18.024 SPRING OF 2008
DER. DERIVATIVES OF VECTOR FIELDS

Derivatives of vector fields. Derivative theory for vector fields is a straightforward extension of that for scalar fields. Given $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ a vector field, in components, $\mathbf{f}$ consists of m scalar fields of n variables. That is,

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})),$$

and each $f_i : D \subset \mathbb{R}^n \to \mathbb{R}$ is a scalar field, where $i = 1, \dots, m$. It seems natural to define the differentiability of $\mathbf{f}$, thus, componentwise.

We recall that if $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is a scalar field and $\mathbf{a} \in \mathrm{int(D)}$ then the *derivative* of f at $\mathbf{a}$ (if exists) is given by the gradient

$$\nabla f(\mathbf{a}) = (D_1 f(\mathbf{a}, D_2 f(\mathbf{a}), \dots, D_n f(\mathbf{a})).$$

For our purposes, it is convenient to understand the derivative of f as a *row matrix* rather than as a vector. We will write as

$$Df(\mathbf{a}) = [D_1 f(\mathbf{a}, D_2 f(\mathbf{a}), \dots, D_n f(\mathbf{a})].$$

Then, f is differentiable at $\mathbf{a}$ (with the total derivative $Df(\mathbf{a})$ if

(*) $$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = Df(\mathbf{a})\mathbf{h} + \|\mathbf{h}\| E(\mathbf{h}; \mathbf{a}),$$

and $|E(\mathbf{h}; \mathbf{a})| \to 0$ as $\|\mathbf{h}\| \to 0$. Here the first term on the right side is matrix multiplication of $Df(\mathbf{a})$ and the column matrix

$$\mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}.$$

Definition 1. Let $\mathbf{f} : D \in \mathbb{R}^m$ and $D \subset \mathbb{R}^n$ be open. In scalar form,

$$\mathbf{f}(\mathbf{x}) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)).$$

We say that f is *differentiable* at $\mathbf{a} \in D$ if each of the component functions $f_1, \dots, f_m$ is differentiable at $\mathbf{a}$ (in the sense discussed in [Apo, Section 8.11] or above).

Furthermore, by the *derivative* of $\mathbf{f}$ at $\mathbf{a}$ we mean the matrix

$$D\mathbf{f}(\mathbf{a}) = \begin{bmatrix} D_1 f_1(\mathbf{a}) & D_2 f_1(\mathbf{a}) & \cdots & D_n f_1(\mathbf{a}) \\ D_1 f_2(\mathbf{a}) & D_2 f_2(\mathbf{a}) & \cdots & D_n f_2 \mathbf{a}) \\ \vdots & \vdots & & \vdots \\ D_1 f_m(\mathbf{a}) & D_2 f_m(\mathbf{a}) & \cdots & D_n f_m \end{bmatrix}.$$

In other words, $D\mathbf{f}(\mathbf{a})$ is the matrix whose ith row is given as the derivative $Df_i(\mathbf{a})$ of the ith coordinate function of $\mathbf{f}$. It is often called the *jacobian matrix** of $\mathbf{f}$. Another notation for this matrix is

$$\frac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_n)}.$$

*It is named after the Prussian mathematician Carl Gustav Jacob Jacobi (1804-1851).

In [Apo, Section 8.18], the differentiability of a vector field $\mathbf{f}$ is given, alternatively, by directly extending the Taylor's formula (*) in the vector-field setting. Our definition then is equivalent to that given in [Apo], which is stated below.

Theorem 2. *The vector field* $\mathbf{f}$ *is differentiable at* $\mathbf{a}$ *if and only if*

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = D\mathbf{f}(\mathbf{a})\mathbf{h} + \|\mathbf{h}\|\mathbf{E}(\mathbf{h};\mathbf{a})$$

and $\|\mathbf{E}(\mathbf{h};\mathbf{a})\| \to 0$ *as* $\|\mathbf{h}\| \to 0$. *Here,* $\mathbf{f}, \mathbf{h}$ *and* $\mathbf{E}$ *are written as column matrices.*

Proof. We observe that both sides of the equation represent column matrices. Considering the ith entries of these matrices, we obtain the following equation

$$f_i(\mathbf{a}+\mathbf{h}) - f_i(\mathbf{a}) = Df_i)(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\|\mathbf{E}_i(\mathbf{h};\mathbf{a}).$$

Now if $\mathbf{f}$ is differentiable at $\mathbf{a}$ if and only if each f_i is, and f_i is differentiable at $\mathbf{a}$ if and only if $|\mathbf{E}_i(\mathbf{h};\mathbf{a})| \to 0$ as $\|\mathbf{h}\| \to 0$. But, $|\mathbf{E}_i(\mathbf{h})| \to 0$ as $\|\mathbf{h}\| \to 0$ for each i if and only if $\|\mathbf{E}(\mathbf{h})\| \to 0$ as $\|\mathbf{h}\| \to 0$, and we are done. □

The following result is immediate from our definition of differenentiability.

Theorem 3. *If* $\mathbf{f}$ *is differentiable at* $\mathbf{a}$ *then* $\mathbf{f}$ *is continuous at* $\mathbf{a}$.

The chain rule for derivatives of vector fields. Before considering the chain rule for vector fields, let us take the chain rule for scalar fields we have already proved [Apo, Section 8.15] and reformulate it in terms of matrices.

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field defined in an open ball about $\mathbf{a}$, and let $\mathbf{x} : I \subset \mathbb{R} \to \mathbb{R}^n$ be a vector-valued function defined in an open interval about t_0. Let $\mathbf{x}(t_0) = \mathbf{a}$ and $\mathbf{x}(I) \subset D$. If f is differentiable at $\mathbf{a}$ and if $\mathbf{x}$ is differentiable at t_0, then we have shown in Theorem 8.8 of [Apo] that $f \circ \mathbf{x}$ is differentiable at t_0 and its derivative is given by

$$\frac{d}{dt} f(\mathbf{x}(t_0)) = \nabla f(\mathbf{x}(t_0)) \cdot \mathbf{x}'(t_0).$$

We now write this formula in scalar form as

$$\frac{d}{dt}(f \circ \mathbf{x}) = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt};$$

or, we write it in matrix form as

$$\frac{d}{dt}(f \circ \mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{bmatrix}.$$

Recalling the definition of the Jacobian matrix Df the latter formula is recognized as

$$\frac{d}{dt}(f(\mathbf{x}(t)) = Df(\mathbf{x}(t))D\mathbf{x}(t).$$

(Note that Df is a row matrix while $D\mathbf{x}$ is by definition a column matrix.) This is the form of the chain rule which is useful to extend to vector fields.

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^p$ and $\mathbf{g} : \mathbb{R}^p \to \mathbb{R}^m$ be vector fields such that the composition field $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ is defined in an open ball about $\mathbf{a} \in \mathbb{R}^n$. Let $\mathbf{f}(\mathbf{a}) = \mathbf{b} \in \mathbb{R}^p$. We write these fields as

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, \ldots, x_n), \qquad \mathbf{g}(\mathbf{y}) = \mathbf{g}(y_1, \ldots, y_p).$$

If $\mathbf{f}$ and $\mathbf{g}$ are differentiable at $\mathbf{a}$ and $\mathbf{b}$, respectively, then partial derivatives of $\mathbf{h}$ exist at $\mathbf{a}$ by applying the chain rule for each $h_i(\mathbf{x}) = g_i(\mathbf{f}(\mathbf{x}))$ as

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial g_i}{\partial y_1}\frac{\partial f_1}{\partial x_j} + \cdots + \frac{\partial g_i}{\partial y_p}\frac{\partial f_p}{\partial x_j}.$$

This is recognized as matrix multiplication

$$[D_1 g_i \; D_2 g_i \; \cdots \; D_p g_i] \begin{bmatrix} D_j f_1 \\ \vdots \\ D_j f_p \end{bmatrix}.$$

In other words, its multiplication of the ith row of $D\mathbf{g}$ and the jth column of $D\mathbf{f}$. Thus, the Jacobian matrix of $\mathbf{h}$ is expected to satisfy the matrix equation

$$D\mathbf{h}(\mathbf{a}) = D\mathbf{g}(\mathbf{b})D\mathbf{f}(\mathbf{a}).$$

Not exactly. If $\mathbf{f}$ and $\mathbf{g}$ are differentiable then we know that the partial derivatives of the composite function $\mathbf{h}$ exist. But, the mere *existence* of the partial derivatives of the function h_i does not imply the differentiability of h_i, and thus, one needs to provide a separate argument that $\mathbf{h}$ is differentiable.

One may avoid giving a separate proof of the differentiability of $\mathbf{h}$ by assuming a stronger hypothesis, namely, that both $\mathbf{f}$ and $\mathbf{g}$ are *continuously differentiable*. Then, the formula,

$$D_j h_i(\mathbf{x}) = \sum_{l=1}^{p} D_l g_i(\mathbf{f}(\mathbf{x})) D_j f_l(\mathbf{x}),$$

which we have proved, shows that $D_j h_i$ is continuous, that is, $\mathbf{h}$ is continuously differentiable. Therefore, $\mathbf{h}$ is differentiable.

We summarize these.

Theorem 4. *Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^p$ and $\mathbf{g} : \mathbb{R}^p \to \mathbb{R}^m$ be vector fields defined in open balls about $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^p$, respectively. Let $\mathbf{f}(\mathbf{a}) = \mathbf{b}$. If $\mathbf{f}$ and $\mathbf{g}$ are continuously differentiable on their respective domains, then $\mathbf{h}(\mathbf{x}) = \mathbf{g} \circ \mathbf{f}(\mathbf{x})$ is continuously differentiable on its domain and*

$$D\mathbf{h}(\mathbf{x}) = D\mathbf{g}(\mathbf{f}(\mathbf{x}) \cdot D\mathbf{f}(\mathbf{x}).$$

Most applications of the chain rule we will use in future are in the C^1-class setting, and thus this theorem will suffice. But, one may remove the C^1 condition. Please read the statement and its proof of [Apo, Theorem 8.11].

Example 5 (Polar coordinates)**.** Given a function $f(x, y)$ defined on the (x, y) plane, we introduce polar coordinates

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$

and f becomes a function of r and θ as $\phi(r, \theta) = f(r\cos\theta, r\sin\theta)$. Then,

$$\frac{\partial \phi}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

and

$$\frac{\partial \phi}{\partial \theta} = -r\frac{\partial f}{\partial x}\sin\theta + r\frac{\partial f}{\partial y}\cos\theta.$$

The second derivative is computed as

$$\begin{aligned}
\frac{\partial^2\phi}{\partial r^2} &= \frac{\partial}{\partial r}\left(\frac{\partial \phi}{\partial r}\right) = \frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right) \\
&= \frac{\partial}{\partial r}\left(\frac{\partial f}{\partial x}\right)\cos\theta + \frac{\partial f}{\partial x}\frac{\partial(\cos\theta)}{\partial r} + \frac{\partial}{\partial r}\left(\frac{\partial f}{\partial y}\right)\sin\theta + \frac{\partial f}{\partial y}\frac{\partial(\sin\theta)}{\partial r} \\
&= \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\cos\theta + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\sin\theta\right)\cos\theta + \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\cos\theta + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\sin\theta\right)\sin\theta \\
&= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial x \partial y}\sin\theta\cos\theta.
\end{aligned}$$

An analogous formula for $\partial^2\phi/\partial\theta^2$ is worked in [Apo, Example 3] and $\partial^2\phi/\partial r\partial\theta$ is in [Apo, Section 8.22, Exercise 5].

Derivatives of inverses. Recall that if $f(x)$ is a differentiable real-valued function of a single variable x, and if $f'(x) > 0$ for $x \in [a,b]$ (or $f'(x) < 0$), then f is strictly increasing (or decreasing), and it has an inverse, say g. Furthermore, g is differentiable and its derivative satisfies

$$g'(f(x)) = \frac{1}{f'(x)}.$$

Part of this theory extends to vector fields.

Theorem 6. *Let* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ *be defined in an open ball of* $\mathbf{a}$ *and* $\mathbf{f}(\mathbf{a}) = \mathbf{b}$. *Suppose that* $\mathbf{f}$ *has an inverse, say* $\mathbf{g}$.

If $\mathbf{f}$ *is differentiable at* $\mathbf{a}$ *and if* $\mathbf{g}$ *is differentiable at* $\mathbf{b}$ *then*

$$D\mathbf{g}(\mathbf{b}) = (D\mathbf{f}(\mathbf{a}))^{-1}.$$

In interpretation, the jacobian matrix of the inverse function is just the inverse of the jacobian matrix.

Proof. We observe that $\mathbf{g} \circ \mathbf{f}(\mathbf{x}) = \mathbf{x}$. Since $\mathbf{f}$ and $\mathbf{g}$ are differentiable and so is the composite function $\mathbf{g} \circ \mathbf{f}$, by the chain rule, it follows that

$$D\mathbf{g}(\mathbf{b})D\mathbf{f}(\mathbf{a}) = I_n,$$

which implies the assertion. □

The theorem shows that in order for the differentiable function $\mathbf{f}$ to have a differentiable inverse, it is *necessary* that the Jacobian matrix $D\mathbf{f}$ have rank n. Roughly speaking, it is also *sufficient* for $\mathbf{f}$ to have an inverse. More precisely, one has the following *Inverse Function Theorem*, which is one of the fundamental results in mathematical analysis.

Theorem 7 (Inverse function Theorem)**.** *If* $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ *is of* C^1 *in a neighborhood of* $\mathbf{a}$ *and* $D\mathbf{f}(\mathbf{a})$ *has rank* n*, then there is an open ball* $B \in \mathbb{R}^n$ *about* $\mathbf{a}$ *such that* $\mathbf{f} : B \to C = \mathbf{f}(B)$ *is one-to-one and onto. Furthermore, the inverse function* $\mathbf{g} : C \to B$ *of* $\mathbf{f}$ *is continuously differentiable and* $D\mathbf{g}(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}$.

The proof and further application of the inverse function theorem is learned in 18.100A.

Example 8 (Invertibility of the polar coordinate transformation)**.** Let

$$\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2, \qquad \mathbf{f}(r,\theta) = (x,y)$$

be the differentiable transformation from the (x,y)-plane to (r,θ) plane by polar coordinates:

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$

Its Jacobian matrix is computed as

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Note that its determinant is r and thus, it is of rank 2 provided that $r \neq 0$. According to the inverse function theorem,thus, away from the origin ($r = 0$), the polar coordinate transformation $\mathbf{f}$ is invertible locally, and its inverse is given by $\mathbf{g}(x,y) = (r,\theta)$. Moreover,

$$\frac{\partial(r,\theta)}{\partial(x,y)} = \frac{\partial(x,y)}{\partial(r,\theta)}^{-1} = \frac{1}{r}\begin{bmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

REFERENCES

[Apo] T. Apostol, *Calculus*, vol. II, Second edition, Wiley, 1967.

©2008 BY VERA MIKYOUNG HUR
E-mail address: verahur@math.mit.edu