

# Handbook of Applied Mathematics

Zeyu Wang<sup>1</sup>

May 14, 2024

---

<sup>1</sup>Email: zeyuwang@zuua.zju.edu.cn

# Contents

Contents .....	i
Mathematical Foundation .....	1
1 Analysis .....	2
1.1 Calculus .....	2
1.1.1 Gauss-Green Theorem (Divergence Theorem) .....	2
1.2 Important Inequalities .....	2
1.2.1 Fundamental inequality .....	2
1.2.2 Triangle inequality .....	2
1.2.3 Bernoulli inequality .....	2
1.2.4 Jensen's inequality .....	2
1.2.5 Cauchy-Schwarz inequality .....	3
1.2.6 Hölder's inequality .....	3
1.2.7 Young's inequality .....	3
1.2.8 Minkowski inequality .....	4
1.3 Special Functions .....	4
1.3.1 Gaussian function .....	4
1.3.2 Dirac delta function .....	4
1.3.3 Gamma function .....	4
1.3.4 Beta Function .....	5
2 Algebra .....	6
2.1 Linear Space .....	6
2.1.1 Linear map .....	6
2.2 Metric Space .....	7
2.2.1 Completeness & Compactness .....	8
2.2.2 Cover .....	8
2.2.3 Cantor's intersection Theorem .....	8
2.2.4 Cluster point .....	8
2.3 Normed Space .....	9
2.4 Inner Product Space .....	9
2.4.1 Orthonormal system .....	9
2.5 Banach Space .....	10
2.6 Hilbert Space .....	10
2.7 Single Variable Polynomial .....	10
2.8 Orthogonal Polynomial .....	11
2.8.1 Legendre polynomial .....	11
2.8.2 Chebyshev polynomial of the first kind .....	11
2.8.3 Chebyshev polynomial of the second kind .....	12
2.8.4 Laguerre polynomial .....	12
2.8.5 Hermite polynomial (probability theory form) .....	12

---

3	Ordinary Differential Equation .....	14
3.1	General Theory .....	14
4	Partial Differential Equation .....	16
4.1	Poisson Equation .....	16
4.2	Heat Equation .....	16
4.3	Wave Equation .....	16
5	Probability Theory .....	17
5.1	Discrete random variables .....	17
5.2	Continuous random variables .....	17
5.3	Characteristic functions .....	17
5.4	Probability limit theorems .....	17
6	Stochastic Process .....	18
6.1	Poisson process .....	18
6.2	Markov chain .....	18
7	Statistics .....	19
8	Graph .....	20
8.1	Shortest Path .....	20
8.2	Matching .....	20
8.3	Network Flow .....	20
8.4	Tree .....	20
9	Combinatorics .....	21
9.1	Generating function .....	21
9.2	Inclusion–exclusion principle .....	21
9.3	Special Numbers .....	21
9.3.1	Catalan number .....	21
9.3.2	Stirling number .....	21
	Scientific Computing .....	22
10	Interpolation .....	23
10.1	Polynomial Interpolation .....	23
10.1.1	Lagrange formula .....	23
10.1.2	Newton formula .....	23
10.1.3	Neville-Aitken algorithm .....	24
10.1.4	Hermite interpolation .....	24
10.1.5	Approximation .....	24
10.1.6	Error analysis .....	25
10.2	Spline .....	25
10.2.1	Cubic spline .....	25
10.2.2	B-spline .....	26
10.2.3	Error analysis .....	26
11	Integration .....	27
11.1	Newton-Cotes Formulas .....	27
11.1.1	Midpoint rule .....	27
11.1.2	Trapezoidal rule .....	28
11.1.3	Simpson’s rule .....	28

---

11.2	Gauss Formulas .....	28
12	Optimization .....	29
12.1	One-dimensional Line Search .....	29
12.2	Unconstrained Optimization .....	30
13	Initial Value Problem .....	32
13.1	Linear Multistep Method .....	32
13.2	Runge-Kutta Method .....	32
13.3	Theoretical analysis .....	33
13.3.1	Error analysis .....	33
13.3.2	Stability .....	33
13.3.3	Convergence .....	34
13.4	Important Methods .....	34
13.4.1	Forward Euler's method .....	34
13.4.2	Backward Euler's method .....	35
13.4.3	Trapezoidal method .....	35
13.4.4	Midpoint method (Leapfrog method) .....	35
13.4.5	Heun's third-order RK method .....	35
13.4.6	Classical fourth-order RK method .....	35
13.4.7	TR-BDF2 method .....	36
14	Number Theory .....	37
14.1	Prime Number .....	37
14.1.1	Primality testing .....	37
14.1.2	Sieves .....	37
	Machine Learning .....	38
15	Regression .....	39
15.1	Linear Regression .....	39
16	Decision Tree .....	40
17	Support Vector Machine .....	41
18	Cluster .....	42
19	Neural Networks .....	43

# **Part 1**

# **Mathematical Foundation**

# Chapter 1

## Analysis

### 1.1 Calculus

#### 1.1.1 Gauss-Green Theorem (Divergence Theorem)

**Theorem 1.1. (Gauss-Green Theorem (Divergence Theorem))** For a bounded open set  $\Omega \in \mathbb{R}^n$  that  $\partial\Omega \in C^1$  and a function  $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \dots, F_n(\mathbf{x})) : \overline{\Omega} \rightarrow \mathbb{R}^n$  satisfies  $\mathbf{F}(\mathbf{x}) \in C^1(\Omega) \cap C(\overline{\Omega})$ ,

$$\int_{\Omega} \operatorname{div} \mathbf{F}(\mathbf{x}) d\mathbf{x} = \int_{\partial\Omega} \mathbf{F}(\mathbf{x}) \cdot \mathbf{n} dS(x),$$

where  $\mathbf{n}$  is outward pointing unit normal vector at  $\partial\Omega$ .

### 1.2 Important Inequalities

#### 1.2.1 Fundamental inequality

**Theorem 1.2. (Fundamental inequality)**

$$\forall x, y \in \mathbb{R}^+, \frac{2}{\frac{1}{a} + \frac{1}{b}} \leq \sqrt{ab} \leq \frac{a+b}{2} \leq \sqrt{\frac{a^2 + b^2}{2}}, \text{ equality holds iff } a = b.$$

#### 1.2.2 Triangle inequality

**Theorem 1.3. (Triangle inequality)**

$$\begin{aligned} a, b \in \mathbb{C}, \quad ||a| - |b|| \leq |a \pm b| \leq |a| + |b|, \\ \mathbf{a}, \mathbf{b} \in \mathbb{R}^n, \quad ||\mathbf{a}| - |\mathbf{b}|| \leq \|\mathbf{a} \pm \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|. \end{aligned}$$

#### 1.2.3 Bernoulli inequality

**Theorem 1.4. (Bernoulli inequality)**

$$\begin{aligned} \forall x \in (-1, +\infty), \forall a \in [1, +\infty), \quad (1+x)^a &\geq 1+ax, \\ \forall x \in (-1, +\infty), \forall a \in (0, 1), \quad (1+x)^a &\leq 1+ax, \\ \forall x \in (-1, +\infty), \forall a \in (-1, 0), \quad (1+x)^a &\geq 1+ax, \\ \forall x_i \in \mathbb{R}, i \in \{1, \dots, n\}, \quad \prod_{i=1}^n (1+x_i) &\geq 1 + \sum_{i=1}^n x_i, \\ \forall y \geq x > 0, \quad (1+x)^y &\geq (1+y)^x. \end{aligned}$$

#### 1.2.4 Jensen's inequality

**Theorem 1.5. (Jensen's inequality)** For a real convex function  $f(x) : [a, b] \rightarrow \mathbb{R}$ , numbers  $x_1, \dots, x_n \in [a, b]$  and weights  $a_1, \dots, a_n$ , the Jensen's inequality can be start as

$$\frac{\sum_{i=1}^n a_i f(x_i)}{\sum_{i=1}^n a_i} \geq f\left(\frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n a_i}\right).$$

And for concave function  $f$ ,

$$\frac{\sum_{i=1}^n a_i f(x_i)}{\sum_{i=1}^n a_i} \leq f\left(\frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n a_i}\right).$$

Equality holds iff  $x_1 = \dots = x_n$  or  $f$  is linear on  $[a, b]$ .

### 1.2.5 Cauchy–Schwarz inequality

**Theorem 1.6. (Cauchy–Schwarz inequality)**

**Discrete form.** For real numbers  $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}, n \geq 2$

$$\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 \geq \left( \sum_{i=1}^n a_i b_i \right)^2.$$

Equality holds iff  $\frac{a_1}{b_1} = \dots = \frac{a_n}{b_n}$  or  $a_i = 0$  or  $b_i = 0$ .

**Inner product form.** For a inner product space  $V$  with a norm induced by the inner product,

$$\forall \mathbf{a}, \mathbf{b} \in V \quad \|\mathbf{a}\| \cdot \|\mathbf{b}\| \geq |\langle \mathbf{a}, \mathbf{b} \rangle|.$$

Equality holds iff  $\exists k \in \mathbb{R}$ , s.t.  $k\mathbf{a} = \mathbf{b}$  or  $\mathbf{a} = k\mathbf{b}$ .

**Probability form.** For random variables  $X$  and  $Y$ ,

$$\sqrt{E(X^2)} \cdot \sqrt{E(Y^2)} \geq |E(XY)|.$$

Equality holds iff  $\exists k \in \mathbb{R}$ , s.t.  $kX = Y$  or  $X = kY$ .

**Integral form.** For integrable functions  $f, g \in L^2(\Omega)$ ,

$$\int_{\Omega} f^2(x) dx + \int_{\Omega} g^2(x) dx \geq \left( \int_{\Omega} f(x)g(x) dx \right)^2.$$

Equality holds iff  $\exists k \in \mathbb{R}$ , s.t.  $kf(x) = g(x)$  or  $f(x) = kg(x)$ .

### 1.2.6 Hölder's inequality

**Theorem 1.7. (Hölder's inequality)**

**Discrete form.** For real numbers  $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}, n \geq 2$  and  $p, q \in [1, +\infty)$  that  $\left(\frac{1}{p}\right) + \left(\frac{1}{q}\right) = 1$ ,

$$\left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}} \geq \left( \sum_{i=1}^n a_i b_i \right).$$

Equality holds iff  $\exists c_1, c_2 \in \mathbb{R}, c_1^2 + c_2^2 \neq 0$ , s.t.  $c_1 a_i^p = c_2 b_i^q$ .

**Integral form.** For functions  $f \in L^p(\Omega), g \in L^q(\Omega)$  and  $p, q \in [1, +\infty)$  that  $\left(\frac{1}{p}\right) + \left(\frac{1}{q}\right) = 1$ ,

$$\left( \int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}} + \left( \int_{\Omega} |g(x)|^q dx \right)^{\frac{1}{q}} \geq \int_{\Omega} f(x)g(x) dx.$$

### 1.2.7 Young's inequality

**Theorem 1.8. (Young's inequality)** For  $p, q \in [1, +\infty)$  that  $\frac{1}{p} + \frac{1}{q} = 1$ ,

$$\forall a, b \in \mathbb{R}^*, \frac{a^p}{p} + \frac{b^q}{q} \geq ab.$$

Equality holds iff  $a^p = b^q$ .

### 1.2.8 Minkowski inequality

**Theorem 1.9. (Minkowski inequality)** For a metric space  $S$ ,

$$\forall f, g \in L^p(S), p \in [1, +\infty], \|f\|_p + \|g\|_p \geq \|f + g\|_p.$$

For  $p \in (1, +\infty)$ , equality holds iff  $\exists k \geq 0$ , s.t.  $f = kg$  or  $kf = g$ .

## 1.3 Special Functions

### 1.3.1 Gaussian function

**Definition 1.10.** A **Gaussian function**, or a **Gaussian**, is a function of the form

$$f(x) = a \exp\left(-\frac{(x-b)^2}{2c^2}\right),$$

where  $a \in \mathbb{R}^+$  is the height of the curve's peak,  $b \in \mathbb{R}$  is the position of the center of the peak and  $c \in \mathbb{R}^+$  is the standard deviation or the Gaussian root mean square width.

**Theorem 1.11.** The integral of a Gaussian is

$$\int_{-\infty}^{+\infty} a \exp\left(-\frac{(x-b)^2}{2c^2}\right) dx = ac\sqrt{2\pi}.$$

**Definition 1.12.** A **normal distribution** or a **Gaussian distribution** is a continuous probability distribution of the form

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

where  $\mu$  is the mean and  $\sigma$  is the standard deviation.

### 1.3.2 Dirac delta function

**Definition 1.13.** The **Dirac delta function** centered at  $\bar{x}$  is

$$\delta(x - \bar{x}) = \lim_{\varepsilon \rightarrow 0} f_{\bar{x},\varepsilon}(x - \bar{x}),$$

where  $f_{\bar{x},\varepsilon}$  is a normal distribution with its mean at  $\bar{x}$  and its standard deviation as  $\varepsilon$ .

**Theorem 1.14.** The Dirac delta function satisfies

$$\delta(x - \bar{x}) = \begin{cases} +\infty, & x = \bar{x} \\ 0, & x \neq \bar{x} \end{cases} \quad \int_{-\infty}^x \delta(x - \bar{x}) dx = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where  $H(x) = \int_{-\infty}^x \delta(x - \bar{x}) dx$  is called **Heaviside function** or **step function**.

**Theorem 1.15.** If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuous, then

$$\int_{-\infty}^{+\infty} \delta(x - \bar{x}) f(x) dx = f(\bar{x}).$$

### 1.3.3 Gamma function

**Definition 1.16.** The **Gamma function** defined on  $\mathbb{C}$  is

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt,$$

where  $\text{Re}(z) > 0$ .



**Theorem 1.17.** The Gamma function satisfies

$$\begin{aligned}\forall x \in \mathbb{C}, \quad \Gamma(x+1) &= x\Gamma(x), \\ \forall n \in \mathbb{N}^*, \Gamma(n) &= (n-1)!. \end{aligned}$$

**Theorem 1.18.** The Gamma function satisfies

$$\forall x \in (0, 1), \Gamma(1-x)\Gamma(x) = \frac{\pi}{\sin(\pi x)},$$

which implies

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

### 1.3.4 Beta Function

**Definition 1.19.** For  $p, q \in \mathbb{R}^+$ , the **Beta function** is defined as

$$B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx.$$

**Theorem 1.20.** The Beta function satisfies

$$\forall p, q \in \mathbb{R}^+, B(p, q) = B(q, p) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

**Theorem 1.21.** The Beta function satisfies

$$\begin{aligned}\forall p > 0, \forall q > 1, B(p, q) &= \frac{q-1}{p+q-1} B(p, q-1), \\ \forall p > 1, \forall q > 0, B(p, q) &= \frac{p-1}{p+q-1} B(p-1, q), \\ \forall p > 1, \forall q > 1, B(p, q) &= \frac{(p-1)(q-1)}{(p+q-1)(p+q-2)} B(p-1, q-1). \end{aligned}$$

# Chapter 2

## Algebra

### 2.1 Linear Space

**Definition 2.1. (Linear Space)** A linear space over a field  $\mathbb{F}$  is a nonempty set  $V$  with a addition and a scalar multiplication that satisfies

- (1) Associativity of addition:  $\forall \mathbf{x}, \mathbf{y} \in V, \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ ,
- (2) Commutativity of addition:  $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ ,
- (3) Identity element of addition:  $\exists \mathbf{0} \in V, \forall \mathbf{x}, \mathbf{x} + \mathbf{0} = \mathbf{x}$ ,
- (4) Inverse elements of addition:  $\forall \mathbf{x} \in V, \exists \mathbf{y} \in V, \text{ s.t. } \mathbf{x} + \mathbf{y} = \mathbf{0}$ ,
- (5) Compatibility of multiplication:  $\forall \mathbf{x} \in V, a, b \in \mathbb{F}, (ab)\mathbf{x} = a(b\mathbf{x})$ ,
- (6) Identity element of multiplication:  $\exists 1 \in \mathbb{F}, \forall \mathbf{x} \in V, 1\mathbf{x} = \mathbf{x}$ ,
- (7) Distributivity:  $\forall \mathbf{x} \in V, a, b \in \mathbb{F}, (a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ ,
- (8) Distributivity:  $\forall \mathbf{x}, \mathbf{y} \in V, a \in \mathbb{F}, a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$ .

**Notation 2.2.** The **dimension** of a linear space  $V$  is written as  $\dim(V)$ .

**Definition 2.3.** Denoted by  $V_1, \dots, V_n$  linear spaces over a field  $\mathbb{F}$ , the **product of linear spaces** is defined as

$$V_1 \times \dots \times V_n = \{(\mathbf{v}_1, \dots, \mathbf{v}_n) : \mathbf{v}_1 \in V_1, \dots, \mathbf{v}_n \in V_n\},$$

which is also a linear space over  $\mathbb{F}$ .

**Definition 2.4.** Given a linear space  $V$ , a subspace  $U \subset V$  and  $\mathbf{v} \in V$ , the **coset** (or **affine subset**) is defined as

$$\bar{\mathbf{v}} = \{\mathbf{w} \in V : \mathbf{w} = \mathbf{v} + \mathbf{u}, \mathbf{u} \in U\}.$$

**Definition 2.5.** Given a linear space  $V$  and a subspace  $U \subset V$ , the **quotient space** is defined as

$$V/U = \{\mathbf{v} + U : \mathbf{v} \in V\}.$$

#### 2.1.1 Linear map

**Definition 2.6.** Denoted by  $V$  and  $W$  the linear spaces over a field  $\mathbb{F}$ , a function  $f : V \rightarrow W$  is called a linear map between  $V$  and  $W$  if it satisfies

- (1) Additivity:  $\forall \mathbf{x}, \mathbf{y} \in V, f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ ;
- (2) Homogeneity:  $\forall \mathbf{x} \in V, \forall k \in \mathbb{F}, f(k\mathbf{x}) = kf(\mathbf{x})$ .

**Notation 2.7.** Denoted by  $\mathcal{L}(V, W)$  the set of all linear maps between  $V$  and  $W$  (it also be written as  $\mathcal{L}(V)$  if  $V = W$ ).

**Theorem 2.8.** For linear space  $V, W$  over a field  $\mathbb{F}$  and linear maps  $f, g \in \mathcal{L}(V, W)$ , if we define

$$\forall \mathbf{x} \in V, \forall k \in \mathbb{F}, (f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \text{ and } (kf)(\mathbf{x}) = kf(\mathbf{x}),$$

then  $\mathcal{L}(V, W)$  is a linear space.

**Theorem 2.9.** For a linear map  $f \in \mathcal{L}(V, W)$ ,  $f(\mathbf{0}) = f(0\mathbf{v}) = 0f(\mathbf{v}) = \mathbf{0}$ .

**Theorem 2.10.** Given  $\mathbf{v}_1, \dots, \mathbf{v}_n$  the basis of linear space  $V$  and  $\mathbf{w}_1, \dots, \mathbf{w}_n$  the basis of linear space  $W$ , then there exists the only linear map  $f \in \mathcal{L}(V, W)$  such that

$$\forall i \in \{1, \dots, n\}, f(\mathbf{v}_i) = \mathbf{w}_i.$$

**Definition 2.11.** For a linear map  $f \in \mathcal{L}(V, W)$ , the **kernal** (or **null space**) of  $f$  is defined as

$$\ker(f) = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\},$$

where  $\ker(f)$  is a subspace of  $V$  and the number  $\dim(\ker(f))$  is the **nullity** of  $f$  which also written as  $\text{nullity}(f)$

**Definition 2.12.** For a linear map  $f \in \mathcal{L}(V, W)$ , the **image** of  $f$  is defined as

$$\text{im}(f) = \{\mathbf{w} \in W : \mathbf{w} = f(\mathbf{v}), \mathbf{v} \in V\},$$

where  $\text{im}(f)$  is a subspace of  $W$  and the number  $\dim(\text{im}(f))$  is the **dimension** (or **rank**) of  $f$  which also written as  $\text{rank}(f)$

**Theorem 2.13. (Rank–nullity theorem)** For a linear map  $f \in \mathcal{L}(V, W)$ ,

$$\dim(\ker(f)) + \dim(\text{im}(f)) = \dim(V).$$

**Definition 2.14.** A **isomorphism** is a invertible linear map.

**Definition 2.15.** Two linear spaces are called **isomorphic** if there exists a invertible linear map between them.

**Theorem 2.16.** Two linear spaces  $V, W$  over a field  $\mathbb{F}$  are isomorphic iff  $\dim(V) = \dim(W)$ .

**Theorem 2.17.** For a linear space  $V$  that  $\dim(V) < +\infty$  and a linear map  $f \in \mathcal{L}(V)$ , the following statements are equivalent:

- (1)  $f$  is invertible;
- (2)  $f$  is injective;
- (3)  $f$  is surjective.

## 2.2 Metric Space

**Definition 2.18. (Metric)** For a nonempty set  $X$ , the **metric** is a function  $d : X \times X \rightarrow \mathbb{R}$  that satisfies

- (1) Positive definiteness:  $\forall \mathbf{x}, \mathbf{y} \in X, d(\mathbf{x}, \mathbf{y}) \geq 0, d(\mathbf{x}, \mathbf{y}) \Leftrightarrow \mathbf{x} = \mathbf{y}$ ,
- (2) Symmetry:  $\forall \mathbf{x}, \mathbf{y} \in X, d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ ,
- (3) Triangle inequality:  $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z})$ ,

**Definition 2.19. (Metric space)** A **metric space** is a set  $X$  provided with a metric.

**Notation 2.20. (Neighbourhood)** For a metric space  $X$ , the **neighbourhood** of  $\mathbf{x} \in X$  with radius  $\varepsilon > 0$  is defined as

$$U_X(\mathbf{x}, \varepsilon) = \{t : d(\mathbf{x}, t) < \varepsilon, t \in X\}.$$

**Notation 2.21. (Punctured neighbourhood)** For a metric space  $X$ , the **punctured neighbourhood** of  $\mathbf{x} \in X$  with radius  $\varepsilon > 0$  is defined as

$$U_X^\circ(\mathbf{x}, \varepsilon) = U_X(\mathbf{x}, \varepsilon) \setminus \{\mathbf{x}\} = \{t : d(\mathbf{x}, t) < \varepsilon, t \in X \setminus \{\mathbf{x}\}\}.$$

### 2.2.1 Completeness & Compactness

**Theorem 2.22. (Cauchy's convergence test)** A sequence  $\{\mathbf{x}_n\}$  in a metric space  $X$  is convergent (or said a **cauchy sequence**) iff

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \text{ s.t. } \forall m, n > N, \|\mathbf{x}_n - \mathbf{x}_m\| < \varepsilon.$$

**Definition 2.23. (Completeness)** A metric space  $X$  is **complete** iff all cauchy sequence of  $X$  is convergent in  $X$ .

**Theorem 2.24. (Supremum and infimum principle)** For a nonempty set  $X$ , if the upper/lower bound of  $X$  exists, then the supremum/infimum of  $X$  exists.

**Theorem 2.25. (The monotone bounded convergence Theorem)** For a bounded sequence  $\{\mathbf{x}_n\}$ , if it is increased, then

$$\lim_{n \rightarrow \infty} \mathbf{x}_n = \sup\{\mathbf{x}_n : n \in \mathbb{N}\}.$$

If it is decreased, then

$$\lim_{n \rightarrow \infty} \mathbf{x}_n = \inf\{\mathbf{x}_n : n \in \mathbb{N}\}.$$

### 2.2.2 Cover

**Definition 2.26. (Cover)** For a metric space  $S \subseteq X$ , A **cover** of  $S$  is a set of open sets  $\{D_n\}$  satisfies

$$\forall \mathbf{x} \in X, \exists D_n, \text{ s.t. } \mathbf{x} \in D_n.$$

**Definition 2.27. (Compactness)** A metric space  $X$  is called **compact** if every open cover of  $X$  has a finite subcover.

### 2.2.3 Cantor's intersection Theorem

**Theorem 2.28. (Cantor's intersection Theorem)** For a decreasing sequence of nested non-empty compact, closed subsets  $S_n \subseteq X, n \in \mathbb{N}$  of a metric space, if  $\{S_n\}$  satisfies

$$S_0 \supset S_1, \dots, \supset S_n \supset \dots,$$

then

$$\bigcap_{k=0}^{\infty} S_k \neq \emptyset.$$

where there is only one point  $\mathbf{x} \in \bigcap_{k=0}^{\infty} S_k$  for a complete metric space.

**Corollary 2.29.** For decreasing sequence of nested non-empty compact, closed subsets  $S_n \in X, n \in \mathbb{N}$  of a complete metric space and  $\{\mathbf{x}\} = \bigcap_{k=0}^{\infty} S_k$ , then

$$\forall \varepsilon > 0, \exists N > 0, \text{ s.t. } \forall n > N, X_n \subset U_X(x, \varepsilon).$$

### 2.2.4 Cluster point

**Definition 2.30. (Cluster point)** For a metric space  $S \subseteq X$ , the **cluster point** of  $S$  is the point  $\mathbf{x} \in X$  satisfies

$$\forall \varepsilon > 0, U_X^\circ(\mathbf{x}, \varepsilon) \cap S \neq \emptyset.$$

**Theorem 2.31.** For a convergent sequence  $\{\mathbf{x}_n : n \in \mathbb{N}, \forall i \neq j, \mathbf{x}_i \neq \mathbf{x}_j\} \subseteq X$ , the point  $x = \lim_{n \rightarrow \infty} \mathbf{x}_n$  is a cluster point of  $X$ .

**Theorem 2.32. (Bolzano–Weierstrass Theorem)** For a metric sapce  $X$  and a bounded infinite subset  $S \subseteq X$ , there exists at least one cluster point of  $X$ .

## 2.3 Normed Space

**Definition 2.33. (Norm)** For a linear space  $V$  over a field  $\mathbb{F}$ , the **norm** is a function  $\|\cdot\| : V \rightarrow \mathbb{F}$  that satisfies

- (1) Positive definiteness:  $\forall \mathbf{x} \in V, \|\mathbf{x}\| \geq 0, \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$ ,
- (2) Absolute homogeneity:  $\forall \mathbf{x} \in V, k \in \mathbb{F}, \|k\mathbf{x}\| = |k| \|\mathbf{x}\|$ ,
- (3) Triangle inequality:  $\forall \mathbf{x}, \mathbf{y} \in V, \|\mathbf{x}\| + \|\mathbf{y}\| \geq \|\mathbf{x} + \mathbf{y}\|$ ,

**Definition 2.34. (Normed space)** A **normed space** is a linear space  $V$  over the the field  $\mathbb{F}$  with a norm.

## 2.4 Inner Product Space

**Definition 2.35. (Inner product)** For a linear space  $V$  over a field  $\mathbb{F}$ , the **inner product** on  $V$  is a function  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$  that satisfies

- (1) Positive definiteness:  $\forall \mathbf{x} \in V, \langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = 0$ ,
- (2) Conjugate symmetry:  $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ ,
- (3) Linearity in the first argument:  $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V, a, b \in \mathbb{F}, \langle a\mathbf{x} + b\mathbf{z}, \mathbf{y} \rangle = a\langle \mathbf{x}, \mathbf{y} \rangle + b\langle \mathbf{z}, \mathbf{y} \rangle$ .

**Definition 2.36. (Inner product space)** An **inner product space** is a linear space  $V$  over the field  $\mathbb{F}$  with an inner product.

**Theorem 2.37.** Given a inner product space  $V$  and the norm defined as  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$  satisfies

$$\forall \mathbf{x}, \mathbf{y} \in V, \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2 \|\mathbf{x}\|^2 + 2 \|\mathbf{y}\|^2.$$

### 2.4.1 Orthonormal system

**Definition 2.38.** A subset  $W$  of an inner product space  $V$  is called **orthonormal** if

$$\forall \mathbf{u}, \mathbf{v} \in S, \langle \mathbf{u}, \mathbf{v} \rangle = \begin{cases} 0, & u \neq v \\ 1, & u = v. \end{cases}$$

**Definition 2.39.** The **Gram-Schmidt process** takes in a finite or infinite independent list  $(\mathbf{u}_1, \mathbf{u}_2, \dots)$  and output two other lists  $(\mathbf{v}_1, \mathbf{v}_2, \dots)$  and  $(\mathbf{u}_1^*, \mathbf{u}_2^*, \dots)$  by

$$\mathbf{v}_{n+1} = \mathbf{u}_{n+1} - \sum_{i=1}^n \langle \mathbf{u}_{n+1}, \mathbf{u}_i^* \rangle \mathbf{u}_i^*,$$

$$\mathbf{u}_{n+1}^* = \frac{\mathbf{v}_{n+1}}{\|\mathbf{v}_{n+1}\|},$$

with the recursion basis as  $\mathbf{v}_1 = \mathbf{u}_1$ .

**Definition 2.40.** Let  $(\mathbf{u}_1^*, \mathbf{u}_2^*, \dots)$  be a finite or infinite orthonormal list. The **orthogonal expansion** or **Fourier expansion** for an arbitrary  $\mathbf{w}$  is the series

$$\sum_{i=1}^n \langle \mathbf{w}, \mathbf{u}_i^* \rangle \mathbf{u}_i^*,$$

where the constants  $\langle \mathbf{w}, \mathbf{u}_i^* \rangle$  are known as the **Fourier coefficients** of  $\mathbf{w}$  and the term  $\langle \mathbf{w}, \mathbf{u}_i^* \rangle \mathbf{u}_i^*$  is the **projection** of  $\mathbf{w}$  on  $\mathbf{u}_i^*$ .

**Theorem 2.41. (Minimum properties of Fourier expansions)** Let  $\mathbf{u}_1^*, \mathbf{u}_2^*, \dots$  be an orthonormal system and let  $\mathbf{w}$  be arbitrary. Then

$$\forall a_1, \dots, a_n, \left\| \mathbf{w} - \sum_{i=1}^n \langle \mathbf{w}, \mathbf{u}_i^* \rangle \mathbf{u}_i^* \right\| \leq \left\| \mathbf{w} - \sum_{i=1}^n a_i \mathbf{u}_i^* \right\|,$$

where  $\left\| \mathbf{w} - \sum_{i=1}^n a_i \mathbf{u}_i^* \right\|$  is minimized only when  $a_i = \langle \mathbf{w}, \mathbf{u}_i^* \rangle$ .

**Theorem 2.42. (Bessel inequality)** Let  $\mathbf{u}_1^*, \mathbf{u}_2^*, \dots$  be an orthonormal system and let  $\mathbf{w}$  be arbitrary. Then

$$\sum_{i=1}^n |\langle \mathbf{w}, \mathbf{u}_i^* \rangle|^2 \leq \|\mathbf{w}\|^2.$$

## 2.5 Banach Space

**Definition 2.43. (Banach space)** A **Banach space** is a complete normed vector space.

## 2.6 Hilbert Space

**Definition 2.44. (Hilbert space)** A **Hilbert space** is a inner product space that is also ce with respect to the distance function induced by the inner product.a complete metric space.

## 2.7 Single Variable Polynomial

**Definition 2.45.** Denoted by  $\mathbb{V}$  a linear space and  $x$  the variable, a **(single variable) polynomial** over  $\mathbb{V}$  is defined as

$$p_{n(x)} = \sum_{i=0}^n c_i x^i,$$

where  $c_0, \dots, c_n \in \mathbb{V}$  are constants that called the **coefficients of the polynomial**.

**Definition 2.46.** Given a polynomial  $p(x) = \sum_{i=0}^n c_i x^i$  where  $c_n \neq 0$ , the degree of  $p(x)$  is marked as  $\deg(p(x)) = n$ . In particular, the degree of zero polynomial  $p(x) = 0$  is  $\deg(0) = -\infty$ .

**Theorem 2.47.** Denoted by  $\mathbb{P}_n = \{p : \deg(p) \leq n\}$  the set of polynomials with degree no more than  $n$  ( $n \geq 0$ ), and  $\mathbb{P} = \bigcup_{n=0}^{\infty} \mathbb{P}_n$  the set contains all polynomials, then  $\mathbb{P}_n$  is a linear space and satisfies

$$\{0\} = \mathbb{P}_0 \subset \mathbb{P}_1 \subset \dots \subset \mathbb{P}_n \subset \dots \mathbb{P}$$

**Theorem 2.48. (Vieta's formulas)** Given a polynomial  $p \in \mathbb{P}_n$  with the coefficients being real or complex numbers, denoted by  $x_1, \dots, x_n$  the complex roots, then

$$\begin{cases} x_1 + \cdots + x_n = -c_{n-1}, \\ \sum_{i=1}^n \sum_{j=i+1}^n x_i x_j = c_{n-2}, \\ \quad \quad \quad \dots \\ \prod_{i=1}^n x_i = (-1)^n c_0, \end{cases}$$

where  $c_n = 1$  WLOG.

## 2.8 Orthogonal Polynomial

**Definition 2.49.** Given a weight function  $\rho(x) : [a, b] \rightarrow \mathbb{R}^+$ , satisfies

$$\int_a^b \rho(x) dx > 0, \int_a^b x^k \rho(x) dx > 0 \text{ exists.}$$

The set of **orthogonal polynomials** on  $[a, b]$  with the weight function  $\rho(x)$  is defined as

$$\{p_i, i \in \mathbb{N}\} \subset L_\rho([a, b]) = \left\{ f(x) : \int_a^b f^2(x) \rho(x) dx < \infty \right\}.$$

where  $\{p_i, i \in \mathbb{N}\}$  are calculate from  $\{x^n, n \in \mathbb{N}\}$  using the Gram-Schmidt process with the inner product

$$\forall f, g \in L_\rho([a, b]), \langle f, g \rangle = \int_a^b \rho(x) f(x) g(x) dx.$$

**Theorem 2.50.** Orthogonal polynomials  $p_{n-1}(x), p_n(x), p_{n+1}(x)$  satisfies

$$p_{n+1}(x) = (a_n + b_n x) p_n(x) + c_n p_{n-1}(x).$$

where  $a_n, b_n, c_n$  are depends on  $[a, b]$  and  $\rho$ .

**Theorem 2.51.** The orthogonal polynomial  $p_n(x)$  on  $[a, b]$  with the weight function  $\rho(x)$  has  $n$  roots on  $(a, b)$ .

### 2.8.1 Legendre polynomial

**Definition 2.52.** The **Legendre polynomial** is defined on  $[-1, 1]$  with the weight function  $\rho(x) = 1$ .

**Theorem 2.53.** The Legendre polynomials  $\{p_i(x), i \in \mathbb{N}\}$  satisfies

$$\int_{-1}^1 p_i(x) p_j(x) dx = \begin{cases} \frac{2}{2i+1}, & i = j \\ 0, & i \neq j. \end{cases}$$

**Theorem 2.54.** The Legendre polynomial  $p_{n-1}, p_n, p_{n+1}$  satisfies

$$p_{n+1}(x) = \frac{2n+1}{n+1} x p_n(x) - \frac{n}{n+1} p_{n-1}(x).$$

**Example 2.55.** The first three terms of Legendre polynomials is

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = \frac{3}{2}x^2 - \frac{1}{2}.$$

### 2.8.2 Chebyshev polynomial of the first kind

**Definition 2.56.** The **Chebyshev polynomial of the first kind** is defined on  $[-1, 1]$  with the weight function  $\rho(x) = \frac{1}{\sqrt{1-x^2}}$ .

**Theorem 2.57.** The Chebyshev polynomials of the first kind  $\{p_i(x), i \in \mathbb{N}\}$  satisfies

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} p_i(x) p_j(x) dx = \begin{cases} \pi & i = j = 0 \\ \frac{\pi}{2} & i = j \neq 0 \\ 0 & i \neq j. \end{cases}$$

**Theorem 2.58.** The Chebyshev polynomial of the first kind  $p_{n-1}, p_n, p_{n+1}$  satisfies

$$p_{n+1}(x) = 2xp_n(x) - p_{n-1}(x).$$

**Example 2.59.** The first three terms of Chebyshev polynomials of the first kind is

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = 2x^2 - 1.$$

### 2.8.3 Chebyshev polynomial of the second kind

**Definition 2.60.** The **Chebyshev polynomial of the second kind** is defined on  $[-1, 1]$  with the weight function  $\rho(x) = \sqrt{1-x^2}$ .

**Theorem 2.61.** The Chebyshev polynomials of the second kind  $\{p_i(x), i \in \mathbb{N}\}$  satisfies

$$\int_{-1}^1 \sqrt{1-x^2} p_i(x) p_j(x) dx = \begin{cases} \frac{\pi}{2}, & i = j \\ 0, & i \neq j. \end{cases}$$

**Theorem 2.62.** The Chebyshev polynomial of the second kind  $p_{n-1}, p_n, p_{n+1}$  satisfies

$$p_{n+1}(x) = 2xp_n(x) - p_{n-1}(x).$$

**Example 2.63.** The first three terms of Chebyshev polynomials of the second kind is

$$p_0(x) = 1, \quad p_1(x) = 2x, \quad p_2(x) = 4x^2 - 1.$$

### 2.8.4 Laguerre polynomial

**Definition 2.64.** The **Laguerre polynomial** is defined on  $[0, +\infty)$  with the weight function  $\rho(x) = x^\alpha e^{-x}$ .

**Theorem 2.65.** The Laguerre polynomial  $\{p_i(x), i \in \mathbb{N}\}$  satisfies

$$\int_0^{+\infty} x^\alpha e^{-x} p_i(x) p_j(x) dx = \begin{cases} \frac{\Gamma(n+\alpha+1)}{n!}, & i = j \\ 0, & i \neq j. \end{cases}$$

**Theorem 2.66.** For  $\alpha = 0$ , the Laguerre polynomial  $p_{n-1}, p_n, p_{n+1}$  satisfies

$$p_{n+1}(x) = (2n+1-x)p_n(x) - n^2 p_{n-1}(x).$$

**Example 2.67.** For  $\alpha = 0$ , the first three terms of Laguerre polynomial is

$$p_0(x) = 1, \quad p_1(x) = -x + 1, \quad p_2(x) = x^2 - 4x + 2.$$

### 2.8.5 Hermite polynomial (probability theory form)



**Definition 2.68.** The **Hermite polynomial** is defined on  $(-\infty, +\infty)$  with the weight function  $\rho(x) = \left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{x^2}{2}}$ .

**Theorem 2.69.** The Hermite polynomial  $\{p_i(x), i \in \mathbb{N}\}$  satisfies

$$\int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} p_i(x) p_j(x) dx = \begin{cases} n!, & i = j \\ 0, & i \neq j. \end{cases}$$

**Theorem 2.70.** For  $\alpha = 0$ , the Hermite polynomial  $p_{n-1}, p_n, p_{n+1}$  satisfies

$$p_{n+1}(x) = xp_n(x) - np_{n-1}(x).$$

**Example 2.71.** For  $\alpha = 0$ , the first three terms of Hermite polynomial is

$$p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2 - 1.$$

## Chapter 3

# Ordinary Differential Equation

**Definition 3.1.** Given a function  $F$ , an **explicit ordinary differential equation** of order  $n$  takes the form

$$\mathbf{F}(\mathbf{u}^{(n-1)}, \dots, \mathbf{u}', \mathbf{u}, t) = \mathbf{u}^{(n)},$$

an **implicit ordinary differential equation** of order  $n$  takes the form

$$\mathbf{F}(\mathbf{u}^{(n)}, \dots, \mathbf{u}', \mathbf{u}, t) = \mathbf{0},$$

**Definition 3.2.** An ODE is **autonomous** if it does not depend on the variable  $x$ .

**Definition 3.3.** A ODE is **linear** if can be written as

$$\sum_{i=0}^n A_i(t) \mathbf{u}^{(i)} + \mathbf{r}(t) = \mathbf{0},$$

where  $A_i(t)$  and  $r(t)$  are continuous functions of  $t$ .

**Definition 3.4.** A linear ODE is **homogeneous** if  $\mathbf{r}(t) = \mathbf{0}$ , and there is always the trivial solution  $\mathbf{u} \equiv \mathbf{0}$ .

**Definition 3.5.** An ODE is **separable** if can be written as

$$P_1(x)Q_1(y) = P_2(x)Q_2(y)\frac{dy}{dx}.$$

**Definition 3.6.** For initial value  $(\mathbf{u}_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ ,  $T \geq t_0$  and  $\mathbf{f} : \mathbb{R}^n \times [t_0, T] \rightarrow \mathbb{R}^n$ , the **initial value problem** (IVP) is to find  $\mathbf{u}(t) \in C^1([t_0, T])$  satisfies

$$\mathbf{u}' = \mathbf{f}(\mathbf{u}, t), \quad \mathbf{u}(t_0) = \mathbf{u}_0.$$

**Theorem 3.7.** Given an IVP, denoted by  $u_0 = u$ ,  $u_i, i = 1, \dots, n$  the  $i$ th derivative of  $u$ , then the ODE

$$\mathbf{F}(\mathbf{u}^{(n-1)}, \dots, \mathbf{u}', \mathbf{u}, t) = \mathbf{u}^{(n)}$$

can be written as an IVP,

$$\begin{pmatrix} \mathbf{u}'_0 \\ \vdots \\ \mathbf{u}'_{n-2} \\ \mathbf{u}'_{n-1} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{n-1} \\ \mathbf{F}(\mathbf{u}_{n-1}, \dots, \mathbf{u}_1, \mathbf{u}_0, t) \end{pmatrix}.$$

### 3.1 General Theory

**Theorem 3.8. (Peano existence theorem)** Given an IVP with an open set  $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}$ , if  $\mathbf{f}(\mathbf{u}, t) \in C(\Omega)$  and  $(\mathbf{u}_0, t_0) \in \Omega$ , then there is a local solution  $\tilde{\mathbf{u}} : U \rightarrow \mathbb{R}^n$  satisfies the IVP, where  $U$  is a neighbourhood of  $t_0$  in  $\mathbb{R}$ .

**Theorem 3.9. (Picard–Lindelöf theorem)** Given an IVP with an open set  $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}$ , if  $\mathbf{f}(\mathbf{u}, t) : \Omega \rightarrow \mathbb{R}^n$  is continuous in  $t$  and Lipschitz continuous in  $\mathbf{u}$  and  $(\mathbf{u}_0, t_0) \in \Omega$ , then there is a unique local solution  $\tilde{\mathbf{u}} : U \rightarrow \mathbb{R}^n$  satisfies the IVP, where  $U$  is a neighbourhood of  $t_0$  in  $\mathbb{R}$ .

**Theorem 3.10. (Comparison theorem)** Given two IVPs

$$\mathbf{u}'_1 = \mathbf{f}_1(\mathbf{u}_1, t), \quad \mathbf{u}_1(t_0) = \mathbf{u}_0,$$

$$\mathbf{u}'_2 = \mathbf{f}_2(\mathbf{u}_2, t), \quad \mathbf{u}_2(t_0) = \mathbf{u}_0,$$

and a open set  $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}$ , if for all  $(\mathbf{u}, t) \in \Omega$ ,  $\mathbf{f}_1(\mathbf{u}, t) < \mathbf{f}_2(\mathbf{u}, t)$ , then

$$\begin{cases} \mathbf{u}_1(t) > \mathbf{u}_2(t), & t > t_0, (\mathbf{u}_1(t), t), (\mathbf{u}_2(t), t) \in \Omega, \\ \mathbf{u}_1(t) < \mathbf{u}_2(t), & t < t_0, (\mathbf{u}_1(t), t), (\mathbf{u}_2(t), t) \in \Omega, \end{cases}$$

## Chapter 4

### Partial Differential Equation

#### 4.1 Poisson Equation

#### 4.2 Heat Equation

#### 4.3 Wave Equation

# Chapter 5

## Probability Theory

**5.1 Discrete random variables**

**5.2 Continous random variables**

**5.3 Characteristic functions**

**5.4 Probability limit theorems**

# Chapter 6

## Stochastic Process

### 6.1 Poisson process

### 6.2 Markov chain

# Chapter 7

## Statistics

# Chapter 8

## Graph

### 8.1 Shortest Path

### 8.2 Matching

### 8.3 Network Flow

### 8.4 Tree



# Chapter 9

## Combinatorics

### 9.1 Generating function

### 9.2 Inclusion–exclusion principle

### 9.3 Special Numbers

#### 9.3.1 Catalan number

#### 9.3.2 Stirling number

# **Part 2**

## **Scientific Computing**

# Chapter 10

## Interpolation

### 10.1 Polynomial Interpolation

#### 10.1.1 Lagrange formula

**Definition 10.1.** To interpolate given points  $(x_0, f(x_0)), \dots, (x_n, f(x_n))$ , the Lagrange formula is

$$p_n(x) = \sum_{i=0}^n f(x_i) l_i(x),$$

where the **elementary Lagrange interpolation polynomial** (or **fundamental polynomial**) for pointwise interpolation  $l_k(x)$  is

$$l_k(x) = \prod_{i=0, i \neq k}^n \frac{x - x_i}{x_k - x_i}.$$

In particular, for  $n = 0$ ,  $l_0(x) = 1$ .

#### 10.1.2 Newton formula

**Definition 10.2.** The  $k$ th divided difference ( $k \in \mathbb{N}^+$ ) on the **table of divided differences**

$$\begin{array}{l|llll} x_0 & f[x_0] & & & \\ x_1 & f[x_1] & f[x_0, x_1] & & \\ x_2 & f[x_2] & f[x_1, x_2] & f[x_0, x_1, x_2] & \\ x_3 & f[x_3] & f[x_2, x_3] & f[x_1, x_2, x_3] & f[x_0, x_1, x_2, x_3] \\ \dots & \dots & \dots & \dots & \dots \end{array}$$

where the **divided differences** satisfy

$$\begin{aligned} f[x_0] &= f(x_0), \\ f[x_0, \dots, x_k] &= \frac{f[x_1, \dots, x_k] - f[x_0, \dots, x_{k-1}]}{x_k - x_0}. \end{aligned}$$

**Corollary 10.3.** Suppose  $(i_0, \dots, i_k)$  is a permutation of  $(0, \dots, k)$ . Then

$$f[x_0, \dots, x_k] = f[x_{i_0}, \dots, x_{i_k}].$$

**Theorem 10.4.** For distinct points  $x_0, \dots, x_n$  and  $x$ , we have

$$f(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, \dots, x_n] \prod_{i=0}^{n-1} (x - x_i) + f[x_0, \dots, x_n, x] \prod_{i=0}^n (x - x_i).$$

**Definition 10.5.** The **Newton formula** for interpolating the points  $(x_0, f(x_0)), \dots, (x_n, f(x_n))$  is

$$p_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + \dots + f[x_0, \dots, x_n] \prod_{i=0}^{n-1} (x - x_i).$$

### 10.1.3 Neville-Aitken algorithm

**Definition 10.6.** Denote  $p_0^{[i]}(x) = f(x_i)$  for  $i = 0, \dots, n$ . For all  $k = 0, \dots, n-1$  and  $i = 0, \dots, n-k-1$ , define

$$p_{k+1}^{[i]}(x) = \frac{(x - x_i)p_k^{[i+1]}(x) - (x - x_{i+k+1})p_k^{[i]}(x)}{x_{i+k+1} - x_i}.$$

Then each  $p_k^{[i]}(x)$  is the interpolating polynomial for the function  $f$  at the points  $x_i, \dots, x_{i+k}$ . In particular,  $p_n^{[0]}(x)$  is the interpolating polynomial of degree  $n$  for the function  $f$  at the points  $x_0, \dots, x_n$ .

### 10.1.4 Hermite interpolation

**Definition 10.7.** Given distinct points  $x_0, \dots, x_k$  in  $[a, b]$ , non-negative integers  $m_0, \dots, m_k$ , and a function  $f \in C^M[a, b]$  where  $M = \max_{i=0, \dots, k} (m_i)$ , the **Hermite interpolation problem** seeks a polynomial  $p(x)$  of the lowest degree satisfies

$$\forall i \in \{0, \dots, k\}, \forall \mu \in \{0, \dots, m_i\}, p^{(\mu)}(x_i) = f^{(\mu)}(x_i).$$

**Definition 10.8. (Generalized divided difference)** Let  $x_0, \dots, x_k$  be  $k+1$  pairwise distinct points with each  $x_i$  repeated  $m_i + 1$  times; write  $N = k + \sum_{i=0}^k m_i$ . The  $N$ th divided difference associated with these points is the coefficient of  $x^N$  in the polynomial  $p$  that uniquely solves the Hermite interpolation problem.

**Corollary 10.9.** The  $n$ th divided difference at  $n+1$  “confluent” (i.e. identical) points is

$$f[x_0, \dots, x_0] = \frac{1}{n!} f^{(n)}(x_0),$$

where  $x_0$  is repeated  $n+1$  times on the left-hand side.

### 10.1.5 Approximation

**Definition 10.10.** Given condition functions  $c_0, \dots, c_k : \mathbb{P}_n \rightarrow \mathbb{R}^+$ , the **Approximation problem** seeks a polynomial  $p_n(x)$  of the given degree  $n$  satisfies a unconstrained optimization

$$\min_{p_n \in \mathbb{P}_n} \sum_{i=0}^k c_i(p_n^{(m_i)}).$$

where condition function  $c(p)$  includes but is not limited to

$$|p^{(m)}(x)|, (p_n^{(m)}(x))^2, \int_a^b |p^{(m)}| dx, \int_a^b (p^{(m)})^2 dx.$$

**Example 10.11.** For non-negative integers  $m_0, \dots, m_k$  and condition functions  $c_i(p_n) = (p_n^{(m_i)}(x))^2$ , denote by

$$p_n(x) = \sum_{i=0}^n c_i x^i$$

the polynomial of the given degree  $n$ , then the  $m$ th derivative of  $p_n$  is

$$p_n^{(m)}(x) = \sum_{i=m}^n \frac{i!}{(i-m)!} c_i x^{i-m}.$$

All above implies the least squares system

$$\begin{cases} p_n^{(m_0)}(x) = \sum_{i=m_0}^n \frac{i!}{(i-m_0)!} c_i x^{i-m_0} = 0, \\ \dots \\ p_n^{(m_k)}(x) = \sum_{i=m_k}^n \frac{i!}{(i-m_k)!} c_i x^{i-m_k} = 0, \end{cases}$$

which can be solved by algorithms such as Householder transformation.

### 10.1.6 Error analysis

**Theorem 10.12.** Let  $f \in C^n[a, b]$  and suppose that  $f^{(n+1)}(x)$  exists at each point of  $(a, b)$ . Let  $p_n(x) \in \mathbb{P}_n$  denote the unique polynomial that coincides with  $f$  at  $x_0, \dots, x_n$ . Define

$$R_n(f; x) = f(x) - p_n(x),$$

as the **Cauchy remainder** of the polynomial interpolation.

If  $a \leq x_0 < \dots < x_n \leq b$ , then there exists some  $\xi \in (a, b)$  satisfies

$$R_n(f; x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i)$$

where the value of  $\xi$  depends on  $x, x_0, \dots, x_n$  and  $f$ .

**Theorem 10.13.** For the Hermite interpolation problem, denote  $N = k + \sum_{i=0}^k m_i$ . Denote by  $p_N(x) \in \mathbb{P}_N$  the unique solution of the problem. Suppose  $f^{(N+1)}(x)$  exists in  $(a, b)$ . Then there exists some  $\xi \in (a, b)$  satisfies

$$R_N(f; x) = \frac{f^{(N+1)}(\xi)}{(N+1)!} \prod_{i=0}^k (x - x_i)^{m_i+1}.$$

## 10.2 Spline

**Definition 10.14.** Given nonnegative integers  $n, k$ , and a strictly increasing sequence  $a = x_1 < \dots < x_N = b$ , the set of **spline** functions of degree  $n$  and smoothness class  $k$  relative to the partition  $\{x_i\}$  is

$$\mathbb{S}_n^k = \left\{ s : s \in C^k[a, b]; \forall i \in \{1, \dots, N-1\}, s|_{[x_i, x_{i+1}]} \in \mathbb{P}_n \right\},$$

where  $x_i$  is the **knot** of the spline.

### 10.2.1 Cubic spline

**Definition 10.15. (Boundary conditions of splines)** The followings are common boundary conditions of cubic splines.

- The **complete cubic spline**  $s$  satisfies  $s'(a) = f'(a), s'(b) = f'(b)$ ;
- The **cubic spline with specified second derivatives**  $s$  satisfies  $s''(a) = f''(a), s''(b) = f''(b)$ ;
- The **natural cubic spline**  $s$  satisfies  $s''(a) = s''(b) = 0$ ;
- The **not-a-knot cubic spline**  $s$  satisfies  $s'''(x)$  exists at  $x = x_2$  and  $x = x_{N-1}$ .
- The **periodic cubic spline**  $s$  satisfies  $s(a) = s(b), s'(a) = s'(b), s''(a) = s''(b)$ .

**Theorem 10.16.** Denote  $m_i = s'(x_i), M_i = s''(x_i)$  for  $s \in \mathbb{S}_3^2$ , then

$$\forall i = 2, 3, \dots, N-1, \quad \lambda_i m_{i-1} + 2m_i + \mu_i m_{i+1} + 1 = 3\mu_i f[x_i, x_{i+1}] + 3\lambda_i f[x_{i-1}, x_i],$$

$$\forall i = 2, 3, \dots, N-1, \quad \mu_i M_{i-1} + 2M_i + \lambda_i m_{i+1} = 6f[x_{i-1}, x_i, x_{i+1}],$$

where

$$\mu_i = \frac{x_i - x_{i-1}}{x_{i+1} - x_{i-1}}, \quad \lambda_i = \frac{x_{i+1} - x_i}{x_{i+1} - x_{i-1}}.$$

In particular,  $m_i$  and  $M_i$  should be replaced to the derivatives given at the boundary.

**Theorem 10.17.** Cubic spline  $s \in \mathbb{S}_3^2$  from the linear system of  $\lambda_i, \mu_i, m_i, M_i$  and the boundary conditions.

### 10.2.2 B-spline

**Definition 10.18.** B-splines are defined recursively by

$$B_i^{n+1}(x) = (x - x_{i-1})(x_{i+n} - x_{i-1})B_i^n(x) + \frac{x_{i+n+1} - x}{x_{i+n+1} - x_{i-1}}B_{i+1}^n(x),$$

where recursion base is the B-spline of degree zero

$$B_i^0(x) = \begin{cases} 1, & x \in (x_{i-1}, x_i], \\ 0, & \text{otherwise.} \end{cases}$$

**Theorem 10.19.** The  $\{B_i^n(x)\}$  forms a basis of  $\mathbb{S}_n^{n-1}$ .

**Definition 10.20.** For  $N \in \mathbb{N}^*$ , the **support** of a  $B_i^n(x)$  is

$$\text{supp } \{B_i^n(x)\} = \overline{\{x \in \mathbb{R} : B_i^n(x) \neq 0\}} = [x_{i-1}, x_{i+n}].$$

**Theorem 10.21. (Integrals of B-splines)** The average of a B-spline over its support only depends on its degree,

$$\frac{1}{t_{i+n} - t_{i-1}} \int_{t_{i-1}}^{t_{i+n}} B_i^n(x) dx = \frac{1}{n+1}.$$

**Theorem 10.22. (Derivatives of B-splines)** For  $n \geq 2$ , we have

$$\forall x \in \mathbb{R}, \quad \frac{d}{dx} B_i^n(x) = \frac{nB_i^{n-1}(x)}{x_{i+n-1} - x_{i-1}} - \frac{nB_{i+1}^{n-1}(x)}{x_{i+n} - x_i}.$$

For  $n = 1$ , it holds for all  $x$  except  $x_{i-1}, t_i, t_{i+1}$ , where the derivative of  $B_i^1(x)$  is not defined.

### 10.2.3 Error analysis

**Theorem 10.23.** Suppose a function  $f \in C^4[a, b]$ , is interpolated by a complete cubic spline or a cubic spline with specified second derivatives at its end points. Then

$$\forall m = 0, 1, 2, |f^{(m)}(x) - s^{(m)}(x)| \leq c_m h^{4-m} \max_{x \in [a, b]} |f^{(4)}(x)|,$$

where  $c_0 = \frac{1}{16}, c_1 = c_2 = \frac{1}{2}$  and  $h = \max_{i=1, \dots, N-1} |x_{i+1} - x_i|$ .

# Chapter 11

## Integration

**Definition 11.1.** A **weighted quadrature formula**  $I_n(f)$  is a linear function

$$I_n(f) = \sum_{i=1}^n w_i f(x_i),$$

which approximates the integral of a function  $f \in C[a, b]$ ,

$$I(f) = \int_a^b \rho(x) f(x) dx,$$

where the weight function  $\rho \in [a, b]$  satisfies  $\forall x \in (a, b), \rho(x) > 0$ . The points  $\{x_i\}$  at which the integrand  $f$  is evaluated are called nodes or abscissas, and the multipliers  $\{w_i\}$  are called weights or coefficients.

**Definition 11.2.** A weighted quadrature formula has (polynomial) **degree of exactness**  $d_E$  iff

$$\begin{aligned} \forall f \in \mathbb{P}_{d_E}, \quad E_n(f) &= 0, \\ \exists g \in \mathbb{P}_{d_E+1}, \quad \text{s.t. } E_n(g) &\neq 0 \end{aligned}$$

where  $\mathbb{P}_d$  denotes the set of polynomials with degree no more than  $d$ .

**Theorem 11.3.** A weighted quadrature formula  $I_n(f)$  satisfies  $d_E \leq 2n - 1$ .

**Definition 11.4.** The **error** or **remainder** of  $I_n(f)$  is

$$E_n(f) = I(f) - I_n(f),$$

where  $I_n(f)$  is said to be convergent for  $C[a, b]$  iff

$$\forall f \in C[a, b], \quad \lim_{n \rightarrow +\infty} E_n(f) = 0.$$

**Theorem 11.5.** Let  $x_1, \dots, x_n$  be given as distinct nodes of  $I_n(f)$ . If  $d_E \geq n - 1$ , then its weights can be deduced as

$$\forall k \in \{1, \dots, n\}, w_k = \int_a^b \rho(x) l_k(x) dx,$$

where  $l_k(x)$  is the elementary Lagrange interpolation polynomial for pointwise interpolation applied to the given nodes.

### 11.1 Newton-Cotes Formulas

**Definition 11.6.** A **Newton-Cotes formula** is a formula based on approximating  $f(x)$  by interpolating it on uniformly spaced nodes  $x_1, \dots, x_n \in [a, b]$ .

#### 11.1.1 Midpoint rule

**Definition 11.7.** The **midpoint rule** is a formula based on approximating  $f(x)$  by the constant  $f\left(\frac{a+b}{2}\right)$ .

For  $\rho(x) \equiv 1$ , it is simply

$$I_M(f) = (b-a)f\left(\frac{a+b}{2}\right).$$

**Theorem 11.8.** For  $f \in C^2[a, b]$ , with weight function  $\rho \equiv 1$ , the error (remainder) of midpoint rule satisfies

$$\exists \xi \in [a, b], \text{ s.t. } E_M(f) = \frac{(b-a)^3}{24} f''(\xi).$$

**Corollary 11.9.** The midpoint rule has  $d_E = 1$ .

### 11.1.2 Trapezoidal rule

**Definition 11.10.** The **trapezoidal rule** is a formula based on approximating  $f(x)$  by the straight line that connects  $(a, f(a))$  and  $(b, f(b))$ .

For  $\rho(x) \equiv 1$ , it is simply

$$I_T(f) = \frac{b-a}{2}(f(a) + f(b)).$$

**Theorem 11.11.** For  $f \in C^2[a, b]$ , with weight function  $\rho \equiv 1$ , the error (remainder) of trapezoidal rule satisfies

$$\exists \xi \in [a, b], \text{ s.t. } E_T(f) = -\frac{(b-a)^3}{12} f''(\xi).$$

**Corollary 11.12.** The trapezoidal rule has  $d_E = 1$ .

### 11.1.3 Simpson's rule

**Definition 11.13.** The **Simpson's rule** is a formula based on approximating  $f(x)$  by the quadratic polynomial that goes through the points  $(a, f(a))$ ,  $\left(\frac{a+b}{2}, f\left(\frac{a+b}{2}\right)\right)$  and  $(b, f(b))$ .

For  $\rho(x) \equiv 1$ , it is simply

$$I_S(f) = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right).$$

**Theorem 11.14.** For  $f \in C^4[a, b]$ , with weight function  $\rho \equiv 1$ , the error (remainder) of Simpson's rule satisfies

$$\exists \xi \in [a, b], \text{ s.t. } E_S(f) = -\frac{(b-a)^5}{2880} f^{(4)}(\xi).$$

**Corollary 11.15.** The Simpson's rule has  $d_E = 3$ .

## 11.2 Gauss Formulas

**Theorem 11.16.** For an interval  $[a, b]$  and a weight function  $\rho : [a, b] \rightarrow \mathbb{R}$ , the nodes for gauss formula  $I_n(f)$  is the root of the  $n$ th order orthogonal polynomial on  $[a, b]$  with the weight function  $\rho(x)$ .

**Theorem 11.17.** A Gauss formula  $I_n(f)$  has  $d_E = 2n - 1$ .



# Chapter 12

## Optimization

### 12.1 One-dimensional Line Search

**Definition 12.1.** Given a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , a initial point  $\mathbf{x}$  and a direction  $\mathbf{d}$ , denoted by  $\varphi(\alpha) = f(\mathbf{x} + \alpha\mathbf{d})$ , a **one-dimensional line search** method solves the problem

$$\varphi(\alpha) = \min_{t \in \mathbb{R}^+} \varphi(t).$$

**Method 12.2. (Success-failure method)** For a one-dimensional line search problem, the **success-failure method** is an inexact one-dimensional line search method to solve the interval  $[a, b] \in [0, +\infty)$  that exact solution  $\alpha^* \in [a, b]$ , where we

- (1) Choose initial value  $\alpha_0 \in [0, +\infty)$ ,  $h_0 > 0$ ,  $t > 0$  (commonly choose  $t = 2$ ), calculate  $\varphi(\alpha_0)$  and let  $k = 0$ ;
- (2) Let  $\alpha_{k+1} = \alpha_k + h_k$  and calculate  $\varphi(\alpha_{k+1})$ , if  $\varphi(\alpha_{k+1}) < \varphi(\alpha_k)$ , then go to (3), otherwise go to (4);
- (3) Let  $h_{k+1} = th_k$ ,  $\alpha = \alpha_k$ ,  $k = k + 1$ , and go to (2);
- (4) If  $k = 0$ , then let  $h_k = -h_k$  and go to (2), otherwise stop and the solution  $[a, b]$  satisfies  $a = \min\{\alpha, \alpha_k\}$ ,  $b = \max\{\alpha, \alpha_k\}$ .

**Definition 12.3.** A general form of one-dimensional line search method is the following three steps:

- (1) **Initialization:** given initial point  $\mathbf{x}$  and acceptable error  $\varepsilon > 0$ ,  $\delta > 0$ ;
- (2) **Iteration:** calculate the direction  $\mathbf{d}$  and step size  $\alpha$  that  $f(\mathbf{x} + \alpha\mathbf{d}) = \min_{t \in \mathbb{R}^+} f(\mathbf{x} + t\mathbf{d})$  and let  $\mathbf{x} = \mathbf{x} + \alpha\mathbf{d}$ ;
- (3) **Stop condition:** if  $\|\nabla f(\mathbf{x})\| \leq \varepsilon$  or  $U_{\mathbb{R}^n}(\mathbf{x}, \delta)$  includes the exact solution, then the current  $\mathbf{x}$  is the solution.

where the iteration step are repeated until  $\mathbf{x}$  satisfies the stop condition.

**Definition 12.4.** Given a method, denoted by  $\{\mathbf{x}_k\}$  the sequence of the iteration and  $\mathbf{x}^*$  the exact solution, the method is **(Q-)linear convergence** if

$$\lim_{k \rightarrow \infty} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} \in (0, 1),$$

the method is **(Q-)sublinear convergence** if

$$\lim_{k \rightarrow \infty} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 1,$$

the method is **(Q-)superlinear convergence** if

$$\lim_{k \rightarrow \infty} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} = 0.$$

For a superlinear convergence method, the method is  $r$ -order linear convergence if

$$\lim_{k \rightarrow \infty} \frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|^r} \in [0, +\infty),$$

where when  $r = 2$  is called **(Q-)quadratic convergence**.

**Remark 12.5.** There is another **R-convergence** for judging a sequence which use another Q-convergence sequence as the boundary of  $\{\|\mathbf{x}_k - \mathbf{x}^*\|\}$ , but is not needed here.

**Method 12.6. (Golden section method)** Given the initial point  $\mathbf{x}$ , an interval  $[a, b]$  and  $\delta > 0$ ,

- The iteration step is:
  - (1) Calculate the two testing points  $\lambda = a + (1 - k)(b - a)$  and  $\mu = a + k(b - a)$  where  $k = \frac{\sqrt{5}-1}{2}$  is the golden ratio;
  - (2) If  $\varphi(\lambda) > \varphi(\mu)$ , let  $a = \lambda$ , otherwise let  $b = \mu$ .
- The stop condition is  $b - a \leq \delta$ ;
- The solution is  $\mathbf{x} + \frac{a+b}{2}\mathbf{d}$ .

**Theorem 12.7.** The golden section method is a **linear convergent** method.

**Method 12.8. (Fibonacci method)** Given the initial point  $\mathbf{x}$ , an interval  $[a, b]$  and  $\delta > 0$ ,

- The  $k$ -th iteration step is:
  - (1) Calculate the two testing points  $\lambda = a + \frac{F_k}{F_{k+2}}(b - a)$  and  $\mu = a + \frac{F_{k+1}}{F_{k+2}}(b - a)$  where  $F_k$  is the  $k$ -th fibonacci number and  $k$ ;
  - (2) If  $\varphi(\lambda) > \varphi(\mu)$ , let  $a = \lambda$ , otherwise let  $b = \mu$ .
- The stop condition is  $b - a \leq \delta$ ;
- The solution is  $\mathbf{x} + \frac{a+b}{2}\mathbf{d}$ .

**Theorem 12.9.** The Fibonacci method is a **linear convergent** method.

**Method 12.10. (Bisection method)** Given the initial point  $\mathbf{x}$ , an interval  $[a, b]$  and  $\delta > 0$ ,

- The iteration step is:
  - (1) Calculate the midpoint  $m = \frac{a+b}{2}$  and  $\varphi(m)$ ;
  - (2) If  $\nabla f(m) \cdot \mathbf{d} < 0$ , let  $a = m$ , otherwise let  $b = m$ .
- The stop condition is  $b - a \leq \delta$ ;
- The solution is  $\mathbf{x} + \frac{a+b}{2}\mathbf{d}$ .

**Theorem 12.11.** The bisection method is a **linear convergent** method.

**Method 12.12. (Newton's method)** Given the initial point  $\mathbf{x}$  and  $\varepsilon > 0$ ,

- The iteration step is:
  - (1) Calculate  $(\nabla^2 f(\mathbf{x}))^T \cdot \mathbf{d}$  and  $(\nabla f(\mathbf{x}))^T \cdot \mathbf{d}$ ;
  - (2) Let  $\mathbf{x} = \mathbf{x} - \frac{(\nabla f(\mathbf{x}))^T \cdot \mathbf{d}}{(\nabla^2 f(\mathbf{x}))^T \cdot \mathbf{d}}$ ;
- The stop condition is  $(\nabla f(\mathbf{x}))^T \cdot \mathbf{d} \leq \varepsilon$ ;
- The solution is  $\mathbf{x}$ .

**Theorem 12.13.** The Newton's method is a **quadratic convergent** method.

## 12.2 Unconstrained Optimization

**Definition 12.14.** Given a convex function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ , a **unconstrained optimization** method solves the problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

by

- (1) **Initialization:** given initial point  $\mathbf{x}$  and acceptable error  $\varepsilon > 0$ ,  $\delta > 0$ ;
- (2) **Iteration:** calculate the direction  $\mathbf{d}$  and step size  $\alpha$ , then let  $\mathbf{x} = \mathbf{x} + \alpha\mathbf{d}$ ;
- (3) **Stop condition:** if  $\|\nabla f(\mathbf{x})\| \leq \varepsilon$  or  $U_{\mathbb{R}^n}(\mathbf{x}, \delta)$  includes the exact solution, then the current  $\mathbf{x}$  is the solution.

**Method 12.15. (Gradient descent method)** Given the initial point  $\mathbf{x}$  and  $\varepsilon > 0$ ,

- The iteration step is:
  - (1) Calculate  $\mathbf{d} = -\nabla f(\mathbf{x})$  and step size  $\alpha$  by a line search method;
  - (2) Let  $\mathbf{x} = \mathbf{x} + \alpha\mathbf{d}$ ;
- The stop condition is  $\|\nabla f(\mathbf{x})\| \leq \varepsilon$ ;
- The solution is  $\mathbf{x}$ .

**Theorem 12.16.** The gradient descent method is a **linear convergent** method.

**Method 12.17. (Quasi-Newton method)** Given the initial point  $\mathbf{x}$ ,  $\varepsilon > 0$  and a matrix  $H \in \mathbb{R}^{n \times n}$  (usually the identity matrix),

- The  $k$ -th iteration step is:
  - (1) Calculate  $\mathbf{d}_k = -H_k \nabla f(\mathbf{x}_k)$  and step size  $\alpha_k$  by a line search method;
  - (2) Let  $\mathbf{x}_k = \mathbf{x}_k + \alpha_k \mathbf{d}_k$  and  $H_{k+1} = r_k(H_k)$  where the function  $r_k$  is a **update** depends on  $\mathbf{x}_k$ ,  $\mathbf{x}_{k+1}$ ,  $\nabla f(\mathbf{x}_k)$  and  $\nabla f(\mathbf{x}_{k+1})$ ;
- The stop condition is  $\|\nabla f(\mathbf{x}_k)\| \leq \varepsilon$ ;
- The solution is  $\mathbf{x}_k$  that satisfies the stop condition.

**Definition 12.18.** Let  $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$  and  $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$ , the **Symmetric Rank-1 update (SR1)** is

$$H_{k+1} = H_k + \frac{(\mathbf{s}_k - H_k \mathbf{y}_k)(\mathbf{s}_k - H_k \mathbf{y}_k)^T}{(\mathbf{s}_k - H_k \mathbf{y}_k)^T \mathbf{y}_k}.$$

The **DFP update** is a rank-2 update defined as

$$H_{k+1} = H_k + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} - \frac{H_k \mathbf{y}_k \mathbf{y}_k^T H_k}{\mathbf{y}_k^T H_k \mathbf{y}_k}.$$

The **BFGS update** is a rank-2 update defined as

$$H_{k+1} = H_k + \left(1 + \frac{\mathbf{y}_k^T H_k \mathbf{y}_k}{\mathbf{s}_k^T \mathbf{y}_k}\right) \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k} - \frac{\mathbf{s}_k \mathbf{y}_k^T H_k + H_k \mathbf{y}_k \mathbf{s}_k^T}{\mathbf{s}_k^T \mathbf{y}_k}.$$

**Theorem 12.19.** The Quasi-Newton method is a **superlinear convergent** method.

**Method 12.20. (Newton's method)** Given the initial point  $\mathbf{x}$  and  $\varepsilon > 0$ ,

- The iteration step is:
  - (1) Calculate  $\mathbf{d} = -(\nabla^2 f(\mathbf{x}))^{-1} \nabla f(\mathbf{x})$  and step size  $\alpha$  by a line search method;
  - (2) Let  $\mathbf{x} = \mathbf{x} + \alpha\mathbf{d}$ ;
- The stop condition is  $\|\nabla f(\mathbf{x})\| \leq \varepsilon$ ;
- The solution is  $\mathbf{x}$ .

**Theorem 12.21.** The Newton's method is a **quadratic convergent** method.

# Chapter 13

## Initial Value Problem

**Notation 13.1.** To numerically solve the IVP, we are given initial condition  $\mathbf{u}_0 = \mathbf{u}(t_0)$ , and want to compute approximations  $\{\mathbf{u}_k, k = 1, 2, \dots\}$  such that

$$\mathbf{u}_k \approx \mathbf{u}(t_k),$$

where  $k$  is the uniform time step size and  $t_n = nk$ .

### 13.1 Linear Multistep Method

**Definition 13.2.** For solving the IVP, an  $s$ -step **linear multistep method** (LMM) has the form

$$\sum_{j=0}^s \alpha_j \mathbf{u}_{n+j} = k \sum_{j=0}^s \beta_j \mathbf{f}(\mathbf{u}_{n+j}, t_{n+j}),$$

where  $\alpha_s = 1$  is assumed WLOG.

**Definition 13.3.** An LMM is **explicit** if  $\beta_s = 0$ , otherwise it is **implicit**.

### 13.2 Runge-Kutta Method

**Definition 13.4.** An  $s$ -stage **Runge-Kutta method** (RK) is a one-step method of the form

$$\begin{aligned} \mathbf{y}_i &= \mathbf{f} \left( \mathbf{u}_n + k \sum_{j=1}^s a_{ij} \mathbf{y}_j, t_n + c_i k \right), \\ \mathbf{u}_{i+1} &= \mathbf{u}_i + k \sum_{j=1}^s b_j \mathbf{y}_j, \end{aligned}$$

where  $i = 1, \dots, s$  and  $a_{ij}, b_j, c_i \in \mathbb{R}$ .

**Definition 13.5.** The textsf{Butcher tableau} is one way to organize the coefficients of an RK method as follows

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array}$$

The matrix  $A = (a_{ij})_{s \times s}$  is called the RK matrix and  $\mathbf{b} = (b_1, \dots, b_s)^T$ ,  $\mathbf{c} = (c_1, \dots, c_s)^T$  are called the RK weights and the RK nodes.

**Definition 13.6.** An  $s$ -stage **collocation method** is a numerical method for solving the IVP, where we

- (1) choose  $s$  distinct collocation parameters  $c_1, \dots, c_s$ ,
- (2) seek  $s$ -degree polynomial  $p$  satisfying
 
$$\forall i = 1, 2, \dots, s, \quad \mathbf{p}(t_n) = \mathbf{u}_n \quad \text{and} \quad \mathbf{p}'(t_n + c_i k) = \mathbf{f}(\mathbf{p}(t_n + c_i k), t_n + c_i k),$$
- (3) set  $\mathbf{u}_{n+1} = \mathbf{p}(t_{n+1})$ .

**Theorem 13.7.** The s-stage collocation method is an s-stage IRK method with

$$a_{ij} = \int_0^{c_i} l_j(\tau) d\tau, \quad b_j = \int_0^1 l_j(\tau) d\tau,$$

where  $i, j = 1, \dots, s$  and  $l_k(\tau)$  is the elementary Lagrange interpolation polynomial.

### 13.3 Theoretical analysis

**Definition 13.8.** A function  $\mathbf{f} : \mathbb{R}^n \times [0, +\infty) \rightarrow \mathbb{R}^n$  is **Lipschitz continuous** in its first variable over some domain

$$\Omega = \{(\mathbf{u}, t) : \|\mathbf{u} - \mathbf{u}_0\| \leq a, t \in [0, T]\}$$

iff

$$\exists L \geq 0, \text{ s.t. } \forall (\mathbf{u}, t) \in \Omega, \quad \|\mathbf{f}(\mathbf{u}, t) - \mathbf{f}(\mathbf{v}, t)\| \leq L \|\mathbf{u} - \mathbf{v}\|.$$

#### 13.3.1 Error analysis

**Definition 13.9.** The **local truncation error**  $\tau$  is the error caused by replacing continuous derivatives with numerical formulas.

**Definition 13.10.** A numerical formulas is **consistent** if  $\lim_{k \rightarrow 0} \tau = 0$ .

#### 13.3.2 Stability

**Definition 13.11.** The **region of absolute stability** (RAS) of a numerical method, applied to

$$\mathbf{u}' = \lambda \mathbf{u}, \quad \mathbf{u}_0 = \mathbf{u}(t_0),$$

is the region  $\Omega$  that

$$\forall \mathbf{u}_0, \quad \forall \lambda k \in \Omega, \quad \lim_{n \rightarrow +\infty} \mathbf{u}_n = 0.$$

**Definition 13.12.** The **stability function** of a one-step method is a function  $R : \mathbb{C} \rightarrow \mathbb{C}$  that satisfies

$$\mathbf{u}_{n+1} = R(z) \mathbf{u}_n$$

for the  $\mathbf{u}' = \lambda \mathbf{u}$  where  $\text{Re}(E(\lambda)) \leq 0$  and  $z = k\lambda$ .

**Definition 13.13.** A numerical method is **stable** or **zero stable** iff its application to any IVP with  $\mathbf{f}(\mathbf{u}, t)$  Lipschitz continuous in  $\mathbf{u}$  and continuous in  $t$  yields

$$\forall T > 0, \quad \lim_{k \rightarrow 0, Nk=t} \|\mathbf{u}_n\| < \infty.$$

**Definition 13.14.** A numerical method is **A( $\alpha$ )-stable** if the region of absolute stability  $\Omega$  satisfies

$$\{z \in \mathbb{C} : \pi - \alpha \leq \arg(z) \leq \pi + \alpha\} \subseteq \Omega.$$

**Definition 13.15.** A numerical method is **A-stable** if the region of absolute stability  $\Omega$  satisfies

$$\{z \in \mathbb{C} : \text{Re}(z) \leq 0\} \subseteq \Omega.$$

**Definition 13.16.** A one-step method is **L-stable** if it is A-stable, and its stability function satisfies

$$\lim_{z \rightarrow \infty} |R(z)| = 0.$$

**Definition 13.17.** An one-step method is **I-stable** iff its stability function satisfies

$$\forall y \in \mathbb{R}, |R(yi)| \leq 1.$$

**Definition 13.18.** An one-step method is **B-stable** (or **contractive**) if for any contractive ODE system, every pair of its numerical solutions  $\mathbf{u}_n$  and  $\mathbf{v}_n$  satisfy

$$\forall n \in \mathbb{N}, \|u_{n+1} - v_{n+1}\| \leq \|u_n - v_n\|.$$

**Definition 13.19.** An RK method is **algebraically stable** iff the RK weights  $b_1, \dots, b_s$  are nonnegative, the **algebraic stability matrix**  $M = (b_i a_{ij} + b_i a_{ji} - b_i b_j)_{s \times s}$  is positive semidefinite.

**Theorem 13.20.** The order of accuracy of an implicit A-stable LMM satisfies  $p \leq 2$ . An explicit LMM cannot be A-stable.

**Theorem 13.21.** No ERK method is A-stable.

**Theorem 13.22.** An RK method is A-stable if and only if it is I-stable and all poles of its stability function  $R(z)$  have positive real parts.

**Theorem 13.23.** If an A-stable RK method with a nonsingular RK matrix  $A$  is stiffly accurate, then it is L-stable.

**Theorem 13.24.** If an A-stable RK method with a nonsingular RK matrix  $A$  satisfies

$$\forall i \in \{1, \dots, s\}, \quad a_{i1} = b_i,$$

then it is L-stable.

**Theorem 13.25.** B-stable one-step methods are A-stable.

**Theorem 13.26.** An algebraically stable RK method is B-stable and A-stable.

### 13.3.3 Convergence

**Definition 13.27.** A numerical method is convergent iff its application to any IVP with  $\mathbf{f}(\mathbf{u}, t)$  Lipschitz continuous in  $\mathbf{u}$  and continuous in  $t$  yields

$$\forall T > 0, \quad \lim_{k \rightarrow \infty, nk=T} \mathbf{u}_n = \mathbf{u}(T).$$

**Theorem 13.28.** A numerical method is convergent iff it is consistent and stable.

## 13.4 Important Methods

### 13.4.1 Forward Euler's method

**Definition 13.29.** The **forward Euler's method** solves the IVP by

$$\mathbf{u}_{n+1} = \mathbf{u}_n + k\mathbf{f}(\mathbf{u}_n, t_n).$$

**Theorem 13.30.** The region of absolute stability for forward Euler's method is

$$\{z \in \mathbb{C} : |1 + z| \leq 1\}.$$

### 13.4.2 Backward Euler's method

**Definition 13.31.** The **backward Euler's method** solves the IVP by

$$\mathbf{u}_{n+1} = \mathbf{u}_n + k\mathbf{f}(\mathbf{u}_{n+1}, t_{n+1}).$$

**Theorem 13.32.** The region of absolute stability for backward Euler's method is

$$\{z \in \mathbb{C} : |1 - z| \geq 1\}.$$

### 13.4.3 Trapezoidal method

**Definition 13.33.** The **trapezoidal method** solves the IVP by

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \frac{k}{2}(\mathbf{f}(\mathbf{u}_n, t_n) + \mathbf{f}(\mathbf{u}_{n+1}, t_{n+1})).$$

**Theorem 13.34.** The region of absolute stability for trapezoidal method is

$$\left\{ z \in \mathbb{C} : \left| \frac{2+z}{2-z} \right| \geq 1 \right\}.$$

### 13.4.4 Midpoint method (Leapfrog method)

**Definition 13.35.** The **midpoint method (Leapfrog method)** solves the IVP by

$$\mathbf{u}_{n+1} = \mathbf{u}_{n-1} + 2k\mathbf{f}(\mathbf{u}_n, t_n).$$

**Theorem 13.36.** The region of absolute stability for midpoint method is

$$\left\{ z \in \mathbb{C} : \left| z \pm \sqrt{1 + z^2} \right| \leq 1 \right\} \stackrel{?}{=} \{0\}.$$

### 13.4.5 Heun's third-order RK method

**Definition 13.37.** The **Heun's third-order formula** is an ERK method of the form

$$\begin{cases} \mathbf{y}_1 &= \mathbf{f}(\mathbf{u}_n, t_n), \\ \mathbf{y}_2 &= \mathbf{f}\left(\mathbf{u}_n + \frac{k}{3}\mathbf{y}_1, t_n + \frac{k}{3}\right), \\ \mathbf{y}_3 &= \mathbf{f}\left(\mathbf{u}_n + \frac{2k}{3}\mathbf{y}_2, t_n + \frac{2k}{3}\right), \\ \mathbf{u}_{n+1} &= \mathbf{u}_n + \frac{k}{4}(\mathbf{y}_1 + 3\mathbf{y}_3). \end{cases} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ \hline \frac{1}{4} & 0 & \frac{3}{4} & \end{array}$$

### 13.4.6 Classical fourth-order RK method

**Definition 13.38.** The **classical fourth-order RK method** is an ERK method of the form

$$\left\{ \begin{array}{l} \mathbf{y}_1 = \mathbf{f}(\mathbf{u}_n, t_n), \\ \mathbf{y}_2 = \mathbf{f}\left(\mathbf{u}_n + \frac{k}{2}\mathbf{y}_1, t_n + \frac{k}{2}\right), \\ \mathbf{y}_3 = \mathbf{f}\left(\mathbf{u}_n + \frac{k}{2}\mathbf{y}_2, t_n + \frac{k}{2}\right), \\ \mathbf{y}_4 = \mathbf{f}(\mathbf{u}_n + k\mathbf{y}_3, t_n + k), \\ \mathbf{u}_{n+1} = \mathbf{u}_n + \frac{k}{6}(\mathbf{y}_1 + 2\mathbf{y}_2 + 2\mathbf{y}_3 + \mathbf{y}_4). \end{array} \right. \quad \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

### 13.4.7 TR-BDF2 method

**Definition 13.39.** The **TR-BDF2 method** is an one-step method of the form

$$\left\{ \begin{array}{l} \mathbf{u}_* = \mathbf{u}_n + \frac{k}{4}(\mathbf{f}(\mathbf{u}_n, t_n) + \mathbf{f}(\mathbf{u}_*, t_n + \frac{k}{2})), \\ \mathbf{u}_{n+1} = \frac{1}{3}(4\mathbf{u}_* - \mathbf{u}_n + k\mathbf{f}(\mathbf{u}_{n+1}, t_{n+1})). \end{array} \right.$$



# Chapter 14

## Number Theory

### 14.1 Prime Number

**Definition 14.1.** A **prime number** (or a **prime**) is a natural number greater than 1 that is not a product of two smaller natural numbers.

**Definition 14.2.** A **composite number** (or a **composite**) is a natural number greater than 1 that is a product of two smaller natural numbers.

#### 14.1.1 Primality testing

**Theorem 14.3.** For a integer  $n \in \mathbb{N}$ , if it is a product of two natural number  $a$  and  $b$  that  $a \leq b$ , then

$$1 \leq a \leq \sqrt{n} \leq b \leq n.$$

**Method 14.4. (Trial division)** Given a integer  $n$ , the **trial division method** divides  $n$  by each integer from 2 up to  $\sqrt{n}$ . Any such integer dividing  $n$  evenly establishes  $n$  as composite, otherwise it is prime.

**Theorem 14.5. (Fermat's little theorem)** For a prime number  $p$  and a number  $a$  that  $\gcd(a, p) = 1$ , then  $a^{p-1} \equiv 1 \pmod{p}$

**Method 14.6.** The **Miller-Rabin** algorithm is a method of primality testing, where given a number  $n$ , where we

- (1) determine directly for small numbers such as  $p = 2$ .
- (2) factorize the number  $p = u \times 2^t$ ;
- (3) choose a number  $a$  that  $\gcd(a, p) = 1$ , and calculate  $a^u, a^{u \times 2}, a^{u \times 2^2}, \dots, a^{u \times 2^{t-1}}$ ;
- (4) if  $a^u \equiv 1 \pmod{p}$ , or  $\exists a^{u \times k}, k < t$  that  $a^{u \times k} \equiv p - 1 \pmod{p}$  then  $p$  passes the test, otherwise,  $p$  is a composite number;
- (5) repeat above steps to eliminate error.

For numbers less than  $2^{32}$ , choose  $a \in \{2, 7, 61\}$  is enough, for numbers less than  $2^{64}$ , choose  $a \in \{2, 325, 9375, 28178, 450775, 9780504, 1795265022\}$  is enough.

#### 14.1.2 Sieves

**Method 14.7. (Sieve of Eratosthenes)** Given a upper limit  $n$ , the **sieve of Eratosthenes** solves all the prime numbers up to  $n$  by marking composite numbers, where we

- (1) create a list of consecutive integers from 2 to  $n$ :  $\{2, 3, 4, \dots, n\}$ ;
- (2) initially, let  $p = 2$ , the smallest prime number;
- (3) enumerate the multiples of  $p$  by counting in increments of  $p$  from  $2p$  to  $n$ , and mark them in the list;
- (4) find the smallest number in the list greater than  $p$  that is not marked;
- (5) if there was no such number, the method is terminated and the numbers remaining not marked in the list are all the primes below  $n$ , otherwise let  $p$  now equal the new number which is the next prime, and repeat from step (3).

# **Part 3**

## **Machine Learning**

# Chapter 15

## Regression

### 15.1 Linear Regression

**Definition 15.1.** Given a data set  $\{(\mathbf{x}_i, y_i), i \in \{1, \dots, m\}\}$  where  $\mathbf{x}_i \in \mathbb{R}^n$ , the linear regression seeks  $\tilde{\mathbf{w}} \in \mathbb{R}^n$  and  $\tilde{b} \in \mathbb{R}$  such that

$$f(\mathbf{x}_i) = \tilde{\mathbf{w}}^T \mathbf{x}_i + \tilde{b} \approx y_i.$$

In general, we choose mean square error to estimate the error between  $f(\mathbf{x}_i)$  and  $y_i$ , which implies

$$(\tilde{\mathbf{w}}, \tilde{b}) = \arg \min_{\mathbf{w} \in \mathbb{R}^n, b \in \mathbb{R}} \sum_{i=1}^m (f(\mathbf{x}_i) - y_i)^2 = \arg \min_{\mathbf{w} \in \mathbb{R}^n, b \in \mathbb{R}} \sum_{i=1}^m (\mathbf{w}^T \mathbf{x}_i + b - y_i)^2.$$

**Theorem 15.2.** Given a data set  $\{(\mathbf{x}_i, y_i), i \in \{1, \dots, m\}\}$  where  $\mathbf{x}_i \in \mathbb{R}^n$ , let

$$X = \begin{pmatrix} \mathbf{x}_1^T & 1 \\ \vdots & 1 \\ \mathbf{x}_m^T & 1 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

if  $X^T X$  is invertible, the solution of linear regression can be written as

$$\begin{pmatrix} \mathbf{w} \\ b \end{pmatrix} = (X^T X)^{-1} X^T \mathbf{y}.$$

## Chapter 16

### Decision Tree

## Chapter 17

# Support Vector Machine

# Chapter 18

## Cluster

# Chapter 19

## Neural Networks