第六章 多元函数的极值问题

§4.1 普通极值问题

设 $f(x_1,\cdots,x_n)$ 是集合 $S\subset\mathbf{R}^n$ 上的函数，如果对 $P_0=(x_1^0,\cdots,x_n^0)$，存在 P_0 在 $\mathbf{R}^n$ 中的邻域 U，使得 $\forall P=(x_1,\cdots,x_n)\in S\bigcap U$，恒有 $f(x_1,\cdots,x_n)\le f(x_1^0,\cdots,x_n^0)$ ($f(x_1,\cdots,x_n)\ge f(x_1^0,\cdots,x_n^0)$)，则 $f(x_1^0,\cdots,x_n^0)$ 称为 $f(x_1,\cdots,x_n)$ 在 S 上的局部极大值（极小值），P_0 称为 $f(x_1,\cdots,x_n)$ 的局部极大值（极小值）点. 如果 S 是开集，则 P_0 称为普通极值点. 否则称为条件极值点.

定理 1: 如果 $P_0=(x_1^0,\cdots,x_n^0)$ 是 $f(x_1,\cdots,x_n)$ 的普通极值点，且 $f(x_1,\cdots,x_n)$ 在 P_0 存在偏导，则 $\dfrac{\partial f(x_1^0,\cdots,x_n^0)}{\partial x_i}=0,\ i=1,\cdots,n.$.

证明：P_0 是内点，因而 x_1^0 是一元函数 $f(x_1,x_2^0,\cdots,x_n^0)$ 的极值点. 因此 $\dfrac{\partial f(x_1^0,\cdots,x_n^0)}{\partial x_1}=0$.

定义： 设 $f(x_1,\cdots,x_n)$ 在区域 D 上处处有偏导. 如果在点 $P_0=(x_1^0,\cdots,x_n^0)$ 成立 $\dfrac{\partial f(x_1^0,x_2^0,\cdots,x_n^0)}{\partial x_i}=0,\ i=1,\cdots,n$，则称 P_0 为 $f(x_1,\cdots,x_n)$ 的判别点.

如果 P_0 是 $f(x_1,\cdots,x_n)$ 的极值点，则其是 $f(x_1,\cdots,x_n)$ 的判别点. 但反之并不成立.

例：令 $f(x,y)=x^2-y^2$，则 $\dfrac{\partial f(0,0)}{\partial x}=0,\dfrac{\partial f(0,0)}{\partial y}=0$. 但 $(0,0)$ 并不是 $f(x,y)$ 的极值点.

与一元函数相同，我们需要利用 $f(x_1,\cdots,x_n)$ 在判别点处的二阶 Taylor 展开来讨论所给判别点是否极值点以及是什么样的极值点. 为此我们需要下面引理.

引理：设 n 阶对称矩阵 A 是正定（负定）的，则存在 $e>0$，使得对任意 $(x_1,\cdots,x_n)$，恒

有

$(x_1,\cdots,x_n)A(x_1,\cdots,x_n)^t \ge e\left(x_1^2+\cdots+x_n^2\right)\left((x_1,\cdots,x_n)A(x_1,\cdots,x_n)^t \le -e\left(x_1^2+\cdots+x_n^2\right)\right)$.

证明： $\mathbf{R}^n$ 中单位球面 $S_n=\left\{(x_1,\cdots,x_n)\middle| x_1^2+\cdots+x_n^2=1\right\}$是有界闭集，因而是紧集. S_n 上的函数$(x_1,\cdots,x_n)A(x_1,\cdots,x_n)^t$连续且处处不为零，因而在$S_n$上达到最小值，设为$e$. 则对任意$(x_1,\cdots,x_n)\ne 0$，恒有

$$\frac{(x_1,\cdots,x_n)}{\sqrt{x_1^2+\cdots+x_n^2}}A\frac{(x_1,\cdots,x_n)^t}{\sqrt{x_1^2+\cdots+x_n^2}} \ge e\ .$$

引理得证.

定理 2： 设$P_0=(x_1^0,\cdots,x_n^0)$是$f(x_1,\cdots,x_n)$在区域D内的判别点. 如果$f(x_1,\cdots,x_n)$在P_0的Hessi矩阵$H_f(P_0)$是正定的，则P_0是$f(x_1,\cdots,x_n)$的严格极小点；如果$H_f(P_0)$是负定的，则P_0是$f(x_1,\cdots,x_n)$的严格极大点；如果$H_f(P_0)$是不定的，则P_0不是$f(x_1,\cdots,x_n)$的极值点.

证明： 设$H_f(P_0)=\left(\dfrac{\partial^2 f(x_1^0,\cdots,x_n^0)}{\partial x_i \partial x_j}\right)$正定，取$e>0$满足上面引理. 将$f$在$P_0$点作二阶 Taylor 展开，由$(x_1^0,\cdots,x_n^0)$是判别点得

$$\begin{aligned}
&f(x_1,\cdots,x_n)-f(x_1^0,\cdots,x_n^0)=\frac{1}{2}d^2 f(x_1^0,\cdots,x_n^0)+o\left(\sum_{i=1}^n (x_i-x_i^0)^2\right)\\
&=\frac{1}{2}\left(x_1-x_1^0,\cdots,x_n-x_n^0\right)H_f(P_0)\left(x_1-x_1^0,\cdots,x_n-x_n^0\right)^t+o\left(\sum_{i=1}^n (x_i-x_i^0)^2\right)\\
&\ge\left(\frac{e}{2}+o\right)\left(\sum_{i=1}^n (x_i-x_i^0)^2\right)
\end{aligned}$$

由于o在$(x_1,\cdots,x_n)$趋于$(x_1^0,\cdots,x_n^0)$时是无穷小，因此存在$(x_1^0,\cdots,x_n^0)$的邻域U，使得$(x_1,\cdots,x_n)\in U$时$\dfrac{e}{2}+o>0$，得$f(x_1,\cdots,x_n)>f(x_1^0,\cdots,x_n^0)$. P_0是$f(x_1,\cdots,x_n)$的严格极小点.

如果$H_f(P_0)$不定，则存在 n 维向量$a\ne 0$和$b\ne 0$，使得$aH_f(P_0)a^t>0$，而$bH_f(P_0)b^t<0$. 令$P_1=P_0+ta$, $P_2=P_0+tb$，则t充分小时P_1和P_2都在D内，且

$$f(P_1)-f(P_0)=t^2aH_f(P_0)a^t+o\cdot t^2\left(\|a\|^2\right)$$
$$=t^2\left(aH_f(P_0)a^t+o\left(\|a\|^2\right)\right).$$

t 充分小时，$aH_f(P_0)a^t+o\left(\|a\|^2\right)>0$，因此 P_0 不是极大点．同理

$$f(P_2)-f(P_0)=t^2\left(bH_f(P_0)b^t+o\left(\|b\|^2\right)\right).$$

在 t 充分小时小于零，因此 P_0 不是极小点．得 P_0 不是极值点.

由定理 2，仅当判别点 $P_0=(x_1^0,\cdots,x_n^0)$ 的 Hessi 矩阵 $H_f(P_0)$ 是半正定或半负定时，我们不能判定 P_0 是否是 f 的极小或极大点，这是必须 $\det H_f(P_0)=0$.

利用定理 2，我们需要判定 Hessi 矩阵 $H_f(P_0)$ 的正定或负定性．对此我们需要下面《线性代数》中给出的定理.

定理 3：对称矩阵 $A=(a_{ij})_{i,j=1}^n$ 是正定（负定）的充分必要条件是 A 的主子行列式

$$\det(a_{ij})_{i,j=1}^k>0 \qquad \left((-1)^k\det(a_{ij})_{i,j=1}^k>0\right)$$

对 $k=1,\cdots,n$ 成立.

例：求 $f(x,y)=x^2-xy+y^2-2x+y$ 的极值点.

解：$\dfrac{\partial f(x,y)}{\partial x}=2x-y-2,\quad \dfrac{\partial f(x,y)}{\partial y}=2y-x+1$．解得判别点为 $P_0=(1,0)$，而

$$H_f(P_0)=\begin{pmatrix}\dfrac{\partial^2 f(P_0)}{\partial x^2} & \dfrac{\partial^2 f(P_0)}{\partial x\partial y}\\ \dfrac{\partial^2 f(P_0)}{\partial x\partial y} & \dfrac{\partial^2 f(P_0)}{\partial y^2}\end{pmatrix}=\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix}.$$

一阶主子行列式 $(a_{11})=2>0$，二阶主子行列式 $\begin{vmatrix}2 & -1\\ -1 & 2\end{vmatrix}=3>0$，因此 $H_f(P_0)$ 正定，$P_0=(1,0)$ 是 $f(x,y)$ 的极小点，$f(1,0)=-1$ 是 $f(x,y)$ 的极小值.

§4.2 条件极值问题

设 $F_1(x_1,\cdots,x_n),\cdots,F_m(x_1,\cdots,x_n)$ 是区域 $D\subset\mathbf{R}^n$ 上的函数．定义 S 为

$$S=\{(x_1,\cdots,x_n)\,|\,F_1(x_1,\cdots,x_n)=\cdots=F_m(x_1,\cdots,x_n)=0\}.$$

设 $f(x_1,\cdots,x_n)$ 是 S 上的函数，$f(x_1,\cdots,x_n)$ 在 S 上的极值问题称为 $f(x_1,\cdots,x_n)$ 对约束条件 $F_1(x_1,\cdots,x_n)=0,\cdots,F_m(x_1,\cdots,x_n)=0$ 的条件极值．即 $P_0=(x_1^0,\cdots,x_n^0)\in S$ 称为 $f(x_1,\cdots,x_n)$ 对约束条件 $F_1(x_1,\cdots,x_n)=0,\cdots,F_m(x_1,\cdots,x_n)=0$ 的条件极大（极小）值．如果存在 P_0 在 $\mathbf{R}^n$ 中的一个邻域 U，使得 $\forall P=(x_1,\cdots,x_n)\in S\bigcap U$，恒有 $f(x_1,\cdots,x_n)\le f(x_1^0,\cdots,x_n^0)$ （$f(x_1,\cdots,x_n)\ge f(x_1^0,\cdots,x_n^0)$）．

例：证明在所有周长相同的三角形中，等边三角形面积最大．

证明：设三角形三条边为 x,y,z．周长固定时 x,y,z 并非自由变量，受条件 $x+y+z=2p$ 的约束，其中 p 为常数．由面积公式知，三角形面积为

$$s=\sqrt{p(p-x)(p-y)(p-z)}\,.$$

因此我们需求 $f(x,y,z)=(p-x)(p-y)(p-z)$ 在条件 $x+y+z=2p$ 下的最大值．

由 $x+y+z=2p$ 解出 $z=2p-x-y$．这时 x,y 是自由变量，在一个开集上变化．代入 $f(x,y,z)$，条件极值问题化为

$$h(x,y)=(p-x)(p-y)(x+y-p)$$

的普通极值问题．

解方程组

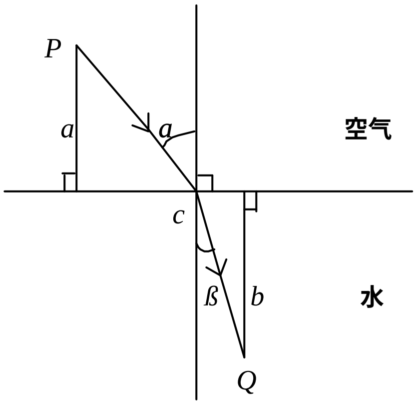
$$\begin{cases}h_x(x,y)=(p-y)(2p-2x-y)=0\\ h_y(x,y)=(p-x)(2p-x-2y)=0,\end{cases}$$

得在 $x>0,\ y>0$ 时只有一个解 $x=\dfrac{2}{3}p,\ y=\dfrac{2}{3}p$．但由问题知，最大值存在，而判别点唯一．因此判别点只能是最大点，得 $x=\dfrac{2}{3}p,\ y=\dfrac{2}{3}p,\ z=\dfrac{2}{3}p$ 时三角形面积最大．

上例中我们是通过求解约束条件的方程，得到自由变量，代入求极值的函数，将条件极值问题化为普通极值问题．但有时直接解约束条件的方程是困难的，但通过微分约束条件，解出某些变量的微分用另一些变量的微分来表示，再代入求极值函数的微分中，从而求得其在约束条件下的判别点．

例：设光在空气和水中的速度分别为 v_1, v_2，设光由空气射入水中时入射角为 α，折射角为 β，假定光总在最短时间由一个点运动到另一点. 证明

$$\frac{\sin\alpha}{v_1}=\frac{\sin\beta}{v_2}.$$



证明：如图. 设 P 为光源，Q 为光到达的点，a,b 为 P,Q 到水面的距离. 设 h 为 P,Q 到水面投影点之间的距离，c 是水面上任取的一个光的入射点. 问题是 c 在什么位置时，光由 P 到 Q 所需时间最短.

光在空气和水中总是以直线传播，因此 c 点确定后，入射角 α 和折射角 β 因而确定，这时所需时间为

$$t=\frac{a}{v_1\cos\alpha}+\frac{b}{v_2\cos\beta},$$

需求 t 对 α,β 的极小值. 但 α,β 并非独立变量，受条件

$$a\operatorname{tg}\alpha+b\operatorname{tg}\beta=h$$

的约束. 问题化为条件极值.

对 $a\operatorname{tg}\alpha+b\operatorname{tg}\beta=h$ 微分，得 $a\dfrac{d\alpha}{\cos^2\alpha}+b\dfrac{d\beta}{\cos^2\beta}=0$，因而

$$d\beta=-\frac{b\cos^2\beta}{a\cos^2\alpha}d\alpha,$$

而

$$\begin{aligned}dt&=\frac{-a\sin\alpha}{v_1\cos^2\alpha}d\alpha+\frac{-a\sin\beta}{v_2\cos^2\beta}d\beta\\&=-\frac{a\sin\alpha}{v_1\cos^2\alpha}d\alpha+\frac{-a\sin\beta}{v_2\cos^2\beta}\left(-\frac{b\cos^2\beta}{a\cos^2\alpha}\right)d\alpha.\end{aligned}$$

但判别点的一阶微分为零，解得 α,β 满足 $\dfrac{\sin\alpha}{v_1}=\dfrac{\sin\beta}{v_2}$ 时 $dt=0$. 这时判别点是唯一的，而由问题知极小点存在，其必在判别点处达到.

上面两例都是先求得判别点唯一，再根据问题性质说明极大极小值的存在性而得到极

值问题的解．这在实际中是常用的.

§4.3 Lagrange 乘子法

首先我们考虑函数 $f(x,y,z)$ 在条件 $j(x,y,z)=0$ 下的极值问题．设 $f(x,y,z)$ 和 $j(x,y,z)$ 都是 C^1 的函数．令

$$\Sigma=\{(x,y,z)\,|\,j(x,y,z)=0\}.$$

设 $P_0=(x_0,y_0,z_0)\in\Sigma$ 是 $f(x,y,z)$ 在 Σ 上的一个极值点．任取 Σ 中过 P_0 的曲线 $g(t)=(x(t),y(t),z(t))$，设 $g(0)=P_0$，则 $t=0$ 是 $f(x(t),y(t),z(t))$ 的极值点．因而由费马定理知

$$\left.\frac{df(x(t),y(t),z(t))}{dt}\right|_{t=0}=\frac{\partial f(x_0,y_0,z_0)}{\partial x}x'(0)+\frac{\partial f(x_0,y_0,z_0)}{\partial y}y'(0)$$
$$+\frac{\partial f(x_0,y_0,z_0)}{\partial z}z'(0)=0,$$

即 $\mathrm{grad}(f)(P_0)$ 与切向量 $(x'(0),y'(0),z'(0))$ 垂直．从而 $\mathrm{grad}(f)(P_0)$ 与 Σ 在 P_0 点的切面垂直．假设 $\mathrm{grad}(j)(P_0)\neq 0$，由隐函数定理及 Σ 的切面公式知，$\mathrm{grad}(j)(P_0)$ 是 Σ 在 P_0 点切面的法向量，得存在 l，使得

$$\mathrm{grad}(f)(P_0)=l\,\mathrm{grad}(j)(P_0).$$

反之如果上式成立，则将 $f(x,y,z)$ 限制在 Σ 上每一条过 P_0 的曲线时，P_0 都是其判别点．因此可以认为 P_0 也是 $f(x,y,z)$ 在 Σ 上的判别点．为求 $f(x,y,z)$ 在 Σ 上的极值点，我们需先求其判别点，即我们需先解方程 $\mathrm{grad}(f)(P_0)=l\,\mathrm{grad}(j)(P_0)$．定义函数

$$F(x,y,z,l)=f(x,y,z)+l\,j(x,y,z),$$

则求 $f(x,y,z)$ 在 Σ 上的判别点与求

$$dF(x,y,z,l)=\frac{\partial F}{\partial x}dx+\frac{\partial F}{\partial y}dy+\frac{\partial F}{\partial z}dz+\frac{\partial F}{\partial l}dl=0$$

的点是等价的，而第二种形式的方程与求 $F(x,y,z,l)$ 的普通极值的判别点是一致的．这样从形式上我们将条件极值的问题化为普通极值的问题.

上面的方法称为 Lagrange 乘子法，其可以推广到更一般的形式．为此我们先给出下面定义．设$j_1(x_1,\cdots,x_n),\cdots,j_m(x_1,\cdots,x_n)$都是$C^1$的函数，令

$$\Sigma=\{(x_1,\cdots,x_n)\,|\,j_1(x_1,\cdots,x_n)=\cdots=j_m(x_1,\cdots,x_n)=0\}.$$

设在 $P_0=(x_1^0,\cdots,x_n^0)$ 处 $\operatorname{rank}\left(\dfrac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_n)}(P_0)\right)=m$，不妨设 $\dfrac{\partial(j_1,\cdots,j_m)}{\partial(x_1,\cdots,x_m)}(P_0)\neq 0$．由隐函数定理知，存在$P_0$的邻域$U$，使在$U$上，$\Sigma\bigcap U$ 可表示为

$$\left(h_1(x_{m+1},\cdots,x_n),\cdots,h_m(x_{m+1},\cdots,x_n),x_{m+1},\cdots,x_n\right)$$

的形式，其中$h_1(x_{m+1},\cdots,x_n),\cdots,h_m(x_{m+1},\cdots,x_n)$是$(x_{m+1}^0,\cdots,x_n^0)$的$C^1$的函数．对于函数 $f(x_1,\cdots,x_n)$，其对于Σ的条件极值在P_0的邻域U上化为

$$f\left(h_1(x_{m+1},\cdots,x_n),\cdots,h_m(x_{m+1},\cdots,x_n),x_{m+1},\cdots,x_n\right)$$

的普通极值问题．我们称 $P_0=(x_1^0,\cdots,x_n^0)$为 $f(x_1,\cdots,x_n)$ 在Σ上的条件极值的判别点．如果$(x_{m+1}^0,\cdots,x_n^0)$是上面函数对普通极值的判别点,

定理 1: 假设如上，$P_0=(x_1^0,\cdots,x_n^0)$ 为 $f(x_1,\cdots,x_n)$ 在Σ上的判别点的充分必要条件是存在 $l_1,\cdots,l_m$，使得

$$\operatorname{grad}(f)(P_0)=(-1)l_1\operatorname{grad}(j_1)(P_0)+\cdots+(-1)l_m\operatorname{grad}(j_m)(P_0).$$

几何说明：与本节开始时的讨论相同，P_0是 $f(x_1,\cdots,x_n)$ 在Σ上的判别点的充分必要条件是 $\operatorname{grad}(f)(P_0)$ 与Σ在 P_0 点的切面垂直，即$\operatorname{grad}(f)(P_0)$ 在 P_0 点的法空间内．但$\{\operatorname{grad}(j_1)(P_0),\cdots,\operatorname{grad}(j_m)(P_0)\}$是$\Sigma$在$P_0$点的法空间的线性基，得定理 1.

证明： $P_0=(x_1^0,\cdots,x_n^0)$ 为 $f(x_1,\cdots,x_n)$ 在 Σ 上条件极值的判别点等价于 $\widetilde{P}_0=(x_{m+1}^0,\cdots,x_n^0)$是 $f\left(h_1(x_{m+1},\cdots,x_n),\cdots,h_m(x_{m+1},\cdots,x_n),x_{m+1},\cdots,x_n\right)$的判别点，即对 $j=m+1,\cdots,n$，在P_0点有

$$\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}\frac{\partial h_i}{\partial x_j}+\frac{\partial f}{\partial x_j}=0,$$

表示成矩阵形式为:

$$\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_m}\right)\frac{D(h_1,\cdots,h_m)}{D(x_{m+1},\cdots,x_n)}+\left(\frac{\partial f}{\partial x_{m+1}},\cdots,\frac{\partial f}{\partial x_n}\right)=0. \tag{1}$$

而由

$$j_i\left(h_1(x_{m+1},\cdots,x_n),\cdots,h_m(x_{m+1},\cdots,x_n),x_{m+1},\cdots,x_n\right)=0,\ i=1,\cdots,m,$$

得在 P_0 点，对 $i=1,\cdots,m$ 有

$$\sum_{l=1}^{m}\frac{\partial j_i}{\partial x_l}\frac{\partial h_l}{\partial x_j}+\frac{\partial j_i}{\partial x_j}=0,\ j=m+1,\cdots,n,$$

表示成矩阵形式，即

$$\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\cdot\frac{D(h_1,\cdots,h_m)}{D(x_{m+1},\cdots,x_n)}+\frac{D(j_1,\cdots,j_m)}{D(x_{m+1},\cdots,x_n)}=0. \tag{2}$$

已知 $\dfrac{\partial(j_1,\cdots,j_m)}{\partial(x_1,\cdots,x_m)}(P_0)\neq 0$，解得

$$\frac{D(h_1,\cdots,h_m)}{D(x_{m+1},\cdots,x_n)}=(-1)\left[\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}\frac{D(j_1,\cdots,j_m)}{D(x_{m+1},\cdots,x_n)}.$$

代入（1）式，得 P_0 是 $f(x_1,\cdots,x_n)$ 对 Σ 的判别点等价于

$$\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_m}\right)(-1)\left[\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}\frac{D(j_1,\cdots,j_m)}{D(x_{m+1},\cdots,x_n)}+\left(\frac{\partial f}{\partial x_{m+1}},\cdots,\frac{\partial f}{\partial x_n}\right)=0.$$

显然 $\left(\dfrac{\partial f}{\partial x_1},\cdots,\dfrac{\partial f}{\partial x_m}\right)(-1)\left[\dfrac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}\dfrac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}+\left(\dfrac{\partial f}{\partial x_1},\cdots,\dfrac{\partial f}{\partial x_m}\right)=0$，与上式结合得

$$\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_m}\right)(-1)\left[\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_n)}+\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_n}\right)=0. \tag{3}$$

令

$$(l_1,\cdots,l_m)=(-1)\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_m}\right)\left[\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}, \tag{4}$$

则上式为

$$-(l_1,\cdots,l_m)\begin{pmatrix}\mathrm{grad}(j_1)\\ \vdots\\ \mathrm{grad}(j_m)\end{pmatrix}=\mathrm{grad}(f).$$

反之，如果 $(l_1,\cdots,l_m)$ 存在，则其必由（4）式给出，代入则（3）式成立，P_0 是 $f(x_1,\cdots,x_n)$

在 Σ 上的判别点. 证毕.

利用定理 1, 引入变量 $(l_1,\cdots,l_m)$, 称为 Lagrange 乘子. 定义一个新函数

$$F(x_1,\cdots,x_n;l_1,\cdots,l_m)=f(x_1,\cdots,x_n)+l_1 j_1(x_1,\cdots,x_n)+\cdots+l_m j_m(x_1,\cdots,x_n),$$

则求 $f(x_1,\cdots,x_n)$ 在 Σ 上条件极值判别点问题化为求 $F(x_1,\cdots,x_n;l_1,\cdots,l_m)$ 的普通极值的判别点的问题, 即

定理 2: $P_0=(x_1^0,\cdots,x_n^0)$ 为 $f(x_1,\cdots,x_n)$ 对 Σ 的条件极值的判别点的充分必要条件是存在 $(l_1^0,\cdots,l_m^0)$, 使得 $(x_1^0,\cdots,x_n^0;l_1^0,\cdots,l_m^0)$ 是 $F(x_1,\cdots,x_n;l_1,\cdots,l_m)$ 的普通极值的判别点.

定理 2 为求条件极值提供了一个实际可应用的方法. 当然其得到的仅是判别点, 要说明所得判别点是否极值点以及是什么极值点还需用 Hessi 矩阵.

假设上面讨论的函数都是 C^2 的. 设 $(x_1^0,\cdots,x_n^0;l_1^0,\cdots,l_m^0)$ 是 $F(x_1,\cdots,x_n;l_1,\cdots,l_m)$ 的判别点. 令 $P_0=(x_1^0,\cdots,x_n^0)$, 将 $(l_1^0,\cdots,l_m^0)$ 固定, 考虑 $F(x_1,\cdots,x_n;l_1^0,\cdots,l_m^0)$ 在 P_0 的 Hessi 矩阵

$$\left(\frac{\partial^2 F}{\partial x_i \partial y_j}(P_0)\right)_{1\le i,j\le n}.$$

分两种情况.

1) $\left(\frac{\partial^2 F}{\partial x_i \partial y_j}(P_0)\right)$ 正定或负定. 这时 P_0 是 $F(x_1,\cdots,x_n;l_1^0,\cdots,l_m^0)$ 的极值点. 而限制在 Σ 上时 $F(x_1,\cdots,x_n;l_1^0,\cdots,l_m^0)\equiv f(x_1,\cdots,x_n)$. 因此 P_0 也是 $f(x_1,\cdots,x_n)$ 的相应的条件极值点.

2) $\left(\frac{\partial^2 F}{\partial x_i \partial y_j}(P_0)\right)$ 不定或半正定或半负定. 这时 P_0 可能不是 $F(x_1,\cdots,x_n;l_1^0,\cdots,l_m^0)$ 的极值点. 但不能因此判定 P_0 不是 $f(x_1,\cdots,x_n)$ 的极值点. 设 $\frac{\partial(j_1,\cdots,j_m)}{\partial(x_1,\cdots,x_m)}(P_0)\neq 0$, 利用 $dj_i=0,\ i=1,\cdots,m$, 解出

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_m \end{pmatrix} = (-1)\left[\frac{D(j_1,\cdots,j_m)}{D(x_1,\cdots,x_m)}\right]^{-1}\frac{D(j_1,\cdots,j_m)}{D(x_{m+1},\cdots,x_n)}\begin{pmatrix} dx_{m+1} \\ \vdots \\ dx_n \end{pmatrix}.$$

代入二次式 $\sum_{i,j=1}^{n}\frac{\partial^2 F(P_0)}{\partial x_i \partial x_j}dx_i dx_j$，**得到** $dx_{m+1},\cdots,dx_n$ **的二次式** $\sum_{i,j=m+1}^{n} a_{ij}dx_i dx_j$． **如果其系数矩阵** (a_{ij}) **是正定的**，**则** P_0 **是** $f(x_1,\cdots,x_n)$ **在** Σ **上的极小点**；**如果** (a_{ij}) **是负定的**，**则** P_0 **是** $f(x_1,\cdots,x_n)$ **在** Σ **上的极大点**.

例：设 $f(x,y,z)=x+y+z$，**求其在曲面** $xyz=c^3$ **上的极值点**.

解：令 $F(x,y,z,l)=f(x,y,z)+l(xyz-c^3)$，**则由** $dF(x,y,z,l)=0$ **解得**

$$x=y=z=c,\ l=-\frac{1}{c^2}.$$

如果 $c=0$，**得** $f(x,y,z)$ **无判别点**，**因而无极值点**.

设 $c\neq 0$，**则在** $l=-\frac{1}{c^2}, P_0=(c,c,c)$ **处**

$$\left(\frac{\partial^2 F(P_0)}{\partial x \partial y}\right)=\begin{pmatrix} 0 & -\frac{1}{c} & -\frac{1}{c} \\ -\frac{1}{c} & 0 & -\frac{1}{c} \\ -\frac{1}{c} & -\frac{1}{c} & 0 \end{pmatrix}$$

不是正定也不是负定的.

对 $xyz-c^3=0$ **微分**，**在** $P_0=(c,c,c)$ **处**，**解得** $dz=(-dx-dy)$． **代入**

$$\sum_{i,j=1}^{3}\frac{\partial^2 F(P_0)}{\partial x_i \partial x_j}dx_i dx_j=-\frac{2}{c}\left[dxdy+dydz+dzdx\right]$$

中得二次式

$$\begin{aligned} &-\frac{2}{c}\left[dxdy+dy(-dx-dy)+(-dx-dy)dx\right] \\ &=-\frac{2}{c}\left[-dx^2-dxdy-dy^2\right] \\ &=(dx,dy)\begin{pmatrix} \frac{2}{c} & \frac{1}{c} \\ \frac{1}{c} & \frac{2}{c} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix}. \end{aligned}$$

因此 $c>0$ 时 $\begin{pmatrix} \frac{2}{c} & \frac{1}{c} \\ \frac{1}{c} & \frac{2}{c} \end{pmatrix}$ 正定，(c,c,c) 是 $f(x,y,z)$ 在 $xyz=c^3$ 上的极小点；$c<0$ 时 $\begin{pmatrix} \frac{2}{c} & \frac{1}{c} \\ \frac{1}{c} & \frac{2}{c} \end{pmatrix}$ 负定，(c,c,c) 是 $f(x,y,z)$ 的极大点.

§4.4 最小二乘法

作为极值理论的应用，下面我们给出一个在实际问题中求解未知函数常用的最小二乘法.

设我们知道一个函数关系 $y=f(x)$，并且通过采样求得这函数在 $n+1$ 个点上的函数值 $y_i=f(x_i), i=0,\cdots,n$. 求这个函数.

显然满足条件的函数 $y=f(x)$ 有无穷多，所以问题无意义. 改为求一 m 次多项式

$$y=a_0+a_1x+\cdots+a_mx^m$$

使误差尽可能小. 这里我们假设 n 可以取得很大，但由于计算量的关系，m 比较小.

令

$$\begin{cases} a_0+a_1x_0+\cdots+a_mx_0^m-y_0=d_0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_0+a_1x_n+\cdots+a_mx_n^m-y_n=d_n, \end{cases}$$

问题化为求 $(a_0,a_1,\cdots,a_m)$，使得 $\sum\limits_{i=0}^{n}d_i^2$ 达到最小.

定义函数 $f(a_0,a_1,\cdots,a_m)=\sum\limits_{i=0}^{n}d_i^2$. 令

$$y=(y_0,y_1,\cdots,y_n),\quad a=(a_0,a_1,\cdots,a_m),\quad d=(d_0,d_1,\cdots,d_n).$$

令

$$G=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ x_0^2 & x_1^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ x_0^m & x_1^m & \cdots & x_n^m \end{pmatrix}_{(m+1)\times(n+1)},$$

利用矩阵形式，$f(a_0,a_1,\cdots,a_m)$ 可表示为

$$
\begin{aligned}
f(a_0,a_1,\cdots,a_m) &= (aG-y,aG-y)\\
&= (aG,aG)-(aG,y)-(y,aG)+(y,y)\\
&= aGG^Ta^T-aGy^T-yG^Ta^T+yy^T,
\end{aligned}
$$

其中G^T,a^T,y^T表示其转置. 直接计算得

$$
\begin{aligned}
df(a_0,a_1,\cdots,a_m) &= daGG^Ta^T+aGG^Tda^T-daGy^T-yG^Tda^T\\
&= da\left[GG^Ta^T+GG^Ta^T-Gy^T-Gy^T\right].
\end{aligned}
$$

因此判别点方程$df(a_0,a_1,\cdots,a_m)=0$的解为$GG^Ta^T=Gy^T$.

引理 1: 设$x_0,x_1,\cdots,x_n$两两互不相同且$m\le n$，则GG^T是可逆矩阵.

证明：显然对称矩阵GG^T是半正定的. 要证$\det(GG^T)\neq 0$，仅需证对于任意$c=(c_0,c_1,\cdots,c_m)\neq 0$时，$cG\neq 0$. 事实上，如果$cG=0$，则$x_0,x_1,\cdots,x_n$是多项式

$$
P(x)=c_0+c_1x+\cdots+c_mx^m
$$

的根. 但$m\le n$，其不可能有$n+1$个不同根. 所以，$cG\neq 0$，$cGG^Tc^T=\|cG\|^2>0$，GG^T正定.

由引理得$a^*=\left((GG^T)^{-1}Gy^T\right)^T$是判别点. 希望证明其是极小点. 对

$$
df=2da\left[GG^Ta^T-Gy^T\right]
$$

微分得

$$
d^2f=2daGG^Tda.
$$

因此$f(a_0,a_1,\cdots,a_m)$的 Hessi 矩阵是正定的，得$a^*=\left((GG^T)^{-1}Gy^T\right)^T$是$f(a_0,a_1,\cdots,a_m)$的极小点. 由于其是唯一的，因而是最小点.

这样我们得到所求的多项式是存在并且唯一的. 由于这种解的存在唯一性，因而其也是有实际意义的.

习题

1．设$x=(x_1,\cdots,x_m)$，$f(x)$在$a=(a_1,\cdots,a_m)$去极小（大）值，并存在$\dfrac{\partial^2 f(a)}{\partial x_i^2}(i=1,2,$

$\cdots, m)$. **求证：**

$$\frac{\partial^2 f(a)}{\partial x_i^2} \geq 0 (\leq 0).$$

2．**求下列函数的极大值点和极小值点**:

(1) $f(x,y) = x^2(y-1)^2$;

(2) $f(x,y) = xy(x^2+y^2-1)$;

(3) $f(x,y) = \left(\sqrt{x^2+y^2}-1\right)^2$;

(4) $f(x,y) = 3x^2y - x^4 - 2y^2$;

(5) $f(x,y) = \sin x + \sin y + \sin(x+y)$;

(6) $f(x,y) = \mathrm{tg}x + \mathrm{tg}y - \mathrm{tg}(x+y)$.

3．**用隐函数微分法求隐函数** $z = z(x,y)$ **的的极大值和极小值**:

(1) $x^2+y^2+z^2-2x-2y-4z-10=0$;

(2) $(x+y)^2+(y+z)^2+(z+x)^2=3$;

(3) $x^2+y^2+z^2-xz-yz+2x+2y+2z-2=0$;

(4) $z^2+xyz-x^2-xy^2-9=0$.

4．**设** $f(x,y) = 3x^2y - x^4 - 2y^2$. **证明**: $(0,0)$ **不是它的极值点**, **但沿过**$(0,0)$ **的每条直线**, $(0,0)$ **都是它的极大点**.

5．**作容积为**V **的开口长方形容器**, **问尺寸怎样时**, **用料最省**.

6．**在椭球面**$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$**的内接长方体中**, **求体积为最大的那个长方体**.

7．**设** $f(x) \in C(\mathbf{R}^m)$. **求证** $f(x)$ **在条件** $(a,x)=1,\ x_i \geq 0,\ i=1,2,\cdots,m$ **下存在最大值与最小值**, **其中** $a=(a_1,\cdots,a_m)$, $a_i > 0\ (i=1,2,\cdots,m)$.

8．**求下列条件极大值和条件极小值**:

(1) $x^2+y^2=1$, **求** $f(x,y) = ax^2+2hxy+by^2$ **的极值**;

(2) $ax+by+cz=k$, **求** $f(x,y,z) = x^l y^m z^n$ **的极值**, **其中** l,m,n **为正整数**, a,b,c,k

为常数;

（3） $\left(x^2+y^2+z^2\right)^2=a^2x^2+b^2y^2+c^2z^2$, $lx+my+nz=0$，求

$$f(x,y,z)=x^2+y^2+z^2$$

的极值;

（4） $lx+my+nz=0$，$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$，求 $f(x,y,z)=\dfrac{x^2}{a^4}+\dfrac{y^2}{b^4}+\dfrac{z^2}{c^4}$ 的极值.

9． $x^2+y^2+z^2\le 1$，求 $x^3+y^3+z^3-2xyz$ 的最大值和最小值.

10．求 $f(x_1,\cdots,x_n)=x_1\cdots x_n$ 在条件 $\dfrac{1}{x_1}+\cdots+\dfrac{1}{x_n}=\dfrac{1}{a}\ (x_i>0,\ a>0)$ 之下的极值，并证明:

当 $a_i>0\ (i=1,2,\cdots,n)$ 时,

$$n\left(\frac{1}{a_1}+\cdots+\frac{1}{a_n}\right)^{-1}\le (a_1a_2\cdots a_n)^{\frac{1}{n}}.$$

11．求圆的外切三角形中面积最小者.

12．长为 a 的铁丝切成两段，一段围成一个正方形，另一段围成一个圆．这两段的长各为多少时，使它们所围正方形和圆形面积之和最大.

13．求一点 O，使与一个凸四边形的四顶点距离之和为最小.

14．设 $f(x)\in C^{(1)}(\mathbf{R}^m,\mathbf{R}^m)$，存在常数 $a>0$，使得对任意 $x,y\in\mathbf{R}^m$ 有

$$|f(x)-f(y)|\ge a|x-y|.$$

证明:

（1） $\det Df(x)\ne 0$， $x\in\mathbf{R}^m$;

（2） $j(x)=|f(x)-y|^2=\left(f(x)-y,\ f(x)-y\right)$ 在 $\mathbf{R}^m$ 存在最小值，无最大值;

（3） $f(\mathbf{R}^m)=\mathbf{R}^m$.