

AD-A147 389

MONOTONE OPERATORS IN MATHEMATICAL PROGRAMMING(U)
GEORGIA INST OF TECH ATLANTA SCHOOL OF MATHEMATICS
J E SPINGARN AUG 84 AFOSR-TR-84-0902 AFOSR-80-0195

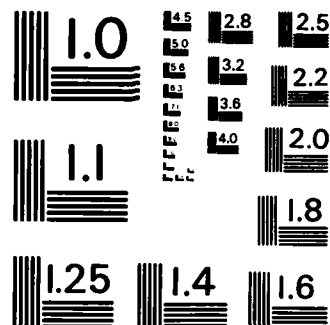
1/1

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR. 84-0902

FINAL REPORT

12

AD-A147 309

MONOTONE OPERATORS IN MATHEMATICAL PROGRAMMING

By

Jonathan E. Spingarn
Principal Investigator

Submitted to

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
BOLLING AIR FORCE BASE
WASHINGTON, DC 20332

Under

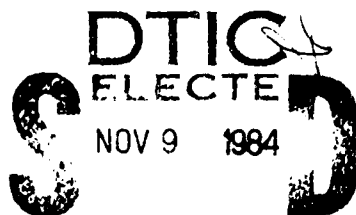
Grant No. AFOSR-80-0195

Approved for public release;
distribution unlimited.

August 1984

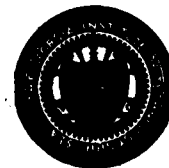
GEORGIA INSTITUTE OF TECHNOLOGY

A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF MATHEMATICS
ATLANTA, GEORGIA 30332



B

1984



84 10 23 154

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b RESTRICTIVE MARKINGS	
2a SECURITY CLASSIFICATION AUTHORITY		3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE		5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 84-0902	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research	
6a NAME OF PERFORMING ORGANIZATION Georgia Institute of Technology	6b OFFICE SYMBOL (if applicable)	7b. ADDRESS (City, State, and ZIP Code) Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332	
6c ADDRESS (City, State, and ZIP Code) Atlanta GA 30332	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-80-0195		
8a NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b OFFICE SYMBOL (if applicable) NII	10 SOURCE OF FUNDING NUMBERS	
8c ADDRESS (City, State, and ZIP Code) Bolling AFB DC 20332		PROGRAM ELEMENT NO. 611027	PROJECT NO. 1304
		TASK NO. A6	WORK UNIT ACCESSION NO.
11 TITLE (Include Security Classification) MONOTONE OPERATORS IN MATHEMATICAL PROGRAMMING			
12 PERSONAL AUTHOR(S) Jonathan E. Spingarn			
13a TYPE OF REPORT Final	13b TIME COVERED FROM 15/6/83 TO 14/6/84	14 DATE OF REPORT (Year, Month, Day) AUG 84	15 PAGE COUNT 19
16 SUPPLEMENTARY NOTATION			
17 COSATI CODES		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
		Monotone multifunction; proximal point algorithm; decomposition algorithm; convexity; complementarity.	
19 ABSTRACT (Continue on reverse if necessary and identify by block number)			
An algorithm was investigated for solving problems where the object is to find $x \in A$ and $y \in T(x)$, where T is a maximal monotone multifunction. An algorithm was described for finding a feasible point for a system of linear inequalities. For inconsistent systems, the feasible point algorithm was shown to generate a sequence converging at a linear rate to the set of least-square solutions. A primal-dual decomposition method was investigated to solve the separable convex programming problem.			
20 DISTRIBUTION AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS		21 ABSTRACT SECURITY CLASSIFICATION	
22a NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff		22b TELEPHONE (Include Area Code) 5027	22c OFFICE SYMBOL

MONOTONE OPERATORS IN MATHEMATICAL PROGRAMMING

FINAL REPORT

Grant No. AFOSR-80-0195

August 1984

Submitted to

Air Force Office of Scientific Research
Bolling Air Force Base, D.C. 20332

by

GEORGIA INSTITUTE OF TECHNOLOGY
School of Mathematics

DTIC
ELECTE
NOV 9 1984
B

Jonathan E. Spingarn

Dr. Jonathan E. Spingarn
Principal Investigator

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release under E.O. 11734-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

ABSTRACT

For T a maximal monotone operator on a Hilbert space H and A a closed subspace of H , the "partial inverse" T_A of T with respect to A was introduced. T_A is maximal monotone. The proximal point algorithm, as it applies to T_A , was shown to result in a simple procedure, the "method of partial inverses", for solving problems in which the object is to find $x \in A$ and $y \in A^\perp$ such that $y \in T(x)$. This method was shown to specialize to give new algorithms for solving several optimization and equilibrium problems.

An algorithm was described for finding a feasible point for a system of linear inequalities. If the solution set has nonempty interior, termination was shown to occur after a finite number of iterations. The algorithm is a projection type method, similar to the relaxation methods of Agmon, Motzkin, and Schoenberg. It differs from previous methods in that it solves for a certain "dual" solution in addition to a primal solution. It is a special case of the method of partial inverses.

The feasible point algorithm was shown, for inconsistent systems, to generate a sequence converging at a linear rate to the set of least square solutions.

A primal-dual decomposition method was investigated to solve the separable convex programming problem. Convergence to a solution and Lagrange multiplier vector from an arbitrary starting point was demonstrated. The method was shown to be equivalent to the method of

partial inverses. In the nonseparable case, it was shown to specialize to a known method, the proximal method of multipliers. Conditions were provided which guarantee linear convergence of the algorithm.

For families of nonlinear programming problems, new conditions were established which guarantee uniqueness of the global optimizer to be a generic property.

Accession For		
NTIS	ORNL	✓
DTIC	TR	
Unannounced		
Justification		
By		
Distribution		
Availability Codes		
Dist	Avail and/or	Special
A-1		

0210
Dist
Availability

RESEARCH SUMMARY

Many optimization problems are equivalent to a problem of finding a zero of a maximal monotone operator T on a Hilbert space H :

$$(1) \quad \text{to find } x \in H \text{ such that } 0 \in T(x).$$

The known "proximal point algorithm" for solving (1) takes an arbitrary starting point $x_0 \in H$ and a sequence (c_n) of positive real numbers and determines a sequence (x_n) by repeatedly applying the proximal mapping:

$$(2) \quad x_{n+1} = (I + c_n T)^{-1}(x_n).$$

A vast array of problems can be regarded as special instances of (1). Examples are convex or linear programming, monotone complementarity problems, network flow problems, variational inequalities, and systems of linear equations or inequalities. Nevertheless, the known uses of the proximal iteration are few. By far, the most important application is found in the method of multipliers of Hestenes and Powell. This algorithm for solving convex programming problems was known for some time before Rockafellar showed it to be a special instance of the proximal point algorithm. Outside of this, there are few serious uses. Although it is theoretically possible to execute, as a practical matter the iteration (2) can be performed only in a few cases.

There are many problems which can be expressed in the form (1),

are not easily solved by the proximal point algorithm, but have the interesting property that T can be "decomposed" into simpler monotone multifunctions that are easy to approximate. (To "approximate" means to execute the rule (2), i.e., to evaluate the proximal mapping). Our research in [1] has shown that new algorithms can be developed which exploit such decomposition.

The principal new idea that enabled us to extend the powers of the proximal point algorithm and derive such decomposition methods is our notion of the "partial inverse" of a monotone mapping. If A is a closed subspace of H and $B=A^\perp$, then each $x \in H$ can be written uniquely as $x = x_A + x_B$ with $x_A \in A$ and $x_B \in B$. If T is a multifunction, the partial inverse of T with respect to A is the multifunction $T_A : H \rightarrow H$ defined by $v \in T_A(u)$ iff there exists x and y in H with $y \in T(x)$, $u = x_A + y_B$ and $v = x_B + y_A$. T_A is (maximal) monotone iff T is (maximal) monotone.

We showed that T_A could be used as a vehicle for introducing introducing duality into and solving many problems. Often, a problem can be expressed in the following form (for suitable choices of T and A):

$$(3) \quad \begin{array}{l} \text{to find } x \in A \text{ such that there exists } y \in A \\ \text{with } 0 \in T_A(x+y), \end{array}$$

where y is a "dual variable" and T is maximal monotone. If z could be found such that $0 \in T_A(z)$ then $x=z_A$ would solve

(3). In [1] we introduced the idea of solving (3) by finding such z via the proximal point algorithm. Theoretically, this can definitely be done; the fact that it is also practical was demonstrated in [1]. There it was shown that this leads to the following iterative procedure, the "method of partial inverses" to solve (3):

ALGORITHM 1 ("method of partial inverses" [1]).

Initialize: Choose arbitrary $x_0 \in A$ and $y_0 \in B$.

Iteration k ($k=0,1,\dots$):

Proximal step: Find $x'_k, y'_k \in H$ such that

$$x_k + y_k = x'_k + y'_k \quad \text{and} \quad \frac{1}{c_k} (y'_k)_A + (y'_k)_B \in T((x'_k)_A + \frac{1}{c_k} (x'_k)_B).$$

Projection step: Let $x_{k+1} = (x'_k)_A$ and $y_{k+1} = (y'_k)_B$.

In [2], we applied the partial inverse approach to the problem of solving a system of linear inequalities:

(4) to find $x \in R^d$ such that $\langle x, u_i \rangle \leq b_i, \quad i=1,\dots,n$

(where $0 \neq u_i \in R^d$ and $b_i \in R$). Let $C_i = \{x : \langle x, u_i \rangle \leq b_i\}$ and $C = C_1 \cap \dots \cap C_n$. It is a straightforward matter to apply Algorithm 1 to solve (4). Our principal accomplishment in [2] was to establish that that algorithm actually terminates after a finite number of iterations, provided the solution set has nonempty interior.

The algorithm so obtained is a new addition to the family of known

"projection" methods. These solve (4) by computing a sequence of projections onto the halfspaces $C_i := \{ x : \langle x, u_i \rangle \leq b_i \}$. The best known of these are the "relaxation" methods of Agmon and Motzkin and Schoenberg. According to the simplest of these, a sequence (x_k) is generated by taking x_{k+1} to be the projection of x_k onto the furthest halfspace C_i . Another possibility is to project in turn in some fixed cyclical order onto the sets C_i . In another variation studied by Motzkin and Schoenberg, one takes x_{k+1} to be the reflection of x_k across the furthest hyperplane. This last method has the surprising property that a solution is always found in a finite number of steps if the solution set has nonempty interior. In a more recent study by Goffin, classes of problems (4) were identified for which finite termination in the relaxation method occurs without reflection. However, the reflection method is the only one of these known methods for which finite termination always occurs under the sole assumption that the feasible set has nonempty interior, so our result proving finite convergence of the algorithm is quite surprising.

The algorithm we have proposed to solve (4) is:

ALGORITHM 2.

Start : Choose arbitrary $x_0, y_{01}, \dots, y_{0n} \in \mathbb{R}^d$
 with $y_{01} + \dots + y_{0n} = 0$.

Step k ($k=0,1,\dots$): Compute

$$x'_{ki} = \text{proj}_{C_i}(x_k + y_{ki}), \quad i=1,\dots,n$$

$$y'_{ki} = x_k + y_{ki} - x'_{ki}, \quad i=1,\dots,n$$

and update

$$x_{k+1} = \frac{1}{n} \sum_{i=1}^n x'_{ki}$$

$$y_{k+1,i} = y'_{ki} - \frac{1}{n} \sum_{j=1}^n y'_{kj}, \quad i=1,\dots,n$$

From results we proved in [1] it follows that regardless of the choice of starting values $x_0, y_{01}, \dots, y_{0n}$,

either $x_k \rightarrow x$ and $y_{ki} \rightarrow y_i$ with $x \in C$,

y_i normal to C_i at x , and $y_1 + \dots + y_n = 0$,

or $|(x_k + y_{k1}, \dots, x_k + y_{kn})| \rightarrow \infty$ and (4) is inconsistent.

Our main result from [2] regarding Algorithm 2 states that termination occurs after a finite number of iterations if the interior of C is nonempty. More precisely, $\text{int}(C) \neq \emptyset$ implies for some k that

$$x_k = x_{k+1} = \dots$$

$$0 = y_{k1} = y_{k+1,1} = \dots$$

.

.

.

$$0 = y_{kn} = y_{k+1,n} = \dots$$

with $x_k \in C$.

In [3], we investigated the behavior of Algorithm 2 in the cases where $\text{int}(C) = \emptyset$ or where $C = \emptyset$. If $\text{int}(C) = \emptyset$ but $C \neq \emptyset$, we showed that the sequence (x_k) converges to a solution and that the distance to the solution set approaches zero at a linear rate. Even in the case where $C = \emptyset$, we got convergence at a linear rate to the set of least-square solutions of (4), and (x_k) converges to one particular least-square solution. More precisely, we have proven in [3] the following (X denotes the set of least square solutions)

THEOREM. Let sequences (x_k) , (y_{ki}) , (x'_{ki}) , (y'_{ki}) be produced by the Algorithm 2. The sequences (x_k) and (x'_{ki}) ($i=1, \dots, n$) converge to limits

- i. $x_k \rightarrow x_\infty$
- ii. $x'_{ki} \rightarrow x'_{\infty i}$, $i=1, \dots, n$,

where

- i. $x_\infty = \frac{1}{n} \sum_i x'_{\infty i}$ and $x'_{\infty i} \in C_i$
- ii. $x'_{\infty i} = \text{proj}_{C_i}(x_\infty)$

$x_\infty \in X$ and the vector $(x'_{\infty 1} - x_\infty, \dots, x'_{\infty n} - x_\infty)$ equals the vector of smallest norm in the set

$$\{(w_1 - \frac{1}{n} \sum w_i, \dots, w_n - \frac{1}{n} \sum w_i) : w_1 \in C_1, \dots, w_n \in C_n\}.$$

This vector also equals the element $\hat{v}$ of smallest norm in $\text{range}(T_A)$.
For the sequences (y_{ki}) , (y'_{ki}) , one has

$$\lim_{k \rightarrow \infty} (y_{k1}, \dots, y_{kn})/k = \lim_{k \rightarrow \infty} (y'_{k1}, \dots, y'_{kn})/k = -\hat{v}.$$

In [4], we applied the partial inverse approach to obtain a new decomposition algorithm for the solution of separable convex programming problems.

The convex programming problem is

$$(5) \quad \begin{aligned} &\text{to minimize } f_0(x) \text{ subject to } x \in C \text{ and} \\ &f_i(x) \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

We assume the functions f_i ($i=0, \dots, m$) to be finite-valued convex: $R^d \rightarrow R$, and $C \subset R^d$ to be closed convex. In the separable case of (5),

$$(6) \quad \begin{aligned} &\text{there are functions } f_{ij} : R^{d_j} \rightarrow R \quad (0 \leq i \leq m, 1 \leq j \leq n) \\ &(d = d_1 + \dots + d_n) \text{ such that for each } i, \\ &f_i(x) = \sum_{j=1}^n f_{ij}(x_j) \quad (x = (x_1, \dots, x_n), \quad x_j \in R^{d_j}) \\ &\text{and there are closed convex sets } C_j \subset R^{d_j} \\ &\text{such that } C = C_1 \times \dots \times C_n \subset R^d. \end{aligned}$$

Our method is closely related to a family of "dual" methods for the solution of (5). The prototype for such methods, the classical dual approach, involves the Lagrangian function

$$L(x,y) = f_0(x) + \sum_{i=1}^m y_i f_i(x) \quad (x \in C, y \geq 0)$$

and the concave dual objective function

$$g(y) = \inf_{x \in C} L(x,y) \quad (y \geq 0).$$

The dual problem to (5) is

(7) to maximize $g(y)$ subject to $y \geq 0$.

Under mild assumptions, the infimum in (5) equals the supremum in (7) and a solution $\bar{y} \geq 0$ to (7) exists. Assuming the existence of such $\bar{y}$, it is known that $\bar{x}$ solves (5) if, and only if, $(\bar{x}, \bar{y})$ is a saddle-point for L .

In a typical dual approach, one minimizes $L(\cdot, y_k)$ over C for a sequence of values $y_k \geq 0$, obtaining a sequence x_k which hopefully converges to an optimum while y_k converges to a dual optimum. Several versions of this strategy have been suggested in the literature.

One very valuable characteristic of the classical approach is that it leads to decomposition algorithms. For the separable problem (6),

$$\begin{aligned} L(x,y) &= f_0(x) + \sum_i y_i f_i(x) \\ &= \sum_j (f_{0j}(x_j) + \sum_i y_i f_{ij}(x_j)). \end{aligned}$$

Thus the minimization in x of $L(x,y)$ over C decomposes into the n separate minimizations of $L_j(x_j, y) = f_{0j}(x_j) + \sum_i y_i f_{ij}(x_j)$ over C_j .

This "Lagrangian decomposition" has been exploited by numerous authors. By replacing a d -dimensional constrained problem with a sequence of less constrained problems of lower dimensions $d_1, \dots, d_n$, it offers a great advantage. Some authors have investigated ways of applying this technique in situations where separability is absent, manufacturing separability by the device of replacing functions with their linear approximations (linear functions are always separable).

However, the classical approach has several potential drawbacks. First, for some $y \geq 0$, $L(x, y)$ may fail to achieve its minimum on C . Worse, $g(y)$ may have the value $-\infty$. Such values of y must be ruled out, so the dual problem, in addition to the nice constraint $y \geq 0$, has the possibly nasty constraint $g(y) > -\infty$. It is possible to generate a sequence x_k failing to be a minimizing sequence for (5) even with y_k being a maximizing sequence for the dual. All of these inconveniences can be ruled out by imposing appropriate assumptions. The most serious problem is that one is severely restricted in the choice of a method to maximize g . Each evaluation of the function $g(y)$ requires that the function $L(x, y)$ be minimized in x over C . Thus any method requiring many evaluations of the function g is impractical. Fortunately, the minimization in x of $L(\cdot, y)$ yields, at no extra cost, a subgradient for g at y , a fact which motivates Uzawa's method, a steepest ascent approach to the maximization of g . The Dantzig-Wolfe algorithm can be viewed as the approach whereby g is maximized via a cutting plane method.

One attractive strategy for avoiding some of the pitfalls

of the classical dual approach is offered by Rockafellar's proximal method of multipliers, a modification of the multiplier method of Hestenes and Powell. At each step, one has $\bar{x}$ and $\bar{y}$ at hand and minimizes the "augmented Lagrangian"

$$(8) \quad \rho(x) = f_0(x) + \frac{1}{2} |x - \bar{x}|^2 + \sum_i \max^2\{0, f_i(x) + \bar{y}_i\}$$

over the set C to obtain the next $\bar{x}$. The strong convexity of $\rho(x)$ guarantees the existence of a unique minimum and the next multiplier $\bar{y}$ is chosen according to a simple update rule. Global convergence is guaranteed under remarkably weak assumptions: f_i finite convex, C closed convex, and existence of a solution-multiplier pair.

Unfortunately, the penalty function (8) cannot be written as a sum of n functions $\rho_j(x_j)$, so the augmented Lagrangian approach does not directly yield a decomposition algorithm. This is the principal disadvantage of the augmented Lagrangian approach used for decomposition purposes. Several strategies for dealing with this problem have been discussed in the literature. One way around this difficulty is to either replace $\rho(x)$ by a linear approximation or rely on a method that minimizes $\rho(x)$ that uses linear approximations, such as Frank-Wolfe.

In the nonseparable case ($n=1$), our method specializes to the proximal method of multipliers. In the separable case, (8) is replaced by a function of the form $\sum \rho_j$, where each ρ_j is a function only of x_j and looking very much like (8). The user is free

to minimize the strongly convex function ρ_j over the set C_j (this is the only constraint in the subproblem minimization) by any method desired. The advantages of the proximal method of multipliers are retained. One still has existence and uniqueness of a minimum in each subproblem minimization, and global convergence to a solution and multiplier under the same minimal assumptions, even when the minimization is performed only approximately according to the stopping criterion we provide. The update rule for the multipliers is just as simple as in the proximal multiplier method.

There are two basic approaches to the solution of the separable problem in a hierarchical or multi-level fashion. These are resource-directive and price-directive methods. In the resource directive approach, one iteratively determines values u_{ij} such that the solutions of the problems

$$(9) \quad \min f_j(x_j) \quad \text{subject to } x_j \in C_j \quad \text{and} \quad f_{ij}(x_j) \leq u_{ij}$$

($j=1, \dots, n$) solve (5). In the price-directive approach, one iteratively determines "prices" $y_1, \dots, y_m$ such that the optimal solutions of the problems

$$(10) \quad \min f_{0j}(x_j) + \sum y_i f_{ij}(x_j) \quad \text{subject to } x_j \in C_j$$

($j=1, \dots, n$), also solves (5).

Our method differs from existing methods in that it iteratively minimizes a function depending both on prices $y_1, \dots, y_m$ and resources u_{ij} . The prices converge to values such that the solutions to (9) are solutions to (5) and the allocations

converge to values such that the solutions to (10) solve (5). The method is in this sense both price- and resource-directive.

Let $\eta > 0$ be an arbitrary constant. If we define $\rho_j: R^{d_j} \rightarrow R$ by

$$\rho_j(q'_j) = \begin{cases} f_{0j}(q'_j) + \frac{1}{2\eta} |q'_j - x_j|^2 + \frac{1}{2\eta} \sum_i \max^2\{0, f_{ij}(q'_j) - u_{ij} + \eta y_i\} & (\text{if } q'_j \in C_j) \\ +\infty & (\text{otherwise}) \end{cases}$$

the decomposition algorithm we have introduced is

ALGORITHM 3.

Initialize: Start with arbitrary $x = (x_1, \dots, x_n) \in R^d$, $y \in R^m$, and $u \in R^{mn}$ such that $\sum_j u_{ij} = 0$, $i=1, \dots, m$.

Iteration k ($k=0, 1, \dots$):

Minimization step: For $j=1, \dots, n$, find the unique q'_j to minimize $\rho_j(q'_j)$ (subject to the implicit constraint $q'_j \in C_j$).

Update: Let $u'_{ij} = \max\{u_{ij} - \eta y_i, f_{ij}(q'_j)\}$ and then

$$x_j^+ = q'_j, \quad u_{ij}^+ = u'_{ij} - \frac{1}{n} \sum_k u'_{ik}, \quad y_i^+ = y_i + \frac{1}{\eta n} \sum_k u'_{ik}.$$

Our principal result regarding convergence of this algorithm is:

THEOREM. Let Algorithm 3 be applied to the convex programming problem (5). The algorithm is equivalent to the proximal point algorithm in the sense that $(x^+, u^+ + y^+) = (I + (\eta \partial F)_A)^{-1}(x, u + y)$, and F is a certain lower semicontinuous convex function. Suppose the minimization step is performed accurately enough so that

$$\text{in step } k: \text{ for each } j, \text{ dist}(0, \partial p_j(q_j^n)) \leq \frac{\epsilon_k}{\eta} \quad (\sum \epsilon_k < \infty)$$

holds. If the generated sequence of iterates $(x, u + y)$ is unbounded, there exists no solution-multiplier pair. If the sequence is unbounded and the Slater condition is satisfied, then the convex programming problem (5) has no solution. If the sequence is bounded, then x , y , and u converge, respectively, to a solution, Lagrange multiplier, and optimal allocation for the convex programming problem. Convergence occurs at a linear rate provided the strong second-order optimality conditions are satisfied for the problem (5).

In [5], we discussed the parameterized family

$$(Q_p) \quad \begin{aligned} &\text{to minimize } f_0(x, p) \text{ subject to } x \in C \text{ and} \\ &f_i(x, p) \leq 0 \quad \text{for } i \in I = \{1, \dots, s\} \\ &= 0 \quad \text{for } i \in J = \{s+1, \dots, m\}. \end{aligned}$$

By a "generic" property of the family (Q_p) is meant a property which holds for all problems (Q_p) except possibly for values of p in a subset of Lebesgue measure zero.

Rockafellar and Fujiwara have (independently) given examples of

constrained families (Q_p) having the property that for almost all p , (Q_p) either has a unique global minimum, or no global minimum at all. Using the transversality theorem, we showed that their results can be extended to more general classes of parameterizations.

Fujiwara considered the family

$(Q_{u,v})$ to minimize $f(x) - x \cdot u$ subject to $g(x) = b + v$

$(f : \mathbb{R}^n \rightarrow \mathbb{R}, g : \mathbb{R}^n \rightarrow \mathbb{R}^m)$ and showed, assuming sufficient differentiability of the functions f and g , that for all v , $(Q_{u,v})$ has at most one global solution for almost all u . This implies, by Fubini's theorem, that for almost all (u,v) , $(Q_{u,v})$ has at most one globally optimal solution. He proved a similar result for problems where the equality constraint is replaced with an inequality. Fujiwara obtained his result by applying a theorem of Araujo and Mas-Colell.

Rockafellar obtained a similar result for the more general family

$(Q_{w,u,v})$ to minimize $f_0(v,x) + w \cdot x$ over all x satisfying

$$\begin{aligned} f_i(v,x) + u_i &\leq 0 && \text{for } i=1,\dots,s \\ &= 0 && \text{for } i=s+1,\dots,m. \end{aligned}$$

Only the linear perturbation w of the objective function has a real role in his proof of global uniqueness; like Fujiwara, he holds the other parameters fixed, shows the result to hold for almost all w and then invokes Fubini's Theorem to show that the result holds for almost all values of the parameters combined. The principal tool he uses to obtain almost sure uniqueness is the

fact that a convex function is differentiable except on possibly a set of measure zero (although convexity of the functions f_i is not assumed). The Mas-Colell theorem employed by Fujiwara is proved by quite direct and elementary means. In a more recent paper, Fujiwara has established a generic global uniqueness property for constrained problems under a compactness assumption on the constraint set. This result, however, deals with "generic" properties which hold for all functions f_i in an open dense set with respect to the strong Whitney C^2 -topology.

Our result on the generic global uniqueness property can be regarded as a generalization of the Mas-Colell theorem to constrained problems. Our use of the transversality theorem simplifies the proof considerably, though at the expense of slightly stronger differentiability assumptions. We considered families of the form

$$\begin{aligned} (Q_{p,q}) \quad & \text{to minimize } f_0(x,p,q) \text{ subject to} \\ & f_i(x,q) \leq 0, \quad i=1,\dots,s, \\ & \quad \quad \quad = 0, \quad i=s+1,\dots,m, \end{aligned}$$

where the assumptions on f_i are as before and the parameters vary over some open set in Euclidean space. The criterion we used to establish generic global uniqueness was:

$$(11) \quad \text{for all } x_1 \neq x_2, \text{ and all } q, \text{ the function} \\ p \mapsto f_0(x_1,p,q) - f_0(x_2,p,q) \text{ is of rank one at all } p$$

The Rockafellar and Fujiwara families considered above are easily seen to satisfy criterion (11).

THEOREM. Let the family $(Q_{p,q})$ satisfy the criteria (11) and the function $q \rightarrow (f_1(x,q), \dots, f_m(x,q))$ is of full rank m for all x at every q .

Then for almost all (p,q) , $(Q_{p,q})$ has at most one global optimizer. In fact, for almost all (p,q) , $f_0(\cdot, p, q)$ cannot achieve the same value at any two distinct critical points (points satisfying the first-order optimality conditions along with some y).

PUBLICATIONS

- [1] "Partial inverse of a monotone operator," Applied Mathematics and Optimization 10 (1983) 247-265.
- [2] "A primal-dual projection method for solving systems of linear inequalities", accepted by Linear Algebra and its Applications.
- [3] "A projection method for least square solutions to overdetermined systems of linear inequalities", submitted.
- [4] "Applications of the method of partial inverses to convex programming", accepted by Mathematical Programming, currently being revised.
- [5] "Multifunctions associated with parameterized classes of constrained optimization problems", submitted.

END

FILMED

12-84

DTIC