2018 SE102 Multivariable Calculus

Hyosang Kang (hyosang@dgist.ac.kr)
Assistant Professor, School of Undergraduate Studies
Daegu-Gyeongbuk Institute of Science and Technology (DGIST)
http://klein.dgist.ac.kr

(last modified: August 28, 2019)

Contents

1 Vectors and vector spaces

A *vector* is one of the most important object in mathematics. In this section, we give concrete concept of a vector. We also introduce abstract definition of vectors and vector spaces, and matrices.

- What *is* a vector?
- What are the examples of vectors and vector spaces?
- Why is the concept of a vector important?
- How do we understand a matrix as a vector?
- How to solve systems of linear equations by using matrices?

1.1 Vectors

Definition 1.1. A (n-dimensional) **vector** is a n-tuple of real numbers

$$\mathbf{a} = (a_1, a_2, \cdots, a_n)$$

with the following operations.

- (vector sum) For a vector $\mathbf{b} = (b_1, b_2, \cdots, b_n)$,
$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n).$$
- (scalar multiplication) For $k \in \mathbf{R}$, we define a vector $k\mathbf{a}$ as
$$k\mathbf{a} = (ka_1, ka_2, \cdots, ka_n).$$

The set of all (n-dimensional) vectors is called the (n-dimensional) **vector space**, and denoted by $\mathbf{R}^n$.

Example 1.2. Let us observe some examples of vectors.

1. Let $O = (0,0,0)$ be the origin and $P = (a_1, a_2, a_3)$ a point in 3-dimensional space. Then the arrow $\overrightarrow{OP}$ can be represented by the **position vector**
$$\overrightarrow{OP} = (a_1, a_2, a_3)$$
The scalar multiplication is a dilation, and the vector sum is a superposition. For points $Q = (b_1, b_2, b_3)$ and $R = (a_1+b_1, a_2+b_2, a_3+b_3)$, the vector $\overrightarrow{PR}$ is represented by the same vector as $\overrightarrow{OQ}$. Thus we have the following additive operation.
$$\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OP} + \overrightarrow{PR} = \overrightarrow{OR}$$
2. Let $c : (-\varepsilon, \varepsilon) \to \mathbf{R}^2$ be a parametrization of a curve on a plane. Suppose that each coordinate functions are given by differentiable function $x(t), y(t)$ defined on $(-\varepsilon, \varepsilon)$.
$$c(t) = (x(t), y(t))$$
Let us define $c'(t)$ as
$$c'(t) = (x'(t), y'(t))$$
We can consider $c'(t)$ as a vector, which represents the **velocity** at $c(t)$. The x, y-components of the vector $c'(t)$ represent the projective speed of $c(t)$ on x, y-axis respectively. We can decompose the velocity vector into the sum of horizontal and vertical velocity of $c(t)$.
$$c'(t) = (x'(t), 0) + (0, y'(t))$$
3. A function is merely an assignment of an object to another object. Such objects can either be a number, a point, or a collection of numbers (e.g. a set). For example, let us consider a function f sending every point in $\mathbf{R}^2$ to a point in $\mathbf{R}^3$ as follows.
$$f(u, v) = (x(u, v), y(u, v), z(u, v))^1 \qquad (1.2.1)$$
The variables x, y, z are dependent to the variables u, v. The functions f (and also x, y, z as functions) contain(s) two or more independent or dependent variables. Such functions are are called **multivariable functions** (cf. Definition 2.1).

In Definition 1.1, we defined a vector as an ordered collection of real numbers. Thus, if we write $\mathbf{a} = (u, v)$ and $\mathbf{b} = (x, y, z)$, the equation (1.2.1) can be written as
$$f(\mathbf{a}) = \mathbf{b}$$
A multivariable function can be viewed as a *single variable* function which assigns a vector to another vector[2]. Note that we did not assume that the *vector sum* and *scalar multiplication* properties of vectors (cf Definition 1.1) are preserved. Thus it is hard to say that f is a function from a vector space to another vector space[3]. However, we will still write f as $f : \mathbf{R}^2 \to \mathbf{R}^3$, where $\mathbf{R}^2, \mathbf{R}^3$ as *sets of all points* in $2, 3$-dimensional plane respectively.

Definition 1.3. For a (n-dimensional) vector $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, the value

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

is called the **norm** of $\mathbf{a}$. A vector with norm 1 is called a **unit** vector. The norm $\|\mathbf{a}\|$ is 0 if and only if $\mathbf{a}$ is a **zero vector**, that is,

$$\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a}$$

for any vector $\mathbf{a}$. For nonzero vector $\mathbf{a}$,

$$\mathbf{u} = \mathbf{a}/\|\mathbf{a}\|$$

is called the **normalization** of $\mathbf{a}$.

Example 1.4. For each $i = 1, \cdots, n$, the vector

$$\mathbf{e}_i = (0, \cdots, 1, \cdots, 0)$$

[1] Later, we will see that the graph (cf. Example **??**) of such f is called a *surface* in $\mathbf{R}^3$.

[2] By this reason (and more), the Multivariable Calculus is often called the *Vector Calculus*.

[3] We need the function to be *linear* (cf. Definition **??**).

with 1 at ith place is called a (unit) **basis vector**. Especially, 3-dimensional basis vectors are denoted by

$$\mathbf{i} = (1,0,0), \quad \mathbf{j} = (0,1,0), \quad \mathbf{k} = (0,0,1)$$

A vector $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ can be decomposed as a linear sum of basis vectors.

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n$$

Thus the set

$$\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$$

is called the **basis**, and generates all n-dimensional vectors.

Definition 1.5. The **inner product** (also called **dot product**) of two vectors **a**, **b** is an operation defined by

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

If $\mathbf{a} \cdot \mathbf{b} = 0$, then we say **a**, **b** are **orthogonal**.

Proposition 1.6. *The inner product satisfies the following.*

1. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$
2. $\mathbf{a} \cdot (k\mathbf{b}) = k(\mathbf{a} \cdot \mathbf{b})$
3. $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$

Theorem 1.7. *Let* **a**, **b** *be nonzero* 2*-dimensional vectors. If* θ *is the angle between* **a** *and* **b**, *then*

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cdot \cos\theta$$

Example 1.8. Let us parametrize a line l in $\mathbf{R}^3$.

1. Suppose that l passes through $P = (x_0, y_0, z_0)$ and parallel to $\mathbf{a} = (a_1, a_2, a_3)$. Then the paramtric equations of l is

$$l(t) = P + t\mathbf{a} \tag{1.8.1}$$

   Equivalently, equation (1.8.1) can be written as a *symmetric form* as follows.

$$\frac{x - x_0}{a_1} = \frac{y - y_0}{a_2} = \frac{z - z_0}{a_3} \tag{1.8.2}$$

2. Suppoe that l passes through two points P, Q. By subsitute $\mathbf{a} = \overrightarrow{OP} - \overrightarrow{OQ}$ to (1.8.1) or (1.8.2), we get a parametric equation of l.

Remark 1.9. As norm measures the *size* of a vector, the inner product measure the *direction*. For example, the direction of a 2-dimensional vector **a** is determined by $0 \leq \theta_1, \theta_2 \leq \pi$ satisfying

$$\mathbf{a} \cdot \mathbf{e}_1 = \|\mathbf{a}\| \cos\theta_1$$

$$\mathbf{a} \cdot \mathbf{e}_2 = \|\mathbf{a}\| \cos\theta_2$$

In order to determine the direction of a 3-dimensional vector, say **a**, we need three angles $0 \leq \alpha, \beta, \gamma \leq \pi$ satisfying

$$\cos\alpha = \frac{\mathbf{a} \cdot \mathbf{i}}{\|\mathbf{a}\|}, \quad \cos\beta = \frac{\mathbf{a} \cdot \mathbf{j}}{\|\mathbf{a}\|}, \quad \cos\gamma = \frac{\mathbf{a} \cdot \mathbf{k}}{\|\mathbf{a}\|} \tag{1.9.1}$$

Such quantities are called the **direction cosines**.

Definition 1.10. For nonzero vectors **a** and **b** with the same dimension, the vector defined by

$$\mathrm{proj}_{\mathbf{b}}\mathbf{a} = \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}\right)\mathbf{b}$$

is called the **projection of a onto b**.

1.2 Matrices

Definition 1.11. A $n \times m$ (n-by-m) **matrix** is a collection of nm numbers (or functions) arranged in the following way.

$$A = (a_{ij})_{n\times m} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

The indices i, j of an entry a_{ij} represents the row and column indices respectively.

Example 1.12.

1. A $n \times m$ matrix is called **square** matrix if $n = m$.

2. If A is a square matrix and $a_{ij} = 0$ for all $i \neq j$, then A is called **diagonal**.

$$A = \begin{pmatrix} a_{11} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & a_{nn} \end{pmatrix}$$

3. If the diagonal entries of a diagonal matrix are all 1, then it is called the **identity matrix**, and denoted by I_n.

$$I_n = \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix}$$

4. If A is a square matrix and $a_{ij} = 0$ for all $i < j$ ($i < j$, respectively), then A is called **lower triangle** (**upper triangle**, respectively).

$$\begin{pmatrix} * & & \mathbf{0} \\ & \ddots & \\ * & & * \end{pmatrix}, \quad \begin{pmatrix} * & & * \\ & \ddots & \\ \mathbf{0} & & * \end{pmatrix}$$

5. If every entry is 0, then it is called the **zero matrix**, and denoted by $\mathbf{0}$.

Definition 1.13. Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times m$ matrix. Then we define

$$A + B = (a_{ij} + b_{ij})$$

$$k \cdot A = (ka_{ij})$$

Let C be a $m \times l$ matrices, then $A \cdot C$ is a $n \times l$ matrix whose entries are

$$A \cdot B = (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj})_{1\leq i\leq n, 1\leq j\leq l}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1l} \\ c_{21} & \cdots & c_{2j} & \cdots & c_{2l} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{ml} \end{pmatrix}$$

Example 1.14. There is more to the matrix multiplication than meets the eye. For example, suppose Bob, Larry, and Joanna worked in a fruits store for three days. Table 1 shows *how many* fruits each sold in total, and Table 2 shows *how much* was the fruits on each day. To compute the *total revenue* each person sold on each day, we do the matrix multiplication as below. The i, j-entries of the resulting matrix As a result represents the total revenue sold by the person i at the day j (cf. [1]).

$$\begin{pmatrix} 38 & 25 & 10 \\ 15 & 22 & 15 \\ 8 & 70 & 27 \end{pmatrix} \begin{pmatrix} 1.19 & 1.45 & 0.99 \\ 1.70 & 0.99 & 2.1 \\ 2.19 & 3.5 & 1.29 \end{pmatrix}$$

	Apple	Orange	Banana
Bob	38	25	10
Larry	15	22	15
Joanna	8	70	27

Table 1: The volumn of sales of each person per items

	Day1	Day2	Day3
Apple	\$1.19	\$1.45	\$.99
Orange	\$1.70	\$0.99	\$2.1
Banana	\$2.19	\$3.5	\$1.29

Table 2: The prices of items per day

A deeper, and more mathematical, thus important meaning of matrix multiplication will be discovered when we visit in the *chain rule*(cf. §5.3) in later sections.

Proposition 1.15. *Let A, B, C be matrices. Whenever the operations are valid, the following holds.*

1. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$
2. $A \cdot (B + C) = A \cdot B + A \cdot C$
3. $(B + C) \cdot A = B \cdot A + C \cdot A$
4. $k \cdot (A \cdot B) = (k \cdot A) \cdot B = A \cdot (k \cdot B)$

Moreover, the ***transpose*** *of a matrix $A = (a_{ij})$ defined by*

$$A^T = (a_{ji})$$

satisfies the following.

5. $\left(A^T\right)^T = A$

6. $k \cdot A^T = (k \cdot A)^T$

7. $(A+B)^T = A^T + B^T$

8. $(AB)^T = B^T A^T$

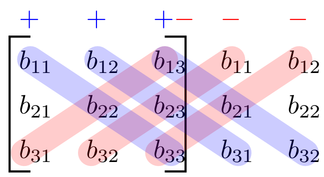
Definition 1.16. The **determinant** of 2×2 matrix A is defined as follows.

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

The **determinant** of 3×3 matrix B is defined as follows.

$$\begin{aligned}\det B &= \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} \\ &= b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} \\ &\quad - b_{13}b_{22}b_{31} - b_{11}b_{23}b_{32} - b_{12}b_{21}b_{33}\end{aligned}$$

An easy way to remember the formula is the following.



Definition 1.17. Let $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ be two 3-dimensional vectors. The **cross-product** of $\mathbf{a}, \mathbf{b}$ is a vector defined by

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$$

An easy way to remember the formula is the following.

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Proposition 1.18. *Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be 3-dimensional vectors and k a constant. The following identity holds.*

1. $\mathbf{a}\times\mathbf{0} = \mathbf{0}\times\mathbf{a} = \mathbf{0}$

2. $\mathbf{a}\times\mathbf{a} = \mathbf{0}$

3. $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$

4. $\mathbf{a}\times(k\mathbf{b}) = (k\mathbf{a})\times\mathbf{b} = k(\mathbf{a}\times\mathbf{b})$

5. $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b} + \mathbf{a}\times\mathbf{c}$

6. $\mathbf{a}\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\mathbf{a}\times\mathbf{b}) = 0$ *(This shows that the cross product $\mathbf{a}\times\mathbf{b}$ is normal to the plane spanned by $\mathbf{a}$ and $\mathbf{b}$.)*

7. $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{b}\cdot\mathbf{c})\mathbf{a}$

Theorem 1.19. *Let $\mathbf{a}, \mathbf{b}$ be two 3-dimensional vectors. Then*

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

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$.

Proposition 1.20. *Let us denote $|A|$ by the determinant of a matrix A. Then the following holds*

1. *If A has a row (or a column) whose entries are all zero, then $|A| = 0$.*

2. *Let B be the matrix obtained by interchanging two rows (or columns) of A. Then $|B| = -|A|$.*

3. *Let B be a matrix obtained by multiply c on a row (or column) followed by adding it to another row (or column). Then $|B| = |A|$.*

Remark 1.21. In general, the determinant of $n\times n$ matrix is defined as a function

$$\det : M(n;\mathbf{R}) \longrightarrow \mathbf{R}$$

satisfying the three properties of Proposition 1.20. One can prove that such function is unique, and the computation is inductively defined. A more detailed argument can be found in [2].

Definition 1.22. Let A be a $n\times n$ matrix. A matrix B satisfying

$$A\cdot B = B\cdot A = I_n$$

is called the **inverse of** A, denoted by $B = A^{-1}$. If an inverse matrix A^{-1} exists, then A is said to be **non-singular**. Otherwise, it is called **singular**.

Theorem 1.23. *A matrix A is singular if and only if $\det A = 0$.*

Proposition 1.24. *If A is a 2×2 matrix, then A^{-1} is*

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}$$

For a 3×3-matrix B, the inverse is given by

$$B^{-1} = \frac{1}{\det B}\begin{pmatrix} c_{11} & -c_{21} & c_{31} \\ -c_{12} & c_{22} & -c_{32} \\ c_{13} & -c_{23} & c_{33} \end{pmatrix} \tag{1.24.1}$$

*where each c_{ij}, called the **cofactor**, is the determinant of 2×2-matrix obtained by deleting ith row and jth column. For example,*

$$c_{21} = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = \begin{vmatrix} b_{12} & b_{13} \\ b_{32} & b_{33} \end{vmatrix}$$

Notice that the row and column indices are switched in (1.24.1).

1.3 Vector spaces

In this section, we investigate abstract notion of vector spaces.

Definition 1.25. A **vector space** V is a set of element called **vectors** satisfying the following properties:

1. (Zero vector) V contains the **zero vector** $\mathbf{0}$, which is a unique vector satisfying $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.

2. (Vector sum) For any two vectors $\mathbf{v}$, $\mathbf{w} \in V$, the vector $\mathbf{v} + \mathbf{w}$ lies in V.

3. (Scalar multiplication) For any $k \in \mathbf{R}$ and $v \in V$, the vector $k\mathbf{v}$ lies in V.

A subset $V \subset \mathbf{R}^n$ is called a **vector subspace** if it is a vector space itself.

Example 1.26. Let us observe some examples of vector spaces.

1. Let $\mathcal{C}^1(\mathbf{R})$ be the set of all differentiable functions on $\mathbf{R}$ whose derivatives are continuous on $\mathbf{R}$. Since f, g are such functions so is the function $h = f + g$. Also for any $k \in \mathbf{R}$, the function kf is differentiable and its derivative is continuous. Thus $\mathcal{C}^1(\mathbf{R})$ is a vector space. Likewise, we can define vectors spaces $\mathcal{C}^n(\mathbf{R})$, $\mathcal{C}^\infty(\mathbf{R})$[4].

2. For each constant $k \in \mathbf{R}$, let V_k be the set of all points on the line $y = kx$ in $\mathbf{R}^2$.

$$V_k = \{(x, y) \mid y = kx\}$$

   Let us identify each point with position vector as in Example 1.2. Then V_k is a vector subspace of $\mathbf{R}^2$. Let V_∞ be the vertical line $x = 0$. Then V_∞ is also a vector subspace of $\mathbf{R}^2$[5]

3. Let P be the set of all point on the plane

$$ax + by + cz = 0$$

   in $\mathbf{R}^3$. By identifying points in P as position vectors, we can say P is a vector subspace of $\mathbf{R}^3$, orthogonal to $\mathbf{n} = (a, b, c)$ (cf. Exercise 3).

4. Let V, W be vector subspace of $\mathbf{R}^n$. Then so is $V \cap W$. For example, let V, W be vector subspace for two planes in $\mathbf{R}^3$ passing through the origin $\mathbf{0}$. Then $V \cap W$ is either a line (if V, W are transversal) or a plane (if $V = W$).

Definition 1.27. Let V a vector space and $\mathbf{v}_1, \cdots, \mathbf{v}_m \in V$. We say the vector space subspace

$$W = \{a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m \mid a_1, \cdots, a_m \in \mathbf{R}\}$$

is **spanned** by $\mathbf{v}_1, \cdots, \mathbf{v}_m$, and denote by $W = \text{span}\langle \mathbf{v}_1, \cdots, \mathbf{v}_m \rangle$.

Example 1.28. In Definition 1.1, we defined the vector space $\mathbf{R}^n$ to be the set of all position vectors of the form $(x_1, \cdots, x_n)$. Note that this vector can be decomposed as follows.

$$\begin{aligned}(x_1, \cdots, x_n) &= (x_1, 0, \cdots, 0) + \cdots + (0, \cdots, 0, x_n) \\ &= x_1(1, 0, \cdots, 0) + \cdots + x_n(0, \cdots, 0, 1)\end{aligned}$$

Let $\mathbf{e}_i = (0, \cdots, 1, \cdots, 0)$ where all the coordinates are zero except the i-th coordinate which is 1. We can immediately see that the vector $(x_1, \cdots, x_n)$ can be written as a linear combination of n unit vectors $\mathbf{e}_1, \cdots, \mathbf{e}_n$, and such expression is unique.

Definition 1.29. We say a vector space V is **spanned by** the set of vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$, if every $v \in V$ can be expressed a linear combination of v_i's:

$$v = a_1\mathbf{v}_1 + \cdots a_n\mathbf{v}_n$$

The set $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is said to be **linearly independent** if the coefficients satisfying

$$a_1\mathbf{v}_1 + \cdots a_n\mathbf{v}_n = \mathbf{0}$$

is the trivial ones, namely $a_1 = \cdots = a_n = 0$.

Definition 1.30. Let $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ be the set of linear independent vectors which spans the vector space V. Such set is called the **basis** of V, and its element is called the **basis vector**. The **dimension** of V is the number of basis vectors which spans V.

Remark 1.31. The linear independency allows us to define the *dimension* of a vector space. Given a vector space, There are (infinitely) many choices of basis. However, the size of basis, i.e. the dimension of the vector space does not depend on the choice of basis. (Exercies 4). The linear independency also plays an important role in *rank* of a matrix.

---

[4] In Analysis, we introduce a l_p-norm on $\mathcal{C}^n(\mathbf{R})$ as follows:

$$\|f\|_p = \int_{\mathbf{R}} |f(x)|^p dx$$

This a standard example of a *Hilbert space* which is an important object in Quantum Physics.

[5] The set of all such subspaces is called a *real projective space of dimension* 1 and denoted by

$$\mathbf{RP}^1 = \{V_k \mid k \in \mathbf{R}\} \cup \{V_\infty\}$$

Similarly, we can construct a real projective space dimension n, by collecting all 1-dimensional subspaces, i.e. lines in $\mathbf{R}^{n+1}$, and it gives a new kinds of space whose *topology* is quite different from $\mathbf{R}^n$. (Can you imagine the shape of $\mathbf{RP}^2$?)

Definition 1.32. Given a $n \times n$ matrix A, let $a_1, \cdots, \mathbf{a}_n$ be the column vectors of $m \times n$ matrix A. The dimension of $V = \mathrm{span}\langle \mathbf{a}_1, \cdots \mathbf{a}_n \rangle$ is called the **rank** of A, and denoted by $\mathrm{rank} A$[6].

Proposition 1.33. *The determinant of A is nonzero if and only if the set of all row (or column) vectors of A is linearly independent.*

[6] One can define a rank of a matrix using row vectors. Thus we may distinguish the *row rank* from the *column rank*. A remarkable fact is that such quantity always coincides: Let $\mathbf{a}^1, \cdots, \mathbf{a}^m$ be the row vectors of the matrix A in Definition 1.32. Then

$$\mathrm{rank} A = \dim \mathrm{span}\langle \mathbf{a}^1, \cdots, \mathbf{a}^m \rangle$$

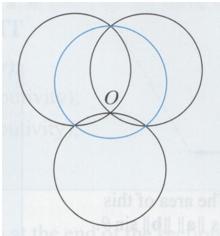
Since the proof is beyond our understanding, so we suggest to read [2]

1.4 Exercies

Most of exercise problems are taken from [?].

Vectors

1. Suppose that three circles of equal radius r intersects in a single point O. Let A, B, C be the remaining intersection points .



   (a) Show that A, B, C lie on a circle of the same radius r.

   (b) Show that O is the orthocenter(수심) of triangle ΔABC.

2. (a) Show that the vectors $\|\mathbf{b}\|\mathbf{a} + \|\mathbf{a}\|\mathbf{b}$ and $\|\mathbf{b}\|\mathbf{a} - \|\mathbf{a}\|\mathbf{b}$ are orthogonal.

   (b) Show that $\|\mathbf{b}\|\mathbf{a} + \|\mathbf{a}\|\mathbf{b}$ bisects the angle between $\mathbf{a}$ and $\mathbf{b}$.

3. Show that the direction cosine α, β, γ satisfy

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

4. Let $\mathbf{a}$, $\mathbf{b}$ be nonzero 2-dimensional vectors which are not parallel each other.

   (a) Show that the vector $\mathrm{proj}_{\mathbf{b}}\mathbf{a}$ is the parallel to $\mathbf{b}$.

   (b) Show that the vector $\mathbf{a} - \mathrm{proj}_{\mathbf{b}}\mathbf{a}$ is orthogonal to $\mathbf{b}$.

5. Let $\mathbf{a} = (1, 2, 1)$, $\mathbf{b} = (2, 1, 2)$, and $\mathbf{u} = (0, 1, -1)$. Suppose that the vector $\mathbf{u}$ can be decomposed by

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$$

   where $\mathbf{u}_1$ is parallel to $\mathbf{a}$, $\mathbf{u}_2$ is parallel to $\mathbf{b}$, and $\mathbf{u}_3$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. Find the vector $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ explicitly.

6. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors in $\mathbf{R}^3$.

   (a) Show that the area of parallelepiped bounded by $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is

$$|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$$

   (b) Let $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$. Show that the area of parallelogram bounded by $\mathbf{a}, \mathbf{b}$ is

$$\sqrt{(a_2b_3 - a_3b_2)^2 + (a_1b_3 - a_3b_1)^2 + (a_1b_2 - a_2b_1)^2}$$

7. Find the distance between $P = (2, 1, 3)$ and the line $l(t) = (2, 3, -2) + t(-1, 1, -2)$.

8. Find the distance between two parallel planes

$$2x - 2y + z = 5, \quad 2x - 2y + z = 20$$

Matrices

9. Prove that

$$\det B = b_{11}\begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} - b_{12}\begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix} + b_{13}\begin{vmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix}$$

10. Evaluate the determinant of

$$\begin{pmatrix} 1 & 0 & -1 \\ -2 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix}$$

11. Verify the followings.

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

12. Find the inverse matrix of

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Parametric equations

1. Find the distance between two skew lines

$$l_1(t) = (0, 5, -1) + t(2, 1, 3)$$
$$l_2(t) = (-1, 2, 0) + t(1, -1, 0)$$

2. Suppose that H is parallel to the linearly independent vectors $\mathbf{u}, \mathbf{v}$, and passes through the point $P = (x_0, y_0, z_0$. Show that the parametric equations for $(x, y, z) \in H$ is

$$(x, y, z) = (x_0, y_0, z_0) + s\mathbf{u} + t\mathbf{v} \qquad (1.33.1)$$

   where s, t are parameters.

3. Suppose that H passes through the point $P = (x_0, y_0, z_0)$ and perpendicular to the vector $\mathbf{u} = (u_1, u_2, u_3)$. Then for every point $X = (x, y, z) \in H$, $\overrightarrow{XP}$ is perpendicular to $\mathbf{u}$, i.e.

$$\mathbf{u} \cdot \overrightarrow{XP} = 0$$

   Show that the parametric equation for H is

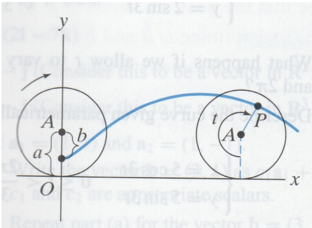
$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad (1.33.2)$$

   and rewrite with parameters s, t.

4. Give a set of parametric equations for the line with symmetric form

$$\frac{x+5}{5} = \frac{y-1}{7} = \frac{z+10}{-2}$$

5. Find a set of parametric equation for the curtate cycloid with $a = 3$, $b = 2$.



6. Suppose that H contains three distinct points A, B, C in $\mathbf{R}^3$, which do not lie on the same straight line. Then the vector $AB \times AC$ is the perpendicular to H. By setting $P = A$, and $\mathbf{u} = \overrightarrow{AB}$, and using (1.33.2), obtain the parametric equation for H.

7. Suppose that $l_1(t) = \mathbf{a}_1 t + \mathbf{b}_1$ and $l_2(t) = \mathbf{a}_2 t + \mathbf{b}_2$, and l_1, l_2 are skewed, i.e. $\mathbf{a}_1, \mathbf{a}_2$ are not parallel. Show that the distance D between l_1, l_2 is

$$D = \frac{|(\mathbf{a}_1 \times \mathbf{a}_2) \cdot (\mathbf{b}_1 - \mathbf{b}_2)|}{\|\mathbf{a}_1 \times \mathbf{a}_2\|}$$

Challenging problem

1. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be 3-dimensional vectors. Prove that
   (a) $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})$
   (b) $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} + (\mathbf{b}\times\mathbf{c})\times\mathbf{a} + (\mathbf{c}\times\mathbf{a})\times\mathbf{b} = 0$
   (c) $(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{d}) = [\mathbf{a}\cdot(\mathbf{c}\times\mathbf{d})]\mathbf{b} - [\mathbf{b}\cdot(\mathbf{c}\times\mathbf{d})]\mathbf{a}$
   (d) $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{b}\times\mathbf{c})\times(\mathbf{c}\times\mathbf{a}) = [\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})]^2$

2. Show that in a $n \times n$ matrix A, if any row (or column) is a constant multiple of another row (or column), then $\det A = 0$.

3. Show that the set of all $n \times m$ matrices

$$M(n, m; \mathbf{R}) = \{A = (a_{ij}) \,|\, 1 \le i \le n, 1 \le j \le m\}$$

can be viewed as a vector space $\mathbf{R}^{nm}$.

4. Show that the dimension of a vector space does not depend on the choice of basis.

2 Multiple integrals

In this section, we introduce the notion of integrals on a multivariable functions. As in single variable case, the integration of multivariable function can be defined by formal Riemann sum. We focus on the concept of integrals on functions and vector fields. (The techniques of multiple integration will be covered in Section §**??**.) The following questions may arise from this section:

- How can we generalize the integration on single variable functions?
- What is the meaning of double and triple integrals?
- How can we integrate function over planar surfaces?
- What is a flux? How can we compute flux over planar surfaces?

2.1 Multivariable functions

The figures in this section are obtained by WolframAlpha©.

Definition 2.1. A function is called **multivariable** if it consists of more than two independent or dependent variables.

In general a multivariable function f consists of n independent variables and m dependent variables.

$$f(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_m) \qquad (2.1.1)$$

The variable y_j is the dependent variableof a function f_j with n independent variables $x_1, x_2, \cdots, x_n$. Thus we can also write the function f as m-tuple of real-valued function f_j's. $(j = 1, \cdots, m)$

$$y_j = f_j(x_1, x_2, \cdots, x_n) \qquad (2.1.2)$$

Definition 2.2. Let $f(x, y) = z$ be a function defined on a set $D \subset \mathbf{R}^2$. The **graph** of f is the set in $\mathbf{R}^3$ defined by

$$G(f) = \{(x, y, f(x, y)) \,|\, (x, y) \in D\}$$

Example 2.3. Here are some examples of multivariable functions and their graphs.

1. A *line* l in $\mathbf{R}^3$ is parametrized by a function $l : \mathbf{R} \to \mathbf{R}^3$ defined as follows. Let l passes through a point $P \in \mathbf{R}^3$ with direction v. Let $\mathbf{a} = \overrightarrow{OP}$ be the position vector of the point P. Then each point $l(t)$ on the line l can be parameterized by $t \in \mathbf{R}$ as follows.
$$l(t) = \mathbf{a} + \mathbf{v}t$$
2. A *plane* H in $\mathbf{R}^3$ is parametrized by a function $H : \mathbf{R}^2 \to \mathbf{R}^3$ defined as follows. There are three different ways to formulate the function H.
   (a) Suppose that H passes through a point $P = (x_0, y_0, z_0) \in \mathbf{R}^3$ and is normal to the vector $\mathbf{n} = (a, b, c)$. Let $\overrightarrow{OP} = (x_0, y_0, z_0)$ be the position vector of the point P. If a point $Q = (x, y, z)$ lies on H, then the coordinate x, y, z must satisfies the following equation.
$$(\overrightarrow{OQ} - \overrightarrow{OP}) \cdot n = 0$$
$$(x - x_0, y - y_0, z - z_0) \cdot (a, b, c) = 0$$
$$ax + by + cz = ax_0 + by_0 + cz_0$$
   Since $\mathbf{n} \neq \mathbf{0}$, one of a, b, c is nonzero. Suppose $c \neq 0$. Then
$$H : \mathbf{R}^2 \to \mathbf{R}^3$$
$$(x, y) \mapsto (x, y, (cz_0 - a(x - x_0) - b(y - y_0))/c)$$
   is a parametrization of the plane H.

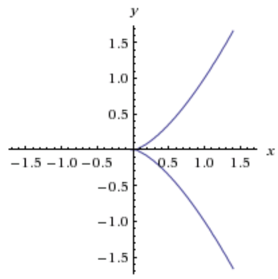


Figure 2.1: A 2-dimensional curve

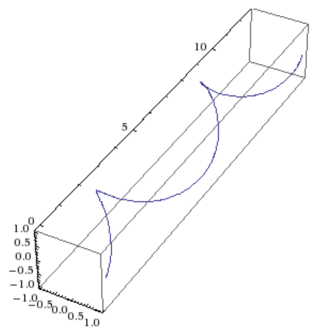


Figure 2.2: A 3-dimensional curve

   (b) Suppose that H passes through the point $P = (x_0, y_0, z_0)$ and is parallel to two vectors $\mathbf{v} = (a_1, a_2, a_3)$, $\mathbf{w} = (b_1, b_2, b_3)$. Then the vector $\mathbf{n} = \mathbf{v} \times \mathbf{w}$ is normal to H and by using the first method, we can find the parametric equation of H.
   (c) Let H be a plane passing through three points P, Q, R which are not lying on the same line. Then $\mathbf{v} = \overrightarrow{PQ}$ and $\mathbf{w} = \overrightarrow{PR}$ are the vectors parallel to H, and we can use the second method.
3. Let $I \subset \mathbf{R}$ be an interval. A *parametric curve* is a function $c : I \to \mathbf{R}^n$ of the following form.
$$c(t) = (x_1(t), \cdots, x_n(t)), \quad t \in I$$
   The graph of $c(t)$ is often called the *trajectory* of a curve.
   (a) Figure 2.1 shows the graph of $f(t) = (t^2, t^3)$.
   (b) Figure 2.2 shows the graph of $f(t) = (\cos t, t, \sin t)$.
4. A *surface* is the graph of a function $z = f(x, y)$.
   (a) The graph of $z = x^2 - y^2$ is called *saddle*. (Figure 2.3)
   (b) The graph of $z = x(x^2 - y^2)$ is called *Monkey's saddle*. (Figure 2.4)

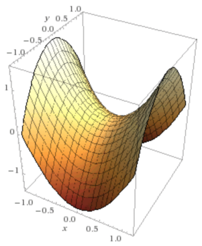


Figure 2.3: A saddle

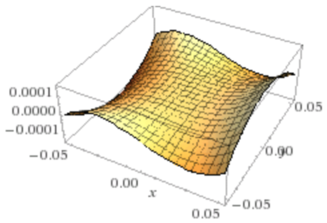
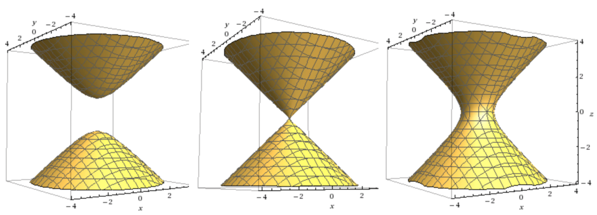


Figure 2.4: A monkey's saddle

5. We cannot draw the graph of a function $w = f(x, y, z)$ with three independent variables. (We need 4-dimensional space!) We use the **level set** instead.

$$L_c(f) = \{(x, y, z) \in \mathbf{R}^3 \mid f(x, y, z) = c\} \tag{2.3.1}$$

The followings are the level sets of $f(x, y, z) = x^2 + y^2 - z^2$ with $c = -1, 0, 1$.



(a) $c = -1$ (b) $c = 0$ (c) $c = 1$

Figure 2.5: The level sets

We will observe two important example of multivariable functions: a *vector field* and a *linear transformation*.

Definition 2.4. Let V be a vector space and D be a open subset of $\mathbf{R}^n$. A **vector field** $\mathbf{F} : D \to V$ is a function which assigns each point $(x_1, \cdots, x_n) \in D$ a vector

$$\mathbf{F}(x_1, x_2, \cdots, x_n) \in V.$$

For example, if $V = \mathbf{R}^m$, then $\mathbf{F}$ is simply written as

$$\mathbf{F}(x_1, \cdots, x_n) = (f_1, \cdots, f_m) \tag{2.4.1}$$

where each $f_i : \mathbf{R}^n \to \mathbf{R}$ is a real-valued function. One might think there is no difference between Equation 2.4.1 and the usual expression of multivariable function as in Equation 2.1.1 and 2.1.2. Beyond the syntax, we must remember that the value of $\mathbf{F}$ is now a *vector*, thus we can take a *vector addition* and *scalar multiplicaion* on this values. For example, if $\mathbf{G} : \mathbf{R}^n \to V$ is another vector fields, we can define a new vector field $\mathbf{H} = \mathbf{F} + \mathbf{G}$ or $\mathbf{H} = c\mathbf{F}$ for some constant $c \in \mathbf{R}$.

Suppose $D = \mathbf{R}^n$ and $V = \mathbf{R}^n$. The vector field $\mathbf{F} : \mathbf{R}^n \to \mathbf{R}^n$ cannot be composed with itself, because the target space is a vector space and the domain is an ordinary n-dimensional space. (There is no *vector addition* and *scalar multiplicaion* in the domain!) By replacing the domain as a vector space, we get a definition of a *linear transformation*.

Definition 2.5. Let V, W be vector spaces. A function $T : V \to W$ is called a **linear transformation** if it satisfies the following.

1. $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$.
2. $T(cv) = cT(v)$ for all $v \in V$ and $c \in \mathbf{R}$.

A linear transformation is a map between vector spaces which preserves the *vector addition* and *scalar multiplicaion*[7] Every linear transformation can be written as a matrix in the following way. Let $\mathcal{B} = \{v_1, \cdots, v_n\}$ and $\mathcal{B}' = \{w_1, \cdots, w_m\}$ be basis of V and W. For each $i = 1, \cdots, n$, the vector $T(v_i)$ can be written as linear combinations of w_j's, namely.

$$T(v_i) = a_{1i}w_1 + a_{2i}w_2 + \cdots a_{mi}w_m \tag{2.5.1}$$

Let us identify T as a following $m \times n$ matrix.

$$[T]_{\mathcal{B}\to\mathcal{B}'} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.5.2}$$

The notation $[\,]_{\mathcal{B}\to\mathcal{B}'}$ simply indicate such expression depends on the choice of basis $\mathcal{B}, \mathcal{B}'$. We will explain why this makes sense. Let $v = x_1v_1 + x_2v_2 + \cdots x_nv_n$ be a vector in V. Since we fixed the basis $\mathbf{B}$ for V, we may identify v as a n-dimensional column vector

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{2.5.3}$$

[7] *spoiler alert!* The derivatives of multivariable functions are considered to be linear transformations. (cf. §??)

Again, the notation $[\,]_{\mathcal{B}}$ indicates the expression depends on the choice of $\mathcal{B}$. Let $w = T(v)$. From Definition 2.5 and Equation 2.5.1, we get

$$\begin{aligned} w &= T(x_1v_1 + x_2v_2 + \cdots + x_nv_n) \\ &= x_1T(v_1) + x_2T(v_2) + \cdots + x_nT(v_n) \\ &= x_1(a_{11}w_1 + \cdots + a_{m1}w_m) + \\ &\qquad \cdots + x_n(a_{1n}w_1 + \cdots + a_{mn}w_m) \\ &= (a_{11}x_1 + a_{1n}x_n)w_1 + \cdots (a_{m1}x_1 + \cdots + a_{mn}x_n)w_m \end{aligned}$$

Thus we can identify the vector w with the following column vector.

$$[w]_{\mathcal{B}'} = \begin{bmatrix} a_{11}x_1 + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} \tag{2.5.4}$$

Now compare the expression $w = T(v)$ with the following matrix multiplication:

$$[w]_{\mathcal{B}'} = [T]_{\mathcal{B}\to\mathcal{B}'}[v]_{\mathcal{B}}$$

(By substituting Equation (2.5.2), (2.5.3), and (2.5.4), we see that the equality holds.)

The upshot of such matrix identification of the linear transformation is that the composition becomes simplely the matrix multiplication. Let V, W, U be vector spaces and $T : V \to W$, $S : W \to U$ be linear transformations. Let $\mathcal{B}, \mathcal{B}', \mathcal{B}$ be basis of V, W, U respectively. Then the composition $S \circ T : V \to U$ is represented by

$$[S \circ T]_{\mathcal{B}\to\mathcal{B}''} = [S]_{\mathcal{B}'\to\mathcal{B}''}[T]_{\mathcal{B}\to\mathcal{B}'}$$

2.2 Multiple integrals

Definition 2.6. (Iterated integrals) Let $f(x,y)$ be a two variable function define on a rectangular domain $D=[a,b]\times[c,d]$. The **iterated integral** $\int_c^d\int_a^b f(x,y)dxdy$ on D is defined as follows.

$$\int_c^d \underbrace{\left[\int_a^b f(x,y)dx\right]}_{\text{consider } y \text{ as a constant}} dy.$$

Iterated integral is merely an integration of two or more variable function by its variables in ordered way. Note that once we take the integral over x, there is no x term in the expression $\int_a^b f(x,y)dx$. Thus we can put

$$g(y)=\int_a^b f(x,y)dx$$

The geometric meaning of $g(y)$ is the following. Let us consider a graph S of $z=f(x,y)$. Take a section of the surface S at $y=y_0\in[c,d]$, which is a graph of a single variable function $z=f(x,y_0)$. $g(y_0)$ the integral of $h(x)=f(x,y_0)$ over $[a,b]$ in this section. Now the function $z=g(y)$ is, again, a single variable function, and the iterated integral $\int_c^d\int_a^b f(x,y)dxdy$ is simply integration of $g(y)$ over $[c,d]$.

$$\int_c^d\int_a^b f(x,y)dxdy=\int_c^d g(y)dy$$

The iterated integral is not a generalization of integration of single variable function in general. Let us introduce how to integrate two variable function over a rectangular domain $[a,b]\times[c,d]\in\mathbf{R}^2$.

Definition 2.7. Let $f(x,y)$ be a function defined on a rectangular region $D=[a,b]\times[c,d]$. Let us subdivide the intervals $[a,b]$ (respectively $[c,d]$) by n (m, respectively) intervals.

$$a=x_0<x_1<\cdots<x_n=b$$

$$c=y_0<y_1<\cdots<y_m=d$$

so that the region D is subdivided by nm rectangular regions $D_{ij}=[x_{i-1},x_i]\times y_{i-1},y_i]$. For each $i=1,\cdots,n$ and $j=1,\cdots,m$, let us choose a point $(x_i^*,y_j^*)\in D_{ij}$ and set

$$\Delta x_i=x_i-x_{i-1},\quad \Delta y_j=y_j-y_{j-1}$$

The sum

$$\sum_{i=1}^n\sum_{j=1}^m f(x_i^*,y_j^*)\Delta x_i\Delta y_j \tag{2.7.1}$$

is called the **Riemann sum** of $f(x,y)$ with respect to the subdivision D_{ij}'s. Let us take the limit by taking further subdivision so that $\Delta x_i,\Delta y_j\to\infty$. (This requires $n,m\to\infty$). If the limit exists, we denote it by $\iint_D fdA$.

$$\iint_D fdA=\lim\sum_{i=1}^n\sum_{j=1}^m f(x_i^*,y_i^*)\Delta x_i\Delta y_j \tag{2.7.2}$$

This is called the **double integral** of f over D.

The double integral $\iint_D fdA$ is the volume under the graph of $z=f(x,y)$ and Figure [?] shows the approximation process of Riemann sum (2.7.1). Each summand in the Riemann sum represents the volume of a cube over the subdomain $[x_{i-1},x_i]\times[y_{j-1},y_j]$ with height $f(x_i^*,y_j*)$. This approximates the volume under the graph of $z=f(x,y)$ over the subdomain $[x_{i-1},x_i]\times[y_{j-1},y_j]$. As we increase the number of subdivisions, the error becomes zero. In

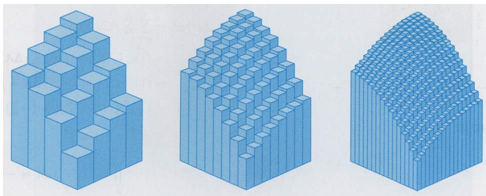


Figure 2.6: Approximation of volume by rectangles [3, p977]

single variable case, a function $y=f(x)$ is continuous at $x=x_0$ if $f(x)\to f(x_0)$ as $x\to x_0$. We can define continuity of two variable function $f(x,y)$ likewise[8]: a function $f(x,y)$ is *continuous* at the point (x_0,y_0) if $f(x,y)\to f(x_0,y_0)$ as $(x,y)\to(x_0,y_0)$. Note that there are two paths for $x\to x_0$, and each path gives the notion of *left* and *right* limit of $f(x)$. The limit $\lim_{x\to x_0} f(x)$ exists only if two limits coincides. (This is the definition of the limit.) However, there are infinitely many paths for $(x,y)\to(x_0,y_0)$. The limit $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ exists only if $f(x,y)$ converges to the same real number for every path.

Example 2.8. Let $f(x,y)=\dfrac{y}{1+xy}$. Figure 2.7 show that the function $f(x,y)$ is continuous on $D=[0,1]\times[0,1]$. (Although $f(x,y)$ is not continuous at the entire plane $\mathbf{R}^2$.)

Theorem 2.9. *If f is continuous on the region $D=[a,b]\times[c,d]$, then the double integral $\iint_D fdA$ exists.*

[8] A rigorous definition by ϵ-δ statement will be given later.

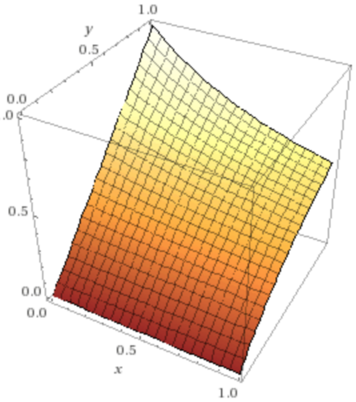


Figure 2.7: The graph of $\dfrac{y}{1+xy}$

Example 2.10. The function $f(x,y)$ on $[0,1]\times[0,1]$ defined by

$$f(x,y)=\begin{cases}1 & x \text{ or } y \text{ is rational}\\ 0 & \text{otherwise}\end{cases}$$

is not integrable on $[0,1]\times[0,1]$.

The following theorem connects the iterated integrals and double integrals.

Theorem 2.11 (Fubini I)**.** *Let $f(x,y)$ be a continuous function defined on $D=[a,b]\times[c,d]$. Then*

$$\iint_D f\,dA=\int_a^b\int_c^d f(x,y)dydx=\int_c^d\int_a^b f(x,y)dxdy$$

Example 2.12. Let us compute the following double integral.

$$\iint_{[0,1]\times[0,1]}\frac{y}{1+xy}dxdy$$

By Fubini's theorem,

$$\begin{aligned}\iint_{[0,1]\times[0,1]}\frac{y}{1+xy}dxdy&=\underbrace{\int_0^1\left[\int_0^1\frac{y}{1+xy}dx\right]dy}_{①}\\&=\underbrace{\int_0^1\left[\int_0^1\frac{y}{1+xy}dy\right]dx}_{②}\end{aligned}$$

In the second iterated integral, we should compute $\displaystyle\int_0^1\frac{y}{1+xy}dy$ first, which is hard. Thus we take the first iterated integral.

$$\begin{aligned}\int_0^1\left[\int_0^1\frac{y}{1+xy}dx\right]dy&=\int_0^1\ln(1+xy)\Big|_{x=0}^1dy\\&=\int_0^1\ln(1+y)dy\\&=y\ln(1+y)\Big|_0^1-\int_0^1\frac{y}{1+y}dy\\&=\ln 2-\int_0^1 1-\frac{1}{1+y}dy\\&=\ln 2-\left(1-\ln(1+y)\Big|_0^1\right)\\&=2\ln 2-1\end{aligned}$$

Remark 2.13. The *continuity* in Theorem 2.11 is crucial. The following example is a function which Fubini's theorem fails.

$$f(x,y)=\begin{cases}\dfrac{x^2-y^2}{(x^2+y^2)^2} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}\tag{2.13.1}$$

Figure 2.8 shows that the graph of $z=f(x,y)$ is discontinuous at $(0,0)$. As $(x,y)\to(0,0)$, $f(x,y)$ may or may not converge to 0. For example, if (x,y) follows the parametric line $c(t)=(t,t)$, then $f(x,y)$ converges to 0. However, if (x,y) follows the x or y-axis, then $f(x,y)$ diverges. Let us compute

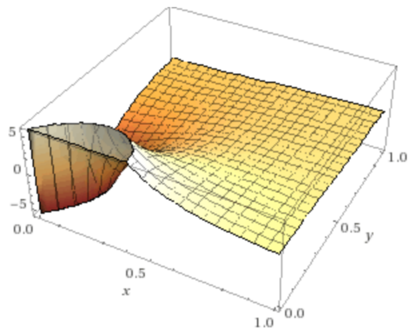


Figure 2.8: The graph of the function (2.13.1)

$\displaystyle\int_0^1\int_0^1 f(x,y)dydx$ first.

$$\begin{aligned}\int_0^1\int_0^1\frac{x^2-y^2}{(x^2+y^2)^2}dydx&=\int_0^1\frac{y}{x^2+y^2}\Big|_{y=0}^1dx\\&=\int_0^1\frac{1}{1+x^2}dx=\tan^{-1}x\Big|_0^1=\frac{\pi}{4}\end{aligned}$$

The iterated integral $\int_0^1 \int_0^1 f(x,y)dxdy$ is

$$\begin{aligned}\int_0^1 \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}dxdy &= \int_0^1 \frac{-x}{x^2+y^2}\bigg|_{x=0}^1 dx \\ &= \int_0^1 \frac{-1}{1+y^2}dy \\ &= -\tan^{-1} y\Big|_0^1 = -\frac{\pi}{4}\end{aligned}$$

and it does not coincide with $\int_0^1 \int_0^1 f(x,y)dydx$.

We can define a double integral over non-rectangular domain as follows.

Definition 2.14. Let $f(x,y)$ be defined on a <u>bounded</u> region D in $\mathbf{R}^2$. Suppose that D lies on a large rectangular domain, say $D \subset [a,b]\times[c,d]$. Let us define a new function $F(x,y)$ as follows (see Figure 2.9).

$$F(x,y) = \begin{cases} f(x,y) & (x,y)\in D \\ 0 & (x,y)\notin D \end{cases} \tag{2.14.1}$$

The **definite integral** of f over the domain D is

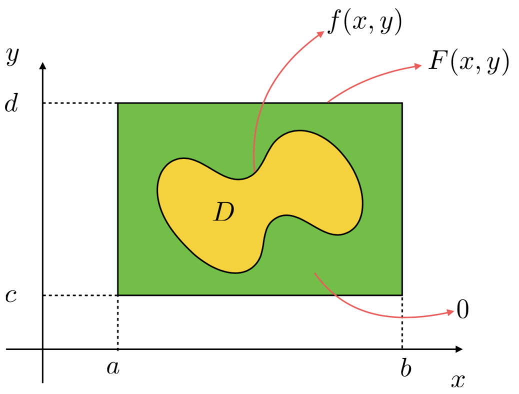


Figure 2.9: The domain of $F(x,y)$

defined as follows.

$$\iint_D f(x,y)dxdy = \iint_{[a,b]\times[c,d]} F(x,y)dxdy$$

Another version of Fubini's theorem allows us to compute the double integral over a general bounded domain.

Theorem 2.15 (Fubini II). *Let $f(x,y)$ be a continuous function defined on D. If $D = \{(x,y)\,|\,a\le x\le b,\ g_1(x)\le y\le g_2(x)\}$, then*

$$\iint_D f(x,y)dxdy = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)dydx \tag{2.15.1}$$

Similarly, if $D = \{(x,y)\,|\,c\le y\le d,\ h_1(y)\le x\le h_2(y)\}$, then

$$\iint_D f(x,y)dxdy = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)dxdy \tag{2.15.2}$$

Equation (2.15.1) and (2.15.2) can be visualized in Figure 2.10. Suppose that $D\subset[a,b]\times[c,d]$. From the Fubini I (Theorem 2.11) and Definition 2.14,

$$\begin{aligned}\iint_D f(x,y)dxdy &= \int_a^b \left[\int_c^d F(x,y)dy\right]dx \\ &= \int_c^d \left[\int_a^b F(x,y)dx\right]dy\end{aligned}$$

Let us write the first iterated integral in the Riemann sum formulation.

$$\int_a^b \left[\int_c^d F(x,y)dy\right]dx \approx \sum_{i=1}^n \left[\int_c^d F(x_i^*,y)dy\right]\Delta x_i$$

Each summand $\int_c^d F(x_i^*,y)dy\Delta x_i$ is the volume under the graph of $z = F(x,y)$ over the rectangular region $[x_{i-1},x_i]\times[c,d]$. From the formulation of F (Equation (2.14.1)), we get

$$\int_c^d F(x_i^*,y)dy = \int_{g_1(x_i^*)}^{g_2(x_i^*)} f(x_i^*,y)dy$$

Thus we get Equation (2.15.1). Equation (2.15.2) can be derived in a similar way.

Example 2.16. Let us compute $\iint_D e^{-y^2}dxdy$ where D is the triangular region whose vertices are $(0,0)$, $(0,1)$, and $(1,1)$. First, let us identify the domain for Equation (2.15.1).

$$D = \{(x,y)\,|\,0\le x\le 1,\ x\le y\le 1\}$$

Then we need to compute

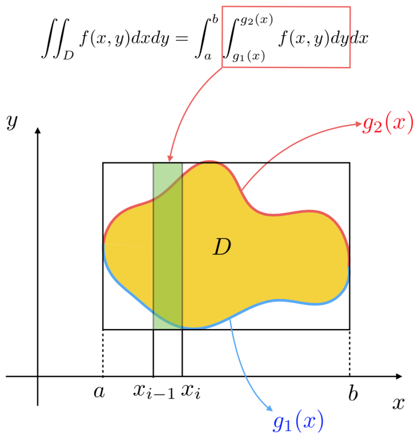
$$\int_0^1 \left[\int_x^1 e^{-y^2}dy\right]dx,$$

but the first integration $\int_x^1 e^{-y^2}dy$ not easy to compute. Thus let us use Equation (2.15.2). The domain D can be identified as follows.

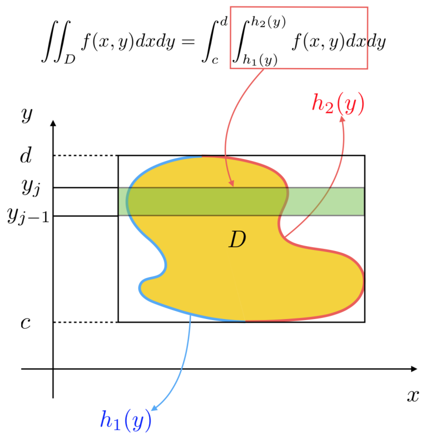
$$D = \{(x,y)\,|\,0\le y\le 1,\ 0\le x\le y\}$$

Equation 2.15.2 gives the following result.

$$\begin{aligned}\iint_D e^{-y^2}dxdy &= \int_0^1 \left[\int_0^y e^{-y^2}dx\right]dy \\ &= \int_0^1 \left[xe^{-y^2}\Big|_{x=0}^y\right]dy = \int_0^1 ye^{-y^2}dy \\ &= -\frac{1}{2}e^{-y^2}\Big|_0^1 = \frac{1}{2}\left(1-e^{-1}\right)\end{aligned}$$



(a) Equation (2.15.1)



(b) Equation (2.15.2)

Figure 2.10: Visualization of Fubini II

Definition 2.17. Let $f(x,y,z)$ be a function defined on the boxed domain $D = [a,b]\times[c,d]\times[e,f]$. The **triple integral** of $f(x,y,z)$ over D is denoted by

$$\iiint_{[a,b]\times[c,d]\times[e,f]} f(x,y,z)dxdydz$$

While the double integral $\iint_D f(x,y)dxdy$ represents the volume under the graph of $z = f(x,y)$, the **triple integral** $\iiint_V fdxdydz$ is the *mass* of volume V with density $f(x,y,z)$. One can also compute the triple integral over the bounded domain which is not necessarily a rectangular. Let V be a bounded domain in $\mathbf{R}^3$ and assume that $D \subset [a,b]\times[c,d]\times[e,f]$. Define

$$F(x,y,z) = \begin{cases} f(x,y,z) & \text{if } (x,y,z)\in V \\ 0 & \text{otherwise} \end{cases}$$

Then the triple integral over V is defined as

$$\iiint_V f(x,y,z)dxdydz = \iiint_{[a,b]\times[c,d]\times[e,f]} F(x,y,z)dxdydz$$

Theorem 2.18 (Fubini III)**.** *Let $f(x,y,z)$ be a continuous function defined on the region V.*

$$\begin{aligned} V =&\{(x,y,z)\,|\,a\le x\le b,\ g_1(x)\le y\le g_2(x),\\ &h_1(x,y)\le z\le h_2(x,y)\}\end{aligned}$$

Then the following holds.

$$\iiint_V f(x,y,z)dxdydz = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)dzdydx \tag{2.18.1}$$

Let us show why Theorem 2.18 holds. An analogue of Theorem 2.11 can be stated for the triple integrals as follows.

$$\begin{aligned}&\iiint_{[a,b]\times[c,d]\times[e,f]} F(x,y,z)dxdydz\\ &= \int_a^b\int_c^d\left[\int_e^f F(x,y,z)dz\right]dydx\\ &\approx \sum_{i=1}^n\sum_{j=1}^m\left[\int_e^f F(x_i^*,y_j^*,z)dz\right]\Delta x_i\Delta y_j\end{aligned}$$

Note that each summand represents the *mass* of cube $[x_{i-1},x_i]\times[y_{j-1},y_j]\times[z_{k-1},z_k]$ (Figure 2.11). Note that $F(x_i^*,y_j^*,z) = f(x_i^*,y_j^*,z)$ for $h_1(x_i^*,y_j^*)\le z\le h_2(x_i^*,y_j^*)$ and zero otherwise. Thus

$$\int_e^f F(x_i^*,y_j^*,z)dz = \int_{h_1(x_i^*,y_j^*,z)}^{h_2(x_i^*,y_j^*,z)} f(x_i^*,y_j^*,z)dz$$

and this gives the following equality.

$$\begin{aligned}&\iiint_{[a,b]\times[c,d]\times[e,f]} F(x,y,z)dxdydz\\ &= \int_a^b\int_c^d\left[\int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)dz\right]dydx\end{aligned}$$

Similarly, we can replace the second integration $\int_c^d$ by $\int_{g_1(x)}^{g_2(x)}$, and this shows Equation (2.18.1).

Example 2.19. Let V be a parallelopiped region bounded by 6 planes : $2x = y$, $2x = y+2$, $y = 0$, $y = 4$, $z = 0$, $z = 3$. Let us compute

$$\iiint_V \frac{2x-y}{2} + \frac{z}{3}dxdydz$$

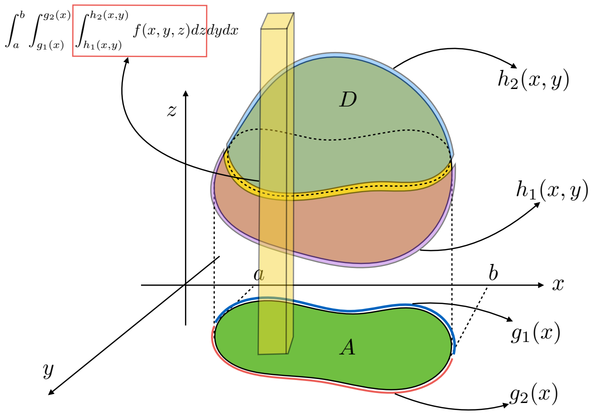


Figure 2.11: The triple integral over D

2.3 Surface integrals on planar surfaces

We can consider a double integral $\iint_D f dxdy$ as the *mass* of the planar area D on xy-plane with density $f(x,y)$. What about a region on a general plane in $\mathbf{R}^3$? How can be find the mass of a planar region in $\mathbf{R}^3$ in general? Let us try to generalize the notion of double integral in Definition 2.7 as follows.

Definition 2.20. Let S be a planar parallelogram in $\mathbf{R}^3$ and $f(x,y,z)$ be a function defined on S. Let us subdivide S into nm sub-parallelograms S_{ij} as follows. Pick a vector $P \in S$ and two vectors $\mathbf{v}, \mathbf{w} \in \mathbf{R}^3$ which is parallel to the boundary of S. Choose a, b such that every point on $p \in S$ is parametrized by

$$p(s,t) = \overrightarrow{OP} + s\mathbf{v} + t\mathbf{w}$$

for $(s,t) \in [0,a] \times [0,b]$. ($a|\mathbf{v}, b|\mathbf{w}|$ are the lengths of two transverse sides of S.) Let s_i, t_j be the sequence defined by

$$0 = s_0 \le s_1 \le \cdots \le s_n = a$$

$$0 = t_0 \le t_1 \le \cdots \le t_m = b$$

Then the sub-parallelogram S_{ij} is the set

$$S_{ij} = \{p(s,t) \,|\, x_{i-1} \le s \le x_i \, y_{j-1} \le t \le y_j\} \tag{2.20.1}$$

The **surface integral** of $f(x,y,z)$ over S, denoted by $\iint_S f(x,y,z)dA$, is the limit of the Riemann sum

$$\iint_S f(x,y,z)dA = \lim_{\Delta S \to 0} \sum_{i=1}^{n} \sum_{j=1}^{m} f(p(s_i^*, t_j^*))\Delta S_{ij} \tag{2.20.2}$$

where $s_i^* \in [x_{i-1}, x_i]$ and $t_j^* \in [y_{j-1}, y_j]$ are arbitrary points and ΔS_{ij} is the area of sub-parallelogram S_{ij}.

One might ask why we use dA instead or $dxdy$ (or $dxdydz$!). First, note that the we use dA instead of $dxdy$, since the surface S may not lie on the xy-plane. Secondly, dA symbolically means the infinitesimal area of the sub-parallelogram[9] ΔS. We do need x, y, z coordinate for the integrand f, because f is defined on the points in $\mathbf{R}^3$. However, the values of x, y, z are not independent because the point (x,y,z) must lie in the surface S. (At least one coordinate must be dependent to the other two.)

Computing surface integral over a planar region requires to use of double integral. We will explain this by the following example.

Example 2.21. Let us compute $\iint_S xzdA$ for the parallelogram $S = \{(x,y,z,) \,|\, x+y+z = 2, 0 \le x \le 1, 0 \le y \le 1\}$ (Figure 2.12). The surface S contains the point $P = (0,0,2)$ and $\mathbf{v} = (1,0,-1)$ and $\mathbf{w} = (0,1,-1)/\sqrt{2}$ are parallel vectors to S. (Note that S is normal to $\mathbf{n} = (1,1,1)$.) For $(s,t) \in [0,\sqrt{2}] \times [0,\sqrt{2}]$, the point $p(s,t) = (0,0,2) + s(1,0,-1) + t(0,1,-1)$ always lies on S. Let us subdivide the interval $[0,2]$ into n, m subintervals, and thus S into nm sub-parallelograms as in Equation (2.20.1). Each region ΔS_{ij} is bounded by two vectors $\mathbf{v}\Delta s_i$ and $\mathbf{w}\Delta t_j$. The area ΔS_{ij} is then

$$\Delta S_{ij} = |\mathbf{v} \times \mathbf{w}|\Delta s_i \Delta t_j = \sqrt{3}\Delta s_i \Delta t_j$$

Thus the Riemann sum in Equation (2.20.2) is written as

$$\lim_{n,m\to\infty} \sum_{i=1}^{n}\sum_{j=1}^{m} \sqrt{3}s_i^*(2 - s_i^* - t_j^*)\Delta s_i \Delta t_j$$

From Equation (2.7.2) from the definition of double integral, above is the double integral of $f(x,y) = \sqrt{3}x(2-x-y)$ over the rectangular region $[0,\sqrt{2}] \times [0,\sqrt{2}]$. By Fubini's theorem (Theorem 2.11),

$$\begin{aligned}
\iint_S xzdA &= \int_0^{\sqrt{2}}\int_0^{\sqrt{2}} \sqrt{3}x(2-x-y)dxdy \\
&= \int_0^{\sqrt{2}} 3x^2 - x^3 - \frac{3}{2}x^2y\Big|_{x=0}^{\sqrt{2}} dy \\
&= \int_0^{\sqrt{2}} 6 - 3\sqrt{2} - 3ydy = 6\sqrt{2} - 9
\end{aligned}$$

Now we want to integrate a *vector field* instead of a function. Such integral is called a *flux*. We will take a moment to ponder about the concept of a flux.

Imagine a fluid flow along a tube with constant velocity $\mathbf{v}$. We want to measure the amount of water flowing per second, namely a *flux*. Such amount is certainly a constant value and one way to measure this is to pouring the water through the end of hose. Another way to measure the flux is taking a section of the tube and measure the amount of water flow across this section over 1 second. If the section S is orthogonal to the direction of fluid, that is the normal vector $\mathbf{n}$ to the section is parallel to $\mathbf{v}$, then

[9] We do often replace $dxdy$ by dA for double integrals. However, we should not confuse the notation $dxdy$ in double and iterated integral. The symbol $dxdy$ in double integral represents the infinitesimal area, as dA does. Meanwhile, $dxdy$ or $dydx$ in iterated integral means the order of integration.

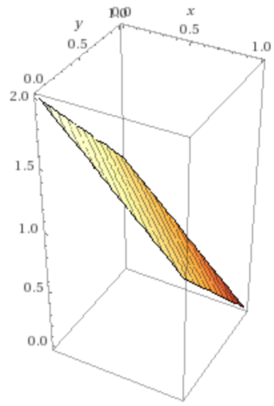


Figure 2.12: The plane $x + y + z = 2$

the flux is $|V|\text{area}(S)$. Obviously, the flux must not change by configuration of the section as long as the section *seals* the tube. This leads us to the following definition of a flux.

Definition 2.22. Let S be a bounded planar surface in $\mathbf{R}^3$ with unit normal direction $\mathbf{n}$. Let $\mathbf{F}$ be a vector field defined on S. The **flux** of F across S to the direction of $\mathbf{n}$ is denoted by $\iint_S \mathbf{F} \cdot d\mathbf{A}$, and defined as follows.

$$\iint_S \mathbf{F} \cdot d\mathbf{A} = \iint_S \mathbf{F} \cdot \mathbf{n} dA \tag{2.22.1}$$

The symbol $d\mathbf{S}$ respresents the infinitesimal area dS of a surface and its normal direction $\mathbf{n}$. Note that $\mathbf{F} \cdot \mathbf{n}$ is a scalar value defined on the surface S. Thus Equation (2.22.1) is simply a double integration of the function $\mathbf{F} \cdot \mathbf{n}$ on S. Also note that the sign of the flux changes by direction of $\mathbf{n}$. One should set the normal vector $\mathbf{n}$ points the *outward* direction of $\mathbf{F}$ across S.

Let us observe an easy (but important) example of a flux. Suppose that $\mathbf{F} = (0, 0, 1)$ is a constant vector field defined on $\mathbf{R}^3$. First, let us compute the flux of $\mathbf{F}$ across the horizontal plane S_1: $z = 1$, $0 \le x, y \le 1$. Here, we set the normal vector of S_1 to be of *upward* direction, namely $\mathbf{n} = \mathbf{k} = (0, 0, 1)$.

$$\begin{aligned}\iint_{S_1} \mathbf{F} d\mathbf{A} &= \iint_{S_1} (0,0,1) \cdot (0,0,1) dA \\ &= \text{area } S_1 = 1\end{aligned}$$

Next, let S_2 be a plane with a slope, say $z = x$, $0 \le x, y \le 1$. Let us compute the flux of $\mathbf{F}$ across S_2 to the upward direction. This means that we should set the normal vector of S_2 to be $\mathbf{n} = (-1, 0, 1)/\sqrt{2}$. Before we start computation, we know that the result should be the same as above.

$$\begin{aligned}\iint_{S_2} \mathbf{F} d\mathbf{A} &= \iint_{S_2} (0,0,1) \cdot (-1,0,1)/\sqrt{2} dA \\ &= \frac{1}{\sqrt{2}} \text{area } S_2 = \frac{1}{\sqrt{2}}\sqrt{2} = 1\end{aligned}$$

Example 2.23. Let S be the plane $2x+4y+z = 8$ bounded by $x = 0$, $y = 0$, $z = 0$. Let us find the flux of the vector field $\mathbf{F} = (x, -y, 0)$ across the surface S with the *upward* direction $\mathbf{n} = (2, 4, 1)/\sqrt{21}$. Since the surface in *not* a parallelogram, we need different approach to compute the integral.

Let $\tilde{S}$ be the plane $2x + 4y + z = 8$ over the rectangular domain $[0, 4] \times [0, 2]$ and

$$\tilde{\mathbf{F}}(x, y, z) = \begin{cases} \mathbf{F}(x,y,z) & (x,y,z) \in S \\ (0,0,0) & \text{otherwise} \end{cases} \tag{2.23.1}$$

Then

$$\iint_S \mathbf{F} \cdot d\mathbf{A} = \iint_{\tilde{S}} \tilde{\mathbf{F}} \cdot d\mathbf{A}$$

and we will compute the right hand-side. Let us subdivide the domain $[0, 4] \times [0, 2]$ into nm rectangular subdomains as in Definition 2.7.

$$0 = x_0 \le x_1 \le \cdots \le x_n = 4$$

$$0 = y_0 \le y_1 \le \cdots \le y_m = 2$$

Let S_{ij} be the sub-parallelogram of S over the domain $D_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$. Note that $\mathbf{v} = (1, 0, -2)$, $\mathbf{w} = (0, 1, -4)$ are vectors parallel to the side of S. Therefore,

$$\Delta S_{ij} = \left| \frac{\mathbf{v}}{|\mathbf{v}|} \times \frac{\mathbf{w}}{|\mathbf{w}|} \right| \Delta x_i \Delta y_j = \frac{\sqrt{21}}{\sqrt{5}\sqrt{17}} \Delta x_i \Delta y_j$$

Thus Equation (2.20.2) can be written as follows.

$$\begin{aligned}\iint_{\tilde{S}} \tilde{\mathbf{F}} \cdot d\mathbf{A} &= \lim \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{\mathbf{F}} \cdot \frac{(2,4,1)}{\sqrt{21}} \Delta S_{ij} \\ &= \frac{1}{\sqrt{5}\sqrt{17}} \lim \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{\mathbf{F}} \cdot (2,4,1) \Delta x_i \Delta y_j\end{aligned}$$

But then, by using Equation (2.23.1) and idea of Theorem 2.15,

$$\begin{aligned}\iint_S \mathbf{F} \cdot d\mathbf{A} &= \frac{1}{\sqrt{5}\sqrt{17}} \int_0^2 \int_0^{2-x/2} 2x - 4y dx dy \\ &= \frac{1}{\sqrt{5}\sqrt{17}} \left(-8x + 4x^2 - \frac{x^3}{6} \Big|_0^2 \right) \\ &= -\frac{4}{3\sqrt{5}\sqrt{17}}\end{aligned}$$

The negative value means that the total flux across the surface is *opposite* direction of $\mathbf{n}$, that is, the fluid (or whichever the vector field represents) flows *inward*.

2.4 A concise version of Gauss' theorem

Consider a vector field defined on $\mathbf{R}s$. How can we compute the flux across a sphere? How does it change with respect to the radius of the sphere? Let us start with a simple example.

Example 2.24. Let

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

be a vector field defined on $\mathbf{R}^3 \backslash \{(0,0,0)\}$. Let S_r be the sphere of radius r centered at $(0,0,0)$. Suppose that $r = 1$. At each point $p \in S$, the vector $\mathbf{F}(p)$ is normal to the surface S, and its length is $|\mathbf{F}(p)| = 1$. Thus

$$\iint_{S_1} \mathbf{F} \cdot d\mathbf{A} = \iint_{S_1} dA = 4\pi$$

In general, the vector $\mathbf{F}(p)$ is normal to S, but its length is $1/r$. Thus

$$\iint_{S_r} \mathbf{F} \cdot d\mathbf{A} = \iint_{S_r} \frac{1}{r} dA = \frac{4\pi r^2}{r} = 4\pi r.$$

Now let us repeat above with

$$\mathbf{F} = \left(\frac{x}{(x^2+y^2+z^2)^{3/2}}, \frac{y}{(x^2+y^2+z^2)^{3/2}}, \frac{z}{(x^2+y^2+z^2)^{3/2}} \right).$$

The vector $\mathbf{F}(p)$ is normal to S, and has length $1/r^2$. Thus

$$\iint_{S_r} \mathbf{F} \cdot d\mathbf{A} = \iint_{S_r} \frac{1}{r^2} dA = \frac{4\pi r^2}{r^2} = 4\pi.$$

Note that the flux of the vector

$$\mathbf{F} = \left(\frac{x}{(x^2+y^2+z^2)^{3/2}}, \frac{y}{(x^2+y^2+z^2)^{3/2}}, \frac{z}{(x^2+y^2+z^2)^{3/2}} \right)$$

is constant for each radius r of the spherical surface. It means that for $r < R$, the flux across the sphere S_r and the flux across the sphere S_R is the same. In other words, the number of particles moving along the vector field $\mathbf{F}$ passing through the surfaces S_r and S_R per given time are the same, thus no particles remain between S_r and S_R.

Theorem 2.25 (Gauss' law)**.** *Let $\mathbf{E}(q)$ be the electric field caused by a electric charge q at $(0,0,0)$. Let $\Phi(q)$ be the electric flux of $\mathbf{E}(q)$ across the closed surface S, that is, $\Phi(q) = \iint_S \mathbf{E}(q) \cdot d\mathbf{A}$. Then*

$$\Phi(q) = \frac{q}{\epsilon_0}$$

where ϵ_0 is the electric constant. *($\epsilon_0 = 8.854 \times 10^{-12}\, F \cdot m^{-1}$)*

We assume that the electric field must depend on $r = |\mathbf{r}| = \sqrt{x^2+y^2+z^2}$. That is the electric force must be equal at points with the same radius from the source charge. Example 2.24 shows that the electric field $\mathbf{E}(q)$ must be of the form

$$\mathbf{E}(q) = \frac{q}{4\pi\epsilon_0} \frac{(x,y,z)}{(x^2+y^2+z^2)^{3/2}}$$

Note that Gauss' theorem works for any type of closed surfaces. (We will prove it in §4.3.) If we use Gauss' theorem, we can find some value of multiple integrals which is hard to compute by hands.

Example 2.26. Let $D = [-1,1] \times [-1,1]$. Let us compute $\iint_D \frac{1}{(x^2+y^2+1)^{3/2}} dA$. Note that using Fubini's theorem (cf. Theorem 2.11) is hopeless:

$$\iint_D \frac{1}{(x^2+y^2+1)^{3/2}} dA = \int_{-1}^{1} \int_{-1}^{1} \frac{1}{(x^2+y^2+1)^{3/2}} dxdy$$

Let S be the boundary of the cube $[-1,1] \times [-1,1] \times [-1,1]$ and

$$\mathbf{F} = \left(\frac{x}{(x^2+y^2+z^2)^{3/2}}, \frac{y}{(x^2+y^2+z^2)^{3/2}}, \frac{z}{(x^2+y^2+z^2)^{3/2}} \right).$$

By the Gauss theorem, $\iint_S \mathbf{F} \cdot d\mathbf{A} = 4\pi$. Using the symmetry of the cube, we know that the flux across the upper side of S is $\pi/3$. On the upper side of S, that is, when $z = 1$, the vector field $\mathbf{F}$ becomes

$$\mathbf{F} = \left(\frac{x}{(x^2+y^2+1)^{3/2}}, \frac{y}{(x^2+y^2+1)^{3/2}}, \frac{1}{(x^2+y^2+1)^{3/2}} \right).$$

Since the normal direction of the surface is $\mathbf{n} = (0,0,1)$, we have

$$\frac{\pi}{3} = \int_{-1}^{1} \int_{-1}^{1} \frac{1}{x^2+y^2+1} dxdy$$

2.5 Exercises

1. Compute the following.

   (a) $\displaystyle\iint_{[-1,1]\times[0,\pi]} xy\cos y dA.$

   (b) $\displaystyle\iint_{[0,1]\times[0,2]} \frac{xy^3}{x^2+1} dA.$

   (c) $\displaystyle\iint_{[1,2]\times[1,2]} \frac{1}{xy} dA.$

2. Compute the following.

   (a) $\displaystyle\int_0^1\int_0^{s^2} \cos(s^3)dtds.$

   (b) $\displaystyle\int_0^1\int_0^{y^3} 3y^4e^{xy}dxdy$

   (c) $\displaystyle\iint_D y^2dA$ where D is the triangular whose vertices are $(0,1)$, $(2,2)$, and $(4,1)$.

   (d) $\displaystyle\iint_D (2x-y)dA$ where D is the disk with radius 2 and center $(0,0)$.

3. Sketch the regions and compute the integrals.

   (a) $\displaystyle\int_0^1\int_0^{3x^2} 3dydx$

   (b) $\displaystyle\int_0^2\int_0^{y^2} ydxdy$

   (c) $\displaystyle\int_0^2\int_0^{x^2} ydydx$

   (d) $\displaystyle\int_{-1}^3\int_x^{2x+1} xydydx$

   (e) $\displaystyle\int_0^2\int_{x^2/4}^{x/2} x^2+y^2dydx$

   (f) $\displaystyle\int_0^4\int_0^{2\sqrt{y}} x\sin y^2dxdy$

   (g) $\displaystyle\int_0^{\pi}\int_0^{\sin x} y\cos xdydx$

   (h) $\displaystyle\int_0^1\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 3dydx$

   (i) $\displaystyle\int_{-1}^1\int_0^{\sqrt{1-y^2}} 3dxdy$

   (j) $\displaystyle\int_0^1\int_{-e^x}^{e^x} y^3dydx$

4. Integrate the function $f(x,y)=1-xy$ over the trianglar region whose vertices are $(0,0)$, $(2,0)$, $(0,2)$.

5. Integrate the function $f(x,y)=3xy$ over the region bounded by $y=32x^3$ and $y=\sqrt{x}$.

6. Integrate the function $f(x,y)=x+y$ over the region bounded by $x+y=2$ and $y^2-2y-x=0$.

7. Compute

$$\int_1^2\int_{\sqrt{x}}^{x}\sin\left(\frac{\pi x}{2y}\right)dydx+\int_2^4\int_{\sqrt{x}}^{2}\sin\left(\frac{\pi x}{2y}\right)dydx$$

Triple integrals

7. Find the volume under the surface $z=2\sin x\cos y$ over the region $[0,\frac{\pi}{2}]\times[0,\frac{\pi}{4}]$.

8. If $f(x,y,z)=1$, the triple integral $\displaystyle\iiint_D dV$ is the volume of the region D. Compute the following.

   (a) The volume under the graph $z=10+x^3+3y^2$ on $[0,1]\times[0,2]$.

   (b) The volume of paraboloid $\dfrac{x^2}{4}+\dfrac{y^2}{9}+z=1$ on $[-1,1]\times[-2,2]$.

   (c) The volume bounded by $z=x\sec^2 y$, $z=0$, $x=0$, $x=2$, $y=0$, and $y=\pi/4$.

   (d) The volume under the plane $z=2-x-y$ on $[0,1]\times[0,1]$.

   (e) The volume under the graph $z=2\sin x\cos y$ on $[0,\pi/2]\times[0,\pi/4]$.

9. Evaluate the triple integrals.

   (a) $\displaystyle\iiint_{[-1,1]\times[0,2]\times[1,3]} xyzdV$

   (b) $\displaystyle\int_{-1}^2\int_1^{z^2}\int_0^{y+z} 3yz^2dxdydz$

   (c) $\displaystyle\int_1^3\int_0^z\int_1^{xz} (x+2y+z)dydxdz$

   (d) $\displaystyle\int_0^1\int_{1+y}^{2y}\int_z^{y+z} zdxdzdy$

10. Integrate the given function over the indicated region W.

    (a) $f(x,y,z)=2x-y+z$; W is the region bounded by the cylinder $z=y^2$, the xy-plane, and the planes $x=0$, $x=1$,$y=-2$, $y=2$.

    (b) $f(x,y,z)=y$; W is the region bounded by the plane $x+y+z=2$, the cylinder $x^2+z^2=1$, and $y=0$.

(c) $f(x, y, z) = 8xyz$; W is the region bounded by the cylinder $y = x^2$, the plane $y + z = 9$, and the xy-plane.

11. Change the order of integration of

$$\int_0^1 \int_0^1 \int_0^{x^2} f(x, y, z)dzdxdy$$

to give five other equivalent iterated integrals.

12. Change the order of integration of

$$\int_0^2 \int_0^x \int_0^y f(x, y, z)dzdydx$$

to give five other equivalent iterated integrals.

3 Line integrals

A line integral is another generalization of integration of single variable function. Instead of taking domains in x-axis, one may take a curve in $\mathbf{R}^2$, $\mathbf{R}^3$ (or $\mathbf{R}^n$ in general) as a domain of integration for multivariable function. We can also take a line integral of vector fields, which represents the *work* done by the vector field.

Along with integration, we can also generalize[10] the differentiation of single variable function to multivariable function, called the directional, or partial derivatives.

Once we introduce the notion of partial derivatives, several important operations such as *curl* and *divergence* on vector fields will be introduced. This will allow us to introduce the notion of conservative vector fields, which is important concept of describing the conservation of energy.

Important questions are :

- What is the geometric meaning of line integrals?
- What does the partial derivatives represent in the graph of multivariable function?
- How can we use curl and divergence to characterize the vector fields?
- What are the conditions for a vector field to be conservative?

[10]Although this generalization is different from the actual definition of differentiability of multivariable function which will be covered in §5.

3.1 Line integrals

The integration of single variable function may exists if the function is continuous except for finitely many points. In general, the integration of multivariable function on a curve my defined if the function is not only discontinuous only at finitely many points, but also the curve is *piecewise* smooth on $\mathbf{R}^n$.

Definition 3.1. Given an interval $I \subset \mathbf{R}$, a curve C parametrized by $c : I \to \mathbf{R}^n$

$$c(t) = (x_1(t), x_2(t), \cdots, x_n(t))$$

is **piecewise** $\mathcal{C}^n$ if all coordinate function $x_i(t)$ are $\mathcal{C}^n$[11] on the interval I except for finitely many points.

Let us define integration on curves on xy-plane. (The line integral on curves in $\mathbf{R}^n$ can be defined in a similar way.)

Definition 3.2. Let C be a piecewise $\mathcal{C}^1$ curve on $\mathbf{R}^2$ parametrized by $c : [a, b] \to \mathbf{R}^2$. Let $c'(t) = (x'(t), y(t))$ be the velocity vector at the point $c(t) = (x(t), y(t))$. For the function $f(x, y)$ defined on C, the **line integral** of $f(x, y)$ along C is

$$\int_C f ds = \int_a^b f(c(t))|c'(t)|dt$$

Note that the line integral depends on the shape of the curve, but *not* on the parametrization of the curve. Let C be a curve parametrized by $c : [a, b] \to \mathbf{R}^2$. Suppose that a one-to-one correspondence $h : [c, d] \to [a, b]$ gives a *re-parametrization* $c \circ h$ of C. Let us write $t = h(\tau)$. Then

$$\begin{aligned}
&\int_c^d f(c \circ h(\tau))|(c \circ h)'(\tau)d\tau \\
&= \int_c^d f(c(t))c'(t)h'(\tau)d\tau \\
&= \int_{t=a}^b f(c(t))dt \\
&= \int_C f(x, y)ds
\end{aligned}$$

Example 3.3. Let us compute the line integral

$$\int_C (2 + x^2 y)ds \tag{3.3.1}$$

where C is the unit circle centered at the origin with counter clockwise orientation. First we parametrize the curve C as follows.

$$c(t) = (\cos(t), \sin(t)), \quad t \in [0, 2\pi]$$

Then $|c'(t)| = 1$ and

$$\int_C (2 + x^2 y)ds = \int_0^{2\pi} (2 + \cos^2(t)\sin(t))dt = 4\pi.$$

One might already have expected the value of the integral in (3.3.1) as twice the length of the curve. *The line integral* $\int_C f ds$ *of a function* f *represent the area under the graph of* $f(x, y)$ *along the curve* C. By the symmetry of the circle, the integral of $x^2 y$ on upper semi-circle is negative of the integral of $x^2 y$ on lower semi-circle.

Similar to the surface integral, we can define a line integral of a vector field.

Definition 3.4. Let $\mathbf{F}(x, y) = (P(x, y), Q(x, y))$ be a vector field defined on a domain $D \subset \mathbf{R}^2$. Let C be a piecewise $\mathcal{C}^1$ curve in D parametrized by $c : [a, b] \to \mathbf{R}^2$. The **line integral** of $\mathbf{F}$ along C (with an orientation) is

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{s} &= \int_a^b \mathbf{F}(x(t), y(t)) \cdot c'(t)dt \\
&= \int_a^b P(x(t), y(t))x'(t) + Q(x(t), y(t))y'(t)dt
\end{aligned}$$

If we denote the vector field $\mathbf{F}$ as

$$\mathbf{F}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j},$$

then the line integral $\int_C \mathbf{F} \cdot d\mathbf{s}$ is written by

$$\int_C \mathbf{F} \cdot d\mathbf{s} = \int_C P dx + Q dy$$

Note also that the line integral of a vector field does not depend on the re-parametrization. Let $f : [c, d] \to [a, b]$ be a one-to-one $\mathcal{C}^1$-function and the composition $\tilde{c} = c \circ f$,

$$\int_C \mathbf{F} \cdot d\mathbf{s} = \int_a^b \mathbf{F}(c(t)) \cdot c'(t)dt = \int_c^d \mathbf{F}(\tilde{c}(t)) \cdot \tilde{c}'(t)dt$$

However, the line integral of a vector field *does* depends on the *orientation* of the curve. Suppose that we reverse the parametrization and let $d(t) = c(\tau)$, $\tau = b - (t - a)$ be the parametrization of C. Then $d'(t) = -c'(\tau)$ and

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{s} &= \int_a^b \mathbf{F}(x(t), y(t)) \cdot d'(t)dt \\
&= \int_b^a \mathbf{F}(x(\tau), y(\tau)) \cdot (-c'(\tau))(-d\tau) \\
&= -\int_a^b \mathbf{F}(x(\tau), y(\tau)) \cdot c'(\tau)d\tau \\
&= -\int_{-C} \mathbf{F} \cdot d\mathbf{s}
\end{aligned}$$

[11] A function $x(t)$ is called $\mathcal{C}^n$ if the n-th derivative $x^{(n)}$ is continuous.

Example 3.5. Let $\mathbf{F}(x, y) = \frac{1}{2}(-y, x)$, and C be a circle of radius of r centered at the origin. Let us compute the line integral of $\mathbf{F}$ with counter clockwise orientation of C. Let $c(t) = (r * \cos t, r \sin t)$, $t \in [0, 2\pi]$ be a parametrization of C. Then

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{s} &= \int_0^{2\pi} \frac{1}{2}(-\sin t, \cos t) \cdot (-r \sin t, r \cos t) dt \\ &= \int_0^{2\pi} \frac{1}{2} r \sin^2 t + r \cos^2 t dt \\ &= \pi r\end{aligned}$$

In many cases, we say C is *positively oriented* if it is counter-clockwise. However, the positivity of the orientation of a curve depends on the region bounded by C. For example, the boundary of a simply connected region is positively oriented if it is counter-clockwise. If the region is multiply-connected, then the boundary of a 'hole' is positively oriented if it is clockwise. The rule is : *Suppose that you are walking along the edge of the region. You are walking in positive direction if your* <u>*left*</u> *arm is toward to the region.* We often use the simbol $\oint$ if the line integral is taken over a *closed* curve.

Example 3.6. Let $\mathbf{a}$ be a vector field defined by

$$\mathbf{a} = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) \tag{3.6.1}$$

Let us compute $\int_C \mathbf{a} \cdot d\mathbf{s}$ where C is the curve parametrized by $c(t) = (\cos nt, \sin nt)$, $0 \leq t \leq 2\pi$ for some integer n. By definition of line integral,

$$\begin{aligned}&\int_C \mathbf{a} \cdot d\mathbf{s} \\ &= \int_0^{2\pi} (-\sin nt, \cos nt) \cdot (-n \sin nt, n \cos nt) dt \\ &= \int_0^{2\pi} n dt \\ &= 2n\pi\end{aligned}$$

The line integral gives an angular displacement of start point and end point. In §3.3, we will show that the line integral of the vector field (3.6.1) does not depends on the contour C (as long as C does not passes through the origin), but only depends on the angular displacement between the start and end points.

3.2 Partial derivatives

Definition 3.7. Let $f(x,y)$ be a function defined on a region $D \subset \mathbf{R}^2$ and $(x_0,y_0) \in D$. Let $\mathbf{u} = (a,b)$ a *unit* vector. If the limit

$$D_{\mathbf{u}} f(x_0,y_0) = \lim_{t\to 0} \frac{f(x_0+at, y_0+bt) - f(x_0,y_0)}{t}$$

exists, then it is called the **u-directional derivative** of f at (x_0,y_0).

Example 3.8. Let $f(x,y) = x^2+y^2$. The $(1,0)$-directional derivative of f at $(\frac{1}{2},0)$ is

$$\begin{aligned} D_{(1,0)} f\left(\tfrac{1}{2},0\right) &= \lim_{t\to 0} \frac{f(\frac{1}{2}+t,0) - f(\frac{1}{2},0)}{t} \\ &= \lim_{t\to 0} \frac{\left(\frac{1}{2}+t\right)^2 - (\frac{1}{2})^2}{t} = 1 \end{aligned}$$

For a single-variable function $y = f(x)$, the differential $f'(x_0)$ represents rate of change of $f(x)$ as x approaches to x. The directional derivative naturally extends this concept of differentials. Let

$$c(t) = (x_0,y_0) + t\mathbf{u}$$

Then $c(t)$ approaches to (x_0,y_0) as $t \to 0$. The value of f along c, $f \circ c(t)$ is a single-variable function, thus differential $(f\circ c)'(0)$ can be found. It is not a coincidence that such differential is identical to the $\mathbf{u}$-directional derivative.

$$(f\circ c)'(0) = D_{\mathbf{u}} f(x_0,y_0)$$

The directional derivative is the *rate of chage of f along straight line passing through (x_0,y_0) parallel to* $\mathbf{u}$. Geometrically, it is the slope of the tangent line to the curve obtained by intersection the graph of f and the plane P parallel to $(\mathbf{u},0)$ and $\mathbf{k}$, and containing $(x_0,y_0,0)$. For example, the directional derivative in Example 3.8 can be computed in the following way. Let $c(t)$ be a curve passing $(\frac{1}{2},0)$ with direction $\mathbf{u}$

$$c(t) = (\tfrac{1}{2},0) + t\mathbf{u} = (\tfrac{1}{2}+t,0)$$

Then

$$D_{(1,0)} f\left(\tfrac{1}{2},0\right) = \left.\frac{d}{dt}\right|_{t=0} f(c(t)) = \left.\frac{d}{dt}\right|_{t=0} \left(\frac{1}{2}+t\right)^2 = 1$$

In general, let $c(t)$ be a curve in $\mathbf{R}^2$ passing through (x_0,y_0) at $t=0$, and $c'(0) = \mathbf{u}$. Then $(f\circ c)'(0) = D_{\mathbf{u}} f(x_0,y_0)$ (Exercise ??. Since $c(t)$ may not be a straight line, this implies that the directional derivative does not depends on the shape of a curve, but only on the infinitesimal velocity at the point (x_0,y_0) where the curve passes through.

Let us fix vectors $\mathbf{u},\mathbf{v}$. Suppose that the $\mathbf{u}$-directional derivative $D_{\mathbf{u}} f(x,y)$ exists for all points in a region $D \subset \mathbf{R}^2$. Then

$$g(x,y) = D_{\mathbf{u}} f(x,y)$$

is again a two-variable function defined on D. Suppose that $g(x,y)$ has $\mathbf{v}$-direnctional derivative $D_{\mathbf{v}} g(x_0,y_0)$ at (x_0,y_0). Then we can define a *second* directional derivative of f.

$$D_{\mathbf{v}} D_{\mathbf{u}} f(x_0,y_0) = D_{\mathbf{v}} g(x_0,y_0).$$

Geometrically, the $D_{\mathbf{u}} f$ measure the slope of the graph of f along $\mathbf{u}$-direction. Thus there is a unique vector tagent to the graph of f at $(x_0,y_0,f(x_0,y_0))$ whose projection onto $\mathbf{R}^2$ is parallel to $\mathbf{u}$. The $D_{\mathbf{v}} D_{\mathbf{u}} f$ measure how much such tangent vector *changes* along v-direction at (x_0,y_0). For example, $D_{\mathbf{u}} D_{\mathbf{u}} f$ is the acceleration f in $\mathbf{u}$-direction.

Example 3.9. Let $f(x,y) = x^3+5x^2y+y^3$ and $\mathbf{u} = (\frac{3}{5},\frac{4}{5})$. Then the directional derivative $D_{\mathbf{u}} f(x,y)$ is

$$\begin{aligned} D_{\mathbf{u}} f(x,y) &= \lim_{t\to 0} \frac{f\left(x+\frac{3}{5}t, y+\frac{4}{5}t\right) - f(x,y)}{t} \\ &= \frac{29}{5}x^2 + 6xy + \frac{12}{5}y^2 \end{aligned}$$

Let $g(x,y) = \frac{29}{5}x^2 + 6xy + \frac{12}{5}y^2$. The *second* directional derivative $D_{\mathbf{u}} D_{\mathbf{u}} f$ is

$$\begin{aligned} D_{\mathbf{u}} D_{\mathbf{u}} f(x,y) &= D_{\mathbf{u}} g(x,y) \\ &= \lim_{t\to 0} \frac{g(x+\frac{3}{5}t, y+\frac{4}{5}t) - f(x,y)}{t} \end{aligned}$$

Definition 3.10. Let $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ be the orthonormal vectors in $\mathbf{R}^2$. We denote D_x, D_y for $\mathbf{e}_1$, $\mathbf{e}_2$-directional derivatives with respectively The $D_x f(x_0,y_0)$, $D_y f(x_0,y_0)$ are called the **partial derivatives** of $f(x,y)$ at (x_0,y_0). Conventionally, the partial derivatives are denoted as follows.

$$D_x f(x_0,y_0) = f_x(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)$$

$$D_y f(x_0,y_0) = f_y(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0)$$

We write $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ for the partial derivatives as functions.

Example 3.11. In many cases[12], $f_x(x_0,y_0)$ can be computed by taking the 'usual' derivative with x while treating y as a constant. For example, the partial derivative of $f(x,y) = x^2+3xy^{1/2}$ are

$$f_x = 2x + 3y^{1/2}, \quad f_y = \frac{3}{2y^{1/2}}$$

[12] The simple method of computing partial derivative by treating one variable as a constant fails if the function is not $\mathcal{C}^1$, that is all partial derivatives are not continuous.

The second partial derivatives are denoted as

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = D_{xx} f = D_x(D_x f)$$

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = D_{xy} f = D_x(D_y f)$$

$$\frac{\partial^2 f}{\partial y \partial x} = f_{xy} = D_{yx} f = D_y(D_x f)$$

$$\frac{\partial^2 f}{\partial y^2} = f_{yy} = D_{yy} f = D_y(D_y f)$$

Although f_{xx} and f_{yy} seems obviously distinct functions, but one may believe that f_{xy} and f_{yx} are the same functions. However, $f_{xy} \neq f_{yx}$ in general as we will see below.

Example 3.12. Let

$$f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Let us find $D_{yx} f(0,0)$. First, we compute $D_x f(x,y)$. For $(x,y) \neq (0,0)$, we take the 'usual' partial derivative (cf. Example 3.11).

$$D_x f(x,y) = \frac{(x^4 + 4x^2y^2 - y^4)y}{(x^2+y^2)^2} \tag{3.12.1}$$

The formula (3.12.1) is not defined at $(0,0)$. We have to compute $D_x f(0,0)$ by definition.

$$D_x f(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = 0$$

Thus

$$D_x f(x,y) = \begin{cases} \dfrac{(x^4 + 4x^2y^2 - y^4)y}{(x^2+y^2)^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases} \tag{3.12.2}$$

Then

$$D_{yx} f(0,0) = \lim_{t \to 0} \frac{D_x f(0,t) - D_x f(0,0)}{t} = -1$$

Now try computing $D_{xy} f(0,0)$. Is it same as $D_{yx} f(0,0)$?

Example 3.12 tells us that we cannot change the order of partial derivatives in general. The next theorem tells when could we interchange the order.

Theorem 3.13 (Clairaut)**.** *If the partial derivatives f_{xy}, f_{yx} are continuous at (x_0, y_0), then*

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$$

3.3 Green's theorem

Definition 3.14. Let $\varphi(x,y)$ be a function defined on D. Then the **gradient** of φ at (x_0,y_0) is a vector defined by

$$\nabla\varphi(x_0,y_0) = (\varphi_x(x_0,y_0), \varphi_y(x_0,y_0))$$

The vector field $\nabla\varphi$ is called the **gradient vector field** of φ. Conversely, let $\mathbf{F} : \mathbf{D} \to \mathbf{R}^2$ be a vector field defined on D. A function $\varphi(x,y)$ satisfying

$$\nabla\varphi = (\varphi_x, \varphi_y) = \mathbf{F}$$

is called the **potential function** of $\mathbf{F}$.

Theorem 3.15. *Let $\varphi(x,y)$ be a potential function of a vector field $\mathbf{F}$ defined on D. Then for any piecewise $\mathcal{C}^1$ curve C from p_0 to p_1 lying in D,*

$$\int_C \mathbf{F}\cdot d\mathbf{s} = \varphi(p_1) - \varphi(p_0)$$

holds. That is, the line integral only depends on the start and end points.

Proof. Let $c(t) = (x(t), y(t))$, $a \le t \le b$ be a parametrization of C from $p_0 = c(a)$ to $p_1 = c(b)$. Note that

$$\int_C \mathbf{F}\cdot d\mathbf{s} = \int_C \varphi_x dx + \varphi_y dy = \int_a^b \frac{\partial\varphi}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial\varphi}{\partial y}\frac{\partial y}{\partial t} dt$$

By the Chain rule (Theorem 5.3), we getn

$$\int_a^b \frac{\partial\varphi}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial\varphi}{\partial y}\frac{\partial y}{\partial t} dt = \int_a^b \frac{d}{dt}\varphi(x(t),y(t))dt = \varphi(p_1) - \varphi(p_0)$$

□

Corollary 3.16. *If a vector field $\mathbf{F}$ admits a potential function on D, then*

$$\oint_C \mathbf{F}\cdot d\mathbf{s} = 0$$

for any closed curve C in D.

Example 3.17. Let $D = \{(x,y) \,|\, x, y > 0\}$ be the first quadrant on $\mathbf{R}^2$. The function

$$\theta(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$$

defined on D is a potential function of $\mathbf{a}$ in (3.6.1). In Example 3.6, we show that $\oint_C \mathbf{a}\cdot d\mathbf{s} \neq 0$ for circular closed path $\{x^2+y^2 = 1\}$. First it seems that this contradicts to Corollary 3.16, but it is not true. The circular path C does *not* lie on the domain D.

Definition 3.18. Let $\mathbf{F}(x,y) = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$. If

$$P_y = Q_x$$

then $\mathbf{F}$ is called a **closed** vector field.

Theorem 3.19. *If a $\mathcal{C}^1$-vector field $\mathbf{F}$ admits a potential function, then it is closed.*

Proof. Suppose that $\mathbf{F} = \langle P, Q\rangle = \nabla\varphi$. Then

$$Q_x - P_y = \varphi_{yx} - \varphi_{xy} = 0.$$

□

Example 3.20. The converse of Theorem 3.19 is not true in general. On $\mathbf{R}^2 \setminus \{(0,0)\}$, the vector field $\mathbf{a}$ in (3.6.1) is closed. But it does not admit any potential function. To show this, suppose that $\mathbf{a}$ has a potential function on $\mathbf{R}^2 \setminus \{(0,0)\}$. Then $\oint_C \mathbf{a}\cdot d\mathbf{s}$ must be 0 for any closed curve C. However Example 3.6 shows this is not true for the circular contour with center at $(0,0)$.

Theorem 3.21 (Poincare lemma)**.** *Let D be a convex region in $\mathbf{R}^2$ and $\mathbf{F}$ a vector field defined on D. If $\mathbf{F}$ is closed, then $\mathbf{F}$ admits a potential function. Furthermore, if D is connected, then the potential function is unique up to constant.*

Proof. Let p_0 be a point in D. For $p = (x,y)$ in D, let us define a function $\varphi(x,y)$ as follows.

$$\varphi(x,y) = \int_{p_0}^{p} \mathbf{F}\cdot d\mathbf{s} = \int_{p_0}^{p} Pdx + Qdy \tag{3.21.1}$$

The function $\varphi(x,y)$ is well-defined due to Theorem 3.15. To show that φ is a potential function of $\mathbf{F}$, let us first show that $\varphi_x = P$. Suppose that the path in the integral (3.21.1) near p is given by $c(t) = (x+t, y)$, $\epsilon \in (-\epsilon, 0]$. Then

$$\varphi_x = \frac{\partial}{\partial x}\int_{(x-\Delta x, y)}^{(x,y)} Pdx = P.$$

We can show $\varphi_y = Q$ in a similar way. □

Theorem 3.22 (Green)**.** *Let D be a connected region in $\mathbf{R}^2$ whose boundary is piecewise $\mathcal{C}^1$ curve C. Let $\mathbf{F}(x,y) = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$ be a vector field defined on D. Then*

$$\oint_C \mathbf{F}\cdot d\mathbf{s} = \iint_D Q_x - P_y dA$$

where the curve C is positively oriented.

Proof. Suppose that the region D is given by

$$D = \{(x, y) \,|\, a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

Then

$$\begin{aligned}
\oint_C P dx &= \int_a^b P(x, g_1(x)) dx + \int_b^a P(x, g_2(x)) dx \\
&= \int_a^b P(x, g_1(x)) - P(x, g_2(x)) dx \\
&= -\iint_D P_y dA
\end{aligned}$$

Similarly, $\int_C Q dy = \iint_D Q_x dA$. □

Example 3.23. Let us compute the line integral

$$\oint_C y^2 dx + x^2 dy$$

where C is the boundary of the square with vertices $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ counter clockwise. Applying Green's theorem to $\mathbf{F} = \langle y^2, x^2 \rangle$, we get

$$\begin{aligned}
\oint_C y^2 dx + x^2 dy &= \iint_D 2x - 2y dA \\
&= \int_0^1 \int_0^1 2x - 2y dx dy \\
&= 0.
\end{aligned}$$

3.4 Conservative vector fields

Definition 3.24. Let $\mathbf{F}(x,y) = (f(x,y), g(x,y))$ be a 2-dimensional $\mathcal{C}^1$-vector field. The **curl** of $\mathbf{F}$ is the function

$$\text{curl}\mathbf{F} = g_x - f_y.$$

The **divergence** of $\mathbf{F}$ is the function

$$\text{div}\mathbf{F} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y}.$$

Let us consider the vector field $\mathbf{F}$ as a 'flow' of a fluid. Suppose we drop a ink to a shallow water which flows along $\mathbf{F}$. The divergence of $\mathbf{F}$ is a infinitesimal rate of change of area of the ink droplet proportional to the initial area. That is, if we put $A(t)$ be the area of the ink at t time after the drop at (x_0, y_0), then

$$\text{div}\mathbf{F}(x_0, y_0) = \lim_{\Delta t \to 0} \frac{A'(0)}{A(0)}$$

For example, let

$$\mathbf{r}(x,y) = (x,y)$$

be the **position vector field**. The divergence of $\mathbf{r}$ is 2. This means that if a fluid flows along the vector field $\mathbf{r}$, the area of ink increases at the rate of twice of the initial area of the ink. To verify this, let $(x_0, y_0) = (r_0 \cos\theta_0, r_0 \sin\theta_0)$ be the center of disk domain of radius r_0. Suppose that a ink is dropped at some portion of the disk D

$$D = \{(r\cos\theta, r\sin\theta) \,|\, r_0 \le r \le r_0 + \Delta r,\ \theta_0 \le \theta \le \theta_0 + \Delta\theta\}$$

Let $D(t)$ be the region spreaded by the ink after the time t, and $A(t)$ be the area of $D(t)$. At $t = 0$, the area $A(0)$ is

$$A(0) = \frac{1}{2}(b-a)^2\phi$$

After a short time Δt, the area is approximately changes to

$$A(\Delta t) = \frac{1}{2}(b-a)^2(1+\Delta t)^2\phi$$

Thus the rate of change is

$$\begin{aligned}\frac{A(\Delta t) - A(0)}{\Delta t} &\approx 2\frac{1}{2}(b-a)^2\phi = 2A(0) \\ &= (\text{div}\,\mathbf{a})A(0)\end{aligned}$$

Thus the divergence div$\mathbf{F}$ is the ratio of $A'(0)$ to $A(0)$.

The curl of $\mathbf{F}$ measures the ratio between the area of a small circular region and the work done by the vector field $\mathbf{F}$ along the boundary of the region. That is, let C_r be a circle at $p \in D$ of radius r, then

$$\text{rot}\mathbf{F}(p) = \lim_{r\to 0} \frac{1}{\pi r^2} \oint_{C_r} \mathbf{F} \cdot d\mathbf{s}.$$

Definition 3.25. If div$\mathbf{F} = 0$, then $\mathbf{F}$ is called **incompressible**. If rot$\mathbf{F} = \mathbf{0}$, then $\mathbf{F}$ is said to be **irrotational**.

Example 3.26. In fluid dynamics, the *Euler equation*

$$\begin{aligned}\frac{\partial \mathbf{F}}{\partial t} &= -\nabla f + g \\ \nabla \cdot \mathbf{F} &= 0\end{aligned}$$

assumes that the fluid is *incompressible*. The *irrotationality* of a vector fields tells us that an object flowing along the vector field $\mathbf{F}$ does not 'spin'.

Theorem 3.27. *A vector field $\mathbf{F}$ is irrotational if it admits a potential function φ, that is, $\mathbf{F} = \nabla\varphi$.*

Proof. □

Theorem 3.28. *Let D be a simply connected region and $\mathbf{F}$ defined on D. Then the following hold.*

$$\oint_C \mathbf{F} \cdot d\mathbf{s} \iff \mathbf{F} = \nabla\varphi \iff \nabla \times \mathbf{F} = 0$$

Theorem 3.29 (2-dimensional divergence theorem)**.** *Let D be a connected region in $\mathbf{R}^2$ and $C = \partial D$ a boundary curve. Let $\mathbf{n}$ be a 2-dimensional vector field along C which is perpendicular (i.e. normal) to the curve. Then*

$$\oint_C \mathbf{F} \cdot \mathbf{n} ds = \iint_D \nabla \cdot \mathbf{F} dA$$

Now, let us generalize curl and divergence to 3-dimensional vector spaces. Let the symbol ∇ (called the *nabla*) be the differential operator

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

Definition 3.30. For a 3-variable function $f = f(x,y,z)$, the **gradient** ∇f is a vector field defined by

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

For a 3-dimensional vector field $\mathbf{F} = (P, Q, R)$, the **curl** $\nabla \times \mathbf{F}$ and the **divergence** $\nabla\mathbf{F}$ are defined as follows.

$$\begin{aligned}\nabla \times \mathbf{F} &= (R_y - Q_z, P_z - R_x, Qx - P_y) \\ &= \begin{vmatrix} i & j & k \\ \partial/\partial_x & \partial/\partial_y & \partial/\partial_z \\ P & Q & R \end{vmatrix} \\ \nabla\mathbf{F} &= P_x + Q_y + R_z\end{aligned}$$

Note that the curl $\nabla \times \mathbf{F}$ is a vector, while the divergence $\nabla\mathbf{F}$ is a scalar.

Theorem 3.31. *The following identities hold.*

1. $\nabla \times (\nabla f) = 0$
2. $\nabla \cdot (\nabla \times \mathbf{F}) = 0$
3. $\nabla \cdot (\mathbf{F} + \mathbf{G}) = \nabla \cdot \mathbf{F} + \nabla \cdot \mathbf{G}$
4. $\nabla \times (\mathbf{F} + \mathbf{G}) = \nabla \times \mathbf{F} + \nabla \times \mathbf{G}$
5. $\nabla \cdot (f\mathbf{F}) = f\nabla \cdot \mathbf{F} + \mathbf{F} \cdot \nabla f$
6. $\nabla \times (f\mathbf{F}) = f\nabla \times \mathbf{F} + \nabla f \times \mathbf{F}$
7. $\nabla \cdot (\mathbf{F} + \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$
8. $\nabla \cdot (\nabla f \times \nabla g) = 0$
9. *Denote* $\nabla^2 = \nabla \cdot \nabla$. *Then*

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \cdot \mathbf{F}$$

Definition 3.32. Let $\mathbf{F}$ be a 3-dimensional vector field.

- If the line integral $\int_C \mathbf{F} \cdot d\mathbf{s}$ only depends on the start and end point, i.e.

$$\oint_C \mathbf{F} \cdot d\mathbf{s} = 0$$

  for any closed curve C, then $\mathbf{F}$ is called **conservative**.

- If a function $\varphi(x, y, z)$ satisfies

$$\nabla\varphi = \mathbf{F},$$

  then φ is called a **potential function** of $\mathbf{F}$.

- If $\nabla \times \mathbf{F} = \mathbf{0}$, then $\mathbf{F}$ is called **closed**.

3.5 Exerciese

1. Evaluate
$$\int_C xy^2 dx - xydy$$
where C is the semicircular arc from $(0,2)$ to $(0,-2)$ clockwise.

2. Show that (3.12.2) is not continuous at $(0,0)$. This explains why $f_{xy}(0,0) \neq f_{yx}(0,0)$.

3. Recall that for single variable function $f(x)$, the existence of derivative $f'(x_0)$ implies that $f(x)$ is continuous atx_0. However, the existence of directional derivative $D_{\mathbf{u}}f(x_0,y_0)$ does *not* imply that $f(x,y)$ is continous at (x_0,y_0). Find examples satisfying the following.

   (a) A function $f(x,y)$ which is continuous at $(0,0)$ but $D_{\mathbf{u}}f(0,0)$ does not exists for all $\mathbf{u}$.

   (b) A function $f(x,y)$ which has $D_{\mathbf{u}}f(0,0)$ for all $\mathbf{u}$, but is not continuous at $(0,0)$.

   Therefore, the continuity and existence of directional (or partial) derivatives have no implication to each others.

4. Show that $\mathbf{F}(x,y) = (x^2+y^2, x^2-y^2)$ has no potential function.

5. Find the line integrals using Green's theorem. (All closed curves are oriented counter-clockwise.)

   (a) Let C is the circle at the origin with radius 2. Evaluate
   $$\oint_C (x-y)dx + (x+y)dy$$

   (b) Let C be the boundary of the region bounded by $y = x^2$ and $x = y^2$. Evaluate
   $$\oint_C (y + e^{\sqrt{x}})dx + (2x + \cos^2 y)dy$$

   (c) Let C be the circle $(x-2)^2 + (y-3)^2 = 1$. Evaluate
   $$\int_C \left(y - \log(x^2+y^2)\right) dx + \left(2\tan^{-1}(y/x)\right) dy$$

6. Use Green's theorem to the following questions.

   (a) Find the area between the ellipse $\dfrac{x^2}{9} + \dfrac{y^2}{4} = 1$ and the circle $x^2 + y^2 = 25$.

   (b) Find the area enclosed by the hypocycloid
   $$c(t) = (a\cos^3 t, a\sin^3 t), \quad 0 \le t \le 2\pi$$

   (c) Let (a_1,b_1), (a_2,b_2), $\cdots$, (a_n,b_n) be the successive vertices of n-polygon. Show that the area inside the polygon is
   $$\frac{1}{2}\left(\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} + \cdots + \begin{vmatrix} a_{n-1} & b_{n-1} \\ a_n & b_n \end{vmatrix} + \begin{vmatrix} a_n & b_n \\ a_1 & b_1 \end{vmatrix}\right)$$

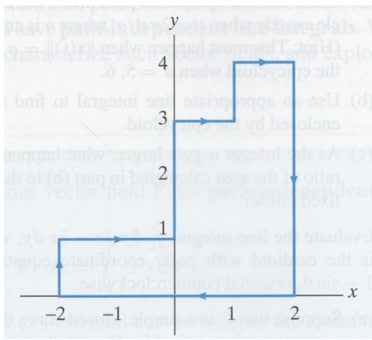
7. Show that the vector field $\mathbf{F} = (-y, x)$ is incompressible.

8. Let $\mathbf{F} = \langle 3y, -4x\rangle$. Find $\oint_{\partial D} \mathbf{F}\cdot d\mathbf{s}$ for the following region D.

   (a) $D = \{(x,y) \,|\, x^2 + y^2 \le 4\}$

   (b) $D = \{(x,y,) \,|\, x^2 + 2y^2 \le 4\}$

9. Evaluate $\int_C (x^4y^5 - 2y)dx + (3x + x^5y^4)dy$ where C is as shown below.



10. Evaluate $\oint_C 5ydx - 3xdy$ where C is the cardioid $r = 1 - \sin\theta$ oriented counter-clockwise.

11. Determine the value of
$$\oint_C \frac{xdx + ydy}{x^2+y^2}$$
where C is a simple closed curve enclosing the origin. (Note that Green's theorem does not apply since $\mathbf{F} = \left\langle \dfrac{x}{x^2+y^2}, \dfrac{y}{x^2+y^2}\right\rangle$ is not defined on the origin.)

12. Let $\mathbf{F}, \mathbf{G}$ be vector fields. Show that

   (a) $\nabla\cdot(\mathbf{F}+\mathbf{G}) = \mathbf{G}\cdot(\nabla\times\mathbf{F}) - \mathbf{F}\cdot(\nabla\times\mathbf{G})$

   (b) $\nabla\times(\nabla\times\mathbf{F}) = \nabla(\nabla\cdot\mathbf{F}) - \nabla\cdot\nabla\cdot\mathbf{F}$

   (c) $\nabla\times(\mathbf{F}\times\mathbf{G}) = (\nabla\cdot\mathbf{G})\mathbf{F} - (\nabla\cdot\mathbf{F})\mathbf{G} + (\mathbf{G}\cdot\nabla)\cdot\mathbf{F} - (\mathbf{F}\cdot\nabla)\cdot\mathbf{G}$

4 Stokes' theorem

Stokes' theorem is one of the most important theorem in integral theory. Not only it connects the line integrals to surface integrals, the general form of Stokes' theorem states that we can compute the integrals on $(n-1)$-form to n-form (cf. §**??**). Important questions are:

- What is a surface integral?
- How does Stokes' theorem related to Green's theorem?
- How does Stokes' theorem related to Divergence theorem?
- What is a Gauss' law? How do we express it?

4.1 Surface integrals

There are two types of surface integrals: a surface integral of a real-valued function and a surface integral of a vector field. In both cases, we need to parametrize the surface.

Definition 4.1. Let D be a region in $\mathbf{R}^2$ and S a suface in $\mathbf{R}^3$. A one-to-one correspondence $X : D \to S$

$$X(u,v) = (x(u,v), y(u,v), z(u,v))$$

is called the **parametrization** of S if every partial derivative of x, y, z is continuous on D.

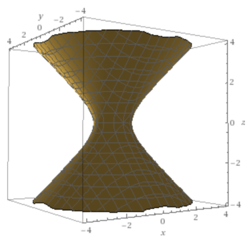
Example 4.2. The parametrization of the surface

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

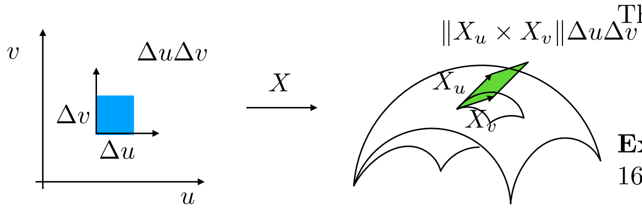
is given as follows.

$$X(u,v) = (\cosh(u)\cos(v), \cosh(u)\sin(v), \sinh(u)), \quad u \in \mathbf{R}, v \in [0, 2\pi]$$

The surface is called the *one-sheeted hyperbola*.



Note that in Definition 4.1, the area $[u, u+\Delta u] \times [v, v+\Delta v]$ is mapped to a curved surface on S. The area on S can be approximated by the parallelogram made by $X_u \Delta u$ and $X_v \Delta v$. Therefore, the area of the image is $\|X_u \times X_v\|$ times bigger than the corresponding area in uv-plane.



Thus it is logical tot define the surface integral as follows.

Definition 4.3. Let $f(x,y,z)$ be a function defined on a surface S, parametrized by $X : D \to S$. The **surface integral** of f on S is defined by

$$\iint_S f dS = \iint_D (f \circ X)(u,v) \|\mathbf{X}_u \times \mathbf{X}_v\| dudv \tag{4.3.1}$$

Example 4.4. Let S be the surface defined by the graph of $z = \sqrt{x^2+y^2}$ over the disk $x^2+y^2 \le 1$. Let us evaluate

$$\iint_S z dS$$

Let

$$X(r,\theta) = (r\cos\theta, r\sin\theta, r), \quad 0 \le r \le 1, 0 \le \theta 2\pi$$

be a parametrization of S. Then

$$X_r = (\cos\theta, \sin\theta, 1), \quad X_\theta = (-r\sin\theta, r\cos\theta, 0)$$

$$X_r \times X_\theta = (-r\cos\theta, -r\sin\theta, r)$$

Thus

$$\iint_S z dS = \int_0^1 \int_0^{2\pi} r\sqrt{2} r d\theta dr = \sqrt{2}\pi$$

Example 4.5. If $f(x,y,z)$ is a density function, the surface integral $\iint_S f dS$ is the mass of the surface. Furthermore, we can compute the *center of mass* of a surface. The center of mass may not lie on the surface, but it represent the mass as if its mass is concentrated in one point. For example, let us find the center of mass $(\bar{x}, \bar{y}, \bar{z})$ of the surface $z = \sqrt{1-x^2-y^2}$ over the xy-plane, provided that the density is constant. By symmetry, we know that $\bar{x} = \bar{y} = 0$. To obtain $\bar{z}$, we should compute

$$\bar{z} = \iint_S z dS$$

Let $X(r,\theta) = (r\cos\theta, r\sin\theta, \sqrt{1-r^2})$, $0 \le r \le 1$, $0 \le \theta \le 2\pi$ be a parametrization of S. Then

$$X_r = \left(\cos\theta, \sin\theta, \frac{-r}{\sqrt{1-r^2}}\right), \quad X_\theta = (-r\sin\theta, r\cos\theta, 0)$$

$$X_r \times X_\theta = \left(\frac{r^2\cos\theta}{\sqrt{1-r^2}}, \frac{r^2\sin\theta}{\sqrt{1-r^2}}, r\right).$$

Thus

$$\iint_S z dS = \int_0^1 \int_0^{2\pi} \sqrt{1-r^2} \frac{r^2}{\sqrt{1-r^2}} = \frac{2\pi}{3}$$

Example 4.6. Let S be the sphere $x^2+y^2+z^2 = 16$. Let us verify that

$$\iint_S \frac{dS}{\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}} = 8\pi$$

Before we begin to compute the surface integral directly, we must observe that

$$\begin{aligned}&\iint_S \frac{dS}{\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}} \\ &= \iint_S \frac{dS}{\sqrt{x^2+y^2+(z-\sqrt{14})^2}}\end{aligned}$$

Let $X(\theta, \varphi)$ be the spherical coordinate (cf. Definition **??**) which parametrizes the sphere S as

$$X(\theta, \varphi) = (4\sin\theta\cos\varphi, 4\sin\theta\sin\varphi, 4\cos\theta)$$

for $0 \le \theta \le \pi, 0 \le \varphi \le 2\pi$. Then

$$\begin{aligned}
&\iint_S \frac{dS}{\sqrt{x^2+y^2+(z-\sqrt{14})^2}} \\
&= \int_0^\pi \int_0^{2\pi} \frac{16\sin\theta}{\sqrt{16\sin^2\theta + (4\cos\theta - \sqrt{14})^2}} d\varphi d\theta \\
&= \int_0^\pi \int_0^{2\pi} \frac{16\sin\theta}{\sqrt{30 - 8\sqrt{14}\cos\theta}} d\varphi d\theta \\
&= \frac{4\pi}{\sqrt{14}} \sqrt{30 - 8\sqrt{14}\cos\theta}\Big|_0^\pi \\
&= \frac{4\pi}{\sqrt{14}} \left(\sqrt{30+8\sqrt{14}} - \sqrt{30-8\sqrt{14}} \right) \\
&= 8\pi
\end{aligned}$$

Now, let us compute the surface integral of a vector field. As we have observed in §2.3, the surface integral of a vector field represent the *flux* (cf. Definition 2.22). In order to make sense what flux means, we need to determine the *orientation* of a surface.

Definition 4.7. Let $X : D \to S$ be a parametrization of S. If the vector field

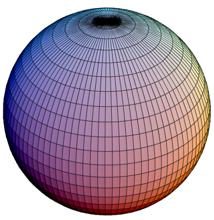
$$\mathbf{n} = \frac{\mathbf{X}_u \times \mathbf{X}_v}{\|\mathbf{X}_u \times \mathbf{X}_v\|} \tag{4.7.1}$$

is continuous, then we say S **has an orientation** and $\mathbf{n}$ is called an **orientation**.

Example 4.8. The vector field

$$\mathbf{A} = \frac{(x, y, z)}{\sqrt{x^2+y^2+z^2}}$$

is an orientation of the sphere $x^2 + y^2 + z^2 = 1$.

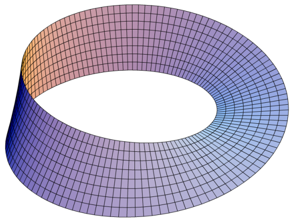


A surface with an orientation always admits *two* orientation.

Example 4.9. There are surfaces with no orientation. The Möbius band, parametrized by

$$\begin{aligned}
X(r, \theta) = ((3 + r\cos(\theta/2))\cos\theta\,, \\
(3 + r\cos(\theta/2))\sin\theta, \\
r\sin(\theta/2))\,, \quad 0 \le \theta \le 2\pi, -1 \le r \le 1
\end{aligned}$$

does not have an orientation.



Definition 4.10. Let S be an oriented surface, and $\mathbf{F}$ a vector field defined on S. Then the **surface integral** of $\mathbf{F}$ is defined by

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_S \mathbf{F} \cdot \mathbf{n} dS$$

and called the **flux** of $\mathbf{F}$ on S.

Note that $\mathbf{F} \cdot \mathbf{n}$ is a scalar function. Thus for a parametrization $X : D \to S$ of S,

$$\begin{aligned}
\iint_S \mathbf{F} \cdot \mathbf{n} dS &= \iint_D (\mathbf{F} \circ X)(u, v) \frac{\mathbf{X}_u \times \mathbf{X}_v}{\|\mathbf{X}_u \times \mathbf{X}_v\|} \|\mathbf{X}_u \times \mathbf{X}_v\| \\
&= \iint_D (\mathbf{F} \circ X)(u, v) \cdot (\mathbf{X}_u \times \mathbf{X}_v) dudv
\end{aligned}$$

Example 4.11. Let the orientation $\mathbf{n}$ of the surface $z = 1 - x^2 - y^2$, $z \ge 0$ is upward direction, i.e. $\mathbf{n} \cdot \mathbf{k} \ge 0$. Let us compute the surface integral of

$$\mathbf{F} = \frac{(x, y, z)}{x^2+y^2+z^2}$$

over the surface S. Let

$$X(x, y) = (x, y, 1 - x^2 - y^2), \quad x^2 + y^2 \le 1$$

be a parametrization of S. Then

$$\begin{aligned}
\iint_S \mathbf{F} \cdot d\mathbf{S} &= \iint_D \frac{(x, y, 1-x^2-y^2)}{x^2+y^2+(1-x^2-y^2)^2} \cdot (-2x, 2y, 1) dxdy \\
&= \iint_D \frac{1 - 3x^2 + y^2}{x^2+y^2+(1-x^2-y^2)^2} dxdy
\end{aligned}$$

where $D = \{x^2 + y^2 \le 1\}$.

4.2 Stokes theorem

Green's theorem connects line integral over a planar curve to a surface integral over a planar surface. Stokes' theorem generalizes Green's theorem to 3-dimensional space.

Theorem 4.12 (Stokes)**.** *Let S be an oriented surface with a piecewise continuous boundary C. For $\mathbf{F}$ be a continuous vector field defined on S. Then*

$$\oint_C \mathbf{F} \cdot d\mathbf{s} = \iint_S \nabla \times \mathbf{F} \cdot d\mathbf{S}$$

where C and S are positively oriented[13].

Proof. Let $\mathbf{F} = \langle P, Q, R\rangle$. Suppose that the surface S is given by the graph of a function $z = f(x, y)$ on a bounded domain $D \subset \mathbf{R}^2$. Let $X(x, y) = (x, y, f(x, y))$ be the parametrization of S. Then

$$\begin{aligned}
&\iint_S \nabla \times \mathbf{F} d\mathbf{S} \\
&= \iint_D -(R_y - Q_z)f_x - (P_z - R_x)f_y + (Q_x - P_y)dA \\
&= \iint_D \frac{\partial}{\partial x}(Q + Rf_y) - \frac{\partial}{\partial y}(P + Rf_x)\, dA
\end{aligned}$$

Let C' be a planar curve which bounds that area D. By Green's theorem, the last integral becomes

$$\begin{aligned}
\oint_{C'} (P + Rf_x)dx + (Q + Rf_y)dy &= \oint_C Pdx + Qdy + Rdz \\
&= \oint_C \mathbf{F} \cdot d\mathbf{s}
\end{aligned}$$

□

Remark 4.13. The Stokes' theorem provides the meaning of curl $\nabla \times \mathbf{F}$ of a vector field $\mathbf{F}$. Suppose that the surface S is planar disk with sufficiently small radius r centered at (x_0, y_0, z_0). Then

$$\oint_C \mathbf{F} \cdot d\mathbf{s} \approx \text{area} S \cdot (\nabla \times \mathbf{F})(x_0, y_0) \circ \mathbf{n}$$

$$(\nabla \times \mathbf{F})(x_0, y_0) \circ \mathbf{n} \approx \frac{1}{\text{area} S} \oint_C \mathbf{F} \cdot d\mathbf{s}$$

where $\mathbf{n}$ is the orientation of S. Therefore, the curl of a vector field $\mathbf{F}$ at the point p on the surface S has a projection onto the orthogonal direction of a surface S equal to the work done by $\mathbf{F}$ along the neighboring boundary of the point p on S.

Example 4.14. Let S be the surface bounded by $z = 1 - x^2 - y^2$, $z \geq 0$ with upward orientation ($\mathbf{n} \cdot \mathbf{k} \geq 0$). Let us confirm that Stokes' theorem holds for $\mathbf{F} = (y, -x, 0)$. First, the line integral along the boundary curve $c(t) = (\cos(t), \sin(t), 0)$, $t \in [0, 2\pi]$.

$$\begin{aligned}
\oint_C \mathbf{F} \cdot d\mathbf{s} &= \int_0^{2\pi} \langle \sin t, -\cos t, 0\rangle \cdot \langle -\sin t, \cos t, 0\rangle dt \\
&= -\int_0^{2\pi} dt = -2\pi
\end{aligned}$$

Let $D = \{x^2 + y^2 \leq 1\}$. Since $\nabla \times \mathbf{F} = (0, 0, -2)$, we get

$$\begin{aligned}
\iint_S \nabla \times \mathbf{F} \cdot d\mathbf{S} &= \iint_S (0, 0, -2) \cdot d\mathbf{S} \\
&= \iint_D -2dA \\
&= -2\pi
\end{aligned}$$

[13] The boundary is **postively** oriented if the direction is counter-clockwise with $\mathbf{n}$ being *upward.*

4.3 Divergence theorem

Theorem 4.15. *Let V be a region in $\mathbf{R}^3$ whose boundary $S = \partial V$ is a closed surface.*[14] *For a vector field $\mathbf{F}$ defined on V,*

$$\iint_S \mathbf{F}\cdot d\mathbf{S} = \iiint_V \nabla\cdot\mathbf{F}dV$$

where the orientation of S is the outward direction.

Proof. Note that for $\mathbf{F} = (P, Q, R)$,

$$\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_S P\mathbf{i}\cdot\mathbf{n}dS + \iint_S Q\mathbf{j}\cdot\mathbf{n}dS + \iint_S R\mathbf{k}\cdot\mathbf{n}dS$$

Suppose that the volume V is given by

$$V = \{(x, y, z) \mid h_1(x, y) \le z \le h_2(x, y),\ (x, y) \in D\}.$$

where D is the region in $\mathbf{R}^2$ on which the volume V is defined. Then

$$\begin{aligned}
&\iint_S R\mathbf{k}\cdot\mathbf{n}dS \\
&= \iint_{S_1} R\mathbf{k}\cdot\mathbf{n}dS + \iint_{S_2} R\mathbf{k}\cdot\mathbf{n}dS \\
&= \iint_D R(x, y, h_2(x, y)) - R(x, y, h_1(x, y))dxdy \\
&= \iiint_V RdV
\end{aligned}$$

Similarly, we can show that $\iint_S P\mathbf{i}\cdot\mathbf{n}dS = \iiint_V PdV$ and $\iint_S Q\mathbf{i}\cdot\mathbf{n}dS = \iiint_V QdV$. □

Example 4.16. Let us compute the surface integral $\iint_S \mathbf{F}\cdot d\mathbf{S}$ where $\mathbf{F} = (z^2, \frac{1}{3}x^3 + \tan z, z + y^2)$ and S is the closed surface $x^2 + y^2 + z^2 = 1$. Since $\nabla\cdot\mathbf{F} = 1$, by divergence theorem,

$$\iint_S \mathbf{F}\cdot d\mathbf{S} = \iiint_{x^2+y^2+z^2\le 1} dV = \frac{4}{3}\pi$$

Example 4.17. Let S be a parabola $x^2 + y^2 + z = 2$ above the plane $z = 1$. Let us find the flux of $\mathbf{F} = (z\tan^{-1}(y^2), z^3\ln(x^2+1), z)$ to the upward direction of S. Here, we can apply the divergence theorem on the volume V bounded by $1 \le z \le 2 - x^2 - y^2$, $x^2 + y^2 \le 1$. Let S' be the surface $\{x^2 + y^2 \le 1\} \cap \{z = 1\}$. Since $\nabla\cdot\mathbf{F} = 1$,

$$\begin{aligned}
\iint_S \mathbf{F}\cdot d\mathbf{S} &= \iiint_V dV - \iint_{S'} \mathbf{F}\cdot d\mathbf{S} \\
&= \iint_D 1 - x^2 - y^2 dxdy + \iint_D dA \\
&= \frac{\pi}{4} + \pi = \frac{5}{4}\pi
\end{aligned}$$

where D is the region $x^2 + y^2 \le 1$ on xy-plane.

[14] A surface is called **closed** if it has no boundary curve.

Example 4.18. For example, let

$$\mathbf{F} = \frac{(x, y, z)}{(x^2 + y^2 + z^2)^{3/2}}$$

and S be the surface $z = 4 - x^2 - y^2$, $z \ge 0$ with orientation $\mathbf{n}\cdot\mathbf{k} \ge 0$. Let us use divergence theorem to compute the flux $\iint_S \mathbf{F}\cdot d\mathbf{S}$. If we let $V = \{z \le 4 - x^2 - y^2, z \ge 0\}$, we cannot use divergence theorem becaus $\mathbf{F}$ is not defined at the origin $(0, 0, 0)$. Instead, let $V = \{\sqrt{4 - x^2 - y^2} \le z \le 4 - x^2 - y^2\}$. Then $\mathbf{F}$ is well-defined on V. Since $\nabla\cdot F = 0$, we have

$$\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_{z=\sqrt{4-x^2-y^2}} \mathbf{F}\cdot d\mathbf{S}$$

Since $\mathbf{F}$ is orthogonal to the upper hemisphere $z = \sqrt{4 - x^2 - y^2}$, $z \ge 0$,

$$\begin{aligned}
\iint_{z=\sqrt{4-x^2-y^2}} \mathbf{F}\cdot d\mathbf{S} &= \iint_{z=\sqrt{4-x^2-y^2}} \mathbf{F}\cdot\mathbf{n}dS \\
&= \iint_{z=\sqrt{4-x^2-y^2}} \frac{1}{4}dS = 2\pi
\end{aligned}$$

Hyosang Kang (hyosang@dgist.ac.kr)

4.4 Exercises

1. Does $\mathbf{F} = (2xy, x^2 + 2yz, y^2)$ admit a potential function? If so, find one.

2. Does the vector field

$$\mathbf{F} = \Big(\log x + \sec^2(x+y), \sec^2(x+y) + \frac{y^2}{y^2+z^2}, \frac{z}{y^2+z^2}\Big)$$

   admits a potential function?

3. Is there a vector field $\mathbf{G}$ satisfying $\nabla \times \mathbf{G} = (x \sin y, \cos y, z - xy)$? If so, find one.

4. Is $\mathbf{F} = (2x-3, -z, \cos z)$ a conservative field?

5. Find the parametrization of $x^2 - y^2 - z^2 = 1$.

6. Let C be the boundary curve of the surface bounded by $2x + y + z = 2$ and $x, y, z \geq 0$. Compute the line integral $\oint_C \mathbf{F} \cdot d\mathbf{s}$ where $\mathbf{F} = (xz, xy, 3xz)$ and C is oriented counter-clockwise.

7. Let S be the surface bounded by $z = e^{-x^2-y^2}$ and $z \geq 1/e$. Let $\mathbf{n}$ be the orientation of S satisfying $\mathbf{n} \cdot \mathbf{k} \geq 0$. Find the flux of

$$\mathbf{F} = (e^{x+y} - xe^{y+z}, e^{y+z} - e^{x+y}, 2)$$

   on S to the direction of $\mathbf{n}$.

8. Compute Example 2.23 again by using Stokes theorem. (Note that it only works for $\nabla \cdot \mathbf{F} = 0$)

9. Let C be the intersection of $z = 1 - 2(x^2 + y^2)$ and $z = x^2 - y^2$ oriented counter-clockwise. Find $\oint_C \mathbf{F} \cdot d\mathbf{s}$ where

$$\mathbf{F} = (y \cos(x) - yz, \sin x, e^z)$$

10. Let $S = \partial V$ be a closed surface. Prove the following.

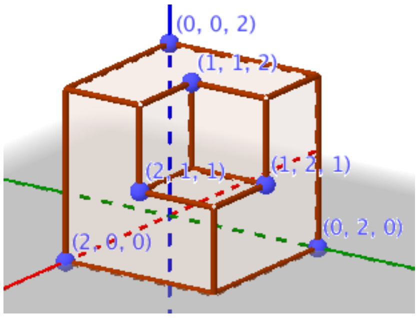
    (a) For any constant vector field $\mathbf{C}$,

$$\iint_S \mathbf{C} \cdot d\mathbf{S} = 0$$

    (b) For any vector field $\mathbf{F}$,

$$\iint_S \nabla \times \mathbf{F} \cdot d\mathbf{S} = 0$$

11. Find the flux of $\mathbf{F} = (x, y, z)$ on the surface below. (It is a surface of a cube with a corner removed.)



12. Let C be a curve parametrized by $c(t) = (\cos t, \sin t, \sin 2t)$. Evaluate

$$\oint_C (y^3 + \cos x)dx + (\sin y + z^2)dy + xdz$$

    using Stoke's theorem. (Hint: the curve C lies on the surface $z = 2xy$.)

13. Let S be a surface bounded by $z = e^{1-x^2-y^2}$ and $z \geq 1$. Let us orient S upward. Evaluate

$$\int_S \langle x, y, 2-2z \rangle \cdot d\mathbf{S}$$

    using Gauss's theorem.

14. (15' Final) Let

$$\mathbf{A} = \frac{\langle x, y, z \rangle}{(x^2+y^2+z^2)^{3/2}}$$

    Let S be a surface bounded by $z = 4 - x^2 - y^2$, $z \geq 0$, orieted upward. Find

$$\iint_S \mathbf{A} \cdot d\mathbf{S}$$

.

5 Derivatives

In this section, we define continuity and differentiability of multivariable functions. As in single variable calculus, we will use limit to define continuity and differentiability of multivariable functions. However, the difference arises due to the existence of two or more independent variables in multivariable functions. For example, given a point (x_0, y_0) in $\mathbf{R}^2$, there are infinitely many paths to take a limit of a point to (x_0, y_0). We will find that the general definition of continuity applies not only to a single variable functions but also to multivariable functions in §5.1. Definitions of directional derivatives such as partial derivatives which will be explained in §5.2. The section §5.3 covers the most important subject in this section, the *chain rule*. It gives how the derivative of a composition of two (or more) multivariable functions can be computed by derivatives of component functions.

- How do we define continuity and differentiability of $f(x, y)$?
- What is the meaning of derivative of $f(x, y)$?
- What is the chain rule?
- How do we generalize the continuity and differentiability to n-variable function $f(x_1, \cdots, x_n)$?

5.1 Continuity of $f(x, y)$

Intuitively, we can say that a function $f(x, y)$ has a limit L at (x_0, y_0) if the value of the function approaches to L when (x, y) *gets close to* (x_0, y_0), that is $(x, y) \to (x_0, y_0)$. But how do we define $(x, y) \to (x_0, y_0)$ rigorously?

Definition 5.1. Let $f(x, y)$ be a two-variable function defined on the entire plane $\mathbf{R}^2$. A constant L is called the **limit of f at** (x_0, y_0) if for any (arbitrary small) $\epsilon > 0$, there exists a (small) $\delta > 0$ such that whenever the distance between (x, y) and (x_0, y_0) is less than δ, the inequality

$$|f(x, y) - L| < \epsilon$$

holds. In such case, we simply write

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

Definition 5.2. The function $f(x, y)$ is **continuous at** (x_0, y_0) if

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$$

At this point, we need to check the difference between continuity of $f(x)$ and $f(x, y)$. We say $f(x)$ is continuous at x_0 if

$$\lim_{x\to x_0} f(x) = f(x_0)$$

The limit $\lim\limits_{x\to x_0}$ assumes that both right and left limits exists and equals to each others. Since there are only two ways to approach x_0 in one-dimensional space $\mathbf{R}$, the concept of the limit is intuitively clear. However, the limit $\lim\limits_{(x,y)\to(x_0,y_0)}$ is not intuitive and this makes two kinds of problems. First, there are infinitely many paths to approach (x_0, y_0) in $\mathbf{R}^2$. Second, it is impossible to prove that the value of a function $f(x, y)$ approach to the same value (of $f(x_0, y_0)$ if we want to prove the continuity) for every path. Definition 5.1 not only solves these problems but also gives general definition of continuity which applies both single and multivariable functions.

Example 5.3. Let us show that

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \tag{5.3.1}$$

is continuous at $(0, 0)$. For any $\epsilon > 0$, let $\delta = \epsilon$. Then whenever $\sqrt{x^2 + y^2} < \delta$, $|x|, |y| < \delta$ holds. Thus

$$\left| \frac{xy}{\sqrt{x^2 + y^2}} - 0 \right| = \frac{|x| \cdot |y|}{\sqrt{x^2 + y^2}} \leq |y| < \delta = \epsilon$$

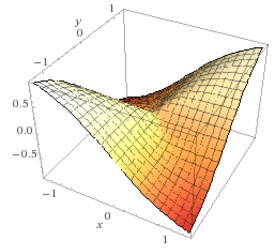


Figure 5.1: The graph of (5.3.1)

We say that a function $f(x)$ is continuous if its graph is continuously connected. Likewise, a two-variable functions $f(x, y)$ is continuous if its graph looks *connected*. That is, the graph of a continuous function $f(x, y)$ has no hole or *cliff*. Therefore, an easy way to prove the discontinuity of $f(x, y)$ is to find a path to (x_0, y_0) where the value of function does not converges to $f(x_0, y_0)$.

Proposition 5.4. *A function $f(x, y)$ is not continuous at (x_0, y_0) if there is a path $c(t) = (x(t), y(t))$ which converges to (x_0, y_0) while $f \circ c(t)$ does not converges to $f(x_0, y_0)$.*

Example 5.5. Let us prove that the function

$$f(x, y) = \begin{cases} \dfrac{x^4 - 4y^2}{x^2 + 2y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \tag{5.5.1}$$

is not continuous at $(0, 0)$. First, let us approach $(0, 0)$ along $c(t) = (0, t)$. Then the value of $f(x, y)$ goes to

$$f(0, t) = \frac{-4t^2}{2t^2} = -2 \to -2$$

which is different from $f(0, 0) = 0$. Thus $f(x, y)$ is not continuous at $(0, 0)$.
Meanwhile, if we approach $(0, 0)$ along $c(t) = (t, 0)$, then the value of $f(x, y)$ goes to

$$f(t, 0) = \frac{t^4}{t^2} = t^2 \to 0$$

Figure 5.5 shows that there are cliff at the point $(0, 0)$, which causes the discontinuity.

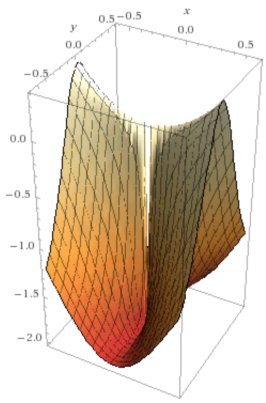


Figure 5.2: The graph of (5.5.1)

5.2 Differentiability of $f(x, y)$

There are two ways to define differentiability of $f(x, y)$. One uses the linear approximation, and the other uses the vector. First, we gives the definition of continuity using linear approximations.

Definition 5.6. Let $f_x(x_0, y_0,\ f_y(x_0, y_0)$ be the partial derivatives of $f(x, y)$. Then the function

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

is called the **linear approximation** of $f(x, y)$ at (x_0, y_0).

Note that the linear approximation $L(x, y)$ is a linear function, that is a polynomial of degree 1. The equation $z = L(x, y)$ gives the *tangent plane* to the graph of $z = f(x, y)$ at $(x_0, y_0, f(x_0, y_0)$. To see this, we observe that the the normal vector $\mathbf{n}$ of the surface $z = f(x, y)$ at (x_0, y_0) is

$$\mathbf{n} = (1, 0, f_x) \times (0, 1, f_y) = (-f_x, -f_y, 1)$$

where the partial derivatives are computed at $(x_0, y_0$. Since the tangent plane passes through $(x_0, y_0, f(x_0, y_0)$, the equation of tangent plane is

$$-f_x(x - x_0) - f_y(y - y_0) + z - z_0 = 0$$

which is the same as $z = L(x, y)$.

Definition 5.7. Let $L(x, y)$ be the linear approximation of $f(x, y)$ at (x_0, y_0). The function $f(x, y)$ is **differentiable** at (x_0, y_0) if the following holds.

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|f(x, y) - L(x, y)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0 \quad (5.7.1)$$

Let us emphasize that the differentiability assumes the existence of linear approximation. That is, a function $f(x, y)$ must have partial derivatives f_x, f_y at (x_0, y_0) in order to be differentiable. Intuitively, the differentiability of $f(x, y)$ implies that the value of function is well-approximated by the linear function $L(x, y)$ near (x_0, y_0). It means that the graph of a differentiable function is *smooth* near (x_0, y_0) so that it almost looks like a plane. But what does it looks like if it is not differentiable?

Exercise 5.8. The existence of the linear approximation $L(x, y)$ does not imply that $f(x, y)$ is differentiable. Find an example of a function $f(x, y)$ which has the linear approximation, but not differentiable at $(0, 0)$.

Exercise 5.9. Let us show that the function

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

is not differentiable at $(0, 0)$. It is the same function in Example 5.3 and it is continuous at $(0, 0)$ as in Figure 5.3.1. The partial derivatives at $(0, 0)$ are

$$f_x(0, 0) = \lim_{t\to 0} \frac{f(t, 0) - f(0, 0)}{t - 0} = \lim_{t\to 0} frac0t = 0$$

$$f_y(0, 0) = \lim_{t\to 0} \frac{f(0, t) - f(0, 0)}{t - 0} = \lim_{t\to 0} frac0t = 0$$

Thus the linear approximation is $L(x, y) = 0$. However, the limit

$$\lim_{(x,y)\to(0,0)} \frac{|xy|/\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)} \frac{|xy|}{x^2 + y^2}$$

does not exists. (Exercise 1).

Example above tells us that if the graph of a function $f(x, y)$ has a tip or a sharp point, then it is not differentiable at that point. The following theorem rigorously explains the intuitive meaning of differentiability.

Theorem 5.10. *Suppose that there are two function $\epsilon_1 = \epsilon_1(x, y)$, $\epsilon_2 = \epsilon_2(x, y)$ satisfying*

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= f_x(x_0, y_0)(x - x_0) \\ &\quad + f_y(x_0, y_0)(y - y_0) + \epsilon_1(x - x_0) + \epsilon_2(y - y_0). \end{aligned}$$

If $\epsilon_1, \epsilon_2 \to 0$ as $(x, y) \to (x_0, y_0)$, then $f(x, y)$ is differentiable at (x_0, y_0).

Proof. By assumption,

$$f(x, y) - L(x, y) = \epsilon_1(x - x_0) + \epsilon_2(y - y_0)$$

Thus

$$\begin{aligned} &\lim_{(x,y)\to(x_0,y_0)} \frac{|f(x, y) - L(x, y)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &= \lim_{(x,y)\to(x_0,y_0)} \frac{\epsilon_1(x - x_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &\quad + \lim_{(x,y)\to(x_0,y_0)} \frac{\epsilon_2(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \end{aligned} \quad (5.10.1)$$

Since

$$\left| \frac{(x - x_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \right| < 1,$$

$$\left| \frac{(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \right| < 1,$$

both limits in (5.10.1) becomes 0. □

Theorem 5.10 gives a geometric meaning of differentiability. The function $f(x, y)$ is differentiable at (x_0, y_0) if the linear approximation $L(x, y)$ is a 'good' approximation, meaning that the value of $L(x, y)$ gets close to $f(x_0, y_0)$ 'faster' then the point (x, y) gets close to (x_0, y_0).

Now, let us translate (5.7.1) in terms of vectors. Let $\mathbf{x} = (x, y)$ and $\mathbf{x}_0 = (x_0, y_0)$. Then the numerator of (5.10.1) is

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = \|\mathbf{x} - \mathbf{x}_0\|,$$

and its denominator is

$$f(x, y) - L(x, y) = f(\mathbf{x}) - f(\mathbf{x}_0) - (f_x, f_y) \circ (\mathbf{x} - \mathbf{x}_0).$$

Let us write the inner product in the last equality by matrix multiplication:

$$(f_x, f_y) \circ (\mathbf{x} - \mathbf{x}_0) = \begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

where $\mathbf{x}$ and $\mathbf{x}_0$ are considered as 2×1 matrices. Then (5.7.1) becomes

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}}{\|\mathbf{x} - \mathbf{x}_0\|} = 0 \tag{5.10.2}$$

At this point, we must recall that we defined the differentiability of single variable function in a similar way: a derivative $f'(x_0)$ is given by

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

By taking $f'(x_0)$ to the left-hand side, equation becomes

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0 \tag{5.10.3}$$

It is obvious that (5.10.2) is similar to (5.10.3). Moreover, the matrix $\begin{bmatrix} f_x & f_y \end{bmatrix}$ corresponds to the derivative $f'(x_0)$ in (5.10.3). It teaches us how to understand the derivative $f'(x_0, y_0)$ of $f(x, y)$. We will explain more detail in §5.4.

5.3 The chain rule

Theorem 5.11 (The chain rule)**.** *Let $f(x,y)$ be a two variable function and and $c(t) = (x(t), y(t))$ be a parametrization of a curve in $\mathbf{R}^2$ which is differentiable at t_0. If $f(x,y)$ is differentiable at $(x_0, y_0) = (x(t_0), y(t_0))$, then the composite*

$$F(t) = (f \circ c)(t) = f(x(t), y(t))$$

is differentiable at t_0 and its differential is given by

$$F'(t_0) = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0) \tag{5.11.1}$$

Proof. Let us define $g_1(t), g_2(t)$ as follows.

$$g_1(t) = \frac{x(t) - x_0 - x'(t_0)(t - t_0)}{t - t_0}$$

$$g_2(t) = \frac{y(t) - y_0 - y'(t_0)(t - t_0)}{t - t_0}$$

Since we assumed that $x(t)$, $y(t)$ are differentiable at t_0, we have

$$\lim_{t \to t_0} g_1(t) = \lim_{t \to t_0} g_2(t) = 0$$

and thus $g_1(t)$, $g_2(t)$ are continuous on $\mathbf{R}$. Next, let us define $F(x,y)$ as follows.

$$F(x,y) = \frac{f(x,y) - L(x,y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}$$

Since we assumed that $f(x,y)$ is differentiable at (x_0, y_0), we have

$$\lim_{(x,y) \to (x_0, y_0)} F(x,y) = 0$$

and $F(x,y)$ is continuous on $\mathbf{R}^2$. Note that

$$x(t) - x_0 = (x'(t_0) + g_1(t))(t - t_0),$$

$$y(t) - y_0 = (y'(t_0) + g_2(t))(t - t_0),$$

$$f(x,y) - f(x_0, y_0) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + F(x,y)\sqrt{(x - x_0)^2 + (y - y_0)^2}.$$

Thus

$$\begin{aligned}
&\frac{f(x(t), y(t)) - f(x_0, y_0)}{t - t_0} \\
&\quad = \frac{f_x(x_0, y_0)(x(t) - x_0) + f_y(x_0, y_0)(y(t) - y_0) + F(x(t), y(t))\sqrt{(x(t) - x_0)^2 + (y(t) - y_0)^2}}{t - t_0} \\
&\quad = f_x(x_0, y_0)(x'(t_0) + g_1(t)) + f_y(x_0, y_0)(y'(t_0) + g_2(t)) \\
&\qquad\qquad + F(x(t), y(t))\sqrt{\left(\frac{x(t) - x_0}{t - t_0}\right)^2 + \left(\frac{y(t) - y_0}{t - t_0}\right)^2} \\
&\quad = f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0) \\
&\qquad\qquad + f_x(x_0, y_0)\underbrace{g_1(t)}_{0} + f_y(x_0, y_0)\underbrace{g_2(t)}_{0} + \underbrace{F(x(t), y(t))}_{0}\sqrt{\underbrace{\left(\frac{x(t) - x_0}{t - t_0}\right)^2}_{x'(t_0)} + \underbrace{\left(\frac{y(t) - y_0}{t - t_0}\right)^2}_{y'(t_0)}}
\end{aligned}$$

As $t \to 0$, each term goes to the vales in underbraces. Thus $f(x(t), y(t))$ is differentiable at $t = t_0$ and its differential is $f_x(x_0, y_0)x'(t_0) + f_y(x_0, y_0)y'(t_0)$. □

Definition 5.12. Let $f(x,y)$ be a function defined on $\mathbf{R}^2$. The vector

$$\nabla f(x_0,y_0) = (f_x(x_0,y_0), f_y(x_0,y_0))$$

is called the **gradient** of $f(x,y)$ at (x_0,y_0).

We can write the chain rule (5.11.1) using the gradient of $f(x,y)$.

$$(f\circ c)'(t_0) = \nabla f(c(t_0))\cdot c'(t_0) \tag{5.12.1}$$

The rate of change of $f(x,y)$ at (x_0,y_0) to the direction of $c'(t_0)$ given by the inner product of gradient of $f(x,y)$ at $(x_0,y_0$ and $c'(t_0)$. That is, the gradient measures how much $f(x,y)$ changes to each direction. More precisely, the gradient $\nabla f(x_0,y_0)$ is a linear map

$$\nabla f(x_0,y_0) : \mathbf{R}^2 \to \mathbf{R}$$

such that

$$\nabla f(x_0,y_0)(\mathbf{v}+\mathbf{w}) = \nabla f(x_0,y_0)\cdot\mathbf{v} + \nabla f(x_0,y_0)\cdot\mathbf{w}$$

for all $\mathbf{v}$, $\mathbf{w}$ in $\mathbf{R}^2$.

Theorem 5.13. *Let $\mathbf{X}:\mathbf{R}^2\to\mathbf{R}^2$ be a function defined by*

$$\mathbf{X}(u,v) = (x(u,v), y(u,v)).$$

Let $x_0 = x(u_0,v_0)$, $y_0 = y(u_0,v_0)$ and assume that $x(u,v)$, $y(u,v)$ are differentiable at (u_0,v_0). Let $f:\mathbf{R}^2\to\mathbf{R}$ be a function which is differentiable at (x_0,y_0). Then the composite function

$$F(u,v) = f(x(u,v), y(u,v))$$

is differentiable at (u_0,v_0) and

$$\begin{aligned} F_u(u_0,v_0) &= f_x(x_0,y_0)x_u(u_0,v_0) + f_y(x_0,y_0)y_u(u_0,v_0),\\ F_v(u_0,v_0) &= f_x(x_0,y_0)x_v(u_0,v_0) + f_y(x_0,y_0)y_v(u_0,v_0) \end{aligned} \tag{5.13.1}$$

Proof. Note that

$$F_u(u_0,v_0) = \frac{d}{du}\Big|_{u=u_0} f(x(u,v_0), y(u,v_0)),$$

$$F_v(u_0,v_0) = \frac{d}{dv}\Big|_{v=v_0} f(x(u_0,v), y(u_0,v)).$$

Thus (5.13.1) follows from (5.11.1). □

We can write equation (5.13.1) in a matrix form:

$$\begin{bmatrix} F_u & F_u \end{bmatrix} = \begin{bmatrix} f_x & f_y \end{bmatrix}\cdot\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}. \tag{5.13.2}$$

Note that $F = f\circ\mathbf{X}$. We can rewrite (5.13.2) as

$$\nabla(f\circ X) = \nabla f\cdot\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}$$

As we change the coordinates of f via the map $\mathbf{X}$, the differential in terms of the new coordinate changes by the multiplication of the matrix $\begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}$.

Example 5.14. The function
$$T(r,\theta) = (r\cos\theta, r\sin\theta)$$
maps a polar coordinate (r,θ) to the corresponding cartesian coordinate. Let us express a function $f(x,y)$ in terms of polar coordinate: $\bar{f} = f \circ T(r,\theta)$. Then the partial derivatives in polar coordinates are

$$\bar{f}_r = f_x x_r + f_y y_r = f_x \cos\theta + f_y \sin\theta$$

$$\bar{f}_\theta = f_x x_\theta + f_y y_\theta = -r f_x \sin\theta + r f_y \cos\theta$$

In matrix form,
$$\begin{bmatrix} \bar{f}_r & \bar{f}_\theta \end{bmatrix} = \begin{bmatrix} f_x & f_y \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

The values of $\bar{f}$ and f are the same, but expressed in different coordinate systems.

5.4 Differentials

Definition 5.15. Let $f(x_1, x_2, \cdots, x_n)$ be a multivariable function defined on $\mathbf{R}^n$. The **gradient** of f at the point $\mathbf{a} = (a_1, \cdots, a_n)$ is a n-dimensional vector

$$\nabla f(a_1, \cdots, a_n) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \cdots, \frac{\partial f}{\partial x_1}(\mathbf{a})\right)$$

Definition 5.16. A multivariable function $f : \mathbf{R}^n \to \mathbf{R}$ is **differentiable** at $\mathbf{a} = (a_1, \cdots, a_n)$ if the gradient $\nabla f(\mathbf{a})$ satisfies

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|} = 0 \quad (5.16.1)$$

where the vectors $\mathbf{x}$, $\mathbf{a}$ are considered as $n \times 1$ matrices.

We can consider a multivariable function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ as a m-tuple of real-valued function $f_1, \cdots, f_m$ with n variables:

$$\mathbf{f}(x_1, \cdots, x_n) = (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))$$

The differentiability of $\mathbf{f}$ is defined from the differentiability of each f_i.

Definition 5.17. A multivariable function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ is **differentiable** at $\mathbf{a} = (x_1, \cdots, x_n)$ if the matrix

$$\mathbf{df}(\mathbf{a}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{bmatrix}$$

satisfies

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|f(\mathbf{x}) - \mathbf{df}(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0$$

. where the vectors $\mathbf{x}$, $\mathbf{a}$ are considered as $n \times 1$ matrices. The matrix $\mathbf{df}(\mathbf{a})$ is called the **differential** of $\mathbf{f}$ at $\mathbf{a}$. We often denote $\mathbf{f}'$ for $\mathbf{df}$.

Note the differential $df(\mathbf{a})$ of the real-valued function $f : \mathbf{R}^n \to \mathbf{R}$ is a $1 \times n$ matrix:

$$df(\mathbf{a}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{a}) \end{bmatrix}.$$

Thus, the dot product in (5.16.1) can be rewritten as a matrix multiplication:

$$\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}$$

For this reason, the gradient $\nabla f(\mathbf{a})$ can be identified with the differential $df(\mathbf{a})$, and we often write the differential $\mathbf{df}(\mathbf{a})$, $df(\mathbf{a})$ as $\nabla \mathbf{f}(\mathbf{a})$ or $\nabla f(\mathbf{a})$. Let us also note that each row of the derivative $\mathbf{df}$ is the gradient $\nabla f_i(\mathbf{a})$:

$$\mathbf{df}(\mathbf{a}) = \begin{bmatrix} \nabla f_1(\mathbf{a}) \\ \vdots \\ \nabla f_m(\mathbf{a}) \end{bmatrix}$$

Definition 5.18. In general, let S be a subset of $\mathbf{R}^n$ and $\mathbf{a}$ be a point on S. The **tangent space** $T_{\mathbf{a}}S$ is the vector space consists of all *tangent* vectors of S at $\mathbf{a}$.

Note that the tangent space $T_{\mathbf{a}}\mathbf{R}^n$ is a n-dimensional vector space $\mathbf{R}^n$. The derivative $\mathbf{df}(\mathbf{a})$ is a linear map $\mathbf{df}(\mathbf{a}) : \mathbf{R}^n \to \mathbf{R}^m$: for each n-dimensional vector $\mathbf{v} = (v_1, \cdots, v_n)$, $\mathbf{df}(\mathbf{a})\mathbf{v}$ is a m-dimensional vector given by

$$d\mathbf{f}(\mathbf{a})\mathbf{v} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

In the single-variable Calculus, we identify derivative $f'(a)$ as a slope of the graph of $y = f(x)$ at $x = a$. We also emphasized that $f'(a)$ is a rate of change of value of $f(x)$ at $x = a$. From this observation, the derivative $df(\mathbf{a})\mathbf{v}$ gives the rate of change of $f(\mathbf{x})$ at $\mathbf{x} = \mathbf{a}$ with respect to the *direction of* $\mathbf{v}$. (Then what does the differential $\mathbf{df}(\mathbf{a})$ represent?)

Example 5.19. Let $f(x, y, z) = x + 2y + 3z$ and S be the graph of $z = xy$. The differential $df(p)$ at the point $p = (1, 1, 1)$ on S is

$$df(1, 1, 1) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}.$$

The tangent plane of S at p is

$$x + y - z = 1$$

and the tangent space T_pS is spanned by

$$\mathbf{v}_1 = (1, 0, 1), \quad \mathbf{v}_2 = (0, 1, 1).$$

The parametric curve $c(t) = (t, 1, t)$ lies on S, and satisfies $c(1) = p$ and $c'(1) = \mathbf{v}_1$. By the chain rule (cf. Theorem 5.11),

$$(f \circ c)'(1) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 4$$

This means that the rate of change of f along a curve passing through p in direction of $\mathbf{v}_1$ is 4. Likewise, the rate of change in direction of $\mathbf{v}_2$ is 5. For any tangent vector $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$, the rate of change of f in direction of $\mathbf{v}$ is

$$df(p) \cdot \mathbf{v} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \left(a \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right) = 4a + 5b.$$

Theorem 5.20. *The function* $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ *defined by*

$$\mathbf{f}(x_1, \cdots, x_n) = (y_1(x_1, \cdots, x_n), \cdots, y_m(x_1, \cdots, x_n))$$

is differentiable if and only if each function

$$y_j = y_j(x_1, x_2, \cdots, x_n), \quad j = 1, \cdots, m$$

is differentiable.

Theorem 5.21 (Chain rule)**.** *Suppose that* $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ *is differentiable at* $\mathbf{a} \in \mathbf{R}^n$ *and* $\mathbf{g} : \mathbf{R}^m \to \mathbf{R}^l$ *is differentiable at* $\mathbf{f}(\mathbf{a}) \in \mathbf{R}^m$*. Then* $\mathbf{g} \circ \mathbf{f}$ *is differentiable at* $\mathbf{a}$ *and*

$$\mathbf{d}(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) = \mathbf{dg}(\mathbf{f}(\mathbf{a}))\mathbf{df}(\mathbf{a}) \tag{5.21.1}$$

Note that differential $d(\mathbf{g} \circ \mathbf{f})(\mathbf{a})$ is $l \times n$ matrix. Thus (5.21.1) is simply a matrix multiplication. It states that the differential of a composition is a *multiplication* of each differentials.

5.5 Exercise

1. Show that the limit
$$\lim_{(x,y)\to(0,0)} \frac{|xy|}{x^2+y^2}$$
does not exist.

2. Show that the function
$$f(x,y) = \begin{cases} \dfrac{y^2}{|x-y|} & x \neq y \\ 0 & x = y \end{cases}$$
is discontinuous at $(0,0)$.

3. Let us number the following statements.

   1. The partial derivatives f_x, f_y are continuous at (x_0, y_0).
   2. The function f is differentiable at (x_0, y_0).
   3. The directional derivative $D_{\mathbf{u}} f(x_0, y_0)$ exists for every $\mathbf{u}$.
   4. The function f is continuous at (x_0, y_0).

   (a) Show that the following implications hold.

$$\begin{tikzcd} & & 3 \\ 1 \arrow[r] & 2 \arrow[ur] \arrow[dr] & \\ & & 4 \end{tikzcd}$$

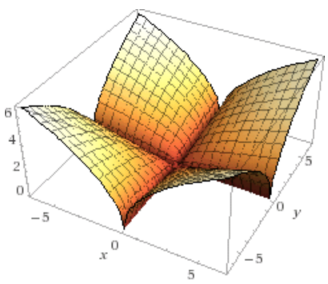
   (b) Find the counter-examples for

   i. $2 \not\to 1$

   ii. $3 \not\to 2$

   iii. $4 \not\to 2$

4. The figure below is the graph of $f(x,y) = \sqrt{|xy|}$. Find the equation of a tangent plane of the graph $z = f(x,y)$ at $(1,1,1)$. (Can you find the tangent plane at $(0,0,0)$?)



5. Let $f(x,y) = (x^2 - y^2, 2xy)$.

   (a) Draw the image of $D = [1,2] \times [1,2]$.

   (b) Find the differential $f'(1,1)$.

   (c) If $D = [1, 1+\epsilon] \times [1, \epsilon]$ and $\epsilon \to 0$, compare the limit
$$\lim_{\epsilon\to 0} \frac{\operatorname{area} f(D)}{\operatorname{area}(D)}$$
   and the differential $f'(1,1)$.

6. The *Laplacian* Δ, also denoted by ∇^2, is an operator defined by
$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$
By using the chain rule (cf. Theorem 5.13), show that Δ can be rewritten as
$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

6 Change of coordinates

In this section, we introduce a technique of multiple integration by change of coordinates. In §6.1, we introduce the Jacobian, which is the most important concept in the change of coordinate. The Jacobian measures the ratio between the areas of domain and target domain. In §**??**, we introduce important coordinate systems such as polar, spherical, and cylinderical systems. Finally in §6.2, we introduce techniques of change of coordinates in integration.

- What is a Jacobian? What does it represent?
- How we use Jacobian in change of coordinates in integration?

6.1 Jacobians

Definition 6.1. Let D be a connected subset in $\mathbf{R}^2$, and $\mathbf{X} : D \to \mathbf{R}^2$ be an *injective* differentiable function defined by

$$X(u,v) = (x(u,v), y(u,v)), \quad (u,v) \in D.$$

Then the value determinant

$$J = \left| \det \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix} \right| = |x_u y_v - x_v y_u|$$

is called the **Jacobian** of X.

Note that the Jacobian is a real-valued function defined on D, and is the absolute value of the determinant of the differential $d\mathbf{X}$.

Example 6.2. Let $X : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear function defined by

$$X(u,v) = (au + bv, cu + dv), \quad (u,v) \in \mathbf{R}^2$$

where $ad - bc \neq 0$. Then the Jacobian is a constant.

$$J = |ad - bc|.$$

The map X is one-to-one correspondence. It sends the unit square $[0,1] \times [0,1]$ onto the parallelogram with vertices at $(0,0)$, (a,c), (b,d), and $(a+b, c+d)$. Note that the area of parallelogram is the same as the Jacobian $J = |ad - bc|$. If we take a rectagular domain D, then the area of the image $\mathbf{X}(D)$ is J times greater than the area of D.

Example 6.3. Let $X : (\mathbf{R}_+)^2 \to \mathbf{R}^2$ be a function defined by

$$X(u,v) = (u^2 - v^2, uv), \quad u, v > 0. \tag{6.3.1}$$

The $\mathbf{R}_+$ represents the set of all positive real numbers, thus $(\mathbf{R}_+)^2$ is the first quadrant of $\mathbf{R}^2$. The rectangular area $D = [a,b] \times [c,d]$ on $(\mathbf{R}_+)^2$ ($0 < a < b$, $0 < c < d$) is mapped to a region bounded by

$$y^2 = -a^2x + a^4, \quad y^2 = -b^2x + b^4, \quad y^2 = c^2x + c^4, \quad y^2 = d^2x + d^4$$

For example, Figure 6.1 shows the boundaries of the D and $\mathbf{X}(D)$ for $a = c = 1$, $c = d = 2$. Note that the area of the image is larger than the area of the domain.

At $(1,1)$, the Jacobian is

$$J = \left| \det \begin{bmatrix} 2u & -2v \\ v & u \end{bmatrix}_{(1,1)} \right| = \left| \det \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix} \right| = 4.$$

If we decrease the area of the domain, then the area of the image is approximately 4 times larger than the area of the domain. (cf. Figure 6.2) To verify, let

$$D = [u_0, u_0 + \Delta u] \times [v_0, v_0 + \Delta v]$$

be a domain in $(\mathbf{R}_+)^2$. The vectors

$$\mathbf{X}_u(u_0, v_0) = (2u_0, v_0), \quad \mathbf{X}_v(u_0, v_0) = (-2v_0, u_0)$$

are tangent to the boundary curves of the image $\mathbf{X}(D)$ at the point $\mathbf{X}(u_0, v_0)$. The area of the image $\mathbf{X}(D)$ is approximated by the area of the parallelogram whose sides are bounded by vectors $\mathbf{X}_u(u_0, v_0)\Delta u$ and $\mathbf{X}_v(u_0, v_0)\Delta v$. Thus

$$\begin{aligned} \frac{\text{area}\mathbf{X}(D)}{\text{area}(D)} &\approx \frac{\|\mathbf{X}_u(u_0,v_0)\Delta u \times \mathbf{X}_v(u_0,v_0)\Delta v\|}{\Delta u \Delta v} \\ &= \|\mathbf{X}_u(u_0,v_0) \times \mathbf{X}_v(u_0,v_0)\| = J(u_0,v_0) \end{aligned}$$

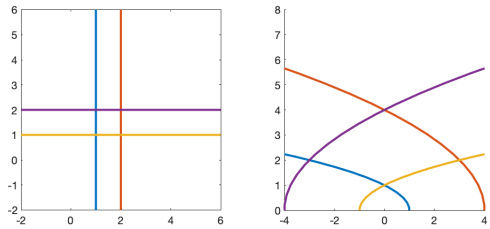


Figure 6.1: Image of (6.3.1) on $[1,2] \times [1,2]$

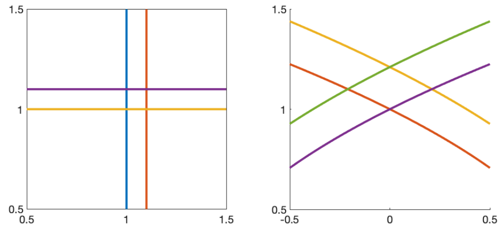


Figure 6.2: Image of (6.3.1) on $[1,1.1] \times [1,1.1]$

Figure 6.2 give closer view of Figure 6.1 near $(1,1)$.

Since $\mathbf{X}$ is one-to-one on $(\mathbf{R}_+)^2$, we can consider its inverse $\mathbf{X}^{-1} : \mathbf{R} \times \mathbf{R}_+ \to \mathbf{R}^2$. Let us write

$$\mathbf{X}^{-1}(x,y) = (u(x,y), v(x,y)), \quad y > 0 \tag{6.3.2}$$

Although we can write $u(x,y)$, $v(x,y)$ explicitly (cf. Exercise 1), let us compute the Jacobian of $\mathbf{X}^{-1}$ directly by using the chain rule. Since $\mathbf{X} \circ \mathbf{X}^{-1}$ is a identity function $(x,y) \mapsto (x,y)$, by the chain rule (cf. Theorem 5.21),

$$d\mathbf{X} \cdot d(\mathbf{X}^{-1}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{6.3.3}$$

The $\cdot$ on the left-hand side is the matrix multiplication, because both $d\mathbf{X}$ and $d\mathbf{X}^{-1}$ are 2×2 matrices.

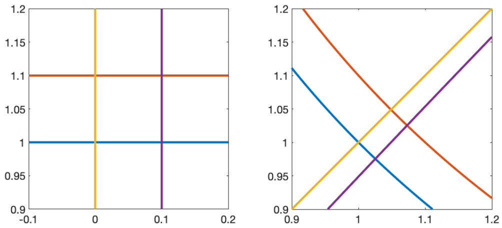


Figure 6.3: Image of the inverse (6.3.2) on $[0, .1] \times [1, 1.1]$

Thus $d(\mathbf{X}^{-1})$ is the inverse matrix of $d\mathbf{X}$:

$$d(\mathbf{X}^{-1}) = \begin{bmatrix} 2u & -2v \\ v & u \end{bmatrix}^{-1} = \frac{1}{2u^2 + 2v^2} \begin{bmatrix} u & 2v \\ -v & 2u \end{bmatrix}$$

Note that the coordinates are given by u, v which are the dependent variables of $\mathbf{X}^{-1}$. Since $\mathbf{X}^{-1}$ is one-to-one, we can find a unique point (x, y) corresponds to (u, v) by $\mathbf{X}^{-1}$. For example, $(0, 1)$ on xy-plane corresponds to the point $(1, 1)$ on the uv-plane. Thus the Jacobian of $\mathbf{X}^{-1}$ at $(x, y) = (0, 1)$ is

$$\begin{aligned} |d(\mathbf{X}^{-1})(0,1)| &= \frac{1}{(2u^2 + 2v^2)^2} \left| \det \begin{bmatrix} u & 2v \\ -v & 2u \end{bmatrix} \right|_{(u,v)=(1,1)} \\ &= \frac{1}{16} \left| \det \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} \right| = \frac{1}{4} \end{aligned}$$

Figure 6.3 shows the closer view at $(0, 1)$ on xy-plane. We can find that the area of the target domain is reduced approximately $1/4$ of the original domain.

6.2 Integration by substitution

Definition 6.4. Let D be a region in $\mathbf{R}^2$ and $T: \mathbf{R}^2 \to \mathbf{R}^2$ be an injective function. If T is a one-to-one correspondence onto its image, we call T as a **transformation**.

Example 6.5. Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be a function defined by

$$T(x, y) = (x - y, x + y)$$

For the domain $D = [0, 1] \times [0, 1]$, the image $T(D)$ is a parallelogram. To observe this, let us identify the image of the boundaries of D. For example, the left-side boundary curve

$$c(t) = (0, t), \quad t \in [0, 1]$$

is mapped to

$$T \circ c(t) = (-t, t), \quad t \in [0, 1]$$

which is a line segment from $(0, 0)$ to $(-1, 1)$. Likewise, we can show that each boundary of D is mapped to a line segment, which consists a parallelogram with vertices $(0, 0)$, $(1, 1)$, $(0, 2)$, and $(-1, 1)$. Note that the Jacobian

$$J = \det \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = 2$$

is the ratio of two areas D and $f(D)$.
Also note that the function T is one-to-one on the entire region $\mathbf{R}^2$. In general, if $u(x, y), v(x, y)$ are linear functions (i.e. polynomial of degree 1) which are not linearly independent, then the map

$$T(x, y) = (u(x, y), v(x, y))$$

is one-to-one on $\mathbf{R}^2$ (Exercise 2). Moreover, the Jacobian J of T is always constant. It means that the linear transformation preserves the ratio of the areas in the domain and its image.
The Jacobian may be negative or even zero. (What would it imply that the Jacobian is zero? (cf. Exercise **??** For example, the linear function

$$T_1(x, y) = (-x + y, x + y)$$

has a negative Jacobian -2. This means that the orientation of the image is reversed. The vertices $(0, 0)$, $(1, 0)$, $(1, 1)$, $(0, 1)$ in D are mapped to $(0, 0)$, $(-1, 1)$, $(0, 2)$, $(1, 1)$ in $f(D)$. Although the region $T_1(D)$ and $T(D)$ are the same parallelogram, the orientation of corresponding vertices are reversed.

Theorem 6.6 (Integration by substitution). *Let $T: D \to T(D)$*

$$T(u, v) = (x(u, v), y(u, v))$$

be a transformation, and $f(x, y)$ be a continuous function defined on $T(D)$. Then

$$\iint_{T(D)} f(x, y)dxdy = \iint_D f \circ T(u, v)|J_T|dudv \tag{6.6.1}$$

where

$$J_T = \det \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}$$

is the Jacobian of T. (Note that we must multiply the <u>absolute</u> value of J.)

Example 6.7. Let D_1 be the region bounded by $x - y = 1$, $x - y = 2$, $x + y = 2$, $x + y = 4$. Let us compute

$$\iint_{D_1} (x + y)e^{x^2 - y^2} dxdy$$

Let $u(x, y) = x - y$ and $v(x, y) = x + y$. Then the integrand can be written as

$$(x + y)e^{x^2 - y^2} = ve^{uv}$$

which we can compute the integral. Thus it is easy to change the coordinate xy to uv. We emphasize that the transformation in (6.10.1) maps (u, v) to (x, y). (*Not* (x, y) to (u, v)!) Since $x = (u + v)/2$ and $y = (v - u)/2$, we need

$$T(u, v) = \left(\frac{u}{2} + \frac{v}{2}, \frac{-u}{2} + \frac{v}{2}\right)$$

whose Jacobian is

$$J = \det \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \frac{1}{2}$$

Since $D_1 = T(D)$ where $D = \{1 \le u \le 2, 2 \le v \le 4\}$,

$$\begin{aligned} &\iint_{D_1} (x + y)e^{x^2 - y^2} dxdy \\ &= \int_2^4 \int_1^2 ve^{uv} \frac{1}{2} dudv \\ &= \frac{1}{2} \int_2^4 e^{uv}\Big|_{u=1}^2 dv \\ &= \frac{1}{2} \int_2^4 e^{2v} - e^v dv \\ &= \frac{1}{2}\left(e^8 - 2e^4 + e^2\right) \end{aligned}$$

We could compute the Jacobian in the following way: Since the inverse

$$T^{-1}(x, y) = (u(x, y), v(x, y)) = (x - y, x + y).$$

is also a transformation, the Jabobian J is the reciprocal of the Jacobian of T^{-1}. (cf. (6.3.3))

$$J = \frac{1}{J_{T^{-1}}} = \frac{1}{\det \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} = \frac{1}{2}.$$

Example 6.8. The function $T : R \times [0, 2\pi] \to \mathbf{R}^2$ defined by

$$T(r, \theta) = (r\cos\theta, r\sin\theta) \tag{6.8.1}$$

is called the **polar coordinate** of $\mathbf{R}^2$. The Jacobian is

$$J = \det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r.$$

Although T is not one-to-one when $r = 0,\ 2\pi$, we can consider T as a transformation because the areas of the lines $r = 0$, $r = 2\pi$ are zero and they do not affect the value of the double integral. Let $D = \{(x, y) \,|\, x^2 + y^2 \leq 1\}$ be the unit disk. Let us compute

$$\iint_D |x|e^{-x^2-y^2}dxdy$$

By using Theorem 6.6, we have

$$\begin{aligned}\iint_0^{2\pi}\int_0^1 r|\cos\theta|e^{-r^2}rdrd\theta &= 4\int_0^1 re^{-r^2}dr \\ &= 2\left(1 - \frac{1}{e}\right).\end{aligned}$$

Remark 6.9. We must be careful in choosing variables in method of change of variables. The notation $J_T = \dfrac{\partial(u, v)}{\partial(x, y)}$ is useful to remember the subsitution rule:

$$\iint_D f(u, v)dudv = \iint_{T(D)} f(u, v)\left|\frac{\partial(u, v)}{\cancel{\partial(x, y)}}\right| \cancel{dxdy}$$

It is useful to know that the Jacobian of the inverse $T^{-1}(u, v) = (x(u, v), y(u, v))$ is

$$J_{T^{-1}} = \left|\frac{\partial(x, y)}{\partial(u, v)}\right| = \frac{1}{J_T} = \frac{1}{\left|\frac{\partial(u,v)}{\partial(x,y)}\right|}$$

Theorem 6.10 (Integration by substitution)**.** *Let V be a region in $\mathbf{R}^3$ and $T : V \to T(V)$*

$$T(u, v, w) = (x(u, v, w), y(u, vv), z(u, v, w))$$

*be a one-to-one correspondence. (We call it a **transformation**.) Let $f(x, y, z)$ be a continuous function defined on $T(V)$. Then*

$$\begin{aligned}&\iint_{T(V)} f(x, y, z)dxdydz \\ &\quad = \iint_V f \circ T(u, v, w)|J_T|dudvdw \end{aligned} \tag{6.10.1}$$

where

$$J_T = \det\begin{bmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{bmatrix}$$

is the Jacobian of T.

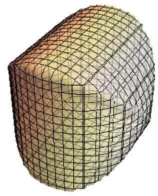


Figure 6.4: Volume between two orthogonal cylinders

Example 6.11. The function $T : R\times[0, 2\pi]\times\mathbf{R} \to \mathbf{R}^3$ defined by

$$T(r, \theta, z) = (r\cos\theta, r\sin\theta, z)$$

is called the **cylinderical coordinate** of $\mathbf{R}^3$. The Jacobian of the cylinderical coordinate is

$$J_T = \det\begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = r$$

Let us find the volume V between two cylinders $x^2+y^2 \leq 1$, $y^2 + z^2 \leq 1$. (cf. Figure 6.4) The projection of V onto xy-plane is the unit disk $x^2 + y^2 \leq 1$. The coordinate z in V ranges from $-\sqrt{1 - y^2}$ to $\sqrt{1 - y^2}$. Thus

$$\begin{aligned}V &= \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\iint_{x^2+y^2\leq 1} dxdydz \\ &= \int_0^{2\pi}\int_0^1\int_{-\sqrt{1-r^2\cos^2\theta}}^{\sqrt{1-r^2\cos^2\theta}} rdzdrd\theta \\ &= \int_0^{2\pi}\int_0^1 2r\sqrt{1-r^2\cos^2\theta}drd\theta \\ &= \int_0^{2\pi}\frac{4}{3}\left(\frac{1}{\cos^2\theta} - \frac{|\sin\theta|^3}{\cos^2\theta}\right)d\theta \\ &= \frac{4}{3}\int_0^{2\pi}|\sin\theta| + \underbrace{\left(\frac{1}{\cos^2\theta} - \frac{|\sin\theta|}{\cos^2\theta}\right)}_{\text{continuous at } \theta=\frac{\pi}{2},\frac{3\pi}{2}}d\theta \\ &= \frac{16}{3}\end{aligned}$$

Using symmetry, we could compute above as follows.

$$\begin{aligned}
V &= 8\int_0^{\pi/2}\int_0^1\int_0^{\sqrt{1-r^2\cos^2\theta}} rdzdrd\theta \\
&= 8\int_0^{\pi/2}\int_0^1 r\sqrt{1-r^2\cos^2\theta}drd\theta \\
&= 8\int_0^{\pi/2}\frac{1}{3}\left(\frac{1}{\cos^2\theta}-\frac{|\sin\theta|^3}{\cos^2\theta}\right)d\theta \\
&= \frac{8}{3}\int_0^{\pi/2}|\sin\theta| + \underbrace{\left(\frac{1}{\cos^2\theta}-\frac{|\sin\theta|}{\cos^2\theta}\right)}_{\text{continuous at }\theta=\frac{\pi}{2}} d\theta \\
&= \frac{16}{3}
\end{aligned}$$

Example 6.12. The **spherical transformation**

$$T(r,\theta,\phi) = (r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi)$$

is also useful trasnformation. (The Jacobian is $J_T = r^2\sin\phi$.) For example, let us compute

$$\iiint_V \sqrt{1-x^2-y^2-z^2}dxdydz$$

where V is the unit ball centered at the origin $(0,0,0)$.

Remark 6.13. As we have seen in Remark 6.9, the Jacobian of the inverse transformation T^{-1} is the reciprocal of Jacobian J_T.

$$J_T = \frac{\partial(u,v,w)}{\partial(x,y,z)} = \frac{1}{J_{T^{-1}}} = \frac{1}{\frac{\partial(x,y,z)}{\partial(u,v,w)}}$$

6.3 Indefinite integrals

Definition 6.14. The integral $\iint_D f(x,y)dxdy$ is called **improper** if it satisfies one of the following.

- the region D is unbounded, or
- the function diverges at some point in D.

Example 6.15. Let $D = \mathbf{R}^2$ be the entire 2-dimensional plane. Let us compute

$$\iint_{\mathbf{R}^2} e^{-x^2-y^2}dxdy$$

By polar coordinate $T(r,\theta) = (r\cos\theta, r\sin\theta)$,

$$T^{-1}(D) = [0,\infty) \times [0, 2\pi].$$

Thus by change of coordinates,

$$\begin{aligned}\iint_{\mathbf{R}^2} e^{-x^2-y^2}dxdy &= \int_0^\infty \int_0^{2\pi} e^{-r^2} r d\theta dr \\ &= 2\pi \left(\frac{-1}{2}e^{-r^2}\right)\Big|_0^\infty \\ &= \pi\end{aligned}$$

Definition 6.16. The **gamma function** $\Gamma : \mathbf{R} \to \mathbf{R}$ is defined by

$$\Gamma(p) = \int_0^\infty x^{p-1}e^{-x}dx$$

Example 6.17. Let us compute $\Gamma\left(\frac{1}{2}\right)$.

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) &= \int_0^\infty x^{-1/2}e^{-x}dx = \int_0^\infty \frac{e^{-x}}{\sqrt{x}}dx \\ &= 2\int_0^\infty e^{-u^2}du \qquad (u = \sqrt{x}) \\ &= \sqrt{\pi}\end{aligned}$$

Definition 6.18. The **beta function** $B(x,y) : \mathbf{R}^2 \to \mathbf{R}$ is defined by

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$$

Proposition 6.19. *The following properties hold.*

1. $B(x,y) = B(y,x)$
2. $B(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$

Proof. By the substitution $s = 1-t$,

$$B(x,y) = \int_1^0 (1-s)^{x-1}s^{y-1}(-ds) = B(y,x).$$

For the second equality, observe that

$$\begin{aligned}B(x,y)\Gamma(x+y) &= \int_0^1 t^{x-1}(1-t)^{y-1}dt \int_0^\infty s^{x+y-1}e^{-x-y}ds \\ &= \int_0^1\int_0^\infty (st)^{x-1}(s(1-t))^{y-1}se^{-x-y}dsdt\end{aligned}$$

Let $u = st$ and $v = s(1-t)$. Then the transformation $T(u,v) = (s,t)$ satisfies

$$T : [0,\infty) \times [0,\infty) \to [0,\infty) \times [0,1],$$

and its Jacobian is

$$J_T = \frac{\partial(s,t)}{\partial(u,v)} = \frac{1}{\begin{vmatrix} t & s \\ 1-t & -s\end{vmatrix}} = \frac{1}{s}. \quad (s > 0)$$

Thus above integral becomes

$$\int_0^\infty \int_0^\infty u^{x-1}v^{y-1}e^{-x-y}dudv = \Gamma(x)\Gamma(y)$$

□

6.4 Exercise

1. Let $\mathbf{X} : (\mathbf{R}_+)^2 \to \mathbf{R}^2$ be a function defined by

$$\mathbf{X}(u, v) = (u^2 - v^2, uv)$$

   (a) Show that $\mathbf{X}$ is one-to-one.

   (b) Show that the target domain of $\mathbf{X}$ is $\mathbf{R} \times \mathbf{R}_+$.

   (c) Find $u(x, y)$, $v(x, y)$ explicitly for the inverse $\mathbf{X}^{-1}$.

$$\mathbf{X}^{-1}(x, y) = (u(x, y), v(x, y))$$

   (d) Compute the Jacobian $J_{\mathbf{X}^{-1}}$ and confirm that it is the *reciprocal* of $J_{\mathbf{X}}$.

2. Show that if $u(x, y), v(x, y)$ are linear functions (i.e. polynomial of degree 1) which are not linearly independent, then the map

$$T(x, y) = (u(x, y), v(x, y))$$

   is a transformation (cf. Definition 6.4).

3. Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be a function with Jacobian J. If $J(p) = 0$ at some $p \in \mathbf{R}^2$, can f be injective on a neighborhood of p?

4. Compute

$$\int_0^1 \int_0^{1-x} e^{y/(x+y)} dy dx$$

5. Let D be a domain bounded by $x^2 - y^2 = 1$, $x^2 - y^2 = 9$, $xy = 2$, and $xy = 4$. Compute

$$\int_D x^2 + y^2 dx dy$$

6. Let $D = \{(x, y) \,|\, x^2 + xy + y^2 \le 1\}$. Compute

$$\int_D e^{-(x^2+xy+y^2)} dx dy$$

7. Using the polar coordinate (cf. Example 6.8), compute the following.

   (a) The area outside $r = 1$ and inside $r = 1 + \cos\theta$.

   (b) The area inside $r^2 = 4\cos^2\theta$.

8. Find the volume between $z = 9 - x^2 - y^2$ and $z = 3x^2 + 3y^2 - 16$.

9. Find the area of the region bounded by three cylinders $x^2 + y^2 \le 1$, $y^2 + z^2 \le 1$, and $x^2 + z^2 \le 1$. (cf. Figure 6.5).

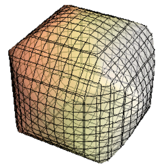


Figure 6.5: The volume between three orthogonal cylinders

10. Find the volume bounded by $x^2 + y^2 \le 1$, $x^2 + z^2 \le 1$, and $z \ge 1/2$.

11. Compute

$$\iiint_V \sqrt{x^2 + y^2 + z^2} e^{x^2+y^2+z^2} dx dy dz$$

   where V is the region between the spheres $x^2 + y^2 + z^2 = a^2$ and $x^2 + y^2 + z^2 = b^2$ $(0 < a < b)$.

12. Prove that

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

7 Applications

7.1 Taylor series

Definition 7.1. Suppose that $f(x,y)$ has continuous second partial derivatives at (x_0,y_0). Then the polynomial

$$\begin{aligned} Q(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) \qquad (7.1.1)\\ &+ f_y(x_0,y_0)(y-y_0) + \frac{1}{2}f_{xx}(x_0,y_0)(x-x_0)^2\\ &+ f_{xy}(x_0,y_0)(x-x_0)(y-y_0)\\ &+ \frac{1}{2}f_{yy}(x_0,y_0)(y-y_0)^2 \end{aligned}$$

is called the **Taylor polynomial** of f of second degree at (x_0,y_0).

Remark 7.2. Let

$$\Delta\mathbf{x} = (\Delta x, \Delta y) = (x-x_0, y-y_0)$$

and denote the *gradient operator* ∇ in vector notation:

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

Let us define the *multiplication* of differential operators as follows:

$$\left(\frac{\partial}{\partial x}\frac{\partial}{\partial x}\right)f = \frac{\partial^2 f}{\partial x^2}, \quad \left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right)f = \frac{\partial^2 f}{\partial x \partial y}$$

By the assumption and Theorem 3.13, $f_{xy}(x_0,y_0) = f_{yx}(x_0,y_0)$, thus (7.1.1) simplifies as

$$Q(x,y) = \sum_{n=0}^{2}\frac{1}{n!}(\Delta\mathbf{x}\cdot\nabla)^n f(x_0,y_0)$$

This generalizes the Taylor polynomial for single variable function:

$$P_k(x) = \sum_{n=0}^{k}\frac{f^{(n)}(x_0)\Delta x^n}{n!}$$

where $x = x_0 + \Delta x$. For single-variable function $\nabla = \frac{\partial}{\partial x}$. Thus $(\Delta x\cdot\nabla)^n f(x_0) = \Delta x^n \nabla^n f(x_0)$.

The k-th order Taylor polynomial for two-variable function is similarly defined. In fact the Taylor polynomial $P_k f$ of any order is well-defined:

$$P_k f(\mathbf{x}) = \sum_{n=0}^{k}\frac{1}{n!}(\Delta\mathbf{x}\cdot\nabla)^n f(\mathbf{x}_0)$$

Theorem 7.3 (Taylor)**.** *Let $f(x,y)$ is a function whose third partial derivatives $f_{xxx}, f_{xyx}, \cdots, f_{yyy}$ are all continuous on a rectangular region*

$$D = \{(x,y)\,|\,|x-x_0|, |y-y_0| \le \epsilon\}$$

Then for each $(x,y)\in D$, there exists a constant $0\le c\le 1$ satisfying

$$f(x,y) = Q(x,y) + R_2(x,y)$$

where

$$R_2(x,y) = \frac{1}{3!}(\Delta\mathbf{x}\cdot\nabla)^3 f(\mathbf{x}_0 + c\Delta\mathbf{x})$$

Remark 7.4. The rectangular region D in Theorem 7.3 is a generalization of the interval $|x-x_0|\le\epsilon$ appears in Taylor theorem for single-variable function.

> Let f be a $\mathcal{C}^{k+1}$-function on an interval $I = (x_0-\epsilon, x_0+\epsilon)$. Then for $x, c\in I$, there exists a constant ξ between x and c such that
>
> $$f(x) = P_k(x) + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-c)^{k+1}$$

Here $x = x_0 + \Delta x$ and $\xi = x_0 + c\Delta x$ for some $0\le c\le 1$.

Example 7.5.

1. We must remember that the choice of c *depends* on the choice of x, x_0. This means that *any* function (with partial derivatives) can be approximated by polynomial-type functions $Q_2(x,y)$, $R_2(x,y)$. If (x,y) is sufficiently close to (x_0,y_0), as degree of Taylor polynomial increases, the the degree of *error term* $R_n(x,y)$ also increases, thus decreases the error $|R_n(x,y)|$ to zero.

2. We can only give an rough answer to the question how big should the degree k should be in order to have a small error.

3. The graph of Taylor polynomial $z = Q(x,y)$ gives rough sketch of the graph of $z = f(x,y)$.

7.2 Maxima and minima

Definition 7.6. Let $f(x, y)$ be a function $f(x, y)$ defined on a region D.

- A point (x_0, y_0) is said to be **local maximal** (**minimal**, respectively) at $\mathbf{x}_0 = (x_0, y_0)$ if there exists a (sufficiently small) $\epsilon > 0$ such that for all $\mathbf{x} = (x, y)$ satisfying $\|\mathbf{x} - \mathbf{x}_0\| < \epsilon$,

$$f(x_0, y_0) \geq f(x, y)$$

($f(x_0, y_0) \leq f(x, y)$, respectively) holds. Such points are called **extremal**.

- A point (x_0, y_0) is called a **critical point** if it satisfies one of the following.

  1. $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$
  2. f_x or f_y does not exist at (x_0, y_0)
  3. f is discontinuous at (x_0, y_0).

  A critical point which is *not* an extremal point is called a **saddle point**.

Example 7.7. Let us find the critical points of

$$f(x, y) = xy - x^2y - xy^2$$

Since $f(x, y)$ is differentiable everywhere, critical points are obtained by solving two equations.

$$f_x = y - 2xy - y^2 = 0, \quad f_y = x - x^2 - 2xy = 0$$

There are four criticial points:

$$(0, 0), (0, 1), (1, 0), \left(\frac{1}{3}, \frac{1}{3}\right).$$

The Figure 7.1 shows the critical points on the graph of $f(x, y)$.

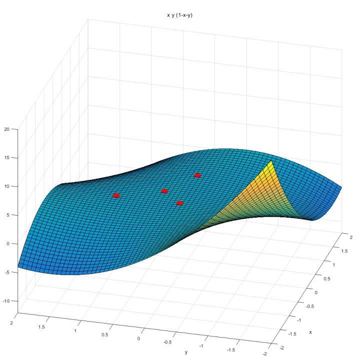
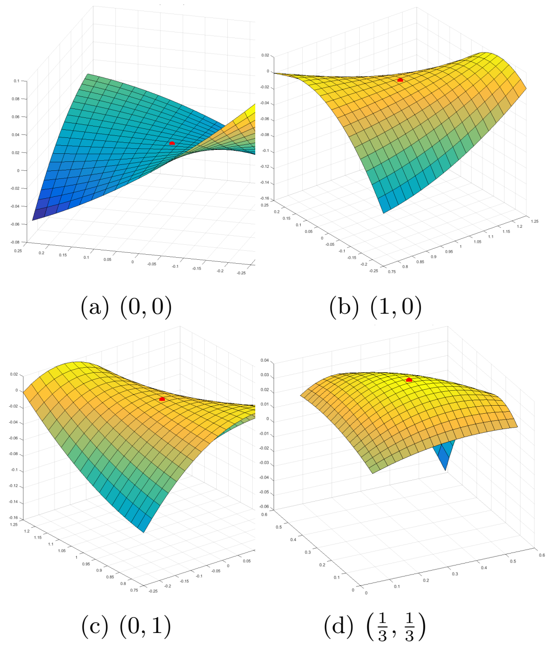


Figure 7.1: The critical points in Example 7.7

Figure 7.7 zooms into each points, thus determines the nature of critical points. Note that $(0, 0), (0, 1), (1, 0)$ are saddle points, while $\left(\frac{1}{3}, \frac{1}{3}\right)$



(a) $(0, 0)$ (b) $(1, 0)$ (c) $(0, 1)$ (d) $\left(\frac{1}{3}, \frac{1}{3}\right)$

Figure 7.2: The nature of critical points in Example 7.7

is local maximum. These nature is not obvious by looking at the global picture of the graph of $f(x, y)$.

However, we could determine the nature of critical point by numeric computation. For example, let us consider $(0, 0)$. The Taylor polynomial of the second degree of $f(x, y)$ at $(0, 0)$ is

$$Q(x, y) = xy$$

. This means the graph of $z = f(x, y)$ is *approximated* by the graph $z = xy$ near $(0, 0)$. The point $(0, 0)$ is a saddle point on $z = xy$, thus it has the same nature on the graph $z = f(x, y)$.

For the point $\left(\frac{1}{3}, \frac{1}{3}\right)$, we have

$$Q(x, y) = -\frac{1}{3}\left(x - \frac{1}{3}\right)^2 - \frac{1}{3}\left(x - \frac{1}{3}\right)\left(y - \frac{1}{3}\right) - \frac{1}{3}\left(y - \frac{1}{3}\right)^2$$

Let us replace $X = x - \frac{1}{3}$, $Y = y - \frac{1}{3}$. Then

$$Q(x, y) = -\frac{1}{3}\left(X + \frac{Y}{2}\right)^2 - \frac{1}{4}Y^2$$

Thus $Q(x, y)$ has local maximum at $X = 0$, $Y = 0$, in other words, $x = \frac{1}{3}$, $y = \frac{1}{3}$.

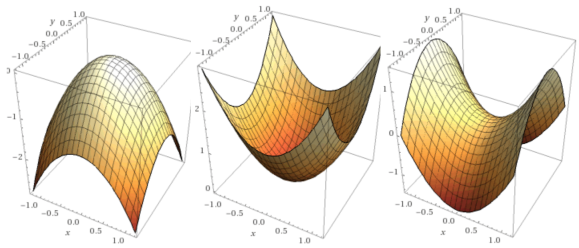
Remark 7.8.

1. Let $f(x, y)$ is differentiable at (x_0, y_0). Suppose that (x_0, y_0) is a critical point $f(x, y)$ satisfying the condition 1 in Definition 7.6. Then

the linear approximation (x_0, y_0) is the plane $z = f(x_0, y_0)$.

2. Further suppose that (x_0, y_0) is a saddle point. Then there exists a curve $c : (-\epsilon, \epsilon) \to \mathbf{R}^2$ with $c(0) = (x_0, y_0)$ such that composition $F(t) = (f \circ c)(t)$ is a inflection point.

3. Graph near critical points of condition 1 is classified into three types below.



(a) Local max (b) Local min (c) Saddle

Figure 7.3: The classification of critical points

Each graph corrsponds to the following functions

$$z = -x^2 - y^2, \quad z = x^2 + y^2, \quad z = x^2 - y^2$$

Definition 7.9. Suppose that all second partial derivatives of a function $f(x, y)$ are continuous on a region R containing (x_0, y_0). Then

$$\Delta_f = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2$$

is called the **discriminant** of f.

Remark 7.10. Note that the quadratic part (terms of degree 2) of $Q(x, y)$ is written as

$$\frac{1}{2} \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \underbrace{\begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}}_{=A} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{7.10.1}$$

where $\Delta x = x - x_0, \Delta y = y - y_0$. Then $\Delta_f = \det A$.

Theorem 7.11 (Hesse)**.** *Let f be a function satisfying conditions in Definition 7.9 and (x_0, y_0) a critical point of f.*

- *If $\Delta > 0$ and $f_{xx}(x_0, y_0) > 0$, then $f(x_0, y_0)$ is a local minimum.*
- *If $\Delta > 0$ and $f_{xx}(x_0, y_0) < 0$, then $f(x_0, y_0)$ is a local maximum.*
- *If $\Delta < 0$, then $f(x_0, y_0)$ is a saddle point.*
- *If $\Delta = 0$, then we cannot determine local extremity by this method.*

Remark 7.12.

1. We will give rough idea why Theorem (7.11) is true. Since Taylor polynomial $Q(x, y)$ approximates $f(x, y)$, the graph of $z = f(x, y)$ is roughly the same as the graph of $Q(x, y)$. The *curvature* of the surface is determined by the quadratic term (7.10.1). Since A is a symmetric matrix, it can be diagonalized[15], and normalized into the following three forms:

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} \pm 1 & 0 \\ 0 & \mp 1 \end{bmatrix}$$

Each matrix corresponds to the local maximum, local minimum, and saddle point respectively.

2. The matrix A in (7.10.1) is called the *Hessian* of f at (x_0, y_0). It is an important object in *Differential Topology*, a branch of mathematics which studies invariants of topological spaces by their differentiable structures.

3. The Hessian is the *Jacobian* (cf. Definition 6.1) of the gradient ∇f.

Example 7.13 (Least square method)**.** Suppose that a set of data is given by

$$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$$

We want to find a line $y = mx + b$ which approximates these data. If the sum of squares of vertical distances between data and the line $y = m_0 x + b_0$ is minimum among all possible lines $y = mx + b$, then we would say that $y = m_0 x + b_0$ best approximates the data. In other words, we want to find m, b such that

$$d(m, b) = \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$

is minimum. The critical point is

$$m_0 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

$$b_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

[15] This fact will be proved in *Linear Algebra*

The Hessian of $d(m, b)$ at (m_0, b_0) is the following.

$$\begin{bmatrix} 2\sum_{i=1}^{n} x_i^2 & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2n \end{bmatrix}$$

The determinant, thus becomes the discriminant of f, always satisfies

$$\Delta_f(m_0, b_0) = 4n\sum_{i=1}^{n} x_i^2 - 4\left(\sum_{i=1}^{n} x_i\right)^2 > 0$$

Since $2\sum_{i=1}^{n} x_i^2 > 0$ is obvious, (m_0, b_0) is a local minimum point, as well as global minimum point.

7.3 Lagrange multiplier

Proposition 7.14. *Let $c = f(x_0, y_0)$ and $L_c(f)$ be the level curve on the xy-plane. (cf. (2.3.1)). Then the gradient vector $\nabla f(x_0, y_0)$ (cf. Example (**??**)) is perpendicular to the tangent line at (x_0, y_0) of the level curve $L_c(f)$.*

Theorem 7.15 (Lagrange multiplier)**.** *Let $g(x, y)$, $f(x, y)$ be differentiable functions. Let (x_0, y_0) be a point on the level set $g(x, y) = c$ where $f(x, y)$ is locally extremal. If $\nabla g(x_0, y_0) \neq \mathbf{0}$, then there exists λ such that*

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0) \tag{7.15.1}$$

Example 7.16. The Lagrange multiplier is a systematic method of finding maxima or minima of a function with a constraints:

- $f(x, y)$ is the target function which we want to maximize or minimize.
- $c = g(x, y)$ is the constraint.
- Find (x_0, y_0) satisfying (7.20.1) for some λ.

For example, let us find the point on the circle $x^2 + y^2 = 10$ at which the function $f(x, y) = 3x + y$ is maximal or minimal.

Corollary 7.17. *The gradient $\nabla f(x_0, y_0)$ is the direction where the value of function $f(x, y)$ changes the fastest from (x_0, y_0).*

Proof. Let $\mathbf{u} = (a, b)$ be the unit vector which represent the direction at (x_0, y_0). In order to maximize or minimize the rate of change $f(x, y)$, we need to maximize or minimize the directional derivative $\nabla_{\mathbf{u}} f(x_0, y_0)$. Note that

$$\nabla_{\mathbf{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathbf{u} = f_x(x_0, y_0)a + f_y(x_0, y_0)b$$

Thus we can apply Lagrange multiplier with the following ingredients:

- The target function:
$$F(a, b) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$$
- The constraints: $G(a, b) = a^2 + b^2 = 1$ ($\mathbf{u}$ must be a unit vector)

With a simple calculation we obtain

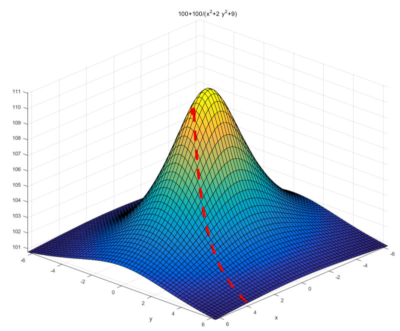
$$\mathbf{u} = (a, b) = \pm(f_x(x_0, y_0), f_y(x_0, y_0)) = \pm\nabla f(x_0, y_0)$$

Depending on the direction, the function changes the fastest in positive or negative directions. □

Exercise 7.18. Let

$$f(x, y) = 100 + \frac{100}{x^2 + 2y^2 + 9}$$

be a function whose graph represent the contour of a mountain. Suppose that a water flow from the point $(1, 0, 110)$ down the valley in a direction whose slope is the *steepest*. Find the trajectory of the water path.



Example 7.19. In economic theory, a *utility function* is a function which measures a usefulness of certain products. If there are multiple products $x_1, x_2, \cdots, x_n$, then the utility function is a multivariable function

$$u(x_1, \cdots, x_n)$$

In order to produce products, constraints always follows due to resources, times, and labor issue. Suppose that the total *cost* of making products $x_1, x_2, \cdots, x_n$ is given by

$$c(x_1, \cdots, x_n) = C$$

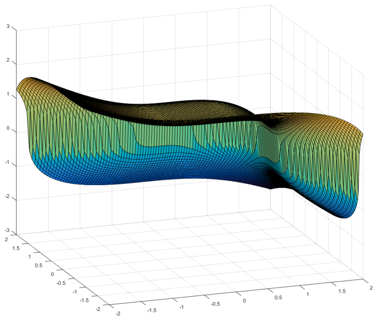
Then we can apply Lagrange multiplier to solve how to maximize the utility.

Example 7.20. Let $g(x, y, z)$, $f(x, y, z)$ be differentiable functions. Suppose that (x_0, y_0, z_0) is a local extremal of $f(x, y, z)$ on the level set $g(x, y, z) = c$. If $\nabla g(x_0, y_0) \neq \mathbf{0}$, then there exists λ such that

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0) \tag{7.20.1}$$

Exercise 7.21. Find the point on the graph $xy^2z^3 = 2$ which is the closest to the origin.

Exercise 7.22. Find the minimal and maximal value of $f(x, y, z) = x^3 + y^3 + z^3$ on the sphere $x^2 + y^2 + z^2 = 1$ on the first octant. (Figure below shows a level set of f.)



7.4 Exercise

1. Identify the critical points and classify them as local maxima, minima, or saddles.

   (a) $f(x, y) = xy - x^2 - 5y^2 + y - 1$

   (b) $f(x, y) = xy + \dfrac{2}{x} + \dfrac{2}{y}$

   (c) $f(x, y) = \cos x \cos y$

   (d) $f(x, y, z) = x^3 + x^2 z - x^2 + y^2 + z^2$

   (e) $f(x, y, z) = e^y(x^2 + y^2 - z^2)$

2. Find local and global extremes.

   (a) $f(x, y) = e^{x^2+2y^2}$

   (b) $f(x, y) = x^3 + y^3 - xy + 5$

   (c) $f(x, y, z) = x^3 + 3x^2 + e^{y^2+1} + z^2$

   (d) $f(x, y, z) = 2 - x^2y^2 + z^4$

3. Find the local extremes of $f(x, y)$ with the given constraints.

   (a) $f(x, y) = xy$, $2x + 3y = 5$

   (b) $f(x, y) = 2x + y^2 + x^2$, $2x^2 + y^2 = 2$

   (c) $f(x, y, z) = 2x + y^2 - z^2$, $x^2 + y^2 + z^2 = 4$

   (d) $f(x, y, z) = x^2 + y^2 + z^2$, $x + y - z = 1$

   (e) $f(x, y, z) = 2x + y^2$, $x - 2y = 0$, $x + z = 0$

   (f) $f(x, y, z) = xy + yz$, $x^2 + y^2 = 1$, $yz = 1$

4. Find the local extremes of $f(x, y) = x^2 + xy + y^2$ on the disk $D = \{x^2 + y^2 \leq 1\}$.

5. Find the nearest and farthest points on the ellipse obtained by intersecting the cylinder $x^2 + y^2 = 4$ and the plane $2x + 2y + z = 2$.

8 Additional Topics

8.1 Gauss elimination

Definition 8.1. Two systems of linear equations of the form

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{8.1.1}$$

are **equivalent** if one can be transformed into another by the finite steps of the following operations.

1. Multiply a nonzero constant on the both sides of the i-th equation.
2. Subtract (or add) the i-th equation from the j-th equation side-by-side.

Example 8.2. From now on, let us assume that $n \geq m$ for simplicity. Two operations above is the canonical operation we use to find the *solution*, i.e. the n-tuple of $(x_1, \cdots, x_n)$ satisfying the system. Such operation do not change the set of solution satisfying the system, thus our goal is to *simplify* the system so that the solution is apparent.

1. The system of linear equations (8.1.1) can be written in the following matrix form.

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{8.2.1}$$

   Let us write (8.2.1) as

$$A \cdot \mathbf{x} = \mathbf{b}$$

   where $\mathbf{x}, \mathbf{b}$ are column vectors. Note that two operations above corresponds the the following operation on the matrix A.

   (a) Multiply a nonzero constant on a row i.

   (b) Replace the j-th row by subtracting (or adding) the i-row.

   After finite operations, the resulting matrix A gives the system of linear equations equivalent to the original system. There is one more operation on A which produces an equivalent system:

   (c) Interchage the i-th and j-th rows.

   The operations (a)-(c) are called the **elementary row operations**.

2. Let us use the matrix representaion to find the solution of system (8.1.1). As we apply elementary row operation on A, the right-hand side vector $\mathbf{b}$ must be change accordingly. Thus we use $m \times (n+1)$ matrix instead.

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right) \tag{8.2.2}$$

   The Vert ical bar is only for distingulish the entries in A and $\mathbf{b}$. After finite steps of elementary row operations, we can transform (8.2.2) into the following form:

$$\left(\begin{array}{ccccccc|c} a'_{11} & a'_{12} & \cdots & a'_{1m} & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \cdots & a'_{2m} & \cdots & a'_{2n} & b'_2 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a'_{mm} & \cdots & a'_{mn} & b'_m \end{array}\right) \tag{8.2.3}$$

   Since $n \neq m$, depending on the situation, there can be (1) exactly one solution (only if $n = m$), (2) infinitely many solutions, or (3) no solution at all. The

Definition 8.3. The matrix of the form

$$\begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1m} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2m} & \cdots & a'_{2n} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a'_{mm} & \cdots & a'_{mn} \end{pmatrix}$$

is said to be of **row echelon form**. The first nonzero entry on i-th row, called *pivot*, must be at prior column than the pivot on $i+1$-th row. The method of finding solutions of a system of linear equations using row echelon form as in Example 8.2 is called the **Gauss elimination**.

Exercise 8.4. Let E_{ij} be the $n \times n$ matrix whose entries are all zero except 1 at (i, j). Prove that the elementary row operations are equivalent to the following matrix multiplications

$$(I_n + cE_{ii}) \cdot A, \quad (I_n - E_{ji}) \cdot A, \quad (I_n - E_{ii} + E_{ij} - E_{jj} + E_{ji}) \cdot A$$

Example 8.5. The Gauss elimination has a further process. By the elementary row operations, we can transform (8.2.3) further into the following form

$$\left(\begin{array}{ccccccc|c} 1 & 0 & \cdots & 0 & \cdots & a'_{1n} & b'_1 \\ 0 & 1 & \cdots & 0 & \cdots & a'_{2n} & b'_2 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & \cdots & a'_{mn} & b'_m \end{array}\right).$$

Note that the first nonzero entry on each row is the pivot normalized to 1, thus *may not* be on the diagonal. The only guaranteed zeros entries are (1) entries before the pivot on each row, and (2) entries directly above each pivot. Suppose there is k ($k \le m$) rows with nonzero pivots, and the column of i-th pivot is n_i. Thus the system (8.1.1) is equivalent to

$$x_{n_l} + \sum_{\substack{j \neq n_i \\ 1 \le i \le k}} a'_{lj} x_j = b'_l, \quad 1 \le l \le k$$

From this equation, we can conclude that $x_{n_1}, \cdots, x_{n_k}$ is *dependent* to $n-k$ other variables. That is, the solution of (8.2.3) has $n-k$ degrees of freedom.

Definition 8.6. A $m \times n$ matrix A satisfying the following properties is said to be of **reduced row echelon form**.

1. Let n_i be the column of the pivot of i-th row. Then
$$1 \le n_1 \le \cdots \le n_k$$
If $k < m$, then all entries in $k+1, \cdots, m$-th row are zeros.

2. All entries of n_i-th column are zeros except i-th row.

Proposition 8.7. *Suppose that A is invertible $n \times n$ matrix. Then the system of linear equation $A \cdot \mathbf{x} = \mathbf{b}$ has the unique solution*

$$\mathbf{x} = A^{-1} \cdot \mathbf{b}.$$

8.2 Vector spaces

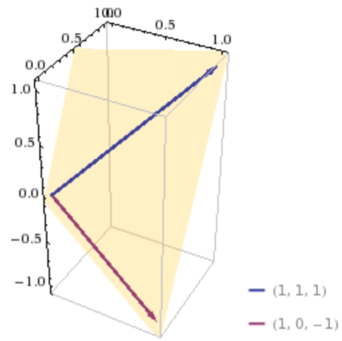
Definition 8.8. A **vector space** V is a set of elements, called *vectors*, satsifying the following properties.

1. For any two vectors $\mathbf{v}, \mathbf{w}$, the sum $\mathbf{v}+\mathbf{w}$ is also a vector and satisfy
$$\mathbf{v}+\mathbf{w} = \mathbf{w}+\mathbf{v}$$
$$(\mathbf{v}+\mathbf{w})+\mathbf{r} = \mathbf{v}+(\mathbf{w}+\mathbf{r})$$
2. For any scalar k, such as real numbers, $k \cdot \mathbf{v}$ is also a vector and satisfy
$$k(\mathbf{v}+\mathbf{w}) = k\mathbf{v}+k\mathbf{w}$$
3. There exists a unique zero vector $\mathbf{0}$ satisfying
$$\mathbf{0}+\mathbf{v} = \mathbf{v}+\mathbf{0} = \mathbf{v}$$
for every vector $\mathbf{v}$ in V.

Definition 8.9. Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ be vectors in $\mathbf{R}^n$ and $c_1, c_2, \cdots, c_n$ real numbers. Then the vector
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$
is called the **linear combination** (or linear sum) of $\mathbf{v}_i$'s. A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is said to be **linearly independent** if the constants $c_1, c_2, \cdots, c_n$ satisfy
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$
if and only if $c_1 = c_2 = \cdots = c_n = 0$.

Definition 8.10. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be a set of (n-dimensional) vectors. A vector space **spanned by** S is the set of all linear combinations of $\mathbf{v}_i$'s.
$$V = \mathrm{span}\langle S\rangle = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \,|\, c_i \in \mathbf{R}\}$$
Note that V is a vector space by Definition 8.8. If S is linearly independent, then it is called the **basis** for V, and the elements $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are called **basis vectors** The size of S, namely n, is called the **dimension** of V, and denoted by
$$n = \dim_{\mathbf{R}} V$$

Example 8.11. Show that two sets of vectors
$$S_1 = \{(1,1,1), (1,0,-1)\}, \quad S_2 = \{(0,1,2), (2,1,0)\}$$
spans the same vector space in $\mathbf{R}^3$ of dimension 2.



Remark 8.12. A finite dimensional vector space can be identified as $\mathbf{R}^n$. However, a vector space is called *infinite dimensional* if no finite collection of vectors can span the vector space. Here are some examples of infinite dimensional vector spaces.

1. The set of all continuous function on $[0,1]$.
$$\mathcal{C}([0,1]) = \{f : [0,1] \to \mathbf{R} \,|\, f \text{ is continuous on } [0,1]\}$$
2. The set of infinite-tuples.
$$\mathbf{R}^\infty = \{(x_1, x_2, x_3, \cdots) \,|\, x_i \in \mathbf{R},\ x_i = 0 \text{ except for finitely many } i\text{'s}\}$$

Definition 8.13. Let A be a $n \times n$ square matrix.
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \tag{8.13.1}$$
Let $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n$ be the n column vectors from A.
$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \cdots, \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}$$
The **rank** of A is the maximal number of linearly independent $\mathbf{c}_i$'s.

Example 8.14. Find the rank of
$$A = \begin{pmatrix} -2 & 1 & -1 \\ -1 & 1 & 2 \\ 2 & -1 & -1 \end{pmatrix}$$

Remark 8.15. Two questions arise naturally.

Q1 Why does rank of a matrix always defined uniquely?

Q2 What if we use row vectors, instead of column vectors, to define the rank?

To answer the first question, we observe the following identity.

$$\text{rank } A = \dim_{\mathbf{R}} \text{span}\langle S\rangle$$

where $S = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_n\}$. Thus the uniqueness of rank A is equivalent to the uniqueness of dimension. Suppose that two sets S_1, S_2 of (n-dimensional) vectors

$$\begin{aligned} S_1 &= \{v_1, \cdots, v_k\} \\ S_2 &= \{w_1, \cdots, w_l\}, \quad l \neq k \end{aligned}$$

spans the same vector space V. If $k > l$, then S_1 cannot be linearly independent. Thus the dimension is uniquely determined.

Let $R = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n\}$ be the set of row vectors of A in (8.13.1).

$$\begin{aligned} \mathbf{r}_1 &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{pmatrix} \\ \mathbf{r}_2 &= \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix} \\ &\vdots \\ \mathbf{r}_n &= \begin{pmatrix} a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \end{aligned}$$

Then Q2 is equivalent to

$$\dim_{\mathbf{R}} \text{span}\langle S\rangle \stackrel{?}{=} \dim_{\mathbf{R}} \text{span}\langle R\rangle$$

This identity is always true. This follow from the simple algorithm to compute the rank of a matrix using an *echelon form*. An echelon form

$$\begin{pmatrix} 0 & 0 & \cdots & * \\ \vdots & \vdots & \iddots & \vdots \\ 0 & * & \cdots & * \\ 0 & * & \cdots & * \end{pmatrix}$$

can be obtained by the following *column operations*:

1. replace a column with itself by multiplying a constant,
2. interchaning two columns, and
3. multiply a column by a scalar, and then subtract from another column.

The number of nonzero entries, $*$, on the diagonal is the rank. One can obtain an echelon form

$$\begin{pmatrix} * & \cdots & * & * \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & * & * \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

by taking a *row operations*, which is similarly defined as above. The crucial fact is that either way, the number of nonzero diagonal entries are the same. That is,

$$\dim_{\mathbf{R}} \text{span}\langle S\rangle = \text{rank } A = \dim_{\mathbf{R}} \text{span}\langle R\rangle$$

Theorem 8.16. *A $n \times n$ matrix A nonsingular if and only if* rank $A = n$.

8.3 Eigenvalues and eigenvectors

Definition 8.17. Let A be a $n \times n$-matrix. If a nonzero vector $\mathbf{v}$ and a contant λ satisfy

$$A \cdot \mathbf{v} = \lambda \mathbf{v},$$

then $\mathbf{v}$ is called the **eigenvector**, and λ the **eigenvalue** of A. Here we consider vector $\mathbf{v}$ as a $n \times 1$-matrix.

Example 8.18. Show that the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are eigenvectors of

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

What are the eigenvalues?

Theorem 8.19. *A constant λ is a eigenvalue of A if and only if it is a root of*

$$\det(A - \lambda I_n) = 0 \tag{8.19.1}$$

*The polynomial (8.19.1) is called the **characteristic polynomial** of A.*

Example 8.20. The fundamental theorem of algebra states that every polynomial has at least one root. Thus there are always n eigenvalues (real and complex) counting multiplicity. Let us see how to get the eigenvectors for each eigenvalue.

1. Let
$$A = \begin{pmatrix} -1 & 0 \\ 1 & 3 \end{pmatrix}$$
The characteristic polynomial is $(-1-\lambda)(3-\lambda) = 0$. Thus $\lambda = -1, 3$ are two eigenvalues. If $\mathbf{v}$ is an eigenvector for -1, then
$$A \cdot \mathbf{v} = -\mathbf{v} \Longrightarrow (A + I) \cdot \mathbf{v} = 0$$
That is,
$$\begin{pmatrix} 0 & 0 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
Thus $v_1 + 4v_2 = 0$ and $\mathbf{v} = (-4, 1)$ is an eigenvector. Similarly, we get $(0, 1)$ as an eigenvector for 3.

2. Let
$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$
Since characteristic polynomial is $(\lambda - 2)^2 = 0$, there is only one eigenvalue $\lambda = 2$. The corresponding eigenvector $\mathbf{v}$ satisfies
$$\begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
so $\mathbf{v} = (1, 1)$ is an eigenvector. Note that such eigenvector is unique up to a scalar multiplication.

3. Let
$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$
The characteristic polynomial is $(\lambda - 1)^2(\lambda - 4) = 0$. For $\lambda = 1$, the eigenvector $\mathbf{v}$ must satisfy
$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
There are two linearly independent eigenvectors. For example,
$$\mathbf{v}_1 = (-1, 1, 0), \quad \mathbf{v}_2 = (-1, 0, 1)$$
are linearly independent eigenvectors for $\lambda = 1$.

4. The eigenvalue can be a complex number. Let
$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$
Then the characteristic polynomial is $\lambda^2 - 2\lambda + 2 = 0$, whose roots are $1 \pm i$[16]. The eigenvector corresponding a complex eigenvalue may have complex entries. For $\lambda = 1 + i$, the eigenvector $\mathbf{v}$ satisfies
$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
Then $\mathbf{v} = (i, 1)$ is an eigenvector for $\lambda = 1 + i$. We can check that the conjuagte $\overline{\mathbf{v}} = (-i, 1)$ is an eigenvector for $\lambda = 1 - i$.

Proposition 8.21. *Let λ be a complex eigenvalue for A and $\mathbf{v}$ a corresponding eigenvector. Then $\overline{\lambda}$ is also an eigenvalue for A and $\overline{\mathbf{v}}$ is a corresponding eigenvector.*

Definition 8.22. Let $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ be a vector whose entries x_i's are functions of t. Let us write

$$\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix} = \frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

[16] Note that complex roots always come as pair.

The system of linear DE

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + f_1(t) \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + f_2(t) \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{b2}x_2 + \cdots + a_{bn}x_n + f_n(t)
\end{aligned}$$

can be written as

$$\frac{d\mathbf{x}}{dt} = A \cdot \mathbf{x} + \mathbf{f} \tag{8.22.1}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

and $\mathbf{f} = (f_1, f_2, \cdots, f_n)$. The system of DE (8.22.1) is called **homogeneous** if $\mathbf{f} = \mathbf{0}$, or **nonhomogeneous** otherwise.

8.4 Solving system of homogeneous DE

Definition 8.23. Let

$$\frac{d\mathbf{x}}{dt} = A \cdot \mathbf{x} \tag{8.23.1}$$

be a system of homogeneous DE. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of A (counting multiplicity), and $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ corresponding eigenvectors. Then the vectors

$$\mathbf{x}_1 = e^{\lambda_1 t}\mathbf{v}_1, \quad \mathbf{x}_2 = e^{\lambda_2 t}\mathbf{v}_2, \quad \cdots, \quad \mathbf{x}_n = e^{\lambda_n t}\mathbf{v}_n$$

are called the **solution vectors**.

Proposition 8.24. *Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ be solution vectors for (8.23.1). Then for any constants $c_1, c_2, \cdots, c_n$,*

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

is also a solution.

Definition 8.25. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ be solution vectors. If

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = 0$$

hold if and only if $c_1 = c_2 = \cdots = c_n = 0$, then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are called **linearly independent**. Otherwise, it is called **linearly dependent**.

Proposition 8.26. *Let*

$$\mathbf{x}_1 = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix}, \quad \cdots, \quad \mathbf{x}_n = \begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{pmatrix}$$

be solution vectors where each entry x_{ij} is an analytic[17] *function. Then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ is linearly independent if and only if the* ***Wronskian***

$$W(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) = \begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix}$$

is nonzero.

Theorem 8.27. *Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be eigenvalues for A counting multiplicity. Suppose that corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are linearly independent. Then*

$$\begin{aligned} \mathbf{x} &= c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n \\ &= c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 + \cdots + c_n e^{\lambda_n t}\mathbf{v}_n \end{aligned}$$

is <u>the</u> general solution of (8.23.1).

[17] A function $x(t)$ is *analytic* if its Taylor expansion always converges to $x(t)$.

Example 8.28. Find the general solution of

$$\begin{aligned} \frac{dx}{dt} &= 2x \\ \frac{dy}{dt} &= x + 3y - 2z \\ \frac{dz}{dt} &= -x + z \end{aligned}$$

Theorem 8.29. *Suppose that λ_1 be an eigenvalue for A with multiplicity m*[18] *($m < n$). Furthermore, suppose that $\mathbf{v}_1$ is the only eigenvector corresponding to λ_1. Then for $\mathbf{v}_2, \mathbf{v}_3, \cdots, \mathbf{v}_m$ satisfying*

$$\begin{aligned} (A - \lambda_1 I)\mathbf{v}_1 &= \mathbf{0} \\ (A - \lambda_1 I)\mathbf{v}_2 &= \mathbf{v}_1 \\ &\vdots \\ (A - \lambda_1 I)\mathbf{v}_n &= \mathbf{v}_n \end{aligned}$$

The vectors

$$\begin{aligned} \mathbf{x}_1 &= e^{\lambda_1 t}\mathbf{v}_1 \\ \mathbf{x}_2 &= e^{\lambda_1 t}\mathbf{v}_2 + te^{\lambda_1 t}\mathbf{v}_1 \\ &\vdots \\ \mathbf{x}_m &= e^{\lambda_1 t}\mathbf{v}_m + te^{\lambda_1 t}\mathbf{v}_{m-1} + \cdots + \frac{t^m}{(m-1)!}e^{\lambda_1 t}\mathbf{v}_1 \end{aligned}$$

are solution vectors for (8.23.1).

Example 8.30. Find the general solution of

$$\begin{aligned} \frac{dx}{dt} &= x + y \\ \frac{dy}{dt} &= -x + 3y \end{aligned}$$

Theorem 8.31. *Suppose that λ_1 be a complex eigenvalues of A, and $\mathbf{v}_1$ a corresponding (complex) eigenvectors. Then*

$$\mathbf{x}_1 = e^{\lambda_1 t}\mathbf{v}_1, \quad \mathbf{x}_2 = \overline{\mathbf{x}_1} = e^{\overline{\lambda_1} t}\mathbf{v}_2$$

are (complex) solution vectors for (8.23.1). In terms of real functions, if $\lambda = \alpha + i\beta$ and $\mathbf{v}_1 = \mathbf{p} + i\mathbf{q}$,

$$\begin{aligned} \mathbf{t}_1 &= e^{\alpha t}(\mathbf{p}\cos\beta t + \mathbf{q}\sin\beta t) \\ \mathbf{t}_2 &= e^{\alpha t}(\mathbf{p}\cos\beta t - \mathbf{q}\sin\beta t) \end{aligned}$$

are solution vectors.

Example 8.32. Find the general solution of

$$\begin{aligned} \frac{dx}{dt} &= 3x - y \\ \frac{dy}{dt} &= 2x + y \end{aligned}$$

[18] That is, counting multiplicity, the spectrum of eigenvalues is

$$\lambda_1 = \lambda_2 = \cdots = \lambda_m, \lambda_{m+1}, \cdots, \lambda_n$$

8.5 Methods of solving systems of nonhomogeneous DEs

Theorem 8.33. *Let*

$$\frac{d\mathbf{x}}{dt} = A \cdot \mathbf{x} + \mathbf{f} \qquad (8.33.1)$$

be a system of nonhomogeneous DE. A solution $\mathbf{x}_u$ *satisfying*

$$\frac{d\mathbf{x}}{dt} = A \cdot \mathbf{x}$$

is called a ***complementary solution*** *of* (8.33.1). *A solution* $\mathbf{x}_p$ *without any parameters satisfying* (8.33.1) *is called a* ***particular solution****. If* $\mathbf{x}_u$ *is the general complementary solution, then the general solution of* (8.33.1) *is given by*

$$\mathbf{x} = \mathbf{x}_u + \mathbf{x}_p$$

Example 8.34 (Method of determined coefficients)**.** Find the general solution of

$$\begin{aligned} \frac{dx}{dt} &= x + 3y - 3 \\ \frac{dy}{dt} &= 3x + y + 2 \end{aligned}$$

Theorem 8.35. *Let*

$$\mathbf{x}_u = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

be the complementary solution of (8.33.1). *Let us denote*

$$\Phi = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

Then

$$\mathbf{x}_p = \Phi \cdot \int \Phi^{-1} \cdot \mathbf{f} dt$$

is a particular solution of (8.33.1).

Example 8.36. Find the general solution of

$$\begin{aligned} \frac{dx}{dt} &= 3x + 2y - e^{-t} \\ \frac{dy}{dt} &= x + 2y - 2 \end{aligned}$$

References

[1] E. Kreyszig. *Advanced engineering mathematics.* John Wiley & Sons, Inc., tenth edition, 2011.

[2] J. R. Munkres. *Analysis on manifolds.* Addison-Wesley Publishing Company, 1991.

[3] J. Stewart. *Calculus, early transcendentals.* Brooks Cole, eighth edition, 2015.