4 Lebesgue 积分

我们分几步来定义 Lebesgue 积分。

4.1 简单函数的积分

我们之前定义过简单函数

$$f=\sum_{k=1}^{N}a_k\chi_{E_k}$$

其中 a_k 为常数，E_k 为可测集合。对于每个简单函数，我们总可以写成”典范表达”：即是要求 a_k 互不相同，E_k 互不相交。由于简单函数 f 的值域是有限集 $\{c_1,c_2,\ldots,c_l\}$，因而 $\sum\limits_{k=1}^{l}c_k\chi_{f^{-1}(c_l)}$ 即为典范表达。另外一方面，对于 f 的任何表达式，可以通过有限步如下操作得到典范表达：细分某个 $E_k=E_{k1}\cup E_{k2}, a_k\chi_{E_k}=a_k\chi_{E_{k1}}+a_k\chi_{E_{k2}}$. 有限步这样的操作可使得所有的 E_{kl} 要么相等或者不相交。如果 $E_{kl}=E_{k'l'}$, 则进行操作 $a_k\chi_{E_{kl}}+a_{k'}\chi_{E_{k'l'}}=(a_k+a_{k'})\chi_{E_{kl}}$. 这样有限步操作后可使得所有的 E_{kl} 互不相交。最后的操作就是如果 $a_k=a_l$, 那么作操作 $a_k\chi_{E_k}+a_l\chi_{E_l}=a_k\chi_{E_k\cup E_l}$. 可以验证每步操作和函数

$$\sum_{k=1}^{N}a_k m(E_k)$$

都不变（当 $E_k=E_{k1}\cup E_{k2}, E_{k1}\cap E_{k2}=\emptyset$ 时，$m(E_k)=m(E_{k1})+m(E_{k2})$; 对第二种操作, $E_{kl}=E_{k'l'}$, 那么 $a_k m(E_{kl})+a_{k'}m(E_{k'l'})=(a_k+a_{k'})m(E_{kl})$; 对第三种操作，$a_k=a_l$, 那么有 $a_k m(E_k)+a_l m(E_l)=a_l m(E_k\cup E_l)$, 这里注意 $E_k\cap E_l=\emptyset$）. 所以对于简单函数我们可以定义积分

$$\int\sum_{k=1}^{N}a_k\chi_{E_k}=\sum_{k=1}^{N}a_k m(E_k).$$

并且定义跟简单函数的表达式无关，也就是跟 a_k, E_k 的选取无关。容易知道简单函数积分具备下列性质：

- 线性，如果 f, g 都是简单函数，那么
$$\int(af+bg)=a\int f+b\int g.$$
- 积分区域的可叠加性，对于两个不想交的可测集 E, F,
$$\int\limits_{E\cup F}f=\int\limits_{E}f+\int\limits_{F}f.$$
此时在可测集 E 上的积分可以理解为 $f\chi_E$（任然为简单函数）的积分。
- 单调性：如果 $f\leq g$，那么 $\int f\leq\int g$. 这等价于如果函数 $f\geq 0$，那么积分非负。对于非负的简单函数，典范表达的系数 a_k 非负，因而积分非负。
- 三角不等式：如果 f 是简单函数，那么 $|\int f|\leq\int|f|$.

4.2 有界函数在有限可测集上的积分

考虑 E 为任意固定的可测集, $m(E) < \infty$. 这节我们考虑所有 E 上的有界可测函数的积分。相当于说对于一般的可测函数 f，我们做截断 $f\chi_E\chi_{E_N}$，其中 $E_n = \{x : |f| < N\}$. 由简单函数逼近定理, 存在 φ_k 几乎处处收敛到 f. 如果 $\{f \neq 0\} \subset E$, $m(E) < \infty$ 并且 $|f| \leq M$，我们可以要求 $\{\varphi_k \neq 0\} \subset E$, $|\varphi_k| \leq M$. 对于这样的函数 f，定义积分

$$\int f = \lim_{k\to\infty} \int \varphi_k.$$

为了说明此定义是良定义，我们需要证明上式右边的极限存在并且跟简单函数列 φ_k 的选取无关。想法就是用 Egorov 定理。记 $I_k = \int \varphi_k$. 对任意 $\varepsilon > 0$ 存在闭集 $A_\epsilon \subset E$ 使得 φ_k 在 A_ε 上一致收敛到 f 并且 $m(E\backslash A_\varepsilon) < \varepsilon$. 这样存在 N 使得

$$|\varphi_n(x) - \varphi_m(x)| < \varepsilon, \quad \forall n, m > N, x \in A_\varepsilon.$$

故而当 $n, m > N$ 时

$$\begin{aligned}|I_n - I_m| &\leq \int_{A_\varepsilon} |\varphi_n - \varphi_m| + \int_{E\backslash A_\varepsilon} |\varphi_n - \varphi_m| \\ &\leq m(A_\varepsilon)\varepsilon + 2M\varepsilon.\end{aligned}$$

这就表明 I_k 是 Cauchy 列，因而收敛。这也就表明如果有两个简单函数列 $\{\varphi_k\}$, $\{\phi_k\}$ 都是几乎处处收敛到 f，那么 $\{\varphi_1, \phi_1, \varphi_2, \phi_2, \ldots\}$ 也是几乎处处收敛到 f。上面的讨论就表明

$$\lim_{k\to\infty} \int \varphi_k = \lim_{k\to\infty} \int \phi_k.$$

这样对于有界的并且支集在一个有限测度集上的函数我们定义了积分。这样函数的积分同样具备如下性质：

- 线性：$\int(af + bg) = a\int f + b\int g$. 这是因为如果 φ_k 几乎处处收敛到 f，ϕ_k 几乎处处收敛到 g，那么 $a\varphi_k + b\phi_k$ 几乎处处收敛到 $af + bg$。而简单函数的积分满足线性性质，因而可知这样的函数也满足线性性质。

- 积分区域的可叠加性，对于两个不相交的可测集 E, F,

$$\int_{E\cup F} f = \int_E f + \int_F f.$$

- 单调性：如果 $f \leq g$，那么 $\int f \leq \int g$. 这等价于如果函数 $f \geq 0$，对于非负的函数 f，可以取 φ_k 也都非负，因而积分非负。

- 三角不等式：如果 f 是简单函数，那么 $|\int f| \leq \int |f|$.

4.3 一般非负函数

现在考虑一般非负函数的积分。记 $\mathcal{L}_0$ 为 $\mathbb{R}^d$ 上所有支集在一个有限测度集上并且有界的函数组成的函数空间，也就是 $m(\{f \neq 0\}) < \infty$ 并且 $|f| \leq M$(这个常数 M 不是固定的). 我们可以定义积分

$$\int f = \sup\{\int g, \quad 0 \leq g \leq f, g \in \mathcal{L}_0\}$$

由此定义的积分显然满足单调性以及积分区域的可叠加性，只需要对线性性中的加法

$$\int f + g = \int f + \int g.$$

给予说明。不妨假设 f, g 的积分都有限，否则上述等号显然。根据定义存在 $f_n, g_n \in \mathcal{L}_0$ 使得

$$\int f = \lim_{k\to\infty} \int f_k, \quad \int g = \lim_{k\to\infty} \int g_k.$$

这表明 $f_k + g_k \leq f + g$ 并且 $f_k + g_k \in \mathcal{L}_0$。因而显然

$$\int f + g \geq \lim \int (f_k + g_k) = \int f + \int g.$$

另外取 $\varphi_k \in \mathcal{L}_0$ 使得

$$\int (f + g) = \lim_{k\to\infty} \int \varphi_k.$$

构造 $f_k = \min\{\varphi_k, f\}$, $g_k = \varphi_k - f_k$, 由于 $f_k \leq \varphi_k$, 因而 $f_k \in \mathcal{L}_0$, 进而 $g_k \in \mathcal{L}_0$. 注意到 $0 \leq f_k \leq f$, 对于 g_k, 在 x 处如果 $\varphi_k(x) \geq f(x)$, 那么 $f_k(x) = f(x)$，因而由条件 $\varphi_k(x) \leq f(x) + g(x)$ 可知 $g_k(x) = \varphi_k(x) - f_k(x) = \varphi_k(x) - f(x) \leq g(x)$; 如果 $\varphi_k(x) < f(x)$, 那么 $f_k(x) = \varphi_k(x)$, 故而 $0 \leq g_k(x) = 0 \leq g(x)$. 这样

$$\int \varphi_k = \int f_k + \int g_k \leq \int f + \int g$$

因而有 $\int (f+g) \leq \int f + \int g$. 这就证明了一般非负函数积分的线性性 (系数非负)。

4.4 一般函数的积分以及可积空间

对于可测函数 f, 如果 $|f|$ 的积分有限，则我们称 f 是可积的。在可测集 E 上所有可积函数组成的空间记作 $L^1(E)$. 对可积函数 f, 我们分解 $f = f_+ - f_-$ 为正部与负部，对于 f_+, f_- 我们可以同上定义积分，由于 $|f| = f_+ + f_-$，因而 f_+, f_- 都是 Lebesgue 可积的，这样可以定义积分

$$\int f = \int f_+ - \int f_-.$$

容易验证以上定义的积分满足如下性质：

- 线性，如果 f, g 都是可积的，那么对任意实数 a, b 有 $af+bg$ 可积，并且
$$\int (af+bg) = a\int f + b\int g.$$
- 积分区域的可叠加性，对于两个不相交的可测集 E, F,
$$\int_{E\cup F} f = \int_E f + \int_F f.$$
- 单调性：如果 $f \le g$，那么 $\int f \le \int g$.
- 三角不等式：如果 f 是可积，那么 $|\int f| \le \int |f|$.

4.5 Lebesgue 积分与 Riemann 积分

现在我们来看我们定义的 Lebesgue 积分确实是 Riemann 积分的推广。

定理 4.1. *有界闭区域上的函数 f 是 Riemann 可积的当且仅当 f 有界并且不连续点是零测集。*

证明: 对于区域 $[a, b]$ 的任何分割 $a = x_0 < x_1 < x_2 < \dots x_n = b$, 定义
$$\omega_i = \sup\{|f(x) - f(y)|, x, y \in [x_{i-1}, x_i]\}, \quad \Delta_i = |x_i - x_{i-1}|, \quad \Delta = \max\{\Delta_i\}$$
因而 f 是 Riemann 可积等价于
$$\lim_{\Delta\to 0}\sum_i \omega_i\Delta_i = 0.$$
再定义
$$\omega_\varepsilon(f) = \{x : \lim_{\delta\to 0}\sup_{y\in[x-\delta,x+\delta]} |f(y) - f(x)| \ge \varepsilon\}$$
先假设如果 f 的不连续点集测度为正，那么存在 $\varepsilon_0 > 0$ 使得 $\omega_{\varepsilon_0}(f) > 0$. 对任意分割，我们有
$$\sum_i \omega_i\Delta_i \ge \sum_{[x_{i-1},x_i]\cap\omega_{\varepsilon_0}(f)\ne\emptyset} \omega_i\Delta_i \ge \varepsilon_0 \sum_{[x_{i-1},x_i]\cap\omega_{\varepsilon_0}(f)\ne\emptyset} (x_i - x_{i-1}) \ge \varepsilon_0 m(\omega_\varepsilon(f)).$$
这表明 f 不是 Riemann 可积的。因而如果 f 是 Riemann 可积的，那么不连续点集必定是零测集。

现在我们假设 $\omega_\varepsilon(f)$ 是零测集，我们证明 f 是 Riemann 可积的。对任意 $\varepsilon > 0$，存在闭集 $A_\varepsilon \subset [a, b]$ 使得 f 在 A_ε 上连续。对 $x_0 \in A_\varepsilon$, 存在 $\delta > 0$ 使得
$$|f(x) - f(y)| < \varepsilon, \quad \forall x, y \in (x_0 - \delta, x_0 + \delta).$$
由于 A_ε 为紧集,因而存在有限个开区间 $(x_l - \frac{1}{2}\delta_l, x_l + \frac{1}{2}\delta_l)$ 覆盖住 A_ε。取 $\delta = \min\{\frac{1}{3}\delta_l\}$。这表明对任意 $x_0 \in A_\varepsilon$,必定有某个 $x_l \in A_\varepsilon$ 使得 $x_0 \in (x_l - \frac{1}{2}\delta_l, x_l + \frac{1}{2}\delta_l)$. 这表明 $[x_0 - \delta, x_0 + \delta] \subset (x_l - \delta_l, x_l + \delta_l)$, 因而有
$$|f(x) - f(y)| < \varepsilon, \quad \forall x, y \in [x_0 - \delta, x_0 + \delta].$$

这个结论当然也可以用反正法来证明，事实上否则存在 $\varepsilon_0 > 0$, 对正整数 n, 存在 $z_n \in A_\varepsilon$, $x_n, y_n \in [z_n - \frac{1}{n}, z_n + \frac{1}{n}]$ 使得

$$|f(x_n) - f(y_n)| \geq \varepsilon_0.$$

由于 A_ε 为紧集，因而存在子列（不妨仍然假设为 z_n）收敛 $z_n \to z_0$, 根据 x_n, y_n 的关系，可知 $x_n \to z_0$, $y_n \to z_0$. 而 f 在 z_0 处连续，与上不等式矛盾。

取 $[a, b]$ 的分割使得 $\Delta < \frac{1}{2}\delta$. 现在我们来考虑和 $\sum\limits_i \omega_i \Delta_i$. 如果区间 $[x_{i-1}, x_i]$ 与 A_ε 的交集非空，含有某个点 $y_0 \in [x_{i-1}, x_i] \cap A_\varepsilon$，那么对于任意 $x, y \in [y_0 - \delta, y_0 + \delta]$ 都有 $|f(x) - f(y)| < \varepsilon$. 注意到 $[x_{i-1}, x_i] \subset [y_0 - \delta, y_0 + \delta]$，故而 $\omega_i < \varepsilon$. 这样我们可以估计

$$\begin{aligned}\sum_i \omega_i \Delta_i &= \sum_{[x_{i-1}, x_i] \cap A_\varepsilon \neq \emptyset} \omega_i \Delta_i + \sum_{[x_{i-1}, x_i] \cap A_\varepsilon = \emptyset} \omega_i \Delta_i \\ &\leq \varepsilon(b-a) + 2Mm([a, b] \backslash A_\varepsilon)\end{aligned}$$

这里 M 为 f 在 $[a, b]$ 上的上届。这就表明如果 f 的不连续点为零测集并且 f 是有界的那么 f 在 $[a, b]$ 上 Riemann 可积。 □

定理 4.2. *如果 f 在区间 $[a, b]$ 上 Riemann 可积，那么 f 也是 Lebesgue 可积，并且积分相等。*

证明: 如果 f 是 Riemann 可积的，根据定义存在阶梯函数 $-M \leq \varphi_k \leq f \leq \phi_k \leq M$ 使得 φ_k 单调上升，ϕ_k 单调下降，这里 $|f| \leq M$, 并 Riemann 积分为

$$\int_{[a,b]}^{\mathcal{R}} f = \lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k = \lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \phi_k.$$

根据定义对于阶梯函数我们有 Riemann 积分与 Lebesgue 积分相同。另外一方面 $\lim\limits_{k \to \infty} \varphi_k = \underline{f}$, $\lim\limits_{k \to \infty} \phi_k = \bar{f}$ 均为有界可测函数。这样可知

$$\lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k = \int_{[a,b]}^{\mathcal{L}} \underline{f}, \quad \lim_{k \to \infty} \int_{[a,b]}^{\mathcal{R}} \phi_k = \int_{[a,b]}^{\mathcal{L}} \bar{f}.$$

这表明 $\int_{[a,b]}^{\mathcal{L}} \bar{f} - \underline{f} = 0$，故而 $\underline{f} = \bar{f}, a.e.$。又因为 $\underline{f} \leq f \leq \bar{f}$，可知 $f = \underline{f} = \bar{f}, a.e.$，因此 f 是 Lebesgue 可积的，并且 $\int_{[a,b]}^{\mathcal{L}} f = \int_{[a,b]}^{\mathcal{R}} f$. □

4.6 Lebesgue 控制收敛定理

既然我们已经知道 Lebesgue 积分是 Riemann 积分的推广，对于比较好的函数（Riemann 可积函数），积分的计算可以转换成经典的 Riemann 积分。而对于比较差的函数，我们只能通过好的函数逼近。这里就有个问题是当用好函数去逼近一个差函数时，对应的积分是否也收敛，也就是下列积分与极限是否可交换：

$$\lim_{k \to \infty} \int f_k = \int \lim_{k \to \infty} f_k?$$

一般来说上面等式不成立,在经典的 Riemann 积分意义下我们已经知道很多例子。即使在 Lebesgue 积分的意义下，例子可以很容易找到，比如 $f_n = \chi_{[n,n+1]}$，$f_n = \frac{1}{n}\chi_{[0,n]}$，$f_n = n\chi_{[0,\frac{1}{n}]}$。对于 Lebesgue 积分而言我们的目的是去找一个更加容易用的判别准则，这就是 Lebesgue 积分理论重要的 **Lebesgue 控制收敛定理**。

我们先来看一个简单的情形，就是函数 f_n 一致有界并且定义在一个有限测度集 E 上。

命题 4.3 (有限测度一致有界收敛定理). *考虑有限测度集 E 上的可测函数列 f_n，几乎处处收敛到函数 f，并且假设 f_n 几乎处处一致有界，i.e., $|f_n| \le M, a.e.$。那么必定有*

$$\int_E |f_n - f| \to 0.$$

特别的有 $\lim\limits_{n\to\infty} \int_E f_n = \int_E f$.

证明: 既然是在有限测度集上，因而可以用 Egorov 定理。对任意 $\epsilon > 0$，存在闭集 $A_\epsilon \subset E$ 使得 f_n 在 A_ϵ 上一致收敛到 f，并且 $m(E\backslash A_\epsilon) < \epsilon$。因而对于足够大的 n，我们有

$$|f_n - f| < \epsilon, \quad \forall x \in A_\varepsilon.$$

这样我们可以估计

$$\begin{aligned}\int_E |f_n - f| &\le \int_{A_\varepsilon} |f_n - f| + \int_{E\backslash A_\varepsilon} |f_n - f| \\ &\le \varepsilon m(A_\varepsilon) + 2Mm(E\backslash A_\varepsilon) \\ &\le m(E)\varepsilon + 2M\varepsilon.\end{aligned}$$

此即证明了定理结论。 □

对非负函数，我们有 Fatou 引理。

命题 4.4 (Fatou 引理). *对于非负函数列 f_n，我们有*

$$\int \liminf_{n\to\infty} f_n \le \varliminf_{n\to\infty} \int f_n.$$

证明: 任取 $g \in \mathcal{L}_0$ 使得 $0 \le g \le \liminf f_k$。再取 $g_k = \min\{g, f_k\}$。显然 $g_n \in \mathcal{L}_0$。我们需要证明 $g_k \to g$。否则存在 x_0 以及 $\varepsilon_0 > 0$ 以及子列 $\{g_{n'}\}$ 使得 $g(x_0) - g_{n'}(x_0) \ge \varepsilon_0$。根据定义 $g(x_0) - f_{n'}(x_0) \ge \varepsilon_0$。这与 $g \le \liminf f_k$ 矛盾。由于 $g_k \le g$，$g_k \to g$，$g \in \mathcal{L}_0$，因而根据支集有限测度的有限函数积分定义可知 $\int g = \lim \int g_n$。因而有

$$\int g = \lim \int g_n \le \liminf \int f_k.$$

由 g 的任意性，可知 $\int \liminf f_k \le \liminf \int f_k$。 □

作为一个推论，我们有单调上升函数列的收敛定理。

命题 4.5 (Beppo-Levi)**.** *对于单调上升的非负函数列 $f_n \to f$，我们有*

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

证明*:* 显然有 $\lim\limits_{n\to\infty} \int f_n \le \int f$。根据 Fatou 引理

$$\int f \le \liminf_{n\to\infty} \int f_n = \lim_{n\to\infty} \int f_n.$$

故而 Beppo-Levi 收敛定理。 □

作为一个推论对于正项函数级数，可以逐项积分:

$$\int \sum_{k=1}^{\infty} a_k(x) = \sum_{k=1}^{\infty} \int a_k(x).$$

作为简单应用，我们证明两个关于积分的性质。

命题 4.6. *假设 f 可积，那么对任意 $\varepsilon > 0$ 有*

- *存在有限测度集 B 使得*

$$\int_{B^c} |f| < \varepsilon.$$

- **积分的绝对连续性***存在 $\delta > 0$ 使得对任意可测集 E 使得 $m(E) < \delta$ 都有*

$$\int_E |f| < \varepsilon.$$

一个等价形式是

$$\lim_{k\to\infty} \int_{E_k} |f| = 0, \quad \text{如果} \lim_{k\to\infty} m(E_k) = 0.$$

证明*:* 不妨假设 f 非负。第一问只需要考虑函数列 $f^N = f\chi_{\{|x|\le N, f(x)\le N\}}$ 即可。对于积分的绝对连续性，可以取 N 足够大使得

$$\int f - f^N < \frac{\varepsilon}{2}.$$

然后再取 $\delta = \frac{\varepsilon}{2N}$，便有对任意可测集 $E, m(E) < \delta$，我们有

$$\int_E f \le \int_E f - f^N + \int_E f^N \le \frac{\varepsilon}{2} + N\delta \le \varepsilon.$$

证明二：反证存在 $\varepsilon_0 > 0$ 以及可测集 E_k 使得

$$m(E_k) \le 2^{-k}, \quad \int_{E_k^c} f \le \int f - \varepsilon_0.$$

考虑函数 $f\chi_{E_k^c}$ 的下极限，我们有

$$\liminf_{k\to\infty} f\chi_{E_k^c} = f\chi_{E^c}, \quad E = \limsup E_k.$$

注意到 $\sum\limits_{k=1}^{\infty} m(E_k) < 1$，因而根据 Borel-Cantelli 定理可知 E 为零测集。这样由 Fatou 引理可得

$$\int f\chi_{E^c} = \int f = \int \liminf_{k\to\infty} f\chi_{E_k^c} \le \liminf_{k\to\infty} \int f\chi_{E_k^c} = \liminf_{k\to\infty} \int_{E_k^c} f \le \int f - \varepsilon_0$$

矛盾！ $\square$

最后我们证明 Lebesgue 控制收敛定理，是上面所有收敛定理的推广。

命题 4.7 (Lebesgue 控制收敛定理)**.** *假设可测函数列 f_n 几乎处处收敛到 f，并且存在可积函数 g 使得*

$$|f_n| \le g, a.e.$$

那么有

$$\lim_{n\to\infty} \int |f_n - f| = 0.$$

特别的有积分收敛 $\lim\limits_{n\to\infty} \int f_n = \int f$.

证明: 对函数列 $g - f_n$ 以及 $g + f_n$ 用 Fatou 引理可得

$$\int 2g \le \liminf_{n\to\infty} \int 2g - |f - f_n|.$$

此即表明

$$\limsup_{n\to\infty} \int |f_n - f| \le 0.$$

$\square$

对于有限测度一致有界收敛定理，可以直接选取控制函数 $g(x) = M$。因为是在有限测度集合上，这样的控制函数显然可积。对于 Beppo Levi 单调升函数收敛定理，可以直接选取 f 作为控制函数。在开始我们举的三个例子：

- 对 $f_n = \chi_{[n,n+1]} \to 0$。可以选取 $g(x) = 1$ 做为控制函数，但是在整个实轴上 g 不可积，因而控制收敛定理不成立。
- 对 $f_n = \frac{1}{n}\chi_{[0,n]} \to 0$，可取控制函数 $g(x) = \min\{\frac{1}{x}, 1\}\chi_{(0,\infty)}$，同样 $g(x)$ 在 $(0,\infty)$ 上不可积.
- 对 $f_n = n\chi_{[0,\frac{1}{n}]} \to 0$，可取控制函数 $g(x) = \frac{1}{x}\chi_{(0,1)}$，但此控制函数不可积。

例子. 设 $f(x)$, $f_n(x)$ 是实值可积函数，如果对任意可测集 E 都有

$$\lim_{n\to\infty}\int\limits_E f_n = \int\limits_E f.$$

那么有 $\liminf f_n \le f \le \limsup f_n, a.e.$

证明: 由于 f_n, f 可积，不防假设 f_n，f 是有限值函数。又由于线性性，不防假设 $f=0$。我们只需要证明 $\limsup f_n \ge 0$。对任意 $\varepsilon>0$，作集合

$$E_n^\varepsilon = \{x : f_n(x) < -\varepsilon\}.$$

我们观察到

$$\liminf E_n^\varepsilon \subset \{\limsup f_n \le -\varepsilon\} \subset \liminf E_n^{\frac{1}{2}\varepsilon}.$$

事实上，如果 $x\in\liminf E_n^\varepsilon$，那么根据定义 $x\in E_n^\varepsilon, n>N$。也即是 $f_n(x)\le-\varepsilon, \forall n>N$。从而有 $\limsup f_n(x)\le-\varepsilon$。反过来如果 $\limsup f_n(x)\le-\varepsilon$，那么根据定义显然存在 N 使得 $\sup\limits_{k\ge N} f_k < -\frac{\varepsilon}{2}$。因而有 $x\in\liminf E_n^{\frac{\varepsilon}{2}}$。根据条件我们有

$$0=\lim\int\limits_{\bigcap\limits_{k\ge N} E_k^\varepsilon} f_n \le -\varepsilon m(\bigcap_{k\ge N} E_k^\varepsilon).$$

故而有 $\bigcap\limits_{k\ge N} E_k^\varepsilon$ 为零测集，这表明 $\liminf E_n^\varepsilon$ 为零测集。因而 $\{\limsup f_n < -\varepsilon\}$ 对任意 $\varepsilon>0$ 均为零测集。故 $\limsup f_n \ge 0, a.e.$. □

4.7 可积函数空间 L^1

这节我们整体来看可测函数空间。对于欧式空间 $\mathbb{R}^n$ 中的可测集 E，E 上所有的可积函数组成的空间记作 $L^1(E)$。如果是全空间，通常我们也简写为 L^1。首先可测函数在一个零测集上改变取值不影响可测性，并且不改变积分值。因而我们可以规定一个等价类，如果两个函数几乎处处相等，那么就属于同一个等价类里。另外，如果函数可积，这也表明取值无穷的点集是零测集，因而对于可积函数来说，可以假设函数取有限值。

之前的讨论已经证明了 $L^1(E)$ 是线性空间，只不过不同于有限维线性空间，比如 $\mathbb{R}^n$，$L^1(E)$ 是无限维的线性空间。另外，积分可以在这个线性空间上定义一个范数

$$\|f\| = \int_E |f|.$$

这只需要验证

- $\|f\|=0$, 那么有 $f=0, a.e.$，这就是为什么我们需要考虑等价类。
- $\|cf\| = |c|\|f\|$, 线性性。

- $\|f+g\| \leq \|f\|+\|g\|$, 三角不等式。

有了范数，我们可以自然的定义距离 $d(f,g)=\|f-g\|$。有了距离我们可以在这个线性空间上定义开集，闭集，邻域等拓扑概念。特别的，我们有极限、收敛的概念。如果 $\|f_k-f\| \to 0$，那么我们说函数列 f_k 在 L^1 的意义下收敛到 f, 记作 $f_k \stackrel{L^1}{\to} f$。

接下来考虑的是 $L^1(E)$ 是否是完备的赋范线性空间，也即是 $L^1(E)$ 是否是 Banach 空间。所谓完备性是指空间里的 Cauchy 列是否有极限。

定理 4.8 (Riesz-Fischer)**.** *$L^1(E)$ 是完备的赋范线性空间。*

证明: 我们需要证明对于 Cauchy 列 f_n，一定存在极限。取子列 f_{k_l} 使得

$$\|f_{k_l}-f_{k_{l-1}}\| \leq 2^{-l}, \quad l \geq 2.$$

定义函数

$$f=f_{k_1}+(f_{k_2}-f_{k_1})+(f_{k_3}-f_{k_2})+\ldots.$$

可知

$$\int_E |f| \leq \int_E |f_{k_1}|+|f_{k_2}-f_{k_1}|+|f_{k_3}-f_{k_2}|+\ldots<\infty.$$

这表明 $f \in L^1(E)$ 并且 $f_{k_l} \stackrel{L^1}{\to} f$. 而 f_k 是 Cauchy 列，故 $f_k \stackrel{L^1}{\to} f$. □

如同多项式函数在连续函数空间中稠密一样，我们希望找到一些特殊稠密子空间。

命题 4.9 (可积函数空间中的典型稠密子空间)**.** *以下函数空间在 $L^1(E)$ 中稠密：*

(1) 简单函数空间;

(2) 阶梯函数空间;

(3) 具有紧支集的连续函数空间 C_c;

(4) 无穷次可微的且有紧支集的函数空间 C_c^∞。

证明: 对于 (1): 假设 $f \in L^1(E)$，可分解 $f=f^+-f^-$。因而存在非负单调升简单函数 $\varphi_k^+ \to f^+$, $\varphi_k^- \to f^-$。由 Beppo-Levi 以及定义可知 $\|\varphi_k^+-f^+\| \to 0$, $\|\varphi_k^- - f^-\| \to 0$。这样定义 $\varphi_k=\varphi^+-\varphi_k^-$，可知 $\|\varphi_k-f\| \to 0$。

对于 (2): 这个想法也是很自然，首先可以用简单函数 g 逼近 f 使得 $\|f-g\|<\frac{\varepsilon}{2}$。可假设

$$g=\sum_{k=1}^N a_k \chi_{E_k}.$$

同样的可以不防假设 E_k 互不相交，a_k 互不相同。这样可取有限个方体并 I_k 使得

$$m(E_k \Delta I_k) < 2^{-k-1}(1+|a_k|)^{-1}\varepsilon.$$

这样定义

$$\varphi = \sum_{k=1}^{N} a_k \chi_{I_k}.$$

可知

$$\|\varphi - g\| \leq \sum_{k=1}^{N} |a_k| m(E_k \Delta I_k) \leq \sum_{k=1}^{N} |a_k| 2^{-k-1}(1+|a_k|)^{-1}\varepsilon < \frac{\varepsilon}{2}.$$

再用三角不等式可知 $\|\varphi - f\| \leq \varepsilon$ 并且 φ 是阶梯函数。

对于 (3): 根据上面的证明过程，我们只需要证明具有紧支集的连续函数可以逼近一个方体的示性函数即可。在一维，方体就是单个区间 (a,b). 那么可以取连续函数（或者光滑函数）g 使得 g 在 $[a+2\varepsilon, b-2\varepsilon]$ 上恒等于 1，在 $(a, a+\varepsilon)$ 以及 $(b-\varepsilon, b)$ 上为 0，在 $[a+\varepsilon, a+2\varepsilon]$ 和 $[b-2\varepsilon, b-\varepsilon]$ 上线性。这样可知 $\|g-(b-a)\| \leq 4\varepsilon(b-a)$. 这表明在一维的时候可以用具有紧支集的连续函数逼近方体的示性函数。高维对每个分量作乘积即可。 □

现在我们讨论可积函数在变换下的不变性。

- **积分的平移不变性**：假设 $f \in L^1(\mathbb{R}^d)$，对任意 $h \in \mathbb{R}^d$，定义平移变换 $\tau_h f(x) = f(x+h)$。那么函数 $\tau_h f \in L^1(\mathbb{R}^n)$，范数 $\|\tau_h f\| = \|f\|$，积分也相同，$\int \tau_h f = \int f$。
- **积分的伸缩不变性**：假设 $f \in L^1(\mathbb{R}^d)$，对任意 $\delta > 0$，定义伸缩变换 $D_\delta f(x) = f(\delta x)$. 那么函数 $D_\delta f \in L^1(\mathbb{R}^n)$，范数 $\|D_\delta f\| = \delta^{-d}\|f\|$，积分 $\int D_\delta f = \delta^{-d}\int f$。
- **积分的反射不变性**：假设 $f \in L^1(\mathbb{R}^d)$，定义反射变换 $Rf(x) = f(-x)$. 那么函数 $Rf \in L^1(\mathbb{R}^n)$，范数 $\|Rf\| = \|f\|$，积分 $\int Rf = \int f$。

证明的想法就是可以用阶梯函数逼近可积函数，对于方体的示性函数，上述变换下显然积分不变。这样由线性性可知阶梯函数也满足上述性质。最后用 Lebesgue 控制收敛定理可知对一般的可积函数都满足上述变换积分保持不变。

平移和伸缩还满足连续性。

命题 4.10 (平移伸缩变换的连续性). *如果 f 可积，那么有*

$$\lim_{h\to 0} \|\tau_h f - f\| = 0, \quad \lim_{\delta \to 1} \|D_\delta f - f\| = 0.$$

证明: 对任意 $\varepsilon > 0$ 存在简单函数 g 使得 $\|g - f\| < \frac{\varepsilon}{3}$ 并且

$$g = \sum_{k=1}^{N} a_k \chi_{I_k}.$$

这里可以要求 I_k 互不相交。这样由于平移和伸缩的不变性，可知

$$\|\tau_h g - \tau_h f\| < \frac{\varepsilon}{3}, \quad \|D_\delta g - D_\delta h\| < \delta^{-d}\frac{\varepsilon}{3}.$$

对于简单函数 g 我们可以得到

$$\|\tau_h g - g\| = \int |\sum_{k=1}^{N} a_k(\chi_{I_k} - \chi_{I_k+\{h\}})| \leq C|h|\sum_{k=1}^{N}|I_k| = C|h|(\|f\| + \frac{\varepsilon}{3}).$$

这样当 $|h|$ 足够小是有

$$\|\tau_h f - f\| \leq \|\tau_h g - g\| + \|f - g\| + \|\tau_h f - \tau_h g\| < \varepsilon.$$

同样的对于伸缩变换，我们有

$$\|D_\delta g - g\| = \int |\sum_{k=1}^{N} a_k(\chi_{\delta I_k} - \chi_{I_k})| \leq C|\delta - 1|\|g\|.$$

$\square$

4.8 习题

北京大学数学分析 II 实验班作业 1

本次作业的提交时间和地点为 **3** 月 **5** 日习题课上，逾期视作零分。

1. 对非负函数 f，如果 $\int f = 0$，那么必定有 $f = 0, a.e.$.

2. 对于可积函数 f，如果对于任意开集 O 都有 $\int_O f = 0$，那么 $f = 0, a.e.$.

3. **Chebychev 不等式**如果 f 可积，那么对任意 $\alpha > 0$ 有
$$m(\{x : |f(x)| > \alpha\}) \le \frac{1}{\alpha}\int |f|.$$

4. 对于非负可测函数 f，定义
$$E_k = \{x : f(x) > 2^k\}, \quad F_k = \{x : 2^k < f(x) \le 2^{k+1}\}.$$
证明 f 可积等价于 $\sum\limits_{k=-\infty}^{\infty} 2^k m(E_k)$ 收敛或者 $\sum\limits_{k=-\infty}^{\infty} 2^k m(F_k)$ 收敛。划分值域还可以用其他趋于无穷数列，比如 $\{n\}$.

5. 考虑可测集集合 E 上的非负可测函数 f，如果 $m(E) < \infty$，那么 f 可积等价于
$$\sum_{n=0}^{\infty} m(\{x \in E : f(x) \ge n\}) < +\infty.$$
注意此题与上题的区别，上题中不需要要求 E 是有限测度，原因在于分割的足够细致（靠近 0 的地方）。而此题的分割在 0 处的跨度为 1，因而需要要求全集 E 有限测度。

6. *此题是关于 Lebesgue 积分与广义积分的关系*：回忆一下广义积分的定义，对于一般的函数，一维与高维有本质上的区别。证明如下结论：
   (a) 在 $\mathbb{R}^d, d \ge 2$, 如果 f 是广义可积的 (可只证明无穷积分或者奇异积分)，那么 f 也必定是 Lebesgue 可积的，并且积分相同。
   (b) 在 $\mathbb{R}$ 上，如果 f 的广义积分存在并且是 Lebesgue 可积的，那么 Lebesgue 积分等于广义积分。
   (c) 在 $(0,1)$ 上，证明函数 $f(x) = \frac{1}{x}\sin(\frac{1}{x^3})$ 广义积分存在，但不是 Lebesgue 可积的。

7. 对于一般的非负函数 f (不一定可测)，定义 f 的 Lebesgue 上、下积分
$$\overline{\int} f = \inf\{\int g, \quad f \le g, \quad g\text{是简单函数}\}, \quad \underline{\int} f = \sup\{\int g, \quad 0 \le g \le f, \quad g\text{是简单函数}\}.$$
证明如果 f 可测，那么 f 的 Lebesgue 积分 (课上定义) 等于 $\underline{\int} f$. 因而可用下积分来定义 Lebesgue 积分。

8. 如果 f 非负有界并且有紧支集，证明 f 可测等价于 f 的上下积分相等。

9. 考虑 $\mathbb{R}$ 上的有界函数 $f(x)$，其不连续点集记为 D。如果 D 只有可数个极限点，证明 f 在区间 $[a,b]$ 上是 Riemann 可积的。

10. 如果可测函数列 $f_k(x)$ 依测度收敛到 $f(x)$，并且有可积函数 F 使得

$$|f_k(x)| \leq F(x).$$

那么有 f 可积，并且

$$\lim_{k\to\infty} \int f_k = \int f.$$

提示：利用 Lebesgue 控制收敛定理。

11. 此题是 *Fatou* 引理更加一般的形式称为 *Fatou-Lebesgue* 定理: 给定可积函数列 f_n，如果存在可积函数 $g(x)$ 使得 $|f_n(x)| \leq g(x)$，那么 $\liminf f_n$ 以及 $\limsup f_n$ 均可积，并且有

$$\int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n \leq \limsup_{n\to\infty} \int f_n \leq \int \limsup_{n\to\infty} f_n.$$

12. 记所有可测函数等价类 (几乎处处相等的意义下) 组成的空间为 $L^0(E)$. 如果 $0 < m(E) < \infty$，定义 $L^0(E)$ 上的距离函数

$$d(f,g) = \int_E \frac{|f-g|}{1+|f-g|}.$$

证明在此距离下收敛等价于依测度收敛。

13. 如果 f 在实轴 $\mathbb{R}$ 上可积，定义 $F(x) = \int_{-\infty}^{x} f(t)dt$. 证明 $F(x)$ 是 $\mathbb{R}$ 上的连续函数。

14. 如果 f 在 $\mathbb{R}^n$ 上可积，E 是 $\mathbb{R}^n$ 上的紧集，证明

$$\lim_{|y|\to\infty} \int_{E+\{y\}} f = 0.$$

15. 证明如果 f_k 在 L^1 中收敛到 f，那么 f_k 依测度收敛到 f, 即是 $f_k \xrightarrow{L^1} f \Longrightarrow f_k \xrightarrow{m} f$.

16. 如果 f_k 几乎处处或者依测度收敛到 f, 并且 $\|f_k\| \to \|f\|$这里指 L^1 范数，那么有 $f_k \xrightarrow{L^1} f$.

17. 如果 $f_k \xrightarrow{L^1} f$，并且 $\sum\limits_{k=1}^{\infty} \|f_k - f\| < \infty$，证明 f_k 近一致收敛到 f.

18. 证明伸缩变换是连续的

$$\lim_{\delta\to 1} \|D_\delta f - f\| = 0.$$

19. 证明 $L^1(\mathbb{R}^d)$ 是可分的距离空间，即是证明存在可数稠密子集。

20. 假设 $f \in L^1(\mathbb{R})$, $a > 0$，证明级数

$$\sum_{n=-\infty}^{\infty} f(\frac{x}{a} + n)$$

在 $\mathbb{R}$ 上几乎处处绝对收敛，和函数 $S(x)$ 是以 a 为周期的周期函数且 $S(x) \in L^1((0,a))$.

4.10 Fubini 定理及应用

Fubini 定理是重积分化成累次积分在 Lebesgue 积分理论的对应。考虑 $\mathbb{R}^d$ 上的可积函数 $f(x,y)$，其中 $x\in\mathbb{R}^{d_1}$, $y\in\mathbb{R}^{d_2}$, $d=d_1+d_2$。固定 x 时，对 y 作积分，我们想知道 $\int_{\mathbb{R}^{d_2}} f(x,y)$ 是否是 $\mathbb{R}^{d_1}$ 上的可积函数，并且是否有如下关系

$$\int f(x,y)=\int_{\mathbb{R}^{d_1}}\left(\int_{\mathbb{R}^{d_2}} f(x,y)\right).$$

如下 Fubini 定理给了这个问题肯定答案.

定理 4.11 (Fubini 定理)**.** *给定 $\mathbb{R}^d$ 上的可积函数 $f(x,y)$，那么*

(1) 对几乎所有的 x，作为 y 的函数 $f(x,y)$ 是 $\mathbb{R}^{d_2}$ 上的可积函数；

(2) 对 y 的积分 $\int_{\mathbb{R}^{d_2}} f(x,y)$ 定义了 $\mathbb{R}^{d_1}$ 上的可积函数；

(3) 积分满足如下关系式

$$\int_{\mathbb{R}^d} f(x,y)=\int_{\mathbb{R}^{d_1}}\left(\int_{\mathbb{R}^{d_2}} f(x,y)\right).$$

证明过程也是标准的：基本的想法就是如果 $f(x,y)$ 是一个方体的示性函数，那么结论是成立。再由线性性，阶梯函数也满足定理。如果能再证明满足定理的函数列的极限也满足定理，那么由阶梯函数在可积函数里的稠密性，可知定理结论成。为了让叙述更加清楚，我们换一种证明过程。记 $\mathcal{F}$ 是所有满足上述 Fubini 定理的可积函数。$\mathcal{F}$ 满足如下性质

命题 4.12. *$\mathcal{F}$ 是线性空间，并且***非负单调序列***的极限，如果可积则也属于 $\mathcal{F}$。*

证明: $\mathcal{F}$ 显然对数乘封闭，我们先证明对加法封闭。如果 $f,g\in\mathcal{F}$, 那存在 $\mathbb{R}^{d_1}$ 中的零测集 X_f, X_g 使得对任意 $x\notin X_f\cup X_g$, $f(x,y)$，$g(x,y)$ 作为 y 的函数是 $\mathbb{R}^{d_2}$ 中的可积函数。特别的，对任意 $x\notin X_f\cup X_g$, $f(x,y)+g(x,y)$ 作为 y 的函数在 $\mathbb{R}^{d_2}$ 中可积。由假设积分 $\int_{\mathbb{R}^{d_2}} f(x,y)$, $\int_{\mathbb{R}^{d_2}} g(x,y)$ 作为 x 函数是 $\mathbb{R}^{d_1}$ 上的可积函数。注意到固定 x 时

$$\int_{\mathbb{R}^{d_2}} f(x,y)+g(x,y)=\int_{\mathbb{R}^{d_2}} f(x,y)+\int_{\mathbb{R}^{d_2}} g(x,y).$$

这就表明 $f+g\in\mathcal{F}$.

现在我们假设非负函数列 $f_k\in\mathcal{F}$，$f_k\to f$, 并且 f 是 $\mathbb{R}^d$ 上的可积函数。根据条件存在零测集 X_k 使得 $x\notin X_k$, $f_k(x,y)$ 作为 y 的函数是 $\mathbb{R}^{d_2}$ 上的可积函数。当 $x\notin\bigcup\limits_{k=1}^{\infty}X_k$，可知作为 y 的函数 $f(x,y)$ 是 $\mathbb{R}^{d_2}$ 上的可测函数且满足

$$\lim_{k\to\infty}\int_{\mathbb{R}^{d_2}} f_k(x,y)=\int_{\mathbb{R}^{d_2}} f(x,y).$$

这里我们需要注意的是上面右边积分不一定有限，如果有限才能说明作为 y 的函数是可积的！再根据假设以及 Beppo-Levi 定理（单调非负函数的积分极限，单调增则直接用 Beppo-Levi，单调降的话可用 Lebesgue 控制收敛定理，因为此时第一个函数就是控制函数）

$$\int_{\mathbb{R}^d} f_k(x,y) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f_k(x,y) \right), \quad \int_{\mathbb{R}^d} f_k(x,y) \to \int_{\mathbb{R}^d} f(x,y),$$
$$\int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f_k(x,y) \right) \to \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y) \right).$$

由于 $f(x,y)$ 作为 $\mathbb{R}^d$ 上的函数可积,可知作为 x 的函数 $\int_{\mathbb{R}^{d_2}} f(x,y)$ 在 $\mathbb{R}^{d_1}$ 上可积,因而 $\int_{\mathbb{R}^{d_2}} f(x,y)$ 几乎处处有限。故而对于几乎处处的 x，作为 y 的函数 $f(x,y)$ 在 $\mathbb{R}^{d_2}$ 上是可积的。并且积分满足 Fubini 定理结论。 □

由上述线性性以及非负单调列封闭，我们现在只需要验证简单集合的示性函数属于 $\mathcal{F}$ 即可。

命题 4.13. *$\mathcal{F}$ 包含以下集合:*

(1) 方体的示性函数属于 $\mathcal{F}$;

(2) 有限测度开集的示性函数属于 $\mathcal{F}$;

(3) 有限测度 G_δ 集的示性函数属于 $\mathcal{F}$;

(4) 零测集的示性函数属于 $\mathcal{F}$;

(5) 有限测度可测集的示性函数属于 $\mathcal{F}$.

证明: 对于 (1)，只需要考虑 $\mathbb{R}^d$ 中的单独一个方体 I。可知 $I = I_x \times I_y$，其中 I_x 是 $\mathbb{R}^{d_1}$ 中的方体, I_y 是 $\mathbb{R}^{d_2}$ 中的方体。因而 $\chi_I(x,y) = \chi_{I_x}(x)\chi_{I_y}(y)$. 自然的对所有的 x，函数 $\chi_{I_x}(x)\chi_{I_y}(y)$ 是 $\mathbb{R}^{d_2}$ 上的可积函数。并且 $|I| = |I_x||I_y| = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_{I_x}(x)\chi_{I_y}(y) \right)$.

对于 (2), 有限开集可以写成是可数个不相交的方体的并, $O = \bigcup\limits_{k=1}^{\infty} I_k$(注意，这里的方体可以是半开半闭的). 这样 $\lim\limits_{k\to\infty} \chi_{E_k} = \chi_O$, 这里 $E_k = \bigcup\limits_{j=1}^{k} E_j$. 这样 χ_{E_k} 单调升收敛到 χ_O。故而 $\chi_O \in \mathcal{F}$.

对于 (3), 对于 G_δ 集 O, 存在有限测度下降开集列 $E_k \to O$. 这样可知 $\chi_{E_k} \to \chi_O$, 由于 E_k, O 均为有限测度，因而 $\chi_{E_1\setminus E_k}$ 单调升到 $\chi_{E_1\setminus O}$. 故 $\chi_{E_1\setminus O} \in \mathcal{F}$。因为 E_1 为开集，这就表明 $\chi_O \in \mathcal{F}$.

对于 (4), 任何零测集都是更大的某个 G_δ 零测集的子集。因而只要证明如果 E 为零测集, 并且 $\chi_E \in \mathcal{F}$, 那么对于任何子集 $A \subset E$，也有 $\chi_A \in \mathcal{F}$. 根据假设, 对几乎处处的 x, χ_E 作为 y 的函数在 $\mathbb{R}^{d_2}$ 上可积，并且积分满足

$$0 = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_E(x,y) \right).$$

这表明作为 x 的函数 $\int_{\mathbb{R}^{d_2}} \chi_E(x,y) = 0, a.e.$（对 x）. 因而对几乎处处的 x, $\chi_E(x,y) = 0, a.e.$(对 y). 特别对 $A \subset E$, 对几乎处处的 x, $\chi_A(x,y) = 0, a.e.$(对 y). 这样，对几乎处处的 x，$\chi_A(x,y)$ 作为 y 的函数是 $\mathbb{R}^{d_2}$ 上的可积函数，并且 $\int_{\mathbb{R}^{d_2}} \chi_A(x,y) = 0, a.e.$. 这就表明 $\chi_A \in \mathcal{F}$.

最后对于 (5)，对任何有限测度可测集 E，都可以写成 $E = G\backslash Z$，其中 G 为 G_δ 集，而 Z 为零测集。因而 $\chi_E = \chi_G - \chi_Z$, 根据线性规则，可知 $\chi_E \in \mathcal{F}$。 □

现在我们就可以证明 Fubini 定理。非负可积函数是一列单调升的简单函数的极限，因而属于 $\mathcal{F}$。对于一般的可积函数，可以分成正部和负部的差，因而一般可积函数也满足 Fubini 定理。

对于一般的非负可测函数，还有与 Fubini 定理相配的 Tonelli 定理.

定理 4.14 (Tonelli 定理)**.** *对于 $\mathbb{R}^d$ 上的非负可测函数 $f(x,y)$，有*

(1) 对于几乎所有的 x，$f(x,y)$ 作为 y 的函数是 $\mathbb{R}^{d_2}$ 上的可测函数。

(2) 作为 x 的函数，$\int_{\mathbb{R}^{d_2}} f(x,y)$ 是 $\mathbb{R}^{d_1}$ 上的可测函数。

(3) 积分满足如下等式（可以是无穷）

$$\int_{\mathbb{R}^d} f(x,y) = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} f(x,y) \right).$$

如果 f 可积，Tonelli 定理就是 Fubini 定理的推论。所以 Tonelli 定理主要说的是可测性质。考虑 $f(x,y)\chi_{|x|+|y|<k}\chi_{f<|k|}$，单调升到 f。再由 Fubini 定理，可知 Tonelli 定理成立。

Tonelli 定理有个自然的推论.

命题 4.15. *如果 E 是 $\mathbb{R}^d$ 上的可测集，那么对几乎所有的 x，集合 $E^x = \{y : (x,y) \in E\}$ 是 $\mathbb{R}^{d_2}$ 上的可测集。并且作为 x 的函数 $m(E^x)$ 是 $\mathbb{R}^{d_1}$ 上的可测函数，满足*

$$m(E) = \int_{\mathbb{R}^{d_1}} m(E^x).$$

这个命题的逆命题不成立。考虑 $\mathscr{N} \times [0,1]$，其中 $\mathscr{N}$ 是 $\mathbb{R}$ 中的不可测函数。对于 $x \in \mathbb{R}$，E^x 要么为空集，要么为 $[0,1]$，均为可测集。但 E 不是可测集。

如果承认选择公理和连续统假设，我们还可以构造一个不可测集集 $E \subset [0,1] \times [0,1]$，使得 E^x, E^y 都是可测集。连续统假设是说最小的不可数基数与实数集基数相同，我们利用此来构造 $\mathbb{R}$ 的全序 $\prec$ 使得 $\{x : x \prec y\}$ 是可数集。事实上，由选择公理存在 $\mathbb{R}$ 上的良序 $\prec$, 构造集合 $X = \{y : \text{集合}\{x : x \prec y\}\text{是不可数的}\}$。如果 X 是空集，$\prec$ 刚好就满足所需条件。否则由良序关系，集合 X 中存在最小元 $\bar{y}$。此时集合 $\{x : x \prec \bar{y}\}$ 不可数，因而与 $\mathbb{R}$ 同样基数，这样 $\{x : x \prec \bar{y}\}$ 中的良序关系就给出了 $\mathbb{R}$ 上满足要求的良序关系。有了这样的良序 $\prec$，我们再来定义集合 $E = \{(x,y) : 0 \le x \le 1, 0 \le y \le 1, x \prec y\}$. 固定 y，集合 E^y 是可数集，因而 $m(E^y) = 0$；固定 y, 集合 $[0,1]\backslash E^y$ 为可数集，因而 $m(E^y) = 1$。这样可知 E 不可测，否则满足积分等式关系。

我们现在来讨论 Fubini 定理的应用。本质上来看，$\mathbb{R}^d$ 上的测度可以看成是低维空间 $\mathbb{R}^{d_1}$，$\mathbb{R}^{d_2}$ 上测度的乘积。我们有如下事实

命题 4.16. *我们有*

(1) 如果 E_1, E_2 分别是 $\mathbb{R}^{d_1}$, $\mathbb{R}^{d_2}$ 中的可测集, 那么 $E = E_1 \times E_2 = \{(x, y) \in \mathbb{R}^d, x \in E_1, y \in E_2\}$ 是 $\mathbb{R}^d$ 中的可测集，并且有 $m(E) = m(E_1)m(E_2)$。

(2) 反过来，如果 $E = E_1 \times E_2$ 是 $\mathbb{R}^d$ 中的可测集，那么 E_1, E_2 分别是 $\mathbb{R}^{d_1}$, $\mathbb{R}^{d_2}$ 中的可测集，并且有 $m(E) = m(E_1)m(E_2)$，或者 E_1, E_2 中有一个为零测集，并且此时有 E 为零测集

(3) 作为推论，如果 $f(x)$ 是 $\mathbb{R}^{d_1}$ 上的可测函数，那么 $g(x, y) = f(x)$ 作为 $\mathbb{R}^d$ 上函数可测。

证明*:* 为了证明这些结论，首先由习题可知外测度满足关系:

$$m^*(E) \leq m^*(E_1) \times m^*(E_2).$$

这里只需要注意到关系式 $\cup E_k \times \cup F_l = \cup(E_k \times F_l)$。一个直接的推论便是如果 E_1 或者 E_2 为零测集，那么 E 必定也为零测集。特别的如果 E_1, E_2 均为零测集，那么 E 为零测集。

接下里我们证明如果 E_1, E_2 均为可测集，并且不是零测集，那么 $E = E_1 \times E_2$ 为可测集。可不妨假设 E_1, E_2 测度有限。由可测集性质，存在 G_δ 集 F_1, F_2 以及零测集 Z_1, Z_2 使得

$$E_1 = F_1 \backslash Z_1, \quad E_2 = F_2 \backslash Z_2.$$

显然 $F_1 \times F_2$ 为 $\mathbb{R}^d$ 中的 G_δ 集，因而可测。注意到

$$(F_1 \times F_2) \backslash (Z_1 \times F_2 \cup F_1 \times Z_2) \subset E \subset F_1 \times F_2$$

由上述结论可知 $Z_1 \times F_2$, $F_1 \times Z_2$ 均为零测集，故而 E 可测。

最后如果 $E = E_1 \times E_2$ 可测，并且 E_1, E_2 不为零测集，那么由 Tonelli 定理，考虑函数 χ_E 为 $\mathbb{R}^d$ 上的可测函数，根据定义 $\chi_E = \chi_{E_1}\chi_{E_2}$. 由 Tonelli 定理可知，对几乎所有的 x，作为 $\mathbb{R}^{d_1}$ 上的函数 $\chi_{E_1}\chi_{E_2}$ 是 $\mathbb{R}^{d_2}$ 上的可测函数，这样由定义可知集合 E_2 为 $\mathbb{R}^{d_2}$ 上的可测集。然后作积分

$$\int_{\mathbb{R}^{d_2}} \chi_{E_1}\chi_{E_2} = \chi_{E_1} m(E_2).$$

根据 Tonelli 定理, 作为 x 的函数 $m(E_2)\chi_{E_1}$ 是 $\mathbb{R}^{d_1}$ 上的可测函数，注意到 $m(E_2) > 0$, 因而可知 E_1 也为可测集。并且有等式

$$m(E) = \int_{\mathbb{R}^d} \chi_E = \int_{\mathbb{R}^{d_1}} \left(\int_{\mathbb{R}^{d_2}} \chi_{E_1}\chi_{E_2} \right) = m(E_2)m(E_1).$$

这就表明 (1) 中可测集测度关系。并证明 (2) 中结论。

对于 (3)，只需要验证 $\{(x, y)|g(x) > t\}$ 为 $\mathbb{R}^d$ 中的可测集即可，这只需要注意到

$$\{(x, y)|f(x) > t\} = \{x|f(x) > t\} \times \mathbb{R}^{d_2}.$$

因而为可测集。 $\square$

在之前的作业中我们已经知道对线性变换 $L : \mathbb{R}^d \to \mathbb{R}^d$ 以及可测集 E, 有 $L(E)$ 可测。现在我们还可以得到测度的变换关系。

命题 4.17 (集合测度在线性变换下的转换). 给定线性变换 $L:\mathbb{R}^d\to\mathbb{R}^d$ 以及 $\mathbb{R}^d$ 上的可测集 E, 那么 $L(E)$ 可测，并且

$$m(L(E))=|detL|m(E).$$

特别的，可测集在旋转变换下的测度不变，即是 $L\in O(n)$，那么 $m(O(E))=m(E)$.

证明: 我们先做一个简化，只需要证明上述关系对方体成立即可。这样由于开集是可数个方体的并可知开集也满足上面关系，因而可知闭集也满足（可只考虑有界集合）。对一般可测集 E 以及任意 $\varepsilon>0$，存在闭集 $F\subset E$ 以及开集 $G\supset E$ 使得 $m(G\backslash F)<\varepsilon$. 因而

$$\begin{aligned}|detL|(m(E)-\varepsilon)&\leq|detL|m(F)=m(L(F))\leq m(L(E))\\&\leq m(L(G))=|detL|m(G)\leq|detL|(m(E)+\varepsilon).\end{aligned}$$

另外线性变换 L 对应的矩阵（不妨假设为 L）可分解为一个上三角阵和一个下三角阵的乘积，故而可只考虑 L 为上三角阵的情形。而一个上三角阵还可分解为一个对角线为 1 的上三角阵和一个对角阵的乘积。对角阵的情形比较容易，就是每个方向上做伸缩变换，即是 $L=(\delta_1,\ldots,\delta_n)$. 此时由于方体在这样的变换下任然是方体，因而方体的体积变换满足上述规律。

基于这些讨论，我们现在只需要考虑 E 为单个方体，L 为一个对角线为 1 的上三角阵。假设 $L=(L_{ij})$, $L_{ii}=1$, $L_{ij=0},i>j$. 对方体 E 作伸缩变换和平移还可不防假设 E 为单位方体 $\{(x_1,x_2,\ldots,x_n),0\leq x_i\leq 1\}$. 这样

$$L(E)=\{(x_1+L_{12}x_2+\ldots L_{1n}x_n,x_2+L_{23}x_3+\ldots+L_{2n}x_n,\ldots,x_n)|0\leq x_i\leq 1\}$$

这样根据定义以及 Fubini 定理

$$m(L(E))=\int\chi_{L(E)}=\int_{\mathbb{R}^{n-1}}\left(\int_{0\leq x_1\leq 1}\chi_{L(E)}\right)=\int_{\mathbb{R}^{n-1}}\chi_{L^{(1)}(E)}$$

这里

$$L^{(1)}(E)=\{(x_2+L_{23}x_3+\ldots+L_{2n}x_n,\ldots,x_n)|0\leq x_i\leq 1,2\leq i\leq n\}.$$

这样一直用 n 次 Fubini 定理可知

$$m(L(E))=1=m(E).$$

这就证明了对单位方体以及对角线为 1 的上三角阵 L，有 $m(L(E))=|detL|m(E)$。上述讨论表明对一般可测集 E 以及线性变换 L，测度转换关系式均成立。 □

由上面测度的变换关系，可以得到积分的在线性变换下的换元公式: 考虑非退化线性变换 $L:\mathbb{R}^d\to\mathbb{R}^d$ 以及 $\mathbb{R}^d$ 上的可积函数 $f(x)$，那么 $f(L(x))$ 也是 $\mathbb{R}^d$ 上的可积函数，并且

$$\int f(L(x))=\frac{1}{|detL|}\int f(x).$$

这里根据公式，非退化的条件是必要的。事实上如果 L 退化，函数 $f(L(E))$ 不一定可测。

接下来我们给出 Fubini 定理的应用。首先回忆卷积的定义：如果 f, g 是 $\mathbb{R}^d$ 上的光滑有紧支集的函数，那么我们可以定义卷积

$$f * g = \int f(x-y)g(y)dy.$$

可知 $f * g$ 也是具有紧支集的光滑函数。事实上卷积的定义可以推广到更加一般的情形即是 f, $g \in L^1$. 首先由习题结论 $f(x-y)$ 作为 $\mathbb{R}^{2d}$ 上的函数可测，并且由命题4.16的结论可知 $g(y)$ 作为 $\mathbb{R}^{2d}$ 上的函数也可测，因而 $f(x-y)g(y)$ 作为 $\mathbb{R}^{2d}$ 上的函数可测。由 Tonelli 定理以及积分的平移不变性可知

$$\int_{\mathbb{R}^{2d}} |f(x-y)||g(y)|dxdy = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-y)||g(y)|dx \right) dy = \|f\|_{L^1}\|g\|_{L^1} < \infty.$$

这就表明 $f(x-y)g(y)$ 在 $\mathbb{R}^{2d}$ 上可积。因而由 Fubini 定理可知对几乎所有的 x，$f(x-y)g(y)$ 作为 y 的函数在 $\mathbb{R}^d$ 上可积，并且积分 $f * g$ 是 $\mathbb{R}^d$ 上的可积函数。因而

$$\|f * g\|_{L^1} \leq \|f\|_{L^1}\|g\|_{L^1}.$$

如果 f, g 均为非负可积函数，上式等号成立。这表明卷积给出了 $L^1 \times L^1$ 到 L^1 的一个双线性连续泛函。卷积还满足如下性质

- 交换律：$f * g = g * f$;
- 结合律：$f * (g * h) = (f * g) * h$;
- 分配律：$f * (g + h) = f * g + f * h$.

现在我们可以用 Fubini 定理证明

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 (1-x)^{p-1}x^{q-1}dx, \quad \Gamma(s) = \int_0^\infty e^{-x}x^{s-1}dx, \quad p, q, s > 0.$$

事实上，由 Fubini 定理

$$\begin{aligned}
&\int_0^\infty e^{-x}x^{p-1}dx \cdot \int_0^\infty e^{-y}y^{q-1}dy \\
&= \iint_{x,y>0} e^{-(x+y)}(x+y)^{p+q-1}(1-\frac{y}{x+y})^{p-1}(\frac{y}{x+y})^{q-1}d(\frac{y}{x+y})d(x+y) \\
&= B(p,q)\Gamma(p+q)
\end{aligned}$$

利用 Fubini 定理，我们还可以计算 $\mathbb{R}^n$ 中半径为 r 的球的体积。由测度的平移以及伸缩变换性质，可知半径为 r 的球的体积为 $c_n r^n$。对单位球，根据 Fubini 定理有

$$c_n = \int_{-1}^1 c_{n-1}(1-x^2)^{\frac{n-1}{2}}dx = c_{n-1}\int_0^1 (1-x)^{\frac{n-1}{2}}x^{-\frac{1}{2}}dx = c_{n-1}B(\frac{n+1}{2}, \frac{1}{2}) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n+2}{2})}$$

由此我们还能得到半径为 r 的球面面积 S_n(我们还没有严格定义面积，类似于一维的曲线长度，高维曲面的面积可用平面逼近)。对 $\delta > 0$，近似的有

$$c_n(r+\delta)^n - c_n r^n \approx \delta \cdot S_n$$

特别的有 $\mathbb{R}^n$ 中半径为 r 的球面面积为 $nc_n r^{n-1}$。作为一个推论，对于径向对称的函数 $f(x) = g(|x|)$，积分可转换成

$$\int_{\mathbb{R}^n} f(x)dx = nc_n \int_0^\infty g(r)r^{n-1}dr.$$

之前我们计算过

$$\int_0^\infty e^{-x^2}dx = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{\sqrt{\pi}}{2}.$$

用 Fubini 定理

$$\int_0^\infty e^{-x^2}dx \cdot \int_0^\infty e^{-y^2}dy = \frac{1}{4}\int_{\mathbb{R}^2} e^{-(x^2+y^2)}dxdy = \frac{2\pi}{4}\int_0^\infty e^{-r^2}rdr = \frac{\pi}{4}.$$

这样也可以得到上面的积分公式。

例子 (丘赛 2022). 定义积分

$$I_n = \int_{[0,1]^n} \frac{n}{\frac{1}{x_1} + \dots \frac{1}{x_n}} dx_1 dx_2 \dots dx_n.$$

证明极限 $\lim\limits_{n\to\infty} I_n$ 的极限存在。

显然 I_n 是非负数列并且有上界，只需要证明

$$I_{n+m} \leq \frac{n}{m+n}I_n + \frac{m}{n+m}I_m$$

即可。

4.11 习题

北京大学数学分析 II 实验班作业 2

本次作业的提交时间和地点为 **3 月 12 日**习题课上，逾期视作零分。

1. 此题是 Riemann-Lebesgue 引理的推广。假设 f 在 $[-\pi,\pi]$ 上可积，这样可定义 Fourier 级数
$$f \sim \sum_{n\in\mathbb{Z}} \hat{f}_n e^{inx}, \quad \hat{f}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$
证明 $\hat{f}_n$ 有界，并且 $\lim\limits_{n\to\infty} \hat{f}_n = 0$.

2. 上题对应的连续情形是如下结论：假设 $f \in L^1(\mathbb{R}^d)$，定义 f 的 Fourier 变换
$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}dx.$$
证明 $\hat{f}$ 一致连续，有界，并且 $\lim\limits_{|\xi|\to\infty} \hat{f}(\xi) = 0$.

3. 如果 $f \in L^1(\mathbb{R})$，对于正整数 n 以及任意常数 $p > 0$，证明
$$\lim_{n\to\infty} n^{-p} f(nx) = 0, \quad a.e. x \in \mathbb{R}.$$

4. 假设 f 是 $\mathbb{R}^d$ 上的可积函数。对 $\alpha > 0$, 定义 $E_\alpha = \{x : |f(x)| > \alpha\}$. 证明
$$\int_{\mathbb{R}^d} |f(x)|dx = \int_0^\infty m(E_\alpha)d\alpha.$$

5. 在 $\mathbb{R}^2$ 中考虑集合 $A = \{0\} \times [0,1]$, $B = \mathcal{N} \times \{0\}$, 这里 $\mathcal{N}$ 是 $[0,1]$ 上的不可测集。证明 $A+B = \{x+y | x \in A, y \in B\}$ 是不可测的。此题表明对于可测集 A, B, 和集 $A+B$ 或者差集 $A-B$ 不一定是可测集。

6. 如果 $E_1 \subset \mathbb{R}^{d_1}$, $E_2 \subset \mathbb{R}^{d_2}$，那么乘积集 $E = E_1 \times E_2$ 在 $\mathbb{R}^{d_1+d_2}$ 中的外测度满足关系
$$m^*(E) \le m^*(E_1) \times m^*(E_2).$$
其实等号也成立。用等测包和 Fubini 定理。

7. 定义 $\delta = (\delta_1, \dots, \delta_d)$, $\delta_j \neq 0$, $1 \le j \le d$. 对于可测集 E，定义
$$E^\delta = \{(\delta_1 x_1, \dots, \delta_d x_d), (x_1, \dots, x_d) \in E\}.$$
证明 E^δ 可测，并且 $m(E^\delta) = |\delta_1| \dots |\delta_d| m(E)$.

8. 假设 f 是 $\mathbb{R}^d$ 上的可测函数，考虑图像 $\Gamma = \{(x,y) \in \mathbb{R}^d \times \mathbb{R}, y = f(x)\}$. 证明 Γ 是 $\mathbb{R}^{d+1}$ 上的零测集。之前的习题中表明如果 $f(x)$ 是连续的，那么图像是零测集。此题表明可以把连续函数推广到一般可测函数。

9. 给定 $\mathbb{R}^d$ 上的非负实值函数 $f(x)$，考虑 $\mathbb{R}^{d+1}$ 中图像下面的区域

$$\Omega_f = \{(x,y)|0 \le y \le f(x)\}.$$

证明 f 可测当且仅当 Ω_f 是 $\mathbb{R}^{d+1}$ 中的可测集，并且 $m(\Omega_f) = \int_{\mathbb{R}^d} f$.

10. 考虑非退化线性变换 $L: \mathbb{R}^d \to \mathbb{R}^d$ 以及 $\mathbb{R}^d$ 上的可积函数 $f(x)$，那么 $f(L(x))$ 也是 $\mathbb{R}^d$ 上的可积函数，并且

$$\int f(L(x)) = \frac{1}{|detL|} \int f(x).$$

11. 如果线性变换 L 退化，f 是 $\mathbb{R}^d$ 上的可测函数，举例说明 $f(L(x))$ 不一定可测。

12. 如果 f 是 $\mathbb{R}^d$ 上的可测函数，证明 $g(x,y) = f(x-y)$ 作为 $\mathbb{R}^{2d}$ 上的函数可测。

13. 证明在 L^1 中不存在一个函数 I 使得 $I * f = f, \forall f \in L^1$.

14. 如果 f 可积，可定义 Fourier 变换 $\hat{f}(\xi) = \int f(x) e^{-2\pi i x \cdot \xi} dx$. 证明如果 f, g 都可积那么

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

5 Lebesgue 微分定理

我们来回顾微积分里最重要的微积分基本定理：

- 函数 f 在 $[a,b]$ 上 Riemann 可积，在 $x_0 \in (a,b)$ 处连续，那么 $F(x) = \int_a^x f(y)dy$ 在 x_0 处可微，并且 $F'(x_0) = f(x_0)$.
- 如果 $F(x)$ 是定义在 $[a,b]$ 上的可微函数，并且 $F'(x)$ 在 $[a,b]$ 上 Riemann 可积，那么 $F(b) - F(a) = \int_a^b F'(x)dx$(这个用 Lagarangian 中值定理，再根据 Riemann 积分的定义).

这章的主要目的是讨论 Lebesgue 积分理论中上述微积分基本定理在怎样的条件下成立。

5.1 Lebesgue 微分定理

第一个问题相对而言容易一些，并且在高维有自然的推广。在 Riemann 积分的情况下, 第一个结论可以写成：如果 f 在 x 处连续，那么

$$\lim_{x\in I, |I|\to 0} \frac{1}{|I|}\int_I f(y)dy = f(x).$$

在高维可以将区间 I 换成开球 B（注意这里不可以是一般的开集，因为高维包含 x 的开集可以很复杂）。由于 Riemann 可积的函数几乎处处连续，因而上述极限对几乎处处的 x 成立。此结论对 Lebesgue 可积函数也成立

定理 5.1. *[Lebesgue 微分定理] 如果 $f \in L^1(\mathbb{R}^d)$，那么有*

$$\lim_{x\in B, m(B)\to 0} \frac{1}{m(B)}\int_B f = f(x), a.e.$$

如果 f 在 x 处连续，显然上面的平均极限存在并且等于 $f(x)$. 至于一般的可积函数，想法便是证明极限不存在或者极限不等于 $f(x)$ 的点集是零测集，而为了证明此集合是零测集，我们可以放缩一点，证明此集合的测度可以任意小。这样可用连续函数 $g(x)$ 逼近 $f(x)$, 连续函数 $g(x)$ 的平均趋于 $g(x)$，$g(x)$ 与 $f(x)$ 的差可由 $\|g-f\|$ 的 L^1 范数来控制，比较麻烦的是 $g-f$ 的平均。为了控制这一项，我们引入 Hardy-Littlewood 极大函数。

如果 $f \in L^1(\mathbb{R}^d)$，定义函数 $\mathcal{M}f(x)$

$$\mathcal{M}f(x) = \sup_{x\in B} \frac{1}{m(B)}\int_B |f|.$$

同样的这里 B 是开球。极大函数有如下性质

命题 5.2 (Hardy-Littlewood). *如果 $f \in L^1$，那么对应的极大函数 $\mathcal{M}f$ 满足如下性质：*

(1) $\mathcal{M}f$ 可测；

(2) 对几乎处处的 x，有 $|f(x)| \le \mathcal{M}f(x) < \infty$.

(3) 存在常数 A 使得对任意 $\alpha>0$，有

$$m(\{x:\mathcal{M}f(x)>\alpha\})\leq \frac{A}{\alpha}\|f\|_{L^1}.$$

这里最重要的是第三条。由 Tchebychev 不等式可知

$$m(\{|f(x)|>\alpha\})\leq \frac{1}{\alpha}\|f\|.$$

这表明尽管 f 的极大函数比 f 大，但在测度的意义 f 与 $\mathcal{M}f$ 是等价的。

证明*:* 对任意 $\alpha>0$, 定义 $E_\alpha=\{x:\mathcal{M}f>\alpha\}$。如果 $x\in E_\alpha$，那么根据定义存在包含 x 的开球 B_x 使得

$$\frac{1}{m(B_x)}\int_{B_x}|f|>\alpha.$$

这表明 $B_x\subset E_\alpha$，因而 E_α 为开集，故而 $\mathcal{M}f$ 可测。

对第二条，由于 f 可积，因而 f 几乎处处有限，故只需要对任意 $N>0,\varepsilon_0>0$，以下集合

$$\{x:N>|f(x)|>\mathcal{M}f(x)+\varepsilon_0\}$$

是零测集即可。否则对任意 $0<\lambda<1$，存在集合 E 以及开球 B 使得

$$E\subset B,\quad |E|>\lambda|B|.$$

再根据定义对任意 $x\in E$ 有

$$\mathcal{M}f(x)\geq \frac{1}{|B|}\int_B|f(y)|dy\geq \frac{1}{|B|}\int_E\mathcal{M}f(y)+\varepsilon_0dy=\frac{1}{|B|}\int_E\mathcal{M}f(y)dy+\frac{|E|\varepsilon_0}{|B|}.$$

根据第一问 $\mathcal{M}f$ 的可测性，在 E 上再积分可得

$$\int_E\mathcal{M}f(x)\geq \frac{|E|}{|B|}\int_E\mathcal{M}f(y)dy+\frac{|E|^2\varepsilon_0}{|B|}$$

而注意到在 E 上有 $\mathcal{M}f<N$，因而

$$(1-\frac{|E|}{|B|})|E|N\geq \frac{|E|^2\varepsilon_0}{|B|},\quad (1-\lambda)N\geq\lambda\varepsilon_0,\quad \forall 0<\lambda<1.$$

矛盾！这就表明 $|f(x)|\leq\mathcal{M}f(x), a.e.$。第二条结论剩下的只需要证明极大函数也是几乎处处有限的，注意到

$$\{\mathcal{M}f=+\infty\}\subset\{\mathcal{M}f>N\}.$$

故而极大函数的几乎处处有限性是结论 (3) 的推论。我们剩下只需要证明 (3)。对 $x\in E_\alpha$，总存在包含 x 的开球 B_x 使得

$$m(B_x)\leq\frac{1}{\alpha}\int_{B_x}|f|.$$

这样我们得到覆盖 E_α 的开球簇 $\{B_x\}_{x\in E_\alpha}$。现在的问题是如何取出合适的不相交的开球使得能盖住足够大的 E_α。这章我们会用到不同形式的覆盖引理，简单起见，我们先做一个简化：为了估计开集 E_α 测度大小，我们只需要估计任何包含在 E_α 中的紧集的测度大小。由开覆盖定理，紧集可被有限个开集覆盖住，因而对这有限个开集，能否取出足够大的不相交的开集。我们有如下有限情形的覆盖引理：

引理 5.3. *对于 $\mathbb{R}^d$ 中有限个开球 $\{B_1, B_2, \ldots, B_N\}$。存在其中的 l 个互不相交的开球 $\{B_{n_1}, \ldots, B_{n_l}\}$ 使得*

$$m(\bigcup_{j=1}^{N} B_j) \le 3^d m(\bigcup_{i=1}^{l} B_{n_i})$$

证明的想法就是取出足够大的但互不相交的开球使得取出的开球半径放大三倍能盖住所有开球的并：自然的想法就是从半径最大的开球选取。不妨假设 B_j 的半径 r_j 是递减的，这样我们最先取出 $B_{n_1} = B_1$, 然后在与 B_1 不相交的开球里取半径最大的开球作为 B_{n_2}. 一般的如果取出 $B_{n_1}, \ldots, B_{n_k}$，那么在所有与这些已经取出的开球不相交的开球里取出半径最大者作为 $B_{n_{k+1}}$. 这个过程有限步便结束，这样可得互不相交的开球 $B_{n_1}, \ldots, B_{n_l}$. 记 $3B_j$ 为半径是 B_j 三倍的同心球。我们只需要证明

$$\bigcup_{j=1}^{N} B_j \subset \bigcup_{j=1}^{l} 3B_{n_j}$$

即可。否则存在 $x \in B_j$ 但 $x \notin 3B_{n_i}, \forall 1 \le i \le l$. 根据选取规则，$B_j$ 必定跟某个被选出的球 B_{n_i} 相交，不妨假设 B_{n_i} 是所有被选出的开球中与 B_j 相交的半径最大球（如果半径相同，则取下标最小的）。因而 B_j 与 $B_{n_1}, \ldots, B_{n_{i-1}}$ 不相交，这表明 $r_j \le r_{n_i}$，这样必然有 $B_j \subset 3B_{n_i}$. 这就证明了引理。

回到命题结论 (3) 的证明，对任意紧集 $K \subset E_\alpha$, 存在有限覆盖 $B_1, \ldots, B_N$. 由上述覆盖引理，存在互不相交的开球 $B_{n_1}, \ldots, B_{n_l}$ 使得 $m(\bigcup_{j=1}^{N} B_j) \le 3^d m(\bigcup_{i=1}^{l} B_{n_i})$. 这样

$$m(K) \le m(\bigcup_{j=1}^{N} B_j) \le 3^d \sum_{i=1}^{l} m(\bigcup_{i=1}^{l} B_{n_i}) \le \frac{3^d}{\alpha} \|f\|_{L^1}.$$

□

现在我们就可以证明 Lebesgue 微分定理。对任意 $\varepsilon > 0$, 取连续函数 $g \in C_0$ 使得 $\|g - f\|_{L^1} < \varepsilon$. 由 g 的连续性，对任意 $\varepsilon_0 > 0$, 存在 $\delta > 0$ 使得当 $x \in B$, $m(B) < \delta$ 时，有 $|g(y) - g(x)| < \varepsilon_0$, $\forall y \in B$. 这样根据定义可知

$$\begin{aligned}\frac{1}{m(B)} \int_B |f(y) - f(x)| dy &\le \frac{1}{m(B)} \int_B |g(y) - g(x)| + |f(y) - g(y)| dy + |g(x) - f(x)| \\ &\le \mathcal{M}(f - g)(x) + |g(x) - f(x)| + \varepsilon_0.\end{aligned}$$

对任意 $\alpha > 0$, 定义

$$A_\alpha = \{x : |g(x) - f(x)| > \alpha\}, \quad E_\alpha = \{x : \mathcal{M}(f-g)(x) > \alpha\}.$$

上述命题导出 $m(E_\alpha) \le \frac{C}{\alpha}\|f-g\|_{L^1} \le \frac{C\varepsilon}{\alpha}$. 再由 Tchebychev 不等式可知 $m(A_\alpha) \le \frac{1}{\alpha}\|g-f\|_{L^1} \le \frac{\varepsilon}{\alpha}$. 因而

$$m(\{x : \mathcal{M}(f-g)(x) + |g(x) - f(x)| + \varepsilon_0 > 3\varepsilon_0\}) \le \frac{(1+C)\varepsilon}{\varepsilon_0}.$$

因而对任意 $\varepsilon_0 > 0, \varepsilon > 0$ 都有

$$m(\{x : \limsup_{x\in B, m(B)\to 0} \frac{1}{m(B)} \int_B |f(y) - f(x)|dy > 3\varepsilon_0\}) \le \frac{(1+C)\varepsilon}{\varepsilon_0}.$$

由 $\varepsilon, \varepsilon_0$ 的任意性可知

$$m(\{x : \lim_{x\in B, m(B)\to 0} \frac{1}{m(B)} \int_B |f(y) - f(x)|dy \neq 0\}) = 0.$$

这就证明了 Lebesgue 微分定理。

我们把满足 $|f(x)| < \infty$ 并且

$$\lim_{x\in B, m(B)\to 0} \frac{1}{m(B)} \int_B |f(y) - f(x)|dy = 0$$

的点 x 组成的集合称为 Lebesgue 集。Lebesgue 微分定理告诉我们，如果 $f \in L^1$, 那么几乎所有的点都属于 Lebesgue 集。事实上条件还可以放松到更大的函数空间：注意到 Lebesgue 点集的定义是局部性质，因而可以只要求 f 是局部可积的，即是对任意有限测度集 E，$f\chi_E$ 可积即可。典型的例子有 $(1+|x|)^{-d}$. 我们把所有局部可积的函数组成的空间记为 $L^1_{loc}(\mathbb{R}^d)$. 显然 Lebesgue 微分定理的结论可以推广到所有 $f \in L^1_{loc}(\mathbb{R}^d)$.

特别的考虑可测集 E，那么有

$$\frac{1}{m(B)} \int \chi_E = \frac{m(B\cap E)}{m(B)}$$

称为 E 在 B 中的密度。如果有

$$\lim_{x\in B, m(B)\to 0} \frac{m(B\cap E)}{m(B)} = 1$$

则称 x 为 E 的密度点（也成稠密点或者全密点）。Lebesgue 微分定理有如下推论

命题 5.4 (Lebesgue 密度定理). *对可测集 E, 几乎所有 E 中的点都是 E 的密度点。几乎所有 E 之外的点都不是 E 的密度点。*

这节的最后我们来看与卷积相关的应用。在 Lebesgue 微分定理里取球 $B(x,\varepsilon)$ 为中心在 x 半径为 ε 的球，那么有

$$\frac{1}{B(x,\varepsilon)}\int_{B(x,\varepsilon)} f = c_d^{-1}\varepsilon^{-d}\int_{B(0,\varepsilon)} f(x-y)dy = \int f(x-y)c_d^{-1}\varepsilon^{-d}\chi_{B(0,\varepsilon)}dy.$$

这里 c_d 为 $\mathbb{R}^d$ 中单位球的体积。定义 $K_\varepsilon = c_d^{-1}\varepsilon^{-d}\chi_{B(0,\varepsilon)}$，那么 Lebesgue 微分定理告诉我们 $K_\varepsilon * f$ 几乎处处收敛到 $f(x)$ 当 $\varepsilon \to 0$ 时。我们称 $K_\varepsilon(x)$ 为卷积核，一般的如果卷积核满足如下条件：

$$\int K_\varepsilon = 1, \quad |K_\varepsilon| \le A\min\{\varepsilon^{-d}, \varepsilon|x|^{-d-1}\}$$

其中 A 为某个常数，那么我们称 K_ε 为恒同逼近。用处是希望 $K_\varepsilon * f$ 在某种意义下趋于 f. 上面积分为 1 的条件原因在于 $\|f * K_\varepsilon\|_{L^1} = \|f\|_{L^1}\|K_\varepsilon\|_{L^1}$. 逐点的条件主要是要求无穷远卷积核衰减足够快，而在 0 附近一致有界。

我们来看几个例子.

- 一般的可取非负光滑函数 $\varphi(x)$ 使得支集在 $\{|x| \le 2\}$ 并且积分为 1。这样可定义恒同逼近 $\varphi_\varepsilon = \varepsilon^{-d}\varphi(\varepsilon^{-1}x)$, 这样的核我们称为 Friedrichs 光滑化核。

- 热核：考虑热方程初值问题

  $$(\partial_t - \Delta)u(t,x) = 0, \quad u(0,x) = u_0.$$

  这里 $\Delta = \sum\limits_{i=1}^{d}\partial_{x_i}^2$ 是 $\mathbb{R}^d$ 上的 Laplace 算子。方程的解可以写成

  $$u(t,x) = H_t(x) * u_0(x)$$

  其中 $H_t(x)$ 是热核

  $$H_t(x) = (4\pi t)^{-\frac{d}{2}}e^{-\frac{|x|^2}{4t}}.$$

  可以验证当 $t > 0$ 时候 $(\partial_t - \Delta)H_t(x) = 0$。为了说明 $H_t * u_0$ 满足热方程初值问题，还需要证明 $H_t * u_0 \to u_0$ 当 $t \to 0$ 时，此时取 $\varepsilon = \sqrt{t}$ 可知 $H_t(x)$ 是恒同逼近。

- Poisson 核：考虑边值问题

  $$\Delta u(x,y) = 0, \quad u(x,0) = u_0(x), \quad x \in \mathbb{R}^d, y \in \mathbb{R}.$$

  这里 Δ 是 $\mathbb{R}^{d+1}$ 中的 Laplace 算子。这里方程的解可以写成

  $$u(x,y) = P_y(x) * u_0(x)$$

  其中 $P_y(x)$ 为 Poisson 核

  $$P_y(x) = \frac{c_d y}{(y^2+|x|^2)^{\frac{d+1}{2}}}, \quad c_d = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}}$$

同样的可验证当 $y^2+|x|^2\neq 0$ 时 $\Delta P_y(x)=0$. 因而在上半平面 $y>0$ 上有 $\Delta P_y(x)*u_0=0$. 为了验证边界值，需要证明 $P_y(x)*u_0(x)\to u_0(x)$ 当 $y\to 0$ 时。此时取 $\varepsilon=y$ 可知 $P_y(x)$ 为恒同逼近。

由于卷积具有光滑函数的效果，如果 K_ε 是恒同逼近，那么 $K_\varepsilon*f$ 在 L^1 的意义下确实收敛到 f。

命题 5.5. 如果 K_ε 为恒同逼近，那么对任意可积函数 f，有 $K_\varepsilon*f$ 可积，并且

$$\lim_{\varepsilon\to 0}\|K_\varepsilon*f-f\|_{L^1}=0.$$

特别的如果 K_ε 光滑，那么 $K_\varepsilon*f$ 也是光滑函数，这表明光滑函数在可积函数空间中是稠密。

证明: 注意到

$$\begin{aligned}\|K_\varepsilon*f-f\|_{L^1}&=\int|\int K_\varepsilon(y)(f(x-y)-f(x))dy|dx\\&\leq\int|K_\varepsilon(y)|\|\tau_{-y}f(x)-f(x)\|_{L^1_x}dy.\end{aligned}$$

这里 $\tau_h f$ 是平移算子。由于平移算子的连续性，可知对任意 $\varepsilon_0>0$，存在 $\delta>0$ 使得

$$\|\tau_h f-f\|_{L^1}<\varepsilon_0,\quad\forall|h|<\delta.$$

这样

$$\begin{aligned}&\int|K_\varepsilon(y)|\|\tau_{-y}f(x)-f(x)\|_{L^1_x}dy\\&=\int_{|y|<\delta}|K_\varepsilon(y)|\|\tau_{-y}f(x)-f(x)\|_{L^1_x}dy+\int_{|y|\geq\delta}|K_\varepsilon(y)|\|\tau_{-y}f(x)-f(x)\|_{L^1_x}dy\\&\leq\varepsilon_0\int_{|y|<\delta}|K_\varepsilon(y)|dy+\int_{|y|>\delta}|K_\varepsilon(y)|2\|f\|_{L^1}dy\\&\leq\varepsilon_0\int_{|y|<\delta}A\varepsilon^{-d}dy+2\|f\|_{L^1}\int_{|y|>\delta}A\varepsilon|y|^{-d-1}dy\\&\leq\varepsilon_0 Ac\varepsilon^{-d}\delta^d+2A\|f\|_{L^1}\varepsilon c\delta^{-1}.\end{aligned}$$

注意到 ε 和 ε_0 可以任意小，δ 依赖于 ε_0. 因而只要取 ε_0 以及 ε 使得

$$\varepsilon_0=\varepsilon^{d+1}\delta^{-d-1}$$

便知

$$\int|K_\varepsilon(y)|\|\tau_{-y}f(x)-f(x)\|_{L^1_x}dy\leq(1+2\|f\|_{L^1})Ac\varepsilon_0^{\frac{1}{d+1}}.$$

此即表明命题结论。 □

除了上面的 L^1 收敛，有时候也关心逐点收敛。如果 f 一致连续，很容易证明逐点收敛。另外一方面，由于可积函数在几乎处处的意义下相同，因而在逐点收敛方面我们最多可期望 $K_\varepsilon*f$ 几乎处处收敛到 f. 事实上在 Lebesgue 集上，逐点收敛确实成立。

命题 5.6. *如果 f 可积，K_ε 为恒同逼近，那么在 f 的 Lebesgue 点 x 处有*

$$\lim_{\varepsilon\to 0}(K_\varepsilon * f)(x) = f(x).$$

证明: 定义

$$\omega(r) = \frac{c_d}{m(B(0,r))}\int_{B(0,r)} |f(x-y)-f(x)|dy, \quad \forall r>0.$$

这里 c_d 为单位球的体积。如果 x 为 f 的 Lebesgue 点，那么有

$$\lim_{r\to 0}\omega(r) = 0.$$

这样我们可以估计

$$\begin{aligned}
|K_\varepsilon * f(x) - f(x)| &\le \int |K_\varepsilon(y)||f(x-y)-f(x)|dy \\
&\le \int_{B(0,r)} |K_\varepsilon(y)||f(x-y)-f(x)|dy + \sum_{k\ge 0}\int_{2^k r\le |y|<2^{k+1}r} |K_\varepsilon(y)||f(x-y)-f(x)|dy \\
&\le A\varepsilon^{-d}r^d\omega(r) + \sum_{k\ge 0}\int_{2^k r\le |y|<2^{k+1}r} A\varepsilon|y|^{-d-1}|f(x-y)-f(x)|dy \\
&\le A\varepsilon^{-d}r^d\omega(r) + \sum_{k\ge 0} A\varepsilon 2^{-k(d+1)}r^{-d-1}2^{(k+1)d}r^d\omega(2^{k+1}r) \\
&\le A\varepsilon^{-d}r^d\omega(r) + A2^d\varepsilon r^{-1}\sum_{k\ge 0}2^{-k}\omega(2^{k+1}r).
\end{aligned}$$

注意到当 $r>0$ 时 $\omega(r)$ 是关于 r 的连续函数。另一方面，

$$\omega(r) \le \frac{1}{r^d}\int |f(x-y)| + |f(x)| \le \frac{\|f\|_{L^1}}{r^d} + |f(x)|.$$

这样可取 $\varepsilon = r$ 便有

$$|K_\varepsilon * f(x) - f(x)| \le A\omega(\varepsilon) + A2^d\sum_{k\ge 0}2^{-k}\omega(2^{k+1}\varepsilon) \to 0.$$

$\square$

5.2 习题

北京大学数学分析 II 实验班作业 3

本次作业的提交时间和地点为 **3 月 19 日**习题课上，逾期视作零分。

1. 此题的目的是定义弱 L^1 空间 L^1_w: 如果 f 可测，定义

$$\lambda_f(t) = m(\{x : |f(x)| > t\}).$$

再定义

$$\|f\|_{L^1_w} = \sup_{t>0}(t\lambda_f(t)).$$

L^1_w 是所有满足 $\|f\|_{L^1_w} < \infty$ 的函数组成的空间。证明 $L^1 \subset L^1_w$, 即是证明如果 f 可积，那么 $f \in L^1_w$.

2. 证明 L^1_w 是线性空间，但上题定义的 $\|\cdot\|_{L^1_w}$ 不是范数，而是拟范数，即是证明

$$\|f+g\|_{L^1_w} \leq 2(\|f\|_{L^1_w} + \|g\|_{L^1_w}), \quad \forall f, g \in L^1_w$$

并且这个不等式右边的常数 2 不能更小。这里回忆范数满足右边常数为 1 的三角不等式。

3. 更进一步我们不能在 L^1_w 中找到一个与 $\|\cdot\|_{L^1_w}$ 范等价的范数，即是证明不存在 L^1_w 的范数 $\|\cdot\|$ 以及常数 C 使得

$$C^{-1}\|f\| \leq \|f\|_{L^1_w} \leq C\|f\|, \quad \forall f \in L^1_w.$$

4. 如果 f 可积，并且不是几乎处处为 0, 那么存在常数 $c > 0$ 使得

$$\mathcal{M}f(x) > c|x|^{-d}, \quad \forall |x| > 1.$$

这表明 $\mathcal{M}f(x)$ 不是可积的。

5. 上题表明极大函数不可积，原因在于无穷远处衰减太慢。此题表明存在局部不可积的极大函数：考虑函数

$$f = x^{-1}|\log x|^{-2}\chi_{\{0<x<\frac{1}{2}\}}.$$

证明 f 可积但存在常数 $c > 0$ 使得

$$\mathcal{M}f(x) > cx^{-1}|\log x|^{-1}, \quad \forall 0 < x < \frac{1}{2}.$$

特别的对应的极大函数在 0 处不可积。

6. 如果 E 是 $[0,1]$ 上的可测集满足对任意 $[0,1]$ 上的区间 I 以及某个 $\alpha > 0$ 都有 $m(E \cap I) \geq \alpha m(I)$，证明 $m(E) = 1$.

7. 假设 E 为 $\mathbb{R}^d$ 中的可测集，对任意 $x \in \mathbb{R}^d$, 回忆距离函数 $d(x,E) = \inf_{y\in E} |x-y|$. 证明对几乎所有的 $x \in E$ 有

$$\lim_{y\to 0} \frac{d(x+y,E)}{|y|} = 0.$$

8. 课上我们称满足如下条件

$$\int K_\varepsilon = 1, \quad |K_\varepsilon| \le A\min\{\varepsilon^{-d}, \varepsilon|x|^{-d-1}\}$$

的卷积核 K_ε 为恒同逼近。证明热核

$$H_t = (4\pi t)^{-\frac{d}{2}} e^{-\frac{|x|^2}{4t}}$$

为恒同逼近 (取 $\varepsilon = \sqrt{t}$).

9. 证明 Poisson 核

$$P_y(x) = \frac{c_d y}{(y^2+|x|^2)^{\frac{d+1}{2}}}, \quad c_d = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}}$$

为恒同逼近 (取 $\varepsilon = y$).

10. 如果 K_ε 为恒同逼近，证明存在常数 c 使得对所有可积函数 f 有

$$\sup_{\varepsilon>0} |(K_\varepsilon * f)(x)| \le c\mathcal{M}f(x), \quad \forall x \in \mathbb{R}^d.$$

5.3 单调函数的 Lebesgue 微分定理

接下来我们讨论第二个问题：此时我们给定函数 $F(x)$，那么何时 $F(x)$ 的导数 Lebesgue 可积并且满足等式 $\int_{[a,b]} F'(x)dx = F(b) - F(a)$. 这个问题相对第一个问题更加复杂，我们只讨论一维的情形，高维的情形可参考 Evans–Gariepy, Measure Theory and Fine Properties of Functions. 我们先来看几个例子：

- Heaviside 函数 $H(x) = \chi_{(0,\infty)}$：这个函数在 0 处有跳跃，导数在 $x \neq 0$ 处为 0, 但 $H(1) - H(0) = 1$。因而在经典 Lebesgue 积分的意义下，微积分基本定理不成立（不过在广义函数意义下成立）。
- Cantor-Lebesgue 函数 $F(x)$: 回忆此函数的定义可知在 Cantor 集合 $\mathcal{C}$ 之外 $F' = 0$。因而 F 几乎处处可导，并且导数为 0. 不过同样的 $F(1) - F(0) = 1$ 不满足微积分基本定理。
- 考虑函数 $F(x) = x^2 \sin \frac{1}{x^2}$，$F(0) = 0$. 当 $x \neq 0$ 时 $F'(x) = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}$ 而 $F'(0) = 0$. 因而 $F(x)$ 处处可微，但 $F'(x)$ 不是 Lebesgue 可积的。

我们先来讨论函数的可微性。为此我们把导数分解成 Dini 导数。定义

$$\Delta_h(F)(x) = \frac{F(x+h) - F(x)}{h}, \quad h \neq 0.$$

然后定义

$$D^+(F)(x) = \limsup_{h \to 0, h > 0} \Delta_h(F)(x), \quad D_+(F)(x) = \liminf_{h \to 0, h > 0} \Delta_h F(x),$$
$$D^-(F)(x) = \limsup_{h \to 0, h < 0} \Delta_h(F)(x), \quad D_-(F)(x) = \liminf_{h \to 0, h < 0} \Delta_h F(x).$$

显然 $F(x)$ 在 x 处的导数存在当且仅当四个 Dini 导数有限且相等。我们还可以定义上下导数：

$$\overline{D}(F)(x) = \max\{D^+(F)(x), D^-(F)(x)\}, \quad \underline{D}(F)(x) = \min\{D_+(F)(x), D_-(F)(x)\}.$$

我们这节这要证明下面的单调函数 Lebesgue 微分定理。

定理 5.7 (单调函数的 Lebesgue 微分定理)**.** *F 是定义在 $[a,b]$ 上的单调递增函数，那么导数 $F'(x)$ 几乎处处存在，并且是非负 Lebesgue 可积的满足*

$$\int_{[a,b]} F'(x) \leq F(b) - F(a).$$

类似于 Heaviside 函数，如果函数 F 有跳跃点，等式显然不成立，而 Cantor-Lebesgue 函数说明即使是连续的单调函数，等式也不成立。

定理的证明分为如下几步：首先可知单调递增函数的不连续点是可数个，因而可把一般单调增函数分解成跳跃函数和连续单调函数的和。第二步证明跳跃函数的导数几乎处处为 0（类似于

Cantor-Lebesgue 函数，由无穷多段阶梯函数组成）。最后证明连续增函数的几乎处处可导数，并且满足积分不等式。

第一步：单调增函数存在左右极限

$$F(x+0)=\lim_{h\to 0,h>0}F(x+h),\quad F(x-0)=\lim_{h\to 0,h>0}F(x-h).$$

$F(x)$ 在 x 处连续等价于 $F(x+0)=F(x-0)$. 由之前习题结论可知单调增函数的不连续点集至多是可数集，因而不妨假设 F 的不连续点为 $\{x_n\}$. 记 $\alpha_n=F(x+0)-F(x-0)$ 称为 F 在 x 处的跃度。这样可知 $F(x_n)=F(x-0)+\beta_n$, $0\le\beta_n\le\alpha_n$. 定义 $j_n(x)=0$ 当 $x<x_n$ 时，$j_n(x)=\alpha_n$ 当 $x>x_n$ 时, $j_n(x_n)=\beta_n$. 再定义函数 $J(x)=J_F(x)$

$$J(x)=J_F(x)=\sum j_n(x).$$

注意到 $j_n(x)\le\alpha_n$ 并且 $\sum\alpha_n\le F(b)-F(a)<\infty$，故而上述级数绝对收敛。这样做的目的是将 F 的间断点分解出来。

引理 5.8. 如果 F 是定义在 $[a,b]$ 上的单调增函数，那么 $F(x)-J_F(x)$ 为连续单调增函数。

证明: 首先我们证明 $F(x)-J_F(x)$ 连续:

$$\lim_{h\to 0,h>0}(F(x\pm h)-J_F(x\pm h))=F(x\pm 0)-\lim_{h\to 0,h>0}J_F(x\pm h).$$

在 $F(x)$ 的连续点 x 处有 $F(x+0)=F(x-0)$. 根据定义此时有

$$J_F(x)=\sum j_n(x)=\sum_{x_n<x}j_n(x)+\sum_{x_n>x}j_n(x)=\sum_{x_n<x}\alpha_n.$$

由于 $x_n\ne x$ 可知

$$\lim_{h\to 0,h>0}J_F(x\pm h)=J_F(x).$$

这样可得在 F 的连续点处 $F(x)-J_F(x)$ 也连续。

在 x_n 处，对 $h>0$ 可知

$$J_F(x_n+h)=\sum_{x_k<x_n}\alpha_k+\alpha_n+\sum_{x_n<x_k\le x_n+h}j_k(x_n+h),$$
$$J_F(x_n-h)=\sum_{x_k<x_n-h}\alpha_k+\sum_{x_k=x_n-h}j_k(x_n-h).$$

对任意 $x_k>x_n$ 都存在足够小的 h 使得 $x_k>x_n+h$. 因而

$$\lim_{h\to 0,h>0}\sum_{x_n<x_k\le x_n+h}j_k(x_n+h)=0.$$

另外一方面显然

$$\sum_{x_k \le x_n - h} j_k(x_n - h) \le \sum_{x_k < x_n} \alpha_k, \quad \lim_{h\to 0, h>0} \sum_{x_k < x_n - h} \alpha_k \ge \sum_{x_k < x_n} \alpha_k.$$

这表明

$$\lim_{h\to 0, h>0} F(x+h) - J_F(x+h) = \lim_{h\to 0, h>0} (F(x-h) - J_F(x-h)).$$

单调性: 对任意 $a \le x < y \le b$, 如果 x 为 F 的连续点，那么有

$$J_F(y) - J_F(x) = \sum_{x_k \le y} j_k(y) - \sum_{x_k < x} j_k(x) = \sum_{x < x_k \le y} j_k(y) \le F(y) - F(x).$$

如果 $x = x_n < y$ 那么有

$$\begin{aligned} J_F(y) - J_F(x) &= \sum_{x_k \le y} j_k(y) - \sum_{x_k < x_n} \alpha_k - \beta_n = \sum_{x_n \le x_k \le y} j_k(y) - \beta_n \\ &\le F(y) - F(x_n + 0) + \alpha_n - \beta_n \le F(y) - F(x_n). \end{aligned}$$

□

上述引理表明单调增函数可以分解为一个跳跃函数 $J(x)$ 和一个连续的单调函数的和。接下来我们证明跳跃函数几乎处处可导并且导数为 0.

命题 5.9. *假设 $J(x)$ 为跳跃函数，那么 $J(x)$ 几乎处处的导数为 0.*

证明*:* 由于 $J(x)$ 单调增，四个 Dini 导数存在并且非负。为证上述命题只需要证明 $\{\overline{D}(J)(x) > 0\}$ 为零测集即可。对任意 $\varepsilon > 0$, 考虑集合

$$E_\varepsilon = \{x : \overline{D}(J)(x) > \varepsilon\}.$$

我们只需要证明 E_ε 为零测集。对任意 $\delta > 0$，取 N 足够大使得

$$\sum_{k>N} \alpha_k < \frac{\varepsilon\delta}{10}.$$

并考虑跳跃函数

$$J_N(x) = \sum_{k>N} j_k(x)$$

以及集合

$$E_{\varepsilon,N} = \{x : \overline{D}(J_N)(x) > \varepsilon\}.$$

我们观察到

$$E_\varepsilon \subset E_{\varepsilon,N} \cup \{x_1, x_2, \dots, x_N\}.$$

因而只需要证明 $E_{\varepsilon,N}$ 的外测度足够小即可。对任意 $y \in E_{\varepsilon,N}$, 存在任意小的 h 使得 $\frac{J_N(y+h)-J_N(y)}{h} > \varepsilon$. 特别的由单调性，可取任意小的 $h > 0$ 使得

$$J_N(y+h) - J_N(y-h) > \varepsilon h.$$

这表明有开区间簇 $(y-h, y+h)$ 盖住 $E_{\varepsilon,N}$。我们希望能找到不相交的开区间 $(y_j - h_j, y_j + h_j)$ 使得能盖住 $E_{\varepsilon,N}$ 的固定部分，同时用 $J_N(b) - J_N(a)$ 的小性来控制这些互不相交的区间的总长度。为此我们需要如下的覆盖引理: 我们可以找到互不相交的开区间 $(y_j - h_j, y_j + h_j)$ 使得

$$m^*(E_{\varepsilon,N}) \le m(\bigcup_{y \in E_{\varepsilon,N}} (y-h, y+h)) \le 5 \sum_{j=1}^{\infty} 2h_j.$$

由上面分析可知

$$h_j < \varepsilon^{-1}(J_N(y_j + h_j) - J_N(y_j - h_j)).$$

由于区间 $(y_j - h_j, y_j + h_j)$ 互不相交，可得

$$\sum_j h_j \le \varepsilon^{-1} \sum_j (J_N(y_j + h_j) - J_N(y_j - h_j)) \le \varepsilon^{-1}(J_N(b) - J_N(a)) \le \frac{\delta}{10}.$$

故而 $m^*(E_\varepsilon) \le m^*(E_{\varepsilon,N}) < \delta$. 这就证明了跳跃函数几乎处处可导并且导数为 0. □

引理 5.10. 假设 B_α 为 $\mathbb{R}^d$ 中一族半径有界的球，则存在其中的互不相交的子簇 B_j 使得

$$m(\bigcup_\alpha B_\alpha) \le 5^d m(\bigcup_{j=1}^{\infty} B_j).$$

证明: 证明与之前有限个球的证明类似。记 $r(B)$ 为球 B 的半径, 由条件

$$R_0 = \sup_\alpha r(B_\alpha) < \infty.$$

首先取 B_1 为一个球使得 $\frac{1}{2}R < r(B_1) \le R$. 然后去掉所有与 B_1 相交的球, 在剩下的球中做同样的操作得到 B_2，这样可以得到球 B_j。如果 $\sum\limits_{j=1} m(B_j) = \infty$ 那么引理显然成立。否则，$m(B_j) \to 0$(如果只取出了有限个球，那么可定义 B_j 为空集). 我们证明在这种情况下把这些选出的球扩大 5 倍能盖住所有原来的球。任取 B_α 中的某个球 B。由于 $r(B_j) \to 0$, 存在 N 使得

$$r(B_j) < \frac{1}{2} r(B), \quad \forall j > N.$$

由 B_N 的选取可知 B 要么在之前被选出要么在之前某一步的选取过程中被去掉。如果在某一步被选出显然有 $B \subset B$. 否则假设在第 k 步被去掉，那么根据规则 B 与 B_k 相交。由于 B_k 被选出，可知

$$r(B_k) > \frac{1}{2} r(B).$$

这表明将 B_k 扩大 5 倍能盖住 B. □

接下来我们证明连续单调增函数的可微性。假设 f 是定义在 $[a,b]$ 上单调函数，我们证明 f 在 $[a,b]$ 上几乎处处可微，为此我们只需要证明

命题 5.11. 如果 F 是定义在闭区间 $[a,b]$ 上的单调增有界函数，那么

$$D^+(F)(x) < \infty, a.e., \quad D^+(F)(x) \leq D_-(F)(x), a.e.$$

特别的如果考虑 $-F(-x)$ 可知 $D^-(F) \leq D_+(F)(x), a.e.$。这样可推出 $D^+ \leq D_- \leq D^- \leq D_+, a.e.$, 因而四个 $Dini$ 导数几乎处处相等，也就是 $F(x)$ 几乎处处可微。

证明: 我们先证明第一部分 $D^+(F) < \infty, a.e.$. 对任意 $\gamma > 0$ 定义集合

$$E_\gamma = \{x \in [a,b] : D^+(F)(x) > \gamma\}.$$

由于 F 单调增当 $h \in [\frac{1}{n+1}, \frac{1}{n}]$ 时

$$\frac{n}{n+1}\frac{F(x+\frac{1}{n+1})-F(x)}{\frac{1}{n+1}} \leq \frac{F(x+h)-F(x)}{h} \leq \frac{n+1}{n}\frac{F(x+\frac{1}{n})-F(x)}{\frac{1}{n}}$$

因而根据定义有

$$D^+(F)(x) = \limsup_{h\to 0, h>0}\frac{F(x+h)-F(x)}{h} = \limsup_{n\to\infty}\frac{F(x+\frac{1}{n})-F(x)}{\frac{1}{n}}.$$

这就表明 $D^+(F)$ 是可测函数，故而 E_γ 是可测集。我们接下来证明

$$\lim_{\gamma\to\infty} m(E_\gamma) = 0$$

证明的基本想法也是用一种覆盖引理：如果 $x \in E_\gamma$ 那么存在 $h > 0$ 使得 $F(x+h)-F(x) \geq \gamma h$. 由于 F 是连续的,因而可以找到开区间 $(x-h, x+h)$ 盖住 x 使得 $F(x+h)-F(x-h) > \frac{1}{3}\gamma h$（不同的 h). 这样可得 E_γ 的开区间覆盖,再在这些开区间里选取一些互不相交的便可知 $\frac{\gamma}{3}m(E_\gamma) \leq F(b)-F(a)$。我们也可用如下的 Riesz 日出引理。

引理 5.12 (Riesz 日出引理)**.** 假设 G 是 $\mathbb{R}$ 上的实值连续函数，定义 $E = \{x : \exists h > 0\ s.t. G(x+h) > G(x)\}$. 那么 E 是开集，因而可以写成可数个开区间的并 $E = \bigcup\limits_{k=1}^{\infty}(a_k, b_k)$ 并且如果 (a_k, b_k) 是有限的开区间那么有 $G(a_k) = G(b_k)$. 特别的如果 G 只是定义在某个闭区间 $[a,b]$ 上，那么结论仍然成立，只是对于 $a_k = a$ 的情形有 $G(a) \leq G(b_k)$.

引理的证明:对任意 $x \in E$,存在 $h > 0$ 使得 $G(x+h) > G(x)$。由于 G 的连续性可知存在 $\delta > 0$ 使得对任意 $|y-x| < \delta$ 有 $|G(y)-G(x)| < \frac{1}{2}(G(x+h)-G(x))$。特别的对任意 $|y-x| < \min\{\delta, h\}$ 有 $G(x+h) > G(y)$，故而根据定义有 $y \in E$. 这就表明集合 E 是开集。现在假设 $(a,b) \subset E$ 是 E 中的一个极大开区间，即是 $a \notin E$, $b \notin E$. 根据定义显然有 $G(a) \geq G(b)$. 如果 $G(a) > G(b)$，由于 $b \notin E$，故而有 $G(x) \leq G(b), \forall x > b$. 取足够小的 $0 < \epsilon < b-a$ 使得 $G(a+\epsilon) > G(b)$. 考虑 G 在区间 $[a+\epsilon, b]$ 上的极大值点 x_ϵ, 由条件可知 $G(b) < G(x_\epsilon) \leq G(a)$. 由于 $x_\epsilon \in [a+\epsilon, b)$ 因而根据

定义存在 $h>0$ 使得 $G(x_\epsilon+h)>G(x_\epsilon)$. 由于 $G(x)\le G(x),\forall x>b$, 这就表明 $x_\epsilon+h\in(a+\epsilon,b)$ 与 x_ϵ 在 $[a+\epsilon,b]$ 上取极大值矛盾。

对函数 $F(x)-\gamma x$ 用上述日出引理，可知 $E_\gamma\subset E=[a,b_0)\cup\bigcup\limits_{k=1}^{\infty}(a_k,b_k)$ 并且

$$F(a_k)-\gamma a_k\le F(b_k)-\gamma b_k$$

这样

$$m(E_\gamma)\le m(E)=\sum_{k=0}^{\infty}b_k-a_k\le\gamma^{-1}\sum_{k=0}^{\infty}(F(b_k)-F(a_k))\le\gamma^{-1}(F(b)-F(a)).$$

由此可知 $m(\{D^+(F)(x)=\infty\})\le m(E_\gamma)\to 0,\gamma\to\infty.$

最后我们来证明命题的第二部分。只需要证明对任意的有理数 $r<R$，集合

$$E_{r,R}=\{x\in[a,b]:D^+(F)(x)>R,D_-(F)(x)<r\}$$

为零测集即可。前面已经证明了 $D^+(F)$ 是可测函数，同样的道理可知 $D_-(F)$ 也为可测函数，因而 $E_{r,R}$ 是可测集。如果 $E_{r,R}$ 对某对 r,R 不是零测集，根据可测集的性质，不妨假设 $E_{r,R}\subset[c,d]\subset[a,b]$ 并且 $d-c<\frac{R}{r}m(E_{r,R})$。对任意 $x\in E_{r,R}$, 存在 $h>0$ 使得

$$F(x-h)-F(x)>-rh.$$

在区间 $[-d,-c]$ 上对函数 $G(x)=F(-x)+rx$ 用日出引理，可知集合

$$-E_{r,R}\subset\{x:\exists h>0\ s.t. G(x+h)>G(x)\}=\cup(-b_k,-a_k)$$

其中开区间 $(-b_k,-a_k)$ 互不相交，并且满足

$$G(-b_k)\le G(-a_k),\quad F(b_k)-rb_k\le F(a_k)-ra_k$$

也即是

$$E_{r,R}\subset\bigcup_{k=1}^{\infty}(a_k,b_k),\quad F(b_k)-F(a_k)\le r(b_k-a_k).$$

再利用 $D^+(F)$ 可以给出 $F(b_k)-F(a_k)$ 的一个下界估计。我们在区间 (a_k,b_k) 上对函数 $G(x)=F(x)-Rx$ 用日出引理，同样的道理可以找到不相交的开区间 $(a_{k,l},b_{k,l})$ 使得

$$(a_k,b_k)\cap E_{r,R}\subset\bigcup_{l=1}^{\infty}(a_{k,l},b_{k,l}),\quad R(b_{k,l}-a_{k,l})\le F(b_{k,l})-F(a_{k,l})$$

这样由 F 的单调性，可知

$$\sum_l(b_{k,l}-a_{k,l})\le R^{-1}\sum_l F(b_{k,l})-F(a_{k,l})\le R^{-1}(F(b_k)-F(a_k))\le\frac{r}{R}(b_k-a_k)$$

再对 k 作和，便知

$$m(E_{r,R})\le\sum_{k,l}(b_{k,l}-a_{k,l})\le\frac{r}{R}\sum_k(b_k-a_k)\le\frac{r}{R}(d-c)<m(E_{r,R}).$$

矛盾！这就表明集合 $E_{r,R}$ 为零测集。 □

有了这些准备，我们现在就可以证明单调增函数的 Lebesgue 微分定理：上面讨论说明 F 几乎处处可微，并且 F' 非负。只需验证不等式

$$\int_a^b F'(x)dx \leq F(b) - F(a).$$

为此延拓 F 到整个实轴使得 $F(x) = F(a), x < a$, $F(x) = F(b), x > b$。在考虑函数

$$f_n(x) = n(F(x + \frac{1}{n}) - F(x))$$

可知 $f_n(x) \to F'(x), a.e.$. 根据定义 $f_n(x)$ 非负，故可用 Fatou 引理

$$\begin{aligned}\int_{[a,b]} F'(x)dx &\leq \liminf_{n\to\infty} \int_{[a,b]} n(F(x + \frac{1}{n}) - F(x))dx \\ &\leq \liminf_{n\to\infty} n(\frac{F(b)}{n} - \int_a^{a+\frac{1}{n}} F(x)dx) \\ &\leq F(b) - F(a).\end{aligned}$$

最后一步由于

$$F(a) \leq n \int_a^{a+\frac{1}{n}} F(x)dx.$$

当然还可以用分解定理，将 F 分解成跳跃函数和连续单调函数的和。跳跃函数的导数几乎处处为 0。

5.4 习题

北京大学数学分析 II 实验班作业 4

本次作业的提交时间和地点为 **3 月 29 日**习题课上 (注意特殊时间)，逾期视作零分。

1. 证明函数 $F(x) = x^2 \sin(x^{-2})$, $x \neq 0$, $F(0) = 0$ 导数存在，但 F' 在 $[0, 1]$ 上不可积。

2. 如果 F 在 $[a, b]$ 上连续，如果对任意 $a \leq x \leq b$ 都有 $D^+(F)(x) \geq 0$，证明 F 是单调增的。

3. 如果 F 是定义在 $[a, b]$ 上的连续函数，证明 Dini 导数

$$D^+(F)(x) = \limsup_{h>0, h\to 0} \frac{F(x+h) - F(x)}{h}$$

是可测的。

4. 考虑跳跃函数 $J(x) = \sum_{n=1}^{\infty} j_n(x)$. 证明

$$\limsup_{h\to 0} \frac{J(x+h) - J(x)}{h}$$

是可测的。

5. 如果 F 为 $[a, b]$ 上的单调增函数，J_F 为对应的跳跃函数，证明

$$\sum_{a<x_k<b} j_k(x) \leq F(b-0) - F(a+0), \quad \forall a \leq x \leq b.$$

6. 如果 $J(x) = \sum_{k=1} j_k(x)$ 为跳跃函数，并且

$$E_\epsilon = \{x : \bar{D}(J)(x) > \epsilon\}, \quad E_{\epsilon,N} = \{x : \bar{D}(J_N(x)) > \epsilon\}, \quad J_N(x) = \sum_{k>N} j_k(x).$$

证明

$$E_\epsilon \subset E_{\epsilon,N} \cup \{x_1, x_2, \ldots, x_N\}.$$

7. 构造一个单调增函数使得不连续点集恰好为有理数集。

8. 在 $\mathbb{R}$ 上，定义单边极大函数

$$\mathcal{M}_+ f(x) = \sup_{h>0} \frac{1}{h} \int_x^{x+h} |f(y)| dy.$$

记 $E_\alpha^+ = \{x \in \mathbb{R} : \mathcal{M}_+ f(x) > \alpha\}$, 证明

$$m(E_\alpha^+) = \frac{1}{\alpha} \int_{E_\alpha^+} |f(y)| dy.$$

5.5 绝对连续函数与微积分基本定理

上节我们介绍的有界变差函数是两个单调增函数的差，与可求长曲线有关，因而一般而言有界变差函数仍然不满足微积分基本定理。我们来回忆一下微积分基本定理的内容：我们要求 F 至少在 $[a,b]$ 上几乎处处可微，并且 F' 是 Lebesgue 可积的，满足

$$F(x)-F(a)=\int_{[a,x]}F'.$$

由积分的绝对连续性我们知道 $F(x)$ 是连续的，实际上 F 是绝对连续的，而这也是充分条件。

定义. *我们称定义在 $[a,b]$ 上的函数 F 是绝对连续函数，如果对任意 $\epsilon>0$, 存在 $\delta>0$ 使得对任意两两不相交区间 (a_k,b_k), $k=1,2,\ldots,n$ 满足 $\sum\limits_{k=1}^{n}(b_k-a_k)<\delta$ 都有*

$$\sum_{k=1}^{n}|F(b_k)-F(a_k)|<\epsilon.$$

所有绝对连续的函数组成空间 $AC([a,b])$.

我们熟知的除了 C^1 函数是绝对连续函数之外，还有 Lipschitz 函数。事实上如果 F 满足 Lipschitz 条件

$$|F(x)-F(y)|\le M|x-y|$$

那么显然有

$$\sum_{k=1}^{n}|F(a_k)-F(b_k)|\le M\sum_{k=1}^{n}|a_k-b_k|<M\delta.$$

绝对连续函数空间 $AC=AC([a,b])$ 具有下列性质：

- 绝对连续函数是一致连续的,因而也是连续的。并且绝对连续函数是有界变差函数满足 $T_F([a,b])=\int_{[a,b]}|F'|$.
- 如果 F 绝对连续，那么对任意零测集 N, $F(N)$ 也是零测集，这样由 F 的连续性，可知 F 将可测集映为可测集。
- AC 为线性空间: 如果 $F,G\in AC$, 那么对任意常数 a, b 有 $aF+bG\in AC$.
- AC 是一个代数：如果 $F,G\in AC$，那么 $FG\in AC$.
- AC 在范数 $\|F\|_{AC}=\sup|F|+\int|F'|$ 下是完备的可分的 Banach 空间。

现在我们就可以叙述一般的微积分基本定理。

定理 5.13 (微积分基本定理)**.** 如果 $F \in AC([a,b])$, 那么 F 几乎处处可微, 并且导数 $F' \in L^1([a,b])$ 满足

$$F(x) - F(a) = \int_{[a,x]} F'.$$

上述微积分基本定理可导出分部积分公式

定理 5.14 (分部积分公式)**.** 如果 $F, G \in AC([a,b])$, 那么

$$\int_{[a,b]} F'G + \int_{[a,b]} FG' = F(b)G(b) - F(a)G(a).$$

由于绝对连续函数必定是有界变差函数，而有界变差函数可写成两个单调增函数的差，因而由单调函数的可微性便知绝对连续函数几乎处处可微，并且导数可积。这样可定义

$$\tilde{F}(x) = F(a) + \int_{[a,x]} F'.$$

再由 Lebesgue 微分定理可知

$$\tilde{F}' = F', a.e.$$

注意到 F, $\tilde{F}$ 均为绝对连续函数，为了证明 $\tilde{F} = F$，我们只需要证明下面命题.

命题 5.15. 如果 $F \in AC([a,b])$ 并且 $F' = 0, a.e.$，那么 F 为常数。

只需要证明 $F(a) = F(b)$ 即可，想法就是绝对连续控制了导数不存在的部分，而导数为零的部分需要用下面的 Vitali 覆盖定理。

定义. 我们称 $\mathcal{B} = \{B_\alpha, B_\alpha\text{为}\mathbb{R}^d\text{中的闭球}\}$ 为 $\mathbb{R}^d$ 中的集合 E 的 *Vitali* 覆盖，如果对任意 $x \in E$ 以及 $\eta > 0$ 存在 $B_\alpha \in \mathcal{B}$ 使得 $x \in B_\alpha$, $m(B_\alpha) < \eta$。换言之，E 中的点被任意小的球盖住。

我们有下列 Vitali 覆盖定理.

定理 5.16 (Vitali)**.** 如果 $E \subset \mathbb{R}^d$ 并且 $m^*(E) < \infty$，对任意 E 的 *Vitali* 覆盖 $\mathcal{B}$ 以及 $\delta > 0$, 存在 $\mathcal{B}$ 中的有限个两两不交的球 $B_1, B_2, \ldots, B_n$ 使得

$$m^*(E \backslash \bigcup_{k=1}^{n} B_k) < \delta.$$

证明*:* 证明的想法是反复利用之前的 Vitali 引理。对任意 $\epsilon > 0$，可取开集 $A \supset E$ 使得 $m(A) \le m^*(E) + \epsilon$. 然后去掉所有 $\mathcal{B}$ 中的那些不完全包含在 A 中或者半径大于 1 的球。剩下的球仍然构成 E 的 Vitali 覆盖（这里注意到 $\mathcal{B}$ 中的球为闭球，而 A 为开集）。按照之前 Vitali 引理的做法，我们可以选取至多可数个互不相交的球 B_k 使得

$$r(B_k) > \frac{1}{2} \sup\{r(B) : B \in \mathcal{B}, B \cap B_j = \emptyset, j < k\}$$

且有

$$\bigcup_{B\in\mathcal{B}} B \subset \bigcup_k 5B_k.$$

这里的 $\mathcal{B}$ 是去掉了所有不全包含在 A 中或者半径大于 1 的球。如果只取出了有限个球 B_k，那么必定有 $E \subset \cup B_j$. 否则假设 $x \in E$, 但 $x \notin B_k$，由于 B_k 为闭集并且 $\mathcal{B}$ 中的球构成 E 的 Vitali 覆盖，因而存在 $B \in \mathcal{B}$ 使得 $x \in B$ 但 $B \cap B_k = \emptyset$. 这与 B_k 的选取矛盾（也就是可以选出更小的球出来）！这样可知选出了可数个球 B_k. 由于 B_k 互不相交，故而

$$\sum_k m(B_k) \leq m(A) < \infty.$$

这样存在 N 使得

$$\sum_{k>N} m(B_k) < 5^{-d}\epsilon.$$

同样的对任意 $x \in E \backslash \bigcup\limits_{k=1}^N B_k$，存在 $B_x \in \mathcal{B}$ 使得 $x \in B_x$, $B_x \cap \bigcup\limits_{k=1}^N B_k = \emptyset$. 根据选取规则必定有 B_l 使得 $B_x \cap B_l \neq \emptyset$(不防假设 B_l 是使得 l 最小的球)。显然 $l \geq N+1$, 并且有

$$r(B_l) > \frac{1}{2} r(B_x).$$

这表明 $B_x \subset 5B_l$. 从而有

$$E \backslash \bigcup_{k=1}^N B_k \subset \bigcup_{k>N} 5B_k.$$

这表明

$$m^*(E \backslash \bigcup_{k=1}^N B_k) \leq 5^d \sum_{k>N} m(B_k) < \epsilon.$$

这就证明了定理结论。 □

证明的过程表明

$$m^*(E / \bigcup_{k=1}^\infty B_k) = 0, \quad \sum_{k\geq 1} m(B_k) \leq m^*(E) + \epsilon.$$

现在我们就可以证明命题5.15。考虑集合 $E = \{x \in [a,b] : F'(x) = 0\}$。对任意 $x \in E$ 以及任意 $\varepsilon > 0$，存在 $\delta(x) > 0$ 使得

$$|F(y) - F(x)| < \epsilon |y - x|, \quad |y| \leq \delta(x).$$

也即是存在区间 $[x-h, x+h], h < \delta(x)$ 盖住 E，因而所有闭区间 $[x-h, x+h], h < \delta(x)$ 构成 E 的 Vitali 覆盖。由 Vitali 覆盖定理，存在有限个不相交区间 $[a_k, b_k] = [x_k - h_k, x_k + h_k]$，$1 \leq k \leq n$ 使得

$$m(E \backslash \bigcup_{k=1}^n [x_k - h_k, x_k + h_k]) < \epsilon.$$

可不妨假设

$$a \leq a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n \leq b.$$

这样

$$\begin{aligned}|F(b)-F(a)| &\leq |F(b)-F(b_n)| + |F(a_1)-F(a)| + \sum_1^{n-1} |F(a_{k+1})-F(b_k)| + \sum_{k=1}^{n} |F(b_k)-F(a_k)| \\ &\leq \varepsilon(b-a) + |F(b)-F(b_n)| + |F(a_1)-F(a)| + \sum_1^{n-1} |F(a_{k+1})-F(b_k)|.\end{aligned}$$

由于 $m(E) = b-a$, 可知

$$(b-b_n) + (a_1-a) + \sum_1^{n-1}(a_{k+1}-b_k) \leq m(E \backslash \bigcup_{k=1}^{n}[x_k-h_k, x_k+h_k]) < \epsilon.$$

这样根据绝对连续性可知 $F(b) = F(a)$. 这就证明了命题5.15. 从而完成了微积分基本定理的证明。

Lebesgue 微积分基本定理还有如下对应 Riemann 情形的形式：

定理 5.17. 如果 F 在 $[a,b]$ 上处处可微，并且 F' 可积，那么 F 是绝对连续函数，因而

$$\int_a^b F'(x)dx = F(b) - F(a).$$

定理的证明依赖于下面关于像集的估计。

命题 5.18. 如果 F 在 $[a,b]$ 实值函数，对任意 $E \subset [a,b]$, 如果 F 在 E 上可微并且 $|F'|_E \leq M$，那么有

$$m^*(F(E)) \leq M m^*(E).$$

证明: 固定任意 $\varepsilon > 0$, 首先存在开集 B 使得 $E \subset B$ 并且 $m(B) < m^*(E) + \varepsilon$. 对任意 $x \in E$，存在 $\delta > 0$ 使得

$$\frac{F(x+h)-F(x)}{h} \leq M + \varepsilon, \quad \forall |h| \leq \delta.$$

这表明 x 能被任意小的区间 $(x-h, x+h)$ 盖住。因而这些区间构成 E 的 Vitali 覆盖, 故存在可数个不相交的区间 $I_i = (x_i - h_i, x_i + h_i)$ 使得

$$m^*(E \backslash \bigcup_{i=1}^{\infty} I_i) = 0, \quad \sum_{i=1}^{\infty} m(I_i) \leq m^*(E) + \varepsilon.$$

在区间 I_i 上我们有

$$\begin{aligned}m^*(F(I_i)) &\leq |F(y_i) - F(z_i)| \leq |F(y_i) - F(x_i)| + |F(x_i) - F(z_i)| \\ &\leq (M+\varepsilon)(|y_i - x_i| + |x_i - z_i|) \leq 2(M+\varepsilon)h_i.\end{aligned}$$

这里用到 F 在区间 I_i 上是有界的。这样

$$m^*(F(E\cap\cup I_i))\le\sum m^*(F(E\cap I_i))\le(M+\varepsilon)\sum m(I_i)\le(M+\varepsilon)(m^*(E)+\varepsilon).$$

由于

$$m^*(F(E))\le m^*(F(E\cap\cup I_i))+m^*(F(E\backslash\cup I_i))$$

而 $E\backslash\cup I_i$ 为零测集，因而只需要证明命题对于 E 为零测集成立即可，即是要证明 $F(E)$ 也为零测集。定义

$$E_n=\{x\in E:|F(y)-F(x)|\le(M+\varepsilon)|y-x|,\forall|y-x|<\frac{1}{n}\}.$$

由条件可知 E_n 单调升到 E，故而 $F(E_n)$ 单调升到 $F(E)$. 因而只需要证明 $F(E_n)$ 为零测集即可。由于 E_n 为零测集，可以找到可数个不相交的开区间 $J_{n,k}$ 使得

$$E_n\subset\cup J_{n,k},\quad\sum m(J_{n,k})<\varepsilon.$$

而对于任意 $x,y\in E_n\cap J_{n,k}$ 有

$$|F(x)-F(y)|\le(M+\varepsilon)|x-y|.$$

这样有

$$m^*(F(E_n))\le\sum m^*(F(E_n\cap J_{n,k}))\le\sum(M+\varepsilon)m(J_{n,k})\le\varepsilon(M+\varepsilon).$$

由 ε 的任意性可知 $F(E_n)$ 为零测集，因而 $F(E)=\cup F(E_n)$ 也为零测集。

□

现在我们可以证明上面定理。由于积分的绝对连续性，只需要证明

$$|F(b)-F(a)|\le\int_a^b|F'|dx.$$

事实上对任意 $\varepsilon>0$，定义集合

$$E_n=\{x\in[a,b]:(n-1)\varepsilon\le|F'(x)|<n\varepsilon\}.$$

因而

$$m^*(F(E_n))\le n\varepsilon m(E_n)\le\varepsilon m(E_n)+\int_{E_n}|F'(x)|dx.$$

这样由 F 的连续性可知

$$\begin{aligned}|F(b)-F(a)|&\le m(F([a,b]))\le\sum m^*(F(E_n))\\&\le\sum(\varepsilon m(E_n)+\int_{E_n}|F'(x)|dx)\\&\le\varepsilon(b-a)+\int_a^b|F'(x)|dx.\end{aligned}$$

之前我们知道单调函数可以分解为一个跳跃函数和连续单调升函数的和，现在我们还可以对连续单调增函数作进一步的分解。

定理 5.19 (Lebesgue 分解定理)**.** 对于定义在 $[a,b]$ 上的单调增函数 F，我们有分解 $F = F_{AC} + F_C + J$ 使得每个函数都是单调增并且

- F_{AC} 绝对连续；
- F_C 连续，并且 $F_C' = 0, a.e.$;
- $J(x) = \sum\limits_{k=1}^{\infty} j_n(x)$，其中 $0 \le j_n(x_n) \le \alpha_n$，$j_n(x) = 0$，$x < x_n$; $j_n(x) = \alpha_n$，$x > x_n$，并且 $\sum \alpha_n < \infty$.

进一步的在差一个常数的意义下，上述分解唯一，即是如果 $F = \tilde{F}_{AC} + \tilde{F}_C + \tilde{J}$ 也满足上述三个条件，那么 $F_{AC} - \tilde{F}_{AC}$, $F_C - \tilde{F}_C$, $J - \tilde{J}$ 为常数。

证明: 只需要取 $F_{AC}(x) = \int_{[a,x]} F'$, $J = J_F(x)$, $F_C = F - F_{AC} - J_F$ 即可。唯一性见习题。 □

我们回头再来看可求长曲线，之前的结论告诉我们可求长曲线等价于分量 $x(t), y(t)$ 都是有界变差函数。但通常而言，下列等式不成立

$$L(\gamma) = \int_{[a,b]} \sqrt{|x'(t)|^2 + |y'(t)|^2}.$$

但是如果 $x(t)$, $y(t)$ 是绝对连续函数，那么有

命题 5.20. 给定曲线 $\gamma(t) = (x(t), y(t)), a \le t \le b$. 如果 $x(t)$, $y(t)$ 是绝对连续函数，那么 $\gamma(t)$ 可求长，并且

$$L(\gamma) = \int_{[a,b]} \sqrt{|x'(t)|^2 + |y'(t)|^2}.$$

证明: 由微积分基本定义可知

$$|\gamma(t_i) - \gamma(t_{i-1})| = |\int_{[t_i, t_{i-1}]} \gamma'(t)| \le \int_{[t_i, t_{i-1}]} |\gamma'(t)|.$$

故而由可求长曲线的定义可知

$$L(\gamma) \le \int_{[a,b]} |\gamma'(t)| = \int_{[a,b]} \sqrt{|x'(t)|^2 + |y'(t)|^2}.$$

为证明反向不等式，对任意 $\epsilon > 0$，可取 $[a,b]$ 上的向量值阶梯函数 g 使得 $\gamma' = g + h$ 并且 $\int_{[a,b]} |h(t)| < \epsilon$. 定义 $G(t) = \int_{[a,t]} g$, $H(t) = \int_{[a,t]} h$. 这样由绝对连续性可知 $\gamma = G + H$, 特别的

$$T_\gamma(a,b) \ge T_G(a,b) - T_H(a,b).$$

由于 G, H 都是绝对连续向量值函数，这样由 h 的选取可知

$$T_H(a,b) < \epsilon.$$

为了估计 $T_G(a,b)$，由于 g 为 $[a,b]$ 上的阶梯函数，这表明存在分割 $a = t_0 < t_1 < \ldots < t_n = b$ 使得 g 在每个区间 (t_i, t_{i+1}) 上都是常数。这表明

$$T_G(a,b) \geq \sum_{i=1}^n |G(t_i) - G(t_{i-1})| = \sum_{i=1}^n \int_{[t_{i-1},t_i]} |g| = \int_{[a,b]} |g|.$$

这样便得

$$T_\gamma(a,b) \geq T_G(a,b) - \epsilon \geq \int_{[a,b]} |\gamma'| - 2\epsilon.$$

由 ϵ 的任意性可知定理结论成立。 □

这个结论给出了可求长曲线可写成导数积分的充分条件。尽管对于一般可求长曲线不能写成导数积分形式，但是对于可求长曲线，总是可以重新选取弧长参数。定义 $s(t) = \int_{[a,t]} |\gamma'|$, 这样有严格单调映射 $s : [a,b] \to [0, L(\gamma)]$. 这样曲线可以重新表示为

$$\tilde{\gamma}(s) = \gamma \circ s^{-1}$$

并且有 $\tilde{\gamma}(s)$ 是 Lipschitz 连续的

$$|\tilde{\gamma}(s) - \tilde{\gamma}(s')| \leq |s - s'|.$$

故而 $\tilde{\gamma}(s)$ 是绝对连续，因而可以表示成导数积分的形式。

最后我们讨论与绝对连续函数相关的一维的积分换元公式。我们证明如下结论

定理 5.21. *如果 $\varphi : [a,b] \to [c,d]$ 是严格单调增的绝对连续，那么对任意 $[c,d]$ 上的可积函数 f，有*

$$\int_c^d f(y)dy = \int_a^b f(\varphi(x))\varphi'(x)dx.$$

证明: 类似于之前的积分换元公式，只需要证明对任意可测集 $E \subset [a,b]$ 集合 $\varphi(E)$ 也是可测集并且

$$m(\varphi(E)) = \int_E \varphi'.$$

这个只要依次说明：

- 开区间 (α, β) 满足上述关系，这个由映射的绝对连续性是显然的。
- 由此可知任何开集也满足上述关系，并且任何 G_δ 集也满足上式关系。

- 如果 E 零测集，对任意 $\varepsilon > 0$, 可以找到开集，从而是可数个互不相交的开区间 (a_k, b_k) 并，使得

$$E \subset \bigcup_{k=1}^{\infty}(a_k, b_k), \quad \sum_{k=1}^{\infty}(b_k - a_k) < \epsilon$$

  而由绝对连续性可知

$$\varphi(E) \subset \bigcup_{k=1}^{\infty}(\varphi(a_k), \varphi(b_k))$$

  可以任意小，从而 $\varphi(E)$ 为零测集。这表明零测集也满足上述关系式。对一般可测集可以表示成一个 G_δ 集与一个零测集的差，从而可知对一般可测集上式也成立。

□

在这章的最后，我们来总结不同连续的概念。

- 连续：函数 f 在 (a, b) 上连续如果对于任意 $x \in [a, b]$, 任意的 $\epsilon > 0$, 存在 $\delta > 0$ 使得 $|f(y) - f(x)| < \epsilon$, $\forall |y - x| < \delta$.

- 一致连续：函数 f 在 (a, b) 上一致连续，如果对任意 $\epsilon > 0$, 存在 $\delta > 0$ 使得 $|f(y) - f(x)| < \epsilon$, $\forall |x - y| < \delta$. 在闭区间上，一致连续与连续等价。

- 绝对连续：函数 F 在 $[a, b]$ 上绝对连续如果对任意 $\varepsilon > 0$，存在 $\delta > 0$，对任意有限个不相交的区间 $(a_i, b_i) \subset [a, b]$, $1 \leq i \leq N$, 有

$$\sum_{i=1}^{N}|F(a_i) - F(b_i)| < \epsilon, \quad \sum_{i=1}^{N}(b_i - a_i) < \delta.$$

- Lipschitz 连续：F 在 $[a, b]$ 上是 Lipschitz 连续，如果存在常数 M 使得

$$|F(x) - F(y)| \leq M|x - y|, \quad \forall x, y \in [a, b].$$

  易知 Lipschitz 连续的函数一定是绝对连续的。

- α-Hölder 连续 $(0 < \alpha < 1)$: 存在常数 M 使得

$$|F(x) - F(y)| \leq M|x - y|^{\alpha}, \quad \forall x, y \in [a, b]$$

  这样的函数组成的空间通常记作 $C^{\alpha}([a, b])$. $\alpha = 0$ 的情形可以理解为通常的连续函数，$\alpha = 1$ 就是 Lipschitz 函数空间，不过注意记号的区别 C^1 通常记作连续可微函数空间，而不是 Lipschitz 函数空间。

我们来看一些典型例子.

- 函数 $f = |x|$ 是 Lipschitz 连续的，但在 0 处不可微；

- 在 $[0,1]$ 上的函数 x^β, $0<\beta<1$ 是绝对连续的，并且是 C^α, $0<\alpha\le\beta$;
- $[0,\frac{1}{2}]$ 上的函数 $f(x)=|\log x|^{-1}$, $f(0)=0$, 是绝对连续函数，但不是 C^α 连续的，对任意 $0<\alpha<1$;
- Cantor-Lebesgue 函数几乎处处可微的连续函数，但不是绝对连续的，但是是 C^α, $\alpha\le\frac{\ln 2}{\ln 3}$;
- Hardy 在 1916 年证明了对所有 $0<a<1$ 以及 $b>a^{-1}$, Weierstrass 函数

  $$W_{a,b}(x)=\sum_{n=0}^{\infty}a^n\cos(b^n\pi x)$$

  是 $C^{-\frac{\ln a}{\ln b}}$ 连续的，但处处不可微。
- Tagaki 函数：令 $s(x)=\min\limits_{n\in\mathbb{Z}}|x-n|$, 定义

  $$T_w(x)=\sum_{n=0}^{\infty}w^n s(2^n x),\quad 0\le x\le 1.$$

  当 $0<w<\frac{1}{2}$ 时，$T_w(x)$ 几乎处处可微。但在 $w=\frac{1}{2}$ 处，$T_{\frac{1}{2}}(x)$ 不是 Lipschitz 的，也不是绝对连续的，但属于 C^α, $0<\alpha<1$, 事实上有个更强的结论

  $$|T_{\frac{1}{2}}(x+h)-T_{\frac{1}{2}}(x)|\le 2h\log_2\frac{1}{h},\quad 0<h<1.$$

  这个函数的图像像牛奶冻，因而也得名”牛奶冻曲线”. 在特殊的点处，还能得到

  $$T_{\frac{1}{4}}(x)=2x(1-x)$$

  参见 Hata-Yamaguti(1984)。

5.6 习题

北京大学数学分析 II 实验班作业 5

本次作业的提交时间和地点为 **4 月 2 日**习题课上，逾期视作零分。

1. 如果 F 是定义在 $[a,b]$ 上的可微函数，并且 F' 是 Riemann 可积的，证明 F 是有界变差函数，并且

$$T_F([a,b]) = \int_a^b |F'(x)|dx.$$

2. 如果 $F \in BV([a,b])$，定义 $G(x) = T_F([a,x])$，那么 G 几乎处处可微并且 $G'(x) = |F'(x)|, a.e.$

3. 根据定义直接证明 Cantor-Lebesgue 函数不是绝对连续的。

4. 绝对连续函数的复合不一定是绝对连续函数：证明 $g \circ f$ 不是绝对连续函数，但 $g = y^{\frac{1}{3}}$, $f(x) = x^3 \sin(x^{-1})$ 都是绝对连续函数。

5. 证明 $AC([a,b])$ 是 $BV([a,b])$ 的线性子空间：即是证明绝对连续函数空间是线性空间并且是有界变差函数.

6. 证明 $AC([a,b])$ 在范数 $\|f\| = \sup|f| + \|f'\|_{L^1}$ 下是完备的 Banach 空间。

7. 证明 $AC([a,b])$ 是可分的：即是存在可数稠密子集。

8. 证明如果有界变差函数 F 满足

$$\int_a^b |F'|dx = T_F(a,b)$$

那么 F 是绝对连续函数。

9. 如果 F 是 $[a,b]$ 上连续的凸函数，证明 $F \in AC([a,b])$。

10. 如果 f 在 $[0,1]$ 上绝对连续，定义 $f_h(x) = \frac{f(x+h)-f(x)}{h}$. 这里 $h > 0$, $f(x) = f(1)$ 如果 $x > 1$. 证明 $\lim\limits_{h\to\infty} \|f_h - f'\|_{L^1} = 0$.

11. 证明如果 F 满足 Lipschitz 条件

$$|F(x) - F(y)| \le M|x-y|$$

等价于 F 绝对连续并且 $|F'| \le M, a.e.$.

12. 证明绝对连续函数满足 Lusin(N) 性质：如果 $N \subset [a,b]$ 为零测集，F 为 $[a,b]$ 上的绝对连续函数，那么 $F(N)$ 为零测集。进一步证明 F 将可测集映为可测集。

13. 如果 F 是 $[a,b]$ 上的连续有界变差函数并且把零测集映为零测集，那么 F 为绝对连续函数。

14. 完成 Lebesgue 分解定理的证明：对于定义在 $[a,b]$ 上的单调增函数 F，我们有分解 $F = F_{AC} + F_C + J$ 使得每个函数都是单调增并且 F_{AC} 绝对连续；F_C 连续，并且 $F_C' = 0, a.e.$; $J(x) = \sum_{k=1}^{\infty} j_n(x)$, 其中 $0 \leq j_n(x_n) \leq \alpha_n$, $j_n(x) = 0$, $x < x_n$; $j_n(x) = \alpha_n$, $x > x_n$, 并且 $\sum \alpha_n < \infty$. 进一步的在差一个常数的意义下，上述分解唯一，即是如果 $F = \tilde{F}_{AC} + \tilde{F}_C + \tilde{J}$ 也满足上述三个条件，那么 $F_{AC} - \tilde{F}_{AC}$, $F_C - \tilde{F}_C$, $J - \tilde{J}$ 为常数。

15. 假设 F 是 $\mathbb{R}$ 上的处处可微的单调增函数，满足

$$\lim_{x \to -\infty} F(x) = A, \quad \lim_{x \to +\infty} F(x) = B.$$

证明

$$\int_{\mathbb{R}} F'(x) dx = B - A.$$

16. 在 $[0,1]$ 上考虑函数

$$F(x) = \begin{cases} x^a \sin(x^{-b}), & 0 < x \leq 1, \\ 0, & x = 0. \end{cases}$$

对怎样的 α(并证明)，F 是 $[0,1]$ 上的 C^α 连续函数？

1 方向导数和微分

高维的积分理论我们已经熟悉，微分理论相对而言复杂很多。从现在开始我们讨论定义在欧式空间 $\mathbb{R}^n$ 或者 $\mathbb{R}^n$ 中的开区域 Ω (连通开集，这里连通指的是任何两点都能连接一条连续的曲线) 上的函数 f。特别的 $(x, f(x))$，$x \in \Omega$ 组成 $\mathbb{R}^{n+1}$ 中的图或者曲面，因而多元微积分主要的目的是研究高维空间中低维曲面的几何等性质。

如同在一维，$\mathbb{R}^n$ 中的点 x 可以对应数组 $x = (x_1, x_2, \ldots, x_n)$，$x_i \in \mathbb{R}$。高维空间中没有除法结构，因而高维只能固定某一个方向来定义导数。

定义 1.1 (方向导数)**.** 给定函数 $f : \Omega \to \mathbb{R}$，$x_0 \in \Omega$，$v \in \mathbb{R}^n$。如果下面的极限

$$\lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h}$$

存在，我们就称函数 f 在 x_0 处沿 v 的**（方向）导数**存在，并将它记作

$$\left(\nabla_v f\right)(x_0) = \lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h}.$$

习惯上，我们还把 $\nabla_v f$ 写成 $\dfrac{\partial f}{\partial v}$。

对于标准坐标系，定义坐标方向 $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$（第 i 个位置为 1）。根据定义 $\nabla_{e_i} f$ 可以看成是固定其他分量，对 x_i 求一元导数，这个特殊的方向导数（依赖于坐标系）称为**偏导数**并记作

$$\frac{\partial f}{\partial x_i}(x_0) = \left(\nabla_{e_i} f\right)(x_0).$$

e_i 和 $\frac{\partial}{\partial x_i}$ 都可以表示坐标系的第 i 个方向，因而可以约定它们等同，这也是我们称 $\frac{\partial}{\partial x_i}$ 为向量场的原因。

例子. 我们有两个基本的例子

1) 假设函数 $f : \mathbb{R}^2 \to \mathbb{R}$ 只依赖于 x-坐标，我们习惯上将它写成 $f(x, y) = f(x)$，那么，

$$\frac{\partial f}{\partial y} \equiv 0.$$

为此，只需要按定义计算即可：

$$\frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{f(x, y + h) - f(x)}{h} = 0.$$

2) 考虑 $\mathbb{R}^2$ 上定义的函数

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

那么，$f(x, y)$ 的两个偏导数 $\frac{\partial f}{\partial x}$ 和 $\frac{\partial f}{\partial y}$ 在原点处均为 0。另外，如果 $v = (v_1, v_2)$ 其中 $v_1 v_2 \neq 0$，那么

$$\nabla_v f(0) = \lim_{h \to 0} \frac{v_1 v_2}{h(v_1^2 + v_2^2)}$$

是不存在的。由此可见，偏导数存在不能保证其它的方向导数存在。

导数从一维到高维不能直接推广，但微分 (即是用线性函数在局部逼近函数) 则可以很自然的推广到高维。

定义 1.2 (函数的微分)**.** 给定函数 $f:\Omega\to\mathbb{R}$ 和 $x_0\in\Omega$。如果存在线性映射 $A:\mathbb{R}^n\to\mathbb{R}$，使得对于 $v\to 0$ 时，其中 $v\in\mathbb{R}^n$，我们有

$$f(x_0+v)=f(x_0)+A(v)+o(|v|),$$

也就是说

$$\lim_{v\to 0}\frac{|f(x_0+v)-f(x_0)-A(v)|}{|v|}=0,$$

其中向量的长度 $|v|$ 可以取标准欧式距离，我们就称 f 在 x_0 处可微并且称线性映射 A 是 f 在 x_0 的**微分**。我们通常将 A 写成下面的样子：

$$df\big|_{x=x_0}=df(x_0):\mathbb{R}^n\to\mathbb{R}.$$

如果 f 在 Ω 的每个点处都可微，我们就称 f 是 Ω 上的**可微函数**。

如果 f 在 $x_0\in\Omega$ 处可微，那么显然 f 在 x_0 处连续。

命题 1.3 (微分与方向导数之间的关系)**.** 假设 $f:\Omega\to\mathbb{R}$ 在 x_0 处可微，那么 f 在 x_0 的任意方向导数都存在。特别地在局部坐标系下，对于 $v=(v_1,\cdots,v_n)$ 有

$$df(x_0)(v)=(\nabla_v f)(x_0)=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x_0)v_j=(\nabla f)(x_0)\cdot v.$$

这里 $\nabla f=(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\ldots,\frac{\partial f}{\partial x_n})$ 为 f 在 x_0 处的梯度向量。df 是微分形式的写法，∇f 是向量写法，微分形式和向量在上面关系下有自然的对偶。

证明: 根据微分的定义，存在线性映射 A 使得

$$f(x_0+hv)-f(x_0)=A(hv)+o(|hv|),\quad \forall h\in\mathbb{R},\quad v\in\mathbb{R}^n.$$

特别的固定 v，有方向导数

$$(\nabla_v f)(x_0)=\lim_{h\to 0}\frac{hA(v)+o(|hv|)}{h}=A(v)=\sum_{i=1}^{n}v_i\cdot A(e_i).$$

由于 A 是线性映射，在局部坐标系下，根据偏导数的定义有

$$df(x_0)=A=(\nabla f)(x_0)=(\frac{\partial}{\partial x_1}f,\frac{\partial}{\partial x_2}f,\ldots,\frac{\partial}{\partial x_n}f)$$

命题得证。 □

之前的例子表明偏导数存在不一定表明方向导数存在，即使是方向导数都存在也不一定表明函数可微，甚至都不一定连续。

例子. 考虑函数在 $\mathbb{R}^2$ 定义函数：

$$f(x,y)=\begin{cases}\dfrac{y^2}{x}, x\neq 0\\ 0, \quad x=0.\end{cases}$$

那么，f 在 $(0,0)$ 处的任意一个方向导数都存在：假设 $v=(v_1,v_2)\in\mathbb{R}^2$ 且 $v\neq 0$。那么，我们有

1) 如果 $v_1=0$，那么

$$\nabla_v f(0)=\lim_{t\to 0}\frac{0-0}{t}=0.$$

2) 如果 $v_1\neq 0$，那么

$$\nabla_v f(0)=\lim_{t\to 0}\frac{t\frac{v_2^2}{v_1}}{t}=\frac{v_2^2}{v_1}.$$

然而，f 在 $(0,0)$ 处不可微，这是因为 f 在 $(0,0)$ 不连续：可以选取点列 $\left\{\left(\dfrac{1}{n^2},\dfrac{1}{n}\right)\right\}_{n\geqslant 1}$，函数在这些点上取值不收敛到 0。

这个例子也说明，只考虑方向导数并不能对函数在一个点附近的行为有效地进行控制。

然而，如果方向导数的具有连续性，情况就大不相同。直观上，连续性允许我们从一点的信息出发理解这点附近的情况。

命题 1.4. 给定开区域 $\Omega\subset\mathbb{R}^n$ 和 $x_0\in\Omega$，f 是在 Ω 上定义的函数。假设 f 的所有偏导数 $\dfrac{\partial f}{\partial x_i}$ 在 Ω 上存在且连续，其中 $i=1,2,\cdots,n$，那么 f 在 x_0 处可微。

证明*:* 在局部坐标系下，记

$$u_j=(v_1,v_2,\ldots,v_j,0,\ldots,0),\quad u_0=0,\quad j=1,2,\ldots,n.$$

注意到 u_j 与 u_{j-1} 只在第 j 个位置上不同，用拉格朗日中值定理以及偏导数的定义可知

$$\begin{aligned}f(x_0+v)-f(x_0)-(\nabla f)(x_0)\cdot v&=\sum_{j=1}^n f(x_0+u_j)-f(x_0+u_{j-1})-\frac{\partial f}{\partial x_j}(x_0)v_j\\&=\sum_{j=1}^n\frac{\partial f}{\partial x_j}(x_0+\tilde{u}_j)v_j-\frac{\partial f}{\partial x_j}(x_0)v_j\end{aligned}$$

再根据偏导数的连续性，可知函数 f 在 x_0 处可微。 □

上面证明表明，尽管拉格朗日中值定理对多元函数不成立，但只要偏导数在一个邻域内成立，那么有

$$f(x_0+v)-f(x_0)=\sum_{j=1}^n v_j\cdot\frac{\partial f}{\partial x_j}(x_0+\tilde{u}_j).$$

特别的作为一个推论

命题 1.5. 如果函数 f 在开区域 Ω 上可微，并且 $df(x)=0$，那么 f 是常数。

一元微积分中我们可以利用导数来研究函数的极值问题，多元函数也有类似结论。

命题 1.6. 假设 $f:\Omega\to\mathbb{R}$ 在 x_0 处可微并且取最大值 *(局部最大或最小值即可)*，那么 $df(x_0)=0$。

证明: 根据定义存在线性映射 A 使得

$$f(x_0+v)-f(x_0)=A(v)+o(v)\geq 0,\quad f(x_0-v)-f(x_0)=-A(v)+o(v)\geq 0,$$

因而必定有 $A(v)=0$，根据定义便有 $df(x_0)=0$。 □

1.1 映射的微分

给定函数 $f:\Omega\to\mathbb{R}$，我们上次课定义它的微分 $df(x)$(或者 ∇f), 可看成是 Ω 到 $\mathbb{R}^n$ 的向量值函数。如果需要求 f 的高阶比如 2 阶导数，那么就等价于求多元向量值函数 df 的一阶导数。

定义 1.7 (微分)**.** 假设 Ω 是 $\mathbb{R}^n$ 中的区域，Ω' 是 $\mathbb{R}^m$ 中的区域，给定映射 $f:\Omega\to\Omega'$。如果存在线性映射

$$df\big|_{x=x_0}=df(x_0):\mathbb{R}^n\to\mathbb{R}^m,$$

使得对于 $\mathbb{R}^n$ 中的 $v\to 0$ 时，我们有

$$f(x_0+v)=f(x_0)+df(x_0)v+o(v),$$

即

$$\lim_{v\to 0}\frac{|f(x_0+v)-f(x_0)-df(x_0)v|}{|v|}=0,$$

我们就称 f 在 x_0 处可微并且称线性映射 $df(x_0)$ 是 f 在 x_0 的**微分**。如果 f 在 Ω 的每个点处都可微，我们就称 f 是 Ω 上面的**可微映射**。

注意到微分的定义并不依赖于坐标系的选取，因而可以推广到更加一般的 Banach 空间中 (特别的可以是无穷维的)。对于欧式空间之间的映射，其微分可以用矩阵来表达。

命题 1.8 (微分的计算)**.** 假设 f 是从 $\mathbb{R}^n$ 中的区域 Ω 到 $\mathbb{R}^m$ 中区域 Ω' 的映射，在标准坐标系下，f 可写成是向量值函数

$$f(x)=(f_1(x),f_2(x),\cdots,f_m(x)),\quad x\in\Omega.$$

那么 f 在 $x_0\in\Omega$ 处可微等价于每个分量函数 f_j 在 x_0 处可微。特别地，如果 f 在 x_0 处可微，那么映射 $df(x_0):\mathbb{R}^n\to\mathbb{R}^m$ 可以用 $m\times n$ 的矩阵

$$\Big(\frac{\partial f_j}{\partial x_i}(x_0)\Big)_{\substack{j=1,\cdots,m\\ i=1,\cdots,n}}$$

来表示（我们将这个矩阵称作是 f 在 x 处的 **Jacobi 矩阵**，并记作 $\mathrm{Jac}(f)$ 或者 $\mathbf{J}(f)$。

证明: 等价性是显然的。根据微分的定义，如果 f 可微，取分量，可知分量函数 f_j 可微。如果每个分量函数 f_j 可微，那么根据微分的定义，向量值函数 f 也可微。

如果 f 在 x_0 处可微，那么存在从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射 A，因而可用 $m\times n$ 的矩阵来表示。考虑分量函数 f_j，因而

$$f_j(x_0+te_i)=f_j(x_0)+\sum_{i=1}^{n}A_{ji}v_i+o(te_i)=f_j(x_0)+tA_{ji}+o(te_i).$$

让 $t\to 0$，可知

$$A_{ji}=\frac{\partial f_j}{\partial x_i}(x_0).$$

□

注记. 上述命题表明，映射可求微分等价于其分量可求微分，所以，我们可以通过继续对分量求微分来引入 k-次可导的概念（就是每次求完微分之后这个微分的每个分量都能再求微分）。所以，我们可以定义 $C^k(\Omega,\mathbb{R}^m)$，这是 k 次微分仍然连续的映射的空间。根据上次课程用偏导数判定微分存在性的定理，我们知道只要 f 的连续 k 次偏导数（可能是沿着不同方向的）存在并且连续，那么映射就是 C^k 的。这是一个非常方便有效的判断方式。

我们现在研究复合映射的微分，也就是所谓的链式法则。

命题 1.9. （链式法则）假设 $\Omega_j\subset\mathbb{R}^{m_j}$ 是开集，其中 $j=1,2,3$，$f:\Omega_1\to\Omega_2$，$g:\Omega_2\to\Omega_3$ 是映射。假设 f 在点 $x_1\in\Omega_1$ 处可微，g 在点 $x_2=f(x_1)\in\Omega_2$ 处可微，那么复合映射 $g\circ f$ 在 x_1 处可微，并且

$$\big(d(g\circ f)\big)(x_1)=(dg)(f(x_1))\circ df(x_1).$$

注记. 上述映射的复合可以用下面的交换图来表示：

$$\begin{tikzcd}
\Omega_1 \arrow[r, "f"] \arrow[dr, "g\circ f"'] & \Omega_2 \arrow[d, "g"] \\
 & \Omega_3
\end{tikzcd}$$

那么，它们所对应的微分（在线性的层次上）也可以用类似的交换图来表示：

$$\begin{tikzcd}
\mathbb{R}^{m_1} \arrow[r, "df(x_1)"] \arrow[dr, "d(g\circ f)(x_1)"'] & \mathbb{R}^{m_2} \arrow[d, "dg(f(x_1))"] \\
 & \mathbb{R}^{m_3}
\end{tikzcd}$$

证明: 链式法则的推导与 1 维的情形如出一辙：令 $x_2=f(x_1)\in\Omega_2$，按照定义有

$$f(x_1+h)=f(x_1)+df(x_1)h+\delta(h),$$
$$g\big(f(x_1)+\ell\big)=g\big(f(x_1)\big)+dg(x_2)\ell+\Delta(\ell),$$

其中 $h \in \mathbb{R}^{m_1}$，$\ell \in \mathbb{R}^{m_2}$，$\lim\limits_{h\to 0}\dfrac{|\delta(h)|}{|h|} = \lim\limits_{\ell\to 0}\dfrac{|\Delta(\ell)|}{|\ell|} = 0$。据此，我们有

$$\begin{aligned} g(f(x_1+h)) - g(f(x_1)) &= g\big(f(x_1) + df(x_1)h + \delta(h)\big) - g(f(x_1)) \\ &= dg(x_2)\big(df(x_1)h + \delta(h)\big) + \Delta(df(x_1)h + \delta(h)) \\ &= \underbrace{dg(x_2)\big(df(x_1)h\big)}_{=dg(x_2)\circ df(x_1)(h)} + df(x_2)\big(\delta(h)\big) + \Delta(df(x_1)h + \delta(h)). \end{aligned}$$

所以，

$$\begin{aligned} &\frac{|g(f(x_1+h)) - g(f(x_1)) - dg(x_2)\circ df(x_1)(h)|}{|h|} \\ =&\frac{|df(x_2)\big(\delta(h)\big)|}{|h|} + \frac{|\Delta(df(x_1)h + \delta(h))|}{|h|} \\ \leqslant& C\Big|\frac{\delta(h)}{h}\Big| + \underbrace{\frac{|\Delta(df(x_1)h+\delta(h))|}{|df(x_1)h+\delta(h)|}}_{o(1)} \cdot \underbrace{\frac{|df(x_1)h+\delta(h)|}{|h|}}_{\leqslant C_1}. \end{aligned}$$

由此可见，这是一个 $o(1)$ 项，按照微分的定义，

$$d(g\circ f)(x_1) = dg(x_2)\circ df(x_1).$$

这就完成了证明。 $\square$

作为推论，我们可以计算反函数（逆映射）的微分：

推论 1.10. 给定区域 $\Omega_1 \subset \mathbb{R}^{n_1}$ 和 $\Omega_2 \subset \mathbb{R}^{n_2}$ 和可微映射 $f : \Omega_1 \to \Omega_2$。假设 f 是双射并且其逆映射 $f^{-1} : \Omega_2 \to \Omega_1$ 是可微的，那么

- $n_1 = n_2$；
- $df(x)$ 是可逆的（等价于 $\mathrm{Jac}(f)(x)$ 的行列式是非零的）。

此时，对于任意的 $y \in \Omega_2$，我们有

$$df^{-1}(y) = \Big(df\big|_{x=f^{-1}(y)}\Big)^{-1}.$$

证明: 我们令 $\Omega_3 = \Omega_1$，$g = f^{-1}$，$x_1 = x$，$x_2 = y$，$g\circ f = \mathrm{Id}$，其中

$$\mathrm{Id} : \Omega_1 \to \Omega_1, \quad x \mapsto x,$$

是单位映射，它的微分在每个点处都是单位映射（线性）。根据链式法则，我们就有

$$\mathrm{Id} = dg(y)\circ df(x).$$

根据矩阵的秩的理论，我们知道 $n_1 \leqslant n_2$。用 f^{-1} 替换 f，我们就得到 $n_2 \leqslant n_1$。这就证明了维数的部分。上面的等式已经蕴含了逆映射的微分的计算。 $\square$

例子 (指数映射的微分). 上个学期我们对于 $n\times n$ 的矩阵定义了指数映射

$$\exp:\mathbf{M}_n(\mathbb{R})\to\mathbf{M}_n(\mathbb{R}),\ A\mapsto e^A=\sum_{k=0}^{\infty}\frac{A^k}{k!}.$$

我们现在计算它的微分 $d\exp$。固定 $A\in\mathbf{M}_n(\mathbb{R})$，我们要找到

$$d\exp(A):\mathbf{M}_n(\mathbb{R})\to\mathbf{M}_n(\mathbb{R}),$$

其中，我们把 $\mathbf{M}_n(\mathbb{R})$ 视作是 $\mathbb{R}^{n^2}$。对于任意较小的 $V\in\mathbf{M}_n(\mathbb{R})$，我们有

$$e^{A+V}-e^V=\sum_{n=0}^{\infty}\frac{1}{n!}\left((A+V)^n-A^n\right)$$

现在强行展开 $(A+V)^n-A^n$（注意矩阵 A 和 V 的乘法未必交换）。通过将 V 的二次项（以及更高次数的项）放到一起，我们得到

$$(A+V)^n-A^n=\sum_{k=0}^{n}A^kVA^{n-1-k}+Q_n(V).$$

二项式展开的一共不超过 2^n 项，所以 $Q_n(V)$ 中至多有 2^n 项。我们上学期证明过（无论你选取什么样的范数），存在常数 c（依赖于范数），使得对任意的 $n\times n$ 的矩阵 A 和 B，我们都有

$$\|A\cdot B\|\leqslant c\|A\|\|B\|.$$

上述 $Q_n(V)$ 的一个通项形如 $AAVVAA\cdots AA$，这是一个由 n 个 A 和 V 排出来的长度为 n 的字符串，其中至少有 2 个 V。我们可以要求 $\|V\|\leqslant\|A\|$，因为最终我们会令 $V\to 0$（除非 $A=0$，此时 $Q_n(V)=V^n$，下面的结论仍然成立），所以

$$\|AAVVAA\cdots AA\|\leqslant c^n\|A\|\|A\|\|V\|\|V\|\|A\|\|A\|\cdots\|A\|\|A\|.$$

那么，我们得到

$$\|Q_n(V)\|\leqslant 2^n\times\left(c^n\|V\|^2\|A\|^{n-2}\right).$$

从而，我们有

$$\begin{aligned}&\Big\|\exp(A+V)-\exp(A)-\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)\Big\|\\ \leqslant&\sum_{n=0}^{\infty}\Big\|\frac{1}{n!}Q_n(V)\Big\|\leqslant\left(\sum_{n=0}^{\infty}\frac{(2c\|A\|)^n}{n!}\right)\|V\|^2=e^{2c\|A\|}\|V\|^2.\end{aligned}$$

那么，我们注意到右端的项是 $o(\|V\|)$ 并且 $\displaystyle\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)$ 是收敛的。所以，

$$d\exp(A)(V)=\sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{n}A^kVA^{n-1-k}\Big)$$

特别地，如果 A 和 V 可交换，那么 $d\exp(A)(V)=\exp(A)V$。我们还有

$$d\exp(0)=\mathrm{Id}.$$

有了 Jacobi 矩阵的概念，现在我们就可以讨论积分换元公式了。

定理 1.11. 如果映射 $\varphi : U \to V$ 以及逆映射 $\varphi^{-1} : V \to U$ 都是连续可微的，并且 *Jacobi* 行列式 $J_\varphi(x) = det\left(\frac{\partial \varphi_j(x)}{\partial x_i}\right)_{1\le i,j\le d} \neq 0$。如果 f 是 $\varphi(U)$ 上的可积函数，那么有下列积分换元公式

$$\int_{\varphi(U)} f(y)dy = \int_U f(\varphi(x))|J_\varphi|dx.$$

注记. 更一般的结论是 φ, φ^{-1} 是 *Lipschitz* 的。由 *Rademacher* 定理可知 φ 几乎处处可微，上述积分换元公式同样成立，可参见 *L. Evans and R. Gariepy* 的 *Measure theory and fine properties of functions.*

关于积分变换公式还有如下 *Guzman* 的结果：假设 E 是 $\mathbb{R}^d$ 中的可测集，$T : E \to T(E)$. 如果 T 以及逆变换 T^{-1} 将 E 与 $T(E)$ 中的可测集映为可测集，并且 $T(E)$ 的测度有限，那么存在 E 上的可积函数 J_T 使得当 $f \in L^1(T(E))$ 时有

$$\int_{T(E)} f dy = \int_E f(T(x))J_T(x)dx.$$

证明: 只需要证明对任意的方体 $I \subset U$ 有

$$m(\varphi(I)) = \int_I |J_\varphi|dx$$

以及对任意的零测集 $E \subset U$ 有 $m(\varphi(E)) = 0$ 即可。零测集的像也为零测集只需要映射 φ 是 Lipschitz 即可。下面只对 I 是方体映射 φ 是连续可微的映射验证上式成立即可。由于 φ 连续可微，存在常数 M 使得在 I 上

$$M^{-1} \le \|d\varphi\|, \|d\varphi^{-1}\|, |J_\varphi| \le M.$$

对任意 $\epsilon > 0$，将 I 分成足够小的方块 I_j 使得

$$|\varphi(x) - \varphi(x_j) - d\varphi(x_j)(x - x_j)| \le M\epsilon|x - x_j|, \quad x \in I_j$$

这里 x_j 为方块 I_j 的中心。由之前线性变换的结果可知，存在常数 K（不依赖于 ϵ）使得

$$m(\varphi(I_j)) \le (|J_\varphi(x_j)| + MK\epsilon)m(I_j)$$

这样可知

$$m(\varphi(I)) \le MK\epsilon m(I) + \sum_j |J_\varphi(x_j)|m(I_j).$$

让 $\epsilon \to 0$ 并且将分割足够细可知由 Lebesgue 积分的定义

$$m(\varphi(I)) \le \int_I |J_\varphi|dx.$$

至于反向不等式可以考虑逆映射 φ^{-1}. 由上述不等式可以得到对于非负可积函数 f 总有

$$\int_{\varphi(U)} f(y)dy \leq \int_U f(\varphi(x))|det(d\varphi)(x)|dx$$

这样考虑映射 $\varphi^{-1}:\varphi(U)\to U$ 同样也满足上述不等式。从而有

$$m(E) \leq \int_{\varphi(E)} |det(d\varphi^{-1})(y)|dy \leq \int_E |det(d\varphi^{-1})(\varphi(x))||det(d\varphi)(x)|dx$$

由于 $\varphi^{-1}\varphi(x)=x$, 可知

$$|det(d\varphi^{-1})(\varphi(x))||det(d\varphi)(x)| = 1.$$

这就说明对任意可测集 $E\subset U$ 都有

$$m(E) \leq \int_{\varphi(E)} |det(d\varphi^{-1})(y)|dy \leq \int_E |det(d\varphi^{-1})(\varphi(x))||det(d\varphi)(x)|dx = m(E).$$

因而有等式成立。也即是

$$m(\varphi(E)) = \int_E |J_\varphi(x)|dx.$$

当然也可以直接证明反向不等式。 $\square$

在欧式空间 $\mathbb{R}^n$ 上的积分除了标准的笛卡尔坐标，通常还可以用极坐标。以 $\mathbb{R}^3$ 为例，考虑坐标变换

$$x = r\sin\theta\sin\varphi, \quad y = r\sin\theta\cos\varphi, \quad z = r\cos\theta.$$

这里 $r\geq 0$, $\theta\in[0,\pi]$, $\varphi\in[0,2\pi]$。自然这个坐标变换不是从 $[0,\infty)\times[0,\pi]\times[0,2\pi]$ 的同胚映射，但可以限定在开区域 $(0,\infty)\times(0,\pi)\times(0,2\pi)$ 上是到像集的同胚，并且像集的余集为 $\mathbb{R}^3$ 上的零测集 (也可以分象限来考虑)。此时 Jacobi 矩阵为

$$\begin{pmatrix} \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

其行列式的绝对值为 $r^2\sin\theta$。这表明在极坐标下有

$$\int_{\mathbb{R}^3} f = \int_0^\infty\int_0^\pi\int_0^{2\pi} f(r,\theta,\varphi)r^2\sin\theta drd\theta d\varphi.$$

1.2 习题

数学分析二作业 6

本次作业的提交时间和地点为 **4 月 9 日**的习题课上，逾期视作零分。

1.（微分的唯一性）考虑映射 $f:\Omega\subset\mathbb{R}^n\to\mathbb{R}^m$。假设存在两个线性映射 $A_i:\mathbb{R}^n\to\mathbb{R}^m$ （$i=1,2$），使得对 $i=1$ 和 2 和 $v\to 0$ 时，都有

$$f(x_0+v)=f(x_0)+A_i(v)+o(v).$$

证明，$A_1=A_2$。

2. 考察函数 $f(x,y)=\begin{cases}(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}, & \text{若}(x,y)\neq(0,0),\\ 0, & \text{若}(x,y)=(0,0).\end{cases}$ 试计算 f 的偏导数并证明 f 的两个偏导数在 $(0,0)$ 处均不连续但是 f 在 $(0,0)$ 处可微。

3. 假设函数 f 在 Ω 上的所有偏导数 $\dfrac{\partial f}{\partial x_i}$ 都存在并且一致有界。证明 f 在 Ω 上连续。f 是否在 Ω 上可微？如果是，请证明；否则给出反例。

4. $\Omega\subset\mathbb{R}^2$ 为开集，我们用 (x,y) 来表示 $\mathbb{R}^2$ 上的坐标。函数 $f:\Omega\to\mathbb{R}$ 的偏导数 $\dfrac{\partial f}{\partial x}$ 和 $\dfrac{\partial f}{\partial y}$ 处处存在。如果偏导数 $\dfrac{\partial f}{\partial y}$ 在 Ω 上连续。证明，f 在 Ω 上可微。

5. 考虑在 $\mathbf{M}_n(\mathbb{R})$ 上定义的行列式函数

$$\det:\mathbf{M}_n(\mathbb{R})\to\mathbb{R},\quad A\mapsto\det(A).$$

任意给定 $A\in\mathbf{M}_n(\mathbb{R})$，试计算 $d\det\big|_{X=A}$。

6. 计算下列函数的偏导数：

$$(1)\ f(x,y,z)=x^{y^z},\qquad (2)\quad f(x,y,z)=\tan\frac{xy}{z^2}$$

7. 假设 f 为可微函数，用 f 的偏导数表达下列多元函数的偏导数：

$$(1)\ u(x,y,z)=f(x,xy,xyz)\qquad\qquad (2)\quad u(x,y)=f(\log x+\frac{1}{y})$$

8. 求如下坐标变换 f 的 Jacobi 矩阵 $J(f)$ 并计算 $\det J(f)$：

$$(1) f:\mathbb{R}^2\to\mathbb{R}^2,\quad (u,v)\mapsto(e^u\cos v,e^u\sin v);$$
$$(2) f:\mathbb{R}^2\backslash(0,0)\to\mathbb{R}^2\backslash(0,0),\quad (u,v)\mapsto(\frac{u}{u^2+v^2},\frac{v}{u^2+v^2}).$$

9. 我们考虑 $\mathbb{R}^3$ 上的柱面坐标系：$x = r\cos\theta, y = \sin\theta, z = z$。这个坐标变换用映射来写就是

$$\Phi : \mathbb{R}_{>0} \times (0, 2\pi) \times \mathbb{R} \to \mathbb{R}^3, \quad (r, \theta, z) \mapsto (r\cos\theta, r\sin\theta, z).$$

设 f 是 $\mathbb{R}^3$ 上的二次可微函数。当 $(x, y) \neq (0, 0)$ 时，试通过计算来证明：

$$\begin{aligned}
&\frac{\partial f}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y},\\
&\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 = \left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \theta}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2,\\
&\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}.
\end{aligned}$$

10. 我们考虑 $\mathbb{R}^3$ 上的球坐标系 $(x, y, z) = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$。证明，

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 = \left(\frac{\partial f}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial \theta}\right)^2 + \frac{1}{r^2\sin^2\theta}\left(\frac{\partial f}{\partial \varphi}\right)^2,$$

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2}.$$

11. 对于常数 $k \in R$, 我们称 $\mathbb{R}^n$ 上的函数 f 为 k **次齐次的**，如果对任意的 $\lambda > 0$，$x \neq 0$，有 $f(\lambda x) = \lambda^k f(x)$。证明线性函数是齐次函数。

12. 假设 f 是可微的。证明：f 是 k 次齐次函数当且仅当它满足 Euler 等式 $\displaystyle\sum_{i=1}^n x_i\frac{\partial f}{\partial x_i} = kf$。

13. 假设可微函数 f 是 k 次齐次函数。证明，对任意 $v \in \mathbb{R}^n$，$\nabla_v f$ 是 $k-1$ 次齐次函数。

14. 令 $f(x_1, \cdots, x_n) = \det\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ \vdots & & \vdots \\ x_1^{n-1} & \cdots & x_n^{n-1}\end{pmatrix}$。证明，$\displaystyle\sum_{i=1}^n x_i\frac{\partial f}{\partial x_i} = \frac{n(n-1)}{2}f$ 和 $\displaystyle\sum_{i=1}^n \frac{\partial f}{\partial x_i} = 0$。

2 偏导数的可交换性和 Taylor 展开

上节课我们讲到偏导数连续即表明可微，因而可以定义 k 次连续可微的映射，通常记作空间 $C^k(\Omega)$。进一步如果任意次连续可微，那么这样的映射称为是**光滑的**，特别的如果是函数的话，那么就称之为**光滑函数**。光滑函数组成的空间记为 $C^\infty(\Omega)$。

定义 (微分同胚). 对于 $\mathbb{R}^n$ 中的两个开区域 U, V，如果存在双射映射 $f: U \to V$，使得 f 以及 f 的逆映射 $f^{-1}: V \to U$ 都是光滑的，我们就称 U 和 V 是**微分同胚的**（光滑同胚）。我们用 $C^\infty(U, V)$ 记作所有从 U 到 V 之间的光滑映射。

例子. 考虑如下映射

$$f: \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^3.$$

很明显，f 是光滑的并且是双射。但是 $f^{-1}(y) = \sqrt[3]{y}$ 不是光滑的。

微分同胚的两个区域从某种意义来讲可以认为是等价的。

引理 2.1. *假设 Φ 是从 U 到 V 的同胚映射。对于非负整数 k，如果 $f \in C^k(V)$，那么 $\Phi^* f = f(\Phi)$ 是 U 上的 k 次连续可微的函数。这里 $\Phi^* f$ 称为 f 在 Φ 下的拉回 (pullback)。*

证明: 对 k 做归纳。$k = 0$ 是显然，假设命题对 n 成立。考虑 V 上的 $n+1$ 次连续可微函数 f。对 $f(\Phi)$ 求一阶偏导数

$$\frac{\partial f(\Phi(x))}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial f(\Phi(x))}{\partial y_j} \cdot \frac{\partial \Phi^j(x)}{\partial x_i}$$

由于 Φ 是光滑同胚，因而 $\frac{\partial \Phi^j(x)}{\partial x_i}$ 是 U 上的光滑函数。根据假设 $\frac{\partial f(y)}{\partial y_j}$ 是 V 上的 n 次连续可微函数，这样根据归纳假设 $\frac{\partial f(\Phi(x))}{\partial y_j}$ 是 U 上的 n 次连续可微函数，因而 $\frac{\partial f(\Phi(x))}{\partial x_i}$ 是 U 上的 n 次连续可微函数，也即是 $f(\Phi)$ 是 U 上的 $n+1$ 次连续可微函数。 □

接下来我们介绍平坦的欧式空间的重要性质，偏导数的可交换性。

定理 2.2 (Clairaut-Schwarz). *给定 $\mathbb{R}^n$ 上的开集 Ω（$n \geqslant 2$）和函数 $f: \Omega \to \mathbb{R}$ 是函数，$i, j \in \{1, 2, \cdots, n\}$ 是两个不同的指标。假设在 Ω 上，函数 $\frac{\partial f}{\partial x_i}(x)$，$\frac{\partial f}{\partial x_j}(x)$ 和 $\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(x)$ 存在并且*连续。*那么，$\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(x)$ 也存在并且对任意的 $x \in \Omega$，我们有*

$$\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_j})(x) = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(x).$$

证明: 不妨假设 $n = 2$，$x_i = x, x_j = y$。固定点 (x_0, y_0)，对 t, s 足够小使得 $(x_0 + t, y_0 + s) \in \Omega$，根据假设 $\frac{\partial}{\partial x}(\frac{\partial f}{\partial y})$ 存在并且连续。特别由 Fubini 定理我们有

$$\int_0^{t_0} \int_0^{s_0} \frac{\partial}{\partial x} \left(\frac{\partial f(x_0 + t, y_0 + s)}{\partial y} \right) ds dt = \int_0^{s_0} \int_0^{t_0} \frac{\partial}{\partial x} \left(\frac{\partial f(x_0 + t, y_0 + s)}{\partial y} \right) dt ds.$$

等式右边用微积分基本定理可计算得到

$$f(x_0+t_0,y_0+s_0)-f(x_0+t_0,y_0)-f(x_0,y_0+s_0)+f(x_0,y_0).$$

根据偏导数的定义

$$\begin{aligned}\frac{\partial}{\partial y}(\frac{\partial f}{\partial x})(x_0,y_0) &= \lim_{s_0\to 0} s_0^{-1}(\frac{\partial f}{\partial x}(x_0,y_0+s_0)-\frac{\partial f}{\partial x}(x_0,y_0))\\ &= \lim_{s_0\to 0} s_0^{-1}\lim_{t_0\to 0} t_0^{-1}(f(x_0+t_0,y_0+s_0)-f(x_0+t_0,y_0)-f(x_0,y_0+s_0)+f(x_0,y_0)).\end{aligned}$$

用积分表达式可知上述极限存在并且等于 $\frac{\partial}{\partial x}(\frac{\partial f}{\partial y})(x_0,y_0)$。 □

注记. 基于这个命题，我们可以引入一个方便的记号来记多重的偏导数：令 $\alpha=(\alpha_1,\cdots,\alpha_n)$ 为一个多重指标，也就是说对于每个 $i\leqslant n$，有 $\alpha_i\in\mathbb{Z}_{\geqslant 0}$，我们用 $\partial^\alpha f$ 表示 α_1 个 $\dfrac{\partial}{\partial x_1}$，$\alpha_2$ 个 $\dfrac{\partial}{\partial x_2}$，$\cdots$，$\alpha_n$ 个 $\dfrac{\partial}{\partial x_n}$ 作用在 f 上，即

$$\partial^\alpha f=\left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}\circ\left(\frac{\partial}{\partial x_{n-1}}\right)^{\alpha_{n-1}}\circ\cdots\circ\left(\frac{\partial}{\partial x_1}\right)^{\alpha_1}(f).$$

当 f 足够多次连续可微时，上述命题保证了这些算子的复合不依赖于它们作用的顺序。我们还令

$$|\alpha|=\alpha_1+\cdots+\alpha_n\leqslant k,$$

传统上我们还把上面的多重偏导数 ∂^α 记作

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots x_n^{\alpha_n}}.$$

比如说，我们经常看到

$$\frac{\partial^2 f}{\partial x^2},\quad \frac{\partial^2 f}{\partial x\partial y}.$$

例子 (Clairaut-Schwarz 的反例)**.** 如果 *Clairaut-Schwarz* 定理中的连续性不成立，那么命题可能并不成立。考察函数

$$f(x,y)=\begin{cases}\frac{xy(x^2-y^2)}{x^2+y^2}, & (x,y)\neq(0,0);\\ 0, & (x,y)=(0,0).\end{cases}$$

那么，我们有

$$\frac{\partial f}{\partial x}(x,y)=\frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2}.$$

从而，我们有

$$\frac{\partial}{\partial y}(\frac{\partial f}{\partial x})(0,0)=\lim_{y\to 0}\frac{\frac{\partial f}{\partial x}(0,y)-\frac{\partial f}{\partial x}(0,0)}{y}=-1.$$

类似地（利用对称性），我们有

$$\frac{\partial}{\partial x}\big(\frac{\partial f}{\partial y}\big)(0,0)=1.$$

这表明 *Clairaut-Schwarz* 定理并不成立。请思考定理中的哪个条件没有被满足。

作为一个应用，接下来我们介绍高维函数的 Taylor 公式，基本想法就是化归为一维情形。

定理 2.3. 假设 $\Omega\subset\mathbb{R}^n$ 是凸的开集，f 在 Ω 上 $k+1$ 次连续可微。那么，对于任意的 $x,y\in\Omega$，其中 $x=(x_1,\cdots,x_n)$，$y=(y_1,\cdots,y_n)$，存在 $\theta\in[0,1]$，使得

$$f(y)=\sum_{|\alpha|\leqslant k}\frac{\partial^\alpha f(x)}{\alpha!}(y-x)^\alpha+\sum_{|\alpha|=k+1}\frac{\partial^\alpha f\big(x+\theta(y-x)\big)}{\alpha!}(y-x)^\alpha,$$

其中，

$$(x-y)^\alpha=(x_1-y_1)^{\alpha_1}\cdots(x_n-y_n)^{\alpha_n},\quad \alpha!=\alpha_1!\cdots\alpha_n!.$$

证明*:* 我们考虑 x 与 y 之间的连线并把函数 f 限制到这条线上。用分析的语言写，我们考虑函数

$$g:(-\varepsilon,1+\varepsilon)\to\mathbb{R},\quad g(t)=f\big(x+t(y-x)\big).$$

这里用到了 Ω 的凸性。我们首先来计算 g 的 k-次导数（$k\geqslant 0$）：

$$g^{(k)}(t)=\sum_{|\alpha|=k}\frac{k!}{\alpha!}(\partial^\alpha f)\big(x+t(y-x)\big)(y-x)^\alpha.$$

这个式子可以用归纳法来证明：$k=0$ 是显然的。假设

$$g^{(k-1)}(t)=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}(\partial^\alpha f)\big(x+t(y-x)\big)(y-x)^\alpha.$$

根据方向导数的性质，对任意的函数 f，我们有

$$\frac{d}{dt}f\big(x+tv\big)\big|_{t=0}=v^i\frac{\partial f}{\partial x_i}(x).$$

所以，

$$\begin{aligned}
g^{(k)}(t)&=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}\Big[(\partial^\alpha f)\big(x+t(y-x)\big)\Big]'(y-x)^\alpha\\
&=\sum_{|\alpha|=k-1}\frac{(k-1)!}{\alpha!}\Big[\sum_{m=1}^n(y_m-x_m)\Big(\frac{\partial}{\partial x_m}(\partial^\alpha f)\Big)\big(x+t(y-x)\big)\Big](y-x)^\alpha\\
&=\sum_{|\beta|=k}\frac{k!}{\beta!}(\partial^\beta f)\big(x+t(y-x)\big)(y-x)^\beta.
\end{aligned}$$

最后一个等号需要搞清楚如下的组合性质：为了从前一步的某个 $|\alpha|=k-1$ 得到固定的 β，其中 $|\beta|=k$，有如下 n 种可能性：

$$\alpha_1=(\beta_1-1,\cdots,\beta_n)\longrightarrow\beta,\ \alpha_2=(\beta_1,\beta_2-1,\cdots,\beta_n)\longrightarrow\beta,\cdots,\alpha_n=(\beta_1,\cdots,\beta_n-1)\longrightarrow\beta.$$

所以，所求的系数应该是（这个等价于 $\beta_1+\cdots+\beta_n=k$）：

$$\sum_{m=1}^{n}\frac{(k-1)!}{\alpha_m!}=\frac{k!}{\beta!}.$$

现在对 g 用 Taylor 公式（在 0 和 1）之间，我们有 $\theta\in[0,1]$，使得

$$g(1)=\sum_{\ell\leqslant k}\frac{g^{(\ell)}(0)}{\ell!}+\frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

将 $g(t)$ 的值代入即可。 □

3 反函数定理和隐函数定理

对于一元函数，如果函数连续单调，那么存在反函数，并且也是连续单调的。多元函数没有单调性，但一元函数的单调性可用导数来判断。具体来说如果连续可微函数 f 在一点导数非零，那么局部上存在反函数。其在高维的推广可用 Jacobi 矩阵的行列式来判断。

首先我们看一个简单情形。如果 U, V 是欧式空间 $\mathbb{R}^n$ 中的有界开区域，f 是 U 到 V 的连续双射，在闭包 $\bar{U}$ 是连续的单射，那么逆映射是连续的。这只需要注意到 f 将闭集映为闭集，开集映为开集即可。现在假设 f 是连续可微的，不加其他条件的前提下逆映射 f^{-1} 不一定是可微的 (比如上节开始的例子 $y=x^3$)。一个充分的条件是 Jacobi 矩阵非退化，即是行列式非零。

引理 3.1. *假设 f 是从 U 到 V 的双射，连续可微，并且 Jacobi 矩阵非退化。如果 f 在闭包 $\bar{U}$ 是连续的单射，那么逆映射 f^{-1} 是从 V 到 U 的连续可微映射。*

显然如果 f 和逆映射都是连续可微的，由链式法则可知 Jacobi 矩阵非退化。这个引理说明如果逆映射存在，并且映射本身是非退化的连续映射，那么逆映射也一定是连续可微的。

证明: 逆映射的连续性上面已经讨论过。固定一点 $x_0\in U$, $y_0=f(x_0)\in V$。根据假设存在非退化矩阵 A 使得

$$f(x_0+v)=f(x_0)+Av+o(v)$$

假设 $E(\delta)$ 满足

$$f^{-1}(y_0+\delta)=x_0+A^{-1}\delta+E(\delta).$$

根据连续性可知 $\lim\limits_{|\delta|\to 0}|E(\delta)|=0$。这样根据 f 的可微性质，我们得到

$$y_0+\delta=f(x_0+A^{-1}\delta+E(\delta))=y_0+A(A^{-1}\delta+E(\delta))+o(A^{-1}\delta+E(\delta)).$$

这样就可以得到关于 $E(\delta)$ 的等式

$$E(\delta)+o(A^{-2}\delta+A^{-1}E(\delta))=0.$$

此即表明

$$\lim_{|\delta|\to 0}\frac{|E(\delta)|}{|\delta|}=0$$

也即是逆映射 f^{-1} 是连续可微的。 □

更一般的非退化的连续可微映射在局部上就能保证逆映射的存在性。

定理 3.2 (反函数定理)**.** $\Omega\subset\mathbb{R}^n$ 是开集，$f:\Omega\to\mathbb{R}^n$ 是 C^1 的。如果对于点 $x_0\in\Omega$，f 在此点的微分 $df(x_0)$ 是可逆的，那么 f 在 x_0 的局部上是 C^1 的微分同胚，即存在开集 $U\subset\Omega$，$x_0\in U$ 和开集 $V\subset\mathbb{R}^n$，使得 f 在 U 上的限制给出的 $f:U\to V$ 双射并且 f 与 f^{-1} 均为 C^1。

证明: 首先我们可以做平移以及线性变换使得 $x_0=0$, $f(x_0)=0$, $df(x_0)=I$（单位矩阵）。由于 f 连续可微，因而存在 $\delta>0$ 使得

$$|f(v)-v|<\epsilon_0|v|,\quad \|J(f)(v)-I\|<\epsilon_0,\quad \forall|v|\le\delta$$

这里 ϵ_0 是某个小的常数，矩阵的范数可取所有元素绝对值之和。用微积分基本定理可得

$$f(u)-f(v)=\int_0^1 J(f)(tu+(1-t)v)(u-v)dt,\quad \forall u,v\in\mathbb{R}^n,|u|,|v|\le\epsilon_0.$$

注意到

$$\|\int_0^1 J(f)(tu+(1-t)v)dt-I\|<\epsilon_0$$

因而当 ϵ_0 足够小时，矩阵 $\int_0^1 J(f)(tu+(1-t)v)dt$ 非退化并且有估计

$$|f(u)-f(v)-(u-v)|\le\epsilon_0|u-v|.$$

这就表明 f 在 $\bar{B}(0,\delta)$ 上是单射。

接下来只需要证明 $f(\bar{B}(0,\delta))$ 包含 $f(0)=0$ 的一个邻域即可。取 $\delta_1>0$ 足够小。对任意 $|v|\le\delta_1$，我们需要找到 $|u|\le\delta$ 使得

$$f(u)=v=u+o(u),\quad u=v-o(u),\quad o(u)=f(u)-u$$

因而可以迭代来寻找这样的 u。归纳定义

$$u_0=v,\quad u_{k+1}=v-o(u_k).$$

对足够小的 $\epsilon_0 < \frac{1}{2}$ 以及足够小的 $\delta_1 < \frac{1}{2}\delta$, 归纳证明

$$|u_{k+1}| = |v - o(u_k)| \le |v| + |o(u_k)| \le \delta_1 + \epsilon_0|\delta| \le \delta.$$

此外

$$|u_{k+1} - u_k| = |o(u_k) - o(u_{k-1})| = |f(u_k) - u_k - f(u_{k-1}) + u_{k-1}| \le \epsilon_0|u_k - u_{k-1}|$$

因而 $u_k \in \bar{B}(0,\delta)$ 并且是 Cauchy 列。这样可知 $u_k \to u$。由 f 的连续性，可知 $u = v - o(u)$。 □

注记. 反函数定理对任意的（无限维）完备赋范线性空间都成立，有兴趣的同学可以尝试将上述证明做推广。

我们还可以要求反函数定理中的映射有更高的正则性：

推论 3.3. 假设 $k \geqslant 2$ 是整数。在反函数定理的假设中，如果我们进一步要求 $f \in C^k(\Omega)$，那么得到的逆映射 $f^{-1} \in C^k(V)$。

证明: 根据链式法则，f^{-1} 的微分是 $d(f^{-1})(y) = (df)^{-1}(f^{-1}(y))$，其中 $y \in V$。令 $g = f^{-1}$。我们在 U 和 V 上任意选取坐标 (x_i) 和 (y_j)，假设 f 和 g 的 Jacobi 矩阵分别是 A 和 B，即

$$df(x) = A(x) = (A_{ij}(x_1, \cdots, x_n)), \ \ dg(y) = B(y) = (B_{ij}(y_1, \cdots, y_n)).$$

根据要求 $A_{ij} \in C^{k-1}(U)$。然而，$B(y) = A^{-1}(g(y))$。根据逆矩阵的计算，我们知道 B_{ij} 形如

$$B_{ij}(y) = \sum \frac{A_{i'j'}(g(y))A_{i''j''}(g(y)) \cdots A_{i'''j'''}(g(y))}{\det A(y)}.$$

在求 ℓ 次导数的时候，分母上只会出现 $\det A$，而导数都出现在分子上面，而由链式法导致的 g 的导数至多出现到 $\ell - 1$ 阶。据此，我们知道 $B_{ij}(y)$ 是 C^{k-1} 的，从而 $g = f^{-1}$ 是 C^k 的。 □

注记. 如果我们把反函数定理条件中映射改为 C^∞ 的，那么结论中的函数和映射也是 C^∞，这从反函数定理的叙述就可以看出来。

作为反函数定理的推论，接下来我们讨论更加一般的隐函数定理。假设 $\Omega \subset \mathbb{R}^{n+p}$ 是开集。为了定理的叙述方便，我们将 $\mathbb{R}^{n+p}$ 写成乘积结构 $\mathbb{R}^n \times \mathbb{R}^p$。在 $\mathbb{R}^n$ 上面我们用坐标 $x = (x_1, \cdots, x_n)$，在 $\mathbb{R}^p$ 上面我们用坐标 $y = (y_1, \cdots, y_p)$。对于 Ω 上的映射 / 函数 f，我们用 $f(x,y)$ 这样的二元函数来表示，其中 $x \in \mathbb{R}^n$，$y \in \mathbb{R}^p$。为了行文方便，我们引入下面（不标准的）符号：将 y 固定，我们就可以将 $f(x,y)$ 看作是 x 的函数，即

$$f(\cdot, y) : \mathbb{R}^n \to \mathbb{R}^p,$$

从而我们可以对 x 的微分，我们将这个微分记作 $d_x f(x,y)$；类似地，我们可以定义 $d_y f(x,y)$。

定理 3.4 (隐函数定理)**.** 假设映射 f 是从非空开集 $\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$ 到 $\mathbb{R}^p$ 的连续可微映射。如果存在 $(x^*, y^*) \in \Omega$，使得 $f(x^*, y^*) = 0$ 并且 $\big(d_y f\big)(x^*, y^*)$ 可逆，那么存在 x^* 的开领域 $U \subset \mathbb{R}^n$，y^* 的开领域 $V \subset \mathbb{R}^p$ 以及连续可微的映射 $\phi : U \to V$ $(U \times V \subset \Omega)$，满足对任意 $x \in U$ 有

$$f(x, \phi(x)) = 0, \quad d\phi(x) = -(d_y f(x, \phi(x)))^{-1} \circ d_x f(x, \phi(x)).$$

此外如果 $f(x, y) = 0$，$x \in U,\, y \in V$，那么有 $y = \phi(x)$。

隐函数定理可以理解为有 p 个有 $n+p$ 个未知变量的方程，如果取一阶线性逼近 (作微分) 并且能找到 $p \times p$ 阶矩阵非退化，那么对应的变量可以唯一 (局部上) 的表示为其他变量的函数。

证明: 我们利用反函数定理来证明隐函数定理。定义映射

$$F : \Omega \to \mathbb{R}^n \times \mathbb{R}^p,\ (x, y) \mapsto F(x, y) = (x, f(x, y)).$$

很明显，F 是连续可微的。我们可以计算 F 在 (x, y) 处微分，用 Jacobi 矩阵来可以表达成：

$$dF = \begin{pmatrix} \mathrm{id} & 0 \\ d_x f & d_y f \end{pmatrix},$$

这是一个 $(n+p) \times (n+p)$ 的方阵。由于 $(d_y f)(x^*, y^*)$ 可逆，所以分块矩阵 $(dF)(x^*, y^*)$ 也是可逆的。对映射 F 在 (x^*, y^*) 处用反函数定理：存在 (x^*, y^*) 的开领域 $\widetilde{U} \subset \Omega$，$x^*$ 的开领域 U 以及 $0 \in \mathbb{R}^p$ 的开领域 V_1 使得 F 是 $\widetilde{U}$ 到 $U \times V_1$ 的 C^1 的微分同胚。特别的存在连续可微函数 T 使得

$$F^{-1}(x, z) = (x, T(x, z)), \quad \forall x \in U, \quad z \in V_1.$$

这样可以得到连续可微映射 $\phi : U \to \mathbb{R}^p$ 满足 $\phi(x) = T(x, 0)$。根据定义有

$$f(x, y) = 0, (x, y) \in \widetilde{U} \Leftrightarrow (x, y) = F^{-1}(x, 0) = (x, \phi(x)).$$

特别的有 $f(x, \phi(x)) = 0$ 对任意 $x \in U$ 成立，并且 $y^* \in \phi(U)$。这样可取 y^* 的开邻域 V(适当缩小 U) 使得 $U \times V \subset \widetilde{U}$，并且 ϕ 是从 U 到 V 的连续可微映射。至于 ϕ 的微分，用链式法则对 $f(x, \phi(x))$ 求微分可知

$$df_x(x, y) + df_y(x, y) \cdot d\phi(x) = 0, \text{ 其中 } y = \phi(x).$$

在 (x^*, y^*) 的附近，$df_y(x, y)$ 是可逆的，所以 $d\phi(x) = -(d_y f(x, y))^{-1} d_x f(x, y)$。 □

3.1 习题

数学分析二作业 7

本次作业的提交时间和地点为 **4 月 23 日**的习题课上，逾期视作零分。

1. 考虑映射函数 $f:\mathbb{R}\to\mathbb{R}$，其中
$$f(x)=\begin{cases}x+x^2\sin\left(\frac{\pi}{x}\right), & x\neq 0;\\ 0, & x=0.\end{cases}$$
证明，f 在 $\mathbb{R}$ 上可微分并且 $df(0)\neq 0$ 但是在 0 附近任意小的邻域上 f 都不是单射。请对比反函数定理说明反函数定理中哪一个条件没有被满足。

2. 如果 f,g 是区域 Ω 上的 k 次连续可微的函数，证明 fg 也是 Ω 上的 k 次连续可微函数。

3. 考虑映射
$$f:\mathbb{R}^2\to\mathbb{R}^2,\quad (x,y)\mapsto(e^x\cos(y),e^x\sin(y)).$$
证明，f 在 $(x,y)\neq(0,0)$ 处都满足反函数定理的要求但是 f 不是单射也不是满射。

4. 我们考虑映射
$$f:\mathbb{R}^2\to\mathbb{R}^2,\quad (x,y)\mapsto(2x-y+x^2y-2y^5,x+3y-4x^2y^2).$$
证明，存在 0 的开邻域 U 和 V，使得 $f:U\to V$ 是微分同胚。

5. 假设 $f:\mathbb{R}^n\to\mathbb{R}^n$ 是连续可微的。如果存在 $a>0$，使得对任意的 $x,y\in\mathbb{R}^n$，我们都有
$$|f(x)-f(y)|\geqslant a|x-y|,$$
那么，$f:\mathbb{R}^n\to\mathbb{R}^n$ 是 C^1 的微分同胚。(提示：证明 f 在任何一点非退化，像集是闭集也是开集)

6. 假设 $f:\mathbb{R}\to\mathbb{R}$ 是 C^∞ 的并且存在 $0\leqslant\alpha<1$，使得 $|f'(x)|\leqslant\alpha$。定义
$$F:\mathbb{R}^2\to\mathbb{R}^2,\quad (x,y)\mapsto(x+f(y),y+f(x)).$$
证明，$F:\mathbb{R}^2\to\mathbb{R}^2$ 是微分同胚。(提示：可以利用上题结论)

7. $f,g\in C^1(\mathbb{R})$ 是实值函数，我们假设对任意的 $x,y\in\mathbb{R}$，都有 $f'(x)\neq g'(y)$。我们定义 $F:\mathbb{R}^2\to\mathbb{R}^2$：
$$F:(x,y)\mapsto(x+y,f(x)+g(y))$$
证明 F 的像集 $F(\mathbb{R}^2)$ 是开集并且 $F:\mathbb{R}^2\to F(\mathbb{R}^2)$ 是微分同胚。

8. 记 $\mathbf{M}_n$ 为 $n\times n$ 的实系数矩阵的全体。对于 $A\in\mathbf{M}_n$，定义
$$\|A\|=\sup_{x\in\mathbb{R}^n,x\neq 0}\frac{|A\cdot x|}{|x|}.$$
其中 $|x|$ 是欧式空间 $\mathbb{R}^n$ 上的标准距离。证明这是一个范数并且和课上用的范数 $\|\cdot\|_1$(所有元素绝对值之和) 等价，即是存在只依赖于 n 的常数 C 使得对所有的 $A\in\mathbf{M}_n$ 都有 $C^{-1}\|A\|\leq\|A\|_1\leq C\|A\|$。

9. 令 $\mathbf{GL}_n(\mathbb{R})$ 是 $\mathbf{M}_n$ 中可逆矩阵的全体。证明，$\mathbf{GL}_n(\mathbb{R})$ 是 $\mathbf{M}_n$（等同于 $\mathbb{R}^{n^2}$）中的开集并且映射
$$\mathrm{inv}:\mathbf{GL}_n(\mathbb{R})\to\mathbf{GL}_n(\mathbb{R}),\quad A\mapsto A^{-1}$$
是 C^1 的映射。进步一证明，对于 $X\in\mathbf{M}_n$，
$$d\,\mathrm{inv}(A)(X)=-A^{-1}\cdot X\cdot A^{-1}.$$

10. 我们定义平方运算 $\Theta:\mathbf{M}_n\to\mathbf{M}_n$，$A\mapsto A^2$。证明 Θ 是连续可微的并计算它的微分 $d\Theta(A)$。据此证明，存在 $\varepsilon>0$，当 $\|B-\mathbf{I}_{n\times n}\|<\varepsilon$ 时，存在 $A\in\mathbf{M}_n$，使得 $A^2=B$。

11. 假设 $n=2$，$U=\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)$，是否存在 $\mathbf{I}_{2\times 2}$ 的邻域 $\mathcal{I}\subset\mathbf{M}_2$，$U$ 的邻域 $\mathcal{U}\subset\mathbf{M}_2$ 以及连续可微的映射 $\Psi:\mathcal{I}\to\mathcal{U}$，使得 $\Psi(\mathbf{I}_{2\times 2})=U$ 并且对任意的 $A\in\mathcal{I}$，都有 $\Psi(A)^2=A$?（提示：考虑 $H=\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$ 并计算 $d\Psi(\mathbf{I}_{2\times 2})H$）

12. 我们定义立方运算 $\mathcal{C}:\mathbf{M}_n\to\mathbf{M}_n$，$A\mapsto A^3$。证明 $\mathcal{C}$ 是连续可微的并计算它的微分 $d\mathcal{C}(A)$。将 $d\mathcal{C}(A)$ 视为是 $n^2\times n^2$ 的矩阵，进一步证明
$$\|d\mathcal{C}(A)-3\mathbf{I}_{n^2\times n^2}\|\leqslant 6\|A-\mathbf{I}_{n\times n}\|+3\|A-\mathbf{I}_{n\times n}\|^2.$$

13. 令 $B_{\frac{1}{3}}(\mathbf{I}_{n\times n})=\{X\in\mathbf{M}_n|\|X-\mathbf{I}_{n\times n}\|<\dfrac{1}{3}\}$。证明，$\mathcal{C}\big(B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\big)$ 是 $\mathbf{M}_n$ 中的开集并且
$$\mathcal{C}:B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\to\mathcal{C}\big(B_{\frac{1}{3}}(\mathbf{I}_{n\times n})\big)$$
是微分同胚。（提示：可以用题 5 的结论）

14. a 和 b 是实数，我们定义映射
$$f:\mathbb{R}^2\to\mathbb{R}^2,\quad(x,y)\mapsto(x+a\sin y,y+b\sin x).$$
证明：f 是 $\mathbb{R}^2$ 到自身的微分同胚当且仅当 $ab\in(-1,1)$。(提示：对于满射，限制在边长为 2π 的正方形中即可)。

15. 假设 $f(\lambda,x)$ 是 $\mathbb{R}^{m+n}$ 到 $\mathbb{R}^n$ 的连续可微映射，这里 $\lambda\in\mathbb{R}^m$, $x\in\mathbb{R}^n$。假设存在常数 $k<1$ 使得对任意 $\lambda\in\mathbb{R}^m$, 作为 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的连续可微映射 $g(x)=f(\lambda,x)$ 满足 $\|dg(x)\|<k$。证明对任意的 $\lambda\in\mathbb{R}^m$，存在唯一的 $x_\lambda\in\mathbb{R}^n$ 使得 $f(\lambda,x_\lambda)=x_\lambda$。特别的这定义出映射：
$$\varphi:\mathbb{R}^m\to\mathbb{R}^n,\lambda\mapsto x_\lambda.$$

16. 证明上题定义的 φ 是 C^1 的并计算它的微分。(提示：先证明 φ 是连续的，再利用 f 的连续可微性质)

17. 证明，对任意的 $t \in \mathbb{R}$，存在唯一的 C^1 映射 $t \mapsto (x(t), y(t)) \in \mathbb{R}^2$，使得

$$\begin{cases} x &= \frac{1}{2}\sin(x+y) + t - 1, \\ y &= \frac{1}{2}\cos(x-y) - t + \frac{1}{2}. \end{cases}$$

进一步证明 $(x(t), y(t)) = (0,0)$ 当且仅当 $t = 1$ 并计算 $x'(1)$ 和 $y'(1)$。

3.2 习题课补充

对于实数集 $\mathbb{R}$，我们考虑上面的一个 C^1 的群结构 $(\mathbb{R}, \star)$。按照定义，所谓的群结构指的是存在 $\mathfrak{e} \in \mathbb{R}$ 和 C^1 的映射

$$f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto x \star y,$$

满足

a) 对任意的 $x \in \mathbb{R}$，都有 $\mathfrak{e} \star x = x \star \mathfrak{e} = x$；

b) 对任意的 $x, y, z \in \mathbb{R}$，都有 $(x \star y) \star z = x \star (y \star z)$；

c) 对任意的 $x \in \mathbb{R}$，存在唯一的 $y \in \mathbb{R}$，使得 $x \star y = y \star x = \mathfrak{e}$。

我们习惯上把 $\mathfrak{e}$ 称作是单位元并且把 c) 中的 y 记作是 x^{-1}（请不要和倒数搞混）。我们在 $\mathbb{R} \times \mathbb{R}$ 上用 (x_1, x_2) 作为坐标并用 $\partial_1 f$ 和 $\partial_2 f$ 来简记 f 的偏导数。另外，在 $\mathbb{R}$ 上我们所熟知的加法和乘法我们用符号 $+$ 和 $\times$ 来表示。

1. 证明，对任意的 $x, y, z \in \mathbb{R}$，我们有

$$(\partial_2 f)(x \star y, \mathfrak{e}) = (\partial_2 f)(x, y) \times (\partial_2 f)(y, \mathfrak{e})$$

2. 证明，对任意的 $x \in \mathbb{R}$，$(\partial_2 f)(x, \mathfrak{e}) > 0$。

3. 假设 φ 是 $(\mathbb{R}, \star)$ 和 $(\mathbb{R}, +)$ 这两个群之间的 C^1-群同态，即

$$\varphi : \mathbb{R} \to \mathbb{R}$$

是 C^1 的并且对于任意的 $x, y \in \mathbb{R}$，我们都有

$$\varphi(x \star y) = \varphi(x) + \varphi(y).$$

证明，存在常数 $a \in \mathbb{R}$，使得

$$\varphi(x) = a \int_{\mathfrak{e}}^{x} \frac{d\tau}{(\partial_2 f)(\tau, \mathfrak{e})}.$$

4. 证明，对任意的 $a \neq 0$，映射

$$\varphi_a : \mathbb{R} \to \mathbb{R}, \quad x \mapsto a \int_{\mathfrak{e}}^{x} \frac{d\tau}{(\partial_2 f)(\tau, \mathfrak{e})},$$

从 $(\mathbb{R}, \star)$ 到 $(\mathbb{R}, +)$ 的 C^1-群同构，即 φ 是群同态并且是 C^1-的微分同胚（它的逆仍然是群同态）。

5. 证明，$\mathbb{R}$ 上任何一个 C^1 的群结构 $(\mathbb{R}, \star)$ 都是交换的，即对任意的 $x, y \in \mathbb{R}$，$x \star y = y \star x$。

6. 证明，映射

$$f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad \big((x_1, y_1), (x_2, y_2)\big) \mapsto x \star y = (x_1 + x_2, y_1 + e^{x_1} y_2),$$

是 $\mathbb{R}^2$ 上的 C^1 的群结构 $(\mathbb{R}, \star)$，它并不交换。

4 欧式空间中的子流形

我们先给出子流形的定义：

定义 4.1 (d-维子流形). 假设 $M \subset \mathbb{R}^n$ 是非空子集。如果存在整数 $d \geqslant 0$,使得对任意的 $x \in M$,存在开集 $U \subset \mathbb{R}^n$,$x \in U$ 以及 $\mathbb{R}^n$ 中的开集 V 以及微分同胚 $\Phi : U \to V$ 使得 $\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$（后面 $n-d$ 个坐标为 0），我们就称 M 是 $\mathbb{R}^n$ 的一个 d-维的**（微分）子流形**。其中 d 称作是 M 的**维数**，记作 $\dim M$，$n-d$ 称为余维数。

注记. （子流形上的局部行为）

1) 由于 $V \cap (\mathbb{R}^d \times \{0\})$ 是 $\mathbb{R}^d$ 中的开集，所以，$U \cap M$ 与 $\mathbb{R}^d$ 中的一个开集是微分同胚的，所以 M 的每个点的附近（与 $\mathbb{R}^n$ 中的某个开集 U 的交）都与 $\mathbb{R}^d$ 中的某个开集是微分同胚的。我们通常说，局部上 M 是 $\mathbb{R}^d$，因为我们可以进一步取更小的邻域，使得 $U \cap M$ 与 $\mathbb{R}^d$ 中的开球是微分同胚的。

2) 由于在 $V \cap \mathbb{R}^d$ 上可以选取 $y_1, \cdots, y_d$ 作为坐标，所以 $\{\Phi^* y_i\}_{i \leqslant d}$ 可以作为 $M \cap U$ 上的一个局部坐标。所以，我们可以**局部上**用 d 个坐标来标记 M 上的点。当然，我们可能有很多种不同的方式来选取这样的坐标。

我们再回过头来看隐函数定理。如果 $f : \mathbb{R}^{n+p} \to \mathbb{R}^p$ 的光滑映射 (高阶光滑性可保持)，$f(x^*, y^*) = 0$，并且 Jacobi 矩阵 $\big(d_y f\big)(x^*, y^*)$ 可逆。隐函数定理的证明告诉我们存在 (x^*, y^*) 的领域 $\tilde{U}$, $(x^*, 0)$ 的领域 $\tilde{V}$ 以及微分同胚 $F : \tilde{U} \to \tilde{V}$，使得 $f^{-1}(0) \cap \tilde{U} = \tilde{V} \cap \{\mathbb{R}^n \times 0\}$。根据子流形的定义，$f^{-1}(0)$ 在 (x^*, y^*) 附近是 n 维子流形。换句话说，子流形可以通过映射 f 来表达。

推论 4.2 (子流形的判断准则). 假设 $f : \Omega \to \mathbb{R}^p$ 是光滑映射，其中 $\Omega \subset \mathbb{R}^n$ 是开集，$n \geqslant p$。对于 $c \in \mathbb{R}^p$，我们把 c 上**纤维**定义为的 c 在 f 下的逆像：

$$f^{-1}(c) = \{x \in \mathbb{R}^n | f(x) = c\}.$$

如果对任意的 $x \in f^{-1}(c)$，$\mathrm{Rank} df(x) = p$。那么，纤维 $f^{-1}(c)$ 是余维数为 p 的子流形。

注记. 我们注意到定理本身的叙述不需要任何的坐标系，而证明本身很好地解释了不依赖于坐标和选取坐标系之间的关系。我们应该和线性代数的情况做类比：给定 n 维线性空间上的 p 个齐次线性方程，它们的公共零点是 $n-p$ 的线性子空间当且仅当这 p 个线性方程的秩是 p。

我们现在给出隐函数定理的几个应用，首先是用来判定子流形，现在我们只需要做一些简单的代数计算即可：

1) $\mathbb{R}^n$ 中的球面 $\mathbf{S}^{n-1} = \{x \in \mathbb{R}^n | x_1^2 + \cdots + x_n^2 = 1\}$ 是一个 $n-1$ 维的子流形。

球面可以被视作是函数

$$f(x_1, \cdots, x_n) = (x_1)^2 + (x_2)^2 + \cdots + (x_n)^2 - 1$$

的零点集。所以，只要验证它的微分 $df(x)$ 的秩为 1 即可。然而，用矩阵来写

$$\begin{aligned} df(x) &= \left(\frac{\partial f}{\partial x_1}(x), \cdots, \frac{\partial f}{\partial x_n}(x)\right) \\ &= 2(x_1, \cdots, x_n). \end{aligned}$$

由于在 $\mathbf{S}^{n-1}$ 上，$|x|=1$，上面的向量显然非零，所以这是光滑子流形。这比把球面写成函数的图像要方便多。

2) 我们考虑 $\mathbb{R}^4 = \mathbb{C}^2$ 中的曲面

$$\mathbf{T}^2 = \Big\{(z_1, z_2)\big||z_1| = 1, |z_2| = 1\Big\}.$$

如果我们用 (x, y, z, w) 作为坐标，那么这个曲面实际上是

$$f(x, y, z, w) = x^2 + y^2 - 1, \quad g(x, y, z, w) = z^2 + w^2 - 1$$

这两个方程的公共零点集。我们需要计算

$$\mathbb{R}^4 \to \mathbb{R}^2, \quad (x, y, z, w) \mapsto \big(f(x, y, z, w), g(x, y, z, w)\big),$$

的 Jacobi 矩阵并判断它的秩。这个矩阵显然是

$$\begin{pmatrix} 2x & 2y & 0 & 0 \\ 0 & 0 & 2z & 2w \end{pmatrix}.$$

由于在 $\mathbf{T}^2$ 上面，$x^2+y^2=1$，$z^2+w^2=1$，所以上面这个矩阵的秩是 2，从而 $\mathbf{T}^2$ 是 2 维的子流形。它实际上是一个环面（甜甜圈）。

在微积分课程里，我们遇到的子流形通常是维数为 1，2 或者余维数为 1 的子流形。

定义 4.3. 假设 $M \subset \mathbb{R}^n$ 是子流形。如果 $\dim M = 1$，我们就称之为 $\mathbb{R}^n$ 中的一条**光滑曲线**；如果 $\dim M = 2$，我们就称之为 $\mathbb{R}^n$ 中的一个**光滑曲面**；如果 $\dim M = n-1$，我们就称之为 $\mathbb{R}^n$ 中的一个**光滑超曲面**。

注记. 这里曲线的概念和上个学期所讲的曲线略有不同：之前是所谓的参数化曲线，即曲线是一个映射 $\gamma : I \to \mathbb{R}^n$，其中 $I \subset \mathbb{R}$ 是一个区间。这里的曲线只是 $\mathbb{R}^n$ 中的一个子集，然而根据隐函数定理的结论，我们在局部上可以给它一个参数化。

我们关心的是如下的对象，他们有着更特殊的要求：

引理 4.4. 假设 $f : \mathbb{R}^n \to \mathbb{R}$ 是光滑函数，如果对于任意的 $x_0 \in f^{-1}(0)$，$df(x_0) \neq 0$，那么 $f^{-1}(0)$ 是超曲面。我们把这种超曲面称为由一个方程整体定义的光滑超曲面。

特别地，在 $\mathbb{R}^3$ 中，我们有

引理 4.5. 假设 $f(x,y,z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意满足 $f(x_0,y_0,z_0)=0$ 的点 (x_0,y_0,z_0)，我们有

$$\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)(x_0,y_0,z_0)\neq 0,$$

那么

$$\left\{(x,y,z)\in\mathbb{R}^3\middle|f(x,y,z)=0\right\}$$

是光滑曲面。我们把这种超曲面称为由一个方程整体定义的光滑曲面。

例子. $\mathbb{R}^3$ 中的柱面由如下方程

$$x^2+y^2=1$$

定义，这是一个光滑曲面：我们取 $f(x,y,z)=x^2+y^2-1$，此时，

$$\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)=(2x,2y,0)\neq 0,$$

我们还有用两个方程定义的曲线：

引理 4.6. 假设 $f(x,y,z)$ 和 $g(x,y,z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意同时满足 $f(x_0,y_0,z_0)=0$ 和 $g(x_0,y_0,z_0)=0$ 的点 (x_0,y_0,z_0)，如果如下两个向量线性不相关：

$$\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)(x_0,y_0,z_0),\quad \left(\frac{\partial g}{\partial x},\frac{\partial g}{\partial y},\frac{\partial g}{\partial z}\right)(x_0,y_0,z_0),$$

那么

$$\left\{(x,y,z)\in\mathbb{R}^3\middle|f(x,y,z)=0,g(x,y,z)=0\right\}$$

是光滑曲线。我们把这种曲线称为由两个方程整体定义的光滑曲线。

在结束关于隐函数定理的几何讨论之前，我们给出隐函数定理的一个对偶版本。这个版本本身和隐函数是无关的，之所以是对偶的，是因为隐函数定理研究从 $\mathbb{R}^n$ 到 $\mathbb{R}^p$ 的映射，其中 $n\geqslant p$，其结论说映射的逆项或者纤维是子流形。而这个定理（证明非常简单，只要验证定义）研究从 $\mathbb{R}^n$ 到 $\mathbb{R}^{p'}$ 的映射，其中 $n\leqslant p$，其结论说映射的像是子流形。

定理 4.7 (参数化子流形). $\Omega\subset\mathbb{R}^n$ 是开集，$f:\Omega\to\mathbb{R}^{n+p}$ 是光滑映射。假设对于点 $x^*\in\Omega$，有 $\mathrm{rank}\ df(x^*)=n$，那么，存在开集 $U\subset\mathbb{R}^n$，$x^*\in U$，使得 $f(U)\subset\mathbb{R}^{n+p}$ 是 n 维的子流形。

证明: 在 Ω 上选取坐标系统 $x_1,\cdots,x_n$，在 $\mathbb{R}^{n+p}$ 上选取坐标坐标系统 $y_1,\cdots,y_n,z_1,\cdots,z_p$。我们把 f 用坐标分量写成

$$f(x)=(f_1(x),\cdots,f_{n+p}(x)),$$

其中每个 $f_i(x)$ 都是 Ω 上的光滑函数。由于有 $\mathrm{rank}\ df(x^*)=n$，通过调换 $\{x_i\}$ 坐标的指标，我们不妨假设

$$\det\left(\frac{\partial f_i}{\partial x_j}\right)_{i,j\leqslant n}(x^*)\neq 0.$$

取 f 的前 n 个分量得到映射 $F(x)=(f_1(x),f_2(x),\cdots,f_n(x))$。根据上述行列式非零的条件，$F$ 为局部微分同胚（反函数定理）。令 U 为 x 在 Ω 中的开集，$F:U\to F(U)$ 为这个微分同胚并令 $V=F(U)$。此时，我们考察 $f(U)\subset\mathbb{R}^{n+p}$。对于 $x\in U$，令 $y=F(x)$，$f(x)\in\mathbb{R}^{n+p}$ 可以写成

$$\begin{aligned}(f_1(x),\cdots,f_n(x);f_{n+1}(x),\cdots,f_{n+p}(x))&=(y;f_{n+1}(x),\cdots,f_{n+p}(x))\\&=(y,f_{n+1}(F^{-1}(y),f_{n+2}(F^{-1}(y)),\cdots,f_{n+p}(F^{-1}(y)).\end{aligned}$$

所以，它可以被视作是 V 上定义的向量值函数 $y\mapsto f_{n+1}(F^{-1}(y),f_{n+2}(F^{-1}(y)),\cdots,f_{n+p}(F^{-1}(y))$ 的图像，这是一个用 y 做为坐标来参数化的子流形。 $\square$

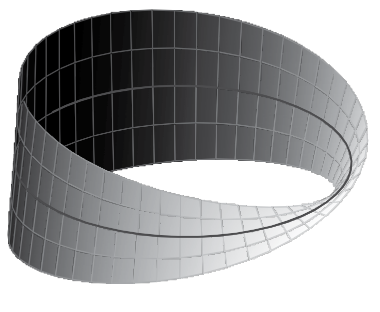
最后我们讨论 $\mathbb{R}^3$ 中的一个经典曲面：$\mathcal{M}$**öbius 带**。我们用参数的方式可以把它写作

$$\mathcal{M}=\left\{(x,y,z)\in\mathbb{R}^3\middle|\left\{\begin{array}{ll}x=(1+r\cos\frac{\theta}{2})\cos\theta, & \\ y=(1+r\cos\frac{\theta}{2})\sin\theta, & \theta\in[0,2\pi],r\in(-1,1)\\ z=r\sin\frac{\theta}{2} & \end{array}\right.\right\}.$$

适当改变一下书写方式更有启发：

$$(x,y,z)=(\cos\theta,\sin\theta,0)+(r\cos\frac{\theta}{2}\cos\theta,r\cos\frac{\theta}{2}\sin\theta,r\sin\frac{\theta}{2}).$$

当 $r=0$ 时，我们得到一个用 θ 来参数化的基本圆周 $(\cos\theta,\sin\theta,0)$。给定 ϑ，就给定了这个圆周上的一个点 $(\cos\theta,\sin\theta,0)$。固定这个 θ，当 r 在 $(-1,1)$ 之间变化的时候，我们就在这个点插上了一个小线段并且这个线段的方向的变化是 θ 的一半。所以，当 θ 从 0 变化到 2π，基本圆周上的点旅行了一周又回到了原来的点，但是它上面所插的小线段只转动了 $180°$。



我们下面证明一个不明显的命题，为此，我们先选择好的记号：把 $r=0$ 光滑圆周写成参数形式：

$$\gamma:[0,2\pi]\to\mathbb{R}^3,\quad\theta\mapsto(\cos\theta,\sin\theta,0).$$

记作

$$e(\theta)=(\cos\theta,\sin\theta,0),\quad e_0=(0,0,1).$$

据此，我们把 M 的参数化写成：

$$\mathcal{M}=\left\{(x,y,z)=e(\theta)+r\cos\frac{\theta}{2}e(\theta)+r\sin\frac{\theta}{2}e_0,\theta\in[0,2\pi],r\in(-1,1)\right\}.$$

根据刚刚证明的定理，这是 $\mathbb{R}^3$ 中的一个光滑曲面。

命题 4.8. 不存在光滑函数 $f\in C^\infty(\mathbb{R}^3)$，使得

$$\mathcal{M}=\big\{(x,y,z)\big|f(x,y,z)=0\big\}$$

且 $df(x,y,z)$ 在 $\mathcal{M}$ 上的每点处不为零，即 $\mathcal{M}$ 不能只用一个方程零点来整体定义。

注记. 我们需要运用关于由一个方程定义的光滑曲面的特殊性质（基于连续函数的介值定理）：给定 $f\in C^\infty(\mathbb{R}^3)$，假设对任意的 $x\in\Omega$，$df(x)\neq 0$，$f^{-1}(0)\subset\Omega$ 是光滑子流形并且由一个方程定义。任意选定 $x\in\Omega$，考虑 x 处的邻域 U（足够小）。由于 $df(x)\neq 0$，我们不妨选 $v\in\mathbb{R}^n$，使得

$$\nabla_v f(x)=df(x)(v)>0.$$

从而，在当 $\varepsilon>0$ 足够小时，$p_\pm=x\pm\varepsilon v\in U$ 并且 $f(p_+)>0$，$f(p_-)<0$。我们声明，任何一条连接 p_+ 和 p_- 的曲线都要经过 $f^{-1}(0)$，即对任意的连续映射 $\gamma:[0,1]\to X$，其中起点 $\gamma(0)=p_-$ 和终点 $\gamma(1)=p_+$)，一定存在 $t_0\in(0,1)$，使得 $\gamma(t_0)\in f^{-1}(0)$（等价于说 $f(\gamma(t_0))=0$)。实际上，$f(\gamma(t))$ 为连续函数，按照 $p_\pm$ 的选取方式，我们有

$$f(\gamma(0))=f(p_-)<0,\quad f(\gamma(1))=f(p_+)>0,$$

所以连续函数的介值定理就给出了 t_0。

我们现在证明关于 Möbius 带的命题：

证明*:* 假设 f 是 $\mathcal{M}$ 的定义方程，满足在 $\mathcal{M}$ 上 df 不为 0。根据定义有

$$f(e(\theta)+r\cos\frac{\theta}{2}e(\theta)+r\sin\frac{\theta}{2}e_0)=0,\quad\forall\theta\in[0,2\pi],\quad r\in(-1,1).$$

对 θ 求导数并在 $r=0$ 处取值可得

$$\nabla f(e(\theta))\cdot e'(\theta)=0$$

同样的对 r 求导数并在 $r=0$ 处取值可得

$$\nabla f(e(\theta))\cdot e_1(\theta)=0,\quad e_1(\theta)=\cos\frac{\theta}{2}e(\theta)+\sin\frac{\theta}{2}e_0$$

现在取

$$\nu(\theta)=\cos\frac{\theta}{2}e_0-\sin\frac{\theta}{2}e(\theta)$$

注意到 $\nu(\theta)$，$e'(\theta), e_1(\theta)$ 两两互相垂直并且是单位向量，由于 df 非零，这表明

$$\nabla f(e(\theta)) \cdot \nu(\theta) \neq 0, \quad \forall \theta \in [0, 2\pi].$$

注意到

$$\nabla f(e(0)) \cdot \nu(0) = \nabla f(1,0,0) \cdot e_0, \quad \nabla f(e(2\pi)) \cdot \nu(2\pi) = -\nabla f(1,0,0) \cdot e_0.$$

由连续性，可知道存在 $\theta \in [0, 2\pi]$ 使得

$$\nabla f(e(\theta)) \cdot \nu(\theta) = 0$$

矛盾！ □

5　子流形的切空间

我们已经证明如果 $f : \mathbb{R}^n \to \mathbb{R}^m$ 满足正确的非退化条件，那么反函数定理等价于说对于 $0 \in \mathbb{R}^m$，它的逆像 $f^{-1}(0)$ 是 $\mathbb{R}^n$ 中的子流形，其中，codim $f^{-1}(0) = m$。我们注意到，$0 \in \mathbb{R}^m$ 是 0-维的子流形，它在 $\mathbb{R}^n$ 中的余维数是也是 m。如果 $S \subset \mathbb{R}^m$ 是子流形，我们同样地可以考虑它的逆像

$$f^{-1}(S) = \left\{x \in \mathbb{R}^n \middle| f(x) \in S\right\}.$$

我们有如下漂亮的定理：

命题 5.1 (子流形的原像)**.** *给定开集 $\Omega \subset \mathbb{R}^n$ 和光滑映射 $f : \Omega \to \mathbb{R}^m (n \geq m)$，$S \subset \mathbb{R}^m$ 是子流形。如果对任意的 $x \in f^{-1}(S)$，rank $df(x) = m$，那么 $f^{-1}(S) \subset \Omega$ 是余维数为 codim S 的子流形。*

证明： 假设 $\dim S = s$。由于子流形的判断是局部性质，为叙述便利，不妨假设存在 $\mathbb{R}^m$ 上的微分同胚 Φ 使得 $S = \Phi^{-1}(\mathbb{R}^s \times \{0\})$。定义从 $\mathbb{R}^n$ 到 $\mathbb{R}^{m-s}$ 的映射

$$F(x) = ((\Phi \circ f)_{s+1}, (\Phi \circ f)_{s+2}, \ldots, (\Phi \circ f)_m)$$

由于 Φ 是微分同胚，df 满秩 (为 m)，因而可知 dF 的秩为 $m - s$。这样由于

$$f^{-1}(S) = (\Phi \circ f)^{-1}(\mathbb{R}^s \times \{0\}) = F^{-1}(0)$$

可知 $f^{-1}(S)$ 是 $n - (m - s)$ 维的子流形。 □

我们以下总是假定 $M^d \subset \mathbb{R}^n$ 是一个 d 维光滑子流形。我们要发展子流形上的微分学，这是我们对多元微分学的自然推广，在实际中有着重要的应用。为此，首先定义最重要的一个几何对象：切空间。我们应该做如下的类比：所谓的微分，是映射在一点处的线性逼近，反函数 / 隐函数定理讲的是映射在一点处的线性逼近决定了映射自身的若干性质。我们应该把切空间视作是流形在一点处用“线性”子流形进行逼近。

首先，我们要考虑的是通过一点 $p \in M$ 的**参数化的 C^∞-曲线**（可以要求是 C^1 的），即

$$\gamma:(-\varepsilon,\varepsilon)\to\mathbb{R}^n,\ \ t\mapsto\gamma(t),\ \varepsilon>0,\ \ \gamma(0)=p.$$

我们将

$$\gamma'(0)=\frac{d}{dt}\gamma\Big|_{t=0}$$

称作该曲线在 p 处的**切向量**。如果 $\gamma\big((-\varepsilon,\varepsilon)\big)\subset M$，我们就说 γ 是M **上的参数曲线**。所谓的 M 在 p 的切空间，就是由在 M 上的通过 p 曲线的曲线的切向量所构成的空间：

定义 5.2. *微分子流形在 $p\in M$ 处的***切空间** *T_pM 是 $\mathbb{R}^n$ 的子空间，它的定义如下：*

$$T_pM=\{\gamma'(0)\in\mathbb{R}^n\big|\gamma \text{ 是经过 } p \text{ 点的 } M \text{ 上的一条参数曲线}\}.$$

命题 5.3 (切空间在微分同胚下的推进 (pushforward))**.** *若 $\Phi:U\to V$ 是微分同胚，$M\subset U$ 是子流形，$p\in U$，其中 U 和 V 是 $\mathbb{R}^n$ 中的开集。那么，$\Phi(M)$ 是子流形并且 $\Phi(p)$ 的切空间为*

$$T_{\Phi(p)}\Phi(M)=d\Phi\Big|_{x=p}(T_pM).$$

证明: 由于 Φ 是微分同胚，显然 $\Phi(M)\subset V$ 是子流形。考虑 M 上经过 p 点的一条参数曲线

$$\gamma:(-\varepsilon,\varepsilon)\to M,\ \ t\mapsto\gamma(t).$$

这样可以得到 $\Phi(M)$ 上经过 $\Phi(p)$ 的参数化曲线 $\Phi(\gamma(t))$。由链式法则可得

$$\frac{d}{dt}\Big|_{t=0}\Phi(\gamma(t))=d\Phi\gamma'(0)$$

这表明

$$d\Phi\Big|_{x=p}(T_pM)\subset T_{\Phi(p)}\Phi(M).$$

由于 Φ 是微分同胚，考虑逆映射 Φ^{-1} 可知

$$d\Phi^{-1}\Big|_{x=\Phi(p)}(T_{\Phi(p)}\Phi(M))\subset T_pM.$$

由于 $d\Phi^{-1}\Big|_{x=\Phi(p)}d\Phi\Big|_{x=p}=I$ 单位矩阵，故而命题成立。 □

根据这个结论我们有

定理 5.4. *T_pM 是 $\mathbb{R}^n$ 的一个 d 维线性子空间。*

证明: 根据子流形的定义，存在 p 的领域 U, $\mathbb{R}^n$ 的开集 V 以及微分同胚 $\Phi:U\to V$ 使得 $M\cap U=\Phi^{-1}(V\cap\mathbb{R}^d\times\{0\})$。根据上面的命题

$$T_pM=d\Phi^{-1}T_{\Phi(p)}\Phi(M)$$

注意到 $\Phi(M\cap U)=V\cap\mathbb{R}^d\times\{0\}$。在 $\Phi(p)$ 这点附近，子流形是 d-维超平面 $\mathbb{R}^d\times\{0\}$，其切空间为 d-维线性子空间 $\mathbb{R}^d\times\{0\}$，特别的有 T_pM 是 d-维线性子空间。 □

如果子流形是通过光滑映射来给出，那么切空间可以通过映射的微分计算出。

命题 5.5. 假设 f 是从 $\mathbb{R}^{d+n} \to \mathbb{R}^n$ 的光滑映射，并且 $df(p)$ 的秩为 n。记 $M = f^{-1}(f(p))$，那么 $T_pM = \ker\, df(p)$。

证明: 可构造同胚映射 $F: \mathbb{R}^{d+n} \to \mathbb{R}^{d+n}$ 使得 $F(x, y) = (x, f(x, y))$（局部 p 点的邻域即可）。这样有

$$T_pM = (dF^{-1})T_{F(p)}F(M) \Leftrightarrow dF(T_pM) = T_{F(p)}F(M) = \mathbb{R}^d \times \{0\}$$

根据 F 的定义，可知 $v \in T_pM$ 等价于 $df \cdot v = 0$ (参见隐函数定理证明中 dF 的表达式)。 $\square$

例子 (超曲面的切空间与法向量的概念)**.** 假设 M 是 $\mathbb{R}^n$ 中由一个函数 f 整体定义的超曲面（R^3 中的曲面是我们最常见的例子），它的定义函数为 $f \in C^\infty(\mathbb{R}^n)$。此时，我们在 $\mathbb{R}^n$ 选定坐标系。按照定义，对于 $p = (x_1, \cdots, x_n)$，那么

$$T_pM = \ker\, df(p) = \left\{v \in \mathbb{R}^n \middle| df(p)(v) = 0\right\}.$$

用坐标系（矩阵的语言）来写，我们有

$$v = (v_1, \cdots, v_n), \quad df(p) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}\right)(p),$$

所以，

$$df(p)(v) = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(p) = 0.$$

我们假定在 $\mathbb{R}^n$ 中取定一个 *Euclid* 内积 $\langle \cdot, \cdot \rangle$，使得 $\left\{\dfrac{\partial}{\partial x_i}\right\}_{i \leqslant n}$ 恰好是标准正交基。此时，我们用如下的符号表示上述向量

$$\nabla f(p) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n}\right)(p).$$

那么，上面的关系就可以表达为

$$\langle v, \nabla f(p) \rangle = 0.$$

换句话说，$\nabla f(p)$ 是切平面的**法向量**。由于 $\nabla f(p) \neq 0$，我们经常把这个向量正规化成长度为 1 的向量：

$$\frac{\nabla f(p)}{|\nabla f(p)|}.$$

我们习惯上把 $\nabla f(p)$ 称作 f 在 p 处的**梯度**。在绝大多数的应用中，我们都（无意识地）自动假定在 $\mathbb{R}^n$ 中有上述 *Euclid* 内积。

利用上面的讨论，我们很容易计算 $\mathbb{R}^3$ 中曲面的切空间，比如说，我们来看柱面的例子：

例子. 柱面的定义方程是 $f = 0$，其中

$$f(x,y,z) = x^2 + y^2 - 1$$

对任意给定的 (x,y,z)，很容易算出

$$\nabla f(p) = (2x, 2y, 0)$$

所以，我们知道 $(-x,y,0)$ 和 $(0,0,1)$ 都和这个点垂直，所以，这两个向量就张成的切空间。特别地，我们也可以从图形或者这个计算上清楚的看出，切空间 T_pM 是依赖于它的基准点 p。和之前线性子空间的切空间的例子对比，我们强调说不同点的切空间是不同的。我们之所以认为它们有时候是一样的，是因为我们把它们都看成了 $\mathbb{R}^n$ 的子空间所以它们有可能相同。

5.1 习题

数学分析二作业 8

本次作业的提交时间和地点为 **5 月 14 日**的习题课上，逾期视作零分。

1. $M \subset \mathbb{R}^n$ 和 $N \subset \mathbb{R}^m$ 是光滑子流形。证明，$M \times N \subset \mathbb{R}^{n+m}$ 也是光滑子流形。

2. 假设 $\Omega \subset \mathbb{R}^n$ 是开集，$f_1, \cdots, f_k$ 是给定的 k 个光滑函数，其中 $k \leqslant n$。令

$$Z_i = \{x \in \Omega | f_i(x) = 0\}.$$

假设 $x \in \bigcap\limits_{i \leqslant k} Z_i$ 并且 $df_1(x), \cdots, df_k(x)$ 线性无关。证明，存在 x 附近的开集 U，使得对任意的 $\ell \leqslant k$，$\bigcap\limits_{i \leqslant \ell} Z_i$ 是维数为 $n - \ell$ 的子流形。

3. 假设 $f(x, y, z)$ 和 $g(x, y, z)$ 是 $\mathbb{R}^3$ 上的光滑函数，如果对于任意同时满足 $f(x_0, y_0, z_0) = 0$ 和 $g(x_0, y_0, z_0) = 0$ 的点 (x_0, y_0, z_0)，如果如下两个向量线性不相关：

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)(x_0, y_0, z_0), \quad \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right)(x_0, y_0, z_0),$$

证明，

$$\left\{(x, y, z) \in \mathbb{R}^3 \middle| f(x, y, z) = 0, g(x, y, z) = 0\right\}$$

是光滑曲线。

4. $\Omega \subset \mathbb{R}^n$ 是开集，$f : \Omega \to \mathbb{R}^m$ 是光滑映射。证明，f 的图像

$$\Gamma_f = \left\{(x, y) \in \mathbb{R}^{n+m} \middle| x \in \mathbb{R}^n, y \in \mathbb{R}^m, y = f(x)\right\}$$

是 $\mathbb{R}^{n+m}$ 中的 n 维光滑子流形。

5. 证明，如下两个方程定义出 $\mathbb{R}^3$ 中的一条光滑曲线：

$$x^2 + y^2 + z^2 = 14, \quad x^3 + y^3 + z^3 = 36.$$

对于曲线上的每一点，计算它的切空间。

6. 在 $\mathbb{R}^2$ 中考虑如下的集合

$$M = \left\{(x, y) \middle| x^3 + y^3 - 3xy = 1\right\}.$$

证明 M 是光滑曲线并计算在任何给定点 $(x_0, y_0) \in M$ 处的切空间。

7. 假设 $M \subset \mathbb{R}^n$ 是子流形并且 $\dim M \geqslant 1$，$p \in M$。那么，存在光滑曲线（作为映射）

$$\gamma : (-1, 1) \to \mathbb{R}^n,$$

使得 $\gamma(0) = p$，$\gamma'(0) \neq 0$ 并且 $\gamma\big((-1, 1)\big) \subset M$。

8. 令 $V=\big\{(x,y)\in\mathbb{R}^2\big|\ y=|x|\big\}$ 为 $\mathbb{R}^2$ 中的锥。证明，V 不是 $\mathbb{R}^2$ 的光滑子流形。

9. 对任意的 $c\in\mathbb{R}^n$，我们可以定义一个为 n 的实系数多项式 P_c，即我们有映射

$$\mathbb{R}^n\to\mathbb{R}[X],\quad c\mapsto X^n+c_1X^{n-1}+\cdots+c_n.$$

给定 $b\in\mathbb{R}^n$，我们假设 P_b 有一个实数根 $x_b\in\mathbb{R}$ 并且这个根的重数是 1。证明，存在 b 在 $\mathbb{R}^n$ 中的开邻域 U 和 x_b 在 $\mathbb{R}$ 中的开邻域 V，使得对任意的 $c\in U$，P_c 在 V 中恰有一个根 $z(c)$ 并且重数是 1。进一步证明映射

$$V\to\mathbb{R},\quad c\mapsto z(c)$$

是光滑的。

10.（根对系数的光滑依赖性）考虑实系数多项式 $P(X)=X^n+c_1X^{n-1}+\cdots+c_n=0$。如果对 $(c_1^*,\cdots,c_n^*)\in\mathbb{R}^n$, $P(X)$ 恰好有 n 个两两不同的实根 $z_1(c_1^*,\cdots,c_n^*)<\cdots<z_n(c_1^*,\cdots,c_n^*)$。证明，存在 $(c_1^*,\cdots,c_n^*)\in\mathbb{R}^n$ 的开邻域 Ω 和 Ω 上的光滑函数 $z_1,\cdots,z_n$，使得对任意的 $(c_1,\cdots,c_n)\in\Omega$，我们有

$$z_1(c_1,\cdots,c_n)<z_2(c_1,\cdots,c_n)<\cdots<z_n(c_1,\cdots,c_n)$$

并且

$$P(z_1(c_1,\cdots,c_n),z_2(c_1,\cdots,c_n),\cdots,z_n(c_1,\cdots,c_n))=0.$$

11. 考虑 $n\times n$ 的实矩阵 $A=(A_{ij})_{i,j\leqslant n}$，假设 A 有 n 个不同的实特征值 $\lambda_1<\lambda_2<\cdots<\lambda_n$。证明，存在 $\varepsilon>0$，使得对任意的 $n\times n$ 的实矩阵 $B=(B_{ij})_{i,j\leqslant n}$，如果对每对指标 (i,j)，$|A_{ij}-B_{ij}|<\varepsilon$，那么 B 有 n 个不同的实特征值 $\lambda_1(B)<\lambda_2(B)<\cdots<\lambda_n(B)$。进一步证明，如果将每个 λ_i 看做是 B 的系数 B_{ij} 的函数，这些函数在区域 $|B_{ij}-A_{ij}|<\varepsilon$（其中 $i,j\leqslant n$）中是光滑函数。

12. $\mathbf{M}_n(\mathbb{R})$ 是全体 $n\times n$ 的实系数矩阵。我们定义**特殊线性群**

$$\mathbf{SL}_n(\mathbb{R})=\{A\in\mathbf{M}_n(\mathbb{R})\big|\det A=1\},$$

和正交群

$$\mathbf{O}(n)=\{A\in\mathbf{M}_n(\mathbb{R})\big|A\cdot A^T=\mathbf{I}_n\},$$

其中 $\mathbf{I}_n$ 是单位矩阵，A^T 是 A 的转置。证明，$\mathbf{SL}_n(\mathbb{R})$ 是 $\mathbf{M}_n(\mathbb{R})=\mathbb{R}^{n^2}$ 中的光滑子流形并计算它的维数。（提示：将行列式映射 $A\mapsto\det A$ 视作是 $\mathbb{R}^{n^2}$ 上的光滑函数）

13. 证明，$\mathbf{SL}_n(\mathbb{R})$ 在 $\mathbf{I}_n$ 处的切空间是

$$\mathfrak{sl}_n(\mathbb{R}):=\{a\in\mathbf{M}_n(\mathbb{R})\big|\mathrm{tr}a=0.\}$$

14. 证明，$\mathbf{O}(n)$ 是 $\mathbf{M}_n(\mathbb{R})=\mathbb{R}^{n^2}$ 中的光滑子流形并计算它的维数（提示：令 $\mathrm{Sym}_n(\mathbb{R})$ 是全体对称的 $n\times n$ 的实系数矩阵，考虑映射 $f:\mathbf{M}_n(\mathbb{R})\to \mathrm{Sym}_n(\mathbb{R}),\ A\mapsto A\cdot A^T-\mathbf{I}_n$）。

15. 证明，$\mathbf{O}_n(\mathbb{R})$ 在 $\mathbf{I}_n$ 处的切空间是

$$\mathfrak{o}(n):=\{a\in \mathbf{M}_n(\mathbb{R})\big| a+a^T=0.\}$$

16. 令 $G=\mathbf{SL}_n(\mathbb{R})$ 或者 $\mathbf{O}(n)$。证明，映射

$$G\to G,\ g\mapsto g^{-1}$$

是光滑的。

17. 令 $G=\mathbf{SL}_n(\mathbb{R})$（或 $\mathbf{O}(n)$），$\mathfrak{g}=\mathfrak{sl}_n(\mathbb{R})$（或 $\mathfrak{o}(n)$），证明，指数映射 exp 把 $\mathfrak{g}$ 映射到 G 中去。

6 子流形之间的光滑映射

给定两个子流形 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$，其中，M 和 N 的维数是任意的，我们定义它们之间的光滑映射 $f: M \to N$ 和以及 f 的微分 df。最常见的例子中通常 $N = \mathbb{R}$，这实际上是 M 上的光滑函数。

定义 6.1. 假设 M, N 分别是 $\mathbb{R}^m$，$\mathbb{R}^n$ 中的子流形，f 是从 M 到 N 的映射。对任意 $p \in M$，按子流形的定义，存在包含 p 的开集 $U \subset \mathbb{R}^m$，$\mathbb{R}^m$ 中的开集 V 以及微分同胚 $\Phi: U \to V$ 使得 $\Phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$。如果映射

$$f_\Phi = \left(\Phi^{-1}\right)^* f = f \circ \Phi^{-1}$$

是在 $V \cap (\mathbb{R}^d \times \{0\}) \subset \mathbb{R}^d$ 上定义并在 $\mathbb{R}^n$ 中取值的 C^∞ 映射：

$$\begin{tikzcd}
V \cap (\mathbb{R}^d \times \{0\}) \arrow[r, "\Phi^{-1}"] \arrow[dr, "f_\Phi"'] & M \cap U \arrow[d, "f"] \\
 & N \subset \mathbb{R}^n
\end{tikzcd}$$

我们就称 f 是 M 与 N 之间的 C^∞ 的映射并且将所有这样的映射的全体记作 $f \in C^\infty(M, N)$。

首先我们需要说明上面的定义跟映射 Φ 的选取无关。对 $p \in M$，对于任意包含 p 的开集 $U' \subset \mathbb{R}^m$，$\mathbb{R}^m$ 中的开集 V' 以及微分同胚 $\Phi': U' \to V'$，使得 $\Phi(U' \cap M) = V' \cap (\mathbb{R}^d \times \{0\})$。考虑 $\widetilde{U} = U \cap U'$(可微性质是局部性质，因而只需要在 p 点邻域内验证即可)，此时，上述的 Φ 和 Φ' 在 $\widetilde{U}$ 上都能定义，我们令 $\widetilde{V} = \Phi(\widetilde{U})$，$\widetilde{V}' = \Phi'(\widetilde{U})$，我们就有如下的交换图表：

$$\begin{tikzcd}
\widetilde{V} \cap (\mathbb{R}^d \times \{0\}) \arrow[r, "\Phi^{-1}"] \arrow[rr, bend left, "\Psi = \Phi'^{-1} \circ \Phi"] \arrow[dr, "f_\Phi"'] & M \cap \widetilde{U} \arrow[d, "f"] & \widetilde{V}' \cap (\mathbb{R}^d \times \{0\}) \arrow[l, "\Phi'^{-1}"'] \arrow[dl, dashed, "f_{\Phi'}"] \\
 & N \subset \mathbb{R}^n &
\end{tikzcd}$$

很明显，$f_{\Phi'}$ 是 f_Φ 与 Ψ 的复合，而 Ψ 是两个微分同胚的复合也光滑，所以 $f_{\Phi'}$ 光滑。

下面的判别准则更加实用。

引理 6.2 (f 局部上是 $\mathbb{R}^m$ 上映射的限制)**.** f 是从子流形 M 到 N 的光滑映射等价于对任意 $p \in M$，存在包含 p 的开集 $U \subset \mathbb{R}^m$ 以及 C^∞ 的映射 $F: U \to \mathbb{R}^n$，使得

$$f\big|_{U \cap M} = F\big|_{U \cap M}.$$

证明*:* 如果存在光滑映射 F 使得 f 是 F 在子流形上的限制，考虑如下的交换图表：

$$\begin{tikzcd}
V \cap (\mathbb{R}^d \times \{0\}) \arrow[r, "\Phi^{-1}"] \arrow[dr, "f_\Phi"'] & M \cap U \arrow[r, "\iota"] \arrow[d, "f"] & U \arrow[dl, "F"] \\
 & N \subset \mathbb{R}^n &
\end{tikzcd}$$

其中，映射 ι 是包含映射，即

$$\iota : M \cap U \to U, \;\; x \mapsto x.$$

由于 F 是光滑的，因而 $f_\Phi = F \circ \iota \circ \Phi^{-1}$ 是光滑的。

反过来如果 f 是光滑映射，考虑如下的交换图表：

$$\begin{tikzcd}
V \arrow[r, "\Phi^{-1}"] \arrow[ddr, dashed, bend left, "\mathring{F}"] & U \arrow[dd, dashed, bend left, "F"] \\
V \cap (\mathbb{R}^d \times \{0\}) \arrow[u, "\iota"] \arrow[r, "\Phi^{-1}"] \arrow[dr, "f_\Phi"'] & M \cap U \arrow[u, "\iota"'] \arrow[d, "f"] \\
& N \subset \mathbb{R}^n
\end{tikzcd}$$

用坐标来看，$f_\Phi(x)$ 是 $\mathbb{R}^d$ 到 $\mathbb{R}^n$ 的光滑函数，定义 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 的函数 $K(x,y) = f_\Phi(x)$，可知该函数光滑。这样通过复合光滑同胚 Φ 可得到 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 的光滑映射 $F = K(\Phi)$ 满足 f 即是 F 在 M 上的限制。 $\square$

例子. 考虑 $M = N = \mathbf{S}^2$，这是 $\mathbb{R}^3$ 中的单位球面。考虑正交矩阵 $A \in \mathbf{O}(3)$，即 3×3 的方阵 A，使得 $A^T \cdot A = \mathbf{I}$，其中 $\mathbf{I}$ 是单位矩阵。我们可以将 A 看作是 $\mathbb{R}^3$ 上的线性（光滑）映射：

$$A : \mathbb{R}^3 \to \mathbb{R}^3, \;\; x \mapsto A \cdot x.$$

由于正交矩阵 A 保持向量的长度，所以

$$A : \mathbf{S}^2 \to \mathbf{S}^2.$$

这是 $\mathbf{S}^2$ 到自身的光滑映射，我们通常把它称为是 $\mathbf{S}^2$ 上的一个旋转。

当 $N = \mathbb{R}$ 时，我们就定义了 M 上的光滑函数 $C^\infty(M)$。与 $\mathbb{R}^n$ 中一个区域上的光滑函数类似，$C^\infty(M)$ 是一个线性空间。

对于子流形之间的光滑映射，我们可以定义微分：

定义 6.3. 假设 $M \subset \mathbb{R}^m$，$N \subset \mathbb{R}^n$ 是子流形，$f : M \to N$ 是光滑映射。对任意的 $v \in T_pM$，按定义，存在（不止一条）参数化的曲线

$$\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^m,$$

使得 $\gamma(0) = p$ 且 $\gamma'(0) = v$。通过和 f 复合，我们得到 N 上过 $f(p)$ 点的曲线：

$$f \circ \gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n, \;\; t \mapsto f(\gamma(t)).$$

我们把这条曲线的在 $f(p)$ 处切向量定义为

$$df(p)(v) := \frac{d}{dt}\Big|_{t=0} f(\gamma(t)) \in T_{f(p)}N.$$

从而，我们定义了

$$df(p) : T_pM \to T_{f(p)}N.$$

根据子流形光滑映射的性质，f 可由看成是 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 的光滑映射 F 在子流形 M 的限制。特别的根据定义

$$\frac{d}{dt}\Big|_{t=0} f(\gamma(t)) = \frac{d}{dt}\Big|_{t=0} F(\gamma(t)) = dF(p)\gamma'(0).$$

这表明上面的定义跟曲线 $\gamma(t)$ 的选取无关，只与 $\gamma'(0)$ 有关，并且 $df(p)$ 可以看成是 $dF(p)$ 在切空间 T_pM 上的限制，因而是线性映射。

在微分学这一部分，我们最后引进的一个对象叫做向量场，向量场在经典的物理中有着无比重要的应用，我们在后面的课程中会逐步展示这些相关的例子。我们来定义一个子流形上全体切向量的集合：

定义 6.4 (向量场与切丛)**.** 假设 $M \subset \mathbb{R}^n$ 是子流形。对任意的 $p \in M$，T_pM 是在这一点处的 M 的切向量的全体。我们把这些不同点的切向量放在一起定义为 M 的**切丛**，我们强调这是所有的 T_pM 的无交并集，即不同点处的切向量是不同的：

$$TM = \coprod_{p \in M} T_pM.$$

我们有自然的投影映射 $\pi : TM \to M$，对于 $p \in M$，这个映射把 T_pM 中的元素全部映射成 p，即 $T_pM = \pi^{-1}(p)$。

如果对于每个点 $p \in M$，我们都指定一个切向量 $X(p) \in T_pM$，我们就得到了一个映射

$$X : M \to TM, \quad p \mapsto X(p),$$

我们把这样一个映射称作是 M 上的一个**（切）向量场**。

注记 (向量场光滑性的一种折衷的定义)**.** 由于所讨论的子流形 M 在 $\mathbb{R}^n$ 中，对任意的 p，我们总可以将 $X(p)$ 看作是 $\mathbb{R}^n$ 中的向量，所以我们可以把 X 看作是映射 $X : M \to \mathbb{R}^n$（要求 $X(p) \in T_pM$）。此时，我们就可以谈论 X 的光滑性。如果映射 $X : M \to \mathbb{R}^n$ 是光滑的，我们就称 X 为**光滑的（切）向量场**。我们将 M 上的光滑切向量的全体记作 $\Gamma(M, TM)$。

对任意的 M 上的光滑函数 f，任意的 M 上的光滑切向量场 X 和 Y，我们可以定义它们之间的乘法和加法：

$$(fX)(p) = f(p)X(p), \quad (X+Y)(p) = X(p) + Y(p).$$

我们有如下简单的命题：

命题 6.5. 假设 $M \subset \mathbb{R}^n$ 是子流形，那么，$\Gamma(M, TM)$ 是 $C^\infty(M)$-模，即对任意的 M 上的光滑函数 f，任意的 M 上的光滑切向量场 X 和 Y，fX 和 $X+Y$ 都是 M 上的光滑向量场。

由于每点 p 处对应切空间 T_pM，局部上切丛 TM 可以看成是 $\mathbb{R}^n \times \mathbb{R}^n$ 的 $2d$ 维子流形。

定理 6.6. 假设 $M \subset \mathbb{R}^n$ 是子流形，那么

$$TM = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \big| x \in M, v \in T_xM\}$$

是 $\mathbb{R}^{2n}$ 的子流形并且 $\dim TM = 2\dim M$。

证明: 令 $d=\dim M$。根据子流形的定义 M 局部上可以看成是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的光滑同胚映射 φ 在 $\mathbb{R}^d\times\{0\}$ 上的像。注意到 $\mathbb{R}^d\times\{0\}$ 是 $\mathbb{R}^n$ 上的光滑子流形，在任何一点 $(x,0)$ 处的切空间为 $\mathbb{R}^d\times\{0\}$，特别的 $\mathbb{R}^d\times\{0\}$ 的切丛是 $\mathbb{R}^n\times\mathbb{R}^n$ 中 $2d$ 维的光滑子流形。定义映射

$$T\varphi:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n,\quad T\varphi(x,v)=(\varphi(x,y),d\varphi\cdot v)$$

显然 $T\varphi$ 是光滑同胚映射，切丛 TM(局部上) 即是 $\mathbb{R}^d\times\{0\}$ 的切丛在 $T\varphi$ 下的像，因而是光滑子流形。 □

假设 $M\subset\mathbb{R}^m$，$N\subset\mathbb{R}^n$ 是子流形，$f:M\to N$ 是光滑映射。对于固定的点 p，我们知道微分 $df(p)$ 是如下的线性映射：

$$df(p):T_pM\to T_{f(p)}N.$$

当 p 变化时，我们得到映射：

$$df:TM\to TN.$$

我们将它称作是 f 的**微分**或者**切映射**。由于 TM 和 TN 都是子流形，这是子流形之间的映射，我们还可以证明 $df:TM\to TN$ 是光滑映射，我们把证明留作作业。

子流形局部上都是平凡的，但向量场和其上定义的光滑函数则是整体的，蕴含子流形的几何拓扑等与局部坐标系无关的性质。假设 $M\subset\mathbb{R}^m$ 是子流形，X 是 M 上的光滑向量场，对于 $f\in C^\infty(M)$，$p\in M$，我们可以定义 f 在 p 处的方向导数为

$$\nabla_{X(p)}f(p)=\frac{d}{dt}\Big|_{t=0}f(\gamma(t))$$

这里 $X(p)\in T_pM$ 由曲线 $\gamma(t)$ 所定义。当 p 变化时，我们就得到了 M 上的函数 $X(f)=\nabla_Xf$。

命题 6.7. *上面定义的 $X(f)$ 是光滑函数，即*

$$X:C^\infty(M)\to C^\infty(M).$$

证明: 按照定义，如果 X 给定，$X(f)$ 只与 f 和 X 在 M 上的值有关系。所以，我们任意选取 $F\in C^\infty(\mathbb{R}^n)$ 和 $\mathfrak{X}\in C^\infty(\mathbb{R}^n;\mathbb{R}^n)$，使得 $F\big|_M=f$，$\mathfrak{X}\big|_M=X$。那么，上面的定义就是在计算 F 在 $\mathbb{R}^n$ 沿着 $\mathfrak{X}$ 的方向导数，所以 $\nabla_{\mathfrak{X}}F$ 光滑的，它的限制就给出了 ∇_Xf，从而光滑。 □

如果 M 上的可微函数有极值，那么我们也有

命题 6.8. *假设 $M\subset\mathbb{R}^m$ 是子流形，$f\in C^\infty(M)$，p 是 f 的极值点（局部最大或者最小）。那么，$df(p)=0$。*

证明: 根据微分的定义，我们要证明对任意的 $v\in T_pM$，$\nabla_vf(p)=0$ 即可。实际上，我们考虑 M 上通过 p 的曲线

$$\gamma:(-\varepsilon,\varepsilon)\to M,\ \ t\mapsto\gamma(t),$$

其中 $\gamma(0)=p$，$\gamma'(0)=v$。我们有

$$\frac{d}{dt}\Big|_{t=0} f\circ\gamma = (\nabla_v f)(p).$$

由于 p 是 $f\circ\gamma$ 的最大值，所以 $(f\circ\gamma)'(0)=0$，这表明对任意的 $v\in T_pM$，$(\nabla_v f)(p)=0$。 □

利用上面证明的命题，我们可以用几何的眼光来看所谓的 **Lagrange 乘子法**，这是多元微积分在求极值方面的重要方法。

我们先把要讨论的问题说清楚：假设 $\mathbb{R}^n$ 上有光滑函数 $f(x_1,\cdots,x_n)$，我们想要计算它的最大值（比如说），然而，这个问题是所谓的带有**约束条件的**。所谓的约束条件指的是点 $x\in\mathbb{R}^n$ 必须满足 m 个方程：

$$\begin{cases} g_1(x_1,x_2,\cdots,x_n)=0,\\ g_2(x_1,x_2,\cdots,x_n)=0,\\ \qquad\cdots\quad\cdots,\\ g_m(x_1,x_2,\cdots,x_n)=0. \end{cases}$$

我们通常要求 $m\leqslant n$（不能有太多的约束）。我们用更简洁的语言来叙述这个问题。令

$$g:\mathbb{R}^n\to\mathbb{R}^m,\ x\mapsto g(x)=(g_1(x),\cdots,g_m(x)),$$

我们要找到 $x\in g^{-1}(0)$，使得 x 是 g 的局部极大值。以下我们进一步假定 g 是光滑的，$M=g^{-1}(0)$ 并且对任意 $x\in M$，$\mathrm{rank} dg(p)=m$。此时，M 是 $\mathbb{R}^n$ 中的子流形。那么，我们的约束条件极值问题等价于求函数 f 在子流形 M 上的极值。

经典的 Lagrange 乘子法是这样说的：为了解决上面的极值问题，我们应该考虑 $\mathbb{R}^n\times\mathbb{R}^m$ 上的函数

$$L(x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_m)=f(x_1,\cdots,x_n)-\sum_{i=1}^m\lambda_i g_i(x_1,\cdots,x_n),$$

其中实变量 $\lambda_1,\cdots,\lambda_m$ 被称为 **Lagrange 乘子**。有些文献上将上面的函数称作是这个极值问题的 **Lagrange 函数**。Lagrange 乘子法的结论说的是 f 的约束条件极值问题等价于 L 的无条件极值问题，即我们要找 $(x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_m)\in\mathbb{R}^n\times\mathbb{R}^m$ 来实现 L 的极值就可以了。

假设 $x\in M$ 是 f 在 M 上的一个局部极值，那么对任意的 $v\in T_xM$，我们有 $df(x)(v)=(\nabla_v f)(x)=0$，即 $df\big(T_pM\big)\equiv 0$，这表明，$T_pM\subset\ker df(p)$。然而，$T_pM=\ker dg(x)$，所以，我们得到如下的结论：$f\big|_M$ 在 $x\in M$ 处取局部极值的必要条件是

$$\ker df(x)\supset\ker dg(x)=\bigcap_{j=1}^m\ker dg_j(x).$$

这表明在 x 处，存在常数 $\lambda_1,\lambda_2,\cdots,\lambda_m$，使得

$$df(x)=\lambda_1\cdot dg_1(x)+\lambda_2\cdot dg_2(x)+\cdots+\lambda_m\cdot dg_m(x).$$

用矩阵的语言来写，对任意的 $i \leqslant n$，我们有

$$\frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) - \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_i}(x_1, \cdots, x_n) = 0.$$

换句话说，为了找到 f 在约束下的极值，一个必要条件就是找到 $(x_1, \cdots, x_n)$ 满足约束和 m 个实数 $\lambda_1, \cdots, \lambda_m$，使得上面的 n 个式子成立。为了记忆这个等式，我们可以考虑 Lagrange 函数

$$L(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m) = f(x_1, \cdots, x_n) - \sum_{i=1}^{m} \lambda_i g_i(x_1, \cdots, x_n)$$

的微分，它对 x_i 偏导数恰好为上面的 n 个方程而对 λ_i 的偏导数恰好给出约束条件，所以，极值点的必要条件可以等价地写成：

$$dL(x, L) = 0.$$

这就是传统的 Lagrange 乘子法给出的求极值的必要条件。

我们可以给出 Lagrange 乘子法的几个经典的应用。

例子. 如果只有一个约束方程 $g : \mathbb{R}^n \to \mathbb{R}$，我们假设 $g(x) = 0$，我们要求 $f(x)$ 的最大值，此时，上述条件变成了

$$\frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) - \lambda \frac{\partial g}{\partial x_i}(x_1, \cdots, x_n) = 0,$$

其中 $i = 1, 2 \cdots, n$。用梯度来表示，这等价于在 $g^{-1}(0)$ 上找一个点，使得 $\nabla f(x)$ 与 $\nabla g(x)$ 共线。

特别的我们令 $f(x) = d(x, z)^2 = (x_1 - z_1)^2 + \cdots + (x_n - z_n)^2$，即我们要在曲面上 $g(x) = 0$ 上找一个点 x，使得该点到一个给定的点的距离是最短的。中学的经验告诉我们我们需要从 z 点到这个曲面做垂线，我们用 Lagrange 乘子法来解读这个问题。首先，我们计算 ∇f

$$\nabla f(x) = 2(x_1 - z_1, x_2 - z_2, \cdots, x_n - z_n).$$

这是从 x 点到 z 的连线。我们已经证明过 $\nabla g(x)$ 是和曲面 $g^{-1}(0)$ 在这点处的切平面垂直的，所以 $\nabla f(x)$ 与 $\nabla g(x)$ 共线等价于说从 z 到 x 的连线和该曲垂直。

我们利用这个例子也可以看到，这一类问题的解可能不是唯一的，比如说如果 $g(x) = x_1^2 + \cdots + x_n^2 - 1$，这定义了单位球面。如果我们选取 $z = 0$，那么任意一个 x 都是这个极值问题的解。

例子 (一个初等几何问题)**.** 给定平面三角形 $\triangle ABC$（非退化），P 是三角形内部的点，从 P 点向三条边作垂线得到 A'，B' 和 C'，我们要找到这样的 P 使得三条线段长度的乘积

$$|PA'||PB'||PC'|$$

最大。

假设 $\triangle ABC$ 面积为 S 而 $\triangle PBC$，$\triangle PAC$ 和 $\triangle PAB$ 的面积分别为 u，v 和 w。通过简单的计算面积，我们知道这个问题等价于求如下函数的最大值：

$$f(u,v,w)=\frac{8uvw}{|AB||BC||CA|},\quad u+v+w=S.$$

当然，我们要求 u,v,w 都是正数。根据 Lagrange 乘子法，我们要求

$$\nabla f-\lambda\nabla g=0 \quad\Leftrightarrow\quad (vw,uw,uv)=\lambda(1,1,1).$$

这说明 $uv=vw=wz$，所以 $u=v=w$，这表明 f 取最大值时，该点为三角形的重心。

例子 (Hadamard 不等式). 我们考虑这样的函数

$$f:\overbrace{\mathbb{R}^n\times\mathbb{R}^n\times\cdots\times\mathbb{R}^n}^{n\text{ 个}}\to\mathbb{R},\quad (v_1,v_2,\cdots,v_n)\mapsto\det(v_1,\cdots,v_n).$$

上面最右边是把 n 个向量排成一排所组成的矩阵。我们在 $\mathbb{R}^n$ 上用标准的 *Euclid* 内积，假设 $|v_1|=|v_2|=\cdots=|v_n|=1$，我们想求 f 的最大值。

我们观察到 f 是定义在 $\mathbb{R}^{n^2}$ 中的函数，其中，我们取坐标 $v_{ij}\in\mathbb{R}$，$1\leqslant i,j\leqslant n$。我们现在有 n 个约束函数

$$g_i(v)=-1+\sum_{j=1}^n v_{ij}^2.$$

很明显，g_i 们的公共零点定义了微分子流形，它实际上是 $\overbrace{\mathbf{S}^{n-1}\times\mathbf{S}^{n-1}\times\cdots\times\mathbf{S}^{n-1}}^{n\text{ 个}}$，这是一个紧集（有界闭集），所以连续函数 f 在它上面有最大值。我们现在用 Lagrange 乘子法，存在 $\lambda_1,\cdots,\lambda_n$，使得对于任意一个固定的 $1\leqslant i_0\leqslant n$，对任意的 $1\leqslant j\leqslant n$，我们有

$$\frac{\partial f}{\partial v_{i_0j}}-\lambda_{i_0}\frac{\partial g_{i_0}}{v_{i_0j}}=0\quad\Leftrightarrow\quad v_{i_0j}\text{ 的余子式与 }v_{i_0j}\text{ 成比例。}$$

如果我们用 v_{ij}^* 表示 v_{ij} 的余子式，那么（一个矩阵如果有两行一样的话它的行列式值就是 0)，那么当 $i\neq i'$ 时，上面的比例关系表明

$$\sum_{j=1}^n v_{ij}v_{i'j}=0.$$

这表明这些 v_i 两两垂直。此时，矩阵 (v_{ij}) 是正交矩阵，所以 f 的最大值是 1 且等号成立当且仅当 $v_1,\cdots,v_n$ 两两垂直。据此，我们就得到一般情况下的 Hadamard 不等式：

命题 6.9 (Hadamard). 对任意的非零向量 $v_1,v_2,\cdots,v_n\in\mathbb{R}^n$，我们有

$$|\det(v_1,v_2,\cdots,v_n)|\leqslant|v_1||v_2|\cdots|v_n|.$$

上面的不等式取等号当且仅当这些向量两两之间垂直。

最后我们还可以把反函数定理推广到子流形上。

定理 6.10 (反函数定理). 假设 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是两个 d 维子流形，$f: M \to N$ 是光滑映射，$p \in M$，$f(p) = q$。假设

$$df(p): T_pM \to T_{f(p)}N$$

是可逆的，那么存在 $\mathbb{R}^m$ 中包含 p 的开邻域 $U \subset M$ 和 $\mathbb{R}^n$ 中包含 q 的开邻域 $V \subset M$，使得

$$f\big|_{U\cap M}: U \cap M \to V \cap N$$

是微分同胚，即 $f\big|_{U\cap M}$ 有逆并且也是光滑的。

证明: 我们把问题化到 $\mathbb{R}^d$ 的情况。首先，在 p 处取开集 U，微分同胚 $\Phi: U \to U' \subset \mathbb{R}^m$，使得 Φ 将 $U \cap M$ 映射为 $U' \cap \mathbb{R}^d \times \{0\}$；在 q 处取开集 V，微分同胚 $\Psi: V \to V' \subset \mathbb{R}^m$，使得 Ψ 将 $V \cap N$ 映射为 $V' \cap \mathbb{R}^d \times \{0\}$。我们考虑如下的交换图表：

$$\begin{tikzcd}
U' \cap (\mathbb{R}^d \times \{0\}) \arrow[r, "\widehat{f}"] & V' \cap (\mathbb{R}^d \times \{0\}) \\
M \cap U \arrow[u, "\Phi"] \arrow[r, "f"] & N \cap V \arrow[u, "\Psi"]
\end{tikzcd}$$

其中，$\widehat{f} = \Psi \circ f \circ \Phi^{-1}$。

我们令 $\widehat{p} = \Phi(p)$，$\widehat{q} = \Psi(q)$，那么，$\widehat{f}(\widehat{p}) = \widehat{q}$。根据链式法则，我们知道

$$d\widehat{f}(\widehat{p}) = d\Psi(q) \circ df(p) \circ d\Phi^{-1}(\widehat{p})$$

是线性同构，所以我们可以在上述交换图表的第一层用反函数定理（此时，我们用 $\mathbb{R}^d$）中的反函数定理。然后用 Ψ 和 Φ 来复合就得到下面一层上的反函数定理，细节留给同学们自己来验证。 □

隐函数定理的几何形式更容易推广到子流形的情况：

命题 6.11 (子流形的原像). 假设 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是子流形，$f: M \to N$ 是光滑映射，$S \subset N$ 也是 $\mathbb{R}^n$ 的子流形。如果对任意的 $x \in f^{-1}(S)$，$\operatorname{rank} df(x) = \dim N$，那么 $f^{-1}(S) \subset M$ 是子流形并且其维数满足等式

$$\dim M - \dim f^{-1}(S) = \dim N - \dim S.$$

由于这个命题和后面课程的关系不大，我们在此略去它的证明。实际上，证明很简单，我们也可以仿照隐函数定理的方法化成某个 $\mathbb{R}^\ell$ 的情形。

6.1 习题

数学分析二作业 9

本次作业的提交时间和地点为 **5 月 14 日**的习题课上，逾期视作零分。

1. 假设 $M \subset \mathbb{R}^m$，$N \subset \mathbb{R}^n$ 是子流形，$f: M \to N$ 是光滑映射。证明，$d: TM \to TN$ 是子流形之间的光滑映射。

2. 假设实数 $a_{ij} = a_{ji}$，求二次型 $f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ 在单位球面 $\mathbf{S}^{n-1}$ 上的条件极值。

3. 求椭球 $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ 的最大体积的内接长方体。

4. f 是 $\mathbb{R}$ 上的非负的连续函数并且如下的反常积分的值为 1：
$$\int_{-\infty}^{\infty} f(x)dx = 1.$$
假设 $[a, b]$ 是使得 $\int_a^b f(x)dx = \frac{1}{2}$ 的长度最短的区间。证明，
$$f(a) = f(b).$$

5. 假定 a 是给定的正实数，考虑 n 元函数
$$f(x_1, \cdots, x_n) = \sum_{k \leqslant n} x_k \log x_k.$$
试求其最小值，其中我们要求 $x_1 + \cdots + x_n = a$，$x_i > 0$。

6. 考虑 $\mathbb{R}^3$ 中的以原点为中心的单位球面 $\mathbb{S}^2$，其北极为 $N = (0, 0, 1)$，南极为 $S = (0, 0, 1)$。记 Σ 为平面 $z = -1$, 对任意的 $p \in S$，$p \neq N$，N 到 p 的连线与 Σ 恰好相交于一个点 $\pi(p)$，这样我们有一一对映
$$\pi: \mathbf{S}^2 - \{N\} \to \Sigma, \quad p \mapsto \pi(p).$$
这个映射被称作是从北极到南极切平面的**球极投影**。证明，上面定义的球极投影是光滑的共形映射，即是这个映射保持角度：对于任意的 $v, w \in T_p\mathbf{S}^2$，我们有
$$\frac{\langle d\pi(p)(v), d\pi(p)(w) \rangle}{|d\pi(p)(v)||d\pi(p)(w)|} = \frac{\langle v, w \rangle}{|v||w|}.$$
这里的内积是 $\mathbb{R}^3$ 中的标准内积。

6.2 习题课补充

假设 $M_1, M_2 \subset M \subset \mathbb{R}^n$ 均为是微分子流形，如果对所有的 $p \in M_1 \cap M_2$，我们都有

$$T_pM_1 + T_pM_2 = T_pM,$$

我们就称 M_1 与 M_2 (作为 M 的子流形) **横截相交**（特别地，如果 $M_1 \cap M_2 = \emptyset$，它们也是横截相交），记作 $M_1 \pitchfork M_2$。我们的目标是证明，如果 $M_1 \pitchfork M_2$，那么 $M_1 \cap M_2$ 是 M 的子流形。

1. 试找出微分子流形 $M_1, M_2 \subset \mathbb{R}^n$，使得 $M_1 \cap M_2$ 不是子流形。

2. 假设 L_1, L_2 是 $V = \mathbb{R}^n$ 的线性子空间并且 $L_1 + L_2 = V$。证明，$L_1 \cap L_2$ 是 V 的线性子空间并且

$$\operatorname{codim} L_1 \cap L_2 = \operatorname{codim} L_1 + \operatorname{codim} L_2.$$

3. $C_1, C_2 \subset \mathbb{R}^2$ 是光滑曲线。证明，$C_1 \pitchfork C_2$ 的充要条件是 C_1, C_2 在它们的所有交点上都不相切（即它们在相交处的切空间不同）。

4. $C_1, C_2 \subset \mathbb{R}^3$ 是光滑曲线。证明，$C_1 \pitchfork C_2$ 的充要条件是 $C_1 \cap C_2 = \emptyset$。

5. 横截相交性的概念可以延拓到映射：假设 $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是微分子流形，$f : N \to M$ 是光滑映射，$S \subset M$ 是子流形。如果对所有的 $p \in f^{-1}(S)$，都有

$$df(T_pN) + T_{f(p)}S = T_{f(p)}M,$$

   我们称光滑映射 f 与 S 是**横截相交的**并记作 $f \pitchfork S$。

   微分子流形 $M_1, M_2 \subset M$，我们用 $i_1 : M_1 \subset M$ 表示包含映射。证明，$M_1 \pitchfork M_2$ 等价于 $i_1 \pitchfork M_2$ 。

6. 证明，如果 $M_1, M_2 \subset M \subset \mathbb{R}^n$ 是子流形并且横截相交，那么 $M_1 \cap M_2$ 是 M 的子流形并且

$$\text{codim} M_1 + \text{codim} M_2 = \text{codim} M_1 \cap M_2.$$

   （提示：在局部上用函数的零点集来定义子流形）

7. $M \subset \mathbb{R}^m$ 和 $N \subset \mathbb{R}^n$ 是微分子流形，$S \subset M$ 是子流形，$f : N \to M$ 是光滑映射并且 $f \pitchfork S$。证明，$f^{-1}(S)$ 是 N 的子流形并且

$$\dim N - \dim f^{-1}(S) = \dim M - \dim S.$$

   (提示: 如果局部上 $S = g^{-1}(0)$，其中 $g : \mathbb{R}^m \to \mathbb{R}^{m-\dim S}$, 那么 $f^{-1}(S) = (g \circ f)^{-1}(0)$）

8. 如果微分子流形 $M_1, M_2 \subset M$ 横截相交, 那么

$$T_p(M_1 \cap M_2) = T_pM_1 \cap T_pM_2.$$

9. 如果 N 是紧的, $f \pitchfork S$, 证明: 对于任意的光滑映射

$$F : N \times [0,1] \to M,$$

满足 $F(x,0) = f(x)$ (F 称为 f 的一个光滑同伦), 存在 $\varepsilon > 0$, 使得对任意的 $t \in [0,\varepsilon)$, 都有 $f_t \pitchfork S$, 其中 $f_t(x) = F(x,t)$。

(这说明横截性在小形变下是稳定的)

10. 证明: 设 $f : [0,1] \to \mathbb{R}^3$ 是不自相交的光滑曲线, $S \subset \mathbb{R}^3$ 是光滑曲面。证明，对于 $\varepsilon > 0$, 存在 f 的光滑同伦 f_t, 使得

- $f_1 \pitchfork S$;
- $\sup\limits_{t,x\in[0,1]} |f_t(x) - f(x)| < \varepsilon$。

7 Hesse 矩阵与凸函数

我们假设 f 是 Ω 上至少两次连续可微的函数，对于给定的点 $p\in\Omega$，一阶微分可用梯度向量 ∇f 来表示，二阶导数就变成是 ∇f 的 Jacobi 矩阵

$$\nabla^2 f(p)=H_f(p)=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(p)\right)=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2}(p) & \frac{\partial^2 f}{\partial x_1\partial x_2}(p) & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}(p)\\ \frac{\partial^2 f}{\partial x_2\partial x_1}(p) & \frac{\partial^2 f}{\partial x_1^2}(p) & \cdot & \frac{\partial^2 f}{\partial x_2\partial x_n}(p)\\ \cdots & \cdot & \cdots & \cdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1}(p) & \frac{\partial^2 f}{\partial x_n\partial x_2}(p) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(p)\end{pmatrix}$$

称为 f 在点 p 处的 Hesse 矩阵，显然这是个对称的方阵。另外，我们的 Taylor 公式在 2 阶的时候可以写成

$$f(x)=f(x_0)+\nabla f(x_0)\cdot(x-x_0)+\frac{1}{2}(x-x_0)\nabla^2 f(x_0)(x-x_0)^T+o(|x-x_0|^2).$$

由于是对称矩阵总是可以对角化，也即是存在 n 个特征值，因而可利用 Hesse 矩阵的正定性来判定函数在 x_0 处的极值情况。与一维情形类似，我们有

命题 7.1. 给定 $f\in C^2(\Omega)$，其中 $\Omega\subset\mathbb{R}^d$ 是开集，如果 $x_0\in\Omega$ 是 f 的最小值点，那么，$\nabla f(x_0)=0$ 并且 $\nabla^2 f(p_0)$ 是半正定的，也即是特征值非负。

证明: 否则假设 $\nabla^2 f(x_0)$ 存在负特征值 $\lambda<0$，对应特征向量 v，考虑 $x=x_0+tv$ 并利用 Taylor 公式

$$\begin{aligned}f(x)&=f(x_0)+\frac{1}{2}tv\nabla^2 f(x_0)(tv)^T+o(|tv|^2)\\&=f(x_0)+tv\cdot\lambda tv+o(t^2|v|^2)\end{aligned}$$

这里一阶项之前已经证明过为 0, 由于 $f(x)\geq f(x_0)$ 对任意足够小的 t 成立，因而有

$$\frac{o(|tv|^2)}{|tv|^2}\geq-\lambda>0$$

对任意足够小的 t 成立，这不可能！ □

这个证明可以用来证明一个接近于上述命题逆命题的命题：

命题 7.2. 给定 $f\in C^2(\Omega)$，其中 $\Omega\subset\mathbb{R}^d$ 是开集，如果 $x_0\in\Omega$ 满足 $\nabla f(x_0)=0$ 并且 $\nabla^2 f(p_0)$ 是正定的，那么，x_0 是 f 的局部最小值点，即存在开集 $U\subset\Omega$，使得 x_0 是 f 在 Ω 上的最小值。

证明: 由于 $\nabla^2 f(p_0)$ 正定，因而最小特征值 $\lambda_1>0$，特别的对任意 $v\in\mathbb{R}^n$ 有

$$v\nabla^2 f(p_0)v^T\geq\lambda_1|v|^2.$$

根据 Taylor 公式，我们有

$$\begin{aligned}f(x) &= f(x_0) + \frac{1}{2}(x-x_0)\nabla^2 f(x_0)(x-x_0)^T + o(|x-x_0|^2)\\ &\geqslant f(x_0) + \frac{1}{2}\lambda_1|x-x_0|^2 + o(|x-x_0|^2)\end{aligned}$$

根据 $o(|x-x_0|^2)$ 的定义，存在 $\varepsilon > 0$，当 $|x-x_0| < \varepsilon$ 时（我们就选取 $U = B_\varepsilon(x_0)$），我们有 $o(|x-x_0|^2) < \frac{1}{2}\lambda_1\frac{|x-x_0|^2}{2}$。从而

$$f(x) > f(x_0) + \lambda_1\frac{|x-x_0|^2}{4} > f(x_0)$$

这说明 x_0 是 U 中的最小值点。 □

接下来我们将上学期学习过的凸函数推广到高维。给定凸集 $\Omega \subset \mathbb{R}^n$（即任意两点的连线属于 Ω），$f : \Omega \to \mathbb{R}$，如果对任意的 $x, y \in \Omega$，任意的 $t \in [0,1]$，都有

$$f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y).$$

我们就称 f 是**凸函数**。（如果 $-f$ 是凸函数，就称 f 是**凹函数**）

另外，如果对任意的 $x, y \in \Omega$，对任意的 $t \in (0,1)$，都有

$$f\big(tx + (1-t)y\big) < tf(x) + (1-t)f(y).$$

我们就称 f 是**严格凸的**。

同样的高维凸集上的凸函数也满足 Jensen 不等式

命题 7.3 (Jensen 不等式)**.** *假设 $f : \Omega \to \mathbb{R}$ 是凸函数，其中 $\Omega \subset \mathbb{R}^n$ 为凸集，那么对任意的 $x_1, \cdots, x_n \in \Omega$ 和任意的 $t_1, \cdots, t_n \in [0,1]$，其中 $t_1 + t_2 + \cdots + t_n = 1$，我们有*

$$f(t_1x_1 + \cdots + t_nx_n) \leqslant t_1f(x_1) + \cdots + t_nf(x_n).$$

例子. *我们先看几个凸函数的例子*

1) 线性函数 $f(x) = a_1x_1 + \cdots + a_nx_n + b$ 是凸函数。

2) 假设 $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ 是任意一个范数，那么，这是一个凸函数，因为

$$\|tx + (1-t)y\| \leqslant \|tx\| + \|(1-t)y\| = t\|x\| + (1-t)\|y\|.$$

3) 假设 A 是一个 $n \times n$ 的半正定矩阵 $(vAv^T \geq 0)$，那么，二次型

$$f(x) = xAx^T$$

是凸函数。实际上，我们可以验证代数恒等式：

$$(1-t)f(x) + tf(y) - f((1-t)x + ty) = t(1-t)f(x-y).$$

右边的值是非负的。

同一维情形，凸函数是连续函数，实际上局部是 Lipschitz 的。

定理 7.4. 假设 $\Omega \subset \mathbb{R}^n$ 是凸的开集，$f:\Omega \to \mathbb{R}$ 是凸函数，那么，f 是连续函数。

证明: 不妨假设原点 $0 \in \Omega$，并且标准向量 $\pm e_i$ 都在 Ω 中（任取 Ω 中的正交标架即可），我们只需要证明 f 在原点处连续即可。对任意 $x = \sum\limits_{i=1}^n \lambda_i \bar{e}_i$，$\bar{e}_i = e_i$ 或者 $-e_i$ 使得 $\lambda_i \geq 0$，只要 $|x|$ 足够小便有 $\sum\limits_{i=1}^n \lambda_i < 1$，记 $\lambda_0 = 1 - \sum\limits_{i=1}^n \lambda_i$。由 Jensen 不等式

$$f(x) = f(\lambda_0 \cdot 0 + \sum_{i=1}^n \lambda_i \bar{e}_i) \leq \lambda_0 f(0) + \sum_{i=1}^n \lambda_i f(\bar{e}_i).$$

这样可得

$$f(x) - f(0) \leq \sum_{i=1}^n \lambda_i (f(\bar{e}_i) - f(0)) \leq C|x|.$$

这里注意到 $|x|^2 = \sum\limits_{i=1}^n \lambda_i^2$，常数 C 只依赖于 $f(\pm e_i)$ 以及 $f(0)$ 的常数。

为了估计 $f(0) - f(x)$，我们需要把 0 写成 x 与 $\pm e_i$ 的线性组合

$$0 = \frac{1}{1 + \sum\limits_{i=1}^n \lambda_i} x + \sum_{i=1}^n \frac{\lambda_i}{1 + \sum\limits_{i=1}^n \lambda_i} (-\bar{e}_i)$$

这样根据 Jensen 不等式我们有

$$f(0) \leq \frac{1}{1 + \sum\limits_{i=1}^n \lambda_i} f(x) + \sum_{i=1}^n \frac{\lambda_i}{1 + \sum\limits_{i=1}^n \lambda_i} f(-\bar{e}_i)$$

特别的有

$$f(0) - f(x) \leq \sum_{i=1}^n \lambda_i (f(-\bar{e}_i) - f(0)) \leq C|x|$$

□

如果 f 具有一阶导数的话，我们可以用线性函数来刻画凸函数。

命题 7.5. 假设 $\Omega \subset \mathbb{R}^n$ 是凸的开集，$f:\Omega \to \mathbb{R}$ 是可微函数。那么 f 是凸函数等价于

$$f(x) \geq f(x_0) + df(x_0)(x - x_0) \quad \forall x \in \Omega.$$

证明: 首先假设 f 是凸函数，根据定义有

$$f(tx + (1-t)x_0) = f(x_0) + df(x_0)(t(x - x_0)) + o(t(x - x_0)) \leq tf(x) + (1-t)f(x_0)$$

这样可知

$$f(x) \geq f(x_0) + df(x_0)(x - x_0) + \frac{o(t(x - x_0))}{t|x - x_0|}|x - x_0|$$

对任意 $0 < t \leq 1$ 成立，让 $t \to 0$，可得 $f(x) \geq f(x_0) + df(x_0)(x - x_0)$。

反过来如果 $f(x) \geq f(x_0) + df(x_0)(x - x_0)$ 对任意 $x, x_0 \in \Omega$ 成立，那么有

$$\begin{aligned} f(x) &\geq f(tx + (1-t)x_0) + df(tx + (1-t)x_0)((1-t)(x - x_0)), \\ f(x_0) &\geq f(tx + (1-t)x_0) + df(tx + (1-t)x_0)(t(x_0 - x)), \quad \forall 0 \leq t \leq 1 \end{aligned}$$

此即表明 $tf(x) + (1-t)f(x_0) \geq f(tx + (1-t)x_0)$。 □

如果我们的函数是 2 阶可微的，类似于 1 维的情况，我们也有对凸函数的二阶导数的判定：

命题 7.6. *假设 $\Omega \subset \mathbb{R}^n$ 是凸的开集，$f : \Omega \to \mathbb{R}$ 是 C^2 的函数，那么，f 是凸函数当且仅当 f 在每个点处的 Hesse 矩阵都是半正定的。进一步，如果在每个点处 Hesse 矩阵都是正定的，那么 f 是严格凸的。*

证明: 首先假设 f 是凸函数，根据上面命题结论，在 x_0 处将 f 进行 Taylor 展开有

$$\begin{aligned} f(tx + (1-t)x_0) &= f(x_0) + tdf(x_0)(x - x_0) + t^2 \frac{1}{2}(x - x_0)\nabla^2 f(x_0)(x - x_0)^T + o(t^2|x - x_0|^2) \\ &\geq f(x_0) + tdf(x_0)(x - x_0). \end{aligned}$$

让 $t \to 0$，可得

$$(x - x_0)\nabla^2 f(x_0)(x - x_0)^T \geq 0.$$

这就表明 Hesse 矩阵是半正定的。

反过来如果 f 的 Hesse 矩阵是半正定 (或者正定) 的，固定 $x, x_0 \in \Omega$，定义关于 t 的函数

$$H(t) = f(x_0 + t(x - x_0)) - f(x_0) - tdf(x_0)(x - x_0).$$

计算可得

$$H'(t) = df(x_0 + t(x - x_0))(x - x_0) - df(x_0)(x - x_0), \quad H''(t) = (x - x_0)\nabla f(x_0 + t(x - x_0))(x - x_0)^T$$

这表明函数 $H(t)$ 是凸函数（或者严格凸函数），注意到 $H(0) = H'(0) = 0$，特别的有 $H(t) \geq 0$(或者 $H(t) > 0$)，对任意 $0 < t \leq 1$。这就表明函数 f 是凸函数 (或者严格凸函数)。 □

7.1 习题

数学分析二作业 10

本次作业的提交时间和地点为 **5 月 21 日**的习题课上，逾期视作零分。

1. $\Omega \subset \mathbb{R}^n$ 为凸集，f 和 g 为 Ω 上的凸函数。试证明，

$$\max(f, g) : \Omega \to \mathbb{R}, \quad x \mapsto \max(f(x), g(x))$$

也是凸函数。

2. 在 $\mathbb{R}^3$ 中考虑方程

$$z^3 + 2z + e^{z-x-y^2} = \cos(x - y + z)$$

所定义的集合 Σ。证明，存在开集 $U \subset \mathbb{R}^2$ 使得 $0 \in U$ 以及函数 $\varphi : U \to \mathbb{R}$，使得 $\Sigma \cap \{(x, y, z)|(x, y) \in U\}$ 是函数 $\varphi(x, y)$ 的图像 $\{(x, y, z)|z = \varphi(x, y), (x, y) \in U\}$。进一步计算 $\Sigma \cap \{(x, y, z)|(x, y) \in U\}$ 在 $(0, 0, \varphi(0, 0))$ 处的切空间并计算 φ 的 Hesse 矩阵。

3. (Hadamard 引理) $\Omega \subset \mathbb{R}^n$ 是凸的开集，$0 \in \Omega$，函数 $f \in C^\infty(\Omega)$ 且 $f(0) = 0$。证明，存在光滑向量值函数 $g : \Omega \to \mathbb{R}^n$，使得对任意的 $x \in \Omega$，都有

$$f(x) = x \cdot g(x), \quad g(0) = df(0).$$

4. 我们仍然假设 f 是区域 Ω 上的光滑函数满足 $f(0) = 0$，$df(0) = 0$(这样的点称为 f 的临界点)，如果 $\nabla^2 f(0)$ 非退化，则称 0 为 f 的非退化临界点。证明，存在 0 的开邻域 $U \subset \Omega$ 以及 U 上 $n \times n$ 光滑对称可逆矩阵函数 $H(x)$(矩阵每个元素是光滑函数，对固定的 x, $H(x)$ 是对称可逆 $n \times n$ 矩阵) 使得对任意 $x \in U$ 有 $f(x) = xH(x)x^T$。(提示：证明类似上题)

5. 假设 U 是 $\mathbb{R}^n$ 中包含原点的开邻域，$H(x)$ 是定义在 U 上的 $n \times n$ 的对称光滑矩阵实函数。如果 $H(0)$ 非退化并且有 n 个不同的特征值，证明存在 0 的邻域 $V \subset U$，使得 $H(x)$ 在 V 上每一点 x 处都存在 n 个不同的特征值 $\lambda_i(x)$ 以及对应的单位特征向量 $v_i(x)$, $1 \le i \le n$ 使得 $\lambda_i(x)$，$v_i(x)$ 是 V 上的光滑 (向量值) 函数。(提示：可利用上次习题结论)

6. (Morse 引理) 假设 f 是 $\mathbb{R}^n$ 中包含原点的区域 Ω 上的光滑函数，满足 $f(0) = 0$，$df(0) = 0$，$\nabla^2 f(0)$ 非退化，即是 0 为 f 的非退化临界点。证明存在非负整数 $m \le n$，存在 0 的开邻域 V 和 $U \subset \Omega$ 以及光滑同胚映射 $\varphi : V \to U$ 使得

$$\varphi(0) = 0, f(\varphi(y)) = -\sum_{i=1}^{m} y_i^2 + \sum_{i=m+1}^{n} y_i^2, \quad \forall y \in V.$$

这里的非负整数 m 称为非退化临界点 0 的**指标**，只依赖于 $\nabla^2 f(0)$，与局部坐标系的选取无关。特别的在 U 上 0 是 f 唯一的非退化临界点，也即是非退化临界点是孤立的。(提示：可以先做线性映射使得 $\nabla^2 f(0)$ 有 n 个不同的特征值)

8 子流形上的积分：第一型曲线和曲面积分

子流形以及子流形上的微分理论我们已近有个初步的了解，在子流形上定义积分理论归结到定义一种测度。但子流形或者更加一般的流形是通过局部坐标系来定义的，子流形上的测度必须与坐标系的选取无关。考虑最简单的曲线，曲线上两点之间的距离 (限制在曲线上) 可以定义成两点之间的弧长，近似等于折线的长度之和。如果曲线连续可微，那么有

$$L(\gamma) = \int_0^1 |\gamma'(t)|dt.$$

换句话说曲线上的长度是通过局部折线来逼近的。

接下来我们考虑 $\mathbb{R}^n$ 中的 d 维的子流形，我们先来看平面的情形。

例子. 假设 $v_1, v_2, \ldots, v_d$ 是 $\mathbb{R}^n$ 上 d 个线性无关的向量，张成 $\mathbb{R}^n$ 中 d 维平面。任取与 $v_i, 1 \le i \le d$ 垂直的单位向量 v_j，$j = d+1, \ldots, n$ 使得 v_i，$1 \le i \le n$ 构成 $\mathbb{R}^n$ 的一组基，并且 $v_{d+1}, \ldots, v_n$ 互相垂直。假设 A 是 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性变换使得 $Ae_i = v_i$。注意到 A 将 $\mathbb{R}^n$ 中的单位方体 $E = \{0 \le x_i \le 1, 1 \le i \le n\}$ 映为 v_1, v_2，$\ldots, v_n$ 张成的平行多面体 $A(E)$。根据之前的测度理论有

$$m(A(E)) = |det(A)|m(E) = |det(A)|$$

记 G 为 v_1, v_2, $\ldots, v_d$ 构成的 $d \times n$ 矩阵，根据 v_{d+1}，$\ldots, v_n$ 的选取可知

$$AA^T = \begin{pmatrix} GG^T & 0 \\ 0 & I_{(n-d)\times(n-d)} \end{pmatrix}$$

特别的有 $det(A) = \sqrt{det(GG^T)} = \sqrt{det(v_i \cdot v_j)_{1 \le i,j \le d}}$。另一方面，根据 v_{d+1}，$\ldots, v_n$ 的选取，$v_1, v_2, \ldots, v_d$ 张成的平行多面体的体积为 $\sqrt{det(GG^T)}$。

现在我们考虑一般的子流形 $M \subset \mathbb{R}^n$，我们假设 $\dim M = d < n$。为了定义子流形上的积分，我们需要研究对任意 $\mathbb{R}^n$ 中的开集 $\tilde{U}$，$\tilde{U} \cap M$ 的面积或体积。由测度的平移不变形，可不妨假设 $0 \in M \cap \tilde{U}$。根据子流形的定义，存在光滑同胚映射 $\tilde{\Phi} : \tilde{V} \to \tilde{U}$，其中 $\tilde{V}$ 是 $\mathbb{R}^n$ 中的开集，使得 $\tilde{\Phi}(\tilde{V} \cap (\mathbb{R}^d \times \{0\})) = \tilde{U} \cap M$，$\tilde{\Phi}(0) = 0$。记 $U = \tilde{U} \cap M$，$V = \{x \in \mathbb{R}^d, (x, 0) \in \tilde{V}\}$，$\Phi = \tilde{\Phi}|_V : V \to \mathbb{R}^n$, 根据假设 $rank(d\Phi) = d$。考虑 V 中的方体 I

$$I = \{(x_1, x_2, \ldots, x_d), \quad 0 \le x_i \le \epsilon_i\}$$

这里 ϵ_i 是小的常数。特别的有

$$\Phi(\epsilon_i e_i) - \Phi(0) = d\Phi(0)\epsilon_i e_i + o(\epsilon_i).$$

这样 $\Phi(I)$ 的面积可近似的看成是 $\mathbb{R}^n$ 中向量 $\Phi(0) + \epsilon_i d\Phi(0)e_i$ 张成的平行多面体的面积。根据上面的讨论有

$$m(\Phi(I)) \approx \prod_{i=1}^{d} \epsilon_i \sqrt{det(d\Phi(0)^T d\Phi(0))} = \sqrt{det(d\Phi(0)^T d\Phi(0))}m(I).$$

当 V 上的分割足够小，由积分的定义可知

$$m(U) = m(\Phi(V)) = \int_V \sqrt{det(d\Phi(x)^T d\Phi(x))}dx$$

这样在子流形上定义出来的面积跟同胚映射 $\tilde{\Phi}$ 的选取无关 (相当于对 V 做坐标变换)。事实上如果有两个同胚映射 $\tilde{\Phi}_i$，$i = 1, 2$，类似的有 Φ_i, V_i。$\tilde{\Phi}_1, \tilde{\Phi_2}$ 诱导了 V_1 到 V_2 的同胚映射 H。这样有

$$\Phi_1 = \Phi_2 \circ H$$

由链式法则

$$d\Phi_2(H(y)) = d\Phi_2 dH, \quad d\Phi_2(H(y))^T d\Phi_2(H(y)) = dH^T d\Phi_2^T d\Phi_2 dH.$$

根据坐标变换公式可知

$$\begin{aligned}\int_{V_2} \sqrt{det(d\Phi_2(x)^T d\Phi_2(x))}dx &= \int_{V_1} \sqrt{det(d\Phi_2)^T d\Phi_2)}|_{H(y)}|J_H(y)|dy \\ &= \int_{V_1} \sqrt{det(d\Phi_2(H(y))^T d\Phi_2(H(y)))}dy.\end{aligned}$$

作为总结，我们可以定义子流形上的曲面测度。

定义. *假设 M 是 $\mathbb{R}^n$ 中的 d 维光滑子流形 (连续可微的子流形即可)，Φ 是 M 的局部坐标化，即是存在 $\mathbb{R}^d$ 中开集 V 使得 Φ 是 V 到 $\mathbb{R}^n$ 的光滑映射，$rank(d\Phi) = d$，$\Phi(x_0) = y_0 \in M$。那么 y_0 附近的曲面测度可定义为*

$$d\sigma(y) = \sqrt{\det\left(d\Phi^T d\Phi\right)}dx,$$

对于 M 上定义的非负可测函数 f(通过同胚映射 Φ 转化成 $\mathbb{R}^d$ 上的可测性)，可定义子流形上 f 的积分

$$\int_U f d\sigma = \int_V f\left(\Phi(x)\right)\sqrt{\det\left(d\Phi^T d\Phi\right)}dx.$$

注记. *上述定义可看成是积分坐标变换公式的自然推广,注意到积分坐标变化公式中 $|J_\Phi| = |det(d\Phi)|$，而 $d\Phi$ 即是 Φ 在一点的微分。对应的子流形，$d\Phi$ 可以看成是从 V 中一点 x 的切空间 T_xV 到 M 中的点 $\Phi(x)$ 处的切空间 $T_{\Phi(x)}M$ 的非退化线性映射，特别的可对应到 $d \times d$ 的非退化矩阵 $A(x)$。注意到*

$$|det(A)|^2 = det(d\Phi^T d\Phi).$$

这也可以解释上述子流形上曲面测度定义的合理性。

例子. *1 维的子流形即是我们熟知的曲线，考虑曲线的参数化*

$$\gamma : (-1, 1) \to \mathbb{R}^n,$$

其中对任意的 $t\in(-1,1)$，$\gamma'(t)\neq 0$。我们固定 $-1<a<b<1$，对于 γ 上定义的函数 f, 我们可以定义 $\gamma(a)$ 和 $\gamma(b)$ 之间的曲线段 C_a^b 上 f 的积分

$$\int_{C_a^b} f d\sigma = \int_a^b f(\gamma(t))|\gamma'(t)|dt.$$

这就是第一型曲线积分。

我们再看余一维的子流形。

例子. 给定光滑函数

$$f:\mathbb{R}^{n-1}\to\mathbb{R},$$

它的图像所定义的 $\mathbb{R}^n$ 中的超曲面 $\Gamma_f=\left\{(x,f(x))\middle|x\in\mathbb{R}^{n-1}\right\}$，显然 Γ_f 有如下的参数化：

$$\Phi:\mathbb{R}^{n-1}\to\mathbb{R}^n,\quad x\mapsto(x,f(x)).$$

所以我们有

$$d\Phi^T d\Phi = \left(I,(\nabla f)^T\right)\begin{pmatrix} I \\ \nabla f\end{pmatrix} = I+(\nabla f)^T\nabla f$$

$1+|\nabla f|^2$ 为其特征值，对应的特征向量为 ∇f，并且与 ∇f 垂直的向量也为特征向量，对应特征值 1。这样其行列式为

$$det(d\Phi^T d\Phi)=1+|\nabla f|^2.$$

因而我们所求的曲面测度为

$$d\sigma=\sqrt{1+|\nabla f|^2}dx.$$

那么，对于任意的 $\mathbb{R}^n$ 上定义的函数 φ，我们有

$$\int_{\Gamma_f}\varphi d\sigma=\int_{\mathbb{R}^{n-1}}\varphi\big(x,f(x)\big)\sqrt{1+|\nabla f(x)|^2}dx.$$

我们回忆 $\mathbb{R}^n$ 中半径为 1 的球的体积 c_n 为

$$c_n=\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n+2}{2})}.$$

半径为 1 的球 $\mathbb{S}^{n-1}$ 的表面积 ω_{n-1} 为 nc_n，这可以通过对半径为 r 的球的体积对 r 求导数得到。对 $\mathbb{R}^n$ 中的积分，我们还可以将其写成球坐标系下的积分，面积微元转换关系为

$$dx=r^{n-1}drd\omega$$

这里 $d\omega$ 表示 $\mathbb{R}^n$ 中单位球的面积微元。现在我们取 $\mathbb{R}^n$ 中的球坐标系来严格说明这一点

$$x_1=r\sin\theta_{n-1}\sin\theta_{n-2}\ldots\sin\theta_1,\quad x_i=r\sin\theta_{n-1}\sin\theta_{n-2}\ldots\sin\theta_i\cos\theta_{i-1},\quad 2\le i\le n,$$

这里 $r>0,\theta_1,\theta_2,\dots\theta_{n-2}\in(0,\pi),\theta_{n-1}\in(0,2\pi)$。记映射 $F_n(r,\theta_1,\theta_2,\dots,\theta_{n-1})=(x_1,x_2,\dots,x_n)$。根据积分换元公式，我们需要求 Jacobi 矩阵 dF_n 的行列式。根据定义有

$$dF_n=\begin{pmatrix} \sin\theta_{n-1}dF_{n-1}(r,\theta_1,\dots,\theta_{n-2}) & \cos\theta_{n-1}F_{n-1}^T \\ \cos\theta_{n-1}\quad 0 & -r\sin\theta_{n-1} \end{pmatrix}.$$

根据定义，注意到 dF_{n-1} 的第一列为 $r^{-1}F_{n-1}^T$。因而根据行列式的定义有

$$\begin{aligned}\det(dF_n)&=-r\sin\theta_{n-1}\sin^{n-1}\theta_{n-1}\det(dF_{n-1})+(-1)^{n+1}\cos^2\theta_{n-1}\sin^{n-2}\theta_{n-1}r(-1)^{n-2}\det(dF_{n-1})\\ &=-r\sin^{n-2}\theta_{n-1}\det(dF_{n-1})\end{aligned}$$

这样根据归纳法可得

$$dx=r^{n-1}\sin^{n-2}\theta_{n-1}\sin^{n-3}\theta_{n-2}\dots\sin\theta_2drd\theta_1d\theta_2\dots d\theta_{n-1}.$$

$\mathbb{R}^n$ 中的单位球面可由映射 F_n 在球面 $r=1$ 上的限制，记为 F_n^S。注意到 dF_n 可以写成

$$dF_n=(r^{-1}F_n^T,dF_n^S)$$

这里 dF_n^S 是 $n\times(n-1)$ 的矩阵，根据子流形测度的定义，我们要求 $(dF_n^S)^TdF_n^S$ 的行列式。注意到

$$dF_n^TdF_n=\begin{pmatrix} r^{-2}F_n\cdot F_n & r^{-1}F_n(dF_n^S)^T \\ r^{-1}(dF_n^S)^TF_n & (dF_n^S)^TdF_n^S \end{pmatrix}$$

注意到单位球面上有

$$x_1^2+x_2^2+\dots+x_n^2=1,\quad F_n^S\cdot F_n^S=1,\quad F_n^S=r^{-1}F_n$$

微分可得

$$2F_n^SdF_n^S=0,\quad r^{-1}F_ndF_n^S=0,\quad (r^{-1}F_n)\cdot(r^{-1}F_n)=1.$$

特别的根据行列式的定义有

$$\sqrt{\det((dF_n^S)^TdF_n^S)}=\det(dF_n)$$

这样 $\mathbb{R}^n$ 中的单位球面的测度可表示为

$$d\mathbb{S}^{n-1}=\sin^{n-2}\theta_{n-1}\sin^{n-3}\theta_{n-2}\dots\sin\theta_2d\theta_1d\theta_2\dots d\theta_{n-1}$$

在欧式空间 $\mathbb{R}^n$ 中，在球坐标下，体积微元可表示为

$$dx=r^{n-1}drd\mathbb{S}^{n-1}$$

我们还可以计算单位球的面积

$$
\begin{aligned}
&\int_{\theta_1,\ldots,\theta_{n-2}\in(0,\pi)}\int_0^{2\pi}\sin^{n-2}\theta_1\ldots\sin\theta_{n-2}d\theta_1\ldots d\theta_{n-2}d\theta_{n-1}\\
&=2\pi 2^{n-2}\int_0^{\frac{\pi}{2}}\sin\theta d\theta\cdot\ldots\int_0^{\frac{\pi}{2}}\sin^{n-2}\theta d\theta\\
&=2\pi B(\frac{n-1}{2},\frac{1}{2})B(\frac{n-2}{2},\frac{1}{2})\ldots B(\frac{3}{2},\frac{1}{2})\\
&=\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}=nc_n.
\end{aligned}
$$

这与之前所求的单位球面面积一致。

最后我们来看一个经典例题。

例子 (Archimedes). 考虑 $\mathbf{S}^2\subset\mathbb{R}^3$ 是标准的单位球面，对于 $a,b\in[-1,1]$，我们令

$$\mathbf{S}^2(a,b)=\big\{(x,y,z)\in\mathbf{S}^2\big|a\leqslant z\leqslant b\big\}.$$

我们令 $\mathbf{C}$ 为与 z-轴平行的圆柱面，并且这个圆柱面的直径是 2（恰好可以套在 $\mathbb{S}^2$ 上）。对于 $a,b\in[-1,1]$，我们令

$$\mathbf{C}(a,b)=\big\{(x,y,z)\in\mathbb{R}^3\big|x^2+y^2=1,a\leqslant z\leqslant b\big\}.$$

那么 $\mathbf{S}^2(a,b)$ 的面积和 $\mathbf{C}(a,b)$ 的面积相等。

对于 $\mathbf{S}^2(a,b)$，可用球面坐标

$$\mathbf{S}^2(a,b)=(\sin\theta\sin\varphi,\sin\theta\cos\varphi,\cos\theta),\quad\varphi\in[0,2\pi),\quad\theta\in[\arccos a,\arccos b].$$

根据上面计算的曲面测度可知 $\mathbf{S}^2(a,b)$ 的面积为

$$\sigma(\mathbf{S}^2(a,b))=\int_0^{2\pi}\int_{\arccos a}^{\arccos b}\sin\theta d\theta d\varphi=2\pi(b-a).$$

接下来计算 $\mathbf{C}(a,b)$ 的面积，可用柱面坐标

$$\mathbf{C}(a,b)=(\cos\theta,\sin\theta,z),\quad\theta\in[0,2\pi),\quad z\in[a,b].$$

计算可得曲面测度 $d\sigma=d\theta dz$，因而 $\mathbf{C}(0,b)$ 的面积为

$$\sigma(\mathbf{C}(0,b))=\int_0^{2\pi}\int_a^b d\theta dz=2\pi(b-a).$$

8.1 习题

数学分析二作业 11

本次作业的提交时间和地点为 **5 月 28 日**的习题课上，逾期视作零分。

1. 计算 $\mathbb{R}^3$ 中曲线 C 的长度，其中参数曲线 C 由 $t \mapsto (a\cos(t), b\sin(t), bt)$ 给出，其中 $t \in [0, 2\pi]$；计算 $\int_C \frac{z^2}{x^2+y^2} ds$，其中 ds 为 C 上的子流形测度。

2. 假设 C 是 $\mathbb{R}^3$ 中 $x^2+y^2+z^2=a^2$ 与 $x+y+z=0$ 所截出的曲线，试计算 $\int_C xyds$。

3. C 是参数曲线 $\big(a(t-\sin t), a(1-\cos t)\big)$，其中 $t \in [0, 2\pi]$。计算积分 $\int_C y^2 ds$。

4. 平面上的曲线 C 在极坐标系下的方程为 $r = f(\theta)$，其中 f 是连续可微的，$\alpha\theta\beta$。证明，C 的长度为
$$\int_\alpha^\beta \sqrt{|f(\theta)|^2 + |f'(\theta)|^2} d\theta.$$

5. 试计算 $x^2+y^2+z^2=a^2$ 与 $x^2+y^2=ax$ 相交得到的曲线在球面上围出的面积。

6. 试计算参数曲面 Σ 的面积，其中 Σ 如下定义：
$$[0,2\pi]\times[0,2\pi] \to \mathbb{R}^3,\quad (\vartheta,\varphi) \mapsto \big((b+a\cos\vartheta)\cos\varphi, (b+a\cos\vartheta)\sin\varphi, a\sin\vartheta\big),$$
其中，我们要求 $0<a<b$。

7. C 是参数曲线 $\big(a(t-\sin t), a(1-\cos t)\big)$，其中 $t \in [0, 2\pi]$。将 C 在 $\mathbb{R}^3$ 中绕着 x 轴旋转一周得到的曲面是 Σ，试计算 Σ 的面积。

8. C 是平面上的抛物线段 $x = y^2$，其中 $x \in [0, a]$。将 C 在 $\mathbb{R}^3$ 中绕着 x 轴旋转一周得到的曲面是 Σ，试计算 Σ 的面积。

9. Σ 是锥面 $z^2 = x^2+y^2$ 被柱面 $x^2+y^2=2ax$ 所截下的部分，计算曲面积分 $\int_\Sigma x^2y^2+y^2z^2+z^2x^2 d\sigma$.

10. Σ 是 $\mathbb{R}^3$ 中的柱面 $x^2+y^2=a^2$，其中 $z\in[0,h]$，计算曲面积分 $\int_\Sigma x^4+y^4 d\sigma$.

11. Σ 是 $\mathbb{R}^3$ 中的半球面 $x^2+y^2+z^2=1$，其中 $z \geqslant 0$，计算曲面积分 $\int_\Sigma x+y+z d\sigma$.

12. Σ 是 $\mathbb{R}^3$ 中的球面 $x^2+y^2+z^2=1$，计算曲面积分 $\int_\Sigma x^2 d\sigma$.

13. Σ 是 $\mathbb{R}^3$ 中的球面上小帽子 $x^2+y^2+z^2=1$，$z \geqslant h > 0$，计算曲面积分 $\int_\Sigma \frac{1}{z} d\sigma$.

14. Σ 是 $\mathbb{R}^3$ 中的参数曲面：$(u,v)\mapsto \big(u\cos v, u\sin v, v\big)$，其中 $(u,v)\in[0,a]\times[0,2\pi]$。计算曲面积分 $\displaystyle\int_\Sigma z d\sigma$.

15. Σ 是 $\mathbb{R}^3$ 中的球面 $x^2+y^2+z^2=1$，f 是 $\mathbb{R}$ 上的连续函数。证明，

$$\int_\Sigma f(ax+by+cz)d\sigma = 2\pi\int_{-1}^{1} f((a^2+b^2+c^2)^{\frac{1}{2}}t)dt.$$

16. 设物体的质量由密度函数 ρ 给出，那么在区域 V 内的总质量为 $\displaystyle\int_V \rho(x)dx$，$V$ 的质心坐标 $(\bar{x}_1,\cdots,\bar{x}_n)$ 由下式确定:

$$\bar{x}_i = \frac{\int_V x_i\rho(x)dx}{\int_V \rho(x)dx}.$$

求由抛物面 $x^2+y^2=2az$ 及球面 $x^2+y^2+z^2=3a^2$ 围成的均质（$\rho\equiv 1$）物体的质心坐标。

9 Stokes 公式

我们之前提到过 Stokes 公式是微积分基本定理在高维的推广，与一维核心不同点在于高维的区域比较复杂。这里我们只做最基本的讨论，更一般的情形我们会做个大概的讲解。为了定义边界上的积分，我们需要边界具有一定的正则性。基于之前的讨论，我们希望边界是余一维的子流形。

定义 (有界带边区域). 我们称 $\mathbb{R}^n$ 中的紧集 Ω 是**有界带边区域**，如果对任意的边界点 $x \in \partial\Omega$，存在开集 $U, V \subset \mathbb{R}^n$，$x \in U$，以及连续同胚 $\Phi : U \to V$，使得

$$\Phi(U \cap \Omega) = V \cap \{x_n \geqslant 0\}, \quad \Phi(U \cap \partial\Omega) = V \cap \{x_n = 0\}$$

这里 x_n 是 $\mathbb{R}^n$ 中坐标架的第 n 个分量，这里注意到 Ω 可以分为内点集 $\mathring{\Omega}$ 和边界点集 $\partial\Omega$ 的无交并。如果同胚 Φ 是连续可微的，那么称 x 为边界上的正则点，否则称为奇异点。

显然 $\partial\Omega$ 上的正则点构成余维数为 1 的光滑 (C^1 的) 子流形，因而可以选取一个单位向量与在该点处的切空间垂直并指向区域的外部，这样的向量称为单位外法向量。

引理 9.1. 假设 Ω 是一个有界带边区域，对于任意的正则点 $p \in \partial\Omega$，存在 $\mathbb{R}^n$ 中唯一的单位向量 $\nu(p)$ 以及 $\varepsilon > 0$，使得

$$\nu(p) \perp T_p\partial\Omega, \quad |\nu(p)| = 1, \quad p - t\nu(p) \in \Omega, \quad \forall 0 \leq t < \varepsilon.$$

证明: 按照定义，存在开集 $p \in U \subset \mathbb{R}^n$，$V \in \mathbb{R}^n$ 和微分同胚 $\Phi : U \to V$，使得 $p \in U$ 并且 $\Phi(\partial\Omega \cap U) = V \cap \{y_n = 0\}, \Phi(U \cap \Omega) = V \cap \{y_n \geq 0\}$。记 f 为 Φ 的第 n 个分量函数，这样有

$$\Omega \cap U = \big\{x \in U \big| f(x) \geqslant 0\big\}, \quad \partial\Omega \cap U = \big\{x \in U \big| f(x) = 0\big\}.$$

由于 Φ 是微分同胚，特别的 ∇f 非零。根据切空间的定义，

$$T_p\partial\Omega = \{v \in \mathbb{R}^n, \nabla f(p) \cdot v = 0.\}$$

这样与切空间垂直的向量为 $\nabla f(p)$。为了确定单外法向量的符号，根据微分的定义有

$$f(p - tv) = f(p) - t\nabla f(p) \cdot v + o(tv), \quad \forall v \in \mathbb{R}^n$$

为了让 $p - t\nu \in \Omega$ 即是 $f(p - t\nu) > 0$，$\nu(p)$ 只能为

$$\nu(p) = -\frac{\nabla f(p)}{|\nabla f(p)|}.$$

□

我们先来看一个简单的例子。考虑 $\mathbb{R}^n$ 中的方体 $I = [a_1, b_1] \times \ldots \times [a_n, b_n] = I_{n-1} \times [a_n, b_n]$，$\phi$ 是定义在 $\mathbb{R}^n$ 上的连续可微函数。根据 Fubini 定义以及微积分基本定理，我们有

$$\begin{aligned}
\int_I \frac{\partial \phi}{\partial x_n} dx &= \int_{I_{n-1}} \int_{a_n}^{b_n} \partial_{x_n} \phi(x_1, \ldots, x_{n-1}, x_n) dx_n dx_1 dx_2 \ldots dx_{n-1} \\
&= \int_{I_{n-1}} \phi(x_1, \ldots, x_{n-1}, b_n) dx_1 \ldots dx_{n-1} - \int_{I_{n-1}} \phi(x_1, \ldots, x_{n-1}, a_n) dx_1 \ldots dx_{n-1}
\end{aligned}$$

注意到 (I_{n-1}, b_n) 是方体 I 的边界，对应的单位外法向量为 $(0,\ldots,0,1)$，而 (I_{n-1}, b_n) 对应的单位外法向量为 $(0,\ldots,0,-1)$。其他边界对应的单位外法向量第 n 个分量为 0。特别的上面等式可以写成

$$\int_I \frac{\partial\phi}{\partial x_n}dx = \int_{\partial I}\phi(x)\nu_n d\sigma$$

这里 ν_n 是边界上点对应的单位外法向量的第 n 个分量，$d\sigma$ 为边界上的测度。不过这里的问题是边界上并不是所有点都可以定义单位外法向量，比如在方体 I 的边上 (平行与某个坐标轴的边界的边界，是 $n-2$ 维的方体)，原因在于方体的边是连续的但不是 C^1 的。不过由于方体的边相对边界是零测集，因而并不影响上面右边的积分。

一般的 Stokes 公式就是研究上述等式对什么样的区域成立。首先等式右边表明边界几乎是 C^1 的，也就是边界上的奇异点相对边界来说是个零测集，这样才能定义单位外法向量以及边界上的测度。我们在这里不追求一般情形的讨论，但会证明几个相关的重要结论和方法，从我们熟知的满足 Stokes 公式的区域出发得到足够多的满足我们需求的区域。这里我们称 $\mathbb{R}^n$ 中的有界带边区域 Ω 满足 Stokes 公式如果对任意 $\mathbb{R}^n$ 上的连续可微函数 Φ 和 $1\le i\le n$ 都有

$$\int_\Omega \frac{\partial\phi}{\partial x_i}dx = \int_{\partial\Omega}\phi(x)\nu_i d\sigma$$

这里 ν_i 是边界正则点处对应的单位外法向量的第 i 个分量，$d\sigma$ 是边界对应的测度。

首先我们介绍微分几何里重要的局部化工具单位分解定理，这是将常数函数 1 分解成至多可数个有紧支集的光滑函数之和。不失一般性，我们在这里只介绍 $\mathbb{R}^n$ 中从属紧集开覆盖的单位分解定理。

引理 9.2. *假设 K 是 $\mathbb{R}^n$ 中的紧集，U 是包含 K 的开集，那么存在光滑函数 f 使得 f 的支集 $\mathrm{supp}f = \{x\in\mathbb{R}^n, f(x)\neq 0\}$ 在 U 中，并且 $f(x)>0,\ x\in K$。*

证明: 根据条件 $dist(K,U^c)>0$，记

$$\delta = \frac{1}{4}dist(K,U^c),\quad U_1 = \{x\in\mathbb{R}^n, dist(x,K)<\delta\}.$$

可知 $U_1\subset U$ 是开集。取非负光滑截断函数 $\rho(x)$ 使得 $\rho(x)=1$，$|x|\le 1$；$\rho(x)=0$，$|x|\ge 2$。定义

$$f(x) = \frac{dist(x,U_1^c)}{dist(x,K)+dist(x,U_1^c)},\quad g(x) = \int_{|y|\le 2} f(x-\delta y)\rho(y)dy.$$

可知 $g(x)$ 光滑 (写成卷积的形式)。当 $x\in U^c$ 时，注意到

$$dist(x-\delta y,K)\ge dist(x,K)-\delta|y|\ge 4\delta-2\delta>\delta$$

这样根据定义有 $x-\delta y\notin U_1$，即是 $g(x)$ 的支集在 U 中。当 $x\in K$ 时，

$$dist(x-\delta y,U_1^c)\ge dist(K,U_1^c)-dist(x-\delta y,K)\ge dist(K,U_1^c)-\delta|y|>0,\quad \forall|y|<1.$$

这就表明 g 在 K 上为正。 □

现在我们可以构造从属紧集开覆盖的单位分解定理。

引理 9.3 (紧集上单位分解)**.** *假设 K 是 $\mathbb{R}^n$ 上的紧集，$\{U_1, U_2, \ldots, U_k\}$ 是 K 的开覆盖。那么存在光滑函数 $f_1(x)$, …, $f_k(x)$ 使得*

$$1 = f_1(x) + \ldots f_k(x), \quad \mathrm{supp} f_i(x) \subset U_i, \quad 1 \le i \le k, \quad x \in K.$$

证明: 对 $1 \le i \le k$，$\delta > 0$，定义开集 $K_i^\delta = \{x \in U_i, dist(x, U_i^c)\} > \delta$。注意到 $\overset{k}{\underset{i=1}{\cup}} K_i^{\frac{1}{m}}$ 也构成 K 的开覆盖，因而存在正整数 N 使得 $K \subset \overset{k}{\underset{i=1}{\cup}} K_i^{\frac{1}{N}}$。根据上面的引理存在光滑函数 g_i 使得 g_i 的支集在 U_i 中，并且 g_i 在 $K_i^{\frac{1}{N}}$ 的闭包上为正。类似于上述引理可以构造非负光滑函数 $g(x)$ 使得 $g(x)$ 在 K 上为 0，在 $\overset{k}{\underset{i=1}{\cup}} K_i^{\frac{1}{N}}$ 的闭包之外为正。考虑和函数

$$G(x) = g_1(x) + \ldots g_k(x) + g(x)$$

可知 $G(x)$ 在 $\mathbb{R}^n$ 是正的光滑函数，最后取 $f_i(x) = G(x)^{-1} g_i(x)$ 即可，此时注意到 $g(x)$ 在 K 上为 0。 □

单位分解定理告诉我们为了证明 Stokes 公式，只需要对支集在任意小的方体上的函数 ϕ 成立即可。接下来我们证明如果 Φ 是从方体 I 到 $\Phi(I)$ 的连续可微同胚，那么有界区域 $\Phi(I)$ 也满足 Stokes 公式。

命题 9.4. *假设 Φ 是从方体 I 到 Ω 的连续可微的微分同胚，那么 Ω 也满足 Stokes 公式。*

证明: 由于 Φ 是从 I 到 $\Omega = \Phi(I)$ 的连续可微的微分同胚，特别的 $\partial\Omega = \Phi(\partial I)$。利用积分换元公式有

$$\int_{\Phi(I)} \nabla_x \varphi dx = \int_I \nabla_y \varphi(\Phi(y))(d\Phi)^{-1} |J_\Phi| dy$$

这里我们用 y 表示 I 上的坐标，$x = \Phi(y)$ 为 Ω 上的坐标。此外我们不妨假设 $J_\Phi > 0$，记矩阵 $d\Phi = A$, 这样有

$$A_{kj} A^{ji} = A^{kj} A_{ji} = \delta_{ki}, \quad A^*_{ij} = |A| A^{ji}$$

这里同一个上下标表示对指标从 1 到 n 求和，$|A|$ 表示 A 的行列式，A^*_{ij} 表示元素 A_{ij} 对应的代数余子式的行列式，$\delta_{ii} = 1$，$\delta_{ij} = 0$ 当 $i \ne j$ 时。根据行列式的定义有

$$\partial_{y_j} |A| = \partial_{y_j} A_{kl} A^*_{kl} = \partial_{y_j} A_{kl} A^{lk} |A|.$$

我们又注意到

$$A_{ij} = \frac{\partial \Phi_i(y)}{\partial y_j}$$

根据导数的可交换性有 $\partial_{y_k}A_{ij}=\partial_{y_j}A_{ik}$。这样计算可得

$$\begin{aligned}\partial_{y_j}(A^{ji}|A|)&=\partial_{y_j}A^{ji}|A|+A^{ji}\partial_{y_j}A_{kl}A^{lk}|A|\\&=\partial_{y_j}A^{ji}|A|+A^{ji}\partial_{y_l}A_{kj}A^{lk}|A|\\&=\partial_{y_j}A^{ji}|A|-A_{kj}\partial_{y_l}A^{ji}A^{lk}|A|=0\end{aligned}$$

这就表明

$$\begin{aligned}\int_{\Phi(I)}\nabla_{x_i}\varphi dx&=\int_I\nabla_{y_j}\varphi(\Phi(y))(d\Phi)^{ji}|J_\Phi|dy\\&=\int_I\nabla_{y_j}\Big(\varphi(\Phi(y))(d\Phi)^{ji}J_\Phi\Big)dy\\&=\int_{\partial I}\varphi(\Phi(y))(d\Phi)^{ji}J_\Phi\nu_jd\sigma\end{aligned}$$

这里 ν 是 ∂I 的单位外法向量，$d\sigma$ 是 ∂I 的曲面测度。由于 I 是方体，边界 ∂I 是由 $2n$ 个 $\mathbb{R}^{n-1}$ 中的方体组成。由对称性，考虑 I 的一部分边界 $(\partial I)_n=\{(z,y_n)|z\in U\subset\mathbb{R}^{n-1}\}$，并假设在这部分边界上单位外法向量为 $(0,1)$，即是当 t 足够小时 $(z,y_n-t)\in I$。

注意到 Φ 将 I 映到 Ω，根据上面引理，$\Phi((\partial I)_n)$ 处的单位外法向量为

$$\tilde{\nu}=\frac{\nabla\Phi_n^{-1}}{|\nabla\Phi_n^{-1}|}=\frac{A^{n\cdot}}{|A^{n\cdot}|}=J_\Phi\frac{(A^{n1},A^{n2},\ldots,A^{nn})}{|(A_{1n}^*,A_{2n}^*,\ldots,A_{nn}^*)|}$$

这里我们仍然记 $A=d\Phi$, $\cdot$ 取下标 1 到 n，从而构成一个向量。再根据曲面测度的定义可知，$\Phi((\partial I)_n)$ 上的曲面测度为

$$d\sigma=\sqrt{d_z\Phi(z,y_n)^Td_z\Phi(z,y_n)}dz$$

我们需要用到一个代数的等式：给定 $\mathbb{R}^n$ 中的线性无关向量 $v_1,v_2,\ldots,v_{n-1}$，张成 $n-1$ 维平行多面体 L。考虑单位向量 e_i（第 i 个分量为 1，其余为 0），由于 e_i 代表的 n 个点到平行多面体 L 的距离平方和为 1，特别的有 e_i 与 $v_1,v_2,\ldots,v_{n-1}$ 张成的 n 维平行多面体的体积平方和等于 L 的体积的平方，即是为 $det(v_i\cdot v_j)_{1\le i,j\le n-1}$。另一方面 $v_1,v_2,\ldots,v_{n-1},e_i$ 张成的平行多面体的体积为可看成是 $v_1,v_2,\ldots,v_{n-1}$ 构成的 $(n-1)\times n$ 矩阵去掉第 i 列的余子式的行列式，即是

$$d_z\Phi(z,y_n)^Td_z\Phi(z,y_n)=\sum_{i=1}^n|A_{in}^*|^2$$

从而有

$$\int_{(\partial I)_n}\varphi(\Phi(y))(d\Phi)^{ji}J_\Phi\nu_jdz=\int_{(\partial I)_n}\varphi(\Phi(y))(d\Phi)^{ni}J_\Phi dz=\int_{\Phi(\partial I)_n}\varphi(x)\tilde{\nu}_id\sigma$$

□

注记. 上面的证明表明如果 Φ 是从有界区域 Ω_1 到 Ω_2 的连续微分同胚，并且 Ω_1 满足 *Stokes* 公式，那么 Ω_2 也满足 *Stokes* 公式。这是因为 Ω_1 满足 *Stokes* 公式表明边界的奇异点相对于边界来说是零测集，不影响 *Stokes* 公式右边的积分。而边界上的正则点集构成开集，因而利用单位分解定理，只需要验证正则点的某个开领域内边界上积分的相等关系。根据定义，正则点邻域与带边上半平面微分同胚，即归化为上面讨论情形。

最后我们对边界点全是正则点的有界带边区域 Ω 证明其满足 Stokes 公式。根据单位分解定理，对任意 Ω 的有限开覆盖 $U_1, U_2, \ldots, U_N$，存在支集在 U_i 上的光滑函数 χ_i 使得在 Ω 上有

$$1 = \chi_1 + \chi_2 + \ldots \chi_N.$$

这样只需要对连续可微函数 $\varphi\chi_j$ 证明其满足 Stokes 公式，换句话说，利用单位分解定理，只需对支集在任意开集 U_j 中的连续可微函数 φ 来验证 Stokes 公式。

注意到我们可以用任意小的有限个方块来覆盖 Ω，取方块 I，对支集在 I 上的函数 φ。如果 $I \subset \Omega$，由于方块满足 Stokes 公式，因而有

$$\int_\Omega \frac{\partial \varphi}{\partial x_n} dx = \int_I \frac{\partial \varphi}{\partial x_n} dx = \int_{\partial I} \varphi \nu_n(I) d\sigma(\partial I) = 0 = \int_{\partial \Omega} \varphi \nu_n(\Omega) d\sigma(\partial \Omega)$$

这里我们不妨取 $i = n$。接下来只需要对与边界 $\partial\Omega$ 相交的方块 I 以及支集在 I 上的连续可微函数 φ 验证 Stokes 公式即可。首先我们有

$$\int_\Omega \frac{\partial \varphi}{\partial x_n} dx = \int_{\Omega\cap I} \frac{\partial \varphi}{\partial x_n} dx, \quad \int_{\partial\Omega} \varphi \nu_n d\sigma = \int_{\partial\Omega\cap I} \varphi \nu_n d\sigma = \int_{\partial\Omega\cap I} \varphi \nu_n d\sigma = \int_{\partial(\Omega\cap I)} \varphi \nu_n d\sigma$$

这样我们只需要证明区域 $\Omega \cap I$ 满足 Stokes 公式即可。根据正则点的定义，取方块 I 足够小，使得 $\partial\Omega \cap I$ 局部上可以表示为图，即是存在 $1 \leq m \leq n, \mathbb{R}^{n-1}$ 上的方块 I_{n-1} 以及可微函数 $\rho(z)$，$z = (y_1, \ldots, y_{m-1}, y_{m+1}, \ldots, y_n) \in I_{n-1}$ 使得

$$\Omega \cap I = \{(y_1, y_2, \ldots, y_n) | z \in I_{n-1}, \quad c_m \leq y_m \leq \rho(z)\}$$

这里 c_m 为常数，我们还要求在 I_{n-1} 上 $c_m < \rho(z)$。构造微分同胚

$$(y_1, y_2, \ldots, y_n) \mapsto (y_1, \ldots, y_{m-1}, \frac{y_m - c_m}{\rho(z) - c_m}, y_{m+1}, \ldots, y_n)$$

特别的这是有界带边区域 $\Omega \cap I$ 到方块 $I_{n-1} \times [0, 1]$ 的微分同胚，根据上面的结论有界带边区域 $\Omega \cap I$ 满足 Stokes 公式。

9.1 习题

数学分析二作业 12

本次作业的提交时间和地点为 **6 月 4** 日的习题课上，逾期视作零分。

1. （Co-Area 公式）$\Omega \subset \mathbb{R}^n$ 是开集，$F: U \to \mathbb{R}$ 是光滑映射。假设对任意的 $x \in U$，$dF(x) \neq 0$（从而，$\Sigma_t = F^{-1}(t)$ 是余 1 维子流形）。那么，对任意的 U 上的可积函数 f，我们有
$$\int_\Omega f(x)dx = \int_{\mathbb{R}} \left(\int_{\Sigma_t} \frac{f}{|\nabla F|} d\sigma_t \right) dt,$$
其中，$d\sigma_t$ 是 Σ_t 上的子流形测度。

2. 给定 $\mathbb{R}^n$ 上的光滑函数 $f: \mathbb{R}^n \to \mathbb{R}$，考虑 $\mathbb{R}^{n+1}$ 中的区域
$$\Omega = \big\{(x, x_n) \big| x \in \mathbb{R}^n, x_n \in \mathbb{R}, x_n \leqslant f(x)\big\}.$$
证明边界上的单位外法向量为
$$(-\frac{\nabla f}{\sqrt{1+|\nabla f|^2}}, \frac{1}{\sqrt{1+|\nabla f|^2}}).$$

3. 假设 $K \subset U_1 \subset U_2$ 为 $\mathbb{R}^n$ 中的有界集合，K 为紧集，U_1, U_2 为开集。证明存在非负光滑函数 f 使得 f 在 K 上为 0，在 $U_2 \backslash U_1$ 上为正。

4. 假设 ρ 是 $\mathbb{R}^3$ 上非负可积的函数（如果一个物体的质量函数是 ρ 的话），它所对应的引力势能 $\Phi(x)$ 和引力场 $G(x)$ 由下式给出:
$$\Phi(x) = -\int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|} dy, \quad G(x) = -\nabla\Phi(x)$$
我们假设区域 $V \subset \mathbb{R}^3$ 是开区域，$\rho_0 > 0$ 是常数，
$$\rho(x) = \begin{cases} \rho_0, & x \in V \\ 0, & x \in \mathbb{R}^3 \setminus V \end{cases}$$
试计算如下情形下的 $\Phi(x)$ 和 $G(x)$：
   1) $V = \{x \in \mathbb{R}^3 : |x|^2 < R^2\}$ 是球体。
   2) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 < R^2, 0 < x_3 < h\}$ 是圆柱体。
   3) $V = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x_3^2 > \frac{h^2}{R^2}(x_1^2 + x_2^2), 0 < x_3 < h\}$ 是圆锥体。

5. (Newton 壳层 (shell) 定理：关于球对称密度分布的引力）假设
$$\rho(x) = \begin{cases} f(|x|), & r < |x| < R; \\ 0, & |x| > R \text{ 或 } |x| < r. \end{cases}$$
其中 f 是在 $[r, R])$ 上定义的非负可积函数。证明:

1) 对球外一点 ($|x| > R$) 产生的引力与球的所有质量集中于球心的质点产生的引力一样。

2) 球壳对球壳内任何一点 ($|x| < r$) 产生的引力 (合力) 为零。

10 第二型曲线曲面积分

前一节我们对有界带边光滑区域证明了 Stokes 公式，假设 $\varphi_1, \ldots, \varphi_n$ 是 $\mathbb{R}^n$ 上的连续可微函数，那么根据 Stokes 公式有

$$\int_\Omega \sum_{i=1}^n \frac{\partial \varphi_i}{\partial x_i} dx = \int_{\partial\Omega} \sum_{i=1}^n \varphi_i \nu_i d\sigma.$$

记 $X = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ 为向量值函数，可以看成是 $\mathbb{R}^n$（Ω）上的向量场 (几何量，与坐标系无关)。记 $\mathrm{div}(X)$ 为向量场 X 的散度，定义为

$$\mathrm{div}(X) = \sum_{i=1}^n \frac{\partial X_i}{\partial x_i}.$$

这是定义在 $\mathbb{R}^n$ 上的函数。二阶微分算子 Laplace 算子 Δ 可表示为 $\mathrm{div}\nabla$，即是

$$\Delta f = \mathrm{div}\nabla f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = \mathrm{tr}(H_f).$$

Stokes 公式的等价形式是如下的散度定理。

定理 10.1 (散度定理)**.** *给定有界光滑带边区域 $\Omega \subset \mathbb{R}^n$，我们用 ν 表示 $\partial\Omega$ 的单位外法向量，$d\sigma$ 为 $\partial\Omega$ 的子流形测度。那么，对任意 $\mathbb{R}^n$ 上 C^1 的向量场 X，我们有*

$$\int_\Omega \mathrm{div}X \; dx = \int_{\partial\Omega} X \cdot \nu d\sigma.$$

这里在固定点 x 处，我们将 X, ν 都看成是 $\mathbb{R}^n$ 中的向量。

当 $n = 2$ 时，散度定理或者 Stokes 公式被称作是 Green 公式。

推论 10.2 (Green 公式)**.** *如果 $\Omega \subset \mathbb{R}^2$ 是有界带边区域满足 Stokes 公式，那么对任意的 $\mathbb{R}^2$ 上的连续可微函数 P 和 Q 有*

$$\int_\Omega (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}) dxdy = \int_{\partial\Omega} (Q, -P) \cdot \nu d\sigma.$$

注记. *特别的如果取 $Q = x$ 或者 $P = -y$，那么左边就给出了区域 Ω 的面积，而右边则转化成了边界 $\partial\Omega$ 上的曲线积分。*

由于 ν 是区域 Ω 的单位外法向量，因而右边的曲线积分蕴含了符号（定向）的选取，这个符号是由所围成的区域 Ω 来决定的。更一般的散度定理是左边区域 Ω 可以是高维空间的可定向曲面，选取的定向决定了右边边界上积分的定向。现在我们来看右边曲线积分的另一种表达方式。注意到边界是连续的曲线，选取参数 $\gamma(t) = (x(t), y(t))$ 使得围成的区域在曲线的左边，这样有

$$\nu(\gamma(t)) d\sigma = \frac{(y'(t), -x'(t))}{|\gamma'(t)|} |\gamma'(t)| dt = (dy, -dx).$$

此时右边曲线积分可以写成等价形式

$$\int_{\partial\Omega}(Q,-P)\cdot\nu d\sigma=\int_{\partial\Omega}Pdx+Qdy.$$

对任意连续可微曲线 $\gamma(t)$ 以及向量场 X，我们将曲线积分

$$\int_{\gamma}X\cdot\nu d\sigma=\int_{\gamma}\sum_{i=1}^{n}X_i dx_i$$

称为第二型曲线积分。如果曲线 γ 是闭曲线，那么积分记为 $\oint_\gamma X\cdot d\gamma$，其方向 (定向) 的选取满足曲线所围成的区域在曲线的左边。

当 $n=3$ 时，散度定理或者 Stokes 公式被称作是 Gauss-Ostrogradsky 公式。

推论 10.3 (Gauss-Ostrogradsky). *如果 $\Omega\subset\mathbb{R}^3$ 是有界带边区域满足 Stokes 公式，那么对任意的 $\mathbb{R}^3$ 上的连续可微函数 P，Q 和 R，我们有*

$$\int_{\Omega}\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}dxdydz=\int_{\partial\Omega}(P,Q,R)\cdot\nu d\sigma.$$

类似的右边曲面上的积分定向的选取取决于所围成的区域 Ω 方向的选取 ($\mathbb{R}^n$ 上的方向取定，比如一维右边是正向)。任取 $\partial\Omega$ 的参数化 $\gamma(u,v)=(x(u,v),y(u,v),z(u,v))$，两个切向方向为 γ_u 以及 γ_v。根据曲面测度的定义以及单位外法向量的定义有

$$\begin{aligned}\nu d\sigma&=\frac{\gamma_u\times\gamma_v}{|\gamma_u\times\gamma_v|}|\gamma_u\times\gamma_v|dudv\\&=(y_uz_v-z_uy_v,z_ux_v-x_uz_v,x_uy_v-y_ux_v)dudv\\&=(dydz,dzdx,dxdy)\end{aligned}$$

这里我们用到坐标变换公式 $dydz=(y_uz_v-z_uy_v)dudv$。这样

$$\int_{\partial\Omega}(P,Q,R)\cdot\nu d\sigma=\oint_{\partial\Omega}Pdydz+Qdzdx+Rdxdy$$

这样的积分我们称为曲面上的第二型曲面积分 (不一定是闭曲面)，带圆圈表示曲面是闭曲面，因而需要选定曲面的正向。在局部坐标系下 $\gamma_u\times\gamma_v$ 指向区域的外部。

10.1 微分形式

为了更好的运用上面介绍的 Stokes 公式，这节我们对微分形式作简要的介绍。假设 $\Omega\subset\mathbb{R}^n$ 为开区域，可取坐标系 $x=(x_1,x_2,\ldots,x_n)$。任何一点处的切空间可由 $\{\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}\}$ 张成，即是 $\frac{\partial}{\partial x_i}$ 可看成是沿着坐标函数 x_i 的切向量。这样 Ω 上的光滑向量场可以表示为

$$X=\sum_{i=1}^{n}X^i(x)\frac{\partial}{\partial x_i}$$

其中 $X^i(x)$ 为 Ω 上的光滑函数。假定 $dx_i(\frac{\partial}{\partial x_j}) = \delta_{ij}$(Kronecker 符号)，那么每点的切空间的对偶空间就可以由 dx_i 来张成。余切丛 (切空间换成对偶空间) 的一个截面就是 Ω 上的 $1-$ 阶**微分形式**，在坐标下可以表示为

$$\omega = \sum_{i=1}^{n} \omega^i(x) dx_i, \quad \omega^i(x) \in C^\infty(\Omega).$$

特别的第二型曲线积分可以看成是 $1-$ 阶微分形式的积分。为了类似得理解第二型曲面积分，我们需要定义高阶的微分形式。对于 $1 \le k \le n$，$k-$ 阶微分形式可由

$$dx_{i_1} \wedge dx_{i_2} \wedge \ldots \wedge dx_{i_k}, \quad 1 \le i_1 < i_2 < \ldots < i_k \le n$$

生成。为了叙述方便我们也允许 dx_{i_p}, dx_{i_q} 互换位置，但满足反对称性，即对任意 $i_1 \le p < q \le i_k$ 有

$$dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge \ldots \wedge dx_{i_q} \wedge \ldots \wedge dx_{i_k} = (-1)^{q-p} dx_{i_1} \wedge \ldots \wedge dx_{i_q} \wedge \ldots \wedge dx_{i_p} \wedge \ldots \wedge dx_{i_k}.$$

反对称性表明 $dx_i \wedge dx_i = 0$。这样一般的 $k-$ 阶微分形式在局部坐标系下可以表示为

$$\omega = \sum_{1 \le i_1, \ldots, i_k \le n} a^{i_1 i_2 \ldots i_k}(x) dx_{i_1} \wedge dx_{i_2} \wedge \ldots \wedge dx_{i_k}, \quad a^{i_1 i_2 \ldots i_k}(x) \in C^\infty(\Omega).$$

全体记作 $\Lambda^k(\Omega)$，显然同阶微分形式可以有线性关系

$$f\omega_1 + g\omega_2 \in \Lambda^k(\Omega), \quad \forall f, g \in C^\infty(\Omega), \quad \forall \omega_1, \omega_2 \in \Lambda^k(\Omega).$$

任意两个微分形式可以做乘法，即是 $\omega_1 \in \Lambda^{k_1}(\Omega), \omega_2 \in \Lambda^{k_2}(\Omega)$，我们可以定义 $\omega_1 \wedge \omega_2 \in \Lambda^{k_1+k_2}(\Omega)$ (当 $k > n$ 时，$\Lambda^k(\Omega)$ 由 0 函数构成)。乘法定义满足对 ω_1, ω_2 线性，因而只需要定义各自基的乘法

$$(dx_{i_1} \wedge \ldots \wedge dx_{i_{k_1}}) \wedge (dx_{j_1} \wedge \ldots \wedge dx_{j_{k_2}}) = dx_{i_1} \wedge \ldots \wedge dx_{i_{k_1}}) \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_{k_2}}$$

根据反对称性，可以交换顺序使得下标严格单调增 (重复下标即为 0)。

对于微分形式，还有一个重要的算子**外微分**$d : \Lambda^k(\Omega) \to \Lambda^{k+1}(\Omega)$。首先我们要求外微分满足线性性质，即是

$$d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2, \quad \forall \omega_1, \omega_2 \in \Lambda^k(\Omega).$$

这样我们只需要定义

$$d(f dx_{i_1} \wedge \ldots \wedge dx_{i_k}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_k}, \quad \forall f \in C^\infty(\Omega).$$

微分形式的优点在于微分形式的不变性。对于坐标变换 $(x_1, \ldots, x_n) \to (y_1, \ldots, y_n)$, 我们知道测度变换公式

$$dy = |J| dx, \quad J = det(\frac{\partial y_i}{\partial x_j})_{1 \le i,j \le n}$$

一个重要的事实是 (这里不加证明，同学们可以自行验证)

$$dy_1 \wedge dy_2 \wedge \ldots \wedge dy_n = J dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n.$$

这里将 y_i 看成是 x 的函数，dy_i 是 y_i 的外微分，即是一阶微分形式。换句话说我们可以用 $n-$ 阶微分形式 $dx_1 \wedge \ldots \wedge dx_n$ 表示 Ω 上的体积微元 $d\text{vol} = dx$(可能相差一个符号)。

定理 10.4. 对任意 $\omega \in \Lambda^k(\Omega)$，有 $d^2\omega = d(d\omega) = 0$。

证明: 证明基于微分形式的反对称性以及 Ω 上导数的可交换性。用 f_i 表示 f 在 x_i 的偏导数，根据定义有

$$d^2(f dx_{i_1} \wedge \ldots dx_{i_k}) = f_{ji} dx_j \wedge dx_i \wedge dx_{i_1} \wedge \ldots dx_{i_k} = 0$$

□

由此我们称满足 $d\omega = 0$ 的微分形式 ω 为**闭的微分形式**，称形如 $\omega = d\omega_1$ 的微分形式 ω 为**恰当的微分形式**。上述定理表明 Ω 上恰当的微分形式一定是闭的。

定理 10.5 (Poincare)**.** $\mathbb{R}^n$ 上闭的微分形式一定是恰当的。

证明: 我们简要给出证明。首先我们看 $n-$ 阶微分形式 $\omega = f dx_1 \wedge \ldots \wedge dx_n$，定义函数

$$g(x_1, x_2, \ldots, x_n) = \int_0^{x_1} f(t, x_2, \ldots, x_n) dt + f(0, x_2, \ldots, x_n)$$

根据定义有

$$d(g dx_2 \wedge \ldots dx_n) = \frac{\partial g}{\partial x_1} dx_1 \wedge \ldots \wedge dx_n = \omega.$$

对于一般的情形将 ω 写成

$$\omega = \omega_1 + dx_1 \wedge \omega_2, \quad \omega_2 \in \Lambda^{k-1}(\Omega).$$

这里 ω_1, ω_2 里都不含 dx_1。同上面的取法，可以构造 ω_3 使得

$$d\omega_3 = dx_1 \wedge \omega_2 - \omega_4 + \omega_1$$

满足 ω_3, ω_4 里不包含 dx_1。特别的有

$$\omega - d\omega_3 = \omega_4.$$

根据假设 ω_4 是不包含 dx_1 的闭的微分形式，特别的系数与 x_1 无关。这就归结为了 $n-1$ 维的情形 (dx_1 不出现，系数函数与 x_1 无关)。一直这样操作下去可知 ω 是恰当的微分形式。 □

注记. 这个结论对更一般的区域 Ω 也成立，比如 Ω 单连通的星型区域。这里单连通指的是任何闭曲线可以在 Ω 内连续的缩成一个点；星型区域表示存在一个点 $P \in \Omega$ 使得任意 $Q \in \Omega$，有 $tP + (1-t)Q \in \Omega$, $t \in [0, 1]$。

一个特列是单连通区域 $\Omega \subset \mathbb{R}^n$ 上的 1− 阶微分形式 $\omega = \sum_{i=1}^{n} \omega^i(x)dx_i$，那么 ω 存在原函数 $f \in C^\infty(\Omega)$，即是 $\omega = df$，等价于 $d\omega = 0$。显然如果 ω 存在原函数，那么 ω 是恰当的，自然是闭的；反过来，如果 ω 是闭的，为了说明 ω 存在原函数，只需要说明 ω 沿曲线积分与路径无关，即是曲线积分 $\int_\gamma \omega$ 只与起始点 $\gamma(0), \gamma(1)$ 相关，这样就可以定义函数

$$f(x) = \int_\gamma \omega, \quad \gamma(0) = P, \quad \gamma(1) = x.$$

这里 $P \in \Omega$ 是固定点。注意到曲线与积分路径无关等价于对任意闭曲线 γ 有 $\oint_\gamma \omega = 0$（此结论与区域的拓扑性质无关，任何连通区域即可）。根据上述 Poincare 定理的证明可知局部上 ω 存在原函数，也即是局部上对任意闭曲线 γ 都有 $\oint_\gamma \omega = 0$。由于区域是单连通的，因而任意闭曲线 γ 可以连续的形变成足够小的闭曲线，这样就可以得到 ω 在任意闭曲线上的积分为 0。

用微分形式的语言，Stokes 公式还可以推广到 $\mathbb{R}^n$ 中的 $k+1$ 维的带边子流形 D 上。

定理 10.6 (Stokes 公式). *假设 D 是 $\mathbb{R}^n$ 中的 $k+1$ 维光滑带边可定向子流形，对任意 $\omega \in \Lambda^k(\mathbb{R}^n)$ 有*

$$\int_D d\omega = \int_{\partial D} \omega.$$

首先我们需要定义 $k-$ 阶微分形式 ω 在 $k-$ 维子流形 D 上的积分，这个过程是标准的。先用单位分解，可以假设 ω 的支集足够小使得在支集附近子流形 D 有坐标 $y \in D_k \subset \mathbb{R}^k$。此时在 D 上坐标函数 x_i 可以是 y 的函数。对任意 $k-$ 阶微分形式 $fdx_{i_1} \wedge \ldots dx_{i_k}$，可展成 D_k 上关于 y 坐标的 $k-$ 阶微分形式。根据上面的讨论，我们可以定义 D_k 上 $k-$ 阶微分形式的积分，此积分即可定义成是子流形 D 上 $k-$ 阶微分形式 $fdx_{i_1} \wedge \ldots dx_{i_k}$ 的积分。但把局部积分拼起来的时候需要用到子流形 D 是可定向的，原因在于 $k-$ 阶微分形式定义的积分符号是不确定的 (交换微分形式中的 dx_i 和 dx_j，积分会变号)。可定向保证了局部定义的积分可以选一个统一的符号。子流形 D 的定向的选取也决定了边界 ∂D（作为 $k-1$ 维子流形）定向的选取。至于子流形以及边界 ∂D 的正则性问题，则需要用到 Hausdorff 测度的概念，简单来说就是 D 上不正则的点相对 D 来说是零测集。这里我们对细节以及详细的证明不作更多讨论。

我们现在回过头来看格林高斯公式。对于格林公式，假设 $D \subset \mathbb{R}^2$ 是带边区域，其边界 ∂D 是闭曲线。这样 Stokes 公式表明

$$\int_{\partial D} \omega = \int_D d\omega, \quad \int_{\partial D} Pdx + Qdy = \int_D \frac{Q}{\partial x} - \frac{\partial P}{\partial y} dxdy.$$

这里边界曲线 ∂D 的正向为逆时针方向 (区域在曲线左边)。

类似的对于高斯公式，假设 $D \subset \mathbb{R}^3$ 是有界带边区域，对于任意 2-形式 $\omega = Pdxdy + Qdydz + Rdzdx$ 有

$$\int_D d\omega = \int_D \frac{\partial P}{\partial z} + \frac{\partial Q}{\partial x} + \frac{\partial R}{\partial y} dxdydz = \oint_{\partial D} \omega$$

最后我们讨论一种情形是 $\mathbb{R}^3$ 中的有界带边二维曲面 D，我们假设曲面 D 是可定向的 (比如 Mobius 带就不可定向)，也就是是可以选取曲面 D 的单位外法向量 ν，并且边界 ∂D 是由有限条闭曲线组成。对任意 $1-$ 形式 $\omega = Pdx + Qdy + Rdz$，我们有

$$\int_{\partial D} \omega = \int_D d\omega = \int_D (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy + (\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})dydz + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})dzdx$$

这里注意到 $2-$ 形式 $d\omega$ 在 D 上的积分即是第二型曲面积分，定向的选取与开始曲面的定向选取一致。左边是第二型曲线积分，方向的选取是右手法则（拇指指向 ν，四指并拢指向的方向满足围成的区域在 D 中）。

10.2 习题

数学分析二作业 13

本次作业的提交时间和地点为 **6 月 11 日**提交，具体提交方式和时间可以和助教商量。

1. 心脏线 $\{(x,y) \in \mathbb{R}^2 : (x^2+y^2)^2 + 4ax(x^2+y^2) - 4a^2y^2 = 0\}$ 围成的区域面积.
2. 双纽线 $(x^2+y^2)^2 = 2a^2(x^2-y^2)$ 围成的面积。
3. 曲线 $(x^3+y^3)^2 = x^2+y^2$ 的内部区域面积。
4. 曲线 $\frac{x^3}{a^3} + \frac{y^3}{b^3} = \frac{x^2}{h^2} + \frac{y^2}{k^2}$ 与坐标轴 $x=0, y=0$ 围成的面积。
5. 曲线 $(2x+y)^4 = x^2 - 2y^2$ 围成的面积。
6. 曲线 $(\frac{x}{a} + \frac{y}{b})^5 = \frac{x^2y^2}{c^4}$ $(x>0, y>0)$ 围成的面积.
7. 曲面 $(x^2+y^2+z^2)^3 = \frac{a^6z^2}{x^2+y^2}$ 所围立体图形的体积。
8. 曲面 $(\frac{x^2}{a^2} + \frac{y^2}{b^2})^2 + \frac{z^4}{c^4} = 1$ 内部的体积。
9. 曲面 $(\frac{x}{a} + \frac{y}{b} + \frac{z}{c})^4 = \frac{xyz}{abc}$ $(x,y,z>0)$ 所围立体图形的体积。
10. $(\frac{x}{a} + \frac{y}{b})^2 + (\frac{z}{c})^2 = 1$ 所围立体图形的体积。
11. 曲面 $z = x^2+y^2, z = 2(x^2+y^2), xy = a^2, xy = 2a^2, x = 2y, 2x = y$ $(x>0, y>0)$ 所围立体图形的体积。
12. 计算下列曲线围成的面积

$$x^{2n+1} + y^{2n+1} = 2x^ny^n, \quad n \text{ 为正整数}.$$

提示：用极坐标。

13. （三维梯度、散度的几何定义）假设 X 是 $\mathbb{R}^3$ 上的光滑向量场，φ 是 $\mathbb{R}^3$ 上的光滑实值函数，$B_\varepsilon(x)$ 为以 $x \in \mathbb{R}^3$ 为球心以 $\varepsilon > 0$ 为半径的球。证明

$$\nabla\varphi(x) = \lim_{\varepsilon\to 0} \frac{1}{m(B_\varepsilon(x))} \int_{\partial B_\varepsilon(x)} \varphi \cdot \nu \, d\sigma, \quad \operatorname{div}(X) = \lim_{\varepsilon\to 0} \frac{1}{m(B_\varepsilon(x))} \int_{\partial B_\varepsilon(x)} X \cdot \nu \, d\sigma.$$

14. 令 Ω 为 $\mathbb{R}^3$ 中的有界光滑带边区域，ν 为 $\partial\Sigma$ 的单位外法向量。对于任意的 $\mathbb{R}^3$ 中的光滑函数 f，证明，

$$\int_\Omega \triangle f \, dxdydz = \int_{\partial\Omega} \nabla_\nu f \, d\sigma.$$

15. 假设一个粒子 P 从 A 点运动到 B 点，它的的运动轨迹（$\mathbb{R}^n$ 中）由参数曲线

$$\gamma : [a,b] \to \mathbb{R}^n$$

给出，其中 $\gamma(a) = A$，$\gamma(b) = B$。粒子受到的力由向量场 $F(\gamma(t))$ 给出，那么我们定义粒子 P 在路径 γ 上做的**功**由第二型曲线积分

$$\int_\gamma F \cdot d\gamma$$

给出。给定平面 $\mathbb{R}^2$ 上的力场

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

试计算 F 分别沿路径 γ_1 和 γ_2 所做的功，其中

$$\gamma_1 : [0,2\pi] \to \mathbb{R}^2 - \{(0,0)\}, \quad t \mapsto (\cos t, \sin t),$$
$$\gamma_2 : [0,2\pi] \to \mathbb{R}^2 - \{(0,0)\}, \quad t \mapsto (2+\cos t, \sin t).$$

16. F 是区域 $U \subset \mathbb{R}^n$ 上的光滑向量场（力场），如果存在函数 $P \in C^\infty(U)$ ，使得 $F = -\nabla P$，我们就称 F 是**保守的**并把 P 称为 F 的**势函数**。

   1) 证明保守力做的功只依赖于路径的起点和终点而与路径的选取无关，即若 $F = -\nabla P$，并且 $\gamma_1, \gamma_2 : [a,b] \to \mathbb{R}^n$，满足 $\gamma_1(a) = \gamma_2(a), \gamma_1(b) = \gamma_2(b)$，那么

   $$\int_{\gamma_1} F \cdot d\gamma_1 = \int_{\gamma_2} F \cdot d\gamma_2.$$

   2) 如果 $F = -\nabla P$ 是保守力，质点的运动满足 Newton 第二定律

   $$m\ddot{x}(t) = F(x(t)),$$

   其中，$\ddot{x}$ 代表两次导数，m 是常数，代表粒子的质量。证明质点运动满足能量守恒，即动能与势能之和为常数:

   $$E(t) = \frac{1}{2}m|\dot{x}|^2 + P(x(t)) = E(t_0).$$

17. 1) 假设空间 $\mathbb{R}^3$ 有 N 个点电荷，它们的位置和电量分别为 $x_i \in \mathbb{R}^3$ 和 $q_i \in \mathbb{R}$, 其中 $i = 1, \cdots, N$。根据 Coulomb 定律，它们在 $x \in \mathbb{R}^3$ 处产生的电场（向量场）为

$$E(x) = \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} \frac{x - x_i}{|x - x_i|^3},$$

其中 $\varepsilon_0 > 0$ 是介电常数。试证明 Gauss 定律：$\Omega \subset \mathbb{R}^3$ 是有界光滑带边区域，那么

$$\oiint_{\partial\Omega} E \cdot d\vec{\sigma} = \frac{Q}{\varepsilon_0},$$

其中 $Q = \sum_{i=1}^{N} q_i$ 是 Ω 内的总电荷并且我们假设对任意的 i，$x_i \in \mathring{\Omega}$。(提示：将奇点 x_i 挖掉，然后用 Stokes 公式)。

2) 假设 $\rho \in C^{\infty}(\mathbb{R}^3, \mathbb{R})$ 是电荷密度函数，根据 Coulomb 定律，它在 $x \in \mathbb{R}^3$ 处产生的电场（向量场）为（如果可积的话）

$$E(x) = \frac{1}{4\pi\varepsilon_0} \int_{\mathbb{R}^3} \frac{\rho(y)(x - y)}{|x - y|^3} dy.$$

试证明 Gauss 定律：如果 $\Omega \subset \mathbb{R}^3$ 是有界光滑带边区域，$\mathrm{supp}\rho \in \mathring{\Omega}$，那么

$$\oiint_{\partial\Omega} E \cdot d\vec{\sigma} = \frac{Q}{\varepsilon_0},$$

其中 $Q = \int_{\Omega} \rho(x) dx$ 是 Ω 内的总电荷。(提示：可以使用 Fubini 定理)

3) 证明，如果电荷密度函数 $\rho \in C_0^{\infty}(\mathbb{R}^3)$ 是具有紧支集的光滑函数，那么静电场

$$E(x) = \frac{1}{4\pi\varepsilon_0} \int_{\mathbb{R}^3} \frac{\rho(y)(x - y)}{|x - y|^3} dy$$

是无旋的，即 $\nabla \times E = 0$。

以下几题是关于 Laplace 算子 Δ 的基本解

18. 定义 $\mathbb{R}^3$ 上的函数

$$E(x) = -\frac{1}{4\pi|x|}.$$

证明对 $\mathbb{R}^3$ 上 Lebesgue 可积函数 f，卷积

$$\Phi_f(x) = E * f(x) = -\frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{f(y)}{|x - y|} dy.$$

是 $\mathbb{R}^3$ 上的局部可积函数，即是 $\Phi_f(x)$ 在任何有界可测集上可积。进一步证明，如果 f 有界并且可积，那么 Φ_f 连续。

19. 对 $\mathbb{R}^3$ 上的连续可微函数 G, H 有

$$G\Delta H - H\Delta G = \mathrm{div}(G\nabla H - H\nabla G).$$

20. 证明对 $\mathbb{R}^3$ 上任意光滑有紧支集的函数 $\varphi(x)$，我们有

$$\int_{\mathbb{R}^3} E(x)\triangle\varphi(x)dx = \varphi(0).$$

（提示：注意到函数 E 在原点不连续，可以将原点的一个邻域扣掉，再用 Stokes 公式）

21. 假设 $f \in C^2(\mathbb{R}^3)$ 并且 f 在 $\mathbb{R}^3$ 上 Lebesgue 可积。证明，$\Phi_f(x) \in C^2(\mathbb{R}^3)$ 并且

$$\triangle\Phi_f(x) = f.$$

22. 证明，如果 f 是 Lebesgue 可积的，那么存在一个仅依赖于 f 的常数 C，使得对任意的 $\lambda > 0$，我们有

$$m\big(x \in \mathbb{R}^3 \big| |\Phi_f(x)| > \lambda\big) \leqslant \frac{C}{\lambda^3}.$$