Differentiability in General

We've defined differentiability for scalar-valued functions, $\mathbb{R}^n \to \mathbb{R}$.

Definition 1. *Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. For a point $\vec{a} \in \mathbb{R}^n$, define*

$$h(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}).$$

We say that f is differentiable *at $\vec{a}$ if*

$$\lim_{\vec{x} \to \vec{a}} \frac{f(\vec{x}) - h(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

If f is differentiable, we say that $h(\vec{x})$ is the tangent hyperplane *to f at $\vec{a}$.*

If any of the partial derivatives of f do not exist, or the above limit does not exist or is not 0, then f is not differentiable at $\vec{a}$.

We'd now like to define differentiability for vector-valued functions, $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$.

In order to define differentiability for scalar-valued functions $f : \mathbb{R}^n \to \mathbb{R}$, we organized our partial derivatives into a vector, the gradient of f.

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n} \right)$$

We would like to do something similar for a vector-valued function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$, and organize all of the partial derivatives into a single object. However, for a function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$, we not only have partial derivatives with respect to all of the different variables, we have partial derivatives of all of the component functions with respect to all of the different variables! This leads us to the derivative matrix.

The derivative matrix

Definition 2. *Let $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$, and suppose all of the partial derivatives of $\vec{f}$ exist. Write $\vec{f}$ in terms of its component functions,*

$$\vec{f}(x_1, ..., x_n) = (f_1(x_1, .., x_n), ..., f_m(x_1, ..., x_n)).$$

Learning outcomes: Understand the general definition of differentiability. Compute derivative matrices, and use them to show that functions are differentiable.

Author(s): Melissa Lynn

We define the derivative matrix *of $\vec{f}$ to be the $m \times n$ matrix with $\dfrac{\partial f_i}{\partial x_j}$ as the ij th entry. That is,*

$$
D\vec{f}(x_1, ..., x_n) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}.
$$

It can be difficult to remember whether the variable or the component changes across the rows or columns. Here are a couple of ways to remember which way it goes:

- The derivative matrix represents a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, hence is an $m \times n$ matrix.

- The rows are gradients of the component functions.

Example 1. *We'll find the derivative matrix of the function $\vec{f}(x, y, z) = (x^2 + yz, xyz)$. Since $\vec{f} : \mathbb{R}^3 \to \mathbb{R}^2$, the derivative matrix will be a 2×3 matrix.*

$$
D\vec{f}(x, y, z) = \begin{pmatrix} \dfrac{\partial}{\partial x}(x^2 + yz) & \dfrac{\partial}{\partial y}(x^2 + yz) & \dfrac{\partial}{\partial z}(x^2 + yz) \\ \dfrac{\partial}{\partial x}(xyz) & \dfrac{\partial}{\partial y}(xyz) & \dfrac{\partial}{\partial z}(xyz) \end{pmatrix}
$$

$$
= \begin{pmatrix} 2x & z & y \\ yz & xz & xy \end{pmatrix}
$$

Problem 1 *Find the derivative matrix of each of the following functions.*

$\vec{f}(x, y) = (x^2 + y^2, \sin(xy), e^{x+y})$

$$
D\vec{f}(x, y) = \begin{pmatrix} \boxed{2x} & \boxed{2y} \\ \boxed{y\cos(xy)} & \boxed{x\cos(xy)} \\ \boxed{e^{x+y}} & \boxed{e^{x+y}} \end{pmatrix}
$$

$\vec{g}(x, y, z) = (x^2 z + yz^2, x + y + z, x + y^2 + z)$

$$
D\vec{g}(x, y, z) = \begin{pmatrix} \boxed{2xz} & \boxed{z^2} & \boxed{x^2 + 2yz} \\ \boxed{1} & \boxed{1} & \boxed{1} \\ \boxed{1} & \boxed{2y} & \boxed{1} \end{pmatrix}
$$

Differentiability

We can now generalize our definition of differentiability for scalar-valued functions, by replacing the gradient with the derivative matrix.

Definition 3. *Consider a function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$. For a point $\vec{a} \in \mathbb{R}^n$, define*

$$\vec{h}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x} - \vec{a}).$$

We say that $\vec{f}$ is differentiable *at $\vec{a}$ if*

$$\lim_{\vec{x} \to \vec{a}} \frac{\vec{f}(\vec{x}) - \vec{h}(\vec{x})}{\|\vec{x} - \vec{a}\|} = \vec{0}.$$

If any of the partial derivatives of $\vec{f}$ do not exist, or the above limit does not exist or is not 0, then $\vec{f}$ is not differentiable at $\vec{a}$.

Note one of the quirks of multivariable differentiation: if the derivative matrix exist, it's still possible for the function to not be differentiable. We've seen this with functions before, when the partial derivatives of a function exist, and yet the function is not differentiable. For example, we've worked with the piecewise function defined by

$$f(x, y) = \begin{cases} 1 - |y| & \text{if } |y| \leq |x| \\ 1 - |x| & \text{if } |y| > |x| \end{cases}.$$

We found that the partial derivatives $f_x(0,0)$ and $f_y(0,0)$ both exist and are 0. However, the function f is not differentiable at $(0,0)$.

For vector-valued functions, we can also reduce differentiability to differentiability of its component functions.

Theorem 1. *A function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable if and only if its component functions are all differentiable.*

This theorem can quickly be proved from the definitions of differentiability in these two cases.

A criterion for differentiability

Checking differentiability using the limit definitions that we've found can be a huge pain! It would be much nicer if we could tell if a function is differentiable just by looking at the partial derivatives. Fortunately, this is possible in some cases.

Theorem 2. *Consider a function $\vec{f} \colon \mathbb{R}^n \to \mathbb{R}^m$. Suppose that all of the partial derivatives, $\partial f_i / \partial x_j$, exist and are continuous in a neighborhood of a point $\vec{a}$. Then $\vec{f}$ is differentiable at $\vec{a}$.*

An analogous result holds for scalar-valued functions. This theorem requires a very important note: its converse is *false*. That is, if one or more of the partial derivatives of a function is discontinuous, it's still possible that the function is differentiable. In this case, you would probably need to resort to the limit definition to determine differentiability.

Example 2. *We'll use the above theorem to show that $\vec{g}(x, y, z) = (x^2 z + yz^2, x + y + z, x + y^2 + z)$ is differentiable at all points in $\mathbb{R}^3$.*

First, we find all of the partial derivatives of g.

$$\frac{\partial g_1}{\partial x} = \boxed{2xz}$$

$$\frac{\partial g_1}{\partial y} = \boxed{z^2}$$

$$\frac{\partial g_1}{\partial z} = \boxed{x^2 + 2yz}$$

$$\frac{\partial g_2}{\partial x} = \boxed{1}$$

$$\frac{\partial g_2}{\partial y} = \boxed{1}$$

$$\frac{\partial g_2}{\partial z} = \boxed{1}$$

$$\frac{\partial g_3}{\partial x} = \boxed{1}$$

$$\frac{\partial g_3}{\partial y} = \boxed{2y}$$

$$\frac{\partial g_2}{\partial z} = \boxed{1}$$

All of these functions are polynomials, hence continuous at all points $\mathbb{R}^3$. Since the partial derivatives of $\vec{g}$ all exist and are continuous on $\mathbb{R}^3$, by the theorem above, $\vec{g}$ is differentiable at all points in $\mathbb{R}^3$.