The Fibre of P-matrices: The Recursive Construction of All Matrices with Positive Principal Minors

Michael J. Tsatsomeros and Yueqiao Faith Zhang

Department of Mathematics and Statistics
Washington State University
Pullman, WA 99164
(tsat@wsu.edu, yueqiao.zhang@wsu.edu)

March 20, 2019

Abstract

P-matrices have positive principal minors and include many well-known matrix classes (positive definite, totally positive, M-matrices etc.) How does one construct a generic P-matrix? Specifically, is there a characterization of P-matrices that lends itself to the tractable construction of every P-matrix? To answer these questions positively, a recursive method is employed that is based on a characterization of rank-one perturbations that preserve the class of P-matrices.

Keywords: P-matrix, principal minors, rank-one perturbation

AMS Subject Classifications: 15A48, 15A15, 15A23, 68Q15
1 Introduction

An \( n \times n \) complex matrix \( A \in M_n(\mathbb{C}) \) is called a P-matrix if all the principal minors of \( A \) are positive. The P-matrices encompass such notable classes as the (hermitian) positive definite matrices, the M-matrices and the totally positive matrices. The study of P-matrices originated in the context of the classes above in the work of Ostrowski, Fan, Koteljanskij, Gantmacher & Krein, Taussky, Fiedler & Ptak, Tucker, and Gale & Nikaido, among others. Indeed, P-matrices play an important role in a wide range of applications, including the linear complementarity problem, global univalence of maps, linear differential inclusion problems, interval matrices and computational complexity; see [11] for a review.

Of particular interest are two related problems:

1. Recognize whether or not a given matrix is a P-matrix (P-problem).
2. Provide a constructive characterization of P-matrices, namely, a method that can generate every P-matrix, apart from the well-known classes mentioned above or the various P-matrix construction methods reviewed in [11].

Both problems are central to computational challenges arising in the Linear Complementarity Problem (LCP)\(^1\) (see e.g., [2]), as exemplified by the following facts:

- LCP has a unique solution if and only if the coefficient matrix is a P-matrix [2].
- The problem of deciding if a given matrix is a P-matrix is co-NP-complete [3].
- The complexity of solving the LCP when the coefficient matrix is a P-matrix is presently unknown. If the problem of solving the LCP with a P-matrix were NP-hard, then the complexity classes NP and co-NP would coincide [7].

One interesting method to construct real P-matrices is to form products \( BC^{-1} \), where \( B, C \in M_n(\mathbb{R}) \) are strictly row diagonally dominant matrices\(^2\) with positive diagonal entries. This was first observed in [6] and is also the subject of study in [8] and [9], where such matrices are referred to as ‘hidden prdd’. At an Oberwolfach meeting [10], C.R. Johnson stated the above result and raised the question whether or not all real P-matrices can be factored this way, namely, they are hidden prdd. A counterexample was provided in [8], along with a polynomial algorithm to detect hidden prdd matrices. Further study of related classes was pursued in [9].

Thus, we still seek to understand the ‘fibre’ of a P-matrix and solve problem (2) above. We shall see that its solution is intrinsically related to problem (1). Using a recursion

---

\(^1\)Recall that the LCP defined by \( A \in M_n(\mathbb{R}) \) and a vector \( q \in \mathbb{R}^n \), is the problem of finding (entrywise) nonnegative vectors \( x \) and \( y \) such that \( y = Ax + q \) and \( x^Ty = 0 \).

\(^2\)\( B = [b_{ij}] \in M_n(\mathbb{C}) \) is strictly row diagonally dominant if \( |b_{ii}| > \sum_{j \neq i} |b_{ij}| \) (i = 1, 2, \ldots, n).
based on rank-one perturbations of P-matrices, we shall be able to reverse the steps of
a recursive algorithm that detects P-matrices [12] in order to construct every P-matrix.

2 Notation and preliminaries

For a positive integer $n$, let $\langle n \rangle = \{1, 2, \ldots, n\}$. For $\alpha \subseteq \langle n \rangle$, $|\alpha|$ denotes the cardinality of $\alpha$ and $\overline{\alpha} = \langle n \rangle \setminus \alpha$. For $\alpha \subseteq \langle n \rangle$ with $|\alpha| = k$ and its elements arranged in ascending order, we let $x[\alpha]$ denote the vector in $\mathbb{C}^k$ obtained from the entries of $x \in \mathbb{C}^n$ indexed by $\alpha$. For $\alpha \subseteq \langle n \rangle$ with $|\alpha| = k$ and its elements arranged in ascending order, we let $x[\alpha]$ denote the vector in $\mathbb{C}^k$ obtained from the entries of $x \in \mathbb{C}^n$ indexed by $\alpha$. We let $A[\alpha, \beta]$ be the submatrix of $A \in M_n(\mathbb{C})$ whose rows and columns are indexed by $\alpha, \beta \subseteq \langle n \rangle$, respectively; the elements of $\alpha, \beta$ are assumed to be in ascending order. When a row or column index set is empty, the corresponding submatrix is considered vacuous and by convention has determinant equal to 1. We abbreviate $A[\alpha, \alpha]$ by $A[\alpha]$ and refer to it as a principal submatrix of $A$. Thus, a matrix $A \in M_n(\mathbb{C})$ is a P-matrix if $\det A[\alpha] > 0$ for all $\alpha \subseteq \langle n \rangle$.

Given $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ such that $A[\alpha]$ is invertible, $A/A[\alpha]$ denotes the Schur complement of $A[\alpha]$ in $A$, that is,

$$A/A[\alpha] = A[\overline{\alpha}] - A[\overline{\alpha}, \alpha]A[\alpha]^{-1}A[\alpha, \overline{\alpha}].$$

The following result is well-known; see e.g., [1], [11].

**Lemma 2.1.** Let $A \in M_n(\mathbb{C})$ be a P-matrix and $\alpha \subseteq \langle n \rangle$. Then $A/A[\alpha]$ is a P-matrix.

3 Rank-one perturbations in constructing P-matrices

The following theorem provides a characterization of P-matrices based on the Schur complement of a diagonal entry.

**Theorem 3.1** ([11], [12]). Let $A \in M_n(\mathbb{C})$, $\alpha \subseteq \langle n \rangle$ with $|\alpha| = 1$. Then $A$ is a P-matrix if and only if $A[\alpha]$, $A[\overline{\alpha}]$, and $A/A[\alpha]$ are P-matrices.

The theorem above was used in [12] to develop the algorithm P-TEST that detects whether a given matrix $A \in M_n(\mathbb{C})$ is a P-matrix or not. P-TEST proceeds recursively to examine a binary tree of principal submatrices and Schur complements of decreasing orders for membership in the P-matrices. This is until the matrices to be actually checked are all $1 \times 1$ (scalars). To date, this is the most efficient known comprehensive algorithm for the detection of complex P-matrices. The algorithm is of $O(2^n)$ time complexity [4, 12], representing an improvement by a factor of $n^3$ over calculating all the principal minors of an $n \times n$ matrix. The following is an implementation of this recursive algorithm for the P-problem suggested by Theorem 3.1.
Algorithm P-TEST P(A)

1. Input $A = [a_{ij}] \in M_n(C)$
2. If $a_{11} \neq 0$ output 'Not a P-matrix’ stop
3. Evaluate $A/A[\alpha]$, where $\alpha = \{1\}$
4. Call $P(A[\alpha])$ and $P(A/A[\alpha])$
5. Output ‘This is a P-matrix’

Subsequent to the results in [12], the ideas in Theorem 3.1 and P-TEST were used in [4] for the purposes of computing all the principal minors of a given arbitrary matrix in an efficient and systematic way. They were also used in [5] to solve the Principal Minor Assignment problem, namely, to construct an $n \times n$ matrix (or decide that none exists) with a given set of $2^n - 1$ principal minors.

Next, we will observe that the above ideas also allow one to construct recursively an arbitrary member (and thus all members) of the class of P-matrices of any given size.

Corollary 3.2. Let $\hat{A} \in M_n(C)$ be a P-matrix, $a \in C$ and $x, y \in C^n$. Then the following are equivalent.

(i) $A = \begin{bmatrix} a & x^T \\ -y & \hat{A} \end{bmatrix}$ is a P-matrix.

(ii) $a > 0$ and $\hat{A} + \frac{1}{a}yx^T$ is a P-matrix.

Proof. The equivalence follows from Theorem 3.1 and the fact that $\hat{A} + \frac{1}{a}yx^T$ is the Schur complement of $a = A[\{1\}]$ in $A$. 

Corollary 3.2 suggests the following recursive process to construct an $n \times n$ complex P-matrix, $n \geq 2$.

Algorithm P-CON

Choose $A_1 > 0$

For $k = 1: n - 1$, given the $k \times k$ P-matrix $A_k$

1. Choose $(x^{(k)}, y^{(k)}) \in C^k \times C^k$ and $a_k > 0$ such that $A_k + \frac{1}{a_k}y^{(k)}(x^{(k)})^T$ is a P-matrix

2. Form the $(k + 1) \times (k + 1)$ matrix $A_{k+1} = \begin{bmatrix} a_k & (x^{(k)})^T \\ -y^{(k)} & A_k \end{bmatrix}$

Output $A = A_n$ is a P-matrix

The recursive nature of every P-matrix is formally shown in the following theorem.

Theorem 3.3. Every matrix constructed via P-CON is a P-matrix. Conversely, every P-matrix $A \in M_n(C)$ can be constructed via P-CON.
Proof. By Corollary 3.2, each of the matrices $A_{k+1}$ ($k = 1, 2, \ldots, n - 1$) in P-CON, including $A_1$, is a P-matrix. We use induction to prove the converse. The base case is trivial. Assume every P-matrix in $M_{n-1}(\mathbb{C})$ can be constructed via P-CON. Let $A \in M_n(\mathbb{C})$ be any P-matrix partitioned as

$$A = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{22} \in M_{n-1}(\mathbb{C})$. By inductive hypothesis, $A_{22}$ is a P-matrix constructible via P-CON. Since $A$ is a P-matrix, by Corollary 3.2, $A/a_{11} = A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$ is a P-matrix and $A_n = A$ is constructible via P-CON with

$$a_{n-1} = a_{11} > 0, \quad x^{(n-1)} = A_{12}^T \text{ and } y^{(n-1)} = -A_{21}. \quad \square$$

4 Constructing P-matrices

Our further development and application of P-CON in this section will be guided by the following considerations.

1. The choice of $(x^{(k)}, y^{(k)}) \in \mathbb{C}^k \times \mathbb{C}^k$ and $a_k > 0$ in P-CON must be made such that $A_k + \frac{1}{a_k} y^{(k)}(x^{(k)})^T$ has positive principal minors. Given that the primary interest in applications concerns real P-matrices, we will offer an implementation of P-CON that constructs real P-matrices. Therefore, we execute Step 1 of P-CON by choosing a pair of real vectors $(x^{(k)}, y^{(k)})$ randomly, and subsequently choose $a_k > 0$ sufficiently large to ensure $A_k + \frac{1}{a_k} y^{(k)}(x^{(k)})^T$ is a P-matrix.

2. Theorem 3.3 and P-CON deal with complex P-matrices. To proceed with construction of non-real P-matrices it is implicit in the condition that $A_k + \frac{1}{a_k} y^{(k)}(x^{(k)})^T$ be a P-matrix that, although the $j$-th entries of $x^{(k)}$, $y^{(k)}$ can be non-real, their product must be real. We will illustrate such a construction in Example 4.5.

Pursuant to consideration (1) above, the lemma and theorem that follow facilitate the choice of $a_k$ in the implementation of P-CON. First is the well-known Matrix Determinant Lemma with proof for completeness.

**Lemma 4.1.** Let $A \in M_n(\mathbb{C})$ be invertible and $y, x \in \mathbb{C}^n$. Then $\det(A + yx^T) = (x^T A^{-1}y + 1) \det(A)$.  

5
Proof. Since

\[
\begin{bmatrix}
I & 0 \\
x^T & 1
\end{bmatrix}
\begin{bmatrix}
I + yx^T & y \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
x^T & 1
\end{bmatrix}^{-1} =
\begin{bmatrix}
I & 0 \\
x^T & 1
\end{bmatrix}
\begin{bmatrix}
I + yx^T & y \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
x^T & 1
\end{bmatrix}
\]

we have

\[
\det \begin{bmatrix}
I + yx^T & y \\
0 & 1
\end{bmatrix} = \det \begin{bmatrix}
I & y \\
x^T & y + 1
\end{bmatrix},
\]

i.e., \(\det(I + yx^T) = x^T y + 1\). Thus,

\[
\det(A + yx^T) = \det(A) \det(I + A^{-1}yx^T) = (x^T A^{-1} y + 1) \det(A).
\]

\[
\square
\]

Theorem 4.2. Let \(A \in M_n(\mathbb{C})\) be a P-matrix, \(y, x \in \mathbb{C}^n\) and \(a > 0\). Then \(A + \frac{1}{a}yx^T\) is a P-matrix if and only if for every \(\alpha \subseteq \langle n \rangle\), we have \((x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \in \mathbb{R}\) and

\[
a > \max_{\alpha \subseteq \langle n \rangle} \left\{ - (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \right\}.
\]

Proof. For every \(\alpha \subseteq \langle n \rangle\), \(\det(A[\alpha]) > 0\) since \(A\) is a P-matrix. Notice that the principal submatrices of \(A + \frac{1}{a}yx^T\) are of the form \(A[\alpha] + \frac{1}{a}y[\alpha](x[\alpha])^T\).

First we prove sufficiency: Suppose that for every \(\alpha \subseteq \langle n \rangle\), \((x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \in \mathbb{R}\) and

\[
a > \max_{\alpha \subseteq \langle n \rangle} \left\{ - (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \right\}.
\]

By Lemma 4.1, for any \(\alpha \subseteq \langle n \rangle\), we have

\[
\det(A[\alpha] + \frac{1}{a}y[\alpha](x[\alpha])^T) = \left( \frac{1}{a} (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] + 1 \right) \det(A[\alpha]) > 0
\]

Therefore, \(A + \frac{1}{a}yx^T\) is a P-matrix.

Next we prove necessity. If \(A + \frac{1}{a}yx^T\) is a P-matrix, then for every \(\alpha \subseteq \langle n \rangle\),

\[
\det(A[\alpha] + \frac{1}{a}y[\alpha](x[\alpha])^T) = \left( \frac{1}{a} (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] + 1 \right) \det(A[\alpha]) > 0
\]

which necessitates that \((x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \in \mathbb{R}\) and \(\frac{1}{a} (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] + 1 > 0\), i.e.,

\[
a > \max_{\alpha \subseteq \langle n \rangle} \left\{ - (x[\alpha])^T (A[\alpha])^{-1} y[\alpha] \right\}.
\]

\[
\square
\]
Next, we incorporate the bound in Theorem 4.2 in Matlab code that constructs real P-matrices. We include the codes for both PCON and PTEST.

**PCON**

```matlab
function [A]=pcon(N)

% Input N is the size of the desirable real P-matrix to be generated

A=rand(1); % or A=abs(normrnd(0,1)); random 1x1 P-matrix
for j=1:N-1
    x=-1+2.*rand(j,1); % random entries in [-1,1]; or use x=normrnd(0,1,[j 1]);
    y=-1+2.*rand(j,1); % or use y=normrnd(0,1,[j 1]);
    [m,n]=size(A);
    v=1:n;
    a=0.01; % or a=abs(normrnd(0,1));
    for k=1:n
        C = nchoosek(v,k);
        [p,q]=size(C);
        for i=1:p
            B=A(C(i,:), C(i,:));
            b=-(x(C(i,:)).'inv(B)*y(C(i,:)));
            if b>a
                a=b;
            end
        end
    end
    a=1.01*a; % or a=(1+abs(normrnd(0,1)))*a;
    A=[a x.'; -y A];
end
```

**PTEST**

```matlab
function [r] = ptest(A)

% Return r=1 if 'A' is a P-matrix (r=0 otherwise).

n = length(A);
if ~A(1,1)>0, r = 0;
elseif n==1, r = 1;
else
    b = A(2:n,2:n);
    d = A(2:n,1)/A(1,1);
    c = b - d*A(1,2:n);
    r = ptest(b) & ptest(c);
end
```
Remark 4.3. The following remarks clarify the functionality of the implementation of P-CON.

- The complexity of P-CON is exponential in $n$ because of the computation of the lower bound for $a = a_k$ in Theorem 4.2, which requires a maximum be computed among all $n$-choose-$k$ principal submatrices of $A$ of size $k \times k$ ($k = 1, 2, \ldots, n$).

- In the provided code above, random choices are uniformly distributed; normal distribution commands are commended out.

- The lower bound for the parameter $a_k$ provided by Theorem 4.2 is strict, hence our choice of $a_k$ is larger than (but kept close to) the lower bound to minimize the chance of diagonal dominance.

- Given that P-matrices are preserved under positive scaling of the rows and columns, there is no loss of generality in restricting the (uniformly distributed) random choice of the entries of $x^{(k)}$ and $y^{(k)}$ to be in $[-1, 1]$.

- We have experimented with normal and uniform distributions for the random choice of the parameters and vector entries. We have, however, observed no discernible difference in the nature of the generated P-matrices.

- As desirable, in the experiments we have run, the matrices generated display no symmetry, no sign pattern, and no diagonal dominance, because none of these traits are imposed by P-CON.

We conclude by illustrating the functionality of P-CON with some generated examples of real P-matrices. We also include an example of a non-real P-matrix generated via P-CON.

Example 4.4. The first two examples are P-matrices generated by execution of P-CON with random variables that are uniformly distributed.

\[
\begin{bmatrix}
0.8944 & -0.1366 & 0.9951 & 0.6232 \\
0.0287 & 0.0101 & -0.7969 & -0.2183 \\
-0.7889 & 0.8908 & 0.0101 & 0.8127 \\
0.7249 & -0.0026 & -0.2578 & 0.5102
\end{bmatrix},
\begin{bmatrix}
5.5491 & 0.0613 & 0.6648 & 0.1950 & -0.3294 \\
0.4015 & 0.4512 & -0.6777 & 0.5162 & 0.7422 \\
0.0948 & 0.2984 & 40.4328 & -0.1956 & 0.2413 \\
0.1547 & -0.3711 & 0.6913 & 0.2783 & -0.1952 \\
0.2808 & 0.4117 & 0.2373 & -0.9657 & 0.6841
\end{bmatrix}
\]

In the next two examples random variables were normally distributed.

\[
\begin{bmatrix}
2.7122 & 1.1093 & -0.8637 & 0.0774 \\
1.2141 & 1.3140 & 0.3129 & -0.8649 \\
1.1135 & 0.0301 & 0.3185 & -1.7115 \\
0.0068 & 0.1649 & 0.1022 & 1.3703
\end{bmatrix},
\begin{bmatrix}
5.7061 & 2.5260 & 1.6555 & 0.3075 & -1.2571 \\
0.8655 & 1.7977 & -0.1332 & -0.7145 & 1.3514 \\
0.1765 & 0.2248 & 1.6112 & 0.9642 & 0.5201 \\
-0.7914 & 0.5890 & 0.0200 & 1.6021 & -0.9792 \\
1.3320 & 0.2938 & 0.0348 & 1.1564 & 0.8314
\end{bmatrix}
\]
Example 4.5. In this example, we explicitly illustrate the construction of a $4 \times 4$ non-real P-matrix using P-CON. The choices are made deliberately to satisfy the necessary conditions and are not randomly generated.

Choose $A_1 = 2$, $x^{(1)} = 2 - i$, $y^{(1)} = 4 + 2i$ and $a_1 = 1$ so that

$$A_1 + \frac{1}{a_1} y^{(1)}(x^{(1)})^T = 12$$

is a $1 \times 1$ P-matrix. Thus $A_2 = \begin{bmatrix} 1 & 2 - i \\ -4 - 2i & 2 \end{bmatrix}$ is a P-matrix.

Choose $x^{(2)} = \begin{bmatrix} -1 - 2i \\ 3 - i \end{bmatrix}$, $y^{(2)} = \begin{bmatrix} \frac{1}{2} - i \\ 3 + i \end{bmatrix}$ and $a_2 = 3$ so that

$$A_2 + \frac{1}{a_2} y^{(2)}(x^{(2)})^T = \begin{bmatrix} 0.1667 & 2.1667 - 2.1667i \\ -4.3333 - 4.3333i & 5.3333 \end{bmatrix}$$

is a P-matrix. Thus $A_3 = \begin{bmatrix} 3 & -1 - 2i & 3 - i \\ -\frac{1}{2} + i & 1 & 2 - i \\ -3 - i & -4 - 2i & 2 \end{bmatrix}$ is a P-matrix.

Choose $x^{(3)} = \begin{bmatrix} -i \\ 1 + i \end{bmatrix}$, $y^{(3)} = \begin{bmatrix} 2i \\ 1 - i \end{bmatrix}$ and $a_3 = 3.5$ so that

$$A_3 + \frac{1}{a_3} y^{(3)}(x^{(3)})^T = \begin{bmatrix} 3.5714 & -1.5714 - 1.4286i & 3.0000 - 0.8095i \\ -0.7857 + 0.7143i & 1.5714 & 2.0952 - 1.0952i \\ -3.0000 - 0.8095i & -4.1905 - 2.1905i & 1.9365 \end{bmatrix}$$

is a P-matrix. Thus $A_4 = \begin{bmatrix} 3.5 & -i & 1 + i & \frac{1}{3} \\ -2i & 3 & -1 - 2i & 3 - i \\ -1 + i & -\frac{1}{2} + i & 1 & 2 - i \\ \frac{2}{3} & -3 - i & -4 - 2i & 2 \end{bmatrix}$ is a P-matrix.

5 Discussion

Several years ago, Siegfried Rump asked the first author how to generate arbitrary P-matrices for the sake of testing P-matrix algorithms that were being developed. The desire was to avoid matrices that share the common traits of P-matrices belonging to the usually studied classes (positive definite, M-matrices, totally positive, lmrdd and other related matrix classes). This desire raised our interest in the problem of generating all P-matrices. Theorem 3.3 resolves this issue theoretically, allowing for the practical generation of arbitrary P-matrices. We provided an algorithmic implementation of
that can indeed produce arbitrary complex P-matrices, and provided Matlab code that can be used to generate arbitrary real P-matrices.

Acknowledgment The authors acknowledge with thanks the comments provided by the Handling Editor and an anonymous referee, which helped improve our results and clarify our efforts.

References


