Pure Mathematical Sciences, Vol. 10, 2021, no. 1, 57 - 66
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/pms.2021.91277

Hypernormal Matrices

Bob Roohparvar

Batteroo, 2075 De La Cruz Suite 205
Santa Clara CA 95050, USA

Farzan Roohparvar

Department of Computer Science
California State University, East Bay
Hayward, CA 94542, USA

Massoud Malek

Department of Mathematics
California State University, East Bay
Hayward, CA 94542, USA

Abstract

An $n \times n$ real matrix A is *hypernormal* if $A\,P\,A^t = A^t P\,A$, for all permutation matrices P. We shall explain how to construct hypernormal matrices.

Keywords: Hypernormal, permutation, doubly stochastic, normal, symmetric, skew-symmetric, mmhermitian,and ske-hermitian matrices. The Euclidean norm of a vector and a matrix. Birkhoff's theorem. Eigenvalues and eigenvectors. The trace of a matrix.

Introduction

The set of all $n \times n$ real matrices is denoted by $M_n(\mathbb{R})$. In all that follows, we assume that the matrix $A \in M_n(\mathbb{R})$. The zero column vector and zero matrix are denoted by θ and Z, respectively. A permutation matrix P is obtained from the identity matrix I, by interchanging some of its rows or columns. We denote the set of all $n \times n$ permutation matrices by $\mathcal{P}_n$. The fact that there are $n!$ permutation matrices in $M_n(\mathbb{R})$, we conclude that $\mathcal{P}_n$ is a subspace of $M_n(\mathbb{R})$ of dimension $n!$ of $M_n(\mathbb{R})$.

A matrix A is said to be normal, if $A A^t = A^t A$. It is also characterized by the fact that both A and A^t share the same eigenvectors, i.e., if $A v = \lambda v$, then $A^t v = \bar{\lambda} v$. The spectral theorem states that a matrix is normal if and only if it is unitarily similar to a diagonal matrix. This implies that a non-symmetric normal matrix has complex eigenvalues.

The matrix A is called *hypernormal*, if $A P A^t = A^t P A$ for all permutation matrices P. Since the identity matrix I_n is a permutation matrix, we conclude that hypernormal matrices are normal. Clearly any symmetric or skew-symmetric matrix is hypernormal.

For any matrix $A \in M_n(\mathbb{R})$ we define the following subspace of $M_n(\mathbb{R})$:

$$H_n(A) = \{X \in M_n(\mathbb{R}) : A X A^t = A^t X A\}.$$

If A is hypernormal, then $\mathcal{P}_n$ is a subspace of $H_n(A)$ and if A is symmetric or skew-symmetric, then $H_n(A) = M_n(\mathbb{R})$.

A complex matrix H is said to be *hermitian* if it is equal to its conjugate transpose, i.e., $H^* = H$. The diagonal entries of hermitian matrices are all real.

The trace of A is the sum of its diagonal entries and is denoted by $tr(A)$. It is well known that the trace of A equals the sum of all its eigenvalues. The sum of all the entries of the matrix $A = (a_{ij})$ is denoted by $\sigma(A)$.

The Euclidean norm of a vector $v = (v_1, v_2, \ldots, v_n)^t$ is $|| v || = \left[\sum_i | v_i |^2 \right]^{1/2}$.

The Euclidean norm $A = (a_{ij})$ is $|| A || = \left[\sum_{i,j} | a_{ij} |^2 \right]^{1/2} = \sqrt{\text{tr}(A A^t)}$.

Let $e_n = (1, 1, \ldots, 1)^t$ and define the matrix $J_n = \frac{1}{n} e_n e_n^t$. The matrix A with non-negative entries is said to be *doubly stochastic*, if $A e_n = A^t e_n = e_n$. The set of all $n \times n$ doubly stochastic matrices is denoted by Ω_n.

In [1] Birkhoff proved that every doubly stochastic matrix is a convex combination of permutation matrices. This means that they are linear combinations of permutation matrices, where the scalar coefficients are non-negative and their sums equal one.

A matrix $A \in M_n(\mathbb{R})$ is said to be a generalized doubly stochastic matrix, if

$$A\,e_n = A^t e_n \quad \text{or} \quad A\,J_n = J_n\,A\,.$$

The set of all $n \times n$ generalized doubly stochastic matrices is denoted by $\widehat{\Omega}_n$. This set can be partitioned into

$$\widehat{\Omega}_n(s) = \{S \in M_n \,:\, S\,J_n = J_n\,S = s\,J_n\}\,.$$

Let $B = (\,b_{i,j}) \in \widehat{\Omega}_n\,(s)$, with $b_0 = min\{\,b_{i,j}\}$. If $b_0 \geq 0$, then the matrix $C = \dfrac{1}{s}\,B$ is doubly stochastic; otherwise, the matrix $B = \dfrac{n\,|\,a_0\,|\,J_n + B}{n\,|\,a_0\,| + s}$ is doubly stochastic.

Hence by Birkhoff's theorem C is a convex combination of permutation matrices. Since J_b is doubly dtochastic, we conclude that B is a linear combination of permutation matrixes. Thus if A is hypernermal, then $\widehat{\Omega}_n \subset H_n(\,A\,)$.

Results

Note that the matrix $S = A - \sigma\,(A\,)\,J_n$ is skew-symmetric if and only if $A + A^t = 2\,\sigma\,(A\,)\,J_n$.

Theorem 1. *If A is a non-symmetric hypernormal matrix, then $S = A - \sigma\,(A\,)\,J_n$ is skew-symmetric.*

Proof. If A is skew-symmetric, then $\sigma\,(A\,) = 0$. Hence $S = A$ is skew-symmetric.

Suppose now that A is not skew-symmetric. Let $\lambda = a + b\,i$ be a complex eigenvalue of A.

Define the hermitian matrix

$$B_\lambda = \frac{1}{b\,i}\,(\lambda\,A^t - \bar{\lambda}\,A) = \frac{(a + b\,i)}{b\,i}\,A^t - \frac{(a - b\,i)}{b\,i}\,A = \frac{a}{b\,i}\,(\,A^t - A\,) + (\,A^t + A\,)$$

with

$$\sigma\,(B_\lambda) = \frac{a}{b\,i}\,\sigma\,(\,A^t - A\,) + \sigma\,(\,A^t + A\,) = 0 + \sigma\,(\,A^t + A\,) = 2\,\sigma\,(A\,)\,.$$

Let $v = (v_1\,, v_2\,, \ldots\,, v_n)^t$ be an eigenvector of A, corresponding to the complex eigenvalue $\lambda = a + b\,i$. From the equality $A^t P\,A = A\,P\,A^t$, we conclude that

$$b\,i\,B_\lambda\,P\,v = (\lambda\,A^t - \bar{\lambda}\,A)\,P\,v = (A^tP)\,(\lambda\,v) - (A\,P)\,(\bar{\lambda}\,v) =$$
$$A^t\,P\,(A\,v) - A\,P\,(A^t v) = (A^tP\,A - A\,P\,A^t)\,v = Z\,v = \theta.$$

This implies that for any permutation matrix P, the vectors $P\,v$ and $w = v - P\,v$ are eigenvectors of the matrix B_λ corresponding to the eigenvalue zero.

For $i \neq j$, define the permutation matrix $P_{i,j}$, obtain from interchanging the i and j rows of the identity matrix I_n.

Clearly all the components of the complex eigenvectors v cannot be the same, thus for some $i \neq j$, the eigenvector $w = v - P_{i,j}\,v$ of B_λ associated to its zero eigenvalue has exactly two nonzero components $w_i = -\,w_j$. For any permutation matrix P, the vector $P\,w$ is also an eigenvector of B_λ associated with its zero eigenvalue. Hence all the columns of the hermitian matrix B_λ are equal. Since the diagonal entries of hermitian matrices are real; it follows that A must be real. Thus $B_\lambda = A^t + A = 2\,\sigma\,(A)\,J_n$ and $a = 0$, which makes $S = A - \sigma\,(A)\,J_n$ skew-symmetric and the complex eigenvalue λ pure imaginary. ■

Next we obtain some equivalent conditions, whenever the matrix $S = A - \sigma\,(A)\,J_n$ is skew-symmetric.

Theorem 2. *Suppose $\sigma\,(A) \neq 0$ but the matrix $S = A - \sigma\,(A)\,J_n$ is skew-symmetric. Then the following conditions are equivalent:*

(i) A *is normal;*

(ii) $A \in \widehat{\Omega}_n$ *;*

(iii) $\widehat{\Omega}_n \subset H_n\,(A)$ *;*

(iv) A *is hypernormal.*

Proof. If A is normal, then since S is skew-symmetric (*i.e.*, $A + A^t = 2\,\sigma\,(A)\,J_n$), we have

$2\,\sigma\,(A)\,A\,J_n = A\,(\,2\,\sigma\,(A)\,J_n) = A\,(\,A{+}A^t) = (\,A{+}A^t)\,A = (\,2\,\sigma\,(A)\,J_n)\,A = 2\,\sigma\,(A)\,J_n\,A\,.$

The fact that $\sigma\,(A) \neq 0$ implies that $A \in \widehat{\Omega}_n$.

If $A \in \widehat{\Omega}_n$, then since both A and $\sigma\,(A)\,J_n$ are members of $\widehat{\Omega}_n$, we conclude that $S \in \widehat{\Omega}_n$. Now, for all $X \in \Omega_n$, we have

$$\begin{aligned} A\,X\,A^t &= (S + \sigma(A)\,J_n)\,X\,(S + \sigma(A)J_n)^t \\ &= -S\,X\,S - \sigma(A)\,J_n X\,S + \sigma(A)\,S\,X\,J_n + \sigma(A)^2 J_n\,X\,J_n \\ &= -S\,X\,S + \sigma(A)^2 J_n\,; \\ A^t X\,A &= (S + \sigma(A)\,J_n)^t X\,(S + \sigma(A)J_n) \\ &= -S\,X\,S - \sigma(A) S\,X\,J_n + \sigma(A)\,J_n\,X\,S + \sigma(A)^2 J_n X\,J_n \\ &= -S\,X\,S + \sigma(A)^2\,J_n\,. \end{aligned}$$

Thus $A\,X\,A^t = A^t X\,A$.

If *(iii)* holds, then *(iv)* follows from the fact that any permutation matrix is doubly stochastic.

If *(iv)* holds, then *(i)* follows from the fact that any hypernormal matrix is normal. ■

Note that, if $A \in \widehat{\Omega}_n$, then $S \in \widehat{\Omega}_n$;

Combining the results of Theorem 1 and Theorem 2, we obtain the following result:

Corollary 1. If A is a non-symmetric hypernormal matrix with $\sigma(A) \neq 0$, then $A \in \widehat{\Omega}_n$.

Corollary 2. A hypernormal matrix A is either symmetric or has at most one nonzero real eigenvalues.

Proof. Suppose A is not symmetric, then by Theorem 1, the matrix $S = A - \sigma(A)J_n$ is skew-symmetric, therefore all its nonzero eigenvalues are pure imaginary. If A is not skew-symmetric, then by Theorem 2, $A \in \widehat{\Omega}_n$. This implies that A and $A - \lambda_0 J_n$ can be simultaneously diagonalized by the same unitary matrix. Therefore any eigenvalue $\lambda_1 \neq \sigma(A)$ of A is also an eigenvalue of the skew-symmetric matrix S with nonzero pure imaginary eigenvalues. ■

Remark 1. Since the rank of a normal matrix is the number of its nonzero eigenvalues, then according to Corollary 2, a hypernormal matrix of even rank is either symmetric or skew-symmetric. Also, according to the previous results, if $A = (a_{ij})$ is a $n \times n$ non-symmetric hypernormal matrix, then for all $k = 1, \cdots, n$, we have $A_{kk} = \frac{1}{n}\sigma(A)$, and tr $(A) = n$.

Corollary 3. If a normal matrix A is hypernormal, then A^2 is symmetric.

Proof. If $A \in H_n$ is not symmetric, then by Theorem 2, $A + A^t = 2\lambda_0 J_n$. Hence

$$A^2 = (2\sigma(A)J_n - A^t)^2 = 4\sigma(A)^2 J_n - 4\sigma(A)^2 J_n + (A^t)^2 = (A^2)^t.$$

Remark 3. Consider the non-symmetric normal matrix $A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and the permutation matrix $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Although A with a unique nonzero real eigenvalue is the direct sum of two hypernormal matrices, and A^2 is symmetric, but $A_{11} = a_{22} \neq a_{33}$, $1 = \mathrm{tr}A \neq \sigma(A) = \frac{1}{3}$, and

$$APA^t = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} = A^tPA.$$

According to Corollary 1, The only hypernormal matrices of order $n \leq 2$ are either symmetric or skew-symmetric. In the next theorem and its corollaries we study the set $H_n(A)$, where A is a non-symmetric and non-skew-symmetric hypernormal matrix of order grater than or equal to three.

Theorem 3. Let A be $n \times n$ hypernormal matrix where $n \geq 3$, $\sigma(A) \neq 0$ and the matrix $S = A - \sigma(A)J_n$ is skew-symmetric. Then

$$H_n(A) = \{X \ : X, X^t \in Z_n(S)\ \}.$$

Proof. For any $X \in M_n(\mathbb{R})$, we have

$$\begin{aligned} AXA^t &= (S + \sigma(A)J_n)(S + \sigma(A)J_n)^t \\ &= -SXS - \sigma(A)J_nXS + \sigma(A)SXJ_n + \sigma(A)^2J_nXJ_n \\ \text{and} \quad A^tXA &= (S + \sigma(A)J_n)^tX(S + \sigma(A)J_n) \\ &= -SXS - \sigma(A)SXJ_n + \sigma(A)J_nXS + \sigma(A)^2J_nXJ_n. \end{aligned}$$

Thus $AXA^t = A^tXA$ if and only if $J_nXS = SXJ_n$.
By Theorem 1 (v), $S \in \widehat{\Omega}_n(0)$, so if $J_nXS = SXJ_n$, then
$J_nXS = J_nJ_nXS = J_nSXJ_n = Z = SXJ_n$. Hence

$$AXA^t = A^tXA, \ \text{ if and only if } XS \text{ and } SX \text{ are both in } \widehat{\Omega}_n(0).$$

Notice that X and X^t are in $Z_n(S)$, if and only if,

$$SXJ_n = Z \text{ and } (J_nXS)^t = -S(X^tJ_n) = Z.$$

This clearly completes the proof. ∎

Corollary 4. Let A be the matrix defined in Theorem 3. Then

$$\text{rank } S = n-1 \text{ if and only if } H_n(A) = \widehat{\Omega}_n.$$

Proof. Since $J_n \in Z_n(S)$, the proof then follows from Theorem 3 and the fact that the rank of a skew-symmetric matrix is the number of its nonzer eigenvalues. ∎

Remark 3. According to this corollary, $H_3(A) = \widehat{\Omega}_3$. Also if n is even, then $\widehat{\Omega}_n$ is strictly contained in $H_n(A)$.

Corollary 5. For any integer $n \geq 3$,

$$\widehat{\Omega}_n = \bigcap_{A \in H_n} H_n(A).$$

Proof. Let $S = P_n - P_n^t$, where $P_n = (p_{ij})$ is the full cycle permutation matrix (i.e., $1 = p_{n1} = p_{i,i+1}$, for $i = 1, n-1$). Then $S \in \Sigma' \cap \widehat{\Omega}_n(0)$ and according to Theorem 1, for any real λ, the matrix $A = S + \lambda J_n$ is hypernormal. Let $X \in M_n(I\!R)$, such that $X, X^t \in Z_n(S)$
(i.e., $SXJ_n = (P - P^t)XJ_n = Z$ and $SX^tJ_n = (P-P^t)X^tJ_n = Z$).
This implies that $X \in \widehat{\Omega}_n$. We conclude the proof by using Theorem 3. ∎

In the remaining part of this paper we use norms to characterize hypernormal matrices.

Lemma 1. Let $A \in M_n(\mathbb{R})$, then A is symmetric (skew-symmetric) if and only if $||A||^2 = \text{tr } A^2$ ($||A||^2 = -\text{tr } A^2$).

Proof. Let $\lambda_k = a_k + ib_k,\ k = 1, \cdots, n$, be the eigenvalues of the matrix A. The Schur triangularization theorem states that there exists a unitary matrix U such that $U^tAU = T$ is upper triangular and has the form

$$T = \begin{pmatrix} \lambda_1 & t_{1,2} & \dots & t_{1,n} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots\dots & & \lambda_n \end{pmatrix}.$$

Note that

$$||A||^2 = ||T||^2 = \sum_{k=1}^{n} |\lambda_k|^2 + \sum_{i,j\ j>i} ||t_{i,j}||^2 = \sum_{k=1}^{n} \left(a_k^2 + b_k^2\right) + \sum_{i,j\ j>i} ||t_{i,j}||^2 .$$

Since all non-real eigenvalues of a real matrix come in conjugate pairs, it follows that

$$\operatorname{tr} A^2 = \sum_{k=1}^{n} \lambda_k^2 = \sum_{k=1}^{n} \left(a_k^2 - b_k^2 + 2a_k b_k i\right) = \sum_{k=1}^{n} \left(a_k^2 - b_k^2\right) \ .$$

Now if $||A||^2 = |\operatorname{tr} A^2|$, then every T_k is zero; this way T becomes a diagonal matrix and

$$\sum_{k=1}^{n} \left(a_k^2 + b_k^2\right) = |\sum_{k=1}^{m} \left(a_k^2 - b_k^2\right)| \ .$$

If $||A||^2 = \operatorname{tr} A^2$ ($||A||^2 = -\operatorname{tr} A^2$), then all the b_k's (a_k's) must be zero; thus A is symmetric (skew-symmetric). The converse is obvious. ∎

Theorem 4. For any positive integer n,

$$H_n = \{ A \in M_n(\mathbb{R}) \ : \ ||A - \sigma(A)J_n||^2 = |\operatorname{tr} A^2 - \sigma(A)^2|\}.$$

Proof. Since $\operatorname{tr}(B + C) = \operatorname{tr} B + \operatorname{tr} C$, $\operatorname{tr}(J_n) = 1$, and

$$\operatorname{tr}(AJ_n) = \operatorname{tr}(J_n A) = \frac{1}{n}\left(\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}\right)\right) = \frac{1}{n}(n\,\sigma(A)) = \sigma(A) \ ,$$

we have

$$\operatorname{tr}(A - \sigma(A)J_n)^2 = \operatorname{tr}(A^2 - \sigma(A)AJ_n - \sigma(A)J_n A + \sigma(A)^2 J_n) = \operatorname{tr} A^2 - \sigma(A)^2.$$

Thus according to Lemma 1,

$$||A - \sigma(A)J_n||^2 = |\operatorname{tr} A^2 - \lambda_0^2| \ \text{ if and only if } \ S = A - \sigma(A)J_n \in \Sigma_n \cup \Sigma'_n.$$

The result then follows from Theorem 2 and the fact that the matrix $S = A - \sigma(A)J_n$ is symmetric if and only if A is symmetric. ∎

Our last result concerns nonskew-symmetric hypernormal matrices.

Corollary 6. Let $A \in M_n(\mathbb{R})$ with $Ae_n = e_n$. Then A is hypernormal if and only if $||A - J_n||^2 = |\operatorname{tr} A^2 - 1|$.

Proof. If A is hypernormal, then A is normal. Hence $Ae_n = A^t e_n = e_n$ i.e., $A \in \widehat{\Omega}_n(1)$. Theorem 4 then implies that $||A - J_n||^2 = |\operatorname{tr} A^2 - 1|$.

Conversely, if $||A - J_n||^2 = |\operatorname{tr} A^2 - 1|$, then by Lemma 1, $A - J_n$ is either symmetric or skew-symmetric; in both cases A becomes normal. Thus $A \in \widehat{\Omega}_n(1)$ and by Theorem 4, A is hypernormal. ∎

We conclude this paper by constructing a hypernormal matrix $A \notin \Sigma_n \cup \Sigma'_n$.

Let $S = \begin{pmatrix} 0 & 4 & -1 & -3 \\ -4 & 0 & 3 & 1 \\ 1 & -3 & 0 & 2 \\ 3 & -1 & -2 & 0 \end{pmatrix} \in \Sigma'_n \cap \widehat{\Omega}_n(0)$. Then the matrix

$$A = S + 4J_4 = \begin{pmatrix} 1 & 5 & 0 & -2 \\ -3 & 1 & 4 & 2 \\ 2 & -2 & 1 & 3 \\ 4 & 0 & -1 & 1 \end{pmatrix}$$

is hypernormal with $\sigma(A) = \operatorname{tr} A = 4$, $a_{kk} = 1 = \frac{1}{4}\sigma(A)$, $k = 1,2,3,4$, $A + A^t = 2\sigma(A)J_4$,

$$\operatorname{tr} A^2 - \sigma(A)^2 = \operatorname{tr} \begin{pmatrix} -22 & 10 & 22 & 6 \\ 10 & -22 & 6 & 22 \\ 22 & 6 & -10 & -2 \\ 6 & 22 & -2 & -10 \end{pmatrix} - 16 = -64 - 16 = -80,$$

and $\sigma(A)\, J_4||^2 = ||S||^2 = 80$. It is not difficult to show that the matrix

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 2 \end{pmatrix} \in H_n(A).$$

According to Remark 2, A is singular. Observe that A^2 is symmetric.
It is important that we choose the skew-symmetric S in $\widehat{\Omega}_n(0)$. Consider for example, the matrix $S = \begin{pmatrix} 0 & -3 & 1 \\ 3 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} \in \Sigma'_n$ which is not a member of $\widehat{\Omega}_n(0)$.

Notice that $A = S + 3J_3 = \begin{pmatrix} 1 & -2 & 2 \\ 4 & 1 & -1 \\ 0 & 3 & 1 \end{pmatrix}$ is not in $\widehat{\Omega}_n$, and therefore is not hypernormal. However, we have $tr\,(A) = \sigma(A) = 3$, $a_{11} = a_{22} = a_{33} = 1 = \frac{1}{3}\sigma(A)$, $A + A^t = 2\sigma(A)J_3$,

$$\operatorname{tr} A^2 - \sigma(A)^2 = \operatorname{tr} \begin{pmatrix} -7 & 2 & 6 \\ 8 & -10 & 6 \\ 12 & 6 & -2 \end{pmatrix} - 9 = -19 - 9 = -28,$$

and $||A - \sigma(A)J_3||^2 = 28$. Observe that A^2 is not symmetric.

References

[1] G. Birkhoff, Tres observaciones sobre el algebra lineal, *Universidad Nacional de Tucuman*, Revista Serie A, **5** (1946), 147-151.

[2] M. Malek, Hypernormal Matrices, *Linear and Multilinear Algebra*, **36** (4) (1994), 255-266. https://doi.org/10.1080/03081089408818299

[3] R. Sinkhorn, Concerning Hypernormal Stochastic Matrices, *Linear Algebra and its Applications*, **141** (1990), 275-281. https://doi.org/10.1016/0024-3795(90)90325-7

Received: October 5, 2021; Published: November 5, 2021