On σ-Triangular Matrices

Han, Yong-hyeon

σ-삼각행열에 관하여

韓　　溶　　鉉

要　　約

본 논문에서는 n차 정방행열들의 집합 M_n위에서의 한 관계(relation)를 정의하고, 그의 성질을 연구하였다. 또한, 위의 관계를 이용하여 σ-삼각행열을 정의하고, 삼각행열에 대하여 성립하는 여러 정리들이 σ-삼각행열에 대해서도 성립함을 보였다.

I. Introduction

In this note we will define and study a relation on matrices, and a matrix of certain type. The latter will be called a σ-triangular matrix. This matrix has many properties that a triangular matrix has.

II. Prerequisite Theory

The set of all $n\times n$ real matrices forms a ring, and it will be denoted by M_n. A matrix in M_n is denoted by boldface uppercase letter-for example,

$$\mathbf{A}=\begin{pmatrix} a^1_1 & a^2_1 & \cdots & a^n_1 \\ a^1_2 & a^2_2 & \cdots & a^n_2 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a^1_n & a^2_n & \cdots & a^n_n \end{pmatrix},\ \text{or compactly}$$

$$\mathbf{A}=\left[a^j_i\right].$$

For a matrix $\mathbf{A}$, $\mathbf{A}_k$ and $\mathbf{A}^l$ always mean the k-th row and the l-th column of $\mathbf{A}$, respectively [6], in other words

$$\mathbf{A}_k=\left[a^1_k,\ a^2_k,\ \ldots\ a^n_k\right]\ \text{and}$$

$$\mathbf{A}^l=\begin{pmatrix} a^l_1 \\ a^l_2 \\ \cdot \\ \cdot \\ a^l_n \end{pmatrix}.$$

The transpose of $\mathbf{A}$ is denoted by $\mathbf{A}'$. In this note, δ^j_i always means the Kronecker delta; $\delta^j_i=1$ if $i=j$, 0 if $i\neq j$. Obviously, $\delta^j_i=\delta^i_j$, and $\mathbf{I}=\left[\delta^j_i\right]$ is the identity matrix in M_n.

By a permutation on $X=\{1, 2, \ldots n\}$, we mean a bijective function, that is, one-one and onto function on X. It is well-known that the set S_n of all permutations on X forms a group with respect to the composition of functons, which is called the symmetric group on n symbols [5].

By a permutation matrix we mean a matrix which has entries 0 and 1 in such a way that there is one 1 in each row and column. The set P_n of all $n\times n$ permutation matrices forms a group with respect to the matrix multiplication. The following theorem is somewhat obvious, and we will omit the proof of it.

Theorem 1 The function $\varphi: S_n\longrightarrow P_n$ which assigns each σ in S_n to $\varphi(\sigma)=\left[\delta^j_{\sigma^{-1}(i)}\right]$ is a group isomorphism.

We will denote $\varphi(\sigma)$ by $\mathbf{P}_\sigma$. Obviously, $\left[\delta^j_{\sigma^{-1}(i)}\right]=\left[\delta^{\sigma(j)}_i\right]$.

Theorem 2 Let $\mathbf{A}=\left[a^j_i\right]$ be a matrix in M_n and σ be a permutation in S_n. Then

$$\mathbf{P}_\sigma\mathbf{A}=\left[a^j_{\sigma^{-1}(i)}\right]\ \text{and}\ \mathbf{A}\mathbf{P}_\sigma'=\left[a^{\sigma^{-1}(j)}_i\right].$$

proof) $\mathbf{P}_\sigma\mathbf{A}=\left[\delta^j_{\sigma^{-1}(i)}\right]\left[a^j_i\right]$

$$=\left[\sum_{k=1}^{n}\delta^k_{\sigma^{-1}(i)}a^j_k\right]$$

$$=\left[a^j_{\sigma^{-1}(i)}\right].$$

$$\mathbf{A}\mathbf{P}_\sigma'=\left[a^j_i\right]\left[\delta^j_{\sigma^{-1}(i)}\right]'$$

$$= \left[a^{j}_{i}\right] \left[\delta^{i}_{\sigma^{-1}(j)}\right]$$
$$= \left[a^{j}_{i}\right] \left[\delta^{\sigma^{-1}(j)}_{i}\right]$$
$$= \left[\sum_{k=1}^{n} a^{k}_{i} \ \delta^{\sigma^{-1}(j)}_{k}\right]$$
$$= \left[a^{\sigma^{-1}(j)}_{i}\right].$$

Since $(\mathbf{P}_\sigma \mathbf{A})_i = \left[a^{j}_{\sigma^{-1}(i)}\right] = \mathbf{A}_{\sigma^{-1}(i)}$, we have $\mathbf{A}_i = (\mathbf{P}_\sigma \mathbf{A})_{\sigma(i)}$.

This asserts that $\mathbf{P}_\sigma\mathbf{A}$ can be obtained from **A** by permuting the rows of **A** according to σ. Similarly, $\mathbf{A}\mathbf{P}_\sigma'$ is obtained from **A** by permuting the columns of **A** according to σ.

Theorem 3 Let σ be a permutation in S_n, then $\mathbf{P}_\sigma' = \mathbf{P}_{\sigma^{-1}}$.

proof) $\mathbf{P}_\sigma' = \left[\delta^{\sigma(j)}_{i}\right]' = \left[\delta^{\sigma(i)}_{j}\right] = \left[\delta^{j}_{\sigma(i)}\right]$
$= \mathbf{P}_{\sigma^{-1}}$.

Theorem 4 Matrix $\mathbf{P}_\sigma$ is orthogonal, that is, $\mathbf{P}_\sigma' = (\mathbf{P}_\sigma)^{-1}$ [3].

proof) $\mathbf{P}_\sigma'\mathbf{P}_\sigma = \left[\delta^{\sigma(j)}_{i}\right]' \left[\delta^{\sigma(j)}_{i}\right]$
$= \left[\delta^{j}_{\sigma(i)}\right] \left[\delta^{\sigma(j)}_{i}\right]$
$= \left[\delta^{\sigma(j)}_{\sigma(i)}\right]$
$= \left[\delta^{j}_{i}\right]$
$= \mathbf{I}$;

the identity matrix in M_n. Hence $\mathbf{P}_\sigma'$ is the inverse of $\mathbf{P}_\sigma$[3].

We will define a unary operation on M_n, which is called a transposition on M_n.

Definition 1 Let σ be a permutation in M_n. By a transposition t_σ on M_n we mean a function t_σ; $M_n \longrightarrow M_n$ which assigns each **A** in M_n to $t_\sigma(\mathbf{A}) = \mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma'$, or equivalently to $\mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma^{-1}$.

By the associativity of matrix multiplication and Theorem 1, it is noticed that $t_\sigma(\mathbf{A})$ is obtained from **A** by permuting the rows and columns of **A** according to σ successively, or vice versa.

Theorem 5 A transposition t_σ; $M_n \longrightarrow M_n$ is a ring isomorphism (that is, an automorphism) and $t_\sigma^{-1} = t_{\sigma^{-1}}$.

proof) Let **A** and **B** be any matrices in M_n. Then

$t_\sigma(\mathbf{A}+\mathbf{B}) = \mathbf{P}_\sigma(\mathbf{A}+\mathbf{B})\mathbf{P}_\sigma'$
$= \mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma' + \mathbf{P}_\sigma\mathbf{B}\mathbf{P}_\sigma'$
$= t_\sigma(\mathbf{A}) + t_\sigma(\mathbf{B})$,

and $t_\sigma(\mathbf{AB}) = \mathbf{P}_\sigma\mathbf{AB}\mathbf{P}_\sigma^{-1}$
$= \mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma^{-1}\mathbf{P}_\sigma\mathbf{B}\mathbf{P}_\sigma^{-1}$
$= t_\sigma(\mathbf{A})\ t_\sigma(\mathbf{B})$.

Hence, t_σ is a homomorphism. Since

$(t_\sigma^{-1}t_\sigma)(\mathbf{A}) = t_{\sigma^{-1}}(t_\sigma(\mathbf{A}))$
$= \mathbf{P}_\sigma^{-1}\mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma^{-1}(\mathbf{P}_{\sigma^{-1}})^{-1}$
$= \mathbf{P}_\sigma^{-1}\mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma^{-1}\mathbf{P}_\sigma$
$= \mathbf{A}$,

and since

$(t_\sigma t_\sigma^{-1})\ (\mathbf{A}) = \mathbf{A}$,

we have

$t_\sigma^{-1}t_\sigma = t_\sigma t_\sigma^{-1} = id$;

the identity function on M_n.

Therefore, t_σ is bijective and $t_\sigma^{-1} = t_{\sigma^{-1}}$.

We will denote a transposition t_σ and $t_\sigma(\mathbf{A})$ by σ and $\sigma\mathbf{A}$, respectively, since it is always easy to distinguish the transposition σ from the permutation σ.

Since

$$\sigma\mathbf{A} = \left[\delta^{j}_{\sigma^{-1}(i)}\right] \left[a^{j}_{i}\right] \left[\delta^{\sigma^{-1}(j)}_{i}\right]$$
$$= \left[a^{j}_{\sigma^{-1}(i)}\right] \left[\delta^{\sigma^{-1}(j)}_{i}\right]$$
$$= \left[a^{\sigma^{-1}(j)}_{\sigma^{-1}(i)}\right],$$

every diagonal element of **A** is transposed to diagonal entries of $\sigma\mathbf{A}$, and every off-diagonal elements of **A** is transposed to off-diagonal entries of $\sigma\mathbf{A}$. Hence we get

Theorem 6 The product of diagonal elements of **A** is invariant under transposition σ, and the sum of diagonal elements of **A**, that is, tr(**A**) is invariant under σ [3].

proof) Obvious.

Theorem 7 Let σ be a transposition on M_n

and let **A** be a matrix in M_n. Then

a) $\sigma\mathbf{A}'=(\sigma\mathbf{A})'$, and $\sigma(-\mathbf{A})=-\sigma\mathbf{A}$,

b) if **A** is symmetric, then $\sigma\mathbf{A}$ is symmetric,

c) if **A** is skew-symmetric, then $\sigma\mathbf{A}$ is skew-symmetric,

d) if **A** is a diagonal matrix, that is, $\left[a_i^{\;j}\right]=0$ for $i\neq j$, then $\sigma\mathbf{A}$ is a diagonal matrix,

e) the determinant of **A** is invariant under σ.

proof) a) $\sigma\mathbf{A}'=\sigma\left[a_i^{\;j}\right]'=\sigma\left[a_j^{\;i}\right]=\left[a_{\sigma^{-1}(j)}^{\;\sigma^{-1}(i)}\right]$

$=\left[a_{\sigma^{-1}(i)}^{\;\sigma^{-1}(j)}\right]'=(\sigma\mathbf{A})'$

$\sigma(-\mathbf{A})'=\sigma\left[-a_i^{\;j}\right]=\left[-a_{\sigma^{-1}(i)}^{\;\sigma^{-1}(j)}\right]=-\left[a_{\sigma^{-1}(i)}^{\;\sigma^{-1}(j)}\right]$

$=-\sigma\mathbf{A}$.

b) If $\mathbf{A}'=\mathbf{A}$, then

$(\sigma\mathbf{A})'=\sigma\mathbf{A}'=\sigma\mathbf{A}$.

c) If $\mathbf{A}'=-\mathbf{A}$, then

$(\sigma\mathbf{A})'=\sigma\mathbf{A}'=\sigma(-\mathbf{A})=-\sigma\mathbf{A}$.

d) Obvious.

e) Since $\mathbf{P}_\sigma$ and $\mathbf{P}_\sigma'$ are orthogonal,

$|\mathbf{P}_\sigma|=|\mathbf{P}_\sigma'|=1$ [3]. Hence

$|\sigma\mathbf{A}|=|\mathbf{P}_\sigma\mathbf{A}\mathbf{P}_\sigma'|=|\mathbf{P}_\sigma||\mathbf{A}||\mathbf{P}_\sigma'|$

$=|\mathbf{A}|$.

Definition 2 Two matrices **A** and **B** in M_n is said to be transposable if and only if $\mathbf{B}=\sigma\mathbf{A}$ for some σ on M_n.

Immediately we have,

Theorem 8 Transposability is an equivalence relation on M_n.

proof) Obvious.

II. The σ-Triangular Matrices

We will define a matrix which is very similar to triangular matrix. We will define $t_i^{\;j}$ by $t_i^{\;j}=0$ if $i>j$, 1 if $i\leq j$ for $i, j=1, 2, \ldots n$. Then

$$\mathbf{T}=\left[t_i^{\;j}\right]=\begin{pmatrix}1 & 1 & \ldots & 1\\ 0 & 1 & \ldots & 1\\ \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \ldots & 1\end{pmatrix}$$

is an $n\times n$ upper triangular matrix, and every $n\times n$ upper triangular matrix is of the form $\left[a_i^{\;j}\right]=\left[a_i^{\;j}t_i^{\;j}\right]$.

Definition 3 Let σ be a permutation in S_n. A matrix $\mathbf{A}=\left[a_i^{\;j}\right]$ is said to be a σ-triangular matrix if and only if $\left[a_i^{\;j}\right]=\left[a_i^{\;j}\,t_{\sigma(i)}^{\;\sigma(j)}\right]$.

Theorem 9 A matrix **A** in M_n is a σ-triangular matrix if and only if $\sigma\mathbf{A}$ is an upper triangular matrix, or equivalently, $\mathbf{A}=\sigma^{-1}\mathbf{B}$ for an upper triangular matrix **B**.

proof) Obvious.

The following theorems which is true for triangular matrices are also true for σ-triangular matrices. The set of all σ-triangular matrices will be denoted by T_σ,

Theorem 10 If σ is a permutation in S_n, then T_σ forms a subring of M_n.

proof) Let **A** and **B** be any σ-triangular matrices. Then

$\mathbf{A}-\mathbf{B}=\left[a_i^{\;j}\,t_{\sigma(i)}^{\;\sigma(j)}\right]-\left[b_i^{\;j}\,t_{\sigma(i)}^{\;\sigma(j)}\right]$

$=\left[\left(a_i^{\;j}-b_i^{\;j}\right)t_{\sigma(i)}^{\;\sigma(j)}\right]$

is in T_σ.

$\mathbf{AB}=\sigma^{-1}\,\sigma\,\mathbf{AB}$

$=\sigma^{-1}(\sigma\mathbf{A}\sigma\mathbf{B})$.

Since $\sigma\mathbf{A}$ and $\sigma\mathbf{B}$ are upper triangular matrices, their product is also an upper triangular matrix [4]. Therefore by Theorem 9, **AB** is in T_σ.

Hence T_σ forms a subring of M_n [7].

Theorem 11 If **A** is a σ-triangular matrix, then the determinant of **A** is the product of diagonal elements of **A**.

proof) Since **A** is a σ-triangular matrix, $\sigma\mathbf{A}$ is an upper triangular matrix. The determinant of $\sigma\mathbf{A}$ is the product of diagonal elements of $\sigma\mathbf{A}$ [4], or the product of diagonal elements of A (Theorem 6). But by Theorem 7, the determinant of σA is the determinant of A.

Corollary 11 A σ-triangular matrix A is nonsingular if and only if each diagonal element of A is nonzero.

proof) Obvious.

Theorem 12 If a σ-triangular matrix A is nonsingular, then A^{-1} is also a σ-triangular matrix.

proof) By the fact that $\sigma A^{-1}=(\sigma A)^{-1}$ and that the inverse of nonsingular upper triangular matrix is an upper triangular matrix [4], σA^{-1} is an upper triangular matrix. Hence A^{-1} is a σ-triangular matrix.

Theorem 13 The characteristic roots of an $n \times n$ σ-triangular matrix $A=\left[a^{j}_{i}\,t^{\sigma(j)}_{\sigma(i)}\right]$ are equal to $a^1_1, a^2_2, \ldots a^n_n$.

proof) The characteristic equation is $|A-\lambda I|=0$ [3]; but $A-\lambda I$ is a σ-triangular matrix, and hence, by Theorem 11, we get

$|A-\lambda I|=(a^1_1-\lambda)\ (a^2_2-\lambda)\cdots(a^n_n-\lambda)$.

Set this result equal to zero and the conclusion follows.

Theorem 14 If A is a σ-triangular matrix and if $A'A=AA'$, then A is a diagonal matrix.

proof) Since A is a σ-triangular matrix, σA is an upper triangular matrix. But, since $(\sigma A)'\sigma A=\sigma(A'A)=\sigma(AA')=\sigma A(\sigma A)'$, σA is a diagonal matrix([4] Th. 8,6,10). Hence, by Theorem 7, A is a diagonal matrix

Theorem 15 Let A be a σ-triangular matrix. The i-th diagonal element of A^p is $(a^i_i)^p$.

proof) Since $A^p={}^{\sigma^{-1}}(\sigma A)^p$, the i-th diagonal element of A^p is equal to the $\sigma(i)$-th diagonal element of$(\sigma A)^p$. But σA is an upper triangular matrix and hence the $\sigma(i)$-th diagonal element of $(\sigma A)^p=\left(\left[a^{\sigma^{-1}(i)}_{\sigma^{-1}(i)}\right]\right)^p$ is $\left(a^{\sigma^{-1}(\sigma(i))}_{\sigma^{-1}(\sigma(i))}\right)^p$ $=\left(a^i_i\right)^p$ ([4] Th. 8,6,11).

Theorem 16 Let A be a nonsingular σ-triangular matrix and denote A^{-1} by B; then $a^i_i\, b^i_i=1$ for $i=1, 2, \ldots n$.

proof) Since $A^{-1}=B$ we get

$$\sigma A^{-1}=\sigma\ B,$$

and hence

$(\sigma A)^{-1}=\sigma B$ or $\left[a^{\sigma^{-1}(i)}_{\sigma^{-1}(i)}\right]^{-1}=\left[b^{\sigma^{-1}(i)}_{\sigma^{-1}(i)}\right]$.

Since σA and σB are upper triangular matrices, we get

$a^{\sigma^{-1}(i)}_{\sigma^{-1}(i)}\, b^{\sigma^{-1}(i)}_{\sigma^{-1}(i)}=1$ for $i=1, 2, \ldots n$ ([4] Th.8, 6, 12).

Therefore $a^i_i\, b^i_i=1$ for $i=1, 2, \ldots n$.

References

[1] Ayres, Frank, Jr., 1965. Modern Algebra, Schaum's Outline Series, McGraw-Hill, New York,

[2] Burton, David M., 1970. A First Course in Ring and Ideals, Addison-Wesley, Reading, Mass.

[3] Finkbeiner, Daniel T., II, 1960. Introduction to Matrices and Linear Transformations, Freeman, San Francisco.

[4] Graybill, Franklin A., 1969. Introduction to Matrices with Applications in Statistics, Wadsworths, Belmont.

[5] Halmos, Paul R., 1974. Finite-Dimensional Vector Spaces, Springer-Verlag, New York.

[6] Lang, Serge, 1971. Linear Algebra, Addison-Wesley, Reading, Mass.

[7] Miller, Kenneth S., 1957. Elements of Modern Abstract Algebra, Happer & Row, New York.