# 你永远不知道一个强迫症能干出什么事情

倪兴程[1]

2024 年 10 月 18 日

[1]Email: 19975022383@163.com

# Todo list

# 目录

# Chapter 1

# 概率测度

## 1.1 集合

**Theorem 1.1** (De-Morgan law)**.** 设 $\{A_n,\ n \in I\}$ 是一个集族，则有如下 *De-Morgan law*：

$$\left(\underset{n\in I}{\cup} A_n\right)^c = \underset{n\in I}{\cap} A_n^c, \quad \left(\underset{n\in I}{\cap} A_n\right)^c = \underset{n\in I}{\cup} A_n^c$$

*Proof.* 对于第一个等式有：

$$x \in \left(\underset{n\in I}{\cup} A_n\right)^c \Leftrightarrow \forall\, n \in I,\ x \notin A_n \Leftrightarrow x \in \underset{n\in I}{\cap} A_n^c$$

对于第二个等式有：

$$x \in \left(\underset{n\in I}{\cap} A_n\right)^c \Leftrightarrow \exists\, n \in I,\ x \notin A_n \Leftrightarrow x \in \underset{n\in I}{\cup} A_n^c$$

□

**Definition 1.1.** 设 $\{A_n, n \in \mathbb{N}^+\}$ 为一个集合序列，

1. 若 $A_n \subset A_{n+1},\ \forall\, n \in \mathbb{N}^+$，则称 $\{A_n\}$ 为单调递增的集合序列，记为 $A_n \uparrow$，定义：

$$\lim_{n\to+\infty} A_n = \overset{+\infty}{\underset{n=1}{\cup}} A_n$$

2. 若 $A_n \supset A_{n+1},\ \forall\, n \in \mathbb{N}^+$，则称 $\{A_n\}$ 为单调递减的集合序列，记为 $A_n \downarrow$，定义：

$$\lim_{n\to+\infty} A_n = \overset{+\infty}{\underset{n=1}{\cap}} A_n$$

单调递增和单调递减的集合序列统称为单调的集合序列。

**Definition 1.2.** 设 $\{A_n, n \in \mathbb{N}^+\}$ 为一个集合序列，其上下极限定义如下：

$$\underline{\lim_{n\to+\infty}} A_n = \overset{+\infty}{\underset{n=1}{\cup}}\ \overset{+\infty}{\underset{k=n}{\cap}} A_k, \quad \overline{\lim_{n\to+\infty}} A_n = \overset{+\infty}{\underset{n=1}{\cap}}\ \overset{+\infty}{\underset{k=n}{\cup}} A_k$$

若：

$$\underline{\lim_{n\to+\infty}} A_n = \overline{\lim_{n\to+\infty}} A_n$$

则认为 $\{A_n, n\in\mathbb{N}^+\}$ 极限存在，记：

$$\lim_{n\to+\infty} A_n = \varliminf_{n\to+\infty} A_n = \varlimsup_{n\to+\infty} A_n$$

**Theorem 1.2.** 设 $\{A_n, n\in\mathbb{N}^+\}$ 为一个集合序列，其上下极限具有如下等价定义：

$$\varliminf_{n\to+\infty} A_n = \{x : \exists\, N\in\mathbb{N}^+,\ \forall\, n\geqslant N,\ x\in A_n\}$$
$$\varlimsup_{n\to+\infty} A_n = \{x : \forall\, N\in\mathbb{N}^+,\ \exists\, n\geqslant N,\ x\in A_n\}$$

*Proof.* 对于下极限来讲：

$$x\in \varliminf_{n\to+\infty} A_n \Leftrightarrow \exists\, n\in\mathbb{N}^+,\ x\in \bigcap_{k=n}^{+\infty} A_k \Leftrightarrow \exists\, N\in\mathbb{N}^+,\ \forall\, n\geqslant N,\ x\in A_n$$

对于上极限来讲：

$$x\in \varlimsup_{n\to+\infty} A_n \Leftrightarrow \forall\, n\in\mathbb{N}^+,\ x\in \bigcup_{k=n}^{+\infty} A_k \Leftrightarrow \forall\, N\in\mathbb{N}^+,\ \exists\, n\geqslant N,\ x\in A_n$$ □

**Theorem 1.3.** 设 $\{A_n, n\in\mathbb{N}^+\}$ 为一个集合序列，则其下极限包含于上极限，即：

$$\varliminf_{n\to+\infty} A_n \subseteq \varlimsup_{n\to+\infty} A_n$$

*Proof.* 利用上下极限的等价定义可直接得出结论。 □

### 1.1.1 重要集族

#### $\pi$ 系、半环、环、域

**Definition 1.3.** 如果 $X$ 上的非空集族 $\mathscr{A}$ 对交的运算是封闭的，即：

$$\forall\, A, B\in\mathscr{A},\ A\cap B\in\mathscr{A}$$

则称 $\mathscr{A}$ 是一个 $\pi$ 系。

**Definition 1.4.** 如果 $X$ 上的非空集族 $\mathscr{A}$ 满足：

*1.* 对交的运算封闭；

*2.* 若 $A, B\in\mathscr{A}$ 且 $B\subseteq A$，则存在有限个两两不交的 $\{C_i\in\mathscr{A} : i=1,2,\ldots,n\}$，使得：

$$A\backslash B = \bigcup_{i=1}^{n} C_i$$

则称 $\mathscr{A}$ 为半环 *(semiring)*。

**Definition 1.5.** 如果 $X$ 上的非空集族 $\mathscr{A}$ 对并和差的运算是封闭的，即对任意的 $A, B\in\mathscr{A}$：

*1.* $A\cup B\in\mathscr{A}$；

2. $A\backslash B \in \mathscr{A}$。

则称 $\mathscr{A}$ 为环 *(ring)*。

**Definition 1.6.** 如果 $X$ 上的非空集族 $\mathscr{A}$ 对交和补的运算是封闭的，且 $X$ 也在其中，即：

1. $\forall\, A, B \in \mathscr{A}$，$A \cap B \in \mathscr{A}$；

2. $\forall\, A \in \mathscr{A}$，$A^c \in \mathscr{A}$；

3. $X \in \mathscr{A}$。

则称 $\mathscr{A}$ 为域 *(field of sets)* 或代数 *(algebra of sets)*。

**Theorem 1.4.** 半环必是 $\pi$ 系，环必是半环，域必是环。

*Proof.* (1) 半环必是 $\pi$ 系可直接由半环的定义得出。

(2) 设 $\mathscr{A}$ 是一个环，$A, B \in \mathscr{A}$，由集合的运算可得：

$$A \cap B = (A \cup B) \setminus [(A\backslash B) \cup (B\backslash A)]$$

因为 $\mathscr{A}$ 是一个环，所以 $A\backslash B, B\backslash A \in \mathscr{A}$，$[(A\backslash B) \cup (B\backslash A)] \in \mathscr{A}$，$A \cup B \in \mathscr{A}$，所以 $A \cap B = (A \cup B) \setminus [(A\backslash B) \cup (B\backslash A)] \in \mathscr{A}$，即 $\mathscr{A}$ 对交的运算是封闭的。

因为 $\mathscr{A}$ 是一个环，对差的运算封闭，所以取 $C = A\backslash B$ 即有 $A\backslash B = C \in \mathscr{A}$。

(3) 设 $\mathscr{A}$ 是一个域，$A, B \in \mathscr{A}$，由集合的运算可得：

$$A \cup B = (A^c \cap B^c)^c \in \mathscr{A}$$
$$A\backslash B = A \cap B^c \in \mathscr{A}$$

即 $\mathscr{A}$ 是一个环。 □

### 单调系、$\lambda$ 系、$\sigma$ 环、$\sigma$ 域

**Definition 1.7.** 如果集族 $\mathscr{A}$ 中的所有单调序列 $\{A_n\}$ 都有 $\lim\limits_{n\to+\infty} A_n \in \mathscr{A}$，则称 $\mathscr{A}$ 为单调系 *(monotone class)*。

**Definition 1.8.** 如果 $X$ 上的集族 $\mathscr{A}$ 满足：

1. $X \in \mathscr{A}$；

2. 若 $A, B \in \mathscr{A}$，$B \subseteq A$，则有 $A\backslash B \in \mathscr{A}$；

3. 单调递增集合序列 $\{A_n\}$ 的极限 $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{A}$。

则称 $\mathscr{A}$ 为 $\lambda$ 系。

**Property 1.1.1.** $\lambda$ 系对补封闭。

*Proof.* 设 $\mathscr{A}$ 是一个 $\lambda$ 系，任取 $A \in \mathscr{A}$，由 $\lambda$ 系的定义可知 $\mathscr{A}$ 对差封闭，所以有 $A^c = X \backslash A \in \mathscr{A}$，即 $\lambda$ 系对补封闭。 □

**Definition 1.9.** 如果 $X$ 上的集族 $\mathscr{A}$ 满足：

1. 若 $A, B \in \mathscr{A}$，则 $A \backslash B \in \mathscr{A}$；
2. 若 $A_n \in \mathscr{A}, \forall n \in \mathbb{N}^+$，则 $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{A}$。

则称 $\mathscr{A}$ 为 $\sigma$ 环。

**Definition 1.10.** 如果 $X$ 上的集族 $\mathscr{A}$ 满足：

1. $X \in \mathscr{A}$；
2. 若 $A \in \mathscr{A}$，则 $A^c \in \mathscr{A}$；
3. 若 $A_n \in \mathscr{A}, \forall n \in \mathbb{N}^+$，则 $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{A}$。

则称 $\mathscr{A}$ 为 $\sigma$ 域。

**Theorem 1.5.** $\sigma$ 域是域。

*Proof.* 设 $\mathscr{A}$ 是一个 $\sigma$ 域，$A, B \in \mathscr{A}$。由 $\sigma$ 域的定义，$\mathscr{A}$ 对补的运算封闭并且 $X \in \mathscr{A}$。因为：

$$A \cap B = (A^c \cup B^c \cup \varnothing \cup \cdots)^c$$

由 $\sigma$ 域的定义，$A \cap B \in \mathscr{A}$。综上，$\mathscr{A}$ 是一个域。 □

**Theorem 1.6.** $\lambda$ 系是单调系，$\sigma$ 域是 $\lambda$ 系。

*Proof.* (1) 设 $\mathscr{A}$ 是一个 $\lambda$ 系。由 $\lambda$ 系的定义，$\mathscr{A}$ 中单调递增的集合序列必在 $\mathscr{A}$ 中有极限。任取 $\mathscr{A}$ 中的单调递减序列 $\{A_n\}$，由性质 1.1.1可知 $\{A_n^c\}$ 是 $\mathscr{A}$ 中的一个单调递增序列，于是有：

$$\lim_{n \to +\infty} A_n^c = \bigcup_{n=1}^{+\infty} A_n^c \in \mathscr{A}$$

所以：

$$\bigcap_{n=1}^{+\infty} A_n = \left(\bigcup_{n=1}^{+\infty} A_n^c\right)^c = X \setminus \left(\bigcup_{n=1}^{+\infty} A_n^c\right) \in \mathscr{A}$$

即 $\{A_n\}$ 在 $\mathscr{A}$ 中有极限。由 $\{A_n\}$ 的任意性，$\mathscr{A}$ 中单调递减的集合序列也必在 $\mathscr{A}$ 中有极限。综上，$\mathscr{A}$ 是一个单调系。由 $\mathscr{A}$ 的任意性，$\lambda$ 系是单调系。

(2) 因为 $\sigma$ 域是域，所以 $\sigma$ 域是环，因此对差的运算封闭。显然 $\sigma$ 域满足 $\lambda$ 系定义中的 (1) 和 (3)。 □

### 集族的关系总结

上面提到的集族之间有如下关系：

$$\sigma\text{ 域} \subset \text{域} \subset \text{环} \subset \text{半环} \subset \pi\text{ 系}$$

$$\sigma\text{ 域} \subset \lambda\text{ 系} \subset \text{单调系}$$

**Theorem 1.7.** 一个对可列并运算封闭的环是 $\sigma$ 环。

*Proof.* 环对差的运算封闭。 □

**Theorem 1.8.** 一个包含 $X$ 的环是域。

*Proof.* 设 $\mathscr{A}$ 是一个环且 $X \in \mathscr{A}$。任取 $A, B \in \mathscr{A}$ 可得：

$$A \cap B = (A \cup B) \backslash [(A \backslash B) \cup (B \backslash A)]$$

因为 $\mathscr{A}$ 是一个环，所以 $A \backslash B \in \mathscr{A}$，$B \backslash A \in \mathscr{A}$，$(A \backslash B) \cup (B \backslash A) \in \mathscr{A}$，$A \cup B \in \mathscr{A}$，于是 $A \cap B \in \mathscr{A}$。由 $A, B$ 的任意性，$\mathscr{A}$ 对交的运算封闭。

因为 $A^c = X \backslash A$，所以 $A^c \in \mathscr{A}$。由 $A$ 的任意性，$\mathscr{A}$ 对补的运算封闭。

综上，$\mathscr{A}$ 是一个域，即一个包含 $X$ 的环是域。 □

**Theorem 1.9.** 一个既是单调系又是域的集族必是 $\sigma$ 域。

*Proof.* 设 $\mathscr{A}$ 既是单调系又是域。因为 $\mathscr{A}$ 是一个域，所以对补的运算封闭且 $X \in \mathscr{A}$。任取 $A_n \in \mathscr{A}$，$n \in \mathbb{N}^+$，因为域是环，所以 $\mathscr{A}$ 对有限并封闭，即 $\bigcup\limits_{i=1}^{n} A_i \in \mathscr{A}$，$\forall\, n \in \mathbb{N}^+$。因为 $\mathscr{A}$ 是单调系，所以单调递增集合序列：

$$\left\{B_n = \bigcup_{i=1}^{n} A_i\right\}$$

的极限：

$$\lim_{n\to+\infty} B_n = \bigcup_{n=1}^{+\infty} B_n = \bigcup_{n=1}^{+\infty} \bigcup_{i=1}^{n} A_i = \bigcup_{n=1}^{+\infty} A_n \in \mathscr{A}$$

由 $\{A_n\}$ 的任意性，$\mathscr{A}$ 对可列并封闭。综上，$\mathscr{A}$ 是一个 $\sigma$ 域，即一个既是单调系又是域的集族必是 $\sigma$ 域。 □

**Theorem 1.10.** 一个既是 $\lambda$ 系又是 $\pi$ 系的集族必是 $\sigma$ 域。

*Proof.* 设 $\mathscr{A}$ 既是 $\lambda$ 系又是 $\pi$ 系。因为 $\mathscr{A}$ 是 $\lambda$ 系，所以 $X \in \mathscr{A}$。任取 $A \in \mathscr{A}$，则 $A^c = X \backslash A$，由 $\lambda$ 系定义的第二个条件，$A^c \in \mathscr{A}$。由 $A$ 的任意性，$\mathscr{A}$ 对补的运算封闭。又因为 $\mathscr{A}$ 是一个 $\pi$ 系，所以 $\mathscr{A}$ 对交的运算封闭。综上可知 $\mathscr{A}$ 是一个域。因为 $\lambda$ 系是单调系，所以 $\mathscr{A}$ 既是域又是单调系，于是 $\mathscr{A}$ 是 $\sigma$ 域。 □

**Theorem 1.11.** 一个包含 $X$ 的 $\sigma$ 环是 $\sigma$ 域。

*Proof.* 设 $\mathscr{A}$ 是 $\sigma$ 域，$A \in \mathscr{A}$。因为 $X \in \mathscr{A}$，$A^c = X \backslash A$，而 $\sigma$ 环对差的运算封闭，所以 $\mathscr{A}$ 对补的运算封闭。又因为 $\sigma$ 环对可列并封闭，所以 $\mathscr{A}$ 是 $\sigma$ 域，即一个包含 $X$ 的 $\sigma$ 环是 $\sigma$ 域。 □

### 1.1.2 集族的生成

**Definition 1.11.** 设 $\mathscr{A},\mathscr{B}$ 是 $X$ 上的集族，$\mathscr{B}$ 是环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域）。若：

1. $\mathscr{A}\subset\mathscr{B}$；

2. 对 $X$ 上任意的另一环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域）$\mathscr{C}$，若 $\mathscr{A}\subset\mathscr{C}$，就有 $\mathscr{B}\subset\mathscr{C}$。

则称 $\mathscr{B}$ 是由集族 $\mathscr{A}$ 生成的环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域），即由集族 $\mathscr{A}$ 生成的环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域）是包含 $\mathscr{A}$ 的最小的环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域），将由集族 $\mathscr{A}$ 生成的环、单调系、$\lambda$ 系和 $\sigma$ 域分别记作 $r(\mathscr{A})$, $m(\mathscr{A})$, $l(\mathscr{A})$, $\sigma(\mathscr{A})$。

**Theorem 1.12.** 由任何集族 $\mathscr{A}$ 生成的环、单调系、$\lambda$ 系和 $\sigma$ 域都存在。

*Proof.* 设 $\mathscr{B}$ 为 $X$ 的所有子集构成的集族，则 $\mathscr{B}$ 是一个 $\sigma$ 域，所以 $\mathscr{B}$ 是一个环（或单调系，或 $\lambda$ 系）并且有 $\mathscr{A}\subset\mathscr{B}$。把所有包含集族 $\mathscr{A}$ 的环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域）的全体记为 $\mathbf{A}$，则 $\mathscr{B}\in\mathbf{A}$，于是 $\mathbf{A}$ 非空。记：

$$\mathscr{C}=\bigcap_{\mathscr{D}\in\mathbf{A}}\mathscr{D}$$

下证 $\mathscr{C}$ 是一个环（或单调系，或 $\lambda$ 系，或 $\sigma$ 域）。

(1) 任取 $A,B\in\mathscr{C}$，则对任意的 $\mathscr{D}\in\mathbf{A}$，有 $A,B\in\mathscr{D}$。因为 $\mathscr{D}$ 是一个环，所以 $A\cup B\in\mathscr{D}$, $A\backslash B\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$A\cup B\in\mathscr{C}$, $A\backslash B\in\mathscr{C}$，所以 $\mathscr{C}$ 是一个环。

(2) 任取单调集合序列 $\{A_n\}\subset\mathscr{C}$，则对任意的 $\mathscr{D}\in\mathbf{A}$，有 $\{A_n\}\subset\mathscr{D}$。因为 $\mathscr{D}$ 是单调系，所以 $\lim\limits_{n\to+\infty}A_n\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$\lim\limits_{n\to+\infty}A_n\in\mathscr{C}$。由 $\{A_n\}$ 的任意性，$\mathscr{C}$ 是一个单调系。

(3) 因为任意的 $\mathscr{D}\in\mathbf{A}$ 都是 $\lambda$ 系，所以 $X\in\mathscr{C}$。任取 $A,B\in\mathscr{C}$, $B\subset A$，则对任意的 $\mathscr{D}\in\mathbf{A}$，有 $A\backslash B\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$A\backslash B\in\mathscr{C}$。由 $A,B$ 的任意性，$\mathscr{C}$ 对包含关系的差运算封闭。任取单调递增集合序列 $\{A_n\}\subset\mathscr{C}$，则对任意的 $\mathscr{D}\in\mathbf{A}$，有 $\{A_n\}\subset\mathscr{D}$。因为 $\mathscr{D}$ 是 $\lambda$ 系，所以 $\mathscr{D}$ 是单调系，于是 $\lim\limits_{n\to+\infty}A_n\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$\lim\limits_{n\to+\infty}A_n\in\mathscr{C}$。由 $\{A_n\}$ 的任意性，$\mathscr{C}$ 对单调递增集合序列的极限是封闭的。综上，$\mathscr{C}$ 是一个 $\lambda$ 系。

(4) 因为任意的 $\mathscr{D}\in\mathbf{A}$ 都是 $\sigma$ 域，所以 $X\in\mathscr{C}$。任取 $A\in\mathscr{C}$，则对任意的 $\mathscr{D}$，有 $A\in\mathscr{D}$。因为 $\mathscr{D}$ 是 $\sigma$ 域，所以 $A^c\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$A^c\in\mathscr{C}$。由 $A$ 的任意性，$\mathscr{C}$ 对补的运算封闭。任取集合序列 $\{A_n\}\subset\mathscr{C}$，则对任意的 $\mathscr{D}$，有 $\{A_n\}\in\mathscr{D}$。因为 $\mathscr{D}$ 是 $\sigma$ 域，所以 $\bigcup\limits_{n=1}^{+\infty}A_n\in\mathscr{D}$。由 $\mathscr{D}$ 的任意性，$\bigcup\limits_{n=1}^{+\infty}A_n\in\mathscr{C}$。由 $\{A_n\}$ 的任意性，$\mathscr{C}$ 对可列并的运算封闭。综上，$\mathscr{C}$ 是一个 $\sigma$ 域。 □

**Theorem 1.13.** 如果 $\mathscr{A}$ 是半环，则：

$$r(\mathscr{A})=\bigcup_{n=1}^{+\infty}\left\{\bigcup_{i=1}^{n}A_i:A_i\in\mathscr{A};\ A_i\cap A_j=\varnothing,\forall\, i\neq j\right\}$$

*Proof.* 令：

$$\mathscr{B}=\bigcup_{n=1}^{+\infty}\left\{\bigcup_{i=1}^{n}A_i:A_i\in\mathscr{A};\ A_i\cap A_j=\varnothing,\forall\, i\neq j\right\}$$

由 $\mathscr{B}$ 的定义，$\mathscr{A}\subset\mathscr{B}$（$n=1$）。因为环对有限并封闭，所以包含 $\mathscr{A}$ 的环必然包含 $\mathscr{B}$。若证得 $\mathscr{B}$ 是一个环，则可得到 $r(\mathscr{A})=\mathscr{B}$。

任取 $A,B\in\mathscr{B}$，则存在 $m,n\in\mathbb{N}^+$ 和互不相交的 $A_1,A_2,\ldots,A_m\in\mathscr{A}$、互不相交的 $B_1,B_2,\ldots,B_n\in\mathscr{A}$ 使得：

$$A=\bigcup_{i=1}^{m}A_i,\ B=\bigcup_{i=1}^{n}B_i$$

于是：

$$A\backslash B=A\cap B^c=\bigcup_{i=1}^{m}(A_i\cap B^c)=\bigcup_{i=1}^{m}\bigcap_{j=1}^{n}(A_i\cap B_j^c)=\bigcup_{i=1}^{m}\bigcap_{j=1}^{n}[A_i\backslash(A_i\cap B_j)]$$

因为 $\mathscr{A}$ 是半环，由定理 1.4和 $\pi$ 系的定义可知 $\mathscr{A}$ 对交封闭，所以对任意的 $i,j$ 有 $A_i\cap B_j\in\mathscr{A}$ 且 $A_i\cap B_j\subset A_i$，于是存在互不相交的 $C_{ij1},C_{ij2},\ldots,C_{ijr_{ij}}\in\mathscr{A}$ 使得：

$$A_i\backslash(A_i\cap B_j)=\bigcup_{k=1}^{r_{ij}}C_{ijk}$$

于是：

$$\begin{aligned}
A\backslash B&=\bigcup_{i=1}^{m}\bigcap_{j=1}^{n}[A_i\backslash(A_i\cap B_j)]=\bigcup_{i=1}^{m}\bigcap_{j=1}^{n}\bigcup_{k=1}^{r_{ij}}C_{ijk}\\
&=\bigcup_{i=1}^{m}(C_{111}\cup C_{112}\cdots\cup C_{11r_{11}})\cap(C_{121}\cup C_{122}\cdots\cup C_{12r_{12}})\cdots\\
&=\bigcup_{i=1}^{m}\left\{\bigcup_{j=1}^{r_{11}}[C_{11j}\cap(C_{121}\cup C_{122}\cdots\cup C_{12r_{12}})]\right\}\cap(C_{131}\cup C_{132}\cdots\cup C_{13r_{13}})\cdots\\
&=\bigcup_{i=1}^{m}\left[\bigcup_{j=1}^{r_{11}}\bigcup_{k=1}^{r_{12}}(C_{11j}\cap C_{12k})\right]\cap(C_{131}\cup C_{132}\cdots\cup C_{13r_{13}})\cdots\\
&=\bigcup_{i=1}^{n}\bigcup_{(k_1,k_2,\ldots,k_n)\in\prod_{k=1}^{n}\{1,2,\ldots,r_{ik}\}}\bigcap_{j=1}^{n}C_{ijk_j}
\end{aligned}$$

因为 $\mathscr{A}$ 是半环，对交封闭，所以 $\bigcap_{j=1}^{n}C_{ijk_j}\in\mathscr{A}$。因为 $C_{ij1},C_{ij2},\ldots,C_{ijr_{ij}}$ 互不相交，所以 $\bigcap_{j=1}^{n}C_{ijk_j}$ 互不相交（给定一组 $k_1^{(1)},k_2^{(1)},\ldots,k_n^{(1)}$，得到 $\bigcap_{j=1}^{n}C_{ijk_j^{(1)}}$。考虑 $A\backslash B$ 中另一个参与并集运算的 $\bigcap_{j=1}^{n}C_{ijk_j^{(2)}}$，它所对应的 $k_1^{(2)},k_2^{(2)},\ldots,k_n^{(2)}$ 必然不同于 $k_1^{(1)},k_2^{(1)},\ldots,k_n^{(1)}$，于是存在 $a,b$ 使得 $k_a^{(1)}\neq k_b^{(2)}$。注意到 $C_{ijk_a^{(1)}}\cap C_{ijk_b^{(2)}}=\varnothing$，所以 $\bigcap_{j=1}^{n}C_{ijk_j^{(1)}}$ 和 $\bigcap_{j=1}^{n}C_{ijk_j^{(2)}}$ 不相交。由 $k_1^{(1)},k_2^{(1)},\ldots,k_n^{(1)}$ 和 $k_1^{(2)},k_2^{(2)},\ldots,k_n^{(2)}$ 的任意性即可得结论），于是有 $A\backslash B\in\mathscr{B}$，即 $\mathscr{B}$ 对差封闭。

考虑：

$$A\cup B=B\cup(A\backslash B)=B\cup\left(\bigcup_{i=1}^{n}\bigcup_{(k_1,k_2,\ldots,k_n)\in\prod_{k=1}^{n}\{1,2,\ldots,r_{ik}\}}\bigcap_{j=1}^{n}C_{ijk_j}\right)$$

因为 $B\cap(A\backslash B)=\varnothing$，所以：

$$B\cap\left(\bigcup_{i=1}^{n}\bigcup_{(k_1,k_2,\ldots,k_n)\in\prod_{k=1}^{n}\{1,2,\ldots,r_{ik}\}}\bigcap_{j=1}^{n}C_{ijk_j}\right)=\varnothing$$

于是 $A\cup B\in\mathscr{B}$。

综上，$\mathscr{B}$ 对并和差封闭，所以 $\mathscr{B}$ 是一个环。 □

**Theorem 1.14.** 若 $\mathscr{A}$ 是域，则 $\sigma(\mathscr{A})=m(\mathscr{A})$。

*Proof.* 因为 $\sigma(\mathscr{A})$ 是包含 $\mathscr{A}$ 的 $\sigma$ 域，所以 $\sigma(\mathscr{A})$ 也是包含 $\mathscr{A}$ 的单调系。因为 $m(\mathscr{A})$ 是包含 $\mathscr{A}$ 的最小的单调系，所以 $m(\mathscr{A})\subset\sigma(\mathscr{A})$。

下证 $\sigma(\mathscr{A})\subset m(\mathscr{A})$，由定理 1.9可得一个既是单调系又是域的集族必是 $\sigma$ 域，所以证得 $m(\mathscr{A})$ 是一个域即可得到 $\sigma(\mathscr{A})\subset m(\mathscr{A})$。又因为 $\mathscr{A}$ 是域，所以 $X\in\mathscr{A}$，同时 $X\in m(\mathscr{A})$，由定理 1.8可得一个包含 $X$ 的环是域，所以只需证明 $m(\mathscr{A})$ 是一个环。

对任意的 $A\in\mathscr{A}$，有 $A\in m(\mathscr{A})$。令：

$$\mathscr{B}_A=\{B:B,A\cup B,A\backslash B\in m(\mathscr{A})\}$$

任取 $\mathscr{B}_A$ 中一个单调不减序列 $\{B_n\}$，则对任意的 $n\in\mathbb{N}^+$，有 $B_n\in m(\mathscr{A})$，于是 $\bigcup\limits_{n=1}^{+\infty}B_n\in m(\mathscr{A})$。考虑：

$$A\cup\left(\bigcup_{n=1}^{+\infty}B_n\right)=\bigcup_{n=1}^{+\infty}(A\cup B_n)$$

则 $\{A\cup B_n\}$ 也是一个单调不减序列，同时对任意的 $n\in\mathbb{N}^+$，有 $A\cup B_n\in m(\mathscr{A})$，所以：

$$A\cup\left(\bigcup_{n=1}^{+\infty}B_n\right)\in m(\mathscr{A})$$

考虑：

$$A\backslash\left(\bigcup_{n=1}^{+\infty}B_n\right)=\bigcap_{n=1}^{+\infty}(A\backslash B_n)$$

则 $\{A\backslash B_n\}$ 是一个单调不增序列，同时对任意的 $n\in\mathbb{N}^+$，有 $A\backslash B_n\in m(\mathscr{A})$，所以：

$$A\backslash\left(\bigcup_{n=1}^{+\infty}B_n\right)\in m(\mathscr{A})$$

由 $\{B_n\}$ 的任意性，$\mathscr{B}_A$ 对单调不减序列的极限封闭。

任取 $\mathscr{B}_A$ 中的一个单调不增序列 $\{C_n\}$，则对任意的 $n\in\mathbb{N}^+$，有 $C_n\in m(\mathscr{A})$，于是 $\bigcap\limits_{n=1}^{+\infty}C_n\in m(\mathscr{A})$。考虑：

$$A\cup\left(\bigcap_{n=1}^{+\infty}B_n\right)=\bigcap_{n=1}^{+\infty}(A\cup C_n)$$

则 $\{A\cup C_n\}$ 也是一个单调不增序列，同时对任意的 $n\in\mathbb{N}^+$，有 $A\cup C_n\in m(\mathscr{A})$，所以：

$$A\cup\left(\bigcap_{n=1}^{+\infty}C_n\right)\in m(\mathscr{A})$$

考虑：

$$A\backslash\left(\bigcap_{n=1}^{+\infty}C_n\right)=\bigcup_{n=1}^{+\infty}(A\backslash C_n)$$

则 $\{A\backslash C_n\}$ 是一个单调不减序列，同时对任意的 $n\in\mathbb{N}^+$，有 $A\backslash C_n\in m(\mathscr{A})$，所以：

$$A\backslash\left(\bigcap_{n=1}^{+\infty}C_n\right)\in m(\mathscr{A})$$

由 $\{C_n\}$ 的任意性，$\mathscr{B}_A$ 对单调不增序列的极限封闭。

综上，$\mathscr{B}_A$ 是一个单调系。因为 $\mathscr{A}$ 是一个域，由定理 1.4可知 $\mathscr{A}$ 是一个环，所以 $\mathscr{A}$ 对并和差封闭，即 $\mathscr{A}$ 中任意元素与 $A$ 的并集和差集都在 $\mathscr{A}$ 中，根据生成的定义，$\mathscr{A}\in m(\mathscr{A})$，即 $\mathscr{A}$ 中任意元素与 $A$ 的并集和差集都在 $m(\mathscr{A})$ 中，再加上它们本身也都在 $m(\mathscr{A})$ 中，所以 $\mathscr{A}\subset\mathscr{B}_A$。由 $m(\mathscr{A})$ 的定义，$m(\mathscr{A})\subset\mathscr{B}_A$，于是：

$$A\in\mathscr{A},\ B\in m(\mathscr{A})\Rightarrow A\cup B, A\backslash B\in m(\mathscr{A})$$

对任意的 $B\in m(\mathscr{A})$，令：

$$\mathscr{D}_B=\{A: A, A\cup B, A\backslash B\in m(\mathscr{A})\}$$

任取 $\mathscr{D}_A$ 中一个单调不减序列 $\{A_n\}$，则对任意的 $n\in\mathbb{N}^+$，有 $A_n\in m(\mathscr{A})$，于是 $\bigcup\limits_{n=1}^{+\infty}A_n\in m(\mathscr{A})$。考虑：

$$B\cup\left(\bigcup_{n=1}^{+\infty}A_n\right)=\bigcup_{n=1}^{+\infty}(B\cup A_n)$$

则 $\{B\cup A_n\}$ 也是一个单调不减序列，同时对任意的 $n\in\mathbb{N}^+$，有 $B\cup A_n\in m(\mathscr{A})$，所以：

$$B\cup\left(\bigcup_{n=1}^{+\infty}A_n\right)\in m(\mathscr{A})$$

考虑：

$$B\backslash\left(\bigcup_{n=1}^{+\infty}A_n\right)=\bigcap_{n=1}^{+\infty}(B\backslash A_n)$$

则 $\{B\backslash A_n\}$ 是一个单调不增序列，同时对任意的 $n\in\mathbb{N}^+$，有 $B\backslash A_n\in m(\mathscr{A})$，所以：

$$B\backslash\left(\bigcup_{n=1}^{+\infty}A_n\right)\in m(\mathscr{A})$$

由 $\{A_n\}$ 的任意性，$\mathscr{D}_B$ 对单调不减序列的极限封闭。

任取 $\mathscr{D}_B$ 中的一个单调不增序列 $\{D_n\}$，则对任意的 $n\in\mathbb{N}^+$，有 $D_n\in m(\mathscr{A})$，于是 $\bigcap\limits_{n=1}^{+\infty}D_n\in m(\mathscr{A})$。考虑：

$$B\cup\left(\bigcap_{n=1}^{+\infty}D_n\right)=\bigcap_{n=1}^{+\infty}(B\cup D_n)$$

则 $\{B\cup D_n\}$ 也是一个单调不增序列，同时对任意的 $n\in\mathbb{N}^+$，有 $B\cup D_n\in m(\mathscr{A})$，所以：

$$B\cup\left(\bigcap_{n=1}^{+\infty}D_n\right)\in m(\mathscr{A})$$

考虑：

$$B\backslash\left(\bigcap_{n=1}^{+\infty}D_n\right)=\bigcup_{n=1}^{+\infty}(B\backslash D_n)$$

则 $\{B\backslash D_n\}$ 是一个单调不减序列，同时对任意的 $n\in\mathbb{N}^+$，有 $B\backslash D_n\in m(\mathscr{A})$，所以：

$$B\backslash\left(\bigcap_{n=1}^{+\infty}D_n\right)\in m(\mathscr{A})$$

由 $\{D_n\}$ 的任意性，$\mathscr{D}_B$ 对单调不增序列的极限封闭。

综上，$\mathscr{D}_B$ 是一个单调系。由证明 $\mathscr{B}_A$ 是单调系时 $A$ 的任意性以及：

$$A \in \mathscr{A},\ B \in m(\mathscr{A}) \Rightarrow A \cup B, A \backslash B \in m(\mathscr{A})$$

可知 $\mathscr{A} \subset \mathscr{D}_B$，由生成的定义，$m(\mathscr{A}) \subset \mathscr{D}_B$，再根据 $\mathscr{D}_B$ 定义中 $B$ 的任意性可得：

$$A, B \in m(\mathscr{A}) \Rightarrow A \cup B, A \backslash B \in m(\mathscr{A})$$

所以 $m(\mathscr{A})$ 是一个环。 □

**Corollary 1.1.** 如果 $\mathscr{A}$ 是域，$\mathscr{B}$ 是单调系，则：

$$\mathscr{A} \subset \mathscr{B} \Rightarrow \sigma(\mathscr{A}) \subset \mathscr{B}$$

并且该推论与上一定理等价。

*Proof.* **(1) 必要性**：因为 $\mathscr{A}$ 是域，所以 $\sigma(\mathscr{A}) = m(\mathscr{A})$。因为 $\mathscr{B}$ 是包含 $\mathscr{A}$ 的单调系，由生成的定义，$\sigma(\mathscr{A}) \subset \mathscr{B}$。

**(2) 充分性**：由 $\mathscr{B}$ 的任意性和生成的定义直接可得。 □

**Theorem 1.15.** 如果 $\mathscr{A}$ 是 $\pi$ 系，则 $\sigma(\mathscr{A}) = l(\mathscr{A})$。

*Proof.* 由定理 1.6可知 $\sigma(\mathscr{A})$ 是一个 $\lambda$ 域，由生成的定义可得 $\mathscr{A} \subset \sigma(\mathscr{A})$，于是 $l(\mathscr{A}) \subset \sigma(\mathscr{A})$。下证 $\sigma(\mathscr{A}) \subset l(\mathscr{A})$。由生成的定义 $\mathscr{A} \subset l(\mathscr{A})$，若证得 $l(\mathscr{A})$ 是一个 $\sigma$ 域，则可得 $\sigma(\mathscr{A}) \subset l(\mathscr{A})$。由定理 1.10可知只需证明 $l(\mathscr{A})$ 是一个 $\pi$ 系。

对任意的 $A \in \mathscr{A}$，令：

$$\mathscr{B}_A = \{B : B, A \cap B \in l(\mathscr{A})\}$$

因为 $l(\mathscr{A})$ 是 $\lambda$ 系，所以 $X \in l(\mathscr{A})$，而 $A \cap X = A \in \mathscr{A}$，由生成的定义，$A \cap X \in l(\mathscr{A})$，于是 $X \in \mathscr{B}_A$。

任取 $B, C \in \mathscr{B}_A$ 且 $B \subset C$，则有 $B, C \in l(\mathscr{A})$，于是 $C \backslash B \in l(\mathscr{A})$。注意到：

$$A \cap (C \backslash B) = (A \cap C) \backslash (A \cap B)$$

因为 $B, C \in \mathscr{B}_A$，所以 $A \cap B, A \cap C \in l(\mathscr{A})$。因为 $B \subset C$，所以 $A \cap B \subset A \cap C$，由 $\lambda$ 系的定义可得 $(A \cap C) \backslash (A \cap B) \in l(\mathscr{A})$，即 $A \cap (C \backslash B) \in l(\mathscr{A})$，所以 $C \backslash B \in \mathscr{B}_A$。

任取 $\mathscr{B}_A$ 中的一个单调不减的集合列 $\{B_n\}$，则有 $B_n \in l(\mathscr{A}), A \cap B_n \in l(\mathscr{A})$ 对 $n \in \mathbb{N}^+$ 成立，于是 $\{B_n\}$ 是 $l(\mathscr{A})$ 中单调不减的集合列。由 $\lambda$ 系的定义可知：

$$\bigcup_{n=1}^{+\infty} B_n \in l(\mathscr{A})$$

考虑：

$$A \cap \left( \bigcup_{n=1}^{+\infty} B_n \right) = \bigcup_{n=1}^{+\infty} (A \cap B_n)$$

因为 $\{B_n\}$ 单调不减，所以 $\{A\cap B_n\}$ 是 $l(\mathscr{A})$ 中单调不减的集合列，由 $\lambda$ 系的定义可得：

$$A\cap\left(\bigcup_{n=1}^{+\infty}B_n\right)\in l(\mathscr{A})$$

综上，$\mathscr{B}_A$ 是一个 $\lambda$ 系。因为 $\mathscr{A}$ 是一个 $\pi$ 系，所以 $\mathscr{A}$ 对交封闭，于是对任意的 $D\in\mathscr{A}$ 有 $D,A\cap D\in l(\mathscr{A})$，即 $\mathscr{A}\subset\mathscr{B}_A$，由生成的定义可得 $l(\mathscr{A})\subset\mathscr{B}_A$。这说明：

$$A\in\mathscr{A},B\in l(\mathscr{A})\Rightarrow A\cap B\in l(\mathscr{A})$$

对任意的 $B\in l(\mathscr{A})$，令：

$$\mathscr{C}_B=\{A:A,B\cap A\in l(\mathscr{A})\}$$

因为 $l(\mathscr{A})$ 是 $\lambda$ 系，所以 $X\in l(\mathscr{A})$，而 $B\cap X=B\in l(\mathscr{A})$，由生成的定义，$B\cap X\in l(\mathscr{A})$，于是 $X\in\mathscr{C}_B$。

任取 $D,E\in\mathscr{C}_B$ 且 $D\subset E$，则有 $D,E\in l(\mathscr{A})$，于是 $E\backslash D\in l(\mathscr{A})$。注意到：

$$B\cap(D\backslash E)=(B\cap D)\backslash(B\cap E)$$

因为 $D,E\in\mathscr{C}_B$，所以 $B\cap D,B\cap E\in l(\mathscr{A})$。因为 $D\subset E$，所以 $B\cap D\subset B\cap E$，由 $\lambda$ 系的定义可得 $(B\cap D)\backslash(B\cap E)\in l(\mathscr{A})$，即 $B\cap(D\backslash E)\in l(\mathscr{A})$，所以 $D\backslash E\in\mathscr{C}_B$。

任取 $\mathscr{C}_B$ 中的一个单调不减的集合列 $\{C_n\}$，则有 $C_n\in l(\mathscr{A}),B\cap C_n\in l(\mathscr{A})$ 对 $n\in\mathbb{N}^+$ 成立，于是 $\{C_n\}$ 是 $l(\mathscr{A})$ 中单调不减的集合列。由 $\lambda$ 系的定义可知：

$$\bigcup_{n=1}^{+\infty}C_n\in l(\mathscr{A})$$

考虑：

$$B\cap\left(\bigcup_{n=1}^{+\infty}C_n\right)=\bigcup_{n=1}^{+\infty}(B\cap C_n)$$

因为 $\{C_n\}$ 单调不减，所以 $\{B\cap C_n\}$ 是 $l(\mathscr{A})$ 中单调不减的集合列，由 $\lambda$ 系的定义可得：

$$B\cap\left(\bigcup_{n=1}^{+\infty}C_n\right)\in l(\mathscr{A})$$

综上，$\mathscr{C}_B$ 是一个 $\lambda$ 系。由：

$$A\in\mathscr{A},B\in l(\mathscr{A})\Rightarrow A\cap B\in l(\mathscr{A})$$

可知 $\mathscr{A}\subset\mathscr{C}_B$，所以 $l(\mathscr{A})\subset\mathscr{C}_B$。由 $\mathscr{C}_B$ 的定义可知 $l(\mathscr{A})$ 对交封闭，所以 $l(\mathscr{A})$ 是一个 $\pi$ 系。 □

**Corollary 1.2.** 如果 $\mathscr{A}$ 是 $\pi$ 系，$\mathscr{B}$ 是 $\lambda$ 系，则：

$$\mathscr{A}\subset\mathscr{B}\Rightarrow\sigma(\mathscr{A})\subset\mathscr{B}$$

并且该推论与上一定理等价。

*Proof.* **(1) 必要性**：因为 $\mathscr{A}$ 是 $\pi$ 系，所以 $\sigma(\mathscr{A})=l(\mathscr{A})$。因为 $\mathscr{B}$ 是包含 $\mathscr{A}$ 的 $\lambda$ 系，由生成的定义，$\sigma(\mathscr{A})\subset\mathscr{B}$。

**(2) 充分性**：由 $\mathscr{B}$ 的任意性和生成的定义直接可得。 □

### 1.1.3 $\mathbb{R}$ 上开集与闭集的构造

**Definition 1.12.** 设 $E$ 是 $\mathbb{R}$ 上的开集，如果开区间 $(\alpha,\beta)\subset E$ 且 $\alpha,\beta\notin E$，则称 $(\alpha,\beta)$ 为 $E$ 的构成区间 *(component interval)*。

**Theorem 1.16.** $\mathbb{R}$ 上任一非空开集 $E$ 可以表示为至多可列个不相交的构成区间的并集。

*Proof.* 该定理的证明分为如下三步：

1. $E$ 的任意两个不同的构成区间不相交；
2. $E$ 中的任意一点必含在一个构成区间中；
3. $E$ 的所有构成区间的并集为 $E$ 且构成区间至多可列。

(1) 任取 $E$ 的两个不同的构成区间 $(\alpha_1,\beta_1)$, $(\alpha_2,\beta_2)$，若这两个构成区间相交，则 $\alpha_1,\beta_1,\alpha_2,\beta_2$ 这四个点至少有一个点在另一个构成区间内，从而在 $E$ 中，这与构成区间的定义矛盾，所以 $(\alpha_1,\beta_1)$, $(\alpha_2,\beta_2)$ 不相交。由 $(\alpha_1,\beta_1)$, $(\alpha_2,\beta_2)$ 的任意性可得 $E$ 的任意两个不同的构成区间必不相交。

(2) 任取 $x\in E$，记：

$$\mathscr{A}=\{(\alpha,\beta):x\in(\alpha,\beta)\subset E\}$$

因为 $E$ 是开集，所以 $\mathscr{A}\neq\varnothing$。取：

$$\alpha_0=\inf\{\alpha:(\alpha,\beta)\in\mathscr{A}\},\ \beta_0=\sup\{\beta:(\alpha,\beta)\in\mathscr{A}\}$$

作开区间 $(\alpha_0,\beta_0)$，显然有 $x\in(\alpha_0,\beta_0)$。下面证明 $(\alpha_0,\beta_0)$ 是一个构成区间。

任取 $x_1\in(\alpha_0,\beta_0)$，则 $\alpha_0<x_1<\beta_0$。由上下确界的定义，存在 $(\alpha_1,\beta_1),(\alpha_2,\beta_2)\in\mathscr{A}$ 使得 $x_1\in(\alpha_1,\beta_2)$，其中 $\alpha_0<\alpha_1<\alpha_2<\beta_2<\beta_1<\beta_0$（若不满足，则 $(\alpha_1,\beta_1),(\alpha_2,\beta_2)$ 无交集，$x$ 必然不可能同时存在于这两个区间之中，这与 $\mathscr{A}$ 的定义矛盾）。因为 $(\alpha_1,\beta_1),(\alpha_2,\beta_2)\in\mathscr{A}$，所以 $(\alpha_1,\beta_1),(\alpha_2,\beta_2)\subset E$，于是 $(\alpha_1,\beta_2)\subset E$，$x_1\in E$，即 $(\alpha_0,\beta_0)\subset E$。

若 $\alpha_0\in E$，因为 $E$ 是一个开集，所以必然存在一个 $\varepsilon>0$ 使得 $(\alpha_0-\varepsilon,\alpha_0+\varepsilon)\subset E$，取 $\beta_3\in(x,\beta_0)$，则 $(\alpha_0-\varepsilon,\beta_3)\subset E$ 且 $x\in(\alpha_0-\varepsilon,\beta_3)$，那么就有 $(\alpha_0-\varepsilon,\beta_3)\in\mathscr{A}$，这与 $\alpha_0$ 是下确界矛盾，于是 $\alpha_0\notin E$。同理，$\beta_0\notin E$。

综上，$(\alpha_0,\beta_0)$ 是一个构成区间。

由先前 $x$ 的任意性可得对于任意的 $x\in E$，$x$ 必含在 $E$ 的一个构成区间中，具体的构成区间由上述 $\alpha_0,\beta_0$ 的产生过程给出。

(3) 由 (2) 可知 $E$ 中任意一点必含在一个构成区间中，对 $E$ 中所有的点取其对应的构成区间的并集即可得到所有构成区间的并集为 $E$。由有理数在实数系中的稠密性，各构成区间必含有一个有理数。由 (1) 可得不同的构成区间不相交，于是每个构成区间可由其中包含的一个有理数来表示，而有理数是可列的，所以 $E$ 的构成区间至多可列。 □

**Corollary 1.3.** $\mathbb{R}$ 上的闭集是从 $\mathbb{R}$ 上挖掉至多可列个互不相交的开区间所得到的集合。

*Proof.* 设闭集 $E\subset\mathbb{R}$，由**??**可知 $E^c$ 是一个开集，则 $E^c$ 可表示为其构成区间的并集，于是 $E=\mathbb{R}\backslash E^c$ 是从 $\mathbb{R}$ 上挖掉至多可列个互不相交的开区间所得到的集合。 □

### 1.1.4 Borel $\sigma$ 域

**Definition 1.13.** 设 $\mathbb{R}$ 上所有有限开区间构成的集合为 $\mathcal{C}$，称 $\sigma(\mathcal{C})$ 为博雷尔 $\sigma$ 域 *(Borel $\sigma$-field)*，记作 $\mathcal{B}$。$\mathcal{B}$ 中的元素被称为博雷尔集 *(Borel set)*。

**Lemma 1.1.** $\mathbb{R}$ 上任意开集可以表示为至多可列个有限开区间的并集。

*Proof.* 设 $E$ 是 $\mathbb{R}$ 上的一个开集，取 $E$ 的一个构成区间 $(\alpha,\beta)$。

**(1)** $\alpha=-\infty,\beta\in\mathbb{R}$：此时有：

$$(\alpha,\beta)=\bigcup_{n=1}^{+\infty}(\beta-n,\beta)$$

**(2)** $\alpha=-\infty,\beta=+\infty$：此时有：

$$(\alpha,\beta)=\bigcup_{n=1}^{+\infty}(2-n,1+n)$$

**(3)** $\alpha,\beta\in\mathbb{R}$：此时其自身即为有限开区间。

**(4)** $\alpha\in\mathbb{R},\beta=+\infty$：此时有：

$$(\alpha,\beta)=\bigcup_{n=1}^{+\infty}(\alpha,\alpha+n)$$

上述结果表明开集的构成区间可由至多可列个有限开区间的并集表示，由定理 1.16可知开集至多由可列个构成区间的并集表示，因为，所以 $\mathbb{R}$ 上任意开集可以表示为至多可列个有限开区间的并集。 □

可列个可列是可列

**Theorem 1.17.** $\mathbb{R}$ 中的开集、闭集和区间（有限或无穷）都是 *Borel* 集。

*Proof.* 由引理 1.1和 $\sigma$ 域对可列并封闭可得开集是 Borel 集。

$\mathbb{R}$ 上的闭集可表示为实直线挖去一些开区间后剩下的部分，而开区间都是开集，从而都是 Borel 集，由引理 1.1可知它们也都能表示为至多可列个有限开区间的并集。可将差运算转化为交与补的运算，由定理 1.1又可将交运算转化为并与补的运算，因为 $\sigma$ 域对可列并与补封闭，所以闭集是 Borel 集。

开区间是开集，闭区间是闭集，它们都是 Borel 集，剩余区间可以由这二者进行并与补的运算得到，因为 $\sigma$ 域对可列并与补封闭，所以区间也是 Borel 集。 □

**Theorem 1.18.** $\mathcal{B}$ 有如下等价定义：

1. 设 $\mathbb{R}$ 上所有开集构成的集合为 $\mathcal{O}$，则 $\sigma(\mathcal{O})=\mathcal{B}$；
2. $\mathcal{B}=\sigma(\{(a,b]:a,b\in\mathbb{R}\})$；
3. $\mathcal{B}=\sigma(\{(-\infty,a]:a\in\mathbb{R}\})$。

*Proof.* (1) 因为有限开区间都是开集，所以 $\mathcal{C}\subset\mathcal{O}$，于是有 $\mathcal{C}\subset\sigma(\mathcal{O})$，由生成的定义可得 $\mathcal{B}=\sigma(\mathcal{C})\subset\sigma(\mathcal{O})$。

由引理 1.1可知对 $\mathbb{R}$ 中的任意一个开集，它都可以表示为至多可列个有限开区间的并集，所以 $\mathcal{O} \subset \sigma(\mathcal{C}) = \mathcal{B}$，由生成的定义可得 $\sigma(\mathcal{O}) \subset \mathcal{B}$。

综上，$\sigma(\mathcal{O}) = \mathcal{B}$。

(2) 对于任意的 $(a, b]$，$a, b \in \mathbb{R}$，取 $(-\infty, a], (-\infty, b]$ 即有 $(a, b] = (-\infty, b] \backslash (-\infty, a] \in \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$，所以有 $\{(a, b] : a, b \in \mathbb{R}\} \subset \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$。由生成的定义，$\sigma(\{(a, b] : a, b \in \mathbb{R}\}) \subset \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$。

对于任意的 $(-\infty, a]$，$a \in \mathbb{R}$，因为 $\sigma$ 域对可列并封闭，并且 $(a-n, a] \in \{(a, b] : a, b \in \mathbb{R}\}$ 对任意 $n \in \mathbb{N}^+$ 成立，所以：

$$(-\infty, a] = \bigcup_{n=1}^{+\infty} (a-n, a] \in \sigma(\{(a, b] : a, b \in \mathbb{R}\})$$

由生成的定义，$\sigma(\{(-\infty, a] : a \in \mathbb{R}\}) \subset \sigma(\{(a, b] : a, b \in \mathbb{R}\})$。

综上，$\sigma(\{(-\infty, a] : a \in \mathbb{R}\}) = \sigma(\{(a, b] : a, b \in \mathbb{R}\})$。

对 $(-\infty, a]$ 取补集即可得到 $(a, +\infty)$，于是 $\{(a, +\infty) : a \in \mathbb{R}\} \subset \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$，因为所有有限开区间都可以由 $\{(a, +\infty) : a \in \mathbb{R}\}$ 中两个元素的差集来表示，所以 $\mathcal{C} \subset \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$。由生成的定义，$\mathcal{B} = \sigma(\mathcal{C}) \subset \sigma(\{(-\infty, a] : a \in \mathbb{R}\})$。

注意到对任意的 $(a, +\infty), a \in \mathbb{R}$ 都可由有限开区间的可列并表示，于是有 $\{(a, +\infty) : a \in \mathbb{R}\} \subset \sigma(\mathcal{C})$。由生成的定义，$\sigma(\{(a, +\infty) : a \in \mathbb{R}\}) \subset \sigma(\mathcal{C})$ □

证明未完成，考虑到 borel 代数的等价定义实在太多

**Definition 1.14.** 定义：

$$\mathcal{B}_{\overline{\mathbb{R}}} = \sigma(\mathcal{B}, -\infty, +\infty)$$

**Theorem 1.19.** $\mathcal{B}_{\overline{\mathbb{R}}}$ 有如下等价定义：

$$\begin{aligned}
\mathcal{B}_{\overline{\mathbb{R}}} &= \sigma([-\infty, a) : a \in \mathbb{R}) \\
&= \sigma([-\infty, a] : a \in \mathbb{R}) \\
&= \sigma((a, +\infty] : a \in \mathbb{R}) \\
&= \sigma([a, +\infty] : a \in \mathbb{R})
\end{aligned}$$

*Proof.* (1) 对任意的 $a \in \mathbb{R}$，有：

$$[-\infty, a) = \{-\infty\} \cup (-\infty, a) \in \mathcal{B}_{\overline{\mathbb{R}}}$$

由生成的定义：

$$\sigma([-\infty, a) : a \in \mathbb{R}) \subset \mathcal{B}_{\overline{\mathbb{R}}}$$

□

## 1.2 测度空间

### 1.2.1 集函数与测度

**Definition 1.15.** 设 $\mathscr{A}$ 是 $X$ 上的一个集族。称定义在 $\mathscr{A}$ 上并且取值非负的函数为非负集函数。

**Definition 1.16.** 设 $\mathscr{A}$ 是 $X$ 上的集族，$\mu$ 是定义在其上的非负集函数。

1. 如果对 $\mathscr{A}$ 中任意互不相交的 $A_1, A_2, \ldots, A_n$，有：

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i)$$

则称 $\mu$ 具有有限可加性 *(finite additivity)*；

2. 如果对 $\mathscr{A}$ 中任意的 $A_1, A_2, \ldots, A_n$，有：

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mu(A_i)$$

则称 $\mu$ 具有次有限可加性 *(finite subadditivity)*；

3. 如果对任意的 $A, B \in \mathscr{A}$，$A \subset B$，$B \backslash A \in \mathscr{A}$，只要 $\mu(A) < +\infty$，就有：

$$\mu(B \backslash A) = \mu(B) - \mu(A)$$

则称 $\mu$ 具有可减性 *(reducibility)*；

4. 若对任意的 $A, B \in \mathscr{A}$，$A \subset B$，有 $\mu(A) \leqslant \mu(B)$，则称 $\mu$ 具有单调性 *(monotonicity)*；

5. 若对 $\mathscr{A}$ 中任意互不相交的集合序列 $\{A_n\}$，只要 $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{A}$，就有：

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mu(A_n)$$

则称 $\mu$ 具有可列可加性 *(countably additivity)*；

6. 若对 $\mathscr{A}$ 中任意的集合序列 $\{A_n\}$，只要 $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{A}$，就有：

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) \leqslant \sum_{n=1}^{+\infty} \mu(A_n)$$

则称 $\mu$ 具有次可列可加性 *(countably subadditivity)*；

7. 若对任意的 $\{A_n\} \subset \mathscr{A}$ 且 $A_n \uparrow A \in \mathscr{A}$，有：

$$\mu(A) = \lim_{n \to +\infty} \mu(A_n)$$

则称 $\mu$ 具有下连续性 *(continuity from below)*；

8. 若对任意的 $\{A_n\} \subset \mathscr{A}$ 且 $A_n \downarrow A$，$\mu(A_1) < +\infty$，有：

$$\mu(A) = \lim_{n \to +\infty} \mu(A_n)$$

则称 $\mu$ 具有上连续性 *(continuity from above)*。

**Definition 1.17.** 设 $\mathscr{A}$ 是 $X$ 上的集族，$\varnothing \in \mathscr{A}$。如果 $\mathscr{A}$ 上的非负集函数 $\mu$ 满足：

1. $\mu(\varnothing) = 0$；
2. $\mu$ 具有可列可加性。

则称 $\mu$ 为 $\mathscr{A}$ 上的测度 *(measure)*。如果对任意的 $A \in \mathscr{A}$ 有 $\mu(A) < +\infty$，则称测度 $\mu$ 是有限的；如果对任意的 $A \in \mathscr{A}$，存在 $\mathscr{A}$ 中的集合序列 $\{A_n\}$，满足 $\mu(A_n) < +\infty$, $\forall\, n \in \mathbb{N}^+$，使得 $A \subset \bigcup\limits_{n=1}^{+\infty} A_n$，则称测度 $\mu$ 是 $\sigma$ 有限的。

**Theorem 1.20.** 测度具有有限可加性与可减性。

*Proof.* 设 $\mu$ 是 $\mathscr{A}$ 上的测度。

(1) 任取互不相交的 $A_1, A_2, \ldots, A_n \in \mathscr{A}$，由 $\mu$ 的可列可加性可得：

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \mu(A_1 \cup A_2 \cdots \cup A_n \cup \varnothing \cdots) = \sum_{i=1}^{n} \mu(A_i) + 0 = \sum_{i=1}^{n} \mu(A_i)$$

由 $A_1, A_2, \ldots, A_n$ 的任意性，$\mu$ 具有有限可加性。

(2) 任取 $A, B \in \mathscr{A}$，$A \subset B$，$B \backslash A \in \mathscr{A}$，$\mu(A) < +\infty$，显然 $(B \backslash A) \cap A = \varnothing$，由测度的有限可加性：

$$\mu(B) = \mu[(B \backslash A) \cup A] = \mu(B \backslash A) + \mu(A)$$

即：

$$\mu(B \backslash A) = \mu(B) - \mu(A)$$

于是 $\mu$ 具有可减性。 □

**Definition 1.18.** 设 $X$ 是一个集合，$\mathscr{F}$ 是 $X$ 的一些子集生成的 $\sigma$ 域，$\mu$ 是 $\mathscr{F}$ 上的测度。称 $(X, \mathscr{F}, \mu)$ 为测度空间 *(measure space)*。若 $A \in \mathscr{F}$ 且 $\mu(A) = 0$，则称 $A$ 为零测集 *(null set)*。若 $\mathscr{F}$ 中零测集的子集还属于 $\mathscr{A}$，则称测度空间 $(X, \mathscr{F}, \mu)$ 是完全测度空间 *(complete measure space)*。若测度空间 $(X, \mathscr{F}, P)$ 满足 $P(X) = 1$，则称其为概率空间 *(probability space)*，对应的 $P$ 叫做**概率测度**，$\mathscr{F}$ 中的元素叫做事件 *(event)*，$P(A)$ 叫做事件 $A$ 发生的概率。

### 半环上的测度

**Theorem 1.21.** 半环 $\mathscr{A}$ 上有有限可加性的非负集函数 $\mu$ 必有单调性和可减性。

*Proof.* (1) 设 $A, B \in \mathscr{A}$，$A \subset B$。因为 $\mathscr{A}$ 是一个半环，所以存在互不相交的 $C_1, C_2, \ldots, C_n \in \mathscr{A}$ 使得：

$$B \backslash A = \bigcup_{i=1}^{n} C_i$$

由 $\mu$ 的有限可加性可得：

$$\mu(B) = \mu\left[A \cup \left(\bigcup_{i=1}^{n} C_i\right)\right] = \mu(A) + \sum_{i=1}^{n} \mu(C_i) \geqslant \mu(A)$$

所以 $\mu$ 有单调性。

(2) 由 (1) 可得：

$$\mu(B\backslash A)=\mu\left(\bigcup_{i=1}^{n}C_i\right)=\sum_{i=1}^{n}\mu(C_i)=\mu(B)-\mu(A)$$

于是 $\mu$ 有可减性。 □

**Theorem 1.22.** 半环 $\mathscr{A}$ 上有可列可加性的非负集函数 $\mu$ 具有次可列可加性、下连续性和上连续性。

*Proof.* 因为 $\mu$ 具有可列可加性，所以：

$$\mu(\varnothing)=\sum_{n=1}^{+\infty}\mu(\varnothing)$$

于是 $\mu(\varnothing)=0$ 或 $\mu(\varnothing)=+\infty$。

**(1)** $\mu(\varnothing)=0$：此时 $\mu$ 是 $\mathscr{A}$ 上的测度。

**下连续性**：设 $\{A_n\}$ 是 $\mathscr{A}$ 中一个单调递增的集合序列且有 $A_n\uparrow A$，$A_0=\varnothing$。由 $\mu$ 的可列可加性和有限可加性（定理 1.20）可得：

$$\begin{aligned}\mu(A)&=\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\mu\left[\bigcup_{n=1}^{+\infty}(A_n\backslash A_{n-1})\right]=\sum_{n=1}^{+\infty}\mu(A_n\backslash A_{n-1})\\&=\lim_{n\to+\infty}\left[\sum_{i=1}^{n}\mu(A_i\backslash A_{i-1})\right]=\lim_{n\to+\infty}\mu(A_n)\end{aligned}$$

下连续性得证。

**上连续性**：设 $\{A_n\}$ 是 $\mathscr{A}$ 中一个单调递减的集合序列且有 $A_n\downarrow A$，$\mu(A_1)<+\infty$。由 $\mu$ 的可列可加性可得：

$$\mu(A_n)=\mu\left\{A\cup\left[\bigcup_{i=n}^{+\infty}(A_i\backslash A_{i+1})\right]\right\}=\mu(A)+\sum_{i=n}^{+\infty}\mu(A_i\backslash A_{i+1})$$

因为 $\mathscr{A}$ 是一个半环且 $\{A_n\}$ 是 $\mathscr{A}$ 中的集合序列，所以存在互不相交的 $C_1,C_2,\ldots,C_{k_n}\in\mathscr{A}$ 使得：

$$A_n\backslash A_{n+1}=\bigcup_{i=1}^{k_n}C_i$$

由 $\mu$ 的有限可加性可得：

$$\mu(A_n)=\mu(A)+\sum_{i=n}^{+\infty}\mu(A_i\backslash A_{i+1})=\mu(A)+\sum_{i=n}^{+\infty}\mu\left(\bigcup_{j=1}^{k_i}C_j\right)=\mu(A)+\sum_{i=n}^{+\infty}\sum_{j=1}^{k_i}\mu(C_j)$$

注意到：

$$\mu(A_1)=\mu(A)+\sum_{i=1}^{+\infty}\sum_{j=1}^{k_i}\mu(C_j)<+\infty$$

所以级数：

$$\sum_{i=1}^{+\infty}\sum_{j=1}^{k_i}\mu(C_j)$$

收敛。由级数收敛的必要性条件可得：

级数收敛的必要条件

$$\lim_{n\to+\infty}\left[\sum_{i=n}^{+\infty}\sum_{j=1}^{k_i}\mu(C_j)\right]=0$$

于是有：

$$\lim_{n\to+\infty}\mu(A_n)=\mu(A)+\lim_{n\to+\infty}\left[\sum_{i=n}^{+\infty}\sum_{j=1}^{k_i}\mu(C_j)\right]=\mu(A)$$

上连续性得证。

**次可列可加性**：设 $\{A_n\}$ 是 $\mathscr{A}$ 中的一个集合序列，由生成的定义可知 $\{A_n\}$ 也是 $r(\mathscr{A})$ 中的一个集合序列，令 $A_0=\varnothing$。由环的定义可得：

$$\bigcup_{i=1}^{n-1}A_i\in r(\mathscr{A}),\ A_n\backslash\bigcup_{i=1}^{n-1}A_i\in\mathscr{A}$$

因为环也是半环（定理 1.4），再根据定理 1.13可得存在互不相交的 $C_{n1},C_{n2},\dots,C_{nk_n}\in\mathscr{A}$ 使得：

$$A_n\backslash\bigcup_{i=1}^{n-1}A_i=\bigcup_{i=1}^{k_n}C_{ni}$$

同理，存在互不相交的 $D_{n1},D_{n2},\dots,D_{nl_n}\in\mathscr{A}$ 使得：

$$A_n\backslash\bigcup_{i=1}^{k_n}C_{ni}=\bigcup_{i=1}^{l_n}D_{ni}$$

显然 $C_{n1},C_{n2},\dots,C_{nk_n},D_{n1},D_{n2},\dots,D_{nl_n}$ 互不相交，同时有：

$$A_n=\left(\bigcup_{i=1}^{k_n}C_{ni}\right)\bigcup\left(\bigcup_{i=1}^{l_n}D_{ni}\right)$$

由 $\mu$ 的可列可加性与有限可加性（定理 1.20）：

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)&=\mu\left(\bigcup_{n=1}^{+\infty}\bigcup_{i=1}^{k_n}C_{ni}\right)=\sum_{n=1}^{+\infty}\sum_{i=1}^{k_n}\mu(C_{ni})\\&\leqslant\sum_{n=1}^{+\infty}\sum_{i=1}^{k_n}\mu(C_{ni})+\sum_{n=1}^{+\infty}\sum_{i=1}^{l_n}\mu(D_{ni})\\&=\sum_{n=1}^{+\infty}\left[\mu\left(\bigcup_{i=1}^{k_n}C_{ni}\right)+\mu\left(\bigcup_{i=1}^{l_n}D_{ni}\right)\right]\\&=\sum_{n=1}^{+\infty}\mu\left[\left(\bigcup_{i=1}^{k_n}C_{ni}\right)\bigcup\left(\bigcup_{i=1}^{l_n}D_{ni}\right)\right]=\sum_{n=1}^{+\infty}\mu(A_n)\end{aligned}$$

次可列可加性得证。

**(2)** $\mu(\varnothing)=+\infty$：此时显然满足所有条件。 □

**Theorem 1.23.** 半环上的测度具有单调性、可减性、次可列可加性、下连续性和上连续性。

*Proof.* 测度具有非负性和可列可加性，由定理 1.22可知半环上的测度具有次可列可加性、下连续性和上连续性。由定理 1.20可知可列可加性蕴含有限可加性，所以根据定理 1.21可知半环上的测度具有单调性和可减性。 □

**Theorem 1.24.** 设 $\mu$ 是环 $\mathscr{A}$ 上的非负集函数，则：

$$\begin{aligned}&(1)\mu\text{可列可加}\\ \Leftrightarrow&(2)\mu\text{次可列可加且有限可加}\\ \Leftrightarrow&(3)\mu\text{下连续且有限可加}\\ \Rightarrow&(4)\mu\text{上连续}\end{aligned}$$

*Proof.* 由定理 1.4可知环是半环，所以根据定理 1.23可得三个必要性成立，下分别证明两个充要性。

(1) 设 $\{A_n\}$ 是 $\mathscr{A}$ 中的一个互不相交的集合序列。由 $\mu$ 的次可列可加性可得：

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)\leqslant\sum_{n=1}^{+\infty}\mu(A_n)$$

由定理 1.21可知 $\mu$ 具有单调性，因为 $\mu$ 具有有限可加性，所以对任意的 $m\in\mathbb{N}^+$ 有：

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)\geqslant\mu\left(\bigcup_{n=1}^{m}A_n\right)=\sum_{n=1}^{m}\mu(A_n)$$

于是由极限的不等式性可得：

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)\geqslant\sum_{n=1}^{+\infty}\mu(A_n)$$

所以有：

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\sum_{n=1}^{+\infty}\mu(A_n)$$

即 $\mu$ 可列可加。

(2) 设 $\{A_n\}$ 是 $\mathscr{A}$ 中的一个互不相交的集合序列，显然有：

$$\bigcup_{i=1}^{n}A_i\uparrow\bigcup_{n=1}^{+\infty}A_n$$

由 $\mu$ 的有限可加性与下连续性可得：

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\lim_{n\to+\infty}\left[\mu\left(\bigcup_{i=1}^{n}A_i\right)\right]=\lim_{n\to+\infty}\left[\sum_{i=1}^{n}\mu(A_i)\right]=\sum_{n=1}^{+\infty}\mu(A_n)$$

即 $\mu$ 可列可加。 □

### 1.2.2 外测度

**Definition 1.19.** 设 $X$ 是一个集合，$\tau$ 是 $X$ 的所有子集构成的集族 $\mathscr{A}$ 到 $\overline{\mathbb{R}}$ 上的函数，如果：

1. $\tau(\varnothing)=0$；
2. 若 $A\subset B$ 且 $A,B\in\mathscr{A}$，则有 $\tau(A)\leqslant\tau(B)$；
3. $\tau$ 具有次可列可加性。

则称 $\tau$ 为 $X$ 上的外测度 *(exterior measure)*。

**Theorem 1.25.** 外测度具有次有限可加性。

*Proof.* 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，$\tau$ 是 $X$ 上的外测度。任取 $A_1,A_2,\dots,A_n\in\mathscr{A}$，由 $\tau$ 的次可列可加性可得：

$$\tau\left(\overset{n}{\underset{i=1}{\cup}}A_i\right)=\tau(A_1\cup A_2\cdots\cup A_n\cup\varnothing\cdots)\leqslant\sum_{i=1}^{n}\tau(A_i)+0=\sum_{i=1}^{n}\tau(A_i)$$

由 $A_1,A_2,\dots,A_n$ 的任意性，$\tau$ 具有次有限可加性。 □

**Theorem 1.26.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，则 $\mathscr{A}$ 上的测度 $\tau$ 一定是 $X$ 上的外测度。

*Proof.* 显然 $\mathscr{A}$ 是一个 $\sigma$ 域。由定理 1.5可知 $\mathscr{A}$ 是一个域，根据定理 1.4可知 $\mathscr{A}$ 是一个半环，所以 $\tau$ 是半环上的测度。由定理 1.23可知 $\tau$ 具有单调性和次可列可加性。因为 $\tau$ 是一个测度，所以 $\tau(\varnothing)=0$。综上，$\tau$ 是 $X$ 上的外测度。 □

**Theorem 1.27.** 设 $\mathscr{A}$ 是一个包含 $\varnothing$ 的集族，$\mu$ 是 $\mathscr{A}$ 上的一个非负集函数且满足 $\mu(\varnothing)=0$，若对于任意的 $A\in\mathscr{A}$ 有：

$$\tau(A)=\inf\left\{\sum_{n=1}^{+\infty}\mu(B_n):B_n\in\mathscr{A},\ A\subset\overset{+\infty}{\underset{n=1}{\cup}}B_n\right\}$$

则 $\tau$ 是一个外测度，称 $\tau$ 为由 $\mu$ 生成的外测度。

*Proof.* (1) 因为 $\varnothing=\overset{+\infty}{\underset{n=1}{\cup}}\varnothing$，所以：

$$0\leqslant\tau(\varnothing)\leqslant\sum_{n=1}^{+\infty}\mu(\varnothing)=0$$

所以 $\tau(\varnothing)=0$。

(2) 设 $A\subset B$ 且 $A,B\in\mathscr{A}$，对于满足条件：

$$B\subset\overset{+\infty}{\underset{n=1}{\cup}}B_n$$

的 $\{B_n\}$，自然有：

$$A\subset\overset{+\infty}{\underset{n=1}{\cup}}B_n$$

所以：

$$\tau(A) \leqslant \sum_{n=1}^{+\infty} \mu(B_n)$$

对右边取下确界即有 $\tau(A) \leqslant \tau(B)$。

(3) 设 $\{A_n\}$ 是 $\mathscr{A}$ 中的一个集合序列。若存在 $n_0 \in \mathbb{N}^+$ 使得 $\tau(A_{n_0}) = +\infty$，则由 $\tau$ 的定义和 (2) 可得：

$$\tau\left(\bigcup_{n=1}^{+\infty} A_n\right) \leqslant +\infty = \tau(A_{n_0}) \leqslant \sum_{n=1}^{+\infty} \tau(A_n)$$

即 $\tau$ 具有次可列可加性。

若 $\tau(A_n) < +\infty$ 对 $n \in \mathbb{N}^+$ 都成立，任取 $\varepsilon > 0$，对任意的 $n \in \mathbb{N}^+$，存在 $\mathscr{A}$ 中的一个集合序列 $\{B_{ni}\}$ 使得：

$$A_n \subset \bigcup_{i=1}^{+\infty} B_{ni},\ \sum_{i=1}^{+\infty} \mu(B_{ni}) < \tau(A_n) + \frac{\varepsilon}{2^n}$$

于是：

$$\sum_{n=1}^{+\infty}\sum_{i=1}^{+\infty} \mu(B_{ni}) < \sum_{n=1}^{+\infty} \tau(A_n) + \varepsilon$$

由 $\{B_{n_i}\}$ 的取法，显然：

$$\tau\left(\bigcup_{n=1}^{+\infty} A_n\right) \leqslant \sum_{n=1}^{+\infty}\sum_{i=1}^{+\infty} \mu(B_{ni})$$

由 $\varepsilon$ 的任意性可得：

$$\tau\left(\bigcup_{n=1}^{+\infty} A_n\right) \leqslant \sum_{n=1}^{+\infty} \tau(A_n)$$

即 $\tau$ 具有次可列可加性。

综上，$\tau$ 是一个外测度。 □

**Definition 1.20** (Caratheodory condition)**.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，$\tau$ 是 $X$ 上的外测度。称满足条件：

$$\tau(T) = \tau(T \cap A) + \tau(T \cap A^c),\ \forall\, T \in \mathscr{A}$$

的集合 $A \in \mathscr{A}$ 为 $\tau$可测集 *(measurable set)*。将由所有 $\tau$ 可测集构成的集族记作 $\mathscr{A}_\tau$。

**Lemma 1.2.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，$\tau$ 是 $X$ 上的外测度。集合 $E \in \mathscr{A}_\tau$ 的充要条件是对与 $\forall\, A \subset E,\ \forall\, B \subset E^c$，总有：

$$\tau(A \cup B) = \tau(A) + \tau(B)$$

*Proof.* 必要性：对任意的 $A \subset E,\ \forall\, B \subset E^c$，取 $T = A \cup B$，因为 $E \in \mathscr{A}_\tau$，那么对于这个 $T$，应有：

$$\tau(A \cup B) = \tau(T) = \tau(T \cap E) + \tau(T \cap E^c) = \tau(A) + \tau(B)$$

充分性：对任意的 $T \in \mathscr{A}$，$\exists\, A \subset E, \exists\, B \subset E^c$，使得 $T = A \cup B$，那么就有：

$$\tau(T) = \tau(A \cup B) = \tau(A) + \tau(B) = \tau(T \cap E) + \tau(T \cap E^c)$$

由 $T$ 的任意性，$E \in \mathscr{A}_\tau$。 □

**Property 1.2.1.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，$\tau$ 是 $X$ 上的外测度。$\mathscr{A}_\tau$ 具有如下性质：

1. $\varnothing, X \in \mathscr{A}_\tau$；
2. $S \in \mathscr{A}_\tau$ 的充要条件是 $S^c \in \mathscr{A}_\tau$；
3. 设 $S_1, S_2, \ldots, S_n$ 是 $\tau$ 可测集，则有 $\bigcup\limits_{i=1}^{n} S_i \in \mathscr{A}_\tau$，并且当 $S_1, S_2, \ldots, S_n$ 互不相交时可得：
$$\tau\left(\bigcup_{i=1}^{n} S_i\right) = \sum_{i=1}^{n} \tau(S_i)$$
4. 设 $S_1, S_2, \ldots, S_n$ 是 $\tau$ 可测集，则 $\bigcap\limits_{i=1}^{n} S_i \in \mathscr{A}_\tau$；
5. 设 $S_1, S_2 \in \mathscr{A}_\tau$，则 $S_1 \backslash S_2 \in \mathscr{A}_\tau$；
6. 设 $\{S_n\}$ 是一列 $\tau$ 可测集，则 $\bigcup\limits_{n=1}^{+\infty} S_n \in \mathscr{A}_\tau$，并且当 $\{S_n\}$ 互不相交时可得：
$$\tau\left(\bigcup_{n=1}^{+\infty} S_i\right) = \sum_{n=1}^{+\infty} \tau(S_i)$$
7. 设 $\{S_n\}$ 是一列 $\tau$ 可测集，则 $\bigcap\limits_{n=1}^{+\infty} S_n \in \mathscr{A}_\tau$。

*Proof.* (1) 代入定义直接可得。

(2) 若 $S \in \mathscr{A}_\tau$，对任意的 $T \in \mathscr{A}$，则有：

$$\begin{aligned}\tau(T) = \tau(T \cap S) + \tau(T \cap S^c) &= \tau[T \cap (S^c)^c] + \tau(T \cap S^c) \\ &= \tau(T \cap S^c) + \tau[T \cap (S^c)^c]\end{aligned}$$

(3) 因为 $S_1 \in \mathscr{A}_\tau$，对任意的 $T$ 都有：

$$\tau(T) = \tau(T \cap S_1) + \tau(T \cap S_1^c) \tag{1.1}$$

因为 $S_2 \in \mathscr{A}_\tau$，对于 $\tau(T \cap S_1^c)$ 有：

$$\tau(T \cap S_1^c) = \tau[(T \cap S_1^c) \cap S_2] + \tau[(T \cap S_1^c) \cap S_2^c] \tag{1.2}$$

将(1.2)式代入(1.1)式，再由定理 1.1，得到：

$$\begin{aligned}\tau(T) &= \tau(T \cap S_1) + \tau[(T \cap S_1^c) \cap S_2] + \tau[(T \cap S_1^c) \cap S_2^c] \\ &= \tau(T \cap S_1) + \tau[(T \cap S_1^c) \cap S_2] + \tau[T \cap (S_1 \cup S_2)^c]\end{aligned}$$

由于 $T\cap S_1\subset S_1$，$(T\cap S_1^c)\cap S_2\subset S_1^c$，满足引理 1.2条件，因此上式的前两项可以合并：

$$\begin{aligned}\tau(T\cap S_1)+\tau[(T\cap S_1^c)\cap S_2]&=\tau[(T\cap S_1)\cup(T\cap S_1^c\cap S_2)]\\&=\tau\{T\cap[S_1\cup(S_1^c\cap S_2)]\}\\&=\tau\{T\cap[(S_1\cup S_1^c)\cap(S_1\cup S_2)]\}\\&=\tau[T\cap(S_1\cup S_2)]\end{aligned}$$

那么就有：

$$\tau(T)=\tau[T\cap(S_1\cup S_2)]+\tau[T\cap(S_1\cup S_2)^c]$$

由 $T$ 的任意性，$S_1\cup S_2\in\mathscr{A}_\tau$。

当 $S_1\cap S_2=\varnothing$ 时，显然 $S_2\subset S_1^c$，那么就有 $T\cap S_2\subset S_1^c$，由引理 1.2：

$$\begin{aligned}\tau[T\cap(S_1\cup S_2)]&=\tau[(T\cap S_1)\cup(T\cap S_2)]\\&=\tau(T\cap S_1)+\tau(T\cap S_2)\end{aligned}$$

由数学归纳法，$\bigcup\limits_{i=1}^{n}S_i\in\mathscr{A}_\tau$ 且当 $S_1,S_2,\dots,S_n$ 互不相交时可得：

$$\tau\left(\bigcup_{i=1}^{n}S_i\right)=\sum_{i=1}^{n}\tau(S_i)$$

(4) 由定理 1.1以及 (2)(3) 直接得到：

$$\bigcap_{i=1}^{n}S_i=\left(\bigcup_{i=1}^{n}S_i^c\right)^c\in\mathscr{A}_\tau$$

(5) $S_1\backslash S_2=S_1\cap S_2^c$，由 (2)(4) 可知 $S_1\backslash S_2\in\mathscr{A}_\tau$。

(6) 设 $\{S_n\}$ 互不相交。由 (3) 可得对任意的 $n$，$\bigcup\limits_{i=1}^{n}S_i\in\mathscr{A}_\tau$，那么对任意的 $T\in\mathscr{A}$，就有（第一行到第二行利用外测度的单调性，第二行到第三行利用 (3)）：

$$\begin{aligned}\tau(T)&=\tau\left[T\cap\left(\bigcup_{i=1}^{n}S_i\right)\right]+\tau\left[T\cap\left(\bigcup_{i=1}^{n}S_i\right)^c\right]\\&\geqslant\tau\left[T\cap\left(\bigcup_{i=1}^{n}S_i\right)\right]+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\\&=\sum_{i=1}^{n}\tau(T\cap S_i)+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\end{aligned}$$

令 $n\to+\infty$，有（第一行到第二行利用极限的不等式性，第二行到第三行利用外测度的次可列可加性）：

$$\begin{aligned}\tau(T)&\geqslant\sum_{i=1}^{n}\tau(T\cap S_i)+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\\&\geqslant\sum_{i=1}^{+\infty}\tau(T\cap S_i)+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\\&\geqslant\tau\left[\bigcup_{i=1}^{+\infty}(T\cap S_i)\right]+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\\&=\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)\right]+\tau\left[T\cap\left(\bigcup_{n=1}^{+\infty}S_n\right)^c\right]\end{aligned}\tag{1.3}$$

又因：

$$T=\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\right]\cup\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)^c\right]$$

由外测度的次有限可加性（定理 1.25）可得：

$$\tau(T)\leqslant\tau\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\right]+\tau\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)^c\right]$$

因此：

$$\tau(T)=\tau\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\right]+\tau\left[T\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)^c\right]$$

由 $T$ 的任意性，$\overset{+\infty}{\underset{n=1}{\cup}}S_n$ 可测。

令 $T=\overset{+\infty}{\underset{n=1}{\cup}}S_n$，代入公式 (1.3) 式，则：

$$\begin{aligned}\tau\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)&\geqslant\sum_{i=1}^{+\infty}\tau\left[\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\cap S_i\right]+\tau\left[\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\cap\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)^c\right]\\&=\sum_{i=1}^{+\infty}\tau(S_i)\end{aligned}$$

但是由外测度的次可列可加性有：

$$\tau\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)\leqslant\sum_{n=1}^{+\infty}\tau(S_n)$$

因此：

$$\tau\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n\right)=\sum_{n=1}^{+\infty}\tau(S_n)$$

若 $\{S_n\}$ 不满足互不相交，则 $\overset{+\infty}{\underset{n=1}{\cup}}S_n$ 可被表示为互不相交的可数个集合的并：

$$\overset{+\infty}{\underset{n=1}{\cup}}S_n=S_1\cup(S_2\backslash S_1)\cup[S_3\backslash(S_1\cup S_2)]\cdots$$

由 (5) 和之前 $\{S_n\}$ 互不相交时的论述即可得 $\overset{+\infty}{\underset{n=1}{\cup}}S_n\in\mathscr{A}_\tau$。

(7) 由定理 1.1和 (2) 可知：

$$\overset{+\infty}{\underset{n=1}{\cap}}S_n=\left(\overset{+\infty}{\underset{n=1}{\cup}}S_n^c\right)^c\in\mathscr{A}_\tau$$

□

**Theorem 1.28.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 的所有子集构成的集族，$\tau$ 是 $X$ 上的外测度。$(X,\mathscr{A}_\tau,\tau)$ 是一个完全测度空间。

*Proof.* 由性质 1.2.1(6) 和外测度的定义可知 $\tau$ 是 $\mathscr{A}_\tau$ 上的测度。任取 $A\in\mathscr{A}$ 满足 $\tau(A)=0$，对任何的 $E\in\mathscr{A}$，由外测度的单调性和非负性可得 $\tau(E\cap A)=0$，于是由外测度的单调性可得：

$$\tau(E)\geqslant\tau(E\cap A^c)=\tau(E\cap A)+\tau(E\cap A^c)$$

因为 $E=(E\cap A)\cup(E\cap A^c)$，由外测度的次可列可加性可得：

$$\tau(E)\leqslant\tau(E\cap A)+\tau(E\cap A^c)$$

于是有：

$$\tau(E)=\tau(E\cap A)+\tau(E\cap A^c)$$

由 $E$ 的任意性，$A\in\mathscr{A}_\tau$，即 $\tau$ 的零测集都属于 $\mathscr{A}_\tau$。由外测度的单调性，零测集的子集都是零测集，所以 $\tau$ 的零测集的子集都属于 $\mathscr{A}_\tau$，$(X,\mathscr{A}_\tau,\tau)$ 是一个完全测度空间。 □

### 1.2.3 测度的扩张

**Definition 1.21.** 设 $\mu,\tau$ 分别为集族 $\mathscr{A},\mathscr{B}$ 上的测度，并且有 $\mathscr{A}\subset\mathscr{B}$。如果对任意的 $A\in\mathscr{A}$ 都有 $\mu(A)=\tau(A)$，则称 $\tau$ 是 $\mu$ 在 $\mathscr{B}$ 上的扩张 *(extension)*。

**Lemma 1.3.** 设 $X$ 是一个集合，$\mathscr{A}$ 是 $X$ 上的一个 $\pi$ 系，如果 $\sigma(\mathscr{A})$ 上的测度 $\mu,\tau$ 满足：

1. 对任意的 $A\in\mathscr{A}$，有 $\mu(A)=\tau(B)$；
2. 存在 $\mathscr{A}$ 中互不相交的集合序列 $\{A_n\}$ 使得：

$$\bigcup_{n=1}^{+\infty}A_n=X,\ \mu(A_n)<+\infty,\ \forall\, n\in\mathbb{N}^+$$

则对任何的 $A\in\sigma(\mathscr{A})$，有 $\mu(A)=\tau(A)$。

*Proof.* 对任意的 $B\in\mathscr{A}$ 且 $\mu(B)<+\infty$，令：

$$\mathscr{C}=\{A\in\sigma(\mathscr{A}):\mu(A\cap B)=\tau(A\cap B)\}$$

下证明 $\mathscr{C}$ 是一个 $\lambda$ 系。

(1) 因为 $\sigma(\mathscr{A})$ 是一个 $\sigma$ 域，所以 $X\in\sigma(\mathscr{A})$，于是有 $\mu(X\cap B)=\mu(B)$, $\tau(X\cap B)=\tau(B)$。由条件 (1)，因为 $B\in\mathscr{A}$，所以 $\mu(X\cap B)=\tau(X\cap B)=\mu(B)=\tau(B)$，于是 $X\in\mathscr{C}$。

(2) 任取 $C,D\in\mathscr{C}$ 且有 $C\subset D$，于是有 $\mu(C\cap B)=\tau(C\cap B)$, $\mu(D\cap B)=\tau(D\cap B)$。考虑 $D\backslash C$，则：

$$\mu[(D\backslash C)\cap B]=\mu[(D\cap B)\backslash(C\cap B)]$$

因为 $C\subset D$，所以 $C\cap B\subset D\cap B$，由测度的有限可加性即可得：

$$\mu[(D\backslash C)\cap B]=\mu[(D\cap B)\backslash(C\cap B)]=\mu(D\cap B)-\mu(C\cap B)$$

因为 $B\in\mathscr{A}$，所以 $D\cap B,C\cap B\in\mathscr{A}$，由测度的有限可加性即可得到：

$$\mu[(D\backslash C)\cap B]=\mu(D\cap B)-\mu(C\cap B)=\tau(D\cap B)-\tau(C\cap B)=\tau[(D\backslash C)\cap B]$$

所以 $D\backslash C\in\mathscr{C}$。

(3) 任取 $\mathscr{C}$ 中一个单调不减的集合序列 $\{C_n\}$，则有 $\mu(C_n \cap B) = \tau(C_n \cap B)$ 对 $n \in \mathbb{N}^+$ 成立。令 $C_0 = \varnothing \in \mathscr{C}$，由测度的可列可加性、有限可加性即可得到：

$$\begin{aligned}
\mu\left[\left(\bigcup_{n=1}^{+\infty} C_n\right) \cap B\right] &= \mu\left\{\left[\bigcup_{n=1}^{+\infty}(C_n \backslash C_{n-1})\right] \cap B\right\} = \mu\left\{\bigcup_{n=1}^{+\infty}[(C_n \backslash C_{n-1}) \cap B]\right\} \\
&= \sum_{n=1}^{+\infty} \mu[(C_n \backslash C_{n-1}) \cap B] = \sum_{n=1}^{+\infty} \mu[(C_n \cap B) \backslash (C_{n-1} \cap B)] \\
&= \sum_{n=1}^{+\infty}[\mu(C_n \cap B) - \mu(C_{n-1} \cap B)] = \sum_{n=1}^{+\infty}[\tau(C_n \cap B) - \tau(C_{n-1} \cap B)] \\
&= \sum_{n=1}^{+\infty} \tau[(C_n \cap B) \backslash (C_{n-1} \cap B)] = \sum_{n=1}^{+\infty} \tau[(C_n \backslash C_{n-1}) \cap B] \\
&= \tau\left\{\bigcup_{n=1}^{+\infty}[(C_n \backslash C_{n-1}) \cap B]\right\} = \tau\left\{\left[\bigcup_{n=1}^{+\infty}(C_n \backslash C_{n-1})\right] \cap B\right\} \\
&= \tau\left[\left(\bigcup_{n=1}^{+\infty} C_n\right) \cap B\right]
\end{aligned}$$

于是：

$$\bigcup_{n=1}^{+\infty} C_n \in \mathscr{C}$$

综上，$\mathscr{C}$ 是一个 $\lambda$ 系。

因为 $\mathscr{A}$ 是一个 $\pi$ 系，对交封闭，所以 $\mathscr{A} \subset \mathscr{C}$。由推论 1.2可知 $\sigma(\mathscr{A}) \subset \mathscr{C}$，于是任意的 $A \in \sigma(\mathscr{A})$ 对任意的 $B \in \mathscr{A}$ 且 $\mu(B) < +\infty$，有：

$$\mu(A \cap B) = \tau(A \cap B)$$

取条件 (2) 中的集合序列 $\{A_n\}$，由测度的可列可加性可得：

$$\begin{aligned}
\mu(A) &= \mu\left[A \cap \left(\bigcup_{n=1}^{+\infty} A_n\right)\right] = \mu\left[\bigcup_{n=1}^{+\infty}(A_n \cap A)\right] = \sum_{n=1}^{+\infty} \mu(A_n \cap A) = \sum_{n=1}^{+\infty} \tau(A_n \cap A) \\
&= \tau\left[\bigcup_{n=1}^{+\infty}(A_n \cap A)\right] = \tau\left[A \cap \left(\bigcup_{n=1}^{+\infty} A_n\right)\right] = \tau(A)
\end{aligned}$$

□

**Theorem 1.29.** 对于半环 $\mathscr{A}$ 上的测度 $\mu$，存在 $\sigma(\mathscr{A})$ 上 $\mu$ 的扩张 $\tau$。若存在 $\mathscr{A}$ 中互不相交的集合序列 $\{A_n\}$ 使得：

$$\bigcup_{n=1}^{+\infty} A_n = X,\ \mu(A_n) < +\infty,\ \forall\, n \in \mathbb{N}^+$$

则 $\tau$ 是唯一的且是由 $\mu$ 生成的外测度。

*Proof.* 取 $\tau$ 为由 $\mu$ 生成的外测度。证明分成以下三步：

1. 证明对于任意的 $A \in \mathscr{A}$，有 $\mu(A) = \tau(A)$；

2. 证明 $\tau$ 是 $\sigma(\mathscr{A})$ 上的测度；

(a) 对任何的 $A, B \in \mathscr{A}$，有：

$$\tau(B) \geqslant \tau(B \cap A) + \tau(B \cap A^c)$$

(b) 设 $\mathscr{F}$ 是 $X$ 的所有子集构成的集族，$\mathscr{F}_\tau$ 是 $\tau$ 可测集，则 $\mathscr{A} \subset \mathscr{F}_\tau$。

3. 满足定理条件时 $\tau$ 是唯一的。

(1) 取任意的 $A \in \mathscr{A}$，对任意满足 $A \subset \bigcup\limits_{n=1}^{+\infty} A_n$ 的 $\mathscr{A}$ 中的集合序列 $\{A_n\}$，由半环上测度的次可列可加性和单调性（定理 1.23）可得：

$$\mu(A) = \mu\left[A \cap \left(\bigcup_{n=1}^{+\infty} A_n\right)\right] = \mu\left[\bigcup_{n=1}^{+\infty} (A \cap A_n)\right] \leqslant \sum_{n=1}^{+\infty} \mu(A \cap A_n) \leqslant \sum_{n=1}^{+\infty} \mu(A_n)$$

由下确界的不等式性可得：

$$\mu(A) \leqslant \inf\left\{\sum_{n=1}^{+\infty} \mu(A_n) : A_n \in \mathscr{A},\ A \subset \bigcup_{n=1}^{+\infty} A_n\right\} = \tau(A)$$

再取 $B_1 = A,\ B_n = \varnothing,\ \forall\, n \geqslant 2$，有：

$$\mu(A) = \sum_{n=1}^{+\infty} \mu(B_n) \geqslant \inf\left\{\sum_{n=1}^{+\infty} \mu(A_n) : A_n \in \mathscr{A},\ A \subset \bigcup_{n=1}^{+\infty} A_n\right\} = \tau(A)$$

于是有 $\mu(A) = \tau(A)$。

(2.a) 因为 $\mathscr{A}$ 是半环，所以存在互不相交的 $\{C_n\} \subset \mathscr{A}$ 使得：

$$B \cap A^c = B \backslash A = \bigcup_{i=1}^{n} C_i$$

由 (1) 可得 $\mu(B) = \tau(B)$，根据半环对交的封闭性与测度的有限可加性可得：

$$\begin{aligned}\tau(B) = \mu(B) &= \mu[(B \cap A) \cup (B \cap A^c)] = \mu(B \cap A) + \mu(B \cap A^c) \\ &= \mu(B \cap A) + \mu\left(\bigcup_{i=1}^{n} C_i\right) = \mu(B \cap A) + \sum_{i=1}^{n} \mu(C_i)\end{aligned}$$

因为：

$$B \cap A^c \subset \bigcup_{i=1}^{n} C_i$$

而半环对包含的差封闭

没有证完，涉及到半环的现代式定义

(2.b) 即要证对任意的 $A \in \mathscr{A}$ 和任意的 $B \in \mathscr{F}$，有：

$$\tau(B) = \tau(B \cap A) + \tau(B \cap A^c)$$

□

## 1.3 可测映射与可测函数

### 1.3.1 可测映射

**Definition 1.22.** 称 $X$ 和其上的一个 $\sigma$ 域 $\mathscr{A}$ 为可测空间 *(measurable space)*，记为 $(X,\mathscr{A})$。

**Definition 1.23.** 设 $(X,\mathscr{A})$ 和 $(Y,\mathscr{B})$ 为可测空间，$f$ 是一个 $X$ 到 $Y$ 的映射。如果 $f^{-1}(\mathscr{B})\subset\mathscr{A}$，则称 $f$ 为从 $(X,\mathscr{A})$ 到 $(Y,\mathscr{B})$ 的可测映射 *(measurable map)*，称 $\sigma(f)=f^{-1}(\mathscr{B})$ 为使映射 $f$ 可测的最小 $\sigma$ 域。

**Lemma 1.4.** 设 $X,Y$ 为两个集合，$f$ 为一个 $X$ 到 $Y$ 的映射，$\mathscr{A}\subset Y$，则：

$$\sigma[f^{-1}(\mathscr{A})]=f^{-1}[\sigma(\mathscr{A})]$$

*Proof.* 先证 $f^{-1}[\sigma(\mathscr{A})]$ 是一个 $\sigma$ 域。

(1) 因为 $\sigma(\mathscr{A})$ 是一个 $\sigma$ 域，所以 $Y\in\sigma(\mathscr{A})$，于是 $X\in f^{-1}[\sigma(\mathscr{A})]$。

(2) 任取 $A\in f^{-1}[\sigma(\mathscr{A})]$，设 $f(A)=B$。由**??**(3) 可得，$A^c=[f^{-1}(B)]^c=f^{-1}(B^c)$。因为 $\sigma(\mathscr{A})$ 是一个 $\sigma$ 域，$B\in\sigma(\mathscr{A})$，所以 $B^c\in\sigma(\mathscr{A})$，所以 $A^c\in f^{-1}[\sigma(\mathscr{A})]$。由 $A$ 的任意性，$f^{-1}[\sigma(\mathscr{A})]$ 对补的运算封闭。

(3) 任取集合序列 $\{A_n\}\subset f^{-1}[\sigma(\mathscr{A})]$，$f(A_n)=B_n\in\sigma(\mathscr{A})$，由**??**(4) 可得：

$$\bigcup_{n=1}^{+\infty}A_n=\bigcup_{n=1}^{+\infty}f^{-1}(B_n)=f^{-1}\left(\bigcup_{n=1}^{+\infty}B_n\right)$$

因为 $\sigma(\mathscr{A})$ 是一个 $\sigma$ 域，$B_n\in\sigma(\mathscr{A}),\ \forall\, n\in\mathbb{N}^+$，所以：

$$\bigcup_{n=1}^{+\infty}B_n\in\sigma(\mathscr{A})$$

于是 $\bigcup\limits_{n=1}^{+\infty}A_n\in f^{-1}[\sigma(\mathscr{A})]$。由 $\{A_n\}$ 的任意性，$f^{-1}[\sigma(\mathscr{A})]$ 对可列并的运算封闭。

综上，$f^{-1}[\sigma(\mathscr{A})]$ 是一个 $\sigma$ 域。由生成的定义，$\mathscr{A}\subset\sigma(\mathscr{A})$，由**??**(2) 可得 $f^{-1}(\mathscr{A})\subset f^{-1}[\sigma(\mathscr{A})]$，即 $f^{-1}[\sigma(\mathscr{A})]$ 是一个包含 $f^{-1}(\mathscr{A})$ 的 $\sigma$ 域，所以 $\sigma[f^{-1}(\mathscr{A})]\subset f^{-1}[\sigma(\mathscr{A})]$。

令：

$$\mathscr{B}=\{B\subset Y:f^{-1}(B)\in\sigma[f^{-1}(\mathscr{A})]\}$$

下证 $\mathscr{B}$ 是一个 $\sigma$ 域。

(1) 因为 $\sigma[f^{-1}(\mathscr{A})]$ 是一个 $\sigma$ 域，所以 $X\in\sigma[f^{-1}(\mathscr{A})]$。由**??**(1) 可得 $f^{-1}(Y)=X\in\sigma[f^{-1}(\mathscr{A})]$，所以 $Y\in\mathscr{B}$。

(2) 任取 $B\in\mathscr{B}$。由**??**(3) 可得 $f^{-1}(B^c)=[f^{-1}(B)]^c$。因为 $\sigma[f^{-1}(\mathscr{A})]$ 是一个 $\sigma$ 域，$f^{-1}(B)\in\sigma[f^{-1}(\mathscr{A})]$，所以 $f^{-1}(B^c)=[f^{-1}(B)]^c\in\sigma[f^{-1}(\mathscr{A})]$，即 $B^c\in\mathscr{B}$。由 $B$ 的任意性，$\mathscr{B}$ 对补的运算封闭。

(3) 任取 $\{B_n\}\subset\mathscr{B}$。由**??**(4) 可得：

$$f^{-1}\left(\bigcup_{n=1}^{+\infty}B_n\right)=\bigcup_{n=1}^{+\infty}f^{-1}(B_i)$$

因为 $B_n \in \mathscr{B}, \forall\, n \in \mathbb{N}^+$，所以 $f^{-1}(B_n) \in \sigma[f^{-1}(\mathscr{A})], \forall\, n \in \mathbb{N}^+$。因为 $\sigma[f^{-1}(\mathscr{A})]$ 是一个 $\sigma$ 域，所以：

$$f^{-1}\left(\bigcup_{n=1}^{+\infty} B_n\right) = \bigcup_{n=1}^{+\infty} f^{-1}(B_n) \in \sigma[f^{-1}(\mathscr{A})]$$

于是 $\bigcup\limits_{n=1}^{+\infty} B_n \in \mathscr{B}$。由 $\{B_n\}$ 的任意性，$\mathscr{B}$ 对可列并的运算封闭。

综上，$\mathscr{B}$ 是一个 $\sigma$ 域。因为 $f^{-1}(\mathscr{A}) \subset \sigma[f^{-1}(\mathscr{A})]$，所以 $\mathscr{A} \subset \mathscr{B}$。由生成的定义，$\sigma(\mathscr{A}) \subset \mathscr{B}$。任取 $C \in f^{-1}[\sigma(\mathscr{A})]$，则存在 $A \in \sigma(\mathscr{A})$ 使得 $C = f^{-1}(A)$，因为 $\sigma(\mathscr{A}) \subset \mathscr{B}$，所以 $A \in \mathscr{B}$，于是 $C = f^{-1}(A) \in \sigma[f^{-1}(\mathscr{A})]$。由 $C$ 的任意性，$f^{-1}[\sigma(\mathscr{A})] \subset \sigma[f^{-1}(\mathscr{A})]$。

综上，$f^{-1}[\sigma(\mathscr{A})] = \sigma[f^{-1}(\mathscr{A})]$。 □

**Theorem 1.30.** 设 $\mathscr{B}$ 是 $Y$ 上的任一集族，$(X, \mathscr{A}), (Y, \sigma(\mathscr{B}))$ 是两个可测空间，则映射 $f$ 为一个 $(X, \mathscr{A})$ 到 $(Y, \sigma(\mathscr{B}))$ 的可测映射的充分必要条件为：

$$f^{-1}(\mathscr{B}) \subset \mathscr{A}$$

*Proof.* 由可测映射的定义：

$$f \text{ 为一个 } (X, \mathscr{A}) \text{ 到 } (Y, \sigma(\mathscr{B})) \text{ 的可测映射} \Leftrightarrow f^{-1}[\sigma(\mathscr{B})] \subset \mathscr{A} \Leftrightarrow \sigma[f^{-1}(\mathscr{B})] \subset \mathscr{A}$$

**(1) 必要性：**若 $\sigma[f^{-1}(\mathscr{B})] \subset \mathscr{A}$，则 $f^{-1}(\mathscr{B}) \subset \sigma[f^{-1}(\mathscr{B})] \subset \mathscr{A}$。

**(2) 充分性：**若 $f^{-1}(\mathscr{B}) \subset \mathscr{A}$，因为 $\mathscr{A}$ 是一个 $\sigma$ 域，由生成的定义可知 $\sigma[f^{-1}(\mathscr{B})] \subset \mathscr{A}$。 □

**Theorem 1.31.** 设 $g$ 是可测空间 $(X, \mathscr{A})$ 到可测空间 $(Y, \mathscr{B})$ 的可测映射，$f$ 是可测空间 $(Y, \mathscr{B})$ 到可测空间 $(Z, \mathscr{C})$ 的可测映射，则 $f \circ g$ 是 $(X, \mathscr{A})$ 到 $(Z, \mathscr{C})$ 的可测映射。

*Proof.* 因为 $g$ 是可测空间 $(X, \mathscr{A})$ 到可测空间 $(Y, \mathscr{B})$ 的可测映射，所以 $g^{-1}(\mathscr{B}) \subset \mathscr{A}$。因为 $f$ 是可测空间 $(Y, \mathscr{B})$ 到可测空间 $(Z, \mathscr{C})$ 的可测映射，所以 $f^{-1}(\mathscr{C}) \subset \mathscr{B}$。由**??**(2) 可得：

$$(f \circ g)^{-1}(\mathscr{C}) = g^{-1}[f^{-1}(\mathscr{C})] \subset g^{-1}(\mathscr{B}) \subset \mathscr{A}$$ □

### 1.3.2 可测函数

#### 可测函数的定义

**Definition 1.24.** 从可测空间 $(X, \mathscr{A})$ 到可测空间 $(\overline{\mathbb{R}}, \mathcal{B}_{\overline{\mathbb{R}}})$ 的可测映射称为 $(X, \mathscr{A})$ 上的可测函数 *(measurable function)*。特别的，从可测空间 $(X, \mathscr{A})$ 到可测空间 $(\mathbb{R}, \mathcal{B})$ 的可测映射称为 $(X, \mathscr{A})$ 上的有限值可测函数或随机变量 *(random variable, r.v.)*。

**Theorem 1.32.** $f$ 是可测空间 $(X, \mathscr{A})$ 上的可测函数的充要条件为：

1. 对任意的 $a \in \mathbb{R}$，$\{f < a\} \in \mathscr{A}$。

2. 对任意的 $a \in \mathbb{R}$，$\{f \leqslant a\} \in \mathscr{A}$。

3. 对任意的 $a \in \mathbb{R}$，$\{f > a\} \in \mathscr{A}$。

4. 对任意的 $a\in\mathbb{R}$，$\{f\geqslant a\}\in\mathscr{A}$。

*Proof.* (1) 由定理 1.19中 $\mathcal{B}_{\overline{\mathbb{R}}}$ 的等价定义 (1) 以及定理 1.30可得：

$$\begin{aligned} f\text{是可测函数} &\Leftrightarrow f^{-1}(\mathcal{B}_{\overline{\mathbb{R}}})\subset\mathscr{A}\\ &\Leftrightarrow \forall\,\alpha\in\mathbb{R},\ f^{-1}\Big([-\infty,\alpha)\Big)\in\mathscr{A}\\ &\Leftrightarrow \forall\,\alpha\in\mathbb{R},\ \{f<a\}\in\mathscr{A}\end{aligned}$$

(2)(3)(4) 同理，只需在 (1) 的第一行到第二行的过程中使用定理 1.19涉及到的对应的 $\mathcal{B}_{\overline{\mathbb{R}}}$ 的等价定义。 □

**Corollary 1.4.** 设 $f(x)$ 和 $g(x)$ 为可测空间 $(X,\mathscr{A})$ 上的可测函数，则 $\{f<g\},\{f\leqslant g\},\{f=g\}\in\mathscr{A}$。

*Proof.* 因为：

$$\{f<g\}=\bigcup_{r\in\mathbb{Q}}[\{f<r\}\cap\{g>r\}]$$

而：

$$\{f<r\}\cap\{g>r\}=\Big(\{f\geqslant r\}\cup\{g\leqslant r\}\Big)^c$$

由定理 1.32可知右式在 $\mathscr{A}$ 中，于是有：

$$\{f<g\}\in\mathscr{A}$$

由 $f(x)$ 与 $g(x)$ 的对称性，$\{g<f\}\in\mathscr{A}$，那么就有 $\{f\leqslant g\}=\{g<f\}^c\in\mathscr{A}$。因为：

$$\begin{aligned}\{f=g\}&=\{f\leqslant g\}\backslash\{f<g\}\\ &=\{f\leqslant g\}\cap\{f<g\}^c\\ &=\Big(\{f\leqslant g\}^c\cup\{f<g\}\Big)^c\end{aligned}$$

所以 $\{f=g\}\in\mathscr{A}$。 □

### 可测函数的运算

**Theorem 1.33.** 可测空间 $(X,\mathscr{A})$ 上集合 $A\in\mathscr{A}$ 的指示函数 $I_A$ 是可测函数。

**Lemma 1.5.** $(X,\mathscr{A})$ 是一个可测空间。对任何 $a,b\in\overline{\mathbb{R}}$ 和任何的 $A,B\in\mathscr{A}$，只要 $a+b$ 有意义，那么 $aI_A+bI_B$ 是可测函数。

**Theorem 1.34.** 若 $f,g$ 是可测空间 $(X,\mathscr{A})$ 上的可测函数，则：

1. 对任意的 $\alpha,\beta\in\overline{\mathbb{R}}$，若对任意的 $x\in X$，$\alpha f+\beta g$ 有意义，则 $\alpha f+\beta g$ 是可测函数；
2. $fg$ 是可测函数；
3. 若对任意的 $x\in X$，有 $g(x)\neq 0$，则 $f/g$ 是可测函数。

**Theorem 1.35.** $\{f_n(x)\}$ 是可测空间 $(X,\mathscr{A})$ 上的一列可测函数，则：

$$\inf_n f_n(x),\ \sup_n f_n(x),\ \underline{\lim}_{n\to+\infty} f_n(x),\ \overline{\lim}_{n\to+\infty} f_n(x)$$

也是可测函数。

### 可测函数与简单函数

**Definition 1.25.** 对于空间 $X$，如果存在有限个互不相交的集合 $\{A_i\subset X: i=1,2,\ldots,n\}$ 满足：

$$\bigcup_{i=1}^{n} A_i = X$$

则称 $\{A_i\subset X: i=1,2,\ldots,n\}$ 为 $X$ 的一个**有限分割**。如果 $A_i\in\mathscr{A},\ \forall\, i=1,2,\ldots,n$，则称 $\{A_i\subset X: i=1,2,\ldots,n\}$ 为可测空间 $(X,\mathscr{A})$ 的一个**有限可测分割**。当 $\{A_n\in\mathscr{A}\}$ 是一个互不相交可列集合且满足：

$$\bigcup_{n=1}^{+\infty} A_n = X$$

时，称 $\{A_n\}$ 为可测空间 $(X,\mathcal{A})$ 的一个**可列可测分割**。

**Definition 1.26.** 对于可测空间 $(X,\mathscr{A})$ 上的函数 $\varphi: X\to\mathbb{R}$，如果存在有限可测分割 $\{A_i\in\mathscr{A}: i=1,2,\ldots,n\}$ 和 $\{a_i\in\mathbb{R}: i=1,2,\ldots,n\}$ 使得：

$$\varphi(x)=\sum_{i=1}^{n} a_i I_{A_i}(x)$$

其中 $I_{A_i}(x)$ 为表示 $x$ 是否在 $A_i$ 中的指示函数，则称 $\varphi(x)$ 为简单函数 *(simple function)*。

**Property 1.3.1.** 简单函数具有如下性质：

1. 简单函数是可测函数；

2. 设 $\varphi(x),\psi(x)$ 为简单函数，可分别表示为：

$$\varphi(x)=\sum_{i=1}^{m} c_i I_{E_i}(x),\quad \psi(x)=\sum_{j=1}^{n} d_j I_{F_j}(x)$$

   则对任意的 $\alpha,\beta\in\mathbb{R}$，$\alpha\varphi(x)+\beta\psi(x)$ 也是简单函数，且可以表示为：

$$\alpha\varphi(x)+\beta\psi(x)=\sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha c_i+\beta d_j) I_{E_i\cap F_j}(x)$$

3. 如果 $f$ 是可测空间 $(X,\mathscr{A})$ 上的可测函数，则 $f$ 为简单函数的充分必要条件为它的值域是有限个实数组成的集合。

*Proof.* (1) 由可测函数的定义立即可得。

(2) 因为：

$$\begin{aligned}\alpha\varphi(x)+\beta\psi(x)&=\alpha\sum_{i=1}^{m}c_iI_{E_i}(x)+\beta\sum_{j=1}^{n}d_jI_{F_j}(x)\\&=\alpha\sum_{i=1}^{m}c_i\sum_{j=1}^{n}I_{E_i\cap F_j}(x)+\beta\sum_{j=1}^{n}d_j\sum_{i=1}^{m}I_{E_i\cap F_j}(x)\\&=\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha c_iI_{E_i\cap F_j}(x)+\sum_{j=1}^{n}\sum_{i=1}^{m}\beta d_jI_{E_i\cap F_j}(x)\\&=\sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha c_i+\beta d_j)I_{E_i\cap F_j}(x)\end{aligned}$$

并且有 $\{E_i\cap F_j : i=1,2,\dots,m,\ j=1,2,\dots,n\}$ 是 $X$ 的有限可测分割，由此可知结论成立。

(3) 必要性由简单函数的定义即可立即得到。设 $f$ 的值域为 $\{a_i : i=1,2,\dots,n\}$，因为 $f$ 是可测函数，由定理 1.32可得 $\{f=a_i\}\in\mathscr{A}$，于是 $f$ 可表为：

$$f(x)=\sum_{i=1}^{n}a_iI_{\{f=a_i\}}(x),\quad X=\bigcup_{i=1}^{n}\{f=a_i\},\quad \{f=a_i\}\cap\{f=a_j\}=\varnothing,\ \forall\, i\neq j$$

所以 $f$ 为简单函数，充分性得证。 □

**Definition 1.27.** 设 $f(x)$ 是可测空间 $(X,\mathscr{A})$ 上的可测函数，令：

$$f^+(x)=\max\{f(x),0\}=\begin{cases}f(x), & f(x)\geqslant 0\\ 0, & f(x)<0\end{cases}$$

$$f^-(x)=-\min\{f(x),0\}=\begin{cases}-f(x), & f(x)\leqslant 0\\ 0, & f(x)>0\end{cases}$$

分别称 $f^+(x)$ 和 $f^-(x)$ 为 $f(x)$ 的正部和负部。

**Theorem 1.36.** 设 $f(x)$ 是可测空间 $(X,\mathscr{A})$ 上的可测函数，则 $f^+(x)$ 和 $f^-(x)$ 也是可测函数。

*Proof.* 由可测函数的定义，$g(x)=0$ 是一个可测函数。因为 $\max\{f(x),0\}=\sup\{f(x),g(x)\}$，由定理 1.35可知此时 $f^+(x)$ 是可测函数。同理，$f^-(x)$ 也是可测函数。 □

**Theorem 1.37.** *(1) 若 $f(x)$ 是可测空间 $(X,\mathscr{A})$ 上的非负可测函数，则存在简单函数列 $\{\varphi_n(x)\}$，使得对任意 $x\in X$，$\varphi_n(x)\leqslant\varphi_{n+1}(x)$，$n\in\mathbb{N}^+$，且 $\lim\limits_{n\to+\infty}\varphi_n(x)=f(x)$。*

*(2) 若 $f(x)$ 是可测空间 $(X,\mathscr{A})$ 上的可测函数，则存在可测简单函数列 $\{\varphi_n(x)\}$，使得对任意 $x\in X$，$\lim\limits_{n\to+\infty}\varphi_n(x)=f(x)$。若 $f(x)$ 有界，则上述收敛可以是一致收敛。*

*Proof.* (1) 对任意的 $n \in \mathbb{N}^+$，将 $[0,n]$ 分为 $n2^n$ 份，令：

$$E_{nj} = \left\{x \in X : \frac{j-1}{2^n} \leqslant f(x) < \frac{j}{2^n}\right\}, \quad j = 1,2,\dots,n2^n$$
$$E_n = \{x : f(x) \geqslant n\}, \quad n = 1,2,\dots$$

作函数列：

$$\varphi_n(x) = \begin{cases} \dfrac{j-1}{2^n}, & x \in E_{nj} \\ n, & x \in E_n \end{cases} \quad j = 1,2,\dots,n2^n;\ n = 1,2,\dots$$

则 $\varphi_n(x)$ 是简单函数，并且有：

$$\varphi_n(x) \leqslant \varphi_{n+1}(x) \leqslant f(x), \quad \forall\, n \in \mathbb{N}^+$$

设 $x \in X$，若 $f(x) < +\infty$，则当 $n > f(x)$ 时有：

$$0 \leqslant f(x) - \varphi_n(x) \leqslant 2^{-n}$$

若 $f(x) = +\infty$，则 $\varphi_n(x) = n,\ n = 1,2,\dots$，因此 $\lim\limits_{n\to+\infty} \varphi_n(x) = f(x)$。

(2) $f(x) = f^+(x) - f^-(x)$，若 $f(x)$ 是可测函数，由定理 1.36可知 $f^+(x), f^-(x)$ 也是可测函数。由 (1)，存在可测简单函数列 $\{\varphi_n^+(x)\}$ 和 $\{\varphi_n^-(x)\}$，使得对任意的 $x \in X$，有：

$$\lim_{n\to+\infty} \varphi_n^+(x) = f^+(x), \lim_{n\to+\infty} \varphi_n^-(x) = f^-(x)$$

令 $\varphi_n(x) = \varphi_n^+(x) - \varphi_n^-(x)$，由定理 1.34可知 $\{\varphi_n(x)\}$ 是可测简单函数列，且对任意的 $x \in X$，$\lim\limits_{n\to+\infty} \varphi_n(x) = f(x)$。

若 $f(x)$ 有界，设 $\sup\limits_{x\in E}\{|f(x)|\} = M$，则由 (1) 的证明过程，当 $n > M$ 时有：

$$\sup_{x\in X}\{|f^+(x) - \varphi_n^+(x)|\} \leqslant \frac{1}{2^n}$$
$$\sup_{x\in X}\{|f^-(x) - \varphi_n^-(x)|\} \leqslant \frac{1}{2^n}$$

因此：

$$\begin{aligned}
\sup_{x\in X}\{|f(x) - \varphi_n(x)|\} &= \sup_{x\in X}\{|f^+(x) - f^-(x) - \varphi_n^+(x) + \varphi_n^-(x)|\} \\
&\leqslant \sup_{x\in X}\{|f^+(x) - \varphi_n^+(x)| + |f^-(x) - \varphi_n^-(x)|\} \\
&\leqslant \sup_{x\in X}\{|f^+(x) - \varphi_n^+(x)|\} + \sup_{x\in X}\{|f^-(x) - \varphi_n^-(x)|\} \\
&\leqslant \frac{1}{2^{n-1}}
\end{aligned}$$

所以 $\{\varphi_n(x)\}$ 在 $X$ 上一致收敛于 $f(x)$。 □

### 1.3.3 可测函数的收敛性

**Lemma 1.6.** 若 $f,g$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数，则对于任意的 $\varepsilon>0$，$\{|f-g|\geqslant\varepsilon\}\in\mathscr{F}$。

*Proof.* 因为：

$$\{|f-g|\geqslant\varepsilon\}=\{f-g\geqslant\varepsilon\}\cup\{f-g\leqslant-\varepsilon\}=\{f\geqslant g+\varepsilon\}\cup\{f\leqslant g-\varepsilon\}$$

因为 $g$ 是可测函数，由定理 1.32可得 $g+\varepsilon$ 和 $g-\varepsilon$ 也是可测函数。由推论 1.4可知 $\{f\geqslant g+\varepsilon\},\{f\leqslant g-\varepsilon\}\in\mathscr{F}$。因为 $\mathscr{F}$ 是一个 $\sigma$ 域，所以：

$$\{|f-g|\geqslant\varepsilon\}=\{f\geqslant g+\varepsilon\}\cup\{f\leqslant g-\varepsilon\}\in\mathscr{F}$$ □

#### 几乎处处收敛

**Definition 1.28.** 设 $\{f_n\}$ 和 $f$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数，如果：

$$\mu\left(\left\{\lim_{n\to+\infty}f_n\neq f\right\}\right)=0$$

则称可测函数列 $\{f_n\}$ *a.e.* 以 $f$ 为极限，记为 $f_n\xrightarrow{a.e.}f$。若此时还有 $f$ 有限 *a.e.* 于 $X$，则称 $\{f_n\}$ *a.e.* 收敛到 $f$。若 $(X,\mathscr{F},P)$ 是概率空间，称 $\{f_n\}$ *a.e.* 收敛到 $f$ 为 $\{f_n\}$ 几乎必然收敛到 $f$，记作 $f_n\xrightarrow{a.s.}f$

**Theorem 1.38.** 设 $\{f_n\}$ 和 $f$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数，$\{f_n\}$ *a.e.* 收敛于 $f$ 的充分必要条件为对任意的 $\varepsilon>0$，有[1]：

$$\mu\left(\varlimsup_{n\to+\infty}\{|f_n-f|\geqslant\varepsilon\}\right)=0$$

*Proof.* 若 $\lim\limits_{n\to+\infty}f_n(x)\neq f(x)$，则有：

$$\exists\, k\in\mathbb{N}^+,\ \forall\, N\in\mathbb{N}^+,\ \exists\, n\geqslant N,\ |f_n(x)-f(x)|\geqslant\frac{1}{k}$$

于是：

$$\left\{\lim_{n\to+\infty}f_n\neq f\right\}=\bigcup_{k=1}^{+\infty}\bigcap_{m=1}^{+\infty}\bigcup_{n=m}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}$$

**(1) 充分性：**此时由半环上测度的次可列可加性（定理 1.23）可得：

$$\begin{aligned}\mu\left(\varlimsup_{n\to+\infty}\{|f_n-f|\geqslant\varepsilon\}\right)&=\mu\left(\bigcup_{k=1}^{+\infty}\bigcap_{m=1}^{+\infty}\bigcup_{n=m}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)\\&\leqslant\sum_{k=1}^{+\infty}\mu\left(\bigcap_{m=1}^{+\infty}\bigcup_{n=m}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)=0\end{aligned}$$

[1]从今往后，对任何的 $\varepsilon>0$ 和 $n\in\mathbb{N}^+$，$f_n(x)-f(x)$ 没有定义的 $x\in X$ 也计入集合 $\{|f_n-f|\geqslant\varepsilon\}$

**(2) 必要性**：对任意取定的 $\varepsilon > 0$，若：

$$x \in \overline{\lim_{n \to +\infty}} \{|f_n - f| \geqslant \varepsilon\}$$

则：

$$\forall\, m \in \mathbb{N}^+,\ \exists\, n \geqslant m,\ |f_n(x) - f(x)| \geqslant \varepsilon$$

即：

$$\lim_{n \to +\infty} f_n(x) \neq f(x)$$

于是对这个 $\varepsilon$，有：

$$\overline{\lim_{n \to +\infty}} \{|f_n - f| \geqslant \varepsilon\} \subset \left\{ \lim_{n \to +\infty} f_n \neq f \right\}$$

因为 $f_n, f$ 都是可测函数，由引理 1.6可知 $\{|f_n - f| \geqslant \varepsilon\} \in \mathscr{F}$。因为 $\mathscr{F}$ 是一个 $\sigma$ 域，对可列并封闭，所以：

$$\overline{\lim_{n \to +\infty}} \{|f_n - f| \geqslant \varepsilon\} = \bigcap_{m=1}^{+\infty} \bigcup_{n=m}^{+\infty} \{|f_n - f| \geqslant \varepsilon\} \in \mathscr{F}$$

由测度的单调性可得：

$$\mu\left( \overline{\lim_{n \to +\infty}} \{|f_n - f| \geqslant \varepsilon\} \right) = 0$$

□

### 几乎一致收敛

**Definition 1.29.** 设 $\{f_n\}$ 和 $f$ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果对任意的 $\varepsilon > 0$，存在 $A \in \mathscr{F}$ 使得 $\mu(A) < \varepsilon$ 且：

$$\lim_{n \to +\infty} \sup_{x \notin A} |f_n(x) - f(x)| = 0$$

则称 $\{f_n\}$ 几乎一致收敛到 $f$，记为 $f_n \xrightarrow{a.u.} f$。

**Theorem 1.39.** 设 $\{f_n\}$ 和 $f$ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数，$f_n \xrightarrow{a.u.} f$ 的充分必要条件为对任意的 $\varepsilon$ 有：

$$\lim_{m \to +\infty} \mu\left( \bigcup_{n=m}^{+\infty} \{|f_n - f| \geqslant \varepsilon\} \right) = 0$$

*Proof.* **(1) 必要性**：因为 $\{f_n\} \xrightarrow{\text{a.u.}} f$，所以：

$$\forall\, \delta > 0,\ \exists\, A \in \mathscr{F},\ \mu(A) < \delta,\ \forall\, \varepsilon > 0,\ \exists\, m \in \mathbb{N}^+,\ \forall\, n > m,\ \sup_{x \notin A} |f_n(x) - f(x)| < \varepsilon$$

即：

$$\forall\, \delta > 0,\ \exists\, A \in \mathscr{F},\ \mu(A) < \delta,\ \forall\, \varepsilon > 0,\ \exists\, m \in \mathbb{N}^+,\ A^c \subset \bigcap_{n=m}^{+\infty} \{|f_n - f| < \varepsilon\}$$

于是：

$$\forall\, \delta > 0,\ \exists\, A \in \mathscr{F},\ \mu(A) < \delta\ \forall\, \varepsilon > 0,\ \exists\, m \in \mathbb{N}^+,\ \bigcup_{n=m}^{+\infty} \{|f_n - f| \geqslant \varepsilon\} \subset A$$

因为 $f_n$ 和 $f$ 都是可测函数，由引理 1.6可知 $\{|f_n-f|\geqslant\varepsilon\}\in\mathscr{F}$。因为 $\mathscr{F}$ 是一个 $\sigma$ 域，对可列并封闭，所以：

$$\bigcup_{n=m}^{+\infty}\{|f_n-f|\geqslant\varepsilon\}\in\mathscr{F}$$

由测度的单调性可得：

$$\mu\left(\bigcup_{n=m}^{+\infty}\{|f_n-f|\geqslant\varepsilon\}\right)<\mu(A)<\delta$$

所以：

$$\forall\,\delta>0,\ \forall\,\varepsilon>0,\ \exists\,m\in\mathbb{N}^+,\ \mu\left(\bigcup_{n=m}^{+\infty}\{|f_n-f|\geqslant\varepsilon\}\right)<\delta$$

即对任意的 $\varepsilon>0$，有（上式前面的两个任意用集合语言来描述就是两个交运算，这是可以交换顺序的，所以有了如下结论）：

$$\lim_{m\to+\infty}\mu\left(\bigcup_{n=m}^{+\infty}\{|f_n-f|\geqslant\varepsilon\}\right)=0$$

**(2) 充分性**：对任意的 $\delta>0$，由所给条件，对 $k\in\mathbb{N}^+$ 有：

$$\lim_{m\to+\infty}\mu\left(\bigcup_{n=m}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)=0$$

于是存在 $\{m_k\}$ 使得：

$$\mu\left(\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)<\frac{\delta}{2^k}$$

取：

$$A=\bigcup_{k=1}^{+\infty}\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}$$

由引理 1.6可得 $A\in\mathscr{F}$，于是根据半环上测度的次可列可加性（定理 1.23）可得：

$$\begin{aligned}\mu(A)&=\mu\left(\bigcup_{k=1}^{+\infty}\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)\\&\leqslant\sum_{k=1}^{+\infty}\mu\left(\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)<\delta\end{aligned}$$

注意到：

$$\begin{aligned}A^c&=\left(\bigcup_{k=1}^{+\infty}\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)^c=\bigcap_{k=1}^{+\infty}\left(\bigcup_{n=m_k}^{+\infty}\left\{|f_n-f|\geqslant\frac{1}{k}\right\}\right)^c\\&=\bigcap_{k=1}^{+\infty}\bigcap_{n=m_k}^{+\infty}\left\{|f_n-f|<\frac{1}{k}\right\}\end{aligned}$$

所以若 $x\notin A$，则对任意的 $k\in\mathbb{N}^+$，当 $n\geqslant m_k$ 时就有：

$$|f_n(x)-f(x)|<\frac{1}{k}$$

由上确界的不等式性，此时即：

$$\forall\,k\in\mathbb{N}^+,\ \exists\,m_k\in\mathbb{N}^+,\ \forall\,n\geqslant m_k,\ \sup_{x\notin A}|f_n(x)-f(x)|\leqslant\frac{1}{k}$$

也即：

$$\lim_{n\to+\infty}\sup_{x\notin A}|f_n(x)-f(x)|=0$$

所以 $f_n\xrightarrow{\text{a.u.}}f$。 □

### 依测度收敛

**Definition 1.30.** 设 $\{f_n\}$ 和 $f$ 时测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数。如果对任意的 $\varepsilon > 0$ 都有：

$$\lim_{n\to+\infty} \mu(\{|f_n - f| \geqslant \varepsilon\}) = 0$$

则称可测函数列 $\{f_n\}$依测度收敛 *(convergent in measure)*到 $f$，记为 $f_n \xrightarrow{\mu} f$。若 $(X, \mathscr{F}, P)$ 是概率空间，称 $f_n \xrightarrow{P} f$ 为 $\{f_n\}$依概率收敛 *(convergent in probability)*到 $f$。

### 依分布收敛

**Theorem 1.40.** 设 $X = \mathbb{R}$，$\mathscr{A} = \{(a, b] : a, b \in \mathbb{R}\}$，$F$ 为 $\mathbb{R}$ 上非降右连续实值函数。对任意的 $a, b \in \mathbb{R}$，令：

$$\mu\Big((a, b]\Big) = \begin{cases} F(b) - F(a), & a < b \\ 0, & a \geqslant b \end{cases}$$

则 $\mu$ 是 $\mathscr{A}$ 上的测度。

**Definition 1.31.** 称 $\mathbb{R}$ 上非降右连续实值函数 $F$ 为。若 $F$ 还满足：

$$\lim_{x\to-\infty} F(x) = 0, \ \lim_{x\to+\infty} F(x) = 1$$

则称 $F$ 为分布函数 *(distribution function, d.f.)*。

**Theorem 1.41.** 设 $f$ 是概率空间 $(X, \mathscr{F}, P)$ 上的随机变量，对任意的 $x \in \mathbb{R}$，令：

$$F(x) = P(f \leqslant x)$$

则 $F$ 是一个分布函数。

*Proof.* (1) 由条件可得：

$$F(x) = P(f \leqslant x) = P(\{f \leqslant x\})$$

是一个非降的实值函数。

(2) 注意到：

$$\{f \leqslant x\} = \lim_{n\to+\infty} \left\{f \leqslant x + \frac{1}{n}\right\}$$

考虑集族：

$$\left\{X_n = \left\{f \leqslant x + \frac{1}{n}\right\}\right\}$$

则 $\{X_n\}$ 是一个单调不增的集族，有：

$$\lim_{n\to+\infty} X_n = \bigcap_{n=1}^{+\infty} X_n = \bigcap_{n=1}^{+\infty} \left\{f \leqslant x + \frac{1}{n}\right\}$$

由半环上测度的上连续性（定理 1.23）可得：

$$\begin{aligned}F(x)=P(\{f\leqslant x\})&=P\left(\lim_{n\to+\infty}\left\{f\leqslant x+\frac{1}{n}\right\}\right)=P\left(\lim_{n\to+\infty}X_n\right)\\&=\lim_{n\to+\infty}P(X_n)=\lim_{n\to+\infty}P\left(\left\{f\leqslant x+\frac{1}{n}\right\}\right)=\lim_{h\to+0}F(x+h)\end{aligned}$$

所以 $F(x)$ 是右连续的。

(3) 注意到 $f$ 是一个实值函数，由半环上测度的上连续性与下连续性（定理 1.23）可得：

$$\lim_{x\to-\infty}F(x)=\lim_{x\to-\infty}P(f\leqslant x)=P(f\leqslant-\infty)=P(\varnothing)=0$$
$$\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}P(f\leqslant x)=P(f\leqslant+\infty)=P(X)=1$$

□

**Definition 1.32.** 设 $f$ 是概率空间 $(X,\mathscr{F},P)$ 上的随机变量，$F$ 是一个分布函数。若对任意的 $x\in\mathbb{R}$，都有：

$$F(x)=P(f\leqslant x)$$

则称 *r.v.* $f$ 的分布函数是 $F$，也说成 $f$ 服从 $F$，记为 $f\sim F$。若 $f$ 是从概率空间 $(X,\mathscr{F},P)$ 到可测空间 $(Y,\mathscr{A})$ 的可测映射，称：

$$P(f^{-1}A),\ \forall\ A\in\mathscr{A}$$

为 $f$ 的概率分布 *(probability distribution)*。

**Definition 1.33.** 设 $\{f_n\sim F_n\}$ 是概率空间 $(X,\mathscr{F},P)$ 上的随机变量序列，$F$ 是一个分布函数。若对于 $F$ 的每一个连续点 $x$ 都有：

$$\lim_{n\to+\infty}F_n(x)=F(x)$$

则称 $\{f_n\}$依分布收敛 *(convergent in distribution)*，记为 $f_n\xrightarrow{d}F$。若此时概率空间 $(X,\mathscr{F},P)$ 上的随机变量 $f\sim F$，则称随机变量序列 $\{f_n\}$ 依分布收敛到 $f$，记为 $f_n\xrightarrow{d}f$。

## 收敛性之间的关系

**Theorem 1.42.** 设 $\{f_n\}$ 和 $f$ 为测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数，则：

1. $f_n\xrightarrow{a.u.}f$ 可推出 $f_n\xrightarrow{a.e.}f$ 和 $f_n\xrightarrow{\mu}f$；
2. 若 $\mu(X)<+\infty$，则 $f_n\xrightarrow{a.u.}f$ 等价于 $f_n\xrightarrow{a.e.}f$；
3. $f_n\xrightarrow{\mu}f$ 当且仅当对 $\{f_n\}$ 的任一子列，存在该子列的子列 $\{f_{n_k}\}$ 使得 $f_{n_k}\xrightarrow{a.u.}f$；
4. $f_n\xrightarrow{P}f$ 可推出 $f_n\xrightarrow{d}f$。

*Proof.* (1) 对任意的 $n \in \mathbb{N}^+$ 和任意的 $\varepsilon > 0$，由推论 1.4可得 $\{|f_n - f| \geqslant \varepsilon\}$ 是可测集，因为 $\mathscr{F}$ 是一个 $\sigma$ 域，对可列并、可列交封闭，所以：

$$\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\} \in \mathscr{F},\ \bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\} \in \mathscr{F}$$

因为：

$$\{|f_n - f| \geqslant \varepsilon\} \subset \bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}$$
$$\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\} \subset \bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}$$

由测度的单调性可得：

$$\mu(\{|f_n - f| \geqslant \varepsilon\}) \leqslant \mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right)$$
$$\mu\left(\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) \leqslant \mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right)$$

因为 $f_n \xrightarrow{\text{a.u.}} f$，由极限的不等式性和测度的非负性可得：

$$0 \leqslant \lim_{n\to+\infty}\mu(\{|f_n - f| \geqslant \varepsilon\}) \leqslant \lim_{n\to+\infty}\mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) = 0$$
$$0 \leqslant \mu\left(\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) \leqslant \lim_{n\to+\infty}\mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) = 0$$

所以：

$$\lim_{n\to+\infty}\mu(\{|f_n - f| \geqslant \varepsilon\}) = 0,\ \mu\left(\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) = 0$$

即 $f_n \xrightarrow{\mu} f$，$f_n \xrightarrow{a.e.} f$。

(2) 设 $f_n \xrightarrow{a.e.} f$，令：

$$\left\{X_n = \bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right\}$$

显然 $\{X_n\}$ 是一个单调不增序列，其极限为：

$$\lim_{n\to+\infty} X_n = \bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}$$

由半环上测度的上连续性可得（定理 1.23）：

$$\mu\left(\lim_{n\to+\infty} X_n\right) = \lim_{n\to+\infty}\mu(X_n)$$

于是：

$$\mu\left(\bigcap_{n=1}^{+\infty}\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) = \lim_{n\to+\infty}\mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right)$$

所以：

$$\lim_{n\to+\infty}\mu\left(\bigcup_{m=n}^{+\infty}\{|f_m - f| \geqslant \varepsilon\}\right) = 0$$

即 $f_n \xrightarrow{a.u.} f$，再结合 (1) 即可得出结论。

(3)

(4) 记 $F$ 为 $f$ 的分布函数，$F_n$ 为 $f_n$ 的分布函数。因为 $f_n \xrightarrow{P} f$，所以对任意的 $\varepsilon > 0$，有：

$$\lim_{n\to+\infty} P(|f_n - f| \geqslant \varepsilon) = 0$$

对任意的 $x \in X$、任意的 $\varepsilon > 0$ 和任意的 $n \in \mathbb{N}^+$，有：

$$\begin{aligned}F_n(x) &= P(f_n \leqslant x)\\ &\leqslant P(f_n \leqslant x, |f_n - f| < \varepsilon) + P(f_n \leqslant x, |f_n - f| \geqslant \varepsilon)\\ &\leqslant P(f \leqslant x + \varepsilon) + P(|f_n - f| \geqslant \varepsilon)\end{aligned}$$

第二行到第三行第一式的变化是因为：

$$|f_n - f| < \varepsilon \Leftrightarrow -\varepsilon < f_n - f < \varepsilon \to f < f_n + \varepsilon$$

而 $f_n \leqslant x$，于是变为 $f \leqslant x + \varepsilon$。该条件比原条件宽松，所以是小于等于号。由极限的不等式性可得：

$$\lim_{n\to+\infty} F_n(x) \leqslant P(f \leqslant x + \varepsilon) + \lim_{n\to+\infty} P(|f_n - f| \geqslant \varepsilon) = P(f \leqslant x + \varepsilon)$$

由 $\varepsilon$ 的任意性可得：

$$\lim_{n\to+\infty} F_n(x) \leqslant F(x)$$

对任意的 $x \in X$、任意的 $\varepsilon > 0$ 和任意的 $n \in \mathbb{N}^+$，又有：

$$\begin{aligned}P(f \leqslant x - \varepsilon) &\leqslant P(f \leqslant x - \varepsilon, |f_n - f| < \varepsilon) + P(f \leqslant x - \varepsilon, |f_n - f| \geqslant \varepsilon)\\ &\leqslant P(f_n \leqslant x) + P(|f_n - f| \geqslant \varepsilon)\end{aligned}$$

由极限的不等式性可得：

$$P(f \leqslant x - \varepsilon) \leqslant \lim_{n\to+\infty} P(f_n \leqslant x) + \lim_{n\to+\infty} P(|f_n - f| \geqslant \varepsilon) = \lim_{n\to+\infty} F_n(x)$$

于是：

$$P(f \leqslant x - 0) \leqslant \lim_{n\to+\infty} F_n(x)$$

若 $F(x)$ 在 $x$ 处连续，就有：

$$P(f \leqslant x - 0) = F(x - 0) = F(x)$$

所以：

$$\lim_{n\to+\infty} F_n(x) = F(x) \tag{1.4}$$

即 $f_n \xrightarrow{d} f$。 □

# 1.4 积分论

## 1.4.1 非负简单函数的积分

**Definition 1.34.** 设 $\varphi(x)$ 为测度空间 $(X,\mathscr{F},\mu)$ 上的一个非负简单函数，即 $X$ 可表示为有限个互不相交的集合 $E_1,E_2,\ldots,E_n\in\mathscr{F}$ 的并，且在 $E_i(i=1,2,\ldots,n)$ 上 $\varphi(x)=c_i\geqslant 0$，即：

$$\varphi(x)=\sum_{i=1}^{n}c_iI_{E_i}(x)$$

其中 $I_{E_i}(x)$ 为表示 $x$ 是否在 $E_i$ 中的示性函数。对于任意的 $A\in\mathscr{F}$，将 $\varphi(x)$ 在 $A$ 上的积分定义为：

$$\int_A\varphi(x)\,\mathrm{d}\mu=\sum_{i=1}^{n}c_i\mu(A\cap E_i)$$

**Property 1.4.1.** 设 $\varphi(x)$, $\psi(x)$ 为测度空间 $(X,\mathscr{F},\mu)$ 上的非负简单函数，可分别表示为：

$$\varphi(x)=\sum_{i=1}^{m}c_iI_{E_i}(x),\quad \psi(x)=\sum_{j=1}^{n}d_jI_{F_j}(x)$$

则：

1. $\varphi(x)$ 的所有表达式在任意的 $A\in\mathscr{F}$ 上的积分值相同；

2. 对于任意的 $A\in\mathscr{F}$，有：
$$\int_A\varphi(x)\,\mathrm{d}\mu\geqslant 0$$

3. 若 $A\in\mathscr{F}$ 且 $\mu(A)=0$，则有：
$$\int_A\varphi(x)\,\mathrm{d}\mu=0$$

4. 设 $A,B\in\mathscr{F}$ 且 $A\cap B=\varnothing$，则：
$$\int_{A\cup B}\varphi(x)\,\mathrm{d}\mu=\int_A\varphi(x)\,\mathrm{d}\mu+\int_B\varphi(x)\,\mathrm{d}\mu$$

5. 对任意的 $A\in\mathscr{F}$ 和 $\alpha,\beta\in\mathbb{R}$ 且 $\alpha,\beta\geqslant 0$：
$$\int_A[\alpha\varphi(x)+\beta\psi(x)]\,\mathrm{d}\mu=\alpha\int_A\varphi(x)\,\mathrm{d}\mu+\beta\int_A\psi(x)\,\mathrm{d}\mu$$

6. 取 $A\in\mathscr{F}$，若对任意的 $x\in A$ 有 $\varphi(x)\geqslant\psi(x)$，则有：
$$\int_A\varphi(x)\,\mathrm{d}\mu\geqslant\int_A\psi(x)\,\mathrm{d}\mu$$

7. 设 $\{A_n\}\subset\mathscr{F}$，$A_n\uparrow E\in\mathscr{F}$，则：
$$\lim_{n\to+\infty}\left[\int_{A_n}\varphi(x)\,\mathrm{d}\mu\right]=\int_E\varphi(x)\,\mathrm{d}\mu$$

8. 取 $A\in\mathscr{F}$，若非负简单函数列 $\varphi_n(x)\uparrow$ 且对任意的 $x\in A$ 有 $\lim\limits_{n\to+\infty}\varphi_n(x)\geqslant\psi(x)$，则有：

$$\lim_{n\to+\infty}\left[\int_A\varphi_n(x)\,\mathrm{d}\mu\right]\geqslant\int_A\psi(x)\,\mathrm{d}\mu$$

*Proof.* (1) 由性质 1.3.1(3)，将 $\varphi(x)$ 表示为：

$$\varphi(x)=\sum_{k=1}^{p}a_kI_{\{f=a_k\}}(x)$$

其中 $\{a_k:k=1,2,\ldots,p\}$ 为 $\varphi(x)$ 的值域，所以 $p\leqslant m$。对任意的 $i$ 和 $k$，显然有：

$$E_i\subset\{f=a_k\}\quad\text{或}\quad E_i\cap\{f=a_k\}=\varnothing$$

当 $E_i\subset\{f=a_k\}$ 时有 $c_i=a_k$。由测度的有限可加性可得：

$$\begin{aligned}
\sum_{i=1}^{m}c_i\mu(A\cap E_i)&=\sum_{i=1}^{m}c_i\mu[(A\cap E_i)\cap X]=\sum_{i=1}^{m}c_i\mu\left[(A\cap E_i)\cap\left(\bigcup_{k=1}^{p}A_k\right)\right]\\
&=\sum_{i=1}^{m}c_i\mu\left[\bigcup_{k=1}^{p}(A\cap E_i\cap A_k)\right]=\sum_{i=1}^{m}c_i\sum_{k=1}^{p}\mu(A\cap E_i\cap A_k)\\
&=\sum_{i=1}^{m}\sum_{E_i\subset\{f=a_k\}}c_i\mu(A\cap E_i\cap A_k)=\sum_{i=1}^{m}\sum_{E_i\subset\{f=a_k\}}a_k\mu(A\cap E_i\cap A_k)\\
&=\sum_{i=1}^{m}\sum_{k=1}^{p}a_k\mu(A\cap E_i\cap A_k)=\sum_{k=1}^{p}\sum_{i=1}^{m}a_k\mu(A\cap A_k\cap E_i)\\
&=\sum_{k=1}^{p}a_k\sum_{i=1}^{m}\mu(A\cap A_k\cap E_i)=\sum_{k=1}^{p}a_k\mu\left[\bigcup_{i=1}^{m}(A\cap A_k\cap E_i)\right]\\
&=\sum_{k=1}^{p}a_k\mu\left[(A\cap A_k)\cap\left(\bigcup_{i=1}^{m}E_i\right)\right]=\sum_{k=1}^{p}a_k\mu[(A\cap A_k)\cap X]\\
&=\sum_{k=1}^{p}a_k\mu(A\cap A_k)
\end{aligned}$$

(2) 由非负简单函数积分的定义和测度的非负性直接可得。

(3) 由非负简单函数积分的定义和测度的非负性以及半环上测度的单调性（定理 1.23）直接可得。

(4) 由非负简单函数积分的定义、测度的有限可加性可得：

$$\begin{aligned}
\int_{A\cup B}\varphi(x)\,\mathrm{d}\mu&=\sum_{i=1}^{n}c_i\mu[(A\cup B)\cap E_i]=\sum_{i=1}^{n}c_i\mu[(A\cap E_i)\cup(B\cap E_i)]\\
&=\sum_{i=1}^{n}c_i[\mu(A\cap E_i)+\mu(B\cap E_i)]=\sum_{i=1}^{n}c_i\mu(A\cap E_i)+\sum_{i=1}^{n}\mu(B\cap E_i)\\
&=\int_A\varphi(x)\,\mathrm{d}\mu+\int_B\varphi(x)\,\mathrm{d}\mu
\end{aligned}$$

(5) 由性质 1.3.1(2) 可得 $\alpha\varphi(x)+\beta\psi(x)$ 也是非负简单函数。由非负简单函数积分的定义和测度的有限可加性可得：

$$\begin{aligned}
\int_A [\alpha\varphi(x)+\beta\psi(x)]\,\mathrm{d}\mu &= \sum_{i=1}^{m}\sum_{j=1}^{n}(\alpha c_i+\beta d_j)\mu[A\cap(E_i\cap F_j)] \\
&= \sum_{i=1}^{m}\alpha c_i\left[\sum_{j=1}^{n}\mu(A\cap E_i\cap F_j)\right]+\sum_{j=1}^{n}\beta d_j\left[\sum_{i=1}^{m}\mu(A\cap E_i\cap F_j)\right] \\
&= \sum_{i=1}^{m}\alpha c_i\mu(A\cap E_i)+\sum_{j=1}^{n}\beta d_j\mu(A\cap F_j) \\
&= \alpha\sum_{i=1}^{m}c_i\mu(A\cap E_i)+\beta\sum_{j=1}^{n}d_j\mu(A\cap F_j) \\
&= \alpha\int_A\varphi(x)\,\mathrm{d}\mu+\beta\int_A\psi(x)\,\mathrm{d}\mu
\end{aligned}$$

(6) 因为 $\varphi(x),\psi(x)$ 是非负简单函数，由性质 1.3.1(2) 可知 $\varphi(x)-\psi(x)$ 也是非负简单函数。根据 (5)(2) 可得：

$$\begin{aligned}
\int_A\varphi(x)\,\mathrm{d}\mu = \int_A[\psi(x)+\varphi(x)-\psi(x)]\,\mathrm{d}\mu &= \int_A\psi(x)\,\mathrm{d}\mu+\int_A[\varphi(x)-\psi(x)]\,\mathrm{d}\mu \\
\geqslant \int_A\psi(x)\,\mathrm{d}\mu
\end{aligned}$$

(7) 由非负简单函数积分的定义、极限的线性性和半环上测度的下连续性（定理 1.23）可得：

$$\lim_{n\to+\infty}\left[\int_{A_n}\varphi(x)\,\mathrm{d}\mu\right]=\lim_{n\to+\infty}\left[\sum_{i=1}^{m}c_i\mu(A_n\cap E_i)\right]=\sum_{i=1}^{m}c_i\mu(E\cap E_i)=\int_E\varphi(x)\,\mathrm{d}\mu$$

(8) 对任意的 $\alpha\in(0,1)$，记 $A_n(\alpha)=\{\varphi_n\geqslant\alpha\psi\}\cap A$。由性质 1.3.1(1) 可知 $\{\varphi_n\},\psi(x)$ 是可测函数，根据定理 1.34(1) 可得 $\alpha\psi(x)$ 也是可测函数。由推论 1.4可知 $A_n(\alpha)\in\mathscr{F}$。设 $\varphi_n(x)$ 可表示为：

$$\varphi_n(x)=\sum_{k=1}^{p_n}a_{nk}I_{E_{nk}}$$

其中 $\{E_{nk}\}$ 是 $X$ 的有限可测分割。因为：

$$\varphi_nI_{A_n(\alpha)}=\sum_{k=1}^{p_n}a_{nk}I_{E_{nk}\cap A_n(\alpha)},\ E_{nk}\cap A_n(\alpha)\in\mathscr{F}$$

所以 $\varphi_nI_{A_n(\alpha)}$ 也是一个非负简单函数。同理，$\psi I_{A_n(\alpha)}$ 也是一个非负简单函数。因为 $\varphi\geqslant\varphi_nI_{A_n(\alpha)}\geqslant\alpha\psi I_{A_n(\alpha)}$，由 (6) 和 (5) 可得：

$$\begin{aligned}
\int_A\varphi_n(x)\,\mathrm{d}\mu\geqslant\int_A\varphi_n(x)I_{A_n(\alpha)}\,\mathrm{d}\mu &\geqslant\int_A\alpha\psi(x)I_{A_n(\alpha)}\,\mathrm{d}\mu \\
&=\alpha\int_A\psi(x)I_{A_n(\alpha)}\,\mathrm{d}\mu=\alpha\sum_{j=1}^{n}d_j\mu[A\cap F_j\cap A_n(\alpha)]
\end{aligned}$$

因为 $\varphi_n\uparrow$ 且 $\lim\limits_{n\to+\infty}\varphi_n(x)\geqslant\psi(x)$ 对任意 $x\in A$ 成立，所以 $A_n(\alpha)\uparrow A$，即 $A\cap F_j\cap A_n(\alpha)\uparrow A\cap F_j$。由极限的不等式性、线性性和半环上测度的下连续性（定理 1.23）可得：

$$\begin{aligned}\lim_{n\to+\infty}\left[\int_A\varphi_n(x)\,\mathrm{d}\mu\right]&\geqslant\alpha\sum_{j=1}^{n}d_j\lim_{n\to+\infty}\mu[A\cap F_j\cap A_n(\alpha)]\\&=\alpha\sum_{j=1}^{n}d_j\mu(A\cap F_j)=\alpha\int_A\psi(x)\,\mathrm{d}\mu\end{aligned}$$

再取 $\alpha\to1$ 即可得到结论。注意上式两个 $n$ 的区别，不想再去写另外的字母了，懒。同时需要注意这里必须要引入 $\alpha$，否则等于的情况就可能不成立。 □

### 1.4.2 非负可测函数的积分

**Definition 1.35.** 设 $f(x)$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的一个非负可测函数，对于任意的 $A\in\mathscr{F}$，将 $f(x)$ 在 $A$ 上的积分定义为：

$$\int_A f(x)\,\mathrm{d}x=\sup_{\varphi(x)}\left\{\int_A\varphi(x)\,\mathrm{d}x:\varphi(x)\text{是非负简单函数，且}\ \forall\,x\in A,\ \varphi(x)\leqslant f(x)\right\}$$

若 $\int_A f(x)\,\mathrm{d}x<+\infty$，则称 $f(x)$ 在 $A$ 上可积。

**Property 1.4.2.** 设 $f(x)$ 和 $g(x)$ 为测度空间 $(X,\mathscr{F},\mu)$ 上的非负可测函数，则：

1. 若 $f(x)$ 是非负简单函数，则其在非负简单函数下定义的积分值与在非负可测函数下定义的积分值相同；
2. 对于任意的 $A\in\mathscr{F}$，$\int_A f(x)\,\mathrm{d}\mu\geqslant0$；
3. 若 $\mu(A)=0$ 且 $A\in\mathscr{F}$，则 $\int_A f(x)\,\mathrm{d}\mu=0$；
4. 对于任意的 $A\in\mathscr{F}$，若 $\{f_n\}$ 是非负简单函数列且 $f_n\uparrow f$，则：

$$\begin{aligned}\int_A f(x)\,\mathrm{d}\mu&=\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]\\&=\lim_{n\to+\infty}\left\{\sum_{j=0}^{n2^n-1}\frac{j}{2^n}\mu\left[\left\{\frac{j}{2^n}\leqslant f<\frac{j+1}{2^n}\right\}\cap A\right]+n\mu[\{f\geqslant n\}\cap A]\right\}\end{aligned}$$

5. 设 $A,B\in\mathscr{F}$ 且 $A\cap B=\varnothing$，则：

$$\int_{A\cup B}f(x)\,\mathrm{d}\mu=\int_A f(x)\,\mathrm{d}\mu+\int_B f(x)\,\mathrm{d}\mu$$

条件还需要再考虑

6. 对任意的 $A\in\mathscr{F}$，任取 $\alpha,\beta\in\mathbb{R}$，有：

$$\int_A[\alpha f(x)+\beta g(x)]\,\mathrm{d}\mu=\alpha\int_A f(x)\,\mathrm{d}\mu+\beta\int_A g(x)\,\mathrm{d}\mu$$

7. 若 $f(x)\leqslant g(x)$ *a.e.* 于 $A\in\mathscr{F}$，则 $\int_A f(x)\,\mathrm{d}\mu\leqslant\int_A g(x)\,\mathrm{d}\mu$；

8. 若 $f(x)=g(x)$ *a.e.* 于 $A\in\mathscr{F}$，则 $\int_A f(x)\,\mathrm{d}\mu=\int_A g(x)\,\mathrm{d}\mu$；

9. 取 $A\in\mathscr{F}$，若 $\int_A f(x)\,\mathrm{d}x<+\infty$，则 $f(x)$ 有限 *a.e.* 于 $A$；

10. 取 $A\in\mathscr{F}$，$\int_A f(x)\,\mathrm{d}\mu=0$ 的充分必要条件为 $f(x)=0$ *a.e.* 于 $A$。

*Proof.* (1) 由非负可测函数积分的定义和性质 1.4.1(6) 直接可得。

(2) 由非负可测函数积分的定义和性质 1.4.1(2) 直接可得。

(3) 由非负可测函数积分的定义和性质 1.4.1(3) 直接可得。

(4) 由非负可测函数积分的定义和所给条件可知对任意的 $n\in\mathbb{N}^+$ 有：

$$\int_A f_n(x)\,\mathrm{d}\mu\leqslant\int_A f(x)\,\mathrm{d}\mu$$

由极限的不等式性可得：

$$\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]\leqslant\int_A f(x)\,\mathrm{d}\mu$$

对任意满足 $\varphi\leqslant f$ 的非负简单函数 $\varphi(x)$，有：

$$\lim_{n\to+\infty} f_n=f\geqslant\varphi$$

于是由性质 1.4.1(8) 可得：

$$\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]\geqslant\int_A \varphi(x)\,\mathrm{d}\mu$$

由上确界的不等式性可得：

$$\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]\geqslant\int_A f(x)\,\mathrm{d}\mu$$

于是就有：

$$\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]=\int_A f(x)\,\mathrm{d}\mu$$

由定理 1.37(1) 可得到积分值的具体表示。

(5) 设 $\varphi(x)$ 是 $A\cup B$ 上任一满足 $\varphi\leqslant f$ 的非负简单函数，于是由性质 1.4.1(4) 可得：

$$\int_{A\cup B}\varphi(x)\,\mathrm{d}x=\int_A\varphi(x)\,\mathrm{d}x+\int_B\varphi(x)\,\mathrm{d}x\leqslant\int_A f(x)\,\mathrm{d}x+\int_B f(x)\,\mathrm{d}x$$

由上确界的不等式性可得：

$$\int_{A\cup B} f(x)\,\mathrm{d}x\leqslant\int_A f(x)\,\mathrm{d}x+\int_B f(x)\,\mathrm{d}x$$

另一方面：

$$\int_{A\cup B} f(x)\,\mathrm{d}x\geqslant\int_{A\cup B}\varphi(x)\,\mathrm{d}x=\int_A\varphi(x)\,\mathrm{d}x+\int_B\varphi(x)\,\mathrm{d}x$$

由上确界的不等式性又可得：

$$\int_{A\cup B} f(x)\,\mathrm{d}x\geqslant\int_A f(x)\,\mathrm{d}x+\int_B f(x)\,\mathrm{d}x$$

所以：

$$\int_{A\cup B} f(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x + \int_B f(x)\,\mathrm{d}x$$

(6) 取非负简单函数列 $\{f_n\},\{g_n\}$ 满足 $f_n\uparrow f, g_n\uparrow g$，由极限的线性性质可得：

$$\lim_{n\to+\infty}(\alpha f_n+\beta g_n) = \alpha\lim_{n\to+\infty} f_n + \beta\lim_{n\to+\infty} g_n = \alpha f+\beta g$$

于是 $\alpha f_n+\beta g_n\uparrow \alpha f+\beta g$。由 (2)、极限的线性性质和性质 1.4.1(5) 可得：

$$\begin{aligned}
\int_A [\alpha f(x)+\beta g(x)]\,\mathrm{d}\mu &= \lim_{n\to+\infty}\left\{\int_A [\alpha f_n(x)+\beta g_n(x)]\,\mathrm{d}\mu\right\}\\
&= \lim_{n\to+\infty}\left[\alpha\int_A f_n(x)\,\mathrm{d}\mu+\beta\int_A g_n(x)\,\mathrm{d}\mu\right]\\
&= \alpha\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]+\beta\lim_{n\to+\infty}\left[\int_A g_n(x)\,\mathrm{d}\mu\right]\\
&= \alpha\int_A f(x)\,\mathrm{d}\mu+\beta\int_A g(x)\,\mathrm{d}\mu
\end{aligned}$$

(7) 令 $A_1=\{f\leqslant g\}$，$A_2=\{f>g\}$，因为 $f,g$ 都是非负简单函数，由性质 1.3.1(1) 可知 $f,g$ 都可测，所以根据推论 1.4可得 $A_1,A_2\in\mathscr{F}$，同时有：

$$A_1\cap A_2=\varnothing,\ A_1\cup A_2=A,\ \mu(A_2)=0$$

由 (4)(3) 可得：

$$\begin{aligned}
\int_A f(x)\,\mathrm{d}\mu = \int_{A_1\cup A_2} f(x)\,\mathrm{d}\mu = \int_{A_1} f(x)\,\mathrm{d}\mu + \int_{A_2} f(x)\,\mathrm{d}\mu = \int_{A_1} f(x)\,\mathrm{d}\mu\\
\int_A g(x)\,\mathrm{d}\mu = \int_{A_1\cup A_2} g(x)\,\mathrm{d}\mu = \int_{A_1} g(x)\,\mathrm{d}\mu + \int_{A_2} g(x)\,\mathrm{d}\mu = \int_{A_1} g(x)\,\mathrm{d}\mu
\end{aligned}$$

对于满足 $\varphi\leqslant f$ 的非负简单函数 $\varphi(x)$，必然也有 $\varphi\leqslant g$，于是由非负可测函数积分的定义可得：

$$\int_{A_1} f(x)\,\mathrm{d}\mu \leqslant \int_{A_1} g(x)\,\mathrm{d}\mu$$

也即：

$$\int_A f(x)\,\mathrm{d}\mu \leqslant \int_A g(x)\,\mathrm{d}\mu$$

(8) 由 (7) 立即可得。

(9) 令 $A_\infty=\{f=+\infty\}$。对任意的 $n\in\mathbb{N}^+$，令：

$$\varphi_n(x)=\begin{cases} n, & x\in A_\infty\\ 0, & x\in A\backslash A_\infty\end{cases}$$

因为 $f$ 是可测函数，由定理 1.32可得 $A_\infty\in\mathscr{F}$，因此 $\varphi_n(x)$ 是非负简单函数。由非负可测函数积分的定义可得：

$$\int_A f(x)\,\mathrm{d}\mu \geqslant \int_A \varphi_n(x)\,\mathrm{d}\mu = n\mu(A_\infty)\geqslant 0$$

所以：

$$\forall\, n \in \mathbb{N}^+,\ 0 \leqslant \mu(A_\infty) \leqslant \frac{1}{n}\int_A f(x)\,\mathrm{d}\mu$$

因为 $\int_A f(x)\,\mathrm{d}\mu < +\infty$，所以 $\mu(A_\infty) = 0$，即 $f(x)$ 有限 a.e. 于 $A$。

(10) 必要性：对任意的 $n \in \mathbb{N}^+$，令：

$$A_n = \left\{ f \geqslant \frac{1}{n} \right\}, \quad \varphi_n(x) = \begin{cases} \frac{1}{n}, & x \in A_n \\ 0, & x \in A \backslash A_n \end{cases}$$

因为 $f$ 是可测函数，所以 $A_n \in \mathscr{F}$，因此 $\varphi_n(x)$ 是非负简单函数。于是：

$$0 = \int_A f(x)\,\mathrm{d}x \geqslant \int_A \varphi_n(x)\,\mathrm{d}x = \frac{1}{n}\mu(A_n) \geqslant 0$$

所以对任意的 $n \in \mathbb{N}^+$, $\mu(A_n) = 0$。因为：

$$\{f > 0\} = \bigcup_{n=1}^{+\infty} A_n$$

由半环上测度的次可列可加性（定理 1.23）以及测度的非负性可得 $\mu(f > 0) = 0$，即 $f(x) = 0$ a.e. 于 $A$。

充分性：函数 $g(x) = 0,\ \forall\, x \in A$ 的积分为 0，由 (7) 立即可证得充分性。 □

### 1.4.3 一般可测函数的积分

**Definition 1.36.** 设 $f(x)$ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数，$A \in \mathscr{F}$。若 $\int_A f^+(x)\,\mathrm{d}\mu$ 和 $\int_A f^-(x)\,\mathrm{d}\mu$ 中至少一个有限，则称 $f(x)$ 在 $A$ 上**积分存在**，将 $f(x)$ 在 $A$ 上的积分定义为：

$$\int_A f(x)\,\mathrm{d}\mu = \int_A f^+(x)\,\mathrm{d}\mu - \int_A f^-(x)\,\mathrm{d}\mu$$

若 $\int_A f^+(x)\,\mathrm{d}\mu$ 和 $\int_A f^-(x)\,\mathrm{d}\mu$ 都有限，则称 $f(x)$ 在 $A$ 上**可积**。

**Property 1.4.3.** 设 $f(x)$ 和 $g(x)$ 为测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数，则：

1. 若 $A \in \mathscr{F}$，$\mu(A) = 0$，则 $A$ 上的任何实值函数 $f(x)$ 都在 $A$ 上可积，并且有 $\int_A f(x)\,\mathrm{d}\mu = 0$；
2. 若 $f$ 在 $A \in \mathscr{F}$ 上积分存在，则 $|\int_A f(x)\,\mathrm{d}\mu| \leqslant \int_A |f(x)|\,\mathrm{d}\mu$；
3. 若 $f$ 在 $A \in \mathscr{F}$ 上积分存在（可积），则 $f$ 在 $A$ 的满足 $B \in \mathscr{F}$ 的子集 $B$ 上也积分存在（可积）；
4. $f$ 在 $A \in \mathscr{F}$ 上可积的充分必要条件为 $|f|$ 可积；
5. 若 $f$ 在 $A \in \mathscr{F}$ 上可积，则 $|f| < +\infty$ *a.e.* 于 $A$；
6. 若 $f, g$ 在 $A \in \mathscr{F}$ 上积分存在，对 $\alpha, \beta \in \mathbb{R}$，$\int_A \alpha f(x)\,\mathrm{d}\mu + \int_A \beta g(x)\,\mathrm{d}\mu$ 有意义，则 $\alpha f + \beta g$ 有定义 *a.e.* 于 $A$，其积分存在且：

$$\int_A [\alpha f(x) + \beta g(x)]\,\mathrm{d}\mu = \alpha \int_A f(x)\,\mathrm{d}\mu + \beta \int_A g(x)\,\mathrm{d}\mu$$

7. 若 $f,g$ 在 $A\in\mathscr{F}$ 上积分存在且 $f\leqslant g$ a.e. 于 $A$，则：

$$\int_A f(x)\,\mathrm{d}\mu\leqslant\int_A g(x)\,\mathrm{d}\mu$$

8. 若 $f,g$ 在 $A\in\mathscr{F}$ 上积分存在且 $f=g$ a.e. 于 $A$，则：

$$\int_A f(x)\,\mathrm{d}\mu=\int_A g(x)\,\mathrm{d}\mu$$

9. 若 $f=0$ a.e. 于 $A\in\mathscr{F}$，则 $\int_A f(x)\,\mathrm{d}\mu=0$；若 $\int_A f(x)\,\mathrm{d}\mu=0$ 且 $f\geqslant 0$ a.e. 于 $A$，则 $f=0$ a.e. 于 $A$；

10. $f,g$ 都是 $X$ 上的可积函数且对任意的 $A\in\mathscr{F}$ 有 $\int_A f(x)\,\mathrm{d}\mu\leqslant\int_A g(x)\,\mathrm{d}\mu$,则 $f\leqslant g$ a.e. 于 $X$；

11. $f,g$ 都是 $X$ 上的可积函数且对任意的 $A\in\mathscr{F}$ 有 $\int_A f(x)\,\mathrm{d}\mu=\int_A g(x)\,\mathrm{d}\mu$,则 $f=g$ a.e. 于 $X$；

*Proof.* (1) 任选 $A$ 上的一个实值函数 $f(x)$。因为 $A$ 是一个非空零测集，由性质 1.4.2(3) 可知：

$$\int_A f^+(x)\,\mathrm{d}\mu=\int_A f^-(x)\,\mathrm{d}\mu=0$$

于是：

$$\int_A f(x)\,\mathrm{d}\mu=\int_A f^+(x)\,\mathrm{d}\mu-\int_A f^-(x)\,\mathrm{d}\mu=0$$

(2) 由性质 1.4.2(6)(2)、绝对值的三角不等式可得：

$$\begin{aligned}\left|\int_A f(x)\,\mathrm{d}\mu\right|&=\left|\int_A[f^+(x)-f^-(x)]\,\mathrm{d}\mu\right|=\left|\int_A f^+(x)\,\mathrm{d}\mu-\int_A f^-(x)\,\mathrm{d}\mu\right|\\&\leqslant\int_A f^+(x)\,\mathrm{d}\mu+\int_A f^-(x)\,\mathrm{d}\mu=\int_A[f^+(x)+f^-(x)]\,\mathrm{d}\mu\\&=\int_A|f(x)|\,\mathrm{d}\mu\end{aligned}$$

(3) 任取 $B\subset A$ 且 $B\in\mathscr{F}$。因为 $f(x)$ 在 $A$ 上积分存在,所以 $\int_A f^+(x)\,\mathrm{d}\mu$ 和 $\int_A f^-(x)\,\mathrm{d}\mu$ 至少有一个有限。设 $\int_A f^+(x)\,\mathrm{d}x$ 有限，另一种情况可对称讨论。由性质 1.4.2(5)(2) 可得：

$$+\infty>\int_A f^+(x)\,\mathrm{d}\mu=\int_B f^+(x)\,\mathrm{d}x+\int_{A\backslash B}f^+(x)\,\mathrm{d}\mu\geqslant\int_B f^+(x)\,\mathrm{d}\mu$$

故 $f(x)$ 在 $B$ 上积分存在。由 $B$ 的任意性，命题成立。

(4) 由性质 1.4.2(6) 可得：

$$\begin{aligned}f\text{可积}&\Leftrightarrow\int_A f^+(x)\,\mathrm{d}\mu,\int_A f^-(x)\,\mathrm{d}\mu\in\mathbb{R}\\&\Leftrightarrow\int_A f^+(x)\,\mathrm{d}\mu+\int_A f^-(x)\,\mathrm{d}\mu\in\mathbb{R}\\&\Leftrightarrow\int_A[f^+(x)+f^-(x)]\,\mathrm{d}\mu=\int_A|f(x)|\,\mathrm{d}\mu\in\mathbb{R}\end{aligned}$$

(5) 因为 $f$ 是一个可测函数，由定理 1.36可知 $f^+, f^-$ 也是可测函数。由定理 1.32可得 $\{f^+=+\infty\},\{f^-=+\infty\}\in\mathscr{F}$。因为 $f$ 可积，所以：

$$\int_A f^+(x)\,\mathrm{d}\mu,\int_A f^-(x)\,\mathrm{d}\mu<+\infty$$

由性质 1.4.2(9) 可得 $\mu[\{f^+=+\infty\}\cap A]=\mu[\{f^-=+\infty\}\cap A]=0$，于是由测度的有限可加性可得：

$$\begin{aligned}\mu(\{|f|=+\infty\}\cap A)&=\mu\Big\{[\{f^-=+\infty\}\cap A]\cup[\{f^+=+\infty\}\cap A]\Big\}\\&=\mu[\{f^-=+\infty\}\cap A]+\mu[\{f^+=+\infty\}\cap A]=0\end{aligned}$$

即 $|f|<+\infty$ a.e. 于 $A$。

(6)

(7) 因为 $f\leqslant g$ a.e. 于 $A$，所以 $f^+\leqslant g^+$ a.e. 于 $A$，$f^-\geqslant g^-$ a.e. 于 $A$。由性质 1.4.2(7) 可得：

$$\int_A f(x)\,\mathrm{d}\mu=\int_A f^+(x)\,\mathrm{d}\mu-\int_A f^-(x)\,\mathrm{d}\mu\leqslant\int_A g^+(x)\,\mathrm{d}\mu-\int_A g^-(x)\,\mathrm{d}\mu=\int_A g(x)\,\mathrm{d}\mu$$

(8) 由 (7) 立即可得。

(9) 第一个结论由性质 1.4.2(10) 显然成立，下证第二个结论。若此时不满足 $f=0$ a.e. 于 $A$，则 $\mu(\{f>0\}\cap A)>0$，于是：

$$\mu\left[\left(\bigcup_{n=1}^{+\infty}\left\{f>\frac{1}{n}\right\}\right)\cap A\right]>0$$

考虑集合序列：

$$A_n=\left(\bigcup_{i=1}^{n}\left\{f>\frac{1}{i}\right\}\right)\cap A=\left\{f>\frac{1}{n}\right\}\cap A,\quad\forall\, n\in\mathbb{N}^+$$

显然 $\{A_n\}$ 是一个单调递增序列，且由 $\sigma$ 域的定义以及定理 1.32可得 $A_n\in\mathscr{F}$ 对任意的 $n\in\mathbb{N}^+$ 成立。于是：

$$\lim_{n\to+\infty}A_n=\bigcup_{n=1}^{+\infty}A_n=\left(\bigcup_{n=1}^{+\infty}\left\{f>\frac{1}{n}\right\}\right)\cap A\in\mathscr{F}$$

由半环上测度的下连续性（定理 1.23）可得：

$$\lim_{n\to+\infty}\mu(A_n)=\mu\left(\lim_{n\to+\infty}A_n\right)>0$$

由极限的性质可知存在 $N\in\mathbb{N}^+$ 使得 $\mu(A_N)>0$，即 $\mu\left(\{f>\frac{1}{N}\}\cap A\right)>0$。依次根据 (8) 和性质 1.4.2(6)、非负简单函数积分的定义、性质 1.4.2(7)、非负简单函数积分的定义可得：

$$\begin{aligned}\int_A f(x)\,\mathrm{d}\mu&=\int_A f(x)I_{\{f\geqslant0\}}\,\mathrm{d}\mu=\int_A f(x)(I_{\{f>0\}}+I_{\{f=0\}})\,\mathrm{d}\mu\\&=\int_A f(x)I_{\{f>0\}}\,\mathrm{d}\mu+\int_A f(x)I_{\{f=0\}}\,\mathrm{d}\mu=\int_A f(x)I_{\{f>0\}}\,\mathrm{d}\mu\\&\geqslant\int_A f(x)I_{A_N}\,\mathrm{d}\mu\geqslant\int_A\frac{1}{N}I_{A_N}\,\mathrm{d}\mu=\frac{\mu(A_N)}{N}>0\end{aligned}$$

矛盾。

(10) 取 $\{f>g\}$，因为 $f,g$ 都是可测函数，由推论 1.4可知 $\{f>g\}\in\mathscr{F}$。根据 (3) 可知：

$$\int_{\{f>g\}} f(x)\,\mathrm{d}\mu, \int_{\{f>g\}} g(x)\,\mathrm{d}\mu \in \mathbb{R}$$

于是由条件和 (6) 有：

$$\int_{\{f>g\}} f(x)\,\mathrm{d}\mu - \int_{\{f>g\}} g(x)\,\mathrm{d}\mu = \int_{\{f>g\}} [f(x)-g(x)]\,\mathrm{d}\mu \leqslant 0$$

因为 $f>g$，由 (7) 可得：

$$\int_{\{f>g\}} [f(x)-g(x)]\,\mathrm{d}\mu \geqslant 0$$

所以有：

$$\int_{\{f>g\}} [f(x)-g(x)]\,\mathrm{d}\mu = 0$$

因为 $f>g$ 在 $\{f>g\}$ 上恒成立，由 (9) 可得 $f-g=0$ a.e. 于 $\{f>g\}$，所以 $\mu(\{f>g\})=0$，即 $f\leqslant g$ a.e. 于 $X$。

(11) 由 (10) 立即可得。 □

**Theorem 1.43** (积分的绝对连续性)**.** *设 $f(x)$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可积函数，对于任意的 $\varepsilon>0$，$\exists\delta>0$，使得对于任意的 $A\subset\mathscr{F}$，只要 $\mu(A)<\delta$，就有：*

$$\int_A |f(x)|\,\mathrm{d}\mu < \varepsilon$$

*Proof.* 由 $f$ 可积和性质 1.4.3(4) 可知 $|f|$ 可积。由性质 1.4.2可知存在非负简单函数列 $\{f_n\}$ 满足 $f_n\uparrow|f|$ 且：

$$\int_A |f(x)|\,\mathrm{d}\mu = \lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

所以对于任意的 $\varepsilon>0$，存在 $N\in\mathbb{N}^+$ 使得：

$$\int_A |f(x)|\,\mathrm{d}\mu < \frac{\varepsilon}{2} + \int_A f_N(x)\,\mathrm{d}\mu$$

取 $f_N$ 在 $A$ 上的最大值 $M$，由非负简单函数积分的定义可得：

$$\int_A |f(x)|\,\mathrm{d}\mu < \frac{\varepsilon}{2} + M\mu(A)$$

所以对于这个 $\varepsilon$ 而言，只要取 $\delta<\dfrac{\varepsilon}{2M}$ 即可。 □

**note 1.1.** *以上定理之所以被称之为绝对连续性，是因为该定理表明，如果将集合看作自变量，用差集的测度定义集合之间的距离，固定的函数 $f$ 在该集合上的积分为因变量，则这个从自变量到因变量的函数 $g$ 一定是绝对连续的。*

**Theorem 1.44** (Levi theorem)**.** *设 $f(x),\{f_n\}$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数。若 $f,\{f_n\}$ a.e. 非负于 $A\in\mathscr{F}$ 且 $f_n\uparrow f$ a.e.，$f,\{f_n\}$ 在 $A$ 上的积分都存在，则：*

$$\int_A f_n(x)\,\mathrm{d}\mu \uparrow \int_A f(x)\,\mathrm{d}\mu$$

*Proof.* 由性质 1.4.3(8)，可仅对 $f,\{f_n\}$ 是非负可测函数且 $f_n \uparrow f$ 讨论。对每个 $n \in \mathbb{N}^+$ 作非负简单函数列 $\{f_{nm}\}$ 使得当 $m \to +\infty$ 时有 $f_{nm} \uparrow f_n$，令 $g_k(x) = \max\limits_{1\leqslant n\leqslant k} f_{nk}(x)$。

**(1) $g_k$ 是非负简单函数：** □

**Theorem 1.45.** 设 $f(x)$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数。若 $f$ 在任意的 $A \in \mathscr{F}$ 上的积分都存在，则对任一可列可测分割 $\{A_n\}$ 有：

$$\int_X f(x)\,\mathrm{d}\mu = \sum_{n=1}^{+\infty}\int_{A_n} f(x)\,\mathrm{d}\mu$$

*Proof.* 构造函数列：

$$f_n(x) = f(x) I_{\bigcup\limits_{i=1}^{n} A_i}$$

则有：

$$f_n^+ \uparrow f^+,\ f_n^- \uparrow f^-$$

由定理 1.33和定理 1.34(2) 可知 $f_n$ 是可测函数，于是由定理 1.36可得 $f_n^+, f_n^-$ 是非负可测函数。根据定理 1.44可知：

$$\int_X f_n^+(x)\,\mathrm{d}\mu \uparrow \int_X f^+(x)\,\mathrm{d}\mu,\quad \int_X f_n^-(x)\,\mathrm{d}\mu \uparrow \int_X f^-(x)\,\mathrm{d}\mu$$

而由性质 1.4.2(5)(10) 可得：

$$\begin{aligned}
\int_X f_n^+(x)\,\mathrm{d}\mu &= \int_{\bigcup\limits_{i=1}^{n} A_n} f_n^+(x)\,\mathrm{d}\mu + \int_{X\setminus\bigcup\limits_{i=1}^{n} A_n} f_n^+(x)\,\mathrm{d}\mu \\
&= \int_{\bigcup\limits_{i=1}^{n} A_n} f^+(x)\,\mathrm{d}\mu = \sum_{i=1}^{n}\int_{A_i} f^+(x)\,\mathrm{d}\mu \\
\int_X f_n^-(x)\,\mathrm{d}\mu &= \int_{\bigcup\limits_{i=1}^{n} A_n} f_n^-(x)\,\mathrm{d}\mu + \int_{X\setminus\bigcup\limits_{i=1}^{n} A_n} f_n^-(x)\,\mathrm{d}\mu \\
&= \int_{\bigcup\limits_{i=1}^{n} A_n} f^-(x)\,\mathrm{d}\mu = \sum_{i=1}^{n}\int_{A_i} f^-(x)\,\mathrm{d}\mu
\end{aligned}$$

于是由极限的线性性质可得：

$$\begin{aligned}
\int_X f(x)\,\mathrm{d}\mu &= \int_X f^+(x)\,\mathrm{d}\mu - \int_X f^-(x)\,\mathrm{d}\mu \\
&= \lim_{n\to+\infty}\left[\int_X f_n^+(x)\,\mathrm{d}\mu\right] - \lim_{n\to+\infty}\left[\int_X f_n^-(x)\,\mathrm{d}\mu\right] \\
&= \lim_{n\to+\infty}\left\{\sum_{i=1}^{n}\left[\int_{A_i} f^+(x)\,\mathrm{d}\mu\right]\right\} - \lim_{n\to+\infty}\left\{\sum_{i=1}^{n}\left[\int_{A_i} f^-(x)\,\mathrm{d}\mu\right]\right\} \\
&= \lim_{n\to+\infty}\left\{\sum_{i=1}^{n}\left[\int_{A_i} f^+(x)\,\mathrm{d}\mu - \int_{A_i} f^-(x)\,\mathrm{d}\mu\right]\right\} \\
&= \lim_{n\to+\infty}\left\{\sum_{i=1}^{n}\left[\int_{A_i} f(x)\,\mathrm{d}\mu\right]\right\} = \sum_{n=1}^{+\infty}\left[\int_{A_n} f(x)\,\mathrm{d}\mu\right]
\end{aligned}$$

□

**Theorem 1.46** (Fatou Lemma)**.** *设 $\{f_n\}$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数列，$A\in\mathscr{F}$，$\{f_n\}$ 在 $A$ 上 a.e. 非负，则：*

$$\int_A\left[\varliminf_{n\to+\infty} f_n(x)\right]\mathrm{d}\mu \leqslant \varliminf_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

*Proof.* 令 $g_k(x)=\inf\limits_{n\geqslant k} f_n(x)$，则 $g_k\uparrow \varliminf\limits_{n\to+\infty} f_n$。因为 $\{f_n\}$ 是可测函数列，由定理 1.35可知 $\{g_k\}$ 也是可测函数列。因为 $\{f_n\}$ 在 $A$ 上 a.e. 非负，根据半环上测度的次可列可加性、单调性（定理 1.23）和测度的非负性可得：

$$\mu\left[\bigcup_{n=1}^{+\infty}(\{f_n<0\}\cap A)\right]\leqslant\sum_{n=1}^{+\infty}\mu(\{f_n<0\}\cap A)=0$$

所以 $\{g_k\}$ 非负 a.e. 于 $A$，于是 $\varliminf\limits_{n\to+\infty} f_n$ 非负 a.e. 于 $A$。由定理 1.35可知 $\varliminf\limits_{n\to+\infty} f_n$ 是可测函数，所以由定理 1.44可得：

$$\int_A\left[\varliminf_{n\to+\infty} f_n(x)\right]\mathrm{d}\mu=\int_A\left[\lim_{k\to+\infty} g_k(x)\right]\mathrm{d}\mu=\lim_{k\to+\infty}\left[\int_A g_k(x)\,\mathrm{d}\mu\right]$$

因为：

$$g_k(x)\leqslant f_n(x),\ \forall\, n\geqslant k$$

由性质 1.4.3(7) 可得：

$$\int_A g_k(x)\,\mathrm{d}\mu\leqslant\int_A f_n(x)\,\mathrm{d}\mu,\ \forall\, n\geqslant k$$

所以：

$$\int_A g_k(x)\,\mathrm{d}\mu\leqslant\inf_{n\geqslant k}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

由极限的不等式性可得：

$$\lim_{k\to+\infty}\left[\int_A g_k(x)\,\mathrm{d}\mu\right]\leqslant\lim_{k\to+\infty}\left\{\inf_{n\geqslant k}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]\right\}=\varliminf_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

即：

$$\int_A\left[\varliminf_{n\to+\infty} f_n(x)\right]\mathrm{d}\mu\leqslant\varliminf_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

□

**Corollary 1.5.** *设 $\{f_n\}$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数列，$A\in\mathscr{F}$。*

1. *若存在上述测度空间上的在 $A$ 上可积的函数 $g$ 使得 $f_n\geqslant g$ a.e. 于 $A$ 对任意的 $n\in\mathbb{N}^+$ 成立，则 $\varliminf\limits_{n\to+\infty} f_n(x)$ 的积分存在且：*

$$\int_A\left[\varliminf_{n\to+\infty} f_n(x)\right]\mathrm{d}\mu\leqslant\varliminf_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

2. *若存在上述测度空间上的在 $A$ 上可积的函数 $g$ 使得 $f_n\leqslant g$ a.e. 于 $A$ 对任意的 $n\in\mathbb{N}^+$ 成立，则 $\varlimsup\limits_{n\to+\infty} f_n(x)$ 的积分存在且：*

$$\int_A\left[\varlimsup_{n\to+\infty} f_n(x)\right]\mathrm{d}\mu\geqslant\varlimsup_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

*Proof.* (1) 构造 a.e. 非负的可测函数列 $\{h_n = f_n - g\}$（定理 1.34(1)），由性质 1.4.3(8) 可将 $\{h_n\}$ 就看做非负可测函数。根据下极限的性质、性质 1.4.2(6) 和定理 1.46可得：

$$\int_A \left[\varliminf_{n\to+\infty} h_n(x)\right] \mathrm{d}\mu \leqslant \varliminf_{n\to+\infty} \left[\int_A h_n(x)\,\mathrm{d}\mu\right]$$

由下极限的性质：

$$\varliminf_{n\to+\infty} f_n(x) = \varliminf_{n\to+\infty} [h_n(x) + g(x)] = \varliminf_{n\to+\infty} h_n(x) + g(x)$$

同时由定理 1.35可得 $\varliminf\limits_{n\to+\infty} h_n(x)$ 是一个非负可测函数。因为 $g(x)$ 可积，由性质 1.4.3(6) 可得：

$$\int_A \left[\varliminf_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu = \int_A \left[\varliminf_{n\to+\infty} h_n(x) + g(x)\right] \mathrm{d}\mu = \int_A \left[\varliminf_{n\to+\infty} h_n(x)\right] \mathrm{d}\mu + \int_A g(x)\,\mathrm{d}\mu$$

同理可得：

$$\begin{aligned}
\varliminf_{n\to+\infty} \left[\int_A f(x)\,\mathrm{d}\mu\right] &= \varliminf_{n\to+\infty} \left\{\int_A [h_n(x) + g(x)]\,\mathrm{d}\mu\right\} \\
&= \varliminf_{n\to+\infty} \left[\int_A h_n(x)\,\mathrm{d}\mu + \int_A g(x)\,\mathrm{d}\mu\right] \\
&= \varliminf_{n\to+\infty} \left[\int_A h_n(x)\,\mathrm{d}\mu\right] + \int_A g(x)\,\mathrm{d}\mu
\end{aligned}$$

于是：

$$\int_A \left[\varliminf_{n\to+\infty} h_n(x)\right] \mathrm{d}\mu + \int_A g(x)\,\mathrm{d}\mu \leqslant \varliminf_{n\to+\infty} \left[\int_A h_n(x)\,\mathrm{d}\mu\right] + \int_A g(x)\,\mathrm{d}\mu$$

$$\int_A \left[\varliminf_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu \leqslant \varliminf_{n\to+\infty} \left[\int_A f_n(x)\,\mathrm{d}\mu\right]$$

(2) 构造 a.e. 非负的可测函数列 $\{h_n = g - f_n\}$，与 (1) 的证明类似。 □

**Theorem 1.47** (Lebesgue 控制收敛定理)**.** *设 $\{f_n\}$ 是测度空间 $(X, \mathscr{F}, \mu)$ 上的可测函数列。若存在 $A \in \mathscr{F}$ 上的非负可积函数 $g$ 使得对任意的 $n \in \mathbb{N}^+$ 有 $|f_n| \leqslant g$ a.e. 于 $A$，则 $f_n \xrightarrow{a.e.} f$ 或 $f_n \xrightarrow{\mu} f$ 蕴含：*

$$\lim_{n\to+\infty} \left[\int_A f_n(x)\,\mathrm{d}\mu\right] = \int_A f(x)\,\mathrm{d}\mu$$

*Proof.* **(1) a.e.** 由性质 1.4.3(8)、极限与上下极限的关系、推论 1.5(1)、上下极限的性质和推论 1.5(2) 可得：

$$\begin{aligned}
\int_A f(x)\,\mathrm{d}\mu &= \int_A \left[\lim_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu = \int_A \left[\varliminf_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu \\
&\leqslant \varliminf_{n\to+\infty} \left[\int_A f_n(x)\,\mathrm{d}\mu\right] \leqslant \varlimsup_{n\to+\infty} \left[\int_A f_n(x)\,\mathrm{d}\mu\right] \\
&\leqslant \int_A \left[\varlimsup_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu = \int_A \left[\lim_{n\to+\infty} f_n(x)\right] \mathrm{d}\mu \\
&= \int_A f(x)\,\mathrm{d}\mu
\end{aligned}$$

于是有：

$$\underline{\lim}_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]=\overline{\lim}_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]=\int_A f(x)\,\mathrm{d}\mu$$

由极限与上下极限的关系可得：

$$\lim_{n\to+\infty}\left[\int_A f_n(x)\,\mathrm{d}\mu\right]=\int_A f(x)\,\mathrm{d}\mu$$

**(2)** $\mu$ 由定理 1.42(3)(1) 和 (1) 立即可得。 □

**Corollary 1.6** (Lebesgue 有界收敛定理)**.** 设 $\{f_n\}$ 和 $f$ 是测度空间 $(X,\mathscr{F},\mu)$ 上的可测函数，$A\in\mathscr{F}$ 且 $\mu(A)<+\infty$。若存在 $M>0$ 使得对任意的 $n\in\mathbb{N}^+$ 有 $|f_n|\leqslant M$ *a.e.* 于 $A$，则 $f_n\xrightarrow{a.e.}f$ 或 $f_n\xrightarrow{\mu}f$ 蕴含：

$$\lim_{n\to+\infty}\left[\int_A f(x)\,\mathrm{d}\mu\right]=\int_A f(x)\,\mathrm{d}\mu$$

*Proof.* 令 $g\equiv M$，于是 $g$ 在 $A$ 上可积。由定理 1.47直接可得结论。 □

**Theorem 1.48.** 设 $f$ 是由测度空间 $(X,\mathscr{F},\mu)$ 到可测空间 $(Y,\mathscr{C})$ 上的可测映射，对于任意的 $A\in\mathscr{C}$，令 $\nu(A)=\mu[f^{-1}(A)]$。

1. $(Y,\mathscr{C},\nu)$ 是一个测度空间；
2. 对 $(Y,\mathscr{C},\nu)$ 上的任何可测函数 $g$ 和任意 $B\in\mathscr{C}$，只要

$$\int_B g(y)\,\mathrm{d}\nu,\quad \int_{f^{-1}(B)} g\circ f\,\mathrm{d}\mu$$

之一有意义，则二者一定相等。

*Proof.* (1) 因为 $\mu$ 是测度，所以 $\mu$ 具有非负性，从而 $\nu$ 也是一个非负集函数。由**??**(1) 和 $\mu(\varnothing)=0$ 可知 $\nu(\varnothing)=0$。任取 $\mathscr{C}$ 中互不相交的集合序列 $\{A_n\}$ 且满足 $\bigcup\limits_{n=1}^{+\infty}A_n\in\mathscr{C}$，由**??**(4) 可知：

$$\nu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\mu\left[f^{-1}\left(\bigcup_{n=1}^{+\infty}A_n\right)\right]=\mu\left[\bigcup_{n=1}^{+\infty}f^{-1}(A_n)\right]$$

因为 $f$ 是一个映射，$\{A_n\}$ 互不相交，所以 $\{f^{-1}(A_n)\}$ 也互不相交。由测度的可列可加性：

$$\nu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\sum_{n=1}^{+\infty}\mu[f^{-1}(A_n)]=\sum_{n=1}^{+\infty}\nu(A_n)$$

所以 $\nu$ 是可测空间 $(Y,\mathscr{C})$ 上的一个测度，即 $(Y,\mathscr{C},\nu)$ 是一个测度空间。

(2) **非负简单函数**：取 $(Y,\mathscr{C},\nu)$ 上的非负简单函数 $g$：

$$g(y)=\sum_{i=1}^{n}a_iI_{A_i}(y),\quad a_i\geqslant 0,\ A_i\in\mathscr{C},\ i=1,2,\ldots,n,\ \bigcup_{i=1}^{n}A_i=Y$$

于是：

$$\int_B g(y)\,\mathrm{d}\nu=\sum_{i=1}^{n}a_i\nu(A_i\cap B)=\sum_{i=1}^{n}a_i\mu[f^{-1}(A_i\cap B)]$$

由??(4) 可得：

$$\begin{aligned}\int_{f^{-1}(B)} g\circ f\,\mathrm{d}\mu &= \int_{f^{-1}(B)} g[f(x)]\,\mathrm{d}\mu = \int_{f^{-1}(B)} \sum_{i=1}^{n} a_i I_{A_i}[f(x)]\,\mathrm{d}\mu \\ &= \sum_{i=1}^{n} a_i \int_{f^{-1}(B)} I_{A_i}[f(x)]\,\mathrm{d}\mu = \sum_{i=1}^{n} a_i \int_{f^{-1}(B)} I_{f^{-1}(A_i)}\,\mathrm{d}\mu \\ &= \sum_{i=1}^{n} a_i\mu[f^{-1}(A_i)\cap f^{-1}(B)] = \sum_{i=1}^{n} a_i\mu[f^{-1}(A_i\cap B)]\end{aligned}$$

即：

$$\int_B g(y)\,\mathrm{d}\nu = \int_{f^{-1}(B)} g\circ f\,\mathrm{d}\mu$$

**非负可测函数**：取 $(Y,\mathscr{C},\nu)$ 上的非负可测函数 $g$，由定理 1.37(1) 可知存在非负简单函数列 $\{g_n\}$ 使得 $g_n\uparrow g$。由非负简单函数时的情形我们得到：

$$g_n\circ f = g_n[f(x)] = \sum_{i=1}^{j_n} a_{ni} I_{A_{ni}}[f(x)] = \sum_{i=1}^{j_n} a_{ni} I_{f^{-1}(A_{ni})},\quad \bigcup_{i=1}^{j_n} A_{ni} = Y$$

其中 $a_{ni}>0,\ A_{ni}\in\mathscr{C}$。由可测函数的定义，$f^{-1}(A_{ni})\in\mathscr{F}$，所以 $g_n\circ f$ 也是一个非负简单函数。因为 $g_n\uparrow g$，所以有 $g_n\circ f\uparrow g\circ f$，由定理 1.44和非负简单函数时的结论可得：

$$\begin{aligned}\int_B g(y)\,\mathrm{d}\nu &= \int_B \left[\lim_{n\to+\infty} g_n(y)\right]\mathrm{d}\nu = \lim_{n\to+\infty}\left[\int_B g_n(y)\,\mathrm{d}\nu\right] \\ &= \lim_{n\to+\infty}\left[\int_{f^{-1}(B)} g_n\circ f\,\mathrm{d}\mu\right] = \int_{f^{-1}(B)}\left[\lim_{n\to+\infty} g_n\circ f\right]\mathrm{d}\mu \\ &= \int_{f^{-1}(B)} g\circ f\,\mathrm{d}\mu\end{aligned}$$

**一般可测函数**：取 $(Y,\mathscr{C},\nu)$ 上的一般可测函数 $g$，由非负可测函数时的情形可得：

$$\begin{aligned}\int_B g(y)\,\mathrm{d}\nu &= \int_B g^+(y)\,\mathrm{d}\nu - \int_B g^-(y)\,\mathrm{d}\nu \\ &= \int_{f^{-1}(B)} g^+\circ f\,\mathrm{d}\mu - \int_{f^{-1}(B)} g^-\circ f\,\mathrm{d}\mu \\ &= \int_{f^{-1}(B)} (g\circ f)^+\,\mathrm{d}\mu - \int_{f^{-1}(B)} (g\circ f)^- dif\mu \\ &= \int_{f^{-1}(B)} g\circ f\,\mathrm{d}\mu\end{aligned}$$

□

**note 1.2.** *以上定理是积分的变量替换定理，和 Riemann 积分中的换元积分法是一样的。若对 $g(x)$ 的直接积分不太好求，可以将 $g(x)$ 变为 $g[f(x)]$ 来求积分。*

## 1.5 $L_p$ 与 $L_\infty$

### 1.5.1 $L_p$

**Definition 1.37.** 设 $(X,\mathscr{F},\mu)$ 是一个测度空间，$p>1$，$E\in\mathscr{F}$，$f,g$ 是 $E$ 上的可测函数。定义等价关系 $\sim$：$f\sim g\Leftrightarrow f=g$ *a.e.* 于 $E$，将商空间：

$$\left\{f:\int_E |f(x)|^p\,\mathrm{d}x<+\infty\right\}/\sim$$

称之为 $E$ 上的 $L_p(X,\mathscr{F},\mu)$ 空间，简记为 $L_p(E)$。

**Property 1.5.1.** 设 $(X,\mathscr{F},\mu)$ 是一个测度空间，$p>1$，$E\in\mathscr{F}$，则 $L_p(E)$ 是一个线性空间。

*Proof.* 由**??**和性质 1.4.2(6) 直接得到。 □

#### $L_p(X,\mathscr{F},\mu)$ 上的距离

**Definition 1.38.** 设 $(X,\mathscr{F},\mu)$ 是一个测度空间，$p>1$，$E\in\mathscr{F}$。在 $L_p(E)$ 中定义元素 $x=x(t)$ 和元素 $y=y(t)$ 之间的距离为：

$$\rho(x,y)=\left[\int_E |x(t)-y(t)|^p\,\mathrm{d}\mu\right]^{\frac{1}{p}}$$

则 $(L_p(E),\rho)$ 是一个度量空间。

下证明上式定义的距离满足距离公理：

*Proof.* (1) $\rho\in R$：由**??**可得：

$$|x(t)-y(t)|^p\leqslant\Big[|x(t)|+|y(t)|\Big]^p\leqslant 2^{p-1}\Big[|x(t)|^p+|y(t)|^p\Big]$$

于是：

$$\rho(x,y)\leqslant\left[2^{p-1}\int_E |x(t)|^p\,\mathrm{d}\mu+2^{p-1}\int_E |y(t)|^p\,\mathrm{d}\mu\right]^{\frac{1}{p}}$$

由 $L_p(X,\mathscr{F},\mu)$ 空间定义，$\rho\in\mathbb{R}$。(2) 非负性由性质 1.4.2(2)(10) 直接可得；(3) 对称性直接可得；(4) 三角不等式：设 $x(t),y(t),z(t)\in L_p(E)$。$p=1$ 时可由绝对值的三角不等式立即得到，$p>1$ 时，由 Minkowski 不等式（即**??**）可得到：

$$\left[\int_E |x(t)-z(t)|^p\,\mathrm{d}\mu\right]^{\frac{1}{p}}\leqslant\left[\int_E |x(t)-y(t)|^p\,\mathrm{d}\mu\right]^{\frac{1}{p}}+\left[\int_E |y(t)-z(t)|^p\,\mathrm{d}\mu\right]^{\frac{1}{p}}$$

即：

$$\rho(x,z)\leqslant\rho(x,y)+\rho(y,z)$$ □

### $L_p(X, \mathscr{F}, \mu)$ 上的范数

**Definition 1.39.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$p > 1$，$E \in \mathscr{F}$。在 $L_p(E)$ 中定义元素 $x = x(t)$ 的范数为：

$$||x||_p = \left[\int_E |x(t)|^p \,\mathrm{d}t\right]^{\frac{1}{p}}$$

则 $L_p(X, \mathscr{F}, \mu)(E)$ 成为一个赋范线性空间。

*Proof.* (1) 由 $L_p(X, \mathscr{F}, \mu)$ 的定义即可得到 $||x|| \in \mathbb{R}$。(2) 非负性由性质 1.4.2(2)(10) 直接得到。(3) 数乘由性质 1.4.2(6) 直接得到，(4) 三角不等式的证明可由 Minkowski 不等式（即**??**）直接得到。 □

## 1.5.2 $L_\infty$

**Definition 1.40.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$f$ 是 $\mathscr{F}$ 上的一个可测函数，令：

$$G(f) = \Big\{c > 0 : \mu(\{x : |f(x)| > c\}) = 0\Big\}$$

称：

$$||f||_\infty = \begin{cases} \inf G(f), & G(f) \neq \varnothing \\ +\infty, & G(f) = \varnothing \end{cases}$$

为 $f$ 的无穷范数，也称其为 $f$ 的本性上确界 *(essential supremum)*。

**Theorem 1.49.** 无穷范数具有如下等价定义：

$$||f||_\infty = \inf_{me=0} \sup_{x \in E \setminus e} |f(x)|$$

*Proof.* 等价性由定义可直接证得，略去。 □

**Definition 1.41.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$，$f, g$ 是 $E$ 上的可测函数。定义等价关系 $\sim$：$f \sim g \Leftrightarrow f = g$ *a.e.* 于 $E$，称商空间 $\{f : ||f||_\infty < +\infty\}/\sim$ 为 $E$ 上的 $L_\infty$ 空间，简记为 $L_\infty(E)$。

**Property 1.5.2.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$，则 $L_\infty(E)$ 是一个线性空间。

*Proof.* 任取 $x, y \in L_\infty(E)$ 和 $\alpha, \beta \in \mathbb{R}$。因为 $x, y$ 各自除了一个零测集外有界，由由半环上测度的次有限可加性（定理 1.23中的次可列可加性推得这两个零测集的并集也是零测集，于是 $\alpha x + \beta y$ 在除了一个零测集以外的空间上作的是 $\mathbb{R}$ 中的加减乘除，于是 $\alpha x + \beta y$ 除了一个零测集外有界，所以 $\alpha x + \beta y \in L_\infty(E)$，即 $L_\infty(E)$ 是一个线性空间。 □

### $L_\infty$ 上的距离

**Definition 1.42.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$。在 $L_\infty(E)$ 中定义元素 $x = x(t)$ 和元素 $y = y(t)$ 之间的距离为：

$$\rho(x, y) = ||x - y||_\infty$$

则 $(L_\infty(E), \rho)$ 是一个度量空间。

下证明上式定义的距离满足距离公理：

*Proof.* (1) $\rho \in R$:因为 $L_\infty$ 是一个线性空间,所以 $x-y \in L_\infty(E)$,由定义可得 $||x-y||_\infty \in \mathbb{R}$。

(2) 非负性：由无穷范数的定义：

$$\|x - y\|_\infty \geqslant 0.$$

且有 $\|x - y\|_\infty = 0$ 当且仅当有 $|x(t) - y(t)| = 0$ a.e. 于 $E$，即 $x = y$。

(3) 对称性显然。

(4) 三角不等式：取任意 $x, y, z \in L_\infty(E)$。由无穷范数等价定义以及下确界定义，对任意的 $\varepsilon > 0$，存在 $E_1, E_2 \subset E$, $\mu(E_1) = \mu(E_2) = 0$，使得：

$$\sup_{t \in E \backslash E_1} |x(t) - z(t)| \leqslant \rho(x, z) + \frac{\varepsilon}{2}, \quad \sup_{t \in E \backslash E_2} |z(t) - y(t)| \leqslant \rho(z, y) + \frac{\varepsilon}{2}$$

由半环上测度的次有限可加性（定理 1.23中的次可列可加性推得），$\mu(E_1 \cup E_2) \leqslant \mu(E_1) + \mu(E_2) = 0$，所以：

$$\begin{aligned}
\rho(x, y) &\leqslant \sup_{t \in E \backslash (E_1 \cup E_2)} |x(t) - y(t)| \\
&\leqslant \sup_{t \in E \backslash (E_1 \cup E_2)} |x(t) - z(t)| + \sup_{t \in E \backslash (E_1 \cup E_2)} |z(t) - y(t)| \\
&\leqslant \sup_{t \in E \backslash E_1} |x(t) - z(t)| + \sup_{t \in E \backslash E_2} |z(t) - y(t)| \\
&\leqslant \rho(x, z) + \rho(z, y) + \varepsilon
\end{aligned}$$

由 $\varepsilon$ 的任意性可得到：

$$\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$$

□

### $L_\infty$ 上的范数

**Definition 1.43.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$。在 $L_\infty(E)$ 中定义元素 $x = x(t)$ 的范数为：

$$||x|| = ||x||_\infty$$

则 $L_\infty(E)$ 成为一个赋范线性空间。

下证明上式定义的范数满足范数定义：

*Proof.* (1) $||x|| \in \mathbb{R}$ 由 $L_\infty(E)$ 空间的定义直接可得。(2) 非负性和 (3) 数乘由无穷范数的定义是显然的。

(4) 三角不等式：由本性上确界定义可得：

$$|x(t)| \leqslant ||x||_\infty,\ |y(t)| \leqslant ||y||_\infty$$

a.e. 于 $E$，所以：

$$|x(t)+y(t)| \leqslant |x(t)| + |y(t)| \leqslant ||x||_\infty + ||y||_\infty$$

a.e. 于 $E$。于是 $||x(t)||_\infty + ||y(t)||_\infty \in G(x+y)$，所以：

$$||x+y||_\infty = \inf G(x+y) \leqslant ||x||_\infty + ||y||_\infty \qquad \square$$

### 1.5.3 收敛性

**Theorem 1.50.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$，$1 \leqslant p \leqslant +\infty$。若 $\{f_n\} \subset L_p(E)$ 且：

$$m > n,\ \lim_{n \to +\infty} ||f_m - f_n||_p = 0$$

则存在 $f \in L_p(E)$ 使得 $\{f_n\}$ 依范数收敛于 $f$，即在 $L_p(E)$ 中引入范数导出的距离时，$L_p(E)$ 是一个 *Banach* 空间。

*Proof.* □

**Definition 1.44.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$，$1 \leqslant p \leqslant +\infty$，$\{f_n\} \subset L_p(E)$，$f \in L_p(E)$。若：

$$\lim_{n \to +\infty} ||f_n - f||_p = 0$$

则称 $\{f_n\}$（$p$ 阶）**平均收敛**到 $f$，记为 $f_n \xrightarrow{L_p} f$。

**Theorem 1.51.** 设 $(X, \mathscr{F}, \mu)$ 是一个测度空间，$E \in \mathscr{F}$，$1 \leqslant p \leqslant +\infty$，$\{f_n\} \subset L_p(E)$，$f \in L_p(E)$。

1. 若 $f_n \xrightarrow{L_p} f$，则 $f_n \xrightarrow{\mu} f$ 且 $||f_n||_p \to ||f||_p$；
2. 若 $f_n \xrightarrow{a.e.} f$ 或 $f_n \xrightarrow{\mu} f$，则：

$$||f_n||_p \to ||f||_p \Leftrightarrow f_n \xrightarrow{L_p} f$$

*Proof.* (1) 设 $f_n \xrightarrow{L_p} f$，则对任意的 $\varepsilon > 0$，由性质 1.4.2(7) 可得：

$$\begin{aligned}
\mu(|f_n - f| \geqslant \varepsilon) &= \int_{\{|f_n - f| \geqslant \varepsilon\}} \mathrm{d}\mu \leqslant \int_{\{|f_n - f| \geqslant \varepsilon\}} \frac{|f_n - f|^p}{\varepsilon^p} \mathrm{d}\mu \\
&\leqslant \frac{1}{\varepsilon^p} \int_{-\infty}^{+\infty} |f_n - f|^p \,\mathrm{d}\mu = \frac{||f_n - f||_p^p}{\varepsilon^p}
\end{aligned}$$

而 $||f_n - f|| \to 0$，所以：

$$\lim_{n\to+\infty} \mu(||f_n - f|| \geqslant \varepsilon) = 0$$

即 $f_n \xrightarrow{\mu} f$。由**??**可得：

$$\Big| ||f_n||_p - ||f||_p \Big| \leqslant ||f_n - f||_p$$

于是 $||f_n||_p \to ||f||_p$。

(2) 充分性由 (1) 直接得到，下证必要性。 □

## 1.6 不定积分

### 1.6.1 符号测度

**Definition 1.45.** 设 $(X, \mathscr{F})$ 是一个可测空间，称从 $\mathscr{F}$ 到 $\overline{R}$ 的集函数 $\varphi$ 为符号测度 *(signed measure)*，若 $\varphi$ 满足：

*1.* $\varphi(\varnothing) = 0$；

*2.* $\varphi$ 具有可列可加性。

若对任意的 $A \in \mathscr{F}$ 有 $|\varphi(A)| < +\infty$，则称 $\varphi$ 是有限的；若存在 $X$ 的可列可测分割 $\{A_n\} \subset \mathscr{F}$ 满足对任意的 $n \in \mathbb{N}^+$ 有 $|\varphi(A_n)| < +\infty$，则称 $\varphi$ 是 $\sigma$ 有限的。

**Property 1.6.1.** 设 $(X, \mathscr{F})$ 是一个可测空间，其上的符号测度 $\varphi$ 具有如下性质：

*1.* $\varphi$ 具有有限可加性；

*2.* $\varphi$ 只可能出现以下两种情况中的一种：

$$\forall\ A \in \mathscr{F},\ -\infty < \varphi(A) \leqslant +\infty$$
$$\forall\ A \in \mathscr{F},\ -\infty \leqslant \varphi(A) < +\infty$$

*3.* 若 $A, B \in \mathscr{F}$，$B \subseteq A$ 且 $|\varphi(A)| < +\infty$，则 $|\varphi(B)| < +\infty$；

*4.* 若 $\{A_n\}$ 两两不交且满足：

$$\left|\varphi\left(\bigcup_{n=1}^{+\infty} A_n\right)\right| < +\infty$$

则有：

$$\sum_{n=1}^{+\infty} |\varphi(A_n)| < +\infty$$

*Proof.* (1) 由符号测度的定义显然可得。

(2) 设 $A, B \in \mathscr{F}$ 且 $\varphi(A) = +\infty, \varphi(B) = -\infty$，则由 (1) 可得：

$$\varphi(A \cup B) = \varphi(A) + \varphi(B\backslash A) = \varphi(B) + \varphi(A\backslash B)$$

要使得上式有意义，$\varphi(A \cup B)$ 必须既等于 $+\infty$ 又等于 $-\infty$，矛盾。

(3) 由 (1) 可得 $\varphi(A) = \varphi(B) + \varphi(A \backslash B)$，当 $|\varphi(A)| < +\infty$ 时，上式有意义必须满足 $|\varphi(B)| < +\infty$。

(4) 记：

$$A_n^+ = \begin{cases} \varnothing, & \varphi(A_n) \leqslant 0 \\ A_n, & \varphi(A_n) > 0 \end{cases} \quad A_n^- = \begin{cases} A_n, & \varphi(A_n) \leqslant 0 \\ \varnothing, & \varphi(A_n) > 0 \end{cases}$$

则：

$$\bigcup_{n=1}^{+\infty} A_n = \left(\bigcup_{n=1}^{+\infty} A_n^+\right) \cup \left(\bigcup_{n=1}^{+\infty} A_n^-\right)$$

由 (3) 可得：

$$\left|\varphi\left(\bigcup_{n=1}^{+\infty} A_n^+\right)\right| < +\infty, \quad \left|\varphi\left(\bigcup_{n=1}^{+\infty} A_n^-\right)\right| < +\infty$$

因为 $\{A_n\}$ 互不相交，由 $\{A_n^+\}, \{A_n^-\}$ 的构造方式显然二者内部互不相交且二者之间也互不相交，于是：

$$\begin{aligned} \sum_{n=1}^{+\infty} |\varphi(A_n)| &= \sum_{n=1}^{+\infty} [|\varphi(A_n^+)| + |\varphi(A_n^-)|] = \sum_{n=1}^{+\infty} \varphi(A_n^+) + \sum_{n=1}^{+\infty} |\varphi(A_n^-)| \\ &= \sum_{n=1}^{+\infty} \varphi(A_n^+) + \left|\sum_{n=1}^{+\infty} \varphi(A_n^-)\right| = \varphi\left(\bigcup_{n=1}^{+\infty} A_n^+\right) + \left|\varphi\left(\bigcup_{n=1}^{+\infty} A_n^-\right)\right| < +\infty \end{aligned} \qquad \square$$

# Chapter 2

# 随机变量的数字特征

## 2.1 期望

**Definition 2.1.** 设 $X$ 是一个随机变量，具有分布函数 $F(x)$。若 $X$ 的积分有限，则称：

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x \,\mathrm{d}F(x)$$

为 $X$ 的数学期望 *(mathematical expectation)*。

## 2.2 方差

**Property 2.2.1.** 设 $X, Y$ 是随机变量，则：

*1.* $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$；

*2.* $\mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}[\mathrm{E}(X|Y)]$；

*3.* $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) \pm \mathrm{Cov}(X, Y) + \mathrm{Var}(Y)$，若 $X, Y$ 不相关，则有 $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$；

*Proof.* (1) 设 $\mathrm{E}(X) = \mu$，由方差的定义：

$$\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \mathrm{E}(X^2 - 2\mu X + \mu^2) = \mathrm{E}(X^2) - 2\mu^2 + \mu^2 = \mathrm{E}(X^2) - \mu^2$$

(2) 由 (1) 可得：

$$\begin{aligned}
\mathrm{E}[\mathrm{Var}(X|Y)] &= \mathrm{E}\{\mathrm{E}(X^2|Y) - [\mathrm{E}(X|Y)]^2\} \\
&= \mathrm{E}[\mathrm{E}(X^2|Y)] - \mathrm{E}\{[\mathrm{E}(X|Y)]^2\} \\
&= \mathrm{E}(X^2) - \mathrm{E}\{[\mathrm{E}(X|Y)]^2\} \\
\mathrm{Var}[\mathrm{E}(X|Y)] &= \mathrm{E}\{[\mathrm{E}(X|Y)]^2\} - \{\mathrm{E}[\mathrm{E}(X|Y)]\}^2 \\
&= \mathrm{E}\{[\mathrm{E}(X|Y)]^2\} - [\mathrm{E}(X)]^2
\end{aligned}$$

于是：

$$\mathrm{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}[\mathrm{E}(X|Y)] = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \mathrm{Var}(X)$$

(3) 由方差的定义可得：

$$\begin{aligned}
\mathrm{Var}(X \pm Y) &= \mathrm{E}[X \pm Y - \mathrm{E}(X \pm Y)]^2 \\
&= \mathrm{E}\{[X - \mathrm{E}(X) \pm [Y - \mathrm{E}(Y)]]\}^2 \\
&= \mathrm{E}\{[X - \mathrm{E}(X)]^2 \pm 2[X - \mathrm{E}(X)][Y - \mathrm{E}(Y)] + [Y - \mathrm{E}(Y)]^2\} \\
&= \mathrm{Var}(X) \pm 2\,\mathrm{Cov}(X, Y) + \mathrm{Var}(Y)
\end{aligned}$$

□

## 2.3　矩

### 2.3.1　原点矩

**Definition 2.2.** 设 $X$ 是一个随机变量，$n \in \mathbb{N}^+$。若数学期望：

$$\mu_n = \mathrm{E}(X^n)$$

存在，则称 $\mu_n$ 为 $X$ 的 $n$ 阶原点矩 *(raw moment)*。

**Definition 2.3.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量，$\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{N}$。若数学期望：

$$\mu_{\alpha_1,\alpha_2,\ldots,\alpha_n} = \mathrm{E}\left(\prod_{i=1}^{n} \mathbf{X}_i^{\alpha_i}\right)$$

存在，则称 $\mu_{\alpha_1,\alpha_2,\ldots,\alpha_n}$ 为 $\mathbf{X}$ 的阶数为 $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ 的原点矩。

**Theorem 2.1.** 设 $X$ 是一个随机变量，$m \in \mathbb{N}^+$。若 $X$ 的 $m$ 阶原点矩 $\mu_m$ 存在，则 $X$ 具有所有不超过 $m$ 阶的原点矩。

*Proof.* 取任意的 $n < m$ 且 $n \in \mathbb{N}^+$，则显然：

$$|x^n| \leqslant \begin{cases} 1, & |x| \leqslant 1 \\ |x^m|, & |x| > 1 \end{cases}$$

于是 $|x^n| \leqslant 1 + |x^m|$。因为 $X$ 有 $m$ 阶中心矩，所以：

$$\int_{-\infty}^{+\infty} |x^m| p(x)\,\mathrm{d}x < +\infty$$

于是：

$$\int_{-\infty}^{+\infty} |x|^n p(x)\,\mathrm{d}x \leqslant \int_{-\infty}^{+\infty} (|x|^m + 1) p(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} |x|^m p(x)\,\mathrm{d}x + 1 < +\infty$$

所以 $X$ 具有 $n$ 阶原点矩。由 $n$ 的任意性，结论成立。 □

### 2.3.2 中心矩

**Definition 2.4.** 设 $X$ 是一个随机变量，$\mu_1 = \mathrm{E}(X)$，$n \in \mathbb{N}^+$。若数学期望：

$$\nu_n = \mathrm{E}[(X - \mu_1)^n]$$

存在，则称 $\nu_n$ 为 $X$ 的 $n$ 阶中心矩 *(central moment)*。

**Definition 2.5.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量，$\mu = \mathrm{E}(\mathbf{X})$，$\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{N}$。若数学期望：

$$\nu_{\alpha_1,\alpha_2,\ldots,\alpha_n} = \mathrm{E}\left[\prod_{i=1}^{n} (\mathbf{X}_i - \mu_i)^{\alpha_i}\right]$$

存在，则称 $\nu_{\alpha_1,\alpha_2,\ldots,\alpha_n}$ 为 $\mathbf{X}$ 的阶数为 $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ 的中心矩。

**Theorem 2.2.** 随机变量 $X$ 的中心矩 $\nu_n$ 与原点矩 $\mu_n$ 之间存在如下关系：

$$\nu_n = \sum_{i=0}^{n} \binom{n}{i} \mu_i (-\mu_1)^{n-i}$$

*Proof.* 由中心矩的定义可得：

$$\nu_n = \mathrm{E}[(X - \mu_1)^n] = \mathrm{E}\left[\sum_{i=0}^{n} \binom{n}{i} X^i (-\mu_1)^{n-i}\right] = \sum_{i=0}^{n} \binom{n}{i} \mu_i (-\mu_1)^{n-i}$$

□

## 2.4 协方差

**Definition 2.6.** 随机向量 $\mathbf{X}$ 与随机向量 $\mathbf{Y}$ 的协方差 *(covariance)* 矩阵定义为：

$$\mathrm{Cov}(\mathbf{X}, \mathbf{Y}) = \mathrm{E}\Big[\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y} - \mathrm{E}(\mathbf{Y})\Big)^T\Big]$$

若 $\mathbf{X} = \mathbf{Y}$，则可将 $\mathrm{Cov}(\mathbf{X}, \mathbf{Y})$ 简写为 $\mathrm{Cov}(\mathbf{X})$。

**Definition 2.7.** 设 $X, Y$ 是两个随机变量，则：

1. 若 $\mathrm{Cov}(X, Y) > 0$，称 $X, Y$ 正相关 *(positively correlated)*；
2. 若 $\mathrm{Cov}(X, Y) < 0$，称 $X, Y$ 负相关 *(negatively correlated)*；
3. 若 $\mathrm{Cov}(X, Y) = 0$，称 $X, Y$。

**Property 2.4.1.** 协方差矩阵具有如下性质：

1. $X$ 是一个 $n$ 维随机向量，则 $\mathrm{tr}[\mathrm{Cov}(\mathbf{X})] = \sum_{i=1}^{n} \mathrm{Var}(\mathbf{X}_i)$；
2. $X$ 是一个 $n$ 维随机向量，则 $\mathrm{Cov}(\mathbf{X})$ 是半正定的对称矩阵；
3. 设 $A$ 和 $B$ 分别为 $p \times n$ 和 $q \times m$ 非随机矩阵，$\mathbf{X}$ 和 $\mathbf{Y}$ 分别为 $n$ 维、$m$ 维随机向量，则：

$$\mathrm{Cov}(A\mathbf{X}, B\mathbf{Y}) = A\,\mathrm{Cov}(\mathbf{X}, \mathbf{Y})B^T$$

*4.* 若 $\mathbf{X}$ 是一个常数项量，$\mathbf{Y}$ 是一个随机向量，则 $\mathrm{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$；

*5.* 设 $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ 为随机向量，则：

$$\begin{aligned}\mathrm{Cov}(\mathbf{X}+\mathbf{Y}, \mathbf{Z}) &= \mathrm{Cov}(\mathbf{X}, \mathbf{Z}) + \mathrm{Cov}(\mathbf{Y}, \mathbf{Z})\\ \mathrm{Cov}(\mathbf{X}, \mathbf{Y}+\mathbf{Z}) &= \mathrm{Cov}(\mathbf{X}, \mathbf{Y}) + \mathrm{Cov}(\mathbf{X}, \mathbf{Z})\end{aligned}$$

*6.* $\mathrm{Cov}(\mathbf{X}) = \mathrm{E}(\mathbf{X}\mathbf{X}^T) - [\mathrm{E}(\mathbf{X})][\mathrm{E}(\mathbf{X})]^T$。

*Proof.* (1) $\mathrm{Cov}(\mathbf{X})$ 在 $(i, i)$ 位置上的元素为：

$$\mathrm{E}\Big[\Big(\mathbf{X}_i - \mathrm{E}(\mathbf{X}_i)\Big)\Big(\mathbf{X}_i - \mathrm{E}(\mathbf{X}_i)\Big)^T\Big] = \mathrm{E}\Big[\Big(\mathbf{X}_i - \mathrm{E}(\mathbf{X}_i)\Big)^2\Big] = \mathrm{Var}(\mathbf{X}_i)$$

所以 $\mathrm{tr}[\mathrm{Cov}(\mathbf{X})] = \sum\limits_{i=1}^{n} \mathrm{Var}(\mathbf{X}_i)$。

(2) 因为：

$$\begin{aligned}\mathrm{Cov}(\mathbf{X})_{(i,j)} &= \mathrm{E}\Big[\Big(\mathbf{X}_i - \mathrm{E}(\mathbf{X}_i)\Big)\Big(\mathbf{X}_j - \mathrm{E}(\mathbf{X}_j)\Big)^T\Big]\\ &= \mathrm{E}\Big[\Big(\mathbf{X}_j - \mathrm{E}(\mathbf{X}_j)\Big)\Big(\mathbf{X}_i - \mathrm{E}(\mathbf{X}_i)\Big)^T\Big]\\ &= \mathrm{Cov}(\mathbf{X})_{(j,i)}\end{aligned}$$

所以 $\mathrm{Cov}(\mathbf{X})$ 是一个对称矩阵。

取 $n$ 维非随机向量 $c$，设 $Y = c^T\mathbf{X}$ 则有：

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Var}(c^T\mathbf{X})\\ &= \mathrm{E}\Big[\Big(c^T\mathbf{X} - \mathrm{E}(c^T\mathbf{X})\Big)\Big(c^T\mathbf{X} - \mathrm{E}(c^T\mathbf{X})\Big)\Big]\\ &= \mathrm{E}\Big[\Big(c^T\mathbf{X} - c^T\,\mathrm{E}(\mathbf{X})\Big)\Big(c^T\mathbf{X} - c^T\,\mathrm{E}(\mathbf{X})\Big)^T\Big]\\ &= \mathrm{E}\Big\{c^T\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)\Big[c^T\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)\Big]^T\Big\}\\ &= c^T\,\mathrm{E}\Big[\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)^T\Big]c\\ &= c^T\,\mathrm{Cov}(\mathbf{X})c \geqslant 0\end{aligned}$$

由 $c$ 的任意性，$\mathrm{Cov}(\mathbf{X})$ 是半正定的。

(3) 类似于 (2) 中的推导，有：

$$\begin{aligned}\mathrm{Cov}(A\mathbf{X}, B\mathbf{Y}) &= \mathrm{E}\Big[\Big(A\mathbf{X} - \mathrm{E}(A\mathbf{X})\Big)\Big(B\mathbf{Y} - \mathrm{E}(B\mathbf{Y})\Big)^T\Big]\\ &= A\,\mathrm{E}\Big[\Big(\mathbf{X} - \mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y} - \mathrm{E}(\mathbf{Y})\Big)^T\Big]B^T\\ &= A\,\mathrm{Cov}(\mathbf{X}, \mathbf{Y})B^T\end{aligned}$$

(4) 由协方差的定义直接可得；

(5) 由可得：

期望的线性性

$$
\begin{aligned}
\operatorname{Cov}(\mathbf{X}+\mathbf{Y},\mathbf{Z}) &= \mathrm{E}\left[\Big(\mathbf{X}+\mathbf{Y}-\mathrm{E}(\mathbf{X}+\mathbf{Y})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\left[\Big(\mathbf{X}+\mathbf{Y}-\mathrm{E}(\mathbf{X})-\mathrm{E}(\mathbf{Y})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T+\Big(\mathbf{Y}-\mathrm{E}(\mathbf{Y})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\Big[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\Big]+\mathrm{E}\Big[\Big(\mathbf{Y}-\mathrm{E}(\mathbf{Y})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\Big] \\
&= \operatorname{Cov}(\mathbf{X},\mathbf{Z})+\operatorname{Cov}(\mathbf{Y},\mathbf{Z}) \\
\operatorname{Cov}(\mathbf{X},\mathbf{Y}+\mathbf{Z}) &= \mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y}+\mathbf{Z}-\mathrm{E}(\mathbf{Y}+\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y}+\mathbf{Z}-\mathrm{E}(\mathbf{Y})-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y}-\mathrm{E}(\mathbf{Y})\Big)^T+\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Y}-\mathrm{E}(\mathbf{Y})\Big)^T\right]+\mathrm{E}\left[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{Z}-\mathrm{E}(\mathbf{Z})\Big)^T\right] \\
&= \operatorname{Cov}(\mathbf{X},\mathbf{Y})+\operatorname{Cov}(\mathbf{X},\mathbf{Z})
\end{aligned}
$$

(6) 显然：

$$
\begin{aligned}
\operatorname{Cov}(\mathbf{X}) &= \mathrm{E}\Big[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)^T\Big] = \mathrm{E}\Big[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\mathbf{X}^T-\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\,\mathrm{E}(\mathbf{X})^T\Big] \\
&= \mathrm{E}\Big[\Big(\mathbf{X}-\mathrm{E}(\mathbf{X})\Big)\mathbf{X}^T\Big]-\mathrm{E}[\mathbf{X}-\mathrm{E}(\mathbf{X})]\,\mathrm{E}(\mathbf{X})^T = \mathrm{E}(\mathbf{X}\mathbf{X}^T)-[\mathrm{E}(\mathbf{X})][\mathrm{E}(\mathbf{X})]^T \qquad \square
\end{aligned}
$$

## 2.5 二次型

**Definition 2.8.** $\mathbf{X}$ 是一个 $n$ 维随机向量，$A=(a_{ij})$ 为 $n$ 阶非随机实对称阵，则随机变量：

$$\mathbf{X}^T A\mathbf{X}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\mathbf{X}_i\mathbf{X}_j$$

称为 $\mathbf{X}$ 的二次型。

### 随机变量二次型的均值

**Theorem 2.3.** $\mathbf{X}$ 是一个 $n$ 维随机向量，$\mathrm{E}(\mathbf{X})=\mu$，$\operatorname{Cov}(\mathbf{X})=\Sigma$，则：

$$\mathrm{E}(\mathbf{X}^T A\mathbf{X})=\mu^T A\mu+\operatorname{tr}(A\Sigma)$$

*Proof.*

$$\begin{aligned}
\mathrm{E}(\mathbf{X}^T A\mathbf{X}) &= \mathrm{E}[(\mathbf{X}-\mu+\mu)^T A(\mathbf{X}-\mu+\mu)] \\
&= \mathrm{E}[(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)] + \mathrm{E}[(\mathbf{X}-\mu)^T A\mu] + \mathrm{E}[\mu^T A(\mathbf{X}-\mu)] + \mathrm{E}(\mu^T A\mu) \\
&= \mathrm{E}\{\mathrm{tr}[(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)]\} + \mu^T A\mu \\
&= \mathrm{E}\{\mathrm{tr}[A(\mathbf{X}-\mu)(\mathbf{X}-\mu)^T]\} + \mu^T A\mu \\
&= \mathrm{tr}\,\mathrm{E}[A(\mathbf{X}-\mu)(\mathbf{X}-\mu)^T] + \mu^T A\mu \\
&= \mathrm{tr}\{A\,\mathrm{E}[(\mathbf{X}-\mu)(\mathbf{X}-\mu)^T]\} + \mu^T A\mu \\
&= \mathrm{tr}(A\Sigma) + \mu^T A\mu
\end{aligned}$$

第二行到第三行利用到了 $\mathrm{E}(\mathbf{X}) = \mu$ 以及 $(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu) = \mathrm{tr}[(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)]$，后式成立是因为 $(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)$ 是一个标量，标量的迹自然等于自身。第三行到第四行使用到了**??**(3)。 □

### 独立随机变量二次型的方差

**Theorem 2.4.** 设随机变量 $X_i$，$i=1,2,\ldots,n$ 相互独立，$\mathrm{E}(X_i)=\mu_i$，$\mathrm{Var}(X_i)=\sigma^2$，$\nu_k^{(i)} = \mathrm{E}[(X_i-\mu_i)^k]$，$\mathbf{X}=(X_1,X_2,\ldots,X_n)^T$，$\mu=(\mu_1,\mu_2,\ldots,\mu_n)^T$，$A=(a_{ij})$ 为 $n$ 阶非随机实对称阵，$a=(a_{11},a_{22},\ldots,a_{nn})^T$，$b=(\nu_3^{(1)}a_{11},\nu_3^{(2)}a_{22},\ldots,\nu_3^{(n)}a_{nn})^T$，则：

$$\mathrm{Var}(\mathbf{X}^T A\mathbf{X}) = \sum_{i=1}^n a_{ii}^2\nu_4^{(i)} + \sigma^4[2\,\mathrm{tr}(A^2) - 3a^T a] + 4\sigma^2\mu^T A^2\mu + 4\mu^T Ab$$

*Proof.* 由性质 2.2.1(1) 可得：

$$\mathrm{Var}(\mathbf{X}^T A\mathbf{X}) = \mathrm{E}[(\mathbf{X}^T A\mathbf{X})^2] - [\mathrm{E}(\mathbf{X}^T A\mathbf{X})]^2$$

由题设可知：

$$\mathrm{E}(\mathbf{X}) = \mu,\ \mathrm{Var}(\mathbf{X}) = \sigma^2 I$$

根据定理 2.3可得：

$$\begin{aligned}
[\mathrm{E}(\mathbf{X}^T A\mathbf{X})]^2 &= [\mathrm{tr}(A\sigma^2 I) + \mu^T A\mu]^2 = [\sigma^2\,\mathrm{tr}(A) + \mu^T A\mu]^2 \\
&= \sigma^4[\mathrm{tr}(A)]^2 + 2\sigma^2\,\mathrm{tr}(A)\mu^T A\mu + (\mu^T A\mu)^2
\end{aligned}$$

同时：

$$\begin{aligned}
(\mathbf{X}^T A\mathbf{X})^2 &= [(\mathbf{X}-\mu+\mu)^T A(\mathbf{X}-\mu+\mu)]^2 \\
&= [(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu) + 2\mu^T A(\mathbf{X}-\mu) + \mu^T A\mu]^2 \\
&= [(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)]^2 + 4[\mu^T A(\mathbf{X}-\mu)]^2 + (\mu^T A\mu)^2 \\
&\quad + 4(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)\mu^T A(\mathbf{X}-\mu) + 2(\mathbf{X}-\mu)^T A(\mathbf{X}-\mu)\mu^T A\mu \\
&\quad + 4\mu^T A(\mathbf{X}-\mu)\mu^T A\mu
\end{aligned}$$

令 $\mathbf{Y}=\mathbf{X}-\mu$，则有 $\mathrm{E}(\mathbf{Y})=\mathbf{0}$，再由定理 2.3可得：

$$\begin{aligned}\mathrm{E}[(\mathbf{X}^TA\mathbf{X})^2]&=\mathrm{E}[(\mathbf{Y}^TA\mathbf{Y})^2]+4\,\mathrm{E}[(\mu^TA\mathbf{Y})^2]+(\mu^TA\mu)^2\\&\quad+4\,\mathrm{E}(\mathbf{Y}^TA\mathbf{Y}\mu^TA\mathbf{Y})+2\mu^TA\mu\sigma^2\,\mathrm{tr}(A)\end{aligned}$$

考虑：

$$\begin{aligned}\mathrm{E}[(\mathbf{Y}^TA\mathbf{Y})^2]&=\mathrm{E}\left(\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ij}a_{kl}\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l\right)\\&=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ij}a_{kl}\,\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)\end{aligned}$$

作分类讨论：

1. $i,j,k,l$ 互不相同，则 $\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)=E(\mathbf{Y}_i)E(\mathbf{Y}_j)E(\mathbf{Y}_k)E(\mathbf{Y}_l)=0$；
2. $i,j,k,l$ 中存在某两个值相同：
   - 此时另外两个不同，则 $\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)=0$；
   - 此时另外两个也相同（即 $i=j,k=l$ 或 $i=k,j=l$ 或 $i=l,j=k$），则 $\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)=\sigma^4$。
3. $i,j,k,l$ 中存在某三个值相同，则 $\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)=0$；
4. $i,j,k,l$ 相同，则 $\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)=\nu_4^{(i)}$。

于是：

$$\begin{aligned}\mathrm{E}[(\mathbf{Y}^TA\mathbf{Y})^2]&=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ij}a_{kl}\,\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_k\mathbf{Y}_l)\\&=\sum_{i=1}^n a_{ii}^2\nu_4^{(i)}+\sigma^4\left(\sum_{i\neq k}a_{ii}a_{kk}+\sum_{i\neq j}a_{ij}^2+\sum_{i\neq j}a_{ij}a_{ji}\right)\\&=\sum_{i=1}^n a_{ii}^2\nu_4^{(i)}+\sigma^4\left(\sum_{i\neq k}a_{ii}a_{kk}+2\sum_{i\neq j}a_{ij}^2\right)\end{aligned}$$

因为：

$$\sum_{i\neq k}a_{ii}a_{kk}=[\mathrm{tr}(A)]^2-a^Ta$$

$$\sum_{i\neq j}a_{ij}^2=\mathrm{tr}(AA^T)-a^Ta=\mathrm{tr}(A^2)-a^Ta$$

所以：

$$\mathrm{E}[(\mathbf{Y}^TA\mathbf{Y})^2]=\sum_{i=1}^n a_{ii}^2\nu_4^{(i)}+\sigma^4\{[\mathrm{tr}(A)]^2+2\,\mathrm{tr}(A^2)-3a^Ta\}$$

由定理 2.3和??(3) 可得：

$$\begin{aligned}\mathrm{E}[(\mu^T A\mathbf{Y})^2] &= \mathrm{E}(\mu^T A\mathbf{Y}\mu^T A\mathbf{Y}) = \mathrm{E}(\mathbf{Y}^T A\mu\mu^T A\mathbf{Y}) = \operatorname{tr}(A\mu\mu^T A\sigma^2 I)\\ &= \sigma^2 \operatorname{tr}(A\mu\mu^T A) = \sigma^2 \operatorname{tr}(\mu^T A^2\mu) = \sigma^2\mu^T A^2\mu\end{aligned}$$

注意到：

$$\begin{aligned}\mathrm{E}(\mathbf{Y}^T A\mathbf{Y}\mu^T A\mathbf{Y}) &= \mathrm{E}\left(\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{Y}_i\mathbf{Y}_j\sum_{k=1}^n\sum_{l=1}^n a_{kl}\mu_k\mathbf{Y}_l\right)\\ &= \mathrm{E}\left(\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ij}a_{kl}\mu_k\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_l\right)\\ &= \sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ij}a_{kl}\mu_k\,\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_l)\end{aligned}$$

和之前的讨论类似，可以得到：

$$\mathrm{E}(\mathbf{Y}_i\mathbf{Y}_j\mathbf{Y}_l) = \begin{cases}\nu_3^{(i)}, & i=j=l\\ 0, & \text{其他情况}\end{cases}$$

于是有：

$$\mathrm{E}(\mathbf{Y}^T A\mathbf{Y}\mu^T A\mathbf{Y}) = \sum_{i=1}^n\sum_{k=1}^n a_{ii}\nu_3^{(i)}a_{ki}\mu_k$$

令 $b = (\nu_3^{(1)}a_{11}, \nu_3^{(2)}a_{22}, \ldots, \nu_3^{(n)}a_{nn})^T$，则：

$$\mathrm{E}(\mathbf{Y}^T A\mathbf{Y}\mu^T A\mathbf{Y}) = \sum_{i=1}^n\sum_{k=1}^n a_{ii}\nu_3^{(i)}a_{ki}\mu_k = \mu^T Ab$$

将以上求得的期望值全部代入，即可得到：

$$\begin{aligned}\mathrm{E}[(\mathbf{X}^T A\mathbf{X})^2] &= \mathrm{E}[(\mathbf{Y}^T A\mathbf{Y})^2] + 4\,\mathrm{E}[(\mu^T A\mathbf{Y})^2] + (\mu^T A\mu)^2\\ &\quad + 4\,\mathrm{E}(\mathbf{Y}^T A\mathbf{Y}\mu^T A\mathbf{Y}) + 2\mu^T A\mu\sigma^2\operatorname{tr}(A)\\ &= \sum_{i=1}^n a_{ii}^2\nu_4^{(i)} + \sigma^4\{[\operatorname{tr}(A)]^2 + 2\operatorname{tr}(A^2) - 3a^T a\}\\ &\quad + 4\sigma^2\mu^T A^2\mu + (\mu^T A\mu)^2 + 4\mu^T Ab + 2\mu^T A\mu\sigma^2\operatorname{tr}(A)\end{aligned}$$

于是：

$$\begin{aligned}\operatorname{Var}(\mathbf{X}^T A\mathbf{X}) &= \mathrm{E}[(\mathbf{X}^T A\mathbf{X})^2] - [\mathrm{E}(\mathbf{X}^T A\mathbf{X})]^2\\ &= \sum_{i=1}^n a_{ii}^2\nu_4^{(i)} + \sigma^4\{[\operatorname{tr}(A)]^2 + 2\operatorname{tr}(A^2) - 3a^T a\}\\ &\quad + 4\sigma^2\mu^T A^2\mu + (\mu^T A\mu)^2 + 4\mu^T Ab + 2\mu^T A\mu\sigma^2\operatorname{tr}(A)\\ &\quad - \sigma^4[\operatorname{tr}(A)]^2 - 2\sigma^2\operatorname{tr}(A)\mu^T A\mu - (\mu^T A\mu)^2\\ &= \sum_{i=1}^n a_{ii}^2\nu_4^{(i)} + \sigma^4[2\operatorname{tr}(A^2) - 3a^T a] + 4\sigma^2\mu^T A^2\mu + 4\mu^T Ab\end{aligned}$$

□

## 2.6 矩母函数

**Definition 2.9.** 设 $X$ 是一个随机变量。称:

$$M_X(t) = \mathrm{E}(e^{tX})$$

为 $X$ 的矩母函数 *(moment-generating function, m.g.f.)*，其中 $t \in \mathbb{R}$。

**Definition 2.10.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量。称:

$$M_{\mathbf{X}}(t) = \mathrm{E}(e^{t^T\mathbf{X}})$$

为 $\mathbf{X}$ 的矩母函数，其中 $t \in \mathbb{R}^n$。

**Property 2.6.1.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量，则其矩母函数 $M_{\mathbf{X}}(t)$ 具有如下性质：

1. $M_{\mathbf{X}}(\mathbf{0}) = 1$；
2. $M_{\mathbf{X}}(t) \geqslant e^{t^T\mu}$，其中 $\mu$ 是 $\mathbf{X}$ 的均值向量；
3. 矩母函数与概率分布之间存在一个双射，即 $M_{\mathbf{X}}(t) = M_{\mathbf{Y}}(t)$ 当且仅当 $\mathbf{X}$ 与 $\mathbf{Y}$ 具有相同的概率分布；
4. 设 $m$ 维随机向量 $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ 彼此独立，$\alpha_i$ 为常数，$\beta_i$ 为 $m$ 维常数向量，则 $\mathbf{Y} = \sum\limits_{i=1}^{n}(\alpha_i\mathbf{X}_i + \beta_i)$ 的特征函数为：
$$M_{\mathbf{Y}}(t) = \prod_{i=1}^{n} e^{t^T\beta_i} M_{\mathbf{X}_i}(\alpha_i t)$$
5. $M_X^{(n)}(0) = \mu_n$，其中 $X$ 是一个随机变量，$\mu_n$ 是 $X$ 的 $n$ 阶原点矩；
6. $M_{\mathbf{X}}(t)$ 有如下幂级数展开：
$$M_{\mathbf{X}}(t) = \sum_{(m_1,m_2,\ldots,m_n)\in\mathbb{N}^n} \mu_{m_1,m_2,\ldots,m_n} \prod_{i=1}^{n} \frac{t_i^{m_i}}{m_i!}$$

*Proof.* (1)$M_{\mathbf{X}}(\mathbf{0}) = \mathrm{E}(e^0) = 1$。

Jensen 不等式链接

(2) 由 Jensen 不等式直接可得。

(3)

(4) 由矩母函数定义可得：

$$M_{\mathbf{Y}}(t) = \mathrm{E}(e^{t^T\mathbf{Y}}) = \mathrm{E}\left(\exp\left\{t^T\sum_{i=1}^{n}(\alpha_i\mathbf{X}_i + \beta_i)\right\}\right) = \mathrm{E}\left(\prod_{i=1}^{n} e^{\alpha_i t^T\mathbf{X}_i}\right)\prod_{i=1}^{n} e^{t^T\beta_i}$$

因为 $\mathbf{X}_i$ 互相独立，所以 $\alpha_i\mathbf{X}_i$ 也相互独立，于是有：

$$M_{\mathbf{Y}}(t) = \mathrm{E}\left(\prod_{i=1}^{n} e^{\alpha_i t^T\mathbf{X}_i}\right)\prod_{i=1}^{n} e^{t^T\beta_i} = \prod_{i=1}^{n}\mathrm{E}\left(e^{\alpha_i t^T\mathbf{X}_i}\right)\prod_{i=1}^{n} e^{t^T\beta_i} = \prod_{i=1}^{n} e^{t^T\beta_i} M_{\mathbf{X}_i}(\alpha_i t)$$

(5) 将 $e^{tX}$ 展开为幂级数：

$$M_X(t)=\mathrm{E}(e^{tX})=\mathrm{E}\left(\sum_{n=0}^{+\infty}\frac{t^nX^n}{n!}\right)$$

于是：

$$M_X^{(n)}(t)=\mathrm{E}\left(X^n+\sum_{m=n+1}^{+\infty}\frac{t^mX^m}{m!}\right)=\mu_n+\sum_{m=1}^{+\infty}\frac{t^m}{m!}\mu_m$$

所以：

$$M_X^{(n)}(0)=\mathrm{E}(X^n)=\mu_n$$

(6) 由可得：

期望的线性性质，Lebesgue 积分

$$\begin{aligned}
M_{\mathbf{X}}(t)&=\mathrm{E}(e^{t^T\mathbf{X}})=\mathrm{E}\left(\exp\left\{\sum_{i=1}^n t_i\mathbf{X}_i\right\}\right)=\mathrm{E}\left[\sum_{m=0}^{+\infty}\frac{1}{m!}\left(\sum_{i=1}^n t_i\mathbf{X}_i\right)^m\right]\\
&=\sum_{m=0}^{+\infty}\frac{1}{m!}\mathrm{E}\left[\left(\sum_{i=1}^n t_i\mathbf{X}_i\right)^m\right]=\sum_{m=0}^{+\infty}\frac{1}{m!}\mathrm{E}\left(\sum_{\sum_{i=1}^n m_i=m}\frac{m!}{m_1!m_2!\cdots m_n!}\prod_{i=1}^n(t_i\mathbf{X}_i)^{m_i}\right)\\
&=\sum_{m=0}^{+\infty}\frac{1}{m!}\sum_{\sum_{i=1}^n m_i=m}\frac{m!}{m_1!m_2!\cdots m_n!}\mathrm{E}\left[\prod_{i=1}^n(t_i\mathbf{X}_i)^{m_i}\right]\\
&=\sum_{m=0}^{+\infty}\sum_{\sum_{i=1}^n m_i=m}\frac{1}{m_1!m_2!\cdots m_n!}\mathrm{E}\left(\prod_{i=1}^n\mathbf{X}_i^{m_i}\right)\prod_{i=1}^n t_i^{m_i}\\
&=\sum_{(m_1,m_2,\dots,m_n)\in\mathbb{N}^n}\mu_{m_1,m_2,\dots,m_n}\prod_{i=1}^n\frac{t_i^{m_i}}{m_i!}
\end{aligned}$$

□

## 2.7 累积量生成函数

**Definition 2.11.** 设 $X$ 是一个随机变量。称 $K_X(t)=\log M_X(t)$ 为 $X$ 的累积量生成函数 *(cumulant-generating function, c.g.f.)*，其中 $t\in\mathbb{R}$。

**Definition 2.12.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量。称 $K_{\mathbf{X}}(t)=\log M_{\mathbf{X}}(t)$ 为 $\mathbf{X}$ 的累积量生成函数，其中 $t\in\mathbb{R}^n$。

**Definition 2.13.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量。因为：

$$\begin{aligned}
M_X(t)&=\sum_{(m_1,m_2,\dots,m_n)\in\mathbb{N}^n}\frac{1}{m_1!m_2!\cdots m_n!}\prod_{i=1}^n t_i^{m_i}\mu_{m_1,m_2,\dots,m_n}\\
&=1+\sum_{\substack{(m_1,m_2,\dots,m_n)\in\mathbb{N}^n\\(m_1,m_2,\dots,m_n\neq\mathbf{0})}}\frac{1}{m_1!m_2!\cdots m_n!}\prod_{i=1}^n t_i^{m_i}\mu_{m_1,m_2,\dots,m_n}
\end{aligned}$$

由对数函数的幂级数展开可得：

$$
\begin{aligned}
K_{\mathbf{X}}(t) &= \log\left(1+\sum_{\substack{(m_1,m_2,\ldots,m_n)\in\mathbb{N}^n\\(m_1,m_2,\ldots,m_n\neq\mathbf{0})}}\frac{1}{m_1!m_2!\cdots m_n!}\prod_{i=1}^{n}t_i^{m_i}\mu_{m_1,m_2,\ldots,m_n}\right)\\
&=\sum_{j=1}^{+\infty}(-1)^{j+1}\frac{1}{j}\left(\sum_{\substack{(m_1,m_2,\ldots,m_n)\in\mathbb{N}^n\\(m_1,m_2,\ldots,m_n\neq\mathbf{0})}}\frac{1}{m_1!m_2!\cdots m_n!}\prod_{i=1}^{n}t_i^{m_i}\mu_{m_1,m_2,\ldots,m_n}\right)^j
\end{aligned}
$$

**Property 2.7.1.**

## 2.8 特征函数

**Definition 2.14.** 设 $X$ 是一个随机变量。称:

$$\varphi_X(t)=\mathrm{E}(e^{itX})$$

为 $X$ 的特征函数 *(characteristic function, c.f.)*，其中 $t\in\mathbb{R}$。

**Definition 2.15.** 设 $\mathbf{X}$ 是一个 $n$ 维随机向量。称:

$$\varphi_{\mathbf{X}}(t)=\mathrm{E}(e^{it^T\mathbf{X}})$$

为 $\mathbf{X}$ 的特征函数，其中 $t\in\mathbb{R}^n$。

**Definition 2.16.** 设 $\mathbf{X}$ 是一个 $m\times n$ 随机矩阵。称:

$$\varphi_{\mathbf{X}}(t)=\mathrm{E}\Big[\exp\Big(i\,\mathrm{tr}(t^T\mathbf{X})\Big)\Big]$$

为 $\mathbf{X}$ 的特征函数，其中 $t\in M_{m\times n}(\mathbb{R})$。

**Property 2.8.1.** 设 $X,Y,X_1,X_2,\ldots,X_n$ 是随机变量，$\alpha_1,\alpha_2,\ldots,\alpha_n$，$\beta_1,\beta_2,\ldots,\beta_n$ 为常数，则：

1. $X$ 的特征函数 $\varphi_X(t)$ 存在；
2. $|\varphi_X(t)|\leqslant\varphi_X(0)=1$；
3. $\varphi_X(-t)=\overline{\varphi_X(t)}$；
4. 若 $X_1,X_2,\ldots,X_n$ 相互独立，则 $Y=\sum\limits_{k=1}^{n}(\alpha_kX_k+\beta_k)$ 的特征函数为：

$$\varphi_Y(t)=\prod_{k=1}^{n}e^{it\beta_k}\varphi_{X_k}(\alpha_kt)$$

5. $X_1, X_2, \dots, X_n$ 相互独立的充分必要条件为：

$$\varphi_{X_1,\dots,X_n}(t_1, t_2, \dots, t_n) = \prod_{i=1}^{n} \varphi_{X_i}(t_i)$$

6. 特征函数与概率分布之间存在一个双射，即 $\varphi_X(t) = \varphi_Y(t)$ 当且仅当 $X$ 与 $Y$ 具有相同的概率分布。

7. 若 $\mathrm{E}(X^n)$ 存在，则 $\varphi_X^{(n)}(t)$ 存在，且对 $1 \leqslant k \leqslant n$ 有：

$$\mathrm{E}(X^k) = i^{-k}\varphi_X^{(k)}(0)$$

特别的：

$$\mathrm{E}(X) = -i\varphi_X'(0),\ \mathrm{Var}(X) = -\varphi_X''(0) + [\varphi_X'(0)]^2$$

8. 若 $\varphi_X(t)$ 在 $t = 0$ 处最高有 $n$ 阶导数，如果 $n$ 为奇数，则 $X$ 具有所有不超过 $n-1$ 阶的原点矩；若 $n$ 为偶数，则 $X$ 具有所有不超过 $n$ 阶的原点矩；

9. $\varphi_X(t)$ 在 $\mathbb{R}$ 上一致连续；

10. $\varphi_X(t)$ 是半正定的，即对任意的 $n \in \mathbb{N}^+$ 及任意的 $t = (t_1, t_2, \dots, t_n)^T \in \mathbb{R}^n$ 和任意的 $c = (c_1, c_2, \dots, c_n)^T \in \mathbb{C}^n$，令 $A = [\varphi_X(t_i - t_j)] \in M_n(\mathbb{C})$，则有：

$$c^T A\overline{c} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i\overline{c_j}\varphi_X(t_i - t_j) \geqslant 0$$

*Proof.* (1) 因为：

$$e^{itX} = \cos(tX) + i\sin(tX)$$

所以 $|e^{itX}| = 1$，于是：

链接 Lebesgue 积分性质

$$\left|\mathrm{E}(e^{itX})\right| = \left|\int_{-\infty}^{+\infty} e^{itx}p(x)\,\mathrm{d}x\right| \leqslant \int_{-\infty}^{+\infty} |e^{itx}|p(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1$$

所以 $\varphi_X(t)$ 存在。

(2) 可以发现：

$$\varphi_X(0) = \int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1$$

再由 (1) 的证明过程即可得出结论。

(3) 因为：

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \mathrm{E}[\cos(tX) + i\sin(tX)] = \mathrm{E}[\cos(tX)] + i\,\mathrm{E}[\sin(tX)]$$

所以：

$$\overline{\varphi_X(t)} = \mathrm{E}[\cos(tX)] - i\,\mathrm{E}[\sin(tX)] = \mathrm{E}[\cos(-tX)] + i\,\mathrm{E}[\sin(-tX)] = \varphi_X(-t)$$

(4) 因为 $X_k$ 相互独立，所以 $e^{it(\alpha_k X_k+\beta_k)}$ 之间也相互独立，$k=1,2,\ldots,n$，于是有：

$$\begin{aligned}\varphi_Y(t) &= \mathrm{E}\left[\exp\left(it\sum_{k=1}^{n}(\alpha_k X_k+\beta_k)\right)\right] = \mathrm{E}\left(\prod_{k=1}^{n} e^{it(\alpha_k X_k+\beta_k)}\right)\\ &= \prod_{k=1}^{n}\mathrm{E}[e^{it(\alpha_k X_k+\beta_k)}] = \prod_{k=1}^{n} e^{it\beta_k}\,\mathrm{E}(e^{it\alpha_k X_k}) = \prod_{k=1}^{n} e^{it\beta_k}\varphi_{X_k}(\alpha_k t)\end{aligned}$$

(5) **必要性**：因为 $X_k$ 相互独立，所以 $e^{it_k X_k}$ 相互独立，$k=1,2,\ldots,n$。由随机向量特征函数的定义可得：

$$\begin{aligned}\varphi_{X_1,\ldots,X_n}(t_1,t_2,\ldots,t_n) &= \mathrm{E}\left[\exp\left(i\sum_{k=1}^{n}t_k X_k\right)\right] = \mathrm{E}\left(\prod_{k=1}^{n} e^{it_k X_k}\right)\\ &= \prod_{k=1}^{n}\mathrm{E}(e^{it_k X_k}) = \prod_{k=1}^{n}\varphi_{X_k}(t_k)\end{aligned}$$

**充分性**：因为：

$$\begin{aligned}\varphi_{X_1,\ldots,X_n}(t_1,t_2,\ldots,t_n) &= \mathrm{E}\left[\exp\left(i\sum_{k=1}^{n}t_k X_k\right)\right]\\ &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left(i\sum_{k=1}^{n}t_k x_k\right)p(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n\end{aligned}$$

$$\begin{aligned}\prod_{i=1}^{n}\varphi_{X_i}(t_i) &= \prod_{k=1}^{n}\mathrm{E}(e^{it_k X_k})\\ &= \prod_{k=1}^{n}\int_{-\infty}^{+\infty} e^{it_k x_k}p(x_k)\,\mathrm{d}x_k\\ &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left(i\sum_{k=1}^{n}t_k x_k\right)p(x_1)p(x_2)\cdots p(x_n)\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n\end{aligned}$$

若两式相等，则有：

$$p(x_1,x_2,\ldots,x_n) = p(x_1)p(x_2)\cdots p(x_n)$$

链接独立性条件

由可得 $X_k$，$k=1,2,\ldots,n$ 相互独立。

(6)

(7) 因为 $\mathrm{E}(X^n)$ 存在，所以：

$$\int_{-\infty}^{+\infty}|x|^n p(x)\,\mathrm{d}x < +\infty$$

于是：

$$\left|\int_{-\infty}^{+\infty} i^n x^n e^{itx}p(x)\,\mathrm{d}x\right| \leqslant \int_{-\infty}^{+\infty}|x|^n p(x)\,\mathrm{d}x < +\infty$$

所以：

$$\varphi_X^{(n)}(t) = \int_{-\infty}^{+\infty} i^n x^n e^{itx}p(x)\,\mathrm{d}x$$

存在。由定理 2.1可知对 $1 \leqslant k \leqslant n$ 有 $\mathrm{E}(X^k)$ 存在，于是：

$$\varphi_X^{(k)}(0) = \int_{-\infty}^{+\infty} i^k x^k p(x)\,\mathrm{d}x = i^k \int_{-\infty}^{+\infty} x^k p(x)\,\mathrm{d}x = i^k\,\mathrm{E}(X^k)$$

也存在。

(8) 注意到：

$$\varphi_X^{(n)}(t) = \int_{-\infty}^{+\infty} i^n x^n e^{itx} p(x)\,\mathrm{d}x$$

因为 $\varphi_X(t)$ 在 $t=0$ 处最高具有 $n$ 阶导数，于是：

$$|\varphi_X^{(n)}(0)| = \left|\int_{-\infty}^{+\infty} i^n x^n p(x)\,\mathrm{d}x\right| = \left|\int_{-\infty}^{+\infty} x^n p(x)\,\mathrm{d}x\right| < +\infty$$

当 $n = 2k+1,\ k \in \mathbb{N}$ 时，有：

$$\int_{-\infty}^{+\infty} |x|^n p(x)\,\mathrm{d}x > |\varphi_X^{(n)}(0)| = \left|\int_{-\infty}^{+\infty} x^n p(x)\,\mathrm{d}x\right|$$

所以 $\mathrm{E}(X^n)$ 不一定存在。当 $n = 2k,\ k \in \mathbb{N}^+$ 时，有：

需要证明对小于的都存在

$$|\varphi_X^{(n)}(0)| = \left|\int_{-\infty}^{+\infty} x^n p(x)\,\mathrm{d}x\right| = \int_{-\infty}^{+\infty} |x|^n p(x)\,\mathrm{d}x < +\infty$$

存在，于是 $\mathrm{E}(X^n)$ 存在。由定理 2.1可知，此时 $X$ 具有所有不超过 $n$ 阶的原点矩。

(9) 对任意的 $t, h \in \mathbb{R}$ 和 $a > 0$，有：

$$\begin{aligned}
|\varphi(t+h) - \varphi(t)| &= \left|\int_{-\infty}^{+\infty} [e^{i(t+h)x} - e^{itx}] p(x)\,\mathrm{d}x\right| \\
&= \left|\int_{-\infty}^{+\infty} (e^{ihx} - 1) e^{itx} p(x)\,\mathrm{d}x\right| \\
&\leqslant \int_{-\infty}^{+\infty} |(e^{ihx} - 1) e^{itx}| p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} |e^{ihx} - 1||e^{itx}| p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} |e^{ihx} - 1| p(x)\,\mathrm{d}x \\
&= \int_{-a}^{a} |e^{ihx} - 1| p(x)\,\mathrm{d}x + \int_{|x|\geqslant a} |e^{ihx} - 1| p(x)\,\mathrm{d}x \\
&\leqslant \int_{-a}^{a} |e^{ihx} - 1| p(x)\,\mathrm{d}x + \int_{|x|\geqslant a} (|e^{ihx}| + 1) p(x)\,\mathrm{d}x \\
&= \int_{-a}^{a} |e^{ihx} - 1| p(x)\,\mathrm{d}x + 2\int_{|x|\geqslant a} p(x)\,\mathrm{d}x
\end{aligned}$$

对于任意的 $\varepsilon > 0$，可以先选定一个充分大的 $a$，使得：

$$2\int_{|x|\geqslant a} p(x)\,\mathrm{d}x < \frac{\varepsilon}{2}$$

对任意的 $x \in [-a, a]$，只要取 $\delta = \dfrac{\varepsilon}{2a}$，则当 $|h| < \delta$ 时，就有：

$$\begin{aligned}
|e^{ihx} - 1| &= \left|e^{ihx} - e^{i\frac{hx}{2}} e^{i\frac{-hx}{2}}\right| = \left|e^{i\frac{hx}{2}}(e^{i\frac{hx}{2}} - e^{i\frac{-hx}{2}})\right| \\
&= \left|e^{i\frac{hx}{2}}\right| \left|e^{i\frac{hx}{2}} - e^{i\frac{-hx}{2}}\right| \\
&= \left|e^{i\frac{hx}{2}} - e^{i\frac{-hx}{2}}\right| \\
&= \left|\cos\frac{hx}{2} + i\sin\frac{hx}{2} - \cos\frac{-hx}{2} - i\sin\frac{-hx}{2}\right| \\
&= \left|2i\sin\frac{hx}{2}\right| = 2\left|\sin\frac{hx}{2}\right| \leqslant 2\left|\frac{hx}{2}\right| \leqslant ha < \frac{\varepsilon}{2}
\end{aligned}$$

于是对任意的 $t \in \mathbb{R}$，有：

$$|\varphi(t+h) - \varphi(t)| < \int_{-a}^{a} \frac{\varepsilon}{2} p(x)\,\mathrm{d}x + 2\int_{|x|\geqslant a} p(x)\,\mathrm{d}x < \frac{\varepsilon}{2}\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x + \frac{\varepsilon}{2} = \varepsilon$$

即 $\varphi_X(t)$ 在 $\mathbb{R}$ 上一致连续。

(10) 显然：

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} c_i\bar{c}_j\varphi_X(t_i - t_j) &= \sum_{k=1}^{n}\sum_{j=1}^{n} c_k\bar{c}_j \int_{-\infty}^{+\infty} e^{i(t_k - t_j)x} p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \sum_{k=1}^{n}\sum_{j=1}^{n} c_k\bar{c}_j e^{i(t_k - t_j)x} p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \left(\sum_{k=1}^{n} c_k e^{it_k x}\right)\left(\sum_{j=1}^{n} \bar{c}_j e^{-it_j x}\right) p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \left(\sum_{k=1}^{n} c_k e^{it_k x}\right)\left(\sum_{j=1}^{n} \overline{c_k e^{it_k x}}\right) p(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \left|\sum_{k=1}^{n} c_k e^{it_k x}\right|^2 p(x)\,\mathrm{d}x
\end{aligned}$$

□

## 2.9 Fisher 信息量

**Definition 2.17.** 设 $\mathbf{X}$ 是测度空间 $(X, \mathcal{F}, \mu)$ 上的一个随机向量，其分布由 $n$ 维参数 $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ 决定，$\mathbf{X}$ 的概率函数为 $f(\mathbf{X}; \theta)$。若 $f(\mathbf{X}; \theta)$ 满足如下正则条件：

*1.* $f(\mathbf{X}; \theta)$ 关于 $\theta$ 的偏导数 *a.e.* 存在；

*2.* 对 $f(\mathbf{X}; \theta)$ 在 $X$ 上的积分关于 $\theta$ 任一分量求导时都可以交换求导与积分的顺序；

*3.* $f(\mathbf{X}; \theta)$ 的定义域与 $\theta$ 无关。

则称：

$$[I(\theta)]_{ij} = \mathrm{E}\left[\left(\frac{\partial \ln f(\mathbf{X}; \theta)}{\partial \theta_i}\right)\left(\frac{\partial \ln f(\mathbf{X}; \theta)}{\partial \theta_j}\right)\right]$$

为 $\mathbf{X}$ 的*Fisher 信息矩阵 (Fisher information matrix, FIM)*。

**note 2.1.** *Fisher* 信息量（即一维情况的信息矩阵）用来表明随机变量 $\mathbf{X}$ 携带的关于参数 $\theta$ 的信息。如果它比较大，表示平均下来 $\theta$ 的微小变化会给 $\mathbf{X}$ 的分布带来较大的变化，即 $\mathbf{X}$ 的分布很依赖 $\theta$ 的具体取值，所以携带了较多关于 $\theta$ 的信息。

**Property 2.9.1.** 设 $\mathbf{X}$ 是测度空间 $(X, \mathcal{F}, \mu)$ 上的一个随机向量，其 *Fisher* 信息矩阵具有如下性质：

1. *Fisher* 信息矩阵可以看作协方差矩阵：

$$I(\theta) = \operatorname{Cov}\left[\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta}\right]$$

2. 若 $\ln f(\mathbf{X};\theta)$ 有关于 $\theta$ 的所有二阶导数，且对该二阶导数在 $X$ 上的积分关于 $\theta$ 任一分量求导时都可以交换求导与积分的顺序，则：

$$[I(\theta)]_{ij} = -\operatorname{E}\left[\frac{\partial^2 \ln f(\mathbf{X};\theta)}{\partial \theta_i \partial \theta_j}\right]$$

*Proof.* (1) 由正则条件可得：

$$\begin{aligned}\operatorname{E}\left[\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i}\right] &= \int_X \frac{\frac{\partial f(\mathbf{X};\theta)}{\partial \theta_i}}{f(\mathbf{X};\theta)} f(\mathbf{X};\theta)\,\mathrm{d}\mu = \int_X \frac{\partial f(\mathbf{X};\theta)}{\partial \theta_i}\,\mathrm{d}\mu \\ &= \frac{\partial}{\partial \theta_i}\int_X f(\mathbf{X};\theta)\,\mathrm{d}\mu = \frac{\partial 1}{\partial \theta_i} = 0\end{aligned}$$

(2) 由 (1) 和正则条件可得：

$$\begin{gathered}\frac{\partial}{\partial \theta_j}\operatorname{E}\left[\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i}\right] = 0 \\ \frac{\partial}{\partial \theta_j}\int_X \frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i} f(\mathbf{X};\theta)\,\mathrm{d}\mu = 0 \\ \int_X \frac{\partial}{\partial \theta_j}\left[\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i} f(\mathbf{X};\theta)\right]\mathrm{d}\mu = 0 \\ \int_X \left[\frac{\partial^2 \ln f(\mathbf{X};\theta)}{\partial \theta_i \partial \theta_j} f(\mathbf{X};\theta) + \frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i}\frac{\partial f(\mathbf{X};\theta)}{\partial \theta_j}\right]\mathrm{d}\mu = 0 \\ \int_X \left[\frac{\partial^2 \ln f(\mathbf{X};\theta)}{\partial \theta_i \partial \theta_j} f(\mathbf{X};\theta) + \frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i}\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_j} f(\mathbf{X};\theta)\right]\mathrm{d}\mu = 0 \\ \operatorname{E}\left[\left(\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i}\right)\left(\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_j}\right)\right] = -\operatorname{E}\left[\frac{\partial^2 \ln f(\mathbf{X};\theta)}{\partial \theta_i \partial \theta_j}\right]\end{gathered}$$

其中倒数第三行到倒数第二行是因为：

$$\frac{\partial \ln f(\mathbf{X};\theta)}{\partial \theta_i} f(\mathbf{X};\theta) = \frac{\frac{\partial f(\mathbf{X};\theta)}{\partial \theta_i}}{f(\mathbf{X};\theta)} f(\mathbf{X};\theta) = \frac{\partial f(\mathbf{X};\theta)}{\partial \theta_i}$$

□

# Chapter 3

# 时间序列分析

## 3.1 平稳时间序列

**Definition 3.1.** 如果时间序列 $\{X_t\} = \{X_t : t \in \mathbb{N}\}$ 满足：

1. 对任何的 $t \in \mathbb{N}$，有 $\mathrm{E}(X_t^2) < +\infty$；
2. 对任何的 $t \in \mathbb{N}$，有 $\mathrm{E}(X_t) = \mu$；
3. 对任何的 $t, s \in \mathbb{N}$，有 $\mathrm{Cov}(X_t, X_s) = \mathrm{E}[(X_t - \mu)(X_s - \mu)] = \gamma(t - s)$，其中 $\gamma(t - s)$ 是 $t - s$ 的实值函数，被称为 $\{X_t\}$ 的自协方差函数 *(auto-covariance function)*。

则称 $\{X_t\}$ 是一个平稳时间序列 *(stationary time series)*。

### 3.1.1 平稳时间序列的性质

#### 自协方差函数

**Property 3.1.1.** 平稳时间序列的自协方差函数具有如下基本性质：

1. 对称性：$\forall\, n \in \mathbb{N},\ \gamma(n) = \gamma(-n)$；
2. 半正定性：对任意的 $n \in \mathbb{N}^+$，$n$ 阶自协方差矩阵：
$$\Gamma_n = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{pmatrix}$$
是半正定矩阵；
3. 有界性：$|\gamma(n)| \leqslant |\gamma(0)|$ 对所有的 $n \in \mathbb{N}$ 成立。

*Proof.* (1) 由协方差的定义：

$$\gamma(n) = \mathrm{E}[(X_{t+n} - \mu)(X_t - \mu)] = \mathrm{E}[(X_t - \mu)(X_{t+n} - \mu)] = \gamma(-n)$$

(2) 任取 $n$ 维实数向量 $\alpha = (a_1, a_2, \ldots, a_n)^T$ 有：

$$\begin{aligned}
a^T \Gamma_n a &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \gamma_{i-j} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \,\mathrm{E}[(X_i - \mu)(X_j - \mu)] \\
&= \mathrm{E}\left[\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j (X_i - \mu)(X_j - \mu)\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n} a_i (X_i - \mu)\right]^2 \geqslant 0
\end{aligned}$$

(3) 对任意的 $n \in \mathbb{N}$，由**??**可得：

$$\begin{aligned}
|\gamma(n)| &= \Big|\mathrm{E}[(X_{n+1} - \mu)(X_1 - \mu)]\Big| \\
&\leqslant \sqrt{\mathrm{E}[(X_{n+1} - \mu)^2]\,\mathrm{E}[(X_1 - \mu)^2]} \\
&= \sqrt{\gamma(0)\gamma(0)} = |\gamma(0)|
\end{aligned}$$ □

**Definition 3.2.** 任何满足上述三个基本性质的实数序列都被称为**非负定序列**。

## 线性变换

**Theorem 3.1.** 平稳时间序列 $\{X_t\}$ 经过线性变换后得到的还是平稳时间序列。

*Proof.* 只需证明 $\{Y_t = aX_t + b : t \in \mathbb{N}\}$ 对任意的 $a, b \in \mathbb{R}$ 是平稳序列。设 $\mathrm{E}(X_t) = \mu$。

(1) 对于线性变换后时间序列的期望，有：

$$\mathrm{E}(Y_t) = \mathrm{E}(aX_t + b) = a\mu + b$$

(2) 对于线性变换后时间序列的二阶原点矩，有：

$$\mathrm{E}(Y_t^2) = \mathrm{Var}(Y_t) + [\mathrm{E}(Y_t)]^2 = a^2\gamma(0) + (a\mu + b)^2 < +\infty$$ □

(3) 对于线性变换后时间序列的协方差，有：

$$\mathrm{Cov}(Y_t, Y_s) = \mathrm{E}[(aX_t + b - a\mu - b)(aX_s + b - a\mu - b)] = a^2\gamma(t - s)$$

### 自相关函数

**Definition 3.3.** $\{X_t\}$ 是一个平稳时间序列，称平稳时间序列：

$$Y_t = \frac{X_t - \mu}{\sqrt{\gamma(0)}},\ t \in \mathbb{N}$$

为 $\{X_t\}$ 的标准化序列。称 $\{Y_t\}$ 的自协方差函数 $\rho(t)$ 为 $\{X_t\}$ 的自相关函数 *(auto-correlation function)*。因为自协方差函数都是非负定序列，所以 $\rho(t)$ 也是非负定序列。

**Theorem 3.2.** 设 $\{X_t\}$ 的自协方差函数为 $\gamma(t)$，$\{Y_t\}$ 是 $\{X_t\}$ 的标准化序列同时它的自协方差函数为 $\rho(t)$，则：

$$\rho(t) = \frac{\gamma(t)}{\gamma(0)},\ t \in \mathbb{N}$$

*Proof.* 显然 $\mu_Y = \mathrm{E}(Y_t) = 0$，由自协方差函数的定义：

$$\rho(t) = \mathrm{E}[(Y_t - \mu_Y)(Y_0 - \mu_Y)] = \mathrm{E}(Y_t Y_0) = \mathrm{E}\left[\left(\frac{X_t - \mu}{\sqrt{\gamma(0)}}\right)\left(\frac{X_0 - \mu}{\sqrt{\gamma(0)}}\right)\right] = \frac{\gamma(t)}{\gamma(0)} \qquad \square$$

### 正交与不相关

**Definition 3.4.** 设 $\{X_t\}$ 和 $\{Y_t\}$ 是平稳时间序列。

1. 若对任何的 $s, t \in \mathbb{N}$，有 $\mathrm{E}(X_t Y_s) = 0$，则称 $\{X_t\}$ 和 $\{Y_t\}$ **正交**。
2. 若对任何的 $s, t \in \mathbb{N}$，有 $\mathrm{Cov}(X_t, Y_s) = 0$，则称 $\{X_t\}$ 和 $\{Y_t\}$ **不相关**。

**Theorem 3.3.** 对于期望为 0 的平稳时间序列，正交性和不相关性等价。

*Proof.* 若 $\mu = 0$，则：

$$\mathrm{Cov}(X_t, X_s) = \mathrm{E}[(X_t - \mu)(X_s - \mu)] = \mathrm{E}(X_t X_s) \qquad \square$$

**Theorem 3.4.** 设 $\gamma_X(t)$ 和 $\gamma_Y(t)$ 分别是平稳时间序列 $\{X_t\}$ 和 $\{Y_t\}$ 的自协方差函数。记 $\mu_X = \mathrm{E}(X_t)$，$\mu_Y = \mathrm{E}(Y_t)$，定义：

$$Z_t = X_t + Y_t,\ t \in \mathbb{N}$$

则：

1. 若 $\{X_t\}$ 和 $\{Y_t\}$ 正交，则 $\{Z_t\}$ 是平稳时间序列，有自协方差函数：

$$\gamma_Z(t) = \gamma_X(t) + \gamma_Y(t) - 2\mu_X\mu_Y,\ t \in \mathbb{N}$$

2. 若 $\{X_t\}$ 和 $\{Y_t\}$ 不相关，则 $\{Z_t\}$ 是平稳时间序列，有自协方差函数：

$$\gamma_Z(t) = \gamma_X(t) + \gamma_Y(t),\ t \in \mathbb{N}$$

*Proof.* 因为对任意的 $t \in \mathbb{N}$，有 $(X_t+Y_t)^2 \leqslant 2X_t^2+2Y_t^2$，所以：

$$\mathrm{E}(Z_t^2) \leqslant 2\,\mathrm{E}(X_t^2)+2\,\mathrm{E}(Y_t^2) < +\infty$$

显然：

$$\mathrm{E}(Z_t)=\mathrm{E}(X_t)+\mathrm{E}(Y_t)=\mu_X+\mu_Y$$

与 $t$ 无关。因为：

$$\begin{aligned}\mathrm{Cov}(Z_t,Z_s)&=\mathrm{Cov}(X_t+Y_t,X_s+Y_s)\\&=\mathrm{Cov}(X_t,X_s)+\mathrm{Cov}(X_t,Y_s)+\mathrm{Cov}(Y_t,X_s)+\mathrm{Cov}(Y_t,Y_s)\end{aligned}$$

(1) 由于 $\{X_t\}$ 和 $\{Y_t\}$ 正交，所以 $\mathrm{E}(X_tY_s)=\mathrm{E}(X_sY_t)=0$，于是：

$$\begin{aligned}\mathrm{Cov}(Z_t,Z_s)&=\gamma_X(t-s)+\gamma_Y(t-s)+\mathrm{E}(X_tY_s)-\mathrm{E}(X_t)\,\mathrm{E}(Y_s)+\mathrm{E}(X_sY_t)-\mathrm{E}(Y_t)\,\mathrm{E}(X_s)\\&=\gamma_X(t-s)+\gamma_Y(t-s)-2\mu_X\mu_Y\end{aligned}$$

所以 $\mathrm{Cov}(Z_t,Z_s)$ 只与 $(t-s)$ 有关。综上，$\{Z_t\}$ 是平稳时间序列，且

$$\gamma_Z(t)=\gamma_X(t)+\gamma_Y(t)-2\mu_X\mu_Y,\ t\in\mathbb{N}$$

(2) 因为 $\{X_t\}$ 和 $\{Y_t\}$ 不相关，所以 $\mathrm{Cov}(X_t,Y_s)=0=\mathrm{Cov}(Y_t,X_s)=0$，于是：

$$\mathrm{Cov}(Z_t,Z_s)=\mathrm{Cov}(X_t,X_s)+\mathrm{Cov}(Y_t,Y_s)=\gamma_X(t-s)+\gamma_Y(t-s)$$

所以以 $\mathrm{Cov}(Z_t,Z_s)$ 只与 $(t-s)$ 有关。综上，$\{Z_t\}$ 是平稳时间序列，且

$$\gamma_Z(t)=\gamma_X(t)+\gamma_Y(t),\ t\in\mathbb{N}$$ □

### 线性相关

**Definition 3.5.** 对于随机变量 $X_1,X_2,\ldots,X_n$,若存在非零的 $n$ 维实向量 $\alpha=(a_1,a_2,\ldots,a_n)^T$ 使得：

$$\mathrm{Var}\left[\sum_{i=1}^n a_i(X_i-\mu)\right]=0$$

则称 $X_1,X_2,\ldots,X_n$ 线性相关。

**Lemma 3.1.** 对称阵 $A\in M_n(\mathbb{R})$ 退化的充分必要条件为存在 $\mathbb{R}^n$ 中的非零向量 $\alpha$ 使得 $\alpha^TA\alpha=0$。

*Proof.* **(1) 充分性**：由实对称矩阵的正交相似，有 $Q^TAQ=B$，其中 $Q$ 是正交矩阵，$B=\mathrm{diag}\{b_1,b_2,\ldots,b_n\}$ 是主对角线上为 $A$ 特征值的对角矩阵。假设此时 $A$ 可逆，则它的所有特征值都不为 0（否则就有 $|A|=|Q^TBQ|=|Q^T|\,|B|\,|Q|=|B|=0$）。由题设存在非零向量 $\alpha$ 使得 $\alpha^TA\alpha=0$，因为 $Q$ 是正交矩阵，所以 $Q^{-1}$ 可逆，于是 $Q^{-1}x=\mathbf{0}$ 只有零解，所以 $Q^{-1}\alpha\neq\mathbf{0}$。设 $Q^{-1}\alpha=\beta=(y_1,y_2,\ldots,y_n)^T$，则有：

$$\beta^TB\beta=\alpha^T(Q^{-1})^TBQ^{-1}\alpha=\alpha^T(Q^T)^{-1}BQ^{-1}\alpha=\alpha^TA\alpha=0$$

注意到：

$$\beta^T B\beta = \sum_{i=1}^{n} b_i y_i^2$$

所以有：

$$\sum_{i=1}^{n} b_i y_i^2 = 0$$

但是此时 $b_i$ 都不是 0，若上式成立需要 $y_i$ 都为 0，这就与 $\beta \neq \mathbf{0}$ 矛盾，所以 $A$ 退化。

**(2) 必要性：** 如果 $A$ 退化，则存在非零向量 $\alpha$ 使得 $A\alpha = \mathbf{0}$，显然此时 $\alpha^T A\alpha = 0$。 □

**Theorem 3.5.** $\{X_t\}$ 是一个平稳时间序列，$\Gamma_n$ 是 $\{X_t\}$ 的 $n$ 阶自协方差矩阵，则 $\Gamma_n$ 退化的充要条件是对任意的 $t \in \mathbb{N}$，$X_{t+1}, X_{t+2}, \ldots, X_{t+n}$ 线性相关。

*Proof.* 由引理 3.1可知 $\Gamma_n$ 退化的充要条件是存在非零的 $n$ 维实向量 $\alpha = (a_1, a_2, \ldots, a_n)^T$ 使得 $\alpha^T\Gamma_n\alpha = 0$，而：

$$\begin{aligned}
\alpha^T\Gamma_n\alpha &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \,\mathrm{E}[(X_i-\mu)(X_j-\mu)] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \,\mathrm{E}[(X_{t+i}-\mu)(X_{t+j}-\mu)] \\
&= \mathrm{E}\left[\sum_{i=1}^{n} a_i(X_{t+i}-\mu)\right]^2 = \mathrm{E}\left[\sum_{i=1}^{n} a_i(X_{t+i}-\mu)\right]^2 + 0 \\
&= \mathrm{E}\left[\sum_{i=1}^{n} a_i(X_{t+i}-\mu)\right]^2 + \left\{\mathrm{E}\left[\sum_{i=1}^{n} a_i(X_{t+i}-\mu)\right]\right\}^2 = \mathrm{Var}\left[\sum_{i=1}^{n} a_i(X_{t+i}-\mu)\right]
\end{aligned}$$

由随机变量线性相关的定义结论得证。 □

**Theorem 3.6.** $\{X_t\}$ 是一个平稳时间序列，$\Gamma_n$ 是 $\{X_t\}$ 的 $n$ 阶自协方差矩阵。若 $\Gamma_n$ 退化，只要 $m > n$，则有 $\Gamma_m$ 退化，也即若 $X_{t+1}, X_{t+2}, \ldots, X_{t+n}$ 线性相关，只要 $m > n$，则有 $X_{t+1}, X_{t+2}, \ldots, X_{t+m}$ 线性相关。

*Proof.* 由定义 3.5与定理 3.5可直接得出，只需在方差公式中取 $a_{n+1} = a_{n+2} = \cdots = a_m = 0$ 即可。 □

### 3.1.2 线性平稳序列

#### 白噪声

**Definition 3.6.** 设 $\{\varepsilon_t\}$ 是一个平稳时间序列。如果对任何的 $s, t \in \mathbb{N}$，有：

$$\mathrm{E}(\varepsilon_t) = \mu,\ \mathrm{Cov}(\varepsilon_t, \varepsilon_s) = \begin{cases} \sigma^2, & t = s \\ 0, & t \neq s \end{cases}$$

则称 $\{\varepsilon_t\}$ 是白噪声 *(white noise)*，记作 $WN(\mu, \sigma^2)$。当 $\{\varepsilon_t\}$ 是独立序列时，称 $\{\varepsilon_t\}$ 是**独立白噪声**；当 $\mu = 0$ 时，称 $\{\varepsilon_t\}$ 是**零均值白噪声**；当 $\mu = 0$, $\sigma^2 = 1$ 时，称 $\{\varepsilon_t\}$ 是**标准白噪声**；当 $\varepsilon_t$ 服从正态分布时，称 $\{\varepsilon_t\}$ 是**正态白噪声**。

## 有限滑动平均

**Definition 3.7.** $\{\varepsilon_t\} = \{\varepsilon_t : t \in \mathbb{Z}\}$ 是 $WN(0, \sigma^2)$。称：

$$X_t = a_0\varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t-q}$$

是白噪声 $\{\varepsilon_t\}$ 的有限滑动平均 *(finite moving average)*，其中 $q \in \mathbb{N}$，$a_0, a_1, \ldots, a_q$ 为常数。

**Theorem 3.7.** $\{\varepsilon_t\} = \{\varepsilon_t : t \in \mathbb{Z}\}$ 是 $WN(0, \sigma^2)$，则该白噪声的有限滑动平均：

$$X_t = a_0\varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t-q}$$

构成的序列 $\{X_t\}$ 具有如下均值与自协方差函数：

$$\mathrm{E}(X_t) = 0,\ \gamma(n) = \begin{cases} \sigma^2 \sum\limits_{i=0}^{q-n} a_i a_{i+n}, & 0 \leqslant n \leqslant q \\ 0, & n > q \end{cases},\ t, n \in \mathbb{Z}$$

*Proof.* 由有限滑动平均的定义，对于 $\{X_t\}$ 的期望有：

$$\mathrm{E}(X_t) = \mathrm{E}\left(\sum_{i=0}^{q} a_i\varepsilon_{t-i}\right) = \sum_{i=0}^{q} a_i \,\mathrm{E}(\varepsilon_{t-i}) = 0,\ t \in \mathbb{Z}$$

对于 $\{X_t\}$ 的自协方差函数，当 $n > q$ 时，$t + n - i > t - j$ 恒成立，由白噪声的定义可得 $\mathrm{E}(\varepsilon_{t+n-i}\varepsilon_{t-j}) = 0,\ \forall\, i, j = 0, 1, 2, \ldots, q$，于是 $\gamma(n) = 0$。当 $0 \leqslant n \leqslant q$ 时，有：

$$\begin{aligned}
\gamma(n) &= \mathrm{E}(X_{t+n}X_t) \\
&= \mathrm{E}[(a_n\varepsilon_t + a_{n+1}\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t+n-q})(a_0\varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_{q-n}\varepsilon_{t+n-q})] \\
&= \mathrm{E}(a_n a_0\varepsilon_t^2 + a_{n+1}a_1\varepsilon_{t-1}^2 + \cdots + a_q a_{q-n}\varepsilon_{t+n-q}^2) \\
&= \sum_{i=0}^{q-n} a_i a_{i+n}\,\mathrm{E}(\varepsilon_{t-i}^2) = \sum_{i=0}^{q-n} a_i a_{i+n}\sigma^2
\end{aligned}$$

综上可得：

$$\gamma(n) = \begin{cases} \sigma^2 \sum\limits_{i=0}^{q-n} a_i a_{i+n}, & 0 \leqslant n \leqslant q \\ 0, & n > q \end{cases},\ n \in \mathbb{N}$$

□

## 单边滑动平均

**Definition 3.8.** $\{\varepsilon_t\} = \{\varepsilon_t : t \in \mathbb{Z}\}$ 是 $WN(0, \sigma^2)$。称：

$$X_t = \sum_{i=0}^{+\infty} a_i\varepsilon_{t-i},\ t \in \mathbb{Z}$$

是白噪声 $\{\varepsilon_t\}$ 的**单边滑动平均**，其中 $a_0, a_1, \ldots$ 为常数。它表明当前的观测 $X_t$ 只与 $t$ 时刻以及之前时刻的白噪声相关，与 $t$ 时刻之后的白噪声无关。

### 无穷滑动平均

**Definition 3.9.** $\{\varepsilon_t\} = \{\varepsilon_t : t \in \mathbb{Z}\}$ 是 $WN(0, \sigma^2)$，$\{a_n\} \in l^1$。称：

$$X_t = \sum_{i=-\infty}^{+\infty} a_i \varepsilon_{t-i}$$

是白噪声 $\{\varepsilon_t\}$ 的无穷滑动平均 *(infinite moving average)*。

**Theorem 3.8.** $\{\varepsilon_t\} = \{\varepsilon_t : t \in \mathbb{Z}\}$ 是 $WN(0, \sigma^2)$，则该白噪声的无穷滑动平均：

$$X_t = \sum_{i=-\infty}^{+\infty} a_i \varepsilon_{t-i}$$

构成的序列 $\{X_t\}$ 是平稳序列且具有如下均值与自协方差函数：

$$\mathrm{E}(X_t) = 0,\ \gamma(n) = \sigma^2 \sum_{i=-\infty}^{+\infty} a_i a_{i+n}$$

*Proof.* 由性质 1.4.2(4)(6) 可得：

$$\begin{aligned}\mathrm{E}\left(\sum_{i=-\infty}^{+\infty} |a_i \varepsilon_{t-i}|\right) &= \mathrm{E}\left[\lim_{n\to+\infty}\left(\sum_{i=-n}^{n} |a_i||\varepsilon_{t-i}|\right)\right] = \lim_{n\to+\infty} \mathrm{E}\left(\sum_{i=-n}^{n} |a_i||\varepsilon_{t-i}|\right)\\ &= \lim_{n\to+\infty}\left[\sum_{i=-n}^{n} |a_i| \,\mathrm{E}(|\varepsilon_{t-i}|)\right] = \sum_{i=-\infty}^{+\infty} |a_i| \,\mathrm{E}(|\varepsilon_{t-i}|)\end{aligned}$$

由**??**可得：

$$\mathrm{E}(|\varepsilon_{t-i}|) = \left|\mathrm{E}(|\varepsilon_{t-i}| \cdot 1)\right| \leqslant \sqrt{\mathrm{E}(\varepsilon_{t-i}^2)\,\mathrm{E}(1)} = \sigma$$

于是：

$$\mathrm{E}\left(\sum_{i=-\infty}^{+\infty} |a_i \varepsilon_{t-i}|\right) = \sum_{i=-\infty}^{+\infty} |a_i| \,\mathrm{E}(|\varepsilon_{t-i}|) \leqslant \sigma \sum_{i=-\infty}^{+\infty} |a_i| < +\infty$$

由性质 1.4.2(9) 可得对任意的 $t \in \mathbb{Z}$，$X_t$ a.e. 有限，即：

$$X_t = \sum_{i=-\infty}^{+\infty} a_i \varepsilon_{t-i}$$

右式 a.e. 收敛。取控制函数 $\sum_{i=-\infty}^{+\infty} |a_i \varepsilon_{t-i}|$，由定理 1.47和性质 1.4.3(6) 可得

$$\mathrm{E}(X_t) = \mathrm{E}\left[\lim_{n\to+\infty}\left(\sum_{i=-n}^{n} a_i \varepsilon_{t-i}\right)\right] = \lim_{n\to+\infty}\left[\mathrm{E}\left(\sum_{i=-n}^{n} a_i \varepsilon_{t-i}\right)\right] = 0$$

对 $t, s \in \mathbb{Z}$ 定义：

$$\varphi_n = \sum_{i=-n}^{n} a_i \varepsilon_{t-i},\quad \psi_n = \sum_{j=-n}^{n} a_j \varepsilon_{s-j}$$

则有 $\varphi_n\psi_n \to X_tX_s$，因为对任意的 $t,s\in\mathbb{Z}$，$X_t$ a.e. 有限，所以 $X_tX_s$ a.e. 有限。由性质 1.4.2(4)(6) 可得：

$$\mathrm{E}\left(\sum_{i=-\infty}^{+\infty}\sum_{j=-\infty}^{+\infty}|a_ia_j\varepsilon_{t-i}\varepsilon_{s-j}|\right)=\sum_{i=-\infty}^{+\infty}\sum_{j=-\infty}^{+\infty}|a_ia_j|\,\mathrm{E}(|\varepsilon_{t-i}\varepsilon_{s-j}|)$$

由**??**可得：

$$\mathrm{E}(|\varepsilon_{t-i}\varepsilon_{s-j}|)=\Big|\mathrm{E}(|\varepsilon_{t-i}\varepsilon_{s-j}|)\Big|\leqslant\sqrt{\mathrm{E}(\varepsilon_{t-i}^2)\,\mathrm{E}(\varepsilon_{s-j}^2)}=\sigma^2$$

于是：

$$\mathrm{E}\left(\sum_{i=-\infty}^{+\infty}\sum_{j=-\infty}^{+\infty}|a_ia_j\varepsilon_{t-i}\varepsilon_{s-j}|\right)\leqslant\sigma^2\sum_{i=-\infty}^{+\infty}\sum_{j=-\infty}^{+\infty}|a_ia_j|=\sigma^2\left(\sum_{i=-\infty}^{+\infty}|a_i|\right)^2<+\infty$$

取控制函数 $\sum\limits_{i=-\infty}^{+\infty}\sum\limits_{j=-\infty}^{+\infty}|a_ia_j\varepsilon_{t-i}\varepsilon_{s-j}|$，由定理 1.47和性质 1.4.3(6) 可得：

$$\begin{aligned}\mathrm{Cov}(X_t,X_s)=\mathrm{E}(X_tX_s)=\mathrm{E}\left(\lim_{n\to+\infty}\varphi_n\psi_n\right)&=\lim_{n\to+\infty}\mathrm{E}(\varphi_n\psi_n)\\&=\lim_{n\to+\infty}\mathrm{E}\left[\left(\sum_{i=-n}^{n}a_i\varepsilon_{t-i}\right)\left(\sum_{j=-n}^{n}a_j\varepsilon_{s-j}\right)\right]\\&=\lim_{n\to+\infty}\mathrm{E}\left(\sum_{i=-n}^{n}\sum_{j=-n}^{n}a_ia_j\varepsilon_{t-i}\varepsilon_{s-j}\right)\\&=\lim_{n\to+\infty}\left[\sum_{i=-n}^{n}\sum_{j=-n}^{n}a_ia_j\,\mathrm{E}(\varepsilon_{t-i}\varepsilon_{s-j})\right]\\&=\sum_{i=-\infty}^{+\infty}a_ia_{i-(t-s)}\sigma^2=\sigma^2\sum_{i=-\infty}^{+\infty}a_ia_{i+(t-s)}\end{aligned}$$

上式最后一步是因为 $t-i=s-j$ 时期望才不为 0，并且 $n$ 是逐渐变大趋于无穷，所以也不用考虑 $n$ 为定值时 $t,s$ 相差过大导致索引越界的问题。由协方差公式可以看出其只与 $t-s$ 相关。

注意到：

$$\mathrm{Var}(X_t)=\sigma^2\sum_{i=-\infty}^{+\infty}a_i^2$$

因为 $l^1\subset l^2$，$\{a_n\}\in l^1$，所以上式也收敛。

综上，$\{X_t\}$ 是一个平稳序列。 □

### 线性平稳序列

**Definition 3.10.** $\{\varepsilon_t\}=\{\varepsilon_t:t\in\mathbb{Z}\}$ 是 $WN(0,\sigma^2)$，$\{a_n\}\in l^2$。称：

$$X_t=\sum_{i=-\infty}^{+\infty}a_i\varepsilon_{t-i}$$

构成的序列 $\{X_t\}$ 为线性平稳序列 *(linearly stationary series)*。

-

**Theorem 3.9.** 线性平稳序列是平稳序列，且有：

$$\mathrm{E}(X_t) = 0,\ \gamma(n) = \sigma^2 \sum_{i=-\infty}^{+\infty} a_i a_{i+n}$$

*Proof.* 在 $L^2$ 空间中定义内积 $(X,Y) = \mathrm{E}(XY)$，定义：

$$\varphi_n = \sum_{i=-n}^{n} a_i \varepsilon_{t-i}$$

则对 $m < n$, $n \to +\infty$ 有：

$$||\varphi_n - \varphi_m||^2 = \left\| \sum_{i=m+1}^{n} a_i \varepsilon_{t-i} + \sum_{i=-n}^{-m-1} a_i \varepsilon_{t-i} \right\|^2 = \sigma^2 \left( \sum_{i=m+1}^{n} a_i^2 + \sum_{i=-n}^{-m-1} a_i^2 \right) \to 0$$

由定理 1.50可知 $X_t \in L_2$。由内积的连续性和性质 1.4.3(6) 可得：

$$\mathrm{E}(X_t) = (X_t, 1) = \lim_{n\to+\infty} (\varphi_n, 1) = \lim_{n\to+\infty} \mathrm{E}(\varphi_n) = 0$$

$$\begin{aligned}
\mathrm{Cov}(X_t, X_s) = \mathrm{E}(X_t X_s) &= \lim_{n\to+\infty} \left( \sum_{i=-n}^{n} a_i \varepsilon_{t-i}, \sum_{i=-n}^{n} a_i \varepsilon_{s-i} \right) \\
&= \lim_{n\to+\infty} \mathrm{E} \left( \sum_{i=-n}^{n} a_i \varepsilon_{t-i} \sum_{j=-n}^{n} a_j \varepsilon_{s-j} \right) = \lim_{n\to+\infty} \mathrm{E} \left( \sum_{i=-n}^{n} \sum_{j=-n}^{n} a_i a_j \varepsilon_{t-i} \varepsilon_{s-j} \right) \\
&= \sigma^2 \sum_{i=-\infty}^{+\infty} a_i a_{i+(t-s)}
\end{aligned}$$

上式倒数第二步到最后一步和定理 3.8中是一样的，同时平稳性的分析也与之一样，故省略。 □

**Theorem 3.10.** 对于线性平稳序列的自协方差函数 $\gamma(n)$，有：

$$\lim_{n\to\infty} \gamma(n) = 0$$

*Proof.* 由**??**可得：

$$\begin{aligned}
|\gamma(n)| = \sigma^2 \Big| \sum_{i=-\infty}^{+\infty} a_i a_{i+n} \Big| &\leqslant \sigma^2 \sum_{i=-\infty}^{+\infty} |a_i a_{i+n}| \\
&= \sigma^2 \sum_{|i|\leqslant n/2} |a_i||a_{i+n}| + \sigma^2 \sum_{|i|>n/2} |a_i||a_{i+n}| \\
&\leqslant \sigma^2 \left( \sum_{|i|\leqslant n/2} a_i^2 \sum_{|i|\leqslant n} a_{i+n}^2 \right)^{\frac{1}{2}} + \sigma^2 \left( \sum_{|i|>n/2} a_i^2 \sum_{|i|>n} a_{i+n}^2 \right)^{\frac{1}{2}} \\
&\leqslant \sigma^2 \left( \sum_{i=-\infty}^{+\infty} a_i^2 \sum_{|i|\leqslant n/2} a_{i+n}^2 \right)^{\frac{1}{2}} + \sigma^2 \left( \sum_{i=-\infty}^{+\infty} a_i^2 \sum_{|i|>n/2} a_{i+n}^2 \right)^{\frac{1}{2}}
\end{aligned}$$

注意到 $|i| \leqslant \frac{n}{2}$ 时 $\frac{n}{2} \leqslant i+n \leqslant \frac{3n}{2}$，所以有：

$$\sum_{|i|\leqslant n/2} a_{i+n}^2 \leqslant \sum_{|i|>n/2} a_i^2$$

结合 $\{a_n\} \in l^2$ 即可得：

$$|\gamma(n)| \leqslant 2\sigma^2 \left( \sum_{i=-\infty}^{+\infty} a_i^2 \sum_{|i|>n/2} a_i^2 \right)^{\frac{1}{2}} = 2\sigma^2 \left( \sum_{i=-\infty}^{+\infty} a_i^2 \right)^{\frac{1}{2}} \left( \sum_{|i|>n/2} a_i^2 \right)^{\frac{1}{2}} \to 0 \qquad \square$$

### 3.1.3 平稳序列的谱密度

**Definition 3.11.** 设平稳序列 $\{X_t\}$ 的自协方差函数为 $\gamma(n)$。

1. 若存在 $[-\pi, \pi]$ 上单调不减且右连续的函数 $F(\lambda)$ 使得：

   $$\gamma(n) = \int_{[-\pi,\pi]} e^{in\lambda} \,\mathrm{d}F(\lambda), \quad F(-\pi) = 0,\ k \in \mathbb{Z}$$

   则称 $F(\lambda)$ 为 $\{X_t\}$ 或 $\{\gamma(n)\}$ 的谱分布函数 *(spectral distribution function)*，简称为**谱函数**；

2. 若存在 $[-\pi, \pi]$ 上的非负函数 $f(\lambda)$ 使得：

   $$\gamma(n) = \int_{[-\pi,\pi]} f(\lambda) e^{in\lambda} \,\mathrm{d}\lambda, \quad n \in \mathbb{Z}$$

   则称 $f(\lambda)$ 为 $\{X_t\}$ 或 $\{\gamma(n)\}$ 的谱密度函数 *(spectral density function)*，简称为**谱密度**。

## 3.2 本科时间序列

$$\bar{x} = \frac{\sum\limits_{t=1}^{n} x_t}{n}$$

$$\hat{\gamma}(k) = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{n-k}, \ \hat{\gamma}(0) = s^2$$

$$\hat{\rho}_k = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}$$

acf(x, lag=) 虚线为自相关系数 2 倍标准差位置

#### 平稳性检验

1. 时序图观察
2. 自相关系数图 acf 函数，应呈现出迅速衰减向 0

**白噪声检验**

同均值同方差不相关

**Theorem 3.11.** 如果一个时间序列是白噪声，得到一个观察期数为 $n$ 的观察序列 $\{x_t\}$，那么有：

$$\hat{\rho}_k \sim \mathrm{N}\left(0, \frac{1}{n}\right), \ \forall\, k \neq 0$$

近似成立

**推导 3.1.** 构建假设：

1. 原假设：延迟期数小于或等于 $m$ 期的序列值之间相互独立，即：

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_m = 0$$

2. 备择假设：延迟期数小于或等于 $m$ 期的序列值之间有相关性，即：

$$\text{至少存在某个} \rho_k \neq 0,\ k \leqslant m$$

确定这里是有相关性而不是不独立吗？

构建 $Q$ 统计量：

$$Q = n \sum_{k=1}^{m} \hat{\rho}_k^2$$

若原假设成立，则 $\hat{\rho}_k^2$ 之间也彼此独立，于是：

$$Q = \sum_{k=1}^{m} (\sqrt{n}\hat{\rho}_k)^2 \sim \chi_m^2$$

当原假设不成立时，$Q$ 统计量的值应该偏大，于是拒绝域取 $\chi_m^2$ 分布的上 $\alpha$ 分位点。

*Box* 和 *Ljung* 为了弥补小样本情况时 $Q$ 统计量效果较差的问题，推导出了 *LB* 统计量：

$$LB = n(n+2) \sum_{k=1}^{m} \left( \frac{\hat{\rho}_k^2}{n-k} \right) \sim \chi_m^2$$

两种检验的代码为 *Box.test(x, type=, lag=)*

## 3.3 线性差分方程理论

### 3.3.1 差分与位移

**Definition 3.12.** 设 $f(x)$ 是定义在 $\mathbb{R}$ 上的函数，称 $\Delta f(x) = f(x) - f(x-1)$ 为 $f(x)$ 在 $x$ 处的 1 阶**差分**，称 $\Delta$ 为**差分算子**，对 $n-1$ 阶差分后的函数再进行一次 1 阶差分运算称为 $n$ 阶差分，记 $\Delta^n f(x)$ 为 $f(x)$ 的 $n$ 阶差分，则：

$$\Delta^n f(x) = \Delta^{n-1} f(x) - \Delta^{n-1} f(x-1)$$

将 $\mathcal{B} f(x) = f(x-1)$ 称为 $f(x)$ 在 $x$ 处的 1 步**位移**，称 $\mathcal{B}$ 为位移算子，对 $n-1$ 步位移后的函数再进行一次 1 步位移运算称为 $n$ 步位移，记 $\mathcal{B}^n f(x)$ 为 $f(x)$ 的 $n$ 步位移，则：

$$\mathcal{B}^n f(x) = f(x-n)$$

**Property 3.3.1.** 设 $f(x), g(x)$ 为定义在 $\mathbb{R}$ 上的函数，$\alpha, \beta$ 是任意常数，$\mathcal{B}$ 为位移算子。差分算子 $\Delta$ 和位移算子 $\mathcal{B}$ 具有如下性质：

1. 差分算子是线性算子，即：
$$\Delta[\alpha f(x) + \beta g(x)] = \alpha\Delta f(x) + \beta\Delta g(x)$$
2. $\Delta[f(x)g(x)] = \Delta f(x)g(x) + \Delta g(x)\mathcal{B}f(x)$；
3. $\Delta\left[\dfrac{f(x)}{g(x)}\right] = \dfrac{\Delta f(x)\mathcal{B}g(x) - \Delta g(x)\mathcal{B}f(x)}{g(x)\mathcal{B}g(x)}$；
4. $\mathcal{B}\alpha = \alpha$；
5. 位移算子是线性算子，即：
$$\mathcal{B}[\alpha f(x) + \beta g(x)] = \alpha\mathcal{B}f(x) + \beta\mathcal{B}g(x)$$
6. 对于多项式 $\varphi(x) = \sum\limits_{i=0}^{m} a_i x^i$，$\psi(x) = \sum\limits_{i=0}^{n} b_i x^i$ 的乘积 $A(x) = \varphi(x)\psi(x)$，有：
$$A(\mathcal{B})f(x) = \varphi(\mathcal{B})[\psi(\mathcal{B})f(x)] = \psi(\mathcal{B})[\varphi(\mathcal{B})f(x)]$$
7. 差分算子与位移算子具有如下关系：
$$\Delta^n = (I - \mathcal{B})^n = \sum_{i=0}^{n} \binom{n}{i}(-1)^i\mathcal{B}^i$$
$$\mathcal{B}^n = (I - \Delta)^n = \sum_{i=0}^{n} \binom{n}{i}(-1)^i\Delta^i$$
$$\Delta\mathcal{B} = \mathcal{B}\Delta$$

*Proof.* (1)(4)(5)(7) 是显然的。

(2) 注意到：
$$\begin{aligned}
\Delta[f(x)g(x)] &= f(x)g(x) - f(x-1)g(x-1) \\
&= f(x)g(x) - f(x-1)g(x) + g(x)f(x-1) - g(x-1)f(x-1) \\
&= [f(x) - f(x-1)]g(x) + [g(x) - g(x-1)]f(x-1) \\
&= \Delta f(x)g(x) + \Delta g(x)\mathcal{B}f(x)
\end{aligned}$$

(3) 注意到：
$$\begin{aligned}
\Delta\left[\frac{f(x)}{g(x)}\right] &= \frac{f(x)}{g(x)} - \frac{f(x-1)}{g(x-1)} = \frac{f(x)g(x-1) - f(x-1)g(x)}{g(x)g(x-1)} \\
&= \frac{f(x)g(x-1) - f(x-1)g(x-1) + f(x-1)g(x-1) - f(x-1)g(x)}{g(x)\mathcal{B}g(x)} \\
&= \frac{[f(x) - f(x-1)]g(x-1) - [g(x) - g(x-1)]f(x-1)}{g(x)\mathcal{B}g(x)} \\
&= \frac{\Delta f(x)\mathcal{B}g(x) - \Delta g(x)\mathcal{B}f(x)}{g(x)\mathcal{B}g(x)}
\end{aligned}$$

(6) 注意到：

$$A(x)=\varphi(x)\psi(x)=\left(\sum_{i=0}^{m}a_ix^i\right)\left(\sum_{i=0}^{n}b_ix^i\right)=\sum_{i=0}^{m}\sum_{j=0}^{n}a_ib_jx^{i+j}$$

所以：

$$\begin{aligned}A(\mathcal{B})f(x)&=\sum_{i=0}^{m}\sum_{j=0}^{n}a_ib_j\mathcal{B}^{i+j}f(x)=\sum_{i=0}^{m}\sum_{j=0}^{n}a_ib_j\mathcal{B}^i\mathcal{B}^jf(x)\\&=\sum_{i=0}^{m}a_i\mathcal{B}^i\left[\sum_{j=0}^{n}b_j\mathcal{B}^jf(x)\right]=\varphi(\mathcal{B})[\psi(\mathcal{B})f(x)]\end{aligned}$$

同理可证 $A(\mathcal{B})f(x)=\psi(\mathcal{B})[\varphi(\mathcal{B})f(x)]$。 □

### 3.3.2 线性差分方程

**Definition 3.13.** 称方程：

$$x_m+a_1(m)x_{m-1}+a_2(m)x_{m-2}+\cdots+a_n(m)x_{m-n}=f(m)$$

为关于 $\{x_m\}$ 的 $n$ 阶线性差分方程，其中 $f(x),a_1(x),a_2(x),\ldots,a_n(x)$ 为定义在 $\mathbb{Z}$ 上的函数，且 $a_n(x)\neq 0$ 在 $x\in\mathbb{Z}$ 恒成立。当 $f(x)=0$ 时，称上述方程为关于 $\{x_m\}$ 的 $n$ 阶齐次线性差分方程。

**note 3.1.** 这里的线性指的是方程关于 $\{x_m\}$ 是线性的。$x_m-\sin(x_{m-1})$ 不是一个线性方程。

**推导 3.2.** 定义算子：

$$L(x_m)=I+a_1(m)\mathcal{B}x_m+a_2(m)\mathcal{B}^2x_m+\cdots+a_n(m)\mathcal{B}^nx_m$$

则线性差分方程可表示为：

$$L(x_m)=f(m)$$

**Theorem 3.12.** 初值问题：

$$L(x_m)=f(m),\ x_0=y_0,\ x_1=y_1,\ \cdots,\ x_{n-1}=y_{n-1}$$

有唯一的解。

*Proof.* 注意到关系：

$$\begin{gathered}x_m+a_1(m)x_{m-1}+a_2(m)x_{m-2}+\cdots+a_n(m)x_{m-n}=f(m),\ \forall\, m\geqslant n\\x_m=-\frac{x_{m+n}}{a_n(m+n)}-\frac{a_1(m)x_{m+n-1}}{a_n(m+n)}-\cdots-\frac{a_{n-1}x_{m+1}}{a_n(m+n)}+\frac{f(m)}{a_n(m+n)},\ \forall\, m<0\end{gathered}$$

于是当如上 $n$ 个初始值给定时，$\{x_m\}$ 的所有值都可以由上述递推关系唯一得到。 □

**Definition 3.14.** 若存在不全为 0 的常数 $c_1, c_2, \ldots, c_k$ 使得序列 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 满足：

$$c_1 x_m^{(1)} + c_2 x_m^{(2)} + \cdots + c_k x_m^{(k)} = 0, \ \forall\, m \in \mathbb{Z}$$

则称这些序列**线性相关**，否则就称这些序列**线性无关**。

**Definition 3.15.** 称矩阵：

$$C(m) = \begin{pmatrix} x_m^{(1)} & x_m^{(2)} & \cdots & x_m^{(k)} \\ \mathcal{B}x_m^{(1)} & \mathcal{B}x_m^{(2)} & \cdots & \mathcal{B}x_m^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{B}^{k-1}x_m^{(1)} & \mathcal{B}^{k-1}x_m^{(2)} & \cdots & \mathcal{B}^{k-1}x_m^{(k)} \end{pmatrix}$$

为序列 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 的 ***Casorati* 矩阵**，将 $\det C(m)$ 称为其 ***Casorati* 行列式**。

**Theorem 3.13.** 若序列 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 线性相关，则其 *Casorati* 行列式 $\det C(m)$ 在 $m \in \mathbb{Z}$ 上恒为 0。

*Proof.* 因为 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 线性相关，所以存在不全为 0 的常数 $c_1, c_2, \ldots, c_k$ 使得

$$c_1 x_m^{(1)} + c_2 x_m^{(2)} + \cdots + c_k x_m^{(k)} = 0, \ \forall\, m \in \mathbb{Z}$$

于是以 $C(m)$ 为系数矩阵的线性方程组有非零解。由可知 $\det C(m) = 0$ 在 $m \in \mathbb{Z}$ 上恒成立。 链接线性方程组理论

□

**Corollary 3.1.** 若存在 $m \in \mathbb{Z}$ 使得 $\det C(m) \neq 0$，则序列 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 线性无关。

### 齐次线性差分方程解的一般理论

**Theorem 3.14.** 如果 $\{x_m^{(1)}\}, \{x_m^{(2)}\}, \ldots, \{x_m^{(k)}\}$ 是 $n$ 阶常系数齐次线性差分方程 $L(x_m) = 0$ 的 $k$ 个解，则它们的线性组合：

$$\{x_m = c_1 x_m^{(1)} + c_2 x_m^{(2)} + \cdots + c_k x_m^{(k)}\}$$

也是解，其中 $c_i$ 为任意常数，$i = 1, 2, \ldots, k$。

*Proof.* 因为 $x_m^{(j)}$, $j = 1, 2, \ldots, k$ 是 $L(x_m) = 0$ 的解，所以：

$$x_m^{(j)} + a_1 x_{m-1}^{(j)} + \cdots + a_n x_{m-n}^{(j)} = 0$$

于是：

$$\sum_{j=1}^{k} x_m^{(j)} + a_1 \sum_{j=1}^{k} x_{m-1}^{(j)} + \cdots + a_n \sum_{j=1}^{k} x_{m-n}^{j} = 0$$

即：

$$x_m + a_1 x_{m-1} + \cdots + a_n x_{m-n} = 0$$

□

**Theorem 3.15.** $n$ 阶齐次线性差分方程 $L(x_m)=0$ 一定存在 $n$ 个线性无关的解。

*Proof.* 由定理 3.12可知 $L(x_m)=0$ 满足初值条件：

$$\begin{cases} x_0=1,\ x_1=0,\ \dots,\ x_{n-1}=0 \\ x_0=0,\ x_1=1,\ \dots,\ x_{n-1}=0 \\ \cdots\cdots\cdots \\ x_0=0,\ x_1=0,\ \dots,\ x_{n-1}=1 \end{cases}$$

的解一定存在，由推论 3.1可知这 $n$ 个解线性无关。 □

**Theorem 3.16.** 设 $\{x_m^{(1)}\},\{x_m^{(2)}\},\dots,\{x_m^{(n)}\}$ 是 $n$ 阶齐次线性差分方程 $L(x_m)=0$ 的 $n$ 个线性无关的解，则方程的通解可以表示为：

$$x_m=c_1x_m^{(1)}+c_2x_m^{(2)}+\cdots+c_nx_m^{(n)},\ \forall\, m\in\mathbb{Z}$$

其中 $c_1,c_2,\dots,c_n$ 是任意常数。

*Proof.* 因为 $\{x_m^{(1)}\},\{x_m^{(2)}\},\dots,\{x_m^{(n)}\}$ 线性无关，由定理 3.13可知存在 $m\in\mathbb{Z}$ 使得 Casorati 行列式 $\det C(m)\neq 0$，于是此时的 Casorati 矩阵 $C(n)$ 可逆。任取 $L(x_m)=0$ 的一个解 $\{y_m\}$，则 $y_m,\mathcal{B}y_m,\dots,\mathcal{B}^{n-1}y_m$ 是确定的数，于是关于 $(c_1,c_2,\dots,c_n)^T$ 的线性方程组：

$$\begin{pmatrix} x_m^{(1)} & x_m^{(2)} & \cdots & x_m^{(n)} \\ \mathcal{B}x_m^{(1)} & \mathcal{B}x_m^{(2)} & \cdots & \mathcal{B}x_m^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{B}^{n-1}x_m^{(1)} & \mathcal{B}^{n-1}x_m^{(2)} & \cdots & \mathcal{B}^{n-1}x_m^{(n)} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix}=\begin{pmatrix} y_m \\ \mathcal{B}y_m \\ \vdots \\ \mathcal{B}^{n-1}y_m \end{pmatrix}$$

存在唯一解 $(c_1^*,c_2^*,\dots,c_n^*)^T$，即在这 $n$ 个位置处 $\{y_m\}$ 可以由 $\{x_m^{(1)}\},\{x_m^{(2)}\},\dots,\{x_m^{(n)}\}$ 线性表出。令：

$$z_m=c_1^*x_m^{(1)}+c_2^*x_m^{(2)}+\cdots+c_n^*x_m^{(n)},\ \forall\, m\in\mathbb{Z}$$

则 $\{z_m\}$ 在这 $n$ 个位置处的值等于 $\{y_m\}$ 对应位置上的值，且 $\{z_m\}$ 可以由 $\{x_m^{(1)}\},\{x_m^{(2)}\},\dots,\{x_m^{(n)}\}$ 线性表出。而由定理 3.12可知 $L(x_m)=0$ 的解由任意的 $n$ 个初始值唯一确定，于是 $\{y_m\}$ 就是 $\{z_m\}$，所以它也可以由 $\{x_m^{(1)}\},\{x_m^{(2)}\},\dots,\{x_m^{(n)}\}$ 线性表出。由 $\{y_m\}$ 的任意性，结论成立。 □

**Corollary 3.2.** $n$ 阶齐次线性差分方程的所有解构成一个 $n$ 维线性空间。

**Definition 3.16.** 称 $n$ 阶齐次线性差分方程的任意 $n$ 个线性无关的解为其**基本解组**。

### 非齐次线性差分方程解的一般理论

**Theorem 3.17.** $n$ 阶非齐次线性差分方程 $L(x_m)=f(m)$ 任意两个解的差是 $n$ 阶齐次线性差分方程 $L(x_m)=0$ 的解。

**Theorem 3.18.** $n$ 阶非齐次线性差分方程 $L(x_m)=f(m)$ 任意一个解与 $n$ 阶齐次线性差分方程 $L(x_m)=0$ 任意一个解的和还是 $L(x_m)=f(m)$ 的解。

**Theorem 3.19.** 设 $\{x_m^{(1)}\},\{x_m^{(2)}\},\ldots,\{x_m^{(n)}\}$ 是 $n$ 阶齐次线性差分方程 $L(x_m)=0$ 的基本解组，$\{y_m\}$ 是 $n$ 阶非齐次线性差分方程 $L(x_m)=f(m)$ 的一个解，则 $L(x_m)=f(m)$ 的通解可表示为：

$$x_m=y_m+c_1x_m^{(1)}+c_2x_m^{(2)}+\cdots+c_nx_m^{(n)},\ \forall\ m\in\mathbb{Z}$$

*Proof.* 对任意的 $\{y_m\}$，由定理 3.17和定理 3.16可知 $L(x_m)=f(m)$ 的任意一个解可以表示为上述形式。由 $\{y_m\}$ 的任意性可得出结论。 □

### 3.3.3 $n$ 阶常系数线性差分方程

**Definition 3.17.** 若 $n$ 阶线性差分方程：

$$x_m+a_1(m)x_{m-1}+a_2(m)x_{m-2}+\cdots+a_n(m)x_{m-n}=f(m)$$

中 $a_1(x),a_2(x),\ldots,a_n(x)$ 都是常数，则称上述方程为关于 $\{x_m\}$ 的 $n$ 阶常系数线性差分方程。

#### 常系数齐次线性差分方程解的一般理论

**推导 3.3.** 取一个一阶常系数齐次线性差分方程 $x_m+a_1x_{m-1}=0$，可以看出该方程的通解为 $x_m=C(-a_1)^m,\ \forall\ m\in\mathbb{Z}$，其中 $C$ 为任意常数。受此启发，我们对于 $n$ 阶常系数齐次线性差分方程寻找指数形式的解。

**Definition 3.18.** 称：

$$\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}+\cdots+a_{n-1}\lambda+a_n=0$$

为 $n$ 阶常系数齐次线性差分方程：

$$x_m+a_1x_{m-1}+a_2x_{m-2}+\cdots+a_nx_{m-n}=0$$

的**特征方程**，等号左边关于 $\lambda$ 的多项式被称为**特征多项式**，记为 $l(\lambda)$。

**Theorem 3.20.** 设 $\lambda_1,\lambda_2,\ldots,\lambda_s$ 是 $n$ 阶常系数齐次线性差分方程 $L(x_m)=0$ 的特征方程 $l(\lambda)=0$ 的解，其重数分别为 $r_1,r_2,\ldots,r_s$，则 $L(x_m)=0$ 的通解可以表示为：

$$x_m=\sum_{i=1}^{s}\sum_{j=0}^{r_i-1}c_{ij}m^j\lambda_i^m,\ \forall\ m\in\mathbb{Z}$$

其中 $c_{ij}$，$i=1,2,\ldots,s$，$j=0,1,\ldots,r_i-1$ 为常数。

*Proof.* **情况一:**此时特征方程有 $n$ 个互不相同的实根,分别设为 $\lambda_1, \lambda_2, \ldots, \lambda_n$。将 $\{\lambda_i^m\}$, $i = 1, 2, \ldots, n$ 代入 $L(x_m) = 0$ 中可以发现:

$$\begin{aligned}&\lambda_i^m + a_1\lambda_i^{m-1} + a_2\lambda_i^{m-2} + \cdots + a_n\lambda_i^{m-n}\\ =&\lambda_i^{m-n}(\lambda_i^n + a_1\lambda_i^{n-1} + a_2\lambda_i^{n-2} + \cdots + a_n) = 0\end{aligned}$$

所以它们就是 $L(x_m) = 0$ 的解。由 $n$ 阶线性差分方程的定义,$a_n \neq 0$,所以 $\lambda_i \neq 0$。于是有:

$$\det C(0) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1^{-1} & \lambda_2^{-1} & \cdots & \lambda_n^{-1} \\ \lambda_1^{-2} & \lambda_2^{-2} & \cdots & \lambda_n^{-2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{-(n-1)} & \lambda_2^{-(n-1)} & \cdots & \lambda_n^{-(n-1)} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (\lambda_j - \lambda_i) \neq 0$$

由推论 3.1可知这 $n$ 个解线性无关,所以构成一个基本解组。

**情况二:**此时特征方程存在重实根。假设根 $\lambda_i$ 的重数为 $r_i$,代入可知 $\{\lambda_i^m\}$ 是 $L(x_m) = 0$ 的解。下证明对 $1 \leqslant b \leqslant r_i - 1$ 且 $b \in \mathbb{N}^+$,$\{m^b\lambda_i^m\}$ 都是 $L(x_m) = 0$ 的解。设 $a_0 = 1$,则:

$$\begin{aligned}&m^b\lambda_i^m + a_1(m-1)^b\lambda_i^{m-1} + \cdots + a_n(m-n)^b\lambda_i^{m-n}\\ =&\sum_{k=0}^{n} a_k(m-k)^b\lambda_i^{m-k} = \sum_{k=0}^{n} a_k \sum_{j=0}^{b} \binom{b}{j} m^j(-k)^{b-j}\lambda_i^{m-k}\\ =&\sum_{k=0}^{n}\sum_{j=0}^{b} a_k \binom{b}{j} m^j(-k)^{b-j}\lambda_i^{m-k} = \sum_{j=0}^{b}\sum_{k=0}^{n} a_k \binom{b}{j} m^j(-k)^{b-j}\lambda_i^{m-k}\\ =&\sum_{j=0}^{b} \binom{b}{j} m^j \sum_{k=0}^{n} a_k(-k)^{b-j}\lambda_i^{m-k} = \lambda_i^m \sum_{j=0}^{b} \binom{b}{j} m^j \sum_{k=0}^{n} a_k(-k)^{b-j}\lambda_i^{-k}\end{aligned}$$

取函数 $f(\lambda_i) = \sum\limits_{k=0}^{n} a_k\lambda_i^{-k}$,则:

$$\sum_{k=0}^{n} a_k(-k)^{b-j}\lambda_i^{-k}$$

可以表示为 $f(\lambda_i), f'(\lambda_i), \ldots, f^{(b-j)}(\lambda_i)$ 的线性组合。考虑特征多项式 $l(\lambda_i)$:

$$l(\lambda_i) = \lambda_i^n + a_1\lambda_i^{n-1} + a_2\lambda_i^{n-2} + \cdots + a_{n-1}\lambda_i + a_n = \sum_{k=0}^{n} a_k\lambda_i^{n-k} = \lambda_i^n f(\lambda_i)$$

于是有:

$$f(\lambda_i) = \frac{l(\lambda_i)}{\lambda_i^n}$$

所以 $f^{(b-j)}(\lambda_i)$ 可以由 $l(\lambda_i), l'(\lambda_i), \ldots, l^{(b-j)}(\lambda_i)$ 线性表出。于是:

$$\begin{aligned}&m^b\lambda_i^m + a_1(m-1)^b\lambda_i^{m-1} + \cdots + a_n(m-n)^b\lambda_i^{m-n}\\ =&\lambda_i^m \sum_{j=0}^{b} \binom{b}{j} m^j \sum_{k=0}^{n} a_k(-k)^{b-j}\lambda_i^{-k} = \lambda_i^m \sum_{j=0}^{b} \binom{b}{j} m^j g[l(\lambda_i), l'(\lambda_i), \ldots, l^{(b-j)}(\lambda_i)]\end{aligned}$$

其中 $g$ 是 $l(\lambda_i), l'(\lambda_i), \ldots, l^{(b-j)}(\lambda_i)$ 的线性函数。因为 $\lambda_i$ 是特征方程的 $r_i$ 重根，所以：

$$l(\lambda) = (\lambda - \lambda_i)^{r_i} h(\lambda) = 0$$

求导可知 $l(\lambda_i) = l'(\lambda_i) = \cdots = l^{(b-j)}(\lambda_i) = 0$，所以 $g = 0$，即：

$$m^b \lambda_i^m + a_1(m-1)^b \lambda_i^{m-1} + \cdots + a_n(m-n)^b \lambda_i^{m-n} = 0$$

所以 $\{m^b \lambda_i^m\}$, $b = 0, 1, 2, \ldots, r_i - 1$ 都是 $L(x_m) = 0$ 的解。

若 $\{m^b \lambda_i^m\}$, $b = 0, 1, 2, \ldots, r_i - 1$ 线性相关，则存在不全为 0 的 $c_1, c_2, \ldots, c_{r_i}$ 使得：

$$c_1 \lambda_i^m + c_2 m \lambda_i^m + \cdots + c_{r_i} m^{r_i - 1} \lambda_i^m = \lambda_i^m (c_1 + c_2 m + \cdots + c_{r_i} m^{r_i - 1}) = 0$$

对任意的 $m \in \mathbb{Z}$ 成立。由 $n$ 阶线性差分方程的定义，$a_n \neq 0$，所以 $\lambda_i \neq 0$，于是需要上式中关于整数 $m$ 的多项式恒等于 0，此时应有 $c_1 = c_2 = \cdots = c_{r_i} = 0$，矛盾，所以 $\{m^b \lambda_i^m\}$, $b = 0, 1, 2, \ldots, r_i - 1$ 线性无关。

不同 $\lambda_i$ 之间的线性无关性涉及到广义 Vandermonde 行列式，以后再写

**情形三：**此时特征方程存在一对共轭复根 $a + bi, a - bi$，令：

$$\rho = \sqrt{a^2 + b^2},\ \theta = \arctan\left(\frac{b}{a}\right)$$

则：

$$(a + bi)^m = (\rho e^{i\theta})^m = \rho^m [\cos(m\theta) + i \sin(m\theta)],\ \forall\, m \in \mathbb{Z}$$
$$(a - bi)^m = (\rho e^{-i\theta})^m = \rho^m [\cos(m\theta) - i \sin(m\theta)],\ \forall\, m \in \mathbb{Z}$$

都是 $L(x_m)$ 的解。

综上可得出定理的结论。 □

### 常系数齐次线性差分方程解的收敛性

**推导 3.4.** 设 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 为 $n$ 阶常系数齐次线性差分方程 $L(x_m) = 0$ 的特征方程 $l(\lambda) = 0$ 的解。

***(1)*** **$\lambda_i$ 都在单位圆内：**此时存在 $\alpha$ 使得：

$$\min\{|\lambda_i| : i = 1, 2, \ldots, s\} < \alpha < 1$$

于是 $L(x_m) = 0$ 的任何解 $\{x_m\}$ 满足：

$$|x_m| = \left| \sum_{i=1}^{s} \sum_{j=0}^{r_i - 1} c_{ij} m^j \lambda_i^m \right| \leqslant \sum_{i=1}^{s} \sum_{j=0}^{r_i - 1} |c_{ij} m^j|\, |\lambda_i^m| \leqslant \sum_{i=1}^{s} \sum_{j=0}^{r_i - 1} |c_{ij} m^j| \alpha^m$$

由指数函数 $\alpha^m$ 与幂函数 $m^j$ 的收敛速度比较可得：

$$\lim_{m \to +\infty} |x_m| = 0$$

此时称 $\{x_m\}$ **以负指数收敛到** 0。

***(2)*** $\lambda_i$ **在单位圆上**：设 $\lambda_i = a + bi$，由定理 3.20可知此时 $L(x_m) = 0$ 有解：

$$x_m = \cos(m\theta) + i\sin(m\theta),\ \theta = \arctan\left(\frac{b}{a}\right),\ \forall\, m \in \mathbb{Z}$$

当 $\theta = \dfrac{\pi}{4}$ 时，这个解不收敛，是一个周期解。

***(3)*** $\lambda_i$ **在单位圆外**：显然此时存在发散于 $+\infty$ 的解。

### 常系数非齐次线性差分方程解的一般理论

**Theorem 3.21.** 设 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 是 $n$ 阶常系数齐次线性差分方程 $L(x_m) = 0$ 的特征方程 $l(\lambda) = 0$ 的互异根，其重数分别为 $r_1, r_2, \ldots, r_s$，$n$ 阶常系数非齐次线性差分方程 $L(x_m) = f(m)$ 的一个特解为 $\{y_m\}$，则 $L(x_m) = f(m)$ 的通解可以表示为：

$$x_m = y_m + \sum_{i=1}^{s}\sum_{j=0}^{r_i-1} c_{ij} m^j \lambda_i^m,\ \forall\, m \in \mathbb{Z}$$

其中 $c_{ij}$, $i = 1, 2, \ldots, s$, $j = 0, 1, \ldots, r_i - 1$ 为常数。

*Proof.* 由定理 3.19和定理 3.20可直接得到。 □

## 3.4 ARIMA

### 3.4.1 AR 模型

**Definition 3.19.** 如果 $\{\varepsilon_t\}$ 是 $\mathrm{WN}(0, \sigma^2)$，实数 $a_1, a_2, \ldots, a_p (a_p \neq 0)$ 使得多项式 $A(z) = 0$ 的根都在单位圆外：

$$A(z) = 1 - \sum_{i=1}^{p} a_i z^i \neq 0, \quad \forall\, |z| \leqslant 1$$

则称 $p$ 阶常系数线性差分方程：

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t$$

为 $p$ **阶自回归模型**，简记为 $\mathrm{AR}(p)$ 模型。称满足 $\mathrm{AR}(p)$ 模型的平稳时间序列 $\{X_t\}$ 为 $\mathrm{AR}(p)$ 序列，称 $a_1, a_2, \ldots, a_p$ 为 $\mathrm{AR}(p)$ 模型的**自回归系数**，多项式 $A(z) = 0$ 的根都在单位圆外这一条件被称为**稳定性条件**，分别称 $A(z)$ 和 $A(\mathcal{B})$ 为 $\mathrm{AR}(p)$ 模型的**特征多项式**和**自回归系数多项式**。可以用 $A(\mathcal{B})$ 将模型改写为 $A(\mathcal{B})X_t = \varepsilon_t$。

**推导 3.5.** $A(z) = 0$ 的解为 $p$ 阶常系数线性差分方程特征方程：

$$\lambda^p - a_1\lambda^{p-1} - \cdots - a_{p-1}\lambda - a_p = 0$$

的解的倒数。取特征方程的任一根 $\lambda_j$，则：

$$A\left(\frac{1}{\lambda_j}\right)=1-\sum_{i=1}^{p}a_i\frac{1}{\lambda_j^i}=\frac{1}{\lambda_j^p}(\lambda_j^p-a_1\lambda_j^{p-1}-\cdots-a_{p-1}\lambda_j-a_p)=0$$

所以上述稳定性条件即为要求 $p$ 阶常系数线性差分方程特征方程的根都在单位圆内。

**Theorem 3.22.** 设 AR$(p)$ 模型的特征多项式 $A(z)=0$ 有 $s$ 个互异根 $z_1,z_2,\ldots,z_s$，根的重数分别为 $r_1,r_2,\ldots,r_s$，$1<\rho<\min\limits_i\{z_i\}$，则：

1. AR$(p)$ 模型的唯一平稳解是：

$$X_t=\sum_{i=0}^{+\infty}\psi_i\varepsilon_{t-i}$$

   其中 $\psi_i$ 为 $A^{-1}(z)$ 在 $\{z:|z|\leqslant\rho\}$ 内展开的幂级数的系数，它具有递推公式：

$$\psi_0=1,\quad \psi_j=\sum_{i=1}^{p}a_i\psi_{j-i},\ \forall\ j\geqslant 1$$

2. AR$(p)$ 模型的通解为：

$$\sum_{i=0}^{+\infty}\psi_i\varepsilon_{t-i}+\sum_{i=1}^{s}\sum_{j=0}^{r_i-1}c_{ij}t^jz_i^{-t}$$

   其中 $c_{ij}$ 为任意常数；

3. AR$(p)$ 模型的任一解都以负指数阶的速度收敛到平稳解，$\min\limits_i z_i$ 越大，收敛越快；

*Proof.* (1) 由复变函数的知识，$A^{-1}(z)$ 在 $\{z:|z|\leqslant\rho\}$ 内解析，即 $A^{-1}(z)$ 有如下展开：

$$A^{-1}(z)=\sum_{i=0}^{+\infty}\psi_iz^i,\quad |z|\leqslant\rho$$

且这个级数是绝对收敛的。由绝对收敛性可知 $|\psi_iz^i|\to 0$，即 $\psi_i=o(\rho^{-i})$，所以 $\{\psi_i\}\in l_1$，由定理 3.8可知 $\{X_t\}$ 是平稳序列。

令 $a_0=-1$，对 $k<0$ 定义 $\psi_k=0$。注意到：

$$\begin{aligned}A(\mathcal{B})X_t&=\left(1-\sum_{i=1}^{p}a_i\mathcal{B}^i\right)X_t=\left(-a_0-\sum_{i=1}^{p}a_i\mathcal{B}^i\right)X_t=-\sum_{i=0}^{p}a_i\mathcal{B}^iX_t\\&=-\sum_{i=0}^{p}a_iX_{t-i}=-\sum_{i=0}^{p}a_i\sum_{j=0}^{+\infty}\psi_j\varepsilon_{t-i-j}=-\sum_{j=0}^{+\infty}\sum_{i=0}^{p}a_i\psi_j\varepsilon_{t-i-j}\\&=-\sum_{j=0}^{+\infty}\sum_{i=0}^{p}a_i\psi_{j-i}\varepsilon_{t-j}=-\sum_{j=0}^{+\infty}\left(\sum_{i=0}^{p}a_i\psi_{j-i}\right)\varepsilon_{t-j}\end{aligned}$$

因为 $\{\psi_i\}\in l_1$，所以对 $|z|\leqslant 1$ 级数 $\sum\limits_{j=0}^{+\infty}\psi_jz^j$ 是良定义的。由：

$$\begin{aligned}1=A(z)A^{-1}(z)&=\left(1-\sum_{i=1}^{p}a_iz^i\right)A^{-1}(z)=-\sum_{i=0}^{p}a_iz^i\sum_{j=0}^{+\infty}\psi_jz^j\\&=-\sum_{i=0}^{p}\sum_{j=0}^{+\infty}a_iz^i\psi_jz^j=-\sum_{j=0}^{+\infty}\sum_{i=0}^{p}a_i\psi_jz^{i+j}=-\sum_{j=0}^{+\infty}\left(\sum_{i=0}^{p}a_i\psi_{j-i}\right)z^j\end{aligned}$$

对比系数可得（递推公式）：

$$-\sum_{i=0}^{p} a_i\psi_{-i} = 1, \quad -\sum_{i=0}^{p} a_i\psi_{j-i} = 0,\ \forall\, j \geqslant 1$$

于是：

$$A(\mathcal{B})X_t = -\sum_{j=0}^{+\infty}\left(\sum_{i=0}^{p} a_i\psi_{j-i}\right)\varepsilon_{t-j} = -\sum_{i=0}^{p} a_i\psi_{-i}\varepsilon_t - \sum_{j=1}^{+\infty}\left(\sum_{i=0}^{p} a_i\psi_{j-i}\right)\varepsilon_{t-j} = \varepsilon_t$$

所以 $\{X_t\}$ 是解。

设还有另一平稳解 $\{Y_t\}$，即 $A(\mathcal{B})Y_t = \varepsilon_t$ 且 $A^{-1}(\mathcal{B})$ 存在，则：

$$Y_t = A^{-1}(\mathcal{B})\varepsilon_t = A^{-1}(\mathcal{B})A(\mathcal{B})X_t = X_t$$

综上，$\{X_t\}$ 是 AR($p$) 模型唯一的平稳解。

(2) 由 (1)、定理 3.19与定理 3.20即可得通解为：

$$X_t + \sum_{i=1}^{s}\sum_{j=0}^{r_i-1} c_{ij}t^j z_i^t = \sum_{i=0}^{+\infty}\psi_i\varepsilon_{t-i} + \sum_{i=1}^{s}\sum_{j=0}^{r_i-1} c_{ij}t^j z_i^{-t}$$

(3) 由 (1)(2) 可得对于 AR($p$) 模型的任一解 $\{Y_t\}$ 有：

$$|X_t - Y_t| = \left|\sum_{i=1}^{s}\sum_{j=0}^{r_i-1} c_{ij}t^j z_i^{-t}\right| \leqslant O\left[\left(\min_i\{z_i\}\right)^{-t}\right]$$

□

**note 3.2.** 上述定理给了我们一个产生 AR($p$) 序列的方式。先任意选择 $p$ 个初始值，然后根据自回归系数产生序列 $\{Y_t\}$。因为任意的 $\{Y_t\}$ 都以负指数阶的速度收敛到平稳解，取一个较大的 $m$ 然后令 $X_t = Y_{m+t}$ 即可得到近似的 AR($p$) 序列 $\{X_t\}$。

**Property 3.4.1.** AR($p$) 序列 $\{X_t\}$ 具有如下性质：

1. 对任意的 $i \geqslant 1$ 且 $i \in \mathbb{N}^+$，$X_t$ 与 $\varepsilon_{t+i}$ 不相关；
2. $\mathrm{Var}(X_t) = \sigma^2 \sum\limits_{i=0}^{+\infty}\psi_i^2$；
3. *(Yule-Walker 方程)* $\{X_t\}$ 的自协方差函数满足：

$$\boldsymbol{\gamma}_n = \Gamma_n\alpha,\ \gamma(0) = \boldsymbol{\gamma}_n^T\alpha + \sigma^2, \quad n \geqslant p$$

$$\boldsymbol{\gamma}_n = \begin{pmatrix}\gamma(1)\\ \gamma(2)\\ \vdots\\ \gamma(n)\end{pmatrix},\ \Gamma_n = \begin{pmatrix}\gamma(0) & \gamma(1) & \cdots & \gamma(n-1)\\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2)\\ \vdots & \vdots & \ddots & \vdots\\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0)\end{pmatrix}$$

$$\alpha = (a_1, a_2, \ldots, a_p, 0, 0, \ldots, 0)^T$$

4. AR($p$) 序列的自协方差函数和自相关函数满足和 AR($p$) 模型相对应的常系数齐次线性差分方程：

$$
\begin{aligned}
\gamma(n) &= a_1\gamma(n-1) + a_2\gamma(n-2)\cdots + a_p\gamma(n-p) \\
\rho(n) &= a_1\rho(n-1) + a_2\rho(n-2)\cdots + a_p\rho(n-p)
\end{aligned}
$$

5. $\{X_t\}$ 的自协方差函数与自相关函数具有拖尾性，即 $\gamma(n)$ 和 $\rho(n)$ 始终不为 0，且二者的模随着 $n$ 的增大呈指数衰减；

*Proof.* (1) 因为 $\mathrm{E}(\varepsilon_{t+i}) = 0$，由性质 1.4.2(9) 可知 $\varepsilon_{t+i}$ a.e. 有限，由定理 3.8可知 $X_t$ a.e. 有限，于是 $X_t\varepsilon_{t+i}$ 良定义。根据性质 1.4.2(4)(6) 可知：

$$
\begin{aligned}
\mathrm{E}\left(\sum_{j=0}^{+\infty}|\psi_j\varepsilon_{t-j}\varepsilon_{t+i}|\right) &= \lim_{n\to+\infty}\mathrm{E}\left(\sum_{j=0}^{n}|\psi_j\varepsilon_{t-j}\varepsilon_{t+i}|\right) = \lim_{n\to+\infty}\left[\sum_{j=0}^{n}\psi_j\,\mathrm{E}(|\varepsilon_{t-j}\varepsilon_{t+i}|)\right] \\
&= \sum_{j=0}^{+\infty}|\psi_j|\,\mathrm{E}(|\varepsilon_{t-j}\varepsilon_{t+i}|)
\end{aligned}
$$

由**??**可知：

$$
\mathrm{E}(|\varepsilon_{t-j}\varepsilon_{t+i}|) = \Big|\mathrm{E}(|\varepsilon_{t-j}\varepsilon_{t+i}|)\Big| \leqslant \sqrt{\mathrm{E}(\varepsilon_{t-j}^2)\,\mathrm{E}(\varepsilon_{t+i}^2)} = \sigma^2
$$

因为 $\{\psi_i\} \in l^2$，所以：

$$
\mathrm{E}\left(\sum_{j=0}^{+\infty}|\psi_j\varepsilon_{t-j}\varepsilon_{t+i}|\right) \leqslant \sigma^2\sum_{j=0}^{+\infty}|\psi_j| < +\infty
$$

取控制函数 $\sum\limits_{j=0}^{+\infty}|\psi_j\varepsilon_{t-j}\varepsilon_{t+i}|$，由定理 1.47和性质 1.4.2(6) 可得：

$$
\begin{aligned}
\mathrm{E}(X_t\varepsilon_{t+i}) &= \mathrm{E}\left(\sum_{j=0}^{+\infty}\psi_j\varepsilon_{t-j}\varepsilon_{t+i}\right) = \mathrm{E}\left[\lim_{n\to+\infty}\left(\sum_{j=0}^{n}\psi_j\varepsilon_{t-j}\varepsilon_{t+i}\right)\right] \\
&= \lim_{n\to+\infty}\mathrm{E}\left(\sum_{j=0}^{n}\psi_j\varepsilon_{t-j}\varepsilon_{t+i}\right) = \lim_{n\to+\infty}\left[\sum_{j=0}^{n}\psi_j\,\mathrm{E}(\varepsilon_{t-j}\varepsilon_{t+i})\right] = 0
\end{aligned}
$$

(2) 由定理 3.22(1) 立即可得。

(3) 由 AR($p$) 模型的定义可得：

$$
\begin{pmatrix} X_t \\ X_{t+1} \\ \vdots \\ X_{t+n-1} \end{pmatrix} = \begin{pmatrix} X_{t-1} & X_{t-2} & \cdots & X_{t-n} \\ X_t & X_{t-1} & \cdots & X_{t-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t+n-2} & X_{t+n-3} & \cdots & X_{t-1} \end{pmatrix}\alpha + \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \\ \vdots \\ \varepsilon_{t+n-1} \end{pmatrix}
$$

于是：

$$X_{t-1}\begin{pmatrix} X_t \\ X_{t+1} \\ \vdots \\ X_{t+n-1} \end{pmatrix} = X_{t-1}\begin{pmatrix} X_{t-1} & X_{t-2} & \cdots & X_{t-n} \\ X_t & X_{t-1} & \cdots & X_{t-n+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t+n-2} & X_{t+n-3} & \cdots & X_{t-1} \end{pmatrix}\alpha_n + X_{t-1}\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \\ \vdots \\ \varepsilon_{t+n-1} \end{pmatrix}$$

$$\boldsymbol{\gamma}_n = \Gamma_n \alpha, \quad n \geqslant p$$

第二行是对第一行取期望的结果。对于 $\gamma(0)$，由 (1) 可得：

$$\begin{aligned}\gamma(0) = \mathrm{E}(X_t^2) = \mathrm{E}\left(\sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t\right)^2 &= \mathrm{E}\left(\sum_{i=1}^{p} a_i X_{t-i}\right)^2 + \mathrm{E}(\varepsilon_t)^2 \\ = \alpha^T \Gamma_n \alpha + \sigma^2 = \alpha^T \boldsymbol{\gamma}_n + \sigma^2, \quad n \geqslant p\end{aligned}$$

(4) 由 (3) 和定理 3.2立即可得。

(5) 由定理 3.20和 (4) 可知 $\gamma(n)$ 和 $\rho(n)$ 的通解都具有如下形式：

$$\sum_{i=1}^{s}\sum_{j=0}^{r_i-1} c_{ij} n^j \lambda_i^n, \ \forall\, n \in \mathbb{N}$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 为对应常系数齐次线性差分方程的特征方程的根，$c_{ij}$ 为任意常数。若 $c_{ij}$ 全为 0，则 $\gamma(0), \rho(0)$ 都为 0，所以它们不可能全为 0. 由 AR($p$) 序列的定义，$|\lambda_i| < 1$，因为 $a_p \neq 0$，所以 $\lambda_i \neq 0$，由此可得 $\gamma(n)$ 和 $\rho(n)$ 始终不为 0。由指数函数与幂函数的收敛速度比较可知 $\gamma(n)$ 和 $\rho(n)$ 的模随着 $n$ 的增大将以指数阶的速度减小。 □

### 3.4.2 MA 模型

**Definition 3.20.** 设 $\{\varepsilon_t\}$ 是 $\mathrm{WN}(0, \sigma^2)$，实数 $b_1, b_2, \ldots, b_q (b_q \neq 0)$ 使得多项式 $B(z) = 0$ 的根都在单位圆内：

$$B(z) = 1 + \sum_{i=1}^{q} b_i z^i \neq 0, \quad \forall\, |z| < 1$$

则称：

$$X_t = \varepsilon_t + \sum_{i=1}^{q} b_i \varepsilon_{t-i}, \quad t \in \mathbb{Z}$$

为 $q$ **阶滑动平均模型**，简记为 MA($q$) 模型。称满足 MA($q$) 模型的平稳时间序列 $\{X_t\}$ 为 MA($q$) 序列。若 $B(z) \neq 0$ 对 $|z| \leqslant 1$ 成立，则称对应的 MA($q$) 模型为**可逆的** MA($q$) **模型**，相应的 MA($q$) 序列为**可逆的** MA($q$) **序列**。