Matrices and Linear Operators

Huiqiang Jiang

1. Fundamentals of Linear Spaces

A *field* K is a nonempty set together with two operations, usually called addition and multiplication, and denoted by $+$ and $\cdot$ respectively, such that the following axioms hold:
1. Closure of K under addition and multiplication: $a, b \in K \Longrightarrow a + b,\ a \cdot b \in K$;
2. Associativity of addition and multiplication: For any $a, b, c \in K$,

$$a + (b + c) = (a + b) + c,\ a \cdot (b \cdot c) = (a \cdot b) \cdot c;$$

3. Commutativity of addition and multiplication: For any $a, b \in K$,

$$a + b = b + a,\ a \cdot b = b \cdot a;$$

4. Existence of additive and multiplicative identity elements: There exists an element of K, called the additive identity element and denoted by 0, such that for all $a \in K$, $a + 0 = a$. Likewise, there is an element, called the multiplicative identity element and denoted by 1, such that for all $a \in K$, $a \cdot 1 = a$. To exclude the trivial ring, the additive identity and the multiplicative identity are required to be distinct.
5. Existence of additive inverses and multiplicative inverses: For every $a \in K$, there exists an element $-a \in K$, such that

$$a + (-a) = 0.$$

Similarly, for any $a \in K \setminus \{0\}$, there exists an element $a^{-1} \in K$, such that $a \cdot a^{-1} = 1$. We can define subtraction and division operations by

$$a - b + a + (-b) \text{ and } \frac{a}{b} = a \cdot b^{-1} \text{ if } b \neq 0.$$

6. Distributivity of multiplication over addition: For any $a, b \in K$,

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

Examples of field: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$. In our lecture, K will be either $\mathbb{R}$ or $\mathbb{C}$, the elements in K are called scalars.

A *linearspace* X over a field K is a set in which two operations are defined: Addition, denoted by $+$ such that

$$x, y \in X \Longrightarrow x + y \in X$$

and scalar multiplication such that

$$a \in K \text{ and } x \in X \Longrightarrow ax \in X.$$

These two operations satisfy the following axioms:
1. Associativity of addition:

$$x + (y + z) = (x + y) + z;$$

2. Commutativity of addition:

$$x + y = y + x;$$

3. Identity element of addition: There exists an element $0 \in X$, called the zero vector, such that $x + 0 = x$ for all $x \in X$.
4. Inverse elements of addition: For every $x \in X$, there exists an element $-x \in X$, called the additive inverse of x, such that

$$x + (-x) = 0.$$

5. Compatibility (Associativity) of scalar multiplication with field multiplication: For any $a, b \in K$, $x \in X$,

$$a\,(bx) = (ab)\,x.$$

6. Identity element of scalar multiplication: $1x = x$.
7. Distributivity of scalar multiplication with respect to vector addition:

$$a\,(x + y) = ax + ay.$$

8. Distributivity of scalar multiplication with respect to field addition:

$$(a + b)\,x = ax + bx.$$

The elements in a linear space are called vectors.

Remark 1. *Zero vector is unique.*

Remark 2. $(-1)\,x = -x$.

Example 1. $\mathbb{R}^n$, $\mathbb{C}^n$.

Example 2. *Polynomials with real coefficients of order at most* n.

Definition 1. *A one-to-one correspondence between two linear spaces over the same field that maps sums into sums and scalar multiples into scalar multiples is called an isomorphism.*

Example 3. *The linear space of real valued functions on* $\{1, 2, \cdots, n\}$ *is isomorphic to* $\mathbb{R}^n$.

Definition 2. *A subset* Y *of a linear space* X *is called a subspace if sums and scalar multiples of elements of* Y *belong to* Y.

The set $\{0\}$ consisting of the zero element of a linear space X is a subspace of X. It is called the trivial subspace.

Definition 3. *The sum of two subsets* Y *and* Z *of a linear space* X*, is the set defined by*

$$Y + Z = \{y + z \in X : y \in Y, z \in Z\}.$$

The intersection of two subsets Y *and* Z *of a linear space* X*, is the set defined by*

$$Y \cap Z = \{x \in X : x \in Y, x \in Z\}.$$

Proposition 1. *If* Y *and* Z *are two linear subspaces of* X*, then both* $Y + Z$ *and* $Y \cap Z$ *are linear subspaces of* X*.*

Remark 3. *The union of two subspaces may not be a subspace.*

Definition 4. *A linear combination of* m *vectors* $x_1, \cdots, x_m$ *of a linear space is a vector of the form*

$$\sum_{j=1}^{m} c_j x_j \text{ where } c_j \in K.$$

Given m vectors $x_1, \cdots, x_m$ of a linear space X, the set of all linear combinations of $x_1, \cdots, x_m$ is a subspace of X, and that it is the smallest subspace of X containing $x_1, \cdots, x_m$. This is called the subspace spanned by $x_1, \cdots, x_m$.

Definition. A set of vectors x, , ... , x, in X span the whole space X if everyx in X can be expressed as a linear combination of x,, ... ,x,,,.

Definition 5. *The vectors* $x_1, \cdots, x_m$ *are called linearly dependent if there exist scalars* $c_1, \cdots, c_m$*, not all of them are zero, such that*

$$\sum_{j=1}^{m} c_j x_j = 0.$$

The vectors $x_1, \cdots, x_m$ *are called linearly independent if they are not dependent.*

Definition 6. *A finite set of vectors which span* X *and are linearly independent is called a basis for* X*.*

Proposition 2. *A linear space which is spanned by a finite set of vectors has a basis.*

Definition 7. *A linear space* X *is called finite dimensional if it has a basis.*

Theorem 1. *All bases for a finite-dimensional linear space* X *contain the same number of vectors. This number is called the dimension of* X *and is denoted as* $\dim X$*.*

Proof. The theorem follows from the lemma below. □

Lemma 1. *Suppose that the vectors* $x_1, \cdots, x_n$ *span a linear space* X *and that the vectors* $y_1, \cdots, y_m$ *in* X *are linearly independent. Then* $m \leq n$*.*

Proof. Since $x_1, \cdots, x_n$ span X, we have

$$y_1 = \sum_{j=1}^{n} c_j x_j.$$

We claim that not all c_j are zero, otherwise $y_1 = 0$ and $y_1, \cdots, y_m$ must be linearly dependent. Suppose $c_k \neq 0$, then x_k can be expressed as a linear combination of y_k and the remaining x_j. So the set consisting of the x_j's, with x_k replaced by y_k span X. If $m \geq n$, repeat this step $n - 1$ more times and conclude that $y_1, \cdots, y_n$ span X. If $m > n$, this contradicts the linear independence of the vectors $y_1, \cdots, y_m$. □

We define the dimension of the trivial space consisting of the single element 0 to be zero.

Theorem 2. *Every linearly independent set of vectors $y_1, \cdots, y_m$ in a finite dimensional linear space X can be completed to a basis of X.*

Theorem 3. *Let X be a finite dimensional linear space over K with $\dim X = n$, then X is isomorphic to K^n.*

Theorem 4. *(a) Every subspace Y of a finite-dimensional linear space X is finite dimensional.*
(b) Every subspace Y has a complement in X, that is, another subspace Z such that every vector x in X can be decomposed uniquely as

$$x = y + z,\ y \in Y,\ z \in Z.$$

Furthermore $\dim X = \dim Y + \dim Z$.

X is said to be the direct sum of two subspaces Y and Z that are complements of each other. More generally X is said to be the direct sum of its subspaces $Y_1, \cdots, Y_m$ if every x in X can be expressed uniquely as

$$x = \sum_{j=1}^{m} y_j \text{ where } y_j \in Y_j.$$

This relation is denoted as

$$X = Y_1 \oplus Y_2 \oplus \cdots \oplus Y_m.$$

If X is finite dimensional and

$$X = Y_1 \oplus Y_2 \oplus \cdots \oplus Y_m,$$

then

$$\dim X = \sum_{j=1}^{m} \dim Y_j.$$

Definition 8. *For X a linear space, Y a subspace, we say that two vectors x_1, x_2 in X are congruent modulo Y, denoted*

$$x_1 \equiv x_2 \ mod\, Y$$

if $x_1 - x_2 \in Y$.

Congruence mod Y is an equivalence relation, that is, it is
(i) symmetric: if $x_1 \equiv x_2$, then $x_2 \equiv x_1$.
(ii) reflexive: $x \equiv x$ for all x in X.
(iii) transitive: if $x_1 \equiv x_2$ and $x_2 \equiv x_3$, then $x_1 \equiv x_3$.

We can divide elements of X into congruence classes mod Y. The congruence class containing the vector x is the set of all vectors congruent with X; we denote it by $\{x\}$.

The set of congruence classes can be made into a linear space by defining addition and multiplication by scalars, as follows:

$$\{x\} + \{y\} = \{x + y\},$$
$$a\{x\} = \{ax\}.$$

The linear space of congruence classes defined above is called the quotient space of X mod Y and is denoted as X/Y.

REMARK 4. *X/Y is not a subspace of X.*

THEOREM 5. *If Y is a subspace of a finite-dimensional linear space X; then*

$$\dim Y + \dim (X/Y) = \dim X.$$

PROOF. Let $x_1, \cdots, x_m$ be a basis for Y, $m = \dim Y$. This set can be completed to form a basis for X by adding $x_{m+1}, \cdots, x_n$, $n = \dim X$. We claim that $\{x_{m+1}\}, \cdots, \{x_n\}$ form a basis for X/Y by verifying that they are linearly independent and span the whole space X/Y. □

THEOREM 6. *Suppose X is a finite-dimensional linear space, U and V two subspaces of X. Then we have*

$$\dim (U + V) = \dim U + \dim V - \dim (U \cap V).$$

PROOF. If $U \cap V = \{0\}$, then $U + V$ is a direct sum and hence

$$\dim (U + V) = \dim U + \dim V.$$

In general, let $W = U \cap V$, we claim $U/W + V/W = (U + V)/W$ is a direct sum and hence

$$\dim (U/W) + \dim (V/W) = \dim ((U + V)/W).$$

Applying Theorem 5, we have

$$\dim (U + V) = \dim U + \dim V - \dim (U \cap V).$$

□

DEFINITION 9. *The Cartesian sum $X_1 \oplus X_2$ of two linear spaces X_1, X_2 over the same field is the set of pairs (x_1, x_2) where $x_i \in X_i$, $i = 1, 2$. $X_1 \oplus X_2$ is a linear space with addition and multiplication by scalars defined componentwisely.*

THEOREM 7.

$$\dim X_1 \oplus X_2 = \dim X_1 + \dim X_2.$$

More generally, we can define the Cartesian sum $\oplus_{k=1}^m X_k$ of m linear spaces X_1, X_2, $\cdots$, X_m, and we have

$$\dim \oplus_{k=1}^m X_k = \sum_{k=1}^m \dim X_k.$$

2. Dual Spaces

Let X be a linear space over a field K. A scalar valued function $l : X \to K$ is called linear if

$$l(x+y) = l(x) + l(y)$$

for all x, y in X, and $l(kx) = kl(x)$ for $\forall x \in X$ and $\forall k \in K$.

The set of linear functions on a linear space X forms a linear space X', the dual space of X, if we define

$$(l+m)(x) = l(x) + m(x)$$

and

$$(kl)(x) = k(l(x)).$$

Theorem 8. *Let X be a linear space of dimension n. Under a chosen basis $x_1, \cdots, x_n$, the elements x of X can be represented as arrays of n scalars:*

$$x = (c_1, \cdots, c_n) = \sum_{k=1}^{n} c_k x_k.$$

Let $a_1, \cdots, a_n$ be any array of n scalars; the function l defined by

$$l(x) = \sum_{k=1}^{n} a_k c_k$$

is a linear function of X. Conversely, every linear function l of X can be so represented.

Theorem 9.

$$\dim X' = \dim X.$$

We write

$$(l, x) \equiv l(x)$$

which is a bilinear function of l and x. The dual of X' is X'', consisting of all linear functions on X'.

Theorem 10. *The bilinear function $(l, x) = l(x)$ gives a natural identification of X with X''. The map $x \mapsto x^{**}$ is an isomorphism where*

$$(x^{**}, l) = (l, x)$$

for any $l \in X^*$.

Definition 10. *Let Y be a subspace of X. The set of linear functions that vanish on Y, that is, satisfy*

$$l(y) = 0 \text{ for all } y \in Y,$$

is called the annihilator of the subspace Y; it is denoted by $Y^\perp$.

Theorem 11. *$Y^\perp$ is a subspace of X' and*

$$\dim Y^\perp + \dim Y = \dim X.$$

Proof. We shall establish a natural isomorphism $T : Y^{\perp} \to (X/Y)'$: For any $l \in Y^{\perp} \subset X'$, we define for any $\{x\} \in X/Y$,

$$(Tl)(\{x\}) = l(x).$$

Then $Tl \in (X/Y)'$ is well defined. One can verify that T is an isomorphism. Hence

$$\dim Y^{\perp} = \dim\left((X/Y)'\right) = \dim(X/Y) = \dim X - \dim Y.$$

□

The dimension of $Y^{\perp}$ is called the codimension of Y as a subspace of X.

$$\operatorname{codim} Y + \dim Y = \dim X.$$

Since $Y^{\perp}$ is a subspace of X', its annihilator, denoted by $Y^{\perp\perp}$, is a subspace of X''.

Theorem 12. *Under the natural identification of X'' and X, for every subspace Y of a finite-dimensional space X,*

$$Y^{\perp\perp} = Y.$$

Proof. Under the identification of X'' and X, $Y \subset Y^{\perp\perp}$. Now $\dim Y^{\perp\perp} = \dim Y$ implies $Y^{\perp\perp} = Y$. □

More generally, let S be a subset of X. The annihilator of S is defined by

$$S^{\perp} = \{l \in X' : l(x) = 0 \text{ for any } x \in S\}.$$

Theorem 13. *Let S be a subset of X.*

$$S^{\perp} = (\operatorname{span} S)^{\perp}.$$

Theorem 14. *Let $t_1, t_2, \cdots, t_n$ be n distinct real numbers. For any finite interval I on the real axis, there exist n numbers $m_1, m_2, \cdots, m_n$ such that*

$$\int_I p(t)\,dt = \sum_{k=1}^{n} m_k p(t_k)$$

holds for all polynomials p of degree less than n.

3. Linear Mappings

Let X, U be linear spaces over the same field K. A mapping $T : X \to U$ is called linear if it is additive:

$$T(x+y) = T(x) + T(y), \text{ for any } x, y \in X.$$

and if it is homogeneous:

$$T(kx) = kT(x) \text{ for any } k \in K \text{ and } x \in X.$$

For simplicity, we often write $T(x) = Tx$.

EXAMPLE 4. *Isomorphisms are linear mappings.*

EXAMPLE 5. *Differentiation is linear.*

EXAMPLE 6. *Linear functionals are linear mappings.*

THEOREM 15. *The image of a subspace of X under a linear map T is a subspace of U. The inverse image of a subspace of U, is a subspace of X.*

DEFINITION 11. *The range of T is the image of X under T; it is denoted as R_T. The nullspace of T is the inverse image of $\{0\}$, denoted as N_T. R_T and N_T are subspaces of X.*

DEFINITION 12. $\dim R_T$ *is called the rank of the mapping T and* $\dim N_T$ *is called the nullity of the mapping T.*

THEOREM 16 (Rank-Nullity Theorem). *Let $T : X \to U$ be linear. Then*

$$\dim R_T + \dim N_T = \dim X.$$

PROOF. Define $\tilde{T} : X/N_T \to R_T$ so that

$$\tilde{T}\{x\} = Tx.$$

Then T is well defined and it is an isomorphism. Hence

$$\dim R_T = \dim X/N_T = \dim X - \dim N_T.$$

□

COROLLARY 1. *Let $T : X \to U$ be linear.*
(a) Suppose $\dim U < \dim X$, *then $Tx = 0$ for some $x \neq 0$.*
(b) Suppose $\dim U = \dim X$, *and the only vector satisfying $Tx = 0$ is $x = 0$. Then $R_T = U$ and T is an isomorphism.*

COROLLARY 2. *Suppose $m < n$, then for any real numbers t_{ij}, $1 \le i \le m$, $1 \le j \le n$, the system of linear equations*

$$\sum_{j=1}^{n} t_{ij} x_j = 0, 1 \le i \le m$$

has a nontrivial solution.

COROLLARY 3. *Given n^2 real numbers t_{ij}, $1 \le i, j \le n$, the inhomogeneous system of linear equations*

$$\sum_{j=1}^{n} t_{ij} x_j = u_i, 1 \le i \le n$$

has a unique solution for any u_i, $1 \leq i \leq n$ *iff the homogeneous system*

$$\sum_{j=1}^{n} t_{ij} x_j = 0, 1 \leq i \leq n$$

has only the trivial solution.

We use $L(X,U)$ to denote the collection of all linear maps from X to U. Suppose that $T, S \in L(X,U)$, we define their sum $T+S$ by

$$(T+S)(x) = Tx + Sx \text{ for any } x \in X$$

and we define, for $k \in K$, kT by

$$(kT)(x) = kTx \text{ for any } x \in X.$$

Then $T+S$, $kT \in L(X,U)$ and $L(X,U)$ is a linear space.

Let $T \in L(X,U)$ and $S \in L(U,V)$, we can define the composition of T with S by

$$S \circ T(x) = S(Tx).$$

Note that composition is associative: if $R \in L(V,Z)$, then

$$R \circ (S \circ T) = (R \circ S) \circ T.$$

THEOREM 17. *(i) The composite of linear mappings is also a linear mapping. (ii) Composition is distributive with respect to the addition of linear maps, that is,*

$$(R+S) \circ T = R \circ T + S \circ T$$

whenever the compositions are defined.

REMARK 5. *We use* ST *to denote* $S \circ T$*, called the multiplication of* S *and* T*. Note that* $ST \neq TS$ *in general.*

DEFINITION 13. *A linear map is called invertible if it is 1-to-1 and onto, that is, if it is an isomorphism. The inverse is denoted as* T^{-1}.

THEOREM 18. *(i) The inverse of an invertible linear map is linear.*
(ii) If S *and* T *are both invertible, and if* $ST = S \circ T$ *is defined, then* ST *also is invertible, and*

$$(ST)^{-1} = T^{-1} S^{-1}.$$

DEFINITION 14. *Let* $T \in L(X,U)$*, the transpose* $T' \in L(U', X')$ *of* T *is defined by*

$$(T'(l))(x) = l(Tx) \text{ for any } l \in U' \text{ and } x \in X.$$

We could use the dual notation to rewrite the above identity as

$$(T'l, x) = (l, Tx).$$

THEOREM 19. *Whenever defined, we have*

$$(ST)' = T'S',$$
$$(T+R)' = T' + R',$$
$$\left(T^{-1}\right)' = (T')^{-1}.$$

EXAMPLE 7. *Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be defined by $y=Tx$ where*

$$y_{i=}\sum_{j=1}^{n}t_{ij}x_j.,1\leq i\leq m.$$

Identifying $(\mathbb{R}^n)^{'}=\mathbb{R}^n$ and $(\mathbb{R}^m)^{'}=\mathbb{R}^m$, $T':\mathbb{R}^m\to\mathbb{R}^n$ is defined by $v=T'u$ where

$$v_j=\sum_{i=1}^{m}t_{ij}u_i,1\leq j\leq n.$$

THEOREM 20. *Let $T\in L(X,U)$. Identifying $X''=X$ and $U''=U$. We have $T''=T$.*

THEOREM 21. *Let $T\in L(X,U)$.*

$$R_T^{\perp}=N_{T'},$$
$$R_T=N_{T'}^{\perp}.$$

THEOREM 22. *Let $T\in L(X,U)$.*

$$\dim R_T=\dim R_{T'}.$$

COROLLARY 4. *Let $T\in L(X,U)$. Suppose that $\dim X=\dim U$, then*

$$\dim N_T=\dim N_{T'}.$$

We now consider $L(X,X)$ which forms an algebra if we define the multiplication as composition.

The set of invertible elements of $L(X,X)$ forms a group under multiplication. This group depends only on the dimension of X, and the field K of scalars. It is denoted as $GL(n,K)$ where $n=\dim X$.

Given an invertible element $S\in L(X,X)$, we assign to each $M\in L(X,X)$ the element $M_S=SMS^{-1}$. This assignment $M\to M_S$ is called a similarity transformation; M is said to be similar to M_S.

THEOREM 23. *(a) Every similarity transformation is an automorphism of $L(X,X)$:*

$$(kM)_S=kM_S,$$
$$(M+K)_S=M_S+K_S,$$
$$(MK)_S=M_SK_S.$$

(b) The similarity transformations form a group with

$$(M_S)_T=M_{TS}.$$

THEOREM 24. *Similarity is an equivalence relation; that is, it is:*
(i) Reflexive. M is similar to itself.
(ii) Symmetric. If M is similar to K, then K is similar to M.
(iii) Transitive. If M is similar to K, and K is similar to L, then M is similar to L.

THEOREM 25. *If either A or B in $L(X,X)$ is invertible, then AB and BA are similar.*

DEFINITION 15. *A linear mapping $P\in L(X,X)$ is called a projection if it satisfies $P^2=P$.*

Theorem 26. *Let $P \in L(X, X)$ be a projection. Then*

$$X = N_P \oplus R_P.$$

And P restricted on R_P is the identity map.

Definition 16. *The commutator of two mappings A and B of X into X is $AB - BA$. Two mappings of X into X commute if their commutator is zero.*

4. Matrices

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $y = Tx$ where

$$y_{i=} \sum_{j=1}^{n} t_{ij} x_j., 1 \leq i \leq m.$$

Then T is a linear map. On the other hand, every map $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ can be written in this form. Actually, t_{ij} is the ith component of Te_j, where $e_j \in \mathbb{R}^n$ has jth component 1, all others 0.

We write

$$T = (t_{ij})_{m \times n} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{pmatrix},$$

which is called an m by n $(m \times n)$ matrix, m being the number of rows, n the number of columns. A matrix is called a square matrix if $m = n$. The numbers t_{ij} are called the entries of the matrix T.

A matrix T can be thought of as a row of column vectors, or a column of row vectors:

$$T = (c_1, \cdots, c_n) = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}$$

where

$$c_j = \begin{pmatrix} t_{1j} \\ \vdots \\ t_{mj} \end{pmatrix} \text{ and } r_i = (t_{i1}, \cdots, t_{in}).$$

Thus

$$Te_j^n = c_j = \sum_{i=1}^{m} t_{ij} e_i^m$$

where we write vectors in $\mathbb{R}^m$ as column vectors.

Since matrices represent linear mappings, the algebra of linear mappings induces a corresponding algebra of matrices:

$$T + S = (t_{ij} + s_{ij})_{m \times n},$$
$$kT = (kt_{ij})_{m \times n}.$$

If $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $S \in L(\mathbb{R}^m, \mathbb{R}^l)$, then the product $ST = S \circ T \in L(\mathbb{R}^n, \mathbb{R}^l)$. For $e_j \in \mathbb{R}^n$,

$$STe_j^n = St_{ij} e_i^m = t_{ij} Se_i^m = t_{ij} s_{ki} e_k^l = \left(\sum_{i=1}^{m} s_{ki} t_{ij} \right) e_k^l,$$

hence $(ST)_{kj} = \sum_{i=1}^{m} s_{ki}t_{ij}$ which is the product of kth row of S and jth column of T.

Remark 6. *If $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $x \in \mathbb{R}^n$, we can also view Tx as the product of two matrices.*

We can write any $n \times n$ matrix A in 2×2 block form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} is an $k \times k$ matrix and A_{22} is an $(n-k) \times (n-k)$ matrix. Product of block matrices follows the same formula.

The dual of the space $\mathbb{R}^n$ of all column vectors with n components is the space $(\mathbb{R}^n)'$ of all row vectors with n components. Here for $l \in (\mathbb{R}^n)'$ and $x \in \mathbb{R}^n$,

$$lx = \sum_{i=1}^{n} l_i x_i.$$

Let $T \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $T' \in L\left((\mathbb{R}^m)', (\mathbb{R}^n)'\right)$. Identifying $(\mathbb{R}^n)' = \mathbb{R}^n$ and $(\mathbb{R}^m)' = \mathbb{R}^m$, $T' : \mathbb{R}^m \to \mathbb{R}^n$ has matrix representation T^T, called the transpose of matrix T,

$$\left(T^T\right)_{ij} = T_{ji}.$$

Theorem 27. *Let $T \in L(\mathbb{R}^n, \mathbb{R}^m)$. The range of T consists of all linear combinations of the columns of the matrix T.*

The dimension of R_T is called the column rank of T and $\dim R_{T^T}$ is called the row rank of T. Hence, the column rank and the row rank of T are the same.

Any $T \in L(X, U)$ can be represented by a matrix once we choose bases for X and U. A choice of basis in X defines an isomorphism $B : X \to \mathbb{R}^n$, and similarly we have isomorphism $C : U \to \mathbb{R}^m$. We have $M = CTB^{-1} \in L(\mathbb{R}^n, \mathbb{R}^m)$ which can be represented by a matrix.

If $T \in L(X, X)$, and $B : X \to \mathbb{R}^n$ is an isomorphism, then we have $M = BTB^{-1} \in L(\mathbb{R}^n, \mathbb{R}^n)$ is a square matrix. Let $C : X \to \mathbb{R}^n$ be an isomorphism, then $N = CTC^{-1}$ is another square matrix representing T. Since

$$N = CB^{-1}MBC^{-1},$$

M, N are similar. Simiar matrices represnts the same mapping under different bases.

Definition 17. *Invertible and singular matrices.*

Definition 18. *Unit matrix I.*

Definition 19. *Upper triangular matrix, lower triangular matrix and diagonal matrix.*

Definition 20. *A square matrix T is called a tridiagonal matrix if $t_{ij} = 0$ whenever $|i-j| > 1$.*

Gaussian elimination.

5. Determinant and Trace

A simplex in $\mathbb{R}^n$ is a polyhedron with $n+1$ vertices. The simplex is ordered if we have an order for the vertices. We denote the first vertix to be the origin and denote the rest in its order of $a_1, a_2, \cdots, a_n$. Our goal is to define the volumn of an ordered simplex.

An ordered simplex S is called degenerate if it lies on an $(n-1)$-dimensional subspace. An ordered nondegenerate simplex

$$S = (0, a_1, a_2, \cdots, a_n)$$

is called positively oriented if it can be deformed continuously and nondegenerately into the standard ordered simplex $(0, e_1, e_2, \cdots, e_n)$, where e_j is the jth unit vector in the standard basis of $\mathbb{R}^n$. Otherwise, we say it is negatively oriented.

For a nondegenerate oriented simplex S we define $O(S) = +1$ (-1) if it is positively (negatively) oriented. For a degenerate simplex S, we set $O(S) = 0$.

The volume of a simplex S is given inductively by the elementary formula

$$\mathrm{Vol}(S) = \frac{1}{n}\mathrm{Vol}(Base) \times Altitude.$$

And the signed volume of an ordered simplex S is

$$\Sigma(S) = O(S)\,\mathrm{Vol}(S).$$

We view $\Sigma(S)$ as a function of vectors $(a_1, a_2, \cdots, a_n)$:
1. $\Sigma(S) = 0$ if $a_j = a_k$ for some $k = j$.
2. $\Sigma(S)$ is linear on a_j if we fix other vertices.

Definition 21. *The determinant of a square matrix $A = (a_1, a_2, \cdots, a_n)$ is defined by*

$$D(a_1, a_2, \cdots, a_n) = n!\Sigma(S)$$

where $S = (0, a_1, a_2, \cdots, a_n)$.

Theorem 28. *(i) $D(a_1, a_2, \cdots, a_n) = 0$ if $a_j = a_k$ for some $k = j$.*
(ii) $D(a_1, a_2, \cdots, a_n)$ is a multilinear function of its arguments.
(iii) Normalization: $D(e_1, e_2, \cdots, e_n) = 1$.
(iv) D is an alternating function of its arguments, in the sense that if a_i and a_j are interchanged, $i \neq j$, the value of D changes by the factor (-1).
(v) If $a_1, a_2, \cdots, a_n$ are linearly dependent, then $D(a_1, a_2, \cdots, a_n) = 0$.

Proof. (iv)

$$\begin{aligned} D(a,b) &= D(a,a) + D(a,b) = D(a, a+b) \\ &= D(a, a+b) - D(a+b, a+b) \\ &= -D(b, a+b) = -D(b,a). \end{aligned}$$

□

Next we introduce the concept of permutation. A permutation is a mapping p of n objects, say the numbers $1, 2, \cdots, n$ onto themselves. Permutations are invertible and they form a group with compositions. These groups, except for $n = 2$, are noncommutative.

Let $x_1, \cdots, x_n$ be n variables; their discriminant is defined to be

$$P(x_1, \cdots, x_n) = \prod_{i<j}(x_i - x_j).$$

Let p be any permutation. Clearly,

$$\prod_{i<j}\left(x_{p(i)}-x_{p(j)}\right)$$

is either $P(x_1,\cdots,x_n)$ or $-P(x_1,\cdots,x_n)$.

Definition 22. *The signature $\sigma(p)$ of a permutation p is defined by*

$$P\left(x_{p(1)},\cdots,x_{p(n)}\right)=\sigma(p)\,P(x_1,\cdots,x_n).$$

Hence, $\sigma(p)=\pm 1$.

Theorem 29.

$$\sigma(p_1\circ p_2)=\sigma(p_1)\,\sigma(p_2).$$

Proof.

$$\begin{aligned}\sigma(p_1\circ p_2)&=\frac{P\left(x_{p_1p_2(1)},\cdots,x_{p_1p_2(n)}\right)}{P(x_1,\cdots,x_n)}\\&=\frac{P\left(x_{p_1p_2(1)},\cdots,x_{p_1p_2(n)}\right)}{P\left(x_{p_2(1)},\cdots,x_{p_2(n)}\right)}\cdot\frac{P\left(x_{p_2(1)},\cdots,x_{p_2(n)}\right)}{P(x_1,\cdots,x_n)}\\&=\sigma(p_1)\,\sigma(p_2).\end{aligned}$$

□

Given any pair of indices, $j\neq k$, we can define a permutation p such that $p(i)=i$ for $i\neq j$ or k, $p(j)=k$ and $p(k)=j$. Such a permutation is called a transposition.

Theorem 30. *The signature of a transposition t is -1. Every permutation p can be written as a composition of transpositions.*

Proof. By induction. □

We have $\sigma(p)=1$ if p is a composition of even number of transpositions and p is said to be an even permutation. We have $\sigma(p)=-1$ if p is a composition of even number of transpositions and p is said to be an odd permutation.

Theorem 31. *Assume that for $1\leq k\leq n$,*

$$a_k=\begin{pmatrix}a_{1k}\\ \vdots\\ a_{nk}\end{pmatrix}\in\mathbb{R}^n.$$

The determinant

$$D(a_1,\cdots,a_n)=\sum_p\sigma(p)\,a_{p(1)1}a_{p(2)2}\cdots a_{p(n)n}$$

where the summation is over all permutations.

PROOF.

$$\begin{aligned} D(a_1, \cdots, a_n) &= D\left(\sum_{j=1}^{n} a_{j1} e_j, \cdots, \sum_{j=1}^{n} a_{jn} e_j\right) \\ &= \sum_{1 \le j_k \le n, 1 \le k \le n} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} D(e_{j_1}, \cdots, e_{j_n}) \\ &= \sum_{p} \sigma(p)\, a_{p(1)1} a_{p(2)2} \cdots a_{p(n)n} \end{aligned}$$

□

REMARK 7. *Determinant is defined by properties 1,2,3 in Theorem 28.*

DEFINITION 23. *Let $A = (a_1, \cdots, a_n)$ be an $n \times n$ matrix where a_k, $1 \le k \le n$ are column vectors. Its determinant, denoted as $\det A$, is defined by*

$$\det A = D(a_1, \cdots, a_n).$$

THEOREM 32.

$$\det A^T = \det A.$$

THEOREM 33. *Let A, B be two $n \times n$ matrices.*

$$\det(BA) = \det A \det B.$$

PROOF. Let $A = (a_1, \cdots, a_n)$.

$$\det(BA) = D(Ba_1, \cdots, Ba_n).$$

Assuming that $\det B \neq 0$, we define

$$C(a_1, \cdots, a_n) = \frac{\det(BA)}{\det B} = \frac{D(Ba_1, \cdots, Ba_n)}{\det B}.$$

We verify that C satisfies properties 1,2,3 in Theorem 28. Hence $C = D$.

When $\det B = 0$, we could do approximation $B(t) = B + tI$. □

COROLLARY 5. *An $n \times n$ matrix A is invertible iff $\det A \neq 0$.*

THEOREM 34 (Laplace expansion). *For any $j = 1, \cdots, n$,*

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

Here A_{ij} is the (ij)th minor of A.,i.e., is the $(n-1) \times (n-1)$ matrix obtained by striking out the ith row and jth column of A.

PROOF. The jth column $a_j = \sum a_{ij} e_i$. Hence,

$$\begin{aligned} \det A &= \sum_{i=1}^{n} a_{ij} D(a_1, \cdots, a_{j-1}, e_i, a_{j+1}, \cdots, a_n) \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \end{aligned}$$

where we used the lemma below. □

LEMMA 2. *Let A be a matrix whose jth column is e_i. Then*

$$\det A = (-1)^{i+j} \det A_{ij}.$$

PROOF. Suppose $i = j = 1$. We define

$$C(A_{11}) = \det \begin{pmatrix} 1 & 0 \\ 0 & A_{11} \end{pmatrix} = \det A$$

where the second equality follows from the basic properties of det. Then $C = \det A_{11}$ since it satisfies the properties 1,2,3 in Theorem 28. General cases follow similarly or could be deduced from the case $i = j = 1$. □

Let $A_{n\times n}$ be invertible. Then

$$Ax = u$$

has a unique solution. Write $A = (a_1, \cdots, a_n)$ and $x = \sum x_j e_j$, we have

$$u = \sum x_j a_j.$$

We consider

$$A_k = (a_1, \cdots, a_{k-1}, u, a_{k+1}, \cdots, a_n).$$

Then

$$\begin{aligned} \det A_k &= \sum x_j \det(a_1, \cdots, a_{k-1}, a_j, a_{k+1}, \cdots, a_n) \\ &= x_k \det A, \end{aligned}$$

hence

$$x_k = \frac{\det A_k}{\det A}.$$

Since

$$\det A_k = \sum_j (-1)^{j+k} u_j \det A_{jk},$$

we have

$$x_k = \sum_{j=1}^{n} (-1)^{j+k} u_j \frac{\det A_{jk}}{\det A}.$$

Comparing it with $x = A^{-1}u$, we have proved

THEOREM 35. *The inverse matrix A^{-1} of an invertible matrix A has the form*

$$\left(A^{-1}\right)_{kj} = (-1)^{j+k} \frac{\det A_{jk}}{\det A}.$$

DEFINITION 24. *The trace of a square matrix A, denoted as* $\operatorname{tr} A$*, is the sum of the entries on its diagonal:*

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}.$$

THEOREM 36. *(i) Trace is a linear functional on matrices.*
(ii) Trace is commutative: $\operatorname{tr} AB = \operatorname{tr} BA$.

DEFINITION 25.

$$\operatorname{tr} AA^T = \sum_{i,j=1}^{n} (a_{ij})^2$$

The square root of the double sum on the right is called the Euclidean, or Hilbert-Schmidt, norm of the matrix A.

The matrix A is called similar to the matrix B if there is an invertible matrix S such that $A = SBS^{-1}$.

THEOREM 37. *Similar matrices have the same determinat and the same trace.*

REMARK 8. *The determinant and trace of a linear map T can be defined as the determinant and trace of a matrix representing T.*

If $A = (a_{ij})_{n\times n}$ is an upper triangular square matrix, we have

$$\det A = \prod_{k=1}^{n} a_{kk}.$$

More generally, if A is an upper triangular block matrix, i.e., $A = (A_{ij})_{k\times k}$ where A_{ii} is an $n_i \times n_i$ matrix and $A_{ij} = O$ if $i > j$, then we can show that

$$\det A = \prod_{j=1}^{k} \det A_{jj}.$$

REMARK 9. *If A, B, C, D are $n \times n$ matrices, we may not have*

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det D - \det C \det B.$$

Another guess

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det (AD - CB)$$

is also false in general. For example,

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

THEOREM 38. *Let A, B, C, D be $n \times n$ matrices and $AC = CA$. Then*

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det (AD - CB).$$

PROOF. We first assume that A is invertible. Since

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & O \\ C & D - CA^{-1}B \end{pmatrix},$$

we have

$$\begin{aligned} \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \det A \det \left(D - CA^{-1}B\right) \\ &= \det \left(AD - ACA^{-1}B\right) \\ &= \det (AD - CB). \end{aligned}$$

If A is singular, we consider $A_\varepsilon = A + \varepsilon I$ and then send $\varepsilon \to 0$. □

THEOREM 39. *Let A, B be $n \times n$ matrices. Then*

$$\det (I - AB) = \det (I - BA).$$

REMARK 10. *In general, it is not true that*

$$\det (A - BC) = \det (A - CB).$$

6. Determinants of Special Matrices

Definition 26. *Let $n \geq 2$. Given n scalars $a_1, \cdots, a_n$, the Vandermonde matrix $V(a_1, \cdots, a_n)$ is a square matrix whose columns form a geometric progression:*

$$V(a_1, \cdots, a_n) = \begin{pmatrix} 1 & \cdots & 1 \\ a_1 & & a_n \\ \vdots & & \vdots \\ a_1^{n-1} & \cdots & a_n^{n-1} \end{pmatrix}.$$

Theorem 40.

$$\det V(a_1, \cdots, a_n) = \prod_{j>i} (a_j - a_i).$$

Proof. $\det V$ is a polynomial in the a_i of degree less than or equal to $n(n-1)/2$. Whenever $a_i = a_j$ for some $i \neq j$, we have

$$\det V = 0,$$

Hence, $a_i - a_j$ is a factor of $\det V$. Hence $\det V$ is divisible by the product

$$\prod_{j>i} (a_j - a_i)$$

which is also a polynomial in the a_i of degree equal to $n(n-1)/2$. Hence for some constant c_n.

$$\det V(a_1, \cdots, a_n) = c_n \prod_{j>i} (a_j - a_i).$$

We consider the coefficient of a_n^{n-1}, we have

$$c_n \prod_{n>j>i} (a_j - a_i) = \det V(a_1, \cdots, a_{n-1}) = c_{n-1} \prod_{n>j>i} (a_j - a_i)$$

hence $c_n = c_{n-1}$. Since $c_2 = 1$, we have $c_n = 1$. □

Example 8. *Let*

$$p(s) = x_1 + x_2 s + \cdots x_n s^{n-1}$$

be a polynomial in s. Let $a_1, \cdots, a_n$ be n distinct numbers, and let $p_1, \cdots, p_n$ be n arbitrary complex numbers; we wish to choose the coefficients $x_1, \cdots, x_n$, so that

$$p(a_j) = p_j, \ 1 \leq j \leq n.$$

Then we have

$$V(a_1, \cdots, a_n) x = p.$$

Definition 27. *Given $2n$ scalars $a_1, \cdots, a_n$, $b_1, \cdots, b_n$. The Cauchy matrix*

$$C(a_1, \cdots, a_n; b_1, \cdots, b_n)$$

is the $n \times n$ matrix whose ij-th element is

$$\frac{1}{a_i + b_j}.$$

THEOREM 41.

$$\det(a_1,\cdots,a_n;b_1,\cdots,b_n)=\frac{\prod_{j>i}(a_j-a_i)\prod_{j>i}(b_j-b_i)}{\prod_{i,j}(a_i+b_j)}.$$

PROOF. We have

$$\begin{aligned}&\prod_{i,j}(a_i+b_j)\det(a_1,\cdots,a_n;b_1,\cdots,b_n)\\&=\det T_{ij}\end{aligned}$$

where

$$T_{ij}=\prod_{k\neq j}(a_i+b_k).$$

One can show that for some constant c_n,

$$\prod_{i,j}(a_i+b_j)\det(a_1,\cdots,a_n;b_1,\cdots,b_n)=c_n\prod_{j>i}(a_j-a_i)\prod_{j>i}(b_j-b_i).$$

Let $a_n=b_n=d$, the coefficient of d^{2n-2} is

$$c_n\prod_{n>j>i}(a_j-a_i)\prod_{n>j>i}(b_j-b_i).$$

On the other hand, since for $1\leq ij<n$,

$$\begin{aligned}T_{nn}&=\prod_{k\neq n}(d+b_k)=d^{n-1}+\cdots,\\T_{nj}&=\prod_{k\neq j}(d+b_k)=2d^{n-1}+\cdots,\\T_{in}&=\prod_{k\neq n}(a_i+b_k),\\T_{ij}&=\prod_{k\neq j}(a_i+b_k)=(a_i+d)\prod_{k\neq j,k\neq n}(a_i+b_k),\end{aligned}$$

we see the coefficient of d^{2n-2} is decided by

$$\begin{aligned}&d^{n-1}\det(T_{ij})_{(n-1)\times(n-1)}\\&=d^{2(n-1)}\prod_{i,j\leq n-1}(a_i+b_j)\det(a_1,\cdots,a_{n-1};b_1,\cdots,b_{n-1})\\&=d^{2(n-1)}c_{n-1}\prod_{j>i}(a_j-a_i)\prod_{j>i}(b_j-b_i).\end{aligned}$$

Hence, we conclude $c_n=c_{n-1}$. Since $c_2=1$ we conclude $c_n=1$. □

EXAMPLE 9. *If we consider*

$$\begin{aligned}a_1=1,a_2=2,a_3=3,\\b_1=1,b_2=2,b_3=3.\end{aligned}$$

We have

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}.$$

And

$$\det A = \frac{(2)^2}{2 \times 3 \times 4 \times 3 \times 4 \times 5 \times 4 \times 5 \times 6} = \frac{1}{43200}.$$

EXAMPLE 10. *Consider*

$$T = \left(\frac{1}{1 + a_i a_j} \right)$$

for n *given scalars* $a_1, \cdots, a_n$.

7. Spectral Theory

Let A be an $n \times n$ matrix.

Definition 28. *Suppose that for a nonzero vector v and a scalar number λ,*

$$Av = \lambda v.$$

Then λ is called an eigenvalue of A and v is called an eigenvector of A corresponding to eigenvalue λ.

Let v be an eigenvector of A corresponding to eigenvalue λ. We have for any positive integer k,

$$A^k v = \lambda^k v.$$

And more generally, for any polynomial p,

$$p(A)\, v = p(\lambda)\, v.$$

Theorem 42. *λ is an eigenvalue of A if and only if*

$$\det(\lambda I - A) = 0.$$

The polynomial

$$p_A(\lambda) = \det(\lambda I - A)$$

is called the characteristic polynomial of the matrix A.

Theorem 43. *Eigenvectors of a matrix A corresponding to distinct eigenvalues are linearly independent.*

Proof. Let λ_k, $1 \leq k \leq K$ be distinct eigenvalues of A and v_k, $1 \leq k \leq K$ be corresponding eigenvectors. We prove by induction in K. The case $K = 1$ is trivial. Suppose the result holds for $K = N$. Now for $K = N + 1$, suppose for constants c_k, $1 \leq k \leq N + 1$, we have

$$\sum_{k=1}^{N+1} c_k v_k = 0. \tag{7.1}$$

Applying A, we have

$$\sum_{k=1}^{N+1} c_k \lambda_k v_k = 0$$

which implies

$$\sum_{k=1}^{N} c_k (\lambda_k - \lambda_{N+1})\, v_k = 0.$$

Since v_k, $1 \leq k \leq N$ are linearly independent and $\lambda_k - \lambda_{N+1} \neq 0$ for $1 \leq k \leq N$, we have $c_k = 0$, $1 \leq k \leq N$, and (7.1) implies $c_{N+1} = 0$. Hence the result holds for $K = N + 1$ too. □

Corollary 6. *If the characteristic polynomial p_A of the $n \times n$ matrix A has n distinct roots, then A has n linearly independent eigenvectors which forms a basis.*

Suppose A has n linearly independent eigenvectors v_k, $1 \leq k \leq n$ corresponding to eigenvalues λ_k, $1 \leq k \leq n$. Then A is similar to the diagonal matrix $\Lambda = [\lambda_1, \cdots, \lambda_n]$. Actually, let $S = (v_1, \cdots, v_n)$, we have

$$A = S \Lambda S^{-1}.$$

Example 11. *The Fibonacci sequence* $f_0, f_1, \cdots$ *is defined by the recurrence relation*

$$f_{n+1} = f_n + f_{n-1}$$

with the starting data $f_0 = 0$, $f_1 = 1$. *Let*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix},$$

we have

$$\begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} = A \begin{pmatrix} f_{n-1} \\ f_n \end{pmatrix}.$$

Simple computation yields

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2},$$
$$v_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.$$

Since

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} v_1 - \frac{1}{\sqrt{5}} v_2,$$

we have

$$\begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\lambda_1^n}{\sqrt{5}} v_1 - \frac{\lambda_2^n}{\sqrt{5}} v_2$$

which implies

$$f_n = \frac{\lambda_1^n}{\sqrt{5}} - \frac{\lambda_2^n}{\sqrt{5}}.$$

Theorem 44. *Denote by* λ_k, $1 \leq k \leq n$, *the eigenvalues of* A, *with the same multiplicity they have as roots of the characteristic equation of* A. *Then*

$$\sum_{k=1}^{n} \lambda_k = \operatorname{tr} A \text{ and } \prod_{k=1}^{n} \lambda_k = \det A.$$

Proof. The first identity follows from the expansion

$$\begin{aligned} p_A(\lambda) = \det(\lambda I - A) &= \sum_p \sigma(p) \prod_{k=1}^{n} \left(\lambda \delta_{p_k k} - a_{p_k k}\right) \\ &= \lambda^n - (\operatorname{tr} A) \lambda^{n-1} + \cdots. \end{aligned}$$

The second identity follows from

$$(-1)^n \prod_{k=1}^{n} \lambda_k = p_A(0) = \det(-A) = (-1)^n \det(A).$$

□

Theorem 45 (Spectral Mapping Theorem). *(a) Let* q *be any polynomial,* A *a square matrix,* λ *an eigenvalue of* A. *Then* $q(\lambda)$ *is an eigenvalue of* $q(A)$.
(b) Every eigenvalue of $q(A)$ *is of the form* $q(\lambda)$, *where* λ *is an eigenvalue of* A.

Proof. (a) Let v be the eigenvector of A corresponding to λ. Then

$$q(A)v = q(\lambda)v.$$

(b) Let μ be an eigenvalue of $q(A)$, then

$$\det(\mu I - q(A)) = 0.$$

Suppose the roots of $q(\lambda) - \mu = 0$ is given by λ_i, then

$$q(\lambda) - \mu = c\prod(\lambda - \lambda_i),$$

which implies

$$\prod \det(\lambda_i I - A) = 0.$$

Hence, for some λ_i, $\det(\lambda_i I - A) = 0$. Hence $\mu = q(\lambda_i)$ where λ_i is an eigenvalue of A. □

If in particular we take $q = p_A$, we conclude that all eigenvalues of $p_A(A)$ are zero. In fact a little more is true.

Theorem 46 (Cayley-Hamilton). *Every matrix A satisfies its own characteristic equation*

$$p_A(A) = 0.$$

Proof. Let

$$Q(s) = sI - A$$

and $P(s)$ defined as the matrix of cofactors of $Q(s)$, i.e.

$$P_{ij}(s) = (-1)^{i+j} D_{ji}(s)$$

where D_{ij} is the determinant of the ij-th minor of $Q(s)$. Then we have

$$P(s)Q(s) = (\det Q(s))I = p_A(s)I.$$

Since the coefficients of Q commutes with A, we have

$$P(A)Q(A) = p_A(A)I,$$

hence $p_A(A) = 0$. □

Lemma 3. *Let*

$$P(s) = \sum P_k s^k, P(s) = \sum Q_k s^k, R(s) = \sum R_k s^k$$

be polynomials in s where the coefficients P_k, Q_k and R_k are $n \times n$ matrices. Suppose that

$$P(s)Q(s) = R(s)$$

and matrix A commutes with each Q_k, then we have

$$P(A)Q(A) = R(A).$$

Definition 29. *A nonzero vector u is said to be a generalized eigenvector of A corresponding to eigenvalue λ, if*

$$(A - \lambda I)^m u = 0$$

for some positive integer m.

Theorem 47 (Spectral Theorem). *Let A be an $n \times n$ matrix with complex entries. Every vector in $\mathbb{C}^n$ can be written as a sum of eigenvectors of A, genuine or generalized.*

Proof. Let x be any vector; the $n+1$ vectors $x, Ax, A^2x, \cdots, A^n x$ must be linearly dependent; therefore there is a polynomial p of degree less than or equal to n such that

$$p(A)x = 0$$

We factor p and rewrite this as

$$\prod_j (A - r_j I)^{m_j} x = 0.$$

All invertible factors can be removed. The remaining r_j are all eigenvalues of A. Applying Lemma 4 to $p_j(s) = (s - r_j)^m$, we have a decomposition of x as a sum of generalized eigenvectors. □

Lemma 4. *Let p and q be a pair of polynomials with complex coefficients and assume that p and q have no common zero.*
(a) There are two polynomials a and b such that

$$ap + bq = 1.$$

(b) Let A be a square matrix with complex entries. Then

$$N_{pq} = N_p \oplus N_q.$$

Here N_p, N_q, and N_{pq} are the null spaces of $p(A)$, $q(A)$, and $p(A)q(A)$.
(c) Let p_k $k = 1, 2, \cdots, m$ be polynomials with complex coefficients and assume that p_k have no common zero. Then

$$N_{p_1 \cdots p_m} = \oplus_{k=1}^m N_{p_k}.$$

Here N_{p_k} is the null space of $p_k(A)$.

Proof. (a) Denote by $\mathcal{P}$ all polynomials of the form $ap + bq$. Among them there is one, nonzero, of lowest degree; call it d. We claim that d divides both p and q; for suppose not; then the division algorithm yields a remainder r, say

$$r = p - md.$$

Since p and d belong to $\mathcal{P}$ so does r; since r has lower degree than d, this is a contradiction. We claim that d has degree zero; for if it had degree greater than zero, it would, by the fundamental theorem of algebra, have a root. Since d divides p and q, and p and q have no common zero, d is a nonzero constant. Hence $1 \in \mathcal{P}$.

(b) From (a), There are two polynomials a and b such that

$$a(A)p(A) + b(A)q(A) = I.$$

For any x, we have

$$x = a(A)p(A)x + b(A)q(A)x \overset{\triangle}{=} x_2 + x_1.$$

Here it is easy to verify that if $x \in N_{pq}$ then $x_1 \in N_q$, $x_2 \in N_p$. Now suppose $x \in N_p \cap N_q$, the above formula implies

$$x = a(A)p(A)x + b(A)q(A)x = 0.$$

Hence $N_{pq} = N_p \oplus N_q$. □

We denote by $\mathcal{P}_A$ the set of all polynomials p which satisfy $p(A) = 0$. $\mathcal{P}_A$ forms a linear space. Denote by $m = m_A$ a nonzero polynomial of smallest degree in $\mathcal{P}_A$; we claim that all p in $\mathcal{P}_A$ are multiples of m. Except for a constant factor, which we fix so that the leading coefficient of m_A is 1, $m = m_A$ is unique. This polynomial is called the minimal polynomial of A.

To describe precisely the minimal polynomial we return to generalized eigenvector. We denote by $N_m = N_m(\lambda)$ the nullspace of $(A - \lambda I)^m$. The subspaces N_m, consist of generalized eigenvectors; they are indexed increasingly, that is,

$$N_1 \subset N_2 \subset N_3 \subset \cdots .$$

Since these are subspaces of a finite-dimensional space, they must be equal from a certain index on. We denote by $d = d(\lambda)$ the smallest such index, that is,

$$N_d = N_{d+k} \text{ for any } k \geq 1,$$
$$N_{d-1} \neq N_d.$$

$d(\lambda)$ is called the index of the eigenvalue λ.

Remark 11. *A maps N_d into itself. i.e., N_d is an invariant subspace under A.*

Theorem 48. *Let $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues of A with index $d_j = d(\lambda_j)$, $1 \leq j \leq k$. Then*
(1)

$$\mathbb{C}^n = \oplus_{j=1}^k N_{d_j}(\lambda_j),$$

(2) and the minimal polynomial

$$m_A = \prod_{j=1}^k (s - \lambda_j)^{d_j}.$$

Proof. (1). It follows from the spectral theorem and Lemma 4.
(2) Since

$$x = \sum_{j=1}^k x_j$$

where $x_j \in N_{d_j}(\lambda_j)$, we have

$$\prod_{j=1}^k (A - \lambda_j I)^{d_j} x = \sum_{j=1}^k \left(\prod_{i=1}^k (A - \lambda_i I)^{d_i} \right) x_j = 0.$$

Hence

$$\prod_{j=1}^k (A - \lambda_j I)^{d_j} = 0.$$

On the other hand, if

$$\prod_{j=1}^k (A - \lambda_j I)^{e_j} = 0$$

where $e_1 < d_1$ and $e_j \leq d_j$, we conclude

$$\mathbb{C}^n = \oplus_{j=1}^k N_{e_j}(\lambda_j)$$

which is impossible. □

Theorem 49. *Suppose the pair of matrices A and B are similar. Then A and B have the same eigenvalues: $\lambda_1, \cdots, \lambda_k$. Furthermore, the nullspaces*

$$N_m(\lambda_j) = \text{null space of } (A - \lambda_j I)^m,$$
$$M_m(\lambda_j) = \text{null space of } (B - \lambda_j I)^m$$

have for all j and m the same dimensions.

Proof. Let S be invertible such that

$$A = SBS^{-1},$$

then we have for any m and λ,

$$(A - \lambda I)^m = S(B - \lambda I)^m S^{-1}$$

□

Since $A - \lambda I$ maps N_{i+1} into N_i, $A - \lambda I$ defines a map from N_{i+1}/N_i into N_i/N_{i-1} for any $i \geq 1$ where $N_0 = \{0\}$.

Lemma 5. *The map*

$$A - \lambda I : N_{i+1}/N_i \to N_i/N_{i-1}$$

is one-to-one. Hence

$$\dim(N_{i+1}/N_i) \leq \dim(N_i/N_{i-1})$$

Proof. Let $B = A - \lambda I$. If $\left\{B\{x\}_{N_{i+1}/N_i}\right\}_{N_i/N_{i-1}} = \{0\}$, then $Bx \in N_{i-1}$, hence $x \in N_i$ and $\{x\}_{N_{i+1}/N_i} = \{0\}_{N_{i+1}/N_i}$. □

Next, we construct Jordan Canonical form of a matrix.

Let 0 be an eigenvalue of A. We want to construct a basis of $N_d = N_d(0)$.

Step I. Let $l_0 = \dim(N_d/N_{d-1})$, we construct $x_1, x_2, \cdots, x_{l_0}$ such that $\{x_1\}, \{x_2\}, \cdots, \{x_{l_0}\}$ form a basis of N_d/N_{d-1}.

Step II. Let $l_1 = \dim(N_{d-1}/N_{d-2}) \geq l_0$, we construct $Ax_1, Ax_2, \cdots, Ax_{l_0}, x_{l_0+1}, \cdots, x_{l_1}$ such that their quotient classes form a basis of N_{d-1}/N_{d-2}.

Step III. We continue this process until we reach N_1. We thus constructed a basis of N_d.

Step IV. It is illuminating to arrange the basis elements in a table:

$$\begin{array}{cccc}
x_1 & Ax_1 & \cdots & & A^{d-1}x_1 \\
\vdots & & & & \vdots \\
x_{l_0} & Ax_{l_0} & \cdots & & A^{d-1}x_{l_0} \\
x_{l_0+1} & Ax_{l_0+1} & \cdots & A^{d-2}x_{l_0+1} & \\
\vdots & & & & \\
x_{l_1} & Ax_{l_1} & \cdots & A^{d-2}x_{l_1} & \\
\vdots & & & & \\
x_{l_{d-2}+1} & & & & \\
\vdots & & & & \\
x_{l_{d-1}} & & & &
\end{array}$$

Noticing

$$\begin{aligned}\dim N_d =& \sum_{k=0}^{d-1} l_k \\ =& dl_0 + (d-1)(l_1 - l_0) + (d-2)(l_2 - l_1) + \cdots 1 \times (l_{d-1} - l_{d-2}) .\end{aligned}$$

Here in the above table, the span of the vectors in each row is invariant under A. And A restricted to each row has matrix representation of the form

$$J_m = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

which is called a Jordan block where

$$J_m(i,j) = 1 \text{ if } j = i+1 \text{ and } 0 \text{ otherwise.}$$

THEOREM 50. *Any matrix A is similar to its Jordan canonical form which consists diagonal blocks of the form*

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

where λ is the eigenvalue of A.

THEOREM 51. *If A and B have the same eigenvalues $\{\lambda_j\}$, and if*

$$\dim N_m(\lambda_j) = \dim M_m(\lambda_j)$$

holds for any m, j, then A and B are similar.

REMARK 12. *The dimension of $N_{d(\lambda)}(\lambda)$ equals the multiplicity of λ as the root of the characteristic equation of A.*

THEOREM 52. *Suppose*

$$AB = BA.$$

Then there is a basis in $\mathbb{C}^n$ which consists of eigenvectors and generalized eigenvectors of both A and B. Theorem remains true if A, B are replaced by any number of pairwise commuting linear maps.

PROOF. Let $\{\lambda_j\}_{j=1}^k$ be distinct eigenvalues of A, $d_j = d(\lambda_j)$ be the index and $N_j = N_{d_j}(\lambda_j)$ the null space of $(A - \lambda_j I)^{d_j}$. Then

$$\mathbb{C}^n = \oplus N_j.$$

For any x, we have

$$B (A - \lambda_j I)^{d_j} x = (A - \lambda_j I)^{d_j} Bx.$$

Hence $B : N_j \to N_j$. Applying the spectral decomposition thereom to $B : N_j \to N_j$, we can find a basis of N_j consists of eigenvectors and generalized eigenvectors of B. □

THEOREM 53. *Every square matrix A is similar to its transpose A^T.*

PROOF. Recall that
$$\dim N_A = \dim N_{A^T}.$$
Since the transpose of $A - \lambda I$ is $A^T - \lambda I$ it follows that A and A^T have the same eigenvalues. The transpose of $(A - \lambda I)^m$ is $\left(A^T - \lambda I\right)^m$; therefore their nullspaces have the same dimension. Hence A and A^T are similar. □

THEOREM 54. *Let λ, μ be two distinct eigenvalues of A. Suppose u is an eigenvector of A w.r.t. λ and suppose v is an eigenvector of A^T w.r.t. μ. Then $u^T v = 0$.*

PROOF. $v^T A u = u^T A^T v = \lambda v^T u = \mu u^T v.$ □

8. Euclidean Structure

DEFINITION 30. *A Euclidean structure in a linear space X over $\mathbb{R}$ is furnished by a real-valued function of two vector arguments called a scalar product and denoted as (x, y), which has the following properties:*
(i) (x, y) is a bilinear function; that is, it is a linear function of each argument when the other is kept fixed.
(ii) It is symmetric:
$$(x, y) = (y, x).$$
(iii) It is positive: $(x, x) > 0$ except for $x = 0$. The Euclidean length (or the norm) of x is defined by
$$\|x\| = \sqrt{(x, x)}.$$

A scalar product is also called an inner product, or a dot product.

DEFINITION 31. *The distance of two vectors x and y in a linear space with Euclidean norm is defined as $\|x - y\|$.*

THEOREM 55 (Schwarz Inequality). *For all x, y,*
$$|(x, y)| \leq \|x\| \|y\|.$$

PROOF. Consider the quadratic function
$$q(t) = \|x + ty\|^2.$$
□

Suppose $x, y \neq 0$, we can define the angle θ between x, y by
$$\cos\theta = \frac{(x, y)}{\|x\| \|y\|}.$$

COROLLARY 7.
$$\|x\| = \max_{\|y\|=1} (x, y).$$

THEOREM 56 (Triangle inequality).
$$\|x + y\| \leq \|x\| + \|y\|.$$

PROOF.
$$(x + y, x + y) = \|x\|^2 + 2(x, y) + \|y\|^2 \leq (\|x\| + \|y\|)^2.$$
□

DEFINITION 32. *Two vectors x and y are called orthogonal (perpendicular), denoted by $x \perp y$ if $(x, y) = 0$.*

Pythagorian theorem

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2$$

holds iff $x \perp y$.

DEFINITION 33. *Let X be a finite-dimensional linear space with a Eulerian structure. A basis $x_1, \cdots, x_n$ is called orthononnal if*

$$(x_j, x_k) = \delta_{jk},\ 1 \leq j, k \leq n.$$

THEOREM 57 (Gram-Schmidt). *Given an arbitrary basis $y_1, \cdots, y_n$ in a finite dimensional linear space equipped with a Euclidean structure, there is a related orthonormal basis $x_1, \cdots, x_n$ such that x_k is a linear combination of $y_1, \cdots, y_k$ for all k.*

PROOF. Let

$$x_1 = \frac{y_1}{\|y_1\|}.$$

Suppose $x_1, \cdots, x_k$ is defined, we define

$$x_{k+1} = \frac{y_{k+1} - \sum_{j=1}^{k} (x_j, y_{k+1}) x_j}{\left\| y_{k+1} - \sum_{j=1}^{k} (x_j, y_{k+1}) x_j \right\|}.$$

□

Under an orthonormal basis $x_1, \cdots, x_n$, suppose $x = \sum a_j x_j$, we have

$$a_j = (x, x_j).$$

Suppose $x = \sum a_j x_j$ and $y = \sum b_j x_j$, we have

$$(x, y) = \sum_{j=1}^{n} a_j b_j.$$

In particular,

$$\|x\|^2 = \sum_{j=1}^{n} a_j^2.$$

The mapping

$$x \mapsto (a_1, \cdots, a_n)$$

carries the space X with a Euclidean structure into $\mathbb{R}^n$ and carries the scalar product of X into the standard scalar product of $\mathbb{R}^n$.

Since the scalar product is bilinear, for y fixed (x, y) is a linear function of x. Conversely, we have the following theorem.

THEOREM 58. *Every linear function $l(x)$ on a finite-dimensional linear space X with Euclidean structure can be written in the form*

$$l(x) = (x, y)$$

where y is some element of X. The mapping $l \mapsto y$ is an isomorphism of the Euclidean space X with its dual.

DEFINITION 34. *Let X be a finite-dimensional linear space with Euclidean structure, Y a subspace of X. The orthogonal complement of Y, denoted as $Y^{\perp}$, consists of all vectors z in X that are orthogonal to every y in Y.*

Recall that we denoted by $Y^{\perp}$ the set of linear functionals that vanish on Y. The notation $Y^{\perp}$ introduced above is consistent with the previous notation when the dual of X is identified with X. In particular, $Y^{\perp}$ is a subspace of X.

THEOREM 59. *For any subspace Y of X,*

$$X = Y \oplus Y^{\perp}.$$

In the decomposition $x = y + y^{\perp}$, the component y is called the orthogonal projection of x into Y, denoted by

$$y = P_Y x.$$

THEOREM 60. *The mapping P_Y is linear and*

$$P_Y^2 = P_Y.$$

THEOREM 61. *Let Y be a linear subspace of the Euclidean space X, x some vector in X. Then among all elements z of Y, the one closest in Euclidean distance to x is $P_Y x$.*

We turn now to linear mappings of a Euclidean space X into another Euclidean space U. Since a Euclidean space can be identified in a natural way with its own dual, the transpose of a linear map A of such a space X into U maps U into X. To indicate this distinction, the transpose of a map A of Euclidean X into U is called the adjoint of A and is denoted by A^*.

Here is the full definition of the adjoint A^*:

$$\begin{aligned} A' &: U' \to X', \\ (A'T)\, x &= T\,(Ax) \\ T &\to y; A'T \to A^* y \\ T\,(Ax) &= (Ax, y) \\ (A'T)\, x &= (x, A^* y) \\ A^* &: U \to X \\ (x, A^* y) &= (Ax, y) \end{aligned}$$

THEOREM 62. *(i) If A and B are linear mappings of X into U, then*

$$(A+B)^* = A^* + B^*.$$

(ii) If A is a linear map of X into U, while C is a linear map of U into V, then

$$(CA)^* = A^* C^*.$$

(iii) If A is a 1-to-1 mapping of X onto U, then

$$\left(A^{-1}\right)^* = \left(A^*\right)^{-1}.$$

(iv) $(A^*)^* = A$.

PROOF. (iii) follows from (ii) applied to $A^{-1}A = I$ and the observation that $I^* = I$. □

When we take X to be $\mathbb{R}^n$ and U to be $\mathbb{R}^m$ with their standard Euclidean structures, and interpret A and A^* as matrices, they are transposes of each other.

There are many situations where quantities $x_1, \cdots, x_n$ cannot be measured directly, but certain linear combinations of them can. Suppose that m such linear combinations have been measured. We can put all this information in the form of a matrix equation

$$Ax = p$$

where $p_1, \cdots, p_m$ are the measured values, and A is an $m \times n$ matrix. When $m > n$, such a system of equations is overdetermined and in general does not have a solution. In such a situation, we seek that vector x that comes closest to satisfying all the equations in the sense that makes $\|Ax - p\|^2$ as small as possible.

THEOREM 63. *Let A be an $m \times n$ matrix, $m > n$, and suppose that A has only the trivial nullvector* 0. *The vector x that minimizes $\|Ax - p\|^2$ is the unique solution z of*

$$A^*Az = A^*p.$$

REMARK 13. *If A has a nontrivial nullvector, then the minimizer can't be unique.*

PROOF. The equation has a unique solution iff the null space of A^*A is $\{0\}$. Suppose $A^*Ay = 0$, then

$$0 = (A^*Ay, y) = \|Ay\|^2,$$

hence y is a nullvector of A and $y = 0$.

The range of A is an n dimensional subspace. Let z be the minimizer, $Az - p$ should be perpendicular to R_A, i.e., for any $y \in \mathbb{R}^n$

$$(Az - p, Ay) = 0,$$

$$(A^*Az - A^*p, y) = 0.$$

Hence we conclude

$$A^*Az - A^*p = 0.$$

□

THEOREM 64.

$$P_Y^* = P_Y.$$

PROOF.

$$(x_1, P_Y x_2) = (P_Y x_1, P_Y x_2) = (P_Y x_1, x_2).$$

□

DEFINITION 35. *A mapping M of a Euclidean space into itself is called an isometry if it preserve the distance of any pair of points.*

Any isometry is the composite of one that maps zero to zero and a translation.

THEOREM 65. *Let M be an isometric mapping of a Euclidean space into itself that maps zero to zero. Then*
(i) M is linear.
(ii) M is invertible and its inverse is an isometry.
(iii) $M^*M = I$ and $\det M = \pm 1$.
Conversely, if M is a linear mapping satisfying $M^*M = I$, then M is an isometry.

Proof. (i). Writing $Mx = x'$. Then we have for any x,

$$\|x'\| = \|x\| .$$

Since

$$\|x' - y'\| = \|x - y\| ,$$

we have

$$(x', y') = (x, y) ,$$

i.e., M preserves the scalar product. Let $z = x + y$, since

$$\|z' - x' - y'\| = \|z - x - y\| = 0,$$

we conclude $z' = x' + y'$. Let $z = cy$, we have

$$\|z' - cy'\| = \|z - cy\| = 0,$$

hence, $z' = cy'$.

(ii) Trivial.

(iii) For any x, y,

$$(x, y) = (Mx, My) = (x, M^*My) ,$$

hence $M^*M = I$. $\det M = \pm 1$ follows from $\det M^* = \det M$.

Conversely, if $M^*M = I$, then for any x, y,

$$(Mx, My) = (x, M^*My) = (x, y)$$

hence M preserves the scalar product and distance. □

Remark 14. *A mapping that preserves distances also preserves volume.*

Definition 36. *A matrix that maps $\mathbb{R}^n$ onto itself isometrically is called orthogonal.*

The orthogonal matrices of a given order form a group under matrix multiplication. Clearly, composites of isometries are isometric, and so are their inverses.

The orthogonal matrices whose determinant is plus 1 form a subgroup, called the special orthogonal group.

Proposition 3. *A matrix M is orthogonal iff its columns are pairwise orthogonal unit vectors.*

Proposition 4. *A matrix M is orthogonal if its rows are pairwise orthogonal unit vectors.*

Proof. $M^* = M^{-1}$. □

An orthogonal matrix maps an orthonormal basis onto another orthonormal basis.

We use the norm to measure the size of a linear mapping A of one Euclidean space X into another Euclidean space U.

Definition 37. *The norm of A is defined by*

$$\|A\| = \sup_{\|x\|=1} \|Ax\| .$$

THEOREM 66. *Let A be a linear mapping from the Euclidean space X into the Euclidean space U. Then*
(i) For any $z \in X$,

$$\|Az\| \leq \|A\| \, \|z\| .$$

(ii)

$$\|A\| = \sup_{\|x\|=1, \|z\|=1} (x, Az) .$$

REMARK 15. $|a_{ij}| \leq \|A\|$.

THEOREM 67. *(i) For any scalar k,*

$$\|kA\| = |k| \, \|A\| .$$

(ii) For any pair of linear mappings A and B of X into U,

$$\|A + B\| \leq \|A\| + \|B\| .$$

(iii) Let A be a linear mapping of X into U, and let C be a linear mapping of U into V; then

$$\|CA\| \leq \|C\| \, \|A\| .$$

(iv) $\|A^*\| = \|A\|$.

THEOREM 68. *Let A be a linear mapping of a finite-dimensional Euclidean space X into itself that is invertible. Denote by B another linear mapping of X into X close to A in the sense of the following inequality:*

$$\|A - B\| < \frac{1}{\|A^{-1}\|}.$$

Then B is invertible.

PROOF. Denote $A - B = C$, so that $B = A - C$. Factor B as

$$B = A(I - A^{-1}C) = A(I - S)$$

where $S = A^{-1}C$. It suffices to show that $I - S$ is invertible which is equivalent to show that the nullspace of $I - S$ is trivial. Suppose that $(I - S)x = 0$, i.e., $Sx = x$ for some $x \neq 0$. Hence $\|S\| \geq 1$. On the other hand,

$$\|S\| \leq \left\|A^{-1}\right\| \|C\| < 1,$$

contradiction. □

Recall convergence and Cauchy sequence for real numbers. A basic property of real numbers is that every Cauchy sequence of numbers converges to a limit. This property of real numbers is called completeness. A second basic notion about real numbers is local compactness: Every bounded sequence of real numbers contains a convergent subsequence.

DEFINITION 38. *A sequence of vectors $\{x_k\}$ in a linear space X with Euclidean structure converges to the limit x:*

$$\lim_{k \to \infty} x_k = x$$

if $\|x_k - x\|$ tends to zero as $k \to \infty$.

THEOREM 69. *A sequence of vectors $\{x_k\}$ in a Euclidean space X is called a Cauchy sequence if $\|x_k - x_j\| \to 0$ as k and $j \to \infty$.*
(i) Every Cauchy sequence in a finite-dimensional Euclidean space converges to a limit.
(ii) In a finite-dimensional Euclidean space every bounded sequence contains a convergent subsequence.

COROLLARY 8.

$$\|A\| = \max_{\|x\|=1} \|Ax\| .$$

THEOREM 70. *Let X be a linear space with a Euclidean structure, and suppose that it is locally compact-that is, that every bounded sequence $\{x_k\}$ of vectors in X has a convergent subsequence. Then X is finite dimensional.*

DEFINITION 39. *A sequence $\{A_n\}$ of mappings converges to a limit A if*

$$\lim_{n\to\infty} \|A_n - A\| = 0.$$

PROPOSITION 5. *In finite dimensional spaces, $\{A_n\}$ converges to A iff for all x, $A_n x$ converges to Ax.*

PROOF. To show A_n is bounded, let e_i be a basis,

$$A_n e_i \text{ is bounded,}$$

For any x, we have

$$\|Ax\| = \left\|\sum (x, e_i) A_n e_i\right\| \le C \left|\sum (x, e_i)\right| \le C.$$

□

Complex Euclidean structure:

DEFINITION 40. *A complex Euclidean structure in a linear space X over $\mathbb{C}$ is furnished by a complex valued function of two vector arguments, called a scalar product and denoted as (x, y), with these properties:*
(i) (x, y) is a linear function of x for y fixed.
(ii) Conjugate symmetry: for all x, y,

$$(x, y) = \overline{(y, x)}.$$

(iii) Positivity: $(x, x) > 0$ for all $x \neq 0$

REMARK 16. *For x fixed, (x, y) is a skew linear function of y. i.e.,*

$$(x, y_1) + (x, y_2) = (x, y_1 + y_2),$$
$$(x, ky) = \bar{k}(x, y).$$

The norm in X is still defined by

$$\|x\|^2 = (x, x).$$

From

$$\|x + y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\,(x, y) + \|y\|^2,$$

one can prove the Schwarz inequality

$$|(x, y)| \le \|x\| \|y\| .$$

Triangle inequality

$$\|x+y\| \le \|x\| + \|y\|$$

thus follows.

We define the adjoint A^* of a linear map A of an abstract complex Euclidean space into itself by relation

$$(x, A^* y) = (Ax, y)$$

as before. Then $A^* = \overline{A^T}$.

We define isometric maps of a complex Euclidean space as in the real case.

DEFINITION 41. *A linear map of a complex Euclidean space into itself that is isometric is called unitary.*

PROPOSITION 6. *M is a unitary map if and only if*

$$M^* M = I.$$

For a unitary map M, $|\det M| = 1$.

Recall that the norm of A is defined by

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

THEOREM 71. *Let $A : \mathbb{R}^n \to \mathbb{R}^m$ or $A : \mathbb{C}^n \to \mathbb{C}^m$ be represented by $A = (a_{ij})_{m\times n}$, then*

$$\|A\| \le \left(\sum |a_{ij}|^2\right)^{\frac{1}{2}}.$$

PROOF. For any $x \in \mathbb{C}^n$,

$$\begin{aligned}
\|Ax\|^2 &= \sum_i \left|\sum_j a_{ij} x_j\right|^2 \\
&\le \sum_i \left(\left(\sum_j |a_{ij}|^2\right)\left(\sum_j |x_j|^2\right)\right) \\
&= \left(\sum_{i,j} |a_{ij}|^2\right)\|x\|^2.
\end{aligned}$$

□

The quantity $\left(\sum |a_{ij}|^2\right)^{\frac{1}{2}}$ is called the Hilbert-Schmidt norm of the matrix A. We have

$$\sum |a_{ij}|^2 = \operatorname{tr} A^* A = \operatorname{tr} AA^*.$$

DEFINITION 42. *The spectral radius $r(A)$ of a linear mapping A of a linear space into itself is*

$$r(A) = \max |\lambda_j|$$

where the λ_j range over all eigenvalues of A.

PROPOSITION 7.

$$\|A\| \ge r(A).$$

$$A = SBS^{-1}$$
$$\|Ax\| = \left\|SBS^{-1}x\right\| \le \|S\| \left\|S^{-1}\right\| \|B\| \|x\|$$

On the other hand, we have

THEOREM 72. *Gelfand's formula*

$$r(A) = \lim_{k\to\infty} \left(\left\|A^k\right\|\right)^{\frac{1}{k}}.$$

LEMMA 6. *If* $r(A) < 1$, *then*

$$\lim_{k\to\infty} A^k = 0.$$

PROOF. For a $s \times s$ Jordan block,

$$J_s(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix},$$

we have

$$J^k = \begin{pmatrix} \lambda^k & \binom{k}{1}\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \cdots & \binom{k}{s+1}\lambda^{k-s+1} \\ 0 & \lambda^k & \binom{k}{1}\frac{k}{1}\lambda^{k-1} & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \binom{k}{2}\lambda^{k-2} \\ \vdots & \ddots & \ddots & \lambda^k & \binom{k}{1}\lambda^{k-1} \\ 0 & \cdots & 0 & 0 & \lambda^k \end{pmatrix}.$$

If $|\lambda| < 1$,

$$\lim J_s^k(\lambda) = 0.$$

Let $J = S^{-1}AS$ be the Jordan canonical form of A, we have

$$\lim J^k = 0$$

if $r(A) < 1$. Hence

$$\lim A^k = \lim SJ^kS^{-1} = 0.$$

□

Now we prove the Gelfand's formula.

PROOF.
$$\left\|A^k\right\| \ge r\left(A^k\right) = (r(A))^k.$$

Hence,

$$r(A) \le \liminf_{k\to\infty} \left(\left\|A^k\right\|\right)^{\frac{1}{k}}.$$

To show

$$r(A) \ge \limsup_{k\to\infty} \left(\left\|A^k\right\|\right)^{\frac{1}{k}},$$

we first assume that A is in its Jordan form. Let

$$A_\varepsilon = \frac{A}{r(A) + \varepsilon},$$

$$\left\| A_\varepsilon^k \right\| < 1$$

implies

$$\left\| A^k \right\| < \left(r\left(A \right) + \varepsilon \right)^k,$$

hence

$$\limsup_{k \to \infty} \left(\left\| A^k \right\| \right)^{\frac{1}{k}} \leq r\left(A \right) + \varepsilon.$$

□

COROLLARY 9. *Suppose* $A_1 A_2 = A_2 A_1$*, then*

$$r\left(A_1 A_2 \right) \leq r\left(A_1 \right) r\left(A_2 \right).$$

PROOF. $r\left(A_1 A_2 \right) \leq \left\| \left(A_1 A_2 \right)^k \right\|^{\frac{1}{k}} \leq \left\| A_1^k \right\|^{\frac{1}{k}} \left\| A_1^k \right\|^{\frac{1}{k}}$
Letting $k \to \infty$, we have

$$r\left(A_1 A_2 \right) \leq r\left(A_1 \right) r\left(A_2 \right).$$

Without the assumption,

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$A_1 A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, A_2 A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

□

Let T denote a 3×3 matrix, its columns x, y, and z:

$$T = (x, y, z)$$

For fixed x, y, $\det T$ is linear in z, hence, there exists w such that

$$\det T = (w, z).$$

It is easy to check that w is a bilinear function of x and y, and we write

$$w = x \times y.$$

9. Spectral Theory of Self-Adjoint Mappings of a Euclidean Space into Itself

A linear mappings A of a real or complex Euclidean space into itself is said to be self-adjoint if

$$A^* = A.$$

Recall

$$(Ax, y) = (x, A^* y) = (x, Ay).$$

When A acts on a real Euclidean space, any matrix representing it in an orthonormal system of coordinates is symmetric, i.e.

$$a_{ji} = a_{ij}.$$

Similarly, when A acts on a complex Euclidean space, any matrix representing it in an orthonormal system of coordinates is hermitian, i.e.,

$$a_{ji} = \overline{a_{ij}}.$$

THEOREM 73. *A self-adjoint map H of complex Euclidean space X into itself has real eigenvalues and a set of eigenvectors that form an orthonormal basis of X.*

PROOF. The eigenvectors and generalized eigenvectors of H span X. So it suffices to show:
(a) H has only real eigenvalues.

$$Hx = \lambda x,$$

$(Hx, x) = (x, Hx)$

$$\lambda = \bar{\lambda}.$$

(b) H has no generalized eigenvectors, only genuine ones. WLOG, suppose $\lambda = 0$ is an eigenvalue, suppose $H^2 x = 0$, then we have

$$0 = \left(H^2 x, x\right) = (Hx, Hx),$$

hence $Hx = 0.$, i.e., $N_1(\lambda) = N_2(\lambda)$.

(c) Eigenvectors of H corresponding to different eigenvalues are orthogonal. $Hx = \lambda x, Hy = \mu y,$

$$(Hx, y) = (x, Hy)$$

implies

$$(\lambda - \mu)(x, y) = 0.$$

□

COROLLARY 10. *Any Hermitian matrix can be diagonalized by a unitary matrix.*

PROOF. Let $U = (x_1, \cdots, x_n)$ where $\{x_i\}$ is an orthonormal basis consists of eigenvectors of H. Suppose $Hx_i = \lambda_i x_i$, then we have

$$HU = (\lambda_1 x_1, \cdots, \lambda_n x_n) = U[\lambda_1, \cdots, \lambda_n].$$

i.e.,

$$H = U\Lambda U^*.$$

□

THEOREM 74. *A self-adjoint map H of real Euclidean space X into itself has real eigenvalues and a set of eigenvectors that form an orthonormal basis of X.*

PROOF. Pick an orthonormal basis of X and let A be the matrix representing H. Then A is a real symmetric matrix and we have $A : \mathbb{R}^n \to \mathbb{R}^n$ if we identify X with $\mathbb{R}^n$. We now extend A naturally as a map $\tilde{A} : \mathbb{C}^n \to \mathbb{C}^n$. Then $\tilde{A}$ is hermitian and Theorem 73 applies. We claim $\sigma(A) = \sigma\left(\tilde{A}\right)$ and

$$N_A(\lambda) = \operatorname{Re} N_{\tilde{A}}(\lambda).$$

Since

$$\dim_{\mathbb{R}} N_A(\lambda) = \dim_{\mathbb{C}} N_{\tilde{A}}(\lambda),$$

we conclude

$$\mathbb{R}^n = \oplus_{\lambda \in \sigma(A)} N_A(\lambda)$$

and we can find a set of eigenvectors that form an orthonormal basis of $\mathbb{R}^n$. □

COROLLARY 11. *Any symmetric matrix can be diagonalized by an orthogonal matrix.*

PROOF. Let $Q = (x_1, \cdots, x_n)$ where $\{x_i\}$ is an orthonormal basis consists of eigenvectors of H. Suppose $Hx_i = \lambda_i x_i$, then we have

$$HU = (\lambda_1 x_1, \cdots, \lambda_n x_n) = Q[\lambda_1, \cdots, \lambda_n].$$

i.e.,

$$H = Q\Lambda Q^*.$$

□

Next, we consider the real quadratic function

$$q(y) = h_{ij} y_i y_j = (y, Hy)$$

where $h_{ij} = h_{ji}$. If we change variable $x = Ly$, then

$$q\left(L^{-1}x\right) = q(y) = (y, Hy) = \left(L^{-1}x, HL^{-1}x\right) = \left(x, \left(L^{-1}\right)^T HL^{-1}x\right) = (x, Mx)$$

where

$$H = L^T ML.$$

DEFINITION 43. *Two symmetric matrices A and B are called congruent if there exists an invertible matrix S such that*

$$A = S^T BS.$$

THEOREM 75. *There exists an invertible matrix L, such that*

$$q\left(L^{-1}x\right) = \sum d_i x_i^2$$

for some constants d_i. i.e., any symmetric matrix is congruent to a diagonal matrix.

PROOF.

$$\begin{aligned} q(y) &= (y, Hy) = (y, Q\Lambda Q^* y) \\ &= (Q^* y, \Lambda Q^* y). \end{aligned}$$

Let $x = Q^* y$, we have $y = Qx$, and

$$q(Qx) = (x, \Lambda x) = \sum_i \lambda_i y_i^2.$$

□

REMARK 17. *We could choose L so that $d_i = 0$ or ± 1.*

Apparently such invertible matrices are not unique. However, we have

THEOREM 76 (Sylvester's Law of Inertia). *Let L be an invertible matrix such that*

$$q\left(L^{-1}x\right) = \sum d_i x_i^2$$

for some constants d_i. Then the number of positive, negative, and zero terms d_i equal to the number of positive, negative, and zero eigenvalues of H.

PROOF. We denote by p_+, p_-, and p_0 the number of terms d_i that are positive, negative, and zero, respectively. Let S be a subspace, we say that q is positive on the subspace S if $q(u) > 0$ for every u in S. We claim that

$$p_+ = \max_{q>0 \text{ on } S} \dim S.$$

The right hand side is independent of the choice of coordinates, so we could choose the coordinate x_i such that

$$q = \sum d_i x_i^2.$$

We also assume that $d_1, \cdots, d_{p_+}$ are positive and the rest nonpositive. It is obvious

$$p_+ \leq \max_{q>0 \text{ on } S} \dim S.$$

We claim that the equality holds. Let S be any subspace whose dimension exceeds p_+. For any vector u in S, define Pu as the vector whose x_i components are the same as the first p_+ components of u, and the rest of the components are zero. The dimension p_+ of the target space of this map is smaller than the dimension of the domain space S. Hence there is a nonzero vector v in the nullspace of P. By definition of P, the first p+ of the z-components of this vector are zero. But then it follows that $q(y) < 0$; this shows that q is not positive on S. □

Definition 44. *The set of eigenvalues of H is called the spectrum of H.*

The spectral theorem for self-adjoint maps asserts that the whole space X can be decomposed as the direct sum of pairwise orthogonal eigenspaces:

$$X = \oplus_{j=1}^{k} N(\lambda_j).$$

where λ_j, $1 \leq j \leq k$ are distinct eigenvalues of H. That means that each $x \in X$ can be decomposed uniquely as the sum

$$x = \sum x_j$$

where $x_j \in N(\lambda_j)$. Applying H, we have

$$Hx = \sum \lambda_j x_j.$$

Let P_j be the orthogonal projection of X onto $N(\lambda_j)$, we have

$$I = \sum P_j$$

and

$$H = \sum \lambda_j P_j.$$

Theorem 77. *(a)*

$$P_j P_k = 0 \text{ for } j \neq k \text{ and } P_j^2 = P_j.$$

(b) Each P_j is self-adjoint.

Definition 45. *A decomposition of the form*

$$I = \sum P_j$$

where the P_j satisfy

$$P_j P_k = 0 \text{ for } j \neq k \text{ and } P_j^2 = P_j.$$

is called a resolution of the identity.

Definition 46. *Let H be a self-adjoint map,*

$$H = \sum \lambda_j P_j.$$

is called the spectral resolution of H.

THEOREM 78. *Let X be a complex Euclidean space, $H : X \to X$ a self-adjoint linear map. Then there is a resolution of the identity*

$$I = \sum P_j$$

in the sense

$$P_j P_k = 0 \text{ for } j \neq k \text{ and } P_j^2 = P_j, P_j^* = P_j$$

which gives a spectral resolution

$$H = \sum \lambda_j P_j.$$

of H.

We have for any polynomial,

$$p(H) = \sum p(\lambda_j) P_j$$

Let $f(\lambda)$ be any real valued function defined on the spectrum of H. We define

$$f(H) = \sum f(\lambda_j) P_j.$$

For example,

$$e^H = \sum e^{\lambda_j} P_j.$$

THEOREM 79. *Let H, K be two self-adjoint matrices that commute. Then they have a common spectral resolution. i.e., there exists a resolution of the identity*

$$I = \sum P_j$$

such that

$$H = \sum \lambda_j P_j,$$
$$K = \sum \mu_j P_j.$$

DEFINITION 47. *A linear mapping A of Euclidean space into itself is called anti-self-adjoint if $A^* = -A$.*

If A is anti-self-adjoint, iA is self-adjoint.

THEOREM 80. *Let A be an anti-self-adjoint mapping of a complex Euclidean space into itself. Then*
(a) The eigenvalues of A are purely imaginary.
(b) We can choose an orthonormal basis consisting of eigenvectors of A.

We introduce now a class of maps that includes self-adjoint, anti-self-adjoint, and unitary maps as special cases.

DEFINITION 48. *A mapping N of a complex Euclidean space into itself is called normal if it commutes with its adjoint:*

$$N^* N = N N^*.$$

THEOREM 81. *A normal map N has an orthonormal basis consisting of eigenvectors.*

PROOF. If N and N^* commute, so do

$$H = \frac{N + N^*}{2}$$

and

$$A = \frac{N - N^*}{2}$$

Clearly, H is adjoint and A is anti-self-adjoint. According to Theorem 6 applied to H and $K = iA$, they have a common spectral resolution, so that there is an orthonormal basis consisting of common eigenvectors of both H and A. But since

$$N = H + A, N^* = H - A,$$

it follows that these are also eigenvectors of N as well as of N^*. □

THEOREM 82. *Let U be a unitary map of a complex Euclidean space into itself, that is, an isometric linear map.*
(a) There is an orthonormal basis consisting of genuine eigenvectors of U.
(b) The eigenvalues of U are complex numbers of absolute value $= 1$.

Let H be self-adjoint. The quotient

$$R_H(x) = \frac{(x, Hx)}{(x, x)}$$

is called the Rayleigh quotient of H.

R_H is a homogeneous function of x of degree zero, i.e., for any scalar k,

$$R_H(kx) = R_H(x).$$

We can view R_H as a real continuous map on the unit sphere. Hence, R_H has a maximum and a minimum. Let x with $\|x\| = 1$ be a critical point, we then have for any y,

$$\begin{aligned}
&\left.\frac{d}{dt}\right|_{t=0} R_H(x + ty) \\
&= \left.\frac{d}{dt}\right|_{t=0} \frac{(x, Hx) + 2t\,\mathrm{Re}\,(x, Hy) + t^2 (y, Hy)}{\|x\|^2 + 2t\,\mathrm{Re}\,(x, y) + t^2 \|y\|^2} \\
&= \frac{2\,\mathrm{Re}\,(x, Hy)}{\|x\|^2} - \frac{(x, Hx)\, 2\,\mathrm{Re}\,(x, y)}{\|x\|^4} \\
&= 2\,\mathrm{Re}\,(x, Hy) - (x, Hx)\, 2\,\mathrm{Re}\,(x, y) \\
&= 2\,\mathrm{Re}\,(Hx, y) - 2\,\mathrm{Re}\,((x, Hx)\, x, y) = 0.
\end{aligned}$$

Hence, we have

$$Hx = (x, Hx)\, x.$$

i.e., x is an eigenvector of H with eigenvalue $R_H(x) = (x, Hx)$.

Actually,

$$\max_{x \neq 0} R_H(x) = \max_{\lambda \in \sigma(H)} \lambda \text{ and } \min R_H(x) = \min_{\lambda \in \sigma(H)} \lambda.$$

More generally, every eigenvector x of H is a critical point of R_H; that is, the first derivatives of $R_H(x)$ are zero when x is an eigenvector of H. Conversely, the eigenvectors are the only critical points of $R_H(x)$. The value of the Rayleigh quotient at an eigenvector is the corresponding eigenvalue of H.

REMARK 18. *Rayleigh quotient gives a direct proof of the spectral decomposition theorem for real or complex self-adjoint maps.*

More generally, we have

THEOREM 83 (Minimax principle). *Let H be a self-adjoint map of a Euclidean space X of finite dimension. Denote the eigenvalues of H, arranged in increasing order, by $\lambda_1, \cdots, \lambda_n$. Then*

$$\lambda_j = \min_{\dim S = j} \max_{x \in S, x \neq 0} \frac{(x, Hx)}{(x, x)}$$

where S denotes linear subspaces of X.

PROOF. Let x_j be an orthonormal basis of X such that $Hx_j = \lambda_j x_j$. For any

$$x = \sum c_j x_j,$$

we have

$$\frac{(x, Hx)}{(x, x)} = \frac{\sum |c_j|^2 \lambda_j}{\sum |c_j|^2}.$$

For any $1 \leq j \leq k$, let S_j be the subspace spanned by $x_1, \cdots, x_j$, then

$$\min_{\dim S = j} \max_{x \in S, x \neq 0} \frac{(x, Hx)}{(x, x)} \leq \max_{x \in S_j, x \neq 0} \frac{(x, Hx)}{(x, x)} \leq \lambda_j.$$

On the other hand, for any S with $\dim S = j$, we consider the projection of S onto S_{j-1}. The null space has dimension at least 1, so we can pick $x^* \neq 0$, such that

$$x^* = \sum_{k \geq j} c_k x_k.$$

Hence,

$$\max_{x \in S, x \neq 0} \frac{(x, Hx)}{(x, x)} \geq \frac{(x^*, Hx^*)}{(x^*, x^*)} = \frac{\sum_{k \geq j} |c_k|^2 \lambda_k}{\sum_{k \geq j} |c_k|^2} \geq \lambda_j.$$

Since S is arbitrary j dimensional subspace, we have

$$\min_{\dim S = j} \max_{x \in S, x \neq 0} \frac{(x, Hx)}{(x, x)} \geq \lambda_j.$$

□

DEFINITION 49. *A self-adjoint mapping M of a Euclidean space X into itself is called positive if for all nonzero x in X,*

$$(x, Mx) > 0.$$

Let H, M be two self-adjoint mappings and M is positive. We consider a generalization of the Rayleigh quotient:

$$R_{H,M}(x) = \frac{(x, Hx)}{(x, Mx)}.$$

THEOREM 84. *Let X be a finite-dimensional Euclidean space, let H and M be two self-adjoint mappings of X into itself, and let M be positive. Then there exists a basis $x_1, \cdots, x_n$ of X where each x_j satisfies an equation of the form*

$$Hx_j = \lambda M x_j$$

where λ *is real, and*

$$(x_i, Mx_j) = 0$$

for $i \neq j$.

THEOREM 85. *Let* H *and* M *be self-adjoint,* M *positive. Then all the eigenvalues* $M^{-1}H$ *are real. If* H *is also positive, all eigenvalues of* $M^{-1}H$ *are positive.*

We recall the definition of the norm of a linear mapping A of a Euclidean space X into itself

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

When the mapping is normal, we have

THEOREM 86. *Suppose* N *is a normal mapping of a Euclidean space* X *into itself. Then*

$$\|N\| = r(N).$$

PROOF. Let x_j be an orthonormal basis of X such that $Nx_j = \lambda_j x_j$. Then we have for any $x = \sum c_j x_j \neq 0$,

$$\frac{\|Nx\|}{\|x\|} = \sqrt{\frac{\sum |\lambda_j|^2 |c_j|^2}{\sum |c_j|^2}} \leq r(N).$$

□

THEOREM 87. *Let* A *be a linear mapping of a finite-dimensional Euclidean space* X *into another finite-dimensional Euclidean space* U. *The norm*

$$\|A\| = \sqrt{r(A^*A)}.$$

PROOF.

$$\|Ax\|^2 = (Ax, Ax) = (x, A^*Ax) \leq \|x\| \|A^*Ax\| \leq \|A^*A\| \|x\|^2 = r(A^*A) \|x\|^2,$$

hence,

$$\|A\| \leq \sqrt{r(A^*A)}.$$

On the other hand, let λ be an eigenvalue of A^*A such that $\lambda = r(A^*A) \geq 0$, and x be corresponding eigenvector, we have

$$\|Ax\|^2 = (Ax, Ax) = (x, A^*Ax) = \lambda \|x\|^2 = r(A^*A) \|x\|^2.$$

Hence,

$$\|A\| = \sqrt{r(A^*A)}.$$

□

10. Schur Decomposition

THEOREM 88. *Let* A *be an* $n \times n$ *matrix. There exists a unitary matrix* U *and an upper triangular matrix, such that*

$$A = UTU^*.$$

PROOF. Let λ_1 be an eigenvalue of A and $N_1 = N(A - \lambda_1 I)$. Let $\dim N_1 = k_1$ and $x_1, x_2, \cdots, x_{k_1}$ be an orthonormal basis of N_1 and we complete it to a basis of X: $x_1, x_2, \cdots, x_n$. Then we have

$$Ax_j = \lambda_1 x_j, 1 \le j \le k_1,$$
$$Ax_j = \sum_{k=1}^{n} b_{kj} x_k, k_1 + 1 \le j \le n.$$

Hence,

$$A(x_1, x_2, \cdots, x_n) = (x_1, x_2, \cdots, x_n) \begin{pmatrix} \lambda_1 I_{k_1} & B_{12} \\ 0 & B_{22} \end{pmatrix}$$

where

$$\begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix} = \begin{pmatrix} b_{1,k_1+1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{n,k_1+1} & \cdots & b_{n,n} \end{pmatrix}.$$

Let $U_1 = (x_1, x_2, \cdots, x_n)$, we have

$$A = U_1 \begin{pmatrix} \lambda_1 I_{k_1} & B_{12} \\ 0 & B_{22} \end{pmatrix} U_1^*,$$

i.e.,

$$U_1^* A U_1 = \begin{pmatrix} \lambda_1 I_{k_1} & B_{12} \\ 0 & B_{22} \end{pmatrix}$$

Let λ_2 be an eigenvalue of B_{22} and $N_1 = N(B_{22} - \lambda_2 I)$, we can continue this process to obtain U_2 so that

$$B_{22} = U_2 \begin{pmatrix} \lambda_2 I_{k_2} & C_{12} \\ 0 & C_{22} \end{pmatrix} U_2^*$$

Hence

$$\begin{aligned} &\begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2^* \end{pmatrix} U_1^* A U_1 \begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2 \end{pmatrix} \\ &= \begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2^* \end{pmatrix} \begin{pmatrix} \lambda_1 I_{k_1} & B_{12} \\ 0 & B_{22} \end{pmatrix} \begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 I_{k_1} & B_{12} U_2 \\ 0 & U_2^* B_{22} U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 I_{k_1} & \multicolumn{2}{c}{B_{12} U_2} \\ 0 & \lambda_2 I_{k_2} & C_{12} \\ & 0 & C_{22} \end{pmatrix}. \end{aligned}$$

$$A = U_1 \begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2 \end{pmatrix} \begin{pmatrix} \lambda_1 I_{k_1} & \multicolumn{2}{c}{B_{12} U_2} \\ 0 & \lambda_2 I_{k_2} & C_{12} \\ & 0 & C_{22} \end{pmatrix} \begin{pmatrix} I_{k_1} & 0 \\ 0 & U_2^* \end{pmatrix} U_1^*.$$

This process can be continued until we obtain an upper triangular matrix. □

When we have a pair of commuting matrices, we have

THEOREM 89. *Let A, B be a pair of commuting matrices, then A and B can be simultaneously upper triangularized by a unitary matrix.*

PROOF. Let λ_1 be an eigenvector of A and $N = N(A - \lambda I)$. Then N is invariant under B. Let x_1 with $\|x_1\| = 1$ be an eigenvector of $B : N \to N$ with eigenvalue μ_1. Then x_1 is an eigenvector of both A and B. Let $x_1, \cdots, x_n$ be an orthonormal basis and $U_1 = (x_1, \cdots, x_n)$, then we have

$$A = U_1 \begin{pmatrix} \lambda_1 & A_{12} \\ 0 & A_{22} \end{pmatrix} U_1^*,$$
$$B = U_1 \begin{pmatrix} \mu_1 & B_{12} \\ 0 & B_{22} \end{pmatrix} U_1^*.$$

A, B commute implies A_{22} and B_{22} commute, so this process can be continued. □

Theorem 90. *Let A, B be a pair of commuting matrices. We have*

$$r(A+B) \leq r(A) + r(B).$$

Proof. Under an orthonormal basis, A, B are both upper triangular matrices. □

11. Calculus of Vector- and Matrix-Valued Functions

Let $A(t)$ be a matrix-valued function of the real variable t defined in some open interval I. We say that $A(t)$ is continuous at $t_0 \in I$ if

$$\lim_{t \to t_0} \|A(t) - A(t_0)\| = 0.$$

We say that A is differentiable at t_0, with derivative $\dot{A}(t_0) = \frac{d}{dt} A(t)$, if

$$\lim_{h \to 0} \left\| \frac{A(t_0 + h) - A(t_0)}{h} - \dot{A}(t_0) \right\| = 0.$$

Remark 19. *Different norms for finite dimensional spaces are equivalent.*

Remark 20. *$A(t)$ is continuous if and only if all of its components are continuous.*

Remark 21. *$A(t)$ is differentiable if and only if all of its components are differentiable.*

Remark 22. *If $A(t) = (a_{ij}(t))$, then $\dot{A}(t) = (\dot{a}_{ij}(t))$.*

Basic properties:

(1)
$$\frac{d}{dt}(A(t) + B(t)) = \frac{d}{dt} A(t) + \frac{d}{dt} B(t).$$

(2)
$$\frac{d}{dt}(A(t) B(t)) = \left(\frac{d}{dt} A(t)\right) B(t) + A(t) \frac{d}{dt} B(t).$$

(3)
$$\frac{d}{dt}(x(t), y(t)) = \left(\frac{d}{dt} x(t), y(t)\right) + \left(x(t), \frac{d}{dt} y(t)\right).$$

Theorem 91. *Let $A(t)$ be a matrix-valued function, differentiable and invertible. Then $A^{-1}(t)$ also is differentiable, and*

$$\frac{d}{dt} A^{-1}(t) = -A^{-1} \dot{A} A^{-1}.$$

PROOF.

$$A^{-1}(t+h) - A^{-1}(t) = A^{-1}(t+h)\left(A(t) - A(t+h)\right)A^{-1}(t).$$

□

In general, the chain rule fails for matrix valued functions. For example, suppose A is a square matrix,

$$\frac{d}{dt}A^2 = \dot{A}A + A\dot{A} \neq 2A\dot{A}$$

if A and $\dot{A}$ don't commute. We have for any positive integer k

$$\frac{d}{dt}A^k = \sum_{j=1}^{k} A^{j-1}\dot{A}A^{k-j}.$$

THEOREM 92. *Let p be any polynomial, let $A(t)$ be a square matrix-valued function that is differentiable.*
(a) If for a particular value of t the matrices $A(t)$ and $\dot{A}(t)$ commute, then

$$\frac{d}{dt}p(A) = p'(A)\dot{A}.$$

(b) Even if $A(t)$ and $\dot{A}(t)$ do not commute, a trace of the chain rule remains:

$$\frac{d}{dt}\operatorname{tr} p(A) = \operatorname{tr}\left(p'(A)\dot{A}\right).$$

We extend now the product rule to multilinear functions $M(x_1, \cdots, x_k)$. Suppose $x_1, \cdots, x_k$ are differentiable vector functions. Then $M(x_1, \cdots, x_k)$ is differentiable and

$$\frac{d}{dt}M(x_1, \cdots, x_k) = M(\dot{x}_1, \cdots, x_k) + \cdots + M(x_1, \cdots, \dot{x}_k).$$

In particular, for the determinant $D(x_1, \cdots, x_n)$, we have

$$\frac{d}{dt}D(x_1, \cdots, x_n) = D(\dot{x}_1, \cdots, x_n) + \cdots + D(x_1, \cdots, \dot{x}_n).$$

If $X(0) = I$, we have

$$\frac{d}{dt}\left(\det X(t)\right)\Big|_{t=0} = \operatorname{tr}\dot{X}(0).$$

THEOREM 93. *Let $Y(t)$ be a differentiable square matrix-valued function. Then for those values of t for which $Y(t)$ is invertible,*

$$\frac{d}{dt}\ln(\det Y) = \operatorname{tr}\left(Y^{-1}\dot{Y}\right).$$

PROOF. For any fixed t_0, we have

$$Y(t) = Y(t_0)\left(Y^{-1}(t_0)Y(t)\right).$$

□

Next, we consider

$$e^A = \sum_{n=0}^{\infty}\frac{A^n}{n!}$$

which converges for any square matrix A.

Theorem 94. *a) If A and B are commuting square matrices,*

$$e^{A+B} = e^A e^B.$$

(b) If A and B do not commute, then in general

$$e^{A+B} \neq e^A e^B.$$

(c) If $A(t)$ depends differentiably on t, so does $e^{A(t)}$.
(d) If for a particular value of t, $A(t)$ and $\dot{A}(t)$ commute, then

$$\frac{d}{dt} e^{At} = e^A \dot{A}.$$

(e) If A is anti-self-adjoint, $A^* = -A$, then e^A is unitary.

To calculate e^A, we could use Jordan canonical form $A = SJS^{-1}$, then

$$e^A = Se^J S^{-1}.$$

This can be used to show

$$\det e^A = e^{\operatorname{tr} A}.$$

Theorem 95. *The eigenvalues depend continuously on the matrix in the following sense: If*

$$\lim_{j\to\infty} A_j = A,$$

then $\sigma(A_j) \to \sigma(A)$ in the sense that for every $\varepsilon > 0$, there is a k such that all eigenvalues of A_j, $j \geq k$ are contained in discs of radius ε centered at the eigenvalues of A.

Proof. Roots of polynomials depend continuously on the coefficients. □

Theorem 96. *Let $A(t)$ be a differentiable square matrix-valued function of the real variable t. Suppose that $A(0)$ has an eigenvalue λ_0 of multiplicity one. Then for t small enough, $A(t)$ has an eigenvalue $\lambda(t)$ that depends differentiably on t, and $\lambda(0) = \lambda_0$.*

Proof. Let

$$p(\lambda, t) = \det(\lambda I - A(t)).$$

The assumption that λ_0 is a simple root of $p(s, 0)$ implies

$$p(\lambda_0, 0) = 0,$$

$$\frac{\partial}{\partial\lambda} p(\lambda_0, 0) \neq 0.$$

From the implicit function theorem, the equation $p(\lambda, t) = 0$ has a solution $\lambda = \lambda(t)$ in a neighborhood of $t = 0$ that depends differentiably on t. □

Next we show that under the same conditions as in the above theorem, the eigenvector pertaining to the eigenvalue $\lambda(t)$ can be chosen to depend differentiably on t.

Theorem 97. *Let $A(t)$ be a differentiable matrix-valued function of t, $\lambda(t)$ be a continuous function such that $\lambda(t)$ is an eigenvalue of $A(t)$ of multiplicity one. Then we can choose an eigenvector $v(t)$ of $A(t)$ pertaining to the eigenvalue $\lambda(t)$ to depend differentiably on t.*

We need the following lemma:

Lemma 7. *Let A be an $n \times n$ matrix, $p = p_A$ its characteristic polynomial, λ some simple root of p. Then at least one of the $(n-1) \times (n-1)$ principal minors of $A - \lambda I$ has nonzero determinant, where the i-th principal minor is the matrix remaining when the ith row and ith column of λ are removed. Moreover, suppose the i-th principal minor of $A - \lambda I$ has nonzero determinant, then the i-th component of an eigenvector v of A pertaining to the eigenvalue λ is nonzero.*

Proof. WLOG, we assume $\lambda = 0$. Hence,

$$p(0) = 0, \, p'(0) \neq 0.$$

We write

$$A = (c_1, \cdots, c_n)$$

and denote by $e_1, \cdots, e_n$ the standard unit vectors. Then

$$sI - A = (se_1 - c_1, \cdots, se_n - c_n).$$

Hence

$$p'(0) = \sum_{j=1}^{n} (-c_1, \cdots, e_j, \cdots - c_n)$$
$$= (-1)^{n-1} \sum_{j=1}^{n} \det A_j$$

where A_j is j-th principal minor of A. $p'(0) \neq 0$ implies that at least one of the determinants of A_j is nonzero.

Now suppose the i-th principal minors of A has nonzero determinant. Denote by $v^{(i)}$ the vector obtained from the eigenvector v by omitting the i-th component, and by A_i the ith principal minor of A. Then $v^{(i)}$ satisfies

$$A_i v^{(i)} = 0$$

Since $\det A_i \neq 0$, we have $v^{(i)} = 0$ and hence $v = 0$, a contradiction. □

Now we prove Theorem 97:

Proof. Suppose $\lambda(0) = 0$ and $\det A_i(0) \neq 0$. Then for any t small, we have $\det(A_i(t) - \lambda(t) I) \neq 0$ and hence the i-th component of $v(t)$ is not equal to 0. We set it equal to 1 as a way of normalizing $v(t)$. For the remaining components we have now an inhomogeneous system of equations:

$$(A_i(t) - \lambda(t) I) v^{(i)}(t) = -c_i^{(i)}(t)$$

where $c_i^{(i)}(t)$ is the vector obtained from i-th column of $A_i(t) - \lambda(t) I$, c_i, by omitting the i-th component. So we have

$$v^{(i)}(t) = -(A_i(t) - \lambda(t) I)^{-1} c_i^{(i)}(t).$$

Since all terms on the right depend differentiably on t, so does $v^{(i)}(t)$ and $v(t)$. To obtain a differentiable $v(t)$ beyond a small interval, one could normalize the eigenvectors to unit length. □

We show next how to actually calculate the derivative of the eigenvalue $\lambda(t)$ and the eigenvector $v(t)$ of a matrix function $A(t)$ when $\lambda(t)$ is a simple root of the characteristic polynomial of $A(t)$. We start with the eigenvector equation

$$Av = \lambda v$$

which leads to

$$\dot{A}v + A\dot{v} = \lambda' v + \lambda \dot{v}.$$

Let u be an eigenvector of A^T:

$$A^T u = \lambda u.$$

We have

$$\left(u, \dot{A}v\right) + (u, A\dot{v}) = (u, \lambda' v) + (u, \lambda \dot{v}),$$

i.e.,

$$\left(u, \dot{A}v\right) = \lambda' (u, v).$$

We claim that $(u, v) \neq 0$. Suppose on the contrary that $(u, v) = 0$; we claim that then the equation

$$\left(A^T - \lambda I\right) w = u$$

has a solution. Consider $T = A^T - \lambda I$, we have

$$u \in N_{T'}^{\perp} = R_T.$$

Hence,

$$\lambda' = \frac{\left(u, \dot{A}v\right)}{(u, v)}.$$

And now

$$(A - \lambda I)\dot{v} = \left(\lambda' I - \dot{A}\right) v$$

which has a solution since

$$\left(\lambda' I - \dot{A}\right) v \in N_T^{\perp} = R_{T'}$$

The solution is unique if we require $v(t)$ be unit vectors and hence $(\dot{v}, v) = 0$.

When the eigenvalue λ is not simple, the situation is much more complicated.

For example, let

$$A(t) = \begin{pmatrix} 0 & 1 \\ t & 0 \end{pmatrix},$$

we have

$$p(\lambda, t) = \lambda^2 - t,$$

hence we can't write $\lambda(t)$ as an differentiable function of t even though $A(t)$ is smooth.

The next example shows that eigenvectors could be discontinued when $A(t)$ is smooth. let

$$A(t) = \begin{pmatrix} b & c \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & c(t) \\ c(t) & 1 + c(t) k(t) \end{pmatrix}$$

be a real symmetric matrix valued functions where

$$c(t) = \exp\left(-\frac{1}{|t|}\right), k(t) = \sin\frac{1}{t}.$$

Then $A(t)$ is smooth and $A(0) = I$. We have

$$\lambda = \frac{b + d \pm \sqrt{(b-d)^2 + 4c^2}}{2} = \frac{2 + ck \pm c\sqrt{k^2 + 4}}{2}$$

Let $v(t) = \begin{pmatrix} x \\ y \end{pmatrix}$ be an eigenvector corresponding to λ_1, then

$$(b - \lambda_1) x + cy = 0,$$

i.e.,

$$\left(\frac{ck + c\sqrt{k^2+4}}{2}\right) x = cy,$$

i.e.,

$$\frac{k + \sqrt{k^2+4}}{2} x = y.$$

It is easy to see that

$$\lim_{t \to 0} v(t)$$

does't exist.

Nonetheless, one can find sufficient conditions on $A(t)$ so that a differentiable function $v(t)$ of eigenvectors exists. See for example Theorem 12 on page 137 of Lax's book.

There are further results about differentiability of eigenvectors, the existence of higher derivatives, but since these are even more tedious, we shall not pursue them, except for one observation, due to Rellich.

Suppose $A(t)$ is an analytic function of t:

$$A(t) = \sum_{k=0}^{\infty} A_k t^k$$

where each A_k is a self-adjoint matrix. Then the characteristic polynomial of $A(t)$ is analytic in t. The characteristic equation

$$p(s,t) = 0$$

defines s as a function of t. Near a value of t where the roots of p are simple, the roots $\lambda(t)$ are regular analytic functions of t; near a multiple root the roots have an algebraic singularity and can be expressed as power series in a fractional power of t:

$$\lambda(t) = \sum_{k=0}^{\infty} r_j t^{\frac{k}{j}}.$$

On the other hand, we know that for real t, the matrix $A(t)$ is self-adjoint and therefore all its eigenvalues are real. Since fractional powers of t have complex values for real t, we can deduce that the eigenvalues $\lambda(t)$ are regular analytic functions of t.

12. Matrix Inequalities

We recall the definition of a positive mapping:

Definition 50. *A self-adjoint linear mapping H from a real or complex Euclidean space into itself is called positive if*

$$(x, Hx) > 0 \text{ for any } x \neq 0.$$

And we write $H > O$.

Similarly, we can define nonnegativity, negativity and nonpositivity of a self-adjoint map.

Basic properties:

(1) The identity map

$$I > 0.$$

(2) If $A, B > 0$ then so is $A + B$ and kA if $k > 0$.
(3) If H is positive and Q is invertible, then

$$Q^* HQ > 0.$$

(4) H is positive iff all its eigenvalues are positive.
(5) Every positive mapping is invertible.
(6) Every positive mapping has a positive square root, uniquely determined.
Proof: To see uniqueness, suppose $S^2 = T^2$, where $S > 0$, $T > 0$. We consider, for any x,

$$\begin{aligned}\operatorname{Re}\{(x, (S+T)(S-T)x)\} &= \operatorname{Re}\{(x, TSx) - (x, STx)\} \\ &= \operatorname{Re}\{(Tx, Sx) - (Sx, Tx)\} = 0.\end{aligned}$$

Now for any eigenvalue μ of $S - T$ with eigenvector x, then μ is real, and

$$(x, (S+T)(S-T)x) = \mu(x, (S+T)x)$$

must be real too, hence, $\mu = 0$ since $(x, (S+T)x) > 0$. So we have $S - T = 0$ and $S = T$.
(7) The set of all positive maps is an open subset of the space of all self-adjoint maps.
(8) The boundary points of the set of all positive maps are nonnegative maps that are not positive.

Remark 23. *In general, there are many solutions to $S^2 = H$ for given $H > 0$.*

Definition 51. *Let M and N be two self-adjoint mappings of a Euclidean space into itself. We say that M is less than N, denoted as $M < N$ if $N - M$ is positive.*

The relation ">" defines a partial order of self-adjoint mappings.
Similarly, we can define the relation $M \leq N$.

Proposition 8. *(1). Additive Property: If $M_1 < N_1$, and $M_2 < N_2$, then*

$$M_1 + M_2 < N_1 + N_2;$$

(2). Transitive Property: If $L < M$ and $M < N$, then $L < N$;
(3). Multiplicative Property: If $M < N$ and Q is invertible, then

$$Q^* MQ < Q^* NQ.$$

Theorem 98. *$M > N > 0$ implies*

$$0 < M^{-1} < N^{-1}$$

Proof. Method I: If $N = I$, trivial. In general, write $N = R^2$ where $R > 0$.

$$M > R^2$$

iff

$$R^{-1} M R^{-1} > I.$$

Hence,

$$RM^{-1}R < I$$

which is equivalent to

$$M^{-1} < R^{-2} = N^{-1}.$$

Method II: Define

$$A(t) = tM + (1-t)N.$$

For $t \in [0,1]$, we have $A(t) > 0$. Since

$$\frac{d}{dt} A^{-1}(t) = -A^{-1}\dot{A}A^{-1} < 0$$

we have

$$N^{-1} = A^{-1}(0) > A^{-1}(1) = M^{-1}.$$

□

Lemma 8. *Let $A(t)$ be a differentiable function of the real variable whose values are self-adjoint mappings; the derivative $\frac{d}{dt}A$ is then also self-adjoint. Suppose that $\frac{d}{dt}A$ is positive; then $A(t)$ is an increasing function, that is,*

$$A(s) < A(t) \text{ whenever } s < t.$$

Proof. For any $x \neq 0$, we have

$$\frac{d}{dt}(x, Ax) = \left(x, \dot{A}x\right) > 0.$$

□

The product of two self-adjoint mappings is not, in general, self-adjoint. We introduce the symmetrized product

$$S = AB + BA$$

of two self-adjoint mappings A and B. In general, $A, B > 0$ would not imply $S > 0$. For any $x \neq 0$, we have

$$(x, Sx) = 2\,\mathrm{Re}\,(Ax, Bx).$$

Take $x = e_1$, we have

$$\mathrm{Re}\,(Ax, Bx) = (a_1, b_1)$$

where a_1, b_1 are the first column of A, B respectively. Choose for example

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

we would have

$$(e_1, Se_1) = 0.$$

However, we have

Theorem 99. *Let A and B be self-adjoint such that $A > O$ and the symmetrized product*

$$S = AB + BA > O.$$

Then $B > O$.

Proof. Define

$$B(t) = B + tA.$$

Then for any $t > 0$,

$$\begin{aligned} S(t) &= AB(t) + BA(t) \\ &= AB + BA + 2tA^2 > 0. \end{aligned}$$

Since $B(t) > 0$ for t sufficiently large and $B(t)$ depends continuously on t, if $B = B(0)$ were not positive, there would be some nonnegative value t_0 such that

$B(t_0)$ lies on the boundary of the set of positive mappings. Hence, $0 \in \sigma(B(t_0))$, let $x \neq 0$ be a corresponding eigenvector, we have

$$(x, S(t_0)x) = 2\,\mathrm{Re}\,(Ax, B(t_0)x) = 0$$

which contradicts to $S(t_0) > 0$. □

THEOREM 100. *Suppose*

$$O < M < N,$$

then

$$O < \sqrt{M} < \sqrt{N}.$$

PROOF. Let

$$A(t) = tN + (1-t)M > 0$$

and

$$R(t) = \sqrt{A(t)} > 0.$$

Then we have $R^2 = A$ and

$$\dot{R}R + R\dot{R} = \dot{A} = N - M > 0.$$

Hence

$$\dot{R} > 0$$

which implies $R(0) < R(1)$, i.e., $\sqrt{M} < \sqrt{N}$. □

On the other hand, we have

EXAMPLE 12. *Let $M = A + B$ and $N = A - B$, we would have*

$$M - N = 2A,$$
$$M^2 - N^2 = 2(AB + BA).$$

Let

$$A = t\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

when t large, we have $M > N > 0$, but $M^2 > N^2$ doesn't hold.

EXAMPLE 13. *Let*

$$A = \begin{pmatrix} 2 & 2 \\ 2 & 9 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix},$$

we have $A > B > 0$, but $A^2 > B^2$ doesn't hold.

COROLLARY 12. *Suppose*

$$O < M < N,$$

then for any positive integer k,

$$O < M^{\frac{1}{2^k}} < N^{\frac{1}{2^k}}.$$

For any $A > 0$, write $A = U\Lambda U^*$, we can define

$$\log A = U(\log\Lambda)U^*.$$

We can verify that $A = e^{\log A}$.

LEMMA 9. *For any $A > 0$,*

$$\log A = \lim_{m\to\infty} m\left(A^{\frac{1}{m}} - 1\right).$$

THEOREM 101. *Suppose*

$$O < M < N,$$

then

$$\log M \leq \log N.$$

DEFINITION 52. *A real-valued function $f(s)$ defined for $s > 0$ is called a monotone matrix function if all pairs of self-adjoint mappings M, N satisfying $O < M < N$ also satisfy $f(M) < f(N)$ where $f(M)$, $f(N)$ are defined by the spectral resolution.*

Hence, the functions

$$f(s) = -\frac{1}{s}, \log s, s^{\frac{1}{2^k}}$$

are monotone matrix functions while $f(s) = s^2$ is not.

REMARK 24. *Let P_n be the class of monotone matrix function of order n. Our definition of monotone matrix function is $\cap_{n\geq 1} P_n$.*

Carl Loewner has proved the following beautiful theorem.

THEOREM 102. *Every monotone matrix function can be written in the form of*

$$f(s) = as + b - \int_0^\infty \frac{(st-1)\, dm(t)}{s+t}$$

where a is nonnegative, b is real, and $m(t)$ is a nonnegative measure on $[0, \infty)$ for which the integral converges. Moreover, a function f can be written in this form if and only if it can be extended as an analytic function in the upper half-plane, and has a positive imaginary part there.

Next, we describe a method for constructing positive matrices.

DEFINITION 53. *Let $f_1, \cdots, f_n$ be an ordered set of vectors in a Euclidean space. The matrix G with entries*

$$G_{ij} = (f_i, f_j)$$

is called the Gram matrix of the set of vectors.

THEOREM 103. *(i) Every Gram matrix is nonnegative.*
(ii) The Gram matrix of a set of linearly independent vectors is positive.
(iii) Every positive matrix can be represented as a Gram matrix.

PROOF. (i), (ii) follows from

$$(x, Gx) = \sum x_i \overline{G_{ij}} \overline{x_j} = \sum x_i (f_j, f_i) \bar{x}_j$$
$$= \sum (\bar{x}_j f_j, \bar{x}_i f_i) = \left\| \sum \bar{x}_i f_i \right\|^2.$$

(iii) Let $H = (h_{ij})$ be positive. Let $f_j = \sqrt{H} e_j$, we have

$$(f_i, f_j) = \left(\sqrt{H} e_i, \sqrt{H} e_j\right) = h_{ij}.$$

Alternatively, we could also define a scalar product $(,)_H$ by

$$(x, y)_H = (x, Hy).$$

Then

$$h_{ij} = (e_i, e_j)_H.$$

□

Example 14. *Take the Euclidean space to consist of real-valued functions on the interval* $[0,1]$*, with the scalar product*

$$(f,g)=\int_0^1 f(x)\,g(x)\,dx.$$

Choose $f_i(x)=t^{i-1}$*,* $i=1,\cdots,n$*, the associated Gram matrix is* $G=(G_{ij})_{n\times n}$ *with*

$$G_{ij}=\frac{1}{i+j-1}.$$

Definition 54. *The Hadamard product of two* $m\times n$ *matrices* A,B *is*

$$A\circ B=(A_{ij}B_{ij})_{m\times n}.$$

Theorem 104 (Schur). *Let* $A=(A_{ij})$ *and* $B=(B_{ij})$ *be* $n\times n$ *positive matrices. Then* $M=A*B$ *where*

$$M_{ij}=A_{ij}B_{ij}$$

is also a positive matrix.

Proof. Suppose

$$A_{ij}=(f_i,f_j)=\sum_k f_i^k\overline{f_j^k},$$
$$B_{ij}=(g_i,g_j)=\sum_l g_i^l\overline{g_j^l}$$

where $\{f_i\}$ $\{g_j\}$ are two sets of linearly independent vectors.

Let $h_i=\left(f_i^k g_i^l\right)_{n\times n}\in\mathbb{R}^{n^2}$, then

$$M_{ij}=\sum_{k,l} f_i^k\overline{f_j^k}g_i^l\overline{g_j^l}=(h_i,h_j).$$

We claim h_i, $1\le i\le n$ are linearly independent in $\mathbb{R}^{n^2}$. Suppose

$$0=\sum c_ih_i=\sum_i c_i\left(f_i^k g_i^l\right)_{n\times n},$$

we conclude for any l,

$$\sum_i\left(c_ig_i^l\right)f_i=0,$$

hence $c_ig_i^l=0$ for each i,l and hence $c_ig_i=0$ for each i. So we conclude $c_i=0$. □

The determinant of every positive matrix is positive. Moreover, determinant is a log concave function on positive definite matrices.

Theorem 105. *Let* A *and* B *be self-adjoint, positive* $n\times n$ *matrices. Then for any* $t\in[0,1]$*,*

$$\det(tA+(1-t)B)\ge(\det A)^t(\det B)^{1-t}.$$

Proof. For any invertible and differentiable matrix function $Y(t)$, we have

$$\begin{aligned}&\frac{d}{dt}\log(\det Y(t))\\&=\operatorname{tr}Y^{-1}Y(t).\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{d^2}{dt^2}\log\left(\det Y\left(t\right)\right)\\
&=\operatorname{tr}\frac{d}{dt}Y^{-1}\dot{Y}\left(t\right)\\
&=\operatorname{tr}\left(-Y^{-1}\dot{Y}Y^{-1}\dot{Y}+Y^{-1}\ddot{Y}\left(t\right)\right).
\end{aligned}$$

Now let $Y\left(t\right)=tA+\left(1-t\right)B$, we have

$$\frac{d^2}{dt^2}\log\left(\det Y\left(t\right)\right)=-\operatorname{tr}\left(Y^{-1}\dot{Y}Y^{-1}\dot{Y}\right)=-\operatorname{tr}\left(Y^{-1}\dot{Y}\right)^2\leq 0$$

since the eigenvalues of the product of a positive matrix and a self-adjoint matrix are real.

Second proof: Define $C=B^{-1}A$, it suffices to show

$$\det\left(tC+\left(1-t\right)I\right)\geq\left(\det C\right)^t.$$

Since C is the product of two positive matrices, it has positive eigenvalues λ_j. Now

$$\det\left(tC+\left(1-t\right)I\right)=\prod_j\left(t\lambda_j+\left(1-t\right)\right)\geq\prod_j\lambda_j^t=\left(\det C\right)^t$$

follows from the fact that

$$t\lambda+\left(1-t\right)\geq\lambda^t$$

holds for any $\lambda>0$ and $0\leq t\leq 1$ since λ^t is a convex function in t. □

Next we give a useful estimate for the determinant of a positive matrix.

THEOREM 106. *Let $H>0$. Then*

$$\det H\leq\prod H_{jj}.$$

PROOF. Let $D=\left[\sqrt{H_{11}},\cdots,\sqrt{H_{nn}}\right]>0$ and

$$B=D^{-1}HD^{-1}>0.$$

Then $B_{ii}=1$ and

$$\frac{\det H}{\prod H_{jj}}=\det B.$$

It suffices to show $\det B\leq 1$ which follows from the arithmetic geometric mean inequality

$$\det B=\prod\lambda_j\leq\left(\frac{\sum\lambda_j}{n}\right)^n=\left(\frac{\operatorname{tr}B}{n}\right)^n=1$$

where λ_j, $1\leq j\leq n$ are eigenvalues of B. □

THEOREM 107. *Let T be any $n\times n$ whose columns are $c_1,\cdots,c_n$. Then*

$$\left|\det T\right|\leq\prod\left\|c_j\right\|.$$

PROOF. We consider $H=T^*T$. We have

$$\left|\det T\right|^2=\det H\leq\prod H_{jj}=\prod\left\|c_j\right\|^2.$$

□

In the real case, the above theorem has an obvious geometrical meaning: among all parallelepipeds with given side lengths $\|c_j\|$, the one with the largest volume is rectangular.

THEOREM 108. *Let H be an $n \times n$ real, symmetric, positive matrix. Then*

$$\frac{\pi^{\frac{n}{2}}}{\sqrt{\det H}} = \int_{\mathbb{R}^n} e^{-(x,Hx)} dx.$$

PROOF. Let $y = \sqrt{H}x$, then

$$\int_{\mathbb{R}^n} e^{-\|y\|^2} \frac{1}{\det \sqrt{H}} dy = \frac{\pi^{\frac{n}{2}}}{\sqrt{\det H}}.$$

□

Now we present a number of interesting and useful results on eigenvalues. Let A be a self-adjoint map of a Euclidean space U into itself. We denote by $p_+(A)$ the number of positive eigenvalues of A, and denote by $p_-(A)$ the number of its negative eigenvalues. Then $p_+(A) =$ maximum dimension of subspace S of U such that (u, Au) is positive on S. $p_-(A) =$ maximum dimension of subspace S of U such (Au, u) is negative on S.

THEOREM 109. *Let A be a self-adjoint map of a Euclidean space U into itself, and let V be a subspace of U such that*

$$\dim V = \dim U - 1.$$

Denote by P orthogonal projection onto V. Then PAP is a self-adjoint map of U into U that maps V into V; we denote by B the restriction of PAP to V. Then we have:
(a)

$$p_+(A) - 1 \le p_+(B) \le p_+(A)$$

and

$$p_-(A) - 1 \le p_-(B) \le p_-(A).$$

(b) The eigenvalues of B separate the eigenvalues of A.

PROOF. (a) Let T be a subspace of V of dimension $p_+(B)$ on which B is positive. Then for any $v \in T$, $v \neq 0$,

$$(v, Av) = (Pv, APv) = (v, PAPv) = (v, Bv) > 0.$$

Hence, A is positive on T and

$$p_+(A) \ge \dim T = p_+(B).$$

On the other hand, let S be a subspace of U of dimension $p_+(A)$ on which A is positive. Let $T = S \cap V$, then

$$\dim T \ge \dim S - 1.$$

For any $v \in T$, $v \neq 0$,

$$(v, Bv) = (v, PAPv) = (v, Av) > 0.$$

Hence B is positive on T, and

$$p_+(B) \ge \dim T \ge \dim S - 1 = p_+(A) - 1.$$

(b) We have for any c,

$$\begin{aligned} & \left|\{\lambda \in \sigma(A) : \lambda > c\}\right| \\ & \geq \left|\{\lambda \in \sigma(B) : \lambda > c\}\right| \\ & \geq \left|\{\lambda \in \sigma(A) : \lambda > c\}\right| - 1. \end{aligned}$$

□

COROLLARY 13. *Let A be an $n \times n$ self-adjoint matrix. For any $1 \leq i \leq n$, the eigenvalues of the i-th principal minor of A separate the eigenvalues of A.*

PROOF. Let V be the subspace with i-th component 0. □

THEOREM 110. *Let M and N denote self-adjoint $n \times n$ matrices satisfying $M < N$. Denote the eigenvalues of M by $\lambda_1 \leq \cdots \leq \lambda_n$, and those of N by $\mu_1 \leq \cdots \leq \mu_n$. Then for $1 \leq j \leq n$,*

$$\lambda_j < \mu_j.$$

PROOF.

$$\begin{aligned} \lambda_j &= \min_{\dim S = j} \max_{x \in S, x \neq 0} \frac{(x, Mx)}{(x, x)}, \\ \mu_j &= \min_{\dim S = j} \max_{x \in S, x \neq 0} \frac{(x, Nx)}{(x, x)}. \end{aligned}$$

Denote by T the subspace of dimension j such that

$$\mu_j = \max_{x \in T, x \neq 0} \frac{(x, Nx)}{(x, x)}.$$

Since $M < N$,

$$\begin{aligned} \mu_j &= \max_{x \in T, x \neq 0} \frac{(x, Nx)}{(x, x)} \\ &> \max_{x \in T, x \neq 0} \frac{(x, Mx)}{(x, x)} \geq \lambda_j. \end{aligned}$$

□

REMARK 25. *$M \leq N$ implies $\lambda_j \leq \mu_j$.*

THEOREM 111. *Let M and N be two self-adjoint $n \times n$ matrices. Denote the eigenvalues of M by $\lambda_1 \leq \cdots \leq \lambda_n$, and those of N by $\mu_1 \leq \cdots \leq \mu_n$. Then for $1 \leq j \leq n$,*

$$\left|\lambda_j - \mu_j\right| \leq \|M - N\|.$$

PROOF. Let $d = \|M - N\|$. We have

$$N - dI \leq M \leq N + dI.$$

□

Wielandt and Hoffman have proved the following interesting result.

THEOREM 112. *Let M, N be self-adjoint $n \times n$ matrices. Denote the eigenvalues of M by $\lambda_1 \leq \cdots \leq \lambda_n$, and those of N by $\mu_1 \leq \cdots \leq \mu_n$. Then*

$$\sum \left|\lambda_j - \mu_j\right|^2 \leq \|M - N\|_2^2$$

where $\|\cdot\|_2$ is the Hilbert-Schmidt norm.

Proof.

$$\|M-N\|_2^2 = \operatorname{tr}(M-N)^2 = \operatorname{tr} M^2 - 2\operatorname{tr} MN + \operatorname{tr} N^2$$
$$= \sum \lambda_j^2 - 2\operatorname{tr} MN + \sum \mu_j^2.$$

So it suffices to show

$$\operatorname{tr} MN \le \sum_{j=1}^n \lambda_j \mu_j.$$

We could assume that $M = [\lambda_1, \cdots, \lambda_n]$, then

$$\operatorname{tr} MN = \sum_{j=1}^n \lambda_j N_{jj}.$$

Let Γ be the collection of all self-adjoint $n \times n$ matrices with eigenvalues $\mu_1 \le \cdots \le \mu_n$. Then Γ is compact so there exists $N_m \in \Gamma$ such that

$$\operatorname{tr} MN_m = \max_{N \in \Gamma} \operatorname{tr} MN.$$

We claim that $MN_m = N_m M$. Let $A = N_m M - MN_m$. Then A is anti-self-adjoint and hence e^{At} is unitary so we have

$$N(t) = e^{At} N_m e^{-At} \in \Gamma.$$

Now

$$\frac{d}{dt} \operatorname{tr} MN(t)$$
$$= \operatorname{tr}\left(MAe^{At}N_m e^{-At} - Me^{At}N_m Ae^{-At}\right)$$
$$= \operatorname{tr} Me^{At}(AN_m - N_m A)e^{-At}.$$

Hence

$$0 = \left.\frac{d}{dt} \operatorname{tr} MN(t)\right|_{t=0}$$
$$= \operatorname{tr} M(AN_m - N_m A)$$
$$= \operatorname{tr} A^2$$

and we conclude $A = 0$.

Now M and N_m can be simultaneously unitary diagonalized. Hence

$$\operatorname{tr} MN \le \operatorname{tr} MN_m = \sum_{j=1}^n \lambda_j \tilde{\mu}_j \le \sum_{j=1}^n \lambda_j \mu_j$$

where $\tilde{\mu}_j$ is a rearrangement of μ_j. □

Theorem 113. *Denote by $\lambda_{\min}(H)$ the smallest eigenvalue of a self-adjoint mapping H in a Euclidean space. Then $\lambda_{\min}(H)$ is a concave function of H, that is, for $0 < t < 1$,*

$$\lambda_{\min}(tL + (1-t)M) \ge t\lambda_{\min}(L) + (1-t)\lambda_{\min}(M).$$

Similarly, $\lambda_{\max}(H)$ is a convex function of H.

PROOF. Let $x \neq 0$ be a unit vector such that

$$\begin{aligned}
&\lambda_{\min}(tL + (1-t)M) \\
&= (x, (tL + (1-t)M)x) \\
&= t(x, Lx) + (1-t)(x, Mx) \\
&\geq t\lambda_{\min}(L) + (1-t)\lambda_{\min}(M).
\end{aligned}$$

□

Let Z be linear mapping of a complex Euclidean space into itself.

$$H = \frac{Z + Z^*}{2}$$

is called the self-adjoint part of Z, and

$$A = \frac{Z - Z^*}{2}$$

is called the anti-self-adjoint part of Z. And we have

$$Z = H + A.$$

THEOREM 114. *Suppose*

$$H = \frac{Z + Z^*}{2} > 0.$$

Then the eigenvalues of Z have positive real part.

PROOF. For any vector h,

$$\operatorname{Re}(h, Zh) = \frac{1}{2}[(h, Zh) + (Zh, h)] = (h, Hh) \geq 0.$$

Now let $\lambda \in \sigma(Z)$ and h be a unit length eigenvector, then

$$\operatorname{Re}\lambda = \operatorname{Re}(h, Zh) \geq 0.$$

□

The decomposition of an arbitrary Z as a sum of its self-adjoint and anti-self-adjoint parts is analogous to writing a complex number as the sum of its real and imaginary parts, and the norm is analogous to the absolute value.

Let a denote any complex number with positive real part; then

$$w = \frac{1 - az}{1 + \bar{a}z}$$

maps the right half-plane $\operatorname{Re} z > 0$ onto the unit disc $|w| < 1$. Analogously, we claim the following:

THEOREM 115. *Let a be a complex number with $\operatorname{Re} a > 0$. Let Z be a mapping whose self-adjoint part $Z + Z^*$ is positive. Then*

$$W = (I - aZ)(I + \bar{a}Z)^{-1}$$

is a mapping of norm less than 1. Conversely, $\|W\| < 1$ implies that $Z + Z^* > 0$.

PROOF. $Z + Z^* > 0$ implies $\operatorname{Re}\lambda > 0$ for any $\lambda \in \sigma(Z)$. Hence $I + \bar{a}Z$ has nonzero eigenvalues and is invertible. For any vector $x \neq 0$, we write

$$x = (I + \bar{a}Z)\, y,$$

hence

$$(I - aZ)\, y = Wx.$$

Now

$$\|Wx\|^2 < \|x\|^2$$

iff

$$\|(I - aZ)\, y\|^2 < \|(I + \bar{a}Z)\, y\|^2$$

iff

$$-(y, aZy) - (aZy, y) < (y, \bar{a}Zy) + (\bar{a}Zy, y)$$

iff

$$(y, Zy) + (Zy, y) > 0$$

iff

$$(y, (Z + Z^*)\, y) > 0.$$

□

13. Polar decomposition and singular value decomposition

Complex number z can be written in the form of $z = re^{i\theta}$ where $r \geq 0$ and $\left|e^{i\theta}\right| = 1$. Mappings of Euclidean spaces have similar decompositions.

THEOREM 116 (Polar Decomposition). *Let A be a linear mapping of a complex Euclidean space into itself. Then A can be factored as*

$$A = RU$$

where R is a nonnegative self-adjoint mapping, and U is unitary. When A is invertible, R is positive.

PROOF. We first assume A is invertible. Let

$$R = \sqrt{AA^*} > 0$$

and

$$U = R^{-1}A.$$

Then U is unitary since

$$U^*U = A^*R^{-1}R^{-1}A = A^*\,(AA^*)^{-1}\,A = I.$$

When A is not invertible, we can still define

$$R = \sqrt{AA^*} \geq 0.$$

For any x, we have

$$\|Rx\|^2 = (Rx, Rx) = (A^*x, A^*x) = \|A^*x\|^2\,.$$

For any $u = Rx \in R_R$, the range of R, we can define $Vu = A^*x$. Then V is an isometry on R_R; therefore it can be extended to the whole space as a unitary mapping. Since

$$A^* = VR$$

we have

$$A = RV^* = RU$$

where $U = V^*$ is unitary. □

REMARK 26. *A can also be factored as*

$$A = UR$$

where R is a nonnegative self-adjoint mapping, and U is unitary.

REMARK 27. *When A is real, we can choose R to be real symmetric and $U = Q$ to be orthogonal. Polar decomposition implies that A is a composition of a map that stretches the space along a set of orthogonal bases, represented by R, and a rotation represented by Q.*

REMARK 28. *Every nonnegative mapping has a nonnegative square root, uniquely determined. To see uniqueness, suppose $S^2 = T^2$, where $S \geq 0$, $T \geq 0$. We consider, for any x,*

$$\begin{aligned} \operatorname{Re}\{(x, (S+T)(S-T)x)\} &= \operatorname{Re}\{(x, TSx) - (x, STx)\} \\ &= \operatorname{Re}\{(Tx, Sx) - (Sx, Tx)\} = 0. \end{aligned}$$

Now suppose $\mu \neq 0$ is an eigenvalue of $S - T$ with eigenvector x, then μ is real, and

$$(x, (S+T)(S-T)x) = \mu (x, (S+T)x)$$

must be real too, hence, $(x, (S+T)x) = 0$ which implies $Sx = Tx = 0$. So we have $(S-T)x = 0$ and $\mu = 0$, a contradiction. Hence any eigenvalue of $S - T$ must be zero which implies $S - T = 0$.

According to the spectral representation theorem, the self-adjoint map R can be expressed as

$$R = WDW^*,$$

where D is diagonal and W is unitary. Hence

$$A = WDW^*U = WDV$$

where $V = W^*U$ is unitary. Such a decomposition is called the singular value decomposition of the mapping A. The diagonal entries of D are called the singular values of A; they are the nonnegative square roots of the eigenvalues of AA^*.

Singular value decomposition can be generalized to any linear map A from one Euclidean space into another Euclidean space.

THEOREM 117 (Singular Value Decomposition). *Let A be an $m \times n$ matrix. There exist an $m \times m$ unitary matrix U and an $n \times n$ unitary matrix V such that*

$$A = U\Sigma V^*$$

where Σ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal.

PROOF. WLOG, we assume $m > n$.
Since

$$A^*A \geq 0.$$

There exists an $n \times n$ unitary matrix V such that

$$A^*A = VDV^*$$

where D is a diagonal matrix with nonnegative diagonal entries. Now for any $x \in \mathbb{C}^n$,

$$\begin{aligned}\|AVx\|^2 &= (AVx, AVx) = (Vx, A^*AVx) \\ &= (Vx, VDV^*Vx) = (Vx, VDx) \\ &= (x, Dx) = \left\|D^{\frac{1}{2}}x\right\|^2.\end{aligned}$$

Hence, there exists an isometry U from the range of $D^{\frac{1}{2}}$ into $\mathbb{C}^m$ which maps $D^{\frac{1}{2}}x$ to AVx. Hence, there exists a unitary map U from $\mathbb{C}^m$ to $\mathbb{C}^m$ such that

$$U\begin{pmatrix} D^{\frac{1}{2}} \\ 0 \end{pmatrix} = AV,$$

i.e.,

$$A = U\begin{pmatrix} D^{\frac{1}{2}} \\ 0 \end{pmatrix}V^* = U\Sigma V^*.$$

□

Given a singular value decomposition of A:

$$A = U\Sigma V^*.$$

We have

$$\begin{aligned}AA^* &= U\Sigma\Sigma^*U^*, \\ A^*A &= V\Sigma^*\Sigma V^*.\end{aligned}$$

Hence the columns of V are eigenvectors of A^*A which are called right-singular vectors and the columns of U are eigenvectors of AA^* which are called left-singular vectors. The non-zero elements of Σ are the square roots of the non-zero eigenvalues of A^*A or AA^* which are called non-zero singular values of A.

Remark 29. *When A is real, we could choose U, V to be real.*

Remark 30. *Suppose $m > n$. To construct the singular value decomposition, we first calculate the eigenvalues and eigenvectors of AA^* to obtain positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ of A and the unitary map V. Writing $U = (u_1, \cdots, u_m)$, we have*

$$u_k = \frac{1}{\sigma_k}Av_k,\ 1 \leq k \leq r,$$

and the unitary map U is obtained by completing u_k, $1 \leq k \leq r$ into an orthonormal basis.

Remark 31. *We could also write singular value decomposition in the form of*

$$A = \sum_{k=1}^{r} \sigma_k u_k v_k^*,$$

a sum of rank one matrices.

Similar to eigenvalues, singular values have a variational structure. Let A be a real $m \times n$ matrix. We define for $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ the bilinear form

$$\sigma(u, v) = (u, Av).$$

Since unit balls are compact, σ achieves its maximum for $\|u\| = \|v\| = 1$.

THEOREM 118. *Let*

$$\lambda_1 = \max_{\|u\|=\|v\|=1} \sigma(u,v) = \sigma(u_1, v_1)$$

where $\|u_1\| = \|v_1\| = 1$*. Then* λ_1 *is the largest singular value and* u_1*,* v_1 *are left and right-singular vectors w.r.t.* λ*. Moreover,*

$$\begin{aligned} Av_1 &= \lambda_1 u_1, \\ A^* u_1 &= \lambda_1 v_1. \end{aligned}$$

PROOF. λ_1 is the largest singular value follows from the observation

$$\lambda_1 = \max_{\|u\|=\|v\|=1} \sigma(u,v) = \max_{\|u\|=\|v\|=1} (u, \Sigma v).$$

For any vectors g, h, we have

$$f(s,t) = \frac{\sigma(u_1 + sg, v_1 + th)}{\|u_1 + sg\| \, \|v_1 + th\|} = \frac{(u_1 + sg, Av_1 + tAh)}{\|u_1 + sg\| \, \|v_1 + th\|}$$

achieves its maximum at $(0,0)$. Hence

$$\begin{aligned} f_s(s,t)|_{(s,t)=(0,0)} &= (g, Av_1) + \lambda_1 \left(\frac{d}{ds} \frac{1}{\|u_1 + sg\|} \bigg|_{s=0} \right) \\ &= (g, Av_1) - \lambda_1 (g, u_1) \end{aligned}$$

since

$$\frac{d}{ds} \|u_1 + sg\|^2 \bigg|_{s=0} = 2(u_1, g).$$

Similarly,

$$\begin{aligned} f_t(s,t)|_{(s,t)=(0,0)} &= (u_1, Ah) + \lambda_1 \left(\frac{d}{ds} \frac{1}{\|v_1 + th\|} \bigg|_{s=0} \right) \\ &= (u_1, Ah) - \lambda_1 (h, v_1). \end{aligned}$$

Hence

$$\begin{aligned} Av_1 &= \lambda_1 u_1, \\ A^* u_1 &= \lambda_1 v_1. \end{aligned}$$

It then follows

$$\begin{aligned} A^* A v_1 &= \lambda_1^2 v_1, \\ A A^* u_1 &= \lambda_1^2 u_1. \end{aligned}$$

□

14. More Properties of Positive Definite Matrices

Let $A > 0$ be an $n \times n$ matrix. For $1 \le k \le n$, $k \times k$ submatrix of A obtained by deleting any $n-k$ columns and the same $n-k$ rows from A is called a k-th-order principal submatrix of A. The determinant of a $k \times k$ principal submatrix is called a k-th order principal minor of A.

THEOREM 119. *Any principal submatrix of a positive matrix* A *is positive.*

The leading principal minor is the determinant of the leading principal submatrix obtained by deleting the last $n-k$ rows and the last $n-k$ columns.

Theorem 120 (Sylvester's criterion). *A is positive definite iff all of its n leading principal minors are strictly positive.*

Proof. Using 13. □

Now we want to generalize Theorem 106 to block matrices.

Theorem 121. *Suppose*

$$A = \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} > 0$$

where B and D are square matrices of dimension k and l. Then $B > 0$ and $D > 0$.

$$\det A \leq \det B \det D.$$

We need the following lemma

Lemma 10. *Suppose $A \geq 0$ and $B \geq 0$. We have*

$$\det(A + B) \geq \det A + \det B.$$

Proof. First we show that for any $B \geq 0$,

$$\det(I + B) \geq 1 + \det B.$$

Let $\lambda_k \geq 0$ be eigenvalues of B. We have

$$\det(I + B) = \prod_{k=1}^{n} (1 + \lambda_k) \geq 1 + \prod_{k=1}^{n} \lambda_k = 1 + \det B.$$

Next, we assume that $A > 0$ and $B \geq 0$. We have

$$\begin{aligned} \det(A + B) &= \det A \det\left(I + A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) \\ &\geq \det A \left(1 + \det\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)\right) \\ &= \det A + \det B. \end{aligned}$$

Finally, for $A \geq 0$ and $B \geq 0$, we take for every $\varepsilon > 0$, $A_\varepsilon = A + \varepsilon I > 0$. Then

$$\det(A_\varepsilon + B) \geq \det A_\varepsilon + \det B.$$

Letting $\varepsilon \to 0$, we have

$$\det(A + B) \geq \det A + \det B.$$

□

Now we are ready to prove Theorem 121:

Proof of Theorem 121. Let

$$Q = \begin{pmatrix} I_k & -B^{-1}C \\ 0 & I_l \end{pmatrix},$$

we have

$$\begin{aligned} Q^* A Q &= \begin{pmatrix} I_k & 0 \\ -C^* B^{-1} & I_l \end{pmatrix} \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} \begin{pmatrix} I_k & -B^{-1}C \\ 0 & I_l \end{pmatrix} \\ &= \begin{pmatrix} B & C \\ 0 & D - C^* B^{-1} C \end{pmatrix} \begin{pmatrix} I_k & -B^{-1}C \\ 0 & I_l \end{pmatrix} \\ &= \begin{pmatrix} B & 0 \\ 0 & D - C^* B^{-1} C \end{pmatrix}. \end{aligned}$$

Hence

$$\det A = \det (Q^* A Q) = \det B \det \left(D - C^* B^{-1} C\right).$$

Now since $C^* B^{-1} C \geq 0$ and $D - C^* B^{-1} C > 0$, we have

$$\begin{aligned} \det D &= \det \left(\left(D - C^* B^{-1} C\right) + C^* B^{-1} C\right) \\ &\geq \det \left(D - C^* B^{-1} C\right) + \det \left(C^* B^{-1} C\right) \\ &\geq \det \left(D - C^* B^{-1} C\right). \end{aligned}$$

Hence, we have

$$\det A = \det B \det \left(D - C^* B^{-1} C\right) \leq \det B \det D.$$

□

More generally, suppose the block matrix

$$A = (A_{ij})_{l \times l} > 0$$

where A_{kk}, $1 \leq k \leq l$ are square matrices. Then

$$\det A \leq \prod_{k=1}^{l} \det A_{kk}.$$

15. Idempotent and Nilpotent Matrices

An $n \times n$ matrix N is said to be nilpotent if for some $k \geq 1$, $N^k = 0$. The smallest such k is called the degree of the nilpotent matrix N.

Example 15. *Any triangular matrix with* 0*s along the main diagonal is nilpotent.*

Theorem 122. *Let N be an $n \times n$ matrices. Then the following statements are equivalent*
1. N is nilpotent.
2. The eigenvalues of N are all zeros.
3. The characteristic polynomial of N is λ^n.
4. The minimal polynomial of N is λ^k for some $1 \leq k \leq n$.
5. $\operatorname{tr}\left(N^k\right) = 0$ for any $k \geq 1$.

Proof. To see $\operatorname{tr}\left(N^k\right) = 0$ for any $1 \leq k \leq n$ implies that the eigenvalues of N are all zeros, we consider the eigenvalues $\lambda_1, \cdots, \lambda_n$ of N. Then

$$p_k = \sum_{j=1}^{n} \lambda_j^k = 0,\ 1 \leq k \leq n.$$

Let $e_k (\lambda_1, \cdots, \lambda_n)$, $0 \leq k \leq n$ be the elementary symmetric polynomial that is the sum of all distinct products of k distinct variables. Hence,

$$\begin{aligned} e_0 &= 1, \\ e_1 &= \sum_{j=1}^{n} \lambda_j, \\ &\cdots \\ e_n &= \prod_{j=1}^{n} \lambda_j. \end{aligned}$$

Newton's identities

$$ke_k = \sum_{j=1}^{k} (-1)^{j-1} e_{k-j} p_j,\ 1 \le k \le n$$

imply that $e_k = 0$ for $1 \le k \le n$. Hence

$$\prod_{j=1}^{n} (\lambda - \lambda_j) = \sum_{j=0}^{n} (-1)^j e_j \lambda^{n-j} = \lambda^n$$

and $\lambda_j = 0$ for $1 \le j \le n$. □

Theorem 123 (Jordan–Chevalley decomposition). *Every matrix can be written as a sum of a diagonalizable (semi-simple) matrix and nilpotent matrix.*

Proof. From Jordan canonical form. □

An $n \times n$ matrix P is said to be idempotent (projection) if $P^2 = P$. The minimal polynomial of P has three possibilities

$$\lambda, \lambda - 1, \lambda (\lambda - 1)$$

and hence it is always diagonalizable and its eigenvalues are either 0 or 1 and $\operatorname{tr} P = \operatorname{rank} P$. A nonsingular idempotent matrix must be identity matrix. In general, P is a projection onto its range.

Proposition 9. *$P^2 = P$ iff*

$$(I - P)^2 = I - P.$$

An $n \times n$ matrix A is said to be an involution if $A^2 = I$. The minimal polynomial of P has three possibilities

$$\lambda + 1, \lambda - 1, (\lambda - 1)(\lambda + 1).$$

and hence it is always diagonalizable and its eigenvalues are either -1 or 1. A is involutory iff

$$\frac{1}{2}(A + I)$$

is idempotent.

Given a unit vector $v \in \mathbb{R}$, $n \ge 2$. A reflection about the hyperplane through the origin is represented by a matrix

$$R_v = I - 2vv^T.$$

A reflection matrix is an orthogonal matrix with determinant -1 and with eigenvalues $-1, 1, \cdots, 1$. It has minimal polynomial

$$\lambda^2 - 1$$

and hence it is diagonalizable.

Lemma 11. *Let $u \ne v$ be two vectors in $\mathbb{R}^n$ such that $\|u\| = \|v\|$. There exists a unique reflection R such that $Ru = v$.*

PROOF. We assume $\|u\| = \|v\| = 1$. Let $w = \frac{u-v}{\|u-v\|}$, and $R = I - 2ww^T$, then

$$\begin{aligned} Ru &= u - \frac{2(u-v)}{\|u-v\|^2}(u, u-v) \\ &= \frac{(2-2(u,v))u - 2(1-(u,v))u + 2v(1-(u,v))}{\|u-v\|^2} \\ &= v. \end{aligned}$$

To see the uniqueness, we observe that $Rv = R^2 u = u$ and $R(u-v) = v - u$. □

THEOREM 124 (Cartan-Dieudonné). *Every orthonal mapping on $\mathbb{R}^n$ which is not identity can be written as a composition of at most n reflections.*

PROOF. We proceed by induction on n. When $n = 1$, trivial. Suppose the theorem is ture for $n-1$, $n \geq 2$. Let $Q \neq I$ be an orthogonal map on $\mathbb{R}^n$.
Case 1. 1 is an eigenvalue of Q with a unit eigenvector w. Let $H = w^\perp$ be the hyperplane perpendicular to w. Then $Q : H \to H$ and $Q|_H$ is an orthogonal map which is not identity. Hence from the induction,

$$Q|_H = R_1 \circ R_2 \circ \cdots R_s$$

where $s \leq n-1$ and R_k, $1 \leq k \leq s$ are reflections on H. We can extend R_k to a reflection on X, still denoted by R_k, satisfying $R_k(w) = w$. Then we have

$$Q = R_1 \circ R_2 \circ \cdots R_s.$$

Case II. 1 is not an eigenvalue of Q. Pick any nonzero $w \in X$, then $v = Qw - w \neq 0$. Let

$$R_v = I - 2\frac{vv^*}{\|v\|^2}.$$

Then from the above lemma,

$$R_v Q w = w.$$

Since $R_v Q$ is an orthogonal map with eigenvalue 1, we have

$$R_v Q = R_1 \circ R_2 \circ \cdots R_s$$

for some $s \leq n-1$, and

$$Q = R_v \circ R_1 \circ R_2 \circ \cdots R_s.$$

The theorem thus follows from induction. □

REMARK 32. *Every orthogonal map Q is the composition of k hyperplane reflections, where*

$$k = n - \dim(\ker(Q - I))$$

and that this number is minimal.

16. Tensor Product

EXAMPLE 16. *Let $P_m(x)$ be the space of polynomials in x of degree less than m, and $P_n(y)$ be the space of polynomials in y of degree less than n. A natural basis for $P_m(x)$ is $1, x, \cdots, x^{m-1}$ and a natural basis for $P_n(y)$ is $1, y, \cdots, y^{n-1}$. Their tensor product $P_m(x) \otimes P_n(y)$ is the space of polynomials in x and y, of degree less than m in x, less than n in y. Hence $P_m(x) \otimes P_n(y)$ has a basis $x^i y^j$, $0 \leq i \leq m-1$, $0 \leq j \leq n-1$. And* $\dim(P_m(x) \otimes P_n(y)) = mn$.

In general, suppose $\{e_i\}_{i=1}^m$ is a basis of the linear space U and $\{f_j\}_{j=1}^n$ is a basis of the linear space V. The tensor product $U \otimes V$ can be defined as a linear space with basis $e_i \otimes f_j$, $1 \le i \le m$, $1 \le j \le n$.

Such definition is not good since it uses basis vectors.

We can define $U \otimes V$ in an invariant manner. Take the collection of all formal finite sums

$$\sum \alpha_i u_i \otimes v_i$$

where u_i and v_i are arbitrary vectors in U and V, respectively and α_i is a scalar. These sums form a linear space $F(U \times V)$.

The linear space Z spanned by sums of the forms

$$(u_1 + u_2) \otimes v - u_1 \otimes v - u_2 \otimes v,$$
$$u \otimes (v_1 + v_2) - u \otimes v_1 - u \otimes v_2,$$

and

$$(\alpha u) \otimes v - \alpha u \otimes v,$$
$$u \otimes (\alpha v) - \alpha u \otimes v$$

are called the space of null sums.

Definition 55. *The tensor product $U \otimes V$ of two finite-dimensional linear spaces U and V is the quotient space $F(U \times V)$ modulo the space Z of null sums.*

For simplicity, we still use $\sum \alpha_i u_i \otimes v_i$ to denote the equivalent class of $\sum \alpha_i u_i \otimes v_i \in F(U \times V)$ in $U \otimes V$. It is easy to verify that in $U \otimes V$,

$$(u_1 + u_2) \otimes v = u_1 \otimes v + u_2 \otimes v,$$
$$u \otimes (v_1 + v_2) = u \otimes v_1 + u \otimes v_2,$$

and

$$(\alpha u) \otimes v = u \otimes (\alpha v) = \alpha u \otimes v.$$

Theorem 125. *Let $\{e_i\}_{i=1}^m$ and $\{f_j\}_{j=1}^n$ be basis in U and V respectively. Then $e_i \otimes f_j$, $1 \le i \le m$, $1 \le j \le n$ form a basis for $U \otimes V$.*

Proof. It is easy to see that $e_i \otimes f_j$, $1 \le i \le m$, $1 \le j \le n$ span $U \otimes V$. Now we need to verify these mn vectors are linearly independent. Let $\{e_i^*\}_{i=1}^m$ and $\left\{f_j^*\right\}_{j=1}^n$ be the dual basis in U' and V', we define for any fixed k, l, a map

$$\phi_{kl} \in (F(U \times V))'$$

such that

$$\phi_{kl}\left(\sum \alpha_i u_i \otimes v_i\right) = \sum \alpha_i e_k^*(u_i) f_l^*(v_i).$$

Since $Z \in \ker \phi_{kl}$, ϕ_{kl} induces a map $\tilde{\phi}_{kl} \in (U \otimes V)'$. Now suppose

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} e_i \otimes f_j = 0,$$

applying ϕ_{kl} we have $c_{kl} = 0$. Hence $e_i \otimes f_j$, $1 \le i \le m$, $1 \le j \le n$ are linearly independent and they form a basis for $U \otimes V$. $\square$

When U and V are equipped with real Euclidean structure, there is a natural way to equip $U \otimes V$ with Euclidean structure. Choose orthonormal bases $\{e_i\}_{i=1}^m$, $\{f_j\}_{j=1}^n$ in U and V respectively, we declare $e_i \otimes f_j$, $1 \leq i \leq m$, $1 \leq j \leq n$ to be an orthonormal basis for $U \otimes V$. It remains to be shown that this Euclidean structure is independent of the choice of the orthonormal bases; this is easily done, based on the following lemma.

LEMMA 12. *Let $u_1, u_2 \in U$ and $v_1, v_2 \in V$, then*

$$(u_1 \otimes v_1, u_2 \otimes v_2) = (u_1, u_2)(v_1, v_2).$$

Finally, let's rewrite the proof of Schur's theorem on Hadamard product of positive matrices.

PROOF. Suppose $A, B > 0$. Then A, B can be written as Gram matrices:

$$A_{ij} = (u_i, u_j)$$

where $u_i \in U$, $1 \leq i \leq n$, are linearly independent, and

$$B_{ij} = (v_i, v_j)$$

where $v_i \in V$, $1 \leq i \leq n$, are linearly independent. Now define $g_i = u_i \otimes v_i$, then $g_i \in U \otimes V$, $1 \leq i \leq n$ are linearly independent and

$$(g_i, g_j) = A_{ij} B_{ij}.$$

Hence $A*B$ is a Gram matrix of linearly independent vectors and hence positive. □

17. Gershgorin's Circle Theorem

THEOREM 126. *Let A be an $n \times n$ matrix with complex entries. Decompose A as*

$$A = D + F$$

where $D = [d_1, \cdots, d_n]$ is the diagonal matrix equal to the diagonal of A; F has zero diagonal entries. Let f_i be the i-th row of F. Define the circular disc C_i to consist of all complex numbers z satisfying

$$|z - d_i| \leq \|f_i\|_1, \ 1 \leq i \leq n.$$

Here the 1-norm of a vector f is the sum of the absolute values of its components. Then every eigenvalue of A is contained in one of the discs C_i.

PROOF. Let u be an eigenvector of A, normalized as $\|u\|_\infty = 1$, where the ∞-norm is the maximum of the absolute value of the components of u. Clearly, $|u_j| \leq 1$ for each j and $|u_i| = 1$ for some i. The i-th component of

$$Au = \lambda u$$

can be written as

$$d_i u_i + f_i u = \lambda u_i$$

which implies

$$|z - d_i| \leq \|f_i\|_1.$$

□

Remark 33. *If C_i is disjoint from all the other Gershgorin discs, then C_i contains exactly one eigenvalue of A. This follows from the Kato's theorem on the continuity of roots for polynomials, which implies that for continuous matrix function $A(t)$, $t \in [a,b] \subset \mathbb{R}$, there exists n contiunous functions $\lambda_k(t)$, $t \in [a,b]$, $1 \leq k \leq n$ such that for each $t \in [a,b]$, $\lambda_k(t)$, $1 \leq k \leq n$ are eigenvalues of $A(t)$ with the same multiplicity.*

18. The motion of a rigid body

An isometry is a mapping of a Euclidean space into itself that preserves distances and an isometry M that preserves the origin is linear and satisfies

$$M^* M = I.$$

The determinant of such an isometry is ± 1 and for all rigid body motions is 1.

Theorem 127 (Euler). *A nontrivial isometry M of three-dimensional real Euclidean space with determinant 1 is a rotation; it has a uniquely defined axis of rotation and angle of rotation θ.*

Proof. The eigenvalues of M has two possibilities: All real; one real and two conjugate complex pairs. In the all real case, the eigenvalue is either $1,1,1$ or $1,-1,-1$, the $1,1,1$ case implies $M = I$ and the $1,-1,-1$ case is a rotation of π angle. In the latter case, the real eigenvalue must be 1 and we claim M is a rotation about the axis in v direction where v_1 is a unit eigenvector with eigenvalue 1. Let v_1, v_2, v_3 be an orthonormal basis, M under such basis has the form

$$M = \begin{pmatrix} 1 & & \\ & \alpha & -\beta \\ & \beta & \alpha \end{pmatrix}.$$

Since $\alpha^2 + \beta^2 = 1$, suppose $\alpha = \cos\theta$, $\beta = \sin\theta$, then M is a clockwise rotation of an angle θ about v_1. □

Remark 34. *θ can be calculated using*

$$\operatorname{tr} M = 1 + 2\cos\theta.$$

We turn now to rigid motions which keep the origin fixed and which depend on time t, that is, functions $M(t)$ whose values are rotations. Differentiate

$$MM^* = I,$$

we have

$$M_t M^* + M M_t^* = 0.$$

Let $A(t) = M_t M^*$, we have

$$A + A^* = 0,$$

i.e., A is anti-symmetric. Since

$$M_t = AM,$$

$A(t)$ is called the infinitesimal generator of the motion.

On the other hand, if $A(t)$ is antisymmetric and $M(t)$ is the unique solution to

$$M_t = AM,$$
$$M(0) = I,$$

then $M(t)$ is a rotation for each t. To see this, since $A = M^{-1}M_t$ and $A + A^* = 0$, we have

$$M^{-1}M_t + M_t^* (M^*)^{-1} = 0$$

i.e.

$$M_t M^* + M M_t^* = 0.$$

REMARK 35. *When A is a constant anti-symmetric matrix, we have*

$$M(t) = e^{At}.$$

REMARK 36. *In general, the solution to*

$$M_t = A(t) M,$$
$$M(0) = I,$$

is not given by

$$e^{\int_0^t A(s)ds}$$

unless $A(t)$ and $A(s)$ commute for all s and t.

19. Convexity

Let X be a linear space over the reals. For any pair of vectors x, y in X, the line segment with endpoints x and y is defined as the set of points in X of form

$$sx + (1-s)y, \; 0 \le s \le 1.$$

DEFINITION 56. *A set K in X is called convex if, whenever x and y belong to K, all points of the line segment with endpoints x, y also belong to K.*

Let l be a linear function in X; then the sets $l(x) = c$, called a hyperplane, $l(x) < c$, called an open half-space, $l(x) \le c$, called a closed half-space, are all convex sets.

EXAMPLE 17. *X the space of real, self-adjoint matrices, K the subset of positive matrices.*

EXAMPLE 18. *X the space of all polynomials with real coefficients, K the subset of all polynomials that are positive at every point of the interval $(0, 1)$.*

THEOREM 128. *(a) The intersection of any collection of convex sets is convex. (b) The sum of two convex sets is convex, where the sum of two sets K and H is defined as the set of all sums $x + y$, x in K, y in H.*

EXAMPLE 19. *A triangle in the plane is the intersection of three half-planes and hence is convex.*

DEFINITION 57. *A point x is called an interior point of a set S in X if for every y in X, $x + ty$ belongs to S for all sufficiently small positive t. A convex set K in X is called open if every point in it is an interior point.*

DEFINITION 58. *Let K be an open convex set that contains the vector 0. We define its gauge function (Minkowski functional) $p_K = p$ as follows: For every x in X,*

$$p(x) = \inf \left\{ r > 0 : \frac{x}{r} \in K \right\}.$$

Example 20. *Let K be the open ball of radius a centered at the origin in a Euclidean space X, then K is a convex set and*

$$p(x) = \frac{\|x\|}{a}.$$

Theorem 129. *(a) The gauge function p of an open convex set K that contains the origin is well-defined for every x.*
(b) p is positive homogeneous:

$$p(ax) = ap(x) \text{ for } a > 0.$$

(c) p is subadditive:

$$p(x+y) \leq p(x) + p(y).$$

(d) $p(x) < 1$ iff x is in K.

Proof. (a) Since $0 \in K$ and K is open, $\left\{r > 0 : \frac{x}{r} \in K\right\}$ is not empty.
(b) Follows from definition.
(c) For any $\varepsilon > 0$, let $s = p(x) + \varepsilon$ and $t = p(y) + \varepsilon$. We have

$$\frac{x}{s}, \frac{y}{t} \in K.$$

Hence the convexity of K implies

$$\frac{x+y}{s+t} = \frac{s}{s+t}\frac{x}{s} + \frac{t}{s+t}\frac{y}{t} \in K.$$

By definition of p,

$$p(x+y) \leq s + t = p(x) + p(y) + 2\varepsilon.$$

Since $\varepsilon > 0$ can be chosen arbitrarily small, we have

$$p(x+y) \leq p(x) + p(y).$$

(d) Follows from the open and convexity of K. □

Theorem 130. *Let p be a positive homogeneous, subadditive function. Then the set K consisting of all x for which $p(x) < 1$ is convex and open.*

Proof. Trivial. □

Theorem 131. *Let K be an open convex set, and let y be a point not in K. Then there is an open half-space containing K but not y.*

Proof. It is a consequence of Hahn Banach Theorem. We define a linear function l on $Y = span\{y\}$ so that $l(y) = 1$. Then

$$l(ty) \leq p_K(ty) \text{ for any } t \in \mathbb{R},$$

i.e., $l \leq p_K$ on Y. Extend l to the whole space by Hahn_Banach theorem. Then we have $l(y) = 1$ and $l(x) \leq p_K(x) < 1$ for any $x \in K$. □

Theorem 132 (Finite dimensional Hahn-Banach theorem). *Let p be a real-valued positive homogeneous subadditive function defined on a linear space X over $\mathbb{R}$. Let U be a subspace of X on which a linear function l is defined, satisfying*

$$l(u) \leq p(u) \text{ for all } u \in U.$$

Then l can be extended to all of X so that $l(x) \leq p(x)$ is satisfied for all x.

PROOF. Pick $z \notin U$. Need to define $l(z) = a$ so that

$$l(u+z) = l(u) + a \leq p(u+z),$$

i.e.,

$$a \leq p(u+z) - l(u)$$

for any $u \in U$. We also need

$$l(v-z) = l(v) - a \leq p(v-z),$$

for any $v \in U$, i.e.,

$$a \geq l(v) - p(v-z).$$

Such a can be chosen if for any $u, v \in U$

$$l(v) - p(v-z) \leq p(u+z) - l(u),$$

which is equivalent to

$$p(u+z) + p(v-z) \geq l(u+v)$$

which holds since

$$p(u+z) + p(v-z) \geq p(u+v) \geq l(u+v).$$

For such a, the linear extension l defined on $V = \{tz + u : t \in \mathbb{R},\, u \in U\}$ by

$$l(u+tz) = l(u) + ta$$

satisfies $l(x) \leq p(x)$ for all $x \in V$. The theorem follows by induction. □

THEOREM 133. *Let K and H be open convex sets that are disjoint. Then there is a hyperplane that separates them. That is, there is a linear function l and a constant d such that*

$$l(x) < d \text{ on } K,\ l(y) > d \text{ on } H.$$

PROOF. $K - H$ is an open and convex set. Since K and H are disjoint, $K - H$ does not contain the origin. With $y = 0$, there is a linear function l that is negative on $K - H$, i.e.,

$$l(x-y) < 0 \text{ for } x \in K \text{ and } y \in H.$$

We can rewrite this as

$$l(x) < l(y) \text{ for } x \in K \text{ and } y \in H.$$

Hence there is a number d such that

$$l(x) \leq d \leq l(y) \text{ for } x \in K \text{ and } y \in H$$

Since both K and H are open, the sign of equality cannot hold. □

DEFINITION 59. *Let $S \subset X$. Its support function q_S on the dual X' of X is defined by*

$$q_S(l) = \sup_{x \in S} l(x) \text{ for any } l \in X'.$$

PROPOSITION 10. *(a) q_S is subadditive in l:*

$$q_S(l+m) \leq q_S(l) + q_S(m).$$

(b)

$$q_{S+T}(l) = q_S(l) + q_T(l).$$

(c)

$$q_{S \cup T}(l) = \max\{q_S(l), q_T(l)\}.$$

Theorem 134. *Let K be an open convex set, q_K its support function. Then $x \in K$ iff*

$$l(x) < q_K(l) \text{ for any } l \in X'.$$

Proof. From the definition, for any $x \in K$, $l(x) \le q_K(l)$ for any $l \in X'$. Since K is open, equality can't hold. On the other hand, if $y \notin K$, there exists l such that $l(y) = 1$ but $l(x) < 1$ for any $x \in K$ and hence $l(y) \ge q_K(l)$ for such l. □

Remark 37. *The above theorem describes open convex sets as intersections of half-spaces.*

Definition 60. *A convex set K in X is called closed if every open segment $(1-t)x + ty$, $0 < t < 1$, that belongs to K has its endpoints x and y in K.*

Remark 38. *The intersection of closed convex sets is a closed convex set.*

Theorem 135. *Let K be a closed, convex set, and $y \notin K$. Then there is a closed half-space that contains K but not y.*

Theorem 136. *Let K be a closed, convex set, q_K its support function. Then $x \in K$ iff*

$$l(x) \le q_K(l) \text{ for any } l \in X'.$$

Remark 39. *The above theorem describes closed convex sets as intersections of half-spaces.*

Definition 61. *Let S be an arbitrary set in X. The closed convex hull of S is defined as the intersection of all closed convex sets containing S.*

Theorem 137. *The closed convex hull of any set S is the set of points x satisfying*

$$l(x) \le q_S(l) \text{ for any } l \in X'.$$

Proof. Let K be the collection of x such that

$$l(x) \le q_S(l) \text{ for any } l \in X'.$$

Then K is closed and convex. Let $S \subset K'$, where K' is a closed and convex set. Since

$$q_S(l) \le q_{K'}(l) \text{ for any } l \in X'.$$

We have for any $x \in K$,

$$l(x) \le q_S(l) \le q_{K'}(l) \text{ for any } l \in X'.$$

Hence $x \in K'$. So we have $K \subset K'$ and K is the closed convex hull of S. □

Let $x_1, \cdots, x_m$ denote m points in X, and $p_1, \cdots, p_m$, denote m nonnegative numbers whose sum is 1. Then

$$x = \sum p_j x_j$$

is called a convex combination of $x_1, \cdots, x_m$.

Remark 40. *Let $x_1, \cdots, x_m$ belong to a convex set, then so does any convex combination of them.*

Definition 62. *A point of a convex set K that is not an interior point is called a boundary point of K.*

DEFINITION 63. *Let K be a closed, convex set. A point e of K is called an extreme point of K if it is not the interior point of a line segment in K. That is, x is not an extreme point of K if*

$$x = \frac{y+z}{2},\ y, z \in K,\ y \neq z.$$

REMARK 41. *An extreme point must be a boundary point but a boundary point may not be an extreme point.*

DEFINITION 64. *A convex set K is called bounded if it does not contain a ray, that is, a set of points of the form $x + ty$, $t > 0$.*

THEOREM 138 (Caratheodory). *Let K be a nonempty closed bounded convex set in X, $\dim X = n$. Then every point of K can be represented as a convex combination of at most $(n+1)$ extreme points of K.*

PROOF. We prove this inductively on the dimension of X. The theorem obviously holds for $\dim X = 1$. Suppose the theorem holds for $\dim X \leq n-1$. Now for $\dim X = n$, we distinguish two cases:

(i) K has no interior points. WLOG, we assume $0 \in K$. We claim that K does not contain n linearly independent vectors; for if it did, the convex combination of these vectors and the origin would also belong K; but these points constitute an n-dimensional simplex, full of interior points. Let $m < n$ be the largest number of linearly independent vectors in K, then K is contained in an m-dimensional subspace of X. By the induction hypothesis, theorem holds for K.

(ii) K has interior points. Denote by K_0 the set of all interior points of K. It is easy to show that K_0 is convex and open and we further assume WLOG $0 \in K_0$. We claim that K has boundary points; for, since K is bounded, any ray issuing from any interior point of K intersects K in an interval; since K is closed, the other endpoint is a boundary point y of K. Let y be a boundary point of K. We apply Theorem 131 to K_0 and y; there is a linear functional l such that

$$l(y) = 1,$$
$$l(x_0) < 1 \text{ for all } x_0 \in K_0.$$

We claim that $l(x_1) \leq 1$ for all $x_1 \in K$. Pick any interior point $x_0 \in K_0$; then all points x on the open segment bounded by x_0 and x_1, are interior points of K, and so $l(x) < 1$. It follows that at the endpoint x_1, $l(x_1) \leq 1$. Let

$$K_1 = \{x \in K : l(x) = 1\}.$$

Then K_1 is bounded and convex. Since $y \in K_1$, so K_1 is nonempty. We claim that every extreme point e of K_1, is also an extreme point of K; for suppose that

$$e = \frac{z+w}{2}$$

for some $z, w \in K$, then we have

$$1 = l(e) = \frac{l(z) + l(w)}{2} \leq 1$$

which implies $l(z) = l(w) = 1$ and hence $z, w \in K_1$. But since e is an extreme point of K_1, $z = w$. This proves that extreme points of K_1, are extreme points of K. Since K_1 lies in a hyperplane of dimension $= n-1$, it follows from the induction assumption that every point in K_1, in particular $y \in K_1$, can be written as a convex combination of n extreme points of K_1, and hence n extreme points of K. This proves that every boundary point of K can be written as a convex combination of n extreme points of K. Now let x_0 be an interior point of K. We take any extreme point e of K and look at the intersection of the line through x_0 and e with K which is a closed interval. Since e is an extreme point of K, e is one of the end points; denote the other end point by y which clearly is a boundary point of K. Since by construction x_0 lies on this interval, it can be written in the form

$$x_0 = py + (1-p)\,e$$

for some $0 < p < 1$. Since y can be written as a convex combination of n extreme points of K we conclude that x_0 can be written as a convex combination of $n+1$ extreme points. The proof is then complete. □

Definition 65. *An $n \times n$ matrix $S = (s_{ij})$ is called doubly stochastic if*

$$s_{ij} \geq 0,$$
$$\sum_i s_{ij} = 1 \text{ for each } j,$$
$$\sum_j s_{ij} = 1 \text{ for each } i.$$

The doubly stochastic matrices form a bounded, closed convex set in the space of all $n \times n$ matrices.

Given a permutation p of the integers $(1, \cdots, n)$, the permutation matrix P associated with p is the matrix (p_{ij}) such that

$$p_{ij} = 1 \text{ if } j = p(i) \text{ and } p_{ij} = 0 \text{ otherwise.}$$

Theorem 139 (Denes Konig, Garrett Birkhoff). *The permutation matrices are the extreme points of the set of doubly stochastic matrices.*

Proof. Let's show that any extreme point must be a permutation matrix. Suppose S is not a permutation matrix. Then there exists $i_0 j_0$ such that $s_{i_0 j_0} \in (0,1)$. Theres exists $j_1 \neq j_0$ such that $s_{i_0 j_1} \in (0,1)$ which implies the existence of $i_1 \neq i_0$ such that $s_{i_1 j_1} \in (0,1)$. There must be a minimal closed chain

$$i_k j_k \to i_k j_{k+1} \to i_{k+1} j_{k+1} \to \cdots \to i_{m-1} j_m \to i_m j_m = i_k j_k.$$

A minimal chain contain either two or zero points in any row or column. Define N such that its entries are alternatively 1 and -1 at the chain positions and zero otherwise. Then S is the middle point of

$$S \pm \varepsilon N$$

which, for ε sufficiently small, are two distinct doubly stochastic matrices. Hence S can't be extreme point. □

Corollary 14. *Every doubly stochastic matrix can be written as a convex combination of permutation matrices.*

Theorem 140 (Helly). *Let X be a linear space of dimension n over the reals. Let $\{K_j\}_{j=1}^N$ be a collection of N convex sets in X. Suppose that every subcollection of $n+1$ sets K has a nonempty intersection. Then all K in the whole collection have a common point.*

Proof. We argue by induction on N. It is trivial when $N = n + 1$. Suppose $N > n + 1$ and that the assertion is true for $N - 1$. Then for any set K_j, there exists

$$x_j \in \bigcap_{1 \le i \le N, i \ne j} K_i.$$

We claim that there exists N numbers $a_1, \cdots, a_N$ not all equal to zero such that

$$\sum_{j=1}^{N} a_j x_j = 0 \text{ and } \sum_{j=1}^{N} a_j = 0$$

since we have $n + 1$ equations for the $N > n + 1$ unknowns. All a_j can not be of the same sign so we could renumber them so that $a_1, \cdots, a_k$ are positive, the rest nonpositive. Let

$$a = \sum_{j=1}^{k} a_j = - \sum_{j=k+1}^{N} a_j > 0.$$

We define

$$y = \frac{1}{a} \sum_{j=1}^{k} a_j x_j = -\frac{1}{a} \sum_{j=k+1}^{N} a_j x_j.$$

Then y is a common point for all K in the whole collection. □

20. Positive matrices

Definition 66. *A real $n \times n$ matrix P is called entrywise positive if all its entries p_{ij} are positive real numbers.*

Theorem 141 (Perron). *Every positive matrix P has a dominant eigenvalue, denoted by $\lambda(P)$ which has the following properties:*
(i) $\lambda(P) > 0$ and the associated eigenvector h has positive entries.
(ii) $\lambda(P)$ is a simple eigenvalue.
(iii) Every other eigenvalue μ of P is less than $\lambda(P)$ in absolute value.
(iv) P has no other eigenvector with nonnegative entries.

Proof. (i) Let $p(P)$ be the set of all nonnegative numbers λ for which there is a nonnegative vector $x \ne 0$ such that

$$Px \ge \lambda x, x \ge 0.$$

We claim that $p(P) = [0, \lambda^*]$ for some $\lambda^* > 0$. It suffices to show that $p(P)$ is bounded and closed. The boundedness follows from

$$\lambda \|x\|^2 \le x^T P x \le \max_{\mu \in \sigma\left(\frac{P+P^T}{2}\right)} \mu \|x\|^2.$$

$p(P)$ is closed follows from the standard compactness argument.

We claim that λ^* is the dominant eigenvalue. Let $h \ge 0$, $h \ne 0$ be such that

$$Ph \ge \lambda^* h.$$

We claim $Ph = \lambda^* h$. If not, then for some i, we have

$$\sum P_{ij} h_j > \lambda^* h_i,$$
$$\sum P_{kj} h_j \ge \lambda^* h_k \text{ for } k \ne i.$$

Now consider $x = h + \varepsilon e_i$, we have for $\varepsilon > 0$ sufficiently small,

$$Px > \lambda^* x$$

which contradicts with the maximality of λ^*. Hence λ^* is an eigenvalue, and we write $\lambda(P) = \lambda^*$.

Now $h > 0$ follows from $Ph > 0$.

(ii) To see $\lambda(P)$ is a simple eigenvalue. First, the eigenspace of $\lambda(P)$ must have dimension 1, otherwise it must contain an nonzero element with some zero entry. Next, we need to show that $\lambda(P)$ has no generalized eigenvectors. Suppose

$$(P - \lambda^* I)\, y = h$$

holds for some y, we can assume that y is positive by replacing y by $y + Kh$. Hence for $y > 0$, we have

$$Py = \lambda^* y + h > \lambda^* y$$

which is impossible.

(iii) Let $\mu \neq \lambda^*$ be another eigenvalue with eigenvector y. We have for each i,

$$\sum P_{ij} |y_j| \geq \left| \sum P_{ij} y_j \right| = |\mu y_i| = |\mu| \, |y_i| \, .$$

Hence $|\mu| \in P(p)$ and $|\mu| \leq \lambda^*$. If $|\mu| = \lambda^*$, then $|y_j| = ch_j$ for some $c > 0$ and for each i,

$$\sum_j P_{ij} |y_j| = \left| \sum_j P_{ij} y_j \right| .$$

Hence $y_j = ce^{i\theta} h_j$ for each j and $\mu = \lambda^*$, a contradiction.

(iv) Recall that the product of eigenvectors of P and its transpose P^T pertaining to different eigenvalues is zero. Since p^T also is positive, the eigenvector pertaining to its dominant eigenvalue, which is the same as that of P, has positive entries. Since a positive vector does not annihilate a nonnegative vector, P has no other eigenvector with nonnegative entries. □

Definition 67. *An $n \times n$ matrix $S = (s_{ij})$ is called a stochastic matrix if*

$$s_{ij} \geq 0 \textit{ for each } 1 \leq i, j \leq n$$

and

$$\sum_{i=1}^{n} s_{ij} = 1 \textit{ for each } 1 \leq j \leq n.$$

Theorem 142. *Let S be an entrywise positive stochastic matrix.*
(i) The dominant eigenvalue $\lambda(S) = 1$.
(ii) Let $x \neq 0$ be any nonnegative vector; then

$$\lim_{N \to \infty} S^N x = ch$$

where $h > 0$ is the dominant eigenvector and $c > 0$ is a constant depending on x.

Proof. (i) S^T is entrywise positive. Since

$$S^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

Perron theorem implies that 1 is the dominant eigenvalue of S^T, and hence S since they share the same set of eigenvalues.

(ii) Assuming that all eigenvectors of S are genuine, we can write

$$x = \sum_{j=1}^{n} c_j h_j$$

where h_1 is the dominant eigenvector and $\{h_j\}$ are eigenvectors forming an orthonormal basis. Then we have

$$S^N x = \sum_{j=1}^{n} c_j \lambda_j^N h_j \to c_1 h_1 \text{ as } N \to \infty.$$

To see $c_1 > 0$, denote $\xi = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, we have

$$\left(S^N x, \xi\right) = \left(x, \left(S^T\right)^N \xi\right) = (x, \xi),$$

hence

$$(c_1 h_1, \xi) = c_1 (h_1, \xi) = (x, \xi) > 0.$$

In general, if h_j is a generalized eigenvector for eigenvalue λ_j with $|\lambda_j| < 1$, we still have

$$\lim_{N \to \infty} S^N h_j = 0.$$

Actually, restricting S to the invariant space $N_{d_j}(\lambda_j)$, it suffices to show that for each y,

$$\lim_{N \to \infty} T^N y = 0$$

if all eigenvalues of T has absolute value smaller than 1. □

EXAMPLE 21. *Let x_j be the population size of the jth species, $1 \le j \le n$; suppose that during a unit of time (a year, a day, a nanosecond) each individual of the collection changes (or gives birth to) a member of the other species according to the probabilities s_{ij}. If the population size is so large that fluctuations are unimportant, the new size of the population of the ith species will be*

$$y_i = \sum_{j=1}^{n} s_{ij} x_j,$$

i.e.,

$$y = Sx.$$

The above theorem claims that as $N \to \infty$, such populations tend to a steady distribution that does not depend on where the population started from.

Perron's theorem fails if the matrix is only assumed to be entrywise nonnegative. For example, we could consider matrices

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

The first one has a dominant eigenvalue; the second has plus or minus 1 as eigenvalues, neither dominated by the other; the third has 1 as a double eigenvalue.

Theorem 143 (Frobenius). *Every nonnegative $n \times n$ matrix $F \neq 0$, has an eigenvalue $\lambda(F)$ with the following properties:*
(i) $\lambda(F)$ is nonnegative, and the associated eigenvector has nonnegative entries:

$$Fh = \lambda(F)\, h, h \geq 0.$$

(ii) Every other eigenvalue μ is less than or equal to $\lambda(F)$ in absolute value:

$$|\mu| \leq \lambda(F).$$

(iii) If $|\mu| = \lambda(F)$, then μ is of the form

$$\mu = e^{\frac{2\pi i k}{m}} \lambda(F)$$

where k and m are positive integers, $m \leq n$.

Proof. (i), (ii) follows from the continuity of eigenvalues.
(iii) Interested students could check the textbook by Lax for the proof. □

21. Normed linear spaces

Definition 68. *Let X be a linear space over $\mathbb{R}$. A function, denoted by $\|\cdot\|$, from X to $\mathbb{R}$ is called a norm if it satisfies:*
(i) Positivity: $\|x\| > 0$ for $x \neq 0$, $\|0\| = 0$.
(ii) Subadditivity(Triangle inequality):

$$\|x + y\| \leq \|x\| + \|y\|$$

for any $x, y \in X$.
(iii) Homogeneity: for any real number k, $\|kx\| = |k| \|x\|$.

Definition 69. *A linear space with a norm is called a normed linear space.*

Normed linear space is a metric space with distance defined for any $x, y \in X$ by

$$d(x, y) = \|x - y\|.$$

Hence, the open unit ball is defined by

$$\{x \in X : \|x\| < 1\}$$

and the closed unit ball is defined by

$$\{x \in X : \|x\| \leq 1\}.$$

And in general, we can define the open ball

$$B(y, r) = \{x \in X : \|x - y\| < r\}.$$

Remark 42. *The balls are convex. Let $K = B(0,1)$ and p_K its gauge function, then for any $x \in X$,*

$$p_K(x) = \|x\|.$$

Example 22. *Given $p \geq 1$, we can define for any $x \in \mathbb{R}^n$,*

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}}.$$

When $p = \infty$, we can define for any $x \in \mathbb{R}^n$,

$$\|x\|_\infty = \max_{1 \leq j \leq n} |x_j|.$$

LEMMA 13 (Hölder's inequality). *Suppose* $1 \le p, q \le \infty$ *and*

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then for any $x, y \in \mathbb{R}^n$,

$$(x, y) \le \|x\|_p \|y\|_q .$$

COROLLARY 15. *Suppose* $1 \le p, q \le \infty$ *and*

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then for any $x \in \mathbb{R}^n$,

$$\|x\|_p = \max_{\|y\|_q = 1} (x, y) .$$

PROOF. Choose y such that

$$y_j = \frac{\operatorname{sgn} x_j \, |x_j|^{p/q}}{\|x\|_p^{p/q}},$$

then $\|y\|_q = 1$ and

$$(x, y) = \|x\|_p .$$

□

THEOREM 144. *For* $1 \le p \le \infty$, $\|\cdot\|_p$ *is a norm on* $\mathbb{R}^n$ *called the* l^p *norm.*

PROOF. One only need to verify subadditivity which is obvious if $p = \infty$.

$$\begin{aligned} \|x + y\|_p &= \max_{\|z\|_q = 1} (x + y, z) \le \max_{\|z\|_q = 1} (x, z) + \max_{\|z\|_q = 1} (y, z) \\ &= \|x\|_p + \|y\|_p . \end{aligned}$$

□

REMARK 43. *For any* $x \in \mathbb{R}^n$,

$$\|x\|_\infty = \lim_{p \to \infty} \|x\|_p .$$

And $\|\cdot\|_2$ *is the standard Euclidean norm.*

DEFINITION 70. *Two norms in a linear space* X, $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *are called equivalent if there are constants* $c \ge 1$ *such that for all* $x \in X$,

$$\|x\|_1 \le c \|x\|_2 \text{ and } \|x\|_2 \le c \|x\|_1 .$$

THEOREM 145. *In a finite-dimensional linear space, all norms are equivalent.*

PROOF. WLOG, we assume $X = \mathbb{R}^n$. Let $\|\cdot\|$ be an arbitrary norm and $\|\cdot\|_2$ be the standard Euclidean norm. It suffices to show these two norms are equivalent.

(a)

$$\begin{aligned} \|x\| &\le \sum_{j=1}^n \|x_j e_j\| = \sum_{j=1}^n |x_j| \, \|e_j\| \le \left(\sum_{j=1}^n |x_j|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n \|e_j\|^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{j=1}^n \|e_j\|^2 \right)^{\frac{1}{2}} \|x\|_2 . \end{aligned}$$

(b) We claim that $\|\cdot\|$ is a continuous function in $\mathbb{R}^n$. To see this, suppose $\lim_{k\to\infty} x_{(k)} \to x$, we have

$$\begin{aligned}|\|x\| - \|x_k\|| &\leq \|x - x_k\| \\ &\leq \left(\sum_{j=1}^{n} \|e_j\|^2\right)^{\frac{1}{2}} \|x - x_k\|_2 \to 0\end{aligned}$$

as $k \to \infty$. $\|\cdot\|$ is positive on the compact set

$$\{x : \|x\|_2 = 1\},$$

hence,

$$\inf_{\|x\|_2=1} \|x\| = m > 0$$

which implies

$$\|x\| \geq m \|x\|_2 .$$

□

Definition 71. *A sequence* $\{x_j\}$ *in a normed linear space is called convergent to x, denoted by*

$$\lim_{j\to\infty} x_j = x,$$

if

$$\lim_{j\to\infty} \|x_j - x\| = 0.$$

In a finite-dimensional linear space, convergence is independent of the choice of norms.

Definition 72. *A set S in a normed linear space is called closed if it contains the limits of all convergent sequences in S.*

Definition 73. *A set S in a normed linear space is called bounded if it is contained in some ball.*

Definition 74. *Cauchy sequence.*

Theorem 146. *Let X be a finite-dimensional normed linear space X.*
(i) Completeness: Every Cauchy sequence converges to a limit.
(ii) Local compactness: Every bounded infinite sequence has a convergent subsequence.

Theorem 147 (Riesz). *Let X be a normed linear space that is locally compact. Then X is finite dimensional.*

Proof. Suppose $\dim X = \infty$; we shall construct a bounded sequence $\{y_k\} \subset X$ which contains no convergent subsequence. Indeed, we shall construct a sequence $\{y_n\}$ recursively such that

$$\|y_k\| = 1 \text{ for each } k,$$
$$\|y_k - y_l\| \geq 1 \text{ for any } k \neq l.$$

Choose any unit vector $\|y_1\|$. Suppose $\{y_j\}_{j=1}^{k-1}$ have been chosen; denote by Y the space spanned by them. Since X is infinite-dimensional, there exists $x \in X\backslash Y$ and $y_0 \in Y$

$$d = \inf_{y\in Y} \|x - y\| = \|x - y_0\| .$$

Let

$$y_k = \frac{x - y_0}{d},$$

then $\|y_k\| = 1$ and for any $y \in Y$,

$$\|y_k - y\| = \frac{1}{d} \|x - y_0 - dy\| \geq 1.$$

□

LEMMA 14. *Let Y be a finite-dimensional subspace of a normed linear space X. Let $x \in X \backslash Y$. Then*

$$d = \inf_{y \in Y} \|x - y\| > 0$$

and the infimum is achieved.

PROOF. There exists $y_k \in Y$, such that

$$\lim_{k \to \infty} \|x - y_k\| = d.$$

Hence $\{y_k\}$ is a bounded sequence in Y and it contains a subsequence converging to $y_0 \in Y$. Hence

$$d = \|x - y_0\|$$

Since $x \notin Y$, $d > 0$. □

THEOREM 148. *Let X be a finite-dimensional normed linear space, and T be a linear function defined on X. Then there is a constant c such that*

$$|Tx| \leq c \|x\| \text{ for any } x \in X$$

and T is continuous.

PROOF.

$$|Tx| = |(x, y)| \leq \|x\|_2 \|y\|_2 \leq c \|x\|.$$

□

THEOREM 149. *Let X be a finite-dimensional normed linear space. X' is a normed space with the norm defined by*

$$\|T\| = \sup_{\|x\|=1} |Tx|.$$

Recall that we can identify X'' with X

$$x^{**}(l) = l(x).$$

THEOREM 150. *Let X be a finite-dimensional normed linear space. The norm induced in X'' by the induced norm in X' is the same as the original norm in X.*

PROOF.

$$\|x^{**}\| = \sup_{\|l\|=1} |x^{**}(l)| = \sup_{\|l\|=1} |l(x)| \leq \|x\|.$$

On the other hand, we can construct, using Hanh-Banach theorem $\|l\| = 1$ and $|l(x)| = 1$. □

Let X and Y be a pair of finite-dimensional normed linear spaces over the reals.

LEMMA 15. *For any linear map $T : X \to Y$, there is a constant c such that for all $x \in X$,*

$$\|Tx\|_Y \le c\,\|x\|_X\,.$$

Hence T is continuous.

THEOREM 151. *The linear maps between X and Y form a normed linear space $L(X,Y)$ with the norm*

$$\|T\| = \sup_{\|x\|_X=1} \|Tx\|_Y\,.$$

Let $T \in L(X,Y)$ and $T' \in L(Y',X')$ be its transpose so we have for any $l \in Y'$,

$$(T'l, x) = (l, Tx)\,.$$

THEOREM 152.

$$\|T'\| = \|T\|\,.$$

PROOF. Let $l \in Y'$ and $m = T'l \in X'$.

$$\|T'l\|_{X'} = \sup_{\|x\|=1} |(T'l,x)| = \sup_{\|x\|=1} |(l,Tx)| \le \|l\|_{Y'} \sup_{\|x\|=1} \|Tx\| = \|T\|\,\|l\|_{Y'}\,,$$

hence

$$\|T'\| \le \|T\|\,.$$

Similarly, we have

$$\|T''\| \le \|T'\|\,.$$

Since $T'' = T$, we conclude

$$\|T'\| = \|T\|\,.$$

□

Let T be a linear map of a linear space X into Y, S another linear map of Y into another linear space Z. Then we can define the product ST as the composite mapping of T followed by S.

THEOREM 153. *Suppose X, Y, and Z above are normed linear spaces; then*

$$\|ST\| \le \|S\|\,\|T\|\,.$$

A mapping T of one linear space X into another linear space Y is called invertible if it maps X onto Y, and is one-to-one. In this case T has an inverse, denoted as T^{-1}.

THEOREM 154. *Let X and Y be finite-dimensional normed linear spaces of the same dimension, and let T be a linear mapping of X into Y that is invertible. Let S be another linear map of X into Y. Suppose that*

$$\|S - T\| < \frac{1}{\|T^{-1}\|}.$$

Then S is invertible.

PROOF. It suffices to show that S is one-to-one. Suppose

$$Sx_0 = 0$$

for some $x_0 \in X$, we have

$$Tx_0 = (T - S)\,x_0,$$

hence

$$x_0 = T^{-1}\left((T-S)\,x_0\right).$$

So

$$\|x_0\| \leq \left\|T^{-1}\right\| \|T-S\| \|x_0\|,$$

hence $x_0 = 0$. ☐

REMARK 44. *Above theorem holds for infinite dimensional normed linear spaces which are complete.*

DEFINITION 75. *Let X, Y be a pair of normed linear spaces. A sequence $\{T_n\}$ of linear maps of X into Y is said to converge to the linear map T, denoted as*

$$\lim_{n\to\infty} T_n = T$$

if

$$\lim_{n\to\infty} \|T_n - T\| = 0.$$

THEOREM 155. *Let X be a complete normed linear space, R a linear map of X into itself with $\|R\| < 1$. Then $S = I - R$ is invertible, and*

$$S^{-1} = \sum_{k=0}^{\infty} R^k.$$

THEOREM 156. *Let X and Y be complete normed linear spaces, and let T be a linear mapping of X into Y that is invertible. Let S be another linear map of X into Y. Suppose that*

$$\|S - T\| < \frac{1}{\|T^{-1}\|}.$$

Then S is invertible.

PROOF.

$$S = S - T + T = T\left(I - T^{-1}(T-S)\right).$$

☐