Several Variable Differential Calculus

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

1 Definitions and Examples

The idea behind differential calculus is to approximate the functions $f\colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ at $x \in U$ by an affine function of the type $\varphi(y) = A(y-x) + f(x)$. The derivative of f at x is to be thought of as the first order linear approximation of the given function in a neighbourhood of x. The first few exercises will make these vague ideas clear.

Definition 1. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be a map with U open. We say f is *differentiable* at $x \in U$ if there exists a linear map $A\colon \mathbb{R}^m \to \mathbb{R}^n$ such that for any $\varepsilon > 0$, there is a $\delta > 0$ with the property that if $\|h\| < \delta$, then

$$\| f(x+h) - f(x) - Ah \| \le \varepsilon \|h\| .$$

Ex. 2. Prove that if f is differentiable at x, then such an A as above is unique. *Hint:* Let A and B both do the job, consider for any unit vector v consider

$$\| A(tv) - B(tv) \| \le \| f(x+tv) - f(x) - A(tv) \| + \| f(x+tv) - f(x) - B(tv) \| \le 2\varepsilon |t|$$

for all t with $|t| < \varepsilon$.

Remark 3. Notice that we make crucial use of the fact that U is open in proving Ex. 2.

This A is called the (Frechet or total) *derivative* of f at x. It is denoted by $Df(x)$.

Ex. 4. Compute the derivative of the constant map $f(x) = c$, c a fixed vector in $\mathbb{R}^n$.

Ex. 5. Compute the derivative of f where f is the restriction to an open set U of a linear map $A\colon \mathbb{R}^m \to \mathbb{R}^n$.

Ex. 6. In the notation of Def. 1, prove that f is differentiable at x iff there exists a linear map $A\colon \mathbb{R}^m \to \mathbb{R}^n$ such that

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

The meaning of Ex. 6 is that if we have a candidate A for the derivative $Df(x)$ an consider the "error term" $f(x+h) - f(x) - Ah$ then this error term goes to zero in $\mathbb{R}^n$ much faster than h as $h \to 0$ in $\mathbb{R}^m$. One usually uses Landau's notation to formulate this more precisely. See at the end of this section.

Ex. 7. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^n$ for $n \in \mathbb{N}$. Find $Df(x)$. *Hint:* Consider

$$\begin{aligned} f(x+h) - f(x) &= (x+h)^n - x^n \\ &= nxh + \sum_{k=1}^{n} \binom{n}{k} x^{n-k} h^k \\ &= nxh + h \cdot \text{ a bounded function} \end{aligned}$$

for $|h| \leq 1$. Thus $Df(x)h = nxh$.

Ex. 8. Compute the derivative $Df(x)$ where $f\colon \mathbb{R} \to \mathbb{R}$ is given by $f(x) = e^x$. *Hint:* Use the series definition of e^x and the fact that $e^{x+y} = e^x e^y$.

More generally, solve:

Ex. 9. Let $f\colon (a,b) \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Then f is differentiable with respect to the above definitions iff f is differentiable in the calculus sense (that is, $f'(x)$ exists) and we have

$$Df(x)h = f'(x)h,$$

where $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ is the "usual" derivative.

Ex. 10. Let $f\colon M(n,\mathbb{R}) \to M(n,\mathbb{R})$ be given by $f(X) = X^2$. Compute $Df(X)H$ *Hint:* Expand $f(X+H) - f(X)$ and collect terms "linear in H".

Ex. 11. Do the same as above for $f(X) = X^k$ for $k \in \mathbb{N}$.

Ex. 12. Let $f\colon M(n,\mathbb{R}) \times M(n,\mathbb{R})$ be given by $f(X,Y) = XY$, the matrix product. Compute the derivative $Df(A,B)$.

Ex. 13. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \langle x, x\rangle$. Compute the derivative $Df(x)$. [Hint: Expand $f(x+h) - f(x)$ and collect the "linear terms" in h.]

Ex. 14. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \langle Ax, x\rangle$, where $A\colon \mathbb{R}^n \to \mathbb{R}$ is a linear map. Compute $Df(x)$.

The essence of following proposition is that the concept of differentiablity of f at x and the derivative $Df(x)$ remain the same even if we change the norms on the domain and/or on the range.

Proposition 15. *Let the notation be as in Def. 1. Let* $|\cdot|$ *be any norms on* $\mathbb{R}^m$ *and* $\mathbb{R}^n$*. Then* $f\colon U \subset (\mathbb{R}^m, \|\,\|) \to (\mathbb{R}^n, \|\,\|)$ *is differentiable at* x *with derivative* A *iff* $f\colon U \subset (\mathbb{R}^m, ||) \to (\mathbb{R}^n, ||)$ *with derivative* B *and* $A = B$*.*

Proof. We are supposed to show that

$$\lim_{\|h\|\to 0} \frac{\|f(x+h) - f(x) - Ah\|}{\|h\|} = 0 \quad \text{iff} \quad \lim_{|h|\to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0.$$

Since any norm on $\mathbb{R}^k$ is equivalent to the euclidean norm there exist constant C_i and C_i', $(i = 1,2)$, such that the following hold:

$$\begin{aligned} C_1|x| &\leq \|x\| \leq && C_2|x| \text{ for all } x \in \mathbb{R}^m \\ C_1'|y| &\leq \|y\| \leq && C_2|y| \text{ for all } y \in \mathbb{R}^n. \end{aligned}$$

The result follows from the following observation:

$$\frac{C_1'|f(x+h)-f(x)-Ah|}{C_2|h|} \leq \frac{\|f(x+h)-f(x)-Ah\|}{\|h\|} \leq \frac{C_2'|f(x+h)-f(x)-Ah|}{C_1|h|}.$$

□

Ex. 16. Let $f\colon \mathbb{R}^k \to \mathbb{R}^m \times \mathbb{R}^n$ be given. Write $f(x) = (f_1(x), f_2(x)) \in \mathbb{R}^m \times \mathbb{R}^n$ in an obvious notation. Prove that f is differentiable at x iff f_i are differentiable at x and we have

$$Df(x)h = \begin{pmatrix} Df_1(x)h \\ Df_2(x)h \end{pmatrix}.$$

In particular, $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is differentiable iff each f_i is differentiable and

$$Df(x)h = \begin{pmatrix} Df_1(x)h \\ \vdots \\ Df_n(x)h \end{pmatrix}.$$

Here $f = (f_1, \ldots, f_n)$.

Ex. 17. Compute the derivative of $f(x,y) = \langle Ax, By\rangle$ where $A\colon \mathbb{R}^m \to \mathbb{R}^k$ and $B\colon \mathbb{R}^n \to \mathbb{R}^k$ are linear.

Ex. 18. Let $B\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ be a bilinear map, that is, $x \mapsto B(x, y_0$ and $y \mapsto B(x_0, y)$ are linear on $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively. Compute $DB(x,y)(h,k)$. *Hint:* $B(x+h, y+k) - B(x,y) = B(x,k) + B(h,y) + B(h,k)$. The first two terms are "linear" in h, k and so we define $DB(x,y)(h,k) := B(x,k) + B(h,y)$. To show $B(h,k)$ goes to zero much faster than (h,k) as $(h,k) \to 0$, write B in terms of a basis.

Ex. 19. Extend the last exercise to multi-linear maps: $f\colon \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k} \to \mathbb{R}^m$.

Ex. 20. Compute the derivative of $f\colon M(n,\mathbb{R}) \to \mathbb{R}$, where $f(X) = \det X$ at the identity I. *Hint:* Think of det as a multi-linear map on the space of column vectors of $X \in M(n,\mathbb{R})$.

Ex. 21. Let $f, g\colon \mathbb{R}^m \to \mathbb{R}$ be differentiable at $x \in \mathbb{R}^m$. Show that $\phi(y) = f(y)g(y)$ is differentiable at $y = x$.

Ex. 22. If $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}$ is differentiable at x, then f is continuous at x. More precisely, show that f is locally Lipschitz at x. That is, prove that there exists a constant $c > 0$ and a $\delta > 0$ such that if $\|y - x\| < \delta$, then $\|f(y) - f(x)\| < c\,\|x - y\|$.

Ex. 23. 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.

Three Very Special Cases

There are three very special cases of the set-up $f\colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ which one should be thoroughly familiar with. They are

1. $m = 1 = n$. This case has been successfully dealt with in Ex. 9. In this case the Frechet derivative $Df(x)$ and the classical (calculus) derivative are related by $Df(x)h = f'(x)h$.

2. $m = 1$ and n arbitrary. In this case we think of f as a differentiable curve in $\mathbb{R}^n$ joining the points $f(s)$ and $f(t)$ for $s, t \in U$. As a mnemonic we use the letter c in place of f. In this case there is a notion of tangent vector (or the velocity vector) $c'(t)$ at the point t to the curve c given by $c'(t) := (c_1'(t), \ldots, c_n'(t))$. Note that

$$c'(t) = \lim_{h \to 0} \frac{c(t+h) - c(t)}{h}.$$

How are $Dc(t)$ and $c'(t)$ related? We leave it to the reader to show $c'(t) = Dc(t)(1)$.

3. $n = 1$ and m arbitrary. In this case the map $Df(x)$ is a linear map from $\mathbb{R}^m \to \mathbb{R}$. We know any such linear map $\varphi \colon \mathbb{R}^m \to \mathbb{R}$ arises as the inner product with a fixed vector $u := (\varphi(e_1), \ldots, vfi(e_m))$: $\varphi(x) = \langle x, v \rangle$. Thus Df is represented by a unique vector, denoted by $\nabla f(x)$ or $\operatorname{grad} f(x)$, so that $Df(x)(h) = \langle h, \nabla f(x) \rangle$.

It is very essential that the reader understands these three special cases very well as they will be repeatedly used in the sequel.

2 Directional and Partial Derivatives

The single most important trick in calculus of several variables is to reduce the problems to one-variable setup.

Definition 24. Fix any $v \in \mathbb{R}^m$, $x \in U$. We claim that there exists $\varepsilon > 0$ such that $x + tv \in U$ for $|t| < \varepsilon$. For, since U is open there exists $r > 0$ such that $B(x, r) \subset U$. If we take $\varepsilon := r / \|v\|$ (which is ∞ if $v = 0$) then the claim follows. Let $f \colon U \subseteq \mathbb{R}^m \to \mathbb{R}$ be a map. Consider the one-variable function $g(t) := f(x + tv)$ for $t \in (-\varepsilon, \varepsilon)$. The basic trick in several variable calculus is to reduce the problem to this function whenever feasible. Note that $g(0) = f(x)$. We may ask whether this function g is differentiable at 0. That is, whether the limit

$$\lim_{t \to 0} \frac{f(0+t) - g(0)}{t}$$

exists. This is the same as requiring that $\lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$ exists. If the limit exists, we denote it by $D_v f(x)$ and call it the *directional derivative* of f at x in the *direction* of v. A particular choice of v is any standard basis vector e_i of $\mathbb{R}^m$. In this case $D_{e_i} f(x)$ is usually denoted by $\frac{\partial f}{\partial x_i}(x)$ or $D_i f(x)$ and called the *i-th partial derivative* of f at x.

The geometric meaning behind this definition is as follows. Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be a function with directional directive at $D_v f(x)$ at x. The geometric interpretation of $D_v f(x)$ is that is the slope of the tangent line at $(x, f(x))$ to the curve formed by the intersection of the graph of f with the plane that contains x and $x + v$ and parallel to the z-axis.

Proposition 25. *Let $f \colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ be differentiable at x. Then f has directional derivatives at x in all directions and $D_v f(x)$ is given by*

$$D_v f(x) = Df(x)(v). \tag{1}$$

Proof. □

Ex. 26. The converse of the above proposition is not true in general. Consider $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & \text{otherwise.} \end{cases}$$

Then all its directional derivatives at $(0,0)$ exist. However, f is not even continuous at $(0,0)$ (and hence is certainly not differentiable at $(0,0)$). [Hint: Approach $(0,0)$ along the parabola $y = x^2$.]

Ex. 27. Find the directional derivative $D_v f(x)$ of the functions as indicated.
(i) $f(x,y,z) = xyz$ where $x = (1,0,0)$ and $v = (\cos\alpha \sin\beta, \sin\alpha \sin\beta, \cos\beta)$.
(ii) $f(x,y) = e^x \sin y$ where $x = (1,0)$ and $v = (\cos\alpha, \sin\alpha)$.

Ex. 28. Prove the **Chain Rule:** Let $f :: U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and $g : V \subseteq \mathbb{R}^n \to \mathbb{R}^k$ be differentiable at $x \in U$ and $y = f(x) \in V$. Prove that $\phi = g \circ f : U \to \mathbb{R}^k$ is differentiable and $D\phi(x) = Dg(y) \circ Df(x)$. [Hint: Let $A = Df(x)$, $B = Dg(y)$.

$$\begin{aligned} \| g \circ f(x+h) - g \circ f(x) - B \circ Ah \| \\ \leq \| g(f(x+h)) - g(f(x)) - B(f(x+h) - f(x)) \| \\ + \| B[f(x+h) - f(x) - Ah] \| . \end{aligned}$$

You need Ex. **??** and Ex. 22 and the fact that $\|Tv\| \leq \|T\| \cdot \|v\|$ for a linear map T.

Ex. 29. Use the chain rule to prove the differentiability of $x \mapsto \|x\|$ on $\mathbb{R}^n \setminus \{0\}$.

Ex. 30. Let $f, g\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f\colon (x,y) \mapsto x+y$ and $g\colon (x,y) \mapsto xy$. Compute the derivatives of f and g.

Ex. 31. Do Ex. 21 using Ex. 30 and the chain rule.

Ex. 32. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}$ be differentiable at x. Then $Df(x)\colon \mathbb{R}^m \to \mathbb{R}$ is a linear map. Hence by Riesz representation theorem, there exists a unique vector $u \in \mathbb{R}^m$ such that $Df(x)v = \langle v, u \rangle$ for all $v \in \mathbb{R}^m$. This u has coordinates $\phi(e_i)$:

$$u = \sum_{i=1}^{m} \phi(e_i)e_i = \begin{pmatrix} \phi(e_1) \\ \vdots \\ \phi(e_m) \end{pmatrix} \in \mathbb{R}^m.$$

Prove that this vector, generally denoted by $\operatorname{grad} f(x)$ and called the *gradient* of $f$at x, is given by

$$\text{grad } f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_m}(x). \end{pmatrix}$$

Hence $Df(x)(v) = \sum_{i=1}^m v_i \frac{\partial f}{\partial x_i}(x) = \langle v, \operatorname{grad} f(x) \rangle$.

Ex. 33 (An interpretation of Ex. 32)**.** Since $Df(x)$ is linear, we can write it as a $(1 \times m)$-matrix with respect to the standard basis of $\mathbb{R}^m$. Thus $Df(x) = (\frac{\partial f_1}{\partial x_1}(x), \dots, \frac{\partial f_m}{\partial x_m}(x))$ (as matrices). This matrix is called the Jacobian of f at x. We have

$$Df(x)h = (\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_m}(x)) \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x) h_i.$$

Ex. 34. Find the gradient of each of the following functions at the indicated points:
(i) $f(x) = \|x\|^2$ for an arbitrary $x \in \mathbb{R}^n$.
(ii) $f(x) = \|x\|^\alpha$ for $0 \neq x \in \mathbb{R}^n$.
(iii) $f(x, y) = x + y + z$ at $x = (1, 2, 3)$.

Ex. 35. We generalise the previous exercise: Let $f = (f_1, \dots, f_n)\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be differentiable at x. Show that $Df(x)$ has the following matrix representation with respect to the standard basis of $\mathbb{R}^m$ and $\mathbb{R}^n$:

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \cdots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Hint: Either imitate the proof of the earlier two exercises or recall that

$$Df(x)(h) = (Df_1(x)h, \dots, Df_n(x)h).$$

This matrix is known as the *Jacobian matrix* of f at x.

Ex. 36. The matrix form of the chain rule is as follows: Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be given by

$$\begin{aligned} f(x_1, \dots, x_n) &= (y_1, \dots, y_n) = (y_1(x_1, \dots, x_m), \dots, y_n(x_1, \dots, x_m)) \text{ and} \\ g(y_1, \dots, y_n) &= (z_1(y_1, \dots, y_n), \dots, z_k(y_1, \dots, y_n)). \end{aligned}$$

Then the Jacobian matrix $J(g \circ f)$ of $D(g \circ f)$ is $J(g) \circ J(f)$ and symbolically written as

$$\frac{\partial z_i}{\partial x_j} = \sum_r \frac{\partial z_i}{\partial y_r} \cdot \frac{\partial y_r}{\partial x_j}.$$

Ex. 37. Consider $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$f(\begin{pmatrix} u \\ v \end{pmatrix}) = \begin{pmatrix} u + v \\ u - v \\ u^2 - v^2 \end{pmatrix}$$

and $g\colon \mathbb{R}^3 \to \mathbb{R}$ given by $g(x, y, z) = x^2 + y^2 + z^2$. Find the Jacobian matrix of $D(g \circ f)$ at $\begin{pmatrix} a \\ b \end{pmatrix}$.

Ex. 38. Let $f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r \end{pmatrix}$ and $w = g(x, y, z) = \sqrt{x^2 + y^2 + z^2}$. Find $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial \theta}$ using the chain rule. Check the result by direct substitution.

Ex. 39. Let $f, g : (a,b) \to \mathbb{R}^n$ be differentiable. Let $\phi(t) := \langle f(t), g(t)\rangle$. Compute $\phi'(t)$. (Do you recognize this as a special case of something we did earlier on?)

Ex. 40. Let $f\colon \mathbb{R}^m \to \mathbb{R}^k$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ be differentiable. Let $\phi(x,y) := \langle f(x), g(y)\rangle$. Show that ϕ is differentiable on $\mathbb{R}^m \times \mathbb{R}^n$.

Ex. 41. Let $c\colon (a,b) \to \mathbb{R}^n$ be differentiable. We think of c as a curve in $\mathbb{R}^n$. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be differentiable. Prove that $g(t) := f \circ c(t)$ is differentiable and $g'(t) = \langle \text{grad } f(c(t)), c'(t)\rangle$. Here $c'(t) = \begin{pmatrix} c_1'(t) \\ \vdots \\ c_n'(t) \end{pmatrix} = Dc(t)(1)$ is the tangent vector to c at t. Note that $g'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(c(t)) \cdot c_i'(t)$.

Ex. 42. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is said to be *homogeneous of degree k* if $f(tx) = t^k x$ for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Let f be homogeneous of degree k and differentiable on $\mathbb{R}^n$. Show that

$$D_x f(x) = \langle x, \text{grad } f(x)\rangle = \sum x_i \frac{\partial f}{\partial x_i}(x) = k f(x).$$

This is known as Euler's theorem. Prove also the converse. *Hint for both:* Consider $g(t) = f(tx)$ for the first part and $t^{-k} g(t)$ for the converse.

Ex. 43. Find the derivatives of the following functions:
(1) $f(x,y) = x^y$.
(2) $f(x,y) = \sin(xy)$.
(3) $f(x,y) = \int_a^{x+y} g$.
(4) $f(x,y) = \int_a^{xy} g$.
(5) $f(x,y) = \int_x^y g$.

In (3) to (5), assume that $g\colon \mathbb{R} \to \mathbb{R}$ is continuous.

Ex. 44. Compute the Jacobian matrix of the following functions:
(1) $(x,y) \mapsto (e^x \cos y, e^x \sin y)$.
(2) $(x,y) \mapsto (x+y, xy, x-y)$.
(3) $x \in \mathbb{R}^n \mapsto \langle Ax, x\rangle$ where $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is linear. Compare the result with Ex. 14.

Ex. 45. Let $c\colon (a,b) \to \mathbb{R}^n$ be differentiable such that $\| c(t) \| = 1$ for $t \in (a,b)$. Prove that $c'(t)$ is perpendicular to $c(t)$ for $t \in (a,b)$. Interpret this result geometrically in terms of spheres and tangent planes.

Ex. 46. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be differentiable. Let 0 be a value of f so that $f^{-1}(0)$ is non-empty. Let $c\colon (a,b) \to \mathbb{R}^n$ be a differentiable curve such that $c(t) \in f^{-1}(0)$ for all $t \in (a,b)$. Show that $\langle c'(t), \text{grad}\, f(c(t))\rangle = 0$. Specialize to $f\colon \mathbb{R}^3 \to \mathbb{R}$ and understand the geometry behind this exercise.

3 Mean Value Theorem

Ex. 47. Let $f\colon \mathbb{R} \to \mathbb{R}^2$ be defined by $f(t) = (\cos t, \sin t)$. Prove that the mean value theorem in the form $f(y) - f(x) = Df(z)(y-x)$ for some $z \in (x,y)$ **cannot** hold when $x = 0$ and $y = 2\pi$.

Ex. 48. Prove the mean value theorem in the following form: Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be differentiable. Let $x, y \in U$ be such that the line joining x, y is contained in U. That is, $x + t(y - x) \in U$ for $t \in [0, 1]$. Hence observe that $x + t(y - x) \in U$ for $(-\varepsilon, 1 + \varepsilon)$ for some $\varepsilon > 0$. Let $a \in \mathbb{R}^n$ be fixed. Show that there exists $t \in [0, 1)$ such that $\langle f(y) - f(x), a\rangle = \langle Df(x + t(y - x))(y - x), a\rangle$. *Hint:* Define $g(t) := \langle a, f(x + t(y - x))\rangle$. Apply the mean value of one-variable calculus.

Ex. 49. Let U be open and convex in $\mathbb{R}^n$. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be differentiable. Assume that $\sup_{z\in U} \|Df(z)\| < c < \infty$ for some $c > 0$. Then show that f is Lipschitz on U. *Hint:* For $x \in \mathbb{R}^n$, $\|x\| = \sup_{\|a\|=1} |\langle x, a\rangle|$.

Ex. 50. An important special case of Ex. 49 is when f continuously differentiable. That is, when $x \mapsto Df(x)$ as a map from U into $L(\mathbb{R}^m, \mathbb{R}^n) \cong M_{n\times m}(\mathbb{R})$ is continuous. Let K be a compact convex set in U. Then there exists a $c > 0$ such that $\|f(x) - f(y)\| \le c\|x - y\|$ for $x, y \in K$.

Ex. 51. Let U be open and $f\colon U \to \mathbb{R}^n$ be differentiable. Let $Df(x) = 0$ for all $x \in U$. Show that f is locally constant on U. hence, if we further assume that U is connected, conclude that f is a constant.

Ex. 52. Give an example to show that connectedness of U is necessary in Ex. 51.

Ex. 53. Compute the Jacobian matrix of the map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(r, \theta) = (r\cos\theta, r\sin\theta)$.

Ex. 54. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be continuously differentiable. Let $x \in U$ be such that $Df(x)$ is one-one. Show that there exists a neighborhood of x in U on which f is one-one. What is the significance of Ex. 53 for the present exercise?

Ex. 55. Let $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be such that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are continuous on U. Show that $Df(x)$ exists on U and that $x \mapsto Df(x)$ is continuous. Thus f is continuously differentiable. *Hint:* Estimate

$$\begin{aligned}
&\left\| f(x+h, y+k) - f(x,y) - \frac{\partial f}{\partial x}h - \frac{\partial f}{\partial y}k \right\| \\
&= \left\| f(x+h, y+k) - f(x, y+k) - \frac{\partial f}{\partial x}(x, y+k) \right\| \\
&\quad + \left\| \frac{\partial f}{\partial x}(x, y+k) - \frac{\partial f}{\partial x}(x,y) \right\| + \text{ other terms.}
\end{aligned}$$

Ex. 56. This is extension of Ex. 55 and included here for records' sake. We say $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^1$ if all its partial derivatives exists and are continuous. Then f is continuously differentiable on U.

Ex. 57. This is the most general form of Ex. 55 and Ex. 56. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be such that $\frac{\partial f_i}{\partial x_j}$ exist and are continuous on U for $1 \le i \le n$ and $1 \le j \le m$. Then f is continuously differentiable on U.

Ex. 58. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be differentiable. Let $p \in U$ and $c\colon (-\varepsilon, \varepsilon) \to U$ be a differentiable curve such that $c(0) = p$ and $c'(0) = v$. (One such curve is $c(t) = p + tv$.) Show that if $g(t) := f \circ c(t)$, then

$$g'(0) = \frac{d}{dt} f \circ c(t)|_{t=0} = D_v f(p) = Df(p)v.$$

The moral of this exercise is that to compute $Df(p)$ it is enough to know $Df(p)v$ for all $v \in \mathbb{R}^m$ and to know that we can use any curve c with initial data $c(0) = p$ and $c'(0) = v$.

Ex. 59. Let $f\colon M(n,\mathbb{R}) \times M(n,\mathbb{R}) \to M(n,\mathbb{R})$ given by $f(A,B) = AB$. Find $Df(A,B)$.

Ex. 60. Let $f\colon \mathbb{R} \to M(n,\mathbb{R})$ be given by $f(t) = e^{tA}$ for a fixed $A \in M(n,\mathbb{R})$. Find $Df(t)$.

Ex. 61. To illustrate the use of Ex. 58, find the derivative of $f\colon GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ given by $f(X) = X^{-1}$ in three ways:
(1) Use "binomial" type expansion for $f(X+H) = (X+H)^{-1}$.
(2) Use the chain rule for the map $X \mapsto (X, X^{-1}) \to I$.
(3) Use $Df(A)H = D_H f(A) = \frac{d}{dt}(Ae^{tA^{-1}H})$ as $c(t) = Ae^{tA^{-1}H}$ has the required initial data $c(0) = A$ and $c'(0) = H$.

Ex. 62. Given $F\colon \mathbb{R}^n \to \mathbb{R}^n$, the problem of finding a function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that grad $f = F$ is equivalent to solving the following system of equations for f:

$$\frac{\partial f}{\partial x_1} = F_1, \ldots, \frac{\partial f}{\partial x_n} = F_n.$$

(i) For $n = 2$, this system has a solution f, then f must have both of the forms:

$$\begin{aligned} f(x,y) &= \int F_1(x,y)dx + c_1(y) \\ f(x,y) &= \int F_2(x,y)dy + c_2(x). \end{aligned}$$

(ii) Find f, if grad $f(x,y) = (x^2 + 2xy, 2xy + x^2)$.

Ex. 63. Let $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$. Assume that $D_1 f$, $D_2 f$, $D_1 D_2 f$ and $D_2 D_1 f$ exist and are continuous. Then $D_1 D_2 f = D_2 D_1 f$. Here $D_i f = D_{e_i} f$ are the partial derivatives. *Hint:* Consider $g_1(x) = f(x, y+k) - f(x,y)$ and $g_2(y) = f(x+h, y) - f(x,y)$. Apply mean value theorem to $g_1(x+h) - g_1(x)$ and $g_2(y+k) - g_2(y)$ and use continuity of $D_1 D_2 f$ and $D_2 D_1 f$.

Ex. 64. (i) If $f(x,y) = \log(x^2+y^2)$, show that $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$.
(ii) Let $f(x) = \frac{1}{\|x\|}$ on $\mathbb{R}^3 \setminus \{0\}$. Show that $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial z^2} = 0$.
(iii) More generally, if $f(x) = \|x\|^{(n-2)/n}$, then $\sum \frac{\partial^2 f}{\partial x_i^2} = 0$, for $n \geq 3$.

Ex. 65. Let

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

Compute $D_1 f$, $D_2 f$, $D_1 D_2 f$, $D_2 D_1 f$ at $(0,0)$.

Ex. 66. Find the partial derivatives of the following functions: (i) $f(x,y) = e^{xy}$. (ii) $f(x,y) = \int_x^{x+y} g(t)dt$, $g : \mathbb{R} \to \mathbb{R}$ continuous. (iii) $f(x,y) = \int_a^{xy} g(t)dt$, $g : \mathbb{R} \to \mathbb{R}$ continuous. (iv) $f(x,y) = f_1(x) f_2(y)$, f_i differentiable. (v) $f(x,y) = g(xy)$.

Ex. 67. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given. If $D_1 f = 0$, show that f is independent of the first variable. If $D_1 f = 0 = D_2 f$, show that f is a constant.

Ex. 68. We say that $f\colon U\subseteq\mathbb{R}^m\to\mathbb{R}$ is $\mathcal{C}^k$ for $k\in\mathbb{N}$ if all its partial derivatives of order less than or equal to k exist and are continuous. Thus f is $\mathcal{C}^2$ on $\mathbb{R}^2$ if D_1f, D_2f, D_1D_2f, D_2D_1f, D_1^2f, D_2^2f all exist and are continuous. It follows from earlier exercises that if $f\colon\mathbb{R}^m\to\mathbb{R}$ is $\mathcal{C}^k$, then "all the mixed partial derivatives of same type" are the same. (This exercise is more for the record than for solving!)

Definition 69. $f\colon U\subseteq\mathbb{R}^m\to\mathbb{R}^n$ is $\mathcal{C}^k$ if each coordinate function f_i is $\mathcal{C}^k$ for $1\le i\le n$.

Ex. 70. Let $f\colon U\subseteq\mathbb{R}^m\to\mathbb{R}$ be $\mathcal{C}^k$ and let $x\in U$, $v\in\mathbb{R}^m$ and $c(t):=x+tv$. Show that if g is given by $g(t)=f\circ c(t)$, then g is defined in an interval around 0 in $\mathbb{R}$ and that it is $\mathcal{C}^k$. Compute $g'(t)$, $g''(t)$ and more generally, $g^{(r)}(t)$ for $0\le r\le k$.

4 Taylor's Theorem

Ex. 71. Let $f\colon[a,b]\to\mathbb{R}$ be such that $f^{(n+1)}$ exists and is continuous on $[a,b]$. That is, f is $\mathcal{C}^{(n+1)}$ on $[a,b]$. Then for $x\in[a,b]$, we have

$$f(x)=f(a)+\sum_1^n\frac{f^{(k)}(a)}{k!}(x-a)^k+R_n(x) \tag{2}$$

where $R_n(x)=\frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n dt$. *Hint:* Use induction and integration by parts. Eq. 2 is known as *Taylor's expansion*.

Ex. 72. Write the Taylor expansion for the following functions: (i) $\log x$ at $a=1$. (ii) $\sqrt{1-x}$ at $a=0$. (iii) e^x, $\sin x$, $\cos x$ at $a=0$.

Ex. 73. Let $f\colon[a,b]\to\mathbb{R}$ be continuous and $g:[a,b]\to\mathbb{R}$ be Riemann integrable on $[a,b]$ with $g(x)\ge 0$ for all $x\in[a,b]$. Then there exists a $c\in[a,b]$ such that

$$\int_a^b f(x)g(x)=f(c)\int_a^b g(x)dx.$$

This is known as the *first mean value theorem* of Riemann integration.

Ex. 74. Let the notation be as in Ex. 71. Use the first mean value theorem of Riemann integration to conclude that there exists $c\in[a,x]$ such that

$$f(x)=f(a)+\sum_1^n\frac{f^{(k)}(a)}{k!}(x-a)^k+R_n(x).$$

where

$$R_n(x)=\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Ex. 75. You may also prove Ex. 74 directly as follows: Consider

$$F(t)=f(t)+\sum_1^n\frac{f^{(k)}(t)}{k!}(x-t)^k+M(x-t)^{n+1}$$

where M is chosen so that $F(a)=f(x)$. This is possible for $x\ne a$. Observe that $F(x)=F(a)$ and and apply Rolle's theorem.

Ex. 76. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$, and $x \in \mathbb{R}^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$ and $\alpha! := \alpha_1! \cdots \alpha_n!$ and $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. We let

$$D^\alpha := D_1^\alpha \cdots D_n^{\alpha_n} = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

Given $\alpha_1, \dots, \alpha_n$ with $|\alpha| = N$, show that there are $N!/\alpha!$ different N-tuples $(i_1, \dots, i_N)$ in which 1 occurs α_1-times, 2 occurs α_2 times etc.

Ex. 77. We rewrite the remainder term $R_n(x)$ in Ex. 71. Assume $a = 0$. In the integral, change the variable: $t = ux$ and get

$$R_n(x) = \frac{x^{n+1}}{n!} \int_0^1 f^{(n+1)}(ux)(1-u)^n dt.$$

Ex. 78. Let f be a $\mathcal{C}^{N+1}$ function on an open convex neighborhood of 0 in $\mathbb{R}^n$. Then

$$f(x) = \sum_{|\alpha| \leq N} \frac{1}{k!} D^\alpha f(0) x^\alpha + R_N(x)$$

where

$$R_N(x) = (N+1) \sum_{|\alpha|=N+1} \frac{x^\alpha}{\alpha!} \int_0^1 D^\alpha f(tx)(1-t)^N dt.$$

Hint: Consider $g(t) = f(tx)$ and use Ex. 70, Ex. 77 and Ex. 76.

Ex. 79. Let $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^{N-1}$ (i.e., $N-1$ times continuously differentiable). Let $a \in U$. Show that we can write

$$f(a+h) = f(a) + \sum_{k=1}^N \frac{f^{(k)}(a)}{k!}(h) + \frac{1}{N!} f^{(N+1)}(x)(h)$$

where $f^{(k)}(a)h = \sum_{i_1 \cdots i_k = 1}^n D_{i_1} D_{i_2} \cdots D_{i_k} f(a) h_{i_1} \cdots h_{i_k}$ and $h \in [a, a+h)$, the line joining a and $a+h$. (We assume $[a, a+h] \subseteq U$). *Hint:* Consider $g(t) := f(a+th)$ and apply the one variable Taylor's theorem as in Ex. 74.

Ex. 80. Let f be C^k in a neighbourhood of the origin in $\mathbb{R}^n$ such that $f(0) = 0$. Then there exist C^{k-1} functions f_i for $1 \leq i \leq n$ (in a possibly small neighbourhood) such that $f(x) = \sum_i x_i f_i(x)$.

Ex. 81. Let the notation be as above. Let

$$T_N(x) = f(x) + \sum_{k=1}^N \frac{f^{(k)}(a)}{k} h.$$

We call T_N the *N-th Taylor polynomial* of T at x. It has the property that

$$\lim_{y \to x} \frac{f(y) - T_N(y)}{\|y - x\|^N} = 0.$$

Ex. 82. Show that the N-th Taylor polynomial of a $\mathcal{C}^{N+1}$ function f at a is unique. *Hint:* Note that if T and S are such polynomials then

$$\lim_{y\to x} \frac{Ty - Sy}{\|y - x\|^N} = 0$$

Write $Ty - Sy = P_k(y) + R(y)$, where P_k is the polynomial consisting of terms of lowest degree k that actually occurs in $T - S$. Observe that

$$\lim_{t\to 0} \frac{P_k(ty_0) + R(ty_0)}{|ty_0|^k} = 0$$

for y_0 with $P_k(y_0) \neq 0$.

Ex. 83. Write the polynomial $x^2y + x^3 + y^3$ in powers of $(x - 1)$ and $(y + 1)$.

Ex. 84. Find the Taylor expansion of $f(x) = (\sum x_i)^N$ at $x = 0$.

Ex. 85. Find the best second degree approximation to the function $f(x, y) = xe^y$ at $(2, 0)$.

Ex. 86. Find the Taylor expansion of $f(x, y) = e^{xy}\sin(x + y)$ at $(0, 0)$ (i) by computing derivatives. (ii) by using Taylor expansion of e^{xy} and $\sin(x + y)$.

Ex. 87. Write the second order Taylor expansion of a $\mathcal{C}^2$ function as follows:

$$\begin{aligned} f(a+h) - f(a) &= \langle \text{grad } f(a), h\rangle + \frac{1}{2}D^2 f(x)h \\ &= \langle \text{grad } f(a), h\rangle + \frac{1}{2}D^2 f(a)h + \|h\|^2 E(h) \end{aligned}$$

where

$$\begin{aligned} \|h\|^2 E(h) &= \frac{1}{2}[D^2 f(x)h - D^2 f(a)h] \\ &= \frac{1}{2}\sum_{i,j=1}^{n} [D_iD_j f(x) - D_iD_j f(a)]h_ih_j. \end{aligned}$$

Conclude that $|E(h)| \to 0$ as $\|h\| \to 0$.

5 Maxima and Minima

Definition 88. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ be function. A point $a \in U$ is said to be a *local maximum* if there exists an open set B containing a such that $f(a) \geq f(x)$ for al $x \in B$.

A *local minimum* is similarly defined.

Ex. 89. Let $f\colon U \subseteq \mathbb{R}^m \to \mathbb{R}$ be differentiable and x be a point of local maximum or local minimum. Show that $D_v f(x) = 0$ for all $v \in \mathbb{R}^m$. Hence conclude that grad $f(x) = 0$. *Hint:* Consider $g(t) := f(x + tv)$ and apply the one variable result.

Ex. 90. Let $f: U \subseteq \mathbb{R}^m \to \mathbb{R}$ be differentiable. Show that in the direction v in which f has the maximum absolute value, the directional derivative at x is along grad $f(x)$. Use this to understand the geometry behind Ex. 89.

Ex. 91. Consider $f(x, y) = x^2 + y^2$. Then grad $f(0, 0) = 0$. So there is no indication of a direction of maximum increase of f at $(0, 0)$. Is this reasonable? What happens at $(0, 0)$? Carry out similar exercise when $f(x, y) = xy$ and $f(x, y) = x^2 - y^2$.

Definition 92. e say a point a in the domain of a differentiable (real valued) function is a *critical point* if the gradient of f at a is zero.

Ex. 93. Any point of local maximum or local minimum is a critical point. Give an example of a critical point which is neither a local maximum or a local minimum.

Ex. 94. Find the critical points of the following functions:

(a) $(x + y)e^{-xy}$	(d) $y^2 - x^3$	(g) $x \sin y$
(b) $xy + xz$	(e) $e^{-\lVert x \rVert^2}$	(h) $(x - y)^4$.
(c) $x^2 + y^2 + z^2$	(f) $x^2 y^2$	(i) $x^2 + y^2 + z^2 + xy$.

Ex. 95. Let $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^2$. Assume that $a \in U$ is a critical point of f. Let

$$Q(h) := \frac{1}{2} \left\langle D^2 f(a)h, h \right\rangle = \frac{1}{2} \sum_{i,j} D_i D_j f(a) h_i h_j.$$

(i) If $Q(h) > 0$ for all $h \neq 0$, then f has a local minimum at a. (ii) If $Q(h) < 0$ for all $h \neq 0$, then f has a local maximum at a. (iii) If $Q(h)$ is indefinite, then a is said to be a *saddle point* of f. In a neighborhood of a, we can find points x, y such that $f(x) < f(a) < f(y)$.

Definition 96. The matrix $D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)$ of a $\mathcal{C}^2$ function f is called the *Hessian* of f at a. A symmetric matrix $A = (a_{ij})$ is said to be *positive definite* (*negative definite*) iff $\langle Ax, x \rangle > 0$ (respectively $\langle Ax, x \rangle < 0$) for $x \neq 0$. Thus a critical point a of f is a point of local minimum (maximum) if the Hessian $H_f(a)$ of f at a is positive (respectively negative) definite.

Ex. 97. There are two well-known criteria for the positive definiteness of a symmetric matrix $A = (a_{ij})$. (i) All the eigen values of A are positive. (ii) All the matrices $(a_{ij})_{1 \leq i,j \leq k}$ for $1 \leq k \leq n$ have positive determinants.

Prove the second criterion for $n = 2$ and the first criterion for all n.

Ex. 98. Classify the critical points of Ex. 94 as maximum, minimum, or neither.

Ex. 99. ind and classify the critical points of the following functions: (a) $f(x, y) := x^2 - xy + y^2$. (b) $f(x, y) := x^3 - 3x^2 + y^2$.

Definition 100. critical point of a C^2-function is said to be *nondegenerate* if the Hessian $D^2 f(a) := (D_i D_j f(a))$ of f at a is nonsingular.

Ex. 101. Show that a nondegenerate critical point a of a C^2-function is isolated, that is, a has a open ball U around it in which there are no critical points of f other than a. *Hint:* Let x be another critical point in U. Apply mean value theorem to $D_i f$ to find

$$0 = \sum_j D_j D_i f(y_i)(x_j - a_j), \qquad 1 \leq i \leq n, \tag{3}$$

where $y_j \in U$. Show that if U is small enough $\det(D_i D_j f(y_i)) \neq 0$ and consequently the system of linear equations Eq. 3 has only one solution $x - a = 0$.

6 Smooth Functions with Compact Support

Definition 102. The *support* of a function f on U is defined to be the closure of the set $\{x \in U : f(x) \neq 0\}$.

Ex. 103. Let $f\colon [a,b] \to \mathbb{R}$ be continuous and C^1 on (a,b). Assume that $\lim_{x\to a_+} f'(x) = l$ and $\lim_{x\to b_-} f'(x) = m$ exist. Show that f is C^1 on $[a,b]$.

Ex. 104. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ be continuous on U and C^1 on $U \setminus \{a\}$ for $a \in U$. Assume that $\ell_i := \lim_{x\to a} D_i f(x)$ exists for $1 \leq i \leq n$. Prove that $D_i f(a) = \ell_i$ and that f is C^1 on U.

Ex. 105. Consider $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(t) = \begin{cases} 0 & \text{for } t \leq 0 \\ \exp(-1/t) & \text{for } t > 0 \end{cases}.$$

f is differentiable on $\mathbb{R} \setminus \{0\}$. (a) Observe that $e^x > \frac{x^k}{k!}$ for $k \in \mathbb{N}$. (b) Prove that $f(x) < k! x^k$ for $k \in \mathbb{N}$ and hence conclude that f is continuous at $x = 0$. (c) Prove by induction that $f^{(k)}(x) = p_k(x^{-1}) f(x)$ for some polynomial of degree less than or equal to $k+1$ (for $x \neq 0$). Note that

$$\begin{aligned} |\left[f^{(k)}(x) - f^{(k)}(0)\right] x| &= \| f(x) x^{-1} p_k(x^{-1}) \| \\ &\leq n! x^{n-k} \end{aligned}$$

Conclude that $f^{(k+1)}(0)$ exists and hence f is infinitely differentiable on all of $\mathbb{R}$.

Ex. 106. Carry out a similar analysis to conclude that $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(t) = \begin{cases} e^{-\frac{1}{t^2}} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

is infinitely differentiable.

Ex. 107. Let f be as in Ex. 106. Let $\varepsilon > 0$ be given. Define $g_\varepsilon(t) := f(t)/(f(t) + f(\varepsilon - t))$ for $t \in \mathbb{R}$. Then g_ε is differentiable, $0 \leq g_\varepsilon \leq 1$, $g_\varepsilon(t) = 0$ iff $t \leq 0$ and $g_\varepsilon(t) = 1$ iff $t \geq \varepsilon$.

Ex. 108. Let f, g be as in Ex. 107. For $r > 0$ and $x \in \mathbb{R}^n$, define $\varphi(x) := 1 - g_\varepsilon(\|x\| - r)$. Then φ is smooth and has the following properties: (i) $0 \leq \varphi \leq 1$, (ii) $\varphi(x) = 1$ iff $\|x\| \leq r$ and $\varphi(x) = 0$ iff $\|x\| \geq r + \varepsilon$.

Ex. 109. Let $\psi(u) = u^{-k} e^u$ for $u > 0$. Show that

$$\begin{aligned} \psi'(u) &= (u-k) u^{-k-1} e^u \\ \psi''(u) &= [u^2 - 2ku + k(k+1)] u^{-k-2} e^u \end{aligned}$$

Show that the expression in the brackets has a minimum when $u = k$ and is positive at $u = k$. Hence $\psi''(u) > 0$ and for any u_0

$$\psi(u) \geq \psi(u_0) + \psi'(u_0)(u - u_0).$$

If $u_o > k$, then $\psi'(u_0) > 0$ and the right hand side above tends to infinity as $n \to \infty$. Hence conclude that $\psi(u) \to \infty$ as $u \to \infty$.

Ex. 110. Use the above exercise to prove that f as defined in Ex. 106 is smooth.

Ex. 111. Let $0 < a < b$. Consider the functions $f_a \colon R \to \mathbb{R}$ given by $f_a(t) = \exp(-1/(t-a))$ for $t \geq a$ and 0 otherwise. and $g_b \colon \mathbb{R} \to \mathbb{R}$ given by $g_b(t) = \exp(1/(t-b))$ for $t \leq b$ and 0 otherwise. Then the product φ of these functions is a smooth function which is 0 outside the interval $[a, b]$. Set $\eta(x) := \varphi(\|x\|)$ for $x \in \mathbb{R}^n$. List the properties of η.

Ex. 112. Let φ be as in Ex. 111. Define h on $\mathbb{R}$ as follows.

$$h(x) := \left(\int_x^b \varphi(t)\, dt\right)\left(\int_a^b \varphi(t)\, dt\right)^{-1}.$$

Then h is smooth with $h(x) \leq 1$ for $x \leq a$ and $h(x) = 0$ if $x \geq b$. Define $\psi(x) := h(\sum_i x_i^2)$ for $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$, then $\psi(x) = 1$ for $x \in B(0, a)$ and $\psi(x) = 0$ for $\|x\| \geq b$.

Ex. 113. If K is a compact set in $\mathbb{R}^n$ and U is an open set containing K then there exists a smooth function f on $\mathbb{R}^n$ which is 1 on K and 0 outside U (i.e., 0 on $\mathbb{R}^n \setminus U$).