微积分 A (2)

姚家燕

第 6 讲

第 5 讲回顾: 复合向量值函数的微分与求导

- 矩阵的范数, 微分的链式法则 (矩阵表示):

$$\mathrm{d}(\vec{f}\circ\vec{g})(X_0)=\mathrm{d}\vec{f}(\vec{g}(X_0))\circ\mathrm{d}\vec{g}(X_0),$$

$$J_{\vec{f}\circ\vec{g}}(X_0)=J_{\vec{f}}(\vec{g}(X_0))\cdot J_{\vec{g}}(X_0),$$

$$\frac{\partial(f_i(g_1,\ldots,g_m))}{\partial x_j}=\frac{\partial f_i}{\partial y_1}(*)\frac{\partial g_1}{\partial x_j}+\frac{\partial f_i}{\partial y_2}(*)\frac{\partial g_2}{\partial x_j}+\cdots+\frac{\partial f_i}{\partial y_m}(*)\frac{\partial g_m}{\partial x_j}.$$

$$\partial_j\big(f_i(g_1,\ldots,g_m)\big)=\partial_1\big(f_i(*)\big)\partial_j g_1+\partial_2\big(f_i(*)\big)\partial_j g_2+\cdots+\partial_m\big(f_i(*)\big)\partial_j g_m.$$

回顾: 隐函数定理

- 隐函数定理 (两个变量的方程):

  设函数 $F(x,y)$ 为 $\mathscr{C}^{(1)}$ 类使得 $F(x_0,y_0)=0$, $\frac{\partial F}{\partial y}(x_0,y_0)\neq 0$. 则方程 $F(x,y)=0$ 在局部上有 $\mathscr{C}^{(1)}$ 类的解 $y=f(x)$, 并且

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x,f(x))}{\frac{\partial F}{\partial y}(x,f(x))}.$$

- 隐函数定理 (多个变量的方程):

假设函数 $F(x_1, x_2, \ldots, x_n, y)$ 为 $\mathscr{C}^{(1)}$ 类使得

$F(X_0, y_0) = 0$, $\frac{\partial F}{\partial y}(X_0, y_0) \neq 0$. 则方程

$$F(x_1, x_2, \ldots, x_n, y) = 0$$

局部上有 $\mathscr{C}^{(1)}$ 类解 $y = f(x_1, x_2, \ldots, x_n)$, 且

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

- 隐函数定理 (多个变量的方程组):

  假设 $F_i(x_1,\ldots,x_n,y_1,\ldots,y_m)$ $(1\leqslant i\leqslant m)$ 为 $\mathscr{C}^{(1)}$ 类使得 $F_i(X_0,Y_0)=0$ $(1\leqslant i\leqslant m)$, $\frac{D(F_1,\ldots,F_m)}{D(y_1,\ldots,y_m)}(X_0,Y_0)\neq 0$. 则方程组

  $F_i(x_1,\ldots,x_n,y_1,\ldots,y_m)=0$ $(1\leqslant i\leqslant m)$

  在局部上有 $\mathscr{C}^{(1)}$ 类解

$$y_i=f_i(x_1,x_2,\ldots,x_n)\ (1\leqslant i\leqslant m),$$

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

回顾: 反函数定理

假设 $X=\vec{g}(Y)$ 为 $\mathscr{C}^{(1)}$ 类使得 $X_0=\vec{g}(Y_0)$ 且 $J_{\vec{g}}(Y_0)$ 可逆. 则局部上存在 $\mathscr{C}^{(1)}$ 反函数 $Y=\vec{f}(X)$, 并且 $J_{\vec{f}}(X)=\left(J_{\vec{g}}(\vec{f}(X))\right)^{-1}$, 即

$$\frac{\partial(f_1,f_2,\ldots,f_n)}{\partial(x_1,x_2,\ldots,x_n)}(X)=\left(\frac{\partial(g_1,g_2,\ldots,g_n)}{\partial(y_1,y_2,\ldots,y_n)}(\vec{f}(X))\right)^{-1}.$$

回顾: 曲面的切平面与法线

- 平面与直线的各种表示方程.
- 平面与法线, 直线与法平面的关系.
- $S: z = f(x, y)$ 在点 (x_0, y_0, z_0) 的切平面:

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

相应的法线方程为

$$\frac{x - x_0}{\frac{\partial f}{\partial x}(x_0, y_0)} = \frac{y - y_0}{\frac{\partial f}{\partial y}(x_0, y_0)} = \frac{z - z_0}{-1}.$$

- $S:\begin{cases} x=f_1(u,v)\\ y=f_2(u,v)\\ z=f_3(u,v)\end{cases}$ 在参数 (u_0,v_0) 所对应的点处的切平面的参数方程为:

$$\begin{pmatrix} x-x_0\\ y-y_0\\ z-z_0\end{pmatrix}=\frac{\partial(f_1,f_2,f_3)}{\partial(u,v)}(u_0,v_0)\begin{pmatrix} u-u_0\\ v-v_0\end{pmatrix},$$

该切平面也可以表示成:

$$\frac{D(f_2,f_3)}{D(u,v)}(u_0,v_0)(x-x_0)+\frac{D(f_3,f_1)}{D(u,v)}(u_0,v_0)(y-y_0)$$
$$+\frac{D(f_1,f_2)}{D(u,v)}(u_0,v_0)(z-z_0)=0,$$

法线方程为 $\frac{x-x_0}{\frac{D(f_2,f_3)}{D(u,v)}(u_0,v_0)}=\frac{y-y_0}{\frac{D(f_3,f_1)}{D(u,v)}(u_0,v_0)}=\frac{z-z_0}{\frac{D(f_1,f_2)}{D(u,v)}(u_0,v_0)}$.

- $S: F(x,y,z)=0$ 在点 P_0 处的法向量为

  梯度 $\mathrm{grad}F(P_0)$, 相应的切平面方程为

$$\frac{\partial F}{\partial x}(P_0)(x-x_0)+\frac{\partial F}{\partial y}(P_0)(y-y_0)+\frac{\partial F}{\partial z}(P_0)(z-z_0)=0,$$

  上述切平面的法线方程为

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

回顾：空间曲线及切线和法平面

(1) 空间曲线的参数表示法:

$$\Gamma: \begin{cases} x = x(t), \\ y = y(t), \quad t \in [\alpha, \beta]. \\ z = z(t), \end{cases}$$

若上述函数在点 $t = t_0$ 处可微, 则称曲线 Γ 在相应点 $P_0(x_0, y_0, z_0)$ 处可微, 相应切线方程为

$$\begin{cases} x - x_0 = x'(t_0)(t - t_0), \\ y - y_0 = y'(t_0)(t - t_0), \\ z - z_0 = z'(t_0)(t - t_0). \end{cases}$$

该切线也可表述成

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

这里要假设 $(x'(t_0), y'(t_0), z'(t_0))$ 不为零向量.

我们将过点 P_0 且与上述切线垂直的平面称为 Γ 在点 P_0 处的法平面, 其方程为

$$x'(t_0)(x - x_0) + y'(t_0)(y - y_0) + z'(t_0)(z - z_0) = 0.$$

(2) 空间曲线的隐函数表示法:

$$\Gamma: \begin{cases} F_1(x,y,z)=0, \\ F_2(x,y,z)=0. \end{cases}$$

设 F_1, F_2 在点 $P_0(x_0, y_0, z_0)$ 可微且 $\text{grad}F_1(P_0)$, $\text{grad}F_2(P_0)$ 不为零, 则曲线 Γ 在该点的切线为

$$\begin{cases} \frac{\partial F_1}{\partial x}(P_0)(x-x_0)+\frac{\partial F_1}{\partial y}(P_0)(y-y_0)+\frac{\partial F_1}{\partial z}(P_0)(z-z_0)=0, \\ \\ \frac{\partial F_2}{\partial x}(P_0)(x-x_0)+\frac{\partial F_2}{\partial y}(P_0)(y-y_0)+\frac{\partial F_2}{\partial z}(P_0)(z-z_0)=0. \end{cases}$$

该切线的方向为

$$\vec{T} = \mathrm{grad} F_1(P_0) \times \mathrm{grad} F_2(P_0) = \begin{pmatrix} \frac{D(F_1,F_2)}{D(y,z)}(P_0) \\ \frac{D(F_1,F_2)}{D(z,x)}(P_0) \\ \frac{D(F_1,F_2)}{D(x,y)}(P_0) \end{pmatrix}.$$

只有当 $\vec{T} \neq \vec{0}$ 时, 上述方程组才的确给出一条直线. 此时 Jacobi 矩阵 $\frac{\partial(F_1,F_2)}{\partial(x,y,z)}(P_0)$ 的秩等于 2. 借助 $\vec{T}$, 我们也可得到切线的另外一个表述:

$$\frac{x-x_0}{\frac{D(F_1,F_2)}{D(y,z)}(P_0)} = \frac{y-y_0}{\frac{D(F_1,F_2)}{D(z,x)}(P_0)} = \frac{z-z_0}{\frac{D(F_1,F_2)}{D(x,y)}(P_0)}.$$

第 6 讲

§8. Taylor 公式

回顾: 我们称函数 $F:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ 为双线性型, 如果 $\forall X_0, Y_0\in\mathbb{R}^n$, 函数 $Y\to F(X_0,Y)$ 和函数 $X\to F(X,Y_0)$ 均为 $\mathbb{R}^n$ 上的线性函数.

记 $X=(x_1,\ldots,x_n)^T$, $Y=(y_1,\ldots,y_n)^T$. 则

$$F(X,Y)=F\Big(\sum_{i=1}^n x_i\mathbf{e}_i,Y\Big)=\sum_{i=1}^n x_iF(\mathbf{e}_i,Y)$$

$$=\sum_{i=1}^n x_iF\Big(\mathbf{e}_i,\sum_{j=1}^n y_j\mathbf{e}_j\Big)=\sum_{i=1}^n\sum_{j=1}^n x_iy_jF(\mathbf{e}_i,\mathbf{e}_j).$$

令 $a_{ij}=F(\mathbf{e}_i,\mathbf{e}_j)$, $A=(a_{ij})_{1\leqslant i,j\leqslant n}$, 则

$$\begin{aligned} F(X,Y) &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j F(\mathbf{e}_i,\mathbf{e}_j) \\ &= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} y_j \\ &= \sum_{i=1}^{n} x_i (AY)_i \\ &= X^T AY. \end{aligned}$$

定理 1. 设 $X_0 \in \mathbb{R}^n$, $r > 0$, $f \in \mathscr{C}^{(2)}(B(X_0, r))$. 则 $\forall X \in B(X_0, r)$, $\exists \theta \in (0, 1)$ 使得

$$
\begin{aligned}
f(X) = f(X_0) &+ \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(X_0)(x_j - x_j^{(0)}) \\
&+ \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_\theta)(x_i - x_i^{(0)})(x_j - x_j^{(0)}) \\
&= f(X_0) + J_f(X_0)\,\Delta X + \frac{1}{2!}(\Delta X)^T H_f(X_\theta)\,\Delta X,
\end{aligned}
$$

其中 $H_f(X_\theta) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(X_\theta)\right)_{1 \leqslant i,j \leqslant n}$ 为 f 在点 X_θ 的海赛矩阵, $\Delta X = X - X_0$, $X_\theta = X_0 + \theta(X - X_0)$.

证明: 固定 $X \in \mathring{B}(X_0, r)$. $\forall t \in [0,1]$, 定义

$$F(t) = f(X_0 + t(X - X_0)).$$

由于 f 为二阶连续可导, 则由复合法则可知 F 亦如此, 进而由单变量函数带 Lagrange 余项的 Taylor 展式可知, $\exists \theta \in (0,1)$ 使得

$$F(1) = F(0) + F'(0) + \frac{1}{2!}F''(\theta).$$

$\forall t \in [0,1]$, 令 $X_t = X_0 + t(X - X_0)$.

由题设可知 $F(0)=f(X_0)$, $F(1)=f(X)$, 并且

$$
\begin{aligned}
F'(t) &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(X_t)(x_j - x_j^{(0)}) = J_f(X_t)(X - X_0). \\
F''(t) &= \sum_{j=1}^{n} (x_j - x_j^{(0)}) \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial x_j}(X_t)\right) \\
&= \sum_{j=1}^{n} (x_j - x_j^{(0)}) \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(X_t) \cdot (x_i - x_i^{(0)}) \\
&= \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t) \cdot (x_i - x_i^{(0)})(x_j - x_j^{(0)}).
\end{aligned}
$$

令 $\Delta X = X - X_0$. 则我们有

$$\begin{aligned} F(1) &= F(0) + F'(0) + \frac{1}{2!}F''(\theta) \\ &= f(X_0) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(X_0)(x_j - x_j^{(0)}) \\ &\quad + \frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_\theta) \cdot (x_i - x_i^{(0)})(x_j - x_j^{(0)}) \\ &= f(X_0) + J_f(X_0)\,\Delta X + \frac{1}{2!}(\Delta X)^T H_f(X_\theta)\,\Delta X. \end{aligned}$$

故所证结论成立.

评注

- 该式为带 Lagrange 余项的一阶 Taylor 展式.
- 由于 f 为 $\mathscr{C}^{(2)}$ 类, 则 H_f 连续. 由夹逼原理与复合极限法则知, 当 $X \to X_0$ 时, 我们有

$$H_f(X_0 + \theta(X - X_0)) = H_f(X_0) + \vec{o}(1).$$

进而可得带 Peano 余项的二阶 Taylor 展式:

$$\begin{aligned} f(X) &= f(X_0) + J_f(X_0)\,\Delta X \\ &\quad + \frac{1}{2!}(\Delta X)^T H_f(X_0)\Delta X + o(\|\Delta X\|^2). \end{aligned}$$

- $H_f(X) = J_{\mathrm{grad}f}(X)$.
- 更一般地, 若 f 为 $\mathscr{C}^{(m+1)}$ 类, 则有

$$f(X) = \sum_{k=0}^{m} \frac{1}{k!}\Big(\sum_{j=1}^{n}(x_j - x_j^{(0)})\frac{\partial}{\partial x_j}\Big)^k f(X_0)$$

$$+\frac{1}{(m+1)!}\Big(\sum_{j=1}^{n}(x_j - x_j^{(0)})\frac{\partial}{\partial x_j}\Big)^{m+1} f(X_\theta),$$

其中 $\theta \in (0,1)$, 并且 $X_\theta = X_0 + \theta(X - X_0)$.
人们通常将 $\sum\limits_{k=0}^{m} \frac{1}{k!}\Big(\sum\limits_{j=1}^{n}(x_j - x_j^{(0)})\frac{\partial}{\partial x_j}\Big)^k f(X_0)$
称为 f 在点 X_0 处的 m 阶 Taylor 多项式.

例 1. $\forall (x,y)\in\mathbb{R}^2$, 令 $f(x,y)=\sin(x+y)$. 求 f 在原点处一阶的带 Lagrange 余项的 Taylor 公式以及二阶的带 Peano 余项的 Taylor 公式.

解: 由于 f 为初等函数, 故二阶连续可导且

$$\begin{aligned} J_f(x,y) &= (\cos(x+y), \cos(x+y)), \\ H_f(x,y) &= \begin{pmatrix} -\sin(x+y) & -\sin(x+y) \\ -\sin(x+y) & -\sin(x+y) \end{pmatrix}, \end{aligned}$$

于是 $J_f(0,0)=(1,1)$, $H_f(0,0)=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

则所求一阶带 Lagrange 余项的 Taylor 公式为

$$\begin{aligned} f(x,y) &= x+y+\frac{1}{2}(x,y)H_f(\theta x,\theta y)\begin{pmatrix} x \\ y \end{pmatrix} \\ &= x+y-\frac{1}{2}(x+y)^2\sin(\theta(x+y)),\ \ \theta\in(0,1). \end{aligned}$$

而所求二阶带 Peano 余项的 Taylor 公式为

$$f(x,y)=x+y+o(x^2+y^2),\quad (x,y)\to(0,0).$$

作业题: 第 1.8 节第 81 页第 1 第 (2) 小题, 题目改为求 "一阶带 Lagrange 余项和二阶带 Peano 余项的 Taylor 公式".

§9. 极值与条件极值

定义 1. 设 $\Omega \subseteq \mathbb{R}^n$, $X_0 \in \Omega$, 而 $f : \Omega \to \mathbb{R}$.

- 如果 $\exists r > 0$ 使得 $\forall X \in B(X_0, r) \subseteq \Omega$, 均有 $f(X) \geqslant f(X_0)$, 那么称点 X_0 为 f 的 (局部) 极小值点, 而称 $f(X_0)$ 为 (局部) 极小值.
- 如果 $\exists r > 0$ 使得 $\forall X \in B(X_0, r) \subseteq \Omega$, 均有 $f(X) \leqslant f(X_0)$, 那么称点 X_0 为 f 的 (局部) 极大值点, 而称 $f(X_0)$ 为 (局部) 极大值.
- 极小值点和极大值点合称极值点.

- 若 $\forall X \in \Omega$, 均有 $f(X) \geqslant f(X_0)$, 则称点 X_0 为 f 的最小值点, 而称 $f(X_0)$ 为最小值.
- 若 $\forall X \in \Omega$, 均有 $f(X) \leqslant f(X_0)$, 则称点 X_0 为 f 的最大值点, 而称 $f(X_0)$ 为最大值.
- 最小值点和最大值点合称最值点.

注: 极值点不一定是最值点, 而最值点也不一定是极值点. 若最值点为内点, 则它为极值点.

定理 1. 假设 $\Omega \subseteq \mathbb{R}^n$, X_0 为 Ω 的内点, 而函数 $f:\Omega \to \mathbb{R}$ 在该点可导. 若 X_0 为 f 的极值点, 则 $J_f(X_0) = \vec{0}$, 即 $\frac{\partial f}{\partial x_j}(X_0) = 0\ (1 \leqslant j \leqslant n)$.

证明: 由于 X_0 为 Ω 的内点, 于是 $\exists r > 0$ 使得 $B(X_0, r) \subset \Omega$. 对任意整数 $1 \leqslant j \leqslant n$, 设 $\vec{e}_j$ 为沿第 j 个坐标轴正向的单位向量. $\forall t \in (-r, r)$, 定义 $F(t) = f(X_0 + t\vec{e}_j)$. 则由题设可知函数 F 在点 $t = 0$ 处可导, 并且 F 还在该点处取极值, 从而由 Fermat 定理可知 $0 = F'(0) = \frac{\partial f}{\partial x_j}(X_0)$.

评注

- 若 $J_f(X_0) = \vec{0}$, 则称点 X_0 为函数 f 的驻点. 由上述定理可知, 若 f 在点 X_0 处可导且取极值, 则该点为 f 的驻点, 但反过来不成立.
- 利用类似的方法可以证明, 多变量函数也有与单变量函数一样的微分中值定理. 下面来介绍高维的 Rolle 定理.

定理 2. 设 $\Omega \subseteq \mathbb{R}^n$ 为有界开区域, 而 $f \in \mathscr{C}(\bar{\Omega})$ 在 Ω 内可微并且在 $\partial\Omega$ 上取常值, 则 f 在 Ω 内必有驻点, 也即 $\exists \xi \in \Omega$ 使得 $J_f(\xi) = \vec{0}$.

证明: 如果 f 在 $\bar{\Omega}$ 上为常值函数, 则 f 在 Ω 内也会为常值函数, 从而 $\forall \xi \in \Omega$, 均有 $J_f(\xi) = \vec{0}$. 现在假设 f 不为常值函数. 由于 $\bar{\Omega}$ 为有界闭集且 $f \in \mathscr{C}(\bar{\Omega})$, 则 f 在 $\bar{\Omega}$ 上有最大值和最小值, 但 f 在 $\partial\Omega$ 上取常值, 故必有最值点 (记作 ξ) 属于 Ω, 该点也为 f 的极值点, 则 $J_f(\xi) = \vec{0}$.

回顾: 设 $A=(a_{ij})_{1\leqslant i,j\leqslant n}$ 为实对称矩阵, 特征根为 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$, 则存在正交矩阵 B 使得

$$A=B^T\begin{pmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \vdots & \vdots & \vdots & 0\\ 0 & 0 & \dots & \lambda_n\end{pmatrix}B.$$

$\forall X\in\mathbb{R}^n$, 记 $BX=(y_1,\dots,y_n)^T$. 则

$$\begin{aligned} X^TAX &= \sum_{j=1}^n \lambda_j y_j^2 \geqslant \lambda_1\|Y\|^2=\lambda_1 Y^TY\\ &= \lambda_1 X^TB^TBX=\lambda_1X^TX=\lambda_1\|X\|^2. \end{aligned}$$

同理可以证明 $X^TAX\leqslant\lambda_n\|X\|^2$.

- 称矩阵 A 为正定, 如果 $\forall X \in \mathbb{R}^n \setminus \{\vec{0}\}$, 我们均有 $X^TAX > 0$, 这等价于说 $\lambda_1 > 0$, 进而也等价于说 A 的所有主子式均 > 0.
- 称矩阵 A 为负定, 如果 $\forall X \in \mathbb{R}^n \setminus \{\vec{0}\}$, 我们均有 $X^TAX < 0$, 这等价于说 $\lambda_n < 0$, 进而等价于说 A 的所有偶数阶顺序主子式 > 0, 而所有奇数阶顺序主子式 < 0.

- 称矩阵 A 半正定, 如果 $\forall X \in \mathbb{R}^n \setminus \{\vec{0}\}$, 均有 $X^T AX \geqslant 0$, 而这又等价于说 $\lambda_1 \geqslant 0$.
- 称矩阵 A 半负定, 如果 $\forall X \in \mathbb{R}^n \setminus \{\vec{0}\}$, 均有 $X^T AX \leqslant 0$, 而这又等价于说 $\lambda_n \leqslant 0$.
- 称矩阵 A 不定, 若它不属于上述的四种情形, 即 $\exists Y_1, Y_2 \in \mathbb{R}^n$ 使 $Y_1^T AY_1 < 0$, $Y_2^T AY_2 > 0$, 而这又等价于说 $\lambda_1 < 0$, $\lambda_n > 0$.

判断实对称矩阵定性的基本方法

计算其特征值，根据其特征值的符号来判断:

(1) 正定当且仅当所有特征值均 > 0;

(2) 负定当且仅当所有特征值均 < 0;

(3) 半正定当且仅当所有特征值均 $\geqslant 0$;

(4) 半负定当且仅当所有特征值均 $\leqslant 0$;

(5) 不定当且仅当既有正特征值也有负特征值.

定理 3. 设 $X_0 \in \mathbb{R}^n$, $r > 0$, 而 $f: B(X_0, r) \to \mathbb{R}$ 为二阶连续可微且 $J_f(X_0) = \vec{0}$.

(1) 若 $H_f(X_0)$ 正定, 则 X_0 为 f 的极小值点.

(2) 若 $H_f(X_0)$ 负定, 则 X_0 为 f 的极大值点.

证明: (1) 由 Taylor 公式可知, 当 $\Delta X \to \vec{0}$ 时,

$$\begin{aligned} f(X_0 + \Delta X) = {} & f(X_0) + J_f(X_0)\,\Delta X \\ & + \frac{1}{2!}(\Delta X)^T H_f(X_0)\Delta X + o(\|\Delta X\|^2). \end{aligned}$$

设 λ 为 $H_f(X_0)$ 的最小特征值. 则由 $H_f(X_0)$ 的正定性可知 $\lambda > 0$. 于是我们有

$$
\begin{aligned}
& f(X_0 + \Delta X) - f(X_0) \\
= \ & \frac{1}{2!}(\Delta X)^T H_f(X_0)\Delta X + o(\|\Delta X\|^2) \\
\geqslant \ & \frac{\lambda}{2}\|\Delta X\|^2 + o(\|\Delta X\|^2) \\
= \ & \Big(\frac{\lambda}{2} + o(1)\Big)\|\Delta X\|^2.
\end{aligned}
$$

因 $\lim\limits_{\Delta X\to\vec{0}} o(1)=0$, 则 $\exists\delta\in(0,r)$ 使 $\forall X\in\mathring{B}(X_0,\delta)$,

$$
\begin{aligned}
f(X)-f(X_0) &\geqslant (\frac{\lambda}{2}+o(1))\|X-X_0\|^2\\
&\geqslant \frac{\lambda}{4}\|X-X_0\|^2\geqslant 0,
\end{aligned}
$$

故 X_0 为函数 f 的 (严格) 极小值点.

(2) 如果 $H_f(X_0)$ 为负定, 那么 $H_{-f}(X_0)$ 为正定,

故 X_0 为 $-f$ 的极小值点, 即为 f 的极大值点.

评注

- 如果海赛矩阵 $H_f(X_0)$ 为半正定或者半负定, 我们一般无法判断点 X_0 是否为 f 的极值点. 为此考虑单变量函数 $f(x)=x^3$, 我们有

$$f'(0)=f''(0)=0,$$

但 0 不是 f 的极值点. 而 $g(x)=x^4$ 也满足 $g'(0)=g''(0)=0$ 且 0 为 g 的极小值点.

- 如果 $H_f(X_0)$ 为不定, 则 $\exists Y_1, Y_2 \in \mathbb{R}^n$ 使得

$$Y_1^T H_f(X_0) Y_1 < 0, \quad Y_2^T H_f(X_0) Y_2 > 0.$$

于是当 $t \to 0$ 时, 我们有

$$\begin{aligned} & f(X_0 + tY_1) - f(X_0) \\ = & \frac{1}{2} Y_1^T H_f(X_0) Y_1 t^2 + o(t^2) \\ = & \Big(\frac{1}{2} Y_1^T H_f(X_0) Y_1 + o(1)\Big) t^2 < 0. \end{aligned}$$

当 $t \to 0$ 时, 我们同样也有

$$
\begin{aligned}
& f(X_0 + tY_2) - f(X_0) \\
= \;& \frac{1}{2} Y_2^T H_f(X_0) Y_2 t^2 + o(t^2) \\
= \;& \Big(\frac{1}{2} Y_2^T H_f(X_0) Y_2 + o(1)\Big) t^2 \\
> \;& 0,
\end{aligned}
$$

由此可知 X_0 不是 f 的极值点.

判断二阶连续可导函数极值点的方法

- 求一阶偏导数, 确定驻点.
- 求二阶偏导数以便得到海赛矩阵.
- 判断海赛矩阵在驻点处的性态:

  正定则为极小值点; 负定则为极大值点; 不定则不为极值点; 半正定或半负定则需要采用另外的方法来处理.

例 1. $\forall (x,y,z)\in\mathbb{R}^3$, 定义

$$f(x,y,z)=x^3+y^2+z^2+6xy+2z.$$

求 f 的极值点.

解: 由于 f 为初等函数, 故 f 为 $\mathscr{C}^{(2)}$ 类且

$$J_f(x,y,z)=(3x^2+6y,2y+6x,2z+2).$$

于是函数 f 的驻点满足:

$$3x^2+6y=0,\ 2y+6x=0,\ 2z+2=0,$$

由此求得驻点为

$$P_1 = (6, -18, -1),\ P_2 = (0, 0, -1).$$

又函数 f 的海赛矩阵为

$$H_f(x, y, z) = \begin{pmatrix} 6x & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

于是我们有

$$H_f(P_1) = \begin{pmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},\ H_f(P_2) = \begin{pmatrix} 0 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

$H_f(P_1)$ 的顺序主子式均 >0, 故 $H_f(P_1)$ 正定, 从而 P_1 为函数 f 的极小值点. 由于 $H_f(P_2)$ 的特征多项式为

$$0=\begin{vmatrix} \lambda & -6 & 0 \\ -6 & \lambda-2 & 0 \\ 0 & 0 & \lambda-2 \end{vmatrix}=(\lambda-2)(\lambda^2-2\lambda-36),$$

于是 $H_f(P_2)$ 的特征根为 $2, 1+\sqrt{37}, 1-\sqrt{37}$, 故 $H_f(P_2)$ 为不定, 从而 P_2 不是 f 的极值点.

作业题: 第 1.9 节第 93 页第 1 题第 (2), (4) 题

例 2. 设隐函数 $z=z(x,y)$ 由方程

$$F(x,y,z)=2x^2+2y^2+z^2+8xz-z+8=0$$

确定, 求其极值点.

解: 由隐函数定理可知, $z(x,y)$ 的驻点满足

$$0=\frac{\partial z}{\partial x}=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}=-\frac{4x+8z}{2z+8x-1},$$

$$0=\frac{\partial z}{\partial y}=-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}=-\frac{4y}{2z+8x-1}.$$

于是 $y=0$, $x=-2z$. 带入隐函数方程可得

$$\begin{cases} x_1 = \frac{16}{7} \\ y_1 = 0 \\ z_1 = -\frac{8}{7} \end{cases}, \quad \begin{cases} x_2 = -2 \\ y_2 = 0 \\ z_2 = 1 \end{cases}.$$

在驻点处, 我们有

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\Big(-\frac{4x+8z}{2z+8x-1}\Big)$$
$$= -\frac{4(2z+8x-1)-8(4x+8z)}{(2z+8x-1)^2} = \frac{4}{14z+1}.$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\Big(-\frac{4x+8z}{2z+8x-1}\Big) = 0.$$

同样我们也有

$$\frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y}\Big(-\frac{4y}{2z+8x-1}\Big) = -\frac{4}{2z+8x-1} = \frac{4}{14z+1}.$$

由此立刻可得

$$H_z\Big(\frac{16}{7}, 0\Big) = \begin{pmatrix} -\frac{4}{15} & 0 \\ 0 & -\frac{4}{15} \end{pmatrix}, \quad H_z\Big(-2, 0\Big) = \begin{pmatrix} \frac{4}{15} & 0 \\ 0 & \frac{4}{15} \end{pmatrix}.$$

则 $(\frac{16}{7}, 0)$ 为极大值点, 而 $(-2, 0)$ 为极小值点.

作业题: 求由方程 $x^2 + \frac{y^2}{4} + \frac{z^2}{9} - \frac{2}{3}z = 0$ 所确定的隐函数 $z = z(x, y)$ 的极值.

谢谢大家!