Derivatives of Vector-Valued Functions

Outline

1. Components
Consider a function general vector-valued function \( f: \mathbb{R}^m \to \mathbb{R}^n \). Such a function can be written as
\[
f(x_1, \ldots, x_m) = (f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)),
\]
where each \( f_i: \mathbb{R}^m \to \mathbb{R} \). The real-valued functions \( f_1, \ldots, f_n \) are called the components of \( f \).

For example, if \( f: \mathbb{R}^2 \to \mathbb{R}^3 \) is the function
\[
f(x, y) = (x^2 + y^2, x^2 - y^2, 2xy)
\]
then \( f \) 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 \( f: \mathbb{R}^m \to \mathbb{R}^n \), the partial derivative of \( f \) with respect to \( x_i \) is the vector
\[
\frac{\partial f}{\partial x_i} = \left( \frac{\partial f_1}{\partial x_i}, \ldots, \frac{\partial f_n}{\partial x_i} \right).
\]
We sometimes use subscripts to denote partial derivatives. For example, if \( f(x, y, z) \) is vector-valued function on \( \mathbb{R}^3 \), then \( f_y \) would denote the partial derivative of \( f \) with respect to \( y \).

3. Parameter Curves
Given a function \( f: \mathbb{R}^m \to \mathbb{R}^n \), a parameter curve for \( f \) is a curve \( \gamma: \mathbb{R} \to \mathbb{R}^n \) of the form
\[
\gamma(t) = f(p + te_i)
\]
where \( p \) is a point in \( \mathbb{R}^m \), and \( e_i \) is a unit vector in the \( x_i \) direction. Note that \( \gamma \) is obtained obtained from \( f(x_1, \ldots, x_m) \) 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 \( f \).

The partial derivatives of a function \( f \) at a point \( p \) can be interpreted as the tangent vectors to the parameter curves through \( f(p) \). Specifically, if \( \gamma \) is the parameter curve defined above, then
\[
\gamma'(0) = \frac{\partial f}{\partial x_i}(p).
\]
4. The Derivative Matrix
   If $f: \mathbb{R}^m \to \mathbb{R}^n$ the derivative of $f$ at a point $p$ is the matrix
   \[
   D_p f = \begin{bmatrix}
   \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_m}(p) \\
   \vdots & \ddots & \vdots \\
   \frac{\partial f_n}{\partial x_1}(p) & \cdots & \frac{\partial f_n}{\partial x_m}(p)
   \end{bmatrix}
   \]
   That is, $D_p f$ is the matrix whose columns are the partial derivatives $\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_m}(p)$.
   
   For example, if $f: \mathbb{R}^2 \to \mathbb{R}^3$ is the function $f(x,y) = (x^2 + y^2, x^2 - y^2, 2xy)$, then
   \[
   D_{(x,y)} f = \begin{bmatrix}
   2x & 2y \\
   2x & -2y \\
   2y & 2x
   \end{bmatrix}
   \]
   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 $D_p f$ is denoted $Df_p$, $Df(p)$, $df_p$, or $f'(p)$. 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 $D_p f$ to mean the derivative of $f$ at $p$, and $D_p f(v)$ to mean the product of the matrix $D_p f$ with a vector $v$.

5. Special Cases of the Derivative
   For a parametric curve $\gamma: \mathbb{R} \to \mathbb{R}^n$, the derivative of $\gamma$ is the same as the tangent vector $\dot{\gamma}$:
   \[
   D_t \gamma = \dot{\gamma}(t) = \begin{bmatrix}
   \dot{\gamma}_1(t) \\
   \vdots \\
   \dot{\gamma}_n(t)
   \end{bmatrix}
   \]
   
   For a real-valued function $f: \mathbb{R}^m \to \mathbb{R}$, the derivative $D_p f$ is the transpose of the gradient vector $\nabla f$:
   \[
   D_p f = \begin{bmatrix}
   \frac{\partial f}{\partial x_1}(p) & \cdots & \frac{\partial f}{\partial x_m}(p)
   \end{bmatrix}^T = \nabla f(p)^T.
   \]

6. Directional Derivatives
   Let $f: \mathbb{R}^m \to \mathbb{R}^n$ be a differentiable function, and let $e_i$ be a unit vector in the $x_i$ direction. Then:
   \[
   D_p f(e_i) = \frac{\partial f}{\partial x_i}(p).
   \]
   That is, the $i$’th column of $D_p f$ is the partial derivative on the right.
More generally, if $v$ is any vector in $\mathbb{R}^m$, then the product

$$D_p f(v)$$

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

The directional derivative $D_p f(v)$ can be interpreted as a tangent vector to a certain parametric curve. Specifically, let $\gamma: \mathbb{R} \to \mathbb{R}^n$ be the curve

$$\gamma(t) = f(p + tv).$$

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

$$\dot{\gamma}(0) = D_p f(v).$$

### 7. Differentials

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

$$df = f(p + dp) - f(p).$$

Then the vectors $df$ and $dp$ are related by the formula

$$df \approx D_p f(dp).$$

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

### 8. The Chain Rule

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

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

$$D_p(f \circ g) = (D_q f)(D_p g).$$

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

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

$$\dot{\gamma}(0) = D_p f(\dot{\gamma}(0))$$

That is, the tangent vector to $\dot{\gamma}$ is the product of the derivative matrix $D_p f$ with the tangent vector to $\gamma$. 


Practice Problems

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

   (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 \( D_{(2, 3\pi/4)} f \). Add the the vectors \( f_r(2, 3\pi/4) \) and \( f_\theta(2, 3\pi/4) \) to your drawing as tangent vectors to parameter curves.
   
   (c) Let \( \gamma : \mathbb{R} \to \mathbb{R}^2 \) be a regular curve, and suppose that \( \gamma(0) = (2, 3\pi/4) \) and \( \dot{\gamma}(0) = (3, 1) \). Find the tangent vector to the curve \( f \circ \gamma \) at the point \((f \circ \gamma)(0)\).

2. Let \( f : \mathbb{R}^2 \to \mathbb{R}^3 \) be a differentiable function, and suppose that \( f(3, 9) = (5, 3, 1) \) and
\[
D_{(3, 9)} f = \begin{bmatrix} 2 & 1 \\ 3 & 5 \\ 1 & 2 \end{bmatrix}.
\]

   (a) Estimate \( f(3.02, 9.05) \).
   
   (b) Compute the directional derivative of \( f \) in the direction of the vector \( (1, 1) \).
   
   (c) Let \( \gamma : \mathbb{R} \to \mathbb{R}^3 \) be the curve \( \gamma(t) = f(t, t^2) \). Compute \( \dot{\gamma}(3) \).

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

   (a) Compute the matrix \( D_{(\theta, \phi)} f \).
   
   (b) At each point on the unit sphere, let \( \hat{\theta} \) be a unit tangent vector pointing in the direction of increasing \( \theta \), and let \( \hat{\phi} \) be a unit tangent vector pointing in the direction of increasing \( \phi \). Use your answer to part (a) to find formulas for \( \hat{\theta} \) and \( \hat{\phi} \).

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

   (a) Let \( f : \mathbb{R} \to \mathbb{R} \) be the function \( f(t) = \psi(\gamma(t)) \). Compute \( f'(3) \).
   
   (b) Let \( g : \mathbb{R}^3 \to \mathbb{R}^3 \) be the function \( g(x, y, z) = \gamma(\psi(x, y, z)) \). Compute \( D_{(1, 2, 5)} g \).