MATH 104: Multivariable Calculus (brief notes)

Truong-Son Van

Contents

Spring 2024

Disclaimer

This is class notes for Multivariable Calculus at Fublbright University Vietnam. I claim no originality in this work as it is mostly taken from the reference books. However, all errors and typos are solely mine.

Chapter 1

Vectors & Matrices

1.1 Basics

Reading: Stewart Chapter 12, Thomas Calculus Chapter 12, Active Calculus Chapter 9

You should be able to answer the following questions after reading this section:

- What is a vector?
- What does it mean for two vectors to be equal?
- How do we add two vectors together and multiply a vector by a scalar?
- How do we determine the magnitude of a vector?
- What is a unit vector
- How do we find a unit vector in the direction of a given vector?

Typically, we talk about 3-dimensional vectors (as discussed in Stewart and Thomas). However, since talking about n-dimensional vectors doesn't require much more effort, we will talk about n-dimensional vectors instead.

Definition 1.1. An n-dimensional Euclidean space $\mathbb{R}^n$ is the Cartesian product of n Euclidean spaces $\mathbb{R}$.

Definition 1.2. An n-dimensional vector $\mathbf{v} \in \mathbb{R}^n$ is a tuple

$$\mathbf{v} = \langle v_1, \dots, v_n \rangle , \tag{1.1}$$

where $v_i \in \mathbb{R}$.

In dimensions less than or equal to 3, we represent a vector geometrically by an arrow, whose length represents the magnitude.

Remark. A point in $\mathbb{R}^n$ is also represented by an n-tuple but with round brackets. A vector connecting two points $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ can be constructed as

$$\mathbf{x} = \langle b_1 - a_1, \dots, b_n - a_n \rangle .$$

We denote the above vector as $\vec{AB}$ where A is the tail (initial point) and B is the tip/head (terminal point). We denote $\mathbf{0}$ to be the zero vector, i.e.,

$$\mathbf{0} = \langle 0, \dots, 0 \rangle .$$

Definition 1.3. The length of a vector $\mathbf{v}$ (denoted by $|\mathbf{v}|$) is defined to be

$$|\mathbf{v}| = \sqrt{v_1^2 + \dots + v_n^2} . \tag{1.2}$$

Definition 1.4. A unit vector is a vector that has magnitude 1.

Exercise 1.1. Turn a vector $\mathbf{v} \in \mathbb{R}^n$ into a unit vector with the same direction.

Rules to manipulate vectors

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $c, d \in \mathbb{R}$. Then,

$$c(\mathbf{a} + \mathbf{b}) = \langle ca_1 + cb_1, \ldots, ca_n + cb_n \rangle = c\mathbf{a} + c\mathbf{b}\,,$$

and

$$(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}\,.$$

These formulas are deceptively simple. Make sure you understand all the implications.

Because of this rule, sometimes it is good to write vectors in terms of elementary vectors:

$$\mathbf{u} = u_1\mathbf{e_1} + \cdots + u_n\mathbf{e_n}\,,$$

where $e_i = \langle 0, \ldots, 1, \ldots, 0 \rangle$ is the vector which has zero at all entries except that the i^{th} entry is 1.

In 3D,

$$\mathbf{e_1} = \mathbf{i}\,, \qquad \mathbf{e_2} = \mathbf{j}\,, \qquad \mathbf{e_3} = \mathbf{k}\,.$$

Properties of vector operations

Read the book

(Make sure you understand the geometric intepretation)

1.2 Products

1.2.1 Dot product

- How is the dot product of two vectors defined and what geometric information does it tell us?
- How can we tell if two vectors in $\mathbb{R}^n$ are perpendicular?
- How do we find the projection of one vector onto another?

Definition 1.5. The dot product of vectors $\mathbf{u} = \langle u_1, \ldots, u_n \rangle$ and $\mathbf{v} = \langle v_1, \ldots, v_n \rangle$ in $\mathbb{R}^n$ is the scalar

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n\,.$$

Properties of dot product

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Then,

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$,
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = (\mathbf{u} \cdot \mathbf{w}) + (\mathbf{v} \cdot \mathbf{w})$,
3. If c is a scalar, then $(c\mathbf{u}) \cdot \mathbf{w} = c(\mathbf{u} \cdot \mathbf{w})$.

Theorem 1.1 (Law of cosine)**.** *If θ is the angle between the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, then*

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}| \cos\theta\,.$$

Corollary 1.1. *Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal to each other if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.*

Projection

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. The component of $\mathbf{u}$ in the direction of $\mathbf{v}$ is the scalar

$$\mathrm{comp}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|}\,,$$

and the projection of $\mathbf{u}$ onto $\mathbf{v}$ is the vector

$$\mathrm{proj}_{\mathbf{v}}\mathbf{u} = \left(\mathbf{u}\cdot\frac{\mathbf{v}}{|\mathbf{v}|}\right)\frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\,.$$

1.2.2 3D special: Cross product

This concept is very specific to $\mathbb{R}^3$. It will not make sense in other dimensions.

Definition 1.6. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. The cross product of $\mathbf{a}$ and $\mathbf{b}$ is defined to be

$$\mathbf{a}\times\mathbf{b} = \langle a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1\rangle\,.$$

Theorem 1.2. *Let θ be the angle between $\mathbf{a}$ and $\mathbf{b}$. Then,*

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta\,.$$

Theorem 1.3. *The vector $\mathbf{a}\times\mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.*

1.2.3 Distance from a point

We can use the cross and dot products to measure the distance of one point to either a plane or a line.

Let $P \in \mathbb{R}^n$ and $\vec{r}(t) = R_0 + t\vec{v}$ be a line. Then the distance from P to $\vec{r}(t)$ is

$$Dist = \frac{|\vec{R_0P}\times\vec{v}|}{|\vec{v}|}$$

1.3 Matrices

A matrix is an 2 dimensional array with rows and columns.

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix}$$

Another way to write out matrix A is

$$A = (A_{ij})$$

where the first index i represents the row and the second index j represents the column.

1.3.1 Operations on matrices

1. Addition: let A and B be two matrices with same dimension $m\times n$. Then $A+B$ is an $m\times n$ matrix such that
$$[A+B]_{ij} = A_{ij} + B_{ij}.$$
2. Scalar multiplication: let A be a $m\times n$ matrix, c is a constant scalar. then cA is a $m\times n$ matrix such that
$$((cA)_{ij}) = (cA_{ij}).$$
3. Matrix multiplication: let A be $m\times n$ matrix and B be $n\times l$ matrix. Then the multiplication AB is a $m\times l$ matrix such that
$$[AB]_{ij} = \sum_k A_{ik}B_{kj}.$$

1.3.2 Linear transformation

A linear transformation is a function $f : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(a\vec{u} + b\vec{v}) = af(\vec{u}) + bf(\vec{v})$$

for all $a, b \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$.

It turns out that every linear transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ can be represented as a $m \times n$ matrix.

Chapter 2

Some basic equations in $\mathbb{R}^3$

Just to build some toy examples for the future, we will play with some basic equations in three dimensions.

2.1 Equations for lines

A line is a collection of points that is parallel to a vector and goes through a specific point. To capture this idea, we have the following representation for a line

$$L = \{\mathbf{r}(t)\,|\,\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}, t \in \mathbb{R}\}\,,$$

where r_0 is the initial position and $\mathbf{v}$ is the direction. The equation for $\mathbf{r}(t)$ is called a **vector equation for a line** L.

Let $\mathbf{v} = \langle v_1, v_2, v_3\rangle$ and $\mathbf{r}_0 = (x_0, y_0, z_0)$. The **parametric equations** of L is the following system of equations

$$\begin{aligned} x &= x_0 + v_1 t\,, \\ y &= y_0 + v_2 t\,, \\ z &= z_0 + v_3 t\,. \end{aligned}$$

This leads to the **symmetric equations** of L

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}\,.$$

Definition 2.1. Two lines are parallel if their directional vectors are parallel (scalar multiple of each other).

Two lines that are not parallel and don't intersect each other are said to be skew.

2.2 Equations for planes

A plane is a collection of points that is perpendicular to one specific direction represented by a some vector called a **normal vector**. Note that due to scaling, there are more than one normal vector. To capture this idea, we have the following representation of a plane

$$P = \{\mathbf{r}\,|\,\mathbf{n}\cdot(\mathbf{r} - \mathbf{r}_0) = 0\}\,.$$

This is called a **vector equation for the plane** P.

Multiplying things out, we have the **scalar equation of the plane** P with normal vector $\mathbf{n} = \langle n_1, n_2, n_3 \rangle$ through a point $P_0(x_0, y_0, z_0)$

$$n_1(r_1 - x_0) + n_2(r_2 - y_0) + n_3(r_3 - z_0) = 0\,.$$

The equation of the form

$$ax + by + cz + d = 0$$

is called a **linear equation**.

Definition 2.2. Two planes are said to be parallel if their normal vectors are parallel. If two planes are not parallel, they intersect in a straight line and the angle between the two planes is defined to be the angle between the two normal vectors.

2.3 Cylinders

Definition 2.3. A cylinder is a surface that consists of all lines (called **rulings**) that are parallel to a given line.

Example 2.1.

1. $z = x^2$
2. $x^2 + y^2 = 1$

2.4 Quadric surfaces

Definition 2.4. A quadric surface is the graph of a second-degree equation in three variables x, y and z. The equation that represents these surfaces is

$$Ax^2 + By^2 + Cz^2 + Dz = E\,.$$

Example 2.2.

1. Ellipsoid
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\,.$$
2. Hyperbolic paraboloid
$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = \frac{z}{c}\,.$$
3. Elliptical cone
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}\,.$$

Read the books for more surfaces and pictures.

Chapter 3

Functions in higher dimensions

Reading: Stewart Chapter 12, 13, Thomas Calculus Chapter 12, 13, Active Calculus Chapter 9

3.1 Functions of several variables

Definition 3.1. A function of several variables is a function $f : D \to C$ where $D \subseteq \mathbb{R}^m$ and $C \subseteq \mathbb{R}^n$, where $m \geq 2$ and $n \geq 1$.

$$f(x) = (f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m)) .$$

D is called the domain of f and C is called the codomain of f.

The domain of f is where each of the component f_i of f is defined.

Example 3.1. The following are some examples of multivariable functions

1. $f(x, y) = x^2 - 2xy + y^2$
2. $f(x, y, z) = \frac{1}{1-xy^2}$

3.2 Vector functions

3.2.1 Limit, continuity and differentiation

The expression in the vector equation for a line is an example of a function that maps from $\mathbb{R}$ to $\mathbb{R}^n$. There's no one who would stop us from considering more general kinds of function.

Definition 3.2. A **vector function** (**vector-valued function**) is a function that has the codomain that belongs to $\mathbb{R}^n$ where $n \geq 2$. In other words, $f : D \to \mathbb{R}^n$.

Example 3.2. The following are some examples of vector functions.

1. $\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}$
2. $\mathbf{f}(t) = \langle \cos(t), \sin(t), t \rangle$

Note that my definition is more general than that in the book. However, **In this course, whenever we talk about vector valued function, we will only refer to that which has one dimensional domain ($D \subseteq \mathbb{R}$).**

By and large, there's nothing different between a vector function and a one-variable scalar function. All the concepts such as limit, continuity and differentiability are applied to each coordinate the same way as in one dimensional case.

Theorem 3.1. *Let $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$, given by $\mathbf{r}(t) = \langle r_1(t), \dots, r_n(t) \rangle$. Then, $\mathbf{r}$ is said to be continuous at t_0 if*

$$\mathbf{r}(t_0) = \lim_{t \to t_0} \mathbf{r}(t)\,,$$

where

$$\lim_{t \to t_0} \mathbf{r}(t) = \langle \lim_{t \to t_0} r_1(t), \dots, \lim_{t \to t_0} r_n(t) \rangle\,.$$

Furthermore, we can define the derivative of $\mathbf{r}$

$$\frac{d}{dt}\mathbf{r}(t) = \mathbf{r}'(t) = \lim_{h \to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}$$

if this limit exists.

When $\mathbf{r} : I \to \mathbb{R}^n$ (I is an interval in $\mathbb{R}$) is continuous, we call it a **space curve** (to describe the intuitive picture of what a curve should look like in our mind).

Geometrically, if $\mathbf{r}'(t)$ exists and $\mathbf{r}'(t) \neq \mathbf{0}$, it represents the **tangent vector** of the curve $\mathbf{r}$ at t.

Definition 3.3. A **parametric equation** for a curve is an equation of the form

$$x = x(t)\,, \quad y = y(t)\,, \quad z = z(t)\,.$$

Typical differentiation rules apply.

Theorem 3.2 (Differentiation rules)**.**

1. $(\mathbf{u}(t) + \mathbf{v}(t))' = \mathbf{u}'(t) + \mathbf{v}'(t)$
2. $(c\mathbf{u}(t))' = c\mathbf{u}'(t)$
3. $(f(t)\mathbf{u}(t))' = f'(t)\mathbf{u}(t) + f(t)\mathbf{u}'(t)$
4. $(\mathbf{u}(t) \cdot \mathbf{v}(t))' = \mathbf{u}'(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \mathbf{v}'(t)$
5. $(\mathbf{u}(t) \times \mathbf{v}(t))' = \mathbf{u}'(t) \times \mathbf{v}(t) + \mathbf{u}(t) \times \mathbf{v}'(t)$
6. $(\mathbf{u}(f(t)))' = \mathbf{u}'(f(t))f'(t)$

3.2.2 Integrals

There are different ways to play with integrals for vector functions, each has its own interpretation and physical applications.

3.2.2.1 Indefinite integral

$$\int_a^b \mathbf{r}(t)\,dt = \left\langle \int_a^b r_1(t)\,dt, \int_a^b r_2(t)\,dt, \int_a^b r_3(t)\,dt \right\rangle$$

3.2.2.2 Arc Length and curvature

Definition 3.4. The length the curve $\mathbf{r} : [a,b] \to \mathbb{R}^n$ is defined to be

$$L = \int_a^b |\mathbf{r}'(t)|\;dt\,.$$

If one wants to keep track the length of the curve $\mathbf{r} : [a,b] \to \mathbb{R}^n$ made by an airplane at any time t, one uses the **arc length function**

$$\ell(t) = \int_a^t |\mathbf{r}'(u)|\;du\,.$$

Re-parametrize with respect to arc length

The nice thing about $\ell(t)$ is that it is a strictly increasing function with respect to t, given that $\mathbf{r}'$ is non-zero for all t. Therefore, letting $s = \ell(t)$, we can talk about the inverse of ℓ, $\ell^{-1} : [0, L] \to [a, b]$

$$t = \ell^{-1}(s)\,.$$

Therefore, we can re-write

$$\mathbf{r}(t) = \mathbf{r}(\ell^{-1}(s))\,.$$

Theorem 3.3.

$$\left|\frac{dr(t)}{ds}\right| = 1\,.$$

Thus,

$$l(s) = \int_0^s \left|\frac{d}{ds}\mathbf{r}(t)\right| dt = s\,.$$

Because of the unchanging nature of the arc-length (with respect to the parametrization), it is used to define a geometric quantity of a space curve called **curvature**.

Definition 3.5 (curvature)**.** Let $\mathbf{T}(t)$ be the unit tangent vector of the curve $\mathbf{r} : [a, b] \to \mathbb{R}^3$. The curvature of $\mathbf{r}(t(s))$ is defined to be

$$\kappa(s) = \left|\frac{d\mathbf{T}(t(s))}{ds}\right|\,.$$

To convert this into the parameter t, we write $s = s(t)$ and use chain rule to get.

Theorem 3.4. *We have that*

$$\kappa(s(t)) = \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|}\,.$$

3.2.3 Space curve in $\mathbb{R}^3$ and motion in space

Read the book. This part is not required but it is so beautiful, you may want to read it as an exercise at home (to test how much you understand what we've been discussing so far).

3.3 Activity: on osculating circle and curvature

For those who are interested in the geometrical meaning of the curvature without having to accept from the book that the curvature is the inverse of the radius of the osculating circle, please take a look at https://github.com/sonv/MultiCalc/blob/main/Writing/latexbuild/osculating.pdf.

Chapter 4

Partial derivatives

4.1 Limits and continuity

The following definition is from Stewart.

Definition 4.1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then we say that the limit of $f(x)$ as x approaches a is L and we write

$$\lim_{x \to a} f(x) = L$$

if for every number $\epsilon > 0$ there is a corresponding number $\delta > 0$ such that $|f(x,y) - L| < \epsilon$ if $|x - a| < \delta$.

Finding if a function has limit as a point in higher dimension is not as simple as the case for 1 dimension.

Determining whether a multivariable function has a limit sometimes is an art and it requires a lot of experiences and practice. However, there are certain rules that could help us.

Theorem 4.1. *Let L, M and k be real numbers and that*

$$\lim_{x \to a} f(x,y) = L\,, \qquad \lim_{x \to a} g(x,y) = M\,.$$

We then have

1. $\lim\limits_{x \to a}(f(x) + g(x)) = L + M$,
2. $\lim\limits_{x \to a}(kf(x)) = kL$,
3. $\lim\limits_{x \to a}(f(x)g(x)) = LM$,
4. $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$ *if* $M \neq 0$,
5. $\lim\limits_{x \to a} f(x)^p = L^p$ *for* $p > 0$,

Strategy to find out that a two-variable function does NOT have a limit.

If $\lim_{(x,y)\to(a,b)} f(x,y) = L_1$ as $(x,y) \to (a,b)$ along a path C_1, and $\lim_{(x,y)\to(a,b)} f(x,y) = L_2$ as $(x,y) \to (a,b)$ along a path C_2, where $L_1 \neq L_2$, then $\lim_{(x,y)\to(a,b)} f(x,y)$ does not exist.

Example 4.1. $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$ does not exist.

$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ does not exist.

$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^4+y^4}$ does not exist.

$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = 0$.

4.2 Partial derivatives

given a function $f : \mathbb{R}^n \to \mathbb{R}^m$, the partial derivative with respect to the jth variable x_j of the ith output at $a \in \mathbb{R}^n$ is

$$\frac{\partial}{\partial x_j} f_i(a) = \lim_{h\to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)}{h} .$$

Notations. The partial derivatives sometimes have different notations:

$$\partial_j f_i(a) = \partial_{x_j} f_i(a) = \frac{\partial}{\partial x_j} f_i(a).$$

From here, one can define higher partial derivatives such as the following

$$\partial^3_{123} f_i(a) = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\frac{\partial}{\partial x_3} f_i(a) .$$

Note that the power over the symbol ∂ represents the order of derivatives.

Some important notations

Let $f : D \to \mathbb{R}$ be a function. We write the following, if exist,

$$\nabla f = \begin{bmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{bmatrix}$$

$$\Delta f = \partial^2_{x_1} f + \dots \partial^2_{x_n} f .$$

4.3 Differentiability

Definition 4.2 (Differentiability)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$. f is said to be differentiable at $a \in \mathbb{R}^n$ if there exists a linear transformation $[Df]_a$ such that for every vector $\mathbf{h} \in \mathbb{R}^n$

$$\lim_{|\mathbf{h}|\to 0} \frac{f(a + \mathbf{h}) - f(a) - [Df]_a \mathbf{h}}{|\mathbf{h}|} = 0.$$

For $f : \mathbb{R}^n \to \mathbb{R}^m$, $[Df]_a$ is a $m \times n$ matrix given by

$$[Df]_a = \left[\frac{\partial}{\partial x_j} f_i(a) \right] .$$

This is called the *Jacobian matrix* of f at a.

For some good intuition, please go to https://mathinsight.org/differentiability_multivariable_definition.

Theorem 4.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$. If the partial derivatives $\partial_j f_i$ exist near $a \in \mathbb{R}^n$ and are continuous at a, then f is differentiable at a.*

Theorem 4.3. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$. If f is differentiable at a then f is continuous at a.*

4.4 Chain rule

Theorem 4.4. *Let $f : \mathbb{R}^n \to \mathbb{R}^l, g : \mathbb{R}^m \to \mathbb{R}^n$ be differentiable functions. Then,*

$$[D(f \circ g)]_a = [Df]_{g(a)}[Dg]_a.$$

Here's a special case

Theorem 4.5. *Let $f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^m \to \mathbb{R}^n$ be differentiable functions. Then,*

$$z(y_1, \dots, y_m) = (f \circ g)(y_1, \dots, y_m)$$

is differentiable and

$$\frac{\partial z}{\partial y_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \frac{\partial g_j}{\partial y_i} .$$

4.5 Directional derivative

Definition 4.3. Let $\mathbf{u} \in \mathbb{R}^n$. The directional derivative of $f : \mathbb{R}^n \to \mathbb{R}$ at $a \in \mathbb{R}^n$ in the direction of $\mathbf{u}$ is the following limit (if exists)

$$D_{\mathbf{u}} f(a) = \lim_{h \to 0} \frac{f(a + h\mathbf{u}) - f(a)}{h} .$$

How can one compute directional derivative?

Theorem 4.6. *If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable then*

$$D_{\mathbf{u}} f(a) = \nabla f(a) \cdot \mathbf{u} .$$

4.6 Tangent planes

Let's think about tangent planes in a more systematic way, based on the definition of a plane learned in the first chapter.

Recall the c-level surface of a function $f(x, y, z)$ is the collection

$$\{(x, y, z) | f(x, y, z) = c\} .$$

Definition 4.4. The tangent plane at the point $P(x_0, y_0, z_0)$ on the c-level surface of a differentiable f is the plane through P_0, normal to $\nabla f(x_0, y_0, z_0)$.

Chapter 5

Optimization

5.1 First and second derivative tests

Read Stewart Chapter 14, Thomas Chapter 14,

We will study multivariable scalar functions

$$f : D \to \mathbb{R}\,,$$

where $D \subseteq \mathbb{R}^n$, $n \geq 2$.

Definition 5.1. A function $f : D \to \mathbb{R}$ has a **local maximum** at $\mathbf{x_0}$ if $f(\mathbf{x_0}) \geq f(\mathbf{x})$ for $\mathbf{x} \in B_\delta(\mathbf{x_0})$ for small enough δ. f has a **global maximum** at $\mathbf{x_0}$ if $f(\mathbf{x_0}) \geq f(\mathbf{x})$ for $\mathbf{x} \in D$. f has a **local (global) minimum** at $\mathbf{x_0}$ if $-f$ has a local (global) maximum at $\mathbf{x_0}$

Theorem 5.1 (First derivative test)**.** *Let $f : D \to \mathbb{R}$ be a function. If $\mathbf{x_0}$ is a local minimum and f has partial derivatives at $\mathbf{x_0}$. Then*

$$\partial_{x_i} f(\mathbf{x}_0) = 0\,.$$

The converse is not true, as having $\nabla f(\mathbf{x}_0) = \mathbf{0}$ does not mean that f has a local minimum at $\mathbf{x}_0$.

Exercise 5.1. Think of a function that the converse to the above theorem is not true.

This leads to the following notion.

Definition 5.2. $\mathbf{x}_0$ is said to be a **critical point** of $f : D \to \mathbb{R}$ if

$$\nabla f(\mathbf{x}_0) = 0$$

or one of the partial derivatives $\partial_{x_i} f(\mathbf{x}_0)$ fails to exist.

Please pay attention about the "fail to exist" condition.

Theorem 5.2 (Second derivative test for functions of 2 variables)**.** *Suppose the second partial derivatives of f are continuous near (a, b) and suppose that (a, b) is a critical point of f. Let*

$$D = f_{xx}(a, b) f_{yy}(a, b) - f_{xy}(a, b)^2\,.$$

1. *If $D > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local minimum.*
2. *If $D > 0$ and $f_{xx}(a, b) < 0$, then $f(a, b)$ is a local maximum.*
3. *If $D < 0$, then $f(a, b)$ is neither a local maximum nor local minimum.*
4. *If $D = 0$, then we cannot conclude.*

Theorem 5.3 (Extreme value theorem)**.** *If f is continuous on a* closed *and* bounded *set D. Then, f attains an absolute minimum and an absolute maximum in D.*

5.1.1 Algorithm to find absolute maxima and minima on closed bounded regions

1. Find the values of f at the critical points of f in D.
2. Find the extreme values of f on the boundary of D.
3. The largest of the values from steps 1 and 2 is the absolute maximum value; the smallest of these values is the absolute minimum value.

5.2 Constrained optimization

Constrained optimization takes various forms, depending on the assumptions. We will deal with the most straight forward form. The problem we will study is the following:

Maximize/minimize a function $f : D \to \mathbb{R}$, subject to a constraint (side condition) of the form $g(\mathbf{x}) = k$, for some constant $k \in \mathbb{R}$.

Theorem 5.4 (Method of Lagrange Multiplier)**.** *Suppose the maximum/minimum values of f exist and $\nabla g(\mathbf{x}) \neq 0$ where $g(\mathbf{x}) = k$. To find the maximum and minimum values of f subject to constraint $g(\mathbf{x}) = k$, we do the following:*

1. *Find all values of $\mathbf{x}$ and $\lambda \in \mathbb{R}$ such that*

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) \, ,$$

   and

$$g(\mathbf{x}) = k \, .$$

2. *Evaluate f at all the points $\mathbf{x}$ that result from step 1. The largest of these values is the maximum of f; the smallest is the minimum value of f.*