People's Democratic Republic of Algeria
Ministry of higher education and scientific research
Ahmed Zabana Relizane University
Faculty of science and technology

جامعة أحمد زبانة-غليزان
Ahmed Zabana Relizane University

ALGEBRA (4)

Lessons and Exercises.

Course intended primarily for students of ”Licence2”

Mehdi SLIMANE

Mathematics department

Licence 2

June 10, 2025

Table of contents

Introduction

This course is an introduction to algebra 4 which builds on the idea of linear algebra. We study the properties of mappings of several variables that are linear in each variable separately.

Chapters one and two are reviews of vector spaces, linear transformations and the inner product spaces. Then we discuss Linear forms and Duality in chapter three. Afterward some applications about Bilinear forms are given in chapter four.

Finally, in chapter five treats the Quadratic and Hermitian forms and their classifications.

Chapter 1

Review of vector spaces and matrices

Chapter contents

1.1 Vector spaces

Definition 1.1.1 Field

A **Field** is a set $\mathbb{F} \neq \emptyset$ with two operations $+$ and $\cdot$ satisfying the following properties

(1) $x + y = y + x$ for all x, y in $\mathbb{F}$.

(2) $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbb{F}$.

(3) there is a unique element 0 (zero) in F such that $x + 0 = x$ for every x in $\mathbb{F}$.

(4) to each x in $\mathbb{F}$ there corresponds a unique element $(-x)$ in $\mathbb{F}$ such that $x + (-x) = 0$.

(5) $xy = yx$ for all x, y in $\mathbb{F}$.

(6) $(xy)z = x(yz)$ for all $x, y, z \in \mathbb{F}$.

(7) There is a unique non-zero element 1 (one) in $\mathbb{F}$ such that $x1 = x$, for every $x \in \mathbb{F}$.

(8) To each $x \neq 0$ in $\mathbb{F}$ there corresponds a unique element x^{-1} in F such that $xx^{-1} = 1$.

Definition 1.1.2 Characteristic of a Field

The smallest positive whole number n such that the sum of the multiplicative identity added to itself n times equals the additive identity. If no such n exists, the field is said to have characteristic zero.

Definition 1.1.3 Vector space

A **vector space** over a field $\mathbb{F}$ is a set V with two operations $+$ and $\cdot$ satisfying the following properties for all $u, v, w \in V$ and $a, b \in \mathbb{F}$:

(1) $u + v \in V$.

(2) $u + v = v + u$.

(3) $(u + v) + w = u + (v + w)$.

(4) there is a special vector $0_V \in V$ such that $u + 0_V = v$ for all u in V.

(5) for every $u \in v$ there exists $w = -v \in V$ such that $v + w = 0_V$.

(6) $a \cdot v \in V$.

(7) $(a + b) \cdot v = a \cdot v + b \cdot v$.

(8) $a \cdot (u + v) = a \cdot u + a \cdot v$.

(9) $(ab) \cdot v = a \cdot (b \cdot v)$.

(10) $1 \cdot v = v$ for all $v \in V$.

1.2 Some examples of vector spaces

Let $\mathbb{F}$ be a field.

(A) The set $\mathbb{F}^n = \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{F}\}$ is a vector space over $\mathbb{F}$:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n);$$

$$b(a_1, \ldots, a_n) = (ba_1, \ldots, ba_n).$$

(B) The set $\mathbb{F}[X]$ of polynomials with coefficients in $\mathbb{F}$ is a vector space over $\mathbb{F}$.

(C) The set $\mathbb{F}_n[X]$ of polynomials of degree less than or equal n form a vector space over $\mathbb{F}$.

(D) The space of functions from a set to a field. let S be any non-empty set. Let V be the set of all functions from the set S into $\mathbb{F}$. The sum of two vectors f and g in V is the vector $f + g$, i.e., the function from S into F, defined by

$$(f + g)(s) = f(s) + g(s).$$

The product of the scalar c and the function f is the function cf defined by

$$(cf)(s) = cf(s).$$

1.3 Vector subspaces

Definition 1.3.1

Let V be a vector space. A non empty subset U of V is a subspace if and only if U is closed under the addition and scalar multiplication on V. That is:

(1) For all $u_1 \in U, u_2 \in U$, $u_1 + u_2 \in U$

(2) For any scalar $k \in \mathbb{F}$ and $u \in U$, $ku \in U$.

Proposition 1.3.2

Let V be a vector space over a field $\mathbb{F}$ and let U be a subset of V. Then U is a subspace of V if and only if U is also a vector space over $\mathbb{F}$ under the operations of V.

Example 1.3.3

(1) If V is any vector space, V is a subspace of V; the subset $\{0_V\}$ consisting of the zero vector alone is a subspace of V, called the zero sub?space of V.

(2) In $\mathbb{F}^n$, the set of n-tuples $(x_1, \ldots, x_n)$ with $x_1 = 0$ is a subspace of $\mathbb{F}^n$.

(3) In $\mathbb{F}^n$, the set of n-tuples $(x_1, \ldots, x_n)$ with $x_1 = 1$ is not a subspace of $\mathbb{F}^n$.

(4) The space of polynomial functions over the field $\mathbb{F}$ is a subspace of the space of all functions from $\mathbb{F}$ into $\mathbb{F}$.

Proposition 1.3.4

Let V be a vector space. Then

(1) $0_V \in U$ for every subspace U of V.

(2) The intersection of any collection of subspaces of V is a subspace of V.

Definition 1.3.5

If S_1, $S_2, \ldots, S_k$ are subsets of a vector space V, the set of all sums

$$v_l + v_2 + \cdots + v_k$$

of vectors v_j in S_j is called the sum of the subsets $S_1, S_2, ..., S_k$ and is denoted by or $S_1 + S_2 + \cdots + S_k$ or

$$\sum_{j=1}^{k} S_i.$$

Proposition 1.3.6

If $W_1, W_2, \ldots, W_k$ are subspaces of V, then the sum

$$W = W_1 + W_2 + \cdots + W_k$$

is a subspace of V which contains each of the subspaces W_i.

Definition 1.3.7 Linear combination

Any summand of the form $a_1v_1 + \cdots + a_nv_n$ is called a **linear combination** of $v_1, \ldots, v_n$.

Definition 1.3.8 Span

Let V be a vector space over $\mathbb{F}$ and let $v_1, \ldots, v_n$ be elements of V. Then the subset $\{a_1v_1 + \ldots + a_nv_n \mid a_1, \ldots, a_n \in F\}$ is called the subspace of V **spanned** by $v_1, \ldots, v_n$. It's denoted by $\text{span}\{v_1, \ldots, v_n\}$. If $\text{span}\{v_1, \ldots, v_n\} = V$, we say that $\{v_1, \ldots, v_n\}$ spans V.

Definition 1.3.9 Linearly independent

A set of vectors is said to be linearly dependent over the field F if there are vectors $v_1, \ldots, v_n$ from S and elements $a_1, \ldots, a_n$ from F, not all zero, such that

$$a_1v_1 + \cdots + a_nv_n = 0.$$

A set of vectors that not linearly dependent over F is called **linearly independent**.

Example 1.3.10

The most basic linearly independent set in $\mathbb{F}^n$ is the set of standard unit vectors $e_1 = (1, 0, 0, ..., 0), e_2 = (0, 1, 0, ..., 0), ..., e_n = (0, 0, 0, ..., 1)$.

These vectors span $\mathbb{F}^n$ since every vector $v = (x_1, x_2, \dots, x_n)$ in $\mathbb{F}^n$ can be expressed as $v = x_1e_1 + x_2e_2 + \cdots + x_ne_n$ which is a linear combination of $e_1, e_2, \dots, e_n$.

$$\mathbb{F}^n = \operatorname{span}\{e_1, e_2, \dots, e_n\}.$$

1.4 Basis, dimension and coordinates

Definition 1.4.1 Basis

Let V be a vector space over F. A subset B of V is called a **basis** for V if B is linearly independent over F and every element of V is a linear combination of elements of B.

Proposition 1.4.2

All bases of the same vector space have the same size.

Definition 1.4.3 Dimension

A vector space V that has a basis consisting of n elements is said to have dimension n. We write $\dim V = n$.

For completeness, the trivial vector space $\{0\}$ is said to be spanned by the empty set and to have dimension 0.Every vector space has a basis. A vector space that has a finite basis is called finite dimensional; otherwise, it is called infinite dimensional.

Definition 1.4.4 Coordinate

Let V is a n-dimensional vector space over $\mathbb{F}$ and $B = \{v_1, \dots, v_n\}$ is an ordered basis for V.

Given a vector v in V, there is a unique n-tuple $(\alpha_l, \dots, \alpha_n)$ of scalars in $\mathbb{F}$ such that:

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

The n-tuple is unique, because if v we also have

$$v = \sum_{i=1}^{n} \beta_i v_i$$

We obtain:

$$\sum_{i=1}^{n} (\alpha_i - \beta_i) v_i = 0,$$

and the linear independence of the α_i tells us that $\alpha_i = \beta_i)$ for each i.

The vector $(\alpha_l, \dots, \alpha_n)$ in $\mathbb{F}^n$ is called the coordinate vector of v relative to B; it is denoted by

$$(v)_B = (\alpha_l, \dots, \alpha_n).$$

or

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Theorem 1.4.5 Incomplete basis theorem

Let V be an n-dimensional vector space. Suppose that the family of vectors $\mathcal{S} = \{u_1, u_2, ..., u_r\}$ is linearly independent. Then there exist in V vectors $\{u_{r+1}, u_{r+2}, ..., u_n\}$ such that the family $\{u_1, u_2, ..., u_n\}$ is basis for V.

Proof. Suppose that $\dim V = n$. If $\mathcal{S}$ is a linearly independent set that is not already a basis for V, then $\mathcal{S}$ fails to span V, so there is some vector u_{r+1} in V that is not in $\operatorname{span}(\mathcal{S})$. We can insert u_{r+1} into $\mathcal{S}$, and the resulting set $\mathcal{S}'$ will still be linearly independent. If $\mathcal{S}'$ spans V, then S' is a basis for V, and we are finished. If S' does not span V, then we can insert an appropriate vector u_{r+2} into $\mathcal{S}'$ to produce a set $\mathcal{S}''$ that is still linearly independent. We can continue inserting vectors in this way until we reach a set with n linearly independent vectors in V. This set will be a basis $\mathcal{B} = \{u_1, u_2, ..., u_n\}$ for V. ☐

1.5 Linear Transformations

Definition 1.5.1 Linear Transformation (Linear map)

Let V, W be two vector spaces over the same field $\mathbb{F}$. A function $T\colon V \longrightarrow W$ is called a *linear transformation* from V to W if the following hold for all vectors u, v in V and for all scalars $k \in \mathbb{F}$.

(1) $T(u + v) = T(u) + T(v)$ (additivity)

(2) $T(ku) = kT(u)$ (homogeneity)

Note 1.5.2

We denote the set of all such linear transformations, from V to W, by $\mathcal{L}(V, W)$.

Definition 1.5.3 Linear operator

If V and W are the same, we call a linear transformation from V to V a *linear operator*. We denote the set of all such linear operator on V, by $\mathcal{L}(V)$.

Proposition 1.5.4 Linear transformation

A function $T\colon V \longrightarrow W$ is a linear transformation if and only if for all vectors v_1, v_2 in V and for any scalar k we have

$$T(k\,v_1 + v_2) = k\,T(v_1) + T(v_2)$$

Example 1.5.5 Identity and zero transformations

If V is any vector space, the **identity transformation** I, defined by $I(v) = v$, is a linear operator on V. The **zero transformation** O, defined by $O(v) = 0$ for all $v \in V$, is a linear operator on V.

Proposition 1.5.6

If T is a linear transformation, then

(a) $T(0) = 0$

(b) $T(-v) = -T(v)$

(c) $T(u - v) = T(u) - T(v)$

Definition 1.5.7 Composition (or product) of two linear transformations

Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$ where U is another $\mathbb{F}$-vector space. The **composition** TS is given by

$$(ST)(u) = S\big(T(u)\big) \quad \text{for all} \quad u \in U$$

$$\begin{tikzcd} U \arrow[r, "T"] \arrow[rr, bend left, "ST"] & V \arrow[r, "S"] & W \end{tikzcd}$$

Definition 1.5.8 Invertible linear transformation

Let $T \in \mathcal{L}(V, W)$. We say T is **invertible** provided there exists some $S \in \mathcal{L}(W, V)$ so that $ST : V \longrightarrow V$ is the identity map on V and $TS : W \longrightarrow W$ is the identity map on W. We call S an **inverse** of T.

As a consequence of the next lemma, we are able to refer to *the* inverse of T which we denote by T^{-1}.

Lemma 1.5.9

Let $T \in \mathcal{L}(V, W)$. If T is invertible, then its inverse is unique.

Proof. Assume S and S' are both inverses for T. Then

$$S = SI_W = STS' = I_V S' = S'$$

where I_V and I_W are the identity maps on V and W respectively. ☐

Theorem 1.5.10 Dimension of $\mathcal{L}(V,W)$

Let V be an n-dimensional vector space over the field $\mathbb{F}$, and let W be an m-dimensional vector space over $\mathbb{F}$. Then the space $\mathcal{L}(V,W)$ is finite-dimensional and has dimension mn:

$$\dim \mathcal{L}(V,W) = (\dim V) \times (\dim W).$$

Proof. Let $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ be a basis for V and $\mathcal{B}' = \{w_1, w_2, \dots, w_m\}$ a basis for W.

For all $1 \leq p \leq n$ and $1 \leq q \leq m$. Consider the linear transformation $f_{p,q} \in \mathcal{L}(V,W)$ defined by:

$$f_{p,q}(v_i) = \begin{cases} 0 & \text{if} \quad i \neq p \\ w_q & \text{if} \quad i = p \end{cases}$$

That means:

$$f_{p,q}(v_i) = \delta_{ip} w_q,$$

where

$$\delta_{ip} = \begin{cases} 0 & \text{if} \quad i \neq p \\ 1 & \text{if} \quad i = p \end{cases}$$

The claim is that the mn transformations $f_{p,q}$ form a basis for $\mathcal{L}(V,W)$.

Let T be a linear transformation from V into W, and $a_{1j}, ..., a_{mj}$ the coordinates of the vector $T(v_j)$ in the ordered basis $\mathcal{B}'$.
That means:

$$T(v_j) = \sum_{q=1}^{m} a_{qj} w_q$$

Let

$$U = \sum_{q=1}^{m} \sum_{p=1}^{n} a_{qp} f_{p,q}$$

We wish to show that: $T = U$.
For all $j = 1, ..., n$, we have:

$$\begin{aligned} U(v_j) &= \sum_{q=1}^{m} \sum_{p=1}^{n} a_{pq} f_{p,q}(v_j) \\ &= \sum_{q=1}^{m} \sum_{p=1}^{n} a_{qp} \delta_{jp} w_q \\ &= \sum_{q=1}^{m} a_{qj} w_q \\ &= T(v_j) \end{aligned}$$

Then $T = U$. This shows that the linear transformations $f_{p,q}$ span $\mathcal{L}(V,W)$.

We must prove that they are independent. Assume that:

$$\sum_{q=1}^{m}\sum_{p=1}^{n} k_{qp} f_{p,q} = 0$$

Then for all $j = 1, ..., n$

$$\sum_{q=1}^{m} k_{qj} w_q = 0$$

and the independence of the basis B' implies that $k_{qj} = 0$ for all $j = 1, ..., n$ and $q = 1, ..., m$.

Hence the set

$$\{f_{p,q}\}_{\substack{1\le p\le n\\ 1\le q\le m}}$$

form a basis for $\mathcal{L}(V, W)$.

Finally

$$\dim \mathcal{L}(V, W) = nm = (\dim V) \times (\dim W).$$

□

1.6 Kernel and range of a transformation

Definition 1.6.1 Kernel and range of a linear transformation

Let $T \colon V \longrightarrow W$ is a linear transformation.

- The set $\ker T$ of all vectors v in V for which $T(v) = 0$ is called the **kernel (or Null Space)** of T.

$$\ker T = \{v \in V \mid T(v) = 0\}$$

- The set $R(T)$ of all outputs (images) $T(v)$ of vectors in V via the transformation T is called the **range** of T.

$$\operatorname{rang} T = \{T(v) \mid v \in V\}$$

Clearly $\ker T$ is a vector subspace of V and $\operatorname{rang} T$ is a vector subspace of W.

Definition 1.6.2 Nullity and rank

If V and W are *finite* dimensional vector spaces and $T \colon V \longrightarrow W$ is a linear transformation, then we call

- $\dim \ker T =$ nullity of T
- $\dim \operatorname{rang} T =$ rank of T

Theorem 1.6.3

If V and W are finite-dimensional vector spaces and $T \colon V \longrightarrow W$ is a linear transformation, then

$$\text{rank } (T) + \text{ nullity } (T) = \dim(V)$$

Definition 1.6.4 One-to-one, onto, bijective

- A function $f\colon X \longrightarrow Y$ is called *one-to-one* (or injective) if $f(x) = f(x')$ imply $x = x'$.
- A function $f\colon X \longrightarrow Y$ is called *onto* (or surjective) if for every y in Y there is at least one x in X such that $f(x) = y$.
- A linear transformation that is both injective and surjective is called *isomorphism* (or bijective).

Proposition 1.6.5

- A linear transformation $T\colon V \longrightarrow W$ is one-to-one if and only if $\ker(T) = \{0\}$.
- A linear transformation $T\colon V \longrightarrow W$ is onto if and only if $\operatorname{rang} T = W$.

Definition 1.6.6

We say two vector spaces V and W are **isomorphic** and write $V \cong W$, if there exists $T \in \mathcal{L}(V, W)$ which is both injective and surjective. We call such a T an **isomorphism**.

Theorem 1.6.7

Two finite-dimension vector spaces V and W are isomorphic if and only if they have the same dimension.

Proof. Assume V and W are isomorphic. This means there exists a linear map $T : V \longrightarrow W$ that is both surjective and injective. Theorem 1.6.3 immediately implies that $\dim V = \dim W$. For the reverse direction, let $\mathcal{B}_V = \{v_1, \ldots, v_n\}$ be a basis for V and $\mathcal{B}_W = \{w_1, \ldots, w_n\}$ be a basis for W. As every vector $v \in V$ can be written (uniquely) as

$$v = a_1v_1 + \cdots + a_nv_n$$

for $a_i \in \mathbb{F}$, we may define a function $T : V \longrightarrow W$ by

$$Tv = a_1w_1 + \cdots + a_nw_n.$$

Observe that the uniqueness of our representation of v implies that T is a well-defined function. Moreover, a straightforward check reveals that T is indeed a linear map. It only remains to show that T is an isomorphism. To see that T is injective, let that $v \in nulT$ and let $b_i \in \mathbb{F}$ be such that

$$v = b_1v_1 + \cdots + b_nv_n.$$

This means

$$0_W = Tv = b_1w_1 + \cdots + b_nw_n.$$

Since $\mathcal{B}_W$ is an independent set, it follows that all our scalars b_i must be 0 and, in turn, $v = 0$. This shows that $\ker T = \{0_V\}$, i.e., T is injective.

To see that T is also surjective, note that any vector $w \in W$ can be written as

$$w = c_1w_1 + \cdots + c_mw_m,$$

for some choice of scalars c_i (why?). Now consider the vector $c_1v_1 + \cdots + c_mv_m \in V$ and observe that

$$T(c_1v_1 + \cdots + c_mv_m) = c_1w_1 + \cdots + c_mw_m = w.$$

This shows that T is surjective. ☐

1.7 Direct-Sum

Definition 1.7.1

Let $U_1, \dots, U_n$ be subspaces of V. and W a subspace of V. We say that W is sum of the U_i's, and write

$$W = U_1 + \dots + U_n$$

provided that for every $w \in W$ there exist $u_i \in U_i$, $1 \leq i \leq n$, with

$$w = \sum_{i=1}^{n} u_i.$$

Example 1.7.2

Consider the subspaces of $\mathbb{R}^3$:

$$U_1 = \{(x, 0, z) \mid x, z \in \mathbb{R}\} \quad \text{and} \quad U_1 = \{(0, y, z) \mid y, z \in \mathbb{R}\}.$$

Remark that every vector $v = (x, y, z) \in \mathbb{R}^3$, can be written as sum of a vector in U_1 and a vector in U_2, for example:

$$v = (x, y, z) = (x, 0, z) + (0, y, 0) \quad \text{or} \quad v = (x, y, z) = (x, 0, 0) + (0, y, z).$$

Therefore $\mathbb{R}^3 = U_1 + U_2$

Definition 1.7.3

Let $U_1, \dots, U_n$ be subspaces of V and W a subspace of V. We say that W is direct sum of the U_i's, and write

$$W = \bigoplus_{i=1}^{n} U_i,$$

provided that for every $w \in W$ there exist **unique** $u_i \in U_i$, $1 \leq i \leq n$, with

$$w = \sum_{i=1}^{n} u_i.$$

Example 1.7.4

Let $V = \mathbb{R}^2$, $U_1 = \{(x, x) \mid x \in \mathbb{R}\}$ and $U_2 = \{(y, -y) \mid y \in \mathbb{R}\}$.

(1) Show that $\mathbb{R}^2 = U_1 + U_2$.

(2) Is $\mathbb{R}^2 = U_1 \oplus U_2$?

solution

(1) Let $(a, b) \in \mathbb{R}^2$, we will find $(x, x) \in U_1$ and $(y, -y) \in U_2$ such that

$$(a, b) = (x, x) + (y, -y) \tag{1.1}$$

That is

$$a = x + y \text{ and } b = x - y$$

Adding and subtracting the two equation we obtain

$$2x = a + b \quad \text{and} \quad 2y = a - b$$

Then we can divide by 2 to obtain the solution $x = \frac{a+b}{2}$ and $y = \frac{a-b}{2}$. So for all $(a, b) \in \mathbb{R}^2$:

$$(a, b) = \left(\frac{a+b}{2}, \frac{a+b}{2}\right) + \left(\frac{a-b}{2}, \frac{b-a}{2}\right).$$

Hence $\mathbb{R}^2 = U_1 + U_2$.

(2) As the equation (1.1) has a unique solution, $\mathbb{R}^2 = U_1 \oplus U_2$.

1.8 A formal definition of the determinant of a matrix

Definition 1.8.1 Permutation

A permutation of the set $\{1, \dots n\}$ is any ordered way to write down the symbols $\{1, \dots n\}$. We denote the set of all this permutations by $\mathcal{S}_n$.

Example 1.8.2

The collection of all permutations of the string $(1, 2, 3)$ is the set

$$\mathcal{S}_3 = (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1).$$

Given a permutation π, we refer to the k-th entry of π by writing $\pi(k)$. For example, if $\pi = (2, 3, 4, 1)$, we would interpret $\pi(2)$ to be the second entry of π, which is 3.

Note 1.8.3

Take any permutation. We claim that it can be created by the following process:

(1) Start with the permutation $(1, 2, 3, \dots n)$.

(2) Repeatedly pick pairs of elements in the permutation we have, and swap them.

(3) By carefully choosing the pairs in step 2 above, we can get to any other permutation.

The signature of the permutation $\operatorname{sgn}(\sigma)$ is defined as follows:

$$\operatorname{sgn}(\sigma) = \begin{cases} 1 & \text{If the total number of swaps is even} \\ -1 & \text{If the total number of swaps is odd.} \end{cases}$$

Example 1.8.4

The permutation $(2,3,4,1)$ has signature

$$\operatorname{sgn}(2,3,4,1)=-1.$$

Remark that

$$(1,2,3,4)\xrightarrow{\text{switch 1,2}}(2,1,3,4)\xrightarrow{\text{switch 1,3}}(2,3,1,4)\xrightarrow{\text{switch 1,4}}(2,3,4,1).$$

Definition 1.8.5

Let A be a $n\times n$ matrix, of the form

$$\begin{pmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{pmatrix}$$

$$\det(A)=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{1,\sigma(1)}\cdot a_{2,\sigma(2)}\cdot\dots\cdot a_{n,\sigma(n)}.$$

Example 1.8.6

If A is a square matrix of order 3×3 , then

$$\begin{aligned}\det A =&\operatorname{sgn}(1,2,3)\cdot a_{11}a_{22}a_{33}+\operatorname{sgn}(1,3,2)\cdot a_{11}a_{23}a_{32}+\operatorname{sgn}((2,1,3))\cdot a_{12}a_{21}a_{33}\\ &+\operatorname{sgn}(2,3,1)\cdot a_{12}a_{23}a_{31}+\operatorname{sgn}(3,1,2)\cdot a_{13}a_{21}a_{32}+\operatorname{sgn}(3,2,1)\cdot a_{13}a_{22}a_{31},\end{aligned}$$

which if you calculate the signatures is just

$$\begin{aligned}\det A =&1\cdot a_{11}a_{22}a_{33}+(-1)\cdot a_{11}a_{23}a_{32}+(-1)\cdot a_{12}a_{21}a_{33}\\ &+1\cdot a_{12}a_{23}a_{31}+1\cdot a_{13}a_{21}a_{32}+(-1)\cdot a_{13}a_{22}a_{31}.\end{aligned}$$

Hence

$$\det A=a_{11}a_{22}a_{33}-\cdot a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+\cdot a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}.$$

1.9 Matrix of a linear transformation

Consider the following data:

- An n-dimensional vector space V over $\mathbb{F}$ with a basis $\mathcal{B}=\{u_1,u_2,\dots,u_n\}$.
- An m-dimensional vector space W over $\mathbb{F}$ with a basis $\mathcal{B}'=\{v_1,v_2,\dots,v_m\}$.
- A linear transformation $T\colon V\longrightarrow W$.

Definition 1.9.1

The matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{B}'$ is the $m \times n$ matrix $[T]_{\mathcal{B}',\mathcal{B}}$ defined by

$$[T]_{\mathcal{B}',\mathcal{B}} = \big[\, [T(u_1)]_{\mathcal{B}'} \,|\, [T(u_2)]_{\mathcal{B}'} \,|\, \dots \,|\, [T(u_n)]_{\mathcal{B}'} \big]$$

Relative to these bases.
More precisely, we have the following relation:

$$\Big[T(v)\Big]_{\mathcal{B}'} = [T]_{\mathcal{B}',\mathcal{B}} \cdot [v]_{\mathcal{B}}$$

Theorem 1.9.2

Let A be a square matrix and and let $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be the matrix transformation $T_A(x) = Ax$. Then the following statements are equivalent:

(1) A is invertible.

(2) The columns of A

(3) $Ax = b$ has a unique solution for each b in $\mathbb{R}^n$.

(4) $Ax = 0$ has a unique solution $x = 0$.

(5) T_A is invertible.

(6) T_A is one-to-one.

(7) T_A is onto

Example 1.9.3

Let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be defined by $T(x, y) = (2x - 3y, x + 2y)$.
Compute the matrix A of T relative to standard basis $\mathbf{S} = \{e_1, e_2\}$ of $\mathbb{R}^2$.
Solution We have

$$T(\mathbf{e_1}) = T(1,0) = (2,1)$$

and

$$T(\mathbf{e_2}) = T(0,1) = (-3,2),$$

so the standard matrix for T is

$$[T]_{\mathbf{S}} = [T(e_1)|T(e_2)] = \begin{pmatrix} 2 & -3 \\ 1 & 2 \end{pmatrix}$$

Proposition 1.9.4 Linear isomorphisms on finite-dimensional dimension vector spaces

Let V and W be two finite-dimensional vector spaces over a field $\mathbb{F}$ of the same dimension. If $T \colon V \longrightarrow W$ is a linear trnasformation and if $\mathcal{B}$ is (resp $\mathcal{B}'$)a basis for V (resp. W), then the following are equivalent:

(a) T is one-to-one.

(b) T is onto.

(c) T is bijective.

(d) $[T]_{\mathcal{B}',\mathcal{B}}$ is invertible,

(e) $\det[T]_{\mathcal{B}',\mathcal{B}} \neq 0$.

Moreover, if these conditions hold, then

$$[T^{-1}]_{\mathcal{B},\mathcal{B}'} = [T]^{-1}_{\mathcal{B}',\mathcal{B}}$$

1.10 Transition matrix

Theorem 1.10.1 Change of coordinates formula

Let $\mathcal{B} = \{v_1, \dots, v_n\}$ and $B' = \{v'_1, \dots, v'_n\}$ be two ordered bases of V. Then there is a unique, necessarily invertible, $n \times n$ matrix P with entries in $\mathbb{F}$ such that for all vector $v \in V$:

(i) $[v]_{\mathcal{B}} = P_{\mathcal{B}' \longrightarrow \mathcal{B}}[v]_{\mathcal{B}'}$

(ii) $[v]_{\mathcal{B}'} = P_{\mathcal{B} \longrightarrow \mathcal{B}'}[v]_{\mathcal{B}}$

The columns of $P_{\mathcal{B}' \longrightarrow \mathcal{B}}$ are given by $[v'_j]_{\mathcal{B}}$.

The matrix

$$P_{\mathcal{B}' \longrightarrow \mathcal{B}} = \big[\, [v'_1]_{\mathcal{B}} \quad | \quad [v'_2]_{\mathcal{B}} \quad | \cdots | \quad [v'_n] \,\big].$$

is called the transition matrix from B' to B.

Remark 1.10.2. Remark that : $(P_{\mathcal{B} \longrightarrow \mathcal{B}'}) \times (P_{\mathcal{B}' \longrightarrow \mathcal{B}}) = I_n$.

Example 1.10.3

Consider the bases $\mathcal{B} = \{u_1, u_2\}$ and $\mathcal{B}' = \{u'_1, u'_2\}$ for $\mathbb{R}^2$, where $u_1 = (1,0), u_2 = (0,1), u'_1 = (1,1), u'_2 = (2,1)$

(a) Find the transition matrix $P_{\mathcal{B}' \longrightarrow \mathcal{B}}$ from $\mathcal{B}'$ to $\mathcal{B}$.

(b) Find the transition matrix $P_{\mathcal{B} \longrightarrow \mathcal{B}'}$ from $\mathcal{B}$ to $\mathcal{B}'$.

(c) Let v be a vector in $\mathbb{R}^2$ such that $[v]_{\mathcal{B}'} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Find $[v]_{\mathcal{B}}$.

Solution.

$$P_{\mathcal{B}' \longrightarrow \mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad P_{\mathcal{B} \longrightarrow \mathcal{B}'} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

$$[v]_{\mathcal{B}} = (P_{\mathcal{B}' \longrightarrow \mathcal{B}})[v]_{\mathcal{B}'} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

1.11 Exercises set

Exercise 1.11.1

Determine whether the vectors

$$v_1 = (1,2,2,-1), v_2 = (4,9,9,-4) \quad \text{and} \quad v_3 = (5,8,9,-5)$$

are linearly dependent or linearly independent in $\mathbb{R}^4$.

Solution. The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1v_1 + k_2v_2 + k_3v_3 = 0.$$

Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{cases} k_1 + 4k_2 + 5k_3 = 0 \\ 2k_1 + 9k_2 + 8k_3 = 0 \\ 2k_1 + 9k_2 + 9k_3 = 0 \\ -k_1 - 4k_2 - 5k_3 = 0 \end{cases}$$

This system has only the trivial solution $k_1 = 0$, $k_2 = 0$, $k_3 = 0$. We conclude that v_1,v_2, and v_3 are linearly independent.

Exercise 1.11.2

Determine whether the vectors $v_1 = (1,-2,3), v_2 = (5,6,-1) \quad \text{and} \quad v_3 = (3,2,1)$ are linearly independent or linearly dependent in $\mathbb{R}^3$.

Solution. The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1v_1 + k_2v_2 + k_3v_3 = 0$$

Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{cases} k_1 + 5k_2 + 3k_3 = 0 \\ -2k_1 + 6k_2 + 2k_3 = 0 \\ 3k_1 - k2 + k_3 = 0 \end{cases}$$

Thus, our problem reduces to determining whether this system has nontrivial solutions. There are various ways to do this; one possibility is to simply solve the system, which yields $k_1 = \frac{1}{2}t, k_2 = \frac{1}{2}t, k_3 = t$.

This shows that the system has nontrivial solutions and hence that the vectors are linearly dependent.

Exercise 1.11.3

Let V be a vector space of dimension n over a field $\mathbb{F}$, and $\mathcal{B} = \{v_1, \dots, v_n\}$ a basis of V. Show that the map $\psi : V \longrightarrow \mathbb{F}^n$ defined by $\psi(v) = [v]_{\mathcal{B}}$ is an isomorphism

Solution. First we show that ψ is linear. Let $\lambda \in \mathbb{F}$ and u, w two vectors in V. As $\mathcal{B}$ is a basis for V, the vectors u and v can be written uniquely as

$$u = \sum_{i=1}^{n} \alpha_i v_i \quad \text{and} \quad w = \sum_{i=1}^{n} \beta_i v_i$$

Then

$$\lambda u + w = \sum_{i=1}^{n} (\lambda \alpha_i + \beta_i) v_i$$

Hence

$$\begin{aligned} \psi(\lambda u + w) &= [\lambda u + w]_{\mathcal{B}} \\ &= \begin{pmatrix} \lambda\alpha_1 + \beta_1 \\ \lambda\alpha_2 + \beta_2 \\ \vdots \\ \lambda\alpha_n + \beta_n \end{pmatrix} \\ &= \lambda \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= \lambda [u]_{\mathcal{B}} + [w]_{\mathcal{B}} \\ &= \lambda \psi(u) + \psi(w). \end{aligned}$$

Since V and $\mathbb{F}^n$ has the same dimension ($\dim V = \dim \mathbb{F}^n = n$), to prove that ψ is bijective, it suffices to prove for example that is injective.

Let $u = \sum_{i=1}^{n} \alpha_i v_i \in V$. We have:

$$\begin{aligned} \psi(u) = 0_{\mathbb{F}^n} &\iff \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\ &\iff \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0 \\ &\iff u = 0. \end{aligned}$$

So $\operatorname{Ker}(\psi) = \{0\}$ and hence ψ is injective. Therefore it is an isomorphism of vector spaces.

Exercise 1.11.4

Let $f \in \mathcal{L}(U, V)$ and $g \in \mathcal{L}(V, W)$ where U, V, W are $\mathbb{F}$-vector spaces. Show that $gf \in \mathcal{L}(U, W)$.

Solution. Let $\alpha \in \mathbb{F}$ and $u_1, u_2 \in U$. We have:

$$\begin{aligned} (gf)(\lambda u_1 + u_2) &= g\big(f(\lambda u_1 + u_2)\big) \\ &= g\big(\lambda f(u_1) + f(u_2)\big) \\ &= \lambda g\big(f(u_1)\big) + g\big(f(u_2)\big) \\ &= \lambda (gf)(u_1) + (gf)(u_2). \end{aligned}$$

Exercise 1.11.5

Consider the bases $\mathcal{B} = \{u_1, u_2\}$ and $\mathcal{B}' = \{u'_1, u'_2\}$ for $\mathbb{R}^2$, where $u_1 = (1,0), u_2 = (0,1), u'_1 = (1,1), u'_2 = (2,1)$

(a) Find the transition matrix $P_{\mathcal{B}' \longrightarrow \mathcal{B}}$ from $\mathcal{B}'$ to $\mathcal{B}$.

(b) Find the transition matrix $P_{\mathcal{B} \longrightarrow \mathcal{B}'}$ from $\mathcal{B}$ to $\mathcal{B}'$.

(c) Let v be a vector in $\mathbb{R}^2$ such that $[v]_{\mathcal{B}'} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Find $[v]_{\mathcal{B}}$.

Solution.

$$P_{\mathcal{B}' \longrightarrow \mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad P_{\mathcal{B} \longrightarrow \mathcal{B}'} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

$$[v]_{\mathcal{B}} = (P_{\mathcal{B}' \longrightarrow \mathcal{B}})[v]_{\mathcal{B}'} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$$

Exercise 1.11.6

Let $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear operator given by:

$$T(x, y, z) = (2x + z, -2x + y, -x + 2y + z).$$

(1) What is the matrix of T with respect to the standard basis $\mathcal{S}$ of $\mathbb{R}^3$?

(2) What is the matrix of T with respect to the ordered basis $\mathcal{B} = \{v_1, v_2, v_3\}$, where

$$v_1 = (1, 0, 1)\ ,\ v_2 = (1, 1, 0)\ ,\ v_3 = (0, 1, 1)\ .$$

(3) Find $[T]_{\mathcal{B},\mathcal{S}}$ the matrix for T relative to the bases $\mathcal{S}$ and $\mathcal{B}$.

(4) Find $[T]_{\mathcal{S},\mathcal{B}}$ the matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{S}$.

Exercise 1.11.7

Show that the following maps ∂, T and L are linear:

(1) Let $\mathcal{D}$ be the vector space of all differentiable function $f : \mathbb{R} \longrightarrow \mathbb{R}$ and $\mathcal{F}$ the space of all function $g : \mathbb{R} \longrightarrow \mathbb{R}$. Define the map $\partial : \mathcal{D} \longrightarrow \mathcal{F}$, by $\partial f = f'$.

(2) Let $\mathcal{C}$ be the space of continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$. Define $T : \mathcal{C} \longrightarrow \mathcal{C}$ by $T(f) = xf(x)$.

(3) The map $L : \mathcal{C} \longrightarrow \mathcal{R}$ given by

$$L(f) = \int_0^1 f\ dx.$$

Solution. (1) Clearly ∂ is a linear map since

$$\partial(f + g) = (f + g)' = f' + g' = \partial f + \partial g$$

and

$$\partial(af) = (af)' = af' = a\partial f.$$

In particular the map D form $\mathbb{F}[X]$ into $\mathbb{F}[X]$ defined by

$$D(a_0 + a_1X + \cdots + a_nX^n) = a_1 + 2a_2X + \cdots + na_nX^{n-1}.$$

Is a linear operator.

(2) Let $\mathcal{C}$ be the space of continuous functions $f : \mathbb{R} \longrightarrow \mathbb{R}$. An example of a linear map on this space is the function $T : \mathcal{C} \longrightarrow \mathcal{C}$ given by $T(f) = xf(x)$.

(3) The map $L : \mathcal{C} \longrightarrow \mathcal{R}$ given by

$$L(f) = \int_0^1 f\ dx.$$

is linear.

Chapter 2

Inner product spaces

Chapter contents

Throughout this chapter we consider only real or complex vector spaces, that is, vector spaces over the field of real numbers or the field of complex numbers.

$$\mathbb{F} = \mathbb{R} \quad \text{or} \quad \mathbb{F} = \mathbb{C}$$

2.1 Inner Products

Definition 2.1.1 Inner Products

Let V be a vector space over $\mathbb{F}$. An **inner product** is a function $\langle -, - \rangle \colon V \times V \longrightarrow \mathbb{F}$ such that for all vectors v, u, w in V and scalars a, b in $\mathbb{F}$:

(1) $\langle v, v \rangle \geq 0$ with equality iff $v = \mathbf{0}_V$.

(2) $\langle v, u \rangle = \overline{\langle u, v \rangle}$, where the bar denoting complex conjugation; **Conjugate symmetric**

(3) $\langle av + bu, w \rangle = a\langle v, w \rangle + b\langle u, w \rangle$. **Linearity in the first component**

- Notice that conjugate symmetry implies that $\langle u, u\rangle \in \mathbb{R}$ even if $\mathbb{F} = \mathbb{C}$ since

$$\langle u, u\rangle = \overline{\langle u, u\rangle}.$$

Example 2.1.2

(1) $\mathbb{R}^n$ with the dot product:

$$\langle (a_1, \dots, a_n), (b_1, \dots, b_n) \rangle = a_1 b_1 + \cdots + a_n b_n.$$

(2) $\mathbb{C}^n$ with the standard inner product:

$$\langle (a_1, \dots, a_n), (b_1, \dots, b_n) \rangle = a_1 \bar{b}_1 + \cdots + a_n \bar{b}_n.$$

(3) If W is a subspace of an inner product space V, then the inner product of V restricted to W gives an inner product on W.

Example 2.1.3

(1) $V = (\mathcal{C}[0,1], \mathbb{C})$, the set of continuous complex valued functions on $[0,1]$ with inner product

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}dx,$$

(2) $V = \mathbb{F}^{n\times n}$, the space of all $n \times n$ matrices over $\mathbb{F}$ with inner product

$$\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij}.$$

Definition 2.1.4 Inner product space

An inner product space is a real or complex vector space, together with a specified inner product on that space.

- A finite-dimensional real inner product space is often called a **Euclidean space**.
- A complet inner product space is often referred to as a **unitary space**.

Definition 2.1.5 Norm of a vector

Let V be an inner product space. For all vector v, we define the norm of v by

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

Theorem 2.1.6

If V is an inner product space, then for any vectors v, u in V and any scalar $a \in \mathbb{F}$, we have

(a) $\|u\| \geq 0$

(b) $\|au\| = |a|\,\|u\|$

(c) $\|u\| = 0 \Leftrightarrow u = 0$

(d) (Cauchy-Schwarz inequality):
$$|\langle u, v\rangle| \leq \|u\|\,\|v\|\,.$$

(e) (Triangle inequality):
$$\|u + v\| \leq \|u\| + \|v\|$$

Proof. Statements **(a)**, **(b)** and **(c)** follow immediately from the definition. Let u and v be two vectors in V, and $c \in \mathbb{F}$:

(d) Consider $u - cv$ and notice that

$$\begin{aligned} 0 &\leq \|u - cv\|^2 \\ &= \langle u - cv, u - cv\rangle \\ &= \|u\|^2 - \langle cv, u\rangle - \langle u, cv\rangle + \|cv\|^2 \\ &= \|u\|^2 - 2\,\mathrm{Re}\,\bar{c}\,\langle u, v\rangle + |c|^2\,\|v\|^2\,. \end{aligned}$$

Notice that if we take $c = \frac{\langle u,v\rangle}{\|v\|^2}$ then

$$0 \leq \|u\|^2 - 2\frac{|\langle u, v\rangle|^2}{\|v\|^2} + \frac{|\langle u, v\rangle|^2}{\|v\|^2} = \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2},$$

Therefore
$$|\langle u, v\rangle|^2 \leq \|u\|^2\|v\|^2$$

Hence
$$|\langle u, v\rangle| \leq \|u\|\,\|v\|\,.$$

(e)

$$\begin{aligned} \|u + v\|^2 &= \langle\, u + v, u + v\,\rangle \\ &= \langle\, u, u\,\rangle + \langle\, u, v\,\rangle + \langle\, v, u\,\rangle + \langle\, v, v\,\rangle \\ &= \|u\|^2 + \langle\, u, v\,\rangle + \langle\, v, u\,\rangle + \|v\|^2 \\ &= \|u\|^2 + \langle\, u, v\,\rangle + \overline{\langle\, u, v\,\rangle} + \|v\|^2 \\ &= \|u\|^2 + 2\,\mathrm{Re}\,\langle\, u, v\,\rangle + \|v\|^2 \end{aligned}$$

Remark that $a \leq \sqrt{a^2 + b^2} = |a + bi|$ and so $\mathrm{Re}\,\langle u, v\rangle \leq |\langle u, v\rangle| \leq \|u\|\,\|v\|\,.$

Therefore

$$\|u + v\|^2 \leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2$$

So

$$\|u+v\|^2 \leq (\|u\|+\|v\|)^2$$

Hence

$$\|u+v\| \leq \|u\|+\|v\|$$

□

Apply the Cauchy-Schwarz inequality to the inner products given in Example 2.1.2 **(2)** and Example 2.1.3 **(1)**, we get:

$$\sum_{i=1}^{n} x_i\bar{y}_i \leq \left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}|y_i|^2\right)^{\frac{1}{2}},$$

and

$$\left|\int_0^1 f(x)\overline{g(x)}dx\right| \leq \left(\int_0^1 |f(x)|^2dx\right)^{\frac{1}{2}}\left(\int_0^1 |f(x)|^2dx\right)^{\frac{1}{2}}$$

Definition 2.1.7

Let V be an inner product space.

- Vectors u and v in V are **orthogonal** $(u \perp v)$ if $\langle u, v\rangle = 0$.
- A subset $S \subseteq V$ is **orthogonal** if any two distinct vectors in S are orthogonal.
- A vector u in V is a **unit vector** if $\|u\| = 1$.
- A subset $S \subseteq V$ is **orthonormal** if S is orthogonal and consists entirely of unit vectors.

Note 2.1.8

Note that :

- $S = \{u_1, \ldots, u_k\}$ is orthonormal iff $\langle u_i, u_j\rangle = \delta_{ij}$.
- We can make an orthonormal set from an orthogonal set by replacing each vector u by $\frac{1}{\|u\|}u$. This will not change the orthogonality since $\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle = \frac{1}{\|x\|\|y\|}\langle x, y\rangle$ since $\|y\| \in \mathbb{R}$. We call this process **normalizing** the set.

Proposition 2.1.9

If V is an inner product space and $S \subseteq V$ is orthogonal subset of nonzero vectors, then S is linearly independent.

Proof. We first note that if S is not the set consisting only of zero, then zero cannot be in S. Suppose that

$$S = \{u_1, ..., u_k\}$$

and

$$a_1u_1 + \cdots + a_ku_k = \mathbf{0}_V$$

for scalars $a_1, \ldots, a_k$ and vectors $u_1, \ldots, u_k$ in S.

Then we see that

$$0 = \langle a_1 u_1 + \cdots + a_k u_k, u_i \rangle = a_i \left\| u_i \right\|^2$$

and since

$$\left\| u_i \right\|^2 \neq 0,$$

we must have $a_i = 0$. This can be done for all i. □

2.2 Orthonormal bases

Definition 2.2.1

Let V be an inner product space. A subset of V is an **orthonormal basis** for V if it is an ordered basis that is orthonormal.

Theorem 2.2.2

Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of nonzero vectors. If $w \in \operatorname{Span} S$, then

$$w = \sum_{i=1}^{k} \frac{\langle w, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

In addition, if S is orthonormal, then the denominators are all 1. That means:

$$w = \sum_{i=1}^{k} \langle w, v_i \rangle v_i.$$

Proof. Since $w \in \operatorname{Span} S$, we must have that there exist scalars $a_1, \ldots, a_k$ such that

$$w = \sum_{i=1}^{k} a_i v_i.$$

We can now take the inner product with v_j for $j = 1, \ldots, k$ and find that

$$\begin{aligned} \langle w, v_j \rangle &= \left\langle \sum_{i=1}^{k} a_i v_i, v_j \right\rangle \\ &= \sum_{i=1}^{k} a_i \langle v_i, v_j \rangle \\ &= a_j \langle v_j, v_j \rangle \end{aligned}$$

and so (since $\|v_j\| \neq 0$), $a_j = \dfrac{\langle w, v_j \rangle}{\langle v_j, v_j \rangle}$.

$$w = \sum_{i=1}^{k} \frac{\langle w, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

□

Corollary 2.2.3

Let $v_1, \dots, v_n$ be an orthonormal basis of an inner product space V and $v, w \in V$. Then:

- **Parseval's identity**:

$$\langle v, w \rangle = \sum_{i=1}^{n} \langle v, v_i \rangle \langle v_i, w \rangle.$$

- **Bessel's equality**:

$$\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i \rangle|^2.$$

Theorem 2.2.4

Let W be a finite dimensional subspace of the inner product space V. Then for a vector $y \in V$, there is a unique vector $u \in W$ that minimizes $\|y-w\|^2$ for all $w \in W$.

Proof. Suppose there is a $u \in W$ such that $\langle w, y-u \rangle = 0$ for any $w \in W$. Then if $w \in W$ (and hence so is $u-w$),

$$\begin{aligned}
\|y-w\|^2 &= \|u+(y-u)-w\|^2 \\
&= \langle u-w+(y-u), u-w+(y-u) \rangle \\
&= \|u-w\|^2 + \langle u-w, y-u \rangle + \langle y-u, u-w \rangle + \|y-u\|^2 \\
&= \|u-w\|^2 + \|y-u\|^2 \\
&\geq \|y-u\|^2 .
\end{aligned}$$

We can do this if W is finite dimensional using the following theorem. ☐

Definition 2.2.5

The orthogonal complement of W, written $W^{\perp}$ (pronounced "W perp"), is the set of all vectors $v \in V$ such that $\langle v, w \rangle = 0$ for all $w \in W$.

Proposition 2.2.6

$W^{\perp}$ is a vector space.

Proof. It is straightforward to see that $\langle \mathbf{0}_V, w \rangle = 0$ for all $w \in W$, so $\mathbf{0}_V \in W^{\perp}$.

Let $v, u \in W^{\perp}$ and $c \in \mathbb{F}$.

Then

$$\langle cv+u, w \rangle = c \langle v, w \rangle = \langle u, w \rangle = 0$$

so

$$cv+u \in W^{\perp}.$$

☐

Theorem 2.2.7 Gram-Schmidt process

Let V be an inner product space and $S = \{w_1, \ldots, w_n\}$ be a linearly independent subset of V. Define $S' = \{v_1, \ldots, v_n\}$ by $v_1 = w_1$ and

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j$$

for $k = 2, \ldots, n$. Then S' is an orthogonal set of nonzero vectors such that $\operatorname{Span} S' = \operatorname{Span} S$.

Proof. We show inductively that v_{k+1} is orthogonal to $v_1, \ldots, v_k$. It is clear that

$$\langle v_2, v_1\rangle = \left\langle w_2 - \frac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1, v_1 \right\rangle = \langle w_2, v_1\rangle - \frac{\langle w_2, v_1\rangle}{\langle v_1, v_1\rangle} \langle v_1, v_1\rangle = 0.$$

We then can use the inductive hypothesis to assume $\langle v_i, v_j\rangle = 0$ for $i, j \leq k$ and see that

$$\langle v_k, v_i\rangle = \left\langle w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j\rangle}{\langle v_j, v_j\rangle} v_j, v_i \right\rangle = \langle w_k, v_i\rangle - \frac{\langle w_k, v_i\rangle}{\langle v_i, v_i\rangle} \langle v_i, v_i\rangle = 0.$$

Thus S' is orthogonal. Hence S' is linearly independent and since each element of S' is in the span of S, $\operatorname{Span} S' \subseteq \operatorname{Span} S$, and hence $\operatorname{Span} S' = \operatorname{Span} S$ (since they have the same dimension). ☐

Theorem 2.2.8

Suppose that $S = \{v_1, \ldots, v_k\}$ is an orthonormal set in a n-dimensional inner product space V. Then

(1) S can be extended to an orthonormal basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V.

(2) If $W = \operatorname{Span} S$, then $S_1 = \{v_{k+1}, \ldots, v_n\}$ is an orthonormal basis for $W^\perp$.

(3) If W is any subspace of V, then $\dim V = \dim W + \dim W^\perp$.

Proof. By the replacement theorem, S can be extended into a basis, and then the Gram-Schmidt process can be used to turn this into an orthogonal set. Then normalizing gives an orthonormal set. S_1 is clearly a linearly independent subset of $W^\perp$. Since $\{v_1, \ldots, v_n\}$ is a basis, any vector in $W^\perp$ can be written as a linear combination of these vectors. However, since $w \in W^\perp$ satisfies $\langle w, v_i\rangle = 0$ for $i = 1, \ldots, k$, w is in the span of S_1, hence S_1 is a basis. The dimension statement is clear now that we know that S is a basis for S, S' is a basis for $W^\perp$, and $\{v_1, \ldots, v_n\}$ is a basis for V. ☐

Proposition 2.2.9 Polarization Identities for real inner product spaces

Let V be a real inner product space and v, w two vectors in V. We have:

$$\langle\, v, w \,\rangle = \frac{1}{4} \|v + w\|^2 - \frac{1}{4} \|v - w\|^2 .$$

Proposition 2.2.10 Polarization Identities for complex inner product spaces

Let V be a complex inner product space and v, w two vectors in V. We have:

$$\langle v, w \rangle = \frac{1}{4} \|v+w\|^2 - \frac{1}{4} \|v-w\|^2 + \frac{i}{4} \|v+iw\|^2 - \frac{i}{4} \|v-iw\|^2 .$$

Proof. Exercise for students. Hint.

$$\|v \pm w\|^2 = \|v\|^2 \pm 2\operatorname{Re}\langle v, w \rangle + \|w\|^2 .$$

and

$$\operatorname{Im}\langle v, w \rangle = \operatorname{Re} -i\langle v, w \rangle = \operatorname{Re}\langle v, iw \rangle$$

☐

2.3 Exercises set

Exercise 2.3.1

Let $\mathbb{F} = \mathbb{C}$. Show that if $\langle -, - \rangle$ is an inner product, then

$$\langle v, au + bw \rangle = \bar{a}\langle v, u \rangle + \bar{b}\langle v, w \rangle.$$

Solution. By definition, we know that for all $u, v \in V$ and $a, b \in \mathbb{F}$, we have

$$\langle v, u \rangle = \overline{\langle u, v \rangle}.$$

$$\begin{aligned}
\langle v, au + bw \rangle &= \overline{\langle au + bw, v \rangle} \\
&= \overline{a\langle u, v \rangle + b\langle w, v \rangle} \\
&= \bar{a}\,\overline{\langle u, v \rangle} + \bar{b}\,\overline{\langle w, v \rangle} \\
&= \bar{a}\langle v, u \rangle + \bar{b}\langle v, w \rangle.
\end{aligned}$$

Exercise 2.3.2

For $u = (u_1, u_2)$ and $v = (v_1, v_2)$ in $\mathbb{R}^2$, let

$$\langle u, v \rangle = u_1v_1 - u_2v_1 - u_1v_2 + 4u_2v_2.$$

Show that this function define an inner product on $\mathbb{R}^2$.

Solution. Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ in $\mathbb{R}^2$ and $a, b \in \mathbb{R}$. Then

(1) $\langle v, v \rangle = v_1^2 - v_2v_1 - v_1v_2 + 4v_2^2 = (v_1 - v_2)^2 + 3v_2^2 \geq 0$

Clearly

$$\begin{aligned}
\langle v, v \rangle = 0 &\iff (v_1 - v_2)^2 + 3v_2^2 = 0 \\
&\iff (v_1 - v_2)^2 = 0 \quad \text{and} \quad 3v_2^2 = 0 \\
&\iff v_1 - v_2 = 0 \quad \text{and} \quad v_2 = 0 \\
&\iff v_1 = 0 \quad \text{and} \quad v_2 = 0 \\
&\iff v = 0.
\end{aligned}$$

(2)

$$\begin{aligned}\langle\, v, u\,\rangle &= v_1u_1 - v_2u_1 - v_1u_2 + 4v_2u_2\\ &= u_1v_1 - u_2v_1 - u_1v_2 + 4u_2v_2\\ &= \langle\, u, v\,\rangle\\ &= \overline{\langle\, u, v\,\rangle}.\end{aligned}$$

(3) Let $w = (w_1, w_2)$ in $\mathbb{R}^2$. Then $av + bu = (av_1 + bu_1, av_2 + bu_2)$. Therefore

$$\begin{aligned}\langle\, av + bu, w\,\rangle &= (av_1 + bu_1)w_1 - (av_2 + bu_2)w_1 - (av_1 + bu_1)w_2 + 4(av_2 + bu_2)w_2\\ &= av_1w_1 - av_2w_1 - av_1w_2 + 4av_2w_2 + bw_1u_1 - bu_2w_1 - bu_1w_2 + 4bu_2w_2\\ &= a\langle\, v, w\,\rangle + b\langle\, u, w\,\rangle.\end{aligned}$$

Hence, the function $\langle\, -, -\,\rangle$ define an inner product on $\mathbb{R}^2$.

Exercise 2.3.3

Apply Cauchy-Schwarz inequality to show that for all x_1, x_2, y_1 and y_2 in $\mathbb{R}$,

$$|x_1y_1 + x_2y_2| \le \sqrt{(x_1^2 + x_2^2)(y_1^2 + y_2^2)}.$$

and

$$|x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2| \le \sqrt{(x_1^2 - 2x_1x_2 + 4x_2^2)(y_1^2 - 2y_1y_2 + 4y_2^2)}.$$

Solution. Consider on $\mathbb{R}^2$ the following real inner product : for $u = (x_1, x_2)$ and $v = (y_1, y_2)$ in $\mathbb{R}^2$,

$$\langle\, u, v\,\rangle = x_1y_1 + x_2y_2$$

By Cauchy-Schwarz inequality :

$$|\langle u, v\rangle| \le \|u\|\,\|v\|\,.$$

Hence

$$|x_1y_1 + x_2y_2| \le \sqrt{(x_1^2 + x_2^2)(y_1^2 + y_2^2)}.$$

Similarly, when we consider the following real inner product on $\mathbb{R}^2$:

$$\langle\, u, v\,\rangle = x_1y_1 - 2x_1y_2 + 4x_2y_2$$

we get :

$$|x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2| \le \sqrt{(x_1^2 - 2x_1x_2 + 4x_2^2)(y_1^2 - 2y_1y_2 + 4y_2^2)}.$$

Exercise 2.3.4 Polarization Identities for real inner product spaces

Let V be a real inner product space and v, w two vectors in V. Prove that:

$$\langle\, v, w\,\rangle = \frac{1}{4}\,\|v + w\|^2 - \frac{1}{4}\,\|v - w\|^2\,.$$

Solution. For all v, w two vectors in V, we have

$$\begin{aligned}\|v+w\|^2 &= \langle v+w, v+w\rangle \\ &= \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle \\ &= \|v\|^2 + \|w\|^2 + \langle v,w\rangle + \overline{\langle v,w\rangle}\end{aligned}$$

Since the inner product is considered real, $\overline{\langle v,w\rangle} = \langle v,w\rangle$. Therefore

$$\|v+w\|^2 = \|v\|^2 + \|w\|^2 + 2\langle v,w\rangle. \tag{2.1}$$

Replacing w by $-w$ in the previous equality, we obtain:

$$\|v-w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v,w\rangle. \tag{2.2}$$

From (2.1) and (2.2), we obtain

$$4\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 .$$

Consequently,

$$\langle v,w\rangle = \frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 .$$

Exercise 2.3.5 Polarization Identities for complex inner product spaces

Let V be a complex inner product space and v, w two vectors in V. Prove that:

$$\langle v,w\rangle = \frac{1}{4}\|v+w\|^2 - \frac{1}{4}\|v-w\|^2 + \frac{i}{4}\|v+iw\|^2 - \frac{i}{4}\|v-iw\|^2 .$$

Solution. Clearly for v, w in V, we have

$$\begin{cases}\|v+w\|^2 = \|v\|^2 + 2\operatorname{Re}\langle v,w\rangle + \|w\|^2 \\ \|v-w\|^2 = \|v\|^2 - 2\operatorname{Re}\langle v,w\rangle + \|w\|^2\end{cases} \tag{2.3}$$

Therefore

$$4\operatorname{Re}\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 \tag{2.4}$$

Replacing w by iw in the equation (2.3), we get

$$\begin{cases}\|v+iw\|^2 = \|v\|^2 + 2\operatorname{Re}\langle v,iw\rangle + \|iw\|^2 \\ \|v-iw\|^2 = \|v\|^2 + 2\operatorname{Re}\langle v,-iw\rangle + \|-iw\|^2\end{cases}$$

So

$$\begin{cases}\|v+iw\|^2 = \|v\|^2 + 2\operatorname{Re}-i\langle v,w\rangle + \|w\|^2 \\ \|v-iw\|^2 = \|v\|^2 + 2\operatorname{Re}i\langle v,w\rangle + \|w\|^2\end{cases}$$

Using the fact that

$$\operatorname{Im}\langle v,w\rangle = \operatorname{Re}-i\langle v,w\rangle = \operatorname{Re}\langle v,iw\rangle$$

we obtain

$$\begin{cases}\|v+iw\|^2 = \|v\|^2 + 2\operatorname{Im}\langle v,w\rangle + \|w\|^2 \\ \|v-iw\|^2 = \|v\|^2 - 2\operatorname{Im}\langle v,w\rangle + \|w\|^2 .\end{cases}$$

Hence

$$4\operatorname{Im}\langle v,w\rangle = \|v+iw\|^2 - \|v-iw\|^2 \tag{2.5}$$

Form (2.4) and (2.5), we obtain

$$4\operatorname{Re}\langle v,w\rangle + 4i\operatorname{Im}\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 + i\|v+iw\|^2 - i\|v-iw\|^2 . \tag{2.6}$$

Exercise 2.3.6

Suppose V is a real inner product space.

(1) Show that $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ for any $u, v \in V$.

(2) Show that if $\|u\| = \|v\|$, then $u+v$ is orthogonal to $u-v$.

Solution. **(1)** For any $u, v \in V$, $\langle u+v, u-v\rangle = \langle u, u\rangle - \langle u, v\rangle + \langle v, u\rangle - \langle v, v\rangle = \|u\|^2 - \|v\|^2$.

(2) If $\|u\| = \|v\|$, since $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2 = 0$, $u+v$ is orthogonal to $u-v$.

Exercise 2.3.7

Let $\mathcal{B} = \{u_1, u_2, u_3\}$ be a basis for the Euclidean inner product space $\mathbb{R}^3$, where

$$u_1 = (1, -2, 1), \quad u_2 = (1, 0, 1) \quad \text{and} \quad u_3 = (-2, 0, 1).$$

(1) Use the Gram-Schmidt process to transform the basis $\mathcal{B}$ into an orthogonal basis $\mathcal{B}' = \{v_1, v_2, v_3\}$.

(2) Normalize the basis B' to obtain an orthonormal basis $\mathcal{B}'' = \{w_1, w_2, w_3\}$ for $\mathbb{R}^3$.

(3) Find $\mathcal{B}''^*$ the dual basis of $\mathcal{B}''$.

Solution.
(1) Apply Gram-Schmidt process to obtain an orthogonal basis for R^3.

$$v_1 = u_1 = (1, -2, 1).$$

$$\begin{aligned} v_2 &= u_2 - \frac{< u_2, v_1 >}{\|v_1\|^2} v_1 \\ &= (1, 0, 1) - \frac{1}{3}(1, -2, 1) \\ &= \left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) \end{aligned}$$

$$\begin{aligned} v_3 &= u_3 - \frac{< u_3, v_1 >}{\|v_1\|^2} v_1 - \frac{< u_3, v_2 >}{\|v_2\|^2} v_2 \\ &= (-2, 0, 1) + \frac{1}{6}(1, -2, 1) + \frac{2/3}{4/3}(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}) \\ &= (-2, 0, 1) + \frac{1}{6}(1, -2, 1) + \frac{1}{3}(1, 1, 1) \\ &= \left(\frac{-3}{2}, 0, \frac{3}{2}\right) \end{aligned}$$

(2) Normalize the basis B'

$$\begin{aligned} w_1 &= \frac{v_1}{\|v_1\|} = \left(\frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) = \left(\frac{\sqrt{6}}{6}, \frac{-\sqrt{6}}{3}, \frac{\sqrt{6}}{6}\right). \\ w_2 &= \frac{v_2}{\|v_2\|} = \frac{\sqrt{3}}{2}\left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right) = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right) \\ w_3 &= \frac{v_3}{\|v_3\|} = \frac{\sqrt{2}}{3}\left(\frac{-3}{2}, 0, \frac{3}{2}\right) = \left(\frac{-\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right) \end{aligned}$$

(3) Using Theorem 2.2.2, for all $v = (x_1, x_2, x_3) \in \mathbb{R}^3$:

$$\begin{aligned} v = (x_1, x_2, x_3) &= \langle v, w_1 \rangle w_1 + \langle v, w_2 \rangle w_2 + \langle v, w_3 \rangle w_3 \\ &= \frac{\sqrt{6}}{6}(x_1 - 2x_2 + x_3)w_1 + \frac{\sqrt{3}}{3}(x_1 + x_2 + x_3)w_2 + \frac{\sqrt{2}}{2}(-x_1 + x_3)w_3 \end{aligned}$$

Then $\mathcal{B}''^* = \{f_1, f_2, f_3\}$ where:

$$\begin{aligned} f_1(x_1, x_2, x_3) &= \frac{\sqrt{6}}{6}(x_1 - 2x_2 + x_3) \\ f_2(x_1, x_2, x_3) &= \frac{\sqrt{3}}{3}(x_1 + x_2 + x_3) \\ f_3(x_1, x_2, x_3) &= \frac{\sqrt{2}}{2}(-x_1 + x_3) \end{aligned}$$

Exercise 2.3.8

Let $V = \mathcal{M}_{n\times n}(\mathbb{R})$ be the real vector space of $n \times n$ matrices. Consider the following inner product on V defined by

$$\langle A, B \rangle = \operatorname{tr}(A^{\mathrm{t}} B).$$

Let

$$\mathcal{S}_n = \{A \in V \mid A^{\mathrm{t}} = A\} \quad \text{and} \quad \mathcal{A}_n = \{A \in V \mid A^{\mathrm{t}} = -A\}$$

(1) Show that for all $A \in V$: $A^{\mathrm{t}} + A \in \mathcal{S}_n$ and $A - A^{\mathrm{t}} \in \mathcal{A}_n$.

(2) Show that, every matrix $A \in V$ can be written as $A = X + Y$ where $X \in \mathcal{S}_n$ and $Y \in \mathcal{A}_n$.

(3) Deduce that $V = \mathcal{S}_n \oplus \mathcal{A}_n$.

(4) Show that $\mathcal{S}_n^{\perp} = \mathcal{A}_n$.

(5) Using Cauchy-Schwarz inequality, show that for all matrix $A \in V$: $\operatorname{tr}(A) \le \sqrt{n}\,\sqrt{\operatorname{tr}(A^{\mathrm{t}}A)}$.

(6) Deduce that, if $A \in V$ is an orthogonal matrix, then $\operatorname{tr}(A) \le n$.

Solution. **(1)** for all $A \in V$, we have $(A^{\mathrm{t}} + A)^{\mathrm{t}} = (A^{\mathrm{t}})^{\mathrm{t}} + A^{\mathrm{t}} = A + A^{\mathrm{t}}$, so $A^{\mathrm{t}} + A \in \mathcal{S}_n$ and similarly we have $A - A^{\mathrm{t}} \in \mathcal{A}_n$.

(2) Clearly

$$A = \underbrace{\frac{1}{2}(A^{\mathrm{t}} + A)}_{X} + \underbrace{\frac{1}{2}(A - A^{\mathrm{t}})}_{Y}.$$

(3) From the previous question, we get $V = \mathcal{S}_n + \mathcal{A}_n$. Since the square matrix which is both symmetric and anti-symmetric matrix is the zero matrix, we obtain $V = \mathcal{S}_n \oplus \mathcal{A}_n$.

(4) Let $A \in \mathcal{A}_n$. For all $B \in \mathcal{S}_n$, we have

$$\langle A, B \rangle = \operatorname{tr}(A^{\mathrm{t}} B) = \operatorname{tr}(A\,B).$$

On other hand, we have

$$\langle A, B \rangle = \langle B, A \rangle = \operatorname{tr}(B^{\mathrm{t}} A) = \operatorname{tr}(-BA) = -\operatorname{tr}(AB).$$

Therefore $\langle A, B \rangle = 0$ for all $B \in \mathcal{S}_n$. Hence $\mathcal{A}_n \subseteq \mathcal{S}_n^{\perp}$

From (3), we obtain

$$\dim \mathcal{A}_n = \dim \mathcal{S}_n^{\perp}.$$

Therefore

$$\mathcal{S}_n^{\perp} = \mathcal{A}_n.$$

(5) Using Cauchy-Schwarz inequality, we get for all matrix $A \in V$:

$$\langle\, I_n, A \,\rangle \leq \|I_n\| \, \|A\|$$

Hence

$$\mathrm{tr}(A) \leq \sqrt{n}\, \sqrt{\mathrm{tr}(A^{\mathrm{t}}A)}.$$

(6) As $A \in V$ is an orthogonal matrix, $A^t A = I_n$. So $\mathrm{tr}(A) \leq \sqrt{n}\, \sqrt{\mathrm{tr}(I_n)}$. That means

$$\mathrm{tr}(A) \leq n.$$

Chapter 3

Linear Forms and Duality

Chapter contents

3.1 Introduction

Recall that the set of linear maps of a vector space E into a vector space F on the same field K is a vector space over K denoted $\mathcal{L}(E, F)$. It has dimension $\dim E \cdot \dim F$ and is isomorphic to the space of matrices $M_{\dim F \times \dim E}(K)$.

Linear forms are special types of linear maps, also called covectors. They are important in decomposing quadratic forms into sums of squares, i.e., diagonalization.

Definition 3.1.1

A **linear form** is a linear map from the vector space E to the field K (viewed as a vector space over itself). Its kernel is called a *hyperplane*.

From the dimension theorem, a linear form is either zero or surjective. In the latter case, its kernel is supplementary to a vector line.

Example 3.1.2

The trace is a linear form on the space of square matrices of order n. Thus, the subspace of zero-trace matrices is a hyperplane and has dimension $n^2 - 1$. Its supplementary is the subspace of scalar matrices.

Definition 3.1.3

The space of linear forms $\mathcal{L}(E, K)$ is called the **dual space** of E, denoted E^*.

3.2 Matrix Representation

Let $\{v_1, \ldots, v_n\}$ be a basis of the vector space E, and let $\phi \in E^*$. Then, the matrix representing ϕ is a $1 \times n$ row matrix with entries $\phi(v_i) \in K$. For

$$x = x_1 v_1 + \cdots + x_n v_n, \quad \text{we have} \quad \phi(x) = x_1 \phi(v_1) + \cdots + x_n \phi(v_n),$$

which can be written in matrix form as:

$$\phi(x) = \begin{bmatrix} \phi(v_1) & \cdots & \phi(v_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Thus, every rank-1 matrix can be identified with a linear form.

3.3 Dual Basis and Antedual Basis

From the above, E and E^* have the same dimension and are thus isomorphic.

Theorem 3.3.1

For every basis $\{v_1, \ldots, v_n\}$ of E, there exists a unique basis $\{\phi_1, \ldots, \phi_n\}$ of E^* such that for all i, j:

$$\phi_i(v_j) = \delta_{ij}.$$

This basis is called the **dual basis**, sometimes denoted $\{v_1^*, \ldots, v_n^*\}$.

Proof. From the matrix representation, it is clear that a linear form is entirely determined by the image of each vector from the basis $\{v_1, \cdots, v_n\}$. Thus for each fixed i, the n equations $\varphi_i(v_j) = \delta_{ij}, j = \overline{1, n}$ uniquely define the form φ_i.

Now, let us prove that $\{\varphi_1, \cdots, \varphi_n\}$ isa basis for E^*. Since E^* has the same dimension as E, it is sufficient to prove that the n forms are free:

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, such that

$$\alpha_1 \varphi_1 + \cdots + \alpha_n \varphi_n = 0$$

then, for $j = 1, \ldots, n$, we have

$$\begin{aligned} 0 &= (\alpha_1 \varphi_1 + \cdots + \alpha_n \varphi_n)(v_j) \\ &= \alpha_1 \varphi_1(v_j) + \cdots + \alpha_j \varphi_j(v_j) + \cdots + \alpha_n \varphi_n(v_j) \\ &= \alpha_1 \cdot 0 + \cdots + \alpha_j \cdot 1 + \cdots + \alpha_n \cdot 0 = \alpha_j \end{aligned}$$

(In the other words, for each v_i from the basis of E, we correspond the unique form φ_i from the basis of E^*). Therefore, for

$$x = x_1 v_1 + \cdots + x_n v_n \in E$$

we have $v_i^*(x) = x_i$, which gives

$$x = v_1^*(x) v_1 + \cdots + v_n^*(x) v_n$$

i.e. the coordinates of a vector x in E in the given basis are the images of x by the dual basis. For $\varphi \in E^*$, such that

$$\varphi = \alpha_1 v_1^* + \ldots + \alpha_n v_n^*$$

we have,

$$\varphi(x) = \alpha_1 v_1^*(x) + \cdots + \alpha_n v_n^*(x)$$

In the other hand, from equation (1), we have

$$\varphi(x) = v_1^*(x) \varphi(v_1) + \cdots + v_n^*(x) \varphi(v_n)$$

Since $\varphi(x)$ is uniquely represented, then Equations (2) and (3) give

$$\alpha_i = \varphi(v_i), \text{ for } i = 1, \ldots, n.$$

□

Thus, we have the following corollary:

Corollary 3.3.2

Let $\{v_1, \cdots, v_n\}$ be a basis for the vector space E and $\{v_1^*, \cdots, v_n^*\}$ be the dual basis for the dual vector space E^*. Then, the form- coordinates of a vector $x \in E$ and its dual $x^* = \varphi \in E^*$ in the related bases are

$$x = \begin{pmatrix} v_1^*(x) \\ \vdots \\ v_n^*(x) \end{pmatrix}, \varphi = \varphi(v_1) \quad \cdots \quad \varphi(v_n)$$

Example 3.3.3

Let E be the space of 2×2 trace-zero matrices with basis:

$$e_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then the dual basis $\{e_1^*, e_2^*, e_3^*\}$ satisfies:

$$e_i^* \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & -a_{11} \end{bmatrix} \right) = a_{1i}.$$

Therefore, we can represent such matrix by form- coordinates in the canonical basis by the following column:

$$A = \begin{pmatrix} e_1^*(A) \\ e_2^*(A) \\ e_3^*(A) \end{pmatrix}$$

Also, we represent a linear form φ on the space of null trace matrices in the dual basis by the row vector:

$$\varphi = \begin{pmatrix} \varphi(e_1) & \varphi(e_2) & \varphi(e_3) \end{pmatrix}$$

For example, if $\varphi =$ trace, then, $\varphi = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ is the null form. It is the restriction of the trace on the space of square matrices to its kernal.

Example 3.3.4

Let $A \in GL_n(\mathbb{K})$, then the columns of A constitutes a basis for $\mathbb{K}^n$. The dual basis is given by the rows of its inverse.

$$\text{In fact, let } A = \begin{pmatrix} C_1 & \cdots & C_i & \cdots & C_n \end{pmatrix} \text{ and } A^{-1} = \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix}.$$

From equality $A^{-1}A = I_n$, it deduced that $L_i(C_j) = \delta_{ij}$. i.e. $L_i = C_i^*$, $i = 1, \ldots, n$. Therefore the R_i constitute the dual basis. Hence, to find the antedual basis of the dual basis, we construct the row matrix of the given dual basis, then, we calculate its inverse. The columns of the inverse matrix constitute the antedual basis..

3.4 Orthogonality with Respect to Duality

Let F and F^* be two subvector spaces of E and E^* respectively. We let to students to verify that the following sets are subvector spaces of E and E^* respectively:

$$F^{\perp} = \{\varphi \in E^*, \forall v \in F, \varphi(v) = 0\}$$
$$(F^*)^{\perp} = \{v \in E, \forall \varphi \in F^*, \varphi(v) = 0\}$$

Definition 3.4.1

Let F and F^* be two subvector spaces of E and E^* respectively. The space $F^{\perp}$ (resp. $\left.(F^*)^{\perp}\right)$ is called the orthogonal of F (resp. F^*) respected to the duality.

The subspace of $\mathbb{K}^n$ of the solutions of an homogenous linear system is the orthogonal of the linear forms defining this system. For example, giving the following system:

$$\begin{cases} x_1 + 2x_2 - x_3 = 0 \\ 2x_1 - x_2 + x_4 = 0 \\ x_3 + x_4 = 0 \end{cases}$$

Then $(F^*)^{\perp} = \{(3x, x, 5x, -5x), x \in \mathbb{R}\}$ is the orthogonal of $F^* = \{\varphi_1, \varphi_2, \varphi_3\}$, where the φ_i are the rows of the system matrix:

$$\begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Note that $\dim F^{*}+\dim\left(F^{*}\right)^{\perp}=\dim E=4$. Thus, we can mention the following theorem:

Theorem 3.4.2

Let F be a subvector space of the vector space E. Then, the following relation holds:

$$\dim F+\dim F^{\perp}=\dim E$$

(The same property holds if we exchange E by E^{*}). Indeed, the theorem is an immediate result of the solutions of a linear system of p equations with n unknown coefficients. The space of solutions is of dimension equal to $n-p$. 5

3.5 Exercise series

Exercise 3.5.1 Lagrange Interpolation

Let $\mathbb{R}_n[X]$ be the space of real polynomials of degree $\leq n$, and let $a_0,\ldots,a_n$ be distinct real numbers.

1. Show that the Lagrange polynomials

$$L_i(X)=\prod_{j\neq i}\frac{X-a_j}{a_i-a_j}$$

   form a basis for $\mathbb{R}_n[X]$.

2. Show that the evaluation maps $P\longmapsto P(a_i)$ define a basis for $(\mathbb{R}_n[X])^{*}$.

Exercise 3.5.2

Let $a_0,\ldots,a_n\in\mathbb{R}$ be distinct real numbers and $\varphi_0,\ldots,\varphi_n$ be linear forms on $E=\mathbb{R}_n[X]$ given by the relations:

$$\varphi_i(P)=P\left(a_i\right)\text{ for all } i=0,\ldots,n$$

1. Prove that the set $\{\varphi_0,\ldots,\varphi_n\}$ is a dual basis for E and determine its antedual basis.
2. Deduce the same result for $\varphi_i(P)=P(i)$ for $i=0,\ldots,n$.
3. Same question as in 1) for $n=2$, where φ_i are defined by

$$\varphi_i(P)=\int_0^1 x^i P(x)dx \text{ for all } i=0,\ldots,n$$

4. Same question for $n=2$ and the φ_i are defined by

$$\varphi_0(P)=P(1),\varphi_1(P)=P'(1),\varphi_2(P)=\int_0^1 P(x)dx$$

3.6 Solutions of the series

Solution of exercise 1 First of all, we notice that since the $a_0, \ldots, a_n \in \mathbb{R}$ are all distinct, then the polynomials L_i are well defined and all distinct. 1. Since $|A| = n+1 = \dim E$, then to prove that the set of the polynomials

$$A = \{L_0, \ldots, L_n\}$$

is a basis for $E = \mathbb{R}_{\mathrm{n}}[X]$, it is sufficient to prove that A is free. Let $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$ such that

$$\alpha_0 L_0 + \ldots + \alpha_n L_n = 0$$

That means the polynomial $\alpha_0 L_0 + \ldots + \alpha_n L_n$ is the zero polynomial. Therefore, for every $a_j \in \{a_0, \ldots, a_n\}$, the polynomial evaluated in a_j vanishes. It follows that

$$\alpha_0 L_0 (a_j) + \ldots + \alpha_n L_n (a_j) = 0$$

By using the definition of L_i, that gives

$$L_i (a_j) = \left\{ \begin{array}{ll} 1 & \text{for } j = i \\ 0 & \text{for all } j \neq i \end{array} \right. .$$

Hence, for $j = 0$, we have

$$\alpha_0 L_0 (a_0) + \ldots + \alpha_n L_n (a_0) = 0 = \alpha_0 \times 1 + \alpha_1 \times 0 + \ldots + \alpha_n \times 0 = \alpha_0$$

By the same manner we get

$$\alpha_1 L_1 (a_1) = 0 = \alpha_1, \ldots, \alpha_n L_n (a_n) = 0 = \alpha_n$$

Consequently, the set A is free. 2. For all $i = 0, \ldots, n$, let φ_i be the linear forms defined by

$$\varphi_i : E_{P \longrightarrow P(a_i)} \underset{E_i}{E_i}.$$

7

Since $\dim E^* = \dim E = n + 1$, then to prove that the set

$$A^* = \{\varphi_0, \ldots, \varphi_n\}$$

is a basis for E^*, it is sufficient to prove that it is free. Let $\alpha_0, \ldots, \alpha_{\mathrm{r}1} \in \mathbb{R}$ such that

$$\alpha_0 \varphi_0 + \ldots + \alpha_n \varphi_n = 0$$

That means the linear form

$$\varphi = \alpha_0 \varphi_0 + \ldots + \alpha_n \varphi_n$$

is the zero linear form. Therefore, for every $P \in \mathbb{R}_n[X]$, we get $\varphi(P) = 0$. It follows that

$$\alpha_0 \varphi_0(P) + \ldots + \alpha_n \varphi_n(P) = 0 = \alpha_0 P (a_0) + \ldots + \alpha_n P (a_n)$$

Since $L_j \in \mathbb{R}_n[X]$, then, for $P = L_j$, the previous equality gives

$$0 = \alpha_0 L_j (a_0) + \ldots + \alpha_n L_j (a_n)$$

Since

$$L_j (a_i) = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for all } i \neq j \end{cases}$$

then, for all $j = 0, \ldots, n$, we get

$$0=\alpha_0 L_j\left(a_0\right)+\ldots+\alpha_n L_j\left(a_n\right)=\alpha_j L_j\left(a_j\right)=\alpha_j$$

which means that the the set $A^*=\{\varphi_0, \ldots, \varphi_n\}$ is free. Solution of exercise 2 1. Let $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$ such that

$$\alpha_0 \varphi_0+\ldots+\alpha_n \varphi_n=0$$

Then, for every $P \in \mathbb{R}_n[X]$, we get $\varphi(P)=0$. It follows that

$$\alpha_0 \varphi_0(P)+\ldots+\alpha_n \varphi_n(P)=0=\alpha_0 P\left(a_0\right)+\ldots+\alpha_n P\left(a_n\right)$$

Therefore, for P is one of the elements of the canonical basis $\{1, X, \ldots, X^n\}$, we get

$$\varphi_i(P)=\varphi_i\left(X^j\right)=a_i^j \text{ for } j=1, \ldots, n$$

Thus we have the following system:

$$\begin{cases}\alpha_0+\alpha_1+\ldots+\alpha_n & =0 \\ \alpha_0 a_0+\alpha_1 a_1+\ldots+\alpha_n a_n & =0 \\ \vdots & \vdots \\ \alpha_0 a_0^n+\alpha_1 a_1^n+\ldots+\alpha_n a_n^n & =0\end{cases}$$

By putting the system in matrix form, we have

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_0 & a_1 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots \\ a_0^n & a_1^n & \cdots & a_n^n \end{pmatrix}\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The matrix

$$A=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_0 & a_1 & \cdots & a_n \\ \vdots & \vdots & \vdots & \vdots \\ a_0^n & a_1^n & \cdots & a_n^n \end{pmatrix}$$

is a Vendermand matrix, its determinant is

$$\det A=\prod_{j \neq i}\left(a_i-a_j\right)$$

Since all the a_i are distinct, then $\det A \neq 0$. Therefore System (4) has only zero as solution, i.e. $\alpha_i=0$ for all $i=1, \ldots, n$. Which means that the φ_i are linearly independent. Therefore the set $\{\varphi_0, \ldots, \varphi_n\}$ is a basis for $(\mathbb{R}_n[X])^*$. The antedual basis is constituted of the polynomials $P_0, \ldots, P_n$ such that

$$\varphi_i\left(P_j\right)=\begin{cases}1 & \text{for } i=j \\ 0 & \text{for all } i \neq j\end{cases}$$

9

That gives

$$P_j\left(a_i\right)=\begin{cases}1 & \text{for } i=j \\ 0 & \text{for all } i \neq j\end{cases}$$

First method: For all $j=0, \ldots, n$, let P_j be as follows

$$P_j(X)=p_{0j}+\ldots+p_{jj} X^j+\ldots+p_{nj} X^n$$

Then for a fixed j and $i = 0, \ldots, n$, we have

$$\begin{aligned} P_j(a_0) &= 0 = p_{0j} + a_0 p_{1j} + a_0^2 p_{2j} + \ldots + a_0^n p_{nj} \\ P_j(a_1) &= 0 = p_{0j} + a_1 p_{1j} + a_1^2 p_{2j} + \ldots + a_1^n p_{nj} \\ &\vdots \\ P_j(a_j) &= 1 = p_{0j} + a_j p_{1j} + a_j^2 p_{2j} + \ldots + a_j^n p_{nj} \\ &\vdots \\ P_j(a_n) &= 0 = p_{0j} + a_n p_{1j} + a_n^2 p_{2j} + \ldots + a_n^n p_{nj} \end{aligned}$$

That means we have every fixed j, we have a linear system of the form

$$\begin{cases} p_{0j} + a_0 p_{1j} + a_0^2 p_{2j} + \ldots + a_0^n p_{nj} & = 0 \\ \vdots & \vdots \\ p_{0j} + a_j p_{1j} + a_j^2 p_{2j} + \ldots + a_j^n p_{nj} & = 1 \\ \vdots & \vdots \\ p_{0j} + a_n p_{1j} + a_n^2 p_{2j} + \ldots + a_n^n p_{nj} & = 0 \end{cases}$$

By putting the previous system in matrix form we have

$$\begin{pmatrix} 1 & a_0 & \cdots & a_0^n \\ 1 & a_1 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & \cdots & a_n^n \end{pmatrix} \begin{pmatrix} p_{0j} \\ \vdots \\ p_{jj} \\ \vdots \\ p_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}.$$

The system matrix A is a vendermand matrix, which means that it is invertible, Then the system has the unique solution

$$\begin{pmatrix} p_{0j} \\ \vdots \\ p_{jj} \\ \vdots \\ p_{nj} \end{pmatrix} = \begin{pmatrix} 1 & a_0 & \cdots & a_0^n \\ 1 & a_1 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & a_n & \cdots & a_n^n \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

That implies that the entries of each column of the inverse matrix represents the coefficients p_{ij} of corresponding polynomial P_j for all $j = 0, \ldots, n$. (for students: by using the matrix form in (5) determine the antedual basis for $n = 2$). Second method: Since

$$P_j(a_i) = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for all } i \neq j \end{cases},$$

Then, for a fixed j and all $i = 0, \ldots, n$, with $i \neq j$, the a_i are roots of the polynomial P_j while a_j is not. That yields to

$$P_j(X) = \left(\prod_{i \neq j} (X - a_i) \right) Q_j(X)$$

and

$$P_j(a_j) = \left(\prod_{i \neq j} (a_j - a_i) \right) Q_j(a_j) = 1.$$

Therefore

$$Q_j\left(a_j\right)=\frac{1}{\prod_{i\neq j}\left(a_j-a_i\right)}$$

That means that $Q_j(X)$ is a constant. Otherwise it would be a fraction, while $Q_j(X)$ is a polynomial. By replacing the value of $Q_j(X)$ in relation (6), we get

$$P_j(X)=\frac{\prod_{i\neq j}\left(X-a_i\right)}{\prod_{i\neq j}\left(a_j-a_i\right)}=\prod_{i\neq j}\frac{X-a_i}{a_j-ai}$$

11

Therefore, the antedual basis is constituted of Lagrange interpolations. 2. It is sufficient to take $a_i=i$ for all $i=0,\ldots,n$, then

$$P_j(X)=\prod_{i\neq j}\frac{X-i}{j-i}$$

3. For

$$\varphi_i(P)=\int_0^1 x^iP(x)dx$$

Let $\alpha_0,\ldots,\alpha_n\in\mathbb{R}$ such that

$$\alpha_0\varphi_0+\ldots+\alpha_n\varphi_n=0$$

Then, for every $P\in\mathbb{R}_n[X]$, we get $\varphi(P)=0$. It follows that

$$\alpha_0\varphi_0(P)+\ldots+\alpha_n\varphi_n(P)=0=\alpha_0\int_0^1 x^0P(x)dx+\ldots+\alpha_n\int_0^1 x^nP(x)dx$$

Therefore, for P is one of the elements of the canonical basis $\{1,X,\ldots,X^n\}$, we get the following system:

$$\begin{cases}\alpha_0\int_0^1 X^0dx+\ldots+\alpha_n\int_0^1 X^ndx & =0\\ \alpha_0\int_0^1 Xdx+\ldots+\alpha_n\int_0^1 X^{n+1}dx & =0\\ \vdots & \vdots\\ \alpha_0\int_0^1 X^ndx+\ldots+\alpha_n\int_0^1 X^{2n}dx & =0\end{cases}$$

which gives

$$\begin{cases}\alpha_0+\frac{1}{2}\alpha_1+\ldots+\frac{1}{n+1}\alpha_n & =0\\ \frac{1}{2}\alpha_0+\frac{1}{3}\alpha_1+\ldots+\frac{1}{n+2}\alpha_n & =0\\ \vdots & \vdots\\ \frac{1}{n+1}\alpha_0+\frac{1}{n+2}\alpha_1\ldots+\frac{1}{2n+1}\alpha_n & =0\end{cases}$$

System (7) in matrix form becomes

$$\begin{pmatrix}1 & \frac{1}{2} & \cdots & \frac{1}{n+1}\\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+2}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n+1}\end{pmatrix}\begin{pmatrix}\alpha_0\\ \alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}$$

The matrix of the system is a Hilbert matrix $H=(h_{ij})$ of order $n+1$, such that the entries $h_{ij}=\frac{1}{i+j-1}$. The determinant of the Hilbert matrix is given by the following relation:

$$\det H = \frac{c_n^4}{c_{2n}}$$

where

$$c_n = \prod_{i=1}^{n-1} i^{n-i} = \prod_{i=1}^{n-1} i!$$

That means $\det H \neq 0$. Therefore, System (7) has only zero as solution, i.e. $\alpha_i = 0$ for all $i = 1, \ldots, n$. Which means that the φ_i are linearly independent. Therefore the set $\{\varphi_0, \ldots, \varphi_n\}$ is a basis for $(\mathbb{R}_n[X])^*$. Let us now find the antedual basis of $\{\varphi_0, \ldots, \varphi_n\}$. Let

$$P_j(X) = p_{0j} + p_{1j}X + \ldots + p_{nj}X^n, \text{ for all } j = 1, \ldots, n$$

such that

$$\varphi_i(P_j) = \int_0^1 X^i P_j(X) dx = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for all } i \neq j \end{cases}$$

Relations (8) and (9) yield to the following system in matrix form:

$$\begin{pmatrix} 1 & \frac{1}{2} & \cdots & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \frac{1}{j} & \frac{1}{j+1} & \cdots & \cdots & \frac{1}{n+j} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \cdots & \frac{1}{2n+1} \end{pmatrix} \begin{pmatrix} p_{0j} \\ \vdots \\ p_{jj} \\ \vdots \\ p_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}.$$

13

Since the matrix of the system is the Hilbert matrix, then it is invertible, which allows us to find all the polynomial P_j of the antedual basis by the relation:

$$\begin{pmatrix} p_{0j} \\ \vdots \\ p_{jj} \\ \vdots \\ p_{nj} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \frac{1}{j} & \frac{1}{j+1} & \cdots & \cdots & \frac{1}{n+j} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \cdots & \frac{1}{2n+1} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

Let the inverse matrix be $H^{-1} = \left(h'_{ij}\right)$

$$H^{-1} = \begin{pmatrix} h'_{11} & h'_{12} & \cdots & \cdots & h'_{1(n+1)} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ h'_{j1} & h'_{j2} & \cdots & \cdots & h'_{j(n+1)} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ h'_{(n+1)1} & h'_{(n+1)2} & \cdots & \cdots & h'_{(n+1)(n+1)} \end{pmatrix}.$$

Then,

$$\begin{pmatrix} h'_{1j} \\ \vdots \\ h'_{jj} \\ \vdots \\ h'_{nj} \end{pmatrix} = \begin{pmatrix} h'_{11} & h'_{12} & \cdots & \cdots & h'_{1(n+1)} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ h'_{j1} & h'_{j2} & \cdots & \cdots & h'_{j(n+1)} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ h'_{(n+1)1} & h'_{(n+1)2} & \cdots & \cdots & h'_{(n+1)(n+1)} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \text{ for all } j = 1, \ldots, n,$$

i.e. the elements of the antedual basis are the columns of the inverse of the system matrix.
For example, for $n=2$, we have

$$H=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}, H^{-1}=\begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}$$

The elements of the antedual basis are:

$$P_0:\begin{pmatrix} p_{00} \\ p_{10} \\ p_{20} \end{pmatrix}=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} 9 \\ -36 \\ 30 \end{pmatrix}$$
$$P_1:\begin{pmatrix} p_{01} \\ p_{11} \\ p_{21} \end{pmatrix}=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}=\begin{pmatrix} -36 \\ 192 \\ -180 \end{pmatrix}$$
$$P_2:\begin{pmatrix} p_{02} \\ p_{12} \\ p_{22} \end{pmatrix}=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}=\begin{pmatrix} 30 \\ -180 \\ 180 \end{pmatrix}$$

Which means:

$$\begin{aligned} P_0(X) &= 30X^2-36X+9 \\ P_1(X) &= -180X^2+192X-36 \\ P_2(X) &= 180X^2-180X+30 \end{aligned}$$

For

$$\varphi_0(P)=P(1), \varphi_1(P)=P'(1), \varphi_2(P)=\int_0^1 P(x)dx$$

let $\alpha_0,\alpha_1,\alpha_2\in\mathbb{R}$ such that

$$\alpha_0\varphi_0+\alpha_1\varphi_1+\alpha_2\varphi_2=0$$

Then, for every $P\in\mathbb{R}_2[X]$, we have

$$\begin{aligned} \alpha_0\varphi_0(P)+\alpha_1\varphi_1(P)+\alpha_2\varphi_2(P)=0 \\ \alpha_0P(1)+\alpha_1P'(1)+\alpha_2\int_0^1 P(X)dx=0 \end{aligned}$$

Therefore, for P is one of the elements of the canonical basis $\left\{1,X,X^2\right\}$, we get the following system:

$$\begin{cases} \alpha_0+\alpha_2 & =0 \\ \alpha_0+\alpha_1+\frac{1}{2}\alpha_2 & =0 \\ \alpha_0+2\alpha_1+\frac{1}{3}\alpha_2 & =0 \end{cases}$$

15

The system matrix $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & \frac{1}{2} \\ 1 & 2 & \frac{1}{3} \end{pmatrix}$ is of determinant $=\frac{1}{3}\neq 0$, which means that the system has only zero solution. Therefore $\{\varphi_0,\varphi_1,\varphi_2\}$ is a basis for $(\mathbb{R}_2[X])^*$. Now, let $\{P_0,P_1,P_2\}$ be the antedual basis for ($\mathbb{R}_2[X]$). Then,

$$\varphi_i\left(P_j\right)=\varphi_i\left(p_{0j}+p_{1j}X+p_{2j}X^2\right)=\begin{cases} 1 & \text{for } i=j \\ 0 & \text{for all } i\neq j \end{cases}$$

we get the following systems in matrix form

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} p_{00} \\ p_{10} \\ p_{20} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} p_{01} \\ p_{11} \\ p_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} p_{02} \\ p_{12} \\ p_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$
$$P_0 : \begin{pmatrix} p_{00} \\ p_{10} \\ p_{20} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ -3 \end{pmatrix}$$
$$P_1 : \begin{pmatrix} p_{01} \\ p_{11} \\ p_{21} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -2 \\ \frac{3}{2} \end{pmatrix}$$
$$P_2 : \begin{pmatrix} p_{02} \\ p_{12} \\ p_{22} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \\ 3 \end{pmatrix}$$

Note that The matrix in Systems (11) is the transpose $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & \frac{1}{2} & \frac{1}{3} \end{pmatrix}$ of matrix $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & \frac{1}{2} \\ 1 & 2 & \frac{1}{3} \end{pmatrix}$ in System (10).

Its rows represent the dual basis $\{\varphi_0, \varphi_1, \varphi_2\}$, while the columns of its inverse $\begin{pmatrix} -2 & \frac{1}{2} & 3 \\ 6 & -2 & -6 \\ -3 & \frac{3}{2} & 3 \end{pmatrix}$ represent the antedual basis $\{P_0, P_1, P_2\}$:

$$\begin{aligned} P_0(X) &= -3X^2 + 6X - 2 \\ P_1(X) &= \frac{3}{2}X^2 - 2X + \frac{1}{2} \\ P_2(X) &= 3X^2 - 6X + 3 \end{aligned}$$

Chapter 4

Bilinear forms

Chapter contents

We study, in this chapter, the bilinear forms on finite dimensional vector spaces over a field $\mathbb{F}$. Moreover, we discussed to symmetric forms and their reduction of to a diagonal form in the case when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

4.1 Linear Functionals

Definition 4.1.1 Linear Functional (or 1-form)

Let V be a vector space. Define $V^* = \mathcal{L}(V, \mathbb{F})$. V^* is called the **dual space** of V. The elements of V^* are called **linear functional**. So a linear functional ϕ on V is a linear transformation $\phi : V \longrightarrow \mathbb{F}$.

Example 4.1.2

Let $\mathbb{F}$ be a field and let $a_1, ..., a_n$ be scalars in $\mathbb{F}$. Define a function $f : \mathbb{F}^n \longrightarrow \mathbb{F}$ by

$$f(x_1, ..., x_n) = a_1 x_1 + \cdots + a_n x_n.$$

Then f is a linear functional on $\mathbb{F}^n$.

Every linear functional on $\mathbb{F}^n$ is of this form, for some scalars $a_i, ..., a_n$.

That is immediate from the definition of linear functional because:

$$\begin{aligned}
f(x_1, \ldots, x_n) &= f(\sum_{i=1}^{n} x_i e_i) \\
&= \sum_{i=1}^{n} f(x_i e_i) \\
&= \sum_{i=1}^{n} x_i f(e_i) \\
&= \sum_{i=1}^{n} a_i x_i \\
&= a_1 x_1 + \cdots + a_n x_n.
\end{aligned}$$

Example 4.1.3

Let n be a positive integer and $\mathbb{F}$ a field. The trace function $\mathrm{tr} : \mathbb{F}^{n\times n} \longrightarrow \mathbb{F}$ is a linear functional. Recall that if $A = (a_{ij}) \in \mathbb{F}^{n\times n}$:

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Example 4.1.4

Let $[a, b]$ be a closed interval on the real line and let $\mathcal{C}([a, b])$ be the space of continuous real-valued functions on $[a, b]$. Then the function $L : \mathcal{C}([a, b]) \longrightarrow \mathbb{R}$ defined by

$$L(f) = \int_a^b f(t)\, dt$$

is a linear functional.

Proposition 4.1.5 Dimension of V^*

Suppose that $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis for the finite dimensional vector space V. Define $f_i \in V^*$ by

$$f_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Then, the set $\mathcal{B}^* = \{f_1, f_2, \ldots, f_n\}$ form a basis for V^*. Therefore $\dim V^* = \dim V$.

Proof. See Exercise 4.8.1. ☐

Definition 4.1.6 Dual basis

The set $\mathcal{B}^*$ in the previous proposition is called the dual basis of $\mathcal{B}$.

Theorem 4.1.7

Let V be a finite-dimensional vector space over the field $\mathbb{F}$, and let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for V. Let $\mathcal{B}^* = \{f_1, ..., f_n\}$ be the dual basis of $\mathcal{B}$:

$$f_i(v_j) = \delta_{ij}.$$

Then, for each linear functional f on V we have

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

and for each vector v in V we have

$$v = \sum_{i=1}^{n} f_i(v) v_i.$$

Proof. We have, for all $j = 1, .., , n$:

$$(\sum_{i=1}^{n} f(v_i) f_i)(v_j) = \sum_{i=1}^{n} f(v_i) f_i(v_j) = \sum_{i=1}^{n} f(v_i)\delta_{ij} = f(v_j)$$

Then

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

Let $v \in V$, then this vector can be expressed as $v = c_1 v_1 + \cdots + c_n v_n$. Then for all $j = 1, .., , n$, we have:

$$f_j(v) = f_j(c_1 v_1 + \cdots + c_n v_n) = c_1 f_j(v_1) + \cdots + c_j f_j(v_j) + \cdots + c_n f_j(v_n) = c_j,$$

Hence

$$v = \sum_{i=1}^{n} f_i(v) v_i.$$

□

Proposition 4.1.8

Let V be an n-dimensional vector space and $x_1, \dots, x_k \in V$ linearly independent vectors with $k < n$. Then there exists $f \in V^*$ and $y \notin \operatorname{span}\{x_1, \dots, x_k\}$ such that

$$f(y) = 1 \quad \text{and} \quad f(x) = 0 \quad \text{for all} \quad x \in \operatorname{span}\{x_1, \dots, x_k\}.$$

4.2 Bilinear maps

Definition 4.2.1 Bilinear maps

Let U, V, W be vector spaces over a field $\mathbb{F}$. A map $f : U \times V \longrightarrow W$ is *bilinear* if it is linear in each variable:

$$\begin{aligned} f(au_1 + u_2, v) &= af(u_1, v) + f(u_2, v) \\ f(u, av_1 + v_2) &= af(u, v_1) + f(u, v_2), \end{aligned}$$

for all $u, u_1, u_2 \in U$, $v, v_1, v_2 \in V$ and $a \in \mathbb{F}$.

We will sometimes write $\langle u, v \rangle$ for $f(u, v)$ if f is clear from context.

Note 4.2.2

We denote the set of all $\mathbb{F}$-bilinear map $f : U \times V \longrightarrow W$ by $\mathrm{Bil}_{\mathbb{F}}(U \times V, W)$.

Example 4.2.3

(1) Matrix multiplication is bilinear:

$$(A, B) \longmapsto AB : \mathcal{M}_{m\times n}(\mathbb{F}) \times \mathcal{M}_{n\times k}(\mathbb{F}) \longrightarrow \mathcal{M}_{m\times k}(\mathbb{F}).$$

(2) Composition of maps is bilinear:

$$(\psi, \phi) \longmapsto \psi \circ \phi : \mathcal{L}(U, W) \times \mathcal{L}(V, U) \longrightarrow \mathcal{L}(V, W).$$

Proposition 4.2.4

For any bilinear map $f : U \times V \longrightarrow W$, we have:

$$f(u, 0) = f(0, v) = 0, \qquad \text{for all} \quad u \in U \,\text{and}\, v \in V.$$

Indeed,

$$f(u, 0) = f(u, 0 + 0) = f(u, 0) + f(u, 0)$$

and similarly for $f(0, v)$.

Definition 4.2.5 Bilinear pairing

Let U and V be vector spaces over a field $\mathbb{F}$. A bilinear map $U \times V \longrightarrow \mathbb{F}$ is called a bilinear pairing.

Definition 4.2.6 Bilinear form

Let V be vector spaces over a field $\mathbb{F}$. A bilinear map $V \times V \longrightarrow \mathbb{F}$ is called a bilinear form.

Example 4.2.7

Consider the functions $\mathrm{S}, \mathrm{T} : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ defined as follows: for any $x, y \in \mathbb{R}$,

$$\mathrm{S}(x, y) = x + y, \qquad \mathrm{T}(x, y) = xy.$$

Clearly S is linear but not multilinear, and T is multilinear and not linear.

Note 4.2.8

We denote the set of all $\mathbb{F}$-bilinear forms on V by $\mathrm{Bil}_{\mathbb{F}}(V)$.

Example 4.2.9

(1) Evaluation $(f, v) \longmapsto f(v) : V^* \times V \longrightarrow \mathbb{F}$ is a bilinear pairing.

(2) Let $A \in M_{m\times n}(\mathbb{F})$. Then mapping $B_A : \mathbb{F}^m \times \mathbb{F}^n \longrightarrow \mathbb{F}$ by

$$f_A(x, y) = x^{\mathrm{t}} A y$$

is a bilinear pairing.

We denote by $\mathrm{Bil}(V, V)$ the set of all bilinear forms on V. Note that any scalar multiple of a bilinear form or any sum of two bilinear forms is again a bilinear form. This gives $\mathrm{Bil}(V, V)$ the structure of a vector space over $\mathbb{F}$.

Definition 4.2.10 Special important bilinear forms

Let $f : V \times V \longrightarrow \mathbb{F}$ be a bilinear form. We say that f is:

(1) **Nondegenerate** if $f(u, v) = 0$ for all $u \in V$ implies that $v = 0$.

(2) **Symmetric** if $f(u, v) = f(v, u)$ for all $u, v \in V$.

(3) **Anti-symmetric (skew-symmetric)** if $f(u, v) = -f(v, u)$ for all $u, v \in V$.

(4) **Alternating** if $f(v, v) = 0$ for all $v \in V$.

Example 4.2.11

(1) $V = \mathbb{R}^2$. The following map

$$\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \longmapsto x_1 y_1 + x_2 y_2$$

is a symmetric form on $\mathbb{R}^2 \times \mathbb{R}^2$.

(2) Let $V = \mathcal{C}([-1, 1], \mathbb{R})$. The map

$$\begin{aligned} \mathcal{C}([-1,1],\mathbb{R}) \times \mathcal{C}([-1,1],\mathbb{R}) &\longrightarrow \mathbb{R} \\ (f, g) &\longmapsto \int_{-1}^{1} f(t) g(t) dt \end{aligned}$$

is a symmetric form.

(3) In general, every real inner product is a symmetric bilinear form.

4.3 Bilinear forms and matrices

Definition 4.3.1

Let V be a vector space over $\mathbb{F}$ with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ and let $f : V \times V \longrightarrow \mathbb{F}$ be a bilinear form. The matrix of f with respect to $\mathcal{B}$ is $A = (a_{ij}) \in M_{n\times n}(\mathbb{F})$ given by

$$a_{ij} = f(v_i, v_j),$$

for $1 \leq i, j \leq n$.

Note 4.3.2

Let V be a vector space over $\mathbb{F}$ with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ and let $f : V \times V \longrightarrow \mathbb{F}$ be a bilinear form. We denote $[f]_{\mathcal{B}}$ to the matrix of f with respect to the basis $\mathcal{B}$.

Proposition 4.3.3

Let $f : V \times V \longrightarrow \mathbb{F}$ be a bilinear form with matrix A with respect to $\mathcal{B} = \{v_1, \ldots, v_n\}$. Then f is completely determined by A: if $v = \sum_{i=1}^{n} x_i v_i$ and $w = \sum_{j=1}^{n} y_j v_j$ then

$$f(v, w) = \sum_{i,j=1}^{n} x_i y_j a_{ij},$$

Proof. Using the bilinearity of f:

$$f(v, w) = \sum_{i,j=1}^{n} x_i y_j f(v_i, v_j) = \sum_{i,j=1}^{n} x_i y_j a_{ij}.$$

□

Example 4.3.4

Let $V = \mathbb{R}^2$ and $\mathcal{B} = \{e_1, e_2\}$ the standard basis of V. Consider the following symmetric form

$$\begin{aligned} f : \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) &\longmapsto 3x_1y_1 - 2x_2y_2 + x_1y_2 + x_2y_1 \end{aligned}$$

The matrix of f relative to the standard basis $\mathcal{B}$ is

$$[f]_{\mathcal{B}} = \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix}$$

Proposition 4.3.5

Let V be a vector space over a field $\mathbb{F}$ and $f \in \mathrm{Bil}_{\mathbb{F}}(V)$, and $\mathcal{B}$ an ordered basis of V. Then,

(1) $[\ ]_{\mathcal{B}} : \mathrm{Bil}_{\mathbb{F}}(V) \longrightarrow \mathcal{M}_n(\mathbb{F})$ is an isomorphism of $\mathbb{F}$-vector spaces.

(2) Let $A \in \mathcal{M}_n(\mathbb{F})$ and $f_A \in \mathrm{Bil}_{\mathbb{F}}(\mathbb{F}^n)$ be the bilinear form defined by the matrix A. Then $[f_A]_{\mathcal{S}} = A$, where $\mathcal{S}$ is the standard basis of $\mathbb{F}^n$.

(3) Let $f \in \mathrm{Bil}_{\mathbb{F}}(\mathbb{F}^n)$ and $A = [f]_{\mathcal{S}}$, then, $f = f_A$.

Proof. This is a homework. ☐

Definition 4.3.6

Let f be a symmetric bilinear form on a vector space V.

(1) We say that $\mathbf{u}, \mathbf{v} \in V$ are **orthogonal** with respect to f if $f(\mathbf{u}, \mathbf{v}) = 0$.

(2) If $W \subseteq V$ is a subspace of V, we define the **orthogonal complement of** W **in** V to be

$$W^{\perp} := \{\mathbf{v} \in V \ : \ f(\mathbf{v}, \mathbf{w}) = 0 \text{ for all } \mathbf{w} \in W\}.$$

Lemma 4.3.7

Let $f \in \mathrm{Bil}_{\mathbb{F}}(V)$ and $\mathcal{B} = \{v_1, \ldots, v_n\}$ an ordered basis of V. Then, for any $u, v \in V$, we have

$$[u]^{\mathrm{t}}_{\mathcal{B}}\, [f]_{\mathcal{B}}\, [v]_{\mathcal{B}} = f(u, v).$$

Moreover, if $A \in \mathcal{M}(\mathbb{F})$ is such that

$$[u]^{\mathrm{t}}_{\mathcal{B}}\, A\, [v]_{\mathcal{B}} = f(u, v),$$

then $A = [f]_{\mathcal{B}}$.

Proof. Let $u, v \in V$ and suppose that

$$u = \sum_{i=1}^{n} \alpha_i v_i \quad \text{and} \quad v = \sum_{j=1}^{n} \beta_j v_j.$$

so that

$$[u]^{t}_{\mathcal{B}} = [\alpha_1, \ldots, \alpha_n] \quad \text{and} \quad [v]_{\mathcal{B}} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Then, we have

$$\begin{aligned} f(u,v) &= f\Big(\sum_{i=1}^{n}\alpha_i v_i,\ \sum_{j=1}^{n}\beta_j v_j\Big) \\ &= \sum_{i=1}^{n}\alpha_i f\Big(v_i,\ \sum_{j=1}^{n}\beta_j v_j\Big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\beta_j f(v_i,\ v_j) \end{aligned}$$

Also, we see that

$$\begin{aligned} [u]_{\mathcal{B}}^{\mathsf{t}}\,[f]_{\mathcal{B}}\,[v]_{\mathcal{B}} &= [\alpha_1,\dots,\alpha_n]\,[f]_{\mathcal{B}}\begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\beta_j f(v_i,\ v_j). \end{aligned}$$

Let $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{F})$ such that

$$[u]_{\mathcal{B}}^{\mathsf{t}}\,A\,[v]_{\mathcal{B}} = f(u,v),$$

Then for all i and j, we have

$$[v_i]_{\mathcal{B}}^{\mathsf{t}}\,A\,[v_j]_{\mathcal{B}} = f(v_i,v_j),$$

Hence

$$e_i^{\mathsf{t}}\,A\,e_j = f(v_i,v_j),$$

Finally, we get $a_{ij} = f(v_i,v_j)$, that means $A = [f]_{\mathcal{B}}$. ☐

Proposition 4.3.8 Bilinear form: change of basis formula

Let V be finite-dimensional vector space over a field $\mathbb{F}$ and $f \in \mathrm{Bil}_{\mathbb{F}}(V)$. If $\mathcal{B}$ and $\mathcal{B}'$ be two ordered bases of V , then

$$P^{\mathsf{t}}\ [f]_{\mathcal{B}}\ P\ =\ [f]_{\mathcal{B}'},$$

where $P = P_{\mathcal{B}'\longrightarrow\mathcal{B}}$.

Proof. Let $u, v \in V$, and $P = P_{\mathcal{B}'\longrightarrow\mathcal{B}}$.

We know that

$$[u]_{\mathcal{B}} = P[u]_{\mathcal{B}'} \quad\text{and}\quad [v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$$

We have:

$$\begin{aligned} f(u,v) &= [u]_{\mathcal{B}}^{\mathsf{t}}\,[f]_{\mathcal{B}}\,[v]_{\mathcal{B}} \\ &= \big(P[u]_{\mathcal{B}'}\big)^{\mathsf{t}}\,[f]_{\mathcal{B}}P\,[v]_{\mathcal{B}'} \\ &= [u]_{\mathcal{B}'}^{\mathsf{t}}\ (P^{\mathsf{t}}[f]_{\mathcal{B}}P)\ [v]_{\mathcal{B}'} \end{aligned}$$

Therefore

$$P^{\mathsf{t}}\ [f]_{\mathcal{B}}\ P\ =\ [f]_{\mathcal{B}'},$$

☐

4.4 Rank and radical

Definition 4.4.1 Radical

Let $f : V \times V \longrightarrow \mathbb{F}$ be a symmetric bilinear form. The radical $\operatorname{rad} f$ of f is the vector subspace of V given by

$$\operatorname{rad} f := \{v \in V \mid f(v, v') = 0, \text{ for all } v' \in V\} = V^{\perp}.$$

Definition 4.4.2 Rank

Let $f : V \times V \longrightarrow \mathbb{F}$ be a symmetric bilinear form such that V is finite-dimensional, we define the rank of f by

$$\operatorname{rank} f =: \dim V - \dim \operatorname{rad} f.$$

Here is how to understand both the rank and the radical of f.

Proposition 4.4.3 Bilinear symmetric form and dual space

Let f be a bilinear symmetric form on a vector space V. Define the map $\sigma_f : V \longrightarrow V^*$ by

$$\sigma_f(v)(w) = f(v, w),$$

for $v, w \in V$. Then

(1) $\sigma_f(v) \in V^*$ since f is linear in the second slot.

(2) $\sigma_f : V \longrightarrow V^*$ is linear since f is linear in the first slot.

(3) $\ker \sigma_f = \{v \in V \mid \sigma_f(v) = 0\} = \{v \in V \mid f(v, w) = 0 \text{ for all } w \in V\} = \operatorname{rad} f$.

Thus $\operatorname{rad} f \leq V$ and $\operatorname{rank} f = \operatorname{rank} \sigma_f$ when V is finite-dimensional.

Moreover f is non-degenerate if and only if σ_f one-to-one or, when V is finite-dimensional, is an isomorphism.

(4) Let f have matrix $A = (a_{ij})$ with respect to a basis $v_1, \ldots, v_n$ of V. Then

$$\sigma_f(v_j)(v_i) = f(v_j, v_i) = a_{ji} = a_{ij},$$

where we used the symmetry of A in the last equality. It follows that

$$\sigma_f(v_j) = \sum_{i=1}^{n} a_{ij} v_i^*$$

so that A is the matrix of σ_f with respect to the dual bases $\{v_1, \ldots, v_n\}$ and $\{v_1^*, \ldots, v_n^*\}$ of V and V^*.

Lemma 4.4.4

Let $f : V \times V \longrightarrow \mathbb{F}$ be a symmetric bilinear form on a finite-dimensional vector space V with matrix A with respect to some basis of V. Then $\operatorname{rank} f = \operatorname{rank} A$. In particular, f is non-degenerate if and only if $\det A \neq 0$.

Example 4.4.5

We contemplate some symmetric bilinear forms on $\mathbb{F}^3$:

(1) $f(x,y) = x_1y_1 + x_2y_2 - x_3y_3$. With respect to the standard basis, we have

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

so that $\operatorname{rank} f = 3$.

(2) $g(x,y) = x_1y_2 + x_2y_1$. Here the matrix with respect to the standard basis is

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

so that g has rank 2 and radical $\operatorname{span}\{e_3\}$.

4.5 Classification of symmetric bilinear forms

In this section, we consider that $\mathbb{F}$ is a field of characteristic not equal 2, (i.e. $1+1 \neq 0$).

Lemma 4.5.1

Let $f : V \times V \longrightarrow \mathbb{F}$ be a symmetric bilinear form such that $f(v,v) = 0$, for all $v \in V$. Then $f \equiv 0$.

Proof. Let $v, w \in V$. We show that $f(v,w) = 0$. We know that $f(v+w, v+w) = 0$ and expanding out gives us

$$0 = f(v,v) + 2f(v,w) + f(w,w) = 2f(v,w).$$

Since $2 \neq 0$ in $\mathbb{F}$, $f(v,w) = 0$. ☐

Theorem 4.5.2 Diagonalization Theorem

Let f be a symmetric bilinear form on a finite-dimensional vector space over $\mathbb{F}$. Then there is a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ of V with respect to which the matrix of f is diagonal:

$$f(v_i, v_j) = 0, \quad \text{for all} \quad 1 \leq i \neq j \leq n.$$

We call $\{v_1, \dots, v_n\}$ a diagonalising basis for f.

Proof. We will prove this theorem by using the proof by induction on $\dim V$.

(1) Clearly the hypothesis holds if $\dim V = 1$.

(2) Now suppose it holds for all vector spaces of dimension at most $n-1$ and that f is a symmetric bilinear form on a vector space V with $\dim V = n$.

There are two possibilities: if $f(v,v) = 0$, for all $v \in V$, then, by the previous lemma, $f(v,w) = 0$, for all $v, w \in V$, and any basis is trivially diagonalizing.

Otherwise, there is $v_1 \in V$ with $f(v_1, v_1) \neq 0$ and we set

$$U := \operatorname{span} v_1, \quad W := \{v \mid f(v_1, v) = 0\} \leq V.$$

We have:

(a) $U \cap W = \{0\}$: if $\lambda v_1 \in W$ then $0 = f(v_1, \lambda v_1) = \lambda f(v_1, v_1)$ forcing $\lambda = 0$.

(b) $V = U + W$: for $v \in V$, write

$$v = \tfrac{f(v_1,v)}{f(v_1,v_1)} v_1 + (v - \tfrac{f(v_1,v)}{f(v_1,v_1)} v_1).$$

The first summand is in U while

$$f\big(v_1, v - \tfrac{f(v_1,v)}{f(v_1,v_1)} v_1\big) = f(v_1, v) - f(v_1, v) = 0$$

so the second summand is in W.

We conclude that $V = U \oplus W$. We therefore apply the inductive hypothesis to $f|_{W \times W}$ (the restriction of f on $W \times W$) to get a basis $\{v_2, \dots, v_n\}$ of W with $f(v_i, v_j) = 0$, for $2 \leq i \neq j \leq n$.

Now $\{v_1, \dots, v_n\}$ is a basis of V and, further, since $v_j \in W$, for $j > 1$, $f(v_1, v_j) = 0$ so that

$$f(v_i, v_j) = 0, \quad \text{for all} \quad 1 \leq i \neq j \leq n.$$

Thus the inductive hypothesis holds at $\dim V = n$ and so the theorem is proved. ☐

Remark 4.5.3. We can do a little better if $\mathbb{F}$ is $\mathbb{C}$ or $\mathbb{R}$: when $B(v_i, v_i) \neq 0$, either

(1) If $\mathbb{F} = \mathbb{C}$, replace v_i with $v_i/\sqrt{f(v_i, v_i)}$ to get a diagonalising basis with each $f(v_i, v_i)$ either 0 or 1.

(2) If $\mathbb{F} = \mathbb{R}$, replace v_i with $v_i/\sqrt{|f(v_i, v_i)|}$ to get a diagonalising basis with each $f(v_i, v_i)$ either 0, 1 or -1.

Example 4.5.4

Let $f : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ be a symmetric bilinear form such that its matrix in the standard basis of $\mathbb{R}^3$ is

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Find a diagonalising basis for f.

Solution: First notes that $A_{11} \neq 0$ so take $v_1 = e_1$. We seek v_2 among y such that

$$0 = f(v_1, y) = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} Ay = \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} y = y_1 + 2y_2 + y_3.$$

We try $v_2 = (1, -1, 1)$ for which

$$f(v_2, y) = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix} Ay = \begin{pmatrix} 0 & 3 & 0 \end{pmatrix} y = 3y_2$$

In particular, $f(v_2, v_2) = -3 \neq 0$ so we can carry on.

Now seek v_3 among y such that $f(v_1, y) = f(v_2, y) = 0$, that is:

$$\begin{cases} y_1 + 2y_2 + y_3 = 0 \\ 3y_2 = 0. \end{cases}$$

A solution is given by $v_3 = (1, 0, -1)$ and $f(v_3, v_3) = -1$.
We have therefore arrived at the diagonalising basis

$$\{(1,0,0),(1,-1,1),(1,0,-1).\}$$

We can verify that:

$$\begin{pmatrix}1 & 0 & 0\\ 1 & -1 & 1\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}1 & 2 & 1\\ 2 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\ 0 & -1 & 0\\ 0 & 1 & -1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ & -3 & 0\\ 0 & 0 & -1\end{pmatrix}$$

Note that: starting from a different v_1 would give a different, equally correct answer.

4.6 Sylvester's Theorem

Let f be a symmetric bilinear form on a real finite-dimensional vector space. We know that there is a diagonalising basis $v_1, \dots, v_n$ with each $f(v_i, v_i) \in \{\pm 1, 0\}$ and would like to know how many of each there are. We give a complete answer.

Definition 4.6.1 Positive and negative definite symmetric bilinear forms

Let f be a symmetric bilinear form on a real vector space V.
Say that f is positive definite if $f(v, v) > 0$, for all $v \in V \setminus \{0\}$.
Say that f is negative definite if $f(v, v) < 0$ is for all $v \in V \setminus \{0\}$.

Definition 4.6.2 Signature of symmetric real bilinear forms

If V is finite-dimensional real vector space, the signature of f is the pair (p, q) where

$$\begin{aligned} p &= \max\{\dim U \mid U \leq V \text{ with } f|_{U\times U} \text{ positive definite}\}\\ q &= \max\{\dim W \mid W \leq V \text{ with } f|_{W\times W} \text{ negative definite}\}. \end{aligned}$$

We write $\operatorname{sgn}(f) = p - q$..

Remark 4.6.3. A symmetric bilinear form f on V is positive definite if and only if it is an inner product on V.

Theorem 4.6.4 Sylvester's Law of Inertia

Let f be a symmetric bilinear form of signature (p, q) on a finite-dimensional real vector space Then:

(a) $p + q = \operatorname{rank} f$;

(b) any *diagonal* matrix representing f has p positive entries and q negative entries.

Proof. Set $K = \operatorname{rad} f$, $r = \operatorname{rank} f$ and $n = \dim V$ so that $\dim K = n - r$.

Let $U \leq V$ be a p-dimensional subspace on which B is positive definite and W a q-dimensional subspace on which f is negative definite.

First note that $U \cap K = \{0\}$ since $f(k, k) = 0$, for all $k \in K$. Thus, by the dimension formula,

$$\dim(U + K) = \dim U + \dim K = p + n - r.$$

Moreover, if $v = u + k \in U + K$, with $u \in U$ and $k \in K$, then $f(v, v) = f(u + k, u + k) = f(u, u) \geq 0$.

From this we see that $W \cap (U + K) = \{0\}$: if $w \in W \cap (U + K)$ then $f(w, w) \geq 0$ by what we just proved but also $f(w, w) \leq 0$ since $w \in W$. Thus $f(w, w) = 0$ and so, by definiteness on W, $w = 0$. Thus

$$\dim W + (U + K) = \dim W + \dim(U + K) = q + n + p - r \leq \dim V = n$$

so that $p + q \leq r$.

Now let $v_1, \ldots, v_n$ be a diagonalising basis of f with $\hat{p}$ positive entries on the diagonal of the corresponding matrix representative A of f and $\hat{q}$ negative entries. Then f is positive definite on the $\hat{p}$-dimensional space $\operatorname{span} v_i \mid f(v_i, v_i) > 0$ (exercise!). Thus $\hat{p} \leq p$. Similarly, $\hat{q} \leq q$.

However $r = \operatorname{rank} A$ is the number of non-zero entries on the diagonal, that is $r = \hat{p} + \hat{q}$. We therefore have

$$r = \hat{p} + \hat{q} \leq p + q = r$$

so that $p = \hat{p}$, $q = \hat{q}$ and $p + q = r$. □

Example 4.6.5

Find the rank and signature of $f = f_A$ where

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Solution: we have already found a diagonalising basis $v_1 = (1, 0, 0)$, $v_2 = (1, -1, 1)$, $v_3 = (1, 0, -1)$ so we need only count how many $f(v_i, v_i)$ are positive and how many negative. In this case, $f(v_1, v_1) = 1 > 0$ while $f(v_2, v_2) = -3 < 0$ and $f(v_3, v_3) = -1 < 0$. Thus the signature is $(1, 2)$ while $\operatorname{rank} f = 1 + 2 = 3$.

Remark 4.6.6.

(a) Here is a useful sanity check: symmetric bilinear B of signature (p, q) on an n-dimensional V has $p, q, p + q \leq n$ (since $p, q, p + q$ are all dimensions of subspaces of n-dimensional V or V^*).

(b) A symmetric bilinear form of signature $(n, 0)$ on a real n-dimensional vector space is simply an inner product.

4.7 Nondegenerate bilinear forms

We will now introduce the important notion of nondegeneracy of a bilinear form. Nondegenerate bilinear forms arise throughout mathematics. For example, an inner product is an example of a nondegenerate bilinear form.

Definition 4.7.1 Nondegenerate bilinear form

Let V be a finite dimensional $\mathbb{F}$-vector space, $f \in \mathrm{Bil}_{\mathbb{F}}(V)$. We say that f is nondegenerate if the following property holds:

$$f(u, v) = 0 \text{ for every } u \in V \quad \Longrightarrow \quad v = 0_V.$$

If f is not nondegenerate, then we say that f is degenerate.

Lemma 4.7.2

$f \in \mathrm{Bil}_{\mathbb{F}}(V)$ and $\mathcal{B} = \{v_1, \ldots, v_i\}$ be a basis for V. Then, f is nondegenerate if and only if $[f]_{\mathcal{B}}$ is an invertible matrix.

Proof. Suppose that f is nondegenerate. We will show that $A = [f]_{\mathcal{B}}$ is invertible by showing that $\ker T_A = \{\mathbf{0}\}$. So, suppose that $\mathbf{x} \in \mathbb{K}^n$ is such that

$$A\mathbf{x} = \mathbf{0}.$$

Then, for every $\mathbf{y} \in \mathbb{K}^n$ we have

$$0 = \mathbf{y}^{\mathrm{t}}\mathbf{0} = \mathbf{y}^{\mathrm{t}}A\mathbf{x} = f_A(\mathbf{y}, \mathbf{x}).$$

As $[-]_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n$ is an isomorphism we have $\mathbf{x} = [u]_{\mathcal{B}}$ for some unique $v \in V$. Moreover, if $\mathbf{y} \in \mathbb{K}^n$ then there is some unique $u \in V$ such that $\mathbf{y} = [v]_{\mathcal{B}}$. Hence, we have just shown that

$$0 = f_A(\mathbf{y}, \mathbf{x}) = [u]_{\mathcal{B}}^{\mathrm{t}}[f]_{\mathcal{B}}[v]_{\mathcal{B}} = f(u, v),$$

Therefore, since f is nondegenerate

$$f(u, v) = 0, \text{ for every } u \in V \Longrightarrow v = 0_V,$$

Hence, $\mathbf{x} = [u]_{\mathcal{B}} = \mathbf{0}$ so that $\ker T_A = \{\mathbf{0}\}$ and A must be invertible. Conversely, suppose that $A = [f]_{\mathcal{B}}$ is invertible. We want to show that f is nondegenerate so that we must show that if

$$f(u, v) = 0, \qquad \text{for every } u \in V,$$

then $v = 0_V$. Suppose that $f(u, v) = 0$, for every $u \in V$. Then, by Lemma 3.1.6, this is the same as

$$0 = f(u, v) = [u]_{\mathcal{B}}^{\mathrm{t}}A[v]_{\mathcal{B}}, \text{ for every } u \in V.$$

In particular, if we consider $e_i = [v_i]_{\mathcal{B}}$ then we have

$$0 = e_i^{\mathrm{t}}A[v]_{\mathcal{B}}, \text{ for every } i \Longrightarrow A[v]_{\mathcal{B}} = \mathbf{0}.$$

As A is invertible this implies that $[v]_{\mathcal{B}} = \mathbf{0}$ so that $v = 0_V$, since $[-]_{\mathcal{B}}$ is an isomorphism.

□

Proposition 4.7.3

Let V be a $\mathbb{F}$-vector space, $f \in \mathrm{Bil}_{\mathbb{K}}(V)$ a nondegenerate bilinear form. Then, f induces an isomorphism of $\mathbb{F}$-vector spaces

$$\sigma_f : V \longrightarrow V^*; v \longmapsto \sigma_f(v),$$

where

$$\sigma_f(v) : V \longrightarrow \mathbb{F}; u \longmapsto \sigma_f(v)(u) = f(u, v).$$

Proof. Clearly σ_f is well-defined, ie, that σ_f is $\mathbb{F}$-linear and $\sigma_f(v) \in V^*$, for every $v \in V$.

Since we know that $\dim V = \dim V^*$ it suffices to show that σ_f is injective. So, suppose that $v \in \ker \sigma_f$. Then, $\sigma_f(v) = 0 \in V^*$, so that $\sigma_f(v)$ is the zero linear form. Hence, we have $\sigma_f(v)(u) = 0$, for every $u \in V$. Thus, using nondegeneracy of f we have

$$0 = \sigma_f(v)(u) = f(u, v), \text{ for every } u \in V, \Longrightarrow v = 0_V.$$

Hence, σ_f is injective and the result follows. ☐

Remark 4.7.4.

(1) We could have also defined an isomorphism

$$\hat{\sigma}_f : V \longrightarrow V^*,$$

where

$$\hat{\sigma}_f(v)(u) = f(v, u), \text{ for every } u \in V.$$

(2) If f is symmetric then we have $\sigma_f = \hat{\sigma}_f$

(3) The converse of the previous proposition : suppose that σ_f induces an isomorphism

$$\sigma_f : V \longrightarrow V^*.$$

Then, f is nondegenerate. This follows because σ_f is injective.

Definition 4.7.5 Left (right) f-complement

Let $f \in \mathrm{Bil}_{\mathbb{F}}(V)$. Let $E \subset V$ be a nonempty subset. Then, we define the (right) f-complement of E in V to be the set

$$E_r^{\perp} = \{v \in V \mid f(u, v) = 0 \text{ for every } u \in E\}$$

this is a subspace of V

Similarly, we define the (left) f-complement of E in V to be the set

$$E_l^{\perp} = \{v \in V \mid f(v, u) = 0, \text{ for every } u \in E\};$$

Remark 4.7.6. It's clear that if f is symmetric or anti-symmetric, we have

$$E_l^{\perp} = E_r^{\perp}.$$

In this case we write $E^{\perp}$.

Proposition 4.7.7

Let $f \in \mathrm{Bil}_K(V)$ be (anti-)symmetric and nondegenerate, $U \subset V$ a subspace of V. Then,

$$\dim U + \dim U^{\perp} = \dim V.$$

Proof. As f is nondegenerate we can consider the isomorphism

$$\sigma_f : V \longrightarrow V^*,$$

We will show that

$$\sigma_f\left(U^{\perp}\right) = \operatorname{ann}_{V^*}(U) = \{\alpha \in V^* \mid \alpha(u) = 0, \text{ for every } u \in U\}.$$

Indeed, suppose that $w \in U^{\perp}$. Then, for every $u \in U$, we have

$$\sigma_f(w)(u) = f(u, w) = 0,$$

so that $\sigma_f(w) \in \operatorname{ann}_{V^*}(U)$. Conversely, let $\alpha \in \operatorname{ann}_{V^*}(U)$. Then, $\alpha = \sigma_f(w)$, for some $w \in V$, since σ_f is an isomorphism. Hence, for every $u \in U$, we must have

$$0 = \alpha(u) = \sigma_f(w)(u) = f(u, w),$$

so that $w \in U^{\perp}$ and $\alpha = \sigma_f(w) \in \sigma_f\left(U^{\perp}\right)$. Hence,

$$\dim U^{\perp} = \dim \sigma_B\left(U^{\perp}\right) = \dim \operatorname{ann}_{V^*}(U) = \dim V - \dim U.$$

□

4.8 Exercises set

Exercise 4.8.1

Suppose that $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis for the finite dimensional vector space V. For all $1 \leq i \leq n$, let $f_i \in V^* = \mathcal{L}(V, \mathbb{F})$ given by

$$f_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

(1) Show that $\mathcal{B}^* = \{f_1, f_2, \dots, f_n\}$ form a basis for V^*.

(2) Deduce that $\dim V^* = \dim V$.

Solution. **(1)** Let $\alpha_1, \dots, \alpha_n$ be scalars such that

$$\sum_{i=1}^{n} \alpha_i f_i = 0.$$

Then for all $r \in \{1, ..., n\}$, we have

$$\sum_{i=1}^{n} \alpha_i f_i(v_r) = 0.$$

So

$$\sum_{i=1}^{n} \alpha_i \delta_{ij} = 0.$$

So $\alpha_r = 0$. Therefore the set $\{f_1, \dots, f_n\}$ is linearly independent. In addition, for all $h \in V^*$, we have

$$h = \sum_{i=1}^{n} h(v_i) f_i.$$

(2) From **(1)**, we obtain

$$\dim V^* = |\mathcal{B}^*| = |\mathcal{B}| = \dim V.$$

Exercise 4.8.2

Let $f : U \times V \longrightarrow W$ be a bilinear mapping. Show that

$$f(u, 0) = f(0, v) = 0$$

for all $u \in U$ and $v \in V$.

Solution. Let $u \in U$ and $v \in V$ be two arbitrary vectors. Then

$$f(u, 0) = f(u, 0 + 0) = f(u, 0) + f(u, 0),$$

and

$$f(0, v) = f(0 + 0, v) = f(0, v) + f(0, v).$$

Hence $f(u, 0) = f(0, v) = 0$.

Exercise 4.8.3

Show that the following are bilinear maps:

(1) Matrix multiplication $M : \mathcal{M}_{n\times p}(\mathbb{F}) \times : \mathcal{M}_{p\times m}(\mathbb{F}) \longrightarrow \mathcal{M}_{n\times m}(\mathbb{F})$, $M(A, B) = AB$.

(2) Evaluation mapping: $E : V^* \times V \longrightarrow \mathbb{F}$, $E(f, v) = f(v)$.

(3) $T : \mathcal{M}_2(\mathbb{Q}) \times \mathcal{M}_2(\mathbb{Q}) \longrightarrow \mathbb{Q}$, $T(A, B) = \mathrm{tr}(AB)$.

Solution. **(1)** Clearly, for all $\alpha \in \mathbb{F}$, $A_1, A_2 \in \mathcal{M}_{n\times p}(\mathbb{F})$ and $B_1, B_2\mathcal{M}_{p\times m}(\mathbb{F})$, we have:

$$\begin{aligned} M(A_1 + \alpha A_2, B_1) &= (A_1 + \alpha A_2)B_1 \\ &= (A_1B_1 + (\alpha A_2)B_1 \\ &= (A_1B_1 + \alpha(A_2B_1) \\ &= M(A_1, B_1) + \alpha M(A_2, B_1) \end{aligned}$$

Similarly, we have:

$$M(A_1, B_1 + \alpha B_2) = M(A_1, B_1) + \alpha M(A_1, B_2).$$

(2) For all $\alpha \in \mathbb{F}$, $u, v \in V$ and $f, g \in V^*$, we have:

$$\begin{aligned} E(u + \alpha v, f) &= f(u + \alpha v) \\ &= f(u) + \alpha f(v) \\ &= E(u, f) + \alpha E(v, f), \end{aligned}$$

and

$$\begin{aligned} E(u, f + \alpha f) &= (f + \alpha f)(u) \\ &= f(u) + (\alpha f)(u) \\ &= f(u) + \alpha(f(u)) \\ &= E(u, f) + \alpha E(u, g). \end{aligned}$$

Exercise 4.8.4

Let V and W be $\mathbb{F}$-vector spaces. For $f \in V^*$ and $g \in W^*$, we consider the mapping $\phi : V \times W \longrightarrow \mathbb{F}$ defined by

$$\phi(v, w) = f(v)g(w).$$

Show that ϕ is bilinear form on $V \times W$.

Solution. For all $\alpha \in \mathbb{F}$, $v_1, v_2 \in V$, and $w_1, w_2 \in W$, we have:

$$\begin{aligned} \phi(v_1 + \alpha v_2, w_1) &= f(v_1 + \alpha v_2)g(w_1) \\ &= \big(f(v_1) + \alpha f(v_2)\big)g(w_1) \\ &= f(v_1)g(w_1) + \alpha f(v_2)g(w_1) \\ &= \phi(v_1, w_1) + \alpha\phi(v_2, w_1). \end{aligned}$$

Similarly,

$$\phi(v_1, w_1 + \alpha w_2) = \phi(v_1, w_1) + \alpha\phi(v_1, w_2).$$

Exercise 4.8.5 composition between linear and bilinear is bilinear

Let U, V, W_1 and W be vector spaces over a field $\mathbb{F}$, and $f : U \times V \longrightarrow W_1$ a bilinear mapping. Show that for each linear map $g : W_1 \longrightarrow W_2$ the composition $g \circ f$ is bilinear.

Solution. Let $F = g \circ f : U \times V \longrightarrow W_2$. Then for all $u_1, u_2 \in U$, v_1, v_2 and $\alpha \in \mathbb{F}$:

$$\begin{aligned} F(u_1 + \alpha u_2, v_1) &= (g \circ f)(u_1 + \alpha u_2, v_1) \\ &= g\big(f(u_1 + \alpha u_2, v_1)\big) \\ &= g\big(f(u_1, v_1) + \alpha f(u_2, v_1)\big) \\ &= g\big(f(u_1, v_1)\big) + \alpha g\big(f(u_2, v_1)\big) \\ &= F(u_1, v_1) + \alpha F(u_2, v_1). \end{aligned}$$

Similarly, we can prove the linearity fir the second argument, that means:

$$F(u_1, v_1 + \alpha v_2) = F(u_1, v_1) + \alpha F(u_1, v_2).$$

Exercise 4.8.6

Let V and W be vector spaces over a field $\mathbb{F}$, and $f : V \times V \longrightarrow W$ is both bilinear and linear. Show that f is the zero map.

Solution. For all $v_1, v_2 \in V$, we have:

$$f(v_1, v_2) = f(v_1 + 0, 0 + v_1) = f\Big((v_1, 0) + (0, v_2)\Big)$$

Using the linearity of f, we get

$$f(v_1, v_2) = f(v_1, 0) + f(0, v_2).$$

Since f is considered bilinear $f(v_1, 0) = f(0, v_2) = 0$ (see Exercise 4.8.2). Therefore $f(v_1, v_2) = 0$ for all $v_1, v_2 \in V$. Hence $f = 0$.

Exercise 4.8.7

Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for a finite-dimensional $\mathbb{F}$-vector space V, and $f \in \operatorname{Bil}_{\mathbb{F}}(V)$. Show that f is symmetric if and only if

$$f(v_i, v_j) = f(v_j, v_i), \quad \text{for all} \quad 1 \leq i, j \leq n,$$

Solution. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis for V. By definition, it's clear that, if f is symmetric, then

$$f(v_i, v_j) = f(v_j, v_i) \quad \text{for all} \quad 1 \leq i, j \leq n.$$

Conversely, let u and v be two vectors in V, then

$$u = \sum_{i=1}^{n} \alpha_i v_i \quad \text{and} \quad v = \sum_{j=1}^{n} \beta_j v_j.$$

Using the bilinearity of f, we get

$$\begin{aligned} f(u,v) &= f\Big(\sum_{i=1}^{n}\alpha_i v_i,\ \sum_{j=1}^{n}\beta_j v_j\Big) \\ &= \sum_{i=1}^{n} f\Big(\alpha_i v_i,\ \sum_{j=1}^{n}\beta_j v_j\Big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} f\Big(\alpha_i v_i,\ \beta_j v_j\Big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j f(v_i,\ v_j) \end{aligned}$$

Similarly, we can show that

$$f(v,u) = \sum_{j=1}^{n}\sum_{i=1}^{n} \beta_j\alpha_i f(v_j,\ v_i)$$

Since $f(v_i,v_j) = f(v_j,v_i)$, for all $1 \le i,j \le n$, we obtain $f(u,v) = f(v,u)$.

Exercise 4.8.8

Consider the bilinear form $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ is given by

$$f(x,y) = 2x_1y_1 + 3x_1y_2 + y_1x_2$$

where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

Let $\mathcal{S} = \{e_1, e_2\}$ be the standard basis of $\mathbb{R}^2$, and $\mathcal{B} = \{v_1, v_2\}$ such that

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(1) Find $[f]_{\mathcal{S}}$, $[f]_{\mathcal{B}}$ and $P = P_{\mathcal{S}\longrightarrow\mathcal{B}}$.

(2) Verify that

$$P^{t}\ [f]_{\mathcal{B}}\ P \ = \ [f]_{\mathcal{S}},$$

Solution. **(1)** Let us write the matrix of f in the standard basis.

$$f(e_1,e_1) = 2, \quad f(e_1,e_2) = 3, \quad f(e_2,e_1) = 1, \quad f(e_2,e_2) = 0$$

hence the matrix of f in the standard basis is

$$[f]_{\mathcal{S}} = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$$

Similarly,

$$f(v_1,v_1) = 6, \quad f(v_1,v_2) = 3, \quad f(v_2,v_1) = 5, \quad f(v_2,v_2) = 2$$

Hence

$$[f]_{\mathcal{B}} = \begin{pmatrix} 6 & 3 \\ 5 & 2 \end{pmatrix}$$

(2) By definition

$$P_{\mathcal{B}\longrightarrow\mathcal{S}} = [[v_1]_{\mathcal{S}} \mid [v_2]_{\mathcal{S}}] = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Therefore

$$P = P_{\mathcal{S}\longrightarrow\mathcal{B}} = \left(P_{\mathcal{B}\longrightarrow\mathcal{S}}\right)^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

So

$$P^t\ [f]_{\mathcal{B}}\ P\ = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 6 & 3 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} = [f]_{\mathcal{S}}.$$

Exercise 4.8.9

Consider the bilinear form

$$f : \mathbb{Q}^3 \times \mathbb{Q}^3 \longrightarrow \mathbb{Q}; (x, y) \longmapsto x_1y_2 + x_3y_2 + x_2y_1.$$

Is f nondegenerate?

Solution. We have

$$A = [f]_{\mathcal{S}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

which is non-invertible. Hence f is degenerate.

Exercise 4.8.10

Let $V = \mathbf{P}_2(\mathbb{R})$ be a vector space over $\mathbb{R}$ of polynomials of degree at most 2 with coefficients in $\mathbb{R}$. For $f, g \in V$ define the bilinear form $\phi : V \times V \longrightarrow \mathbb{R}$ by:

$$\psi(f, g) = \int_{-1}^{1} xf(x)g(x)\ dx.$$

(1) Is ψ non-degenerate or degenerate?

(2) Give the matrix A associated to ψ relative to the standard basis $\mathcal{B} = \{1, x, x^2\}$ of V.

(3) Find a basis of V for which the matrix associated to ψ is diagonal.

(4) Find the rank and signature of ψ.

Solution. **(1)** Let $f = a + bx + cx^2 \in V$ such that $\psi(f, g) = 0$ for all $g \in V$. Then

$$\psi(f, 1) = \psi(f, x) = \psi(f, x^2) = 0$$

That means

$$\int_{-1}^{1} ax + bx^2 + cx^3\ dx = 0, \quad \int_{-1}^{1} ax^2 + bx^3 + cx^4\ dx = 0, \quad \text{and} \quad \int_{-1}^{1} ax^3 + bx^4 + cx^5\ dx = 0.$$

Therefore

$$\begin{cases} \frac{2b}{3} = 0 \\ \frac{2a}{3} + \frac{2c}{5} = 0 \\ \frac{2b}{5} = 0 \end{cases}$$

So $b = 0$ and $3c = -5a$. Take for example $(a, b, c) = (-3, 0, 5)$, that means $f = -3 + 5x^2$. Then

$$\psi(-3 + 5x^2, g) = 0 \quad \text{for all } g \in V.$$

Hence ψ is degenerate.

(2) Let $f_1 = 1$, $f_2 = x$ and $f_3 = x^2$. By definition

$$A = [\psi]_{\mathcal{B}} = \begin{pmatrix} \psi(f_1, f_1) & \psi(f_1, f_2) & \psi(f_1, f_3) \\ \psi(f_2, f_1) & \psi(f_2, f_2) & \psi(f_2, f_3) \\ \psi(f_3, f_1) & \psi(f_3, f_2) & \psi(f_3, f_3) \end{pmatrix}$$

After calculation, we get

$$A = [\psi]_{\mathcal{B}} = \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix}$$

(3) The matrix A is denationalization :

$$\begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Replace $R_1 \longrightarrow 3R_1$, $R_2 \longrightarrow 15R_1$ and $R_3 \longrightarrow 5R_1$

$$\begin{pmatrix} 0 & 2 & 0 \\ 10 & 0 & 6 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Replace $C_1 \longrightarrow 3C_1$, $C_2 \longrightarrow 15C_1$ and $C_3 \longrightarrow 5C_1$

$$\begin{pmatrix} 0 & 30 & 0 \\ 30 & 0 & 30 \\ 0 & 30 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $R_3 \longrightarrow R_3 + (-1)R_1$:

$$\begin{pmatrix} 0 & 30 & 0 \\ 30 & 0 & 30 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $C_3 \longrightarrow C_3 + (-1)C_1$:

$$\begin{pmatrix} 0 & 30 & 0 \\ 30 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & -3 \\ 0 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $C_1 \longrightarrow C_1 + C_2$:

$$\begin{pmatrix} 30 & 30 & 0 \\ 30 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & -3 \\ 15 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $R_1 \longrightarrow R_1 + R_2$:

$$\begin{pmatrix} 60 & 30 & 0 \\ 30 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 15 & 0 \\ 0 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & -3 \\ 15 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $C_2 \longrightarrow C_2 - (1/2)C_1$:

$$\begin{pmatrix} 60 & 0 & 0 \\ 30 & -15 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 15 & 0 \\ 0 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & \frac{-3}{2} & -3 \\ 15 & \frac{15}{2} & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $R_2 \longrightarrow R_2 - (1/2)R_1$:

$$\begin{pmatrix} 60 & 0 & 0 \\ 0 & -15 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 15 & 0 \\ \frac{-3}{2} & \frac{15}{2} & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & \frac{-3}{2} & -3 \\ 15 & \frac{15}{2} & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $R_2 \longrightarrow 2R_2$:

$$\begin{pmatrix} 60 & 0 & 0 \\ 0 & -30 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 15 & 0 \\ -3 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \begin{pmatrix} 3 & \frac{-3}{2} & -3 \\ 15 & \frac{15}{2} & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Replacing $C_2 \longrightarrow 2C_2$:

$$\underbrace{\begin{pmatrix} 60 & 0 & 0 \\ 0 & -60 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{D} = \underbrace{\begin{pmatrix} 3 & 15 & 0 \\ -3 & 15 & 0 \\ -3 & 0 & 5 \end{pmatrix}}_{P^{\mathsf{t}}} \begin{pmatrix} 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{2}{5} & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 3 & -3 & -3 \\ 15 & 15 & 0 \\ 0 & 0 & 5 \end{pmatrix}}_{P}$$

Hence

$$D = P^{\mathsf{t}}AP.$$

Let

$$q_1 = 3 + 15x, \quad q_2 = -3 + 15x, \quad \text{and} \quad q_3 = -3 + 5x^2.$$

If we take $\mathcal{B}' = \{q_1, q_2, q_3\}$, then the matrix of ψ relative to this basis is:

$$[\psi]_{\mathcal{B}} = \begin{pmatrix} 60 & 0 & 0 \\ 0 & -60 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Verify that $\psi(q_1, q_1) = 60$, $\psi(q_2, q_2) = -60$, $\psi(q_3, q_3) = 0$, and $\psi(q_1, q_2) = \psi(q_2, q_3) = \psi(q_1, q_3) = 0$.

(4) We have already found a diagonalising basis $\mathcal{B}' = \{q_1, q_2, q_3\}$, so we need only count how many $\psi(q_i, q_i)$ are positive and how many negative. In this case, $\psi(q_1, q_1) = 60 > 0$ while $\psi(q_2, q_2) = -60 < 0$ and $\psi(q_3, q_3) = 0$. Thus the signature is $(1, 1)$ while $\operatorname{rank} \psi = 1 + 1 = 2$.

Exercise 4.8.11

Consider the bilinear form

$$f : \mathcal{M}_2(\mathbb{Q}) \times \mathcal{M}_2(\mathbb{Q}) \longrightarrow \mathbb{Q}; (X, Y) \longmapsto \operatorname{tr}(XY).$$

(1) Show that f is nondegenerate.

(2) FInd the matrix of f relative to the standard basis of $\mathcal{M}_2(\mathbb{Q})$

(3) Find a basis of V for which the matrix associated to f is diagonal.

(4) Find the rank and signature of f.

Solution. **(1)** Let

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

be the standard basis of $\mathcal{M}_2(\mathbb{Q})$.

Suppose that

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \in \mathcal{M}_2(\mathbb{Q})$$

such that

$$f(X, Y) = 0, \text{ for every } Y \in \mathcal{M}_2(\mathbb{Q}).$$

Then, in particular, we have

$$f\left(X, e_i\right) = 0, i \in \{1, 2, 3, 4\}$$

Hence,

$$\begin{cases} x_{11} = f\left(X, e_1\right) = 0, \\ x_{12} = f\left(X, e_3\right) = 0, \\ x_{21} = f\left(X, e_2\right) = 0, \\ x_{22} = f\left(X, e_4\right) = 0. \end{cases}$$

So that $X = 0_2 \in \mathcal{M}_2(\mathbb{Q})$.

(2)

$$[f]_{\mathcal{S}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(3) We have

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

Replacing $R_2 \longrightarrow R_2 + R_3$:

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

Replacing $C_2 \longrightarrow C_2 + C_3$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $R_3 \longrightarrow 2R_3$ and $C_3 \longrightarrow 2C_3$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $R_3 \longrightarrow R_3 + (-1)R_2$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $C_3 \longrightarrow C_3 + (-1)C_2$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{D} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P^{\mathrm{t}}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{P}.$$

Put

$$w_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad w_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

Then the matrix of f relative to the basis $\{w_1, w_2, w_3, w_4\}$ is D.

The signature of f is $(3, 1)$ while $\operatorname{rank} f = 3 + 1 = 4$.

Exercise 4.8.12

Consider the following symmetric bilinear form $B = f_A : \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}$ where

$$A = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix}$$

Find a matrix P and a diagonal matrix D such that $P^{\mathrm{t}} A P = D$.

Solution. Method 1:

$$\begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $R_1 \longrightarrow R_1 + R_3$

$$\begin{pmatrix} 1 & 2 & 1 & 2 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $C_1 \longrightarrow C_1 + C_3$

$$\begin{pmatrix} 2 & 2 & 1 & 2 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 2 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Replacing $R_2 \longrightarrow R_2 + (-1)R_1$

$$\begin{pmatrix}2 & 2 & 1 & 2\\ 0 & -2 & -1 & -1\\ 1 & 0 & 0 & 2\\ 2 & 1 & 2 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $C_2 \longrightarrow C_2 + (-1)C_1$

$$\begin{pmatrix}2 & 0 & 1 & 2\\ 0 & -2 & -1 & -1\\ 1 & -1 & 0 & 2\\ 2 & -1 & 2 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & -1 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $R_3 \longrightarrow 2R_3$ and $C_3 \longrightarrow 2C_3$

$$\begin{pmatrix}2 & 0 & 2 & 2\\ 0 & -2 & -2 & -1\\ 2 & -2 & 0 & 4\\ 2 & -1 & 4 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & -1 & 2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $R_3 \longrightarrow R_3 + (-1)R_1$

$$\begin{pmatrix}2 & 0 & 2 & 2\\ 0 & -2 & -2 & -1\\ 0 & -2 & -2 & 2\\ 2 & -1 & 4 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ -1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & -1 & 2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $C_3 \longrightarrow C_3 + (-1)C_1$

$$\begin{pmatrix}2 & 0 & 0 & 2\\ 0 & -2 & -2 & -1\\ 0 & -2 & -2 & 2\\ 2 & -1 & 2 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ -1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & -1 & 0\\ 0 & 1 & 0 & 0\\ 1 & -1 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

$R_4 \longrightarrow R_4 + (-1)R_1$

$$\begin{pmatrix}2 & 0 & 0 & 2\\ 0 & -2 & -2 & -1\\ 0 & -2 & -2 & 2\\ 0 & -1 & 2 & -2\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ -1 & 0 & 1 & 0\\ -1 & 0 & -1 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & -1 & 0\\ 0 & 1 & 0 & 0\\ 1 & -1 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

$C_4 \longrightarrow C_4 + (-1)C_1$

$$\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & -2 & -2 & -1\\ 0 & -2 & -2 & 2\\ 0 & -1 & 2 & -2\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ -1 & 0 & 1 & 0\\ -1 & 0 & -1 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & -1 & -1\\ 0 & 1 & 0 & 0\\ 1 & -1 & 1 & -1\\ 0 & 0 & 0 & 1\end{pmatrix}$$

$R_3 \longrightarrow R_3 + (-1)R_2$

$$\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & -2 & -2 & -1\\ 0 & 0 & 0 & 3\\ 0 & -1 & 2 & -2\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ 0 & -1 & 2 & 0\\ -1 & 0 & -1 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & -1 & -1\\ 0 & 1 & 0 & 0\\ 1 & -1 & 1 & -1\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $C_3 \longrightarrow C_3 + (-1)C_2$

$$\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & -2 & 0 & -1\\ 0 & 0 & 0 & 3\\ 0 & -1 & 3 & -2\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 1 & -1 & 0\\ 0 & -1 & 2 & 0\\ -1 & 0 & -1 & 1\end{pmatrix}\begin{pmatrix}0 & 2 & 1 & 0\\ 2 & 0 & 0 & 1\\ 1 & 0 & 0 & 2\\ 0 & 1 & 2 & 0\end{pmatrix}\begin{pmatrix}1 & -1 & 0 & -1\\ 0 & 1 & -1 & 0\\ 1 & -1 & 2 & -1\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Replacing $R_4 \longrightarrow (-2)R_4$ and $C_4 \longrightarrow (-2)C_4$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & -6 \\ 0 & 2 & -6 & -8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 2 & 0 & 2 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 2 & 2 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Replacing $R_4 \longrightarrow R_4 + R_2$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & -6 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 1 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & -1 & 2 & 2 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Replacing $C_4 \longrightarrow C_4 + C_2$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & -6 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 1 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Replacing $R_3 \longrightarrow R_3 + (-1)R_4$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -6 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -2 & 1 & 2 \\ 1 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Replacing $C_3 \longrightarrow C_3 + (-1)C_4$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ -1 & -2 & 1 & 2 \\ 1 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 1 & -2 & 1 \\ 1 & -1 & 1 & 1 \\ 0 & 0 & 2 & -2 \end{pmatrix}$$

Method 2:

We need to start with v_1 with $B(v_1, v_1) \neq 0$. Those diagonal zeros say that none of the standard basis will do so let us try $v_1 = (1, 1, 0, 0)$ for which $B(v_1, v_1) = 4$.

Now seek v_2 among the y with

$$0 = B(v_1, y) = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} A\mathbf{y} = \begin{pmatrix} 2 & 2 & 1 & 1 \end{pmatrix} \mathbf{y} = 2y_1 + 2y_2 + y_3 + y_4.$$

We take $v_2 = (0, 0, 1, -1)$ with

$$B(v_2, y) = \begin{pmatrix} 0 & 0 & 1 & -1 \end{pmatrix} A\mathbf{y} = \begin{pmatrix} 1 & -1 & -2 & 2 \end{pmatrix} \mathbf{y} = y_1 - y_2 - 2y_3 + 2y_4.$$

We need to start with v_1 with $f(v_1, v_1) \neq 0$. Those diagonal zeros say that none of the standard basis will do so let us try $v_1 = (1, 1, 0, 0)$ for which $f(v_1, v_1) = 4$. Now seek v_2 among the y with

$$0 = f(v_1, y) = (1\,1\,0\,0)Ay = (2\,2\,1\,1)y = 241 + 242 + 43 + 34.$$

We take $v_2 = (0, 0, 1, -1)$ with

$$B(v_2, y) = (0\,0\,1\ -1)Ay = (1\ -1\ -2\,2)8 = 11 - 42 - 243 + 241.$$

Then $B(v_2, v_2) = -4$ and we seek v_3 among the y with $B(v_1, y) = B(v_2, y) = 0$, that is:

$$\begin{aligned} 2y_1 + 2y_2 + y_3 + y_4 &= 0 \\ y_1 - y_2 - 2y_3 + 2y_4 &= 0. \end{aligned}$$

One solution is $v_3 = (-3, 5, -4, 0)$ with

$$B(v_3, y) = \begin{pmatrix} -3 & 5 & -4 & 0 \end{pmatrix} A\mathbf{y} = 3\begin{pmatrix} 2 & -2 & -1 & -1 \end{pmatrix}\mathbf{y} = 3(2y_1 - 2y_2 - y_3 - y_4).$$

Thus $B(v_3, v_3) = -36$ and we need to find $v_4 = y$ with $B(v_1, y) = B(v_2, y) = B(v_3, y) = 0$:

$$\begin{aligned} 2y_1 + 2y_2 + y_3 + y_4 &= 0 \\ y_1 - y_2 - 2y_3 + 2y_4 &= 0 \\ 2y_1 - 2y_2 - y_3 - y_4 &= 0. \end{aligned}$$

A solution is $v_4 = (0, 4, -5, -3)$ with $B(v_4, v_4) = 36$. We now have a diagonalising basis with $B(v_i, v_i) = 4, -4, -36, 36$ so B has signature $(2, 2)$ and so has rank 4.

After all this linear equation solving it is probably good to check our answer: let P have the v_j as columns and check that $P^T AP$ is diagonal:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -3 & 5 & -4 & 0 \\ 0 & 4 & -5 & -3 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & -3 & 0 \\ 1 & 0 & 5 & 4 \\ 0 & 1 & -4 & -5 \\ 0 & -1 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & -36 & 0 \\ 0 & 0 & 0 & 36 \end{pmatrix}$$

Exercise 4.8.13

Let $f : V \times V \longrightarrow \mathbb{F}$ be a symmetric bilinear form. Show that

$$\operatorname{rad} f := \{v \in V \mid f(v, v') = 0 \text{ for all } v' \in V\}.$$

is a vector subspace of V.

Solution. Since $f(0, v) = 0$ for all $v \in V$, $0 \in \operatorname{rad} f$, so $\operatorname{rad} f \neq \emptyset$. Let $v_1, v_2 \in \operatorname{rad} f$ and $\alpha \in \mathbb{F}$. Then for all $v \in V$

$$f(v_1 + \alpha v_2, v) = f(v_1, v) + \alpha f(v_2, v) = 0.$$

Hence $v_1 + \alpha v_2 \in \operatorname{rad} f$.

Exercise 4.8.14

Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and $f \in \operatorname{Bil}_{\mathbb{F}}(V)$. Show that, if f is nondegenerate, then

$$f(u, v) = 0 \quad \text{for every } v \in V \quad \Longrightarrow \quad u = 0_V.$$

Solution. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. We know that f is nondegenerate if and only if $\det[f]_{\mathcal{B}} \neq 0$. Assume that f is nondegenerate, and let g the bilinear form on V defined by

$$g(u, v) = f(v, u) \quad \text{for all } u, v \in V.$$

Clearly g is a symmetric bilinear form and

$$[g]_{\mathcal{B}} = [f]_{\mathcal{B}}^{\mathsf{t}}.$$

Therefore $\det[g]_{\mathcal{B}} = \det[f]_{\mathcal{B}}^{\mathsf{t}} = \det[f]_{\mathcal{B}} \neq 0$. That means g is nondegenerate, and hence

$$g(v, u) = 0 \quad \text{for every } v \in V \quad \Longrightarrow \quad u = 0_V.$$

Which is give the requested implication.

Exercise 4.8.15

Let $E \subset V$ be a nonempty subset and $f \in \mathrm{Bil}_{\mathbb{F}}(V)$ be (anti-)symmetric. Show that

$$E^{\perp} = (\mathrm{Span}_{\mathbb{F}}\, E)^{\perp}.$$

Solution. Obviously, we have

$$(\mathrm{Span}_{\mathbb{F}}\, E)^{\perp} \subset E^{\perp},$$

since if $f(u, v) = 0$, for every $u \in \mathrm{Span}_{\mathbb{F}}(E)$, then this must also hold for those $u \in E$. Hence,

$$v \in \mathrm{Span}_{\mathbb{F}}(E)^{\perp} \Longrightarrow v \in E^{\perp}.$$

Conversely, if $v \in E^{\perp}$, so that $f(\mathbf{e}, v) = 0$, for every $\mathbf{e} \in E$, then if $w = c_1 e_1 + \ldots + c_k e_k \in \mathrm{Span}_{\mathbb{F}}(E)$ for some $e_i \in E$, then

$$f(w, v) = f\left(c_1 e_1 + \ldots c_k e_k, w\right) = c_1 f\left(e_1, v\right) + \ldots + c_k f\left(e_k, w\right) = 0 + \ldots + 0 = 0.$$

Exercise 4.8.16

Let V be a $\mathbb{F}$-vector space and $f \in \mathrm{Bil}_{\mathbb{F}}(V)$. Suppose that $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$ is an ordered basis of V and $\mathcal{B}^* = \{v_1^*, \ldots, v_n^*\} \subset V^*$ is the dual basis. Define the linear mapping $\sigma_f : V \longrightarrow V^*; v \longmapsto \sigma_f(v)$, by

$$\sigma_f(v)(u) = f(u, v) \quad \text{for all } u, v \in V.$$

Show that $[\sigma_f]_{\mathcal{B}}^{\mathcal{B}^*} = [f]_{\mathcal{B}}$.

Solution. By definition,

$$[\sigma_f]_{\mathcal{B}}^{\mathcal{B}^*} = \left[[\sigma_f(v_1)]_{\mathcal{B}^*} \cdots [\sigma_f(v_n)]_{\mathcal{B}^*} \right].$$

Now, for each $i, \sigma_f(v_i) \in V^*$ is a linear form on V so we need to know what it does to elements of V. Suppose that

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n \in V$$

Then,

$$\sigma_f(v_i)(v) = f\left(\sum_{k=1}^{n} \lambda_k v_k, v_i\right) = \sum_{k=1}^{n} \lambda_k f(v_k, v_i)$$

and

$$\left(\sum_{j=1}^{n} f(v_j, v_i) v_j^*\right)(v) = \left(\sum_{j=1}^{n} f(v_j, v_i) v_j^*\right)\left(\sum_{k=1}^{n} \lambda_k v_k\right) = \sum_{k=1}^{n} \lambda_k f(v_k, v_i),$$

so that we must have

$$\sigma_f(v_i) = \sum_{j=1}^{n} f(v_j, v_i) v_j^*$$

Hence,

$$[\sigma_f]_{\mathcal{B}}^{\mathcal{B}^*} = [f]_{\mathcal{B}}$$

It is now clear that B is nondegenerate precisely when the morphism σ_B is an isomorphism.

Chapter

5

Quadratic and Hermitian forms

Chapter contents

This chapter gives the basic properties of Hermitian and quadratic forms.

5.1 Real and complex symmetric bilinear forms

Throughout this section we consider only real or complex vector spaces, that is, vector spaces over the field of real numbers or the field of complex numbers.

$$\mathbb{F} = \mathbb{R} \quad \text{or} \quad \mathbb{F} = \mathbb{C}$$

Proposition 5.1.1 Polarisation identity

Let $f \in \mathrm{Bil}_{\mathbb{F}}(V)$ be a symmetric bilinear form. Then, for any $u, v \in V$, we have

$$f(u,v) = \frac{1}{2}\Big(f(u+v, u+v) - f(u,u) - f(v,v) \Big).$$

Proof. Left as an exercise for the reader. ☐

Corollary 5.1.2

Let $f \in \mathrm{Bil}_{\mathbb{F}}(V)$ be a nonzero symmetric bilinear form. Then, there exists nonzero $v \in V$ such that

$$f(v,v) \neq 0.$$

Proof. Suppose that the result does not hold: that is, for every $v \in V$ we have $f(v,v) = 0$. Then, using the polarisation identity, we get, for every $u, v \in V$,

$$f(u,v) = \frac{1}{2}(f(u+v, u+v,) - f(u,u) - f(v,v)) = \frac{1}{2}(0-0-0) = 0.$$

Hence, we must have that $f = 0$ is the zero bilinear form, which contradicts our assumption on f. Hence, ther must exist some $v \in V$ such that $f(v,v) \neq 0$. ☐

Theorem 5.1.3 Classification of nondegenerate symmetric bilinear forms over $\mathbb{C}$

Let $f \in \mathrm{Bil}_{\mathbb{C}}(V)$ be symmetric and nondegenerate. Then, there exists an ordered basis $\mathcal{B} \subset V$ such that

$$[f]_{\mathcal{B}} = I_{\dim V}.$$

Proof. By the previous corollary, there exists some nonzero $v_1 \in V$ such that

$$f(v_1, v_1) \neq 0$$

(we know that f is nonzero since it is nondegenerate).

Let

$$E_1 = \mathrm{Span}_{\mathbb{C}}\{v_1\}$$

and consider

$$E_1^{\perp} = \{w \in V \mid f(w, v_1) = 0\}.$$

We have

$$E_1 \cap E_1^{\perp} = \{0_V\}$$

indeed, let $x \in E_1 \cap E_1^{\perp}$. Then, $x = cv_1$, for some $c \in \mathbb{C}$. As $x \in E_1^{\perp}$ we must have

$$0 = f(x, v_1) = f(cv_1, v_1) = cf(v_1, v_1)$$

so that $c = 0$ (as $(f(v_1, v_1) \neq 0)$). Thus, by Proposition 4.7.7, we obtain

$$V = E_1 \oplus E_1^{\perp}.$$

Moreover, f restricts to a nondegenerate symmetric bilinear form on $E_1^{\perp}$: indeed, the restriction is

$$f_{|E_1^{\perp}} : E_1^{\perp} \times E_1^{\perp} \longrightarrow \mathbb{C}; (u, u') \longmapsto f(u, u'),$$

and this is a symmetric bilinear form. We need to check that it is nondegenerate. Suppose that $w \in E_1^{\perp}$ is such that, for every $z \in E_1^{\perp}$ we have

$$f(z, w) = 0.$$

Then, for any $v \in V$, we have $v = cv_1 + z, z \in E_1^{\perp}, c \in \mathbb{C}$, so that

$$f(v, w) = f(cv_1 + z, w) = cf(v_1, w) + f(z, w) = 0 + 0 = 0,$$

where we have used the assumption on w and that $w \in E_1^\perp$. Hence, using nongeneracy of f on V we see that $w = 0$. Hence, we have that f is also nondegenerate on $E_1^\perp$.

As above, we can now find $v_2 \in E_1^\perp$ such that $f(v_2, v_2) \neq 0$ and, if we denote $E_2 = \mathrm{Span}_{\mathbb{C}}\{v_2\}$, then

$$E_1^\perp = E_2 \oplus E_2^\perp,$$

where $E_2^\perp$ is the f-complement of E_2 in $E_1^\perp$. Hence, we have

$$V = E_1 \oplus E_2 \oplus E_2^\perp.$$

Proceeding in the manner we obtain

$$V = E_1 \oplus \cdots \oplus E_n$$

where $n = \dim V$, and where $E_i = \mathrm{Span}_{\mathbb{C}}\{v_i\}$. Moreover, by construction we have that

$$f(v_i, v_j) = 0, \text{ for } i \neq j.$$

Define

$$b_i = \frac{1}{\sqrt{f(v_i, v_i)}} v_i$$

we know that the square root $\sqrt{f(v_i, v_i)}$ exists (and is nonzero) since we are considering $\mathbb{C}$-scalars. Then, it is easy to see that

$$f(b_i, b_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Finally, since

$$V = \mathrm{Span}_{\mathbb{C}}\{b_1\} \oplus \cdots \oplus \mathrm{Span}_{\mathbb{C}}\{b_n\},$$

we have that $\mathcal{B} = \{b_1, \ldots, b_n\}$ is an ordered basis such that

$$[f]_{\mathcal{B}} = I_n.$$

□

Corollary 5.1.4

Let $A \in \mathrm{GL}_n(\mathbb{C})$ be a symmetric matrix. Then, there exists $P \in \mathrm{GL}_n(\mathbb{C})$ such that

$$P^{\mathrm{t}} A P = I_n.$$

Since A is an invertible matrix the bilinear form $f_A \in \mathrm{Bil}_{\mathbb{C}}(\mathbb{C}^n)$ is symmetric and nondegenerate.

Theorem 5.1.5 Sylvester's law of inertia

Let V be an $\mathbb{R}$-vector space, $f \in \mathrm{Bil}_{\mathbb{R}}(V)$ a nondegenerate symmetric bilinear form. Then, there is an ordered basis $\mathcal{B} \subset V$ such that $[f]_{\mathcal{B}}$ is a diagonal matrix

$$[f]_{\mathcal{B}} = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix},$$

where $d_i \in \{1, -1\}$.

Proof. The proof is similar to the proof of the previous theorem: we determine $v_1, \ldots, v_n \in V$ such that

$$V = \operatorname{Span}_{\mathbb{R}}\{v_1\} \oplus \cdots \oplus \operatorname{Span}_{\mathbb{R}}\{v_n\}$$

and with $f(v_i, v_j) = 0$, whenever $i \neq j$.

However, we now run into a problem: what if $f(v_i, v_i) < 0$? We can't find a real square root of a negative number so we can't proceed as in the complex case. However, if we define

$$\delta_i = \sqrt{|f(v_i, v_i)|}, \text{ for every } i$$

then we can obtain a basis $\mathcal{B} = (b_1, \ldots, b_n)$, where we define

$$b_i = \frac{1}{\delta_i} v_i$$

Then, we see that

$$f(b_i, b_j) = \begin{cases} 0, i \neq j \\ \pm 1, i = j \end{cases}$$

and $[f]_{\mathcal{B}}$ is of the required form.

□

Remark 5.1.6. If p is the number of 1's appearing on the diagonal and q the number of -1's appearing on the diagonal, then

$$\operatorname{sgn}(f) = p - q.$$

5.2 Quadratic forms

Definition 5.2.1 Quadratic form

A quadratic form on a vector space V over $\mathbb{F}$ is a function $Q : V \longrightarrow \mathbb{F}$ of the form

$$Q(v) = f(v, v),$$

for all $v \in V$, where $f : V \times V \longrightarrow \mathbb{F}$ is a symmetric bilinear form.

Remark 5.2.2. For $v \in V$ and $\lambda \in \mathbb{F}$,

$$Q(\lambda v) = f(\lambda v, \lambda v) = \lambda^2 Q(v)$$

so Q is emphatically not a linear function!

Example 5.2.3

Here are two quadratic forms on $\mathbb{F}^3$:

(1) $Q(x) = x_1^2 + x_2^2 - x_3^2 = f_A(x, x)$ where

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

(2) $Q(x) = x_1x_2 = f_A(x,x)$ where

$$A = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Proposition 5.2.4 Polarisation of quadratic forms

Let $Q : V \longrightarrow \mathbb{F}$ be a quadratic form with $Q(v) = f(v,v)$ for a symmetric bilinear form f. Then

$$f(v,w) = \tfrac{1}{2}\big(Q(v+w) - Q(v) - Q(w)\big),$$

for all $v, w \in V$. f is called the polarisation of Q.

Proof. Expand out to get

$$Q(v+w) - Q(v) - Q(w) = f(v,w) + f(w,v) = 2f(v,w).$$

$\square$

Here is how to do polarisation in practice: any quadratic form $Q : \mathbb{F}^n \longrightarrow \mathbb{F}$ is of the form

$$Q(x) = \sum_{1\le i\le j\le n} q_{ij}x_ix_j = x^T \begin{pmatrix} q_{11} & & \frac{1}{2}q_{ji} \\ & \ddots & \\ \frac{1}{2}q_{ij} & & q_{nn} \end{pmatrix} x = x^{\mathrm{t}}Ax$$

so that the polarisation is f_A where

$$A_{ij} = A_{ji} = \begin{cases} q_{ii} & \text{if } i = j; \\ \frac{1}{2}q_{ij} & \text{if } i < j. \end{cases}$$

Example 5.2.5

Let $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ be given by

$$Q(x) = x_1^2 + 2x_2^2 + 2x_1x_2 + x_1x_3.$$

Let us find the polarisation f of Q, that is, we find A so that $f = f_A$: we have $q_{11} = 1$. $q_{22} = 2$, $q_{12} = 2$ and $q_{13} = 1$ with all other q_{ij} vanishing so

$$A = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 2 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

Definition 5.2.6 Rank and signature of quadratic forms

Let Q be a quadratic form on a finite-dimensional vector space V over $\mathbb{F}$.
The rank of Q is the rank of its polarisation.
If $\mathbb{F} = \mathbb{R}$, the *signature* of Q is the signature of its polarisation.

Theorem 5.2.7

Let Q be a quadratic form with rank r polarisation on a finite-dimensional vector space over $\mathbb{F}$.

(1) When $\mathbb{F} = \mathbb{C}$, there is a basis $\{v_1, \ldots, v_n\}$ of V such that

$$Q\Big(\sum_{i=1}^{n} x_i v_i\Big) = x_1^2 + \cdots + x_r^2.$$

(2) When $\mathbb{F} = \mathbb{R}$ and Q has signature (p, q), there is a basis $\{v_1, \ldots, v_n\}$ of V such that

$$Q\Big(\sum_{i=1}^{n} x_i v_i\Big) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2.$$

Example 5.2.8

Find the signature of $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ given by

$$Q(x) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 + 4x_2x_3.$$

Q has polarisation $f = f_A$ with

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Solution: exploit the zero in the $(1, 2)$-slot of A to see that $e_1, e_2, y = (-1, -2, 1)$ is a diagonalising basis and so gives us a diagonal matrix representing f with $Q(e_1) = Q(e_2) = 1 > 0$ and $Q(y) = -4 < 0$ along the diagonal. So the signature is $(2, 1)$.

Here are two alternative techniques:

(1) Orthogonal diagonalisation yields a diagonal matrix representing B with the eigenvalues of A down the diagonal so we just count how many positive and negative eigenvalues there are.

In fact, A has eigenvalues 1 and $1 \pm \sqrt{5}$. Since $\sqrt{5} > 2$, $1 - \sqrt{5} < 0$ and we again conclude that the signature is $(2, 1)$.

(2) Write Q as a linear combination of linearly independent squares and then count the number of positive and negative coefficients. In fact,

$$\begin{aligned} Q(x) &= x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 + 4x_2x_3 \\ &= (x_1 + x_3)^2 + x_2^2 + 4x_2x_3 = (x_1 + x_3)^2 + (x_2 + 2x_3)^2 - 4x_3^2. \end{aligned}$$

But now we need to check that $x_1 + x_3, x_2 + 2x_3, x_3$ are linearly independent linear functionals on $\mathbb{R}^3$. Here comes to the rescue and says we only need show that $(\ker x_1 + x_2) \cap (\ker x_2 + 2x_3) \cap (\ker x_3) = \{0\}$. But $x_3 = 0 = x_1 + x_3 = x_2 + 2x_3$ rapidly implies that each $x_i = 0$ and we are done. The coefficients of these squares are $1, 1, -4$ and so, once more, we get that the signature is $(2, 1)$.

Example 5.2.9

Determine the rank and signature of the quadratic form $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ given by

$$Q(x, y, z) = 2xy + 2yz.$$

by reducing it to its canonical form.

Solution:
Clearly

$$Q(x,y,z) = 2xy + 2yz = \frac{1}{2}\Big((x+y+z)^2 - (x-y+z)^2 \Big)$$

Hence, the matrix for the canonical form is

$$A = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So, the rank is 2. Further, the signature is (1,1) and $\operatorname{sgn}(Q) = 0$.

5.3 Hermitian forms

Definition 5.3.1 Hermitian form

Let V be a $\mathbb{C}$-vector space. A function $H : V \times V \longrightarrow \mathbb{C}$ is called a Hermitian form on V if

(HF1) for any $u, v, w \in V$ and $\lambda \in \mathbb{C}$, $H(u + \lambda v, w) = H(u, w) + \lambda H(v, w)$,

(HF2) for any $u, v \in V$, we have $H(u, v) = \overline{H(v, u)}$, **(Hermitian symmetric)**.

Where the bar denoting complex conjugation, that means if $z = a + b\mathrm{i}$ is a complex number ($\mathrm{i}^2 = -1$), then by definition $\bar{z} = a - b\mathrm{i}$.

Note 5.3.2

We denote the set of all Hermitian forms on V by $\operatorname{Herm}(V)$.

Example 5.3.3

(1) The function $H_1 : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ defined by

$$H_1(z, w) = \sum_{i=1}^{n} z_i \overline{w_i},$$

where $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ and $w = \begin{pmatrix} w_1 \\ \vdots \\ w_2 \end{pmatrix}$,
is a Hermitian form on $\mathbb{C}^n$.

(2) The function $H_2 : \mathbb{C}^2 \times \mathbb{C}^2 \longrightarrow \mathbb{C}$ defined by

$$H_2(z, w) = z_1 w_1 + \mathrm{i} z_2 w_1 - \mathrm{i} z_1 w_2.$$

is a Hermitian form on $\mathbb{C}^2$.

(3) The function $H_3 : \mathbb{C}^2 \times \mathbb{C}^2 \longrightarrow \mathbb{C}$ defined by

$$H_3(z, w) = z_1 w_1 + z_2 w_2$$

is not a Hermitian form on $\mathbb{C}^2$. Take for example $z = \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$H_3(z, w) = 1 + \mathrm{i}$$

and

$$\overline{H_3(w, z)} = \overline{1 + \mathrm{i}} = 1 - \mathrm{i}$$

So

$$H_3(z, w) \neq \overline{H_3(w, z)}$$

Definition 5.3.4 Hermitian matrix

Recall that a square matrix $A = (a_{ij})$ is called Hermitian matrix if $a_{ij} = \overline{a_{ji}}$ for all for all indices i and j,

Definition 5.3.5 Skew-Hermitian matrix

Recall that a square matrix $A = (a_{ij})$ is called skew-Hermitian matrix if $a_{ij} = -\overline{a_{ji}}$ for all for all indices i and j,

Note 5.3.6

The conjugate transpose of a matrix A is denoted by A^{h}

Remark 5.3.7.

(1) Let A be a complex square martix. Then

- **(a)** A Hermitian $\iff$ $A^{\mathsf{h}} = A$.
- **(b)** A skew-Hermitian $\iff$ $A^{\mathsf{h}} = -A$

(2) If A is a real matrix, then

- **(a)** A Hermitian $\iff$ A is symmetric (i.e. $A^{\mathsf{t}} = A$).
- **(b)** A skew-Hermitian $\iff$ A is skew symmetric (i.e. $A^{\mathsf{t}} = -A$).

Example 5.3.8

The following matrix A is Hermitian

$$A = \begin{pmatrix} 1 & 1-\mathrm{i} & 2-3\mathrm{i} \\ 1+\mathrm{i} & 4 & 2\mathrm{i} \\ 2+3\mathrm{i} & -2\mathrm{i} & 0 \end{pmatrix}$$

The following matrix B is skew-Hermitian

$$B = \begin{pmatrix} -\mathrm{i} & 2+\mathrm{i} \\ -2+\mathrm{i} & 0 \end{pmatrix}$$

because

$$A^{\mathsf{h}} = \begin{pmatrix} \overline{-\mathrm{i}} & \overline{2+\mathrm{i}} \\ \overline{-2+\mathrm{i}} & \overline{0} \end{pmatrix}^{\mathsf{t}} = \begin{pmatrix} \mathrm{i} & 2-\mathrm{i} \\ -2-\mathrm{i} & 0 \end{pmatrix}^{\mathsf{t}} = \begin{pmatrix} i & -2-\mathrm{i} \\ 2-\mathrm{i} & 0 \end{pmatrix} = -A$$

Proposition 5.3.9 Hermitian properties of matrices

Let $A, B \in \mathcal{M}_n(\mathbb{C})$ and $\lambda \in \mathbb{C}$. Then

(1) $(A+B)^{\mathsf{h}} = A^{\mathsf{h}} + B^{\mathsf{h}}$

(2) $(\lambda A)^{\mathsf{h}} = \overline{\lambda} A^{\mathsf{h}}$.

(3) $(AB)^{\mathsf{h}} = B^{\mathsf{h}} A^{\mathsf{h}}$.

(4) $(A^{\mathsf{h}})^{\mathsf{h}} = A$.

(5) If A is invertible, we have $(A^{\mathsf{h}})^{-1} = (A^{-1})^{\mathsf{h}}$.

Proposition 5.3.10 Hermitian properties of matrices

(1) The sum of two Hermitian matrices is Hermitian.

(2) The inverse of an invertible Hermitian matrix is Hermitian as well.

(3) The sum of a square matrix and its conjugate transpose $\left(A + A^{\mathsf{h}}\right)$ is Hermitian.

(4) The difference of a square matrix and its conjugate transpose $\left(A - A^{\mathsf{h}}\right)$ is skew-Hermitian.

(5) The product of two Hermitian matrices A and B is Hermitian if and only if $AB = BA$.

(6) if A and B are Hermitian, then ABA is Hermitian.

Definition 5.3.11

Let V be a vector space over $\mathbb{C}$ with basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ and let $H : V \times V \longrightarrow \mathbb{F}$ be a Hermitian form.

We define $[H]_{\mathcal{B}}$ the matrix of H with respect to $\mathcal{B}$ by $[H]_{\mathcal{B}} = (a_{ij}) \in M_{n\times n}(\mathbb{F})$ given by

$$a_{ij} = H(v_i, v_j), \quad \text{for} \quad 1 \le i, j \le n.$$

The Hermitian symmetric property of a Hermitian form implies that

$$[H]_{\mathcal{B}} = \overline{[H]_{\mathcal{B}}}^{\mathsf{t}}.$$

Remark 5.3.12. f V is a vector space over $\mathbb{C}$ with basis $\mathcal{B} = \{v_1, \dots, v_n\}$ and $H : V \times V \longrightarrow \mathbb{C}$ a Hermitian form with matrix $A = [H]_{\mathcal{B}}$ with respect to $\mathcal{B}$. Then H is completely determined by A: if $v = \sum_{i=1}^{n} x_i v_i$ and $w = \sum_{j=1}^{n} y_j v_j$ then

$$H(v, w) \sum_{i,j=1}^{n} x_i \overline{y_j} H(v_i, v_j) = \sum_{i,j=1}^{n} x_i \overline{y_j} a_{ij} = \sum_{i,j=1}^{n} \overline{y_j} a_{ij} x_i = x^{\mathsf{t}} A \overline{y} = y^{\mathsf{h}} A x.$$

Lemma 5.3.13

Let $H \in \operatorname{Herm}(V)$ and $\mathcal{B} = \{v_1, \dots, v_n\}$ an ordered basis of V. Then, for any $u, v \in V$, we have

$$H(u, v) = [u]_{\mathcal{B}}^{\mathsf{t}} \, [H]_{\mathcal{B}} \, \overline{[v]_{\mathcal{B}}} = [v]_{\mathcal{B}}^{\mathsf{h}} \, [H]_{\mathcal{B}} \, \overline{[u]_{\mathcal{B}}}.$$

Moreover, if $A \in \mathcal{M}_n(\mathbb{C})$ is such that

$$[u]_{\mathcal{B}}^{\mathsf{t}} \, A \, \overline{[v]_{\mathcal{B}}} = H(u, v),$$

then $A = [H]_{\mathcal{B}}$.

Proposition 5.3.14 Hermitian form and chage of basis

Let $H \in \operatorname{Herm}(V)$, $\mathcal{B}, \mathcal{B}'$ two ordered bases of V. If $P = P_{\mathcal{B} \longrightarrow \mathcal{B}'}$ is the change of coordinate matrix from $\mathcal{B}$ to $\mathcal{B}'$, then

$$P^{\mathsf{h}} [H]_{\mathcal{B}'} P = [H]_{\mathcal{B}}.$$

Proof. Let $u, v \in V$, and $P = P_{\mathcal{B}' \longrightarrow \mathcal{B}}$. We know that

$$[u]_{\mathcal{B}} = P[u]_{\mathcal{B}'} \quad \text{and} \quad [v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$$

We have:

$$\begin{aligned} H(u, v) &= [v]_{\mathcal{B}}^{\mathsf{h}} [H]_{\mathcal{B}} [u]_{\mathcal{B}} \\ &= \left(P[v]_{\mathcal{B}'}\right)^{\mathsf{h}} [H]_{\mathcal{B}} P[u]_{\mathcal{B}'} \\ &= [v]_{\mathcal{B}'}^{\mathsf{h}} P^{\mathsf{h}} [H]_{\mathcal{B}} P[u]_{\mathcal{B}'} \end{aligned}$$

Therefore

$$P^{\mathsf{h}} \; [H]_{\mathcal{B}} \; P \; = \; [H]_{\mathcal{B}'},$$

□

5.4 Classification of Hermitian forms

Definition 5.4.1 Nondegenerate Hermitian form

Let $H \in \mathrm{Herm}(V)$. We say that H is nondegenerate if $[H]_{\mathcal{B}}$ is invertible, for any basis $\mathcal{B}$ of V. The previous lemma ensures that this notion of nondegeneracy is well-defined (ie, does not depend on the choice of basis $\mathcal{B}$).

Theorem 5.4.2 Classification of Hermitian forms

Let V be a $\mathbb{C}$-vector space, $n = \dim V$ and $H \in \mathrm{Herm}(V)$ be nondegenerate Hermetian form on V. Then, there is an ordered basis $\mathcal{B}$ of V such that

$$[H]_{\mathcal{B}} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

where $d_i \in \{1, -1\}$.

Corollary 5.4.3

If $u, v \in V$ with

$$[u]_{\mathcal{B}} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad \text{and} \quad [v]_{\mathcal{B}} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

, we have

$$H(u, v) = \sum_{i=1}^{n} d_i u_i \overline{v_i}.$$

Definition 5.4.4 Sesquilinear form

Let V be a $\mathbb{C}$-vector space. A function $H : V \times V \longrightarrow \mathbb{C}$ is called a sesquilinear form on V if for any $u, v, w \in V$ and $\lambda \in \mathbb{C}$, we have

(1) $H(u + \lambda v, w) = H(u, w) + \lambda H(v, w)$,

(2) $H(w, u + \lambda v) = H(w, u) + \overline{\lambda} H(w, v)$,

Definition 5.4.5 Positive and positive definite of Hermitian forms

Given a complex vector space V, a Hermitian form $H \in \mathrm{Herm}(V)$ is called

(1) Positive if $H(u, u) \geq 0$ for all $u \in V$

(2) Positive definite if $H(u, u) > 0$ for all $u \in V$

Definition 5.4.6 Hermitian space (or unitary space)

A pair $\langle V, H \rangle$ is called Hermitian space, where V is a $\mathbb{C}$-vector space and H is a Hermitian form on V such that $[H]_B = I_n$, for some basis B.

Definition 5.4.7 Positive and positive definite of Hermitian forms

A Hermitian space $\langle V, H \rangle$ is called:

(1) pre-Hilbert space if H is positive.

(2) Hilbert space if H is positive definite.

5.5 Exercises Set

Exercise 5.5.1

Let B be a nonzero real symmetric bilinear form on V with quadratic form Q. Show that $Q(v) \neq 0$ for some $v \in V$.

Solution. For all $u, v \in V$, we have

$$Q(u+v) = Q(u) + 2B(u,v) + Q(v).$$

If $Q(v) = 0$ for all $v \in V$, we get from previous equality

$$0 = 0 + 2B(u,v) + 0 \quad \text{for all} \quad u, v \in V.$$

Therefore $B = 0$.

$$0 = Q(u+v) = Q(u) + 2B(u,v) + Q(v) = 2B(2,v).$$

Which is a contradiction. So $Q(v) \neq 0$ for some $v \in V$.

Exercise 5.5.2

Let V be a complex vector space, and $H : V \times V \longrightarrow \mathbb{C}$ a nonzero Hermitian form on V. Let $v_1, v_2 \in V$ such that $c = H(v_1, v_2) \neq 0$. Let $v_3 = c\, v_2$.

(1) Show that $H(v_1, v_3)$ is a nonzero real number.

(2) Show that, there exists $v \in V$ such that $H(v,v) \neq 0$.

Solution.

(1) $H(v_1, v_3) = H(v_1, cv_2) = \bar{c}H(v_1, v_2) = \bar{c}c \in \mathbb{R} \setminus \{0\}$.

(2) We have

$$\begin{aligned} H(v_1+v_3, v_1+v_3) &= H(v_1,v_1) + H(v_1,v_3) + H(v_3,v_1) + H(v_3,v_3) \\ &= H(v_1,v_1) + H(v_1,v_3) + \overline{H(v_1,v_3)} + H(v_3,v_3) \\ &= H(v_1,v_1) + 2H(v_1,v_3) + H(v_3,v_3) \quad \text{Because } H(v_1,v_3) \text{ is a real number.} \end{aligned}$$

So

$$H(v_1+v_3, v_1+v_3) = H(v_1,v_1) + 2H(v_1,v_3) + H(v_3,v_3)$$

Since the term $2H(v_1,v_3)$ isn't zero, at least one of the three other terms in the last equation isn't zero.

Exercise 5.5.3

Let $B \in \mathrm{Bil}_{\mathbb{R}}(V)$ be a real nondegenerate symmetric bilinear form with quadratic form Q such that $Q(u) = 0$ for some nonzero vector $u \in V$.

(1) Show that there exists $w \in V$ such that $B(u,w) = \frac{1}{2}$.

(2) Show that, for all $x \in \mathbb{R}$, $Q(xu + w) = x + Q(w)$.

(3) Deduce that $Q(V) = \mathbb{R}$.

Solution. **(1)** Let u be a nonzero vector in V such that $Q(u) \neq 0$. As B is nondegenerate, there exists $w_1 \in V$ such that $B(u, w_1) \neq 0$. Let

$$w = \frac{1}{2B(u, w_1)} w_1$$

So

$$B(u, w) = B\Big(u, \frac{1}{2B(u, w_1)} w_1\Big) = \frac{1}{2B(u, w_1)} B(u, w_1) = \frac{1}{2}.$$

(2) For all $x \in \mathbb{R}$,

$$\begin{aligned} Q(xu + w) &= B(xu + w, xu + w) \\ &= x^2 Q(u) + 2xB(u, w) + Q(w) \\ &= x + Q(w). \end{aligned}$$

(3) W deduce from **(2)** : for all $x \in \mathbb{R}$,

$$x = Q\Big((x - Q(w))u \,+\, w\Big).$$

Hence $Q(V) = \mathbb{R}$.

Exercise 5.5.4 Hermitian form is anti-linear in the second argument

Show that, if H is a Hermitian form on V, then

$$H(u, v + bw) = H(u, v) + \bar{b} H(v, w),$$

any $u, v, w \in V$ and $b \in \mathbb{C}$.

Solution. By definition, we know that for all $u, v \in V$ and $b \in \mathbb{F}$, we have

$$\begin{aligned} H(u, v + bw) &= \overline{H(v + bw, u)} \\ &= \overline{H(v, u) + bH(w, u)} \\ &= \ \overline{H(v, u)} + \bar{b}\, \overline{H(w, u)} \\ &= H(u, v) + \bar{b} H(u, w). \end{aligned}$$

Exercise 5.5.5

Show that the determinant of a Hermitian matrix is a real number.

Solution. If A is Hermitian, then $A = \overline{A^t}$, so $\det(A) = \det \overline{A^t}$. Therefore $\det A = \overline{\det A^t}$. Since $\det A = \det A^t$, we get $\det A = \overline{\det A}$. Hence $\det A$ is a real number.

Exercise 5.5.6

Let $H \in \mathrm{Herm}(V)$. Show that for all $u, v \in V$ and $\alpha, \beta \in C$, we have

$$H(\alpha u + \beta v, \alpha u + \beta v) = |\alpha|^2 H(u, u) + 2\Re(\alpha \overline{\beta} H(u, v)) + |\beta|^2 H(v, v).$$

Solution.

$$\begin{aligned}
H(\alpha u+\beta v,\alpha u+\beta v) &= \alpha H(u,\alpha u+\beta v)+\beta H(v,\alpha u+\beta v)\\
&= \alpha H(u,\alpha u)+\alpha H(u,\beta v)+\beta H(v,\alpha u)+\beta H(v,\beta v)\\
&= \alpha\overline{\alpha}H(u,u)+\alpha\overline{\beta}H(u,v)+\beta\overline{\alpha}H(v,u)+\beta\overline{\beta}H(v,v)\\
&= \alpha\overline{\alpha}H(u,u)+\alpha\overline{\beta}H(u,v)+\beta\overline{\alpha}\overline{H(u,v)}+\beta\overline{\beta}H(v,v)\\
&= |\alpha|^2H(u,u)+2\Re(\alpha\overline{\beta}H(u,v))+|\beta|^2H(v,v).
\end{aligned}$$

Exercise 5.5.7 First polarization identities for sesquilinear form

Show that for any sesquilinear form $H : V\times V\longrightarrow\mathbb{C}$, we have

$$4H(u,v)=H(u+v,u+v)-H(u-v,u-v)+\mathrm{i}H(u+\mathrm{i}v,u+\mathrm{i}v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v),$$

Solution. Let Φ be the quadratic form associated with H:

$$\Phi(x)=H(x,x)\quad\text{for all } x\in V.$$

For any $\alpha,\beta\in\mathbb{C}$, we have

$$\begin{aligned}
\Phi(\alpha x+\beta y) &= H(\alpha x+\beta y,\alpha x+\beta y)\\
&= |\alpha|^2\Phi(x)+\alpha\bar{\beta}H(x,y)+\bar{\alpha}\beta H(y,x)+|\beta|^2\Phi(y).
\end{aligned}$$

Using this equality subsequently for $\alpha=\beta=1; \alpha=1$ and $\beta=-1; \alpha=1$ and $\beta=i; \alpha=1$ and $\beta=-i$; we get

$$\begin{aligned}
\Phi(x+y) &= \Phi(x)+H(x,y)+H(y,x)+\Phi(y)\\
-\Phi(x-y) &= -\Phi(x)+H(x,y)+H(y,x)-\Phi(y)\\
i\Phi(x+iy) &= i\Phi(x)+H(x,y)-H(y,x)+i\Phi(y)\\
-i\Phi(x-iy) &= -i\Phi(x)+H(x,y)-H(y,x)-i\Phi(y).
\end{aligned}$$

By adding these equalities we obtain:

$$4H(u,v)=H(u+v,u+v)-H(u-v,u-v)+\mathrm{i}H(u+\mathrm{i}v,u+\mathrm{i}v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v),$$

Exercise 5.5.8 Second polarization identities for sesquilinear form

Show that for any sesquilinear form $H : V\times V\longrightarrow\mathbb{C}$, we have

$$2H(u,v)=(1+\mathrm{i})\big(H(u,u)+H(v,v)\big)-H(u-v,u-v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v).$$

Solution. From Exercise 5.5.7 :

$$4H(u,v)=H(u+v,u+v)-H(u-v,u-v)+\mathrm{i}H(u+\mathrm{i}v,u+\mathrm{i}v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v).$$

Then

$$\begin{aligned}
4H(u,v) &= H(u,u)+H(u,v)+H(v,u)+H(v,v)-H(u-v,u-v)\\
&\quad +\mathrm{i}H(u,u)+\mathrm{i}H(u,\mathrm{i}v)+\mathrm{i}H(\mathrm{i}v,u)+\mathrm{i}H(\mathrm{i}v,\mathrm{i}v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v)\\
&= H(u,u)+H(u,v)+\cancel{H(v,u)}+H(v,v)-H(u-v,u-v)\\
&\quad +\mathrm{i}H(u,u)+H(u,v)-\cancel{H(v,u)}+\mathrm{i}H(v,v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v).
\end{aligned}$$

Hence

$$2H(u,v)=(1+\mathrm{i})\big(H(u,u)+H(v,v)\big)-H(u-v,u-v)-\mathrm{i}H(u-\mathrm{i}v,u-\mathrm{i}v).$$

Exercise 5.5.9

Show that a sesquilinear form $H : V \times V \longrightarrow \mathbb{C}$ is Hermitian if and only if $H(v,v) \in \mathbb{R}$ for all $v \in V$.

Solution. Clearly, if H is Hermitian, then for all $v \in V$, $H(v,v) = \overline{H(v,v)}$. Therefore $H(v,v) \in \mathbb{R}$.

Conversely, suppose that H is a sesquilinear form such that $H(v,v) \in \mathbb{R}$ for all $v \in V$. To prove that H is Hermitian, we need only to

$$H(u+v,u+v) = H(u,u) + H(u,v) + H(v,u) + H(v,v)$$

and

$$H(u-v,u-v) = H(u,u) - H(u,v) - H(v,u) + H(v,v)$$

So

$$H(u,v) + H(v,u) = a \in \mathbb{R}. \tag{5.1}$$

Also

$$H(\mathrm{i}u,v) + H(v,\mathrm{i}u) = b \in \mathbb{R}$$

that is

$$\mathrm{i}H(u,v) - \mathrm{i}H(v,u) = b$$

Multiplying by i,

$$H(u,v) - H(v,u) = \mathrm{i}b \tag{5.2}$$

From (5.1) and (5.2), we get

$$H(u,v) = \frac{\alpha + i\beta}{2}$$

and

$$H(v,u) = \frac{\alpha - i\beta}{2},$$

which means $H(u,v) = \overline{H(v,u)}$, for any $u,v \in V$, as required.

Exercise 5.5.10

Let $\langle V, H \rangle$ be a Hermitian space. Show that for any linear map $f \in \mathcal{L}(V)$ such that $H(f(v),v) = 0$ for all $v \in V$, then $f = 0$.

Solution. We have, for all $u,v \in V$ and $\alpha \in \mathbb{C}$

$$\begin{aligned} 0 = H(f(u+\alpha v), u+\alpha v) &= H(f(u)+\alpha f(v), u+\alpha v) \\ &= \overline{\alpha} H(f(u),v) + \alpha H(f(v),u). \end{aligned}$$

In particular, when $\alpha = 1$ or $\alpha = \mathrm{i}$, we get

$$H(f(u),v)) + H(f(v),u) = 0$$

and

$$iH(f(u),v)) - iH(f(v),u) = 0$$

Therefore

$$H(f(u),v)) = 0 \quad \text{for all } u,v \in V$$

Since H is nondegenerate, $f(u) = 0$ for all $u \in V$.

Bibliography

[1] Basic Algebra. Along with a Companion Volume Advanced Algebra. Digital Second Edition, 2016.

[2] Basic Algebra. Anthony W. Knapp. Along with a Companion Volume Advanced Algebra. Digital Second Edition, (2016).

[3] Multilinear algebra D. G. Northcott. Cambridge University press (2008).

[4] Advanced Linear Algebra, Steven Roman, Graduate Texts in Mathematics, Springer (2007).

[5] Linear Algebra . Kenneth Hoffman. Prentice-HaLL, INc., Englewood Cliffs, New Jersey.