VECTOR CALCULUS
DIFFERENTIABILITY OF VECTOR FIELDS

PAUL L. BAILEY

ABSTRACT. We have studied three varieties of multivariable functions. Paths are continuous functions defined on an interval, and taking values in higher dimensional space. Scalar fields are functions which take multiple variables, but output a single real number at each point. Vector fields are functions which map vectors to vectors.

We have discussed various aspects of what it means to be differentiable in higher dimensions, but Thomas' otherwise excellent book never formally defines derivatives for vector fields. We remedy that here.

1. TOPOLOGICAL IDEAS IN HIGHER DIMENSIONS

In order to precisely and conveniently define limits in higher dimensions, we first give names to some concepts regarding open sets.

Definition 1. Let $p \in \mathbb{R}^n$. The *ball of radius* ϵ around p is

$$B_\epsilon(p) = \{x \in \mathbb{R}^n \mid |x - p| < \epsilon\},$$

where $|x - p|$ is the distance between x and p.

Let $D \subset \mathbb{R}^n$ and let $x \in \mathbb{R}^n$.

We say that x is an *interior point* of D if there exists $\epsilon > 0$ such that

$$B_\epsilon(x) \subset D.$$

An interior point of D is necessarily in D.

We say that x is a *boundary point* of D if for every $\epsilon > 0$, $B_\epsilon(x)$ contains points that are in D as well as points that are not in D.

Definition 2. Let $U \subset \mathbb{R}^n$.

We say that U is *open* if every point in U is an interior point of U.

Let $p \in \mathbb{R}^n$. A *neighborhood* of p is a set which contains an open set which contains p. A *deleted neighborhood* of p is a set of the form $U \setminus \{p\}$, where U is a neighborhood of p.

Date: March 3, 2020.

2. Limits and Continuity for Vector Fields

Let A and B be sets. The notation $f : A \to B$ means that f is a function from A into B. The notation $f : D \subset A \to B$ means that $D \subset A$ and $f : D \to B$.

Let $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^m$. We would like to define what it means for $\vec{F}$ to be differentiable at a point $\vec{p}_0 \in \mathbb{R}^n$. Since this involves limits, and limits require open sets, we begin there.

Definition 3. Let $\vec{F}$ be defined on a deleted neighborhood U of $\vec{p}_0$ in $\mathbb{R}^n$.

We say that the *limit* of $\vec{F}$ as $\vec{p}$ approaches $\vec{p}_0$ equals $\vec{q} \in \mathbb{R}^n$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that $0 < |\vec{p} - \vec{p}_0| < \delta$ implies $|\vec{F}(\vec{p}) - \vec{q}| < \epsilon$.

If such an $\vec{q}$ exists, we write

$$\lim_{\vec{p} \to \vec{p}_0} \vec{F}(\vec{p}) = \vec{q}.$$

Definition 4. Let $\vec{F}$ be defined on a neighborhood U of $\vec{p}_0$ in $\mathbb{R}^n$.

We say that $\vec{F}$ is *continuous* at $\vec{p}_0$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that $|\vec{p} - \vec{p}_0| < \delta$ implies $|\vec{F}(\vec{p}) - \vec{F}(\vec{p}_0)| < \epsilon$.

Note that by setting $m = n = 1$ above, we have the standard definitions of these concepts from single variable Calculus.

3. Differentiability in Single Variable Calculus

Recall the definition of differentiability for functions $f : \mathbb{R} \to \mathbb{R}$.

Definition 5. Let $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ is an interval. Let $x_0 \in I$ be an interior point. We say that f is *differentiable at* x_0 if the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. Such a limit is called the *derivative* of f at x_0, and is denoted $f'(x_0)$.

Note that by setting $h = x - x_0$, the derivative may be written

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h}.$$

We can create a function from this by letting x_0 vary; to suggest this in the notation, we replace x_0 with x to get

$$f'(x) = \frac{f(x + h) - f(x)}{h}.$$

4. Differentiability of Paths

Definition 6. A *path in* $\mathbb{R}^m$ is a continuous function $\vec{r} : I \to \mathbb{R}^m$, where $I \subset \mathbb{R}$ is an interval. A *curve* in $\mathbb{R}^m$ is the image of a path.

Let $\vec{r} : I \to \mathbb{R}^m$ be a path, and let $t_0 \in I$ be an interior point. We say that $\vec{r}$ is *differentiable* at $\vec{r}$ if

$$\lim_{t \to t_0} \frac{\vec{r}(t) - \vec{r}(t_0)}{t - t_0}$$

exists. Such a limit is called the *derivative* of $\vec{r}$ at t_0, and is denoted $\vec{r}'(t_0)$.

Letting t_0 vary, we replace $t - t_0$ with h, then replace t_0 with t, to obtain

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

We often view a path as a particle in motion; we normally use the variable t to denote the domain, because we are visualizing t to represent time. The derivative of a path at a point is a vector, and is called the *velocity vector* of $\vec{r}$ at time t. Let $\vec{v}(t) = \vec{r}'(t)$. We know that

- $\vec{v}(t)$ is tangent to the curve at the point $\vec{r}(t)$;
- $\vec{v}(t)$ points in the direction of motion;
- $|\vec{v}(t)|$ is the *speed* of the particle.

If $\vec{r} : I \to \mathbb{R}^m$ is a path in $\mathbb{R}^m$, then there exist functions $r_1, \ldots, r_m : I \to \mathbb{R}$ such that

$$\vec{r}(t) = \langle r_1(t), r_2(t), \ldots, r_m(t) \rangle.$$

We call $r_1, \ldots, r_m$ the *coordinate functions* of $\vec{r}$. We have seen that, in order to compute the derivative, we simply compute the vector of coordinate functions:

$$\vec{r}'(t) = \langle r_1'(t), r_2'(t), \ldots, r_m'(t) \rangle.$$

5. Differentiability of Scalar Fields

Definition 7. A *scalar field* is a continuous function $f : D \subset \mathbb{R}^n \to \mathbb{R}$.

We have discussed these for $n = 2$ and $n = 3$, using coordinate systems (x, y) and (x, y, z). In order to take this into arbitrarily higher dimensions, we need to reassign our notation. Thus let $x_1, \dots, x_m$ denote the distance from the origin along the coordinates axes in $\mathbb{R}^n$; we use these as coordinate variables, and say our coordinate system is $\vec{x} = (x_1, x_2, \dots, x_n)$.

Definition 8. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field, and let $\vec{x} = \langle x_1, x_2, \dots, x_n \rangle \in D$ be an interior point. The *partial derivative* of f at $\vec{x}$ in the direction x_i is

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \dots, x_i + h, \dots, x_n) - f(x_1, \dots, x_n)}{h},$$

where $x_i + h$ is in the i^{th} slot.

The partial derivative is the rate of change of f in the direction of the x_i axis. To visualize it, fix all variables except x_i, to create a function $f_i : D_i \to \mathbb{R}$ given by $f_i(x) = f(x_1, \dots, x, \dots, x_n)$, and $D_i = \{x \in \mathbb{R} \mid (x_1, \dots, x, \dots, x_n) \in D\}$, where x is in the i^{th} slot. Then in the graph of f_i, the partial derivative is the slope of the tangent line at the appropriate point.

Definition 9. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field, and let $\vec{x} = \langle x_1, x_2, \dots, x_n \rangle \in D$ be an interior point. Let $\vec{v} \in \mathbb{R}^n$. The *directional derivative* of f is the direction $\vec{v}$ at the point $\vec{x}$ is

$$D_{\vec{v}}(\vec{x}) = \lim_{h \to 0} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h},$$

where $\vec{u} = \dfrac{\vec{v}}{|\vec{v}|}$ is a unit vector in the direction of $\vec{v}$.

The directional derivative is the rate of change of f in the direction of $\vec{v}$.

Definition 10. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field.
The *gradient* of f is the function

$$\nabla f : D \to \mathbb{R}^n \quad \text{given by} \quad \nabla f = \left\langle \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\rangle.$$

Since each of the components of the gradient, $\dfrac{\partial f}{\partial x_i}$, is a function defined on the interior of the domain of f, the vector of these components is actually a vector field. Let $\vec{x} = \langle x_1, x_2, \dots, x_n \rangle \in D$. Then gradient of f at $\vec{x}$ is

$$\nabla f(\vec{x}) = \left\langle \frac{\partial f}{\partial x_1}(\vec{x}), \dots, \frac{\partial f}{\partial x_n}(\vec{x}) \right\rangle,$$

which is a vector whose dimension is that of the domain of f.

If we plug the gradient into the directional derivative, we maximize it. Thus the gradient vector lives in the domain, and points in the direction of maximum increase of f.

6. Geometric Interpretation of Derivatives

Differentiation is the process of estimating a function with a linear function; a function is differentiable if it is approximately linear.

In single variable calculus, we called a function linear if it was of the form

$$L(x) = mx + b.$$

Then, given a function f which is differentiable at x_0, we claim that f is approximated by its "linearization",

$$L(x) = f'(x_0)(x - x_0) + f(x_0).$$

In this section, we wish to focus on linearity in the sense of linear algebra, as follows.

Definition 11. Let $T : \mathbb{R}^n \to \mathbb{R}^m$. We say that T is *linear* if

(T1) $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$, for all $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$;
(T2) $T(a\vec{v}) = aT(\vec{v})$, for all $\vec{v} \in \mathbb{R}^n$ and all $a \in \mathbb{R}$.

It can be shown that a linear function maps the origin to the origin. This leads to the first and most important change in point of view between differentiation in one dimension and in higher dimensions.

In the case that $f : I \subset \mathbb{R} \to \mathbb{R}$, we saw the derivative as the slope of the tangent line. Instead, let us translate the point (x_0, y_0), where $y_0 = f(x_0)$, to the origin, and consider how the function $f(x + x_0) - y_0$ approximates a linear function at the origin.

Let $T : \mathbb{R} \to \mathbb{R}$ be linear, and let $T(1) = m$. Then for any $x \in \mathbb{R}$, **(T2)** implies $T(x) = T(x \cdot 1) = xT(1) = xm = mx$. That is, T is simply multiplication by $T(1)$. So to say that f is approximately linear means that it is approximated by a function which multiplies everything by the same number. The single variable definition we used was

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

By properties of limits, this is the same as

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

If we set $m = f'(x_0)$ and $T(x) = mx$, then $T(h) = mh = hf'(x)$, so the above is equivalent to

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

7. Differentiability of Vector Fields

Definition 12. Let $\vec{F} : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Let $\vec{p}_0$ be an interior point of D. We say that $\vec{F}$ is *differentiable at* $\vec{p}_0$ if there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\vec{h} \to 0} \frac{|\vec{F}(\vec{p}_0 + \vec{h}) - \vec{F}(\vec{p}_0) - T(\vec{h})|}{|\vec{h}|} = 0.$$

If such a T exists, it is unique, and is called the *derivative* of $\vec{F}$ at $\vec{p}_0$.

Let $\vec{p}_0$ be a point in $\mathbb{R}^n$. In order to visualize a linear transformation which behaves "locally" near $\vec{p}_0$, we consider the "tangent space" of $\mathbb{R}^n$ at $\vec{p}_0$. This is the set all vectors in $\mathbb{R}^n$, translated so that their tails are situated at $\vec{p}_0$; then if $\vec{p}_0$ is an interior point in the domain of a differentiable vector field $\vec{F}$ and $\vec{q}_0 = \vec{F}(\vec{p}_0)$, we view T as a linear transformation from the tangent space at $\vec{p}_0$ to the tangent space at $\vec{q}_0$.

There are several ways to formally define the notion of tangent space, the most useful for us being the following.

Definition 13. Let $\vec{p}_0 \in \mathbb{R}^n$. Consider the set of all paths in $\mathbb{R}^n$ which are differentiable at $\vec{p}_0$. We say that two of these paths are *equivalent* if their velocity vectors are equal at $\vec{p}_0$.

A *tangent vector* to $\vec{p}_0$ is an equivalence class of paths. The *tangent space* at $\vec{p}_0$ is the set of tangent vectors. This is identified with a copy of $\mathbb{R}^n$.

Each path $\vec{r}$ through $\vec{p}_0$ is mapped by $\vec{F}$ to a path through $\vec{q}_0 = \vec{F}(\vec{p}_0)$, via composition. Say that $\vec{r}(t_0) = \vec{p}_0$. Then $(\vec{F} \circ \vec{r})(t_0) = \vec{q}_0$ is a path through $\vec{q}_0$, and we view T as mapping $\vec{r}'(t_0)$ to $(\vec{F} \circ \vec{r})'(t_0)$. If two paths are equivalent, this process will map them to equivalent tangent vectors in the range.

From this perspective, it is possible to compute the matrix of T. The j^{th} standard basis vector for the tangent space of $\vec{p}_0 = (p_1, \ldots, p_n)$ are represented by path of the form

$$\vec{r}_j(t) = \langle p_1, \ldots, p_j + t, \ldots, p_n \rangle = \vec{p} + t\vec{e}_j.$$

The corresponding tangent vector is $\vec{r}'(t) = \vec{e}_j$. Let $\vec{F} = \langle F_1, \ldots, F_m \rangle$, where $F_i : \mathbb{R}^n \to \mathbb{R}$ is the i^{th} component function of $\vec{F}$. Recall that if $g : \mathbb{R}^n \to \mathbb{R}$, then the derivative of the composition is given by the chain rule:

$$\frac{d}{dt}(g \circ \vec{r}) = \sum_{j=1}^{n} \frac{\partial g}{\partial x_j} \frac{dx_j}{dt},$$

where $\vec{r}(t) = \langle x_1(t), \ldots, x_n(t) \rangle$. Applied to our case,

$$\frac{d}{dt}(F_i \circ \vec{r}_j) = \frac{\partial F_i}{\partial x_j} \frac{dx_j}{dt},$$

since the rest of the summands in the chain rule are zero. Thus

$$\vec{e}_j \quad \mapsto \quad \left\langle \frac{\partial F_1}{\partial x_j}, \frac{\partial F_2}{\partial x_j}, \ldots, \frac{\partial F_m}{\partial x_j} \right\rangle.$$

We put these vectors into the columns of a matrix to see that the matrix of T is

$$A_T = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix}$$

Example 1. Let $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$ by given by $\vec{F}(x, y) = x^2 - y^2, 2xy$. Find the derivative of $\vec{F}$, and use it to estimate the level of area distortion caused by $\vec{F}$ at the point $(1, 2)$. Also, estimate $\vec{F}(1.23, 2.34)$.

Solution. The derivative is the transformation given by the matrix of partials, evaluated at the point $(1, 2)$, which is

$$A = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{bmatrix} \Bigg|_{(1,2)} = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix} \Bigg|_{(1,2)} = \begin{bmatrix} 2 & -4 \\ 4 & 2 \end{bmatrix}.$$

The area distortion is $\det A = 4 + 8 = 12$. Locally, $\vec{F}$ multiplies area by the factor of approximately 12 near the point $(1, 2)$.

Note that $\vec{F}(1, 2) = (-3, 4)$. Now consider the small vector $\vec{h} = \langle 0.23, 0.34 \rangle$, residing in the tangent space at $(1, 2)$. Then $A\vec{h} = \langle 2(0.23) - 4(0.34), 4(0.23) + 2(0.34) \rangle = \langle -0.9, 1.6 \rangle$ is a tangent vector at $(-3, 4)$, which indicates how far from $(-3, 4)$ is $\vec{F}(1, 2)$. That is,

$$\vec{F}(1.23, 2.34) \approx \vec{F}(1, 2) + A\vec{h} = (-3, 4) + \langle -0.9, 1.6 \rangle = (-3.9, 5.6).$$

□

8. The Derivative is a Linear Transformation

Let us rederive the Jacobian matrix from the perspective of an arbitrary tangent vector.

Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Then $F = \langle F_1, F_2, \dots, F_m \rangle$, where $F_j : D \to \mathbb{R}$ is the j^{th} component function of F.

Let $\vec{p} \in D$ be an interior point. Let $\vec{q} = F(\vec{p})$. Let $\alpha : \mathbb{R} \to \mathbb{R}^n$ for a path in $\mathbb{R}^n$ with $\alpha(t_0) = \vec{p}$; the equivalence class of α is a tangent vector at $\vec{p}$. The derivative of F maps $\alpha'(t_0)$ to a tangent vector at $\vec{q}$ given as $(F \circ \alpha)'(t_0)$.

Let $x_1, \dots, x_n$ be variables for $\mathbb{R}^n$. Now

$$\alpha' = \frac{d}{dt}\,\alpha = \left\langle \frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right\rangle,$$

and

$$(F \circ \alpha)' = \left\langle \frac{d}{dt}\,F_1 \circ \alpha, \dots, \frac{d}{dt}\,F_n \circ \alpha \right\rangle.$$

From the change rule,

$$\frac{d}{dt}\,F_i \circ \alpha = \frac{\partial F_i}{\partial x_1}\frac{x_1}{dt} + \dots + \frac{\partial F_i}{\partial x_n}\frac{x_n}{dt} = \nabla F_i \cdot \alpha',$$

so we obtain

$$(F \circ \alpha)' = \langle \nabla F_1 \cdot \alpha', \dots, \nabla F_m \cdot \alpha' \rangle.$$

Thus we have a well-defined mapping T from the tangent space at $\vec{p}$ to the tangent space at $\vec{q}$, which may be rewritten as a matrix product:

$$T(\alpha') = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} \alpha' \end{bmatrix} = \begin{bmatrix} \nabla F_1 \\ \nabla F_2 \\ \vdots \\ \nabla F_m \end{bmatrix} \begin{bmatrix} \alpha' \end{bmatrix}.$$

BASIS Scottsdale
Email address: `paul.bailey@basised.com`