Derivatives of Vector-Valued Functions

Outline

1. **Components**

Consider a function general vector-valued function $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$. Such a function can be written as

$$\mathbf{f}(x_1,\ldots,x_m) \;=\; \big(f_1(x_1,\ldots,x_m),\ldots,f_n(x_1,\ldots,x_m)\big),$$

where each $f_i\colon \mathbb{R}^m \to \mathbb{R}$. The real-valued functions $f_1,\ldots,f_n$ are called the **components** of $\mathbf{f}$.

For example, if $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^3$ is the function

$$\mathbf{f}(x,y) \;=\; (x^2+y^2, x^2-y^2, 2xy)$$

then $\mathbf{f}$ has components $f_1(x,y) = x^2+y^2$, $f_2(x,y) = x^2-y^2$, and $f_3(x,y) = 2xy$.

2. **Partial Derivatives**

If $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$, the **partial derivative** of $\mathbf{f}$ with respect to x_i is the vector

$$\frac{\partial \mathbf{f}}{\partial x_i} \;=\; \left(\frac{\partial f_1}{\partial x_i},\,\ldots,\,\frac{\partial f_n}{\partial x_i}\right).$$

We sometimes use subscripts to denote partial derivatives. For example, if $\mathbf{f}(x,y,z)$ is vector-valued function on $\mathbb{R}^3$, then $\mathbf{f}_y$ would denote the partial derivative of $\mathbf{f}$ with respect to y.

3. **Parameter Curves**

Given a function $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$, a **parameter curve** for $\mathbf{f}$ is a curve $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^n$ of the form

$$\begin{aligned}\boldsymbol{\gamma}(t) \;&=\; \mathbf{f}(\mathbf{p} + t\mathbf{e}_i) \\ &=\; \mathbf{f}(p_1,\ldots,p_{i-1},p_i+t,p_{i+1},\ldots,p_m).\end{aligned}$$

where $\mathbf{p}$ is a point in $\mathbb{R}^m$, and $\mathbf{e}_i$ is a unit vector in the x_i direction. Note that $\boldsymbol{\gamma}$ is obtained obtained from $\mathbf{f}(x_1,\ldots,x_m)$ by varying x_i and holding the other x's constant. Thus parameter curves can be thought of as the images of the "gridlines" under the function $\mathbf{f}$.

The partial derivatives of a function $\mathbf{f}$ at a point $\mathbf{p}$ can be interpreted as the tangent vectors to the parameter curves through $\mathbf{f}(\mathbf{p})$. Specifically, if $\boldsymbol{\gamma}$ is the parameter curve defined above, then

$$\dot{\boldsymbol{\gamma}}(0) \;=\; \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{p}).$$

4. **The Derivative Matrix**

If $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$ the **derivative** of $\mathbf{f}$ at a point $\mathbf{p}$ is the matrix

$$D_{\mathbf{p}}\mathbf{f} \;=\; \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{p}) & \cdots & \dfrac{\partial f_1}{\partial x_m}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(\mathbf{p}) & \cdots & \dfrac{\partial f_n}{\partial x_m}(\mathbf{p}) \end{bmatrix}$$

That is, $D_{\mathbf{p}}\mathbf{f}$ is the matrix whose columns are the partial derivatives $\dfrac{\partial \mathbf{f}}{\partial x_1}(\mathbf{p}), \ldots, \dfrac{\partial \mathbf{f}}{\partial x_m}(\mathbf{p})$.

For example, if $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^3$ is the function $\mathbf{f}(x,y) = (x^2+y^2, x^2-y^2, 2xy)$, then

$$D_{(x,y)}\mathbf{f} \;=\; \begin{bmatrix} 2x & 2y \\ 2x & -2y \\ 2y & 2x \end{bmatrix}$$

Note that the derivative is actually a function that takes a point (x,y) as input and outputs a matrix of numbers.

In some books, the derivative $D_{\mathbf{p}}\mathbf{f}$ is denoted $D\mathbf{f}_{\mathbf{p}}$, $D\mathbf{f}(\mathbf{p})$, $d\mathbf{f}_{\mathbf{p}}$, or $\mathbf{f}'(\mathbf{p})$. In addition, the derivative is sometimes referred to as the **Jacobian**, in which case it may be denoted with a J instead of a D. We will use the notation $D_{\mathbf{p}}\mathbf{f}$ to mean the derivative of $\mathbf{f}$ at $\mathbf{p}$, and $D_{\mathbf{p}}\mathbf{f}(\mathbf{v})$ to mean the product of the matrix $D_{\mathbf{p}}\mathbf{f}$ with a vector $\mathbf{v}$.

5. **Special Cases of the Derivative**

For a parametric curve $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^n$, the derivative of $\boldsymbol{\gamma}$ is the same as the tangent vector $\dot{\boldsymbol{\gamma}}$:

$$D_t\boldsymbol{\gamma} \;=\; \dot{\boldsymbol{\gamma}}(t) \;=\; \begin{bmatrix} \dot{\gamma}_1(t) \\ \vdots \\ \dot{\gamma}_n(t) \end{bmatrix}.$$

For a real-valued function $f\colon \mathbb{R}^m \to \mathbb{R}$, the derivative $D_{\mathbf{p}}f$ is the *transpose* of the gradient vector ∇f:

$$D_{\mathbf{p}}f \;=\; \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{p}) & \cdots & \dfrac{\partial f}{\partial x_m}(\mathbf{p}) \end{bmatrix} \;=\; \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{p}) \\ \vdots \\ \dfrac{\partial f}{\partial x_m}(\mathbf{p}) \end{bmatrix}^T \;=\; \nabla f(\mathbf{p})^T.$$

6. **Directional Derivatives**

Let $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$ be a differentiable function, and let $\mathbf{e}_i$ be a unit vector in the x_i direction. Then:

$$D_{\mathbf{p}}\mathbf{f}(\mathbf{e}_i) \;=\; \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{p}).$$

That is, the i'th column of $D_{\mathbf{p}}\mathbf{f}$ is the partial derivative on the right.

More generally, if $\mathbf{v}$ is any vector in $\mathbb{R}^m$, then the product

$$D_{\mathbf{p}}\mathbf{f}(\mathbf{v})$$

is called the **directional derivative** of $\mathbf{f}$ in the direction of $\mathbf{v}$. This is something like a "partial derivative" in the direction of the vector $\mathbf{v}$.

The directional derivative $D_{\mathbf{p}}(\mathbf{v})$ can be interpreted as a tangent vector to a certain parametric curve. Specifically, let $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^n$ be the curve

$$\boldsymbol{\gamma}(t) \;=\; \mathbf{f}(\mathbf{p} + t\mathbf{v}).$$

That is, $\boldsymbol{\gamma}$ is the image under $\mathbf{f}$ of a straight line in the direction of $\mathbf{v}$. Then

$$\dot{\boldsymbol{\gamma}}(0) \;=\; D_{\mathbf{p}}\mathbf{f}(\mathbf{v}).$$

7. Differentials

The derivative of a function $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$ can also be thought of in terms of **differentials**. Specifically, let $\mathbf{p}$ be a point in $\mathbb{R}^m$, with corresponding value $\mathbf{f}(\mathbf{p})$. Now, suppose we move from $\mathbf{p}$ to a nearby point $\mathbf{p} + \mathbf{dp}$, and let $\mathbf{df}$ denote the corresponding change in the value of $\mathbf{f}$

$$\mathbf{df} \;=\; \mathbf{f}(\mathbf{p} + \mathbf{dp}) - \mathbf{f}(\mathbf{p}).$$

Then the vectors $\mathbf{df}$ and $\mathbf{dp}$ are related by the formula

$$\mathbf{df} \;\approx\; D_{\mathbf{p}}\mathbf{f}(\mathbf{dp}).$$

That is, the change in $\mathbf{f}$ is roughly the product of the matrix $D_{\mathbf{p}}\mathbf{f}$ with the vector $\mathbf{dp}$. Note that, since we cannot divide vectors, we cannot interpret $D_{\mathbf{p}}\mathbf{f}$ as the "ratio" of $\mathbf{df}$ to $\mathbf{dp}$.

8. The Chain Rule

Suppose we have a pair of differentiable functions $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{g}\colon \mathbb{R}^k \to \mathbb{R}^m$. Since the codomain of $\mathbf{g}$ is the same as the domain of $\mathbf{f}$, we can form the composition $\mathbf{f} \circ \mathbf{g}\colon \mathbb{R}^k \to \mathbb{R}^n$.

In this case, the **Chain Rule** gives us a formula for the derivative of $\mathbf{f} \circ \mathbf{g}$. According to the rule, if $\mathbf{p} \in \mathbb{R}^k$ and $\mathbf{q} = \mathbf{g}(\mathbf{p})$, then

$$D_{\mathbf{p}}(\mathbf{f} \circ \mathbf{g}) \;=\; (D_{\mathbf{q}}\mathbf{f})(D_{\mathbf{p}}\mathbf{g}).$$

That is, the derivative matrix for $\mathbf{f} \circ \mathbf{g}$ at $\mathbf{p}$ is the product of the derivative matrix for $\mathbf{f}$ at $\mathbf{q}$ and the derivative matrix for $\mathbf{g}$ at $\mathbf{p}$. This is simply a matrix form of the Chain Rule for partial derivatives.

As a special case, let $\mathbf{f}\colon \mathbb{R}^m \to \mathbb{R}^n$, let $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^m$ is a parametric curve in $\mathbb{R}^m$, and let $\mathbf{p} = \boldsymbol{\gamma}(0)$. Then the composition $\tilde{\boldsymbol{\gamma}} = \mathbf{f} \circ \boldsymbol{\gamma}$ is a parametric curve in $\mathbb{R}^n$, and

$$\dot{\tilde{\boldsymbol{\gamma}}}(0) \;=\; D_{\mathbf{p}}\mathbf{f}\big(\dot{\boldsymbol{\gamma}}(0)\big)$$

That is, the tangent vector to $\tilde{\boldsymbol{\gamma}}$ is the product of the derivative matrix $D_{\mathbf{p}}\mathbf{f}$ with the tangent vector to $\boldsymbol{\gamma}$.

Practice Problems

1. Let $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the polar coordinates transformation $\mathbf{f}(r,\theta) = (r\cos\theta, r\sin\theta)$.

   (a) Make a drawing showing the parameter curves $\theta = C$ for $C \in \{0, \pi/4, \pi/2, 3\pi/4, \pi\}$, as well as the curves $r = C$ for $C \in \{1,2,3\}$.

   (b) Compute $D_{(2,3\pi/4)}\mathbf{f}$. Add the the vectors $\mathbf{f}_r(2,3\pi/4)$ and $\mathbf{f}_\theta(2,3\pi/4)$ to your drawing as tangent vectors to parameter curves.

   (c) Let $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^2$ be a regular curve, and suppose that $\boldsymbol{\gamma}(0) = (2,3\pi/4)$ and $\dot{\boldsymbol{\gamma}}(0) = (3,1)$. Find the tangent vector to the curve $\mathbf{f}\circ\boldsymbol{\gamma}$ at the point $(\mathbf{f}\circ\boldsymbol{\gamma})(0)$.

2. Let $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^3$ be a differentiable function, and suppose that $\mathbf{f}(3,9) = (5,3,1)$ and

$$D_{(3,9)}\mathbf{f} \;=\; \begin{bmatrix} 2 & 1 \\ 3 & 5 \\ 1 & 2 \end{bmatrix}.$$

   (a) Estimate $\mathbf{f}(3.02, 9.05)$.

   (b) Compute the directional derivative of $\mathbf{f}$ in the direction of the vector $(1,1)$.

   (c) Let $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^3$ be the curve $\boldsymbol{\gamma}(t) \;=\; \mathbf{f}(t,t^2)$. Compute $\dot{\boldsymbol{\gamma}}(3)$.

3. Let $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the map $f(\theta,\phi) = (\cos\theta\cos\phi,\, \sin\theta\cos\phi,\, \sin\phi)$.

   (a) Compute the matrix $D_{(\theta,\phi)}\mathbf{f}$.

   (b) At each point on the unit sphere, let $\widehat{\boldsymbol{\theta}}$ be a unit tangent vector pointing in the direction of increasing θ, and let $\widehat{\boldsymbol{\phi}}$ be a unit tangent vector pointing in the direction of increasing ϕ. Use your answer to part (a) to find formulas for $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\phi}}$.

4. Let $\psi\colon \mathbb{R}^3 \to \mathbb{R}$, and suppose that $\psi(1,2,5) = 3$ and $\nabla\psi(1,2,5) = (1,3,2)$. Let $\boldsymbol{\gamma}\colon \mathbb{R} \to \mathbb{R}^3$ be a regular curve, and suppose that $\boldsymbol{\gamma}(3) = (1,2,5)$ and $\dot{\boldsymbol{\gamma}}(3) = (4,5,2)$.

   (a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be the function $f(t) = \psi(\boldsymbol{\gamma}(t))$. Compute $f'(3)$.

   (b) Let $\mathbf{g}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the function $\mathbf{g}(x,y,z) = \boldsymbol{\gamma}(\psi(x,y,z))$. Compute $D_{(1,2,5)}\mathbf{g}$.