G-37-614

FINAL REPORT

PROXIMAL POINT ALGORITHMS IN MATHEMATICAL PROGRAMMING

Submitted by

Jon Spingarn
School of Mathematics

Submitted to

National Science Foundation
1800 G Street, NW
Washington, DC 20550

Contracting through

Georgia Tech Research Corporation

February 1987

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

FINAL REPORT

Proximal Point Algorithm in Mathematical Programming

Submitted by

Jon Spingarn
School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332

Submitted to

National Science Foundation
1800 G Street, NW
Washington, DC 20550

Contracted through

Georgia Tech Research Corporation
Georgia Institute of Technology
Atlanta, Georgia 30332

February 1987

NATIONAL SCIENCE FOUNDATION
Washington, D.C. 20550

FINAL PROJECT REPORT
NSF FORM 98A

PLEASE READ INSTRUCTIONS ON REVERSE BEFORE COMPLETING

PART I–PROJECT IDENTIFICATION INFORMATION

1. Institution and Address	2. NSF Program	3. NSF Award Number
Georgia Institute of Tech. School of Mathematics Atlanta, GA 30332	Math. Sci. Applied Math.	DMS-8506712
	4. Award Period From 6/15/85 To 11/30/87	5. Cumulative Award Amount $33,089

6. Project Title

Proximal Point Algorithm in Mathematical Programming

PART II–SUMMARY OF COMPLETED PROJECT *(FOR PUBLIC USE)*

For a nonexpansive piecewise isometric mapping q on R^d having a fixed point, we have studied the iterates of $(I+q)/2$. Such iterates were shown to behave, eventually, as if q were actually an isometry. Under certain circumstances, a finite number of iterations are guaranteed to produce a fixed point. Examples dealing with systems of linear inequalities and with network flows have been examined. A new constructive proof has been presented that for any real anti-symmetric matrix A, there exists $x \geq 0$ such that $Ax \geq 0$ and $x+Ax > 0$.

Some properties of iterates of nonexpansive mappings on R^d and especially of nonexpansive piecewise isometries (foldings) have been explored. For a nonexpansive q, the set S if all cluster points of q-sequences $x, q(x), q^2(x), \ldots$ is nonempty only if q has a fixed point. S is closed, convex, $q|S$ is an isometry, and all q-sequences converge to S. A characterization was presented for the class of all foldings having the property that the set S absorbs every q-sequence after only finitely many iterations.

A solution approach has been proposed for a certain class of network equilibrium problems. The method of solution is a specialization to a network setting of the method of partial inverses. Application is discussed to the computation of economic equilibria.

PART III–TECHNICAL INFORMATION *(FOR PROGRAM MANAGEMENT USES)*

1. ITEM *(Check appropriate blocks)*	NONE	ATTACHED	PREVIOUSLY FURNISHED	TO BE FURNISHED SEPARATELY TO PROGRAM: Check (√)	TO BE FURNISHED SEPARATELY TO PROGRAM: Approx. Date
a. Abstracts of Theses					
b. Publication Citations					
c. Data on Scientific Collaborators					
d. Information on Inventions					
e. Technical Description of Project and Results					
f. Other *(specify)*					

2. Principal Investigator/Project Director Name *(Typed)*	3. Principal Investigator/Project Director Signature	4. Date
Jon Spingarn		2/13/87

TABLE OF CONTENTS

I. Iterates of Nonexpansive Mappings

A. Summary

Working with Dr. Jim Lawrence (George Mason University), we have studied nonexpansive piecewise isometric mappings (foldings) and the behavior of sequences of iterates of such mappings. The results of our research to date are contained in [1] and [2].

If $q: R^d \to R^d$ is a nonexpansive mapping having a fixed point, it is known that any sequence

(1) $$x_{k+1} = (q(x_k) + x_k)/2$$

converges to a fixed point of q [8]. In [1], the principal goal was to exhibit a class of mappings, the nonexpansive piecewise isometric mappings, or "foldings", for which the behavior of the iteration (1) is especially favorable.

The iteration (1) is equivalent to the "proximal point algorithm" for finding a zero of a maximal monotone multifunction [9], [10], [11]. If T is maximal monotone on R^d, the proximal point algorithm generates a sequence

(2) $$x_{k+1} = p(x_k), \quad p = (I + T)^{-1}.$$

Furthermore, the mapping $q = 2p - I$ is nonexpansive so (1) and (2) are equivalent.

In [1], the iteration (1) was studied in detail for the case where q is a folding. It was shown that the

sequence (1) behaves, eventually, as if q were actually an isometry.

It was also shown in [1] that this fact has many surprising consequences, especially the finite termination of several algorithms. As examples, we described algorithms for solving systems of linear inequalities and network flow problems. We also obtained a constructive and simple new proof of the fact that for any antisymmetric real $n \times n$ matrix A, there exists $x \geq 0$ such that $Ax \geq 0$ and $x + Ax > 0$.

The paper [2] was motivated by a desire to determine whether or not the iterates

$$x_{k+1} = q(x_k) \tag{3}$$

of a folding q coincide, as do the iterates (1), after finitely many steps, with the iterates of some isometry. The answer turned out to be negative. This outcome is consistent with the fact that for a general nonexpansive mapping q (not necessarily a folding), the iterates of $(I+q)/2$ have long been known to behave more nicely than the iterates of q; for instance, any sequence (1) converges to a fixed point of q (if one exists) [8], but the same is not in general true for (3).

It was shown in [2] that much can still be said about the iterates of (3) when q is nonexpansive, and still more can be said if q is a folding. In [2], we first analyzed the structure of foldings. We showed that a folding induces a decomposition of the underlying space into finitely many polyhedral convex sets (the "folds" of q). If q has a fixed point, we showed that there is a unique fold whose interior contains a fixed point, and that the fixed point set is the intersection of this fold with an affine set.

In [2], we defined the "cluster set" consisting of all cluster points of q-sequences, and showed this set to be nonempty only if q possesses a fixed point. If q has a fixed point, we demonstrated that all q-sequences converge to S (though not necessarily to a particular point in S). Further, S is closed and convex, $q(S) = S$, and the restriction of q to S is an isometry.

We then turned to the question of characterizing all foldings having the property: for every x there exists k such that $q^k(x) \in S$. As a particularly illuminating example of a folding that fails to have this property, consider the folding $q: R^2 \to R^2$ defined by $q(x,y) = s(r(x,y))$, where r is a rotation about the origin by an angle that is an irrational multiple of pi and

$$s(x,y) = \begin{cases} (x,y) & \text{if } x < 1 \\ (2-x,y) & \text{if } x \geq 1. \end{cases}$$

It is easy to see that for any $(x,y) \in R^2$, the iterates $q^k(x,y)$ move arbitrarily close to the cluster set $S = \{(x,y): x^2 + y^2 \leq 1\}$ and that the restriction of q to S coincides with the isometry r. But it followed from our results in [2] that there do exist points (x,y) for which $q^k(x,y) \notin S$ for all k. Further, we presented in [2] a characterization of the class of foldings q whose cluster set absorbs all q-sequences.

B. The Proximal Iterates of a Folding -- Summary of the Main Results from [1]

A multifunction $r: R^d \rightrightarrows R^d$ assigns a subset r(x) of R^d to each $x \in R^d$. The set of x such that r(x) is nonempty is dom(r), the domain of r. We may identify r with its graph, namely the set $\{(x,y): y \in r(x)\}$. For the sake of simplicity, we will use the same symbol to represent a multifunction and its graph. For example, the notations $y \in r(x)$ and $(x,y) \in r$ will be used interchangeably. If for all (x,y) and (x',y') in r

i. $|y-y'| \le |x-x'|$, then r is nonexpansive

ii. $|y-y'|^2 \le |x-x'|^2 - |(x-y) - (x'-y')|^2$, then r is proximal

iii. $\langle x-x', y-y' \rangle \ge 0$, then r is monotone.

If r is nonexpansive, then r is maximal nonexpansive if (the graph of) r is not properly contained in (the graph of) another nonexpansive r'. In an identical manner, we can define maximal proximal and maximal monotone.

The mapping $(x,y) \overset{\alpha}{\mapsto} (x,2y-x)$ induces a one-to-one correspondence of the class of proximal multifunctions onto the class of nonexpansive multifunctions. By this we mean that the image under α of (the graph of) a proximal multifunction is (the graph of) a nonexpansive multifunction and the image under α^{-1} of a nonexpansive multifunction is proximal. Likewise, $(x,y) \overset{\beta}{\mapsto} (x+y,x-y)$ carries monotone onto nonexpansive multifunctions. Of course, the composition $(x,y) \overset{\alpha^{-1}\beta}{\mapsto} (x+y,x)$ carries monotone onto proximal multifunctions.

In the following, we take q, p, and T to be corresponding nonexpansive, proximal, and monotone multifunctions i.e., $q = \alpha(p) = \beta(T)$. From the definitions, it is apparent that

$$p = (I+q)/2 = (I+T)^{-1}.$$

The correspondences α and β clearly preserve maximality. Thus q is maximal nonexpansive iff p is maximal proximal iff T is maximal monotone. Since $\beta(x,0) = \alpha(x,x) = (x,x)$, the fixed point set for p, the fixed point set for q, and the set of zeros of T coincide.

A nonexpansive multifunction q is obviously single-valued on its domain. A theorem of Kirszbraun [12] states that q is maximal if, and only if, $\mathrm{dom}(q) = R^d$. Thus "maximal nonexpansive" and "nonexpansive function on R^d" have the same meaning. Since $\mathrm{dom}(p) = \mathrm{dom}(q)$, the following are equivalent:

i. q is a nonexpansive function: $R^d \to R^d$

ii. p is a proximal function: $R^d \to R^d$

iii. T is a maximal monotone multifunction: $R^d \rightrightarrows R^d$.

If these equivalent conditions hold, it follows by a result of Krasnoselski [8] that for any x, the sequence

$$(4) \quad ((I+q)/2)^n(x) = p^n(x) = ((I+T)^{-1})^n, \quad n = 0,1,2,\ldots$$

converges to a fixed point of p and q (a zero of T) if one exists. Specifically, Krasnoselski's result asserts that the iterates of $(I+q)/2$ converge to a fixed point of q for any nonexpansive function q having a fixed point.

Before continuing, some historical remarks are in order. In the literature, "proximal" mappings have often been called "firmly nonexpansive" or the "resolvent of the monotone operator T". The correspondence between nonexpansive and monotone multifunctions was observed by Minty [13] who exploited it to show that for T maximal monotone on a Hilbert space, $(I+T)^{-1}$ is a function defined on the whole space. The term "proximal mapping" was first introduced by Moreau [10] to describe the function

$$x \mapsto \arg_z \min\{f(z) + \tfrac{1}{2}|z-x|^2\},$$

where f is lower semicontinuous convex. Moreau's "proximal mapping" is actually $(I+\partial f)^{-1}$, ∂f being the subdifferential of f (∂f is maximal monotone [14]). It was observed by Martinet [9] that Moreau's "proximal mapping" is in fact a "proximal function" as we have defined the term here, and that this implies convergence of its iterates. The correspondence between the class of proximal multifunctions and the class of nonexpansive multifunctions has also been observed before.

The iteration (4) has also been called the proximal point algorithm, especially when the emphasis has been placed on convergence to a zero of T. This name was introduced by Rockafellar [11], [15], in his study of algorithms for the solution of convex programming problems.

In [1], we investigated the behavior of the proximal point algorithm on nonexpansive mappings that are piecewise isometries. We call these mappings "foldings". To define this notion precisely, suppose that q is a nonexpansive function on R^d and let I denote the collection of all convex sets $K \subset R^d$ such that the restriction $q|K$ is an isometry. Let I be partially ordered by set inclusion. Every singleton trivially belongs to I, and I is closed under chain unions, so every point of R^d is contained in a maximal element of I. If $K \in I$ and $L \in I$ and $\mathrm{int}(K) \cap \mathrm{int}(L) \neq \emptyset$ then $\mathrm{conv}(K \cup L) \in I$. Thus, if K and L are distinct maximal elements of I, then $\mathrm{int}(K) \cap \mathrm{int}(L) = \emptyset$. If q has only a finite number of folds, then q is a folding and the maximal elements of I are the folds of q. If q is a folding, it is not hard to see that each fold is closed, polyhedral, and has nonempty interior.

The class of foldings is closed under composition.

For suppose q and q' are foldings with (locally finite) folds K_i $(i \in I)$ and K'_j $(j \in J)$, respectively. Then $q' \circ q$ is an isometry on each of the convex sets

$$L_{ij} := K_i \cap q^{-1}(q(K_i) \cap K'_j)$$

and these sets L_{ij} are locally finite. Thus $q' \circ q$ is a folding (although the sets L_{ij} are not necessarily its folds).

Via the correspondences α and β, composition of non-expansive functions induces corresponding operators on the classes of proximal functions and maximal monotone multi-functions. Thus, if p_1 and p_2 are proximal functions and T_1 and T_2 are maximal monotone multifunctions, it is natural to define two new operations:

$$p_1 * p_2 = \alpha^{-1}(\alpha(p_1) \circ \alpha(p_2))$$

and

$$T_1 \circledast T_2 = \beta^{-1}(\beta(T_1) \circ \beta(T_2)).$$

Since the class of foldings is closed under composition, it follows that the class of proximal mappings that correspond to foldings is closed under $*$, and the class of maximal monotone multifunctions that correspond to foldings is closed under $\circledast$. Since composition is associative, the operations $*$ and $\circledast$ are also associative.

In [1], we have investigated the behavior of sequences $x_{n+1} = p(x_n)$, where p is the proximal function corresponding to a folding $q = \alpha(p) = 2p-I$.

An observation will simplify the discussion. Suppose $\bar{x}$ is a fixed point for q (for p). Since folds are closed, there is $r > 0$ such that the open ball $B(\bar{x};r)$ is contained in the finite union of all the folds containing $\bar{x}$. For all x in $B(0;r)$, the line segment $[\bar{x},\bar{x}+x]$ is entirely contained in each fold that contains both $\bar{x}$ and $\bar{x}+x$ since folds are convex. It follows that for all x in $B(0;r)$ the restriction of q to $[\bar{x},\bar{x}+x]$ is an isometry. Defining $\hat{q}(x) = q(\bar{x}+x) - \bar{x}$ and $\hat{p}(x) = p(\bar{x}+x) - \bar{x}$, we then have $\hat{q}(0) = \hat{p}(0) = 0$, $\hat{q}$ and $\hat{p}$ are positively homogeneous on $B(0;r)$, and $\hat{q} = \alpha(\hat{p})$. Now, $\hat{q}$ and $\hat{p}$ are merely the translation of q and p for which the fixed point at $\bar{x}$ is displaced to the origin. Any statement about the behavior of q or p near $\bar{x}$ translates into a statement about $\hat{q}$ or $\hat{p}$ near 0. For this reason, we lose no generality if in discussing the behavior of q (or p) near a fixed point, we assume that fixed point to be 0 and q (or p) to be positively homogeneous on a neighborhood. And when we consider a sequence of iterates $x_{k+1} = p(x_k)$ converging to that fixed point, we may as well assume q (or p) to be positively homogeneous on the entire space since only local behavior is relevant.

The following three theorems were proved in [1] to describe the behavior of the proximal iterates of a folding.

Theorem 1 below can be interpreted as saying that the proximal iteration $x_{k+1} = p(x_k)$, when applied to a folding having a fixed point, spirals towards a fixed point $\bar{x}$. It states that for arbitrary x_0, there is a subspace $V(x_0)$ such that for all k sufficiently large, the set $\{x_k-\bar{x}, x_{k+1}-\bar{x},\ldots\}$ positively spans $V(x_0)$. (If $V(x_0) = \{0\}$, this says that $x_k = \bar{x}$ and the iteration terminates.)

The second theorem states that the restriction of q to the subspace described in the first theorem is an isometry with a unique fixed point.

The subspace described in the first two theorems depends on the starting point x. The third theorem establishes the existence of a subspace not depending on x, but having some of the same properties:

Theorem 1. Let $q = \alpha(p)$ be a positively homogeneous folding, $p^k(x) \to 0$. Let $x_k = p^k(x)$, $C_k = \text{cone}\{x_k, x_{k+1},\ldots\}$, and $L_k = \text{span}\{x_k, x_{k+1},\ldots\}$. There is a subspace $V(x)$ of R^d and $K > 0$ such that $L_k = C_k = V(x)$ for all $k \geq K$.

Theorem 2. Let $q = \alpha(p)$ be a positively homogeneous folding on R^d. Suppose $x_k = p^k(x) \to 0$, and let $V = V(x)$. Then $q|V$ is an isometry, 0 is the only fixed point of q in V, and $p|V$ is a linear isomorphism.

Theorem 3. Let $q = \alpha(p)$ be a positively homogeneous folding on R^d. There is a subspace $\hat{V} \subset \text{lin } N_F$ such that q is an isometry on $\hat{V}$, 0 is the only fixed point of q in $\hat{V}$, and if $x \in R^d$ is such that $p^k(x) \to 0$ then there is $K > 0$ (possibly depending on x) such that $p^k(x) \in \hat{V}$ for all $k \geq K$.

C. Applications of the Results of [1]

1. Polyhedral Convex Functions. In [1], we established the following, which asserts that every polyhedral convex function gives rise, in a natural way, to a folding:

Theorem 4. If $g: R^d \to R \cup \{\infty\}$ is a proper polyhedral convex function, then $\beta(\partial g)$ is a folding on R^d.

Using our results on foldings, we were then able in [1] to prove that the proximal point algorithm can be used to minimize a convex polyhedral function in only finitely many steps:

Theorem 5. Let $g: R^d \to R \cup \{\infty\}$ be a proper polyhedral convex function that achieves a minimum value, and let p be the proximal mapping $(I+\partial g)^{-1}$. Then for aribtrary x, there is k such that $p^k(x)$ is a minimizer for g. In other words, the proximal point algorithm terminates after a finite number of iterations with a minimizer.

The following presents an important class of problems for which the proximal point algorithm always finds a solution in only finitely many iterations:

Theorem 6. Let A be a subspace of R^d, $B = A^{\perp}$, $g: R^d \to R \cup \{\infty\}$ a proper polyhedral convex function, $p = \pi_A * p_g$. If the fixed point set F for p is nonempty, then it has the form $F = C+D$ with $C \subset A$ and $D \subset B$. If C has nonempty interior relative to A, or if D has nonempty interior relative to B, then for each $z \in R^d$, there is some m such that $p^m(z) \in F$.

If $A = R^d$, then $p_g = \pi_A * p_g$ and $D = B = \{0\}$. Since D (trivially) has nonempty interior relative to B, Theorem 6 implies that the iteration $x_{k+1} = p_g(x_k)$ must yield a fixed point (provided one exists) after a finite number of iterations.

If $K \subset R^d$ is a polyhedral convex set, then ψ_K, the characteristic function of K, is a polyhedral convex function. Applying Theorem 6 to the case $g = \psi_K$, we obtain the following

Theorem 7. Let $K \subset R^d$ be a nonempty polyhedral convex set, $A \subset R^d$ a linear subspace, $B = A^\perp$, $p = \pi_A * \pi_K$, $z_{k+1} = p(z_k)$. The set of fixed points of p is nonempty if, and only if, $A \cap K \neq \emptyset$. If $A \cap K$ has nonempty interior with respect to A, or if $A \cap K \neq \emptyset$ and $B \cap K^o$ has nonempty interior with respect to B, then for some m, z_m is a fixed point for p, $\pi_A(z_m) \in A \cap K$, and $\pi_B(z_m) \in B \cap K^o$.

2. Systems of Inequalities. We will now describe two iterative schemes for finding a feasible point for a system of linear inequalities.

Consider a system

$$\langle x,u_i\rangle \leq b_i \ , \quad i = 1,\ldots,n$$

of linear inequalities ($u_i \in R^d$, $b \in R$). For each i, let $C_i = \{x: \langle x,u_i \rangle \leq b_i\}$, p_i = projection onto C_i, and $q_i = \alpha(p_i)$. The q_i are clearly foldings. Also, define $C = C_1 \cap \ldots \cap C_n$, $p = p_1 * \ldots * p_n$, and $q = q_1 \circ \ldots \circ q_n$ $(= \alpha(p))$.

The first scheme we wish to suggest for solving the system is to iterate the proximal mapping p:

<u>Theorem 8</u>. If C has nonempty interior, then for each $x \in R^d$ there is $K > 0$ such that $p^K(x) = p^{K+1}(x) = \ldots \in C$.

This scheme for solving a system of linear inequalities is very closely related to the methods of Agmon [16] and Motzkin and Schoenberg [17]. In the method of "successive projection with relaxation parameter = 2" (also a finitely terminating algorithm [17, Theorem 1]), one takes $x_{k+1} = q(x_k)$ instead of $x_{k+1} = p(x_k)$.

We now describe a second method for solving the system. This latter method has been studied in some detail in [5], [6].

Define $K = C_1 \times \ldots \times C_n$. K is a polyhedral convex set in R^{dn}. Define the subspaces

$$A = \{(x_1,\ldots,x_n) \in R^{dn}: x_1 = \ldots = x_n\}$$

$$B = \{(y_1,\ldots,y_n) \in R^{dn}: y_1 + \ldots + y_n = 0\}$$

of R^{dn} and note that $B = A^{\perp}$. Clearly

$$A \cap K = \{(x,\ldots,x): x \in C\}$$

so solving the system is equivalent to finding a point in the intersection of A amd K. Let π_A and π_K denote the orthogonal projection mappings onto A and K, respectively. The algorithm is described in the following

Theorem 9 [5]. Suppose C has nonempty interior. If a sequence (z_k) is generated by the rule $z_{k+1} = (\pi_A * \pi_K)(z_k)$, then $\pi_A(z_m) \in A \cap K$ for some m.

3. Network Flows. Let G be a network (directed graph). Abstractly, G consists of a finite set N of "nodes", a finite set A of "arcs", and an incidence matrix $E = (e_{ij})$, where

$$e_{ij} = \begin{cases} 1 & \text{if } i \text{ is the initial node of arc } j \\ -1 & \text{if } i \text{ is the terminal node of arc } j \\ 0 & \text{otherwise} \end{cases}$$

Let us suppose for each arc $j \in A$, there is a polyhedral convex set $K_j \subset R^d$. A flow in G is a function $x: A \to R^d$. We will call x a circulation, and write $x \in C$, if it is conservative at each node, that is,

$$\sum_{j \in A} e_{ij} x(j) = 0 \quad \text{for all } i \in N.$$

The set C is a subspace of the vector space of all flows in G. Define $K = \prod_{j \in A} K_j$; in other words, K is the set of all flows x such that $x(j) \in K_j$ for all $j \in A$. The feasible circulation problem is

(P) $\qquad$ to find $x \in K \cap C$.

This is a problem of finding a point in the intersection of a polyhedral convex set and a subspace. Applying Theorem 7, we obtain

<u>Theorem 10</u>. If $\mathrm{int}(K) \cap C \neq \emptyset$, any sequence $x_{n+1} = (\pi_C * \pi_K)(x_n)$ terminates in a finite number of iterations with a flow x_m such that $\pi_C(x_m) \in K \cap C$.

The regularity condition $\mathrm{int}(K) \cap C \neq \emptyset$ means there exists a circulation x such that $x(j) \in \mathrm{int}\ K_j$ for all j. For each $a: A \to R^d$ (a may be regarded as a point in $R^{d|A|}$), define the perturbed problem

(P_a) $\qquad$ to find $x \in (K+a) \cap C$.

Thus $(P_0) = (P)$. The regularity condition can be expected to hold for most problems in the following sense:

<u>Theorem 11</u>. Suppose $\mathrm{int}\ K_j \neq \emptyset$ for all j. The set of all parameter values a which fail to satisfy one of the following conditions

i. $\mathrm{int}(K+a) \cap C \neq \emptyset$ (so (P_a) satisfies the hypothesis for Theorem 10;

ii. $(K+a) \cap C = \emptyset$ (so (P_a) is infeasible);

forms the boundary of a convex subset $R^{d|A|}$ and is thus a set of measure zero.

Theorem 11 means that, except for a belonging to a small set, either the perturbed problem (P_a) is solvable in a finite number of steps by the algorithm, or (P_a) has no solution and the iterates will diverge.

4. Linear Programming Duality. As a final application, we proved in [1] that if A is a real $d \times d$ antisymmetric matrix, there exists a vector $x \geq 0$ such that $Ax \geq 0$ and $x+Ax > 0$. While this is not a new result, our proof is novel, and provides an elegant application of our results on foldings.

D. The Simple Iterates of a Folding -- Summary of the Main Results from [2]

In [2], we proved several properties of foldings, the first being

Theorem 12. The folds of a folding q are closed, convex and have nonempty interior.

Folds are not necessarily disjoint, but their interiors cannot overlap, as the following shows:

Theorem 13. If P and Q are distinct folds then $P \cap \mathrm{int}(Q) = \emptyset$.

The fact that folds are convex, have nonoverlapping interiors, and their union is the whole space implies that they are convex polyhedral sets:

Theorem 14. Folds are polyhedral convex sets.

In the next theorem, it was shown that the fixed point set of a folding has a very special structure; it is the intersection of an affine flat and a fold.

<u>Theorem 15</u>. Let q: $R^d \to R^d$ be a folding with nonempty fixed point set F. Then there is an affine flat A and a fold Q such that $F = A \cap Q$.

The following is a characterization of linear isometries:

<u>Theorem 16</u>. Let q: $R^d \to R^d$ be positively homogeneous ($q(tx) = tq(x)$ for $t \geq 0$) and a folding. Then q is a (linear) isometry if, and only if, its fixed point set F is a subspace. If q is a linearly isometry, then q maps $F^{\perp}$ onto $F^{\perp}$.

We have seen already that the fixed point set of a folding is the intersection of a fold and a flat. The next result shows that the flat must meet the interior of the fold.

<u>Theorem 17</u>. Let q: $R^d \to R^d$ be a folding. If q has a fixed point, there is a unique fold P whose interior contains a fixed point.

If q: $R^d \to R^d$, we define the <u>cluster set</u> of q to be the set S of all cluster points of q-sequences $x, q(x), \ldots$

If q is nonexpansive and has a fixed point, then every q-sequence has a cluster point. On the other hand, if the cluster set of q is nonempty, Theorem 18 shows that q must have a fixed point. If q is nonexpansive and has a fixed point, Theorem 19 demonstrates that S is closed, convex, $q(S) = S$, and q is distance preserving on S. Also, all q-sequences converge to S in the sense that $\mathrm{dist}(q^k(x),S) \to 0$ for all $x \in R^d$. Thoerem 20 gives a sufficient condition guaranteeing that a folding q have the property: for every x there is some finite k such that $q^k(x) \in S$. Theorem 21 asserts the necessity of this same condition.

<u>Theorem 18</u>. A nonexpansive mapping $q\colon R^d \to R^d$ possesses a fixed point if and only if its cluster set is nonempty.

<u>Theorem 19</u>. Let $q\colon R^d \to R^d$ be a nonexpansive mapping with a nonempty fixed point set F. Then

i. for any $z \in S$, z is a cluster point of the sequence $z, q(z), q^2(z), \ldots$

ii. $q(S) = S$

iii. q is an isometry on S

iv. S is closed and convex

v. for any $x \in R^d$, and $\varepsilon > 0$, there exists K such that $\mathrm{dist}(q^k(x),S) < \varepsilon$ for all $k \geq K$.

If q is a folding, and P is the unique fold whose interior meets F, then also

vi. $F \cap \mathrm{int}(S)$ is nonempty

vii. $S \subset P$.

Theorem 20. Let $q: R^d \to R^d$ be a folding having a nonempty fixed point set F, P the fold whose interior meets F, i the isometry that agrees with q on P, S the set of all cluster points of q-sequences, and W the linear subspace parallel to the flat consisting of all points having periodic orbit under i. If $P = P + W^{\perp}$, then for every $x \in R^d$, there is some k such that $q^k(x) \in S$.

Theorem 21. Let q be a folding on R^d having a nonempty fixed point set F. Let P be the fold whose interior meets F, S the set of all cluster points of q-sequences. Let i denote the isometry that agrees with q on P, and W the subspace parallel to the flat consisting of all points having periodic orbit under i. If $P \neq P + W^{\perp}$, then there exists $y \notin S$ such that $q^k(y) \notin S$ for all k.

II. Solution of Network Equilibrium Problems

In a forthcoming paper [3], a new solution approach will be proposed for a certain class of network equilibrium problems. Applications to the computation of spatial economic equilibria will also be discussed.

The simplest problem to be considered will be the single-location supply-demand equilibrium problem. Here, the goal is to balance the supply and demand of a vector of consumable goods in a competitive market environment. Given a demand vector d, the suppliers are assumed to choose a production plan x so as to solve a certain convex programming problem

$$S(d) \quad \text{to minimize } f_0(x)$$
$$\text{subject to } f_i(x) + d_i \leq 0, \quad i = 1,\ldots,m,\ x \in C.$$

Solution of this problem also yields a vector $P(d)$ of shadow prices (Kuhn-Tucker vectors). The consumers are represented by a demand function $Q(p)$ which specifies the quantities of goods consumers are willing to purchase at unit prices p. The problem is then

$$\text{to determine } p \text{ and } d \text{ such that } p \in P(d) \text{ and } q \in Q(p).$$

This type of formulation is found in the PIES model [18] and others. The solution approach we will propose has some strong theoretical advantages over existing methods, the main one being global convergence under the assumption that the multifunction $-Q$ be maximal monotone. Explicit knowledge of the function $P(d)$ is not required; each iteration requires the solution of a convex programming problem for which the only constraint is $x \in C$. Not only are sequences produced converging to equilibrium values of p and d, but a sequence x_k is produced as well converging to an optimal value for x. Each iteration requires also one evaluation of the function $(I-Q)^{-1}$ which is known to be single-valued because of the fact that $-Q$ is maximal monotone.

More generally, our model will incorporate the spatial or multi-location equilibrium problem. Under this model, there are a finite number of countries or locations, each one having its own demand function $Q_j(p_j)$ and its own convex programming problem $S_j(d_j)$. Goods may be shipped between locations, and when this is done a certain transportation cost is incurred. It is required that the p_j are shadow prices for the associated convex programming problems $S_j(d_j)$, $q_j \in Q_j(p_j)$, and the markets are balanced in the sense that no goods are shipped in a manner that does not make sense for either producers or consumers.

The method of solution proposed here is a specialization to a network setting of the "method of partial inverses" introduced in [1]. Given a maximal monotone multifunction $T: H \rightrightarrows H$ on a Hilbert space H (which for our purposes will always be a finite dimensional Euclidean space equipped with the standard inner product) and a closed subspace $A \subset H$, the method of partial inverses is a procedure for solving the following problem (taking $B = A^{\perp}$)

(5) $\qquad$ to find $x \in A$ and $y \in B$ such that $y \in T(x)$.

The projection of x onto the subspace A or B will be denoted as x_A or x_B, respectively. To solve (5), the method of partial inverses constructs sequences $x_n \in A$ and $y_n \in B$ in such a way that

(6) $$x_{n+1} = (x_n')_A \quad \text{and} \quad y_{n+1} = (y_n')_B,$$
$$\text{where } x_n' \text{ and } y_n' \text{ are chosen so that}$$
$$x_n' + y_n' = x_n + y_n \quad \text{and} \quad y_n' \in T(x_n').$$

The existence and uniqueness of x_n' and y_n' having the above property is established in [1]. The main result from [1] regarding the convergence of the algorithm is the following:

Theorem 22. Let x_k and y_k be sequences of iterates produced by the method of partial inverses. It will always happen either that

i. $x_k \to \bar{x}$ and $y_k \to \bar{y}$ for some solution $\bar{x},\bar{y}$, or that
ii. $|x_k + y_k| \to \infty$ and (5) has no solutions.

The distance from $x_k + y_k$ to the set $\{x+y\colon x,y \text{ solves } (5)\}$ is nonincreasing.

A network G is a triple (N,A,e). The finite sets N and A consist of the nodes and arcs of G, and the incidence function e maps $N \times A$ into $\{+1,-1,0\}$, where

$$e(i,j) = \begin{cases} +1 & \text{if i is the "initial" node of arc j} \\ -1 & \text{if is the "terminal" node of arc j} \\ 0 & \text{otherwise} \end{cases}$$

Each arc is required to have exactly one initial and one terminal node, and these must be distinct. Let E be the $k \times n$ matrix whose (i,j)-entry is $e(i,j)$ ($k = |N|$ and $n = |A|$).

A flow in G is a function $x\colon A \to R^m$. If $m > 1$, x is a "multicommodity" flow. The divergence $y = \text{div } x$ of the flow x is the function $y\colon N \to H$ defined for each node i by

$$y(i) = \sum_{j \in A} e(i,j)x(j).$$

In matrix notation, we can write $y = Ex = \text{div } x$. If $\text{div } x = 0$, we say that x is a <u>circulation</u> in G. The set C of all circulations in G is a subspace of the vector space flows in G. A <u>potential</u> in G is a function $u: N \to H$. A potential determines, in a natural way, a function $v = \Delta u: A \to H$ called the <u>tension function</u> on A which is the <u>differential</u> of u. If i is the initial and i' the terminal node of arc j, then

$$\Delta u(j) = v(j) = u(i') - u(i),$$

or, in matrix form, $v = \Delta u = -E'u$. The tension v is called a <u>differential</u> if $v = \Delta u$ for some potential u. The set D of all differentials in G is also a subspace of the space of flows in G.

Since D is the range of E' and C is the kernel of E, we have the important relationship

$$C = D^{\perp}, \quad C^{\perp} = D.$$

An arbitrary flow x can thus be written in a unique way as the sum of a circulation and a differential:

$$x = x_C + x_D.$$

Both of the economic equilibrium problems mentioned above can be phrased in a network setting. Because the supply and demand functions arise in different manners, it turns out to be convenient to consider networks that contain two classes of arcs. Suppose that the set J of arcs of G is divided into two distinct classes J_1 and J_2. With each arc $j \in J_1$ let there be an associated maximal monotone operator $T_j\colon R^m \rightrightarrows R^m$. With each arc $j \in J_2$, let there be an associated family (P_{u_j}) of optimization problems, where, for each $u_j = (u_{1j},\ldots,u_{mj}) \in R^m$, the problem (P_{u_j}) is

(P_{u_j}) to minimize $f_{0j}(x_j)$ over $x_j \in R^{d_j}$
subject to the constraints $x_j \in C_j$ and
$f_{1j}(x_j) + u_{1j} \le 0,\ldots,f_{mj}(x_j) + u_{mj} \le 0$

where the functions $f_{ij}\colon R^{d_j} \to R$ are convex and $C_j \subset R^{d_j}$ is a nonempty closed convex set. Given such a network, it makes sense to consider the following problem (taking $d_j = 0$ for $j \in J_1$, and $d = \Sigma d_j$):

(Q) to find $u = \pi u_j \in C$, $y = \pi y_j \in D$, and $x = \pi x_j \in R^d$ such that

(a) for each $j \in J_1$, $y_j \in T_j(u_j)$ and

(b) for each $j \in J_2$, x_j solves the problem (P_{u_j}) and y_j is a Kuhn-Tucker vector.

To say that y_j is a Kuhn-Tucker vector for (P_{u_j}) means that

$y_{1j},\ldots,y_{mj} \ge 0$ and the infimum of the function
$f_{0j}(x_j') + \sum_i y_{ij}(f_{ij}(x_j') + u_{ij})$ over all $x_j' \in C_j$
is finite and equal to the optimal value in (P_{u_j}).

For each $j \in J_2$, define

$$F_j(x_j,u_j) = \begin{cases} f_{0j}(x_j) & \text{if } f_{1j}(x_j)+u_{1j} \le 0,\ldots,f_{mj}(x_j)+u_{mj} \le 0 \\ & \quad \text{and } x_j \in C_j \\ +\infty & \text{otherwise} \end{cases}$$

and let $T_j = \partial F_j$.

Let $H = \{(x,u): x \in R^d$ and u is a flow in $G\}$. Let $A = \{(x,u) \in H: u \in C\}$ and $B = \{(v,y) \in H: v = 0$ and $y \in D\}$. It is clear that A and B are complementary subspaces of H. Define a maximal monotone multifunction $T: H \rightrightarrows H$ by declaring

$(v,y) \in T(x,u)$ if, and only if $y_j \in T_j(u_j)$ for all $j \in J_1$ and $(v_j,y_j) \in T_j(x_j,u_j)$ for all $j \in J_2$.

The network equilibrium problem (Q) is then equivalent to the problem

to find $(x,u) \in A$ and $(v,y) \in B$ such that $(v,y) \in T(x,u)$.

This problem can be solved by the method of partial inverses (6). In order that it be possible to implement this procedure, we know that it would be required that routines be available to perform each of the following tasks:

(1) given $(x,u) \in H$, to compute the projection of (x,u) onto A and B, and

(2) given $(x,u) \in A$ and $(0,y) \in B$, to determine (x',u')

and (v',y') such that

$$(x'+v',u'+y') = (x,u+y)$$

and

$$(v',y') \in T(x',u').$$

Procedures for performing each of these tasks will be investigated in [3]. The resulting algorithm for solving (Q) is the following:

Initialization: Start with an arbitrary flow x, $u \in C$, $y \in D$.

Step 1: (a) For each arc $j \in J_1$, find u_j' and y_j' such that

$$u_j' + y_j' = u_j + y_j \quad \text{and } y_j' \in T_j(u_j')$$

(b) For each arc $j \in J_2$, find x_j' to minimize the function

$$f_{0j}(x_j') + \frac{1}{2}\Sigma|x_{ij}'-x_{ij}|^2 + \frac{1}{2}\Sigma\min^2\{0,f_{ij}(x_j') + u_{ij} + y_{ij}\}$$

subject to the constraint $x_j \in C_j$, and let

$$u_{ij}' = \begin{cases} u_{ij} + y_{ij} & \text{if } f_{ij}(x_j') + u_{ij} + y_{ij} \geq 0 \\ -f_{ij}(x_j') & \text{if } f_{ij}(x_j') + u_{ij} + y_{ij} \leq 0 \end{cases}$$

and

$$y_{ij}' = u_{ij} + y_{ij} - u_{ij}'$$

Step 2: Update x, u, and y as follows:

$$x^+ = x', \quad u^+ = (u')_C, \quad y^+ = (y')_D$$

and repeat Step 1.

According to Theorem 22, we also have

<u>Theorem 23</u>. Let x_k, u_k, and y_k be iterates produced by the above algorithm. It will always happen either that

i. $x_k \to \bar{x}$, $u_k \to \bar{u}$, and $y_k \to \bar{y}$ for some solution $(\bar{x},\bar{u},\bar{y})$ to (Q),

or that

ii. $|(x_k,u_k+y_k)| \to \infty$ and (Q) has no solution.

References

Papers containing results of this project:

1. J. E. Spingarn and J. Lawrence, "On fixed points of nonexpansive piecewise isometric mappings," to appear in Proc. London Math. Soc.

2. J. E. Spingarn and J. Lawrence, "On iterates of non-expansive mappings and foldings," submitted.

3. J. E. Spingarn, "On the computation of spatial economics equilibria," in progress.

Papers by the principal investigator closely related to this project:

4. "Partial inverse of a monotone operator," Applied Mathematics and Optimization 10 (1983) 247-265.

5. "A primal-dual projection method for solving systems of linear inequalities," Linear Algebra and Its Applications, 65: 45-62 (1985).

6. "A projection method for least square solutions to overdetermined systems of linear inequalities," accepted by Linear Algebra and Its Applications.

7. "Applications of the method of partial inverses to convex programming: decomposition," Mathematical Programming, 32 (1985) 199-223.

Other papers referenced in this report:

8. M. A. Krasnoselski, Two observations about the method of successive approximations, Uspehi. Math. Nauk. 10 (1955) 123-127.

9. B. Martinet, Détermination approchée d'un point fixe d'une application pseudo-contractante. Cas de l'application prox, C. R. Acad. Sci. Paris (Ser. A) 274 (1972) 163-165.

10. J.-J. Moreau, Proximité et dualité dans un espace Hilbertian, Bull. Soc. Math. France 93 (1965) 273-299.

11. R. T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877-898.

12. M. D. Kirszbraun, Über die zusammenziehende and Lipschitzsche transformationen, Fundamenta Mathematicae 22 (1934) 77-108.

13. G. J. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Math. J. 29 (1962) 341-346.

14. R. T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math. 33 (1970) 209-216.

15. R. T. Rockafellar, Augmented Lagrangians and the proximal point algorithm in convex programming, Math. of O. R. 1 (1976) 97-116.

16. S. Agmon, The relaxation method for linear inequalities, Canad. J. Math. 6 (1954) 382-392.

17. T. S. Motzkin, I. J. Schoenberg, The relaxation method for linear inequalities, Canad. J. Math. 6 (1954) 393-404.

18. B. Ahn, Computation of Market Equilibria for Policy Analysis: The Project Independence Evaluation System Approach, Garland Publ., New York 1979.