Real matrices and some functional analysis

Marianna Bolla, DSc. Prof. BME Math. Inst.

September 15, 2020

1 Metric, normed vector, and Euclidean spaces

Theory of real matrices will be discussed in the more general framework of linear operators between Hilbert spaces. Linear operators, sometimes between infinite dimensional spaces, will be intensively used throughout the book. For this purpose, we first introduce the notion of some abstract spaces. We suppose that the reader is familiar with the notion of a *vector space* (in other words, linear space). Here we only consider real vector spaces, for the elements of which the operation of addition and the multiplication with a real number is defined together with the usual commutativity, associativity, and distributivity properties; further, there is a distinguished zero element which is, in fact, equal to the real zero times any element of the vector space.

Definition 1 *$(\mathcal{X}, d)$ is a metric space if a metric d is defined on the set $\mathcal{X}$ by the following properties: $\forall x, y, z \in \mathcal{X}$*

- $d(x, y) = d(y, x) \geq 0$,
- $d(x, y) = 0$ *if and only if* $x = y$,
- $d(x, y) \leq d(x, z) + d(z, y)$ *(triangle inequality).*

By means of the metric d, a natural topology can be established via the notion of open and closed sets; further, convergence facts and continuity of real valued functions on $\mathcal{X}$ can be stated. We say that the set $K \subset \mathcal{X}$ is *complete* if every Cauchy sequence in it converges to a limit belonging to K. The metric space $(\mathcal{X}, d)$ is complete if $\mathcal{X}$ itself is complete. The set $K \subset \mathcal{X}$ is *compact* if every (infinite) sequence in K has (at least one) point of accumulation in K, and it is *precompact* if for any $r > 0$ can be covered with finitely many spheres of radius r. Any compact set in a metric space is closed; further, the compactness of the subset K of a metric space is equivalent to any of the following facts:

- K is precompact and complete;
- Any open cover of K has a finite cover. (If we cover K with open sets, then a finite number of them can be selected which also cover K).

In a complete metric space a set is compact if and only if it is precompact and closed. The metric space $(\mathcal{X}, d)$ itself is compact if $\mathcal{X}$ is compact. The Weierstrass theorem states the following.

Theorem 1 *Any continuous function $\mathcal{X} \to \mathbb{R}$ defined on a compact, non-empty metric space is bounded, and it takes on a maximal and minimal value.*

A metric space is called *separable* if it contains a countable dense set.

A *normed vector space* is a vector space endowed with a metric compatible with its linear structure.

Definition 2 *$(\mathcal{X}, \|.\|)$ is a normed vector space if $\mathcal{X}$ is a vector space and the norm $\|\cdot\|$ defined on it satisfies the following conditions: $\forall x, y \in \mathcal{X}$ and $c \in \mathbb{R}$*

- $\|x\| \geq 0$,
- $\|x\| = 0$ *if and only if* $x = 0$,
- $\|cx\| = |c| \cdot \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality).*

A normed vector space is also a metric space with the metric $d(x, y) = \|x - y\|$. A complete normed vector space is called *Banach space*, see [14]. It can be shown that every finite dimensional normed vector space is complete, and hence, it is a Banach space. A subspace of a normed vector space consists of a subset of $\mathcal{X}$ which is a subspace of the original vector space together with the same norm restricted to the subset.

Definition 3 *A real Euclidean space (in other words, inner product space) is a vector space $\mathcal{X}$ endowed with a bilinear function $\langle ., . \rangle$ (called inner product) which satisfies the following requirements: $\forall x, y, z \in \mathcal{X}$ and $a, b \in \mathbb{R}$*

- $\langle x, x \rangle \geq 0$,
- $\langle x, x \rangle = 0$ *if and only if* $x = 0$,
- $\langle x, y \rangle = \langle y, x \rangle$,
- $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$.

We will only consider real Euclidean spaces, and simply call them Euclidean spaces. Any Euclidean space is a normed vector space with the natural norm

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

This norm satisfies the *Cauchy–Schwarz inequality* as follows.

Proposition 1 *For any two elements x, y in a Euclidean space*

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

holds, where equality is attained if and only if x and y are linearly dependent.

If we think of x and y as vectors, and neither of them is the zero element of the underlying Euclidean space, then their linear dependence means that $y = cx$ with some real constant $c \neq 0$. Visually, they are parallel vectors (may be of opposite direction). The meaning of two- or three-dimensional perpendicular vectors can also be generalized. We say that x and y are *orthogonal* if $\langle x, y \rangle = 0$ (the zero element is at the same time orthogonal and parallel to any other

element of the Euclidean space). The notion of angle extends to the following: the *angle* between the non-zero elements x and y is

$$\cos^{-1} \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Further, for given non-zero x, any element y can uniquely be decomposed as $y = y_1 + y_2$, where y_1 is parallel and y_2 is orthogonal to x. Namely,

$$y_1 = \frac{\langle x, y \rangle}{\|x\|^2} x$$

and $y_2 = y - y_1$.

2 Hilbert spaces

A *Hilbert space* is a complete Euclidean space. In fact, every Euclidean space can be completed (in other words, imbedded) into a Hilbert space of which it is a dense subspace. Therefore, Euclidean spaces are also called pre-Hilbert spaces. This distinction between a Hilbert and pre-Hilbert space makes sense only in the infinite dimensional case; every finite dimensional Euclidean space is a Hilbert space, and of course a normed vector space, Banach space, and metric space, at the same time. Whenever we write $\mathbb{R}^n$ we think of the n-dimensional real vector space in all of these senses. Every subspace of a Euclidean space is again a Euclidean space with the induced inner product, while every closed subspace of a Hilbert space is a Hilbert space.

It can be shown that all separable Hilbert spaces are isomorphic as they are, in fact, L^2 spaces. We can think of them as countably infinite dimensional vectors. Let us see some examples.

(a) Countably infinite dimensional vectors $x = (x_1, x_2, \dots)$ form a so-called l^2 space if $\sum_{i=1}^{\infty} x_i^2 < \infty$. In this case $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i$ (which is finite by the Cauchy–Schwarz inequality) and $\|x\| = \sqrt{\sum_{i=1}^{\infty} x_i^2}$.

(b) The set $C([a, b])$ of continuous, real valued functions on the closed interval $[a, b]$ form a Euclidean space with the inner product

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx, \quad f, g \in C([a, b].$$

However, this space is not complete, therefore $C([a, b])$ is not a Hilbert space, see the example on page 6 of [5].

(c) The $L^2(\mathcal{X})$ space of real-valued, square-integrable functions with respect to some finite measure μ on a compact set $\mathcal{X}$ is a separable Hilbert space with the inner product

$$\langle f, g \rangle = \int_{\mathcal{X}} f(x) g(x)\, d\mu(x), \quad f, g \in L^2(\mathcal{X}).$$

The induced norm is $\|f\| = \sqrt{\int_{\mathcal{X}} f^2(x)\, d\mu(x)}$.

It can be shown that if μ is the Lebesgue measure on $\mathbb{R}$, then $L^2([a, b])$ is the completion of the pre-Hilbert space $C([a, b])$ into a Hilbert space.

The separability of a Hilbert space $\mathcal{H}$ means that – as a vector space – it is spanned by countably many elements of it, the linear independence of which can be assumed (such a set is called a basis); or else – by virtue of the Schmidt orthogonalization procedure – there exists an *orthonormal basis* $b_1, b_2, \dots \in \mathcal{H}$ such that $\langle b_i, b_j \rangle = \delta_{ij}$ $(\forall i, j)$, where δ_{ij} is the Kronecker-delta. For example, in (a), the infinite dimensional vectors $e_1, e_2, \dots$ form an orthonormal basis if the ith component of e_i is 1, and the others are all zeros (this is the canonical basis); in (c), there are many orthonormal bases, they can be constructed by means of orthogonal systems of polynomials or trigonometric functions.

The linearity of a Hilbert space (which is defined over $\mathbb{R}$) is somehow compatible with the set of real-valued linear functions – called functionals – on it. A *linear functional* F on the Hilbert space $\mathcal{H}$ assigns the real number $F(x)$ to any $x \in \mathcal{H}$ and this assignment is linear. The continuity (in the usual sense) of such an F is equivalent to its boundedness, i.e. there is a constant C such that

$$|F(x)| \le C\|x\|, \quad \forall x \in \mathcal{H}.$$

The *dual* of a Hilbert space consists of the continuous linear functionals on it, which also form a Hilbert space. The famous Riesz–Fréchet representation theorem, published by the two authors in the same issue of the C. R. Acad. Sci. Paris (see [8, 3]), states the following.

Theorem 2 (Riesz–Fréchet representation theorem) *A Hilbert space and its dual are isomorphic. Namely, the effect of any $F : \mathcal{H} \to \mathbb{R}$ continuous linear functional can be written as*

$$F(x) = \langle x, y \rangle, \quad \forall x \in \mathcal{H}$$

with an appropriate $y \in \mathcal{H}$. Moreover, y is uniquely determined by F.

In the sequel, we will investigate linear maps between two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ which are bounded, i.e. there is a constant $C \ge 0$ such that

$$\|Ax\| \le C\|x\|, \quad \forall x \in \mathcal{H}. \tag{1}$$

The $A : \mathcal{H} \to \mathcal{H}'$ bounded linear maps are called *operators*, and the boundedness of them also implies their continuity in the usual sense. The *operator norm* of A is the smallest bound C in (1), and by the linearity, it has the equivalent form

$$\|A\| = \sup_{\|x\|_{\mathcal{H}}=1} \|Ax\|_{\mathcal{H}'}.$$

We will be mainly interested in *completely continuous* (with other words, *compact*) operators. We will see that they can be well approximated by means of finite rank ones, hence the important facts about them are reminiscent of those of linear algebra. Here the *rank* of an operator is the dimension of its range (image space) in $\mathcal{H}'$.

Definition 4 *The operator $A : \mathcal{H} \to \mathcal{H}'$ is said to be compact or completely continuous if it maps bounded sets of $\mathcal{H}$ into precompact sets of $\mathcal{H}'$; or equivalently, for any bounded sequence $(x_n) \subset \mathcal{H}$, the image sequence $(Ax_n) \subset \mathcal{H}'$ has a convergent subsequence.*

Frequently, $\mathcal{H} = \mathcal{H}'$, and in this situation, we can define the *identity* operator I such that $Ix = x$, $\forall x \in \mathcal{H}$. The identity operator in an infinite dimensional Hilbert space is not compact, since the sequence (e_n) of an orthonormal basis is bounded, still, it does not have a convergent subsequence.

Any operator between finite dimensional Hilbert spaces is compact, and it can be identified with a matrix as we will see in the linear algebra part. Analogously, an operator between the separable Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ is described with a special infinite matrix in the following way. Let $(e_i) \subset \mathcal{H}$ and $(f_j) \subset \mathcal{H}'$ be canonical bases and for the real numbers a_{ij} the condition

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}^2 < \infty \tag{2}$$

hold. Then the formula

$$A\left(\sum_{i=1}^{\infty} c_i e_i\right) = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} a_{ji} c_i\right) f_j$$

defines a compact operator $A : \mathcal{H} \to \mathcal{H}'$. Such an operator is called *Hilbert–Schmidt operator*. Soon, we will see that there are compact operators which are not of Hilbert–Schmidt type.

Definition 5 *The adjoint A^* of the operator $A : \mathcal{H} \to \mathcal{H}'$ is defined by*

$$\langle Ax, y\rangle_{\mathcal{H}'} = \langle x, A^* y\rangle_{\mathcal{H}}, \quad \forall x \in \mathcal{H},\, y \in \mathcal{H}'.$$

Therefore, A^* is a $\mathcal{H}' \to \mathcal{H}$ operator, and if necessary, we will denote in the lower index that in which Hilbert space the inner product is understood. The inner product $\langle Ax, y\rangle_{\mathcal{H}}$ is called *bilinear form* as it is linear (and also continuous) in both variables x and y.

Definition 6 *The operator $A : \mathcal{H} \to \mathcal{H}$ is said to be self-adjoint if $A = A^*$.*

Recall that we only deal with Hilbert spaces over $\mathbb{R}$, therefore we sometimes call a self-adjoint operator symmetric, especially in the finite rank case.

Compact operators can be decomposed as weighted sums of special rank one operators, where the weights tend to zero, and therefore keeping the most important ones, we obtain low rank approximation of them. More precisely, the following statements are true.

The *spectral theorem of Hilbert and Schmidt* states a very important property of compact, self-adjoint operators in a separable Hilbert space.

Theorem 3 *Let $\mathcal{H}$ be a separable Hilbert space and $A : \mathcal{H} \to \mathcal{H}$ be a compact, self-adjoint operator. Then there exist an orthonormal basis $(\psi_i) \subset \mathcal{H}$ and a sequence (λ_i) of real numbers such that*

$$A\psi_i = \lambda_i \psi_i, \quad i = 1, 2, \ldots. \tag{3}$$

Further, if $\mathcal{H}$ has infinite dimension, then $\lim_{i\to\infty} \lambda_i = 0$.

The element $\psi_i \in \mathcal{H}$ in (3) is called *eigenfunction* corresponding to the *eigenvalue* λ_i $(i = 1, 2, \dots)$. If there are no multiple eigenvalues then the unit-norm eigenfunctions are unique up to orientation (multiplication with ± 1). To multiple eigenvalues there corresponds a unique eigenspace of dimension equal to the multiplicity of this eigenvalue. Within the eigenspace any orthonormal system can embody the eigenfunctions corresponding to the equal eigenvalues.

The Hilbert–Schmidt theorem gives rise to the following *spectral decomposition* (briefly SD) of the self-adjoint, compact operator A:

$$A = \sum_{i=1}^{\infty} \lambda_i \langle ., \psi_i \rangle \psi_i. \tag{4}$$

Consequently, the effect of A on $x \in \mathcal{H}$ can be written as

$$Ax = \sum_{i=1}^{\infty} \lambda_i \langle x, \psi_i \rangle \psi_i$$

meaning the pointwise convergence of the sequence $\sum_{i=1}^{n} \lambda_i \langle x, \psi_i \rangle \psi_i$. This implies the convergence of (4) in operator norm too. That is, with the notation $A_n = \sum_{i=1}^{n} \lambda_i \langle ., \psi_i \rangle \psi_i$, the relation $\|A - A_n\| \to 0$ also holds as $n \to \infty$. Thus, in the self-adjoint case, we have shown that compact operators can be approached by finite rank ones.

Note that the spectral theorem naturally extends to non-compact, self-adjoint $\mathcal{H} \to \mathcal{H}$ operators which can be decomposed as $A + cI$, where A is compact, self-adjoint and c is a non-zero constant. Then the eigenvalues of $A + cI$ are the numbers $\lambda_i + c$ with the eigenfunctions ψ_i $(i = 1, 2, \dots)$, and the spectrum of $A + cI$ converges to the constant c.

The whole machinery can be extended to general compact operators as follows.

Theorem 4 *Let $\mathcal{H}$ and $\mathcal{H}'$ be separable Hilbert spaces and $A : \mathcal{H} \to \mathcal{H}'$ be a compact operator. Then there exist orthonormal bases $(\psi_i) \subset \mathcal{H}$ and $(\phi_i) \subset \mathcal{H}'$ together with a sequence (s_i) of nonnegative real numbers such that*

$$A\psi_i = s_i \phi_i, \quad A^* \phi_i = s_i \psi_i, \quad i = 1, 2, \dots. \tag{5}$$

Further, if $\mathcal{H}$ and $\mathcal{H}'$ have infinite dimension, then $\lim_{i \to \infty} s_i = 0$.

The elements $\psi_i \in \mathcal{H}$ and $\phi_i \in \mathcal{H}'$ in (5) are called *singular function pairs* (or left and right singular functions) corresponding to the *singular value* s_i $(i = 1, 2, \dots)$. About the uniqueness the same can be said as in case of the SD. We remark that the singular values of a self-adjoint operator are the absolute values of its eigenvalues. In case of a positive eigenvalue, the left and right singular functions are the same (they coincide with the corresponding eigenfunction with any, but the same orientation). In case of a negative eigenvalue, the left and right singular functions are opposite (any of them can be the corresponding eigenfunction of bivalent orientation). In case of a zero singular value the orientation is immaterial as it does not contribute to the decomposition of the underlying operator.

The above theorem gives rise to the following *singular value decomposition* (briefly, SVD) of the compact operators A and A^*:

$$A = \sum_{i=1}^{\infty} s_i \langle ., \psi_i \rangle_{\mathcal{H}} \phi_i, \quad A^* = \sum_{i=1}^{\infty} s_i \langle ., \phi_i \rangle_{\mathcal{H}'} \psi_i. \tag{6}$$

Consequently, the effect of A on $x \in \mathcal{H}$ can be written as

$$Ax = \sum_{i=1}^{\infty} \lambda_i \langle x, \psi_i \rangle \phi_i,$$

while the effect of A^* on $y \in \mathcal{H}'$ as

$$A^* y = \sum_{i=1}^{\infty} \lambda_i \langle y, \phi_i \rangle_{\mathcal{H}'} \phi_i.$$

The convergence in (6) is also meant in spectral norm. That is, with the notation $A_n = \sum_{i=1}^{n} s_i \langle ., \psi_i \rangle_{\mathcal{H}} \phi_i$, the convergence fact $\|A - A_n\| \to 0$ gives rise to a finite rank approximation of a general compact operator.

For a Hilbert–Schmidt operator, $\sum_{i=1}^{\infty} s_i^2 < \infty$ (also $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ if it is self-adjoint) in accord with the requirement (2). These convergences imply that $\lim_{i\to\infty} s_i = 0$ (also $\lim_{i\to\infty} \lambda_i = 0$ in the self-adjoint case). Therefore a Hilbert-Schmidt operator is always compact, but the converse is not true: if only $\lim_{i\to\infty} s_i = 0$ holds and $\sum_{i=1}^{\infty} s_i^2 = \infty$, then our operator is compact but not Hilbert–Schmidt.

It can be shown that for a general compact operator with SVD in (6)

$$\|A\| = \max_i s_i$$

and for a self-adjoint one with SD in (4)

$$\|A\| = \max_i |\lambda_i|.$$

Therefore, the operator norm of compact operators is also called *spectral norm.*

We will frequently use the following propositions for compact operators.

Proposition 2 *Let $A : \mathcal{H} \to \mathcal{H}'$ be a compact operator with SVD in (6). Assume that its singular values are enumerated in non-increasing order ($s_1 \geq s_2 \geq \dots$). Then*

$$\max_{\|x\|_{\mathcal{H}}=1,\, \|y\|_{\mathcal{H}'}=1} \langle Ax, y \rangle = s_1$$

and it is attained with the choice $x = \psi_1$ and $y = \phi_1$ (uniquely if $s_1 > s_2$). This was the $k = 1$ case. Further, for $k = 2, 3, \dots$

$$\max_{\substack{\|x\|_{\mathcal{H}}=1,\, \|y\|_{\mathcal{H}'}=1 \\ \langle x, \psi_i \rangle_{\mathcal{H}} = 0 \,(i=1,\dots,k-1) \\ \langle y, \phi_i \rangle_{\mathcal{H}'} = 0 \,(i=1,\dots,k-1)}} \langle Ax, y \rangle = s_k$$

and it is attained with the choice $x = \psi_k$ and $y = \phi_k$ (uniquely if $s_k > s_{k+1}$).

A bit more is true.

Proposition 3 *Let $A : \mathcal{H} \to \mathcal{H}'$ be a compact operator with SVD in (6). Assume that its singular values are enumerated in non-increasing order and $k > 0$ is an integer such that $s_k > s_{k+1}$. Then*

$$\max_{\substack{\langle x_i, x_j\rangle_{\mathcal{H}} = \delta_{ij} \\ \langle y_i, y_j\rangle_{\mathcal{H}'} = \delta_{ij}}} \sum_{i=1}^{k} \langle Ax_i, y_i\rangle = \sum_{i=1}^{k} s_i$$

and it is attained with the choice $x_i = \psi_i$ and $y_i = \phi_i$ $(i = 1, \dots, k)$.

Proposition 4 *Let $A : \mathcal{H} \to \mathcal{H}$ be a self-adjoint compact operator with SD in (4). Assume that its eigenvalues are enumerated in non-increasing order $(\lambda_1 \geq \lambda_2 \geq \dots)$. Then*

$$\max_{\|x\|=1} \langle Ax, x\rangle = \lambda_1$$

and it is attained with the choice $x = \psi_1$ (uniquely if $\lambda_1 > \lambda_2$). This was the $k = 1$ case. Further, for $k = 2, 3, \dots$

$$\max_{\substack{\|x\|=1 \\ \langle x, \psi_i\rangle = 0 \, (i=1,\dots,k-1)}} \langle Ax, x\rangle = \lambda_k$$

and it is attained with the choice $x = \psi_k$ (uniquely if $\lambda_k > \lambda_{k+1}$).

A bit more is true.

Proposition 5 *Let $A : \mathcal{H} \to \mathcal{H}$ be a self-adjoint compact operator with SD in (4). Assume that its eigenvalues are enumerated in non-increasing order and $k > 0$ is an integer such that $\lambda_k > \lambda_{k+1}$. Then*

$$\max_{\langle x_i, x_j\rangle = \delta_{ij}} \sum_{i=1}^{k} \langle Ax_i, x_i\rangle = \sum_{i=1}^{k} \lambda_i$$

and it is attained with the choice $x_i = \psi_i$ $(i = 1, \dots, k)$.

Integral operators between L^2 spaces (see [9, 10]) are important examples of Hilbert–Schmidt operators. In Example (c) we introduced the $\mathcal{H} = L^2(\mathcal{X})$ space of real-valued, square-integrable functions with respect to some finite measure μ on the compact set $\mathcal{X}$. Likewise, consider $\mathcal{H}' = L^2(\mathcal{Y})$ with the finite measure ν, and let $K : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a *kernel* such that for it

$$\int_{\mathcal{X}} \int_{\mathcal{Y}} K^2(x, y)\, \mu(dx)\, \nu(dy) < \infty \tag{7}$$

holds. With the kernel K, the *integral operator* $A_K : H' \to H$ is defined in the following way: to the function $g \in \mathcal{H}'$ the operator A_K assigns the function $f \in \mathcal{H}$ such that

$$f(x) = (A_K g)(x) = \int_{\mathcal{Y}} K(x, y) g(y)\, \nu(dy), \quad x \in \mathcal{X}.$$

It is easy to see that A_K is linear, further it is a Hilbert–Schmidt operator, therefore compact with

$$\|A_K\| \leq \|K\|_2 \tag{8}$$

where $\|K\|_2$ is the squareroot of the finite expression in (7), that is the L^2-norm of K in the product space.

Say, A_K has the SVD

$$A_K = \sum_{i=1}^{\infty} s_i \langle ., \phi_i \rangle_{H'} \psi_i \tag{9}$$

where for the singular values

$$\sum_{i=1}^{\infty} s_i^2 = \|K\|_2^2 < \infty$$

also holds, implying that $\lim_{i\to\infty} s_i = 0$ (if s_i's are really countably infinitely many).

It is easy to see that the adjoint of A_K is the integral operator $A_K^* : \mathcal{H} \to \mathcal{H}'$ with SVD

$$A_K^* = \sum_{i=1}^{\infty} s_i \langle ., \psi_i \rangle_H \phi_i,$$

and

$$\|A_K\| = \|A_K^*\| = s_1.$$

Remark that $\|K\|_2 = (\sum_{i=1}^{\infty} s_i^2)^{1/2}$ is called the *Hilbert–Schmidt norm* of A_K (in finite dimension it will be called Frobenius norm) denoted by $\|A_K\|_2$, and for it, $\|A_K\|_2 \geq \|A_K\|$ due to Inequality (8).

With a *symmetric kernel* $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ ($K(x, y) = K(y, x)\ \forall x, y$), the induced integral operator A_K becomes self-adjoint, therefore it also admits an SD

$$A_K = \sum_{i=1}^{\infty} \lambda_i \langle ., \psi_i \rangle \psi_i$$

with real eigenvalues such that $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ and corresponding orthonormal eigenvectors $\psi_1, \psi_2, \ldots$. Under some additional condition, the kernel itself can be expanded, see the upcoming Mercer theorem.

Definition 7 *The integral operator A_K corresponding to the symmetric kernel K is said to be positive semidefinite if it has all nonnegative eigenvalues.*

The SD of the self-adjoint A_K gives rise to the following, so-called Karhunen–Loève expansion of K: $K(x, y) = \sum_{i=1}^{\infty} \lambda_i \psi_i(x) \psi_i(y)$, where the convergence is understood in the L^2-norm of the product space. The following theorem states more, namely, uniform pointwise convergence.

Theorem 5 (Mercer theorem) *If K is a symmetric, continuous kernel of a positive semidefinite integral operator on $L^2(\mathcal{X})$, where $\mathcal{X}$ is some compact space, then it can be expanded into the following uniformly convergent series:*

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \psi_i(x) \psi_i(y), \quad \forall x, y \in \mathcal{X}$$

by the eigenfunctions ψ_i and the eigenvalues $\lambda_i \geq 0$ of the integral operator induced by the kernel K.

Finally, we will use the following *uniform boundedness principle* concerning a collection of linear operators between Hilbert spaces. In fact, the forthcoming Banach–Steinhaus theorem is stated more generally, for normed vector spaces.

Theorem 6 (Banach–Steinhaus theorem) *Let $\mathcal{X}$ be a Banach space and $\mathcal{Y}$ a normed vector space; further, let $\mathcal{A}$ be a collection of $\mathcal{X} \to \mathcal{Y}$ continuous linear operators. If for all $x \in \mathcal{X}$ we have*

$$\sup_{A \in \mathcal{A}} \|Ax\| < \infty,$$

then

$$\sup_{A \in \mathcal{A}} \|A\| < \infty$$

too.

3 Matrices

From now on, we will confine ourselves to finite dimensional real Euclidean spaces which are also Hilbert spaces. Linear operations between these spaces can be described by matrices of real entries. To stress that the elements of the Euclidean space are finite dimensional vectors, we will use bold-face lower-case letters, further vectors are treated as column-vectors. The inner product of the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is therefore written with matrix multiplication, like $\mathbf{x}^T\mathbf{y}$, where ^{T} stands for the transposition, hence $\mathbf{x}^T$ is a row-vector. Matrices will be denoted by bold-face upper-case letters. An $m \times n$ matrix $\boldsymbol{A} = (a_{ij})$ of real entries a_{ij}'s corresponds to an $\mathbb{R}^n \to \mathbb{R}^m$ linear transformation (operator). Its transpose, $\boldsymbol{A}^T$, is an $n \times m$ matrix. An $n \times n$ matrix is called quadratic and it maps $\mathbb{R}^n$ into itself. The identity matrix is denoted by $\boldsymbol{I}$ or $\boldsymbol{I}_n$ if we want to refer to its size.

The quadratic matrix $\boldsymbol{A}$ is *symmetric* if $\boldsymbol{A} = \boldsymbol{A}^T$ and *orthogonal* if $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}$. The orthogonal matrix $\boldsymbol{P}$ is a *permutation matrix* if, in each row and column, exactly one of its entries differs from 0, and this non-zero entry is 1.

The $n \times n$ matrix $\boldsymbol{A}$ has an inverse if and only if its determinant, $|\boldsymbol{A}| \neq 0$, and its inverse is denoted by $\boldsymbol{A}^{-1}$. In this case, the linear transformation corresponding to $\boldsymbol{A}^{-1}$ undoes the effect of the $\mathbb{R}^n \to \mathbb{R}^n$ transformation corresponding to $\boldsymbol{A}$, i.e. $\boldsymbol{A}^{-1}\mathbf{y} = \mathbf{x}$ if and only if $\boldsymbol{A}\mathbf{x} = \mathbf{y}$ for any $\mathbf{y} \in \mathbb{R}^n$. It is important that in case of an invertible (*regular*) matrix $\boldsymbol{A}$, the *range* (or image space) of $\boldsymbol{A}$ – denoted by $\mathcal{R}(\boldsymbol{A})$ – is the whole $\mathbb{R}^n$, and in exchange, the kernel of $\boldsymbol{A}$ (the subspace of vectors that are mapped into the zero vector by $\boldsymbol{A}$) consists of the only $\mathbf{0}$.

Note that for an $m \times n$ matrix $\boldsymbol{A}$, its range is

$$\mathcal{R}(\boldsymbol{A}) = \mathrm{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$$

where $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the column vectors of $\boldsymbol{A}$ for which fact the notation $\boldsymbol{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ will be used; further, $\mathrm{Span}\{\ldots\}$ is the subspace spanned by the vectors in its argument. The *rank* of $\boldsymbol{A}$ is the dimension of its range:

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{dim}\mathcal{R}(\boldsymbol{A}),$$

and it is also equal to the maximum number of linearly independent rows of $\boldsymbol{A}$; trivially, $\mathrm{rank}(\boldsymbol{A}) \leq \min\{m, n\}$. In case of $m = n$, $\boldsymbol{A}$ is regular if and only if $\mathrm{rank}(\boldsymbol{A}(= n$, and *singular*, otherwise.

An orthogonal matrix $\boldsymbol{A}$ is always regular and $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$; further its rows (or columns) constitute a complete orthonormal set in $\mathbb{R}^n$. Let k $(1 \leq k < n)$ be an integer; an $n \times k$ matrix $\boldsymbol{A}$ is called *suborthogonal* if its columns form (a not complete) orthonormal set in $\mathbb{R}^n$. For such an $\boldsymbol{A}$, the relation $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_k$ holds, but $\boldsymbol{A}\boldsymbol{A}^T \neq \boldsymbol{I}_n$. In fact, the $n \times n$ matrix $\boldsymbol{P} = \boldsymbol{A}\boldsymbol{A}^T$ is symmetric and idempotent ($\boldsymbol{P}^2 = \boldsymbol{P}$), hence, it corresponds to the orthogonal projection onto $\mathcal{R}(\boldsymbol{A})$. The *trace* of the $n \times n$ matrix $\boldsymbol{A}$ is

$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}.$$

How the above matrix–matrix and matrix–scalar functions will look like if the underlying matrix is a product? If $\boldsymbol{A}$ and $\boldsymbol{B}$ can be multiplied together ($\boldsymbol{A}$ is $m \times n$ and $\boldsymbol{B}$ is $n \times k$ type), then their product corresponds to the succession of linear operations $\boldsymbol{B}$ and $\boldsymbol{A}$ in this order, therefore

$$(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$$

and if $\boldsymbol{A}$ and $\boldsymbol{B}$ are regular $n \times n$ matrices, then so is $\boldsymbol{A}\boldsymbol{B}$, and

$$(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}.$$

Further, $(\boldsymbol{A}^{-1})^T = (\boldsymbol{A}^T)^{-1}$, and vice versa. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are $n \times n$ matrices, then

$$|\boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|.$$

Therefore, the determinant of the product of several matrices of the same size does not depend on the succession of the matrices, however, the matrix multiplication is usually not commutative. The trace is commutative in the following sense: if $\boldsymbol{A}$ is an $n \times k$ and $\boldsymbol{B}$ is a $k \times n$ matrix, then

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}).$$

For several factors, the trace is accordingly, cyclically commutative:

$$\mathrm{tr}(\boldsymbol{A}_1\boldsymbol{A}_2 \ldots \boldsymbol{A}_n) = \mathrm{tr}(\boldsymbol{A}_2 \ldots \boldsymbol{A}_n\boldsymbol{A}_1) = \cdots = \mathrm{tr}(\boldsymbol{A}_n\boldsymbol{A}_1 \ldots \boldsymbol{A}_{n-1})$$

when, of course, the sizes of the factors are such that the successive multiplications in $\boldsymbol{A}_1 \ldots \boldsymbol{A}_n$ can be performed and the number of rows in $\boldsymbol{A}_1$ is equal to the number of columns in $\boldsymbol{A}_n$. Further,

$$\mathrm{rank}(\boldsymbol{A}\boldsymbol{B}) \leq \min\{\mathrm{rank}(\boldsymbol{A}), \mathrm{rank}(\boldsymbol{B})\},$$

consequently, the rank cannot be increased in course of matrix multiplications.

Given an $n \times n$ symmetric real matrix $\boldsymbol{A}$, the quadratic form in the variables $x_1, \ldots, x_n$ is the homogeneous quadratic function of these variables:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = \mathbf{x}^T\boldsymbol{A}\mathbf{x},$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$, hence the matrix multiplication results in a scalar. The possible signs of a quadratic form (with different $\mathbf{x}$'s) characterize the underlying matrix. Accordingly, they fall into exactly one of the following categories.

Definition 8 *Let $\boldsymbol{A}$ be $n \times n$ symmetric real matrix.*

- *$\boldsymbol{A}$ is positive (negative) definite if $\mathbf{x}^T\boldsymbol{A}\mathbf{x} > 0$ ($\mathbf{x}^T\boldsymbol{A}\mathbf{x} < 0$), $\forall \mathbf{x} \neq \mathbf{0}$.*
- *$\boldsymbol{A}$ is positive (negative) semidefinite if $\mathbf{x}^T\boldsymbol{A}\mathbf{x} \geq 0$ ($\mathbf{x}^T\boldsymbol{A}\mathbf{x} \leq 0$), $\forall \mathbf{x} \in \mathbb{R}^n$, and $\mathbf{x}^T\boldsymbol{A}\mathbf{x} = \mathbf{0}$ for at least one $\mathbf{x} \neq \mathbf{0}$.*
- *$\boldsymbol{A}$ is indefinite if $\mathbf{x}^T\boldsymbol{A}\mathbf{x}$ takes on both positive and negative values (with different, non-zero $\mathbf{x}$'s).*

The positive and negative definite matrices are all regular, whereas the positive and negative semidefinite ones are singular. The indefinite matrices can be either regular or singular. To more easily characterize the definiteness of symmetric matrices, we will use their eigenvalues.

In analogy with the spectral theory of compact operators, the notion of an eigenvalue and eigenvector is introduced: λ is an eigenvalue of the $n \times n$ real matrix $\boldsymbol{A}$ with corresponding eigenvector $\mathbf{u} \neq \mathbf{0}$ if $\boldsymbol{A}\mathbf{u} = \lambda\mathbf{u}$. If $\mathbf{u}$ is an eigenvector of $\boldsymbol{A}$, it is easy to see that for $c \neq 0$, $c\mathbf{u}$ is also an eigenvector with the same eigenvalue. Therefore, it is better to speak about *eigen-direction*s instead of eigenvectors; or else, we will consider specially normalized, e.g. unit-norm eigenvectors, when only the orientation is bivalent. It is well known that an $n \times n$ matrix $\boldsymbol{A}$ has exactly n eigenvalues (with multiplicities) which are (possibly complex) roots of the characteristic polynomial $|\boldsymbol{A} - \lambda\boldsymbol{I}|$. Knowing the eigenvalues, the corresponding eigenvectors are obtained by solving the system of linear equations $(\boldsymbol{A} - \lambda\boldsymbol{I})\mathbf{u} = \mathbf{0}$ which must have a non-trivial solution due to the choice of λ. In fact, there are infinitely many solutions (in case of single eigenvalues they are constant multiples of each other). An eigenvector corresponding to a complex eigenvalue must also have complex coordinates, but in case of our main interest (the symmetric matrices) this cannot occur.

The notion of an eigenvalue and eigenvector extends to matrices of complex entries in the same way. As for the allocation of the eigenvalues of a quadratic matrix (even of complex entries), the following result is known.

Theorem 7 (Gersgorin disc theorem) *Let $\boldsymbol{A}$ be an $n \times n$ matrix of entries $a_{ij} \in \mathbb{C}$. The Gersgorin disks of $\boldsymbol{A}$ are the following regions of the complex plane:*

$$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|\}, \quad i = 1, \ldots n.$$

Let $\lambda_1, \ldots, \lambda_n$ denote the (possibly complex) eigenvalues of $\boldsymbol{A}$. Then

$$\{\lambda_1, \ldots, \lambda_n\} \subset \cup_{i=1}^n D_i.$$

Furthermore, any connected component of the set $\cup_{i=1}^n D_i$ contains as many eigenvalues of $\boldsymbol{A}$ as the number of discs that form this component.

We will introduce the notion of *normal matrices* which admit a spectral decomposition (briefly, SD) similar to that of compact operators. The real matrix $\boldsymbol{A}$ is called normal if $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$. Among real matrices, only the symmetric, anti-symmetric ($\boldsymbol{A}^T = -\boldsymbol{A}$), and orthogonal matrices are normal. Normal matrices have the following important spectral property: to their eigenvalues there corresponds an orthonormal set of eigenvectors; choosing this as a new

basis, the matrix becomes *diagonal* (all the off-diagonal entries are zeros). Here we only state the analogous version of the Hilbert–Schmidt theorem 3 for symmetric matrices which, in addition, have all real eigenvalues, and consequently, eigenvectors of real coordinates.

Theorem 8 *The $n \times n$ symmetric, real matrix $\boldsymbol{A}$ has real eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ (with multiplicities), and the corresponding eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ can be chosen such that they constitute a complete orthonormal set in $\mathbb{R}^n$.*

This so-called Spectral Decomposition theorem implies the following SD of the $n \times n$ symmetric matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^T, \tag{10}$$

where $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix containing the eigenvalues – called *spectrum* – in its main diagonal, while $\boldsymbol{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ is the orthogonal matrix containing the corresponding eigenvectors of $\boldsymbol{A}$ in its columns in the order of the eigenvalues. Of course, permuting the eigenvalues in the main diagonal of $\boldsymbol{\Lambda}$, and the columns of $\boldsymbol{U}$ accordingly, will lead to the same SD, however – if not otherwise stated – we will enumerate the real eigenvalues in non-increasing order. About the uniqueness of the above SD we can state the following: the unit-norm eigenvector corresponding to a single eigenvalue is unique (up to orientation), whereas to an eigenvalue with multiplicity m there corresponds a unique m-dimensional so-called *eigen-subspace* within which any orthonormal set can be chosen for the corresponding eigenvectors.

It is easy to verify that for the eigenvalues of the symmetric matrix $\boldsymbol{A}$

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(\boldsymbol{A}) \quad \text{and} \quad \prod_{i=1}^{n} \lambda_i = |\boldsymbol{A}|$$

hold. Therefore $\boldsymbol{A}$ is singular if and only if it has a 0 eigenvalue, and

$$r = \operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{\Lambda}) = |\{i \,:\, \lambda_i \neq 0\}|;$$

moreover, $\mathcal{R}(\boldsymbol{A}) = \operatorname{Span}\{\mathbf{u}_i \,:\, \lambda_i \neq 0\}$. Therefore, the SD of $\boldsymbol{A}$ simplifies to

$$\sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

Its spectrum also determines the definiteness of $\boldsymbol{A}$ in the following manner.

Proposition 6 *Let $\boldsymbol{A}$ be $n \times n$ symmetric real matrix.*

- *$\boldsymbol{A}$ is positive (negative) definite if and only if all of its eigenvalues are positive (negative).*
- *$\boldsymbol{A}$ is positive (negative) semidefinite if and only if all of its eigenvalues are nonnegative (nonpositive), and its spectrum includes the zero.*
- *$\boldsymbol{A}$ is indefinite if its spectrum contains at least one positive and one negative eigenvalue.*

The matrix of an orthogonal projection $\boldsymbol{P}_F$ onto the r-dimensional subspace $F \subset \mathbb{R}^n$ has the following SD (only the $r < n$ case is of importance, since in the $r = n$ case $\boldsymbol{P}_F = \boldsymbol{I}_n$):

$$\boldsymbol{P}_F = \sum_{i=1}^{r} \mathbf{u}_i \mathbf{u}_i^T = \boldsymbol{A}\boldsymbol{A}^T,$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_r$ is any orthonormal set in F which is the eigen-subspace corresponding to the eigenvalue 1 of multiplicity r. Note that the eigenspace corresponding to the other eigenvalue 0 of multiplicity $n - r$ is the orthogonal complementary subspace $F^\perp$ of F in $\mathbb{R}^n$, but it has no importance, as only the eigenvectors in the first r columns of $\boldsymbol{U}$ enter into the above SD of $\boldsymbol{P}_F$. With the notation $\boldsymbol{A} = (\mathbf{u}_1, \ldots, \mathbf{u}_r)$, the SD of $\boldsymbol{P}_F$ simplifies to $\boldsymbol{A}\boldsymbol{A}^T$, indicating that $\boldsymbol{A}$ is a suborthogonal matrix.

Now the analogue of Theorem 4 for rectangular matrices is formulated.

Theorem 9 *Let $\boldsymbol{A}$ be an $m \times n$ rectangular matrix of real entries,* $\operatorname{rank}(\boldsymbol{A}) = r \leq \min\{m, n\}$. *Then there exist an orthonormal set $(\mathbf{v}_1, \ldots, \mathbf{v}_r) \subset \mathbb{R}^m$ and $(\mathbf{u}_1, \ldots, \mathbf{u}_r) \subset \mathbb{R}^n$ together with the positive real numbers $s_1 \geq s_2 \geq \cdots \geq s_r > 0$ such that*

$$\boldsymbol{A}\mathbf{u}_i = s_i \mathbf{v}_i, \quad \boldsymbol{A}^* \mathbf{v}_i = s_i \mathbf{u}_i, \quad i = 1, 2, \ldots, r. \tag{11}$$

The elements $\mathbf{v}_i \in \mathbb{R}^m$ and $\mathbf{u}_i \in \mathbb{R}^n$ $(i = 1, \ldots, r)$ in (11) are called *relevant singular vector pairs* (or left and right singular vectors) corresponding to the *singular value* s_i $(i = 1, 2, \ldots, r)$. The transformations in (11) give a one-to one mapping between $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\boldsymbol{A}^T)$, all the other vectors of $\mathbb{R}^n$ and $\mathbb{R}^m$ are mapped into the zero vector of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. However, the left and right singular vectors can appropriately be completed into a complete orthonormal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset \mathbb{R}^m$ and $\{\mathbf{u}_1, \ldots \mathbf{u}_n\} \subset \mathbb{R}^n$, respectively, such that, the so introduced extra vectors in the kernel subspaces in $\mathbb{R}^m$ and $\mathbb{R}^n$ are mapped into the zero vector of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. With the orthogonal matrices $\boldsymbol{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ and $\boldsymbol{U} = (\mathbf{u}_1, \ldots \mathbf{u}_n)$, the following SVD of $\boldsymbol{A}$ and $\boldsymbol{A}^T$ holds:

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{S}\boldsymbol{U}^T = \sum_{i=1}^{r} s_i \mathbf{v}_i \mathbf{u}_i^T \quad \text{and} \quad \boldsymbol{A}^T = \boldsymbol{U}\boldsymbol{S}^T\boldsymbol{V}^T = \sum_{i=1}^{r} s_i \mathbf{u}_i \mathbf{v}_i^T, \tag{12}$$

where $\boldsymbol{S}$ is an $m \times n$ so-called *generalized diagonal matrix* which contains the singular values $s_1, \ldots, s_r$ in the first r positions of its main diagonal (starting from the upper left corner) and zeros otherwise. We remark that there are other equivalent forms of the above SVD depending on, whether $m < n$ or $m \geq n$. For example, in the $m < n$ case, $\boldsymbol{V}$ can be an $m \times m$ orthogonal, $\boldsymbol{S}$ an $m \times m$ diagonal, and $\boldsymbol{U}$ an $n \times m$ suborthogonal matrix with the same relevant entries. About the uniqueness of the SVD the following can be stated: to a single positive singular value there corresponds a unique singular vector pair (of course, the orientation of the left and right singular vectors can be changed at the same time). To a positive singular value of multiplicity say $k > 1$ a k-dimensional left and right so-called *isotropic subspace* corresponds, within which, any k-element orthonormal sets can embody the left and right singular vectors with orientation such that the requirements in (11) are met.

We also remark that the singular values of a symmetric matrix are the absolute values of its eigenvalues. In case of a positive eigenvalue, the left and right

singular vectors are the same (they coincide with the corresponding eigenvector with any, but the same orientation). In case of a negative eigenvalue, the left and right side singular vectors are opposite (any of them is the corresponding eigenvector which have a bivalent orientation). In case of a zero singular value the orientation is immaterial, as it does not contribute to the SVD of the underlying matrix. Numerical algorithms for SD and SVD of real matrices are presented in [4, 15].

Assume that the $m \times n$ matrix $\boldsymbol{A}$ of rank r has SVD (12). It is easy to see that the matrices $\boldsymbol{A}\boldsymbol{A}^T$ and $\boldsymbol{A}^T\boldsymbol{A}$ are positive semidefinite (possibly, positive definite) matrices of rank r, and their SD is

$$\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{V}(\boldsymbol{S}\boldsymbol{S}^T)\boldsymbol{V}^T = \sum_{i=1}^{r} s_i^2 \mathbf{v}_i \mathbf{v}_i^T \quad \text{and} \quad \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{U}(\boldsymbol{S}^T\boldsymbol{S})\boldsymbol{U}^T = \sum_{i=1}^{r} s_i^2 \mathbf{u}_i \mathbf{u}_i^T$$

where the diagonal matrices $\boldsymbol{S}\boldsymbol{S}^T$ and $\boldsymbol{S}^T\boldsymbol{S}$ both contain the numbers $s_1^2, \ldots, s_r^2$ in the leading positions of their main diagonals as non-zero eigenvalues.

These facts together also imply that the only positive singular value of a suborthogonal matrix is the 1 with multiplicity of its rank.

Definition 9 *We say that the $n \times n$ symmetric matrix $\boldsymbol{G} = (g_{ij})$ is a Gram-matrix if its entries are inner products: there is a dimension $d > 0$ and vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$ such that*

$$g_{ij} = \mathbf{x}_i^T \mathbf{x}_j, \quad i, j = 1, \ldots n.$$

Proposition 7 *The symmetric matrix $\boldsymbol{G}$ is a Gram-matrix if and only if it is positive semidefinite or positive definite.*

Proof 1 *If $\boldsymbol{G}$ is a Gram-matrix, then it can be decomposed as $\boldsymbol{G} = \boldsymbol{A}\boldsymbol{A}^T$, where $\boldsymbol{A}^T = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. With this,*

$$\mathbf{x}^T\boldsymbol{G}\mathbf{x} = \mathbf{x}^T\boldsymbol{A}\boldsymbol{A}^T\mathbf{x} = (\boldsymbol{A}^T\mathbf{x})^T(\boldsymbol{A}^T\mathbf{x}) = \|\boldsymbol{A}^T\mathbf{x}\|^2 \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Conversely, if $\boldsymbol{G}$ is positive semidefinite (or positive definite) with rank $r \leq n$, then its SD – using (10) – can be written as

$$\boldsymbol{G} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

Let the $n \times r$ matrix $\boldsymbol{A}$ be defined as

$$\boldsymbol{A} = (\sqrt{\lambda_1}\mathbf{u}_1, \ldots, \sqrt{\lambda_r}\mathbf{u}_r). \tag{13}$$

Then the row vectors of the matrix $\boldsymbol{A}$ will be r-dimensional vectors reproducing $\boldsymbol{G}$. Of course, such a decomposition is not unique: first of all, instead of $\boldsymbol{A}$ the matrix $\boldsymbol{A}\boldsymbol{Q}$ will also do, where $\boldsymbol{Q}$ is an arbitrary $r \times r$ orthogonal matrix (obviously, $\mathbf{x}_i$'s can be rotated); and $\mathbf{x}_i$'s can also be put in a higher ($d > r$) dimension with attaching any (but the same) number of zero coordinates to them.

The spectral norm (operator norm) of an $m \times n$ real matrix $\boldsymbol{A}$ of rank r, with positive singular values $s_1 \geq \cdots \geq s_r > 0$, is

$$\|\boldsymbol{A}\| = \max_{\|\mathbf{x}\|=1} \|\boldsymbol{A}\mathbf{x}\| = s_1,$$

and its *Frobenius norm*, denoted by $\|.\|_2$, is

$$\|\boldsymbol{A}\|_2 = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2} = \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^T)} = \sqrt{\mathrm{tr}(\boldsymbol{A}^T\boldsymbol{A})} = \left(\sum_{i=1}^{r} s_i^2\right)^{1/2}.$$

The Frobenius norm is sometimes called Euclidean norm and corresponds to the Hilbert–Schmidt norm of operators between separable Hilbert spaces. Accordingly, the operator $\mathbb{R}^n \to \mathbb{R}^m$ corresponding to $\boldsymbol{A}$, can be considered as an integral operator with kernel (a_{ij}), hence, $\|\boldsymbol{A}\|_2$ is the L^2-norm of this kernel, which is expressed in the above formula. For a symmetric real matrix $\boldsymbol{A}$,

$$\|\boldsymbol{A}\| = \max_{\|\mathbf{x}\|=1} \|\boldsymbol{A}\mathbf{x}\| = \max_i |\lambda_i| \quad \text{and} \quad \|\boldsymbol{A}\|_2 = \left(\sum_{i=1}^{r} \lambda_i^2\right)^{1/2}.$$

Obviously, for a real matrix $\boldsymbol{A}$ of rank r,

$$\|\boldsymbol{A}\| \le \|\boldsymbol{A}\|_2 \le \sqrt{r}\|\boldsymbol{A}\|. \tag{14}$$

More generally, a matrix norm is called *unitary invariant* if

$$\|\boldsymbol{A}\|_{\mathrm{un}} = \|\boldsymbol{Q}\boldsymbol{A}\boldsymbol{R}\|_{\mathrm{un}}$$

with any $m \times m$ and $n \times n$ orthogonal matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively. It is easy to see that a unitary invariant norm of a real matrix merely depends on its singular values (or eigenvalues if it is symmetric). For example, the spectral and Frobenius norms are such, and the *Schatten norm* (sometimes called trace norm) defined by

$$\|\boldsymbol{A}\|_4 = \left(\sum_{i=1}^{r} s_i^4\right)^{\frac{1}{4}} \tag{15}$$

is also unitary invariant.

By means of SD or SVD we are able to define so-called *generalized inverses* of singular square or rectangular matrices: in fact, any matrix that undoes the effect of the underlying linear transformation between the ranges of $\boldsymbol{A}^T$ and $\boldsymbol{A}$ will do. A generalized inverse is far not unique as any transformation operating on the kernels can be added. However, the following *Moore–Penrose inverse* is uniquely defined and it coincides with the usual inverse if exists.

Definition 10 *The Moore–Penrose inverse of the $n \times n$ symmetric matrix with SD (10) is*

$$\boldsymbol{A}^+ = \sum_{i=1}^{r} \frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^T = \boldsymbol{U}\boldsymbol{\Lambda}^+\boldsymbol{U}^T,$$

where $\boldsymbol{\Lambda}^+ = \mathrm{diag}(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_r}, 0, \ldots, 0)$ is the diagonal matrix containing the reciprocals of the non-zero eigenvalues, otherwise zeros, in its main diagonal.

The Moore–Penrose inverse of the $m \times n$ real matrix is the $n \times m$ matrix $\boldsymbol{A}^+$ with SVD (12)

$$\boldsymbol{A}^+ = \sum_{i=1}^{r} \frac{1}{s_i}\mathbf{u}_i\mathbf{v}_i^T = \boldsymbol{U}\boldsymbol{S}^+\boldsymbol{V}^T,$$

where $\boldsymbol{S}^+$ is $n \times m$ generalized diagonal matrix containing the reciprocals of the non-zero singular values of $\boldsymbol{A}$ in the leading positions, otherwise zeros, in its main diagonal.

Note that, analogously, any analytic function f of the symmetric real matrix $\boldsymbol{A}$ can be defined by its SD, $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, in the following way:

$$f(\boldsymbol{A}) := \boldsymbol{U} f(\boldsymbol{\Lambda}) \boldsymbol{U}^T \tag{16}$$

where $f(\boldsymbol{\Lambda}) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$, of course, only if every eigenvalue is in the domain of f. In this way, for a positive semidefinite (or positive definite) $\boldsymbol{A}$, its squareroot is

$$\boldsymbol{A}^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^T, \tag{17}$$

and for a regular $\boldsymbol{A}$ its inverse is obtained by applying the $f(x) = x^{-1}$ function to it:

$$\boldsymbol{A}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^T.$$

For a singular $\boldsymbol{A}$, the Moore–Penrose inverse is obtained by using $\boldsymbol{\Lambda}^+$ instead of $\boldsymbol{\Lambda}^{-1}$. Accordingly, for a positive semidefinite matrix, its $-1/2$ power is defined as the squareroot of $\boldsymbol{A}^+$.

We will frequently use the following propositions, called *separation theorems* for singular values and eigenvalues. These are special cases of Propositions 2, 3, 4, and 5. About the applications of these theorems in multivariate statistical analysis see [7].

Proposition 8 *Let $\boldsymbol{A}$ be an $m \times n$ real matrix with SVD in (12). Assume that its non-zero singular values are enumerated in non-increasing order ($s_1 \geq s_2 \geq \ldots s_r > 0$). Then*

$$\max_{\substack{\mathbf{x}\in\mathbb{R}^n,\, \mathbf{y}\in\mathbb{R}^m \\ \|\mathbf{x}\|=1,\, \|\mathbf{y}\|=1}} \mathbf{y}^T \boldsymbol{A}\mathbf{x} = s_1$$

and it is attained with the choice $x = \mathbf{u}_1$ and $y = \mathbf{v}_1$ (uniquely if $s_1 > s_2$). This was the $k = 1$ case. Further, for $k = 2, 3, \ldots, r$

$$\max_{\substack{\mathbf{x}\in\mathbb{R}^n,\, \mathbf{y}\in\mathbb{R}^m \\ \|x\|=1,\, \|y\|=1 \\ \mathbf{x}^T\mathbf{u}_i=0\,(i=1,\ldots,k-1) \\ \mathbf{y}^T\mathbf{v}_i=0\,(i=1,\ldots,k-1)}} \mathbf{y}^T \boldsymbol{A}\mathbf{x} = s_k$$

and it is attained with the choice $\mathbf{x} = \mathbf{u}_k$ and $y = \mathbf{v}_k$ (uniquely if $s_k > s_{k+1}$).

A bit more is true.

Proposition 9 *Let $\boldsymbol{A}$ be an $m \times n$ real matrix with SVD in (12). Assume that its singular values are enumerated in non-increasing order and $k > 0$ is an integer such that $s_k > s_{k+1}$. Then*

$$\max_{\substack{\boldsymbol{X}\text{ is } n\times k,\, \boldsymbol{Y}\text{ is } m\times k \\ \boldsymbol{X}^T\boldsymbol{X}=\boldsymbol{I}_k,\, \boldsymbol{Y}^T\boldsymbol{Y}=\boldsymbol{I}_k}} \operatorname{tr}(\boldsymbol{Y}^T\boldsymbol{A}\boldsymbol{X}) = \max_{\substack{\mathbf{x}_i^T\mathbf{x}_j=\delta_{ij} \\ \mathbf{y}_i^T\mathbf{y}_j=\delta_{ij}}} \sum_{i=1}^{k} \mathbf{y}_i^T \boldsymbol{A}\mathbf{x}_i = \sum_{i=1}^{k} s_i$$

and it is attained with the suborthogonal matrices $\boldsymbol{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$ and $\boldsymbol{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ such that $\mathbf{x}_i = \mathbf{u}_i$ and $\mathbf{y}_i = \mathbf{v}_i$ $(i = 1, \ldots, k)$.

Proposition 10 *Let $\boldsymbol{A}$ be $n\times n$ real symmetric matrix with SD in (10). Assume that its eigenvalues are enumerated in non-increasing order ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$). Then*

$$\max_{\|\mathbf{x}\|=1} \mathbf{x}^T \boldsymbol{A}\mathbf{x} = \lambda_1$$

and it is attained with the choice $\mathbf{x} = \mathbf{u}_1$ (uniquely if $\lambda_1 > \lambda_2$). This was the $k = 1$ case. Further, for $k = 2, 3, \ldots, n$

$$\max_{\substack{\mathbf{x}\in\mathbb{R}^n,\, \|\mathbf{x}\|=1 \\ \mathbf{x}^T\mathbf{u}_i=0\ (i=1,\ldots,k-1)}} \mathbf{x}^T \boldsymbol{A}\mathbf{x} = \lambda_k$$

and it is attained with the choice $\mathbf{x} = \mathbf{u}_k$ (uniquely if $\lambda_k > \lambda_{k+1}$).

A bit more is true.

Proposition 11 *Let $\boldsymbol{A}$ be $n\times n$ real symmetric matrix with SD in (10). Assume that its eigenvalues are enumerated in non-increasing order and $k > 0$ is an integer such that $\lambda_k > \lambda_{k+1}$. Then*

$$\max_{\substack{\boldsymbol{X} \text{ is } n\times k \\ \boldsymbol{X}^T\boldsymbol{X}=\boldsymbol{I}_k}} \operatorname{tr}(\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}) = \max_{\substack{\mathbf{x}_i\in\mathbb{R}^n\ (i=1,\ldots,k) \\ \mathbf{x}_i^T\mathbf{x}_j=\delta_{ij}}} \sum_{i=1}^{k} \mathbf{x}_i^T \boldsymbol{A}\mathbf{x}_i = \sum_{i=1}^{k} \lambda_i$$

and it is attained with the suborthogonal matrix $\boldsymbol{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$ such that $\mathbf{x}_i = \mathbf{u}_i$ $(i = 1, \ldots, k)$.

Applying the above statements for the symmetric real matrix $-\boldsymbol{A}$, similar statements can be proved. Note that these minima and maxima are attained as we optimize over convex sets.

Proposition 12 *Let $\boldsymbol{A}$ be $n\times n$ real symmetric matrix with SD in (10). Assume that its eigenvalues are enumerated in non-increasing order ($\lambda_1 \geq \cdots \geq \lambda_{n-1} \geq \lambda_n$). Then*

$$\min_{\|x\|=1} \mathbf{x}^T \boldsymbol{A}\mathbf{x} = \lambda_n$$

and it is attained with the choice $\mathbf{x} = \mathbf{u}_n$ (uniquely if $\lambda_n < \lambda_{n-1}$). This was the $k = 1$ case. Further, for $k = 2, 3, \ldots, n$

$$\min_{\substack{\mathbf{x}\in\mathbb{R}^n,\, \|\mathbf{x}\|=1 \\ \mathbf{x}^T\mathbf{u}_i=0\ (i=n-k+2,\ldots,n)}} \mathbf{x}^T \boldsymbol{A}\mathbf{x} = \lambda_{n-k+1}$$

and it is attained with the choice $\mathbf{x} = \mathbf{u}_{n-k+1}$ (uniquely if $\lambda_{n-k+2} < \lambda_{n-k+1}$).

Proposition 13 *Let $\boldsymbol{A}$ be $n\times n$ real symmetric matrix with SD in (10). Assume that its eigenvalues are enumerated in non-increasing order and $k > 0$ is an integer such that $\lambda_{n-k+1} < \lambda_{n-k}$. Then*

$$\min_{\substack{\mathbf{X} \text{ is } n\times k \\ \mathbf{X}^T\mathbf{X}=\boldsymbol{I}_k}} \operatorname{tr}(\mathbf{X}^T\boldsymbol{A}\mathbf{X}) = \min_{\substack{\mathbf{x}_i\in\mathbb{R}^n\ (i=1,\ldots,k) \\ \mathbf{x}_i^T\mathbf{x}_j=\delta_{ij}}} \sum_{i=1}^{k} \mathbf{x}_i^T \boldsymbol{A}\mathbf{x}_i = \sum_{i=n-k+1}^{n} \lambda_i$$

and it is attained with the suborthogonal matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$ such that $\mathbf{x}_i = \mathbf{u}_{n-k+i}$ $(i = 1, \ldots, k)$.

Many of the above propositions follow from the forthcoming so-called *minimax principle*.

Theorem 10 (Courant–Fischer–Weyl theorem) *Let $\boldsymbol{A}$ be an $n \times n$ symmetric real matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Then*

$$\lambda_k = \max_{\substack{F \subset \mathbb{R}^n \\ \dim(F)=k}} \min_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\|=1}} \mathbf{x}^T \boldsymbol{A} \mathbf{x} = \min_{\substack{F \subset \mathbb{R}^n \\ \dim(F)=n-k+1}} \max_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\|=1}} \mathbf{x}^T \boldsymbol{A} \mathbf{x} \quad (k = 1, \ldots, n).$$

The statement naturally extends to singular values of rectangular matrices.

Theorem 11 *Let $\boldsymbol{A}$ be an $m \times n$ real matrix with positive singular values $s_1 \geq \cdots \geq s_r$, where $r = \operatorname{rank}(\boldsymbol{A})$. Then*

$$s_k = \max_{\substack{F \subset \mathbb{R}^n \\ \dim(F)=k}} \min_{\mathbf{x} \in F} \frac{\|\boldsymbol{A}\mathbf{x}\|}{\|\mathbf{x}\|} \quad (k = 1, \ldots, r).$$

Theorem 10 is, in turn, implied by the upcoming separation theorem. In the sequel, we will denote by $\lambda_i(.)$ the ith largest eigenvalue of the symmetric matrix in the argument (they are enumerated in non-increasing order).

Theorem 12 (Cauchy–Poincaré separation theorem) *Let $\boldsymbol{A}$ be an $n \times n$ symmetric real matrix and $\boldsymbol{B}$ be an $n \times k$ suborthogonal matrix ($k \leq n$). Then*

$$\lambda_i(\boldsymbol{A}) \geq \lambda_i(\boldsymbol{B}^T \boldsymbol{A} \boldsymbol{B}) \geq \lambda_{i+n-k}(\boldsymbol{A}), \quad i = 1, \ldots, k.$$

The first inequality is attained with equality if $\boldsymbol{B}$ contains the eigenvectors corresponding to the k largest eigenvalues of $\boldsymbol{A}$ in its columns; whereas, the second inequality is attained with equality if $\boldsymbol{B}$ contains the eigenvectors corresponding to the k smallest eigenvalues of $\boldsymbol{A}$ in its columns.

Note that the first inequality makes sense for a k such that $\lambda_k > \lambda_{k+1}$, whereas the second inequality makes sense for a k such that $\lambda_{n-k+1} < \lambda_{n-k}$.

The Cauchy–Poincaré theorem implies the following important inequalities due to H. Weyl.

Theorem 13 (Weyl's perturbation theorem) *Let $\boldsymbol{A}$ and $\boldsymbol{C}$ be $n \times n$ symmetric matrices. Then*

$$\begin{aligned} \lambda_j(\boldsymbol{A} + \boldsymbol{C}) &\leq \lambda_i(\boldsymbol{A}) + \lambda_{j-i+1}(\boldsymbol{C}) \quad \text{if} \quad i \leq j, \\ \lambda_j(\boldsymbol{A} + \boldsymbol{C}) &\geq \lambda_i(\boldsymbol{A}) + \lambda_{j-i+n}(\boldsymbol{C}) \quad \text{if} \quad i \geq j. \end{aligned}$$

The above inequalities give rise to the following perturbation result for symmetric matrices we will intensively use in the lessons. Here we consider symmetric matrices such that $\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{C}$, where $\boldsymbol{C}$ is a 'small' perturbation on $\boldsymbol{B}$.

Theorem 14 *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $n \times n$ symmetric matrices. Then*

$$|\lambda_i(\boldsymbol{A}) - \lambda_i(\boldsymbol{B})| \leq \|\boldsymbol{A} - \boldsymbol{B}\|, \quad i = 1, \ldots, n.$$

A similar statement is valid for rectangular matrices.

Theorem 15 *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ real matrices with singular values $s_1(\boldsymbol{A}) \geq \cdots \geq s_{\min\{m,n\}}(\boldsymbol{A})$ and $s_1(\boldsymbol{B}) \geq \cdots \geq s_{\min\{m,n\}}(\boldsymbol{B})$. Then*

$$|s_i(\boldsymbol{A}) - s_i(\boldsymbol{B})| \leq \|\boldsymbol{A} - \boldsymbol{B}\|, \quad i = 1, \ldots, \min\{m, n\}.$$

Applying the above theorems for rank k matrices $\boldsymbol{B}$ we can solve the following optimization problems stated in a more general form, for rectangular matrices.

Theorem 16 *Let $\boldsymbol{A}$ be an arbitrary $m \times n$ real matrix with SVD $\sum_{i=1}^{r} s_i \mathbf{u}_i \mathbf{v}_i^T$, where r is the rank of $\boldsymbol{A}$. Then for any positive integer $k \leq r$ such that $s_k > s_{k+1}$,*

$$\min_{\substack{\boldsymbol{B} \text{ is } m \times n \\ \operatorname{rank}(\boldsymbol{B})=k}} \|\boldsymbol{A} - \boldsymbol{B}\| = s_{k+1} \quad \textit{and} \quad \min_{\substack{\boldsymbol{B} \text{ is } m \times n \\ \operatorname{rank}(\boldsymbol{B})=k}} \|\boldsymbol{A} - \boldsymbol{B}\|_2 = \left(\sum_{i=k+1}^{r} s_i^2 \right)^{1/2}$$

hold, and both minima are attained with the matrix $\boldsymbol{B}_k = \sum_{i=1}^{k} s_i \mathbf{u}_i \mathbf{v}_i^T$.

Note that $\boldsymbol{B}_k$ is called the *best rank k approximation* of $\boldsymbol{A}$, and the aforementioned theorem guarantees that it is the best approximation both in spectral and Frobenius norm. In fact, it is true for any unitary invariant norm:

$$\min_{\substack{\boldsymbol{B} \text{ is } m \times n \\ \operatorname{rank}(\boldsymbol{B})=k}} \|\boldsymbol{A} - \boldsymbol{B}\|_{\text{un}} = \|\boldsymbol{A} - \boldsymbol{B}_k\|_{\text{un}}.$$

The next theorem is about perturbation of eigenvectors of symmetric matrices. In fact, the original theorem, due to [2], applies to complex self-adjoint (Hermitian) matrices and to spectral subspaces spanned by eigenvectors corresponding to a set of eigenvectors separated from the others, see [12]. Here we follow the formalism of [1].

Theorem 17 (Perturbation of spectral subspaces) *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be symmetric matrices; S_1 and S_2 are subsets of $\mathbb{R}$ or $\mathbb{C}$ such that $\operatorname{dist}(S_1, S_2) = \delta > 0$. Let $\boldsymbol{P_A}(S_1)$ and $\boldsymbol{P_B}(S_2)$ be orthogonal projections onto the subspace spanned by the eigenvectors of the matrix in the lower index, corresponding to the eigenvalues within the subset in the argument. Then with any unitary invariant norm:*

$$\|\boldsymbol{P}_A(S_1)\boldsymbol{P}_B(S_2)\|_{\text{un}} \leq \frac{c_1}{\delta} \|\boldsymbol{P}_A(S_1)(\boldsymbol{A} - \boldsymbol{B})\boldsymbol{P}_B(S_2)\|_{\text{un}} \leq \frac{c_1}{\delta} \|\boldsymbol{A} - \boldsymbol{B}\|_{\text{un}}$$

where c_1 is a constant.

Remark 1 *In another context, [13] proves that $c_1 = \pi/2$. When S_1 and S_2 are separated by an annulus, then the constant improves to $c_1 = 1$; further, with the Frobenius norm, $c_1 = 1$ will always do, see [1]. If $\boldsymbol{P}_A(S_1)$ and $\boldsymbol{P}_B^{\perp}(S_2)$ project onto subspaces of the same dimension, then either the spectral- or the Frobenius norm of $\boldsymbol{P}_A(S_1)\boldsymbol{P}_B(S_2)$ can be expressed in terms of the sines of the so-called canonical (principal) angles between these subspaces and $\|\boldsymbol{P}_A(S_1)\boldsymbol{P}_B(S_2)\|_2$ is considered as the distance between them. This is why the special case of Theorem 17, when Frobenius norm is used, is called Davis–Kahan $\sin(\boldsymbol{\theta})$ theorem.*

We will also need the following statements concerning relations between singular values, eigenvalues, and traces.

Proposition 14 *Let $\boldsymbol{A}$ be an $n \times n$ matrix and $\boldsymbol{Q}$ be an $n \times n$ orthogonal one. Then $\operatorname{tr}(\boldsymbol{AQ})$ is maximal if $\boldsymbol{AQ}$ is symmetric, and in this case the trace of this symmetric matrix is equal to the sum of the singular values of $\boldsymbol{A}$.*

Proposition 15 *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $n \times n$ symmetric, positive semidefinite matrices with eigenvalues $\lambda_i(\boldsymbol{A})$'s and $\lambda_i(\boldsymbol{B})$'s. Then*

$$\operatorname{tr}(\boldsymbol{AB}) \leq \sum_{i=1}^{n} \lambda_i(\boldsymbol{A}) \cdot \lambda_i(\boldsymbol{B}),$$

with equality if and only if $\boldsymbol{A}$ and $\boldsymbol{B}$ commute, i.e. $\boldsymbol{AB} = \boldsymbol{BA}$.

Note that a necessary and sufficient condition for $\boldsymbol{A}$ and $\boldsymbol{B}$ commute is that they have the same system of eigenvectors (possibly, eigenspaces).

Proposition 16 *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $n \times n$ real matrices with singular values $s_1(\boldsymbol{A}) \geq \cdots \geq s_n(\boldsymbol{A}) \geq 0$ and $s_1(\boldsymbol{B}) \geq \cdots \geq s_n(\boldsymbol{B}) \geq 0$. Then*

$$\prod_{i=1}^{k} s_i(\boldsymbol{AB}) \leq \prod_{i=1}^{k} [s_i(\boldsymbol{A}) \cdot s_i(\boldsymbol{B})], \quad k = 1, \ldots, n.$$

Especially, for $k = 1$, this implies that

$$s_{\max}(\boldsymbol{AB}) \leq s_{\max}(\boldsymbol{A}) \cdot s_{\max}(\boldsymbol{B}),$$

which is not surprising, since the maximal singular value is the operator norm of the matrix.

The next part will be devoted to the Perron–Frobenius theory of matrices with nonnegative entries. First we define the notion of the irreducibility for a quadratic matrix, and a similar notion for rectangular matrices.

Definition 11 *A quadratic matrix $\boldsymbol{A}$ is called reducible if there exists an appropriate permutation of its rows and columns, or equivalently, the exists a permutation matrix (see Section 3) $\boldsymbol{P}$ such that, with it, $\boldsymbol{A}$ can be transformed into the following block-matrix form:*

$$\boldsymbol{PAP}^T = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{D} & \boldsymbol{C} \end{pmatrix} \quad \text{or} \quad \boldsymbol{PAP}^T = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{D} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix},$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are quadratic matrices, whereas $\boldsymbol{O}$ is the zero matrix of appropriate size. A quadratic matrix is called irreducible if it is not reducible.

Note that the eigenvalues of a quadratic matrix are unaffected under the same permutation of its rows and columns, while the coordinates of the corresponding eigenvectors are subject to the same permutation. Since in Definition 11, the same permutation is applied to the rows and columns, and the spectrum of the involved block-matrix consists of the spectra of $\boldsymbol{B}$ and $\boldsymbol{C}$, the SD of a reducible matrix can be traced back to the SD of some smaller matrices. Now we want to create a similar notion for rectangular matrices in terms of their SVD.

Definition 12 *An $m \times n$ real matrix $\boldsymbol{A}$ is called decomposable if there exist appropriate permutations of its rows and columns, or equivalently, the exist permutation matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ of sizes $m \times m$ and $n \times n$, respectively, such that, with them, $\boldsymbol{A}$ can be transformed into the following block-matrix form:*

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}^T = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}.$$

The real matrix $\boldsymbol{A}$ is non-decomposable if it is not decomposable.

Note that the singular values of a real matrix are unaffected under appropriate permutation of its rows and columns, while the coordinates of the corresponding singular vectors are subject to the same permutations. Since in Definition 12, the singular spectrum of the matrix $\boldsymbol{A}$ is composed of the singular spectra of $\boldsymbol{B}$ and $\boldsymbol{C}$, only SVD of non-decomposable matrices is of importance.

When we apply Definition 12 to a symmetric matrix $\boldsymbol{A}$, the matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ of Definition 11 are also symmetric, whereas $\boldsymbol{D}$ becomes the zero matrix. As the singular values of a symmetric matrix are the absolute values of its eigenvalues, a reducible symmetric matrix is also decomposable. However, the converse is not true: there exist quadratic (even symmetric) real matrices which are decomposable, still irreducible. The easiest example is the

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

matrix which has eigenvalues $1, -1$ and singular values $1, 1$. The former ones cannot be concluded from the spectra of smaller matrices, while the latter ones can, as the above matrix can be transformed into the identity matrix with interchanging its rows or columns (but not both of them). Though, for positive semidefinite matrices, the two notions are equivalent. We also remark that the rectangular matrix $\boldsymbol{A}$ is non-decomposable if and only if the symmetric matrix $\boldsymbol{A}\boldsymbol{A}^T$ (or, equivalently, $\boldsymbol{A}^T\boldsymbol{A}$) is irreducible.

The block-matrix of the following statement (see e.g. [1]) is also decomposable and irreducible at the same time.

Proposition 17 *Let $\boldsymbol{A}$ be an $m \times n$ real matrix of rank r. Then the $(m+n) \times (m+n)$ symmetric matrix*

$$\widetilde{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{O} \end{pmatrix}$$

has the following spectrum: its largest and smallest non-zero eigenvalues are

$$\lambda_i(\widetilde{\boldsymbol{A}}) = -\lambda_{n+m-i+1}(\widetilde{\boldsymbol{A}}) = s_i(\boldsymbol{A}), \quad i = 1, \ldots, r,$$

while the others are zeros, where $\lambda_i(.)$ and $s_i(.)$ denote the ith largest eigenvalue and singular value of the matrix in the argument, respectively.

Hence, the spectrum of $\widetilde{\boldsymbol{A}}$ is symmetric about the 0, and its rank is $2r$; moreover, the relevant eigenvectors can be composed of the singular vector pairs of $\boldsymbol{A}$.

The subsequent theorems apply to matrices of nonnegative entries.

Theorem 18 (Frobenius theorem) *Any irreducible, quadratic real matrix of nonnegative entries has a single positive eigenvalue among its maximum absolute value ones with corresponding eigenvector of all positive coordinates.*

Remark 2 *More precisely, there may be $k \geq 1$ complex eigenvalues of maximum absolute value r, allocated along the circle of radius r in the complex plane. In fact, those complex numbers are vertices of a regular k-gone, but the point is that exactly one of these vertices is allocated on the positive part of the real axis, see [11] for the proof.*

The Perron theorem is the specialized version of the Frobenius theorem, applicable to matrices of strictly positive entries.

Theorem 19 (Perron theorem) *Any irreducible, quadratic real matrix of positive entries has only one maximum absolute value eigenvalue which is positive with multiplicity one, and the corresponding eigenvector has all positive coordinates.*

As a byproduct of the proof of the above theorems, the following useful bounds for the maximum absolute value positive eigenvalue – guaranteed by the Frobenius theorem – can be obtained.

Proposition 18 *Let $\boldsymbol{A}$ be an irreducible $n \times n$ real matrix of nonnegative entries and introduce the following notation for the maxima and minima of the row-sums of $\boldsymbol{A}$:*

$$m := \min_{i \in \{1,\ldots,n\}} \sum_{j=1}^{n} a_{ij} \quad \text{and} \quad M := \max_{i \in \{1,\ldots,n\}} \sum_{j=1}^{n} a_{ij}.$$

Then the single positive eigenvalue λ with maximum absolute value admits the following lower and upper bound:

$$m \leq \lambda \leq M,$$

where either the lower or the upper bound is attained if and only if $m = M$, i.e. the row-sums of $\boldsymbol{A}$ have a constant value.

Finally, we introduce the *Kronecker-sum* and *Kronecker-product* of matrices.

Definition 13 *Let $\boldsymbol{A}_i$ be $n_i \times n_i$ matrix $(i = 1, \ldots, k)$, $n := \sum_{i=1}^{k} n_i$. The Kronecker-sum of $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ is the $n \times n$ block-diagonal matrix $\boldsymbol{A}$ the diagonal blocks of which are the matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ in this order. We use the notation $\boldsymbol{A} = \boldsymbol{A}_1 \oplus \cdots \oplus \boldsymbol{A}_k$ for it.*

Definition 14 *Let $\boldsymbol{A}$ be $p \times n$ and $\boldsymbol{B}$ be $q \times m$ real matrix. Their Kronecker-product, denoted by $\boldsymbol{A} \otimes \boldsymbol{B}$, is the following $pq \times nm$ block-matrix: it has pn blocks each of which is a $q \times m$ matrix such that the block indexed by (i, j) is the matrix $a_{ij}\boldsymbol{B}$ $(i = 1, \ldots, p;\ j = 1, \ldots, n)$.*

This product is associative, for the addition distributive, but usually not commutative. If $\boldsymbol{A}$ is $n \times n$ and $\boldsymbol{B}$ is $m \times m$ quadratic matrix, then

$$|\boldsymbol{A} \otimes \boldsymbol{B}| = |\boldsymbol{A}|^m \cdot |\boldsymbol{B}|^n;$$

further, if both are regular, then so is their Kronecker-product. Namely,

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}.$$

It is also useful to know that – provided $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric – the spectrum of $\boldsymbol{A} \otimes \boldsymbol{B}$ consists of the real numbers

$$\alpha_i \beta_j \quad (i = 1, \ldots, n;\ j = 1, \ldots, m),$$

where α_i's and β_j's are the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

References

[1] Bhatia R 1997 *Matrix Analysis*. Springer.

[2] Davis C and Kahan WM 1970 The rotation of eigenvectors by a perturbation, III. *SIAM J. Numer. Anal.* **7**, 1–46.

[3] Fréchet M 1907 Sur les ensembles de fonctions et les opérations linéaires. *C. R. Acad. Sci. Paris* **144**, 1414–1416.

[4] Golub GH and Van Loan CF 1996 *Matrix Computations*, 3rd edn. Johns Hopkins University Press, Baltimore, MD.

[5] Komornik V 2003 *Lecture notes on real analysis II (in Hungarian)*. Typotex, Budapest.

[6] Petz D 2008 *Quantum Information Theory and Quantum Statistics*. Springer.

[7] Rao CR 1979 Separation theorems for singular values of matrices and their applications in multivariate analysis. *J. Multivariate Anal.* **9**, 362–377.

[8] Riesz F 1907 Sur une espéce de géométrie analytique des systémes de fonctions sommables. *C. R. Acad. Sci. Paris* **144**, 1409–1411.

[9] Riesz F 1909 Sur les opérations fonctionnelles linéaires. *C. R. Acad. Sci. Paris* **149**, 974–977.

[10] Riesz F and Sz.-Nagy B 1952 *Leçons d'analyse fonctionnelle*. Academic Publishing House, Budapest.

[11] Rózsa P 1974 *Linear Algebra and Its Applications (in Hungarian)*. Műszaki Könyvkiadó, Budapest.

[12] Stewart GW and Sun J-g. 1990 *Matrix Pertubation Theory*. Academic Press.

[13] Sz.-Nagy B 1953 Über die Ungleichnung von H. Bohr. *Math. Nachr.* **9**, 255–259.

[14] Von Neumann J 1935 On complete topological spaces. *Trans. Amer. Math. Soc.* **37**, 1–20.

[15] Wilkinson JH 1965 *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford.