MA3101: Analysis III

Solutions to exercises from Michael Spivak's *Calculus on Manifolds*

Satvik Saha
19MS154

*Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.*

Contents

Chapter 1

Functions on Euclidean Space

1.1 Norm and Inner Product

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

Solution. Let $x_i = (0, \ldots, 0, x^i, 0, \ldots, 0)$, i.e. $x_i \in \mathbb{R}^n$ such that the ith component is x^i, while the remaining are all zero. Then, by repeated (but finitely many!) applications of the Triangle inequality, we have

$$|x| = \left|\sum_{i=1}^{n} x_i\right| \leq |x_1| + \left|\sum_{i=2}^{n} x_i\right| \leq \cdots \leq |x_1| + \cdots + |x_n|.$$

Since $|x_i| = |x^i|$, the desired result follows.

Problem 1-2. When does equality hold for the following, where $x, y \in \mathbb{R}^n$?

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

Solution. Clearly, equality holds when at least one of $x, y = 0$. Otherwise, we have equality whenever $x = \lambda y$ for some real $\lambda > 0$: this is because the proof uses the inequality

$$\langle x, y\rangle \leq |x| \cdot |y|.$$

Cauchy Schwarz gives $|\langle x, y\rangle| \leq |x| \cdot |y|$, with equality whenever $x = \lambda y$ for some $\lambda \in \mathbb{R}$. For $\langle x, y\rangle = |\langle x, y\rangle|$, we also demand $\langle x, y\rangle \geq 0$, i.e. $\lambda\langle y, y\rangle \geq 0$, hence $\lambda > 0$.

Problem 1-3. Prove that $|x - y| \leq |x| + |y|$. When does equality hold?
Solution. Use $|y| = |-y|$ in the Triangle inequality to obtain

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

Equality holds whenever at least one of $x, y = 0$, or when $x = \lambda y$ for some $\lambda < 0$.

Problem 1-4. Prove that
$$||x| - |y|| \leq |x - y|.$$

Solution. Note that the Triangle Inequality gives

$$|x| = |x - y + y| \leq |x - y| + |y|, \qquad |y| = |y - x + x| \leq |y - x| + |x|.$$

Rearranging and using $|x - y| = |y - x|$, we have

$$|x| - |y| \leq |x - y|, \qquad -(|x| - |y|) = |y| - |x| \leq |x - y|.$$

Together, this gives[1]

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

Problem 1-5. Prove and interpret geometrically the "triangle inequality"

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

Solution. Set $a = z - x$, $b = z - y$, $c = y - x$, and note that $a = b + c$ so we immediately have $|a| = |b + c| \leq |b| + |c|$.

Consider a triangle whose vertices in $\mathbb{R}^n$ are represented by the position vectors x, y, z. The quantities $|z - x|, |z - y|, |y - x|$ are the lengths of the three sides of the triangle, hence the inequality guarantees that the sum of the lengths of any two sides of a triangle is always at least the length of the remaining one.

Problem 1-6. Let f and g be integrable on $[a, b]$.

(a) Prove that

$$\left|\int_a^b f \cdot g\right| \leq \left(\int_a^b f^2\right)^{1/2} \cdot \left(\int_a^b g^2\right)^{1/2}.$$

(b) If equality holds, must $f = \lambda g$ for some $\lambda \in \mathbb{R}$? What if f and g are continuous?

(c) Show that

$$|\sum_{i=1}^n x^i y^i| \leq |x| \cdot |y|$$

is a special case of (a).

Solution.

(a) Examine the polynomial in λ,

$$0 \leq \int_a^b (f - \lambda g)^2 = \lambda^2 \int_a^b g^2 - 2\lambda \int_a^b f \cdot g + \int_a^b f^2 = A\lambda^2 - 2B\lambda + C.$$

Clearly, $A, C \geq 0$. If both $A, C = 0$, then $-2B\lambda \geq 0$ for all $\lambda \in \mathbb{R}$ forces $B = 0$, hence the inequality is trivial. If $A = 0$, we see that $-2B\lambda + C \geq 0$ for all λ forces $B = 0$ again. Finally, if $C = 0$, swap the roles of f and g to argue that $B = 0$ yet again.

Thus, we can assume that $A, C > 0$, hence the polynomial has non-zero roots if any. Rewrite the polynomial inequality as

$$0 \leq (\sqrt{A}\lambda - B/\sqrt{A})^2 - B^2/A + C.$$

Choosing $\lambda = B/A$ immediately gives $B^2 \leq AC$, which is the desired inequality.

When $B^2 = AC$, this polynomial actually has a root at $\lambda = B/A$. Otherwise, when $B^2 - AC < 0$, the polynomial has negative discriminant, and thus admits no real root.

[1] Note that $||x| - |y||$ is at least one of $|x| - |y|$ or $|y| - |x|$.

(b) Equality does *not* demand f and g to be linearly dependent. Define the functions $u_x\colon [a,b] \to \mathbb{R}$,

$$u_x(t) = \begin{cases} 1, & \text{if } t = x, \\ 0, & \text{if } t \neq x. \end{cases}$$

Then, setting $f = u_a$, $g = u_b$ gives equality; indeed, we can choose any two such functions with point discontinuities at distinct points within $[a,b]$.

If f and g are to be continuous, then equality does force the linear dependence of f and g. Either $A, C = 0$ forcing the corresponding $f, g = 0$; if not, we have equality precisely when

$$0 = \int_a^b (f - Bg/A)^2.$$

By setting $h = f - Bg/A$, we must have $h = 0$. We have used the elementary fact that for *continuous functions*[2],

$$\int_a^b f^2 = 0 \implies f = 0.$$

(c) For $x, y \in \mathbb{R}^n$, consider the functions $s_z\colon [0,n] \to \mathbb{R}$,

$$s_z(t) = z^i \text{ where } i = \lceil t \rceil,$$

and $s_z(0) = 0$. It is clear that

$$\int_0^n s_x s_y = \sum_{i=1}^n \int_{i-1}^i x^i y^i = \sum_{i=1}^n x^i y^i, \qquad \int_0^n s_z^2 = \sum_{i=1}^n (z^i)^2 = |z|^2.$$

This immediately gives

$$|\sum_{i=1}^n x^i y^i| \leq |x| \cdot |y|.$$

Problem 1-7. A linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is *norm preserving* if $|T(x)| = |x|$, and *inner product preserving* if $\langle Tx, Ty \rangle = \langle x, y \rangle$.

(a) Prove that T is norm preserving if and only if T is inner product preserving.

(b) Prove that such a linear transformation T is 1-1, and T^{-1} is of the same sort.

Solution.

(a) If T is inner product preserving, then $|x| = \sqrt{\langle x, x \rangle}$ immediately shows that T is norm preserving.

If T is norm preserving, then the identity

$$|x + y|^2 = |x|^2 + |y|^2 + 2\langle x, y \rangle$$

shows that T is inner product preserving. To see this, note that replacing $x \mapsto Tx$, $y \mapsto Ty$, $x + y \mapsto T(x + y)$ does not change the normed quantities, hence subtracting directly gives $\langle x, y \rangle = \langle Tx, Ty \rangle$.

(b) We claim that T is bijective and invertible, i.e. T has full rank. It is sufficient to show that its null space is trivial: the Rank-Nullity Theorem will guarantee the rest. Indeed, suppose that $Tx = 0$; taking norms gives $|x| = 0$ forcing $x = 0$.

Given any $x \in \mathbb{R}^n$, denote $x' = T^{-1}x$, hence the norm preserving nature of T gives $|x'| = |Tx'|$, i.e. $|T^{-1}x| = |x|$.

[2]If $f(\alpha) \neq 0$ at some point, then there is a non-empty neighbourhood of α on which $f(x)^2 > 0$, making the integral of f^2 strictly positive.

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

$$\angle(x, y) = \arccos \frac{\langle x, y \rangle}{|x| \cdot |y|}.$$

The linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is *angle preserving* if T is 1-1, and for $x, y \neq 0$ we have $\angle(Tx, Ty) = \angle(x, y)$.

(a) Prove that if T is norm preserving, then T is angle preserving.

(b) 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.

(c) What are all angle preserving $T\colon \mathbb{R}^n \to \mathbb{R}^n$?

Solution.

(a) This follows immediately from the fact that if T is norm preserving, it is also inner product preserving, hence all the terms in the angle formula are invariant under T.

(b) First suppose that all $Tx_i = \pm\lambda x_i$, then

$$|Tx|^2 = \sum_{i=1}^{n} |(Tx)^i|^2 = \sum_{i=1}^{n} | \pm \lambda x^i|^2 = \lambda^2 |x|^2,$$

hence $|Tx| = |\lambda x|$. Additionally,

$$\langle Tx, Ty \rangle = \sum_{i=1}^{n} (\pm\lambda x^i) \cdot (\pm\lambda y^i) = \sum_{i=1}^{n} \lambda^2 x^i y^i = \lambda^2 \langle x, y \rangle.$$

This gives

$$\angle(Tx, Ty) = \arccos \frac{\langle Tx, Ty \rangle}{|Tx| \cdot |Ty|} = \arccos \frac{\lambda^2 \langle x, y \rangle}{|\lambda x| \cdot |\lambda y|} = \arccos \frac{\langle x, y \rangle}{|x| \cdot |y|} = \angle(x, y).$$

Next, suppose that T is angle preserving, and that all $Tx_i = \lambda_i x_i$. Clearly, all $\lambda_i \neq 0$ since T is 1-1. Without loss of generality, let all $|x_i| = 1$ (normalize the given basis and observe that the same relations hold). Examine the angle between some $x_i + x_j$ and $x_i - x_j$: note that $\langle x_i + x_j, x_i - x_j \rangle = |x_1|^2 - |x_j|^2 = 0$ by expanding, so

$$\angle(x_i + x_j, x_i - x_j) = \arccos \frac{\langle x_i + x_j, x_i - x_j \rangle}{|x_i + x_j| \cdot |x_i - x_j|} = \frac{\pi}{2}.$$

$$\angle(T(x_i + x_j), T(x_i - x_j)) = \arccos \frac{\langle \lambda_i x_i + \lambda_j x_j, \lambda_i x_i - \lambda_j x_j \rangle}{|\lambda_i x_i + \lambda_j x_j| \cdot |\lambda_i x_i - \lambda_j x_j|}.$$

Thus, we demand

$$\langle \lambda_i x_i + \lambda_j x_j, \lambda_i x_i - \lambda_j x_j \rangle = 0,$$

This immediately gives $\lambda_i^2 - \lambda_j^2 = 0$, hence $\lambda_i = \pm\lambda_j$ as desired.

(c) In the standard basis $e_1, \dots, e_2$, the angle between any two distinct basis vectors is $\pi/2$, thus we demand $\langle Te_i, Te_j \rangle = 0$ for all $i \neq j$. Thus, the columns of the matrix representation of T in the standard basis are normal. We also want $|Te_i| = |Te_j|$: note that $\angle(e_i, e_i + e_j) = \pi/4$, thus $\cos(\pi/4) = 1/\sqrt{2}$ gives

$$\frac{1}{2} = \frac{|\langle Te_i, Te_i + Te_j \rangle|^2}{|Te_i|^2 \cdot |Te_i + Te_j|^2},$$

which expands to

$$2 \cdot |Te_i|^4 = |Te_i|^2 \cdot (|Te_i|^2 + |Te_j|^2),$$

hence $|Te_i| = |Te_j|$. Thus, the matrix representation of T is an orthogonal matrix multiplied by a non-zero scalar, i.e. T is a non-zero scalar multiple of a norm preserving map. This is both necessary and sufficient to ensure that T is angle preserving (note that without the scalar multiple, T is norm preserving hence angle preserving).

Problem 1-9. If $0 \leq \theta < \pi$, let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ 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 $\angle(x, Tx) = \theta$.

Solution. Note that $\det T = \cos^2\theta + \sin^2\theta = 1$, hence T is 1-1.

The map T is also norm preserving, hence angle preserving. To see this, it is sufficient to show that it preserves the norms of the basis vectors, and indeed $|Te_1|^2 = |Te_2|^2 = \cos^2\theta + \sin^2\theta = 1$. This means that T is inner product preserving, hence $\langle Te_1, Te_2\rangle = 0$. Now, for any $x = x^1e_1 + x^2e_2$, we have

$$|Tx|^2 = (x^1)^2 + (x^2)^2 + 2x^1x^2\langle Te_1, Te_2\rangle = (x^1)^2 + (x^2)^2 = |x|^2.$$

Finally, we can calculate

$$\begin{aligned} Tx &= (x^1\cos\theta + x^2\sin\theta)e_1 + (-x^1\sin\theta + x^2\cos\theta)e_2, \\ \langle x, Tx\rangle &= (x^1)^2\cos\theta + x^1x^2\sin\theta - x^2x^1\sin\theta + (x^2)^2\cos\theta = |x|^2\cos\theta, \\ \angle(x, Tx) &= \arccos\frac{\langle x, Tx\rangle}{|x|\cdot|Tx|} = \arccos\cos\theta = \theta. \end{aligned}$$

Problem 1-10. If $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, show that there is a number M such that $T(h) \leq M|h|$ for $h \in \mathbb{R}^n$.

Solution. Let a_{ij} be the entries of the matrix representation of T in the standard basis, and let $h = (h^1, \ldots, h^n) \in \mathbb{R}^n$. Then,

$$(Th)^i = \sum_{j=1}^n a_{ij}h^j \leq |h|\sum_{j=1}^n |a_{ij}|.$$

This is because each component $|h^j| \leq |h|$. Using $|x| \leq \sum_{i=1}^n |x^i|$, write

$$|Th| \leq \sum_{i=1}^n \left(|h|\sum_{j=1}^n |a_{ij}|\right) = |h|\sum_{i,j}|a_{ij}|.$$

Thus, setting $M = \sum_{i,j}|a_{ij}|$ completes the proof.

Problem 1-11. If $x, y \in \mathbb{R}^n$ and $z, w \in \mathbb{R}^m$, show that

$$\langle(x,z),(y,w)\rangle = \langle x,y\rangle + \langle z,w\rangle, \qquad |(x,z)| = \sqrt{|x|^2 + |z|^2}.$$

Solution. We have

$$\begin{aligned} (x,z) &= (x^1, \ldots, x^n, z^1, \ldots, z^m), \\ (y,w) &= (y^1, \ldots, y^n, w^1, \ldots, w^m), \\ \langle(x,z),(y,w)\rangle &= x^1y^1 + \cdots + x^ny^n + z^1w^1 + \cdots + z^mw^m = \langle x,y\rangle + \langle z,w\rangle, \\ |(x,z)|^2 &= (x^1)^2 + \cdots + (x^n)^2 + (z^1)^2 + \cdots + (z^m)^2 = |x|^2 + |z|^2. \end{aligned}$$

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 $\varphi_x \in (\mathbb{R}^n)^*$ by $\varphi_x(y) = \langle x, y\rangle$. Define $T \colon \mathbb{R}^n \to (\mathbb{R}^n)^*$ by $T(x) = \varphi_x$. Show that T is a 1-1 linear transformation and conclude that every $\varphi \in (\mathbb{R}^n)^*$ is φ_x for a unique $x \in \mathbb{R}^n$.

Solution. First, show that T is injective. If some $Tx = 0$, then $\varphi_x(y) = \langle x, y\rangle = 0$ for all $y \in \mathbb{R}^n$, hence $x = 0$ (we can successively set $y = e_i$ to deduce this).

Next, show that T is surjective. Let $\varphi \in (\mathbb{R}^n)^*$ be a linear functional on $\mathbb{R}^n$. Then, set $x^i = \varphi(e_i)$, and note that for any $y \in \mathbb{R}^n$, we must have

$$\varphi(y) = \varphi(y^1 e_1 + \cdots + y^n e_n) = y^1 x^1 + \cdots + y^n x^n = \langle x, y\rangle.$$

Thus, we have found $x \in \mathbb{R}^n$ such that $T(x) = \varphi$.

Now, if this $\varphi = \varphi_x = \varphi_z$ for some $x, z \in \mathbb{R}^n$, then

$$0 = \varphi_x(y) - \varphi_z(y) = \langle x - z, y\rangle$$

for all $y \in \mathbb{R}^n$, hence $x - z = 0$ or $x = z$.

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

$$|x + y|^2 = |x|^2 + |y|^2.$$

Solution. Calculate

$$|x + y|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle = |x|^2 + |y|^2.$$

1.2 Subsets of Euclidean Space

Problem 1-14. Prove that the union of any (even infinite) number of open sets is open. Prove that the intersection of two (and hence of finitely many) open sets is open. Give a counterexample for infinitely many open sets.

Solution. Let $\{U_\alpha\}$ be a collection of open sets, and let x be a point in their union. Thus, $x \in U_\alpha$ for some α, hence there is an open rectangle A around x contained within U_α. This immediately means shows that A is contained within the union of the open sets.

Let U_1, U_2 be open, and let x be a point in their intersection. Find open rectangles A_1 and A_2 around x, each contained within U_1 and U_2 respectively. Enumerate

$$A_1 = [a_1^1, b_1^1] \times \cdots \times [a_1^n, b_1^n],$$

$$A_2 = [a_2^1, b_2^1] \times \cdots \times [a_2^n, b_2^n].$$

By definition, we have $a_i^j \le x^j \le b_i^j$ for all coordinates. Define $a^j = \max(a_1^j, a_2^j)$ and $b^j = \min(b_1^j, b_2^j)$ for each j, and define the closed rectangle

$$A = [a^1, b^1] \times \cdots \times [a^n, b^n].$$

We now have $x \in A \subseteq A_1 \cap A_2 \subseteq U_1 \cap U_2$ as desired. It follows that the finite union of closed sets is closed.

Examine the closed intervals $[0, 1 - 1/n] \subseteq \mathbb{R}$. Their intersection is the set $[0, 1)$, which is not closed because the complement $(-\infty, 0) \cup [1, \infty)$ is not open (there is no open interval around 1 contained within this set).

Problem 1-15. Prove that $\{x \in \mathbb{R}^n : |x - a| < r\}$ is open.

Solution. Note that this set is precisely $\{x + a \in \mathbb{R}^n : |x| < r\}$ by definition, so without loss of generality suppose that $a = 0$ (or redefine $x - a \mapsto x$). Pick arbitrary x within this set, hence set

$$\epsilon = r - |x| > 0.$$

Define the open rectangle around x as follows.

$$A = (x^1 - \epsilon/\sqrt{n}, x^1 + \epsilon/\sqrt{n}) \times \cdots \times (x^n - \epsilon/\sqrt{n}, x^n + \epsilon/\sqrt{n}).$$

Then, for any $y \in A$, we have

$$x^i - \frac{\epsilon}{\sqrt{n}} < y^i < x^i - \frac{\epsilon}{\sqrt{n}}, \qquad (y^i - x^i)^2 \leq \frac{\epsilon}{n},$$

which when summed over i gives $|y - x|^2 < \epsilon^2$. Thus,

$$|y| = |y - x + x| \leq |y - x| + |x| < \epsilon + |x| < r.$$

Problem 1-16. Find the interior, exterior, and boundary of the sets

(a) $\{x \in \mathbb{R}^n : |x| \leq 1\}$.
(b) $\{x \in \mathbb{R}^n : |x| = 1\}$.
(c) $\{x \in \mathbb{R}^n : \text{each } x^i \text{ is rational}\}$.

Solution.

(a) We have seen by the previous exercise that the set $\{x \in \mathbb{R}^n : |x| < 1\}$ is open, and hence is part of the interior of the given set. Now, we show that there are no more points in the interior. Note that the remaining points in $\mathbb{R}^n$ can be partitioned into those satisfying $|x| = 1$, and $|x| > 1$. The latter do not belong to the given set, so we need only concern ourselves with the former. Pick $x \in \mathbb{R}^n$, $|x| = 1$. Then, let A be an arbitrary open rectangle around x,

$$A = (a^1, b^1) \times \cdots \times (a^n, b^n).$$

Define

$$r = \min(x^1 - a^i, \ldots, x^n - a^n, b^1 - x^1, \ldots, b^n - x^n) > 0,$$

and define $y = (1 + r/2)x$. It is clear that $y \in A$, but $|y| = 1 + r/2 > 1$. Note that each

$$y^i = x^i + \frac{1}{2} r x^i \leq x^i + \frac{1}{2} r |x^i| \leq x^i + r < x^i + (b^i - x^i) = b^i,$$

and

$$y^i = x^i + \frac{1}{2} r x^i \geq x^i - \frac{1}{2} r |x^i| \geq x^i - r > x^i - (x^i - a^i) = a^i.$$

We have chosen r as the radius of the largest sphere centred at x inscribed within the rectangle A.

We claim that the exterior is $\{x \in \mathbb{R}^n : |x| > 1\}$, i.e. this set is the interior of the complement of the given set, which happens to be $\{x \in \mathbb{R}^n : |x| > 1\}$. To show that this set is open, we perform a construction identical to that in Exercise 1-15, except with $\epsilon = |x| - 1 > 0$, whence $|y| = |x - x + y| \geq |x| - |y - x| > |x| - \epsilon = 1$. Again, we are not concerned with points $|x| \leq 1$ since those lie outside the set whose interior we are examining.

We claim that the boundary is $\{x \in \mathbb{R}^n : |x| = 1\}$. We have already shown that if $|x| = 1$ and A is an open rectangle around x, we can define $r > 0$ as before whence $(1 + r/2)x$ and $(1 - r/2)x$ are points from the exterior and the set itself respectively. Any point $|x| < 1$ lies in the open interior and $|x| > 1$ lies in the open exterior, hence there are open rectangles around such points which do not intersect the other set (the exterior and interior respectively).

(b) We claim that the interior of the given set is empty; to see this, we invoke the previous construction yet again, whereby given any point satisfying $|x| = 1$ and any open rectangle around it, there exists points within that open rectangle such that $|y_1| < 1$ and $|y_2| > 1$.

The previous constructions also suffice to show that $\{x \in \mathbb{R}^n : |x| \neq 1\}$ is the exterior of the given set; given a point satisfying $|x| < 1$, this is in the interior of $\{x \in \mathbb{R}^n : |x| < 1\}$ and hence there is an open rectangle around x contained within $\{x \in \mathbb{R}^n : |x| < 1\}$. There is an analogous case for when $|x| > 1$.

The boundary of the given set is $\{x \in \mathbb{R}^n : |x| = 1\}$. Again, we cannot choose any point from either the interior, nor the exterior which are open. Also, given any point satisfying $|x| = 1$ and an open rectangle A around x, we have $x \in A$ from the set itself and some point from the rectangle such that $|y| > 1$, i.e. not from the set.

(c) We claim that the interior and exterior are both empty. This is because given any open rectangle whatsoever, each of the component intervals (a^i, b^i) will contain at least one rational point and one irrational point. Thus, any open rectangle will contain a point from the set (all coordinates rational), and a point not from the set (at least one irrational coordinate). This also shows that the boundary of the given set is all of $\mathbb{R}^n$.

Problem 1-17. Construct a set $A \subset [0, 1] \times [0, 1]$ such that A contains at most one point on each horizontal and each vertical line but boundary $A = [0, 1] \times [0, 1]$.
Solution. As suggested, divide the rectangle into its four quadrants (splitting it evenly along the $x = 1/2$ and $y = 1/2$ lines), and pick one rational point (both coordinates are rational numbers) from each of them: this can be done without repeating any of the x or y coordinates since there are infinitely many of them to choose from, and we are only choosing four at this step. Now, repeat the same procedure on each of the four quadrants, and keep going. At any stage, we have 4^n sub-rectangles, and choose 4^{n+1} new points all without their coordinates colliding: again, this is possible since there are infinitely many choices, and we have only exhausted finitely many of them. The set A is the union of all the intermediate sets obtained at each stage.

We can show that the boundary of A is the entire rectangle $[0, 1] \times [0, 1]$. Pick an arbitrary point x from this rectangle, and consider an arbitrary open rectangle B containing it. Now, since every subdivided rectangle of side $1/2^n$ always contains a point from A by construction, and B must entirely contain such a rectangle for sufficiently large n, we see that B must contain a point from A. In addition, B must also contain some irrational point, hence a point outside A. This shows that x is in the boundary of A.

Problem 1-18. If $A \subset [0, 1]$ is the union of open intervals (a_i, b_i) such that each rational number in $(0, 1)$ is contained in some (a_i, b_i), show that boundary $A = [0, 1] - A$.

Solution. First, let x be a point from the boundary of A. We know that every open interval around x contains some point from A, some point outside A. This forces $x \in [0, 1]$ since if $x < 0$, the interval $(-2x, 0)$ has an empty intersection with A; there is an analogous case ruling out $x > 1$. Furthermore, if $x \in A$, then x belongs to some open interval (a_i, b_i), and this interval has an empty intersection with A^c. Thus, $x \in [0, 1] - A$.

Next, let x be a point from $[0,1]-A$; we claim that x is in the boundary of A. Pick a non-empty open interval $I=(x-\delta,x+\delta)$. Note that I must contain at least one rational number from $(0,1)$, and every rational number from $(0,1)$ is contained within A; this gives a non-empty intersection of I with A. On the other hand, $x\notin A$, i.e. I has a non-empty intersection with A^c.

Problem 1-19. If A is a closed set that contains every rational number $r\in[0,1]$, show that $[0,1]\subset A$.

Solution. Suppose not, i.e. there exists $x\in[0,1]$, $x\notin A$. Since A is closed, its complement is open; since $x\in A^c$, we can choose a non-empty interval $(x-\delta,x+\delta)$ which does not intersect A. However, every interval around x must contain some rational number from $[0,1]$ (indeed from $(0,1)$), and hence must intersect A.

Problem 1-20. Prove that every compact subset of $\mathbb{R}^n$ is closed and bounded.

Solution. Let $K\subset\mathbb{R}^n$ be compact, and pick $x\in K^c$. For each $y\in K$, note that $|x-y|>0$, so set $\delta_y=|x-y|/3$ and define the open balls $U_y=\{z\in\mathbb{R}^n:|z-x|<\delta_y\}$ and $V_y=\{z\in\mathbb{R}^n:|z-y|<\delta_y\}$. It is clear that $U_y\cap V_y=\emptyset$. Furthermore, the collection of all V_y form an open cover of K, whose compactness guarantees that we can extract a finite subcover $V_{y_1},\dots,V_{y_k}$. Set $U=U_{y_1}\cap\cdots\cap U_{y_k}$ which is non-empty and open. Now, $x\in U$, and

$$U\cap K\subseteq U\cap(V_{y_1}\cup\cdots\cup V_{y_k})=(U\cap V_{y_1})\cap\dots(U\cap V_{y_k})=\emptyset.$$

The last inference follows from the fact that each $U_{y_i}\cap V_{y_i}=\emptyset$. This shows that $x\in U\subset K^c$, hence K^c is open, i.e. K is closed.

Next, given $y\in K$, construct the open rectangles $A_y=(y^1-1,y^1+1)\times\cdots\times(y^n-1,y^n+1)$. Since all such A_y form an open cover of K, we can extract a finite subcover $A_{y_1},\dots,A_{y_k}$. For each index i, set $a^i=\min(y^i_1,\dots,y^i_k)-1$ and $b^i=\max(y^i_1,\dots,y^i_k)+1$, and set $A=(a^1,b^1)\times\cdots\times(a^n,b^n)$. This is a bounded open rectangle which clearly contains K, hence K is bounded.

Problem 1-21.

(a) If A is closed and $x\notin A$, prove that there is a number $d>0$ such that $|y-x|\geq d$ for all $y\in A$.

(b) If A is closed, B is compact, and $A\cap B=\emptyset$, prove that there is $d>0$ such that $|y-x|\geq d$ for all $y\in A$ and $x\in B$.

(c) Give a counterexample in $\mathbb{R}^2$ if A and B are closed but neither is compact.

Solution.

(a) Since A is closed, A^c is open. This means that for $x\notin A$, there exists an open rectangle around x which does not intersect A. Let $U=(a^1,b^1)\times\cdots\times(a^n,b^n)$, and set $d=\min(x^1-a^1,\dots,x^n-a^n,b^1-x^1,\dots,b^n-x^n)>0$. Then, if $|x-y|<d$, we have each $|x^i-y^i|<d$, hence each $y^i\in(a^i,b^i)$ so $y\in U$. In other words, if $y\in A$, then $y\notin U$ so $|x-y|\geq d$.

(b) We see that B is compact, hence closed and bounded. Given $y\in B\subset A^c$, we use the previous construction to obtain $d_y>0$ such that $|x-y|\geq d_y$ for all $x\in A$. Thus, the open set $B_y=\{x\in\mathbb{R}^n:|x-y|<d_y\}$ has an empty intersection with A. Since the collection of all such B_y forms an open cover of B which is compact, we can extract a finite subcover

$B_{y_1}, \ldots, B_{y_k}$. Each B_{y_i} has an empty intersection with A, which means that their union $B_{y_1} \cup \cdots \cup B_{y_k}$ also has an empty intersection with A. In fact, given any $x \in A$ and $y \in B$, we must have $y \in B_{y_i}$ for some i; but $B_{y_i} \cap A = \emptyset$ means $x \notin B_{y_i}$ so $|x - y| \geq d_{y_i}$. Setting $d = \min(d_1, \ldots, d_k) > 0$ completes the construction.

(c) Let $A = \{(x, y) \in \mathbb{R}^2 : y = 0\}$ and $B = \{(x, y) \in \mathbb{R}^2 : x > 0, xy = 1\}$. Indeed, the distance between $(x, 0) \in A$ and $(x, 1/x) \in B$ when $x > 0$ is just $1/x$, which can be made arbitrarily small.

Problem 1-22. If U is open and $C \subset U$ is compact, show that there is a compact set D such that $C \subset$ interior D and $D \subset U$.

Solution. Since U^c is closed and C is compact with $U^c \cap C = \emptyset$, we can pick $d > 0$ such that $|x - y| > d$ for all $x \in U^c$, $y \in C$. Define the set $D = \{x \in \mathbb{R}^n : |x - y| \leq d/2 \text{ for some } y \in C\}$. Clearly, $D \cap U^c = \emptyset$ (any point $x \in D \cap U^c$ satisfies $|x - y| \leq d/2$ for some $y \in C$, as well as $|x - y| > d$ for all $y \in C$, which is impossible), hence $D \subset U$. Also, given any point $y \in C \subset D$, the open set $U = \{x \in \mathbb{R}^n : |x - y| < d/2\}$ is contained within D by definition, hence $C \subset$ interior D. We further claim that D is compact; it is sufficient to show that it is closed and bounded, after which the Heine-Borel theorem guarantees the rest.

To see that D is closed, pick $x \in D^c$. Then, $|x - y| > d/2$ for all $y \in C$. Also, we have $x \in D^c \subset C^c$ and C^c is open, so there exists an open set V around x such that $V \cap C = \emptyset$.

1.3 Functions and Continuity

Problem 1-23. If $f \colon A \to \mathbb{R}^m$ and $a \in A$, show that $\lim_{x \to a} f(x) = b$ if and only if $\lim_{x \to a} f^i(x) = b^i$ for $i = 1, \ldots, m$.

Solution. First, suppose that $\lim_{x \to a} f(x) = b$. Then, given $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$ with $|x - a| < \delta$, we have $|f(x) - b| < \epsilon$. This immediately gives

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

for each index $i = 1, \ldots, m$, hence we have the limits $\lim_{x \to a} f^i(x) = b^i$.

Next, suppose that each $\lim_{x \to a} f^i(x) = b^i$. Thus, given an index i and $\epsilon > 0$, there exists $\delta_i > 0$ such that for all $x \in A$ with $|x - a| < \delta_i$, we have $|f^i(x) - b^i| < \epsilon/n$. Set $\delta = \min(\delta_1, \ldots, \delta_n) > 0$, whence for all $x \in A$, $|x - a| < \delta$, we have

$$|f(x) - b| \leq \sum_{i=1}^{n} |f^i(x) - b^i| < \sum_{i=1}^{n} \epsilon/n = \epsilon.$$

This gives $\lim_{x \to a} f(x) = b$.

Problem 1-24. Prove that $f \colon A \to \mathbb{R}^m$ is continuous at a if and only if each f^i is.

Solution. First, let f be continuous at a, whence $\lim_{x \to a} f(x) = f(a)$. This immediately gives $\lim_{x \to a} f^i(x) = (f(a))^i = f^i(a)$, hence each f^i is continuous at a.

Next, let each f^i be continuous at a, whence each $\lim_{x \to a} f^i(x) = f^i(a)$. This immediately gives $\lim_{x \to a} f(x) = (f^1(a), \ldots, f^n(a)) = f(a)$, hence f is continuous at a.

Problem 1-25. Prove that a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous.

Solution. Recall that we can find $M > 0$ such that $|Tx| \le M|x|$ for all $x \in \mathbb{R}^m$. Thus, given $a \in \mathbb{R}^n$ and $\epsilon > 0$, set $\delta = \epsilon/M > 0$, whence for all $x \in \mathbb{R}^n$ with $|x-a| < \delta$, we have

$$|Tx - Ta| = |T(x-a)| \le M|x-a| < M \cdot \epsilon/M = \epsilon.$$

Problem 1-26. Let $A = \{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } 0 < y < x^2\}$.

(a) Show that every straight line through $(0, 0)$ contains an interval around $(0, 0)$ which is in $\mathbb{R}^2 - A$.

(b) Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by $f(x) = 0$ if $x \notin A$ and $f(x) = 1$ if $x \in A$. For $h \in \mathbb{R}^2$, define $g_h\colon \mathbb{R} \to \mathbb{R}$ by $g_h(t) = f(th)$. Show that each g_h is continuous at 0, but f is not continuous at $(0, 0)$.

Solution.

(a) Note that the vertical line $x = 0$ is entirely contained in $\mathbb{R}^2 - A$, as is the horizontal line $y = 0$. Otherwise, consider a line parametrised by the points $(t, \lambda t)$ for some fixed $\lambda \neq 0$. Again, if $\lambda < 0$, we trivially see that the line is in $R^2 - A$ (given a point $(t, \lambda t)$, either $t \le 0$, or $t > 0$ so $\lambda t < 0$; either case excludes the point from A). Thus, assume that $\lambda > 0$. Then, for all $t \in (-\lambda, \lambda)$, we either have $t \le 0$, or $t > 0$ and $\lambda t > t^2$, hence this interval of our line lies in $\mathbb{R}^n - A$.

(b) Given $h = (x, y) \in R^2$, the points th for $t \in \mathbb{R}$ define a line (unless $h = 0$ in which case g_0 is the constant zero function which is clearly continuous). If $y = 0$ or $x = 0$, we have seen that the line is either vertical or horizontal, making g_h the constant zero function again. Otherwise, the point $th = (tx, ty) = (tx, tx \cdot y/x)$, which can be parametrised as the line $(t', \lambda t')$ for $\lambda = y/x$. Using the previous result, the points corresponding to $t' \in (-\lambda, \lambda)$ all lie outside A, so $g_h(t) = f(th) = 0$ for all $|t'| < \lambda$, i.e. $|t| < |y|/x^2$. This immediately shows that g_h is continuous at 0.

On the other hand, the limit $\lim_{x\to 0} f(x)$ does not exist. This is because along the sequence of points $(0, 1/n) \to (0, 0)$, we have $f(0, 1/n) = 0$ but along the sequence of points $(1/n, 1/2n^2) \to 0$, we have $f(1/n, 1/2n^2) = 1$. Thus, f is not continuous at $(0, 0)$.

Problem 1-27. Prove that $\{x \in \mathbb{R}^n : |x - a| < r\}$ is open by considering the function $f\colon \mathbb{R}^n \to \mathbb{R}$ with $f(x) = |x - a|$.

Solution. We claim that f is continuous, from which it follows that the pre-image of the open interval $(-r, r)$ which is precisely the given set is open.

Given $b \in \mathbb{R}^n$ and $\epsilon > 0$, set $\delta = \epsilon > 0$. Then for all $x \in \mathbb{R}^n$ with $|x - b| < \delta$, we have

$$|f(x) - f(b)| = ||x-a| - |b-a|| \le |(x-a) - (b-a)| = |x-b| < \epsilon.$$

This gives the continuity of f at every $b \in \mathbb{R}^n$.

Problem 1-28. If $A \subset \mathbb{R}^n$ is not closed, show that there is a continuous function $f\colon A \to \mathbb{R}$ which is unbounded.

Solution. Since A is not closed, i.e. $\mathbb{R}^n - A$ is not open, there exists some $x_0 \in \mathbb{R}^n - A$ such that every open set around x_0 intersects A. Thus, for any $x \in A$, we have $|x - x_0| > 0$; however, given any $\delta > 0$ open ball B around x_0, there exists $x \in A \cap B$ hence $|x - x_0| < \delta$. Thus, define

$f\colon A \to \mathbb{R}$, $x \mapsto 1/|x - x_0|$. This is well defined and unbounded on A since given arbitrary $M > 0$, we can find $x \in A$ such that $|x - x_0| < 1/M$, hence $f(x) > M$.

To see that f is continuous on A, pick $a \in A$ and examine $\lim_{x\to a} f(x)$. Note that we have already shown that $\lim_{x\to a} |x - x_0| = |a - x_0| > 0$. This immediately gives $\lim_{x\to a} f(x) = 1/|x - a|$, via results from real analysis (the reciprocal of a non-zero real continuous function is continuous).

Problem 1-29. If A is compact, prove that every continuous function $f\colon A \to \mathbb{R}$ takes on a maximum and a minimum value.

Solution. Since the continuous image of a compact set is compact, we see that $f(A) \subset \mathbb{R}$ is a compact set, which means that it is closed and bounded. In other words, its supremum and infimum exist: furthermore, these are contained within $f(A)$. Let $\alpha = \sup f(A)$, and suppose that $\alpha \notin f(A)$. Since $\mathbb{R} - f(A)$ is open, there is an open interval $(\alpha - \delta, \alpha + \delta)$ which does not intersect $f(A)$. Thus, the number $\alpha - \delta$ is also an upper bound for $f(A)$ (if $x \in f(A)$ such that $x > \alpha - \delta$, then $x \notin (\alpha - \delta, \alpha + \delta)$ forces $x > \alpha + \delta > \alpha$, a contradiction). This contradicts the fact that α is the lowest upper bound of $f(A)$. Thus, $\alpha \in f(A)$, so there is some $x \in A$ such that $f(x) = \alpha$. We can also show that $\beta = inf f(A) \in f(A)$ analogously.

Problem 1-30. Let $f\colon [a, b] \to \mathbb{R}$ be an increasing function. If $x_1, \dots, x_n \in [a, b]$ are distinct, show that

$$\sum_{i=1}^{n} o(f, x_i) < f(b) - f(a).$$

Solution. Without loss of generality, let $a \le x_1 < x_2 < \cdots < x_n \le b$. Note that f maps $[a, b]$ to a compact interval, hence f is bounded. Now, for some x_i, examine the set

$$A_{i,\delta} = \{f(x) : x \in [a, b], |x - x_i| < \delta\}.$$

Since f is increasing, this set is bounded above by $f(x_i + \delta)$ and bounded below by $f(x_i - \delta)$ (for this to make sense, extend f by setting $f(x) = f(b)$ for all $x > b$, and $f(x) = f(a)$ for all $x < a$). Furthermore, as δ decreases to 0, $f(x_i + \delta)$ decreases and $f(x_i - \delta)$ increases, thus the limit of the supremum remains at most $f(x_i + \delta)$ and the limit of the infimum remains at least $f(x_i - \delta)$. Thus, the oscillation $o(f, x_i)$ is at most $f(x + \delta) - f(x - \delta)$. Now, set $r = \min(x_2 - x_1, x_3 - x_2, \dots, x_n - x_{n-1})$, $\delta = r/2$, and $y_i = x_i + \delta$; the latter immediately gives $y_i \le x_{i+1} - \delta$. This gives the inequalities

$$\begin{aligned}
o(f, x_1) &\le f(x_1 + \delta) - f(x_1 - \delta) \le f(y_1) - f(a), \\
o(f, x_2) &\le f(x_2 + \delta) - f(x_2 - \delta) \le f(y_2) - f(y_1), \\
&\vdots \qquad\quad \vdots \\
o(f, x_{n-1}) &\le f(x_{n-1} + \delta) - f(x_{n-1} - \delta) \le f(y_{n-1}) - f(y_{n-2}), \\
o(f, x_n) &\le f(x_n + \delta) - f(x_n - \delta) \le f(b) - f(y_{n-1}).
\end{aligned}$$

Adding them up gives the desired inequality,

$$\sum_{i=1}^{n} o(f, x_i) < f(b) - f(a).$$

Chapter 2

Differentiation

2.1 Basic definitions

Problem 2-1. Prove that if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then it is continuous at a.

Solution. We have
$$\lim_{x\to a} \frac{|f(x) - f(a) - Df(a)(x-a)|}{|x-a|} = 0.$$
Since $Df(a)$ is a linear transformation, there exists $M > 0$ such that $|Df(a)(h)| < M|h|$ for all $h \in \mathbb{R}^n$. Thus, given $\epsilon > 0$, there exists $\delta > 0$ such that for all $0 < |x-a| < \delta_0$, we have
$$\frac{|f(x) - f(a) - Df(a)(x-a)|}{|x-a|} < \epsilon, \qquad ||f(x) - f(a)| - |Df(a)(x-a)|| \le \epsilon|x-a|.$$
Thus,
$$|f(x) - f(a)| \le |Df(a)(x-a)| + \epsilon|x-a| < M|x-a| + \epsilon|x-a|.$$
Choose $\delta = \min(\delta_0, \epsilon/(M+\epsilon))$, whence for all $|x-a| < \delta$, we have
$$|f(x) - f(a)| < (M+\epsilon)\cdot\frac{\epsilon}{M+\epsilon} = \epsilon.$$

Problem 2-2. A function $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $g\colon \mathbb{R} \to \mathbb{R}$ such that $f(x,y) = g(x)$. What is $f'(a,b)$ in terms of g'?

Solution. First, suppose that f is independent of the second variable. Define $g\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x,0)$. Then, we can see that $f(x,y) = f(x,0) = g(x)$ for all $x, y \in \mathbb{R}$.

Next, suppose that $g\colon \mathbb{R} \to \mathbb{R}$ such that $f(x,y) = g(x)$. Then, given any $x \in \mathbb{R}$ and $y_1, y_2 \in \mathbb{R}$, we have $f(x,y_1) = g(x) = f(x,y_2)$, hence f is independent of the second variable.

We claim that $f'(a,b) = g'(a)$. To see this, note that if f is differentiable, we have
$$\lim_{(x,y)\to(a,b)} \frac{|f(x,y) - f(a,b) - f'(a,b)(x-a, y-b)|}{|(x-a, y-b)|} = 0.$$
Since the limit $\lim_{(x,y)\to(a,b)} |(x-a,y-b)| = 0$ also exists, we can multiply and write
$$\lim_{(x,y)\to(a,b)} |f(x,y) - f(a,b) - f'(a,b)(x-a,y-b)| = 0.$$
Using $f(x,y) = g(x)$, write
$$\lim_{(x,y)\to(a,b)} |g(x) - g(a) - f'(a,b)(x-a,y-b)| = 0.$$

Now, if g is also differentiable, we repeat the same process as above to write

$$\lim_{(x,y)\to(a,b)} |g(x) - g(a) - g'(a)(x-a)| = 0,$$

nothing that the additional $y \to b$ does not affect the result. Thus, the linear transformation $(x, y) \mapsto g'(a)x$ satisfies the role of $f'(a, b)$; since the derivative of a function is unique, this means that this is the only choice of $f'(a, b)$.

Problem 2-3. Define when a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ 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?

Solution. A function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is independent of the first variable if for each $y \in \mathbb{R}$, we have $f(x_1, y) = f(x_2, y)$ for all $x_1, x_2 \in \mathbb{R}$.

The functions which are independent in both variables are precisely the constant functions. Given such a function, we see that $f(x, y) = f(x, 0) = f(0, 0)$ for all $x, y \in \mathbb{R}$.

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 $f\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x) = \begin{cases} |x| \cdot g(x/|x|), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

(a) If $x \in \mathbb{R}^2$ and $h\colon \mathbb{R} \to \mathbb{R}$ is defined by $h(t) = f(tx)$, show that h is differentiable.

(b) Show that f is not differentiable at $(0, 0)$ unless $g = 0$.

Solution.

(a) This is trivially true when $x = 0$ since $h = 0$. Otherwise, set $g(x/|x|) = \alpha$. This gives $h(0) = f(0) = 0$; for $t > 0$ we have $h(t) = f(tx) = |tx| \cdot g(tx/|tx|) = t|x| \cdot g(x/|x|) = \alpha t|x|$, and for $t < 0$ we have $h(t) = -t|x| \cdot g(-x/|x|) = t|x| \cdot g(x/|x|) = \alpha t|x|$. Thus, $h(t) = (\alpha|x|)t$, which is linear and hence clearly differentiable, with $h'(t) = \alpha|x|$.

We can verify that indeed,

$$\lim_{h\to 0} \frac{h(t+h) - h(t) - \alpha|x|h}{|h|} = \lim_{h\to 0} \frac{\alpha|x|(t+h) - \alpha|x|t - \alpha|x|h}{|h|} = 0.$$

(b) It is clear that $f = 0$ when $g = 0$. If not, then $g(x_0) = \alpha_0 \neq 0$ for some $x_0 \in \mathbb{R}^2$, $|x_0| = 1$. Since $g(1, 0) = g(0, 1) = 0$ and $g(-1, 0) = g(0, -1) = 0$, we can say that neither coordinate of $x_0 = (x_1, x_2)$ is zero. Suppose that f is indeed differentiable at $(0, 0)$; then

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

Note that $f(0, 0) = 0$. We can choose specific sequences along which to take this limit: first, hold $k = 0$, $h \to 0^+$, and note that $f(h, 0) = |h| \cdot g(1, 0) = 0$. This gives

$$|Df(0, 0)(1, 0)| = \lim_{h\to 0} \frac{|Df(0, 0)(h, 0)|}{|h|} = 0.$$

Similarly, holding $h = 0$, $k \to 0^+$ yields

$$|Df(0, 0)(0, 1)| = \lim_{h\to 0} \frac{|Df(0, 0)(0, k)|}{|k|} = 0.$$

Using the linearity of $Df(0,0)$, we see that $Df(0,0) = 0$ identically. Now note that as $t \to 0$, $tx_0 \to 0$ so we should also have

$$\lim_{t\to 0} \frac{|f(tx_0) - Df(0,0)(tx_0)|}{|tx_0|} = 0.$$

We have shown that $Df(0,0)(tx_0) = 0$, however $f(tx_0) = |tx_0| \cdot g(tx_0/|tx_0|) = \alpha t \neq 0$ when $t \neq 0$. Thus, the limit

$$\lim_{t\to 0} \frac{|f(tx_0)|}{|tx_0|} = \lim_{t\to 0} \frac{|\alpha t|}{|t|} = |\alpha| \neq 0.$$

This is a contradiction, this f cannot be differentiable at $(0,0)$.

Problem 2-5. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

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

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

Solution. We construct the continuous real-valued function g on the unit circle $g(x,y) = x|y|$. This is indeed continuous since the desired limits are trivial. Then, $g(1,0) = g(0,1) = 0$, and $g(-x,-y) = -g(x,y)$. In addition, f is defined exactly as described in problem 2-4. This directly shows that f is not differentiable at $(0,0)$.

Problem 2-6. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = \sqrt{|xy|}$. Show that f is not differentiable at $(0,0)$.

Solution. Suppose that to the contrary, f is differentiable at $(0,0)$. This means that there exists $Df(0,0)$ such that

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

Now, $f(0,0) = 0$. By holding $k = 0$, $h \to 0$, we get $f(h,0) = 0$ so

$$|Df(0,0)(1,0)| = \lim_{h\to 0} \frac{|Df(0,0)(h,0)|}{|h|} = 0.$$

Similarly, holding $h = 0$, $k \to 0$ gives

$$|Df(0,0)(0,1)| = \lim_{h\to 0} \frac{|Df(0,0)(0,k)|}{|k|} = 0.$$

Thus, $Df(0,0) = 0$ identically. On the other hand, note that $(t,t) \to 0$ as $t \to 0$, hence we must have

$$\lim_{t\to 0} \frac{|f(t,t) - Df(0,0)(t,t)|}{|(t,t)|} = 0.$$

However, this actually evaluates to

$$\lim_{t\to 0} \frac{\sqrt{|t^2|}}{\sqrt{2t^2}} = \frac{1}{\sqrt{2}},$$

which is a contradiction.

Problem 2-7. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a function such that $|f(x)| \leq |x|^2$. Show that f is differentiable at 0.

Solution. We claim that the zero function gives the derivative of f at 0. To verify this, first note that $|f(0)| \leq 0$ forces $f(0) = 0$; we now calculate the limit

$$0 \leq \lim_{h\to 0} \frac{|f(h) - f(0)|}{|h|} = \lim_{h\to 0} \frac{|f(x)|}{|x|} \leq \lim_{h\to 0} \frac{|h|^2}{|h|} = \lim_{h\to 0} |h| = 0.$$

The squeeze theorem guarantees that this limit exists, and is equal to zero.

Problem 2-8. Let $f\colon \mathbb{R} \to \mathbb{R}^2$. 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

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

Solution. First suppose that f is differentiable at a. Then,

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

Write $Df(a) = (d^1, d^2)$, and recall that $|x^i| \leq |x|$ for each component, thus the squeeze theorem applied to

$$0 \leq \lim_{h\to 0} \frac{|f^i(a+h) - f^i(a) - d^i(h)|}{|h|} \leq \lim_{h\to 0} \frac{|(f(a+h) - f(a) - d(h)|}{|h|} = 0$$

gives

$$\lim_{h\to 0} \frac{|f^i(a+h) - f^i(a) - d^i(h)|}{|h|} = 0.$$

This immediately shows that f^1 and f^2 are differentiable, with $(f^i)'(a)$ equal to the ith component of $Df(a)$.

Now suppose that each f^i is differentiable, hence we have the limits

$$\lim_{h\to 0} \frac{|f^i(a+h) - f^i(a) - d^i(h)|}{|h|} = 0.$$

Recall that $|x| \leq \sum_i |x^i|$, thus the squeeze theorem applied to

$$0 \leq \lim_{h\to 0} \frac{|f(a+h) - f(a) - d(h)|}{|h|} \leq \sum_{i=1}^{2} \lim_{h\to 0} \frac{|f^i(a+h) - f^i(a) - d^i(h)|}{|h|} = 0$$

gives

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - d(h)|}{|h|} = 0.$$

Here, we have denoted $d = (d^1, d^2) = Df(a)$.

In short, we have used the squeeze theorem on the chain of inequalities

$$0 \leq |x^i| \leq |x| \leq \sum_i |x^i|.$$

Problem 2-9. Two functions $f, g\colon \mathbb{R} \to \mathbb{R}$ are *equal up to nth order* at a if

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

(a) 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 order first order at a.

(b) If $f'(a), \dots, f^{(n)}(a)$ exist, show that f and the function g defined by

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

are equal up to nth order at a.

Solution.

(a) First suppose that f is differentiable at a. Set $g(x) = f(a) + f'(a)(x-a)$, and calculate

$$\lim_{h\to 0} \frac{f(a+h) - [f(a) + f'(a)h]}{h} = \lim_{h\to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0,$$

with the last inference following directly from the definition of the derivative $f'(a)$. Thus, f and g are equal up to the first order at a.

Next, suppose that there exists $g(x) = a_0 + a_1(x-a)$ which is equal to f up to the first order at a. This means that the limit

$$\lim_{h\to 0} \frac{f(a+h) - [a_0 + a_1 h]}{h} = 0, \qquad \lim_{h\to 0} \frac{f(a+h) - a_0}{h} = a_1.$$

Since $\lim_{h\to 0} h = 0$, we can multiply this with the first limit and get

$$\lim_{h\to 0} f(a+h) - a_0 - a_1 h = 0, \qquad \lim_{h\to 0} f(a+h) = a_0.$$

This is not enough to ensure that f is differentiable at a! We also require $f(a) = g(a)$ at minimum. As a counterexample, define $f, g\colon \mathbb{R} \to \mathbb{R}$, $f(0) = 1$, $f(x) = 0$ for all $x \neq 0$, $g(x) = 0$. Then, we indeed have equality up to the first order since

$$\lim_{h\to 0} \frac{f(0+h) - g(0+h)}{h} = 0,$$

but f is not differentiable at 0.

Assuming that $f(a) = g(a)$, we immediately have $f(a) = a_0$, hence f is continuous at a, and also differentiable at a by the existence of the second and first limits, with $f'(a) = a_1$.

(b) First note that we have $f(a) = g(a)$, $f'(a) = g'(a)$, $\dots$, $f^{(n)}(a) = g^{(n)}(a)$. This is simply because

$$\frac{d^i}{dx^i}(x-a)^i = i!.$$

When the power of $x-a$ is lower than i, the term vanishes, and when it is higher than i, a factor of $x-a$ remains hence this term vanishes as well at $x = a$.

As suggested, examine the limit

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

Note that the numerator and denominator are differentiable $n-1$ times at a, all of them continuous (recall that $f^{(n)}(a)$ exists). Also, after differentiating the numerator up to $n-1$ times, the limit of this numerator as $x \to a$ collapses to 0 – we have $f^{(i)}(x) \to f^{(i)}(a)$ by the continuity of the derivatives, and the constant term in the polynomial on the right is precisely $f^{(i)}(a)$, with the remaining terms vanishing because of the still present factors of $x-a$. Thus, we can apply L'Hôpital's rule by differentiating $n-1$ times, after which we have reduced this limit to

$$\lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{n!(x-a)} = \frac{f^{(n)}(a)}{n!}.$$

What we have shown is that

$$\lim_{h\to 0} \frac{f(a+h) - g(a+h) + f^{(n)}(a)h^n/n!}{h^n} = \frac{f^{(n)}(a)}{n!}.$$

Separating the final term from the left hand side and cancelling, we have

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

as desired.

2.2 Basic theorems

Problem 2-10. Find f' for the following.

(a) $f(x,y,z) = x^y$.

(b) $f(x,y,z) = (x^y, z)$.

(c) $f(x,y) = \sin(x\sin(y))$.

(d) $f(x,y,z) = \sin(x\sin(y\sin(z)))$.

(e) $f(x,y,z) = x^{y^z}$.

(f) $f(x,y,z) = x^{y+z}$.

(g) $f(x,y,z) = (x+y)^z$.

(h) $f(x,y) = \sin(xy)$.

(i) $f(x,y) = [\sin(xy)]^{\cos 3}$.

(j) $f(x,y) = (\sin(xy), \sin(x\sin(y)), x^y)$.

Solution.

(a) We rewrite this as $f(x,y,z) = \exp(y\ln x)$, so $f = \exp\circ(\pi^2 \cdot (\ln\circ\pi^1))$. Thus,

$$\begin{aligned} f'(a,b,c) &= \exp'(b\ln a) \cdot [\ln a \cdot (\pi^2)'(a,b,c) + b \cdot (\ln\circ\pi^1)'(a,b,c)] \\ &= \exp(b\ln a) \cdot [\ln a \cdot (0,1,0) + b \cdot \ln'(a) \cdot (\pi^1)'(a,b,c)] \\ &= \exp(b\ln a) \cdot [(0,\ln a,0) + b \cdot \frac{1}{a} \cdot (1,0,0)] \\ &= (ba^{b-1}, a^b \ln a, 0). \end{aligned}$$

(b) We write $f = (\exp\circ(\pi^2 \cdot \ln\circ\pi^2), \pi^3)$. Calculating derivatives component-wise, using $(\pi^3)'(a,b,c) = (0,0,1)$ together with the previous result gives

$$f'(a,b,c) = \begin{pmatrix} ba^{b-1} & a^b\ln a & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(c) We write $f = \sin \circ (\pi^1 \cdot \sin \circ \pi^2)$. Thus,

$$\begin{aligned} f'(a,b) &= \sin'(a \sin b) \cdot [\sin b \cdot (\pi^1)'(a,b) + a \cdot \sin'(b) \cdot (\pi^2)'(a,b)] \\ &= \cos(a \sin b) \cdot [\sin b \cdot (1,0) + a \cos b \cdot (0,1)] \\ &= (\cos(a \sin b) \sin(b), a \cos(a \sin b) \cos(b)). \end{aligned}$$

(d) We write $f = \sin \circ (\pi^1 \cdot \sin \circ (\pi^2 \cdot \sin \circ \pi^3))$. Thus,

$$\begin{aligned} f'(a,b,c) &= \sin'(a \sin(b \sin c)) \cdot [\sin(b \sin c) \cdot (\pi^1)'(a,b,c) + \\ &\qquad a \cdot \sin'(b \sin c) \cdot [\sin c \cdot (\pi^2)'(a,b,c) + b \cdot \sin'(c) \cdot (\pi^3)'(a,b,c)]] \\ &= \cos(a \sin(b \sin c)) \cdot [\sin(b \sin c) \cdot (1,0,0) + \\ &\qquad a \cos(b \sin c) \cdot [\sin c \cdot (0,1,0) + b \cos c \cdot (0,0,1)]] \\ &= (\cos(a \sin(b \sin c)) \sin(b \sin c), \\ &\qquad a \cos(a \sin(b \sin c)) \cos(b \sin c) \sin c, \\ &\qquad ab \cos(a \sin(b \sin c)) \cos(b \sin c) \cos c). \end{aligned}$$

(e) We write $f(x,y,x) = \exp(\exp(z \ln y) \ln x)$. Since $f(x,y,z) = x^{(y^z)}$, we can reuse our work from (a): define $g(x,y) = x^y$, so $f = g \circ (\pi^1, g \circ (\pi^2, \pi^3))$. Thus,

$$\begin{aligned} f'(a,b,c) &= g'(a,b^c) \cdot ((\pi^1)'(a,b,c), g'(b,c) \cdot ((\pi^2)'(a,b,c), (\pi^3)'(a,b,c))) \\ &= (b^c a^{b^c-1}, a^{b^c} \ln a) \cdot ((1,0,0), (cb^{c-1}, b^c \ln b) \cdot ((0,1,0),(0,0,1))) \\ &= (b^c a^{b^c-1}, a^{b^c} \ln a) \cdot ((1,0,0), (0, cb^{c-1}, b^c \ln b)) \\ &= (b^c a^{b^c-1}, a^{b^c} \ln a \cdot cb^{c-1}, a^{b^c} \ln a \cdot b^c \ln b). \end{aligned}$$

(f) We write $f(x,y,z) = x^y \cdot x^z$, hence $f = g \circ (\pi^1, \pi^2) \cdot g \circ (\pi^1, \pi^3)$. Thus,

$$\begin{aligned} f'(a,b,c) &= g'(a,b) \cdot ((1,0,0),(0,1,0)) \cdot g(a,c) + g'(a,c) \cdot ((1,0,0),(0,0,1)) \cdot g(a,b) \\ &= (ba^{b-1}, a^b \ln a, 0) \cdot a^c + (ca^{c-1}, 0, a^c \ln c) \cdot a^b \\ &= ((b+c)a^{b+c-1}, a^{b+c} \ln a, a^{b+c} \ln a). \end{aligned}$$

(g) We write $f = g \circ (\pi^1 + \pi^2, \pi^3)$. Thus,

$$\begin{aligned} f'(a,b,c) &= g'(a+b,c) \cdot ((\pi^1 + \pi^2)'(a,b,c), (\pi^3)'(a,b,c)) \\ &= (c(a+b)^{c-1}, (a+b)^c \ln(a+b)) \cdot ((1,1,0),(0,0,1)) \\ &= (c(a+b)^{c-1}, c(a+b)^{c-1}, (a+b)^c \ln(a+b)). \end{aligned}$$

(h) We write $f = \sin \circ (\pi^1 \cdot \pi^2)$. Thus,

$$\begin{aligned} f'(a,b) &= \sin'(ab) \cdot [a \cdot (0,1) + b \cdot (1,0)] \\ &= \cos(ab) \cdot (b,a) \\ &= (b \cos(ab), a \cos(ab)). \end{aligned}$$

(i) Let the function from the previous exercise be labelled h; then we write $f = g \circ (h, \cos 3)$. Thus,

$$\begin{aligned} f'(a,b) &= g'(h(a,b), \cos 3) \cdot (h'(a,b), (0,0)) \\ &= (\cos 3 (\sin ab)^{(\cos 3)-1}, (\sin ab)^{\cos 3} \ln \sin ab) \cdot ((b \cos(ab), a \cos(ab)), (0,0)) \\ &= ((\cos 3) b \cos(ab)(\sin ab)^{(\cos 3)-1}, (\cos 3) a \cos(ab)(\sin ab)^{(\cos 3)-1}). \end{aligned}$$

(j) We combine the previously computed derivatives of the components to write

$$f'(a,b) = \begin{pmatrix} b \cos ab & a \cos ab \\ \cos(a \sin b) \sin(b) & a \cos(a \sin b) \cos(b) \\ ba^{b-1} & a^b \ln a \end{pmatrix}$$

Problem 2-11. Find f' for the following, where $g\colon \mathbb{R} \to \mathbb{R}$ is continuous.

(a) $f(x,y) = \int_a^{x+y} g$.

(b) $f(x,y) = \int_a^{xy} g$.

(c) $f(x,y,z) = \int_{xy}^{\sin(x\sin(y\sin z))} g$.

Solution. Define $h\colon \mathbb{R} \to \mathbb{R}$,

$$h(x) = \int_a^x g,$$

whence we can deduce that $h' = g$.

(a) We write $f = h \circ (\pi^1 + \pi^2)$, whence

$$f'(b,c) = h'(b+c) \cdot (1,1) = (g(b+c), g(b+c)).$$

(b) We write $f = h \circ (\pi^1 \cdot \pi^2)$, whence

$$f'(b,c) = h'(bc) \cdot (c,b) = (cg(bc), bg(bc)).$$

(c) In general, for some $f = h \circ s$, we have

$$f'(x) = (h' \circ s)(x) \cdot s'(x) = g(s(x)) \cdot s'(x).$$

We write

$$f(x,y,z) = -\int_0^{xy} g + \int_0^{\sin(x\sin(y\sin z))} g,$$

hence we can reuse our previous work to write

$$f'(a,b,c) = -(bg(ab), ag(ab), 0) + g(\sin(a\sin(b\sin c))) \cdot f_d'(a,b,c),$$

where $f_d'(a,b,c)$ refers to the derivative computed in problem 2-10(d).

Problem 2-12. A function $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ 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

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

(a) Prove that if f is bilinear, then

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

(b) Prove that $Df(a,b)(x,y) = f(a,y) + f(x,b)$.

(c) Show that the formula

$$Dp(a,b)(x,y) = bx + ay$$

where $p(x,y) = x \cdot y$ is a special case of (b).

Solution. Let $\{e_1, \ldots, e_n\}$ and $\{e_1, \ldots, e_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. Then, given any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, we have

$$f(x, y) = \sum_{i=1}^{n} x^i f(e_i, y) = \sum_{i=1}^{n} \sum_{j=1}^{m} x^i y^j f(e_i, e_j).$$

Thus, a bilinear map is completely determined by the values of all $f(e_i, e_j)$. By setting $M' = \max |f(e_i, e_j)|$, we see that

$$|f(x, y)| \leq \sum_{i=1}^{n} |x^i| \sum_{j=1}^{m} |y^i| \cdot M' \leq M' \sum_{i=1}^{n} |x^i| \cdot (m|y|) \leq M \cdot (n|x|) \cdot (m|y|).$$

In other words, we can always choose $M > 0$ such that

$$|f(x, y)| \leq M \cdot |x| \cdot |y|.$$

(a) Recall that when evaluating the limit

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

we can write $|f(h, k)| \leq M|h||k| \leq M(|h|^2 + |k|^2)$ (we have applied a loose AM-GM inequality). Thus,

$$0 \leq \lim_{(h,k) \to 0} \frac{|f(h, k)|}{|(h, k)|} \leq \lim_{(h,k) \to 0} \frac{M(|h|^2 + |k|^2)}{\sqrt{|h|^2 + |k|^2}} = \lim_{(h,k) \to 0} M\sqrt{|h|^2 + |k|^2} = 0,$$

hence the desired limit exists and is equal to zero by the squeeze theorem.

(b) We claim that

$$\lim_{(h,k) \to 0} \frac{|f(a + h, b + k) - f(a, b) - [f(a, k) + f(h, b)]|}{|(h, k)|} = 0.$$

Expand

$$f(a + h, b + k) = f(a, b + k) + f(h, b + k) = f(a, b) + f(a, k) + f(h, b) + f(h, k).$$

This means that our claim is equivalent to

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

which is indeed true.

(c) This follows directly from the fact that p is bilinear: note that

$$\begin{aligned} p(ax, y) &= ax \cdot y = ap(x, y) = x \cdot ay = p(x, ay) \\ p(x_1 + x_2, y) &= (x_1 + x_2)y = x_1 y + x_2 y = p(x_1, y) + p(x_2, y), \\ p(x, y_1 + y_2) &= x(y_1 + y_2) = xy_1 + xy_2 = p(x, y_1) + p(x, y_2). \end{aligned}$$

Problem 2-13. Define $\mathrm{IP}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $\mathrm{IP}(x, y) = \langle x, y\rangle$.

(a) Find $D(\mathrm{IP})(a, b)$ and $(\mathrm{IP})'(a, b)$.

(b) If $f, g\colon \mathbb{R} \to \mathbb{R}^n$ are differentiable and $h\colon \mathbb{R} \to \mathbb{R}$ is defined by $h(t) = \langle f(t), g(t)\rangle$, show that
$$h'(a) = \langle f'(a)^\top, g(a)\rangle + \langle f(a), g'(a)^\top\rangle.$$

(c) If $f\colon \mathbb{R} \to \mathbb{R}^n$ is differentiable and $|f(t)| = 1$ for all t, show that $\langle f'(t)^\top, f(t)\rangle = 0$.

(d) Exhibit a differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ such that the function $|f|$ defined by $|f|(t) = |f(t)|$ is not differentiable.

Solution.

(a) Note that IP is bilinear, hence
$$D(\mathrm{IP})(a, b)(x, y) = \mathrm{IP}(a, y) + \mathrm{IP}(x, b) = \sum_{i=1}^{n} b^i x^i + \sum_{i=1}^{n} a^i y^i.$$

Thus,
$$(\mathrm{IP})'(a, b) = (b, a).$$

(b) Write $h = \mathrm{IP} \circ (f, g)$, thus
$$\begin{aligned} h'(t) &= (\mathrm{IP})'(f(t), g(t)) \cdot (f'(t), g'(t)) \\ &= (g(t), f(t)) \cdot (f'(t), g'(t)) \\ &= \langle g(t), f'(t)^\top\rangle + \langle f(t), g'(t)^\top\rangle. \end{aligned}$$

(c) Note that by setting $f = g$, we have $h(t) = \langle f(t), f(t)\rangle = |f(t)|^2 = 1$, thus the derivative of this constant function must be $h'(t) = 0$. On the other hand, we know that
$$h'(t) = 2\langle f(t), f'(t)^\top\rangle,$$
which immediately gives $\langle f(t), f'(t)^\top\rangle = 0$.

(d) Set $f(t) = t$, whence $|f|$ is not differentiable at 0.

Problem 2-14. Let E_i, $i = 1, \ldots, k$ be Euclidean spaces of various dimensions. A function $f\colon E_1 \times \cdots \times E_k \to \mathbb{R}^p$ is called *multilinear* if for each choice of $x_j \in E_j$, $j \neq i$ the function $g\colon E_i \to \mathbb{R}^p$ defined by
$$g(x) = f(x_1, \ldots, x_{i-1}, x, x_{x+1}, \ldots, x_k)$$
is a linear transformation.

(a) If f is multilinear and $i \neq j$, show that for $h = (h_1, \ldots, h_k)$ with $h_l \in E_l$, we have
$$\lim_{h\to 0} \frac{|f(a_1, \ldots, h_i, \ldots, h_j, \ldots, a_k|}{|h|} = 0.$$

(b) Prove that
$$Df(a_1, \ldots, a_k)(x_1, \ldots, x_k) = \sum_{i=1}^{k} f(a_1, \ldots, x_i, \ldots, a_k).$$

Solution. Like before, a multilinear map is completely determined by the values of $f(e_{j_1}, \dots, e_{j_k})$, where each set e_{j_i} forms a basis of E_i.

$$f(x_1, \dots, x_k) = \sum_{j_1, \dots, j_k} x_1^{j_1} \dots x_k^{j_k} f(e_{j_1}, \dots, e_{j_k}).$$

Thus, by setting M' as the maximum of the norm of these 'base vectors' of the range and using the triangle equality like before, we can again pick $M > 0$ such that

$$|f(x_1, \dots, x_k)| \leq M \cdot |x_1| \dots |x_k|.$$

(a) Again, write

$$0 \leq \frac{|f(a_1, \dots, h_i, \dots, h_j, \dots, a_k)|}{|h|} \leq \frac{M \cdot |a_1| \dots |h_i| \dots |h_j| \dots |a_k|}{|(h_i, h_j)|}.$$

Take the limit $h \to 0$, and note that we already know that as $h_i, h_j \to 0$, we have $|h_i| \cdot |h_j| / |(h_1, h_j)| \to 0$ from our work on bilinear functions. Thus, the squeeze theorem gives the desired limit.

(b) Denote the target expression on the right hand side by $d(a)(x)$ for short. Then, we claim that

$$\lim_{h \to 0} \frac{|f(a_1 + h_1, \dots, a_k + h_k) - f(a_1, \dots, a_k) - d(a)(h)|}{|h|} = 0.$$

Expand

$$\begin{aligned} f(a_1 + h_1, \dots, a_k + h_k) &= f(a_1, a_2 + h_2, \dots, a_k + h_k) + f(h_1, a_2 + h_2, \dots, a_k + h_k) \\ &\quad \vdots \qquad\qquad \vdots \\ &= \sum_{\delta_i \in \{0,1\}} f((1 - \delta_1)a_1 + \delta_1 h_1, \dots, (1 - \delta_k)a_k + \delta_k h_k). \end{aligned}$$

In other words, we have all 2^k possible terms with either a_i or h_i in the ith parameter. Coming back to the numerator of our limit, we have subtracted away the (single) term with all $\delta_i = 0$, and those (k) terms with exactly one $\delta_i = 1$. Using the triangle inequality, we see that our limit expression is bounded above by

$$\sum_{\substack{\delta_i \in \{0,1\} \\ \text{At least 2 of } \delta_i = 1}} \frac{|f((1 - \delta_1)a_1 + \delta_1 h_1, \dots, (1 - \delta_k)a_k + \delta_k h_k)|}{|h|}.$$

Take the limit $h \to 0$. Now, we already know that for those terms with exactly 2 of $\delta_i = 1$,

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

For those terms with 3 or more of $\delta_i = 1$, we perform exactly the same process as in (a), using the fact that with $j \geq 3$,

$$0 \leq \lim_{h \to 0} \frac{|h_{i_1}| \dots |h_{i_j}|}{|h|} \leq \lim_{h \to 0} |h_{i_3}| \dots |h_{i_j}| \cdot \lim_{h \to 0} \frac{|h_{i_1}| \cdot |h_{i_2}|}{|h|} = 0.$$

The latter follows because the individual limits exist and are zero. Thus, all the terms in our sum vanish as $h \to 0$. The squeeze theorem now guarantees that the desired limit also exists, and is equal to zero.

Problem 2-15. Regard an $n \times n$ matrix as a point in the n-fold product $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ by considering each row as a member of $\mathbb{R}^n$.

(a) Prove that $\det\colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ is differentiable, and

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

(b) If $a_{ij}\colon \mathbb{R} \to \mathbb{R}$ are differentiable and $f(t) = \det(a_{ij}(t))$, show that

$$f'(t) = \sum_{j=1}^{n} \det \begin{pmatrix} a_{11}(t) & \ldots & a_{1n}(t) \\ \vdots & & \vdots \\ a'_{j1}(t) & \ldots & a'_{jn}(t) \\ \vdots & & \vdots \\ a_{n_1}(t) & \ldots & a_{nn}(t) \end{pmatrix}.$$

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

$$\sum_{j=1}^{n} a_{ji}(t) s_j(t) = b_i(t), \qquad i = 1, \ldots, n.$$

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

Solution.

(a) There is nothing to prove, since det is a multilinear map in the rows of a given matrix by construction. The rest follows directly from the formula for the derivative of a multilinear map, derived in the previous exercise.

(b) Denote $a_j = (a_{j1}, \ldots, a_{jn})$, where each $a_1, \ldots, a_n\colon \mathbb{R} \to \mathbb{R}^n$ is a row of the matrix; this also gives $a'_j = (a'_{j1}, \ldots, a'_{jn})$. Then, $f = \det \circ (a_1, \ldots, a_n)$, so the chain rule directly gives

$$f'(t) = (\det)'(a_1(t), \ldots, a_n(t))(a'_1(t), \ldots, a'_n(t)) = \sum_{j=1}^{n} \det \begin{pmatrix} a_1(t) \\ \vdots \\ a'_j(t) \\ \vdots \\ a_n(t) \end{pmatrix}.$$

(c) Note that we can write this system of equations as

$$\begin{pmatrix} a_{11}(t) & \ldots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n_1}(t) & \ldots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{pmatrix} = \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix}.$$

Denote this as

$$A(t)s(t) = b(t).$$

The matrix on the left is given to be invertible. Thus, we can write the vector $s(t) = A^{-1}(t)b(t)$, which when expanded gives each $s_j(t)$ as a polynomial expression of all $a_{ij}(t)$ and $b_i(t)$. Applying the summation, product, quotient, and chain rules shows that each $s_j(t)$ must be differentiable.

By differentiating each equation, we have

$$\sum_{j=1}^{n} a'_{ji}(t)s_j(t) + a_{ij}(t)s'_j(t) = b'_i(t),$$

which we can concisely write again as

$$A'(t)s(t) + A(t)s'(t) = b'(t).$$

This immediately gives

$$\begin{aligned} s'(t) &= A^{-1}(t)[b'(t) - A'(t)s(t)] \\ &= A^{-1}(t)[b'(t) - A'(t)[A^{-1}(t)b(t)]] \\ &= A^{-1}(t)b'(t) - A^{-1}(t)A'(t)A^{-1}(t)b(t). \end{aligned}$$

Problem 2-16. Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and has a differentiable inverse $f^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n$. Show that $(f^{-1})'(a) = [f'(f^{-1}(a))]^{-1}$.

Solution. Use $f \circ f^{-1} = x$ (the identity function, whose derivative is itself). Thus, the chain rule gives

$$f'(f^{-1}(a)) \cdot (f^{-1})'(a) = 1,$$

the right hand side being the identity matrix. Left multiplying by $[f'(f^{-1}(a))]^{-1}$ immediately gives the result.

2.3 Partial Derivatives

Problem 2-17. Find the partial derivatives of the following functions.

(a) $f(x, y, z) = x^y$.

(b) $f(x, y, z) = z$.

(c) $f(x, y) = \sin(x \sin(y))$.

(d) $f(x, y, z) = \sin(x \sin(y \sin(z)))$.

(e) $f(x, y, z) = x^{y^z}$.

(f) $f(x, y, z) = x^{y+z}$.

(g) $f(x, y, z) = (x + y)^z$.

(h) $f(x, y) = \sin(xy)$.

(i) $f(x, y) = [\sin(xy)]^{\cos 3}$.

Solution. Note that we could simply reuse our work from problem 2-10, with the knowledge that the partial derivatives are merely the entries of the Jacobian. Here, we apply the product and chain rules in one step each.

For brevity, denote $D_i f(x) \equiv d_i$.

(a)

$$d_1 = yx^{y-1}, \qquad d_2 = x^y \ln y, \qquad d_3 = 0.$$

(b)
$$d_1 = d_2 = 0, \qquad d_3 = 1.$$

(c)
$$d_1 = \sin(y)\cos(x\sin y), \qquad d_2 = x\cos(y)\cos(x\sin y).$$

(d)
$$\begin{aligned} d_1 &= \sin(y\sin z)\cos(x\sin(y\sin z)), \\ d_2 &= x\sin(z)\cos(y\sin z)\cos(x\sin(y\sin z)), \\ d_3 &= xy\cos(z)\cos(y\sin z)\cos(x\sin(y\sin z)). \end{aligned}$$

(e)
$$d_1 = y^z x^{y^z-1}, \qquad d_2 = x^{y^z}\ln(x)\cdot zy^{z-1}, \qquad d_3 = x^{y^z}\ln(x)\cdot y^z\ln(y).$$

(f)
$$d_1 = (y+z)x^{y+z-1}, \qquad d_2 = x^{y+z}\ln x, \qquad d_3 = x^{y+z}\ln x.$$

(g)
$$d_1 = z(x+y)^{z-1}, \qquad d_2 = z(x+y)^{z-1}, \qquad d_3 = (x+y)^z\ln(x+y).$$

(h)
$$d_1 = y\cos(xy), \qquad d_2 = x\cos(xy).$$

(i)
$$d_1 = \cos(3)(\sin(xy))^{\cos(3)-1}\cdot y\cos(xy), \qquad d_2 = \cos(3)(\sin(xy))^{\cos(3)-1}\cdot x\cos(xy).$$

Problem 2-18. Find the partial derivatives of the following, where $g\colon \mathbb{R}\to\mathbb{R}$ is continuous.

(a) $f(x,y) = \int_a^{x+y} g$.

(b) $f(x,y) = \int_x^y g$.

(c) $f(x,y) = \int_a^{xy} g$.

(d) $f(x,y) = \int_a^{\int_b^y g} g$.

Solution. Define $h_c\colon \mathbb{R}\to\mathbb{R}$,
$$h_c(x) = \int_c^x g,$$
whence we can deduce that $h_c' = g$.

(a) We compute
$$D_1 f(x,y) = D_1 h(x+y) = h'(x+y)\cdot D_1(x+y) = g(x+y),$$
$$D_2 f(x,y) = D_2 h(x+y) = h'(x+y)\cdot D_2(x+y) = g(x+y).$$

(b) Write
$$f(x,y) = -\int_0^x g + \int_0^y g = h(y) - h(x).$$

Thus,
$$D_1 f(x,y) = -h'(x) = -g(x), \qquad D_2 f(x,y) = h'(y) = g(y).$$

(c)
$$D_1f(x,y) = D_1h(xy) = h'(xy) \cdot D_1(xy) = g(xy) \cdot y,$$
$$D_2f(x,y) = D_2h(xy) = h'(xy) \cdot D_2(xy) = g(xy) \cdot x.$$

(d) Write $f(x,y) = h_a(h_b(y))$. Thus, $D_1f(x,y) = 0$, and
$$D_2f(x,y) = h_a'(h_b(y)) \cdot h_b'(y) = g\left(\int_b^y g\right) \cdot g(y).$$

Problem 2-19. If
$$f(x,y) = x^{x^{x^{x^y}}} + (\log x)(\arctan(\arctan(\arctan(\sin(\cos xy) - \log(x+y))))),$$
find $D_2f(1,y)$.

Solution. Since we are evaluating the partial derivative with respect to the second variable at points $(1,y)$, we can restrict our attention to the line $x = 1$, along which we have $f(1,y) = 1$. This means that $D_2f(1,y) = 0$.

We have to justify this carefully by nothing that $D_2f(a,b)$ gives the derivative of the function 'sliced' along the $x = a$ line, at the point $x = a, y = b$. The given function is constant along the $x = 1$ line, which means that the partial derivative along this line is zero.

Problem 2-20. Find the partial derivatives of f in terms of the derivatives of g and h.

(a) $f(x,y) = g(x)h(y)$.
(b) $f(x) = g(x)^{h(y)}$.
(c) $f(x,y) = g(x)$.
(d) $f(x,y) = g(y)$.
(e) $f(x,y) = g(x+y)$.

Solution.

(a)
$$D_1f(x,y) = g'(x)h(y), \qquad D_2f(x,y) = g(x)h'(y).$$

(b)
$$D_1f(x,y) = h(y)g(x)^{h(y)-1} \cdot g'(x), \qquad D_2f(x,y) = g(x)^{h(y)} \log g(x) \cdot h'(y).$$

(c)
$$D_1f(x,y) = g'(x), \qquad D_2f(x,y) = 0.$$

(d)
$$D_1f(x,y) = 0, \qquad D_2f(x,y) = g'(y).$$

(e)
$$D_1f(x,y) = g'(x+y), \qquad D_2f(x,y) = g'(x+y).$$

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

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

(a) Show that $D_2 f(x,y) = g_2(x,y)$.

(b) How should f be defined so that $D_1 f(x,y) = g_1(x,y)$?

(c) Find a function $f \colon \mathbb{R}^2 \to \mathbb{R}$ 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 $D_2 f(x,y) = x$.

Solution.

(a) Note that the first term is purely a function of x, and hence vanishes upon applying D_2. Differentiating the second one with respect to the second variable while holding x constant, we see that $g_2(x,t) \equiv \tilde{g_2}(t)$ is really a function of one variable t, so the Fundamental Theorem of Calculus (regarding antiderivatives) gives us

$$D_2 f(x,y) = \frac{d}{dy}\int_0^y g_2(x,t)\,dt = g_2(x,y).$$

(b) We may have defined

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

(c) The required maps are

$$(x,y) \mapsto \frac{1}{2}(x^2+y^2), \qquad (x,y) \mapsto xy.$$

Problem 2-22. If $f \colon \mathbb{R}^2 \to \mathbb{R}$ and $D_2 f = 0$, show that f is independent of the second variable. If $D_1 f = D_2 f = 0$, show that f is constant.

Solution. Fix $x \in \mathbb{R}$, and let $h_x \colon \mathbb{R} \to \mathbb{R}$, $y \mapsto f(x,y)$. Then, we can see that $h'_x(y) = D_2 f(x,y) = 0$ by definition, hence h_x is a constant function. Thus, $h(y_1) = h(y_2)$ for any $y_1, y_2 \in \mathbb{R}$, i.e. $f(x,y_1) = f(x,y_2)$.

An analogous argument can be made to show that if $D_1 f = 0$, then f is independent of its first variable. Thus, when both $D_1 f = D_2 f = 0$, f is independent of both its variables, and we have shown that such a function must be constant, with $f(x,y) = f(0,0)$ everywhere.

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

(a) If $f \colon A \to \mathbb{R}$ and $D_1 f = D_2 f = 0$, show that f is constant.

(b) Find a function $f \colon A \to \mathbb{R}$ such that $D_2 f = 0$ but f is not independent of the second variable.

Solution.

(a) It is clear that the function $g(y) = f(-1,y)$ is constant for all $y \in \mathbb{R}$, since $g'(y) = D_2 f(-1,y) = 0$ is well-defined everywhere. Thus, all $f(-1,y_1) = f(-1,y_2)$. Given $(x_1,y_1), (x_2,y_2)$, there are three possible cases.

Case I: Both $y_1 = y_2 = 0$. Thus, we must have both $x_1, x_2 < 0$. Since $D_1 f(x, 0) = 0$ along the entire line joining $(x_1, 0)$ and $(x_2, 0)$, we must have $f(x_1, 0) = f(x_2, 0)$.

Case II: Exactly one of $y_1, y_2 = 0$, say $y_1 = 0$. Then, $x_1 < 0$, so repeating the process from the previous case, $f(x_1, y_1) = f(x_1, 0) = f(-1, 0)$. Since $y_2 \neq 0$, we have $D_1 f(x, y_2) = 0$ for all $x \in \mathbb{R}$, hence $f(x, y_2)$ is constant. This gives $f(x, y_2) = f(-1, y_2) = f(-1, 0)$. Together, $f(x_1, y_1) = f(x_2, y_2)$.

Case III: Neither $y_1, y_2 = 0$. Then, we repeat the same process as in the previous case to write $f(x_1, y_1) = f(-1, y_1) = f(-1, 0)$ and $f(x_2, y_2) = f(-1, y_2) = f(-1, 0)$ to conclude that $f(x_1, y_1) = f(x_2, y_2)$.

(b) Define $f\colon A \to \mathbb{R}$,

$$f(x, y) = \begin{cases} 0, & \text{if } x < 0, \\ 1, & \text{if } x \geq 0, y > 0, \\ 2, & \text{if } x \geq 0, y < 0. \end{cases}$$

It is easily checked that $D_2 f = 0$, yet $f(1, 1) = 1 \neq 2 = f(1, -1)$. Note that in the region $x < 0$, we have $f(x, y) = 0$ so $D_2 f(x, y) = 0$ there. In the region $x \geq 0$ and $y > 0$, we have $f(x, y) = 1$ so $D_2 f(x, y) = 0$ there; similarly when $x \geq 0$ and $y < 0$, $D_2 f(x, y) = 0$.

Problem 2-24. Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by

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

(a) Show that $D_2 f(x, 0) = x$ for all x and $D_1 f(0, y) = -y$ for all y.

(b) Show that $D_{1,2} f(0, 0) \neq D_{2,1} f(0, 0)$.

Solution.

(a) When $(x, y) \neq 0$, we can calculate by brute force,

$$D_2 f(x, y) = \frac{(x^3 - 3xy^2)(x^2 + y^2) - (x^3 y - xy^3)(2y)}{(x^2 + y^2)^2}$$
$$D_2 f(x, 0) = \frac{x^3 \cdot x^2}{(x^2)^2} = x,$$

$$D_1 f(x, y) = \frac{(3x^2 y - y^3)(x^2 + y^2) - (x^3 y - xy^3)(2x)}{(x^2 + y^2)^2}$$
$$D_1 f(0, y) = \frac{-y^3 \cdot y^2}{(y^2)^2} = -y.$$

At $(0, 0)$, we use the basic definition to calculate

$$D_1 f(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0,$$

and similarly $D_2 f(0, 0) = 0$.

(b) We simply iterate

$$D_{1,2} f(0, 0) = D_2(D_1 f(x, y))(0, 0) = D_2(D_1 f(0, y))(0, 0) = -1,$$
$$D_{2,1} f(0, 0) = D_1(D_2 f(x, y))(0, 0) = D_1(D_2 f(x, 0))(0, 0) = +1.$$

In the first case, we can restrict our attention to the line $x = 0$ while calculating the outermost D_2, since we are interested in the variation of $D_1 f(x, y)$ only along that line. The second case is similar.

Problem 2-25. Define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} e^{-x^{-2}}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}.$$

Show that f is a C^∞ function, and $f^{(i)}(0) = 0$ for all i.

Solution. It is clear that when $x \neq 0$, we can choose an arbitrarily small neighbourhood around x such that it does not contain 0; then, the derivatives of f can be simply evaluated via the chain rule,

$$\begin{aligned} f(x) &= \exp(-x^{-2}), \\ f'(x) &= \exp(-x^{-2}) \cdot 2x^{-3}, & x^3 f'(x) &= \exp(-x^{-2}) \cdot 2 \\ x^3 f''(x) + 3x^2 f'(x) &= \exp(-x^{-2}) \cdot 2x^{-3}, & x^6 f''(x) &= \exp(-x^{-2}) \cdot (2 - 6x^2) \\ \vdots \quad & \quad \vdots & \vdots \quad & \quad \vdots \end{aligned}$$

It is clear that f is infinitely differentiable on this neighbourhood, because at each stage, we have a composition of exponential and polynomial functions, each of which are differentiable. Suppose that

$$x^{3n} f^{(n)}(x) = \exp(-x^{-2}) \cdot p_n(x)$$

for some polynomial p_n. Then,

$$\begin{aligned} x^{3n} f^{(n+1)}(x) + 3n x^{3n-1} f^{(n)}(x) &= \exp(-x^{-2}) \cdot (p_n'(x) + 2x^{-3} p_n(x)), \\ x^{3n+3} f^{(n+1)}(x) + 3nx^2 \cdot x^{3n} f^{(n)}(x) &= \exp(-x^{-2}) \cdot (x^3 p_n'(x) + 2p_n(x)), \end{aligned}$$

whence

$$x^{3n+3} f^{(n+1)}(x) = \exp(-x^{-2}) \cdot (x^3 p_n'(x) - 3nx^2 p_n(x) + 2p_n(x)).$$

It is clear that the degree of p_n jumps up by 2 at each step, thus the degree of each p_n is precisely $2n - 2$. To verify this, if the leading term of p_n at any stage is $a_n x^{2n-2}$, then the new leading term must be $2(n-1)a_n x^{2n} - 3n a_n x^{2n} = -(n+2)a_n x^{2n}$. In other words, the leading term is never 'cancelled out', and the leading coefficient is always $(-1)^{n+1} \cdot (n+1)!$.

Now, the series definition of exp gives

$$\exp(x^{-2}) = 1 + \frac{1}{x^2} + \cdots + \frac{1}{n! x^{2k}} + \cdots > \frac{1}{n! x^{2k}}$$

hence

$$0 < \exp(-x^{-2}) < n! x^{2k}.$$

for all $k \in \mathbb{N}$.

First, we show that $f'(0) = 0$ by evaluating the limit

$$\lim_{x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0} \frac{\exp(-x^{-2})}{x} \leq \lim_{x \to 0} \frac{x^2}{x} = 0,$$

hence the result follows by the squeeze theorem. Now if $f^{(n)}(0) = 0$, then we can show that $f^{(n+1)}(0) = 0$ by evaluating

$$\lim_{x \to 0} \frac{f^{(n)}(x) - f^{(n)}(0)}{x} = \lim_{x \to 0} \frac{\exp(-x^{-2}) \cdot p_n(x)}{x \cdot x^{3n}}.$$

Here, we choose $k = 3n$, whence

$$0 < \frac{\exp(-x^{-2}) \cdot |p_n(x)|}{|x \cdot x^{3n}|} < \frac{(3n)! x^{6n} \cdot |p_n(x)|}{|x^{3n+1}|}.$$

The numerator has degree $6n + (2n - 2) = 8n - 2 > 3n + 1$ for all $n \geq 1$. Thus, taking limits $x \to 0$, we see that $f^{(n+1)}(0) = 0$ by the squeeze theorem.

Problem 2-26. Let

$$f(x) = \begin{cases} e^{-(x-1)^{-2}} \cdot e^{-(x+1)^{-2}}, & \text{if } x \in (-1,1), \\ 0, & \text{if } x \notin (-1,1). \end{cases}$$

(a) Show that $f\colon \mathbb{R} \to \mathbb{R}$ is a C^∞ function which is positive on $(-1,1)$ and 0 elsewhere.

(b) Show that there is a C^∞ function $g\colon \mathbb{R} \to [0,1]$ such that $g(x) = 0$ for $x \le 0$ and $g(x) = 1$ for $x \ge \epsilon$.

(c) If $a \in \mathbb{R}^n$, define $g\colon \mathbb{R}^n \to \mathbb{R}$ by

$$g(x) = f([x^1 - a^1]/\epsilon) \cdot \cdots \cdot f([x^n - a^n]/\epsilon).$$

Show that g is a C^∞ function which is positive on

$$(a^1 - \epsilon, a^1 + \epsilon) \times \cdots \times (a^n - \epsilon, a^n + \epsilon)$$

and zero elsewhere.

(d) If $A \subset \mathbb{R}^n$ is open and $C \subset A$ is compact, show that there is a non-negative C^∞ function $f\colon A \to \mathbb{R}$ such that $f(x) > 0$ for $x \in C$ and $f = 0$ outside of some closed set contained in A.

(e) Show that we can choose such an f so that $f\colon A \to [0,1]$ and $f(x) = 1$ for $x \in C$.

Solution.

(a) Consider the function $h\colon \mathbb{R} \to \mathbb{R}$ defined by

$$h(x) = \begin{cases} e^{-x^{-2}}, & \text{if } x > 0, \\ 0, & \text{if } x \le 0. \end{cases}.$$

The proof that this is C^∞ is identical to the previous problem. Now, observe that

$$f(x) = h(1-x) \cdot h(1+x)$$

The product and chain rules now guarantee that f is C^∞, with all $f^{(n)}(\pm 1) = 0$; the latter fact can be shown directly from

$$f^{(n)}(x) = \sum_{i=0}^{n} \binom{n}{i} (-1)^i h^{(i)}(1-x) h^{(n-i)}(1+x).$$

Furthermore, note that exp is always positive, hence f is positive on $(-1,1)$.

(b) Define $\tilde{f}(x) = f(2x/\epsilon - 1)$; then $\tilde{f}$ is also C^∞, strictly positive on $(0,\epsilon)$, and zero elsewhere. Now, set

$$\alpha = \int_0^\epsilon \tilde{f} > 0, \qquad g(x) = \frac{1}{\alpha} \int_0^x \tilde{f}$$

for $x > 0$, and $g = 0$ elsewhere. Then, observe that g is also C^∞ – the first derivative is simply $\tilde{f}/\alpha$, which is also C^∞. Furthermore, when $x \ge \epsilon$, the integral sees no contribution from $\epsilon \le t \le x$, since $f(t) = 0$ there. Thus, $g(x) = \alpha/\alpha = 1$ for $x \ge \epsilon$.

(c) Since f is C^∞, any partial derivative of g is a product of terms of the form $f^{(i)}([x^j - a^j]/\epsilon)$ (along with some constant factors of ϵ), all of which are continuous and differentiable (indeed, all C^∞). This shows that g is also C^∞.

Note that if any component $x^j \notin (a^j - \epsilon, a^j + \epsilon)$, then either $(x^j - a^j)/\epsilon \le -1$, or $(x^j - a^j)/\epsilon \ge 1$. In either case, $f([x^j - a^j]/\epsilon) = 0$, so $g(x) = 0$ outside the rectangle A. Otherwise, all the contributions by f in the product are positive, so g is positive within the rectangle A.

(d) Choose an open cover of the compact set C using non-empty open rectangles (say, open rectangles of side 1, with each point in C getting its own rectangle); from this, extract a finite subcover of open rectangles $A_1, \dots, A_k$. Construct functions $g_1, \dots, g_k$ as in the previous part such that each g_i is positive on the corresponding A_i, and zero elsewhere. Finally, set $f\colon A \to \mathbb{R}$, $f = g_1 + \cdots + g_k$. This is clearly positive on C, and zero outside the union of $A_1, \dots, A_k$; this union is open, hence its complement within A is closed.

Note that f is continuous, hence maps the compact set C to a compact interval. In other words, f attains its minimum on C; this minimum cannot be 0 since f is positive. Thus, $f \geq \epsilon > 0$ on C.

(e) Let $g\colon \mathbb{R} \to \mathbb{R}$ be C^∞, $g(x) = 0$ for $x \leq 0$, $g(x) > 0$ on $(0, \epsilon)$, and $g(x) = 1$ for $x \geq \epsilon$ (we can do this via part (b)). Also choose f from the previous part, and have $f \geq \epsilon$ on C (simply rescale if necessary). Then, set $F = g \circ f$, whence we have normalized $F(x) = g(f(x)) = 1$ when $x \in C$, and $F = 0$ wherever $f = 0$.

Problem 2-27. Define $g, h\colon \{x \in \mathbb{R}^2 : |x| \leq 1\} \to \mathbb{R}^3$ by

$$\begin{aligned} g(x,y) &= (x, y, \sqrt{1 - x^2 - y^2}), \\ h(x,y) &= (x, y, -\sqrt{1 - x^2 - y^2}). \end{aligned}$$

Show that the minimum of f on $\{x \in \mathbb{R}^3 : |x| = 1\}$ is either the maximum of $f \circ g$ or the maximum of $f \circ h$ on $\{x \in \mathbb{R}^2 : |x| \leq 1\}$.

Solution. Suppose that f attains a maximum at (x_0, y_0, z_0); then, $x_0^2 + y_0^2 + z_0^2 = 1$, so $z_0 = \pm\sqrt{1 - x_0^2 - y_0^2}$. Thus, the point (x_0, y_0, z_0) is one of $g(x_0, y_0)$ or $h(x_0, y_0)$. In this way, note that the collection of all possible points $g(x,y)$ together with $h(x,y)$ gives precisely the unit sphere $\{x \in \mathbb{R}^3 : |x| = 1\}$: no more (clearly each output point has unit norm), no less (we have just shown that every point on the unit sphere satisfies $x^2 + y^2 + z^2 = 1$, and hence has a pre-image either in g or h). Thus, the maximum $f(x_0, y_0, z_0) = \alpha$ is at least as large as every possible value $f \circ g(x,y)$ or $f \circ h(x,y)$. This shows that the maximum of f is attained at the maximum of at least one of $f \circ g$ or $f \circ h$.

2.4 Derivatives

Problem 2-28. Find expressions for the partial derivatives of the following functions.

(a) $F(x,y) = f(g(x)k(y), g(x) + h(y))$.

(b) $F(x,y,z) = f(g(x+y), h(y+z))$.

(c) $F(x,y,z) = f(x^y, y^z, z^x)$.

(d) $F(x,y) = f(x, g(x), h(x,y))$.

Solution.

(a) Set $a = (g(x)k(y), g(x) + h(y))$, whence

$$\begin{aligned} D_1F(x,y) &= D_1f(a) \cdot g'(x)k(y) + D_2f(a) \cdot g'(x), \\ D_2F(x,y) &= D_1f(a) \cdot g(x)k'(y) + D_2f(a) \cdot h'(y). \end{aligned}$$

(b) Set $a = (g(x+y), h(y+z))$, whence

$$\begin{aligned} D_1F(x,y,z) &= D_1f(a) \cdot g'(x+y), \\ D_2F(x,y,z) &= D_1f(a) \cdot g'(x+y) + D_2f(a) \cdot h'(y+z), \\ D_3F(x,y,z) &= D_2f(a) \cdot h'(y+z). \end{aligned}$$

(c) Set $a = (x^y, y^z, z^x)$, whence

$$\begin{aligned} D_1F(x,y,z) &= Df_1(a)\cdot yx^{y-1} + D_3f(a)\cdot z^x\log z,\\ D_2F(x,y,z) &= Df_2(a)\cdot zy^{z-1} + D_1f(a)\cdot x^y\log x,\\ D_3F(x,y,z) &= Df_3(a)\cdot xz^{x-1} + D_2f(a)\cdot y^z\log y. \end{aligned}$$

(d) Set $a = (x, g(x), h(x,y))$, whence

$$\begin{aligned} D_1F(x,y,z) &= Df_1(a) + D_2f(a)\cdot g'(x) + D_3f(a)\cdot D_1h(x,y),\\ D_2F(x,y,z) &= D_3f(a)\cdot D_2h(x,y). \end{aligned}$$

Problem 2-29. Let $f\colon \mathbb{R}^n \to \mathbb{R}$. For $x \in \mathbb{R}^n$, the limit

$$\lim_{t\to 0}\frac{f(a+tx)-f(a)}{t},$$

if it exists, is denoted $D_xf(a)$, and is called the *directional derivative* of f at a, in the direction x.

(a) Show that $D_{e_i}f(a) = D_if(a)$.
(b) Show that $D_{tx}f(a) = tD_xf(a)$.
(c) If f is differentiable at a, show that $D_xf(a) = Df(a)(x)$ and therefore

$$D_{x+y}f(a) = D_xf(a) + D_yf(a).$$

Solution.

(a) Note that the definitions directly give

$$D_{e_i}f(a) = \lim_{t\to 0}\frac{f(a+te_i)-f(a)}{t} = \lim_{t\to 0}\frac{f(a^1,\dots,a^i+t,\dots,a^n)-f(a)}{t} = D_if(a).$$

(b) Note that

$$D_{tx}f(a) = \lim_{h\to 0}\frac{f(a+htx)-f(a)}{h} = \lim_{ht\to 0}\frac{f(a+(ht)x)}{ht}\cdot t = D_xf(a)\cdot t.$$

(c) If f is differentiable at a, then we know that $Df(a)$ is the unique linear transformation satisfying

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

Specifically, $tx \to 0$ as $t \to 0$, so

$$\lim_{t\to 0}\frac{|f(a+tx)-f(a)-Df(a)(tx)|}{|t||x|} = 0.$$

Now, we know that

$$\lim_{t\to 0}\frac{f(a+tx)-f(a)}{t} = D_xf(a), \qquad \lim_{t\to 0}\frac{|f(a+tx)-f(a)-D_xf(a)t|}{|t|} = 0.$$

Thus,

$$\begin{aligned} 0 &\le |Df(a)(x) - D_xf(a)|\\ &= \frac{|Df(a)(tx) - D_xf(a)t|}{|t|}\\ &\le \frac{|f(a+tx)-f(a)-Df(a)(tx)|}{|t|} + \frac{|f(a+tx)-f(a)-D_xf(a)t|}{|t|}. \end{aligned}$$

Taking the limit $t \to 0$, the right hand side vanishes, hence $Df(a)(x) = D_xf(a)$. The linearity of $Df(a)$ immediately gives

$$D_{x+y}f(a) = D_xf(a) + D_yf(a).$$

Problem 2-30. Let f be defined as in Problem 2-4. Show that $D_x f(0,0)$ exists for all x, but if $g \neq 0$, then $D_{x+y} f(0,0) = D_x f(0,0) + D_y f(0,0)$ is not true for all x and y.

Solution. Without loss of generality, let $x \in \mathbb{R}^2$, $|x| = 1$ (this will later give the existence of all remaining $D_{tx} f(0,0) = t D_x f(0,0)$). Then, we compute the limit

$$D_x f(0,0) = \lim_{t \to 0} \frac{f(0+tx) - f(0)}{t} = \lim_{t \to 0} \frac{|tx| \cdot g(tx/|tx|)}{t} = \lim_{t \to 0} g(x) = g(x).$$

Note that here, we have successfully 'cancelled' the factors of t because $g(-x) = -g(x)$.

Suppose that $g \neq 0$: we already know that this forces some $g(x) = \alpha \neq 0$ where neither $x^1, x^2 = 0$. Now,

$$D_{e_1} f(0,0) = g(e_1) = 0, \qquad D_{e_2} f(0,0) = g(e_2) = 0.$$

Write

$$D_{x^1 e_1 + x^2 e_2} f(0,0) = D_x f(0,0) = \alpha \neq 0,$$

and

$$D_{x^1 e_1} f(0,0) + D_{x^2 e_2} f(0,0) = x^1 D_{e_1} f(0,0) + x^2 D_{e_2} f(0,0) = 0.$$

This demonstrates that $D_x f(0,0)$ is not linear in x in this case.

Problem 2-31. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined as in Problem 1-26. Show that $D_x f(0,0)$ exists for all x, although f is not even continuous at $(0,0)$.

Solution. We have shown in 1-26 that given any straight line through the origin, there is an interval around $(0,0)$ on that line which is in $\mathbb{R}^2 - A$, i.e. $f = 0$ on that interval. Thus, given any vector $x \in \mathbb{R}^2$, we have shown that there is an interval $(-\delta, \delta)$ on which $f(tx) = 0$ for all $t \in (-\delta, \delta)$. This immediately shows that the limit

$$D_x f(a) = \lim_{t \to 0} \frac{f(0+tx) - f(0)}{t} = 0.$$

On the other hand, we recall that f was not continuous at the origin.

Problem 2-32.

(a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x^2 \sin(1/x), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Show that f is differentiable at 0 but f' is not continuous at 0.

(b) Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} (x^2+y^2)\sin(1/\sqrt{x^2+y^2}), & \text{if } (x,y) \neq 0, \\ 0, & \text{if } (x,y) = 0. \end{cases}$$

Show that f is differentiable at $(0,0)$ but $D_i f$ is not continuous at $(0,0)$.

Solution.

(a) We compute the limit

$$\lim_{h\to 0}\frac{|f(h)-f(0)|}{|h|}=\lim_{h\to 0}\frac{|h^2\sin(1/h)|}{|h|}=\lim_{h\to 0}|h\sin(1/h)|\leq\lim_{h\to 0}|h|=0.$$

The last inference follows since sin is bounded. Thus, we see that the first limit exists and is equal to zero, hence $f'(0)=0$.

We compute $f'(x)$ where $x\neq 0$ using the chain rule, thus obtaining

$$f'(x)=\begin{cases}2x\sin(1/x)-\cos(1/x), & \text{if } x\neq 0,\\ 0, & \text{if } x=0.\end{cases}$$

Clearly, as $x\to 0$, we have already established that

$$\lim_{x\to 0}x\sin(1/x)=0.$$

However, the limit $\lim_{x\to 0}\cos(1/x)$ does not exist: along the sequence $x_n=1/2n\pi\to 0$, we have $\cos(2n\pi)=1$ but along the sequence $x_n=1/(2n\pi+\pi/2)$, we have $\cos(2n\pi+\pi/2)=0$. Thus, f' is not continuous at 0.

(b) We compute the limit

$$\lim_{(h,k)\to 0}\frac{|f(h,k)-f(0,0)|}{|h|}=\lim_{h\to 0}\frac{|h|^2\sin(1/|h|)}{|h|}=\lim_{h\to 0}|h|\cdot|\sin(1/|h|)|\leq\lim_{h\to 0}|h|=0.$$

We have written the function f in terms of the norm and proceeded similar to the previous part here. Thus, we have shown that $Df(0,0)=0$.

We can compute the partial derivatives at $(0,0)$ by examining the limits

$$D_1f(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{h^2\sin(1/h)}{h}=0,$$

and similarly $D_2f(0,0)=0$. Elsewhere, we use the chain rule to write

$$D_1f(x,y)=\begin{cases}2x\sin(1/\sqrt{x^2+y^2})-x/\sqrt{x^2+y^2}\cdot\cos(1/\sqrt{x^2+y^2}), & \text{if } (x,y)\neq 0,\\ 0, & \text{if } (x,y)=0.\end{cases}$$

$$D_2f(x,y)=\begin{cases}2y\sin(1/\sqrt{x^2+y^2})-y/\sqrt{x^2+y^2}\cdot\cos(1/\sqrt{x^2+y^2}), & \text{if } (x,y)\neq 0,\\ 0, & \text{if } (x,y)=0.\end{cases}$$

In the first case, look at the restricted function

$$D_1f(x,0)=\begin{cases}2x\sin(1/|x|)-x/|x|\cdot\cos(1/|x|), & \text{if } x\neq 0,\\ 0, & \text{if } x=0.\end{cases}$$

This is very similar to the previous part: note that $\lim_{x\to 0}x\sin(1/|x|)=0$, however the limit $\lim_{x\to 0}x/|x|\cdot\cos(1/|x|)$ does not for exactly the same reasons as given before. Thus, D_1f cannot be continuous. An analogous argument shows that D_2f cannot be continuous.

Problem 2-33. Show that the continuity of $D_1 f^i$ at a may be eliminated from the hypothesis of Theorem 2-8.

> If $f\colon \mathbb{R}^n \to \mathbb{R}^m$, then $Df(a)$ exists if all $D_j f^i(x)$ exist in an open set containing a and if each function $D_j f^i$ is continuous at a.

Solution. We proceed exactly as in the proof given in the book, breaking $f(a+h) - f(a) - \sum_i D_i f(a) h^i$ into n terms. However, we need not apply the Mean Value theorem to control the *first* term, because we already have

$$\lim_{h\to 0} \frac{|f(a^1 + h^1, a^2, \ldots, a^n) - f(a^1, \ldots, a^n) - D_1 f(a) h^1|}{|h|} = 0$$

by definition.

Problem 2-34. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called *homogeneous* of degree m if $f(tx) = t^m f(x)$ for all x. If f is also differentiable, show that

$$\sum_{i=1}^n x^i D_i f(x) = m f(x).$$

Solution. As suggested, set $g(t) = f(tx) = t^m f(x)$, and note that $g'(t) = m t^{m-1} f(x)$. On the other hand, we have

$$g'(t) = f'(tx) \cdot x = \sum_{i=1}^n x^i D_i f(tx).$$

Simply set $t = 1$ to retrieve the desired formula.

Problem 2-35. If $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable and $f(0) = 0$, prove that there exist $g_i\colon \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \sum_{i=1}^n x^i g_i(x).$$

Solution. As suggested, set $h_x(t) = f(tx)$, then $h_x'(t) = f'(tx) \cdot x = \sum_i x^i D_i f(tx)$. Now,

$$\int_0^1 h_x'(t)\, dt = h_x(1) - h_x(0) = f(x) - f(0) = f(x).$$

Simply set

$$g_i(x) = \int_0^1 D_i f(tx)\, dt,$$

whence

$$f(x) = \int_0^1 h_x'(t)\, dt = \sum_{i=1}^n \int_0^1 x^i D_i f(tx)\, dt = \sum_{i=1}^n x^i g_i(x).$$

2.5 Inverse Functions

Problem 2-36. Let $A \subset \mathbb{R}^n$ be an open set and $f\colon A \to \mathbb{R}^n$ a continuously differentiable 1-1 function such that $\det f'(x) \neq 0$ for all x. Show that $f(A)$ is an open set and $f^{-1}\colon f(A) \to A$ is differentiable. Show also that $f(B)$ is open for any open set $B \subset A$.

Solution. Pick $y \in f(A)$, corresponding to which there is one $x \in A$, $f(x) = y$. Now, f is continuously differentiable on A, and $\det(f'(x)) \neq 0$. Thus, the Inverse Function theorem guarantees the existence of open set $V \subset A$, $W \subset f(A)$ where $x \in V$, $y \in W$, such that the restriction $f\colon V \to W$ has a continuous inverse. Thus, $f(A)$ is indeed open. Indeed, the same process can be applied to any open set $B \subset A$, by considering the restriction of f to B and proceeding exactly as before; thus $f(B)$ is also open.

Now, we have been given differentiable $f^{-1}\colon W \to V$, hence the inverse function is differentiable at every $y \in f(A)$. (Note that the inverse function is of course unique since f is 1-1, thus the seemingly different inverses on the open sets W corresponding to each choice of y are in fact just restrictions of the inverse $f^{-1}\colon f(A) \to A$).

Problem 2-37.

(a) Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function. Show that f is *not* 1-1.

(b) Generalize this result to the case of a continuously differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$.

Solution.

(a) As suggested, suppose that f is 1-1, hence we cannot have $Df = 0$ identically on any open set. In other words, we cannot have $D_1 f = 0$ or $D_2 f = 0$ identically on any open set. Pick some (a, b) such that $D_1 f(a, b) \neq 0$; since f is continuously differentiable, $D_1 f$ is also continuous, hence there is an open set A around (a, b) where $D_1 f(x, y) \neq 0$. Define $g(x, y) = (f(x, y), y)$ on A, and note that

$$g'(x, y) = \begin{pmatrix} D_1 f(x, y) & D_2 f(x, y) \\ 0 & 1 \end{pmatrix}, \qquad \det g'(x, y) = D_1 f(x, y) \neq 0.$$

Furthermore, g is continuously differentiable on A, and g is 1-1 because if $g(x_1, y_1) = g(x_2, y_2)$, we must have $f(x_1, y_1) = f(x_2, y_2)$ hence $(x_1, y_1) = (x_2, y_2)$ from the 1-1 nature of f. Thus, the previous exercise guarantees that $g(A)$ is open. Now, examine $g(a, b) = (f(a, b), b)$; pick an open rectangle around this point contained in $g(A)$, and pick $b' \neq b$ from the second coordinate. Thus, the point $(f(a, b), b') \in g(A)$. However, this is impossible since

$$g(x, y) = (f(a, b), b') \implies y = b = b'.$$

Thus, f cannot be 1-1.

(b) We proceed as before: suppose that f is 1-1, hence all its partial derivatives cannot be zero everywhere. Pick $a \in \mathbb{R}^n$ such that $D_1 f^1(a) \neq 0$ (WLOG) in a neighbourhood A around a. Then, define $g\colon A \to \mathbb{R}^n$,

$$g(x) = (f^1(x), x^2, \ldots, x^n).$$

Like before, we can show that $\det g'(x) = D_1 f^1(x) \neq 0$ on A, g is continuously differentiable on A, and g is 1-1 on A. Like before, $g(A)$ must be open, hence we ought to be able to pick $(g(a), a') \in g(A)$ where $a' \neq (a^2, \ldots, a^n)$, but this would contradict the fact that g is 1-1.

Problem 2-38.

(a) If $f\colon \mathbb{R} \to \mathbb{R}$ satisfies $f'(a) \neq 0$ for all $a \in \mathbb{R}$, show that f is 1-1 (on all of $\mathbb{R}$).

(b) Define $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x, y) = (e^x \cos y, e^x \sin y)$. Show that $\det f'(x, y) \neq 0$ for all (x, y) but f is not $1 - 1$.

Solution.

(a) Suppose that f is not 1-1, i.e. $f(a) = f(b)$ for some $a \neq b$. Then, we can apply the Mean Value theorem to find c between a and b such that $0 = f(b) - f(a) = f'(c)(b - a)$. This forces $f'(c) = 0$, a contradiction.

(b) Note that this corresponds to the complex function $z \mapsto \exp z$, which is its own derivative but is in fact periodic, with period $2\pi i$.

We simply compute

$$f'(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}, \qquad \det f'(x, y) = e^{2x} \cos^2 y + e^{2x} \sin^2 y = e^{2x} > 0.$$

On the other hand, $f(0, 0) = f(0, 2\pi) = (1, 0)$.

Problem 2-39. Use the function $f \colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x/2 + x^2 \sin(1/x), & \text{if } x \neq 0. \\ 0, & \text{if } x = 0, \end{cases}$$

to show that the continuity of the derivative cannot be eliminated from the hypothesis of Theorem 2-11 (the Inverse Function theorem).

Solution. We show that f satisfies the remaining hypotheses: using our previous work on $x^2 \sin(1/x)$, we directly compute the derivative

$$f'(x) = \begin{cases} 1/2 + 2x \sin(1/x) - \cos(1/x), & \text{if } x \neq 0, \\ 1/2, & \text{if } x = 0. \end{cases}$$

Note that this is *not* continuous at 0: recall that $\cos(1/x)$ is not continuous there. However, we do have $f'(0) = 1/2 \neq 0$.

We can check that f is not invertible on any neighbourhood of 0: note that any neighbourhood of 0 will always contain points of the form $2/2n\pi$ for sufficiently high n, and thus all points between that and 0. Now,

$$f\left(\frac{1}{2n\pi - \pi/2}\right) = \frac{1}{4n\pi - \pi} - \frac{1}{(2n\pi - \pi/2)^2},$$
$$f\left(\frac{1}{2n\pi + \pi/2}\right) = \frac{1}{4n\pi + \pi} + \frac{1}{(2n\pi + \pi/2)^2}.$$

Call these values x and y, so that $x > y > 0$. Now, calculate

$$f(x) - f(y) = \frac{1}{2}(x - y) - (x^2 + y^2) = \frac{1}{2}(x - y) - (x - y)^2 - 2xy.$$

Calculate

$$x - y = \frac{\pi}{4n^2\pi^2 - \pi^2/4}, \qquad xy = \frac{1}{4n^2\pi^2 - \pi^2/4}.$$

Set $a = xy = 1/(4n^2\pi^2 - \pi^2/4)$. Then, $x - y = \pi a$, so

$$f(x) - f(y) = \frac{1}{2}\pi a - \pi^2 a^2 - 2a = \left(\frac{\pi}{2} - 2\right) a - \pi^2 a^2 < 0.$$

This means that $f(x) < f(y)$. On the other hand, $f(x) > 0$ because $f(x) = x/2 - x^2 = x(1/2 - x) > 0$ when $0 < x < 1/2$.

Now, $0 < f(x) < f(y)$, so the continuity of f and the Intermediate Value theorem guarantee that $f(z) = f(x)$ for some $0 < z < y$. Thus, we have found $z < x$, $f(z) = f(x)$ which means that f cannot be invertible in any neighbourhood of 0.

2.6 Implicit Functions

Problem 2-40. Use the implicit function theorem to redo Problem 2-15(c).

Solution. Set $f\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$,

$$f^i(t,x) = \sum_{j=1}^{n} a_{ji}(t) \cdot x^j - b_i(t),$$

in other words,

$$f(t,x) = A(t) \cdot x - b(t).$$

We have been given $\det(a_{ij}(t)) \neq 0$, and this is precisely the determinant of the last n columns of the Jacobian of f, i.e. the matrix $[D_{1+j}f^i]$ for $1 \leq i,j \leq n$. We also know that $f(t,s(t)) = 0$, i.e. setting $x = s(t)$ gives a solution for any t. Assuming that f is *continuously* differentiable, we are guaranteed the uniqueness of the solution x corresponding to some given t, hence the uniqueness of $s\colon \mathbb{R} \to \mathbb{R}^n$, as well as its differentiability by the implicit function theorem.