5. Derivatives Part 1

In this lecture, we will discuss

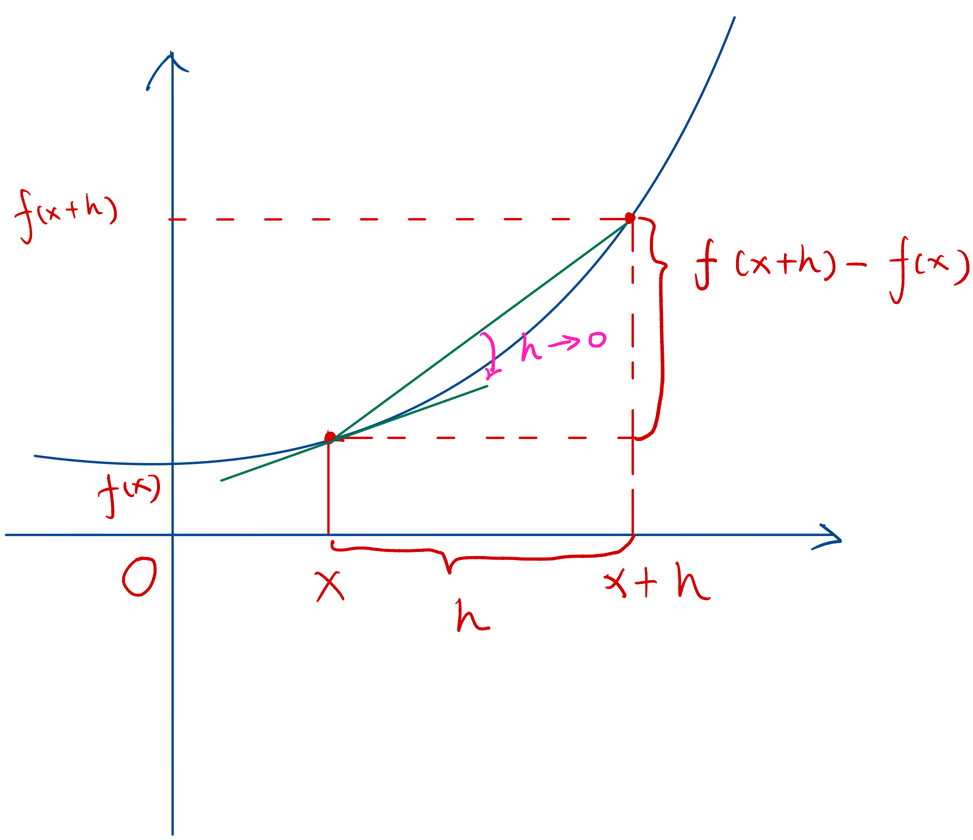
- Partial Derivatives
- Derivative of a Function of Several Variables
  - Jacobian matrix of a vector-valued function (notations: $D\mathbf{F}(x)$, $\mathbf{J}_{\mathbf{F}}$, $\nabla\mathbf{F}$)
  - Gradient of $f(\mathbf{x}) : U \subseteq \mathbb{R}^m \to \mathbb{R}$, or gradient of f, or $\nabla f(\mathbf{x})$
  - Differentiability of a Vector-Valued Function
  - Derivative of a Vector-Valued Function
- Differential
  - Review the differential of $f : \mathbb{R} \to \mathbb{R}$
  - Differential of $f : \mathbb{R}^2 \to \mathbb{R}$

Review

Recall that the derivative $f'(x)$ of a function $f(x)$ is defined as a limit of difference quotients

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

provided that the limit exists. The number $f'(x_0)$ is the slope of the tangent to the graph of $f(x)$ at the point $(x_0, f(x_0))$.



Partial Derivatives

For the rigorous definitions, we discuss the following notions. In the homework problems, we won't worry much about the subtly.

Recall the the following definition and examples of open ball. in Lecture 4

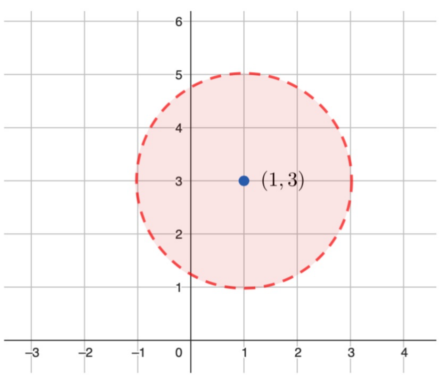
Definition. Open Balls in $\mathbb{R}^m$
The open ball $B(\mathbf{a}, r) \subseteq \mathbb{R}^m$ with center $\mathbf{a} = (a_1, \ldots, a_m)$ and radius $r (r > 0)$ is the set of all points $\mathbf{x}$ in $\mathbb{R}^m$ whose distance from a fixed point a is smaller than r. In symbols,

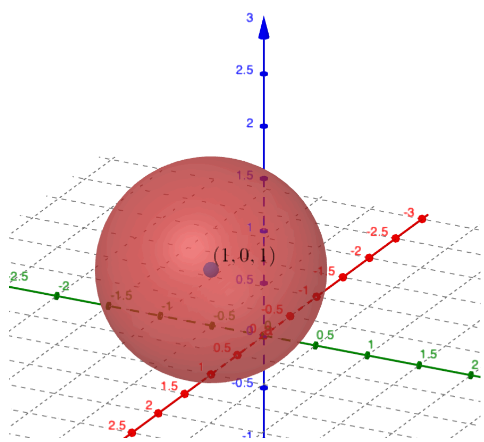
$$B(\mathbf{a}, r) = \{\mathbf{x} \in \mathbb{R}^m \mid \|\mathbf{x} - \mathbf{a}\| < r\},$$

where $\mathbf{x} = (x_1, \ldots, x_m)$ and $\|\mathbf{x} - \mathbf{a}\| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_m - a_m)^2}$.

For example,

- In $\mathbb{R}^2$, the open ball $B((1,3), 2)$ contains all points in $\mathbb{R}^2$ whose distance from $(1, 3)$ is strictly smaller than 2.
- In $\mathbb{R}^3$, the open ball $B((1,0,1), 1)$ contains all points in $\mathbb{R}^3$ whose distance from $(1, 0, 1)$ is strictly smaller than 1.





- In $\mathbb{R}$, the open ball $B((2), 1)$ contains all points in $\mathbb{R}$ whose distance from 2 is strictly smaller than 1.

  Note it is simply the open interval $(1, 3)$.

Definition. Open Sets in $\mathbb{R}^m$
A subset $U \subseteq \mathbb{R}^m$ is called open if for every point $\mathbf{a} \in U$ there exists a real number $\epsilon > 0$ such that $B(\mathbf{a}, \epsilon)$ is contained in U.

For example, open intervals, e.g. $(0, 1)$ etc. are examples of open sets in $\mathbb{R}$.

In the following discussion, we assume U is an open set unless otherwise stated.

Definition. Partial Derivative for $f : \mathbb{R}^2 \to \mathbb{R}$

Let $f(x, y)$ be a real-valued function with domain U in $\mathbb{R}^2$, and let (a, b) be a point in U. Then, the partial derivative of f at (a, b) with respect to x, denoted by $\frac{\partial f}{\partial x}(a, b)$, is defined as

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

and the partial derivative of f at (a, b) with respect to y, denoted by $\frac{\partial f}{\partial y}(a, b)$, is defined as

$$\frac{\partial f}{\partial y}(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h}.$$

In other words, $\frac{\partial f}{\partial x}$ can be obtained by regarding y as a constant, and applying standard rules for differentiating functions of x.

Example 1.

Find the first partial derivatives of the function $f(x, y) = \frac{\sin(xy^2)}{x^2 + 1}$.

ANS: Treating y as a constant and taking derivative of $f(x,y)$ with respect to x, we have.

$$f_x(x,y) = \frac{\partial f}{\partial x}(x,y) = \frac{y^2\cos(xy^2)\cdot(x^2+1) - (2x)\sin(xy^2)}{(x^2+1)^2}$$

Quotient Rule: $\left[\frac{f(x)}{g(x)}\right]' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$

Chain Rule: $[f(g(x))]' = f'(g(x))g'(x)$

and treating x as a constant and taking derivative of $f(x,y)$ with respect to y, we have

$$f_y(x,y) = \frac{\partial f}{\partial y}(x,y) = \frac{\frac{\partial(xy^2)}{\partial y}}{x^2+1} \cos(xy^2)$$

$$= \frac{2xy}{x^2+1}\cos(xy^2)$$

Note the definition of partial derivatives can be generalized to f.

Definition. Partial Derivative

Let $f: U \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ be a real-valued function of m variables $x_1, \ldots, x_m$, defined on an open set U in $\mathbb{R}^m$. The partial derivative of f with respect to x_i (or with respect to the i th variable, $i=1, \ldots, m)$ is a real-valued function $\partial f / \partial x_i$ of m variables, defined by

$$\frac{\partial f}{\partial x_i}\left(x_1, \ldots, x_m\right)=\lim _{h \rightarrow 0} \frac{f\left(x_1, \ldots, x_i+h, \ldots, x_m\right)-f\left(x_1, \ldots, x_i, \ldots, x_m\right)}{h},$$

provided that the limit exists.

Example 2. (Related to WebWork HW#2)

(1) Compute the partial derivatives f_x, f_z for $f(x, y, z)=\ln \left(x+y+z^2\right)$

(2) Compute $f_z(1,0, \pi)$.

ANS: (1) By the chain rule,

$$f_x(x,y,z) = \frac{\partial f}{\partial x}(x,y,z) = \frac{1}{x+y+z^2} \cdot 1 = \frac{1}{x+y+z^2}$$

$$f_z(x,y,z) = \frac{\partial f}{\partial z}(x,y,z) = \frac{1}{x+y+z^2} \cdot 2z = \frac{2z}{x+y+z^2}$$

(2) $$f_z(1,0,\pi) = \frac{2\pi}{1+0+\pi^2} = \frac{2\pi}{1+\pi^2}$$

Derivative of a Function of Several Variables

We start by defining the Jacobian matrix of a vector-valued function $\mathbf{F} : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$.

Note our current WebWork exercises are mostly concerned with special case $n = 1$, that is, $\mathbf{F}$ is a real-valued function.

We will return to this general definition of the Jacobian matrix later in this course.

- Let $\mathbf{F}$ be a vector-valued function $\mathbf{F} : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$.
- Recall that $\mathbf{F}$ can be written in terms of its components as

$$\mathbf{F}\left(x_1, \ldots, x_m\right) = \left(F_1\left(x_1, \ldots, x_m\right), F_2\left(x_1, \ldots, x_m\right), \ldots, F_n\left(x_1, \ldots, x_m\right)\right),$$

- or as $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots, F_n(\mathbf{x}))$, where $\mathbf{x} = (x_1, \ldots, x_m)$.
- In other words, we can describe a vector-valued function $\mathbf{F}$ using n real-valued functions of m variables.

Roughly speaking, ***Jacobian matrix*** of a vector-valued function of several variables is the matrix of all its first-order partialderivatives.

space

Definition. Jacobian Matrix $D\mathbf{F}(\mathbf{x})$

By $D\mathbf{F}(\mathbf{x})$ we denote the $n \times m$ matrix of partial derivatives of the components of $\mathbf{F}$ evaluated at $\mathbf{x}$ (provided that all partial derivatives exist at $\mathbf{x}$). Thus,

$$D\mathbf{F}(x) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}) & \frac{\partial F_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial F_1}{\partial x_m}(\mathbf{x}) \\ \frac{\partial F_2}{\partial x_1}(\mathbf{x}) & \frac{\partial F_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial F_2}{\partial x_m}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(\mathbf{x}) & \frac{\partial F_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial F_n}{\partial x_m}(\mathbf{x}) \end{bmatrix} \tag{1}$$

The matrix $D\mathbf{F}(\mathbf{x})$ has n rows and m columns (the number of rows is the number of component functions of $\mathbf{F}$, and the number of columns equals the number of variables).

Remark about the notation for the Jacobian matrix. We will follow the notation from the book as $D\mathbf{F}(\mathbf{x})$. Sometimes, it is also written as $\mathbf{J}_{\mathbf{F}}$, $\nabla\mathbf{F}$, and $\dfrac{\partial\left(F_1, \ldots, F_m\right)}{\partial\left(x_1, \ldots, x_n\right)}$ in other places.

Example 3 Let $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^4$ be given by $\mathbf{F}(x, y, z) = \left(e^{x+yz}, x^2 + 1, \sin(y+z), 4y\right)$. Compute $D\mathbf{F}(\mathbf{x})$.

ANS: We have

$$F_1(x,y,z) = e^{x+yz}, \; F_2(x,y,z) = x^2+1, \; F_3(x,y,z) = \sin(y+z), \; F_4(x,y,z) = 4y.$$

Thus by definition.

$$D\vec{F}(x,y,z) = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} \\ \frac{\partial F_4}{\partial x} & \frac{\partial F_4}{\partial y} & \frac{\partial F_4}{\partial z} \end{bmatrix}_{4\times 3}$$

$$= \begin{bmatrix} e^{x+yz} & ze^{x+yz} & ye^{x+yz} \\ 2x & 0 & 0 \\ 0 & \cos(y+z) & \cos(y+z) \\ 0 & 4 & 0 \end{bmatrix}$$

Special Cases of $D\mathbf{F}(\mathbf{x})$

- $n = m = 1$
  - If $f(x) : \mathbb{R} \to \mathbb{R}$, then $Df(x)$ is a 1×1 matrix whose entry is the derivative of the only component f to the only variable x.
  - Thus $Df(x)$ is the usual derivative $f'(x)$.
- $n = 1$
  - Assume that $f(\mathbf{x}) : U \subseteq \mathbb{R}^m \to \mathbb{R}$ is a real-valued function of m variables. Then $Df(\mathbf{x})$ is the $1 \times m$ matrix

$$Df(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}) \quad \frac{\partial f}{\partial x_2}(\mathbf{x}) \cdots \frac{\partial f}{\partial x_m}(\mathbf{x}) \right],$$

    whose only row consists of partial derivatives of f with respect to all variables $x_1, \ldots, x_m$, evaluated at $\mathbf{x} = (x_1, \ldots, x_m)$.
  - $Df(\mathbf{x})$ is called the *gradient* of f at $\mathbf{x}$, and is denoted by $\nabla f(\mathbf{x})$.

Example 4 (Related to WebWork HW#5)

Compute $\nabla f(2, 1, -1)$ if $f(x, y, z) = xy \ln\left(z^2 + xy\right)$.

ANS: We know

$$\nabla f(x,y,z) = \left(\frac{\partial f}{\partial x}(x,y,z), \ \frac{\partial f}{\partial y}(x,y,z), \ \frac{\partial f}{\partial z}(x,y,z) \right)$$

$$f_x(x,y,z) = \frac{\partial f}{\partial x}(x,y,z) = y\ln(z^2+xy) + xy \frac{y}{z^2+xy}$$

(product rule and chain rule)

$$f_y(x,y,z) = \frac{\partial f}{\partial y}(x,y,z) = x\ln(z^2+xy) + xy \frac{x}{z^2+xy}$$

$$f_z(x,y,z) = \frac{\partial f}{\partial z}(x,y,z) = 2xyz \cdot \frac{1}{z^2+xy}$$

Thus $\nabla f(x,y,z) = (f_x, f_y, f_z)$

$$\nabla f(2,1,-1) = (f_x(2,1,-1), f_y(2,1,-1), f_z(2,1,-1))$$

$$= \left(\ln 3 + \frac{2}{3}, \ 2\ln 3 + \frac{4}{3}, \ -\frac{4}{3} \right)$$

Definition. Differentiability of a Vector-Valued Function

A vector-valued function $\mathbf{F} = (F_1, \ldots, F_n) : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$, defined on an open set $U \subseteq \mathbb{R}^m$, is differentiable at $\mathbf{a} \in U$ if

(a) all partial derivatives of the components $F_1, \ldots, F_n$ of $\mathbf{F}$ exist at $\mathbf{a}$, and

(b) the Jacobian matrix of partial derivatives $D\mathbf{F}(\mathbf{a})$ of $\mathbf{F}$ at a satisfies

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a}) - D\mathbf{F}(\mathbf{a})(\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0 \tag{2}$$

Definition Derivative of a Vector-Valued Function

If a vector-valued function $\mathbf{F}$ satisfies the conditions (a) and (b) of the above definition, then the Jacobian matrix $D\mathbf{F}(\mathbf{a})$ of partial derivatives given by Eq(1) is called the ***derivative of* $\mathbf{F}$ *at* $\mathbf{a}$**.

Differential

Let $f : \mathbb{R} \to \mathbb{R}$. Recall in calculus, we have the differential of f by

$$df(x) = f'(x)dx$$

This notion can be generalized to $f : \mathbb{R}^2 \to \mathbb{R}$.

Definition. Differential of a function $f : \mathbb{R}^2 \to \mathbb{R}$ at a point

Choose a point (a, b) in the domain of a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$. The differential of the function $f(x, y)$ at (a, b) is defined as

$$df = f_x(a, b)dx + f_y(a, b)dy.$$

Example 5 (Related to WebWork HW#6)

Find the gradient ∇f of the function f given the differential.

$$df = \frac{y}{x^2 + y^2}dx + \frac{-x}{x^2 + y^2}dy$$

Recall
$\nabla f = (f_x, f_y)$

ANS: By the definition of differential of f at (x, y), we know

$$df = f_x(x, y)dx + f_y(x, y)dy$$

By comparing with the given df, we have

$$f_x(x,y) = \frac{y}{x^2+y^2}$$

$$f_y(x,y) = \frac{-x}{x^2+y^2}$$

Thus $\nabla f = (f_x(x,y), f_y(x,y))$

$$= \left(\frac{y}{x^2+y^2}, \frac{-x}{x^2+y^2}\right)$$

Exercise 6 (Related to WebWork HW#9)

Suppose $f(x,t) = x^t + t\ln x$.

(1) At any point (x,t), compute the differential df.

(2) At the point $(1,2)$, compute differential df.

(3) At the point $(1,2)$ with $dx = 0.5$ and $dt = -0.3$, compute df.

ANS: (1) Recall $df = f_x(x,t)\,dx + f_t(x,t)\,dt$.

Keeping t fixed as a constant, we have

$$f_x(x,t) = \frac{\partial f}{\partial x}(x,t) = tx^{t-1} + \frac{t}{x}$$

Keeping x fixed as a constant, we have

$$f_t(x,t) = \frac{\partial f}{\partial t}(x,t) = x^t\ln x + \ln x \quad \text{(Recall } (a^t)' = a^t\ln a\text{)}$$

Thus $df = (tx^{t-1} + \frac{t}{x})dx + (x^t\ln x + \ln x)\,dt$

(2) At the point $(1,2)$, $(x=1, t=2)$

$$df = 4dx$$ Note $\ln x = 0$ at $x=1$, thus the coefficient for dt is 0.

(3) At the point $(1,2)$ with $dx = 0.5$, $dt = -0.3$, we have

$$df = 4\cdot \cancel{dx}^{0.5} = 2$$

Exercise 7

Find the gradient of the function $f(x,y,z) = xe^{yz^2}$, at the point $(1,1,2)$.

ANS: We have

$$\nabla f(x,y,z) = (f_x(x,y,z), f_y(x,y,z), f_z(x,y,z))$$

$$= (e^{yz^2}, xz^2e^{yz^2}, 2xze^{yz^2})$$

Thus $\nabla f(1,1,2) = (e^4, 4e^4, 4e^4)$

$$= e^4\vec{i} + 4e^4\vec{j} + 4e^4\vec{k}$$