Chapter 5
Matrices with Real or Complex Entries

A matrix $M \in \mathbf{M}_{n\times m}(K)$ is an element of a vector space of finite dimension n^2. When $K = \mathbb{R}$ or $K = \mathbb{C}$, this space has a natural topology, that of K^{nm}. Therefore we may manipulate such notions as open and closed sets, and continuous and differentiable functions.

5.1 Special Matrices

5.1.1 Hermitian Adjoint

When considering matrices with complex entries, a useful operation is complex conjugation $z \mapsto \bar{z}$. One denotes by $\bar{M}$ the matrix obtained from M by conjugating the entries. We then define the *Hermitian adjoint* matrix of M by

$$M^* := (\bar{M})^T = \overline{M^T}.$$

One has $m^*_{ij} = \overline{m_{ji}}$ and $\det M^* = \overline{\det M}$. The map $M \mapsto M^*$ is an *antiisomorphism*, which means that it is antilinear (meaning that $(\lambda M)^* = \bar{\lambda} M^*$) and bijective. In addition, we have the product formula

$$(MN)^* = N^* M^*.$$

If M is nonsingular, this implies $(M^*)^{-1} = (M^{-1})^*$; this matrix is sometimes denoted M^{-*}.

The interpretation of the Hermitian adjoint is that if we endow $\mathbb{C}^n$ with the canonical scalar product

$$\langle x, y\rangle = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n,$$

and with the canonical basis, then M^* is the matrix of the adjoint $(u_M)^*$; that is,

$$\langle Mx, y\rangle = \langle x, M^* y\rangle, \qquad \forall x, y \in \mathbb{C}^n.$$

D. Serre, *Matrices,* Graduate Texts in Mathematics 216,
DOI 10.1007/978-1-4419-7683-3_5, © Springer Science+Business Media, LLC 2010

5.1.2 Normal Matrices

Definition 5.1 *A matrix $M \in \mathbf{M}_n(C)$ is* normal *if M and M^* commute: $M^*M = MM^*$.*

If M has real entries, this amounts to having $MM^T = M^T M$.

Because a square matrix M always commutes with M, $-M$, or M^{-1} (assuming that the latter exists), we can define sub-classes of normal matrices. The following statement serves also as a definition of such classes.

Proposition 5.1 *The following matrices $M \in \mathbf{M}_n(\mathbb{C})$ are normal.*

- Hermitian *matrices, meaning that $M^* = M$*
- Skew-Hermitian *matrices, meaning that $M^* = -M$*
- Unitary *matrices, meaning that $M^* = M^{-1}$*

The Hermitian, skew-Hermitian, and unitary matrices are thus normal. One verifies easily that H is Hermitian (respectively, skew-Hermitian) if and only if x^*Hx is real (respectively, pure imaginary) for every $x \in \mathbb{C}^n$.

For real-valued matrices, we have instead

Definition 5.2 *A square matrix $M \in \mathbf{M}_n(\mathbb{R})$ is*

- Symmetric *if $M^T = M$*
- Skew-symmetric *if $M^T = -M$*
- Orthogonal *if $M^T = M^{-1}$*

We denote by $\mathbf{H}_n$ the set of Hermitian matrices in $\mathbf{M}_n(\mathbb{C})$. It is an $\mathbb{R}$-linear subspace of $\mathbf{M}_n(\mathbb{C})$, but not a $\mathbb{C}$-linear subspace, becausee iM is skew-Hermitian when M is Hermitian. If $M \in \mathbf{M}_{n\times m}(\mathbb{C})$, the matrices $M+M^*$, $i(M^*-M)$, MM^*, and M^*M are Hermitian. One sometimes calls $\frac{1}{2}(M+M^*)$ the *real part* of M and denotes it $\Re M$. Likewise, $\frac{1}{2i}(M-M^*)$ is the *imaginary part* of M and is denoted $\Im M$. Both are Hermitian and we have

$$M = \Re M + \mathrm{i}\Im M.$$

This terminology anticipates Chapter 10.

A matrix M is unitary if u_M is an isometry, that is $\langle Mx, My\rangle \equiv \langle x,y\rangle$. This is equivalent to saying that $\|Mx\| \equiv \|x\|$. The set of unitary matrices in $\mathbf{M}_n(\mathbb{C})$ forms a multiplicative group, denoted by $\mathbf{U}_n$. Unitary matrices satisfy $|\det M| = 1$, because $\det M^*M = |\det M|^2$ for every matrix M and $M^*M = I_n$ when M is unitary. The set of unitary matrices whose determinant equals 1, denoted by $\mathbf{SU}_n$ is obviously a normal subgroup of $\mathbf{U}_n$.

A matrix with *real* entries is orthogonal (respectively, symmetric, skew-symmetric) if and only if it is unitary, Hermitian, or skew-Hermitian.

5.1.3 Matrices and Sesquilinear Forms

Given a matrix $M \in \mathbf{M}_n(\mathbb{C})$, the map

$$(x,y) \mapsto \langle x,y\rangle_M := \sum_{j,k} m_{jk}\overline{x_j}y_k = x^* My,$$

defined on $\mathbb{C}^n \times \mathbb{C}^n$, is a sesquilinear form. When $M = I_n$, this is nothing but the scalar product. It is Hermitian if and only if M is Hermitian. It follows that $M \mapsto \langle\cdot,\cdot\rangle_M$ is an isomorphism between $\mathbf{H}_n$ and the set of Hermitian forms over $\mathbb{C}^n$. We say that a Hermitian matrix H is *degenerate* (respectively, *nondegenerate*) if the form $\langle\cdot,\cdot\rangle_H$ is so. Nondegeneracy amounts to saying that $x \mapsto Hx$ is one-to-one. In other words, we have the following.

Proposition 5.2 *A Hermitian matrix H is degenerate (respectively, nondegenerate) if and only if* $\det H = 0$ *(respectively, $\neq 0$).*

We say that the Hermitian matrix H is *positive-definite* if $\langle\cdot,\cdot\rangle_H$ is so. Then $\sqrt{\langle\cdot,\cdot\rangle_H}$ is a norm over $\mathbb{C}^n$. If $-H$ is positive-definite, we say that H is *negative-definite*. We denote by $\mathbf{HPD}_n$ the set of positive-definite Hermitian matrices. If H and K are positive-definite, and if λ is a positive real number, then $\lambda H + K$ is positive-definite. Therefore $\mathbf{HPD}_n$ is a convex cone in $\mathbf{H}_n$. This cone turns out to be open. The Hermitian matrices H for which $\langle\cdot,\cdot\rangle_H$ is a positive-semidefinite Hermitian form over $\mathbb{C}^n$ are called *positive-semidefinite* Hermitian matrices. They also form a convex cone $\mathbf{H}_n^+$. If $H \in \mathbf{H}_n^+$ and ε is a positive real number, then $H + \varepsilon I_n$ is positive-definite. Because $H + \varepsilon I_n$ tends to H as $\varepsilon \to 0^+$, we see that the closure of $\mathbf{HPD}_n$ is $\mathbf{H}_n^+$.

One defines similarly, among the real symmetric matrices, the positive-definite, respectively, positive-semidefinite, ones. Again, the positive-definite real symmetric matrices form an open cone in $\mathbf{Sym}_n(\mathbb{R})$, denoted by $\mathbf{SPD}_n$, whose closure $\mathbf{Sym}_n^+$ is made of positive-semidefinite ones.

The cone $\mathbf{HPD}_n$ defines an order over $\mathbf{H}_n$: we write $K > H$ when $K - H \in \mathbf{HPD}_n$, and more generally $K \geq H$ if $K - H$ is positive-semidefinite. The fact that

$$(K \geq H \geq K) \Longrightarrow (K = H)$$

follows from the next lemma.

Lemma 6. *Let H be Hermitian. If $x^*Hx = 0$ for every $x \in \mathbb{C}^n$, then $H = 0_n$.*

Proof. Using (1.1), we have $y^*Hx = 0$ for every $x, y \in \mathbb{C}^n$. Therefore $Hx = 0$ for every x, which means $H = 0_n$. □

We likewise define an ordering on real-valued symmetric matrices, referring to the ordering on real-valued quadratic forms.[1]

If U is unitary, the matrix U^*MU is similar to M, and we say that they are *unitary similar*. If M is normal, Hermitian, skew-Hermitian, or unitary, and if U is unitary, then U^*MU is still normal, Hermitian, skew-Hermitian, or unitary. When

[1] We warn the reader that another order, a completely different one, although still denoted by the same symbol $\geq$, is defined in Chapter 8. The latter concerns general $n \times m$ real-valued matrices, whereas the present one deals only with symmetric matrices. In practice, the context is never ambiguous.

$O \in \mathbf{O}_n(\mathbb{R})$ and $M \in \mathbf{M}_n(\mathbb{R})$, we again say that $O^T MO$ and M are *orthogonally similar*.

We notice that Lemma 6 implies the following stronger result.

Proposition 5.3 *Let $M \in \mathbf{M}_n(\mathbb{C})$ be given. If $x^*Mx = 0$ for every $x \in \mathbb{C}^n$, then $M = 0_n$.*

Proof. We decompose $M = H + \mathrm{i}K$ into its real and imaginary parts. Recall that H, K are Hermitian. Then

$$x^*Mx = x^*Hx + \mathrm{i}x^*Kx$$

is the decomposition of a complex number into real and imaginary parts. From the assumption, we therefore have $x^*Hx = 0$ and $x^*Kx = 0$ for every x. Then Lemma 6 tells us that $H = K = 0_n$. □

5.2 Eigenvalues of Real- and Complex-Valued Matrices

Let us recall that $\mathbb{C}$ is algebraically closed. Therefore the characteristic polynomial of a complex-valued square matrix has roots if $n \geq 1$. Therefore every endomorphism of a nontrivial $\mathbb{C}$-vector space possesses eigenvalues. A real-valued square matrix may have no eigenvalues in $\mathbb{R}$, but it has at least one in $\mathbb{C}$. If n is odd, $M \in \mathbf{M}_n(\mathbb{R})$ has at least one real eigenvalue, because P_M is real of odd degree.

5.2.1 Unitary Trigonalization

If $K = \mathbb{C}$, one sharpens Theorem 3.5.

Theorem 5.1 (Schur) *If $M \in \mathbf{M}_n(\mathbb{C})$, there exists a unitary matrix U such that U^*MU is upper-triangular.*

One also says that every matrix with complex entries is *unitarily trigonalizable*.

Proof. We proceed by induction over the size n of the matrices. The statement is trivial if $n = 1$. Let us assume that it is true in $\mathbf{M}_{n-1}(\mathbb{C})$, with $n \geq 2$. Let $M \in \mathbf{M}_n(\mathbb{C})$ be a matrix. Because $\mathbb{C}$ is algebraically closed, M has at least one eigenvalue λ. Let X be an eigenvector associated with λ. By dividing X by $\|X\|$, we can assume that X is a unit vector. One can then find a unitary basis $\{X^1, X^2, \dots, X^n\}$ of $\mathbb{C}^n$ whose first element is X. Let us consider the matrix $V := (X^1 = X|X^2|\cdots|X^n)$, which is unitary, and let us form the matrix $M' := V^*MV$. Because

$$VM'\mathbf{e}^1 = MV\mathbf{e}^1 = MX = \lambda X = \lambda V\mathbf{e}^1,$$

one obtains $M'\mathbf{e}^1 = \lambda\mathbf{e}^1$. In other words, M' has the block-triangular form:

$$M' = \begin{pmatrix} \lambda & \cdots \\ 0_{n-1} & N \end{pmatrix},$$

where $N \in \mathbf{M}_{n-1}(\mathbb{C})$. Applying the induction hypothesis, there exists $W \in \mathbf{U}_{n-1}$ such that W^*NW is upper-triangular. Let us denote by $\hat{W}$ the (block-diagonal) matrix $\mathrm{diag}(1,W) \in \mathbf{U}_n$. Then $\hat{W}^*M'\hat{W}$ is upper-triangular. Hence, $U = V\hat{W}$ satisfies the conditions of the theorem. □

A useful consequence of Theorem 5.1 is the following.

Corollary 5.1 *The set of diagonalizable matrices is a dense subset in* $\mathbf{M}_n(\mathbb{C})$.

Remark

The set of real matrices diagonalizable within $\mathbf{M}_n(\mathbb{R})$ is not dense in $\mathbf{M}_n(\mathbb{R})$. For instance, the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

whose eigenvalues $\pm i$ are nonreal, is interior to the set of nondiagonalizable matrices; this is a consequence of Theorem 5.2 of continuity of the spectrum. The set of real matrices diagonalizable within $\mathbf{M}_n(\mathbb{C})$ is dense in $\mathbf{M}_n(\mathbb{R})$, but the proof is more involved.

Proof. The triangular matrices with pairwise distinct diagonal entries are diagonalizable, because of Proposition 3.17, and form a dense subset of the triangular matrices. Conjugation preserves diagonalizability and is a continuous operation. Thus the closure of the diagonalizable matrices contains the matrices conjugated to a triangular matrix, that is, all of $\mathbf{M}_n(\mathbb{C})$. □

5.2.2 *The Spectrum of Special Matrices*

Proposition 5.4 *The eigenvalues of Hermitian matrices, as well as those of real symmetric matrices, are real.*

Proof. Let $M \in \mathbf{M}_n(\mathbb{C})$ be an Hermitian matrix and let λ be one of its eigenvalues. Let us choose an eigenvector X: $MX = \lambda X$. Taking the Hermitian adjoint, we obtain $X^*M = \bar{\lambda}X^*$. Hence,

$$\lambda X^*X = X^*(MX) = (X^*M)X = \bar{\lambda}X^*X,$$

or

$$(\lambda - \bar{\lambda})X^*X = 0.$$

However, $X^*X = \sum_j |x_j|^2 > 0$. Therefore, we are left with $\bar{\lambda} - \lambda = 0$. Hence λ is real. □

We leave it to the reader to show, as an exercise, that the eigenvalues of skew-Hermitian matrices are purely imaginary.

Proposition 5.5 *The eigenvalues of the unitary matrices, as well as those of real orthogonal matrices, are complex numbers of modulus one.*

Proof. As before, if X is an eigenvector associated with λ, one has

$$|\lambda|^2\|X\|^2 = (\lambda X)^*(\lambda X) = (MX)^*MX = X^*M^*MX = X^*X = \|X\|^2,$$

and therefore $|\lambda|^2 = 1$. □

5.2.3 Continuity of Eigenvalues

We study the continuity of the spectrum as a function of the matrix. The spectrum is an n-uplet $(\lambda_1,\dots,\lambda_n)$ of complex numbers. Mind that each eigenvalue is repeated according to its algebraic multiplicity. At first glance, $\mathrm{Sp}(M)$ seems to be a well-defined element of $\mathbb{C}^n$, but this is incorrect because there is no way to define a natural ordering between the eigenvalues; thus another n-uplet $(\lambda_{\sigma(1)},\dots,\lambda_{\sigma(n)})$ describes the same spectrum for every permutation σ. For this reason, the spectrum of M must be viewed as an element of the quotient space $A_n := \mathbb{C}^n/\mathscr{R}$, where $a\mathscr{R}b$ is true if and only if there exists a permutation σ such that $b_j = a_{\sigma(j)}$ for all j. There is a natural topology on A_n, given by the distance

$$d(\dot{a},\dot{b}) := \min_{\sigma\in S_n}\max_{1\le j\le n}|b_j - a_{\sigma(j)}|.$$

The metric space (A_n,d) is complete.

The way to study the continuity of

$$\begin{array}{c} M \mapsto \mathrm{Sp}(M) \\ \mathbf{M}_n(\mathbb{C}) \to A_n \end{array}$$

is to split this map into

$$\begin{array}{ccccc} M & \mapsto & P_M & \mapsto & \mathrm{Sp}(M) = \mathrm{root}(P_M), \\ \mathbf{M}_n(\mathbb{C}) & \to & \mathbf{Unit}_n & \to & A_n, \end{array}$$

where $\mathbf{Unit}_n$ is the affine space of monic polynomials of degree n, and the map *root* associates with every element of $\mathbf{Unit}_n$ its set of roots, counted with multiplicities. The first arrow is continuous, because every coefficient of P_M is a combination of minors, thus is polynomial in the entries of M. There remains to study the continuity of *root*. For the sake of completeness, we prove the following.

Lemma 7. *The map* $\mathrm{root} : \mathbf{Unit}_n \to A_n$ *is continuous.*

Proof. Let $P \in \mathbf{Unit}_n$ be given. Let $a_1, \ldots, a_r$ be the distinct roots of P, m_j their multiplicities, and ρ the minimum distance between them. We denote by D_j the open disk with center a_j and radius $\rho/2$, and C_j its boundary. The union of the C_js is a compact set on which P does not vanish. The number

$$\eta := \inf\left\{ |P(z)|\,;\, z \in \bigcup_j C_j \right\}$$

is thus strictly positive.

The affine space $\mathbf{Unit}_n$ is equipped with the distance $d(Q,R) := \|Q-R\|$ deriving from one of the (all equivalent) norms of $\mathbb{C}_{n-1}[X]$. For instance, we can take

$$\|q\| := \sup\left\{ |q(z)|\,;\, z \in \bigcup_j C_j \right\}, \qquad q \in \mathbb{C}_{n-1}[X].$$

If $d(P,Q) < \eta$, then we have

$$|P(z) - Q(z)| < |P(z)|, \qquad \forall z \in C_j, \forall j = 1, \ldots, r.$$

Rouché's theorem asserts that if two holomorphic functions f and g on a disk D, continuous over $\overline{D}$, satisfy $|f(z) - g(z)| < |f(z)|$ on the boundary of D, then they have the same number of zeroes in D, counting with multiplicities. In our case, we deduce that Q has exactly m_j roots in D_j. This sums up to $m_1 + \cdots + m_r = n$ roots in the (disjoint) union of the D_js. Because its degree is n, Q has no other roots. Therefore $d(\mathrm{root}(P), \mathrm{root}(Q)) < \rho$.

This proves the continuity of $Q \mapsto \mathrm{root}(Q)$ at P. □

As a corollary, we have the following fundamental theorem.

Theorem 5.2 *The map* $\mathrm{Sp} : \mathbf{M}_n(\mathbb{C}) \to A_n$ *is continuous.*

One often invokes this theorem by saying that *the eigenvalues of a matrix are continuous functions of its entries.*

5.2.4 Regularity of Simple Eigenvalues

The continuity result in Theorem 5.2 cannot be improved without further assumptions. For instance, the eigenvalues of

$$\begin{pmatrix} 0 & 1 \\ s & 0 \end{pmatrix}$$

are $\pm\sqrt{s}$, and thus are not differentiable functions, at least at the origin. It turns out that the obstacle to the regularity of eigenvalues is the crossing of two or more eigenvalues. But simple eigenvalues are analytic (thus $\mathscr{C}^\infty$) functions of the entries of the matrix.

Theorem 5.3 *Let λ_0 be an algebraically simple eigenvalue of a matrix $M_0 \in \mathbf{M}_n(\mathbb{C})$. Then there exists an open neighbourhood $\mathscr{M}$ of M_0 in $\mathbf{M}_n(\mathbb{C})$, and two analytic functions*

$$M \mapsto \Lambda(M), \qquad M \mapsto X(M)$$

over $\mathscr{M}$, such that

- *$\Lambda(M)$ is an eigenvalue of M.*
- *$X(M)$ is an eigenvector, associated with $\Lambda(M)$.*
- *$\Lambda(M_0) = \lambda_0$.*

Remarks

- From Theorem 5.2, if M is close to M_0, there is exactly one eigenvalue of M close to λ_0. Theorem 5.3 is a statement about this eigenvalue.
- The theorem is valid as well in $\mathbf{M}_n(\mathbb{R})$, with the same proof.

Proof. Let X_0 be an eigenvector of M_0 associated with λ_0. We know that λ_0 is also a simple eigenvalue of M_0^T. Thanks to Proposition 3.15, an eigenvector Y_0 of M_0^T (associated with λ_0) satisfies $Y_0^T X_0 \neq 0$. We normalize Y_0 in such a way that $Y_0^T X_0 = 1$.

Let us define a polynomial function F over $\mathbf{M}_n(\mathbb{C}) \times \mathbb{C} \times \mathbb{C}^n$, with values in $\mathbb{C} \times \mathbb{C}^n$, by

$$F(M, \lambda, x) := (Y_0^T x - 1, Mx - \lambda x).$$

We have $F(M_0, \lambda_0, X_0) = (0,0)$.

The differential of F with respect to (λ, x), at the base point (M_0, λ_0, X_0), is the linear map

$$(\mu, y) \overset{\delta}{\mapsto} (Y_0^T y, (M_0 - \lambda_0)y - \mu X_0).$$

Let us show that δ is one-to-one. Let (μ, y) be such that $\delta(\mu, y) = (0,0)$. Then $\mu = \mu Y_0^T X_0 = Y_0^T (M_0 - \lambda_0) y = 0^T y = 0$. After that, there remains $(M_0 - \lambda_0) y = 0$. Inasmuch as λ_0 is simple, this means that y is colinear to X_0; now the fact that $Y_0^T y = 0$ yields $y = 0$.

Because δ is a one-to-one endomorphism of $\mathbb{C} \times \mathbb{C}^n$, it is an isomorphism. We may then apply the implicit function theorem to F: there exist neighborhoods $\mathscr{M}$, $\mathscr{V}$, and $\mathscr{W}$ and analytic functions $(\Lambda, X) : \mathscr{M} \to \mathscr{V}$ such that

$$\begin{pmatrix} (M, \lambda, x) \in \mathscr{W} \\ F(M, \lambda, x) = (0,0) \end{pmatrix} \Longleftrightarrow \begin{pmatrix} M \in \mathscr{M} \\ (\lambda, x) = (\Lambda(M), X(M)) \end{pmatrix}.$$

Notice that $F = 0$ implies that (λ, x) is an eigenpair of M. Therefore the theorem is proved. □

5.3 Spectral Decomposition of Normal Matrices

We recall that a matrix M is *normal* if M^* commutes with M. For real matrices, this amounts to saying that M^T commutes with M. Because it is equivalent for an Hermitian matrix H to be zero or to satisfy $x^*Hx = 0$ for every vector x, we see that M is normal if and only if $\|Mx\|_2 = \|M^*x\|_2$ for every vector, where $\|x\|_2$ denotes the standard Hermitian (Euclidean) norm (take $H = MM^* - M^*M$).

Theorem 5.4 *In* $\mathbf{M}_n(\mathbb{C})$*, a matrix is normal if and only if it is unitarily diagonalizable:*

$$(M^*M = MM^*) \Longleftrightarrow (\exists U \in \mathbf{U}_n; \quad M = U^{-1}\operatorname{diag}(d_1,\ldots,d_n)U).$$

This theorem contains the following properties.

Corollary 5.2 *Unitary, Hermitian, and skew-Hermitian matrices are unitarily diagonalizable.*

Observe that among normal matrices one distinguishes each of the above families by the nature of their eigenvalues. Those of unitary matrices have modulus one, and those of Hermitian matrices are real. Finally, those of skew-Hermitian matrices are purely imaginary.

Proof. A diagonal matrix is obviously normal. If U is unitary, a matrix M is normal if and only if U^*MU is normal: we deduce that unitarily diagonalizable matrices are normal.

We now prove the converse. We proceed by induction on the size n of the matrix M. If $n = 0$, there is nothing to prove. Otherwise, if $n \geq 1$, there exists an eigenpair (λ, x):

$$Mx = \lambda x, \quad \|x\|_2 = 1.$$

Because M is normal, $M - \lambda I_n$ is, too. From the above, we see that $\|(M^* - \bar{\lambda})x\|_2 = \|(M-\lambda)x\|_2 = 0$, and hence $M^*x = \bar{\lambda}x$. Let V be a unitary matrix such that $V\mathbf{e}^1 = x$. Then the matrix $M_1 := V^*MV$ is normal and satisfies $M_1\mathbf{e}^1 = \lambda\mathbf{e}^1$. Hence it satisfies $M_1^*\mathbf{e}^1 = \bar{\lambda}\mathbf{e}^1$. This amounts to saying that M_1 is block-diagonal, of the form $M_1 = \operatorname{diag}(\lambda, M')$. Obviously, M' inherits the normality of M_1. From the induction hypothesis, M', and therefore M_1 and M, are unitarily diagonalizable. □

One observes that the same matrix U diagonalizes M^*, because $M = U^{-1}DU$ implies $M^* = U^*D^*U^{-1*} = U^{-1}D^*U$, because U is unitary.

Let us consider the case of a positive-semidefinite Hermitian matrix H. If $HX = \lambda X$, then $0 \leq X^*HX = \lambda\|X\|^2$. The eigenvalues are thus nonnegative. Let $\lambda_1, \ldots, \lambda_p$ be the nonzero eigenvalues of H. Then H is unitarily similar to

$$D := \operatorname{diag}(\lambda_1, \ldots, \lambda_p, 0, \ldots, 0).$$

From this, we conclude that $\operatorname{rk} H = p$. Let $U \in \mathbf{U}_n$ be such that $H = UDU^*$. Defining the vectors $X_\alpha = \sqrt{\lambda_\alpha}U_\alpha$, where the U_α are the columns of U, we obtain the following statement.

Proposition 5.6 *Let $H \in \mathbf{M}_n(\mathbb{C})$ be a positive-semidefinite Hermitian matrix. Let p be its rank. Then H has p real, positive eigenvalues, and the eigenvalue $\lambda = 0$ has multiplicity $n - p$. There exist p column vectors X_α, pairwise orthogonal, such that*

$$H = X_1 X_1^* + \cdots + X_p X_p^*.$$

Finally, H is positive-definite if and only if $p = n$ (in which case, $\lambda = 0$ is not an eigenvalue).

5.4 Normal and Symmetric Real-Valued Matrices

The situation is a bit more involved if M, a normal matrix, has real entries. Of course, one can consider M as a matrix with complex entries and diagonalize it on a unitary basis, but if M has a nonreal eigenvalue, we quit the field of real numbers when doing so. We prefer to allow orthonormal bases consisting of only *real* vectors. Some of the eigenvalues might be nonreal, thus one cannot in general diagonalize M. The statement is thus the following.

Theorem 5.5 *Let $M \in \mathbf{M}_n(\mathbb{R})$ be a normal matrix. There exists an orthogonal matrix O such that OMO^{-1} is block-diagonal, the diagonal blocks being 1×1 (those corresponding to the real eigenvalues of M) or 2×2, the latter being matrices of direct similitude:*[2]

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad (b \neq 0).$$

Likewise, $OM^T O^{-1}$ is block-diagonal, the diagonal blocks being eigenvalues or matrices of direct similitude.

Proof. One again proceeds by induction on n. When $n \geq 1$, the proof is the same as in the previous section whenever M has at least one real eigenvalue.

If this is not the case, then n is even. Let us first consider the case $n = 2$. Then

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

This matrix is normal, therefore we have $b^2 = c^2$ and $(a-d)(b-c) = 0$. However, $b \neq c$, because otherwise M would be symmetric, and hence would have two real eigenvalues. Hence $b = -c$ and $a = d$.

Now let us consider the general case, with $n \geq 4$. We know that M has an eigenpair (λ, z), where λ is not real. If the real and imaginary parts of z were colinear, M would have a real eigenvector, hence a real eigenvalue, a contradiction. In other words, the real and imaginary parts of z span a plane P in $\mathbb{R}^n$. As before, $Mz = \lambda z$ implies $M^T z = \bar{\lambda} z$. Hence we have $MP \subset P$ and $M^T P \subset P$. Now let V be an orthogonal

[2] A similitude is an endomorphism of a Euclidean space that preserves angles. It splits as aR, where R is orthogonal and a is a scalar. It is direct if its determinant is positive.

matrix that maps the plane $P_0 := \mathbb{R}\mathbf{e}^1 \oplus \mathbb{R}\mathbf{e}^2$ onto P. Then the matrix $M_1 := V^T MV$ is normal and satisfies

$$M_1 P_0 \subset P_0, \quad M_1^T P_0 \subset P_0.$$

This means that M_1 is block-diagonal. Of course, each diagonal block (of sizes 2×2 and $(n-2)\times(n-2)$) inherits the normality of M_1. Applying the induction hypothesis, we know that these blocks are unitarily similar to a block-diagonal matrix whose diagonal blocks are direct similitudes. Hence M_1 and M are unitarily similar to such a matrix. □

Corollary 5.3 *Real symmetric matrices are diagonalizable over* $\mathbb{R}$*, through orthogonal conjugation. In other words, given* $M \in \mathbf{Sym}_n(\mathbb{R})$*, there exists an* $O \in \mathbf{O}_n(\mathbb{R})$ *such that* OMO^{-1} *is diagonal.*

In fact, because the eigenvalues of M are real, OMO^{-1} has only 1×1 blocks. We say that real symmetric matrices are *orthogonally diagonalizable*.

The interpretation of this statement in terms of quadratic forms is the following. For every quadratic form Q on $\mathbb{R}^n$, there exists an orthonormal basis $\{e_1,\dots,e_n\}$ in which this form can be written with at most n squares:[3]

$$Q(x) = \sum_{i=1}^{n} a_i x_i^2.$$

Replacing the basis vector e_j by $|a_j|^{1/2} e_j$, one sees that there also exists an orthogonal basis in which the quadratic form can be written

$$Q(x) = \sum_{i=1}^{r} x_i^2 - \sum_{j=1}^{s} x_{j+r}^2,$$

with $r+s \le n$. This quadratic form is nondegenerate if and only if $r+s=n$. The pair (r,s) is unique and called the *signature* or the *Sylvester index* of the quadratic form. In such a basis, the matrix associated with Q is

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}.$$

[3] In solid mechanics, when Q is the matrix of inertia, the vectors of this basis are along the *inertia axes*, and the a_j, which then are positive, are the momenta of inertia.

5.5 Functional Calculus

Given a square matrix $A \in \mathbf{M}_n(\mathbb{C})$ and a function $f : \mathscr{U} \to \mathbb{C}$, we should like to define a matrix $f(A)$, in such a way that the maps $A \mapsto f(A)$ and $f \mapsto f(A)$ have nice properties.

The case of polynomials is easy. If

$$P(X) = a_0 X^m + a_1 X^{m-1} + \cdots + a_{m-1} X + a_m$$

has complex coefficients, then we define

$$P(A) = a_0 A^m + a_1 A^{m-1} + \cdots + a_{m-1} A + a_m I_n.$$

Remark that this definition does not need the scalar field to be that of complex numbers.

An important consequence of the Cayley–Hamilton theorem is that if two polynomials P and Q are such that the characteristic polynomial P_A divides $Q-P$, then $Q(A) = P(A)$. This shows that what really matters is the behavior of P and of a few derivatives at the eigenvalues of A. By *few derivatives*, we mean that if ℓ is the algebraic multiplicity of an eigenvalue λ, then one only needs to know $P(\lambda), \dots, P^{(\ell-1)}(\lambda)$. Actually, ℓ can be chosen as the multiplicity of the root λ in the minimal polynomial of A. For instance, if $N \in \mathbf{M}_n(k)$ is a nilpotent matrix, then the Taylor formula yields

$$P(N) = P(0) I_n + P'(0) N + \cdots + \frac{1}{(n-1)!} P^{(n-1)}(0) N^{n-1}. \tag{5.1}$$

This suggests the following treatment when f is a holomorphic function. Naturally, we ask that its domain $\mathscr{U}$ contain $\mathrm{Sp} A$. We interpolate f at order n at every point of $\mathrm{Sp} A$, by a polynomial P:

$$P^{(r)}(\lambda) = f^{(r)}(\lambda), \qquad \forall \lambda \in \mathrm{Sp} A, \quad \forall 0 \le r \le n-1.$$

We then define

$$f(A) := P(A). \tag{5.2}$$

In order that this definition be meaningful, we verify that it does not depend upon the choice of the interpolation polynomial. This turns out to be true, because if Q is another interpolation polynomial as above, then $Q-P$ is divisible by

$$\prod_{\lambda \in \mathrm{Sp} A} (X - \lambda)^n,$$

thus by P_A, and therefore $Q(A) = P(A)$.

Proposition 5.7 *The functional calculus with holomorphic functions enjoys the following properties. Below, the domains of functions are such that the expressions make sense.*

- *If f and g match at order n over* $\mathrm{Sp}A$*, then* $f(A) = g(A)$.
- *Conjugation: If M is nonsingular, then* $f(M^{-1}AM) = M^{-1}f(A)M$.
- *Linearity:* $(af + g)(A) = af(A) + g(A)$.
- *Algebra homomorphism:* $(fg)(A) = f(A)g(A)$.
- *Spectral mapping:* $\mathrm{Sp}\, f(A) = f(\mathrm{Sp}A)$.
- *Composition:* $(f \circ g)(A) = f(g(A))$.

Proof. The first property follows directly from the definition, and the linearity is obvious. The conjugation formula is already true for polynomials.

The product formula is true for polynomials. Now, if f and g are interpolated by P and Q, respectively, at order n at every point of $\mathrm{Sp}A$, then fg is likewise interpolated by PQ. This proves the product formula for holomorphic functions.

We now prove the spectral mapping formula. For this, let (λ, x) be an eigenpair: $Ax = \lambda x$. Let P be an interpolation polynomial of f as above. Then

$$f(A)x = P(A)x = a_0A^m x + \cdots + a_m x = (a_0\lambda^m + \cdots + a_m)x = P(\lambda)x = f(\lambda)x.$$

Therefore $f(\lambda) \in \mathrm{Sp}\, f(A)$, which tells us $f(\mathrm{Sp}A) \subset \mathrm{Sp}\, f(A)$.

Conversely, let R be a polynomial vanishing at order n over $f(\mathrm{Sp}A)$. We have $R(f(A)) = (R \circ f)(A) = 0_n$, because $R \circ f$ is flat at order n over $\mathrm{Sp}A$. Replacing A by $f(A)$ and f by R above, we have $R(\mathrm{Sp}\, f(A)) \subset \mathrm{Sp}\, 0_n = \{0\}$. We have proved that if R vanishes at order n over $f(\mathrm{Sp}A)$, then it vanishes at $\mathrm{Sp}\, f(A)$. Therefore $\mathrm{Sp}\, f(A) \subset f(\mathrm{Sp}A)$.

There remains to treat composition. Because of the spectral mapping formula, our assumption is that the domain of f contains $g(\mathrm{Sp}A)$. If f and g are interpolated at order n by P and Q at $g(\mathrm{Sp}A)$ and $\mathrm{Sp}A$ respectively, then $f \circ g$ is interpolated at $\mathrm{Sp}A$ at order n by $P \circ Q$. The formula is true for polynomials, thus it is true for $f \circ g$ too. □

Remark

As long as we are interested in polynomial functions only, Proposition 5.7 is valid in $\mathbf{M}_n(k)$ for an arbitrary field.

5.5.1 The Dunford–Taylor Formula

An alternate definition can be given in terms of a Cauchy integral: the so-called *Dunford–Taylor integral.*

Proposition 5.8 *Let f be holomorphic over a domain $\mathscr{U}$ containing the spectrum of $A \in \mathbf{M}_n(\mathbb{C})$. Let Γ be a positively oriented contour around* $\mathrm{Sp}A$*, contained in $\mathscr{U}$.*

Then we have

$$f(A) = \frac{1}{2i\pi} \int_\Gamma f(z)(zI_n - A)^{-1} dz.$$

Proof. Because of the conjugation property, and thanks to Proposition 3.20, it is enough to verify the formula when $A = \lambda I_n + N$ where N is nilpotent. By translation, it suffices to treat the nilpotent case.

So we assume that A is nilpotent. Therefore Γ is a disjoint union of Jordan curves. It is oriented in the trigonometric sense with index one around the origin. Because of nilpotence, we have

$$(zI_n - A)^{-1} = z^{-1} I_n + z^{-2} A + \cdots + z^{-n} A^{n-1}.$$

Thanks to the Cauchy formula

$$\frac{1}{2i\pi} \int_\Gamma f(z) z^{-m-1}\, dz = \frac{1}{m!} f^{(m)}(0),$$

we obtain

$$\frac{1}{2i\pi} \int_\Gamma f(z)(zI_n - A)^{-1} dz = f(0) I_n + f'(0) A + \cdots + \frac{1}{(n-1)!} f^{(n-1)}(0) A^{(n-1)}.$$

Now let P be a polynomial matching f at order n at the origin. We have

$$\begin{aligned}\frac{1}{2i\pi} \int_\Gamma f(z)(zI_n - A)^{-1} dz &= P(0) I_n + P'(0) A + \cdots + \frac{1}{(n-1)!} P^{(n-1)}(0) A^{(n-1)} \\ &= P(A) = f(A),\end{aligned}$$

where we have used Formula (5.1), and the definition of $f(A)$. □

Definition 5.3 *The factor $(zI_n - A)^{-1}$ appearing in the Dunford–Taylor formula is the* resolvent *of A at z. It is denoted $R(z;A)$. The domain of $z \mapsto R(z;A)$, which is the complement of* $\mathrm{Sp} A$*, is the* resolvent set.

5.5.2 Invariant Subspaces

An important situation occurs when f is an indicator function; that is, $f(z)$ takes its values in $\{0,1\}$. Because f is assumed to be holomorphic, hence continuous, the closures of the subdomains

$$\mathscr{U}_0 := \{z \mid f(z) = 0\}, \qquad \mathscr{U}_1 := \{z \mid f(z) = 1\}$$

are disjoint sets. Because of the multiplicative property, we see that $f(A)^2 = (f^2)(A) = f(A)$. This tells us that $f(A)$ is a *projector*. If E and F denote its kernel and range, we have $\mathbb{C}^n = E \oplus F$.

Again, the multiplicative property tells us that $f(A)$ commutes with A. This implies that both E and F are invariant subspaces for A:

$$A(E) \subset E, \qquad A(F) \subset F.$$

Using Proposition 3.20 and applying Proposition 5.8 blockwise, we see that passing from A to $f(A)$ amounts to keeping the diagonal blocks $\lambda_j I_{n_j} + N_j$ of the Dunford decomposition for which $f(\lambda_j) = 1$, while dropping those for which $f(\lambda_j) = 0$.

In particular, when $\mathscr{U}_1$ contains precisely one eigenvalue λ (which may have some multiplicity), $f(A)$ is called an *eigenprojector*, because it is the projection onto the characteristic subspace $E(\lambda) := \ker(A - \lambda I_n)^n$, parallel to the other characteristic subspaces.

5.5.2.1 Stable and Unstable Subspaces

The following notions are useful in the linear theory of differential equations, especially when studying asymptotic behavior as time goes to infinity.

Definition 5.4 *Let $A \in \mathbf{M}_n(\mathbb{C})$ be given. Its* stable *invariant subspace is the sum of the subspaces $E(\lambda)$ over the eigenvalues of negative real part. The* unstable *subspace is the sum over the eigenvalues of positive real part. At last, the* central *subspace is the sum over the pure imaginary eigenvalues.*

These spaces are denoted, respectively, $S(A)$, $U(A)$, and $C(A)$.

By invariance of $E(\lambda)$, the stable, unstable, and central subspaces are each invariant under A. From the above analysis, there are three corresponding eigenprojectors π_s, π_u, and π_c, given by the Dunford–Taylor formulæ. For instance, π_s is obtained by choosing $f_s \equiv 1$ over $\Re z < -\varepsilon$ and $f_s \equiv 0$ over $\Re z > \varepsilon$, for a small enough positive ε. In other words,

$$\pi_s = \frac{1}{2i\pi} \int_{\Gamma_s} (zI_n - A)^{-1} dz$$

for some large enough circle Γ_s contained in $\Re z < 0$.

Because $f_s + f_u + f_c \equiv 1$ around the spectrum of A, we have

$$\pi_s + \pi_u + \pi_c = f_s(A) + f_u(A) + f_c(A) = \mathbf{1}(A) = I_n.$$

In addition, the properties $f_s f_u \equiv 0$, $f_s f_c \equiv 0$, and $f_c f_u \equiv 0$ around the spectrum yield

$$\pi_s \pi_u = \pi_u \pi_c = \pi_c \pi_s = 0_n.$$

The identities above give, as expected,

$$\mathbb{C}^n = S(A) \oplus U(A) \oplus C(A).$$

We leave the following characterization to the reader.

Proposition 5.9 *The stable (respectively, unstable) subspace of A is the set of vectors $x \in \mathbb{C}^n$ such that the solution of the Cauchy problem*

$$\frac{dy}{dt} = Ay, \qquad y(0) = x$$

tends to zero exponentially fast as time goes to $+\infty$ (respectively, to $-\infty$).

5.5.2.2 Contractive/Expansive Invariant Subspaces

In the linear theory of discrete dynamical systems, that is, of iterated sequences

$$x^{m+1} = \phi(x^m),$$

what matters is the position of the eigenvalues with respect to the unit circle. We thus define the *contractive subspace* as the sum of $E(\lambda)$s over the eigenvalues of modulus less than 1. The sum over $|\lambda| > 1$ is called the *expansive subspace*. At last, the sum over $|\lambda| = 1$ is the *neutral subspace*. Again, $\mathbb{C}^n$ is the direct sum of these three invariant subspaces.

The link with the stable/unstable subspaces can be described in terms of the exponential of matrices, a notion developed in Chapter 10. The contractive (respectively, expansive, neutral) subspace of $\exp A$ coincides with $S(A)$ (respectively, with $U(A)$, $C(A)$).

5.6 Numerical Range

In this paragraph, we denote $\|x\|_2$ the Hermitian norm in $\mathbb{C}^n$. If $A \in \mathbf{M}_n(\mathbb{C})$ and x is a vector, the expression $r_A(x) = x^*Ax$ is a complex number.

Definition 5.5 *The* numerical range *of A is the subset of the complex plane*

$$\mathscr{H}(A) = \{r_A(x) \mid \|x\|_2 = 1\}.$$

The numerical range is obviously compact. It is unitarily invariant:

$$\mathscr{H}(U^*AU) = \mathscr{H}(A), \qquad \forall U \in \mathbf{U}_n,$$

because $x \mapsto Ux$ is a bijection between unitary vectors. It is thus enough to evaluate the numerical range over upper-triangular matrices, thanks to Theorem 5.1.

5.6.1 The Numerical Range of a 2×2 Matrix

Let us begin with the case $n = 2$. As mentioned above, it is enough to consider triangular matrices. By adding μI_n, we shift the numerical range by a complex number μ. Doing so, we may reduce our analysis to the case where $\operatorname{Tr} A = 0$. Next, multiplying A by z has the effect of applying similitude to the numerical range, whose magnitude is $|z|$ and angle is $\operatorname{Arg}(z)$. Doing so, we reduce our analysis to either

$$A = 0_2, \text{ or } A = 2J_2 := \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \text{ or } A = B_a := \begin{pmatrix} 1 & 2a \\ 0 & -1 \end{pmatrix},$$

for some $a \in \mathbb{C}$. At last, conjugating by $\mathrm{diag}(1, e^{i\beta})$ for some real β, we may assume that a is real, nonnegative.

Clearly, $\mathscr{H}(0_2) = \{0\}$. The case of $2J_2$ is quite simple:

$$\mathscr{H}(2J_2) = \{2yz \,|\, |y|^2 + |z|^2 = 1\}.$$

This is a rotationally invariant set, containing the segment $[0,1]$ and contained in the unit disk, by Cauchy–Schwarz. Thus it equals the unit disk.

Let us examine the third case in details.

$$\mathscr{H}(B_a) = \{|y|^2 - |z|^2 + 2a\bar{y}z \,|\, |y|^2 + |z|^2 = 1\}.$$

At fixed moduli $|y|$ and $|z|$, the number $r_{B_a}(x)$ runs over a circle whose center is $|y|^2 - |z|^2$ and radius is $2a|yz|$. Therefore $\mathscr{H}(B_a)$ is the union of the circles $C(r;\rho)$ where the center r is real, and (r,ρ) is constrained by

$$a^2 r^2 + \rho^2 = a^2.$$

Let $\mathscr{E} \in \mathbb{C} \sim \mathbb{R}^2$ be the filled ellipse with foci ± 1 and passing through $z = \sqrt{1+a^2}$. It is defined by the inequality

$$\frac{(\Re z)^2}{1+a^2} + \frac{(\Im z)^2}{a^2} \le 1.$$

If $z = r + \rho e^{i\theta} \in C(r;\rho)$, we have

$$\frac{(\Re z)^2}{1+a^2} + \frac{(\Im z)^2}{a^2} = \frac{r^2}{1+a^2} + \frac{\rho^2}{a^2} + \frac{2r\rho}{1+a^2}\cos\theta - \frac{\rho^2}{a^2(1+a^2)}\cos^2\theta.$$

Considering this expression as a quadratic polynomial in $\cos\theta$, its maximum is reached when the argument equals[4] ra^2/ρ and then it takes the value $r^2 + \rho^2/a^2$. The latter being less than or equal to one, we deduce that z belongs to $\mathscr{E}$. Therefore $\mathscr{H}(A) \subset \mathscr{E}$.

Conversely, let z belong to $\mathscr{E}$. The polynomial $r \mapsto g(r) := (\Im z)^2 + (\Re z - r)^2 + a^2(r^2 - 1)$ is convex and reaches its minimum at $r_0 = (\Re z)/(1+a^2)$, which belongs to $[-1,1]$. We have

$$g(r_0) = (\Im z)^2 + \frac{a^2}{1+a^2}(\Re z)^2 - a^2,$$

which is nonpositive by assumption. Because $g(\pm 1) \ge 0$, we deduce the existence of an $r \in [-1,1]$ such that $g(r) = 0$. This precisely means that $z \in C(r;\rho)$. Therefore $\mathscr{E} \subset \mathscr{H}(A)$.

Finally, $\mathscr{H}(A)$ is a filled ellipse whose foci are ± 1. Its great axis is $\sqrt{1+a^2}$ and the small one is a. Its area is therefore $\pi a\sqrt{1+a^2}$, which turns out to equal

[4] The value ra^2/ρ is not necesssarily within $[-1,1]$, but we don't mind. When $ra^2/\rho > 1$, the circle is contained in the interior of $\mathscr{E}$, whereas if $ra^2/\rho \le 1$, it is interiorly tangent to $\mathscr{E}$.

$|\det[B_a^*, B_a]|^{1/2}$, because we have

$$[B_a^*, B_a] = \begin{pmatrix} -4a^2 & -4a \\ 4a & 4a^2 \end{pmatrix}.$$

We notice that the same formula holds true for the other cases $A = 0_2$ or $A = 2J_2$.

Going backward to a general 2×2 matrix through similitude and conjugation, and pointing out that affine transformations preserve the class of ellipse, while multiplying the area by the Jacobian determinant, we have established the following.

Lemma 8. *The numerical range of a 2×2 matrix is a filled ellipse whose foci are its eigenvalues. Its area equals*

$$\frac{\pi}{4}\,|\det[A^*, A]|^{1/2}.$$

5.6.2 *The General Case*

Let us turn towards matrices of sizes $n \geq 3$ (the case $n = 1$ being trivial). If z, z' belong to the numerical range, we have $z = r_A(x)$ and $z' = r_A(x')$ for suitable unit vectors. Applying Lemma 8 to the restriction of r_A to the plane spanned by x and x', we see that there is a filled ellipse, containing z and z', and contained in $\mathscr{H}(A)$. Therefore the segment $[z, z']$ is contained in the numerical range and $\mathscr{H}(A)$ is convex.

If x is a unitary eigenvector, then $r_A(x) = x^*(\lambda x) = \lambda$. Finally we have the following.

Theorem 5.6 (Toeplitz–Hausdorff) *The numerical range of a matrix $A \in \mathbf{M}_n(\mathbb{C})$ is a compact convex domain. It contains the eigenvalues of A.*

5.6.2.1 The Case of Normal Matrices

If A is normal, we deduce from Theorem 5.4 that its numerical range equals that of the diagonal matrix D with the same eigenvalues $a_1, \dots, a_n$. Denoting $\theta_j = |x_j|^2$ when $x \in \mathbb{C}^n$, we see that

$$\mathscr{H}(A) = \{\theta_1 a_1 + \cdots + \theta_n a_n \,|\, \theta_1, \dots, \theta_n \geq 0 \text{ and } \theta_1 + \cdots + \theta_n = 1\}.$$

This is precisely the convex envelope of $a_1, \dots, a_n$.

Proposition 5.10 *The numerical range of a normal matrix is the convex hull of its eigenvalues.*

5.6.3 The Numerical Radius

Definition 5.6 *The* numerical radius *of* $A \in \mathbf{M}_n(\mathbb{C})$ *is the nonnegative real number*

$$w(A) := \sup\{|z|\,;\, z \in \mathscr{H}(A)\} = \sup_{x \neq 0} \frac{|x^*Ax|}{\|x\|} = \sup_{\|x\|=1} |x^*Ax|.$$

As a supremum of seminorms $A \mapsto |x^*Ax|$, it is a seminorm. Because of Proposition 5.3, $w(A) = 0$ implies $A = 0_n$. The numerical radius is thus a norm. We warn the reader that it is not a *matrix norm* (this notion is developed in Chapter 7) inasmuch as $w(AB)$ is not always less than or equal to $w(A)w(B)$. For instance

$$w(J_2^T J_2) = 1, \text{ and } w(J_2^T) = w(J_2) = \frac{1}{2}.$$

However, the numerical radius satisfies the inequality $w(A^k) \leq w(A)^k$ for every positive integer k (see Exercise 8). A norm with this property is called *superstable*.

Because $\mathbf{M}_n(\mathbb{C})$ is finite dimensional, the numerical radius is equivalent as a norm to any other norm, for instance to matrix norms. However, norm equivalence involves constant factors, which may depend dramatically on the dimension n. It is thus remarkable that the equivalence with the standard operator norm is uniform in n:

Proposition 5.11 *For every* $A \in \mathbf{M}_n(\mathbb{C})$, *we have*

$$w(A) \leq \|A\|_2 \leq 2w(A).$$

Proof. Cauchy–Schwarz gives

$$|x^*Ax| \leq \|x\|_2 \|Ax\|_2 \leq \|A\|_2 \|x\|_2^2,$$

which yields $w(A) \leq \|A\|_2$.

On the other hand, let us majorize $|y^*Ax|$ in terms of $w(A)$. We have

$$4y^*Ax = (x+y)^*A(x+y) - (x-y)^*A(x-y) + \mathrm{i}(x+\mathrm{i}y)^*A(x+\mathrm{i}y) - \mathrm{i}(x-\mathrm{i}y)^*A(x-\mathrm{i}y).$$

The triangle inequality and the definition of $w(A)$ then give

$$\begin{aligned} 4|y^*Ax| &\leq \left(\|x+y\|_2^2 + \|x-y\|_2^2 + \|x+\mathrm{i}y\|_2^2 + \|x-\mathrm{i}y\|_2^2\right) w(A) \\ &= 4(\|x\|_2^2 + \|y\|_2^2) w(A). \end{aligned}$$

If x and y are unit vectors, this means $|y^*Ax| \leq 2w(A)$. If x is a unit vector, we now write

$$\|Ax\|_2 = \sup\{|y^*Ax|\,;\, \|y\|_2 = 1\} \leq 2w(A).$$

Taking the supremum over x, we conclude that $\|A\|_2 \leq w(A)$. $\square$

Specific examples show that each one of the inequalities in Proposition 5.11 can be an equality.

5.7 The Gershgorin Domain

In this section, we use the norm $\|\cdot\|_\infty$ over $\mathbb{C}^n$, defined by

$$\|x\|_\infty := \max_i |x_i|.$$

Let $A \in \mathbf{M}_n(\mathbb{C})$, λ be an eigenvalue and x an associated eigenvector. Let i be an index such that $|x_i| = \|x\|_\infty$. Then $x_i \neq 0$ and the majorization

$$|a_{ii} - \lambda| = \left| \sum_{j\neq i} a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j\neq i} |a_{ij}|$$

gives the following.[5]

Proposition 5.12 (Gershgorin) *The spectrum of A is included in the* Gershgorin domain $\mathscr{G}(A)$*, defined as the union of the* Gershgorin disks

$$D_i(A) := D(a_{ii}; r_i), \qquad r_i := \sum_{j\neq i} |a_{ij}|.$$

Replacing A by its transpose, which has the same spectrum, we have likewise

$$\mathrm{Sp}(A) \subset \mathscr{G}(A^T) = \bigcup_{j=1}^{n} D'_j(A), \qquad D'_j(A) := D_j(A^T) = D(a_{jj}; r'_j), \qquad r'_j := \sum_{i\neq j} |a_{ij}|.$$

One may improve this result by considering the connected components of $\mathscr{G}(A)$. Let G be one of them. It is the union of the D_ks that meet G. Let p be the number of such disks. One has $G = \cup_{i\in I} D_i(A)$ where I has cardinality p.

Theorem 5.7 *There are exactly p eigenvalues of A in G, counted with their multiplicities.*

Proof. For $r \in [0,1]$, we define a matrix $A(r)$ by the formula

$$a_{ij}(r) := \begin{cases} a_{ii}, & j = i, \\ r a_{ij}, & j \neq i. \end{cases}$$

[5] This result can also be deduced from Proposition 7.5: let us decompose $A = D + C$, where D is the diagonal part of A. If $\lambda \neq a_{ii}$ for every i, then $\lambda I_n - A = (\lambda I_n - D)(I_n - B)$ with $B = (\lambda I_n - D)^{-1} C$. Hence, if λ is an eigenvalue, then either λ is an a_{ii}, or $\|B\|_\infty \geq 1$.

It is clear that the Gershgorin domain $\mathscr{G}_r$ of $A(r)$ is included in $\mathscr{G}(A)$. We observe that $A(1) = A$, and that $r \mapsto A(r)$ is continuous. Let us denote by $m(r)$ the number of eigenvalues (counted with multiplicity) of $A(r)$ that belong to G.

Because G and $\mathscr{G}(A) \setminus G$ are compact, one can find a Jordan curve, oriented in the trigonometric sense, that separates G from $\mathscr{G}(A) \setminus G$. Let Γ be such a curve. Inasmuch as $\mathscr{G}_r$ is included in $\mathscr{G}(A)$, the residue formula expresses $m(r)$ in terms of the characteristic polynomial P_r of $A(r)$:

$$m(r) = \frac{1}{2i\pi} \int_\Gamma \frac{P_r'(z)}{P_r(z)}\, dz.$$

Because P_r does not vanish on Γ and $r \mapsto P_r, P_r'$ are continuous, $r \mapsto m(r)$ is continuous. Because $m(r)$ is an integer and $[0,1]$ is connected, $m(r)$ remains constant. In particular, $m(0) = m(1)$.

Finally, $m(0)$ is the number of entries a_{jj} (eigenvalues of $A(0)$) that belong to G. But a_{jj} is in G if and only if $D_j(A) \subset G$. Hence $m(0) = p$, which implies $m(1) = p$, the desired result. □

An improvement of Gershgorin's theorem concerns irreducible matrices.

Proposition 5.13 *Let A be an irreducible matrix. If an eigenvalue of A does not belong to the interior of any Gershgorin disk, then it belongs to every circle $S(a_{ii}; r_i)$.*

Proof. Let λ be such an eigenvalue and x an associated eigenvector. By assumption, one has $|\lambda - a_{ii}| \ge \sum_{j\ne i} |a_{ij}|$ for every i. Let I be the set of indices for which $|x_i| = \|x\|_\infty$ and let J be its complement. If $i \in I$, then

$$\|x\|_\infty \sum_{j\ne i} |a_{ij}| \le |\lambda - a_{ii}|\, \|x\|_\infty = \left| \sum_{j\ne i} a_{ij} x_j \right| \le \sum_{j\ne i} |a_{ij}|\, |x_j|.$$

It follows that $\sum_{j\ne i}(\|x\|_\infty - |x_j|)|a_{ij}| \le 0$, where all the terms in the sum are nonnegative. Each term is thus zero, so that $a_{ij} = 0$ for $j \in J$. Because A is irreducible, J is empty. One has thus $|x_j| = \|x\|_\infty$ for every j, and the previous inequalities show that λ belongs to every circle. □

5.7.1 An Application

Definition 5.7 *A square matrix $A \in \mathbf{M}_n(\mathbb{C})$ is said to be*

1. Diagonally dominant *if*

$$|a_{ii}| \ge \sum_{j\ne i} |a_{ij}|, \quad 1 \le i \le n$$

2. Strongly diagonally dominant *if it is diagonally dominant and in addition at least one of these n inequalities is strict*

3. Strictly diagonally dominant *if the inequality above is strict for every index i*

Corollary 5.4 *Let A be a square matrix. If A is strictly diagonally dominant, or if A is irreducible and strongly diagonally dominant, then A is invertible.*

In fact, either zero does not belong to the Gershgorin domain, or it is not interior to the disks. In the latter case, A is assumed to be irreducible, and there exists a disk D_j that does not contain zero.

Exercises

1. Show that the eigenvalues of skew-Hermitian matrices, as well as those of real skew-symmetric matrices, are pure imaginary.
2. Let $P,Q \in \mathbf{M}_n(\mathbb{R})$ be given. Assume that $P + iQ \in \mathbf{GL}_n(\mathbb{C})$. Show that there exist $a, b \in \mathbb{R}$ such that $aP + bQ \in \mathbf{GL}_n(\mathbb{R})$. Deduce that if $M,N \in \mathbf{M}_n(\mathbb{R})$ are similar in $\mathbf{M}_n(\mathbb{C})$, then these matrices are similar in $\mathbf{M}_n(\mathbb{R})$.
3. Given an invertible matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(\mathbb{R}),$$

define a map h_M from $S^2 := \mathbb{C} \cup \{\infty\}$ into itself by

$$h_M(z) := \frac{az+b}{cz+d}.$$

   a. Show that h_M is a bijection.
   b. Show that $h : M \mapsto h_M$ is a group homomorphism. Compute its kernel.
   c. Let $\mathscr{H}$ be the upper half-plane, consisting of those $z \in \mathbb{C}$ with $\Im z > 0$. Compute $\Im h_M(z)$ in terms of $\Im z$ and deduce that the subgroup

$$\mathbf{GL}_2^+(\mathbb{R}) := \{M \in \mathbf{GL}_2(\mathbb{R}) \mid \det M > 0\}$$

   operates on $\mathscr{H}$.
   d. Conclude that the group $\mathbf{PSL}_2(\mathbb{R}) := \mathbf{SL}_2(\mathbb{R})/\{\pm I_2\}$, called the *modular group*, operates on $\mathscr{H}$.
   e. Let $M \in \mathbf{SL}_2(\mathbb{R})$ be given. Determine, in terms of $\operatorname{Tr} M$, the number of fixed points of h_M on $\mathscr{H}$.

4. Show that $M \in \mathbf{M}_n(\mathbb{C})$ is normal if and only if there exists a unitary matrix U such that $M^* = MU$.
5. Let $d : \mathbf{M}_n(\mathbb{R}) \to \mathbb{R}^+$ be a multiplicative function; that is,

$$d(MN) = d(M)d(N)$$

for every $M,N \in \mathbf{M}_n(\mathbb{R})$. If $\alpha \in \mathbb{R}$, define $\delta(\alpha) := d(\alpha I_n)^{1/n}$. Assume that d is not constant.

a. Show that $d(0_n) = 0$ and $d(I_n) = 1$. Deduce that $P \in \mathbf{GL}_n(\mathbb{R})$ implies $d(P) \neq 0$ and $d(P^{-1}) = 1/d(P)$. Show, finally, that if M and N are similar, then $d(M) = d(N)$.
b. Let $D \in \mathbf{M}_n(\mathbb{R})$ be a diagonal matrix. Find matrices $D_1, \dots, D_{n-1}$, similar to D, such that $DD_1 \cdots D_{n-1} = (\det D)I_n$. Deduce that $d(D) = \delta(\det D)$.
c. Let $M \in \mathbf{M}_n(\mathbb{R})$ be a diagonalizable matrix. Show that $d(M) = \delta(\det M)$.
d. Using the fact that M^T is similar to M, show that $d(M) = \delta(\det M)$ for every $M \in \mathbf{M}_n(\mathbb{R})$.

6. Let $A \in \mathbf{M}_n(\mathbb{C})$ be given, and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues. Show, by induction on n, that A is normal if and only if

$$\sum_{i,j} |a_{ij}|^2 = \sum_1^n |\lambda_\ell|^2.$$

Hint: The left-hand side (whose square root is called *Schur's norm*) is invariant under conjugation by a unitary matrix. It is then enough to restrict attention to the case of a triangular matrix.

7. (Fiedler and Pták [13]) Given a matrix $A \in \mathbf{M}_n(\mathbb{R})$, we wish to prove the equivalence of the following properties:

P1 For every vector $x \neq 0$ there exists an index k such that $x_k(Ax)_k > 0$.
P2 For every vector $x \neq 0$ there exists a diagonal matrix D with positive diagonal elements such that the scalar product (Ax, Dx) is positive.
P3 For every vector $x \neq 0$ there exists a diagonal matrix D with nonnegative diagonal elements such that the scalar product (Ax, Dx) is positive.
P4 The real eigenvalues of all principal submatrices of A are positive.
P5 All principal minors of A are positive.

a. Prove that **Pj** implies **P(j+1)** for every $j = 1, \dots, 4$.
b. Assume **P5**. Show that for every diagonal matrix D with nonnegative entries, one has $\det(A + D) > 0$.
c. Then prove that **P5** implies **P1**.

8. (Berger) We show here that the numerical radius satisfies the *power inequality*. In what follows, we use the *real part* of a square matrix

$$\Re M := \frac{1}{2}(M + M^*).$$

a. Show that $w(A) \leq 1$ is equivalent to the fact that $\Re(I_n - zA)$ is positive-semidefinite for every complex number z in the open unit disc.
b. We now assume that $w(A) \leq 1$. If $|z| < 1$, verify that $I_n - zA$ is nonsingular.
c. If $M \in \mathbf{GL}_n(\mathbb{C})$ has a nonnegative real part, prove that $\Re\left(M^{-1}\right) \geq 0_n$. Deduce that $\Re(I_n - zA)^{-1} \geq 0_n$ whenever $|z| < 1$.

d. Let $m \geq 1$ be an integer and ω be a primitive mth root of unity in $\mathbb{C}$. Check that the formula

$$\frac{1}{1-X^m} = \frac{1}{m} \sum_{k=0}^{m-1} \frac{1}{1-\omega^k X}$$

can be recast as a polynomial identity.
Deduce that

$$(I_n - z^m A^m)^{-1} = \frac{1}{m} \sum_{k=0}^{m-1} (I_n - \omega^k z A)^{-1},$$

whenever $|z| < 1$.

e. Deduce from above that

$$\Re(I_n - z^m A^m)^{-1} \geq 0_n,$$

whenever $|z| < 1$. Going backward, conclude that for every complex number y in the open unit disc, $\Re(I_n - yA^m) \geq 0_n$ and thus $w(A^m) \leq 1$.

f. Finally, prove the power inequality

$$w(M^m) \leq w(M)^m, \qquad \forall M \in \mathbf{M}_n(\mathbb{C}), \forall m \in \mathbb{N}.$$

Note: A norm that satisfies the power inequality is called a *superstable* norm.

9. Given a complex $n \times n$ matrix A, show that there exists a unitary matrix U such that $M := U^* A U$ has a constant diagonal:

$$m_{ii} = \frac{1}{n} \operatorname{Tr} A, \qquad \forall i = 1, \ldots, n.$$

Hint: Use the convexity of the numerical range.
In the Hermitian case, compare with Schur's theorem 6.7.

10. Let $B \in \mathbf{GL}_n(\mathbb{C})$. Verify that the inverse and the Hermitian adjoint of $B^{-1}B^*$ are similar. Conversely, let $A \in \mathbf{GL}_n(\mathbb{C})$ be a matrix whose inverse and the Hermitian adjoint are similar: $A^* = PA^{-1}P^{-1}$.

a. Show that there exists an invertible Hermitian matrix H such that $H = A^* H A$. **Hint:** Look for an H as a linear combination of P and of P^*.

b. Show that there exists a matrix $B \in \mathbf{GL}_n(\mathbb{C})$ such that $A = B^{-1}B^*$. Look for a B of the form $(aI_n + bA^*)H$.

11. a. Show that $|\det(I_n + A)| \geq 1$ for every skew-Hermitian matrix A, and that equality holds only if $A = 0_n$.

b. Deduce that for every $M \in \mathbf{M}_n(\mathbb{C})$ such that $H := \Re M$ is positive-definite,

$$\det H \leq |\det M|$$

by showing that $H^{-1}(M - M^*)$ is similar to a skew-Hermitian matrix. You may use the *square root* defined in Chapter 10.

12. Let $A \in \mathbf{M}_n(\mathbb{C})$ be a normal matrix. We decompose $A = L + D + U$ in strictly lower, diagonal, and strictly upper-triangular parts. Let us denote by ℓ_j the Euclidean length of the jth column of L, and by u_j that of the jth row of U.

    a. Show that

    $$\sum_{j=1}^{k} u_j^2 \le \sum_{j=1}^{k} l_j^2 + \sum_{j=1}^{k} \sum_{m=1}^{j-1} u_{mj}^2, \quad k = 1, \dots, n-1.$$

    b. Deduce the inequality

    $$\|U\|_S \le \sqrt{n-1}\|L\|_S,$$

    for the Schur–Frobenius norm

    $$\|M\|_S := \left(\sum_{i,j=1}^{n} |m_{ij}|^2 \right)^{1/2}.$$

    c. Prove also that

    $$\|U\|_S \ge \frac{1}{\sqrt{n-1}}\|L\|_S.$$

    d. Verify that each of these inequalities is optimal. **Hint:** Consider a circulant matrix.

13. For $A \in \mathbf{M}_n(\mathbb{C})$, define

    $$\varepsilon := \max_{i \ne j} |a_{ij}|, \quad \delta := \min_{i \ne j} |a_{ii} - a_{jj}|.$$

    We assume in this exercise that $\delta > 0$ and $\varepsilon \le \delta/4n$.

    a. Show that each Gershgorin disk $D_j(A)$ contains exactly one eigenvalue of A.
    b. Let $\rho > 0$ be a real number. Verify that A^ρ, obtained by multiplying the ith row of A by ρ and the ith column by $1/\rho$, has the same eigenvalues as A.
    c. Choose $\rho = 2\varepsilon/\delta$. Show that the ith Gershgorin disk of A^ρ contains exactly one eigenvalue. Deduce that the eigenvalues of A are simple and that

    $$d(\mathrm{Sp}(A), \mathrm{diag}(a_{11}, \dots, a_{nn})) \le \frac{2n\varepsilon^2}{\delta}.$$

14. Let $A \in \mathbf{M}_n(\mathbb{C})$ be given. We define

    $$\mathscr{B}_{ij}(A) = \{ z \in \mathbb{C} \mid |(z - a_{ii})(z - a_{jj})| \le r_i(A) r_j(A) \}.$$

    These sets are *Cassini ovals*. Finally, set

    $$\mathscr{B}(A) := \bigcup_{1 \le i < j \le n} \mathscr{B}_{ij}(A).$$

a. Show that $\mathrm{Sp}A \subset \mathscr{B}(A)$.
b. Show that this result is sharper than Proposition 5.12.
c. When $n = 2$, show that in fact $\mathrm{Sp}A$ is included in the boundary of $\mathscr{B}(A)$.

Note: It is tempting to make a generalization from the present exercise and Proposition 5.12, and conjecture that the spectrum is contained in the union of sets defined by inequalities

$$|(z-a_{ii})(z-a_{jj})(z-a_{kk})| \leq r_i(A)r_j(A)r_k(A)$$

and so on. However, the claim is already false with this third-order version.

15. Let I be an interval of $\mathbb{R}$ and $t \mapsto P(t)$ be a map of class $\mathscr{C}^1$ with values in $\mathbf{M}_n(\mathbb{R})$ such that for each t, $P(t)$ is a projector: $P(t)^2 = P(t)$.

a. Show that the rank of $P(t)$ is constant.
b. Show that $P(t)P'(t)P(t) = 0_n$.
c. Let us define $Q(t) := [P'(t), P(t)]$. Show that $P'(t) = [Q(t), P(t)]$.
d. Let $t_0 \in I$ be given. Show that the differential equation $U' = QU$ possesses a unique solution in I such that $U(t_0) = I_n$. Show that $P(t) = U(t)P(t_0)U(t)^{-1}$.

16. Show that the set of projectors of given rank p is a connected subset in $\mathbf{M}_n(\mathbb{C})$.
17. Let E be an invariant subspace of a matrix $M \in \mathbf{M}_n(\mathbb{R})$.

a. Show that $E^\perp$ is invariant under M^T.
b. Prove the following identity between characteristic polynomials:

$$P_M(X) = P_{M|E}(X)P_{M^T|E^\perp}(X). \tag{5.3}$$

18. Prove Proposition 5.9.
19. (Converse of Lemma 8.) Let A and B be 2×2 complex matrices, that have the same spectrum. We assume in addition that

$$\det[A^*, A] = \det[B^*, B].$$

Prove that A and B are unitarily similar. **Hint:** Prove that they both are unitarily similar to the same triangular matrix.
Deduce that two matrices in $\mathbf{M}_2(\mathbb{C})$ are unitarily similar if and only if they have the same numerical range.

20. Prove the following formula for complex matrices:

$$\log\det(I_n + zA) = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k} \mathrm{Tr}(A^k) z^k.$$

Hint: Use an analogous formula for $\log(1+az)$.