Appendix B

Mostly linear algebra: a brief review

We assume that you have taken an introductory linear algebra course in addition to multivariable calculus. The following is only a quick discussion of some terminology and notation used in this book. We do not give proofs of all results.

B.1 Polynomials and sequences

A *polynomial* is an expression of the form

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n. \tag{B.1.1}$$

In this book, the coefficients $a_0, \ldots, a_n$ will always be complex numbers, and z will usually be complex, so that p is a function from $\mathbb{C}$ to $\mathbb{C}$. We will then call p either a polynomial or a complex polynomial. If all the coefficients are real numbers, we will call p a *real polynomial*; if they are rational, we will call it a *rational polynomial*. (There are other possibilities: in coding theory, for instance, the coefficients of a polynomial may be elements of a finite field.) A real polynomial can take on nonreal values. For example, let $p(z) := 2z^2$ for all $z \in \mathbb{C}$. Then p is a real polynomial but $p(1+i)$ is not a real number.

A real polynomial on a nonempty subset $S \subset \mathbb{C}$ is the restriction of a real polynomial to S. A complex polynomial on S is the restriction of a complex polynomial to S.

We denote by $\mathbb{C}[t]$ the collection of all complex polynomials and by $\mathbb{R}[t]$ the collection of all real polynomials.

Since the identity function $i : \mathbb{C} \to \mathbb{C}$ defined by $i(z) := z$ is continuous, Proposition 1.5.14 implies that every polynomial is continuous on all of $\mathbb{C}$. Hence, for any nonempty subset S of $\mathbb{C}$, Proposition 1.5.12 implies that every polynomial function on S is continuous on all of S.

Recall that in calculus, a sequence in $\mathbb{R}$ is a function $f : \mathbb{N} \to \mathbb{R}$.

Definition B.1.1 (Sequence). Let X be any nonempty set. A *sequence in* X is a function

$$f : \mathbb{N} \to X.$$

If f is a sequence in X and $x_n := f(n)$ for each $n \in \mathbb{N}$, we will often denote f by (x_n). This notation is less cumbersome and more intuitive than the functional notation f. So, for instance, "let (x_n) be a sequence in X" means that (x_n) denotes a sequence

$$f : \mathbb{N} \to X, \quad \text{where } f(n) = x_n \quad \text{for all } n \in \mathbb{N}. \tag{B.1.2}$$

We sometimes denote the sequence (x_n) by an infinite tuple $(x_1, x_2, \dots)$. Each x_i is a *term* of the sequence (x_n). The k^{th} term of a sequence (x_n) is x_k.

If $k \neq l$, then x_k and x_l are said to be *different terms* of the sequence even if $x_k = x_l$. Hence, every sequence (x_n) always has infinitely many different terms, whereas the set $\{x_1, x_2, \dots\}$ that contains all the terms of the sequence may be finite. If M is a subset of a set X and (x_n) is a sequence in X, we write $(x_n) \subset M$ to mean that $x_n \in M$ for all n.

B.2 Vector spaces

Let $\mathcal{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. A *vector space over* $\mathcal{K}$ is a triple $(V, +, \bullet)$, where V is a nonempty set and $+ : V \times V \to V$ and $\bullet : \mathcal{K} \times V \to \mathcal{K}$ are functions satisfying certain properties. The elements of V are then called vectors, $\mathcal{K}$ is called the scalar field of V, elements of $\mathcal{K}$ are called scalars, the function $+$ is called the addition operation, and $\bullet$ is called the scalar multiplication operation. For v, w in $V \times V$ and $\alpha \in \mathcal{K}$, we will write $v + w$ for $+(v, w)$ and we write αv for $\bullet(\alpha, v)$. We call $u + v$ the sum of u and v; we call αv the scalar multiplication of v by α. The required properties of $+$ and $\bullet$ are as follows:

1. *Additive identity*: There exists $\mathbf{0}_V \in V$ such that $\mathbf{0}_V + v = v$ for any $v \in V$.
2. *Additive inverse*: For any $v \in V$, there is $-v \in V$ such that $v + (-v) = \mathbf{0}$.
3. *Commutative law for addition*: $v + w = w + v$ for all $v, w \in V$.
4. *Associative law for addition*: For all $v_1, v_2, v_3 \in V$,
$$v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3.$$
5. *Multiplicative identity*: For all $v \in V$, we have $1v = v$.
6. *Associative law for multiplication*: For all $\alpha, \beta \in \mathcal{K}$ and all $v \in V$, $\alpha(\beta v) = (\alpha\beta)v$.
7. *Distributive law for scalar addition*: For all scalars $\alpha, \beta \in \mathcal{K}$ and all $v \in V$, $(\alpha + \beta)v = \alpha v + \beta v$.
8. *Distributive law for vector addition*: For all $\alpha \in \mathcal{K}$ and $v, w \in V$, we have $\alpha(v + w) = \alpha v + \alpha w$.

Although the operations of addition and scalar multiplication defined on different vector spaces can be different, we will always use + to denote the addition operation, and αv to denote the scalar multiplication of v by α.

A real vector space is one whose scalar field is $\mathbb{R}$; a complex vector space is one whose scalar field is $\mathbb{C}$.

When it's clear or irrelevant what the addition and scalar multiplication operations are, we write simply "V is a vector space over $\mathcal{K}$". When $\mathcal{K}$ is clear or irrelevant, we may write "V is a vector space".

It is easy to show that every vector space V has a unique additive identity (the vector $\mathbf{0}_V$ in property 1 above). It is also called the zero vector. The zero vector of every vector space will be denoted by $\mathbf{0}$.

Let $v_1, \dots, v_n$ be vectors in a vector space V. It follows from properties 3 and 4 above that the meaning of $v_1 + \cdots + v_n$ does not depend on the order in which the v_k are written.

Definition B.2.1 (Vector subspace). Let V be a vector space over $\mathcal{K}$ and let M be a nonempty subset of V. Then M is a *vector subspace* of V if for all m_1, m_2 in M and all $\alpha \in \mathcal{K}$, we have $m_1 + m_2 \in M$ and $\alpha m_1 \in M$ – i.e., M is closed under the operations of addition and scalar multiplication defined on V.

A vector subspace M is *nonzero* (or *nontrivial*) if $M \neq \{\mathbf{0}\}$. It is *trivial* if $M = \{\mathbf{0}\}$. Do not confuse the trivial vector subspace with the empty set; the empty set is not a vector subspace. A vector subspace M is *proper* if $M \neq V$.

Notation B.2.2 (Notation for $\mathbb{R}^n$ and $\mathbb{C}^n$). Throughout this book, we will denote elements of $\mathcal{K}^n$ (where $\mathcal{K}$ is either $\mathbb{R}$ or $\mathbb{C}$) either by a bold letter or by a column matrix whose entries are indexed non-bold versions of the same letter: $\mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{v} := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, and so on (sometimes we also write such an element as the transpose of a line matrix). If $\mathbf{x}$ is in $\mathbb{R}^n$, then $x_1, \dots, x_n$ are real numbers; if $\mathbf{x}$ is in $\mathbb{C}^n$, then $x_1, \dots, x_n$ are complex numbers.

We write vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$ as column matrices because (see Theorem B.5.2) every linear transformation from $\mathcal{K}^n$ to $\mathcal{K}^m$ is given by multiplication by a $m \times n$ matrix A; writing these vectors in column form enables us to use the familiar $f(x) = y$ format, as we do in (B.5.4), for example. It would not make sense to write that equation as $\begin{bmatrix} 1 & 4 & -2 \\ 3 & 0 & 2 \end{bmatrix} (1, -1, -2)$.

Let V be a vector space over $\mathcal{K}$. Let A, B be nonempty subsets of V and let $v \in V, \alpha \in \mathcal{K}$. Then

$$v + A := \{\, v + a \mid a \in A \,\} \tag{B.2.1}$$

$$A + B := \{\, a + b \mid a \in A,\, b \in B \,\} \tag{B.2.2}$$

$$\alpha B := \{\, \alpha b \mid b \in B \,\}. \tag{B.2.3}$$

We also define $A + v$ to be $v + A$ and $-B := (-1)B$.

Hence, αB can be obtained from B by "stretching" or "shrinking" each vector in B by a factor of $|\alpha|$. Consequently, if B is an open ball in $\mathbb{R}^2$ with radius $r > 0$, then αB is an open ball in $\mathbb{R}^2$ with radius $|\alpha|\, r$. Although the above operations look like addition and multiplication of numbers, if A is a vector subspace of V, then $A = -A$ and A and $-A$ don't cancel.

Definition B.2.3. (Balanced subset, convex subset). Let V be a vector space with scalar field $\mathcal{K}$.

1. $M \subset V$ is *balanced* if $tM \subset M$ whenever $t \in \mathcal{K}$ and $|t| \leq 1$.
2. $M \subset V$ is *convex* if whenever x, y are in M and $t \in [0, 1]$, then $tx + (1-t)y$ is also in M.

Thus every vector subspace of V is convex and balanced. If M is a subset of $\mathbb{R}^n$, then we can restate the definition of convexity to say that M is convex if for any two points $\mathbf{x}, \mathbf{y} \in M$, the straight line segment joining $\mathbf{x}$ to $\mathbf{y}$ lies entirely in M.

Example B.2.4. Any convex, symmetric subset U containing $\mathbf{0}$ must be balanced. To see this, let t be a scalar with $0 < |t| \leq 1$. Then since U is convex and $\mathbf{0} \in U$, we have $|t|U = |t|U + (1 - |t|)\{\mathbf{0}\} \subset |t|U + (1 - |t|)U \subset U$. Hence, since U is symmetric, $\frac{t}{|t|}U \subset U$ so that

$$tU = |t| \left(\frac{t}{|t|} U \right) \subset |t|U \subset U.$$

If $t = 0$, then we also have $tU \subset U$, because $\mathbf{0} \in U$. Thus, U is balanced.

Product of vector spaces

Assume that $X_1, \ldots, X_n$ are vector spaces over a common scalar field $\mathcal{K}$. Let

$$\mathbf{X} := \{\, (x_1, \ldots, x_n) \mid x_j \in X_j, 1 \leq j \leq n \,\} = X_1 \times \cdots \times X_n. \tag{B.2.4}$$

For $\mathbf{x} := (x_1, \ldots, x_n)$, $\mathbf{y} := (y_1, \ldots, y_n)$ in $\mathbf{X}$ and $\alpha \in \mathcal{K}$, define

$$\begin{aligned} \mathbf{x} + \mathbf{y} &:= (x_1 + y_1, \ldots, x_n + y_n) \quad \text{and} \\ \alpha\mathbf{x} &:= (\alpha x_1, \ldots, \alpha x_n). \end{aligned} \tag{B.2.5}$$

If as usual we denote by $\mathbf{0}$ the zero vector in each X_j, it can be verified easily that $\mathbf{X}$ is a vector space over $\mathcal{K}$ with $(\mathbf{0}, \ldots, \mathbf{0})$ as the zero vector. The vector space $\mathbf{X}$ is called the *product* of the vector spaces $X_1, \ldots, X_n$.

B.3 Linear independence and span

The definitions of *linear combination* and *linear independence* in the context of finite-dimensional vector spaces remain valid in infinite dimensions.