THE TECHNOLOGICAL INSTITUTE

ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS

by

A. Charnes, M. J. L. Kirby* and W. M. Raike

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS

SYSTEMS RESEARCH MEMORANDUM No. 155

The Technological Institute The College of Arts and Sciences
Northwestern University

ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS

by

A. Charnes, M. J. L. Kirby* and W. M. Raike

*
The University of Chicago

April 1966

Part of the research underlying this report was undertaken for the Office of Naval Research, Contract Nonr-1228(10), Project NR 047-021, and for the U. S. Army Research Office - Durham, Contract No. DA-31-124-ARO-D-322 at Northwestern University. Reproduction of this paper in whole or in part is permitted for any purpose of the United States Government.

SYSTEMS RESEARCH GROUP

A. Charnes, Director

1. Introduction

Recently the SIAM Journal has published two papers [1, 8] which present elaborate classifications and sophisticated constructions for two classes of problems in probabilistic programming. Neither the authors of these papers, nor their referees, appear to have recognized that the problem in [1] and the problems in [8] (for example, the "complete" problem) are very special instances of the class of "constrained generalized median" problems first discussed by Charnes, Cooper and Thompson in 1961.[1/] This canonical "median" formulation is not a matter of a choice of nomenclature. Its technical advantage in many problems in probabilistic programming derives from bringing into immediate focus the relevance of the absolute value function.

As we shall demonstrate below by means of the absolute value function we easily (1) obtain, characterize and markedly extend the essential results of [1], (2) interpret these results in linear programming and probabilistic terms, (3) develop gradients and directional derivatives and interpret probabilistically incremental formulae in all generality, (4) place these results in the framework of chance-constrained programming, and (5) reduce all cases to investigation of the behavior of the objective function along a ray.

Before undertaking these developments and in order not to interrupt them, some important assertions and constructions in [1] require correction. The problem in [1] is to choose a vector $X \geq 0$ which maximizes the expected value of the function $p^T X + \sum_{i=1}^{m} g_i(X, b)$,

---

[1/] See reference [4]

where

$$(1) \qquad g_i(X,b) = \begin{cases} (b_i - a^i X)\,\gamma_i & \text{if } b_i - a^i X \leq 0 \\ (b_i - a^i X)\,\delta_i & \text{if } b_i - a^i X \geq 0, \end{cases}$$

and $\gamma_i \geq \delta_i$.

Here $a^i, i=1\ldots m$, is the i^{th} row of a given mxn matrix A,

γ_i and $\delta_i, i=1\ldots m$, are given constants,

p^T is a given $1xn$ vector,

and $b_i, i=1,\ldots,m$, are random variables whose marginal distributions are known.

Thus the problem is

$$(2) \qquad \max \; E(p^T X + \sum_{i=1}^{m} g_i(X,b)) \quad \text{subject to } X \geq 0,$$

where we compute the expectation using the joint distribution of the random variables $b_1,\ldots,b_m$.

In (2), it is assumed that the vector X is to be selected before any observations are made on the random variables $b_i, i=1,\ldots,m$. Thus X is not permitted to be a function of these random variables, but rather it must be a deterministic vector. Such a vector of decision rules is called a zero order decision rule in the customary terminology of chance-constrained programming.[1/] Thus, (2) is a chance-constrained programming problem in which the chance constraints are of the particularly simple form $P(X \geq 0) \geq 1$, in which the function whose expectation we wish to maximize is a piecewise linear function of the vector of decision rules, and in which we seek the optimal zero order rule.

[1/] For results on zero order rules see references [2, 3, 4, 5, 6].

In the mathematical development presented in section 2, we first show that (2) can be converted into a constrained generalized median problem of the type discussed in [4]. This transformation enables us to use standard linear programming methods in obtaining our results. It also permits us to indicate how these results can be extended in two important directions: one is the case in which $g_i(X,b)$ is a piecewise linear function of $b_1, \ldots, b_m$ rather than only b_i as it is in (1); the other is the case in which X is subjected to other linear constraints, say $DX \geq d$, in addition to the constraints $X \geq 0$. Neither of these extensions are discussed in [1], and the techniques used in [1] cannot be readily extended to include these cases.

The main results in [1] are a pair of theorems which give necessary and sufficient conditions for the existence of an optimal solution to (2). We will obtain most of these conditions in a much more direct manner than was used in [1]. In particular, our proofs will require only well-known theorems in linear programming and some elementary inequalities on the absolute value, and will not require concepts such as Kakutani's fixed point theorem which were used in [1]. Neither will we require the elaborate specifications of classes of probability distributions used in [1].

One result obtained in [1] characterizes the situation in which the objective function of (2) is bounded from above, but the supremum of the function is not attained for a finite X. Rather than deriving this particular result, we characterize this situation in a constructive way by showing that it can be related to the behavior of the objective function along a particular ray and by obtaining an explicit expression for its limiting value.

First, however, we must correct some assertions in [1] about the admissible class of random variables. In [1] it is assumed that F_i, the cumulative distribution function of b_i, has at most a finite number of discontinuities in each finite interval. In our development in section 2 it will be clear that such an assumption is never required. We can, for example, admit random variables whose distribution function has probability mass at all the irrational points in an interval (or on the real line).

It is also assumed in [1] that $F_i(\cdot), i=1,\ldots m$, satisfies the following conditions:

$$3(a) \qquad \lim_{z_i \to -\infty} z_i F_i(z_i) = 0,$$

and

$$3(b) \qquad \lim_{z_i \to +\infty} z_i [1-F_i(z_i)] = 0.$$

A remark is then made [1, p. 930] to the effect that 3(a) and 3(b) are "sufficient for (but slightly stronger than) the existence of the first moment of $F_i(\cdot)$." In actual fact, as we prove in theorem 2, 3(a) and 3(b) are <u>necessary</u> for the existance of the first moment of $F_i(\cdot)$, but they are <u>not</u> sufficient. This can be seen by the following example:

Let $F(z) = \begin{cases} 1, & z \geq -e \\ \dfrac{e}{|z| \ln |z|}, & z < -e \end{cases}$

Then

$$\lim_{z \to -\infty} zF(z) = \lim_{z \to -\infty} \frac{-e}{\ln |z|} = 0$$

and $\lim_{z \to +\infty} z[1-F(z)] = 0$, as $F(z) = 1$ for $z \geq -e$

Hence $F(z)$ satisfies conditions 3(a) and 3(b). However, $F(z)$ does not have a

first moment since

$$\int_{-\infty}^{+\infty} z\,dF(z) = e \int_{-\infty}^{-e} z\left(\frac{1+\ln(-z)}{z^2[\ln(-z)]^2}\right) dz$$

$$= -e\int_{e}^{\infty} \frac{1+\ln(z)}{z[\ln(z)]^2}\,dz$$

$$= -e\int_{e}^{\infty} \frac{1}{z[\ln(z)]^2}\,dz + e\left[\ln(\ln(e)) - \lim_{t\to+\infty}\left[\ln(\ln(t))\right]\right]$$

$= -\infty$, since the integral term is negative and $\lim_{t\to+\infty}[\ln(\ln(t))] = \infty$. Thus $E(z)$ does not exist.

Conditions 3(a) and 3(b) are used repeatedly in [1] to guarantee that the integration by parts formula

$$\int_{-\infty}^{x} F_i(z)\,dz = xF_i(x) - \lim_{z\to-\infty} zF_i(z) - \int_{-\infty}^{x} z\,dF_i(z)$$

is valid. But the above example shows that conditions 3(a) and 3(b) are <u>not</u> sufficient to guarantee the validity of this integration formula since both integrals diverge. On the other hand, the existance of $E(z)$ is certainly a sufficient condition for the formula to be meaningful.

2. Mathematical Development

We begin by introducing some notation. Let Y be an n-vector. Then

$$|Y^T| = (|y_1|, \ldots, |y_i|, \ldots, |y_n|)$$

and

$$E|Y^T| = (E|y_1|, \ldots, E|y_i|, \ldots, E|y_n|).$$

We will use $||Y||$ to denote the norm of Y.

From the definition of $g_i(X,b)$ in (1) we have that

$$g_i(X, b) = \frac{\delta_i}{2}\left(|b_i - a^iX| + (b_i - a^iX)\right) + \frac{\gamma_i}{2}\left((b_i - a^iX) - |b_i - a^iX|\right).$$

Hence

$$E\left(g_i(X, b)\right) = \left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i) - \left(\frac{\delta_i + \gamma_i}{2}\right) a^iX + \left(\frac{\delta_i - \gamma_i}{2}\right) E|b_i - a^iX|$$

and

$$E\left(p^TX + \sum_{i=1}^{m} g_i(X, b)\right) = p^TX + \sum_{i=1}^{m} E\left(g_i(X, b)\right) = \sum_{i=1}^{m}\left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i)$$

$$+\left(p^T - \sum_{i=1}^{m}\left(\frac{\delta_i + \gamma_i}{2}\right) a^i\right) X + \sum_{i=1}^{m}\left(\frac{\delta_i - \gamma_i}{2}\right) E|b_i - a^iX| .$$

Thus, dropping the constant term $\sum_{i=1}^{m}\left(\frac{\delta_i + \gamma_i}{2}\right) E(b_i)$ from the objective function, (2) can be written as

(4)
$$\min \quad h(X) = c^TX + \alpha^TE|b - AX|$$

$$\text{subject to} \quad X \geq 0$$

where
$$c_j = \sum_{i=1}^{m}\left(\frac{\delta_i + \gamma_i}{2}\right) a_{ij} - p_j, \quad j = 1, \ldots, n,$$

and
$$\alpha_i = \frac{\gamma_i - \delta_i}{2}, \quad i = 1, \ldots, m.$$

Also $\alpha_i \geq 0$ as we are given that $\gamma_i \geq \delta_i$.

(4) is a constrained generalized median problem of the type discussed in [4]. The constraints are of the particularly simple form $X \geq 0$.

Theorem 1: We first note that $h(X)$ is finite for some $X \geq 0$, in which case $h(X)$ is finite for all $X \geq 0$, if and only if $E|b_i| < \infty$ for all i such that $\alpha_i > 0$ (i.e. $\gamma_i > \delta_i$).

Proof: Using the well-known and elementary inequalities on the absolute value we have that

$$|AX| - |b| \leq \Big|\,|AX| - |b|\,\Big| \leq |b - AX| \leq |b| + |AX|$$

and so applying the expected value operator (which is order- preserving) to both sides of the inequalities we get

$$(c^T X + \alpha^T |AX|) - \alpha^T E|b| \leq h(X) \leq (c^T X + \alpha^T |AX|) + \alpha^T E|b|.$$

This establishes the sufficiency of the stated conditions; necessity is clear.

Theorem 2: A necessary condition for $E|b_i| < \infty$ is that

$$\lim_{t \to \infty} tP\{|b_i| \geq t\} = 0.$$

Proof: Let $\chi_t = \begin{cases} 1, & |b_i| \leq t \\ 0, & |b_i| > t \end{cases}$

Then $E|b_i| = E(|b_i|\chi_t) + E(|b_i|(1 - \chi_t))$.

But by definition of the Lebesgue-Stieltjes (or Radon) integral,

$$E|b_i| = \lim_{t \to \infty} E(|b_i|\chi_t) \text{ when } E|b_i| < \infty .$$

Thus $E\left(|b_i|(1 - \chi_t)\right) \longrightarrow 0$ as $t \longrightarrow \infty$.

But $E\left(|b_i|(1 - \chi_t)\right) = \int_t^\infty |b_i| dF_i(b_i) \geq tP\{|b_i| \geq t\} \geq 0$ for $t \geq 0$.

So when $E|b_i| < \infty$ we have $tP\{|b_i| \geq t\} \to 0$ as $t \to \infty$.

To see that the condition

$$\lim_{t \to +\infty} tP\{|b_i| \geq t\} = 0$$

is equivalent to 3(a) and 3(b), note that

$$P\{|b_i| \geq t\} = P(b_i \geq t) + P(b_i \leq -t) = 1 - F_i(t) + F_i(-t).$$

Hence we have $0 = \lim_{t \to +\infty} \left(t[1 - F_i(t)] + tF_i(-t)\right)$,

and since $F_i(\ )$ is a nonnegative function, 3(a) and 3(b) immediately result.

Theorem 3: $h(X)$ is bounded from below for all $X \geq 0$ if and only if there exists a w satisfying

$$|w| \leq \alpha$$
$$w^T A + c^T \geq 0.$$

Proof: By virtue of the inequality

$$c^T X + \alpha^T |AX| - \alpha^T E|b| \leq h(X) \leq c^T X + \alpha^T |AX| + \alpha^T E|b|$$

obtained in the proof of theorem 1, it is enough to show that $c^T X + \alpha^T |AX|$ is bounded from below for all $X \geq 0$ if and only if there exists a w satisfying $|w| \leq \alpha$ and $w^T A + c^T \geq 0$.

To do this, we first note that the linear programming problem

$$\text{(5)} \qquad \begin{aligned} &\min \; c^T X + \alpha^T |AX| \\ &\text{subject to} \; X \geq 0 \end{aligned}$$

can be rewritten in the following form:

$$\begin{aligned} &\min \; c^T X + \alpha^T (Y^+ + Y^-) \\ &\text{subject to} \; Y^+ - Y^- - AX = 0 \\ &X, Y^+, Y^- \geq 0 . \end{aligned}$$

The dual to this problem is

$$\begin{aligned} &\max \; w^T \cdot 0 \\ &\text{subject to} \; -w^T A \leq c^T \\ &\qquad\qquad w^T \leq \alpha^T \\ &\qquad\qquad -w^T \leq \alpha^T \end{aligned}$$

or

$$\begin{aligned} &\max \; w^T \cdot 0 \\ &\text{subject to} \; w^T A + c^T \geq 0 \\ &\qquad\qquad |w| \leq \alpha . \end{aligned}$$

Since (5) is consistent (the origin is feasible), the extended dual theorem of linear programming (see [7], vol. I, p. 190) states that the objective function of (5) is finite if and only if there is a feasible solution to its dual, i.e. (6). Thus the theorem is proved.

To convert the result of theorem 3 to the corresponding result in [1], we use our definitions of c^T and α. Since $\alpha_i = \frac{\gamma_i - \delta_i}{2}$, we see that $|w_i| \leq \alpha$ means $\frac{\delta_i - \gamma_i}{2} \leq w_i \leq \frac{\gamma_i - \delta_i}{2}$. If we put $\pi_i = w_i + \left(\frac{\gamma_i + \delta_i}{2}\right)$, then the condition $|w_i| \leq \alpha$ becomes $\delta_i \leq \pi_i \leq \gamma_i$. Similarly, since $c_j = \sum_{i=1}^{m} \left(\frac{\delta_i + \gamma_i}{2}\right) a_{ij} - p_j$, we get that $w^TA + c^T = \pi^TA - p^T$, so $w^TA + c^T \geq 0$ becomes $\pi^TA \geq p^T$. Thus we see that theorem 3 is analogous to the result in [1] which says that the optimal value of the objective function to (2) is finite if and only if there exists an m-vector π satisfying

$$\pi^TA \geq p^T$$

and

$$\delta \leq \pi \leq \gamma .$$

We not turn to an elucidation of what is termed in [1] the "insoluble[1/] finite" case. This is the situation we mentioned in the introduction in which $h(X)$ is bounded from below but its infimum is not attained for a finite X, i.e. $h(X) > \inf_{X \geq 0} h(X)$ for all $X \geq 0$.

Theorem 4: As $t \to \infty$, $E|b_i - tq| \sim \begin{cases} tq - E(b_i), & q > 0 \\ E(b_i) - tq, & q < 0 \\ E|b_i|, & q = 0, \end{cases}$

where we use the symbol $\sim$ to mean that the difference between the two quantities tends to zero as $t \to \infty$.

[1/] This terminology is peculiar since we explicitly express the optimal value of the objective function for this case. This is done in the corollary to theorem 4.

Proof: Define the function $\chi_{tq} = \chi_{tq}(b_i) = \begin{cases} 1, & b_i < tq \\ 0, & b_i \geq tq \end{cases}$.

Then $E|b_i - tq| = E[(b_i - tq)(1 - \chi_{tq})] + E[(tq - b_i)\chi_{tq}]$

$= \left(E(b_i) - 2E(b_i\chi_{tq})\right) + tq[P(b_i < tq) - P(b_i \geq tq)]$.

But $E(b_i\chi_{tq}) \sim \begin{cases} 0, & q < 0 \quad \text{as} \quad t \longrightarrow \infty \\ E(b_i), & q > 0 \end{cases}$,

and the second term on the right approaches tq if $q > 0$ and $-tq$ if $q < 0$. Thus the theorem is proved.

Now fix X and define $I^0 = \{i : a^iX = 0\}$. Let $b_i^1 = -b_i \operatorname{sgn}(a^iX)$ and $b_i^2 = \begin{cases} b_i, & i \in I^0 \\ 0, & i \notin I^0 \end{cases}$. Then we have

Corollary: As $t \to \infty$, $h(tX) \sim t[c^TX + \alpha^T|AX|] + \alpha^T[E(b^1) + E|b^2|]$.

Proof: $h(tX) = tc^TX + \alpha^TE|b - tAX| \sim tc^TX + \sum_{i \in I^0} \alpha_i E|b_i|$

$+ \sum_{i \notin I^0} \alpha_i \operatorname{sgn}(a^iX)[ta^iX - E(b_i)]$

$= t[c^TX + \alpha^T|AX|] + \alpha^T[E(b^1) + E|b^2|]$.

Theorem 5: Suppose that $h(X_n) \longrightarrow v_0$ for some sequence $\{X_n\}$ with $X_n \geq 0$, $||X_n|| \longrightarrow \infty$ as $n \longrightarrow \infty$, where $v_0 = \inf_{X \geq 0} h(X)$. Then $c^T\bar{Y} + \alpha^T|A\bar{Y}| = 0$ for some $\bar{Y} \geq 0$ with $||\bar{Y}|| = 1$.

Proof: Since the set $S_1 = \{X: ||X|| = 1, X \geq 0\}$ is compact, the set $\left\{\frac{X_n}{||X_n||}\right\}$ has a limit point $\bar{Y}$ in S_1. Moreover, there exists some subsequence, which we denote by $\{Y_n\}$, such that $Y_n \longrightarrow \bar{Y}$ as $n \longrightarrow \infty$.

Now

(7) $$\frac{1}{||X_n||}\left(c^T X_n + \alpha^T|AX_n|\right) = c^T Y_n + \alpha^T|AY_n| \longrightarrow c^T\bar{Y} + \alpha^T|A\bar{Y}| \quad \text{as} \quad n\to\infty .$$

However, $c^T X_n + \alpha^T|AX_n|$ is bounded from below since $h(X_n) > v_0$. This follows from the inequality $(c^T X + \alpha^T|AX|) - \alpha^T E|b| \le h(X) \le (c^T X + \alpha^T|AX|) + \alpha^T E|b|$. Thus, in particular, $c^T X_n + \alpha^T|AX_n|$ remains bounded while $||X_n|| \longrightarrow \infty$. Hence, from (7), we conclude that the left hand side goes to zero while $c^T\bar{Y} + \alpha^T|A\bar{Y}| \ge 0$. Thus the theorem is proved.

We will now show that when $\{X_n\}$ exists with the properties described above, $h(X)$ must approach its infimum along some <u>ray</u> in the orthant $X \ge 0$. First, however, we establish

<u>Lemma 1</u>: For any vector ξ, $|h(X + \xi) - h(X)| \le |c^T\xi| + \alpha^T|A\xi|$.

<u>Proof</u>:
$$|h(X + \xi) - h(X)| \equiv \left|c^T\xi + \alpha^T E(|b - AX - A\xi| - |b - AX|)\right|$$
$$\le \left|c^T\xi + \alpha^T E|A\xi|\right| \le |c^T\xi| + \alpha^T|A\xi|.$$

<u>Theorem 6</u>: Let $||Y_n|| = 1$, and

let $t_n \longrightarrow \infty$ as $n\to\infty$.

Assume $h(t_n Y_n)\to v_0$ as $n\to\infty$ and $\bar{Y} = \lim_{n\to\infty} Y_n$.

Then $h(t_n\bar{Y})\to v_0$ as $n\to\infty$.

<u>Proof</u>: There are only a finite number of ways an n-tuple can be partitioned into subsets of components such that the components in a subset are either all positive, all zero, or all negative. Thus there exists some subsequence of $\{Y_n\}$, which with a minor abuse of notation we also denote by $\{Y_n\}$, with the property that the components of AY_n are of constant sign for all n.

Let $\delta_{in} = \begin{cases} \frac{1}{t_n^2} y_{in} & \text{if } 0 \neq a^i Y_n \to a^i \bar{Y} = 0 \\ 0 & \text{otherwise,} \end{cases}$

where y_{in} is the i^{th} element of Y_n.

Then the sign of each component of $A(\bar{Y} + \delta_n)$ remains the same for all n sufficiently large. Also

$h(t_n\bar{Y} + t_n\delta_n) \sim h(t_nY_n)$ by theorem 5

and $h(t_nY_n) \to v_0$ by hypothesis.

But, by Lemma 1, $|h(t_n\bar{Y} + t_n\delta_n) - h(t_n\bar{Y})| \leq t_n[|c^T\delta_n| + \alpha^T|A\delta_n|]$ and by our definition of δ_n the expression on the right approaches zero as $t_n \to \infty$. Thus $h(t_n\bar{Y}) \to v_0$ as $n \to \infty$.

Thus we have reduced the problem to one of investigating the behavior of $h(X)$ along rays in the nonnegative orthant. Moreover, by the corollary to theorem 4, the "knife-edge" cases in which $c^TX + \alpha^T|AX| = 0$ have $\alpha^T[E(b_1) + E|b^2|]$ as the optimal value of the objective function. However, it should be expressly noted that these "knife-edge" infinite ray cases can never appear as solutions to practical problems. Rather they exhibit an inadequacy of realistic formulation of the model. We develop their properties here only for completeness of analysis.

We now turn to the development of incremental formulae and their interpretation in probabilistic terms. It should be noted that, again, elementary properties of absolute values quickly and naturally lead to results in all generality.

3. The Increment and the Radial Directional Derivative of $h(X)$.

We now give expressions in terms of the probabilities of the various b versus ΔX events for the increment of $h(X)$ (i.e., $h(X + \xi) - h(X)$) and for the directional derivative $\frac{d}{dt} h(tX)$ whenever the latter exists. The expressions are free of any restrictions as to the character of the distributions of the b.

First, on inspecting figure 1 below, we note that, for $\epsilon > 0$,

$$f(u) = |u - (a + \epsilon)| - |u - a| = \begin{cases} \epsilon & , u < a \\ \epsilon + 2(a - u), & a \leq u < a + \epsilon \\ -\epsilon & , u \geq a + \epsilon . \end{cases}$$

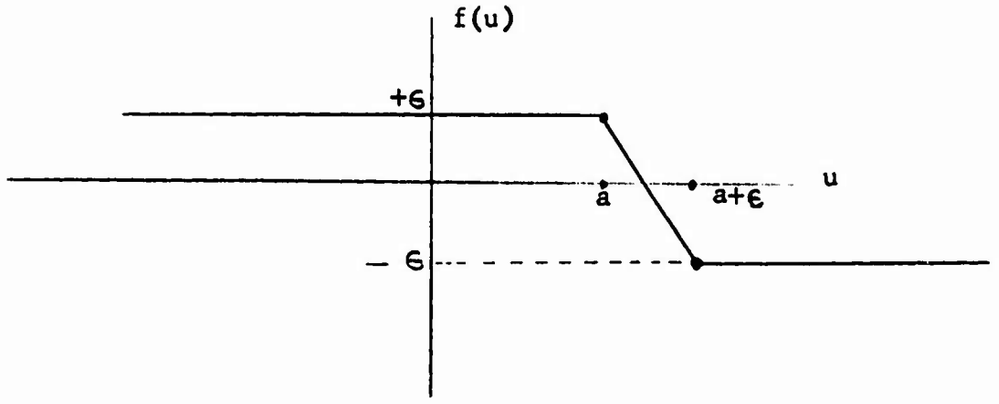

Figure 1

Thus, if u is a random variable with finite expectation,

$$E_u\Big(|u-(a+\epsilon)| - |u-a|\Big) = \epsilon[P(u < a+\epsilon) - P(u \geq a+\epsilon)] + \Big[\epsilon+2a-2\Big(a+\epsilon\theta(\epsilon)\Big)\Big]P(a \leq u < a+\epsilon)$$

for some $0 \leq \theta(\epsilon) \leq 1$ by the mean value theorem. In other words,

$$(8)\quad E_u\Big(|u-(a+\epsilon)| - |u-a|\Big) = \epsilon[P(u < a+\epsilon) - P(u \geq a+\epsilon) - 2\theta(\epsilon)P(a \leq u < a+\epsilon)].$$

But

$$h(X+\xi) - h(X) = c^T\xi + \alpha^T E\Big(|b - AX - A\xi| - |b - AX| \Big) \quad ,$$

so if we use (8) we get, using $P(b \geq AX)$ to denote the vector whose components are $P(b_i \geq a^i x)$,

$$(9)\quad h(X+\xi) - h(X) = c^T\xi + \alpha^T D(A\xi) \Big[P\Big(b < A(X+\xi) \Big) - P(b \geq A(X+\xi)) - 2D\Big(\theta(|A\xi|)\Big) P(AX \leq b < A(X+\xi)) \Big]$$

where $D(y)$ denotes the diagonal matrix whose diagonal consists of the components of the vector y and where $\theta(|A\xi|) = \Big(\theta_1(|a^1\xi|), \ldots, \theta_m(|a^m\xi|)\Big)$ with $0 \leq \theta_i(|a^i\xi|) \leq 1$. Thus we have an expression for the increment of $h(X)$, and thereby an expression for any possible directional variation, in terms of the probabilities of various events. Clearly then, any conditions on the probability distributions involved in b which guarantee that $P(AX \leq b < A(X+\xi)) \longrightarrow 0$ as $||\xi|| \longrightarrow 0$ will guarantee the existence of a gradient of $h(X)$ at X.

We now specialize (9) as follows: In (9) replace X by tX and ξ by ϵX. Then we get

$$h(t+\epsilon)X - h(tX) = c^T(\epsilon X) + \alpha^T D(A\epsilon X) \Big[P\Big(b < A(t+\epsilon)X \Big) - P\Big(b \geq A(t+\epsilon)X \Big) - 2D\Big(\theta(|A\epsilon X|)\Big) P\Big(AtX \leq b < A(t+\epsilon)X \Big) \Big] = \epsilon \Big\{ c^T X + \alpha^T D(AX) \Big[P\Big(b < (t+\epsilon)AX \Big) - P\Big(b \geq (t+\epsilon)AX \Big) - 2D\Big(\theta(|AX|)\Big) P\Big(tAX \leq b < (t+\epsilon)AX \Big) \Big] \Big\} .$$

Then as $t \longrightarrow \infty$ we get, for any fixed $\epsilon > 0$,

$$(10)\qquad \frac{h\Big((t+\epsilon)X\Big) - h(tX)}{\epsilon} \sim c^T X + \alpha^T |AX| \quad .$$

Thus, if the left hand side of (10) has a limit as $\epsilon \longrightarrow 0$, we see that the directional derivative $\frac{d}{dt} h(tX)$ is given by

$$(11)\qquad \frac{d}{dt} h(tX) = c^T X + \alpha^T |AX| \quad .$$

It is clear that computational methods can be based on our expressions (9), (10) and (11), but we shall reserve these developments for another occasion.

4. Extensions

The special model treated in [1] is rarely of any real-world or economic-theoretical significance because many constraints cannot be expressed adequately by "linear" penalty functions. As a simple reminder, note that X must be bounded, although of course the specification of the bounds may involve interrelations between groups of the individual variables. We have chosen to develop our results for clarity and ease of comparison in the context of [1]. However, it should be obvious on reflection that practically all the developments of the preceding sections can be extended to the case where there are additional linear inequality constraints on X in (2).

For example, consider the extension of theorem 3 to the case where X must also satisfy $DX \geq d$. Then our proof of theorem 3 requires only the obvious modification of the dual problem. Thus we get

Theorem 7: If the constraints $DX \geq d$, $X \geq 0$ are consistent, the optimal value of the objective function for the problem

$$\min \quad c^T X + \alpha^T E|b - AX|$$

$$\text{subject to} \quad DX \geq d$$

$$X \geq 0$$

is finite if and only if there exist vectors w and v satisfying

$$-w^T A + v^T D \leq c^T$$

$$|w| \leq \alpha$$

$$v \geq 0 \ .$$

A much more significant extension of the general model follows from observing that the linear constraints $DX \geq d$ can be considered to be the deterministic equivalent constraints for chance constraints $P(DX \leq b) \geq \beta$. This follows from

the fact that since X is a zero order rule, the chance constraints can be replaced by the equivalent set of deterministic constraints $DX \leq F^{-1}(1-\beta)$, where $F^{-1}(1-\beta)$ is the vector of $F_i^{-1}(1-\beta_i)$, the $(1-\beta_i)$ fractile point of the marginal distribution of b_i.

It is also evident that modifying the objective function to the form

$$\min \quad c^T X + E|Gb - AX|,$$

where G is a given mxm matrix, does not affect the proof of our theorems. Thus our results hold for the very general class of constrained generalized median models discussed in [4].

Finally, it should be noted that the so-called "complete" problem[1/] of linear programming under uncertainty is merely another special instance of the constrained generalized median problem, and as such it is susceptible to our analysis. To see this we merely observe that the complete problem can be written as

$$\min \quad c^T X + E\left(\sum_{i=1}^{m} g_i(X,b)\right)$$

$$\text{s.t.} \quad D X = d$$

$$X \geq 0$$

where $g_i(X,b) = \begin{cases} (b_i - a^i X)\gamma_i & \text{if } b_i - a^i X \leq 0 \\ (b_i - a^i X)\delta_i & \text{if } b_i - a^i X \geq 0. \end{cases}$

This formulation is obtained by expressing the second stage decision variables in terms of the first stage decision variables X.

[1/] See [8], page 102.

References

[1] Williams, A. C., "On Stochastic Linear Programming," Journal of the Society for Industrial and Applied Mathematics, Vol. 13, No. 4, December, 1965, pp. 927-40.

[2] Charnes, A. and W. W. Cooper, "Deterministic Equivalents for Optimizing and Satisficing Under Chance Constraints," Operations Research, Vol. 11, No. 1, January-February, 1963, pp. 18-39.

[3] Kirby, M. J. L., "Generalized Inverses and Chance-Constrained Programming," Ph.D. Thesis in Applied Mathematics, Northwestern University, Evanston, Illincis, June, 1965.

[4] Charnes, A., W. W. Cooper, and G. L. Thompson, "Constrained Generalized Medians and Hypermedians as Deterministic Equivalents for Two-Stage Linear Programs Under Uncertainty," Management Science, Vol. 12, No. 1, September, 1965, pp. 83-112. Also O.N.R. Research Memorandum No. 41, Northwestern University, Evanston, Illinois July 6, 1961.

[5] Ben Israel, A., "On Some Problems of Mathematical Programming," Ph.D. Thesis in Engineering Science, Northwestern University, Evanston, Illinois, June, 1962.

[6] Charnes, A. and M. Kirby, "Optimal Decision Rules for the n-Period E-Model of Chance-Constrained Programming," Cahiers du Centre d'Etudes de Recherche Operationnelle. January, 1966.

[7] Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming, Vols. I and II. New York: John Wiley and Sons, Inc., 1961.

[8] Wets, R. J. B., "Programming Under Uncertainty: The Equivalent Convex Program," Journal of the Society for Industrial and Applied Mathematics, Vol. 14, No. 1, Jan. 1966, pp. 89-105.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Northwestern University

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
ON A SPECIAL CLASS OF CONSTRAINED GENERALIZED MEDIAN PROBLEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Research paper

5. AUTHOR(S) (Last name, first name, initial)
Charnes, Abraham; Kirby, Michael J. L.; Raike, William M.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1966	17	8

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
DA 31-124-ARO-D-322 Joint Services N-000-14-66-CO20A03 b. PROJECT NO.	Systems Research Memorandum No. 155
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d	

10. AVAILABILITY/LIMITATION NOTICES

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Army; Joint Services Electronics Program

13. ABSTRACT

By means of elementary properties of the absolute value function, important properties of a special class of "constrained generalized median" problems (and eventually, the most general class, vide Charnes, Cooper, Thompson) such as existence of solutions, gradient and incremented formulae, linear programming and probabilistic interpretations are obtained for all classes of joint distribution functions for which the problems make sense. Results of A. C. Williams and R. Wets obtained by involved arguments and sophisticated constructs appear, when corrected, as special instances of some of the above results but devoid of the interrelations and interpretations herein adduced.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
chance-constrained programming linear programming under uncertainty constrained generalized medians zero-order decision rules						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

Errata Sheet - SRM No. 155

Page 3.	Line 17:	Insert after "specifications":	and restrictions of the
	Line 18:	Insert before "used":	as
Page 4.	Line 5:	Insert after "required":	when differentiability is not the main focus
	Line 16:	First letter should be:	e
Page 5.	Line 12:	Correct "a" in word "existence" to:	e
Page 6.	Line 15:	Remove "We first note that" from Theorem 1; then begin new paragraph above Theorem 1 with:	We first note that
	Last line:	Replace $\|AX\| - \|b\|$ by:	$\pm\|AX\| \mp \|b\|$
Page 7.	Line 3:	Replace left-hand expression by	$(c^T X \pm \alpha^T \|AX\|) \mp \alpha^T E \|b\|$
	Line 4:	Insert after "sufficiency":	and necessity
	Line 4 and 5:	Delete "necessity is clear"	
Page 8.	Line 9:	Delete "linear programming"	
	Line 12:	Insert after "following":	linear programming
Page 9.	Line 15:	Correct "not" to:	now
	1/	Insert after "since":	not only is it not insoluble but
		Insert after "we":	here
		Insert after "explicitly":	write down
Page 11.	Line 13:	Insert after "$A\xi$":	$\|$
Page 16.	Line 12:	Replace "it is susceptible to our" by:	falls under our and the previous analysis in [4]
	Line 13:	Insert quotations around:	"complete"
	1/	Insert :	"Completely slacked" might be more appropriate since 'complete' implies generality rather than the speciality actually involved.