	AD-A145 476 UNCLASSIFIED	CONSTRAINED OPTIMIZATION USING ITERATED PARTIAL 1/ KUHN-TUCKER VECTORS(U) IOWA UNIV IOWA CITY DEPT OF STATISTICS AND ACTURIAL SCIENCE R L DYKSTRA ET AL. MAY 84 TR-105 NOO14-83-K-0249	
İ.		END	
			Ē



MICROCOPY RESOLUTION TEST CHART NATIONAL BUREAU OF STANDARDS 1963-A

CONSTRAINED OPTIMIZATION

USING ITERATED PARTIAL KUHN-TUCKER VECTORS

by

Richard L. Dykstra

and

Peter C. Wollan

Department of Statistics and Actuarial Science The University of Iowa Iowa City, IA 52242

> Technical Report #105 May 1984



OTTE FILE COPY

AD-A145 476

DISTRIBUTION STATEMENT A Approved for public release; Distribution Unlimited

This research was partially supported by ONR Contract N00014-83-K-0249.

84 09 06 053

ABSTRACT

A frequently occurring problem is that of minimizing a convex function subject to a finite set of inequality constraints. Often what makes this problem difficult is the sheer number of constraints. That is, we could solve this problem for a smaller set of constraints, but solving for the total set causes difficulty. Here we discuss an approach which uses our ability to solve these partial problems to lead to a total solution. We will illustrate the method with several examples in the last section of the paper. Our \mathcal{T}_{C} approach will be somewhat heuristic in nature to promote understanding.

Key Words and Phrases: Optimization, Kuhn-Tucker vectors, inequality constraints, iteration, least squares, I-projections.



1. INTRODUCTION.

Let us consider the following convex programming problem: Minimize the convex function $f_0(x)$ (defined on some subset of R^n) over the region C subject to the constraints $f_1(x) \le 0$, $\cdots, f_m(x) \le 0$ where the f_1 , $i = 1, \cdots, m$ are finite convex functions on C. We could of course define $f_0(x)$ to be $+\infty$ for $x \notin C$, and hence assume that our functions are defined over R^n . We shall call $f_0(x)$ the objective function, and refer to $f_1(x)$, $i = 1, \cdots, m$ as the constraint functions.

2

We shall define $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ to be a vector of Kuhn-Tucker coefficients for our problem, or simply a Kuhn-Tucker vector if $\lambda_i \ge 0$, $i = 1, \dots, m$ and if

$$\inf_{\mathbf{x}\in\mathbf{C}} \mathbf{f}_0(\mathbf{x}) + \boldsymbol{\lambda}_1 \mathbf{f}_1(\mathbf{x}) + \cdots + \boldsymbol{\lambda}_m \mathbf{f}_m(\mathbf{x})$$

is finite and equal to the optimal value of our original problem. The existence of Kuhn-Tucker (KT) vectors is guaranteed under mild conditions. (See Rockafellar (1970) for a discussion of this material.) Part of the importance of KT vectors stems from the fact that they can convert a constrained problem into an unconstrained (or at least more simply constrained) problem.

Another important construct is the Lagrangian associated with our problem. It is defined as the function L on

 $\mathbf{R}^{m} \mathbf{x} \mathbf{K}^{n}$ given by

$$L(\boldsymbol{\lambda}, \mathbf{x}) = \begin{cases} f_0(\mathbf{x}) + \boldsymbol{\lambda}_1 f_1(\mathbf{x}) + \dots + \boldsymbol{\lambda}_m f_m(\mathbf{x}), & \mathbf{x} \in \mathbb{C}, \ \boldsymbol{\lambda}_1 \ge 0, \ i = 1, \dots, m \\ -\infty, & \mathbf{x} \in \mathbb{C}, \ \boldsymbol{\lambda}_1 < 0 \ (\mathbf{3}i) \\ +\infty, & \mathbf{x} \notin \mathbb{C}. \end{cases}$$

A vector pair (λ^0, x^0) is said to be a saddle-point of the in

(1.1)
$$L(\lambda, x^0) \leq L(\lambda^0, x^0) \leq L(\lambda^0, x) \quad \forall x, \lambda.$$

Such saddle-points play a big role in the following fundamental theorem (Rockafellar, p. 281).

<u>Theorem</u> 1.1. In order that λ^0 be a KT vector and x^0 be an optimal solution, it is necessary and sufficient that (λ^0, x^0) be a saddle-point of the Lagrangian L. Moreover, this condition folds if: x^0 and λ^0 satisfy

(1.2)
(a)
$$\lambda_{i}^{0} \geq 0$$
, $f_{i}(x^{0}) \leq 0$ and $\lambda_{i}^{0}f_{i}(x^{0}) = 0$, $i = 1, \dots, n$
(b) $0 \in [\partial f_{0}(x^{0}) + \lambda_{1}\partial f_{1}(x^{0}) + \dots + \lambda_{m}\partial f_{m}(x^{0})]$.

The notation $\partial f(x^0)$ indicates the set of subtradients of f at x_0 . Condition (b) implies that x^0 minimizes $f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$. Of course if the f_1 are all differentiable, (1.2)(b) may be replaced by

(1.2) (b')
$$\nabla f_0(x^0) + \lambda_1 \nabla f_1(x^0) + \dots + \lambda_m \nabla f_m(x^0) = 0.$$

Given that one can solve a constrained minimization problem, condition (1.2)(b') can often be used to find a KT vector for the problem. This fact shall prove useful later on.

2. THE METHOD.

The basic idea behind our approach is that one can reduce the number of constraints being considered by modifying the objective function through the use of estimated KT vectors. At each stage, the modified problem is solved, and updated estimates of KT vectors are found. Under fairly general conditions, the solutions to the modified problems must converge to a true global solution.

To be more specific, we assume that our constraint functions are grouped into vectors and given by

> $f_{1}(x) = (f_{11}(x), f_{12}(x), \dots, f_{1m_{1}}(x))$: $f_{k}(x) = (f_{k1}(x), f_{k2}(x), \dots, f_{km_{k}}(x))$

where $\sum_{i=1}^{k} m_i = m$.

We then define our Lagrangian as

$$L(\boldsymbol{\lambda}, \boldsymbol{x}) = \begin{cases} f_0(\boldsymbol{x}) + \boldsymbol{\lambda}'_1 f_1(\boldsymbol{x}) + \dots + \boldsymbol{\lambda}'_k f_k(\boldsymbol{x}), \ \boldsymbol{x} \in C, \ \boldsymbol{\lambda}_{ij} \ge 0, \ \forall ij \\ -\infty, \ \boldsymbol{x} \in C, \ \boldsymbol{\lambda}_{ij} < 0, \ \exists i,j \\ +\infty, \ \boldsymbol{x} \notin C, \end{cases}$$

where $\lambda = (\lambda_1, \dots, \lambda_k)$ and λ_i is an $m_i \times 1$ vector.

We shall use $L_{i}(\lambda, x)$ to denote $L(\lambda, x)$ with λ_{i} set equal to zero, and $L_{i}(\lambda_{i}, x | \lambda)$ to denote $L(\lambda, x)$ considered as a function of λ_{i} and $x (\lambda_{1}, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_{k})$ are regarded as fixed).

Initially, we set $\lambda_j^{(0,j)} = 0$, $j = 1, \dots, k$ and $\lambda^{(1,0)} = (\lambda_1^{(0,1)}, \dots, \lambda_k^{(0,k)})$. Our algorithm is sequentially defined as follows (beginning with n = 1, i = 0):

a) Let $x^{(n,i+1)} (x^{(n+1,1)} \text{ if } i = k)$ denote the solution to

 $\begin{array}{l} \text{Minimize } \mathbb{L}_{i+1}(\lambda^{(n,i)}, x) \\ x: f_{i+1,j}(x) \leq 0 \ \forall j \\ (x: f_{1,j}(x) \leq 0 \ \forall j \ \text{if } i = k) \end{array}$

This is a convex programming problem which we assume we can solve. We let $\lambda_{i+1}^{(n,i+1)}$ ($\lambda_1^{(n+1,1)}$ if i = k) denote a KT vector for the problem. (Condition (1.2) may prove useful in obtaining this KT vector.)

b) We now update our global KT estimate by setting

$$\boldsymbol{\lambda}^{(n,i+1)} = (\boldsymbol{\lambda}_{1}^{(n,1)}, \cdots, \boldsymbol{\lambda}_{i+1}^{(n,i+1)}, \boldsymbol{\lambda}_{i+2}^{(n-1,i+2)}, \cdots, \boldsymbol{\lambda}_{k}^{(n-1,k)}).$$

(If i = k, we set

$$\boldsymbol{\lambda}^{(n+1,1)} = (\boldsymbol{\lambda}_1^{(n+1,1)}, \boldsymbol{\lambda}_2^{(n,2)}, \cdots, \boldsymbol{\lambda}_k^{(n,k)}.)$$

We then replace (n,i) by (n,i+1) ((n+1,1) if i = k) and return to step a).

In many situations, the problems in step a) which must be solved are always of the same form, and hence lead to easily written computer programs for performing these steps. We shall give some explicit examples in Section 4.

3. JUSTIFICATION FOR THE ALGORITHM.

The crucial fact that justifies this procedure is that the Lagrangian can only increase at each step in the algorithm. This follows since (for i < k)

$$L(\lambda^{(n,i)}, x^{(n,i)}) = L_{i}(\lambda_{i}^{(n,i)}, x^{(n,i)} | \lambda^{(n,i)})$$

$$\leq L_{i}(\lambda_{i}^{(n,i)}, x^{(n,i+1)} | \lambda^{(n,i)}) \text{ (by Thm. 1.1)}$$

$$= L_{i+1}(\lambda_{i+1}^{(n,i)}, x^{(n,i+1)} | \lambda^{(n,i+1)})$$

$$\leq L_{i+1}(\lambda_{i+1}^{(n,i+1)}, x^{(n,i+1)} | \lambda^{(n,i+1)}) \text{ (by Thm. 1.1)}$$

$$= L(\lambda^{(n,i+1)}, x^{(n,i+1)}).$$

A similar argument holds if i = k for showing

$$L(\lambda^{(n,k)}, x^{(n,k)}) \leq L(\lambda^{(n+1,1)}, x^{(n+1,1)}).$$

Moreover, we note that if $y \in C$ is any vector such that $f_{ij}(y) \le 0$, $\forall i, j$, then

(3.1)
$$L(\lambda^{(n,1)}, x^{(n,1)}) \leq L(\lambda^{(n,1)}, y) \leq f_0(y),$$

so that $\lim L(\lambda^{(n,1)}, x^{(n,1)})$ exists finite independently of i. Now, since $x^{(n,1)}$ minimizes a convex function, under conditions which guarantee sufficient curvature of $L(x,\lambda)$ (such as $xHx \ge \gamma ||x||^2$ for some $\gamma > 0$ where H is the Hessian of f_0), we know that $x^{(n,1+1)}$ must be close to $x^{(n,1)}$ for sufficiently large n. Now, if $x^{(n,1)}$ must contain a convergent subsequence (such as if C is a bounded region) converging to $x^0 \in C$, then we only need continuity properties of f_0 and f_1 , $i=1,\cdots,m$ to guarantee that $f_1(x^0) \le 0$, $i=1,\cdots,m$ and $f_0(x^0) \le f_0(y)$ for all y in C which satisfy all constraints (by 3.1). Thus x^0 must be a solution to our problem, and since every convergent subsequence converges to x^0 , the algorithm must work correctly.

4. APPLICATIONS.

 Let us first consider least squares problems under linear inequality constraints. Thus, we wish to minimize the objective function

$$f_0(x) = \frac{1}{2} \sum_{1}^{n} (g_i - x_i)^2 w_i$$

subject to the constraints

$$f_{i}(x) = x'a_{i} = \sum_{j=1}^{n} a_{ij}x_{j} \le 0, i = 1, \cdots, m,$$

where w > 0, g,a₁,...,a_m are given $n \ge 1$ vectors such that there exists at least one vector x where $a'_i x \le 0 \forall i$.

Of course, the solution to our problem under a single constraint $a'_i x \le 0$ can be easily found by the expression

(4.1)
$$P_{i}(g) = \begin{cases} g , & \text{if } \sum_{j=1}^{n} a_{ij}g_{j} \leq 0 \\ (g'_{i}, \cdots, g'_{n}), & \text{if } \sum_{j=1}^{n} a_{ij}g_{j} > 0, \end{cases}$$

where

$$g'_{j} = g_{j} - (\sum_{l=1}^{n} g_{l}a_{il})a_{ij}w_{j}^{-1} / \sum_{l=1}^{n} a_{il}^{2}w_{l}^{-1}.$$

The corresponding KT value can then be found from (1.2)(b') as

(4.2)
$$\lambda_{i} = (g_{j} - P_{i}(g)_{j})w_{j}a_{ij}^{-1}$$

where j is any index such that $a_{ij} \neq 0$. Suppose now that we modify our objective function (with g replaced by an arbitrary h) by adding $\lambda_i f_i(x)$. However, the problem

(4.3) Minimize
$$\frac{1}{2} \sum_{j=1}^{n} (h_j - x_j)^2 w_j + \lambda_i f_i(x)$$

 $x' a_n \le 0$

has the same solution as

$$\begin{array}{l} \text{Minimize} & -\sum\limits_{j=1}^{n} x_{j} w_{j} [h_{j} + (P_{i}(g)_{j} - g_{j})] + \frac{1}{2} \sum\limits_{j=1}^{n} x_{i}^{2} w_{i} \\ & x' a_{p} \leq 0 \end{array}$$

which is equivalent to

(4.4) Minimize
$$\frac{1}{2} \sum_{j=1}^{n} (h_j + (P_i(g)_j - g_j) - x_j)^2 w_j$$
.

Thus our adjustment still leaves the problem as the same type of least squares problem, but with the value h modified to now be $h + (P_i(g)-g)$.

We can now apply the method proposed in section 2, which can be expressed as follows:

1) Set $g_{11} = P_1(g)$, and $I_{11} = g_{11}-g$. 2) Set $g_{12} = P_2(g_{11})$, and $I_{12} = g_{12}-g_{11}$. 3) Continue, until $g_{1m} = P_m(g_{1,m-1})$, and $I_{1m} = g_{1m}-g_{1,m-1}$.

- 4) Now set $g_{21} = P_1(g_{1m}-I_{11})$, and $I_{21} = g_{21}-(g_{1m}-I_{11})$.
- 5) Contine. In general, set $g_{nj} = F_j(g_{n,j-1}-I_{n-1,j})$, and $I_{nj} = g_{nj}-(g_{n,j-1}-I_{n-1,j})$ if j > 1, and $g_{n1} = P_1(g_{n-1,m}-I_{n-1,1})$ and $I_{n1} = g_{n1}-(g_{n-1,m}-I_{n-1,1})$.

This scheme is easy to program since the projections are of such simple form, and there is no branching or searching involved. Moreover, $g_{n,j}$ is guaranteed to converge to the true solution. This is a special case of an algorithm given by Dykstra (1983).

2. Depending upon the nature of the linear constraints, it may be possible to easily find the projection under several simultaneous constraints. For example, if

$$f_{i}(x) = x_{i} - x_{i+1}, \quad i = 1, \dots, n-1,$$

the set of vectors which satisfy all n-l constraints are just the nondecreasing vectors. Projections onto these types of regions are quite tractable (see Barlow, Bartholomew, Bremner and Brunk (1972)). Another set of linear inequality constraints which can be simultaneously handled are

$$f_{i}(x) = i^{-1} \sum_{j=1}^{i} x_{j} - x_{i+1}, \quad i = 1, \dots, n-1$$

(Shaked (1979) and Dykstra and Robertson (1983)).

By being able to handle large constraint sets, we can improve the efficiency of our method.

To elaborate, we consider the problem

(4.5) Minimize
$$\frac{1}{2}\sum_{j=1}^{n} (g_{j} - x_{j})^{2} w_{j}$$

x'A₁ <0, i=1, ..., k 1

where $A_i = (a_{i1}, \dots, a_{im_i})$ is an $n \times m_i$ matrix of independent columns. We assume that we can solve (4.5) for any particular i and any g, and will denote the solution by $F_i(g)$. It can be shown that the KT vector for this problem is

$$\lambda_{i}^{m_{i}\times l} = (A_{i}^{\prime}A_{i})^{-l}A_{i}^{\prime}(g \cdot w - P_{i}(g) \cdot w)$$

where x y denotes coordinatewise multiplication.

Interestingly enough, we can repeat the argument used in (4.3) and (4.4) to derive the same type of result. Thus it follows that our earlier stated algorithm is still valid if $P_i(g)$ denotes the projection of g onto $\{x; x'A_i \leq 0\}$.

This extension may prove very useful for situations where there are a great many constraints. For example, Dykstra and Robertson (1982) were able to find least squares projections of rectangular arrays under the constraints of nondecreasing rows and nondecreasing columns for even large arrays. 3. As a final example, we consider the problem of finding I-projections onto an intersection of linear inequality regions.

We will not elaborate on the importance of I-projections, only state that they occur in a myriad of places in many different settings. We refer the reader to Kullback (1959) or Csiszar (1975) for elaboration on their importance. The problem we are dealing with is

(4.6) Minimize
$$\sum_{i=1}^{n} p_i \ln(p_i/r_i)$$

 $a'_i p \le c_i, i=1, \cdots, m \quad i=1$
 $\sum_{i=1}^{n} p_i = 1, p_i \ge 0$

where $r\neq 0$ is a finite, nonnegative, vector and the set of feasible points is not empty. We will define our objective function as

$$f_{0}(p) = \begin{cases} \sum_{i=1}^{n} p_{i} \ln(p_{i}/r_{i}), p_{i} \ge 0, \sum_{i=1}^{n} p_{i} = 1\\ \infty, \text{ elsewhere.} \end{cases}$$

Our constraint functions are

$$f_{i}(p) = (a_{i}-c_{i})'p = \sum_{j=1}^{n} (a_{ij}-c_{j})p_{j}, i = 1, \cdots, m,$$

where c_i is a vector of constants. We note that r can be arbitrarily scaled without changing the problem.

The solution to the problem of minimizing $f_{\Omega}(p)$ subject

to $f_i(p) \leq 0$ is given by

$$\hat{p}_{ij} = r_j e^{-\lambda_i (a_{ij} - c_i)} / \sum_{j=1}^{n} r_j e^{-\lambda_i (a_{ij} - c_i)}$$

where λ_i is the solution to

$$\sum_{j=1}^{n} (a_{ij} - c_i) r_j e^{-\lambda(a_{ij} - c_i)} = 0$$

providing it is nonnegative, and zero otherwise. It also turns out that λ_1 is the KT value associated with the problem.

Now if we modify our objective function (with r replaced by an arbitrary nonnegative, nonzero h) by adding $\lambda_{j}f_{j}(x)$, our problem becomes

$$\begin{array}{ll} \text{Minimize} & f_0(p) + \lambda_i & \sum_{j=1}^n (a_{ij} - c_i) p_j, \\ (a_p - c_p)' p \le 0 & j = 1 \end{array}$$

or

$$\begin{array}{ccc} \underset{(a_{r}-c_{r})' p \leq 0}{\text{Minimize}} & \underset{j=1}{\overset{n}{\Sigma}} p_{j} [\ln p_{j} / h_{j} - \ln e^{-\lambda_{i} (a_{ij} - c_{i})}] \\ p_{j} \geq 0 \quad \forall j, \underset{1}{\overset{n}{\Sigma}} p_{j} = 1 \end{array}$$

or equivalently,

(4.7) Minimize
$$\sum_{\substack{j=1\\ j \geq 0 \\ j \neq j}}^{n} p_j \ln p_j / h_j (\hat{p}_{ij} / r_j).$$

 $p_j \geq 0 \\ \forall j, \sum_{j=1}^{n} p_j = 1$

The key point is that our problem is precisely of the

same form as before, except that we have modified our vector h. Thus we may use our procedure of modifying our objective function using updated estimates of the KT vector, and only have to solve the one type of problem. Setting $\lambda_i = 0$ in (4.7) is equivalent to setting $\hat{p}_i/r = 1$. This scheme is quite effective for finding I-projections under multiple linear inequality constraints.

In summary, this procedure seems to work quite well for situations where many constraints are involved, and partial solutions (solutions under partial constraints) are easily available.

REFERENCES

- Barlow, R.E., Bartholomew, D.J., Bremner, J.M., and Brunk, H.D. (1972). <u>Statistical Inference Under Order Restrictions</u>. John Wiley, New York.
- Csiszar, I. (1975). "I-Divergence Geometry for Probability Distributions and Minimization Problems," <u>Annals of Prob-</u><u>ability</u>, 3, 146-159.
- Dykstra, Richard L. (1983). "An Algorithm for Restricted Least Squares Regression," Journal of the American Statistical Association, 78, 837-842.
- Dykstra, Richard L. and Robertson, Tim (1982). "An Algorithm for Isotonic Regression for Two or More Independent Variables," <u>Annals of Statistics</u>, 10, 708-716.
- Dykstra, Richard L. and Robertson, Tim (1983). "On Testing Monotone Tendencies," Journal of the American Statistical Association, 78, 342-350.
- Kullback, S. (1959). <u>Information Theory and Statistics</u>. John Wiley, New York.
- Rockafellar, R. Tyrrell (1970). <u>Convex Analysis</u>. Princeton University Press, Princeton, NJ.
- Shaked, Moshe (1979). "Estimation of Starshaped Sequences of Poisson and Normal Means," <u>Annals of Statistics</u>, 7, 729-741.

REPORT DOCUMENTATION PAGE			READ INSTRUCTIONS BEFORE COMPLETING FORM	
AEPOAT NUMBER		2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
10)				
TITLE (and Substille)		S. TYPE OF REPORT & PERIOD COVERED		
Constrained O	ptimization Using It			
Kuhn-Tucker V	ectors.		Technical - August 1904	
	•		105	
AUTHOR(4)			B. CONTRACT OR GRANT NUMBER(+)	
Richard L. Dy Peter C. Woll	kstra an		N00014-83-K-0249	
PERFORMING ORG	ANIZATION NAME AND ADDRES	sminl Science	10. PROGRAM ELEMENT, PROJECT, TASK	
Department of	Statistics and Actu	arial belence		
The Universit	y of towa			
10wa CICy, 10	wa)2242			
CONTROLLING OF	FICE NAME AND ADDRESS		12. ANOTE 1084	
Office of Nav	al Research			
800 N. Quincy			15	
AT LINGTON VA	NCY NAME & ADDRESSII dillon	nt Inun Controlling Office)	18. SECURITY CLASS. (of this report)	
			Unclassified	
			134. DECLASSIFICATION/DOWNGRADING SCHEDULE	
DISTRIBUTION S	TATEMENT (al the obstract enter	ed in Block 20, 11 different f	MITED.	
7. DISTRIBUTION S	RATEMENT (of the obsised union	od in Block 20, if different f	MITED.	
. SUPPLEMENTAR	TATEMENT (of the obstract enter	od in Block 20, it dilloroni t	MITED.	
DISTRIBUTION ST	TATEMENT (of the obstract onion	od in Block 20, if different f	MITED.	
7. DISTRIBUTION ST	TATEMENT (of the obsiract union	od in Block 20, if different f	MITED.	
7. DISTRIBUTION ST	TATEMENT (of the obstract onion IV NOTES	ad in Block 20, it different f	MITED.	
DISTRIBUTION ST S. SUPPLEMENTAR D. KEY WORDS (Con Optimization squares, I-	TATEMENT (of the obstract onter by NOTES winus on reverse side if necessary , Kuhn-Tucker vectory projections.	od in Block 20, if different f and identify by block numb s, inequality con	MITED.	
DISTRIBUTION ST S. SUPPLEMENTAR Optimization squares, I-	NATEMENT (of the obstract enter NATEMENT (of the obstract enter Name on reverse side if necessary , Kuhn-Tucker vector projections.	od in Block 20, 11 dilloreni 1 , end identify by block numb s, inequality cor	MITED.	
DISTRIBUTION ST DISTRIBUTION ST DISTRIBUTION ST DESTRIBUTION ST DESTRI	TATEMENT (of the observed enfor TATEMENT (of the observed enfor NY NOTES Winne on reverse of the Winnessee , Kuhn-Tucker vector projections. Wently occurring pro finite set of inequit t is the sheer numbe a smaller set of con Here we discuss an lems to lead to a to examples in the lass pristic in nature to	and identify by block numbers, inequality constraints, but sol blem is that of mality constraints, straints, but sol approach which us tal solution. We tal solution. We	MITED. Tern Report) Astraints, iteration, least Minimizing a convex function S. Often what makes the prob- That is, we could solve the living for the total set causes ses our ability to solve these e will illustrate the method paper. Our approach will be nding.	
DISTRIBUTION ST DISTRIBUTION ST E. SUPPLEMENTAR Optimization squares, I- B. AFSTRACY/Can A freq subject to a lem difficul problem for difficulty. partial prob with several somewhat heu DD 1/44 73 147.	ATEMENT (of the observed enfor TATEMENT (of the observed enfor NUMBER of reverse of the Unseccessive the set of the seccessive the sheer number a smaller set of con Here we discuss an lems to lead to a to examples in the lass written of the sec of a sec to the sec of the seccessive the seccessive seccessive seccessive the seccessive seccessive seccessive the seccessive seccessive the seccessive seccessive the seccessive seccessive seccessive the seccessive seccessive seccessive the seccessive seccessive seccessive the seccessive seccessive seccessive seccessive seccessive the seccessive seccessi	and identify by block numbers, inequality constraints, blem is that of mality constraints, straints, but sol approach which us tal solution. We it section of the promote understan	MITED. Tom Report) Tom Report istraints, iteration, least minimizing a convex function is. Often what makes the prob- . That is, we could solve the living for the total set causes ses our ability to solve these is will illustrate the method paper. Our approach will be nding.	

.

4

. .

