

Chapter 1

Problem 1.1. Prove that $|x| \leq \sum_{i=1}^n |x^i|$.

Proof. On the face of it, this seems to resemble the triangle inequality. However, notice that the $|\cdot|$ is a vector magnitude and the same symbol on the right represents an absolute value of a real number. However, we can use the triangle inequality here:

$$|x|^2 = |x^1 e_1 + x^2 e_2 + \cdots + x^n e_n|^2$$

Where we have expanded x in the usual basis of \mathbb{R}^n . From the triangle inequality applied to this vector sum, we have

$$|x|^2 \leq \left(\sum_{i=1}^n |x^i| |e_i| \right)^2 = \left(\sum_{i=1}^n |x^i| \right)^2$$

Taking the square root, the statement is obtained. ■

Problem 1.3. Prove that $|x - y| \leq |x| + |y|$. When does equality hold?

Proof. Let $x, y \in \mathbb{R}^n$. Then we have

$$|x - y| = |x + (-y)| \leq |x| + |-y| = |x| + |y|$$
■

Problem 1.4. Prove that $||x| - |y|| \leq |x - y|$.

Proof. This is commonly referred to as the reverse triangle inequality. It is rather easily seen as a corollary to the standard triangle inequality. Let $x, y \in \mathbb{R}^n$. Then note that

$$|x| = |(x - y) + y| \leq |x - y| + |y|$$

Upon subtraction, we obtain

$$|x| - |y| \leq |x - y|$$

Repeating this but starting with y , we obtain the similar

$$-(|x| - |y|) \leq |x - y|$$

Taken together, these last two results give

$$||x| - |y|| \leq |x - y|$$
■

Problem 1.5. The quantity $|y - x|$ is called the distance between x and y . Prove and interpret geometrically the "triangle inequality" $|z - x| \leq |z - y| + |y - x|$.

Proof. Let $x, y, z \in \mathbb{R}^n$. Then

$$\begin{aligned} |z - x| &= |(z - y) + (y - x)| \\ &\leq |z - y| + |y - x| \end{aligned}$$

Note that the trick of simultaneously adding and subtracting a quantity — a net addition of 0 — is a standard trick that occurs quite often in mathematics.

The moniker “triangle inequality” is obvious if we interpret x , y , and z as vertices of a triangle. Then $|z - x|$, $|z - y|$, and $|y - x|$ are the three pairwise distances between these points, i.e., the lengths of the sides of the triangle. This statement says that the length of any side of a triangle is *at most* the sum of the other two sides. This makes sense if we just think about constructing triangles; it seems quite impossible to create a triangle with one side longer than both other sides combined. ■

Problem 1.7. A linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is norm preserving if $|T(x)| = |x|$, and inner product preserving if $\langle T(x), T(y) \rangle = \langle x, y \rangle$.

- Prove that T is norm preserving if and only if T is inner product preserving.
- Prove that such a linear transformation T is 1-1 and T^{-1} is of the same sort.

Proof. • Let T be norm preserving. Then, by the polarization identity,

$$\begin{aligned} \langle T(x), T(y) \rangle &= \frac{|T(x) + T(y)|^2 - |T(x) - T(y)|^2}{4} \\ &= \frac{|T(x + y)|^2 - |T(x - y)|^2}{4} \\ &= \frac{|x + y|^2 - |x - y|^2}{4} \\ &= \langle x, y \rangle \end{aligned}$$

hence it is inner product preserving.

Let T be inner product preserving. Then

$$|T(x)|^2 = \langle T(x), T(x) \rangle = \langle x, x \rangle = |x|^2$$

and so it is norm preserving.

- Let T be such a linear transformation. To show that it is injective, let $x, y \in \mathbb{R}^n$ such that $T(x) = T(y)$. Then we want to show that x is equal to y . Thus, consider the distance between these points:

$$|x - y| = |T(x - y)| = |T(x) - T(y)| = 0$$

by the norm preserving property and linearity of T . Hence, $x = y$ and T is injective. Let $u \in \mathbb{R}^n$ be some element in the range of T . Then since T is injective, we can find some $x \in \mathbb{R}^n$ such that $T(x) = u$, i.e., $x = T^{-1}(u)$ (basically, T is surjective on its range). We want to consider the properties of T^{-1} (we are looking at a specific point where we are saying it exists — is there some general theorem about linear injective maps which guarantees its existence?). Well,

$$|u| = |T \circ T^{-1}(u)| = |T(T^{-1}(u))| = |T^{-1}(u)|$$

and we are done. ■

Problem 1.8. If $x, y \in \mathbb{R}^n$ are non-zero, the angle between x and y , denoted $\angle(x, y)$, is defined as $\arccos(\langle x, y \rangle / |x| \cdot |y|)$, which makes sense by Theorem 1-1 (2). The linear transformation T is angle preserving if T is 1-1, and for $x, y \neq 0$ we have $\angle(Tx, Ty) = \angle(x, y)$.

- Prove that if T is norm preserving, then T is angle preserving.
- If there is a basis x_1, \dots, x_n of \mathbb{R}^n and numbers $\lambda_1, \dots, \lambda_n$ such that $Tx_i = \lambda_i x_i$, prove that T is angle preserving if and only if all $|\lambda_i|$ are equal.
- What are all angle preserving $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$?

Proof. • Let T be norm preserving. Then we know from the previous problem that it is also inner product preserving. Hence,

$$\begin{aligned} \angle(Tx, Ty) &= \arccos(\langle Tx, Ty \rangle / |Tx| \cdot |Ty|) \\ &= \arccos(\langle x, y \rangle / |x| \cdot |y|) \\ &= \angle(x, y) \end{aligned}$$

Hence T is angle preserving.

- Let $\{x_i\}_{i=1}^n$ be such a basis and $\{\lambda_i\}_{i=1}^n$ be the numbers in the hypothesis. ■

Problem 1.9. If $0 \leq \theta < \pi$, let $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ have the matrix $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$. Show that T is angle preserving and if $x \neq 0$, then $\angle(x, Tx) = \theta$.

Proof. We could show that this matrix is angle preserving through brute force application of the definition: let $x, y \in \mathbb{R}^2$ arbitrary and show that $\langle Tx, Ty \rangle$ and $\langle x, y \rangle$ return the same value. Alternately, we can consider just a single vector $x \in \mathbb{R}^n$ and show that T is norm preserving, since we know from the problem 1-7 that these are equivalent.

Let $x = (x_1, x_2)^T \in \mathbb{R}^2$. Then

$$\begin{aligned} |T(x)|^2 &= (x_1 \cos \theta + x_2 \sin \theta)^2 + (-x_1 \sin \theta + x_2 \cos \theta)^2 \\ &= (\cos^2 \theta + \sin^2 \theta) x_1^2 + (\cos^2 \theta + \sin^2 \theta) x_2^2 + 2(\cos \theta \sin \theta - \cos \theta \sin \theta) x_1 x_2 \\ &= x_1^2 + x_2^2 \\ &= |x|^2 \end{aligned}$$

■ Thus, T is norm preserving and so also angle preserving.

To look at the angle between x and $T(x)$, we must look at the arc cosine of the following quantity:

$$\begin{aligned} \frac{\langle x, Tx \rangle}{|x| |Tx|} &= \frac{x_1^2 \cos \theta + (\sin \theta - \sin \theta) x_1 x_2 + x_2^2 \cos \theta}{|x|^2} \\ &= \frac{(x_1^2 + x_2^2) \cos \theta}{|x|^2} \\ &= \cos \theta \end{aligned}$$

Hence, the angle between x and Tx is

$$\angle(x, Tx) = \arccos \left(\frac{\langle x, Tx \rangle}{|x| |Tx|} \right) = \arccos(\cos \theta) = \theta$$

Problem 1.10. If $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a linear transformation, show that there is a number M such that $|T(h)| \leq M|h|$ for $h \in \mathbb{R}^m$. *Hint:* Estimate $|T(h)|$ in terms of $|h|$ and the entries in the matrix of T .

Proof. The statement is intuitively obvious: it makes sense that a linear map between two finite-dimensional vector spaces cannot just take a vector anywhere, but instead is bounded above by some scalar multiple of (the magnitude of) that vector. Keep in mind that we can write each component of the matrix-vector product $T(h) = Th$ as a sum (inner product) of a particular row of T with h :

$$[Th]_j = T_{j*}h = \sum_{i=1}^m T_{ji}h_i$$

Taking the Euclidean magnitude of this product, we have

$$\begin{aligned} |T(h)| &= \left(\sum_{j=1}^n [Th]_j^2 \right)^{1/2} \\ &= \left(\sum_{j=1}^n \left(\sum_{i=1}^m T_{ji}h_i \right)^2 \right)^{1/2} \\ &\leq \left(\sum_{j=1}^n \left(\sum_{i=1}^m |T_{ji}| |h_i| \right)^2 \right)^{1/2} \end{aligned}$$

Note that $|h|^2 = |h_1|^2 + \cdots + |h_m|^2$ and so $|h_i| < |h|$ for each $i = 1, \dots, m$. Also, define $\tilde{T}_j = \max \{|T_{j1}|, |T_{j2}|, \dots, |T_{jm}|\}$; that is, \tilde{T}_j is the maximum magnitude of the entries in the j th row of T . Then

$$\begin{aligned} |T(h)| &\leq \left(\sum_{j=1}^n \left(\sum_{i=1}^m \tilde{T}_j |h_i| \right)^2 \right)^{1/2} \\ &\leq \left(\sum_{j=1}^n m^2 |h|^2 \tilde{T}_j^2 \right)^{1/2} \end{aligned}$$

Since both multiplicands in the innermost sum are constant over the index of summation. Similarly, we can define $\tilde{T} = \max \{\tilde{T}_1, \tilde{T}_2, \dots, \tilde{T}_n\}$; that is, \tilde{T} is the maximum entry magnitude in T . Then

$$\begin{aligned} |T(h)| &\leq \left(\sum_{j=1}^n m^2 |h|^2 \tilde{T}^2 \right)^{1/2} \\ &= m\sqrt{n}\tilde{T} |h| \\ &= M |h| \end{aligned}$$

where $M = m\sqrt{n}\tilde{T}$ is a constant involving only the dimensions of the mapping and the entries of T . Hence, the statement is proven. ■

Problem 1.11. If $x, y \in \mathbb{R}^n$ and $z, w \in \mathbb{R}^m$, show that $\langle (x, z), (y, w) \rangle = \langle x, y \rangle + \langle z, w \rangle$ and $|(x, z)| = \sqrt{|x|^2 + |z|^2}$. Note that (x, z) and (y, w) denote points in \mathbb{R}^{n+m} .

Proof. This problem shows that the inner product and the norm behave together well (note that the Euclidean 2-norm is actually induced by the standard Euclidean inner product; this is converse to the exposition provided in this book). First, note that the i th component of (x, z) is

$$(x, z)^i = \begin{cases} x^i, & i = 1, 2, \dots, n \\ z^{i-n}, & i = n+1, n+2, \dots, n+m \end{cases}$$

and similarly for (y, w) .

$$\begin{aligned} \langle (x, z), (y, w) \rangle &= \sum_{i=1}^{n+m} (x, z)^i (y, w)^i \\ &= \sum_{i=1}^n x^i y^i + \sum_{i=n+1}^{n+m} z^{i-n} w^{i-n} \\ &= \sum_{i=1}^n x^i y^i + \sum_{j=1}^m z^j w^j \\ &= \langle x, y \rangle + \langle z, w \rangle \end{aligned}$$

where a change of summation index was used to simplify the second summation. Hence the first statement holds.

Next, we can use the same trick:

$$\begin{aligned} |(x, z)|^2 &= \sum_{i=1}^{n+m} \left((x, z)^i \right)^2 \\ &= \sum_{i=1}^n (x^i)^2 + \sum_{i=n+1}^{n+m} (z^{i-n})^2 \\ &= \sum_{i=1}^n (x^i)^2 + \sum_{j=1}^m (z^j)^2 \\ &= |x|^2 + |z|^2 \end{aligned}$$

and, taking the square root of both sides, the statement holds. ■

Problem 1.12. Let $(\mathbb{R}^n)^*$ denote the dual space of the vector space \mathbb{R}^n . If $x \in \mathbb{R}^n$, define $\phi_x \in (\mathbb{R}^n)^*$ by $\phi_x(y) = \langle x, y \rangle$. Define $T : \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*$ by $T(x) = \phi_x$. Show that T is a 1-1 linear transformation and conclude that every $\phi \in (\mathbb{R}^n)^*$ is ϕ_x for a unique $x \in \mathbb{R}^n$.

Proof. Given a point $x \in \mathbb{R}^n$, we can define a dual vector space element ϕ_x which is (by definition of the dual space) a linear function from \mathbb{R}^n to \mathbb{R} . The problem statement defines a mapping T which actually performs this assignment, i.e., it takes a point in $x \in \mathbb{R}^n$ and returns the dual vector $T(x) = \phi_x$. We want to show that this mapping T assigning dual vectors to vectors in the obvious way, through the inner product, is linear and injective.

First, let $x, y \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$. Then T gives

$$T(\alpha x + \beta y) = \phi_{\alpha x + \beta y}$$

which is, itself, a dual vector, mapping \mathbb{R}^n to \mathbb{R} linearly, through the inner product construction. To further analyze it, let's see how it operates on an arbitrary element $z \in \mathbb{R}^n$:

$$\begin{aligned} T(\alpha x + \beta y)(z) &= \phi_{\alpha x + \beta y}(z) \\ &= \langle \alpha x + \beta y, z \rangle \\ &= \alpha \langle x, z \rangle + \beta \langle y, z \rangle \\ &= \alpha \phi_x(z) + \beta \phi_y(z) \\ &= \alpha T(x)(z) + \beta T(y)(z) \\ &= [\alpha T(x) + \beta T(y)](z) \end{aligned}$$

where we used the linearity of the inner product and the definitions of ϕ and T . Thus, since $z \in \mathbb{R}^n$ arbitrary, we conclude from this last that T is linear.

To prove injectivity, let $x, y \in \mathbb{R}^n$ be two elements in the domain of T such that $T(x) = T(y)$. We must show that $x = y$. Note that $T(x) = \phi_x$ and $T(y) = \phi_y$, both

elements of the dual vector space $(\mathbb{R}^n)^*$. Since these linear maps (dual vectors) are equal, they must agree on every vector in \mathbb{R}^n ; that is,

$$\phi_x(z) = \phi_y(z) \quad \forall z \in \mathbb{R}^n$$

In particular, they must agree on the standard Euclidean basis. Let $e_i = (0, \dots, 1, \dots, 0)^T$ be the i th basis vector, all zeroes except for the i th component being 1. Then

$$\phi_x(e_i) = \langle x, e_i \rangle = x^i$$

and

$$\phi_y(e_i) = \langle y, e_i \rangle = y^i$$

Since these must be equal, we have $x^i = y^i$. Repeating this for every $i = 1, 2, \dots, n$, we see that $x = y$. Thus, T is injective.

The remaining part of the problem is to show that this construction is more general, in the sense that any dual vector in $(\mathbb{R}^n)^*$ is one of these ϕ corresponding to some $x \in \mathbb{R}^n$. That is, any dual vector is equivalent to an inner product with some element of \mathbb{R}^n , and that this particular element is unique.

Let $\phi \in (\mathbb{R}^n)^*$, and $z \in \mathbb{R}^n$. Recall that we know how a linear map acts on an arbitrary vector if we know how it acts on the basis vectors. This is due to linearity: writing $z = \sum_{i=1}^n z^i e_i$, we have

$$\begin{aligned} \phi(z) &= \phi\left(\sum_{i=1}^n z^i e_i\right) \\ &= \sum_{i=1}^n z^i \phi(e_i) \end{aligned}$$

Note that each of the values $\phi(e_i)$ are constant, independent of the $z \in \mathbb{R}^n$ we are operating on. This summation is of the form of a Euclidean inner product. Specifically,

$$\phi(z) = \langle x, z \rangle = \phi_x(z)$$

where $x = (\phi(e_1), \phi(e_2), \dots, \phi(e_n))$. Thus we have identified an $x \in \mathbb{R}^n$ giving rise to our arbitrary dual vector ϕ through this inner product construction. Uniqueness of this x follows as before, by operating this dual vector on the standard basis vectors. ■

Problem 1.13. If $x, y \in \mathbb{R}^n$, then x and y are called perpendicular (or orthogonal) if $\langle x, y \rangle = 0$. If x and y are perpendicular, prove that $|x + y|^2 = |x|^2 + |y|^2$.

Proof. Since the definition of orthogonality requires the inner product, and the target of our prove involves norms, we can easily prove this proposition by expressing our norm of interest in terms of inner products:

$$\begin{aligned} |x + y|^2 &= \langle x + y, x + y \rangle \\ &= \langle x \rangle + 2 \langle x, y \rangle + \langle y, y \rangle \\ &= |x|^2 + 2 \cdot 0 + |y|^2 \\ &= |x|^2 + |y|^2 \end{aligned}$$

■

Problem 1.23. If $f : A \rightarrow \mathbb{R}^m$ and $a \in A$, show that $\lim_{x \rightarrow a} f(x) = b$ if and only if $\lim_{x \rightarrow a} f^i(x) = b^i$ for $i = 1, \dots, m$.

Proof. Assume first that $\lim_{x \rightarrow a} f(x) = b$. Let $\epsilon > 0$. By definition of limit, there exists some $\delta' > 0$ such that $|x - a| < \delta'$ implies $|f(x) - b| < \epsilon$. Since

$$|f(x) - b|^2 = \sum_{i=1}^n |f^i(x) - b^i|^2$$

which is a sum of positive numbers. Hence,

$$|f^i(x) - b^i| \leq |f(x) - b|$$

for every $i = 1, 2, \dots, n$. Thus we have

$$|x - a| < \delta' \implies |f^i(x) - b^i| < \epsilon$$

That is, $\lim_{x \rightarrow a} f^i(x) = b^i$.

Next, assume $\lim_{x \rightarrow a} f^i(x) = b^i$ for all $i = 1, 2, \dots, n$. Let $\epsilon > 0$ be arbitrary. Then for any i , there exists some $\delta_i > 0$ such that

$$|x - a| < \delta_i \implies |f^i(x) - b^i| \leq \frac{\epsilon}{\sqrt{n}}$$

Let $\delta = \min \{\delta_1, \delta_2, \dots, \delta_n\}$. Then we have, for $|x - a| < \delta$,

$$|f(x) - b|^2 = \sum_{i=1}^n |f^i(x) - b^i|^2 \leq \sum_{i=1}^n \left(\frac{\epsilon}{\sqrt{n}} \right)^2 = n \frac{\epsilon^2}{n} = \epsilon^2$$

Taking the square root of both sides (both positive), we have shown the existence of a $\delta > 0$ such that

$$|x - a| < \delta \implies |f(x) - b| < \epsilon$$

That is, $\lim_{x \rightarrow a} f(x) = b$.

■

Problem 1.24. Prove that $f : A \rightarrow \mathbb{R}^m$ is continuous at a if and only if each f^i is.

Proof. This proposition readily holds from the previous problem, 1-23, if we replace b with $f(a)$.

■

Problem 1.25. Prove that a linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuous. *Hint:* Use Problem 1-10.

Proof. Recall that Problem 1-10 allowed us to bound a linear transformation above in the sense that we can find a number M such that

$$|T(h)| \leq M|h| \text{ for } h \in \mathbb{R}^n$$

Let $a \in \mathbb{R}^n$ be arbitrary. We want to show that T is continuous at a , i.e., that $\lim_{x \rightarrow a} T(x) = T(a)$. Let $\epsilon > 0$, arbitrary, and let M be the (positive) number guaranteed by Problem 1-10. Then notice that

$$|T(x) - T(a)| = |T(x - a)| \leq M|x - a|$$

. With this in hand, let $\delta = \epsilon/M$. Then we have

$$|x - a| < \delta \implies |T(x) - T(a)| < M \frac{\epsilon}{M} = \epsilon$$

Thus we have shown that $\lim_{x \rightarrow a} T(x) = T(a)$, that is, T is continuous at a . Since $a \in \mathbb{R}^n$ arbitrary, T is continuous everywhere. ■

Problem 1.26. Let $A = \{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } 0 < y < x^2\}$.

- Show that every straight line through $(0, 0)$ contains an interval around $(0, 0)$ which is in $\mathbb{R}^2 - A$.
- Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $f(x) = 0$ if $x \notin A$ and $f(x) = 1$ if $x \in A$. For $h \in \mathbb{R}^2$ define $g_h : \mathbb{R} \rightarrow \mathbb{R}$ by $g_h(t) = f(th)$. Show that each g_h is continuous at 0, but f is not continuous at $(0, 0)$.

Proof. • hi ■

Problem 1.27. Prove that $\{x \in \mathbb{R}^n : |x - a| < r\}$ is open by considering the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $f(x) = |x - a|$.

Proof. Notice that the set of interest is the open ball of radius r about a . The function f suggested to us takes a point x and returns the distance of x to a . The dependence of the set inclusion on this distance and the function's computation of this distance hint that the two are related somehow. In fact, notice that if we let $D = \{x \in \mathbb{R}^n : |x - a| < r\}$, then $f(D) = [0, r)$. This is clearly not an open set, but notice that $[0, r) \subset (-\infty, r)$ and that $f^{-1}((-\infty, r)) = D$. Hence, if we can show that f is continuous, then D is automatically open, by (the topological) definition of continuity.

Let $z \in \mathbb{R}^n$ be arbitrary. Then we want to show that $\lim_{x \rightarrow z} f(x) = f(z)$. Let $\epsilon > 0$ be arbitrary as well. Then define $\delta = \epsilon$, and we have, for $|x - z| < \delta$,

$$\begin{aligned} |f(x) - f(z)| &= ||x - a| - |z - a|| \\ &\leq |x - a - z + a| \\ &= |x - z| \\ &< \delta \\ &= \epsilon \end{aligned}$$

where the first inequality is from the reverse triangle inequality. Hence, we have that f is continuous at every point in \mathbb{R}^n , and so our conclusion that D is open holds by the (more general topological) definition of continuity: open sets are mapped to by open sets under a continuous map. ■

Problem 1.29. If A is compact, prove that every continuous function $f : A \rightarrow \mathbb{R}$ takes on a maximum and a minimum value.

Proof. Let A be compact and $f : A \rightarrow \mathbb{R}$ continuous. Then we know that the continuous image of a compact set is compact (Theorem 1-9 in the book), and so $f(A)$ is compact. We also know that, in \mathbb{R}^n , the compact sets are exactly the closed and bounded sets (Corollary 1-7 and Problem 1-20), and so $f(A)$ is closed and bounded.

Being closed and bounded, $f(A) \subset \mathbb{R}$ contains its maximum and its minimum values (for example, if it did not, then the maximum value would be a supremum outside the set, and the set would be open). Hence, the statement holds. ■

Proof. Assume by way of contradiction that $f(A)$ does not contain a maximum or a minimum value. Then $f(A)$ is unbounded, but this is a contradiction since the continuous image of a compact set is compact, i.e., closed and bounded. Thus, $f(A)$ must contain its maximum or minimum value. We can do this independently for both the maximum and minimum, obtaining both. ■

Problem 1.30. Let $f : [a, b] \rightarrow \mathbb{R}$ be an increasing function. If $x_1, \dots, x_n \in [a, b]$ are distinct, show that $\sum_{i=1}^n o(f, x_i) < f(b) - f(a)$.

Proof. Let the hypotheses hold. By definition, the oscillation of f at any point x_i is

$$o(f, x_i) = \lim_{\delta \rightarrow 0} [M(a, f, \delta) - m(a, f, \delta)]$$

where M and m are the supremum and infimum of f , respectively, over the x such that $|x - x_i| < \delta$. Assume that we have ordered the $\{x_i\}_{i=1}^n$ in increasing order. Then the sum in question is

$$\begin{aligned} \sum_{i=1}^n o(f, x_i) &= \sum_{i=1}^n \lim_{\delta \rightarrow 0} [M(x_i, f, \delta) - m(x_i, f, \delta)] \\ &= \end{aligned}$$

■

Chapter 2

Problem 2.1. Prove that if $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then it is continuous at a . *Hint:* Use Problem 1-10.

Proof. Recall that Problem 1-10 allowed us to bound a linear map above by a function of its matrix entries. Specifically, for linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, we have a positive number M such that

$$|T(h)| \leq M|h|$$

for all $h \in \mathbb{R}^n$. This may come in handy when talking about derivatives, since a derivative (evaluated at a point) is a linear map.

Assume $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$. That is, there exists linear map $Df(a)$ such that

$$\lim_{h \rightarrow 0} \frac{|f(h+a) - f(a) - Df(a)(h)|}{|h|} = 0$$

The definition of continuity at a involves the limit of the quantity $|f(h+a) - f(a)|$, and since we know f is differentiable at a , we can likely use this definition of differentiability.

Notice that

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{|f(h+a) - f(a)|}{|h|} &= \lim_{h \rightarrow 0} \frac{|f(h+a) - f(a) - Df(a)(h) + Df(a)(h)|}{|h|} \\ &\leq \lim_{h \rightarrow 0} \frac{|f(h+a) - f(a) - Df(a)(h)|}{|h|} + \lim_{h \rightarrow 0} \frac{|Df(a)(h)|}{|h|} \end{aligned}$$

Notice that the first term on the right-hand side is the limit in the definition of the derivative of f at a , and since the derivative exists, this limit is 0. If we let M be the number guaranteed to exist by Problem 1-10 for the linear map $Df(a)$, then we have

$$\lim_{h \rightarrow 0} \frac{|f(h+a) - f(a)|}{|h|} \leq 0 + \lim_{h \rightarrow 0} \frac{M|h|}{|h|} = M$$

We are almost there. Looking at the limit of interest, we have

$$\begin{aligned} \lim_{h \rightarrow 0} |f(h+a) - f(a)| &= \lim_{h \rightarrow 0} \frac{|f(h+a) - f(a)|}{|h|} |h| \\ &= \left(\lim_{h \rightarrow 0} \frac{|f(h+a) - f(a)|}{|h|} \right) \left(\lim_{h \rightarrow 0} |h| \right) \\ &\leq M \cdot 0 \\ &= 0 \end{aligned}$$

and thus $\lim_{h \rightarrow 0} |f(h+a) - f(a)| = 0$, and so f is continuous at $a \in \mathbb{R}^n$ if it is differentiable there. ■

Problem 2.2. A function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is independent of the second variable if for each $x \in \mathbb{R}$ we have $f(x, y_1) = f(x, y_2)$ for all $y_1, y_2 \in \mathbb{R}$. Show that f is independent of the second variable if and only if there is a function $g : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(x, y) = g(x)$. What is $f'(a, b)$ in terms of g' ?

Proof. Assume f is independent of the second variable. Then we can define $g : \mathbb{R} \rightarrow \mathbb{R}$ as $g(x) = f(x, 0)$. Clearly, for any $y \in \mathbb{R}$, we have $f(x, y) = f(x, 0) = g(x)$, where the first equality follows from the independence of the second variable of f .

For the other direction, assume there is a g such that $f(x, y) = g(x)$ for all $x, y \in \mathbb{R}$. Then let $y_1, y_2 \in \mathbb{R}$, arbitrary. Then we have

$$f(x, y_1) = g(x) = f(x, y_2)$$

and so f is independent of the second variable.

We wish to determine the derivative of such an f in terms of the univariate function g . Our experience with partial derivatives from calculus 3 tells us that we should expect the derivative with respect to the second variable is zero, since the function doesn't change as we vary it. Also, the derivative with respect to the first variable is likely the derivative of the univariate function. In all, we guess that

$$f'(a, b) = (g'(a), 0)^T$$

Let us use this in the definition of derivative to verify that we obtain a zero limit:

$$\begin{aligned} \lim_{(h,k) \rightarrow (0,0)} \frac{|f(a+h, b+k) - f(a, b) - f'(a, b)(h, k)|}{|(h, k)|} &= \lim_{(h,k) \rightarrow (0,0)} \frac{|g(a+h) - g(a) - g'(a)(h)|}{|(h, k)|} \\ &\leq \lim_{(h,k) \rightarrow (0,0)} \frac{|g(a+h) - g(a) - g'(a)(h)|}{|h|} \\ &= 0 \end{aligned}$$

where the last equality holds if g is differentiable at a . Thus, the original difference quotient must be zero, confirming that the derivative of f at $(a, b)^T$ is, indeed, $(g'(a), 0)^T$. ■

Problem 2.3. Define when a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is independent of the first variable and find $f'(a, b)$ for such f . Which functions are independent of the first variable and also of the second variable?

Proof. This is quite obvious with some calculus 3 intuition. A function f is independent of the first variable if $f(x_1, y) = f(x_2, y)$ for any $x_1, x_2, y \in \mathbb{R}$. From the previous problem, we are sure that there exists some function $g : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(x, y) = g(y)$ for any $x, y \in \mathbb{R}$. Similar to the previous problem, again, we know that the derivative of such an f is $f'(a, b) = (0, g'(b))^T$.

If a function of two variables is independent of both variables, then let $z = f(0, 0)$. Then for any $(x, y) \in \mathbb{R}^2$, we have

$$f(x, y) = f(0, y) = f(0, 0) = z$$

and so $f(x, y) = z$ for any $x, y \in \mathbb{R}$, hence f is a constant function. Conversely, if f is a constant function then it is trivially independent of both variables. Thus, the constant functions are exactly those which are independent of both variables. ■

Problem 2.4. Let g be a continuous real-valued function on the unit circle $\{x \in \mathbb{R}^2 : |x| = 1\}$ such that $g(0, 1) = g(1, 0) = 0$ and $g(-x) = -g(x)$. Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} |x| \cdot g\left(\frac{x}{|x|}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

- If $x \in \mathbb{R}^2$ and $h : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $h(t) = f(tx)$, show that h is differentiable.
- Show that f is not differentiable at $(0, 0)$ unless $g = 0$. *Hint:* First show that $Df(0, 0)$ would have to be 0 by considering (h, k) with $k = 0$ and then with $h = 0$.

Proof. • Let $h(t) = f(tx)$ for some fixed $x \in \mathbb{R}^2$. Consider first the case where $x = 0$. Then

$$h(t) = f(0) = 0$$

everywhere, and so h is trivially differentiable since it is constant, and $h' = 0$.

Consider now $x \neq 0$. If $t > 0$ then

$$h(t) = |tx| \cdot g\left(\frac{tx}{|tx|}\right) = tx \cdot g\left(\frac{x}{|x|}\right) = tf(x)$$

Alternately, if $x \neq 0$ and $t < 0$, then

$$h(t) = |tx| \cdot g\left(\frac{tx}{|tx|}\right) = -tx \cdot g\left(-\frac{x}{|x|}\right) = tx \cdot g\left(\frac{x}{|x|}\right) = tf(x)$$

Lastly, if $x \neq 0$ and $t = 0$, then

$$h(t) = h(0) = f(0) = 0 = 0 \cdot f(x)$$

Thus, in the case $x \neq 0$, we have shown by exhaustion that $h(t) = f(tx) = tf(x)$ for all $t \in \mathbb{R}$. This is a simple linear function and is differentiable with derivative $h'(t) = f(x)$.

- Assume f is differentiable at $(0, 0)$ with $g \neq 0$. As the hint suggests, we consider taking the derivative along two different paths and show that the value of the derivative is not unique, a contradiction to the fact that differentiable functions have unique derivatives.

First, consider $k = 0$. Then the difference quotient is

$$\begin{aligned}
 0 &= \lim_{(h,0) \rightarrow (0,0)} \frac{|f(h,0) - f(0,0) - Df(0,0) \cdot (h,0)|}{|(h,0)|} \\
 &= \lim_{h \rightarrow 0} \frac{\left| |(h,0)| g\left(\frac{(h,0)}{|(h,0)|}\right) - Df(0,0) \cdot (h,0) \right|}{|h|} \\
 &= \lim_{h \rightarrow 0} \frac{\left| |h| g\left(\frac{h}{|h|}, 0\right) - Df(0,0) \cdot (h,0) \right|}{|h|} \\
 &= \lim_{h \rightarrow 0} \frac{|h| g(\pm 1, 0) - Df(0,0) \cdot (h,0)|}{|h|} \\
 &= \lim_{h \rightarrow 0} \frac{|Df(0,0) \cdot (h,0)|}{|h|}
 \end{aligned}$$

Recall that the linear map $Df(0,0) : \mathbb{R}^2 \rightarrow \mathbb{R}$ can be represented as a 1×2 array, say $Df(0,0) = [a, b]$. Then this last expression becomes

$$0 = \lim_{h \rightarrow 0} \frac{|ah|}{|h|} = \lim_{h \rightarrow 0} \frac{|a| |h|}{|h|} = |a|$$

Hence $|a| = 0$. Similarly, we can show that $|b| = 0$ by the same procedure, but letting $h = 0$. Hence, the derivative is $Df(0,0) = [0, 0]$, and so $Df(0,0)(h,k) = 0$ for any $(h,k) \in \mathbb{R}^2$. In the difference quotient, we have

$$\begin{aligned}
 0 &= \lim_{(h,k) \rightarrow (0,0)} \frac{|f(h,k) - f(0,0) - Df(0,0) \cdot (h,k)|}{|(h,k)|} \\
 &= \lim_{(h,k) \rightarrow (0,0)} \frac{|f(h,k)|}{|(h,k)|} \\
 &= \lim_{(h,k) \rightarrow (0,0)} \frac{\left| |(h,k)| \cdot g\left(\frac{(h,k)}{|(h,k)|}\right) \right|}{|(h,k)|} \\
 &= \lim_{(h,k) \rightarrow (0,0)} \left| g\left(\frac{(h,k)}{|(h,k)|}\right) \right| \\
 &\neq 0
 \end{aligned}$$

since we assumed $g \neq 0$. This is quite clearly a contradiction, and so our assumption that f is differentiable at $(0,0)$ with $g \neq 0$ is incorrect. If $g = 0$, then f is certainly differentiable at $(0,0)$. ■

Problem 2.5. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & (x,y) \neq 0 \\ 0 & (x,y) = 0 \end{cases}$$

Show that f is a function of the kind considered in Problem 2-4 so that f is not differentiable at $(0,0)$.

Problem 2.6. Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined by $f(x, y) = \sqrt{|xy|}$. Show that f is not differentiable at $(0, 0)$.

Proof. As in Problem 2-4, we will assume f is differentiable at $(0, 0)$ and consider different paths toward $(0, 0)$ until we hit a contradiction. Note that $f(0, 0) = 0$ and that we can, as usual, express the derivative as $Df(0, 0) = [a, b]$. By our assumption of differentiability, we have

$$0 = \lim_{(h,k) \rightarrow (0,0)} \frac{|f(h, k) - f(0, 0) - Df(0, 0)(h, k)|}{|(h, k)|}$$

First consider $k = 0$. Then this difference quotient becomes

$$\begin{aligned} 0 &= \lim_{h \rightarrow 0} \frac{|f(h, 0) - Df(0, 0)(h, 0)|}{|h|} \\ &= \lim_{h \rightarrow 0} \frac{|ah|}{|h|} \\ &= |a| \end{aligned}$$

and so $a = 0$. By the symmetry of the function f , we can repeat this to show that $b = 0$. That is, the derivative at 0 must be $Df(0, 0) = [0, 0]$. So far, so good; there is no contradiction.

Next, let us try $h = k = u$ and let $u \rightarrow 0$. The derivative difference quotient becomes

$$\begin{aligned} 0 &= \lim_{u \rightarrow 0} \frac{|f(u, u) - f(0, 0) - Df(0, 0)(u, u)|}{|(u, u)|} \\ &= \lim_{u \rightarrow 0} \frac{|\sqrt{|u^2|}|}{\sqrt{2}|u|^2} \\ &= \frac{1}{\sqrt{2}} \lim_{u \rightarrow 0} \frac{|u|}{|u|} \\ &= 2^{-1/2} \end{aligned}$$

which is certainly not zero. Hence, we have a contradiction, and so we must conclude that f is not differentiable at $(0, 0)$. ■

Problem 2.7. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function such that $|f(x)| \leq |x|^2$. Show that f is differentiable at 0.

Proof. Let f be such a function. Let's think about a special case to try to get some intuition. If we consider $n = 1$, we have $f : \mathbb{R} \rightarrow \mathbb{R}$ and $|f(x)| \leq |x|^2$. If we picture the graph of this function, the function will be bounded within the curves $y = x^2$ and $y = -x^2$. As we move closer to the origin, the function f gets "pinched" by these bounds; further, the derivatives of these bounding curves at the origin are both 0, and so it is reasonable to guess that $f'(0) = 0$. Generalizing to arbitrary n , we will guess that $Df(0) = 0 \in \mathbb{R}^{1 \times n}$.

Let us use the definition and the condition on f to prove this. The difference quotient is the quantity

$$DQ = \frac{|f(0+h) - f(0) - Df(0)(h)|}{|h|}$$

First, note that $|f(x)| \leq |x|^2$ means that $|f(0)| \leq 0$ so that $f(0) = 0$. Then, with our guess that $Df(0) = 0 \in \mathbb{R}^{1 \times n}$, we have

$$DQ = \frac{|f(h)|}{|h|} \leq \frac{|h|^2}{|h|} = |h|$$

Finally, taking the limit of the difference quotient, we obtain

$$\lim_{h \rightarrow 0} DQ \leq \lim_{h \rightarrow 0} |h| = 0$$

Thus, by definition of derivative, f is differentiable at $0 \in \mathbb{R}^n$, and the derivative is $Df(0) = 0 \in \mathbb{R}^{1 \times n}$. ■

Problem 2.8. Let $f : \mathbb{R} \rightarrow \mathbb{R}^2$. Prove that f is differentiable at $a \in \mathbb{R}$ if and only if f^1 and f^2 are, and that in this case

$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \end{pmatrix}$$

Proof. This problem hints very strongly toward the general structure of the derivative (Jacobian) of a multivariable, multivalued function by considering a simple case.

First, assume the function f is differentiable at $a \in \mathbb{R}$. Then in a sense we can decompose the difference quotient involving f into two similar difference quotients involving the component functions f^1 and f^2 :

$$\begin{aligned} \frac{|f(a+h) - f(a) - Df(a)(h)|^2}{|h|^2} &= \frac{|f^1(a+h) - f^1(a) - (Df(a))^1(h)|^2}{|h|^2} \\ &\quad + \frac{|f^2(a+h) - f^2(a) - (Df(a))^2(h)|^2}{|h|^2} \end{aligned}$$

where the superscripts on the right-hand side indicate component functions and elements of the derivative. Since the left hand side is equal to a sum of squares, the left-hand side is greater than either of the right-hand sides. Taking the square root of both sides and then taking the limit as $h \rightarrow 0$ gives the desired results; for instance,

$$0 = \lim_{h \rightarrow 0} \frac{|f(a+h) - f(a) - Df(a)(h)|^2}{|h|^2} \geq \lim_{h \rightarrow 0} \frac{|f^1(a+h) - f^1(a) - (Df(a))^1(h)|^2}{|h|^2}$$

and so the derivative of f^1 is the first component of the derivative of f , namely $(Df(a))^1$. Repeating this process for the other term on the right-hand side, we have the derivative of f^2 is $(Df(a))^2$. Thus, the component function are differentiable, and we have

$$f'(a) = Df(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \end{pmatrix}$$

Conversely, assume that the component functions are differentiable. Then we can use the triangle inequality on the difference quotient for f , similar to what we did above (the difference is that above we actually squared out the components exactly, whereas here we can use the triangle inequality on the linear combination $f = f^1 e_1 + f^2 e_2$ to get $|f| \leq |f^1| + |f^2|$).

$$\frac{|f(a+h) - f(a) - [(f^1)'(a), (f^2)'(a)](h)|}{|h|} \leq \frac{|f^1(a+h) - f^1(a) - (f^1(a))'(h)|}{|h|} + \frac{|f^2(a+h) - f^2(a) - (f^2(a))'(h)|}{|h|}$$

Taking the limit of both sides as $h \rightarrow 0$, the right-hand side goes to zero since we assumed f^1 and f^2 are differentiable at a . Thus, the limit of the left-hand side is zero, and this is the definition of the differentiability of f at a . Also, the derivative we used was exactly the one identified in the other direction of this proof. ■

Problem 2.9. Two functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are equal up to n th order at a if

$$\lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^n} = 0$$

- Show that f is differentiable at a if and only if there is a function g of the form $g(x) = a_0 + a_1(x-a)$ such that f and g are equal up to first order at a .
- If $f'(a), \dots, f^{(n)}(a)$ exist, show that f and the function g defined by

$$g(x) = \sum_{i=0}^n \frac{f^{(i)}(a)}{i!} (x-a)^i$$

are equal up to the n th order at a . *Hint:* the limit

$$\lim_{x \rightarrow a} \frac{f(x) - \sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!} (x-a)^i}{(x-a)^n}$$

may be evaluated by L'Hospital's rule.

Proof. This problem makes concrete the notion of the derivative as a low-order approximation to the function.

- Assume f is differentiable at a . Note the similarity of the definition of "equal up to n th order" to the difference quotient in the derivative. This is highly suggestive, and after playing around a little bit we can guess the function $g(x) = f(a) + Df(a)(x-a)$, i.e., $a_0 = f(a)$ and $a_1 = Df(a)$. Then by the definition of differentiability of f at a , we have

$$\begin{aligned} 0 &= \lim_{h \rightarrow 0} \frac{|f(a+h) - f(a) - Df(a)h|}{|h|} \\ &= \lim_{h \rightarrow 0} \frac{|f(a+h) - g(a+h)|}{|h|} \end{aligned}$$

Thus f and g are equal up to order 1.

Conversely, let there exist some $g(x) = a_0 + a_1(x - a)$ which is equal to f up to order 1. First, we want to show that the terminology "up to order n " is suggestive in that if two functions are equal up to order n , then they are equal up to order k for all $k = 0, 1, \dots, n$. Assume f and g are equal up to order n . Then

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^{n-1}} &= \lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^n} h^n \\ &= \left(\lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^n} \right) \left(\lim_{h \rightarrow 0} h^n \right) \\ &= 0 \cdot 0 \\ &= 0 \end{aligned}$$

Hence f and g are equal up to order $n-1$ if they are equal up to order n . We could continue for all orders under n , but for our problem we don't need this.

We can start the converse now. Assume f and g are equal up to order 1, and $g(x) = a_0 + a_1(x - a)$. Then by the previous paragraph they are equal up to order zero, i.e.,

$$\begin{aligned} 0 &= \lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^0} \\ &= \lim_{h \rightarrow 0} f(a+h) - \lim_{h \rightarrow 0} g(a+h) \\ &= f(a) - a_0 \end{aligned}$$

where we used the fact that differentiability of f at a guarantees continuity of f at a , and we used the particular form (and continuity) of g . Thus, $f(a) = a_0$, as we suspected from the previous problem.

Lastly, since f and g are equal up to order 1, we have

$$\begin{aligned} 0 &= \lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^1} \\ &= \lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - a_1 h}{h} \end{aligned}$$

This last equation says that f is differentiable at a , with derivative $f'(a) = a_1$.

- Assume that the first n derivatives of f at a exist, and let g be defined as in the hypothesis. To show that f and g are equal up to order n , we look at

$$\begin{aligned} L &= \lim_{h \rightarrow 0} \frac{f(a+h) - g(a+h)}{h^n} = \lim_{h \rightarrow 0} \frac{f(a+h) - \sum_{i=0}^n \frac{f^{(i)}(a)}{i!} h^i}{h^n} \\ &= \lim_{x \rightarrow a} \frac{f(x) - \sum_{i=0}^n \frac{f^{(i)}(a)}{i!} (x-a)^i}{(x-a)^n} \end{aligned}$$

where we made a change of variables $x = a + h$. Plugging a into this quotient, we obtain the indeterminate form $0/0$, hence L'Hopital's rule applies:

$$L = \lim_{x \rightarrow a} \frac{f'(x) - \sum_{i=1}^n \frac{f^{(i)}(a)}{i!} i (x-a)^{(i-1)}}{n (x-a)^{(n-1)}}$$

However, if we evaluate this quotient at $x = a$, we again obtain the same indeterminate form, $0/0$. We continue this process of applying L'Hopital's rule, i.e., we repeatedly differentiate the numerator and denominator:

$$\begin{aligned} L &= \lim_{x \rightarrow a} \frac{f^{(2)}(x) - \sum_{i=2}^n \frac{f^{(i)}(a)}{i!} i(i-1)(x-a)^{(i-2)}}{n(n-1)(x-a)^{(n-2)}} \\ &= \dots \\ &= \lim_{x \rightarrow a} \frac{f^{(n-1)}(x) - \sum_{i=n-1}^n \frac{f^{(i)}(a)}{i!} i(i-1)(i-2) \cdots (i-(n-2))(x-a)^{(i-(n-1))}}{n(n-1)(n-2) \cdots (n-(n-1))(x-a)^{(n-(n-2))}} \\ &= \lim_{x \rightarrow a} \frac{f^{(n)}(x) - \sum_{i=n}^n \frac{f^{(i)}(a)}{i!} i(i-1)(i-2) \cdots (i-(n-1))(x-a)^{(i-n)}}{n(n-1)(n-2) \cdots (n-(n-1))(x-a)^{(n-n)}} \end{aligned}$$

Notice that this last equation simplifies quite nicely:

$$\begin{aligned} L &= \lim_{x \rightarrow a} \frac{f^{(n)}(x) - \frac{f^{(n)}(a)}{n!} n! (x-a)^0}{n! (x-a)^0} \\ &= \frac{1}{n!} \lim_{x \rightarrow a} [f^{(n)}(x) - f^{(n)}(a)] \\ &= \frac{1}{n!} [f^{(n)}(a) - f^{(n)}(a)] \\ &= 0 \end{aligned}$$

where the limit is able to be taken because since $f^{(n)}$ is differentiable, it is continuous.¹ Hence, $L = 0$, which is what we needed in order to say that f and g are equal up to order n . ■

Problem 2.10. Use the theorems of this section to find f' for the following:

- $f(x, y, z) = x^y$
- $f(x, y, z) = (x^y, z)$
- $f(x, y) = \sin(x \sin y)$
- $f(x, y, z) = \sin(x \sin(y \sin z))$

¹This is a problem, since we only have guaranteed the first n derivatives to exist, so we cannot be sure that $f^{(n)}$ itself is differentiable. We will need to take a closer look at this.

- $f(x, y, z) = x^{y^z}$
- $f(x, y, z) = x^{y+z}$
- $f(x, y, z) = (x + y)^z$
- $f(x, y) = \sin(xy)$
- $f(x, y) = [\sin(xy)]^{\cos 3}$
- $f(x, y) = (\sin(xy), \sin(x \sin y), x^y)$

Problem 2.11. Find f' for the following (where $g : \mathbb{R} \rightarrow \mathbb{R}$ is continuous):

- $f(x, y) = \int_a^{x+y} g$
- $f(x, y) = \int_a^{x \cdot y} g$
- $f(x, y, z) = \int_{x^y}^{\sin(x \sin(y \sin z))} g$

Problem 2.12. A function $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ is bilinear if for $x, x_1, x_2 \in \mathbb{R}^n$, $y, y_1, y_2 \in \mathbb{R}^m$, and $a \in \mathbb{R}$ we have

$$\begin{aligned} f(ax, y) &= af(x, y) = f(x, ay) \\ f(x_1 + x_2, y) &= f(x_1, y) + f(x_2, y) \\ f(x, y_1 + y_2) &= f(x, y_1) + f(x, y_2) \end{aligned}$$

- Prove that if f is bilinear, then

$$\lim_{(h,k) \rightarrow 0} \frac{|f(h, k)|}{|(h, k)|} = 0$$

- Prove that $Df(a, b)(x, y) = f(a, y) + f(x, b)$
- Show that the formula for $Dp(a, b)$ in Theorem 2-3 is a special case of (b).

Proof. • Assume $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ is bilinear. ■

Problem 2.13. Define $IP : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by $IP(x, y) = \langle x, y \rangle$.

- Find $D(IP)(a, b)$ and $(IP)'(a, b)$.
- If $f, g : \mathbb{R} \rightarrow \mathbb{R}^n$ are differentiable and $h : \mathbb{R} \rightarrow \mathbb{R}$ is defined by $h(t) = \langle f(t), g(t) \rangle$, show that

$$h'(a) = \langle f'(a)^T, g(a) \rangle + \langle f(a), g'(a)^T \rangle$$

(Note that $f'(a)$ is an $n \times 1$ matrix; its transpose $f'(a)^T$ is a $1 \times n$ matrix, which we consider as a member of \mathbb{R}^n .)

- If $f : \mathbb{R} \rightarrow \mathbb{R}^n$ is differentiable and $|f(t)| = 1$ for all t , show that $\langle f'(t)^T, f(t) \rangle = 0$.
- Exhibit a differentiable function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that the function $|f|$ defined by $|f|(t) = |f(t)|$ is not differentiable.

Problem 2.14. Let $E_i, i = 1, \dots, k$ be Euclidean spaces of various dimensions. A function $f : E_1 \times \dots \times E_k \rightarrow \mathbb{R}^p$ is called multilinear if for each choice of $x_j \in E_j, j \neq i$ the function $g : E_i \rightarrow \mathbb{R}^p$ defined by $g(x) = f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_k)$ is a linear transformation.

- If f is multilinear and $i \neq j$, show that for $h = (h_1, \dots, h_k)$, with $h_l \in E_l$, we have

$$\lim_{h \rightarrow 0} \frac{|f(a_1, \dots, h_i, \dots, h_j, \dots, a_k)|}{|h|} = 0$$

Hint: if $g(x, y) = f(a_1, \dots, x, \dots, y, \dots, a_k)$, then g is bilinear.

- Prove that

$$Df(a_1, \dots, a_k)(x_1, \dots, x_k) = \sum_{i=1}^k f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_k)$$

Problem 2.15. Regard an $n \times n$ matrix as a point in the n -fold product $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ by considering each row as a member of \mathbb{R}^n .

- Prove that $\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable and

$$D(\det)(a_1, \dots, a_n)(x_1, \dots, x_n) = \sum_{i=1}^n \det \begin{pmatrix} a_1 \\ \vdots \\ x_i \\ \vdots \\ a_n \end{pmatrix}$$

- If $a_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ are differentiable and $f(t) = \det(a_{ij}(t))$, show that

$$f'(t) = \sum_{j=1}^n \det \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & & \vdots \\ a'_{j1}(t) & \cdots & a'_{jn}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}$$

- If $\det(a_{ij}(t)) \neq 0$ for all t and $b_1, \dots, b_n : \mathbb{R} \rightarrow \mathbb{R}$ are differentiable, let $s_1, \dots, s_n : \mathbb{R} \rightarrow \mathbb{R}$ be the functions such that $s_1(t), \dots, s_n(t)$ are the solutions of the equations

$$\sum_{j=1}^n a_{ji}(t) s_j(t) = b_i(t) \quad i = 1, \dots, n$$

Show that s_i is differentiable and find $s'_i(t)$.

Problem 2.16. Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is differentiable and has a differentiable $f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Show that $(f^{-1})'(a) = [f'(f^{-1}(a))]^{-1}$. *Hint:* $f \circ f^{-1}(x) = x$.

Problem 2.17. Find the partial derivatives of the following functions:

- $f(x, y, z) = x^y$
- $f(x, y, z) = z$
- $f(x, y) = \sin(x \sin y)$
- $f(x, y, z) = \sin(x \sin(y \sin z))$
- $f(x, y, z) = x^{y^z}$
- $f(x, y, z) = x^{y+z}$
- $f(x, y, z) = (x + y)^z$
- $f(x, y) = \sin(xy)$
- $f(x, y) = [\sin(xy)]^{\cos 3}$

Proof. Note: we do not need to use the limit definition of the partial derivative because for each derivative we can treat the multivariate functions as functions of only the variable we are differentiating with respect to.

- $D_1 f(x, y, z) = yx^{y-1}$
 $D_2 f(x, y, z) = \log(x) x^y$
 $D_3 f(x, y, z) = 0$
- $D_1 f(x, y, z) = D_2 f(x, y, z) = 0$
 $D_3 f(x, y, z) = 1$
- $D_1 f(x, y, z) = \sin(y) \cos(x \sin(y))$
 $D_2 f(x, y, z) = x \cos(y) \cos(x \sin(y))$
 $D_3 f(x, y, z) = 0$
- $D_1 f(x, y, z) = \sin(y \sin(z)) \cos(x \sin(y \sin(z)))$
 $D_2 f(x, y, z) = \sin(z) x \cos(y \sin(z)) \cos(x \sin(y \sin(z)))$
 $D_3 f(x, y, z) = y \cos(z) x \sin(y \sin(z)) \cos(x \sin(y \sin(z)))$
- $D_1 f(x, y, z) = y^z x^{y^z-1}$
 $D_2 f(x, y, z) = zy^{z-1} \log(x) x^{y^z}$
 $D_3 f(x, y, z) = \log(x) \log(y) y^z x^{y^z}$
- $D_1 f(x, y, z) = (y + z) x^{y+z-1}$
 $D_2 f(x, y, z) = \log(x) x^{y+z}$
 $D_3 f(x, y, z) = \log(x) x^{y+z}$

- $D_1 f(x, y, z) = z(x + y)^{z-1}$
 $D_2 f(x, y, z) = z(x + y)^{z-1}$
 $D_3 f(x, y, z) = \log(x + y)(x + y)^z$
- $D_1 f(x, y) = y \cos(xy)$
 $D_2 f(x, y) = x \cos(xy)$
- $D_1 f(x, y) = \cos(3) \sin(xy)^{\cos(3)-1} y \cos(xy)$
 $D_2 f(x, y) = \cos(3) \sin(xy)^{\cos(3)-1} x \cos(xy)$

■

Problem 2.18. Find the partial derivatives of the following functions (where $g : \mathbb{R} \rightarrow \mathbb{R}$ is continuous):

- $f(x, y) = \int_a^{x+y} g$
- $f(x, y) = \int_y^x g$
- $f(x, y) = \int_a^{xy} g$
- $f(x, y) = \int_a^{(\int_b^y g)} g$

Recall that the Second Fundamental Theorem of Calculus states that for continuous functions g , for f defined as

$$f(x) = \int_a^x g(t) dt$$

we have

$$f'(x) = g(x)$$

.

Proof. • Note

■

Problem 2.19. If $f(x, y) = x^{x^y} + (\log x)(\arctan(\arctan(\arctan(\sin(\cos xy) - \log(x + y))))))$ find $D_2 f(1, y)$. *Hint:* There is an easy way to do this.

Proof. At $x = 1$, the function to differentiate is $f(1, y) = 0$, and so the derivative is 0, since the function is constant.

■

Problem 2.20. Find the partial derivatives of f in terms of the derivatives of g and h if

- $f(x, y) = f(x)h(y)$
- $f(x, y) = g(x)^{h(y)}$
- $f(x, y) = g(x)$

- $f(x, y) = g(y)$
- $f(x, y) = g(x + y)$

Proof. • $D_1 f(x, y) = h(y) f'(x)$
 $D_2 f(x, y) = f(x) h'(y)$

- $D_1 f(x, y) = g'(x) h(y) g(x)^{h(y)-1}$
 $D_2 f(x, y) = \log(g(x)) h'(y) g(x)^{h(y)}$

- $D_1 f(x, y) = g'(x)$
 $D_2 f(x, y) = 0$

- $D_1 f(x, y) = 0$
 $D_2 f(x, y) = g'(y)$

- $D_1 f(x, y) = g'(x + y)$
 $D_2 f(x, y) = g'(x + y)$

■

Problem 2-21. Let $g_1, g_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuous. Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$f(x, y) = \int_0^x g_1(t, 0) dt + \int_0^y g_2(x, y) dt$$

- Show that $D_2 f(x, y) = g_2(x, y)$.
- How should f be defined so that $D_1 f(x, y) = g_1(x, y)$?
- Find a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ such that $D_1 f(x, y) = x$ and $D_2 f(x, y) = y$. Find one such that $D_1 f(x, y) = y$ and $D_2 f(x, y) = x$.

Proof. • Since the first integral in the sum is a function only of x , it vanishes under the derivative D_2 . The second integral is differentiated via the second fundamental theorem of calculus:

$$D_2 f(x, y) = 0 + g_2(x, y) = g_2(x, y).$$

- By the symmetry of this problem, it is easy to see that f should be defined in this case as

$$f(x, y) = \int_0^x g_1(t, y) dt + \int_0^y g_2(x, 0) dt$$

and a similar calculation as in the first part of this proof confirms this to work.

- Firstly, the function $f(x, y) = \frac{1}{2}(x^2 + y^2)$ has partial derivatives $D_1 f(x, y) = x$ and $D_2 f(x, y) = y$. Next, the function $f(x, y) = xy$ has partial derivatives $D_1 f(x, y) = y$ and $D_2 f(x, y) = x$.

■

Problem 2-22. If $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $D_2f = 0$, show that f is independent of the second variable. If $D_1f = D_2f = 0$, show that f is constant.

Proof. Assume $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $D_2f = 0$. The usual mean value theorem of univariate calculus applies to the function $g(y) = f(x, y)$, which has derivative $g'(y) = D_2f(x, y) = 0$; hence $g(y) = f(x, y)$ is constant (with respect to y) on its domain, and so it is independent of the second variable. ■

Problem 2-23. Let $A = \{(x, y) \in \mathbb{R}^2 \mid x < 0, \text{ or } x \geq 0 \text{ and } y \neq 0\}$.

- If $f : A \rightarrow \mathbb{R}$ and $D_1f = D_2f = 0$, show that f is constant. *Hint:* Note that any two points in A can be connected by a sequence of lines each parallel to one of the axes.
- Find a function $f : A \rightarrow \mathbb{R}$ such that $D_2f = 0$ but f is not independent of the second variable.

Problem 2-24. Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$f(x, y) = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq 0, \\ 0 & (x, y) = 0. \end{cases}$$