Institute of Socio-Economic Planning
Discussion Paper Series

No.158 (82-25)

VARIABLE DIMENSION ALGORITHMS
AND
A CLASS OF PRIMAL-DUAL SUBDIVIDED MANIFOLDS†

Masakazu Kojima*

Yoshitsugu Yamamoto**

August 1982

† To appear in Proceedings of the 3rd Mathematical Programming Symposium, Japan, October 1982.

* Department of Information Sciences, Tokyo Institute of Technology
Oh-Okayama, Meguro, Tokyo 152, Japan

** Institute of Socio-Economic Planning, University of Tsukuba
Sakura, Ibaraki 305, Japan

University of Tsukuba
Sakura, Ibaraki 305
Japan

VARIABLE DIMENSION ALGORITHMS

AND

A CLASS OF PRIMAL-DUAL SUBDIVIDED MANIFOLDS

Masakazu Kojima

Department of Information Sciences, Tokyo Institute of Technology

Yoshitsugu Yamamoto

Institute of Socio-Economic Planning, University of Tsukuba

Abstract We present a brief interpretation of variable dimension algorithms for solving systems of nonlinear equations. The interpretaion is based on a basic model consisting of a one-parameter family of systems of equations and a class of primal-dual subdivided maifolds.

1. Introduction

The purpose of this paper is to give a brief interpretation of variable dimension algorithms (abbreviated by vd algorithms) which were originally proposed for computing fixed points on an n-simplex by Van der Laan and Talman [8] and later extended to solving a system of equations

(1) $f(x) = 0, \qquad x \in R^n$

(Van der Laan and Talman [9,10,11], Todd [15], Todd and Wright [16], Kojima and Yamamoto [6], Saigal [14], Yamamoto [18], etc.). Here f is a continuous map from the n-dimensional Euclidean space R^n into itself. Throughout the paper, we assume that $f: R^n \to R^n$ is C^1

(continuously differentiable) and restrict our attention to "piecewise smooth vd algorithms" (Kojima [5]). See Kojima and Yamamoto [6,7] for "simplicial vd algorithms".

Let us begin with the Newton (-Raphson) method, one of the most popular method for solving systems of equations. Let x^0 be an arbitrary point of R^n or a rough approximate solution of (1) if it is known. Then the iteration of the Newton method is defined by

$$x^{p+1} = x^p - Df(x^p)^{-1}f(x^p) \qquad (p = 0, 1, \dots),$$

where $Df(x)$ is the Jacobian matrix of f at x. It is well-known that if the initial point x^0 is sufficiently close to a non-degenerate solution x^* of (1), i.e. an $x^* \in R^n$ such that $f(x^*) = 0$ and $\det Df(x^*) \neq 0$, then the sequence $\{x^p\}$ converges to x^* quadratically. When the initial point x^0 is far from x^*, however, the convergence is not guaranteed and the Newton method may fail. In other words, the Newton method does not enjoy the global convergence property. "Continuation" is one of the techniques developed for overcoming the lack of this nice property (See, for example, Broyden [2], Ortega and Rheinboldt [13], Iri [4]).

In the followings we assume for simplicity that $x^0 = 0$ and consider the family of systems of equations

$$(2) \qquad f(x) - (1 - s) f(0) = 0, \quad 0 \le s \le 1, \ x \in R^n,$$

where s is a parameter. As s increases from 0 to 1, the system (2) is continuously deformed from the system

$$f(x) - f(0) = 0$$

having a trivial solution $x^0 = 0$ into the system (1) which we want to solve. Thus, starting from the known solution x^0 of (2) with the parameter $s = 0$, we trace solutions of the system (2) until the parameter s hits 1. Then we obtain a solution of the system (1). This is a basic idea of the continuation methods. In many practical problems for which the Newton method fails the continuation methods often succeed with the same initial point.

We now convert the system (2) into the system

(3) $\qquad - f(0) + t\, f(x) = 0, \quad 1 \le t, \ x \in R^n.$

Suppose $0 \le s < 1$. Then $(x, s) \in R^{n+1}$ satisfies (2) if and only if $(x, 1/(1-s)) \in R^{n+1}$ satisfies (3). Specifically, the known solution $(0, 0) \in R^{n+1}$ of (2) is corresponding to the solution $(0, 1) \in R^{n+1}$ of (3). Let S denote the connected component of the solution set of (3) which contains $(0, 1) \in R^{n+1}$. Since the system (3) consists of n scalar equations in $n+1$ variables, it is intuitively clear that S forms a one dimensional curve, which we shall call a path. If the path S "converges" to $(x^*, +\infty)$ for some $x^* \in R^n$, x^* is a solution of (1). In this case a point $\bar{x} \in R^n$ which solves the system (3) with a sufficiently large $\bar{t}$ satisfies

$$\| f(\bar{x}) \| = (1/\bar{t}) \| f(0) \| \fallingdotseq 0;$$

$\bar{x}$ is an approximate solution of (1). Then we may apply the Newton method to the system (1) and the initial point $\bar{x}$ if an approximate solution with higher accuracy is needed.

Let

$$S_x = \{ x \in R^n : (x, t) \in S \ \text{ for some } \ t \in [1, +\infty) \}.$$

Since the path S winds through R^{n+1}, S_x also winds through R^n toward x^*. Generally, we are forced to consume much work to trace a nonlinear path in a higher dimensional space. So, if S_x could lead to a solution x^* of (1) along a straight line, it would be ideal. The system (3) can be rewritten as follows by introducing a redundant variable vector y:

(4) $\quad y + t\, f(x) = 0, \quad (x, y, t) \in L \times T.$

Here

$$L = \{ (x, y) \in R^{2n}: \ x \in R^n, \ \ y = - f(0) \} ,$$

$$T = [1, +\infty),$$

i.e. the variable vector x and the variable t are not restricted while the variable vector y is fixed to $- f(0)$. This combination furnishes one dimensional freedom to the set of solutions of (4). Of course, there are many combinations which furnishe one dimensional freedom. For instance, for given k-dimensional subspace $X \subset R^n$ and (n-k)-dimensional subset $Y \subset R^n$, we may replace L by

$$L = X \times Y = \{ (x, y) \in R^{2n}: \ x \in X, \ \ y \in Y \}.$$

It is needless to say that we cannot reach a solution of (1) unless it happens to lie on the subspace X. But since y plays as a slack variable vector in (4), the work required for tracing the path S of solutions is greatly reduced if k is small. The aim of vd algorithms is to approach a solution of (1) while keeping the variable vector x in lower dimensional subspaces as long as possible. At final stages, however, we have to let the variable vector x freely move around in R^n to approximate a solution whose location is not known beforehand. Thus the dimension of subspaces on which x moves must vary along the path S. In the next section we shall illustrate how to realize such a variable dimension structure by presenting an example.

2. Variable Dimension Structure

Figure 1 shows two subdivided manifolds (see Section 3 for the definition) of R^2.

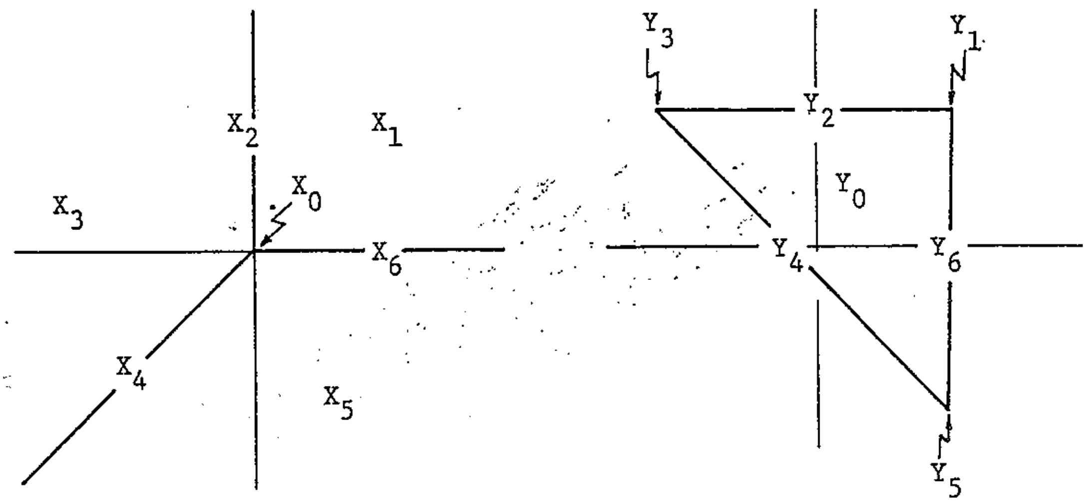
Let

$$\mathcal{P} = \{ X_1, X_3, X_5 \}, \qquad \bar{\mathcal{P}} = \{ X_0, X_1, \ldots, X_6 \}$$

$$\mathcal{D} = \{ Y_0 \}, \qquad \bar{\mathcal{D}} = \{ Y_0, Y_1, \ldots, Y_6 \}.$$

Note that $\dim X_j + \dim Y_j = 2$ for $j = 0, 1, \ldots, 6$.

Example 1.



$X_0 = \{0\}$

X_1, X_3, X_5 : two dimensional polyhedral cones

X_2, X_4, X_6 : half lines

Y_0 : a triangle having 0 in its interior

Y_1, Y_3, Y_5 : vertices of Y_0

Y_2, Y_4, Y_6 : facets of Y_0

Figure 1.

Let

$$L = \bigcup_{j=0}^{6} (X_j \times Y_j).$$

Then L forms a two dimensional piecewise linear manifold (the definition will be given in Section 3) in R^4. Figure 2 illustrates a combinatorial structure of L, i.e. an adjacency relation among two dimensional polyhedral sets $X_j \times Y_j$'s.

Now we consider the system (4) of equations with L defined above and $T = [0, +\infty)$. Obviously, $z^0 = (0, 0, 0) \in X_0 \times Y_0 \times T$ satisfies (4). Let S be the connected component of the solution set of (4) which contains z^0. Since the system (4) consinsts of two scalar equations on the three dimensional piecewise linear manifolds $L \times T$, S forms a piecewise smooth path under a moderate regularity assumption. It runs through the three dimensional manifold $L \times T$ in

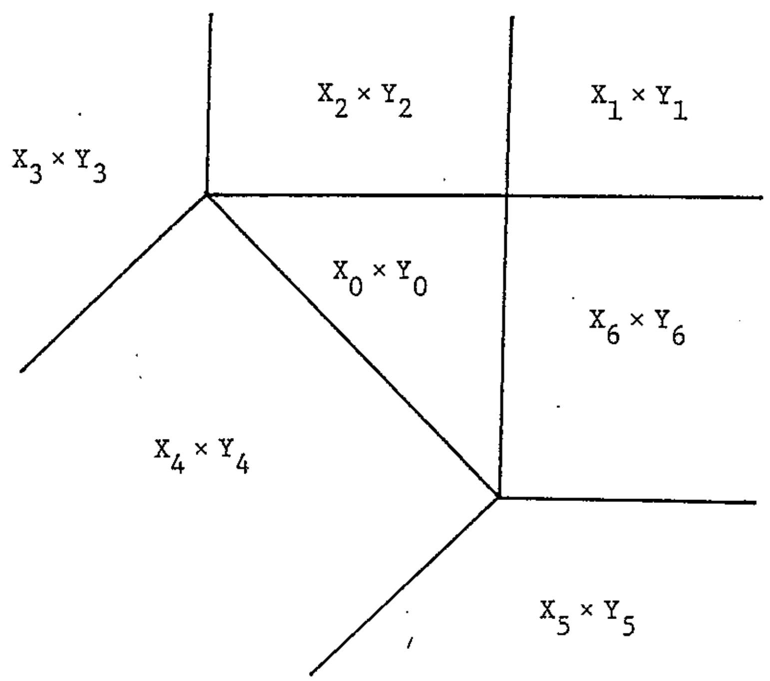


Figure 2. combinatorial structure of L

R^5 all the way from z^0 (S and $L \times T$ are illustrated in Figure 3). But we can observe the variable dimension structure of the path S in the projections S_x and S_y of S on the spaces of the variable vectors x and y. See Figure 4. We see that S_x leads to a solution of (1) along several faces of P with varying dimensions. The system (4) with L and T defined above is decomposed into the family of systems of equations:

$$y + t\, f(x) = 0,\ (x, y, t) \in X_j \times Y_j \times T \quad (j = 0, 1, \ldots, 6).$$

Suppose S moves from a three dimensional polyhedral set $X_i \times Y_i \times T$ into another one $X_j \times Y_j \times T$ traversing their common two dimensional face. Then by the construction of L, either

X_i is a facet of X_j and Y_j is a facet of Y_i

or

X_j is a facet of X_i and Y_i is a facet of Y_j

occurs. In this way the dimensions of X_i and Y_i vary complementarily by unity as the path S moves into a new polyhedral set. Observe that the variable vector x moves along lower dimensional faces of P in earlier several stages of tracing the path S. As stated in Section 1, y serves as a slack variable vector in (4), and hence the computation of S is relatively easy while x moves on lower dimensional faces.

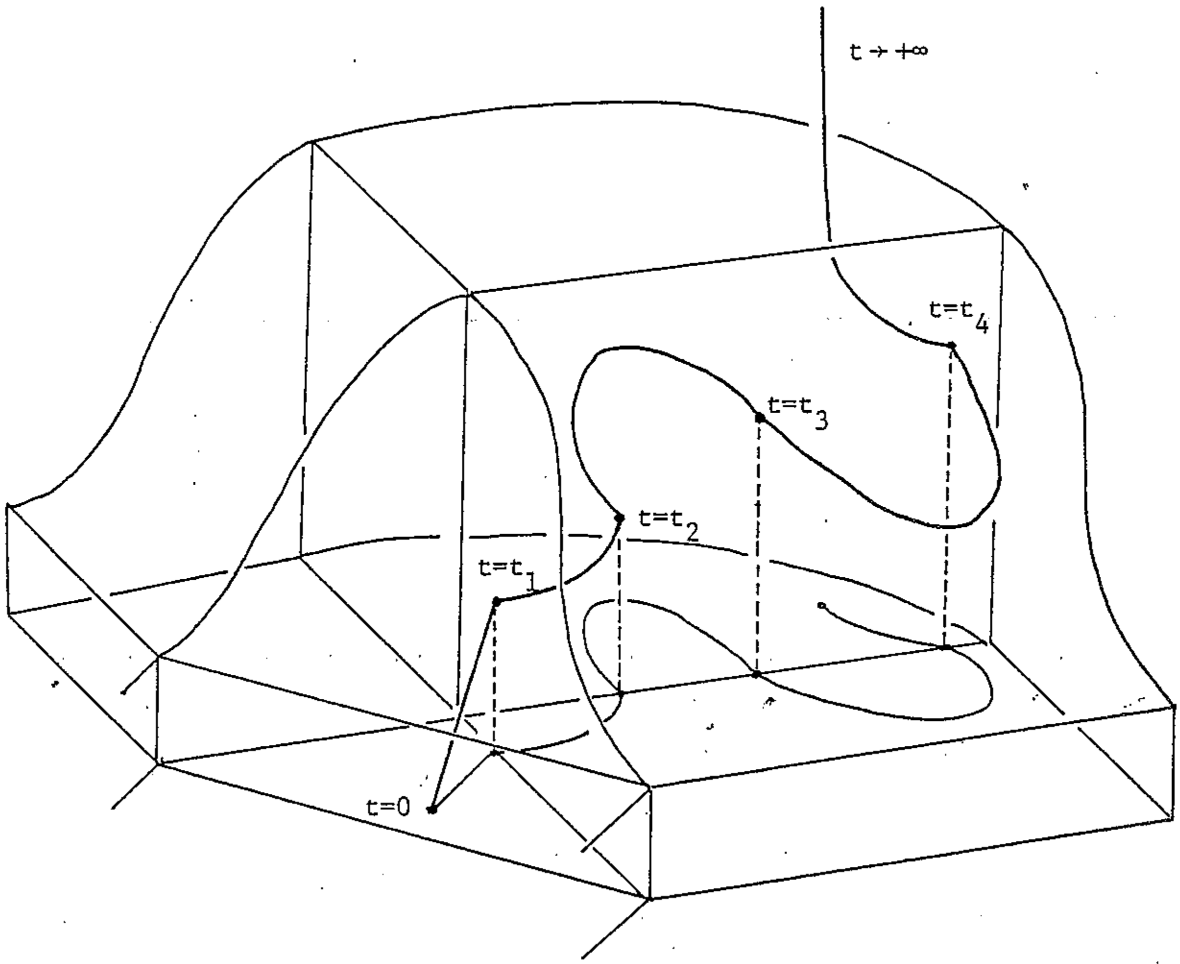


Figure 3. a segment of the manifold $L \times T$ and the path S

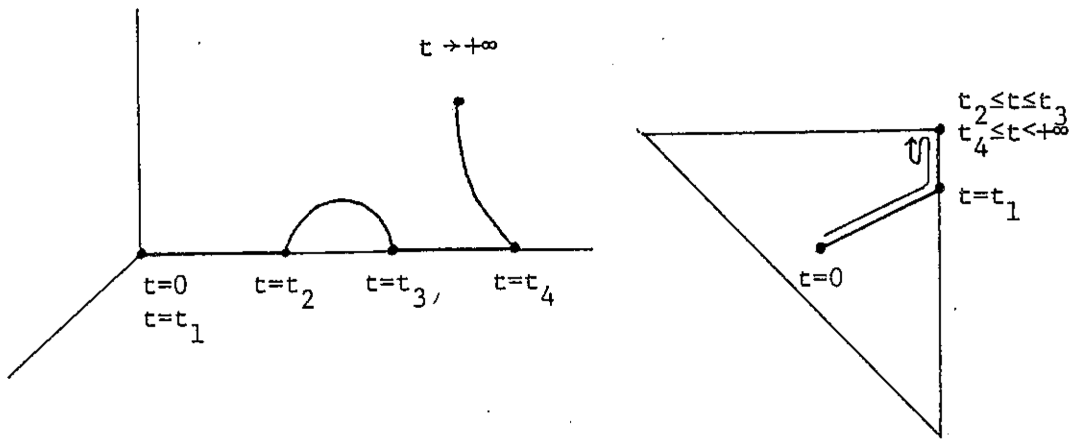


Figure 4. projections S_x and S_y

3. A Class of Primal-Dual Subdivided Manifolds

To give a precise description of subdivided manifolds which compose the variable dimension structure we shall introduce some definitions. We call an m-dimensional convex polyhedral set C in R^k a <u>cell</u> or an m-cell. Let B be a cell such that $B \subset C$. If

$$\lambda x + (1-\lambda) y \in B$$

$$x, y \in C, \quad 0 < \lambda < 1$$

always imply that $x, y \in B$, then we say that B is a <u>face</u> of C and write $B < C$.

Let $\mathcal{M}$ be a finite or countable collection of m-cells in R^k. Let

$$\bar{\mathcal{M}} = \{ B : B < C \text{ for some } C \in \mathcal{M} \} ,$$

$$|\mathcal{M}| = \bigcup \{ C : C \in \mathcal{M} \} .$$

$\mathcal{M}$ is a <u>subdivided</u> m-<u>manifold</u> if it satisfies the following three conditions:

(i) For each pair B, C $\in \mathcal{M}$, either $B \cap C = \phi$ or $B \cap C$ is a common face of B and C.

(ii) Each (m-1)-cell of $\bar{\mathcal{M}}$ lies in at most two m-cells of $\mathcal{M}$.

(iii) $\mathcal{M}$ is locally finite; each point $x \in |\mathcal{M}|$ has a neighborhood that intersects with finitely many m-cells of $\mathcal{M}$.

Define the <u>boundary</u> $\partial\mathcal{M}$ of a subdivided m-manifold $\mathcal{M}$ by

$$\partial\mathcal{M} = \{ B : B \text{ is an (m-1)-cell of } \bar{\mathcal{M}} \text{ that lies in exactly one m-cell of } \mathcal{M} \} .$$

Let M be a subset of R^k. If $M = |\mathcal{M}|$ for some subdivided m-manifold $\mathcal{M}$, then M is an m-dimensional <u>piecewise</u> <u>linear</u> <u>manifold</u>, or simply, an m-<u>manifold</u> and $\mathcal{M}$ is a <u>subdivision</u> of M.

Now we are ready to describe a primal-dual pair of subdivided manifolds (abbriviated by PDM). A triplet $(\mathcal{P}, \mathcal{D}; d)$ is a PDM if it satisfies the following conditions:

(i) $\mathcal{P}$ is a subdivision of R^n.

(ii) $\mathcal{D}$ is a subdivision of a polyhedral subset D of R^n such that each cell of $\mathcal{D}$ is bounded.

(iii) d is an operator from $\bar{\mathcal{P}} \cup \bar{\mathcal{D}}$ into itself such that $X^d \in \bar{\mathcal{D}}$ for every $X \in \bar{\mathcal{P}}$ and $Y^d \in \bar{\mathcal{P}}$ for every $Y \in \bar{\mathcal{D}}$.

(iv) If $Z \in \bar{\mathcal{P}} \cup \bar{\mathcal{D}}$ then $(Z^d)^d = Z$ and $\dim Z + \dim Z^d = n$.

(v) If $X_1, X_2 \in \bar{\mathcal{P}}$ and $X_1 < X_2$, then $X_2^d < X_1^d$.

(v)' If $Y_1, Y_2 \in \bar{\mathcal{D}}$ and $Y_1 < Y_2$, then $Y_2^d < Y_1^d$.

See Section 3 of [6] for the general definition of PDM and its fundamental properties. We call $\mathcal{P}$ (resp. $\mathcal{D}$) the primal (resp. dual) subdivided manifold, d the dual operator and Z^d the dual of Z for each $Z \in \bar{\mathcal{P}} \cup \bar{\mathcal{D}}$. Note that the condition (iv) implies that the dual operator d is one-to-one and onto, and that its inverse is d itself. It is easily seen that the two subdivided manifolds in Example 1 form a PDM by defining the dual operator d by $X_j^d = Y_j$ and $Y_j^d = X_j$ for $j = 0, 1, \ldots, 6$.

Define

$$< \mathcal{P}, \mathcal{D}; d > = \{ X \times X^d : X \in \bar{\mathcal{P}} \},$$

or equivalently

$$< \mathcal{P}, \mathcal{D}; d > = \{ Y^d \times Y : Y \in \bar{\mathcal{D}} \}.$$

Let $\mathcal{L} = < \mathcal{P}, \mathcal{D}; d >$.

Theorem 1. (Theorem 3.1 in Kojima and Yamamoto [7])

L is a subdivided n-manifold with no boundary, i.e., $\partial L = \phi$, and $|L|$ is closed.

We further impose the following two conditions on the PDM (P , $\mathcal{D}$; d):

(vi) $\{0\} \in \bar{P}$, and $0 \in \text{int } Y_0$, where $Y_0 = \{0\}^d \in \mathcal{D}$.

(vii) There exists a positive number α such that

$$x^t y \geq \alpha \| x \| \quad \text{for every} \quad (x,y) \in |L|,$$

where $\| x \| = (x^t x)^{1/2}$. From the condition (vi), we see that (0, c) lies in $|L|$ whenever $\| c \|$ is sufficiently small. This point (0, c) will be used as an initial point of the vd algorithms. The condition (vii) will play an important role when we derive a sufficient condition for the global convergence of the vd algorithms. It is also easy to see that the PDM in Example 1 satisfies the conditions (vi) and (vii).

We shall give three examples of PDM below. Here we restrict ourselves to picturing them. See Section 3 of Kojima and Yamamoto [7] for their precise descriptions.

The PDM's given in Examples 2, 3 and 4 are used in the 2n-method (Van der Laan and Talman [11]), the 2^n-method (or the octahedral method, Wright [17]) and the (3^n-1)-method (Kojima and Yamamoto [7]). They are all easily verified to satisfy the conditions (vi) and (vii) as well as (i) - (v).

Example 2.

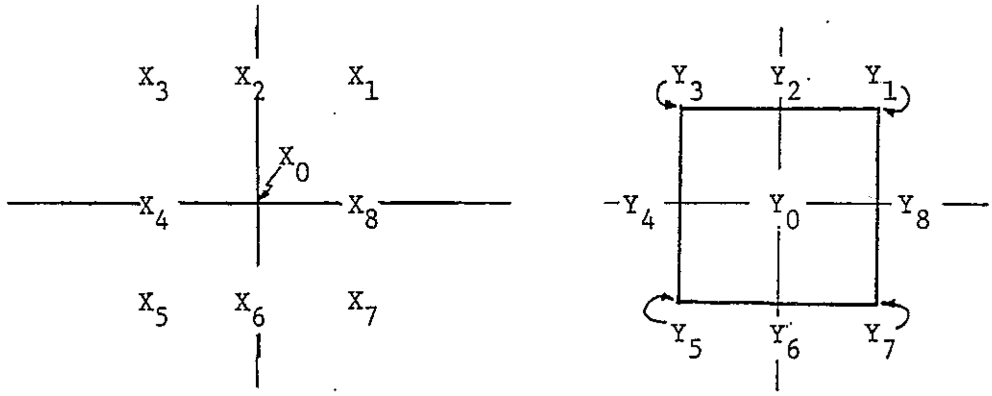


Figure 5.

Example 3.

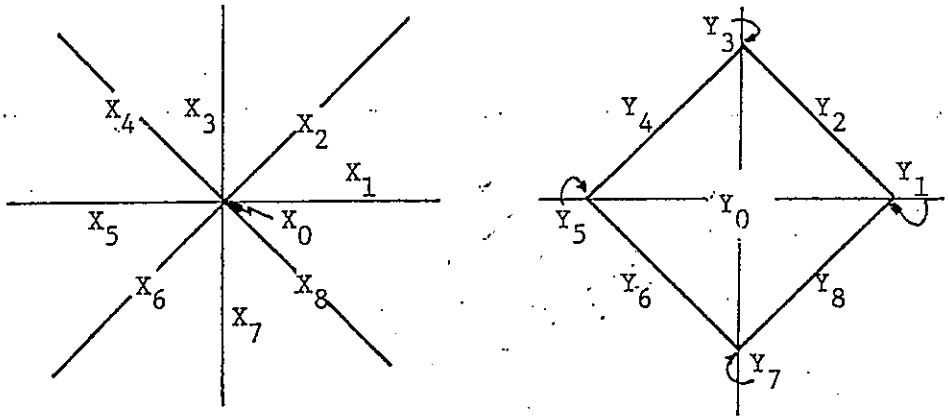


Figure 6.

Example 4.

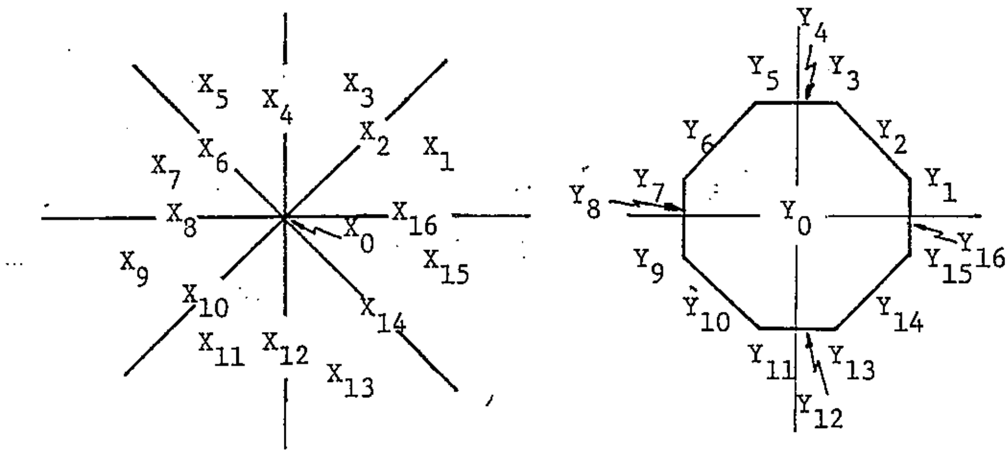


Figure 7.

4. Piecewise Smooth Vd Algorithm

Let $(\mathcal{P}, \mathcal{D}; d)$ be a PDM which satisfies the conditions (i) - (vii) given in Section 3, and let

$$L = \langle \mathcal{P}, \mathcal{D}; d \rangle = \{ X \times X^d : X \in \bar{\mathcal{P}} \}.$$

Define the subdivided (n+1)-manifold

$$M = \{ Z \times T : Z \in L \},$$

where $T = [0, +\infty)$. By Theorem 1, we see

$$(5) \qquad \begin{aligned} \partial M &= \{ Z \times \{0\} : Z \in L \} \\ &= \{ X \times X^d \times \{0\} : X \in \bar{\mathcal{P}} \}. \end{aligned}$$

Now we define the PC^1 (abbreviation of piecewise C^1) map $h : |M| \to R^n$ by

$$h(x, y, t) = y + t\, f(x) \quad \text{for every } (x, y, t) \in |M|.$$

Let c be an interior point of the cell Y_0 of $\mathcal{D}$. By the condition (vi), any point c with sufficiently small norm $\| c \|$ lies in the interior of Y_0. In Section 2 we have taken $c = 0$. We shall consider the system of equations

$$(6) \qquad h(x, y, t) = c, \qquad (x, y, t) \in |M|.$$

It follows from (5) and the condition (vi) in Section 3 that $h^{-1}(c)$ intersects with $|\partial M|$ at a single point $z^0 = (x^0, y^0, t^0) = (0, c, 0)$. This point will serve as an initial point for the piecewise smooth vd algorithm. Let S be a connected component of $h^{-1}(c)$ which contains the initial point z^0. We see that under a moderate regularity assumption $h^{-1}(c)$ is a disjoint union of piecewise smooth paths and loops, and that S is a piecewise smooth path.

Lemma 2. (Lemma 5.3 in Kojima and Yamamoto [7])

Assume the condition (i) - (vi) in Section 3. Then under a moderate regularity assumption S is an unbounded path which is homeomorphic to $[0, +\infty)$.

If there is a bounded subset V of D such that $S \subset R^n \times V \times T$, each $(\bar{x}, \bar{y}, \bar{t}) \in S$ with a sufficiently large $\bar{t}$ satisfies

$$\| f(\bar{x}) \| \le (1/\bar{t}) (\| \bar{y} \| + \| c \|) \fallingdotseq 0;$$

$\bar{x}$ is an approximate solution of (1). In examples in Sections 2 and 3, each polyhedral subset D is bounded in itself, i.e. there is a $\delta > 0$ such that $\| y \| \le \delta$ for any $y \in D$. Therefore to obtain an approximate solution $\bar{x}$ such that $\| f(\bar{x}) \| \le \varepsilon$ we have only to trace the path S until $\bar{t}$ exceeds $(\delta + \| c \|)/\varepsilon$. For tracing the path S numerically, we can employ various predictor-corrector procedures developed in the homotopy continuation methods (Li and Yorke [12], Allgower and Georg [1], Georg [3], etc.). See Kojima [5] for more detail.

Remark When we trace the solution path S of the system (6), we may encounter numerical unstability as the value of t increases. It might be better to employ the system of equations

$$(1-t)\, y + t\, f(x) = 0, \quad (x, y, t) \in |\mathcal{L}| \times [0, 1].$$

In this case we obtain a solution of the system (1) when the variable t attains 1.

The following theorem provides a sufficient condition for the global convergence of the piecewise smooth vd algorithms.

Theorem 3. (Theorem 5.1 in Kojima [5])

In addition to the conditions (i) - (vii), suppose that for some $\mu > 0$ and for every $x \in R^n$ with $\| x \| \geq \mu$ there exists an $\hat{x} \in R^n$ such that

$$\| \hat{x} \| \leq \mu \quad \text{and} \quad (x - \hat{x})^t f(x) > 0$$

(a weaker version of Merrill's condition). Then there exist bounded sets $U \subset R^n$ and $V \subset D$ such that $S \subset U \times V \times T$.

References

[1] E. Allgower and K. Georg, "Simplicial and continuation methods for approximaing fixed points and solutions to systems of equations", SIAM Review 22 (1980) 28-85.

[2] C. G. Broyden, "A new method of solving nonlinear simultaneous equations", Computer Journal 12 (1969) 94-99.

[3] K. Georg, "On tracing an implicitly defined curve by quasi-Newton steps and calculating bifurcation by local perturbations", University of Bonn, Bonn, Jan. 1980.

[4] M. Iri, Suchi Keisan (Asakura, Tokyo, 1981).

[5] M. Kojima, "An introduction to variable dimension algorithms for solving systems of equations", in: E. L. Allgower, K. Glashoff and H. -O. Peitgen, ed. Numerical Solution of Nonlinear Equations, Lecture Notes in Mathematics 878 (Springer, Berin, 1980) pp.199-237.

[6] M. Kojima and Y. Yamamoto, "Variable dimension algorithms: Basic theory, interpretations and extensions of some existing methods", to appear in Mathematical Programming.

[7] M. Kojima and Y. Yamamoto, "A unified approach to several restart fixed point algorithms for their implementation and a new variable dimension algorithm", Discussion Paper Series No.153(82-20), Institute of Socio-Economic Planning, University of Tsukuba, April 1982.

[8] G. van der Laan and A. J. J. Talman, "A restart algorithm for computing fixed point without extra dimension", Mathematical Programming 17 (1979) 74-84.

[9] G. van der Laan and A. J. J. Talman, "A restart algorithm without an artificial level for computing fixed points on unbounded regions", in: H. -O. Peitgen and H. -O. Walther, ed., Functional Differential Equations and Approximation of Fixed Points, Lecture Notes in Mathematics 730 (Springer, Berlin, 1979) pp.247-256.

[10] G. van der Laan and A. J. J. Talman, "Convergence and properties of recent variable dimension algorithms", in: W. Forster, ed., Numerical Solution of Highly Nonlinear Problems, Fixed Point Algorithms and Complementarity Problems (North-Holland, New York, 1980) pp.3-36.

[11] G. van der Laan and A. J. J. Talman, "A class of simplicial restart fixed point algorithms without an extra dimension", Mathematical Programming 20 (1981) 33-48.

[12] T. Y. Li and J. A. Yorke, "A simple reliable numerical algorithm for following homotopy paths", in: S. M. Robinson, ed., Analysis and

Computation of Fixed Points (Academic Press, New York, 1980).

[13] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables (Academic Press, New York, 1970).

[14] R. Saigal, "A homotopy for solving large, sparse and structured fixed point problems", Northwestern University, Evanston, Illinois, Jan. 1981.

[15] M. J. Todd, "Fixed-point algorithms that allow restarting without an extra dimension", Technical Report No.379, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, Sept. 1978.

[16] M. J. Todd and A. H. Wright, "A variable-dimesion simplicial algorithm for antipodal fixed-point theorems", Numerical Functional Analysis and Optimization 2 (1980) 155-186.

[17] A. H. Wright, "The octahedral algorithm, a new simplicial fixed point algorithm", Mathematical Programming 21 (1981) 47-69.

[18] Y. Yamamoto, "A new variable dimension algorithm for the fixed point problem", Discussion Paper Series No.111(81-12), Institute of Socio-Economic Planning, University of Tsukuba, March 1981 (to appear in Mathematical Programming).