目录

1 多维空间与多元函数

对 $u=(u^1,\cdots,u^m)\in\mathbb{R}^m$, 定义 u 的范数 (或模) 为

$$\|u\|=\sqrt{(u^1)^2+\cdots+(u^m)^2}.$$

我们把

$$U(x_0,r)=\{x\in\mathbb{R}^m \mid \|x-x_0\|<r\}$$

称为点 x_0 的 r 邻域.

定理: 设 $\{x_n\}\subseteq\mathbb{R}^m$, 其中 $x_n=(x_n^1,\cdots,x_n^m)$, $a=(a^1,\cdots,a^m)\in\mathbb{R}^m$, 则

$$\lim_{n\to\infty}x_n=a \Longleftrightarrow \lim_{n\to\infty}x_n^i=a^i,\quad i=1,\cdots,m.$$

1.1 多元函数的极限与连续

定义 (序列式定义) : 设 f 是定义在 $D\subseteq\mathbb{R}^m$ 上的 m 元函数, a 为 D 的聚点. 若对 $U^\circ(a)\cap D$ 中任意收敛于 a 的序列 $\{x_n\}$, 都有

$$\lim_{n\to\infty}f(x_n)=A,$$

则称 f 沿集合 D 有极限 A, 记为 $\lim\limits_{x\to a,x\in D}f(x)=A$. 当不至于混淆时, 简记为 $\lim\limits_{x\to a}f(x)=A$.

定义 (ε-δ 式定义) : 设 f 为定义在 $D\subseteq\mathbb{R}^m$ 上的 m 元函数, a 为 D 的聚点. 若对任意的 $\varepsilon>0$, 存在 $\delta>0$, 当 $x\in U^\circ(a,\delta)\cap D$ 时, 有

$$|f(x)-A|<\varepsilon,$$

则称 f 沿集合 D 有极限 A.

定理 ($\mathbb{R}^m$ 的收敛原理) : 设 f 为定义在 $D\subseteq\mathbb{R}^m$ 上的 m 元函数, 则 $\lim\limits_{x\to a}f(x)=A$ 的充要条件为对任意的 $\varepsilon>0$, 存在 $\delta>0$, 当 $x',x''\in U^\circ(a,\delta)\cap D$ 时, 有

$$|f(x')-f(x'')|<\varepsilon.$$

定义 (序列式定义) : 设 f 为定义在 $D\subseteq\mathbb{R}^m$ 上的 m 元函数, $a\in D$. 若 D 中任意收敛于 a 的序列 $\{x_n\}$, 都有

$$\lim_{n\to\infty}f(x_n)=f(a),$$

则称 f 沿集合 D 在点 a 处连续.

定义 (ε-δ 式定义) : 设 f 是定义在 $D\subseteq\mathbb{R}^m$ 上的 m 元函数, $a\in D$. 若对任意的 $\varepsilon>0$, 存在

$\delta > 0$, 当 $x \in U(a, \delta) \cap D$ 时, 有

$$|f(x) - f(a)| < \varepsilon,$$

则称 f 沿集合 D 在点 a 处连续.

注: 由上述两种定义可以看出, 若 a 为 D 的孤立点, 则 f 必在 a 处连续 (因为 (1): 此时收敛于 a 的点列只有 $x_n = a, n = 1, 2, \cdots$, 而该点列必然满足 $\lim\limits_{n\to\infty} f(x_n) = f(a)$. (2) $U(a, \delta) \cap D = \{a\}$, 而 $|f(a) - f(a)| = 0 < \varepsilon$ 必然成立).

定义: 若 f 在 $D \subseteq \mathbb{R}^m$ 中的每一点都连续, 则称 f 在 D 上是连续的.

定理 (多元函数四则运算的连续性) : 设 $D \subseteq \mathbb{R}^m$, $a \in D$, m 元函数 $f(x)$ 与 $g(x)$ 在 $U(a, \eta) \cap D$ 内有定义, $\lambda \in \mathbb{R}$. 若 f 与 g 沿集合 D 在 a 点连续, 那么

$$f(x) + g(x), \quad \lambda f(x), \quad f(x)g(x), \quad \frac{f(x)}{g(x)}$$

也沿集合 D 在 a 点连续. 这里最后的 $\dfrac{f(x)}{g(x)}$ 要求 $g(a) \neq 0$.

1.2 重极限与累次极限

重极限与累次极限的存在性之间没有必然联系:

(1) 重极限存在, 累次极限都不存在, 或者只存在一个.

(2) 重极限不存在, 累次极限存在.

但是, 当存在性得到保证之后, 它们之间就有联系了.

(1) 若重极限与某个累次极限都存在, 则二者必然相等.

(2) 若重极限与两个累次极限都存在, 则它们三者必相等. (是 (1) 的直接推论)

1.3 有界闭集上连续函数的性质

定义: 设 $E \subseteq \mathbb{R}^m$, 若存在 $M > 0$, 对任意的 $x \in E$, 有

$$\|x\| \leqslant M,$$

则称 E 是 $\mathbb{R}^m$ 中的**有界集**.

定义: 设 $E \subseteq \mathbb{R}^m$, 若 E 中任意收敛点列的极限仍在 E 中, 则称 E 是 $\mathbb{R}^m$ 中的**闭集**.

定理 (波尔查诺-魏尔斯特拉斯聚点定理) : 设 $\{x_n\}$ 是 $\mathbb{R}^m$ 中的有界点列, 则 $\{x_n\}$ 存在收敛子列.

定理: 设 K 是 $\mathbb{R}^m$ 中的有界闭集, f 是 K 上的连续函数, 则 f 在 K 上有界.

定理: 设 K 是 $\mathbb{R}^m$ 中的有界闭集, f 是 K 上的连续函数, 则 f 在 K 上可以取到最大值与最小值.

定义: 设 f 是定义在 $D \subseteq \mathbb{R}^m$ 上的 m 元函数, 若对任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 当 $x_1, x_2 \in D$ 时,

只要 $\|x_1 - x_2\| < \delta$, 就有

$$|f(x_1) - f(x_2)| < \varepsilon,$$

则称 f 在 D 上**一致连续**.

定理: 设 K 是 $\mathbb{R}^m$ 中的有界闭集, f 是 K 上的连续函数, 则 f 在 K 上一致连续.

1.4 紧致性

定义: 设 (X, d) 为度量空间, K 为 X 的子集, 若对任意的 $\{x_n\} \subseteq K$, 存在子列 $\{x_{n_k}\}$ 收敛于 K 中一元素, 则称 K 为 (X, d) 中的**列紧集**.

定义: 设 (X, d) 为度量空间, K 为 X 的子集. 若对 K 的任意一族开覆盖, 都存在有限子覆盖, 则称 K 为 X 中的**紧集**.

定理 (连续映射保持紧致性) : 设有度量空间 (X, d_X), (Y, d_Y), $f : X \to Y$ 为连续映射, 则对任意的紧集 $K \subseteq X$, 像集 $f(K)$ 是 Y 中的紧集.

由此可以推得下面的结论.

定理: 设 (X, d) 为度量空间, K 为 (X, d) 中的紧集, f 为 K 上的连续函数, 则

(1) f 在 K 上有界.

(2) f 在 K 上能取到最大值与最小值.

(3) f 在 K 上一致连续.

注: $\mathbb{R}$ 中的闭区间 $[a, b]$ 上的连续函数的有界性、最值性与一致连续性就是因为 $[a, b]$ 是紧集.

注:

(1) 有限维赋范线性空间中, 有界闭集 $\Longleftrightarrow$ 紧集 $\Longleftrightarrow$ 列紧集.

(2) 度量空间中, 紧集 $\Longleftrightarrow$ 列紧集.

(3) 拓扑空间中, 紧集与列紧集稍有区别.

1.5 连通性

闭区间上的连续函数具有性质:

(1) 有界性.

(2) 取得最大值与最小值.

(3) 一致连续性.

(4) 介值性质.

前面我们证明了在紧集上连续的多元函数也具有性质 (1)(2)(3). 但是, 到目前为止, 我们还没有涉及性质 (4), 这是因为保证性质 (4) 的不是定义域的紧致性, 而是连通性.

定义: 设 $E \subseteq \mathbb{R}^m$, $x_0, x_1 \in E$, 若存在连续映射

$$\gamma : [0, 1] \longrightarrow E$$

满足 $\gamma(0)=x_0,\gamma(1)=x_1$, 则称 γ 为 E 中连接 x_0 与 x_1 的一条**路径**.

定义: 设 $E\subseteq\mathbb{R}^m$, 若对任意的 $x_0,x_1\in E$, 在 E 中均存在连接 x_0 与 x_1 两点的路径, 则称 E 是**连通**的.

定理 (连续映射保持连通性) : 设有度量空间 (X,d_X), (Y,d_Y), $f:X\to Y$ 为连续映射, 则对任意的连通集 $E\subseteq X$, 像集 $f(E)$ 是 Y 中的连通集.

由此可以得出下面的结论.

定理 (介值定理) : 设 E 为 $\mathbb{R}^m$ 的路径连通子集, f 为 E 上的连续函数, 则 f 具有介值性质.

2 多元微分学

2.1 偏导数与全微分

一、二元函数的情形

注: (1) 即使函数 $f(x,y)$ 在某一点沿任何方向的方向导数都存在, 也不能保证 $f(x,y)$ 在这一点连续. 例如

$$f(x,y)=\begin{cases}1, & x\in\mathbb{R},\ 0<y<x^2,\\ 0, & \text{其余}\end{cases}$$

在原点处沿任意方向 $\boldsymbol{v}$ 的方向导数均为 0.

(2) 即使函数 $f(x,y)$ 在某一点连续且在该点存在沿任意方向的方向导数, 也不能保证在这一点可微, 例如

$$f(x,y)=\begin{cases}\dfrac{2xy^3}{x^2+y^4}, & x^2+y^2\neq 0,\\ 0, & x^2+y^2=0.\end{cases}$$

由于当 $(x,y)\neq(0,0)$ 时,

$$|f(x,y)|=\left|\frac{2xy^3}{x^2+y^4}\right|=\left|\frac{2xy^2}{x^2+y^4}\right||y|\leqslant\left|\frac{x^2+y^4}{x^2+y^4}\right||y|=|y|,$$

故 $\lim\limits_{(x,y)\to(0,0)}f(x,y)=0=f(0,0)$, 即 $f(x,y)$ 在 $(0,0)$ 处连续. 设方向 $\boldsymbol{v}=(\cos\alpha,\sin\alpha)$, 则

$$\frac{\partial f}{\partial\boldsymbol{v}}(0,0)=\frac{2t\cos\alpha\sin^3\alpha}{\cos^2\alpha+t^2\sin^4\alpha}=0,$$

即 f 在 $(0,0)$ 处沿任意方向的方向导数均为 0. 假设 $f(x,y)$ 在 $(0,0)$ 处可微, 则

$$f(0+\Delta x,0+\Delta y)-f(0,0)=f_x(0,0)\Delta x+f_y(0,0)\Delta y+o(\sqrt{(\Delta x)^2+(\Delta y)^2}),$$

也即

$$\lim_{(\Delta x,\Delta y)\to(0,0)}\frac{f(\Delta x,\Delta y)}{\sqrt{(\Delta x)^2+(\Delta y)^2}}=0.$$

但是, 注意到

$$\lim_{\substack{\Delta x=(\Delta y)^2\\ \Delta y\to 0}} \frac{f(\Delta x,\Delta y)}{\sqrt{(\Delta x)^2+(\Delta y)^2}} \neq 0,$$

矛盾, 故 $f(x,y)$ 在 $(0,0)$ 处不可微.

$$\text{各偏导连续} \quad \Longrightarrow \quad \text{可微} \quad \begin{matrix} \Nearrow \\ \\ \Searrow \end{matrix} \quad \not\Leftarrow \left\{ \begin{matrix} \text{连续} \\ \not\Uparrow \\ \text{各方向导数存在} \end{matrix} \right.$$

二、m 元函数的一般情形

定义: 设 f 是定义在 $D \subseteq \mathbb{R}^m$ 上的函数, $x_0 \in D$, $\boldsymbol{v} \in \mathbb{R}^m$ 且 $\|\boldsymbol{v}\| = 1$, 若极限

$$\lim_{t\to 0^+} \frac{f(x_0+t\boldsymbol{v})-f(x_0)}{t}$$

存在, 则称此极限为函数 f 在 x_0 处沿方向 $\boldsymbol{v}$ 的**方向导数**, 记作

$$\frac{\partial f}{\partial \boldsymbol{v}}(x_0) \quad \text{或} \quad \partial_{\boldsymbol{v}} f(x_0).$$

定义: 设 $f(x) = f(x_1,\cdots,x_m)$ 为 m 元函数. 若极限

$$\lim_{t\to 0} \frac{f(x_0+t\boldsymbol{e}_i)-f(x_0)}{t}$$

存在, 则称此极限为函数 f 在 x_0 处关于变量 x_i 的偏导数, 记为

$$\frac{\partial f}{\partial x_i}(x_0).$$

注: 方向导数中要求 $t \to 0^+$(对应的是射线), 而偏导则要求 $t \to 0$(对应直线), 二者是不同的. (有些教材中方向导数的定义要求 $t \to 0$, 此时考察的是该方向所在直线的变化率)

定理: 设 $\boldsymbol{v} \in \mathbb{R}^m$ 为单位向量, 则

$$\frac{\partial f}{\partial \boldsymbol{v}} = \nabla f \cdot \boldsymbol{v},$$

其中 $\nabla f = \left(\dfrac{\partial f}{\partial x^1},\cdots,\dfrac{\partial f}{\partial x^m}\right)$ 为函数 f 的梯度.

定义: 若

$$f(x_0+\Delta x)-f(x_0) = \sum_{i=1}^m A_i \Delta x_i + o(\|\Delta x\|),$$

则称 f 在 x_0 处可微, 且称

$$\mathrm{d}f(x_0) = \sum_{i=1}^m A_i \mathrm{d}x_i$$

为函数 f 在 x_0 处的**全微分**.

注: 以二元函数为例, 可微的定义即是

$$f(x_0+h,y_0+k)-f(x_0,y_0)=Ah+Bk+o(\sqrt{h^2+k^2}).$$

对此, 我们有两种等价的表示方法:

$$\begin{aligned}f(x_0+h,y_0+k)-f(x_0,y_0)&=Ah+Bk+\varepsilon\sqrt{h^2+k^2},\\ f(x_0+h,y_0+k)-f(x_0,y_0)&=Ah+Bk+\alpha h+\beta k.\end{aligned}$$

其中 $\lim\limits_{(h,k)\to(0,0)}\varepsilon(h,k)=0$, $\lim\limits_{(h,k)\to(0,0)}\alpha(h,k)=0$, $\lim\limits_{(h,k)\to(0,0)}\beta(h,k)=0$.

下面我们证明它们的等价性. 显然 $o(\sqrt{h^2+k^2})$ 等价于 $\varepsilon\sqrt{h^2+k^2}$, 因此我们只需证明 $\varepsilon\sqrt{h^2+k^2}$ 等价于 $\alpha h+\beta k$. 由于

$$\varepsilon\sqrt{h^2+k^2}=\varepsilon\frac{h^2+k^2}{\sqrt{h^2+k^2}}=\frac{\varepsilon h}{\sqrt{h^2+k^2}}h+\frac{\varepsilon k}{\sqrt{h^2+k^2}}k,$$

令 $\alpha=\dfrac{\varepsilon h}{\sqrt{h^2+k^2}}$, $\beta=\dfrac{\varepsilon k}{\sqrt{h^2+k^2}}$, 则 $\varepsilon\sqrt{h^2+k^2}=\alpha h+\beta k$, 其中 $\alpha,\beta\to0((h,k)\to(0,0))$. 反之,

$$\alpha h+\beta k=\frac{\alpha h+\beta k}{\sqrt{h^2+k^2}}\sqrt{h^2+k^2},$$

令 $\varepsilon=\dfrac{\alpha h+\beta k}{\sqrt{h^2+k^2}}$, 则 $\alpha h+\beta k=\varepsilon\sqrt{h^2+k^2}$, 且 $\varepsilon\to0((h,k)\to(0,0))$.

定理 (各偏导连续 $\Longrightarrow$ 可微) : 若 m 元函数 f 的各个偏导

$$\frac{\partial f}{\partial x_i}(x)=\frac{\partial f}{\partial x_i}(x_1,\cdots,x_m)$$

在 x_0 处连续 $(i=1,\cdots,m)$, 则函数 f 在 x_0 处可微.

证明 以二元函数为例. 设 f 的偏导 f_x 与 f_y 在 (x_0,y_0) 的某邻域内连续, 下证 f 在 (x_0,y_0) 处可微. 设 (x_0+h,y_0+k) 属于上述邻域, 由一元函数的有限增量公式 (Lagrange 中值定理) 得

$$\begin{aligned}f(x_0+h,y_0+k)-f(x_0,y_0)&=f(x_0+h,y_0+k)-f(x_0,y_0+k)+f(x_0,y_0+k)-f(x_0,y_0)\\&=f_x(x_0+\theta_1h,y_0+k)h+f_y(x_0,y_0+\theta_2k)k.\end{aligned}$$

其中 $0<\theta_1,\theta_2<1$. 由 f_x 与 f_y 在 (x_0,y_0) 的连续性知

$$\begin{aligned}f_x(x_0+\theta_1h,y_0+k)&=f_x(x_0,y_0)+\alpha,\\ f_y(x_0,y_0+\theta_2k)&=f_y(x_0,y_0)+\beta.\end{aligned}\tag{1}$$

其中 α,β 满足当 $(h,k)\to(0,0)$ 时的极限为 0. 因此

$$f(x_0+h,y_0+k)-f(x_0,y_0)=f_x(x_0,y_0)h+f_y(x_0,y_0)k+\alpha h+\beta k.$$

因此 f 在 (x_0,y_0) 处可微.

定理 (混合偏导数交换求导次序)：设 f_{xy} 与 f_{yx} 均在点 (x_0,y_0) 邻近存在且在 (x_0,y_0) 处连续, 则

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

证明 令

$$\begin{aligned}\varphi(x)&=f(x,y_0+k)-f(x,y_0),\\ \psi(y)&=f(x_0+h,y)-f(x_0,y).\end{aligned}$$

易知

$$\varphi(x_0+h)-\varphi(x_0)=\begin{array}{c}f(x_0+h,y_0+k)-f(x_0+h,y_0)\\ -f(x_0,y_0+k)+f(x_0,y_0)\end{array}=\psi(y_0+k)-\psi(y_0)\ .$$

由于

$$\begin{aligned}\varphi'(x)&=f_x(x,y_0+k)-f_x(x,y_0),\\ \psi'(y)&=f_y(x_0+h,y)-f_y(x_0,y),\end{aligned}$$

故由一元函数的有限增量公式得

$$\begin{aligned}\varphi(x_0+h)-\varphi(x_0)&=\varphi'(x_0+\theta_1h)h=[f_x(x_0+\theta_1h,y_0+k)-f_x(x_0+\theta_1h,y_0)]h\\ &=f_{xy}(x_0+\theta_1h,y_0+\theta_2k)hk,\\ \psi(y_0+k)-\psi(y_0)&=\psi'(y_0+\theta_3k)k=[f_y(x_0+h,y_0+\theta_2k)-f_y(x_0,y_0+\theta_3k)]k\\ &=f_{yx}(x_0+\theta_4h,y_0+\theta_3k)hk.\end{aligned}$$

其中 $0<\theta_1,\theta_2,\theta_3,\theta_4<1$. 因此

$$f_{xy}(x_0+\theta_1h,y_0+\theta_2k)hk=f_{yx}(x_0+\theta_4h,y_0+\theta_3k)hk.$$

当 $h\neq0,k\neq0$ 时, 有

$$f_{xy}(x_0+\theta_1h,y_0+\theta_2k)=f_{yx}(x_0+\theta_4h,y_0+\theta_3k).$$

令 $(h,k)\to(0,0)$, 由 f_{xy} 与 f_{yx} 在 (x_0,y_0) 处的连续性可知

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

2.2 复合函数求导的链式法则

设

$$f(x)=f(x_1,\cdots,x_m)=f(\varphi_1(t),\cdots,\varphi_m(t))=F(t),$$

且 $f(x)$ 在 $x_0=\varphi(t_0)=(\varphi_1(t_0),\cdots,\varphi_m(t_0))$ 处可微. 下面我们来求 $F'(t_0)$.

$$\begin{aligned}F(t)-F(t_0)&=f(\varphi_1(t),\cdots,\varphi_m(t))-f(\varphi_1(t_0),\cdots,\varphi_m(t_0))=f(x)-f(x_0)\\&=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(x_0)(x-x_0)+o(\|x-x_0\|)\\&=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(\varphi(t_0))(\varphi_i(t)-\varphi_i(t_0))+o\left(\sqrt{\sum_{i=1}^m(\varphi_i(t)-\varphi_i(t_0))^2}\right).\end{aligned}$$

由于

$$\varphi_i(t)-\varphi_i(t_0)=\frac{\mathrm{d}\varphi_i}{\mathrm{d}t}(t_0)(t-t_0)+o(t-t_0),$$

故

$$F'(t_0)=\lim_{t\to t_0}\frac{F(t)-F(t_0)}{t-t_0}=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(\varphi_i(t_0))\frac{\mathrm{d}\varphi_i}{\mathrm{d}t}(t_0).$$

当不至于引起混淆时, 上式可以记为

$$\frac{\mathrm{d}F}{\mathrm{d}t}=\sum_{i=1}^m\frac{\partial f}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}t}.$$

由于求偏导可视为固定其他变元, 对某一变量求导, 因此设 $f(x)=f(x_1,\cdots,x_m)$, $x_i(t)=x_i(t_1,\cdots,t_k)$, 则若 $f(x)$ 在 $x_0=(x_0^1,\cdots,x_0^m)$ 可微, $x_i(t)$ 在 $t_0=(t_0^1,\cdots,t_0^k)$ 可微, $x_0^i=x_0^i(t_0)$, 则

$$\frac{\partial f}{\partial t_j}(t_0)=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(x_0)\frac{\partial x_i}{\partial t_j}(t_0).$$

当不至于引起混淆时, 简记为

$$\frac{\partial f}{\partial t_j}=\sum_{i=1}^m\frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t_j}.$$

其中 $j=1,\cdots,k$.

2.3 全微分的形式不变性

设 $f(x)=f(x_1,\cdots,x_m)$, $x_i=x_i(t)=x_i(t_1,\cdots,t_k)$, 则

$$\mathrm{d}f=\sum_{j=1}^k\frac{\partial f}{\partial t_j}\mathrm{d}t_j,\tag{1}$$

又由链式法则知

$$\frac{\partial f}{\partial t_j}=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t_j}.$$

因此

$$\mathrm{d}f=\sum_{j=1}^{k}\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t_j}\mathrm{d}t_j.$$

由于

$$\mathrm{d}x_i=\sum_{j=1}^{k}\frac{\partial x_i}{\partial t_j}\mathrm{d}t_j,$$

故

$$\mathrm{d}f=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\left(\sum_{j=1}^{k}\frac{\partial x_i}{\partial t_j}\mathrm{d}t_j\right)=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\mathrm{d}x_i,$$

即

$$\mathrm{d}f=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\mathrm{d}x_i. \tag{2}$$

由 (1)(2) 两式可以看出, 不论 $x=(x_1,\cdots,x_m)$ 是自变量还是中间变量, 函数 f 的全微分表示都是

$$\mathrm{d}f=\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\mathrm{d}x_i.$$

这一重要事实称为**全微分的形式不变性**.

利用全微分的形式不变性, 我们可以证明以下全微分的运算法则.

定理: 设 $u(x),v(x)$ 为 $D\subseteq\mathbb{R}^m$ 上的可微函数, $\lambda\in\mathbb{R}$, 则

(1) $\mathrm{d}(u(x)+v(x))=\mathrm{d}u(x)+\mathrm{d}v(x)$.

(2) $\mathrm{d}(\lambda u(x))=\lambda\mathrm{d}u(x)$.

(3) $\mathrm{d}(u(x)v(x))=v(x)\mathrm{d}u(x)+u(x)\mathrm{d}v(x)$.

(4) $\mathrm{d}\left(\dfrac{u(x)}{v(x)}\right)=\dfrac{v(x)\mathrm{d}u(x)-u(x)\mathrm{d}v(x)}{(v(x))^2}$.

证明　考察

$$f(u,v)=uv,\quad g(u,v)=\frac{u}{v}.$$

易得

$$\frac{\partial f}{\partial u}=v,\quad \frac{\partial f}{\partial v}=u,\qquad \frac{\partial g}{\partial u}=\frac{1}{v},\quad \frac{\partial f}{\partial v}=-\frac{u}{v^2}.$$

从而

$$\mathrm{d}f(u,v)=\mathrm{d}(uv)=v\mathrm{d}u+u\mathrm{d}v,\quad \mathrm{d}g(u,v)=\mathrm{d}\left(\frac{u}{v}\right)=\frac{1}{v}\mathrm{d}u-\frac{u}{v^2}\mathrm{d}v.$$

由全微分的形式不变性知

$$\mathrm{d}(u(x)v(x)) = v(x)\mathrm{d}u(x) + u(x)\mathrm{d}v(x), \quad \mathrm{d}\left(\frac{u(x)}{v(x)}\right) = \frac{v(x)\mathrm{d}u(x) - u(x)\mathrm{d}v(x)}{(v(x))^2}.$$

例 设 $f(x,y)$ 是 n 阶连续可微函数, $\varphi(t) = f(x+th, y+tk)$, 求 $\varphi^{(n)}(t)$.

解 由链式法则知

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(x+th, y+tk).$$

即是说, 用算子 $\frac{\mathrm{d}}{\mathrm{d}t}$ 作用一下 $\varphi(t)$, 相当于用算子 $\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)$ 作用一下 $f(x+th, y+tk)$. 因此

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(x+th, y+tk),$$

$$\cdots$$

$$\frac{\mathrm{d}^n\varphi}{\mathrm{d}t^n} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x+th, y+tk).$$

由于连续可微函数的求偏导次序可以交换, 因此 $\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n$ 可按照二项式定理展开:

$$\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n = \sum_{i=0}^{n}\binom{n}{i}h^{n-i}k^i\frac{\partial^n}{\partial x^{n-i}\partial y^i}.$$

从而

$$\frac{\mathrm{d}^n\varphi}{\mathrm{d}t^n} = \sum_{i=0}^{n}\binom{n}{i}h^{n-i}k^i\frac{\partial^n}{\partial x^{n-i}\partial y^i}f(x+th, y+tk).$$

例 设 $f(x_1,\cdots,x_m)$ 是 n 阶连续可微函数, $\varphi(t) = f(x_1+th_1,\cdots,x_m+th_m)$, 求 $\varphi^{(n)}(t)$.

解 由链式法则知

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right) f(x_1+th_1,\cdots,x_m+th_m).$$

故用算子 $\frac{\mathrm{d}}{\mathrm{d}t}$ 作用一下 $\varphi(t)$, 相当于用算子 $\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)$ 作用一下 $f(x_1+th_1,\cdots,x_m+th_m)$. 因此

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2} = \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^2 f(x_1+th_1,\cdots,x_m+th_m),$$

$$\cdots$$

$$\frac{\mathrm{d}^n\varphi}{\mathrm{d}t^n} = \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^n f(x_1+th_1,\cdots,x_m+th_m).$$

由于连续可微函数的求偏导次序可以交换, 故 $\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^n$ 可以多项式展开为:

$$\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^n=\sum_{i_1+\cdots+i_m=n}\binom{n}{i_1,\cdots,i_m}h_1^{i_1}\cdots h_m^{i_m}\frac{\partial^n}{\partial x_1^{i_1}\cdots\partial x_m^{i_m}}.$$

从而

$$\frac{\mathrm{d}^n\varphi}{\mathrm{d}t^n}=\sum_{i_1+\cdots+i_m=n}\binom{n}{i_1,\cdots,i_m}h_1^{i_1}\cdots h_m^{i_m}\frac{\partial^n}{\partial x_1^{i_1}\cdots\partial x_m^{i_m}}f(x_1+th_1,\cdots,x_m+th_m).$$

注: 多项式定理:

$$(u_1+\cdots+u_m)^n=\sum_{i_1+\cdots+i_m=n}\binom{n}{i_1,\cdots,i_m}u_1^{i_1}\cdots u_m^{i_m},$$

其中 $i_1,\cdots,i_m$ 非负, $i_1+\cdots+i_m=n$, 多项式系数

$$\binom{n}{i_1,\cdots,i_m}=\frac{n!}{i_1!\cdots i_m!}.$$

显然二项式定理是多项式定理的特殊情况, 二项式系数是多项式系数的特殊情况.

2.4 有限增量公式

定理 (拉格朗日中值定理) : 设 D 为 $\mathbb{R}^m$ 中的开集, $f(x)=f(x_1,\cdots,x_m)$ 在 $\overline{D}$ 上连续, D 上可微, 则对任意的 $a,a+h$, 只要 $[a,a+h]\subseteq\overline{D}$, 则有

$$f(a+h)=f(a)+\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}(a+\theta h)h_i, \tag{1}$$

其中 $a=(a_1,\cdots,a_m)$, $h=(h_1,\cdots,h_m)$, $0<\theta<1$.

注: 设 $a,b\in\mathbb{R}^m$, 则分别称

$$(a,b)=\{a+t(b-a)\,|\,0<t<1\},$$
$$[a,b]=\{a+t(b-a)\,|\,0\leqslant t\leqslant 1\}$$

为连接 a 与 b 的开线段与闭线段.

定理: 设 $f(x)=f(x_1,\cdots,x_m)$ 在开区域 $D\subseteq\mathbb{R}^m$ 上可微, 若

$$\frac{\partial f}{\partial x_1}=\cdots=\frac{\partial f}{\partial x_m}=0,\quad x\in D,$$

则 f 为 D 内的常值函数.

推论: 设 $f(x)=f(x_1,\cdots,x_m)$ 在开区域 $D\subseteq\mathbb{R}^m$ 上可微, 在 $\overline{D}$ 上连续, 若

$$\frac{\partial f}{\partial x_1}=\cdots=\frac{\partial f}{\partial x_m}=0,\quad x\in D,$$

则 f 在 $\overline{D}$ 为常值函数.

2.5 泰勒公式

定理 (带拉格朗日型余项的泰勒公式)：设 D 为 $\mathbb{R}^m$ 中的开集, $f \in C^{n+1}(D)$, $[a, a+h] \subseteq D$, 则

$$f(a+h) = \sum_{i=0}^{n} \frac{1}{i!}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^i f(a) + \frac{1}{(n+1)!}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^{n+1} f(a+\theta h).$$

其中 $a = (a_1, \cdots, a_m)$, $h = (h_1, \cdots, h_m)$, $0 < \theta < 1$.

思路　对一元函数 $\varphi(t) = f(a+th)$, $t \in [0,1]$ 应用泰勒公式, 则有

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + \cdots + \frac{\varphi^{(n)}(0)}{n!}t^n + \frac{\varphi^{(n+1)}(\theta)}{(n+1)!}t^{n+1},$$

其中 $0 < \theta < 1$. 故

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{\varphi''(0)}{2!} + \cdots + \frac{\varphi^{(n)}(0)}{n!} + \frac{\varphi^{(n+1)}(\theta)}{(n+1)!}.$$

由 §2.3 的例题知

$$\varphi^{(k)}(t) = \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^k f(a+th),$$

带入即可得到结论.

定理: (带皮亚诺型余项的泰勒公式)：设 f 在点 $a \in \mathbb{R}^m$ 附近 n 阶连续可微, 则

$$f(a+h) = \sum_{i=0}^{n} \frac{1}{i!}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^i f(a) + o(\|h\|^n).$$

注: 带皮亚诺型余项的泰勒公式反映的是函数 f 在 a 点附近的渐进性质, 而带拉格朗日型余项的泰勒公式反映的是大范围性质.

2.6 重指标表示泰勒公式

定义: 设 $\alpha_1, \cdots, \alpha_m$ 为非负整数, 称

$$\alpha = (\alpha_1, \cdots, \alpha_m)$$

为一个**重指标**.

对于重指标 $\alpha=(\alpha_1,\cdots,\alpha_m)$ 与 $\mathbb{R}^m$ 中的点 $h=(h_1,\cdots,h_m)$, 我们规定

$$\begin{aligned}
|\alpha| &= \alpha_1+\cdots+\alpha_m,\\
\alpha! &= \alpha_1!\cdots\alpha_m!,\\
h^\alpha &= h_1^{\alpha_1}\cdots h_m^{\alpha_m},\\
\partial^\alpha &= \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_m^{\alpha_m}}.
\end{aligned}$$

由此可得

$$\begin{aligned}
\frac{1}{i!}\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^i &= \sum_{i_1+\cdots+i_m=n}\frac{1}{i_1!\cdots i_m!}h_1^{i_1}\cdots h_m^{i_m}\frac{\partial^i}{\partial x_1^{i_1}\cdots\partial x_m^{i_m}}\\
&= \sum_{|\alpha|=i}\frac{h^\alpha}{\alpha!}\partial^\alpha.
\end{aligned}$$

故泰勒公式

$$f(a+h)=\sum_{i=0}^n\frac{1}{i!}\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^i f(a)+R_{n+1}$$

可以改写为

$$\begin{aligned}
f(a+h) &= \sum_{i=0}^n\sum_{|\alpha|=i}\frac{h^\alpha}{\alpha!}\partial^\alpha f(a)+R_{n+1}\\
&= \sum_{|\alpha|=0}^n\frac{\partial^\alpha f(a)}{\alpha!}h^\alpha+R_{n+1}.
\end{aligned}$$

这样一来, 多元泰勒公式就与一元泰勒公式在形式上保持一致了. 而余项则可以表示为

$$R_{n+1}=\sum_{|\beta|=n+1}\frac{\partial^\beta f(a+\theta h)}{\beta!}h^\beta.$$

2.7 隐函数

要确定一个隐函数, 不光要指出 x 的变化范围, 还要指出 y 的变化范围.

定理: 设 $F(x,y)$ 在包含 (x_0,y_0) 的开集 D 上连续可微, 且

$$\begin{aligned}
F(x_0,y_0) &= 0,\\
\frac{\partial F}{\partial y}(x_0,y_0) &\neq 0,
\end{aligned}$$

则方程

$$F(x,y)=0$$

在 (x_0, y_0) 的某邻域内唯一确定了连续可微函数

$$y = y(x),$$

使得在该邻域内

$$F(x, y(x)) = 0.$$

且 $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ 可由

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

求出.

定理: 设函数组

$$F_1(x, y_1, \cdots, y_p), \cdots, F_p(x, y_1, \cdots, y_p)$$

在含有点 $(x_0, y_1^0, \cdots, y_p^0)$ 的开集内连续可微, 且

$$F_i(x_0, y_1^0, \cdots, y_p^0) = 0, \quad i = 1, \cdots, p,$$
$$\frac{\partial(F_1, \cdots, F_p)}{\partial(y_1, \cdots, y_p)}(x_0, y_1^0, \cdots, y_p^0) \neq 0,$$

则在点 $(x_0, y_1^0, \cdots, y_p^0)$ 的邻域内确定了唯一的连续可微函数组

$$y_1 = y_1(x), \cdots, y_p = y_p(x),$$

满足

$$F_i(x, y_1(x), \cdots, y_p(x)) = 0, \quad i = 1, \cdots, p,$$

并且它们的导数可由下列方程组求出:

$$\frac{\partial F_i}{\partial x} + \sum_{j=1}^{p} \frac{\partial F_i}{\partial y_j}\frac{\mathrm{d}y_j}{\mathrm{d}x} = 0, \quad i = 1, \cdots, p.$$

也即

$$\begin{cases} \dfrac{\partial F_1}{\partial x} + \dfrac{\partial F_1}{\partial y_1}\dfrac{\mathrm{d}y_1}{\mathrm{d}x} + \cdots + \dfrac{\partial F_1}{\partial y_p}\dfrac{\mathrm{d}y_p}{\mathrm{d}x} = 0, \\ \qquad\qquad\qquad \cdots \\ \dfrac{\partial F_p}{\partial x} + \dfrac{\partial F_p}{\partial y_1}\dfrac{\mathrm{d}y_1}{\mathrm{d}x} + \cdots + \dfrac{\partial F_p}{\partial y_p}\dfrac{\mathrm{d}y_p}{\mathrm{d}x} = 0, \end{cases}$$

或

$$\begin{pmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_p} \\ \vdots & & \vdots \\ \dfrac{\partial F_p}{\partial y_1} & \cdots & \dfrac{\partial F_p}{\partial y_p} \end{pmatrix} \begin{pmatrix} \dfrac{\mathrm{d}y_1}{\mathrm{d}x} \\ \vdots \\ \dfrac{\mathrm{d}y_p}{\mathrm{d}x} \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial F_1}{\partial x} \\ \vdots \\ \dfrac{\partial F_p}{\partial x} \end{pmatrix}.$$

定理: 设函数组

$$\begin{cases} F_1(x_1, \cdots, x_m, y_1, \cdots, y_p), \\ \cdots \\ F_p(x_1, \cdots, x_m, y_1, \cdots, y_p), \end{cases}$$

在包含 $(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^p)$ 的开集 D 上连续可微. 若

$$F_i(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^p) = 0, \quad i = 1, \cdots, p,$$

$$\frac{\partial(F_1, \cdots, F_p)}{\partial(y_1, \cdots, y_p)}(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^p) \neq 0,$$

则在点 $(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^p)$ 的邻域内确定了唯一的连续可微函数

$$y_1 = y_1(x_1, \cdots, x_m), \cdots, y_p = y_p(x_1, \cdots, x_m)$$

满足

$$F_i(x_1, \cdots, x_m, y_1(x_1, \cdots, x_m), \cdots, y_p(x_1, \cdots, x_m)) = 0, \quad i = 1, \cdots, p.$$

并且它们对变元 x_k 的偏导数 $\dfrac{\partial y_j}{\partial x_k}(j = 1, \cdots, p)$ 可由下述方程组求得:

$$\frac{\partial F_i}{\partial x_k} + \sum_{j=1}^{p} \frac{\partial F_i}{\partial y_j} \frac{\partial y_j}{\partial x_k} = 0, \quad i = 1, \cdots, p$$

也即

$$\begin{cases} \dfrac{\partial F_1}{\partial x_k} + \dfrac{\partial F_1}{\partial y_1}\dfrac{\partial y_1}{\partial x_k} + \cdots + \dfrac{\partial F_1}{\partial y_p}\dfrac{\partial y_p}{\partial x_k} = 0, \\ \cdots \\ \dfrac{\partial F_p}{\partial x_k} + \dfrac{\partial F_p}{\partial y_1}\dfrac{\partial y_1}{\partial x_k} + \cdots + \dfrac{\partial F_p}{\partial y_p}\dfrac{\partial y_p}{\partial x_k} = 0, \end{cases}$$

或

$$\begin{pmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_p} \\ \vdots & & \vdots \\ \dfrac{\partial F_p}{\partial y_1} & \cdots & \dfrac{\partial F_p}{\partial y_p} \end{pmatrix} \begin{pmatrix} \dfrac{\partial y_1}{\partial x_k} \\ \vdots \\ \dfrac{\partial y_p}{\partial x_k} \end{pmatrix} = - \begin{pmatrix} \dfrac{\partial F_1}{\partial x_k} \\ \vdots \\ \dfrac{\partial F_p}{\partial x_k} \end{pmatrix}.$$

2.8 多元函数极值的充分与必要条件

定理 (极值的必要条件) : 设 $f = f(x_1, \cdots, x_m)$ 在点 a 处取极值, 若 f 在点 a 可微, 则

$$\frac{\partial f}{\partial x_1}(a) = 0\,, \cdots, \frac{\partial f}{\partial x_m}(a) = 0.$$

证明 反证法. 不妨设 f 在点 a 取得极大值. 由可微的定义与极值的定义知存在 a 的某邻域 $U(a, \eta)$, 使得 $f(a)$ 为该邻域内的最大值, 且对任意的 $a + h \in U(a, \eta)$, 有

$$f(a+h) - f(a) = \sum_{i=1}^{m} A_i h_i + o(\|h\|), \tag{1}$$

其中 $A_i = \dfrac{\partial f}{\partial x_i}(a)$. 假设 A_i 不全为 0, 令 $\bar{h} = (A_1\varepsilon, \cdots, A_m\varepsilon)$, 其中 $\varepsilon > 0$ 充分小, 满足 $a+\bar{h} \in U(a, \eta)$. 由 (1) 式可知

$$f(a+\bar{h}) - f(a) = \left(\sum_{i=1}^{m} A_i^2\right)\varepsilon + o(\varepsilon).$$

因此当 ε 充分小时,

$$f(a+\bar{h}) - f(a) > 0,$$

与 f 在 a 处取极大值矛盾.

定义: 若 $f = f(x_1, \cdots, x_m)$ 在其可微点 a 处的各个偏导均为 0, 则称 a 为 f 的**临界点**.

由前述定理即得: a 为 f 的极值点的必要条件为 a 为 f 的临界点.

现在假设 $f = f(x_1, \cdots, x_m)$ 在点 $a = (a_1, \cdots, a_m)$ 处是二阶连续可微的, 且 a 为 f 的临界点:

$$\frac{\partial f}{\partial x_1}(a) = 0\,, \cdots, \frac{\partial f}{\partial x_m}(a) = 0.$$

在此条件下, 我们来探讨 f 在点 a 取极值的充分条件.

由于 a 为 f 的临界点, 故由泰勒展式知

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_m\frac{\partial}{\partial x_m}\right)^2 f(a) + o(\|h\|^2) \\ &= \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|^2) \\ &= \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij} h_i h_j + o(\|h\|^2) \end{aligned}$$

其中

$$A_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

定义: 设函数 $f=f(x_1,\cdots,x_m)$ 在点 a 临近是二阶连续可微的, 我们称

$$H_f(a)=\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1\partial x_m}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_m\partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_m^2}(a) \end{pmatrix}$$

为函数 f 在点 a 处的**黑塞矩阵**.

定理 (极值的必要条件) : 设 f 在点 a 邻近具有二阶连续偏导, 则

(1) 若 a 为 f 的极小值点, 则 $H_f(a)$ 半正定.

(2) 若 a 为 f 的极大值点, 则 $H_f(a)$ 半负定.

引理: 设二次型 $Q(\xi)$ 是正定的, 则存在常数 $\sigma>0$, 对任意的 $\xi\in\mathbb{R}^m$, 有

$$Q(\xi)\geqslant\sigma\|\xi\|^2.$$

证明　设

$$S=\{\xi\in\mathbb{R}^m \mid \|\xi\|=1\}$$

为 $\mathbb{R}^m$ 中的单位球面. 由于 S 是 $\mathbb{R}^m$ 中的有界闭集, $Q(\xi)$ 是连续函数, 故 $Q(\xi)$ 在 S 上可以取到最小值. 设

$$Q(\xi_0)=\min_{\xi\in S}Q(\xi)=\sigma.$$

因为 $Q(\xi)$ 正定, 且 $\xi_0\neq 0$, 故 $\sigma>0$. 注意到对任意的 $0\neq\xi\in\mathbb{R}^m$, 有 $\dfrac{\xi}{\|\xi\|}\in S$, 从而

$$Q\left(\frac{\xi}{\|\xi\|}\right)\geqslant\sigma,$$

从而

$$\frac{1}{\|\xi\|^2}Q(\xi)\geqslant\sigma,$$

即

$$Q(\xi)\geqslant\sigma\|\xi\|^2.$$

显然当 $\xi=0$ 时不等式也成立. 综上, 结论得证.

注: 从上述证明过程我们可以看出对于正定二次型 $Q(\xi)$, 有双边不等式成立:

$$\sigma_1\|\xi\|^2\leqslant Q(\xi)\leqslant\sigma_2\|\xi\|^2,$$

其中 σ_1,σ_2 为正常数, $\xi\in\mathbb{R}^m$.

定理 (极值的充分条件) : 设函数 $f=f(x_1,\cdots,x_m)$ 在点 a 临近是二阶连续可微的, a 为 f 的

临界点. 若 f 在点 a 处的黑塞矩阵 $H_f(a)$ 是正定的, 则 a 为 f 的严格极小值点; 若 $H_f(a)$ 是负定的, 则 a 为 f 的严格极大值点; 若 $H_f(a)$ 是不定的, 则 a 不是 f 的极值点.

证明 记 $A_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$. 因为 a 为临界点, 故由泰勒展式得

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij}h_ih_j + o(\|h\|^2) \\ &= \frac{1}{2}Q(h) + o(\|h\|^2) \\ &\geqslant \frac{1}{2}\sigma\,\|h\|^2 + o(\|h\|^2) \\ &= \left(\frac{1}{2}\sigma + o(1)\right)\|h\|^2, \end{aligned}$$

故存在 $\delta > 0$, 当 $0 < \|h\| < \delta$ 时, 有

$$f(a+h) - f(a) \geqslant \frac{1}{4}\sigma\,\|h\|^2 > 0.$$

因此, a 为 f 的严格极小值点.

$H_f(a)$ 负定的情况类似可得. $H_f(a)$ 不定的情况是前述极值必要条件的推论.

2.9 条件极值

本节我们总假定约束条件

$$\begin{cases} g_1(x) = g_1(x_1, \cdots, x_n) = 0, \\ \cdots \\ g_m(x) = g_m(x_1, \cdots, x_n) = 0 \end{cases}$$

的 Jacobi 矩阵

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial g_m}{\partial x_1} & \cdots & \dfrac{\partial g_m}{\partial x_n} \end{pmatrix}$$

在所涉及的范围内是行满秩的, 其中 $m < n$.

处理条件极值的一个巧妙方法, 是拉格朗日提出的待定乘数法.

定理 (拉格朗日乘数法) : 设 $a = (a_1, \cdots, a_n)$ 为目标函数

$$f(x) = f(x_1, \cdots, x_n)$$

在约束条件

$$\begin{cases} g_1(x) = g_1(x_1, \cdots, x_n) = 0, \\ \cdots \\ g_m(x) = g_1(x_1, \cdots, x_n) = 0 \end{cases}$$

下的极值点, 则存在常量 $\lambda^0 = (\lambda_1^0, \cdots, \lambda_m^0)$, 使得 (a, λ^0) 为函数

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

的临界点, 也即

$$\begin{cases} \dfrac{\partial f}{\partial x_j}(a) + \displaystyle\sum_{i=1}^{m} \lambda_i^0 \dfrac{\partial g_i}{\partial x_j}(a) = 0, & j = 1, \cdots, n, \\ g_i(a) = 0, & i = 1, \cdots, m. \end{cases}$$

3 向量函数微分学

3.1 线性映射、矩阵范数

满足:

(1) $\|A\| \geqslant 0$, 且 $\|A\| = 0$ 当且仅当 $A = O$.

(2) $\|cA\| = |c| \, \|A\|$.

(3) $\|A + B\| \leqslant \|A\| + \|B\|$.

(4) $\|AB\| \leqslant \|A\| \, \|B\|$.

则称 $\|\cdot\|$ 为一个矩阵范数.

例 对矩阵 $A = (a_{ij})_{m\times n}$, 定义 F-范数 (**Frobenius 范数**) 为

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}.$$

证明 只验证 $\|AB\|_F \leqslant \|A\|_F \, \|B\|_F$. 设 $B = (b_{ij})_{n\times t}$, 令 $AB = (c_{ij})_{m\times t}$. 由柯西不等式知

$$c_{ij}^2 = \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right)^2 \leqslant \left(\sum_{k=1}^{n} |a_{ik}||b_{kj}| \right)^2 \leqslant \left(\sum_{k=1}^{n} a_{ik}^2 \right) \left(\sum_{k=1}^{n} b_{kj}^2 \right).$$

从而

$$\|AB\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}^2 \right)^{\frac{1}{2}} \leqslant \left[\sum_{i=1}^{m} \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_{ik}^2 \right) \left(\sum_{k=1}^{n} b_{kj}^2 \right) \right]^{\frac{1}{2}} = \left(\sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik}^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{m} \sum_{j=1}^{n} b_{kj}^2 \right)^{\frac{1}{2}} = \|A\|_F \, \|B\|_F.$$

注: 这里用到了结论 $\left(\sum_{i=1}^{m} X_i\right)\left(\sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} X_i Y_j$. $\left(令 X_i = \sum_{k=1}^{n} a_{ik}^2, Y_j = \sum_{k=1}^{n} b_{kj}^2 即可\right)$

例　对矩阵 $A = (a_{ij})_{m\times n}$, 定义 ∞ 范数 (**行范数**) 为

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|.$$

证明　只证明 $\|AB\|_\infty \leqslant \|A\|_\infty \|B\|_\infty$. 设 $B = (b_{ij})_{n\times t}$, 令 $AB = (c_{ij})_{m\times t}$, 则

$$\begin{aligned}
\|AB\|_\infty = \max_i \sum_{j=1}^{n} |c_{ij}| &= \max_i \sum_{j=1}^{t} \left|\sum_{k=1}^{n} a_{ik} b_{kj}\right| \\
&\leqslant \max_i \sum_{j=1}^{t}\sum_{k=1}^{n} |a_{ik}||b_{kj}| = \max_i \left[\sum_{k=1}^{n} |a_{ik}| \left(\sum_{j=1}^{t} |b_{kj}|\right)\right] \\
&\leqslant \max_i \left[\sum_{k=1}^{n} |a_{ik}| \max_k \left(\sum_{j=1}^{t} |b_{kj}|\right)\right] \\
&\leqslant \left(\max_i \sum_{k=1}^{n} |a_{ik}|\right)\left(\max_k \sum_{j=1}^{t} |b_{kj}|\right) = \|A\|_\infty \|B\|_\infty .
\end{aligned}$$

性质 (矩阵范数与向量范数的相容性) : F-范数与 ∞ 范数均满足

$$\|Ax\| \leqslant \|A\| \, \|x\| ,$$

其中 $A = (a_{ij})$ 为 $m \times n$ 矩阵, x 为 n 维列向量.

证明　对 F-范数, 由柯西不等式得

$$\|Ax\|^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 \leqslant \sum_{i=1}^{m}\left[\left(\sum_{j=1}^{n} a_{ij}^2\right)\left(\sum_{j=1}^{n} x_j^2\right)\right] = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)\left(\sum_{j=1}^{n} x_j^2\right) = \|A\| \, \|x\| .$$

对 ∞ 范数, 有

$$\|Ax\| = \max_i \left|\sum_{j=1}^{n} a_{ij}x_j\right| \leqslant \max \sum_{j=1}^{n} |a_{ij}||x_j| \leqslant \max \sum_{j=1}^{n} |a_{ij}|(\max_j |x_j|) = \|A\| \, \|x\| .$$

注: 与行范数类似, 我们可以定义矩阵 $A = (a_{ij})_{m\times n}$ 的列范数

$$\|A\|_1 = \max_j \sum_{i=1}^{m} |a_{ij}|.$$

此外, 我们还可以定义矩阵的 2 范数 (**特征值范数**):

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)},$$

其中 $\lambda_{\max}(A^H A)$ 表示矩阵 $A^H A$ 的最大特征值. 注意到 $A^H A$ 为 Hermite 矩阵, 其特征值必为实数, 因此谈论其特征值的最大值是有意义的. (当 A 为实矩阵时, $A^H A$ 就变为 $A^{\mathrm{T}}A$)

3.2 向量值函数的微分

此后若无特别说明, 矩阵范数均使用 F-范数或 ∞-范数, 即对 $A=(a_{ij})_{m\times n}$,

$$\|A\|_F=\left(\sum_{i=1}^m\sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}\quad \text{或}\quad \|A\|_\infty=\max_i\sum_{j=1}^n|a_{ij}|.$$

向量范数也使用相对应的欧几里得范数或最大值范数, 即对向量 $\boldsymbol{x}=(x_1,\cdots,x_n)$,

$$\|\boldsymbol{x}\|_2=\left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}\quad \text{或}\quad \|\boldsymbol{x}\|_\infty=\max_i|x_i|.$$

并且我们不再对范数加下标进行区分, 因为对于后面所出现的结论而言, 使用何种范数并无本质区别.

定义在 $E\subseteq\mathbb{R}$ 上的一元数值函数 $f(x)$ 在 x_0 处的微分是一个线性函数:

$$\begin{aligned}\mathscr{A}:\mathbb{R}&\longrightarrow\mathbb{R}\\ h&\longmapsto f'(x_0)h\end{aligned}$$

我们将其推广到一般的向量值函数上.

定义: 设 G 为 $\mathbb{R}^n$ 的开子集, 映射 $f:G\to\mathbb{R}^m$, $\boldsymbol{x}_0\in G$, 若存在线性映射 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^m$, 使得

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\|f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-\mathscr{A}\boldsymbol{h}\|}{\|\boldsymbol{h}\|}=0,\tag{1}$$

则称 f 在 $\boldsymbol{x}_0$ 处是**可微的**.

注: (1) 式等价于

$$\|f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-\mathscr{A}\boldsymbol{h}\|=o(\|\boldsymbol{h}\|)$$

或者

$$\|f(\boldsymbol{x})-f(\boldsymbol{x}_0)-\mathscr{A}(\boldsymbol{x}-\boldsymbol{x}_0)\|=o(\|\boldsymbol{x}-\boldsymbol{x}_0\|).$$

命题: 满足上述定义的线性映射 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^m$ 是唯一的.

证明 假设另有线性映射 $\mathscr{B}:\mathbb{R}^n\to\mathbb{R}^m$ 满足上述定义. 任取 $\boldsymbol{h}\in\mathbb{R}^n$ 与充分小的 $\varepsilon>0$, 满足

$\boldsymbol{x}_0 + \varepsilon \boldsymbol{h} \in G$. 下证 $(\mathscr{B} - \mathscr{A})\boldsymbol{h} = \boldsymbol{0}$. 易知

$$\begin{aligned}
\|(\mathscr{B} - \mathscr{A})\boldsymbol{h}\| &= \frac{1}{\varepsilon}\|(\mathscr{B} - \mathscr{A})(\varepsilon\boldsymbol{h})\| = \frac{1}{\varepsilon}\|\mathscr{B}(\varepsilon\boldsymbol{h}) - \mathscr{A}(\varepsilon\boldsymbol{h})\| \\
&= \frac{1}{\varepsilon}\|[\mathscr{B}(\varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) + f(\boldsymbol{x}_0)] + [f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0) - \mathscr{A}(\varepsilon\boldsymbol{h})]\| \\
&\leqslant \frac{1}{\varepsilon}(\|f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0) - \mathscr{B}(\varepsilon\boldsymbol{h})\| + \|f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0) - \mathscr{A}(\varepsilon\boldsymbol{h})\|) \\
&= \|h\|\left(\frac{\|f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0) - \mathscr{B}(\varepsilon\boldsymbol{h})\|}{\|\varepsilon\boldsymbol{h}\|} + \frac{\|f(\boldsymbol{x}_0 + \varepsilon\boldsymbol{h}) - f(\boldsymbol{x}_0) - \mathscr{A}(\varepsilon\boldsymbol{h})\|}{\|\varepsilon\boldsymbol{h}\|}\right).
\end{aligned}$$

令 $\varepsilon \to 0$, 则 $\|((\mathscr{B} - \mathscr{A})(\boldsymbol{h}))\| = 0$, 从而

$$(\mathscr{B} - \mathscr{A})\boldsymbol{h} = \boldsymbol{0}.$$

由 $\boldsymbol{h}$ 的任意性知

$$\mathscr{B} = \mathscr{A}.$$

由此, 我们可以称满足上述定义的线性映射 $\mathscr{A} : \mathbb{R}^n \to \mathbb{R}^m$ 为向量值函数 f 在 $\boldsymbol{x}_0$ 处的**微分**或**全导数**, 记作

$$\mathrm{D}f(\boldsymbol{x}_0) = \mathscr{A}.$$

性质:

(1) $\mathrm{D}(f + g)(\boldsymbol{x}_0) = \mathrm{D}f(\boldsymbol{x}_0) + \mathrm{D}g(\boldsymbol{x}_0)$.

(2) $\mathrm{D}(\lambda f)(\boldsymbol{x}_0) = \lambda\mathrm{D}f(\boldsymbol{x}_0)$.

定理 (可微 $\Longrightarrow$ 连续) : 设向量值函数 $f : G(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$ 在点 $\boldsymbol{x}_0$ 可微, 则存在常数 $\gamma > 0$ 与 $\boldsymbol{x}_0$ 的某邻域 $U(\boldsymbol{x}_0)$, 对任意的 $\boldsymbol{x} \in U(\boldsymbol{x}_0)$, 有

$$\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leqslant \gamma\|\boldsymbol{x} - \boldsymbol{x}_0\|.$$

进而由此得出: 若 f 在 $\boldsymbol{x}_0$ 处可微, 则 f 在 $\boldsymbol{x}_0$ 处连续.

定理 (链式法则) : 若向量值函数 f 在 $\boldsymbol{x}_0$ 处可微, 向量值函数 g 在 $f(\boldsymbol{x}_0)$ 处可微, 则 $g \circ f$ 在 $\boldsymbol{x}_0$ 处可微, 且

$$\mathrm{D}(g \circ f)(\boldsymbol{x}_0) = \mathrm{D}g(f(\boldsymbol{x}_0))\mathrm{D}f(\boldsymbol{x}_0).$$

定理: 开集 $G \subseteq \mathbb{R}^n$ 上的向量值函数

$$f(\boldsymbol{x}) = \begin{pmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{pmatrix}$$

在 $\boldsymbol{x}_0$ 处可微的充要条件为每个分量函数 $f_i(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 处可微. 满足此条件时, f 在点 $\boldsymbol{x}_0$ 处的微分为

$$\mathrm{D}f(\boldsymbol{x}_0)=\begin{pmatrix}\dfrac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\boldsymbol{x}_0)\\ \vdots & & \vdots\\ \dfrac{\partial f_m}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\boldsymbol{x}_0)\end{pmatrix}.$$

注: 我们看到, 向量值函数在 $\boldsymbol{x}_0$ 处的微分 (是一个线性映射) 对 $\mathbb{R}^n$ 中的元素 $\boldsymbol{x}$ 的作用, 相当于用对应的 Jacobi 矩阵乘以向量 $\boldsymbol{x}$, 即

$$\begin{aligned}\mathscr{A}:\mathbb{R}^n&\longrightarrow\mathbb{R}^m\\ \boldsymbol{x}&\longmapsto \mathrm{D}f(\boldsymbol{x}_0)\boldsymbol{x}\end{aligned}$$

这里我们用 $\mathscr{A}$ 表示线性映射, 用 $\mathrm{D}f(\boldsymbol{x}_0)$ 表示 Jacobi 矩阵以示区别. 换言之, 向量值函数的微分具体表现为 Jacobi 矩阵, 这一事实非常重要.

定理 (有限增量估计) : 设 f 为开集 G 上的可微向量值函数, 则对任意的 $[\boldsymbol{a},\boldsymbol{b}]\subseteq G$, 存在 $\boldsymbol{c}\in(\boldsymbol{a},\boldsymbol{b})$, 使得

$$\|f(\boldsymbol{b})-f(\boldsymbol{a})\|\leqslant\|\mathrm{D}f(\boldsymbol{c})\|\,\|\boldsymbol{b}-\boldsymbol{a}\|\,.$$

注: 向量值函数没有一般函数的有限增量公式, 只有上面的有限增量估计.

3.3 向量值函数对部分变元的偏微分

我们把 $\mathbb{R}^{n+p}$ 看做乘积

$$\mathbb{R}^{n+p}=\mathbb{R}^n\times\mathbb{R}^p,$$

又把 $\mathbb{R}^{n+p}$ 中的点表示为

$$(\boldsymbol{x},\boldsymbol{y})=(x_1,\cdots,x_n,y_1,\cdots,y_p).$$

定义: 设开集 $\Omega\subseteq\mathbb{R}^{n+p}$, 向量值函数 $F:\Omega\to\mathbb{R}^m$, $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in\Omega$. 若存在线性映射 $\mathscr{A}_1:\mathbb{R}^n\to\mathbb{R}^m$, 使得

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\|F(\boldsymbol{x}_0+\boldsymbol{h},\boldsymbol{y}_0)-F(\boldsymbol{x}_0,\boldsymbol{y}_0)-\mathscr{A}_1\boldsymbol{h}\|}{\|\boldsymbol{h}\|}=0,$$

则称 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 处对变元 $\boldsymbol{x}$ 可微, 且称 $\mathscr{A}_1$ 为 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 处对变元 $\boldsymbol{x}$ 的偏微分, 记作

$$\mathrm{D}_{\boldsymbol{x}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\mathscr{A}_1.$$

定义: 设开集 $\Omega\subseteq\mathbb{R}^{n+p}$, 向量值函数 $F:\Omega\to\mathbb{R}^m$, $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in\Omega$. 若存在线性映射 $\mathscr{A}_2:\mathbb{R}^p\to\mathbb{R}^m$, 使得

$$\lim_{\boldsymbol{k}\to\boldsymbol{0}}\frac{\|F(\boldsymbol{x}_0,\boldsymbol{y}_0+\boldsymbol{k})-F(\boldsymbol{x}_0,\boldsymbol{y}_0)-\mathscr{A}_2\boldsymbol{k}\|}{\|\boldsymbol{k}\|}=0,$$

则称 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 处对变元 $\boldsymbol{y}$ 可微, 且称 $\mathscr{A}_2$ 为 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 处对变元 $\boldsymbol{y}$ 的偏微分, 记作

$$\mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\mathscr{A}_2.$$

定理: 设开集 $\Omega\subseteq\mathbb{R}^{n+p}$, 向量值函数 $F:\Omega\to\mathbb{R}^m$, $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in\Omega$. 若 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 可微, 则 F 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 对变元 $\boldsymbol{x}$ 与变元 $\boldsymbol{y}$ 的偏微分均存在, 且

$$\begin{aligned}\mathrm{D}_{\boldsymbol{x}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)\boldsymbol{h}&=\mathrm{D}F(\boldsymbol{x}_0,\boldsymbol{y}_0)(\boldsymbol{h},\boldsymbol{0}),\qquad\forall\,\boldsymbol{h}\in\mathbb{R}^n.\\ \mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)\boldsymbol{k}&=\mathrm{D}F(\boldsymbol{x}_0,\boldsymbol{y}_0)(\boldsymbol{0},\boldsymbol{k}),\qquad\forall\,\boldsymbol{k}\in\mathbb{R}^p.\end{aligned}$$

思路 定义线性映射 $\mathscr{A}_1\boldsymbol{h}=\mathscr{A}(\boldsymbol{h},\boldsymbol{0})$, $\mathscr{A}_2\boldsymbol{k}=\mathscr{A}(\boldsymbol{0},\boldsymbol{k})$, 直接验证偏微分定义即可. 例如

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\|F(\boldsymbol{x}_0+\boldsymbol{h},\boldsymbol{y}_0)-F(\boldsymbol{x}_0,\boldsymbol{y}_0)-\mathscr{A}_1\boldsymbol{h}\|}{\|\boldsymbol{h}\|}=\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\|F((\boldsymbol{x}_0,\boldsymbol{y}_0)+(\boldsymbol{h},\boldsymbol{0}))-F(\boldsymbol{x}_0,\boldsymbol{y}_0)-\mathscr{A}(\boldsymbol{h},\boldsymbol{0})\|}{\|(\boldsymbol{h},\boldsymbol{0})\|}=0.$$

注: 上述定理说明当 $F:\Omega(\subseteq\mathbb{R}^{n+p})\to\mathbb{R}^m$ 在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 可微时,

$$\mathrm{D}F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\left(\begin{array}{ccc:ccc}\dfrac{\partial F_1}{\partial x_1}&\cdots&\dfrac{\partial F_1}{\partial x_n}&\dfrac{\partial F_1}{\partial y_1}&\cdots&\dfrac{\partial F_1}{\partial y_p}\\ \vdots&&\vdots&\vdots&&\vdots\\ \dfrac{\partial F_m}{\partial x_1}&\cdots&\dfrac{\partial F_m}{\partial x_n}&\dfrac{\partial F_m}{\partial y_1}&\cdots&\dfrac{\partial F_m}{\partial y_p}\end{array}\right),$$

也即

$$\mathrm{D}F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\left(\begin{array}{c:c}\mathrm{D}_{\boldsymbol{x}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)&\mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)\end{array}\right).$$

即是说, 偏微分 $\mathrm{D}_{\boldsymbol{x}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 与 $\mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 的矩阵就是微分 $\mathrm{D}F(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 的矩阵的两个子块.

3.4 再论隐函数定理

我们将 §2.7 中的最后一个隐函数定理改写为向量值函数的形式.

定理: 设开集 $\Omega\subseteq\mathbb{R}^{m+p}$, 向量值函数 $F:\Omega\to\mathbb{R}^p$ 连续可微, $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in\Omega$, 且

$$\begin{gathered}F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\boldsymbol{0},\\ \det(\mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}_0,\boldsymbol{y}_0))\neq0,\end{gathered}$$

则方程

$$F(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0}$$

在 $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ 的某邻域内唯一确定了连续可微函数

$$\boldsymbol{y}=f(\boldsymbol{x}),$$

使得
$$F(\boldsymbol{x}, f(\boldsymbol{x})) = \boldsymbol{0},$$
且其微分为
$$\mathrm{D}f(\boldsymbol{x}) = -[\mathrm{D}_{\boldsymbol{y}}F(\boldsymbol{x}, f(\boldsymbol{x}))]^{-1}\mathrm{D}_{\boldsymbol{x}}F(\boldsymbol{x}, f(\boldsymbol{x})).$$
用 Jacobi 矩阵来表示就是
$$\begin{pmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial y_p}{\partial x_1} & \cdots & \dfrac{\partial y_p}{\partial x_m} \end{pmatrix} = -\begin{pmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_p} \\ \vdots & & \vdots \\ \dfrac{\partial F_p}{\partial y_1} & \cdots & \dfrac{\partial F_p}{\partial y_p} \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial F_p}{\partial x_1} & \cdots & \dfrac{\partial F_p}{\partial x_m} \end{pmatrix}.$$

3.5 逆映射定理、反函数

定理 (逆映射定理) : 设 $\Omega \subseteq \mathbb{R}^m$ 为开集, 映射 $f : \Omega \to \mathbb{R}^m$ 连续可微, $a \in \Omega$, 且
$$\det \mathrm{D}f(a) \neq 0,$$
则存在点 a 的一个邻域 U, 使得 $f|_U$ 存在逆映射 g(即 $f|_U$ 为双射), 且逆映射 g 也连续可微. 此外,
$$\mathrm{D}g(y) = (\mathrm{D}f(x))^{-1}.$$

证明 令 $b = f(a)$. 构造映射
$$\begin{aligned} F : \Omega \times \mathbb{R}^m &\longrightarrow \quad \mathbb{R}^m \\ (x, y) &\longmapsto f(x) - y \end{aligned}$$
易知 $\mathrm{D}_x F(x, y) = \mathrm{D}f(x)$, $\mathrm{D}_y F(x, y) = -E_m$. 由于
$$F(a, b) = f(a) - b = f(a) - f(a) = 0,$$
$$\det \mathrm{D}_x F(a, b) = \det \mathrm{D}f(a) \neq 0.$$
因此由隐函数定理可知, 存在 a 的某邻域 U, 在其内由 $F(x, y) = 0$ 唯一确定了连续可微的函数
$$x = g(y),$$
且
$$\mathrm{D}g(y) = -(\mathrm{D}_x F(x, y))^{-1}\mathrm{D}_y F(x, y) = (\mathrm{D}f(x))^{-1},$$

即
$$\mathrm{D}g(y) = (\mathrm{D}f(x))^{-1}.$$

注: 从证明的最后我们可以看出互逆映射 (互为反函数) 的 Jacobi 行列式互为倒数, 即
$$\frac{\partial(x_1,\cdots,x_m)}{\partial(y_1,\cdots,y_m)} \cdot \frac{\partial(y_1,\cdots,y_m)}{\partial(x_1,\cdots,x_m)} = 1.$$
这是一元函数的反函数求导公式的推广.

注: 我们解释一下为何 $\mathrm{D}_x F(x,y) = \mathrm{D}f(x)$, $\mathrm{D}_y F(x,y) = -E_m$. 将 $F(x,y) = f(x) - y$ 写成分量的形式即为
$$\begin{pmatrix} F_1(x,y) \\ \vdots \\ F_m(x,y) \end{pmatrix} = \begin{pmatrix} f_1(x) - y_1 \\ \vdots \\ f_m(x) - y_m \end{pmatrix},$$
因此 $\dfrac{\partial F_i}{\partial x_j} = \dfrac{\partial f_i}{\partial x_j}$, $\dfrac{\partial F_i}{\partial y_j} = -\delta_{ij}$, 故
$$\mathrm{D}_x F(x,y) = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_m} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_m} \end{pmatrix} = \mathrm{D}f(x),$$
$$\mathrm{D}_y F(x,y) = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix} = -E_m.$$

例 极坐标变换
$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta \end{cases}$$
的 Jacobi 行列式
$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$
因此在任意点 $(x,y) \neq (0,0)$ 附近存在逆变换 $r = r(x,y), \theta = \theta(x,y)$.

定理: 设 $\Omega \subseteq \mathbb{R}^m$ 为开集, 映射 $f : \Omega \to \mathbb{R}^m$, 若对任意的 $x \in \Omega$, 有
$$\det \mathrm{D}f(x) \neq 0,$$
则 f 为开映射.

4 多元微分学的几何应用

4.1 曲线的切线

设有连续可微曲线 $\boldsymbol{r}(t)=(x(t),y(t),z(t))$, 且 $\boldsymbol{r}'(t)\neq\boldsymbol{0}$. 设 $\boldsymbol{r}(t_0)=P_0=(x_0,y_0,z_0)$ 是曲线上的定点, $\boldsymbol{r}(t)=P$ 是动点. 我们考察沿着割线 P_0P 方向的向量

$$\frac{\boldsymbol{r}(t)-\boldsymbol{r}(t_0)}{t-t_0}.$$

当 $t\to t_0$ 时, 割线 P_0P 的位置应该是曲线在 P_0 处的切线, 这样, 我们求得切向量

$$\boldsymbol{r}'(t_0)=\lim_{t\to t_0}\frac{\boldsymbol{r}(t)-\boldsymbol{r}(t_0)}{t-t_0}.$$

于是, 曲线 $\boldsymbol{r}(t)$ 在 P_0 处的切线方程就是

$$\frac{x-x_0}{x'(t_0)}=\frac{y-y_0}{y'(t_0)}=\frac{z-z_0}{z'(t_0)}.$$

注: 上式是由解析几何的知识得出的. 空间直线由直线上一点与方向数所确定. 这里直线上一点为 $P_0=(x_0,y_0,z_0)$, 方向数为切向量 $\boldsymbol{r}'(t_0)=(x'(t_0),y'(t_0),z'(t_0))$.

再看由隐式给出的曲线

$$\begin{cases} F(x,y,z)=0, \\ G(x,y,z)=0. \end{cases} \tag{1}$$

这里 F,G 连续可微, 且

$$\operatorname{rank}\begin{bmatrix} F_x & F_y & F_z \\ G_x & G_y & G_z \end{bmatrix}=2.$$

由隐函数定理, 我们当然可以解出其中某两个变元作为第三个变元的函数. 这样以后, 把曲线写成显式形式, 再利用前面的结果就能写出切线方程. 但是这里我们不这样做, 而是使用一种更具启发性的做法.

设 $\boldsymbol{r}(t)=(x(t),y(t),z(t))$ 是方程组 (1) 的解, $\boldsymbol{r}(t_0)=P_0=(x_0,y_0,z_0)$ 是该曲线上一点. 将 $\boldsymbol{r}(t)$ 带入方程组 (1) 得

$$\begin{cases} F(x(t),y(t),z(t))=0, \\ G(x(t),y(t),z(t))=0. \end{cases} \tag{2}$$

上式两边对 t 求导得

$$\begin{cases} F_xx'(t)+F_yy'(t)+F_zz'(t)=0, \\ G_xx'(t)+G_yy'(t)+G_zz'(t)=0. \end{cases}$$

从而

$$\begin{cases} F_x(P_0)x'(t_0) + F_y(P_0)y'(t_0) + F_z(P_0)z'(t_0) = 0, \\ G_x(P_0)x'(t_0) + G_y(P_0)y'(t_0) + G_z(P_0)z'(t_0) = 0. \end{cases}$$

也即

$$\begin{cases} \nabla F(P_0) \cdot \boldsymbol{r}'(t_0) = 0, \\ \nabla G(P_0) \cdot \boldsymbol{r}'(t_0) = 0. \end{cases}$$

即是说, 曲线 $\boldsymbol{r}(t)$ 在 P_0 处的切向量 $\boldsymbol{r}'(t_0)$ 与两向量 $\nabla F(P_0)$ 和 $\nabla G(P_0)$ 均正交. 因此, 切向量 $\boldsymbol{r}'(t_0)$ 平行于

$$\nabla F(P_0) \times \nabla G(P_0) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ F_x & F_y & F_z \\ G_x & G_y & G_z \end{vmatrix}_{P_0} = \left(\frac{\partial(F,G)}{\partial(y,z)}(P_0), \frac{\partial(F,G)}{\partial(z,x)}(P_0), \frac{\partial(F,G)}{\partial(x,y)}(P_0) \right).$$

因此曲线 $\boldsymbol{r}(t)$ 在 $P_0 = \boldsymbol{r}(t_0)$ 处的切线方程为

$$\frac{x - x_0}{\frac{\partial(F,G)}{\partial(y,z)}(P_0)} = \frac{y - y_0}{\frac{\partial(F,G)}{\partial(z,x)}(P_0)} = \frac{z - z_0}{\frac{\partial(F,G)}{\partial(x,y)}(P_0)}.$$

4.2 曲面的切平面与法线

设有隐式表示的曲面

$$F(x, y, z) = 0.$$

这里 F 连续可微, 且 $\nabla F \neq \boldsymbol{0}$. 设 $\boldsymbol{r}(t) = (x(t), y(t), z(t))$ 是该曲面上过点 $P_0 = (x_0, y_0, z_0)$ 的任意一条连续可微曲线, 且 $\boldsymbol{r}(t_0) = P_0$, 则

$$F(x(t), y(t), z(t)) \equiv 0.$$

两边对 t 求导得

$$F_x x'(t) + F_y y'(t) + F_z z'(t) = 0.$$

从而

$$F_x(P_0)x'(t_0) + F_y(P_0)y'(t_0) + F_z(P_0)z'(t_0) = 0$$

注意到 $(x'(t_0), y'(t_0), z'(t_0))$ 是曲线 $\boldsymbol{r}(t)$ 在 P_0 处的切向量, 由曲线 $\boldsymbol{r}(t)$ 的任意性知 $\nabla F(P_0)$ 是曲面在 P_0 处的法向量. 进而可得曲面在 P_0 处的切平面为

$$\nabla F(P_0) \cdot (\boldsymbol{r} - \boldsymbol{r}_0) = 0,$$

也即

$$F_x(P_0)(x-x_0)+F_y(P_0)(y-y_0)+F_z(P_0)(z-z_0)=0.$$

5 勒贝格测度

在前面我们学习了多元微积分中的微分理论. 现在, 我们自然要考虑多元积分的相关问题. 我们想要回答的一般问题是: 给定 $\mathbb{R}^n$ 的某个子集 Ω 与函数 $f:\Omega\to\mathbb{R}$, 能否求出 f 在 Ω 上的积分, 从而得到某个数 $\int_\Omega f$? (也可以考虑其他类型的函数, 比如复值函数与向量值函数. 但是只要我们弄清楚如何对实值函数求积分, 对上述其他类型的函数求积分就不再困难, 因为对复值函数与向量值函数求积分可以通过分别求它们每个实值分量的积分来实现.)

对于一维的情形, 我们已经建立了黎曼积分 $\int_{[a,b]} f$ 的概念. 这就回答了当 $\Omega=[a,b]$ 时的问题, 此时 f 是黎曼可积的. 然而, 并非所有函数都是黎曼可积的. 黎曼积分的概念可以推广到更高维的情形, 但是我们只能对那些 "黎曼可积的" 函数求积分, 而这样的函数并不多. (例如, 黎曼可积函数序列的一致极限仍然是黎曼可积的, 但黎曼可积函数序列的逐点极限不一定是黎曼可积的, L^2 极限也不一定是黎曼可积的.)

因此, 我们必须在黎曼积分之外寻求一种真正令人满意的积分概念, 这个概念甚至可以处理间断性非常强的函数. 这就导出了本章与下一章的研究对象——勒贝格积分. 勒贝格积分可以处理很大一类函数, 其中包括所有的黎曼可积函数与其他某些函数. 事实上, 我们可以放心地说, 勒贝格积分本质上能求数学中任何实际需要的函数的积分, 至少能对欧几里得空间中所有绝对可积的函数求积分. (如果使用选择公理, 那么我们还能构造出某些病态函数, 这些函数无法用勒贝格积分来处理, 但它们并不会出现在实际应用中.)

为了知道如何计算积分 $\int_\Omega f$, 我们必须先弄清楚一个更基础、更根本的问题: 应该如何计算 Ω 的长度、面积或者体积 (统一称为测度)? (这是因为若能求出函数 1 在 Ω 上的积分, 那么我们就得到了 Ω 的测度.)

最理想的情况是, 对 $\mathbb{R}^n$ 的每一个子集 Ω, 我们都能指派一个非负数 $m(\Omega)$, 它就是 Ω 的测度. $m(\Omega)$ 可以等于 0(例如单点集或空集), 也可以等于无穷大 (例如整个 $\mathbb{R}^n$). 这个测度应当满足某些合理的特点性质. 例如, 单位立方体 $[0,1]^n$ 的测度应该等于 1. 如果 A 和 B 不相交, 那么应该有 $m(A\cup B)=m(A)+m(B)$. 只要 $A\subseteq B$, 就有 $m(A)\leqslant m(B)$. 对任意的 $x\in\mathbb{R}^n$, 都有 $m(x+A)=m(A)$(也就是说, 集合 A 平移了向量 x 之后, 测度保持不变).

令人惋惜的是, 这样的度量其实根本就不存在. 但是, 我们可以采取一些补救措施, 那就是只测量 $\mathbb{R}^n$ 中一个特定类型的集合, 即可测集. 我们只在这些可测集 Ω 上定义测度 $m(\Omega)$. 一旦我们把注意力集中在可测集上, 那么上述所有性质就都满足了. 另外, 我们在实际生活中遇到的几乎所有的集合都是可测集 (例如所有的开集和闭集都是可测的), 而这对分析学研究已经足够好了.

5.1 目标: 勒贝格测度

设 $\mathbb{R}^n$ 为一个欧几里得空间, 我们希望定义 $\mathbb{R}^n$ 的一类特殊子集 (称为可测集), 对每一个这样的集合 $\Omega \subseteq \mathbb{R}^n$, 都有 $[0, +\infty]$ 中的一个数字 $m(\Omega)$ 与其对应. 可测集满足下列性质:

(1) (博雷尔性质) $\mathbb{R}^n$ 中的每一个开集都是可测集, 每一个闭集也都是可测集.

(2) (互补性) 若 $\Omega \subseteq \mathbb{R}^n$ 为可测集, 则 $\mathbb{R}^n \setminus \Omega$ 也为可测集.

(3) (布尔代数性质) 若 $\{\Omega_j\}_1^n$ 为有限个可测集, 则它们的并集 $\bigcup\limits_{j=1}^{n} \Omega_j$ 与交集 $\bigcap\limits_{j=1}^{n} \Omega_j$ 都是可测集.

(4) (σ 代数性质) 若 $\{\Omega_j\}_1^\infty$ 为可数个可测集, 则它们的并集 $\bigcup\limits_{j=1}^{\infty} \Omega_j$ 与交集 $\bigcap\limits_{j=1}^{\infty} \Omega_j$ 都是可测集.

注意其中有一些性质是多余的. 例如 σ 代数性质可以推出布尔代数性质, 且若每个开集均可测, 则根据互补性可知每个闭集也均可测.

对于 $\mathbb{R}^n$ 中的每一个可测集 Ω, 我们都指派了一个勒贝格测度 $m(\Omega)$, 它满足如下性质:

(5) (空集) 空集的测度为 $m(\varnothing) = 0$.

(6) (非负性) 对任意的可测集 Ω, 都有 $0 \leqslant m(\Omega) \leqslant +\infty$.

(7) (单调性) 若 $A \subseteq B$, 且 A 与 B 均为可测集, 则 $m(A) \leqslant m(B)$.

(8) (有限次可加性) 若 $\{A_j\}_1^n$ 为有限个可测集, 则 $m\left(\bigcup\limits_{j=1}^{n} A_j\right) \leqslant \sum\limits_{j=1}^{n} m(A_j)$.

(9) (有限可加性) 若 $\{A_j\}_1^n$ 为有限个互不相交的可测集, 则 $m\left(\bigcup\limits_{j=1}^{n} A_j\right) = \sum\limits_{j=1}^{n} m(A_j)$.

(10) (可数次可加性) 若 $\{A_j\}_1^\infty$ 为可数个可测集, 则 $m\left(\bigcup\limits_{j=1}^{\infty} A_j\right) \leqslant \sum\limits_{j=1}^{\infty} m(A_j)$.

(11) (可数可加性) 若 $\{A_j\}_1^\infty$ 为可数个互不相交的可测集, 则 $m\left(\bigcup\limits_{j=1}^{\infty} A_j\right) = \sum\limits_{j=1}^{\infty} m(A_j)$.

(12) (规范性) 对单位立方体 $[0,1]^n \subseteq \mathbb{R}^n$, 其测度 $m([0,1]^n) = 1$.

(13) (平移不变性) 若 Ω 为可测集, $x \in \mathbb{R}^n$, 则 $x + \Omega = \{x + y : y \in \Omega\}$ 也为可测集, 且 $m(x + \Omega)$ $= m(\Omega)$.

同样, 在这些性质中也存在一些多余的内容. 例如可数可加性可以推出有限可加性, 有限可加性与非负性又可以推出单调性. 此外, 我们还可以利用可加性推出次可加性.

于是, 本章的目标就可以叙述如下:

定理 (勒贝格测度的存在性) : 存在可测集的概念, 同时还存在一种方法, 使得 $\mathbb{R}^n$ 中的每一个可测集 Ω 都能被指派一个数字 $m(\Omega)$, 且保证性质 (1) ~ (13) 全都成立.

事实上, 勒贝格测度是唯一的. 任何满足公理 (1) ~ (13) 的可测性与测度的概念, 都将极大地与勒贝格测度重合. 当然, 有一些测度只满足上述部分公理. 此外, 我们还可能对欧几里得空间 $\mathbb{R}^n$ 之外的其他区域上的测度感兴趣, 这就引出了测度论, 它自身是一个完整的课题, 对此我们不做讨论. 但要说明一点, 在现代概率论和分析学更深入的研究 (例如广义函数论) 中, 测度的概念非常重要.

5.2 第一步: 外测度

我们尝试用一些盒子来覆盖集合, 这种方法提出了外测度这一概念. 外测度适用于每一个集合, 并且满足除了有限可加性 (9) 与可数可加性 (11) 之外的所有性质 (5) ~ (13). 之后我们会对外测度稍做修改, 从而使其能够满足可加性.

定义 (开盒子) : $\mathbb{R}^n$ 中的集合

$$B = \prod_{i=1}^{n}(a_i, b_i) = \{(x_1, \cdots, x_n) \in \mathbb{R}^n : x_i \in (a_i, b_i), i = 1, \cdots, n\}$$

称为 $\mathbb{R}^n$ 中的一个**开盒子**, 简称盒子, 其中 $b_i \geqslant a_i$. 定义盒子 B 的**体积**为

$$\mathrm{vol}(B) = \prod_{i=1}^{n}(b_i - a_i) = (b_1 - a_1)\cdots(b_n - a_n).$$

注: 后面我们会发现这里定义的开盒子的体积 $\mathrm{vol}(B)$ 就等于其测度 $m(B)$.

例如 $(0,1)^n$ 就是一个盒子, 其体积为 1. 当 $n = 1$ 时, 开盒子与开区间是一样的. 当有某 $a_i = b_i$ 时, 盒子就是体积为 0 的空集, 我们仍然把它看做一个盒子. 有时为了强调我们处理的是 n 维体积, 我们把 $\mathrm{vol}(B)$ 写为 $\mathrm{vol}_n(B)$.

定义 (开盒覆盖) : 设 Ω 为 $\mathbb{R}^n$ 的子集, $\{B_j\}_{j\in J}$ 为一族开盒子. 若

$$\Omega \subseteq \bigcup_{j\in J} B_j,$$

则称开盒子族 $\{B_j\}_{j\in J}$ 覆盖了 Ω.

假设 Ω 被至多可数个开盒子 $\{B_j\}_{j\in J}$ 覆盖, 若我们希望 Ω 是可测的, 其测度满足单调性与次可加性, 且对每个 j 都有 $m(B_j) = \mathrm{vol}(B_j)$, 则

$$m(\Omega) \leqslant m\left(\bigcup_{j\in J} B_j\right) \leqslant \sum_{j\in J} m(B_j) = \sum_{j\in J} \mathrm{vol}(B_j).$$

从而

$$m(\Omega) \leqslant \inf\left\{\sum_{j\in J} \mathrm{vol}(B_j) : \Omega \subseteq \bigcup_{j\in J} B_j, J\text{至多可数}\right\}.$$

受此启发, 我们定义

定义 (外测度) : 设 Ω 为 $\mathbb{R}^n$ 的子集, 我们称

$$m^*(\Omega) = \inf\left\{\sum_{j\in J} \mathrm{vol}(B_j) : \Omega \subseteq \bigcup_{j\in J} B_j, J\text{至多可数}\right\}$$

为集合 Ω 的**外测度**.

因为 $\sum_{j\in J}\mathrm{vol}(B_j)$ 是非负的, 所以 $m^*(\Omega)\geqslant 0$. 有时为了强调使用的是 n 维外测度, 我们把 $m^*(\Omega)$ 写为 $m_n^*(\Omega)$. 由确界原理知对 $\mathbb{R}^n$ 中的任意集合 Ω, 外测度 $m^*(\Omega)$ 都存在, 因此对每一个集合都可以定义外测度的概念.

引理 (外测度的性质)：外测度满足如下的性质:

① (空集) 空集的测度为 $m^*(\varnothing)=0$.

② (非负性) 对任意的可测集 Ω, 都有 $0\leqslant m^*(\Omega)\leqslant +\infty$.

③ (单调性) 若 $A\subseteq B$, 且 A 与 B 均为可测集, 则 $m^*(A)\leqslant m^*(B)$.

④ (有限次可加性) 若 $\{A_j\}_1^n$ 为有限个可测集, 则 $m^*\left(\bigcup_{j=1}^n A_j\right)\leqslant \sum_{j=1}^n m^*(A_j)$.

⑤ (可数次可加性) 若 $\{A_j\}_1^\infty$ 为可数个可测集, 则 $m^*\left(\bigcup_{j=1}^\infty A_j\right)\leqslant \sum_{j=1}^\infty m^*(A_j)$.

⑥ (平移不变性) 若 Ω 为可测集, $x\in\mathbb{R}^n$, 则 $x+\Omega=\{x+y: y\in\Omega\}$ 也为可测集, 且 $m^*(x+\Omega)=m^*(\Omega)$.

闭盒子的外测度也符合我们的期望.

引理 (闭盒子的外测度)：闭盒子

$$B=\prod_{i=1}^n[a_i,b_i]=\{(x_1,\cdots,x_n)\in\mathbb{R}^n: x_i\in[a_i,b_i], i=1,\cdots,n\}$$

的外测度

$$m^*(B)=\prod_{i=1}^n(b_i-a_i).$$

推论: 外测度满足性质: ⑦ (规范性) 对单位立方体 $[0,1]^n\subseteq\mathbb{R}^n$, 其外测度 $m^*([0,1]^n)=1$.

至此, 外测度已经满足了除有限可加性与可数可加性之外的测度的所有性质.

一旦有了闭盒子的外测度, 关于开盒子的相应结论也就容易了.

推论 (开盒子的外测度)：开盒子

$$B=\prod_{i=1}^n(a_i,b_i)=\{(x_1,\cdots,x_n)\in\mathbb{R}^n: x_i\in(a_i,b_i), i=1,\cdots,n\}$$

的外测度

$$m^*(B)=\prod_{i=1}^n(b_i-a_i).$$

证明　不妨设 $b_i>a_i$, 因为若某 $b_i=a_i$, 则结论变为 $m^*(\varnothing)=0$, 前面已证. 取充分小的 $\varepsilon>0$, 使得 $b_i-\varepsilon>a_i+\varepsilon$, 从而

$$\prod_{i=1}^n[a_i+\varepsilon,b_i-\varepsilon]\subseteq\prod_{i=1}^n(a_i,b_i)\subseteq\prod_{i=1}^n[a_i,b_i].$$

由外测度的单调性与闭盒子的外测度, 令 $\varepsilon\to 0$ 即可证得结论.

推论: 开盒子的外测度 $m^*(B)$ 等于体积 $\mathrm{vol}(B)$.

例 考虑实直线 $\mathbb{R}$ 上的外测度.

(1) 由于 $m^*(\mathbb{R}) \geqslant m^*([-r,r]) = 2r$. 令 $r \to +\infty$, 得 $m^*(\mathbb{R}) = +\infty$.

(2) 根据闭盒子 $B = \prod_{i=1}^{n}[a_i, b_i]$ 的外测度 $m^*(B) = \prod_{i=1}^{n}(b_i - a_i)$ 可知单点集 $\{q\}$ 的外测度为 $m^*(\{q\}) = 0$, 注意到 $\mathbb{Q}$ 是可数的, 从而由次可数可加性得

$$m^*(\mathbb{Q}) = m^*\left(\bigcup_{q\in\mathbb{Q}}\{q\}\right) \leqslant \sum_{q\in\mathbb{Q}} m^*(\{q\}) = 0.$$

于是 $m^*(\mathbb{Q}) = 0$. 同理可得每一个可数集的外测度均为 0.

(3) 根据有限次可加性可得

$$m^*(\mathbb{R}) \leqslant m^*(\mathbb{R}\setminus\mathbb{Q}) + m^*(\mathbb{Q}).$$

因为 $m^*(\mathbb{Q}) = 0$, 故 $m^*(\mathbb{R}\setminus\mathbb{Q}) \geqslant m^*(\mathbb{R}) = +\infty$. 类似可得 $[0,1]\setminus\mathbb{Q}$ 的外测度为 1.

例 $\mathbb{R}$ 中的闭盒子 $[0,1]$ 的一维外测度为 1, 但 $\mathbb{R}^2$ 中的闭盒子 $\{(x,0)\in\mathbb{R}^2 : x\in[0,1]\}$ 的外测度为 0. 因此, 一维外测度与二维外测度是有很大区别的. 进一步, 由次可数可加性得 $\mathbb{R}^2$ 中整个 x 轴的外测度为 0.

5.3 外测度是不可加的

由上一节我们看到, 只要外测度满足有限可加性与可数可加性, 那么我们就拥有了一个可用的测度. 但遗憾的是, 这些性质对外测度是不成立的, 即使是一维外测度.

命题 (可数可加性不成立) : 在 $\mathbb{R}$ 中存在可数个互不相交的集合 $\{A_j\}_1^\infty$, 使得

$$m^*\left(\bigcup_{j=1}^{\infty} A_j\right) \neq \sum_{j=1}^{\infty} m^*(A_j).$$

注: 上面的证明用到了选择公理, 这一点是非常必要的.

通过改进上述论证, 我们可以证明外测度也不满足有限可加性.

命题 (有限可加性不成立) : 在 $\mathbb{R}$ 中存在有限个互不相交的集合 $\{A_j\}_1^n$, 使得

$$m^*\left(\bigcup_{j=1}^{n} A_j\right) \neq \sum_{j=1}^{n} m^*(A_j).$$

注: 这些例子与巴拿赫-塔斯基悖论有关. 该悖论说的是我们可以把 $\mathbb{R}^3$ 中的单位球划分为有限多块, 在经过旋转和平移后, 这有限多块能够重新组装成两个完整的单位球! 当然, 这个划分涉及了不可测集.

5.4 可测集

在上一节中, 我们看到一些集合的外测度性状比较糟糕, 特别是它们可以作为有限可加性与可数可加性的反例. 不过, 这些集合是相当病态的, 利用选择公理, 它们被人为地构造出来. 我们希望把这些集合排除在外, 从而使得有限可加性与可数可加性能够成立. 非常幸运, 这是能够实现的. 这归功于一个聪明的定义, 它由康斯坦丁 · 卡拉西奥多里提出.

定义 (勒贝格可测性) : 设 E 为 $\mathbb{R}^n$ 的子集, 若对 $\mathbb{R}^n$ 的任意子集 A, 均有

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c),$$

则称 E 是**勒贝格可测**的, 简称可测的, 这里 $E^c = \mathbb{R}^n \setminus E$ 为 E 在 $\mathbb{R}^n$ 中的补集. 若 E 是可测的, 则我们把 E 的**勒贝格测度**定义为 $m(E) = m^*(E)$. 若 E 不可测, 则 $m(E)$ 无定义.

上述定义即是说, 当我们用 E 把任意的集合 A 划分为两部分时, 可加性都成立. 因此, 我们可以把可测集看做是能使有限可加性成立的集合. 有时我们把 $m(E)$ 写成 $m_n(E)$, 以此强调使用的是 n 维勒贝格测度.

上面的定义使用起来有些困难, 我们很难由这个定义直接验证一个集合是否可测. 但是, 我们将使用这一定义验证可测集的一些有用的性质, 然后利用这些性质来判断可测性, 而不再使用上述定义.

引理一 (可测集的性质) :

(1) (互补性) 若 E 是可测的, 则 $\mathbb{R}^n \setminus E$ 也是可测的.

(2) (平移不变性) 若 E 是可测的, $x \in \mathbb{R}^n$, 则 $x + E$ 也是可测的, 且 $m(x+E) = m(E)$.

(3) (布尔代数性质) 若 $\{E_j\}_1^n$ 可测, 则 $\bigcup_{j=1}^{n} E_j$ 与则 $\bigcap_{j=1}^{n} E_j$ 也是可测的.

(4) 每个开盒子都是可测的, 每个闭盒子也都是可测的.

(5) 每一个外测度为 0 的集合 E 都是可测的.

引理二 (有限可加性) : 设 $\{E_j\}_1^n$ 为有限个互不相交的可测集, A 是任意一个集合 (不一定可测), 则

$$m^*\left(A \cap \bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} m^*(A \cap E_j).$$

因此

$$m\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} m(E_j).$$

推论: 若 $A \subseteq B$ 为两个可测集, 则 $B \setminus A$ 也是可测集. 若还有 $m(A) < +\infty$, 则

$$m(B \setminus A) = m(B) - m(A).$$

证明 $B = A \cup (B \setminus A)$, 且 A 与 $B \setminus A$ 互不相交, 则 $m(B) = m(A) + m(B \setminus A)$. 因 $m(A) < +\infty$,

移项即得结论.

引理 (可数可加性) : 设 $\{E_j\}_1^\infty$ 为可数个互不相交的可测集, 则 $\bigcup\limits_{j=1}^{\infty} E_j$ 也是可测集, 且

$$m\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} m(E_j).$$

引理 (σ 代数性质): 若 $\{E_j\}_1^\infty$ 为可数个可测集, 则它们的并集 $\bigcup\limits_{j=1}^{\infty} E_j$ 与交集 $\bigcap\limits_{j=1}^{\infty} E_j$ 也都是可测集.

我们所期望的最后一条性质是博雷尔性质, 为此需要以下的预备引理.

引理: $\mathbb{R}^n$ 中的任意开集都能写成有限或可数个开盒子的并集.

注: 实际上, 若限制开盒子互不相交, 则我们有结论:

(1) $\mathbb{R}$ 中的任意开集都能写成有限或可数个开区间的并集.

(2) $\mathbb{R}^n$ $(n \geqslant 2)$ 中的开集都能写成可数个互不相交的半开半闭立方体的并集.

引理 (博雷尔性质) : 每一个开集都是勒贝格可测的, 每一个闭集也都是勒贝格可测的.

证明 只需要证明关于开集的结论即可. 设 E 为开集, 则 E 可以写成至多可数个开盒子的并集, 而开盒子是可测的, 因此 E 是可测的. 由互补性可知任意闭集也是可测的.

现在, 关于勒贝格测度的构造及其基本性质就介绍完了, 接下来, 我们将进入构造勒贝格积分的下一步——介绍可积函数的类型.

5.5 可测函数

在黎曼积分的理论中, 我们只能对一类特定的函数——黎曼可积函数求积分. 现在我们可以对更大的一类函数——可测函数求积分. 更准确来说, 我们只能对绝对可积的可测函数求积分.

定义 (可测函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f : \Omega \to \mathbb{R}^m$. 若对任意的开集 $V \subseteq \mathbb{R}^m$, $f^{-1}(V)$ 是可测的, 则称 f 是**可测函数**.

就像之前讨论的那样, 我们在现实生活中处理的绝大多是集合都是可测的. 与之相似, 我们在现实生活中处理的大部分函数也是可测的.

注: 一句话概括可测函数的定义即为: 任意开集的原像都是可测的.

引理 (连续函数是可测的) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f : \Omega \to \mathbb{R}^m$. 若 f 是连续的, 则 f 是可测的.

证明 连续函数的等价条件是任意开集的原像是开集, 可测函数的等价条件为任意开集的原像是可测集. 由于开集必为可测集, 故连续函数必可测 (当然, 前提是定义域为可测集).

引理: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f : \Omega \to \mathbb{R}^m$, 则 f 是可测的当且仅当 f 的每个分量 f_i 是可测的.

遗憾的是, 两个可测函数的复合函数不一定是可测的. 不过, 连续函数作用在可测函数上仍是可

测的.

引理 (连续 $\circ$ 可测 $=$ 可测) : 设可测集 $\Omega \subseteq \mathbb{R}^n$, 开集 $W \subseteq \mathbb{R}^m$, 函数 $f: \Omega \to W$ 是可测的, $g: W \to \mathbb{R}^p$ 是连续的, 则 $g \circ f: \Omega \to \mathbb{R}^p$ 是可测的.

推论 (可测函数的各种运算仍为可测函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f: \Omega \to \mathbb{R}$ 与 $g: \Omega \to \mathbb{R}$ 均为可测函数, 则 $|f|, f \pm g, fg, f/g, \max\{f, g\}, \min\{f, g\}$ 均为可测函数. (当然, 对于 f/g, 要求在 Ω 上有 $g(x) \neq 0$.)

思路 ① 对 $|f|$, 考虑 $g(x) = |x|$, 则 $g \circ f = |f|$ 可测.

② 对 $f + g$, 考虑函数

$$\begin{aligned} h: \Omega &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (f(x), g(x)) \end{aligned}$$

则因为 h 的每个分量都是可测函数, 故 h 是可测函数. 又令

$$\begin{aligned} k: \quad \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (a, b) &\longmapsto a + b \end{aligned}$$

为连续函数, 则 $k \circ h = f + g$ 为可测函数.

③ 其余情况类似可证.

引理: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f: \Omega \to \mathbb{R}$, 则 f 是可测的当且仅当对任意的实数 a, $f^{-1}((a, +\infty))$ 都是可测的.

受此启发, 我们可以把可测函数的概念推广到广义实数系 $\mathbb{R}^* = \mathbb{R} \cup \{+\infty, -\infty\}$ 上.

定义 (广义实数系上的可测函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f: \Omega \to \mathbb{R}^*$. 若对任意的实数 a, $f^{-1}((a, +\infty])$ 都是可测的, 则称 f 是可测的.

引理 (可测函数序列的极限是可测的) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f_n: \Omega \to \mathbb{R}^*$ 均为可测函数 $(n = 1, 2, \cdots)$, 则函数 $\sup f_n, \inf f_n, \overline{\lim_{n\to\infty}} f_n = \lim_{n\to\infty} \sup_{k\geqslant n} f_k = \inf_n \sup_{k\geqslant n} f_k, \underline{\lim_{n\to\infty}} f_n = \lim_{n\to\infty} \inf_{k\geqslant n} f_k = \sup_n \inf_{k\geqslant n} f_k$ 也都是可测的. 特别地, 若 $\{f_n\}$ 逐点收敛于函数 f, 则 f 也是可测函数.

正如你所看到的, 我们对可测函数做的任何事情几乎都能构造出另一个可测函数, 这解释了为什么我们在数学中处理的每一个函数差不多都是可测函数. (实际上, 构造不可测函数的唯一方法就是人为地去构造, 比如使用选择公理.)

6 勒贝格积分

我们首先考虑一类简单的可测函数——简单函数, 证明如何对简单函数求积分, 进而再对所有的可测函数 (或者至少是绝对可积的函数) 求积分.

6.1 简单函数

定义 (简单函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}$ 为可测函数. 若 f 的像集 $f(\Omega)$ 为有限集, 则称 f 是一个**简单函数**. (即是说函数值只取有限个元素 $c_1,\cdots,c_k$.)

例 设 Ω 为 $\mathbb{R}^n$ 的可测子集, E 为 Ω 的可测子集, 特征函数 $\chi_E:\Omega\to\mathbb{R}$ 定义为

$$\chi_E(x)=\begin{cases}1, & x\in E,\\ 0, & x\in\Omega\setminus E.\end{cases}$$

易知 χ_E 是一个可测函数. (一般情况下, $\chi_E(\Omega)=\{0,1\}$. 当 $E=\varnothing$ 时, $\chi_E(\Omega)=\{0\}$. 当 $E=\Omega$ 时, $\chi_E(\Omega)=\{1\}$.)

我们给出简单函数的三个基本性质: 它们构成一个线性空间, 它们是特征函数的线性组合, 它们逼近可测函数. 具体为如下三个引理.

引理 (简单函数构成线性空间) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, 函数 $f:\Omega\to\mathbb{R}$ 与 $g:\Omega\to\mathbb{R}$ 均为简单函数, 则 $f+g$ 与 cf 均为简单函数, 其中 $c\in\mathbb{R}$.

引理 (简单函数是特征函数的线性组合) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}$ 为简单函数, 则存在有限个实数 $c_1,\cdots,c_k$ 与 Ω 中的有限个互不相交的可测集 $E_1,\cdots,E_k$, 使得

$$f=\sum_{i=1}^{k}c_i\chi_{E_i}.$$

引理 (非负简单函数逼近非负可测函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}$ 为非负可测函数, 则存在非负递增的简单函数序列 $\{f_n:\Omega\to\mathbb{R}\}$:

$$0\leqslant f_1(x)\leqslant\cdots\leqslant f_n(x)\leqslant\cdots,\quad x\in\Omega,$$

使得 $\{f_n\}$ 逐点收敛于 f, 即

$$\lim_{n\to\infty}f_n(x)=f(x),\quad x\in\Omega.$$

现在我们来说明如何计算简单函数的积分.

定义 (非负简单函数的勒贝格积分) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}$ 为非负简单函数, $f(\Omega)=\{c_1,\cdots,c_k\}$, 我们把函数 f 在 Ω 上的**勒贝格积分** $\int_\Omega f$ 定义为

$$\int_\Omega f=\sum_{i=1}^{k}c_i m(E_i).$$

其中 $E_i=\{x\in\Omega:f(x)=c_i\}$.

注: 若将简单函数 f 表示为特征函数的线性组合 $f = \sum_{i=1}^{k} c_i \chi_{E_i}$, 则

$$\int_\Omega f = \int_\Omega \sum_{i=1}^{k} c_i \chi_{E_i} = \sum_{i=1}^{k} c_i m(E_i).$$

注: 若 Ω' 为 Ω 的可测子集, 那么通过把 f 限制到 Ω', 我们可以定义 f 在 Ω' 上的勒贝格积分为

$$\int_{\Omega'} f = \int_{\Omega'} f|_{\Omega'} \left(= \int_\Omega f \chi_{\Omega'} \right).$$

有时我们也把 $\int_\Omega f$ 记作 $\int_\Omega f \mathrm{d}m$, 以此强调勒贝格测度 m 的作用. 或者使用一个像 x 那样的虚拟变量, 比如 $\int_\Omega f(x) \mathrm{d}x$.

例 函数 $f : \mathbb{R} \to \mathbb{R}$ 在区间 $[1,2]$ 上等于 3, 区间 $(2,4)$ 上等于 4, 其余处取值为 0, 则

$$\int_\mathbb{R} f = 3 \times m([1,2]) + 4 \times m((2,4)) = 3 \times 1 + 4 \times 2 = 11.$$

例 函数 $g : \mathbb{R} \to \mathbb{R}$ 在 $[0,+\infty)$ 上取值为 1, 其余处取值为 0, 则

$$\int_\mathbb{R} g = 1 \times m([0,+\infty)) = +\infty.$$

因此, 简单函数的积分可以等于无穷.

注: 我们只考虑非负函数的积分就是为了避免出现 $(+\infty) + (-\infty)$ 的无定义形式.

非负简单函数的积分有如下基本性质.

命题: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f : \Omega \to \mathbb{R}$ 与 $g : \Omega \to \mathbb{R}$ 均为非负简单函数, 则

(1) $0 \leqslant \int_\Omega f \leqslant +\infty$. $\int_\Omega f = 0$ 当且仅当 f 几乎处处为 0.

(2) $\int_\Omega (f+g) = \int_\Omega f + \int_\Omega g$.

(3) $\int_\Omega cf = c \int_\Omega f$, 其中 $c \geqslant 0$ 为常数.

(4) 若 $f \leqslant g$, 则 $\int_\Omega f \leqslant \int_\Omega g$.

(5) 若 A, B 为 Ω 的互不相交的可测子集, 则 $\int_{A \cup B} f = \int_A f + \int_B f$.

6.2 非负可测函数

现在我们从非负简单函数的积分过渡到非负可测函数的积分. 有时我们允许可测函数的取值为 $+\infty$.

定义 (从上/下方控制) : 设有非负函数 $f : \Omega \to \mathbb{R}$ 与 $g : \Omega \to \mathbb{R}$, 若

$$f(x) \geqslant g(x), \quad x \in \Omega,$$

则称 f 从上方控制 g, 或 g 从下方控制 f.

定义 (非负可测函数的勒贝格积分)：设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to[0,+\infty]$ 为非负可测函数, 我们定义 f 在 Ω 上的**勒贝格积分** $\displaystyle\int_\Omega f$ 为

$$\int_\Omega f=\sup\left\{\int_\Omega s : s\text{ 为非负简单函数, 且从下方控制 } f\right\}.$$

注: 读者应当将上述定义与下黎曼积分的概念进行比较. 有趣的是, 这里我们并不需要让这个下积分与上积分相等.

注: 若 Ω' 为 Ω 的可测子集, 那么通过把 f 限制到 Ω', 我们可以定义 f 在 Ω' 上的勒贝格积分为

$$\int_{\Omega'} f=\int_{\Omega'} f|_{\Omega'}\left(=\int_\Omega f\chi_{\Omega'}\right).$$

命题: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to[0,+\infty]$ 与 $g:\Omega\to[0,+\infty]$ 为非负可测函数, 则

(1) $0\leqslant\displaystyle\int_\Omega f\leqslant+\infty$. $\displaystyle\int_\Omega f=0$ 当且仅当 f 几乎处处为 0.

(2) $\displaystyle\int_\Omega cf=c\int_\Omega f$, 其中 $c\geqslant 0$.

(3) 若 $f\leqslant g$, 则 $\displaystyle\int_\Omega f\leqslant\int_\Omega g$.

(4) 若在 Ω 上几乎处处都有 $f(x)=g(x)$, 则 $\displaystyle\int_\Omega f=\int_\Omega g$.

(5) 若 Ω' 为 Ω 的可测子集, 则 $\displaystyle\int_{\Omega'} f=\int_\Omega f\chi_{\Omega'}\leqslant\int_\Omega f$.

(6) 若 A,B 为 Ω 的互不相交的可测子集, 则 $\displaystyle\int_{A\cup B} f=\int_A f+\int_B f$.

注: 上述命题中的 (4) 非常有趣, 它说明我们可以修改函数在任意测度为 0 的集合上的值, 而这不会对积分值产生任何影响. 只有正测度的点集才会对积分产生影响.

前面我们已经看到, 积分运算与极限运算并不总是可以交换次序的. 但是, 倘若函数序列是单调递增的, 那么勒贝格积分与极限运算的次序就可以交换.

定理 (勒贝格单调收敛定理)：设 Ω 为 $\mathbb{R}^n$ 的可测子集, $\{f_n:\Omega\to\mathbb{R}\}$ 为单调递增的非负可测函数序列, 即

$$0\leqslant f_1(x)\leqslant\cdots\leqslant f_n(x)\leqslant\cdots,\quad x\in\Omega,$$

那么

$$\int_\Omega \lim_{n\to\infty} f_n=\lim_{n\to\infty}\int_\Omega f_n.$$

勒贝格单调收敛定理非常有用. 例如, 现在我们就可以交换加法与积分的运算次序了.

命题: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to[0,+\infty]$ 与 $g:\Omega\to[0,+\infty]$ 均为非负可测函数, 则

$$\int_\Omega (f+g)=\int_\Omega f+\int_\Omega g.$$

一旦能够交换两个非负函数的加法与积分的运算次序, 根据归纳法, 我们就能够交换任意有限

多个非负函数的加法与积分的运算次序. 更令人惊讶的是, 我们还可以交换无限多个非负函数的求和与积分的次序.

推论 (非负函数的逐项积分定理) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $\{g_n:\Omega\to[0,+\infty]\}$ 为非负可测函数序列, 则

$$\int_\Omega \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty}\int_\Omega g_n.$$

证明 令 $h_n=\sum\limits_{k=1}^{n} g_k$, $\{h_n\}$ 为非负单调递增的函数序列, 则

$$\int_\Omega \sum_{n=1}^{\infty} g_n = \int_\Omega \lim_{n\to\infty} h_n = \lim_{n\to\infty}\int_\Omega h_n = \lim_{n\to\infty}\int_\Omega \sum_{k=1}^{n} g_k = \lim_{n\to\infty}\sum_{k=1}^{n}\int_\Omega g_k = \sum_{n=1}^{\infty}\int_\Omega g_n.$$

其中第二个等号利用了勒贝格单调收敛定理, 倒数第二个等号利用了有限个非负函数的加法与积分的运算次序的可交换性.

勒贝格单调收敛定理说明单调的非负函数序列可以交换极限与积分的次序. 然而, 对于一般的非负函数序列, 等式

$$\int_\Omega \lim_{n\to\infty} f_n = \lim_{n\to\infty}\int_\Omega f_n$$

是不成立的. 对此, 读者可以考查 $\Omega=(0,+\infty)$ 与函数序列

$$f_n(x)=\begin{cases} n, & x\in\left(0,\dfrac{1}{n}\right), \\ 0, & x\in\left[\dfrac{1}{n},+\infty\right). \end{cases}$$

直接计算可得

$$(0=)\int_\Omega \lim_{n\to\infty} f_n \neq \lim_{n\to\infty}\int_\Omega f_n(=1).$$

但是, 如果只考虑下极限, 我们有如下的法图引理.

引理 (法图引理) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $\{f_n:\Omega\to[0,+\infty]\}$ 为非负可测函数序列, 则有

$$\int_\Omega \varliminf_{n\to\infty} f_n \leqslant \varliminf_{n\to\infty}\int_\Omega f_n.$$

思路 因为 $\varliminf\limits_{n\to\infty} f_n=\lim\limits_{n\to\infty}\inf\limits_{k\geqslant n} f_k$, 而函数序列 $\{g_n\}=\{\inf\limits_{k\geqslant n} f_k\}$ 单调, 可运用勒贝格单调收敛定理.

证明 令 $g_n=\inf\limits_{k\geqslant n} f_k$, 则 $\{g_n\}$ 单调, 且 $\lim\limits_{n\to\infty} g_n=\lim\limits_{n\to\infty}\inf\limits_{k\geqslant n} f_k=\varliminf\limits_{n\to\infty} f_n$. 由勒贝格单调收敛定理得

$$\int_\Omega \varliminf_{n\to\infty} f_n = \int_\Omega \lim_{n\to\infty} g_n = \lim_{n\to\infty}\int_\Omega g_n \leqslant \varliminf_{n\to\infty}\int_\Omega f_n.$$

这里最后的不等号是因为 $g_n\leqslant f_n$.

注: 非负函数的逐项积分定理与法图引理没有额外的要求, 只需要函数序列非负即可, 这是非常

好的.

6.3 一般可测函数

现在我们介绍完了关于非负函数的勒贝格积分理论, 接下来我们讨论一般的可测函数. 为了避免出现 $(+\infty)+(-\infty)$ 的情形, 我们将注意力集中在一类特殊的函数——绝对可积的函数上.

定义 (绝对可积函数) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}^*$ 为可测函数, 若

$$\int_\Omega |f| < +\infty,$$

则称 f 是**绝对可积**的.

定义: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}^*$ 为可测函数, 我们分别称称

$$f^+=\max\{f,0\},\quad f^-=-\min\{f,0\}$$

为函数 f 的**正部**和**负部**, 它们都是非负可测的.

容易知道

$$|f|=f^++f^-,\quad f=f^+-f^-.$$

由此可以看出, 若 f 是绝对可积的, 即 $\int_\Omega |f|<+\infty$, 那么由于 f^+ 与 f^- 是非负的, 故 $\int_\Omega f^+$ 与 $\int_\Omega f^-$ 都是有限的. 因此, 我们给出以下定义.

定义 (勒贝格积分) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}^*$ 为绝对可积函数, 则称

$$\int_\Omega f=\int_\Omega f^+-\int_\Omega f^-$$

为 f 的**勒贝格积分**.

注意, 这个定义与前面非负函数的勒贝格积分的定义是一致的. 因为若 f 是非负的, 那么 $f^+=f$, $f^-=0$. 此外, 容易得到以下的三角不等式:

$$\left|\int_\Omega f\right|=\left|\int_\Omega f^+-\int_\Omega f^-\right|\leqslant\int_\Omega f^++\int_\Omega f^-=\int_\Omega |f|.$$

注: 若 Ω' 为 Ω 的可测子集, 那么通过把 f 限制到 Ω', 我们可以定义 f 在 Ω' 上的勒贝格积分为

$$\int_{\Omega'} f=\int_{\Omega'} f|_{\Omega'}\left(=\int_\Omega f\chi_{\Omega'}\right).$$

命题: 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f:\Omega\to\mathbb{R}$ 与 $g:\Omega\to\mathbb{R}$ 均为绝对可积函数, 则

(1) cf 也是绝对可积函数, 且 $\int_\Omega cf=c\int_\Omega f$, 其中 c 为任意实数.

(2) $f+g$ 也是绝对可积函数, 且 $\int_\Omega (f+g)=\int_\Omega f+\int_\Omega g$.

(3) 若在 Ω 上几乎处处都有 $f(x) \leqslant g(x)$, 则 $\displaystyle\int_\Omega f \leqslant \int_\Omega g$.

(4) 若在 Ω 上几乎处处都有 $f(x) = g(x)$, 则 $\displaystyle\int_\Omega f = \int_\Omega g$.

(5) 若 A, B 为 Ω 的互不相交的可测子集, 则 $\displaystyle\int_{A\cup B} f = \int_A f + \int_B f$.

将上述性质 (5) 推广到可数个互不相交的可测子集, 则有如下定理.

定理 (积分区域的可数可加性) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f : \Omega \to \mathbb{R}$ 为绝对可积函数, $\{E_i\}_1^\infty$ 为 Ω 的互不相交的子集, 则

$$\int_{\bigcup\limits_{i=1}^{\infty} E_i} f = \sum_{i=1}^{\infty} \int_{E_i} f.$$

前面我们提到过, 极限运算与积分运算是不能随意交换的. 但是, 如果存在一个绝对可积的函数 F 能够从上方控制住每一个 f_n, 那么我们就能成功地交换极限运算与积分运算的次序了. 这就是下面的极其有用、极其重要的勒贝格控制收敛定理.

定理 (勒贝格控制收敛定理) : 设 Ω 为 $\mathbb{R}^n$ 的可测子集, $\{f_n : \Omega \to \mathbb{R}^*\}$ 为一列逐点收敛的可测函数. 若存在绝对可积函数 $F : \Omega \to \mathbb{R}^*$, 满足

$$|f_n(x)| \leqslant F(x), \quad x \in \Omega, n = 1, 2, \cdots,$$

则有

$$\int_\Omega \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_\Omega f_n.$$

思路 注意到 $\{F + f_n\}$ 与 $\{F - f_n\}$ 非负, 利用法图引理.

证明 设 $f_n \to f$, 因为每个 f_n 可测, 故 f 是可测的. 又因为每个 $|f_n| \leqslant F$, 故 f 是绝对可积的. 下面证明

$$\lim_{n\to\infty} \int_\Omega f_n = \int_\Omega f.$$

因为 $\{F + f_n\}$ 是趋于 $F + f$ 的非负函数序列, 故由法图引理得

$$\int_\Omega (F + f) \leqslant \varliminf_{n\to\infty} \int_\Omega (F + f_n),$$

从而

$$\int_\Omega f \leqslant \varliminf_{n\to\infty} \int_\Omega f_n. \tag{1}$$

又因为 $\{F - f_n\}$ 是趋于 $F - f$ 的非负函数序列, 故又由法图引理得

$$\int_\Omega (F - f) \leqslant \varliminf_{n\to\infty} \int_\Omega (F - f_n),$$

从而 (注意到提出负号之后下极限变为上极限)

$$\int_\Omega f \geqslant \varlimsup_{n\to\infty} \int_\Omega f_n. \tag{2}$$

由 (1)(2) 两式即得

$$\int_\Omega f = \lim_{n\to\infty} \int_\Omega f_n.$$

定理 (逐项积分定理)：设 Ω 为 $\mathbb{R}^n$ 的可测子集, $\{f_n : \Omega \to \mathbb{R}\}$ 为绝对可积函数序列. 若级数 $\sum\limits_{n=1}^{\infty} \int_\Omega |f_n|$ 收敛, 则和函数 $\sum\limits_{n=1}^{\infty} f_n$ 绝对可积, 且

$$\int_\Omega \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_\Omega f_n.$$

证明 由条件知

$$\int_\Omega \left| \sum_{n=1}^{\infty} f_n \right| \leqslant \int_\Omega \sum_{n=1}^{\infty} |f_n| = \sum_{n=1}^{\infty} \int_\Omega |f_n| < +\infty, \tag{1}$$

故 $\sum\limits_{n=1}^{\infty} f_n$ 绝对可积, 其中 (1) 式中间的等号利用了非负函数的逐项积分定理. 令 $g = \sum\limits_{n=1}^{\infty} f_n$, $g_n = \sum\limits_{i=1}^{n} f_i$, $F = \sum\limits_{n=1}^{\infty} |f_n|$. 由 (1) 式得 $\int_\Omega F < +\infty$, 即 F 是绝对可积函数. 因为

$$|g_n(x)| = \left| \sum_{i=1}^{n} f_i(x) \right| \leqslant \sum_{i=1}^{n} |f_i(x)| \leqslant \sum_{i=1}^{\infty} |f_i(x)| = F(x),$$

故由勒贝格控制收敛定理知

$$\int_\Omega \sum_{n=1}^{\infty} f_n = \int_\Omega g = \int_\Omega \lim_{n\to\infty} g_n = \lim_{n\to\infty} \int_\Omega g_n = \lim_{n\to\infty} \int_\Omega \sum_{i=1}^{n} f_i = \lim_{n\to\infty} \sum_{i=1}^{n} \int_\Omega f_i = \sum_{n=1}^{\infty} \int_\Omega f_n.$$

结论得证.

定理 (积分号下求导)：设 Ω 为 $\mathbb{R}^n$ 的可测子集, $f : \Omega \times (a, b) \to \mathbb{R}$ 为 $n+1$ 元函数. 若对任意的 $t \in (a, b)$, $f(x, t)$ 作为 x 的函数在 Ω 上绝对可积; 对任意的 $x \in \Omega$, 作为 t 的函数在 (a, b) 上可导, 且 $\left| \dfrac{\partial}{\partial t} f(x, t) \right| \leqslant F(x)$, 这里 F 是一个绝对可积函数, 则

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega f(x, t) \mathrm{d}x = \int_\Omega \frac{\partial}{\partial t} f(x, t) \mathrm{d}x.$$

6.4　与黎曼积分的比较

勒贝格积分处理的对象是绝对可积函数. 对其而言, 可测函数 f 可积等价于 $|f|$ 可积. 勒贝格积分是黎曼积分的推广, 但不是黎曼反常积分的推广（若只考虑非负函数, 则勒贝格积分也是黎曼

反常积分的推广). 这主要是因为勒贝格积分是绝对收敛的积分, 而黎曼反常积分不一定绝对收敛.

定理: 设 f 为 $[a,b]$ 上的有界函数, 则 f 黎曼可积的充要条件为 f 的不连续点构成一零测集.

附注:

(1) 对于黎曼积分而言, 若函数 f 是可积的, 则 $|f|$ 也一定是可积的. 但是, 若 $|f|$ 可积, 我们不能得到 f 自身是可积的.

(2) 对于黎曼反常积分而言, 若 $|f|$ 可积, 则 f 自身也一定是可积的 (此时称 f 的反常积分是绝对收敛的). 但是若 f 可积, 我们不能得到 $|f|$ 是可积的 (此时称 f 的反常积分是条件收敛的).

6.5 Fubini 定理

定理 (Tonelli 定理, 非负可测函数的情形) : 设 $A\times B$ 为 $\mathbb{R}^{p+q}=\mathbb{R}^p\times\mathbb{R}^q$ 的可测子集, 函数 $f:A\times B\to\mathbb{R}$ 为非负可测函数, 则

(1) 对于几乎处处的 $y\in B$, $f(x,y)$ 作为 x 的函数在 A 上非负可测, 且 $G(y)=\int_A f(x,y)\mathrm{d}x$ 在 B 上非负可测.

(2) 对于几乎处处的 $x\in A$, $f(x,y)$ 作为 y 的函数在 B 上非负可测, 且 $F(x)=\int_B f(x,y)\mathrm{d}y$ 在 A 上非负可测.

(3) $\int_A F(x)\mathrm{d}x=\int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y=\int_B G(y)\mathrm{d}y$, 也即

$$\int_A \mathrm{d}x\int_B f(x,y)\mathrm{d}y=\int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y=\int_B \mathrm{d}y\int_A f(x,y)\mathrm{d}x.$$

定理 (Fubini 定理, 可积函数的情形) : 设 $A\times B$ 为 $\mathbb{R}^{p+q}=\mathbb{R}^p\times\mathbb{R}^q$ 的可测子集, 函数 $f:A\times B\to\mathbb{R}$ 为绝对可积函数, 则

(1) 对于几乎处处的 $y\in B$, $f(x,y)$ 作为 x 的函数在 A 上绝对可积, 且 $G(y)=\int_A f(x,y)\mathrm{d}x$ 在 B 上绝对可积.

(2) 对于几乎处处的 $x\in A$, $f(x,y)$ 作为 y 的函数在 B 上绝对可积, 且 $F(x)=\int_B f(x,y)\mathrm{d}y$ 在 A 上绝对可积.

(3) $\int_A F(x)\mathrm{d}x=\int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y=\int_B G(y)\mathrm{d}y$, 也即

$$\int_A \mathrm{d}x\int_B f(x,y)\mathrm{d}y=\int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y=\int_B \mathrm{d}y\int_A f(x,y)\mathrm{d}x.$$

思路 对 f^+ 与 f^- 应用 Tonelli 定理即可.

Fubini 定理说明了高维积分与低维积分之间的联系, 它使得我们能够将重积分化为累次积分进行计算. 此外, Fubini 定理告诉我们, 只要重积分有限, 那么它就和累次积分相等.

注: (1) Fubini 定理提供了判断一个函数是否绝对可积的方法: 若两个累次积分不相等, 即

$$\int_A \mathrm{d}x\int_B f(x,y)\mathrm{d}y\neq\int_B \mathrm{d}y\int_A f(x,y)\mathrm{d}x,$$

那么 f 在 $A\times B$ 上不是绝对可积的.

(2) 需要注意的是, 即使两个累次积分都存在且相等, 即

$$\int_A \mathrm{d}x \int_B f(x,y)\mathrm{d}y = \int_B \mathrm{d}y \int_A f(x,y)\mathrm{d}x,$$

函数 f 也可能不是绝对可积的.

对于注记 (1), 读者可以考查下面的例子.

例 积分区域 $A\times B=(0,1)\times(0,1)$, 被积函数为

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

直接计算得

$$\begin{aligned}
\int_{(0,1)} \mathrm{d}x \int_{(0,1)} f(x,y)\mathrm{d}y &= \int_0^1 \left(\frac{y}{x^2+y^2}\bigg|_{y=0}^{y=1} \right) \mathrm{d}x = \int_0^1 \frac{1}{1+x^2}\mathrm{d}x = \frac{\pi}{4}, \\
\int_{(0,1)} \mathrm{d}y \int_{(0,1)} f(x,y)\mathrm{d}x &= \int_0^1 \left(-\frac{x}{x^2+y^2}\bigg|_{x=0}^{x=1} \right) \mathrm{d}y = -\int_0^1 \frac{1}{1+y^2}\mathrm{d}y = -\frac{\pi}{4}.
\end{aligned}$$

因此由 Fubini 定理知 f 在 $A\times B$ 上不是绝对可积的.

7 含参量积分

本章我们来考察积分

$$\int_a^b f(x,t)\mathrm{d}x \quad 与 \quad \int_a^{+\infty} g(x,t)\mathrm{d}x.$$

这些积分的被积函数除了与积分变量 x 有关之外, 还依赖于参变元 t. 由这样的积分定义了参变元 t 的函数

$$\varphi(t)=\int_a^b f(x,t)\mathrm{d}x \quad 或 \quad \psi(t)=\int_a^{+\infty} g(x,t)\mathrm{d}x.$$

下面我们就来研究用这种方式定义的函数.

7.1 含参量常义积分

例 计算椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ $(b>a>0)$ 的周长.

解 只需计算位于第一象限的椭圆曲线段 L 的弧长即可. L 的参数方程为

$$\begin{cases} x=a\cos t, \\ y=b\sin t, \end{cases} \quad t\in\left[0,\frac{\pi}{2}\right].$$

从而 L 的弧长为

$$\begin{aligned}\int_L \mathrm{d}s &= \int_0^{\frac{\pi}{2}} \sqrt{a^2\sin^2 t + b^2\cos^2 t}\mathrm{d}t = \int_0^{\frac{\pi}{2}} \sqrt{b^2 - (b^2 - a^2)\sin^2 t}\mathrm{d}t \\ &= b\int_0^{\frac{\pi}{2}} \sqrt{1 - \frac{b^2 - a^2}{b^2}\sin^2 t}\mathrm{d}t =: b\int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\sin^2 t}\mathrm{d}t.\end{aligned}$$

这最后一式就是含参变量 k 的积分, 我们把它叫做第二类完全椭圆积分. 遗憾的是, 该积分的被积函数 $\sqrt{1 - k^2\sin^2 t}$ 无法用初等函数表示, 因此计算此积分时, 通常只能采用数值积分的方法.

下面我们研究含参量积分的分析性质. (对于连续性与可微性, 闭区间 $[a, b]$ 可以换为一般的区间 I, 但对于可积性则必须使用闭区间.)

定理 (连续性) : 设 $f(x, y)$ 在闭矩形 $[a, b] \times [c, d]$ 上连续, 那么函数

$$I(y) = \int_a^b f(x, y)\mathrm{d}x$$

在 $[c, d]$ 上连续.

注: 上述定理即是说极限与积分可以交换次序:

$$\int_a^b \lim_{y\to y_0} f(x, y)\mathrm{d}x = \lim_{y\to y_0}\int_a^b f(x, y)\mathrm{d}x.$$

定理 (可积性, 交换积分次序) : 设 $f(x, y)$ 在 $D = [a, b] \times [c, d]$ 上连续, 则两个累次积分存在且相等, 即

$$\int_c^d \mathrm{d}y\int_a^b f(x, y)\mathrm{d}x = \int_a^b \mathrm{d}x\int_c^d f(x, y)\mathrm{d}y.$$

证明 函数 f 满足重积分化为累次积分的条件, 故

$$\int_c^d \mathrm{d}y\int_a^b f(x, y)\mathrm{d}x = \iint_D f(x, y)\mathrm{d}x\mathrm{d}y = \int_a^b \mathrm{d}x\int_c^d f(x, y)\mathrm{d}y.$$

定理 (可微性, 积分号下求导) : 设函数 f 与 f_y 都在闭矩形 $[a, b] \times [c, d]$ 上连续, 则函数 $I(y) = \displaystyle\int_a^b f(x, y)\mathrm{d}x$ 在 $[c, d]$ 上可导, 且

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^b f(x, y)\mathrm{d}x = \int_a^b \frac{\partial}{\partial y} f(x, y)\mathrm{d}x.$$

证明 任取 $y_0 \in [a, b]$. 下证

$$\lim_{\Delta y\to 0}\frac{I(y_0 + \Delta y) - I(y_0)}{\Delta y} = \int_a^b \frac{\partial}{\partial y} f(x, y_0)\mathrm{d}x.$$

由 Lagrange 中值定理可得

$$\begin{aligned}\lim_{\Delta y\to 0}\frac{I(y_0+\Delta y)-I(y_0)}{\Delta y} &= \lim_{\Delta y\to 0}\int_a^b \frac{f(x,y_0+\Delta y)-f(x,y_0)}{\Delta y}\mathrm{d}x \\ &= \lim_{\Delta y\to 0}\int_a^b f_y(x,y_0+\theta\Delta y)\mathrm{d}x \\ &= \int_a^b \lim_{\Delta y\to 0} f_y(x,y_0+\theta\Delta y)\mathrm{d}x\end{aligned}$$

这里最后一个等号利用了 f_y 的连续性, $0<\theta<1$. 从而

$$\frac{\mathrm{d}I}{\mathrm{d}y}(y_0) = \int_a^b f_y(x,y_0)\mathrm{d}x.$$

因此 $I(y)$ 在 y_0 处可导, 且导数值由上式给出. 由 y_0 的任意性知结论得证.

下面是积分上下限随参量变动的相应结论.

定理 (连续性) : 设 $f(x,y)$ 在闭矩形 $[a,b]\times[c,d]$ 上连续, 那么函数

$$F(y) = \int_{a(y)}^{b(y)} f(x,y)\mathrm{d}x$$

也在 $[c,d]$ 上连续, 其中 $a(y)$ 与 $b(y)$ 为取值于 $[c,d]$ 的连续函数.

定理 (可微性) : 设函数 f 与 f_y 都在闭矩形 $[a,b]\times[c,d]$ 上连续, 则函数

$$F(y) = \int_{a(y)}^{b(y)} f(x,y)\mathrm{d}x$$

在 $[c,d]$ 上可导, 且导数为

$$F'(y) = \int_{a(y)}^{b(y)} f_y(x,y)\mathrm{d}x - f(a(y),y)a'(y) + f(b(y),y)b'(y).$$

这里 $a(y)$ 与 $b(y)$ 是取值于 $[a,b]$ 的可导函数.

思路　利用复合函数求导.

例　设 $b>a>0$, 计算积分

$$I = \int_0^1 \frac{x^b-x^a}{\ln x}\mathrm{d}x.$$

解　首先我们需要明确一点, I 不是瑕积分, 因为

$$\lim_{x\to 0^+}\frac{x^b-x^a}{\ln x} = 0.$$

然后, 注意到

$$f(b)-f(a) = \int_a^b f',$$

因此令 $f(y)=\dfrac{x^y}{\ln x}$, 则

$$\frac{x^b-x^a}{\ln x}=\frac{x^b}{\ln x}-\frac{x^a}{\ln x}=f(b)-f(a)=\int_a^b f'(y)\mathrm{d}y=\int_a^b x^y\mathrm{d}y.$$

从而

$$I=\int_0^1\frac{x^b-x^a}{\ln x}\mathrm{d}x=\int_0^1\left(\int_a^b x^y\mathrm{d}y\right)\mathrm{d}x=\int_a^b\left(\int_0^1 x^y\mathrm{d}x\right)\mathrm{d}y=\int_a^b\frac{1}{1+y}\mathrm{d}y=\ln\frac{1+b}{1+a}.$$

这里我们能够交换积分次序, 因为被积函数在 $[0,1]\times[a,b]$ 上是连续的 (补充定义被积函数在 $(0,y)$ 处的值为 0) .

7.2 含参量反常积分一致收敛的定义

含参量反常积分也分为两种——无穷区间上的含参量反常积分与无界函数的含参量反常积分.

设 I 为一个区间, 若对每一个 $y\in I$, 反常积分

$$\int_a^{+\infty}f(x,y)\mathrm{d}x$$

都收敛, 将由此确定的函数记为 $\Phi(y)$, 则有

$$\Phi(y)=\int_a^{+\infty}f(x,y)\mathrm{d}x,\quad y\in I.$$

同样的, 我们需要讨论函数 $\Phi(y)$ 的连续性、可积性与可微性. 为此, 我们引入含参量反常积分一致收敛的概念.

定义 (含参量反常积分的一致收敛) : 设有逐点收敛的含参量反常积分 $\Phi(y)=\displaystyle\int_a^{+\infty}f(x,y)\mathrm{d}x$, $y\in I$. 若对任意的 $\varepsilon>0$, 存在 $A>a$, 当 $A'>A$ 时, 对任意的 $y\in I$, 有

$$\left|\int_a^{A'}f(x,y)\mathrm{d}x-\Phi(y)\right|<\varepsilon, \tag{1}$$

则称含参量反常积分 $\displaystyle\int_a^{+\infty}f(x,y)\mathrm{d}x$ 在区间 I 上一致收敛于 $\Phi(y)$.

注: 上述定义中的 (1) 式相当于

$$\left|\int_{A'}^{+\infty}f(x,y)\mathrm{d}x\right|<\varepsilon.$$

7.3 含参量反常积分一致收敛的判别

定理 (柯西准则) : $\displaystyle\int_a^{+\infty}f(x,y)\mathrm{d}x$ 一致收敛的充要条件为对任意的 $\varepsilon>0$, 存在 $A>a$, 当 $A_1,A_2>A$ 时, 对任意的 $y\in I$, 有

$$\left|\int_{A_1}^{A_2}f(x,y)\mathrm{d}x\right|<\varepsilon.$$

定理: $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 一致收敛的充要条件为对任一满足 $A_1=a$, $\lim\limits_{n\to\infty} A_n=+\infty$ 的递增序列 $\{A_n\}$, 函数项级数
$$\sum_{n=1}^{\infty}\int_{A_n}^{A_{n+1}} f(x,y)\mathrm{d}x$$
均一致收敛.

证明 ($\Longrightarrow$) 记 $u_n(y)=\int_{A_n}^{A_{n+1}} f(x,y)\mathrm{d}x$. 因为 $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 一致收敛, 故对任意的 $\varepsilon>0$, 存在 $A>a$, 当 $A',A''>A$ 时, 对任意的 $y\in I$, 有
$$\left|\int_{A'}^{A''} f(x,y)\mathrm{d}x\right|<\varepsilon.$$
因为 $\lim\limits_{n\to\infty} A_n=+\infty$, 故对上述 A, 存在 $N>0$, 当 $n>N$ 时, 对任意的正整数 p, 有
$$A_{n+1},A_{n+2},\cdots,A_{n+p}>A.$$
从而对任意的 $y\in I$, 有
$$\begin{aligned}|u_{n+1}(y)+\cdots+u_{n+p}(y)|&=\left|\int_{A_{n+1}}^{A_{n+2}} f(x,y)\mathrm{d}x+\cdots+\int_{A_{n+p}}^{A_{n+p+1}} f(x,y)\mathrm{d}x\right|\\&=\left|\int_{A_{n+1}}^{A_{n+p+1}} f(x,y)\mathrm{d}x\right|<\varepsilon.\end{aligned}$$
由函数项级数的柯西准则知 $\sum\limits_{n=1}^{\infty} u_n(y)=\sum\limits_{n=1}^{\infty}\int_{A_n}^{A_{n+1}} f(x,y)\mathrm{d}x$ 一致收敛.

($\Longleftarrow$) 假设 $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 不一致收敛, 那么存在 $\varepsilon_0>0$, 对任意的 $A>a$, 存在 $A',A''>A$ 与 $y_0\in I$, 使得
$$\left|\int_{A'}^{A''} f(x,y_0)\mathrm{d}x\right|\geqslant\varepsilon_0.$$
取 $a_1=a+1$, 存在 $A_2>A_1>a_1$ 与 $y_1\in I$, 使得 $\left|\int_{A_1}^{A_2} f(x,y_1)\mathrm{d}x\right|\geqslant\varepsilon_0$.

取 $a_3=A_2+1$, 存在 $A_4>A_3>a_3$ 与 $y_2\in I$, 使得 $\left|\int_{A_3}^{A_4} f(x,y_2)\mathrm{d}x\right|\geqslant\varepsilon_0$.

$\cdots$

取 $a_{2n-1}=A_{2n-2}+1$, 存在 $A_{2n}>A_{2n-1}>a_{2n-1}$ 与 $y_n\in I$, 使得 $\left|\int_{A_{2n-1}}^{A_{2n}} f(x,y_n)\mathrm{d}x\right|\geqslant\varepsilon_0$.

这样一来, 我们得到了一列趋于 $+\infty$ 的递增数列 $\{A_n\}$. 令
$$u_n(y)=\int_{A_n}^{A_{n+1}} f(x,y)\mathrm{d}x,$$
则 $\sum\limits_{n=1}^{\infty} u_n(y)$ 不一致收敛, 矛盾. 故假设错误, $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 一致收敛.

下面利用柯西准则导出魏尔斯特拉斯判别法.

定理 (魏尔斯特拉斯判别法)：设反常积分 $\int_a^{+\infty} F(x)\mathrm{d}x$ 收敛, 若对任意的 $(x,y)\in[a,+\infty)\times I$, 有

$$|f(x,y)|\leqslant F(x),$$

则含参量反常积分 $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 一致收敛.

证明 因为 $\int_a^{+\infty} F(x)\mathrm{d}x$ 收敛, 故对任意的 $\varepsilon>0$, 存在 $A>a$, 当 $A_1,A_2>A$ 时, 有

$$\left|\int_{A_1}^{A_2} F(x)\mathrm{d}x\right|=\int_{A_1}^{A_2} F(x)\mathrm{d}x<\varepsilon.$$

从而对任意的 $y\in I$, 有

$$\left|\int_{A_1}^{A_2} f(x,y)\mathrm{d}x\right|\leqslant\int_{A_1}^{A_2}|f(x,y)|\mathrm{d}x\leqslant\int_{A_1}^{A_2} F(x)\mathrm{d}x<\varepsilon.$$

由柯西准则知 $\int_a^{+\infty} f(x,y)\mathrm{d}x$ 一致收敛.

下面我们导出含参量反常积分的狄利克雷判别法与阿贝尔判别法, 它们的证明仍然利用的是积分第二中值定理.

定理 (狄利克雷判别法)：若 $\int_a^{A} f(x,y)\mathrm{d}x$ 一致有界, $g(x,y)$ 关于 x 单调, 关于 y 一致收敛于 0, 则 $\int_a^{+\infty} f(x,y)g(x,y)\mathrm{d}x$ 一致收敛.

证明 设 $\left|\int_a^{A} f(x,y)\mathrm{d}x\right|\leqslant M$, $\forall y\in I$, $\forall A>a$. 因此对任意的 $A_1,A_2>a$, 有

$$\left|\int_{A_1}^{A_2} f(x,y)\mathrm{d}x\right|=\left|\int_{a}^{A_2} f(x,y)\mathrm{d}x-\int_{a}^{A_1} f(x,y)\mathrm{d}x\right|\leqslant 2M.$$

因为 $g(x,y)$ 对 y 一致趋于 0 $(x\to+\infty)$, 因此对任意的 $\varepsilon>0$, 存在 $A>a$, 当 $x>A$ 时, 对任意 $y\in I$, 有

$$|g(x,y)|<\varepsilon.$$

因为 $g(x,y)$ 关于 x 单调, 故由积分第二中值定理, 当 $A_1,A_2>A$ 时, 有

$$\begin{aligned}\left|\int_{A_1}^{A_2} f(x,y)g(x,y)\mathrm{d}x\right|&\leqslant|g(A_1,y)|\left|\int_{A_1}^{\xi} f(x,y)\mathrm{d}x\right|+|g(A_2,y)|\left|\int_{\xi}^{A_2} f(x,y)\mathrm{d}x\right|\\&<2M\varepsilon+2M\varepsilon=4M\varepsilon.\end{aligned}$$

其中 $\xi\in[A_1,A_2]$. 由柯西准则知 $\int_a^{+\infty} f(x,y)g(x,y)\mathrm{d}x$ 一致收敛.

定理 (阿贝尔判别法)： 若 $f(x,y)$ 关于 x 单调, 且一致有界, $\int_a^{+\infty} g(x,y)\mathrm{d}x$ 一致收敛, 则

$\int_a^{+\infty} f(x,y)g(x,y)\mathrm{d}x$ 一致收敛.

证明 设 $|f(x,y)| \leqslant M$, $\forall x \in [a,+\infty)$, $\forall y \in I$. 因为 $\int_a^{+\infty} g(x,y)\mathrm{d}x$ 一致收敛, 故对任意的 $\varepsilon > 0$, 存在 $A > a$, 当 $A_1, A_2 > A$ 时, 对任意的 $y \in I$, 有

$$\left|\int_{A_1}^{A_2} g(x,y)\mathrm{d}x\right| < \varepsilon.$$

因为 $f(x,y)$ 关于 x 单调, 故由积分第二中值定理, 有

$$\begin{aligned}\left|\int_{A_1}^{A_2} f(x,y)g(x,y)\mathrm{d}x\right| &\leqslant |f(A_1,y)|\left|\int_{A_1}^{\xi} g(x,y)\mathrm{d}x\right| + |f(A_2,y)|\left|\int_{\xi}^{A_2} g(x,y)\mathrm{d}x\right| \\ &< M\varepsilon + M\varepsilon = 2M\varepsilon.\end{aligned}$$

其中 $\xi \in [A_1, A_2]$. 由柯西准则知 $\int_a^{+\infty} f(x,y)g(x,y)\mathrm{d}x$ 一致收敛.

注: 细心的读者可能已经发现了, 狄利克雷判别法与阿贝尔判别法都只要求函数关于 x 单调, 这是因为证明过程中所使用的的积分第二中值定理只要求函数关于积分变量是单调的.

例 证明含参量反常积分 $\int_0^{+\infty} \mathrm{e}^{-xy}\frac{\sin x}{x}\mathrm{d}x$ 在 $[0,+\infty)$ 上一致收敛.

证明 令 $g(x,y) = \frac{\sin x}{x}$, 则 $\int_0^{+\infty} g(x,y)\mathrm{d}x$ 对 y 一致收敛. 令 $f(x,y) = \mathrm{e}^{-xy}$, 则 f 关于 x 单调, 且 $|f(x,y)| \leqslant 1$. 由阿贝尔判别法知所给积分在 $[0,+\infty)$ 上一致收敛.

例 证明含参量反常积分 $\int_1^{+\infty} \frac{y\sin xy}{1+y^2}\mathrm{d}y$ 在 $(0,+\infty)$ 上内闭一致收敛.

证明 任取 $[a,b] \subseteq (0,+\infty)$. 令 $f(x,y) = \sin xy$, 则对任意的 $x \in [a,b]$, 有

$$\int_1^A \sin xy\mathrm{d}y = \left|\frac{\cos Ax - \cos x}{x}\right| \leqslant \frac{2}{a}, \quad \forall A > 1.$$

因此 $f(x,y)$ 一致有界. 令 $g(x,y) = \frac{y}{1+y^2}$, 则

$$g_y(x,y) = \frac{1-y^2}{(1+y^2)^2} \leqslant 0,$$

故 $g(x,y)$ 对 y 单调. 又 $g(x,y)$ 关于 x 一致趋于 0 $(y \to +\infty)$, 因此由狄利克雷判别法知所给积分在 $[a,b]$ 上一致收敛. 由 $[a,b]$ 的任意性知所给积分在 $(0,+\infty)$ 上内闭一致收敛.

7.4 含参量反常积分一致收敛的分析性质

定理 (连续性) : 设 $f(x,y)$ 在 $[a,+\infty) \times I$ 上连续. 若 $\Phi(y) = \int_a^{+\infty} f(x,y)\mathrm{d}x$ 在 I 上一致收敛, 则 $\Phi(y)$ 在 I 上连续.

注: 上述定理也可表述为

$$\int_a^{+\infty}\lim_{y\to y_0}f(x,y)\mathrm{d}x=\lim_{y\to y_0}\int_a^{+\infty}f(x,y)\mathrm{d}x.$$

定理 (交换积分次序, 参变量有界) : 设 $f(x,y)$ 在 $[a,+\infty)\times[c,d]$ 上连续. 若 $\int_a^{+\infty}f(x,y)\mathrm{d}x$ 在 $[c,d]$ 上一致收敛, 则

$$\int_c^d\mathrm{d}y\int_a^{+\infty}f(x,y)\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_c^d f(x,y)\mathrm{d}y.$$

当 $[c,d]$ 改为无穷区间 $[c,+\infty)$ 时, 上述定理的条件就不足以保证积分次序可交换了, 但这时我们有以下结论.

定理 (交换积分次序, 参变量无界) : 设 $f(x,y)$ 在 $[a,+\infty)\times[c,+\infty)$ 上连续, 若

(1) $\int_a^{+\infty}f(x,y)\mathrm{d}x$ 在 $[c,+\infty)$ 上内闭一致收敛, $\int_c^{+\infty}f(x,y)\mathrm{d}x$ 在 $[a,+\infty)$ 上内闭一致收敛.

(2) $\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty}|f(x,y)|\mathrm{d}x$ 与 $\int_a^{+\infty}\mathrm{d}x\int_c^{+\infty}|f(x,y)|\mathrm{d}y$ 至少一个存在.

那么积分次序可交换, 即

$$\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty}f(x,y)\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_c^{+\infty}f(x,y)\mathrm{d}y.$$

定理 (积分号下求导) : 设 $f(x,y)$ 与 $f_y(x,y)$ 在在 $[a,+\infty)\times I$ 上连续, $\int_a^{+\infty}f(x,y)\mathrm{d}x$ 收敛, $\int_a^{+\infty}f_y(x,y)\mathrm{d}x$ 一致收敛, 则 $\Phi(y)=\int_a^{+\infty}f(x,y)\mathrm{d}x$ 可导, 且

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^{+\infty}f(x,y)\mathrm{d}x=\int_a^{+\infty}\frac{\partial}{\partial y}f(x,y)\mathrm{d}x.$$

7.5 引入“收敛因子”计算广义积分

引入“收敛因子”是计算广义积分时经常用到的一种方法. 我们以下面的题目为例进行说明, 并借此求出一个重要积分——狄利克雷积分 $\int_0^{+\infty}\frac{\sin x}{x}\mathrm{d}x$ 的值.

例 计算下列积分.

(1) $\int_0^{+\infty}\mathrm{e}^{-\alpha x}\cos\beta x\mathrm{d}x$. (2) $\int_0^{+\infty}\mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x$. (3) $\int_0^{+\infty}\frac{\sin\beta x}{x}\mathrm{d}x$.

其中 $\alpha>0$, $\beta\in\mathbb{R}$.

解 (1) 直接计算即得

$$\int_0^{+\infty}\mathrm{e}^{-\alpha x}\cos\beta x\mathrm{d}x=\frac{\alpha}{\alpha^2+\beta^2}.$$

(2) 记 $I(\beta)=\int_0^{+\infty}\mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x$. 将被积函数记为 $F(x,\beta)$, 对 β 求导得

$$F_\beta(x,\beta)=\mathrm{e}^{-\alpha x}\cos\beta x.$$

由魏尔斯特拉斯判别法知反常积分 $\int_0^{+\infty} F_\beta(x,\beta)\mathrm{d}x$ 一致收敛, 从而可积分号下求导:

$$I'(\beta) = \int_0^{+\infty} F_\beta(x,\beta)\mathrm{d}x = \int_0^{+\infty} \mathrm{e}^{-\alpha x}\cos\beta x\mathrm{d}x = \frac{\alpha}{\alpha^2+\beta^2}.$$

最后一个等号利用了 (1) 的结论. 因此

$$I'(\beta) = \frac{\alpha}{\alpha^2+\beta^2}.$$

求解该微分方程得

$$I(\beta) = \arctan\frac{\beta}{\alpha} + C.$$

带入 $I(0)=0$ 得 $C=0$, 因此

$$I(\beta) = \int_0^{+\infty} \mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x = \arctan\frac{\beta}{\alpha}.$$

(3) 记 $J(\alpha) = \int_0^{+\infty} \mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x$, 被积函数记为 $G(x,\alpha)$. 因为 $G(x,\alpha)$ 在 $[0,+\infty)\times[0,+\infty)$ 连续 (补充定义 $G(0,\alpha)=\beta$), 且 $\int_0^{+\infty} \mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x$ 在 $[0,+\infty)$ 上一致收敛, 故 $J(a)$ 在 $[0,+\infty)$ 上连续, 从而

$$\lim_{\alpha\to 0^+}\int_0^{+\infty} \mathrm{e}^{-\alpha x}\frac{\sin\beta x}{x}\mathrm{d}x = \int_0^{+\infty}\frac{\sin\beta x}{x}\mathrm{d}x.$$

因此

$$\int_0^{+\infty}\frac{\sin\beta x}{x}\mathrm{d}x = \lim_{\alpha\to 0^+}\arctan\frac{\beta}{\alpha} = \begin{cases} \dfrac{\pi}{2}, & \beta > 0, \\ 0, & \beta = 0, \\ -\dfrac{\pi}{2}, & \beta < 0. \end{cases}$$

注: 含参积分 $\int_0^{+\infty}\frac{\sin\beta x}{x}\mathrm{d}x$ 不允许在积分号下求导, 因为那样的话将会得到一个发散的积分

$$\int_0^{+\infty}\cos\beta x\mathrm{d}x.$$

为了克服这一困难, 我们引入了收敛因子 $\mathrm{e}^{-\alpha x}$, 从而计算出了积分. 收敛因子的引入看似将问题变得复杂, 但是却让我们成功解决了这一棘手的积分问题, 它是我们在计算广义积分时经常使用的一种方法.

注: 我们顺带得到了积分

$$\int_0^{+\infty}\frac{\sin x}{x}\mathrm{d}x = \frac{\pi}{2}.$$

并且, 由此积分我们也能反过来得到上述例题 (3) 的结论:

$$\int_0^{+\infty} \frac{\sin \beta x}{x} \mathrm{d}x = \frac{\pi}{2} \operatorname{sgn}(\beta).$$

为此, 只需作变量替换 $y = \beta x$ 即可.

7.6 “相减型积分”的计算

除了上一节所讲的引入收敛因子求积分的方法外, 还有一类积分问题我们要特别注意——“相减型积分”. 请看下面的例题.

例 设 $b > a > 0$, 试求 $\displaystyle\int_0^{+\infty} \frac{\cos ax - \cos bx}{x^2} \mathrm{d}x$.

解 令 $f(y) = -\dfrac{\cos xy}{x^2}$, 则 $f'(y) = \dfrac{\sin xy}{x}$, 从而

$$\frac{\cos ax - \cos bx}{x^2} = f(b) - f(a) = \int_a^b f'(y) \mathrm{d}y = \int_a^b \frac{\sin xy}{x} \mathrm{d}y.$$

从而原积分

$$I = \int_0^{+\infty} \frac{\cos ax - \cos bx}{x^2} \mathrm{d}x = \int_0^{+\infty} \left(\int_a^b \frac{\sin xy}{x} \mathrm{d}y \right) \mathrm{d}x = \int_a^b \left(\int_0^{+\infty} \frac{\sin xy}{x} \mathrm{d}x \right) \mathrm{d}y.$$

这里我们交换了积分次序, 因为 $\displaystyle\int_0^{+\infty} \frac{\sin xy}{x} \mathrm{d}x$ 在 $[a, b]$ 上是一致收敛的 (可用狄利克雷判别法验证). 作变量替换 $u = xy$, 则 $x = \dfrac{u}{y}$, 从而

$$\int_0^{+\infty} \frac{\sin xy}{x} \mathrm{d}x = \int_0^{+\infty} \frac{\sin u}{u} \mathrm{d}u = \frac{\pi}{2}.$$

因此

$$I = \int_a^b \frac{\pi}{2} \mathrm{d}y = \frac{\pi}{2}(b - a).$$

注: 读者应该已经看出来了, 如果被积函数是两个同型函数作差, 我们就可以尝试将这个被积函数利用牛顿-莱布尼兹公式转化为积分, 从而将原积分化为累次积分进行计算.

7.7 Γ 函数

阶乘本来只对非负整数有定义:

$$0! = 1, \quad n! = n \cdot [(n-1)!].$$

借助于含参变量积分定义的 Γ 函数, 可以看成是“阶乘”的推广.

定义: 我们把含参变量积分

$$\Gamma(s)=\int_0^{+\infty}\mathrm{e}^{-x}x^{s-1}\mathrm{d}x$$

称为 Γ 函数.

引理: Γ 函数在 $(0,+\infty)$ 上有定义且在 $(0,+\infty)$ 上连续.

思路 由广义积分那一章最后一节的讨论可知 Γ 函数在 $(0,+\infty)$ 上有定义. 因为 $\Gamma(s)$ 在 $(0,+\infty)$ 内闭一致收敛, 被积函数在 $(0,+\infty)\times(0,+\infty)$ 上连续, 故 $\Gamma(s)$ 在 $(0,+\infty)$ 上连续.

性质:

(1) (正则性) $\Gamma(s)>0, \Gamma(1)=1$.

(2) (阶乘性) 对任意的 $s>0, \Gamma(s+1)=s\Gamma(s)$.

(3) (对数凸性) $\ln\Gamma(s)$ 在 $(0,+\infty)$ 上是凸函数.

证明 (1) 由定义可得 $\Gamma(s)>0$, 直接计算得 $\Gamma(1)=1$.

(2) 直接利用分部积分公式验证即可.

(3) 我们有赫德尔不等式: 设 $f(t),g(t)$ 非负, $\lambda,\mu>0,\lambda+\mu=1$, 则

$$\int_a^b(f(t))^\lambda(g(t))^\mu\mathrm{d}t\leqslant\left(\int_a^b f(t)\mathrm{d}t\right)^\lambda\left(\int_a^b g(t)\mathrm{d}t\right)^\mu.$$

于是当 $[a,b]\subseteq(0,+\infty)$, $s,t>0$ 时,

$$\begin{aligned}\int_a^b\mathrm{e}^{-x}x^{\lambda s+\mu t-1}\mathrm{d}x&=\int_a^b\left(x^{s-1}\mathrm{e}^{-x}\right)^\lambda\left(x^{t-1}\mathrm{e}^{-x}\right)^\mu\mathrm{d}x\\&\leqslant\left(\int_a^b x^{s-1}\mathrm{e}^{-x}\mathrm{d}x\right)^\lambda\left(\int_a^b x^{t-1}\mathrm{e}^{-x}\mathrm{d}x\right)^\mu.\end{aligned}$$

令 $a\to0^+$, $b\to+\infty$, 得

$$\Gamma(\lambda s+\mu t)\leqslant(\Gamma(s))^\lambda(\Gamma(t))^\mu.$$

两边取对数, 即得

$$\ln\Gamma(\lambda s+\mu t)\leqslant\lambda\ln\Gamma(s)+\mu\ln\Gamma(t).$$

因此 $\ln\Gamma(s)$ 为凸函数.

实际上, 上述三条性质完全决定了 Γ 函数. 也就是说, 如果某个定义在 $(0,+\infty)$ 上的函数满足上述三条性质, 那么该函数一定是 Γ 函数, 这一事实最早是由玻尔 (H.Bohr) 与莫勒儒普 (J.Mollerup) 发现的.

定理 (Γ 函数的余元公式) : 对任意的 $x\in(0,1)$, 有

$$\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin\pi x}.$$

例　由余元公式得 $\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)=\pi$, 故

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

7.8　B 函数

与 Γ 函数密切相关的一个二元函数是 B 函数.

定义: 含参变元积分

$$\mathrm{B}(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x$$

称为 B 函数.

由广义积分那一章最后一节的讨论可知 $\mathrm{B}(\alpha,\beta)$ 函数对任意的 $\alpha>0,\beta>0$ 有定义.

引理:

(1) $\mathrm{B}(x,y)>0$, $\mathrm{B}(1,y)=\dfrac{1}{y}$.

(2) $B(x+1,y)=\dfrac{x}{x+y}\mathrm{B}(x,y)$.

(3) 对取定的 y, $\ln\mathrm{B}(x,y)$ 是 x 的凸函数.

由上述引理, 我们就不难证明 B 函数可以用 Γ 函数来表示:

定理: 对任意的 $x>0,y>0$, 有

$$\mathrm{B}(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

思路　取定 $y>0$, 令

$$f(x)=\frac{\Gamma(x+y)\mathrm{B}(x,y)}{\Gamma(y)}.$$

由上述引理可以验证 $f(x)$ 满足 Γ 函数的三条性质, 又因为这三条性质确定了 Γ 函数, 因此

$$f(x)=\Gamma(x).$$

从而结论得证.

推论 (对称性) : 对任意的 $x>0,y>0$, 有 $\mathrm{B}(x,y)=\mathrm{B}(y,x)$.

7.9　Γ 函数与 B 函数的应用

在数学的世界里, 我们经常能够遇见 Γ 函数与 B 函数的身影. 现在就让我们来简单地看一下它们的应用.

例 1　设 $\alpha>-1,\beta>-1$, 试计算积分

$$I=\int_0^{\frac{\pi}{2}}\sin^\alpha x\cos^\beta x\mathrm{d}x.$$

解 作变量替换 $t=\sin^2 x$, 则 $x=\arcsin\sqrt{t}$, 从而

$$\begin{aligned}\int_0^{\frac{\pi}{2}}\sin^\alpha x\cos^\beta x\mathrm{d}x&=\frac{1}{2}\int_0^{\frac{\pi}{2}}(\sin^2 x)^{\frac{\alpha}{2}}(1-\sin^2 x)^{\frac{\beta}{2}}\mathrm{d}x=\frac{1}{2}\int_0^1 t^{\frac{\alpha-1}{2}}(1-t)^{\frac{\beta-1}{2}}\mathrm{d}t\\&=\frac{1}{2}\mathrm{B}\left(\frac{\alpha+1}{2},\frac{\beta+1}{2}\right)=\frac{1}{2}\frac{\Gamma\left(\frac{\alpha+1}{2}\right)\Gamma\left(\frac{\beta+1}{2}\right)}{\Gamma\left(\frac{\alpha+\beta}{2}+1\right)}.\end{aligned}$$

例 2 由上例与 $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$ 知

$$\int_0^{\frac{\pi}{2}}\sin^\alpha x\mathrm{d}x=\int_0^{\frac{\pi}{2}}\cos^\alpha x\mathrm{d}x=\frac{\sqrt{\pi}}{2}\frac{\Gamma\left(\frac{\alpha+1}{2}\right)}{\Gamma\left(\frac{\alpha+2}{2}\right)}.$$

例 3 设 $|\alpha|<1$, 试计算积分

$$I=\int_0^{\frac{\pi}{2}}(\tan x)^\alpha\mathrm{d}x.$$

解

$$\begin{aligned}I=\int_0^{\frac{\pi}{2}}\sin^\alpha x\cos^{-\alpha}x\mathrm{d}x&=\frac{1}{2}\mathrm{B}\left(\frac{1+\alpha}{2},\frac{1-\alpha}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1+\alpha}{2}\right)\Gamma\left(\frac{1-\alpha}{2}\right)\\&=\frac{1}{2}\Gamma\left(\frac{1+\alpha}{2}\right)\Gamma\left(1-\frac{1+\alpha}{2}\right).\end{aligned}$$

由 Γ 函数的余元公式得

$$I=\int_0^{\frac{\pi}{2}}\sin^\alpha x\cos^{-\alpha}x\mathrm{d}x=\frac{\pi}{\sin\frac{1+\alpha}{2}\pi}=\frac{\pi}{\cos\frac{\alpha}{2}\pi}.$$

7.10 n 维球体的体积

例 $\mathbb{R}^n$ 中半径为 r 的球体体积为

$$V_n(r)=\frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}r^n.$$

解 根据

$$V_1(r)=2r,\quad V_2(r)=\pi r^2,\quad V_3(r)=\frac{4}{3}\pi r^3,$$

我们猜测 n 维球体的体积具有形式:

$$V_n(r)=a_n r^n,$$

其中 a_n 与 r 无关. 显然 $n=1$ 时, 结论成立. 假设结论对 $n-1$ 维球体成立. 对 n 维球体

$$B_n(r)=\{(x_1,\cdots,x_n)\in\mathbb{R}^n:x_1^2+\cdots+x_n^2\leqslant r^2\},$$

其体积为

$$\begin{aligned}
V_n(r)&=\int_{B_n(r)}\mathrm{d}x_1\cdots\mathrm{d}x_n=\int_{-r}^{r}\mathrm{d}x_n\int_{B_{n-1}\left(\sqrt{r^2-x_n^2}\right)}\mathrm{d}x_1\cdots\mathrm{d}x_{n-1}\\
&=\int_{-r}^{r}a_{n-1}(r^2-x_n^2)^{\frac{n-1}{2}}\mathrm{d}x_n=2\int_0^r a_{n-1}(r^2-x_n^2)^{\frac{n-1}{2}}\mathrm{d}x_n\\
&\xlongequal{\underline{x_n=\cos t}}2a_{n-1}\int_{\frac{\pi}{2}}^{0}(r^2\sin^2 t)^{\frac{n-1}{2}}\mathrm{d}(r\cos t)\\
&=2a_{n-1}r^n\int_0^{\frac{\pi}{2}}\sin^n t\mathrm{d}t=\left(2a_{n-1}\int_0^{\frac{\pi}{2}}\sin^n t\mathrm{d}t\right)r^n=a_nr^n.
\end{aligned}$$

此外, 我们还得到了递推关系

$$a_n=2a_{n-1}\int_0^{\frac{\pi}{2}}\sin^n t\mathrm{d}t.$$

利用 $a_1=2$, 我们可以求得各个球体积系数 a_n, 进而可以求得各维度球体的体积

$$V_n(r)=a_nr^n.$$

另外, 由上节例 2 知

$$\frac{a_n}{a_{n-1}}=2\int_0^{\frac{\pi}{2}}\sin^n t\mathrm{d}t=\pi^{1/2}\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n+2}{2}\right)}.$$

由数学归纳法可得

$$a_n=\frac{\pi^{n/2}}{\Gamma\left(\dfrac{n}{2}+1\right)},\quad n=1,2,\cdots.$$

8 微分形式与外微分

8.1 微分形式

如无特别说明, 下面所出现的函数的定义域均为 $\mathbb{R}^n$ 或其子区域 D.

在 $\{\mathrm{d}x_1,\mathrm{d}x_2,\cdots,\mathrm{d}x_n\}$ 里面任选 k 个组成有序元, 记为 $\mathrm{d}x_{i_1}\wedge\mathrm{d}x_{i_2}\wedge\cdots\wedge\mathrm{d}x_{i_k}$, 这里 $i_1,i_2,\cdots,i_k\in\{1,2,\cdots,n\}$, 规定:

(1) $\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_r}\wedge\mathrm{d}x_{i_{r+1}}\wedge\cdots\wedge\mathrm{d}x_{i_k}=-\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_{r+1}}\wedge\mathrm{d}x_{i_r}\wedge\cdots\wedge\mathrm{d}x_{i_k}$(**相邻反对称性**);

(2) 若 $i_1,i_2,\cdots,i_k$ 中有两个相同, 则 $\mathrm{d}x_{i_1}\wedge\mathrm{d}x_{i_2}\wedge\cdots\wedge\mathrm{d}x_{i_k}=0$.

由 (2) 可知, 共有 $\binom{n}{k}$ 个非零有序元

$$\mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \cdots \wedge \mathrm{d}x_{i_k}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n.$$

称它们为基本元. 基本元的线性组合

$$\omega = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} g_{i_1, i_2, \cdots, i_k}(\boldsymbol{x}) \mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \cdots \wedge \mathrm{d}x_{i_k}$$

称为 $\mathbb{R}^n$ 中的 $\boldsymbol{k}$ **次微分形式**, 简称 k-形式. 我们把 $\mathbb{R}^n$ 或其子区域 D 上的连续可微函数称为 0-形式. $\mathbb{R}^n$ 中的全体 k-形式所成集合记为 Λ^k. 易知 Λ^k 是一个 $\binom{n}{k}$ 维线性空间. 又把 $\mathbb{R}^n$ 中的全体微分形式所成集合记为 Λ, 即

$$\Lambda = \Lambda^0 + \Lambda^1 + \cdots + \Lambda^n.$$

易知 Λ 是一个 2^n 维线性空间.

8.2 外积运算

对于基本元

$$\mathrm{d}x_I = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \in \Lambda^p, \quad \mathrm{d}x_J = \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \in \Lambda^q, \qquad (p, q \geqslant 1)$$

定义 $\mathrm{d}x_I$ 与 $\mathrm{d}x_J$ 的外积为

$$\mathrm{d}x_I \wedge \mathrm{d}x_J = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \in \Lambda^{p+q}.$$

对于一般的微分形式

$$\omega = \sum_I g_I(\boldsymbol{x}) \mathrm{d}x_I \in \Lambda^p, \quad \eta = \sum_J h_J(\boldsymbol{x}) \mathrm{d}x_J \in \Lambda^q,$$

定义它们的外积为

$$\omega \wedge \eta = \sum_I \sum_J g_I(\boldsymbol{x}) h_J(\boldsymbol{x}) \mathrm{d}x_I \wedge \mathrm{d}x_J \in \Lambda^{p+q}.$$

对于 0-形式 f 与 p-形式 $\omega = \sum\limits_I g_I(\boldsymbol{x}) \mathrm{d}x_I$, 定义它们的外积为

$$f \wedge \omega = \omega \wedge f = f\omega = \sum_I f(\boldsymbol{x}) g_I(\boldsymbol{x}) \mathrm{d}x_I \in \Lambda^p.$$

性质:

(1) (特殊交换律) 设 $\omega \in \Lambda^p$, $\eta \in \Lambda^q$, 则

$$\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega.$$

(2) (分配律) 设 $\omega, \eta, \sigma \in \Lambda$, 则

$$\begin{aligned}(\omega + \eta) \wedge \sigma &= \omega \wedge \sigma + \eta \wedge \sigma, \\ \omega \wedge (\eta + \sigma) &= \omega \wedge \eta + \omega \wedge \sigma.\end{aligned}$$

(3) (结合律) 设 $\omega, \eta, \sigma \in \Lambda$, 则

$$(\omega \wedge \eta) \wedge \sigma = \omega \wedge (\eta \wedge \sigma).$$

此外, 由特殊分配律可知: 若 p 为奇数, 则对于 $\omega \in \Lambda^p$, 有 $\omega \wedge \omega = 0$.

8.3 外微分算子

对于 $D \subseteq \mathbb{R}^n$ 上的连续可微函数 f, 其全微分为

$$\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i.$$

下面我们将其推广到一般的 k-形式上 (称为**外微分运算**, 相应的算子 d 称为**外微分算子**).

对于 k-形式

$$\omega = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} g_{i_1, \cdots, i_k}(\boldsymbol{x}) \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k},$$

定义外微分算子 d 在其上的作用为

$$\begin{aligned}\mathrm{d}\omega &= \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathrm{d}g_{i_1, \cdots, i_k}(\boldsymbol{x}) \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} \\ &= \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \sum_{i=1}^{n} \frac{\partial g_{i_1, \cdots, i_k}}{\partial x_i}(\boldsymbol{x}) \mathrm{d}x_i \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}.\end{aligned}$$

对于 Λ 中的任意元素 $\omega = \omega_0 + \omega_1 + \cdots + \omega_n$, 其中 $\omega_i \in \Lambda^i$, 定义

$$\mathrm{d}\omega = \mathrm{d}\omega_0 + \mathrm{d}\omega_1 + \cdots + \mathrm{d}\omega_n.$$

容易验证, 如此定义的外微分运算

$$\begin{aligned}\mathrm{d} : \Lambda &\longrightarrow \Lambda \\ \omega &\longmapsto \mathrm{d}\omega\end{aligned}$$

具有线性性, 也即

$$\mathrm{d}(\alpha\omega+\beta\eta)=\alpha\mathrm{d}\omega+\beta\mathrm{d}\eta,$$

其中 α,β 为常数, $\omega,\eta\in\Lambda$.

性质:

(1) (外微分的乘法法则) 设 $\omega\in\Lambda^p$, $\eta\in\Lambda^q$, 则

$$\mathrm{d}(\omega\wedge\eta)=\mathrm{d}\omega\wedge\eta+(-1)^p\omega\wedge\mathrm{d}\eta.$$

(2) (外微分算子 d 的幂零性) 对任意的 $\omega\in\Lambda$, 有

$$\mathrm{d}^2\omega=\mathrm{d}(\mathrm{d}\omega)=0.$$

8.4 微分形式与变量替换

例 证明: 对 $\mathbb{R}^n$ 中的 n 个 1-形式 $\omega_i=\sum\limits_{j=1}^{n}a_{ij}(\boldsymbol{x})\mathrm{d}x_j$ $(i=1,\cdots,n)$, 即

$$\begin{cases}\omega_1=a_{11}(\boldsymbol{x})\mathrm{d}x_1+\cdots+a_{1n}(\boldsymbol{x})\mathrm{d}x_n,\\ \qquad\qquad\cdots\\ \omega_n=a_{n1}(\boldsymbol{x})\mathrm{d}x_1+\cdots+a_{nn}(\boldsymbol{x})\mathrm{d}x_n,\end{cases}$$

有

$$\omega_1\wedge\cdots\wedge\omega_n=\det(a_{ij}(\boldsymbol{x}))\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n.$$

证明 易知

$$\omega_1\wedge\cdots\wedge\omega_n=\sum_{i_1\cdots i_n}a_{1i_1}(\boldsymbol{x})\cdots a_{ni_n}(\boldsymbol{x})\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_n},$$

其中 $\sum\limits_{i_1\cdots i_n}$ 表示对所有 $12\cdots n$ 的排列 $i_1\cdots i_n$ 求和 (因为 $i_1,i_2,\cdots,i_n$ 必须两两不同所得项才非零). 令

$$\varepsilon^{i_1\cdots i_n}=\begin{cases}-1, & i_1\cdots i_n\text{为奇排列},\\ 1, & i_1\cdots i_n\text{为偶排列}.\end{cases}$$

则

$$\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_n}=\varepsilon^{i_1\cdots i_n}\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n.$$

因此

$$\begin{aligned}\omega_1\wedge\cdots\wedge\omega_n &= \sum_{i_1\cdots i_n}\varepsilon^{i_1\cdots i_n}a_{1i_1}(\boldsymbol{x})\cdots a_{ni_n}(\boldsymbol{x})\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n\\ &= \left(\sum_{i_1\cdots i_n}\varepsilon^{i_1\cdots i_n}a_{1i_1}(\boldsymbol{x})\cdots a_{ni_n}(\boldsymbol{x})\right)\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n.\end{aligned}$$

也即

$$\omega_1\wedge\cdots\wedge\omega_n = \det(a_{ij}(\boldsymbol{x}))\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n.$$

注: 证明的最后我们用到了结论: $\varepsilon^{i_1\cdots i_n} = (-1)^{\tau(i_1\cdots i_n)}$, 其中 $\tau(i_1\cdots i_n)$ 为排列 $i_1\cdots i_n$ 的逆序数. 这是因为

(1) 一次对换等效于奇数次邻换.

(2) 设 $i_1i_2\cdots i_n$ 为 $12\cdots n$ 的一个排列, 且 $i_1i_2\cdots i_n$ 可以经过 k 次对换变为 $12\cdots n$, 则 $i_1i_2\cdots i_n$ 的奇偶性与 k 的奇偶性相同.

综合 (1)(2) 即得: $i_1\cdots i_n$ 的奇偶性与从 $i_1\cdots i_n$ 变到 $12\cdots n$ 所需邻换次数的奇偶性相同.

例 设有数值函数

$$y_1 = y_1(x_1,\cdots,x_n),\cdots,y_n = y_n(x_1,\cdots,x_n),$$

则

$$\mathrm{d}y_1\wedge\cdots\wedge\mathrm{d}y_n = \frac{\partial(y_1,\cdots,y_n)}{\partial(x_1,\cdots,x_n)}\mathrm{d}x_1\wedge\cdots\wedge\mathrm{d}x_n.$$

证明 因为 $\mathrm{d}y_i = \sum\limits_{j=1}^{n}\dfrac{\partial y_i}{\partial x_j}\mathrm{d}x_j$, 从而由上例即可证得所求结论.

9 曲线积分

9.1 第一型曲线积分

定理: 设曲线 $\gamma:\boldsymbol{r}(t)$ 在 $[\alpha,\beta]$ 上连续可微, 则 γ 可求长, 且弧长为

$$l(\gamma) = \int_\alpha^\beta|\boldsymbol{r}'(t)|\mathrm{d}t.$$

定义: 设 $f(x,y)$ 为定义在平面可求长曲线 L 上的函数, 分割 T 将 L 划分为 n 小段 L_i, 且 L_i 对应的弧长为 Δs_i, 分割 T 的细度 $\|T\| = \max\limits_i\{\Delta s_i\}$. 从每个小弧段上任取一点 (ξ_i,η_i), 若极限

$$J = \lim_{\|T\|\to 0}\sum_{i=1}^{n}f(\xi_i,\eta_i)\Delta s_i$$

存在, 且与分割 T 和点 (ξ_i,η_i) 的取法无关, 则称 J 为函数 $f(x,y)$ 在曲线 L 上的第一型曲线积分,

记作

$$\int_L f(x,y)\,\mathrm{d}s.$$

定理: 设曲线 $L:\boldsymbol{r}=\boldsymbol{r}(t)=(x(t),y(t),z(t)),\ t\in[\alpha,\beta]$ 连续可微, 函数 f 在 L 上连续, 则 f 沿 L 的第一型曲线积分存在, 且

$$\int_L f(x,y,z)\mathrm{d}s=\int_a^\beta f(x(t),y(t),z(t))\sqrt{x'^2(t)+y'^2(t)+z'^2(t)}\,\mathrm{d}t.$$

9.2 第二型曲线积分

定义: 设 L 为空间中可求长的有向曲线, 函数 $P(x,y,z)$ 在 L 上连续, 分割 T 将曲线 L 划分为若干小弧段 $L_1,\cdots,L_n$, 分割点为 $M_i(x_i,y_i,z_i)$, $\Delta x_i=x_i-x_{i-1}$. 在每段 L_i 上任取一点 (ξ_i,η_i,σ_i), 若极限

$$J=\lim_{\|T\|\to 0}\sum_{i=1}^n P(\xi_i,\eta_i,\sigma_i)\Delta x_i$$

存在且与分割 T 和点 (ξ_i,η_i,σ_i) 的取法无关, 则称 J 为函数 $P(x,y,z)$ 沿有向曲线 L 的**对 x 坐标**的曲线积分, 记作 $\int_L P(x,y,z)\mathrm{d}x$, 简记为 $\int_L P\mathrm{d}x$. 即

$$\int_L P\mathrm{d}x=\lim_{\|T\|\to 0}\sum_{i=1}^n P(\xi_i,\eta_i,\sigma_i)\Delta x_i.$$

类似地, 可定义

$$\int_L Q\mathrm{d}y=\lim_{\|T\|\to 0}\sum_{i=1}^n Q(\xi_i,\eta_i,\sigma_i)\Delta y_i,$$
$$\int_L R\mathrm{d}z=\lim_{\|T\|\to 0}\sum_{i=1}^n R(\xi_i,\eta_i,\sigma_i)\Delta z_i.$$

分别称为 $Q(x,y,z)$ 沿曲线 L **对 y 坐标**的曲线积分、$R(x,y,z)$ 沿曲线 L **对 z 坐标**的曲线积分. 以上这些对坐标的曲线积分, 统称为**第二型曲线积分**. 此外, 我们约定

$$\int_L P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z=\int_L P\mathrm{d}x+\int_L Q\mathrm{d}y+\int_L R\mathrm{d}z.$$

第二型曲线积分可以写成向量形式

$$\int_L \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s},$$

其中 $\boldsymbol{F}(x,y,z)=(P(x,y,z),Q(x,y,z),R(x,y,z)),\mathrm{d}\boldsymbol{s}=(\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)$.

对于沿闭有向曲线的积分, 常常把积分号写为 $\oint$, 例如

$$\oint_L P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z\quad 或\quad \oint_L \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s}.$$

注: 平面曲线是空间曲线的特殊情形. 按定义, 沿这样的曲线显然有

$$\int_L R\mathrm{d}z = 0.$$

因此, 对于平面曲线 L, 我们只需考虑 $\int_L P\mathrm{d}x + Q\mathrm{d}y$.

性质: (假定各式右端的积分存在)

(1) 线性性: $\int_L (\alpha f + \beta g)\mathrm{d}x = \alpha\int_L f\mathrm{d}x + \beta\int_L g\mathrm{d}x$, 其中 $f = f(x,y,z), g = g(x,y,z)$.

(2) 可加性: $\int_L f\mathrm{d}x = \int_{L_1} f\mathrm{d}x + \int_{L_2} f\mathrm{d}x$, 其中 L 为有向线段 L_1 与 L_2 首尾相接而成.

(3) 有向性: $\int_{L^-} f\mathrm{d}x = -\int_L f\mathrm{d}x$, 其中 L^- 表示将有向曲线 L 反向后得到的有向曲线.

定理: 设 $L : \boldsymbol{r} = \boldsymbol{r}(t) = (x(t), y(t), z(t))$, $t \in [\alpha, \beta]$ 为空间连续可微有向曲线, 函数 P, Q, R 在曲线 L 上连续, 则

$$\int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = \int_\alpha^\beta [P(x(t),y(t),z(t))x'(t) + Q(x(t),y(t),z(t))y'(t) + R(x(t),y(t),z(t))z'(t)]\mathrm{d}t.$$

9.3 两类曲线积分的联系

设有向曲线 $L : \boldsymbol{r} = \boldsymbol{r}(t)$ 沿正向的单位切向量为 $\boldsymbol{t} = (\cos\alpha, \cos\beta, \cos\gamma)$, 其中 $\cos\alpha, \cos\beta, \cos\gamma$ 分别为 $\boldsymbol{t}$ 与 x 轴, y 轴, z 轴正向夹角的方向余弦, 则两类曲线积分有关系

$$\int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = \int_L (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,\mathrm{d}s.$$

写成向量形式即为

$$\int_L \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s} = \int_L \boldsymbol{F}\cdot\boldsymbol{t}\,\mathrm{d}s.$$

注: 直观推导一下上述结论.

$$\begin{aligned}
\int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z &= \int_L (Px' + Qy' + Rz')\,\mathrm{d}t \\
&= \int_L (P\cos\alpha + Q\cos\beta + R\cos\gamma)\sqrt{x'^2(t) + y'^2(t) + z'^2(t)}\,\mathrm{d}t \\
&= \int_L (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,\mathrm{d}s.
\end{aligned}$$

这里我们用到了单位切向量 $\boldsymbol{t}(t)=\dfrac{\boldsymbol{r}(t)}{|\boldsymbol{r}(t)|}$ 的分量为

$$\begin{aligned}\cos\alpha&=\frac{x'(t)}{\sqrt{x'^2(t)+y'^2(t)+z'^2(t)}},\\ \cos\beta&=\frac{y'(t)}{\sqrt{x'^2(t)+y'^2(t)+z'^2(t)}},\\ \cos\gamma&=\frac{z'(t)}{\sqrt{x'^2(t)+y'^2(t)+z'^2(t)}}.\end{aligned}$$

10 曲面积分

10.1 第一型曲面积分

定理: 设映射 $\boldsymbol{r}=\boldsymbol{r}(u,v)$, $(u,v)\in D$ 是双射且 $\boldsymbol{r}_u\times\boldsymbol{r}_v\neq\boldsymbol{0}$, 则曲面 $S:\boldsymbol{r}=\boldsymbol{r}(u,v)$ 的面积为

$$A=\iint\limits_D\sqrt{EG-F^2}\mathrm{d}u\mathrm{d}v.$$

定义: 设 Σ 为空间可求面积的曲面, f 为 Σ 上的函数. 作分割 T 将曲面 Σ 划分为小曲面块 Σ_i, 对应的面积为 ΔS_i, 分割 T 的细度 $\|T\|=\max\limits_i\{\operatorname{diam}S_i\}$. 从每个 Σ_i 上任取一点 (ξ_i,η_i,σ_i), 若极限

$$J=\lim_{\|T\|\to0}\sum_{i=1}^n f(\xi_i,\eta_i,\sigma_i)\Delta S_i$$

存在且与分割 T 和点 (ξ_i,η_i,σ_i) 的取法无关, 则称此极限为函数 f 沿曲面 Σ 的第一型曲面积分, 记作

$$\iint\limits_\Sigma f(x,y,z)\mathrm{d}S.$$

定理: 设有正则曲面 $\Sigma:\boldsymbol{r}=\boldsymbol{r}(u,v)=(x(u,v),y(u,v),z(u,v))$, $(u,v)\in D$, 函数 f 在曲面 Σ 上连续, 则 f 沿曲面 Σ 的第一型曲面积分存在, 且

$$\iint\limits_\Sigma f(x,y,z)\mathrm{d}S=\iint\limits_D f(x(u,v),y(u,v),z(u,v))\sqrt{EG-F^2}\mathrm{d}u\mathrm{d}v.$$

10.2 第二型曲面积分

定义: 设 Σ 为可定向曲面, 指定侧的单位法向量 $\boldsymbol{n}=(\cos\alpha,\cos\beta,\cos\gamma)$, $P(x,y,z)$ 为定义在 Σ 上的函数, 若第一型曲面积分

$$\iint\limits_\Sigma P(x,y,z)\cos\alpha(x,y,z)\mathrm{d}S=\lim_{\|T\|\to0}P(\xi_i,\eta_i,\sigma_i)\cos\alpha(\xi_i,\eta_i,\sigma_i)\Delta S_i \tag{1}$$

存在, 则称该积分为函数 P 沿曲面 Σ 正侧**对 yz 坐标**的曲面积分, 记作

$$\iint\limits_{+\Sigma} P(x,y,z)\mathrm{d}y\wedge\mathrm{d}z.$$

当不至于引起混淆时, “+Σ” 可以简记为“Σ”.

注: (1) 式就是有向小曲面块 S_i 在 yz 平面上投影的有向面积.

类似地, 我们定义

$$\iint\limits_{\Sigma} Q(x,y,z)\mathrm{d}z\wedge\mathrm{d}x=\iint\limits_{\Sigma} Q(x,y,z)\cos\beta(x,y,z)\mathrm{d}S,$$
$$\iint\limits_{\Sigma} R(x,y,z)\mathrm{d}x\wedge\mathrm{d}y=\iint\limits_{\Sigma} R(x,y,z)\cos\gamma(x,y,z)\mathrm{d}S.$$

它们分别称为 $Q(x,y,z)$ 沿曲面 Σ 正侧对 zx 坐标的曲面积分、$R(x,y,z)$ 沿曲面 Σ 正侧对 xy 坐标的曲面积分. 以上这些对坐标的曲面积分, 统称为第二型曲面积分. 为书写简便, 我们常常将“∧”省去不写, 例如

$$\iint\limits_{\Sigma} P(x,y,z)\mathrm{d}y\mathrm{d}z.$$

此外, 我们约定

$$\iint\limits_{\Sigma} P\mathrm{d}y\wedge\mathrm{d}z+Q\mathrm{d}z\wedge\mathrm{d}x+R\mathrm{d}x\wedge\mathrm{d}y=\iint\limits_{\Sigma} P\mathrm{d}y\wedge\mathrm{d}z+\iint\limits_{\Sigma} Q\mathrm{d}z\wedge\mathrm{d}x+\iint\limits_{\Sigma} R\mathrm{d}x\wedge\mathrm{d}y,$$

也即

$$\iint\limits_{\Sigma} P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y=\iint\limits_{\Sigma} P\mathrm{d}y\mathrm{d}z+\iint\limits_{\Sigma} Q\mathrm{d}z\mathrm{d}x+\iint\limits_{\Sigma} R\mathrm{d}x\mathrm{d}y.$$

注: 由第二型曲面积分的定义可得

$$\iint\limits_{\Sigma} P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y=\iint\limits_{\Sigma}\left(P\cos\alpha+Q\cos\beta+R\cos\gamma\right)\mathrm{d}S.$$

记 $\mathrm{d}\boldsymbol{S}=(\mathrm{d}y\wedge\mathrm{d}z,\mathrm{d}z\wedge\mathrm{d}x,\mathrm{d}x\wedge\mathrm{d}y)=(\mathrm{d}y\mathrm{d}z,\mathrm{d}z\mathrm{d}x,\mathrm{d}x\mathrm{d}y)$, 将上式改写为向量形式即为

$$\iint\limits_{\Sigma}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}=\iint\limits_{\Sigma}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S.$$

其中

$$\boldsymbol{F}(x,y,z)=(P(x,y,z),Q(x,y,z),R(x,y,z)),$$
$$\boldsymbol{n}(x,y,z)=(\cos\alpha(x,y,z),\cos\beta(x,y,z),\cos\gamma(x,y,z)).$$

若曲面 Σ 的单位正法向为

$$\boldsymbol{n} = (\cos\alpha, \cos\beta, \cos\gamma),$$

则 Σ 的负侧 $-\Sigma$ 的单位法向就是

$$-\boldsymbol{n} = (-\cos\alpha, -\cos\beta, -\cos\gamma).$$

因此

$$\iint\limits_{-\Sigma} P\mathrm{d}y \wedge \mathrm{d}z = \iint\limits_{\Sigma} P(x,y,z)(-\cos\alpha(x,y,z))\mathrm{d}S = -\iint\limits_{+\Sigma} P\mathrm{d}y \wedge \mathrm{d}z.$$

记号 $\mathrm{d}y \wedge \mathrm{d}z$ 表示 yOz 平面上的有向面积元. 我们约定以 $\boldsymbol{i}$ 作为 $\mathrm{d}y \wedge \mathrm{d}z$ 的单位正法向, 还约定记号 $\mathrm{d}z \wedge \mathrm{d}y$ 表示以 $-\boldsymbol{i}$ 为单位正法向的同一块面积元, 因此

$$\mathrm{d}z \wedge \mathrm{d}y = -\mathrm{d}y \wedge \mathrm{d}z.$$

同理

$$\mathrm{d}x \wedge \mathrm{d}z = -\mathrm{d}z \wedge \mathrm{d}x,$$
$$\mathrm{d}y \wedge \mathrm{d}x = -\mathrm{d}x \wedge \mathrm{d}y.$$

注: 通常以 $\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$ 表示 $\mathbb{R}^3$ 中正的体积元, 于是

$$\iiint\limits_{\Omega} g(x,y,z)\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \iiint\limits_{\Omega} g(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z,$$
$$\iiint\limits_{\Omega} g(x,y,z)\,\mathrm{d}y \wedge \mathrm{d}x \wedge \mathrm{d}z = -\iiint\limits_{\Omega} g(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

这里的 $\iiint\limits_{\Omega} g(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z$ 表示通常的三重积分.

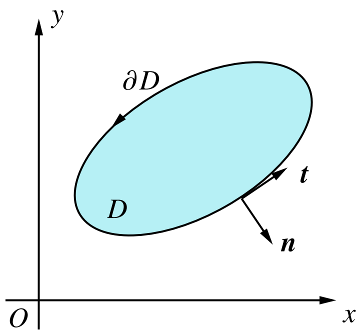
10.3 Green 公式, Gauss 公式, Stokes 公式

对于平面区域 D, 我们给它的边界 ∂D 规定一个正向: 如果一个人沿 ∂D 的这个方向行走时, 区域 D 总是在他左边. 这个定向也称为 D 的**诱导定向**, 带有这样定向的 ∂D 称为 D 的**正向边界**.

定理 (Green 公式): 设 D 为平面上由光滑或分段光滑的简单闭曲线所围的闭区域. 如果函数 $P(x,y), Q(x,y)$ 在 D 上具有连续偏导数, 那么

$$\int_{\partial D} P\mathrm{d}x + Q\mathrm{d}y = \iint\limits_{D} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y,$$

其中 ∂D 取正向, 即诱导定向.



设正向 ∂D 的单位切向量为 $\boldsymbol{t}$, 单位外法向量为 $\boldsymbol{n}$, 则得到 Green 公式的另一种常用表示

$$\begin{aligned}\iint\limits_D \left(\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)\mathrm{d}x\mathrm{d}y &= \int_{\partial D} F\mathrm{d}y - G\mathrm{d}x = \int_{\partial D}[-G\cos(\boldsymbol{t},x)+F\cos(\boldsymbol{t},y)]\mathrm{d}s \\ &= \int_{\partial D}[F\cos(\boldsymbol{n},x)+G\cos(\boldsymbol{n},y)]\mathrm{d}s,\end{aligned}$$

注意 Green 公式并不要求区域是单连通的.

对于封闭曲面 Ω, 我们规定 $\partial\Omega$ 的外侧为其正向, 称为 Ω 的**诱导定向**.

定理 (Gauss 公式): 设 Ω 是 $\mathbb{R}^3$ 中由光滑或分片光滑的双侧封闭曲面所围成的闭区域, 函数 $P(x,y,z),Q(x,y,z)$ 和 $R(x,y,z)$ 在 Ω 上具有连续偏导数, 则

$$\iint\limits_{\partial\Omega} P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y = \iiint\limits_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z,$$

这里 $\partial\Omega$ 的定向为外侧 (即取诱导定向) .

利用算子 ∇, 我们可以将 Gauss 公式可以写为

$$\iint\limits_{\partial\Omega}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S} = \iiint\limits_{\Omega}\nabla\cdot\boldsymbol{F}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

取定曲面 Σ 的侧后, $\partial\Sigma$ 与 Σ 满足右手法则, 则称 $\partial\Sigma$ 的这个定向为 Σ 的**诱导定向**.

定理 (Stokes 公式): 若函数 $P(x,y,z),Q(x,y,z),R(x,y,z)$ 在曲面 Σ 及其边界上具有连续偏导数, 则

$$\begin{aligned}&\int_{\partial\Sigma} P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z \\ =&\iint\limits_{\Sigma}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}y\mathrm{d}z+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}z\mathrm{d}x+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y \\ =&\iint\limits_{\Sigma}\left[\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma\right]\mathrm{d}S,\end{aligned}$$

其中 $\partial\Sigma$ 取诱导定向.

利用算子 ∇, 我们可以将 Stokes 公式可以写为

$$\int_{\partial\Sigma} \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{s} = \iint\limits_{\Sigma} \nabla\times\boldsymbol{F}\,\mathrm{d}\boldsymbol{S}.$$

利用行列式记号, 可以将 Stokes 公式写成

$$\int_{\partial\Sigma} P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = \iint\limits_{\Sigma} \begin{vmatrix} \mathrm{d}y\mathrm{d}z & \mathrm{d}z\mathrm{d}x & \mathrm{d}x\mathrm{d}y \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \iint\limits_{\Sigma} \begin{vmatrix} \cos\alpha & \cos\beta & \cos\gamma \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \mathrm{d}S.$$

将上面的行列式按第一行展开即可得到原本的 Stokes 公式.

10.4 Green 公式与平面区域的面积, Gauss 公式与空间区域的体积

设 D 是 $\mathbb{R}^2$ 中的闭区域, 令 $Q(x,y)=x$, 由 Green 公式得

$$\int_{\partial D} x\,\mathrm{d}y = \iint\limits_{D} \mathrm{d}x\mathrm{d}y = S(D).$$

同理

$$-\int_{\partial D} y\,\mathrm{d}x = \iint\limits_{D} \mathrm{d}x\mathrm{d}y = S(D).$$

因此

$$S(D) = \frac{1}{2}\int_{\partial D} x\,\mathrm{d}y - y\,\mathrm{d}x.$$

当 ∂D 有参数表示 $\boldsymbol{r}(t) = (x(t), y(t))$, $t\in[a,b]$ 时, D 的面积就可表示为

$$S(D) = \frac{1}{2}\int_a^b \left(x\frac{\mathrm{d}y}{\mathrm{d}t} - y\frac{\mathrm{d}x}{\mathrm{d}t}\right)\mathrm{d}t.$$

设 Ω 是 $\mathbb{R}^3$ 中的闭区域, 令 $P(x,y,z)=x$, 则由 Gauss 公式得

$$\iint\limits_{\partial\Omega} x\,\mathrm{d}y\mathrm{d}z = \iiint\limits_{\Omega} \mathrm{d}x\mathrm{d}y\mathrm{d}z = V(\Omega).$$

同理

$$\iint\limits_{\partial\Omega} y\,\mathrm{d}z\mathrm{d}x = \iint\limits_{\partial\Omega} z\,\mathrm{d}x\mathrm{d}y = V(\Omega).$$

因此

$$V(\Omega) = \frac{1}{3}\iint\limits_{\partial\Omega} x\,\mathrm{d}y\mathrm{d}z + y\,\mathrm{d}z\mathrm{d}x + z\,\mathrm{d}x\mathrm{d}y.$$

10.5 用微分形式重新表述 Green 公式, Gauss 公式, Stokes 公式

我们复习一下前面所学: 根据 §10.2 的论述, $\mathrm{d}x\wedge\mathrm{d}y$ 可以看成 $\mathbb{R}^2$ 中的正向面积元, $\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z$ 可以看成 $\mathbb{R}^3$ 中的正向体积元, 进而

$$\iint\limits_D f(x,y)\,\mathrm{d}x\wedge\mathrm{d}y=\iint\limits_D f(x,y)\,\mathrm{d}x\mathrm{d}y,$$

$$\iiint\limits_\Omega g(x,y,z)\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z=\iiint\limits_\Omega g(x,y,z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

此外, $\mathrm{d}y\wedge\mathrm{d}z$ 可以看成 $\mathbb{R}^3$ 中 yOz 平面上的正向面积元, 进而

$$\iint\limits_\Sigma h(x,y,z)\,\mathrm{d}y\wedge\mathrm{d}z=\iint\limits_\Sigma h(x,y,z)\,\mathrm{d}y\mathrm{d}z.$$

下面我们将 Green 公式, Gauss 公式, Stokes 公式完整地表示出来.

Green 公式:

$$\int_{\partial D}P\mathrm{d}x+Q\mathrm{d}y=\iint\limits_D\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\wedge\mathrm{d}y,$$

Gauss 公式:

$$\iint\limits_{\partial\Omega}P\mathrm{d}y\wedge\mathrm{d}z+Q\mathrm{d}z\wedge\mathrm{d}x+R\mathrm{d}x\wedge\mathrm{d}y=\iiint\limits_\Omega\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z.$$

Stokes 公式:

$$\int_{\partial\Sigma}P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z=\iint\limits_\Sigma\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}y\wedge\mathrm{d}z+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}z\wedge\mathrm{d}x+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\wedge\mathrm{d}y.$$

其中 D 是平面区域, Ω 是 $\mathbb{R}^3$ 中的区域, Σ 是 $\mathbb{R}^3$ 中的曲面, 均取诱导定向.

分别对上述三个公式左端的微分形式用外微分算子作用一下, 所得结果恰好为右端对应的微分形式. 因此它们 (实际上还有 Newton-Leibniz 公式) 可以统一写成 (不论维数如何, 只写一重积分号)

$$\int_{\partial D}\omega=\int_D\mathrm{d}\omega\,.$$

这类公式统称为 **Stokes 公式**. 它说明了高次的微分形式 $\mathrm{d}\omega$ 在给定区域上的积分等于低一次的微分形式 ω 在低一维的区域边界上的积分. Stokes 公式是数学分析中最精彩的结论之一, 我们以后还会看到它的广泛应用!

10.6 曲线积分与路径无关的条件

定义: 设 D 为 $\mathbb{R}^2$ 中的区域, 若 D 中任意一条闭曲线在 D 内都能连续收缩成一个点, 则称 D 是**单连通的**.

定理: 设 D 为 $\mathbb{R}^2$ 中的单连通区域, $P(x,y), Q(x,y)$ 在 D 内连续可微, 则下列各条等价:

(1) $\displaystyle\oint_L P\mathrm{d}x + Q\mathrm{d}y = 0$.

(2) $\displaystyle\int_L P\mathrm{d}x + Q\mathrm{d}y$ 与路径无关.

(3) 存在连续可微函数 $U(x,y)$, 使得 $\mathrm{d}U = P\mathrm{d}x + Q\mathrm{d}y$.

(4) $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$.

注: 上述 (1)(2)(3) 的等价不需要单连通这一条件.

定义: 设 G 为 $\mathbb{R}^3$ 中的区域, 若 G 中任意一条闭曲线在 G 内都能连续收缩成一个点, 则称 G 是**单连通的**.

例 (1) 开球体是单连通的. (2) 开球体内部抠掉一个球形小洞之后仍然是单连通的. (3) 开球体上打了一个贯通的圆柱形孔洞之后就不是单连通的了.

定理: 设 G 为 $\mathbb{R}^3$ 中的单连通区域, $P(x,y,z), Q(x,y,z), R(x,y,z)$ 在 G 内连续可微, 则下列各条等价:

(1) $\displaystyle\oint_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = 0$.

(2) $\displaystyle\int_L P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z$ 与路径无关.

(3) 存在连续可微函数 $U(x,y,z)$, 使得 $\mathrm{d}U = P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z$.

(4) $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$, $\dfrac{\partial Q}{\partial z} = \dfrac{\partial R}{\partial y}$, $\dfrac{\partial R}{\partial x} = \dfrac{\partial P}{\partial z}$.

注: 上述 (1)(2)(3) 的等价不需要单连通这一条件.

定义: 设 $\omega \in \Lambda^p$ 为一个 p-形式. 若

$$\mathrm{d}\omega = 0,$$

则称 ω 为一个**闭形式**. 若存在 $(p-1)$-形式 θ, 使得

$$\omega = \mathrm{d}\theta,$$

则称 ω 为一个**恰当形式**.

由于 $\mathrm{d}(\mathrm{d}\theta) = 0$, 故恰当形式必为闭形式.

采用微分形式的术语, 曲线积分与路径无关的条件可以陈述如下.

定理: 设 G 为 $\mathbb{R}^2$ 或 $\mathbb{R}^3$ 中的单连通区域, ω 为 G 上连续可微的 1-形式, 则下列各条等价:

(1) $\displaystyle\oint_L \omega = 0$.

(2) $\int_L \omega$ 与路径无关.

(3) ω 为恰当形式.

(4) ω 为闭形式.

注: 上述 (1)(2)(3) 的等价不需要单连通这一条件.

10.7 关于单连通性的一点注记

我们在 §10.6 中所定义的单连通性实际上是一维单连通性 (注意到我们考虑的是曲线积分). 对于区域的连通性, 我们有多种方式进行定义, 例如:

定义: 若区域 G 内任一闭曲线都能在 G 中连续收缩到 G 内一点, 则称区域 G 是**一维单连通**的.

定义: 若区域 G 内任一闭曲面都能在 G 中连续收缩到 G 内一点, 则称区域 G 是**二维单连通**的.

11 场论介绍

场 (field) 是最重要的物理概念之一. 数学中的场论, 对各种各样的物理场作抽象概括, 进行定性与定量的研究.

如果空间某区域 Ω 中的每一点在时刻 t 都对应着一个确定的数量 $f(x, y, z, t)$(向量 $\boldsymbol{f}(x, y, z, t)$), 那么我们就说在这区域内定义了一个数量场 (向量场). 若一个场不随时间变化而变化, 则称该场为稳定场, 否则称为不稳定场. 这里我们只考虑稳定场.

例如, 热力学中的温度场 $T(x, y, z)$ 与连续体力学中的密度场 $\rho(x, y, z)$ 都是数量场, 而引力场 $\boldsymbol{F}(x, y, z)$ 与电场 $\boldsymbol{E}(x, y, z)$ 都是向量场.

本章我们主要介绍的三个场是: 数量场 f 的梯度场 $\mathbf{grad}\, f = \nabla f$、向量场 $\boldsymbol{F}$ 的散度场 $\operatorname{div} \boldsymbol{F} = \nabla \cdot \boldsymbol{F}$ 与旋度场 $\mathbf{rot}\, \boldsymbol{F} = \nabla \times \boldsymbol{F}$.

容易看出, 所谓的数量场与向量场, 不过是数量值函数与向量值函数的另一种说法而已. 在本章中, 我们总假定所讨论的场连续可微足够多次.

11.1 方向导数与梯度

设在空间区域 Ω 中定义了一个数量场 $f(M) = f(x, y, z)$. 设单位向量 $\boldsymbol{v} = (\cos\alpha, \cos\beta, \cos\gamma)$ 是任意一个方向, 那么 f 在某点 M_0 沿方向 $\boldsymbol{v}$ 的方向导数为

$$\frac{\partial f}{\partial \boldsymbol{v}}(M_0) = \nabla f(M_0) \cdot \boldsymbol{v}.$$

若把向量 $\nabla f(M_0)$ 与 $\boldsymbol{v}$ 之间的夹角记为 θ, 那么

$$\frac{\partial f}{\partial \boldsymbol{v}}(M_0) = \|\nabla f(M_0)\| \cos\theta.$$

由此可以看出, 当 $\theta = 0$ 时, 方向导数有最大值, 且最大值为

$$\|\nabla f(M_0)\|.$$

定义: 设有定义在区域 Ω 上的数量场 $\boldsymbol{f}(x,y,z)$, M 是 Ω 内一点, 我们称

$$\mathbf{grad}\, f(M) = \nabla f(M)$$

为数量场 f 在 M 处的**梯度** (gradient). 让 M 取遍 Ω 内的点, 这样得到的向量场 $\mathbf{grad}\, f$ 称为数量场 f 的梯度场.

换言之, 数量场 f 在 M 处的梯度就是向量 $\nabla f(M)$. 梯度的方向就是方向导数最大的方向, 梯度的模就是方向导数的最大值. 采用梯度的记号, 可以把方向导数表示为

$$\frac{\partial f}{\partial \boldsymbol{v}}(M) = \mathbf{grad}\, f(M) \cdot \boldsymbol{v}.$$

设 C 是任意给定的常数, 我们把满足

$$f(x,y,z) = C \tag{1}$$

的点 (x,y,z) 的集合叫做数量场 f 的一个等值面. 设 M_0 是这等值面上的一个点, 并设

$$\nabla f(M_0) \neq \mathbf{0},$$

那么梯度 $\nabla f(M_0)$ 正好沿着曲面 (1) 在 M_0 处的法线方向. 我们得出结论: 在等值面上, 梯度向量与等值面垂直.

对于定义在平面区域上的数量场 $f(x,y)$, 我们也可以考察它的梯度

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

并且也可以考察它的等值线

$$f(x,y) = C.$$

同样可以得出结论: 在等值线上, 梯度向量与等值线垂直.

等值面与等值线在实际生活中有很多应用. 地形图上的等高线, 气象图上的等温线、等气压线

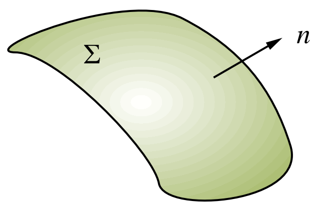
和等降雨量线等都是等值线的例子.

11.2　通量与散度

定义: 设有区域 Ω 内的向量场 $\boldsymbol{F}=\boldsymbol{F}(x,y,z)$, Σ 是 Ω 内的一块有向曲面, $\boldsymbol{n}=\boldsymbol{n}(x,y,z)$ 是曲面 Σ 的单位正法向量, 我们把曲面积分

$$\iint\limits_{\Sigma}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S$$

称为向量场 $\boldsymbol{F}$ 通过曲面 Σ 指定侧的**通量**.



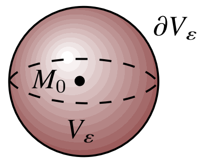
设通量

$$\Phi=\iint\limits_{\Sigma}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S,$$

那么

(1) 若通量 $\Phi>0$, 则向量场 $\boldsymbol{F}$ 沿指定侧穿过曲面 Σ 的流量多于反向穿过的流量; $\Phi<0$ 时, 结论反之.

(2) 若 Σ 为封闭曲面, 定向为外侧, 则: $\Phi>0$ 表明流出量大于流入量; $\Phi=0$ 表明流出量等于流入量; $\Phi<0$ 表明流出量小于流入量.



设 M_0 是 Ω 中的一点, 取以 M_0 为中心, ε 为半径的小球 V_ε. 下面我们证明, 当 $\varepsilon\to0$ 时, 向量场 F 通过 ∂V_ε 外侧的通量与 V_ε 的体积的比值

$$\frac{1}{|V_\varepsilon|}\oiint\limits_{\partial V_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S$$

有确定的极限. 由 Gauss 公式与积分中值定理得

$$\frac{1}{|V_\varepsilon|}\oiint\limits_{\partial V_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S=\frac{1}{|V_\varepsilon|}\iiint\limits_{V_\varepsilon}\nabla\cdot\boldsymbol{F}\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=(\nabla\cdot\boldsymbol{F})(M_\varepsilon).$$

这里 M_ε 是 V_ε 中的一点. 当 $\varepsilon \to 0$ 时, 应有 $M_\varepsilon \to M_0$, 因此

$$\lim_{\varepsilon\to 0}\frac{1}{|V_\varepsilon|}\oiint\limits_{\partial V_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S = (\nabla\cdot\boldsymbol{F})(M_0).$$

我们来解释一下这个极限的物理意义. 如果 $\boldsymbol{F}$ 是流体的速度场, 那么比值

$$\frac{1}{|V_\varepsilon|}\oiint\limits_{\partial V_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{n}\,\mathrm{d}S$$

是单位时间内从 V_ε 中流出的流体的量与 V_ε 的体积之比. 这个比值可以看做 V_ε 中的平均泉源密度 (比值为负时, 其实是漏洞). 当 $\varepsilon \to 0$ 时, 上述比值的极限就是流体在 M_0 处的泉源密度.

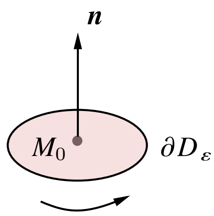
定义: 设有定义在区域 Ω 上的向量场 $\boldsymbol{F} = \boldsymbol{F}(x, y, z) = (P, Q, R)$, M 是 Ω 内一点, 我们称

$$\operatorname{div}\boldsymbol{F}(M) = (\nabla\cdot\boldsymbol{F})(M) = \frac{\partial P}{\partial x}(M) + \frac{\partial Q}{\partial y}(M) + \frac{\partial R}{\partial z}(M)$$

为向量场 $\boldsymbol{F}$ 在点 M 处的**散度** (divergence). 让 M 取遍 Ω 内的点, 这样得到的数量场 $\operatorname{div}\boldsymbol{F}$ 称为向量场 $\boldsymbol{F}$ 的散度场.

由前面的讨论可知, 散度是通量关于体积的变化率, 也即穿出单位体积区域边界的通量. 可用散度判断向量场中某一点是“源”还是“汇”, 以及“源”的强弱和“汇”的大小.

11.3 方向旋量与旋度



设 $\boldsymbol{F} = \boldsymbol{F}(x, y, z)$ 是定义在 Ω 上的向量场, M_0 为 Ω 中的一点, $\boldsymbol{n} = \boldsymbol{n}(x, y, z)$ 为单位向量, 以 M_0 为中心, $\boldsymbol{n}$ 为法向量作小圆面 D_ε. 我们考察当 $\varepsilon \to 0$ 时, 比值

$$\frac{1}{|D_\varepsilon|}\oint_{\partial D_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{t}\,\mathrm{d}s$$

的极限. 根据 Stokes 公式与积分中值定理, 我们有

$$\frac{1}{|D_\varepsilon|}\oint_{\partial D_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{t}\,\mathrm{d}s = \frac{1}{|D_\varepsilon|}\iint\limits_{D_\varepsilon}(\nabla\times\boldsymbol{F})\cdot\boldsymbol{n}\,\mathrm{d}S = [(\nabla\times\boldsymbol{F})\cdot\boldsymbol{n}](M_\varepsilon).$$

这里 M_ε 是 D_ε 中的一点. 当 $\varepsilon \to 0$ 时, 应有 $M_\varepsilon \to M_0$, 因此

$$\lim_{\varepsilon\to 0}\frac{1}{|D_\varepsilon|}\oint_{\partial D_\varepsilon}\boldsymbol{F}\cdot\boldsymbol{t}\,\mathrm{d}s=[(\nabla\times\boldsymbol{F})\cdot\boldsymbol{n}](M_0).$$

定义: 我们称

$$[(\nabla\times\boldsymbol{F})\cdot\boldsymbol{n}](M_0)$$

为向量场 $\boldsymbol{F}$ 在 M_0 处绕方向 $\boldsymbol{n}$ 的**方向旋量**.

我们曾考察过数量场沿哪个方向的方向导数取值最大的情形, 从而导出梯度的概念. 对于方向旋量, 我们可以提出类似的问题——绕哪个方向旋转, 方向旋量达到最大值? 为了回答这个问题, 我们把方向旋量写为

$$(\nabla\times\boldsymbol{F})\cdot\boldsymbol{n}=\|\nabla\times\boldsymbol{F}\|\cos\theta.$$

这里 θ 是向量 $\nabla\times\boldsymbol{F}$ 与向量 $\boldsymbol{n}$ 的夹角. 由此可以看出, 当 $\boldsymbol{n}$ 沿着 $\nabla\times\boldsymbol{F}$ 的方向时, 方向旋量达到最大值, 且最大值为

$$\|\nabla\times\boldsymbol{F}\|.$$

定义: 设有定义在区域 Ω 上的向量场 $\boldsymbol{F}=\boldsymbol{F}(x,y,z)=(P,Q,R)$, M 是 Ω 内一点, 我们称

$$\mathbf{rot}\,\boldsymbol{F}(M)=(\nabla\times\boldsymbol{F})(M)=\left.\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z}\\ P & Q & R\end{vmatrix}\right|_{(M)}$$

为向量场 $\boldsymbol{F}$ 在点 M 处的**旋度** (rotation). 让 M 取遍 Ω 内的点, 这样得到的向量场 $\mathbf{rot}\,\boldsymbol{F}$ 称为向量场 $\boldsymbol{F}$ 的旋度场.

由前面的讨论我们知道, 旋度的方向就是使得方向旋量最大的方向, 旋度的模就是方向旋量的最大值.

11.4 梯度, 散度, 旋度的性质

通过直接计算, 我们可以验证梯度, 散度与旋度具有以下一些性质.

(1) (梯度, 散度, 旋度的线性性)

$$\nabla(\alpha f+\beta g)=\alpha\,\nabla f+\beta\,\nabla g,\qquad \mathbf{grad}\,(\alpha f+\beta g)=\alpha\,\mathbf{grad}\,f+\beta\,\mathbf{grad}\,g.$$

$$\nabla\cdot(\alpha\boldsymbol{F}+\beta\boldsymbol{G})=\alpha\,\nabla\cdot\boldsymbol{F}+\beta\,\nabla\cdot\boldsymbol{G},\qquad \mathrm{div}\,(\alpha\boldsymbol{F}+\beta\boldsymbol{G})=\alpha\,\mathrm{div}\,\boldsymbol{F}+\beta\,\mathrm{div}\,\boldsymbol{G}.$$

$$\nabla\times(\alpha\boldsymbol{F}+\beta\boldsymbol{G})=\alpha\,\nabla\times\boldsymbol{F}+\beta\,\nabla\times\boldsymbol{G},\qquad \mathbf{rot}\,(\alpha\boldsymbol{F}+\beta\boldsymbol{G})=\alpha\,\mathbf{rot}\,\boldsymbol{F}+\beta\,\mathbf{rot}\,\boldsymbol{G}.$$

(2) (梯度, 散度, 旋度的乘法法则)

$$\begin{aligned}
&\nabla(fg) = (\nabla f)g + f(\nabla g), && \mathbf{grad}\,(fg) = (\mathbf{grad}\,f)g + f(\mathbf{grad}\,g),\\
&\nabla\cdot(f\boldsymbol{G}) = (\nabla f)\cdot\boldsymbol{G} + f(\nabla\cdot\boldsymbol{G}), && \mathrm{div}\,(f\boldsymbol{G}) = (\mathbf{grad}\,f)\cdot\boldsymbol{G} + f(\mathrm{div}\,\boldsymbol{G}).\\
&\nabla\times(f\boldsymbol{G}) = (\nabla f)\times\boldsymbol{G} + f(\nabla\times\boldsymbol{G}). && \mathbf{rot}\,(f\boldsymbol{G}) = (\mathbf{grad}\,f)\times\boldsymbol{G} + f(\mathbf{rot}\,\boldsymbol{G}).
\end{aligned}$$

其中 α 与 β 是常数, f 与 g 是数量场, $\boldsymbol{F}$ 与 $\boldsymbol{G}$ 是向量场.

通过直接计算, 我们还可以得到:

性质 (梯度, 散度, 旋度的算子零性) :

(3) $\mathbf{rot}\circ\mathbf{grad} = 0$.

(4) $\mathrm{div}\circ\mathbf{rot} = 0$.

证明 (3) 设 f 为任意一个二阶连续可微的数量场, 则

$$(\mathbf{rot}\circ\mathbf{grad})f = \mathbf{rot}\,(\mathbf{grad}\,f) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ f_x & f_y & f_z \end{vmatrix} = (f_{zy}-f_{yz}, f_{xz}-f_{zx}, f_{yx}-f_{xy}) = \mathbf{0}.$$

故 $\mathbf{rot}\circ\mathbf{grad}$ 为零算子.

(4) 设 $\boldsymbol{F}$ 为任意一个二阶连续可微向量场, 则

$$(\mathrm{div}\circ\mathbf{rot})\boldsymbol{F} = \mathrm{div}\,(\mathbf{rot}\,\boldsymbol{F}) = \mathrm{div}\,(R_y - Q_z, P_z - R_x, Q_x - P_y) = 0.$$

因此 $\mathrm{div}\circ\mathbf{rot}$ 为零算子.

实际上, 公式 (3) 与 (4) 可以统一概括为外微分算子的幂零性:

$$\mathrm{d}\circ\mathrm{d} = 0.$$

为此, 我们指出以下三点:

(a) 设 f 为数量场, 则 $\mathrm{d}f$ 的系数就是 $\mathbf{grad}\,f$ 的分量.

(b) 设 $\boldsymbol{F}=(P,Q,R)$, $\omega = P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z$, 则 $\mathrm{d}\omega$ 的系数就是 $\mathbf{rot}\,\boldsymbol{F}$ 的分量.

(c) 设 $\boldsymbol{F}=(P,Q,R)$, $\theta = P\mathrm{d}y\wedge\mathrm{d}z + Q\mathrm{d}z\wedge\mathrm{d}x + R\mathrm{d}x\wedge\mathrm{d}y$, 则 $\mathrm{d}\theta$ 的系数就是 $\mathrm{div}\,\boldsymbol{F}$.

证明 为方便起见, 我们用 $(\mathbf{rot}\,\boldsymbol{F})_1, (\mathbf{rot}\,\boldsymbol{F})_2, (\mathbf{rot}\,\boldsymbol{F})_3$ 表示 $\mathbf{rot}\,\boldsymbol{F}$ 的三个分量.

(3) 设 f 为数量场, $\boldsymbol{F} = \mathbf{grad}\,f$. 取 $\omega = f_x\mathrm{d}x + f_y\mathrm{d}y + f_z\mathrm{d}z = \mathrm{d}f$, 从而由$\underline{\mathrm{d}\omega = \mathrm{d}(\mathrm{d}f) = 0}$ 得

$$0 = \mathrm{d}\omega = (\mathbf{rot}\,\boldsymbol{F})_1\mathrm{d}y\wedge\mathrm{d}z + (\mathbf{rot}\,\boldsymbol{F})_2\mathrm{d}z\wedge\mathrm{d}x + (\mathbf{rot}\,\boldsymbol{F})_3\mathrm{d}x\wedge\mathrm{d}y.$$

从而 $(\mathbf{rot}\,\boldsymbol{F})_1=(\mathbf{rot}\,\boldsymbol{F})_2=(\mathbf{rot}\,\boldsymbol{F})_3=0$, 也即

$$\mathbf{rot}\,(\mathbf{grad}\,f)=\mathbf{rot}\,\boldsymbol{F}=\mathbf{0}.$$

由 f 的任意性知 $\mathbf{rot}\circ\mathbf{grad}=0$.

(4) 设 $\boldsymbol{F}=(P,Q,R)$, $\omega=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$, 则

$$\theta\overset{\Delta}{=\!=}\mathrm{d}\omega=(\mathbf{rot}\,\boldsymbol{F})_1\mathrm{d}y\wedge\mathrm{d}z+(\mathbf{rot}\,\boldsymbol{F})_2\mathrm{d}z\wedge\mathrm{d}x+(\mathbf{rot}\,\boldsymbol{F})_3\mathrm{d}x\wedge\mathrm{d}y.$$

从而由$\underline{\mathrm{d}\theta=\mathrm{d}(\mathrm{d}\omega)=0}$ 得

$$0=\mathrm{d}\theta=(\mathrm{div}\,(\mathbf{rot}\,\boldsymbol{F}))\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z.$$

因此 $\mathrm{div}\,(\mathbf{rot}\,\boldsymbol{F})=0$. 由 $\boldsymbol{F}$ 的任意性知 $\mathrm{div}\circ\mathbf{rot}=0$.

11.5 Green 第二公式

我们称 $\Delta=\nabla\cdot\nabla$ 为 Laplace 算子.

定理 (**Green 第二公式**):

$$\iiint\limits_{\Omega}(g\Delta f-f\Delta g)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint\limits_{\partial\Omega}\left(g\frac{\partial f}{\partial\boldsymbol{n}}-f\frac{\partial g}{\partial\boldsymbol{n}}\right)\mathrm{d}S.$$

证明 在散度的乘法法则 $\nabla\cdot(f\boldsymbol{G})=(\nabla f)\cdot\boldsymbol{G}+f(\nabla\cdot\boldsymbol{G})$ 中取 $\boldsymbol{G}=\nabla f$, 即得

$$\nabla\cdot(g\nabla f)=(\nabla g)\cdot(\nabla f)+g\Delta f.$$

由 Gauss 公式得

$$\iiint\limits_{\Omega}[(\nabla g)\cdot(\nabla f)+g\Delta f]\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iiint\limits_{\Omega}\nabla\cdot(g\nabla f)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint\limits_{\partial\Omega}g\nabla f\cdot\boldsymbol{n}\,\mathrm{d}S=\iint\limits_{\partial\Omega}g\frac{\partial f}{\partial\boldsymbol{n}}\mathrm{d}S.\tag{1}$$

在上式中交换 f 与 g, 可得

$$\iiint\limits_{\Omega}[(\nabla f)\cdot(\nabla g)+f\Delta g]\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iiint\limits_{\Omega}\nabla\cdot(f\nabla g)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint\limits_{\partial\Omega}f\nabla g\cdot\boldsymbol{n}\,\mathrm{d}S=\iint\limits_{\partial\Omega}f\frac{\partial g}{\partial\boldsymbol{n}}\mathrm{d}S.\tag{2}$$

(1)-(2) 得

$$\iiint\limits_{\Omega}(g\Delta f-f\Delta g)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint\limits_{\partial\Omega}\left(g\frac{\partial f}{\partial\boldsymbol{n}}-f\frac{\partial g}{\partial\boldsymbol{n}}\right)\mathrm{d}S.$$

注: (1) 式与 (2) 式被称为 **Green 第一公式**.

11.6 散度定理

定理 (散度定理) : 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, 向量场 $\boldsymbol{F}(\boldsymbol{x})$ 在 $\overline{\Omega}$ 上连续可微, 则

$$\int_\Omega \operatorname{div} \boldsymbol{F}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{F}(\boldsymbol{x}) \cdot \boldsymbol{n}(x)\,\mathrm{d}S,$$

其中 $\boldsymbol{n}$ 为 $\partial\Omega$ 的单位外法向量.

实际上, 下面的公式都是散度定理的特例.

(1) 牛顿-莱布尼兹公式:

$$\int_a^b u'(x)\mathrm{d}x = u(b) - u(a).$$

(2) Green 公式:

$$\iint\limits_D \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right)\mathrm{d}x\mathrm{d}y = \int_{\partial D} (u\cos\alpha + v\cos\beta)\,\mathrm{d}s.$$

(3) Gauss 公式:

$$\iiint\limits_V \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \iint\limits_{\partial V} (u\cos\alpha + v\cos\beta + w\cos\gamma)\,\mathrm{d}S.$$

其中 α, β, γ 分别为单位外法向量 $\boldsymbol{n}$ 与 x 轴, y 轴, z 轴正向的夹角.

推论 (分部积分公式) : 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, u, v 在 $\overline{\Omega}$ 上连续可微, 则

$$\int_\Omega v\frac{\partial u}{\partial x_i}\mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} uvn_i\mathrm{d}S - \int_\Omega u\frac{\partial v}{\partial x_i}\mathrm{d}\boldsymbol{x}.$$

其中 n_i 是单位外法向量 $\boldsymbol{n}$ 的第 i 个分量.

思路　在散度定理中取 $\boldsymbol{F} = (0, \cdots, 0, uv, 0, \cdots, 0)$.

注: §11.5 中的 Green 第二公式也可以由散度定理推出, 只需取 $\boldsymbol{F} = u\nabla v - v\nabla u$.

附注: 积分中值定理也可以推广到高维重积分与曲面积分.

定理 (重积分与曲面积分的中值定理) : 设 Ω 为 $\mathbb{R}^n$ 中的有界区域, 函数 u 在 $\overline{\Omega}$ 上连续, 则存在 $\xi \in \Omega$, 使得

$$\int_\Omega u(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = u(\boldsymbol{\xi})|\Omega|.$$

若 u 在 $\partial\Omega$ 上连续, 则存在 $\boldsymbol{\eta} \in \partial\Omega$, 使得

$$\int_{\partial\Omega} u(\boldsymbol{x})\mathrm{d}S = u(\boldsymbol{\eta})|\partial\Omega|.$$

11.7 曲线积分与路径无关的条件, 保守场

前面我们已经研究过曲线积分与路径无关的条件, 所得的结论可以用场论的语言陈述如下:

定义: 设向量场 $\boldsymbol{F}$ 在区域 Ω 内连续可微, 若存在 Ω 上的连续可微函数 U, 使得

$$\boldsymbol{F} = \mathbf{grad}\, U,$$

则称 $\boldsymbol{F}$ 是**保守场** (**有势场**), 函数 U 称为向量场 $\boldsymbol{F}$ 的**势函数**.

定义: 设向量场 $\boldsymbol{F}$ 在区域 Ω 内有定义, 若

$$\mathbf{rot}\, \boldsymbol{F} = \mathbf{0},$$

则称 $\boldsymbol{F}$ 是**无旋场**.

由于

$$\mathbf{rot} \circ \mathbf{grad} = 0,$$

因此保守场一定是无旋场.

定理: 设向量场 $\boldsymbol{F}$ 在单连通区域 Ω 内连续可微, 则下列各条等价:

(1) $\displaystyle\oint_L \boldsymbol{F} \cdot \boldsymbol{t}\, \mathrm{d}s = 0$.

(2) $\displaystyle\int_L \boldsymbol{F} \cdot \boldsymbol{t}\, \mathrm{d}s$ 与路径无关.

(3) $\boldsymbol{F}$ 是保守场.

(4) $\boldsymbol{F}$ 是无旋场.

注: 其中 (1)(2)(3) 的等价不需要单连通这一条件.

例　记 $\boldsymbol{r} = (x, y, z)$, $r = \|\boldsymbol{r}\| = \sqrt{x^2 + y^2 + z^2}$. 若向量场 $\boldsymbol{F}(x, y, z)$ 可以表示为

$$\boldsymbol{F} = f(r)\boldsymbol{r},$$

则称 $\boldsymbol{F}$ 是一个**中心场**. 试证明: 中心场必为保守场.

证明　假设存在连续可微函数 U 满足

$$\mathbf{grad}\, U = \boldsymbol{F}.$$

直接计算得

$$\mathbf{grad}\, U = \left(\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}\right) = \left(\frac{\mathrm{d}U}{\mathrm{d}r}\frac{\partial r}{\partial x}, \frac{\mathrm{d}U}{\mathrm{d}r}\frac{\partial r}{\partial y}, \frac{\mathrm{d}U}{\mathrm{d}r}\frac{\partial r}{\partial z}\right) = \frac{\mathrm{d}U}{\mathrm{d}r}\left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z}\right).$$

由于

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r},$$

从而

$$\mathbf{grad}\, U = \left(\frac{1}{r}\frac{\mathrm{d}U}{\mathrm{d}r}\right)(x, y, z) = \frac{1}{r}\frac{\mathrm{d}U}{\mathrm{d}r}\boldsymbol{r}.$$

因此, 要使得 $\mathbf{grad}\, U = f(r)\boldsymbol{r}$, 只需

$$\frac{\mathrm{d}U}{\mathrm{d}r} = rf(r).$$

两边积分得

$$U(r) = \int_1^r \rho f(\rho)\mathrm{d}\rho.$$