14. 多元微分学

14.1. 偏导数与全微分

14.1.1. 偏导数与方向导数

14.1.1.1.偏导数

设 $D \subset \mathbb{R}^d$ 为一个区域， $f: D \mapsto \mathbb{R}$ 为一个 d 元函数，取 $(x_1, x_2, \dots, x_d) \in D$ 。
如果一元函数 $f(x, x_2, x_3, \dots, x_d)$ 在 $x = x_1$ 处可导，则称 $\left.\dfrac{\mathrm{d}\, f(x, x_2, x_3, \dots, x_d)}{\mathrm{d}\, x}\right|_{x=x_1}$ 为 f 在 $(x_1, x_2, \dots, x_d)$ 处对 x 的偏导数，记为 $\dfrac{\partial f}{\partial x}(x_1, x_2, \dots, x_d)$ 或 $f_1(x_1, x_2, \dots, x_d)$ 等.
实际上有

$$\frac{\partial f}{\partial x}(x_1, x_2, \dots, x_d) = \left.\frac{\mathrm{d}\, f(x, x_2, x_3, \dots, x_d)}{\mathrm{d}\, x}\right|_{x=x_1} = \lim_{x \to x_1} \frac{f(x, x_2, x_3, \dots, x_d) - f(x_1, x_2, x_3, \dots, x_d)}{x - x_1}$$

14.1.1.2. 方向导数

14.1.1.2.1. 射线的方向

在 $\mathbb{R}^d$ 上过原点的射线可由其与单位圆的交点唯一确定，也就可以用单位向量代表其方向。
设 $\boldsymbol{v} \in \mathbb{R}^d$ 为单位向量，记 $\boldsymbol{v} = (\cos\theta_1, \cos\theta_2, \dots, \cos\theta_d)$ ， 其中 $\theta_j \in [0, \pi]$ 是向量 $\boldsymbol{v}$ 与 x_j 轴的正向的夹角。那么自然有 $\sum\limits_{i=1}^{d} \cos^2\theta_i = 1$.

14.1.1.2.2. 方向导数

设 $D \subset \mathbb{R}^d$ 为一个区域， $f: D \mapsto \mathbb{R}$ 为一个 d 元函数，取 $\boldsymbol{x} = (x_1, x_2, \dots, x_d) \in D$ ， 方向 $\boldsymbol{v} = (\cos\theta_1, \cos\theta_2, \dots, \cos\theta_d)$ 。
如果 $\lim\limits_{t \to 0^+} \dfrac{f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x})}{t}$ 存在，则称其为 f 在 $\boldsymbol{x}$ 处沿着 $\boldsymbol{v}$ 的方向导数，记为 $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x})$ 。

14.1.1.3. 偏导数与方向导数

记 $\boldsymbol{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ 为第 i 个分量取 1、其余分量取 0 的单位向量
那么偏导数 $f_i(\boldsymbol{x}) = \dfrac{\partial f}{\partial x_i}(\boldsymbol{x}) = \lim\limits_{x \to x_1} \dfrac{f(\boldsymbol{x} + \boldsymbol{e}_i \Delta x_i) - f(\boldsymbol{x})}{\Delta x_i}$ 是双边极限，如果这个极限存在那么沿 $\boldsymbol{e}_i$ 的方向导数一定存在，而沿着 $-\boldsymbol{e}_i$ 的方向导数也存在

14.1.1.4. 方向导数与连续性

所有方向导数都存在且相同也不能保证连续性，因为只是从直线的趋近

14.1.2. 全微分

14.1.2.1. 可微

设区域 $D \subset \mathbb{R}^d$ ， 函数 $f: D \mapsto \mathbb{R}$ ， 取一点 $\boldsymbol{x}_0 \in D$ ， 记 $y = f(\boldsymbol{x})$
另设 $\boldsymbol{x} \in D$ ， 如果有一个常向量
$\boldsymbol{A} = (A_1, A_2, \dots, A_d),\ \text{s.t. } \Delta y = f(\boldsymbol{x}) - f(\boldsymbol{x}_0) = \boldsymbol{A} \cdot \Delta\boldsymbol{x} + o(|\Delta\boldsymbol{x}|), \Delta\boldsymbol{x} \to 0$ ， 其中 $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_0$ ， 则称 $f(\boldsymbol{x})$ 在 $\boldsymbol{x} = \boldsymbol{x}_0$ 处可微，而 $\boldsymbol{A} \cdot \Delta\boldsymbol{x}$ 称为 $f(\boldsymbol{x})$ 在 $\boldsymbol{x} = \boldsymbol{x}_0$ 处的全微分，记为 $\mathrm{d}\, f(\boldsymbol{x}_0)$.

14.1.2.2. 可微的必要条件

设区域 $D \subset \mathbb{R}^d$ ，函数 $f : D \mapsto \mathbb{R}$ ，如果 f 在 $\boldsymbol{x}_0$ 处可微，则 f 在 $\boldsymbol{x}_0$ 处连续，且偏导数 $\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0)$ 都存在，任何方向 $\boldsymbol{v}$ 上的方向导数 $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$ 都存在，且
$\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \dfrac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\cos\theta_1 + \dfrac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\cos\theta_2 + \cdots + \dfrac{\partial f}{\partial x_d}(\boldsymbol{x}_0)\cos\theta_d$.下证

14.1.2.2.1. 可微则连续

由 $\Delta y = \boldsymbol{A} \cdot \Delta \boldsymbol{x} + o(|\Delta \boldsymbol{x}|)$ ，则 $\Delta \boldsymbol{x} \to 0$ 时 $\Delta y = f(\boldsymbol{x}) - f(\boldsymbol{x}_0) \to 0$ ，故连续

14.1.2.2.2. 可微则偏导数存在

以 $\dfrac{\partial f}{\partial x_1}(\boldsymbol{x}_0)$ 为例， $f(x_1 + \Delta x, x_2, \ldots, x_d) - f(x_1, x_2, \ldots, x_d) = \boldsymbol{A} \cdot \Delta \boldsymbol{x} + o(|\Delta \boldsymbol{x}|) = \boldsymbol{A} \cdot \boldsymbol{e}_1 \Delta x_1 + o(|x_1|)$ ，那么 $\dfrac{\partial f}{\partial x_1}(\boldsymbol{x}_0) = \lim\limits_{\Delta x_1 \to 0} \dfrac{f(x_1 + \Delta x, x_2, \ldots, x_d) - f(x_1, x_2, \ldots, x_d)}{\Delta x_1} = \lim\limits_{\Delta x_1 \to 0} \dfrac{A \cdot \boldsymbol{e}_1 \Delta x_1 + o(|x_1|)}{\Delta x_1} = \boldsymbol{A} \cdot \boldsymbol{e}_1$ 存在

14.1.2.2.3. 可微则方向导数存在

$f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x}) = \boldsymbol{A} \cdot t\boldsymbol{v} + o(|t\boldsymbol{v}|) = \boldsymbol{A} \cdot t\boldsymbol{v} + o(|t|)$ ，那么
$\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \lim\limits_{t \to 0} \dfrac{f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x})}{t} = \lim\limits_{t \to 0} \dfrac{\boldsymbol{A} \cdot t\boldsymbol{v} + o(|t|)}{t} = \boldsymbol{A} \cdot \boldsymbol{v}$ ，而上证 $\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = \boldsymbol{A} \cdot \boldsymbol{e}_i$ ，另有
$\boldsymbol{v} = \sum\limits_{i=1}^{d} e_i \cos\theta_i$ ，于是 $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \boldsymbol{A} \cdot \boldsymbol{v} = \boldsymbol{A} \cdot \sum\limits_{i=1}^{d} e_i \cos\theta_i = \sum\limits_{i=1}^{d} \cos\theta_i (\boldsymbol{A} \cdot \boldsymbol{e}_i) = \sum\limits_{i=1}^{d} \cos\theta_i \dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0)$.

14.1.2.3. 可微的充分条件

设区域 $D \subset \mathbb{R}^d$ ，函数 $f : D \mapsto \mathbb{R}$ ，取一点 $\boldsymbol{x}_0 \in D$ ，如果 $\exists\, \delta_0 > 0,\ \text{s.t.}\ \forall\, 1 \le i \le d, f$ 在 $U(\boldsymbol{x}_0, \delta_0)$ 中 $\dfrac{\partial f}{\partial x_i}$ 都存在且 $f(\boldsymbol{x})$ 在 $\boldsymbol{x} = \boldsymbol{x}_0$ 处连续，则 f 在 $\boldsymbol{x}_0$ 处可微。下证

[[]]

14.1.2.4.全微分

设区域 $D \subset \mathbb{R}^d$ ，函数 $f : D \mapsto \mathbb{R}$ ，记 $C^1(D)$ 为 D 上所有满足 $\forall\, 1 \le i \le d, \dfrac{\partial f}{\partial x_i}$ 连续的函数 f 全体

如果 $f \in C^1(D)$ ，则 f 可微且 $\mathrm{d}\, f(\boldsymbol{x}) = \sum\limits_{i=1}^{d} \dfrac{\partial f}{\partial x_i} \mathrm{d}\, x_i$ 连续，故也称 f 为连续可微函数

14.1.3.梯度

14.1.3.1. 梯度

设区域 $D \subset \mathbb{R}^d$ ，函数 $f : D \mapsto \mathbb{R}$ 在 $\boldsymbol{x}_0$ 处可微，设 $\operatorname{grad} f(x_0) = (\dfrac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \dfrac{\partial f}{\partial x_2}(\boldsymbol{x}_0), \ldots, \dfrac{\partial f}{\partial \boldsymbol{x}_n}(x_0))$ ，称为函数 f 在 x_0 处的梯度(向量)

任选一个方向向量 $v = (\cos\theta_1, \cos\theta_2, \ldots, \cos\theta_n)$ ，由上面 14.1.2.2.3. 可微则方向导数存在可知
$\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \sum\limits_{i=1}^{d} \cos\theta_i \dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0)$ ，那么当 $\operatorname{grad} f(\boldsymbol{x}_0) \ne \boldsymbol{0}$ 时， $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$ 取到最大值 $\Leftrightarrow v = \dfrac{\operatorname{grad} f(\boldsymbol{x}_0)}{|\operatorname{grad} f(\boldsymbol{x}_0)|}$

14.1.3.2. 梯度的性质

常值函数 $\operatorname{grad}(\text{Constant}) = \mathbf{0}$
函数加减 $\operatorname{grad}(f \pm g)(\boldsymbol{x_0}) = \operatorname{grad} f(\boldsymbol{x_0}) \pm \operatorname{grad} g(\boldsymbol{x_0})$
函数相乘 $\operatorname{grad}(f \cdot g)(\boldsymbol{x_0}) = \operatorname{grad} f(\boldsymbol{x_0}) g(\boldsymbol{x_0}) + f(\boldsymbol{x_0}) \operatorname{grad} g(\boldsymbol{x_0})$
函数相除 $\operatorname{grad}(\dfrac{f}{g})(\boldsymbol{x_0}) = \dfrac{\operatorname{grad} f(\boldsymbol{x_0}) g(\boldsymbol{x_0}) - f(\boldsymbol{x_0}) \operatorname{grad} g(\boldsymbol{x_0})}{g^2(\boldsymbol{x_0})}$

14.2. 向量函数的导数与全微分

14.2.1. 向量函数可微

14.2.1.1. 向量函数可微

设区域 $D \subset \mathbb{R}^n$ ，函数 $f: D \mapsto \mathbb{R}^m$ ，取一点 $\boldsymbol{x}_0 \in D$ ，记 $y = f(\boldsymbol{x})$
另设 $\boldsymbol{x} \in D$ ，如果有一个常矩阵 $A = A_{m\times n}$ s.t. $\Delta y = f(\boldsymbol{x}) - f(\boldsymbol{x}_0) = A \times \Delta \boldsymbol{x} + \alpha(|\Delta \boldsymbol{x}|)$ 且 $\dfrac{|\alpha(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|} \to 0, \Delta \boldsymbol{x} \to 0$ ，其中 $\Delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_0$ ，则称 $f(\boldsymbol{x})$ 在 $\boldsymbol{x} = \boldsymbol{x}_0$ 处可微，而 $A \times \Delta \boldsymbol{x}$ 称为 $f(\boldsymbol{x})$ 在 $\boldsymbol{x} = \boldsymbol{x}_0$ 处的全微分，记为 $\mathrm{d}\, f(\boldsymbol{x}_0) = A \,\mathrm{d}\, \boldsymbol{x}$ ，其中 $\mathrm{d}\, \boldsymbol{x} = (\mathrm{d}\, x_1, \mathrm{d}\, x_2, \dots, \mathrm{d}\, x_n)^T$.
同样地，如果函数 f 在区域 D 上的每一点都可微，则称 f 在 D 上可微

14.2.1.2. 与分量函数可微的关系

设区域 $D \subset \mathbb{R}^n$ ，函数 $f: D \mapsto \mathbb{R}^m$ ，取一点 $\boldsymbol{x}_0 \in D$ ，则 f 在 $\boldsymbol{x}_0$ 处可微 $\Leftrightarrow$ 对 $\forall\, 1 \le i \le m, f_i$ 在 $\boldsymbol{x}_0$ 处都可微。下证

- 必要性：向量函数可微，则 $\Delta y = A \times \Delta \boldsymbol{x} + \alpha(|\Delta \boldsymbol{x}|)$ ，取矩阵 A 的第 i 行记 $\boldsymbol{\alpha}_i$ ，那么对向量函数的第 i 个分量函数都有 $\Delta y_i = \boldsymbol{\alpha}_i \cdot \Delta \boldsymbol{x} + \alpha_i(\Delta \boldsymbol{x})$ ，而 $|\alpha_i| \le |\alpha|$ 则有 $0 \le \lim\limits_{\boldsymbol{x}\to \mathbf{0}} \dfrac{|\alpha_i(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|} \le \lim\limits_{\boldsymbol{x}\to \mathbf{0}} \dfrac{|\alpha(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|} = 0$ ，则可以记 $\alpha_i(\Delta \boldsymbol{x}) = o(\Delta \boldsymbol{x})$ ，也就是 $\Delta y_i = \boldsymbol{\alpha}_i \cdot \Delta \boldsymbol{x} + o(\Delta \boldsymbol{x})$
- 充分性：分量函数可微，则 $\alpha_i \to 0$ ，那么 $|\alpha| = \sqrt{\sum\limits_{i=1}^{m} \alpha_i^2} \le \sum\limits_{i=1}^{m} |\alpha_i|$ ，其余的思路同上，其中可以得到 A 是 Jacobi 矩阵

$$A = \frac{\partial(f_1, f_2, \dots, f_m)}{\partial(x_1, x_2, \dots, x_n)} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}_{m\times n}$$

14.2.1.3. 连续可微

如果 $\dfrac{\partial f_i}{\partial \boldsymbol{x}_j}, \forall\, 1 \le i \le m, \forall\, 1 \le j \le n$ 都连续，则称函数 f 连续可微，记 $f \in C^1(D, \mathbb{R}^m)$.

14.2.2. 可微的几何意义

14.2.2.1.切平面

以 n=2 为例，设函数 $z = f(x, y)$ 在 (x_0, y_0) 处可微，那么 $\Delta z = A\Delta x + B\Delta y + o(\rho)$ ，其中 $\rho = \sqrt{x^2 + y^2}$ ，那么写为 $z = z_0 + A(x - x_0) + B(y - y_0)$ ，称为函数在 (x_0, y_0) 处的切平面

14.3. 多元函数的求导法

14.3.1. 导数的四则运算法则

设区域 $D \subset \mathbb{R}^d$ ，函数 $f,g: D \mapsto \mathbb{R}$ 在区域 D 内都可微，则 $\forall\, \boldsymbol{x} \in D$ 都有
导数的加减 $(f \pm g)'(\boldsymbol{x}) = f'(\boldsymbol{x}) \pm g'(\boldsymbol{x})$
导数的乘法 $(f \cdot g)'(\boldsymbol{x}) = f'(\boldsymbol{x})g(\boldsymbol{x}) + f(\boldsymbol{x})g'(\boldsymbol{x})$
导数的除法 $(\dfrac{f}{g})'(\boldsymbol{x}) = \dfrac{f'(\boldsymbol{x})g(\boldsymbol{x}) - f(\boldsymbol{x})g'(\boldsymbol{x})}{g^2(\boldsymbol{x})}$

14.3.2. 复合函数的求导法则

设区域 $\Omega \subset \mathbb{R}^m, D \subset \mathbb{R}^n$ 函数 $u: D \to \Omega, y = f(u): \Omega \to \mathbb{R}$ ，设 u 在 $\boldsymbol{x}_0$ 处可微， f 在 $\boldsymbol{u}_0 = u(\boldsymbol{x}_0)$ 处可微，则 $f \circ u: D \to \mathbb{R}$ 在 $\boldsymbol{x}_0$ 处可微，且 $(f \circ u)'(\boldsymbol{x}_0) = f'(u)u'(\boldsymbol{x}_0)$ ，下证
由于 f 在 $\boldsymbol{u}_0$ 处可微，则 $\Delta f = f(\boldsymbol{u}_0 + \Delta \boldsymbol{u}) - f(\Delta \boldsymbol{u}) = f'(\boldsymbol{u}_0)\Delta \boldsymbol{u} + \alpha(\Delta \boldsymbol{u})$
定义 $\beta(\Delta \boldsymbol{u}) = \begin{cases} \dfrac{\alpha(\Delta \boldsymbol{u})}{|\Delta \boldsymbol{u}|} & \Delta \boldsymbol{u} \neq \boldsymbol{0} \\ 0 & \Delta \boldsymbol{u} = \boldsymbol{0} \end{cases}$ ，显然连续，则 $\Delta f = f'(\boldsymbol{u}_0)\Delta \boldsymbol{u} + \beta(\Delta \boldsymbol{u})|\Delta \boldsymbol{u}|$
设 $\boldsymbol{x}_0$ 有增量 $\Delta \boldsymbol{x} \neq 0$ ，定义 $\Delta u = u(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) - u(\boldsymbol{x}_0 = u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})$

$$\begin{aligned}
\Delta(f \circ u) &= (f \circ u)(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) - (f \circ u)(\boldsymbol{x}_0) \\
&= f(\boldsymbol{u}_0 + \Delta \boldsymbol{u}) - f(\boldsymbol{u}_0) \\
&= f'(\boldsymbol{u}_0)\Delta \boldsymbol{u} + \beta(\Delta \boldsymbol{u})|\Delta \boldsymbol{u}| \\
&= f'(\boldsymbol{u}_0)[u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})] + \beta(\Delta \boldsymbol{u})|u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})| \\
&= f'(\boldsymbol{u}_0)u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + f'(\boldsymbol{u}_0)\gamma(\Delta \boldsymbol{x}) + \beta(\Delta \boldsymbol{u})|u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})|
\end{aligned}$$

要证 $\lim\limits_{\Delta \boldsymbol{x} \to \boldsymbol{0}} \dfrac{f'(\boldsymbol{u}_0)\gamma(\Delta \boldsymbol{x}) + \beta(\Delta \boldsymbol{u})|u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|} = 0$ ，显然前者 $\lim\limits_{\Delta \boldsymbol{x} \to \boldsymbol{0}} \dfrac{f'(\boldsymbol{u}_0)\gamma(\Delta \boldsymbol{x})}{|\Delta \boldsymbol{x}|} = 0$

$$\begin{aligned}
\lim_{\Delta \boldsymbol{x} \to \boldsymbol{0}} \frac{\beta(\Delta \boldsymbol{u})|u'(\boldsymbol{x}_0)\Delta \boldsymbol{x} + \gamma(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|} &= \lim_{\Delta \boldsymbol{x} \to \boldsymbol{0}} \left|\beta(\Delta \boldsymbol{u}) \left[u'(\boldsymbol{x}_0)\frac{\Delta \boldsymbol{x}}{|\Delta \boldsymbol{x}|} + \frac{\gamma(\Delta \boldsymbol{x})}{|\Delta \boldsymbol{x}|}\right]\right| \\
&\leq \lim_{\Delta \boldsymbol{x} \to \boldsymbol{0}} |\beta(\Delta \boldsymbol{u})| \left[|u'(\boldsymbol{x}_0)| \cdot \frac{|\Delta \boldsymbol{x}|}{|\Delta \boldsymbol{x}|} + \frac{|\gamma(\Delta \boldsymbol{x})|}{|\Delta \boldsymbol{x}|}\right] \\
&= 0
\end{aligned}$$

14.3.3. 多元向量函数的偏导数

设函数 $f(\boldsymbol{u}) = f(u_1, u_2, \ldots, u_m): \Omega \subset \mathbb{R}^m \to \mathbb{R}$ 在 $\boldsymbol{u}_0 = (u_1^0, u_2^0, \ldots, u_m^0)^T$ 处可微，设向量函数 $u(\boldsymbol{x}) = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}), \ldots, u_n(\boldsymbol{x}))^T: D \subset \mathbb{R}^n \to \mathbb{R}^m$ 在 $\boldsymbol{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0)^T$ 处可微，且 $\boldsymbol{u}_0 = u(\boldsymbol{x}_0)$ ，则复合函数 $f(u(\boldsymbol{x}))$ 在 $\boldsymbol{x}_0$ 处可偏导，且 $\dfrac{\partial f(u(x_i))}{\partial \boldsymbol{x}} = \sum\limits_{j=1}^{m} (\dfrac{\partial f(\boldsymbol{u}_0)}{\partial u_j}) \cdot \dfrac{\partial u_j(\boldsymbol{x}_0)}{\partial x_i}, \forall\, 1 \leq i \leq n$

14.3.4. 高阶偏导数

14.3.4.1. 高阶偏导数

如果函数 $f(\boldsymbol{x})$ 的偏导数 $\dfrac{\partial f}{\partial x_i}$ 本身仍然可偏导，则可以求其偏导数，称为二阶偏导数，记 $\dfrac{\partial}{\partial x_j}(\dfrac{\partial f}{\partial x_i}) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$ 。类似地，可以定义三阶及更高阶的偏导数

14.3.4.2. 求导顺序无关

如果函数 $f(\boldsymbol{x})$ 的两个二阶混合偏导数 $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ 和 $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ 在点 $\boldsymbol{x}_0$ 的邻域内存在，且在点 $\boldsymbol{x}_0$ 处连续，那么在该区域内有 $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0)$

证明二元函数的情况。构造一个点 (x_0, y_0) 为中心、长 h 宽 k 的矩形，定义 $\Delta(h,k) = f(x_0+h, y_0+k) - f(x_0, y_0)$

构造辅助函数 $g(y) = f(x_0+h, y) - f(x_0, y)$ ，则 $\Delta(h,k) = g(y_0+k) - g(y_0) = g'(y_0+\theta k)\cdot k$ ，其中 $0 \le \theta \le 1$

而显然 $g'(y) = f_y(x_0+h, y) - f_y(x_0, y)$ ，也就得到 $\Delta(h,k) = [f_y(x_0+h, y_0+\theta k) - f_y(x_0, y_0+\theta k)]\cdot k$

再构造辅助函数 $\phi(x) = f_y(x, y_0+\theta k)$ ，那么

$f_y(x_0+h, y_0+\theta k) - f_y(x_0, y_0+\theta k) = \phi(x_0+h) - \phi(x_0) = \phi(x_0+\epsilon h)\cdot h$ ，其中 $0 \le \epsilon \le 1$

并且显然有 $\phi'(x) = f_{yx}(x, y_0+\theta k)$ ，于是 $\Delta(h,k) = f_{yx}(x_0+\epsilon h, y_0+\theta k)\cdot hk$

重复先对 x 再对 y 的上述过程，则得到 $\dfrac{\Delta(h,k)}{hk} = f_{yx}(x_0+\epsilon h, y_0+\theta k) = f_{xy}(x_0+\epsilon' h, y_0+\theta' k)$

令 $(h,k) \to (0,0)$ 则由连续性得到 $f_{yx}(x_0,y_0) = f_{xy}(x_0,y_0)$ ，多元函数或高阶偏导数同理

14.3.5. 一阶微分的形式不变性

设 $z = f(u,v)$，且 u, v 是自变量或者是其他变量（如 x, y）的函数。

- 如果 u, v 是自变量，则 $\mathrm{d}\,z = \dfrac{\partial z}{\partial u}\mathrm{d}\,u + \dfrac{\partial z}{\partial v}\mathrm{d}\,v$
- 如果 $u = u(x,y), v = v(x,y)$ ，则根据链式法则， $\mathrm{d}\,z = \dfrac{\partial z}{\partial x}\mathrm{d}\,x + \dfrac{\partial z}{\partial y}\mathrm{d}\,y$ 。但是可以证明这个表达式最终整理回 $\dfrac{\partial z}{\partial u}\mathrm{d}\,u + \dfrac{\partial z}{\partial v}\mathrm{d}\,v$ 的形式

$$
\begin{aligned}
\mathrm{d}\,z &= \frac{\partial z}{\partial x}\mathrm{d}\,x + \frac{\partial z}{\partial y}\mathrm{d}\,y \\
&= (\frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x})\mathrm{d}\,x + (\frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y})\mathrm{d}\,y \\
&= \frac{\partial z}{\partial u}(\frac{\partial u}{\partial x}\mathrm{d}\,x + \frac{\partial u}{\partial y}\mathrm{d}\,y) + \frac{\partial z}{\partial v}(\frac{\partial v}{\partial x}\mathrm{d}\,x + \frac{\partial v}{\partial y}\mathrm{d}\,y) \\
&= \frac{\partial z}{\partial u}\mathrm{d}\,u + \frac{\partial z}{\partial v}\mathrm{d}\,v
\end{aligned}
$$

14.3.6. 高阶微分

14.3.6.1. 二阶微分

二阶微分定义为一阶微分的微分，将一阶微分看作关于 x, y 的函数，再求微分，记 $\mathrm{d}^2 f$

$$
\begin{aligned}
\mathrm{d}^2 f &= \mathrm{d}(\mathrm{d}\,f) \\
&= \mathrm{d}(f_x\,\mathrm{d}\,x + f_y\,\mathrm{d}\,y) \\
&= \mathrm{d}(f_x)\,\mathrm{d}\,x + \mathrm{d}(f_y)\,\mathrm{d}\,y \\
&= (f_{xx}\,\mathrm{d}\,x + f_{xy}\,\mathrm{d}\,y)\,\mathrm{d}\,x + (f_{yx}\,\mathrm{d}\,x + f_{yy}\,\mathrm{d}\,y)\,\mathrm{d}\,y \\
&= f_{xx}(\mathrm{d}\,x)^2 + (f_{xy} + f_{yx})\,\mathrm{d}\,x\,\mathrm{d}\,y + f_{yy}(\mathrm{d}\,y)^2
\end{aligned}
$$

如果二阶混合偏导数连续则 $f_{xy} = f_{yx}$ ，那么 $\mathrm{d}^2 f = f_{xx}(\mathrm{d}\,x)^2 + 2f_{xy}\,\mathrm{d}\,x\,\mathrm{d}\,y + f_{yy}(\mathrm{d}\,y)^2$

14.3.6.2. 高阶微分

N 阶微分递归定义为 n-1 阶微分的微分，记 $\mathrm{d}^n f = \mathrm{d}(\mathrm{d}^{n-1} f)$ ，形式上作二项式展开为

$$\mathrm{d}^n f = \sum_{k=0}^{n} C_n^k \frac{\partial f}{\partial x^k \partial y^{n-k}} (\mathrm{d}\, x)^k (\mathrm{d}\, y)^{n-k}$$

14.4. 泰勒公式

14.4.1. 带拉格朗日余项的泰勒公式

设函数 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$ 的某邻域 $U(\boldsymbol{x}_0, \delta_0)$ 内有 $K+1$ 阶的连续偏导数，则对该邻域内的 $\forall\ \boldsymbol{x}_0 + \boldsymbol{h} = (x_1^0 + h_1, x_2^0 + h_2, \dots, x_n^0 + h_n)$ 都有

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \sum_{k=1}^{K} \frac{1}{k!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^k f(\boldsymbol{x}_0) + \frac{1}{(K+1)!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^{K+1} f(\boldsymbol{x}_0 + \theta \boldsymbol{h}), 0 < \theta < 1$$

其中 $R_{K+1}(\boldsymbol{h}) = \dfrac{1}{(K+1)!} (\sum\limits_{i=1}^{n} h_i \dfrac{\partial}{\partial x_i})^{K+1} f(\boldsymbol{x}_0 + \theta \boldsymbol{h})$ 称为拉格朗日余项，下证

构造一元函数 $\phi(t) = f(\boldsymbol{x_0} + t\boldsymbol{h}), 0 \le t \le 1$ ，则 $\phi(t)$ 在这个邻域也有 $K+1$ 阶的连续偏导数。借由一元函数的泰勒公式

$$\phi(t) = \phi(0) + \sum_{k=1}^{K} \frac{\phi^{(k)}(0)}{k!} t^k + \frac{\phi^{(k+1)}(\theta t)}{(K+1)!} t^{K+1}, 0 < \theta < 1$$

注意到 $\forall\ 1 \le k \le K, \phi^{(k)}(t) = (\sum\limits_{i=1}^{n} h_i \dfrac{\partial}{\partial x_i})^k f(\boldsymbol{x}_0 + t\boldsymbol{h})$ ，代入上式再令 $t = 1$ 即可

14.4.2. 带皮亚诺余项的泰勒公式

设函数 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$ 的某邻域 $U(\boldsymbol{x}_0, \delta_0)$ 内有 $K+1$ 阶的连续偏导数，则对该邻域内的 $\forall\ \boldsymbol{x}_0 + \boldsymbol{h} = (x_1^0 + h_1, x_2^0 + h_2, \dots, x_n^0 + h_n)$ 都有

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \sum_{k=1}^{K} \frac{1}{k!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^k f(\boldsymbol{x}_0) + o(|\boldsymbol{h}|^K)$$

其中 $o(|\boldsymbol{h}|^K)$ 称为皮亚诺余项，下证

由带拉格朗日余项的泰勒公式

$$\begin{aligned} f(\boldsymbol{x}_0 + \boldsymbol{h}) &= f(\boldsymbol{x}_0) + \sum_{k=1}^{K} \frac{1}{k!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^k f(\boldsymbol{x}_0) + \frac{1}{(K+1)!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^{K+1} f(\boldsymbol{x}_0 + \theta \boldsymbol{h}) \\ &= f(\boldsymbol{x}_0) + \sum_{k=1}^{K} \frac{1}{k!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^k f(\boldsymbol{x}_0) + \frac{1}{K!} (\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i})^{K} [f(\boldsymbol{x}_0 + \theta \boldsymbol{h}) - f(\boldsymbol{x}_0)] \end{aligned}$$

记 $R_K(\boldsymbol{h}) = \dfrac{1}{K!} (\sum\limits_{i=1}^{n} h_i \dfrac{\partial}{\partial x_i})^{K} [f(\boldsymbol{x}_0 + \theta \boldsymbol{h}) - f(\boldsymbol{x}_0)]$ ，由于邻域内有 $K+1$ 阶连续偏导数，则

$$\lim_{|\boldsymbol{h}| \to 0} \frac{R_K(\boldsymbol{h})}{|\boldsymbol{h}|^K} = \lim_{|\boldsymbol{h}| \to 0} \frac{1}{K!} (\sum_{i=1}^{n} \frac{h_i}{|\boldsymbol{h}|} \frac{\partial}{\partial x_i})^K [f(\boldsymbol{x}_0 + \theta \boldsymbol{h}) - f(\boldsymbol{x}_0)] = 0$$

14.4.3. 拉格朗日微分中值定理

在带拉格朗日余项的泰勒公式中取 $K = 0$ ，则对区域 D 内有连续偏导数的函数 $f(\boldsymbol{x})$ ，$\forall\ 0 \le t \le 1, \boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) \in D$ ，都有

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &= \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}_0 + \theta(\boldsymbol{x} - \boldsymbol{x}_0))}{\partial x_i} (x_i - x_i^0) \\ &= f'(\boldsymbol{x}_0 + \theta(\boldsymbol{x} - \boldsymbol{x}_0)) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)^T \end{aligned}$$

其推论是，偏导数全为 0 的函数是常值函数

14.4.4. Hessi 矩阵

在带拉格朗日余项的泰勒公式中取 $K=1$ ，则有

$$f(\boldsymbol{x}_0+\boldsymbol{h})=f(\boldsymbol{x}_0)+f'(\boldsymbol{x}_0)\cdot\boldsymbol{h}^{T+}\frac{1}{2}\boldsymbol{h}\cdot H_f(\boldsymbol{x}_0)\cdot\boldsymbol{h}^T+o(|\boldsymbol{h}|^2)$$

其中 Hessi 矩阵 $H_f(\boldsymbol{x}_0)=(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(\boldsymbol{x}_0))_{n\times n}$

14.5. 隐函数存在定理

14.5.1. 单个方程的隐函数存在定理

设二元函数 $F(x,y)$ 在邻域 $U((x_0,y_0),\delta)$ 内满足三个条件： $F(x_0,y_0)=0$ ； $F(x,y),F'_y(x,y)$ 在邻域内连续； $F'_y(x_0,y_0)\neq 0$ ，则 $\exists\, 0<\delta_0<\delta$ ，使 $U(x_0,\delta_0)$ 内存在唯一满足下列条件的函数 $y=f(x)$ ：
$y_0=f(x_0)$ ； $F(x,f(x))=0,\forall\, x\in U(x_0,\delta_0)$ ；如果 $F'_x(x,y)$ 在邻域内连续则 $f(x)$ 在 $U(x_0,\delta_0)$ 内有连续导数，且 $f'(x)=-\dfrac{-F'_x(x,y)}{F'_y(x,y)}$ ，下证

不妨设 $F'_y(x_0,y_0)>0$ ，由 F 的连续性知存在邻域使 $F'_y(x,y)>0$ ，若固定 $x=x_0$ 则函数 $F(x_0,y)$ 关于 y 严格单增，而 $F(x_0,y_0)=0$ ，则存在 $y_1<y_0<y_2$ 使 $F(x_0,y_1)<0<F(x_0,y_2)$ 。再由 F 的连续性知存在更小的邻域内 $F(x,y_1)<0<F(x,y_2)$ ，这个邻域内对每一个固定的 x ，由介值定理都有一个对应的 y 使 $F(x,y)=0$ ，这个对应关系就是要找的函数 $y=f(x)$ ，唯一性由 y 的单调保证。连续性取 ϵ 邻域可证，导数：
任取 $x\in U(x_0,\delta_0)$ ，对充分小的 Δx ，记 $\Delta y=f(x+\Delta x)-f(x)$ ，由 $F(x,y)=0$ 和 $F(x+\Delta x,y+\Delta y)=0$ 得

$$0=F(x+\Delta x,y+\Delta y)-F(x,y)=F'_x(x+\theta\Delta x,y+\theta\Delta y)\Delta x+F'_y(x+\theta\Delta x,y+\theta\Delta y)\Delta y$$

于是 $\dfrac{\Delta y}{\Delta x}=-\dfrac{F'_x(x+\theta\Delta x,y+\theta\Delta y)}{F'_y(x+\theta\Delta x,y+\theta\Delta y)}$ ，令 $\Delta x\to 0$ ，由两个偏导数的连续性得证

14.5.2. 多元函数的隐函数存在定理

设多元函数 $F(\boldsymbol{x},y)$ 在邻域 $U(\boldsymbol{x}_0,\delta)\times U(y_0,\delta)$ 内满足三个条件： $F(\boldsymbol{x}_0,y_0)=0$ ； $F(\boldsymbol{x},y),F'_y(\boldsymbol{x},y)$ 在邻域内连续； $F'_y(\boldsymbol{x}_0,y_0)\neq 0$ ，则 $\exists\, 0<\delta_0<\delta$ ，使 $U(\boldsymbol{x}_0,\delta_0)$ 内存在唯一满足下列条件的函数 $y=f(\boldsymbol{x})$ ：
$y_0=f(\boldsymbol{x}_0)$ ； $F(\boldsymbol{x},f(\boldsymbol{x}))=0,\forall\,\boldsymbol{x}\in U(\boldsymbol{x}_0,\delta_0)$ ；如果各个偏导数在邻域内连续则 $f(\boldsymbol{x})$ 在 $U(\boldsymbol{x}_0,\delta_0)$ 内有各个连续偏导数，且 $\dfrac{\partial f}{\partial x_i}=-\dfrac{F'_{x_i}(\boldsymbol{x},y)}{F'_y(\boldsymbol{x},y)},\forall\, 1\le i\le n$

14.5.3. 隐函数组存在定理

设向量函数 $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{u})=(F_1(\boldsymbol{x},\boldsymbol{u}),F_2(\boldsymbol{x},\boldsymbol{u}),\dots,F_m(\boldsymbol{x},\boldsymbol{u}))$ 在 $U(\boldsymbol{x}_0,\delta)\times U(\boldsymbol{u}_0,\delta)$ 上满足三个条件：
$F_j(\boldsymbol{x}_0,\boldsymbol{u}_0)=0,\forall\, 1\le j\le m$ ；对 $\forall\, 1\le j\le m$ ，函数 $F_j(\boldsymbol{x},\boldsymbol{u})$ 及各个偏导数都连续；
$\dfrac{\partial(F_1,F_2,\dots,F_m)}{\partial(u_1,u_2,\dots,u_m)}(\boldsymbol{x}_0,\boldsymbol{u}_0)\neq 0$ ，则 $\exists\, 0<\delta_0<\delta$ ，使 $U(\boldsymbol{x}_0,\delta_0)$ 内存在唯一满足下列条件的 m 维 n 元向量函数 $\boldsymbol{u}=f(\boldsymbol{x})=(f_1(\boldsymbol{x}),f_2(\boldsymbol{x}),\dots,f_m(\boldsymbol{x}))$ ： $\boldsymbol{u}_0=(f_1(\boldsymbol{x}_0),f_2(\boldsymbol{x}_0),\dots,f_m(\boldsymbol{x}_0))$ ；对 $\forall\, 1\le j\le m,\forall\,\boldsymbol{x}\in U(\boldsymbol{x}_0,\delta_0)$ 都有 $F_j(\boldsymbol{x},f_1(\boldsymbol{x}),\dots,f_m(\boldsymbol{x}))=0$ ； $\boldsymbol{u}=f(\boldsymbol{x})$ 的每个分量函数 $f_j(\boldsymbol{x}),\forall\, 1\le j\le m$ 在 $U(\boldsymbol{x}_0,\delta_0)$ 内都有连续的偏导数，如果记矩阵 $A=(\dfrac{\partial F_i(\boldsymbol{x},\boldsymbol{u})}{\partial x_j})_{m\times n},B=(\dfrac{\partial F_i(\boldsymbol{x},\boldsymbol{u})}{\partial u_j})_{m\times m}$ ，则 $f'(\boldsymbol{x})=-B^{-1}\cdot A$

14.6. 逆映射存在定理

设 $\boldsymbol{y}=(y_1,y_2,\dots,y_n)=(f_1(\boldsymbol{x}),f_2(\boldsymbol{x},\dots,f_n(\boldsymbol{x}))=f(\boldsymbol{x})$ 是区域 $D\subset\mathbb{R}^n$ 到区域 $\Omega\subset\mathbb{R}^n$ 的一个 C^1 映射，并且在 $\boldsymbol{x}_0\in D$ 处满足 $\dfrac{\partial(f_1,f_2,\dots,f_n)}{\partial(x_1,x_2,\dots,x_n)}(\boldsymbol{x}_0)\neq 0$ 。记 $\boldsymbol{y}_0=f(\boldsymbol{x}_0)$ ，则存在一个邻域 $U(\boldsymbol{x}_0,\delta_0)\subset D$ 使得映射 $\boldsymbol{y}=f(\boldsymbol{x})$ 是 $U(\boldsymbol{x}_0,\delta_0)$ 到 $f(U(\boldsymbol{x}_0,\delta_0))$ 的一个 C^1 同胚映射，其中 $f(U(\boldsymbol{x}_0,\delta_0))$ 表示一个依赖于 $U(\boldsymbol{x}_0,\delta_0)$ 的区域且包含 $\boldsymbol{y}_0$ 。下证
$\forall\, 1\le j\le n$ ，记 $F_j(\boldsymbol{x},\boldsymbol{y})=y_j-f_j(\boldsymbol{x})$ ，考虑方程组 $F_j(\boldsymbol{x},\boldsymbol{y})=0,\forall\,1\le j\le n$ ，则满足
$\forall\,1\le j\le n, F_j(\boldsymbol{x}_0,\boldsymbol{y}_0)=0$ 以及 $\exists\,\delta'>0,\ \text{s.t.}\ \forall\,1\le j\le n$ 都有 $F_j(\boldsymbol{x},\boldsymbol{y})$ 在邻域 $U((\boldsymbol{x}_0,\boldsymbol{y}_0),\delta')$ 内有连续偏导数，且 $\dfrac{\partial(F_1,F_2,\dots,F_n)}{\partial(x_1,x_2,\dots,x_n)}(\boldsymbol{x}_0,\boldsymbol{y}_0)=(-1)^n\dfrac{\partial(f_1,f_2,\dots,f_n)}{\partial(x_1,x_2,\dots,x_n)}(\boldsymbol{x}_0)\neq 0$ ，那么由 14.5.3. 隐函数组存在定理知 $\exists\,\boldsymbol{x}=g(\boldsymbol{y})=(g_1(\boldsymbol{y}),g_2(\boldsymbol{y}),\dots,g_n(\boldsymbol{y}))$ 满足 $\boldsymbol{x}_0=g(\boldsymbol{y}_0)$ 以及
$y_j-f_j(g_1(\boldsymbol{y}),g_2(\boldsymbol{y}),\dots,g_n(\boldsymbol{y}))=0,\forall\,1\le j\le n$ 并且 $(g_1(\boldsymbol{y}),g_2(\boldsymbol{y}),\dots,g_n(\boldsymbol{y}))$ 在邻域 $U(\boldsymbol{y}_0,\delta_1)$ 上有连续偏导数。以上所有表示 $\boldsymbol{x}=g(\boldsymbol{y})$ 是 $\boldsymbol{y}=f(\boldsymbol{x})$ 的逆映射。
C^1 同胚映射怎么证的我看不懂。

14.7. 多元函数的极值

14.7.1. 通常的极值问题

14.7.1.1. 极值和极值点

设函数 $u=f(\boldsymbol{x})$ 在区域 D 内有定义，取 $\boldsymbol{x}_0\in D$ ，如果
$\exists\,U(\boldsymbol{x}_0,\delta_0)\in D,\ \text{s.t.}\ \forall\,\boldsymbol{x}\in U(\boldsymbol{x}_0,\delta_0), f(\boldsymbol{x})\le f(\boldsymbol{x}_0)$ ，则称 $u=f(\boldsymbol{x})$ 在 $\boldsymbol{x}_0$ 处取极大值， $\boldsymbol{x}_0$ 称为极大值点，同样可以定义极小值和极小值点，以及严格极值与严格极值点.

14.7.1.2. 极值点性质

设函数 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}_0$ 处取极值且关于 x_i 都可偏导，则 $\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0)=0,\forall\,1\le i\le n$ 。若该点处可微则 $f'(\boldsymbol{x}_0)=0$ ，此时称 $\boldsymbol{x}_0$ 为函数的一个驻点或临界点，非极值点的驻点称鞍点。下证
一元辅助函数 $F_i(x_i)=f(x_1^0,x_2^0,\dots,x_{i-1}^0,x_i,x_{i+1}^0,\dots,x_n^0)$ 在 x_i^0 处取极值，故 $F_i'(x_i^0)=\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0)=0$

14.7.1.3. Hessi 矩阵判定极值

设函数 $f(\boldsymbol{x})$ 在区域 $D\subset\mathbb{R}^n$ 中具有连续的二阶偏导数，且 $f'(\boldsymbol{x}_0)=0$ ，再设 Hessi 矩阵 $H_f(\boldsymbol{x}_0)$ 满秩，则其正定、负定、不定分别对应 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}_0$ 处取极小值、极大值、非极值，下证
在点 $\boldsymbol{x}_0$ 的充分小邻域 $U(\boldsymbol{x}_0,\delta_0)$ 内作泰勒展开

$$\begin{aligned}f(\boldsymbol{x})&=f(\boldsymbol{x}_0)+f'(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^T+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^T+o(|\boldsymbol{x}-\boldsymbol{x}_0|^2)\\&=f(\boldsymbol{x}_0)+\frac{1}{2}\frac{\boldsymbol{x}-\boldsymbol{x}_0}{|\boldsymbol{x}-\boldsymbol{x}_0|}H_f(\boldsymbol{x}_0)\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^T}{|\boldsymbol{x}-\boldsymbol{x}_0|}\cdot|\boldsymbol{x}-\boldsymbol{x}_0|^2+o(|\boldsymbol{x}-\boldsymbol{x}_0|^2)\end{aligned}$$

记 $\boldsymbol{h}=\dfrac{\boldsymbol{x}-\boldsymbol{x}_0}{|\boldsymbol{x}-\boldsymbol{x}_0|}$ 则 $|\boldsymbol{h}|=1$ 。注意到 $\boldsymbol{h}H_f(\boldsymbol{x}_0)\boldsymbol{h}^T$ 是定义在 $S(\boldsymbol{0};1)=\{\boldsymbol{h}\mid|\boldsymbol{h}|=1\}$ 上的连续函数，且定义域是紧集，可以取到最大值 M 和最小值 m

- 当矩阵正定时 $\forall\,\boldsymbol{x}\neq\boldsymbol{x}_0$ 都有 $(\boldsymbol{x}-\boldsymbol{x}_0)H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^T>0$ ，那么 $f(\boldsymbol{x})>f(\boldsymbol{x}_0)+\dfrac{m}{4}|\boldsymbol{x}-\boldsymbol{x}_0|^2>f(\boldsymbol{x})$
- 当矩阵负定时 $\forall\,\boldsymbol{x}\neq\boldsymbol{x}_0$ 都有 $(\boldsymbol{x}-\boldsymbol{x}_0)H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^T<0$ ，那么 $f(\boldsymbol{x})<f(\boldsymbol{x}_0)+\dfrac{M}{4}|\boldsymbol{x}-\boldsymbol{x}_0|^2<f(\boldsymbol{x})$

- 当矩阵不定时有 $m < 0 < M$ 于是存在 $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0) < f(\boldsymbol{x}_2)$

  对于常用的二元函数 $f(x,y)$ ，记其 Hessi 矩阵 $H_f(x_0,y_0) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ ，则用 A 和 $\begin{vmatrix} A & B \\ C & D \end{vmatrix}$ 的正负可以判断极值

- $\begin{vmatrix} A & B \\ C & D \end{vmatrix} > 0$ 取极值
  - $A > 0$ 为极小值
  - $A < 0$ 为极大值
- $\begin{vmatrix} A & B \\ C & D \end{vmatrix} < 0$ 不取极值

14.7.2. 拉格朗日乘数法解决条件极值问题

设函数 $f(\boldsymbol{x})$ 和 $\phi(\boldsymbol{x}) = (\phi_1(\boldsymbol{x}), \phi_2(\boldsymbol{x}), \dots, \phi_m(\boldsymbol{x}))$ 在区域 $D \subset \mathbb{R}^n$ 上各个偏导数连续，点 $\boldsymbol{x}_0$ 为函数 $f(\boldsymbol{x})$ 在条件 $\forall\, 1 \le i \le m, \phi_i(\boldsymbol{x}) = 0$ 下的极值点，且 $\phi'(\boldsymbol{x}_0)$ 的秩为 m，则 $\forall\, 1 \le i \le m, \exists\, \lambda_i$ 使得

$$\begin{cases} \dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0) + \sum\limits_{j=1}^{m} \lambda_j \dfrac{\partial \phi_j}{\partial x_i}(\boldsymbol{x}_0) = 0, & i = 1, 2, \dots, n \\ \phi_j(\boldsymbol{x}_0) = 0, & j = 1, 2, \dots, m \end{cases}$$

下证

$\phi'(\boldsymbol{x}_0)$ 的秩为 m，则其中 m 列组成的行列式非 0，不妨设 $\dfrac{\partial(\phi_1, \phi_2, \dots, \phi_m)}{\partial(x_{n-m+1}, x_{n-m+2}, \dots, x_n)}(\boldsymbol{x}_0) \neq 0$

根据 14.5.3. 隐函数组存在定理，在点 $\boldsymbol{x}_0$ 的某个邻域内唯一确定一组隐函数 $\{x_{n-m+j} = g_j(x_1, x_2, \dots, x_{n-m})\}, \forall\, 1 \le j \le m$ ，且具有连续的偏导数，满足 $x_{n-m+j}^0 = g_j(x_1^0, x_2^0, \dots, x_m^0), \forall\, 1 \le j \le m$ 以及

$$\phi_k(x_1, x_2, \dots, x_{n-m}, g_1(x_1, x_2, \dots, x_m), g_2(x_1, x_2, \dots, x_m), \dots, g_m(x_1, x_2, \dots, x_m)) = 0, \forall\, 1 \le k \le m$$

将得到的隐函数组代入原函数 $f(\boldsymbol{x})$ ，就将条件极值点 $\boldsymbol{x}_0$ 转化为普通极值点 $\boldsymbol{x}_0' = (x_1^0, x_2^0, \dots, x_{n-m}^0)$ ，于是对每一个分量

$$\frac{\partial f}{\partial x_k} = \frac{\partial f(\boldsymbol{x}_0)}{\partial x_k} + \sum_{i=1}^{m} \frac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+i}} \cdot \frac{\partial g_i(\boldsymbol{x}_0')}{\partial x_k} = 0, \forall\, 1 \le k \le m$$

记 $g(\boldsymbol{x}') = (g_1(\boldsymbol{x}'), g_2(\boldsymbol{x}'), \dots, g_m(\boldsymbol{x}'))$ ，其中 $\boldsymbol{x}' = (x_1, x_2, \dots, x_{n-m})$ 则可以将上面的 m 个等式写为方程组

$$\left(\frac{\partial f(\boldsymbol{x}_0)}{\partial x_1}, \frac{\partial f(\boldsymbol{x}_0)}{\partial x_2}, \dots, \frac{\partial f(\boldsymbol{x}_0)}{\partial x_m} \right) + \left(\frac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+1}}, \frac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+2}}, \dots, \frac{\partial f(\boldsymbol{x}_0)}{\partial x_n} \right) g'(\boldsymbol{x}_0') = 0$$

仍然根据 14.5.3. 隐函数组存在定理，对向量函数 $g(\boldsymbol{x}')$ ，其

$$\begin{aligned} g'(\boldsymbol{x}') &= -B^{-1} \times A \\ &= (\frac{\partial \phi(\boldsymbol{x}', g(\boldsymbol{x}'))}{\partial \boldsymbol{x}'})_{m \times m}^{-1} \times (\frac{\partial \phi(\boldsymbol{x}', g(\boldsymbol{x}'))}{\partial g(\boldsymbol{x}')})_{m \times (n-m)} \\ &= - \begin{pmatrix} \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_{n-m+1}} & \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_{n-m+2}} & \cdots & \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_n} \\ \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_{n-m+1}} & \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_{n-m+2}} & \cdots & \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_{n-m+1}} & \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_{n-m+2}} & \cdots & \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_n} \end{pmatrix}^{-1} \times \begin{pmatrix} \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_1} & \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_2} & \cdots & \dfrac{\partial \phi_1(\boldsymbol{x}_0)}{\partial x_{n-m}} \\ \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_1} & \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_2} & \cdots & \dfrac{\partial \phi_2(\boldsymbol{x}_0)}{\partial x_{n-m}} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_1} & \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_2} & \cdots & \dfrac{\partial \phi_m(\boldsymbol{x}_0)}{\partial x_{n-m}} \end{pmatrix} \end{aligned}$$

代入上式，则 $\left(\dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+1}}, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+2}}, \dots, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_n}\right) \times (-B^{-1})$ 是一个 m 维行向量，记为 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ ，则得到两个方程

$$\begin{cases} \left(\dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+1}}, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_{n-m+2}}, \dots, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_n}\right) + \lambda \times B = 0 \\ \left(\dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_1}, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_2}, \dots, \dfrac{\partial f(\boldsymbol{x}_0)}{\partial x_m}\right) + \lambda \times A = 0 \end{cases}$$

另外，构造函数 $F(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{j=1}^{m} \lambda_j \phi_j$ 则转化为 F 的普通极值问题，称拉格朗日乘数法

14.8. 多元微分学的几何应用

14.8.1. 曲线的切线与法平面

14.8.1.1. 简单曲线

设曲线 Γ 由连续映射 $h(t) = (x_1(t), x_2(t), \dot{;} x_n(t)), t \in [\alpha, \beta]$ 所确定
如果 $\forall\, t_1, t_2 \in [\alpha, \beta]$ 都有 $h(t_1) \neq h(t_2)$ 则称 Γ 为一条简单曲线
如果 $\forall\, t_1, t_2 \in [\alpha, \beta)$ 都有 $h(t_1) \neq h(t_2)$ 但是 $h(\alpha) = h(\beta)$ ，则称 Γ 为一条简单闭曲线，又称 **Jordan 曲线**

14.8.1.2. 参数方程组确定的曲线

设空间曲线由参数方程 $\Gamma = \begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases}, \quad t \in (\alpha, \beta)$ 给出，并且 $x(t), y(t), z(t)$ 都是 t 的可微曲线，以及 $x'^2(t) + y'^2(t) + z'^2(t) \neq 0$ ，则这样的曲线局部总是简单曲线
设一个点 $P_0(x(t_0), y(t_0), z(t_0))$ ，其切线定义为曲线上过另外一个点 $P_1(x(t), y(t), z(t))$ 与点 P_0 的割线当 $t \to t_0$ 时的极限位置
其割线方程为 $\dfrac{x - x(t_0)}{x(t) - x(t_0)} = \dfrac{y - y(t_0)}{y(t) - y(t_0)} = \dfrac{z - z(t_0)}{z(t) - z(t_0)}$ ，分母中同时除以 $t - t_0$ ，再取 $t \to t_0$ 则得到
切线方程 $\dfrac{x - x(t_0)}{x'(t_0)} = \dfrac{y - y(t_0)}{y'(t_0)} = \dfrac{z - z(t_0)}{z'(t_0)}$
由这个切线方程可以看出 $(x'(t_0), y'(t_0), z'(t_0))$ 是曲线在点 P_0 处的切向量
同时可以得到其法平面，即以此切向量为法向的平面，其方程为
$x'(t_0)(x - x(t_0)) + y'(t_0)(y - y(t_0)) + z'(t_0)(z - z(t_0)) = 0$

14.8.1.3. 两个方程联立确定的曲线

给定两个可微函数组成的方程组 $\begin{cases} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \end{cases}$ ，其中 $(x, y, z) \in D \subset \mathbb{R}^n$ 是一个区域
给定一个点 $P_0(x_0, y_0, z_0)$ ，当矩阵

$$\begin{pmatrix} F_1'(P_0) \\ F_2'(P_0) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} & \dfrac{\partial F_1}{\partial z} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} & \dfrac{\partial F_2}{\partial z} \end{pmatrix} \Big|_{P_0}$$

的秩为 2 的时候，由 14.5.2. 多元函数的隐函数存在定理 表明其中一个变量可以表示为其余两个变量的隐函数，故确定一条过点 P_0 的曲线
同样记其参数形式 $(x(t), y(t), z(t)), t \in (\alpha, \beta)$ ，其中 $(x(t_0), y(t_0), z(t_0)) = (x_0, y_0, z_0)$
那么由 14.8.1.2. 参数方程确定的曲线 得到曲线在此处的切向量为 $(x'(t_0), y'(t_0), z'(t_0))$
而方程组满足 $\begin{cases} F_1(x(t), y(t), z(t)) = 0 \\ F_2(x(t), y(t), z(t)) = 0 \end{cases}$ ，两个方程两边分别关于 t 求导数，得到

$\frac{\partial F_k(P_0)}{\partial x}x'(t_0)+\frac{\partial F_k(P_0)}{\partial y}y'(t_0)+\frac{\partial F_k(P_0)}{\partial z}z'(t_0)=0, k=1,2$ ，那么切向量和 $F_1'(P_0), F_2'(P_0)$ 两个向量都垂直，故和 $F_1'(P_0)\times F_2'(P_0)$ 平行

记 $A=\frac{\partial(F_1,F_2)}{\partial(y,z)}|_{P_0}, B=\frac{\partial(F_1,F_2)}{\partial(z,x)}|_{P_0}, C=\frac{\partial(F_1,F_2)}{\partial(x,y)}|_{P_0}$ ，则这个向量可以写为

$F_1'(P_0)\times F_2'(P_0)=A\boldsymbol{i}+B\boldsymbol{j}+C\boldsymbol{k}$

因此曲线的切向量可以写为 (A,B,C) ，其切线方程 $\frac{x-x_0}{A}=\frac{y-y_0}{B}=\frac{z-z_0}{C}$ ，法平面方程

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

14.8.2. 曲面的切平面与法线

14.8.2.1. 单个方程确定的曲面

设曲面 S 由方程 $F(x,y,z)=0$ 确定，并且 $F\in C^1(D)$ ，而 $D\subset\mathbb{R}^3$ 为一个区域。设点 $P_0(x_0,y_0,z_0)\in D$ 使得 $F(x_0,y_0,z_0)=0$

设曲面 S 上任意一条过点 P_0 的光滑曲线用参数方程写为 $(x(t),y(t),z(t)), t\in(\alpha,\beta)$ ，使得 $F(x(t),y(t),z(t))\equiv 0, \forall t\in(\alpha,\beta)$ ，并且 $(x(t_0),y(t_0),z(t_0))=(x_0,y_0,z_0)$

在上述方程两边对 t 求导，得到 $F_x'(P_0)x'(t_0)+F_y'(P_0)y'(t_0)+F_z'(P_0)z'(t_0)=0$ 也就是任意曲线的切向量 $(x'(t_0),y'(t_0),z'(t_0))$ 与固定向量 $F'(P_0)$ 正交，那么曲线在这点处的切线都在以 $F'(P_0)$ 为法向量的平面上，自然定义这个平面 $F_x'(x-x_0)+F_y'(y-y_0)+F_z'(z-z_0)=0$ 为点 P_0 处的切平面，而直线 $\frac{x-x_0}{F_x'}=\frac{y-y_0}{F_y'}=\frac{z-z_0}{F_z'}$ 为点 P_0 处的法线

14.8.2.2. 参数方程组确定的曲面

设曲面 S 由参数方程 $\begin{cases}x=x(u,v)\\ y=y(u,v),\\ z=z(u,v)\end{cases}\quad (u,v)\in D$ 确定，其中 $D\subset\mathbb{R}^2$ 是一个区域，且三个函数都具有连续的偏导数。记点 $P_0(x_0,y_0,z_0)=(x(u_0,v_0),y(u_0,v_0),z(u_0,v_0))$

取经过该点的两个特殊曲线

$$\begin{cases}x=x(u,v_0)\\ y=y(u,v_0)\\ z=z(u,v_0)\end{cases} \text{和} \begin{cases}x=x(u_0,v)\\ y=y(u_0,v)\\ z=z(u_0,v)\end{cases}$$

则根据 14.8.1.2. 参数方程组确定的曲线 得到其切向量分别为 $(x_u'(u_0,v_0),y_u'(u_0,v_0),z_u'(u_0,v_0))$ 和 $(x_v'(u_0,v_0),y_v'(u_0,v_0),z_v'(u_0,v_0))$ ，当矩阵

$$\begin{pmatrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v}\end{pmatrix}|_{(u_0,v_0)}$$

的秩为 2 时，仿照 14.8.1.3. 两个方程联立确定的曲线 定义

$A=\frac{\partial(y,z)}{\partial(u,v)}|_{(u_0,v_0)}, B=\frac{\partial(z,x)}{\partial(u,v)}|_{(u_0,v_0)}, C=\frac{\partial(x,y)}{\partial(u,v)}|_{(u_0,v_0)}$ ，那么该点处的切平面方程为

$A(x-x_0)+B(y-y_0)+C(z-z_0)=0$ ，法线方程为 $\frac{x-x_0}{A}=\frac{y-y_0}{B}=\frac{z-z_0}{C}$

14.9. 多元凸函数

14.9.1. 凸函数

设凸域 $D\subset\mathbb{R}^n$ ，函数 $f(x)$ 在 D 内有定义，如果 $\forall \boldsymbol{x}_1,\boldsymbol{x}_2\in D, \forall 0<t<1$ 都有 $f(t\boldsymbol{x}_1+(1-t)\boldsymbol{x}_2)\le tf(\boldsymbol{x}_1)+(1-t)f(\boldsymbol{x}_2)$ ，则称函数 $f(x)$ 在 D 内为凸函数，严格小于则称为严格凸函数

14.9.2. 凸函数的等价定义

设凸域 $D \subset \mathbb{R}^n$ ，函数 $f(x)$ 在 D 内具有连续的二阶偏导数，则下列结论等价
(1) $f(x)$ 在 D 内为凸函数
(2) 对 $\forall \boldsymbol{x_0}, \boldsymbol{x} \in D$ 都有 $f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^T$
(3) 对 $\forall \boldsymbol{x}_0 \in D$ ，函数 $f(x)$ 在 $\boldsymbol{x}_0$ 处的 Hessi 矩阵 $H_f(\boldsymbol{x}_0)$ 半正定
下证

- (1) $\Rightarrow$ (2)：将函数作带皮亚诺余项的一阶泰勒展开

$$\begin{aligned} f(t\boldsymbol{x} + (1-t)\boldsymbol{x}_0) &= f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)) \\ &= f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0) \times t(\boldsymbol{x} - \boldsymbol{x}_0)^T + o(t|\boldsymbol{x} - \boldsymbol{x}_0|) \end{aligned}$$

  代入到凸函数的定义得到

$$\begin{aligned} f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0) \times t(\boldsymbol{x} - \boldsymbol{x}_0)^T + o(t|\boldsymbol{x} - \boldsymbol{x}_0|) &\leq t f(\boldsymbol{x}) + (1-t) f(\boldsymbol{x}_0) \\ f'(\boldsymbol{x}_0) \times t(\boldsymbol{x} - \boldsymbol{x}_0)^T + o(t|\boldsymbol{x} - \boldsymbol{x}_0|) &\leq t[f(\boldsymbol{x}) - f(\boldsymbol{x}_0)] \\ f'(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^T + \frac{1}{t} o(t|\boldsymbol{x} - \boldsymbol{x}_0|) &\leq f(\boldsymbol{x}) - f(\boldsymbol{x}_0) \end{aligned}$$

  以 t 为变量的时候 $\lim\limits_{t \to 0^+} \frac{1}{t} o(t|\boldsymbol{x} - \boldsymbol{x}_0|) = \lim\limits_{t \to 0^+} \frac{o(t)}{t} = 0$ ，得证
- (2) $\Rightarrow$ (3)：取点 $\boldsymbol{x}_0 \in D$ 以及增量 $\Delta \boldsymbol{x} \neq \boldsymbol{0} \in \mathbb{R}^n$ ，取充分小的 t 使得 $\boldsymbol{x}_0 + t\Delta \boldsymbol{x} \in D$ ，将函数作带皮亚诺余项的二阶泰勒展开

$$f(\boldsymbol{x}_0 + t\Delta \boldsymbol{x}) = f(\boldsymbol{x}_0) + t f'(\boldsymbol{x}_0) \Delta \boldsymbol{x}^T + \frac{t^2}{2} \Delta \boldsymbol{x} H_f(\boldsymbol{x}_0) \Delta \boldsymbol{x}^T + o(|t^2 \Delta \boldsymbol{x}^2|)$$

  由 (2)的条件得到后两项

$$\begin{aligned} \frac{t^2}{2} \Delta \boldsymbol{x} H_f(\boldsymbol{x}_0) \Delta \boldsymbol{x}^T + o(|t^2 \Delta \boldsymbol{x}^2|) &\geq 0 \\ \frac{1}{2} \Delta \boldsymbol{x} H_f(\boldsymbol{x}_0) \Delta \boldsymbol{x}^T + \frac{1}{t^2} o(|t^2 \Delta \boldsymbol{x}^2|) &\geq 0 \end{aligned}$$

  以 t 为变量的时候 $\lim\limits_{t \to 0^+} \frac{1}{t^2} o(|t^2 \Delta \boldsymbol{x}^2|) = \lim\limits_{t \to 0^+} \frac{o(t^2)}{t^2} = 0$ ，得证
- (3) $\Rightarrow$ (1)： $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in D$ ，令 $\boldsymbol{x}_0 = t\boldsymbol{x}_1 + (1-t)\boldsymbol{x}_2$ ，由凸区域保证了 $\boldsymbol{x}_0 \in D$ ，于是对函数作带拉格朗日余项的一阶泰勒展开

$$f(\boldsymbol{x}_1) = f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x}_1 - \boldsymbol{x}_0)^T + \frac{1}{2}(\boldsymbol{x}_1 - \boldsymbol{x}_0) H_f(\boldsymbol{x}_0 + \theta(\boldsymbol{x_1} - \boldsymbol{x}_0))(\boldsymbol{x}_1 - \boldsymbol{x}_0)^T, 0 < \theta < 1$$

  由(3)的条件得到 $f(\boldsymbol{x}_1) \geq f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x}_1 - \boldsymbol{x}_0)^T$ ，同理有 $f(\boldsymbol{x}_2) \geq f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x}_2 - \boldsymbol{x}_0)^T$ ，两个不等式乘系数相加得到

$$\begin{aligned} t f(\boldsymbol{x}_1) + (1-t) f(\boldsymbol{x}_2) &\geq t[f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x}_1 - \boldsymbol{x}_0)^T] \\ &\quad + (1-t)[f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)(\boldsymbol{x}_2 - \boldsymbol{x}_0)^T] \\ &= f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)[t(\boldsymbol{x}_1 - \boldsymbol{x}_0)^T + (1-t)(\boldsymbol{x}_2 - \boldsymbol{x}_0)^T] \\ &= f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0)[t\boldsymbol{x}_1 + (1-t)\boldsymbol{x}_2 - \boldsymbol{x}_0]^T \\ &= f(\boldsymbol{x}_0) \\ &= f(t\boldsymbol{x}_1 + (1-t)\boldsymbol{x}_2) \end{aligned}$$