Notes On Advanced Calculus of Several Variables

Thomas Hughes

February 2018

Preface

These are a collection of notes on the topic of multivariable calculus, taken from C.H. Edwards Jr.'s text book. These notes are not intended for professional use and so will not include rigorous citations.

Chapter 1

Euclidean Space and Linear Mappings

1 The Vector Space $\mathbb{R}^n$

As a set, $\mathbb{R}^n$ is simply the collection of all ordered n-tuples of real numbers. That is,

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}^n\}$$

It therefore follows $\mathbb{R}^n = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n-times}$. The elements of $\mathbb{R}^n$ are generally referred to as *vectors*. This means that vectors are simply n-tuples of real numbers, *not* directed line segments, or an equivalence class of directed line segments; notions which are common in introductory treatments of the topic.

The set $\mathbb{R}^n$ is endowed with two operations called *vector addition* and *scalar multiplication*. Given two vectors x and y in $\mathbb{R}^n$, vector addition is a composition map, $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, defined by

$$x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

Scalar multiplication is a map going from $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$ defined as

$$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$$

Without further discussion, we see that it follows that $\langle \mathbb{R}^n, +, \cdot \rangle$ forms a *vector space*, whose definition is assumed knowledge.

Exercises

1.1 Given $(a_1, \ldots, a_n) \in \mathbb{R}^n$, show that the set of all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that $a_1 x_1 + \ldots + a_n x_n = 0$ is a subspace of $\mathbb{R}^n$.

Proof. It is enough to show that if x and y are two vectors in our set, then $\alpha x + \beta y$ must also be in the set. To that end, we compute

$$\alpha x + \beta y = (\alpha x_1 + \beta y_1, \ldots, \alpha x_n + \beta y_n)$$

and we see that

$$a_1(\alpha x_1 + \beta y_1) + \ldots + a_n(\alpha x_n + \beta y_n) = \alpha \left(\sum a_i x_i\right) + \beta \left(\sum a_i y_i\right) = 0$$

as desired. $\square$

1.2 Prove that the intersection of two subspaces of $\mathbb{R}^n$ is also a subspace.

Proof. Let V and W be subspaces of $\mathbb{R}^n$ and let $x, y \in V \cap W$. Since V and W are subspaces, we know

$$\alpha x + \beta y \in V \quad \text{and} \quad \alpha x + \beta y \in W$$

and so $\alpha x + \beta y \in V \cap W$. Therefore $V \cap W$ is a subspace. $\square$

1.3 Given subspaces V and W of $\mathbb{R}^n$, denote by $V + W$ the set of all vectors $v + w$ with $v \in V$ and $w \in W$. Show that $V + W$ is a subspace of $\mathbb{R}^n$.

Proof. Let $v_1 + w_1, v_2 + w_2 \in V + W$. Then, since V and W are subspaces, we get

$$\alpha(v_1 + w_1) + \beta(v_2 + w_2) = (\alpha v_1 + \beta v_2) + (\alpha w_1 + \beta w_2) \in V + W$$

implying $V + W$ is a subspace. ☐

1.4 If V is the set of all $(x, y, z) \in \mathbb{R}^3$ such that $x + 2y = 0$ and $x + y = 3z$, show that V is a subspace of $\mathbb{R}^3$.

Proof. Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be two elements of our set. Then

$$\alpha x_1 + \beta x_2 + 2\alpha y_1 + 2\beta y_2 = \alpha(x_1 + 2y_1) + \beta(x_2 + 2y_2) = 0$$

and

$$\alpha x_1 + \beta x_2 + \alpha y_1 + \beta y_2 = \alpha(x_1 + y_1) + \beta(x_2 + y_2) = 3(\alpha z_1 + \beta z_2)$$

which implies that $\alpha(x_1, y_1, z_1) + \beta(x_2, y_2, z_2)$ is also in our set. Therefore, it is a subspace. ☐

1.5 Let $\mathscr{D}_0$ denote the set of all differentiable real-valued functions on $[0, 1]$ such that $f(0) = f(1) = 0$. Show that $\mathscr{D}_0$ is a vector space with properly defined addition and multiplication. Would this be true if the condition $f(0) = f(1) = 0$ were replaced by $f(0) = 0$ and $f(1) = 1$?

Proof. Verifying that $\mathscr{D}_0$ is a vector space is a trivial matter. Assuming that $f(1) = 1$ would imply that, for $f, g \in \mathscr{D}_0$, $f(1) + g(1) = 1 + 1 = 2$ so that $\mathscr{D}_0$ is not closed wrt addition, implying that it isn't a vector space. ☐

1.6 Given a set S, denote by $\mathscr{F}(S, \mathbb{R})$ the set of all real-valued functions on S, that is, all maps $S \to \mathbb{R}$. Show that $\mathscr{F}(S, \mathbb{R})$ is a vector space with the operations defined properly. Note that $\mathscr{F}(\{1, \ldots, n\}, \mathbb{R})$ can be interpreted as $\mathbb{R}^n$ since the function $\phi \in \mathscr{F}(\{1, \ldots, n\}, \mathbb{R})$ may be regarded as the n-tuple $(\phi(1), \phi(2), \ldots, \phi(n))$.

Proof. Trivial. ☐

2 Subspaces of $\mathbb{R}^n$

In this section, we will develop the notion of the *dimension* of a vector space, and apply our results in particular to $\mathbb{R}^n$. First we define what we mean by linear independence and linear dependence. We say that a collection of vectors $\{v_1, \ldots, v_n\}$ are *linearly dependent* if there exists a linear combination satisfying

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0 \qquad \text{not all } \alpha_i = 0$$

Proposition 2.1 If $\{v_1, \ldots, v_n\}$ is linearly dependent then one of the vectors is a linear combination of the others.

Proof. Suppose $\{v_1, \ldots, v_n\}$ is linearly dependent. Then there exists

$$\alpha_1 v_1 + \ldots + \alpha_i v_i + \ldots + \alpha_n v_n = 0$$

where $\alpha_i \neq 0$. But then

$$-\frac{\alpha_1}{\alpha_i} v_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \ldots - \frac{\alpha_n}{\alpha_i} v_n = v_i$$

□

We say that a collection of vectors is *linearly independent* if they are not linearly dependent. It is obvious that this then means that a collection of vectors is linearly independent when

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

implies $\alpha_i = 0$ for every $1 \leq i \leq n$. From the above proposition, this is equivalent to saying that none of the vectors can be written as a linear combination of the others.

If $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$, then we define $span\{v_1, \ldots, v_k\} = \{\alpha_1 v_1 + \ldots + \alpha_k v_k : \alpha_i \in \mathbb{R}\}$. That is, the span of a set is simply the set of all linear combinations of its elements. We leave it as an exercise to show that the span of a set is a subspace of $\mathbb{R}^n$.

Proposition 2.2 If $\{v_1, \ldots, v_n\}$ is linearly independent and $v \in span\{v_1, \ldots, v_n\}$ then there is a unique linear combination such that

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = v$$

Proof. If

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = v = \beta_1 v_1 + \ldots + \beta_n v_n$$

then

$$(\alpha_1 - \beta_1) v_1 + \ldots + (\alpha_n - \beta_n) v_n = 0$$

which, by independence of $\{v_1, \ldots, v_n\}$, implies $\alpha_i = \beta_i$. □

If $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ is linearly independent and *span* $\{v_1, \ldots, v_k\} = \mathbb{R}^n$, we say it is a *basis* of $\mathbb{R}^n$. That is, $\{v_1, \ldots, v_k\}$ is a basis of $\mathbb{R}^n$ if it is a linearly independent spanning set of $\mathbb{R}^n$.

Proposition 2.3 The collection $\{e_1, \ldots, e_n\} \subset \mathbb{R}^n$, where $e_i = (0, \ldots, 1, \ldots, 0)$ with 1 in the ith component, is a basis of $\mathbb{R}^n$.

Proof. If $x \in \mathbb{R}^n$, then

$$x = (x_1, \ldots, x_n) = x_1 e_1 + \ldots + x_n e_n$$

so that *span* $\{e_1, \ldots, e_n\} = \mathbb{R}^n$. Now, if

$$\alpha_1 e_1 + \ldots + \alpha_n e_n = (0, \ldots, 0)$$

then

$$(\alpha_1, \ldots, \alpha_n) = (0, \ldots, 0)$$

implying $\alpha_i = 0$ for all $1 \leq i \leq n$, so that $\{e_1, \ldots, e_n\}$ is linearly independent. Therefore it is a basis. $\square$

Proposition 2.4 If $\{v_1, \ldots, v_k\}$ is a basis of $\mathbb{R}^n$, then $k = n$.

Proof. Suppose $\{v_1, \ldots, v_k\}$ is a basis of $\mathbb{R}^n$. Then $\{v_1, \ldots, v_k, e_1\}$ is dependent. Thus

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = e_1$$

So there is a first non-zero α_i so that

$$\alpha_i v_i + \alpha_{i+1} v_{i+1} + \ldots + \alpha_k v_k = e_1$$

which implies

$$\frac{1}{\alpha_i} e_i - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \ldots - \frac{\alpha_k}{\alpha_i} v_k = v_i$$

which implies $\{e_1, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k\}$ is a basis of $\mathbb{R}^n$. So, again $\{e_1, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k, e_2\}$ is dependent. Thus there is

$$\beta_1 e_1 + \alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \ldots + \alpha_k e_k = e_2$$

Now if every $\alpha_j = 0$ then we would have $e_1 = e_2$ which is clearly absurd. Thus there is some first non-zero α_j so that

$$-\frac{\beta_1}{\alpha_j} e_1 + \frac{1}{\alpha_j} e_2 - \frac{\alpha_1}{\alpha_j} v_1 - \ldots - \frac{\alpha_{j-1}}{\alpha_j} v_{j-1} - \frac{\alpha_{j+1}}{\alpha_j} v_{j+1} - \ldots - \frac{\alpha_k}{\alpha_j} v_k = v_j$$

which implies $\{e_1, e_2, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k\}$ is a basis of $\mathbb{R}^n$. Continuing in this fashion, we get that either $\{e_1, \ldots, e_n, v_{i_1}, \ldots, v_{i_{k-n}}\}$ or $\{e_1, \ldots, e_k\}$ is a basis of $\mathbb{R}^n$. Note that the former, by our previous proposition, is dependent and therefore is not a basis. For the latter, if $k < n$ then, again by our previous proposition, $e_{k+1} \notin span\,\{e_1, \ldots, e_k\}$ which would imply it is not a basis. Therefore $k = n$. □

Corollary 2.5 Every basis of $\mathbb{R}^n$ has the same number of vectors in it.

Corollary 2.6 Every collection of n linearly independent vectors in $\mathbb{R}^n$ is a basis of $\mathbb{R}^n$.

Note that with some simple modifications, the above proposition can be used to show that if V is a vector space with a finite basis, then every basis of V is finite, and has the same cardinal number. In general, we say that the dimension of a vector space V, denoted $dim\, V$, is the cardinality of a basis of V. In particular, $dim\, \mathbb{R}^n = n$. If there is some vector space V such that for all $n \in \mathbb{N}$ there is a $\{v_1, \ldots, v_n\} \subset V$ which is linearly independent, then V is *infinite-dimensional*. Notice that the space of all continuous real-valued functions is infinite-dimensional since $\{1, x, x^2, \ldots, x^n\}$ is linearly independent for all n.

Exercises

2.1 Why is it true that the vectors $v_1, \ldots, v_k$ are linearly dependent if any one of them is zero?

Proof. If $v_i = 0$, then

$$0v_1 + \ldots + v_i + \ldots + 0v_k = 0$$

which implies that $\{v_1, \ldots, v_k\}$ is dependent. □

2.2 Which of the following sets of vectors are bases for the appropriate space $\mathbb{R}^n$?

1. $(1, 0)$ and $(1, 1)$.
2. $(1, 0, 0), (1, 1, 0)$, and $(0, 0, 1)$.
3. $(1, 1, 1), (1, 1, 0)$, and $(1, 0, 0)$.
4. $(1, 1, 1, 0), (1, 0, 0, 0), (0, 1, 0, 0)$ and $(0, 0, 1, 0)$.

5. $(1,1,1,1),(1,1,1,0),(1,1,0,0)$, and $(1,0,0,0)$.

Proof. 1. This is a basis of $\mathbb{R}^2$ since for all α, it is clear

$$\alpha(1,0) \neq (1,1)$$

and given (x,y)

$$(x-y)(1,0)+y(1,1)=(x-y+y,y)=(x,y)$$

2. This is a basis of $\mathbb{R}^3$ since

$$\alpha(1,0,0)+\beta(1,1,0)+\gamma(0,0,1)=(0,0,0)$$

implies

$$(\alpha+\beta,\beta,\gamma)=(0,0,0)$$

and solving the simultaneous equations gives $\alpha=\beta=\gamma=0$. Furthermore, given (x,y,z)

$$(x-y)(1,0,0)+y(1,1,0)+z(0,0,1)=(x,y,z)$$

3. This is also a basis of $\mathbb{R}^3$ since

$$\alpha(1,1,1)+\beta(1,1,0)+\gamma(1,0,0)=(0,0,0)$$

implies

$$(\alpha+\beta+\gamma,\alpha+\beta,\alpha)=(0,0,0)$$

which implies $\alpha=\beta=\gamma=0$. Furthermore, given (x,y,z)

$$z(1,1,1)+(y-z)(1,1,0)+(x-y)(1,0,0)=(x,y,z)$$

4. This is not a basis of $\mathbb{R}^4$ since

$$(1,0,0,0)+(0,1,0,0)+(0,0,1,0)=(1,1,1,0)$$

5. This is a basis of $\mathbb{R}^4$. You can show this by applying the same techniques used in (3).

□

2.3 Find the dimension of the subspace V of $\mathbb{R}^4$ that is generated by the vectors $(0,1,0,1)$, $(1,0,1,0)$, and $(1,1,1,1)$.

Proof. We notice that

$$(1,0,1,0)+(0,1,0,1)=(1,1,1,1)$$

On the other hand if

$$\alpha(1,0,1,0)+\beta(0,1,0,1)=(0,0,0,0)$$

then

$$(\alpha,\beta,\alpha,\beta)=(0,0,0,0)$$

which implies that $\alpha=\beta=0$. Thus $\{(1,0,1,0),(0,1,0,1)\}$ is linearly independent. Thus the dimension is 2. □

2.4 Show that the vectors $(1,0,0,1),(0,1,0,1),(0,0,1,1)$ form a basis for the subspace V of $\mathbb{R}^4$ which is defined by the equation $x_1+x_2+x_3-x_4=0$.

Proof. Suppose $u \in span\{(1,0,0,1),(0,1,0,1),(0,0,1,1)\}$. Then

$$\begin{aligned} u &= \alpha_1(1,0,0,1)+\alpha_2(0,1,0,1)l+\alpha_3(0,0,1,1) \\ &= (\alpha_1,\alpha_2,\alpha_3,\alpha_1+\alpha_2+\alpha_3) \end{aligned}$$

which implies that if $u=(0,0,0,0)$ then $\alpha_1=\alpha_2=\alpha_3=0$. Furthermore if $\alpha_i = x_i$, then we get

$$span\{(1,0,0,1),(0,1,0,1),(0,0,1,1)\} = \{(x_1,x_2,x_3,x_4)\in\mathbb{R}^4 : x_1+x_2+x_3-x_4=0\}$$

□

2.5 Show that any set $v_1,\ldots,v_k$, of linearly independent vectors in a vector space V can be extended to a basis for V. That is, if $k<n=dim V$, then there exist vectors $v_{k+1},\ldots,v_n$ in V such that $v_1,\ldots,v_n$ is a basis for V.

Proof. As mentioned in the section, a few minor modifications to Proposition 2.4 will show for a finite-dimensional vector space V, every basis of V has $dim V$ number of vectors in it. Thus if $\{v_1,\ldots,v_k\}$ is linearly independent and $k<n$, then it cannot be a basis, and so it cannot span V. Thus there is a vector v_{k+1} such that $\{v_1,\ldots,v_k,v_{k+1}\}$ is linearly independent. Indeed we can find $n-k$ vectors, $v_{k+1},\ldots,v_n$, such that $\{v_1,\ldots,v_k,v_{k+1},\ldots,v_n\}$ is linearly independent in V which would imply that it is a basis of V. □

Chapter 2

Multivariable Differential Calculus

As in the single-variable case, the basic conceit of multivariable differential calculus will be to use linear functions to approximate nonlinear functions. Recall that in the calculus of single-variable functions, we can use the derivative at a point–when it exists–to form the line tangent to the graph of the function at that point to approximate the function. That is, the line

$$y = f'(a)(x-a) + f(a)$$

can be used to obtain approximate values of $f(x)$ when x is sufficiently close to a.

We might also say that $f'(a)$ is a linear approximation to the change $f(x) - f(a)$. To be more precise, let $\Delta f_a(h) = f(a+h) - f(a)$ and $df_a(h) = f'(a)h$. The <u>linear</u> mapping $df_a : \mathbb{R} \to \mathbb{R}$, defined by $df_a(h) = f'(a)h$, is called the *differential* of f at a; it is simply that linear mapping $\mathbb{R} \to \mathbb{R}$ whose matrix is the *derivative* $f'(a)$ of f at a (the matrix of a linear mapping $\mathbb{R} \to \mathbb{R}$ being just a real number). With this terminology we find that when h is small, the linear change $df_a(h)$ is a good approximation to the actual change $\Delta f_a(h)$ in the sense that

$$\lim_{h\to 0} \frac{\Delta f_a(h) - df_a(h)}{h} = \lim_{h\to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0$$

Here $df_a(h)$ will be called the differential of f at a; its $m \times n$ matrix will be called the derivative of f at a, thus preserving the above relationship between the differential (a linear mapping) and the derivative (its matrix).

1 Curves in $\mathbb{R}^n$

We begin with the special case of a mapping $f : \mathbb{R} \to \mathbb{R}^m$. We define

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

so that, should this limit exist, we say that f is differentiable at $a \in \mathbb{R}$. The *derivative* of f, $f'(a)$, is a vector tangent to image curve of f at the point $f(a)$. If we imagine that the curve is simply the motion traced out by the movement of a particle then the length $|f'(a)|$ is the speed at time $t = a$ of the particle. For this reason, we can refer to $f'(a)$ as the *velocity vector* at time $t = a$.

If the derivative mapping $f' : \mathbb{R} \to \mathbb{R}$ is itself differentiable at a, its derivative at a is the *second derivative* $f''(a)$ of f at a. Again, if we take f in physical terms, then $f''(a)$ is a vector which represents the *acceleration* of a particle at time a. We can say that $f''(a)$ is the *acceleration vector*.

Since limits can be taken coordinatewise, it follows that

$$f' = (f_1', \ldots, f_m')$$

That is, the differentiable function $f : \mathbb{R} \to \mathbb{R}^m$ may be differentiated coordinatewise. This leads to the following theorem.

Theorem 1.1 Let $f, g : \mathbb{R} \to \mathbb{R}^m$, and $\phi : \mathbb{R} \to \mathbb{R}$ all be differentiable. Then

$$\begin{aligned}
(f+g)' &= f' + g' \\
(\phi f)' &= \phi' f + \phi f' \\
(f \cdot g)' &= f' \cdot g + f \cdot g' \\
(f \circ \phi)'(t) &= \phi'(t) f'(\phi(t))
\end{aligned}$$

Proof.

1. Notice the ith coordinate of $f+g$ is $f_i + g_i$ which implies that the ith coordinate of $(f+g)'$ is $(f_i+g_i)' = f_i' + g_i'$
2. Notice the ith coordinate of ϕf is ϕf_i which implies that the ith coordinate of $(\phi f)'$ is $(\phi f_i)' = \phi' f_i + \phi f_i'$.
3. Notice the ith coordinate of $f \cdot g$ is $f_i \cdot g_i$ which implies that the ith coordinate of $(f \cdot g)'$ is $(f_i \cdot g_i)' = f_i' \cdot g + f_i \cdot g_i'$.
4. Notice the ith coordinate of $f \circ \phi$ is $f_i \circ \phi$ which implies that the ith coordinate of $(f \circ \phi)'$ is $(f_i \circ \phi)' = \phi' f_i'(\phi)$

□

The *tangent line* at $f(a)$ to the image curve of the differentiable mapping $f : \mathbb{R} \to \mathbb{R}^m$ is, by definition, that straight line which passes through $f(a)$ and is parallel to the tangent vector $f'(a)$. We now inquire as to how well this tangent line approximates the curve close to $f(a)$. That is, how closely does the mapping $h \to f(a) + hf'(a)$ of $\mathbb{R}$ into $\mathbb{R}^m$ (whose image is the tangent line) approximate the mapping $h \to f(a+h)$? Let us write

$$\Delta f_a(h) = f(a+h) - f(a)$$

for the actual change in f from a to $a+h$, and

$$df_a(h) = hf'(a)$$

for the linear (as a function of h) change along the tangent line. Thus we are asking how small the difference vector $\Delta f_a(h) - df_a(h)$ is when h is small. The answer is that it goes to zero faster than h does. That is,

$$\begin{aligned}
\lim_{h\to 0} \frac{\Delta f_a(h) - df_a(h)}{h} &= \lim_{h\to 0} \frac{f(a+h) - f(a) - hf'(a)}{h} \\
&= \left(\lim_{h\to 0} \frac{f(a+h)_f(a)}{h} \right) - f'(a) \\
&= 0
\end{aligned}$$

This already provides the "only if" part of the following theorem

Theorem 1.2 The mapping $f : \mathbb{R} \to \mathbb{R}^m$ is differentiable if and only if there is a linear mapping $L : \mathbb{R} \to \mathbb{R}^m$ such that

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

in which case L is defined by $L(h) = df_a(h) = f'(a)h$.

Proof. Suppose that there is a linear mapping satisfying the above. Then there is a vector $b \in \mathbb{R}^m$ such that L is defined by $L(h) = hb$; we must show that $f'(a)$ exists and that it is equal to b. But

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h\to 0} \left(\frac{f(a+h) - f(a) - hb}{h} \right) + b = b$$

$\square$

If $f : \mathbb{R} \to \mathbb{R}^m$ is differentiable at a, then the linear mapping $df_a : \mathbb{R} \to \mathbb{R}^m$, defined by $df_a(h) = hf'(a)$, is called the *differential* of f at a. Notice that the derivative vector $f'(a)$ is, as a column vector, the matrix of the linear mapping df_a, since

$$df_a(h) = hf'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix} h$$

The following discussion provides some motivation for the notation df_a for the differential of f at a. Let us consider the identity function $\mathbb{R} \to \mathbb{R}$, and write x for its name as well as its value at x. Since its derivative is 1 everywhere, its differential at a is defined by

$$dx_a(h) = 1 \cdot h = h$$

If f is real-valued, and we substitute $h = dx_a(h)$ into the definition of $df_a : \mathbb{R} \to \mathbb{R}$, we obtain

$$df_a(h) = f'(a)h = f'(a)dx_a(h)$$

So the two linear mappings df_a and $f'(a)dx_a$ are equal,

$$df_a = f'(a)dx_a$$

If we now use the Leibniz notation $f'(a) = \frac{df}{dx}$ and drop the subscript a, we obtain the famous formula

$$df = \frac{df}{dx}dx$$

which now not only makes sense, but is true! It is an actual equality of linear mappings of the real line into itself.

Now let f and g be two differentiable functions from $\mathbb{R} \to \mathbb{R}$, and write $h = g \circ f$ for the composition. Then the chain rule gives

$$\begin{aligned} dh_a(t) &= h'(a)t \\ &= g'(f(a))[f'(a)t] \\ &= g'(f(a))[df_a(t)] \\ &= dg_{f(a)}(df_a(t)) \end{aligned}$$

so we see that the single-variable chain rule takes the form

$$dh_a = dg_{f(a)} \circ df_a$$

Exercises

1.1 Let $f : \mathbb{R} \to \mathbb{R}^m$ be a differentiable mapping with $f'(t) \neq 0$ for all $t \in \mathbb{R}$. Let p be a fixed point not on the image curve of f. If $q = f(t_0)$ is the point of the curve closest to p, that is, if $|p - q| \leq |p - f(t)|$ for all $t \in \mathbb{R}$, show that the vector $p - q$ is orthogonal to the curve at q.

Proof. Let all the hypotheses hold and let us take $\phi : \mathbb{R} \to \mathbb{R}$ to be defined by

$$\phi(t) = |p - f(t)|^2 = \sum_{i=1}^{m} (p_i - f_i(t))^2$$

By the hypotheses on p and $f(t)$, we know that $\phi(t)$ is differentiable for all t and has a minimum at t_0. Thus we have

$$\phi'(t_0) = (-2) \sum_{i=1}^{m} (f_i'(t_0)(p_i - f_i(t_0))) = 0$$

which implies

$$0 = \sum_{i=1}^{m} (f_i'(t_0)(p_i - f_i(t_0))) = \langle f'(t_o), p - f(t_0) \rangle = \langle f'(t_o), p - q \rangle$$

so that $p - q$ is orthogonal to the curve at q. □

1.2

a Let $f, g : \mathbb{R} \to \mathbb{R}^n$ be two differentiable curves, with $f'(t) \neq 0$ and $g'(t) \neq 0$ for all $t \in \mathbb{R}$. Suppose the two points $p = f(s_0)$ and $q = g(t_0)$ are closer than any other pair of points on the two curves. Then prove that the vector $p - q$ is orthogonal to both velocity vectors $f'(s_0)$ and $g'(t_0)$.

b Apply the result of (a) to find the closest pair of points on the "skew" straight lines in $\mathbb{R}^3$ defined by $f(s) = (s, 2s, -s)$ and $g(t) = (t+1, t-2, 2t+3)$.

Proof. a If we consider q as a point, then p is the point on f closest to q and thus by the previous exercise we will get that $p - q$ is orthogonal to $f'(s_0)$. A similar argument demonstrates the same for $g'(t_0)$.

b Notice that $dg_s = (1, 2, -1)$ and $df_t = (1, 1, 2)$. If there is a unique 2-tuple (s_0, t_0), such that

$$\langle dg_{s_0}, f(t_0) - g(s_0)\rangle = \langle df_{t_0}, f(t_0) - g(s_0)\rangle = 0$$

then, by (a), we will have our solution. To that end, we solve the simlu-ateneous equations

$$\begin{aligned} \langle (1,2,-1), (t+1-s, t-2-2s, 2t+3+s)\rangle &= 0 \\ \langle (1,1,2), (t+1-s, t-2-2s, 2t+3+s)\rangle &= 0 \end{aligned}$$

which is a linear system in t and s. We get the unique solution

$$(s_0, t_0) = \left(\frac{31}{35}, \frac{396}{35}\right)$$

□

1.3 Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a *conservative* force field on $\mathbb{R}^n$, meaning that there exists a continuously differentiable *potential function* $V : \mathbb{R}^n \to \mathbb{R}$ such that $F(x) = -\nabla V(x)$ for all $x \in \mathbb{R}^n$ [recall that $\nabla V = (\partial V/\partial x_1, \ldots, \partial V/\partial x_n)$]. Call the curve $\phi : \mathbb{R} \to \mathbb{R}^n$ a "quasi-Newtonian particle" if and only if there exist constants $m_1, m_2, \ldots, m_n$, called its "mass components," such that

$$F_i(\phi(t)) = m_i \phi_i''(t) \quad (F = ma)$$

for each $i = 1, \ldots, n$. Thus, with respect to the x_i-direction, it behaves as though it has mass m_i. Define its *kinetic energy* $K(t)$ and *potential energy* $P(t)$ at time t by

$$K(t) = \frac{1}{4}\sum_{i=1}^{n} m_i[\phi_i'(t)]^2, \quad P(t) = V(\phi(t))$$

Now prove that the law of the *conservation of energy* holds for quasi-Newtonian particles, that is, $K + P =$ constant.

Proof. Notice if $V : \mathbb{R}^n \to \mathbb{R}$ is $\mathscr{C}^1$, then

$$\nabla V = (D_1 V_1, \ldots, D_n V) = V'$$

and so

$$F(x) = -V'(x)$$

Now,

$$K(t) = \frac{1}{2}\sum m_i[\phi_i(t)]^2$$
$$K'(t) = \sum m_i(\phi_i'(t))(\phi_i''(t))$$

and

$$\begin{aligned} P(t) &= V(\phi(t)) \\ P'(t) &= V'(\phi(t))\phi'(t) \\ &= -F(\phi(t))\phi'(t) \end{aligned}$$

and since $F_i(\phi(t)) = m_i\phi_i''(t)$ it follows that

$$\begin{aligned} P'(t) &= \begin{bmatrix} -m_1\phi_1''(t) & \dots & -m_n\phi_n''(t) \end{bmatrix} \begin{bmatrix} \phi_1'(t) \\ \vdots \\ \phi_n'(t) \end{bmatrix} \\ &= -\sum m_i(\phi_i'(t))(\phi_i''(t)) \end{aligned}$$

which implies that $(K(t) + P(t))' = K'(t) + P'(t) = 0$. Since $K + P : \mathbb{R} \to \mathbb{R}$ we know that this therefore implies that $K + P = c$ where $c \in \mathbb{R}$ □

1.4 (n-body problem)

1.5 If $f : \mathbb{R} \to \mathbb{R}^m$ is linear, prove that $f'(a)$ exists for all $a \in \mathbb{R}$, with $df_a = f$.

Proof. Notice
$$\lim_{h\to 0} \frac{f(a+h) - f(a) - f(h)}{h} = \lim_{h\to 0} \frac{0}{h} = 0$$
so that by Theorem 1.2 f is differentiable and $df_a = f$. □

1.6 If L_1 and L_2 are two linear mappings from $\mathbb{R}$ to $\mathbb{R}^n$ satisfying Theorem 1.2, prove that $L_1 = L_2$.

Proof. Since both $L_1(h)$ and $L_2(h)$ satisfy the equation in Theorem 1.2, it follows

$$\begin{aligned} 0 &= \lim_{h\to 0} \frac{f(a+h) - f(a) - L_1(h)}{h} - \frac{f(a+h) - f(a) - L_2(h)}{h} \\ &= \lim_{h\to 0} \frac{L_2(h) - L_1(h)}{h} \\ &= \lim_{h\to 0} L_2(1) - L_1(1) \\ &= L_2(1) - L_1(1) \end{aligned}$$

so that $L_2(1) = L_1(1)$ and by linearity it follows $L_2(h) = L_1(h)$. □

1.7 Let $f, g : \mathbb{R} \to \mathbb{R}$ both be differentiable at a.

1. Show that $d(fg)_a = g(a)df_a + f(a)dg_a$.
2. Show that
$$d\left(\frac{f}{g}\right)_a = \frac{g(a)df_a - f(a)dg_a}{(g(a))^2} \qquad g(a) \neq 0$$

Proof. 1. Notice that
$$f(a+h)g(a+h) - f(a)g(a) = \Delta f_a \Delta g_a + g(a)\Delta f_a + f(a)\Delta g_a$$
which implies
$$\begin{aligned} d(fg)_a &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\ &= \lim_{h\to 0} \frac{\Delta f_a(h)\Delta g_a(h)}{h} + \lim_{h\to 0} \frac{g(a)\Delta f_a(h)}{h} + \lim_{h\to 0} \frac{f(a)\Delta g_a(h)}{h} \\ &= 0 + g(a)df_a + f(a)dg_a \end{aligned}$$

2. Notice that if $h(x) = \frac{1}{g(x)}$ then
$$\begin{aligned} dh_a &= \lim_{h\to 0} \frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} \\ &= \lim_{h\to 0} \frac{g(a) - g(a+h)}{g(a+h)g(a)h} \\ &= \lim_{h\to 0} -\left(\frac{g(a+h) - g(a)}{h}\right)\left(\frac{1}{g(a+h)g(a)}\right) \\ &= (-dg_a)\left(\frac{1}{[g(a)]^2}\right) \\ &= -\frac{dg_a}{[g(a)]^2} \end{aligned}$$
and thus $\left(\frac{f}{g}\right)(x) = f(x)[g(x)]^{-1} = f(x)h(x)$ and so
$$\begin{aligned} d\left(\frac{f}{g}\right)_a &= d(fh)_a \\ &= h(a)df_a + f(a)dh_a \\ &= \frac{df_a}{g(a)} - \frac{f(a)dg_a}{[g(a)]^2} \\ &= \frac{g(a)df_a - f(a)dg_a}{[g(a)]^2} \end{aligned}$$

□

1.8 Let $\gamma(t)$ be the position vector of a particle moving with constant acceleration vector $\gamma''(t) = a$. Then show that $\gamma(t) = \frac{1}{2}t^2 a + tv_0 + p_0$ where $p_0 = \gamma(0)$ and $v_0 = \gamma'(0)$. If $a = 0$, conclude that the particle moves along a straight line through p_0 with velocity vector v_0 (the law of inertia).

Proof. We see that

$$(\gamma_1''(t), \ldots, \gamma_n''(t)) = \gamma''(t) = (a_1, \ldots, a_n)$$

which implies

$$\gamma_i'(t) = \int \gamma_i''(t)dt = \int a_i = ta_i + v_0^i$$

which, in turn, implies

$$\gamma_i(t) = \int \gamma_i'(t) = \int ta_i + c_1 = \frac{1}{2}t^2 a_i + tv_0^i + p_0^i$$

Thus

$$\gamma(t) = \frac{1}{2}t^2 a + tv_0 + p_0$$

where, clearly, $p_0 = \gamma(0)$ and $v_0 = \gamma'(0)$. Setting $a = 0$ we see that $\gamma(t) = tv_0 + p_0$ is linear in t, and therefore proceeds in a straight line through point p_0. Since $\gamma'(t) = v_0$, the velocity vector is v_0. □

1.9 Let $\gamma : \mathbb{R} \to \mathbb{R}^n$ be a differentiable curve. Show that $|\gamma(t)|$ is constant if and only if $\gamma(t)$ and $\gamma'(t)$ are orthogonal for all t.

Proof. If $\gamma(t) = 0$ there is nothing to show. Otherwise, if $|\gamma(t)|$ is constant, then the image curve of γ is the n-sphere centered at the origin. If we let p be the origin, then Exercise 1.1 above implies the result. Now, if $\gamma(t)$ and $\gamma'(t)$ are orthogonal for all t, then, letting $\phi(t) = |\gamma(t)|^2$ gives us that, for all t:

$$d\phi_t = 2\langle \gamma(t), \gamma'(t) \rangle = 0$$

which, since $\phi : \mathbb{R} \to \mathbb{R}$ we get the single-variable result that $\phi(t)$ must be constant, which implies that $|\gamma(t)|$ is constant. □

1.10 Suppose that a particle moves around a circle in the plane $\mathbb{R}^2$, of radius r centered at 0, with constant speed v. Deduce from the previous exercise that $\gamma(t)$ and $\gamma''(t)$ are both orthogonal to $\gamma'(t)$, so it follows that $\gamma''(t) = k(t)\gamma(t)$. Substitute this result into the equation obtained by differentiating $\langle \gamma(t), \gamma'(t) \rangle = 0$ to obtain $k = -v^2/r^2$. Thus the acceleration vector always points towards the origin and has constant length v^2/r.

Proof. If $\gamma(t)$ is a circle in the plane, centered at 0 of radius r, then

$$|\gamma(t)| = r$$

for all t. So by the previous exercise we know $\langle \gamma(t), \gamma'(t) \rangle = 0$ for all t. But since the particle has constant speed, by definition, $|\gamma'(t)| = v$ which, again, implies

$$\langle \gamma'(t), \gamma''(t) \rangle = 0$$

Thus $\gamma(t)$ and $\gamma''(t)$ are both orthogonal to $\gamma'(t)$. Thus we can conclude $\gamma''(t) = k(t)\gamma(t)$. Therefore

$$\begin{aligned} 0 &= [\langle \gamma(t), \gamma'(t) \rangle]' \\ &= \sum [\gamma_i'(t)]^2 + \gamma_i(t)^2 k(t) \end{aligned}$$

which implies

$$k(t) = -\frac{|\gamma_i'(t)|^2}{|\gamma_i(t)|^2} = -\frac{v^2}{r^2}$$

□

2 Directional Derivatives and the Differential

Exercises

2.1 If $F : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a, show that F is continuous at a.

Proof. Notice that differentiability of F at a implies

$$F(a + h) = F(a) + dF_a(h) + \epsilon(h)$$

where $\epsilon(h) \to 0$ as $h \to 0$ which implies $\epsilon(h) \to 0$ as $h \to 0$. Thus taking the limit of both sides yields

$$\lim_{h\to 0} F(a + h) = \lim_{h\to 0} F(a) + \lim_{h\to 0} dF_a(h) + \lim_{h\to 0} \epsilon(h) = F(a)$$

□

2.2 If $p : \mathbb{R}^2 \to \mathbb{R}$ is defined by $p(x, y) = xy$, show that p is differentiable everywhere with $dp_{(a,b)} = bx + ay$.

Proof. Notice $D_1p(a, b) = b$ and $D_2p(a, b) = a$, implying that p is continuously differentiable at a. Thus p is differentiable at a and

$$dp_{(a,b)} = \begin{bmatrix} D_1p(a, b) & D_2p(a, b) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = bx + ay$$

□

2.3 If $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by $f(x, y) = \frac{xy^2}{x^2+y^2}$ unless $x = y = 0$, in which case $f(0, 0) = 0$. Show that $D_vf(0, 0)$ exists for all v, but f is *not* differentiable at $(0, 0)$.

Proof. By the definition of the directional derivative, we know that if $v = (0,0)$ there is nothing to show, otherwise

$$\begin{aligned} D_v f(0,0) &= \lim_{t\to 0} \frac{f((0,0)+tv) - f(0,0)}{t} \\ &= \lim_{t\to 0} \frac{t^3 v_1 v_2^2}{t^3(v_1^2+v_2^2)} \\ &= \frac{v_1 v_2^2}{v_1^2+v_2^2} \end{aligned}$$

which exists since v_1 and v_2 are not both zero. Thus $D_1 f(0,0) = D_2 f(0,0) = 0$, and yet $D_{(1,1)} f(0,0) = \frac{1}{2}$. But then $D_{(1,1)} f(0,0) \neq D_1 f(0,0) + D_2 f(0,0)$ so that, by contrapositive, $df_{(0,0)}$ does not exist. $\square$

2.4 Do the same as in the previous problem with the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = (x^{\frac{1}{3}} + y^{\frac{1}{3}})^3$.
2.5 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = x^3 \sin(1/x) + y^2$ for $x \neq 0$ and $f(0,y) = y^2$. Show some stuff about it.
2.6 Use the approximation $\Delta f_a \approx df_a$ to estimate the value of

1. $[(3.02)^2 + (1.97)^2 + (5.98)^2]$
2. $(e^4)^{1/10} = e^{0.4} = e^{1.1^2 - 0.9^2}$

2.7
2.8
2.9 If $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ are both differentiable at $a \in \mathbb{R}^n$, prove directly from the definition that the mapping $h : \mathbb{R}^n \to \mathbb{R}^{m+k}$, defined by $h(x) = (f(x), g(x))$, is differentiable at a.

Proof. If we let $L_a(s) = (df_a(s), dg_a(s))$ then

$$\begin{aligned} \lim_{s\to 0} \frac{|\Delta h_a(s) - L_a(s)|}{|s|} &= \lim_{s\to 0} \frac{|(\Delta f_a(s) - df_a(s), \Delta g_a(s) - dg_a(s))|}{|s|} \\ &\leq \lim_{s\to 0} \frac{|f_a(s) - df_a(s)|}{|s|} + \lim_{s\to 0} \frac{|\Delta g_a(s) - dg_a(s)|}{|\Delta s|} \\ &= 0 \end{aligned}$$

Thus h is differentiable at a and $dh_a = (df_a(s), dg_a(s))$. $\square$

Chapter 3

Successive Approximations and Implicit Functions

Chapter 4

Multiple Integrals

Chapter 5

Line and Surface Integrals; Differential Forms and Stoke's Theorem

Chapter 6

The Calculus of Variations