0 \\
y^{(0)}=1.5
\end{gathered}
$$

Where linear damping $v^{1}=-\left(\frac{\mathrm{M} r-\mathrm{IT}}{\mathrm{M} r}\right) v$ was employed and, $\mathrm{M} r$ is the total number of steps and IT is the current step.

Example 2 (two variables).

$$
f(y)=3\left(y_{1}-2\right)^{2}+\left(y_{2}-2.5\right)^{2}
$$

Sum.

$$
\begin{aligned}
\mathrm{S} & =\left(\mathrm{U}_{2}-\mathrm{L}_{2}\right)\left(\mathrm{U}_{1}{ }^{3}-\mathrm{L}_{1}{ }^{3}\right)-6\left(\mathrm{U}_{1}{ }^{2}-\mathrm{L}_{1}{ }^{2}\right) \\
& +\left(\mathrm{U}_{1}-\mathrm{L}_{1}\right)\left(1 / 3\left(\mathrm{U}_{2}{ }^{3}-\mathrm{L}_{2}{ }^{3}\right)-5 / 2\left(\mathrm{U}_{2}{ }^{2}-\mathrm{L}_{2}{ }^{2}\right)+18.25\left(\mathrm{U}_{2}-\mathrm{L}_{2}\right)\right.
\end{aligned}
$$

| $\begin{aligned} & \omega(1)=0 \\ & \omega(1)=1 \end{aligned}$ | $\begin{gathered} \text { AV } \\ -.1687 \\ .0187 \end{gathered}$ | SGM .0412 .4816 | $\begin{gathered} d \\ -.2065 \\ -.3859 \end{gathered}$ |
| :---: | :---: | :---: | :---: |
|  |  | $y^{(1)}=2.25$ |  |
| ${ }_{\omega}(1)=0$ | -. 3005 | . 0287 | -. 3208 |
| $\omega(1)=1$ | . 3380 | . 5090 | . 0410 |
|  |  | $y^{(2)}=1.8750$ |  |
| $\omega(1)=0$ | -. 2817 | . 0264 | -. 2938 |
| $\omega(1)=1$ | -. 3194 | . 0165 | -. 3248 |
|  |  | $y^{(3)}=2.0625$ |  |
| ! |  | ! | : |
| $\omega^{\omega}(1)=0$ | -. 3333 | . 0284 | -. 3333 |
| $\omega(1)=1$ | -. 3332 | . 0191 | -. 3332 |
|  |  | $y^{(8)}=1.998$ |  |

Runs using the straight mean were made on the following domains :

$$
\begin{array}{ll}
b(1)=-1 & s(1)=1 \\
b(2)=-5 & s(2)=5 \\
b(1)=-2 & s(1)=2 \\
b(2)=-1 & s(2)=1 \\
b(1)=-5 & s(1)=5 \\
b(2)-5 & s(2)=5
\end{array}
$$

In all cases the minimum was attained.

## REFERENCES

[1] GRAVES, G. and A. WHINSTON, «A New Approach to Discrete Mathematical Programming», Management Science, Vol. 15, No. 3, Nov. 1968.
[2] GRAVES, G. and A. WHINSTON, «An Algorithm for the Quadratic Assignment Problem», Management Science, Vol. 17, No. 7, March 1970.
[3] KNOPP, K., Theory and Application of Infinite Series, Second English Edition corresponding to the Fourth German Edition of 1947, Hafner Publishing Company, New York.
[4] FALK, J.E. and R.M. SOLAND, «An Algorithm for Separable Non-Convex Programming Problems», Management Science, Vol. 15, No. 9, May 1969.

# 8th INTERNATIONAL SYMPOSIUM ON MATHEMATICAL PROGRAMMING SPONSORED BY THE MATHEMATICAL PROGRAMMING SOCIETY 

AUGUST 26-31, 1973

STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305, USA

## MEETING ANNOUNCEMENT

The Mathematical Programming Society announces that the 8th International Symposium on Mathematical Programming will be held at Stanford University, August 26-31, 1973.

Contributed papers on theoretical, computational, and applicational aspects of mathematical programming are welcome. Abstracts should be sent before March 1, 1973 to the Chairman of the Programm Committee, Professor George B. Dantzig, Department of Operations Research, Stanford University, Stanford, California 94305, U.S.A.

Further information concerning the Symposium may be obtained from Professor Richard W. Cottle at the same address.

1) Boletin tecnico $\mathrm{N}^{\circ} 5$ Ano 1970 (Instituto Nacional de Tecnologia Industrial, Buenos Aires).
2) INTI (Instituto Nacional de Tecnologia Industrial, Buenos Aires), Ano IX N ${ }^{\circ} 22$.
3) La investigacion industrial y el INTI, Presidencia de la Nacion, Secretaria General, Republica Argentina, 1970.
4) $\mathrm{F} / \mathrm{Q}$ (Fiabilité-Qualité) Bulletin de l'AFCIQ, Paris, Vol. VIII No ${ }^{\circ}$, juin 1972.
5) Zkohomuka u Matematuyeckue Memogbl (Méthodes mathématiques et économiques) (Académie des Sciences d'U.R.S.S., Moscou), Tome VIII $\mathrm{N}^{\circ} 3,1972$.
6) Industrielle Organ:sation (Institut d'organisation industriel.'e de 1 Ecole Polytechnique Fédérale de Zurich), $\mathrm{N}^{\circ}$ 5-6-7, mai, juin, juillet 1972.
7) BIN Revue IBN (Institut Belge de Normalisation - Belgisch Instituut voor Normalisatie), $N^{\text {os }} 3$ à 6, mars à juin 1972 + Catalogue 1972.
8) Acier-Stahl- Steel (Revue Internationale des Applications de l'Acier), No 5 mai 1972.
9) Revue française d'Automatique, Informatique, Recherche Opérationnelle, $\mathrm{N}^{\circ}$ V-I, mars 1972 et R-I, avril 1972.
10) Annales de Sciences économiques appliquées (Université Catholique de Louvain), $\mathrm{N}^{\circ}$ I, mai 1972.
11) Computer Decisions (U.S.A.), Vol. 4, no 1, January 1972.
12) Paninformatic, $\mathrm{N}^{\circ}$ 12, avril, mai 1972.
13) AGIFORS proceedings, 11 year index, Volumes I - XI, 1961-1971.
14) Trabajos de Estadistica y de Investigacion Operativa (Consejo Superior de Investigaciones cientificas, Madrid), Volumen XXIII cuadernos 1 y 2.
15) Giornale degli economisti e annali di economia (Universita Commerciale Luigi Bocconi, Milano), $\mathrm{N}^{\circ}$ 1-2 Gennaio-Febbraio et $\mathrm{N}^{\circ}$ 3-4 MarsoAprile 1972.
16) Bulletin économique pour l'Europe (Nations Unies, Genève), Vol. 23, $\mathrm{N}^{\text {os }} 1$ et 2.
17) Revue de Statistique Appliquée, Vol. 20, Nos 1 et 2, 1972.
18) Operational Research Quarterly, Vol. 23, Nos 2 et 3, 1972.

[^0]:    * Research supported by Army Research Office.

