Lecture 2: Spectral Theorems

This lecture introduces normal matrices. The spectral theorem will inform us that normal matrices are exactly the unitarily diagonalizable matrices. As a consequence, we will deduce the classical spectral theorem for Hermitian matrices. The case of commuting families of matrices will also be studied. All of this corresponds to section 2.5 of the textbook.

1 Normal matrices

Definition 1. A matrix $A \in \mathcal{M}_n$ is called a normal matrix if

$$AA^* = A^*A.$$

Observation: The set of normal matrices includes all the Hermitian matrices ($A^* = A$), the skew-Hermitian matrices ($A^* = -A$), and the unitary matrices ($AA^* = A^*A = I$). It also contains other matrices, e.g. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, but not all matrices, e.g. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Here is an alternate characterization of normal matrices.

Theorem 2. A matrix $A \in \mathcal{M}_n$ is normal iff

$$\|Ax\|_2 = \|A^*x\|_2 \qquad \text{for all } x \in \mathbb{C}^n.$$

Proof. If A is normal, then for any $x \in \mathbb{C}^n$,

$$\|Ax\|_2^2 = \langle Ax, Ax \rangle = \langle x, A^*Ax \rangle = \langle x, AA^*x \rangle = \langle A^*x, A^*x \rangle = \|A^*x\|_2^2.$$

Conversely, suppose that $\|Ax\| = \|A^*x\|$ for all $x \in \mathbb{C}^n$. For any $x, y \in \mathbb{C}^n$ and for $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ chosen so that $\Re(\lambda \langle x, (A^*A - AA^*)y \rangle) = |\langle x, (A^*A - AA^*)y \rangle|$, we expand both sides of $\|A(\lambda x + y)\|_2^2 = \|A^*(\lambda x + y)\|_2^2$ to obtain

$$\|Ax\|_2^2 + \|Ay\|_2^2 + 2\Re(\lambda \langle Ax, Ay \rangle) = \|A^*x\|_2^2 + \|A^*y\|_2^2 + 2\Re(\lambda \langle A^*x, A^*y \rangle).$$

Using the facts that $\|Ax\|_2^2 = \|A^*x\|_2^2$ and $\|Ay\|_2^2 = \|A^*y\|_2^2$, we derive

$$\begin{aligned} 0 &= \Re(\lambda \langle Ax, Ay \rangle - \lambda \langle A^*x, A^*y \rangle) = \Re(\lambda \langle x, A^*Ay \rangle - \lambda \langle x, AA^*y \rangle) = \Re(\lambda \langle x, (A^*A - AA^*)y \rangle) \\ &= |\langle x, (A^*A - AA^*)y \rangle|. \end{aligned}$$

Since this is true for any $x \in \mathbb{C}^n$, we deduce $(A^*A - AA^*)y = 0$, which holds for any $y \in \mathbb{C}^n$, meaning that $A^*A - AA^* = 0$, as desired. $\square$

Before proceeding to the next section, we isolate the following two results.

Lemma 3. Normality is preserved under unitary equivalence.

Proof. Left to the reader. □

Lemma 4. A triangular matrix is normal if and only if it is diagonal.

Proof. It is easy to observe that a diagonal matrix is normal. We now wish to prove that if a triangular matrix $T \in \mathcal{M}_n$ is normal, then it is necessarily diagonal. We proceed by induction on n. For $n = 1$, there is nothing to do. Let us now assume that the result holds up to an integer $n-1$, $n \geq 2$, and let us prove that it also holds for n. Given $T \in \mathcal{M}_n$, we decompose it into blocks and compute the products TT^* and T^*T as follows

$$T = \left[\begin{array}{c|c} t_{1,1} & z^* \\ \hline 0 & \widetilde{T} \end{array}\right], \quad TT^* = \left[\begin{array}{c|c} |t_{1,1}|^2 + \|z\|_2^2 & \text{x} \\ \hline \text{x} & \widetilde{T}\widetilde{T}^* \end{array}\right], \quad T^*T = \left[\begin{array}{c|c} |t_{1,1}|^2 & \text{x} \\ \hline \text{x} & zz^* + \widetilde{T}^*\widetilde{T} \end{array}\right].$$

Since $TT^* = T^*T$, equality in the top-left block implies $z = 0$, and in turn equality in the bottom-right block yields $\widetilde{T}\widetilde{T}^* = \widetilde{T}^*\widetilde{T}$. The matrix $\widetilde{T} \in \mathcal{M}_{n-1}$ is triangular and normal, so it must be diagonal by the induction hypothesis. Taking $z = 0$ into account, we now see that T is itself diagonal. This finishes the proof by induction. □

2 Spectral theorem

The spectral theorem for normal matrices basically states that a matrix A is normal iff it is unitarily diagonalizable — i.e., there exist a unitary matrix U and a diagonal matrix D such that $A = UDU^*$. It is important to remark that the latter is equivalent to saying that there exists an orthonormal basis (the columns of U) of eigenvectors of A (the corresponding eigenvalues being the diagonal elements of D). Additionally, the following result provides an easy-to-check necessary and sufficient condition for normality.

Theorem 5. Given $A \in \mathcal{M}_n$, the following statements are equivalent:

(i) A is normal,

(ii) A is unitarily diagonalizable,

(iii) $\sum_{1 \leq i,j \leq n} |a_{i,j}|^2 = \sum_{1 \leq i \leq n} |\lambda_i|^2$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, counting multiplicities.

Proof. (i) ⇔ (ii). By Schur's theorem, A is unitarily equivalent to a triangular matrix T. Then

$$A \text{ is normal} \underset{\text{Lem.3}}{\iff} T \text{ is normal} \underset{\text{Lem.4}}{\iff} T \text{ diagonal} \iff A \text{ is unitarily diagonalizable.}$$

(ii) ⇒ (iii). Suppose that A is unitarily equivalent to a diagonal matrix D. Note that the diagonal entries of D are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A. Then

$$\sum_{1 \le i,j \le n} |a_{i,j}|^2 = \operatorname{tr}(A^* A) = \operatorname{tr}(D^* D) = \sum_{1 \le i \le n} |\lambda_i|^2.$$

(iii) ⇒ (ii). By Schur's theorem, A is unitarily equivalent to a triangular matrix T. Therefore,

$$\sum_{1 \le i,j \le n} |a_{i,j}|^2 = \operatorname{tr}(A^* A) = \operatorname{tr}(D^* D) = \sum_{1 \le i,j \le n} |t_{i,j}|^2. \tag{1}$$

On the other hand, we have

$$\sum_{1 \le i \le n} |\lambda_i|^2 = \sum_{1 \le i \le n} |t_{i,i}|^2, \tag{2}$$

because the diagonal entries of T are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A. Thus, the equality between (1) and (2) imply that $t_{i,j} = 0$ whenever $i \neq j$, i.e., that T is a diagonal matrix. Hence, A is unitarily diagonalizable. □

As a simple corollary, we obtain the important spectral theorem for Hermitian matrices.

Theorem 6. If a matrix $A \in \mathcal{M}_n$ is Hermitian, then A is unitarily diagonalizable and its eigenvalues are real.

Proof. The first part of the statement holds since Hermitian matrices are normal matrices. For the second part, note that if $A = UDU^*$ for a unitary matrix U and a diagonal matrix D, then $A^* = U\overline{D}U^*$, so if A is Hermitian, then $D = \overline{D}$, i.e., the eigenvalues of A are real. □

3 Commuting families

In this section, we investigate families of matrices $\{A_i, i \in I\} \subseteq \mathcal{M}_n$ such that $A_i A_j = A_j A_i$ for all $i, j \in I$. These are called commuting families. Any family of diagonal matrices is a commuting family, and in fact so is any family of the type SD_iS^{-1} where S is an invertible matrix and the D_i, $i \in I$, are diagonal matrices. The following result is a converse of this statement.

Theorem 7. A commuting family $\mathcal{F} \subseteq \mathcal{M}_n$ of diagonalizable matrices is simultaneously diagonalizable, i.e., there exists an invertible $S \in \mathcal{M}_n$ such that $S^{-1}AS$ is diagonal for all $A \in \mathcal{F}$.

Proof. We proceed by induction on n. For $n = 1$, there is nothing to do. Let us now assume that the result holds up to an integer $n-1$, $n \geq 2$, and let us prove that it also holds for n. Considering a commuting family $\mathcal{F} \subseteq \mathcal{M}_n$ of diagonalizable matrices, we may assume that there is a matrix $M \in \mathcal{F}$ with at least two eigenvalues (otherwise $\mathcal{F}$ contains only multiples of the identity matrix, and the result is clear). Therefore, for some invertible matrix $M \in \mathcal{M}_n$,

$$M' := S^{-1}MS = \left[\begin{array}{c|c|c} \lambda_1 I & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \lambda_k I \end{array}\right] \qquad \text{where } \lambda_1, \ldots, \lambda_k \text{ are all distinct.}$$

For any $A \in \mathcal{F}$, the equality $AM = MA$ gives $A'M' = M'A'$, where $A' := S^{-1}AS$, hence

$$\left[\begin{array}{c|c|c} \lambda_1 A'_{1,1} & \cdots & \lambda_n A'_{1,n} \\ \hline \vdots & \ddots & \vdots \\ \hline \lambda_1 A'_{n,1} & \cdots & \lambda_n A'_{n,n} \end{array}\right] = \left[\begin{array}{c|c|c} \lambda_1 A'_{1,1} & \cdots & \lambda_1 A'_{1,n} \\ \hline \vdots & \ddots & \vdots \\ \hline \lambda_n A'_{n,1} & \cdots & \lambda_n A'_{n,n}. \end{array}\right]$$

By looking at the off-diagonal elements, we conclude that $A'_{i,j} = 0$ whenever $i \neq j$. Therefore, every $A \in \mathcal{F}$ satisfies

$$S^{-1}AS = \left[\begin{array}{c|c|c} A'_{1,1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & A'_{n,n} \end{array}\right].$$

We now observe that each $\mathcal{F}'_i := \{A'_{i,i}, A \in \mathcal{F}\}$ is a commuting family of diagonal matrices with size smaller than n — the commutativity is easy to check, and the diagonalizibility follows from Theorem 1.3.10 in the textbook. By applying the induction hypothesis to each $\mathcal{F}'_i$, we find invertible matrices S_i such that $S_i^{-1}A'_{i,i}S_i =: D_i$ is diagonal for each $i \in [1:k]$. We finally obtain, for every $A \in \mathcal{F}$,

$$\left[\begin{array}{c|c|c} S_1^{-1} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & S_n^{-1} \end{array}\right] S^{-1}AS \left[\begin{array}{c|c|c} S_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & S_n \end{array}\right] = \left[\begin{array}{c|c|c} D_1 & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & D_n \end{array}\right],$$

so that every $A \in \mathcal{F}$ is diagonalizable through a common invertible matrix. This finishes the proof by induction. □

The following theorem is a version of Schur's theorem for commuting matrices.

Theorem 8. A commuting family $\mathcal{F} \subseteq \mathcal{M}_n$ of matrices is simultaneously unitarily triangularizable, i.e., there exists a unitary $U \in \mathcal{M}_n$ such that U^*AU is upper triangular for all $A \in \mathcal{F}$.

Proof. We proceed by induction on n. For $n = 1$, there is nothing to do. Let us now assume that the result holds up to an integer $n-1$, $n \geq 2$, and let us prove that it also holds for n. Given a commuting family $\mathcal{F} \in \mathcal{M}_n$, Lemma 1 from Lecture 0 guarantees that $\mathcal{F}$ possesses a common eigenvector, which can be assumed to be ℓ_2-normalized. Call this vector v_1 and form an orthonormal basis $v = (v_1, v_2, \ldots, v_n)$. Each $A \in \mathcal{F}$ is unitarily equivalent to the matrix of $x \in \mathbb{C}^n \mapsto Ax \in \mathbb{C}^n$ relative to the basis v, i.e..

$$A \underset{\text{unit.}}{\sim} \left[\begin{array}{c|c} a & \mathrm{x} \\ \hline 0 & \widetilde{A} \end{array}\right], \qquad A \in \mathcal{F}.$$

By looking at the product AB and BA for all $A, B \in \mathcal{F}$, we see that $\widetilde{A}\widetilde{B} = \widetilde{B}\widetilde{A}$ for all $A, B \in \mathcal{F}$. Thus, the family $\widetilde{\mathcal{F}} := \{\widetilde{A}, A \in \mathcal{F}\}$ is a commuting family of matrices of size $n-1$. By the induction hypothesis, the family $\widetilde{\mathcal{F}}$ is simultaneously unitarily triangularizable. We can then infer that the family $\mathcal{F}$ is itself simultaneously unitarily triangularizable (see the argument in the proof of Schur's theorem). This finishes the proof by induction. □

Theorem 9. A commuting family $\mathcal{F} \subseteq \mathcal{M}_n$ of normal matrices is simultaneously unitarily diagonalizable, i.e., there exists a unitary $U \in \mathcal{M}_n$ such that U^*AU is diagonal for all $A \in \mathcal{F}$.

Proof. By Theorem 8, there exists a unitary matrix $U \in \mathcal{M}_n$ such that $T_A := U^*AU$ is upper triangular for all $A \in \mathcal{F}$. Then, for each $A \in \mathcal{F}$, T_A is normal (by Lemma 3) and in turn diagonal (by 4). This is the desired result. □

An aside: Fuglede's theorem

Let A and B be two square matrices. Suppose that A and B commute and that A is a normal matrix. Prove that A^* and B commute — this is (a special case of) Fuglede's theorem. Deduce that the product of two commuting normal matrices is also normal.

One needs to prove that $A^*B - BA^* = 0$, knowing that $AB = BA$ and $AA^* = A^*A$. Recall that a square matrix C is zero if and only if $\mathrm{tr}[CC^*] = 0$. Here one has

$$\begin{aligned}\mathrm{tr}[(A^*B - BA^*)(A^*B - BA^*)^*] &= \mathrm{tr}[(A^*B - BA^*)(B^*A - AB^*)] \\ &= \mathrm{tr}[A^*BB^*A] - \mathrm{tr}[A^*BAB^*] - \mathrm{tr}[BA^*B^*A] + \mathrm{tr}[BA^*AB^*].\end{aligned}$$

To conclude, it is enough to remark that

$$\begin{aligned}\mathrm{tr}[A^*BAB^*] &= \mathrm{tr}[A^*ABB^*] = \mathrm{tr}[AA^*BB^*] = \mathrm{tr}[A^*BB^*A], \\ \mathrm{tr}[BA^*AB^*] &= \mathrm{tr}[BAA^*B^*] = \mathrm{tr}[ABA^*B^*] = \mathrm{tr}[BA^*B^*A].\end{aligned}$$

Now assume furthermore that B is normal (i.e. $BB^* = B^*B$). Using what has just been done, it is possible to derive that AB is normal, since

$$(AB)(AB)^* = ABB^*A^* = AB^*BA^* = B^*AA^*B = B^*A^*AB = (AB)^*(AB).$$

4 Exercises

Ex.1: What can be said about the diagonal entries of Hermitian and skew-Hermitian matrices?

Ex.2: Prove that a matrix $A \in \mathcal{M}_n$ is normal iff $\langle Ax, Ay \rangle = \langle A^*x, A^*y \rangle$ for all $x, y \in \mathbb{C}^n$.

Ex.3: Prove that if two matrices $A, B \in \mathcal{M}_n$ commute and have no common eigenvalues, then the difference $A - B$ is invertible.

Ex.4: Prove Lemma 3.

Ex.5: Exercise 8 p. 109.

Ex.6: Prove that the product of two commuting normal matrices is also a normal matrix. Show that the product of two normal matrices can be normal even even if the two matrices do not commute. In general, is it true that the product of two normal matrices (not necessarily commuting) is normal?

Ex.7: Exercise 14 p. 109.

Ex.8: Exercise 20 p. 109.

Ex.9: Exercise 24 p. 110.

Ex.10: What would a spectral theorem for skew-Hermitian matrices look like? Could it be deduced from the spectral theorem for Hermitian matrices?

Ex.11: Generalize Fuglede's theorem by showing that if M and N are two normal matrices such that $MB = BN$ for some matrix B, then $M^*B = BN^*$.