

Chapter 1

Elementary Linear and Multilinear Algebra

This chapter is the only one where results are given either without proof, or with sketchy proofs. A beginner should have a close look at a textbook dedicated to linear algebra, not only reading statements and proofs, but also solving exercises in order to become familiar with all the relevant notions.

1.1 Vectors and Scalars

Scalars are elements of some field k (or K), or sometimes of a ring R . The most common fields are the field of rational numbers \mathbb{Q} , the field of real numbers \mathbb{R} , and the field of complex numbers \mathbb{C} . There are also finite fields, such as $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ (p a prime number). Other interesting fields are $k(X)$ (rational fractions), that of formal Laurent series, or the p -adic field \mathbb{Q}_p . Linear algebra also makes use of the ring of integers \mathbb{Z} or of those of polynomials in one or in several variables $k[X]$ and $k[X_1, \dots, X_r]$. One encounters a lot of other rings in number theory and algebraic geometry, for instance, the Gaussian integers $\mathbb{Z}[i]$.

Inasmuch this book is about matrices, we show that square matrices form a non-commutative ring; this ring can be used as a set of scalars, when we write a large matrix blockwise. This is one of the few instances where the ring of scalars is not Abelian. Another one occurs in Section 4.4.

The digits 0 and 1 have the usual meaning in a field K , with $0 + x = 1 \cdot x = x$. The subring $\mathbb{Z} \cdot 1$, composed of all sums (possibly empty) of the form $\pm(1 + \dots + 1)$ is isomorphic to either \mathbb{Z} or a finite field \mathbb{F}_p . In the latter case, p is a prime number, which we call the *characteristic* of K and denote $\text{char}(K)$. In the former case, we set $\text{char}(K) = 0$.

One says that a nonzero polynomial $P \in K[X]$ *splits* over K if it can be written as a product of the form

$$a \prod_{i=1}^r (X - a_i)^{n_i}, \quad a, a_i \in K, \quad r \in \mathbb{N}, n_i \in \mathbb{N}^*.$$

We may assume that the a_i s are pairwise distinct. Such a factorization is then unique, up to the order of the factors. A field K in which every nonconstant polynomial $P \in K[X]$ admits a root, or equivalently in which every polynomial $P \in K[X]$ splits, is *algebraically closed*. If the field K' contains the field K and if every polynomial $P \in K[X]$ splits in K' , then the set \bar{K} of roots in K' of polynomials in $K[X]$ is an algebraically closed field containing K , and it is the smallest such field, unique up to isomorphism. One calls \bar{K} the *algebraic closure* of K . Every field K admits an algebraic closure, unique up to isomorphism. The fundamental theorem of algebra asserts that $\bar{\mathbb{R}} = \mathbb{C}$. The algebraic closure of \mathbb{Q} , for instance, is the set of *algebraic numbers*; it is the set of complex roots of all polynomials $P \in \mathbb{Z}[X]$.

1.1.1 Vector Spaces

Let K be a field and $(E, +)$ be a commutative group. Because E and K are distinct sets, the symbol $+$ has two meanings, depending on whether it is used for the addition in E or in K . This does not cause any confusion. Let moreover

$$\begin{aligned} (a, x) &\mapsto ax, \\ K \times E &\rightarrow E, \end{aligned}$$

be a map such that

$$(a+b)x = ax + bx, \quad a(x+y) = ax + ay, \quad a(bx) = (ab)x, \quad 1x = x$$

for every $x, y \in E$ and $a, b \in K$. We say that E is a *vector space* over K (or a K -vector space). The elements of E are called *vectors*. In a vector space one always has $0x = 0$ (more precisely, $0_K x = 0_E$).

When $P, Q \subset K$ and $F, G \subset E$, one denotes by PQ (respectively, $P+Q, F+G, PF$) the set of products pq as (p, q) ranges over $P \times Q$ (respectively, $p+q, f+g, pf$ as p, q, f, g range over P, Q, F, G). A subgroup $(F, +)$ of $(E, +)$, which is stable under multiplication by scalars (i.e., such that $KF \subset F$), is again a K -vector space. One says that it is a *linear subspace* of E , or just a subspace. Observe that F , as a subgroup, is nonempty, because it contains 0_E . The intersection of any family of linear subspaces is a linear subspace. The sum $F+G$ of two linear subspaces is again a linear subspace. The trivial formula $(F+G)+H = F+(G+H)$ allows us to define unambiguously $F+G+H$ and, by induction, the sum of any finite family of subsets of E . When these subsets are linear subspaces, their sum is also a linear subspace.

Let I be a set. One denotes by K^I the set of maps $a = (a_i)_{i \in I} : I \rightarrow K$ where only finitely many of the a_i s are nonzero. This set is naturally endowed with a K -vector space structure, with the addition and product laws

$$(a+b)_i := a_i + b_i, \quad (\lambda a)_i := \lambda a_i.$$

Let E be a vector space and let $i \mapsto f_i$ be a map from I to E . A *linear combination* of $(f_i)_{i \in I}$ is a sum

$$\sum_{i \in I} a_i f_i,$$

where the a_i s are scalars, only finitely many of them being nonzero (in other words, $(a_i)_{i \in I} \in K^I$). This sum involves only finitely many nonzero terms, thus makes sense. It is a vector of E . The family $(f_i)_{i \in I}$ is *free*, or *linearly independent*, if every linear combination but the trivial one (when all coefficients are zero) is nonzero. It is a *generating* family if every vector of E is a linear combination of its elements. In other words, $(f_i)_{i \in I}$ is free (respectively, generating) if the map

$$\begin{aligned} K^I &\rightarrow E, \\ (a_i)_{i \in I} &\mapsto \sum_{i \in I} a_i f_i, \end{aligned}$$

is injective (respectively, onto). Finally, one says that $(f_i)_{i \in I}$ is a *basis* of E if it is both free and generating. In that case, the above map is bijective, and it is actually an isomorphism between vector spaces.

If $\mathcal{G} \subset E$, one often identifies \mathcal{G} and the associated family $(g)_{g \in \mathcal{G}}$. The set G of linear combinations of elements of \mathcal{G} is a linear subspace E , called the linear subspace *spanned* by \mathcal{G} . It is the smallest linear subspace E containing \mathcal{G} , equal to the intersection of all linear subspaces containing \mathcal{G} . The subset \mathcal{G} is generating when $G = E$.

1.1.1.1 Dimension of a Vector Space

Let us mention an abstract result.

Theorem 1.1 *Every K -vector space admits at least one basis. Every free family is contained in a basis. Every generating family contains a basis. All the bases of E have the same cardinality (which is called the dimension of E).*

For general vector spaces, this statement is a consequence of the axiom of choice. As such, it is overwhelmingly (but not universally) accepted by mathematicians. Because we are interested throughout this book in finite-dimensional spaces, for which the existence of bases follows from elementary considerations, we prefer to start with the following.

Definition 1.1 *The dimension of a vector space E , denoted by $\dim E$, is the upper bound of the cardinality of free families in E . It may be infinite. If $E = \{0\}$, the dimension is zero.*

If $\dim E < +\infty$, we say that E is finite-dimensional.

When E is finite-dimensional, every free family of cardinal $\dim E$ is contained in a free family that is maximal for the inclusion (obvious); the maximality implies that the latter family is generating, hence is a basis. Next, given a generating family β in E , one may consider free families contained in β . Again, a maximal one is a basis.

Thus β contains a basis. The fact that two bases have the same cardinality is less easy, but still elementary.

The dimension is monotone with respect to inclusion: if $F \subset E$, then $\dim F \leq \dim E$. The equality case is useful.

Proposition 1.1 *Let E be a finite-dimensional vector space, and F be a linear subspace of E .*

If $\dim F = \dim E$, then $F = E$.

Proposition 1.2 *If F, G are two linear subspaces of E , the following formula holds,*

$$\dim F + \dim G = \dim F \cap G + \dim(F + G).$$

Corollary 1.1 *In particular,*

$$\dim F \cap G \geq \dim F + \dim G - \dim E.$$

If $F \cap G = \{0\}$, one writes $F \oplus G$ instead of $F + G$, and one says that the sum of F and G is *direct*. Proposition 1.2 gives

$$\dim F \oplus G = \dim F + \dim G.$$

Let E and F be vector spaces over K . One builds the abstract direct sum of E and F as follows, and one denotes it again $E \oplus F$. Its vectors are those of the Cartesian product $E \times F$, whereas the sum and the multiplication by a scalar are defined by

$$(e, f) + (e', f') = (e + e', f + f'), \quad \lambda(e, f) = (\lambda e, \lambda f).$$

The spaces E and F can be identified with the subspaces $E \times \{0_F\}$ and $\{0_E\} \times F$ of $E \oplus F$, respectively.

Given a set I , the family $(\mathbf{e}^i)_{i \in I}$, defined by

$$(\mathbf{e}^i)_j = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases}$$

is a basis of K^I , called the *canonical basis*. The dimension of K^I is therefore equal to the cardinality of I .

1.1.1.2 Extension of the Scalars

Let L be a field and K a subfield of L . If F is an L -vector space, then F is also a K -vector space. As a matter of fact, L is itself a K -vector space, and one has

$$\dim_K F = \dim_L F \cdot \dim_K L.$$

The most common example (the only one that we consider) is $K = \mathbb{R}$, $L = \mathbb{C}$, for which we have

$$\dim_{\mathbb{R}} F = 2 \dim_{\mathbb{C}} F.$$

Conversely, if G is an \mathbb{R} -vector space, one builds its *complexification* $G \otimes_{\mathbb{R}} \mathbb{C}$ (read G tensor \mathbb{C}) as follows:

$$G \otimes_{\mathbb{R}} \mathbb{C} = G \times G,$$

with the induced structure of the additive group. An element (x, y) of $G \otimes_{\mathbb{R}} \mathbb{C}$ is also denoted $x + iy$. One defines multiplication by a complex number by

$$(\lambda = a + ib, z = x + iy) \mapsto \lambda z := (ax - by, ay + bx).$$

One verifies easily that $G \otimes_{\mathbb{R}} \mathbb{C}$ is a \mathbb{C} -vector space, with

$$\dim_{\mathbb{C}} G \otimes_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{R}} G.$$

Furthermore, G may be identified with an \mathbb{R} -linear subspace of $G \otimes_{\mathbb{R}} \mathbb{C}$ by

$$x \mapsto (x, 0).$$

Under this identification, one has $G \otimes_{\mathbb{R}} \mathbb{C} = G \oplus iG$. In a more general setting, one may consider two fields K and L with $K \subset L$, instead of \mathbb{R} and \mathbb{C} . The extension of scalars from K to L yields the space $G \otimes_K L$, a *tensor product*. We construct the tensor product of arbitrary vector spaces in Section 4.1.

1.2 Linear Maps

Let E, F be two finite-dimensional K -vector spaces. A map $u : E \rightarrow F$ is *linear* (one also speaks of a *homomorphism*) if $u(x + y) = u(x) + u(y)$ and $u(ax) = au(x)$ for every $x, y \in E$ and $a \in K$. One then has $u(0_E) = 0_F$. The preimage $u^{-1}(0_F)$, denoted by $\ker u$, is the *kernel* of u . It is a linear subspace of E . The *range* $u(E)$ is also a linear subspace of F , whose dimension is called the *rank* of u , and denoted by $\operatorname{rk} u$. It is often denoted $R(u)$. Taking a basis $(u(x_i))_{i \in I}$ of $u(E)$, together with a basis $(y_j)_{j \in J}$ of $\ker u$, the x s and the y s form a basis of E , hence comes the following.

Theorem 1.2 *If $u : E \rightarrow F$ is linear, then*

$$\dim E = \dim \ker u + \operatorname{rk} u.$$

The set of homomorphisms from E to F is naturally a K -vector space, with operations

$$(u + v)(x) = u(x) + v(x), \quad (\lambda u)(x) = \lambda u(x).$$

It is denoted $\mathcal{L}(E; F)$. Its dimension equals the product of $\dim E$ and $\dim F$.

If $u \in \mathcal{L}(E; F)$ and $v \in \mathcal{L}(F; G)$ are given, the composition $v \circ u$ is well defined and is linear: $v \circ u \in \mathcal{L}(E; G)$.

1.2.1 Eigenvalues, Eigenvectors

If $F = E$, one denotes $\text{End}(E) := \mathcal{L}(E; E)$; its elements are the *endomorphisms* of E . Therefore $\text{End}(E)$ is an algebra, that is a ring under the laws $(+, \circ)$, a K -vector space, with the additional property that $\lambda(u \circ v) = (\lambda u) \circ v$.

For an endomorphism, Theorem 1.2 reads $n = \dim \ker u + \text{rk } u$. A subspace of E of dimension n equals E itself, therefore we infer the equivalence

$$u \text{ is bijective} \iff u \text{ is injective} \iff u \text{ is surjective.}$$

Replacing u by $\lambda \text{id}_E - u$, we thus have

$$\lambda \text{id}_E - u \text{ is bijective} \iff \lambda \text{id}_E - u \text{ is injective} \iff \lambda \text{id}_E - u \text{ is surjective.}$$

In other words, we face the alternative

- Either there exists $x \in E \setminus \{0\}$ such that $u(x) = \lambda x$. Such a vector is called an *eigenvector* and λ is called an *eigenvalue*,
- Or, for every $b \in E$, the following equation admits a unique solution $y \in E$,

$$u(y) - \lambda y = b.$$

The set of eigenvalues of u is denoted $\text{Sp}(u)$. We show later on that it is a finite set. Notice that an eigenvector is always a nonzero vector.

1.2.2 Linear Forms and Duality

When the target space is the field of scalars, a linear map (i.e., $u : E \rightarrow K$) is called a *linear form*. The set of linear forms is the *dual space* of E , denoted by E' :

$$E' = \mathcal{L}(E; K).$$

The dimension of E equals that of E' . If $\mathcal{B} = \{v^1, \dots, v^n\}$ is a basis of E , then the *dual basis* of E' is $\{\ell_1, \dots, \ell_n\}$ defined by

$$\ell_i(v^j) := \begin{cases} 1, & \text{if } j = i, \\ 0, & \text{if } j \neq i. \end{cases}$$

The ℓ_i s are the coordinate maps over in E in the basis \mathcal{B} , inasmuch as we have

$$x = \sum_{i=1}^n \ell_i(x) v^i, \quad \forall x \in E.$$

In other words, the identity map id_E decomposes as

$$\text{id}_E = \sum_{i=1}^n v^i \ell_i.$$

In the above equality, the ℓ_i s play the role of functions, and the v^i s can be viewed as coefficients. The image of x under id_E is a vector (here, itself), therefore these coefficients are vectors. The same comment applies below.

Every linear map $u : E \rightarrow F$ decomposes as a finite sum $w^1 m_1 + \cdots + w^r m_r$, where the w^j s are vectors of F and the m_i s are linear forms on E . In other words,

$$u(x) = \sum_{i=1}^r m_i(x) w^i, \quad \forall x \in E, \quad \left(\text{equivalently } u = \sum_{i=1}^r w^i m_i \right).$$

This decomposition is highly non unique. The minimal number r in such a sum equals the rank of u . In terms of the tensor product introduced in Section 4.1, we identify $\mathcal{L}(E; F)$ with $F \otimes E'$ and this decomposition reads

$$u = \sum_{i=1}^r w^i \otimes m_i.$$

1.2.2.1 Bidual

The *bidual* of E is the dual space of E' . A vector space E can be identified canonically with a subspace of its bi-dual: given $x \in E$, one defines a linear form over E' by

$$\ell \mapsto \delta_x \ell(x).$$

The map $\delta : x \mapsto \delta_x$ is linear and one-to-one from E to $(E')'$. These spaces have the same dimension and thus δ is an isomorphism. Because it is canonically defined, we identify E with its bi-dual.

1.2.2.2 Polarity

Given a subset S of E , the set of linear forms vanishing identically over S is the *polar set* of S , denoted by S^0 . The following properties are obvious:

- A polar set is a linear subspace of E' .
- If $S \subset T$, then $T^0 \subset S^0$.
- $S^0 = (\text{Span}(S))^0$.

Proposition 1.3 *If F is a subspace of E , then*

$$\dim F + \dim F^0 = \dim E.$$

Proof. Let $\{v^1, \dots, v^r\}$ be a basis of F , which we extend as a basis \mathcal{B} of E . Let $\{\ell^1, \dots, \ell^n\}$ be the dual basis of E' . Then F^0 equals $\text{Span}(\ell^{r+1}, \dots, \ell^n)$ has dimension $n - r$. \square

Corollary 1.2 *A subspace F of E equals its bipolar $(F^0)^0$. Rigorously speaking, $(F^0)^0 = \delta(F)$.*

The bipolar of a subset $S \subset E$ equals $\text{Span}(S)$.

Proof. Obviously, $(F^0)^0 \subset F$. In addition, their dimensions are equal to $n - \dim F^0$ because of Proposition 1.3. Therefore they coincide.

If $S \subset E$, we know that $S^0 = (\text{Span}(S))^0$. Therefore

$$(S^0)^0 = ((\text{Span}(S))^0)^0 = \text{Span}(S).$$

□

1.2.2.3 Adjoint Linear Map

Let $u \in \mathcal{L}(E; F)$ be given. If ℓ is a linear form over F , then $\ell \circ u$ is a linear form over E , thus an element of E' . The map $\ell \mapsto \ell \circ u$ is linear and is denoted by u^* . It is an element of $\mathcal{L}(F'; E')$, called the *adjoint* of u . One has

$$u^*(\ell) = \ell \circ u.$$

Because E and F have finite dimensions, then $(u^*)^*$, an element of $\mathcal{L}(E''; F'')$, is an element of $\mathcal{L}(E; F)$, after the identification of E and F with their bi-duals $(E')' = E$. We prove below that it coincides with u , or more accurately, that the following diagram is commutative.

$$\begin{array}{ccc} E & \xrightarrow{u} & F \\ \delta_E \downarrow & & \downarrow \delta_F \\ E'' & \xrightarrow{u^{**}} & F'' \end{array}$$

We list below the main facts about adjunction.

Proposition 1.4 *We recall that E is a finite-dimensional vector space. Then*

- $(u^*)^* = u$.
- $(\ker u)^0 = R(u^*)$ and $R(u)^0 = \ker u^*$.
- u is injective if and only if u^* is surjective.
- u is surjective if and only if u^* is injective.
- For every $u : E \rightarrow F$ and $v : F \rightarrow G$, one has

$$(v \circ u)^* = u^* \circ v^*.$$

Proof. Let $x \in E$ and $L \in F'$ be given. Then

$$\begin{aligned} (u^{**} \circ \delta_E(x))(L) &= (\delta_x \circ u^*)(L) = \delta_x(u^*(L)) = \delta_x(L \circ u) \\ &= (L \circ u)(x) = L(u(x)) = \delta_{u(x)}(L). \end{aligned}$$

We therefore have

$$(u^{**} \circ \delta_E)(x) = \delta_{u(x)} = \delta_F(u(x)) = \delta_F \circ u(x),$$

whence $u^{**} \circ \delta_E = \delta_F \circ u$.

If $x \in \ker u$ and $\ell \in F'$, then $u^*(\ell)(x) = \ell \circ u(x) = \ell(0) = 0$, whence $R(u^*) \subset (\ker u)^0$. Conversely, let $m \in (\ker u)^0$ be given, that is, a linear form on E , vanishing over $\ker u$. When y is in $R(u)$, m is constant over $u^{-1}(y)$. We thus define a map

$$\begin{aligned} R(u) &\rightarrow K \\ y &\mapsto m(x), \end{aligned}$$

where x is any element in $u^{-1}(y)$. We extend this map as a linear form ℓ over F . It satisfies $m = \ell \circ u = u^*(\ell)$, hence the converse inclusion $(\ker u)^0 \subset R(u^*)$. We deduce immediately that u is injective if and only if u^* is surjective.

Again, if $x \in u$ and $\ell \in \ker u^*$, then $\ell(u(x)) = u^*(\ell)(x) = 0(x) = 0$ shows that $\ker u^* \subset R(u)^0$. Conversely, let $\ell \in F'$ vanish over $R(u)$. Then $u^*(\ell) = u \circ \ell \equiv 0$, hence the equality $R(u)^0 = \ker u^*$. It follows immediately that u is surjective if and only if u^* is injective.

Finally, if $\ell \in G'$, then

$$(v \circ u)^*(\ell) = \ell \circ (v \circ u) = (\ell \circ v) \circ u = v^*(\ell) \circ u = u^*(v^*(\ell)) = (u^* \circ v^*)(\ell).$$

□

1.3 Bilinear Maps

Let E, F , and G be three K -vector spaces. A map $b : E \times F \rightarrow G$ is *bilinear* if the partial maps $x \mapsto b(x, y)$ and $y \mapsto b(x, y)$ are linear from E (respectively, from F) into G . The set of bilinear maps from $E \times F$ into G is a vector space, denoted by $\mathbf{Bil}(E \times F; G)$.

If the target space is K itself, then one speaks of a bilinear form. The set of bilinear forms over $E \times F$ is a vector space, denoted by $\mathbf{Bil}(E \times F)$. Its dimension equals the product of $\dim E$ and $\dim F$. If $F = E$, we simply write $\mathbf{Bil}(E)$.

1.3.1 Bilinear Forms When $F = E$

Let $b \in \mathbf{Bil}(E)$ be given. We say that b is *symmetric* if

$$b(x, y) = b(y, x), \quad \forall x, y \in E.$$

Likewise, we say that b is *skew-symmetric* if

$$b(x, y) = -b(y, x), \quad \forall x, y \in E.$$

Finally, we say that b is *alternating* if it satisfies

$$b(x, x) = 0, \quad \forall x \in E.$$

An alternating form is skew-symmetric, because

$$b(x, y) + b(y, x) = b(x + y, x + y) - b(x, x) - b(y, y).$$

If the characteristic of K is different from 2, the converse is true, because the definition contains the identity $2b(x, x) = 0$. Notice that in characteristic 2, skew-symmetry is equivalent to symmetry. To summarize, there are basically two distinct classes, those of symmetric forms and of alternating forms. The third class of skew-symmetric forms equals either the latter if $\text{char}(K) \neq 2$ or the former if $\text{char}(K) = 2$.

1.3.2 Degeneracy versus Nondegeneracy

We assume that E has finite dimension. Let $b \in \mathbf{Bil}(E)$ be given. We may define two linear maps $b_0, b_1 : E \rightarrow E'$ by the formulæ

$$b_0(x)(y) = b_1(y)(x) := b(x, y), \quad \forall x, y \in E.$$

Recalling that $(E')'$ is identified with E through $y(\ell) := \ell(y)$, the following calculation shows that b_1 is the adjoint of b_0 , and conversely:

$$b_0^*(y)(x) = (y \circ b_0)(x) = y(b_0(x)) = b_0(x)(y) = b(x, y) = b_1(y)(x),$$

whence $b_0^*(y) = b_1(y)$ for all $y \in E$.

Because $\dim E' = \dim E$, and thanks to Proposition 1.4, we thus have an equivalence among the injectivity, surjectivity, and bijectivity of b_0 and b_1 . We say that b is *nondegenerate* if b_0 , or equivalently b_1 , is one-to-one. It is degenerate otherwise. Degeneracy means that there exists a nonzero vector $\bar{x} \in E$ such that $b(\bar{x}, \cdot) \equiv 0$.

When b is symmetric (respectively, skew-symmetric), the maps b_0 and b_1 are identical (respectively, opposite), thus their kernels coincide. It is then called the *kernel* of b . A (skew-)symmetric bilinear form is nondegenerate if and only if $\ker b = \{0\}$.

1.3.3 Bilinear Spaces

Let a nondegenerate bilinear form b be given on a space E , either symmetric or skew-symmetric. We say that (E, b) is a *bilinear space*. If $u \in \mathbf{End}(E)$ is given, we may define an *adjoint* u^* , still an element of $\mathbf{End}(E)$, by the formula

$$b(u(x), y) = b(x, u^*(y)), \quad \forall x, y \in E.$$

An accurate expression of $u^*(y)$ is $(b_1)^{-1}(b_1(y) \circ u)$. The following properties are obvious

$$(\lambda u + v)^* = \lambda u^* + v^*, \quad (v \circ u)^* = u^* \circ v^*, \quad (u^*)^* = u.$$

Mind that this adjoint depends on the bilinear structure; another bilinear form yields another adjoint.

We also say that $u \in \mathbf{End}(E)$ is an isometry if

$$b(u(x), u(y)) = b(x, y), \quad \forall x, y \in E.$$

This is equivalent to saying that $u^* \circ u = \text{id}_E$. In particular, an isometry is one-to-one. One easily checks that the set of isometries is a group for the composition of linear maps.

1.3.4 Quadratic Forms

A *quadratic form* over E is a function $q : E \rightarrow K$ given by a formula

$$q(x) = b(x, x), \quad \forall x \in E,$$

where b is a symmetric bilinear form over E .

When the characteristic of K is different from 2, there is a one-to-one correspondence between quadratic forms and symmetric bilinear forms, because of the reciprocal formula

$$b(x, y) = \frac{1}{2}(q(x+y) - q(x) - q(y)),$$

or as well

$$b(x, y) = \frac{1}{4}(q(x+y) - q(x-y)).$$

We then say that b is the *polar form* of q .

The *kernel* of q is by definition that of b , and q is nondegenerate when b is so. We warn the beginner that the kernel of q is usually different from the set

$$\Gamma(q) := \{x \in E \mid q(x) = 0\}.$$

The latter is a cone, that is a set invariant under the multiplication by scalars, whereas the former is a vector space. For this reason, $\Gamma(q)$ is called the *isotropic cone* of q . The kernel is obviously contained in $\Gamma(q)$ and we have the stronger property that

$$\Gamma(q) + \ker q = \Gamma(q).$$

When q is nondegenerate and $\text{char}(K) \neq 2$, $u \in \mathbf{End}(E)$ is a b -isometry if and only if $q \circ u = q$ (use the correspondence $q \leftrightarrow b$ above). We also say that u is a q -isometry.

1.3.5 Euclidean Spaces

When $K = \mathbb{R}$ is the field of real numbers, the range of a quadratic form may be either \mathbb{R} or \mathbb{R}^+ . The latter situation is especially interesting. We say that q (or b as well) is *positive semidefinite* if $q \geq 0$ over E . We say that q is *positive definite* if moreover $q(x) = 0$ implies $x = 0$; that is, $q(y) > 0$ for every nonzero vector. Then the polar form b is called a *scalar product*. A positive-definite form is always nondegenerate, but the converse statement is false.

Definition 1.2 A pair (E, q) where E is a real vector space and q is a positive definite quadratic form on E is called a Euclidean space.

Proposition 1.5 A scalar product satisfies the Cauchy–Schwarz¹ inequality

$$b(x, y)^2 \leq q(x)q(y), \quad \forall x, y \in E.$$

The equality holds true if and only if x and y are colinear.

Proof. The polynomial

$$t \mapsto q(tx + y) = q(x)t^2 + 2b(x, y)t + q(y)$$

takes nonnegative values for $t \in \mathbb{R}$. Hence its discriminant $4(b(x, y))^2 - 4q(x)q(y)$ is nonpositive. When the latter vanishes, the polynomial has a real root t_0 , which implies that $t_0x + y = 0$. \square

The Cauchy–Schwarz inequality implies immediately

$$q(x + y) \leq \left(\sqrt{q(x)} + \sqrt{q(y)} \right)^2,$$

which means that the square root $\| \cdot \| := q^{1/2}$ satisfies the triangle inequality

$$\|x + y\| \leq \|x\| + \|y\|.$$

¹ In *Cauchy–Schwarz*, the name Schwarz (1843–1921) is spelled without a t .

Because $\|\cdot\|$ is positively homogeneous, it is thus a *norm* over E : every Euclidean space is a normed space. The converse is obviously false.

The space \mathbb{R}^n is endowed with a canonical scalar product

$$\langle x, y \rangle := x_1 y_1 + \cdots + x_n y_n.$$

The corresponding norm is

$$\|x\| = (x_1^2 + \cdots + x_n^2)^{1/2}.$$

It is denoted $\|\cdot\|_2$ in Chapter 7.

1.3.6 Hermitian Spaces

When the scalar field is that of complex numbers \mathbb{C} , the complex conjugation yields an additional structure.

Definition 1.3 *Let E be a complex space, and $\phi : E \times E \rightarrow \mathbb{C}$ be a scalar-valued map. We say that ϕ is a sesquilinear form if it satisfies the following*

Linearity: For every $x \in E$, $y \mapsto \phi(x, y)$ is linear,

Anti-linearity: For every $y \in E$, $x \mapsto \phi(x, y)$ is antilinear, meaning

$$\phi(\lambda x + x', y) = \bar{\lambda} \phi(x, y) + \phi(x', y).$$

Given a sesquilinear form ϕ , the formula $\psi(x, y) := \overline{\phi(y, x)}$ defines another sesquilinear form, in general different from ϕ . The equality case is especially interesting:

Definition 1.4 *An Hermitian form is a sesquilinear form satisfying in addition*

$$\phi(y, x) = \overline{\phi(x, y)}, \quad \forall x, y \in E.$$

For an Hermitian form, the function $q(x) := \phi(x, x)$ is real-valued and satisfies

$$q(\lambda x) = |\lambda|^2 q(x).$$

The form ϕ can be retrieved from q via the formula

$$\phi(x, y) = \frac{1}{4}(q(x+y) - q(x-y) - iq(x+iy) + iq(x-iy)). \quad (1.1)$$

Definition 1.5 *An Hermitian form is said to be positive definite if $q(x) > 0$ for every $x \neq 0$.*

There are also semipositive-definite Hermitian forms, satisfying $q(x) \geq 0$ for every $x \in E$. A semipositive-definite form satisfies the Cauchy–Schwarz inequality

$$|\phi(x, y)|^2 \leq q(x)q(y), \quad \forall x, y \in E.$$

In the positive-definite case, the equality holds if and only if x and y are colinear.

Definition 1.6 *An Hermitian space is a pair (E, ϕ) where E is a complex space and ϕ is a positive-definite Hermitian form.*

As in the Euclidean case, an Hermitian form is called a *scalar product*. An Hermitian space is a normed space, where the norm is given by

$$\|x\| := \sqrt{q(x)}.$$

The space \mathbb{C}^n is endowed with a canonical scalar product

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

The corresponding norm is

$$\|x\| = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$