

Chapter 6

Normal Matrices

Normal matrices are matrices that include Hermitian matrices and enjoy several of the same properties as Hermitian matrices. Indeed, while we proved that Hermitian matrices are unitarily diagonalizable, we did not establish any converse. That is, if a matrix is unitarily diagonalizable, then does it have any special property involving for example its spectrum or its adjoint? As we shall see normal matrices are unitarily diagonalizable.

6.1 Introduction to Normal matrices

Definition 6.1.1. A matrix $A \in M_n$ is called *normal* if $A^*A = AA^*$.

Proposition 6.1.1. $A \in M_n$ is normal if and only if every matrix unitarily equivalent to A is normal.

Proof. Suppose A is normal and $B = U^*AU$, where U is unitary. Then $B^*B = U^*A^*AU = U^*AA^*U = U^*AUU^*A^*U = BB^*$. If U^*AU is normal then it is easy to see that $U^*AA^*U = U^*A^*AU$. Multiply this equation on the right by U^* and on the left by U to obtain $AA^* = A^*A$. \square

Examples.

- (1) Unitary matrices are normal ($U^*U = I = UU^*$).
- (2) Hermitian matrices are normal ($AA^* = A^2 = A^*A$).
- (3) If $A^* = -A$, we have $A^*A = AA^* = -A^2$. Hence matrices for which $A^* = -A$, called *skew-Hermitian*, are normal.

Example 6.1.1. Consider the arbitrary matrix $N \in M_2(R)$, written as $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If we suppose that N is normal then

$$\begin{aligned} N^*N &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}^T \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix} \\ NN^* &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} \end{aligned}$$

From this we conclude that $b^2 = c^2$, or $b = \pm c$. Consider the cases in turn.

(i) If $c = b$, then N is Hermitian and thus normal.

(ii) If $c = -b \neq 0$, then $(N^*N)_{12} = ab + cd = b(a - d)$. On the other hand $(NN^*)_{12} = ac + bd = (d - a)b$. For $b(a - d) = (d - a)b$, we must have $a = d$. This gives that real 2×2 normal matrices are either symmetric or have the form

$$N = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

Note this form includes both rotations and skew-symmetric matrices.

Recall the definition of a unitarily diagonalizable matrix: A matrix $A \in M_n$ is called *unitarily diagonalizable* if there is a unitary matrix U for which U^*AU is diagonal. A simple consequence of this is that if $U^*AU = D$ (where $D = \text{diagonal}$ and $U = \text{unitary}$), then

$$AU = UD$$

and hence A has n orthonormal eigenvectors. This is just a part of the *spectral theorem for normal matrices*.

Theorem 6.1.1 (Spectral theorem for normal matrices). *If $A \in M_n$ has eigenvalues $\lambda_1 \dots \lambda_n$, counted according to multiplicity, the following statements are equivalent.*

- (a) A is normal.
- (b) A is unitarily diagonalizable.
- (c) $\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 = \sum_{j=1}^n |\lambda_j|^2$.
- (d) There is an orthonormal set of n eigenvectors of A .

Proof. (a) \Rightarrow (b). If A is normal, then AA^* is Hermitian and therefore unitarily diagonalizable. Thus $U^*A^*AU = D = U^*AA^*U$. Also, $A, A^*, A^*A = AA^*$ form a commuting family. This implies that eigenvectors of A^*A are also eigenvectors of A . Since A^*A has a complete orthonormal set we know that U^*AU is also diagonal. It is easy to see also that (b) \Rightarrow (a). We also note that (a) \Rightarrow (d).

(b) \Rightarrow (c). Suppose $U^*AU = D$. Then $U^*A^*U = D^*$ and $U^*A^*AU = D^*D$. By Corollary 3.5.2 similarly preserves the trace. We know trace of A^*A is $\text{tr}(A^*A) = \sum_{j=1}^n \sum_{k=1}^n a_{jk}^* a_{kj} = \sum_{j=1}^n \sum_{k=1}^n \bar{a}_{kj} a_{kj} = \sum_{j=1}^n \sum_{k=1}^n |a_{kj}|^2$. Since the trace of D^*D is $\sum |\lambda_j|^2$, the result follows.

(c) \Rightarrow (b). We know that A is unitarily equivalent to a upper triangular matrix T . We also know that if $A \sim B$ are unitarily equivalent $\sum_{ij} |a_{ij}|^2 = \sum_{ij} |b_{ij}|^2$. Application of this equality to the upper triangular matrix T yields

$$\sum_{i,j} |a_{ij}|^2 = \sum |\lambda_j|^2 + \sum_{j>i} |t_{ij}|^2 = \sum |\lambda_j|^2.$$

Thus $t_{ij} = 0$ for $j > i$. Thus A is uniformly diagonalizable.

(d) \Rightarrow (b). Trivial. \square

Corollary 6.1.1. *Let $A \in M_n$ and A is normal. If U is unitary and if U^*AU is upper triangular then U^*AU is diagonal.*

Theorem 6.1.2. *Let $N \in M_n(\mathbb{R})$. Then N is normal if and only if there is a real orthogonal matrix $Q \in M_n(\mathbb{R})$ such that*

$$Q^T N Q = \begin{bmatrix} A_1 & & & \\ & A_2 & & \circ \\ & & \ddots & \\ & \circ & & A_n \end{bmatrix} \quad (1)$$

where A_i is 1×1 (real) or A_i is 2×2 (real) of the form

$$A_i = \begin{bmatrix} \alpha_i & \beta_j \\ -\beta_j & \alpha_i \end{bmatrix}.$$

Proof. First of all, any matrix A of the form given by (1) is normal, and therefore so also is any matrix unitarily similar (real orthogonally similar in this case) to it.

To prove the converse we assume that $N \in M_n(\mathbb{R})$ is normal. We know that N is unitarily diagonalizable. That is, there is a unitary matrix U such that $U^*NU = D$, the diagonal matrix of its eigenvalues. Because N is real, all complex eigenvalues occur in complex conjugate pairs. Arrange them as successive diagonal entries in D . If λ is a real eigenvalue, we can assume without loss of generality that the corresponding eigenvector is real. For complex eigenvalues, the corresponding eigenvectors also occur in conjugate pairs. Thus if $\alpha + i\beta$ is an eigenvalue of N with corresponding eigenvector written in real and complex parts $u = u_r + iu_s$. Since N is real we have that $\alpha - i\beta$ is also an eigenvalue of N with corresponding eigenvector $\bar{u} = u_r - iu_s$. By the fact that N is unitarily diagonalizable, these vectors are orthogonal. This means $\langle u_r, u_s \rangle = 0$.

Replace the eigenvectors $u_r \pm iu_s$ by the real and imaginary parts in U . This gives the matrix Q . Now compute $Q^T N Q$. It is easy to see that compute $Nu_r = \alpha u_r - \beta v_s$ and $Nu_s = \alpha u_s + \beta v_r$. When the first of these vectors $(\alpha u_r - \beta v_s)$ is multiplied by Q^T we obtain the vector $[0, \dots, \alpha, -\beta, 0, \dots, 0]^T$. Multiplication by the second gives the vector $[0, \dots, \beta, \alpha, 0, \dots, 0]^T$. In this way the components $\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$ arise. \square

Corollary 6.1.2. (a) $A \in M_n$ is symmetric if and only if (1) holds with all blocks 1×1 (and real).

(b) $AA^T = I$ if and only if (1) has the form

$$\begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_p & & & \\ & & & A_1 & & * \\ * & & & & \ddots & \\ & & & & & A_k \end{bmatrix}$$

where $\lambda_j = \pm 1$ and $A_j = \begin{bmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \theta_j & \cos \theta_j \end{bmatrix}$ $\theta_j \in \mathbb{R}$.

6.2 Exercises

1. If A and B commute and if A is normal, then A^* and B commute.