NASA CR-151250

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF HOUSTON HOUSTON, TEXAS

(NASA-CR-151250) ON THE CCNVERGENCE OF OPTIMAL LINEAF COMBINATICN PROCEDURES (Houston Univ.) 9 p HC A02/MF A01 CSCL 09B

N77-21865

Unclas

G3/61 22928

ON THE CONVERGENCE OF OPTIMAL LINEAR COMBINATION PROCEDURES

BY WILLIAM TALLY

REPORT #63 FEBRUARY 1977

PREPARED FOR
EARTH OBSERVATION DIVISION, JSC
UNDER
CONTRACT NAS-9-15000

HOUSTON, TEXAS 77004

On the Convergence of

Optimal Linear Combination Procedures

William Tally

Report #63
NAS-9-15000

January 1977

Introduction:

The following algorithm has been suggested by Decell and Smiley in [1] for optimal linear combinations in the feature selection problem.

Let Ψ be a continuous function from M_n^k (see definition 1) into R^1 that is invariant under multiplication on the left by kxk invertible matrices. Then there exists $H_1 \in \mathcal{H}_n$ (see definition 2) such that

$$\Psi([I_k|Z]H_1) = \underset{H \in \mathcal{H}_n}{\text{l.u.b.}} \{\Psi([I_k|Z]H)\}.$$

Now for each positive integer i, let the element $H \in \mathcal{H}_n$ be chosen such that

$$\Psi([I_k|Z]H_i H_{i-1} \cdots H_1) = \underset{H \in \mathcal{H}_n}{\text{l.u.b.}} \Psi([I_k|Z]H \cdot H_{i-1} \cdots H_1)$$

The question of whether or not the above process terminates at an absolute Ψ-extremum (rank k maximal statistic) appeared in [1]. In this paper, we show that there exists a function Ψ as above for which the above process does not terminate at an absolute Ψ-extremum.

Let $H_1, \ldots, H_p$ be the matrices representing Householder transformations. Then for the matrix $[I_k|Z]H_1 \cdots H_p$, let $\Theta([I_k|Z]H_1 \cdots H_p)$ be the span in R^n of the k **row** vectors of that matrix. Suppose that $v_1, \ldots, v_k$ are linearly independent vectors in R^n. Then we show in this paper that there exists some integer $p \leq \min(n, n-k)$ and Householder transformations whose matrices are $H_1, \ldots, H_p$ for which

$\theta([I_k|Z]H_1\cdots H_p) = \mathrm{Span}\{v_1,\ldots,v_k\}$. We also determine the minimum integer p having the above property.

Preliminaries:

Definition 1. Let M_n^k be the set of all kxn rank k matrices.

Definition 2. Let $\mathcal{H}_n$ denote the set of all Householder transformations.

Definition 3. Let $\mathcal{S}_n^k$ denote the collection of all vector subspaces of R^n of dimension k.

Definition 4. Let $S^n = \{x \in R^n \mid \|x\| = 1\}$.

Definition 5. Let $\mathcal{C}$ be a closed subset of R^n and $x \notin \mathcal{C}$. Then there exists $c_x \in \mathcal{C}$ such that $\|x-c_x\| \leq \|x-c\|$ for any $c \in \mathcal{C}$. Let $\rho(x;\mathcal{C}) = \|x-c_x\|$.

Definition 6. Let A and B be elements of $\mathcal{S}_n^k$. Then there exists an element $a^* \in A \cap S^n$ having the property that $\rho(a^*; B\cap S^n) \geq \rho(a; B\cap S^n)$ for all $a \in A\cap S^n$. The number $\rho(a^*; B\cap S^n)$ will be called the distance from A to B and will be denoted by the symbol d(A;B).

Proposition 1. For any elements A, B, and C in $\mathcal{S}_n^k$

i) $d(A;B) \geq 0$ and $d(A;B) = 0$ if and only if $A = B$.

ii) $d(A;C) \leq d(A;B) + d(B;C)$.

iii) For any $\xi > 0$ there exists a $\delta > 0$ such that whenever $d(A;B) < \delta$, then $d(B;A) < \xi$.

Definition 7. For any $P \in \mathcal{S}_n^k$ and $\xi > 0$, let

$$\mathcal{U}_\xi(P) = \{X \in \mathcal{S}_n^k \mid d(X;P) < \xi\}.$$

Definition 8. Let T be the topology on $\mathcal{S}_n^k$ determined by the subbasis $\{\mathcal{U}_\xi(P) \mid \xi > 0 \text{ and } P \in \mathcal{S}_n^k\}$.

Definition 9. Let $\mathcal{C}$ be a closed subset of $\mathcal{S}_n^k$ and let $P \in \mathcal{S}_n^k$. Let $D(P;\mathcal{C}) = \text{g.l.b.}\{d(P;C) \mid C \in \mathcal{C}\}$.

<u>Proposition 2.</u> $(\mathcal{S}_n^k, T)$ is normal.

<u>Proof:</u> Let $\mathcal{A}$ and $\mathcal{B}$ be two closed disjoint subsets of $\mathcal{S}_n^k$. Let $\mathcal{U}_1 = \{P \in \mathcal{S}_n^k \mid D(P;\mathcal{A}) < D(P;\mathcal{B})\}$ and $\mathcal{U}_2 = \{P \in \mathcal{S}_n^k \mid D(P;\mathcal{A}) > D(P;\mathcal{B})\}$. By Proposition 1, we can determine that $\mathcal{U}_1$ and $\mathcal{U}_2$ are both open and are disjoint. This completes the proof.

Definition 10. For any vector $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ in R^n, let $w^U = \begin{pmatrix} w_1 \\ \vdots \\ w_k \end{pmatrix}$

and $w^L = \begin{pmatrix} w_{k+1} \\ \vdots \\ w_n \end{pmatrix}$.

<u>Proposition 3.</u> Suppose that $\{v_1, \ldots, v_k\}$ is a collection of linearly independent vectors in R^n. Let p be the dimension of $\text{Span}\{v_1^L, \ldots, v_k^L\}$ and assume $p > 0$. Then there exists a vector $x \in R^n$ such that $\|x\| = 1$, and if H_x is the Householder transformation determined by x, then the dimension of $\text{Span}\{H_x(v_1)^L, \ldots, H_x(v_k)^L\} = p-1$.

<u>Proof:</u> Case i) Dimension of $\text{Span}\{v_1^U, \ldots, v_k^U\}$ is less than k. We select a vector x^L in $\text{Span}\{v_1^L, \ldots, v_k^L\}$ such that $\|x^L\| = \sqrt{\tfrac{1}{2}}$. Since $[v_i^L - 2(v_i^L \cdot x^L)x^L] \cdot x^L = 0$ for $i=1,\ldots,k$. It follows that the dimension of $\text{Span}\{v_1^L - 2(v_1^L \cdot x^L)x^L, \ldots, v_k^L - 2(v_k^L \cdot x^L)x^L\}$ is p-1. Now by assumption there exists a vector x^U in R^k such that $\|x^U\| = \sqrt{\tfrac{1}{2}}$, and $v_i^U \cdot x^U = 0$ for $i=1,\ldots,k$. Since $v_i^L - 2(v_i \cdot x)x^L = v_i^L - 2(v_i^L \cdot x^L)x^L$, then the dimension of

Span $\left\{v^L-2(v_1^L\cdot x^L)x^L,\ldots,v_k^L-2(v_k^L\cdot x^L)x^L\right\}$ is p-1, for

$$x = \begin{pmatrix} x^U \\ \vdots \\ x^U \end{pmatrix} .$$

Case ii) The dimension of Span$\left\{v_1^U,\ldots,v_k^U\right\}$ = k.
We select a vector x_o^L in Span $\left\{v_1^L,\ldots,v_k^L\right\}$ with $\|x_o^L\| = \sqrt{\tfrac{1}{2}}$.
Then we have that the dimension of
Span $\left\{v_1^L-2(v_1^L\cdot x_o^L)x_o^L,\ldots,v_k^L-2(v_k^L\cdot x_o^L)x_o^L\right\}$ is p-1. We assume then that $x^L = \lambda x_o^L$ for some $\lambda < 1$. We want a vector x^U in R^k such that if $x = \begin{pmatrix} x^U \\ x^L \end{pmatrix}$ then $\|x^U\|^2 + \|x^L\|^2 = 1$ and $v_i^L-2(v_i\cdot x)x^L = v_i^L-2(v_i^L\cdot x_o^L)x_o^L$ for i=1,...,k.

By substituting x_o^L into this equation in place of x^L we can determine that $v_i^U\cdot x^U = (\frac{1-\lambda^2}{\lambda})v_i^L\cdot x_o^L$ for i=1,...,k.
By our assumption we can find a vector x^U satisfying the above equations whenever a choice of λ is made. We observe that if λ approaches 1, then $\|x^U\|$ must approach 0, and $\|x^L\|$ must approach $\sqrt{\tfrac{1}{2}}$ so that if λ approaches 1, then $\|x^U\|^2 + \|x^L\|^2$ must approach $\sqrt{\tfrac{1}{2}}$. If λ approaches 0, then $\|x^U\|$ approaches $+\infty$ and $\|x^L\|$ approaches 0 so $\|x^U\|^2 + \|x^L\|^2$ approaches $+\infty$ as λ approaches 0.
It follows from this that there exists some λ for which $\|x^U\|^2 + \|x^L\|^2 = 1$. Thus we have the dimension of Span$\left\{v_1^L-2(v_1\cdot x)x^L,\ldots,v_k^L-2(v_k\cdot x)x^L\right\}$ is p-1 which is the required condition. This completes the proof of proposition 3.

Definition 11. For any $M \in M_n^k$ let $\theta(M) = \text{Span}\{v_1, \ldots, v_k\}$ where $\{v_1, \ldots, v_k\}$ are the row vectors of M. θ is easily seen to be continuous.

<u>Proposition 4.</u> Suppose that $\theta([I_k|Z]H_1 \ldots H_p) = \text{Span}\{v_1, \ldots, v_k\}$ for Householder transformations $H_1, \ldots, H_p$. Then the dimension of $\text{Span}\{v_1^L, \ldots, v_k^L\}$ cannot exceed p.

<u>Proof:</u> We observe first of all that for any collection of vectors $\{y_1, \ldots, y_m\}$ and any Householder transformation H_x determined by the vector x that

$\text{Span}\{H_x(y_1), \ldots, H_x(y_m)\} \subset \text{Span}\{y_1, \ldots, y_m, x\}$ ..

Now $\theta([I_k|Z]H_1 \ldots H_p) = \text{Span}\{H_p \ldots H_1(e_1), \ldots, H_p \ldots H_1(e_k)\}$ where e_i is the vector with 1 in the i^{th} place and 0 everywhere else. Thus by the above statements,

$\text{Span}\{v_1, \ldots, v_k\} \subset \text{Span}\{e_1, \ldots, e_k, x_1, \ldots, x_p\}$.

It follows that $\text{Span}\{v_1^L, \ldots, v_k^L\} \subset \text{Span}\{x_1^L, \ldots, x_p^L\}$.

Thus the dimension of $\text{Span}\{v_1^L, \ldots, v_k^L\}$ is less than or equal to p. This completes the proof of Proposition 4.

<u>Proposition 5.</u> For linearly independent vectors $\{v_1, \ldots, v_k\}$, if p is the dimension of $\text{Span}\{v_1^L, \ldots, v_k^L\}$ and $p \geq 0$, then there exists Householder transformations $H_1, \ldots, H_p$ such that $\theta([I_k|Z]H_1 \ldots H_p) = \text{Span}\{v_1, \ldots, v_k\}$ and no fewer than p Householder transformations can have this property.

<u>Proof:</u> This is a consequence of Propositions 3 and 4 .

Construction of the map Ψ

Definition 12. For any $P\in\mathcal{S}_n^k$ let $P = \text{Span}\{v_1,\dots,v_k\}$ and define $L(P)$ = the dimension of $\text{Span}\{v_1^L,\dots,v_k^L\}$.

Definition 13. For $0\le p\le n-k$ let $\mathcal{L}_p = \{A\subset\mathcal{S}_n^k \mid L(A)\le p\}$.

Proposition 6. $\mathcal{L}_p$ is closed for $p=0,\dots,n-k$.

Proof: This is a consequence of the fact that if $\{u_1,\dots,u_m\}$ is a collection of vectors in R^{n-k} and q is the dimension of $\text{Span}\{u_1,\dots,u_m\}$ then there exists a real number $\xi>0$ such that if $\|u_i-u_i^*\|$ for $i=1,\dots,m$, then the dimension of $\text{Span}\{u_1^*,\dots,u_m^*\}$ is greater than or equal to q. This completes the proof of Proposition 6.

Now for some $P\in\mathcal{L}_1$ there exists $\xi>0$ such that if $A\in\mathcal{L}_1$, then $\mathcal{U}_\xi(A)$ does not contain P. Let $\mathcal{A}$ be the closure in $\mathcal{S}_n^k$ of $\bigcup_{A\in\mathcal{L}_1}\{\mathcal{U}_\xi(A)\}$. By Urysohns lemma, [2] there exists a continuous function $\Phi_1:\mathcal{S}_n^k\to[0,1]\subset R^1$ such that $\Phi_1(P) = 1$ and $\Phi_1(A)=0$ for any $A\in\mathcal{A}$. Let $I = \text{Span}\{e_1,\dots,e_k\}$. Then $\mathcal{U}_\xi(I)\subset\mathcal{A}$ since $I\in\mathcal{L}_1$. Define a map $\Phi_2:\mathcal{S}_n^k\to[0,\frac{1}{2}]$ by $\Phi_2(X) = 0$ if $X\notin\mathcal{U}_\xi(I)$ and $\Phi_2(X) = \frac{\xi-d(X;I)}{2\xi}$ if $X\in\mathcal{U}_\xi(I)$. Let $\Phi = \Phi_1 + \Phi_2$ and define $\Psi = \Phi\circ\Theta$. We observe that $\mathcal{L}_1 = \Theta(\{[I_k|Z]H \mid H\in\mathcal{H}_n\})$. Also if $\Theta([I_k|Z]H_1) = I$ for some $H_1\in\mathcal{H}_n$ then for any $H\in\mathcal{H}_n$, $\Theta([I_k|Z]H.H_1)\in\mathcal{L}_1$. That Ψ has the desired properties follows from the fact that the function Φ has a maximum value of $\frac{1}{2}$ at I over the set $\mathcal{L}_1$ but Φ has a maximum value of 1 at P over the entire space $\mathcal{S}_n^k$.

REFERENCES

1. Decell, H. P. and Smiley, W. G.III, Householder Transformations and Optimal Linear Combinations, 1974, Report #38, University of Houston Mathematics Department.
2. Royden, H. L., Real Analysis, page 148, 1970, Macmillan Company, London.
3. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, 1958 John Wiley and Sons, Inc., New York.
4. Kullback, Solomon, Information Theory and Statistics, 1968 Dover Publications, New York.
5. Quirein, J. A., "Divergence and Necessary Condition for Extremum" Report #12 NAS-9-12777 University of Houston, Department of Mathematics, Nov. 1972.