DISCUSSION PAPER NO. 332

A PROPERTY OF MATRICES WITH POSITIVE DETERMINANTS

by

M. Kojima */

R. Saigal

June 1978

The work of both authors was supported by Grant MSC77-03472 from the National Science Foundation.

*/ Department of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan. The work of this author was done while visiting Northwestern University.

A Property of Matrices with Positive Determinants

M. Kojima*

R. Saigal

ABSTRACT

Let $\mathcal{P}$ be the set of all $n \times n$ real matrices which have a positive determinant. We show here that at least 2^{n-1} matrices are needed to "see" each matrix in $\mathcal{P}$. Also, any finite subset of $\mathcal{P}$ can be "seen" from a class of at most 2^{n-1} matrices in $\mathcal{P}$.

The work of both authors was supported by Grant MCS77-03472 from the National Science Foundation.

*Department of Information Sciences, Tokyo Institute of Technology, Tokyo, Japan. The work of this author was done while visiting Northwestern University.

A Property of Matrices with Positive Determinants

M. Kojima

R. Saigal

§1. Introduction

Let $\mathcal{P}$ be the set of all $n \times n$ real matrices which have a positive determinant. Given two matrices A and B in $\mathcal{P}$, we say A can be seen from B, and vice-versa, if

$$(1-t)A + tB \text{ is nonsingular for every } t \text{ in } [0,1] \tag{1.1}$$

i.e., the line joining A and B lies inside $\mathcal{P}$. Also, we say a subset $\mathcal{Q}$ can see B if for some A in $\mathcal{Q}$, (1.1) holds. In this note we consider the following question: What is the smallest set of matrices $\mathcal{Q}$ from which each matrix in $\mathcal{P}$ can be seen? It follows readily that $\mathcal{Q} \subset \mathcal{P}$. We give here a partial answer to the above question. We first show that $|\mathcal{Q}| \geq 2^{n-1}$, and also that any finite set of matrices in $\mathcal{P}$ can be seen from a subset of cardinality at most 2^{n-1}.

Certain other properties of $\mathcal{P}$ are known. Eaves [1] showed that $\mathcal{P}$ is path connected, and Todd [7] showed that any two matrices in $\mathcal{P}$ can be seen from a third (and thus considerably strengthened the result of [1]). Todd's work has implifications for understanding the produce of the fixed point algorithms [1],[2],[5],[8]. This work is motivated by the recent use of property (1.1) by Kojima and Saigal [3],[4] in establishing conditions when PL mappings are homeomorphisms, and the use of property (1.1) in establishing PL approximations to diffeomorphisms, Saigal [6]. Since the set of all matrices with negative determinants is a nonsingular linear transformation of $\mathcal{P}$, all our methods are also valid for this set.

§2. Notation, Definitions and Preliminary Results

In this section, we present our notation and establish some preliminary results.

For an $n \times n$ matrix A, by $A(k)$ we represent the $k \times k$ submatrix consisting of the first k rows and k columns of A, for each $k = 1,\ldots,n$, and call these the leading principal submatrices of A; and their determinants the leading principal minors of A. A simple fact about these leading principal minors of A is the following

Lemma 2.1: Let $\mathcal{P}^0 = \{A: \det A(k) \neq 0 \text{ for each } k\}$. Then, $\mathcal{P}^0$ is open and dense in the space of all $n \times n$ matrices.

Proof: $\mathcal{P}^0$ is clearly open. Now, for $A \notin P^0$, we note that for all sufficiently small $\varepsilon > 0$, $A + \varepsilon I \in \mathcal{P}^0$, and thus the denseness follows.

We now state a result, without proof, from the work of Saigal [5]:

Theorem 2.2 (Saigal): Let A and B be $n \times n$ matrices in $\mathcal{P}$. Then

$$\min_{0 \leq t \leq 1} \det[(1-t)A + tB] > 0 \tag{2.1}$$

if and only if

$$\inf_{\lambda \geq 0} \det[B^{-1}A + \lambda I] > 0. \tag{2.2}$$

Proof: See [5, Lemma 3.1.1].

(2.1) characterizes the situation when matrix A can see the matrix B in $\mathcal{P}$, and the theorem states that this is so if and only if $B^{-1}A$ has no negative real eigenvalues. Now, for $\delta > 0$, define an $n \times n$ diagonal matrix:

$$E(\delta) = \begin{bmatrix} \delta & & & \\ & \delta^2 & & \\ & & \ddots & \\ & & & \delta^{2n-1} \end{bmatrix} \tag{2.3}$$

and a polynomial

$$\phi(\lambda,\delta) = \lambda^n + a_1(\delta)\lambda^{n-1} + \cdots + a_{n-1}(\delta)\lambda + a_n(\delta)$$
$$= \det(E(\delta)A + \lambda I). \tag{2.4}$$

Then, we can prove:

Lemma 2.3: Let $\det(E(\delta)A + \lambda I) = \phi(\lambda,\delta)$, and

$$\text{sign} \det A(k) = \varepsilon_k \quad \text{for } k = 1,\ldots,n,$$

where $\varepsilon_k \in \{-1,1\}$. Then, there exist positive constants δ^*, b_k, c_k, $k = 1,\ldots,n$, such that

$$\varepsilon_k b_k \delta^{2^k-1} \le a_k(\delta) \le \varepsilon_k c_k \delta^{2^k-1}. \tag{2.5}$$

Proof: The proof is by induction. For $n = 1$, we see that

$$\phi(\lambda,\delta) = \lambda + \delta \det A$$

and thus (2.5) holds with $b_1 = c_1 = |\det A|$. Now, assume that the result is true for $n = 1,\ldots,r$, and consider the case when A is an $(r+1)\times(r+1)$ matrix. Let

$$A = \left[\begin{array}{c|c} \bar{A} & b \\ \hline a & \gamma \end{array}\right].$$

Then

$$\phi(\lambda,\delta) = \det(E(\delta)A + \lambda I)$$

$$= \det\left[\begin{array}{c|c} \bar{E}(\delta)\bar{A} + \lambda I & \bar{E}(\delta)b \\ \hline \delta^{2^r} a & \delta^{2^r}\gamma + \lambda \end{array}\right]$$

$$= \lambda \det(\bar{E}(\delta)\bar{A} + \lambda I) + \delta^{2^r} \det\left[\begin{array}{c|c} \bar{E}(\delta)\bar{A} + \lambda I & \bar{E}(\delta)b \\ \hline a & \gamma \end{array}\right]$$

Denoting the first term by $\lambda\bar{\phi}(\lambda,\delta)$ and the second term by $\chi(\lambda,\delta)$, we note, from the induction hypothesis, that there exist $\bar{\delta}^*$, $\bar{b}_k$, $\bar{c}_k$, $k = 1,\ldots,r$ such that

$$\bar{\phi}(\lambda,\delta) = \lambda^r + \bar{a}_1(\delta)\lambda^{r-1} + \cdots + \bar{a}_r(\delta) \tag{2.6}$$

and $\qquad \varepsilon_k \bar{b}_k \delta^{2^{k}-1} \leq \bar{a}_k(\delta) \leq \varepsilon_k \bar{c}_k \delta^{2^{k}-1}$ for $k = 1,\ldots,r$, and δ in $(0,\bar{\delta}^*)$.

Also

$$\chi(\lambda,\delta) = \delta^{2^{r+1}-1}\, \varepsilon_{r+1} |\det A| + \delta^{2^{r}} \omega(\lambda,\delta) \tag{2.7}$$

where $\omega(\lambda,\delta)$ is a polynomial of degree r in λ, whose coefficients are polynomials in the variable δ, $\omega(0,\delta) = 0$ for all δ. Now, noting that $\phi(\lambda,\delta)$ is the sum of λ times (2.6) and (2.7), we have our result.

A consequence of Lemma 2.3 is the following:

<u>Theorem 2.4</u>: Let A be an $n \times n$ matrix in $\mathcal{P}$, and let

$$\text{sign} \det A(k) = \varepsilon_k, \quad k = 1,\ldots,n,$$

where $\varepsilon_k \in \{-1,+1\}$. Then

(i) for some positive number $\delta^* > 0$ and all δ in $(0,\delta^*)$

$$\inf_{\lambda \geq 0} \phi(\lambda,\delta) > 0$$

if and only if $\varepsilon_k = +1$ for all $k = 1,\ldots,n-1$.

(ii) If $\varepsilon_p = -1$, then there exists a $\delta^* > 0$ such that for all δ in $(0,\delta^*)$, $\inf_{\lambda \geq 0} \phi(\lambda,\delta) < 0$.

<u>Proof</u>: Now, if $\varepsilon_k = +1$ for every k, then from Lemma 2.3, $\phi(\lambda,\delta) \geq \lambda^n + \sum_{k=1}^{n} b_k \delta^{2^k-1}\lambda^{n-k}$ for some $b_k > 0$, all δ in $(0,\delta^*)$, and $\lambda \geq 0$. Hence, for $\lambda \geq 0$, $\phi(\lambda,\delta) > 0$. Thus, the if part of (i) follows. Now, let $\varepsilon_p = -1$ for some $p < n$, $\varepsilon_n = +1$. From Lemma 2.3, there exist positive numbers c_k and δ^* such that $\phi(\lambda,\delta) \leq \lambda^n + \sum_{k=1}^{n} \varepsilon_k c_k \delta^{2^k-1}\lambda^{n-k}$, for all δ in $(0,\delta^*)$. Now, consider the above for $\lambda(\delta) = \delta^{2^{(p-1)}+\frac{1}{2}}$, for δ in $(0,\delta^*)$. Then

$$\phi(\lambda(\delta),\delta) = \sum_{k=0}^{n} \varepsilon_k c_k \delta^{2^k-1}(\delta^{2^{p-1}+\frac{1}{2}})^{n-k} = \sum_{k=0}^{n} \varepsilon_k c_k \delta^{\omega(k)}$$

where $\omega(x) = (2^x - 1) + (n - x)(2^{(p-1)} + \frac{1}{2})$ for all $x \geq 0$, and $\varepsilon_0 = c_0 = +1$. We note that $\omega(x)$ is convex, and that $\omega(p+1) - \omega(p) = 2^{p-1} - \frac{1}{2} \geq \frac{1}{2}$ and $\omega(p-1) - \omega(p) = \frac{1}{2}$. Hence $\omega(k) - \omega(p) \geq \frac{1}{2}$ for all $k \neq p$. Hence, we have

$$\phi(\lambda(\delta),\delta) = \delta^{\omega(p)}\{\varepsilon_p c_p + \sum_{k \neq p} \varepsilon_k c_k \delta^{\omega(k)-\omega(p)}\}.$$

Since $\varepsilon_p c_p < 0$, we note that $\phi(\lambda(\delta),\delta)$ is negative for all sufficiently small $\delta > 0$, and thus (ii) follows. Also, the only if part of (i) follows from the above.

§3. The Main Theorems:

In this section we prove our main results that if a class $\mathcal{A}$ of matrices can see any matrix in $\mathcal{P}$, then $|\mathcal{A}| \geq 2^{n-1}$, and also that any finite subset of $\mathcal{P}$ can be seen from a subset $\mathcal{A}$ of cardinality at most 2^{n-1}.

For this purpose, consider the class of diagonal matrices

$$\mathfrak{D} = \{D: D_{ii} = +1 \text{ or } -1,\ \ D_{ij} = 0,\ i \neq j\} \subset \mathcal{P}.$$

Then, as can be readily confirmed, $|\mathfrak{D}| = 2^{n-1}$.

We are now ready to prove our main theorem:

Theorem 3.1: Let $\mathcal{A} \subset \mathcal{P}$ be a subset of m $n \times n$ matrices. Assume that $m < 2^{n-1}$. Then, there exists a matrix B in $\mathcal{P}$ such that

$$\min_{0 \leq t \leq 1} \det[(1-t)A + tB] < 0 \quad \text{for all } A \text{ in } \mathcal{A}.$$

Proof: From Lemma 2.1, since $\mathcal{P}^0$ is open and dense, there exists a matrix C such that $CA \in \mathcal{P}^0$ for each $A \in \mathcal{A}$. Now, since $m < 2^{n-1}$, there exists a matrix $D \in \mathfrak{D}$ such that if $A \in \mathcal{A}$ then DCA has at least one negative leading principal minor. Hence, by Theorem 2.4(ii), for each $A \in \mathcal{A}$, there exists a positive number δ_A such that for every $\delta \in (0,\delta_A]$,

$$\min_{\lambda \geq 0} \det[E(\delta)DCA + \lambda I] < 0.$$

Letting $\delta^* = \min \delta_A > 0$, and $B = (E(\delta^*)DC)^{-1}$, we obtain the desired result from Theorem 2.2.

We now prove our other main result:

Theorem 3.2: Let $\mathcal{A} \subset \mathcal{P}$ be a finite set of matrices. Then, there exist $m \leq 2^{n-1}$ matrices $B^1,\ldots,B^m$ in $\mathcal{P}$ such that if $A \in \mathcal{A}$, then

$$\min_{0 \leq t \leq 1} \det[(1-t)A + tB^j] > 0 \quad \text{for some } j.$$

Proof: From Lemma 2.1, there exists a matrix C such that $CA \in \mathcal{P}^0$ for each A in $\mathcal{A}$. Now, classify the matrices in $\{CA: A \in \mathcal{A}\}$ by the equivalence relation $A \sim B$ if and only if $\operatorname{sign} \det A(k) = \operatorname{sign} \det B(k)$ for each $k = 1,\ldots,n-1$. There are at most 2^{n-1} classes generated by the equivalence relation. Now, let $\mathcal{A}_1,\ldots,\mathcal{A}_m$ be these classes. Consider $\mathcal{A}_r$. For $A \in \mathcal{A}_r$, $k = 1,\ldots,n-1$, define

$$\varepsilon_k = \operatorname{sign} \det A(k)$$

and the matrix $D \in \mathcal{D}$ such that $D_{ii} = \varepsilon_{k-1}\varepsilon_k$ with $\varepsilon_0 = 1$. Then, $\operatorname{sign} \det[DA](k) = \varepsilon_k^2 > 0$, for $k = 1,\ldots,n$, and A in $\mathcal{A}_r$. From Theorem 2.4, there exists a δ_A such that for all δ in $(0,\delta_A]$, $\min_{\lambda \geq 0} \det[E(\delta)DA + \lambda I] > 0$. Now, define $\delta^* = \min_{A \in \mathcal{A}_r} \delta_A$, and we note that all matrices in $\mathcal{A}_r$ can be seen from $B_r = [E(\delta^*)D]^{-1}$ (from Theorem 2.2), and we have our result.

As is evident from the proof of Theorem 3.2, the "sign matrices" in $\mathcal{D}$ may not be sufficient, since some scaling matrix $E(\delta)$ is also involved. It can be readily verified that the matrices $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ and $\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ can see all 2×2 matrices $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ in $\mathcal{P}$. The former can see any matrix for which $a+d \geq 0$, and the latter can see any matrix with $(a+d) < 0$. But we conjecture that 2^{n-1} matrices are sufficient to see all $n \times n$ matrices in $\mathcal{P}$. Our approach appears to fail for this case. In the space of 3×3 matrices, for example, the 3×3 sign matrices in $\mathfrak{D}$, namely

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix}$$

cannot see the matrix $\begin{bmatrix} -39 & 20 & 20 \\ 20 & -39 & 20 \\ 20 & 20 & -39 \end{bmatrix}$. (We are grateful to Todd [9] for this example.) Thus, in our approach, scaling is important.

References

[1] Eaves, B. C., "A Short Course in Solving Equations with PL Homotopies," SIAM-AMS Proceedings, 9 (1976), 73-143.

[2] Eaves, B. C. and R. Saigal, "Homotopies for Computation of Fixed Points in Unbounded Regions," Mathematical Programming, 3 (1972), 225-237.

[3] Kojima, M. and R. Saigal, "A Study of PC^1 Homeomorphisms on Subdivided Polyhedrons," Submitted to SIAM Math. Analysis, April 1978.

[4] Kojima, M. and R. Saigal, "On the Relationship between Conditions that Insure a Piecewise Linear Mapping is a Homeomorphism," Submitted to Math. of O.R., May, 1978.

[5] Saigal, R., "On Paths Generated by Fixed Point Algorithms," Mathematics of Operations Research, 1 (1976), 359-380.

[6] Saigal, R., "On Piecewise Linear Approximations to Smooth Mappings," Submitted to Math. of O.R., Dec. 1977.

[7] Todd, M. J., "On the Jacobian of a Function at a Zero Computed by a Fixed Point Algorithm," To appear Math. of O.R.

[8] Todd, M. J., The Computation of Fixed Points and Applications, Springer-Verlag, New York, 1976.

[9] Todd, M. J., Private Communication, June 1, 1978.