DEPARTMENT OF MATHEMATICS
UNIVERSITY OF HOUSTON HOUSTON, TEXAS

NASA CR- 147519

(NASA-CR-147519) SCME RESULTS ON CONTRACTIVE MAPPINGS AS RELATED TO PATTERN RECOGNITION (Houston Univ.) 20 p HC $3.50 CSCL 09B

N76-21953

Unclas
G3/63 21497

SOME RESULTS CN
CONTRACTIVE MAPPINGS AS
RELATED TO PATTERN RECCGNITION
BY D. R. BROWN & M. J. 3'MALLEY
REPORT #49

PREPARED FOR
EARTH OBSERVATION DIVISION, JSC
UNDER
CONTRACT NAS-9-12777

HOUSTON, TEXAS 77004

Some Results on Contractive Mappings as related to Pattern Recognition

by

Dennison R. Brown and Matthew J. O'Malley

Mathematics Department

University of Houston

Houston, Texas

Report #49

October 1975

Some Results on Contractive Mappings as related to Pattern Recognition

Abstract

Several of the presently used techniques in pattern recognition can be reformulated as the problem of determining fixed points of a function of the form: $f(x) = A(x) \cdot x$. If x_0 is a fixed point of f and if f is contractive at x_0, then, for any y belonging to a sufficiently small neighborhood of x_0, the orbit of y will converge to x_0. In this paper we develop several general results regarding contractive mappings and, in particular, we study functions of the form: $f(x) = A(x) \cdot x$.

1. Introduction

Definition (1.1). Let $(X, \|\ \|)$ be a normed linear space and let $f : X \to X$ be a function. We say f is a contractive mapping of X provided there exists a real number α, $0 \le \alpha < 1$, such that $\|f(x) - f(y)\| \le \alpha\|x - y\|$ for all $x, y \in X$.

Definition (1.2). Let S be a set and let $T : S \to S$ be a function. We say that $s_0 \in S$ is a fixed point of T provided $T(s_0) = s_0$.

If $(X, \|\ \|)$ is a normed linear space and if f is a contractive mapping of X having $x_0 \in X$ as a fixed point, then x_0 is the unique fixed point of

f. For, if $x_1, x_2 \in X$ are fixed points of f, then

$$||x_1 - x_2|| = ||f(x_1) - f(x_2)|| \leq \alpha||x_1 - x_2|| .$$

Thus, if $x_1 \neq x_2$, we have that $1 \leq \alpha < 1$, a contradiction.

If we impose the added condition that $(X, || \; ||)$ is complete (that is, a Banach space), then any contractive mapping of X has a fixed point. Precisely stated, we have the following well-known theorem.

<u>Theorem</u> (1.3). Let $(X, || \; ||)$ be a Banach space and let $f : X \to X$ be a contractive mapping of X. Then, for any $z \in X$, the sequence $\{f^n(z)\}_{n=1}^{\infty}$ converges to a fixed point $x_0 \in X$. Moreover, x_0 is the unique fixed point of f.

In many cases it may be that f is not contractive on the whole space X but rather only at certain points of X. We formalize the concept.

<u>Definition</u> (1.4). Let $(X, || \; ||)$ be a normed linear space and let $f : X \to X$ be a function. We say that f is <u>contractive at</u> $x_0 \in X$ if and only if there exist real numbers ϵ and α, $\epsilon > 0$, $0 \leq \alpha < 1$, such that $||f(x) - x_0|| \leq \alpha||x - x_0||$ for all $x \in S_\epsilon(x_0) = \{x \in X \mid ||x - x_0|| < \epsilon\}$. We say that f is <u>locally contractive on</u> X provided there exists at least one point $x_0 \in X$ such that f is contractive at x_0.

Observe that if f is contractive at x_0, then $||f(x_0) - x_0|| \leq \alpha||x_0 - x_0|| = 0$, and hence x_0 is a fixed point of f. The following proposition is of prime importance in [2].

Proposition (1.5). Let $(X, \|\ \|)$ be a normed linear space and let $f : X \to X$ be a function. If f is contractive at $x_0 \in X$, with corresponding real numbers ϵ and α, then, for any $y \in S_\epsilon(x_0)$, $f^n(y) \in S_\epsilon(x_0)$ for each positive integer n, and the sequence $\{f^n(y)\}_{n=1}^{\infty}$ converges to x_0.

Proof. Let $y \in S_\epsilon(x_0)$. Then $\|f(y) - x_0\| \le \alpha\|y - x_0\| < \alpha \cdot \epsilon < \epsilon$; hence $f(y) \in S_\epsilon(x_0)$. Let $k \ge 1$ and suppose that $f^k(y) \in S_\epsilon(x_0)$ and $\|f^k(y) - x_0\| \le \alpha^k\|y - x_0\|$. Then

$$\|f^{k+1}(y) - x_0\| = \|f(f^k(y)) - x_0\| \le \alpha\|f^k(y) - x_0\| \le \alpha \cdot \alpha^k\|y - x_0\|$$

$$= \alpha^{k+1}\|y - x_0\| < \alpha^{k+1} \cdot \epsilon < \epsilon .$$

Therefore, $f^{k+1}(y) \in S_\epsilon(x_0)$ and, since $\alpha^k \to 0$, as $k \to +\infty$, it follows that $\|f^k(y) - x_0\| \to 0$, as $k \to +\infty$. Thus, the sequence $\{f^n(y)\}_{n=1}^{\infty}$ converges to x_0.

2. Conditions for functions to be locally contractive

Throughout this section $(X, \|\ \|)$ will be a finite dimensional real normed linear space and $f : X \to X$ will be a function. Observe that if X is n-dimensional, then f can be represented as:

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \\ f_n(x_1, \ldots, x_n) \end{pmatrix} ,$$

where each $f_i(x_1, \ldots, x_n)$ is a real-valued function defined on X. We call

$f_i(x_1, \ldots, x_n)$ the <u>i-th coordinate function</u> of f. We say that $f \in C'$ in a region D (an open connected subset of X) provided each of the coordinate functions of f possess continuous first partial derivatives in D.

The <u>Frechét derivative of f at the point</u> $p = (p_1, \ldots, p_n) \in X$ is defined to be the matrix of real numbers

$$df(p) = \begin{pmatrix} \frac{\partial f_1(p_1, \ldots, p_n)}{\partial x_1} & \cdots & \frac{\partial f_1(p_1, \ldots, p_n)}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_n(p_1, \ldots, p_n)}{\partial x_1} & \cdots & \frac{\partial f_n(p_1, \ldots, p_n)}{\partial x_n} \end{pmatrix} \quad [1].$$

Observe that $df(p)$ determines a linear transformation $df(p) : X \to X$ defined by $(df(p))(x) = df(p) \cdot x$, where $\cdot$ denotes matrix multiplication. When $df(p)$ is looked at in this manner (that is, as a linear transformation of X into itself), we call $df(p)$ or $df(p) \cdot x$ the <u>differential of f at the point</u> p. Despite its ambiguity, this distinction in terminology should cause no difficulty in the sequel.

Observe that the condition that f is contractive at x_0 is equivalent to the condition: there exist real numbers δ and α, $\delta > 0$, $0 \le \alpha < 1$, such that $||f(x_0 + \Delta x) - f(x_0)|| = ||f(x_0 + \Delta x) - x_0|| \le \alpha||\Delta x||$ whenever $||\Delta x|| < \delta$. The following theorem is well-known. (See, for example, [1; Theorem 17, p.264].)

<u>Theorem</u> (2.1). Let $f \in C'$ in the region $D \subseteq X$, let $p \in D$, and let $df(p)$ be the Frechét derivative of f at the point p. Then

$$f(p + \Delta p) = f(p) + df(p) \cdot \Delta p + R(\Delta p),$$

where $\lim_{\Delta p \to 0} \frac{||R(\Delta p)||}{||\Delta p||} = 0.$

Theorem (2.2). Let $f \in C'$ in the region $D \subseteq X$ and let $x_0 \in D$. Then f is contractive at x_0 if and only if x_0 is a fixed point of f and there exist real numbers δ and β, $\delta > 0$, $0 \leq \beta < 1$, such that $\frac{||df(x_0) \cdot \Delta x||}{||\Delta x||} \leq \beta$ whenever $||\Delta x|| < \delta$.

Proof. ($\longleftarrow$). Suppose that $f(x_0) = x_0$ and that such a δ and β exist. By Theorem (2.1),

$$||df(x_0) \cdot \Delta x|| = ||f(x_0 + \Delta x) - f(x_0) - R(\Delta x)|| \geq ||f(x_0 + \Delta x) - f(x_0)|| - ||R(\Delta x)||.$$

Thus, for $||\Delta x|| < \delta$,

$$1 > \beta \geq \frac{||df(x_0) \cdot \Delta x||}{||\Delta x||} \geq \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||} - \frac{||R(\Delta x)||}{||\Delta x||}, \text{ or}$$

$$\beta + \frac{||R(\Delta x)||}{||\Delta x||} \geq \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||}.$$

Choose $\epsilon > 0$ so that $\beta + \epsilon = \alpha < 1$. Since $\lim_{||\Delta x|| \to 0} \frac{||R(\Delta x)||}{||\Delta x||} = 0$, there exists $\gamma > 0$ such that $\frac{||R(\Delta x)||}{||\Delta x||} < \epsilon$ whenever $||\Delta x|| < \gamma$. Thus, for $\tau = \min\{\delta, \gamma\}$, it follows that if $||\Delta x|| < \tau$, then

$$||f(x_0 + \Delta x) - f(x_0)|| \leq (\beta + \epsilon)||\Delta x|| = \alpha||\Delta x||$$

and hence f is contractive at x_0.

($\longrightarrow$). Suppose f is contractive at x_0. Then x_0 is a fixed point of f and there exist real numbers $\gamma > 0$ and α, $0 \le \alpha < 1$, such that $||f(x_0 + \Delta x) - x_0|| = ||f(x_0 + \Delta x) - f(x_0)|| \le \alpha||\Delta x||$ whenever $||\Delta x|| < \gamma$. By Theorem (2.1),

$$|| df(x_0) \bullet \Delta x|| = ||f(x_0 + \Delta x) - f(x_0) - R(\Delta x)|| \le ||f(x_0 + \Delta x) - f(x_0)|| + ||R(\Delta x)||.$$

Choose $\epsilon > 0$ so that $\alpha + \epsilon = \beta < 1$. Since $\lim\limits_{||\Delta x||\to 0} \dfrac{||R(\Delta x)||}{||\Delta x||} = 0$, there exists $\tau > 0$ such that $\dfrac{||R(\Delta x)||}{||\Delta x||} < \epsilon$ whenever $||\Delta x|| < \tau$. Thus, for $\delta = \min\{\gamma,\tau\}$, it follows that

$$\frac{|| df(x_0) \bullet \Delta x||}{||\Delta x||} \le \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||} + \frac{||R(\Delta x)||}{||\Delta x||} \le \alpha + \epsilon = \beta < 1.$$

Definition (2.3). Let $f \in C'$ in the region $D \subseteq X$, let $x_0 \in D$, and suppose that $df(x_0)$ exists. We define the norm of the linear operator $df(x_0)$ to be:

$$||df(x_0)|| = \sup\{||df(x_0) \bullet x|| \mid ||x|| = 1, x \in X\}.$$

Theorem (2.4). Under the hypothesis of Theorem (2.2), $|| df(x_0) || < 1$ if and only if there exist real numbers $\delta > 0$ and $\beta, 0 \le \beta < 1$, such that $\dfrac{||df(x_0) \bullet \Delta x||}{||\Delta x||} \le \beta$ whenever $||\Delta x|| < \delta$.

Proof. Suppose $||df(x_0)|| < 1$. Then $1 > \beta = ||df(x_0)|| =$

$$= \sup_{||x|| = 1} ||df(x_0) \bullet x|| \ge ||df(x_0) \bullet \frac{\Delta x}{||\Delta x||}|| = \frac{||df(x_0) \bullet \Delta x||}{||\Delta x||} \text{ for all } \Delta x.$$

Conversely, suppose that $\frac{||df(x_0) \cdot \Delta x||}{||\Delta x||} \le \beta < 1$ whenever $||\Delta x|| < \delta$. Let $x \in X$, $||x|| = 1$, and choose $\gamma < \delta$. Then $||\gamma x|| = \gamma ||x|| = \gamma < \delta$, so that

$$||df(x_0) \cdot x|| = ||df(x_0) \cdot \frac{\gamma x}{\gamma}|| = \frac{1}{\gamma} ||df(x_0) \cdot \gamma x|| \le \frac{1}{\gamma} \beta \cdot ||\gamma x|| = \beta .$$

Since x was an arbitrary vector of X of norm 1, it follows that $||df(x_0)|| \le \beta < 1$.

Combining Theorems (2.2) and (2.4), we have the following result.

Result (2.5). Let $f \in C'$ in the region $D \subseteq X$ and $x_0 \in D$. f is contractive at x_0 if and only if x_0 is a fixed point of f and $||df(x_0)|| < 1$.

Theorem (2.6). Under the hypothesis of Result (2.5),

$$\lim_{||\Delta x|| \to 0} \frac{||df(x_0) \cdot \Delta x||}{||\Delta x||}$$ exists if and only if

$$\lim_{||\Delta x|| \to 0} \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||}$$ exists. Moreover, if either limit exits (and hence the other), they are equal.

Proof. By Theorem (2.1)

$$\big|\,||df(x_0) \cdot \Delta x|| - ||f(x_0 + \Delta x) - f(x_0)||\,\big|$$

$$= \big|\,||f(x_0 + \Delta x) - f(x_0) - R(\Delta x)|| - ||f(x_0 + \Delta x) - f(x_0)||\,\big|$$

$$\le ||f(x_0 + \Delta x) - f(x_0) - R(\Delta x) - [f(x_0 + \Delta x) - f(x_0)]||$$

$$= ||R(\Delta x)||,$$

where $\lim_{||\Delta x||\to 0} \frac{||R(\Delta x)||}{||\Delta x||} = 0$. Thus, given $\epsilon > 0$, there exists $\delta > 0$ such that

$$\left| \frac{||df(x_0) \cdot \Delta x||}{||\Delta x||} - \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||} \right| \le \frac{||R(\Delta x)||}{||\Delta x||} < \epsilon,$$

whenever $||\Delta x|| < \delta$. Thus, if one limit exists, then the other exists and the two limits are equal.

<u>Lemma</u> (2.7). Let $(X, || \ ||)$ be a finite dimensional normed linear space and let $A : X \to X$ be a linear operator. If $\lim_{||x||\to 0} \frac{||Ax||}{||x||}$ exists, then $\lim_{||x||\to 0} \frac{||Ax||}{||x||} = ||A||$, where $||A|| = \sup\{||Ax|| \mid x \in X, ||x|| = 1\}$.

<u>Proof</u>. Since the unit ball $= \{x \in X \mid ||x|| = 1\}$ is a compact subspace of X and A is continuous, there exists $x_1 \in X$ such that $||x_1|| = 1$ and $||Ax_1|| = ||A||$. Therefore, $\frac{||Ax_1||}{||x_1||} = ||A||$. Choose $\alpha_n = \frac{1}{n}$ for each positive integer n. Then $\frac{||A(\alpha_n x_1)||}{||\alpha_n x_1||} = \frac{\alpha_n ||Ax_1||}{\alpha_n ||x_1||} = ||A||$ for all n, and hence $\lim_{||\alpha_n x_1||\to 0} \frac{||A(\alpha_n x_1)||}{||\alpha_n x_1||} = ||A||$. Thus, if $\lim_{||x||\to 0} \frac{||Ax||}{||x||}$ exists, then $\lim_{||x||\to 0} \frac{||Ax||}{||x||} = ||A||$.

Observe that under the hypothesis of Result (2.5), $df(x_0)$ exists and is a linear function of X into itself. Thus, by Result (2.5), Theorem (2.6), and Lemma (2.7), we have the following corollary.

Corollary .8). Let $f \in C'$ in the region $D \subseteq X$ and let $x_0 \in D$. If

$$\lim_{||\Delta x|| \to 0} \frac{||df(x_0) \cdot \Delta x||}{||\Delta x||} = L < \infty \quad \text{(equivalently, if}$$

$$\lim_{||\Delta x|| \to 0} \frac{||f(x_0 + \Delta x) - f(x_0)||}{||\Delta x||} = L < \infty), \text{ then } f \text{ is contractive at } x_0$$

if and only if x_0 is a fixed point of f and $L < 1$.

While it does not seem unreasonable to expect that both limits of Theorem (2.6) exist, the following theorem shows that the existence of the limits imposes strong conditions on the linear mapping $df(x_0)$.

Theorem (2.9). Let $(X, (,))$ be a finite dimensional inner product space with induced norm $||x|| = \sqrt{(x,x)}$ and let $A : X \to X$ be a linear operator. Then $A = \alpha U$ for some non-negative real number α and orthogonal matrix U if and only if $\lim_{||x|| \to 0} \frac{||Ax||}{||x||}$ exists.

Proof. Observe that a linear operator B defined on X is orthogonal if and only if $(Bx,Bx) = (x,x)$, or equivalently, if and only if $||Bx|| = ||x||$ for all $x \in X$. Note that the theorem is obviously true if $A = 0$. If $A \neq 0$ and if $\lim_{||x|| \to 0} \frac{||Ax||}{||x||}$ exists, then, by Lemma (2.7), $\lim_{||x|| \to 0} \frac{||Ax||}{||x||} = ||A||$ and

therefore, $\lim_{||x|| \to 0} \frac{||(\frac{1}{||A||}A)x||}{||x||} = 1$. Let $B = \frac{1}{||A||} \cdot A$ and let $y \in X$.

Choose $\alpha_n = \frac{1}{n}$ for each positive integer and consider the sequence of points $\{\alpha_n y\}_{n=1}^{\infty}$. Since $||\alpha_n y|| \to 0$ as $n \to +\infty$, then $\lim_{||\alpha_n y|| \to 0} \frac{||B(\alpha_n y)||}{||\alpha_n y||} = 1$.

But $\dfrac{||B(\alpha_n y)||}{||\alpha_n y||} = \dfrac{||By||}{||y||}$ for each n and hence $\lim\limits_{||\alpha_n y||\to 0} \dfrac{||B(\alpha_n y)||}{||\alpha_n y||} = \dfrac{||By||}{||y||}$.

Therefore, $||By|| = ||y||$, so that B is orthogonal, and thus, $A = ||A||B$ has the required form.

Conversely, suppose that $A = \alpha U$ for some positive real number α and orthogonal matrix U. Then $||Ux|| = ||x||$ for all $x \in X$, and hence

$$\lim_{||x||\to 0} \frac{||Ax||}{||x||} = \lim_{||x||\to 0} \frac{||\alpha Ux||}{||x||} = \alpha \lim_{||x||\to 0} \frac{||Ux||}{||x||} = \alpha \cdot 1 = \alpha \cdot ||U|| = ||\alpha U|| = ||A||.$$

<u>Corollary</u> (2.10). Let $(X, (\,,\,))$ be a finite dimensional inner product space with induced norm $||x|| = \sqrt{(x,x)}$, let $f \in C'$ in the region $D \subseteq X$, and let $x_0 \in D$. If $\lim\limits_{||\Delta x||\to 0} \dfrac{||df(x_0) \cdot \Delta x||}{||\Delta x||}$ exists, then f is contractive at x_0 if and only if x_0 is a fixed point of f and $df(x_0) = \alpha U$ for some real number α, $0 < \alpha < 1$, and orthogonal matrix U.

We end this section with an example to show that the condition that $\lim\limits_{||\Delta x||\to 0} \dfrac{||df(x_0) \cdot \Delta x||}{||\Delta x||}$ exists is required in Corollary (2.10).

<u>Example</u> (2.11). Let E^2 denote real 2-space and let $f : E^2 \to E^2$ be defined by $f(x_1,x_2) = (\dfrac{3x_1}{4}, \dfrac{x_2}{2})$. Since $||f(x_1,x_2) - (0,0)|| \le \frac{3}{4}||(x_1,x_2) - (0,0)||$ for all $(x_1,x_2) \in E^2$, and since $(0,0)$ is clearly a fixed point of f, then f is contractive at $(0,0)$. However, $df(x_1,x_2)$ at any point $(x_1,x_2) \in E^2$ is equal to $\begin{bmatrix} 3/4 & 0 \\ 0 & 1/2 \end{bmatrix}$, and it is easily verified that $\begin{bmatrix} 3/4 & 0 \\ 0 & 1/2 \end{bmatrix}$ cannot be written in the form αU, where α is a scalar and U is orthogonal.

3. A matrix representation of df(x)

Let E^n denote real or complex n space and let $f : E^n \to E^n$ be defined by $f(x) = A(x) \cdot x$, where A is a function from E^n into $M_{n\times n}$, the space of $n \times n$ matrices. For any $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E^n$, write $f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}$. Thus, for each $i = 1, \ldots, n$, $f_i : E^n \to E^1$ is the projection of $f(x)$ onto its ith coordinate. Let $A \in C'$ in the region $R \subseteq E^n$ (hence $f \in C'$ in R). Then, for any point $x \in R$, the Frechét derivative of f is

$$df(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix} .$$

Let
$$A(x) = \begin{pmatrix} a_{11}(x) & \cdots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \cdots & a_{nn}(x) \end{pmatrix} .$$

Therefore, $f_i(x) = \Sigma_{k=1}^{n} a_{ik}(x) \cdot x_k$ for each i and

$$\frac{\partial f_i(x)}{\partial x_j} = \Sigma_{k=1}^{n} \frac{\partial [a_{ik}(x)]}{\partial x_j} \cdot x_k + a_{ij}(x)$$

is the (i,j)th entry of $df(x)$. If we denote the matrix

$$\begin{pmatrix} \sum_{k=1}^{n} \frac{\partial [a_{1k}(x)]}{\partial x_1} \cdot x_k & \cdots & \sum_{k=1}^{n} \frac{\partial [a_{1k}(x)]}{\partial x_n} \cdot x_k \\ \\ \sum_{k=1}^{n} \frac{\partial [a_{nk}(x)]}{\partial x_1} \cdot x_k & \cdots & \sum_{k=1}^{n} \frac{\partial [a_{nk}(x)]}{\partial x_n} \cdot x_k \end{pmatrix}$$

by $B(x)$, then $df(x) = B(x) + A(x)$.

Since $A : E^n \to M_{n \times n}$, it follows that $dA(x)$ can be represented by the $n^2 \times n$ matrix

$$\begin{pmatrix} \frac{\partial [a_{11}(x)]}{\partial x_1} & \cdots & \frac{\partial [a_{11}(x)]}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial [a_{1n}(x)]}{\partial x_1} & \cdots & \frac{\partial [a_{1n}(x)]}{\partial x_n} \\ \frac{\partial [a_{21}(x)]}{\partial x_1} & \cdots & \frac{\partial [a_{21}(x)]}{\partial x_n} \\ \vdots & & \\ \frac{\partial [a_{n1}(x)]}{\partial x_1} & \cdots & \frac{\partial [a_{n1}(x)]}{\partial x_n} \\ \vdots & & \\ \frac{\partial [a_{nn}(x)]}{\partial x_1} & \cdots & \frac{\partial [a_{nn}(x)]}{\partial x_n} \end{pmatrix}$$

If we let $D(x)$ be the $n \times n^2$ matrix

$$\begin{pmatrix} x_1 \cdots x_n & & & \\ & x_1 \cdots x_n & & \\ & & \ddots & \\ & & & x_1 \cdots x_n \end{pmatrix},$$

then $B(x) = D(x)dA(x)$ so that $df(x) = D(x) \cdot dA(x) + A(x)$.

Although $D(x)$ and $dA(x)$ are not self maps, it is straightforward to show that

$$\|df(x)\| \le \|D(x)dA(x)\| + \|A(x)\| \le \|D(x)\| \cdot \|dA(x)\| + \|A(x)\| ,$$

where we take the usual sup norm on the linear functions $df(x)$, $D(x)$, $dA(x)$, and $A(x)$. (For example, $\|D(x)\| = \sup\{\|D(x) \cdot y\|_n \mid \|y\|_{n^2} = 1, y \in E^{n^2}\}$, where $\| \ \|_n$ and $\| \ \|_{n^2}$ denote the Euclidean norms in E^n and E^{n^2}, respectively.)

We next show that $\|D(x)\| \le \|x\|$ for all $x \in E^n$. For, if $y \in E^{n^2}$, then

$$D(x) \cdot y = D(x) \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ y_{n+1} \\ \vdots \\ y_{2n} \\ \vdots \\ y_{n^2} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i y_{n+i} \\ \vdots \\ \sum_{i=1}^{n} x_i y_{n^2-n+i} \end{pmatrix},$$

and hence,

$$||D(x) \cdot y|| = \sqrt{\sum_{k=0}^{n-1} \left(\sum_{i=1}^{n} x_i y_{kn+i} \right)^2} \quad .$$

By the Cauchy-Schwarz-Bunyakovskii inequality,

$$\left(\sum_{i=1}^{n} x_i y_{kn+i} \right)^2 \le \left(\sum_{i=1}^{n} x_i^2 \right) \cdot \left(\sum_{i=1}^{n} y_{kn+i}^2 \right) \quad \text{for any } k = 0,1,\ldots n-1,$$

and hence,

$$\begin{aligned}
||D(x) \cdot y||^2 &= \sum_{k=0}^{n-1} \left(\sum_{i=1}^{n} x_i y_{kn+i} \right)^2 \\
&\le \sum_{k=0}^{n-1} \left[\left(\sum_{i=1}^{n} x_i^2 \right) \cdot \left(\sum_{i=1}^{n} y_{kn+i}^2 \right) \right] \\
&= \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{k=0}^{n-1} \sum_{i=1}^{n} y_{kn+i}^2 \right) \\
&= \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{i=1}^{n^2} y_i^2 \right) \\
&= ||x||^2 \cdot ||y||^2 \quad .
\end{aligned}$$

In particular, if $||y|| = 1$, we have that $||D(x) \cdot y|| \le ||x||$ and hence $||D(x)|| = \sup \{ ||D(x) \cdot y|| \mid ||y|| = 1, y \in E^{n^2} \} \le ||x||$.

Summarizing the results thus far developed in this section, we have the following theorem.

<u>Theorem</u> (3.1). Let $f : E^n \to E^n$ be defined by $f(x) = A(x) \cdot x$, where A is a function from E^n into $M_{n \times n}$. Suppose that $A \in C'$ in the region $R \subseteq E^n$ (hence $f \in C'$ in R). Then, for any point $x \in R$,

$df(x) = D(x) \cdot dA(x) + A(x)$, where $D(x)$ is the $n \times n^2$ matrix

$$\begin{pmatrix} x_1 \cdots x_n & & & \\ & x_1 \cdots x_n & & \\ & & \ddots & \\ & & & x_1 \cdots x_n \end{pmatrix}.$$

Moreover,

$$\|df(x)\| \leq \|D(x)\| \cdot \|dA(x)\| + \|A(x)\| \leq \|x\| \cdot \|dA(x)\| + \|A(x)\|.$$

It should be noted that the above norm inequality is not sharp enough in seeking points at which f is contractive. For, if f is to be contractive at x_0, then x_0 is a fixed point of f. But then $\|x_0\| = \|f(x_0)\| = \|A(x_0) \cdot x_0\| \leq \|A(x_0)\| \cdot \|x_0\|$, so that (if $x_0 \neq 0$) $\|A(x_0)\| \geq 1$. Hence, $\|x_0\| \cdot \|dA(x_0)\| + \|A(x_0)\| \geq 1$ and the above inequality cannot be used to show that $\|df(x_0)\| < 1$.

Finally, we close the paper with some observations relating to the work of B. C. Peters and H. F. Walker in [2]. A major result of [2] is that a function Φ_ϵ is locally contractive at a consistent maximum-likelihood estimate x_0 for sufficiently small values of $\epsilon > 0$. Since $\Phi_\epsilon(x_0) = x_0$, it suffices to show that $\|d\Phi_\epsilon(x_0)\| < 1$. It is easily verified that Φ_ϵ

can be considered as a function from E^n into itself and that Φ_ϵ can be written in the form: $\Phi_\epsilon = (1 - \epsilon)I + \epsilon \cdot f$, where I is the identity map on E^n and $f : E^n \to E^n$ is defined by $f(x) = A(x) \bullet x$, where A is a function from E^n into $M_{n \times n}$. (That is, f is a function of the type considered earlier in this section.)

Since $d\Phi_\epsilon(x) = (1 - \epsilon)I + \epsilon\, df(x)$ at any point x for which the differentials exist, then

$$||d\Phi_\epsilon(x)|| \le (1 - \epsilon) + \epsilon\ ||\ df(x)||$$

Thus, if $||\ df(x)|| < 1$, then

$$||d\Phi_\epsilon(x)|| \le (1 - \epsilon) + \epsilon||\ df(x)|| \le 1 - (1 - ||\ df(x)||)\epsilon < 1.$$

Therefore, since the set of fixed points of f is the set of fixed points of Φ_ϵ for any $\epsilon > 0$, it follows that if f is contractive at y_0, then Φ_ϵ is contractive at y_0 for any $\epsilon > 0$.

Conversely, observe that $df(x) = \frac{1}{\epsilon}[d\Phi_\epsilon(x) + (\epsilon - 1)I]$, so that $||\ df(x)|| \le \frac{1}{\epsilon}\ ||d\Phi_\epsilon(x)|| + \frac{|\epsilon - 1|}{\epsilon}$. Thus, if $||d\Phi_\epsilon(x)|| = \alpha$, then $||\ df(x)|| < 1$ if $\frac{1}{\epsilon}(\alpha + |\epsilon - 1|) < 1$. For $0 < \epsilon \le 1$, $\frac{1}{\epsilon}(\alpha + |\epsilon - 1|) < 1$ is equivalent to $\alpha < 2\epsilon - 1$, and thus, if $||d\Phi_\epsilon(x)|| < 2\epsilon - 1$, then $||\ df(x)|| < 1$. (Note, the assumption $\alpha < 2\epsilon - 1$ implies that $\alpha < 1$ for $0 < \epsilon \le 1$.) Thus, if Φ_ϵ is contractive at y_0 and if $||d\Phi_\epsilon(y_0)|| < 2\epsilon - 1$, then f is contractive at y_0.

Observe that if $\epsilon \ge 1$ and if $||d\Phi_\epsilon(x)|| = \alpha$, then

$||\, df(x)|| \le \frac{1}{\epsilon}(\alpha + \epsilon - 1)$. Thus, since the condition $\frac{1}{\epsilon}(\alpha + \epsilon - 1) < 1$ is equivalent to the condition $\alpha < 1$, it follows that if $||d\Phi_\epsilon(x)|| < 1$, then $||\, df(x)|| < 1$. Therefore, for $\epsilon \ge 1$, f is contractive at y_0 if and only if Φ_ϵ is contractive at y_0. We remark that Peters and Walker have recently shown that Φ_ϵ is contractive at a consistent maximum-likelihood estimate x_0 for $0 < \epsilon < 2$.

We summarize these final comments.

Theorem (3.2). With the notation and hypothesis of the preceding discussion, if f is contractive at y_0, then Φ_ϵ is contractive at y_0 for all $\epsilon > 0$. If $0 < \epsilon < 1$, $||d\Phi_\epsilon(y_0)|| < 2\epsilon - 1$, and y_0 is a fixed point of Φ_ϵ (hence Φ_ϵ is contractive at y_0), then f is contractive at y_0. If $1 \le \epsilon$, f is contractive at y_0 if and only if Φ_ϵ is contractive at y_0.

BIBLIOGRAPHY

1. Buck, R. C., Advanced Calculus, McGraw-Hill, New York, 1965.

2. Peters, B. C. and Walker, H. F., "An iterative procedure for obtaining maximum-likelihood estimates of the parameters for a mixture of normal distributions" NASA technical note, Report 43, prepared under Contract NAS-9-12777.