Rook polynomials to and from permanents

Gi-Sang Cheon\textsuperscript{a,1}, Suk-Geun Hwang\textsuperscript{b,*,2}, Seok-Zun Song\textsuperscript{c,1}

\textsuperscript{a}Department of Mathematics, Daejin University, Pocheon 487-711, South Korea
\textsuperscript{b}Department of Mathematics Education, Kyungpook University, Taegu 702-701, South Korea
\textsuperscript{c}Department of Mathematics, Cheju University, Cheju 690-756, South Korea

Received 30 August 2000; received in revised form 23 June 2001; accepted 22 February 2002

Abstract

In this paper, we find an expression of the rook vector of a matrix \( A \) (not necessarily square) in terms of permanents of some matrices associated with \( A \), and obtain some simple exact formulas for the permanents of all \( n \times n \) Toeplitz band matrices of zeros and ones whose bands are of width not less than \( n - 1 \).

\( \odot \) 2002 Elsevier Science B.V. All rights reserved.

Keywords: Permanent; Rook polynomial; Toeplitz matrix

1. Introduction

For an \( m \times n \) matrix \( A = [a_{ij}] \) with \( m \leq n \), the \textit{permanent} of \( A \), per \( A \), is defined by

\[
\text{per} \ A = \sum_{\sigma} a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{m\sigma(m)},
\]

where the summation runs over all one to one functions of \( \{1,2,\ldots,m\} \) into \( \{1,2,\ldots,n\} \). For \( \alpha \subset \{1,2,\ldots,m\} \) and \( \beta \subset \{1,2,\ldots,n\} \), let \( A[\alpha|\beta] \) denote the submatrix of \( A \) whose rows and columns are indexed by \( \alpha \) and \( \beta \), respectively. It is clear that

\[
\text{per} \ A = \sum_{\beta} \text{per} \ A[1,2,\ldots,m|\beta],
\]

\textsuperscript{*} Corresponding author.
\textit{E-mail address:} sghwang@bh.kyungpook.ac.kr (S.-G. Hwang).
\textsuperscript{1} Supported by Korea Research Foundation Grant(KRF-2000-015-DP0001).
\textsuperscript{2} Supported by Com\textsuperscript{2}MaC-KOSEF.
where the summation runs over all \( m \)-subsets \( \beta \) of \( \{1,2,\ldots,n\} \). For \( k = 1,2,\ldots,m \), let

\[
\sigma_k(A) = \sum_{x,\beta} \text{per} A[x|\beta],
\]

where the summation runs over all \( k \)-subsets \( x \) of \( \{1,2,\ldots,m\} \) and all \( k \)-subsets \( \beta \) of \( \{1,2,\ldots,n\} \). We adopt the natural convention that \( \sigma_0(A) = 1 \) and \( \sigma_k(A) = 0 \) for \( k > m \). Note that \( \sigma_m(A) = \text{per} A \).

Suppose that the \( m \times n \) matrix \( A \) is a \((0,1)\)-matrix. \( A \) may be identified with an \( m \times n \) board in which the square in row \( i \) and column \( j \) has been removed for every pair \( (i,j) \) with \( a_{ij} = 0 \). Then \( \sigma_k(A) \) equals the number of ways to place \( k \) identical rooks on this board so that no rook can attack another, i.e., no two rooks lie on a same row or column. In this sense, \( \sigma_k(A) \) is called the \( k \)th rook number of \( A \) and \( r_k(x) = \sigma_0(A) + \sigma_1(A)x + \cdots + \sigma_m(A)x^m \) is called the rook polynomial of \( A \). The vector \( (\sigma_0(A),\sigma_1(A),\ldots,\sigma_m(A))^T \) is called the rook vector of \( A \).

The evaluation of the permanent of a matrix is known to be a very hard problem. In fact, in 1979, Valiant [4,5] proved that determining the permanent of a \((0,1)\)-matrix is a \#P-complete problem. So it is worthwhile to investigate efficient methods of permanent evaluation for various classes of matrices.

In this paper, we find some relationship between the permanent of a matrix and the rook vectors of some related matrices and make use of it to evaluate permanents of certain types of matrices including a large class of \((0,1)\) Toeplitz matrices.

In the sequel, let \( J_{m,n} \) denote the \( m \times n \) matrix of 1’s which may also be denoted simply by \( J \) in case that the size is clearly seen within the context. Let \( J_n \) denote the matrix \( J_{n,n} \).

### 2. Rook polynomials to permanents

We first discuss how the permanent of a matrix is related to the rook vectors of certain submatrices of it.

For a finite set \( S \) of positive integers and for each \( r = 0,1,\ldots,|S| \), where \( |S| \) stands for the number of elements of \( S \), let \( Q_r(S) \) denote the set of all \( r \)-subsets of \( S \). For \( x \in Q_r(S) \), we denote by \( \overline{x} \) the complement of \( x \) relative to \( S \).

For an \( n \times n \) nonnegative matrix \( A \), we shall call a \( p \times (n + 1 - p) \) zero submatrix of \( A \) a killing zero block (KZB) of \( A \). It is well known as a Frobenius–König theorem that the permanent of a nonnegative square matrix is zero if and only if it has a KZB.

We believe that the following formula has been noted and applied before.

**Theorem 1.** Let \( C \) be a nonnegative square matrix of order \( m + n \) partitioned as

\[
C = \begin{bmatrix} A & X \\ Y & B \end{bmatrix},
\]

where \( A,B \) are nonvacuous square matrices of order \( m,n \), respectively, \( m \leq n \). Then

\[
\text{per} \ C = \sum_{r=0}^m \sum \left( \text{per} A[x|\beta] \text{per} Y[\gamma|\delta] \text{per} A[\overline{x}|\overline{\beta}] \text{per} B[\overline{\gamma}|\overline{\delta}] \right),
\]

(1)
where the second summation runs over all \( z, \delta \in Q_r(1, \ldots, m) \) and all \( \beta, \gamma \in Q_r(m + 1, \ldots, m + n) \).

**Proof.** By Frobenius–König theorem, it follows that a nonzero term in the expansion of \( \text{per} \ C \) uses exactly \( r \) entries of \( X \) if and only if it uses exactly \( r \) entries of \( Y \), for each \( r = 0, 1, \ldots, m \). Suppose that a nonzero term in \( \text{per} \ C \) uses exactly \( r \) entries of \( X \) and exactly \( r \) entries of \( Y \). Among the remaining \( m + n - 2r \) entries in the term, \( m - r \) are those of \( A \) and \( n - r \) are those of \( B \). Letting \( r \) vary from 0 to \( m \), we get (1). \( \Box \)

For square matrices \( A \) and \( B \) of orders \( m \) and \( n \), respectively, let \( A \# B \) denote the \((m + n) \times (m + n)\) matrix defined by

\[
A \# B = \begin{bmatrix} A & J \\ J & B \end{bmatrix}.
\]

For a real number \( x \) and a nonnegative integer \( k \), let \( [x]_k \) denote the number defined by

\[
[x]_k = \begin{cases} 
1 & \text{if } k = 0, \\
x(x - 1) \cdots (x - k + 1) & \text{if } k \geq 1.
\end{cases}
\]

For a positive integer \( n \), let \( D_n = \text{diag}(0!, 1!, \ldots, n!) \).

**Corollary.** Let \( A, B \) be square matrices of order \( m \) and \( n \), respectively, with \( m \leq n \). Then

\[
\text{per}(A \# B) = (\sigma_m(A), \ldots, \sigma_0(A), 0, \ldots, 0) D_n^2(\sigma_n(B), \ldots, \sigma_0(B))^T.
\]

**Proof.** In Theorem 1, letting \( X = J_{m,n}, Y = J_{n,m} \), we have \( \text{per} X[z|\beta] = \text{per} Y[\gamma|\delta] = r! \) for all \( z, \delta \in Q_r(1, \ldots, m) \), all \( \beta, \gamma \in Q_r(m + 1, \ldots, m + n), (r = 0, 1, \ldots, m) \). Thus by Theorem 1,

\[
\text{per}(A \# B) = \sum_{r=0}^{m} r!^2 \sum_{z,\beta,\gamma,\delta} \text{per} A[\bar{z}|\bar{\delta}] \text{per} B[\bar{\gamma}|\bar{\beta}] \\
= \sum_{r=0}^{m} r!^2 \sigma_m-r(A)\sigma_{n-r}(B) \\
= (\sigma_m(A), \ldots, \sigma_0(A), 0, \ldots, 0) \begin{bmatrix} 0!^2 & 0 & \cdots & 0 \\
0 & 1!^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & n!^2 \end{bmatrix} \begin{bmatrix} \sigma_n(B) \\
\sigma_{n-1}(B) \\
\vdots \\
\sigma_0(B) \end{bmatrix}.
\]

\( \Box \)
Our Theorem 1 and its Corollary can often be useful tools for evaluation of permanents of some classes of matrices. For instance, if

\[
A = \begin{bmatrix} I_m & J \\ J & I_n \end{bmatrix}, \quad B = \begin{bmatrix} J_m & J \\ J & I_n \end{bmatrix},
\]

where we assume \( m \leq n \), then

\[
\text{per } A = \sum_{k=0}^{m} k!^2 \sigma_{m-k}(I_m) \sigma_{n-k}(I_n)
\]

\[
= \sum_{k=0}^{m} k!^2 \binom{m}{m-k} \binom{n}{n-k}
\]

\[
= \sum_{k=0}^{m} [m]_k [n]_k,
\]

\[
\text{per } B = \sum_{k=0}^{m} k!^2 \sigma_{m-k}(J_m) \sigma_{n-k}(I_n)
\]

\[
= \sum_{k=0}^{m} k!^2 \binom{m}{m-k} (m-k)! \binom{n}{n-k}
\]

\[
= \sum_{k=0}^{m} [m]_k [m]_{m-k} [n]_k.
\]

Let \( S(n,k) \) denote the Stirling number of the second kind and let \( L_n = [a_{ij}] \) be the strict lower triangular matrix of order \( n \) defined by

\[
a_{ij} = \begin{cases} 
1 & \text{if } i > j, \\
0 & \text{otherwise.}
\end{cases}
\]  

It is well known that the rook polynomial of \( L_n \) is

\[
\sum_{k=0}^{n} S(n,n-k)x^k.
\]

For positive integers \( m, n \) with \( m \leq n \), let \( T_{m,n} \) denote the \((0,1)\) Toeplitz matrix of order \( m + n \) defined by

\[
T_{m,n} = \begin{bmatrix} J & L_m \\ L_n^T & J \end{bmatrix}.
\]
We have, by Corollary to Theorem 1,

\[
\begin{bmatrix}
S(m,0) \\
S(m,1) \\
\vdots \\
S(m,m)
\end{bmatrix}^T
\begin{bmatrix}
0!^2 & 0 & \cdots & 0 & 0 \\
0 & 11^2 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & (n-1)!^2 & 0 \\
\vdots & \vdots & \ddots & 0 & n!^2
\end{bmatrix}
\begin{bmatrix}
S(n,0) \\
S(n,1) \\
\vdots \\
S(n,n)
\end{bmatrix}
\]

\[= \sum_{k=0}^{m} k!^2 S(m,k)S(n,k).\]

3. Rook polynomials from permanents

In this section, we discuss how to find the rook polynomial of a matrix in terms of the permanents of some other related matrices. In this case, it has to be that the permanents of other matrices considered should be rather easily evaluated.

For an \(m \times n\) nonnegative matrix \(A\) with \(m \leq n\), let

\[\tilde{A} = \begin{bmatrix} A \\ O \end{bmatrix}\]

be the square matrix of order \(n\) obtained from \(A\) by putting \(n-m\) zero rows at the bottom. Then \(\sigma_k(\tilde{A}) = \sigma_k(A)\) for all \(k = 0, 1, \ldots\). Let \(Y_0(A) = A\) and

\[Y_k(A) = \begin{bmatrix} J & \tilde{A} \\ J_k & J \end{bmatrix}, \quad y_k(A) = \text{per} Y_k(A) \quad (k = 1, 2, \ldots).\]

Let

\[y_A = \begin{bmatrix} y_0(A) \\ y_1(A) \\ \vdots \\ y_n(A) \end{bmatrix}, \quad r_A = \begin{bmatrix} \sigma_n(A) \\ \vdots \\ \sigma_1(A) \\ \sigma_0(A) \end{bmatrix}\]
and let $\triangle_n$ denote the $(n+1)$ square Pascal matrix

$$
\triangle_n = \begin{bmatrix}
(0) & 0 & 0 & \ldots & 0 \\
(0) & (1) & 0 & \ldots & 0 \\
(1) & (1) & (2) & \ldots & 0 \\
(2) & (1) & (2) & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(n) & (n) & (n) & \ldots & (n) \\
(0) & (1) & (2) & \ldots & (n)
\end{bmatrix}.
$$

**Theorem 2.** For an $m \times n$ nonnegative matrix $A$ with $m \leq n$, we have

$$
y_A = D_n \triangle_n D_r x_A.
$$

**Proof.** Since $\sigma_k(\tilde{A}) = \sigma_k(A)$ for all $i = 0, 1, \ldots$, we have, by Corollary to Theorem 1, that, for each $k = 0, 1, \ldots$,

$$
y_k(A) = \sum_{j=0}^k k!(\binom{k}{j} j! \sigma_{n-j}(A))
$$

and equality (3) follows. $\square$

In [2], it is shown that for every $m \times n$ matrix $A$,

$$
\sum_{k=0}^m \sigma_{n-k}(A)[x]_k = \text{per}[J_{m,x},A].
$$

We would like to point out that equality (3) can also be derived from (4).
Since

\[
\Delta_n^{-1} = \begin{bmatrix}
0 & 0 & 0 & \ldots & 0 & 0 \\
-1 & 1 & 0 & \ldots & 0 & 0 \\
2 & -2 & 2 & \ldots & 0 & 0 \\
0 & -1 & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(-1)^{n-1} & (-1)^{n-1} & (-1)^{n-1} & \ldots & (n-1) & 0 \\
(-1)^n & (-1)^{n+1} & (-1)^{n+1} & \ldots & (n) & (n) \\
\end{bmatrix},
\]

we have:

**Corollary 1.** For an \( m \times n \) nonnegative matrix \( A \), we have

\[
\sigma_{n-k}(A) = \sum_{j=0}^{k} (-1)^{k+j} \frac{1}{k!j!} \binom{k}{j} y_j(A), \quad (k = 0, 1, \ldots).
\]

For positive integers \( m, n \) with \( m \leq n \), an \( m \times n \) \((0, 1)\)-matrix \( A = [a_{ij}] \) with the property

\[
a_{ij} \geq a_{i,j+1}, \quad (i = 1, \ldots, m, \ j = 1, \ldots, n-1),
\]

\[
a_{ij} \leq a_{i+1,j}, \quad (i = 1, \ldots, m-1, \ j = 1, \ldots, n)
\]

is called a **Ferrers matrix**. An \( m \times n \) Ferrers matrix whose row sum vector equals \((b_1, \ldots, b_m)^T\) is denoted by \( F_n(b_1, \ldots, b_m) \) or simply by \( F(b_1, \ldots, b_m) \) in case that \( b_n = m \).

It is well known that

\[
\text{per } F(b_1, \ldots, b_m) = \prod_{i=1}^{n} (b_i - i + 1).
\]

For \( A = F_n(b_1, \ldots, b_m) \), let \( \tilde{A} \) be the \( n \times n \) matrix defined by

\[
\tilde{A} = \begin{bmatrix} O \\ A \end{bmatrix}.
\]
Then $\tilde{A}$ is also a Ferrers matrix and $\sigma_i(\tilde{A}) = \sigma_i(A)$ for all $i = 0, 1, 2, \ldots$. Thus, in dealing with the rook vector of a Ferrers matrix, we need to deal with only square matrices.

For an $n \times n$ Ferrers matrix $A = F_n(b_1, \ldots, b_n)$, let $f_A(t) = \text{per} A[J_n, t]$, i.e.,

$$f_A(t) = \prod_{i=1}^{n} (t + b_i - i + 1).$$

Then, for each $t = 0, 1, \ldots$, $\text{per} Y_t(A) = f_A(t) t!$ so that

$$y_A = D_n \begin{bmatrix} f_A(0) \\ f_A(1) \\ \vdots \\ f_A(n) \end{bmatrix}.$$

Thus, we have, as another corollary to Theorem 2, the following direct formula for the rook vector of a Ferrers matrix.

**Corollary 2.** For the $n \times n$ Ferrers matrix $A = F_n(b_1, \ldots, b_n)$, we have

$$r_A = D_n^{-1} \triangle^{-1} (f_A(0), \ldots, f_A(n))^T.$$

If $A = L_n = F_n(0, 1, \ldots, n-1)$, then

$$\begin{bmatrix} f_A(0) \\ f_A(1) \\ \vdots \\ f_A(n) \end{bmatrix} = \begin{bmatrix} 0^n \\ 1^n \\ \vdots \\ n^n \end{bmatrix}, \quad r_A = \begin{bmatrix} S(n, 0) \\ S(n, 1) \\ \vdots \\ S(n, n) \end{bmatrix}$$

from which we also have the following matrix equation relating the Stirling numbers of the second kind and the binomial coefficients, which is known in the literature in some other form:

$$\begin{bmatrix} S(0, 0) & 0 & 0 & \cdots \\ S(1, 0) & S(1, 1) & 0 & \cdots \\ S(2, 0) & S(2, 1) & S(2, 2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$
4. Permanents of Toeplitz matrices

An $n \times n$ matrix $A = [a_{ij}]$ is called a Toeplitz matrix if $a_{ij} = a_{i+1,j+1}$ for all $i,j$. The evaluation of the permanent of a Toeplitz matrix has long time been one of the main interests in combinatorial matrix theory.

In 1991, Shevelev [3] obtained linear homogeneous relations with constant coefficients for sequences of permanents and determinants of Toeplitz band matrices with complex elements of an arbitrary fixed width. In 1997, Codenotti et al. [1] obtained an algorithm for the permanents of certain very sparse (0,1) Toeplitz matrices. However, there do not exist any exact formulas for the permanent of general Toeplitz matrices up to present time. In this section, we give some precise formulas for a large class of (0,1) Toeplitz band matrices with bandwidth $\geq n - 1$.

For positive integers $n, p, q$ with $p + q \leq n$, let $T_n(p,q)$ denote the $n \times n$ Toeplitz matrix defined by

$$T_n(p,q) = \begin{bmatrix}
J & J & L_q \\
J & J_r & J \\
L_p^T & J & J
\end{bmatrix},$$

where $r = n - p - q \geq 0$ and the central block $J_r$ may be vacuous in case that $r = 0$. Recall that the matrices $L_p$ and $L_q$ are the strictly lower triangular (0,1)-matrices defined by (2).
Now we are ready to prove the following:

**Theorem 3.** Let \( n, p, q \) be positive integers with \( p \leq q, p + q \leq n \). Then

\[
\text{per } T_n(p, q) = (0^p, 1^p, \ldots, \frac{r}{p}, 0, \ldots, 0)(\Delta_k^{-1})^T \Delta_k^{-1} \begin{bmatrix} (r + 0)^p[r + 0] \vspace{1em} \vdots \vspace{1em} (r + k)^p[r + k] \end{bmatrix},
\]

where \( r = n - p - q \) and \( k = n - p \).

**Proof.** Note that \( k = r + q \). Let \( A = L_p^T \) and

\[
B = \begin{bmatrix} J & L_q \\
J_r & J \end{bmatrix}.
\]

Then \( B = F_k(r, r + 1, \ldots, k - 1, k, \ldots, k) \) and

\[
T_n(p, q) = \begin{bmatrix} J & B \\
A & J \end{bmatrix}.
\]

Let \( u_p = (0^p, 1^p, \ldots, \frac{r}{p})^T \) and \( v = (v_0, v_1, \ldots, v_k)^T \) where \( v_i = (r + i)^p[r + i], \ (i = 0, 1, \ldots, k) \). For \( t = 0, 1, \ldots, k \), we have

\[
f_B(t) = \text{per}[J_{k, t}, B] = \text{per}[J_{q + r, t}, B] = (t + r)^p[t + r],
\]

so that

\[
r_B = D_k^{-1} \Delta_k^{-1} v
\]

by Corollary 2 to Theorem 2. Note that \( r_A = D_k^{-1} \Delta_k^{-1} u_p \). If we set

\[
\tilde{r}_A = \begin{bmatrix} r_A \\
0_{k - p} \end{bmatrix}, \quad u = \begin{bmatrix} u_p \\
0_{k - p} \end{bmatrix}
\]

where \( 0_{k - p} \) denotes the \((k - p)\)-vector of zeros, then \( \tilde{r}_A = D_k^{-1} \Delta_k^{-1} u_p \). Since \( T_n(p, q) \) can be transformed into \( A \# B \), we have by Corollary to Theorem 1, that

\[
\text{per } T_n(p, q) = \tilde{r}_A^T D_k^2 r_B = u^T (\Delta_k^{-1})^T D_k^2 D_k^{-1} \Delta_k^{-1} v = u^T (\Delta_k^{-1})^T \Delta_k^{-1} v
\]

and the proof is complete. \( \square \)

With some more restrictions on the parameters \( n, p, q \) the expression of \( \text{per } T_n(p, q) \) gets more elegance. Here are some examples.
Case (i): \( p + q = n \). In this case \( r = 0 \) and

\[
\text{per } T_n(p, q) = \begin{bmatrix} u_p \\ 0_{n-p} \end{bmatrix}^T (\triangle_{n-p}^{-1})^T \triangle_{n-p}^{-1} u_{n-p}.
\]

Case (ii): \( k = 2n \) and \( T_n(p, q) \) is symmetric. In this case we have \( p = q \) and

\[
T_{2k}(p, q) = \begin{bmatrix} J & J & L_p \\ J & T_{2s} & J \\ L^T_p & J & J \end{bmatrix},
\]

where \( s = k - p \). Let

\[
v_{k,p} = \begin{bmatrix} (s + 0)p[s + 0] \\ (s + 1)p[s + 1] \\ \vdots \\ (s + k)p[s + k] \end{bmatrix} = \begin{bmatrix} (k - p + 0)p[k - p + 0] \\ (k - p + 1)p[k - p + 1] \\ \vdots \\ (k - p + k)p[k - p + k] \end{bmatrix}.
\]

Then

\[
\text{per } T_{2k}(p, p) = v^T_{k,p} (\triangle_k^{-1})^T \triangle_k^{-1} v_{k,p} = \| \triangle_k^{-1} v_{k,p} \|_2^2,
\]

where \( \| \cdot \| \) stands for the Euclidean norm.

Moreover if, in addition, \( s = 0 \), i.e., if \( k = p \), then

\[
T_{2k}(k, k) = \begin{bmatrix} J & L_k \\ L_k^T & J \end{bmatrix}.
\]

Since \( v_{k,k} = u_k \) and since

\[
\triangle_k^{-1} u_k = D_k \begin{bmatrix} S(k, 0) \\ S(k, 1) \\ \vdots \\ S(k, k) \end{bmatrix} = \begin{bmatrix} 0!S(k, 0) \\ 1!S(k, 1) \\ \vdots \\ k!S(k, k) \end{bmatrix},
\]

we have

\[
\text{per } T_{2k}(k, k) = \sum_{i=0}^{k} i!^2 S(k, i)^2.
\]
References