Vector Calculus

Thomas Hughes

February 2018

Foreword

This is an attempt to combine the better aspects of different textbooks of vector Calculus. We will principally refer to Spivak's *Calculus on Manifolds* and Hubbard & Hubbard's *Vector Calculus, Linear Algebra, and Differential Forms*. This is not intended for commercial or professional use and will therefore not include citations. However, page numbers and other references will be included at times.

Chapter 1

Derivatives

1.1 Important Theorems and Propositions

Definition 1.1 (pg.64, H&H)
If A is an $n \times m$ matrix, its *length* $|A|$, or the *Minkowski norm*, is the square root of the sum of the squares of all its entries:

$$|A|^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}^2$$

Proposition 1.2 (pg.64, H&H)

1. If A is an $n \times m$ matrix, and $\mathbf{x} \in \mathbb{R}^m$, then

$$|A\mathbf{x}| \leq |A| \, |\mathbf{x}|$$

2. If A is an $n \times m$ matrix, and B is a $m \times k$ matrix, then

$$|AB| \leq |A| \, |B|$$

Proof. 1. If we denote the ith row of A by $\mathbf{r}_i$ then we have

$$|A\mathbf{x}|^2 = \sum_{i=1}^{n} (\mathbf{r}_i \cdot \mathbf{x})^2 \leq \left(\sum_{i=1}^{n} |\mathbf{r}_i|^2 \right) |\mathbf{x}|^2 = |A|^2 \, |\mathbf{x}|^2$$

Taking the square root of both sides gives us the result.

2. If we denote the ith column of B by $\mathbf{c}_i$ then we have

$$|AB|^2 = \sum_{i=1}^{k} (A\mathbf{c}_i)^2 \leq |A|^2 \sum_{i=1}^{k} |\mathbf{c}_i|^2 = |A|^2 \, |B|^2$$

Taking the square root of both sides gives us the result.

□

Proposition 1.3 (pg.76, H&H)
A sequence $(\mathbf{a}_m) = \mathbf{a}_1, \mathbf{a}_2, \ldots$ with $\mathbf{a}_i \in \mathbb{R}^n$ converges to $\mathbf{a}$ iff each coordinate converges; i.e., if for all j with $1 \leq j \leq n$ the sequence of the jth coordinates, $((a_m)_j)$, converges to a_j, the jth coordinate of the limit $\mathbf{a}$.

Proof. If $((a_m)_j) \to a_j$ for each $1 \leq j \leq n$ then, given $\epsilon > 0$, there is M such that $m \geq M$ implies

$$|(a_m)_j - a_j| < \frac{\epsilon}{n} \qquad (\forall 1 \leq j \leq n)$$

and thus $m \geq M$ implies

$$||\mathbf{a}_m - \mathbf{a}|| \leq \sum_{j=1}^{n} |(a_m)_j - a_j| < \sum_{i=1}^{n} \frac{\epsilon}{n} = \epsilon$$

so that $(\mathbf{a}_m) \to \mathbf{a}$.

Now if $(\mathbf{a}_m) \to \mathbf{a}$ then, given $\epsilon > 0$, there is M such that $m \geq M$ implies

$$|(a_m)_j - a_j| \leq ||\mathbf{a}_m - \mathbf{a}|| < \epsilon$$

for all j. $\square$

Proposition 1.4 (pg.82, H&H)
Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$. Let $\mathbf{x}_0 \in \overline{U}$. Then $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f} = \mathbf{a}$ iff for each component, $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i = a_i$.

Proof. If $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i = a_i$ for all i then, given $\epsilon > 0$, there is $\delta > 0$ such that $||\mathbf{x} - \mathbf{x}_0|| < \delta$ implies

$$|f_i(\mathbf{x}) - a_i| < \frac{\epsilon}{m}$$

for every i. Therefore, $||\mathbf{x} - \mathbf{x}_0|| < \delta$ implies

$$||\mathbf{f}(\mathbf{x}) - \mathbf{a}|| \leq \sum_{i=1}^{m} |f_i(\mathbf{x}) - a_i| < \sum_{i=1}^{m} \frac{\epsilon}{m} = \epsilon$$

which implies that $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f} = \mathbf{a}$.

Now if $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f} = \mathbf{a}$, then, given $\epsilon > 0$, there $\delta > 0$ such that, $||\mathbf{x} - \mathbf{x}_0|| < \delta$ implies

$$|f_i(\mathbf{x}) - a_i| \leq ||\mathbf{f}(\mathbf{x}) - \mathbf{a}|| < \epsilon$$

so that $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i = a_i$ for all $1 \leq i \leq m$. $\square$

Proposition 1.5 (pg.87, H&H)
Let $(\mathbf{a}_m)$ be a sequence in $\mathbb{R}^n$. If $\displaystyle\sum_{i=1}^{\infty} ||\mathbf{a}_i||$ converges, then $\displaystyle\sum_{i=1}^{\infty} \mathbf{a}_i$ converges.

Proof. Suppose $\displaystyle\sum_{i=1}^{\infty} ||\mathbf{a}_i||$ converges. If we denote

$$\mathbf{s}_m = \sum_{i=1}^{m} \mathbf{a}_i = \left(\sum_{i=1}^{m} (a_i)_1, \ldots, \sum_{i=1}^{m} (a_i)_n \right)$$

By Proposition 1.3, we know that if $\sum_{i=1}^{\infty} |(a_i)|_k$ converges for each $1 \leq k \leq n$, then $(\mathbf{s}_m)$ will converge as well. We see that for every k, we have

$$\sum_{i=1}^{m} |(a_i)k| \leq \sum_{i=1}^{k} ||\mathbf{a}_i|| < \sum_{i=1}^{\infty} ||\mathbf{a}_i||$$

so that, by boundedness and monotonicity, $\sum_{i=1}^{m} |(a_i)_k|$ converges. Therefore, $(\mathbf{s}_m)$ converges, which is what we wanted to show. $\square$

Corollary 1.6 (pg.88, H&H)
Let A be a square matrix. If $|A| < 1$, the series

$$S = I + A + A^2 + \dots$$

converges to $(I - A)^{-1}$.

Proof. As in the real case,

$$S_n(I - A) = I - A^{n+1}$$

By Proposition 1.2 we have

$$\left|A^{n+1}\right| \leq |A|\,|A^n| \leq \dots \leq |A|^{n+1}$$

so that $|A| < 1$ implies $\lim_{n\to\infty} \left|A^{n+1}\right| = 0$ which, in turn, implies that $\lim_{n\to\infty} A^{n+1} = 0$, where 0 is understood to be the zero matrix. So,

$$S(I - A) = \left(\lim_{n\to\infty} S_n\right)(I - A) = \lim_{n\to\infty} S_n(I - A) = \lim_{n\to\infty} I - A^{n+1} = I$$

Since $I - A$ is square, we conclude that $S = (I - A)^{-1}$. $\square$

Corollary 1.7 (pg.88, H&H)
If $|A| < 1$, then $I - A$ is invertible.

Corollary 1.8 (pg.88)
The set of invertible $n \times n$ matrices is open.

Remark: Now that our context is understood to be vector spaces of arbitrary dimension, we will no longer be using boldface characters to indicated vectors, unless it seems especially appropriate. Also, we will use $|\cdot|$ for all norms, dispensing with our convention of using $||\cdot||$ to indicate norms on spaces of dimension possibly greater than 1.

Definition 1.9 (pg.16, S)
A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *differentiable at* $a \in \mathbb{R}^n$ if there is a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

The transformation λ is said to be the *derivative of f at a* and it is denoted $Df(a)$. The matrix of $Df(a)$, with respect to the standard basis, is denoted by $f'(a)$.

Theorem 1.10 (pg.19, S)
If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a, and $g : \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $f(a)$, then the composition $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a, and

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

so that

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$$

Theorem 1.11 (pg.20, S)

1. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a constant function, then
$$Df(a) = 0$$
2. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then
$$Df(a) = f$$
3. If $f : \mathbb{R}^n \to \mathbb{R}^m$, then f is differentiable at a if and only if each f_i is and
$$Df(a) = (Df_1(a), \dots, Df_m(a))$$
4. If $s(x, y) = x + y$ then
$$Ds(a, b) = s$$
5. If $p(x, y) = xy$ then
$$Dp(a, b) = bx + ay$$
so that $p'(a, b) = [ba]$.

Proof. 1.
$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - 0|}{|h|} = \lim_{h \to 0} \frac{0}{|h|} = 0$$

2.
$$\begin{aligned} \lim_{h \to 0} \frac{|f(a+h) - f(a) - f(h)|}{|h|} &= \lim_{h \to 0} \frac{|f(a) + f(h) - f(a) - f(h)|}{|h|} \\ &= \lim_{h \to 0} \frac{0}{|h|} \\ &= 0 \end{aligned}$$

3. If each $Df_i(a)$ exists, and we let $\lambda(h) = (Df_1(a)(h), \dots, Df_m(a)(h))$ then
$$0 \le \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} \le \sum_{i=1}^{m} \frac{|f_i(a+h) - f_i(a) - Df_i(h)|}{|h|}$$

so that, in the limit, we have that $Df(a) = (Df_1(a), \dots, Df_m(a))$.

Now if $Df(a) = (\lambda_1, \dots, \lambda_m)$ exists, then

$$0 \leq \frac{|f_i(a+h) - f_i(a) - \lambda_i(h)|}{|h|} \leq \frac{|f(a+h) - f(a) - Df(a)(h)|}{|h|}$$

so that, again, in the limit, we have that $(Df_1(a), \dots, Df_m(a)) = Df(a)$.

4. This follows from (2).

□

Definition 1.12 (pg.25, S)
If $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$, the limit

$$\lim_{h \to 0} \frac{f(a_1, a_2, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

if it exists, is denoted $D_i f(a)$ and is called the *ith partial derivative of f at a*. We may also denote $D_i f(a)$ by $\frac{\partial f}{\partial x_i}$ or f_{x_i}.

Definition 1.13 (pg.33, S)
If $f : \mathbb{R}^n \to \mathbb{R}$, then for $v \in \mathbb{R}^n$ the limit

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

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

Proposition 1.14 (pg.109, H&H)
If $Df(a)$ exists, then $D_v f(a)$ exists for all v and

$$D_v f(a) = Df(a)(v)$$

Proof. Suppose $Df(a)$ exists. Then, given v, if $g(h) = a + hv$ we have

$$g(0) = a \text{ and } Dg(0) = v$$

so that, by Theorem 1.10, we get that $f \circ g$ is differentiable at 0. Thus

$$D(f \circ g)(0) = \lim_{h \to 0} \frac{(f \circ g)(h) - (f \circ g)(0)}{h} = \lim_{h \to 0} \frac{f(a + hv) - f(a)}{h} = D_v f(a)$$

and so

$$\begin{aligned} D_v f(a) &= D(f \circ g)(0) \\ &= Df(g(0)) \circ Dg(0) \\ &= Df(a)(v) \end{aligned}$$

□

Corollary 1.15 (pg.n/a)
For $f : \mathbb{R}^n \to \mathbb{R}$, $D_i f(a)$ exists iff $D_{e_i} f(a)$ exists and

$$D_i f(a) = D_{e_i} f(a)$$

Corollary 1.16 (pg.n/a)
For $f : \mathbb{R}^n \to \mathbb{R}$, if $Df(a)$ exists, then $D_i f(a)$ exists for all i and

$$D_i f(a) = Df(a)(e_i)$$

Definition 1.17 (pg.107, H&H)
The *Jacobian Matrix* of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is the $m \times n$ matrix composed of partial derivatives of f evaluated at $a \in \mathbb{R}^n$:

$$[Jf(a)] = \begin{bmatrix} D_1 f_1(a) & \dots & D_n f_1(a) \\ \vdots & & \vdots \\ D_1 f_m(a) & \dots & D_n f_m(a) \end{bmatrix}$$

Theorem 1.18 (pg.114, H&H)
If $Df(a)$ exists then $[Jf(a)]$ exists and

$$f'(a) = [Jf(a)]$$

Proof. If $Df(a)$ exists then, by Theorem 1.11, $Df(a) = (Df_1(a), \dots, Df_m(a))$ so that $Df_j(a)$ exists for all $1 \leq j \leq m$. By Corollary 1.16, this implies $D_i f_j(a)$ exists for all i and j. So $[Jf(a)]$ exists. Furthermore

$$\begin{aligned} [f'(a)]_{ij} &= [f_i'(a)]_{1j} \\ &= f_i'(a) \cdot e_j \\ &= Df(a)(e_j) \\ &= D_j f_i(a) \\ &= [Jf(a)]_{ij} \end{aligned}$$

Therefore $f'(a) = [Jf(a)]$. $\square$

Alas, the converse is not necessarily true. Consider

$$f(x, y) = \begin{cases} \sin\left(\frac{1}{x}\right) & x \neq 0 \text{ and } y \neq 0 \\ 0 & \text{elsewhere} \end{cases}$$

Since f is not even continuous at $(0, 0)$ it is quite easy to show that this implies that f cannot be differentiable there. However, both $D_1 f(0, 0)$ and $D_2 f(0, 0)$ exist. The issue here is that $D_1 f$ is not continuous at $(0, 0)$. Indeed, with the added condition of continuity of the partials, we obtain a criterion for the existence of $Df(x)$.

Definition 1.19 (pg.124, H&H)
A function is continuously differentiable on $U \subset \mathbb{R}^n$ if all its partial derivatives exist and are continuous on U.

Theorem 1.20 (pg.31, S)
If $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^m$ is continuously differentiable on U, then f is differentiable on U.

Proof. Suppose first that $m = 1$. If $a \in U$, then for every $1 \leq i \leq n$ the Mean Value Theorem gives us

$$f(a_1 + h_1, \ldots, a_i + h_i, a_{i+1}, \ldots, a_n) - f(a_1 + h_1, \ldots, a_i, \ldots, a_n) = D_i f(c_i) h_i$$

for some $c_i \in (0, h_i)$. Now, we observe

$$\frac{|f(a+h) - f(a) - \sum_{i=1}^n D_i f(a)(h_i)|}{|h|} =$$
$$\frac{|\sum_{i=1}^n f(a_1 + h_1, \ldots, a_i + h_i, a_{i+1}, \ldots, a_n) - f(a_1 + h_1, \ldots, a_i, \ldots, a_n) - D_i f(a)(h_i)|}{|h|}$$

which, by the above, simplifies to

$$\begin{aligned} \frac{|f(a+h) - f(a) - \sum_{i=1}^n D_i f(a)(h_i)|}{|h|} &= \frac{|\sum_{i=1}^n D_i f(c_i) h_i - D_i f(a)(h_i)|}{|h|} \\ &= \frac{|\sum_{i=1}^n (D_i f(c_i) - D_i f(a)) h_i|}{|h|} \\ &\leq \sum_{i=1}^n |D_i f(c_i) - D_i f(a)| \frac{|h_i|}{|h|} \\ &\leq \sum_{i=1}^n |D_i f(c_i) - D_i f(a)| \end{aligned}$$

So that,

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \sum_{i=1}^n D_i f(a)(h_i)|}{|h|} = \lim_{h \to 0} \sum_{i=1}^n |D_i f(c_i) - D_i f(a)| = 0$$

implying the theorem when $m = 1$. If $m > 1$ then the above argument along with Theorem 1.11 (3) implies the result. $\square$

We should note that the converse of this theorem is not necessarily true. That is, we can find a function that is differentiable, but not continuously differentiable. Consider the single variable case when

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We see that if $x \neq 0$ then

$$f'(x) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$$

and

$$f'(0) = \lim_{h \to 0} \frac{h^2 \sin\left(\frac{1}{h}\right)}{h} = 0$$

which shows that $f'(0)$ exists but it is not continuous there and therefore f is not continuously differentiable. Indeed the situation can get much worse. Consider

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

One can show that $f'(0) = \frac{1}{2}$ which implies that the derivative is positive at 0. And yet, the function is not increasing in any neighborhood of 0. This means that the function looks very little like its best linear approximation about 0; knowing the derivative tells us almost nothing about the function at $x = 0$. For this reason, one should maintain caution when studying functions which are not continuously differentiable.

Also, you might notice that we made use of the Mean Value Theorem from single variable analysis in the proof of Theorem 1.20. We now extend this result to the multivariable case.

Theorem 1.21 (pg.120, H&H)
Let $U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ be differentiable, and the segment $[a, b]$ joining a to b be contained in U. Then there exists $c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a)$$

Proof. Notice that, under the assumptions on f, if we define $v = b - a$ and

$$g(t) = f(a + tv)$$

then $g : [0, 1] \to \mathbb{R}$ is differentiable. By the MVT, we get

$$g(1) - g(0) = g'(k)$$

for some $k \in (0, 1)$. If we let $a + kv = c$, then $c \in (a, b)$ and the above equality suggests that

$$\begin{aligned} f(b) - f(a) &= f'(a + kv) \cdot v \\ &= f'(c)(b - a) \end{aligned}$$

$\square$

Chapter 2

Inverse and Implicit Function Theorems

Definition 2.1 (pg.202, H&H)
Let $f : U \to \mathbb{R}$ be a differentiable mapping. The derivative $f'(x)$ satisfies a Lipschitz condition on a subset $V \subset U$ with Lipschitz ratio M if for all $x, y \in V$

$$|f'(x) - f'(y)| \le M\,|x - y|$$

Proposition 2.2 (pg.n/a)
If f satsifies a Lipschitz condition on V then f is continuously differentiable on V.

Proof. Let $y \in V$ and let $\epsilon > 0$ be given. Choose $\delta = \frac{\epsilon}{M}$ so that $|x - y| < \delta$ implies

$$|f'(x) - f'(y)| \le M\,|x - y| < M\delta = \epsilon$$

so that f' is continuous at y and therefore on V. $\square$

Indeed we can see that f' is uniformly continuous where is satisfies a Lipschitz condition (we may also say where f' is Lipschitz). We consider an example:

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2^2 \\ x_1^2 + x_2 \end{pmatrix}$$

so that

$$f'\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 1 & -2x_2 \\ 2x_1 & 1 \end{bmatrix}$$

This implies that

$$f'\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - f'\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} 0 & -2(x_2 - y_2) \\ 2(x_1 - y_1) & 0 \end{bmatrix}$$

Therefore, calculating the length of the above matrix produces

$$\left|\begin{bmatrix} 0 & -2(x_2 - y_2) \\ 2(x_1 - y_1) & 0 \end{bmatrix}\right| = 2\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = 2\left|\begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix}\right|$$

Thus

$$\left|f'\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - f'\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right| = 2\,|x - y|$$

So that, letting $M = 2$ we get that f is Lipschitz on $\mathbb{R}^2$.

It is not always so easy to find the Lipschitz ratio M. There is a way to compute Lipschitz ratios using high partial derivatives.

Definition 2.3 (pg.204, H&H)
Let $U \subset \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}$ be a differentaible function. If the function $D_i f$ is itself differentiable, then its partial derivative with respect to the jth variable,

$$D_j(D_i f)$$

which we may alternatively denote by

$$D_{i,j} f$$

is called a second partial derivative of f.

We denote the class of all continuously differentiable functions by $\mathcal{C}^1$. The class of all k-times continuously differentiable mappings is denoted by $\mathcal{C}^k$. Note that this means that all kth order partial derivatives exist and are continuous.

Proposition 2.4 (pg.204, H&H)
Let $U \subset \mathbb{R}^n$ be open, and $f : U \to \mathbb{R}^n$ be a $\mathcal{C}^2$ mapping. If $|D_k D_j f_i(x)| \leq c_{i,j,k}$ for all $x \in U$ and all triples of indices $1 \leq i, j, k \leq n$, then

$$|f'(u) - f'(v)| \leq \left(\sum_{1 \leq i,j,k \leq n} (c_{i,j,k})^2 \right)^{\frac{1}{2}} |u - v|$$

Proof. Theorem **??**, which we now refer to simply as the Mean Value Theorem, implies the following

$$\begin{aligned}
(D_j f_i(u) - D_j f_i(v))^2 &= |D_j f_i(u) - D_j f_i(v)|^2 \\
&= |[D_1 D_j f_i(s_{ji}) \ldots D_n D_j f_i(s_{ji})] (u - v)|^2 \\
&\leq |[D_1 D_j f_i(s_{ji}) \ldots D_n D_j f_i(s_{ji})]|^2 |u - v|^2 \\
&\leq \left(\sum_k (c_{i,j,k})^2 \right) |u - v|^2
\end{aligned}$$

where s_{ji} is a point on the segment connecting u to v. Thus

$$|f'(u) - f'(v)|^2 = \sum_{i,j} (D_j f_i(u) - D_j f_i(v))^2 \leq \sum_{i,j} \left(\sum_k (c_{i,j,k})^2 \right) |u - v|^2 = \sum_{i,j,k} (c_{i,j,k})^2 |u - v|^2$$

which implies the result. ∎