Calculus
微积分

Wentao Zhu
Queens' College, University of Cambridge

June 2025

目录

0 Preliminaries 预备知识

0.1 Trigonometry 三角学

0.1.1 Square Identities 平方恒等式

$$\tan^2\theta + 1 \equiv \sec^2\theta, \quad \cot^2\theta + 1 \equiv \csc^2\theta$$

0.1.2 Compound Angle Formulae 两角和差公式

$$\sin(A+B) \equiv \sin A\cos B + \cos A\sin B$$

$$\sin(A-B) \equiv \sin A\cos B - \cos A\sin B$$

$$\cos(A+B) \equiv \cos A\cos B - \sin A\sin B$$

$$\cos(A-B) \equiv \cos A\cos B + \sin A\sin B$$

$$\tan(A+B) \equiv \frac{\tan A + \tan B}{1 - \tan A\tan B}, \quad \tan(A-B) \equiv \frac{\tan A - \tan B}{1 + \tan A\tan B}$$

0.1.3 Double Angel Formulae 倍角公式

$$\sin 2A \equiv 2\sin A\cos A$$

$$\cos 2A \equiv \cos^2 A - \sin^2 A \equiv 2\cos^2 A - 1 \equiv 1 - 2\sin^2 A$$

$$\tan 2A \equiv \frac{2\tan A}{1 - \tan^2 A}$$

0.1.4 Half Angel Formulae 半角公式

$$\sin\frac{A}{2} \equiv \pm\sqrt{\frac{1-\cos A}{2}}, \quad \cos\frac{A}{2} \equiv \pm\sqrt{\frac{1+\cos A}{2}}$$

$$\tan\frac{A}{2} \equiv \frac{\sin A}{1+\cos A} \equiv \frac{1-\cos A}{\sin A}$$

0.1.5 Change Plus/Minus into Product Formulae 和差化积公式

$$\sin A + \sin B \equiv 2\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\sin A - \sin B \equiv 2\cos\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

$$\cos A + \cos B \equiv 2\cos\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\cos A - \cos B \equiv -2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$$

0.1.6 Change Product into Plus/Minus Formulae 积化和差公式

$$\sin A\cos B \equiv \frac{1}{2}[\sin(A+B)+\sin(A-B)]$$
$$\cos A\sin B \equiv \frac{1}{2}[\sin(A+B)-\sin(A-B)]$$
$$\cos A\cos B \equiv \frac{1}{2}[\cos(A+B)+\cos(A-B)]$$
$$\sin A\sin B \equiv -\frac{1}{2}[\cos(A+B)-\cos(A-B)]$$

0.1.7 Induction Formulae 诱导公式

设 $k \in \mathbb{Z}$，有：

$$\sin\left(\frac{k\pi}{2} \pm \alpha\right), \quad \cos\left(\frac{k\pi}{2} \pm \alpha\right)$$

- 若 k 为偶数，函数名不变；若 k 为奇数，正弦变余弦，余弦变正弦。
- 符号由将 α 视为锐角时原函数在对应象限的符号决定。（“奇变偶不变，符号看象限”）

常用特例：

$$\sin(\pi-\alpha)=\sin\alpha, \quad \sin(\pi+\alpha)=-\sin\alpha$$
$$\cos(\pi-\alpha)=-\cos\alpha, \quad \cos(\pi+\alpha)=-\cos\alpha$$
$$\sin\left(\frac{\pi}{2}-\alpha\right)=\cos\alpha, \quad \cos\left(\frac{\pi}{2}-\alpha\right)=\sin\alpha$$

例题 0.1 证明下列恒等式：

(a) $\tan 3x \equiv \dfrac{3\tan x - \tan^3 x}{1-3\tan^2 x}$

(b) $\dfrac{2\sin(x-y)}{\cos(x-y)-\cos(x+y)} \equiv \cot y - \cot x$

(c) $\tan A + \cot A \equiv 2\operatorname{cosec} 2A$

(d) $\tan 2A - \tan A \equiv \tan A \sec 2A$

(e) $\cot 2A + \tan A \equiv \dfrac{1}{2}\operatorname{cosec} A \sec A$

(f) $\sec^2 A \operatorname{cosec}^2 A \equiv 4\operatorname{cosec}^2 2A$

(g) $1-\tan^2 A \equiv \cos 2A \sec^2 A$

(h) $(\cos A + 3\sin A)^2 \equiv 5 - 4\cos 2A + 3\sin 2A$

(i) $\operatorname{cosec} 2A + \cot 2A \equiv \cot A$

(j) $\sin(A+B)\sin(A-B) \equiv \sin^2 A - \sin^2 B$

例题 0.2 证明下列恒等式：

(a) $\dfrac{\sin A}{\cos B}+\dfrac{\cos A}{\sin B}\equiv\dfrac{2\cos(A-B)}{\sin 2B}$

(b) $\dfrac{\cos A+\sin A}{\cos A-\sin A}\equiv\sec 2A+\tan 2A$

(c) $\dfrac{1-\tan^2 A}{1+\tan^2 A}\equiv\cos 2A$

(d) $\dfrac{\cos 2A+\sin 2A-1}{\cos 2A-\sin 2A+1}\equiv\tan A$

(e) $\dfrac{\sin 3A+\sin A}{2\sin 2A}\equiv\cos A$

(f) $\dfrac{\cos 3A-\sin 3A}{1-2\sin 2A}\equiv\cos A+\sin A$

(g) $\dfrac{\cos 2A+9\cos A+5}{4+\cos A}\equiv 2\cos A+1$

(h) $\dfrac{\cos^3 A-\sin^3 A}{\cos A-\sin A}\equiv\dfrac{2+\sin 2A}{2}$

(i) $8\sin^2 x\cos^2 x\equiv 1-\cos 4x$

(j) $(2\sin A+\cos A)^2\equiv\dfrac{1}{2}(4\sin 2A-3\cos 2A+5)$

例题 0.3 计算下列三角函数的值：*(1)* $\sin(150^\circ)$ *(2)* $\cos(225^\circ)$ *(3)* $\tan(300^\circ)$ *(4)* $\csc(\frac{5\pi}{3})$

(1) $\sin(150^\circ)=\sin(180^\circ-30^\circ)=\sin 30^\circ=\frac{1}{2}$
(2) $\cos(225^\circ)=\cos(180^\circ+45^\circ)=-\cos 45^\circ=-\frac{\sqrt{2}}{2}$
(3) $\tan(300^\circ)=\tan(360^\circ-60^\circ)=-\tan 60^\circ=-\sqrt{3}$
(4) $\csc(\frac{5\pi}{3})=\csc(2\pi-\frac{\pi}{3})=-\csc(\frac{\pi}{3})=-\frac{2}{\sqrt{3}}=-\frac{2\sqrt{3}}{3}$

例题 0.4 化简下列表达式：$\dfrac{\sin(2\pi-\alpha)\cdot\cos(\pi+\alpha)}{\tan(\pi-\alpha)\cdot\cot(\frac{\pi}{2}+\alpha)}$

$$
\begin{aligned}
&\frac{\sin(2\pi-\alpha)\cdot\cos(\pi+\alpha)}{\tan(\pi-\alpha)\cdot\cot(\frac{\pi}{2}+\alpha)}\\
=&\frac{(-\sin\alpha)\cdot(-\cos\alpha)}{(-\tan\alpha)\cdot(-\tan\alpha)} \quad \text{应用诱导公式}\\
=&\frac{\sin\alpha\cos\alpha}{\tan^2\alpha}\\
=&\frac{\sin\alpha\cos\alpha}{(\frac{\sin\alpha}{\cos\alpha})^2}\\
=&\frac{\sin\alpha\cos\alpha}{\frac{\sin^2\alpha}{\cos^2\alpha}}\\
=&\sin\alpha\cos\alpha\cdot\frac{\cos^2\alpha}{\sin^2\alpha}\\
=&\frac{\cos^3\alpha}{\sin\alpha}\\
=&\cot\alpha\cdot\cos^2\alpha
\end{aligned}
$$

例题 0.5 已知 $\sin(\pi+\alpha)=\frac{2}{3}$，且 α 在第三象限，求 $\cos(\frac{\pi}{2}-\alpha)$ 的值。

$$
\begin{aligned}
&\text{已知：}\ \sin(\pi+\alpha)=\frac{2}{3}\\
&\Rightarrow \sin\alpha=\frac{2}{3}\\
&\cos(\frac{\pi}{2}-\alpha)=-\sin\alpha=-\frac{2}{3}
\end{aligned}
$$

0.2 Hyperbolic Functions 双曲函数

0.2.1 Exponential Forms of Hyperbolic Functions 双曲函数的指数形式

1. **双曲正弦 (Hyperbolic Sine)**
   - **定义：** $\sinh x=\dfrac{e^x-e^{-x}}{2}$
   - **定义域:** $(-\infty,\infty)$
   - **值域:** $(-\infty,\infty)$
2. **双曲余弦 (Hyperbolic Cosine)**
   - **定义：** $\cosh x=\dfrac{e^x+e^{-x}}{2}$
   - **定义域:** $(-\infty,\infty)$
   - **值域:** $[1,\infty)$
3. **双曲正切 (Hyperbolic Tangent)**
   - **定义：** $\tanh x=\dfrac{\sinh x}{\cosh x}=\dfrac{e^x-e^{-x}}{e^x+e^{-x}}$
   - **定义域:** $(-\infty,\infty)$
   - **值域:** $(-1,1)$
4. **双曲余切 (Hyperbolic Cotangent)**
   - **定义：** $\coth x=\dfrac{\cosh x}{\sinh x}=\dfrac{e^x+e^{-x}}{e^x-e^{-x}}$
   - **定义域:** $(-\infty,0)\cup(0,\infty)$
   - **值域:** $(-\infty,-1)\cup(1,\infty)$
5. **双曲正割 (Hyperbolic Secant)**
   - **定义：** $\operatorname{sech} x=\dfrac{1}{\cosh x}=\dfrac{2}{e^x+e^{-x}}$
   - **定义域:** $(-\infty,\infty)$
   - **值域:** $(0,1]$
6. **双曲余割 (Hyperbolic Cosecant)**
   - **定义：** $\operatorname{csch} x=\dfrac{1}{\sinh x}=\dfrac{2}{e^x-e^{-x}}$
   - **定义域:** $(-\infty,0)\cup(0,\infty)$
   - **值域:** $(-\infty,0)\cup(0,\infty)$

0.2.2 Hyperbolic Identities 双曲恒等式

奥斯本法则 (Osborne's Rule)：要从三角恒等式得到对应的双曲恒等式，可将 cos 换为 cosh，并将所有包含两个 sinh 因子的乘积项的符号改变。

下面举几个例子：

1. **平方和公式**

$$\begin{aligned}&\text{三角恒等式：} && \cos^2 x + \sin^2 x = 1\\ &\text{双曲恒等式：} && \cosh^2 x - \sinh^2 x = 1\end{aligned}$$

2. **和角公式（正弦）**

$$\begin{aligned}&\text{三角恒等式：} && \sin(A+B) = \sin A\cos B + \cos A\sin B\\ &\text{双曲恒等式：} && \sinh(A+B) = \sinh A\cosh B + \cosh A\sinh B\end{aligned}$$

3. **和角公式（余弦）**

$$\begin{aligned}&\text{三角恒等式：} && \cos(A+B) = \cos A\cos B - \sin A\sin B\\ &\text{双曲恒等式：} && \cosh(A+B) = \cosh A\cosh B + \sinh A\sinh B\end{aligned}$$

4. **二倍角公式（余弦）**

$$\begin{aligned}&\text{三角恒等式：} && \cos 2A = \cos^2 A - \sin^2 A\\ &\text{双曲恒等式：} && \cosh 2A = \cosh^2 A + \sinh^2 A\end{aligned}$$

0.2.3 Logarithmic Form of Inverse Hyperbolic Functions 反双曲函数的对数形式

1. **反双曲正弦**
   - **定义：** $\sinh^{-1} x = \ln\left(x + \sqrt{x^2+1}\right)$
   - **定义域：** $(-\infty, \infty)$
   - **值域：** $(-\infty, \infty)$
2. **反双曲余弦**
   - **定义：** $\cosh^{-1} x = \ln\left(x + \sqrt{x^2-1}\right)$
   - **定义域：** $[1, \infty)$
   - **值域：** $[0, \infty)$
3. **反双曲正切**
   - **定义：** $\tanh^{-1} x = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)$
   - **定义域：** $(-1, 1)$
   - **值域：** $(-\infty, \infty)$
4. **反双曲余切**

- **定义：** $\coth^{-1} x = \frac{1}{2}\ln\left(\frac{x+1}{x-1}\right)$
- **定义域：** $(-\infty,-1)\cup(1,\infty)$
- **值域：** $(-\infty,0)\cup(0,\infty)$

5. **反双曲正割**

- **定义：** $\operatorname{sech}^{-1} x = \ln\left(\frac{1+\sqrt{1-x^2}}{x}\right)$
- **定义域：** $(0,1]$
- **值域：** $[0,\infty)$

6. **反双曲余割**

- **定义：** $\operatorname{csch}^{-1} x = \ln\left(\frac{1}{x}+\frac{\sqrt{1+x^2}}{|x|}\right)$
- **定义域：** $(-\infty,0)\cup(0,\infty)$
- **值域：** $(-\infty,0)\cup(0,\infty)$

0.3 Complex Numbers 复数域

0.3.1 Complex Numbers and Complex Conjugates 复数和复共轭

在实数范围内，方程 $x^2=-1$ 无解，因为任何实数的平方均为非负。为克服此限制，我们引入**虚数单位** i，其定义为：

$$i^2=-1$$

由此，$\sqrt{-1}=i$，使得所有负实数都可以开平方。

将实数与虚数单位相结合，便构成了**复数**。任一复数 z 均可表示为：

$$z=a+bi$$

其中 $a,b\in\mathbb{R}$。a 称为**实部**，记作 $\mathrm{Re}(z)$；b 称为**虚部**，记作 $\mathrm{Im}(z)$。

对于复数 $z=a+bi$，我们定义其**复共轭**为：

$$\bar{z}=a-bi$$

复共轭改变了虚部的符号。它在复数的除法、求模以及简化表达式等运算中起着至关重要的作用。

0.3.2 Operations and Properties of Complex Numbers 复数的运算与性质

设 $z_1=a+bi$，$z_2=c+di$。

- **加法/减法**：

$$z_1\pm z_2=(a\pm c)+(b\pm d)i$$

- **乘法**：

$$z_1\cdot z_2=(a+bi)(c+di)=ac+adi+bci+bdi^2=(ac-bd)+(ad+bc)i$$

- **除法**：利用复共轭将分母实数化。

$$\frac{z_1}{z_2}=\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2}$$

设 $z=a+bi$，其共轭为 $\bar{z}=a-bi$，模定义为 $|z|=\sqrt{a^2+b^2}$。
- **基本关系**：

$$z\cdot\bar{z}=(a+bi)(a-bi)=a^2+b^2=|z|^2$$

- **运算性质**：

$$\overline{z_1\pm z_2}=\overline{z_1}\pm\overline{z_2}$$
$$\overline{z_1\cdot z_2}=\overline{z_1}\cdot\overline{z_2}$$
$$\overline{\left(\frac{z_1}{z_2}\right)}=\frac{\overline{z_1}}{\overline{z_2}}$$
$$|z_1\cdot z_2|=|z_1|\cdot|z_2|$$
$$\left|\frac{z_1}{z_2}\right|=\frac{|z_1|}{|z_2|}$$

i 的幂遵循一个周期为 4 的循环：

$$\begin{aligned} i^1&=i, & i^2&=-1,\\ i^3&=-i, & i^4&=1,\\ i^{4k+n}&=i^n \qquad (k\in\mathbb{Z},n=0,1,2,3) \end{aligned}$$

这一性质极大地简化了 i 的高次幂计算。

0.3.3 Geometric Representation of Complex Numbers 复数的几何表示

为了直观地表示复数，我们引入**复平面**（complex plane），又称**阿甘图**（Argand diagram）。它是一个直角坐标系：
- **横轴（实轴）**：表示复数的实部 $\mathrm{Re}(z)$。
- **纵轴（虚轴）**：表示复数的虚部 $\mathrm{Im}(z)$。
复数 $z=a+bi$ 与复平面上的点 (a,b) 一一对应。
在复平面上，点 $z=a+bi$ 也可以用其到原点的距离和与正实轴的夹角来表示。
- **模 (Modulus)**：向量 z 的长度，记为 $r=|z|=\sqrt{a^2+b^2}$。
- **辐角 (Argument)**：向量 z 与正实轴所成的角，记为 $\theta=\arg(z)$。满足 $\tan\theta=\frac{b}{a}$，且通常取主值 $\theta\in(-\pi,\pi]$。

由此得到复数的**极坐标形式**：

$$z=r(\cos\theta+i\sin\theta)$$

设 $z_1=r_1(\cos\theta_1+i\sin\theta_1)$，$z_2=r_2(\cos\theta_2+i\sin\theta_2)$。根据三角恒等式，它们的乘积为：

$$\begin{aligned} z_1z_2&=r_1r_2[(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+i(\cos\theta_1\sin\theta_2+\sin\theta_1\cos\theta_2)]\\ &=r_1r_2[\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)] \end{aligned}$$

结论：复数相乘，**模相乘**，**辐角相加**。

$$|z_1z_2|=|z_1|\cdot|z_2|,\quad \arg(z_1z_2)=\arg(z_1)+\arg(z_2)$$

- 用复数 z_2 乘以 z_1，其几何效果是：
  1. **缩放**：将点 z_1 到原点的距离（模）拉伸或压缩 $|z_2|$ 倍。
  2. **旋转**：将点 z_1 绕原点旋转一个角度 $\arg(z_2)$。

- **结论**：乘法 $z_1 z_2$ 是一个**复合变换**：先进行由 z_2 的模决定的**缩放**，再进行由 z_2 的辐角决定的**旋转**。
- **特例**：当 $|z_2| = 1$ 时，乘法是**纯旋转**。例如，乘以 i ($\arg(i) = \pi/2$) 相当于逆时针旋转 $90°$。

设 $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$，$z_2 = r_2(\cos\theta_2 + i\sin\theta_2) \neq 0$。它们的商为：

$$\begin{aligned}\frac{z_1}{z_2} &= \frac{r_1}{r_2}\left[\frac{\cos\theta_1 + i\sin\theta_1}{\cos\theta_2 + i\sin\theta_2}\right] \\ &= \frac{r_1}{r_2}\left[\frac{(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 - i\sin\theta_2)}{\cos^2\theta_2 + \sin^2\theta_2}\right] \\ &= \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]\end{aligned}$$

结论：复数相除，**模相除**，**辐角相减**。

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \quad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)$$

- 用复数 z_2 去除 z_1，其几何效果是：
  1. **缩放**：将点 z_1 到原点的距离缩放 $1/|z_2|$ 倍。
  2. **旋转**：将点 z_1 绕原点旋转一个角度 $-\arg(z_2)$（即顺时针旋转 $\arg(z_2)$）。
- **结论**：除法 z_1/z_2 是乘法的逆操作，对应着**反向的缩放和旋转**。

下面我们总结一些辐角的运算性质，这些性质是上述几何意义的直接推论。

- $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$：连续两次旋转，总旋转角等于两次旋转角之和。
- $\arg(z_1/z_2) = \arg(z_1) - \arg(z_2)$：旋转的抵消，结果是净旋转角。
- $\arg(\bar{z}) = -\arg(z)$：共轭运算关于实轴对称，相当于关于实轴做镜像，旋转角自然取反。
- $\arg(z^n) = n\arg(z)$：连续进行 n 次相同的旋转。

在复平面上，**乘法/除法实现了旋转与缩放的统一处理**。这正是极坐标形式强大之处，也是后续引入指数形式 $re^{i\theta}$ 的天然动机，因为复数指数函数完美地封装了旋转（$e^{i\theta}$）和缩放（r）这两种基本操作。

0.3.4 Euler's Formula and Euler's Identity 欧拉公式与欧拉恒等式

欧拉公式（Euler's Formula）建立了复数指数函数与三角函数之间的桥梁，是复分析中的基石：

$$e^{i\theta} = \cos\theta + i\sin\theta$$

理解欧拉公式的关键在于跳出实数指数的思维定式，在复平面的几何视角下，将其看作一种旋转运动。

对于实数 a，函数 $f(t) = e^{at}$ 描述了一种**增长**（$a > 0$）或**衰减**（$a < 0$）。当指数变为纯虚数 it 时，函数 $f(t) = e^{it}$ 描述的则是一种**匀速圆周运动**。它在复平面上的轨迹是单位圆。

欧拉公式告诉我们，复数指数函数 e^z 是一个更宏大的统一体，它同时包含了“伸缩”和“旋转”这两种基本运动。对于一个一般的复数 $z = x + iy$，有

$$e^z = e^{x+iy} = e^x \cdot e^{iy}$$

e^x 负责伸缩，e^{iy} 负责旋转。

作为一个特殊的例子，当 $\theta = \pi$ 时，欧拉公式变为：

$$e^{i\pi} = \cos\pi + i\sin\pi = -1 + i\cdot 0 = -1$$

移项后得到著名的**欧拉恒等式**（Euler's Identity）：

$$e^{i\pi} + 1 = 0$$

这个等式将数学中 5 个最重要的常数（$e, i, \pi, 1, 0$）优美地联系在了一起。它描述了一个具体的旋转：从正实轴上的 1 开始，在单位圆上逆时针旋转 π 弧度，正好到达-1。

0.3.5 Exponential Form of Complex Numbers 复数的指数形式

利用欧拉公式，复数的极坐标形式可以简洁地表示为**指数形式**：

$$z = re^{i\theta}$$

其中 $r = |z|$，$\theta = \arg(z)$。

指数形式极大地简化了复数的乘、除和幂运算。

- **乘法**：

$$z_1\cdot z_2 = (r_1e^{i\theta_1})(r_2e^{i\theta_2}) = r_1r_2e^{i(\theta_1+\theta_2)}$$

几何意义：模相乘，辐角相加。

- **除法**：

$$\frac{z_1}{z_2} = \frac{r_1e^{i\theta_1}}{r_2e^{i\theta_2}} = \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)}$$

- **乘方（棣莫弗公式 DeMoivre's Formula）**：

$$z^n = (re^{i\theta})^n = r^ne^{in\theta} = r^n(\cos n\theta + i\sin n\theta)$$

- **开方**：

$$z^{1/n} = r^{1/n}e^{i(\theta+2k\pi)/n}, \quad k = 0, 1, \dots, n-1$$

一个非零复数的 n 次方根有 n 个不同的值。

例题 0.6 把下面的复数写成 $a + ib$ 的形式：$e^{i\frac{\pi}{2}}, e^{i\frac{3\pi}{2}}, e^{i\frac{3\pi}{4}}, e^{i\frac{11\pi}{3}}, e^{1+i}, e^{-1}$。

例题 0.7 把 $1+\sqrt{3}i$ 表示成指数形式，继而求 $(1+\sqrt{3}i)^{30}$。

1 Logical Framework 逻辑框架

1.1 Terminology and Propositional Logic 术语和命题逻辑

命题和基本的逻辑框架是构建数学体系的基础。数学是一个公理化系统。数学理论从公理（被假定为真的命题）出发，通过逻辑规则推导出新命题（定理）。

定义 1.1 命题（*proposition*）是一个非真即假的数学陈述。

- 首先，命题是一个陈述句，不能是疑问句或者祈使句。
- 其次，它必须具有确定的真值，要么是真，要么是假，不能模棱两可。某些命题可能包含对命题的引用，例如“这个陈述是假的”，这类自指命题可能导致悖论，通常需要在形式系统中谨慎处理。数学推理必须建立在无歧义的真假判断上。如果允许命题的真值不确定，数学证明的链条就会断裂。在多值逻辑或模糊逻辑中，命题的真值可以介于真假之间（如概率真值），但这些体系是经典逻辑的扩展，用于特定领域（如计算机科学、人工智能），并不动摇经典数学的基础。

定义 1.2 定义（*definition*）是对数学术语的精确和明确的描述。

定义 1.3 定理（*theorem*）是一个有效的数学结果。

定义 1.4 证明（*proof*）是解释为什么这个结果是有效的。

定义 1.5 引理（*lemma*）是一种初步结果，它的主要作用是帮助我们证明即将出现的重要定理或一般定理。

定义 1.6 推论（*corollary*）是其证明在很大程度上依赖于先前的重要定理或一般定理的结果。

在数学逻辑中，命题可以根据其结构分为简单命题和复合命题。**简单命题**（simple proposition）或者**原子命题**（atomic proposition）是不能再分解为更小命题的基本陈述句，不包含其他命题作为组成部分。例如：“2 是偶数。”，“6 是素数。”都属于原子命题。相对的，**复合命题**（compound proposition）是由简单命题通过**逻辑连接词**（如“非”、“且”、“或”、“如果... 则...”、“当且仅当”...）组合而成的命题。例如：“2 是偶数且 3 是奇数。”，“若 $x > 2$，则 $x^2 > 4$。”，“三角形等边当且仅当等角。”都属于复合命题，它们都包含逻辑连接词。

逻辑符号	**英文名称**	**连接词示例**
$\neg$	negation	not p
$\wedge$	conjunction	p and q
$\vee$	disjunction	p or q
$\Rightarrow$	implication/conditional	If p, then q
$\Leftrightarrow$	biconditional	p if and only if q

表 1: 逻辑符号及其英文名称和连接词示例

在数学逻辑中，给定一个蕴含命题（即“如果 p，那么 q”，记作 $p \Rightarrow q$），我们可以定义其逆命题、否命题和逆否命题。给定原命题（implication）$p \Rightarrow q$，交换其条件和结论得到**逆命题**（converse）$q \Rightarrow p$，否定其条件和结论得到**否命题**（inverse）$\neg p \Rightarrow \neg q$，既交换又否定原命题的条件和结论得到**逆否命题**（contrapositive）$\neg q \Rightarrow \neg p$。

1.2 Truth Value and Truth Table 真值和真值表

真值（truth value）是指一个命题在逻辑上的真假状态。在经典逻辑中，命题的真值只有两种：

- 真（True, T, 1）
- 假（False, F, 0）

基于命题的真值通过特定的逻辑连接词对命题进行组合或变换的规则叫**逻辑运算**（logical operations）。它是数理逻辑、计算机科学和布尔代数的核心概念。**真值表**（truth table）是描述逻辑运算行为的工具，它系统地列出所有可能的输入组合（命题的真值）及对应的输出结果（复合命题的真值）。

表 2: 真值表

p	q	$\neg p$	$\neg q$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q$	$\neg q \Rightarrow \neg p$
T	T	F	F	T	T	T	T	T	T
T	F	F	T	F	T	F	T	F	F
F	T	T	F	F	T	T	F	F	T
F	F	T	T	F	F	T	T	T	T

- $p \Leftrightarrow q$ 的逻辑运算规律可能有些反直觉，只要两个原子命题都是真命题，即便两者之间没有任何联系，$p \Leftrightarrow q$ 也是一个真命题。可以这么理解，蕴含逻辑代表“如果 p，那么 q”的承诺，如果 q 是真的，那么无所谓假设的条件 p 本身是否为真，不影响 q 是一个真命题，因此 $p \Leftrightarrow q$ 应为真。
- 可以注意到当假设的条件 p 本身为假命题时，无所谓 q 的真值，条件命题恒为真，这种真我们叫“**空真**（vacuous truth）”。
- 逻辑蕴含的真值规则是形式化的约定，它剥离了语义关联，仅保留真值关系。这种定义保证了逻辑系统的简洁性和普适性，尽管有时与直觉不符。
- $p \Rightarrow q$ 的等价表述包括“p 是 q 的**充分**（sufficient）**条件**”，“q 是 p 的**必要**（necessary）**条件**”。类似的，$p \Leftrightarrow q$ 的等价表述包括“p 和 q 互为**充（分必）要条件**”，“p **（逻辑）等价于** q”。

1.3 Logical Equivalence 逻辑等价

逻辑等价（logical equivalence）指的是两个命题在所有可能的情况下（即所有真值组合下）具有完全相同的真值。换句话说，如果两个命题 p 和 q 必然同时为真或者同时为假，那么它们就是逻辑等价的，记作 $p \Leftrightarrow q$ 或者 $p \equiv q$。

- 注意，当我们写下“$p \Leftrightarrow q$”来表示 p 与 q 逻辑等价的时候，我们实际上想表达的意思是“$p \Leftrightarrow q$ 这个命题是真的”，但在之前的语境下“$p \Leftrightarrow q$”似乎只是一个真假未知的命题，因此在表述的时候千万不要自我混淆。换句话说，在对象语言（object-level）中，$\Leftrightarrow$ 是一个二元逻辑连接词；在元语言（meta-level）中，它往往被“滥用”来表示逻辑等价这一更强的概念。

表 3: 常见逻辑等价关系

名称	逻辑等价式
双重否定律	$\neg\neg p \Leftrightarrow p$
交换律	$p \wedge q \Leftrightarrow q \wedge p$
结合律	$(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$
分配律	$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$
德摩根反演律 (de Morgan’s Laws)	$\neg(p \vee q) \Leftrightarrow (\neg p) \wedge (\neg q)$
	$\neg(p \wedge q) \Leftrightarrow (\neg p) \vee (\neg q)$
蕴含的等价形式	$(p \Rightarrow q) \Leftrightarrow (\neg p) \vee q$
逆否命题等价	$(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$
等价条件的含义	$(p \Leftrightarrow q) \Leftrightarrow [(p \Rightarrow q) \wedge (q \Rightarrow p)]$

1.4 First-Order Logic: Predicates and Quantifiers 一阶逻辑：谓词和量词

一阶逻辑，又称**一阶谓词逻辑**，是数理逻辑中最基本且广泛使用的形式系统。它扩展了命题逻辑（Propositional Logic），引入了量词和谓词，能够更精细地描述数学结构和自然语言中的逻辑关系。

谓词（predicate）可以理解为一个不完整的命题，它描述对象的性质或对象之间的关系。一般来说，它本身没有确定的真假值，它的真假值取决于一个或多个变量的取值。例如，“x 是素数”是一个谓词，它还不是一个命题，它的真假值取决于 x 这个变量的取值。我们可以记作 $P(x)$。当我们给这个变量 x 赋值以后，这个陈述就变成了命题，继而它的真值就可以被判断了。例如，$P(3)$ 为真，$P(4)$ 为假。当然你也可以构造一个真值是确定的，不取决于变量的取值的谓词。例如，“如果 x 是偶数，那么 $x+1$ 就是奇数（x 是一个整数）”这个谓词 $P(x)$ 的真值对于所有 x 而言都是 T。类似的，“n 是偶数并且 n 是奇数”则是一个对于所有 n 而言真值都为 F 的谓词。

量词 (quantifier) 用来限定变量的范围，告诉谓词在什么情况下成立。主要有两种：

1. **全称量词**（universal quantifier）$\forall$ 读作“对于所有”、“任意”。
2. **存在量词**（existential quantifier）$\exists$ 读作“存在”、“至少有一个”。

谓词和量词使我们可以描述更复杂的逻辑关系：

1. **全称陈述**（universal statement）$\forall n[P(n)]$
2. **存在陈述**（existential statement）$\exists n[P(n)]$

- 证明一个全称陈述是错的我们只需要找到一个反例（counterexample）。
- 证明一个存在陈述是对的我们只需要找到一个例子（example）。
- 全称陈述 $\forall n[P(n)]$ 的否命题是存在陈述 $\exists n[\neg P(n)]$。
- 存在陈述 $\exists n[P(n)]$ 的否命题是全称陈述 $\forall n[\neg P(n)]$。

“一阶”指的是量词仅作用于个体变量（如 $\forall x$），而不能作用于谓词或函数（即不能量化“性质”本身）。例如，“二阶逻辑”允许 $\forall P$，但一阶逻辑不允许。高阶逻辑表达能力更强，但一阶逻辑具有更好的计算性质（如完备性、紧致性）。

1.5 Some Methods of Proof 一些证明方法

一些基本的证明思路：

- 直接证明（direct proof）
- 数学归纳法（proof by induction）
- 证明逆否命题（proof by contraposition）
- 举例/反例（proof by example or counterexample）
- 反证法（proof by contradiction）

之所以要强调“数学”归纳法，是因为一般科学界也有归纳法，如果你观察到 1 只、10 只、100 只、1000 只甚至 10000 只狐狸是橙色的，你可能会得出结论：狐狸是橙色的，但是你观察到的第 10001 只狐狸仍然有可能是别的颜色的，因此普通归纳法在逻辑上是不严谨的。而数学归纳法能够从逻辑上包含所有情况，继而得出可靠的结论。

例题 1.1 (反证法) 用反证法证明 $\sqrt{2}$ 是无理数。

2 Sets 集合

2.1 Naive Set Theory 朴素集合论

朴素集合论（Naive Set Theory）是指在不依赖严格公理系统的情况下，直观地研究集合及其性质的理论。它由德国数学家格奥尔格 · 康托（Georg Cantor）在 19 世纪末创立，是现代集合论的雏形，但后来因发现悖论而被更严格的公理化集合论（如 ZFC 系统）取代。

2.1.1 Sets 集合

一个**集合**（set）简单来说（在朴素集合论的框架下来说）就是 a collection of objects。这些集合里的 objects 被称为集合的**元素**（element），通常是数字，但也可以是函数、向量、矩阵、序列、人物、动物、颜色等任意具体或抽象的对象。我们通常用大写字母表示集合，用小写字母表示集合的元素。以下三个集合是等价的（第一个用了枚举法表示集合，后两个用了描述法）：

$$A = \{1, 2, 3\}$$

$$A = \{n | n \in \mathbb{Z} \ and \ 1 \leq n \leq 3\}$$

$$A = \{n : n \in \mathbb{Z} \ and \ 1 \leq n \leq 3\}$$

一个不包含任何元素的集合叫**空集**（empty set），符号记为 $\varnothing$。两个集合相等当且仅当里面的元素完全相同。如果集合 S 里的每一个元素都是集合 T 的元素，我们称集合 S 是集合 T 的一个**子集**（subset），记作 $S \subseteq T$。$A = B$ 当且仅当 $A \subseteq B$ 并且 $B \subseteq A$。空集是任何集合的子集。

集合 A 所有子集的集合叫 A 的**幂集**（power set）。若集合 A 有 n 个元素，则子集的总数（幂集的元素个数）为 $|\mathcal{P}(A)| = 2^n$，真子集的个数为 $2^n - 1$，非空子集的个数为 $2^n - 1$，非空真子集的个数为 $2^n - 2$。例如，集合 $S = \{0, 1, \{0, 1\}, \{2\}\}$ 的子集共 16 个：

$$\{0\}, \{1\}, \{\{0, 1\}\}, \{\{2\}\},$$

$$\{0, 1\}, \{0, \{0, 1\}\}, \{0, \{2\}\}, \{1, \{0, 1\}\}, \{1, \{2\}\}, \{\{0, 1\}, \{2\}\},$$

$$\{0, 1, \{0, 1\}\}, \{0, 1, \{2\}\}, \{0, \{0, 1\}, \{2\}\}, \{1, \{0, 1\}, \{2\}\},$$

$$\{0, 1, \{0, 1\}, \{2\}\}, \varnothing$$

2.1.2 Set Operations 集合运算

集合的基本运算有**并**、**交**、**差**、**补**四种。

两个集合 A 和 B 的**并集**（union）是包含所有属于 A 或 B 的元素的集合，记为 $A \cup B$。

两个集合 A 和 B 的**交集**（intersection）是仅包含同时属于 A 且 B 的元素的集合，记为 $A \cap B$。

在特定讨论范围内，包含所有相关元素的集合叫**全集**（universal set），通常用 U 或者 Ω 表示。全集中不属于集合 A 的所有元素组成的集合叫做 A 的**补集**（complement set），记作 A^c 或 $\overline{A}$。

两个集合之间的另一个基本运算叫**差集**（difference of sets），记作 $A - B$ 或者 $A \setminus B$，代表所有属于集合 A 但不属于集合 B 的元素的集合。差集有时也叫做补集的一种特殊形式，A^c 也可以被记作 $U \setminus A$ 或者 $U - A$。

我们可以借助**韦恩图**（Venn diagram）来可视化集合。

用集合运算语言表述德摩根反演律：$\overline{A \cap B} = \overline{A} \cup \overline{B}$。

2.1.3 Cartesian Product 笛卡尔乘积集合

在数学中，两个集合 A 和 B 的**笛卡尔积** (Cartesian product) 是一个集合，其中每个元素都是由一个来自 A 的元素和一个来自 B 的元素组成的有序对。如果集合 A 有 a 个元素，集合 B 有 b 个元素，则 A 与 B 的笛卡尔积有 $a \times b$ 个元素。

2.1.4 Finite Sets and Infinite Sets 有限集与无限集

若集合 S 由 n 个元素组成，这里 n 是确定的非负整数，则称集合 S 为**有限集**（finite set）。不是有限集的集合称为**无限集**（infinite set）。常见的无限集包括：自然数集 $\mathbb{N}$，整数集 $\mathbb{Z}$，有理数集 $\mathbb{Q}$，实数集 $\mathbb{R}$，复数集 $\mathbb{C}$ 还有区间（interval）。区间是一类特殊的 $\mathbb{R}$ 的子集：设 $a, b(a < b)$ 是两个实数，则满足不等式 $a < x < b$ 的所有实数 x 的集合称为以 a, b 为端点的**开区间**（open interval），记为 $(a, b) = \{x|a < x < b\}$。满足不等式 $a \leq x \leq b$ 的所有实数 x 的集合称为以 a, b 为端点的**闭区间**（closed interval），记为 $[a, b] = \{x|a \leq x \leq b\}$。满足不等式 $a < x \leq b$ 或 $a \leq x < b$ 的所有实数 x 的集合称为以 a, b 为端点的**半开半闭区间**，分别记为 $(a, b] = \{x|a < x \leq b\}$ 或 $[a, b) = \{x|a \leq x < b\}$。上述几类区间的长度是有限的，称为**有限区间**。除此之外，还有下述几类**无限区间**：$(a, +\infty) = \{x|x > a\}$；$[a, +\infty) = \{x|x \geq a\}$；$(-\infty, b) = \{x|x < b\}$；$(-\infty, b] = \{x|x \leq b\}$。我们也可以把实数集 $\mathbb{R}$ 写成区间的形式：$\mathbb{R} = (-\infty, +\infty)$。

2.2 Cantor's Theorem and Continuum Hypothesis 康托定理和连续统猜想

康托尔当时面临的问题是比较两个东西的大小，可以说康托尔创建集合论就是为了比较大小。试想，想要比较生活中两个“有限集”之间的大小关系，我们真的需要集合论的语言吗？显然是不必要的。因此康托尔面临的显然并不是如“一个点多还是两个点多”这般的问题，而是无限集大小的比较。有关无限集大小的讨论才是朴素集合论的本质核心和主要贡献，比较无限集的大小正是集合论创建之初的根本目的。

对于有限集而言，我们可以简单地用“元素的个数”来衡量集合的大小，但对于无限集而言，我们需要新的度量。我们引出集合的**基数**（cardinality），或者叫**势**的概念来比较无限集的大小。集合 S 的基数记作 $|S|$。当然，对于有限集 A，它的基数就等于其元素的个数 $|A| = n$，例如 $|\varnothing| = 0$。

我们可以定义两个集合的**基数相等**（或者称**等势**）当且仅当存在它们之间的一一对应，也就是**双射**（bijection）。注意，无穷集可能和自己的一个真子集等势。例如，自然数集同全体非负偶数构成的集合等势。另一个更加反直觉的例子是，自然数集与自然数集的笛卡尔积同自然数集本身也是等势的。有没有可能无穷集都是等势的呢？如果是，那么问题就解决了，我们就可以简单地用 $+\infty$ 来一并表示所有无限集的大小了。所以，这个问题本质上是在问，无穷大有没有不同的层次和结构？

很遗憾，康托尔定理说明了一个显然的反例：如果 X 是集合，则 X 与它的幂集 $\mathcal{P}(X)$ 不等势。对于有限集而言，等势需要满足两个集合的元素个数相同，即 $n = 2^n$，显然该方程在自然数范围内是无解的。对于无限集而言，假设存在双射 $f : X \to \mathcal{P}(X)$，考虑集合 $T = \{x \in X | x \notin f(x)\}$，则 T 是 $\mathcal{P}(X)$ 的元素，由于双射，存在唯一的 S 使得 $f(S) = T$。易得 $S \in T$ 不成立且 $S \notin T$ 不成立，矛盾。这意味着我们需要对无穷集进行更精细的刻画。

所谓更精细的刻画，即我们不取双射而取一些更弱的映射：X, Y 是集合，如果存在一个**单射**（injection）$f : X \to Y$，则称 X 的基数/势小于等于 Y 的基数/势，记为 $|X| \leq |Y|$。康托尔定理指出，任何集合 X 的基数严格小于其幂集 $\mathcal{P}(X)$ 的基数，即：$|X| < |\mathcal{P}(X)|$。（注意这意味着不存在基数最大的无穷集，因为假设存在这样一个集合，取它的幂集就会立刻得到一个基数更大的集合。）康托尔的发现说明了无穷并非是单一结构的，而是具有更丰富的层次。

我们把自然数集的势 $|\mathbb{N}|$ 记为 $\aleph_0$（阿列夫 0），把与自然数集等势的集合称为**可数集**（countable set），或者叫**可列集**。自然数集、整数集、有理数集都是可数集。（注意有理数集是整数集与整数集笛卡尔积的一个子集。）我们可以得到以下定理：

- 自然数集并上或除去有限集后与自然数集本身等势。
- $\aleph_0$ 是最小的无穷势。（换言之，对于任何一个无穷集，它的势都大于等于自然数集的势。）
- 两个等势的集合的幂集等势。
- 实数集与自然数集的幂集等势。
- 不是可数集的无穷集称为**不可数集**（uncountable set）。

集合	**势（基数）**	**可数性**
$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{N}\times\mathbb{N}$	$\aleph_0$	Countable
$\mathbb{R}$, $\mathbb{C}$, $\mathcal{P}(\mathbb{N})$	$2^{\aleph_0}$	Uncountable
$\mathcal{P}(\mathbb{R})$, $\mathcal{P}(\mathbb{C})$, $\mathcal{P}(\mathcal{P}(\mathbb{N}))$	$2^{2^{\aleph_0}}$	Uncountable
$\mathcal{P}(\mathcal{P}(\mathbb{R}))$, $\mathcal{P}(\mathcal{P}(\mathbb{C}))$, $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$	$2^{2^{2^{\aleph_0}}}$	Uncountable

表 4: 集合的势与可数性（按势分组）

通过上面的讨论我们发现，有理数集是可数的，实数集是不可数的。根据我们的直觉，似乎没有什么集合介于两者之间且势有本质区别。事实真的是这样吗？我们定义 $\aleph_1$ 是大于 $\aleph_0$ 的最小基数，同理可以定义 $\aleph_n$ 是大于 $\aleph_{n-1}$ 的最小基数。下面这个著名的命题叫**连续统猜想**（Continuum Hypothesis）：$\aleph_1 = 2^{\aleph_0}$。换言之，有理数集和实数集之间的集合要么是可数的，要么和实数集等势。

1874 年，康托提出连续统猜想。1900 年，希尔伯特将连续统猜想列为希尔伯特 23 问题之首。1938 年，哥德尔证明了连续统猜想与 ZFC 公理系统不矛盾。1963 年，科恩证明了连续统猜想与 ZFC 公理系统相互独立（即连续统猜想就是 ZFC 公理系统中一个既不能被证明也不能被证伪/不完备的命题），并因此获得了 1966 年的菲尔兹奖。事实上，广义连续统猜想（即 $\aleph_n = 2^{\aleph_{n-1}}$，即幂集运算恰好产生下一个阿列夫数）是可以推出选择公理的。以下是集合的势的一些其它常识：

- 有限个集合的笛卡尔积的势取决于其中最大的集合的势。
- 有限个集合的并的势取决于其中最大的集合的势。
- 可数个可数集的并仍然是可数的。
- 可数个有限集的笛卡尔积可能是不可数的。

2.3 Russell's Paradox and ZFC Set Theory 罗素悖论和 ZFC 公理系统

19 世纪末，数学日益形式化，但人们开始发现一些自相矛盾的“悖论”，尤其是康托的朴素集合论中过于自由的集合定义会导致逻辑矛盾。其中，最著名的悖论当属**罗素悖论**（Russell’s Paradox）。

之前我们在定义“集合”这个概念的时候只是非常轻率和宽松地描述为“a collection of objects”，这个 objects 似乎可以是任何东西，但事实上这样简单的定义隐藏着巨大的漏洞和问题。罗素悖论告诉我们不加限制的集合概念会导致自相矛盾。假设我们可以随意定义集合——只要我们说“某些

东西满足某个性质"，我们就可以把这些东西组成一个集合 $\{x|P(x)\}$，即朴素集合论下的**概括原则**（Comprehension Principle）。现在考虑这样一个集合：$R=\{x|x\notin x\}$，也就是说，R 是"所有不属于自己的集合"组成的集合。请问，$R\in R$ 成立吗？你会发现，无论你的答案是属于还是不属于，都会导致矛盾，这就是罗素悖论。罗素悖论的一个经典例子是**理发师悖论**（Barber Paradox）：如果一个理发师只给这个村里所有不给自己刮胡子的人刮胡子，那么他应不应该给自己刮胡子？

罗素悖论指出了一个根本问题：不能允许任意性质都定义出一个集合，否则就可能定义出像 R 这样的"**自指**（self-reference）"集合，导致逻辑系统崩溃。一对矛盾 $P\wedge\neg P$ 足以用来证明一切。因为 P 为真，所以 $P\vee Q$ 为真，继而 P 和 Q 至少有一个为真。又因为 $\neg P$ 为真，所以 P 为假，故 Q 为真，证毕。这在哲学中被叫做**爆炸原理**（Principle of Explosion）。这意味着我们必须要对集合里的 objects 加以限制，但需要添加哪些限制呢？策梅洛-弗兰克尔集合论（ZF/ZFC）给出了一个答案。ZFC 是数学中最常用的公理化集合论系统，由以下 9 条公理组成，为现代数学奠定基础。

1. 外延公理（Axiom of Extensionality）
   - 内容：两个集合相等，当且仅当它们的元素完全相同。
   - 意义：定义集合的本质是它的元素，而非描述方式。
2. 分离公理模式（Axiom Schema of Separation）
   - 内容：对任意集合 A 和性质 P，存在子集 $\{x\in A|P(x)\}$。
   - 意义：分离公理是对概括公理的一个改进，也是我们构造一个集合的主要办法。事实上，概括公理是一条非常好的公理，正是因为集合有着相当大的任意性，集合论才能统一诸多数学理论，成为一种奠基的语言。但是罗素悖论说明概括公理有一些太任意了，我们要给它增加一点小小的限制来保证它成立。相比于概括公理，分离公理对 x 的范围做出了一点小小的限制，我们必须首先假设 x 在某一个集合中，然后从这个集合中分离出满足命题 $P(x)$ 的元素。可以看出，分离公理主要用于构造子集这一操作。根据外延公理，集合 $\{x\in A|P(x)\}$ 是唯一的，所以这个符号是良定（well-defined）的。事实上，$\{x\in A|P(x)\}$ 就是朴素集合论中集合的描述法，分离公理保证了这一写法是正确的。

     根据分离公理我们可以得到绝大多数朴素集合论中我们早就熟悉的集合。设 X 是一个集合，根据分离公理 $\{x\in X|x\neq x\}$ 是一个集合，将其记为 $\varnothing_X$，称为依赖于 X 的空集。对于两个集合 X 和 Y，利用外延公理可知 $\varnothing_X=\varnothing_Y$，所以通过不同集合构造出的空集是相等的，我们将其记为 $\varnothing$，称为空集。

     根据分离公理我们也可以很容易地得到我们在朴素集合论中熟悉的集合操作，例如交集、差集和子集。对于两个集合 X 和 Y，根据分离公理，$\{x\in X|x\in Y\}$ 是一个集合，将其记为 $X\cap Y$，称为 X 和 Y 的交集。利用外延公理容易验证 $X\cap Y$ 和 $Y\cap X$ 是同一个集合。事实上，我们还可以在此定义无限多个集合的交集，我们可以任意选择一个集合作为分离公理要求的框定范围，再用外延公理说明这个集合不依赖于最初集合的选择即可。同理，可定义 $X\setminus Y=\{x\in X|x\notin Y\}$，称为 X 和 Y 的差集。

     对于两个集合 X 和 Y，如果 $x\in X\implies x\in Y$，那么称 X 是 Y 的一个子集。分离公理中构造的每一个集合 $\{x\in X|P(x)\}$ 都是集合 X 的子集；另一方面，每一个子集 Y 都可以表示为外延公理 $\{x\in X|x\in Y\}$ 的形式。如果 X 是 Y 的子集且两者并不相等，则称 X 是 Y 的真子集。

     分离公理本质上就是在定义子集，它给到我们一个重要的启示：比一个已知的集合更小的集合我们是可以无条件去定义的。事实上，构造一个比已知集合更大的集合很多时候是不被允许的。

3. 并集公理（Axiom of Union）

- 内容：对任意集合 A，存在其所有元素的并集 $\bigcup A$。
- 意义：支持无限集合的合并。

4. 幂集公理（Axiom of Power Set）

- 内容：对任意集合 A，存在其所有子集的集合 $\mathcal{P}(A)$。
- 意义：事实上，我们目前只知道空集是存在的，如果没有幂集公理，那么集合论中完全可能仅存在一个集合：空集。空集根据分离公理与并集公理仍然只能得到空集。幂集公理可以帮助我们构造包含任意有限个元素的集合。它的另一个重要意义是允许我们讨论更高阶的无限（如实数集的基数），即在集合的势以及自然数集的构造中发挥至关重要的作用。

5. 无穷公理（Axiom of Infinity）

- 内容：存在一个无限集。
- 意义：我们已经构造了有任意有限个元素的集合，我们希望更进一步，在集合论中构造出具有无限多个元素的集合。我们容易发现，这一点并不能通过已有的公理推出。事实上无穷公理有很多等价的写法，我们这里采取幂集的方式：$\exists X(\varnothing \in X) \wedge (x \in X \implies \mathcal{P}(x) \in X)$。事实上无穷公理生成的这个集合就是自然数集，无穷公理是把 Peano 公理嵌入集合论中了。

6. 配对公理（Axiom of Pairing）

- 内容：对任意两个集合 A 和 B，存在集合 $\{A, B\}$ 包含它们。
- 意义：我们常见的集合论操作中还有一个操作未被提及：笛卡尔（Cartesian）积。笛卡尔积事实上是构造了一个有序数对，ZFC 公理首先构造无序数对。至此，朴素集合论中常见的集合都已经被定义完毕。

7. 正则公理（Axiom of Regularity/Foundation）

- 内容：任何非空集合 A 都包含一个与 A 不相交的元素。
- 意义：我们在定义了常见的集合之后，接下来的重点便是如何去避免罗素悖论。回顾罗素悖论，我们已经将概括公理弱化为了分离公理。我们要做的是要防止分离公理“进化”为概括公理。不难发现，如果我们承认所有集合构成一个集合，那么分离公理就会“进化”为概括公理，这是我们不允许的。我们希望表述一个公理，即所有集合放在一起不构成一个集合，但是这种语句显然不太好。对于这个以所有集合为元素的“集合”，它的哪一个元素是最诡异的呢？那便是它自己。事实上罗素悖论也正是取出了这一个集合所导致的。于是我们必须把这一个元素否定掉。同样，在否定掉这一点后，我们也要否定掉所有“套娃”式的定义。正则公理的一个重要推论是不存在集合 X 满足 $X \in X$。这条定理防止分离公理“进化”为概括公理，从而避免了罗素悖论。
  在集合论的语言下，罗素悖论究竟用到了哪些事实呢？事实上只用到了三点：分离公理；所有集合构成一个集合；排中律（反证法）。只要承认上述三点，就一定会产生悖论。其中分离公理是不能修改的（很难想象一个集合的一部分不是一个集合），排中律同样很难否认（秉持直觉主义数学哲学观的数学家们就会否定排中律，从而避免了罗素悖论）。思来想去，只能牺牲“所有集合构成一个集合”这一点。事实上，通过罗素悖论同样可以证

明，所有范畴不构成一个范畴，所有无穷范畴不构成一个无穷范畴...... 这也在侧面说明了数学没有先验的公理这一 Bourbaki 的数学哲学观点。

8. 替代公理模式（Axiom Schema of Replacement）
   - 内容：若一个“可定义的规则”将集合 A 的元素映射到其他对象，则这些映射结果的集合存在。
   - 意义：支持通过映射构造新集合，避免直接定义“过大”的集合。
9. 选择公理（Axiom of Well-Ordering/Choice）
   - 内容：对任意一组非空集合的集合，存在一个选择函数，能从每个集合中各选一个元素。
   - 意义：非构造性证明工具（如证明无限维空间的基存在），但独立于其他公理。

2.4 Gödel's Incompleteness Theorems 哥德尔不完备定理

在数学哲学的发展历程中，不同学派对数学基础的探索形成了一条深刻而富有张力的思想脉络。这条脉络始于 19 世纪末 20 世纪初，当时数学家们正试图为数学建立一个坚实而统一的基础。

逻辑主义的代表人物罗素（Russell）在《数学原理》中提出，数学可以完全归结为纯粹逻辑。他们认为，数学真理本质上就是逻辑真理，所有数学概念——如数、函数——都可以通过逻辑概念定义，而数学定理则可由逻辑规则推导出来。这一雄心勃勃的计划试图消除数学中的直觉成分，使其完全建立在理性演绎之上。

与此同时，**形式主义**的代表希尔伯特（Hilbert）则采取了不同的路径。他并不认为数学必须依赖于逻辑，而是主张通过公理化方法来严格形式化数学。希尔伯特希望建立一个完备且一致的公理系统，使得所有数学命题都能在其中得到证明或否证。他的“证明论”（元数学）试图用有限的方法证明数学系统的无矛盾性，从而确保数学的可靠性。

朴素集合论自信地宣称，集合概念足以统一数学的全部内容，逻辑主义与形式主义可以通过朴素集合论实现等价。然而，罗素悖论的出现，暴露了朴素集合论的内在矛盾，使其陷入危机。为了挽救集合论，ZF/ZFC 公理系统被提出，通过限制集合的构造方式来避免悖论，解决了罗素悖论，使集合论重新成为数学的可靠基础。

集合论声称一切数学归于我，而其中唯一的一个问题罗素悖论还被 ZF 公理给解决了，到这里为止，似乎数学就要迎来大结局。然而，就在逻辑主义和形式主义看似各有所成时，哥德尔的不完备性定理（1931）给了两者致命一击。哥德尔证明了：任何足够强的数学系统如果是无矛盾的，则必然是不完备的（存在无法被证明的真命题），而且其一致性无法在系统内部得到证明。这一结果彻底粉碎了希尔伯特的形式化完全数学之梦，也表明逻辑主义无法彻底消解数学的不可约减性。

在这样的背景下，**结构主义**（以 Bourbaki 学派为代表）提出了一种新的数学哲学观。他们认为，数学的本质不是研究具体的对象，而是研究**抽象结构**。Bourbaki 的《数学原本》试图以结构化的方式重建整个数学体系，强调数学的后验性——即数学不是先验存在的真理，而是人类在探索模式与关系时构建的框架。这一观点既避开了哥德尔定理对基础主义的限制，又为现代数学的多样性和抽象性提供了哲学支持。

2.5 Three Times of Crisis in the History of Mathematics 数学史上的三次危机

本节作为上节内容的补充材料存在。

数学史上的三次重大危机分别出现在古希腊、17-18 世纪和 19 世纪末至 20 世纪初，每一次都动摇了数学的基础，并最终推动了数学的深刻发展。

第一次数学危机为**无理数的发现**（公元前 5 世纪）。古希腊毕达哥拉斯学派认为“万物皆数”，即所有数都可表示为整数或整数比（分数）。然而，希帕索斯（Hippasus）发现边长为 1 的正方形的对角线长度（$\sqrt{2}$）无法表示为分数，打破了这一信念。无理数的存在挑战了当时对数的理解，导致几何学（如欧几里得《几何原本》）暂时取代算术成为数学的基础。欧多克索斯提出比例理论，后来实数系的严格定义（19 世纪戴德金分割等）彻底解决了问题。

第二次数学危机为**微积分的基础问题**（17-18 世纪）。牛顿和莱布尼茨发明的微积分虽实用，但依赖模糊的“无穷小量”概念（如瞬间变化率或无穷小增量），贝克莱主教嘲讽其为“消失的鬼魂”。微积分的逻辑漏洞引发争议，例如无穷小量在计算中是否为零。如果不等于零，那为什么有时我们会把它视为零；如果等于零，那无穷小为什么可以作为分母？第二次数学危机可以看作是一场数学的严谨性与数学的艺术性之间的交锋。19 世纪柯西、魏尔斯特拉斯等人提出极限的严格定义（$\varepsilon-\delta$ 语言），实数理论和分析学的基础得以巩固。

第三次数学危机为**集合论悖论**（19 世纪末-20 世纪初）。康托尔创立集合论作为数学基础，但罗素在 1901 年提出“理发师悖论”，揭示了朴素集合论中的自指矛盾。策梅洛-弗兰克尔（ZF）公理系统排除了已知悖论，哥德尔不完备定理（1931）则表明数学系统无法同时满足完备性和一致性，危机转为对数学基础的深层反思。

3 Mappings and Functions 映射与函数

3.1 Mappings 映射

映射是描述两个集合之间元素对应关系的广义概念。

定义 3.1 设 X, Y 是两个给定的集合，若按照某种规则 f，使得对集合 X 中的每一个元素 x，都可以找到集合 Y 中唯一确定的元素 y 与之对应，则称这个对应规则 f 是集合 X 到集合 Y 的一个**映射**，记为 $f: X \to Y$ 或者 $x \mapsto y = f(x)$。其中 y 称为在映射 f 之下 x 的**像**（***image***），x 称为在映射 f 之下 y 的一个**逆像**（也称为**原像**）。集合 X 称为映射 f 的**定义域**（***domain***），记为 D_f，集合 Y 称为映射 f 的**陪域**（***codomain***），而在映射 f 之下，X 中元素 x 的像 y 的全体称为映射 f 的**值域**（***range***），记为 R_f，即 $R_f = \{y | y \in Y \wedge y = f(x), x \in X\}$。

- 定义域是输入所有可能取值的集合。陪域是输出的目标集合，或者说所有可能的输出范围的上层集合。值域是实际输出结果的集合，是陪域的一个子集。
- 映射要求元素的像必须是唯一的。对于不满足像的唯一性要求的对应规则，一般只要对值域范围稍加限制，就能使它成为映射。映射并不要求逆像也具有唯一性。

定义 3.2 设 f 是集合 X 到集合 Y 的一个映射，若 f 的逆像也具有唯一性，即对 X 中的任意两个不同元素 $x_1 \neq x_2$，它们的像 y_1 与 y_2 也满足 $y_1 \neq y_2$，则称 f 为**单射**（***injection***）；如果映射 f 满足 $R_f = Y$，则称 f 为**满射**（***surjection***）；如果映射 f 既是单射又是满射，则称 f 是**双射**（***bijection***）（又称一一对应）。

- 设 $f: X \to Y$ 是单射，则由定义 3.2 可知，对任意一个 $y \in R_f$，它的逆像 $x \in X$ 是唯一确定的。由定义 3.1 可知，对应关系 $g: R_f \to X$ 构成了 R_f 到 X 上的一个双射，我们把它称为 f 的**逆映射**，记为 f^{-1}，其定义域为 $D_{f^{-1}} = R_f$，值域为 $R_{f^{-1}} = X$。
- 现设有如下两个映射：$g: X \to U_1$ 和 $f: U_2 \to Y$，如果 $R_g \subseteq U_2 = D_f$，那就可以构造出一个新的映射 $f \circ g: X \to Y$，我们将之称为 f 和 g 的**复合映射**。

3.2 Functions 函数

映射强调元素的对应关系，不限制集合类型（可以是数字、图形、抽象对象等）。函数是映射的一种特例，通常指数集到数集的映射。在纯数学中，两者界限逐渐模糊。有些现代教材可能将“函数”推广到任意集合的映射（如“泛函”是函数空间的函数），此时两者可视为同义词。

3.2.1 Real-Valued Functions of a Single Variable 一元实函数

定义 3.3 若在定义 *3.1* 中特殊地取集合 X 为 $\mathbb{R}$ 的子集，集合 $Y = \mathbb{R}$，则映射 $f: X \to Y$ 称为**一元实函数**（*a real-valued function of a single variable*），简称**函数**。

- 由于函数表示的必是实数集合与实数集合之间的对应关系，所以在其映射表示中，第一行是不需要的，只要写成 $y = f(x), x \in X$ 即可。
- 我们称变量 x 为**自变量**（independent variable），变量 y 为**因变量**（dependent variable）。

3.2.2 Euclidean Space 欧几里得空间

欧几里得空间，或简称**欧式空间**（记为 $\mathbb{R}^n$）是指满足欧几里得几何性质的空间。它是一种“平直”的空间（与弯曲的黎曼空间相对）。在我们熟悉的二维平面和三维世界中，欧几里得几何的以下性质成立：

- 两点之间直线最短。
- 平行线永不相交。
- 三角形内角和为 180°。
- 勾股定理成立。

3.2.3 Cartesian Coordinate System 笛卡尔坐标系

笛卡尔坐标系由解析几何之父法国数学家笛卡尔（René Descartes）提出，通过坐标（数字组）表示空间中的点，实现了几何与代数的桥梁（解析几何），让代数可视化，几何参数化。在平面上，笛卡尔坐标系也叫**平面直角坐标系**，是两个相交于原点相互垂直的实数轴。在平面直角坐标系下，二维欧氏空间可看作由 x 轴、y 轴和四个象限构成。

注意，欧几里得空间不依赖于坐标系，它本身具有内在的几何性质（如距离、角度），这些性质不依赖于坐标的选择。欧几里得空间是抽象的几何空间，而坐标系为其提供了具体的数值表示方法，是描述欧几里得空间的工具。通过坐标系，欧几里得空间中的点、向量、曲线、曲面等都可以用坐标和方程表示。笛卡尔坐标系只是其中一种描述方式（还有其他坐标系，如极坐标、球坐标等）。没有笛卡尔坐标系，欧几里得空间仍然存在（但难以计算）；没有欧几里得空间，笛卡尔坐标系就失去了几何背景。事实上，笛卡尔坐标系是 $\mathbb{R}^n$ 上的标准正交坐标系（基向量为单位正交向量）。（之后线性代数部分会学到）

当我们提到 n 维欧几里得空间 $\mathbb{R}^n$ 的时候，如无特殊说明，一般默认选用笛卡尔坐标系作为坐标系。

3.2.4 Graphs of Functions: Cartesian Equations 函数的图像：笛卡尔方程

二维笛卡尔坐标系（平面直角坐标系）可以帮助我们在二维欧几里得空间（欧式平面）中画出一元实函数的图像，直观地展示函数的形状。为了简便，我们一般把装备了二维笛卡尔坐标系的二维欧氏空间统称为**笛卡尔平面**（Cartesian plane），它是一个包含了空间及坐标系的可操作系统。对于二元实函数，我们需要三维空间来可视化它，此时用手画在纸上就比较困难，但我们可以借助计算机来生成函数图像。对于更高维的函数图像，由于已经突破了我们生活的世界和我们的感知能力的维度限制，我们无法完整地画出函数图像，但我们可以通过一些技巧来可视化三元甚至更高维的函数。常用的两个方法如下：

1. 切片法：对于三元函数，我们可以固定输出来创造一个三维空间中的等值面（level surface），也就是降维到二元函数。改变这个常数，我们可以得到一系列等值面。同样的操作如果发生在二元函数，我们会在笛卡尔平面上得到一系列等高线。当然我们也可以固定一个输入同样实现降维，把三元函数变为二元函数，然后在三维空间里画出一个三维曲面。改变输入的值，我们可以得到一系列切片曲面。对于 n 元函数，原理相同，固定 $n-2$ 或 $n-3$ 个变量，将其降维成二维曲线或三维曲面来研究。

2. 增加其他视觉元素来编码信息：最常见的例如在三维空间的一个点上用颜色或亮度来表示第四个变量的值。（当然同样的方法也可以用来在平面上可视化二元函数。）

任何函数的图像，都可以用一个**笛卡尔方程**（Cartesian equation）来表示。在笛卡尔平面中，笛卡尔方程是指用一个包含变量 x 和 y 的代数方程来描述平面几何图形的方法，通常写作 $F(x,y)=0$ 的形式。笛卡尔方程的几何本质是定义了一个点的集合 $S_F \subseteq \mathbb{R}^2$，这个集合包含了所有且仅有那些满足 $F(x,y)=0$ 的点 $(x,y)\in\mathbb{R}^2$。因此，它的核心是描述一个静态的几何图形。

有了笛卡尔方程的定义，我们就可以说，一个函数 $f : X \to Y$ 的图像是笛卡尔方程 $y = f(x), x \in X$ 所定义的 $\mathbb{R}^2$ 的子集。（注意这里的方程可以轻松转化为一般形式，我们只需要定义 $F(x,y) = y - f(x)$。）用形式语言我们可以写：

$$Graph(f) = \{(x, f(x)) \in \mathbb{R}^2 | x \in X\}$$

- 任何一元实函数都可以被用一个 $\mathbb{R}^2$ 中被笛卡尔方程 $y = f(x)$ 定义的图像所表示。然而，并非所有笛卡尔方程都能表示一个函数。判断一个笛卡尔方程是否能定义为一个函数，关键在于垂直线检验（vertical line test）。对函数而言，在其图像上，任何一条垂直于 x 轴的直线，与函数图像至多只有一个交点。如果存在一条垂直线与图形有两个或以上交点，则该图形不能表示为一个函数。例如，圆的方程 $x^2 + y^2 = 4$ 是一个笛卡尔方程，但却不是一个函数。也就是说，笛卡尔方程除了能帮助画出函数图像，还能更一般地给出笛卡尔平面里更多其他图形的代数方程式。

- 注意，垂直线检验是几何角度函数定义的充要条件。如果一个图形是函数的图像，那么它必然通过垂直线检验。（必要性）如果一个图形通过了垂直线检验，那么它必然能表示一个函数。（充分性）

- 我们已经知道了在笛卡尔平面内任意画一个图形，它显然不一定能作为某个函数的图像。但事实上，在笛卡尔平面内任意画一个图形，它也不一定能表示为一个笛卡尔方程（如果我们将其严格定义为一个由有限次初等运算构成的，关于 x 和 y 的多项式方程，或有限个这样的方程构成的方程组）。在严格的数学意义上，由多项式方程定义的图形只是所有可能图形中非常特殊的一类。超越曲线、非可构图形、分形等均无法由这样的多项式方程表示。因此，有极大概率你在纸上随手画的一个图形就没有对应的笛卡尔方程。如果笛卡尔方程泛指包含 x 和 y 的等式，那么理论上任意图形都可以表示为笛卡尔方程，但可能只是对该图形的循环描述，并不提供新的代数洞察。

- 除了笛卡尔方程这种直接写出 x 和 y 关系的图形表示方法，我们也可以用参数方程（parametric equation）来表示，这种方程会直接给出 x 与一个参数 t 的关系以及 y 与 t 的关系，继而间接给出 x 与 y 的关系。

3.2.5 Elementary Functions 初等函数

以下 6 类函数被称为**基本初等函数**（basic elementary function）：

1. 常数函数：$y = c$
2. 幂函数：$y = x^\alpha (\text{e.g.} x, x^2, x^3, \sqrt{x}, \sqrt[3]{x}, \frac{1}{x^2} ...)$
3. 指数函数：$y = a^x (a > 0 \text{且} a \neq 1)$
4. 对数函数：$y = \log_a x (a > 0 \text{且} a \neq 1)$

5. 三角函数：$y=\sin x, y=\cos x, y=\tan x, y=\cot x, y=\sec x, y=\csc x...$

6. 反三角函数：$y=\arcsin x, \arccos x, \arctan x...$

由基本初等函数经过有限次四则运算与复合运算所产生的函数称为**初等函数**（elementary function），例如 $y=ax^2+bx+c, y=\frac{\log_a(1+x)}{\sqrt{x^2+1}}, y=\sin\frac{1}{x}+\cos^2 x$。

初等函数的**自然定义域**（natural domain）是指它的自变量的最大取值范围。一般来说，给出一个函数的具体表达式的同时应该指出它的定义域，否则即表示默认该函数的自然定义域为其定义域。当两个函数不仅函数关系相同，而且定义域也相同时，它们的值域必然相同，它们表示的是相同的函数，此时自变量与因变量采用什么符号是无关紧要的。

3.2.6 Piecewise, Implicit, and Parametric Functions 函数的分段表示、隐式表示与参数表示

设 A, B 是两个互不相交的实数集合，α 和 β 是分别定义在集合 A 和集合 B 上的函数，则

$$f(x)=\begin{cases}\alpha(x), x\in A\\ \beta(x), x\in B\end{cases}$$

是定义在集合 $A\cup B$ 上的函数，这样的表示方法称为函数的分段表示。这里函数 f 是分成两段来表示的，事实上，分段表示可以分成任意有限段，甚至无限多段。分段定义的函数经常在当真实世界的数据被建模的时候出现。

下面我们介绍几个常用的分段表示函数。

1. 符号函数（signal function）sgn x：

$$sgnx=\begin{cases}1, x>0\\ 0, x=0\\ -1, x<0\end{cases}$$

   它的定义域是 $D=(-\infty,+\infty)$，值域是 $R=\{-1,0,1\}$。

2. 绝对值函数（absolute value function）$|x|$：

$$|x|=\begin{cases}x, x\geq 0\\ -x, x<0\end{cases}$$

   它的定义域是 $D=(-\infty,+\infty)$，值域是 $R=[0,\infty)$。

3. 下取整函数（greatest integer function/integer floor function）$\lfloor x\rfloor$：$\lfloor 2.4\rfloor=2$
   它的定义域是 $D=(-\infty,+\infty)$，值域是 $R=\mathbb{Z}$。

4. 上取整函数（least integer function/integer ceiling function）$\lceil x\rceil$：$\lceil 2.4\rceil=3$
   它的定义域是 $D=(-\infty,+\infty)$，值域是 $R=\mathbb{Z}$。

前面所举例子的共同特点是函数形式均为 $y=f(x)$，即因变量 y 单独放在等式的一边，而等式的另一边是只含有自变量 x 的表达式，这称为**函数的显式表示**（explicit form）。而所谓**函数的隐式表示**（implicit form），是指通过方程 $F(x,y)=0$ 来确定变量 y 与 x 之间函数关系的方式，这也是一种重要的函数表示形式。

在表示变量 x 与 y 的函数关系时，我们常常需要引入第三个变量（例如参数 t），通过建立 t 与 x,t 与 y 之间的函数关系，间接地确定 x 与 y 之间的函数关系，这种方法称为**函数的参数表示**（parametric form）。

3.2.7 Basic Properties of Functions 函数的简单特性

1. 有界性（boundedness）

   定义 3.4 *若存在两个常数 m 和 M，使函数 $y=f(x),x\in D$ 满足 $m\le f(x)\le M,x\in D$，则称函数 f 在 D* **有界**。*其中 m 是它的***下界**，*M 是它的***上界**。

   - 注意，当一个函数有界时，它的上界与下界不唯一。由上面定义可知，任意小于 m 的数也是 f 的下界，任意大于 M 的数也是 f 的上界。

   定义 3.5 *若存在常数 $M>0$，使函数 $y=f(x),x\in D$ 满足 $|f(x)|\le M,x\in D$，则称函数 f 在 D* **有界**。

   - 这是有界函数的另一定义，容易证明这两种定义是等价的。

2. 单调性（monotonicity）

   定义 3.6 *对函数 $y=f(x),x\in D$，若对任意 $x_1,x_2\in D$，当 $x_1<x_2$ 时成立 $f(x_1)\le f(x_2)$（或 $f(x_1)<f(x_2)$），则称函数 f 在 D* **单调增加**（或**严格单调增加**）；*若对任意 $x_1,x_2\in D$，当 $x_1<x_2$ 时成立 $f(x_1)\ge f(x_2)$（或 $f(x_1)>f(x_2)$），则称函数 f 在 D* **单调减少**（或**严格单调减少**）。

3. 奇偶性（parity）

   定义 3.7 *设函数 f 的定义域 D 关于原点对称，即 $x\in D\Leftrightarrow -x\in D$。若对一切 $x\in D$，成立 $f(-x)=f(x)$，则称函数 f 是***偶函数**（*even function*）；*若对一切 $x\in D$，成立 $f(-x)=-f(x)$，则称函数 f 是***奇函数**（*odd function*）。

   - 显然，奇函数的图像关于原点对称，偶函数的图像关于 y 轴对称。了解了函数的奇偶性，我们只需在 $D\cap[0,+\infty)$ 上讨论函数的性质，再由对称性推出它在 $D\cap(-\infty,0]$ 上的性质。

   定理 3.1 *定义在 $(-\infty,+\infty)$ 上的任何函数都可以唯一地表示为一个奇函数与一个偶函数的和。*

   证明 3.1

   $$f(x)=\frac{f(x)+f(-x)}{2}+\frac{f(x)-f(-x)}{2}$$

4. 周期性（periodicity）

   定义 3.8 *若存在常数 $T>0$，使得对一切 $x\in D$，成立 $f(x+T)=f(x)$，则称函数 f 是***周期函数**，*T 称为它的***周期**。*若存在满足上述条件的最小的 T，则称它为 f 的***最小周期**。

   - 显然，周期函数 f 的定义域 D 必须满足条件：对一切 $x\in D$，有 $x\pm T\in D$。
   - 并非每个周期函数都有最小周期。

3.2.8 Two Useful Inequalities 两个常用不等式

下面介绍两个简单但重要的不等式，它们不仅在数学分析的证明中频繁出现，而且在其他数学分支中也都有广泛的用途。

定理 3.2（三角不等式） 对于任意实数 a 和 b，都有 $||a|-|b|| \leq |a \pm b| \leq |a|+|b|$。

定义 3.9 设 $a_1, a_2, ..., a_n$ 是 n 个正数，则称 $\frac{a_1+a_2+...+a_n}{n}$ 是它们的**算术平均值**；$\sqrt[n]{a_1 a_2 ... a_n}$ 是它们的**几何平均值**；$\frac{n}{(\frac{1}{a_1}+\frac{1}{a_2}+...\frac{1}{a_n})}$ 是它们的**调和平均值**。

这三个平均值之间成立如下关系：

定理 3.3（平均值不等式） 对任意 n 个正数 $a_1, a_2, ..., a_n$，有

$$\frac{a_1+a_2+...+a_n}{n} \geq \sqrt[n]{a_1 a_2 ... a_n} \geq \frac{n}{(\frac{1}{a_1}+\frac{1}{a_2}+...\frac{1}{a_n})}$$

4 Limits of Sequences 数列极限

4.1 Completeness of the Real Number System 实数系的连续性

4.1.1 The Real Number System 实数系

人类对数的认识是从自然数开始的。若一个集合中的任意两个元素进行了某种运算后，所得的结果仍属于这个集合，我们称该集合对这种运算是**封闭**的。显然，任意两个自然数，其和与积必定还是自然数，即自然数集合 $\mathbb{N}$ 对于加法与乘法运算是封闭的。但是 $\mathbb{N}$ 对于减法运算并不封闭。

当数系由自然数集合扩充到整数集合 $\mathbb{Z}$ 后，关于加法、减法和乘法运算都封闭了，但是关于除法运算是不封闭的，因此数系又由整数集合扩充为有理数集合 $\mathbb{Q}$。

从几何直观上分析，整数系具有“离散性”，而有理数系具有“**稠密性**”。既然 $\mathbb{Q}$ 是稠密的，粗粗想来，它似乎已经是完美的了，其实不然。例如，若用 c 表示边长为 1 的正方形的对角线的长度，这个 c 就无法用有理数来表示，我们已经在之前证明过。换句话说，有理数集合对于开方运算是不封闭的。所以，有理点虽然在坐标轴上密密麻麻，但并没有布满整条直线，其中留有许多“空隙”。

注意到有理数一定能表示成有限小数或无限循环小数，很自然会想到，扩充有理数集合最直接的方式之一，就是把所有无限不循环小数（称为无理数）吸纳进来，我们将全体有理数和全体无理数所构成的集合称为实数集 $\mathbb{R}$。无理点确实填补了有理点在坐标轴上的所有“空隙”，即实数铺满了整个数轴。这样，每个实数都可以在坐标轴上找到自己的对应点，而坐标轴上的每个点又可以通过自己的坐标表示唯一一个实数。实数集合的这一性质称为实数系 $\mathbb{R}$ 的“连续性”。为了强调实数系所特有的这种连续性，$\mathbb{R}$ 又被称为**实数连续统**，而那条表示实数全体的坐标轴又称为**数轴**。

实数系的连续性是分析学的基础，对于我们将要学习的极限论、微积分乃至整个分析学具有无比的重要性。可以说，正是因为有了连续性，实数系才成为数学分析课程的“舞台”。

实数系 $\mathbb{R}$ 的连续性，从几何角度理解，就是实数全体布满整个数轴而没有“空隙”，但从分析学角度阐述，则有多种相互等价的表述方式。在本节中将要讲述的“确界存在定理”就是实数系 $\mathbb{R}$ 连续性的表述之一。

注意，实数集也具有稠密性，即在任何两个不同的实数之间，必然存在另一个实数（实际上存在无穷多个实数）。它的逆否命题“如果在两个实数之间不存在另一个实数，那么它们必定是同一个实数”可以带来一个简单且有趣的结论：$0.\dot{9} = 1$。显然我们无法找到一个介于 $0.\dot{9}$ 和 1 之间的数。事实上，在严格的实数理论中，一个小数表示的定义是：

$$0.a_1a_2a_3... = \sum_{n=1}^{\infty} \frac{a_n}{10^n}$$

对于 $0.\dot{9}$：

$$0.\dot{9} = \sum_{n=1}^{\infty} \frac{9}{10^n} = 9 \times \sum_{n=1}^{\infty} \left(\frac{1}{10}\right)^n = 9 \times \frac{\frac{1}{10}}{1-\frac{1}{10}} = 1$$

从数学结构的角度看，实数系是一个**有序域**，并且满足：

1. **阿基米德性质**（Archimedean property）（对于任意正实数 x, y，存在自然数 n 使得 $nx > y$）
2. **完备性**（completeness）（柯西序列收敛，或具有最小上界性质）

本章我们将重点关注实数系的完备性，在之后说明实数系的完备性与其连续性是等价的。实数系的稠密性可以从阿基米德性质配合有理数的存在性推出。

4.1.2 Maximum and Minimum 最大数与最小数

下面我们讨论实数集的各种子集，简称为数集。

设 S 是一个数集, 如果 $\exists\xi \in S$, 使得 $\forall x \in S$, 有 $x \le \xi$, 则称 ξ 是数集 S 的**最大数**(maximum), 记为 $\xi = \max S$；如果 $\exists\eta \in S$，使得 $\forall x \in S$，有 $x \ge \eta$，则称 η 是数集 S 的**最小数**（minimum），记为 $\eta = \min S$。

当数集 S 是非空有限集时，$\max S$ 与 $\min S$ 显然存在，且 $\max S$ 是这有限个数中的最大者，$\min S$ 是这有限个数中的最小者。但是当 S 是无限集时，最大数及最小数就有可能不存在。

4.1.3 Supremum and Infimum 上确界与下确界

设 S 是一个非空数集，如果 $\exists M \in \mathbb{R}$，使得 $\forall x \in S$，有 $x \le M$，则称 M 是 S 的一个**上界**(upper bound)；如果 $\exists m \in \mathbb{R}$，使得 $\forall x \in S$，有 $x \ge m$，则称 m 是 S 的一个**下界**（lower bound）。当数集 S 既有上界，又有下界时，称 S 为**有界集**。显然，S 为有界集 $\Leftrightarrow \exists X > 0$，使得 $\forall x \in S$，有 $|x| \le X$。

设数集 S 有上界，记 U 为 S 的上界全体所组成的集合，则显然 U 不可能有最大数，我们会证明：U 一定有最小数。设 U 的最小数为 β，就称 β 为数集 S 的**上确界**（supremum），即最小上界（least upper bound），记为 $\beta = \sup S$。

从上面的叙述可知上确界 β 满足下述两个性质：

1. β 是数集 S 的上界。
2. 任何小于 β 的数不是数集 S 的上界。

又假若数集 S 有下界，记 L 为 S 的下界全体所组成的集合，则显然 L 不可能有最小数，同样可以证明：L 一定有最大数。设 L 的最大数为 α，就称 α 为数集 S 的**下确界**（infimum），即最大下界(greatest lower bound)，记为 $\alpha = \inf S$。

类似地，下确界 α 满足下述两个性质：

1. α 是数集 S 的下界。
2. 任何大于 α 的数不是数集 S 的下界。

S	上确界 (Sup)	最大值 (Max)	下确界 (Inf)	最小值 (Min)
$\{0, 1, 9, 7, 6, 1976\}$	1976	1976	0	0
$\{x \in \mathbb{R} : x < 1\}$	1	不存在	不存在	不存在
$\mathbb{R}$	不存在	不存在	不存在	不存在
$\varnothing$	不存在	不存在	不存在	不存在
$\{x \in \mathbb{R} : x \ge 1\}$	不存在	不存在	1	1

表 5: 示例集合的上确界、最大值、下确界与最小值

定理 4.1 (确界存在定理——实数系连续性定理) 非空有上界的数集必有上确界，非空有下界的数集必有下确界。

定理 4.2 非空有界数集的上下确界是唯一的。

确界存在定理反映了实数系连续性这一基本性质，这可以从几何上加以理解：假若实数全体不能布满整条数轴而是留有“空隙”，则“空隙”左边的数集就没有上确界，“空隙”右边的数集就没有下确界。

例题 4.1 设 $T=\{x|x\in\mathbb{Q}\wedge x>0,x^2<2\}$，证明 T 在 $\mathbb{Q}$ 内没有上确界。

例题 4.2 证明 $\sqrt{6}$ 不是有理数。

例题 4.3 $\sqrt{3}+\sqrt{2}$ 是不是有理数？

例题 4.4 求下列数集的最大数、最小数，或证明它们不存在：*(1)* $A=\{x|x\geq 0\}$ *(2)* $B=\{\sin x|0<x<\frac{2\pi}{3}\}$ *(3)* $C=\{\frac{n}{m}|m,n\in\mathbb{N}^+\wedge n<m\}$。

例题 4.5 设数集 S 有上界，则数集 $T=\{x|-x\in S\}$ 有下界，且 $\sup S=-\inf T$。

例题 4.6 对任何非空数集 S，必有 $\sup S\geq\inf S$。当 $\sup S=\inf S$ 时，数集 S 有什么特点？

例题 4.7 设 $S=\{x|x\in\mathbb{Q}\wedge x^2<3\}$，证明：*(1)* S 没有最大数与最小数；*(2)* S 在 $\mathbb{Q}$ 内没有上确界与下确界。

4.2 Limits of Sequences of Real Numbers 数列极限

4.2.1 Sequences in $\mathbb{R}$ and Their Limits 数列与数列极限

数列（sequence in $\mathbb{R}$）是指按正整数编了号的一串实数**序列**（sequence）：$x_1,x_2,...,x_n,...$，通常表示成 $\{x_n\}$，其中 x_n 称为该数列的**通项**。

注意，序列可以是在任意一般的**度量空间**（metric space）中。度量空间指的是装备了某种特定**距离函数**（distance function），或者叫**度量**（metric），的某个空间（space）。序列的每项自然取自该空间的元素，该空间可以是最普通的 $\mathbb{R}$，也可以是某个向量空间 V，也可以是 $\mathbb{R}^n$，也可以是闭区间上连续函数构成的空间 $C[a,b]$，也可以是 $\mathbb{Q}$，也可以是某个开区间 (a,b) 等等。而在分析序列的时候还需要考虑空间配备的距离函数是因为序列的收敛性、柯西性等重要性质的定义都涉及距离的度量，因此如何定义“距离”也非常重要。如果一个空间缺失距离函数，则不是一个度量空间，研究该空间内的序列也几乎没有意义。在本章中，由于我们讨论的场域是实数系 $\mathbb{R}$，我们主要研究 $\mathbb{R}$ 中的序列，也就是由实数构成的序列，为了之后称呼方便我们简称“数列”，其距离函数我们默认为最常规的欧几里得度量（Euclidean metric），即 $d(x,y)=|x-y|$。之后我们会交替使用“数列”和“$\mathbb{R}$ 中的序列”来表达同一个概念。

序列的有些结论是无关乎其所在度量空间而恒成立的，有些则只在部分特定的度量空间中成立（例如完备的度量空间），请务必区分！

定义 4.1 设 $\{x_n\}$ 是一给定数列，L 是一个实常数。如果对于任意给定的 $\varepsilon>0$，可以找到正整数 N，使得当 $n>N$ 时，成立 $|x_n-L|<\varepsilon$，则称数列 $\{x_n\}$ **收敛于**L（或 L 是数列 $\{x_n\}$ 的**极限**），记为 $\lim\limits_{n\to\infty}x_n=L$，有时也记为 $x_n\to L(n\to\infty)$。如果不存在实数 L 使 $\{x_n\}$ 收敛于 L，则称数列 $\{x_n\}$ **发散**。

- 从极限的定义可知，一个数列收敛与否，收敛于哪个数，与这一数列的**前面有限项无关**。也就是说，改变数列前面的有限项，不影响数列的收敛性。例如数列 $10,100,1000,10000,\frac{1}{5},\frac{1}{6},...,\frac{1}{n},...$ 的极限仍然是 0。

- 注意该定义中很重要的一点是“存在 $L\in\mathbb{R}$”，放在更一般的情境下也就是极限必须在序列所在的空间中。这对于我们现在所讨论的 $\mathbb{R}$ 中的序列而言似乎不用担心，因为一串实数趋近于的目标（如果有）好像必然也是实数。事实上在 $\mathbb{R}$ 中我们的确不用担心这个问题，之后我们会讨论到实数系在常规欧几里得度量下具有完备性，柯西数列集就是收敛数列集。

例题 4.8 证明数列 $\{\frac{n}{n+3}\}$ 的极限为 1。

数列 $\{n^2\}: 1, 4, 9, ..., n^2, ...$ 与数列 $\{(-1)^n\}: -1, 1, -1, 1, ...$ 是发散数列。事实上，随着 n 的增加，$x_n = n^2$ 无限增大，而 $x_n = (-1)^n$ 不断地在 1 与-1 两个数值上跳跃，显然不能满足收敛数列的条件。

在收敛的数列中，我们称极限为 0 的数列为**无穷小量**。例如数列 $\{\frac{1}{n}\}, \{\frac{(-1)^n}{n^2+1}\}$ 都是无穷小量。要注意，无穷小量是一个变量，而不是一个"非常小的量"。常数列 $0, 0, 0, ..., 0, ...$ 是一个特殊的无穷小量。

根据数列极限的定义，可直接得到 $\lim_{n\to\infty} x_n = a \Leftrightarrow \{x_n - a\}$ 是无穷小量。

例题 4.9 证明 $\{q^n\}(0 < |q| < 1)$ 是无穷小量。

根据数列极限的定义来证明某一数列收敛，其关键是对任意给定的 $\varepsilon > 0$ 寻找正整数 N。在上面的两例题中，N 都是通过解不等式 $|x_n - a| < \varepsilon$ 而得出的。但在大多数情况下，这个不等式并不容易解。实际上，数列极限的定义并不要求取到最小的或最佳的正整数 N，所以在证明中常常对 $|x_n - a|$ 适度地做一些放大处理，这是一种常用的技巧。

4.2.2 Properties of Sequence Limits 数列极限的性质

定理 4.3 收敛数列的极限是唯一的。(事实上把"数列"改成"序列"仍然成立)

对于数列 $\{x_n\}$，如果存在实数 M，使数列的所有的项都满足 $x_n \leq M, n = 1, 2, 3, ...$，则称 M 是数列 $\{x_n\}$ 的上界。如果存在实数 m，使数列的所有的项都满足 $m \leq x_n, n = 1, 2, 3, ...$，则称 m 是数列 $\{x_n\}$ 的下界。一个数列若既有上界又有下界，则称之为**有界数列**。

定理 4.4 收敛数列必有界。(事实上把"数列"改成"序列"仍然成立)

定理 4.5 设数列 $\{x_n\}, \{y_n\}$ 均收敛，若 $\lim_{n\to\infty} x_n = a, \lim_{n\to\infty} y_n = b$，且 $a < b$，则存在正整数 N，当 $n > N$ 时，成立 $x_n < y_n$。

定理 4.6 (夹逼定理) 若三个数列 $\{x_n\}, \{y_n\}, \{z_n\}$ 从某项开始成立 $x_n \leq y_n \leq z_n, n > N_0$，且 $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = a$，则 $\lim_{n\to\infty} y_n = a$。

注意，以上两个定理以及之后我们会学到的非常重要的单调有界数列收敛定理都是没有一般序列版本的推广的，其原因在于 $\mathbb{R}$ 是具有序完备性的，而一般的度量空间并不天然具备这种"序"结构。注意，"单调"是一个序关系概念，一般的度量空间 (X, d) 只是一个配备了"距离"概念的集合，它本身没有一个天然的、全体的序关系"$\leq$ 或者 $<$"。

4.2.3 Operations on Sequence Limits 数列极限的四则运算

定理 4.7 设 $\lim_{n\to\infty} x_n = a, \lim_{n\to\infty} y_n = b$，则

1. $\lim_{n\to\infty} (\alpha x_n + \beta y_n) = \alpha a + \beta b$
2. $\lim_{n\to\infty} (x_n y_n) = ab$
3. $\lim_{n\to\infty} (\frac{x_n}{y_n}) = \frac{a}{b} (b \neq 0)$

4.3 Infinitely Large Quantities 无穷大量

4.3.1 Infinitely Large Quantities 无穷大量

定义 4.2 若对于任意给定的 $G>0$，可以找到正整数 N，使得当 $n>N$ 时成立 $|x_n|>G$，则称数列 $\{x_n\}$ 是无穷大量，记为 $\lim_{n\to\infty} x_n=\infty$。

如果无穷大量从某一项开始都是正的（或负的），则称其为正无穷大量（或负无穷大量），统称为定号无穷大量。

定理 4.8 设 $x_n\neq 0$，则 $\{x_n\}$ 是无穷大量的等价条件是 $\{\frac{1}{x_n}\}$ 是无穷小量。

关于无穷大量的运算，如下的性质是显然的：同号无穷大量之和仍然是该符号的无穷大量，而异号无穷大量之差是无穷大量，其符号与被减无穷大量的符号相同；无穷大量与有界量之和或差仍然是无穷大量；同号无穷大量之积为正无穷大量，而异号无穷大量之积为负无穷大量。

4.3.2 Indeterminate Forms 待定型/未定式

若分别以 $+\infty,-\infty,\infty,0$ 表示正无穷大量，负无穷大量，无穷大量与无穷小量，则很容易举出例子说明，如 $\infty\pm\infty,(+\infty)-(+\infty),(+\infty)+(-\infty),0\cdot\infty,\frac{0}{0},\frac{\infty}{\infty}$ 等极限，其结果可以是无穷小量，或非零极限，或无穷大量，也可以没有极限。我们称这种类型的极限为**待定型**。

下面介绍的 Stolz 定理将为求某些类型的待定型极限带来很大的方便。在叙述定理前，先给出单调数列的定义。

定义 4.3 如果数列 $\{x_n\}$ 满足 $x_n\leq x_{n+1}, n=1,2,3,...$，则称 $\{x_n\}$ 为**单调增加数列**；若进一步满足 $x_n<x_{n+1}, n=1,2,3,...$，则称 $\{x_n\}$ 为**严格单调增加数列**。

可以类似地定义**单调减少数列**和**严格单调减少数列**。

因为数列前面有限项的变化不会影响它的收敛性，所以下面我们谈到单调数列的场合，都可以将“从某一项开始为单调的数列”统统包括在内。

定理 4.9 (*Stolz* 定理) 设 $\{y_n\}$ 是严格单调增加的正无穷大量，且

$$\lim_{n\to\infty}\frac{x_n-x_{n-1}}{y_n-y_{n-1}}=a$$

其中 a 可以为有限量，$+\infty$ 与 $-\infty$，则

$$\lim_{n\to\infty}\frac{x_n}{y_n}=a$$

4.4 Convergence Criteria 收敛准则

4.4.1 Monotone Convergence Theorem 单调有界数列收敛定理

知道了收敛数列必定有界，而有界数列不一定收敛的结论后，很自然会产生这样两个问题：

1. 对有界数列加上什么条件，就可以保证它必定收敛？

2. 若不对有界数列加任何条件，则能得到怎样的（比收敛稍弱一些的）结论？

我们先来回答第一个问题：只要对有界数列加上单调性，那么它就一定收敛，而极限就是该数列所构成的数集的上确界或下确界。

定理 4.10 单调有界数列必定收敛。

4.4.2 Nested Intervals Theorem 闭区间套定理

定义 4.4 如果一列闭区间 $\{[a_n, b_n]\}$ 满足条件
(1)$[a_{n+1}, b_{n+1}] \subset [a_n, b_n], n = 1, 2, 3, ...$
(2) $\lim\limits_{n\to\infty}(b_n - a_n) = 0$
则称这列闭区间形成一个**闭区间套**。

定理 4.11 (闭区间套定理) 如果 $\{[a_n, b_n]\}$ 形成一个闭区间套，则存在唯一的实数 ξ 属于所有的闭区间 $[a_n, b_n]$，且 $\xi = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$。

4.4.3 Subsequence in $\mathbb{R}$ 子数列

为了回答本节一开始提出的第二个问题，先引入子列的概念。

定义 4.5 设 $\{x_n\}$ 是一个数列，而 $n_1 < n_2 < ... < n_k < n_{k+1} < ...$ 是一列严格单调增加的正整数，则 $x_{n_1}, x_{n_2}, ..., x_{n_k}, ...$ 也形成一个数列，称为数列 $\{x_n\}$ 的子数列，记为 $\{x_{n_k}\}$。

定理 4.12 若数列 $\{x_n\}$ 收敛于 a，则它的任何子数列 $\{x_{n_k}\}$ 也收敛于 a。(一般度量空间普遍结论)

上述定理的推论经常被用来判断一个数列的发散。

推论 4.1 若存在数列 $\{x_n\}$ 的两个子数列分别收敛于不同的极限，则数列 $\{x_n\}$ 必定发散。

4.4.4 Bolzano-Weierstrass Theorem 波尔扎诺-魏尔斯特拉斯定理

现在我们来回答第二个问题：如果只考虑有界数列，则只能得到下面稍弱的结论。

定理 4.13 (*Bolzano-Weierstrass* 定理) 有界数列必有收敛子数列。

- 该定理在一般的度量空间中**不成立**！

当数列无界时，也有与上述定理相对应的结论。

定理 4.14 若 $\{x_n\}$ 是一个无界数列，则存在子数列 $\{x_{n_k}\}$，使得 $\lim\limits_{k\to\infty} x_{n_k} = \infty$。

4.4.5 Cauchy's Convergence Criterion 柯西收敛原理

从序列本身的特征直接判断它是否收敛是个很有意义的重要问题。但"单调有界数列必定收敛"这一定理只是给出了判断数列收敛的一个充分而非必要的条件，事实上，许多收敛的数列并非是单调的。所以，有必要从数列本身出发来寻找其收敛的充分必要条件，为此，先引进柯西数列（Cauchy sequence in $\mathbb{R}$）的概念。

定义 4.6 如果数列 $\{x_n\}$ 具有以下特性：对于任意给定的 $\varepsilon > 0$，存在正整数 N，使得当 $n, m > N$ 时成立 $|x_n - x_m| < \varepsilon$，则称数列 $\{x_n\}$ 是一个**柯西数列**。

- 在很多（尤其是比较早期的）中文数学教材和文献中，"柯西数列" 经常被翻译为 "基本数列"。

定理 4.15 (*Cauchy* 收敛原理) 数列收敛的充分必要条件是：它是柯西数列。

- 注意，**收敛是一个比柯西更强的条件**，收敛蕴含着柯西。易证得**收敛列一定是柯西列**，这是一个在任何度量空间都成立的普遍真理：一个序列如果已经无限接近某个固定的点（收敛），那么它的项彼此之间也必然会无限接近（柯西）。这是由极限的定义和三角不等式直接保证的，不依赖于集合 X 的特性或者度量 d 的选择。但事实上，**柯西列并不是在每一个度量空间中都保证一定收敛的**。可以这样去理解为何收敛是一个相较柯西更强的条件：如果一个序列的项最终会无限接近某个固定的极限点 L，并且这个点 L 必须在所讨论的空间内，那么这个序列是收敛的。而如果一个序列的项彼此之间会无限接近，那么它就是柯西列。它不要求极限点 L 存在于当前空间，只关心序列自身的"内部"性质。(这也是"基本列"翻译的由来）因此柯西性不依赖环境，而收敛性依赖环境。如果一个度量空间中的每一个柯西列都在该空间内收敛，那么这个度量空间被称为**完备**的。在完备空间中，柯西列与收敛列是等价的。实数集 $\mathbb{R}$ 在通常的欧几里得度量下（也就是 Cauchy 收敛原理所讨论的情境）是完备的。相反，有理数集 $\mathbb{Q}$ 在欧几里得度量下是不完备的。例如，用有理数去逼近 $\sqrt{2}$ 的有理数列是柯西列，但其极限 $\sqrt{2}$ 不在 $\mathbb{Q}$ 内，所以该柯西列在 $\mathbb{Q}$ 中不收敛。再例如 $\{(1+\frac{1}{n})^n\}$ 是由有理数构成的柯西列，但其极限 e 也不是有理数。另一个简单的不完备度量空间的例子是装备了欧式度量的开区间 $(0,1)$，数列 $\{\frac{1}{2n}\}$ 是该空间的一个柯西数列，但其极限 0 却不在 $(0,1)$ 内，因此它不是一个收敛数列。

- 有学生可能会有疑问：为什么在收敛的定义里，极限一定得包含在所讨论的空间里呢？既然收敛是指无限趋近但永远达不到，那为啥我们要关心一个永远达不到的点是否被包含在内呢？没错，当我们说一个序列在空间 X 中收敛时，我们不仅仅是在描述这个序列"无限趋近"的行为，我们还在做一个存在性声明——这个趋近的最终目标也存在于我们当前讨论的"宇宙"（即空间 X）中。这很重要是因为这保证了在这个空间内部进行运算和推理的封闭性。如果一个序列的极限会跑到空间外面去，那么在这个空间内讨论"极限"就会不断遇到"未定义"的情况，许多重要的定理（如连续函数的极限仍是函数值、微积分基本定理等）将无法成立。

- 当某个教材说"实数系具有完备性"时，这是一个默认的简写。它完整的表述应是："配备了通常的欧几里得度量的实数集 $\mathbb{R}$ 构成一个完备的度量空间。"实数集配上其他度量不一定完备。

- 在了解了不完备度量空间中的柯西不收敛列的例子后，有学生可能会有疑问：看起来柯西与收敛的唯一差别就是柯西列所趋近的那个"理想点（可以理解为一种"广义极限"）"不在所讨论空间内，其它序列的行为都是一样的？这个问题等价于是否所有不收敛的柯西列最终都会无限趋近于一个该空间内不存在的"理想点"？因为如果是的话，那柯西列距离成为收敛列就仅仅只差该"理想点"存在于所讨论空间中而已。这个问题可以进一步等价于是否存在柯西列虽然内部项在彼此无限接近，但它们不会无限接近某个外部（即便不存在于此空间）的固定点，而是（也许）只是不断波动？答案是不存在！这个问题其实恰好对应了度量空间的完备化这个重要的思想。一个柯西列可能在自己家里（原空间）找不到归宿，但它总是可以在一个更大的、完整的"家园"（完备化空间）里找到唯一的归宿。不存在一个哪都去不了的、纯粹的"流浪"柯西列。这意味着，对于任何度量空间，我们都可以将其完备化（详见下条）。这句话等价于任何柯西列都存在"理想点"。因此我们可以根据这个"理想点"是否在空间里给一个空间的柯西列二分，在的即也是收敛列，不在的即为仅柯西不收敛，但确实有这么一个被无限趋近于的固定点在原空间外部。

- 任何一个度量空间都可以通过"添加所有缺失的极限点"来使其变得完备，这个过程称为完备化。正式来说，对于任何度量空间 X，都存在一个完备的度量空间 $\overline{X}$（称为 X 的完备化），使得 X 是 $\overline{X}$ 的稠密子集。在完备化空间 $\overline{X}$ 中，每一个柯西列都收敛。如果原空间 X 中一

个柯西列不收敛（于 X 中），那是因为它的极限点在 $\overline{X}\setminus X$ 中，即它在 X 的"外部"。$\mathbb{Q}$ 的完备化是 $\mathbb{R}$，$(0,1)$ 的完备化是 $[0,1]$。在完备化的空间中，原来在不完备空间中的那些"不收敛的柯西列"都找到了它们的归宿（极限点），从而变成了收敛列。

4.4.6 Fundamental Theorems of the Real Number System 实数系的基本定理

$$\begin{array}{c}\text{确界存在定理}\\ \downarrow\\ \text{单调有界数列收敛定理}\\ \downarrow\\ \text{闭区间套定理}\\ \downarrow\\ \text{Bolzano-Weierstrass 定理}\\ \downarrow\\ \text{Cauchy 收敛原理}\end{array}$$

上面是本节一些重要定理之间的逻辑推理关系，可以看出，实数系的连续性包含了实数系的完备性。事实上，我们可以证明实数系的完备性也包含了实数系的连续性。也就是说，在实数系中完备性与连续性这两个概念是等价的。

定理 4.16 实数系的完备性等价于实数系的连续性。

因此上述五个定理是等价的，所以这五个定理中的每一个都可以称为是实数系的基本定理。

5 Limits of Functions and Continuity 函数极限与连续函数

5.1 Limits of Functions 函数极限

5.1.1 Definition of the Limit of a Function 函数极限的定义

定义 5.1 设函数 $y=f(x)$ 在点 x_0 的某个去心邻域中有定义,即存在 $\rho>0$,使 $O(x_0,\rho)\backslash\{x_0\}\subset D_f$。如果存在实数 A，对于任意给定的 $\varepsilon>0$，可以找到 $\delta>0$，使得当 $0<|x-x_0|<\delta$ 时，成立 $|f(x)-A|<\varepsilon$，则称 A 是函数 $f(x)$ 在点 x_0 的极限，记为 $\lim\limits_{x\to x_0}f(x)=A$ 或 $f(x)\to A(x\to x_0)$。如果不存在具有上述性质的实数 A，则称函数 $f(x)$ 在点 x_0 的极限不存在。

- 注意，讨论函数在一个点 x_0 处的极限 A 并不要求函数在该点处有定义。如果 A 存在，它也自然未必在该函数的值域中。

- 事实上，极限代表 x 无限趋近 x_0 最后指向的那个理想点，它只依赖于 x_0 附近点的函数值，这与函数就在 $x=x_0$ 时的取值是多少是两件事，即便有时它们的答案是一样的。假设函数在 x_0 处有定义 $y(x_0)=y_0$，并且函数在 x_0 处的极限 A 存在，不意味着 $A=y_0$。因此，极限与函数值是互相独立的概念。例如在思考 $\lim\limits_{x\to1}\frac{x^2-1}{x-1}$ 时，如果混淆两者的概念，会觉得极限似乎不存在，因为极限似乎是一个分母为 0 的值。但事实上讨论 x 在 1 处的极限却是完全合法的，因为函数在该点的去心邻域内显然有定义。"当 $x=1$ 时函数的分母为 0"只是为我们额外提供了"该函数在 $x=1$ 时无定义"的独立信息。当我们求极限的时候，我们是在求一个"理想

点”，这个点完全可能不在函数的值域中；当我们求取值的时候，我们是在找值域内的某一点。同理，在求极限的时候我们是可以对上述函数做分子分母同时约去 $x-1$ 这一项的操作的，由于 x 永远取不到 1，因此分子分母同时约去一个非零项是合法的。

5.1.2 Properties of Function Limits 函数极限的性质

（1）极限的唯一性

定理 5.1 设 A 与 B 都是函数 $f(x)$ 在点 x_0 的极限，则 $A=B$。

（2）局部保序性

定理 5.2 若 $\lim\limits_{x\to x_0} f(x)=A, \lim\limits_{x\to x_0} g(x)=B$，且 $A>B$，则存在 $\delta>0$，当 $0<|x-x_0|<\delta$ 时，成立 $f(x)>g(x)$。

（3）夹逼性

定理 5.3 若存在 $r>0$，使得当 $0<|x-x_0|<r$ 时，成立 $g(x)\leq f(x)\leq h(x)$，且 $\lim\limits_{x\to x_0} g(x)=\lim\limits_{x\to x_0} h(x)=A$，则 $\lim\limits_{x\to x_0} f(x)=A$。

5.1.3 Operations on Function Limits 函数极限的四则运算

定理 5.4 设 $\lim\limits_{x\to x_0} f(x)=A, \lim\limits_{x\to x_0} g(x)=B$，则

1. $\lim\limits_{x\to x_0}(\alpha f(x)+\beta g(x))=\alpha A+\beta B$
2. $\lim\limits_{x\to x_0}(f(x)g(x))=AB$
3. $\lim\limits_{x\to x_0}\frac{f(x)}{g(x)}=\frac{A}{B}(B\neq 0)$

5.1.4 Relationship Between Function Limits and Sequence Limits 函数极限与数列极限的关系

定理 5.5（*Heine* 定理） $\lim\limits_{x\to x_0} f(x)=A$ 的充分必要条件是：对于任意满足条件 $\lim\limits_{n\to\infty} x_n=x_0$，且 $x_n\neq x_0$ 的数列 $\{x_n\}$，相应的函数值数列 $\{f(x_n)\}$ 成立 $\lim\limits_{n\to\infty} f(x_n)=A$。

这一性质被经常用于证明某个函数极限的不存在性。

5.1.5 One-Sided Limits 单侧极限

在函数极限 $\lim\limits_{x\to x_0} f(x)=A$ 的定义中，自变量 x 可以按任意的方式趋于 x_0。但有时候，$f(x)$ 只在 x_0 的一侧（左侧或右侧）有定义，或者需要分别研究 $f(x)$ 在 x_0 两侧的性态，这就有必要引入单侧极限的概念。

定义 5.2 设 $f(x)$ 在 $(x_0-\rho,x_0)$ 有定义（$\rho>0$）。如果存在实数 B，对于任意给定的 $\varepsilon>0$，可以找到 $\delta>0$，使得当 $-\delta<x-x_0<0$ 时，成立 $|f(x)-B|<\varepsilon$，则称 B 是函数 $f(x)$ 在点 x_0 的**左极限**，记为 $\lim\limits_{x\to x_{0-}} f(x)=f(x_{0-})=B$。

定义 5.3 设 $f(x)$ 在 $(x_0,x_0+\rho)$ 有定义（$\rho>0$）。如果存在实数 C，对于任意给定的 $\varepsilon>0$，可以找到 $\delta>0$，使得当 $0<x-x_0<\delta$ 时，成立 $|f(x)-C|<\varepsilon$，则称 C 是函数 $f(x)$ 在点 x_0 的**右极限**，记为 $\lim\limits_{x\to x_{0+}} f(x)=f(x_{0+})=C$。

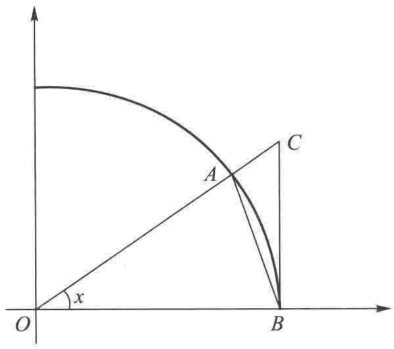


图 1: 第一重要极限的经典证明

显然，函数 $f(x)$ 在 x_0 极限存在的充分必要条件是 $f(x)$ 在 x_0 的左极限与右极限存在并且相等。

5.1.6 Extensions of the Limit Concept 函数极限定义的扩充

在本节中，函数极限是对 $\lim\limits_{x\to x_0} f(x)=A$ 定义的，表述为 $\forall\varepsilon>0,\exists\delta>0,\forall x(0<|x-x_0|<\delta):|f(x)-A|<\varepsilon$，其中 x_0,A 都是有限实数。

但实际上，自变量的极限过程有 6 种情况：$x\to x_0,x_{0^+},x_{0^-},\infty,+\infty,-\infty$，函数值的极限有 4 种情况：$f(x)\to A,\infty,+\infty,-\infty$。

5.1.7 Two Important Limits 第一重要极限与第二重要极限

第一重要极限:

$$\lim_{x\to 0}\frac{\sin x}{x}=1$$

这里介绍两种计算方法，第一种是最经典的利用夹逼定理与几何面积法。（如图）设 $\angle AOB$ 的弧度为 x，$0<x<\frac{\pi}{2}$，由于 $\triangle OAB$ 面积 $<$ 扇形 OAB 面积 $<\triangle OBC$ 面积，可以得到

$$\sin x<x<\tan x,0<x<\frac{\pi}{2}$$

从而有

$$\cos x<\frac{\sin x}{x}<1,0<x<\frac{\pi}{2}$$

应用极限的夹逼性，得到答案等于 1。

第二种方法是后面会学到的洛必达法则，只需要一步即可得到答案。

比较容易混淆的是下面这个极限：

$$\lim_{x\to\infty}\frac{\sin x}{x}=0$$

这个极限用夹逼定理就非常简单了：

$$\frac{-1}{x}\leq\frac{\sin x}{x}\leq\frac{1}{x}$$

$$\lim_{x\to\infty}\frac{-1}{x}\leq\lim_{x\to\infty}\frac{\sin x}{x}\leq\lim_{x\to\infty}\frac{1}{x}$$

容易跟 $\frac{\sin x}{x}$ 函数图像混淆的还有 $\sin\frac{1}{x}$ 的函数图像，注意 $\lim\limits_{x\to 0}\sin\frac{1}{x}$ 相当于 $\lim\limits_{t\to\infty}\sin t$，函数在 1 和 -1 之间无限次振荡，不趋于任何固定值，极限不存在。但思考下面这个变体：

$$\lim_{x\to 0}x\sin\frac{1}{x}=0$$

$$\lim_{x\to\infty} x\sin\frac{1}{x}=1$$

这个函数作为“振荡但有极限”的例子也非常有名。事实上，它可以通过令 $\frac{1}{x}=t$ 实现与第一重要极限的转化。

例题 5.1 (第一重要极限的衍生 1)

$$\begin{aligned}
\lim_{x\to 0}\frac{\tan x}{x} &= \lim_{x\to 0}\frac{\sin x}{x\cos x}\\
&= \lim_{x\to 0}\left(\frac{\sin x}{x}\cdot\frac{1}{\cos x}\right)\\
&= \left(\lim_{x\to 0}\frac{\sin x}{x}\right)\cdot\left(\lim_{x\to 0}\frac{1}{\cos x}\right)\\
&= 1\cdot\frac{1}{1}\\
&= 1
\end{aligned}$$

例题 5.2 (第一重要极限的衍生 2)

$$\begin{aligned}
\lim_{x\to 0}\frac{1-\cos x}{x^2} &= \lim_{x\to 0}\frac{1-\cos x}{x^2}\cdot\frac{1+\cos x}{1+\cos x} && \text{(分子有理化)}\\
&= \lim_{x\to 0}\frac{1-\cos^2 x}{x^2(1+\cos x)}\\
&= \lim_{x\to 0}\frac{\sin^2 x}{x^2(1+\cos x)} && \text{(利用恒等式}\sin^2 x+\cos^2 x=1\text{)}\\
&= \lim_{x\to 0}\left(\frac{\sin x}{x}\right)^2\cdot\frac{1}{1+\cos x}\\
&= \left(\lim_{x\to 0}\frac{\sin x}{x}\right)^2\cdot\lim_{x\to 0}\frac{1}{1+\cos x} && \text{(极限的乘法法则)}\\
&= (1)^2\cdot\frac{1}{1+1} && \text{(已知}\lim_{x\to 0}\frac{\sin x}{x}=1\text{)}\\
&= 1\cdot\frac{1}{2}\\
&= \frac{1}{2}
\end{aligned}$$

第二重要极限:

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e$$

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}}=e$$

例题 5.3

$$\begin{aligned}
\lim_{n\to\infty}\left(1-\frac{3}{n}\right)^{5n} &= \lim_{n\to\infty}\left[\left(1+\frac{-3}{n}\right)^n\right]^5\\
&= \left[\lim_{n\to\infty}\left(1+\frac{-3}{n}\right)^n\right]^5\\
&= \left(e^{-3}\right)^5\\
&= e^{-15}
\end{aligned}$$

第二重要极限（拓展形式）:

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^{bx}=e^{ab}$$

5.2 Continuous Functions 连续函数

5.2.1 Definition of Continuity 连续函数的定义

定义 5.4 设函数 $f(x)$ 在点 x_0 的某个邻域中有定义，并且成立 $\lim\limits_{x\to x_0} f(x) = f(x_0)$，则称函数 $f(x)$ 在点 x_0 **连续**，而称 x_0 是函数 $f(x)$ 的**连续点**。

- 这个定义告诉我们函数 $f(x)$ 在点 x_0 处连续等价于下面三个条件全部成立：
  1. $f(x)$ 在点 x_0 处有定义（即 $x_0 \in dom(f)$）；
  2. $f(x)$ 在点 x_0 处的极限存在；
  3. $\lim\limits_{x\to x_0} f(x) = f(x_0)$。

定义 5.5 若函数 $f(x)$ 在区间 (a,b) 的每一点都连续，则称函数 $f(x)$ 在**开区间**(a,b) **上连续**。

为了讨论函数在闭区间上的连续性，需要**单侧连续**的概念。

定义 5.6 若 $\lim\limits_{x\to x_{0-}} f(x) = f(x_0)$，则称函数 $f(x)$ 在 x_0 **左连续**；若 $\lim\limits_{x\to x_{0+}} f(x) = f(x_0)$，则称函数 $f(x)$ 在 x_0 **右连续**。

定义 5.7 若 $f(x)$ 在 (a,b) 连续，且在左端点 a 右连续，在右端点 b 左连续，则称函数 $f(x)$ **在闭区间**$[a,b]$ **上连续**。

- 区间连续的几何含义是：图形是"一笔画"成的，没有断点或跳跃。

5.2.2 Operations on Continuous Functions 连续函数的四则运算

根据函数极限的四则运算，对于连续函数，也有下述运算规则：

设 $\lim\limits_{x\to x_0} f(x) = f(x_0)$ $\lim\limits_{x\to x_0} g(x) = g(x_0)$，则

1. $\lim\limits_{x\to x_0} (\alpha f(x) + \beta g(x)) = \alpha f(x_0) + \beta g(x_0)$
2. $\lim\limits_{x\to x_0} (f(x)g(x)) = f(x_0)g(x_0)$
3. $\lim\limits_{x\to x_0} \frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x_0)}(g(x_0) \neq 0)$

由上述运算法则，设有有限个函数在某区间连续，则它们之间进行有限次加、减、乘、除四则运算，所得到的函数在该区间除去使分母为零的点后余下的范围连续。

对于常数函数 $f(x) = c$ 与恒等函数 $g(x) = x$，容易从定义出发证明它们的连续性，然后由上述的连续函数的四则运算规律，可以得到

（1）任意多项式 $p_n(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ 在 $(-\infty, +\infty)$ 上连续；

（2）任意有理函数 $Q(x) = \frac{a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + ... + b_1 x + b_0}$ 在其定义域上连续。

5.2.3 Types of Discontinuities 不连续点类型

函数的**不连续点**又称**间断点**。通常将不连续点分成三类。

第一类不连续点：函数在该点的左、右极限都存在但不相等。这一类不连续点又称**跳跃点**，而右极限与左极限之差为函数在该点的**跃度**。

第二类不连续点：函数在该点的左、右极限至少有一个不存在。

第三类不连续点：函数在该点的左、右极限都存在而且相等，但不等于函数在该点的取值或者函数在该点压根无定义。这类不连续点可以通过重新定义在该点的函数值，使之成为函数的连续点，因此第三类不连续点又称为**可去不连续点**或**可去间断点**。

定理 5.6 单调函数的不连续点一定是跳跃点。

5.2.4 Continuity of Inverse Functions 反函数连续性定理

之前我们曾介绍过逆映射，对于函数来说，与之相对应的就是**反函数**。

定理 5.7 (反函数存在性定理) 若函数 $y = f(x), x \in D_f$ 是严格单调增加（减少）的，则存在它的反函数 $x = f^{-1}(y), y \in R_f$，并且 $f^{-1}(y)$ 也是严格单调增加（减少）的。

定理 5.8 (反函数连续性定理) 设函数 $y = f(x)$ 在闭区间 $[a, b]$ 上连续且严格单调增加，$f(a) = \alpha, f(b) = \beta$，则它的反函数 $x = f^{-1}(y)$ 在 $[\alpha, \beta]$ 连续且严格单调增加。

5.2.5 Continuity of Composite Functions 复合函数的连续性

设 $\lim\limits_{x\to x_0} g(x) = u_0, \lim\limits_{u\to u_0} f(u) = A$，对于复合函数 $f \circ g(x)$，我们不能得出 $\lim\limits_{x\to x_0} f \circ g(x) = A$ 的结论。但是当 f 与 g 都是连续函数时，则上述的结论是成立的。

定理 5.9 若 $u = g(x)$ 在点 x_0 连续，$g(x_0) = u_0$，又 $y = f(u)$ 在点 u_0 连续，则复合函数 $y = f \circ g(x)$ 在点 x_0 连续。

定理 5.10 一切初等函数在其定义区间上连续。

5.3 Orders of Infinitesimals and Infinitely Large Quantities 无穷小量与无穷大量的阶

5.3.1 Comparison of Infinitesimals 无穷小量的比较

与数列极限类似，在函数极限中同样也有无穷小量与无穷大量的概念，这里先讨论无穷小量。

定义 5.8 若 $\lim\limits_{x\to x_0} f(x) = 0$，则称当 $x \to x_0$ 时，$f(x)$ 是**无穷小量**。

无穷小量是以零为极限的变量。

设 $u(x), v(x)$ 是两个变量，当 $x \to x_0$ 时，它们都是无穷小量。为了比较两者趋于零的速度快慢，我们讨论 $\frac{u(x)}{v(x)}$ 的极限情况：

（1）若比值的极限是 0，我们称当 $x \to x_0$ 时，$u(x)$ 是关于 $v(x)$ 的**高阶无穷小量**（或 $v(x)$ 是关于 $u(x)$ 的**低阶无穷小量**），记为 $u(x) = o(v(x))(x \to x_0)$。

（2）若存在 $A > 0$，当 x 在 x_0 的某个去心邻域中，成立 $\left|\frac{u(x)}{v(x)}\right| \le A$，则称当 $x \to x_0$ 时，$\frac{u(x)}{v(x)}$ 是有界量，记为 $u(x) = O(v(x))(x \to x_0)$。若又存在 $a > 0$，当 x 在 x_0 的某个去心邻域中，成立 $a \le \left|\frac{u(x)}{v(x)}\right| \le A$，则称当 $x \to x_0$ 时，$u(x)$ 与 $v(x)$ 是**同阶无穷小量**。显然，若 $\lim\limits_{x\to x_0} \frac{u(x)}{v(x)} = c \ne 0$，则 $u(x)$ 与 $v(x)$ 必是同阶无穷小量。

（3）若 $\lim\limits_{x\to x_0} \frac{u(x)}{v(x)} = 1$，称当 $x \to x_0$ 时，$u(x)$ 与 $v(x)$ 是**等价无穷小量**，记为 $u(x) \sim v(x)(x \to x_0)$。上式也可写成 $u(x) = v(x) + o(v(x))(x \to x_0)$。它表示当 $x \to x_0$ 时，$u(x)$ 与 $v(x)$ 并不一定相等，两者相差一个关于 $v(x)$ 的高阶无穷小量。

我们往往选取 $v(x) = (x - x_0)^k$ 作为与 $u(x)$ 进行比较的无穷小量（如果极限过程是 $x \to \infty$，则选取 $v(x) = \frac{1}{x^k}$），这样便于得出 $u(x)$ 作为无穷小量的确切阶数。

我们常用 $u(x)=o(1)(x\to x_0)$ 表示当 $x\to x_0$ 时，$u(x)$ 是无穷小量；用 $u(x)=O(1)(x\to x_0)$ 表示当 $x\to x_0$ 时，$u(x)$ 是有界量。

5.3.2 Comparison of Infinitely Large Quantities 无穷大量的比较

定义 5.9 若 $\lim\limits_{x\to x_0} f(x)=\infty$（或 $\pm\infty$），则称当 $x\to x_0$ 时，$f(x)$ 是**无穷大量**（或**正、负无穷大量**）。

设 $u(x),v(x)$ 是两个变量，当 $x\to x_0$ 时，它们都是无穷大量。为了比较两者趋于零的速度快慢，我们讨论 $\frac{u(x)}{v(x)}$ 的极限情况：

（1）若比值的极限是 ∞，我们称当 $x\to x_0$ 时，$u(x)$ 是关于 $v(x)$ 的**高阶无穷大量**（或 $v(x)$ 是关于 $u(x)$ 的**低阶无穷大量**）。

（2）若存在 $A>0$，当 x 在 x_0 的某个去心邻域中，成立 $\left|\frac{u(x)}{v(x)}\right|\le A$，则称当 $x\to x_0$ 时，$\frac{u(x)}{v(x)}$ 是有界量，记为 $u(x)=O(v(x))(x\to x_0)$。若又存在 $a>0$，当 x 在 x_0 的某个去心邻域中，成立 $a\le\left|\frac{u(x)}{v(x)}\right|\le A$，则称当 $x\to x_0$ 时，$u(x)$ 与 $v(x)$ 是**同阶无穷大量**。显然，若 $\lim\limits_{x\to x_0}\frac{u(x)}{v(x)}=c\neq 0$，则 $u(x)$ 与 $v(x)$ 必是同阶无穷大量。

（3）若 $\lim\limits_{x\to x_0}\frac{u(x)}{v(x)}=1$，称当 $x\to x_0$ 时，$u(x)$ 与 $v(x)$ 是**等价无穷大量**，记为 $u(x)\sim v(x)(x\to x_0)$。

5.3.3 Equivalent Quantities 等价量

所谓等价量，就是指等价无穷小量或等价无穷大量。在极限计算中，等价量起着举足轻重的作用。下面是几个重要的等价量（都是 $x\to 0$）：

- $\ln(1+x)\sim x$
- $e^x-1\sim x$
- $(1+x)^\alpha-1\sim \alpha x$
- $\sin x\sim x$

设一个变量是由几个相互不同阶的成分相加而成的，则当它是无穷大量时，它与阶数最高的那个无穷大量成分等价；当它是无穷小量时，它与阶数最低的那个无穷小量成分等价。

定理 5.11 设 $u(x),v(x)$ 和 $w(x)$ 在 x_0 的某个去心邻域 U 上有定义，且 $\lim\limits_{x\to x_0}\frac{v(x)}{w(x)}=1$，那么

1. 当 $\lim\limits_{x\to x_0}u(x)w(x)=A$ 时，$\lim\limits_{x\to x_0}u(x)v(x)=A$；
2. 当 $\lim\limits_{x\to x_0}\frac{u(x)}{w(x)}=A$ 时，$\lim\limits_{x\to x_0}\frac{u(x)}{v(x)}=A$。

必须注意的是，当计算中出现无穷小量或无穷大量相加或相减时，就不能不加考虑便用等价量直接进行代换。

5.4 Continuous Functions on a Closed Interval 闭区间上的连续函数

闭区间上的连续函数具有一些重要的性质，这些性质是开区间上的连续函数不一定具有的。它们在今后的学习中有重要的应用。

5.4.1 Boundedness Theorem 有界性定理

定理 5.12 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则它在 $[a,b]$ 上有界。

开区间上的连续函数就不一定是有界的。例如 $f(x)=\frac{1}{x}$ 在开区间 $(0,1)$ 上连续，但显然无界。

5.4.2 Extreme Value Theorem 最值定理/极值定理

定理 5.13 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则它在 $[a,b]$ 上必能取到最大值与最小值。

5.4.3 Zero Theorem 零点存在定理

定理 5.14 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，且 $f(a)\cdot f(b)<0$，则一定存在 $\xi\in(a,b)$，使 $f(\xi)=0$。

定理 5.15 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，且 $f([a,b])\subset[a,b]$，则存在 $\xi\in[a,b]$，使 $f(\xi)=\xi$。这样的 ξ 称为 $f(x)$ 的一个**不动点**。

5.4.4 Intermediate Value Theorem 中间值定理/介值定理

定理 5.16 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则它一定能取到最大值和最小值之间的任何一个值。

推论 5.1 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，m 是最小值，M 是最大值，则 $f(x)$ 的值域是闭区间 $R_f=[m,M]$。

5.4.5 Concept of Uniform Continuity 一致连续概念

在第二节中, 我们已经指出, 函数 $f(x)$ 在某个区间 X 上连续, 是指 $f(x)$ 在区间 X 上的每一点连续 (对区间端点是指左连续与右连续). 而 $f(x)$ 在一点 $x_0\in X$ 的连续性, 可以表述为

$$\forall\varepsilon>0,\exists\delta>0,\forall x\in X(|x-x_0|<\delta):|f(x)-f(x_0)|<\varepsilon.$$

需要强调的是, 这里的 $\delta>0$ 与两个因素有关: 它既依赖于 ε, 同时也依赖于所讨论的点 x_0. 也就是说,δ 应表述为 $\delta=\delta(x_0,\varepsilon)$.

这样就产生一个问题: 对任意给定的 $\varepsilon>0$, 能否找到一个只与 ε 有关, 而对区间 X 上一切点都适用的 $\delta=\delta(\varepsilon)>0$? 也就是说, 对区间 X 上任意两点 x',x'', 只要满足 $|x'-x''|<\delta(\varepsilon)$, 就能保证不等式 $|f(x')-f(x'')|<\varepsilon$ 成立?

这一问题的答案是不一定的, 它不仅与所讨论的函数 $f(x)$ 有关, 也与所讨论的区间 X 有关. 如果对给定的点 $x_0\in X$ 与给定的 $\varepsilon>0$, 将所允许的 $\delta(x_0,\varepsilon)$ 的最大值 (或上确界) 记为 $\delta^*(x_0,\varepsilon)$, 则显然, 上述统一的 $\delta=\delta(\varepsilon)>0$ 的存在性等价于对于区间 X 上的一切点 $x_0,\delta^*(x_0,\varepsilon)$ 有非零下确界. 但实际上, 对 $\delta^*(x_0,\varepsilon)$ 在区间 X 上取下确界, 得到的结果有可能是 $\inf_{x_0\in X}\delta^*(x_0,\varepsilon)=0$.

下面我们就这一关于函数在某区间上整体性质的问题作严格的叙述. 下文中提到的区间 X 表示任意一种有限或无限的区间, 如闭区间 $[a,b]$, 开区间 $(a,b),(a,+\infty),(-\infty,b),(-\infty,+\infty)$, 半开半闭区间 $[a,b),(a,b],(-\infty,b],[a,+\infty)$ 等。

定义 5.10 设函数 $f(x)$ 在区间 X 上定义, 若对于任意给定的 $\varepsilon>0$, 存在 $\delta>0$, 只要 $x',x''\in X$ 满足 $|x'-x''|<\delta$, 就成立 $|f(x')-f(x'')|<\varepsilon$, 则称函数 $f(x)$ 在区间 X 上一致连续.

在上面定义中, 若固定 $x''=x_0\in X$, 就得到 $f(x)$ 在点 x_0 的连续性. 由于 x_0 可以是 X 中的任意一点, 于是得到

$$f(x)\text{ 在区间}X\text{ 上一致连续}\Rightarrow f(x)\text{ 在区间}X\text{ 上连续}.$$

长度无限的区间，如 $[a,+\infty)$ 上的连续函数不一定一致连续；长度有限的开区间 (a,b) 上的连续函数也不一定一致连续。但是对于长度有限的闭区间 $[a,b]$ 上的连续函数，我们有下面的著名定理:

定理 5.17 (Heine–Cantor 定理) 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则它在 $[a,b]$ 上一致连续。

定理 5.18 设函数 $f(x)$ 在区间 X 上定义, 则 $f(x)$ 在 X 上一致连续的充分必要条件是: 对任何点列 $\{x'_n\}(x'_n \in X)$ 与 $\{x''_n\}(x''_n \in X)$, 只要满足 $\lim_{n\to\infty}(x'_n - x''_n) = 0$, 就成立 $\lim_{n\to\infty}(f(x'_n) - f(x''_n)) = 0$。

定理 5.19 设函数 $f(x)$ 在开区间 (a,b) 上连续，则 $f(x)$ 在 (a,b) 上一致连续的充分必要条件是: $f(a+)$ 与 $f(b-)$ 存在且有限。

例题 5.4 证明函数 $f(x) = \frac{1}{x}$ 在区间 $(0,1]$ 上连续，但非一致连续。

对任意固定的 $x_0 \in (0,1]$，要证 $\lim\limits_{x\to x_0} f(x) = f(x_0)$。

给定 $\varepsilon > 0$，需找 $\delta > 0$，使得当 $|x - x_0| < \delta$ 且 $x \in (0,1]$ 时，有

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x - x_0|}{xx_0} < \varepsilon.$$

先限制 $x > x_0/2$，取 $\delta_1 = x_0/2$。若 $|x - x_0| < \delta_1$，则

$$x > \frac{x_0}{2} \quad \Rightarrow \quad \frac{1}{x} < \frac{2}{x_0}.$$

于是

$$\frac{|x - x_0|}{xx_0} < \frac{|x - x_0|}{(x_0/2)\cdot x_0} = \frac{2|x - x_0|}{x_0^2}.$$

为使上式 $< \varepsilon$，只需

$$|x - x_0| < \frac{x_0^2\varepsilon}{2}.$$

取

$$\delta = \min\left(\frac{x_0}{2}, \frac{x_0^2\varepsilon}{2}\right) > 0.$$

则当 $|x - x_0| < \delta$ 时，

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| < \varepsilon.$$

此 δ 依赖于 x_0 和 ε，故 f 在 $(0,1]$ 上每点连续。

用反证法思路（一致连续定义的否定）。

取 $\varepsilon_0 = 1$。对任意 $\delta > 0$，取正整数 n 使 $\frac{1}{n} < \delta$。

令

$$x_n = \frac{1}{n}, \quad y_n = \frac{1}{n+1} \in (0,1].$$

则

$$|x_n - y_n| = \frac{1}{n(n+1)} < \frac{1}{n} < \delta,$$

但

$$|f(x_n) - f(y_n)| = |n - (n+1)| = 1 \geq \varepsilon_0.$$

因此 f 在 $(0,1]$ 上不是一致连续的。

5.5 专题练习：求极限

求极限的重要思想：洛必达法则、泰勒展开、构造重要极限、等价无穷小代换、夹逼定理......

例题 5.5 已知 $\lim\limits_{x\to0}(1+x+\frac{f(x)}{x})^{\frac{1}{x}}=e^3$，求 $\lim\limits_{x\to0}(1+\frac{f(x)}{x})^{\frac{1}{x}}$。

首先分析已知极限，

$$\lim_{x\to0}e^{\frac{\ln(1+x+\frac{f(x)}{x})}{x}}=e^3$$

$$\lim_{x\to0}\frac{\ln(1+x+\frac{f(x)}{x})}{x}=3$$

所以分子是分母的同阶无穷小：

$$\lim_{x\to0}\ln\left(1+x+\frac{f(x)}{x}\right)=0$$

$$\lim_{x\to0}\left(1+x+\frac{f(x)}{x}\right)=1$$

$$\lim_{x\to0}\frac{f(x)}{x}=0$$

因此自然想到第二重要极限（自然指数极限）。注意 $\lim\limits_{x\to0}\ln\left(1+x+\frac{f(x)}{x}\right)=0$ 的等价无穷小是：

$$x+\frac{f(x)}{x}$$

代换得：

$$\lim_{x\to0}\frac{x+\frac{f(x)}{x}}{x}=\lim_{x\to0}\left(1+\frac{f(x)}{x^2}\right)=3$$

得到：

$$\lim_{x\to0}\frac{f(x)}{x^2}=2$$

目标极限等于：

$$\lim_{x\to0}\left[\left(1+\frac{f(x)}{x}\right)^{\frac{x}{f(x)}}\right]^{\frac{f(x)}{x^2}}=e^2$$

例题 5.6 已知 $\lim\limits_{x\to0}\dfrac{f(x)}{x^2}$ 存在，且

$$f(x)=\frac{x-\sin x}{x}+x^2\cdot\lim_{x\to0}\frac{f(x)}{1-\cos x}.$$

求 $\lim\limits_{x\to0}\dfrac{f(x)}{x^2}$。

设

$$A=\lim_{x\to0}\frac{f(x)}{x^2},\quad B=\lim_{x\to0}\frac{f(x)}{1-\cos x}.$$

由于 $1-\cos x\sim\frac{x^2}{2}$，有

$$B=\lim_{x\to0}\frac{f(x)}{x^2/2}=2A.$$

代入原式：

$$f(x)=\frac{x-\sin x}{x}+Bx^2=1-\frac{\sin x}{x}+Bx^2.$$

于是

$$\frac{f(x)}{x^2} = \frac{1-\frac{\sin x}{x}}{x^2} + B.$$

注意

$$1-\frac{\sin x}{x} = \frac{x-\sin x}{x}, \quad \frac{1-\frac{\sin x}{x}}{x^2} = \frac{x-\sin x}{x^3}.$$

已知

$$\lim_{x\to 0}\frac{x-\sin x}{x^3} = \frac{1}{6}.$$

所以

$$A = \frac{1}{6} + B = \frac{1}{6} + 2A.$$

解得

$$A - 2A = \frac{1}{6} \quad \Rightarrow \quad -A = \frac{1}{6} \quad \Rightarrow \quad A = -\frac{1}{6}.$$

例题 5.7

例题 5.8

例题 5.9

例题 5.10

6 Differentials 微分

6.1 Differentials and Derivatives 微分和导数

6.1.1 Definition of Differentials 微分的定义

定义 6.1 对函数 $y = f(x)$ 定义域中的一点 x_0，若存在一个只与 x_0 有关，而与 Δx 无关的数 $g(x_0)$，使得当 $\Delta x \to 0$ 时恒成立关系式 $\Delta y = g(x_0)\Delta x + o(\Delta x)$，则称 $f(x)$ 在 x_0 处的微分存在，或称 $f(x)$ 在 x_0 处**可微**。若函数在某一区间上的每一点都可微，则称其在该区间上可微。

- 这表示函数的增量 Δy 可以分成两部分：
  1. 线性主部（linear principal part）：$g(x_0)\Delta x$
  2. 高阶无穷小：$o(\Delta x)$
- 当 $f(x)$ 在 x 处可微且 $\Delta x \to 0$ 时，我们将 Δx 称为自变量的微分，记作 dx，而将 Δy 的线性主部 $g(x)dx$ 称为因变量的微分，记作 dy 或 $df(x)$，这样就有了以下的微分关系式：

$$dy = g(x)dx$$

注意，可微一定连续，连续不一定可微！
函数可微的几何含义是：图形不仅是"一笔画"成的，而且是"光滑"的，没有尖角或棱角。

6.1.2 Differentials and Derivatives 微分和导数

若 $f(x)$ 在 x_0 处可微，则当 $\Delta x \to 0$ 时有关系式

$$\Delta y = g(x_0)\Delta x + o(\Delta x)$$

我们稍加处理

$$g(x_0) = \frac{\Delta y}{\Delta x} - \frac{o(\Delta x)}{\Delta x}$$
$$g(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

定义 6.2 若函数 $y = f(x)$ 在其定义域中的一点 x_0 处极限 $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ 存在，则称 $f(x)$ 在该点**可导**，并称这个极限值为 $f(x)$ 在该点处的**导数**，记为 $f'(x_0)$ 或 $y'(x_0)$, $\frac{df}{dx}\big|_{x=x_0}$, $\frac{dy}{dx}\big|_{x=x_0}$。若函数在某一区间上的每一点都可导，则称其在该区间上可导。

函数 $f(x)$ 的所有可导点的集合是其定义域的子集，导数值 $f'(x)$ 可看成定义在这一子集上的一个新的函数，我们将它称为函数 $f(x)$ 的**导函数**，记为 $f'(x)$ 或 $y'(x)$, $\frac{df}{dx}$, $\frac{dy}{dx}$。导函数一般就简称为导数。

若 $f(x)$ 在 x 处可微，则它必定在 x 处可导，而前面所述的函数 $g(x)$ 不是别的，正是它在这点的导数值 $f'(x)$。于是，差分的无穷小量关系式和微分关系式分别成为

$$\Delta y = f'(x)\Delta x + o(\Delta x)$$

和

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

定理 6.1 函数 $y = f(x)$ 在 x 处可微的充分必要条件是它在 x 处可导。

该定理告诉我们，对于一元函数来说，它在任一点的可微性与可导性是等价的。

6.2 Meaning and Properties of the Derivative 导数的意义和性质

6.2.1 Various Interpretations of the Derivative 导数的多种意义

（1）物理意义：瞬时变化率（instantaneous rate of change）

微积分的发明人之一牛顿最早用导数研究的是如何确定力学中运动物体的瞬时速度问题：$v(t) = \frac{ds}{dt}$。我们可以将“速度”这个概念加以推广——凡是牵涉某个量的变化快慢的，诸如物理学中的光热磁电的各种传导率、化学中的反应速率、经济学中的资金流动速率、人口学中的人口增长速率等等，统统都可以看成是广义的“速度”，因而都可以用导数来表达。换句话说，导数实际上是因变量关于自变量的变化率。

（2）几何意义：切线斜率（slope of tangent）

函数 $y = f(x)$ 在 x_0 的导数 $f(x_0)$ 是曲线在该点切线的斜率。这是将动态的“变化率”在静态图像上的可视化表示。

（3）代数意义：线性近似的系数（coefficient of linear approximation）

从可微定义 $\Delta y = f'(x_0)\Delta x + o(\Delta x)$ 来看，导数 $f'(x_0)$ 是增量比的最佳线性近似系数。

（4）分析意义：极限过程

导数是一个极限，这是导数的形式定义，也是所有其他意义的严格基础。

6.2.2 One-Sided Derivatives 单侧导数

由于导数是一个极限，而极限存在意味着左极限和右极限都存在并且相等，可知函数可导的充分必要条件是相应的左导数和右导数存在并且相等。

6.3 Derivative Rules for Arithmetic Operations and Inverse Functions 导数四则运算和反函数求导法则

6.3.1 Finding the Derivative from the Definition 从定义出发求导函数

计算一个函数的导函数的运算称为对这个函数求导。

一些简单函数可以直接通过导数的定义来求导，当然也可以直接使用微分的定义来求导。

显然，常数函数 $y = C$ 的导数恒等于零。

例题 6.1 求 $y = \sin x$ 的导函数。

例题 6.2 求 $y = \ln x$ 的导函数。

例题 6.3 求 $y = e^x$ 的导函数。

例题 6.4 求幂函数 $y = x^a (x > 0)$ 的导函数，其中 a 为任意实数。

除了少数几个最简单的函数之外，可以直接用定义较方便地求出导数的函数实在是微乎其微，因而就有必要对一般的函数导出一系列的求导运算法则，主要包括四则运算法则、反函数求导法则和复合函数求导法则。

6.3.2 Derivative Rules for Arithmetic Operations 求导的四则运算法则

定理 6.2 设 $f(x)$ 和 $g(x)$ 在某一区间上都是可导的，则对任意常数 c_1 和 c_2，它们的线性组合 $c_1 f(x) + c_2 g(x)$ 也在该区间上可导，且满足如下的线性运算关系：$[c_1 f(x) + c_2 g(x)]' = c_1 f'(x) + c_2 g'(x)$。

定理 6.3 设 $f(x)$ 和 $g(x)$ 在某一区间上都是可导的，则它们的积函数也在该区间上可导，且满足

$$[f(x)\cdot g(x)]' = f'(x)g(x) + f(x)g'(x)$$

相应的微分表达式为

$$d[f(x)\cdot g(x)] = g(x)d[f(x)] + f(x)d[g(x)]$$

定理 6.4 设 $g(x)$ 在某一区间上可导，且 $g(x)\neq 0$，则它的倒数也在该区间上可导，且满足

$$\left[\frac{1}{g(x)}\right]' = -\frac{g'(x)}{[g(x)]^2}$$

相应的微分表达形式为

$$d\left[\frac{1}{g(x)}\right] = -\frac{d[g(x)]}{[g(x)]^2}$$

推论 6.1 设 $f(x)$ 和 $g(x)$ 在某一区间上都是可导的，且 $g(x)\neq 0$，则它们的商函数也在该区间上可导，且满足

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

这一结论的微分形式为

$$d\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)d\left[f(x)\right] - f(x)d\left[g(x)\right]}{[g(x)]^2}$$

6.3.3 Derivative Rule for Inverse Functions 反函数求导法则

定理 6.5 (反函数求导定理) 若函数 $y=f(x)$ 在 (a,b) 上连续、严格单调、可导并且 $f'(x)\neq 0$，记 $\alpha = \min(f(a+), f(b-)), \beta = \max(f(a+), f(b-))$，则它的反函数 $x=f^{-1}(y)$ 在 (α,β) 上可导，且有

$$\left[f^{-1}(y)\right]' = \frac{1}{f'(x)}$$

6.3.4 Derivative Formulas for Basic Elementary Functions 基本初等函数的导数公式

$(C)' = 0$
$(x^a)' = ax^{a-1}$
$(\sin x)' = \cos x$
$(\cos x)' = -\sin x$
$(\tan x)' = \sec^2 x$
$(\cot x)' = -\csc^2 x$
$(\sec x)' = \tan x \sec x$
$(\csc x)' = -\cot x \csc x$
$(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$
$(\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}$
$(\arctan x)' = \dfrac{1}{1+x^2}$
$(\operatorname{arccot} x)' = -\dfrac{1}{1+x^2}$

$(a^x)' = \ln a \cdot a^x$
特别地 $(e^x)' = e^x$
$(\log_a x)' = \dfrac{1}{\ln a}\cdot\dfrac{1}{x}$
特别地 $(\ln x)' = \dfrac{1}{x}$
$(\sinh x)' = \cosh x$
$(\cosh x)' = \sinh x$
$(\tanh x)' = \operatorname{sech}^2 x$
$(\coth x)' = -\operatorname{csch}^2 x$
$(\sinh^{-1} x)' = \dfrac{1}{\sqrt{1+x^2}}$
$(\cosh^{-1} x)' = \dfrac{1}{\sqrt{x^2-1}}$
$(\tanh^{-1} x)' = (\coth^{-1} x)' = \dfrac{1}{1-x^2}$

6.4 Chain Rule and Its Applications 复合函数求导法则及其应用

6.4.1 Chain Rule 复合函数求导法则

定理 6.6 (复合函数求导法则) 设函数 $u = g(x)$ 在 $x = x_0$ 可导，而函数 $y = f(u)$ 在 $u = u_0 = g(x_0)$ 处可导，则复合函数 $y = f(g(x))$ 在 $x = x_0$ 可导，且有

$$[f(g(x))]'_{x=x_0} = f'(u_0)g'(x_0) = f'(g(x_0))g'(x_0)$$

复合函数的求导规则可以写成

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

我们一般称它为**链式法则**（chain rule）。

- 形如

  $$y = f(x) = u(x)^{v(x)}$$

  的函数称为**幂指函数**，对于幂指函数的求导，常采用的方法叫**对数求导法**，计算时，先对两边取对数，令

  $$z(x) = \ln y = v(x) \ln u(x)$$

  ，再分别求导。

6.4.2 Implicit Differentiation and Finding Differentials 隐函数求导与求微分

对方程两边同时求导或求微分的方法，无论对于显函数、可显化的隐函数还是不可显化的隐函数都能有效地使用。

6.5 Higher-Order Derivatives and Higher-Order Differentials 高阶导数和高阶微分

6.5.1 Definition of Higher-Order Derivatives 高阶导数的定义

定义 6.3 设函数 $y = f(x)$ 的 $n-1$ 阶导数 $f^{(n-1)}(x)$ 仍是个可导函数，则它的导数被称为 $f(x)$ 的n **阶导数**，记为 $f^{(n)}(x)$，并称 $f(x)$ 是n **阶可导函数**。

显然，若 $f(x)$ 的 n 阶导数存在，则它的低于 n 阶的导数都存在。

6.5.2 Rules for Higher-Order Derivatives 高阶导数的运算法则

定理 6.7 设 $f(x)$ 和 $g(x)$ 都是 n 阶可导的，则对任意常数 c_1 和 c_2，它们的线性组合 $c_1f(x)+c_2g(x)$ 也是 n 阶可导的，且满足如下线性运算关系

$$[c_1f(x) + c_2g(x)]^{(n)} = c_1f^{(n)}(x) + c_2g^{(n)}(x)$$

定理 6.8 (*Leibniz* 公式) 设 $f(x)$ 和 $g(x)$ 都是 n 阶可导的，则它们的积函数也 n 阶可导，且成立公式

$$[f(x) \cdot g(x)]^{(n)} = \sum_{k=0}^{n} C_n^k f^{(n-k)}(x)g^{(k)}(x)$$

这里 $C_n^k = \frac{n!}{k!(n-k)!}$ 是组合数。

6.5.3 Higher-Order Differentials 高阶微分

我们可以通过与定义高阶导数类似的方法来定义高阶微分。如 dy 是函数 $y = f(x)$ 的一阶微分，则称 dy 的微分 $d(dy) = d^2y$ 为 y 的二**阶微分**...

可以求得

$$d^2y = f''(x)dx^2$$

$$d^ny = f^{(n)}(x)dx^n$$

其中 dx^n 表示 $(dx)^n$。

这个公式建立了高阶导数与高阶微分之间的关系，即 y **的** n **阶微分等于它的** n **阶导数乘上自变量的微分的** n **次方**。

6.6　专题练习：一元函数微分学的计算

（1）基本求导公式与四则运算

例题 6.5 设 $f(x)=\prod_{n=1}^{100}\left(\tan\frac{\pi x^n}{4}-n\right)$，则 $f'(1)=$ ________________。

对于连乘积函数：

$$f(x)=\prod_{n=1}^{N}g_n(x),$$

其导数公式为：

$$f'(x)=f(x)\cdot\sum_{n=1}^{N}\frac{g_n'(x)}{g_n(x)}.$$

其中 $g_n(x)=\tan\left(\frac{\pi x^n}{4}\right)-n$。首先计算 $g_n'(x)$：

$$\frac{d}{dx}\tan\left(\frac{\pi x^n}{4}\right)=\sec^2\left(\frac{\pi x^n}{4}\right)\cdot\frac{\pi}{4}\cdot nx^{n-1}.$$

所以：

$$g_n'(x)=\frac{\pi n x^{n-1}}{4}\cdot\sec^2\left(\frac{\pi x^n}{4}\right).$$

当 $x=1$ 时：

$$\tan\left(\frac{\pi\cdot 1^n}{4}\right)=\tan\frac{\pi}{4}=1,$$

因此：

$$g_n(1)=1-n.$$

特别地，当 $n=1$ 时，$g_1(1)=1-1=0$。

所以：

$$f(1)=g_1(1)\cdot g_2(1)\cdots g_{100}(1)=0.$$

设：

$$A(x)=\tan\left(\frac{\pi x}{4}\right)-1,$$

$$B(x)=\prod_{n=2}^{100}\left(\tan\left(\frac{\pi x^n}{4}\right)-n\right).$$

则 $f(x)=A(x)B(x)$，且 $A(1)=0$。

由乘积法则：

$$f'(1)=A'(1)\cdot B(1).$$

$$A'(x)=\frac{\pi}{4}\sec^2\left(\frac{\pi x}{4}\right).$$

当 $x=1$ 时：

$$\sec^2\left(\frac{\pi}{4}\right)=\frac{1}{\cos^2(\pi/4)}=\frac{1}{(\sqrt{2}/2)^2}=2.$$

所以：

$$A'(1)=\frac{\pi}{4}\cdot 2=\frac{\pi}{2}.$$

$$B(1)=\prod_{n=2}^{100}\left(\tan\left(\frac{\pi}{4}\right)-n\right)=\prod_{n=2}^{100}(1-n).$$

注意 $1-n=-(n-1)$，所以：

$$B(1)=\prod_{n=2}^{100}[-(n-1)]=(-1)^{99}\prod_{n=2}^{100}(n-1)=(-1)^{99}\cdot(1\times 2\times\cdots\times 99).$$

即：

$$B(1)=-99!.$$

$$f'(1)=A'(1)\cdot B(1)=\frac{\pi}{2}\cdot(-99!)=-\frac{\pi\cdot 99!}{2}.$$

（2）复合函数的导数与微分形式不变性

例题 6.6 设 $y=\ln(x+\sqrt{x^2+a^2})(a\neq 0)$，求 $y'|_{x=0}$。

当 $a=1$ 时，就是反双曲正弦函数。本题是复合函数，外层 $y=\ln u$，内层 $u=x+\sqrt{x^2+a^2}$。所求的导数等于外层导数乘以内层导数。因为

$$\begin{aligned} y'&=\frac{1}{x+\sqrt{x^2+a^2}}\cdot(x+\sqrt{x^2+a^2})' \\ &=\frac{1}{x+\sqrt{x^2+a^2}}\left[1+\frac{1}{2\sqrt{x^2+a^2}}\cdot(x^2+a^2)'\right] \\ &=\frac{1}{x+\sqrt{x^2+a^2}}\left(1+\frac{x}{\sqrt{x^2+a^2}}\right)=\frac{1}{\sqrt{x^2+a^2}} \end{aligned}$$

所以

$$y'|_{x=0}=\left.\frac{1}{\sqrt{x^2+a^2}}\right|_{x=0}=\frac{1}{|a|}\quad(a\neq 0)$$

例题 6.7 设

$$f(x)=\begin{cases}2x-1, & |x|\leq 1,\\ \ln\sqrt{x}, & |x|>1,\end{cases}$$

且 $y=f[f(x)]$，求 $\left.\frac{dy}{dx}\right|_{x=e}$。

因为

$$\left.\frac{dy}{dx}\right|_{x=e}=f'[f(x)]f'(x)\Big|_{x=e}=f'[f(e)]f'(e),$$

其中

$$f(e)=\ln\sqrt{x}\Big|_{x=e}=\frac{1}{2},$$

$$f'[f(e)]=f'\left(\frac{1}{2}\right)=(2x-1)'\Big|_{x=\frac{1}{2}}=2,$$

$$f'(e)=(\ln\sqrt{x})'\Big|_{x=e}=\frac{1}{2e},$$

所以

$$\left.\frac{dy}{dx}\right|_{x=e}=2\cdot\frac{1}{2e}=\frac{1}{e}.$$

例题 6.8 设 $y=e^{\sin(\ln x)}$，求 dy 及 $\frac{dy}{dx}$。

由一阶微分形式的不变性，得

$$d[f(u)] = f'(u)du$$

令 $u = g(x)$ 为中间变量，形式不变！

若 u 就是 x，则 $d[f(x)] = f'(x)dx$。这就是一阶微分形式不变性！

$$\begin{aligned} d[e^{\sin(\ln x)}] &= e^{\sin(\ln x)}d[\sin(\ln x)] \\ \Rightarrow d(e^u) &= (e^u)'du = e^u du \\ &= e^{\sin(\ln x)}\cos(\ln x)d(\ln x) \\ &= e^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}dx, \end{aligned}$$

所以

$$dy = e^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}dx, \quad \frac{dy}{dx} = e^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}.$$

（3）分段函数的导数

设 $f(x) = \begin{cases} f_1(x), & x \geq x_0, \\ f_2(x), & x < x_0, \end{cases}$ 其中 $f_1(x), f_2(x)$ 分别在 $x > x_0, x < x_0$ 时可导，则

(a) **在分段点 x_0 处用导数定义求导**：

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f_1(x) - f(x_0)}{x - x_0}, \quad f'_-(x_0) = \lim_{x \to x_0^-} \frac{f_2(x) - f(x_0)}{x - x_0}$$

根据 $f'_+(x_0)$ 是否等于 $f'_-(x_0)$ 来判定 $f'(x_0)$；

(b) **在非分段点用导数公式求导**，即 $x > x_0$ 时，$f'(x) = f'_1(x)$; $x < x_0$ 时，$f'(x) = f'_2(x)$。

例题 6.9 设 $y = \ln|x|, x \neq 0$，求 y'。

此题需要注意，$x = 0$ 是无定义点，无定义自然不连续，不连续一定不可导，所以只需要讨论 $x \neq 0$ 时的导数即可。加绝对值的函数无法直接求导，需要先去绝对值将其写成分段函数。

$$y = \ln|x| = \begin{cases} \ln x, & x > 0, \\ \ln(-x), & x < 0. \end{cases}$$

当 $x > 0$ 时，

$$y' = (\ln x)' = \frac{1}{x};$$

当 $x < 0$ 时，

$$y' = [\ln(-x)]' = \frac{1}{-x} \cdot (-x)' = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.$$

因此

$$y' = (\ln|x|)' = \frac{1}{x} \quad (x \neq 0).$$

例题 6.10 设 $y = |xe^{-x}|$，求 y' 和 y''。

y 的表达式含有绝对值符号，可知其为分段函数，$x = 0$ 为其分段点，则

$$y = |xe^{-x}| = \begin{cases} -xe^{-x}, & x < 0, \\ xe^{-x}, & x \geq 0, \end{cases}$$

所以

$$y' = \begin{cases} e^{-x}(x-1), & x<0, \\ e^{-x}(1-x), & x>0. \end{cases}$$

而在分段点处用导数定义求导：

$$y'_-(0) = \lim_{x\to 0^-} \frac{y(x)-y(0)}{x} = \lim_{x\to 0^-} \frac{-xe^{-x}}{x} = -1,$$

$$y'_+(0) = \lim_{x\to 0^+} \frac{y(x)-y(0)}{x} = \lim_{x\to 0^+} \frac{xe^{-x}}{x} = 1,$$

因为 $y'_-(0) \neq y'_+(0)$，可知 y 在 $x=0$ 处不可导。

所以

$$y'' = \begin{cases} e^{-x}(2-x), & x<0, \\ e^{-x}(x-2), & x>0. \end{cases}$$

(4）反函数的导数

我们知道：

$$y'_x = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{x'_y}$$

这里重点讲一下**反函数的二阶导数求法**：

$$y''_{xx} = \frac{d^2y}{dx^2} = \frac{d(\frac{dy}{dx})}{dx} = \frac{d\left(\frac{1}{x'_y}\right)}{dx} = \frac{d\left(\frac{1}{x'_y}\right)}{dy}\cdot\frac{1}{x'_y} = -\frac{1}{(x'_y)^2}\cdot(x'_y)'_y\cdot\frac{1}{x'_y} = -\frac{x''_{yy}}{(x'_y)^2}\cdot\frac{1}{x'_y} = \boxed{-\frac{x''_{yy}}{(x'_y)^3}}$$

例题 6.11 当 $x>0$ 时，设 $y=f(x)=3x^2+e^x$ 有反函数 $x=\varphi(y)$，则 $\varphi''(3+e)=$ ________。

本题若想得式子 $x=\varphi(y)$ 比较困难，它考查的是反函数求二阶导数知识，可以直接用公式

$$\varphi''(y) = -\frac{f''(x)}{[f'(x)]^3}.$$

注意当 $y=3+e$ 时，$x=1$。

当 $f(x)=3x^2+e^x=3+e$ 时，有 $x=1$，于是

$$\varphi''(y)|_{y=3+e} = -\frac{f''(x)}{[f'(x)]^3}\bigg|_{x=1} = -\frac{6+e^x}{(6x+e^x)^3}\bigg|_{x=1} = -\frac{1}{(6+e)^2}.$$

(5）隐函数求导法

例题 6.12 设函数 $y=y(x)$ 由方程 $y^3+xy^2+x^2y+6=0$ 确定，且 $y'(1)=0$，求 $y''(1)$ 的值。

该方程不容易得到显式 $y=y(x)$，但仍可以求导。只要方程两边对 x 求导即可！

在 $y^3+xy^2+x^2y+6=0$ 两边关于 x 求导，得

$$3y^2y'+y^2+2xyy'+2xy+x^2y'=0,$$

由 $y'(1)=0$，得 $y^2(1)+2\cdot 1\cdot y(1)=0$，解得 $y(1)=-2$ 或 $y(1)=0$（不满足题目所给方程，舍去）。

在 $3y^2y'+y^2+2xyy'+2xy+x^2y'=0$ 两边关于 x 求导，得

$$(3y^2+2xy+x^2)y''+2(3y+x)(y')^2+4(y+x)y'+2y=0,$$

代入 $x=1$，$y(1)=-2$ 与 $y'(1)=0$，解得 $y''(1)=\frac{4}{9}$。

（6）参数方程所确定的函数的导数

我们知道：

$$\frac{dy}{dx}=\frac{dy/dt}{dx/dt}$$

这里重点讲一下**由参数方程确定的函数的二阶导数**：

$$\frac{d^2y}{dx^2}=\frac{d\left(\frac{dy}{dx}\right)}{dx}=\frac{d\left(\frac{dy}{dx}\right)}{dt}\cdot\frac{dt}{dx}$$

例题 6.13 设 $y=y(x)$ 由参数方程

$$\begin{cases}x=\sin t,\\ y=t\sin t+\cos t\end{cases}$$

确定，则

$$\left.\frac{d^2y}{dx^2}\right|_{t=\frac{\pi}{4}}=\underline{\qquad\qquad}.$$

本题是参数方程求导，将 $\frac{dy}{dx}$ 转化为 $\frac{dy/dt}{dx/dt}=t$，二阶导

$$\frac{d^2y}{dx^2}=\frac{d}{dx}\left(\frac{dy}{dx}\right)=\frac{d}{dt}\left(\frac{dy}{dx}\right)\cdot\frac{dt}{dx}.$$

因为

$$\frac{dx}{dt}=\cos t,\quad \frac{dy}{dt}=\sin t+t\cos t-\sin t=t\cos t,$$

$$\frac{dy}{dx}=\frac{dy/dt}{dx/dt}=\frac{t\cos t}{\cos t}=t,$$

$$\frac{d^2y}{dx^2}=\frac{d}{dx}\left(\frac{dy}{dx}\right)=\frac{d}{dt}\left(\frac{dy}{dx}\right)\cdot\frac{dt}{dx}=\frac{1}{\cos t},$$

所以

$$\left.\frac{d^2y}{dx^2}\right|_{t=\frac{\pi}{4}}=\frac{1}{\cos\frac{\pi}{4}}=\sqrt{2}.$$

（7）对数求导法

对于多项相乘、相除、开方、乘方的式子，一般先取对数再求导。

例题 6.14 设函数 $y=y(x)$ 由方程 $xe^{f(y)}=e^y\ln 2$ 确定，其中 f 具有二阶导数，且 $f'\neq 1$，则

$$\frac{d^2y}{dx^2}=\underline{\qquad\qquad\qquad}.$$

本题中有“$e^{f(y)}$”“e^y”这种复杂的形式，可以考虑两边取对数，同时，经分析可知 $e^y\ln 2>0$，所以 $x>0$，满足取对数条件，此时变为简单的隐函数方程，两边直接对 x 求导。因为本题是求二阶导数，所以需要再次求导。

方程 $xe^{f(y)}=e^y\ln 2$ 两端取对数，得

$$\ln x+f(y)=y+\ln(\ln 2).$$

两端关于 x 求导，得

$$\frac{1}{x}+f'(y)\cdot y'=y',$$

即

$$y' = \frac{1}{x[1-f'(y)]}.$$

两端继续关于 x 求导，得

$$-\frac{1}{x^2} + f''(y) \cdot (y')^2 + f'(y) \cdot y'' = y'',$$

整理得

$$y''[1-f'(y)] = -\frac{1}{x^2} + f''(y) \cdot (y')^2,$$

代入 $y' = \dfrac{1}{x[1-f'(y)]}$，得

$$y''[1-f'(y)] = -\frac{1}{x^2} + \frac{f''(y)}{x^2[1-f'(y)]^2},$$

由此可得

$$y'' = \frac{[1-f'(y)]^2 - f''(y)}{x^2[1-f'(y)]^3}.$$

（8）幂指函数求导法

例题 6.15 求函数 $y = x^x (x > 0)$ 的导数。

x^x 是典型的幂指函数，将 x^x 化为 $e^{x\ln x}$，再去求导。

$$\begin{aligned} y' &= (x^x)' = (e^{x\ln x})' \\ &= e^{x\ln x}(x\ln x)' \\ &= x^x(1+\ln x) \quad (x>0) \end{aligned}$$

例题 6.16 求函数 $y = x^{\frac{1}{x}} (x > 0)$ 的导数。

$x^{\frac{1}{x}}$ 是幂指函数，将 $x^{\frac{1}{x}}$ 化为 $e^{\frac{1}{x}\ln x}$，再去求导。

$$\begin{aligned} y' &= \left(x^{\frac{1}{x}}\right)' = \left(e^{\frac{1}{x}\ln x}\right)' \\ &= e^{\frac{1}{x}\ln x}\left(-\frac{1}{x^2} \cdot \ln x + \frac{1}{x} \cdot \frac{1}{x}\right) \\ &= x^{\frac{1}{x}-2}(1-\ln x) \quad (x>0) \end{aligned}$$

（9）高阶导数

求高阶导数主要有三种方法。

（1）归纳法：逐次求导，探索规律，得出通式。

例题 6.17 求 $y = \sin x$ 的 n 阶导数。

（2）莱布尼兹公式：设 $f(x)$ 和 $g(x)$ 都是 n 阶可导的，则它们的积函数也 n 阶可导，且成立公式

$$[f(x) \cdot g(x)]^{(n)} = \sum_{k=0}^{n} C_n^k f^{(n-k)}(x) g^{(k)}(x)$$

这里 $C_n^k = \frac{n!}{k!(n-k)!}$ 是组合数。

例题 6.18 设 $f(x) = xe^x$，则 $f^{(n)}(x) = $ ________________。

$$f^{(n)}(x)=(xe^x)^{(n)}=(e^x)^{(n)}\cdot x+C_n^1(e^x)^{(n-1)}\cdot 1=xe^x+ne^x=(n+x)e^x.$$

（3）泰勒展开式

例题 6.19 设 $f(x)=x^2 2^x$，则当 $n\geq 1$ 时，$f^{(n)}(0)=$ __________。

方法一：泰勒展开有

$$f(x)=x^2 2^x=x^2 e^{x\ln 2}=x^2\sum_{n=0}^{\infty}\frac{(x\ln 2)^n}{n!}=\sum_{n=0}^{\infty}\frac{(\ln 2)^n}{n!}x^{n+2}=\sum_{n=2}^{\infty}\frac{(\ln 2)^{n-2}}{(n-2)!}x^n,$$

又 $f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n$，由泰勒展开式的唯一性，得

$$\frac{(\ln 2)^{n-2}}{(n-2)!}=\frac{f^{(n)}(0)}{n!}.$$

又 $f'(0)=0$，故

$$f^{(n)}(0)=\frac{(\ln 2)^{n-2}}{(n-2)!}n!=n(n-1)(\ln 2)^{n-2}\quad(n=1,2,3,\cdots).$$

方法二：莱布尼兹高阶求导公式有

$$f^{(n)}(x)=C_n^0 x^2(2^x)^{(n)}+C_n^1\cdot 2x\cdot(2^x)^{(n-1)}+C_n^2\cdot 2\cdot(2^x)^{(n-2)}$$

$$=x^2\cdot 2^x\cdot(\ln 2)^n+n\cdot 2x\cdot 2^x\cdot(\ln 2)^{n-1}+\frac{n(n-1)}{2}\cdot 2\cdot 2^x\cdot(\ln 2)^{n-2},$$

当 $x=0$ 时，$f^{(n)}(0)=n(n-1)(\ln 2)^{n-2}\quad(n=1,2,3,\cdots)$。

习题 6.1 设函数 $g(x)$ 可微，$h(x) = e^{1+g(x)}$，$h'(1) = 1$，$g'(1) = 2$，则 $g(1) = (\quad)$。

(A) $\ln 3 - 1$ *(B)* $-\ln 3 - 1$ *(C)* $-\ln 2 - 1$ *(D)* $\ln 2 - 1$

习题 6.2 设 $f(x) = \lim_{t\to 0} x(1-2t)^{-\frac{x}{t}}$，则 $f'(x) =$ ________________。

习题 6.3 已知 $f'(x) = Ae^x$（A 为正常数），则 $f(x)$ 的反函数的二阶导数为 ________________。

习题 6.4 设函数 $y = y(x)$ 由

$$\begin{cases} x = \ln(1+t^2) + 1, \\ y = 2\arctan t - (t+1)^2 \end{cases}$$

确定，则

$$\frac{d^2y}{dx^2} = \underline{\qquad\qquad\qquad}$$

习题 6.5 设函数 $f(x)$ 在 $x = 2$ 的某邻域内具有任意阶导数，且 $f'(x) = e^{f(x)}$，$f(2) = 1$，则当 $n \geq 1$ 时，

$$f^{(n)}(2) = \underline{\qquad\qquad\qquad}$$

习题 6.6 设 $y = y(x)$ 是由方程 $\sin(xy) = \ln\frac{x+e}{y} + 1$ 确定的隐函数，求 $y'(0)$ 的值。

习题 6.7 已知 $g(x)$ 在 $x = 0$ 处二阶可导，且 $g(0) = g'(0) = 0$，设

$$f(x) = \begin{cases} \frac{g(x)}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

证明：$f(x)$ 的导函数在 $x = 0$ 处连续。

习题 6.8 求函数

$$f(x) = x^2 \ln(1+x)$$

在 $x = 0$ 处的 n 阶导数 $f^{(n)}(0)(n \geq 3)$。

7 Mean Value Theorems and Their Applications 微分中值定理及其应用

7.1 Mean Value Theorems 微分中值定理

微分中值定理是一系列中值定理的总称，包含：罗尔定理、拉格朗日中值定理和柯西中值定理。

7.1.1 Extreme Values of Functions and Fermat's Lemma 函数极值与费马引理

牛顿在研究物体运动和莱布尼兹在研究曲线的几何性质的过程中分别独立地发现了微分和导数。但事实上，微分的思想可追溯到费马对函数极值的研究。

定义 7.1 设 $f(x)$ 在 (a,b) 上有定义，$x_0 \in (a,b)$，如果存在点 x_0 的某一个邻域 $O(x_0,\delta) \subset (a,b)$，使得 $f(x) \leq f(x_0), x \in O(x_0,\delta)$，则称 x_0 是 $f(x)$ 的一个**极大值点**，$f(x_0)$ 称为相应的**极大值**。

类似地可以定义 $f(x)$ 的**极小值点**和**极小值**，在不需要区分极大和极小的时候我们将其统称为**极值点**和**极值**。

- 注意，对极值点的定义并不牵涉函数的其他性质，如连续、可微等。

定理 7.1 (费马引理) 设 x_0 是 $f(x)$ 的一个极值点，且 $f(x)$ 在 x_0 处可导，则 $f'(x_0) = 0$。

7.1.2 Rolle theorem 罗尔定理

定理 7.2 (***Rolle* 定理**) 设函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 上可导，且 $f(a) = f(b)$，则至少存在一点 $\xi \in (a,b)$ 使得 $f'(\xi) = 0$。

- 罗尔定理又叫罗尔中值定理。
- 罗尔定理的几何意义是：一条平滑的、连接等高点位的曲线，内部至少有一条水平切线。

7.1.3 Lagrange's Mean Value Theorem 拉格朗日中值定理/均值定理

定理 7.3 设函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 上可导，则至少存在一点 $\xi \in (a,b)$ 使得

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

- 显然，拉格朗日中值定理为罗尔定理的推广。
- 拉格朗日中值定理又叫均值定理，又叫有限增量定理。
- 拉格朗日中值定理的几何意义是：一条平滑的曲线，内部至少有一条切线与端点连线平行。

7.1.4 Analyzing Function Properties Using Lagrange's Mean Value Theorem 用拉格朗日中值定理讨论函数性质

定理 7.4 (拉格朗日中值定理的第一重要引理) 若 $f(x)$ 在 (a,b) 上可导且有 $f'(x) \equiv 0$，则 $f(x)$ 在 (a,b) 上恒为常数。

定理 7.5 (一阶导数与单调性的关系) 设函数 $f(x)$ 在区间 I 上可导，则 $f(x)$ 在 I 上单调增加的充分必要条件是：对于任意 $x \in I$ 有 $f'(x) \geq 0$。

定义 7.2 设函数 $f(x)$ 在区间 I 上定义，若对 I 中的任意两点 x_1 和 x_2，和任意 $\lambda \in (0,1)$，都有 $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$，则称 $f(x)$ 是 I 上的下凸函数。若不等号严格成立，则称其是严格下凸函数。

类似地可以给出上凸函数和严格上凸函数的定义。

定理 7.6 (二阶导数与凸性的关系) 设函数 $f(x)$ 在区间 I 上二阶可导，则 $f(x)$ 在区间 I 上是下凸函数的充分必要条件是：对于任意 $x \in I$ 有 $f''(x) \geq 0$。

定理 7.7 设 $f(x)$ 在区间 I 上连续，$(x_0-\delta, x_0+\delta) \subset I$。
(1) 设 $f(x)$ 在 $(x_0-\delta, x_0)$ 与 $(x_0, x_0+\delta)$ 上二阶可导。若 $f''(x)$ 在 $(x_0-\delta, x_0)$ 与 $(x_0, x_0+\delta)$ 上的符号相反，则点 $(x_0, f(x_0))$ 是曲线 $y=f(x)$ 的拐点；若符号相同则不是拐点。
(2) 设 $f(x)$ 在 $(x_0-\delta, x_0+\delta)$ 上二阶可导，若点 $(x_0, f(x_0))$ 是曲线 $y=f(x)$ 的拐点，则 $f''(x_0)=0$。

- 假设二阶可导，二阶导为 0 是拐点（point of inflection）的必要条件，举例为 0 但不是拐点的：$y=x^4$ 在 $(0,0)$ 点。
- 另外，假设函数在该点二阶不可导，它仍然有可能是拐点，例如：$y=x^{\frac{1}{3}}$。
- 因此，若要通过求二阶导来确定拐点，既要先在候选点中筛查出二阶导为 0 的非拐点，又要考虑 $f''(x)$ 不存在的点。

定理 7.8 (*Jensen* 不等式) 若 $f(x)$ 为区间 I 上的下凸（上凸）函数，则对于任意 $x_i \in I$ 和满足 $\Sigma_{i=1}^{n}\lambda_i = 1$ 的 $\lambda_i > 0 (i=1,2,...,n)$，成立

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i) \quad \left(f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq \sum_{i=1}^{n} \lambda_i f(x_i)\right).$$

特别地，取 $\lambda_i = \frac{1}{n}(i=1,2,\cdots,n)$，就有

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

7.1.5 Cauchy's Mean Value Theorem 柯西中值定理

定理 7.9 (柯西中值定理) 若函数 $f(x)$ 与 $g(x)$ 满足：在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 内可导，且对任意 $x \in (a,b)$，$g'(x) \neq 0$，则在 (a,b) 内至少存在一点 ξ，使得

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

- 显然，当 $g(x)=x$ 时，上式即为拉格朗日公式，所以柯西中值定理为拉格朗日中值定理的推广，因此柯西中值定理又叫拓展中值定理。
- 若换个角度，将 $f(t)$ 和 $g(t)$ 看成 xy 平面上某条曲线 $y=F(x)$ 的参数方程，柯西中值定理也可以看成是拉格朗日中值定理的参数表达形式。
- 柯西中值定理的几何意义是：用参数方程表示的曲线上至少有一点，在这一点处的切线平行于连接两个端点的弦。

7.2 L'Hospital's Rule 洛必达法则

7.2.1 Indeterminate Forms and L'Hospital's Rule 未定式和洛必达法则

定理 7.10 设函数 $f(x)$ 和 $g(x)$ 在 $(a, a+d]$ 上可导（d 是某个正常数），且 $g'(x) \neq 0$。若此时有

$$\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$$

或

$$\lim_{x\to a^+} g(x) = \infty,$$

且 $\lim_{x\to a^+} \frac{f'(x)}{g'(x)}$ 存在（可以是有限数或 ∞），则成立

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{x\to a^+} \frac{f'(x)}{g'(x)}.$$

- 注意洛必达法则使用的三个前提：(1) 未定式 (2) 去心邻域内可导 (3) 求导后极限要存在
- 可以多次重复使用，每次使用之前要注意检查是否符合使用条件。
- 若洛必达后结果是极限不存在，则无法使用洛必达法则，法则失效。
- 若洛必达后求出了极限，不要再使用洛必达，因为此时已不是未定式，后续法则失效。
- 洛必达使用的经验：(1) 及时使用等价无穷小替换 (2) 及时外提极限存在且非 0 的项

7.3 Taylor's Formula 泰勒公式

7.3.1 Taylor's Formula with Peano Remainder 带皮亚诺余项的泰勒公式

我们已经知道如果 $f(x)$ 在 x_0 处可微，那么在 x_0 附近就有

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)$$

这意味着当我们在 $x = x_0$ 附近用一次多项式 $f(x_0) + f'(x_0)(x - x_0)$ 近似替代 $f(x)$ 时，其精确度对于 $x - x_0$ 而言只达到一阶。为了提高精确度，必须考虑用更高次数的多项式作逼近。事实上，当 $f(x)$ 在 x_0 处 n 阶可导时，有如下更精确的估计：

定理 7.11 (带皮亚诺余项的泰勒公式) 设 $f(x)$ 在 x_0 处有 n 阶导数，则存在 x_0 的一个邻域，对于该邻域中的任一点 x，成立

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x),$$

其中余项 $r_n(x)$ 满足

$$r_n(x) = o((x - x_0)^n).$$

上述公式称为 $f(x)$ 在 $x = x_0$ 处的**带皮亚诺余项的泰勒公式**，它的前 $n+1$ 项组成的多项式

$$p_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

称为 $f(x)$ 的 n 次 **泰勒多项式**，余项 $r_n(x) = o((x - x_0)^n)$ 称为 **皮亚诺余项**。

我们希望找到一个关于 $(x-x_0)$ 的 n 次多项式 $p_n(x)$，使得它在 x_0 附近能很好地逼近 $f(x)$。我们将其写作：

$$p_n(x)=a_0+a_1(x-x_0)+a_2(x-x_0)^2+\cdots+a_n(x-x_0)^n$$

其中 $a_0,a_1,\ldots,a_n$ 是待定的系数。

“很好地逼近”的一个自然标准是：让 $p_n(x)$ 和 $f(x)$ 在 x_0 点从 0 阶到 n 阶的导数值都相等。即：

$$\begin{cases} p_n(x_0)=f(x_0) \\ p_n'(x_0)=f'(x_0) \\ p_n''(x_0)=f''(x_0) \\ \vdots \\ p_n^{(n)}(x_0)=f^{(n)}(x_0) \end{cases}$$

现在我们根据上述标准来解出系数 $a_0,a_1,\ldots,a_n$。

首先求 a_0。计算 $p_n(x_0)$：

$$p_n(x_0)=a_0+a_1\cdot 0+a_2\cdot 0+\cdots+a_n\cdot 0=a_0$$

根据 $p_n(x_0)=f(x_0)$，我们得到：

$$a_0=f(x_0)$$

接着求 a_1。先求一阶导数：

$$p_n'(x)=a_1+2a_2(x-x_0)+3a_3(x-x_0)^2+\cdots+na_n(x-x_0)^{n-1}$$

代入 $x=x_0$ 得：

$$p_n'(x_0)=a_1$$

根据 $p_n'(x_0)=f'(x_0)$，我们得到：

$$a_1=f'(x_0)$$

然后求 a_2。求二阶导数：

$$p_n''(x)=2a_2+3\cdot 2\cdot a_3(x-x_0)+\cdots+n(n-1)a_n(x-x_0)^{n-2}$$

代入 $x=x_0$ 得：

$$p_n''(x_0)=2a_2$$

根据 $p_n''(x_0)=f''(x_0)$，我们得到：

$$2a_2=f''(x_0)\quad\Rightarrow\quad a_2=\frac{f''(x_0)}{2!}$$

以此类推，对于一般的系数 a_k，我们求 k 阶导数：

$$p_n^{(k)}(x)=k!\cdot a_k+(\text{一些含有}(x-x_0)\text{ 的项})$$

代入 $x=x_0$ 后，所有含 $(x-x_0)$ 的项都变为 0，只剩下：

$$p_n^{(k)}(x_0)=k!\cdot a_k$$

根据 $p_n^{(k)}(x_0) = f^{(k)}(x_0)$，我们得到：

$$a_k = \frac{f^{(k)}(x_0)}{k!}$$

将所有这些求得的系数代回最初的多项式 $p_n(x)$，我们就得到了著名的泰勒多项式：

$$p_n(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

现在，我们定义余项 $r_n(x)$ 为真实函数值与多项式近似值之间的误差：

$$r_n(x) = f(x) - p_n(x)$$

我们的目标是证明 $r_n(x) = o((x-x_0)^n)$，即：

$$\lim_{x\to x_0} \frac{r_n(x)}{(x-x_0)^n} = 0$$

证明思路是反复使用洛必达法则。观察极限 $\lim_{x\to x_0} \frac{r_n(x)}{(x-x_0)^n}$。当 $x \to x_0$ 时，分子 $r_n(x) \to 0$，分母也 $\to 0$，这是一个 $\frac{0}{0}$ 型未定式。

根据我们之前的构造，由于 $f(x)$ 和 $p_n(x)$ 在 x_0 点的 0 到 n 阶导数都相等，所以它们的差 $r_n(x)$ 满足：

$$r_n(x_0) = r_n'(x_0) = r_n''(x_0) = \cdots = r_n^{(n)}(x_0) = 0$$

接下来，我们对分子分母连续使用 n 次洛必达法则：

$$\begin{aligned}
\lim_{x\to x_0} \frac{r_n(x)}{(x-x_0)^n} &\overset{\text{洛必达}}{=} \lim_{x\to x_0} \frac{r_n'(x)}{n(x-x_0)^{n-1}} \\
&\overset{\text{洛必达}}{=} \lim_{x\to x_0} \frac{r_n''(x)}{n(n-1)(x-x_0)^{n-2}} \\
&\vdots \\
&\overset{\text{第 n 次洛必达}}{=} \lim_{x\to x_0} \frac{r_n^{(n)}(x)}{n!} \\
&= \frac{1}{n!}\left[\lim_{x\to x_0}\left(f^{(n)}(x) - p_n^{(n)}(x)\right)\right]
\end{aligned}$$

注意到 $p_n^{(n)}(x)$ 是一个常数，恒等于 $f^{(n)}(x_0)$。而根据题设，$f(x)$ 在 x_0 处有 n 阶导数，所以 $\lim_{x\to x_0} f^{(n)}(x) = f^{(n)}(x_0)$。

因此，我们最终得到：

$$\lim_{x\to x_0} \frac{r_n(x)}{(x-x_0)^n} = \frac{1}{n!}\left[f^{(n)}(x_0) - f^{(n)}(x_0)\right] = 0$$

这就严格证明了 $r_n(x) = o((x-x_0)^n)$。

7.3.2 Taylor's Formula with Lagrange Remainder 带拉格朗日余项的泰勒公式

注意到带 Peano 余项的泰勒公式只是对余项给出了定性描述，其对应的定量版本如下：

定理 7.12 (带拉格朗日余项的泰勒公式) *设 $f(x)$ 在 $[a,b]$ 上具有 n 阶连续导数，且在 (a,b) 上有 $n+1$ 阶导数．设 $x_0 \in [a,b]$ 为一定点，则对于任意 $x \in [a,b]$，成立*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + r_n(x),$$

其中余项 $r_n(x)$ 满足

$$r_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1},\quad \xi \text{ 在}x \text{ 和}x_0 \text{ 之间}.$$

上述公式称为 $f(x)$ 在 $x=x_0$ 处的**带拉格朗日余项的泰勒公式**. 余项称为 **拉格朗日余项**.

7.3.3 Maclaurin's Formula 麦克劳林公式

首先我们考虑函数 $f(x)$ 在 $x=0$ 处的泰勒公式，即

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+r_n(x),$$

其中 $r_n(x)$ 有皮亚诺余项与拉格朗日余项两种表示形式，即有 $r_n(x)=o(x^n)$，
或 $r_n(x)=\frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1},\theta\in(0,1)$。函数 $f(x)$ 在 $x=0$ 处的泰勒公式又称为函数 $f(x)$ 的 **麦克劳林公式**。

1. 指数函数 e^x (对所有 $x\in\mathbb{R}$ 收敛)

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+R_n(x)$$

$$R_n(x)=\frac{e^{\theta x}}{(n+1)!}x^{n+1}\quad\text{或}\quad e^x=\sum_{k=0}^{n}\frac{x^k}{k!}+o(x^n)$$

2. 自然对数 $\ln(1+x)$ (对 $-1<x\le 1$ 收敛)

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots+(-1)^{n-1}\frac{x^n}{n}+R_n(x)$$

$$R_n(x)=(-1)^n\frac{x^{n+1}}{(n+1)(1+\theta x)^{n+1}}\quad\text{或}\quad \ln(1+x)=\sum_{k=1}^{n}(-1)^{k-1}\frac{x^k}{k}+o(x^n)$$

3. 正弦函数 $\sin x$ (对所有 $x\in\mathbb{R}$ 收敛)

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots+(-1)^m\frac{x^{2m+1}}{(2m+1)!}+R_{2m+1}(x)$$

$$R_{2m+1}(x)=\frac{\sin\left(\theta x+(2m+2)\frac{\pi}{2}\right)}{(2m+2)!}x^{2m+2}\quad\text{或}\quad \sin x=\sum_{k=0}^{m}(-1)^k\frac{x^{2k+1}}{(2k+1)!}+o(x^{2m+1})$$

4. 余弦函数 $\cos x$ (对所有 $x\in\mathbb{R}$ 收敛)

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots+(-1)^m\frac{x^{2m}}{(2m)!}+R_{2m}(x)$$

$$R_{2m}(x)=\frac{\cos\left(\theta x+(2m+1)\frac{\pi}{2}\right)}{(2m+1)!}x^{2m+1}\quad\text{或}\quad \cos x=\sum_{k=0}^{m}(-1)^k\frac{x^{2k}}{(2k)!}+o(x^{2m})$$

5. 正切函数 $\tan x$ (收敛域与展开阶数有关)

$$\tan x=x+\frac{x^3}{3}+\frac{2x^5}{15}+\frac{17x^7}{315}+\cdots+o(x^{2m+1})$$

此展开式通常只给出佩亚诺余项形式。

6. 二项式函数 $(1+x)^\alpha$, $\alpha \in \mathbb{R}$ (收敛域与 α 有关，当 $|x|<1$ 时绝对收敛)

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + R_n(x)$$

$$R_n(x) = \frac{\alpha(\alpha-1)\cdots(\alpha-n)}{(n+1)!}(1+\theta x)^{\alpha-n-1}x^{n+1}$$

或

$$(1+x)^\alpha = \sum_{k=0}^{n}\binom{\alpha}{k}x^k + o(x^n), \quad \text{其中}\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$$

7. 平方根倒数（特殊情况）$\frac{1}{\sqrt{1+x}}$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 + \cdots + o(x^n)$$

8. 反正切函数 $\arctan x$ (对 $|x|\le 1$ 收敛)

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^m\frac{x^{2m+1}}{2m+1} + R_{2m+1}(x)$$

$$R_{2m+1}(x) = (-1)^{m+1}\frac{x^{2m+2}}{2m+2}\cdot\frac{1}{(1+\theta^2x^2)^{m+1}} \quad \text{或} \quad \arctan x = \sum_{k=0}^{m}(-1)^k\frac{x^{2k+1}}{2k+1} + o(x^{2m+1})$$

9. 双曲正弦 $\sinh x$ (对所有 $x\in\mathbb{R}$ 收敛)

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2m+1}}{(2m+1)!} + o(x^{2m+1})$$

10. 双曲余弦 $\cosh x$ (对所有 $x\in\mathbb{R}$ 收敛)

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2m}}{(2m)!} + o(x^{2m})$$

定理 7.13 设 $f(x)$ 在 x_0 的某个邻域有 $n+2$ 阶导数存在，则它的 $n+1$ 次泰勒多项式的导数恰为 $f'(x)$ 的 n 次泰勒多项式。

7.4 Applications of Taylor's Formula 泰勒公式的应用

7.4.1 Approximate Calculation 近似计算

7.4.2 Evaluating Limits 求极限

对于待定式的极限问题，一般可以采用洛必达法则来求。但对于一些求导比较繁琐，特别是要多次使用洛必达法则的情况，泰勒公式往往是比洛必达法则更为有效的求极限工具。

7.4.3 Proving Inequalities 证明不等式

7.4.4 Finding Asymptotic Equations of Curves 求曲线的渐近线方程

7.5 Applications of Differential Calculus 微分的应用

7.5.1 Local (Relative) Extreme Values 极值问题

设 x_0 为 $f(x)$ 的一个极值点（即极大值点或极小值点），如果 $f(x)$ 在 x_0 处可导，则由 Fermat 引理，必有 $f'(x_0)=0$。这就是说，$f(x)$ 的全部极值点必定都在使得 $f'(x)=0$ 和使得 $f'(x)$ 不

存在的点集之中。使 $f'(x)=0$ 的点称为 $f(x)$ 的**驻点**（stationary point）。$f(x)$ 的驻点（即使得 $f'(x)=0$ 的点）与使得 $f'(x)$ 不存在的点统称为 $f(x)$ 的**临界点**（critical point）。所以，我们可以先求出 $f(x)$ 的所有临界点，再进行判别。

定理 7.14 (极值点判定定理) *设函数 $f(x)$ 在 x_0 点的某一邻域中有定义，且 $f(x)$ 在 x_0 点连续。*

(1) 设存在 $\delta>0$，使得 $f(x)$ 在 $(x_0-\delta,x_0)$ 与 $(x_0,x_0+\delta)$ 上可导，

(i) 若在 $(x_0-\delta,x_0)$ 上有 $f'(x)\geq 0$，在 $(x_0,x_0+\delta)$ 上有 $f'(x)\leq 0$，则 x_0 是 $f(x)$ 的极大值点；

(ii) 若在 $(x_0-\delta,x_0)$ 上有 $f'(x)\leq 0$，在 $(x_0,x_0+\delta)$ 上有 $f'(x)\geq 0$，则 x_0 是 $f(x)$ 的极小值点；

(iii) 若 $f'(x)$ 在 $(x_0-\delta,x_0)$ 与 $(x_0,x_0+\delta)$ 上同号，则 x_0 不是 $f(x)$ 的极值点。

(2) 设 $f'(x_0)=0$，且 $f(x)$ 在 x_0 点二阶可导，

(i) 若 $f''(x_0)<0$，则 x_0 是 $f(x)$ 的极大值点；

(ii) 若 $f''(x_0)>0$，则 x_0 是 $f(x)$ 的极小值点；

(iii) 若 $f''(x_0)=0$，则 x_0 可能是 $f(x)$ 的极值点，也可能不是 $f(x)$ 的极值点。

注意 (1) 为极值点的一阶导数检验（First Derivative Test for Local Extrema），(2) 为极值点的二阶导数检验（Second Derivative Test for Local Extrema）。

7.5.2 Global (Absolute) Extreme Values 最值问题

在自然科学、生产技术、经济管理等领域，经常需要研究如何花费最小代价去获取最大收益的问题，这在许多情况下，可以归结为求一个函数在某一范围内的最大值或最小值问题。

根据连续函数的性质，闭区间上的连续函数必定能取到最大值与最小值。需要注意的是，如果去掉函数的连续性或者将闭区间改为开区间，函数有可能取不到最大值或最小值。

函数的最大值与最小值统称为函数的最值，使函数取到最大值（或最小值）的点称为函数的最大值点（或最小值点），也称为函数的最值点。

函数的极大值与极小值反映的是函数的一种局部性质，而函数的最大值与最小值反映的是函数在某一范围内的一种整体性质。

对于一个定义于闭区间 $[a,b]$ 上的函数 $f(x)$ 来说，区间的两个端点 a 与 b 是有可能成为它的最值点的。同时，若最值点属于开区间 (a,b) 的话，那它一定是函数的极值点。因此，我们只要按照前面在极值问题中所述的方法，找出所有 $f(x)$ 的驻点与使 $f'(x)$ 不存在的点，即所有 $f(x)$ 的临界点 (critical points)，再加上区间的端点（endpoints），从中找出使函数取最大值或最小值的点就可以了。

7.5.3 Curve Sketching 函数作图

在指定的坐标系中作出一个函数的图形，从而可以直观地去研究它的某些性质，这是很有实际意义的。现在虽说有了电子计算机和许多数学软件，可以（用描点法）画出各种各样的函数图形，但用分析的方法勾勒出函数的大致形状仍然是数学研究中的重要手段之一。

函数作图的过程一般可分为以下几个步骤：

（1）考察函数 $f(x)$ 的定义域及其在定义域内的连续性，找出函数的不连续点，并以这些点作为分点，将定义域分成若干个区间，使函数在每个区间上连续。

（2）计算 $f'(x)$，找出 $f(x)$ 的驻点与导数不存在的点，从而求出 $f(x)$ 的极值点与极值，并以这些点为分点，对区间进行再划分，使函数在每个区间上保持单调。

(3) 计算 $f''(x)$，找出所有使 $f''(x)=0$ 的点与使 $f''(x)$ 不存在的点，从而求出 $f(x)$ 的拐点，并以这些点为分点，继续对区间进行再划分，使函数在每个区间上保持固定的凸性。
(4) 对上述 (1)、(2)、(3) 三个步骤所得到的结果列出表格，在表格中标出函数在每个分点上的函数值（如果有定义的话），以及函数在每个区间上的单调性与凸性。
(5) 求出曲线 $y=f(x)$ 的渐近线，包括水平渐近线、垂直渐近线和斜渐近线。

通过上述步骤，我们就可以在平面坐标系上标出函数曲线的一些特殊点，如极值点与拐点等（如有需要的话，还可补充计算若干个点上 $f(x)$ 的值并定位于坐标系中），再根据曲线的凸性与渐近线的位置，就可作出函数 $y=f(x)$ 的图像。

须注意的是，在作图之前，应该先考察函数的几何性质如奇偶性、周期性等，如 $f(x)$ 是奇函数或偶函数，那么只要画出一半图形，而另一半可通过对称画出；对于周期函数，只要画出一个周期的图形就可以了，而其余部分可通过周期延拓画出。

7.5.4 Newton's Method 方程的近似求解：牛顿法

求方程的解主要方法有两种：解析方法和数值方法。

解析方法也称为公式法，它是将方程的解表达为方程的系数的函数形式，只要把待求解的方程的系数代入表达式，就可以求出方程的解。如果不考虑运算中的四舍五入所产生的误差，那么在理论上，解析方法所得的解是精确的，我们将这个解称为**解析解**或**精确解**，解析方法也因此被称为精确方法。（但由于真正运算时不可能不产生误差，因此从求解实际问题来说，不存在真正的精确方法。）

例如，对于一元二次方程

$$ax^2+bx+c=0 \quad (a\neq 0),$$

可以得到它的两个解为

$$x_1,x_2=\frac{-b\pm\sqrt{b^2-4ac}}{2a},$$

这就是在用解析方法求解方程。

但十分遗憾的是，除了我们在中学里已学过的简单的三角方程、对数方程和指数方程等情况之外，能精确求解的方程的数量和种类与实际需要求解的问题的数量相比，只能说是九牛之一毛。例如，形如

$$y=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

的多项式，可以算得是最简单的一类函数了，而著名法国数学家 Galois 在一个半世纪之前就证明了，当 $n\geq 5$ 时，对它不存在一般的求根公式。因此，对于更为复杂的超越函数，就更不能指望有什么普遍适用的，可以求得精确解的公式了。

数值方法是一种求近似解的方法。粗略地说，它无意去追究方程的解与系数本质上到底存在着什么样的联系，而只是设法去构造一个可实际计算的过程，并通过运行这个过程产生方程的精确解的一系列近似值。在一定的条件下，这些近似值理论上将收敛于方程的精确解，因此可以用精度较高的近似值来代替精确解，我们称其为**数值解**或**近似解**。由于实际问题中提出的许多多形态异的方程绝大多数都无法找到其解析解，因此，数值方法是用数学工具解决实际问题过程中的一个重要方法。

数值计算中常用的求近似值的方法是迭代法。先将原来的方程

$$f(x) = 0$$

化为等价形式

$$x = F(x)$$

这里的 $F(x)$ 称为**迭代函数**。

取一个适当的初始值 x_0，按

$$x_{k+1} = F(x_k), k = 0, 1, 2, \cdots$$

产生序列 $\{x_k\}$（设每个 x_k 都属于 $F(x)$ 的定义域），这样的计算过程称为迭代。若在理论上成立

$$x_k \to x^* \quad (k \to \infty)$$

那么显然 x^* 就是原方程的解，因此只要在迭代过程中，选取某个合适的 x_k 作为 x，就得到原方程的近似解了。

下面我们利用 Taylor 公式来举一个例子。设 $f(x)$ 在含有 x^* 的某个区间 $[a, b]$ 中具有二阶连续导数，且对于每个 $x \in [a, b]$，都有 $f'(x) \neq 0$。作出 $f(x)$ 在 x 处的 Taylor 公式，由于 x^* 是方程 $f(x) = 0$ 的解，则有

$$f(x^*) = f(x) + f'(x)(x^* - x) + f''(\xi)\frac{(x^* - x)^2}{2} = 0,$$

也即

$$x^* = x - \frac{f(x)}{f'(x)} - \frac{f''(\xi)}{f'(x)} \cdot \frac{(x^* - x)^2}{2},$$

当 $x \to x^*$ 时，上式的最后一项是趋向于零的，因此有

$$\lim_{x \to x^*} \left[x - \frac{f(x)}{f'(x)}\right] = x^* - \frac{f(x^*)}{f'(x^*)} = x^*.$$

这样，

$$F(x) = x - \frac{f(x)}{f'(x)}$$

就是一个满足 $x^* = F(x^*)$ 要求的迭代函数，由此得到迭代公式

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \cdots,$$

这就是著名的 **Newton 迭代法**（简称 **Newton 法**）。

定理 7.15 *设 $f(x)$ 在 $[a, b]$ 中有二阶连续导数，且满足条件*

(1) $f(a) \cdot f(b) < 0$；

(2) $f'(x)$ 在 (a, b) 保号；

(3) $f''(x)$ 在 (a, b) 保号。

取 x_0 是 a 和 b 中满足

$$f(x_0) \cdot f''(x_0) > 0$$

的那一个点，则以 x_0 为初值的 Newton 迭代过程

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \cdots$$

产生的序列 $\{x_k\}$ 单调收敛于方程

$$f(x) = 0$$

在 $[a, b]$ 中的惟一解。

7.5.5 Curvature of a Curve 曲线的曲率

$$K = \frac{|y''|}{(1+y'^2)^{\frac{3}{2}}}$$

8 Indefinite Integrals 不定积分

8.1 Concept and Properties of Indefinite Integrals 不定积分的概念和运算法则

8.1.1 Inverse Operation of Differentiation –The Indefinite Integral 微分的逆运算——不定积分

定义 8.1 若在某个区间上，函数 $F(x)$ 和 $f(x)$ 成立关系

$$F'(x) = f(x),$$

或等价地，

$$d(F(x)) = f(x)dx,$$

则称 $F(x)$ 是 $f(x)$ 在这个区间上的**一个原函数**（*antiderivative*）。

之所以要称“一个”原函数，是由于一个函数若存在原函数，那么它的原函数必定是不唯一的。比如，若 $F(x)$ 是 $f(x)$ 的原函数，即 $F'(x) = f(x)$，那么对任何常数 C，都有 $[F(x)+C]' = f(x)$，由定义可知 $F(x)+C$ 也是 $f(x)$ 的原函数，所以 $f(x)$ 的原函数有无穷多个。

可以验证，$f(x)$ 的所有原函数都具有 $F(x)+C$ 的形式，或者说，它们之间至多相差一个常数。若 $G(x)$ 是 $f(x)$ 的任一个原函数，则 $G'(x) = f(x)$，即 $[F(x) - G(x)]' = 0$，由拉格朗日中值定理的第一重要引理可知，$F(x) - G(x) \equiv C$，即 $G(x) = F(x) + C$。

所以，只要求出了 $f(x)$ 的任意一个原函数 $F(x)$，就可以用 $F(x)+C$ 来代表 $f(x)$ 的全部原函数了。

定义 8.2 一个函数 $f(x)$ 的原函数全体称为这个函数的不定积分，记作 $\int f(x)dx$。

这里，“$\int$”称为**积分号**（integral sign），$f(x)$ 称为**被积函数**（integrand），x 称为**积分变量**（variable of integration）。

8.1.2 Linearity of the Indefinite Integral 不定积分的线性性质

定理 8.1 (线性性) 若函数 $f(x)$ 和 $g(x)$ 的原函数都存在，则对任意常数 k_1 和 k_2，函数 $k_1f(x)+k_2g(x)$ 的原函数也存在，且有

$$\int [k_1 f(x) + k_2 g(x)]dx = k_1 \int f(x)dx + k_2 \int g(x)dx.$$

此式应理解为等式两端所表示的函数族相同。另外，当 $k_1 = k_2 = 0$ 时，等式右端应理解为常数 C。

根据微分与不定积分的关系以及基本初等函数的微分公式，我们可以得到一些最基本的不定积分公式：

微分	不定积分
$d(e^x) = e^x dx$	$\int e^x dx = e^x + C$
$d(\ln x) = \dfrac{dx}{x}$	$\int \dfrac{dx}{x} = \ln \vert x \vert + C$
$d(x^\alpha) = \alpha x^{\alpha-1} dx$	$\int x^\alpha dx = \dfrac{1}{\alpha+1} x^{\alpha+1} + C \quad (\alpha \neq -1)$
$d(\sin x) = \cos x dx$	$\int \cos x dx = \sin x + C$
$d(\cos x) = -\sin x dx$	$\int \sin x dx = -\cos x + C$
$d(\tan x) = \sec^2 x dx$	$\int \sec^2 x dx = \tan x + C$
$d(\cot x) = -\csc^2 x dx$	$\int \csc^2 x dx = -\cot x + C$
$d(\sec x) = \tan x \sec x dx$	$\int \tan x \sec x dx = \sec x + C$
$d(\csc x) = -\cot x \csc x dx$	$\int \cot x \csc x dx = -\csc x + C$
$d(\arcsin x) = \dfrac{dx}{\sqrt{1-x^2}}$	$\int \dfrac{dx}{\sqrt{1-x^2}} = \arcsin x + C$
$d(\arctan x) = \dfrac{dx}{1+x^2}$	$\int \dfrac{dx}{1+x^2} = \arctan x + C$

8.2 Integration by Substitution and by Parts 换元积分法和分部积分法

8.2.1 Integration by Substitution 换元积分法

第一类换元积分法的核心思想是“凑微分”。当我们计算积分 $\int f(x)dx$ 时，如果被积表达式 $f(x)dx$ 可以看作是一个复合函数 $f(g(x))$ 与其内层函数 $g(x)$ 的微分 $g'(x)dx$ 的乘积，即 $f(x)dx = f(g(x)) \cdot g'(x)dx$，那么我们就可以通过变量代换 $u = g(x)$，将原积分转化为对变量 u 的积分。其公式表示为：

$$\int f(g(x))g'(x)dx = \int f(u)du,$$

其中 $u = g(x)$，$du = g'(x)dx$。求出关于 u 的积分后，再将 $u = g(x)$ 代回，即可得到原积分的结果。

这种方法的关键在于识别出被积函数中“隐藏”的复合函数结构及其导数。

例如，计算积分 $\int 2x\cos(x^2)dx$。这里我们观察到 x^2 的导数是 $2x$，正好是余下的因子。因此，令 $u = x^2$，则 $du = 2xdx$，原积分变为：

$$\int \cos(u)du = \sin(u) + C.$$

最后代回 $u = x^2$，得到最终结果：

$$\int 2x\cos(x^2)dx = \sin(x^2) + C.$$

再举一例，$\int \frac{\ln x}{x}dx$。我们注意到 $(\ln x)' = \frac{1}{x}$，因此令 $u = \ln x$，则 $du = \frac{1}{x}dx$，积分变为：

$$\int udu = \frac{u^2}{2} + C = \frac{(\ln x)^2}{2} + C.$$

第二类换元积分法的思路与第一类相反，它适用于被积函数本身不易积分，但通过引入一个新的变量 $x = \varphi(t)$ 后，能简化积分形式的情况。其公式表示为：

$$\int f(x)dx \overset{x=\varphi(t)}{=} \int f(\varphi(t))\varphi'(t)dt.$$

求出关于 t 的积分后，再利用 $x=\varphi(t)$ 的反函数关系 $t=\varphi^{-1}(x)$ 代回原变量 x。

这种方法常用于处理含有根式的积分。经典的例子是含有 $\sqrt{a^2-x^2}$、$\sqrt{a^2+x^2}$ 或 $\sqrt{x^2-a^2}$ 的积分。例如，计算积分 $\int \frac{dx}{\sqrt{a^2-x^2}} \quad (a>0)$。直接积分很困难，我们利用三角恒等式 $1-\sin^2 t=\cos^2 t$ 来消去根号。令 $x=a\sin t$，其中 $t\in(-\frac{\pi}{2},\frac{\pi}{2})$，则 $dx=a\cos t dt$，且 $\sqrt{a^2-x^2}=\sqrt{a^2-a^2\sin^2 t}=a\cos t$。于是积分变为：

$$\int \frac{a\cos t dt}{a\cos t}=\int 1dt=t+C.$$

由于 $x=a\sin t$，即 $t=\arcsin(\frac{x}{a})$，所以最终结果为：

$$\int \frac{dx}{\sqrt{a^2-x^2}}=\arcsin\left(\frac{x}{a}\right)+C.$$

另一个典型例子是 $\int \frac{dx}{\sqrt{x^2+a^2}}$。这里我们利用三角恒等式 $1+\tan^2 t=\sec^2 t$。令 $x=a\tan t$，其中 $t\in(-\frac{\pi}{2},\frac{\pi}{2})$，则 $dx=a\sec^2 t dt$，且 $\sqrt{x^2+a^2}=a\sec t$。积分变为：

$$\int \frac{a\sec^2 t dt}{a\sec t}=\int \sec t dt=\ln|\sec t+\tan t|+C.$$

为了代回 x，我们根据 $x=a\tan t$ 构造一个直角三角形，可得 $\sec t=\frac{\sqrt{x^2+a^2}}{a}$。因此：

$$\int \frac{dx}{\sqrt{x^2+a^2}}=\ln\left|\frac{x}{a}+\frac{\sqrt{x^2+a^2}}{a}\right|+C_1=\ln|x+\sqrt{x^2+a^2}|+C.$$

其中 $C=C_1-\ln a$。

8.2.2 Integration by Parts 分部积分法

分部积分法的理论基础是函数乘积的微分公式。对任意两个可微的函数 $u(x),v(x)$，成立关系式

$$d[u(x)v(x)]=v(x)d[u(x)]+u(x)d[v(x)],$$

两边同时求不定积分并移项，就有

$$\int u(x)d[v(x)]=u(x)v(x)-\int v(x)d[u(x)],$$

也即

$$\int u(x)v'(x)dx=u(x)v(x)-\int v(x)u'(x)dx,$$

这就是**分部积分公式**。

粗略看来，分部积分公式只是把原来需要求的关于 $u(x)v'(x)$ 的不定积分改为求 $u'(x)v(x)$ 的不定积分而已，两者的形式又差不多，似乎并无多大的意义，其实不然。在许多时候，直接求 $\int u(x)v'(x)dx$ 与求 $\int v(x)u'(x)dx$ 相比，难度是不可同日而语的。

8.2.3 Table of Basic Integrals 基本积分表

$$\int x^\alpha dx = \begin{cases} \dfrac{1}{\alpha+1}x^{\alpha+1}+C, & \alpha \neq -1; \\ \ln|x|+C, & \alpha = -1; \end{cases} \qquad \int \ln x dx = x(\ln x - 1) + C$$

$$\int a^x dx = \frac{a^x}{\ln a} + C, \quad \text{特别地} \int e^x dx = e^x + C \qquad \int \sin x dx = -\cos x + C$$

$$\int \cos x dx = \sin x + C \qquad \int \tan x dx = -\ln|\cos x| + C$$

$$\int \cot x dx = \ln|\sin x| + C \qquad \int \sec x dx = \ln|\sec x + \tan x| + C$$

$$\int \csc x dx = \ln|\csc x - \cot x| + C \qquad \int \sinh x dx = \cosh x + C$$

$$\int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin\frac{x}{a} + C \qquad \int \frac{dx}{\sqrt{x^2 \pm a^2}} = \ln|x+\sqrt{x^2\pm a^2}| + C$$

$$\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C \qquad \int \frac{dx}{x^2+a^2} = \frac{1}{a}\arctan\frac{x}{a} + C$$

$$\int \sqrt{a^2-x^2}dx = \frac{1}{2}x\sqrt{a^2-x^2} + \frac{a^2}{2}\arcsin\frac{x}{a} + C \qquad \int \sqrt{x^2\pm a^2}dx = \frac{1}{2}\left(x\sqrt{x^2\pm a^2} \pm a^2\ln|x+\sqrt{x^2\pm a^2}|\right) + C$$

8.3 Integration of Rational Functions 有理函数的不定积分

形如 $\frac{p_m(x)}{q_n(x)}$ 的函数称为有理函数，这里 $p_m(x)$ 和 $q_n(x)$ 分别是 m 次和 n 次多项式。在本节中，我们将通过介绍求一般有理函数的不定积分的方法, 证明这样的一个结论: 有理函数的原函数一定是初等函数。

求有理函数的不定积分是我们在实际应用中经常遇到的问题, 此外, 对于求某些其他类型函数的不定积分, 如无理函数、三角函数的不定积分问题, 也可以通过适当的变换化成求有理函数的不定积分问题而得到解决。

在考虑有理函数的不定积分 $\int \frac{p_m(x)}{q_n(x)}dx$ 时, 我们总假定 $\frac{p_m(x)}{q_n(x)}$ 是真分式, 即成立 $m < n$, 因为不然的话, 可以通过多项式的带余除法, 使得

$$\frac{p_m(x)}{q_n(x)} = p_{m-n}(x) + \frac{r(x)}{q_n(x)},$$

其中 $p_{m-n}(x)$ 是 $m-n$ 次多项式, 而 $r(x)$ 是次数不超过 $n-1$ 的多项式. 这样就得到

$$\int \frac{p_m(x)}{q_n(x)}dx = \int p_{m-n}(x)dx + \int \frac{r(x)}{q_n(x)}dx.$$

由于求 $\int p_{m-n}(x)dx$ 非常容易, 原问题就变为求一个真分式类型的有理函数的不定积分了. 另外, 为了讨论的方便, 我们假定 $q_n(x)$ 的最高次项系数为 1.

由代数学基本定理, 分母多项式 $q_n(x)$ 在复数域上恰有 n 个根. 由于 $q_n(x)$ 是实多项式, 因此它的根要么是实根, 要么是成对出现的共轭复根. 设它的全部实根为 $\alpha_1, \alpha_2, \cdots, \alpha_l$, 其重数分别为 $m_1, m_2, \cdots, m_l$, 全部复根为 $\beta_1 \pm i\gamma_1, \beta_2 \pm i\gamma_2, \cdots, \beta_j \pm i\gamma_j$, 其重数分别为 $n_1, n_2, \cdots, n_j$ ($\sum_{k=1}^{l} m_k + 2\sum_{k=1}^{j} n_k = n$)，记 $\xi_k = -\beta_k$，$\eta_k^2 = \beta_k^2 + \gamma_k^2 (\xi_k^2 < \eta_k^2)$，则在实数域上可将 $q_n(x)$ 因式分解为

$$q_n(x) = \prod_{k=1}^{l}(x-\alpha_k)^{m_k} \cdot \prod_{k=1}^{j}(x^2+2\xi_k x+\eta_k^2)^{n_k}.$$

求有理函数的不定积分 $\int \frac{p_m(x)}{q_n(x)}dx$ 的关键，是将有理函数 $\frac{p_m(x)}{q_n(x)}$ 分解成简单分式之和，再分别求简单分式的不定积分。

定理 8.2 设有理函数 $\frac{p(x)}{q(x)}$ 是真分式, 多项式 $q(x)$ 有 k 重实根 α, 即 $q(x)=(x-\alpha)^k q_1(x), q_1(\alpha)\neq 0$. 则存在实数 λ 与多项式 $p_1(x)$, $p_1(x)$ 的次数低于 $(x-\alpha)^{k-1}q_1(x)$ 的次数, 成立

$$\frac{p(x)}{q(x)}=\frac{\lambda}{(x-\alpha)^k}+\frac{p_1(x)}{(x-\alpha)^{k-1}q_1(x)}.$$

定理 8.3 设有理函数 $\frac{p(x)}{q(x)}$ 是真分式, 多项式 $q(x)$ 有 l 重共轭复根 $\beta\pm i\gamma$, 即

$$q(x)=(x^2+2\xi x+\eta^2)^l q^*(x), q^*(\beta\pm i\gamma)\neq 0,$$

其中 $\xi=-\beta, \eta^2=\beta^2+\gamma^2(\xi^2<\eta^2)$。则存在实数 μ,ν 和多项式 $p^*(x)$, $p^*(x)$ 的次数低于 $(x^2+2\xi x+\eta^2)^{l-1}q^*(x)$ 的次数, 成立

$$\frac{p(x)}{q(x)}=\frac{\mu x+\nu}{(x^2+2\xi x+\eta^2)^l}+\frac{p^*(x)}{(x^2+2\xi x+\eta^2)^{l-1}q^*(x)}.$$

重复应用定理 8.2 与 8.3, 可将有理函数

$$\frac{p_m(x)}{q_n(x)}=\frac{p_m(x)}{\prod_{k=1}^{i}(x-\alpha_k)^{m_k}\cdot\prod_{k=1}^{j}(x^2+2\xi_k x+\eta_k^2)^{n_k}}$$

分解成简单分式之和, 分解的规律是, 若 $q_n(x)$ 含有因子 $(x-\alpha_k)^{m_k}$, 则在和式中就有项

$\dfrac{\lambda_{k_1}}{x-\alpha_k}, \dfrac{\lambda_{k_2}}{(x-\alpha_k)^2}, \cdots, \dfrac{\lambda_{km_k}}{(x-\alpha_k)^{m_k}}$;若$q_n(x)$ 含有因子$(x^2+2\xi_k x+\eta_k^2)^{n_k}$,则在和式中就有项

$$\frac{\mu_{k_1}x+\nu_{k_1}}{x^2+2\xi_k x+\eta_k^2}, \frac{\mu_{k_2}x+\nu_{k_2}}{(x^2+2\xi_k x+\eta_k^2)^2}, \cdots, \frac{\mu_{kn_k}x+\nu_{kn_k}}{(x^2+2\xi_k x+\eta_k^2)^{n_k}},$$

即

$$\frac{p_m(x)}{q_n(x)}=\sum_{k=1}^{i}\sum_{r=1}^{m_k}\frac{\lambda_{kr}}{(x-\alpha_k)^r}+\sum_{k=1}^{j}\sum_{r=1}^{n_k}\frac{\mu_{kr}x+\nu_{kr}}{(x^2+2\xi_k x+\eta_k^2)^r},$$

其中 $\lambda_{kr}, \mu_{kr}, \nu_{kr}$ 可以用待定系数法具体算出来。由不定积分的线性性质, 即知

$$\int\frac{p_m(x)}{q_n(x)}dx=\sum_{k=1}^{i}\sum_{r=1}^{m_k}\lambda_{kr}\int\frac{dx}{(x-\alpha_k)^r}+\sum_{k=1}^{j}\sum_{r=1}^{n_k}\int\frac{\mu_{kr}x+\nu_{kr}}{(x^2+2\xi_k x+\eta_k^2)^r}dx.$$

它所涉及的不定积分只有两种类型:

$$(1)\int\frac{dx}{(x-\alpha)^n}\quad (n\geq 1).$$

可以得出

$$\int\frac{dx}{(x-\alpha)^n}=\begin{cases}\ln|x-\alpha|+C, & n=1,\\ \frac{1}{n-1}\cdot\frac{1}{(x-\alpha)^{n-1}}+C, & n\geq 2.\end{cases}$$

$$(2)\int\frac{\mu x+\nu}{(x^2+2\xi x+\eta^2)^n}dx\quad (n\geq 1, \xi^2<\eta^2).$$

首先, 将原式化成

$$\int\frac{\mu x+\nu}{(x^2+2\xi x+\eta^2)^n}dx=\frac{\mu}{2}\int\frac{2x+2\xi}{(x^2+2\xi x+\eta^2)^n}dx+(\nu-\mu\xi)\int\frac{dx}{(x^2+2\xi x+\eta^2)^n}.$$

等号右端第一项中的积分可利用第一类换元积分方法,

$$\int\frac{2x+2\xi}{(x^2+2\xi x+\eta^2)^n}dx=\int\frac{d(x^2+2\xi x+\eta^2)}{(x^2+2\xi x+\eta^2)^n}=\begin{cases}\ln|x^2+2\xi x+\eta^2|+C, & n=1,\\ \frac{1}{n-1}\cdot\frac{1}{(x^2+2\xi x+\eta^2)^{n-1}}+C, & n\geq 2.\end{cases}$$

等号右端第二项中的积分可配成

$$\int \frac{dx}{(x^2+2\xi x+\eta^2)^n} = \int \frac{d(x+\xi)}{[(x+\xi)^2+(\eta^2-\xi^2)]^n},$$

可以计算得到

$$I_n = \int \frac{dx}{(x^2+2\xi x+\eta^2)^n}\text{的递推表达式}$$

$$I_n = \frac{1}{2(\eta^2-\xi^2)(n-1)}\left[(2n-3)I_{n-1} + \frac{x+\xi}{(x^2+2\xi x+\eta^2)^{n-1}}\right], \quad n \geq 2,$$

$$I_1 = \frac{1}{\sqrt{\eta^2-\xi^2}}\arctan\frac{x+\xi}{\sqrt{\eta^2-\xi^2}} + C.$$

特别地，

$$I_2 = \frac{1}{2(\sqrt{\eta^2-\xi^2})^3}\arctan\frac{x+\xi}{\sqrt{\eta^2-\xi^2}} + \frac{1}{2(\eta^2-\xi^2)}\cdot\frac{x+\xi}{x^2+2\xi x+\eta^2}.$$

至此，在理论上，求有理函数的不定积分问题已经圆满地得到了解决，并且上面的结果也告诉我们，有理函数的原函数一定是初等函数。

9 Definite Integrals 定积分

9.1 The Concept of the Definite Integral and Integrability Conditions 定积分的概念和可积条件

9.1.1 Definition of the Definite Integral 定积分的定义

设 $f(x)$ 是定义于 $[a,b]$ 上的函数，若要求对任意划分 P 和任意 $\xi_i \in [x_{i-1}, x_i]$，极限

$$\lim_{\lambda\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

都存在，则 $f(x)$ 必须是 $[a,b]$ 上的有界函数（自行证明）。为此，在定义函数的定积分时，我们首先要求函数是有界的。

定义 9.1 设 $f(x)$ 是定义在 $[a,b]$ 上的有界函数，在 $[a,b]$ 上任意取分点 $\{x_i\}_{i=0}^{n}$，作成一种划分

$$P: a = x_0 < x_1 < x_2 < \cdots < x_n = b,$$

并任意取点 $\xi_i \in [x_{i-1}, x_i]$。记小区间 $[x_{i-1}, x_i]$ 的长度为 $\Delta x_i = x_i - x_{i-1}$，并令 $\lambda = \max_{1\le i\le n}(\Delta x_i)$，若当 $\lambda \to 0$ 时，极限

$$\lim_{\lambda\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

存在，且极限值既与划分 P 无关，又与对 ξ_i 的取法无关，则称 $f(x)$ 在 $[a,b]$ 上**黎曼可积**（*Riemann integrable*）。和式

$$S_n = \sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

称为**黎曼和**，其极限值 I 称为 $f(x)$ 在 $[a,b]$ 上的**定积分**，记为

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

这里 a 和 b 分别被称为积分的下限和上限。

在上面的定义中，要求 $a < b$。当 $a \ge b$ 时，我们规定

$$\int_a^b f(x)dx = -\int_b^a f(x)dx,$$

并由此得到

$$\int_a^a f(x)dx = 0.$$

它们的几何意义是很明显的。

这一定义也可以用“$\varepsilon-\delta$ 语言”表述如下：

设有定数 I，对任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意一种划分

$$P: a = x_0 < x_1 < x_2 < \cdots < x_n = b,$$

和任意点 $\xi_i \in [x_{i-1}, x_i]$，只要

$$\lambda = \max_{1\le i\le n}(\Delta x_i) < \delta,$$

便有

$$\left|\sum_{i=1}^{n} f(\xi_i)\Delta x_i - I\right| < \varepsilon,$$

则称 $f(x)$ 在 $[a,b]$ 上黎曼可积，称 I 是 $f(x)$ 在 $[a,b]$ 上的定积分。

在不会发生混淆的情况下，一般就把黎曼可积简称为可积（以后我们会知道，还存在其他意义下的积分）。需要特别注意的是，“可积” 要求黎曼和的极限值与划分 P 以及 ξ_i 的取法无关。

9.1.2 Darboux Sums 达布和

记 $f(x)$ 在 $[a,b]$ 上的上确界和下确界分别为 M 和 m，则有

$$m \leq f(x) \leq M.$$

另外，记 $f(x)$ 在 $[x_{i-1}, x_i]$ 的上确界和下确界分别为 M_i 和 $m_i(i = 1, 2, \cdots, n)$，即

$$M_i = \sup\{f(x)|x \in [x_{i-1}, x_i]\} \text{ 和} m_i = \inf\{f(x)|x \in [x_{i-1}, x_i]\},$$

显然，它们与对 $[a,b]$ 所作的划分 P 有关。取定了划分 P 后，定义和式

$$\bar{S}(P) = \sum_{i=1}^{n} M_i\Delta x_i \quad \text{与} \underline{S}(P) = \sum_{i=1}^{n} m_i\Delta x_i,$$

它们分别被称为相应于划分 P 的**达布大和**与**达布小和**（统称为**达布和**），那么显然有

$$\underline{S}(P) \leq \sum_{i=1}^{n} f(\xi_i)\Delta x_i \leq \bar{S}(P).$$

从直观上容易看出，如果对任意一种划分 P，当 $\lambda = \max_{1\leq i\leq n}(\Delta x_i) \to 0$ 时，$\bar{S}(P)$ 和 $\underline{S}(P)$ 的极限都存在并且相等，那么 $f(x)$ 是可积的，反之亦然。下面，我们来严格证明这一结论。先引入以下引理:

引理 9.1 *若在原有划分中加入分点形成新的划分，则大和不增，小和不减。*

以下记 $\bar{\mathbf{S}}$ 是一切可能的划分所得到的达布大和的集合，而 $\underline{\mathbf{S}}$ 是一切可能的划分所得到的达布小和的集合。

引理 9.2 *对任意 $\bar{S}(P_1) \in \bar{\boldsymbol{S}}$ 和 $\underline{S}(P_2) \in \underline{\boldsymbol{S}}$，恒有*

$$m(b-a) \leq \underline{S}(P_2) \leq \bar{S}(P_1) \leq M(b-a).$$

由引理 9.2, $\bar{\mathbf{S}}$ 和 $\underline{\mathbf{S}}$ 都是有界集合，因此分别有下确界和上确界。记 $\bar{\mathbf{S}}$ 的下确界为

$$L = \inf\{\bar{S}(P)|\bar{S}(P) \in \bar{\mathbf{S}}\},$$

$\underline{\mathbf{S}}$ 的上确界为

$$l = \sup\{\underline{S}(P)|\underline{S}(P) \in \underline{\mathbf{S}}\},$$

则对任意 $\bar{S}(P_1) \in \bar{\mathbf{S}}$ 和 $\underline{S}(P_2) \in \underline{\mathbf{S}}$, 有

$$\underline{S}(P_2) \leq l \leq L \leq \bar{S}(P_1).$$

下面我们来证明，当 $\lambda = \max_{1\leq i\leq n}(\Delta x_i) \to 0$ 时，Darboux 大和与 Darboux 小和的极限确实存在，且分别等于它们各自的下确界和上确界。

引理 9.3 (达布定理) *对任意在 $[a,b]$ 上有界的函数 $f(x)$，恒有*

$$\lim_{\lambda\to 0} \bar{S}(P) = L, \quad \lim_{\lambda\to 0} \underline{S}(P) = l.$$

9.1.3 Necessary and Sufficient Conditions for Riemann Integrability 黎曼可积的充分必要条件

定理 9.1 有界函数 $f(x)$ 在 $[a,b]$ 可积的充分必要条件是，对于任意划分 P，当 $\lambda = \max(\Delta x_i) \to 0$ 时，达布大和与达布小和的极限相等，即成立

$$\lim_{\lambda\to 0} \bar{S}(P) = L = l = \lim_{\lambda\to 0} \underline{S}(P).$$

若记

$$\omega_i = M_i - m_i$$

为 $f(x)$ 在 $[x_{i-1}, x_i]$ 上的**振幅**，则定理 9.1 也可以等价地表述为

定理 9.2 有界函数 $f(x)$ 在 $[a,b]$ 可积的充分必要条件是，对任意划分，当 $\lambda = \max_{1\le i\le n}(\Delta x_i) \to 0$ 时，

$$\lim_{\lambda\to 0} \sum_{i=1}^{n} \omega_i \Delta x_i = 0.$$

由上述充分必要条件可以判断某些函数类的可积性。

推论 9.1 闭区间上的连续函数必定可积。

推论 9.2 闭区间上的单调函数必定可积。

定理 9.3 有界函数 $f(x)$ 在 $[a,b]$ 可积的充分必要条件是，对任意给定的 $\varepsilon > 0$，存在着一种划分，使得相应的振幅满足

$$\sum_{i=1}^{n} \omega_i \Delta x_i < \varepsilon.$$

推论 9.3 闭区间上只有有限个不连续点的有界函数必定可积。

9.2 Fundamental Properties of the Definite Integral 定积分的基本性质

定理 9.4 (性质 *1:* 线性性质) 设 $f(x)$ 和 $g(x)$ 都在 $[a,b]$ 上可积，k 和 h 是常数，则函数 $kf(x)+hg(x)$ 在 $[a,b]$ 上也可积，且

$$\int_a^b [kf(x) + hg(x)]dx = k\int_a^b f(x)dx + h\int_a^b g(x)dx.$$

推论 9.4 若 $f(x)$ 在 $[a,b]$ 上可积，而 $g(x)$ 只在有限个点上与 $f(x)$ 的取值不相同，则 $g(x)$ 在 $[a,b]$ 上也可积，并且有

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

定理 9.5 (性质 *2:* 乘积可积性) 设 $f(x)$ 和 $g(x)$ 都在 $[a,b]$ 上可积，则 $f(x)\cdot g(x)$ 在 $[a,b]$ 上也可积。

定理 9.6 (性质 *3:* 保序性) 设 $f(x)$ 与 $g(x)$ 在 $[a,b]$ 上可积，且对任意 $x \in [a,b]$ 有 $f(x) \le g(x)$，则

$$\int_a^b f(x)dx \le \int_a^b g(x)dx.$$

定理 9.7 (性质 *4:* 绝对可积性) 设 $f(x)$ 在 $[a,b]$ 上可积，则 $|f(x)|$ 在 $[a,b]$ 上也可积，且成立

$$\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx.$$

定理 9.8 (性质 *5:* 区间可加性) 设 $f(x)$ 在 $[a,b]$ 上可积，则对任意点 $c \in [a,b]$，$f(x)$ 在 $[a,c]$ 和 $[c,b]$ 上都可积；反过来，若 $f(x)$ 在 $[a,c]$ 和 $[c,b]$ 上都可积，则 $f(x)$ 在 $[a,b]$ 上可积。此时成立

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

定理 9.9 (性质 *6:* 积分第一中值定理) 设 $f(x)$ 和 $g(x)$ 都在 $[a,b]$ 上可积，$g(x)$ 在 $[a,b]$ 上不变号，则存在 $\eta \in [m,M]$，使得

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

这里 M 和 m 分别表示 $f(x)$ 在 $[a,b]$ 的上确界和下确界。特别地，若 $f(x)$ 在 $[a,b]$ 上连续，则存在 $\xi \in [a,b]$，使得

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

定理 9.10 设 $f(x)$ 在对称区间 $[-a,a]$ 上可积，

(1) 若 $f(x)$ 是偶函数，则成立

$$\int_{-a}^a f(x)dx = 2\int_0^a f(x)dx;$$

(2) 若 $f(x)$ 是奇函数，则成立

$$\int_{-a}^a f(x)dx = 0.$$

定理 9.11 设 $f(x)$ 是以 T 为周期的可积函数，则对任意 a，

$$\int_a^{a+T} f(x)dx = \int_0^T f(x)dx.$$

9.3 The Fundamental Theorem of Calculus（FTC）微积分基本定理

9.3.1 FTC Part 1: Derivative of an Integral with Variable Limits 微积分基本定理 1: 变限积分函数求导

设 $f(x)$ 在区间 $[a,b]$ 上可积，由定积分的区间可加性，可知对任意 $x \in [a,b]$，积分

$$\int_a^x f(t)dt$$

存在。当 x 在 $[a,b]$ 中变化时，

$$\int_a^x f(t)dt$$

的值也随之而变化，所以它是定义在 $[a,b]$ 上的关于 x 的函数。这个函数具有如下的重要性质：

定理 9.12 (变上限积分函数求导) 设 $f(x)$ 在 $[a,b]$ 上可积，作函数

$$F(x) = \int_a^x f(t)dt, \quad x \in [a,b],$$

则

(1) $F(x)$ 是 $[a,b]$ 上的连续函数；

(2) 若 $f(x)$ 在 $[a,b]$ 上连续，则 $F(x)$ 在 $[a,b]$ 上可微，且有

$$F'(x) = f(x).$$

例题 9.1

$$\left(\int_{-1}^{x} \sin t \, dt\right)' = \sin x$$

例题 9.2

$$\left(\int_{0}^{x} (\cos t + t^2 - 2) \, dt\right)' = \cos x + x^2 - 2$$

例题 9.3

$$\left(\int_{0}^{x} x^2 \sin t \, dt\right)' = \left(x^2 \int_{0}^{x} \sin t \, dt\right)' = 2x \int_{0}^{x} \sin t \, dt + x^2 \cdot \sin x = -2x\cos x + 2x + x^2 \sin x$$

该公式有 4 种变式，我们逐一来看：

例题 9.4 (变下限积分函数求导)

$$\left(\int_{x}^{0} \cos t dt\right)' = -\cos x$$

例题 9.5 (变上限 + 链式法则)

$$\frac{d}{dx}\left(\int_{a}^{\varphi(x)} f(t)dt\right) = \frac{d\int_{a}^{\varphi(x)} f(t)dt}{d\varphi(x)} \cdot \frac{d\varphi(x)}{dx} = f(\varphi(x)) \cdot \varphi'(x)$$

$$\begin{aligned} F'(x) &= \left(\int_{1}^{x^2+2x} (\sin t + e^t)dt\right)' \\ &= (\sin(x^2+2x) + e^{x^2+2x}) \cdot (2x+2) \\ &= 2(x+1)(\sin(x^2+2x) + e^{x^2+2x}) \end{aligned}$$

例题 9.6 (变下限 + 链式法则)

$$\frac{d}{dx}\left(\int_{\psi(x)}^{b} f(t)dt\right) = -f(\psi(x)) \cdot \psi'(x)$$

例题 9.7 (变上下限 + 链式法则)

$$\frac{d}{dx}\left(\int_{\psi(x)}^{\varphi(x)} f(t)dt\right) = f(\varphi(x)) \cdot \varphi'(x) - f(\psi(x)) \cdot \psi'(x)$$

$$\frac{d}{dx}\left[\int_{x^2}^{e^x} (t^3+2t)dt\right]$$

$$= [(e^x)^3 + 2(e^x)] \cdot e^x - [(x^2)^3 + 2(x^2)] \cdot (2x)$$

$$= (e^{3x} + 2e^x)e^x - (x^6 + 2x^2)(2x)$$

$$= e^{4x} + 2e^{2x} - 2x^7 - 4x^3$$

9.3.2 FTC Part 2: Newton-Leibniz Formula 微积分基本定理 2: 牛顿-莱布尼兹公式

定理 9.13 (牛顿-莱布尼兹公式) 设 $f(x)$ 在 $[a,b]$ 上连续，$F(x)$ 是 $f(x)$ 在 $[a,b]$ 上的一个原函数，则成立

$$\int_a^b f(x)dx = F(b) - F(a).$$

9.3.3 Integration by Parts and Substitution for Definite Integrals 定积分的分部积分法和换元积分法

定理 9.14 设 $u(x), v(x)$ 在区间 $[a,b]$ 上有连续导数，则

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]\Big|_a^b - \int_a^b v(x)u'(x)dx.$$

上式也能写成下列形式

$$\int_a^b u(x)dv(x) = [u(x)v(x)]\Big|_a^b - \int_a^b v(x)du(x).$$

定理 9.15 设 $f(x)$ 在区间 $[a,b]$ 上连续，$x = \varphi(t)$ 在区间 $[\alpha,\beta]$（或区间 $[\beta,\alpha]$）上有连续导数，其值域包含于 $[a,b]$，且满足 $\varphi(\alpha) = a$ 和 $\varphi(\beta) = b$，则

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.$$

9.4 Applications of Definite Integrals in Geometry 定积分在几何计算中的应用

9.4.1 Area of a Plane Region 求平面图形面积

求平面图形面积的题目可以分为两类：X 型和 Y 型。

9.4.2 Volume of a Solid of Revolution 求旋转体体积

9.4.3 Surface Area of a Solid of Revolution 求旋转曲面的面积

9.4.4 Arc Length of a Plane Curve 求平面曲线的弧长

	直角坐标显式方程 $y = f(x),\ x \in [a,b]$	直角坐标参数方程 $\begin{cases} x = x(t) \\ y = y(t) \end{cases},\ t \in [T_1, T_2]$	极坐标方程 $r = r(\theta),\ \theta \in [\alpha,\beta]$
平面图形面积	$\int_a^b f(x)dx$	$\int_{T_1}^{T_2} \lvert y(t)x'(t)\rvert dt$	$\frac{1}{2}\int_\alpha^\beta r^2(\theta)d\theta$
弧长的微分	$ds = \sqrt{1+[f'(x)]^2}dx$	$ds = \sqrt{[x'(t)]^2+[y'(t)]^2}dt$	$ds = \sqrt{r^2(\theta)+r'^2(\theta)}d\theta$
曲线弧长	$\int_a^b \sqrt{1+[f'(x)]^2}dx$	$\int_{T_1}^{T_2} \sqrt{[x'(t)]^2+[y'(t)]^2}dt$	$\int_\alpha^\beta \sqrt{r^2(\theta)+r'^2(\theta)}d\theta$
旋转体体积	$\pi\int_a^b [f(x)]^2dx$	$\pi\int_{T_1}^{T_2} y^2(t)\lvert x'(t)\rvert dt$	$\frac{2}{3}\pi\int_\alpha^\beta r^3(\theta)\sin\theta d\theta$
旋转曲面面积	$2\pi\int_a^b \lvert f(x)\rvert\sqrt{1+[f'(x)]^2}dx$	$2\pi\int_{T_1}^{T_2} \lvert y(t)\rvert\sqrt{x'^2(t)+y'^2(t)}dt$	$2\pi\int_\alpha^\beta r(\theta)\sin\theta\sqrt{r^2(\theta)+r'^2(\theta)}d\theta$

10 Improper Integrals 反常积分

10.1 Concept and Evaluation of Improper Integrals 反常积分的概念和计算

前面讨论 Riemann 积分时，首先假定了积分区间 $[a, b]$ 有限且被积函数 $f(x)$ 在 $[a, b]$ 上有界，但在实际中经常会碰到不满足这两个条件，却确实需要求积分的情况。所以，我们有必要突破 Riemann 积分的限制条件，考虑积分区间无限或被积函数无界的积分问题，这样的积分称为**反常积分**（或**广义积分**），而以前学过的 Riemann 积分相应地称为**正常积分**（或**常义积分**）。

定义 10.1 (第一类反常积分：无穷区间上的积分) 设函数 $f(x)$ 在 $[a, +\infty)$ 有定义，且在任意有限区间 $[a, A] \subset [a, +\infty)$ 上可积，若极限

$$\lim_{A\to+\infty}\int_a^A f(x)dx$$

存在，则称反常积分 $\int_a^{+\infty} f(x)dx$ 收敛（或称 $f(x)$ 在 $[a, +\infty)$ 上可积），其积分值为

$$\int_a^{+\infty} f(x)dx = \lim_{A\to+\infty}\int_a^A f(x)dx;$$

否则称反常积分 $\int_a^{+\infty} f(x)dx$ 发散。

对反常积分 $\int_{-\infty}^a f(x)dx$ 与 $\int_{-\infty}^{+\infty} f(x)dx$ 可类似地给出敛散性定义。注意只有当 $\int_{-\infty}^a f(x)dx$ 和 $\int_a^{+\infty} f(x)dx$ 都收敛时，才认为 $\int_{-\infty}^{+\infty} f(x)dx$ 是收敛的。

例题 10.1

$$\begin{aligned}\int_1^{+\infty}\frac{1}{x}dx &= \lim_{A\to+\infty}\int_1^A\frac{1}{x}dx \\ &= \lim_{A\to+\infty}\left[\ln|x|\right]_1^A \\ &= \lim_{A\to+\infty}(\ln A - \ln 1) \\ &= \lim_{A\to+\infty}\ln A \\ &= +\infty\end{aligned}$$

由于极限不存在（为 $+\infty$），所以该反常积分发散。

定义 10.2 (第二类反常积分：无界函数的积分/瑕积分) 设函数 $f(x)$ 在 $x = b$ 的左邻域无界，若对于任意 $\eta \in (0, b-a)$，$f(x)$ 在区间 $[a, b-\eta]$ 上有界可积，且极限

$$\lim_{\eta\to0^+}\int_a^{b-\eta} f(x)dx$$

存在，则称反常积分 $\int_a^b f(x)dx$ 收敛（或称无界函数 $f(x)$ 在 $[a, b]$ 上可积），其积分值为

$$\int_a^b f(x)dx = \lim_{\eta\to0^+}\int_a^{b-\eta} f(x)dx;$$

否则称反常积分 $\int_a^b f(x)dx$ 发散。

$x = a$ 为瑕点和 $x = c \in (a, b)$ 为瑕点的情况可以类似定义。当 $x = c$ 为瑕点时，注意只有当 $\int_a^c f(x)dx$ 和 $\int_c^b f(x)dx$ 都收敛时，才认为 $\int_a^b f(x)dx$ 是收敛的，且规定

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

例题 10.2

$$\begin{aligned}\int_0^1 \frac{1}{x}dx &= \lim_{\eta\to 0^+}\int_\eta^1 \frac{1}{x}dx \\ &= \lim_{\eta\to 0^+}[\ln|x|]_\eta^1 \\ &= \lim_{\eta\to 0^+}(\ln 1-\ln\eta) \\ &= \lim_{\eta\to 0^+}(-\ln\eta) \\ &= +\infty\end{aligned}$$

由于极限不存在（为 $+\infty$），所以该反常积分发散。

关于定积分的性质，对于反常积分大多相应成立，如线性性质、保序性、区间可加性等；但也有一些性质，如乘积可积性，却不再成立。定积分的一切计算法则，如线性运算、换元积分法、分部积分法等，也都可以平行地用于反常积分。

10.2 Cauchy Principal Value（PV）柯西主值

柯西主值的概念会在复变函数中详细引入，这里仅作简单介绍。

考虑 $y=\frac{1}{x}$ 从 -1 到 1 的积分，注意它是一个发散的反常积分，一个常见的误区是认为它是奇函数，所以根据对称性积分应该为 0。

例题 10.3 由于被积函数 $f(x)=\frac{1}{x}$ 在 $x=0$ 处无界，该点为瑕点。根据定义，需要将积分拆分为两部分：

$$\int_{-1}^1 \frac{1}{x}dx = \int_{-1}^0 \frac{1}{x}dx + \int_0^1 \frac{1}{x}dx$$

$$\begin{aligned}\int_{-1}^0 \frac{1}{x}dx &= \lim_{A\to 0^-}\int_{-1}^A \frac{1}{x}dx \\ &= \lim_{A\to 0^-}[\ln|x|]_{-1}^A \\ &= \lim_{A\to 0^-}(\ln|A|-\ln 1) \\ &= \lim_{A\to 0^-}\ln|A| \\ &= -\infty\end{aligned}$$

所以，$\int_{-1}^0 \frac{1}{x}dx$ 发散。

$$\begin{aligned}\int_0^1 \frac{1}{x}dx &= \lim_{B\to 0^+}\int_B^1 \frac{1}{x}dx \\ &= \lim_{B\to 0^+}[\ln|x|]_B^1 \\ &= \lim_{B\to 0^+}(\ln 1-\ln B) \\ &= \lim_{B\to 0^+}(-\ln B) \\ &= +\infty\end{aligned}$$

所以，$\int_0^1 \frac{1}{x}dx$ 也发散。

由于 $\int_{-1}^0 \frac{1}{x}dx$ 和 $\int_0^1 \frac{1}{x}dx$ 都发散，因此原反常积分 $\int_{-1}^1 \frac{1}{x}dx$ 发散。

为什么这里的正无穷部分和负无穷部分无法抵消？本质原因在于 A 和 B 趋近于 0^- 和 0^+ 的速率未必是相同的。按照反常积分的定义，在计算反常积分时 $A \to 0^-$ 与 $B \to 0^+$ 是独立的两个过程，因此正负两部分面积抵消过后结果可能是正的可能是负的也可能是 0。所以在反常积分中我们规定两部分只要有一部分是发散的我们就认为这个积分是发散的。当然，如果我们人为框定两个极限的趋近是同步的，那么自然这个积分结果就为 0，我们称它为该积分的**柯西主值**。

例题 10.4

$$PV\int_{-1}^{1}\frac{1}{x}\,dx = \lim_{\varepsilon\to 0^+}\left[\int_{-1}^{-\varepsilon}\frac{1}{x}\,dx + \int_{\varepsilon}^{1}\frac{1}{x}\,dx\right].$$

$$\int_{-1}^{-\varepsilon}\frac{1}{x}\,dx = [\ln|x|]_{-1}^{-\varepsilon} = \ln(\varepsilon) - \ln(1) = \ln\varepsilon,$$

$$\int_{\varepsilon}^{1}\frac{1}{x}\,dx = [\ln|x|]_{\varepsilon}^{1} = \ln 1 - \ln\varepsilon = -\ln\varepsilon.$$

$$PV = \lim_{\varepsilon\to 0^+}[\ln\varepsilon + (-\ln\varepsilon)] = \lim_{\varepsilon\to 0^+} 0 = 0.$$

当反常积分收敛时，显然柯西主值等于积分值，而当反常积分发散时，柯西主值也有可能存在，因此柯西主值推广了反常积分的收敛概念。

柯西主值在某些领域中有独到的作用。

10.3 Convergence Tests for Improper Integrals 反常积分的判敛

由于一般的被积函数的原函数并不一定是初等函数，而且即使是初等函数，也往往不容易求出，因此有必要建立反常积分敛散性的判别法。事实上，在理论研究和实际应用中，经常会遇到只需要确定一个反常积分的收敛性，而不一定需要求出它的积分值的情况。

本节我们将介绍一套非常实用的反常积分判敛的思路。首先，我们需要熟悉两个基准工具：p 积分和 q 积分。

p 积分：$\displaystyle\int_a^{+\infty}\frac{1}{x^p}dx \begin{cases}\text{收敛,} & p>1\\ \text{发散,} & p\le 1\end{cases}$

q 积分：$\displaystyle\int_a^{b}\frac{1}{(x-a)^q}dx \begin{cases}\text{收敛,} & 0<q<1\\ \text{发散,} & q\ge 1\end{cases}$

我们通常把瑕点和无穷大并称为“奇点”或者“反常点”。在判断敛散性的时候我们首先要找到积分的奇点，如果有多个，我们需要将积分拆成多个积分的和，使每个子积分只包含一个奇点，继而逐一处理。积分包含无穷大这个奇点是比较明显的，因为在积分上下限里就会体现，但瑕点是需要额外留意的，若无其他特别形式，一般瑕点为使被积函数分母为 0 的点。

利用这两个已知敛散性的积分作比较的标尺，我们可以求目标函数与 p/q 积分被积函数的比值在奇点的极限。拿已知是发散的函数举例，若目标函数比上它的极限是 0，则判断不出敛散性；若极限是无穷大，则说明目标函数还要更加发散，反常积分是发散的；若是其他定值（非零常数）那说明敛散性是同阶的，反常积分发散。同理，与已知收敛的函数做比值，只能判断出比它更收敛或者同阶收敛的函数。

有了这个判别法，我们可以尝试把目标函数与某个 p/q 积分作比求极限，如果判断不出再根据选取的工具是太收敛了还是太发散了进行调整，直到能判断出来。但这个流程的本质是猜和凑，我们完全可以优化一下这个方法，还是同样的原理，只不过我们现在把已知函数的阶数 p/q 当未知变量，反过来建立含极限的方程来求出与目标函数敛散性同阶的 p/q 的值来判断，这样可以最大程度地提高比值判别法的效率。

$$\lim_{x\to\infty}\frac{f(x)}{\frac{1}{x^p}}=x^pf(x)=l$$

$$\lim_{x\to a^+}\frac{f(x)}{\frac{1}{(x-a)^q}}=(x-a)^qf(x)=l$$

$$\lim_{x\to a^-}\frac{f(x)}{\frac{1}{(a-x)^q}}=(a-x)^qf(x)=l$$

当然，并不是所有函数都能均可解得这样的 p/q 值，比较常见的特殊类型有如下三类：$e^x, x\to\infty$、$\ln x, x\to\infty$、$\ln x, x\to 0$。这些遇到了需要单独分析。

例题 10.5 判断 $\int_1^{+\infty}\frac{x^2\arctan x}{1+x^3}dx$ 的敛散性

$$\lim_{x\to+\infty}\frac{f(x)}{1/x^p}=\lim_{x\to+\infty}\frac{x^{p+2}\arctan x}{1+x^3}=l$$

$$x\to+\infty,\sim\frac{x^{p+2}}{x^3}=x^{p-1},p=1$$

因 $p=1$，参考积分发散，故原积分发散。

例题 10.6 判断 $\int_1^{+\infty}\frac{1}{\sqrt[3]{x^4+100}}dx$ 的敛散性

$$\lim_{x\to+\infty}\frac{f(x)}{1/x^p}=\lim_{x\to+\infty}\frac{x^p}{\sqrt[3]{x^4}}=x^{p-\frac{4}{3}}=l$$

解得 $p=\frac{4}{3}>1$，参考积分收敛，故原积分收敛。

例题 10.7 判断 $\int_0^{+\infty}\frac{1}{1+\sqrt{x}}dx$ 的敛散性

$$\lim_{x\to+\infty}\frac{f(x)}{1/x^p}=\lim_{x\to+\infty}\frac{x^p}{\sqrt{x}}=x^{p-\frac{1}{2}}=l$$

解得 $p=\frac{1}{2}<1$，参考积分发散，故原积分发散。

10.4 专题练习：一元函数积分学的计算

（1）不定积分的积分法

例题 10.8 求不定积分

$$\int e^{\frac{\sin\theta}{\cos\theta+\sin\theta}}\cdot\frac{1}{(\cos\theta+\sin\theta)^2}d\theta$$

。

对复杂部分求导：

$$\left(\frac{\sin\theta}{\cos\theta+\sin\theta}\right)'=\frac{\cos\theta(\cos\theta+\sin\theta)-\sin\theta(-\sin\theta+\cos\theta)}{(\cos\theta+\sin\theta)^2}=\frac{1}{(\cos\theta+\sin\theta)^2}$$

故

$$d\left(\frac{\sin\theta}{\cos\theta+\sin\theta}\right)=\frac{1}{(\cos\theta+\sin\theta)^2}d\theta,$$

于是

$$\text{原式}=\int e^{\frac{\sin\theta}{\cos\theta+\sin\theta}}d\left(\frac{\sin\theta}{\cos\theta+\sin\theta}\right)=e^{\frac{\sin\theta}{\cos\theta+\sin\theta}}+C.$$

例题 10.9 求不定积分

$$\int\frac{xe^x}{\sqrt{e^x-1}}dx$$

。

令 $u=\sqrt{e^x-1}$，则 $x=\ln(1+u^2)$，$dx=\frac{2u}{1+u^2}du$，从而

$$\int\frac{xe^x}{\sqrt{e^x-1}}dx=\int\frac{(1+u^2)\ln(1+u^2)}{u}\cdot\frac{2u}{1+u^2}du=2\int\ln(1+u^2)du$$

$$=2u\ln(1+u^2)-\int\frac{4u^2}{1+u^2}du=2u\ln(1+u^2)-4u+4\arctan u+C$$

$$=2x\sqrt{e^x-1}-4\sqrt{e^x-1}+4\arctan\sqrt{e^x-1}+C$$

例题 10.10 设 $f(\ln x)=\frac{\ln(1+x)}{x}$，计算 $\int f(x)dx$。

设 $\ln x=t$，则 $x=e^t$，$f(t)=\frac{\ln(1+e^t)}{e^t}$，故

$$\int f(x)dx=\int\frac{\ln(1+e^x)}{e^x}dx=-\int\ln(1+e^x)d(e^{-x})$$

$$=-e^{-x}\ln(1+e^x)+\int\frac{1}{1+e^x}dx$$

$$=-e^{-x}\ln(1+e^x)+\int\left(1-\frac{e^x}{1+e^x}\right)dx$$

$$=-e^{-x}\ln(1+e^x)+x-\ln(1+e^x)+C$$

$$=x-(1+e^{-x})\ln(1+e^x)+C$$

例题 10.11 计算不定积分 $\int e^{2x}(\tan x+1)^2dx$。

$$
\begin{aligned}
\int e^{2x}(\tan x+1)^2dx &= \int e^{2x}(\sec^2 x+2\tan x)dx \\
&= \int e^{2x}\sec^2 xdx+2\int e^{2x}\tan xdx \\
&= \int e^{2x}d(\tan x)+2\int e^{2x}\tan xdx \\
&= e^{2x}\tan x-2\int e^{2x}\tan xdx+2\int e^{2x}\tan xdx \\
&= e^{2x}\tan x+C
\end{aligned}
$$

（2）定积分的计算

例题 10.12 求

$$
\int_{-1}^{1} x^2\sqrt{1-x^2}dx
$$

。

$$
\begin{aligned}
\text{原式} &= 2\int_0^1 x^2\sqrt{1-x^2}dx = 2\int_0^{\frac{\pi}{2}}\sin^2 t\cos^2 tdt \\
&= 2\int_0^{\frac{\pi}{2}}\sin^2 t(1-\sin^2 t)dt = 2\left(\int_0^{\frac{\pi}{2}}\sin^2 tdt-\int_0^{\frac{\pi}{2}}\sin^4 tdt\right) \\
&= 2\left(\frac{1}{2}\times\frac{\pi}{2}-\frac{3}{4}\times\frac{1}{2}\times\frac{\pi}{2}\right) = 2\left(\frac{\pi}{4}-\frac{3\pi}{16}\right) = 2\times\frac{\pi}{16} = \frac{\pi}{8}.
\end{aligned}
$$

（3）变限积分的计算

例题 10.13 求曲线 $y=\int_0^{\sin x}e^{t^2}dt$ 在点 *(0, 0)* 处的法线方程。

欲求曲线在给定点处的法线方程，应先检查此点是否在曲线上，如果此点在曲线上，再求该点处切线的斜率，最后利用点斜式求法线方程。

易知点 $(0,0)$ 在曲线 $y=\int_0^{\sin x}e^{t^2}dt$ 上。

由于 $y'=e^{\sin^2 x}\cdot\cos x$，$y'|_{x=0}=1$，可知切线斜率 $k=1$，法线斜率为 $-\frac{1}{k}=-1$，因此所求法线方程为 $y=-x$。

（4）反常积分的计算

例题 10.14 计算反常积分

$$
\int_2^3\frac{dx}{\sqrt{x-x^2}}
$$

。

注意到被积函数含有绝对值符号且 $x=1$ 是其无穷间断点，故

原式 =

$$
\int_1^2\frac{dx}{\sqrt{x-x^2}}+\int_1^3\frac{dx}{\sqrt{x^2-x}}
$$

而

$$
\int_1^2\frac{dx}{\sqrt{x-x^2}} = \int_1^2\frac{dx}{\sqrt{\frac{1}{4}-\left(x-\frac{1}{2}\right)^2}} = \arcsin\left(2x-1\right)\Big|_1^2 = \arcsin 1 = \frac{\pi}{2}
$$

$$\int_1^3 \frac{dx}{\sqrt{x^2-x}} = \int_1^3 \frac{dx}{\sqrt{\left(x-\frac{1}{2}\right)^2-\frac{1}{4}}} = \ln\left[\left(x-\frac{1}{2}\right)+\sqrt{\left(x-\frac{1}{2}\right)^2-\frac{1}{4}}\right]_1^3$$
$$= \ln(2+\sqrt{3})$$

因此

$$\int_2^3 \frac{dx}{\sqrt{x-x^2}} = \frac{\pi}{2} + \ln(2+\sqrt{3})$$

习题 10.1 若

$$\int xf(x)dx = \arcsin x + C$$

，则

$$\int \frac{1}{f(x)}dx = \underline{\qquad\qquad\qquad}$$

。

习题 10.2 若

$$f(x) = \frac{1}{1+x^2} + x^3 \int_0^1 f(x)dx$$

，则

$$\int_0^1 f(x)dx = \underline{\qquad\qquad\qquad}$$

。

习题 10.3 极限

$$\lim_{x\to+\infty} \frac{\int_c^x \left(1-\frac{1}{t}\right)^t \cdot e^{ct}dt}{e^{ct}} = \underline{\qquad\qquad\qquad}$$

。

习题 10.4 设 $f(x)$ 的一个原函数为 $\ln^2 x$，则

$$\int xf'(x)dx = \underline{\qquad\qquad\qquad}$$

。

习题 10.5

$$\int \frac{\arcsin\sqrt{x}}{\sqrt{x}}dx = \underline{\qquad\qquad\qquad}$$

。

习题 10.6

$$\int_2^\infty \frac{dx}{(x+7)\sqrt{x-2}} = \underline{\qquad\qquad\qquad}$$

。

习题 10.7 计算

$$\int \arcsin\sqrt{\frac{x}{a+x}}dx$$

（a 是大于 *0* 的常数）。

习题 10.8 求

$$\int \frac{\arctan e^x}{e^x}dx$$

。

习题 10.9 求

$$\int \max\{1,|x|\}dx$$

。

习题 10.10 求定积分

$$\int_0^{\frac{3}{4}} \frac{1}{1+\cos^2 x} dx$$

。

习题 10.11 计算

$$\int_{\frac{\pi}{4}}^{\frac{\pi}{4}} e^{x/2} \frac{\cos x - \sin x}{\sqrt{\cos x}} dx$$

。

习题 10.12 计算

$$I = \int_0^{\pi} \frac{x \sin x}{1+\cos^2 x} dx$$

。

习题 10.13 求

$$\int_{-1}^{1} \frac{x+1}{1+\sqrt[3]{x^2}} dx$$

。

习题 10.14 设

$$f(x) = \begin{cases} \frac{1}{1+\sin x}, & x \geq 0, \\ \frac{1}{1+e^x}, & x < 0, \end{cases}$$

求

$$\int_{-1}^{\frac{\pi}{4}} f(x) dx$$

。

习题 10.15 设 $x \geq -1$，求

$$\int_{-1}^{x} (1-|t|) dt$$

。

习题 10.16 求连续函数 $f(x)$，使它满足

$$\int_0^1 f(tx) dt = f(x) + x \sin x$$

。

习题 10.17 设

$$f(x) = \int_0^1 t|t-x| dt$$

，求 $f'(x)$。

11 Series of Numbers 数项级数

11.1 Convergence of Number Series 数项级数的收敛性

11.1.1 Number Series 数项级数

设 $x_1, x_2, \cdots, x_n, \cdots$ 是无穷可列个实数，我们称它们的“和”

$$x_1 + x_2 + \cdots + x_n + \cdots$$

为**无穷数项级数**（简称**级数**），记为 $\sum_{n=1}^{\infty} x_n$，其中 x_n 称为级数的**通项**或**一般项**。

当然，我们无法直接对无穷多个实数逐一地进行加法运算，所以必须对上述的级数求和给出合理的定义。为此作级数 $\sum_{n=1}^{\infty} x_n$ 的“部分和数列” $\{S_n\}$：

$S_1 = x_1,$

$S_2 = x_1 + x_2,$

$S_3 = x_1 + x_2 + x_3,$

…………

$S_n = x_1 + x_2 + ... + x_n = \displaystyle\sum_{k=1}^{n} x_k,$

…………

定义 11.1 如果部分和数列 $\{S_n\}$ 收敛于有限数 S，则称无穷级数 $\sum_{n=1}^{\infty} x_n$ 收敛，且称它的和为 S，记为

$$S = \sum_{n=1}^{\infty} x_n;$$

如果部分和数列 $\{S_n\}$ 发散，则称无穷级数 $\sum_{n=1}^{\infty} x_n$ 发散。

由上述定义可知，只有当无穷级数收敛时，无穷多个实数的加法才是有意义的，并且它们的和就是级数的部分和数列的极限。所以，级数的收敛与数列的收敛本质上是一回事。

例题 11.1 设 $|q| < 1$，则几何级数（即等比级数）

$$\sum_{n=1}^{\infty} q^{n-1} = 1 + q + q^2 + \cdots + q^n + \cdots$$

是收敛的。这是因为它的部分和数列的通项为

$$S_n = \sum_{k=1}^{n} q^{k-1} = \frac{1-q^n}{1-q}$$

显然，

$$\lim_{n\to\infty} S_n = \frac{1}{1-q}$$

例题 11.2

$$\sum_{n=1}^{\infty} (-1)^{n-1} = 1 - 1 + 1 - \cdots + (-1)^{n-1} + \cdots$$

是发散的。这是因为它的部分和数列的通项为

$$S_n = \begin{cases} 0, & n\text{为偶数}, \\ 1, & n\text{为奇数}. \end{cases}$$

显然 $\{S_n\}$ 是发散的。

例题 11.3 级数

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} + \cdots \quad (p > 0)$$

当 $p > 1$ 时收敛；当 $0 < p \leq 1$ 时发散到正无穷大。

$\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ 称为 **p 级数**($p = 1$ 时又称 $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ 为**调和级数**)，它在判别其他级数的敛散情况时有重要作用。

当 $p = 2$ 时，这个级数的和是一个非常著名且精美的结果，被称为**巴塞尔问题**:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

当级数 $\sum_{n=1}^{\infty} x_n$ 收敛时，我们还可以构作它的“**余和数列**” $\{r_n\}$，其中

$$r_n = \sum_{k=n+1}^{\infty} x_k = x_{n+1} + x_{n+2} + \cdots,$$

设 $\sum_{n=1}^{\infty} x_n = S$，则 $r_n = S - S_n$，显然，这时 $\{r_n\}$ 收敛于 0。

11.1.2 Basic Properties of Series 级数的基本性质

定理 11.1 (级数收敛的必要条件) 若 $\displaystyle\sum_{n=1}^{\infty} x_n$ 收敛，则其通项所构成的数列 $\{x_n\}$ 是无穷小量，即

$$\lim_{n \to \infty} x_n = 0.$$

定理 11.2 (线性性) 设 $\displaystyle\sum_{n=1}^{\infty} a_n = A$, $\displaystyle\sum_{n=1}^{\infty} b_n = B$，$\alpha, \beta$ 是两个常数，则

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha A + \beta B.$$

例题 11.4 求级数 $\displaystyle\sum_{n=1}^{\infty} \frac{4^{n+1} - 3 \cdot 2^n}{5^n}$ 的值。

因为几何级数 $\displaystyle\sum_{n=0}^{\infty} \left(\frac{4}{5}\right)^n$ 与 $\displaystyle\sum_{n=0}^{\infty} \left(\frac{2}{5}\right)^n$ 都收敛，所以有

$$\sum_{n=1}^{\infty} \frac{4^{n+1} - 3 \cdot 2^n}{5^n} = \frac{16}{5} \sum_{n=0}^{\infty} \left(\frac{4}{5}\right)^n - \frac{6}{5} \sum_{n=0}^{\infty} \left(\frac{2}{5}\right)^n = \frac{16}{5} \cdot \frac{1}{1 - \frac{4}{5}} - \frac{6}{5} \cdot \frac{1}{1 - \frac{2}{5}} = 14.$$

定理 11.3 设级数 $\displaystyle\sum_{n=1}^{\infty} x_n$ 收敛，则在它的求和表达式中任意添加括号后所得的级数仍然收敛，且其和不变。

例题 11.5 计算级数 $\displaystyle\sum_{n=1}^{\infty} \frac{2n-1}{2^n}$。

设级数的部分和数列为 $\{S_n\}$，则

$$\begin{aligned} S_n = 2S_n - S_n &= 2\sum_{k=1}^{n} \frac{2k-1}{2^k} - \sum_{k=1}^{n} \frac{2k-1}{2^k} \\ &= \sum_{k=0}^{n-1} \frac{2k+1}{2^k} - \sum_{k=1}^{n} \frac{2k-1}{2^k} = 1 + \sum_{k=1}^{n-1} \frac{1}{2^{k-1}} - \frac{2n-1}{2^n}, \end{aligned}$$

于是

$$\lim_{n\to\infty} S_n = 1 + \sum_{k=0}^{\infty} \frac{1}{2^k} = 3.$$

例题 11.6 讨论下列级数的敛散性。收敛的话，试求出级数之和。

(1) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+2)}$

(2) $\displaystyle\sum_{n=1}^{\infty} \frac{2n}{3n+1}$

(3) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)}$

(4) $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{2^n} - \frac{1}{3^n}\right)$

(5) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt[n]{n}}$

(6) $\displaystyle\sum_{n=1}^{\infty} \frac{5^{n-1}+4^{n+1}}{3^{2n}}$

(7) $\displaystyle\sum_{n=1}^{\infty} \left(\sqrt{n+2} - 2\sqrt{n+1} + \sqrt{n}\right)$

(8) $\displaystyle\sum_{n=1}^{\infty} \frac{2n-1}{3^n}$

(9) $\displaystyle\sum_{n=0}^{\infty} q^n \cos n\theta \quad (|q|<1)$

(1) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+2)}$

裂项：$\displaystyle\frac{1}{n(n+2)} = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+2}\right)$

部分和：

$$S_N = \frac{1}{2}\left(1 + \frac{1}{2} - \frac{1}{N+1} - \frac{1}{N+2}\right) \to \frac{1}{2} \times \frac{3}{2} = \frac{3}{4}$$

收敛，和为 $\dfrac{3}{4}$。

(2) $\displaystyle\sum_{n=1}^{\infty} \frac{2n}{3n+1}$

通项 $a_n \to \dfrac{2}{3} \neq 0$，发散。

(3) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)}$

裂项：

$$\frac{1}{n(n+1)(n+2)} = \frac{1}{2}\left[\frac{1}{n(n+1)} - \frac{1}{(n+1)(n+2)}\right]$$

部分和：

$$S_N = \frac{1}{2}\left[\frac{1}{1\cdot 2} - \frac{1}{(N+1)(N+2)}\right] \to \frac{1}{4}$$

收敛，和为 $\frac{1}{4}$。

(4) $\sum_{n=1}^{\infty}\left(\frac{1}{2^n} - \frac{1}{3^n}\right)$

$$\sum_{n=1}^{\infty}\frac{1}{2^n} = 1, \quad \sum_{n=1}^{\infty}\frac{1}{3^n} = \frac{1}{2}$$

和为 $1 - \frac{1}{2} = \frac{1}{2}$，收敛。

(5) $\sum_{n=1}^{\infty}\frac{1}{\sqrt[n]{n}}$

通项 $a_n = n^{-1/n} \to 1 \neq 0$，发散。

(6) $\sum_{n=1}^{\infty}\frac{5^{n-1}+4^{n+1}}{3^{2n}}$

$$= \frac{1}{5}\sum_{n=1}^{\infty}\left(\frac{5}{9}\right)^n + 4\sum_{n=1}^{\infty}\left(\frac{4}{9}\right)^n = \frac{1}{5}\cdot\frac{5/9}{4/9} + 4\cdot\frac{4/9}{5/9} = \frac{1}{4} + \frac{16}{5} = \frac{69}{20}$$

收敛，和为 $\frac{69}{20}$。

(7) $\sum_{n=1}^{\infty}\left(\sqrt{n+2} - 2\sqrt{n+1} + \sqrt{n}\right)$

$$= (\sqrt{n+2} - \sqrt{n+1}) - (\sqrt{n+1} - \sqrt{n})$$

部分和：

$$S_N = (\sqrt{N+2} - \sqrt{2}) - (\sqrt{N+1} - 1) \to 1 - \sqrt{2}$$

收敛，和为 $1-\sqrt{2}$。

(8) $\sum_{n=1}^{\infty}\frac{2n-1}{3^n}$

$$= 2\sum_{n=1}^{\infty}\frac{n}{3^n} - \sum_{n=1}^{\infty}\frac{1}{3^n} = 2\cdot\frac{3}{4} - \frac{1}{2} = 1$$

收敛，和为 1。

(9) $\sum_{n=0}^{\infty} q^n \cos n\theta \quad (|q|<1)$

用复数法：

$$\sum_{n=0}^{\infty} q^n \cos n\theta = \mathrm{Re}\left(\frac{1}{1-qe^{i\theta}}\right) = \frac{1-q\cos\theta}{1-2q\cos\theta+q^2}$$

收敛。

11.2 Limit Superior and Limit Inferior 上极限与下极限

11.2.1 Limit Superior and Limit Inferior of Sequences 数列的上极限和下极限

定义 11.2 在有界数列 $\{x_n\}$ 中，若存在它的一个子列 $\{x_{n_k}\}$ 使得

$$\lim_{k\to\infty} x_{n_k} = \xi,$$

则称 ξ 为数列 $\{x_n\}$ 的一个极限点。

显然，“ξ 是数列 $\{x_n\}$ 的极限点”也可以等价地表述为：“对于任意给定的 $\varepsilon > 0$，存在 $\{x_n\}$ 中的无穷多个项属于 ξ 的 ε 邻域”。

记

$$E = \{\xi \mid \xi \text{ 是}\{x_n\} \text{ 的极限点}\},$$

则 E 显然是非空的有界集合，因此，E 的上确界 $H = \sup E$ 和下确界 $h = \inf E$ 存在。

定理 11.4 E 的上确界 H 和下确界 h 均属于 E，即

$$H = \max E, \quad h = \min E.$$

定义 11.3 E 的最大值 $H = \max E$ 称为数列 $\{x_n\}$ 的**上极限**，记为

$$H = \varlimsup_{n\to\infty} x_n;$$

E 的最小值 $h = \min E$ 称为数列 $\{x_n\}$ 的**下极限**，记为

$$h = \varliminf_{n\to\infty} x_n.$$

定理 11.5 设 $\{x_n\}$ 是有界数列，则 $\{x_n\}$ 收敛的充分必要条件是

$$\varlimsup_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n.$$

例题 11.7 求数列 $\left\{x_n = \cos\dfrac{2n\pi}{5}\right\}$ 的上极限与下极限。

因为 $x_{5n-4} = x_{5n-1} = \cos\frac{2\pi}{5}$，$x_{5n-3} = x_{5n-2} = -\cos\frac{\pi}{5}$，$x_{5n} = 1$，所以 $\{x_n\}$ 的最大极限点是 1，最小极限点是 $-\cos\frac{\pi}{5}$，即

$$\varlimsup_{n\to\infty} x_n = 1, \quad \varliminf_{n\to\infty} x_n = -\cos\frac{\pi}{5}.$$

例题 11.8 求数列 $\{x_n = n^{(-1)^n}\}$ 的上极限与下极限。

此数列为

$$1,\ 2,\ \frac{1}{3},\ 4,\ \frac{1}{5},\ 6,\ \frac{1}{7},\ 8, \cdots,$$

它没有上界，因而 $\varlimsup\limits_{n\to\infty} x_n = +\infty$。

又由 $x_n > 0$ 且 $\{x_{2n-1}\}$ 的极限为 0，即知

$$\varliminf_{n\to\infty} x_n = 0.$$

例题 11.9 求数列 $\{x_n = -n\}$ 的上极限与下极限。

由于 $\lim\limits_{n\to\infty} x_n = -\infty$，因而

$$\varlimsup_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n = \lim_{n\to\infty} x_n = -\infty.$$

11.2.2 Operations of Limit Superior and Limit Inferior 上极限和下极限的运算

定理 11.6 设 $\{x_n\},\{y_n\}$ 是两数列，则

(1)

$$\varlimsup_{n\to\infty}(x_n+y_n)\le\varlimsup_{n\to\infty}x_n+\varlimsup_{n\to\infty}y_n,\quad \varliminf_{n\to\infty}(x_n+y_n)\ge\varliminf_{n\to\infty}x_n+\varliminf_{n\to\infty}y_n;$$

(2) 若 $\lim\limits_{n\to\infty}x_n$ 存在，则

$$\varlimsup_{n\to\infty}(x_n+y_n)=\lim_{n\to\infty}x_n+\varlimsup_{n\to\infty}y_n,\quad \varliminf_{n\to\infty}(x_n+y_n)=\lim_{n\to\infty}x_n+\varliminf_{n\to\infty}y_n.$$

(要求上述诸式的右端不是待定型，即不为 $(+\infty)+(-\infty)$ 等。)

定理 11.7 设 $\{x_n\},\{y_n\}$ 是两数列，

(1) 若 $x_n\ge 0, y_n\ge 0$，则

$$\varlimsup_{n\to\infty}(x_ny_n)\le\varlimsup_{n\to\infty}x_n\cdot\varlimsup_{n\to\infty}y_n,\quad \varliminf_{n\to\infty}(x_ny_n)\ge\varliminf_{n\to\infty}x_n\cdot\varliminf_{n\to\infty}y_n;$$

(2) 若 $\varlimsup\limits_{n\to\infty}x_n=x,\ 0<x<+\infty$，则

$$\varlimsup_{n\to\infty}(x_ny_n)=\varlimsup_{n\to\infty}x_n\cdot\varlimsup_{n\to\infty}y_n,\quad \varliminf_{n\to\infty}(x_ny_n)=\varliminf_{n\to\infty}x_n\cdot\varliminf_{n\to\infty}y_n.$$

(要求上述诸式的右端不是待定型，即不为 $0\cdot(+\infty)$ 等。)

11.3 Series with Positive Terms 正项级数

11.3.1 Series with Positive Terms 正项级数

定义 11.4 如果级数 $\sum\limits_{n=1}^{\infty}x_n$ 的各项都是非负实数，即

$$x_n\ge 0,\quad n=1,2,\cdots,$$

则称此级数为**正项级数**。

显然，正项级数 $\sum\limits_{n=1}^{\infty}x_n$ 的部分和数列 $\{S_n\}$ 是单调增加的，即

$$S_n=\sum_{k=1}^{n}x_k\le\sum_{k=1}^{n+1}x_k=S_{n+1},\quad n=1,2,\cdots.$$

根据单调数列的性质，立刻可以得到

定理 11.8 (正项级数的收敛原理) 正项级数收敛的充分必要条件是它的部分和数列有上界。

若正项级数的部分和数列无上界，则其必发散到 $+\infty$。

例题 11.10 级数 $\sum\limits_{n=2}^{\infty}\left[\dfrac{1}{\sqrt{n}}\ln\dfrac{n^2}{(n-1)(n+1)}\right]$ 是正项级数。它的部分和数列的通项

$$S_n=\sum_{k=2}^{n+1}\left[\frac{1}{\sqrt{k}}\ln\frac{k^2}{(k-1)(k+1)}\right]<\sum_{k=2}^{n+1}\left(\ln\frac{k}{k-1}-\ln\frac{k+1}{k}\right)=\ln 2-\ln\frac{n+2}{n+1}<\ln 2,$$

所以正项级数 $\sum\limits_{n=2}^{\infty}\left[\dfrac{1}{\sqrt{n}}\ln\dfrac{n^2}{(n-1)(n+1)}\right]$ 收敛。

11.3.2 Comparison Tests 比较判别法

判断一个正项级数是否收敛，最常用的方法是用一个已知收敛或发散的级数与它进行比较。

定理 11.9 (比较判别法) 设 $\sum_{n=1}^{\infty} x_n$ 与 $\sum_{n=1}^{\infty} y_n$ 是两个正项级数，若存在常数 $A>0$，使得

$$x_n \leq A y_n, \quad n=1,2,\cdots,$$

则

(1) 当 $\sum_{n=1}^{\infty} y_n$ 收敛时，$\sum_{n=1}^{\infty} x_n$ 也收敛；

(2) 当 $\sum_{n=1}^{\infty} x_n$ 发散时，$\sum_{n=1}^{\infty} y_n$ 也发散。

注意，由于改变级数有限个项的数值，并不会改变它的收敛性或发散性（虽然在收敛的情况下可能改变它的“和”），所以定理的条件可放宽为：“存在正整数 N 与常数 $A>0$，使得 $x_n \leq A y_n$ 对一切 $n>N$ 成立”。

例题 11.11 判断正项级数 $\sum_{n=1}^{\infty} \frac{n+3}{2n^3-n}$ 的敛散性。

容易看出当 $n>3$ 时成立

$$\frac{n+3}{2n^3-n}<\frac{1}{n^2},$$

于是由 $\sum_{n=1}^{\infty} \frac{1}{n^2}$ 的收敛性，可知 $\sum_{n=1}^{\infty} \frac{n+3}{2n^3-n}$ 收敛。

例题 11.12 判断正项级数 $\sum_{n=1}^{\infty} \sin\frac{\pi}{n}$ 的敛散性。

由于当 $x \in [0,\frac{\pi}{2}]$ 时，成立不等式 $\sin x \geq \frac{2}{\pi}x$，所以当 $n \geq 2$ 时，

$$\sin\frac{\pi}{n} \geq \frac{2}{\pi}\cdot\frac{\pi}{n}=\frac{2}{n},$$

由于 $\sum_{n=1}^{\infty} \frac{1}{n}$ 是发散的，可知 $\sum_{n=1}^{\infty} \sin\frac{\pi}{n}$ 发散。

下述定理是上一个定理的极限形式，它在使用上更为方便。

定理 11.10 (比较判别法的极限形式) 设 $\sum_{n=1}^{\infty} x_n$ 与 $\sum_{n=1}^{\infty} y_n$ 是两个正项级数，且

$$\lim_{n\to\infty}\frac{x_n}{y_n}=l \quad (0 \leq l \leq +\infty),$$

则

(1) 若 $0 \leq l<+\infty$，则当 $\sum_{n=1}^{\infty} y_n$ 收敛时，$\sum_{n=1}^{\infty} x_n$ 也收敛；

(2) 若 $0<l \leq +\infty$，则当 $\sum_{n=1}^{\infty} y_n$ 发散时，$\sum_{n=1}^{\infty} x_n$ 也发散。

所以当 $0<l<+\infty$ 时，$\sum_{n=1}^{\infty} x_n$ 与 $\sum_{n=1}^{\infty} y_n$ 同时收敛或同时发散。

11.3.3 Cauchy's Root Test and d'Alembert's Ratio Test 柯西判别法与达朗贝尔判别法

用比较判别法时，先要对所考虑的级数的收敛性有一个大致估计，进而找一个敛散性已知的合适级数与之相比较。但就绝大多数情况而言，这两个步骤都具有相当难度，因此，理想的判别方法似应着眼于对级数自身元素的分析。

正项等比级数 $\sum\limits_{n=1}^{\infty} q^n$ $(q>0)$ 给我们以很重要的启示。众所周知，$\sum\limits_{n=1}^{\infty} q^n$ 的敛散性只依赖于其后项与前项之比，即公比 q 是否小于 1。直观地类比一下，容易想象，若一个级数 $\sum\limits_{n=1}^{\infty} x_n$ 的后项与前项之比 $\dfrac{x_{n+1}}{x_n}$ 或前 n 项的"平均公比" $\sqrt[n]{x_n}$（记 $x_0=1$，则 $\sqrt[n]{x_n}=\sqrt[n]{\dfrac{x_1}{x_0}\cdot\dfrac{x_2}{x_1}\cdots\dfrac{x_n}{x_{n-1}}}$）存在小于（或大于）1 的极限，则这个级数应该是收敛（或发散）的。若它们的极限不存在，那么可以通过讨论其上（下）极限来得到类似的结论。正是基于这样的思路，产生了如下的 Cauchy 判别法与 d'Alembert 判别法。

定理 11.11 (柯西判别法) 设 $\sum\limits_{n=1}^{\infty} x_n$ 是正项级数，$r=\varlimsup\limits_{n\to\infty}\sqrt[n]{x_n}$，则

(1) 当 $r<1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 收敛；

(2) 当 $r>1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 发散；

(3) 当 $r=1$ 时，判别法失效，即级数可能收敛，也可能发散。

定理 11.12 (达朗贝尔判别法) 设 $\sum\limits_{n=1}^{\infty} x_n$ $(x_n\neq 0)$ 是正项级数，则

(1) 当 $\varlimsup\limits_{n\to\infty}\dfrac{x_{n+1}}{x_n}=\bar{r}<1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 收敛；

(2) 当 $\varliminf\limits_{n\to\infty}\dfrac{x_{n+1}}{x_n}=\underline{r}>1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 发散；

(3) 当 $\bar{r}\geq 1$ 或 $\underline{r}\leq 1$ 时，判别法失效，即级数可能收敛，也可能发散。

推论 11.1 设 $\{x_n\}$ 是正项数列，则

$$\varliminf_{n\to\infty}\frac{x_{n+1}}{x_n}\leq\varliminf_{n\to\infty}\sqrt[n]{x_n}\leq\varlimsup_{n\to\infty}\sqrt[n]{x_n}\leq\varlimsup_{n\to\infty}\frac{x_{n+1}}{x_n}.$$

该引理告诉我们：若一个正项级数的敛散情况可以由达朗贝尔判别法判定，则它一定也能用柯西判别法来判定。但是，能用柯西判别法判定的，却未必能用达朗贝尔判别法判定。

11.3.4 Raabe's Test 拉贝判别法

对某些正项级数 $\sum\limits_{n=1}^{\infty} x_n$，成立 $\lim\limits_{n\to\infty}\dfrac{x_{n+1}}{x_n}=1$（或者说 $\lim\limits_{n\to\infty}\dfrac{x_n}{x_{n+1}}=1$），这时柯西判别法与达朗贝尔判别法都失效，下面给出一种针对这类情况的判别法。

定理 11.13 (拉贝判别法) 设 $\sum\limits_{n=1}^{\infty} x_n$ $(x_n\neq 0)$ 是正项级数，$\lim\limits_{n\to\infty} n\left(\dfrac{x_n}{x_{n+1}}-1\right)=r$，则

(1) 当 $r>1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 收敛；

(2) 当 $r<1$ 时，级数 $\sum\limits_{n=1}^{\infty} x_n$ 发散。

11.3.5 Integral Test 积分判别法

设 $f(x)$ 定义于 $[a,+\infty)$，并且 $f(x)\geq 0$，进一步设 $f(x)$ 在任意有限区间 $[a,A]$ 上 Riemann 可积。

取一单调增加趋于 $+\infty$ 的数列 $\{a_n\}$：

$$a=a_1<a_2<a_3<\cdots<a_n<\cdots,$$

令

$$u_n=\int_{a_n}^{a_{n+1}} f(x)\,dx.$$

定理 11.14 (积分判别法) 反常积分 $\displaystyle\int_a^{+\infty} f(x)\,dx$ 与正项级数 $\displaystyle\sum_{n=1}^{\infty} u_n$ 同时收敛或同时发散于 $+\infty$，且

$$\int_a^{\infty} f(x)\,dx=\sum_{n=1}^{\infty} u_n=\sum_{n=1}^{\infty}\int_{a_n}^{a_{n+1}} f(x)\,dx.$$

特别地，当 $f(x)$ 单调减少时，取 $a_n=n$，则反常积分 $\displaystyle\int_a^{\infty} f(x)\,dx$ 与正项级数

$$\sum_{n=N}^{\infty} f(n)\quad (N=[a]+1)$$

同时收敛或同时发散。

11.4 Series with Arbitrary Terms 任意项级数

11.4.1 Series with Arbitrary Terms 任意项级数

一个级数，如果只有有限个负项或有限个正项，都可以用正项级数的各种判别法来判断它的敛散性。如果一个级数既有无限个正项，又有无限个负项，那么正项级数的各种判别法不再适用。

为此，我们从正项级数转向讨论任意项级数，也就是通项任意地可正或可负的级数。

由于 Cauchy 收敛原理是对敛散性最本质的刻画，为了判断任意项级数的敛散性，我们将关于数列的 Cauchy 收敛原理应用于级数的情况，即可得到

定理 11.15 (级数的柯西收敛原理) 级数 $\displaystyle\sum_{n=1}^{\infty} x_n$ 收敛的充分必要条件是：对任意给定的 $\varepsilon>0$，存在正整数 N，使得

$$|x_{n+1}+x_{n+2}+\cdots+x_m|=\left|\sum_{k=n+1}^{m} x_k\right|<\varepsilon$$

对一切 $m>n>N$ 成立。

定理结论还可以叙述为：对任意给定的 $\varepsilon>0$，存在正整数 N，使得

$$|x_{n+1}+x_{n+2}+\cdots+x_{n+p}|=\left|\sum_{k=1}^{p} x_{n+k}\right|<\varepsilon$$

对一切 $n>N$ 与一切正整数 p 成立。

取 $p=1$，上式即为 $|x_{n+1}|<\varepsilon$，于是就得到级数收敛的必要条件 $\lim\limits_{n\to\infty} x_n=0$。

11.4.2 Leibniz Series (Alternating Series) 莱布尼兹级数

先考虑一类特殊的任意项级数。

定义 11.5 如果级数 $\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty}(-1)^{n+1}u_n\ (u_n > 0)$, 则称此级数为**交错级数**(*Alternating Series*)。

进一步，若级数 $\sum_{n=1}^{\infty}(-1)^{n+1}u_n\ (u_n > 0)$ 满足 $|u_n|$ 单调减少且收敛于 *0*，则称这样的交错级数为 **莱布尼兹级数**。

定理 11.16 (莱布尼兹判别法) 莱布尼兹级数必定收敛。

11.4.3 Abel's Test and Dirichlet's Test 阿贝尔判别法与狄利克雷判别法

引理 11.1 (阿贝尔变换) 设 $\{a_k\},\{b_k\}$ 是两数列，记 $B_k = \sum_{i=1}^{k} b_i\ (k = 1, 2, \dots)$，则

$$\sum_{k=1}^{p} a_k b_k = a_p B_p - \sum_{k=1}^{p-1}(a_{k+1} - a_k)B_k.$$

上式也称为**分部求和公式**。

利用阿贝尔变换即得到如下的阿贝尔引理。

引理 11.2 (阿贝尔引理) 设

(1) $\{a_k\}$ 为单调数列；

(2) $\{B_k\}$ $\left(B_k = \sum_{i=1}^{k} b_i,\ k = 1, 2, \cdots\right)$ 为有界数列，即存在 $M > 0$, 对一切 k, 成立 $|B_k| \le M$,

则

$$\left|\sum_{k=1}^{p} a_k b_k\right| \le M\left(|a_1| + 2|a_p|\right).$$

定理 11.17 (级数的 *A-D* 判别法) 若下列两个条件之一满足，则级数 $\sum_{n=1}^{\infty} a_n b_n$ 收敛：

*(1) (**阿贝尔判别法**)* $\{a_n\}$ 单调有界，$\sum_{n=1}^{\infty} b_n$ 收敛；

*(2) (**狄利克雷判别法**)* $\{a_n\}$ 单调趋于 *0*，$\left\{\sum_{i=1}^{n} b_i\right\}$ 有界。

11.4.4 Absolute Convergence and Conditional Convergence 级数的绝对收敛与条件收敛

由于正项级数的判别法较易使用，很自然地要问，能否利用它对任意项级数的敛散性先做一个粗略的判断呢?

由 Cauchy 收敛原理和三角不等式

$$|x_{n+1} + x_{n+2} + \cdots + x_m| \le |x_{n+1}| + |x_{n+2}| + \cdots + |x_m|,$$

很容易知道：若对一个数项级数 $\sum_{n=1}^{\infty} x_n$ 逐项取绝对值得到新的级数 $\sum_{n=1}^{\infty} |x_n|$，则当 $\sum_{n=1}^{\infty} |x_n|$ 收敛时必有 $\sum_{n=1}^{\infty} x_n$ 收敛。

显然，这个结论的逆命题是不成立的，即不能由 $\sum_{n=1}^{\infty} x_n$ 收敛断言 $\sum_{n=1}^{\infty} |x_n|$ 也收敛。

例如 Leibniz 级数 $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ 收敛，但对每项取绝对值后，得到的是调和级数 $\sum_{n=1}^{\infty} \frac{1}{n}$，它是发散的。

定义 11.6 如果级数 $\sum_{n=1}^{\infty} |x_n|$ 收敛，则称 $\sum_{n=1}^{\infty} x_n$ 为**绝对收敛级数**。如果级数 $\sum_{n=1}^{\infty} x_n$ 收敛而 $\sum_{n=1}^{\infty} |x_n|$ 发散，则称 $\sum_{n=1}^{\infty} x_n$ 为**条件收敛级数**。

由定义，$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ 即是一个条件收敛级数。

$\sum_{n=1}^{\infty} |x_n|$ 的敛散性可以采用正项级数敛散性的判别法来判定。需要指出，虽然一般说来，由 $\sum_{n=1}^{\infty} |x_n|$ 发散并不能得出 $\sum_{n=1}^{\infty} x_n$ 发散，但若用 Cauchy 判别法或 d'Alembert 判别法判断出 $\sum_{n=1}^{\infty} |x_n|$ 发散，则级数 $\sum_{n=1}^{\infty} x_n$ 本身一定发散，这是因为这两个判别法判定发散的依据是级数的通项不趋于 0，即不满足收敛的必要条件。

定理 11.18 若 $\sum_{n=1}^{\infty} x_n$ 绝对收敛，则 $\sum_{n=1}^{\infty} x_n^+$ 与 $\sum_{n=1}^{\infty} x_n^-$ 都收敛；若 $\sum_{n=1}^{\infty} x_n$ 条件收敛，则 $\sum_{n=1}^{\infty} x_n^+$ 与 $\sum_{n=1}^{\infty} x_n^-$ 都发散到 $+\infty$。

11.4.5 Commutative Law for Series 加法交换律

定理 11.19 若级数 $\sum_{n=1}^{\infty} x_n$ 绝对收敛，则它的更序级数 $\sum_{n=1}^{\infty} x'_n$ 也绝对收敛，且和不变，即

$$\sum_{n=1}^{\infty} x'_n = \sum_{n=1}^{\infty} x_n.$$

定理 11.20 设级数 $\sum_{n=1}^{\infty} x_n$ 条件收敛，则对任意给定的 a，$-\infty \le a \le +\infty$，必定存在 $\sum_{n=1}^{\infty} x_n$ 的更序级数 $\sum_{n=1}^{\infty} x'_n$ 满足

$$\sum_{n=1}^{\infty} x'_n = a.$$

11.5 Infinite Products 无穷乘积

11.5.1 Definition of Infinite Products 无穷乘积的定义

设 $p_1, p_2, \cdots, p_n, \cdots$ $(p_n \neq 0)$ 是无穷可列个实数，我们称它们的“积”

$$p_1 \cdot p_2 \cdot \cdots \cdot p_n \cdot \cdots$$

为**无穷乘积**，记为 $\prod_{n=1}^{\infty} p_n$，其中 p_n 称为无穷乘积的通项或一般项。

与级数相类似，需要对上述的无穷乘积给出合理的定义。为此构作无穷乘积 $\prod_{n=1}^{\infty} p_n$ 的“部分积数列”$\{P_n\}$：

$$
\begin{aligned}
&P_1 = p_1,\\
&P_2 = p_1 \cdot p_2,\\
&P_3 = p_1 \cdot p_2 \cdot p_3,\\
&\quad\vdots\\
&P_n = p_1 \cdot p_2 \cdot \cdots \cdot p_n = \prod_{k=1}^{n} p_k,\\
&\quad\vdots
\end{aligned}
$$

定义 11.7 如果部分积数列 $\{P_n\}$ 收敛于一个非零的有限数 P，则称无穷乘积 $\prod_{n=1}^{\infty} p_n$ 收敛，且称 P 为它的积，记为

$$\prod_{n=1}^{\infty} p_n = P.$$

如果 $\{P_n\}$ 发散或 $\{P_n\}$ 收敛于 *0*，则称无穷乘积 $\prod_{n=1}^{\infty} p_n$ 发散。

注意，当 $\lim_{n\to\infty} p_n = 0$ 时，我们称无穷乘积 $\prod_{n=1}^{\infty} p_n$ 发散于 0，而不是收敛于 0。在学习了无穷乘积收敛的充分必要条件后将会知道，它使无穷乘积的敛散性与级数的敛散性统一起来。

定理 11.21 如果无穷乘积 $\prod_{n=1}^{\infty} p_n$ 收敛，则

(1) $\lim_{n\to\infty} p_n = 1$;

(2) $\lim_{m\to\infty} \prod_{n=m+1}^{\infty} p_n = 1.$

11.5.2 Infinite Products and Infinite Series 无穷乘积与无穷级数

定理 11.22 无穷乘积 $\prod_{n=1}^{\infty} p_n$ 收敛的充分必要条件是级数 $\sum_{n=1}^{\infty} \ln p_n$ 收敛。

推论 11.2 若 $a_n > 0$（或 $a_n < 0$），则无穷乘积 $\prod_{n=1}^{\infty}(1+a_n)$ 收敛的充分必要条件是级数 $\sum_{n=1}^{\infty} a_n$ 收敛。

推论 11.3 若级数 $\sum_{n=1}^{\infty} a_n$ 收敛，则无穷乘积 $\prod_{n=1}^{\infty}(1+a_n)$ 收敛的充分必要条件是级数 $\sum_{n=1}^{\infty} a_n^2$ 收敛。

定义 11.8 当级数 $\sum_{n=1}^{\infty} \ln p_n$ 绝对收敛时，称无穷乘积 $\prod_{n=1}^{\infty} p_n$ 绝对收敛。

显然，绝对收敛的无穷乘积必定收敛。由于绝对收敛级数具有可交换性，可知绝对收敛的无穷乘积具有可交换性，而收敛但非绝对收敛的无穷乘积不一定具有可交换性。

定理 11.23 设 $a_n > -1$, $n = 1, 2, \cdots$, 则下述三命题等价:

(1) 无穷乘积 $\prod\limits_{n=1}^{\infty}(1 + a_n)$ 绝对收敛;

(2) 无穷乘积 $\prod\limits_{n=1}^{\infty}(1 + |a_n|)$ 收敛;

(3) 级数 $\sum\limits_{n=1}^{\infty}|a_n|$ 收敛。

12 Series of Functions 函数项级数

12.1 Uniform Convergence of Series of Functions 函数项级数的一致收敛性

12.1.1 Pointwise Convergence 点态收敛

现在我们将级数的概念从数推广到函数上去。设 $u_n(x)$ $(n = 1, 2, 3, \cdots)$ 是具有公共定义域 E 的一列函数，我们将这无穷个函数的“和”

$$u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots$$

称为**函数项级数**，记为 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$。函数项级数的收敛性可以借助数项级数来得到。

定义 12.1 设 $u_n(x)$ $(n = 1, 2, 3, \cdots)$ 在 E 上定义。对于任意固定的 $x_0 \in E$，若数项级数 $\displaystyle\sum_{n=1}^{\infty} u_n(x_0)$ 收敛，则称函数项级数 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 在点 x_0 收敛，或称 x_0 是 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 的收敛点。

函数项级数 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 的收敛点全体所构成的集合称为 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 的**收敛域**。

设 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 的收敛域为 $D \subset E$，则 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 就定义了集合 D 上的一个函数

$$S(x) = \sum_{n=1}^{\infty} u_n(x), \quad x \in D.$$

$S(x)$ 称为 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 的**和函数**。由于这是通过逐点定义的方式得到的，因此称 $\displaystyle\sum_{n=1}^{\infty} u_n(x)$ 在 D 上**点态收敛**于 $S(x)$。

给定一个函数项级数 $\sum_{n=1}^{\infty} u_n(x)$，可以作出它的**部分和函数**

$$S_n(x) = \sum_{k=1}^{n} u_k(x), \quad x \in E;$$

显然，使 $\{S_n(x)\}$ 收敛的 x 全体正是集合 D。因此在 D 上，$\sum_{n=1}^{\infty} u_n(x)$ 的和函数 $S(x)$ 就是其部分和函数序列 $\{S_n(x)\}$ 的极限，即有

$$S(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \sum_{k=1}^{n} u_k(x), \quad x \in D.$$

反过来，若给定一个函数序列 $\{S_n(x)\}(x \in E)$，只要令

$$\begin{cases} u_1(x) = S_1(x), \\ u_{n+1}(x) = S_{n+1}(x) - S_n(x) \quad (n = 1, 2, \cdots), \end{cases}$$

也可得到相应的函数项级数 $\sum_{n=1}^{\infty} u_n(x)$，它的部分和函数序列就是 $\{S_n(x)\}$。

所以，函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 与函数序列 $\{S_n(x)\}$ 的收敛性在本质上完全是一回事。为方便起见，今后我们将经常通过讨论函数序列来研究函数项级数的性质。

12.1.2 Fundamental Problems of Series of Functions (or Function Sequences) 函数项级数（或函数序列）的基本问题

通过前面的学习我们已经知道，若有有限个函数 $u_1(x), u_2(x), \cdots, u_n(x)$ 在 D 上定义且具有某种分析性质，如连续性、可导性和 Riemann 可积性（以下就称可积性）等，则它们的和函数

$$u_1(x) + u_2(x) + \cdots + u_n(x)$$

在 D 上仍保持同样的分析性质，且其和函数的极限（或导数、积分）可以通过对每个函数分别求极限（或导数、积分）后再求和来得到，即成立

(a)

$$\lim_{x\to x_0}[u_1(x) + u_2(x) + \cdots + u_n(x)] = \lim_{x\to x_0} u_1(x) + \lim_{x\to x_0} u_2(x) + \cdots + \lim_{x\to x_0} u_n(x);$$

(b)

$$\frac{d}{dx}[u_1(x) + u_2(x) + \cdots + u_n(x)] = \frac{d}{dx}u_1(x) + \frac{d}{dx}u_2(x) + \cdots + \frac{d}{dx}u_n(x);$$

(c)

$$\int_a^b [u_1(x) + u_2(x) + \cdots + u_n(x)]\, dx = \int_a^b u_1(x)\, dx + \int_a^b u_2(x)\, dx + \cdots + \int_a^b u_n(x)\, dx.$$

这些性质给我们带来了很大的方便。

在研究函数项级数时，我们面对的是无限个 $u_n(x)$ $(n = 1, 2, 3, \cdots)$，它们的和函数 $S(x)$ 大多是不知道的，也就是说，我们只能借助 $u_n(x)$ 的分析性质来间接地获得 $S(x)$ 的分析性质。那么很自然地，我们希望在一定条件下，上述运算法则可以推广到无限个函数求和的情况。

这个问题是函数项级数（或函数序列）研究中的基本问题，其实质是极限（或求导、求积分）运算与无限求和运算在什么条件下可以交换次序（由于求导、求积分与无限求和均可看作特殊的极限运算，因此更一般地，可将其统一视为两种极限运算的交换次序）。下面我们将会看到，仅要求 $\sum_{n=1}^{\infty} u_n(x)$ 在 D 上点态收敛是不够的。

(1) 将性质 (a) 推广到无限个函数的情况，是指当 $u_n(x)$ 在 D 上连续时，和函数

$$S(x) = \sum_{n=1}^{\infty} u_n(x)$$

也在 D 上连续，并且成立

$$\lim_{x\to x_0} \sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \lim_{x\to x_0} u_n(x),$$

即极限运算与无限求和运算可以交换次序（也称函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 可以逐项求极限）。

对于函数序列 $\{S_n(x)\}$ 而言，相应的结论是极限函数 $S(x) = \lim_{n\to\infty} S_n(x)$ 也在 D 上连续，并且成立

$$\lim_{x\to x_0} \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \lim_{x\to x_0} S_n(x),$$

即两种极限运算可以交换次序。

下面的例子说明，在点态收敛的情况下，上述性质不一定成立。

例题 12.1 设 $S_n(x) = x^n$，则 $\{S_n(x)\}$ 在区间 $(-1, 1]$ 上收敛，极限函数为

$$S(x) = \lim_{n\to\infty} S_n(x) = \begin{cases} 0, & -1 < x < 1, \\ 1, & x = 1. \end{cases}$$

虽然对一切 $n, S_n(x)$ 在 $(-1, 1]$ 上连续（也是可导的），但极限函数 $S(x)$ 在 $x = 1$ 不连续（当然更谈不上在 $x = 1$ 可导）。

(2) 将性质 (b) 推广到无限个函数的情况，是指当 $u_n(x)$ 在 D 上可导时，和函数

$$S(x)=\sum_{n=1}^{\infty}u_n(x)$$

也在 D 上可导，并且成立

$$\frac{d}{dx}\sum_{n=1}^{\infty}u_n(x)=\sum_{n=1}^{\infty}\frac{d}{dx}u_n(x),$$

即求导运算与无限求和运算可以交换次序（也称函数项级数 $\sum_{n=1}^{\infty}u_n(x)$ 可以逐项求导）。

对于函数序列 $\{S_n(x)\}$ 而言，相应的结论是极限函数 $S(x)=\lim_{n\to\infty}S_n(x)$ 也在 D 上可导，并且成立

$$\frac{d}{dx}\lim_{n\to\infty}S_n(x)=\lim_{n\to\infty}\frac{d}{dx}S_n(x),$$

即求导运算与极限运算可以交换次序。

上一个例子已说明在点态收敛情况下，和函数（或极限函数）可能不可导；下面将看到，即使和函数（或极限函数）可导，上述两等式也不一定成立。

例题 12.2 设 $S_n(x)=\frac{\sin nx}{\sqrt{n}}$，则 $\{S_n(x)\}$ 在 $(-\infty,+\infty)$ 上收敛，极限函数为 $S(x)=0$，从而导函数 $S'(x)=0$。

由于

$$S_n'(x)=\sqrt{n}\cos nx,$$

因此 $S_n(x)$ 的导函数所构成的序列 $\{S_n'(x)\}$ 并不收敛于 $S'(x)$（例如当 $x=0$, $S_n'(0)=\sqrt{n}\to+\infty$）。

(3) 将性质 (c) 推广到无限个函数的情况，是指当 $u_n(x)$ 在闭区间 $[a,b]\subset D$ 上可积时，和函数 $S(x)=\sum_{n=1}^{\infty}u_n(x)$ 也在 $[a,b]$ 上可积，并且成立

$$\int_a^b\sum_{n=1}^{\infty}u_n(x)dx=\sum_{n=1}^{\infty}\int_a^b u_n(x)dx,$$

即求积分运算与无限求和运算可以交换次序（也称函数项级数 $\sum_{n=1}^{\infty}u_n(x)$ 可以逐项求积分）。

对于函数序列 $\{S_n(x)\}$ 而言，相应的结论是极限函数 $S(x)=\lim_{n\to\infty}S_n(x)$ 也在区间 $[a,b]$ 上可积，并且成立

$$\int_a^b\lim_{n\to\infty}S_n(x)dx=\lim_{n\to\infty}\int_a^b S_n(x)dx,$$

即求积分运算与极限运算可以交换次序。

但下面两例将说明，在点态收敛情况下，和函数（或极限函数）可能不可积；即使可积，上述两等式也不一定成立。

例题 12.3 设

$$S_n(x)=\begin{cases}1, & \text{当}x\cdot n!\ \text{为整数},\\ 0, & \text{当}x\ \text{为其他值},\end{cases}\qquad x\in[0,1].$$

显然，对每一个 $n\in\mathbb{N}^+$，$S_n(x)$ 在 $[0,1]$ 上有界，至多只有有限个不连续点，因而是可积的。

但是，当 x 是无理数时，对一切 n，$S_n(x)=0$，因此 $S(x)=\lim_{n\to\infty}S_n(x)=0$; 当 x 是有理数 $\frac{q}{p}$ $(p\in\mathbb{N}^+,\ q\in\mathbb{N},\ q\le p)$ 时，对于 $n\ge p$，$S_n(x)=1$，因此 $S(x)=\lim_{n\to\infty}S_n(x)=1$。所以，$\{S_n(x)\}$ 的极限函数 $S(x)$ 就是熟知的 *Dirichlet* 函数，它在 $[0,1]$ 上是不可积的。

例题 12.4 设 $S_n(x) = nx(1-x^2)^n$，则 $\{S_n(x)\}$ 在区间 $[0,1]$ 上收敛于极限函数

$$S(x) = 0.$$

显然对任意 $n, S_n(x)$ 与 $S(x)$ 都在 $[0,1]$ 上可积，但是

$$\int_0^1 S_n(x)\,dx = \int_0^1 nx(1-x^2)^n\,dx = -\frac{n}{2}\int_0^1 (1-x^2)^n d(1-x^2) = \frac{n}{2(n+1)} \to \frac{1}{2} \quad (n\to\infty),$$

而

$$\int_0^1 S(x)\,dx = 0.$$

这些例子说明，为了解决这类交换运算次序问题，需要引进比“点态收敛”要求更强的新的收敛概念。

12.1.3 Uniform Convergence of Series of Functions (or Function Sequences) 函数项级数（或函数序列）的一致收敛性

“函数序列 $\{S_n(x)\}$ 在集合 D 上（点态）收敛于 $S(x)$” 是指对于任意 $x_0 \in D$, 数列 $\{S_n(x_0)\}$ 收敛于 $S(x_0)$. 用 “$\varepsilon - N$” 语言来表示的话, 就是: 对任意给定的 $\varepsilon > 0$, 可以找到正整数 N, 当 $n > N$ 时, 成立:

$$|S_n(x_0) - S(x_0)| < \varepsilon.$$

一般说来, 这里的 N 应理解为 $N(x_0, \varepsilon)$, 即 N 不仅与 ε 有关, 而且随着 x_0 的变化而变化. 这意味着在 D 的不同处, $S_n(x)$ 的收敛速度可能大相径庭.

我们希望 $\{S_n(x)\}$ 不仅在 D 上点点收敛于 $S(x)$, 而且在 D 上的收敛速度具有某种整体一致性. 通过分析其 “$\varepsilon - N$” 定义可以比较直观地发现, 要做到这一点, 关键在于存在一个仅与 ε 有关, 而与 x_0 无关的 $N = N(\varepsilon)$.

定义 12.2 设 $\{S_n(x)\}(x \in D)$ 是一函数序列, 若对任意给定的 $\varepsilon > 0$, 存在仅与 ε 有关的正整数 $N(\varepsilon)$, 当 $n > N(\varepsilon)$ 时,

$$|S_n(x) - S(x)| < \varepsilon$$

对一切 $x \in D$ 成立, 则称 $\{S_n(x)\}$ 在 D 上一致收敛于 $S(x)$, 记为 $S_n(x) \overset{D}{\rightrightarrows} S(x)$.

若函数项级数 $\sum_{n=1}^{\infty} u_n(x)(x \in D)$ 的部分和函数序列 $\{S_n(x)\}$, 其中 $S_n(x) = \sum_{k=1}^{n} u_k(x)$, 在 D 上一致收敛于 $S(x)$, 则我们称 $\sum_{n=1}^{\infty} u_n(x)$ 在 D 上一致收敛于 $S(x)$.

采用符号表述的话, 就是 “$S_n(x) \overset{D}{\rightrightarrows} S(x)$” $\Leftrightarrow \forall \varepsilon > 0, \exists N, \forall n > N, \forall x \in D$:

$$|S_n(x) - S(x)| < \varepsilon;$$

和 “$\sum_{n=1}^{\infty} u_n(x)$ 在 D 上一致收敛于 $S(x)$” $\Leftrightarrow \forall \varepsilon > 0, \exists N, \forall n > N, \forall x \in D$:

$$\left|\sum_{k=1}^{n} u_k(x) - S(x)\right| = |S_n(x) - S(x)| < \varepsilon.$$

推论 12.1 若函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 D 上一致收敛，则函数序列 $\{u_n(x)\}$ 在 D 上一致收敛于 $u(x) \equiv 0$.

由于函数项级数的一致收敛性本质上就是部分和函数序列的一致收敛性，下面我们仅对函数序列举例讨论。

定义 12.3 若对于任意给定的闭区间 $[a,b]\subset D$, 函数序列 $\{S_n(x)\}$ 在 $[a,b]$ 上一致收敛于 $S(x)$, 则称 $\{S_n(x)\}$ 在 D 上内闭一致收敛于 $S(x)$.

显然, 在 D 上一致收敛的函数序列必在 D 上内闭一致收敛, 但其逆命题不成立。

定理 12.1 设函数序列 $\{S_n(x)\}$ 在集合 D 上点态收敛于 $S(x)$, 定义 $S_n(x)$ 与 $S(x)$ 的“距离”为

$$d(S_n,S)=\sup_{x\in D}|S_n(x)-S(x)|,$$

则 $\{S_n(x)\}$ 在 D 上一致收敛于 $S(x)$ 的充分必要条件是:

$$\lim_{n\to\infty}d(S_n,S)=0.$$

例题 12.5 设 $S_n(x)=\frac{nx}{1+n^2x^2}$, 则 $\{S_n(x)\}$ 在 $(0,+\infty)$ 上收敛于 $S(x)=0$, 由于

$$|S_n(x)-S(x)|=\frac{nx}{1+n^2x^2}\le\frac{1}{2},$$

等号成立当且仅当 $x=\frac{1}{n}$, 可知

$$d(S_n,S)=\frac{1}{2}\neq 0\quad(n\to\infty),$$

因此 $\{S_n(x)\}$ 在 $(0,+\infty)$ 上不是一致收敛的。

例题 12.6 设 $S_n(x)=(1-x)x^n$, 则 $\{S_n(x)\}$ 在 $[0,1]$ 上收敛于 $S(x)=0$。由

$$|S_n(x)-S(x)|=(1-x)x^n,$$

及

$$[(1-x)x^n]'=x^{n-1}[n-(n+1)x],$$

容易知道 $|S_n(x)-S(x)|$ 在 $x=\frac{n}{n+1}$ 取到最大值, 从而

$$d(S_n,S)=\left(1-\frac{n}{n+1}\right)\left(\frac{n}{n+1}\right)^n=\frac{1}{n+1}\left(1+\frac{1}{n}\right)^{-n}\to 0\quad(n\to\infty),$$

这就说明 $\{S_n(x)\}$ 在 $[0,1]$ 上一致收敛于 $S(x)=0$。

定理 12.2 设函数序列 $\{S_n(x)\}$ 在集合 D 上点态收敛于 $S(x)$, 则 $\{S_n(x)\}$ 在 D 上一致收敛于 $S(x)$ 的充分必要条件是: 对任意数列 $\{x_n\}(x_n\in D)$, 成立

$$\lim_{n\to\infty}(S_n(x_n)-S(x_n))=0.$$

12.2 Tests for and Properties of Uniformly Convergent Series 一致收敛级数的判别与性质

12.2.1 Tests for Uniform Convergence 一致收敛的判别

用定义或上节介绍的两个充要条件定理判断函数项级数（或函数序列）的一致收敛性需要先知道它的和函数（或极限函数），这在许多情况下是难以甚至完全不可能做到的，因此有必要寻找无需事先知道和函数（或极限函数）的判断条件。

我们知道，用“$\varepsilon-N$”定义判断一个数列的极限，需要先知道它的极限值，而用 Cauchy 收敛原理则可以避开这一点。将这个结论用于函数项级数，就有

定理 12.3 (函数项级数一致收敛的 Cauchy 收敛原理) 函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 D 上一致收敛的充分必要条件是：对于任意给定的 $\varepsilon > 0$，存在正整数 $N = N(\varepsilon)$，使

$$|u_{n+1}(x) + u_{n+2}(x) + \cdots + u_m(x)| < \varepsilon$$

对一切正整数 $m > n > N$ 与一切 $x \in D$ 成立。

可以相应写出函数序列一致收敛的 Cauchy 收敛原理：函数序列 $S_n(x)$ 在 D 上一致收敛 $\Leftrightarrow$ $\forall \varepsilon > 0, \exists N, \forall m > n > N, \forall x \in D:$

$$|S_m(x) - S_n(x)| < \varepsilon.$$

定理 12.4 (Weierstrass 判别法) 设函数项级数 $\sum_{n=1}^{\infty} u_n(x) (x \in D)$ 的每一项 $u_n(x)$ 满足

$$|u_n(x)| \leq a_n, \quad x \in D,$$

并且数项级数 $\sum_{n=1}^{\infty} a_n$ 收敛，则 $\sum_{n=1}^{\infty} u_n(x)$ 在 D 上一致收敛。

定理 12.5 设函数项级数 $\sum_{n=1}^{\infty} a_n(x) b_n(x) (x \in D)$ 满足如下两个条件之一，则 $\sum_{n=1}^{\infty} a_n(x) b_n(x)$ 在 D 上一致收敛。

(1) (Abel 判别法) 函数序列 $\{a_n(x)\}$ 对每一固定的 $x \in D$ 关于 n 是单调的，且 $\{a_n(x)\}$ 在 D 上一致有界：

$$|a_n(x)| \leq M, \quad x \in D, \quad n \in \mathbb{N}^+;$$

同时，函数项级数 $\sum_{n=1}^{\infty} b_n(x)$ 在 D 上一致收敛。

(2) (Dirichlet 判别法) 函数序列 $\{a_n(x)\}$ 对每一固定的 $x \in D$ 关于 n 是单调的，且 $\{a_n(x)\}$ 在 D 上一致收敛于 *0*；同时，函数项级数 $\sum_{n=1}^{\infty} b_n(x)$ 的部分和序列在 D 上一致有界：

$$\left| \sum_{k=1}^{n} b_k(x) \right| \leq M, \quad x \in D, \quad n \in \mathbb{N}^+.$$

12.2.2 Properties of Uniformly Convergent Series 一致收敛级数的性质

现在我们可以来回答上一节中提出的关于函数项级数 (或函数序列) 的基本问题，即在什么条件下，和函数 (或极限函数) 仍然保持连续性、可导性、可积性等分析性质。

定理 12.6 (连续性定理) 设函数序列 $\{S_n(x)\}$ 的每一项 $S_n(x)$ 在 $[a, b]$ 上连续，且在 $[a, b]$ 上一致收敛于 $S(x)$，则 $S(x)$ 在 $[a, b]$ 上也连续。

定理 12.7 设函数序列 $\{S_n(x)\}$ 的每一项 $S_n(x)$ 在 $[a, b]$ 上连续，且在 $[a, b]$ 上一致收敛于 $S(x)$，则 $S(x)$ 在 $[a, b]$ 上可积，且

$$\int_a^b S(x) dx = \lim_{n \to \infty} \int_a^b S_n(x) dx.$$

定理 12.8 设函数序列 $\{S_n(x)\}$ 满足 *(1)* $S_n(x)$ *(*$n = 1, 2, \cdots$*)* 在 $[a, b]$ 上有连续的导函数；*(2)* $\{S_n(x)\}$ 在 $[a, b]$ 上点态收敛于 $S(x)$；*(3)* $\{S'_n(x)\}$ 在 $[a, b]$ 上一致收敛于 $\sigma(x)$，则 $S(x)$ 在 $[a, b]$ 上可导，且

$$\frac{d}{dx} S(x) = \sigma(x).$$

定理 12.9 (Dini 定理) 设函数序列 $\{S_n(x)\}$ 在闭区间 $[a, b]$ 上点态收敛于 $S(x)$，如果

(1) $S_n(x) (n = 1, 2, \cdots)$ 在 $[a, b]$ 上连续；

(2) $S(x)$ 在 $[a, b]$ 上连续；

(3) $\{S_n(x)\}$ 关于 n 单调，即对任意固定的 $x \in [a, b]$，$\{S_n(x)\}$ 是单调数列，

则 $\{S_n(x)\}$ 在 $[a, b]$ 上一致收敛于 $S(x)$。

12.2.3 Weierstrass Function 处处不可导的连续函数之例

在数学分析的发展历史上，数学家们一直猜测：连续函数的不可导点至多是可列集。随着级数理论的发展，Weierstrass 利用函数项级数第一个构造出了一个处处连续而处处不可导的函数，为上述猜测做了一个否定的终结。

12.3 Power Series 幂级数

下面我们讨论一类特殊的函数项级数

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + \cdots,$$

这样的函数项级数称为**幂级数**。

显然，幂级数可以看成是一个“无限次多项式”，而它的部分和函数 $S_n(x)$ 是一个 $n-1$ 次多项式。为了方便，我们通常取 $x_0=0$，也就是说讨论

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots,$$

只要对所得的结果做一个平移 $x = t - x_0$，就可以平行推广到 $x_0 \neq 0$ 的情况。

12.3.1 Radius of Convergence of Power Series 幂级数的收敛半径

对于幂级数 $\sum_{n=0}^{\infty} a_n x^n$，我们首先有

$$\varlimsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = \varlimsup_{n\to\infty} \sqrt[n]{|a_n|} \cdot |x|$$

根据数项级数的 Cauchy 判别法，当上面的极限值小于 1 时，$\sum_{n=0}^{\infty} a_n x^n$ 绝对收敛；当上面的极限值大于 1 时，$\sum_{n=0}^{\infty} a_n x^n$ 发散。如果令

$$A = \varlimsup_{n\to\infty} \sqrt[n]{|a_n|},$$

定义

$$R = \begin{cases} +\infty, & \text{当} A = 0, \\ \frac{1}{A}, & \text{当} A \in (0,+\infty), \\ 0, & \text{当} A = +\infty, \end{cases}$$

则我们有

定理 12.10 (Cauchy-Hadamard 定理) 幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 当 $|x| < R(R>0)$ 时绝对收敛；当 $|x| > R$ 时发散。

注意在区间的端点 $x = \pm R$，幂级数收敛与否必须另行判断。

对于 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$，则有平行的结论：幂级数在以 x_0 为中心，以 R 为半径的对称区间内绝对收敛，而在该区间外发散。在区间的端点 $x_0 \pm R$，幂级数的敛散性必须另行判断。

数 R 称为幂级数的**收敛半径**。当 $R=+\infty$ 时，幂级数对一切 x 都是绝对收敛的；当 $R=0$ 时，幂级数仅当 $x = x_0$ 时收敛。

定理 12.11 (d’ Alembert 判别法) 如果对幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 成立

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = A,$$

则此幂级数的收敛半径为 $R = \frac{1}{A}$。

12.3.2 Properties of Power Series 幂级数的性质

Abel 曾系统地研究过幂级数，并建立了 Abel 第一定理与第二定理，其中第一定理是这样的：设 $x_0 = 0$，如果幂级数在点 ξ 收敛，则当 $|x| < |\xi|$ 时幂级数绝对收敛；如果幂级数在点 η 发散，则当 $|x| > |\eta|$ 时幂级数发散。显然，这一结论已包含在 Cauchy-Hadamard 定理之中。下面我们叙述第二定理：

定理 12.12 (Abel 第二定理) *设幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径为 R，则*

(i) $\sum_{n=0}^{\infty} a_n x^n$ 在 $(-R, R)$ 上内闭一致收敛，即在任意闭区间 $[a, b] \subset (-R, R)$ 上一致收敛；

(ii) 若 $\sum_{n=0}^{\infty} a_n x^n$ 在 $x = R$ 收敛，则它在任意闭区间 $[a, R] \subset (-R, R]$ 上一致收敛。

12.4 Power Series Expansion of Functions 函数的幂级数展开

12.4.1 Taylor Series and Remainder Formula 泰勒级数与余项公式

上一节中已展示了幂级数的良好性质。显而易见，如果一个函数在某一区间上能够表示成一个幂级数，将给理论上讨论其性质带来极大的方便，同时也具有重要的应用价值。下面我们就来讨论函数可以表示成幂级数的条件，以及在这些条件满足时如何将函数表示成幂级数。

假设函数 $f(x)$ 在 x_0 的某个邻域 $O(x_0, r)$ 上可表示成幂级数

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad x \in O(x_0, r),$$

也就是说，$\sum_{n=0}^{\infty} a_n (x - x_0)^n$ 在 $O(x_0, r)$ 上的和函数为 $f(x)$。根据幂级数的逐项可导性，$f(x)$ 必定在 $O(x_0, r)$ 上任意阶可导，且对一切 $k \in \mathbb{N}^+$，

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n (x - x_0)^{n-k}.$$

令 $x = x_0$，得到

$$a_k = \frac{f^{(k)}(x_0)}{k!}, \quad k = 0, 1, 2, \cdots,$$

也就是说，系数 $\{a_n\}$ 由和函数 $f(x)$ 唯一确定，我们称它们为 $f(x)$ 在 x_0 的 **Taylor 系数**。

反过来，设函数 $f(x)$ 在 x_0 的某个邻域 $O(x_0, r)$ 上任意阶可导，则我们能求出它在 x_0 的 Taylor 系数 $a_n = \frac{f^{(n)}(x_0)}{n!} (n = 0, 1, 2, \cdots)$，并作出幂级数

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n,$$

这一幂级数称为 $f(x)$ 在 x_0 的 **Taylor 级数**。

现在我们要问：是否存在常数 $\rho (0 < \rho \leq r)$，使得

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

在 $O(x_0, \rho)$ 上收敛于 $f(x)$？答案并不是肯定的。

一个任意阶可导的函数的 Taylor 级数并非一定能收敛于函数本身。

为了寻求一个函数的 Taylor 级数收敛于它本身的条件，我们回忆在 §5.3 中所得到的 Taylor 公式：设 $f(x)$ 在 $O(x_0, r)$ 有 $n + 1$ 阶导数，则

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + r_n(x),$$

其中 $r_n(x)$ 是 n 阶 Taylor 公式的余项。现在我们假定讨论的函数 $f(x)$ 在 $O(x_0,r)$ 上任意阶可导，也就是说，上面的 Taylor 公式对一切正整数 n 成立，于是我们可以断言：

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

在 $O(x_0,\rho)$ $(0<\rho\le r)$ 成立的充分必要条件是：

$$\lim_{n\to\infty}r_n(x)=0$$

对一切 $x\in O(x_0,\rho)$ 成立。

这时，我们才称 $f(x)$ 在 $O(x_0,\rho)$ 可以展开成幂级数（或 Taylor 级数），或者称

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

是 $f(x)$ 在 $O(x_0,\rho)$ 上的**幂级数展开**（或 **Taylor 展开**）。

在前面的章节中，曾导出余项

$$r_n(x)=\frac{f^{(n+1)}(x_0+\theta(x-x_0))}{(n+1)!}(x-x_0)^{n+1},\quad 0<\theta<1,$$

$r_n(x)$ 的这一形式称为 Lagrange 余项。为了讨论各种函数的 Taylor 展开，我们还需要 $r_n(x)$ 的另一形式，即积分形式：

定理 12.13 设 $f(x)$ 在 $O(x_0,r)$ 上任意阶可导，则

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+r_n(x),\quad x\in O(x_0,r),$$

其中

$$r_n(x)=\frac{1}{n!}\int_{x_0}^{x}f^{(n+1)}(t)(x-t)^n dt.$$

对余项 $r_n(x)$ 的积分形式应用积分第一中值定理，考虑到当 $t\in[x_0,x]$（或 $[x,x_0]$）时，$(x-t)^n$ 保持定号，于是就有

$$r_n(x)=\frac{f^{(n+1)}(\xi)}{n!}\int_{x_0}^{x}(x-t)^n dt\quad(\xi\text{在}x_0\text{与}x\text{之间})$$

$$=\frac{f^{(n+1)}(x_0+\theta(x-x_0))}{(n+1)!}(x-x_0)^{n+1},\quad 0\le\theta\le 1,$$

这就是我们已经知道的 Lagrange 余项；如果将 $f^{(n+1)}(t)(x-t)^n$ 看作一个函数，应用积分第一中值定理，则有

$$r_n(x)=\frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!}\int_{x_0}^{x}dt\quad(\xi\text{在}x_0\text{与}x\text{之间})$$

$$=\frac{f^{(n+1)}(x_0+\theta(x-x_0))}{n!}(1-\theta)^n(x-x_0)^{n+1},\quad 0\le\theta\le 1,$$

$r_n(x)$ 的这一形式称为 **Cauchy 余项**。

12.4.2 Taylor Expansions of Elementary Functions 初等函数的泰勒展开

我们先通过讨论使余项 $r_n(x)$ 趋于 0 的 x 的范围，导出基本初等函数的幂级数展开式，然后介绍将一般函数展开成幂级数的一些方法。

(1) $f(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots, \quad x \in (-\infty, +\infty)$.

(2) $f(x) = \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$

$$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots, \quad x \in (-\infty, +\infty).$$

(3) $f(x) = \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$

$$= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots, \quad x \in (-\infty, +\infty).$$

(4) $f(x) = \arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$

$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots, \quad x \in [-1, 1].$$

(5) $f(x) = \ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$

$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n-1} \frac{x^n}{n} + \cdots, \quad x \in (-1, 1].$$

下面我们通过例题介绍幂级数展开的一般方法。

例题 12.7 求 $f(x) = \frac{1}{x^2}$ 在 $x = 1$ 的幂级数展开。

解： 当 $|x-1| < 1$ 时，

$$\frac{1}{x} = \frac{1}{1+(x-1)} = \sum_{n=0}^{\infty} (-1)^n (x-1)^n,$$

对等式两边求导，应用幂级数的逐项可导性，

$$-\frac{1}{x^2} = \sum_{n=1}^{\infty} (-1)^n n (x-1)^{n-1},$$

于是得到

$$\frac{1}{x^2} = \sum_{n=0}^{\infty} (-1)^n (n+1)(x-1)^n, \quad x \in (0, 2).$$

例题 12.8 求 $f(x) = \frac{1}{3+5x-2x^2}$ 在 $x = 0$ 的幂级数展开。

解：

$$\begin{aligned} f(x) = \frac{1}{3+5x-2x^2} &= \frac{1}{(3-x)(1+2x)} = \frac{1}{7}\left(\frac{1}{3-x} + \frac{2}{1+2x}\right) \\ &= \frac{1}{7}\left(\frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{x}{3}\right)^n + 2\sum_{n=0}^{\infty}(-2x)^n\right) \\ &= \frac{1}{7}\sum_{n=0}^{\infty}\left[\frac{1}{3^{n+1}} - (-2)^{n+1}\right] x^n. \end{aligned}$$

由于 $\frac{1}{3-x}$ 的幂级数展开的收敛范围是 $(-3,3)$，$\frac{2}{1+2x}$ 的幂级数展开的收敛范围是 $\left(-\frac{1}{2}, \frac{1}{2}\right)$，因此 $f(x)$ 的幂级数展开在 $\left(-\frac{1}{2}, \frac{1}{2}\right)$ 成立。

设 $f(x)$ 的幂级数展开为 $\sum_{n=0}^{\infty} a_n x^n$，收敛半径为 $R_1, g(x)$ 的幂级数展开为 $\sum_{n=0}^{\infty} b_n x^n$，收敛半径为 R_2，则 $f(x)g(x)$ 的幂级数展开就是它们的 Cauchy 乘积：

$$f(x)g(x) = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} c_n x^n,$$

其中 $c_n = \sum_{k=0}^{n} a_k b_{n-k}, |x| < \min\{R_1, R_2\}$。

当 $b_0 \neq 0$ 时，我们可以通过待定系数法求 $\frac{f(x)}{g(x)}$ 的幂级数展开：设

$$\frac{f(x)}{g(x)} = \sum_{n=0}^{\infty} c_n x^n,$$

则

$$\left(\sum_{n=0}^{\infty} b_n x^n\right)\left(\sum_{n=0}^{\infty} c_n x^n\right) = \sum_{n=0}^{\infty} a_n x^n,$$

分离 x 的各次幂的系数，可依次得到

$$b_0c_0 = a_0 \quad \Rightarrow \quad c_0 = \frac{a_0}{b_0},$$

$$b_0c_1 + b_1c_0 = a_1 \quad \Rightarrow \quad c_1 = \frac{a_1 - b_1c_0}{b_0},$$

$$b_0c_2 + b_1c_1 + b_2c_0 = a_2 \quad \Rightarrow \quad c_2 = \frac{a_2 - b_1c_1 - b_2c_0}{b_0},$$

……

一直继续下去，可求得所有的 c_n。

例题 12.9 求 $e^x \sin x$ 在 $x = 0$ 的幂级数展开（到 x^5）。

解：

$$e^x \sin x = \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right)\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) = x + x^2 + \frac{1}{3}x^3 - \frac{1}{30}x^5 + \cdots,$$

上述幂级数展开对一切 $x \in (-\infty, +\infty)$ 都成立。

例题 12.10 求 $\tan x$ 在 $x = 0$ 的幂级数展开（到 x^5）。

解：由于 $\tan x$ 是奇函数，我们可以令

$$\tan x = \frac{\sin x}{\cos x} = c_1 x + c_3 x^3 + c_5 x^5 + \cdots.$$

于是

$$(c_1 x + c_3 x^3 + c_5 x^5 + \cdots)\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$

比较等式两端 x, x^3 与 x^5 的系数，就可得到

$$c_1 = 1, \quad c_3 = \frac{1}{3}, \quad c_5 = \frac{2}{15},$$

因此

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots.$$

13 Limits and Continuity in Euclidean Space 欧氏空间上的极限和连续

13.1 Fundamental Theorems in Euclidean Space 欧氏空间上的基本定理

到目前为止，我们在数学分析课程中学习的都只是一元函数的分析性质。但在现实生活中，除了非常简单的情况之外，可以仅用一个自变量和一个因变量的变化关系来刻画的问题可以说是比较少的。比如，即便是像物理学中所研究质点运动这么一个相对较为容易的问题，也需要用到三个空间变量 x, y, z 和一个时间变量 t 以及多个函数值（如位置、速度、加速度、动量等），更不用说在化学、生物及社会科学领域产生的远为复杂的情况。这种多自变量和多因变量的变化关系，反映到数学上就是多元函数（或多元函数组，即向量值函数）。

从本节开始我们将转向研究多元函数（组）。多元函数的分析性质无非也是极限理论、连续性、可微性、可积性等，它们与一元函数的相应性质既有紧密联系，又有很大差别。

13.1.1 Distance and Limits in Euclidean Space 欧氏空间上的距离与极限

前面说过，极限理论是整个数学分析的基础。在导出多元函数的极限定义之前，我们先来回忆一下一元函数的情况。

极限定义

$$\lim_{x\to x_0} f(x) = A \Leftrightarrow \forall\varepsilon > 0, \exists\delta > 0, \forall x(0 < |x - x_0| < \delta) : |f(x) - A| < \varepsilon$$

意味着，在自变量的变化过程中，只要 x 与 x_0 充分接近（$x \neq x_0$），函数值 $f(x)$ 就可以与 A 任意接近。而这个“接近”，不管是用符号“$0 < |x - x_0| < \delta$”和“$|f(x) - A| < \varepsilon$”表示，还是用语言“在 x_0 的 δ 去心邻域 $O(x_0, \delta)\backslash\{x_0\}$ 中”和“落在点 A 的 ε 邻域中”表示，实质上都是用绝对值，即一维空间中两点间的距离刻画的。显而易见，若没有距离的概念和定义，就无所谓“接近”或“不接近”，也就没有“收敛”和“发散”。收敛就是距离趋向于零。

对于多元函数（组），上述的 x, x_0（及 $f(x), A$）都是由多个分量组成的，为了研究多元函数的性质，我们先要将“距离”的概念推广至高维空间，定义出类似于“绝对值”那样的度量标准，然后才能在此基础上去相应地定义极限，进而构筑整个多元分析理论。

记 $\mathbb{R}$ 为实数全体，定义 n 个 $\mathbb{R}$ 的 **笛卡尔乘积集**为

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, x_2, \cdots, x_n) \mid x_i \in \mathbb{R}, i = 1, 2, \cdots, n\}.$$

$\mathbb{R}^n$ 中的元素 $x = (x_1, x_2, \cdots, x_n)$ 称为**向量**或**点**，x_i 称为 x 的**第 i 个坐标**。特别地，$\mathbb{R}^n$ 中的零元素记为 $0 = (0, 0, \cdots, 0)$。

设 $x = (x_1, x_2, \cdots, x_n), y = (y_1, y_2, \cdots, y_n)$ 为 $\mathbb{R}^n$ 中任意两个向量，λ 为任意实数，定义 $\mathbb{R}^n$ 中的加法和数乘运算：

$$x + y = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n),$$

$$\lambda x = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n),$$

$\mathbb{R}^n$ 就成为**向量空间**。

如果再在 $\mathbb{R}^n$ 上引入**内积**运算

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{k=1}^{n} x_ky_k,$$

那么它就被称为 **欧式空间**。

容易验证内积满足以下性质：设 $x, y, z \in R^n, \lambda, \mu \in R$，则

(1)（**正定性**）$\langle x, x\rangle \geq 0$，而 $\langle x, x\rangle = 0$ 当且仅当 $x = 0$；

(2)（**对称性**）$\langle x, y\rangle = \langle y, x\rangle$；

(3)（**线性性**）$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$；

(4)（**Schwarz 不等式**）$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$。

定义 13.1 *Euclid* 空间 $\mathbb{R}^n$ 中任意两点 $x = (x_1, x_2, \cdots, x_n)$ 和 $y = (y_1, y_2, \cdots, y_n)$ 的距离定义为

$$|x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2};$$

并称

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{k=1}^{n} x_k^2}$$

为 x 的 *Euclid* 范数（简称范数）。显然，x 的范数 $\|x\|$ 就是 x 到 *0* 的距离（即 x 的模长）。

定理 13.1 距离满足以下性质：

(1) （**正定性**）$|x - y| \geq 0$，而 $|x - y| = 0$ 当且仅当 $x = y$；

(2) （**对称性**）$|x - y| = |y - x|$；

(3)（**三角不等式**）$|x - z| \leq |x - y| + |y - z|$。

定义了距离就可以引入邻域以及收敛的概念。

定义 13.2 设 $a = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n, \delta > 0$, 则点集

$$O(a, \delta) = \{x \in \mathbb{R}^n \mid |x - a| < \delta\}$$

$$= \left\{x \in \mathbb{R}^n \mid \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2} < \delta\right\}$$

称为点 a 的 δ 邻域，a 称为这个邻域的中心，δ 称为邻域的半径。

特别地，$O(a, \delta)$ 在 $\mathbb{R}$ 上就是开区间，在 $\mathbb{R}^2$ 上是开圆盘，在 $\mathbb{R}^3$ 上则是开球。

定义 13.3 设 $\{x_k\}$ 是 $\mathbb{R}^n$ 中的一个点列。若存在定点 $a \in \mathbb{R}^n$，对于任意给定的 $\epsilon > 0$，存在正整数 K，使得当 $k > K$ 时，

$$|x_k - a| < \epsilon \quad (\text{即} x_k \in O(a, \epsilon)),$$

则称点列 $\{x_k\}$ 收敛于 a，记为 $\lim_{k\to\infty} x_k = a$。而称 a 为点列 $\{x_k\}$ 的极限。

一个点列不收敛就称其发散。

记 $x_k = (x_1^k, x_2^k, \cdots, x_n^k) \quad (k = 1, 2, \cdots), a = (a_1, a_2, \cdots, a_n)$，利用不等式

$$|x_j^k - a_j| \leq |x_k - a| = \sqrt{\sum_{i=1}^{n}(x_i^k - a_i)^2} \leq \sum_{i=1}^{n}|x_i^k - a_i|, \quad j = 1, 2, \cdots, n$$

可以得到

定理 13.2 $\lim_{k\to\infty} x_k = a$ 的充分必要条件是 $\lim_{k\to\infty} x_i^k = a_i$ $(i = 1, 2, \cdots, n)$。

根据此定理，我们可以利用对点列分量的讨论，将一维的一些结论平行地推广到高维去。

前面已给出了 $\mathbb{R}$ 中收敛数列的一些性质，如惟一性、有界性、保序性、夹逼性及四则运算法则。由于高维的两个点之间不存在大小关系，因此保序性和夹逼性这两个与比较大小有关的性质已不再有意义了。

有界性牵涉到的是点的模长，我们对高维点集的有界性作如下定义：

定义 13.4 设 S 是 $\mathbb{R}^n$ 上的点集。若存在正数 M，使得对于任意 $x \in S$，

$$\|x\| \leq M,$$

（或等价地，存在正数 M' 使得 $S \subset O(0, M')$）则称 S 为有界集。

可以证明惟一性（收敛点列 $\{x_k\}$ 的极限是惟一的）、有界性（收敛点列 $\{x_k\}$ 必定有界）和极限的线性运算法则在高维情况下依然成立。

13.1.2 Open Sets and Closed Sets 开集与闭集

在一维的情况，闭区间上的许多重要结果，如闭区间套定理、连续函数的若干性质，在开区间是不成立的。因此有理由相应地对高维空间的点集作类似划分。

设 S 是 $\mathbb{R}^n$ 上的点集，它在 $\mathbb{R}^n$ 上的补集 $\mathbb{R}^n \setminus S$ 记为 S^c。对于任意 $x \in \mathbb{R}^n$，从其邻域与 S 的关系来分，无非是下列三种情况之一：

（1）存在 x 的一个 δ 邻域 $O(x, \delta)$ 完全落在 S 中（注意：这时 x 必属于 S），这时称 x 是 S 的**内点**。S 的内点全体称为 S 的**内部**，记为 S°。

（2）存在 x 的一个 δ 邻域 $O(x, \delta)$ 完全不落在 S 中，这时称 x 是 S 的**外点**。

（3）不存在 x 的具有上述性质的 δ 邻域，即 x 的任意 δ 邻域既包含 S 中的点，又包含不属于 S 的点，那么就称 x 是 S 的**边界点**。S 的边界点的全体称为 S 的**边界**，记为 ∂S。

要注意的是，内点必属于 S，外点必不属于 S（或者说必属于 S^c），但边界点可能属于 S，也可能不属于 S。

进一步，若存在 x 的一个邻域，其中只有 x 点属于 S，则称 x 是 S 的**孤立点**。显然，孤立点必是边界点。

若 x 的任意邻域都含有 S 中的无限个点，则称 x 是 S 的**聚点**。S 的聚点的全体记为 S'。显然，S 的内点必是 S 的聚点；S 的边界点，只要不是 S 的孤立点，也必是 S 的聚点。因此 S 的聚点可能属于 S，也可能不属于 S。例如在 $\mathbb{R}$ 中，0 是点集

$$\left\{ \frac{1}{n} \,\middle|\, n = 1, 2, \cdots \right\}$$

的聚点，但它不属于这个点集。

定理 13.3 x 是点集 $S(\subset \mathbb{R}^n)$ 的聚点的充分必要条件是：存在点列 $\{x_k\}$ 满足 $x_k \in S, x_k \neq x (k = 1, 2, \cdots)$，使得 $\lim_{k\to\infty} x_k = x$。

现在我们可以引出“开”和“闭”的概念了。

定义 13.5 设 S 是 $\mathbb{R}^n$ 上的点集。若 S 中的每一个点都是它的内点，则称 S 为开集；若 S 中包含了它的所有的聚点，则称 S 为闭集。

S 与它的聚点全体 S' 的并集称为 S 的闭包，记为 $\overline{S}$。

定理 13.4 $\mathbb{R}^n$ 上的点集 S 为闭集的充分必要条件是 S^c 是开集。

引理 13.1 (De Morgan 公式) 设 S_α 是 $\mathbb{R}^n$ 中的一组（有限或无限多个）子集，则

(1)

$$\left(\bigcup_\alpha S_\alpha\right)^c = \bigcap_\alpha S_\alpha^c;$$

(2)

$$\left(\bigcap_\alpha S_\alpha\right)^c = \bigcup_\alpha S_\alpha^c.$$

定理 13.5

(1) 任意一组开集 S_α 的并集 $\bigcup_\alpha S_\alpha$ 是开集；

(2) 任意一组闭集 T_α 的交集 $\bigcap_\alpha T_\alpha$ 是闭集；

(3) 任意有限个开集 $S_1, S_2, \cdots, S_k$ 的交集 $\bigcap_{i=1}^k S_i$ 是开集；

(4) 任意有限个闭集 $T_1, T_2, \cdots, T_k$ 的并集 $\bigcup_{i=1}^k T_i$ 是闭集。

13.1.3 Fundamental Theorems in Euclidean Space 欧氏空间上的基本定理

定理 13.6 (闭矩形套定理) 设 $\Delta_k = [a_k, b_k] \times [c_k, d_k]$ *($k = 1, 2, \cdots$)* 是 $\mathbb{R}^2$ 上一列闭矩形。如果 *(1)* $\Delta_{k+1} \subset \Delta_k$，即 $a_k \le a_{k+1} < b_{k+1} \le b_k$，$c_k \le c_{k+1} < d_{k+1} \le d_k$，$k = 1, 2, \cdots$; *(2)* $\sqrt{(b_k - a_k)^2 + (d_k - c_k)^2} \to 0 \quad (k \to \infty)$，

则存在惟一的点 $a = (\xi, \eta)$ 属于 $\bigcap_{k=1}^\infty \Delta_k$，且

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = \xi, \quad \lim_{k\to\infty} c_k = \lim_{k\to\infty} d_k = \eta.$$

定理 13.7 (Cantor 闭区域套定理) 设 S_k 是 $\mathbb{R}^n$ 上的非空闭集序列，满足

$$S_1 \supseteq S_2 \supseteq \cdots \supseteq S_k \supseteq S_{k+1} \supseteq \cdots,$$

以及 $\lim_{k\to\infty} \operatorname{diam} S_k = 0$，则存在惟一点属于 $\bigcap_{k=1}^\infty S_k$。

这里

$$\operatorname{diam} S = \sup_{x,y\in S} |x - y|,$$

它称为 S 的直径。

定理 13.8 (Bolzano–Weierstrass 定理) $\mathbb{R}^n$ 上的有界点列 $\{x_k\}$ 中必有收敛子列。

以二维的情况为例，只要先对 $\{x_k\} = \{(x_k, y_k)\}$ 的第一个分量 $\{x_k\}$ 用一维的 Bolzano–Weierstrass 定理，找到其收敛子列 $\{x_{n_k}\}$；再对数列 $\{y_{n_k}\}$ 用一维的 Bolzano–Weierstrass 定理，找到其收敛子列 $\{y_{n_{k_m}}\}$，则 $\{(x_{n_{k_m}}, y_{n_{k_m}})\}$ 就是 $\{x_k\}$ 的收敛子列。

从这个定理立即得到：

推论 13.1 $\mathbb{R}^n$ 上的有界无限点集至少有一个聚点。

定义 13.6 若 $\mathbb{R}^n$ 上的点列 $\{x_k\}$ 满足：对于任意给定的 $\epsilon > 0$，存在正整数 K，使得对任意 $k, l > K$，成立

$$|x_l - x_k| < \epsilon,$$

则称 $\{x_k\}$ 为基本点列（或 *Cauchy* 点列）。

定理 13.9 (Cauchy 收敛原理) $\mathbb{R}^n$ 上的点列 $\{x_k\}$ 收敛的充分必要条件是：$\{x_k\}$ 为基本点列。

13.1.4 Compact Sets 紧集

定义 13.7 设 S 为 $\mathbb{R}^n$ 上的点集。如果 $\mathbb{R}^n$ 中的一组开集 $\{U_\alpha\}$ 满足 $\bigcup_\alpha U_\alpha \supseteq S$，那么称 $\{U_\alpha\}$ 为 S 的一个开覆盖。

如果 S 的任意一个开覆盖 $\{U_\alpha\}$ 中总存在一个有限子覆盖，即存在 $\{U_\alpha\}$ 中的有限个开集 $\{U_{\alpha_i}\}_{i=1}^{p}$，满足 $\bigcup_{i=1}^{p} U_{\alpha_i} \supseteq S$，则称 S 为紧集。

定理 13.10 (Heine-Borel 定理) *$\mathbb{R}^n$ 上的点集 S 是紧集的充分必要条件为：它是有界闭集。*

定理 13.11 *设 S 是 $\mathbb{R}^n$ 上的点集，那么以下三个命题等价：*
(1) S 是有界闭集；
(2) S 是紧集；
(3) S 的任一无限子集在 S 中必有聚点。

Cantor 闭区域套定理，Bolzano-Weierstrass 定理，Cauchy 收敛原理和 Heine-Borel 定理称为 Euclid 空间上的基本定理，它们是相互等价的。

13.2 Multivariable Continuous Functions 多元连续函数

13.2.1 Functions of Several Variables 多元函数

定义 13.8 设 D 是 $\mathbb{R}^n$ 上的点集，D 到 $\mathbb{R}$ 的映射

$$f : D \to \mathbb{R}, \quad x \mapsto z$$

称为 n 元函数，记为 $z = f(x)$。这时，D 称为 f 的定义域，$f(D) = \{z \in \mathbb{R} \mid z = f(x), x \in D\}$ 称为 f 的值域，$\Gamma = \{(x, z) \in \mathbb{R}^{n+1} \mid z = f(x), x \in D\}$ 称为 f 的图像。

13.2.2 Limits of Multivariable Functions 多元函数的极限

定义 13.9 设 D 是 $\mathbb{R}^n$ 上的开集，$x_0 = (x_1^0, x_2^0, \cdots, x_n^0) \in D$ 为一定点，$z = f(x)$ 是定义在 $D \setminus \{x_0\}$ 上的 n 元函数，A 是一个实数。如果对于任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $x \in O(x_0, \delta) \setminus \{x_0\}$ 时，成立

$$|f(x) - A| < \varepsilon,$$

则称当 x 趋于 x_0 时 f 收敛，并称 A 为 f 当 x 趋于 x_0 时的 **（n 重）极限**，记为

$$\lim_{x \to x_0} f(x) = A \quad (\text{或} f(x) \to A \ (x \to x_0)),$$

或

$$\lim_{(x_1, x_2, \cdots, x_n) \to (x_1^0, x_2^0, \cdots, x_n^0)} f(x_1, x_2, \cdots, x_n) = A.$$

对一元函数而言，只要在 x_0 的左、右极限存在且相等，那么函数在 x_0 处的极限就存在。而多元函数就没有这样简单。根据极限存在的定义，要求当 x 以任何方式趋于 x_0 时，函数值都趋于同一个极限。这就为我们判断函数极限的不存在提供了方便，因为若自变量沿不同的两条曲线趋于某一定点时，函数的极限不同或不存在，那么这个函数在该点的极限一定不存在。

13.2.3 Iterated Limits 累次极限

对重极限 $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$，人们很自然会想到的是，能否在一定条件下将重极限 $(x,y)\to(x_0,y_0)$ 分解成为两个独立的极限 $x\to x_0$ 和 $y\to y_0$，再利用一元函数的极限理论和方法逐个处理之？

这后一种极限称为**累次极限**。

定义 13.10 设 D 是 $\mathbb{R}^2$ 上的开集，$(x_0,y_0)\in D$ 为一定点，$z=f(x,y)$ 为定义在 $D\setminus\{(x_0,y_0)\}$ 上的二元函数。如果对于每个固定的 $y\neq y_0$，极限 $\lim\limits_{x\to x_0} f(x,y)$ 存在，并且极限 $\lim\limits_{y\to y_0}\lim\limits_{x\to x_0} f(x,y)$ 存在，那么称此极限值为函数 $f(x,y)$ 在点 (x_0,y_0) 的**先对 x 后对 y 的二次极限**。

同理可定义**先对 y 后对 x 的二次极限** $\lim\limits_{x\to x_0}\lim\limits_{y\to y_0} f(x,y)$。

累次极限存在与重极限存在的关系很复杂。二次极限存在不能保证二重极限存在，二重极限存在同样不能保证二次极限存在。此外一个二次极限存在不能保证另一个二次极限也存在；即使两个二次极限都存在，也不一定相等。也就是说，**两个极限运算不一定可以交换次序**。

13.2.4 Continuity of Multivariable Functions 多元函数的连续性

定义 13.11 设 D 是 $\mathbb{R}^n$ 上的点集，$f:D\to\mathbb{R}^m$ 是一个映射，$x_0\in D$ 为一定点。如果

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

则称函数 f 在点 x_0 连续。用 "ε-δ" 语言来说就是：如果对于任意给定的 $\varepsilon>0$，存在 $\delta>0$，使得当 $x\in O(x_0,\delta)$ 时，成立

$$|f(x)-f(x_0)|<\varepsilon,$$

则称函数 f 在点 x_0 连续。

如果函数 f 在 D 上每一点连续，就称 f 在 D 上连续，或称 f 是 D 上的连续函数。

例题 13.1 函数 $f(x,y)=\sin\sqrt{x^2+y^2}$ 在 $\mathbb{R}^2$ 上连续。

证： 设 (x_0,y_0) 为 $\mathbb{R}^2$ 上的任一点，则有

$$\begin{aligned}
|f(x,y)-f(x_0,y_0)| &= |\sin\sqrt{x^2+y^2}-\sin\sqrt{x_0^2+y_0^2}| \\
&= 2\left|\cos\frac{\sqrt{x^2+y^2}+\sqrt{x_0^2+y_0^2}}{2}\right|\cdot\left|\sin\frac{\sqrt{x^2+y^2}-\sqrt{x_0^2+y_0^2}}{2}\right| \\
&\le 2\left|\sin\frac{\sqrt{x^2+y^2}-\sqrt{x_0^2+y_0^2}}{2}\right| \\
&\le \left|\sqrt{x^2+y^2}-\sqrt{x_0^2+y_0^2}\right| \\
&\le \sqrt{(x-x_0)^2+(y-y_0)^2} \quad \text{(利用三角不等式)}.
\end{aligned}$$

于是，对于任意给定的 $\varepsilon>0$，取 $\delta=\varepsilon$，当 $\sqrt{(x-x_0)^2+(y-y_0)^2}<\delta$ 时就成立

$$|f(x,y)-f(x_0,y_0)|<\varepsilon.$$

这说明 $f(x,y)$ 在点 (x_0,y_0) 连续。由于 (x_0,y_0) 为 $\mathbb{R}^2$ 上的任意一点，所以 $f(x,y)$ 在 $\mathbb{R}^2$ 上连续。

13.2.5 Vector-Valued Functions 向量值函数

定义 13.12 设 D 是 $\mathbb{R}^n$ 上的点集，D 到 $\mathbb{R}^m$ 的映射

$$f : D \to \mathbb{R}^m$$

$$x = (x_1, x_2, \cdots, x_n) \mapsto z = (z_1, z_2, \cdots, z_m)$$

称为 n 元 m 维向量值函数（或多元函数组），记为 $z = f(x)$。D 称为 f 的定义域，$f(D) = \{z \in \mathbb{R}^m \mid z = f(x), x \in D\}$ 称为 f 的值域。

多元函数是 $m = 1$ 的特殊情形。

显然，每个 $z_i (i = 1, 2, \cdots, m)$ 都是 x 的函数 $z_i = f_i(x)$，它称为 f 的第 i 个坐标（或分量）函数，于是，f 可以表达为分量形式

$$\begin{cases} z_1 = f_1(x), \\ z_2 = f_2(x), \\ \vdots \\ z_m = f_m(x), \end{cases} \qquad x \in D.$$

因此 f 又可表示为

$$f = (f_1, f_2, \cdots, f_m).$$

定义 13.13 设 D 是 $\mathbb{R}^n$ 上的开集，$x_0 \in D$ 为一定点，$f : D \setminus \{x_0\} \to \mathbb{R}^m$ 是映射（向量值函数），A 是一个 m 维向量。如果对于任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $x \in O(x_0, \delta) \setminus \{x_0\}$ 时，成立

$$|f(x) - A| < \varepsilon \quad (\text{即} f(x) \in O(A, \varepsilon)),$$

则称 A 为当 x 趋于 x_0 时 f 的极限，并称当 x 趋于 x_0 时 f 收敛。记为

$$\lim_{x \to x_0} f(x) = A \quad \text{或} \quad f(x) \to A \ (x \to x_0).$$

定义 13.14 设 D 是 $\mathbb{R}^n$ 上的开集，$x_0 \in D$ 为一定点。$f : D \to \mathbb{R}^m$ 是映射（向量值函数）。如果 f 满足

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

那么称 f 在 x_0 点连续。用"ε-δ"语言来说就是：如果对于任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $x \in O(x_0, \delta)$ 时，成立

$$|f(x) - f(x_0)| < \varepsilon \quad (\text{即} f(x) \in O(f(x_0), \varepsilon)),$$

则称 f 在点 x_0 连续。

如果映射 f 在 D 上每一点连续，就称 f 在 D 上连续。这时称映射 f 为 D 上的连续映射。

如果我们如前所述将 $f : D \to \mathbb{R}^m$ 表示为 $f = (f_1, f_2, \cdots, f_m)$，那么下面的定理说明了我们可以利用坐标函数来判断 f 的连续性。也就是说，映射（向量值函数）的连续性可以归结到它的坐标函数的连续性上去。

定理 13.12 设 D 是 $\mathbb{R}^n$ 上的开集，$x_0 \in D$ 为一定点。那么映射 $f : D \to \mathbb{R}^m$ 在 x_0 点连续的充分必要条件为：函数 $f_1, f_2, \cdots, f_m$ 在 x_0 点连续。

设 Ω 是 $\mathbb{R}^k$ 上的开集，D 为 $\mathbb{R}^n$ 上的开集。$g: D \to \mathbb{R}^k$ 与 $f: \Omega \to \mathbb{R}^m$ 为映射。并且 g 的值域 $g(D)$ 满足 $g(D) \subseteq \Omega$，则可以定义复合映射

$$f \circ g : D \to \mathbb{R}^m,$$

$$u \mapsto f(g(u)).$$

定理 13.13 如果 g 在 D 上连续，f 在 Ω 上连续，那么复合映射 $f \circ g$ 在 D 上连续。

13.3 Properties of Continuous Functions 连续函数的性质

13.3.1 Continuous Mappings on Compact Sets 紧集上的连续映射

定义 13.15 设 K 是 $\mathbb{R}^n$ 中的点集，$f: K \to \mathbb{R}^m$ 为映射（向量值函数），$x_0 \in K$。如果对于任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $x \in O(x_0, \delta) \cap K$ 时，成立

$$|f(x) - f(x_0)| < \varepsilon \quad (\text{即} f(x) \in O(f(x_0), \varepsilon)),$$

则称 f 在点 x_0 连续。

定理 13.14 连续映射将紧集映射成紧集。

设 $f(x)$ 是 $\mathbb{R}^n$ 中紧集 K 上的连续函数，那么 $f(K)$ 是 $\mathbb{R}$ 中的紧集，因此是有界闭集，并且数集 $f(K)$ 有最大数和最小数。于是就可得到以下结论：

定理 13.15 (有界性定理) 设 K 是 $\mathbb{R}^n$ 中紧集，f 是 K 上的连续函数。则 f 在 K 上有界。

定理 13.16 (最值定理) 设 K 是 $\mathbb{R}^n$ 中紧集，f 是 K 上的连续函数，则 f 在 K 上必能取到最大值和最小值，即存在 $\xi_1, \xi_2 \in K$，使得对于一切 $x \in K$ 成立

$$f(\xi_1) \le f(x) \le f(\xi_2)$$

现在引入一致连续的概念。

定义 13.16 设 K 是 $\mathbb{R}^n$ 中点集，$f: K \to \mathbb{R}^m$ 为映射。如果对于任意给定的 $\varepsilon > 0$，存在 $\delta > 0$，使得

$$|f(x') - f(x'')| < \varepsilon$$

对于 K 中所有满足 $|x' - x''| < \delta$ 的 x', x'' 成立，则称 f 在 K 上一致连续。

显然，一致连续的映射一定是连续的，但反之不然。下面的定理说明了紧集上的连续映射的一致连续性质。

定理 13.17 (一致连续性定理) 设 K 是 $\mathbb{R}^n$ 中紧集，$f: K \to \mathbb{R}^m$ 为连续映射，则 f 在 K 上一致连续。

13.3.2 Connected Sets and Continuous Mappings on Connected Sets 连通集与连通集上的连续映射

定义 13.17 设 S 是 $\mathbb{R}^n$ 中点集，若连续映射

$$\gamma : [0, 1] \to \mathbb{R}^n$$

的值域全部落在 S 中，即满足 $\gamma([0,1]) \subseteq S$，则称 γ 为 S 中的道路，$\gamma(0)$ 与 $\gamma(1)$ 分别称为道路的起点与终点。

若 S 中的任意两点 x, y 之间，都存在在 S 中以 x 为起点，y 为终点的道路，则称 S 为（道路）连通的，或称 S 为连通集。

直观地说，这意味着 S 中任意两点可以用全部位于 S 中的曲线相联结。

显然 $\mathbb{R}$ 上的连通子集为区间，而且 $\mathbb{R}$ 上的连通子集为紧集的充要条件为：它是闭区间。

定义 13.18 连通的开集称为（开）区域。（开）区域的闭包称为闭区域。

定理 13.18 连续映射将连通集映射成连通集。

推论 13.2 连续函数将连通的紧集映射成闭区间。

由此立即得到：

定理 13.19 (中间值定理) 设 K 为 $\mathbb{R}^n$ 中连通的紧集，f 是 K 上的连续函数，则 f 可取到它在 K 上的最小值 m 与最大值 M 之间的一切值。换言之，f 的值域是闭区间 $[m, M]$。

14 Differential Calculus of Multivariable Functions 多元函数的微分学

14.1 Partial Derivatives and Total Differential 偏导数与全微分

14.1.1 Partial Derivatives 偏导数

定义 14.1 设 $D \subset \mathbb{R}^2$ 为开集，

$$z = f(x, y), \quad (x, y) \in D$$

是定义在 D 上的二元函数，$(x_0, y_0) \in D$ 为一定点。如果存在极限

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x},$$

那么就称函数 f 在点 (x_0, y_0) 关于 x **可偏导**，并称此极限为 f 在点 (x_0, y_0) 关于 x 的**偏导数**，记为

$$\frac{\partial z}{\partial x}(x_0, y_0) \left(\text{或} f_x(x_0, y_0), \frac{\partial f}{\partial x}(x_0, y_0) \right).$$

如果函数 f 在 D 中每一点都关于 x 可偏导，则 D 中每一点 (x, y) 与其相应的 f 关于 x 的偏导数 $f_x(x, y)$ 构成了一种对应关系即二元函数关系，它称为 f 关于 x 的**偏导函数**（也称为**偏导数**），记为

$$\frac{\partial z}{\partial x} \left(\text{或} f_x(x, y), \frac{\partial f}{\partial x} \right).$$

类似地可定义 f 在点 (x_0, y_0) 关于 y 的偏导数 $\frac{\partial z}{\partial y}(x_0, y_0)$（或 $f_y(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)$）及关于 y 的偏导函数 $\frac{\partial z}{\partial y} \left(\text{或} f_y(x, y), \frac{\partial f}{\partial y} \right)$.

若 f 在点 (x_0, y_0) 关于 x 和 y 均可偏导，就简称 f 在点 (x_0, y_0) 可偏导。

“可导必定连续”是一元函数中的一条熟知的性质，但对多元函数来讲，类似性质并不成立，即可偏导未必连续。

14.1.2 Directional Derivatives 方向导数

偏导数反映的是二元函数沿 x 轴方向或 y 轴方向的变化率。而在平面 $\mathbb{R}^2$ 上，从一点出发有无穷条射线，当然也可以讨论函数沿任一射线方向的变化率。

定义 14.2 设 $D \subseteq \mathbb{R}^2$ 为开集，

$$z = f(x, y), \quad (x, y) \in D$$

是定义在 D 上的二元函数，$(x_0, y_0) \in D$ 为一定点，$v = (\cos\alpha, \sin\alpha)$ 为一个方向。如果极限

$$\lim_{t \to 0^+} \frac{f(x_0 + t\cos\alpha, y_0 + t\sin\alpha) - f(x_0, y_0)}{t}$$

存在，则称此极限为函数 f 在点 (x_0, y_0) 的沿方向 v 的**方向导数**，记为 $\frac{\partial f}{\partial v}(x_0, y_0)$。

由于 x 轴和 y 轴的正向的方向分别为 $e_1 = (1, 0)$ 和 $e_2 = (0, 1)$，从定义立即得到，函数 $f(x, y)$ 在点 (x_0, y_0) 处关于 x（或 y）可偏导的充分必要条件为 $f(x, y)$ 沿方向 e_1 和 $-e_1$（或方向 e_2 和 $-e_2$）的方向导数都存在且为相反数，且这时成立

$$\frac{\partial f}{\partial x}(x_0,y_0)=\frac{\partial f}{\partial e_1}(x_0,y_0) \quad 或 \quad \frac{\partial f}{\partial y}(x_0,y_0)=\frac{\partial f}{\partial e_2}(x_0,y_0)$$

同样，若将 $\mathbb{R}^n$ 中的单位向量 v（即满足 $\|v\|=1$ 的向量）视为一个方向，就可类似定义 n 元函数的方向导数：设 $D\subset\mathbb{R}^n$ 为开集，$x^0=(x_1^0,x_2^0,\cdots,x_n^0)$ 为 D 中一定点，$v=(v_1,v_2,\cdots,v_n)$ 为一方向。定义 D 上的 n 元函数 $u=f(x_1,x_2,\cdots,x_n)$ 在点 x^0 的沿方向 v 的方向导数为

$$\frac{\partial f}{\partial v}(x_1^0,x_2^0,\cdots,x_n^0)=\lim_{t\to 0+}\frac{f(x_1^0+tv_1,x_2^0+tv_2,\cdots,x_n^0+tv_n)-f(x_1^0,x_2^0,\cdots,x_n^0)}{t}$$

(如果等式右面的极限存在的话)。

14.1.3 Total Differential 全微分

一般地，对于函数 $z=f(x,y)$，记它的**全增量**为

$$\Delta z=f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0),$$

我们引入如下定义：

定义 14.3 设 $D\subset\mathbb{R}^2$ 为开集，

$$z=f(x,y),\quad (x,y)\in D$$

是定义在 D 上的二元函数，$(x_0,y_0)\in D$ 为一定点。

若存在只与点 (x_0,y_0) 有关而与 $\Delta x,\Delta y$ 无关的常数 A 和 B，使得

$$\Delta z=A\Delta x+B\Delta y+o(\sqrt{\Delta x^2+\Delta y^2}),$$

这里 $o(\sqrt{\Delta x^2+\Delta y^2})$ 表示在 $\sqrt{\Delta x^2+\Delta y^2}\to 0$ 时比 $\sqrt{\Delta x^2+\Delta y^2}$ 高阶的无穷小量。则称函数 f 在点 (x_0,y_0) 处是**可微**的，并将其线性主要部分 $A\Delta x+B\Delta y$ 为 f 在点 (x_0,y_0) 处的**全微分**，记为 $dz(x_0,y_0)$ 或 $df(x_0,y_0)$。

若（在 $\sqrt{\Delta x^2+\Delta y^2}\to 0$ 时）将自变量 x,y 的微分 $\Delta x,\Delta y$ 分别记为 dx,dy，那么有全微分形式

$$dz(x_0,y_0)=Adx+Bdy.$$

下面作几点说明。

首先，可以明显看出，如果函数 f 在点 (x_0,y_0) 处可微，则 f 在点 (x_0,y_0) 处是连续的，即**可微必连续**。

其次，若 $\Delta y=0$，便得到

$$f(x_0+\Delta x,y_0)-f(x_0,y_0)=A\Delta x+o(\Delta x),$$

于是

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}=A,$$

所以 $\frac{\partial f}{\partial x}(x_0,y_0)=A$。同理可证 $\frac{\partial f}{\partial y}(x_0,y_0)=B$。因此**可微必可偏导**，同时，得到全微分公式

$$df(x_0,y_0)=\frac{\partial f}{\partial x}(x_0,y_0)dx+\frac{\partial f}{\partial y}(x_0,y_0)dy.$$

定理 14.1 设 $D \subseteq \mathbb{R}^2$ 为开集，$(x_0, y_0) \in D$ 为一定点。如果函数

$$z = f(x, y), \quad (x, y) \in D$$

在 (x_0, y_0) 可微，那么对于任一方向 $\mathbf{v} = (\cos\alpha, \sin\alpha)$，$f$ 在 (x_0, y_0) 点沿方向 $\mathbf{v}$ 的方向导数存在，且

$$\frac{\partial f}{\partial \mathbf{v}}(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\cos\alpha + \frac{\partial f}{\partial y}(x_0, y_0)\sin\alpha.$$

如果函数 f 在开集（或区域）D 上的每一点都是可微的，则称 f 在 D 上可微。此时成立

$$dz = \frac{\partial f}{\partial x}(x, y)dx + \frac{\partial f}{\partial y}(x, y)dy.$$

第三，用同样的思想可以定义一般 n 元函数 $u = f(x_1, x_2, \cdots, x_n)$ 的全微分，并可得到

$$du = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n.$$

如果 $u = f(x_1, x_2, \cdots, x_n)$ 在 $x = (x_1, x_2, \cdots, x_n)$ 点可微，那么

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x_1}\cos\theta_1 + \frac{\partial f}{\partial x_2}\cos\theta_2 + \cdots + \frac{\partial f}{\partial x_n}\cos\theta_n,$$

其中 $v = (\cos\theta_1, \cos\theta_2, \cdots, \cos\theta_n)$ 为一方向，而 θ_i 就是 v 与 x_i 轴正向的夹角。

第四，一元函数的可导与可微是等价的。在高维情形可微必可偏导，但可偏导并不一定可微。事实上，一个函数即使在某一点处连续，且所有方向导数都存在，也不一定在该点可微。

但关于函数的可微性有如下的充分条件:

定理 14.2 设函数 $z = f(x, y)$ 在 (x_0, y_0) 点的某个邻域上存在偏导数，并且偏导数在 (x_0, y_0) 点连续，那么 f 在 (x_0, y_0) 点可微。

例题 14.1 求函数 $z = x^2 + 3xy - y^2$ 在点 $(1, 2)$ 处的全微分。

$$\frac{\partial z}{\partial x} = 2x + 3y, \quad \frac{\partial z}{\partial y} = 3x - 2y.$$

在点 $(1, 2)$ 处:

$$\frac{\partial z}{\partial x}(1, 2) = 2(1) + 3(2) = 8, \quad \frac{\partial z}{\partial y}(1, 2) = 3(1) - 2(2) = -1.$$

所以:

$$dz = 8dx - dy.$$

例题 14.2 求函数 $u = \ln(x^2 + y^2 + z^2)$ 的全微分。

$$\frac{\partial u}{\partial x} = \frac{2x}{x^2 + y^2 + z^2}, \quad \frac{\partial u}{\partial y} = \frac{2y}{x^2 + y^2 + z^2}, \quad \frac{\partial u}{\partial z} = \frac{2z}{x^2 + y^2 + z^2}.$$

所以:

$$du = \frac{2x}{x^2 + y^2 + z^2}dx + \frac{2y}{x^2 + y^2 + z^2}dy + \frac{2z}{x^2 + y^2 + z^2}dz.$$

例题 14.3 设 $z = e^{x^2 y}$，求 dz。

$$\frac{\partial z}{\partial x} = 2xye^{x^2y}, \quad \frac{\partial z}{\partial y} = x^2e^{x^2y}.$$

所以：

$$dz = 2xye^{x^2y}dx + x^2e^{x^2y}dy.$$

例题 14.4 求函数 $f(x, y) = \arctan\left(\frac{x}{y}\right)$ 的全微分。

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

$$\frac{\partial f}{\partial y} = -\frac{x}{y^2} \cdot \frac{1}{1+\left(\frac{x}{y}\right)^2} = -\frac{x}{x^2+y^2}.$$

所以：

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

例题 14.5 设 $w = xyz + \sin(x+y+z)$，求 dw。

$$\frac{\partial w}{\partial x} = yz + \cos(x+y+z), \quad \frac{\partial w}{\partial y} = xz + \cos(x+y+z), \quad \frac{\partial w}{\partial z} = xy + \cos(x+y+z).$$

所以：

$$dw = [yz + \cos(x+y+z)]dx + [xz + \cos(x+y+z)]dy + [xy + \cos(x+y+z)]dz.$$

例题 14.6 计算函数 $z = \frac{x}{x^2+y^2}$ 在点 $(1, 1)$ 处的全微分。

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

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

在点 $(1, 1)$ 处：

$$\frac{\partial z}{\partial x}(1,1) = \frac{1-1}{4} = 0, \quad \frac{\partial z}{\partial y}(1,1) = \frac{-2}{4} = -\frac{1}{2}.$$

所以：

$$dz = 0 \cdot dx - \frac{1}{2}dy = -\frac{1}{2}dy.$$

例题 14.7 求函数 $u = \frac{x}{\sqrt{x^2+y^2+z^2}}$ 的全微分。

令 $r = \sqrt{x^2+y^2+z^2}$，则 $u = \frac{x}{r}$。

$$\frac{\partial u}{\partial x} = \frac{r - x \cdot \frac{x}{r}}{r^2} = \frac{r^2-x^2}{r^3} = \frac{y^2+z^2}{r^3},$$

$$\frac{\partial u}{\partial y} = -\frac{x \cdot \frac{y}{r}}{r^2} = -\frac{xy}{r^3},$$

$$\frac{\partial u}{\partial z} = -\frac{x \cdot \frac{z}{r}}{r^2} = -\frac{xz}{r^3}.$$

所以：

$$du = \frac{y^2+z^2}{r^3}dx - \frac{xy}{r^3}dy - \frac{xz}{r^3}dz.$$

例题 14.8 设 $z=f(x,y)=x^y\ (x>0)$，求 dz。

$$\frac{\partial z}{\partial x}=yx^{y-1},\quad \frac{\partial z}{\partial y}=x^y\ln x.$$

所以:

$$dz=yx^{y-1}dx+x^y\ln xdy.$$

14.1.4 Gradient 梯度

定义 14.4 设 $D\subset\mathbb{R}^2$ 为开集，$(x_0,y_0)\in D$ 为一定点。如果函数 $z=f(x,y)$ 在 (x_0,y_0) 点可偏导，则称向量 $(f_x(x_0,y_0),f_y(x_0,y_0))$ 为 f 在点 (x_0,y_0) 的梯度，记为 $gradf(x_0,y_0)$，即

$$gradf(x_0,y_0)=f_x(x_0,y_0)\mathbf{i}+f_y(x_0,y_0)\mathbf{j}.$$

如果 f 在 (x_0,y_0) 点可微，注意到方向导数公式中的 $\|\mathbf{v}\|=1$，则得到它的另一种表达:

$$\boxed{\frac{\partial f}{\partial \mathbf{v}}(x_0,y_0)=\mathrm{grad}f(x_0,y_0)\cdot\mathbf{v}=\|\mathrm{grad}f(x_0,y_0)\|\cos(\mathrm{grad}f,\mathbf{v}),}$$

其中 $(\mathrm{grad}f,\mathbf{v})$ 表示 $\mathrm{grad}f(x_0,y_0)$ 与 $\mathbf{v}$ 的夹角。

由此可见,函数 f 在其任何一可微点的方向导数的绝对值不会超过它在该点的梯度的模 $\|\mathrm{grad}f\|$，且最大值 $\|\mathrm{grad}f\|$ 在梯度方向达到。这就是说，沿着梯度方向函数值增加最快。同样，f 的方向导数的最小值 $-\|\mathrm{grad}f\|$ 在梯度的反方向达到，或者说，沿着梯度相反方向函数值减少最快。

读者很容易证明梯度的下列基本性质:

(1) 若 $f=c$（c 为常数），则 $\mathrm{grad}f=0$;

(2) 若 α,β 为常数，则 $\mathrm{grad}(\alpha f+\beta g)=\alpha\mathrm{grad}f+\beta\mathrm{grad}g$;

(3) $\mathrm{grad}(f\cdot g)=f\cdot\mathrm{grad}g+g\cdot\mathrm{grad}f$;

(4) $\mathrm{grad}\left(\frac{f}{g}\right)=\frac{g\cdot\mathrm{grad}f-f\cdot\mathrm{grad}g}{g^2}\quad(g\neq0)$。

用同样的思想可以定义一般 n 元函数的梯度：设 $D\subset\mathbb{R}^n$ 为开集，$x^0=(x_1^0,x_2^0,\cdots,x_n^0)\in D$ 为一定点。如果函数 $u=f(x_1,x_2,\cdots,x_n)$ 在 x^0 点可偏导，我们称向量

$$(f_{x_1}(x_1^0,x_2^0,\cdots,x_n^0),f_{x_2}(x_1^0,x_2^0,\cdots,x_n^0),\cdots,f_{x_n}(x_1^0,x_2^0,\cdots,x_n^0))$$

为 f 在点 x^0 的梯度，记为 $\mathrm{grad}f(x_1^0,x_2^0,\cdots,x_n^0)$（或 $\mathrm{grad}f(x^0)$）。

上述关于梯度的基本性质与公式对一般 n 元函数也成立。

14.1.5 Higher-Order Partial Derivatives 高阶偏导数

设 $z=f(x,y)$ 在区域 $D\subset\mathbb{R}^2$ 上具有偏导函数

$$\frac{\partial z}{\partial x}=f_x(x,y)\quad 和\quad \frac{\partial z}{\partial y}=f_y(x,y).$$

那么在 D 上，$f_x(x,y)$ 和 $f_y(x,y)$ 都是 x,y 的二元函数。如果这两个偏导函数的偏导数也存在，则称它们是 $f(x,y)$ 的二阶偏导数。

按照对自变量的求导次序的不同，二阶偏导数有下列四种:

$$\frac{\partial^2 z}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial}{\partial x}(f_x(x,y))=f_{xx}(x,y),$$

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}(f_y(x,y)) = f_{xy}(x,y),$$
$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(f_x(x,y)) = f_{yx}(x,y),$$
$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial y}(f_y(x,y)) = f_{yy}(x,y),$$

其中第二、第三两个二阶偏导数称为混合偏导数。

可类似得到三阶、四阶以至更高阶偏导数。二阶及二阶以上的偏导数统称为高阶偏导数。

同样可对 n 元函数 $u = f(x_1, x_2, \cdots, x_n)$ 定义高阶偏导数。

定理 14.3 如果函数 $z = f(x,y)$ 的两个混合偏导数 f_{xy} 和 f_{yx} 在点 (x_0, y_0) 连续，那么等式

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

成立。

14.1.6 Higher-Order Differentials 高阶微分

设 $z = f(x,y)$ 在区域 $D \subset \mathbb{R}^2$ 上具有连续偏导数，那么它是可微的，并且

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy.$$

若 z 还具有二阶连续偏导数，那么 $\frac{\partial z}{\partial x}$ 与 $\frac{\partial z}{\partial y}$ 也是可微的，从而 dz 可微。我们称 dz 的微分为 z 的二阶微分，记为

$$d^2 z = d(dz).$$

一般地，可在 z 的 k 阶微分 $d^k z$ 的基础上定义它的 $k+1$ 阶微分为（如果存在的话）

$$d^{k+1} z = d(d^k z), \quad k = 1, 2, \cdots.$$

二阶及二阶以上的微分统称为高阶微分。

由于对自变量 x, y 总有

$$d^2 x = d(dx) = 0, \quad d^2 y = d(dy) = 0,$$

于是 $z = f(x,y)$ 的二阶微分为

$$\begin{aligned} d^2 z = d(dz) &= d\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right) \\ &= d\left(\frac{\partial z}{\partial x}\right)dx + \left(\frac{\partial z}{\partial x}\right)d^2 x + d\left(\frac{\partial z}{\partial y}\right)dy + \left(\frac{\partial z}{\partial y}\right)d^2 y \\ &= \left(\frac{\partial^2 z}{\partial x^2}dx + \frac{\partial^2 z}{\partial y \partial x}dy\right)dx + \left(\frac{\partial^2 z}{\partial x \partial y}dx + \frac{\partial^2 z}{\partial y^2}dy\right)dy \\ &= \frac{\partial^2 z}{\partial x^2}dx^2 + 2\frac{\partial^2 z}{\partial x \partial y}dxdy + \frac{\partial^2 z}{\partial y^2}dy^2, \end{aligned}$$

这里 dx^2 和 dy^2 分别表示 $(dx)^2$ 和 $(dy)^2$。

若将 $\frac{\partial}{\partial x}$ 和 $\frac{\partial}{\partial y}$ 看作求偏导数的运算符号，并约定

$$\left(\frac{\partial}{\partial x}\right)^2 = \frac{\partial^2}{\partial x^2}, \quad \left(\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial y}\right) = \frac{\partial^2}{\partial x \partial y}, \quad \left(\frac{\partial}{\partial y}\right)^2 = \frac{\partial^2}{\partial y^2},$$

那么一阶和二阶的微分公式可以分别表示为

$$dz = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)z,$$

$$d^2z = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)^2 z.$$

同样地约定

$$\left(\frac{\partial}{\partial x}\right)^p = \frac{\partial^p}{\partial x^p}, \quad \left(\frac{\partial}{\partial x}\right)^p \left(\frac{\partial}{\partial y}\right)^q = \frac{\partial^{p+q}}{\partial x^p \partial y^q}, \quad \left(\frac{\partial}{\partial y}\right)^q = \frac{\partial^q}{\partial y^q} \quad (p, q = 1, 2, \cdots),$$

读者不难用数学归纳法证明高阶微分公式

$$d^k z = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)^k z, \quad k = 1, 2, \cdots.$$

对于 n 元函数 $u = f(x_1, x_2, \cdots, x_n)$ 可同样定义各阶微分，并且成立

$$d^k u = \left(dx_1\frac{\partial}{\partial x_1} + dx_2\frac{\partial}{\partial x_2} + \cdots + dx_n\frac{\partial}{\partial x_n}\right)^k u, \quad k = 1, 2, \cdots.$$

14.1.7 Derivatives of Vector-Valued Functions 向量值函数的导数

为了方便，在本节中总是将向量记号 x, f 和 y 等视为列向量。将 $\mathbb{R}^n$ 上区域 D 上的 n 元 m 值向量值函数

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

$$x \mapsto y = f(x)$$

写成坐标分量形式

$$\begin{cases} y_1 = f_1(x_1, x_2, \cdots, x_n), \\ y_2 = f_2(x_1, x_2, \cdots, x_n), \\ \vdots \\ y_m = f_m(x_1, x_2, \cdots, x_n), \end{cases}$$

$(x_1, x_2, \cdots, x_n)^T \in D,$

并设点 $x^0 = (x_1^0, x_2^0, \cdots, x_n^0)^T \in D$（记号“T”表示转置）。

将上面关于多元函数的讨论用于 f 的每一个分量函数，即可平行地得到：

1. 若 f 的每一个分量函数 $f_i(x_1, x_2, \cdots, x_n)$ $(i = 1, 2, \cdots, m)$ 都在 x^0 点可偏导，就称向量值函数 f 在 x^0 点可导，并称矩阵

$$\left(\frac{\partial f_i}{\partial x_j}(x^0)\right)_{m \times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n}(x^0) \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n}(x^0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}(x^0) \end{pmatrix}$$

为向量值函数 f 在 x^0 点的导数或**雅可比矩阵**(Jacobian matrix),记为 $f'(x^0)$(或 $df(x^0), J_f(x^0)$)。

注：n 元函数 $z = f(x_1, x_2, \cdots, x_n)$ 是 $m = 1$ 时的特殊情形，所以它在 x^0 点的导数就是

$$f'(x^0) = \left(\frac{\partial f}{\partial x_1}(x^0), \frac{\partial f}{\partial x_2}(x^0), \cdots, \frac{\partial f}{\partial x_n}(x^0)\right).$$

如果向量值函数 f 在 D 上每一点可导，就称 f 在 D 上可导。这时对应关系

$$x \in D \mapsto f'(x) = J_f(x)$$

称为 f 在 D 上的导数，记为 $f'(x)$（或 $Df(x), J_f(x)$）。

2. 若 f 的每一个分量函数 $f_i(x_1, x_2, \cdots, x_n)$ $(i = 1, 2, \cdots, m)$ 的偏导数都在 x^0 点连续，即 f 的雅可比矩阵的每个元素都在 x^0 点连续，则称向量值函数 f 的导数在 x^0 点连续。

如果向量值函数 f 的导数在 D 上每一点连续，则称 f 的导数在 D 上连续。

3. 若存在只与 x^0 有关，而与 Δx 无关的 $m \times n$ 矩阵 A，使得在 x^0 点附近成立

$$\Delta y = f(x^0 + \Delta x) - f(x^0) = A\Delta x + o(\Delta x)$$

（其中 $\Delta x = (\Delta x_1, \Delta x_2, \cdots, \Delta x_n)^T$；$o(\Delta x)$ 是列向量，其模是 $\|\Delta x\|$ 的高阶无穷小量），则称向量值函数 f 在 x^0 点可微，并称 $A\Delta x$ 为 f 在 x^0 点的微分，记为 dy。若将 Δx 记为 $dx(dx = (dx_1, dx_2, \cdots, dx_n)^T)$，那么就有 $dy = Adx$。

如果向量值函数 f 在 D 上每一点可微，则称 f 在 D 上可微。

定理 14.4 *向量值函数 f 在 x^0 点可微的充分必要条件是它的坐标分量函数 $f_i(x_1, x_2, \cdots, x_n)$ $(i = 1, 2, \cdots, m)$ 都在 x^0 点可微。此时成立微分公式*

$$dy = f'(x^0)dx.$$

综合上述三点，我们可以得到以下的统一表述：向量值函数 f 连续、可导和可微就是它的每一个坐标分量函数 $f_i(x_1, x_2, x_3, \cdots, x_n)$ $(i = 1, 2, \cdots, m)$ 连续、可导和可微。

此外，在用雅可比矩阵定义了向量值函数的导数 $f'(x^0)$ 之后，多元函数和向量值函数的微分公式与一元函数的微分公式

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

在形式上就是完全一致的。也就是说，只要将 x, y 和 f 理解为向量，这就是多元函数和向量值函数的微分公式。

例题 14.9 *求下列函数的梯度。*

(1) $z = x^2 + y^2 \sin(xy)$

(2) $z = 1 - \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)$

(3) $u = x^2 + 2y^2 + 3z^2 + 3xy + 4yz + 6x - 2y - 5z$，*在点* $(1, 1, 1)$

$$\frac{\partial z}{\partial x} = 2x + y^2 \cdot \cos(xy) \cdot y = 2x + y^3 \cos(xy)$$

$$\frac{\partial z}{\partial y} = 2y \sin(xy) + y^2 \cdot \cos(xy) \cdot x = 2y \sin(xy) + xy^2 \cos(xy)$$

$$\nabla z = \left(2x + y^3 \cos(xy),\ 2y \sin(xy) + xy^2 \cos(xy)\right)$$

$$\frac{\partial z}{\partial x} = -\frac{2x}{a^2}, \quad \frac{\partial z}{\partial y} = -\frac{2y}{b^2}$$

$$\nabla z = \left(-\frac{2x}{a^2},\ -\frac{2y}{b^2}\right)$$

$$\frac{\partial u}{\partial x} = 2x + 3y + 6, \quad \frac{\partial u}{\partial y} = 4y + 3x + 4z - 2, \quad \frac{\partial u}{\partial z} = 6z + 4y - 5$$

在点 $(1,1,1)$：

$$\frac{\partial u}{\partial x}=2+3+6=11,\quad \frac{\partial u}{\partial y}=4+3+4-2=9,\quad \frac{\partial u}{\partial z}=6+4-5=5$$

$$\nabla u(1,1,1)=(11,9,5)$$

例题 14.10 求函数 $z=xe^{2y}$ 在点 $P(1,0)$ 处的沿从点 $P(1,0)$ 到点 $Q(2,-1)$ 方向的方向导数。

梯度：

$$\nabla z=\left(e^{2y},\ 2xe^{2y}\right),\quad \nabla z(1,0)=(1,2)$$

方向向量：

$$\vec{PQ}=(1,-1),\quad \mathbf{u}=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$$

方向导数：

$$D_{\mathbf{u}}z(1,0)=(1,2)\cdot\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)=-\frac{1}{\sqrt{2}}$$

$$\boxed{-\frac{1}{\sqrt{2}}}$$

例题 14.11 设 $z=x^2-xy+y^2$，求它在点 $(1,1)$ 处的沿方向 $\mathbf{v}=(\cos\alpha,\sin\alpha)$ 的方向导数，并指出：*(1)* 沿哪个方向的方向导数最大？*(2)* 沿哪个方向的方向导数最小？*(3)* 沿哪个方向的方向导数为零？

梯度：

$$\nabla z=(2x-y,\ -x+2y),\quad \nabla z(1,1)=(1,1)$$

方向导数：

$$D_{\mathbf{u}}z(1,1)=(1,1)\cdot(\cos\alpha,\sin\alpha)=\cos\alpha+\sin\alpha$$

$$\cos\alpha+\sin\alpha=\sqrt{2}\sin\left(\alpha+\frac{\pi}{4}\right)$$

(1) 方向导数最大时：

$$\alpha+\frac{\pi}{4}=\frac{\pi}{2}+2k\pi\quad\Rightarrow\quad\alpha=\frac{\pi}{4}+2k\pi$$

(2) 方向导数最小时：

$$\alpha+\frac{\pi}{4}=\frac{3\pi}{2}+2k\pi\quad\Rightarrow\quad\alpha=\frac{5\pi}{4}+2k\pi$$

(3) 方向导数为零时：

$$\cos\alpha+\sin\alpha=0\quad\Rightarrow\quad\tan\alpha=-1\quad\Rightarrow\quad\alpha=\frac{3\pi}{4}+k\pi$$

$$\boxed{\begin{array}{l}D_{\mathbf{u}}z=\cos\alpha+\sin\alpha\\ (1)\ \alpha=\dfrac{\pi}{4}+2k\pi\\ (2)\ \alpha=\dfrac{5\pi}{4}+2k\pi\\ (3)\ \alpha=\dfrac{3\pi}{4}+k\pi\end{array}}$$

例题 14.12 设 $z=e^{\frac{x}{y^2}}$，验证 $2x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}=0$。

易得：

$$\frac{\partial z}{\partial x} = \frac{1}{y^2} e^{\frac{x}{y^2}}$$
$$\frac{\partial z}{\partial y} = -\frac{2x}{y^3} e^{\frac{x}{y^2}}$$

例题 14.13 如果可微函数 $f(x, y)$ 在点 $(1, 2)$ 处的从点 $(1, 2)$ 到点 $(2, 2)$ 方向的方向导数为 *2*，从点 $(1, 2)$ 到点 $(1, 1)$ 方向的方向导数为 -2，求
(1) 这个函数在点 $(1, 2)$ 处的梯度；
(2) 点 $(1, 2)$ 处的从点 $(1, 2)$ 到点 $(4, 6)$ 方向的方向导数。

$$D_{\mathbf{u}_1} f = 2, \quad \mathbf{u}_1 = \frac{(2-1, 2-2)}{\|(1, 0)\|} = (1, 0)$$
$$D_{\mathbf{u}_2} f = -2, \quad \mathbf{u}_2 = \frac{(1-1, 1-2)}{\|(0, -1)\|} = (0, -1)$$

设梯度 $\nabla f(1, 2) = (a, b)$，则：

$$(a, b) \cdot (1, 0) = a = 2, \quad (a, b) \cdot (0, -1) = -b = -2 \Rightarrow b = 2$$

$$\nabla f(1, 2) = (2, 2)$$

方向 $\mathbf{u}_3 = \dfrac{(4-1, 6-2)}{\|(3, 4)\|} = \left(\dfrac{3}{5}, \dfrac{4}{5}\right)$ 方向导数：

$$D_{\mathbf{u}_3} f = (2, 2) \cdot \left(\frac{3}{5}, \frac{4}{5}\right) = \frac{6+8}{5} = \frac{14}{5}$$

$$\boxed{(2, 2)}, \quad \boxed{\frac{14}{5}}$$

例题 14.14 对于函数 $f(x, y) = xy$，在第一象限（包括边界）的每一点，指出函数值增加最快的方向。

梯度：

$$\nabla f = (y, x)$$

在第一象限（包括边界），$x \geq 0, y \geq 0$，梯度方向为 (y, x)，即增加最快的方向与向量 (y, x) 同向。

$$\boxed{\text{方向为}(y, x)}$$

14.2 Chain Rule for Multivariable Functions 多元复合函数的求导法则

14.2.1 Chain Rule 链式法则

设 $z = f(x, y), (x, y) \in D_f$ 是区域 $D_f \subset \mathbb{R}^2$ 上的二元函数，而

$$g : D_g \to \mathbb{R}^2, \quad (u, v) \mapsto (x(u, v), y(u, v))$$

是区域 $D_g \subset \mathbb{R}^2$ 上的二元二维向量值函数。如果 g 的值域 $g(D_g) \subset D_f$，那么可以构造复合函数

$$z = f \circ g = f[x(u, v), y(u, v)], \quad (u, v) \in D_g.$$

复合函数有如下求偏导数的法则。

定理 14.5 (链式法则) 设 g 在 $(u_0,v_0)\in D_g$ 点可导，即 $x=x(u,v),y=y(u,v)$ 在 (u_0,v_0) 点可偏导。记 $x_0=x(u_0,v_0),y_0=y(u_0,v_0)$，如果 f 在 (x_0,y_0) 点可微，那么

$$\frac{\partial z}{\partial u}(u_0,v_0)=\frac{\partial z}{\partial x}(x_0,y_0)\frac{\partial x}{\partial u}(u_0,v_0)+\frac{\partial z}{\partial y}(x_0,y_0)\frac{\partial y}{\partial u}(u_0,v_0),$$

$$\frac{\partial z}{\partial v}(u_0,v_0)=\frac{\partial z}{\partial x}(x_0,y_0)\frac{\partial x}{\partial v}(u_0,v_0)+\frac{\partial z}{\partial y}(x_0,y_0)\frac{\partial y}{\partial v}(u_0,v_0).$$

下面不加证明地把链式法则推至一般情况。设

$$z=f(y_1,y_2,\cdots,y_m),\quad (y_1,y_2,\cdots,y_m)\in D_f$$

为区域 $D_f\subset\mathbb{R}^m$ 上的 m 元函数。又设

$$g:D_g\to\mathbb{R}^m,\quad (x_1,x_2,\cdots,x_n)\mapsto(y_1,y_2,\cdots,y_m)$$

为区域 $D_g\subset\mathbb{R}^n$ 上的 n 元 m 维向量值函数。如果 g 的值域 $g(D_g)\subset D_f$，那么可以构造复合函数

$$z=f\circ g=f[y_1(x_1,x_2,\cdots,x_n),y_2(x_1,x_2,\cdots,x_n),\cdots,y_m(x_1,x_2,\cdots,x_n)],$$

$$(x_1,x_2,\cdots,x_n)\in D_g.$$

定理 14.6 (链式法则) 设 g 在 $x^0\in D_g$ 点可导，即 $y_1,y_2,\cdots,y_m$ 在 x^0 点可偏导，且 f 在 $y^0=g(x^0)$ 点可微，则

$$\frac{\partial z}{\partial x_i}(x^0)=\frac{\partial z}{\partial y_1}(y^0)\frac{\partial y_1}{\partial x_i}(x^0)+\frac{\partial z}{\partial y_2}(y^0)\frac{\partial y_2}{\partial x_i}(x^0)+\cdots+\frac{\partial z}{\partial y_m}(y^0)\frac{\partial y_m}{\partial x_i}(x^0),$$

$i=1,2,\cdots,n$。上式可以用矩阵表示为

$$\left(\frac{\partial z}{\partial x_1},\frac{\partial z}{\partial x_2},\cdots,\frac{\partial z}{\partial x_n}\right)_{x=x^0}=\left(\frac{\partial z}{\partial y_1},\frac{\partial z}{\partial y_2},\cdots,\frac{\partial z}{\partial y_m}\right)_{y=y^0}\begin{pmatrix}\frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n}\end{pmatrix}_{x=x^0},$$

或用向量值函数的导数记号表为

$$(f\circ g)'(x^0)=f'(y^0)\,g'(x^0).$$

例题 14.15 设 $z=f(x,y)$，其中 $x=e^t$，$y=\ln t$，求 $\dfrac{dz}{dt}$。

$$\frac{dz}{dt}=\frac{\partial z}{\partial x}\cdot\frac{dx}{dt}+\frac{\partial z}{\partial y}\cdot\frac{dy}{dt}=f_x(x,y)\cdot e^t+f_y(x,y)\cdot\frac{1}{t}.$$

例题 14.16 设 $z=x^2+y^2$，其中 $x=u\cos v$，$y=u\sin v$，求 $\dfrac{\partial z}{\partial u}$ 和 $\dfrac{\partial z}{\partial v}$。

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u}=2x\cdot\cos v+2y\cdot\sin v=2u(\cos^2 v+\sin^2 v)=2u.$$

$$\frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}=2x\cdot(-u\sin v)+2y\cdot(u\cos v)=-2u^2\cos v\sin v+2u^2\sin v\cos v=0.$$

例题 14.17 设 $w=f(x,y,z)$，其中 $x=r\cos\theta$，$y=r\sin\theta$，$z=z$，求 $\dfrac{\partial w}{\partial r}$ 和 $\dfrac{\partial w}{\partial \theta}$。

$$\frac{\partial w}{\partial r}=\frac{\partial w}{\partial x}\cdot\frac{\partial x}{\partial r}+\frac{\partial w}{\partial y}\cdot\frac{\partial y}{\partial r}+\frac{\partial w}{\partial z}\cdot\frac{\partial z}{\partial r}=f_x\cdot\cos\theta+f_y\cdot\sin\theta+f_z\cdot 0.$$
$$\frac{\partial w}{\partial \theta}=\frac{\partial w}{\partial x}\cdot\frac{\partial x}{\partial \theta}+\frac{\partial w}{\partial y}\cdot\frac{\partial y}{\partial \theta}+\frac{\partial w}{\partial z}\cdot\frac{\partial z}{\partial \theta}=f_x\cdot(-r\sin\theta)+f_y\cdot(r\cos\theta)+f_z\cdot 0.$$

例题 14.18 设 $z=f(x,y)$，其中 $x=u+v$，$y=u-v$，求 $\frac{\partial z}{\partial u}$ 和 $\frac{\partial z}{\partial v}$。

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u}=f_x\cdot 1+f_y\cdot 1=f_x+f_y.$$
$$\frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}=f_x\cdot 1+f_y\cdot(-1)=f_x-f_y.$$

例题 14.19 设 $z=\arctan\left(\frac{y}{x}\right)$，其中 $x=e^{u+v}$，$y=e^{u-v}$，求 $\frac{\partial z}{\partial u}$ 和 $\frac{\partial z}{\partial v}$。

先求 z 对 x,y 的偏导：
$$\frac{\partial z}{\partial x}=\frac{-y}{x^2+y^2},\quad \frac{\partial z}{\partial y}=\frac{x}{x^2+y^2}.$$
再求 x,y 对 u,v 的偏导：
$$\frac{\partial x}{\partial u}=e^{u+v},\quad \frac{\partial x}{\partial v}=e^{u+v},$$
$$\frac{\partial y}{\partial u}=e^{u-v},\quad \frac{\partial y}{\partial v}=-e^{u-v}.$$
于是：
$$\frac{\partial z}{\partial u}=\frac{-y}{x^2+y^2}\cdot e^{u+v}+\frac{x}{x^2+y^2}\cdot e^{u-v}=\frac{e^{u-v}-e^{u+v}}{x^2+y^2}\cdot e^{u}$$
代入 $x=e^{u+v},y=e^{u-v}$，得 $x^2+y^2=e^{2u+2v}+e^{2u-2v}=e^{2u}(e^{2v}+e^{-2v})$。
$$\frac{\partial z}{\partial u}=\frac{-e^{u-v}\cdot e^{u+v}+e^{u+v}\cdot e^{u-v}}{e^{2u}(e^{2v}+e^{-2v})}=0.$$
$$\frac{\partial z}{\partial v}=\frac{-y}{x^2+y^2}\cdot e^{u+v}+\frac{x}{x^2+y^2}\cdot(-e^{u-v})=\frac{-e^{u-v}e^{u+v}-e^{u+v}e^{u-v}}{e^{2u}(e^{2v}+e^{-2v})}=\frac{-2e^{2u}}{e^{2u}(e^{2v}+e^{-2v})}=-\frac{2}{e^{2v}+e^{-2v}}.$$

14.2.2 Invariance of the First Total Differential Form 一阶全微分的形式不变性

本段中总假设讨论的函数满足相应的可微条件。设 $z=f(x,y)$ 为二元函数，那么当 x,y 为自变量时，
$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy.$$
而当 x,y 为中间变量时，如
$$x=x(u,v),\quad y=y(u,v),$$
这时 $dx=x_u du+x_v dv$，$dy=y_u du+y_v dv$，那么由链式法则得
$$\begin{aligned}dz&=\frac{\partial z}{\partial u}du+\frac{\partial z}{\partial v}dv=(z_x x_u+z_y y_u)du+(z_x x_v+z_y y_v)dv\\&=z_x(x_u du+x_v dv)+z_y(y_u du+y_v dv)=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy.\end{aligned}$$

这说明了无论 x,y 是自变量，还是中间变量，一阶微分具有相同的形式，这就是一阶全微分的形式不变性。

对于多元函数 $z=f(y)$，其中 $y=(y_1,y_2,\cdots,y_m)^T$。当 y 为自变量时，一阶全微分形式为

$$dz=f'(y)dy.$$

而当 y 为中间变量 $y=g(x)$ $(x=(x_1,x_2,\cdots,x_n)^T)$ 时，$dy=g'(x)dx$。由链式法则定理，得

$$\begin{aligned}dz&=(f\circ g)'(x)dx=f'(y)g'(x)dx\\&=f'(y)(g'(x)dx)=f'(y)dy.\end{aligned}$$

15 Differential Equations 微分方程

15.1 Linear Difference Equations with Constant Coefficients 常系数线性差分方程

差分方程是序列 $\{y_x\}$ 满足的递推关系。序列 $\{y_x\}$ 是问题的未知量，而 $x \in \mathbb{N}$ 是一个指标。

例题 15.1 等差数列 $\{y_x\}$ 以 a 为首项、公差为 d，它满足差分方程

$$y_{x+1} = y_x + d,$$

以及初始条件

$$y_1 = a.$$

该方程的解可以通过写出前几项并观察规律得到：

$$y_x = a + (x-1)d.$$

例题 15.2 类似地，等比数列 $\{y_x\}$ 以 a 为首项、公比为 r，它满足差分方程

$$y_{x+1} = ry_x,$$

以及初始条件

$$y_1 = a.$$

该方程的解

$$y_x = ar^{x-1}$$

同样可以通过写出前几项并观察规律得到。

例题 15.3 此外，生成等差数列和等比数列的差分方程可以"结合"起来，得到形如

$$y_{x+1} = ry_x + d,$$

的差分方程，以及初始条件

$$y_1 = a.$$

上述例子中给出的方程被称为**常系数线性差分方程 (linear difference equations with constant coefficients)**。这是本课程中我们将要研究的唯一类型的差分方程。相关定义如下：

一个 n 阶**常系数线性差分方程**是形如

$$P(E)y_x = Q(x),$$

的差分方程，其中 $Q(x)$ 是一个**已知**函数，$P(E)$ 是一个关于所谓**移位算子** E 的 n 次**已知**多项式。

移位算子 E 作用于序列 $\{y_x\}$ 的规则如下：

$$E(y_x) = y_{x+1}.$$

算子 E 接收输入序列 $\{a, b, c, d, \ldots\}$ 并返回输出序列 $\{b, c, d, \ldots\}$。实际上，关系 $E(y_x) = y_{x+1}$ 意味着

$$E(y_1) = y_2, \quad 即 \quad E(a) = b,$$

$$E(y_2) = y_3, \quad 即 \quad E(b) = c,$$

$$E(y_3) = y_4, \quad 即 \quad E(c) = d,$$

依此类推。这也意味着

$$E^2(y_x) = E(y_{x+1}) = y_{x+2},$$

$$E^3(y_x) = E^2(y_{x+1}) = E(y_{x+2}) = y_{x+3},$$

依此类推。

例题 15.4 三阶差分方程

$$2y_{x+3} - 4y_{x+2} - 5y_{x+1} + 3y_x = 9x^2$$

可以表示为 $P(E)y_x = Q(x)$ 的形式，其中

$$P(E) = 2E^3 - 4E^2 - 5E + 3 \qquad 且 \qquad Q(x) = 9x^2.$$

形如 $P(E)y_x = Q(x)$ 的差分方程分为两类：

- 若 $Q(x) \neq 0$，则方程 $P(E)y_x = Q(x)$ 称为**非齐次 (non-homogeneous)** 的。
- 若 $Q(x) \equiv 0$，则方程 $P(E)y_x = 0$ 称为**齐次 (homogeneous)** 的。

下面我们将研究这两种情况，从齐次情况开始。

15.1.1 Homogeneous Linear Difference Equations with Constant Coefficients 齐次的常系数线性差分方程

例题 15.5 求差分方程

$$y_{x+2} + 6y_{x+1} + 8y_x = 0$$

的通解（即所有解的集合）。该方程具有形式 $P(E)y_x = 0$，其中多项式 $P(E)$ 为

$$P(E) = E^2 + 6E + 8.$$

注意，此多项式的常数项非零。这个要求将始终施加于差分方程。为了理解为什么这样做是合理的，可将上述中的数字 *8* 替换为 *0*，并说明由此产生的差分方程可被视为一个常数项非零的一阶差分方程。为了求方程 $(E^2 + 6E + 8)y_x = 0$ 的解，我们考虑形如 $y_x = m^x$ 的序列，其中 $m \neq 0$ 是待确定的常数。我们注意到

$$E(m^x) = m^{x+1} = mm^x$$

和

$$E^2 m^x = m^{x+2} = m^2 m^x,$$

因此差分方程 $(E^2 + 6E + 8)y_x = 0$ 变为

$$(m^2 + 6m + 8)m^x = 0.$$

因此，只要 m 是所谓的辅助方程 *(auxiliary equation)* $m^2 + 6m + 8 = 0$ 的解，序列 $y_x = m^x$ 就是 $(E^2 + 6E + 8)y_x = 0$ 的解。这里，辅助方程的解为

$$m_1 = -2 \quad 和 \quad m_2 = -4,$$

于是我们得到两个解

$$y_x = (-2)^x \quad 和 \quad y_x = (-4)^x.$$

我们知道 $y_x=(-2)^x$ 和 $y_x=(-4)^x$ 是方程 $(E^2+6E+8)y_x=0$ 的解。现在我们观察到这个方程是线性的，这意味着解 $y_x=(-2)^x$ 和 $y_x=(-4)^x$ 的任意线性组合也是解：

$$(E^2+6E+8)(\alpha(-2)^x+\beta(-4)^x)=\alpha(E^2+6E+8)(-2)^x+\beta(E^2+6E+8)(-4)^x=\alpha(0)+\beta(0)=0.$$

通过这种方式，我们得到解

$$y_x=\alpha(-2)^x+\beta(-4)^x,$$

它包含两个任意常数。此外，由于差分方程 $(E^2+6E+8)y_x$ 是二阶的，并且我们已经得到一个包含两个任意常数的解，因此我们找到了通解。这基于下面的线性代数论证：

考虑所有序列构成的向量空间 V。向量加法和标量乘法的运算（它们使得所有此类序列的集合 V 成为一个向量空间）定义如下：

$$\forall y_x\in V \text{ 和} \forall z_x\in V:(y+z)_x:=y_x+z_x,$$

$$\forall y_x\in V \text{ 和} \forall \lambda\in\mathbb{R}:(\lambda y)_x:=\lambda y_x.$$

注意，由 $o_x=0$ $(\forall x\in\mathbb{N})$ 给出的序列 o_x 在 V 中扮演零向量的角色。实际上，

$$\forall y_x\in V:(y+o)_x=y_x+o_x=y_x+0=y_x,$$

$$\forall y_x\in V:(0y)_x=o_x.$$

移位算子 $E:V\to V$ 定义为 $E(y_x)=y_{x+1}$，它将向量空间 V 映射到自身。E 将输入向量 $\{a,b,c,d,...\}\in V$ 映射为输出向量 $\{b,c,d,...\}\in V$。这是一个线性变换，因为

$$\forall y_x\in V \text{ 和} \forall z_x\in V:E((y+z)_x)=(y+z)_{x+1}=y_{x+1}+z_{x+1}=E(y_x)+E(z_x),$$

$$\forall y_x\in V \text{ 和} \forall \lambda\in\mathbb{R}:E((\lambda y)_x)=(\lambda y)_{x+1}=\lambda y_{x+1}=\lambda E(y_x).$$

此外，不难证明任何多项式 $P(E):V\to V$ 也是一个从向量空间 V 到自身的线性变换。

因此，求解齐次差分方程

$$P(E)y_x=0$$

等价于寻找线性变换

$$P(E):V\to V$$

的核，即 V 中由所有被映射到零序列 $\{o_x\}\in V$ 的序列构成的向量子空间。这样，求差分方程 $P(E)y_x=0$ 通解的问题就归结为求线性变换 $P(E):V\to V$ 的核的一组基的问题。给定这样一组基，差分方程 $P(E)y_x=0$ 的每个解 $\{y_x\}$ 都将是基向量的线性组合。现在我们需要以下定理（不加证明地陈述）：

定理 15.1 *如果由*

$$P(E)=a_nE^n+a_{n-1}E^{n-1}+\ldots+a_1E+a_0$$

定义的线性变换 $P(E):V\to V$ *满足* $a_0\neq 0$，*那么* $\dim(\ker(P(E)))=n$。

定理意味着齐次方程 $P(E)y_x=0$（其中 $P(E)=a_nE^n+a_{n-1}E^{n-1}+\ldots+a_1E+a_0$）的通解恰好包含 n 个任意常数。在例题的具体情况下，找到了算子 $P(E)=E^2+6E+8$ 的二维核的一组基 $\{(-2)^x,(-4)^x\}$ 后，我们可以确信 $y_x=\alpha(-2)^x+\beta(-4)^x$（其中 α 和 β 是任意常数）是差分方程 $y_{x+2}+6y_{x+1}+8y_x=0$ 的通解。

基于定理求解 $P(E)y_x=0$ 的一般方法如下：

求解方法

设一个 n 阶差分方程的形式为 $P(E)y_x=0$，其中多项式 $P(m)$ 的常数项不为零。我们求解辅助方程 $P(m)=0$：

情况 1：如果该方程的解 $m_1, m_2, \ldots, m_n$ 全部互异，则差分方程 $P(E)y_x=0$ 的通解为

$$y_x=\alpha_1(m_1)^x+\alpha_2(m_2)^x+\ldots+\alpha_n(m_n)^x,$$

其中 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 是任意常数。

情况 2：如果一个解 m 具有代数重数 k，则这个特定的 m 在 y_x 的通解中贡献如下形式的项：

$$y_x=\ldots\left(\beta_1+\beta_2 x+\beta_3 x^2+\ldots+\beta_k x^{k-1}\right) m^x+\ldots,$$

其中 $\beta_1, \beta_2, \ldots, \beta_k$ 是任意常数。我们将在课程后面解释为什么会出现这种形式。

例题 15.6 假设差分方程 $P(E)y_x=0$ 中的多项式 $P(E)$ 可分解如下：

$$P(m)=(m-4)(m-3)^2(m+1)(m+9)^4.$$

那么，y_x 的通解为

$$y_x=\alpha_1 4^x+(\alpha_2+\alpha_3 x)3^x+\alpha_4(-1)^x+(\alpha_5+\alpha_6 x+\alpha_7 x^2+\alpha_8 x^3)(-9)^x,$$

其中 $\alpha_1, \alpha_2, \ldots, \alpha_8$ 是任意常数。

为了完善理论，我们需要考虑辅助方程 $P(m)=0$ 的解可能为非实数的情形。由于多项式 $P(m)$ 的系数都是实数，所以每当出现非实数解 m 时，它必定伴随其共轭复数出现。

共轭复数对

假设 $P(m)=0$ 有一对共轭复数解。设这些解用极坐标指数形式表示为

$$m_1=re^{i\theta} \quad \text{和} \quad m_2=re^{-i\theta}, \quad \text{其中} \quad 0 \leq \theta<\pi.$$

同时为简单起见，假设这对解不是重复的。换句话说，m_1 和 m_2 是互异的解。

根据上述模式，y_x 的通解包含如下形式的项：

$$y_x=\ldots \alpha_1(re^{i\theta})^x+\alpha_2(re^{-i\theta})^x \ldots,$$

其中 α_1 和 α_2 是任意复常数。该项可以等价地表示为

$$y_x=\ldots\left(\alpha_1 e^{i\theta x}+\alpha_2 e^{-i\theta x}\right) r^x \ldots.$$

现在让我们施加序列 y_x 为实数的要求。不难证明，与此要求相容的、关于 α_1 和 α_2 的最一般条件是 $\alpha_1=\overline{\alpha_2}$。这导致如下形式的项：

$$y_x=\ldots\left(\beta_1 \cos(\theta x)+\beta_2 \sin(\theta x)\right) r^x \ldots,$$

其中 β_1 和 β_2 是实任意常数。

换句话说，给定一对共轭复数解 $m_1=re^{i\theta}$ 和 $m_2=re^{-i\theta}$，模长 r 被提升到 x 次幂（在 y_x 的通解中），而辐角 θ 则成为三角函数的自变量。

例题 15.7 假设与差分方程 $P(E)y_x = 0$ 关联的多项式 $P(m)$ 可分解如下：

$$P(m) \equiv (m+5)^2(m-3-2i)(m-3+2i).$$

$3 \pm 2i$ 的模长 r 等于 $r = \sqrt{3^2+2^2} = \sqrt{13}$。角度 θ（在此上下文中我们总可选择 $0 \le \theta < \pi$）通过求解以下方程组得到：

$$\begin{cases} \cos(\theta) = \frac{3}{\sqrt{13}}, \\ \sin(\theta) = \frac{2}{\sqrt{13}}. \end{cases}$$

那么 y_x 的通解为：

$$y_x = (\alpha_1 + \alpha_2 x)(-5)^x + (\alpha_3 \cos(\theta x) + \alpha_4 \sin(\theta x))13^{x/2},$$

其中 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 均为实任意常数。

最后, 在继续讨论非齐次情况 $P(E)y_x = Q(x)$ 之前, 让我们给出一个关于齐次方程 $P(E)y_x = 0$ 解的长程行为的结果。

定理 15.2 *给定一个 n 阶常系数齐次线性差分方程 $P(E)y_x = 0$，所有解 y_x 当 $x \to \infty$ 时趋于 0 的充分必要条件是辅助方程 $P(m) = 0$ 的所有根的模长都小于 1。*

15.1.2 Non-homogeneous Linear Difference Equations with Constant Coefficients 非齐次的常系数线性差分方程

考虑非齐次差分方程

$$P(E)y_x = Q(x),$$

其中 $Q(x)$ 是给定的函数。其通解的构造分为三步：

步骤 1：求出齐次部分

$$P(E)y_x = 0$$

的通解。此通解称为**余序列 (complementary sequence)**，记为 $(CS)_x$。它通过为齐次差分方程开发的方法获得。如果多项式 $P(E)$ 的次数为 n，则 $(CS)_x$ 包含 n 个实任意常数。

步骤 2：求非齐次方程

$$P(E)y_x = Q(x)$$

的一个特解。此解称为**特序列 (particular sequence)**，记为 $(PS)_x$。对于简单的函数 $Q(x)$，将在下面的例题中介绍一种获得特序列的方法。

步骤 3：非齐次方程 $P(E)y_x = Q(x)$ 的通解是特序列与余序列之和：

$$y_x = (PS)_x + (CS)_x.$$

在继续讨论例子之前，让我们确认上述表达式确实满足 $P(E)y_x = Q(x)$。

实际上，由于 $P(E)(CS)_x = 0$ 且 $P(E)(PS)_x = Q(x)$，我们看到差分方程的线性性意味着

$$\begin{aligned} P(E)y_x &= P(E)\big((PS)_x + (CS)_x\big) \\ &= P(E)(PS)_x + P(E)(CS)_x \\ &= Q(x) + 0 \\ &= Q(x). \end{aligned}$$

下面还需证明 $y_x=(PS)_x+(CS)_x$ 给出了 $P(E)y_x=Q(x)$ 的通解。为此，需要证明该方程的任意解 s_x 都可以写成 $s_x=(PS)_x+(CS)_x$ 的形式，其中 $(CS)_x$ 的常数经过适当选择。

假设给定一个解 s_x，考虑序列

$$s_x-(PS)_x.$$

这个序列满足齐次方程 $P(E)y_x=0$，因为

$$P(E)(s_x-(PS)_x)=P(E)s_x-P(E)(PS)_x=Q(x)-Q(x)=0.$$

因此，基于线性代数的论证，我们推断向量

$$s_x-(PS)_x$$

属于线性变换

$$P(E):V\to V$$

的零空间。于是，$s_x-(PS)_x$ 必须是 $P(E)$ 零空间中基向量的线性组合，因此它必须具有 $(CS)_x$ 的形式（其中常数是线性组合 $(CS)_x$ 中适当选择的标量）。这证明了任意解 s_x 都可以写成

$$s_x=(PS)_x+(CS)_x,$$

其中 $(CS)_x$ 中的常数经过适当选择。

例题 15.8 求解差分方程

$$(E^2-5E+6)y_x=5.$$

考虑齐次部分 $(E^2-5E+6)y_x=0$，我们发现辅助方程 $m^2-5m+6=0$ 给出互异根

$$m_1=3,\quad m_2=2.$$

因此余序列为

$$(CS)_x=A(3)^x+B(2)^x,$$

其中 A 和 B 是实任意常数。

对于特序列，我们尝试一个代入非齐次方程后有可能产生关于 x 的恒等式的序列 y_x。回忆一下，这正是我们所说的差分方程的解的含义。对于简单的 $Q(x)$，通常可以通过考虑一个与函数 $Q(x)$ 具有相同一般形式的特序列来实现。这里 $Q(x)=5$，因此我们尝试一个常数序列 $y_x=a$。确实，将 $y_x=a$ 代入方程

$$y_{x+2}-5y_{x+1}+6y_x=5$$

得

$$a-5a+6a=5,$$

这意味着 $a=\frac{5}{2}$，因此一个特序列为

$$(PS)_x=\frac{5}{2}.$$

所以，方程 $(E^2-5E+6)y_x=5$ 的通解为

$$y_x=(PS)_x+(CS)_x=\frac{5}{2}+A(3)^x+B(2)^x,$$

其中 A 和 B 是实任意常数。注意通解包含两个任意常数，正如二阶差分方程应有的情况。

例题 15.9 求解差分方程

$$y_{x+1} = (y_x + D)(1 + r),$$

满足初始条件

$$y_1 = P(1 + r).$$

P 和 D 是存款额，r 是年利率，x 是指示年份的指标，y_x 是累积金额。我们将此差分方程整理为形式

$$y_{x+1} - (1 + r)y_x = D(1 + r).$$

对于齐次部分 $(E - (1 + r))y_x = 0$，辅助方程

$$m - (1 + r) = 0$$

给出单根

$$m = (1 + r).$$

因此余序列为

$$(CS)_x = A(1 + r)^x,$$

其中 A 是任意常数。对于特序列，我们尝试

$$y_x = a,$$

其中 a 是将 y_x 代入非齐次方程 $y_{x+1} - (1 + r)y_x = D(1 + r)$ 后确定的常数。我们得到

$$a - (1 + r)a = D(1 + r),$$

由此得

$$a = -\frac{D}{r}(1 + r).$$

因此，一个特序列为

$$(PS)_x = -\frac{D}{r}(1 + r).$$

差分方程 $y_{x+1} - (1 + r)y_x = D(1 + r)$ 的通解因此为

$$y_x = -\frac{D}{r}(1 + r) + A(1 + r)^x,$$

其中 A 是任意常数。为了确定 A，我们使用初始条件 $y_1 = P(1 + r)$。这给出

$$\left(A - \frac{D}{r}\right)(1 + r) = P(1 + r),$$

这意味着

$$A = P + \frac{D}{r}.$$

因此，我们问题的解为

$$y_x = -\frac{D}{r}(1 + r) + \left(P + \frac{D}{r}\right)(1 + r)^x,$$

或等价地，

$$y_x = P(1 + r)^x + \frac{D}{r}[(1 + r)^x - (1 + r)].$$

15.2 ODEs and PDEs 常微分方程与偏微分方程

一个关于 f 且仅涉及一个自变量的微分方程称为常微分方程。一个关于 f 且涉及两个或更多自变量的微分方程称为偏微分方程。

例题 15.10 方程 $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 2f$ 是 $f(x, y, z)$ 的一个偏微分方程。该方程的一个解是 $f = x^2 + y^2 + z^2$。如果在微分方程中将 f 替换为 $x^2 + y^2 + z^2$，我们得到恒等式 $2x^2 + 2y^2 + 2z^2 = 2x^2 + 2y^2 + 2z^2$。

微分方程的阶是方程中出现的任何导数的最高阶数。

例题 15.11 方程 $x\frac{\partial f}{\partial x} = f^2 + y$ 是 $f(x, y)$ 的一个一阶偏微分方程，而方程 $\frac{d^3 f}{dx^3} = f^5 + 2x^6 - \frac{df}{dx}$ 是 $f(x)$ 的一个三阶常微分方程。

微分方程的次数是最高阶导数在微分方程中出现的代数次数。

例题 15.12 方程 $x\left(\frac{\partial f}{\partial x}\right)^4 + \frac{\partial f}{\partial y} = f$ 是 $f(x, y)$ 的一个四次一阶偏微分方程，而方程 $\left(\frac{d^3 f}{dx^3}\right)^2 = \left(\frac{df}{dx}\right)^5 + f$ 是 $f(x)$ 的一个二次三阶常微分方程。

微分方程的通解是该方程所有解的集合。

15.3 Linear ODEs with Constant Coefficients 常系数线性常微分方程

一个 n 阶常系数线性常微分方程是形如

$$P(D)y = Q(x)$$

的函数 $y(x)$ 的微分方程，其中 $Q(x)$ 是给定函数，$P(D)$ 是关于微分算子 $D = \frac{d}{dx}$ 的 n 次给定多项式。

例题 15.13 三阶线性微分方程

$$2\frac{d^3 y}{dx^3} + \frac{d^2 y}{dx^2} - 5\frac{dy}{dx} + 3y = \sin(x)$$

可以表示为 $P(D)y = Q(x)$ 的形式，其中

$$P(D) = 2D^3 + D^2 - 5D + 3$$

且

$$Q(x) = \sin(x)$$

与差分方程类似，

- 若 $Q(x) \neq 0$，则方程 $P(D)y = Q(x)$ 称为**非齐次**的。
- 若 $Q(x) \equiv 0$，则方程 $P(D)y = 0$ 称为**齐次**的。

下面我们将研究这两种情况，从齐次情况开始。

15.3.1 Homogeneous Linear ODEs with Constant Coefficients 齐次的常系数线性常微分方程

求解方法

给定任何 n 阶形如 $P(D)y=0$ 的常微分方程，我们求解多项式方程 $P(m)=0$，它再次被称为辅助方程。

情况 1：如果解 $m_1,m_2,...,m_n$ 全部互异，则微分方程 $P(D)y=0$ 的通解为

$$y(x)=\alpha_1 e^{m_1 x}+\alpha_2 e^{m_2 x}+\cdots+\alpha_n e^{m_n x},$$

其中 $\alpha_1,\alpha_2,...,\alpha_n$ 是任意常数。

情况 2：如果一个解 m 具有代数重数 k，则这个特定的 m 在 $y(x)$ 的通解中贡献如下形式的项：

$$y(x)=\cdots+(\beta_1+\beta_2 x+\beta_3 x^2+...+\beta_k x^{k-1})e^{mx}+\cdots,$$

其中 $\beta_1,\beta_2,...,\beta_k$ 是任意常数。

例题 15.14 假设微分方程 $P(D)y=0$ 中的多项式 $P(D)$ 使得辅助方程 $P(m)$ 可分解如下：

$$P(m)=(m-4)(m-3)^2(m+1)(m+9)^4.$$

那么，$y(x)$ 的通解为

$$y(x)=\alpha_1 e^{4x}+(\alpha_2+\alpha_3 x)e^{3x}+\alpha_4 e^{-x}+(\alpha_5+\alpha_6 x+\alpha_7 x^2+\alpha_8 x^3)e^{-9x},$$

其中 $\alpha_1,\alpha_2,...,\alpha_8$ 是任意常数。

例题 15.15 求解四阶微分方程

$$\frac{d^4y}{dx^4}-\frac{d^3y}{dx^3}-2\frac{d^2y}{dx^2}=0.$$

多项式 $P(D)$ 为 $P(D)=D^4-D^3-2D^2$，因此关于 m 的辅助方程为

$$m^4-m^3-2m^2=0.$$

我们因式分解该方程的左侧，得到

$$m^4-m^3-2m^2\equiv m^2(m^2-m-2)\equiv m^2(m-2)(m+1),$$

所以解为

$$m=0\ (\text{代数重数 }2),\quad m=2\quad \text{和}\quad m=-1.$$

因此，微分方程 $(D^4-D^3-2D^2)y=0$ 的通解为

$$y(x)=(\alpha_1+\alpha_2 x)e^{0x}+\alpha_3 e^{2x}+\alpha_4 e^{-1x}=\alpha_1+\alpha_2 x+\alpha_3 e^{2x}+\alpha_4 e^{-x},$$

其中 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是任意常数。

最后，通过考虑当辅助方程 $P(m)=0$ 的解为非实数时会发生什么，来完善对微分方程 $P(D)y=0$ 的研究。回想一下，由于多项式 $P(m)$ 的系数都是实数，所以每当出现非实数解 m 时，它必定伴随其共轭复数出现。

共轭复数对

假设 $P(m)=0$ 有一对共轭复数解。设这些解为

$$m_1 = a + ib \quad 和 \quad m_2 = a - ib,$$

其中 a 和 b 是实数。同时为简单起见，假设这对解不是重复的。换句话说，m_1 和 m_2 是互异的解。那么 $y(x)$ 的通解包含如下形式的项:

$$y(x) = \dots \alpha_1 e^{m_1 x} + \alpha_2 e^{m_2 x} \dots,$$

其中 α_1 和 α_2 是任意复常数。利用关系 $m_1 = a + ib$ 和 $m_2 = a - ib$，该项变为

$$y(x) = \dots \alpha_1 e^{ax+ibx} + \alpha_2 e^{ax-ibx} \dots,$$

可以表示为

$$y(x) = \dots (\alpha_1 e^{ibx} + \alpha_2 e^{-ibx}) e^{ax} \dots$$

沿用与差分方程相同的论证，要求 $y(x)$ 为实函数会迫使 α_1 和 α_2 为共轭复数，在这种情况下我们得到

$$y(x) = \dots (\beta_1 \cos(bx) + \beta_2 \sin(bx)) e^{ax} \dots,$$

其中 β_1 和 β_2 是实任意常数。

总结一下，给定一对共轭复数解 m_1 和 m_2，实部 a 成为指数函数的自变量，虚部 b 成为三角函数的自变量。这样的解描述了振幅随时间变化的振荡。

例题 15.16 假设与微分方程 $P(D)y = 0$ 关联的多项式 $P(m)$ 可分解如下:

$$P(m) \equiv (m+5)^2 (m-3-2i)(m-3+2i).$$

那么，$y(x)$ 的通解为

$$y(x) = (\alpha_1 + \alpha_2 x) e^{-5x} + (\alpha_3 \cos(2x) + \alpha_4 \sin(2x)) e^{3x},$$

其中 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 均为实任意常数。

最后，我们给出一个关于 $P(D)y = 0$ 解的长程行为的结果:

给定一个 n 阶常系数齐次线性微分方程 $P(D)y = 0$，所有解 $y(x)$ 当 $x \to \infty$ 时趋于 0 的充分必要条件是辅助方程 $P(m) = 0$ 的所有根都具有负实部。

15.3.2 Non-homogeneous Linear ODEs with Constant Coefficients 非齐次的常系数线性常微分方程

与差分方程类似，非齐次方程 $P(D)y = Q(x)$ 的通解分三步构造:

步骤 1: 求出齐次部分

$$P(D)y(x) = 0$$

的通解。此解称为**余函数 (the complementary function)**，记为 $(CF)(x)$。它可以通过为齐次常微分方程开发的方法获得。如果多项式 $P(D)$ 的次数为 n，则 $(CF)(x)$ 包含 n 个实任意常数。

步骤 2: 然后求非齐次方程

$$P(D)y(x) = Q(x)$$

的一个特解。此解称为**特积分 (a particular integral)**，记为 $(PI)(x)$。我们获得特积分的方法与为差分方程引入的方法类似。

步骤 3: 非齐次方程 $P(D)y(x) = Q(x)$ 的通解是特积分与余函数之和:

$$y(x) = (PI)(x) + (CF)(x).$$

例题 15.17 求解微分方程

$$(D^2 - 3D + 2)y(x) = 5.$$

考虑齐次部分 $(D^2 - 3D + 2)y(x) = 0$，我们发现辅助方程 $m^2 - 3m + 2 = 0$ 给出互异根

$$m_1 = 1, \quad m_2 = 2.$$

因此余函数为

$$(CF)(x) = Ae^x + Be^{2x},$$

其中 A 和 B 是实任意常数。

对于一个特积分，我们尝试一个函数 $(PI)(x)$，它代入非齐次微分方程后有可能产生关于 x 的恒等式。回想一下，这正是我们所说的解的含义。

这里，我们需要满足 $y''(x) - 3y'(x) + 2y(x) = 5$，所以一个常数解 $y(x) = a$ 肯定可行。确实，将 $y(x) = a$ 代入 $y''(x) - 3y'(x) + 2y(x) = 5$，我们得到

$$2a = 5,$$

这意味着一个特积分为

$$(PI)(x) = \frac{5}{2}.$$

因此，方程 $(D^2 - 3D + 2)y(x) = 5$ 的通解为

$$y(x) = (PI)(x) + (CF)(x) = \frac{5}{2} + Ae^x + Be^{2x},$$

其中 A 和 B 是实任意常数。

15.4 More General Forms of ODEs 更一般形式的常微分方程

15.4.1 Separable ODEs 可分离变量常微分方程

一个可分离的常微分方程是可以整理成如下形式的方程：

$$\frac{dy}{dx} = F(x)G(y),$$

其中 $F(x)$ 和 $G(y)$ 是给定的 x 和 y 的函数。注意右边的乘积结构。最熟悉的可分离微分方程情形是 $G(y) = 1$。此时方程变为

$$\frac{dy}{dx} = F(x).$$

其通解由 $F(x)$ 的任一原函数加上一个任意常数给出：

$$y(x) = \int F(x)dx + C.$$

例题 15.18 求解微分方程

$$\frac{dy}{dx} = \frac{y^2}{x^2}.$$

我们看出这是一个可分离方程。我们将其整理为形式

$$\int \frac{dy}{y^2} = \int \frac{dx}{x^2}$$

并进行积分，得到

$$-\frac{1}{y} = -\frac{1}{x} + C.$$

使 y 成为该方程的主项，得到所谓的通解的**显式形式**，即

$$y(x) = \frac{x}{1 - Cx}.$$

15.4.2 Exact ODEs 恰当常微分方程

考虑一个一阶微分方程，它被整理成如下形式：

$$M(x,y)dx + N(x,y)dy = 0,$$

其中 $M(x,y)$ 和 $N(x,y)$ 是给定的 x 和 y 的函数。该方程等价于

$$\frac{dy}{dx} = -\frac{M(x,y)}{N(x,y)},$$

所以它是一个相当一般的一阶微分方程。现在假设存在一个函数 $F(x,y)$，具有以下性质：

$$\frac{\partial F}{\partial x} = M(x,y) \quad \text{且} \quad \frac{\partial F}{\partial y} = N(x,y).$$

注意这样的函数可能实际上并不存在；然而，如果它存在，微分方程 $M(x,y)dx + N(x,y)dy = 0$ 可以表示为

$$\frac{\partial F(x,y)}{\partial x}dx + \frac{\partial F(x,y)}{\partial y}dy = 0.$$

两边除以 dx，我们得到等价方程

$$\frac{\partial F(x,y)}{\partial x} + \frac{\partial F(x,y)}{\partial y}\frac{dy}{dx} = 0,$$

这意味着复合函数 $F(x,y(x))$ 关于 x 的导数等于 0；即，

$$\frac{d}{dx}F(x,y(x)) = 0.$$

因此，对于某个任意常数 C，表达式 $F(x,y(x)) = C$ 给出了微分方程的通解。我们说关系式 $F(x,y) = C$ 隐式地定义了函数 $y(x)$。

在我们进一步发展这个理论之前，考虑一个形式为 $M(x,y)dx + N(x,y)dy = 0$ 的微分方程的例子，对于这个方程，确实存在一个满足

$$\frac{\partial F}{\partial x} = M(x,y) \quad \text{且} \quad \frac{\partial F}{\partial y} = N(x,y)$$

性质的函数 $F(x,y)$。这个例子阐明了“$F(x,y) = C$ 隐式地定义了通解 $y(x)$”这一陈述的含义。

例题 15.19 求解微分方程 $(y + 3x^2)dx + xdy = 0$。

按照前面的方法，我们希望找到一个函数 $F(x,y)$，使得

$$\frac{\partial F}{\partial x} = y + 3x^2 \quad \text{且} \quad \frac{\partial F}{\partial y} = x.$$

我们有一个最简单类型的联立偏微分方程组。让我们通过求解左边的偏微分方程的通解，然后将这个解代入右边的偏微分方程，以获得与两个偏微分方程都一致的解来求解它。

左边偏微分方程的通解为 $F(x,y) = xy + x^3 + g(y)$，其中 $g(y)$ 是任意函数。将这个解代入右边的偏微分方程，我们得到

$$x + g'(y) = x,$$

这意味着 $g(y) = A$，其中 A 是任意常数。因此，我们更新解 $F(x,y) = xy + x^3 + g(y)$，现在变为 $F(x,y) = xy + x^3 + A$。

根据迄今发展的理论，微分方程 $(y + 3x^2)dx + xdy = 0$ 的通解 $y(x)$ 通过令 $F(x,y)$ 等于常数 C 而隐式地得到。意识到 $F(x,y) = xy + x^3 + A$ 中的常数 A 可以被吸收到常数 C 中，我们有

$$xy + x^3 = C.$$

这隐式地定义了关于 x 的通解 $y(x)$。通过将 y 作为该方程的主项，我们得到显式形式的通解 $y(x)$：

$$y(x) = \frac{C}{x} - x^2.$$

你可以确认这个表达式满足微分方程

$$\frac{dy}{dx} = \frac{-y - 3x^2}{x},$$

它对应于展开形式 $(y + 3x^2)dx + xdy = 0$。

现在让我们通过陈述关于给定函数 $M(x, y)$ 和 $N(x, y)$ 的条件（这些条件保证函数 $F(x, y)$ 的存在）来继续理论的发展。如例题所示，函数 $F(x, y)$ 存在，只要它满足偏微分方程组

$$\frac{\partial F}{\partial x} = M(x, y) \quad \text{且} \quad \frac{\partial F}{\partial y} = N(x, y).$$

然而，由于 $F(x, y)$ 的偏导数可交换，上述方程组成立当且仅当

$$\frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial N}{\partial x}.$$

任何形如 $M(x, y)dx + N(x, y)dy = 0$ 且给定的函数 $M(x, y)$ 和 $N(x, y)$ 满足

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

的微分方程称为**恰当常微分方程**。

回顾例题，我们可以确认

$$\frac{\partial M}{\partial y} = 1 \quad \text{且} \quad \frac{\partial N}{\partial x} = 1.$$

所以例题中的微分方程是恰当的，这解释了为什么在那种情况下可以找到函数 $F(x, y)$。

现在我们能够总结求解微分方程

$$M(x, y)dx + N(x, y)dy = 0$$

的方法。

首先检查是否满足

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

如果这个关系不成立，则方程不是恰当的，因此我们需要采用其他方法。如果这个关系成立，则方程是恰当的，并且 $F(x, y)$ 存在。为了找到 $F(x, y)$，我们求解偏微分方程组

$$\frac{\partial F}{\partial x} = M(x, y) \quad \text{且} \quad \frac{\partial F}{\partial y} = N(x, y),$$

由于方程是恰当的，该方程组解的存在性得到保证。找到 $F(x, y)$ 后，令其等于常数 C。关系式 $F(x, y) = C$ 隐式地定义了微分方程 $M(x, y)dx + N(x, y)dy = 0$ 的通解 $y(x)$。如果可能，我们使 y 成为 $F(x, y) = C$ 的主项，以显式形式得到 $y(x)$ 的通解。

15.4.3 Linear ODEs 线性常微分方程

一个关于 y 的线性常微分方程是一个关于函数 $y(x)$ 的方程，它可以整理为形式

$$\frac{dy}{dx} + P(x)y = Q(x),$$

其中 $P(x)$ 和 $Q(x)$ 是给定的函数。为了推导该方程的解，我们将其表示为形式

$$(P(x)y - Q(x))\,dx + (1)dy = 0$$

并检查这个方程是否是恰当的。我们有：

$$\frac{\partial\,(P(x)y - Q(x))}{\partial y} = P(x) \quad \text{和} \quad \frac{\partial(1)}{\partial x} = 0.$$

我们得出结论，除非给定函数 $P(x) = 0$，否则这个方程不是恰当的。在 $P(x) = 0$ 的情况下，方程变为可分离方程 $\frac{dy}{dx} = Q(x)$。我们已经知道如何求解可分离方程，所以 $P(x) = 0$ 的情况没有真正的意义。在 $P(x)$ 非零的情况下，有趣的是，方程

$$(P(x)y - Q(x))\,dx + dy = 0$$

乘以函数 $I(x) = e^{\int P(x)dx}$ 后就变为恰当的。这个函数 $I(x)$ 称为**积分因子 (integrating factor)**。为了确认这一点，考虑等价的微分方程，整理为形式

$$\left(e^{\int P(x)dx}P(x)y - e^{\int P(x)dx}Q(x)\right)dx + e^{\int P(x)dx}dy = 0$$

并进行标准检验。我们有：

$$\frac{\partial\left(e^{\int P(x)dx}P(x)y - e^{\int P(x)dx}Q(x)\right)}{\partial y} = e^{\int P(x)dx}P(x) \quad \text{和} \quad \frac{\partial e^{\int P(x)dx}}{\partial x} = e^{\int P(x)dx}P(x),$$

所以该方程现在是恰当的。因此，让我们将其作为恰当方程来求解。我们知道存在一个函数 $F(x,y)$，使得

$$\frac{\partial F}{\partial x} = e^{\int P(x)dx}P(x)y - e^{\int P(x)dx}Q(x) \quad \text{和} \quad \frac{\partial F}{\partial y} = e^{\int P(x)dx}.$$

让我们先积分右边的偏微分方程，以获得其通解。我们看到它由下式给出：

$$F = e^{\int P(x)dx}y + g(x),$$

其中 $g(x)$ 是任意函数。我们将这个通解代入左边的偏微分方程，以获得与两个偏微分方程都一致的解。我们发现

$$e^{\int P(x)dx}P(x)y + g'(x) = e^{\int P(x)dx}P(x)y - e^{\int P(x)dx}Q(x),$$

这简化为

$$g'(x) = -e^{\int P(x)dx}Q(x).$$

因此，

$$g(x) = -\int e^{\int P(x)dx}Q(x)dx,$$

并且函数 $F(x,y)$ 更新为

$$F(x,y) = e^{\int P(x)dx}y - \int e^{\int P(x)dx}Q(x)dx.$$

最后，我们令 $F(x,y)$ 等于常数 C，以获得线性微分方程 $\frac{dy}{dx} + P(x)y = Q(x)$ 关于函数 $y(x)$ 的通解。这给出

$$e^{\int P(x)dx}y - \int e^{\int P(x)dx}Q(x)dx = C.$$

记积分因子为 $I(x) = e^{\int P(x)dx}$，我们得到一个相当简单的 $y(x)$ 的显式形式通解：

$$y(x) = \frac{1}{I(x)}\left(\int I(x)Q(x)dx + C\right).$$

既然我们已经推导出这个结果，我们可以简单地记住它。因此，我们有以下求解方法：

给定一个关于 y 的线性微分方程，将其整理为形式

$$\frac{dy}{dx} + P(x)y = Q(x)$$

（即，整理使得 $\frac{dy}{dx}$ 的系数等于 1）并计算积分因子 $I(x) = e^{\int P(x)dx}$。然后，$y(x)$ 的通解由下式给出：

$$y(x) = \frac{1}{I(x)} \left(\int I(x)Q(x)dx + C \right).$$

15.4.4 Homogeneous ODEs 齐次常微分方程

回想一下，如果函数 $f(x, y)$ 满足

$$f(\lambda x, \lambda y) = \lambda^n f(x, y),$$

则称其为 n 次齐次函数。

一个 n 次**齐次**常微分方程是一个可以表示为如下形式的微分方程：

$$M(x, y) + N(x, y)\frac{dy}{dx} = 0,$$

其中**给定**的函数 $M(x, y)$ 和 $N(x, y)$ 都是 n 次齐次函数。

为了求解齐次常微分方程，我们用新的因变量 $z(x)$ 替换因变量 $y(x)$，其定义为

$$z(x) = \frac{y(x)}{x}.$$

换句话说，我们根据下式用 x、z 和 $\frac{dz}{dx}$ 表示 y 和 $\frac{dy}{dx}$：

$$y = xz \quad 和 \quad \frac{dy}{dx} = z + x\frac{dz}{dx},$$

并在微分方程中使用这些表达式以消除 y 和 $\frac{dy}{dx}$。可以证明，这总是导致一个关于 $z(x)$ 的可分离常微分方程。求解关于 $z(x)$ 的这个微分方程后，我们利用关系式 $y(x) = xz(x)$ 来获得相应的 $y(x)$ 的解。

例题 15.20 对于 $x > 0$，求解常微分方程

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

满足条件：当 $x = 1$ 时，$y = 7$。

我们观察到这是一个二次齐次常微分方程。我们使用关系式

$$y = xz \quad 和 \quad \frac{dy}{dx} = z + x\frac{dz}{dx}$$

得到一个关于函数 $z(x)$ 的常微分方程，即

$$2x^2\left(z + x\frac{dz}{dx}\right) = x^2 + x^2z^2.$$

我们消去因子 x^2，

$$2\left(z + x\frac{dz}{dx}\right) = 1 + z^2,$$

并将项 $2z$ 移到右边。得到的方程显然是可分离的：

$$2x\frac{dz}{dx} = z^2 - 2z + 1.$$

我们分离变量并积分：

$$2\int \frac{dz}{z^2-2z+1} = \int \frac{dx}{x}.$$

左边被积函数的分母是一个完全平方，所以我们有

$$2\int \frac{dz}{(z-1)^2} = \int \frac{dx}{x},$$

这给出了 $z(x)$ 的隐式形式通解：

$$-\frac{2}{z-1} = \ln(x) + C.$$

注意，因为我们被告知 $x > 0$，所以不需要写 $\ln|x|$。

在应用条件 $(x, y) = (1, 7)$ 之前，让我们找出函数 $y(x)$ 的对应解。实际上，我们可以得到后者的显式形式。为此，我们将 $z(x)$ 作为上述关系式的主项，得到

$$z = 1 - \frac{2}{\ln(x) + C},$$

然后将 z 替换为比值 $\frac{y}{x}$，以获得 $y(x)$ 的通解：

$$y = x\left(1 - \frac{2}{\ln(x) + C}\right).$$

最后，利用当 x 等于 1 时 y 等于 7 的条件，我们得到

$$7 = 1 - \frac{2}{C}.$$

该方程的解为

$$C = -\frac{1}{3}.$$

因此，满足微分方程和给定条件的 $y(x)$ 的特解为

$$y = x\left(1 - \frac{2}{\ln(x) - \frac{1}{3}}\right).$$

15.5 Systems of Difference and Differential Equations 差分与微分方程系统

我们需要用到线性代数的知识来求解差分/微分方程组（系统）。

16 Multiple Integrals 重积分

16.1 Multiple Integrals over Bounded Closed Regions 有界闭区域上的重积分

16.1.1 Double Integrals 二重积分的概念

定义 16.1 设 D 为 $\mathbb{R}^2$ 上的零边界闭区域，函数 $z=f(x,y)$ 在 D 上有界。将 D 用曲线网分成 n 个小区域 $\Delta D_1,\Delta D_2,\cdots,\Delta D_n$，它称为 D 的一个划分，并记所有的小区域 ΔD_i 的最大直径为 λ，即

$$\lambda=\max_{1\le i\le n}|\operatorname{diam}\Delta D_i|.$$

在每个 ΔD_i 上任取一点 (ξ_i,η_i)，记 $\Delta\sigma_i$ 为 ΔD_i 的面积，若 λ 趋于零时，和式

$$\sum_{i=1}^{n}f(\xi_i,\eta_i)\Delta\sigma_i$$

的极限存在且与区域的分法和点 (ξ_i,η_i) 的取法无关，则称 $f(x,y)$ 在 D 上可积，并称此极限为 $f(x,y)$ 在 D 上的二重积分，记为

$$\iint_D f(x,y)\,d\sigma=\lim_{\lambda\to 0}\sum_{i=1}^{n}f(\xi_i,\eta_i)\Delta\sigma_i$$

$f(x,y)$ 称为被积函数，D 称为积分区域，x 和 y 称为积分变量，$d\sigma$ 称为面积元素，

$$\iint_D f(x,y)\,d\sigma$$

也称为积分值。

性质 1 若在已有的划分上添加有限条曲线作进一步划分，则 Darboux 大和不增，Darboux 小和不减。

性质 2 任何一个 Darboux 小和都不大于任何一个 Darboux 大和。因此，若记 $I^*=\inf\{S\}$，$I_*=\sup\{s\}$（这里上、下确界是对所有划分来取的），则有

$$s\le I_*\le I^*\le S.$$

性质 3 $f(x,y)$ 在 D 上可积的充分必要条件是：

$$\lim_{\lambda\to 0}(S-s)=0,$$

即

$$\lim_{\lambda\to 0}\sum_{i=1}^{n}\omega_i\Delta\sigma_i=0.$$

这里 $\omega_i=M_i-m_i$ 是 $f(x,y)$ 在 ΔD_i 上的**振幅**。此时成立

$$\lim_{\lambda\to 0}s=\lim_{\lambda\to 0}S=\iint_D f(x,y)d\sigma.$$

利用这个性质就可得到：

定理 16.1 若 $f(x,y)$ 在零边界闭区域 D 上连续，那么它在 D 上可积。

16.1.2 Multiple Integrals 多重积分

同 $\mathbb{R}^2$ 中定义面积一样，可以在 $\mathbb{R}^n(n \geq 3)$ 中引入体积的概念。若定义 $\mathbb{R}^n$ 中的 n 维闭矩形 $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ 的体积为 $(b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$，那么就可以将 $\mathbb{R}^2$ 上定义面积的叙述完全平移到 $\mathbb{R}^n(n \geq 3)$ 上来定义体积，并同样称边界体积为零的有界区域为零边界区域，而且可以证明光滑曲面片的体积为零，这里就不一一详述了。设 Ω 是 $\mathbb{R}^n(n \geq 3)$ 上的有界区域，其边界是一张或数张无重点的封闭曲面（本章总是如此假定），那么同样可得：有界点集 Ω 是可求体积的充分必要条件为其边界的体积为零，即 Ω 为零边界区域。

同 $\mathbb{R}^2$ 中的原理一样，我们引入 n 重积分的概念：

定义 16.2 设 Ω 为 $\mathbb{R}^n$ 上的零边界闭区域，函数 $u = f(x)$ 在 Ω 上有界。将 Ω 用曲面网分成 n 个小区域 $\Delta\Omega_1, \Delta\Omega_2, \cdots, \Delta\Omega_n$（称为 Ω 的一个划分），记 ΔV_i 为 $\Delta\Omega_i$ 的体积，并记所有的小区域 $\Delta\Omega_i$ 的最大直径为 λ。在每个 $\Delta\Omega_i$ 上任取一点 x_i，若 λ 趋于零时，和式

$$\sum_{i=1}^{n} f(x_i)\Delta V_i$$

的极限存在且与区域的分法和点 x_i 的取法无关，则称 $f(x)$ 在 Ω 上可积，并称此极限为 $f(x)$ 在有界闭区域 Ω 上的 n 重积分，记为

$$\int_{\Omega} f dV = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(x_i)\Delta V_i$$

$f(x)$ 称为被积函数，Ω 称为积分区域，x 称为积分变量，dV 称为体积元素，$\int_{\Omega} f dV$ 也称为积分值。

类似于二维情形可知，若 $f(x)$ 在零边界闭区域 Ω 上连续，那么它在 Ω 上可积。

注意到 $n = 2$ 与 $n > 2$ 的重积分定义并没有本质的区别，今后我们经常把它们一起讨论。为明确起见，通常采用如下记法：

在 $\mathbb{R}^2$ 中，$f(x, y)$ 在 D 上的二重积分记为

$$\iint_D f(x, y) dx dy;$$

在 $\mathbb{R}^3$ 中，$f(x, y, z)$ 在 Ω 上的三重积分记为

$$\iiint_{\Omega} f(x, y, z)\, dx\, dy\, dz \quad \text{或} \quad \iiint_{\Omega} f(x, y, z)\, dV;$$

而在 $\mathbb{R}^n$ 中，$f(x_1, x_2, \cdots, x_n)$ 在 Ω 上的 n 重积分记为

$$\int_{\Omega} f(x_1, x_2, \cdots, x_n)\, dx_1\, dx_2 \cdots dx_n \quad \text{或} \quad \iint_{\Omega} \cdots \int f(x_1, x_2, \cdots, x_n)\, dx_1\, dx_2 \cdots dx_n.$$

16.1.3 Peano Curve 皮亚诺曲线

值得注意的是，一条平面曲线所绘出的图形的面积并不一定是 0。皮亚诺发现，存在将实轴上的闭区间映满平面上的一个二维区域的连续映射。也就是说，这条曲线通过该二维区域的每个点。这种曲线被称为皮亚诺曲线。

16.2 Properties and Evaluation of Multiple Integrals 重积分的性质与计算

16.2.1 Properties of Multiple Integrals 重积分的性质

除非特别声明，本节中总假定考虑的区域是 $\mathbb{R}^n(n \geq 2)$ 中的零边界闭区域。

性质 1（线性性） 设 f 和 g 都在区域 Ω 上可积，α, β 为常数，则 $\alpha f + \beta g$ 在 Ω 上也可积，并且

$$\int_\Omega (\alpha f + \beta g) dV = \alpha \int_\Omega f dV + \beta \int_\Omega g dV.$$

性质 2（区域可加性） 设区域 Ω 被分成两个内点不相交的区域 Ω_1 和 Ω_2，如果 f 在 Ω 上可积，则 f 在 Ω_1 和 Ω_2 上都可积；反之，如果 f 在 Ω_1 和 Ω_2 上可积，则 f 也在 Ω 上可积。此时成立

$$\int_\Omega f dV = \int_{\Omega_1} f dV + \int_{\Omega_2} f dV.$$

性质 3 设被积函数 $f = 1$。当 $n = 2$ 时

$$\iint_\Omega dx dy = \iint_\Omega 1 dx dy = \Omega \text{ 的面积};$$

当 $n \geq 3$ 时

$$\int_\Omega dV = \int_\Omega 1 dV = \Omega \text{ 的体积}.$$

性质 4（保序性） 设 f 和 g 都在区域 Ω 上可积，且满足 $f \leq g$，则成立不等式

$$\int_\Omega f dV \leq \int_\Omega g dV.$$

性质 5 设 f 在区域 Ω 上可积，M 与 m 分别为 f 在 Ω 上的上确界和下确界，则成立不等式

$$mV \leq \int_\Omega f dV \leq MV,$$

其中 V 当 $n = 2$ 时为 Ω 的面积，当 $n > 2$ 时为 Ω 的体积。

性质 5 是性质 4 的直接推论。

性质 6（绝对可积性） 设 f 在区域 Ω 上可积，则 $|f|$ 也在 Ω 上可积，且成立不等式

$$\left| \int_\Omega f dV \right| \leq \int_\Omega |f| dV.$$

性质 7（乘积可积性） 设 f 和 g 都在区域 Ω 上可积，则 $f \cdot g$ 也在 Ω 上可积。

性质 8（积分中值定理） 设 f 和 g 都在区域 Ω 上可积，且 g 在 Ω 上不变号。设 M 与 m 分别为 f 在 Ω 上的上确界和下确界，则存在常数 $\mu \in [m, M]$，使得

$$\int_\Omega f \cdot g dV = \mu \int_\Omega g dV.$$

特别地，如果 f 在 Ω 上连续，则存在 $\xi \in \Omega$，使得

$$\int_\Omega f \cdot g dV = f(\xi) \int_\Omega g dV.$$

16.2.2 Evaluation over Rectangular Regions 矩形区域上的重积分计算

定理 16.2 设二元函数 $f(x,y)$ 在闭矩形 $D=[a,b]\times[c,d]$ 上可积。若积分

$$h(x)=\int_c^d f(x,y)\,dy$$

对于每个 $x\in[a,b]$ 存在，则 $h(x)$ 在 $[a,b]$ 上可积，并有等式

$$\iint_D f(x,y)\,dx\,dy=\int_a^b h(x)\,dx=\int_a^b\left(\int_c^d f(x,y)\,dy\right)\,dx=\int_a^b dx\int_c^d f(x,y)\,dy.$$

定理 16.3 设 $f(x_1,x_2,\cdots,x_n)$ 在 n 维闭矩形 $\Omega=[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ 上可积。记 $\Omega_s=[a_2,b_2]\times\cdots\times[a_n,b_n]$。若积分

$$h(x_1)=\int_{\Omega_s} f(x_1,x_2,\cdots,x_n)\,dx_2\cdots dx_n$$

对于每个 $x_1\in[a_1,b_1]$ 存在，则 $h(x_1)$ 在 $[a_1,b_1]$ 上可积，并成立

$$\int_\Omega f(x_1,x_2,\cdots,x_n)\,dx_1\,dx_2\cdots dx_n=\int_{a_1}^{b_1} h(x_1)\,dx_1=\int_{a_1}^{b_1} dx_1\int_{\Omega_s} f(x_1,x_2,\cdots,x_n)\,dx_2\cdots dx_n.$$

17 Integrals with Parameters 含参变量积分

17.1 Proper Integrals with Parameters 含参变量的常义积分

17.1.1 Definition of Proper Integrals with Parameters 含参变量常义积分的定义

设 $f(x,y)$ 是定义在闭矩形 $[a,b]\times[c,d]$ 上的连续函数，于是对于任意固定的 $y\in[c,d]$, $f(x,y)$ 是 $[a,b]$ 上关于 x 的一元连续函数，因此它在 $[a,b]$ 上的积分存在，且积分值 $\int_a^b f(x,y)\,dx$ 由 y 唯一确定。也就是说，

$$I(y)=\int_a^b f(x,y)\,dx,\quad y\in[c,d]$$

确定了一个关于 y 的一元函数。由于式中的 y 可以看成一个参变量，所以称它为含参变量 y 的积分。同理可定义含参变量 x 的积分

$$J(x)=\int_c^d f(x,y)\,dy,\quad x\in[a,b].$$

它们统称**含参变量常义积分**，一般就称为**含参变量积分**。

17.1.2 Analytic Properties of Proper Integrals with Parameters 含参变量常义积分的分析性质

既然含参变量积分是参变量的函数，就应该研究它的分析性质，诸如连续性、可微性和可积性。

定理 17.1 (连续性定理) 设 $f(x,y)$ 在闭矩形 $D=[a,b]\times[c,d]$ 上连续，则函数

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

在 $[c,d]$ 上连续。

即有

$$\lim_{y\to y_0} I(y)=I(y_0),\quad \forall y_0\in[c,d],$$

即

$$\lim_{y\to y_0}\int_a^b f(x,y)\,dx=\int_a^b \lim_{y\to y_0} f(x,y)\,dx,\quad \forall y_0\in[c,d].$$

即极限运算与积分运算可以交换。

定理 17.2 (积分次序交换定理) 若 $f(x,y)$ 在闭矩形 $[a,b]\times[c,d]$ 上连续，则

$$\int_c^d\left[\int_a^b f(x,y)\,dx\right]dy=\int_a^b\left[\int_c^d f(x,y)\,dy\right]dx.$$

定理 17.3 (积分号下求导定理) 设 $f(x,y), f_y(x,y)$ 都在闭矩形 $[a,b]\times[c,d]$ 上连续，则

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

在 $[c,d]$ 上可导，并且在 $[c,d]$ 上成立

$$\frac{dI(y)}{dy}=\int_a^b f_y(x,y)\,dx.$$

这个定理的结论也可写为

$$\frac{d}{dy}\int_a^b f(x,y)\,dx=\int_a^b \frac{\partial}{\partial y}f(x,y)\,dx.$$

这说明求导运算与积分运算可以交换。

定理 17.4 设 $f(x,y), f_y(x,y)$ 都是闭矩形 $[a,b]\times[c,d]$ 上的连续函数，又设 $a(y), b(y)$ 是在 $[c,d]$ 上的可导函数，满足 $a\le a(y)\le b, a\le b(y)\le b$，则函数

$$F(y)=\int_{a(y)}^{b(y)} f(x,y)\,dx$$

在 $[c,d]$ 上可导，并且在 $[c,d]$ 上成立

$$F'(y)=\int_{a(y)}^{b(y)} f_y(x,y)\,dx+f(b(y),y)b'(y)-f(a(y),y)a'(y).$$

例题 17.1 设 $F(y)=\displaystyle\int_0^1 \frac{\sin(xy)}{x+1}\,dx\ (y>0)$，求 $\displaystyle\lim_{y\to 0^+}F(y)$。

由于 $f(x,y)=\dfrac{\sin(xy)}{x+1}$ 在 $[0,1]\times[0,1]$ 上连续，由连续性定理：

$$\lim_{y\to 0^+}F(y)=\int_0^1 \lim_{y\to 0^+}\frac{\sin(xy)}{x+1}\,dx=\int_0^1 0\,dx=0.$$

例题 17.2 计算 $\displaystyle\int_0^1\left[\int_0^1 \frac{x-y}{(x+y)^3}\,dy\right]dx$。(提示：先交换积分次序)

直接计算内层积分较复杂。考虑交换积分次序：

$$\text{原式}=\int_0^1\left[\int_0^1 \frac{x-y}{(x+y)^3}\,dx\right]dy.$$

先计算内层积分（对 x）：

$$\int_0^1 \frac{x-y}{(x+y)^3}\,dx.$$

令 $u=x+y$，则 $x=u-y$，$dx=du$，当 $x=0$ 时 $u=y$，当 $x=1$ 时 $u=1+y$。

$$\int_y^{1+y}\frac{u-2y}{u^3}\,du=\int_y^{1+y}(u^{-2}-2yu^{-3})\,du=\left[-\frac{1}{u}+\frac{y}{u^2}\right]_y^{1+y}.$$

计算得：

$$\left(-\frac{1}{1+y}+\frac{y}{(1+y)^2}\right)-\left(-\frac{1}{y}+\frac{y}{y^2}\right)=\frac{-1}{1+y}+\frac{y}{(1+y)^2}+\frac{1}{y}-1.$$

化简后为 $\dfrac{1}{y(1+y)^2}$。因此：

$$\text{原式}=\int_0^1\frac{1}{y(1+y)^2}\,dy.$$

分部分式分解：$\dfrac{1}{y(1+y)^2}=\dfrac{1}{y}-\dfrac{1}{1+y}-\dfrac{1}{(1+y)^2}$。

$$\int_0^1\left(\frac{1}{y}-\frac{1}{1+y}-\frac{1}{(1+y)^2}\right)dy=\left[\ln\frac{y}{1+y}+\frac{1}{1+y}\right]_0^1.$$

极限计算得：$\ln\frac{1}{2}+\frac{1}{2}-1=\ln\frac{1}{2}-\frac{1}{2}$。

例题 17.3 设 $I(a)=\int_0^1 \frac{\ln(1+ax^2)}{x^2}\,dx\ (a>0)$，求 $I'(a)$。

令 $f(x,a)=\frac{\ln(1+ax^2)}{x^2}$，则

$$\frac{\partial f}{\partial a}=\frac{x^2}{x^2(1+ax^2)}\cdot\frac{1}{x^2}?$$

正确求导：

$$\frac{\partial}{\partial a}\left(\frac{\ln(1+ax^2)}{x^2}\right)=\frac{1}{x^2}\cdot\frac{x^2}{1+ax^2}=\frac{1}{1+ax^2}.$$

$f(x,a)$ 与 $f_a(x,a)$ 在 $x=0$ 处有可去间断点，定义 $f(0,a)=a$，则它们在 $[0,1]\times[\delta,A]$（任意 $0<\delta<A$）连续。由积分号下求导定理：

$$I'(a)=\int_0^1 \frac{1}{1+ax^2}\,dx=\frac{1}{\sqrt{a}}\arctan(\sqrt{a}).$$

例题 17.4 设 $F(t)=\int_0^{t^2} e^{-tx^2}\,dx$，求 $F'(t)$。

这里积分上限 $b(t)=t^2$，下限 $a(t)=0$，被积函数含 t。

$$F'(t)=\int_0^{t^2}\frac{\partial}{\partial t}e^{-tx^2}\,dx+e^{-t(t^2)^2}\cdot 2t-e^{-t\cdot 0}\cdot 0.$$

$$=\int_0^{t^2}(-x^2)e^{-tx^2}\,dx+2te^{-t^5}.$$

第一项积分：令 $u=x\sqrt{t}$，可算得 $-\frac{1}{2t^{3/2}}\int_0^{t^{5/2}} u^2e^{-u^2}\,du$，但保留积分形式也可。

例题 17.5 计算 $\int_0^1 \frac{x^b-x^a}{\ln x}\,dx\ (0<a<b)$。（提示：考虑含参积分 $\int_a^b x^y\,dy$）

考虑含参积分 $\int_a^b x^y\,dy=\left.\frac{x^y}{\ln x}\right|_{y=a}^{y=b}=\frac{x^b-x^a}{\ln x}$。所以：

$$\frac{x^b-x^a}{\ln x}=\int_a^b x^y\,dy.$$

于是：

$$\int_0^1 \frac{x^b-x^a}{\ln x}\,dx=\int_0^1\left[\int_a^b x^y\,dy\right]dx.$$

交换积分次序：

$$=\int_a^b\left[\int_0^1 x^y\,dx\right]dy=\int_a^b\left[\frac{1}{y+1}\right]dy=\ln\frac{b+1}{a+1}.$$

例题 17.6 设 $\varphi(y)=\int_0^{\frac{\pi}{2}}\ln(y^2-\sin^2x)\,dx\ (y>1)$，求 $\varphi'(y)$。

$$\varphi'(y)=\int_0^{\frac{\pi}{2}}\frac{\partial}{\partial y}\ln(y^2-\sin^2x)\,dx=\int_0^{\frac{\pi}{2}}\frac{2y}{y^2-\sin^2x}\,dx.$$

利用公式 $\int_0^{\frac{\pi}{2}}\frac{dx}{y^2-\sin^2x}=\frac{\pi}{2y\sqrt{y^2-1}}\ (y>1)$：

$$\varphi'(y)=2y\cdot\frac{\pi}{2y\sqrt{y^2-1}}=\frac{\pi}{\sqrt{y^2-1}}.$$

17.2 Euler Integrals 欧拉积分

17.2.1 Beta Function 贝塔函数

形如

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

的含参变量积分称为 **Beta 函数**，或**第一类 Euler 积分**。

先看它的定义域。将 Beta 函数写成

$$B(p,q)=\int_0^{\frac{1}{2}} x^{p-1}(1-x)^{q-1}\,dx+\int_{\frac{1}{2}}^1 x^{p-1}(1-x)^{q-1}\,dx,$$

当 $x\to 0$ 时，$x^{p-1}(1-x)^{q-1}\sim x^{p-1}$，所以只有当 $p>0$ 时右边第一个反常积分收敛。而当 $x\to 1$ 时，$x^{p-1}(1-x)^{q-1}\sim(1-x)^{q-1}$，所以只有当 $q>0$ 时右边第二个反常积分收敛。这说明了

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

对于每对 $(p,q)\in(0,+\infty)\times(0,+\infty)$ 收敛，即 Beta 函数 $B(p,q)$ 的定义域为 $(0,+\infty)\times(0,+\infty)$。

1. 连续性： $B(p,q)$ 在 $(0,+\infty)\times(0,+\infty)$ 上连续。

2. 对称性： $B(p,q)=B(q,p)$，$\forall p>0,q>0$。

3. 递推公式：

$$B(p,q)=\frac{q-1}{p+q-1}B(p,q-1),\quad p>0,q>1.$$

由 $B(p,q)$ 的对称性并结合递推公式可得到，当 $p>1,q>1$ 时，成立

$$B(p,q)=\frac{(p-1)(q-1)}{(p+q-1)(p+q-2)}B(p-1,q-1).$$

4. 其他表示：

(1) 作变量代换 $x=\cos^2\varphi$，得到

$$B(p,q)=2\int_0^{\pi/2}\cos^{2p-1}\varphi\sin^{2q-1}\varphi\,d\varphi.$$

据此可以得到

$$B\left(\frac{1}{2},\frac{1}{2}\right)=\pi.$$

(2) 作变量代换 $x=\frac{1}{1+t}$，得到

$$B(p,q)=\int_0^{\infty}\frac{t^{q-1}}{(1+t)^{p+q}}\,dt=\int_0^1\frac{t^{q-1}}{(1+t)^{p+q}}\,dt+\int_1^{\infty}\frac{t^{q-1}}{(1+t)^{p+q}}\,dt.$$

在最后一个积分中再作变量代换 $t=\frac{1}{u}$，得到

$$\int_1^{\infty}\frac{t^{q-1}}{(1+t)^{p+q}}\,dt=\int_0^1\frac{u^{p-1}}{(1+u)^{p+q}}\,du,$$

于是

$$B(p,q)=\int_0^1\frac{t^{p-1}+t^{q-1}}{(1+t)^{p+q}}\,dt\quad(=B(q,p)).$$

- $B\left(\frac{1}{2},\frac{1}{2}\right)=\pi$

- $B(1,1) = 1$
- $B\left(\frac{1}{2}, 1\right) = 2$
- $B\left(\frac{3}{2}, \frac{1}{2}\right) = \frac{\pi}{2}$
- $B\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{2\pi}{\sqrt{3}}$
- $B\left(\frac{1}{4}, \frac{3}{4}\right) = \pi\sqrt{2}$

17.2.2 Gamma Function 伽马函数

形如

$$\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x}\, dx$$

的含参变量积分称为 **Gamma 函数**，或**第二类 Euler 积分**。

先看它的定义域。将 Gamma 函数写成

$$\Gamma(s) = \int_0^1 x^{s-1} e^{-x}\, dx + \int_1^{+\infty} x^{s-1} e^{-x}\, dx,$$

由反常积分的收敛判别法，当 $s \leq 0$ 时，右边第一个反常积分发散，而当 $s > 0$ 时，两个反常积分都收敛，因此 Gamma 函数 $\Gamma(s)$ 的定义域为 $(0, +\infty)$。

1. 连续性与可导性： $\Gamma(s)$ 在 $(0, +\infty)$ 上连续且可导。

2. 递推公式： $\Gamma(s)$ 满足

$$\Gamma(s+1) = s\Gamma(s), \quad s > 0.$$

特别地，当 $s = n$ 为正整数时，

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \cdots = n!\Gamma(1),$$

而 $\Gamma(1) = \int_0^{\infty} e^{-x}\, dx = 1$，所以

$$\Gamma(n+1) = n!.$$

因而 Gamma 函数可以说是阶乘的推广。

由于 $\Gamma(s) = \frac{\Gamma(s+1)}{s}$ 以及 $\Gamma(1) = 1$，所以

$$\lim_{s \to 0^+} \Gamma(s) = +\infty.$$

3. 其他表示：

(1) 在 $\Gamma(s)$ 的表示式中作变量代换 $x = t^2$，那么

$$\Gamma(s) = 2\int_0^{\infty} t^{2s-1} e^{-t^2}\, dt.$$

据此可知

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^{\infty} e^{-t^2}\, dt = \sqrt{\pi}.$$

(2) 作变量代换 $x = \alpha t$ $(\alpha > 0)$ 可得

$$\Gamma(s) = \alpha^s \int_0^{\infty} t^{s-1} e^{-\alpha t}\, dt.$$

4. 定义域的延拓：

由于等式

$$\Gamma(s) = \frac{\Gamma(s+1)}{s}$$

的右边在 $(-1,0)$ 上有意义，则可以应用上式来定义左边函数 $\Gamma(s)$ 在 $(-1,0)$ 上的值。用同样的方法，再利用 $\Gamma(s)$ 已在 $(-1,0)$ 上定义的值，定义 $\Gamma(s)$ 在 $(-2,-1)$ 上的值。如此继续下去，就可以把 $\Gamma(s)$ 的定义域延拓到

$$(-\infty,+\infty) \setminus \{0,-1,-2,-3,\cdots\}.$$

- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- $\Gamma(1) = 1$
- $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}$
- $\Gamma(2) = 1$
- $\Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}$
- $\Gamma(3) = 2$
- $\Gamma(n+1) = n! \quad (n \in \mathbb{N})$

17.2.3 Relation Between Beta and Gamma Functions 贝塔函数与伽马函数的关系

定理 17.5 *Beta* 函数与 *Gamma* 函数之间具有如下关系：

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p>0, \quad q>0.$$

定理 17.6 ***(Legendre 公式)***

$$\Gamma(s)\Gamma\left(s+\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2^{2s-1}}\Gamma(2s), \quad s>0.$$

定理 17.7 ***(余元公式)***

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}, \quad 0<s<1.$$

定理 17.8 ***(Stirling 公式)*** *Gamma* 函数有如下的渐进估计：

$$\Gamma(s+1) = \sqrt{2\pi s}\left(\frac{s}{e}\right)^s e^{\frac{\theta}{12s}}, \quad s>0,$$

这里 $0<\theta<1$。特别地，当 $s=n$ 为正整数时，

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{\theta}{12n}}.$$

例题 17.7 计算下列积分：

$$\int_0^1 x^4(1-x)^3\,dx.$$

$$\int_0^1 x^4(1-x)^3\,dx = B(5,4) = \frac{\Gamma(5)\Gamma(4)}{\Gamma(9)} = \frac{4!\cdot 3!}{8!} = \frac{24\cdot 6}{40320} = \frac{144}{40320} = \frac{1}{280}.$$

例题 17.8 计算下列积分：

$$\int_0^{\frac{\pi}{2}} \sin^5 x \cos^3 x \, dx.$$

利用三角形式 $B(p,q) = 2\int_0^{\pi/2} \sin^{2p-1} x \cos^{2q-1} x \, dx$：

$$\int_0^{\pi/2} \sin^5 x \cos^3 x \, dx = \frac{1}{2} B(3,2) = \frac{1}{2} \cdot \frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} = \frac{1}{2} \cdot \frac{2! \cdot 1!}{4!} = \frac{1}{2} \cdot \frac{2}{24} = \frac{1}{24}.$$

例题 17.9 计算下列积分：

$$\int_0^{\infty} \frac{x^{1/2}}{(1+x)^4} \, dx.$$

利用公式 $B(p,q) = \int_0^{\infty} \frac{t^{p-1}}{(1+t)^{p+q}} \, dt$：

$$\int_0^{\infty} \frac{x^{1/2}}{(1+x)^4} \, dx = B\left(\frac{3}{2}, \frac{5}{2}\right) = \frac{\Gamma(3/2)\Gamma(5/2)}{\Gamma(4)}.$$

计算：

$$\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma(5/2) = \frac{3}{2} \cdot \frac{1}{2}\sqrt{\pi} = \frac{3}{4}\sqrt{\pi}, \quad \Gamma(4) = 3! = 6.$$

所以：

$$B\left(\frac{3}{2}, \frac{5}{2}\right) = \frac{\frac{1}{2}\sqrt{\pi} \cdot \frac{3}{4}\sqrt{\pi}}{6} = \frac{\frac{3}{8}\pi}{6} = \frac{\pi}{16}.$$

例题 17.10 利用 *Gamma* 函数计算：

$$\int_0^{\infty} x^6 e^{-2x} \, dx.$$

利用公式 $\Gamma(s) = \alpha^s \int_0^{\infty} t^{s-1} e^{-\alpha t} \, dt$：

$$\int_0^{\infty} x^6 e^{-2x} \, dx = \frac{\Gamma(7)}{2^7} = \frac{6!}{128} = \frac{720}{128} = \frac{45}{8}.$$

例题 17.11 证明：

$$B\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{2\pi}{\sqrt{3}}.$$

利用余元公式 $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$：

$$B\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{\Gamma(1/3)\Gamma(2/3)}{\Gamma(1)} = \Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right).$$

由余元公式，取 $s = \frac{1}{3}$：

$$\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right) = \frac{\pi}{\sin(\pi/3)} = \frac{\pi}{\sqrt{3}/2} = \frac{2\pi}{\sqrt{3}}.$$

例题 17.12 计算：

$$\Gamma\left(\frac{1}{2}\right), \quad \Gamma\left(\frac{5}{2}\right), \quad \Gamma\left(-\frac{1}{2}\right).$$

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

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi}.$$

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

例题 17.13 计算积分：

$$\int_0^1 \frac{dx}{\sqrt{-\ln x}}.$$

令 $t = -\ln x$，则 $x = e^{-t}$，$dx = -e^{-t}dt$，当 $x = 0$ 时 $t \to \infty$，当 $x = 1$ 时 $t = 0$：

$$\int_0^1 \frac{dx}{\sqrt{-\ln x}} = \int_\infty^0 \frac{-e^{-t}dt}{\sqrt{t}} = \int_0^\infty t^{-1/2}e^{-t}dt = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

例题 17.14 计算积分：

$$\int_0^\infty e^{-x^2}\,dx.$$

利用公式 $\Gamma(s) = 2\int_0^\infty t^{2s-1}e^{-t^2}dt$：取 $s = \frac{1}{2}$：

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty e^{-t^2}dt = \sqrt{\pi} \quad \Rightarrow \quad \int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

18 Fourier Series 傅里叶级数

18.1 Fourier Series Expansion of Functions 函数的傅里叶级数展开

古往今来，从阿基米德（Archimedes）开始的众多大数学家，一直在孜孜不倦地寻找用简单函数较好地近似代替复杂函数的途径——除了理论上的需要之外，它对实际应用领域的意义更是不可估量。但在微积分发明之前，这个问题一直没能获得本质上的突破。

人们最熟悉的简单函数无非两类：幂函数和三角函数。英国数学家泰勒（Taylor）在 18 世纪初找到了用幂函数的（无限）线性组合表示一般函数 $f(x)$ 的方法，即通过泰勒展开将函数化成幂级数形式

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n,$$

经过理论上的完善之后，它很快成为了微分学（乃至整个函数论）的重要工具之一。这方面内容已在前面有关章节中作了介绍。

但是，函数的泰勒展开在应用中有一定的局限性。首先我们在实际问题中总是（也只能）使用泰勒级数的部分和，即 $f(x)$ 的 n 次泰勒多项式

$$P_n(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

来近似地代替函数 $f(x)$，这时候它要求 $f(x)$ 有至少 n 阶的导数，这一条件对许多实际问题来说是过于苛刻的（特别是在发现了许多不可导甚至不连续的重要函数之后）；同时，一般来说泰勒多项式仅在点 x_0 附近与 $f(x)$ 吻合得较为理想，也就是说，它只有局部性质。为此有必要寻找函数的新的级数展开方法。

形如

$$a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

的函数项级数称为**三角级数**，其中 a_0, a_n 和 b_n $(n=1,2,\cdots)$ 为常数。

19 世纪初，法国数学家和工程师傅里叶在研究热传导问题时，找到了在有限区间上用三角级数表示一般函数 $f(x)$ 的方法，即把 $f(x)$ 展开成所谓的**傅里叶级数**。

与泰勒展开相比，傅里叶展开对于 $f(x)$ 的要求要宽容得多，并且它的部分和在整个区间都与 $f(x)$ 吻合得较为理想。因此，傅里叶级数是比泰勒级数更有力、适用性更广的工具，它在声学、光学、热力学、电学等研究领域极有价值，在微分方程求解方面更是起着基本的作用。

本章只介绍有关傅里叶级数的一些基本知识，大致包括三个方面：

如何将一个给定的函数展开为傅里叶级数（称为傅里叶展开）

傅里叶级数的收敛条件

傅里叶级数的性质及某些相关问题

18.1.1 Fourier Expansion of Functions with Period 2π 周期为 2π 的函数的傅里叶展开

下面我们先来讨论这样一个问题：假定 $f(x)$ 可以表示成如下形式的级数

$$f(x)=a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

也就是说假定等式右边的三角级数收敛于 $f(x)$，该如何来确定三角级数中的系数 a_n 和 b_n？为了回答这个问题，我们先来回忆一下向量的正交分解。

在三维空间中，我们如何表示一个向量 V？我们知道，如果有一组标准正交基 $\{V_1, V_2, V_3\}$（即 $\|V_i\| = 1$ 且 $V_i \cdot V_j = 0$），那么向量 V 可以唯一地表示为：

$$V = a_1 V_1 + a_2 V_2 + a_3 V_3$$

其中系数 a_1, a_2, a_3 就是向量 V 在每个基向量方向上的“坐标”或“权重”。

求解坐标 a_1 我们提供两种视角。从几何直觉视角来看，基向量方向上的权重正是向量 V 在其上的正交投影，自然想到与投影挂钩的内积运算。V 与 V_1 的内积代表 V_1 上的正交投影长度乘以 V_1 本身的长度，而 V_1 作为一个标准正交基向量是单位长度，于是 V 与 V_1 的内积直接就得到了 V 上的正交投影长度。注意，投影长度其实还不直接等于那个维度的坐标，需要除以基向量的长度才得到坐标，但由于这里基向量长度就是 1 所以内积结果直接就等于坐标。因此，我们可以简单总结：对于一个单位正交基，求坐标就是求内积。

$$a_1 = V \cdot V_1$$

请注意这里“单位”与“正交”存在的必要性！若给到的一组基是互相正交但不是单位长度的，则

$$V \cdot V_1 = \|V_1\| \times \|V\| \cos\theta$$

$$\text{投影长度} = \|V\| \cos\theta = \frac{V \cdot V_1}{\|V_1\|}$$

$$a_1 = \frac{\text{投影长度}}{\|V_1\|} = \frac{V \cdot V_1}{\|V_1\|^2} = \frac{V \cdot V_1}{V_1 \cdot V_1}$$

若给到的一组基不是正交基，此时我们无法再使用内积来计算投影长度继而简便地求出坐标，因为内积只得到正交投影，我们需要更复杂的代数方法。

下面我们从纯代数计算的视角把以上所有重新推一遍：

$$V \cdot V_1 = a_1 V_1 \cdot V_1 + a_2 V_2 \cdot V_1 + a_3 V_3 \cdot V_1 = a_1 V_1 \cdot V_1$$

$$a_1 = \frac{V \cdot V_1}{V_1 \cdot V_1}$$

若 V_1 为单位基向量，则

$$a_1 = V \cdot V_1$$

若基向量既不是单位向量也不正交，我们可以通过线性方程组求解：

$$V \cdot V_1 = a_1 V_1 \cdot V_1 + a_2 V_2 \cdot V_1 + a_3 V_3 \cdot V_1$$

$$V \cdot V_2 = a_1 V_1 \cdot V_2 + a_2 V_2 \cdot V_2 + a_3 V_3 \cdot V_2$$

$$V \cdot V_3 = a_1 V_1 \cdot V_3 + a_2 V_2 \cdot V_3 + a_3 V_3 \cdot V_3$$

当然上述过程可以推广到 n 维。

下面我们把上述思想从有限维向量空间推广到无限维函数空间。我们把一个周期为 2π 的函数 $f(t)$ 看作一个“向量”,“基向量”在这里是一系列不同频率的正弦和余弦函数：$\{1, \cos t, \sin t, \cos 2t, \sin 2t, ...\}$。在函数空间中，我们定义两个函数 $f(t)$ 和 $g(t)$ 在区间 $[a, b]$ 上的内积为它们乘积的积分

$$\langle f, g \rangle = \int_a^b f(t) g(t) dt$$

对于复函数，需要取其中一个的共轭，但我们现在讨论的是实函数三角函数，所以暂时忽略共轭。

基函数找到了，是否是正交单位基呢？三角函数系是具有正交性的。很容易验证，在区间 $[-\pi,\pi]$（或任何长度为 2π 的区间）上，这些基函数是两两正交的：

$$\int_{-\pi}^{\pi}\sin(mt)\cos(nt)\,dt=0\quad(\forall m,n)$$

$$\int_{-\pi}^{\pi}\sin(mt)\sin(nt)\,dt=0\quad(m\neq n)$$

$$\int_{-\pi}^{\pi}\cos(mt)\cos(nt)\,dt=0\quad(m\neq n)$$

满足正交基我们就可以用内积来求坐标，但我们先看一下这些基是不是单位基。计算基函数的范数（norm），可以发现这些三角基函数并不是单位基：

$$(1,1)=\int_{-\pi}^{\pi}1\,dt=2\pi\Rightarrow||1||=\sqrt{2\pi}$$

$$(\cos nt,\cos nt)=\int_{-\pi}^{\pi}\cos(nt)\cos(nt)\,dt=\pi\Rightarrow||\cos(nt)||=\sqrt{\pi}$$

$$(\sin nt,\sin nt)=\int_{-\pi}^{\pi}\sin(nt)\sin(nt)\,dt=\pi\Rightarrow||\sin(nt)||=\sqrt{\pi}$$

所以，在求投影坐标时，我们不能只做内积，还必须除以基函数范数的平方。

下面我们开始正式推导系数的公式，我们想将周期为 2π 的函数 $f(t)$ 分解为：

$$f(t)=a_0+\sum_{n=1}^{\infty}[a_n\cos(nt)+b_n\sin(nt)]$$

根据刚才总结的内积公式：

$$a_0=\frac{\langle f(t),1\rangle}{\langle 1,1\rangle}=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\,dt$$

$$a_n=\frac{\langle f(t),\cos(nt)\rangle}{\langle\cos(nt),\cos(nt)\rangle}=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\cos(nt)\,dt,\quad n=1,2,\cdots$$

$$b_n=\frac{\langle f(t),\sin(nt)\rangle}{\langle\sin(nt),\sin(nt)\rangle}=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\sin(nt)\,dt,\quad n=1,2,\cdots$$

如果想把 a_0 的表达式统一进 a_n 中可以把三角级数的常数项写成 $\frac{a_0}{2}$：

$$f(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}[a_n\cos(nt)+b_n\sin(nt)]$$

这样我们就有如下表达式：

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,dx,\quad n=0,1,2,\cdots.$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,dx,\quad n=1,2,\cdots.$$

上面两式称为 **欧拉-傅里叶公式**（Euler-Fourier Formulae）。

设周期为 2π 的函数 $f(x)$ 在 $[-\pi,\pi]$ 上可积或绝对可积，则利用欧拉-傅里叶公式就可求出系数 a_n,b_n，并记

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

右端的三角级数称为 $f(x)$ 的 **傅里叶级数**，相应的 a_n 和 b_n 称为 $f(x)$ 的 **傅里叶系数**。

要特别指出的是，目前在 $f(x)$ 和它的傅里叶级数之间不能用等号而只能用“$\sim$”，因为我们不知道右端的三角级数是否收敛；即使收敛，也不知道它是否收敛到 $f(x)$ 本身。这些问题我们将在下一节讨论。

18.1.2 Sine Series and Cosine Series 正弦级数和余弦级数

由定积分的性质，若 $f(x)$ 是奇函数，那么显然有 $a_n = 0$，而

$$b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx, \quad n = 1,2,\cdots,$$

这时，相应的傅里叶级数为

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin nx.$$

形如 $\sum\limits_{n=1}^{\infty} b_n \sin nx$ 的三角级数称为**正弦级数**。

同样，若 $f(x)$ 是偶函数，那么有 $b_n = 0$ 和

$$a_n = \frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\,dx, \quad n = 1,2,\cdots.$$

相应的 Fourier 级数为

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx.$$

形如 $\frac{a_0}{2} + \sum\limits_{n=1}^{\infty} a_n \cos nx$ 的三角级数称为**余弦级数**。

反过来，在实际问题中，由于某种特殊的用途，也经常需要将一个函数展开成正弦级数或余弦级数。

18.1.3 Fourier Expansion of Functions with Arbitrary Period 任意周期的函数的傅里叶展开

如果 $f(x)$ 的周期为 $2T$，作变换 $x = \frac{T}{\pi}t$，则

$$\varphi(t) = f\left(\frac{T}{\pi}t\right) = f(x)$$

是定义在 $(-\infty,+\infty)$ 上的周期为 2π 的函数。利用前面的结果，有

$$\varphi(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt),$$

代回变量，就有

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos \frac{n\pi}{T}x + b_n \sin \frac{n\pi}{T}x\right).$$

相应的傅里叶系数为

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(t)\cos nt\,dt = \frac{1}{T}\int_{-T}^{T} f(x)\cos\frac{n\pi}{T}x\,dx, \quad n = 0, 1, 2, \cdots,$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \varphi(t)\sin nt\,dt = \frac{1}{T}\int_{-T}^{T} f(x)\sin\frac{n\pi}{T}x\,dx, \quad n = 1, 2, \cdots.$$

18.1.4 Frequency Domain Plot of Periodic Functions 周期函数的频域图

我们已经成功将周期函数分解成了若干正弦函数和余弦函数，现在我们要从纯数学的分析视角转换到一个更贴合实际应用的物理学视角，把周期函数看作波信号，来理解傅里叶级数在做一件什么事。于是我们现在把自变量 x 统一换成时间 t：

$$f(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{T}t + b_n\sin\frac{n\pi}{T}t\right).$$

我们首先明确几个有关波（Wave）的基本概念：

1. **周期**（Period）：周期指每隔最少多久时间波信号会重复，上文我们假设周期为 $2T$。
2. **频率**（Frequency）：我们需要区分两种频率，这里拿 $n = 1$ 时的基波举例，非基波的频率只需乘以相应的 n：$\omega_n = n\omega_0$。

   （a）**基波频率**（f_0）：这是信号完成一次完整循环所需时间的倒数

   $$f_0 = \frac{1}{2T}$$

   它的单位是“周/单位时间”（比如“赫兹（Hz）”，如果 t 的单位是秒）

   （b）**基波角频率**（ω_0）：在数学和信号处理中，我们经常使用角频率（单位是弧度/单位时间），因为它与正弦和余弦函数配合使用更方便。角频率和普通频率的关系是：

   $$\omega_0 = 2\pi \times f_0$$

   根据式子

   $$f(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{T}t + b_n\sin\frac{n\pi}{T}t\right).$$

   基波角频率 $\omega_0 = \frac{\pi}{T}$，可以验证以下关系：

   $$2\pi \times f_0 = 2\pi \times \frac{1}{2T} = \frac{\pi}{T} = \omega_0$$

3. **基波**（Fundamental Wave）与**谐波**（Harmonic Wave）：“基础频率”的波称为基波，在这里就是 $n = 1$ 的项：$a_1\cos(\omega_0 t) + b_1\sin(\omega_0 t)$，其频率就是基波频率。谐波，通俗来讲，就是一个复杂波动中，频率是某个基础频率的整数倍的“简单成分波”。$n = 2$ 的项（$a_2\cos(2\omega_0 t) + b_2\sin(2\omega_0 t)$）被称为二次谐波，其频率是基波频率的 2 倍。$n = 3$ 的项是三次谐波，以此类推。
4. **振幅**（Amplitude）：振幅指的是波振动时，离开平衡位置的最大距离。它衡量了波的强度或能量的大小。

5. **相位**（Phase）：相位描述了在一个周期内，波的起始点或当前所处的阶段位置。它决定了波的波形在时间轴上的水平移动。通常用希腊字母 ϕ 表示相位，单位一般是弧度。考虑以下两个波：

$$y_1(t) = A\sin(\omega t)$$

$$y_2(t) = A\sin(\omega t + \phi)$$

如果 $\phi > 0$，我们称波 $y_2(t)$ 的相位超前于 $y_1(t)$；如果 $\phi < 0$，我们称波 $y_2(t)$ 的相位滞后于 $y_1(t)$。

相位决定了当多个波叠加时，它们是相互增强还是相互抵消。如果两个波的波峰和波谷同时出现，我们称它们是**同相位**，它们叠加后振幅会变大（**建设性干涉**）；如果一个波的波峰遇到另一个波的波谷，我们称它们是**反相位**，它们叠加后振幅会减小，甚至完全抵消（**破坏性干涉**）。

下面回到公式

$$f(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{T}t + b_n\sin\frac{n\pi}{T}t\right).$$

一个更简洁、更物理的写法是将同一个频率 n 的 cos 项和 sin 项合并成一个单一的正弦波（或余弦波）：

$$a_n\cos(\omega_n t) + b_n\sin(\omega_n t) = A_n\sin(\omega_n t + \phi_n)$$

其中 A_n 是第 n 次谐波的合成振幅：

$$A_n = \sqrt{a_n^2 + b_n^2}$$

ϕ_n 是第 n 次谐波的合成相位：

$$\phi_n = -\arctan\left(\frac{b_n}{a_n}\right)$$

它告诉你这个频率成分的波形相对于一个纯余弦波的起始位置。

注意，"**任意相位的两个同频率的正弦和余弦波相加，结果一定是同频率的某个相位的正弦波（或余弦波）**"这个结论可能来得没有那么直接，你可能会担心相加后即便频率相同，波形是否会改变（例如出现两个凸起）。在没有严格证明之前，这样的考虑和担心是合理和严谨的，下面我们证明一下这个结论：

设两个同频率 ω 的波形：

$$A\cos(\omega t) + B\sin(\omega t)$$

我们想把它合并为单一正弦函数形式：

$$R\sin(\omega t + \phi)$$

利用三角恒等式：

$$R\sin(\omega t + \phi) = R\sin(\omega t)\cos\phi + R\cos(\omega t)\sin\phi$$

所以：

$$A = R\sin\phi, B = R\cos\phi$$

6. **直流分量**（DC Component）：$\frac{a_0}{2}$ 是直流分量，代表信号的平均值，可以看作是"频率为 0"的成分。

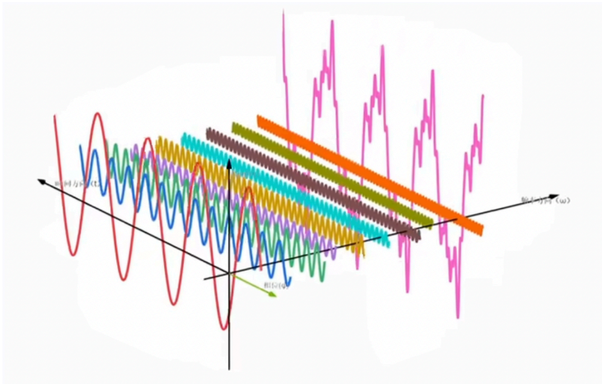


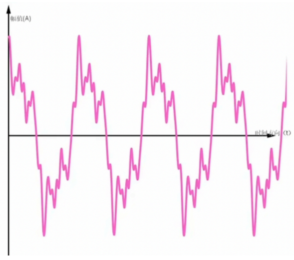
图 2: 周期函数通过傅里叶级数从时域到频域的转换

下面我们把波信号放进空间和坐标系中观察。原始波信号是一个横轴为时间，纵轴为信号强度、振幅或者大小的图像，这个 $f(t)$ 和 t 的平面叫做**时域**。时域是我们最直观的观察世界的方式。在这个视角下，信号被表示为随着时间变化的幅度。时域图的重点是观察信号在时间线上的演变、模式和瞬态特征。

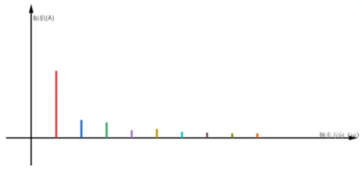
通过傅里叶级数的分解，原始波被分解成若干条不同频率不同相位的正弦波，于是我们可以在原先的二维坐标系中添加第三条轴：（角）频率 ω，把这些正弦波按频率高低排列在频率轴上。0 点处会有一条水平的信号，这是频率为 0 的直流分量，代表信号的平均值。之后在频率为 $\omega_0, 2\omega_0, 3\omega_0, ..., n\omega_0, ...$ 处每个点（按照我们最初傅里叶级数的形式说）其实背后都是两条波，一条 $b_n \sin(\omega_n t)$ 一条 $a_n \cos(\omega_n t)$，它们频率相同，振幅分别由 b_n 和 a_n 决定，相位不同（把标准余弦波看作相位为 $\frac{\pi}{2}$ 的标准正弦波的话，那么 $b_n \sin(\omega_n t)(b_n > 0)$ 的相位为 0，$a_n \cos(\omega_n t)(a_n > 0)$ 的相位为 $\frac{\pi}{2}$，$b_n \sin(\omega_n t)(b_n < 0)$ 的相位为 π，$a_n \cos(\omega_n t)(a_n < 0)$ 的相位为 $\frac{3\pi}{2}$），但我们说过可以把它们相加，得到的谐波仍然是一条正弦波或者余弦波（这里我们选用正弦波），并且拥有相同的频率，它的振幅和相位由 a_n 和 b_n 共同决定。这样我们在每个频率上就都有一条正弦波，它们已经按频率高低排好，但拥有不同的振幅和相位。

于是如果我们转换视角的话就可以画出 A 和 ω 的图，注意在级数形式下这个图是离散的。当然这个图不反映相位的信息，仅知道谐波的频率和振幅还不能将它们复原回原始波信号。我们可以再次转换视角画出 ϕ 和 ω 的图，这个图自然也是离散的。我们称这种横轴为频率的平面为**频域**，其纵轴可以是振幅也可以是相位，前者的图像称为**振幅谱**，后者称为**相位谱**。频域是一个分析的视角，它不关心信号在时间上具体怎么变化，而是关心这个信号内部包含哪些不同频率的正弦波，以及各自的振幅和相位是多少。

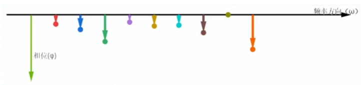
傅里叶级数可以把周期性的波从连续的时域信号转换为离散的频域信号。它给到我们三组信息：分解后的正弦波的频率、振幅以及相位，根据这些信息我们也可以从若干正弦波还原回原始波。由于这个过程是互逆的，傅里叶级数并没有创造本质上新的信息，但是变换后的频域图往往是我们从技术应用角度而言更需要的。频域视角将隐藏在时域中的周期性、谐波特性、噪声成分等清晰地暴露出来，使得滤波、压缩、识别、分析等操作变得异常直观和高效。后续我们会介绍针对非周期性函数的“傅里叶变换”以及它的一些实际应用。在学习了傅里叶级数的复指数形式以及傅里叶变换后，我们会对傅里叶变换“坐标变换”的本质以及它只是将信息从时域“旋转”到了频域来观察这个事实有更深刻的理解。



(a) 时域图



(b) 频域图：振幅谱



(c) 频域图：相位谱

图 3: 时域信号和频域信号

18.2 Convergence Tests for Fourier Series 傅里叶级数的收敛判别法

本节内容相对复杂，超出了本笔记的讨论范畴，感兴趣的可参考陈纪修老师的《数学分析》。但值得一提的是，至今数学家仍未找到一个傅里叶级数收敛的充分必要条件。

18.3 Properties of Fourier Series 傅里叶级数的性质

18.3.1 Analytic Properties of Fourier Series 傅里叶级数的分析性质

为简单起见，假定 $f(x)$ 的周期为 2π。

定理 18.1 *设 $f(x)$ 在 $[-\pi,\pi]$ 上可积或绝对可积，则对于 $f(x)$ 的 Fourier 系数 a_n 与 b_n，有*

$$\lim_{n\to\infty} a_n = 0, \quad \lim_{n\to\infty} b_n = 0.$$

对于函数的 Fourier 级数表示，有必要讨论它的逐项微分和逐项积分问题。关于逐项积分，Fourier 级数有非常好的性质。

定理 18.2 (Fourier 级数的逐项积分定理) *设 $f(x)$ 在 $[-\pi,\pi]$ 上可积或绝对可积，*

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

则 $f(x)$ 的 Fourier 级数可以逐项积分，即对于任意 $c, x \in [-\pi,\pi]$，

$$\int_c^x f(t)\,dt = \int_c^x \frac{a_0}{2}\,dt + \sum_{n=1}^{\infty}\int_c^x (a_n \cos nt + b_n \sin nt)\,dt.$$

定理 18.3 (Fourier 级数的逐项微分定理) *设 $f(x)$ 在 $[-\pi,\pi]$ 上连续，*

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

*$f(-\pi) = f(\pi)$，且除了有限个点外 $f(x)$ 可导。进一步假设 $f'(x)$ 在 $[-\pi,\pi]$ 上可积或绝对可积（**注意：** $f'(x)$ 在有限个点可能无定义，但这并不影响其可积性）。则 $f'(x)$ 的 Fourier 级数可由 $f(x)$ 的 Fourier 级数逐项微分得到，即*

$$f'(x) \sim \frac{d}{dx}\left(\frac{a_0}{2}\right) + \sum_{n=1}^{\infty}\frac{d}{dx}(a_n \cos nx + b_n \sin nx) = \sum_{n=1}^{\infty}(-a_n n \sin nx + b_n n \cos nx).$$

18.4 Fourier Transform (FT) and Fourier Integral 傅里叶变换和傅里叶积分

18.4.1 Fourier Transform and Inverse Fourier Transform 傅里叶变换及其逆变换

以上关于傅里叶级数的论述都是对周期函数而言的，那么对于不具备周期性的函数，又该如何处理呢?

在 $(-\infty,+\infty)$ 上可积的非周期函数 $f(x)$ 可以看成是周期函数的极限情况，处理思路是这样的:

(1) 先取 $f(x)$ 在 $[-T,T]$ 上的部分（即把它视为仅定义在 $[-T,T]$ 上的函数)，再以 $2T$ 为周期，将它延拓为 $(-\infty,+\infty)$ 上的周期函数 $f_T(x)$;

(2) 对得到的周期函数 $f_T(x)$ 作傅里叶展开;

(3) 令 T 趋于无穷大。

下面来导出具体过程。将欧拉公式

$$\cos\theta=\frac{e^{i\theta}+e^{-i\theta}}{2},\quad \sin\theta=\frac{e^{i\theta}-e^{-i\theta}}{2i}=-\frac{i}{2}(e^{i\theta}-e^{-i\theta})$$

代入周期为 $2T$ 的函数 $f_T(x)$ 的傅里叶级数,记 $\frac{\pi}{T}$ 是基波角频率(下面就简称为频率),$\omega_n=\frac{n\pi}{T}$，得到

$$f_T(x)\sim\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos\omega_nx+b_n\sin\omega_nx)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(\frac{a_n-ib_n}{2}e^{i\omega_nx}+\frac{a_n+ib_n}{2}e^{-i\omega_nx}\right).$$

记

$$\begin{gathered}c_0=a_0,\\ c_n=a_n-ib_n=\frac{1}{T}\int_{-T}^{T}f_T(t)e^{-i\omega_nt}\,dt\quad(n=1,2,\cdots),\\ c_{-n}=a_n+ib_n=\overline{c_n},\end{gathered}$$

则得到

$$f_T(x)\sim\frac{c_0}{2}+\frac{1}{2}\sum_{n=1}^{\infty}(c_ne^{i\omega_nx}+c_{-n}e^{-i\omega_nx})=\frac{1}{2}\sum_{n=-\infty}^{\infty}c_ne^{i\omega_nx},$$

这称为**傅里叶级数的复数形式**。将 c_n 的表达式代入，即有

$$f_T(x)\sim\frac{1}{2T}\sum_{n=-\infty}^{\infty}\left[\int_{-T}^{T}f_T(t)e^{-i\omega_nt}\,dt\right]e^{i\omega_nx}.$$

记 $\Delta\omega=\omega_n-\omega_{n-1}=\frac{\pi}{T}$，于是当 $T\to+\infty$ 时 $\Delta\omega\to0$，即得到

$$f(x)=\lim_{T\to+\infty}f_T(x)\sim\lim_{\Delta\omega\to0}\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}\left[\int_{-T}^{T}f_T(t)e^{-i\omega_nt}\,dt\right]e^{i\omega_nx}\Delta\omega.$$

记 $\varphi_T(\omega)=\frac{1}{2\pi}\int_{-T}^{T}f_T(t)e^{-i\omega t}\,dt\,e^{i\omega x}$，则上式可写成

$$f(x)\sim\lim_{\Delta\omega\to0}\sum_{n=-\infty}^{\infty}\varphi_T(\omega_n)\Delta\omega,$$

它看上去很像 Riemann 和的极限形式，不过由于 $\Delta\omega\to0$ 时函数 $\varphi_T(\omega)$ 将随之趋于

$$\varphi(\omega)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}f(t)e^{-i\omega t}\,dt$$

因此这并非真正的 Riemann 和。但是，我们暂且不理会这些，就将它看成 $\varphi(\omega)$ 在 $(-\infty,+\infty)$ 上的“积分”，于是（形式上）有

$$f(x) \sim \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(t)e^{-i\omega t}\,dt\right]e^{i\omega x}\,d\omega.$$

我们称方括号中的函数

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x)e^{-i\omega x}\,dx \quad (\omega \in (-\infty,+\infty))$$

为 f 的**傅里叶变换**（或像函数），记为 $F[f]$，即

$$F[f](\omega) = \hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x)e^{-i\omega x}\,dx,$$

而称函数

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)e^{i\omega x}\,d\omega \quad (x \in (-\infty,+\infty))$$

为 f 的**傅里叶逆变换**（或像原函数），记为 $F^{-1}[f]$，即

$$F^{-1}[f](x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\hat{f}(\omega)e^{i\omega x}\,d\omega.$$

注意这里假设了像函数与像原函数的存在性。

我们称函数

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} f(t)e^{-i\omega t}\,dt\right]e^{i\omega x}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty} f(t)e^{i\omega(x-t)}\,dt$$

为 f 的**傅里叶积分**。容易想到，在一定条件下，它应与 $f(x)$ 相等，但研究这些条件已超出本课程的要求，我们不加证明地给出以下充分条件。

定理 18.4 *设函数 f 在 $(-\infty,+\infty)$ 上绝对可积，且在 $(-\infty,+\infty)$ 中的任何闭区间上分段可导。则 f 的 Fourier 积分满足：对于任意 $x \in (-\infty,+\infty)$ 成立*

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty} f(t)e^{i\omega(x-t)}\,dt = \frac{f(x+)+f(x-)}{2}.$$

定义 18.1 *设函数 f 在 $[a,b]$ 上除有限个点*

$$a = x_0 < x_1 < x_2 < \cdots < x_N = b$$

外均可导，而在 $x_i(i = 0,1,2,\cdots,N)$ 处 f 的左右极限 $f(x_i-)$ 和 $f(x_i+)$ 都存在（在 $x_0 = a$ 只要求右极限存在，在 $x_N = b$ 只要求左极限存在），并且极限

$$\lim_{h\to 0^-}\frac{f(x_i+h)-f(x_i-)}{h}$$

和

$$\lim_{h\to 0^+}\frac{f(x_i+h)-f(x_i+)}{h}$$

都存在（在 $x_0 = a$ 只要求上述第二个极限存在，在 $x_N = b$ 只要求上述第一个极限存在），那么称 f 在 $[a,b]$ 上分段可导。

18.4.2 Properties of the Fourier Transform 傅里叶变换的性质

定理 18.5 (线性性质) 设 c_1, c_2 是常数。若 f, g 的 *Fourier* 变换存在，则

$$F[c_1 f + c_2 g] = c_1 F[f] + c_2 F[g];$$

若 $\hat{f} = F[f], \hat{g} = F[g]$ 的 *Fourier* 逆变换存在，则

$$F^{-1}[c_1 \hat{f} + c_2 \hat{g}] = c_1 F^{-1}[\hat{f}] + c_2 F^{-1}[\hat{g}].$$

定理 18.6 (位移性质) 若函数 f 的 *Fourier* 变换存在，则

$$F[f(x \pm x_0)](\omega) = F[f](\omega) e^{\pm i\omega x_0};$$

若 $\hat{f} = F[f]$ 的 *Fourier* 逆变换存在，则

$$F^{-1}[\hat{f}(\omega \pm \omega_0)](x) = F^{-1}[\hat{f}](x) e^{\mp i\omega_0 x}.$$

注：以上两式常简记为

$$F[f(x \pm x_0)] = F[f] e^{\pm i\omega x_0}, \quad F^{-1}[\hat{f}(\omega \pm \omega_0)] = F^{-1}[\hat{f}] e^{\mp i\omega_0 x}.$$

今后类似的情况也用此种记号，而不再一一明确指出变换的函数取值。

定理 18.7 (时间尺度性)

$$F[f(ax)] = \frac{1}{|a|} \hat{f}\left(\frac{\omega}{a}\right).$$

定理 18.8 (频率尺度性)

$$F\left[\frac{1}{a} f\left(\frac{x}{a}\right)\right] = \hat{f}(a\omega).$$

定理 18.9 (微分性质) *1.* 设函数 $f(x)$ 在 $(-\infty, +\infty)$ 上有连续的导数，且 $f(x)$ 与 $f'(x)$ 在 $(-\infty, +\infty)$ 上绝对可积，若 $\lim_{x\to\infty} f(x) = 0$，则有

$$F[f'] = i\omega \cdot F[f].$$

2. 若 $f(x)$ 和 $xf(x)$ 在 $(-\infty, +\infty)$ 上绝对可积，则

$$F[-ix \cdot f] = (F[f])'.$$

定理 18.10 (积分性质) 设函数 $f(x)$ 和 $\int_{-\infty}^{x} f(t)\,dt$ 在 $(-\infty, +\infty)$ 上绝对可积，则

$$F\left[\int_{-\infty}^{x} f(t)\,dt\right] = \frac{1}{i\omega} F[f].$$

18.4.3 Convolution 卷积

现在引入卷积的概念。

定义 18.2 (卷积) 设函数 f 和 g 在 $(-\infty, +\infty)$ 上定义，且积分

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(t) g(x - t)\,dt$$

存在，则称函数 $f * g$ 为 f 和 g 的卷积。

显然，卷积具有对称性，即 $f*g=g*f$。

建立以下两个定理需要更广泛意义下的积分理论，但由于其重要性，我们仍写出其结论。

定理 18.11 (卷积的 Fourier 变换) 设函数 f 和 g 在 $(-\infty,+\infty)$ 上绝对可积，则有

$$F[f*g]=F[f]\cdot F[g].$$

定理 18.12 (Parseval 等式) 设函数 f 在 $(-\infty,+\infty)$ 上绝对可积，且

$$\int_{-\infty}^{+\infty}[f(x)]^2\,dx$$

收敛。记 f 的 *Fourier* 变换为 $\hat{f}$，则

$$\int_{-\infty}^{+\infty}[f(x)]^2\,dx=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|\hat{f}(\omega)|^2\,d\omega.$$

18.5 Fast Fourier Transform (FFT) 快速傅里叶变换

18.5.1 Discrete Fourier Transform (DFT) 离散傅里叶变换

人们刚开始利用无线电技术传输信号时，是将连续信号进行某种调制处理后直接传送的，本质上传送的还是连续信号 (也叫模拟信号)。这样的传输方式抗干扰能力差，失真严重，尤其是经过长距离传送或多级传递后，信号可能面目全非，质量自然难尽人意。

后来发展了离散的传输方法，它不是传送连续信号本身，而是每隔一段时间 Δt，从信号中提取一个数值脉冲（称为数值抽样），将连续信号转化成数据序列 $x(0),x(1),x(2),\cdots,x(N-1)$，再经编码后发送。只要抽取的时间间隔足够小，这列数据就能很好地反映原信号，接收方通过逆向处理，可以复原出所传递的信号。这种方法称为数字信号传输，具有抗干扰能力强、信号还原质量高、易于加密和解密等优点，问世后便受到广泛的重视，至今方兴未艾。

可以想见的是，为了保证接收的质量，Δt 必须取得很小，即 N 非常之大。因此，直接发送这列数据将会长时间地占用传输设备和线路，这不但需要支付昂贵的费用，在情况紧急时甚至会误事。

所以，在抽样之后需要对数据序列 $x(0),x(1),\cdots,x(N-1)$ 进行简化和压缩，但由于序列中数据的大小是散乱的，因此一方面我们不能随意舍弃某些数据，另一方面压缩的效果也比较差。

后来经研究发现，若对数据序列 $x(0),x(1),\cdots,x(N-1)$ 施以如下的离散 Fourier 变换

$$X(j)=\sum_{n=0}^{N-1}x(n)e^{-2\pi i\frac{nj}{N}}\quad(j=0,1,2,\cdots,N-1,\ i=\sqrt{-1})$$

就可以有效地解决上面的问题。(之所以称它为“离散 Fourier 变换”，在于它可以看成是 Fourier 变换 $\hat{f}(\omega)=\int_{-\infty}^{\infty}f(x)e^{-i\omega x}dx$ 的一种离散的近似形式的推广。)

利用正交关系式

$$\frac{1}{N}\sum_{n=0}^{N-1}e^{-2\pi i\frac{nj}{N}}e^{2\pi i\frac{nk}{N}}=\delta_{j,k}=\begin{cases}1, & j=k,\\ 0, & j\neq k\end{cases}$$

，可以导出离散 Fourier 逆变换

$$x(k)=\frac{1}{N}\sum_{j=0}^{N-1}X(j)e^{2\pi i\frac{kj}{N}},\quad k=0,1,2,\cdots,N-1,$$

这是因为

$$\begin{aligned}\frac{1}{N}\sum_{j=0}^{N-1}X(j)e^{2\pi i\frac{kj}{N}} &= \frac{1}{N}\sum_{j=0}^{N-1}\sum_{n=0}^{N-1}x(n)e^{-2\pi i\frac{nj}{N}}e^{2\pi i\frac{kj}{N}}\\ &= \sum_{n=0}^{N-1}x(n)\left[\frac{1}{N}\sum_{j=0}^{N-1}e^{-2\pi i\frac{nj}{N}}e^{2\pi i\frac{kj}{N}}\right]\\ &= \sum_{n=0}^{N-1}x(n)\delta_{n,k} = x(k).\end{aligned}$$

也就是说，若发送方将 $x(0), x(1), \cdots, x(N-1)$ 作了离散 Fourier 变换后传输出去，接收方可以对收到的数据进行离散 Fourier 逆变换，再现原始信号。

从表面看来，这么做似乎毫无必要，因为变换后的数据长度仍是 N，并没有缩短，况且还要额外支出两次变换的代价。其实不然。

从变换公式容易看出，变换后的序列中的每个 $X(j)$，都包含了原序列中所有信号的信息。因此，即使丢失了某些 $X(j)$，仍可望由其余数据基本正确地还原出原始数据。这当然使得传输过程的抗干扰能力进一步提高，但更重要的是，这可以让我们通过有意剔除某些模较小的数据（通常这类数据数量很大）而使需传输的序列大为缩短。此外，$X(0), X(1), \cdots, X(N-1)$ 的排列将很有规律，模较大的数据往往集中在序列中一两个较窄的范围内，易于作高效的压缩处理。

18.5.2 Fast Fourier Transform (FFT) 快速傅里叶变换

尽管早就发现离散 Fourier 变换有如此诱人的好处，但在一个相当长的时期中，人们对它基本上只限于纸上谈兵。这是因为，做一次变换需要进行 N^2 次复数乘法和 $N(N-1)$ 次复数加法，实际使用中的 N 总是极为巨大的，相应的高昂代价令人望而却步。

一直到 20 世纪 60 年代中期，Cooley 和 Tukey 发现了计算离散 Fourier 变换的高效（同时又特别适合于计算机硬件操作）的方法——快速 Fourier 变换（简称 FFT——Fast Fourier Transform）之后，它才真正获得了生命力。可以毫不夸张地说，基于 FFT 的离散 Fourier 变换技术，是当今信息传输、频谱分析、图像处理、数据压缩等领域中最重要的数学工具之一。目前，国际上任何一个综合的数学软件中，必定含有 FFT 的计算程序。

18.6 Practical Applications of Fourier Transform 傅里叶变换的实际应用

18.6.1 Sound Processing 声音处理

一段嘈杂的音频通过傅里叶变换可以分离出频域图，把低频部分的波傅里叶逆变换回去可以得到男人说话的音频，高频部分可以得到女人说话的音频，而超高频部分则为噪声。这样我们就实现了分离音频，这可以帮助我们音频降噪、给一首歌或一个视频消去人声、给一个歌手的演唱提取干声、制作变声器等等。

18.6.2 Image Processing 图像处理

低频部分表示图像的大致轮廓，高频部分表示图像的细节。如果我们截取图片 A 的低频部分和图片 B 的高频部分把两张图片拼在一起，那么这张融合的图片远看会比较像图片 A，近看会比较像图片 B。

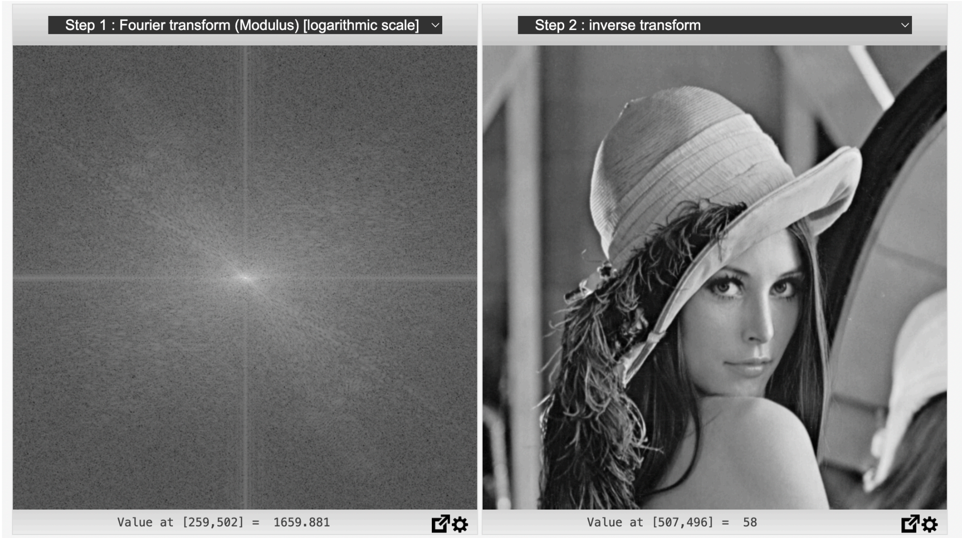


图 4: 图像傅里叶变换与逆变换

18.6.3 Applications in Different Disciplines 在不同学科中的应用

(1). 物理学与工程学

- **量子力学：** 波函数在坐标空间和动量空间之间的变换就是傅里叶变换。
- **光学：** 用于分析光的衍射、成像系统（傅里叶光学）和全息术。
- **热传导与流体力学：** 求解偏微分方程时，傅里叶变换可将方程简化到频域求解。

(2). 医学成像

- **磁共振成像（MRI）：** 采集的信号实质上是频率域的，通过傅里叶逆变换重建出人体断层图像。
- **计算机断层扫描（CT）：** 投影重建算法（如滤波反投影）的核心是傅里叶切片定理。

(3). 金融与经济学

- **时间序列分析：** 通过傅里叶变换分析经济数据的周期性（如商业周期、季节性波动）。
- **期权定价：** 在某些模型中（如傅里叶期权定价方法），利用傅里叶变换高效计算期权价格。

(4). 地质与气象学

- **地震波分析：** 通过傅里叶变换研究地震信号的频谱特征，推断地下结构。
- **气象数据分析：** 分析气候数据的长期趋势和周期性变化（如厄尔尼诺现象）。

19 Laplace Transform 拉普拉斯变换

19.1 Laplace Transform and Practical Application 拉普拉斯变换和实际应用

19.1.1 Definition of Laplace Transform 拉普拉斯变换的定义

拉普拉斯变换是一种积分变换，它将定义在时域 $[0, \infty)$ 上的函数 $f(t)$ 转换为定义在复频域上的函数 $\tilde{f}(s)$。其主要动机在于：

- 将线性常微分方程转换为代数方程，从而简化求解过程。
- 广泛应用于工程、物理（如电路分析、控制理论、信号处理）等领域。
- 能够处理不连续或具有指数增长特性的函数。

定义 19.1 (拉普拉斯变换) 给定函数 $f : [0, \infty) \to \mathbb{R}$，其拉普拉斯变换定义为

$$\tilde{f}(s) := \int_0^\infty f(t) e^{-ts}\, dt,$$

其中 s 为复数（通常取实部足够大以保证积分收敛）。也常记为 $\mathcal{L}\{f\}(s)$ 或 $\mathcal{L}\{f\}$。

以下是几个基本初等函数的拉普拉斯变换及其证明。

命题 19.1 (常数函数的拉普拉斯变换) 对于 $s > 0$，有

$$\mathcal{L}\{1\} = \frac{1}{s}.$$

证明 19.1 直接计算：

$$\mathcal{L}\{1\} = \int_0^\infty e^{-ts}\, dt = \left[-\frac{1}{s} e^{-ts} \right]_0^\infty = 0 - \left(-\frac{1}{s} \right) = \frac{1}{s}.$$

命题 19.2 (指数函数的拉普拉斯变换) 对于 $s > \alpha$，有

$$\mathcal{L}\{e^{\alpha t}\} = \frac{1}{s - \alpha}.$$

证明 19.2

$$\mathcal{L}\{e^{\alpha t}\} = \int_0^\infty e^{\alpha t} e^{-ts}\, dt = \int_0^\infty e^{-(s-\alpha)t}\, dt = \left[-\frac{1}{s-\alpha} e^{-(s-\alpha)t} \right]_0^\infty = \frac{1}{s-\alpha}.$$

命题 19.3 (幂函数的拉普拉斯变换) 对于 $s > 0$ 和 $n \in \mathbb{N}$，有

$$\mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}}.$$

证明 19.3 使用分部积分或递推关系：

$$\mathcal{L}\{t^n\} = \int_0^\infty t^n e^{-ts}\, dt = \frac{n}{s} \mathcal{L}\{t^{n-1}\},$$

递推至 $n = 0$ 的情形即得。

命题 19.4 (三角函数的拉普拉斯变换)

$$\mathcal{L}\{\sin(\alpha t)\} = \frac{\alpha}{\alpha^2 + s^2}, \quad \mathcal{L}\{\cos(\alpha t)\} = \frac{s}{\alpha^2 + s^2}.$$

证明 19.4 利用欧拉公式或分部积分，例如：

$$\begin{aligned}
\mathcal{L}\{\sin(\alpha t)\} &= \int_0^\infty \sin(\alpha t)e^{-ts}\,dt \\
&= \left[-\frac{\cos(\alpha t)}{\alpha}e^{-ts}\right]_0^\infty - \int_0^\infty \left(-\frac{\cos(\alpha t)}{\alpha}\right)(-s)e^{-ts}\,dt \\
&= \frac{1}{\alpha} - \frac{s}{\alpha}\mathcal{L}\{\cos(\alpha t)\}.
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}\{\sin(\alpha t)\} &= \int_0^\infty \sin(\alpha t)e^{-ts}\,dt \\
&= \left[\frac{\sin(\alpha t)}{-s}e^{-ts}\right]_0^\infty - \int_0^\infty \frac{e^{-ts}}{-s}\alpha\cos(\alpha t)dt \\
&= \frac{\alpha}{s}\mathcal{L}\{\cos(\alpha t)\}.
\end{aligned}$$

$$\begin{aligned}
\frac{1}{\alpha} - \frac{s}{\alpha}\mathcal{L}\{\cos(\alpha t)\} &= \frac{\alpha}{s}\mathcal{L}\{\cos(\alpha t)\} \\
\mathcal{L}\{\cos(\alpha t)\} &= \frac{s}{\alpha^2+s^2} \\
\mathcal{L}\{\sin(\alpha t)\} &= \frac{\alpha}{\alpha^2+s^2}
\end{aligned}$$

19.1.2 Existence of Laplace Transform 拉普拉斯变换的存在性

为保证积分收敛，通常对 $f(t)$ 施加以下条件：

定义 19.2 (至多指数增长) 函数 $f(t)$ 在 $[a,\infty)$ 上具有至多指数增长 γ，若存在 $M>0$ 使得

$$|f(t)| \le Me^{\gamma t} \quad \text{对所有} t \in [a,\infty) \text{ 成立}.$$

若存在某个 $\gamma > 0$ 使上式成立，则称 $f(t)$ 具有至多指数增长 *(is of exponential growth at most)*。

定理 19.1 (拉普拉斯变换的存在性) 若 $f(t)$ 连续且具有至多指数增长 γ，则其拉普拉斯变换 $\tilde{f}(s)$ 对所有 $s>\gamma$ 存在。

19.1.3 Properties of Laplace Transform 拉普拉斯变换的性质

拉普拉斯变换具有以下重要性质（均假设变换存在）：

定理 19.2 (线性性) 对任意常数 α, β，

$$\mathcal{L}\{\alpha f + \beta g\} = \alpha\mathcal{L}\{f\} + \beta\mathcal{L}\{g\}.$$

证明 19.5 由积分线性性直接得出。

定理 19.3 (导数的变换) 若 f 连续可微且 f' 具有至多指数增长，则

$$\mathcal{L}\{f'\} = s\mathcal{L}\{f\} - f(0).$$

证明 19.6 分部积分：

$$\mathcal{L}\{f'\} = \int_0^\infty f'(t)e^{-st}\,dt = \left[f(t)e^{-st}\right]_0^\infty + s\int_0^\infty f(t)e^{-st}\,dt = s\mathcal{L}\{f\} - f(0).$$

定理 19.4 (平移性质)

$$\mathcal{L}\{e^{at}f(t)\} = \tilde{f}(s-a).$$

证明 19.7

$$\mathcal{L}\{e^{at}f(t)\} = \int_0^\infty e^{at}f(t)e^{-st}\,dt = \int_0^\infty f(t)e^{-(s-a)t}\,dt = \tilde{f}(s-a).$$

定理 19.5 (伸缩性质) 对 $a, b > 0$,

$$\mathcal{L}\{f(at)\} = \frac{1}{a}\tilde{f}\left(\frac{s}{a}\right), \quad \tilde{f}(bs) = \mathcal{L}\left\{\frac{1}{b}f\left(\frac{t}{b}\right)\right\}.$$

证明 19.8 令 $u = at$ 或 $u = t/b$ 换元即可。

定理 19.6 (乘以 t 的变换) 若 f 具有至多指数增长，则

$$\mathcal{L}\{-tf(t)\} = \frac{d}{ds}\tilde{f}(s).$$

证明 19.9 在积分号下求导：

$$\frac{d}{ds}\tilde{f}(s) = \frac{d}{ds}\int_0^\infty f(t)e^{-st}\,dt = \int_0^\infty f(t)\frac{\partial}{\partial s}e^{-st}\,dt = \int_0^\infty (-tf(t))e^{-st}\,dt = \mathcal{L}\{-tf(t)\}.$$

交换微分与积分的合理性由控制收敛定理保证。

例题 19.1 求出下列函数的拉普拉斯变换：

$$(a)\quad 2e^{3t} - t^2 \qquad (b)\quad 3\cos 5t - \sin 5t \qquad (c)\quad (t+3)^2e^{-t} \qquad (d)\quad e^t \sin t$$

利用基本初等函数的拉普拉斯变换公式以及拉普拉斯变换的线性性，我们可以直接写出 $(a)(b)$ 两问的答案：

$$(a)\quad \frac{2}{s-3} - \frac{2}{s^3} \qquad (b)\quad \frac{3s-5}{s^2+25}$$

$(c)(d)$ 两问都是 $f(t)$ 乘以一项 $e^{\alpha t}$ 的形式，显然需要用到拉普拉斯变换的平移性质：

$$(c)\quad \frac{2}{(s+1)^3} + \frac{6}{(s+1)^2} + \frac{9}{s+1} \qquad (d)\quad \frac{1}{(s-1)^2+1}$$

例题 19.2 求出拉普拉斯变换为下列函数的函数：

$$(a)\frac{s+1}{s^2+4} \quad (b)\frac{5s+2}{s^3} \quad (c)\frac{3s+4}{s^2-2s-3} \quad (d)\frac{1+s+s^2}{s^2(s+1)} \quad (e)\frac{s}{s^2+2s+10}$$

$(a)(b)$ 两问可以直接写出答案：

$$(a)\quad \cos(2t) + \frac{1}{2}\sin(2t) \qquad (b)\quad 5t + t^2$$

$(c)(d)$ 两问需要先进行部分分式展开：

$$(c)\quad \frac{13}{4}e^{3t} - \frac{1}{4}e^{-t} \qquad (d)\quad t + e^{-t}$$

(e) 问需要用到拉普拉斯变换的平移性质：

$$(e)\quad e^{-t}\left(\cos(3t) - \frac{1}{3}\sin(3t)\right)$$

19.1.4 Application: Solving Differential Equations 应用：解微分方程

拉普拉斯变换可将线性常微分方程（含常数系数或某些变系数）转化为代数方程，求解后再逆变换得到原解。

例题 19.3 (常系数线性 ODE) 求解：

$$y'' - 3y' + 2y = 4e^{-t}, \quad y(0) = 1,\ y'(0) = 2.$$

取拉普拉斯变换：

$$(s^2 Y(s) - sy(0) - y'(0)) - 3(sY(s) - y(0)) + 2Y(s) = \frac{4}{s+1}.$$

代入初值得：

$$(s^2 - 3s + 2)Y(s) - s - 2 + 3 = \frac{4}{s+1}.$$

解得：

$$Y(s) = \frac{s^2+3}{(s-1)(s-2)(s+1)} = -\frac{2}{s-1} + \frac{7/3}{s-2} + \frac{2/3}{s+1}.$$

逆变换得：

$$y(t) = -2e^t + \frac{7}{3}e^{2t} + \frac{2}{3}e^{-t}.$$

例题 19.4 (变系数线性 ODE) 求解：

$$ty''(t) - ty'(t) + y(t) = 2, \quad y(0) = 2,\ y'(0) = -4.$$

令 $Y(s) = \mathcal{L}\{y(t)\}$。利用性质 $\mathcal{L}\{tg(t)\} = -\frac{d}{ds}\mathcal{L}\{g\}$：

$$\mathcal{L}\{ty''\} = -s^2 Y'(s) - 2sY(s) + y(0),$$
$$\mathcal{L}\{ty'\} = -sY'(s) - Y(s).$$

代入方程并化简得：

$$s(s-1)Y'(s) + 2(s-1)Y(s) = 2\frac{s-1}{s}.$$

解得：

$$Y(s) = \frac{2}{s} + \frac{c}{s^2}.$$

逆变换得 $y(t) = 2 + ct$，利用 $y'(0) = -4$ 得 $c = -4$，故

$$y(t) = 2 - 4t.$$

19.2 Understanding Laplace Transform 理解拉普拉斯变换

19.2.1 Geometric Insights about the Essence of Laplace Transform 有关拉普拉斯变换本质的几何洞见

拉普拉斯变换可以看作函数 $f(t)$ 在指数衰减（或增长）权函数 e^{-st} 下的投影：

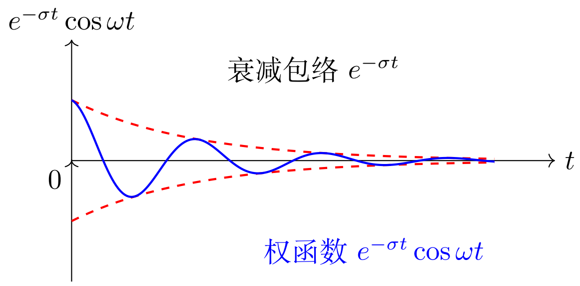
$$\mathcal{L}\{f(t)\}(s) = \int_0^\infty f(t)e^{-st}dt$$

其中 $s = \sigma + i\omega$ 是复数。当 $\sigma > 0$ 时，$e^{-\sigma t}$ 表示指数衰减，$\sigma < 0$ 时表示指数增长。

- **权重函数** e^{-st} 对 $f(t)$ 进行加权：

$$权重 = e^{-(\sigma+i\omega)t} = e^{-\sigma t}(\cos\omega t - i\sin\omega t)$$

- **实部** $e^{-\sigma t}\cos\omega t$ 和 **虚部** $e^{-\sigma t}\sin\omega t$ 形成振荡衰减的权函数
- 变换结果 $\tilde{f}(s)$ 是 $f(t)$ 在这种振荡衰减模式下的“匹配程度”



19.2.2 Laplace Transform and Fourier Transform 拉普拉斯变换与傅里叶变换

傅里叶变换将时域信号分解为不同频率的正弦波：

$$\mathcal{F}\{f(t)\}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

当 $f(t)$ 满足一定条件时，拉普拉斯变换是傅里叶变换的推广：

$$\mathcal{L}\{f(t)\}(\sigma + i\omega) = \mathcal{F}\{f(t)e^{-\sigma t}\}(\omega)$$

即：对 $f(t)$ 先乘以指数衰减 $e^{-\sigma t}$，再作傅里叶变换，就得到拉普拉斯变换。

- **傅里叶变换**：仅在虚轴上考察（$s = i\omega$）

$$\mathcal{L}\{f(t)\}(i\omega) = \int_{0}^{\infty} f(t)e^{-i\omega t}dt$$

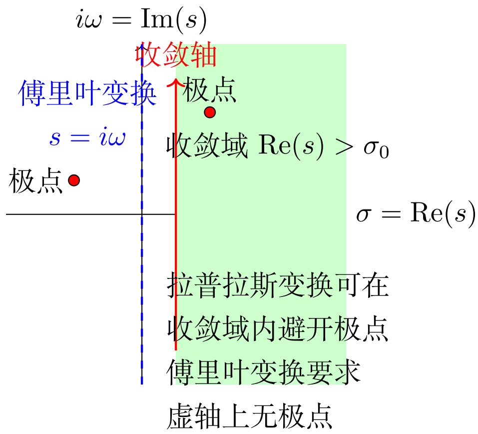
  这是单边傅里叶变换

- **拉普拉斯变换**：在整个复平面上考察（$s = \sigma + i\omega$）

$$\mathcal{L}\{f(t)\}(\sigma + i\omega) = \int_{0}^{\infty} [f(t)e^{-\sigma t}]e^{-i\omega t}dt$$

  相当于先对 $f(t)$ 乘以收敛因子 $e^{-\sigma t}$，再作傅里叶变换

- 傅里叶变换要求 $f(t)$ 绝对可积：$\int |f(t)|dt < \infty$
- 拉普拉斯变换通过选择适当的 σ 使 $f(t)e^{-\sigma t}$ 绝对可积
- 几何上：选择复平面上合适的垂直线 $\mathrm{Re}(s) = \sigma$，使得在这条线右侧变换收敛



- **傅里叶变换**：分析周期/稳态特性，频域分析
- **拉普拉斯变换**：分析瞬态响应，稳定性，收敛性
- 几何上：
  - 傅里叶变换：纯振荡分解（单位圆）
  - 拉普拉斯变换：振荡 + 衰减/增长（复平面）

20 About this Note 关于本笔记的说明

大学数学课程可以大致分为**基础课**和**专业课**。基础课的难度偏低，主要侧重理解概念并且掌握计算，不太要求掌握证明，对于很多话题不会深挖。基础课不仅数学专业的学生需要掌握，也是很多非数学专业（例如其他理科、工科、社科、商科等）学生需要一起学习的基础数学课程，这说明基础课不仅是学习更高级更深入的数学的基石，更是很多其他学科（如经济学、统计学、管理学、物理、化学、计算机科学、数据科学、机器学习等）的关键数学门槛。对于非数学专业的学生而言，掌握这些知识最基本的理解和计算已完全足够，事实上最基础的往往也是最重要的部分。

大学数学的入门内容或者说核心必修内容包括“三件套”：**微积分**（Calculus）、**线性代数**（Linear Algebra）和**概率论与数理统计**（Probability and Mathematical Statistics）。

在面向包含非数学专业的大众理工科学生的**基础课**里，介绍微积分的这门课往往被叫做**《高等数学》**（Advanced Mathematics），简称“《高数》”。也有很多国外大学把这门课直接就叫**《微积分》**。其内容紧承中学数学，包含集合、映射、函数、极限、导数、积分和微分方程等讨论。高数的面向群体很广，学习这门课的学生来自五花八门的专业，也正因如此，这门课的难度设置得比较友好，侧重于计算和解题，帮助工科等的学生快速入门多元微积分基础知识，为后续专业课的学习服务。对于非数学专业学生而言，完全没有必要将微积分相关的知识学得非常深入和细致。除了不纠结理论和证明，国外很多学校甚至会把国内高数课程中一些相对较难或者较进阶的知识点（例如函数极限、重积分、反常积分、拉普拉斯变换、傅里叶级数等）也给删去，只保留最基础的偏应用导向的内容。事实上，只是这些最基础的内容也足以让工科学生取得非常让人安心的数学前置准备。而对于数学专业的学生而言，高等数学这门课的课程安排就显得有些“少儿版”了，其深度和严谨程度有所欠缺。对于数学专业的学生，我们应以更高的标准要求自己，对于同样的内容，应以更严谨更系统的模式去学习，不仅要掌握最基本的计算，更要学会从分析的角度看数学。**专业课**便是专门面向数学专业学生的课程，在专业课里介绍微积分的这门课叫做**《数学分析》**（Mathematical Analysis），简称“《数分》”，又称**《高级微积分》**。数分与高数讨论的是同样的话题，但却从深度、难度上都是一个“升级版”。在高数中被避而不谈直接省略的许多命题和定理的证明过程以及分析步骤，在数分中我们都会详细讨论。在高数中很多没有被严格定义的数学名词在数分中我们也会给出精确完善的定义。这些内容从应用的眼光来看都是既繁琐又无用的部分，但却对应着数学专业的学生需要尽早培养的思维和能力。事实上，想从专业严谨的角度深刻地讨论微积分，就绕不开分析，而数学分析便是数学专业学生分析学的入门课。

类似的，介绍线性代数（主要研究行列式、向量空间、线性变换等）的课在基础课里往往就叫**《线性代数》**，简称“《线代》”，而在专业课里这门课叫**《高等代数》**，简称“《高代》”。但注意，与微积分的情况不同的是，《高等代数》这门课严格来说不仅仅讲授了线性代数的内容，而是包含了三个方面：**线性代数**，**多项式理论**，**群、环、域的基本概念**。当然，线性代数占的比重最大，相较基础课，在同样的内容上《高代》会提供更高的难度。除此之外专业课还在其之后很自然地拓展了两块内容，对代数结构做了一些入门的介绍，让整个课程内容更加丰富，也很好地为之后的进阶代数课做了铺垫：多项式理论研究一元和多元多项式环；群、环、域的基本概念紧密结合多项式理论和线性变换（包括与度量有关的线性变换）理论，水到渠成地介绍一元（多元）多项式环、矩阵环、线性变换环、模 p 剩余类域、正交群、酉群和辛群。自从 1832 年伽罗瓦利用一元高次方程的根的置换群给出了方程有求根公式的充分必要条件之后，代数学的研究对象发生了根本性的转变。研究各种代数系统的结构及其态射（即保持运算的映射）成为现代代数学研究的中心问题。20 世纪，代数学研究结构及其态射的观点已经渗透到现代代数学的各个分支中。因此，在高等代数课程中贯穿研究线性空间和多项式环的结构及其态射这条主线是非常合理的安排，可以帮助学生快速把握整个代数学的精髓。

概率论与数理统计（主要研究随机事件、概率分布、统计推断等）一般就以这个名字命名课程，简称“概统”，只不过作为统计学或者数学学生的专业课时讲授的会更加严谨专业，难度会更高。因此，若有非数学专业学生在学习了高等数学和线性代数的教材后感到题目过于简单，部分内容讲授地过于粗糙简略或者草率，还学有余力，可以学习一下数学分析和高等代数的课程。值得一提的是，市面上针对刚才提到的每一门课的教材因出自不同教授之手，难度自然会有差异，但除了个别教材是极端值外（主要是过难），主流的教材基本上平均难度相差不大。

数学专业的学生在学完了这些入门的专业课后，如果对分析学感兴趣，可以进一步学习**《实分析/实变函数》**（$\mathbb{R}$ 上数学分析）、**《复分析/复变函数》**（$\mathbb{C}$ 上数学分析）、**《泛函分析》**（无限维数学分析）、**《常微分方程》**、**《偏微分方程》**等课程；如果对代数学感兴趣，可以进一步学习**《近世代数/抽象代数》**、**《数论》**、**《离散数学》**等课程；如果对统计学感兴趣，可以进一步学习**《高等概率论》**、**《随机矩阵》**、**《随机过程》**、**《贝叶斯推断》**、**《时间序列》**、**《理论计量经济学》**等课程；如果对几何学感兴趣，可以进一步学习**《解析几何》**、**《微分几何》**、**《点集拓扑》**、**《代数拓扑》**、**《流形》**等课程。当然，统计学本质上也算一种应用数学，其他热门的应用数学课程还包括：**《数值分析》**、**《运筹学与优化》**、**《博弈论》**、**《信息论》**、**《数学建模》**等。其中，**机器学习**算是统计学与计算机科学结合的产物；**计量经济学**算是统计学的一个分支；**理论经济学**算是应用数学的一个分支。但不管怎样，想要在任一这些方向上进行后续的学习或研究，数学入门“三件套”都是你必须扎实掌握的基础知识。

本笔记是微积分入门课程，参考资料既包含简单版微积分的《高等数学》教材，也包含进阶版微积分的《数学分析》教材，固应属于与《数学分析》同一难度水平的课程。为了强调笔记主要内容及尽可能不造成歧义，固将本笔记简单地命名为“微积分”。

21 Reference 参考资料

LSE MA100 LectureNotes

LSE MA103 LectureNotes

LSE MA203 LectureNotes

LSE MA212 slides

Thomas Calculus book

数学分析陈纪修第三版

数学基础 30 讲高等数学分册张宇

宋浩高等数学教学视频

本笔记的内容涵盖以下 LSE 的课程：MA100（微积分部分）, MA103, MA203, MA212（微积分部分）