Lecture 0: Review

This opening lecture is devised to refresh your memory of linear algebra. There are some deliberate blanks in the reasoning, try to fill them all. If you still feel that the pointers are too sketchy, please refer to Chapters 0 and 1 of the textbook for more detailed arguments.

1 Range and null space

Denote by $\mathcal{M}_{m\times n} := \mathcal{M}_{m\times n}(\mathbb{C})$ the set of all matrices with m rows and n columns. Denote by $\mathcal{M}_{m\times n}(\mathbb{R})$ the subset of $\mathcal{M}_{m\times n}$ composed of matrices with only real entries. Denote by $\mathcal{M}_n := \mathcal{M}_{n\times n}$ the set of all square matrices of size $n \times n$, and by $\mathcal{M}_n(\mathbb{R})$ the subset of $\mathcal{M}_n$ composed of matrices with only real entries.

For $A \in \mathcal{M}_{m\times n}$, define its range

$$\operatorname{ran} A := \{Ax, x \in \mathbb{C}^m\},$$

and its null space

$$\ker A := \{x \in \mathbb{C}^n : Ax = 0\}.$$

Verify that these are linear subspaces of $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. Define the rank and the nullity of A by

$$\operatorname{rk} A := \dim(\operatorname{ran} A), \qquad \operatorname{nul} A := \dim(\ker A).$$

They are deduced form one another by the rank-nullity theorem (prove it)

$$\operatorname{rk} A + \operatorname{nul} A = n.$$

Recall that $A \in \mathcal{M}_{m\times n}$ is injective (one-to-one, nonsingular) if $\ker A = \{0\}$, and surjective if $\operatorname{ran} A = \mathbb{C}^m$. Note that a *square* matrix A is injective (or surjective) iff it is both injective and surjective, i.e., iff it is bijective. Bijective matrices are also called invertible matrices, because they are characterized by the existence of a unique square matrix B (the inverse of A, denoted A^{-1}) such that $AB = BA = I$.

2 Trace and determinant

The trace and determinants are functions taking *square* matrices and returning scalars. The trace of $A \in \mathcal{M}_n$ is the sum of its diagonal elements, i.e.,

$$\operatorname{tr} A := \sum_{i=1}^{n} a_{i,i} \quad \text{where} \quad A = [a_{i,j}]_{i,j=1}^{n}.$$

Notice that the trace is linear (i.e., $\mathrm{tr}\,(\lambda A + \mu B) = \lambda \mathrm{tr}\,(A) + \mu \mathrm{tr}\,(B)$) and that (prove it)

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA) \qquad \text{whenever } A \in \mathcal{M}_{m\times n} \text{ and } B \in \mathcal{M}_{n\times m}.$$

As for the determinant, it can be defined in several equivalent ways:

1. As a function of the columns of a marix, it is the only function $f : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ that is linear with respect to each columns ($f(\ldots, \lambda x + \mu y, \ldots) = \lambda f(\ldots, x, \ldots) + \mu f(\ldots, y, \ldots)$), alternating ($f(\ldots, x, \ldots, y, \ldots) = -f(\ldots, y, \ldots, x, \ldots)$), and unit-normalized ($f(I) = 1$). Use this to derive the identity

$$\det(AB) = \det(A)\det(B) \quad \text{for all } A, B \in \mathcal{M}_n.$$

2. $\det A = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$,
where S_n is the set of $n!$ permutations of $\{1, \ldots, n\}$ and $\mathrm{sgn}(\sigma) = (-1)^s$, s = number of pairwise interchanges composing σ (hence the computation rules for 2×2 and 3×3 determinants).
Use this to prove that
$$\det A^\top = \det A \qquad \text{for all } A \in \mathcal{M}_n.$$

3. Laplace expansion with respect to a row or a column, e.g. with respect to the ith row

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j},$$

where $A_{i,j}$ is the submatrix of A obtained by deleting the ith row and the jth column. The matrix $B \in \mathcal{M}_n$ with entries $b_{i,j} := (-1)^{i+j} \det A_{i,j}$ is called the comatrix of A — note that $B^\top$ is also called the adjoint of A (*classical* adjoint, not to be confused with *hermitian* adjoint). Use Laplace expansion to prove that $AB^\top = (\det A)I$. Deduce that $A \in \mathcal{M}_n$ is invertible iff $\det A \neq 0$, in which case give an expression for the inverse of A.

3 Eigenvalues and eigenvectors

Given a *square* matrix $A \in \mathcal{M}_n$, if there exist $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n$, $x \neq 0$, such that

$$Ax = \lambda x,$$

then λ is called an eigenvalue of A and x is called an eigenvector corresponding to the eigenvalue λ. The set of all eigenvectors corresponding to an eigenvalue λ is called the eigenspace corresponding to the eigenvalue λ. Verify that an eigenspace is indeed a linear space. Note

that λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$, i.e., iff λ is a zero of the characteristic polynomial of A defined by

$$p_A(x) := \det(A - xI).$$

Observe that p_A is a polynomial of the form

$$p_A(x) = (-1)^n x^n + (-1)^{n-1} \mathrm{tr}\,(A)\, x^{n-1} + \cdots + \det(A).$$

Since this polynomial can also be written in factorized form as $(\lambda_1 - x) \cdots (\lambda_n - x)$, where $\{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues of A (complex and possibly repeated), we have

$$\mathrm{tr}\,(A) = \lambda_1 + \cdots + \lambda_n, \qquad \det(A) = \lambda_1 \cdots \lambda_n.$$

Verify that the existence of n linearly independent eigenvectors $v_1, \ldots, v_n \in \mathbb{C}^n$ corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$ of $A \in \mathcal{M}_n$ (which occurs in particular if A has n distinct eigenvalues) is equivalent to the existence of a invertible matrix $V \in \mathcal{M}_n$ and of a diagonal matrix $D \in \mathcal{M}_n$ such that

$$A = VDV^{-1}.$$

(What are the relations between the $v_i's$, λ_i's and V, D?) In this case, we say that the matrix A is diagonalizable. More generally, two matrices A and B are called equivalent if there exists an invertible matrix V such that $A = VBV^{-1}$. Note that two similar matrices have the same characteristic polynomial, hence the same eigenvalues (counting multiplicities), and in particular the same trace and determinant.

It is useful to know that a commuting family of diagonalizable matrices is simultaneously diagonalizable in the sense that each matrix in the family is similar to a diagonal matrix via one and the same similarity matrix V. This will be proved in Lecture 2. Another proof strategy relies on the following observation.

Lemma 1. If $\{A_i, i \in I\}$ is a commuting family of matrices in $\mathcal{M}_n$, then there exists $x \in \mathbb{C}^n$ which is an eigenvector for every A_i, $i \in I$.

Proof. We proceed by induction on n. For $n = 1$, there is nothing to do. Let us now assume that the results holds up to an integer $n - 1$, $n \geq 2$, and let us prove that it also holds for n. In a commuting family of matrices in $\mathcal{M}_n$, we pick a matrix A which is not a multiple of I. Let $\lambda \in \mathbb{C}$ be an eigenvector for A and let $\mathcal{E}_\lambda := \{x \in \mathbb{C}^n : Ax = \lambda x\}$ be the corresponding eigenspace, which has dimension $k < n$. We can easily observe that $\mathcal{E}_\lambda$ is stable under the action of any A_i (i.e., $A_i(\mathcal{E}_\lambda) \subseteq \mathcal{E}_\lambda$) for every $i \in I$. It follows that

$$A_i = V \left[\begin{array}{c|c} \widetilde{A}_i & \mathrm{X} \\ \hline 0 & \mathrm{X} \end{array}\right] V^{-1} \qquad \text{for some invertible } V \in \mathcal{M}_n.$$

We can easily observe that the family $\{\widetilde{A}_i, i \in I\}$ is a commuting family in $\mathcal{M}_k$, and the induction hypothesis applies to yield the existence of an eigenvector $\widetilde{x} \in \mathbb{C}^k$ common to every $\widetilde{A}_i$. Then $x := V[\widetilde{x}, 0]^\top$ is an eigenvector common to every A_i. This finishes the inductive proof. $\square$

4 Exercises

Ex.1: Answer the in-line exercises.

Ex.2: We recall that $\operatorname{rk} A^* = \operatorname{rk} A$, where $A^* \in \mathcal{M}_{n\times m}$ denotes the adjoint of a matrix $A \in \mathcal{M}_{m\times n}$. In general, is it true that $\operatorname{nul} A^* = \operatorname{nul} A$? Establish that $\ker A = \ker A^*A$, deduce that $\operatorname{nul} A = \operatorname{nul} A^*A$ and that $\operatorname{rk} A = \operatorname{rk} A^*A = \operatorname{rk} A^* = \operatorname{rk} AA^*$, and finally conclude that $\operatorname{ran} A = \operatorname{ran} AA^*$.

Ex.3: Calculate $\operatorname{tr} A^*A$ and observe that $A = 0$ iff $\operatorname{tr} A^*A = 0$.

Ex.4: For $A, B \in \mathcal{M}_n$, prove that $AB = I$ implies $BA = I$. Is this true if A and B are not square? [Hint: consider the matrices C and $C^\top$, where $C = \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix}$.]

Ex.5: Consider the subset of $\mathcal{M}_n(\mathbb{R})$ defined by

$$\mathcal{G} := \{A \in \mathcal{M}_n(\mathbb{R}) : \det A = \pm 1 \text{ and } a_{i,j} \in \mathbb{Z} \text{ for all } 1 \leq i, j \leq n\}.$$

Prove that $I \in \mathcal{G}$, that $A, B \in \mathcal{G} \Rightarrow AB \in \mathcal{G}$ and that $A \in \mathcal{G} \Rightarrow A^{-1} \in \mathcal{G}$ (in other words, $\mathcal{G}$ is a multiplicative group).

Ex.6: Exercise 5 p. 37.

Ex.7: Given a polynomial $P(x) = c_d x^d + \cdots + c_1 x + c_0$ and a matrix $A \in \mathcal{M}_n$, prove that if λ is an eigenvalue of A, then $P(\lambda)$ is an eigenvalue of $P(A) := c_d A^d + \cdots + c_1 A + c_0 I$. Prove also that if $\lambda \neq 0$, then λ^{-1} is an eigenvalue of A^{-1}.

Ex.8: Exercise 3 p. 54.

Ex.9: Determine the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & t & \cdots & t \\ t & 1 & t & \vdots \\ t & \cdots & \ddots & t \\ t & \cdots & t & 1 \end{bmatrix},$$

and diagonalize it.

Lecture 1: Schur's Unitary Triangularization Theorem

This lecture introduces the notion of unitary equivalence and presents Schur's theorem and some of its consequences. It roughly corresponds to Sections 2.1, 2.2, and 2.4 of the textbook.

1 Unitary matrices

Definition 1. A matrix $U \in M_n$ is called unitary if

$$UU^* = I \quad (= U^*U).$$

If U is a real matrix (in which case U^* is just $U^\top$), then U is called an orthogonal matrix.

Observation: If $U, V \in \mathcal{M}_n$ are unitary, then so are $\bar{U}$, $U^\top$, U^* $(= U^{-1})$, UV.
Observation: If U is a unitary matrix, then

$$|\det U| = 1.$$

Examples: Matrices of reflection and of rotations are unitary (in fact, orthogonal) matrices. For instance, in $3D$-space,

$$\text{reflection along the } z\text{-axis:} \qquad U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \det U = -1,$$

$$\text{rotation along the } z\text{-axis:} \qquad U = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \det U = 1.$$

That these matrices are unitary is best seen using one of the alternate characterizations listed below.

Theorem 2. Given $U \in \mathcal{M}_n$, the following statements are equivalent:

(i) U is unitary,

(ii) U preserves the Hermitian norm, i.e.,

$$\|Ux\| = \|x\| \qquad \text{for all } x \in \mathbb{C}^n.$$

(iii) the columns $U_1, \ldots, U_n$ of U form an orthonormal system, i.e.,

$$\langle U_i, U_j \rangle = \delta_{i,j}, \qquad \text{in other words} \quad U_j^* U_i = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Proof. (i) $\Rightarrow$ (ii). If $U^*U = I$, then, for any $x \in \mathbb{C}^n$,

$$\|Ux\|_2^2 = \langle Ux, Ux\rangle = \langle U^*Ux, x\rangle = \langle x, x\rangle = \|x\|_2^2.$$

(ii) $\Rightarrow$ (iii). Let $(e_1, e_2, \ldots, e_n)$ denote the canonical basis of $\mathbb{C}^n$. Assume that U preserves the Hermitian norm of every vector. We obtain, for $j \in \{1, \ldots, n\}$,

$$\langle U_j, U_j\rangle = \|U_j\|_2^2 = \|Ue_j\|_2^2 = \|e_j\|_2^2 = 1.$$

Moreover, for $i, j \in \{1, \ldots, n\}$ with $i \neq j$, we have, for any complex number λ of modulus 1,

$$\begin{aligned}\Re\langle \lambda U_i, U_j\rangle &= \frac{1}{2}\left(\|\lambda U_i + U_j\|_2^2 - \|U_i\|_2^2 - \|U_j\|_2^2\right) = \frac{1}{2}\left(\|U(\lambda e_i + e_j)\|_2^2 - \|U(e_i)\|_2^2 - \|U(e_j)\|_2^2\right)\\ &= \frac{1}{2}\left(\|\lambda e_i + e_j\|_2^2 - \|e_i\|_2^2 - \|e_j\|_2^2\right) = 0.\end{aligned}$$

This does imply that $\langle U_i, U_j\rangle = 0$ (argue for instance that $|\langle \lambda U_i, U_j\rangle| = \Re\langle \lambda U_i, U_j\rangle$ for a properly chosen $\lambda \in \mathbb{C}$ with $|\lambda| = 1$).
(iii) $\Rightarrow$ (i). Observe that the (i, j)th entry of U^*U is $U_i^*U_j$ to realize that (iii) directly translates into $U^*U = I$. □

According to (iii), a unitary matrix can be interpreted as the matrix of an orthonormal basis in another orthonormal basis. In terms of linear maps represented by matrices A, the change of orthonormal bases therefore corresponds to the transformation $A \mapsto UAU^*$ for some unitary matrix U. This transformation defines the unitary equivalence.

Definition 3. Two matrices $A, B \in \mathcal{M}_n$ are called unitary equivalent if there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$B = UAU^*.$$

Observation: If $A, B \in \mathcal{M}_n$ are unitary equivalent, then

$$\sum_{1\leq i,j\leq n} |a_{i,j}|^2 = \sum_{1\leq i,j\leq n} |b_{i,j}|^2.$$

Indeed, recall form Lect.1-Ex.3 that $\sum_{1\leq i,j\leq n} |a_{i,j}|^2 = \operatorname{tr}(A^*A)$ and $\sum_{1\leq i,j\leq n} |b_{i,j}|^2 = \operatorname{tr}(B^*B)$, and then write

$$\begin{aligned}\operatorname{tr}\left(B^*B\right) &= \operatorname{tr}\left((UAU^*)^*(UAU^*)\right) = \operatorname{tr}\left(UA^*U^*UAU^*\right) = \operatorname{tr}\left(UA^*AU^*\right) = \operatorname{tr}\left(U^*UA^*A\right)\\ &= \operatorname{tr}\left(A^*A\right).\end{aligned}$$

2 Statement of Schur's theorem and some of its consequences

Schur's unitary triangularization theorem says that every matrix is unitarily equivalent to a triangular matrix. Precisely, it reads as follows.

Theorem 4. Given $A \in \mathcal{M}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, counting multiplicities, there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$A = U \begin{bmatrix} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} U^*.$$

Note that such a decomposition is far from unique (see Example 2.3.2 p.80-81). Let us now state a few consequences from Schur's theorem. First, Cayley–Hamilton theorem says that every square matrix annihilates its own characteristic polynomial.

Theorem 5. Given $A \in \mathcal{M}_n$, one has

$$p_A(A) = 0.$$

The second consequence of Schur's theorem says that every matrix is similar to a block-diagonal matrix where each block is upper triangular and has a constant diagonal. This is an important step in a possible proof of Jordan canonical form.

Theorem 6. Given $A \in \mathcal{M}_n$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, there is an invertible matrix $S \in \mathcal{M}_n$ such that

$$A = S \begin{bmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & T_k \end{bmatrix} S^{-1}, \qquad \text{where } T_i \text{ has the form } T_i = \begin{bmatrix} \lambda_i & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

The arguments for Theorems 5 and 6 (see next section) do not use the unitary equivalence stated in Schur's theorem, but merely the equivalence. The unitary equivalence is nonetheless crucial for the final consequence of Schur's theorem, which says that there are diagonalizable matrices arbitrary close to any matrix (in other words, the set of diagonalizable matrices is dense in $\mathcal{M}_n$).

Theorem 7. Given $A \in \mathcal{M}_n$ and $\varepsilon > 0$, there exists a diagonalizable matrix $\widetilde{A} \in \mathcal{M}_n$ such that

$$\sum_{1 \leq i,j \leq n} |a_{i,j} - \widetilde{a}_{i,j}|^2 < \varepsilon.$$

3 Proofs

Proof of Theorem 4. We proceed by induction on $n \geq 1$. For $n = 1$, there is nothing to do. Suppose now the result true up to an integer $n-1$, $n \geq 2$. Let $A \in \mathcal{M}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, counting multiplicities. Consider an eigenvector v_1 associated to the eigenvalue λ_1. We may assume that $\|v_1\|_2 = 1$. We use it to form an orthonormal basis $(v_1, v_2, \ldots, v_n)$. The matrix A is equivalent to the matrix of the linear map $x \mapsto Ax$ relative to the basis $(v_1, v_2, \ldots, v_n)$, i.e.,

$$A = V \left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & & & \\ \vdots & & \widetilde{A} & \\ 0 & & & \end{array}\right] V^{-1}, \tag{1}$$

where $V = \Big[v_1 \Big| \cdots \Big| v_n\Big]$ is the matrix of the system $(v_1, v_2, \ldots, v_n)$ relative to the canonical basis. Since this is a unitary matrix, the equivalence of (1) is in fact a unitary equivalence. Note that $p_A(x) = (\lambda_1 - x) p_{\widetilde{A}}(x)$, so that the eigenvalues of $\widetilde{A} \in \mathcal{M}_{n-1}$, counting multiplicities, are $\lambda_2, \ldots, \lambda_n$. We use the induction hypothesis to find a unitary matrix $\widetilde{W} \in \mathcal{M}_{n-1}$ such that

$$\widetilde{A} = \widetilde{W} \begin{bmatrix} \lambda_2 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \widetilde{W}^*, \qquad \text{i.e.,} \qquad \widetilde{W}^* \widetilde{A} \widetilde{W} = \begin{bmatrix} \lambda_2 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

Now observe that

$$\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \widetilde{W} & \\ 0 & & & \end{array}\right]^* \left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & & & \\ \vdots & & \widetilde{A} & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \widetilde{W} & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \widetilde{W}^* & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & & & \\ \vdots & & \widetilde{A}\widetilde{W} & \\ 0 & & & \end{array}\right]$$

$$= \left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & & & \\ \vdots & & \widetilde{W}^*\widetilde{A}\widetilde{W} & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & \lambda_2 & \cdots & \mathrm{x} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{array}\right].$$

Since $W := \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \widetilde{W} \end{array}\right]$ is a unitary matrix, this reads

$$\left[\begin{array}{c|ccc} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & & & \\ \vdots & & \widetilde{A} & \\ 0 & & & \end{array}\right] = W \begin{bmatrix} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_2 & \cdots & \mathrm{x} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} W^*. \tag{2}$$

Putting the unitary equivalences (1) and (2) together shows the result of Theorem 4 (with $U = VW$) for the integer n. This concludes the inductive proof. □

Now that Schur's theorem is established, we may prove the consequences stated in Section 2.

Proof of Theorem 5. First attempt, valid when A is diagonalizable. In this case, there is a basis $(v_1, \dots, v_n)$ of eigenvectors associated to (not necessarily distinct) eigenvalues $\lambda_1, \dots, \lambda_n$. It is enough to show that the matrix $p_A(A)$ vanishes on each basis vector v_i, $1 \leq i \leq n$. Note that

$$p_A(A) = (\lambda_1 I - A) \cdots (\lambda_n I - A) = [(\lambda_1 I - A) \cdots (\lambda_{i-1} I - A)(\lambda_{i+1} I - A) \cdots (\lambda_n I - A)](\lambda_i I - A),$$

because $(\lambda_i I - A)$ commutes with all $(\lambda_j I - A)$. Then the expected results follows from

$$p_A(A)(v_i) = [\cdots](\lambda_i I - A)(v_i) = [\cdots](0) = 0.$$

Final proof. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A \in \mathcal{M}_n$, counting multiplicities. According to Schur's theorem, we can write

$$A = STS^{-1}, \qquad \text{where} \quad T = \begin{bmatrix} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_2 & \cdots & \mathrm{x} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Since $P(A) = SP(T)S^{-1}$ for any polynomial P, we have in particular $p_A(A) = Sp_A(T)S^{-1}$, so it is enough to show that $p_A(T) = 0$. We compute

$$p_A(T) = (\lambda_1 I - T)(\lambda_2 I - T) \cdots (\lambda_n I - T)$$

$$= \left[\begin{array}{c|c|ccc} 0 & \mathrm{x} & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & \mathrm{x} & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \mathrm{X} & \\ 0 & 0 & & & \end{array}\right] \left[\begin{array}{cc|ccc} \mathrm{x} & \mathrm{x} & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} & \cdots & \mathrm{x} \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \mathrm{X} & \\ 0 & 0 & & & \end{array}\right] \cdots \left[\begin{array}{ccc|c|c} & & & \mathrm{x} & \mathrm{x} \\ & \mathrm{X} & & \vdots & \vdots \\ & & & \mathrm{x} & \mathrm{x} \\ \hline 0 & \cdots & 0 & \mathrm{x} & \mathrm{x} \\ \hline 0 & \cdots & 0 & 0 & 0 \end{array}\right].$$

Using repeatedly the observation about multiplication of block-triangular matrices that

$$\left[\begin{array}{c|c|c} 0 & \mathrm{x} & \mathrm{x} \\ \hline 0 & \mathrm{x} & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} \end{array}\right] \left[\begin{array}{c|c|c} \mathrm{x} & \mathrm{x} & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} \end{array}\right] = \left[\begin{array}{c|c|c} 0 & 0 & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} \\ \hline 0 & 0 & \mathrm{x} \end{array}\right],$$

the zero-block on the top left propagates until we obtain $p_A(T) = 0$ — a rigorous proof of this propagation statement should proceed by induction. □

To establish the next consequence of Schur's theorem, we will use the following result.

Lemma 8. If $A \in \mathcal{M}_m$ and $B \in \mathcal{M}_n$ are two matrices with no eigenvalue in common, then the matrices

$$\left[\begin{array}{c|c} A & M \\ \hline 0 & B \end{array}\right] \qquad \text{and} \qquad \left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right]$$

are equivalent for any choice of $M \in \mathcal{M}_{m \times n}$.

Proof. For $X \in \mathcal{M}_{m\times n}$, consider the matrices S and S^{-1} given by

$$S = \left[\begin{array}{c|c} I & X \\ \hline 0 & I \end{array}\right] \qquad \text{and} \qquad S^{-1} = \left[\begin{array}{c|c} I & -X \\ \hline 0 & I \end{array}\right].$$

We compute

$$S^{-1}\left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right] S = \left[\begin{array}{c|c} A & AX - XB \\ \hline 0 & B \end{array}\right].$$

The result will follow as soon as we can find $X \in \mathcal{M}_{m\times n}$ such that $AX - XB = M$ for our given $M \in \mathcal{M}_{m\times n}$. If we denote by $\mathcal{F}$ the linear map

$$\mathcal{F} : X \in \mathcal{M}_{m\times n} \mapsto AX - XB \in \mathcal{M}_{m\times n},$$

it is enough to show that $\mathcal{F}$ is surjective. But since $\mathcal{F}$ is a linear map from $\mathcal{M}_{m\times n}$ into itself, it is therefore enough to show that $\mathcal{F}$ is injective, i.e., that

$$AX - XB = 0 \overset{?}{\Longrightarrow} X = 0. \tag{3}$$

To see why this is true, let us consider $X \in \mathcal{M}_{m\times n}$ such that $AX = XB$, and observe that

$$\begin{aligned} A^2X &= A(AX) = A(XB) = (AX)B = (XB)B = XB^2, \\ A^3X &= A(A^2X) = A(XB^2) = (AX)B^2 = (XB)B^2 = XB^3, \end{aligned}$$

etc., so that $P(A)X = XP(B)$ for any polynomial P. If we choose $P = p_A$ as the characteristic polynomial of A, Cayley–Hamilton theorem implies

$$0 = Xp_A(B). \tag{4}$$

Denoting by $\lambda_1, \ldots, \lambda_n$ the eigenvalues of A, we have $p_A(B) = (\lambda_1 I - B)\cdots(\lambda_n I - B)$. Note that each factor $(\lambda_i I - B)$ is invertible, since none of the λ_i is an eigenvalue of B, so that $p_A(B)$ is itself invertible. We can now conclude from (4) that $X = 0$. This establishes (3), and finishes the proof. □

We could have given a less conceptual proof of Lemma 8 in case both A and B are upper triangular (see exercises), which is actually what the proof presented below requires.

Proof of Theorem 6. For $A \in \mathcal{M}_n$, we sort its eigenvalues as $\lambda_1, \ldots, \lambda_1; \lambda_2, \ldots, \lambda_2; \ldots; \lambda_k, \ldots, \lambda_k$, counting multiplicities. Schur's theorem guarantees that A is similar to the matrix

$$\begin{bmatrix} T_1 & \mathrm{X} & \cdots & \mathrm{X} \\ 0 & T_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{X} \\ 0 & \cdots & 0 & T_k \end{bmatrix} \qquad \text{where} \qquad T_i = \begin{bmatrix} \lambda_i & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

We now use the Lemma 8 repeatedly in the chain of equivalences

$$A \sim \left[\begin{array}{c|cccc} T_1 & \mathrm{X} & \cdots & \cdots & \mathrm{X} \\ \hline 0 & T_2 & \mathrm{X} & \cdots & \mathrm{X} \\ \vdots & 0 & T_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathrm{X} \\ 0 & 0 & \cdots & 0 & T_k \end{array}\right] \sim \left[\begin{array}{c|cccc} T_1 & 0 & \cdots & \cdots & 0 \\ \hline 0 & T_2 & \mathrm{X} & \cdots & \mathrm{X} \\ \vdots & 0 & T_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathrm{X} \\ 0 & 0 & \cdots & 0 & T_k \end{array}\right] = \left[\begin{array}{cc|ccc} T_1 & 0 & & & \\ 0 & T_2 & & \mathrm{X} & \\ \hline & & T_3 & \cdots & \mathrm{X} \\ & 0 & \vdots & \ddots & \vdots \\ & & 0 & \cdots & T_k \end{array}\right]$$

$$\sim \left[\begin{array}{cc|ccc} T_1 & 0 & & & \\ 0 & T_2 & & 0 & \\ \hline & & T_3 & \cdots & \mathrm{X} \\ & 0 & \vdots & \ddots & \vdots \\ & & 0 & \cdots & T_k \end{array}\right] = \left[\begin{array}{ccc|ccc} T_1 & 0 & 0 & & & \\ 0 & T_2 & 0 & & \mathrm{X} & \\ 0 & 0 & T_3 & & & \\ \hline & & & \ddots & \mathrm{X} \\ & & 0 & 0 & T_k \end{array}\right] \sim \left[\begin{array}{ccc|ccc} T_1 & 0 & 0 & & & \\ 0 & T_2 & 0 & & 0 & \\ 0 & 0 & T_3 & & & \\ \hline & & & \ddots & 0 \\ & & 0 & 0 & T_k \end{array}\right] \sim \cdots$$

$$\sim \begin{bmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & T_k \end{bmatrix}$$

This is the announced result. □

Proof of Theorem 7. Let us sort the eigenvalues of A as $\lambda_1 \geq \cdots \geq \lambda_n$. According to Schur's theorem, there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$A = U \begin{bmatrix} \lambda_1 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} U^*.$$

If $\widetilde{\lambda}_i := \lambda_i + i\eta$ and $\eta > 0$ is small enough to guarantee that $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ are all distinct, we set

$$\widetilde{A} = U \begin{bmatrix} \widetilde{\lambda}_1 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \widetilde{\lambda}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \widetilde{\lambda}_n \end{bmatrix} U^*.$$

In this case, the eigenvalues of $\widetilde{A}$ (i.e., $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$) are all distinct, hence $\widetilde{A}$ is diagonalizable. We now notice that

$$\sum_{1 \leq i,j \leq n} |a_{i,j} - \widetilde{a}_{i,j}|^2 = \mathrm{tr}\, \left((A - \widetilde{A})^* (A - \widetilde{A}) \right).$$

But since $A - \widetilde{A}$ is unitarily equivalent of the diagonal matrix $\mathrm{diag}[\lambda_1 - \widetilde{\lambda}_1, \ldots, \lambda_n - \widetilde{\lambda}_n]$, this quantity equals $\sum_{1 \leq i \leq n} |\lambda_i - \widetilde{\lambda}_i|^2$. It follows that

$$\sum_{1 \leq i,j \leq n} |a_{i,j} - \widetilde{a}_{i,j}|^2 = \sum_{1 \leq i \leq n} i^2 \eta^2 < \varepsilon,$$

provided η is chosen small enough to have $\eta^2 < \varepsilon / \sum_i i^2$. □

4 Exercises

Ex.1: Is the sum of unitary matrices also unitary?

Ex.2: Exercise 2 p. 71.

Ex.3: When is a diagonal matrix unitary?

Ex.4: Exercise 12 p. 72.

Ex.5: Given $A \in \mathcal{M}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, counting multiplicities, prove that there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$A = U \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \mathrm{x} & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \mathrm{x} & \cdots & \mathrm{x} & \lambda_n \end{bmatrix} U^*.$$

Ex.6: Prove that a matrix $U \in \mathcal{M}_n$ is unitary iff it preserves the Hermitian inner product, i.e., iff $\langle Ux, Uy \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^n$.

Ex.7: Given a matrix $A \in \mathcal{M}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, counting multiplicities, and given a polynomial $P(x) = c_d x^d + \cdots + c_1 x + c_0$, prove that the eigenvalues of $P(A)$ are $P(\lambda_1), \ldots, P(\lambda_n)$, counting multiplicities.
(Note: this is not quite the same as Exercise 7 from Lecture 1.)

Ex.8: Exercise 5 p. 97.

Ex.9: For any matrix $A \in \mathcal{M}_n$, prove that

$$\det(\exp(A)) = \exp(\mathrm{tr}\,(A)).$$

(The exponential of a matrix is defined as the convergent series $\exp(A) = \sum_{k \geq 0} \frac{1}{k!} A^k$.)

Ex.10: Given an invertible matrix $A \in \mathcal{M}_n$, show that its inverse A^{-1} can be expressed as a polynomial of degree $\leq n - 1$ in A.

Ex.11: Without using Cayley-Hamilton theorem, prove that if $T \in \mathcal{M}_m$ and $\widetilde{T} \in \mathcal{M}_n$ are two upper triangular matrices with no eigenvalue in common, then the matrices

$$\left[\begin{array}{c|c} T & M \\ \hline 0 & \widetilde{T} \end{array}\right] \quad \text{and} \quad \left[\begin{array}{c|c} T & 0 \\ \hline 0 & \widetilde{T} \end{array}\right]$$

are equivalent for any choice of $M \in \mathcal{M}_{m \times n}$.
[Hint: Observe that you need to show $TX = X\widetilde{T} \Longrightarrow X = 0$. Start by considering the element in the lower left corner of the matrix $TX = X\widetilde{T}$ to show that $x_{m,1} = 0$, then consider the diagonal $i - j = m - 2$ (the one just above the lower left corner) of the matrix $TX = X\widetilde{T}$ to show that $x_{m-1,1} = 0$ and $x_{m,2} = 0$, etc.]

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. □

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

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

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

Proof. (i) $\Leftrightarrow$ (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) $\Rightarrow$ (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

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

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

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

On the other hand, we have

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

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 $A_iA_j = A_jA_i$ 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 $M^*B = BN^*$.

Lecture 3: QR-Factorization

This lecture introduces the Gram–Schmidt orthonormalization process and the associated QR-factorization of matrices. It also outlines some applications of this factorization. This corresponds to section 2.6 of the textbook. In addition, supplementary information on other algorithms used to produce QR-factorizations is given.

1 Gram–Schmidt orthonormalization process

Consider n linearly independent vectors $u_1, \dots, u_n$ in $\mathbb{C}^m$. Observe that we necessarily have $m \geq n$. We wish to 'orthonormalize' them, i.e., to create vectors $v_1, \dots, v_n$ such that

$$(v_1, \dots, v_k) \text{ is an orthonormal basis for } V_k := \operatorname{span}[u_1, \dots, u_k] \quad \text{for all } 1k \in [1:n].$$

It is always possible to find such vectors, and in fact they are uniquely determined if the additional condition $\langle v_k, u_k \rangle > 0$ is imposed. The step-by-step construction is based on the following scheme.

Suppose that $v_1, \dots, v_{k-1}$ have been obtained; search in V_k for a vector

$$\widetilde{v}_k = u_k + \sum_{i=1}^{k-1} c_{k,i}\, v_i \quad \text{such that} \quad \widetilde{v}_k \perp V_{k-1};$$

the conditions $0 = \langle \widetilde{v}_k, v_i \rangle = \langle u_k, v_i \rangle + c_{k,i}$, $i \in [1 : k-1]$, impose the choice $c_{k,i} = -\langle u_k, v_i \rangle$; now that $\widetilde{v}_k$ is completely determined, normalize it to obtain the vector $v_k = \dfrac{1}{\|\widetilde{v}_k\|} \widetilde{v}_k$.

For instance, let us write down explicitly all the steps in the orthonormalization process for the vectors

$$\underline{u_1 = [6,3,2]^\top}, \quad \underline{u_2 = [6,6,1]^\top}, \quad \underline{u_3 = [1,1,1]^\top}.$$

- $\widetilde{v}_1 = u_1, \quad \|\widetilde{v}_1\| = \sqrt{36+3+4} = 7, \quad \underline{v_1 = 1/7\,[6,3,2]^\top};$
- $\widetilde{v}_2 = u_2 + \alpha v_1, \quad 0 = \langle \widetilde{v}_2, v_1 \rangle \Rightarrow \alpha = -\langle u_2, v_1 \rangle = -(16+18+2)/7, \quad \underline{\alpha = -8},$
  $\widetilde{v}_2 = 1/7\,[7 \cdot 6 - 8 \cdot 6, 7 \cdot 6 - 8 \cdot 3, 7 \cdot 1 - 8 \cdot 2]^\top = 1/7\,[-6, 18, -9]^\top = 3/7\,[-2, 6, -3]^\top,$
  $\|\widetilde{v}_2\| = 3/7\,\sqrt{4+36+9} = 3, \quad \underline{v_2 = 1/7\,[-2,6,-3]^\top};$
- $\widetilde{v}_3 = u_3 + \beta v_2 + \gamma v_1, \quad \begin{matrix} 0 = \langle \widetilde{v}_3, v_2 \rangle, \\ 0 = \langle \widetilde{v}_3, v_1 \rangle, \end{matrix} \Rightarrow \begin{matrix} \beta = -\langle u_3, v_2 \rangle = -(-2+6-3)/7, & \underline{\beta = -1/7}, \\ \gamma = -\langle u_3, v_1 \rangle = -(6+3+2)/7, & \underline{\gamma = -11/7}, \end{matrix}$
  $\widetilde{v}_3 = 1/49\,[49+2-66, 49-6-33, 49+3-22]^\top = 1/49\,[-15, 10, 30]^\top = 5/49\,[-3,2,6]^\top,$
  $\|\widetilde{v}_3\| = 5/49\,\sqrt{9+4+36} = 5/7, \quad \underline{v_3 = 1/7\,[-3,2,6]^\top}.$

2 QR-factorization

Theorem 1. For a nonsingular $A \in \mathcal{M}_n$, there exists a unique pair of unitary matrix $Q \in \mathcal{M}_n$ and upper triangular matrix $R \in \mathcal{M}_n$ with positive diagonal entries such that

$$A = QR.$$

The QR-factorization can be used for the following tasks:

- solving linear systems according to

  $$[Ax = b] \iff [Qy = b, \quad y = Rx],$$

  since the system $y = Rx$ is easy to solve [backward substitution], and the system $Qy = b$ is even easier to solve [take $y = Q^*b$];
- calculate the (modulus of the) determinant and find the inverse [$|\det A| = \prod_{i=1}^n r_{i,i}$, $A^{-1} = R^{-1}Q^*$];
- find the Cholesky factorization of a positive definite matrix $B = A^*A \in \mathcal{M}_n$, $A \in \mathcal{M}_n$ being nonsingular (we will later see why every positive definite matrix can be factored in this way), i.e., find a factorization

  $$B = LL^*,$$

  where $L \in \mathcal{M}_n$ is lower triangular with positive diagonal entries [$L = R^*$];
- find a Schur's factorization of a matrix $A \in \mathcal{M}_n$ via the QR-algorithm defined by

  $$\begin{aligned} A_0 &:= A, & A_0 &=: Q_0R_0, \\ A_1 &:= R_0Q_0, & A_1 &=: Q_1R_1, \\ &\vdots & &\vdots \\ A_k &:= R_{k-1}Q_{k-1}, & A_1 &=: Q_kR_k, \\ &\vdots & &\vdots \end{aligned}$$

  Note that A_k is always unitarily equivalent to A. If all the eigenvalues of A have distinct moduli, then A_k tends to an upper triangular matrix T (which is therefore unitarily equivalent to A, see Exercise 3). The eigenvalues of A are read on the diagonal of T.

In the general case of nonsingular or nonsquare matrices, the QR-factorization reads:

Theorem 2. For $A \in \mathcal{M}_{m\times n}$, $m \geq n$, there exists a matrix $Q \in \mathcal{M}_{m\times n}$ with orthonormal columns and an upper triangular matrix $R \in \mathcal{M}_n$ such that

$$A = QR.$$

Beware that the QR-factorization of a rectangular matrix A is not always understood with Q rectangular and R square, but sometimes with Q square and R rectangular, as with the MATLAB command `qr`.

3 Proof of Theorems 1 and 2

Uniqueness: Suppose that $A = Q_1R_1 = Q_2R_2$ where Q_1, Q_2 are unitary and R_1, R_2 are upper triangular with positive diagonal entries. Then

$$M := R_1R_2^{-1} = Q_1^*Q_2.$$

Since M is a unitary (hence normal) matrix which is also upper triangular, it must be diagonal (see Lemma 4 of Lecture 2). Note also that the diagonal entries of M are positive (because the upper triangular matrices R_1 and R_2^{-1} have positive diagonal entries) and of modulus one (because M is a diagonal unitary matrix). We deduce that $M = I$, and consequently that

$$R_1 = R_2, \qquad Q_1 = Q_2.$$

Existence: Let us consider a matrix $A \in \mathcal{M}_{m\times n}$ with $m \geq n$, and let $u_1, \ldots, u_n \in \mathbb{C}^m$ denote its columns. We may assume that $u_1, \ldots, u_n$ are linearly independent (otherwise limiting arguments can be used, see Exercise 4). Then the result is just a matrix interpretation of the Gram–Schmidt orthonormalization process of m linearly independent vectors in $\mathbb{C}^n$. Indeed, the Gram–Schmidt algorithm produces orthonormal vectors $v_1, \ldots, v_n \in \mathbb{C}^m$ such that, for each $j \in [1:n]$,

$$u_j = \sum_{k=1}^{j} r_{k,j}\, v_k = \sum_{k=1}^{n} r_{k,j}\, v_k, \tag{1}$$

with $r_{k,j} = 0$ for $k > j$, in other words, $R := [r_{i,j}]_{i,j=1}^n$ is an $n \times n$ upper triangular matrix. The n equations (1) reduce, in matrix form, to $A = QR$, where Q is the $m \times n$ matrix whose columns are the orthonormal vectors $v_1, \ldots, v_n$.

[To explain the other QR-factorization, let us complete $v_1, \ldots, v_n$ with $v_{m+1}, \ldots, v_n$ to form an orthonormal basis $(v_1, \ldots, v_m)$ of $\mathbb{C}^m$. The analogs of the equations (1), i.e., $u_j = \sum_{k=1}^m r_{k,j}\, v_k$ with $r_{k,j} = 0$ for $k > j$, read $A = QR$, where Q is the $m \times m$ orthogonal matrix with columns $v_1, \ldots, v_m$ and R is an $m \times n$ upper triangular matrix.]

To illustrate the matrix interpretation, observe that the orthonormalization carried out in Section 1 translates into the factorization [identify all the entries]

$$\begin{bmatrix} 6 & 6 & 1 \\ 3 & 6 & 1 \\ 2 & 1 & 1 \end{bmatrix} = \underbrace{\frac{1}{7}\begin{bmatrix} 6 & -2 & -3 \\ 3 & 6 & 2 \\ 2 & -3 & 6 \end{bmatrix}}_{\text{unitary}} \underbrace{\begin{bmatrix} 7 & 8 & 11/7 \\ 0 & 3 & 1/7 \\ 0 & 0 & 5/7 \end{bmatrix}}_{\text{upper triangular}}.$$

An aside: other QR-factorization algorithms

The Gram–Schmidt algorithm has the disadvantage that small imprecisions in the calculation of inner products accumulate quickly and lead to effective loss of orthogonality. Alternative ways to obtain a QR-factorization are presented below on some examples. They are based on the following idea and exploits the fact that the computed product of unitary matrices gives a unitary matrix with acceptable error.

Multiply the (real) matrix A on the left by some orthogonal matrices Q_i which 'eliminate' some entries below the main 'diagonal', until the result is an upper triangular matrix R, thus

$$Q_k \cdots Q_2 Q_1 A = R \quad \text{yields} \quad A = QR, \quad \text{with} \quad Q = Q_1^* Q_2^* \cdots Q_k^*.$$

We mainly know two types of unitary transformations, namely rotations and reflexions. Therefore, the two methods we describe are associated with Givens rotations [preferred when A is sparse] and Householder reflections [preferred when A is dense].

Givens rotations

The matrix

$$\Omega^{[i,j]} := \begin{bmatrix} \ddots & & & i & & j & & \\ & 1 & & \vdots & 0 & \vdots & 0 & \\ i & \cdots & & \cos\theta & \cdots & \sin\theta & \cdots & \\ & 0 & & \vdots & \ddots & \vdots & 0 & \\ j & \cdots & & -\sin\theta & \cdots & \cos\theta & \cdots & \\ & 0 & & \vdots & 0 & \vdots & 1 & \\ & & & & & & & \ddots \end{bmatrix}$$

corresponds to a rotation along the two-dimensional space $\mathrm{span}[e_i, e_j]$. The rows of the matrix $\Omega^{[i,j]}A$ are the same as the rows of A, except for the i-th and j-th rows, which are linear combinations of the i-th and j-th rows of A. By choosing θ appropriately, we may introduce a zero at a prescribed position on one of these rows. Consider for example the matrix $A = \begin{bmatrix} 6 & 6 & 1 \\ 3 & 6 & 1 \\ 2 & 1 & 1 \end{bmatrix}$ of the end of Section 1. We pick $\Omega^{[1,2]}$ so that $\Omega^{[1,2]}A = \begin{bmatrix} 6 & 6 & 1 \\ 0 & \times & \times \\ \times & \times & \times \end{bmatrix}$. Then we pick $\Omega^{[1,3]}$ so that $\Omega^{[1,3]}\Omega^{[1,2]}A = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$. Finally, thanks to the leading zeros in the second and third rows,

we can pick $\Omega^{[2,3]}$ so that $\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}$. The matrix $\left[\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}\right]^*$ is the unitary matrix required in the factorization of A.

Householder reflections

The reflection in the direction of a vector v transforms v into $-v$ while leaving the space $v^\perp$ unchanged. It can therefore be expressed through the Hermitian unitary matrix

$$H_v := I - \frac{2}{\|v\|^2}\, vv^*.$$

Consider the matrix $A = \begin{bmatrix} 6 & 6 & 1 \\ 3 & 6 & 1 \\ 2 & 1 & 1 \end{bmatrix}$ once again. We may transform $u_1 = [6,3,2]^\top$ into $7e_1 = [7,0,0]^\top$ by way of the reflection in the direction $v_1 = u_1 - 7e_1 = [-1,3,2]^\top$. The latter is represented by the matrix

$$H_{v_1} = I - \frac{2}{\|v_1\|^2}\, v_1 v_1^* = I - \frac{1}{7}\begin{bmatrix} 1 & -3 & -2 \\ -3 & 9 & 6 \\ -2 & 6 & 4 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 6 & 3 & 2 \\ 3 & -2 & -6 \\ 2 & -6 & 3 \end{bmatrix}.$$

Then the matrix $H_{v_1}A$ has the form $\begin{bmatrix} 7 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$, where the precise expression for the second column is

$$H_{v_1}u_2 = u_2 - \frac{\langle v_1, u_2\rangle}{7}\, v_1 = u_2 - 2v_1 = \begin{bmatrix} 8 \\ 0 \\ -3 \end{bmatrix}.$$

To cut the argument short, we may observe at this point that the multiplication of $H_{v_1}A$ on the left by the permutation matrix $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ [which can be interpreted as $H_{e_2-e_3}$] exchanges the second and third rows, thus gives an upper triangular matrix. In conclusion, the orthogonal matrix Q has been obtained as

$$(PH_{v_1})^* = H_{v_1}^* P^* = H_{v_1}P = \frac{1}{7}\begin{bmatrix} 6 & 2 & 3 \\ 3 & -6 & -2 \\ 2 & 3 & -6 \end{bmatrix}.$$

4 Exercises

Ex.1: Prove that a matrix $A \in \mathcal{M}_{m\times n}$, $m \leq n$, can be factored as $A = LP$ where $L \in M_m$ is lower triangular and $P \in \mathcal{M}_{m\times n}$ has orthonormal rows.

Ex.2: Prove the uniqueness of the Cholesky factorization of a positive definite matrix.

Ex.3: Exercise 5 p. 117.

Ex.4: Fill in the details of the following argument: for $A \in \mathcal{M}_{m\times n}$ with $m \geq n$, there exists a sequence of matrices $A_k \in \mathcal{M}_{m\times n}$ with linearly independent columns such that $A_k \to A$ as $k \to \infty$; each A_k can be written as $A_k = Q_k R_k$ where $Q_k \in \mathcal{M}_{m\times n}$ has orthonormal columns and $R_k \in \mathcal{M}_n$ is upper triangular; there exists a subsequence (Q_{k_j}) converging to a matrix $Q \in \mathcal{M}_{m\times n}$ with orthonormal columns, and the sequence (R_{k_j}) converges to an upper triangular matrix $R \in \mathcal{M}_n$; taking the limit when $j \to \infty$ yields $A = QR$.

Ex.5: Fill in the numerical details in the section on Givens rotations.

Lecture 4: Jordan Canonical Forms

This lecture introduces the Jordan canonical form of a matrix — we prove that every square matrix is equivalent to a (essentially) unique Jordan matrix and we give a method to derive the latter. We also introduce the notion of minimal polynomial and we point out how to obtain it from the Jordan canonical form. Finally, we make an encounter with companion matrices.

1 Jordan form and an application

Definition 1. A Jordan block is a matrix of the form $J_1(\lambda) = \lambda \in \mathbb{C}$ when $k = 1$ and

$$J_k(\lambda) = \underbrace{\left.\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{bmatrix}\right\} k}_{k} \qquad \text{when } k \geq 2.$$

Theorem 2. For any $A \in \mathcal{M}_n$, there is an invertible matrix $S \in \mathcal{M}_n$ such that

$$A = S \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_k}(\lambda_k) \end{bmatrix} S^{-1} =: SJS^{-1},$$

where $n_1 + n_2 + \cdots + n_k = n$. The numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the (not necessarily distinct) eigenvalues of A. The Jordan matrix J is unique up to permutation of the blocks.

Observation: two matrices close to one another can have Jordan forms far apart, for instance

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & \varepsilon \\ 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Thus, any algorithm determining the Jordan canonical form is inevitably unstable! Try typing `help jordan` in MATLAB...

The Jordan form can be useful when solving a system of ordinary differential equations in the form $[x' = Ax, x(0) = x_0]$. If $A = SJS^{-1}$ is the Jordan canonical form, then the change of unknown functions $\widetilde{x} = S^{-1}x$ transforms the original system to $[\widetilde{x}' = J\widetilde{x}, \widetilde{x}(0) = \widetilde{x}_0]$. Writing

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_k}(\lambda_k) \end{bmatrix} \qquad \text{and} \qquad \widetilde{x} = \begin{bmatrix} u_{(1)} \\ \hline u_{(2)} \\ \hline \vdots \\ \hline u_{(k)} \end{bmatrix},$$

the new system decouples as $u'_{(\ell)} = J_{n_\ell}(\lambda_\ell)u_{(\ell)}$, $\ell \in [1:k]$. Each of these k systems has the form

$$\begin{bmatrix} u'_1 \\ u'_2 \\ \cdots \\ u'_m \end{bmatrix} = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_m \end{bmatrix}.$$

This can be solved by backward substitution: first, use $u_m(t) = \lambda u_m(t)$ to derive

$$u_m(t) = u_m(0)e^{\lambda t}.$$

Then, $u'_{m-1}(t) = \lambda u_{m-1}(t) + u_m(t)$ reads $\dfrac{d}{dt}[u_{m-1}(t)e^{-\lambda t}] = [u'_{m-1}(t) - \lambda u_{m-1}(t)]e^{-\lambda t} = u_m(0)$, so that

$$u_{m-1}(t) = [u_{m-1}(0) + u_m(0)t]e^{\lambda t}.$$

Next, $u'_{m-2}(t) = \lambda u_{m-2}(t) + u_{m-1}(t)$ reads $\dfrac{d}{dt}[u_{m-2}(t)e^{-\lambda t}] = [u'_{m-2}(t) - \lambda u_{m-2}(t)]e^{-\lambda t} = u_{m-1}(0) + u_m(0)t$, so that

$$u_{m-2}(t) = [u_{m-2}(0) + u_{m-1}(0)t + u_m(0)t^2]e^{\lambda t},$$

etc. The whole vector $[u_1, \ldots, u_m]^\top$ can be determined in this fashion.

2 Proof of Theorem 2

Uniqueness: The argument is based on the observation that, for

$$N_k := \underbrace{\left.\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}\right\} k}_{k},$$

we have $N_k^\ell = 0$ if $\ell \geq k$ and

$$N_k^\ell = \begin{bmatrix} & & \overset{\leftarrow \ell \rightarrow}{} & & \\ 0 & \cdots & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 1 & \vdots \\ 0 & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \qquad \text{has rank } k - \ell \quad \text{if } \ell < k.$$

Let $\mu_1 < \ldots < \mu_\ell$ be the distinct values of $\lambda_1, \ldots, \lambda_k$. After permutation,

$$A \sim \left[\begin{array}{ccc|c|cc} J_{n_{1,1}}(\mu_1) & & & & & \\ & \ddots & & & & \\ & & J_{n_{1,h_1}}(\mu_1) & & & \\ \hline & & & \ddots & & \\ \hline & & & & J_{n_{\ell,1}}(\mu_\ell) & \\ & & & & & \ddots \\ & & & & & J_{n_{\ell,h_\ell}}(\mu_\ell) \end{array}\right].$$

We are going to prove that for each μ_i (say, for μ_1) and for each m, the number of blocks $J_m(\mu_i)$ of size m is uniquely determined by A. From

$$A - \mu_1 I \sim \left[\begin{array}{ccc|c|c} N_{n_{1,1}} & & & & \\ & \ddots & & & \\ & & N_{n_{1,h_1}} & & \\ \hline & & & \ddots & \\ \hline & & & & \text{Full Rank} \end{array}\right],$$

we obtain, for each $m \geq 0$,

$$(A - \mu_1 I)^m \sim \left[\begin{array}{ccc|c|c} N_{n_{1,1}}^m & & & & \\ & \ddots & & & \\ & & N_{n_{1,h_1}}^m & & \\ \hline & & & \ddots & \\ \hline & & & & \text{Full Rank} \end{array}\right].$$

Therefore, we derive, for each $m \geq 1$,

$$\begin{aligned} \operatorname{rank}\left((A-\mu_1 I)^{m-1}\right) - \operatorname{rank}\left((A-\mu_1 I)^m\right) &= \begin{pmatrix} 1 & \text{if } m \leq n_{1,1} \\ 0 & \text{otherwise} \end{pmatrix} + \cdots + \begin{pmatrix} 1 & \text{if } m \leq n_{1,h_1} \\ 0 & \text{otherwise} \end{pmatrix} \\ &= [\text{number of Jordan blocks of size } \geq m] =: j_{\geq m}. \end{aligned}$$

Since the numbers $j_{\geq m}$ are completely determined by A, so are the numbers $j_m = j_{\geq m} - j_{\geq m+1}$, which represent the numbers of Jordan blocks of size $= m$. □

Let us emphasize that the previous argument is also the basis of a method to find the Jordan canonical form. We illustrate the method on the example

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 0 & 4 & -1 \end{bmatrix}.$$

The eigenvalues of A are the roots of $\det(xI - A) = (1-x)^3$ — calculation left to the reader — hence 1 is the unique eigenvalue of A. Therefore, there are three possibilities for the Jordan canonical form J of A:

$$J_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

The observation that $\operatorname{rank}(J - I) = \operatorname{rank}(A - I) = 1$ — calculation left to the reader — shows that $J = J_2$ (since $\operatorname{rank}(J_1 - I) = 0$, $\operatorname{rank}(J_2 - I) = 1$, $\operatorname{rank}(J_3 - I) = 2$).

Existence: Let $\mu_1 < \cdots < \mu_\ell$ be the distinct eigenvalues of A. We know (see Lecture 1) that

$$A \sim \begin{bmatrix} T_1 & 0 & \cdots & 0 \\ 0 & T_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & T_\ell \end{bmatrix}, \qquad \text{where} \quad T_i = \begin{bmatrix} \mu_i & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \mu_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \mu_i \end{bmatrix}.$$

It is now enough to show that each matrix T_i is equivalent to a Jordan matrix, i.e., that

$$T = \begin{bmatrix} \mu & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & \mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & \mu \end{bmatrix} \in \mathcal{M}_m \Longrightarrow T \sim J = \begin{bmatrix} J_{m_1}(\mu) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\mu) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{m_k}(\mu) \end{bmatrix},$$

with $m_1 + m_2 + \cdots + m_k = m$, or equivalently that

$$T = \begin{bmatrix} 0 & \mathrm{x} & \cdots & \mathrm{x} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{x} \\ 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathcal{M}_m \Longrightarrow T \sim M = \begin{bmatrix} N_{m_1} & 0 & \cdots & 0 \\ 0 & N_{m_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & N_{m_k} \end{bmatrix}.$$

In other words, we want to find vectors

$$x_1, T(x_1), \cdots, T^{m_1-1}(x_1), \ldots, x_k, T(x_k), \cdots, T^{m_k-1}(x_k)$$

forming a basis of $\mathbb{C}^m$ and with $T^{m_1}(x_1) = 0, \ldots, T^{m_k}(x_k) = 0$. This is guaranteed by the following lemma, stated at the level of linear transformations.

Lemma 3. If V is a vector space of dimension m and if $T : V \to V$ is a linear map with $T^p = 0$ for some integer $p \geq 1$, then there exist integers $m_1, \ldots, m_k \geq 1$ with $m_1 + \cdots + m_k = m$ and vectors $v_1, \ldots, v_k \in V$ such that $(v_1, T(v_1), \ldots, T^{m_1-1}(v_1), \ldots, v_k, T(v_k), \ldots, T^{m_k-1}(v_k))$ forms a basis for V and that $T^{m_1}(v_1) = 0, \ldots, T^{m_k}(v_k) = 0$.

Proof. We proceed by induction on $n \geq 1$. For $n = 1$, there is nothing to do. Let us now suppose the result true up to $m-1$, $m \geq 2$, and let us prove that it holds for the integer m, too. Note that $U := \operatorname{ran} T$ is a vector space of dimension $n < m$ (otherwise T would be surjective, hence bijective, and $T^p = 0$ would be impossible). Consider the restriction of T to U, i.e., the linear map $\widetilde{T} : x \in U \mapsto Tx \in U$, and note that $\widetilde{T}^p = 0$. Applying the induction hypothesis, there exist integers $n_1, \dots, n_\ell \geq 1$ with $n_1 + \cdots + n_\ell = n$ and vectors $u_1, \dots, u_\ell \in U$ such that $(u_1, \widetilde{T}(u_1), \dots, \widetilde{T}^{n_1-1}(u_1), \dots, u_\ell, \widetilde{T}(u_\ell), \dots, \widetilde{T}^{n_\ell-1}(u_\ell))$ forms a basis for U and that $\widetilde{T}^{n_1}(u_1) = 0, \dots, \widetilde{T}^{n_\ell}(u_\ell) = 0$. Since $u_1, \dots, u_\ell \in U = \operatorname{ran} T$, there exist $v_1, \dots, v_\ell \in V$ such that $u_i = Tv_i$. Note that $T^{n_i+1}(v_i) = 0$, so that $(T^{n_1}(v_1), \dots, T^{n_\ell}(v_\ell))$ is a linearly independent system of ℓ vectors in $\ker T$, which has dimension $m - n$. Complete this system with vectors $v_{\ell+1}, \dots, v_{m-n} \in \ker T$ to form a basis for $\ker T$. Now consider the system $(v_1, T(v_1), \dots, T^{n_1}(v_1), \dots, v_\ell, T(v_\ell), \dots, T^{n_\ell}(v_\ell), v_{\ell+1}, \cdots, v_{m-n})$. Observe that this is a linearly independent system. Observe also that the number of vectors in this system is $n_1 + 1 + \cdots + n_\ell + 1 + m - n - \ell = n_1 + \cdots + n_\ell + m - n = m$, so that the system is in fact a basis for V. Finally, observe that $T^{n_1+1}(v_1) = 0, \dots, T^{n_\ell+1}(v_\ell) = 0, T(v_{\ell+1}) = 0, \dots, T(v_{n-m}) = 0$ to conclude that the induction hypothesis is true for the integer m. This finishes the proof. □

3 Minimal polynomial

Let $A \in \mathcal{M}_n$ be given. Consider the set of monic (i.e., having a leading coefficient equal to 1) polynomials p that annihilates A (i.e., such that $p(A) = 0$). According Cayley–Hamilton's theorem, this set contains at least one polynomial, namely the characteristic polynomial p_A (possibly multiplied by $(-1)^n$, depending on the convention). In this set, we can consider a polynomial of minimal degree. It turns out that there is only one such polynomial.

Theorem 4. Given $A \in \mathcal{M}_n$, there exists a unique monic polynomial of minimal degree that annihilates A. This polynomial, called the minimal polynomial of A and denoted q_A, divides all polynomials that annihilate A.

Proof. Let m be a monic polynomial of minimal degree that annihilates A, and let p be an arbitrary polynomial that annihilates A. The Euclidean division of p by m reads

$$p(x) = q(x)m(x) + r(x), \qquad \deg(r) < \deg(m).$$

Note that $p(A) = q(A)m(A) + r(A) = 0$ and $m(A) = 0$ imply $r(A) = 0$. But since $\deg(r) < \deg(m)$, we must have $r = 0$. This means that m divides p.
Now let $\widetilde{m}$ be another monic polynomial of minimal degree that annihilates A. By the previous argument, m divides $\widetilde{m}$ and $\widetilde{m}$ divides m, so that m and $\widetilde{m}$ are constant multiples of each other. Since they are both monic, we deduce that $m = \widetilde{m}$. This means that a monic polynomial of minimal degree that annihilates A is unique. □

Observation: Since the minimal polynomial q_A divides the characteristic polynomial p_A, if $p_A(x) = (x-\mu_1)^{n_1} \cdots (x-\mu_k)^{n_k}$ with $n_1, \ldots, n_k \geq 1$, then $q_A(x) = (x-\mu_1)^{m_1} \cdots (x-\mu_k)^{m_k}$ with $1 \leq m_1 \leq n_1, \ldots, 1 \leq m_k \leq n_k$. (Note that $m_i \geq 1$ holds because, if $x_i \neq 0$ is an eigenvector corresponding to μ_i, then $0 = q_A(A)(x_i) = q_A(\mu_i)x_i$, hence $q_A(\mu_i) = 0$).

Observation: One can read the minimal polynomial out of the Jordan canonical form (since equivalent matrices have the same minimal polynomial): the m_i are the order of the largest Jordan block of A corresponding to the eigenvalue μ_i.

Theorem 5. For $a_0, a_1, \ldots, a_{n-1}$, the matrix

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \ddots & \vdots & -a_1 \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{bmatrix}$$

has characteristic polynomial and minimal polynomial given by

$$p_A(x) = q_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

The matrix A is called the companion matrix of the polynomial $x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$.

Proof. We first determine the characteristic polynomial as

$$p_A(x) = \det(xI - A) = \begin{vmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \ddots & \vdots & a_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & x & a_{n-2} \\ 0 & \cdots & 0 & -1 & x + a_{n-1} \end{vmatrix}$$

Expanding along the last column shows that $p_A(x)$ equals

$$\begin{aligned}(-1)^{n+1}a_0 \cdot (-1)^{n-1} + (-1)^{n+2}a_1 \cdot (-1)^{n-2}x + \cdots + (-1)^{2n-1}a_{n-2} \cdot (-1)x^{n-2} + (-1)^{2n}(a_{n-1}+x) \cdot x^{n-1} \\ = a_0 + a_1 x + \cdots + a_{n-2}x^{n-2} + a_{n-1}x^{n-1} + x^n,\end{aligned}$$

as expected. To verify that $q_A = p_A$, it now suffices to show that q_A cannot have degree $m < n$. Suppose the contrary, i.e., that q_A has the form $q_A(x) = x^m + c_{m-1}x^{m-1} + \ldots + c_1 x + c_0$ with $m < n$. Then

$$0 = q_A(A)(e_1) = (A^m + c_{m-1}A^{m-1} + \ldots + c_1 A + c_0 I)(e_1) = e_{m+1} + c_{m-1}e_m + \cdots + c_1 e_2 + c_0 e_1,$$

which contradicts the linear independence of the basis vectors $e_{m+1}, e_m, \ldots . e_2, e_1$. □

4 Exercises

Ex.1: Prove that if $A \in \mathcal{M}_n$ satisfies $A^p = 0$ for some integer $p \geq 1$, then it satisfies $A^n = 0$.

Ex.2: Exercise 2 p. 129

Ex.3: Exercise 2 p. 139

Ex.4: Exercise 6 p. 140

Ex.5: Exercise 7 p. 140

Ex.6: Exercise 8 p. 140

Ex.7: Exercise 17 p. 141

Ex.8: Exercise 9 p. 149

Ex.9: Exercise 13 p. 150

Lecture 5: Eigenvalues of Hermitians Matrices

This lecture takes a closer look at Hermitian matrices and at their eigenvalues. After a few generalities about Hermitian matrices, we prove a minimax and maximin characterization of their eigenvalues, known as Courant–Fischer theorem. We then derive some consequences of this characterization, such as Weyl theorem for the sum of two Hermitian matrices, an interlacing theorem for the sum of two Hermitian matrices, and an interlacing theorem for principal submatrices of Hermitian matrices.

1 Basic properties of Hermitian matrices

We recall that a matrix $A \in \mathcal{M}_n$ is called Hermitian if $A^* = A$ and skew-Hermitian if $A^* = -A$, and we note that A is Hermitian if and only if iA is skew-Hermitian. We have observed earlier that the diagonal entries of a Hermitian matrix are real. This can also be viewed as a particular case of the following result.

Proposition 1. Given $A \in \mathcal{M}_n$,

$$[A \text{ is Hermitian}] \iff [\langle Ax, x\rangle = x^*Ax \in \mathbb{R} \text{ for all } x \in \mathbb{C}^n].$$

Proof. $\Rightarrow$ If A is Hermitian, then, for any $x \in \mathbb{C}^N$,

$$\langle Ax, x\rangle = \langle x, A^*x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle},$$

so that $\langle Ax, x\rangle \in \mathbb{R}$.
$\Leftarrow$ Suppose that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. For any $u, v \in \mathbb{C}^N$, we have

$$\underbrace{\langle A(u+v), u+v\rangle}_{\in\mathbb{R}} = \underbrace{\langle Au, u\rangle}_{\in\mathbb{R}} + \underbrace{\langle Av, v\rangle}_{\in\mathbb{R}} + \langle Au, v\rangle + \langle Av, u\rangle, \qquad \text{so that } \langle Au, v\rangle + \langle Av, u\rangle \in \mathbb{R}.$$

Taking $u = e_j$ and $v = e_k$ yields

$$a_{k,j} \ + \ a_{j,k} \in \mathbb{R}, \qquad \text{thus } \mathrm{Im}(a_{k,j}) = -\mathrm{Im}(a_{j,k}),$$

then taking $u = ie_j$ and $v = e_k$ yields

$$ia_{k,j} - ia_{j,k} \in \mathbb{R}, \qquad \text{thus } \mathrm{Re}(a_{k,j}) = \mathrm{Re}(a_{j,k}).$$

Altogether, this gives $a_{k,j} = \overline{a_{j,k}}$ for all $j, k \in [1:n]$, i.e., $A = A^*$. □

Proposition 2. Any matrix $A \in \mathcal{M}_n$ can be uniquely written in the form

$$A = H + S, \qquad \text{where } H \in \mathcal{M}_n \text{ is Hermitian and } S \in \mathcal{M}_n \text{ is skew-Hermitian},$$
$$A = H_1 + iH_2, \qquad \text{where } H_1, H_2 \in \mathcal{M}_n \text{ are both Hermitian.}$$

Proof. If $A = H + S$ with H Hermitian and S skew-Hermitian, then $A^* = H^* + S^* = H + S$. By adding and subtracting these two relations, we derive $H = (A + A^*)/2$ and $S = (A - A^*)/2$, hence H and S are uniquely determined. Moreover, with H and S given above, it is readily verified that H is Hermitian, that S is skew-Hermitian, and that $A = H + S$. For the second statement, we use the fact that H_2 is Hermitian if and only if $S := iH_2$ is skew-Hermitian. □

2 Variational characterizations of eigenvalues

We now recall that, according to the spectral theorem, if $A \in \mathcal{M}_n$ is Hermitian, there exists a unitary matrix $U \in \mathcal{M}_n$ and a real diagonal matrix D such that $A = UDU^*$. The diagonal entries of D are the eigenvalues of A, which we sort as

$$\lambda_1^{\uparrow}(A) \le \lambda_2^{\uparrow}(A) \le \cdots \le \lambda_n^{\uparrow}(A).$$

We utilize this notation for the rest of the lecture, although we may sometimes just write $\lambda_j^{\uparrow}$ instead of $\lambda_j^{\uparrow}(A)$ when the context is clear. Note also that the columns $u_1, \ldots, u_n$ of U form an orthonormal system and that $Au_j = \lambda_j^{\uparrow}(A)u_j$ for all $j \in [1:n]$. We start by observing that

$$\lambda_1^{\uparrow}(A)\|x\|_2^2 \le \langle Ax, x\rangle \le \lambda_n^{\uparrow}(A)\|x\|_2^2, \tag{1}$$

with the leftmost inequality becoming an equality if $x = u_1$ and the rightmost inequality becoming an equality if $x = u_n$. The argument underlying the observation (1) will reappear several times (sometimes without explanation), so we spell it out here. It is based on the expansion of $x \in \mathbb{C}^n$ on the orthonormal basis $(u_1, \ldots, u_n)$, i.e.,

$$x = \sum_{j=1}^n c_j u_j \qquad \text{with} \qquad \sum_{j=1}^n c_j^2 = \|x\|_2^2.$$

We now simply write

$$\langle Ax, x\rangle = \langle \sum_{j=1}^n c_j Au_j, \sum_{j=1}^n c_j u_j\rangle = \langle \sum_{j=1}^n c_j \lambda_j^{\uparrow} u_j, \sum_{j=1}^n c_j u_j\rangle = \sum_{j=1}^n \lambda_j^{\uparrow} c_j^2 = \begin{cases} \ge \lambda_1^{\uparrow} \sum_{j=1}^n c_j^2 = \lambda_1^{\uparrow}\|x\|_2^2, \\ \\ \le \lambda_n^{\uparrow} \sum_{j=1}^n c_j^2 = \lambda_n^{\uparrow}\|x\|_2^2. \end{cases} \tag{2}$$

The inequalities (1) (with the cases of equality) can also be expressed as $\lambda_1^{\uparrow} = \min_{\|x\|_2=1}\langle Ax, x\rangle$ and $\lambda_n^{\uparrow} = \max_{\|x\|_2=1}\langle Ax, x\rangle$, which is known as Rayleigh–Ritz theorem. It is a particular case of Courant–Fischer theorem stated below.

Theorem 3. For $A \in \mathcal{M}_n$ and $k \in [1:n]$,

$$\lambda_k^{\uparrow}(A) = \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle = \max_{\dim(V)=n-k+1} \min_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle. \tag{3}$$

Remark. This can also be stated with $\dim(V) \ge k$ and $\dim(V) \ge n-k+1$, respectively, or (following the textbook) as

$$\lambda_k^{\uparrow}(A) = \min_{w_1,\ldots,w_{n-k} \in \mathbb{C}^n} \max_{\substack{x \perp w_1,\ldots,w_{n-k} \\ \|x\|_2=1}} \langle Ax, x\rangle = \max_{w_1,\ldots,w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_k \\ \|x\|_2=1}} \langle Ax, x\rangle.$$

Proof of Theorem 3. We only prove the first equality — the second is left as an exercise. To begin with, we notice that, with $U := \operatorname{span}[u_1, \dots, u_k]$, we have

$$\min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x \rangle \leq \max_{\substack{x \in U \\ \|x\|_2=1}} \langle Ax, x \rangle \leq \lambda_k^{\uparrow}(A),$$

where the last inequality follows from an argument similar to (2). For the inequality in the other direction, we remark that our objective is to show that for any k-dimensional linear subspace V of $\mathbb{C}^n$, there is $x \in V$ with $\|x\|_2 = 1$ and $\langle Ax, x \rangle \geq \lambda_k^{\uparrow}(A)$. Considering the subspace $W := \operatorname{span}[u_k, \dots, u_n]$ of dimension $n - k + 1$, we have

$$\dim(V \cap W) = \dim(V) + \dim(W) - \dim(V \cup W) \geq k + n - k + 1 - n = 1.$$

Hence we may pick $x \in V \cup W$ with $\|x\|_2 = 1$. The inequality $\langle Ax, x \rangle \geq \lambda_k^{\uparrow}(A)$ follows from an argument similar to (2). □

We continue with some applications of Courant–Fischer theorem, starting with Weyl theorem.

Theorem 4. Let $A, B \in \mathcal{M}_n$ be Hermitian matrices. For $k \in [1 : n]$,

$$\lambda_k^{\uparrow}(A) + \lambda_1^{\uparrow}(B) \leq \lambda_k^{\uparrow}(A + B) \leq \lambda_k^{\uparrow}(A) + \lambda_n^{\uparrow}(B).$$

Proof. We use Courant–Fischer theorem and inequality (1) to write

$$\begin{aligned} \lambda_k^{\uparrow}(A + B) &= \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \Big(\langle (A + B)x, x \rangle \Big) = \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \Big(\langle Ax, x \rangle + \underbrace{\langle Bx, x \rangle}_{\leq \lambda_n^{\uparrow}(B)} \Big) \\ &\leq \Big(\min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x \rangle \Big) + \lambda_n^{\uparrow}(B) = \lambda_k^{\uparrow}(A) + \lambda_n^{\uparrow}(B). \end{aligned}$$

This establishes the rightmost inequality. We actually use this result to prove the leftmost inequality by replacing A with $A + B$ and B with $-B$, namely

$$\lambda_k^{\uparrow}(A) = \lambda_k^{\uparrow}(A + B + (-B)) \leq \lambda_k^{\uparrow}(A + B) + \lambda_n^{\uparrow}(-B) = \lambda_k^{\uparrow}(A + B) - \lambda_1^{\uparrow}(B).$$

A rearrangement gives the desired result. □

Corollary 5. For Hermitian matrices $A, B \in \mathcal{M}_n$, if all the eigenvalues of B are nonnegative (i.e., $\langle Bx, x \rangle \geq 0$ for all $x \in \mathbb{C}^n$, or in other words B is positive semidefinite), then,

$$\lambda_k^{\uparrow}(A) \leq \lambda_k^{\uparrow}(A + B) \qquad \text{for all } k \in [1 : n].$$

Weyl theorem turns out to be the particular case $k = 1$ of Lidskii's theorem stated below with the sequences of eigenvalues arranged in nonincreasing order rather than nondecreasing order.

Theorem 6. For Hermitian matrices $A, B \in \mathcal{M}_n$, if $1 \le j_1 < \ldots < j_k \le n$, then

$$\sum_{\ell=1}^{k} \lambda_{j_\ell}^{\downarrow}(A+B) \le \sum_{\ell=1}^{k} \lambda_{j_\ell}^{\downarrow}(A) + \sum_{\ell=1}^{k} \lambda_{\ell}^{\downarrow}(B).$$

Proof. Replacing B by $B - \lambda_{k+1}^{\downarrow}(B)I$, we may assume that $\lambda_{k+1}^{\downarrow}(B) = 0$. By the spectral theorem, there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$B = U \mathrm{diag}\big[\underbrace{\lambda_1^{\downarrow}(B), \ldots, \lambda_k^{\downarrow}(B)}_{\ge 0}, \underbrace{\lambda_{k+1}^{\downarrow}(B), \ldots, \lambda_n^{\downarrow}(B)}_{\le 0}\big] U^*.$$

Let us introduce the Hermitian matrices

$$B^+ := U \mathrm{diag}\big[\lambda_1^{\downarrow}(B), \ldots, \lambda_k^{\downarrow}(B), 0, \ldots, 0\big] U^*, \quad B^- := U \mathrm{diag}\big[0, \ldots, 0, -\lambda_{k+1}^{\downarrow}(B), \ldots, -\lambda_n^{\downarrow}(B)\big] U^*.$$

They have only nonnegative eigenvalues and satisfy $B = B^+ - B^-$. According to Corollary 5,

$$\lambda_j^{\downarrow}(A + B^+) \ge \lambda_j^{\downarrow}(A) \qquad \text{and} \qquad \lambda_j^{\downarrow}(A + B^+) = \lambda_j^{\downarrow}(A + B + B^-) \ge \lambda_j^{\downarrow}(A + B).$$

It follows that

$$\begin{aligned}\sum_{\ell=1}^{k} \big(\lambda_{j_\ell}^{\downarrow}(A+B) - \lambda_{j_\ell}^{\downarrow}(A)\big) &\le \sum_{\ell=1}^{k} \big(\lambda_{j_\ell}^{\downarrow}(A+B^+) - \lambda_{j_\ell}^{\downarrow}(A)\big) \le \sum_{j=1}^{n} \big(\lambda_{j}^{\downarrow}(A+B^+) - \lambda_{j}^{\downarrow}(A)\big) \\ &= \mathrm{tr}\,(A+B^+) - \mathrm{tr}\,(A) = \mathrm{tr}\,(B^+) = \sum_{\ell=1}^{k} \lambda_\ell^{\downarrow}(B).\end{aligned}$$

This is just a rearrangement of the desired result. □

3 Interlacing theorems

The two results of this section locate the eigenvalues of a matrix derived from a matrix A relatively to the eigenvalues of A. They are both consequences of Courant–Fischer theorem.

Theorem 7. Let $A \in \mathcal{M}_n$ be a Hermitian matrix and A_s be an $s \times s$ principal submatrix of A, $s \in [1:n]$. Then, for $k \in [1:s]$,

$$\lambda_k^{\uparrow}(A) \le \lambda_k^{\uparrow}(A_s) \le \lambda_{k+n-s}^{\uparrow}(A).$$

Remark. The terminology of interlacing property is particularly suitable in the case of an $(n-1) \times (n-1)$ principal submatrix $\widetilde{A}$, since we then have

$$\lambda_1^{\uparrow}(A) \le \lambda_1^{\uparrow}(\widetilde{A}) \le \lambda_2^{\uparrow}(A) \le \lambda_2^{\uparrow}(\widetilde{A}) \le \lambda_3^{\uparrow}(A) \le \cdots \le \lambda_{n-1}^{\uparrow}(A) \le \lambda_{n-1}^{\uparrow}(\widetilde{A}) \le \lambda_n^{\uparrow}(A).$$

As for the particular case $s = 1$, it gives

$$\lambda_1^{\uparrow}(A) \le a_{j,j} \le \lambda_n^{\uparrow}(A) \qquad \text{for all } j \in [1:n],$$

which is also a consequence of (1) with $x = e_j$.

Proof of Theorem 7. Suppose that the rows and columns of A kept in A_s are indexed by a set S of size s. For a vector $x \in \mathbb{C}^s$, we denote by $\widetilde{x} \in \mathbb{C}^n$ the vector whose entries on S equal those of x and whose entries outside S equal zero. For a linear subspace V of $\mathbb{C}^s$, we define $\widetilde{V} := \{\widetilde{x}, x \in V\}$, which is a subspace of $\mathbb{C}^n$ with dimension equal the dimension of V. Given $k \in [1:s]$, Courant–Fischer theorem implies that, for all linear subspace V of $\mathbb{C}^s$ with $\dim(V) = k$,

$$\max_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A_s x, x \rangle = \max_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x} \rangle = \max_{\substack{\widetilde{x} \in \widetilde{V} \\ \|\widetilde{x}\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x} \rangle \geq \lambda_k^{\uparrow}(A).$$

Taking the minimum over all k-dimensional subspaces V gives $\lambda_k^{\uparrow}(A_s) \geq \lambda_k^{\uparrow}(A)$. Similarly, for all linear subspace V of $\mathbb{C}^s$ with $\dim(V) = s - k + 1 = n - (k + n - s) + 1$,

$$\min_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A_s x, x \rangle = \min_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x} \rangle = \min_{\substack{\widetilde{x} \in \widetilde{V} \\ \|\widetilde{x}\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x} \rangle \leq \lambda_{k+n-s}^{\uparrow}(A).$$

Taking the maximum over all $(s{-}k{+}1)$-dimensional subspaces V gives $\lambda_k^{\uparrow}(A_s) \leq \lambda_{k+n-s}^{\uparrow}(A)$. □

Theorem 8. Let $A, B \in \mathcal{M}_n$ be Hermitian matrices with $\operatorname{rank}(B) = r$. Then, for $k \in [1 : n{-}2r]$,

$$\lambda_k^{\uparrow}(A) \leq \lambda_{k+r}^{\uparrow}(A + B) \leq \lambda_{k+2r}^{\uparrow}(A).$$

Before turning to the proof, we observe that the $n \times n$ Hermitian matrices of rank r are exactly the matrices of the form

$$B = \sum_{j=1}^{r} \mu_j v_j v_j^*, \qquad \mu_1, \ldots, \mu_r \in \mathbb{R} \setminus \{0\}, \quad (v_1, \ldots, v_r) \text{ orthonormal system.} \tag{4}$$

Indeed, by the spectral theorem, the $n \times n$ Hermitian matrices of rank r are exactly the matrices of the form

$$B = V \operatorname{diag}\big[\mu_1, \ldots, \mu_r, 0, \ldots, 0\big] V^*, \qquad \mu_1, \ldots, \mu_r \in \mathbb{R} \setminus \{0\}, \quad V^* V = I, \tag{5}$$

which is just another way of writing (4).

Proof of Theorem 8. Let $k \in [1 : n - r]$ and let $B \in \mathcal{M}_n$ be as in (4). We use Courant–Fischer theorem to derive

$$\begin{aligned}
\lambda_k^{\uparrow}(A) &= \max_{w_1, \ldots, w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1, \ldots, w_k \\ \|x\|_2 = 1}} \langle Ax, x \rangle \leq \max_{w_1, \ldots, w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1, \ldots, w_k, v_1, \ldots, v_r \\ \|x\|_2 = 1}} \langle Ax, x \rangle \\
&\leq \max_{w_1, \ldots, w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1, \ldots, w_k, v_1, \ldots, v_r \\ \|x\|_2 = 1}} \langle (A + B)x, x \rangle \leq \max_{w_1, \ldots, w_{k+r} \in \mathbb{C}^n} \min_{\substack{x \perp w_1, \ldots, w_{k+r} \\ \|x\|_2 = 1}} \langle (A + B)x, x \rangle \\
&= \lambda_{k+r}^{\uparrow}(A + B).
\end{aligned}$$

This establishes the rightmost inequality. This inequality can also be used to establish the leftmost inequality. Indeed, for $k \in [1 : k - 2r]$, we have $k + r \in [1 : k - r]$, and it follows that $\lambda_{k+r}^{\uparrow}(A + B) \leq \lambda_{k+r+r}^{\uparrow}(A + B + (-B)) = \lambda_{k+2r}^{\uparrow}(A)$. □

4 Exercises

Ex.1: Prove the second inequality in (3).

Ex.2: Verify in details the assertions made in (4) and (5).

Ex.3: Exercise 1 p. 174

Ex.4: Exercise 3 p. 174

Ex.5: Exercise 6 p. 174

Ex.6: Exercise 11 p. 175

Ex.7: Exercise 13 p. 175

Ex.8: Exercise 4 p. 181

Ex.9: Exercise 2 p. 198

Ex.10: Exercise 5 p. 199

Ex.11: Exercise 6 p. 199

Ex.12: Exercise 7 p. 199

Ex.13: Exercise 14 p. 200

Ex.14: Exercise 17 p. 200

Lecture 6: Matrix Norms and Spectral Radii

After a reminder on norms and inner products, this lecture introduces the notions of matrix norm and induced matrix norm. Then the relation between matrix norms and spectral radii is studied, culminating with Gelfand's formula for the spectral radius.

1 Inner products and vector norms

Definition 1. Let V be a vector space over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

(IP_1) $\langle x, x \rangle \geq 0$ for all $x \in V$, with equality iff $x = 0$, [positivity]

(IP_2) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$ for all $\lambda, \mu \in \mathbb{K}$ and $x, y, z \in V$, [linearity]

(IP_3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in V$. [Hermitian symmetry]

Observation: If $\mathbb{K} = \mathbb{R}$, then (IP_3) simply says that $\langle x, y \rangle = \langle y, x \rangle$. This is not the case in the complex setting, where one has to be careful about complex conjugation. For instance, (IP_2) and (IP_3) combine to give, for all $\lambda, \mu \in \mathbb{C}$ and $x, y, z \in V$,

$$\langle x, \lambda y + \mu z \rangle = \overline{\lambda} \langle x, y \rangle + \overline{\mu} \langle x, z \rangle.$$

Observation: On $V = \mathbb{C}^n$, there is the classical inner product defined by

$$\langle x, y \rangle := y^* x = \sum_{j=1}^{n} x_j \overline{y_j}, \qquad x, y \in \mathbb{C}^n. \tag{1}$$

On $V = \mathcal{M}_n$, there is the Frobenius inner product defined by

$$\langle A, B \rangle_F := \operatorname{tr}(B^* A) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{k,\ell} \overline{b_{k,\ell}}, \qquad A, B \in \mathcal{M}_n.$$

Cauchy–Schwarz inequality is a fundamental inequality valid in any inner product space. At this point, we state it in the following form in order to prove that any inner product generates a normed space.

Theorem 2. If $\langle \cdot, \cdot \rangle$ is an inner product on a vector space V, then, for all $x, y \in V$,

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Proof. For $x, y \in V$, choose $\theta \in (-\pi, \pi]$ such that $\mathrm{Re}\langle e^{i\theta} x, y \rangle = |\langle x, y \rangle|$. Consider the function defined for $t \in \mathbb{R}$ by

$$q(t) := \langle e^{i\theta} x + ty, e^{i\theta} x + ty \rangle = \langle e^{i\theta} x, e^{i\theta} x \rangle + 2t \mathrm{Re}\langle e^{i\theta} x, y \rangle + t^2 \langle y, y \rangle = \langle x, x \rangle + 2t |\langle x, y \rangle| + t^2 \langle y, y \rangle.$$

This is a quadratic polynomial with $q(t) \geq 0$ for all $t \in \mathbb{R}$, so its discriminant satisfies

$$\Delta = (2|\langle x, y\rangle|)^2 - 4\langle x, x\rangle\langle y, y\rangle \leq 0,$$

which directly translates into $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$, as desired. □

The general definition of a norm is given below. A normed space is simply a vector space endowed with a norm.

Definition 3. Let V be a vector space over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). A function $\|\cdot\| : V \to \mathbb{R}$ is called a (vector) norm if

(N_1) $\|x\| \geq 0$ for all $x \in V$, with equality iff $x = 0$, [positivity]

(N_2) $\|\lambda x\| = |\lambda|\|x\|$ for all $\lambda \in \mathbb{K}$ and $x \in V$, [homogeneity]

(N_3) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$. [triangle inequality]

As examples, we observe that the expression

$$\|x\|_\infty := \max_{j \in [1:n]} |x_j|$$

defines a norm on $\mathbb{K}^n$. The corresponding vector space is denoted as ℓ_∞^n. The expression

$$\|x\|_1 := |x_1| + |x_2| + \cdots + |x_n|$$

defines a norm on $\mathbb{K}^n$. The corresponding vector space is denoted as ℓ_1^n. More generally, for $p \geq 1$, the expression

$$\|x\|_p := \big[|x_1|^p + |x_2|^p + \cdots + |x_n|^p\big]^{1/p}$$

defines a norm on $\mathbb{K}^n$. The corresponding vector space is denoted as ℓ_p^n. In the case $p = 2$, note that ℓ_2^n is the vector space $\mathbb{K}^n$ endowed with the inner product (1).

Proposition 4. If V is a vector space endowed with an inner product $\langle\cdot,\cdot\rangle$, then the expression $\|x\| := \sqrt{\langle x, x\rangle}$ defines a norm on V.

Proof. The properties (N_1) and (N_2) are readily checked. As for (N_3), consider $x, y \in V$, and use Theorem 2 to obtain

$$\begin{aligned}\langle x + y, x + y\rangle &= \langle x, x\rangle + 2\mathrm{Re}\langle x, y\rangle + \langle y, y\rangle \leq \langle x, x\rangle + 2|\langle x, y\rangle| + \langle y, y\rangle \\ &\leq \langle x, x\rangle + 2\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle} + \langle y, y\rangle = \big(\sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}\big)^2,\end{aligned}$$

so that $\sqrt{\langle x + y, x + y\rangle} \leq \sqrt{\langle x, x\rangle} + \sqrt{\langle y, y\rangle}$, i.e., $\|x + y\| \leq \|x\| + \|y\|$, as expected. □

Now that we know that $\|x\| = \sqrt{\langle x, x\rangle}$ defines a norm on an inner product space, we can state Cauchy–Schwarz inequality in the more familiar form

$$|\langle x, y\rangle| \leq \|x\|\|y\|, \qquad x, y \in V.$$

If V is a finite-dimensional space, then all norms on V are equivalent in the following sense (in fact, this characterizes finite dimension).

Theorem 5. If $\|\cdot\|$ and $\|\cdot\|'$ are two norms on a finite-dimensional vector space V, then there exist constants $c, C > 0$ such that

$$c\|x\| \leq \|x\|' \leq C\|x\| \qquad \text{for all } x \in V.$$

Proof. Fixing a basis $(v_1, \ldots, v_n)$ of V, we are going to show that any arbitrary norm $\|\cdot\|'$ is equivalent to the norm $\|\cdot\|$ defined by

$$\|x\| = \max_{j\in[1:n]} |x_j|, \qquad \text{where} \quad x = \sum_{j=1}^{n} x_j v_j.$$

On the one hand, for any $x \in V$, we have

$$\|x\|' = \Big\| \sum_{j=1}^{n} x_j v_j \Big\|' \leq \sum_{j=1}^{n} \|x_j v_j\|' = \sum_{j=1}^{n} |x_j| \|v_j\|' \leq \max_{j\in[1:n]} |x_j| \ \sum_{j=1}^{n} \|v_j\|' = C\|x\|,$$

where we have set $C := \sum_{j=1}^{n} \|v_j\|'$. On the other hand, let us assume that there is no $c > 0$ such that $\|x\| \leq (1/c)\|x\|'$ for all $x \in V$, so that, for each integer $k \geq 1$, we can find $x^{(k)} \in V$ with $\|x^{(k)}\| > k\|x^{(k)}\|'$. Since we can assume without loss of generality that $\|x^{(k)}\| = 1$, we have $\|x^{(k)}\|' < 1/k$. The sequence $\big(x_1^{(k)}\big)_{k\geq 1}$ of complex numbers is bounded, so we can extract a subsequence $\big(x_1^{(\varphi_1(k))}\big)_{k\geq 1}$ converging to some $x_1 \in \mathbb{C}$; next, the sequence $\big(x_2^{(\varphi_1(k))}\big)_{k\geq 1}$ of complex numbers is bounded, so we can extract a subsequence $\big(x_2^{(\varphi_1(\varphi_2(k)))}\big)_{k\geq 1}$ converging to some $x_2 \in \mathbb{C}$; etc. Setting $\varphi = \varphi_1 \circ \cdots \circ \varphi_n$, we end up with subsequences $\big(x_1^{(\varphi(k))}\big)_{k\geq 1}, \ldots, \big(x_n^{(\varphi(k))}\big)_{k\geq 1}$ such that $x_j^{(\varphi(k))} \underset{k\to\infty}{\longrightarrow} x_j$ for each $j \in [1:n]$. Note that the vectors $x^{(\varphi(k))} := \sum_{j=1}^{n} x_j^{(\varphi(k))} v_j$ and $x = \sum_{j=1}^{n} x_j v_j$ satisfy $\|x - x^{(\varphi(k))}\| = \max_{j\in[1:n]} |x_j - x_j^{(\varphi(k))}| \underset{k\to\infty}{\longrightarrow} 0$, therefore

$$\|x\|' \leq \|x - x^{(\varphi(k))}\|' + \|x^{(\varphi(k))}\|' \leq C\|x - x^{(\varphi(k))}\| + 1/\varphi(k).$$

Taking the limit as $k \to \infty$ yields $\|x\|' = 0$, hence $x = 0$, which contradicts

$$\|x\| = \max_{j\in[1:n]} |x_j| = \max_{j\in[1:n]} \lim_{k\to\infty} |x_j^{(\varphi(k))}| = \lim_{k\to\infty} \max_{j\in[1:n]} |x_j^{(\varphi(k))}| = \lim_{k\to\infty} \|x^{(\varphi(k))}\| = \lim_{k\to\infty} 1 = 1.$$

This proves the existence of the desired constant $c > 0$. □

For instance, we can use Cauchy–Schwarz inequality to derive, for any $x \in \mathbb{C}^n$,

$$\|x\|_1 = \sum_{j=1}^n |x_j| = \sum_{j=1}^n 1 \times |x_j| \le \Big[\sum_{j=1}^n 1^2\Big]^{1/2} \Big[\sum_{j=1}^n |x_j|^2\Big]^{1/2} = \sqrt{n}\|x\|_2,$$

and this inequality is best possible because it turns into an equality for $x = [1, 1, \dots, 1]^\top$. On the other hand, for any $x \in \mathbb{C}^n$, we have $\|x\|_2 \le \|x\|_1$, since

$$\|x\|_2^2 = \sum_{j=1}^n |x_j|^2 \le \Big[\sum_{j=1}^n |x_j|\Big]^2 = \|x\|_1^2,$$

and this inequality is best possible because it turns into an equality for $x = [1, 0, \dots, 0]^\top$. We can more generally compare any ℓ_p-norm with any ℓ_q-norm. The proof is left as an exercise.

Proposition 6. Given $1 \le p < q \le \infty$, for all $x \in \mathbb{K}^n$,

$$\|x\|_q \le \|x\|_p \le n^{1/p-1/q}\|x\|_q,$$

and these inequalities are best possible.

2 Matrix norms

Since $\mathcal{M}_n$ is a vector space, it can be endowed with a *vector* norm. There is one more ingredient making this norm a *matrix* norm.

Definition 7. A function $|||\cdot||| : \mathcal{M}_n \to \mathbb{C}$ is called a matrix norm if

(MN_1) $|||A||| \ge 0$ for all $A \in \mathcal{M}_n$, with equality iff $x = 0$, [positivity]

(MN_2) $|||\lambda A||| = |\lambda|\,|||A|||$ for all $\lambda \in \mathbb{C}$ and $A \in \mathcal{M}_n$, [homogeneity]

(MN_3) $|||A + B||| \le |||A||| + |||B|||$ for all $A, B \in \mathcal{M}_n$. [triangle inequality]

(MN_4) $|||AB||| \le |||A|||\,|||B|||$ for all $A, B \in \mathcal{M}_n$. [submultiplicativity]

As an example, we notice that the Frobenius norm defined by

$$\|A\|_F := \sqrt{\langle A, A\rangle_F} = \sqrt{\mathrm{tr}\,(A^*A)}, \qquad A \in \mathcal{M}_n,$$

is a matrix norm. Indeed, for $A, B \in \mathcal{M}_n$,

$$\|AB\|_F^2 = \mathrm{tr}\,((AB)^*(AB)) = \mathrm{tr}\,(B^*A^*AB) = \mathrm{tr}\,(BB^*A^*A) = \langle A^*A, BB^*\rangle_F \le \|A^*A\|_F\|B^*B\|_F.$$

Now notice that, with $\lambda_1 \ge \cdots \ge \lambda_n \ge 0$ denoting the eigenvalues of the Hermitian matrix $M := A^*A$, we have

$$\|A^*A\|_F = \|M\|_F = \mathrm{tr}\,(M^2) = \sum_{j=1}^n \lambda_j^2 \le \Big[\sum_{j=1}^n \lambda_j\Big]^2 = \mathrm{tr}\,(M)^2 = \|A\|_F^2.$$

Likewise, we have $\|B^*B\|_F \leq \|B\|_F^2$. We deduce $\|AB\|_F^2 \leq \|A\|_F^2\|B\|_F^2$, i.e., $\|AB\|_F \leq \|A\|_F\|B\|_F$, as desired.

Another important example of matrix norms is given by the norm induced by a vector norm.

Definition 8. If $\|\cdot\|$ is a vector norm on $\mathbb{C}^n$, then the induced norm on $\mathcal{M}_n$ defined by

$$|\!|\!| A |\!|\!| := \max_{\|x\|=1} \|Ax\|$$

is a matrix norm on $\mathcal{M}_n$.

A consequence of the definition of the induced norm is that $\|Ax\| \leq |\!|\!| A |\!|\!| \|x\|$ for any $x \in \mathbb{C}^n$. Let us now verify ($\mathbf{MN}_4$) for the induced norm. Given $A, B \in \mathcal{M}_n$, we have, for any $x \in \mathbb{C}^n$ with $\|x\| = 1$,

$$\|ABx\| \leq |\!|\!| A |\!|\!| \|Bx\| \leq |\!|\!| A |\!|\!| \, |\!|\!| B |\!|\!| \|x\| = |\!|\!| A |\!|\!| \, |\!|\!| B |\!|\!|,$$

so taking the maximum over x gives the desired inequality $|\!|\!| AB |\!|\!| \leq |\!|\!| A |\!|\!| \, |\!|\!| B |\!|\!|$.

There are situations where one can give an explicit expression for induced norms, for instance

$$|\!|\!| A |\!|\!|_{\infty\to\infty} := \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{k\in[1:n]} \sum_{\ell=1}^{n} |a_{k,\ell}|, \tag{2}$$

$$|\!|\!| A |\!|\!|_{1\to 1} := \max_{\|x\|_1=1} \|Ax\|_1 \quad = \max_{\ell\in[1:n]} \sum_{k=1}^{n} |a_{k,\ell}|. \tag{3}$$

3 Spectral radius

There is also a (somewhat) explicit expression for the matrix norm induced by the Euclidean norm. It involves the spectral radius of a matrix $M \in \mathcal{M}_n$ defined as

$$\rho(M) := \max\big\{|\lambda|, \lambda \text{ eigenvalue of } M\big\}.$$

Proposition 9. For any $A \in M_n$,

$$|\!|\!| A |\!|\!|_{2\to 2} := \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\rho(A^*A)}.$$

Moreover, if $A \in \mathcal{M}_n$ is Hermitian, then

$$|\!|\!| A |\!|\!|_{2\to 2} = \rho(A).$$

Proof. Note that $\rho(A^*A) = \lambda_1^{\downarrow}$, where $\lambda_1^{\downarrow} \geq \cdots \geq \lambda_n^{\downarrow} \geq 0$ be the eigenvalues of A^*A. Then

$$|\!|\!| A |\!|\!|_{2\to 2}^2 = \max_{\|x\|_2=1} \|Ax\|_2^2 = \max_{\|x\|_2=1} \langle Ax, Ax\rangle = \max_{\|x\|_2=1} \langle A^*Ax, x\rangle = \lambda_1^{\downarrow},$$

where the last equality is a characterization of the largest eigenvalue given in Lecture 5. This implies the first result. For the second result, if A is Hermitian, then $\lambda_1^{\downarrow}, \cdots, \lambda_n^{\downarrow}$ are the eigenvalues of $A^*A = A^2$, that is, $\mu_1^2, \ldots, \mu_n^2$ where $\mu_1, \ldots, \mu_n$ are the eigenvalues of A. In particular, $\sqrt{\lambda_1^{\downarrow}}$ is the largest values among the $|\mu_j|$, i.e., the spectral radius of A. □

We now examine the relation between spectral radius and the other matrix norms. We start with the following observations.

Lemma 10. If $\|\cdot\|$ is a matrix norm on $\mathcal{M}_n$, then, for any $A \in \mathcal{M}_n$,

$$\rho(A) \leq \|A\|.$$

Proof. Let λ be an eigenvalue of A, and let $x \neq 0$ be a corresponding eigenvector. From $Ax = \lambda x$, we have

$$AX = \lambda X, \qquad \text{where } X := \left[\begin{array}{c|c|c|c} x & \cdots & x \end{array}\right] \in \mathcal{M}_n \setminus \{0\}.$$

It follows that

$$|\lambda|\|X\| = \|\lambda X\| = \|AX\| \leq \|A\|\|X\|,$$

and simplifying by $\|X\|(> 0)$ gives $|\lambda| \leq \|A\|$. Taking the maximum over all eigenvalues λ gives the result. □

Lemma 11. Given $A \in \mathcal{M}_n$ and $\varepsilon > 0$, there exists a matrix norm $\|\cdot\|$ such that

$$\|A\| \leq \rho(A) + \varepsilon.$$

Proof. The Jordan canonical form of A is

$$A = S \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_k}(\lambda_k) \end{bmatrix} S^{-1},$$

where $S \in \mathcal{M}_n$ is an invertible matrix, $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of A, and $n_1 + \cdots + n_k = n$. Setting

$$D(\eta) = \begin{bmatrix} D_{n_1}(\eta) & 0 & \cdots & 0 \\ 0 & D_{n_2}(\eta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{n_k}(\eta) \end{bmatrix}, \qquad \text{where} \quad D_m(\eta) = \begin{bmatrix} \eta & 0 & \cdots & 0 \\ 0 & \eta^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \eta^m \end{bmatrix},$$

we calculate (since the multiplication on the left by $D_m(1/\varepsilon)$ multiplies the ith row by $1/\varepsilon^i$ and the multiplication on the right by $D_m(\eta)$ multiplies the jth column by ε^j)

$$D(1/\varepsilon)S^{-1}ASD(\varepsilon) = \begin{bmatrix} B_{n_1}(\lambda_1,\varepsilon) & 0 & \cdots & 0 \\ 0 & B_{n_2}(\lambda_2,\varepsilon) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_{n_k}(\lambda_k,\varepsilon) \end{bmatrix},$$

where

$$B_m(\lambda,\varepsilon) = D_m(1/\varepsilon)J_m(\lambda)D_m(\varepsilon) = \begin{bmatrix} \lambda & \varepsilon & 0 & \cdots & 0 \\ 0 & \lambda & \varepsilon & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & \varepsilon \\ 0 & \cdots & 0 & 0 & \lambda \end{bmatrix}.$$

Let us now define a matrix norm (the necessary verifications are left as an exercise) by

$$|\!|\!| M |\!|\!| := |\!|\!| D(1/\varepsilon)S^{-1}MSD(\varepsilon) |\!|\!|_{1\to 1}, \qquad M \in \mathcal{M}_n. \tag{4}$$

According to (3), we conclude that

$$|\!|\!| A |\!|\!| = \max_{\ell\in[1:n]} \big(|\lambda_\ell| + \varepsilon\big) = \rho(A) + \varepsilon. \qquad \square$$

Lemmas 10 and 11 can be combined to give the following expression for the spectral norm of a matrix $A \in \mathcal{M}_n$:

$$\rho(A) = \inf\big\{ |\!|\!| A |\!|\!|, |\!|\!|\cdot|\!|\!| \text{ is a matrix norm on } \mathcal{M}_n \big\}.$$

The spectral radius can also be expressed via Gelfand's formula below.

Theorem 12. Given any matrix norm $|\!|\!|\cdot|\!|\!|$ on $\mathcal{M}_n$,

$$\rho(A) = \lim_{k\to\infty} |\!|\!| A^k |\!|\!|^{1/k}, \qquad A \in \mathcal{M}_n.$$

Proof. Given $k \geq 0$, we use Lemma 10 to write

$$\rho(A)^k = \rho(A^k) \leq |\!|\!| A^k |\!|\!|, \qquad \text{i.e.,} \quad \rho(A) \leq |\!|\!| A^k |\!|\!|^{1/k}.$$

Taking the limit as $k \to \infty$ gives $\rho(A) \leq \lim_{k\to\infty} |\!|\!| A^k |\!|\!|^{1/k}$. To establish the reverse inequality, we need to prove that, for any $\varepsilon > 0$, there is $K \geq 0$ such that $|\!|\!| A^k |\!|\!|^{1/k} \leq \rho(A) + \varepsilon$ for all $k \geq K$. From Lemma 11, we know that there exists a matrix norm $\|\cdot\|$ on $\mathcal{M}_n$ such that $\|A\| \leq \rho(A) + \varepsilon/2$. Moreover, by the equivalence of the norms on $\mathcal{M}_n$, we know that there exists some constant $C > 0$ such that $|\!|\!| M |\!|\!| \leq C\|M\|$ for all $M \in \mathcal{M}_n$. Then, for any $k \geq 0$,

$$|\!|\!| A^k |\!|\!| \leq C\|A^k\| \leq C\|A\|^k \leq C(\rho(A) + \varepsilon/2)^k,$$
$$|\!|\!| A^k |\!|\!|^{1/k} \leq C^{1/k}(\rho(A) + \varepsilon/2) \underset{k\to\infty}{\longrightarrow} \rho(A) + \varepsilon/2.$$

The latter implies the existence of $K \geq 0$ such that $|\!|\!| A^k |\!|\!|^{1/k} \leq \rho(A)+\varepsilon$ for $k \geq K$, as desired. $\square$

4 Exercises

Ex.1: Prove that $\|\cdot\|_p$ indeed defines a norm for $p \geq 1$, and prove Proposition 6. You will need Hölder's inequality (Cauchy–Schwarz inequality): given $u_1, \ldots, u_n \geq 0$, $v_1, \ldots, v_n \geq 0$, and $p, q \in [1, \infty]$ with $1/p + 1/q = 1$,

$$\sum_{j=1}^{n} u_j v_j \leq \Big[\sum_{j=1}^{n} u_j^p\Big]^{1/p} \Big[\sum_{j=1}^{n} v_j^q\Big]^{1/q}.$$

Ex.2: Exercise 1 p. 262

Ex.3: Exercise 6 p. 263

Ex.4: If V is a real inner product space, prove the polarization formula

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right), \qquad x, y \in V.$$

If V is a complex inner product space, prove the polarization formula

$$\langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right), \qquad x, y \in V.$$

Ex.5: Exercise 7 p. 263

Ex.6: Exercise 8 p. 263

Ex.7: Exercises 4 and 10 p. 263

Ex.8: Exercise 1 p. 267

Ex.9: Prove that $\|x\|_p \underset{p\to\infty}{\longrightarrow} \|x\|_\infty$ for any $x \in \mathbb{K}^n$.

Ex.10: Exercise 2 p. 267

Ex.11: Exercise 4 p. 267

Ex.12: Exercise 3 p. 311

Ex.13: Exercise 5 p. 311

Ex.14: Exercise 7 p. 311 (hence verify (4))

Ex.15: Exercises 10 and 16 p. 312-313

Ex.16: Exercise 11 p. 312

Ex.17: Exercise 15 p. 313

Ex.18: Exercise 19 p. 313

Ex.19: Exercise 20 p. 313

Ex.20: Exercise 21 p. 313

Lecture 7: Positive (Semi)Definite Matrices

This short lecture introduces the notions of positive definite and semidefinite matrices. Two characterizations are given and the existence and uniqueness of square roots for positive semidefinite matrices is proved. Gram matrices are also briefly mentioned along the way.

1 Definitions and characterizations

Definition 1. A positive definite (resp. semidefinite) matrix is a Hermitian matrix $A \in \mathcal{M}_n$ satisfying

$$\langle Ax, x\rangle > 0 \ \ (\text{resp. } \geq 0) \qquad \text{for all } x \in \mathbb{C}^n \setminus \{0\}.$$

We write $A \succ 0$ (resp.$A \succeq 0$) to designate a positive definite (resp. semidefinite) matrix A.

Before giving verifiable characterizations of positive definiteness (resp. semidefiniteness), we make a few observations (stated with $\succ$, but also valid for $\succeq$ provided $>$ is replaced by ≥ 0):

1. If $A, B \succ 0$ and if $t > 0$, then $A + B \succ 0$ and $tA \succ 0$.

2. The eigenvalues of a positive definite matrix are > 0. Indeed, if (λ, x) is an eigenpair, then $\lambda\|x\|_2^2 = \langle \lambda x, x\rangle = \langle Ax, x\rangle > 0$. We derive in particular that $\operatorname{tr}(A) > 0$ and $\det(A) > 0$ for $A \succ 0$.

3. The diagonal entries of a positive definite matrix are >0, since $a_{i,i} = \langle Ae_i, e_i\rangle$ for all $i \in [1{:}n]$.

4. A principal submatrix of $A \succ 0$ satisfies $A_S \succ 0$. Indeed, if the rows and columns of A kept in A_S are indexed by a set S, then for $x \in \mathbb{C}^{\operatorname{card}(S)}$, $\langle A_S x, x\rangle = \langle A\widetilde{x}, \widetilde{x}\rangle > 0$, where $\widetilde{x} \in \mathbb{C}^n$ denotes the vector whose entries on S equal those of x and whose entries outside S equal zero.

5. If $A \succ 0$, then $|a_{i,j}|^2 < a_{i,i}a_{j,j}$ for all $i, j \in [1 : n]$. This is a consequence of the fact that $\begin{bmatrix} a_{i,i} & a_{i,j} \\ \overline{a_{i,j}} & a_{j,j} \end{bmatrix}$ is a principal submatrix of A, so it has positive determinant, i.e., $a_{i,i}a_{j,j} - |a_{i,j}|^2 > 0$.

The characterizations of positive definite matrices stated below are also valid for positive semidefinite matrices, provided $>$ is replaced by ≥ 0.

Theorem 2. For an Hermitian matrix $A \in \mathcal{M}_n$,

$$[A \succ 0] \iff [\det(A_1) > 0, \det(A_2) > 0, \ldots, \det(A_n) > 0] \iff [\text{all eigenvalues of } A \text{ are } > 0],$$

where $A_1 := A_{[1:1]}, A_2 := A_{[1:2]}, \ldots, A_{n-1} := A_{[1:n-1]}, A_n := A_{[1:n]}$ are the leading principal submatrices of A.

Proof. The first implication follows from Observation 4.
For the second implication, assuming that the determinants of all leading principal submatrices are positive, we prove by induction on $k \in [1:n]$ that all the eigenvalues of A_k are positive — the desired result being the case $k = n$. For $k = 1$, this is true because $\lambda_1^\uparrow(A_1) = \det(A_1) > 0$. Next, let us suppose the induction hypothesis true up to $k-1$, $k \geq 2$. By the interlacing property, we have

$$\lambda_1^\uparrow(A_k) \leq \lambda_1^\uparrow(A_{k-1}) \leq \lambda_2^\uparrow(A_k) \leq \lambda_2^\uparrow(A_{k-1}) \leq \cdots \leq \lambda_{n-1}^\uparrow(A_k) \leq \lambda_{n-1}^\uparrow(A_{k-1}) \leq \lambda_n^\uparrow(A_k),$$

and by the induction hypothesis, we have $\lambda_{n-1}^\uparrow(A_{k-1}) \geq \cdots \geq \lambda_1^\uparrow(A_{k-1}) > 0$. It follows that $\lambda_n^\uparrow(A_k) \geq \cdots \geq \lambda_2^\uparrow(A_k) > 0$. In turn, we derive

$$\lambda_1^\uparrow(A_k) = \frac{\det(A_k)}{\lambda_n^\uparrow(A_k) \cdots \lambda_2^\uparrow(A_k)} = \frac{> 0}{> 0} > 0.$$

This shows that all the eigenvalues of A_k are positive. The inductive proof is now complete.
For the third implication, we invoke the spectral theorem for Hermitian matrices to write

$$A = U \mathrm{diag}[\lambda_1, \ldots, \lambda_n] U^*, \qquad \text{with } UU^* = I = U^*U.$$

The assumption that $\lambda_j > 0$ for all $j \in [1:n]$ implies, for $x \in \mathbb{C}^n \setminus \{0\}$, that

$$\langle Ax, x \rangle = \langle U \mathrm{diag}[\lambda_1, \ldots, \lambda_n] U^* x, x \rangle = \langle \mathrm{diag}[\lambda_1, \ldots, \lambda_n] U^* x, U^* x \rangle = \sum_{j=1}^n \lambda_j |(U^*x)_j|^2 > 0,$$

where the strict inequality holds because $U^*x \neq 0$. This proves that $A \succ 0$. □

2 Square roots of positive semidefinite matrices

Theorem 3. For a positive semidefinite matrix $A \in \mathcal{M}_n$, there exists a unique positive semidefinite matrix $B \in \mathcal{M}_n$ such that $B^2 = A$.

Proof. The existence follows from the spectral theorem. Indeed, we have

$$A = U \mathrm{diag}[\lambda_1, \ldots, \lambda_n] U^*, \qquad \text{with } UU^* = I = U^*U,$$

and we know that $\lambda_j \geq 0$ for all $j \in [1:n]$ — see Observation 2 or Theorem 2. We then set

$$B := U \mathrm{diag}[\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}] U^*,$$

and it is clear that $B^2 = A$. As for the uniqueness, let us consider a positive semidefinite matrix $C \in \mathcal{M}_n$ such that $C^2 = A$ and let us prove that $C = B$. Let p be a polynomial such that $p(\lambda_1) = \sqrt{\lambda_1}, \ldots, p(\lambda_n) = \sqrt{\lambda_n}$ — if $\mu_1, \ldots, \mu_k$ are the distinct eigenvalues of A, take

$p(x) = \sum_{j=1}^k \sqrt{\lambda_j} \prod_{i \neq j}[(x-\mu_i)/(\mu_j - \mu_i)]$. Note that C and $A = C^2$ commute, hence C and $p(A)$ commute. Observing that

$$p(A) = Up\big(\mathrm{diag}[\lambda_1, \ldots, \lambda_n]\big)U^* = U\mathrm{diag}[p(\lambda_1), \ldots, p(\lambda_n)]U^* = U\mathrm{diag}[\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}]U^* = B,$$

this means that C and B commute. Then, by the spectral theorem for commuting Hermitian matrices, there exists a unitary $V \in \mathcal{M}_n$ such that

$$B = V\mathrm{diag}[\beta_1, \ldots, \beta_n]V^* \quad \text{and} \quad C = V\mathrm{diag}[\gamma_1, \ldots, \gamma_n]V^*.$$

The equality $B^2 = C^2$ gives $\beta_j^2 = \gamma_j^2$ for all $j \in [1:n]$, and since $\beta_j \geq 0$ and $\gamma_j \geq 0$ (because $B \succeq 0$ and $C \succeq 0$), we derive that $\beta_j = \gamma_j$ for all $j \in [1:n]$, i.e., $B = C$. □

The next statement establishes a relation between positive semidefinite matrices and Gram matrices. The Gram matrix G of a system of vectors $(c_1, \ldots, c_n) \in \mathbb{C}^m$ is defined by

$$G_{i,j} = \langle c_j, c_i \rangle.$$

It can alternatively be written as $G = C^*C$, where $C = \begin{bmatrix} c_1 \end{vmatrix} \cdots \begin{vmatrix} c_n \end{bmatrix}$.

Theorem 4. Given $A \in \mathcal{M}_n$ and $m \geq n$,

$$[A \succeq 0] \iff [A = C^*C \text{ for some } C \in \mathcal{M}_{m \times n}].$$

Proof. $\Rightarrow$ It suffices to take $C := \begin{bmatrix} B \\ \hline 0 \end{bmatrix}$, where B is the square root of A from Theorem 3.
$\Leftarrow$ It is clear that C^*C is Hermitian and that, for $x \in \mathbb{C}^n$, $\langle C^*Cx, x \rangle = \langle Cx, Cx \rangle \geq 0$. □

3 Exercises

Ex.1: Exercise 2 p. 400

Ex.2: Exercise 3 p. 400

Ex.3: Exercise 5 p. 400

Ex.4: Exercise 2 p. 408

Ex.5: Exercise 3 p. 408

Ex.6: Exercise 12 p. 409

Ex.7: Exercise 14 p. 410

Ex.8: Verify that the Gram matrix of a system of vectors is invertible (hence positive definite) if and only if the system of vectors is linearly independent.

Lecture 8: Variations on Geršgorin Theorem

In this lecture, we intend to locate the eigenvalues of a matrix without calculating them. Gešgorin theorem, as well as the more elaborate Ostrovsky theorem, informs us that the eigenvalues belong to the union of certain disks in the complex plane. This is used to show the invertibility of some diagonal dominant matrices.

1 Gešgorin theorem

The first piece of information on the set of eigenvalues of a matrix — called the spectrum — is given below.

Theorem 1. For any $A \in \mathcal{M}_n$,

$$\text{sp}(A) \subseteq \bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j=1, j\neq i}^{n} |a_{i,j}| \Big\}. \tag{1}$$

The disks $D(a_{i,i}, R_i) := \big\{ z \in \mathbb{C} : |z - a_{i,i}| \leq R_i \big\}$ centered at $a_{i,i}$ and of radius $R_i := \sum_{j\neq i} |a_{i,j}|$ are called the Geršgorin disks, and their union is called the Geršgorin region.

Proof. Let λ be an eigenvalue of A. We need to prove that there exists $i \in [1:n]$ such that $|\lambda - a_{i,i}| \leq \sum_{j\neq i} |a_{i,j}|$. Considering an eigenvector $x \neq 0$ associated to λ, we have $Ax = \lambda x$, i.e., $\sum_{j=1}^{n} a_{i,j} x_j = \lambda x_i$ for all $i \in [1:n]$. It follows that

$$|\lambda - a_{i,i}||x_i| = \Big| \sum_{j\neq i} a_{i,j} x_j \Big| \leq \sum_{j\neq i} |a_{i,j}||x_j| \leq \sum_{j\neq i} |a_{i,j}| \|x\|_\infty.$$

Choosing $i \in [1:n]$ with $|x_i| = \|x\|_\infty$ and simplifying by $\|x\|_\infty > 0$ yields the conclusion. □

Applying the Theorem 1 to $A^\top$ (not to A^*!), which has the same spectrum as A, gives

$$\text{sp}(A) \subseteq \bigcup_{j=1}^{n} D(a_{j,j}, C_j), \qquad \text{where } C_j := \sum_{i=1, i\neq j}^{n} |a_{i,j}|. \tag{2}$$

As an example, consider the matrix

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 2 & 1 & 2/3 \\ 3 & 3/2 & 1 \end{bmatrix}. \tag{3}$$

From (1) and (2), we obtain

$$\begin{array}{rcccccccc} \text{sp}(A) & \subseteq & D(1,5/6) & \cup & D(1,8/3) & \cup & D(1,9/2) & = & D(1,9/2), \\ \text{sp}(A) & \subseteq & D(1,5) & \cup & D(1,2) & \cup & D(1,4/3) & = & D(1,5). \end{array}$$

Since $A \mapsto SAS^{-1}$ is another operation that preserves the spectrum, taking $S = \mathrm{diag}[d_1, \dots, d_n]$ for some $d_1, \dots, d_n > 0$ in (1) gives

$$\mathrm{sp}(A) \subseteq \bigcup_{i=1}^{n} \Big\{ z \in \mathbb{C} : |z - a_{i,i}| \leq d_i \sum_{j=1, j\neq i}^{n} \frac{1}{d_j} |a_{i,j}| \Big\}.$$

For instance, we notice that the matrix of (3) can be written as

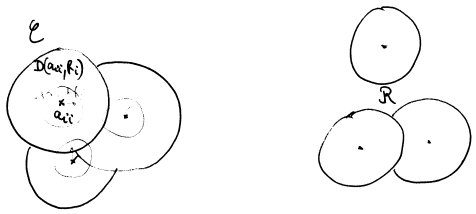
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}. \tag{4}$$

We then deduce from Theorem 1 that

$$\mathrm{sp}(A) \subseteq D(1, 2).$$

Note that (4) also reveals the eigenvalues of A: 0 (of multiplicity 2) and 3 (of multiplicity 1). This example shows that the eigenvalues can lie on the boundary of the Geršgorin region.

If the Geršgorin disks are all distinct, we can say that each one of them contains exactly one eigenvalue. More generally, we establish (not too rigorously) below that, if k Geršgorin disks form a connected region $\mathcal{C}$ disjoint from the region $\mathcal{R}$ formed by the remaining $n - k$ disks, then $\mathcal{C}$ contains exactly k eigenvalues of A.



Indeed, notice that the Geršgorin disks $D(a_{i,i}, R_i)$ corresponding to A contain the Geršgorin disks $D(a_{i,i}, tR_i)$ corresponding to $A_t := (1-t)\mathrm{diag}(A) + tA$ for any $t \in [0, 1]$. Notice also that A_t has k eigenvalues in $\mathcal{C}$ when $t = 0$ (namely the centers of the k disks forming $\mathcal{C}$) and that these eigenvalues are continuous functions of t, since Weyl theorem yields

$$|\lambda_i^{\downarrow}(A_t) - \lambda_i^{\downarrow}(A_{t'})| \leq \rho\big(A_t - A_{t'}\big) = \rho\big((t - t')(A - \mathrm{diag}(A))\big) = |t - t'|\rho\big(A - \mathrm{diag}(A)\big).$$

Thus, when t increases to 1, the k eigenvalues of A_t cannot 'escape' $\mathcal{C}$ and no eigenvalue in $\mathcal{R}$ can 'enter' $\mathcal{C}$ either. Therefore, there are exactly k eigenvalues of $A_1 = A$ in $\mathcal{C}$.

2 Diagonal dominance

A matrix $A \in \mathcal{M}_n$ is called diagonally dominant with respect to the rows if the dominant entry in each row is the diagonal entry, in the sense that

$$|a_{i,i}| \geq \sum_{j \neq i} |a_{i,j}| \qquad \text{for all } i \in [1:n].$$

It is called strictly diagonally dominant with respect to the rows if the previous inequalities are strict, i.e.,

$$|a_{i,i}| > \sum_{j \neq i} |a_{i,j}| \qquad \text{for all } i \in [1:n].$$

Diagonal dominance and strict diagonal dominance with respect to the columns are defined in an obvious way. Using (1) or (2), we easily see that the spectrum of a strictly diagonally dominant matrix does not contain the origin, so that this matrix is invertible. This is not necessarily true for merely diagonally dominant matrices, as shown by the counterexample $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, but it remains with an additional mild condition.

Theorem 2. If $A \in \mathcal{M}_n$ is a diagonally dominant matrix without any zero entry and if there is at least one index i such that $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$, then A is invertible.

Proof. Suppose that A is not invertible, i.e., that there exists $x \in \mathbb{C}^n \setminus \{0\}$ such that $Ax = 0$. This means that $\sum_{j=1}^n a_{i,j} x_j = 0$ for all $i \in [1:n]$. It follows that

$$|a_{i,i}||x_i| = \Big| \sum_{j \neq i} a_{i,j} x_j \Big| \leq \sum_{j \neq i} |a_{i,j}||x_j| \underset{(*)}{\leq} \sum_{j \neq i} |a_{i,j}| \|x\|_\infty \leq |a_{i,i}| \|x\|_\infty.$$

If the index i is chosen so that $|x_i| = \|x\|_\infty$, then all the inequalities above turn into equalities. In particular, equality in $(*)$ yields $|a_{i,j}||x_j| = |a_{i,j}| \|x\|_\infty$ for all $j \in [1:n]$ and since all $|a_{i,j}|$ are nonzero, we deduce that $|x_j| = \|x\|_\infty$ for all $j \in [1:n]$. Next, if the index i is chosen so that $|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|$, we obtain

$$|a_{i,i}| \|x\|_\infty = |a_{i,i}||x_i| = \Big| \sum_{j \neq i} a_{i,j} x_j \Big| \leq \sum_{j \neq i} |a_{i,j}||x_j| \leq \sum_{j \neq i} |a_{i,j}| \|x\|_\infty < |a_{i,i}| \|x\|_\infty,$$

which is impossible. We conclude that A is invertible. □

3 Ostrovsky theorem

Theorem 3. For any $A \in \mathcal{M}_n$ and any $t \in [0,1]$,

$$\text{(5)} \qquad \mathrm{sp}(A) \subseteq \bigcup_{i=1}^n \Big\{ z \in \mathbb{C} : |z - a_{i,i}| \leq R_i^{1-t} C_i^t \Big\},$$

where

$$R_i := \sum_{j \neq i} |a_{i,j}| \qquad \text{and} \qquad C_i := \sum_{j \neq i} |a_{j,i}|.$$

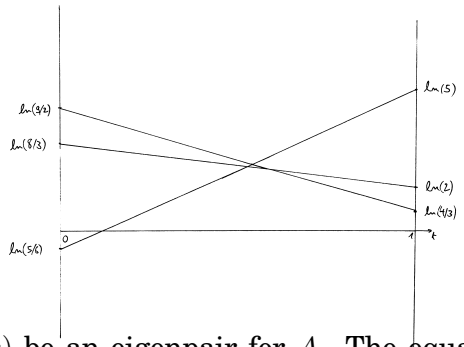
Remark. Choosing $t=0$, we retrieve (1) and choosing $t=1$, we retrieve (2). At first sight, we could nonetheless believe that Theorem 3 is not strictly stronger than (1) and (2) together, because each $R_i^{1-t}C_i^t$ is minimized either at $t=0$ or $t=1$, but this belief is incorrect. Indeed, applying Ostrovsky theorem to the matrix (3) yields

$$\begin{aligned}\mathrm{sp}(A) &\subseteq D\big(1,(5/6)^{1-t}5^t\big)\cup D\big(1,(8/3)^{1-t}2^t\big)\cup D\big(1,(9/2)^{1-t}(4/3)^t\big)\\ &= D\big(1,\max\{(5/6)^{1-t}5^t,(8/3)^{1-t}2^t,(9/2)^{1-t}(4/3)^t\}\big).\end{aligned}$$

This disk is as small as possible after minimizing over $t\in[0,1]$ the logarithm of its radius

$$\max\{(1-t)\ln(5/6)+t\ln(5),(1-t)\ln(8/3)+t\ln(2),(1-t)\ln(9/2)+t\ln(4/3)\},$$

which is a piecewise linear function. The picture shows that the minimum is not taken at $t=0$ nor at $t=1$.



Proof of Theorem 3. Let (λ,x) be an eigenpair for A. The equality $Ax=\lambda x$ translates into $\sum_{j=1}^n a_{i,j}x_j=\lambda x_i$ for all $i\in[1:n]$. It follows that

$$|\lambda-a_{i,i}||x_i|=\Big|\sum_{j\neq i}a_{i,j}x_j\Big|\le\sum_{j\neq i}|a_{i,j}||x_j|=\sum_{j\neq i}|a_{i,j}|^{1-t}|a_{i,j}|^t|x_j|.$$

From Hölder's inequality, which states that, for $u_1,\dots,u_m,v_1,\dots,v_m\ge 0$ and for $p,q\ge 1$ satisfying $1/p+1/q=1$,

$$\sum_{j=1}^m u_jv_j\le\Big(\sum_{j=1}^m u_j^p\Big)^{1/p}\Big(\sum_{j=1}^m v_j^q\Big)^{1/q},$$

applied with $u_j=|a_{i,j}|^{1-t}$, $v_j=|a_{i,j}|^t|x_j|$, $p=1/(1-t)$, and $q=1/t$, we derive

$$|\lambda-a_{i,i}||x_i|\le\Big(\sum_{j\neq i}|a_{i,j}|\Big)^{1-t}\Big(\sum_{j\neq i}|a_{i,j}||x_j|^{1/t}\Big)^t=R_i^{1-t}\Big(\sum_{j\neq i}|a_{i,j}||x_j|^{1/t}\Big)^t. \tag{6}$$

Now suppose that (5) does not hold, i.e., that there is an eigenvalues with $|\lambda-a_{i,i}|>R_i^{1-t}C_i^t$ for all $i\in[1:n]$. In connection with (6), we obtain (provided $|x_i|>0$)

$$R_i^{1-t}C_i^t|x_i|<R_i^{1-t}\Big(\sum_{j\neq i}|a_{i,j}||x_j|^{1/t}\Big)^t,\qquad\text{hence}\qquad C_i|x_i|^{1/t}<\sum_{j\neq i}|a_{i,j}||x_j|^{1/t}.$$

Choosing the index i such that $|x_i|=\|x\|_\infty>0$ yields

$$C_i\|x\|_\infty^{1/t}=C_i|x_i|^{1/t}<\Big(\sum_{j\neq i}|a_{i,j}|\Big)\|x\|_\infty^{1/t}=C_i\|x\|_\infty^{1/t},$$

which is absurd. We conclude that (5) holds. □

4 Exercises

Ex.1: Exercise 2 p. 351

Ex.2: Exercise 3 p. 351

Ex.3: Exercise 5 p. 351

Ex.4: Exercise 9 p. 352