Math Dominates the Universe
数学支配着宇宙

Jingze Xing

Wednesday 31st August, 2022

目录

第一章　微积分 A1 ：一元微积分与常微分方程

1.1　极限

1.1.1　基本事实 & 可避免的错误

- $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{\theta}{n!n}$，$\frac{n}{n+1} < \theta < 1$
- Taylor 展开 $\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$，不可忘记除以 $n!$
- **(Leibniz)** 两个函数之积的高阶导数

$$(f \cdot g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

1.1.2　例题

课本 p.57-9(14)

$$\begin{aligned}
&\lim_{x\to\infty}(\sqrt{x^2+2x} - \sqrt[3]{x^3-x^2}) \\
=&\lim_{x\to\infty}(x\sqrt{1+\frac{2}{x}} - x\sqrt[3]{1-\frac{1}{x}}) \\
=&\lim_{x\to\infty}(\frac{(1+\frac{2}{x})^{\frac{1}{2}}}{\frac{1}{x}}) - \lim_{x\to\infty}(\frac{(1-\frac{1}{x})^{\frac{1}{3}}}{\frac{1}{x}}) \\
=&2 * \frac{1}{2} - (-1) * \frac{1}{3} = \frac{4}{3}
\end{aligned}$$

《数学分析习题课讲义》2.3.2-6

求证：

$$S_n = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p}\ (p > 1) \tag{1.1}$$

收敛

证明.

$$\frac{1}{2^p} + \frac{1}{3^p} < \frac{2}{2^p} = r$$

$$\frac{1}{4^p}+\frac{1}{5^p}+\frac{1}{6^p}+\frac{1}{7^p}<\frac{4}{4^p}=r^2$$
$$\therefore S_n\le S_{2^n-1}<1+r+r^2+\cdots+r^n<\frac{1}{1-r}$$

又 $\because \{S_n\}$单调递增 $\therefore \{S_n\}$收敛.

$\square$

习题课 1-1

求证：

$$\lim_{n\to+\infty}\left(\frac{1^p+2^p+\cdots+n^p}{n^p}-\frac{n}{p+1}\right)=\frac{1}{2}\tag{1.2}$$

证明. 该式本质是 n 的多项式

$$\text{let } x_n=(p+1)(1^p+2^p+\cdots+n^p)-n^{p+1},\ y_n=(p+1)n^p$$

$$\begin{aligned}
\text{原式}&=\lim_{x\to+\infty}\frac{x_{n+1}-x_n}{y_{n+1}-y_n}\\
&=\lim_{x\to+\infty}\frac{(p+1)(n+1)^p-(n+1)^{p+1}+n^{p+1}}{(p+1)((n+1)^p-n^p)}\\
&=\lim_{x\to+\infty}\frac{(p+1)\sum\limits_{k=0}^{p}\binom{p}{k}n^k-\sum\limits_{k=0}^{p}\binom{p+1}{k}n^k}{(p+1)\sum\limits_{k=0}^{p-1}\binom{p}{k}n^k}\\
&=\lim_{x\to+\infty}\frac{(p+1)\sum\limits_{k=0}^{p-2}\binom{p}{k}n^k-\sum\limits_{k=0}^{p-2}\binom{p+1}{k}n^k+\frac{1}{2}p(p+1)n^{p-1}}{(p+1)\sum\limits_{k=0}^{p-2}\binom{p}{k}n^k+p(p+1)n^{p-1}}\\
\text{(只考虑最高次项)}\ &=\frac{1}{2}
\end{aligned}$$

$\square$

习题课 4-6(4)

已知 $f(n)=x^n\ln n$，求 $f^{(n)}(x)$.

解. 直接用 Leibniz 公式会得到复杂的交错和，因此考虑递推.

$$\begin{aligned}
f^{(n)}(x)&=\big(f'(x)\big)^{(n-1)}=\big(nx^{n-1}\ln x+x^{n-1}\big)^{(n-1)}\\
&=\big(nx^{n-1}\ln x\big)^{(n-1)}+(n-1)!=n\cdot\big(x^{n-1}\ln x\big)^{(n-1)}+n!\frac{1}{n}\\
&=n\cdot\big((n-1)x^{n-2}\ln x+x^{n-2}\big)^{(n-2)}+n!\cdot\frac{1}{n}
\end{aligned}$$

$$
\begin{aligned}
&= n(n-1)\cdot\left(x^{n-2}\ln x\right)^{(n-2)} + n!\left(\frac{1}{n-1}+\frac{1}{n}\right)\\
&= n(n-1)\cdot\left((n-2)x^{n-3}\ln x + x^{n-3}\right)^{(n-3)} + n!\left(\frac{1}{n-1}+\frac{1}{n}\right)\\
&= n(n-1)(n-2)\cdot\left(x^{n-3}\ln x\right)^{(n-3)} + n!\left(\frac{1}{n-2}+\frac{1}{n-1}+\frac{1}{n}\right)\\
&= \cdots = n!\left(\ln x + 1 + \frac{1}{2} + \cdots \frac{1}{n-2}+\frac{1}{n-1}+\frac{1}{n}\right)
\end{aligned}
$$

习题课 4-7

定义

$$P_{n,m}(x) = \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1-x^m)^n$$

，求 $P_{n,m}(1)$.

解.

$$
\begin{aligned}
P_{n,m}(1) &= \frac{\mathrm{d}^n}{\mathrm{d}x^n}(1-x)^n(1+x^2+\cdots+x^{m-1})^n\\
\text{（用 Leibniz 定理展开）} &= (-1)^n n!\cdot m^n
\end{aligned}
$$

*$(1-x)^n$ 求至少 n 阶导数才非 0.

微积分 A 期中讲座

定义 $\{a_n\}: a_1 = 1, a_{n+1} = \sin a_n$，求证：$\lim\limits_{n\to\infty} n\cdot a_n^2 = 3$.

证明. 用 stolz 定理表现类等差数列 $\frac{1}{a_n^2}$ 公差趋向 $\frac{1}{3}$.

$$
\begin{aligned}
\lim_{n\to\infty}\frac{\frac{1}{a_n^2}}{n} &= \lim_{n\to\infty}\left(\frac{1}{a_{n+1}^2}-\frac{1}{a_n^2}\right) = \lim_{n\to\infty}\frac{a_n^2-\sin^2 a_n}{a_n^2\sin^2 a_n}\\
&= \lim_{x\to 0}\frac{(x-x+\frac{1}{6}x^3+o(x^3))(x+x-\frac{1}{6}x^3+o(x^3))}{x^4} = \frac{1}{3}
\end{aligned}
$$

□

1.2　Taylor 展开、Cauchy 中值定理

课本 p.125-10

设函数 $y = f(x)$ 在 $(-1,1)$ 内二阶可导，$f''(0)\neq 0$. $\forall x\in(-1,1), x\neq 0$，$\exists\ \theta(x)$ 满足 $f(x)-f(0) = xf'(x\cdot\theta(x))$，证明：$\lim\limits_{x\to 0}\theta(x) = \frac{1}{2}$

证明. 用 Cauchy 中值定理将 $\theta(x)$ 分离出来.

$$
\begin{aligned}
f^{''}(\zeta_x)\cdot\theta(x)x &= f^{'}(x\cdot\theta(x)) - f^{'}(0) \\
f^{''}(\zeta_x)\cdot\theta(x)x &= \frac{f(x)-f(0)}{x} - f^{'}(0) \\
\theta(x) &= \frac{1}{f^{''}(\zeta_x)}\cdot\frac{f(x)-f(0)-f^{'}(0)x}{x^2} \\
&= \frac{1}{f^{''}(\zeta_x)}\cdot\frac{f^{'}(x)-f^{'}(0)}{2x} \\
&= \frac{f^{''}(0)}{2f^{''}(0)} = \frac{1}{2}\ (x\to 0)
\end{aligned}
$$

□

习题课 6-2-1

设 $f(x)$ 在 $[a,b]$ 上连续，在 (a,b) 内二阶可导，证明 $\exists\xi\in(a,b)$，使得

$$
\frac{\frac{f(x)-f(a)}{x-a}-\frac{f(b)-f(a)}{b-a}}{x-b} = \frac{1}{2}f^{''}(\xi)
$$

证明. 令 $g(x)=\frac{f(x)-f(a)}{x-a}$

则 $\frac{\frac{f(x)-f(a)}{x-a}-\frac{f(b)-f(a)}{b-a}}{x-b} = \frac{g(x)-g(b)}{x-b} = g'(\eta) = \frac{f'(\eta)(\eta-a)-(f(\eta)-f(a))}{(\eta-a)^2}$

而 $f(a)=f(x)+f'(x)(a-x)+\frac{f''(\xi)}{2}(a-x)^2$

故 $g'(\eta)=\frac{1}{2}f''(\xi)$ □

1.3 凸函数

- 开区间上的凸函数连续，闭区间上的凸函数未必连续
- 开区间上的凸函数处处存在两个单侧导数，且对于下凸函数，满足 $f'_-(x)\le f'_+(x)$，但是导数不一定存在

1.4 积分

1.4.1 基本定理

- $f(x)$ 在 J 上非一致连续 $\iff$
  $\exists\varepsilon_0>0,\ x_n,x'_n\in J$, 满足 $|x_n-x'_n|<\frac{1}{n}$，但 $|f(x_n)-f(x'_n)|>\varepsilon_0,\forall n\ge 1$
- 有界闭区间上的连续函数、单调函数黎曼可积，可积必有界

- **(Lebesgue)** $[a,b]$ 上的有界函数 $f(x)$ 可积 $\iff$ $f(x)$ 在 $[a,b]$ 上的间断点集为零测集
- **(Cauchy-Schwarz)**
$$\left(\int_a^b f\cdot g\right)^2\le\int_a^b f^2\int_a^b g^2$$
- **(积分中值定理)** 设 $f(x)$ 在 $[a,b]$ 上连续，则
$$\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x$$
- 可积函数 $f(x)$ 的变上限积分 $g(x)=\int_a^x f(t)\mathrm{d}t$ 连续，且若 $f(x)$ 在 x_0 处**连续**，即有 $g'(x_0)=f(x_0)$
- 导函数不一定可积，如：
$$F(x)=\begin{cases}x^2\sin(\frac{1}{x^2}) & x\neq 0\\ 0 & x=0\end{cases}$$
$$F'(x)=\begin{cases}2x\sin(\frac{1}{x^2})-\frac{2}{x}\cos(\frac{1}{x^2}), & x\neq 0\\ 0 & x=0\end{cases}$$
其中 $F'(x)$ 在 $[0,1]$ 上无界，不可积.

1.4.2　常见不定积分

$$\int\frac{1}{\sqrt{x^2+a}}=\ln\left|x+\sqrt{x^2+a}\right|+C$$
$$\int\frac{1}{x^2+a}=\frac{1}{\sqrt{a}}\arctan(\frac{x}{\sqrt{a}})+C$$
$$\int\frac{\mathrm{d}x}{\sin x}=\int\frac{\mathrm{d}(\frac{x}{2})}{\sin\frac{x}{2}\cos\frac{x}{2}}=\int\frac{\mathrm{d}(\frac{x}{2})}{\tan\frac{x}{2}\cos^2\frac{x}{2}}=\int\frac{\mathrm{d}(\tan\frac{x}{2})}{\tan\frac{x}{2}}=\ln|\tan\frac{x}{2}|+C$$
$$\int\frac{\mathrm{d}x}{x\sqrt{x^2-1}}=\int\frac{\mathrm{d}x}{x^2\sqrt{1-\frac{1}{x^2}}}=-\int\frac{\mathrm{d}(\frac{1}{x})}{\sqrt{1-\frac{1}{x^2}}}=\arccos\left(\frac{1}{x}\right)+C$$
$$\begin{aligned}\int\sqrt{x^2-1}\mathrm{d}x&=\int\sqrt{\sec^2 t-1}\mathrm{d}(\sec t)=\int\frac{\sin^2 t}{\cos^3 t}\mathrm{d}t=\int\frac{\sin^2 t\mathrm{d}(\sin t)}{\cos^4 t}\\&=\int\frac{u^2}{(1-u)^2(1+u)^2}\mathrm{d}u=\int\frac{1}{4}\left(\frac{u}{(1-u)^2}-\frac{u}{(1+u)^2}\right)\mathrm{d}u\\&=\frac{1}{4}\int\left(\frac{1}{(1-u)^2}-\frac{1}{1-u}-\frac{1}{1+u}+\frac{1}{(1+u)^2}\right)\mathrm{d}u\\&=\frac{1}{4}\left(-\frac{1}{u-1}-\frac{1}{u+1}+\ln\frac{1-u}{1+u}\right)+C\\(u=\sin t)\quad&=\frac{\sin t}{2\cos^2 t}+\frac{1}{2}\ln\frac{1-\sin t}{\cos t}+C\end{aligned}$$

$$
\begin{aligned}
&(\cos t = \frac{1}{x}) && = \frac{1}{2}x\sqrt{x^2-1} + \frac{1}{2}\ln(x-\sqrt{x^2-1}) + C \\
&\int \frac{\mathrm{d}x}{x^2\sqrt{x^2-1}} && = \frac{\sqrt{x^2-1}}{x} + C \\
&J_n = \int_0^{\frac{\pi}{2}} \cos^n x\mathrm{d}x && = \int_0^{\frac{\pi}{2}} \cos^{n-1} x\mathrm{d}(\sin x) = \cos^{n-1} x \sin x\Big|_0^{\frac{\pi}{2}} + (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2} x \sin^2 x\mathrm{d}x \\
& && = (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2} x\mathrm{d}x - (n-1)\int_0^{\frac{\pi}{2}} \cos^n x\mathrm{d}x = (n-1)J_{n-2} - (n-1)J_n \\
& && \Longrightarrow J_n = \frac{n-1}{n}J_{n-2} \Longrightarrow J_{2m} = \frac{(2m-1)!!}{(2m)!!}\frac{\pi}{2}, J_{2m+1} = \frac{(2m)!!}{(2m+1)!!}
\end{aligned}
$$

1.4.3　例题

Lijun Yang: Nov.18 P25

设 $f(x)$ 在 $[a,b]$ 上连续可微，且 $f(a)=0$. 证明:

$$\int_a^b f^2(x)\mathrm{d}x \leq \frac{1}{2}(b-a)^2 \int_a^b [f'(x)]^2\mathrm{d}x$$

证明.

$$
\begin{aligned}
f(x) &= \int_a^x f'(t)\mathrm{d}t = \int_a^x 1\cdot f'(t)\mathrm{d}t \\
f^2(x) &= \left(\int_a^x 1\cdot f'(t)\mathrm{d}t\right)^2 \\
(\text{Cauchy}) &\leq \left(\int_a^x 1\cdot \mathrm{d}t\right)\cdot\left(\int_a^x [f'(t)]^2\mathrm{d}t\right) \\
&\leq (x-a)\int_a^b [f'(t)]^2\mathrm{d}t \\
\int_a^b f^2(x)\mathrm{d}x &\leq \int_a^b (x-a)\mathrm{d}x \int_a^b [f'(t)]^2\mathrm{d}t \\
&\leq \frac{1}{2}(b-a)^2 \int_a^b [f'(x)]^2\mathrm{d}x
\end{aligned}
$$

□

习题课 8-4-4.4

设 $k,n\in\mathbb{Z}_+$，求证：

$$\int_0^\pi \cos nx \cos kx\mathrm{d}x = \begin{cases} \frac{\pi}{2}, & k=n \\ 0, & k\neq n \end{cases} \tag{1.3}$$

证明.

$$\begin{aligned}\int_0^{\pi}\cos nx\cos kx &= \int_0^{\pi}\frac{1}{2}(\cos((n+k)x)+\cos((n-k)x))\mathrm{d}x\\ &= \frac{1}{2}\int_0^{\pi}\cos((n+k)x)\mathrm{d}x+\frac{1}{2}\int_0^{\pi}\cos((n-k)x)\mathrm{d}x\\ &= \frac{\pi}{2}[n=k]\end{aligned}$$

□

1.5　积分的应用

- 极坐标下的面积
$$S=\int_a^n\frac{1}{2}f^2(\theta)\mathrm{d}\theta$$
- 弧长
$$|\Gamma|=\int_{\alpha}^{\beta}\sqrt{x'(t)^2+y'(t)^2}\mathrm{d}t$$
- 极坐标下曲线的弧长
$$|\Gamma|=\int_{\alpha}^{\beta}\sqrt{r(\theta)^2+r'(\theta)^2}\mathrm{d}\theta$$
- 曲率
$$\kappa=\frac{|x'y''-x''y'|}{[x'^2+y'^2]^{\frac{3}{2}}}$$
或
$$\kappa(x)=\frac{|f''(x)|}{[1+f'(x)^2]^{\frac{3}{2}}}$$
REMARK: 圆的曲率为 $\frac{1}{R}$
- 绕 x,y 轴的旋转体体积
$$V_x=\int_a^b\pi f^2(x)\mathrm{d}x,V_y=\int_a^b 2\pi x\cdot f(x)\mathrm{d}x$$
- 绕 x,y 轴的旋转体表面积
$$S_x=\int_a^b 2\pi f(x)\sqrt{1+f'(x)^2}\mathrm{d}x,S_y=\int_a^b 2\pi x\sqrt{1+f'(x)^2}\mathrm{d}x$$
- 曲线的形心（质量均匀时即为质心）
$$\bar{x}=\frac{\int_a^b x(t)\sqrt{x'^2+y'^2}\mathrm{d}t}{\int_a^b\sqrt{x'^2+y'^2}\mathrm{d}t},\bar{y}=\frac{\int_a^b y(t)\sqrt{x'^2+y'^2}\mathrm{d}t}{\int_a^b\sqrt{x'^2+y'^2}\mathrm{d}t}$$

- 平面图形的形心

$$\bar{x}=\frac{\int_a^b xf(x)\mathrm{d}t}{\int_a^b f(x)\mathrm{d}t},\bar{y}=\frac{\int_a^b \frac{1}{2}f^2(x)\mathrm{d}t}{\int_a^b f(x)\mathrm{d}t}$$

- (Guldin,I,II)
  - 曲线绕直线旋转所得的旋转面的侧面积，等于曲线的弧长，乘以形心绕直线旋转一周的周长
  - 封闭图形绕直线旋转所得的旋转面的体积，等于图形的面积，乘以形心绕直线旋转一周的周长

1.6 广义积分

1.6.1 定义

设 $f(x)$ 在 $[a,b)$ 上有唯一瑕点 b，且 $\forall b'\in(a,b)$，$f(x)$ 在 $[a,b')$ 上可积，则 $f(x)$ 在该区间上**内闭可积**.

1.6.2 Dirichlet 判敛

设有：

1. $f(x)$ 在 $[a,b)$ 上**内闭可积**，且存在 $M>0$，使得 $|\int_a^{b'}f(x)\mathrm{d}x|<M,\forall b'\in[a,b)$
2. $g(x)$ 在 $[a,b)$ 上**单调**且 $\lim\limits_{x\to b^-}g(x)=0$

则广义积分 $\int_a^b f(x)g(x)\mathrm{d}x$ 收敛

1.6.3 Abel 判敛

设有：

1. $f(x)$ 在 $[a,b)$ 上**内闭可积**，且 $\int_a^b f(x)\mathrm{d}x$ 收敛
2. **(ii)** $g(x)$ 在 $[a,b)$ 上**单调有界**

则广义积分 $\int_a^b f(x)g(x)\mathrm{d}x$ 收敛

1.7 常微分方程

1.7.1 常数变易法

1. 求解 $f'(x)+p(x)f(x)=q(x)$,

2. 先解 $f'(x)+p(x)f(x)=0$，得 $f(x)=Ce^{-\int p(x)\mathrm{d}x}$
3. 进一步设原方程解为 $f(x)=C(x)e^{-\int p(x)\mathrm{d}x}$
4. $C'(x)e^{-\int p(x)\mathrm{d}x}=q(x)$
5. $f(x)=e^{-\int p(x)\mathrm{d}x}(\int q(x)e^{\int p(x)\mathrm{d}x}+C)$

1.7.2 特殊可降阶高阶常微分方程

方程中不显含自变量 x，可以表示为 $F(y,\frac{\mathrm{d}y}{\mathrm{d}x},\frac{\mathrm{d}^2y}{\mathrm{d}x^2})=0$.

令 $p=\frac{\mathrm{d}y}{\mathrm{d}x}$，则

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}p}{\mathrm{d}x}=\frac{\mathrm{d}p}{\mathrm{d}y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x}=p\cdot\frac{\mathrm{d}p}{\mathrm{d}y}$$

，问题转化为函数 p 关于自变量 y 的一阶常微分方程 $F(y,p,p\frac{\mathrm{d}p}{\mathrm{d}y})$.

1.7.3 二阶线性常系数齐次方程

设 p,q 为实常数，则对常微分方程

$$y''+py'+qy=0$$

，称 $\lambda^2+p\lambda+q=0$ 为特征方程，令 $\Delta=p^2-4q$，设方程的根为 $\lambda_{1,2}$.

- $\Delta>0$，$y=C_1e^{\lambda_1x}+C_2e^{\lambda_2x}$
- $\Delta=0$，$y=(C_1+C_2x)e^{-\frac{p}{2}x}$
- $\Delta<0$，有二虚根 $\lambda=\alpha\pm i\beta$，$y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x)$

1.7.4 Euler 方程

设 $a_0,a_1,\ldots,a_n$ 为实常数，则方程

$$x^ny^{(n)}+a_{n-1}x^{n-1}y^{(n-1)}+\cdots+a_1xy'+a_0y=0$$

称为 **Euler 方程**，一般作变量替换 $t=\log|x|$ 将方程化为以 t 为自变量的常系数方程，第 k 项转化为：

$$\begin{aligned}a_kx^ky^{(k)}&=a_kx^k\cdot\boldsymbol{y}(t(x))^{(k)}\\&=a_kx^k\cdot(\boldsymbol{y}'(t)\cdot\frac{1}{x})^{(k-1)}\\&=a_kx^k\cdot(y''(t))^{(k-2)}\\&=\ldots\end{aligned}$$

第二章　微积分 A2：多元函数微积分与级数

2.1　多元函数的连续、可导与可微

2.1.1　基本定理

- $f(x,y)$ 在 (x_0,y_0) 连续 $\iff$
  动点 (x,y) 以任意路径趋向 (x_0,y_0) 时的极限全部等于 $f(x_0,y_0)$

- $f(x,y)$在(x_0,y_0)可微 $\iff$

$$\lim_{(\Delta x,\Delta y)\to(0,0)} \frac{f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0)-f'_x(x_0,y_0)\Delta x-f'_y(x_0,y_0)\Delta y}{\sqrt{\Delta x^2+\Delta y^2}}=0$$

- 函数是否连续与偏导数是否存在无关

- 函数可微 $\Longrightarrow$ 偏导数存在
  偏导数连续 $\Longrightarrow$ 函数可微
  o 邻域上偏导数有界 $\Longrightarrow$ 函数可微
  邻域上偏导数只有一个不一致连续 $\Longrightarrow$ 函数可微

- 梯度的定义

$$\mathrm{grad} f(\boldsymbol{X}_0)=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\ldots,\frac{\partial f}{\partial x_n}\right)_{\boldsymbol{X}_0}$$

- $\boldsymbol{J}(f\circ g)=\boldsymbol{J}(f)\cdot\boldsymbol{J}(g)$（一阶微分形式不变）
  $\mathrm{d}f(\boldsymbol{X}_0)=\boldsymbol{J}f(\boldsymbol{X}_0)\mathrm{d}\boldsymbol{X}$

- 多元隐函数的偏导数

$$\frac{\partial y}{\partial x_i}=-\frac{\frac{\partial F}{\partial x_i}(\boldsymbol{X},y)}{\frac{\partial F}{\partial y}(\boldsymbol{X},y)}$$

- 向量值隐函数的 Jacobi 矩阵

$$\boldsymbol{J}f(\boldsymbol{X})=\frac{\partial(y_1,\ldots,y_m)}{\partial(x_1,\ldots,x_n)}=-\left(\frac{\partial(F_1,\ldots,F_m)}{\partial(y_1,\ldots,y_m)}\right)^{-1}\frac{\partial(F_1,\ldots,F_m)}{\partial(x_1,\ldots,x_n)}$$

理解：记 $G_k(\boldsymbol{X}) = f_k(\boldsymbol{X}, y_1(\boldsymbol{X}), \ldots, y_m(\boldsymbol{X})) \equiv 0$

$$\frac{\partial(F_1,\ldots,F_m)}{\partial(y_1,\ldots,y_m)}\frac{\partial(y_1,\ldots,y_m)}{\partial(x_1,\ldots,x_n)} + \frac{\partial(F_1,\ldots,F_m)}{\partial(x_1,\ldots,x_n)} = \frac{\partial(G_1,\ldots,G_m)}{\partial(x_1,\ldots,x_n)} = 0$$

- 隐函数的逆映射：对于 $\boldsymbol{Y} = f(\boldsymbol{x}), f:\mathbb{R}^n \to \mathbb{R}^n$，当 $\det(\boldsymbol{J}f(X_0))$ 非零时，有：

$$\boldsymbol{J}f^{-1}(Y_0) = [\boldsymbol{J}f(Y_0)]^{-1}, Y_0 = f(X_0)$$

★ Taylor 展开

$$f(x,y) = \sum_{k=0}^{n}\frac{1}{k!}\left(\frac{\partial}{\partial x}(x-x_0) + \frac{\partial}{\partial y}(y-y_0)\right)^k f(x_0,y_0) + o\left((\sqrt{(x-x_0)^2+(y-y_0)^2})^n\right)$$

★ **（多元函数的中值定理）** 设 $f:\mathbb{R}^n \to \mathbb{R}$ 为 C^1 函数，则 $\forall a, b \in \Omega$，$\exists\, \theta \in (0,1)$，s.t.

$$f(b) - f(a) = \nabla f(a + (b-a)\theta)\cdot(b-a)$$

* 定理不能推广到向量值函数的情形

2.1.2 极值与条件极值

★ $f(x,y)$ 在 (x_0,y_0) 处有极值的条件

（必要）$\frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = 0$（极值点必为驻点）

（必要）$H(x_0,y_0)$ 非不定

（充分）$H(x_0,y_0)$ 正定（极小值）或负定（极大值）

（充分）在 (x_0,y_0) 的某个邻域上 $H(x_0,y_0)$ 半正定（极小值）或半负定（极大值）

- **（条件极值的必要条件）** 设原问题的 Lagrange 函数对应无约束问题 $\max(\min)\ L(\boldsymbol{x},\boldsymbol{\lambda})$，则原问题的每个解必然对应着无约束问题的一个驻点. 即答案必然在无约束问题的驻点中

2.1.3 几何应用

- 平面及其法线（直线及其法平面）公式
  记 $P_0(x_0,y_0,z_0)$ 处法线方向向量为 $\vec{n} = (A,B,C)$，则法线为

$$\frac{x-x_0}{A} = \frac{y-y_0}{B} = \frac{z-z_0}{C}$$

法平面使得 $\vec{P_0P}$ 与 $\vec{n}$ 垂直，即点积为零

$$A(x-x_0) + B(y-y_0) + C(z-z_0) = 0$$

因此，只要求出法线方向向量即可快速表达法线及切平面

- 空间曲面的表达方式及法向量

- ◦ 空间曲面可以看做是二维空间到三维欧氏空间投影的像，故可以用以 x, y 为自变量确定 z 的隐函数方程表示

$$F(x,y,z)=0, \vec{n}=\mathrm{grad}F(P_0)$$

由于沿任何**在该点切平面内的**方向向量都使函数值保持不变，故均与梯度垂直，即梯度为法向量，表示函数 $F(x,y,z)$ 变化最快的方向.

- ◦ 借助辅助变量 $P=(u,v)$ 表示

$$S:\begin{cases} x=x(u,v) \\ y=y(u,v) \\ z=z(u,v) \end{cases} \quad \text{注意：这与 } F(x,y,z)=0 \text{ 完全不同}$$

化曲为平，切平面的参数方程：$\begin{pmatrix}\Delta x\\ \Delta y\\ \Delta z\end{pmatrix}=\begin{pmatrix}x'_u & x'_v\\ y'_u & y'_v\\ z'_u & z'_v\end{pmatrix}_{P_0}\begin{pmatrix}\Delta u\\ \Delta v\end{pmatrix}$

- 空间曲线的表达方式及切向量

- ◦ $\mathcal{L}:\begin{cases} x=x(t) \\ y=y(t) \\ z=z(y) \end{cases} \quad \vec{n}=(x'(t),y'(t),z'(t))$

- ◦ $\mathcal{L}:\begin{cases} F(x,y,z)=0 \\ G(x,y,z)=0 \end{cases}$ 曲线可以视作**两空间曲面的交**，其**切线**为两**切平面的交**，即**垂直于**两个**法向量**，因而可用叉积确定 $\vec{n}=\mathrm{grad}F(P_0)\times\mathrm{grad}G(P_0)$

2.1.4　例题

课本 p.66-3(4)

以下方程确定了 $z=f(x,y)$：$\begin{cases} x=u\cos v \ (\bigstar) \\ y=u\sin v \ (\text{☆}) \\ z=v \end{cases}$，计算：$\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial x\partial y}$.

解.

将 x, y 视为自变量，将 u, v 视作 $u=\boldsymbol{u}(x,y), v=\boldsymbol{v}(x,y)$，则

$$z=\boldsymbol{z}(u,v)=\boldsymbol{z}(\boldsymbol{u}(x,y),\boldsymbol{v}(x,y))$$
$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x}=\frac{\partial v}{\partial x}$$

对 (★) 与 (☆) 两边依次求 x,y 的偏导

$$1=\frac{\partial u}{\partial x}\cos v-u\frac{\partial v}{\partial x}\sin v$$

$$0=\frac{\partial u}{\partial x}\sin v+u\frac{\partial v}{\partial x}\cos v$$

$$0=\frac{\partial u}{\partial y}\cos v-u\frac{\partial v}{\partial y}\sin v$$

$$1=\frac{\partial u}{\partial y}\sin v+u\frac{\partial v}{\partial y}\cos v$$

从而 $\dfrac{\partial v}{\partial x}=-\dfrac{\sin v}{u},\ \dfrac{\partial u}{\partial x}=\cos v,\ \dfrac{\partial v}{\partial y}=\dfrac{\cos v}{u},\ \dfrac{\partial u}{\partial y}=\sin v$

$$\frac{\partial^2 z}{\partial x^2}=\frac{\partial}{\partial x}\left(-\frac{\sin v}{u}\right)=\frac{\sin v}{u^2}\frac{\partial u}{\partial x}-\frac{\cos v}{u}\frac{\partial v}{\partial x}=\frac{\sin 2v}{u^2}$$

$$\frac{\partial^2 x}{\partial x\partial y}=\frac{\partial}{\partial y}\left(-\frac{\sin v}{u}\right)=\frac{\sin v}{u^2}\frac{\partial u}{\partial y}-\frac{\cos v}{u}\frac{\partial v}{\partial y}=-\frac{\cos 2v}{u^2}$$

注：最后仍可用 u,v 表示.

2.2　（广义）含参积分

2.2.1　累次积分

- $f(x,y)$ 在有界闭域 $\Omega\in\mathbb{R}^2$ 上连续 $\implies$ $f(x,y)$ 在 Ω 上**一致连续**.
- **（积分号下可求极限）**$f(x,y)$ 为矩形有界闭域 D 上的**连续函数**时，积分运算和极限运算可以交换顺序.
- **（积分号下可求导）**设 $f(x)$ 与 $\frac{\partial f}{\partial y}$ 在 $D=[a,b]\times[c,d]$ 上**连续**，则
$$\frac{\mathrm{d}}{\mathrm{d}y}\left[\int_a^b f(x,y)\mathrm{d}x\right]=\int_a^b\left[\frac{\partial f}{\partial y}(x,y)\right]\mathrm{d}x$$
若要将 b 改为 $+\infty$，则需等式右侧的积分关于 $y\in[c,d]$ 一致收敛.
- **（积分号下求积分公式）**$f(x,y)$ 在 $D=[a,b]\times[c,d]$ 上**连续**，则
$$\int_c^d\mathrm{d}y\int_a^b f(x,y)\mathrm{d}x=\int_a^b\mathrm{d}x\int_c^d f(x,y)\mathrm{d}y$$
若要将 b 改为 $+\infty$，则需 $\int_a^{+\infty}f(x,y)\mathrm{d}x$ 关于 $y\in[c,d]$ 一致收敛.

2.2.2　广义积分

- **（连续性定理）**设 $f(x)$ 在 $[a,\infty)\times K\subset\mathbb{R}^2$，其中 K 为某一区间. 那么，若广义含参积分
$$J(y)=\int_a^{+\infty}f(x,y)\mathrm{d}x$$
关于 $y\in K$ 一致收敛，则 $J(y)$ 在区间 K 上连续.

2.2.3 例题

课本 p.110-5(2)

$$\int_0^1 \frac{x^b - x^a}{\ln x}\sin\left(\ln\frac{1}{x}\right)$$

2.3 重积分

2.3.1 基本定理

- 若 $f(x,y)$ 在闭矩形域 K 上可积，则它在 K 上有界.
- $D \subset \mathbb{R}^2$ 为有界集合，则 D 有面积 $\iff$ ∂D 为零测集.
  一个平面集合不可求面积与面积为零是两回事.
- **（平面变换的面积公式）**当 $(u,v) \mapsto (x(u,v), y(u,v))$ **可微、正则且一一对应**时，面积微元
  $$\mathrm{d}x\mathrm{d}y = \left|\frac{D(x,y)}{D(u,v)}\right|\mathrm{d}u\mathrm{d}v$$
  体积微元的变换类似.
- 球坐标系体积微元 $\mathrm{d}x\mathrm{d}y\mathrm{d}z = r^2\sin\varphi\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi$，$\varphi$ 为 $\vec{r}$ 与 z 轴的夹角.
  柱坐标系体积微元 $\mathrm{d}x\mathrm{d}y\mathrm{d}z = r\ \mathrm{d}r\mathrm{d}\theta\mathrm{d}z$
- **Euler-Poisson 积分**
  $$\begin{aligned} I &= \int_0^{+\infty} e^{-t^2}\mathrm{d}t \\ I^2 &= \int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y \\ &= \int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^{+\infty} re^{-r^2}\mathrm{d}r \quad \text{（注意 } x,y>0 \text{ 对应的 } \theta \text{ 范围）} \\ &= \frac{\pi}{2}\left(-\frac{1}{2}e^{-r^2}\right)\Bigg|_0^{+\infty} = \frac{\pi}{4} \\ I &= \frac{\sqrt{\pi}}{2} \end{aligned}$$

2.4 场论初步 (Green, Gauss & Stokes)

2.4.1 曲线积分与曲面积分

- **（第一型曲线积分）**设曲线 C 有正则参数表达 $r(t) = (x(t), y(t), z(t)), t \in [a,b]$，则
  $$\int_C f(x,y,z)\mathrm{d}l \triangleq \int_a^b f(r(t))\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}$$

- **（第二型曲线积分）**设向量场 $\mathbf{F}(x,y,z)=(M(\dots),N(\dots),P(\dots))$，$C^+$ 有正则表示 $r(t)=(x(t),y(t),z(t)),t:a\to b$，则

$$\begin{aligned}
\int_C \mathbf{F}(r)\cdot \mathrm{d}\vec{r} &\triangleq \int_C \mathbf{F}\cdot\tau \mathrm{d}l && \text{定义}\\
&= \int_C \mathbf{F}(r(t))\cdot r'(t)\mathrm{d}t && \text{计算公式}\\
&= \int_C [Mx'(t)+Ny'(t)+Pz'(t)]\mathrm{d}t && \text{内积展开}\\
&= \int_C M\mathrm{d}x+N\mathrm{d}y+P\mathrm{d}z && \text{上式的缩写}
\end{aligned}$$

- **（第一型曲面积分）**设曲面 S 有正则的参数表示 $r=r(u,v),(u,v)\in D$，则

$$\iint_S f(x,y,z)\mathrm{d}S \triangleq \iint_D f(r(u,v))|r_u\times r_v|\mathrm{d}u\mathrm{d}v$$

  对显式曲面 $z=z(x,y)$，面积微元为 $\sqrt{1+z_x^2+z_y^2}$

- **（第二型曲面积分）**设向量场 $\mathbf{F}(x,y,z)=(M(\dots),N(\dots),P(\dots))$

$$\begin{aligned}
\iint_{S^+}\mathrm{d}\vec{S} &\triangleq \iint_S(\mathbf{F}\cdot\vec{n})\mathrm{d}S\\
&= \iint_D \mathbf{F}(r(u,v))\cdot(r_u\times r_v)\mathrm{d}u\mathrm{d}v\\
&= \iint_{S^+} P\mathrm{d}y\wedge\mathrm{d}z+Q\mathrm{d}z\wedge\mathrm{d}x+R\mathrm{d}x\wedge\mathrm{d}y
\end{aligned}$$

  其中 $\mathrm{d}x\wedge\mathrm{d}y$ 为面积微元 $\mathrm{d}S$ 在 xy 平面上的投影

  现考虑其中一个分量，若 S^+ 正法向向上，D 为 S 在 xy 平面上的投影，则

$$\iint_{S^+} R\mathrm{d}x\wedge\mathrm{d}y=\iint_D R(x,y,z(x,y))\mathrm{d}x\mathrm{d}y$$

  将第二型曲面积分转换为二重积分.

- 为什么 $\mathrm{d}\vec{S}=(r_u\times r_v)\mathrm{d}u\mathrm{d}v$：
  S 在 (u,v) 对应的点处近似于平行四边形的小平面，该平面的两条邻边分别为 $r_u\mathrm{d}u$ 与 $r_v\mathrm{d}v$，$\mathrm{d}\vec{S}$ 即对应该小平面的有向面积

2.4.2 向量场

- **散度 (divergence)**：表征在某点处的单位体积内散发出来的矢量的**通量**

$$\mathrm{div}\mathbf{F}=\nabla\cdot\mathbf{F}=\frac{\partial F_x}{\partial x}+\frac{\partial F_y}{\partial y}+\frac{\partial F_z}{\partial z}$$

- **旋度 (curl)**：表示三维向量场对某一点附近的微元造成的**旋转程度**，旋度向量的方向表示向量场在这一点附近**旋转度最大的环量的旋转轴**

$$\mathrm{rot}\mathbf{F} = \nabla \times \mathbf{F} = (\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y})$$

- **单连通 (simply connected)**：Ω 中的任意一条简单闭曲线都可以**连续地在 Ω 中**收缩成一点.

- 几种特殊的场：
  - **（保守场）** $L(A,B)$ 逐段光滑，则 $\int_{L(A,B)} P\mathrm{d}x + Q\mathrm{d}y$ 与积分路径无关，只与起终点有关.

    等价表述：对 D 内任意一条逐段光滑的封闭曲线 l_+,

    $$\oint_{l_+} P\mathrm{d}x + Q\mathrm{d}y = 0$$

  - **（梯度场）** $\mathbf{F} = \nabla u = (u_x, u_y, u_z)$，且 u 与 $\mathbf{F}$ 的定义域相同.
  - **（无旋场）** $\mathrm{rot}\mathbf{F} = 0$，二维平面上，$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$

  保守场 $\iff$ 梯度场 $\implies$ 无旋场（对二维、三维欧氏空间均成立）
  无旋场 & 平面单连通区域 $\iff$ 保守场

- 设 D 为单连通有界闭域，$\mathrm{d}u = P\mathrm{d}x + Q\mathrm{d}y$，则任两点 $A(x_1, y_1), B(x_2, y_2) \in D$，

$$\int_{L(A,B)} P\mathrm{d}x + Q\mathrm{d}y = u(x,y)\Big|_A^B$$

- **（Green 定理）** 在平面有界闭区域 D 上，对于向量场 $\mathbf{F}$，沿闭路 ∂D^+ 的环量（通量）等于闭域 D 上各点旋度（散度）的积分.
  - 环量 (circulation) 形式

    $$\oint_{\partial D^+} (\mathbf{F} \cdot \vec{\tau})\mathrm{d}l = \iint_D \mathrm{rot}\mathbf{F}\ \mathrm{d}x\mathrm{d}y$$

  - 通量 (flux) 形式

    $$\oint_{\partial D^+} (\mathbf{F} \cdot \vec{n})\mathrm{d}l = \iint_D \mathrm{div}\mathbf{F}\ \mathrm{d}x\mathrm{d}y$$

  更一般的，可以记为

$$\oint_{\partial D^+} P\mathrm{d}x + Q\mathrm{d}y = \iint_D \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right)\mathrm{d}x\mathrm{d}y$$

- （**Gauss 定理**）设 $\Omega \subset \mathbb{R}^3$ 为空间有界闭域，$\mathbf{F} = (P, Q, R)$ 为 Ω 上的连续可微向量场，则

$$\iint_{\partial\Omega^+} (\mathbf{F} \cdot \vec{n}) \mathrm{d}S = \iiint_{\Omega} (\mathrm{div}\mathbf{F}) \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

- （**Stokes 定理**）设 S^+ 是空间中的一个定向曲面，分片正则，∂S^+ 为分段正则的空间闭曲线，S^+ 与 ∂S^+ 定向协调，则

$$\int_{\partial S^+} \mathbf{F} \cdot \mathrm{d}\vec{r} = \iint_{S^+} \mathrm{rot}\mathbf{F} \cdot \mathrm{d}S$$

2.5　常数项级数

2.5.1　非负项级数的收敛性

- （**Cauchy 积分判敛法**）设 $f \in C[1, +\infty]$ **非负递减**，$u_n = f(n)$，则
  $\sum_{n=1}^{+\infty}$ 收敛 $\iff$ $\int_1^{+\infty} f(x)\mathrm{d}x$ 收敛

- 设 $\{u_n\}$ 是非负递减数列，则
  $\sum_{n=1}^{+\infty} u_n$ 与 $\sum_{n=1}^{+\infty} 2^n u_{2^n}$ 敛散性相同

- （**根值判敛法**）若对正项级数 $\{u_n\}$ 有

$$\lim_{n \to +\infty} \sqrt[n]{u_n} = \rho$$

  若 $\rho < 1$，则 $\sum_{n=1}^{+\infty} u_n$ 收敛

  若 $\rho > 1$，则 $\sum_{n=1}^{+\infty} u_n$ 发散

  此外，将极限分别更换为上极限与下极限，也有类似的结果

- （**Raabe 判敛法**）若存在 $\rho > 1$，使得当 n 足够大时，有

$$n\left(\frac{u_n}{u_{n+1}} - 1\right) \geq \rho$$

  则 $\sum_{n=1}^{+\infty} u_n$ 收敛.
  若 n 充分大时，有

$$n\left(\frac{u_n}{u_{n+1}} - 1\right) \leq 1$$

  则 $\sum_{n=1}^{+\infty} u_n$ 发散.

- （**Leibniz 定理**）若交错项级数 $\sum_{n=1}^{+\infty} (-1)^n u_n$ 满足 $\{u_n\}$ 非负递减，则该级数收敛，且 $S \leq u_1$.

2.5.2 任意项级数的收敛性

- （**Dirichlet 判敛法**）若 $\sum_{i=1}^{n} u_n$ 有界，v_n 单调趋近于 0，则级数 $\sum_{n=1}^{+\infty} u_n v_n$ 收敛
- （**Abel 判敛法**）若 $\sum_{i=1}^{n} u_n$ 收敛，v_n 单调有界，则级数 $\sum_{n=1}^{+\infty} u_n v_n$ 收敛、
- 无穷乘积 $\prod_{n=1}^{+\infty}(1+a_n)$ 收敛 $\iff$ $\sum_{i=m}^{+\infty} \ln(1+a_n)$ 收敛
  其中 m 为充分大的正整数

2.5.3 函数项级数的收敛性

- （**Weierstrass**）若存在非负常数项级数，使得集合 I 上，

  $$|u_n(x)| \leq M_n, n=1,2,\ldots;\ x \in I$$

  且 $\sum_{n=1}^{+\infty} M_n$ 收敛，则函数项级数 $\sum_{n=1}^{+\infty} u_n(x)$ 在 I 上一致收敛
- （**Dirichlet 判敛法**）若有：
  - $v_n(x)$ 对任意固定的 $x \in I$ **单调**，且在 I 上**一致趋于 0**
  - 部分和数列 $\sum_{n=1}^{+\infty} u_n(x)$ **一致有界**

  $$\left|\sum_{k=1}^{n} u_k(x)\right| \leq M, n=1,2,\ldots;\ x \in I$$

  则函数级数 $\sum_{n=1}^{+\infty} u_n(x)v_n(x)$ 在 I 上一致收敛
- （**Abel 判敛法**）若有：
  - $v_n(x)$ 对任意固定的 $x \in I$ **单调**，且在 I 上**一致有界**，即

    $$|v_n(x)| \leq M, n=1,2,\ldots;\ x \in I$$

  - $\sum_{n=1}^{+\infty} v_n(x)$ 在 I 上**一致收敛**

  则函数级数 $\sum_{n=1}^{+\infty} u_n(x)v_n(x)$ 在 I 上一致收敛

第三章　线性代数

3.1　线性映射和矩阵

3.1.1　线性映射

定义　<u>线性运算</u>指向量的加法与数乘.

定义　<u>向量空间</u>为带有线性运算的集合 $\mathbb{R}^m$ 被称为向量空间.

定义　映射 $f:\mathbb{R}^n \to \mathbb{R}^m$ 若满足

1. 任意 $\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^n$，都有 $f(\boldsymbol{x}+\boldsymbol{x}') = f(\boldsymbol{x}) + f(\boldsymbol{x}')$
2. 任意 $\boldsymbol{x} \in \mathbb{R}^n, k \in \mathbb{R}$, 都有 $f(k\boldsymbol{x}) = kf(\boldsymbol{x})$

则 f 为从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 <u>线性映射</u>.

定义　从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性映射称为<u>线性变换</u>.

定义　若线性映射 f 有 $f(\boldsymbol{e_i}) = \boldsymbol{a_i}, \boldsymbol{e_i} \in \mathbb{R}^n, \boldsymbol{a_i} \in \mathbb{R}^m$，则矩阵 $A = [\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}]$ 即为标准坐标向量下的 <u>线性映射的表示矩阵</u>，且满足 $A\boldsymbol{e_i} = \boldsymbol{a_i}$.

定理　（线性映射的线性运算）若矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 表示 $\mathbb{R}^n \to \mathbb{R}^m$ 的线性映射，则 $\boldsymbol{A}+\boldsymbol{B}$ 与 $k\boldsymbol{A}$ 也是同样范畴上的线性映射.

定理　$\boldsymbol{AB}$ 表示线性映射 $\boldsymbol{A}, \boldsymbol{B}$ 的复合 $\boldsymbol{A} \circ \boldsymbol{B}$. 需要注意，$\boldsymbol{AB} = 0$ 也不能推出 $\boldsymbol{A} = 0$ 或 $\boldsymbol{B} = 0$.

理解　矩阵乘法 $A\boldsymbol{x}$ 视为这样一种运算：按照 $\boldsymbol{x}$ 指定的系数，将 $\boldsymbol{A}$ 的列进行线性组合.

定义　<u>反对称矩阵</u>：$A = -A^{\mathrm{T}}$，反对称矩阵对角线必定为 0.

定义 若一个上三角矩阵对角线上全为 0，则称为严格上三角矩阵.

定义 将阶梯形矩阵，从下向上消元，并单位化主元，得到的矩阵每个非零行上，主变量为 1 而其他列的元素均为 0，称这个矩阵为 行简化阶梯形矩阵.

定理 对于 $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^m$，$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}=trace(\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})$.

3.1.2 线性方程式组

定理 对于方程组 $A\boldsymbol{x}=\boldsymbol{b}$，将 $[A\ \boldsymbol{b}]$ 简化为阶梯型后，若

1. $[A\ \boldsymbol{b}]$ 阶梯数比 A 多 **1**，则方程组无解（$0\neq 1$）
2. $[A\ \boldsymbol{b}]$ 阶梯数与 A 相等，则方程组有解
   (a) 若阶梯数等于未知数个数，则有唯一解
   (b) 若阶梯数小于未知数个数，则有无穷多组解

定义 $A\boldsymbol{x}=\boldsymbol{b}$ 称为齐次线性方程组，$\vec{0}$ 为其平凡解，除此之外的解称为非平凡解.

3.1.3 可逆矩阵

定义 设 A 为 n 阶方阵，若存在 n 阶方阵 B，使得 $AB=BA=I_n$，则 A 为 可逆矩阵或非奇异矩阵，B 为 A 的逆.

定理 以下命题等价

1. A 可逆
2. 任意 $\boldsymbol{b}\in\mathbb{R}^n$，$A\boldsymbol{x}=\boldsymbol{b}$ 的解唯一
3. 其次方程组 $A\boldsymbol{x}=0$ 仅有零解
4. A 对应的阶梯形矩阵有 n 个主元
5. A 对应的行简化阶梯形矩阵是 I_n
6. A 能表示为有限个初等矩阵的乘积（即消元至行简化阶梯形矩阵的逆过程）

定义 若矩阵 $A=[a_{ij}]_{n\times n}$ 对于 $i=1,2,\ldots,n$ 都有 $|a_{ii}|>\sum_{j\neq i}|a_{ij}|$，则称其为 （行）对角占优矩阵.

定理 对角占优矩阵必然可逆.

定义 若矩阵 A 通过若干初等行变换可以变为矩阵 B，则称 A,B 左相抵. 即存在可逆矩阵 P，使得 $PA=B,A=P^{-1}B$. 所有和 A 相抵的矩阵中，最简单的是其行简化阶梯形，它被称

为 A 的左相抵标准形.

定理 左相抵构成等价关系.

定理 **(Sherman-Morrison)** 设 A 为 n 阶可逆方阵，$\boldsymbol{u},\boldsymbol{v}$ 为 n 阶向量，则 $A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆 $\iff 1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}\neq 0$，且此时

$$(A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1}=A^{-1}-\frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}A^{-1}}{1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}} \tag{3.1}$$

若将 $\boldsymbol{u},\boldsymbol{v}$ 改为 $n\times k$ 的矩阵，则类似地有

$$(A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1}=A^{-1}-A^{-1}\boldsymbol{u}(I_k+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u})^{-1}\boldsymbol{v}^{\mathrm{T}}A^{-1} \tag{3.2}$$

3.1.4 LU 分解

定理 若 n 阶方阵 A 仅通过倍加矩阵做行变化即可化为阶梯形，则存在单位下三角矩阵（主对角线均为 1）L 与上三角矩阵 U，使得 $A=LU$，此即 **LU 分解**.

定义 方阵 A 左上角的 $k\times k$ 块为第 k 个顺序主子阵.

定理 可逆矩阵 A 存在 LU 分解，当且仅当 A 的所有顺序主子阵均可逆，此时 LU 分解唯一.（满足在消元过程中不需要行的调换）

定理 若可逆矩阵 A 存在 LU 分解，则存在对角线均不为 0 的对角矩阵 D、单位下三角矩阵 L、单位上三角矩阵 U，满足 $A=LDU$，且该分解**唯一**. 此即 **LDU 分解**.

定理 若可逆对称阵 A 有 LDU 分解，则 $L=U^{\mathrm{T}}$.

定理 可逆矩阵 A 存在分解 $A=PLU$，P 为置换矩阵，显然该分解不唯一.

技巧 将 A 分解成对称阵 $X=\frac{1}{2}(A+A^{\mathrm{T}})$ 与反对称阵 $Y=\frac{1}{2}(A-A^{\mathrm{T}})$.

3.2 子空间和维数

3.2.1 基本概念

定义 映射 $A:\boldsymbol{x}\mapsto A\boldsymbol{x}\in\mathbb{R}^{m\times n}$ 的像集

$$\mathcal{R}(A)=\{A\boldsymbol{x}|\boldsymbol{x}\in\mathbb{R}^n\}\subseteq\mathbb{R}^m \tag{3.3}$$

显然 $\mathcal{R}(A)$ 必是子空间，即 A 的**列（向量）空间**.

命题　$\mathcal{R}(A) = \mathbb{R}^m \iff A$ 为满射.

定义　映射 $A : \boldsymbol{x} \mapsto A\boldsymbol{x} \in \mathbb{R}^{m\times n}$ 的**原像**

$$\mathcal{N}(A) = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} = \boldsymbol{0}\} \in \mathbb{R}^n \tag{3.4}$$

显然 $\mathcal{N}(A)$ 必是子空间，即 A 的**零空间（也称解空间）**.

命题　$\mathcal{N}(A) = \{\boldsymbol{0}\} \iff A$ 是单射 $\iff \forall \boldsymbol{x}_1 \neq \boldsymbol{x}_2,\ A(\boldsymbol{x}_1 - \boldsymbol{x}_2) \neq 0$.

定义　将 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ 的全体线性组合为 $\mathbb{R}^n$ 的子空间，记为

$$\mathrm{span}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) := \{k_1\boldsymbol{a}_1, \ldots, k_n\boldsymbol{a}_n | k_1, \ldots, k_n \in \mathbb{R}\} \tag{3.5}$$

定理　$A = \begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix} \in \mathbb{R}^{n\times n}$ 可逆 $\iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的一组基.

3.2.2　基和维数

定义　向量组 S 的任一**极大线性无关组**中向量的个数称为 S 的**秩 (rank)**；子空间 $\mathcal{M}$ 的一组基的向量数目称为**维数**，记为 $\dim \mathcal{M}$.

定理　**（基存在定理）**给定 $\mathbb{R}^m$ 的子空间 $\mathcal{M} \neq \{0\}$，则 $\mathcal{M}$ 存在一组基，且基向量个数不超过 m.

定理　**（基扩充定理）**若 $\mathcal{M} \subseteq \mathcal{N}$，则 $\mathcal{M}$ 的任意一组基都能够**扩充**成 $\mathcal{N}$ 的一组基.

定理　对于方阵 $A \in \mathbb{R}^{n\times n}$，$A$ 可逆 $\iff A$ 是单射 $\iff A$ 是满射.

3.2.3　矩阵的秩

定义　执行 Gauss 消去法将矩阵化为阶梯型后，主元所在的列对应主变量的系数，称为**主列**；其他列对应自由变量的系数，称为**自由列**.

定理　若矩阵 $A = \begin{bmatrix}\boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n\end{bmatrix}$ 与 $B = \begin{bmatrix}\boldsymbol{b}_1 & \ldots & \boldsymbol{b}_n\end{bmatrix}$ 左相抵，则

1. 部分组 $\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_r}$ 线性无关，当且仅当 $\boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_r}$ 线性无关.
2. $\boldsymbol{a}_j = k_1\boldsymbol{a}_{i_1} + \cdots + k_r\boldsymbol{a}_{i_r} \iff \boldsymbol{b}_j = k_1\boldsymbol{b}_{i_1} + \cdots + k_r\boldsymbol{b}_{i_r}$

二级结论

- $\operatorname{rank}(A)+\operatorname{rank}(B)=\operatorname{rank}\left(\begin{bmatrix}A & O\\ O & B\end{bmatrix}\right)\leq \operatorname{rank}\left(\begin{bmatrix}A & X\\ O & B\end{bmatrix}\right)$. 因而当 A,B 可逆时，上述两个分块矩阵都可逆
- $\operatorname{rank}(A+B)\leq \operatorname{rank}(A)+\operatorname{rank}(B)$
- $\operatorname{rank}(AB)\leq \min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$
- 反对称矩阵的秩必定是偶数
- $\mathcal{R}(AB)\subseteq\mathcal{R}(A)$

定义　$\mathcal{R}(A^{\mathrm{T}})$ 为 A 的**行（向量）空间**；$\operatorname{rank}(A)=\operatorname{rank}(A^{\mathrm{T}})$；$\operatorname{rank}(A)=n$ 定义为**列满秩**，同理有**行满秩**.

定理　对于 $A\in\mathbb{R}^{m\times n}$，

- A 是满射 $\iff$ A 行满秩
- A 是单射 $\iff$ A 列满秩

定义　通过行列初等变换，能将矩阵 A 变换为**相抵标准型** $\begin{bmatrix}I_r & O\\ O & O\end{bmatrix}$.

3.2.4　线性方程组的解

方法　（求零空间的一组基）矩阵 $A\in\mathbb{R}^{m\times n}$，求其行简化阶梯形，考虑依次将各个自由元取为 1，同时维持其他自由元为 0（即只让其中一个自由列在线性组合中系数非零），由于主列都是单位向量，从而容易根据自由列的系数计算出零解. 这样求出的解（由于自由元的取值特点）构成零空间的一组基，继而有 $\dim\mathcal{N}(A)=n-\operatorname{rank}(A)$.

3.3　内积和正交性

3.3.1　基本概念

定义　$\boldsymbol{b}$ 向 $\boldsymbol{a}$ 的**垂直投影**为 $\hat{x}=\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$；向量 $\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}$ 称为向量 $\boldsymbol{b}$ 向直线 $\operatorname{span}(\boldsymbol{a})$ 的投影.

定理　**（Cauchy-Schwarz 不等式）** $|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}|\leq\|a\|\|b\|$.

定义　设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，若它的一组基是**正交向量组**，则称之为 $\mathcal{M}$ 的一组**正交基**；若是**正交单位向量组**，则称为 $\mathcal{M}$ 的**标准正交基**.

定义 **（Gram-Schmidt 正交化）**从 $\mathcal{M}$ 的任意一组基 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ 出发，执行如下操作

$$\widetilde{\boldsymbol{q}}_k = \boldsymbol{a}_k - \sum_{j=1}^{k-1} \frac{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \boldsymbol{a}_k}{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \widetilde{\boldsymbol{q}}_j} \widetilde{\boldsymbol{q}}_j \quad (k = 1, 2, \ldots, r) \tag{3.6}$$

即得到一组正交基，最后再通过单位化，得到一组标准正交基.

3.3.2 正交矩阵和 QR 分解

定义 若方阵 Q 满足 $Q^{\mathrm{T}}Q = I_n$，则称 Q 为 n 阶**正交矩阵**；Q 的行、列向量各自构成 $\mathbb{R}^n$ 的一组标准正交基；多个正交矩阵的积也为正交矩阵.

定理 方阵 Q 的以下叙述等价

1. Q 是正交矩阵
2. Q 为**保距变换**，即 $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$
3. Q 为**保内积变换**，即 $Q\boldsymbol{x} \cdot Q\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}$

保距变换一定也是**保角变换**.

定义 **（Givens 变换）**在 $\boldsymbol{e}_i - \boldsymbol{e}_j$ 平面上转角 θ 的旋转变换的矩阵为

$$\begin{bmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & \cos\theta & & & -\sin\theta & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & \sin\theta & & & \cos\theta & \\ & & & & & & 1 \\ & & & & & & & \ddots \end{bmatrix} \tag{3.7}$$

定义 **（Householder 变换）**关于 $\mathbb{R}^n$ 中的单位法向量所确定的超平面 $\mathcal{N}(\boldsymbol{v}^{\mathrm{T}})$ 进行反射变换的矩阵为

$$H_{\boldsymbol{v}} = I_n - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}} \tag{3.8}$$

不难注意到 $\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$ 是 $\boldsymbol{w}$ 向 $\mathrm{span}(\boldsymbol{v})$ 的投影，因此 $H_{\boldsymbol{v}}$ 的效果是将 $\boldsymbol{w}$ 中与 $\boldsymbol{v}$ 共线的成分反向.

定义 **（QR 分解）**设 $A = \begin{bmatrix} \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n \end{bmatrix}$ 为 n 可逆矩阵. 在对 $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ 进行 Gram-Schmidt 正交化过程的中，本质上是进行多次的列初等变换，故可以借此对 A 进行表达

$$\boldsymbol{a}_k = \widetilde{\boldsymbol{q}}_k + \sum_{j=1}^{k-1} \frac{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \boldsymbol{a}_k}{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \widetilde{\boldsymbol{q}}_j} \widetilde{\boldsymbol{q}}_j \quad (k = 1, 2, \ldots, r) \tag{3.9}$$

$$A = \widetilde{Q}\widetilde{R} = \begin{bmatrix} \widetilde{\boldsymbol{q}}_1 & \dots & \widetilde{\boldsymbol{q}}_n \end{bmatrix} \begin{bmatrix} 1 & \frac{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_2}{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\widetilde{\boldsymbol{q}}_1} & \cdots & \frac{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_n}{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\widetilde{\boldsymbol{q}}_1} \\ & 1 & \ddots & \vdots \\ & & \ddots & \frac{\widetilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\boldsymbol{a}_n}{\widetilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\widetilde{\boldsymbol{q}}_{n-1}} \\ & & & 1 \end{bmatrix} \tag{3.10}$$

可以进一步将 Q 单位化，得到

$$A = Q\ \mathrm{diag}(\|\widetilde{\boldsymbol{q}}_i\|)\widetilde{R} = QR \tag{3.11}$$

定理　设 A 为 n 阶可逆矩阵，则存在唯一的分解 $A = QR$，其中 Q 为正交矩阵，R 为对角元均为正数的上三角矩阵.

定义　若矩阵 Q 满足 $Q^{\mathrm{T}}Q = I_n$，则称之为 列正交矩阵.

定理　对于 $m \times n$ 的矩阵 A，其中 $m \geq n$，则

1. （简化 QR 分解）存在 $m \times n$ 列正交矩阵 Q_1 和具有非负对角元的 n 阶上三角矩阵 R_1，使得 $A = Q_1R_1$
2. （QR 分解）进一步的，存在 m 阶正交矩阵 Q 和 $m \times n$ 矩阵 R，使得 $A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ O \end{bmatrix}$

3.3.3　正交投影

定义　（子空间正交）若子空间 $\mathcal{M}$ 的任意向量与 $\mathcal{N}$ 中的任意向量都正交，则称为 $\mathcal{M}$ 与 $\mathcal{N}$ 正交，记为 $\mathcal{M} \perp \mathcal{N}$.

若 $\mathcal{M},\mathcal{N} \in \mathbb{R}^n, \mathcal{M} \perp \mathcal{N}, \mathcal{M} \cup \mathcal{N} = \mathbb{R}^n$，则称 $\mathcal{N}$ 为 $\mathcal{M}$ 的正交补，记为 $\mathcal{N} = \mathcal{M}^{\perp}$.

定理　对于 $A \in \mathbb{R}^{m\times n}$，有

1. $\mathcal{R}(A^{\mathrm{T}})^{\perp} = \mathcal{N}(A), \mathcal{R}(A)^{\perp} = \mathcal{N}(A^{\mathrm{T}})$（矩阵导出的四个子空间的关系）
2. $\mathcal{R}(A^{\mathrm{T}}A) = \mathcal{R}(A^{\mathrm{T}}), \mathcal{N}(A^{\mathrm{T}}A) = \mathcal{N}(A)$

TBD　正交投影

3.4　行列式

3.4.1　行列式函数

定义　定义在全体 n 阶方阵上的函数 δ，如果满足如下性质：

1. 列多线性性：$\delta(\cdots, k\boldsymbol{a}_i + k'\boldsymbol{a}_i', \cdots) = k\delta(\cdots, \boldsymbol{a}_i, \cdots) + k'\delta(\cdots, \boldsymbol{a}_i', \cdots)$
2. 列反对称性：交换任意两列，函数符号反转
3. 单位化条件：$\delta(I_n) = 1$

则称 δ 为一个 n 阶<u>**行列式函数**</u>，该函数存在且唯一. 容易证明 $\det(AB) = \det(A)\det(B)$.

定义　<u>**Vandermonde 矩阵**</u>及其行列式

$$\begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{vmatrix} = \prod_{1\le j<i\le n}(\lambda_i - \lambda_j) \tag{3.12}$$

3.4.2　行列式的展开式

定义　给定 n 阶方阵 A，令 $A\binom{i}{j}$ 表示从 A 中划去第 i 行第 j 列后得到的 $n-1$ 阶方阵，称 $M_{ij} = \det\left(A\binom{i}{j}\right)$ 为元素 a_{ij} 的<u>**余子式**</u>；而 $C_{ij} = (-1)^{i+j}M_{ij}$ 为元素 a_{ij} 的<u>**代数余子式**</u>.

定理　行列式按第一列展开 $\det(A) = a_{11}C_{11} + \cdots + a_{n1}C_{n1}$

定理　令 A 的第 i 列为 $\boldsymbol{a}_i$，记第 j 列元素的代数余子式组成的向量 $\boldsymbol{c}_j = \begin{bmatrix} C_{1j} \\ \vdots \\ C_{nj} \end{bmatrix}$，则

$$\boldsymbol{a}_{j'}^{\mathrm{T}}\boldsymbol{c}_j = \begin{cases} 0, & j \neq j' \\ \det(A), & j = j' \end{cases} \tag{3.13}$$

定义　对矩阵 $A = [a_{ij}]$，记 $C = [C_{ij}]_{n\times n}$，则 C^{T} 为 A 的<u>**伴随矩阵**</u>，且 $C^{\mathrm{T}}A = \det(A)I_n$.

二级结论　若 A 不可逆，则其伴随矩阵秩为 0 或 1.

3.5　特征值和特征向量

引子　考虑问题 $A\boldsymbol{x} = \lambda\boldsymbol{x}$，即求 $(A - \lambda I_n)\boldsymbol{x} = 0$ 的解，即求 $\det(A - \lambda I_n)$ 的解，即求某个多项式的 0 解. 由代数学基本定理知，在复数域上恰有 n 个根（计算重数）. 因此，接下来考虑的矩阵都是**复矩阵**.

定义　给定方阵 A 若对 $\lambda \in \mathbb{C}$，存在<u>非零</u>向量 $\boldsymbol{x} \in \mathbb{C}^n$，使得 $A\boldsymbol{x} = \lambda x$，则称 λ 为 <u>**特征值**</u>，$\boldsymbol{x}$ 为 <u>**特征向量**</u>. 以 λ 为自变量，则多项式

$$p_A(\lambda) = \det(\lambda I_n - A) \tag{3.14}$$

为矩阵 A 的**特征多项式**.

定理

1. λ_0 是 A 的特征值 $\iff p_A(\lambda_0)=0$
2. $\boldsymbol{x}_0$ 是 A 的特征向量 $\iff \boldsymbol{x}_0\in\mathcal{N}(\lambda_0 I_n - A)$
3. 上（下）三角矩阵的对角线元素即为全部的特征值

定义 给定 n 阶方阵 A 及 A 的一个特征值 $\lambda_0\in\mathbb{C}$,若 λ_0 是 $p_A(\lambda)$ 的 n_0 重根,则称 n_0 为 λ_0 作为 A 的特征值的**代数重数(简称重数)**,称 λ_0 是 A 的 n_0 重特征值($n_0=1$ 时称为**单特征值**).

定理 若 λ_0 是 A 的非实数特征值，则 $\overline{\lambda}_0$ 也是 A 的特征值，且 λ_0 与 $\overline{\lambda}_0$ 的重数相同.

定理 （**Vieta 定理**）对于一元 n 次多项式 $p(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$，其 n 个根 $x_1,\ldots,x_n$ 满足

$$-\frac{a_1}{a_0}=x_1+x_2+\cdots+x_n \tag{3.15}$$

$$(-1)^k\frac{a_k}{a_0}=\sum_{1\le i_1<\cdots<i_k\le n}x_{i_1}x_{i_2}\cdots x_{i_k} \tag{3.16}$$

$$(-1)^n\frac{a_n}{a_0}=x_1x_2\cdots x_n \tag{3.17}$$

即 $a_0(x-x_1)(x-x_2)\cdots(x-x_n)$ 的展开式与 $p(x)$ 完全相等. 故有

$$\lambda_1+\lambda_2+\cdots+\lambda_n=\operatorname{trace}(A) \tag{3.18}$$

第四章　复变函数

4.1　复变函数

4.1.1　函数的定义

给定 $G \in \mathbb{C}$ 及从 G 到 $\mathbb{C}$ 的对应法则 f，满足 $\forall z = x + iy \in G$，都有一个或多个 $\omega = u + iv \in \mathbb{C}$ 与之对应，则称 ω 为关于 z 的函数.

4.1.2　极限的定义

设 $\omega = f(z)$ 在 $B_\varphi^*(z_0) \triangleq \{z \in \mathbb{C} \big| 0 < |z - z_0| < \rho\}$ 上有定义，若 $\exists A \in \mathbb{C}, \forall \varepsilon > 0, \exists \delta = \delta(\varepsilon) > 0, \mathrm{s.t.} 0 < |z - z_0| < \delta \to |f(z) - A| < \varepsilon$，则称 $z \to z_0$ 时，$f(z)$ 以 A 为极限. **注意这意味着沿任意路径逼近得到的极限都是** A.

4.1.3　连续性的定义

若 f 在实心邻域 B_φ 上有定义，且 $\lim\limits_{z \to z_0} f(z) = f(z_0)$，则称 $f(z)$ 在 z_0 连续.

1. 连续函数的和、差、积、商仍是连续函数
2. 设 $g = g(z)$ 连续，$\omega = f(g)$ 在 $g_0 = g(z_0)$ 处连续，则 $\omega = (g \circ f)(z)$ 在 z_0 连续.
3. 闭区域 $\overline{\mathscr{D}}$ 上的连续函数一定能在 $\overline{\mathscr{D}}$ 上取到最小（大）模长.

4.1.4　区域、曲线的定义

点集 $\mathscr{D}$ 称为一个**区域**，如果它是一个开集 且它连通. 没有重点的连续曲线称为**简单曲线**或 **Jordan 曲线**；若仅有曲线起点与终点重合，则为**简单闭曲线**，曲线以逆时针为正向. 若一个区域内任意一条闭曲线的内部都属于该区域，那么该区域为**单连通域**.

4.1.5　复数的辐角

对于 $z \in \mathbb{C} \backslash \{0\}$，定义$\arg(z) \in (-\pi, \pi]$ 为辐角的主值；$\mathrm{Arg}(z) = \arg(z) + 2k\pi, k \in \mathbb{Z}$ 为负数的辐角函数.

4.2 解析函数

4.2.1 导数的定义

若极限 $\lim\limits_{z\to z_0}\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}$ 存在且有限，则 $f(z)$ 在 z_0 可导.

4.2.2 可微与微分

若 $\omega=f(z)$ 在 z_0 的某个邻域内有表达式

$$\Delta\omega=f(z_0+\Delta z)-f(z_0)=\mathscr{A}\Delta z+\rho(\Delta z)\Delta z \tag{4.1}$$

$$\mathscr{A}\in\mathbb{C},\ \lim_{|z_0|\to 0}\rho(z)=0 \tag{4.2}$$

则称 $f(z)$ 在 z_0 可微，$\mathscr{A}\Delta z$ 称为 $f(z)$ 在 z_0 的微分，记为 $\mathrm{d}\omega=\mathscr{A}\Delta z=f'(z_0)\mathrm{d}z$.

函数在一点可导和可微是等价的.

4.2.3 解析函数（或全纯函数、正则函数）

$\forall z_0\in\mathbb{C}$，若 $\omega=f(z)$ 在 z_0 的**某邻域内处处可导**，则称 $f(z)$ 在 z_0 处解析，z_0 为解析点；否则 z_0 为 $f(z)$ 的奇点. 注意可能函数在某一点可导，在其任意邻域上均不可导. 在整个 $\mathbb{C}$ 上都解析的函数称为<u>整函数</u>.

★Lemma

1. 两个解析函数的和、差、积、商仍是解析函数
2. 设 $g=g(z)$ 在 $\mathscr{D}$ 上解析，$\omega=f(g)$ 在 $g(\mathscr{D})$ 上解析，则 $\omega=(f\circ g)(z)$ 在 $\mathscr{D}$ 上解析.

4.2.4 函数可导的充要条件

Cauchy-Riemann 方程：设 $z=x+iy\in D,\omega=f(z)=u(x,y)+iv(x,y)$，则 $f(z)$ 在 z 可导的<u>充要条件</u> 是 <u>$u(x,y)$ 与 $v(x,y)$ 在 (x,y) 可微</u> 且

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x} \tag{4.3}$$

导数为

$$f'(z)=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=-i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y} \tag{4.4}$$

形式导数：用形式变元 $z,\overline{z}$ 表示 x,y，则 $\begin{cases}x=\frac{1}{2}(z+\overline{z})\\ y=\frac{1}{2i}(z-\overline{z})\end{cases}$，则 $\begin{cases}x_z=x_{\overline{z}}=\frac{1}{2}\\ y_z=-y_{\overline{z}}=\frac{1}{2i}\end{cases}$ 进而可以求出 $u_z,u_{\overline{z}},v_z,v_{\overline{z}}$，可以发现

$$f_z=u_z+iv_z=\frac{1}{2}(u_x+u_y)+\frac{i}{2}(v_x-u_y) \tag{4.5}$$

$$f_{\overline{z}} = u_{\overline{z}} + iv_{\overline{z}} = \frac{1}{2}(u_x - v_y) + \frac{i}{2}(u_y + v_x) \tag{4.6}$$

注意到 $f_{\overline{z}} = 0 \iff f$ 满足柯西-黎曼方程.

4.2.5 初等函数

指数函数

定义指数函数为

$$e^z = \exp(z) = e^x(\cos y + i\sin y) \tag{4.7}$$

该函数为整函数，$\exp z' = \exp z$，周期为 $2k\pi i$，值域为 $\mathbb{C}\backslash\{0\}$，满足 $\exp(z_1+z_2) = \exp z_1 \exp z_2$，$|\exp z| = e^x$，$\arg(\exp z) = y$.

对数函数

定义为指数函数的反函数，即

$$\ln(z) = \ln|z| + i\arg(z), \mathrm{Ln}(z) = \ln(z) + 2k\pi i \tag{4.8}$$

，导数 $\frac{\mathrm{d}\ln z}{\mathrm{d}z} = \frac{1}{z}$.

幂函数

定义幂函数为

$$z^b \triangleq \exp(b \cdot \mathrm{Ln}z) = e^{b\ln z} \cdot e^{2bk\pi i}, k \in \mathbb{Z} \tag{4.9}$$

多值性讨论

- $b \in \mathbb{Z}$, $e^{2bk\pi i} \equiv 1$，单值
- $b \in \mathbb{Q}\backslash\mathbb{Z}$, $e^{2k\pi i\frac{m}{n}}$，n 值
- $b \in \mathbb{R}\backslash\mathbb{Q}$, $e^{2(bk)\pi i}$ ，无穷多值
- $b \in \mathbb{C}\backslash\mathbb{R}$, $e^{2(\alpha+i\beta)k\pi i} = e^{-2\beta k\pi} \cdot e^{2\alpha k\pi i}$，仅模长部分便有无穷多值

取同一个第 k 支的情况下有 $(z^b)' = bz^{b-1}, z^{a+b} = z^a \cdot z^b, z^{-a} = \frac{1}{z^a}$

三角函数

根据指数函数的定义进行“逆推”，有

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \tag{4.10}$$

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} \tag{4.11}$$

$\sin z$ 与 $\cos z$ 都是整函数，导数性质、和角公式与实数下相同. 注意该函数**无界**.

4.3 复变函数积分

4.3.1 积分的计算

设 $f(z)$ 沿 C 连续，弧上第 k 段取点 $\zeta_k=\xi_k+i\eta_k$，记 $\delta=\max\{|\Delta x_k+i\Delta y_k|\}$.

$$\begin{aligned}I_n&=\sum_{i=1}^{n}[u(\xi_k+i\eta_k)+iv(\xi_k+i\eta_k)](\Delta x_k+i\Delta y_k)\\&=\sum_{k=1}^{n}[(u\Delta x_k-v\Delta y_k)+i(u\Delta y_k+v\Delta x_k)]\\&\xrightarrow[n\to\infty]{\delta\to 0}\int_c[(u\mathrm{d}x-v\mathrm{d}y)+i(u\mathrm{d}y+v\mathrm{d}x)]\end{aligned}\tag{4.12}$$

4.3.2 积分的性质

积分的复共轭：

$$\overline{\int_c f(z)\mathrm{d}z}=\int_c\overline{f}(z)\overline{\mathrm{d}z}\tag{4.13}$$

若曲线 C 上有 $\left|f(z)\right|\leq M(<+\infty)$，$C$ 的弧长为 L，则

$$\left|\int_c f(z)\mathrm{d}z\right|\leq ML\tag{4.14}$$

（积分控制）.

一个重要的积分 $(n\in\mathbb{Z})$

$$\begin{aligned}I_n&=\oint_c\frac{\mathrm{d}z}{(z-z_0)^{n+1}}=\frac{i}{R^n}\int_0^{2\pi}e^{in\theta}\mathrm{d}\theta\\&=\frac{i}{R^n}\int_0^{2\pi}(\cos n\theta-i\sin n\theta)\mathrm{d}\theta=2\pi i[n=0]\end{aligned}\tag{4.15}$$

4.3.3 柯西-古萨 (Cauchy-Goarsat) 定理

若函数 $f(z)$ 在**单连通区域** $\mathscr{B}$ 内**处处解析**，那么函数 $f(z)$ 沿 $\mathscr{B}$ 内的任何一条封闭曲线 C 的积分为 0.

一种不严谨的理解：基于(4.12)，使用格林公式，以实部为例，变为 $\iint_D(-u_y-v_x)\mathrm{d}x\mathrm{d}y$. 若处处解析，则处处满足柯西-黎曼方程，故 $u_y=-v_x$，因此实部被积变量恒为 0. 虚部同理.（由于 u,v 不一定有一阶**连续**偏导数，故不一能使用 Green 公式.）

4.3.4 复合闭路定理

连续变形原理

在区域内一个解析函数沿闭曲线的积分，不因闭曲线在区域内做连续变形而改变它的值.

复合闭路定理

设 Jordan 闭曲线 $\gamma=\gamma_0+\gamma_1^-+\cdots+\gamma_n^-$ 围成一个 (n+1)-连通区域 $\mathscr{D}$，$\omega=f(z)$ 在其上解析，在 $\overline{\mathscr{D}}$ 上连续，则 $\oint_\gamma f(z)\mathrm{d}z=0$.

4.3.5 原函数与不定积分

设 $\omega=f(z)$ 在单连通域 $\mathscr{D}$ 上解析，定义原函数

$$F(z)=\int_{z_0}^{z} f(\zeta)\mathrm{d}\zeta$$

，则 $F(z)$ 在 $\mathscr{D}$ 上解析，且 $F'(z)=f(z),\forall z\in\mathscr{D}$. 原函数可以有多个，但它们的差恒为常数.

Newton-Leibniz 定理

设 $\omega=f(z)$ 在单连通域 $\mathscr{D}$ 上解析，$G(z)$ 为 $f(z)$ 在 $\mathscr{D}$ 上的一个原函数，则

$$\int_{z_0}^{z_1} f(z)\mathrm{d}z=G(z)\Big|_{z_0}^{z_1} \tag{4.16}$$

分部积分公式

设 $\omega=f(z),\sigma=g(z)$ 在单连通域 D 上解析

$$\int_{z_0}^{z_1} f'(z)g(z)\mathrm{d}z=f(z)g(z)\Big|_{z_0}^{z_1}-\int_{z_0}^{z_1} f(z)g'(z)\mathrm{d}z \tag{4.17}$$

三个等价命题

设 $\omega=f(z)$ 在 n-连通区域 D 上解析，则

(1) $\forall C\subseteq D$ 有 $\oint_C f(z)\mathrm{d}z=0$

(2) $f(z)$ 在 D 上有积分路径无关性

(3) $f(z)$ 在 D 上有原函数

等价，且任意一条成立，牛顿-莱布尼茨定理即可使用.

4.3.6 Cauchy 积分公式

设 $\omega=f(z)$ 在单连通域 D 上解析，在 $\overline{D}$ 上连续，则 $\forall z_0\in D,C\subseteq D$

$$f(z_0)=\frac{1}{2\pi i}\oint_{\partial D}\frac{f(z)}{z-z_0}\mathrm{d}z \tag{4.18}$$

$$= \frac{1}{2\pi i} \oint_c \frac{f(z)}{z - z_0} \mathrm{d}z \tag{4.19}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + Re^{i\theta}) \mathrm{d}\theta \tag{4.20}$$

$$= \frac{1}{\pi R^2} \iint_{|z-z_0| \le R} f(z) \mathrm{d}x \mathrm{d}y \tag{4.21}$$

4.3.7 高阶导数

定理：解析函数 $f(z)$ 的**任意阶导数**仍为解析函数，其 n 阶导数满足

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_c \frac{f(z)}{(z - z_0)^{n+1}} \mathrm{d}z \tag{4.22}$$

其中 C 为围绕 z_0 的任意一条**正向**简单闭曲线，且 C 在单连通的解析区域 $\mathscr{D}$ 上.

莫雷拉 (Morera) 定理（柯西定理的逆定理）

若 $f(z)$ 在**单连通区域** $\mathscr{D}$ 内**连续**，且沿 $\mathscr{D}$ 内任意闭合曲线积分为 0（**路径无关**），则 $f(z)$ 在 $\mathscr{D}$ 内解析.（在任意一个解析函数上修改一个点，函数仍然路径无关，因此连续是必要的.）

4.3.8 代数基本定理

$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 (a_n \ne 0)$ 在 $\mathbb{C}$ 上恰有 n 个零点（记重数）.
只要证明 $\forall n \in \mathbb{N}_+, P_n$ 存在零点即可（从而可以不断地分离因式降阶）.

刘维尔（Liouville）定理

一个有界的正函数必然是常函数

证明. 设 $|f(z)| < M$，根据(4.22)

$$|f'(z_0)| \le \frac{1}{2\pi} \oint_{|z-z_0|=R} \frac{M}{R^2} \mathrm{d}l = \frac{M}{R} \xrightarrow{R\to+\infty} 0$$

故 $f(z_0)$ 为常数 □

代数基本定理证明

假设 $P_n(z)$ 在 $\mathbb{C}$ 上没有零点，设 $f(z) = \frac{1}{P_n(z)}$，则 $f(z)$ 为整函数. 在 $|z| \to +\infty$ 时，显然 $|f(z)| \to \frac{1}{a_n |z|^n} \to 0$ 故存在 R 使得 $\forall |z| > R, f(|z|) < 1$，则在 $\mathbb{C}$ 上，

$$|f(z)| \le \max\{1, \max_{|z| \le R}\{|f(z)|\}\}$$

不等式右侧显然不等于 ∞（有界闭域上连续函数有界），故 $f(z)$ 为常函数，$P_n(z)$ 为常数，这与 $a_n \ne 0$ 矛盾. 故 $P_n(z)$ 必有零点.

4.3.9 解析函数与调和函数

调和函数

设 $\varphi = \varphi(x, y) \in C^2(\mathscr{D})$ 且处处有 $\Delta\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} \equiv 0$，则 φ 为 $\mathscr{D}$ 上的调和函数.

设 $\omega = f(z) = u + iv$ 在 $\mathscr{D}$ 上解析，则 u, v 在 $\mathscr{D}$ 上调和. 称上述 u, v 为 $\mathscr{D}$ 上的**共轭调和函数**. 给出 $\mathscr{D}$ 上调和函数 u，找出其共轭调和函数 $\iff$ 找出解析函数 $f(z)$ 使 $Re\ f(z) = u$.

设 u 为**单连通域**上的调和函数，则必然存在 $f(z)$ 使 $Re\ f(z) = u$. 注意多连通情况下不一定正确.

证明. 这样的 $f(z)$ 必定满足 $f'(z) = u_x - iu_y = U(z)$,$U(z)$ 显然是解析的,故 $f(z) = \int_{z_0}^{z} U(z)\mathrm{d}z$ 即为满足条件的函数. □

例：$u = x^3 - 3xy^2$

❝不定积分法

$f'(z) = (3x^2 - 3y^2) - i(-6xy) = 3(x + iy)^2$

$f(z) = \int 3z^2 \mathrm{d}z = z^3 + c = (x + iy)^3 + C = (x^3 - 3xy^2) + i(3x^2y - y^3) + C$

$v(x) = 3x^2y - y^3 - iC$

由于 $f(z) - u = iv$ 为虚数，故 C 必须为纯虚数.

❝偏积分法

$v_y = u_x = 3x^2 - 3y^2, v_x = -u_y = 6xy$

v 先关于 y 积分，即 $v = 3x^2y - y^3 + g(x)$，则 $v_x = 6xy + g'(x) = 6xy \implies g(x) = C$

4.4 级数

4.4.1 复数项级数

复数项序列

设 $\{x_i\}, \{y_i\}$ 为实数序列，则 $z_n = x_n + iy_n$ 即为复数项序列.

级数

定义 $I = \sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} x_n + i\sum_{n=0}^{\infty} y_n$（不论是否收敛都称为级数），级数的部分和定义为 $S_n = \sum_{k=0}^{n} z_k$.

1. 若 $\lim\limits_{n\to\infty} S_n = A \in \mathbb{C}$，则称 I 收敛
2. 若 $\sum_{n=0}^{\infty} |z_n|$ 收敛，则称 I 绝对收敛
3. 若 I 收敛但不绝对收敛，则称 I 条件收敛

复数项级数与常数项级数的关系

$$\lim_{n\to\infty} z_n = A = \alpha + i\beta \iff \lim_{n\to\infty} x_n = \alpha \wedge \lim_{n\to\infty} y_n = \beta$$

$\lim_{n\to\infty} z_n = A = \alpha + i\beta$ 条件收敛或绝对收敛 $\iff \lim_{n\to\infty} x_n \lim_{n\to\infty} y_n$ 均条件收敛或绝对收敛

敛散判别法

(I) Cauchy 根式判别法

- $\sqrt[n]{|z_n|} < q < 1(\forall n > N)$，则 I 绝对收敛
- 只要满足 $\sqrt[n]{|z_n|} \geq q \geq 1$ 的项有无穷多个，则 I 发散
- 若 $\lim_{n\to\infty} \sqrt[n]{|z_n|} = q$，则 $q < 1$ 时 I 绝对收敛，$q > 1$ 时 I 发散. $q = 1$ 时无法确定，例如 $\frac{1}{n}, \frac{(-1)^n}{n}, \frac{1}{n^2}$ 三者收敛情况均不同

(II) D'Alembert 判别法

- $\left|\frac{z_{n+1}}{z_n}\right| < q < 1(\forall n > N)$ 则 I 绝对收敛
- $\left|\frac{z_{n+1}}{z_n}\right| \geq q \geq 1(\forall n > N)$ 则 I 发散
- 若 $\lim_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| = q$，则 $q < 1$ 时绝对收敛，$q > 1$ 时绝对发散，$q = 1$ 时无法判断

(III) Dirichlet 判别法

4.4.2　幂级数

（复变）函数级数

定义 $\mathscr{D}$ 上的函数列 $\{f_n(z)\}_{n=0}^{\infty}$，其级数为 $I = \sum_{n=0}^{\infty} f_n(z)$，当 z 固定时 I 就变成常数项级数.

幂级数

形如 $I(z) = \sum_{n=0}^{\infty} C_n(z-a)^n$ 或 $I(z) = \sum_{n=0}^{\infty} C_n z^n$ 的级数称为幂级数.

1. 若 i 在 z_0 处收敛，则 $\forall z : |z| < |z_0|$，I 在 z 处收敛
2. 若 i 在 z_0 处发散，则 $\forall z : |z| > |z_0|$，I 在 z 处发散

证明. (1)
易知 $\lim_{n\to\infty} |C_n z_0^n|$ 收敛到 0，故其必有 $|C_n z_0^n| < M(\forall n \in \mathbb{N})$，则 $\sum_{n=0}^{\infty} |C_n z_0^n| \left|\frac{z}{z_0}\right|^n$ 有界递增（并因此收敛），因此 $I(z)$ 绝对收敛. □

收敛半径与收敛圆盘

$I=\sum_{n=0}^{\infty}C_nz^n$ 的<u>收敛半径</u> R 定义为

$$\begin{aligned} R &\triangleq \sup\{|z| : I \text{ is convergent at } z\} \\ &= \sup\{|z| : I \text{ is absolutely convergent at } z\} \\ &= \inf\{|z| : I \text{ is divergent at } z\} \end{aligned}$$

（$R=+\infty$ 时函数级数在 $\mathbb{C}$ 上收敛.）定义 $C_R:|z|=R$ 为收敛圆周，$D_R:|z|\le R$ 为<u>收敛圆盘</u>，考虑 C_R 上的点属于内外哪一侧：

- $I(z)$ 在 C_R 上处处发散：$I(z)=\sum z^n$
- $I(z)$ 在 C_R 上部分收敛：$I(z)=\sum\frac{z^n}{n}$，C_R 上仅 $z=1$ 处发散
- $I(z)$ 在 C_R 上全部收敛：$I(z)=\sum\frac{(-1)^{\lfloor\sqrt{n}\rfloor}}{n}z^n$
- $I(z)$ 在 C_R 上全部绝对收敛：$I(z)=\sum\frac{z^n}{n^2}$，一旦有一个点绝对收敛，则整个圆上所有点都绝对收敛

收敛半径的计算

- 若 $\lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right|=\lambda$，则 $R=\frac{1}{\lambda}$
- 若 $\lim\limits_{n\to\infty}\sqrt[n]{|c_n|}=\lambda$ 或者 $\varlimsup\limits_{n\to\infty}\sqrt[n]{|c_n|}=\lambda$，则 $R=\frac{1}{\lambda}$

幂级数的和函数

设 $I=\sum_{n\ge0}c_n(z-a)^n$ 的收敛半径为 R，则在 $D:|z-a|<R$ 上有

(a) 和函数 $f(z)$ 为解析函数

(b) $f(z)$ 能够逐项求导

$$f'(z)=\sum_{n\ge1}nc_n(z-a)^{n-1}$$

(c) $f(z)$ 能够逐项积分

$$F(z)=\sum_{n\ge0}\frac{c_n}{n+1}(z-a)^{n+1}$$

(d) $f(z)$ 在 C_R 上至少存在一个奇点

4.4.3 Taylor 展开式

设 $\omega=f(z)$ 在单连通域 $\mathscr{D}$ 上解析，$z_0\in\mathscr{D}, d=\inf\limits_{z\in\partial\mathscr{D}}|z-z_0|$，则 $\forall z\in B_d(z_0)$，有

$$f(z)=\sum_{n\geq 0}c_n(z-z_0)^n, c_n=\frac{f^{(n)}(z_0)}{n!}=\frac{1}{2\pi i}\oint\frac{f(z)}{(z-z_0)^{n+1}}\mathrm{d}z \tag{4.23}$$

且该展开式唯一.

4.4.4 解析函数的零点

设 $\omega=f(z)$ 在 $\mathscr{D}$ 上解析

1. 若 $f(z_0)=0$，则称 z_0 为 $f(z)$ 的零点
2. 若 $f(z_0)=f'(z_0)=\cdots=f^{(m-1)}(z_0)=0, f^{(m)}(z_0)\neq 0$，则称 z_0 为 m 级零点
3. 若 $f(z_0)=0$，且某个去心领域 $B^*_\delta(z_0)$ 上 $f(z)$ 恒不为 0，则称 z_0 为 $f(z)$ 的孤立零点

定理　z_0 为 $f(z)$ 的 m 级零点 $\iff \exists B_\delta(z_0)$ 及其上的解析函数 $\varphi(z)$，满足

$$f(z)=(z-z_0)^n\varphi(z), \varphi(z_0)\neq 0$$

定理　设 $f(z)$ 在 $\mathscr{D}$ 上解析，则 $f(z)$ 在 $\mathscr{D}$ 上的所有零点都孤立，除非 $f(z)\equiv 0$.

4.4.5 解析函数的唯一性定理

设 $f(z)$ 与 $g(z)$ 在 $\mathscr{D}$ 上解析，且 $a\in D$，若 $\exists\{z_n\}\in D$，满足

1. $z_n\neq a$
2. $\lim\limits_{n\to\infty}z_n=a$
3. $f(z_n)=g(z_n), \forall n\geq 0$

则 $f(z)=g(z), \forall z\in D$

4.4.6 一般常级数

形如 $I=\sum_{n=-\infty}^{\infty}c_n(z-z_0)^n=\sum_{n\geq 0}c_n(z-z_0)^n+\sum_{n\geq 1}c_{-n}\zeta^n=I_++I_-, \zeta=\frac{1}{z-z_0}$ 的级数称为一般常级数.

若 I_+ 的收敛半径为 R_+，I_- 的收敛半径为 R_-，则称 $|\frac{1}{R_-}<|z-z_0|<R_+|$ 为 I 的<u>收敛圆环域</u> $D(z_0,r_1,r_2)$.

4.4.7 洛朗级数

设 $\omega = f(z)$ 在 $D(z_0, r, R)$ 上解析，则 $\forall z \in D(z_0, r, R)$，有

$$f(z) = \sum_n c_n (z - z_0)^n, c_n = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{(z - z_0)^{n+1}} \mathrm{d}\zeta \tag{4.24}$$

其中 C 为 D 上任意环绕的 Jordan 闭曲线. 该展开唯一，称为 **Laurent 级数**.
注意 $f(z)$ 在 z_0 的导数一般不存在，故不能套用高阶导数公式.

4.5 留数

4.5.1 孤立奇点

定义　若 z_0 为 $f(z)$ 的奇点，且在某个邻域 $B_\delta^*(z_0)$ 内 $f(z)$ 解析，则称 z_0 为 $f(z)$ 的孤立奇点. 若 $z_0 \in \mathbb{C}$ 为孤立奇点，且 $f(z)$ 在上述领域中的洛朗级数为 $f(z) = \sum_{-\infty}^{\infty} c_n (z - z_n)^n$，则

(A) z_0 为可去奇点，若级数中不含负幂项，且 $\lim\limits_{z \to z_0} = c_0$. 若补充定义 $f(z_0) = c_0$，则 $f(z)$ 在 z_0 解析，因而以下命题等价：

- z_0 为可取奇点
- $\lim\limits_{z \to z_0} f(z) = A \in \mathbb{C}$
- $f(z)$ 在某邻域 $B_\delta^*(z_0)$ 内有界

(B) z_0 为 m 级极点，若展式中含有有限的负幂项，且最低负幂项为 $c_{-m}(z - z_0)^{-m}, c_{-m} \neq 0$. 若 $\lim\limits_{z \to z_0} f(z) = \infty$，则必定为极点而不是本性奇点. 以下命题等价

- z_0 为 $f(z)$ 的 m 级极点
- 存在某个 $B_0(z_0)$ 上的解析函数 $g(z)$，满足 $f(z) = g(z)(z - z_0)^{-m}$
- $\lim\limits_{z \to z_0} (z - z_0)^n f(z) = A \in \mathbb{C} \backslash \{0\}$
- z_0 为 $\frac{1}{f(z)}$ 的 m 级零点

(C) z_0 为本性奇点，若展式中含有无穷个负幂项. 注意 $\lim\limits_{z \to z_0} f(z)$ 必不存在.
(Weierstress) $\forall A \in \overline{\mathbb{C}}$，都存在 $\{z_n\}_{n=1}^{\infty} \subseteq B_\delta^*(z_0)$，满足 $\lim\limits_{n \to \infty} z_n = z_0$ 且 $\lim\limits_{n \to \infty} f(z_n) = A$

若 $f(z)$ 在 $R < |z| < \infty$ 内解析，则称 ∞ 为 $f(z)$ 的孤立奇点. 设 $\zeta = \frac{1}{z}$，则 $\varphi(\zeta) = \sum_n c_{-n} \zeta^n = f(z)$，若 0 为 $\varphi(\zeta)$ 的本性/可去奇点或 m 级极点，则 ∞ 为 $f(z)$ 的本性/可去奇点或 m 级极点.

4.5.2 留数

设 z_0 为 $f(z)$ 的孤立奇点，$f(z)$ 在 z_0 的洛朗级数为 $f(z)=\sum_{n=-\infty}^{\infty}c_n(z-z_0)^n$，$C$ 为环绕 z_0 的正向简单曲线，则

$$\oint_C f(z)\mathrm{d}z = 2\pi i c_{-1} \tag{4.25}$$

这是由(4.15)直接得到的. 定义留数为

$$\mathrm{Res}[f(z),z_0]=\frac{1}{2\pi i}\oint_C f(z)\mathrm{d}z = c_{-1} \tag{4.26}$$

4.5.3 留数的计算规则

(A) 若 z_0 为 $f(z)$ 的一级极点，则

$$\mathrm{Res}[f(z),z_0]=\lim_{z\to z_0}(z-z_0)f(z)$$

(B) 若 z_0 为 $f(z)$ 的 m 级极点，则

$$\mathrm{Res}[f(z),z_0]=\frac{1}{(m-1)!}\lim_{z\to z_0}\frac{\mathrm{d}^{m-1}}{\mathrm{d}z^{m-1}}\{(z-z_0)^m f(z)\} \tag{4.27}$$

(C) 设 $f(z)=\frac{P(z)}{Q(z)}$，$P(z),Q(z)$ 在 z_0 解析，若 $P(z_0)\neq 0, Q(z_0)=0, Q'(z_0)\neq 0$，则 z_0 为 $f(z)$ 的一级极点，且（可由 (A) 导出）

$$\mathrm{Res}[f(z),z_0]=\frac{P(z_0)}{Q'(z_0)}$$

(D) 若无穷远也为孤立极点，则无穷远处的留数定义为 $\frac{1}{2\pi i}\oint_{C^-}f(z)\mathrm{d}z$，不难发现，此即所有有限点的留数之和的负数

$$\mathrm{Res}[f(z),\infty]=-\mathrm{Res}\left[f\left(\frac{1}{z}\right)\frac{1}{z^2},0\right] \tag{4.28}$$

4.5.4 留数的应用

(A) 形如 $\int_0^{2\pi}R(\cos\theta,\sin\theta)\mathrm{d}\theta$ 的积分，对于 $z=\cos\theta+i\sin\theta$，可以用 z 反求 $\cos\theta,\sin\theta$.

$$\int_0^{2\pi}R(\cos\theta,\sin\theta)\mathrm{d}\theta=\oint_{|z|=1}R\left[\frac{z^2+1}{2z},\frac{z^2-1}{2iz}\right]\frac{\mathrm{d}z}{iz} \tag{4.29}$$

(B) 形如 $\sum_{-\infty}^{\infty}R(x)\mathrm{d}x$ 的积分，其中 $R(x)$ 为**有理函数**，且分母比分子的次数**至少高二次**，且在实轴上没有奇点. 将其延拓到复数域上，并且沿路径 $(-R,0)\to(R,0)\xrightarrow{x^2+y^2=R^2,y\geq 0}(-R,0), R\to+\infty$ 进行积分，设 z_k 为虚部为正数的全部奇点，则

$$\int_{-\infty}^{\infty}R(x)\mathrm{d}x=2\pi i\sum\mathrm{Res}[R(z),z_k] \tag{4.30}$$

(C) 形如 $\sum_{-\infty}^{\infty} R(x)e^{iax}\mathrm{d}x(a>0)$ 的积分，其中 $R(x)$ 为**有理函数**，且分母比分子的次数**至少高一次**，且在实轴上没有奇点. 将其延拓到复数域上，并且沿路径 $(-R,0) \to (R,0) \xrightarrow{x^2+y^2=R^2, y\geq 0} (-R,0), R \to +\infty$ 进行积分，设 z_k 为虚部为正数的全部奇点，则

$$\int_{-\infty}^{\infty} R(x)e^{iax}\mathrm{d}x = 2\pi i \sum \mathrm{Res}[R(z)e^{iaz}, z_k] \tag{4.31}$$

第五章　概率论

5.1　事件的概率

5.1.1　概率的公理化定义

设 Ω 为样本空间，定义事件集类 $\mathscr{J} \subset 2^{\Omega}$. 定义 $P : \mathscr{J} \to \mathbb{R}$ 满足三条公理

1. $P(A) \geq 0, \forall A \in \mathscr{J}$
2. $P(\Omega) = 1$
3. $P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i), \forall A_i \in \mathscr{J}, A_iA_j = \phi, \forall i \neq j$

则称 P 为<u>概率函数</u>，$(\omega, \mathscr{J}, P)$ 为<u>概率空间</u>

错位排序

A_i 表示第 i 个数恰好在原位上，则每个数**都不在**原位的概率为

$$\begin{aligned} P &= 1 - P(A_1 + A_2 + \cdots + A_n) \\ &= 1 - \sum_{i_1 \leq \cdots \leq i_r} (-1)^{r+1} P(A_{i_1} \ldots A_{i_r}) \\ &= 1 - \sum_{r=1}^{n} (-1)^{r+1} \binom{n}{r} \frac{(n-r)!}{n!} (= \frac{1}{r!}) \\ &= 1 - [1 - \frac{1}{2!} + \frac{1}{3!} + \cdots + (-1)^{n+1} \frac{1}{(n)!}] (\to \frac{1}{e}) \end{aligned} \tag{5.1}$$

5.1.2　条件概率

设 $A, B \in \mathscr{J}$，定义 <u>条件概率</u> 为

$$P(A|B) \triangleq \frac{P(AB)}{P(B)} \quad (P(B) > 0) \tag{5.2}$$

容易证明，

$$\widetilde{P}(A) = P(A|B) : \mathscr{J} \to \mathbb{R}$$

也是概率函数.

5.1.3　独立事件

两个事件的独立

若 $P(AB) = P(A)P(B)$ 则称 A, B 相互独立. 此时有 $P(A) = P(A|B)$，并且能**推出** A **与** B^C **独立**.

多个事件的独立

若对 $A_1, A_2, \ldots$ 为可数个事件，若从中**任取有限个事件** $A_{i_1}, \ldots, A_{i_m}$ 都有

$$P(A_{i_1} \cdots A_{i_m}) = P(A_{i_1}) \cdots P(A_{i_m})$$

，则称 $A_1, A_2, \ldots$ 相互独立.

条件独立

若 $P(AB|E) = P(A|E)P(B|E)$，则称事件 A, B 关于事件 E 条件独立.
注意：条件独立不能推出独立，独立也不能推出条件独立

5.1.4　贝叶斯 (Bayes) 公式

定义 Ω 的一个分割 $\{B\}$ 满足 $\sum_i B_i = \Omega$ 且 $B_iB_j = \phi, \forall i \neq j$. 则

$$P(A) = P(A\sum_i B_i) = \sum_i P(B_i)P(A|B_i) \tag{5.3}$$

$$P(B_j|A) = \frac{P(AB_j)}{P(A)} = \frac{P(B_j)P(A|B_j)}{\sum_i P(B_i)P(A|B_i)} \tag{5.4}$$

(5.4) 即为 Bayes 公式.

5.2　随机变量

5.2.1　1 维随机变量

定义实值函数 $X(\omega) : \Omega \to \mathbb{R}$，该映射给样本空间中的每个试验结果一个对应的数值，随机变量 即定义为试验结果的一个实值函数，也可以通过**随机变量的函数**定义另一个随机变量.（默认用大写字母表示随机变量，小写字母表示实数.）

5.2.2　随机变量的概率

$\forall I \in \mathbb{R}$ 且 I 为可测集，利用数值所对应的事件定义概率函数 $P_X : \mathbb{R} \to \mathbb{R}$

$$P_X(X \in I) \triangleq P(X^{-1}(I))$$

5.2.3 （累积）分布函数 (cdf)

定义为 $F(x) \triangleq P(X \le x), x \in \mathbb{R}$，则 $P(a < X \le b) = F(b) - F(a)$. cdf 有以下性质：

1. $F(x)$ 单调（非严格）递增
2. $\lim\limits_{x\to-\infty} F(x) = 0, \lim\limits_{x\to+\infty} F(x) = 1$
3. $F(x)$ 右连续

以上性质也是 $F(x)$ 成为 cdf 的充要条件.

5.2.4 离散分布

概率质量函数 (pmf)

定义为 $f(x) = P(X = x), \forall x \in \mathbb{R}$.

期望和方差

期望 $E(X) \triangleq \sum_i x_i f(x_i)$；方差 $Var(X) \triangleq \sum_i (x_i - E(X))^2 f(x_i) = E(X - E(X)^2)$.

1. 若用 X 的函数 $g(X)$ 定义 Y，则 $E(g(X)) = \sum_i g(x_i) f(x_i) \neq g(E(X))$
2. $E(X+Y) = E(X) + E(Y)$；若 X, Y 独立，则 $Var(X+Y) = Var(X) + Var(Y)$

5.2.5 常见离散分布

Bernoulli 分布

- $X = \begin{cases} 1, & p \\ 0, & 1-p \end{cases}$，记为 $X \sim B(p)$
- $E(X) = p$，$Var(X) = p(1-p)$

二项分布

- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$，记为 $X \sim B(n, p)$
- X 为 n 次伯努利试验中试验成功的次数
- $E(X) = np$，$Var(X) = np(1-p)$

Poisson 分布

- $f(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$，记为 $X \sim P(\lambda)$
- $E(X) = \lambda$，$Var(X) = \lambda$
- 对于 $X \sim B(n, p)$，n 非常大而 np 较小时，X 近似服从 $P(np)$
- $P(\lambda)$ 多用于当 X 表示一定时间内出现的小概率时间的次数时

5.2.6 连续分布

概率密度函数 (pdf)

若存在 $f \geq 0$，使得 $\forall$ 可测集 $I \subset R$ 都有

$$P(X \in I) = \int_I f(x)\mathrm{d}x$$

，则 X 为连续型随机变量，f 为 X 的概率密度函数

1. $P(X = a) \equiv 0, \forall a \in \mathbb{R}$
2. $E(X) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x$
3. $Var(X) = \int_{-\infty}^{\infty}(x - E(X))^2 f(x)\mathrm{d}x = E(X^2) - E(X)^2$

5.2.7 常见连续分布

均匀分布

- $f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a,b) \\ 0, & otherwise \end{cases}$，记为 $X \sim U(a,b)$
- $E(X) = \frac{a+b}{2}$，$Var(X) = \frac{(b-a)^2}{12}$

正态分布

- $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \forall x \in \mathbb{R}$，记为 $X \sim N(\mu, \sigma^2)$
- $E(X) = \mu, Var(X) = \sigma^2$
- $X \sim N(\mu, \sigma^2) \implies Y = \frac{X-\mu}{\sigma} \sim N(0,1)$

指数分布

- $f = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$
- $E(X) = \frac{1}{\lambda}, Var(X) = \frac{1}{\lambda^2}$
- $F(X) = 1 - e^{-\lambda x}$
- X 一般刻画寿命或者等待时间
- 无记忆性：

$$P(X > x + t | X > x) = \frac{P(X > x + t)}{P(X > x)} = \frac{e^{-\lambda(x+t)}}{e^{-\lambda x}} = e^{-\lambda t} = f(t)$$

5.3 联合分布

5.3.1 随机向量

$(X_1, X_2, \ldots, X_n) : \Omega \to \mathbb{R}^n$，其中 X_i 为随机变量，定义（联合）累积分布函数

$$F(x_1, \ldots, x_n) \triangleq P(X_1 \le x_1, \ldots, X_n \le x_n), (x_1, \ldots, x_n) \in \mathbb{R}^n$$

5.3.2 离散分布

若 $\forall 1 \le i \le n, X_i$ 为离散型随机变量，则 $(X_1, \ldots, X_n)$ 为离散型随机向量. 定义 pmf 为 $f(x_1, \ldots, x_n) \triangleq P(X_1 = x_1, \ldots, X_n = x_n)$.

5.3.3 连续分布

若对一切可测集 $Q \subset \mathbb{R}^n$,$P((X_1, \ldots, X_n) \in Q) = \int_Q f(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n$,则 $(X_1, \ldots, X_n)$ 为连续型随机向量，f 为其 pdf.

以 $n = 2$ 为例，$F(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x, y)\mathrm{d}x\mathrm{d}y$，反过来有 $f(a, b) = \frac{\partial^2 F}{\partial x \partial y}$

二元正态分布

$X, Y \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho), |\rho| < 1$

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\frac{x-\mu_1}{\sigma_1}\frac{y-\mu_2}{\sigma_2}\right]\right\} \tag{5.5}$$

5.3.4 边际分布

对于的连续随机变量，定义 X_i 的边际累积分布函数 (cdf)

$$F_i(x) \triangleq P(X_i \le x) = P(X_i \le x, -\infty < X_j < \infty (j \ne i)) \tag{5.6}$$

以 $n = 2$ 为例，$F_X(x) = \lim\limits_{y \to \infty} P(X \le x, Y \le y) = \lim\limits_{y \to \infty} F(x, y)$.

定义 X_i 的边际密度函数 (pdf)

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y$$

5.3.5 条件分布

对于 $n = 2$ 的连续型随机变量，求其条件密度函数

$$P(X \le x | y \le Y \le y + \mathrm{d}y) = \frac{\int_{-\infty}^{x} \left(\int_{y}^{y+\mathrm{d}y} f(t, s)\mathrm{d}q\right)\mathrm{d}p}{\int_{y}^{y+\mathrm{d}y} f_Y(q)\mathrm{d}q}$$

$$P(X=x|y\le Y\le y+\mathrm{d}y)=\frac{\int_y^{y+\mathrm{d}y}f(t,s)\mathrm{d}q}{\int_y^{y+\mathrm{d}y}f_Y(q)\mathrm{d}q}\to\frac{f(x,y)}{f_Y(y)}(\mathrm{d}y\to 0)$$

$$f_X(x|y)\triangleq\frac{f(x,y)}{f_Y(y)} \tag{5.7}$$

定义 $f_X(x|y)$ 的累积分布函数 (cdf) 为 $F(a|y)=\int_{-\infty}^{a}f_X(x|y)\mathrm{d}x$

(1)（乘法法则）$f(x,y)=f_X(x|y)f_Y(y)=f_Y(y|x)f_X(x)$

(2)（全概率公式）$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_{-\infty}^{+\infty}f_X(x|y)f_Y(y)\mathrm{d}y$

对于二元正态分布，

$$f_Y(y|x)=\frac{1}{\sqrt{2\pi}\sigma_2}\frac{1}{2\sqrt{1-\rho^2}}\exp\left\{-\frac{[y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))]^2}{2(1-\rho^2)\sigma_2^2}\right\} \tag{5.8}$$

即 $\boldsymbol{Z=(Y|x)\sim N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),(1-\rho^2)\sigma_2^2)}$

5.3.6 独立性

若 $F(x_1,\dots,x_n)=F_1(x_1)\cdots F_n(x_n),\forall x_1,\dots,x_n\in\mathbb{R}$，则称 $X_1,\dots,X_n$ 相互独立. 该条件完全等价于 $f(x_1,\dots,x_n)=f_1(x_1)\cdots f_n(x_n)$.

1. 若 $X_1,\dots,X_n$ 相互独立，则 $Y_1=g_1(X_1,\dots,X_m)$ 与 $Y_2=g_2(X_{m+1},\dots,X_n)$ 相互独立
2. 若 pdf $f(x_1,\dots,x_n)=g_1(x_1)\cdots g_n(x_n),\forall x_1,\dots,x_n\in\mathbb{R}$，则 $X_1,X_2,\dots,X_n$ 相互独立，且 f_i 与 g_i 相差常数倍

5.3.7 多个随机变量的函数

★密度函数变换法

设存在由随机变量 $X1,X2$ 到 Y_1,Y_2 的可逆映射 $\begin{cases}Y_1=g_1(X_1,X_2)\\Y_2=g_2(X_1,X_2)\end{cases}$,逆为 $\begin{cases}X_1=h(Y_1,Y_2)\\X_2=h(Y_1,Y_2)\end{cases}$.
对于 X_1,X_2 上的区域 A，若其在 Y_1,Y_2 上的映射为 B，则显然有

$$P((X_1,X_2)\in A)=P((Y_1,Y_2)\in B)$$

$$P((Y_1,Y_2)\in B)=\int_B\mathscr{F}(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2$$

$$P((X_1,X_2)\in A)=\int_A f(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2=\int_B f(h(y_1),h(y_2))\left|\frac{D(x_1,x_2)}{D(y_1,y_2)}\right|\mathrm{d}y_1\mathrm{d}y_2$$

因此

$$\mathscr{F}(Y_1,Y_2)=f(h(Y_1),h(Y_2))\left|\frac{D(x_1,x_2)}{D(y_1,y_2)}\right| \tag{5.9}$$

5.4 随机变量的数学特征

5.4.1 期望

1. 刻画分布的集中趋势
2. $X_1, \ldots, X_n$ 独立时，$E(X_1X_2\cdots X_n) = E(X_1)\cdots E(X_n)$

5.4.2 分位数

$\forall \alpha \in (0,1)$，若 $P(X \le a) \ge \alpha$ 且 $P(X \ge a) \ge 1-\alpha$，则称 $X = a$ 为 X 的 **下α-分位数**.

1. 若 cdf 连续，则 $F(a) = \alpha$
2. 分位数不一定唯一

5.4.3 方差

1. 刻画分布的集中程度
2. $Var(cX) = c^2Var(X)$
3. $Var(X+Y) = Var(X) + Var(Y) + 2E[(X-\mu_1)(Y-\mu_2)]$

5.4.4 协方差与相关系数

定义协方差 $Cov(X,Y) \triangleq E[(X-\mu_1)(Y-\mu_2)]$

1. $Cov(X,X) = Var(X)$
2. $Cov(X,Y) = E(XY) - E(X)E(Y)$
3. $Cov(aX_1 + bX_2 + c, Y) = aCov(X_1,Y) + bCov(X_2,Y)$

定义相关系数 $Corr(X,Y) \triangleq \frac{Cov(X,Y)}{\sigma_1\sigma_2} = E(\frac{X-\mu_1}{\sigma_1}\frac{Y-\mu_2}{\sigma_2}) = \rho$

1. X, Y 独立 $\implies Cov(X,Y) = 0$，反之不一定
2. $|Corr(X,Y)| \le 1$
3. ρ 实际上是**线性**相关系数，不能表达高维的相关关系

5.4.5 矩 (Moment)

定义 将 $E[(X-C)^n]$ 称为 X 关于 C 的 n 阶矩. 当 $C = E(X)$ 时，称为中心矩；当 $C = 0$ 时，称为原点矩；标准化后的矩 $E[\big(\frac{x-\mu}{\sigma}\big)^n]$ 称为 n 阶标注矩.

偏度系数

定义 3 阶标准矩又称偏度系数.

1. 偏度系数小于零，左偏；大于零，右偏
2. 相对于 5 阶以上的奇数阶矩，3 阶矩容易计算且噪声的影响小

峰度系数

定义　4 阶标准矩又称峰度系数.

1. 正态分布峰度恒为 3
2. 超值峰度定义为 $E[\left(\frac{X-\mu}{\sigma}\right)^4]$
3. 超值峰度为正，一般相对于正态分布峰更尖，尾部更扁

5.4.6 矩母函数

定义　(moment generating function, mgf) $M_X(t) \triangleq E(e^{tX})$ 若在 $t=0$ 的某个邻域内 $M_X(t)$ 存在，则称其为 X 的矩母函数，否则称 X 的矩母函数不存在. **注意标明 t 的取值范围.**
对于 $X \sim Exp(\lambda)$，

$$M_X(t) = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}\mathrm{d}x = \frac{\lambda}{\lambda - t}, t < \lambda \tag{5.10}$$

对于 $X \sim N(\mu, \sigma^2)$，

$$M_X(t) = e^{\frac{\sigma^2 t^2}{2} + \mu t} \tag{5.11}$$

对于 $Y = aX + b$，

$$M_Y(t) = E(e^{aX+b}) = e^{tb}E(e^{taX}) = e^{tb}M_X(ta) \tag{5.12}$$

定理　矩母函数确定矩：

$$E(X^n) = M_X^{(n)}(0) \tag{5.13}$$

证明.

$$M_X(t) = \sum_{n\geq 0} M_X^{(n)}(0)\frac{t^n}{n!}$$

$$M_X(t) = E(e^{tx}) = E\left(\sum_{n\geq 0}\frac{(tx)^n}{n!}\right) = \sum_{n\geq 0} E(t^n)\frac{x^n}{n!}$$

由于 $M_X(t)$ 泰勒展开唯一，故 $M_X^{(n)} = E(t^n)$ □

例　对于 $X \sim N(0,1)$，

$$M_X(t) = e^{\frac{t^2}{2}} = \sum_{n\geq 0}\frac{t^{2n}}{2^n n!} = \sum_{n\geq 0}\frac{(2n)!}{2^n n!}\frac{t^{2n}}{(2n)!}$$

故 $E(X^{2n}) = \frac{(2n!)}{2^n n!}, E(X^{2n-1}) \equiv 0$
定理　矩母函数确定分布. 若存在 $a > 0$，使得 $M_X(t) = M_Y(t), t \in (-a, a)$，则 X, Y 同分

布.

REMARK　矩存在的时候，矩母函数不一定存在（即泰勒展式不收敛），因此各阶矩完全相同也无法说明两个随机变量同分布.

例　取服从对数正态分布的变量 Y 于另一变量 Z

$$y = f_1(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{\log^2 x}{2}}, x > 0$$
$$z = f_2(x) = f_1(x)[1 + \sin(2\pi \log x)], x > 0$$

则它们的 n 阶矩存在以下关系：

$$E(Z^n) - E(Y^n) = \int_0^\infty x^n f_1(x) \sin(2\pi \log x) \mathrm{d}x = T$$

令 $t = \log x - n$，则 $x = e^{t+n}$

$$\begin{aligned} T &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{n(t+n)} e^{-\frac{1}{2}(t+n)} e^{-\frac{1}{2}(t+n)^2} \sin(2\pi(t+n)) e^{t+n} \mathrm{d}t \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{1}{2}(n+t)(n+t+1)\right] \mathrm{d}t \\ &= \frac{e^{-\frac{1}{8}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{1}{2}(n+t+\frac{1}{2})^2\right] \mathrm{d}t = 0 \end{aligned}$$

因此对一切 n，Y, Z 的 n 阶矩相等. 但是由于

$$e^{tx} f_1(x) = \frac{1}{\sqrt{2\pi}} e^{tx - \frac{\log^2 x}{2}}$$

对于一切的 $t > 0$ 都有 $\lim\limits_{x\to\infty} e^{tx} f_1(x) = +\infty$，因此无法积分，$M_Y(t)$ 不存在. 故即使各阶矩完全相同，实际上也不是同分布.

独立随机变量和的分布

对于相互独立的 $X_1, X_2, \ldots, X_n$，有

$$M_{X_1+\cdots+X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t) \tag{5.14}$$

对于联合分布 $(X_1, X_2, \ldots, X_n)$，定义矩母函数为多元函数

$$M_{X_1,\ldots,X_n}(t_1, t_2, \ldots, t_n) = E(e^{t_1 X_1 + \cdots + t_n X_n}) \tag{5.15}$$

若该函数在原点的邻域 $B(0, \sigma), \sigma > 0$ 内有定义，则称联合分布的矩母函数存在；若两组联合分布的矩母函数在 $B(0, \sigma)$ 内恒等，则这两组联合分布同分布.

5.4.7　条件期望

定义

$$E(Y|X=x)=\begin{cases}\sum_i y_iP(Y=y_i|X=x)\\ \int_{-\infty}^{\infty}f_Y(y|X=x)\mathrm{d}y\end{cases}$$

定理　$E(Y)=E(E(Y|X))$

定理　$E((Y-g(X))^2)\geq E[(Y-E(Y|X))^2]$，这意味着 $E(Y|X)$ 是均方误差意义下的最优预测. 均方最优：$E[(Y-c)^2]\geq E[(Y-E(Y))^2]$，这是因为 $\frac{\partial E[(Y-c)^2]}{\partial c}=2E[Y-c]\big|_{c=E(Y)}=0.$

5.5　不等式与极限定理

5.5.1　概率不等式

Markov 不等式

若 $Y\geq 0$，则 $\forall a>0$，有

$$P(Y\geq a)\leq\frac{E(Y)}{a}\tag{5.16}$$

证明. 令 $I=\begin{cases}1, Y\geq a\\ 0, Y<a\end{cases}$ ，则不论 Y 的取值，均有 $I\leq\frac{Y}{a}$
故 $P(Y\geq a)=E(I)\leq E(\frac{Y}{a})=\frac{E(Y)}{a}$ □

Chebyshef 不等式

若 $Var(Y)$ 存在，则 $\forall a>0$

$$E(|Y-E(Y)|\geq a)\leq\frac{Var(Y)}{a^2}\tag{5.17}$$

证明. 在(5.16)中代入 $P[|Y-E(Y)|\geq a]=P[(Y-E(Y))^2\geq a^2]=E[Var(Y)\geq a^2]$ 即可. □

Chernoff 不等式

$\forall a>0, t>0$ 有

$$P(Y\geq a)\leq\frac{E(e^{tY})}{e^{ta}}\tag{5.18}$$

5.5.2　大数定律

设 $X_1,X_2,\ldots,X_n$ 独立同分布，$\bar{X}=\frac{1}{n}\sum_{k=1}^n X_k$ ，则

Khinchin 弱大数定律

$\forall\varepsilon>0$ 有

$$\lim_{n\to\infty} P(|\bar{X}-\mu|<\varepsilon)=1 \tag{5.19}$$

证明.

$$P(|\bar{X}-\mu|\geq\varepsilon)\leq\frac{Var(\bar{X})}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}\to 0(n\to\infty)$$

□

若 $P(|\bar{X}-\mu|\geq\varepsilon)\leq\alpha$，则称 α 为置信水平，ε 为精度.

Kolmogorov 强大数定律

$\bar{X}$ 几乎必然收敛到 μ

$$\forall\varepsilon>0,\ P(\lim_{n\to\infty}|\bar{X}-\mu|<\varepsilon)=1 \tag{5.20}$$

5.5.3　中心极限定理 (CLT)

若 $E(X_i),Var(X_i)$ 均存在，则

$$\lim_{n\to\infty} P(\frac{X_1+\cdots+X_n-n\mu}{\sqrt{n}\sigma}\leq x)=\Phi(X)\sim N(0,1),\forall x\in\mathbb{R} \tag{5.21}$$

第六章　随机过程

6.1　Bernoulli 过程与 Poisson 过程

6.1.1　基本概念

定义　令 $(\omega, \mathscr{J}, P)$ 为概率空间,T 为指标集,S 为状态空间（相空间）,$\forall t \in T, \omega \in \Omega, X_t(\omega) \in S$，$X_t$ 为随机变量，$\{X_t : t \in T\}$ 称为一个<u>**随机过程**</u>**(stochastic process)**；$\forall \omega_0 \in \Omega$，$X_t(\omega_0)$ 称为一个<u>**样本轨道**</u>**(sample path)**.

定义　随机变量 $\{N(t) : t \geq 0\}$ 称为<u>**计数过程**</u>，如果满足

1. $N(t) \geq 0$，取整数值
2. $\forall t > s \geq 0, N(t) \geq N(s)$
3. $N(t) - N(s)$ 表示时间 $(s, t]$ 时间内的事件数

一般记第 n 次到达的时间为 Y_n，则有到达时间序列 $\{Y_n\}_{n=1}^{\infty}$；定义 $\{T_n\}_{n=1}^{\infty}$，满足 $T_1 = Y_1$, $T_i = Y_i - Y_{i-1} (i > 1)$.

REMARK:

1. X_t 记为 $X(t)$ 时强调函数性质
2. T 通常解释为时间，可以为离散或者连续
3. X_t 称为过程在 t 时刻的状态，S 可以为离散或者连续
4. $X_t(\omega) = X(t, \omega)$ 视为 (t, ω) 的函数

6.1.2　Bernoulli 过程

定义　T 为离散时间，记为 $\{1, 2, \ldots, n, \ldots\}$. $X_1, X_2, \ldots$ 独立同分布且 $X_i \sim B(p)$，X_i 可以视为第 i 次实验成功与否. $\{X_n\}_{n=1}^{\infty}$ 记为 Bernoulli 过程.

REMARK:

1. $X_1, X_2, \ldots$ 相互独立 $\iff \forall n, X_1, X_2, \ldots, X_n$ 相互独立
2. $\forall$certain n，$\{X_{n+1}, X_{n+2}, \ldots\}$ 仍为 Bernoulli 过程.

6.1.3 Bernoulli 过程的性质

首次到达（相邻两次到达）的分布：几何分布

定义 令 $T=$ 首次试验成功的时间，则 $P(T=m)=p(1-p)^{m-1}(m=1,2,\dots)$

性质 无记忆性

$$P(T-n=m|T>m)=\frac{P(T=n+m,T>n)}{P(T>n)}=\frac{p(1-p)^{n+m}}{1-\sum_{k\geq 1}p(1-p)^{k-1}}=p(1-p)^{m-1}=P(T=m) \tag{6.1}$$

第 k 次到达时间的分布：负二项分布

定义 对于 Bernoulli 过程 $\{T_n\}_{n=1}^{\infty}$，令

$$P(Y_k=m)=P(\text{第 } m \text{ 次成功，且前 } m \text{ 次成功了 } k \text{ 次})=\binom{m-1}{k-1}p^k(1-p)^{m-k}$$

Bernoulli 过程的分裂

若每次到达的时间以 q 的概率保留下来，则保留下来的过程为 Bernoulli 过程，参数为 pq.

Bernoulli 过程的合并

两个独立的 Bernoulli 过程合并，结果仍为 Bernoulli 过程，且参数为 $p+q-pq$.

6.1.4 Poisson 过程

平稳增量过程

定义 如果增量 $X(t+\tau)-X(t)$ 的**分布**仅与 τ 有关，而与 t 无关，则称为平稳增量过程.

独立增量过程

定义 时刻 t 以前发生的事件数 [即 $N(t)$] 必须独立于时刻 t 与 $t+s$ 之间发生的事件数 $[N(t+s)-N(t)]$，则该过程为独立增量过程.

Poisson 过程

定义 1 计数过程 $\{N(t):t\geq 0\}$ 称为 Poisson 过程，如果满足

1. $N(0)=0$
2. 过程有平稳独立增量
3. 存在 $\lambda>0$，满足 $h\to 0$ 时，$P(N(h)=1)=\lambda h+o(h)$
4. $P(N(h)\geq 2)=o(h)$

定义 2

1. $N(0) = 0$
2. 过程有独立增量
3. $\forall s, t \geq 0, P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$

6.1.5 Poisson 过程的性质

Poisson 过程的分裂

定理 假设 Poisson 过程过程中每次发生的事件分为 I 型和 II 型，以概率 p 为 I 型，否则为 II 型. $N_1(t), N_2(t)$ 分别表示 $(0,t]$ 内两种事件的数目，则

1. $N(t) = N_1(t) + N_2(t)$
2. $\{N_1(t) : t \geq 0\}, \{N_2(t) : t \geq 0\}$ 为 Poisson 过程，且到达率分别为 $\lambda p, \lambda(1-p)$
3. 这两个过程相互独立

Poisson 过程的合并

定理 已知 $\{N_1(t) : t \geq 0\}, \{N_2(t) : t \geq 0\}$ 为相互独立的 Poisson 过程，到达率分别为 λ_1, λ_2，令 $N(t) = N_1(t) + N_2(t)$，则 $\{N(t) : t \geq 0\}$ 为 Poisson 过程，其到达率为 $\lambda_1 + \lambda_2$. 且时间 N_1 先发生的概率为 $\frac{\lambda_1}{\lambda_1+\lambda_2}$，事实上，对于每个发生的事件，其属于 N_i 的概率为 $\frac{\lambda_i}{\lambda_1+\lambda_2}$.

条件作用

已知 $[0,t]$ 内发生了一次事件，则该事件发生在 $[0,s](s \leq t)$ 上的概率为

$$\begin{aligned} P(Y_1 \leq s | N(t) = 1) &= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} = \frac{P(N(s) = 1)P(N(t) - N(s) = 0)}{P(N(t) = 1)} \\ &= \frac{\lambda s e^{-\lambda s} \cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t} \end{aligned}$$

$\implies N(t) = 1$ 条件下，$Y_1 \sim U(0,t)$.

定理 $N(t_2) = n$ 的条件下，对于 $t_1 < t_2$，$N(t_1) \sim B(n, \frac{t_1}{t_2})$. 相当于在 $[0, t_2]$ 上均匀分布随机放置 n 个到达点，第 j 次到达即为第 j 阶次序统计量.

定理 $N(t) = n$ 的条件下，事件发生的 n 个时刻 $Y_1, Y_2, \ldots, Y_n$ 的联合 pdf 为

$$f(t_1, \ldots, t_n) = \frac{n!}{t^n} \tag{6.2}$$

证明. 已知条件等价于

$$T_1 = 1, T_2 = t_2 - t_1, \ldots, T_n = t_n - t_{n-1}, T_{n+1} > t - t_n$$

这里 T_i 为间隔事件，$T_i \overset{iid}{\sim} Exp(\lambda)$. $Y_1,\ldots,Y_n$ 的 pdf 就等价于 $T_1,\ldots,T_n$ 的 pdf

$$\begin{aligned} f(t_1,\ldots,t_n|N(t)=n) &= \frac{f(t_1,\ldots,t_n,n)}{P(N(t)=n)} \\ &= \frac{\lambda e^{-\lambda t_1}\cdots\lambda e^{-\lambda(t_n-t_{n-1})}\cdot e^{-\lambda(t-t_n)}}{\frac{(\lambda t)^n}{n!}e^{-\lambda t}} \\ &= \frac{n!}{t^n} \end{aligned}$$

□

因此有生成 Poisson 过程的另一种方式：

1. 根据 $N(t)\sim P(\lambda t)$ 生成 $(0,t]$ 内事件发生数
2. 假定 $N(t)=n$，则取 $U_1,\ldots,U_n \overset{iid}{\sim} U(0,t)$
3. 令 $Y_j = U_{(j)}, j=1,2,\ldots,n$

次序统计量

设有 $X_1,\ldots,X_n$ 独立同分布，令 $X_{(i)}=\min\{X_1,\ldots,X_n\}$ $(1\le i\le n)$，则 $X_{(i)}$ 为 i 阶次序统计量.

设 X_i 的 pdf 为 $f(x)$，cdf 为 $F(x)$，则

$X_{(j)}=x \iff X_1,\ldots,X_n$ 中有 $j-1$ 个取值 $<x$，且有 $n-j$ 个取值 $>x$，可以不严格地推导出 $X_{(j)}$ 的 pdf.

$$P(x\le X_{(j)}\le x+\mathrm{d}x)\approx \binom{n}{j-1,n-j,1}F^{j-1}(x)(1-F(x))^{n-j}f(x)\mathrm{d}x$$

$$f_j(x)=\frac{n!}{(j-1)!(n-j)!}F^{j-1}(x)(1-F(x))^{n-j}f(x) \tag{6.3}$$

由此定义联合分布 pdf ...

6.1.6 Poisson 过程的推广

非齐次 Poisson 过程

定义　计数过程 $\{N(t):t\ge 0\}$ 称为到达率为 $\lambda(t)(\lambda(t)>0)$ 的 Poisson 过程，如果

1. $N(0)=0$
2. 具有独立增量
3. $P(N(t+h)-N(t)=1)=\lambda(t)h+o(h)$
4. $P(N(t+h)-N(t)\ge 2)=o(h)$

若令 $m(t)=\int_0^t \lambda(k)\mathrm{d}k$，则

$$P(N(t+s)-N(t)=n)=e^{m(t+s)-m(t)}\frac{(m(t+s)-m(t))^n}{n!}$$

其中 $m(t+s)-m(t)$ 即为这段时间内的期望事件数. 该过程没有平稳增量性质.

复合 Poisson 过程

定义 $\{N(t):t\geq 0\}$ 为 Poisson 过程，X_i iid 且与 $N(t)$ 相互独立，令 $Z(t)\triangleq\sum_{i=1}^{N(t)}X_i$，则 Z_t 为复合 Poisson 过程. **定理**

1. $Z(t)$ 有独立增量性质
2. 若 $E(X_i^2)\leq\infty$，则 $E(Z(t))=\lambda tE(X_i), Var(Z(t))=\lambda tE(X_i^2)$（用矩母函数证明）

条件 Poisson 过程

定义 设 $\Lambda>0$ 为随机变量，当 $\Lambda=\lambda$ 时，计数过程 $\{N(t):t\geq 0\}$ 为到达率为 λ 的 Poisson 过程，则称 $N(t)$ 为条件 Poisson 过程. $N(t)$ 不是一个 Poisson 过程.

定理 $E(\Lambda)<\infty$，则 $E(N(t))=tE(\Lambda), Var(N(t))=t^2Var(\Lambda)+tE(\Lambda)$

若事件间隔 X_i 的分布不限定为指数分布，则计数过程称为**更新过程**

表 6.1: **Summary**

到达过程	Bernoulli 过程	Poisson 过程
到达时间	离散	连续
到达率	p 每次试验	λ 单位时间
相邻两次到达间隔	几何分布	指数分布
t 内到达次数的分布	二项分布	Poisson 分布
第 k 次到达	负二项分布	Gamma 分布

6.2 离散时间 Markov 链

6.2.1 基本概念

指标集 T 离散，不妨记 $T=\{0,1,2,\ldots\}$.

状态空间 S 离散，不妨记 $T=\{0,1,2,\ldots\}$

定义 (Markov 链) $\{X_n, n=0,1,\ldots\}$ 为随机过程，$X_i\in S$，若 $\forall n\geq 0$ 及任意状态 $i,j,i_0,\ldots,i_{n-1}$，有

$$P(X_{n+1}=j|X_1=1,\ldots,X_{n-1}=i_{n-1},X_n=i)=P(X_{n+1}=j|X_n=i) \tag{6.4}$$

则称 $\{X_n, n=0,1,\dots\}$ 为离散时间 Markov 链. 可以导出：

$$\begin{aligned}&P(X_0=i_0,\dots,X_n=i_n)\\=&P(X_n=i_n|X_0=i_0,\dots,X_{n-1}=i_{n-1})P(X_0=i_0,\dots,X_{n-1}=i_{n-1})\\=&P(X_n=i_n|X_{n-1}=i_{n-1})P(X_0=i_0,\dots,X_{n-1}=i_{n-1})\\=&\dots\\=&P(X_n=i_n|X_{n-1}=i_{n-1})\dots P(X_1=i_1|X_0=i_0)P(X_0=i_0)\end{aligned}$$

定义 $P(X_{n+1}=j|X_n=i)$ 称为 Markov 链的（一步）**转移概率**. 当它与 n 无关时，称 Markov 链关于时间是**齐次** 的，记

$$P_{ij}^{(n)}=P_{ij}=P(X_{n+1}=j|X_n=i) \tag{6.5}$$

将 $P\triangleq(P_{ij})$ 称为**转移概率矩阵**.

1. 状态有限（无限）时，对应称为有（无）限链
2. 路径概率 $P(X_k=i_k)=P_{i_0i_1}\cdots P_{i_{n-1}i_n}$
3. 转移概率矩阵描述了一个 n 个点带自环的完全图

例 (随机游走) $S=\{0,\pm1,\pm2,\dots\}, p\in(0,1)$. $P(X_i=1)=p, P(X_i=0)=1-p$，令 $Y_n=\sum_0^n X_i$，则 $\{Y_n\}$ 称为随机游走模型. $P_{i,i+1}=p, P_{i,i-1}=1-p$.

例 (赌博模型) $S=\{0,1,\dots,n\}$，0 和 n 为**吸收态**，$P_{0,0=1}, P_{n,n}=1$. 此模型称为**具有吸收壁的有限随机游走**.

6.2.2 C-K 方程

定义 n 步转移概率为 $P_{ij}^{(n)}\triangleq P(X_n=j|X_0=i)$，由于时间齐次性，$P(X_{m+n}=j|X_m=i)=P_{ij}^{(n)}$，即与 m 无关. 规定 $P_{ij}^{(n)}=\begin{cases}1, i=j\\0, i\neq j\end{cases}=\delta_{ij}$.

定理 (Chapman-Kolmogorov 方程) $\forall m,n\geq0, \forall i,j\in S$，有

$$P_{ij}^{(m+n)}=\sum_{k\in S}P_{ik}^{(m)}P_{kj}^{(n)} \tag{6.6}$$

证明. $P(X_{m+n}=j|X_0=i_0,\dots,X_m=k)=P(X_{m+n}=j|X_m=k)$ □

定理 若记 $P^{(n)}\triangleq(P_{ij}^{(n)})$，则 $P^{(n)}=P^n, P^{(n+m)}=P^{(n)}P^{(m)}$.

计算 X_n 的边际分布

$$\begin{aligned}P(X_n=j)&=\sum_{i\in S}P(X_0=i)P(X_n=j|X_0=i)\\&=\sum_{i\in S}P(X_0=i)P_{ij}^n\end{aligned} \tag{6.7}$$

矩阵形式：记 $\vec{\beta}_n=(\beta_{n_1},\beta_{n_2}\dots), \beta_{n_i}\triangleq P(X_n=i)$，则 $\vec{\beta}_n=\vec{\beta}_0P^n$.

6.2.3　状态的分类

定义（可达）　　称状态 i **可达**状态 j，若 $\exists n \geq 0, P_{ij}^{(n)} > 0$. 若相互可达，则记为 $i \leftrightarrow j$.

1. 自返性：$i \leftrightarrow i$
2. 对称性：$i \leftrightarrow j \iff j \leftrightarrow i$
3. 传递性

显然，$\leftrightarrow$ 定义了一个 S 上的**等价关系**，将所有状态划分为若干个**等价类**. 若一条 Markov 链上仅有一个类，则称其为**不可约的**.

定义　　记 $f_{ij}^{(n)}$ 表示从 i 经过 n 次转移后，首次到达 j 的概率. 则**首达**概率

$$f_{ij}^{(n)} = P(X_n = j, X_k \neq j, k = 1, 2, \dots, n-1 | X_0 = i)$$

额外定义平凡情况 $f_{ij}^{(0)} = \delta_{ij}$. 注意区别 $P_{ij}^{(n)}$ 与 $f_{ij}^{(n)}$.

定义　　$f_{ij} \triangleq \sum_{n=1}^{\infty} f_{ij}^{(n)}$ 表示从 i 出发经有限步可达 j 的概率.

定义　　若 $f_{ii} = 1$，则称状态 i 为**常返 (recurrent)**，否则 i 为**瞬时 (transient)** 的.

★定理（常返态的等价判定）

- i 为常返态 $\iff \sum_{n=0}^{\infty} P_{ii}^{(n)} = +\infty$
- i 为瞬时态 $\iff \sum_{n=0}^{\infty} P_{ii}^{(n)} = \frac{1}{1-f_{ii}}$

证明. 引理：$P_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)}$

$$\begin{aligned}
\sum_{n=0}^{\infty} P_{ii}^{(n)} &= P_{ii}^{(0)} + \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_{ii}^{(k)} P_{ii}^{(n-k)} \\
&= P_{ii}^{(0)} + \left(\sum_{n=1}^{\infty} f_{ii}^{(k)} \right) \left(\sum_{k=1}^{\infty} P_{ii}^{(k)} \right) \\
&= 1 + f_{ii} \sum_{k=0}^{\infty} P_{ii}^{(k)} \\
&= \frac{1}{1 - f_{ii}} (若\ f_{ii} < 1)
\end{aligned}$$

因此 $\sum_{k=0}^{\infty} P_{ii}^{(k)}$收敛 $\iff f_{ii} < 1$. □

性质（常返态的命运）

1. 令 $I_n = [X_n = i]$，则 $\sum_{n=0}^{\infty}$ 为经过状态 i 的次数，

$$E(\sum_{n=0}^{\infty} I_n | X_0 = i) = \sum_{n=0}^{\infty} P(I_n = 1 | X_0 = i) = \sum_{n=0}^{\infty} P_{ii}^{(n)}$$

为从 i 出发的链回到 i 的期望次数

2. 若 i 常返，则从 i 出发时以概率 1 有限步回到 i
3. 若 i 非常返，则从 i 出发以概率 $P = 1 - f_i i > 0$ 回不到 i，从而从 i 出发的链恰好经过 i 的次数为 k 的概率为 $f_{ii}^{k-1}(1-f_{ii})$（几何分布），所以只能返回有限次，最后永远离开

★**定理（常返态的事实）**

1. 若 $i \leftrightarrow j$，则 i 和 j 同时为常返态或者非常返态
2. 常返态 i 所能到达的一切状态均与 i 相互可达，即**从常返态出发不能到达非常返态**
3. 在一个**有限** Markov 链上，从任意非常返态出发，最终必然到达常返态
4. 一个**有限** Markov 链至少有一个常返态
5. 一个**有限** Markov 链若**不可约**，则所有状态均为常返态

例子　（随机游走）当 $p = \frac{1}{2}$ 时常返，当 $p \neq \frac{1}{2}$ 时非常返.

定理　若 $i \leftrightarrow j$，且 i 为常返态，则 $f_{ji} = 1$

定义　若集合 $\{n | n \geq 1, P_{ii}^{(n)} > 0\}$ 非空，则其最大公约数 $d = d(i)$ 称为状态 i 的**周期**. 若 $d > 1$，则称状态 i 是**周期的**；若 $d = 1$ 则称 i 是**非周期的**. 若链中所有状态的周期都为 d，则称 d 为该**链的周期**. 若链中所有状态周期都为 1，则称该链是**非周期的**；否则称其为**周期的**.

定理　若 $i \leftrightarrow j$，则 i 的周期与 j 相等.

定理　若一个**不可约** Markov 链周期为 d，其状态空间 S 存在唯一的**划分** $\{S_1, S_2, \ldots, S_d\}$，且使得从 S_r 中任意状态出发，任 1 步转移必然进入 S_{r+1} 中. 实际上，若将状态 i 固定在 S_d 中，则有

$$S_r = \{j | \exists n \in \mathbb{N}, s.t. P_{ij}^{(nd+r)} > 0\}$$

对于任意 Markov 链，其状态空间存在划分 $\{C_0, C_1, C_2, \ldots\}$，其中 C_0 为所有非常返态构成的集合，$C_n (n \geq 1)$ 不可约.

6.2.4 稳态性质

定义　$\vec{\beta} = (\beta_j)_{j \in S}$ 为概率分布，若

$$\vec{\beta} P = \vec{\beta} \quad (i.e. \beta_j = \sum_{i \in S} \beta_i P_{ij})$$

则称 $\vec{\beta}$ 为该 Markov 链的**平稳分布 (stationary distribution)**.

若 $\vec{\beta}$ 为 X_0 的分布 (即 $P(X_0 = j) = \beta_j$)，则 $X_n (n \geq 1)$ 的分布都为 $\vec{\beta}$，从而 $\{X_n : n \geq 0\}$ 为**平稳过程**.

平稳分布为边际分布，不是条件分布，一般地有

$$P(X_{n+1} = j | X_n = i) = P_{ij} \neq P(X_{n+1} = j)$$

定义　设 i 为常返态，定义 $\mu_i \triangleq \sum_{n=1}^{\infty} n f_{ii}^{(n)}$ 为由 i 出发再返回 i 所需的平均时间（步数）.

若 $\mu_i < +\infty$，则称 i 为**正常返的 (positive recurrent)**；

若 $\mu_i = +\infty$，则称 i 为**零常返的 (null recurrent)**.

μ_i 越小，返回越频繁.

★**定理** 若 i 常返且周期为 d，则

$$\lim_{n\to\infty} P_{ii}^{(nd)} = \frac{d}{\mu_i} \tag{6.8}$$

当 $\mu_i = +\infty$ 时，$\frac{d}{\mu_i} = 0$.（不证）

定理 i 为零常返态或非常返态 $\iff \lim\limits_{n\to\infty} P_{ii}^{(n)} = 0$

证明. i 为零常返 $\implies \mu_i = +\infty \implies \lim\limits_{n\to\infty} P_{ii}^{(nd)} = 0$

又知 $P_{ii}^{(m)} = 0, m$ 不被 d 整除，故 $\lim\limits_{n\to\infty} P_{ii}^{(n)} = 0$

i 非常返 $\implies \sum_{n=0}^{\infty} P_{ii}^{(n)} < +\infty \implies \lim\limits_{n\to\infty} P_{ii}^{(n)} = 0$ □

定理 $i \leftrightarrow j$，常返，则 i 与 j 同时为正常返或零常返.

定理 若 j 为非常返或零常返，则 $\forall i \in S$ 都有 $\lim\limits_{n\to\infty} P_{ij}^{(n)} = 0$. 有限链不可能又零常返态，从而不可约有限 Markov 链所有状态都是正常返的. 若 Markov 链有零常返态

定义 若 i 正常返且周期为 1，则称 i 为**遍历的 (ergodic)**；若一个 Markov 链中所有状态都是遍历的，则称该链是遍历的.

★**定理** 对于不可约、非周期的 Markov 链，

(1) 若它是遍历的，则 $\pi_j = \lim\limits_{n\to\infty} P_{ij}^{(n)}$ 是该链的唯一平稳分布

(2) 若状态都是非常返的或是零常返的，则平稳分布不存在

6.3 某次课

1. 若它是遍历的，则 $\pi_j = \lim\limits_{n\to\infty}$

$\vec{\pi}P = \vec{\pi}$

证明.

$$\forall M, \sum_{j=0}^{M} P_{ij}^{(n)} \le \sum_{j=0}^{\infty} P_{ij}^{(n)} = 1$$

$$P_{ij}^{(n+1)} = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj} \ge \sum_{k=0}^{M} P_{ik}^{(n)} P_{kj}$$

$$P_{ij}^{(n)} \ge \sum_{k=0}$$

□

定义　若 $\pi_j = \lim\limits_{n\to\infty} P_{ij}^{(n)}$ 存在，则 $\vec{\pi} = (\pi_0, \pi_1, \dots)$ 称为该链的**极限分布**.
性质

1. $\lim\limits_{n\to\infty} P^{(n)} = \lim\limits_{n\to\infty} P^n = \begin{bmatrix} \vec{\pi} \\ \vec{\pi} \\ \vdots \end{bmatrix}$
2. 对于一切初始分布 $\vec{\beta}$，$\lim\limits_{n\to\infty} \vec{\beta} P^n = \vec{\pi}$
3. 一个链具有平稳分布并不意味着有极限分布
4. 若链不可约且遍历，则极限分布式链唯一的平稳分布
5. 可以证明，有限链总存在平稳分布；若其不可约，则平稳分布唯一

6.3.1　可逆性

（平稳分布当状态空间规模非常大时难以计算）
定义　$\pi_j \geq 0, \sum_i \pi_i = 1$，若满足以下可逆性条件：

$$\pi_i P_{ij} = \pi_j P_{ji}, \forall i, j \in S \tag{6.9}$$

则称该链对于 $\vec{\pi}$ 可逆.
定理　若 P 相对于 $\vec{\pi}$ 可逆，则 $\vec{\pi}$ 为链的平稳分布.

证明. $\sum_i \pi_i P_{ij} = \sum_i \pi_j P_{ji} = \pi_j (\sum_i P_{ji}) = \pi_j \implies \vec{\pi} P = \vec{\pi}$ □

6.3.2　特征值与特征向量

练习 5.2.23

给定 m 阶方阵 A_1，n 阶方阵 A_2 和 $m \times n$ 矩阵 B，证明若 A_1 和 A_2 没有相同的特征值，则关于 $m \times n$ 矩阵 X 的矩阵方程 $A_1 X - X A_2 = B$ 有唯一解.

证明. 因为 X 与 B 规格相同，所以只要证明该变换是单射即可证明是双射.
即证：$A_1 X = X A_2$ 仅有零解

$$\begin{aligned} A_1^2 X &= (A_1 X) A_2 = X A_2^2 \\ A_1^3 X &= (A_1 X) A_2^2 = X A_2^3 \\ &\vdots \\ A_1^n X &= (A_1 X) A_2^{n-1} = X A_2^n \end{aligned} \tag{$\bigstar$}$$

考虑上述等式的线性组合，则

$$f(A_1) X = X f(A_2)$$

恒成立.

不妨令 $f = p_A$，则 $f(A_1) = O$. 令 λ_i 为 A_1 的特征值，则 $\det(\lambda_i I_n - A_2) \neq 0$，所以

$$\det(f(A_2)) = \prod_{i=1}^{n} (\lambda_i I_n - A_2) \neq 0 \tag{$\bigstar$}$$

故 $f(A_2)$ 可逆，因此

$$X = f(A_1) X \left[f(A_2)\right]^{-1} = O$$

□

6.3.3 实对称矩阵

对实对称矩阵，A 正定

$\Longleftrightarrow A$ 的特征值都是正数

$\Longleftrightarrow A = TT^{\mathrm{T}}$,$T$ 可逆

$\Longleftrightarrow A = LDL^{\mathrm{T}}$

$\Longleftrightarrow A$ 的顺序主子式都是正数