Principal pivot transforms: properties and applications

Michael J. Tsatsomeros *
Department of Mathematics and Statistics
University of Regina
Regina, Saskatchewan
Canada S4S 0A2

Revised November 1999

Abstract

The principal pivot transform (PPT) is a transformation of the matrix of a linear system tantamount to exchanging unknowns with the corresponding entries of the right-hand side of the system. The notion of the PPT is encountered in mathematical programming, statistics and numerical analysis among other areas. The purpose of this paper is to draw attention to the main properties and uses of PPTs, make some new observations and motivate further applications of PPTs in matrix theory. Special consideration is given to PPTs of matrices whose principal minors are positive.

Key words: linear system, sweep operator, principal submatrix, P-matrix, iterative method

AMS subject classifications: 15A06, 15A09, 15-02, 90C33, 65U05

*Work supported by a grant from the Natural Sciences and Engineering Research Council of Canada.
E-mail: tsat@math.uregina.ca

1 Introduction

Suppose that $A \in M_n(\mathbb{C})$ (the n-by-n complex matrices) is partitioned in blocks as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{1.1}$$

and further suppose that A_{11} is an invertible submatrix. Consider then

$$B = \begin{pmatrix} (A_{11})^{-1} & -(A_{11})^{-1}A_{12} \\ A_{21}(A_{11})^{-1} & A_{22} - A_{21}(A_{11})^{-1}A_{12} \end{pmatrix}. \tag{1.2}$$

The matrices A and B are related as follows: If $x = (x_1^T, x_2^T)^T$ and $y = (y_1^T, y_2^T)^T$ in $\mathbb{C}^n$ are partitioned conformally to A, then

$$A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{if and only if} \quad B\begin{pmatrix} y_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_2 \end{pmatrix}.$$

The operation of obtaining B from A is encountered in several contexts, including mathematical programming, statistics and numerical analysis, which we outline below.

Tucker [23] considers a "combinatorial" equivalence relation among rectangular matrices, which is implicitly determined by a nonsingular (not necessarily principal) submatrix of A. The equivalence class of A comprises the *pivot(al) transforms* of A, namely, all matrices that are permutationally equivalent to B above. When the equivalence relation is determined by a principal submatrix, Tucker [24] uses the term *principal pivot(al) transform* and asserts that if A has positive principal minors (that is, if A is a P-matrix), then so does every principal pivot transform of A. Tucker's motivation for introducing combinatorial equivalence and studying principal pivot transforms is rooted in an effort to generalize Dantzig's simplex method from ordered to general fields. In turn, the domain-range relation between A and B observed by Tucker is later used by Cottle and Dantzig [5] as an important feature of the "principal pivoting algorithm" for the linear complementarity problem, when the coefficient matrix is a real P-matrix. In this algorithm, principal pivot transforms are used to exchange the role of basic and nonbasic variables of the problem and the fact that P-matrices are preserved is applied effectively. Principal pivot transforms have since found similar uses in the context of mathematical programming (see e.g., Pang [19]).

The principal pivot transform, under the name *sweep operator*, plays an important role in statistics, mainly because of conceptual and computational advantages it enjoys in solving least-squares regression problems. The vector b that minimizes $\|Xb - Y\|_2$, where X is $n \times k$ of full rank and Y is $n \times 1$, appears in the principal pivot transform of

$$A = \begin{pmatrix} X^TX & X^TY \\ Y^TX & Y^TY \end{pmatrix}$$

relative to its principal submatrix X^TX, which is given by

$$\begin{pmatrix} (X^TX)^{-1} & -(X^TX)^{-1}X^TY \\ Y^TX(X^TX)^{-1} & Y^TY - Y^TX(X^TX)^{-1}X^TY \end{pmatrix}.$$

Indeed $b = (X^TX)^{-1}X^TY$ is recognizable as the solution of the normal equations associated with this least-squares problem, and $Y^TY - Y^TX(X^TX)^{-1}X^TY$ is the residual sum of squares.

When $X^T X$ is not invertible, this process can be modified to produce particular solutions of the least-squares problem via the corresponding generalized inverse (see section 3 for a description). The first application of the above ideas in statistics seems to be contemporary to Tucker's results on combinatorial equivalence and is attributed to Efroymson [8]. For further references and the historical context in statistics, see Seber [20] and Goodnight [11]. Adaptations of the sweep operator are in use for a variety of statistical computations, for example in solving least-squares problems subject to linear constraints (Neytchev [17]) and in repeated computations of likelihood e.g., in Monte Carlo Markov chain models (see Meehan, Dempster, Brown [15], and Wolfinger, Tobias, Sall [26]).

A fundamental matrix factorization of the principal pivot transform turns up in a discussion of interval nonsingularity and P-matrices in Johnson and Tsatsomeros [13]. We review this factorization in Lemma 3.4 and take the opportunity to provide a proof valid for complex matrices of a result claimed in [13] (see Remark 5.4). In a related vein, Elsner and Szulc [9] introduce a generalization of P-matrices to block P-matrices and show that a certain subclass of the block P-matrices is invariant under principal pivot transformations.

The relation between A and B mentioned earlier prompted Stewart and Stewart [21] to refer to B as the *exchange* of A. The authors use exchanges in order to generate S-orthogonal matrices from hyperbolic Householder transformations, and then apply them to solve the mixed Cholesky updating/downdating problem. In [21] it is also noted that this method of construction of S-orthogonal matrices is a folk result in circuit theory and a reference to Belovitch [1] is made for a special case.

The principal pivot transform also appears under the term *gyration* in Duffin, Hazony, and Morrison [7], and is mentioned in a survey of Schur complements by Cottle [4].

The above varied interest in principal pivot transforms and the luck of a readily accessible comprehensive reference motivate us here to collect, study and present proofs of their main properties. We also initiate a discussion on determinants, characteristic polynomials and eigenvalues of principal pivot transforms from a matrix-theoretic prospective. The relation and parallelism of principal pivot transforms to inversion is also considered, as well as a potential application to iterative techniques for solving linear systems. Last, we discuss matrix classes that are invariant under principal pivot transformations, including the aforementioned P-matrices and S-orthogonal matrices, as well as an interesting subclass of the P-matrices introduced in [19] that contains the M-matrices and their inverses. In addition to the brief account of principal pivot transforms given so far, the subsequent commentary contains more information regarding our sources and motivation.

2 Notation and preliminaries

Let n be a positive integer and $A \in M_n(\mathbb{C})$. The i-th entry of a vector x is denoted by $x(i)$. In the remainder the following notation is also used:

- $\langle n \rangle = \{1, 2, \ldots, n\}$. For any $\alpha \subseteq \langle n \rangle$, the cardinality of α is denoted by $|\alpha|$ and $\bar{\alpha} = \alpha \backslash \langle n \rangle$.
- $A[\alpha, \beta]$ is the submatrix of A whose rows and columns are indexed by $\alpha, \beta \subseteq \langle n \rangle$, respec-

tively; the elements of α, β 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 $A[\alpha]$.

- $A/A[\alpha]$ is the *Schur complement* of an invertible principal submatrix $A[\alpha]$ in A, that is, $A/A[\alpha] = A[\bar{\alpha}] - A[\bar{\alpha}, \alpha]A[\alpha]^{-1}A[\alpha, \bar{\alpha}]$. It is known that $\det(A/A[\alpha]) = \det A/\det(A[\alpha])$.
- $\sigma(A)$ is the spectrum and $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$ is the spectral radius of A.
- $\mathrm{diag}(d_1, \ldots, d_n)$ is the diagonal matrix in $M_n(\mathbb{C})$ with diagonal entries $d_1, \ldots, d_n$.

Definition 2.1 Given $\alpha \subseteq \langle n \rangle$ and provided that $A[\alpha]$ is invertible, we define the *principal pivot transform* (PPT) *of* $A \in M_n(\mathbb{C})$ *relative to* α as the matrix $\mathrm{ppt}\,(A, \alpha)$ obtained from A by replacing

$$A[\alpha] \text{ by } A[\alpha]^{-1}, \quad A[\alpha, \bar{\alpha}] \text{ by } -A[\alpha]^{-1}A[\alpha, \bar{\alpha}],$$
$$A[\bar{\alpha}, \alpha] \text{ by } A[\bar{\alpha}, \alpha]A[\alpha]^{-1}, \quad \text{and } A[\bar{\alpha}] \text{ by } A/A[\alpha].$$

By convention, if $\alpha = \emptyset$, then $\mathrm{ppt}\,(A, \alpha) = A$.

We continue with two comments relevant to the subsequent discussion. First, as with Schur complements (see e.g., Ouellette [18]), the notion of a PPT can be extended to the case of non-invertible principal submatrices by considering generalized inverses. Some work in this direction is presented in Meenakshi [16]. Second, the PPT is related but distinct from the following block representation of the inverse (see [18, (1.9)]): Given an invertible $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ such that $A[\alpha]$ and $A[\bar{\alpha}]$ are invertible, A^{-1} is obtained from A by replacing

$$A[\alpha] \text{ by } (A/A[\bar{\alpha}])^{-1}, \quad A[\alpha, \bar{\alpha}] \text{ by } -A[\alpha]^{-1}A[\alpha, \bar{\alpha}](A/A[\alpha])^{-1},$$
$$A[\bar{\alpha}, \alpha] \text{ by } (A/A[\alpha])^{-1}A[\bar{\alpha}, \alpha]A[\alpha]^{-1}, \quad \text{and } A[\bar{\alpha}] \text{ by } (A/A[\alpha])^{-1}.$$

3 Basic properties of principal pivot transforms

We begin with a formal statement of the basic domain-range exchange property of $\mathrm{ppt}\,(A, \alpha)$, and include a proof sketch for the sake of completeness.

Theorem 3.1 *Let $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ so that $A[\alpha]$ is invertible. Given a pair of vectors $x, y \in \mathbb{C}^n$, define $u, v \in \mathbb{C}^n$ by $u[\alpha] = y[\alpha]$, $u[\bar{\alpha}] = x[\bar{\alpha}]$, $v[\alpha] = x[\alpha]$, $v[\bar{\alpha}] = y[\bar{\alpha}]$. Then $B = \mathrm{ppt}\,(A, \alpha)$ is the unique matrix with the property that for every such x, y, $y = Ax$ if and only if $Bu = v$. Moreover,* $\mathrm{ppt}\,(B, \alpha) = A$.

Proof. Consider the permutation matrix P for which

$$Px = \begin{pmatrix} x[\alpha] \\ x[\bar{\alpha}] \end{pmatrix} \quad \text{and} \quad PAP^T = \begin{pmatrix} A[\alpha] & A[\alpha, \bar{\alpha}] \\ A[\bar{\alpha}, \alpha] & A[\bar{\alpha}] \end{pmatrix}.$$

By the construction outlined in Definition 2.1 and on letting $B = \text{ppt}(A,\alpha)$, we have

$$PBP^T = \begin{pmatrix} A[\alpha]^{-1} & -A[\alpha]^{-1}A[\alpha,\bar{\alpha}] \\ A[\bar{\alpha},\alpha]A[\alpha]^{-1} & A/A[\alpha] \end{pmatrix}.$$

Then, with u and v as prescribed, it can be easily verified that $PAP^T(Px) = Py$ if and only if $Pv = PBP^T(Pu)$, or equivalently, $Ax = y$ if and only if $Bu = v$. To show uniqueness, suppose that $B'u = v$ if and only if $Ax = y$. Then $(B - B')u = 0$ for all u such that $u[\alpha] = y[\alpha] = A[\alpha]x[\alpha] + A[\alpha,\bar{\alpha}]x[\bar{\alpha}]$ and $u[\bar{\alpha}] = x[\bar{\alpha}]$. As $A[\alpha]$ is invertible and x is chosen freely, it follows that $(B - B')u = 0$ for all $u \in \mathbb{C}^n$, that is $B = B'$. To see that $\text{ppt}(B,\alpha) = A$, notice that $\text{ppt}(B,\alpha)x = Ax$ for all $x \in \mathbb{C}^n$. ∎

In [23] it is mentioned that $A^{-1} \in M_n(\mathbb{C})$ can be found with a sequence of at most n PPTs (and by interchanging rows or columns if needed). For example, as indicated in the next theorem, in certain cases the inverse of a matrix is the outcome of consecutive PPTs.

Theorem 3.2 *Let $A \in M_n(\mathbb{C})$ and suppose that there exists a partition of $\alpha \subseteq \langle n \rangle$ into subsets α_i, $i = 1,2,\ldots,k$ so that the sequence of matrices*

$$A_0 = A, \quad A_i = \text{ppt}(A_{i-1},\alpha_i), \; i = 1,2,\ldots,k$$

is well defined (i.e., the matrices $A_{i-1}[\alpha_i]$ are invertible). Then $\text{ppt}(A,\alpha) = A_k$. *In particular, if $\alpha = \langle n \rangle$, then A is invertible and $A^{-1} = A_k$.*

Proof. By Theorem 3.1 (and the notation thereof) applied to each of the A_i in sequence, and since the α_i are mutually disjoint sets whose union is α, we have that $Ax = y$ if and only if $A_k u = v$ for all $x, y \in \mathbb{C}^n$. It follows by the uniqueness of a PPT that $A_k = \text{ppt}(A,\alpha)$. ∎

Recall that a (1)-inverse of a rectangular matrix X is a matrix Y such that $XYX = X$ (see Campbell and Meyer [3]). The following connection between PPTs and (1)-inverses of positive semi-definite matrices is observed in [11].

Proposition 3.3 *Consider $A \in M_n(\mathbb{C})$ so that $A[\langle n \rangle,\alpha]$ consists of $|\alpha|$ linearly independent columns that span the column space of A. Then* $\text{ppt}(A^TA,\alpha)$ *is an* (1)*-inverse of A^TA.*

Proof. Without loss of generality, suppose that $A = [A_1|A_2] \in M_n(\mathbb{C})$, where A_1 consists of $k \le n$ linearly independent columns of A that span its column space. Consider the matrices

$$A^TA = \begin{pmatrix} A_1^TA_1 & A_1^TA_2 \\ A_2^TA_1 & A_2^TA_2 \end{pmatrix} \text{ and } B = \text{ppt}(A^TA,\langle k \rangle) = \begin{pmatrix} (A_1^TA_1)^{-1} & -(A_1^TA_1)^{-1}A_1^TA_2 \\ A_2^TA_1(A_1^TA_1)^{-1} & 0 \end{pmatrix}.$$

(Notice that the trailing principal submatrix of $\text{ppt}(A^TA,\langle k \rangle)$ vanishes.) It is then easily verifiable that $A^TABA^TA = A^TA$. ∎

For a survey of results on generalized inverses of partitioned matrices see [18, section 4.1].

To study further the basic properties of $\text{ppt}(A,\alpha)$, we continue with a useful observation that appears implicitly in the proof of [23, Theorem 4] and in [13].

Lemma 3.4 *Let $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ so that $A[\alpha]$ is invertible. Let T_1 be the matrix obtained from the identity by setting the diagonal entries indexed by α equal to 0. Let $T_2 = I - T_1$ and consider the matrices $C_1 = T_2 + T_1 A,\ C_2 = T_1 + T_2 A$. Then $\operatorname{ppt}(A, \alpha) = C_1 C_2^{-1}$.*

Proof. Without loss of generality, we can assume that $\alpha = \langle k \rangle$. (Otherwise we can apply our argument to a permutation similarity of A). Observe then that

$$C_1 = \begin{pmatrix} I & 0 \\ A[\bar{\alpha}, \alpha] & A[\bar{\alpha}] \end{pmatrix} \text{ and } C_2 = \begin{pmatrix} A[\alpha] & A[\alpha, \bar{\alpha}] \\ 0 & I \end{pmatrix}$$

and thus

$$\begin{aligned} C_1 C_2^{-1} &= \begin{pmatrix} I & 0 \\ A[\bar{\alpha}, \alpha] & A[\bar{\alpha}] \end{pmatrix} \begin{pmatrix} A[\alpha]^{-1} & -A[\alpha]^{-1} A[\alpha, \bar{\alpha}] \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} A[\alpha]^{-1} & -A[\alpha]^{-1} A[\alpha, \bar{\alpha}] \\ A[\bar{\alpha}, \alpha] A[\alpha]^{-1} & A/A[\alpha] \end{pmatrix} = \operatorname{ppt}(A, \alpha). \end{aligned}$$

■

Definition 3.5 In the notation of Lemma 3.4, we refer to $\operatorname{ppt}(A, \alpha) = C_1 C_2^{-1}$ as the *basic factorization* of $\operatorname{ppt}(A, \alpha)$.

In connection to a remark added in proof in [23], the following result sheds more light on the combinatorial relationship between a matrix and (the basic factorization of) its PPTs.

Theorem 3.6 *Let $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ so that $A[\alpha]$ is invertible. Let $B = \operatorname{ppt}(A, \alpha)$ and I be the identity in $M_n(\mathbb{C})$. Then there exists a permutation matrix $P \in M_{2n}(\mathbb{C})$ such that*

$$\begin{pmatrix} -B & I \end{pmatrix} P \begin{pmatrix} I \\ A \end{pmatrix} = 0. \tag{3.1}$$

Moreover, if T_1 and T_2 are as in Lemma 3.4, then

$$P = \begin{pmatrix} T_1 & T_2 \\ T_2 & T_1 \end{pmatrix}.$$

Conversely, if $B = \operatorname{ppt}(A, \alpha)$ for some $\alpha \subseteq \langle n \rangle$, then (3.1) holds for an appropriately defined permutation matrix $P \in M_{2n}(\mathbb{C})$.

Proof. In the notation of Lemma 3.4, we have that $B = \operatorname{ppt}(A, \alpha)$ if and only if

$$\begin{pmatrix} -B & I \end{pmatrix} \begin{pmatrix} C_2 \\ C_1 \end{pmatrix} = 0.$$

The claims of the theorem follow by substituting $C_1 = T_2 + T_1 A$ and $C_2 = T_1 + T_2 A$. That P as above is a permutation matrix follows from the definition of T_1 and T_2. ■

Example 3.7 To illustrate the definitions and observations so far, let $\alpha = \{1, 3\}$ so that

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 8 & 1 \end{pmatrix} \quad \text{and} \quad B = \operatorname{ppt}(A, \alpha) = \begin{pmatrix} -1 & -6 & 1 \\ -1 & -5 & 1 \\ 2 & 4 & -1 \end{pmatrix}.$$

Notice the exchange taking place relative to the index set α in the equations

$$A\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}4\\2\\11\end{pmatrix}\quad\text{and}\quad B\begin{pmatrix}4\\1\\11\end{pmatrix}=\begin{pmatrix}1\\2\\1\end{pmatrix}.$$

The basic factorization of B is $C_1C_2^{-1}$, where

$$C_1=\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix}\quad\text{and}\quad C_2=\begin{pmatrix}1&2&1\\0&1&0\\2&8&1\end{pmatrix}.$$

Also if $\beta=\{2\}$, then

$$\mathrm{ppt}\,(B,\beta)=\begin{pmatrix}.2&1.2&-.2\\-.2&-.2&.2\\1.2&-.8&-.2\end{pmatrix}=A^{-1}.$$

Theorem 3.8 *Let $A\in M_n(\mathbb{C})$ and $\alpha\subseteq\langle n\rangle$ so that $A[\alpha]$ is invertible. Then*

(i) $\det(\mathrm{ppt}\,(A,\alpha))=\dfrac{\det A[\bar\alpha]}{\det A[\alpha]}$, *and*

(ii) if in addition $A[\bar\alpha]$ is invertible, $\mathrm{ppt}\,(A,\alpha)^{-1}=\mathrm{ppt}\,(A,\bar\alpha)$.

Proof. Let $C_1C_2^{-1}$ be the basic factorization of $\mathrm{ppt}\,(A,\alpha)$. The conclusions follow, respectively, from Lemma 3.4 and by directly verifying that $\mathrm{ppt}\,(A,\alpha)^{-1}=C_2C_1^{-1}=\mathrm{ppt}\,(A,\bar\alpha)$. ∎

Note that invertibility of $\mathrm{ppt}\,(A,\alpha)$ does not necessarily imply invertibility of A [16]. A simple counterexample is provided by

$$A=\begin{pmatrix}1&2\\1&2\end{pmatrix}\quad\text{and}\quad\mathrm{ppt}\,(A,\{1\})=\begin{pmatrix}1&-2\\1&0\end{pmatrix}.$$

4 Eigenvalues of principal pivot transforms

We continue with what to our knowledge are new observations on the eigenvalues of PPTs. It turns out that the process of finding a PPT does not, in general, correspond to an easily describable mapping of the eigenvalues. Somewhat elegant descriptions of this mapping, which seem challenging to analyze practically, are mentioned in the following theorem.

Theorem 4.1 *Let $A\in M_n(\mathbb{C})$ and $\alpha\subseteq\langle n\rangle$ so that $A[\alpha]$ is invertible. Let $C_1C_2^{-1}$ be the basic factorization of* $\mathrm{ppt}\,(A,\alpha)$. *Then the following are equivalent:*

(i) $\lambda\in\sigma(\mathrm{ppt}\,(A,\alpha))$

(ii) λ is a finite eigenvalue of the matrix pencil $C_1-\lambda C_2$.

When, in addition, $\lambda\neq0$, then the following condition is also equivalent to (i) and (ii):

(iii) $A - D$ *is singular, where* $D = \operatorname{diag}(d_1, \ldots, d_n)$ *with* $d_i = \lambda^{-1}$ *if* $i \in \alpha$ *and* $d_i = \lambda$ *otherwise.*

Proof. The equivalence of (i) and (ii) follows from Lemma 3.4 and the fact that λ is a finite eigenvalue of the matrix pencil $C_1 - \lambda C_2$ if and only if λ is an eigenvalue of $C_1 C_2^{-1}$. For the equivalence of (ii) and (iii) when $\lambda \neq 0$, observe that up to a permutation similarity of A,

$$C_1 - \lambda C_2 = \begin{pmatrix} I[\alpha] - \lambda A[\alpha] & -\lambda A[\alpha, \bar{\alpha}] \\ A[\bar{\alpha}, \alpha] & A[\bar{\alpha}] - \lambda I[\bar{\alpha}] \end{pmatrix},$$

where I is the identity matrix in $M_n(\mathbb{C})$. Thus, multiplying the leading $|\alpha|$ rows of $C_1 - \lambda C_2$ by $-\lambda^{-1}$, we obtain that (ii) holds if and only if

$$A - \begin{pmatrix} \lambda^{-1}\, I[\alpha] & 0 \\ 0 & \lambda\, I[\bar{\alpha}] \end{pmatrix}$$

is singular. ∎

It is worth pointing out the parallelism between a PPT viewed as 'partial inversion' and its nonzero eigenvalues being the zeros of $\det(A - D)$ in (iii) of the above theorem. A more precise account of $\det(A - D)$ as a function of λ and of its relation to the spectrum of the PPT is given next. Note that unless $\alpha = \emptyset$, $\det(A - D)$ is <u>not</u> a polynomial in λ.

Proposition 4.2 *Let* $A \in M_n(\mathbb{C})$ *and* $\alpha \subseteq \langle n \rangle$ *such that* $A[\alpha]$ *and* $A[\bar{\alpha}]$ *are invertible. Let* λ *be an indeterminate, and let* $D = \operatorname{diag}(d_1, \ldots, d_n)$ *with* $d_i = \lambda^{-1}$ *if* $i \in \alpha$ *and* $d_i = \lambda$ *otherwise. Then*

$$g(\lambda) = (-1)^{|\bar{\alpha}|} \lambda^{|\alpha|}\, \frac{\det(A - D)}{\det A[\alpha]}$$

is the characteristic polynomial of $\operatorname{ppt}(A, \alpha)$*. Moreover, the coefficients of* $g(\lambda)$ *can be expressed as real linear combinations of the principal minors of* A*.*

Proof. Since $A[\alpha]$ and $A[\bar{\alpha}]$ are invertible, we respectively have that $B = \operatorname{ppt}(A, \alpha)$ is well defined and, by Theorem 3.8, nonsingular. It then follows from Theorem 4.1 (iii) that λ is an eigenvalue of B if and only if $\det(A - D) = 0$, where D is as described above. Since D is diagonal, we obtain the following easy to verify determinantal expansion:

$$\det(A - D) = \sum_{\beta \subseteq \langle n \rangle} (-1)^{|\bar{\beta}|} \prod_{i \notin \beta} d_i \det A[\beta]. \tag{4.2}$$

Since $\bar{\beta} = (\bar{\beta} \cap \bar{\alpha}) \cup (\bar{\beta} \cap \alpha)$ and $(\bar{\beta} \cap \bar{\alpha}) \cap (\bar{\beta} \cap \alpha) = \emptyset$, we have

$$\prod_{i \notin \beta} d_i = \lambda^{|\bar{\beta} \cap \bar{\alpha}| - |\bar{\beta} \cap \alpha|}. \tag{4.3}$$

Also notice that $|\alpha| \geq |\bar{\beta} \cap \alpha| \geq |\bar{\beta} \cap \alpha| - |\bar{\beta} \cap \bar{\alpha}|$, i.e.,

$$|\bar{\beta} \cap \bar{\alpha}| - |\bar{\beta} \cap \alpha| \geq -|\alpha|, \tag{4.4}$$

and that

$$|\bar{\beta} \cap \bar{\alpha}| - |\bar{\beta} \cap \alpha| \leq |\bar{\beta} \cap \bar{\alpha}| \leq |\bar{\alpha}|. \tag{4.5}$$

Equalities hold in (4.4) and (4.5) if and only if $\bar{\beta} = \alpha$ and $\beta = \alpha$, respectively. Thus, multiplying the equation in (4.2) by $\lambda^{|\alpha|}$ and using (4.3)-(4.5), we obtain that λ is an eigenvalue of B if and only if λ is a (nonzero) root of the polynomial

$$\lambda^{|\alpha|} \det(A - D) = \sum_{\beta \subseteq \langle n \rangle} (-1)^{|\bar{\beta}|} \lambda^{|\alpha| + |\bar{\beta} \cap \bar{\alpha}| - |\bar{\beta} \cap \alpha|} \det A[\beta]. \tag{4.6}$$

The term of highest degree in (4.6) appears when $\beta = \alpha$ and equals $(-1)^{|\bar{\alpha}|} \lambda^n \det A[\alpha]$. The constant term in (4.6) appears when $\beta = \bar{\alpha}$ and equals $(-1)^{|\alpha|} \det A[\bar{\alpha}]$. Thus, by Theorem 3.8 (i), $g(\lambda)$ in the statement of the theorem is indeed the characteristic polynomial of B and its coefficients are real linear combinations of the principal minors of A as seen by (4.6). ∎

Note that under the assumptions (and as a consequence) of the above proposition, if $A \in M_n(\mathbb{C})$ has real principal minors, then the spectrum of $\mathrm{ppt}\,(A, \alpha)$ is closed under complex conjugation.

Corollary 4.3 *Let $A \in M_n(\mathbb{C})$ and $\alpha \subseteq \langle n \rangle$ so that $A[\alpha]$ is invertible. Then $1 \in \sigma(\mathrm{ppt}\,(A, \alpha))$ (resp., $-1 \in \sigma(\mathrm{ppt}\,(A, \alpha))$) if and only if $1 \in \sigma(A)$ (resp., $-1 \in \sigma(A)$). Also $\mathrm{ppt}\,(A, \alpha)$ is singular if and only if $A[\bar{\alpha}]$ is singular.*

Proof. The results on the ± 1 eigenvalues follow from Proposition 4.2. The singularity condition for $\mathrm{ppt}\,(A, \alpha)$ follows either from Theorem 3.8 (i) or from Theorem 4.1 (ii). ∎

We continue with an application to iterative techniques for solving a linear system $Ax = b$, where $A \in M_n(\mathbb{C})$ is invertible. Such iterative techniques are obtained by expressing the unique solution x as a fixed point of a matrix equation $x = Tx + c$ for an appropriate matrix T. In fact, based on a *splitting* of A into $A = M - N$ and assuming that M is invertible, we take $T = M^{-1}N$ and $c = M^{-1}b$. Then the sequence $\{x_k\}_0^\infty$ generated by $x_k = Tx_{k-1} + c$ for arbitrary x_0 converges to the solution x if and only if $\rho(T) < 1$ (see e.g., Varga [25]). The Jacobi method is obtained when $M = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ and $N = M - A$. In many instances, certain splittings lead to divergent sequences. This may be overcome by considering a PPT $\hat{T}$ of T and an equation $x = \hat{T}x + d$ equivalent to $x = Tx + c$, as suggested by the following result and illustrated by the subsequent example.

Proposition 4.4 *Let $T \in M_n(\mathbb{C})$ and $x, c \in \mathbb{C}^n$. Let $\alpha \subseteq \langle n \rangle$ so that $T[\alpha]$ is invertible. Consider the vector u defined by*

$$u(i) = \begin{cases} c(i) & \text{if } i \in \alpha \\ 0 & \text{otherwise.} \end{cases}$$

Then $x = Tx + c$ if and only if $x = \hat{T}x + d$, where $d = c - (I + \hat{T})u$.

Proof. Let T, $\hat{T}$, x, c, u and d as prescribed. Observe that by Theorem 3.1, $Tx = x - c$ is equivalent to $\hat{T}(x - u) = x - (c - u)$, which in turn is equivalent to $x = \hat{T}x - \hat{T}u + (c - u)$, that is, $x = \hat{T}x + d$. ∎

Example 4.5 Consider the matrix A and the corresponding Jacobi iteration matrix T given by

$$A = \begin{pmatrix} 1 & -3/2 & -1/4 \\ -3/2 & 1 & -5/2 \\ -1/2 & -1/2 & 1 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 0 & 3/2 & 1/4 \\ 3/2 & 0 & 5/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

We find that $\sigma(T) = \{2.1419, -.6419, -1.5\}$. That is, as $\rho(T) > 1$, the Jacobi iteration $x_k = Tx_{k-1} + c$ fails to converge to the solution of a system $Ax = b$. However, if we consider

$$\hat{T} = \text{ppt}\,(T, \{1, 2\}) = \begin{pmatrix} 0 & 2/3 & -5/3 \\ 2/3 & 0 & -1/6 \\ 1/3 & 1/3 & -11/12 \end{pmatrix},$$

then $\sigma(\hat{T}) = \{-1/4, 0, 2/3\}$ and thus $\rho(\hat{T}) = 2/3 < 1$. It follows that the iteration $x_k = \hat{T}x_{k-1} + d$ with d as in Proposition 4.4, converges to the solution of $Ax = b$. In passing we mention that T above satisfies the assumptions of the Stein-Rosenberg theorem in [25] and hence the Gauss-Seidel iteration for A also fails to converge to the solution of the system.

5 Principal pivot transforms of P-matrices

One of the main matrix classes discussed in association with PPTs is the class of *P-matrices*, that is, matrices in $M_n(\mathbb{C})$ all of whose principal minors are positive. Tucker [24] asserts that principal pivot transformations preserve the class of P-matrices. In the case of real P-matrices a simple proof of this assertion can indeed be based on Theorem 3.1 and on the following characterization of real P-matrices: $A \in M_n(\mathbb{R})$ is a P-matrix if and only if for every nonzero $x \in \mathbb{R}^{\mathrm{n}}$, x and Ax have at least one pair of corresponding entries whose product is positive (see Fiedler [10, Theorem 5.22]). Here we present a proof of the assertion in [24] for the general case of complex P-matrices, based on the following well known result.

Lemma 5.1 *Let $A \in M_n(\mathbb{C})$ be a P-matrix. Then A^{-1} is a P-matrix. Moreover, $A/A[\alpha]$ is a P-matrix for all $\alpha \subseteq \langle n \rangle$.*

Proof. Since A is a P-matrix, $A[\alpha]$ is invertible and the block representation of A^{-1} mentioned at the end of section 2 is valid for every $\alpha \subseteq \langle n \rangle$. Therefore every principal submatrix of A^{-1} is of the form $(A/A[\alpha])^{-1}$ for some $\alpha \in \langle n \rangle$ and has determinant $\det A[\alpha]/\det A > 0$. This shows that A^{-1} is a P-matrix and, in turn, that $A/A[\alpha]$ is a P-matrix for every $\alpha \subseteq \langle n \rangle$. ∎

Theorem 5.2 *Let $A \in M_n(\mathbb{C})$ be a P-matrix and $\alpha \subseteq \langle n \rangle$. Then* ppt (A, α) *is a P-matrix.*

Proof. Let A be a P-matrix and consider first the case where α is a singleton; without loss of generality assume that $\alpha = \{1\}$. Let $B = \text{ppt}\,(A, \alpha) = (b_{ij})$. By definition, the principal submatrices of B that do not include entries from the first row of B coincide with the principal submatrices of $A/A[\alpha]$ and thus, by Lemma 5.1, have positive determinants. The principal submatrices of B that include entries from the first row of B are equal to the corresponding principal submatrices of the matrix B' obtained from B using $b_{11} = (A[\alpha])^{-1} > 0$ as the pivot and eliminating the nonzero entries below it. Notice that

$$B' = \begin{pmatrix} 1 & 0 \\ -A[\bar{\alpha}, \alpha] & I \end{pmatrix} \begin{pmatrix} b_{11} & -b_{11}A[\alpha, \bar{\alpha}] \\ A[\bar{\alpha}, \alpha]b_{11} & A/A[\alpha] \end{pmatrix} = \begin{pmatrix} b_{11} & -b_{11}A[\alpha, \bar{\alpha}] \\ 0 & A[\bar{\alpha}] \end{pmatrix}.$$

That is, B' is itself a P-matrix for it is block upper triangular and the diagonal blocks are P-matrices. It follows that all the principal minors of B are positive and thus B is a P-matrix.

Next, consider the case $\alpha = \{i_1, \ldots, i_k\} \subseteq \langle n \rangle$ with $k \geq 1$. By the proof completed so far, the sequence of matrices

$$A_0 = A, \quad A_j = \text{ppt}\,(A_{j-1}, \{i_j\}), \; j = 1, 2, \ldots, k$$

is well defined and comprises P-matrices. Moreover, from the uniqueness of $B = \text{ppt}\,(A, \alpha)$ shown in Theorem 3.1 it follows that $A_k = \text{ppt}\,(A, \alpha) = B$ and thus B is a P-matrix. ∎

The next theorem summarizes the relation between PPTs and P-matrices and follows readily from the previous result.

Theorem 5.3 *Let $A \in M_n(\mathbb{C})$. Then the following are equivalent:*

(i) A is a P-matrix.
(ii) there exists $\alpha \subseteq \langle n \rangle$ such that $\text{ppt}\,(A, \alpha)$ *is a P-matrix.*
(iii) for all $\alpha \subseteq \langle n \rangle$, $\text{ppt}\,(A, \alpha)$ *is a P-matrix.*

Remark 5.4 Theorem 5.3 is stated in similar terms in [13, Theorem 4.4]. However, the proof provided in [13], unless modified, is valid only when A is a real matrix.

The P-matrices include several well studied subclasses of matrices, notably the positive definite matrices, totally positive matrices and the M-matrices. None of these classes is invariant under principal pivot transformations. There is, however, an interesting subclass of the P-matrices that is PPT invariant. This class is introduced in [19] and we describe its basic features next. All inequalities henceforth are entrywise.

First recall that $A \in M_n(\mathbb{R})$ is a *Z-matrix* if its off-diagonal entries are all nonpositive, and is a nonsingular *M-matrix* if, in addition, it is a P-matrix. Equivalently, A is a nonsingular M-matrix if and only if $A = sI - P$, where $P \geq 0$ and $s > \rho(P)$. (See Berman and Plemmons [2] for details on these matrix classes.) The matrix $A \in M_n(\mathbb{R})$ is called a *hidden Z-matrix* in [19] provided that there exist Z-matrices X, Y such that

$$AX = Y \quad \text{and} \quad r^T X + s^T Y > 0$$

for some vectors $r, s \geq 0$. Consider now the class $\mathcal{G}$ consisting of all matrices that are hidden Z-matrices and P-matrices simultaneously. As is shown in [19], $A \in \mathcal{G}$ if and only if

$$A = (s_1 I - P_1)(s_2 I - P_2)^{-1}, \tag{5.1}$$

where s_1, s_2 are scalars, P_1, P_2 are nonnegative matrices, and there exists a vector $u \geq 0$ such that

$$P_1 u < s_1 u \quad \text{and} \quad P_2 u < s_2 u. \tag{5.2}$$

Notice that $\mathcal{G}$ contains the M-matrices (by taking $P_2 = 0, s_2 = 1$) and their inverses (by taking $P_1 = 0, s_1 = 1$). It is also shown in [19] that every PPT (as well as every permutational similarity, Schur complement and principal submatrix) of a matrix in $\mathcal{G}$ is also in $\mathcal{G}$. We include a proof of this result, utilizing the language and the observations herein and in [13].

Theorem 5.5 *Let $A \in M_n(\mathbb{C}) \cap \mathcal{G}$ and $\alpha \subseteq \langle n \rangle$. Then $ppt(A, \alpha) \in \mathcal{G}$.*

Proof. Let A be as described in (5.1) and (5.2) and denote $Y = (s_1 I - P_1)$, $X = (s_2 I - P_2)$ so that $AX = Y$. Let T be the matrix obtained from the identity by setting the diagonal entries indexed by $\alpha \subseteq \langle n \rangle$ equal to 0 and consider the matrices

$$U = TX + (I - T)Y, \quad V = (I - T)X + TY.$$

That is, U and V are obtained from X and Y by exchanging the rows indexed by α. Thus, on letting $B = \text{ppt}\,(A, \alpha)$, we have $B = VU^{-1}$, where U and V are Z-matrices satisfying $Uu > 0$ and $Vu > 0$. It follows that $B \in \mathcal{G}$. ■

A natural question that arises is whether every matrix in $\mathcal{G}$ is the PPT of an M-matrix. (Some related questions are included in the concluding section.) The answer is in the negative and is provided by the following counterexample. Let

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Then A is an H-matrix with positive diagonal entries and thus as is shown in [19], it belongs to $\mathcal{G}$. It can be easily checked, however, that no PPT of A is a Z-matrix.

We continue with a few words on some other matrix classes that are preserved by principal pivot transformations. One such class comprises the *semipositive* matrices, that is, all $A \in M_n(\mathbb{R})$ such that $Ax > 0$ for some $x > 0$. Clearly, by Theorem 3.1, a PPT of a semipositive matrix is semipositive.

Next, recall the S-orthogonal matrices mentioned in the introduction. The matrix $Q \in M_n(\mathbb{R})$ is called S-orthogonal if there exists a signature matrix $S \in M_n(\mathbb{R})$ (i.e., a diagonal matrix S whose diagonal entries are ± 1) such that $Q^T S Q = S$. That is, the columns of Q are orthonormal with respect to the indefinite scalar product $\langle Sx, y \rangle$ defined by S. When $S = I$, then an S-orthogonal matrix is simply an orthogonal matrix. In [21] it is formally shown that S-orthogonal matrices can be constructed for any prescribed signature matrix S as follows. Suppose that $S = \text{diag}(s_1, \ldots, s_n)$ and that $s_i = 1$ for all $i \in \alpha \subseteq \langle n \rangle$ and $s_i = -1$ for all $i \in \bar{\alpha}$. Let $R \in M_n(\mathbb{R})$ be an orthogonal matrix such that $R[\alpha]$ is invertible. Then $Q = \text{ppt}\,(R, \bar{\alpha}]$ exists and is S-orthogonal. It follows that PPTs preserve this generalized orthogonality property.

It is interesting to note that the mapping via PPTs of orthogonal matrices $R \in M_n(\mathbb{R})$ to S-orthogonal matrices discussed above can be implemented by a transformation on the space of n-by-$2n$ matrices of the form

$$\begin{pmatrix} R & I \end{pmatrix} \longrightarrow \begin{pmatrix} (R_{11})^{-1} & 0 \\ R_{21}(R_{11})^{-1} & -I \end{pmatrix} \begin{pmatrix} R & I \end{pmatrix} \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 0 & -I \\ 0 & 0 & I & 0 \\ 0 & -I & 0 & 0 \end{pmatrix} = \begin{pmatrix} I & \text{ppt}\,(R, \langle k \rangle) \end{pmatrix},$$

where we have taken $\alpha = \langle k \rangle$ and partitioned R based on its k-by-k leading principal minor R_{11} for illustrative purposes. A similar transformation is used in the solution of Riccati equations in systems theory that maps symplectic matrices (thought of as embedded in the symplectic pencils) to Hamiltonian matrices (see Laub [14] and references therein).

Last, $A \in M_n(\mathbb{C})$ is an *EP matrix* provided that it commutes with its Moore-Penrose inverse, or equivalently if $R(A) = R(A^*)$ (see [3]). As is shown in [16], under certain assumptions a PPT (as well as a Schur complement) of an EP matrix is an EP matrix.

6 Conclusions and some questions

Most of the work involving PPTs so far focuses on their basic domain/range exchange property and the fact that they preserve P-matrices. There also seems to be very limited cross-discipline awareness of the uses of PPTs. With this work we hope to raise the level of this awareness and the interest in PPTs. As with Schur complements, a lot of fundamental questions can be asked, regarding e.g., rank, inertia, and possible generalizations of PPTs. We conclude with a few questions of personal interest.

It has been shown in Coxson [6] that the important problem of testing for P-matrices is co-NP-complete. In view of Theorem 5.3, we are led to ask: *Is there a computationally advantageous utilization of PPTs to check whether a given matrix is a P-matrix or not?* Some progress in this direction, based on the proof of Theorem 5.2, is reported in Tsatsomeros and Li [22].

As observed in Example 4.5, PPTs in certain instances map the eigenvalues to desired regions, e.g., the open unit disk. *When and how can we choose α so that the eigenvalues of $ppt(A, \alpha)$ lie in a given region of the complex plane?*

The class $\mathcal{G}$ of matrices that are hidden Z and P contains both the M-matrices and their inverses. Thus it is natural to ask whether PPTs can play a role in the inverse M-matrix problem (see e.g., Johnson [12]). To be more specific, it is clear from the results herein and properties of M-matrices that a nonnegative matrix $A \in M_n(\mathbb{R})$ is the inverse of an M-matrix if and only if it can be transformed into a Z-matrix via consecutive PPTs corresponding to some (and thus every) partition of $\langle n \rangle$. That is, M and inverse M-matrices can be thought as 'connected' through paths of PPTs within the class $\mathcal{G}$. *Can certain PPTs of A (other than the inverse itself) be used to characterize it as an inverse M-matrix? Is there a simple identification process of matrices in $\mathcal{G}$ that are inverse M-matrices? When is a nonnegative matrix in $\mathcal{G}$ an inverse M-matrix?*

Acknowledgment

The author thanks anonymous referees for bringing the sweep operator to his attention and for other useful remarks.

References

[1] V. Belovitch. *Classical Network Theory*. Holden Day, San Francisco, 1968.

[2] A. Berman and R. J. Plemmons. *Nonnegative Matrices in Mathematical Sciences*. SIAM, Philadelphia, 1994.

[3] S. L. Campbell and D. Meyer. *Generalized Inverses of Linear Transformations.* Dover, New York, 1991.

[4] R. W. Cottle. Manifestations of the Schur complement. *Linear Algebra and Its Applications*, 8:189-211, 1974.

[5] R. W. Cottle and G. B. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and Its Applications*, 1:103-125, 1968.

[6] G. E. Coxson. The P-matrix problem is co-NP-complete. *Mathematical Programming*, 64:173-178, 1994.

[7] R. J. Duffin and D. Hazony and N. Morrison. Network synthesis through hybrid matrices. *SIAM Journal on Applied Mathematics*, 14:390-413, 1966.

[8] M. A. Efroymson. Multiple regression analysis. In A. Ralston and H. S. Wilf, editors, *Mathematical Methods for Digital Computers, Vol. I*, pages 191-203. John Wiley & Sons, New York, 1960.

[9] L. Elsner and T. Szulc. Block P-Matrices. *Linear and Multilinear Algebra*, 44:1-12, 1998.

[10] M. Fiedler. *Special matrices and their applications in numerical mathematics.* Martinus Nijhoff, Dordrecht, 1986.

[11] J. H. Goodnight. A tutorial on the SWEEP operator. *The American Statistician*, 33(3):149-158, 1979.

[12] C. R. Johnson. Inverse M-matrices. *Linear Algebra and Its Applications*, 47:195-216, 1982.

[13] C. R. Johnson and M. Tsatsomeros. Convex Sets of Nonsingular and P-Matrices. *Linear and Multilinear Algebra*, 38(3): 233-239, 1995.

[14] A. J. Laub. Numerical linear algebra aspects of control design computations. *IEEE Transactions on Automatic Control*, 30(2):97-108, 1985.

[15] P. M. Meehan, A. P. Dempster, and E. N. Brown. A belief function approach to likelihood updating in a Gaussian linear model. *Bayesian Statistics*, 5:685-691, 1996.

[16] A. Meenakshi. Principal Pivot Transforms of an EP Matrix. *C.R. Math. Rep. Acad. Sci. Canada*, 8(2):121-126, 1986.

[17] P. N. Neytchev. SWEEP operator for least-squares subject to linear constraints. *Computational Statistics & Data Analysis*, 20:599-609, 1995.

[18] D. Ouellette. Schur complements and statistics. *Linear Algebra and Its Applications*, 36:187-295, 1981.

[19] Jong-Shi Pang. Hidden Z-Matrices with Positive Principal Minors. *Linear Algebra and Its Applications*, 23:201-215, 1979.

[20] G. A. F. Seber. *Linear Regression Analysis.* John Wiley & Sons, New York, 1977.

[21] M. Stewart and G. W. Stewart. On Hyperbolic Triangularization: Stability and Pivoting. *SIAM Journal on Matrix Analysis and Applications*, 19(4):847-860, 1998.

[22] M. Tsatsomeros and L. Li. A recursive test for P-matrices. *BIT*, to appear.

[23] A. W. Tucker. A combinatorial equivalence of matrices. In *Combinatorial Analysis*, Bellman and Hall, editors, pages 129-140. American Mathematical Society, Providence, 1960.

[24] A. W. Tucker. Principal pivotal transforms of square matrices. *SIAM Review*, 5:305, 1963.

[25] R. S. Varga. *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[26] R. Wolfinger, R. Tobias, and J. Sall. Computing Gaussian likelihoods and their derivatives for general linear mixed models. *SIAM Journal on Scientific Computing*, 15:1294-1310, 1994.