3. Euclidean Spaces

Inner product and norm • Open and closed sets • Continuity • Differentiation.

3.1. **Inner product and norm.** Let $\mathbb{R}$ be the *set* of real numbers and let $\mathbb{R}^m$ $(m \geq 1)$ denote the Cartesian product of m copies of $\mathbb{R}$. The elements of $\mathbb{R}^m$ are ordered m-tuples of real numbers. Thus

$$\mathbb{R}^m := \{x = (x_1, \ldots, x_m) \,:\, x_i \in \mathbb{R}\}.$$

An element of $\mathbb{R}^m$ is often called a *point*. Under the usual operations

$$x + y := (x_1 + y_1, \ldots, x_m + y_m) \quad \text{and} \quad \lambda x := (\lambda x_1, \ldots, \lambda x_m) \qquad (x, y \in \mathbb{R}^m,\ \lambda \in \mathbb{R})$$

$\mathbb{R}^m$ is a *vector space* over $\mathbb{R}$. Hence the elements of $\mathbb{R}^m$ can also be referred to as *vectors*.

Note : The set $\mathbb{R}^m$ may be equipped with various natural structures (e.g., group structure, vector space structure, topological structure, etc.) thus yielding various *spaces*, each such space having the same underlying *set* $\mathbb{R}^m$. We must usually decide from the context which structure is intended.

Many geometric concepts require an extra structure on $\mathbb{R}^m$ that we now define.

Definition 42. The **Euclidean space** $\mathbb{R}^m$ is the above mentioned vector space $\mathbb{R}^m$ together with the **standard inner product** (or dot product)

$$x \bullet y := x_1 y_1 + \cdots + x_m y_m \qquad (x, y \in \mathbb{R}^m).$$

We say that $x, y \in \mathbb{R}^m$ are **orthogonal** if $x \bullet y = 0$. The most important properties of the standard inner product are the following.

Proposition 19. *If x, y, z are vectors in $\mathbb{R}^m$ and $\lambda \in \mathbb{R}$, then*

(IP1) $x \bullet y = y \bullet x$ *(symmetry).*
(IP2) $(\lambda x + y) \bullet z = \lambda x \bullet z + y \bullet z$ *(linearity).*
(IP3) $x \bullet x \geq 0$*, and* $x \bullet x = 0$ *if and only if* $x = 0$ *(positive definiteness).*

Proof. Straightforward computation. □

Definition 43. The **Euclidean norm** $\|x\|$ of $x \in \mathbb{R}^m$ is defined as

$$\|x\| := \sqrt{x \bullet x}.$$

If $m = 1$, then $\|x\|$ is the usual *absolute value* $|x|$ of x. The relationship between the norm and the vector structure of $\mathbb{R}^m$ is very important.

◇ **Exercise 32.** Show that if $x, y \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}$, then

(a) $\|x\| \geq 0$, and $\|x\| = 0$ if and only if $x = 0$ (positivity).
(b) $\|\lambda x\| = |\lambda| \, \|x\|$ (homogeneity).
(c) $x \bullet y = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$ (polarization identity).
(d) $\|x \pm y\|^2 = \|x\|^2 + \|y\|^2$ if and only if $x \bullet y = 0$ (Pythagorean property).

Theorem 20. (Cauchy-Schwarz Inequality) *If $x, y \in \mathbb{R}^m$, then*

$$|x \bullet y| \leq \|x\| \, \|y\|.$$

Equality holds if and only if x and y are linearly dependent.

Proof. If x and y are linearly dependent, equality clearly holds. Why ? If not, then $\lambda x - y \neq 0$ for all $\lambda \in \mathbb{R}$, so

$$\begin{aligned} 0 < \|\lambda x - y\|^2 &= (\lambda x_1 - y_1)^2 + \cdots + (\lambda x_m - y_m)^2 \\ &= (x_1^2 + \cdots + x_m^2)\,\lambda^2 - 2(x_1y_1 + \cdots + x_my_m)\,\lambda + y_1^2 + \cdots + y_m^2. \end{aligned}$$

Therefore the right hand side is a quadratic equation in λ with no real solution, and its discriminant must be negative. Thus

$$4\,(x_1y_1 + \cdots + x_my_m)^2 - 4\left(x_1^2 + \cdots + x_m^2\right)\left(y_1^2 + \cdots + y_m^2\right) < 0$$
$$(x \bullet y)^2 < \|x\|^2 \, \|y\|^2$$

which implies $|x \bullet y| < \|x\| \, \|y\|$. □

The Cauchy-Schwarz Inequality serves in proving several other inequalities (this is Problem 11).

Definition 44. The **standard basis** for $\mathbb{R}^m$ consists of the vectors

$$e_j = (\delta_{1j}, \ldots, \delta_{mj}), \quad j = \overline{1, m}$$

where δ_{ij} equals 1 if $i = j$ and equals 0 if $i \neq j$.

Thus we write

$$x = x_1e_1 + \cdots + x_me_m \qquad (x \in \mathbb{R}^m).$$

With respect to the standard inner product on $\mathbb{R}^m$, the standard basis is **orthonormal**, i.e., $e_i \bullet e_j = \delta_{ij}$ for $i, j = \overline{1, m}$. (Thus $\|e_j\| = 1$, while e_i and e_j for distinct i and j are orthogonal vectors.)

Definition 45. For $x, y \in \mathbb{R}^m$ we define the **Euclidean distance** $d(x, y)$ by

$$d(x, y) := \|x - y\|.$$

From **Exercise 32** and PROBLEM 11 we immediately obtain (for $x, y, z \in \mathbb{R}^m$)

(M1) $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$.
(M2) $d(x, y) = d(y, x)$.
(M3) $d(x, z) \leq d(x, y) + d(y, z)$.

NOTE : (1) More generally, a **metric space** is defined as a set M equipped with a *distance* between its elements satisfying the properties (M1) – (M3). So *the Euclidean space* $\mathbb{R}^m$ *is a metric space*. The notation $d(x, y) = \|x - y\|$ is frequently useful even when we are dealing with the Euclidean space $\mathbb{R}^m$ as a metric space and not using its vector space structure. In particular, $\|x\| = d(x, 0)$.

(2) An *abstract* concept of **Euclidean space** (i.e., a space satisfying the *axioms* of Euclidean geometry) can be introduced. It is defined as a structure $(\mathsf{M}, \mathsf{E}, \Phi)$, consisting of a (non-empty) *set* M, an associated *standard vector space* E (which is a real Euclidean vector space, i.e., a real vector space equipped with a scalar product $(\cdot|\cdot) : \mathsf{E} \times \mathsf{E} \to \mathbb{R}$), and a *structure map*

$$\Phi : \mathsf{M} \times \mathsf{M} \to \mathsf{E}, \quad (x, y) \mapsto \overrightarrow{xy}$$

such that

(ES1) $\overrightarrow{xy} + \overrightarrow{yz} = \overrightarrow{xz}$ for every $x, y, z \in \mathsf{M}$;
(ES2) for every $o \in \mathsf{M}$ and every $v \in \mathsf{E}$, there is a unique $x \in \mathsf{M}$ such that $\overrightarrow{ox} = v$.

Elements of M are called **points**, whereas elements of E are called **vectors**. ($\overrightarrow{ox}$ is the *position vector* of x with the initial point o.) The *dimension* of the space M is the dimension of the vector space E. It turns out that

(i) if we fix an arbitrary point $o \in \mathsf{M}$, there is a one-to-one correspondence between (the space) M and (the associated vector space) E (the mapping $x \mapsto \overrightarrow{ox}$ is a bijection);

(ii) in addition, if we fix an arbitrary (ordered) orthonormal basis $(e_1, e_2, \ldots, e_m)$ of E, the (inner product) spaces E and $\mathbb{R}^m$ are *isomorphic*. In other words, the scalar product on E "is" the dot product: for $v, w \in \mathsf{E}$,

$$\begin{aligned} (v \,|\, w) &= (\, v_1 e_1 + \cdots + v_m e_m \,|\, w_1 e_1 + \cdots + w_m e_m \,) \\ &= v_1 w_1 + \cdots + v_m w_m. \end{aligned}$$

In this sense, we *identify* the abstract m-dimensional Euclidean space M with the (concrete) Euclidean space $\mathbb{R}^m$.

We conclude this section with some important remarks (about notation). The element (vector) $(0, \ldots, 0) \in \mathbb{R}^m$ will usually be denoted simply 0.

If $\tau : \mathbb{R}^m \to \mathbb{R}^m$ is a linear transformation, the matrix of τ with respect to the standard basis of $\mathbb{R}^m$ is the $m \times m$ matrix $T = [t_{ij}]$, where $T(e_j) = \sum_{i=1}^m t_{ij} e_i$ (the coefficients of $T(e_j)$ appear in the j^{th} column of the matrix). If the linear transformation $\sigma : \mathbb{R}^m \to \mathbb{R}^m$ has the matrix S, then (the composite) transformation $\sigma\tau$ has the matrix ST (matrix multiplication).

3.2. **Open and closed sets.** The analog in $\mathbb{R}^m$ of an *open interval* in $\mathbb{R}$ is introduced in the following

DEFINITION 46. For $p \in \mathbb{R}^m$ and $\delta > 0$, we denote the **open ball** of center p and radius δ by

$$\mathcal{B}(p, \delta) := \{x \in \mathbb{R}^m : \|x - p\| < \delta\}.$$

A point p in a set $\mathsf{A} \subseteq \mathbb{R}^m$ is said to be an **interior point** of A if there exists $\delta > 0$ such that $\mathcal{B}(p, \delta) \subseteq \mathsf{A}$. The set of interior points of A is called the **interior** of A and is denoted by $\operatorname{int}(\mathsf{A})$. Note that $\operatorname{int}(\mathsf{A}) \subseteq \mathsf{A}$.

DEFINITION 47. A set $\mathsf{A} \subseteq \mathbb{R}^m$ is said to be **open** (in $\mathbb{R}^m$) if $\mathsf{A} = \operatorname{int}(\mathsf{A})$ (i.e., if every point of A is an interior point of A).

Note that the *empty set* $\emptyset$ satisfies every definition involving conditions on its elements, therefore $\emptyset$ is open. Furthermore, the whole space $\mathbb{R}^m$ is open.

PROPOSITION 21. *The set $\mathcal{B}(p, \delta)$ is open in $\mathbb{R}^m$, for every $p \in \mathbb{R}^m$ and $\delta > 0$.*

Proof. For arbitrary $q \in \mathcal{B}(p, \delta)$ set $\beta = \|q - p\|$, then $\delta - \beta > 0$. Hence $\mathcal{B}(q, \delta - \beta) \subseteq \mathcal{B}(p, \delta)$, because for every $x \in \mathcal{B}(q, \delta - \beta)$

$$\|x - p\| \leq \|x - q\| + \|q - p\| < (\delta - \beta) + \beta = \delta.$$

$\square$

PROPOSITION 22. *For any $\mathsf{A} \subseteq \mathbb{R}^m$, the interior $\operatorname{int}(\mathsf{A})$ is the largest open set contained in A.*

Proof. First, we show that $\operatorname{int}(\mathsf{A})$ is open. If $p \in \operatorname{int}(\mathsf{A})$, there is $\delta > 0$ such that $\mathcal{B}(p, \delta) \subseteq \mathsf{A}$. As in the proof of PROPOSITION 21, we find for any $q \in \mathcal{B}(p, \delta)$ a $\beta > 0$ such that $\mathcal{B}(q, \beta) \subseteq \mathsf{A}$. But this implies $\mathcal{B}(p, \delta) \subseteq \operatorname{int}(\mathsf{A})$, and hence $\operatorname{int}(\mathsf{A})$ is an open set.

Furthermore, if $\mathsf{U} \subseteq \mathsf{A}$ is open, it is clear by definition that $\mathsf{U} \subseteq \operatorname{int}(\mathsf{A})$, thus $\operatorname{int}(\mathsf{A})$ is the largest open set contained in A. $\square$

◇ **Exercise 33.** Show that

(a) the union of any collection of open subsets of $\mathbb{R}^m$ is again open in $\mathbb{R}^m$;
(b) the intersection of finitely many open subsets of $\mathbb{R}^m$ is open in $\mathbb{R}^m$.

Let $\emptyset \neq \mathsf{A} \subseteq \mathbb{R}^m$. An **open neighborhood** of A is an open set containing A, and a **neighborhood** of A is any set containing an open neighborhood of A. A neighborhood of a set $\{p\}$ is also called a neighborhood of the point $p \in \mathbb{R}^m$. (Note that $p \in \mathsf{A} \subseteq \mathbb{R}^m$ is an interior point of A <u>if and only</u> if A is a neighborhood of p.)

DEFINITION 48. A set F is said to be **closed** if its *complement* $\mathsf{F}^c := \mathbb{R}^m \setminus \mathsf{F}$ is open.

The empty set is closed, and so is the entire space $\mathbb{R}^m$.

PROPOSITION 23. *For every $p \in \mathbb{R}^m$ and $\delta > 0$, the set $\overline{\mathcal{B}}(p,\delta) := \{x \in \mathbb{R}^m : \|x-p\| \leq \delta\}$ is closed. ($\overline{\mathcal{B}}(p,\delta)$ is the **closed ball** of center p and radius δ.)*

Proof. For arbitrary $q \in \overline{\mathcal{B}}(p,\delta)^c$ set $\beta = \|p-q\|$, then $\beta - \delta > 0$. So $\mathcal{B}(q, \beta-\delta) \subseteq \overline{\mathcal{B}}(p,\delta)^c$, because by the reverse triangle inequality (this is PROBLEM 11), for every $x \in \mathcal{B}(q,\beta-\delta)$

$$\|p-x\| \geq \|p-q\| - \|x-q\| > \beta - (\beta-\delta) = \delta.$$

This proves that $\overline{\mathcal{B}}(p,\delta)^c$ is open. □

DEFINITION 49. A point $p \in \mathbb{R}^m$ is said to be a **cluster point** of a set $\mathsf{A} \subseteq \mathbb{R}^m$ if for every $\delta > 0$ we have $\mathcal{B}(p,\delta) \cap \mathsf{A} \neq \emptyset$. The set of cluster points of A is called the **closure** of A and is denoted by $\operatorname{cl}(\mathsf{A})$.

PROPOSITION 24. *Let $\mathsf{A} \subseteq \mathbb{R}^m$. Then $\operatorname{cl}(\mathsf{A})^c = \operatorname{int}(A^c)$; in particular, the closure of A is a closed set. Moreover, $\operatorname{int}(\mathsf{A})^c = \operatorname{cl}(\mathsf{A}^c)$.*

Proof. Note that $\mathsf{A} \subseteq \operatorname{cl}(\mathsf{A})$. To say that x is *not* a cluster point of A means that it is an interior point of A^c. Thus $\operatorname{cl}(\mathsf{A})^c = \operatorname{int}(\mathsf{A}^c)$, or $\operatorname{cl}(\mathsf{A}) = \operatorname{int}(\mathsf{A}^c)^c$, which implies that $\operatorname{cl}(\mathsf{A})$ is closed in $\mathbb{R}^m$.

Furthermore, by applying this identity to A^c we obtain that $\operatorname{int}(\mathsf{A})^c = \operatorname{cl}(\mathsf{A}^c)$. □

By taking complements of sets we immediately obtain the following result.

PROPOSITION 25. *For any $\mathsf{A} \subseteq \mathbb{R}^m$, the closure $\operatorname{cl}(\mathsf{A})$ is the smallest closed set containing A.*

From set theory we recall DE MORGAN'S LAWS, which state, for arbitrary collections $(\mathsf{A}_i)_{i\in I}$ of sets $\mathsf{A}_i \subseteq \mathbb{R}^m$, that

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c \quad \text{and} \quad \left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c.$$

In view of these laws and **Exercise 33** we find, by taking complements of sets,

PROPOSITION 26.

(a) *The intersection of any collection of closed subsets of $\mathbb{R}^m$ is again closed in $\mathbb{R}^m$.*
(b) *The union of finitely many closed subsets of $\mathbb{R}^m$ is closed in $\mathbb{R}^m$.*

3.3. **Continuity.** Let $\mathsf{U} \subseteq \mathbb{R}^m$ be an *open* set. A mapping $F : \mathsf{U} \to \mathbb{R}^n$ is **continuous** at $p \in \mathsf{U}$ if given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$F\left(\mathcal{B}(p,\delta)\right) \subseteq \mathcal{B}(F(p),\varepsilon).$$

In other words, F is continuous at p if points arbitrarily close to $F(p)$ are images of points sufficiently close to p. We say that F is **continuous** provided it is continuous at each $p \in \mathsf{U}$.

Note : Equivalently, F is continuous at $p \in \mathsf{U}$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|F(x) - F(p)\| < \varepsilon$ for $\|x - p\| < \delta$. This simply means that $\lim_{x\to p} F(x) = F(p)$.

A mapping $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ determines n $\mathbb{R}$-valued functions (of m variables) as follows. Let $x = (x_1, \dots, x_m) \in \mathsf{U}$ and $F(x) = (y_1, \dots, y_n)$. Then we can write

$$y_1 = F_1(x_1, \dots, x_m), \quad y_2 = F_2(x_1, \dots, x_m), \quad \dots, \quad y_n = F_n(x_1, \dots, x_m).$$

The functions $F_i : \mathsf{U} \to \mathbb{R}$, $i = \overline{1,n}$ are the **component functions** of F. The continuity of the mapping F is equivalent to the continuity of its component functions.

$\diamond$ **Exercise 34.** Prove that a mapping $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is continuous if and only if each component function $F_i : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}$ is continuous.

The following results are standard (and easy to prove).

Proposition 27. *Let $F, G : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be continuous mappings and let $\lambda \in \mathbb{R}$. Then $F + G$, λF, and $F \bullet G$ are each continuous. If $n = 1$ and $G(x) \neq 0$ for all $x \in \mathsf{U}$, then the quotient $\frac{F}{G}$ is also continuous.*

Proposition 28. *Let $F : \mathsf{U} \subseteq \mathbb{R}^\ell \to \mathbb{R}^m$ and $G : \mathsf{V} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be continuous mappings, where U and V are open sets such that $F(\mathsf{U}) \subseteq \mathsf{V}$. Then $G \circ F$ is a continuous mapping.*

$\diamond$ **Exercise 35.** Show that the following mappings are continuous.

(a) The *identity mapping* $\mathbf{1}_{\mathbb{R}^m} : \mathbb{R}^m \to \mathbb{R}^m, \quad x \mapsto x$.
(b) The *norm function* $\nu : \mathbb{R}^m \to \mathbb{R}, \quad x \mapsto \|x\|$.
(c) The i^{th} *natural projection* $\mathrm{pr}_i : \mathbb{R}^m \to \mathbb{R}, \quad x \mapsto x_i$.

Hence derive that every *polynomial function* (in several variables)

$$p_k : \mathbb{R}^m \to \mathbb{R}, \quad x = (x_1, \dots, x_m) \mapsto \sum_{\substack{i_1,\dots,i_m=0 \\ i_1+\cdots+i_m\leq k}}^{k} a_{i_1\dots i_m} x_1^{i_1} \dots x_m^{i_m}$$

is continuous.

Note : More generally, every *rational function* (i.e., a quotient of two polynomial functions) is continuous. It can be shown that *elementary* functions like exp, log, sin, and cos are also continuous.

Linear mappings $L : \mathbb{R}^m \to \mathbb{R}^n$ play an important role in differentiation. Such mappings are continuous.

$\diamond$ **Exercise 36.** Show that every linear mapping $L : \mathbb{R}^m \to \mathbb{R}^n$ is continuous.

In most applications it is convenient to express continuity in terms of neighborhoods instead of open balls.

$\diamond$ **Exercise 37.** Prove that a mapping $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is continuous at $p \in \mathsf{U}$ if and only if given a neighborhood $\mathcal{N}$ of $F(p)$ in $\mathbb{R}^n$ there exists a neighborhood $\mathcal{M}$ of p in $\mathbb{R}^m$ such that $F(\mathcal{M}) \subseteq \mathcal{N}$.

It is often necessary to deal with mappings (functions) defined on arbitrary (i.e., not necessarily open) sets. To extend the previous ideas to this situation, we shall proceed as follows.

Let $F : \mathsf{A} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be a mapping, where A is an *arbitrary* set. We say that F is **continuous** on A provided there exists an open set $\mathcal{U} \subseteq \mathbb{R}^m$ containing A, and a continuous mapping $\overline{F} : \mathcal{U} \to \mathbb{R}^n$ such that (the restriction) $\overline{F}\big|_{\mathsf{A}} = F$. In other words, F is continuous on A if it is the restriction of a continuous mapping defined on an open neighborhood of A.

Note : It is clear that if $F : \mathsf{A} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is continuous and $p \in \mathsf{A}$, then given a neighborhood $\mathcal{N}$ of $F(p)$ in $\mathbb{R}^n$, there exists a neighborhood $\mathcal{M}$ of p in $\mathbb{R}^m$ such that $F(\mathcal{M} \cap \mathsf{A}) \subseteq \mathcal{N}$. For this reason, it is convenient to call the set $\mathcal{M} \cap \mathsf{A}$ a *neighborhood* of p in A.

Example 50. An important class of continuous mappings is formed by the mappings $F : \mathsf{A} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ that are **Lipschitz continuous**, i.e., for which there exists $k > 0$ such that

$$\|F(x) - F(y)\| \leq k\,\|x - y\| \qquad (x, y \in \mathsf{A}).$$

Such a number k is called a *Lipschitz constant* for F. For example, the norm function $\nu : x \mapsto \|x\|$ is a Lipschitz continuous on $\mathbb{R}^m$ with Lipschitz constant 1.

$\diamond$ **Exercise 38.** Consider a mapping $F : \mathsf{A} \to \mathbb{R}^n$, where $\mathsf{A} \subseteq \mathbb{R}^m$ is an *arbitrary* set. Show that the following statements are logically equivalent.

(a) F is continuous.
(b) $F^{-1}(\mathsf{O})$ is open in A for every open set O in $\mathbb{R}^n$. (In particular, if A is open in $\mathbb{R}^m$ then: $F^{-1}(\mathsf{O})$ is open in $\mathbb{R}^m$ for every open set O in $\mathbb{R}^n$.)

(c) $F^{-1}(\mathsf{F})$ is closed in A for every closed set F in $\mathbb{R}^n$. (In particular, if A is closed in $\mathbb{R}^m$ then: $F^{-1}(\mathsf{F})$ is closed in $\mathbb{R}^m$ for every closed set F in $\mathbb{R}^n$.)

(A subset $\mathsf{U} \subseteq \mathsf{A}$ is said to be *open in* A if there is an open set W such that $\mathsf{U} = \mathsf{A} \cap \mathsf{W}$. Likewise, a subset V is said to be *closed in* A if there exists a closed set W such that $\mathsf{V} = \mathsf{A} \cap \mathsf{W}$.)

Definition 51. A set $\mathsf{A} \subseteq \mathbb{R}^m$ is said to be **disconnected** if there exist open sets U and V in $\mathbb{R}^m$ such that

$$\mathsf{A} \cap \mathsf{U} \neq \emptyset, \quad \mathsf{A} \cap \mathsf{V} \neq \emptyset, \quad (\mathsf{A} \cap \mathsf{U}) \cap (\mathsf{A} \cap \mathsf{V}) = \emptyset, \quad (\mathsf{A} \cap \mathsf{U}) \cup (\mathsf{A} \cap \mathsf{V}) = \mathsf{A}.$$

(In other words, A is the union of two disjoint non-empty subsets that are open in A.) The set A is said to be **connected** if A is *not* disconnected.

It is not difficult to prove that *the only connected subsets of* $\mathbb{R}$ *are the intervals: open, closed or half-open* (these include the singletons and the set $\mathbb{R}$ itself). The following result then follows (this is Problem 14):

Theorem 29 (Intermediate Value Theorem). *Let* $\mathsf{A} \subseteq \mathbb{R}^m$ *be connected and let* $F : \mathsf{A} \to \mathbb{R}$ *be a continuous function. Then* $F(\mathsf{A})$ *is an interval in* $\mathbb{R}$*; in particular,* F *takes all values between any two that it assumes.*

Definition 52. We say that a continuous mapping $F : \mathsf{A} \subseteq \mathbb{R}^m \to \mathbb{R}^m$ is a **homeomorphism** onto $F(\mathsf{A})$ if F is one-to-one and the inverse $F^{-1} : F(\mathsf{A}) \subseteq \mathbb{R}^m \to \mathbb{R}^m$ is continuous. In this case A and $F(\mathsf{A})$ are *homeomorphic* sets.

Example 53. Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$F(x_1, x_2, x_3) = (ax_1, bx_2, cx_3), \qquad a, b, c \in \mathbb{R} \setminus \{0\}.$$

F is clearly continuous, and the restriction of F to the (unit) *sphere*

$$\mathbb{S}^2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_1^2 + x_2^2 + x_3^2 = 1\right\}$$

is a continuous mapping $\widetilde{F} : \mathbb{S}^2 \to \mathbb{R}^3$. Observe that $\widetilde{F}(\mathbb{S}^2) = \mathbb{E}$, where $\mathbb{E}$ is the *ellipsoid*

$$\mathbb{E} = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1\right\}.$$

It is also clear that F is one-to-one and that

$$F^{-1}(x_1, x_2, x_3) = \left(\frac{x_1}{a}, \frac{x_2}{b}, \frac{x_3}{c}\right).$$

Thus $\widetilde{F}^{-1} = F^{-1}|_{\mathbb{E}}$ is continuous. Therefore, $\widetilde{F}$ is a *homeomorphism* of the sphere $\mathbb{S}^2$ onto the ellipsoid $\mathbb{E}$.

Note : There is a class of infinite sets, called *compact sets*, that in certain limited aspects behave very much like finite sets. A set $\mathsf{K} \subseteq \mathbb{R}^m$ is said to be **sequentially compact** if every sequence of elements in K contains a subsequence which *converges* to a point in K. (A sequence $(x_k)_{k\in\mathbb{N}}$ of elements $x_k \in \mathbb{R}^m$ is said to be **convergent**, with *limit* $p \in \mathbb{R}^m$, if $\lim_{k\to\infty} \|x_k - p\| = 0$, which is a limit of numbers in $\mathbb{R}$. Recall that this limit means: for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|x_k - p\| < \varepsilon$ for $k \geq N$. In this case we write $\lim_{k\to\infty} x_k = p$.) It follows immediately that *a subset of a sequentially compact set* $\mathsf{K} \subseteq \mathbb{R}^m$ *is sequentially compact if and only if it is closed in* K.

Continuous mapping do not necessarily preserve closed sets; on the other hand, they do preserve (sequentially) compact sets. (In this sense compact and finite sets behave similarly: the image of a finite set under a mapping is a finite set too.) More precisely, *if* $\mathsf{K} \subset \mathbb{R}^m$ *is a sequentially compact set and* $F : \mathbb{R}^m \to \mathbb{R}^n$ *is continuous, then* $F(\mathsf{K}) \subset \mathbb{R}^n$ *is sequentially compact.* The following characterization is very useful: *A set* $\mathsf{K} \subset \mathbb{R}^m$ *is sequentially compact if and only if it is bounded and closed.* (A set $\mathsf{A} \subset \mathbb{R}^m$ is **bounded** if there exists a number $k > 0$ such that $\|x\| \leq k$ for all $x \in \mathsf{A}$; equivalently, if there exists a number $k > 0$ such that $\mathsf{A} \subseteq \overline{\mathcal{B}}(0, k)$.)

There is an alternative, more general, definition of compactness for sets. A subset $\mathsf{K} \subset \mathbb{R}^m$ is said to be **compact** if every open covering of K contains a finite subcovering of K. (A collection $(\mathsf{O}_i)_{i\in I}$ of open sets in $\mathbb{R}^m$ is said to be an *open covering* of a set $\mathsf{K} \subseteq \mathbb{R}^m$ if $\mathsf{K} \subseteq \bigcup_{i\in I} \mathsf{O}_i$.)

In spaces like $\mathbb{R}^m$, however, the two definitions of compactness coincide; this is a consequence of the following result.

(Heine-Borel Theorem) *A set* $\mathsf{K} \subset \mathbb{R}^m$ *is compact if and only if it is bounded and closed.*

3.4. **Differentiation.** Let U be an open subset of $\mathbb{R}^m$ and let $p \in \mathsf{U}$. A function $F : \mathsf{U} \to \mathbb{R}$ is **differentiable** at p if there exists a linear functional $L_p : \mathbb{R}^m \to \mathbb{R}$ such that

$$\lim_{x\to p} \frac{F(x) - F(p) - L_p(x-p)}{\|x-p\|} = 0$$

or, equivalently, if there exist a linear functional $L_p : \mathbb{R}^m \to \mathbb{R}$ and a function $R(\cdot, p)$, defined on an open neighborhood $\mathcal{V}$ of p, such that

$$F(x) = F(p) + L_p(x-p) + \|x-p\| \cdot R(x,p), \qquad x \in \mathcal{V}$$

and

$$\lim_{x\to p} R(x,p) = 0.$$

Then L_p is called a **derivative** (or differential) of F at p. We say that F is **differentiable** provided it is differentiable at each $p \in \mathsf{U}$.

Note : We think of a derivative L_p as a "linear" *approximation* of F near p. By the definition, the error involved in replacing $F(x)$ by $F(p) + L_p(x-p)$ (this is an *affine* map) is negligible compared to the distance from x to p, provided that this distance is sufficiently small.

If $L_p(x) = b_1x_1 + \cdots + b_mx_m$ is a derivative of F at p, then

$$b_i = \frac{\partial F}{\partial x_i}(p) := \lim_{t\to 0} \frac{1}{t} \left(F(p + te_i) - F(p)\right), \quad i = \overline{1, m}.$$

In particular, if F is differentiable at p, these *partial derivatives* exist and the derivative L_p is *unique*. We denote by $DF(p)$ (or sometimes $F'(p)$) *the* derivative of F at p, and write (by a slight abuse of notation)

$$DF(p) = \frac{\partial F}{\partial x_1}(p)(x_1 - p_1) + \frac{\partial F}{\partial x_2}(p)(x_2 - p_2) + \cdots + \frac{\partial F}{\partial x_m}(p)(x_m - p_m).$$

$\diamond$ **Exercise 39.** Show that any linear functional $F : \mathbb{R}^m \to \mathbb{R}$ is differentiable and $DF(p) = F$ for all $p \in \mathbb{R}^m$.

$\diamond$ **Exercise 40.** Prove that any differentiable function $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}$ is continuous.

Note : Mere existence of partial derivatives is *not* sufficient for differentiability (of the function F). For example, the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$F(x_1, x_2) = \frac{x_1 x_2}{x_1^2 + x_2^2} \quad \text{and} \quad F(0,0) = 0$$

is *not* continuous at $(0,0)$, yet both partial derivatives are defined there. However, *if all partial derivatives* $\frac{\partial F}{\partial x_i}$, $i = \overline{1,m}$ *are defined and continuous in a neighborhood of* p, *then* F *is differentiable at* p.

If the function $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}$ has all partial derivatives continuous (on U) we say that F is **continuously differentiable** (or of *class* C^1) on U. We denote this class of functions by $C^1(\mathsf{U})$. (The class of continuous functions on U is denoted by $C^0(\mathsf{U})$.)

Note : We have seen that

$$F \in C^1(\mathsf{U}) \Rightarrow F \text{ is differentiable (on } \mathsf{U}) \Rightarrow \text{all partial derivatives } \frac{\partial F}{\partial x_i} \text{ exist (on } \mathsf{U})$$

but the converse implications may fail. Many results actually need F to be of class C^1 rather than differentiable.

If $r \geq 1$, the class $C^r(\mathsf{U})$ of functions $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}$ that are *r-fold continuously differentiable* (or C^r functions) is specified inductively by requiring that the partial derivatives of F exist and belong to $C^{r-1}(\mathsf{U})$. If F is of class C^r for all r, then we say that F is of *class* C^∞ or simply **smooth**. The class of smooth functions on U is denoted by $C^\infty(\mathsf{U})$.

Note : If $F \in C^r(\mathsf{U})$, then (at any point of U) the value of the partial derivatives of order k, $1 < k \leq r$ is independent of the order of differentiation; that is, if $(j_1, \dots, j_k)$ is a permutation of $(i_1, \dots, i_k)$, then

$$\frac{\partial^k F}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k F}{\partial x_{j_1} \dots \partial x_{j_k}}.$$

We are now interested in extending the notion of differentiability to mappings $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$. We say that F is **differentiable** at $p \in \mathsf{U}$ if its component functions are differentiable at p; that is, by writing

$$F(x_1, \ldots, x_m) = (F_1(x_1, \ldots, x_m), \ldots, F_n(x_1, \ldots, x_m))$$

the functions $F_i : \mathsf{U} \to \mathbb{R}$, $i = \overline{1,n}$ have partial derivatives at $p \in \mathsf{U}$. F is **differentiable** provided it is differentiable at each $p \in \mathsf{U}$. (For the case $m = 1$, we obtain the notion of a differentiable parametrized curve in Euclidean space $\mathbb{R}^n$.)

The class $C^r(\mathsf{U}, \mathbb{R}^n)$, $1 \le r \le \infty$ of C^r-mappings $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is defined in the obvious way. We will be concerned primarily with *smooth* (i.e., of class C^∞) mappings. So if F is a smooth mapping, then its component functions F_i have continuous partial derivatives of all orders and each such derivative is independent of the order of differentiation.

NOTE : Let us define a (geometric) **tangent vector** at $p \in \mathbb{R}^m$ as an ordered pair (p, v). As a matter of notation, we will abbreviate (p, v) as v_p. We think of v_p as the *vector* v with its *initial point* at p. (In other words, $p + v$ is considered as the "position vector" of a point; we shall always picture v_p as the "arrow" from the point p to the point $p + v$.) Clearly, two tangent vectors v_p and w_q are *equal* if $v = w$ and $p = q$. (It is essential to recognize that v_p and v_q are *different* tangent vectors if $p \neq q$.)

The set $\{p\} \times \mathbb{R}^m$ of all tangent vectors at p is denoted by $T_p \mathbb{R}^m$, and is called the **tangent space** of $\mathbb{R}^m$ at p. Thus

$$T_p \mathbb{R}^m := \{v_p = (p, v) \, : \, p, v \in \mathbb{R}^m\}.$$

This set is a *vector space* over $\mathbb{R}$ (obviously isomorphic to $\mathbb{R}^m$ itself) under the natural operations: $v_p + w_p := (v + w)_p$ and $\lambda \, v_p := (\lambda v)_p$. The tangent vectors $(e_i)_p$, $i = \overline{1,m}$ form a basis for $T_p \mathbb{R}^m$. (In fact, as a vector space, $T_p \mathbb{R}^m$ is essentially the same as $\mathbb{R}^m$ itself; the only reason we add T_p is so that the geometric tangent spaces $T_p \mathbb{R}^m$ and $T_q \mathbb{R}^m$ at distinct points p and q be disjoint sets.)

Let v_p be a tangent vector in $\mathbb{R}^m$. One can associate with it the function (parametrized line)

$$\mathbb{R} \ni t \mapsto p + tv \in \mathbb{R}^m.$$

If $F : \mathbb{R}^m \to \mathbb{R}$ is a differentiable function, then $t \mapsto F(p + tv)$ is an ordinary function $\mathbb{R} \to \mathbb{R}$. (The derivative of this function at $t = 0$ tells the *initial rate of change* of F as p moves in the v direction.) The number

$$v_p[F] := \left. \frac{d}{dt} F(p + tv) \right|_{t=0}$$

is called the **directional derivative** of F with respect to v_p. We have

$$v_p[F] = v_1 \frac{\partial F}{\partial x_1}(p) + \cdots + v_m \frac{\partial F}{\partial x_m}(p) \qquad (v = (v_1, \ldots, v_m) \in \mathbb{R}^m).$$

The map $v_p[\cdot] : C^\infty(\mathbb{R}^m) \to \mathbb{R}$, $F \mapsto v_p[F]$ is linear and satisfies the *Leibniz rule* (i.e., $v_p[FG] = v_p[F]G(p) + F(p)v_p[G]$ for $F, G \in C^\infty(\mathbb{R}^m)$); such a mapping is called a **derivation** at p. So *any geometric tangent vector* v_p *defines a derivation* $v_p[\cdot]$ *at* p. In fact, each derivation at p is defined by a *unique* geometric tangent vector (at p). Moreover, for any $p \in \mathbb{R}^m$, the correspondence $v_p \mapsto v_p[\cdot]$ is an isomorphism from the tangent space $T_p\,\mathbb{R}^m$ to the vector space of all derivations on p. It is customary (and convenient) to denote the derivation $(e_i)_p[\cdot]$ by $\frac{\partial}{\partial x_i}\Big|_p$; thus, $\frac{\partial}{\partial x_i}\Big|_p [F] = \frac{\partial F}{\partial x_i}(p)$.

Let $T_p\,\mathbb{R}^m$ be the tangent space to $\mathbb{R}^m$ at p; this vector space can be *identified* with $\mathbb{R}^m$ via

$$v_1 \,\frac{\partial}{\partial x_1}\bigg|_p + \cdots + v_m \,\frac{\partial}{\partial x_m}\bigg|_p \mapsto (v_1, \cdots, v_m).$$

Let $\alpha : \mathsf{U} \subseteq \mathbb{R} \to \mathbb{R}^m$ be a smooth (parametrized) curve with component functions $\alpha_1, \ldots, \alpha_m$. The **velocity vector** (or tangent vector) to α at $t \in \mathsf{U}$ is the element

$$\dot{\alpha}(t) := \left(\frac{d\alpha_1}{dt}(t), \cdots, \frac{d\alpha_m}{dt}(t)\right) \in T_{\alpha(t)}\,\mathbb{R}^m.$$

Example 54. Given a point $p \in \mathsf{U} \subseteq \mathbb{R}^m$ and a tangent vector $v \in T_p\,\mathbb{R}^m$, we can always find a smooth curve $\alpha : (-\varepsilon, \varepsilon) \to \mathsf{U}$ with $\alpha(0) = p$ and $\dot{\alpha}(0) = v$. Simply define $\alpha(t) = p + tv$, $t \in (-\varepsilon, \varepsilon)$. By writing $p = (p_1, \ldots, p_m)$ and $v = (v_1, \ldots, v_m)$, the component functions of α are $\alpha_i(t) = p_i + tv_i$, $i = \overline{1, m}$. Thus α is smooth, $\alpha(0) = p$ and

$$\dot{\alpha}(0) = \left(\frac{d\alpha_1}{dt}(0), \cdots, \frac{d\alpha_m}{dt}(0)\right) = (v_1, \ldots, v_m) = v.$$

We shall now introduce the concept of *derivative* (or differential) of a differentiable mapping. Let $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be a differentiable mapping. To each $p \in \mathsf{U}$ we associate a linear mapping

$$DF(p) : \mathbb{R}^m = T_p\,\mathbb{R}^m \to \mathbb{R}^n = T_{F(p)}\,\mathbb{R}^n$$

which is called the **derivative** (or *differential*) of F at p and is defined as follows. Let $v \in T_p\,\mathbb{R}^m$ and let $\alpha : (-\varepsilon, \varepsilon) \to \mathsf{U}$ be a differentiable curve such that $\alpha(0) = p$ and $\dot{\alpha}(0) = v$. By the chain rule (for functions), the curve $\beta = F \circ \alpha : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ is also differentiable. Then

$$DF(p) \cdot v := \dot{\beta}(0).$$

Note : The above definition of $DF(p)$ does not depend on the choice of the curve which passes through p with tangent vector v, and $DF(p)$ is, in fact, linear. So

$$DF(p) \cdot v = \frac{d}{dt} F(\alpha(t))\bigg|_{t=0} \in T_{F(p)}\,\mathbb{R}^n = \mathbb{R}^n.$$

The derivative $DF(p)$ is also denoted by $T_p F$ and called the *tangent mapping* of F at p.

The matrix of the linear mapping $DF(p)$ (relative to bases $\left(\left.\frac{\partial}{\partial x_1}\right|_p, \ldots, \left.\frac{\partial}{\partial x_m}\right|_p\right)$ of $T_p\,\mathbb{R}^m$ and $\left(\left.\frac{\partial}{\partial y_1}\right|_{F(p)}, \ldots, \left.\frac{\partial}{\partial y_n}\right|_{F(p)}\right)$ of $T_{F(p)}\,\mathbb{R}^n$) is the **Jacobian matrix**

$$\frac{\partial F}{\partial x}(p) = \frac{\partial(F_1, \ldots, F_n)}{\partial(x_1, \ldots, x_m)}(p) := \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_1}{\partial x_m}(p) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(p) & \cdots & \frac{\partial F_n}{\partial x_m}(p) \end{bmatrix} \in \mathbb{R}^{n\times m}$$

of F at p. When $m = n$ this is a square matrix and its determinant is then defined. This determinant is called the **Jacobian** of F at p and is denoted by $J_F(p)$. Thus

$$J_F(p) = \left|\frac{\partial F}{\partial x}(p)\right| := \det \frac{\partial F}{\partial x}(p)\cdot$$

$\diamond$ **Exercise 41.** Let $f : \mathsf{I} \to \mathbb{R}$ and $g : \mathsf{J} \to \mathbb{R}$ be differentiable functions, where I and J are open intervals such that $f(\mathsf{I}) \subseteq \mathsf{J}$. Show that the function $g \circ f$ is differentiable and (for $t \in \mathsf{I}$)

$$(g \circ f)'(t) = g'(f(t)) \cdot f'(t).$$

The standard *chain rule* (for scalar-valued) functions extends to (vector-valued) mappings.

PROPOSITION 30 (GENERAL CHAIN RULE). *Let $F : \mathsf{U} \subseteq \mathbb{R}^\ell \to \mathbb{R}^m$ and $G : \mathsf{V} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be differentiable mappings, where U and V are open sets such that $F(\mathsf{U}) \subseteq \mathsf{V}$. Then $G \circ F$ is a differentiable mapping and (for $p \in \mathsf{U}$)*

$$D(G \circ F)(p) = DG(F(p)) \circ DF(p).$$

Proof. The fact that $G\circ F$ is differentiable is a consequence of the chain rule for functions. Now, let $v \in T_p\,\mathbb{R}^\ell$ be given and let us consider a (differentiable) curve $\alpha : (-\varepsilon, \varepsilon) \to \mathsf{U}$ with $\alpha(0) = p$ and $\dot\alpha(0) = v$. Set $DF(p) \cdot v = w$ and observe that

$$DG(F(p)) \cdot w = \left.\frac{d}{dt}(G \circ F \circ \alpha)\right|_{t=0}.$$

Then

$$\begin{aligned} D(G \circ F)(p) \cdot v &= \left.\frac{d}{dt}(G \circ F \circ \alpha)\right|_{t=0} \\ &= DG(F(p)) \cdot w \\ &= DG(F(p)) \circ DF(p) \cdot v. \end{aligned}$$

□

NOTE : In terms of Jacobian matrices, the general chain rule can be written

$$\frac{\partial (G \circ F)}{\partial x}(p) = \frac{\partial G}{\partial y}(F(p)) \cdot \frac{\partial F}{\partial x}(p)\cdot$$

Thus if $H = G \circ F$ and $y = F(x)$, then

$$\frac{\partial H}{\partial x} = \begin{bmatrix} \frac{\partial G_1}{\partial y_1} & \cdots & \frac{\partial G_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial G_n}{\partial y_1} & \cdots & \frac{\partial G_n}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_\ell} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_\ell} \end{bmatrix}$$

where $\dfrac{\partial G_1}{\partial y_1}, \dots, \dfrac{\partial G_n}{\partial y_m}$ are evaluated at $y = F(x)$ and $\dfrac{\partial F_1}{\partial x_1}, \cdots, \dfrac{\partial F_m}{\partial x_\ell}$ at x. Writing this out, we obtain

$$\frac{\partial H_i}{\partial x_j} = \frac{\partial G_i}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial G_i}{\partial y_m}\frac{\partial y_m}{\partial x_j} \qquad (i = \overline{1,n}\,;\ j = \overline{1,\ell}).$$

◇ **Exercise 42.** Let

$$F(x_1, x_2) = (x_1^2 - x_2^2 + x_1x_2, x_2^2 - 1) \quad \text{and} \quad G(y_1, y_2) = (y_1 + y_2, 2y_1, y_2^2).$$

(a) Show that F and G are differentiable, and that $G \circ F$ exists.

(b) Compute $D(G \circ F)(1, 1)$

  (i) directly

  (ii) using the chain rule.

Note : The precise sense in which the derivative $DF(p)$ of the (differentiable) mapping F at p is an (affine) approximation of F near p is given by the following result (in which $DF(p)$ is interpreted as a linear mapping from $\mathbb{R}^m$ to $\mathbb{R}^n$) : *If the mapping* $F : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}^n$ *is differentiable, then for each* $p \in \mathsf{U}$,

$$\lim_{x \to p} \frac{\|F(x) - F(p) - DF(p) \cdot (x - p)\|}{\|x - p\|} = 0$$

or, equivalently, there exists a (local) map $\epsilon_p : \mathbb{R}^m \to \mathbb{R}^n$ *satisfying, for all* h *with* $p + h \in \mathsf{U}$,

$$(5) \qquad F(p+h) = F(p) + DF(p) \cdot h + \epsilon_p(h) \quad \textit{with} \quad \lim_{h \to 0} \frac{\|\epsilon_p(h)\|}{\|h\|} = 0.$$

The mapping $\mathbb{R}^m \to \mathbb{R}^n$, $x \mapsto F(p) + DF(p) \cdot (x - p)$ is the *best affine approximation* to F at p. (It is the unique affine approximation for which the difference mapping ϵ_p satisfies the estimate (5).)

If $\mathsf{A} \subseteq \mathbb{R}^m$ is an *arbitrary* set, then $C^\infty(\mathsf{A})$ denotes the set of all functions $F : \mathsf{A} \to \mathbb{R}$ such that $F = \overline{F}\big|_{\mathsf{A}}$, where $\overline{F} : \mathcal{U} \to \mathbb{R}$ is a smooth function on some open neighborhood $\mathcal{U}$ of A.

Problems (11–15)

(11) Let $x, y \in \mathbb{R}^m$. Prove the following inequalities.

  (a) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality).

  (b) $|\,\|x\| - \|y\|\,| \leq \|x - y\|$ (reverse triangle inequality).

  (c) $|x_i| \leq \|x\| \leq |x_1| + \cdots + |x_m| \leq \sqrt{m}\,\|x\|, \quad i = \overline{1,m}.$

(12) Let $\tau : \mathbb{R}^m \to \mathbb{R}^m$ be a linear transformation, and let $A \in \mathbb{R}^{m\times m}$ denote its matrix with respect to the standard basis of $\mathbb{R}^m$. Show that the following statements are logically equivalent.
   (a) $\|\tau(x)\| = \|x\|$ for all $x \in \mathbb{R}^m$.
   (b) $\tau(x) \bullet \tau(y) = x \bullet y$ for all $x, y \in \mathbb{R}^m$.
   (c) $A^\top A = \mathbf{1}$ (i.e., the matrix A is orthogonal).

   (Such a linear transformation is called an **orthogonal transformation**.) Hence deduce that such a linear transformation τ is invertible. Is τ^{-1} of the same sort?

(13) Let F be a subset of $\mathbb{R}^m$. Show that the following statements are logically equivalent.
   (a) F is closed.
   (b) $\mathsf{F} = \mathrm{cl}\,(\mathsf{F})$.
   (c) For every sequence $(x_k)_{k\in\mathbb{N}}$ of points $x_k \in \mathbb{R}^m$ that is convergent to a limit, say p, we have $p \in \mathsf{F}$.

(14) Let $\mathsf{A} \subseteq \mathbb{R}^m$ be an arbitrary set.
   (a) Show that the following statements are logically equivalent.
      (i) A is disconnected.
      (ii) There exists a surjective continuous function $\mathsf{A} \to \{0, 1\}$.

      (Recall the definition of the **characteristic function** χ_{A} of a set A: $\chi_{\mathsf{A}}(x) = 1$ if $x \in \mathsf{A}$ and $\chi_{\mathsf{A}}(x) = 0$ if $x \notin \mathsf{A}$.)
   (b) Assume that A is connected and let $F : \mathsf{A} \to \mathbb{R}^n$ be a continuous mapping. Show that $F(\mathsf{A})$ is connected in $\mathbb{R}^n$.
   (c) Let A be connected and let $F : \mathsf{A} \to \mathbb{R}$ be a continuous function. Show that $F(\mathsf{A})$ is an interval in $\mathbb{R}$; in particular, F takes all the values between any two that it assumes.

(15) Show that
      (i) if $\sigma : \mathbb{R}^2 \to \mathbb{R}$ is defined by $\sigma(x, y) = x + y$, then $D\sigma(a, b) = \sigma$.
      (ii) if $\pi : \mathbb{R}^2 \to \mathbb{R}$ is defined by $\pi(x, y) = x \cdot y$, then $D\pi(a, b) \cdot (x, y) = bx + ay$.

   Hence deduce that if the functions $F, G : \mathsf{U} \subseteq \mathbb{R}^m \to \mathbb{R}$ are differentiable at $p \in \mathsf{U}$, then

$$\begin{aligned} D(F + G)(p) &= DF(p) + DG(p) \\ D(F \cdot G)(p) &= G(p)DF(p) + F(p)DG(p). \end{aligned}$$

   If moreover $G(p) \neq 0$, then

$$D\left(\frac{F}{G}\right)(p) = \frac{G(p)DF(p) - F(p)DG(p)}{(G(p))^2}.$$