Review

- 多元函数的无条件极值

Thm. n元函数f在x_0的邻域中可微, x_0为f的极值点,则x_0为f的驻点, 即$\mathrm{grad}f(\mathrm{x}_0)=0.$

Thm. n元函数f在x_0的邻域中二阶连续可微, $\mathrm{grad}f(\mathrm{x}_0)=0,$

(1)若$H_f(\mathrm{x}_0)$正定,则$f(\mathrm{x}_0)$严格极小.

(2)若$H_f(\mathrm{x}_0)$负定,则$f(\mathrm{x}_0)$严格极大.

(3)若$H_f(\mathrm{x}_0)$不定,则$f(\mathrm{x}_0)$不是极值.

§ 10. 条件极值

最简单的条件极值问题：(P_1) $\begin{array}{l}\max(\min)\ f(x,y)\\ s.t.\quad g(x,y)=0\end{array}$

称$f(x,y)$为目标函数，$g(x,y)=0$为约束条件.

求解问题(P_1)的思路：若$g(x,y)=0$确定了隐函数

$$y=y(x), y'(x)=-g'_x(x,y)\big/g'_y(x,y),$$

则原问题(P_1)转化为一元函数的无条件极值问题

$$\max(\min)\ \ \varphi(x)=f\big(x,y(x)\big).$$

Question. 可行性？隐函数的求解.

清華大學

1. 一个约束的条件极值问题

$$(\mathrm{P}_2)\quad \begin{aligned} &\max(\min)\ \ f(\mathrm{x}) \\ &s.t.\quad \varphi(\mathrm{x})=0 \end{aligned}$$

Thm. 设$\Omega\subset\mathbb{R}^n$为开区域,$f,\varphi:\Omega\to\mathbb{R}$均一阶连续可微,且$\mathrm{grad}\varphi(\mathrm{x})\neq 0,\forall\mathrm{x}\in\Omega$.若$\mathrm{x}_0\in\Omega$是条件极值问题$(P_2)$的最值(极值)点,则存在$\lambda_0\in\mathbb{R}$, $s.t.(\mathrm{x}_0,\lambda_0)$为

$$L(\mathrm{x},\lambda)=f(\mathrm{x})+\lambda\varphi(\mathrm{x})$$

的驻点.称$L(\mathrm{x},\lambda)$为Lagrange函数,称λ为Lagrange乘子.

Proof. 已知$\mathrm{grad}\varphi(\mathrm{x})\neq 0,\forall\mathrm{x}\in\Omega$,不妨设$\dfrac{\partial\varphi}{\partial x_n}(\mathrm{x}_0)\neq 0$,则

清華大學

$\varphi(\mathrm{x})=0$在$\mathrm{x}_0=(x_0^{(1)},\cdots,x_0^{(n)})$的邻域中确定了隐函数,

存在$\hat{\mathrm{x}}_0=(x_0^{(1)},\cdots,x_0^{(n-1)})$的邻域$B(\hat{\mathrm{x}}_0,\eta)$,及函数

$$x_n=g(x_1,\cdots,x_{n-1})=g(\hat{\mathrm{x}}),\quad \hat{\mathrm{x}}\in B(\hat{\mathrm{x}}_0,\eta),$$

s.t. $\quad \varphi(\hat{\mathrm{x}},g(\hat{\mathrm{x}}))=\varphi(x_1,\cdots,x_{n-1},g(x_1,\cdots,x_{n-1}))=0,$

且 $$\frac{\partial g}{\partial x_i}=-\frac{\partial\varphi}{\partial x_i}\Big/\frac{\partial\varphi}{\partial x_n},\qquad i=1,2,\cdots,n-1.\quad (*)$$

不妨设x_0是$f(\mathrm{x})$在$\Omega_1=\{\mathrm{x}:\varphi(\mathrm{x})=0\}$上最(极)大值点,

则$\exists\delta\in(0,\eta)$, s.t.

$$f(\mathrm{x})\le f(\mathrm{x}_0),\forall\mathrm{x}\in\mathrm{B}(\mathrm{x}_0,\delta)\bigcap\Omega_1.$$

$g(x_1,\cdots,x_{n-1})$在$\hat{\mathrm{x}}_0=(x_0^{(1)},\cdots,x_0^{(n-1)})$连续, $\exists\delta_1\in(0,\frac{\delta}{2})$, s.t.

$$\left|g(\hat{\mathrm{x}})-x_0^{(n)}\right|<\frac{\delta}{2},\qquad \forall\hat{\mathrm{x}}\in \mathrm{B}(\hat{\mathrm{x}}_0,\delta_1)\subset\mathbb{R}^{n-1}.$$

于是$\forall\hat{\mathrm{x}}\in \mathrm{B}(\hat{\mathrm{x}}_0,\delta_1)$,有

$$\left\|\left(\hat{\mathrm{x}},g(\hat{\mathrm{x}})\right)-\mathrm{x}_0\right\|\le\left\|\hat{\mathrm{x}}-\hat{\mathrm{x}}_0\right\|+\left|g(\hat{\mathrm{x}})-x_0^{(n)}\right|<\frac{\delta}{2}+\frac{\delta}{2}=\delta,$$

即$\left(\hat{\mathrm{x}},g(\hat{\mathrm{x}})\right)\in \mathrm{B}(\mathrm{x}_0,\delta)\bigcap\Omega_1$, 从而

$$f(\hat{\mathrm{x}},g(\hat{\mathrm{x}}))\le f(\mathrm{x}_0)=f(\hat{\mathrm{x}}_0,g(\hat{\mathrm{x}}_0)),\quad \forall\hat{\mathrm{x}}\in \mathrm{B}(\hat{\mathrm{x}}_0,\delta_1).$$

因此,$\hat{\mathrm{x}}_0$是$f(\hat{\mathrm{x}},g(\hat{\mathrm{x}}))=f(x_1,\cdots,x_{n-1},g(x_1,\cdots,x_{n-1}))$的极大值点, 故

$$\frac{\partial f}{\partial x_i}(\mathrm{x}_0)+\frac{\partial f}{\partial x_n}(\mathrm{x}_0)\frac{\partial g}{\partial x_i}(\hat{\mathrm{x}}_0)=0,\quad i=1,2,\cdots,n-1.$$

利用(*),得

$$\frac{\partial f}{\partial x_i}(\mathrm{x}_0)-\frac{\partial f}{\partial x_n}(\mathrm{x}_0)\frac{\frac{\partial \varphi}{\partial x_i}(\mathrm{x}_0)}{\frac{\partial \varphi}{\partial x_n}(\mathrm{x}_0)}=0,\quad i=1,2,\cdots,n-1.$$

于是$\exists\lambda_0\in\mathbb{R}, s.t.$

$$\frac{\partial f}{\partial x_i}(\mathrm{x}_0)+\lambda_0\frac{\partial \varphi}{\partial x_i}(\mathrm{x}_0)=0,\quad i=1,2,\cdots,n,$$

即 $$\frac{\partial L}{\partial x_i}(\mathrm{x}_0,\lambda_0)=0,\quad i=1,2,\cdots,n.$$

而$\frac{\partial L}{\partial \lambda}(\mathrm{x}_0,\lambda_0)=\varphi(\mathrm{x}_0)=0.$故$(\mathrm{x}_0,\lambda_0)$为$L(\mathrm{x},\lambda)$的驻点.□

清華大學

例: 求$f=xy$在圆周$(x-1)^2+y^2=1$上的最大(小)值.

解: $$\max(\min)\ f(x,y)=xy$$
$$s.t.\quad g(x,y)=(x-1)^2+y^2-1=0$$

构造$Lagrange$函数

$$L(x,y,\lambda)=xy+\lambda\left[(x-1)^2+y^2-1\right]$$

求解

$$\begin{cases} L'_x=y-2\lambda(x-1)=0 \\ L'_y=x-2\lambda y=0 \\ L'_\lambda=(x-1)^2+y^2-1=0 \end{cases}$$

得驻点(不需求出λ的值) $(x_1,y_1)=(0,0)$,

$$(x_2,y_2)=\left(\frac{3}{2},\frac{\sqrt{3}}{2}\right),\quad (x_3,y_3)=\left(\frac{3}{2},-\frac{\sqrt{3}}{2}\right).$$

一方面,连续函数$f(x,y)$在有界闭集上能够达到最大(小)值. 另一方面, 达到最大(小)值的点一定对应于$L(x,y,\lambda)$的驻点. 所以最大(小)值一定在上述三点中的某两点达到. 而

$$f(0,0)=0, f\left(\frac{3}{2},\frac{\sqrt{3}}{2}\right)=\frac{3\sqrt{3}}{4}, f\left(\frac{3}{2},-\frac{\sqrt{3}}{2}\right)=-\frac{3\sqrt{3}}{4}.$$

故f在$\left(\frac{3}{2},\frac{\sqrt{3}}{2}\right)$取得最大值$\frac{3\sqrt{3}}{4}$,在$\left(\frac{3}{2},-\frac{\sqrt{3}}{2}\right)$取得最小值$-\frac{3\sqrt{3}}{4}$. □

Question. Lagrange乘子法的几何意义？

$$\begin{aligned}&\max(\min)\ \ f(x,y,z)\\&s.t.\quad g(x,y,z)=0\end{aligned}\qquad (\mathrm{P}_3)$$

其中$g_x'^2+g_y'^2+g_z'^2>0$.(正则性条件)

结论:构造辅助函数

$$L(x,y,z,\lambda)=f(x,y,z)+\lambda g(x,y,z),$$

若(P_3)在(x_0,y_0,z_0)取得极值,则$\exists\lambda_0$, $s.t.(x_0,y_0,z_0,\lambda_0)$是$L(x,y,z,\lambda)$的驻点. 因此在点$(x_0,y_0,z_0,\lambda_0)$处,

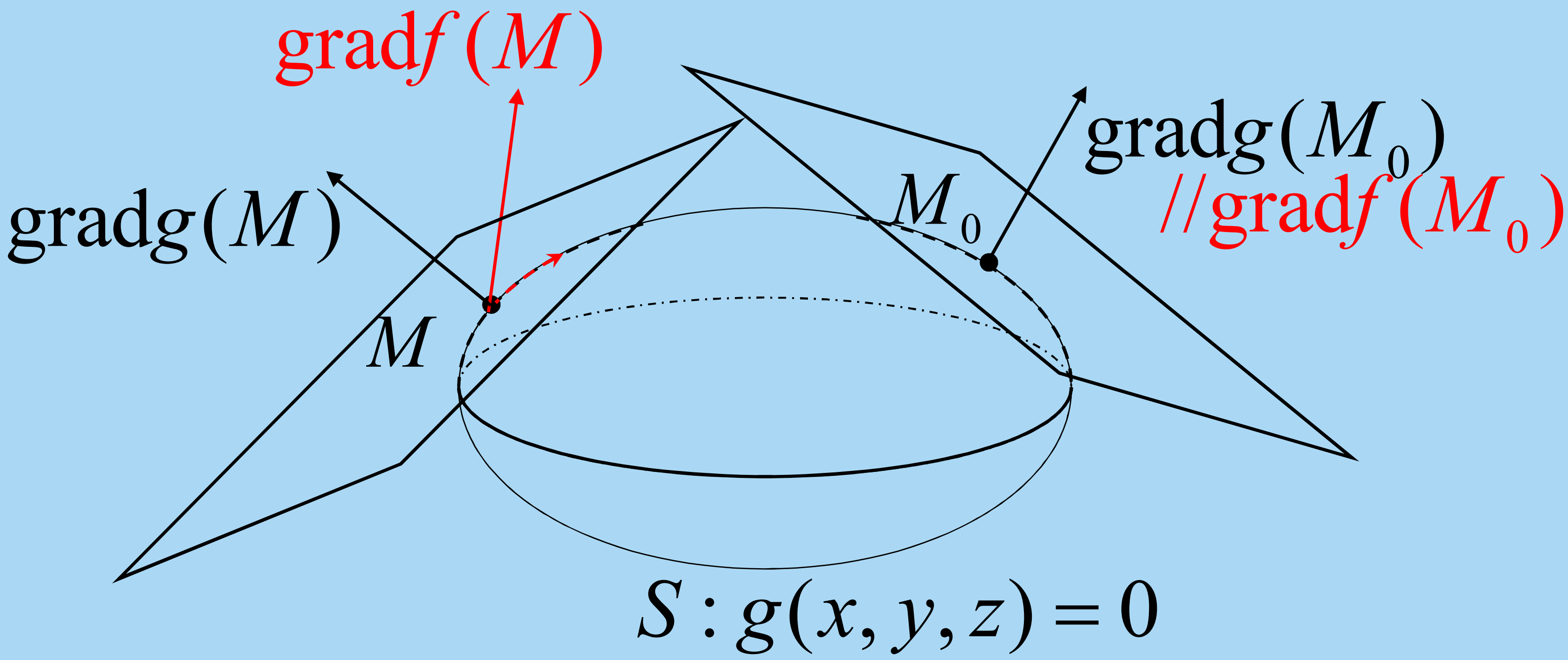
$$\begin{cases} L'_x = f'_x + \lambda_0 g'_x = 0 \\ L'_y = f'_y + \lambda_0 g'_y = 0 \\ L'_z = f'_z + \lambda_0 g'_z = 0 \\ L'_\lambda = g = 0 \end{cases}$$

$\mathrm{grad} f(M_0)$
$= -\lambda_0 \mathrm{grad} g(M_0).$

Remark. (几何解释) 求解(P_3)就是求函数f在曲面S：$g(x,y,z)=0$上的最大(小)值. (P_3)在$M_0(x_0,y_0,z_0)$取得极值,则$\exists\lambda_0, s.t.$

$$\mathrm{grad} f(M_0) = -\lambda_0 \mathrm{grad} g(M_0),$$

即M_0处f增加(减少)最快的方向$\pm\mathrm{grad} f(M_0)$与曲面S在M_0的法向量平行.如图


grad$f(M)$
grad$g(M)$
grad$g(M_0)$
$//$grad$f(M_0)$
M
M_0
$S: g(x, y, z) = 0$

2. 多个约束的条件极值问题

$$\max(\min) \quad f(x,y,z)$$

$$s.t. \quad g(x,y,z)=0 \qquad (\mathrm{P}_4)$$

$$h(x,y,z)=0$$

结论: 构造*Lagrange*函数

$$L(x,y,z,\lambda,\mu)$$

$$=f(x,y,z)+\lambda g(x,y,z)+\mu h(x,y,z).$$

若(P_4)在(x_0,y_0,z_0)取得极值,则$\exists\lambda_0,\mu_0,s.t.$ $(x_0,y_0,z_0,\lambda_0,\mu_0)$为$L(x,y,z,\lambda,\mu)$的驻点.即

$$
(1)\begin{cases}
L'_x = f'_x + \lambda_0 g'_x + \mu_0 h'_x = 0 \\
L'_y = f'_y + \lambda_0 g'_y + \mu_0 h'_y = 0 \\
L'_z = f'_z + \lambda_0 g'_z + \mu_0 h'_z = 0 \\
L'_\lambda = g = 0 \\
L'_\mu = h = 0.
\end{cases}
$$

$$
\begin{aligned}
&-\mathrm{grad}f(M_0) \\
&= \lambda_0 \mathrm{grad}g(M_0) \\
&\quad + \mu_0 \mathrm{grad}h(M_0)
\end{aligned}
$$

因此,求解条件极值问题(P$_4$),可以先求函数L的驻点$(x_0, y_0, z_0, \lambda_0, \mu_0)$(对具体问题不需出$\lambda$, μ的值),再判断(P$_4$)是否在点$M_0(x_0, y_0, z_0)$取得极值.

Remark:(几何解释)条件极值问题(P_4)就是求函数f在曲线

$$\mathrm{L}:\begin{cases} g(x,y,z)=0 \\ h(x,y,z)=0 \end{cases}$$

上的最大(小)值. 设f在L上一点M_0取得极值,则由(1),

$$-\mathrm{grad}f(M_0)=\lambda_0\mathrm{grad}g(M_0)+\mu_0\mathrm{grad}h(M_0),$$

而L在点M_0的切向量T与

$$\mathrm{grad}g(M_0)\times\mathrm{grad}h(M_0)$$

平行. 故函数$f(x,y,z)$在曲线L上一点M_0处取得极值时, $\mathrm{grad}f(M_0)$与L在点M_0的切向量T垂直.

清华大学

4. 例

例1: 求曲面$S_1: z = x^2 + y^2$到平面$\Pi: x + y - 2z = 2$的最短距离.

解: 平面Π外一点(x, y, z)到平面的距离为

$$\frac{1}{\sqrt{6}}\left|x + y - 2z - 2\right|.$$

对条件极值问题

$$\min \quad (x + y - 2z - 2)^2$$

$$s.t. \quad x^2 + y^2 - z = 0$$

构造辅助函数

$$L(x, y, z, \lambda) = (x + y - 2z - 2)^2 + \lambda(x^2 + y^2 - z).$$

清華大學

由
$$\begin{cases} L'_x = 2(x+y-2z-2)+2\lambda x = 0 \\ L'_y = 2(x+y-2z-2)+2\lambda y = 0 \\ L'_z = -4(x+y-2z-2)-\lambda \\ L'_\lambda = x^2+y^2-z = 0 \end{cases}$$

得
$$(x,y,z) = (1/4, 1/4, 1/8).$$

根据题意距离的最小值一定存在,而驻点唯一,故必在$(1/4,1/4,1/8)$处取得最小值:

$$\frac{1}{\sqrt{6}}\left|\frac{1}{4}+\frac{1}{4}-2\times\frac{1}{8}-2\right| = \frac{7}{4\sqrt{6}}. \quad \square$$

例2. 设$\alpha, \beta > 0, 1/\alpha + 1/\beta = 1$.求证, $\forall x, y > 0$,

$$xy \le \frac{1}{\alpha} x^{\alpha} + \frac{1}{\beta} y^{\beta}.$$

分析: 欲证$f(x, y) \ge g(x, y)$.只要证明, $\forall$常数C,条件极值问题

$$\min \ f(x, y)$$
$$s.t \ \ g(x, y) = C$$

的最小值不小于C.

解: 对条件极值问题

$$\min \ f(x, y) = \frac{1}{\alpha} x^{\alpha} + \frac{1}{\beta} y^{\beta}$$
$$s.t \quad xy = C (> 0),$$

Question:
约束条件换成 $f = C$?

令$L(x,y,\lambda)=\frac{1}{\alpha}x^{\alpha}+\frac{1}{\beta}y^{\beta}+\lambda(xy-C)$.

由$\begin{cases} L'_x = x^{\alpha-1}+\lambda y=0 \\ L'_y = y^{\beta-1}+\lambda x=0 \\ L'_\lambda = xy-C=0 \end{cases}$ 得驻点 $\begin{cases} x_0 = C^{\frac{\beta}{\alpha+\beta}}, \\ y_0 = C^{\frac{\alpha}{\alpha+\beta}}. \end{cases}$

$$f(x_0,y_0)=(\frac{1}{\alpha}+\frac{1}{\beta})C^{\frac{\alpha\beta}{\alpha+\beta}}=C^{\frac{\alpha\beta}{\alpha+\beta}}=C.$$

又当$x^2+y^2\to+\infty$时, $f(x,y)\to+\infty$,因此f在曲线

L:$xy=C(x,y>0)$上有最小值,且最小值为C.

由C的任意性,$\frac{1}{\alpha}x^{\alpha}+\frac{1}{\beta}y^{\beta}\geq xy$. □

例3. $A=(a_{ij})$为$n\times n$对称矩阵, $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 求

n元函数

$$f(x_1,x_2,\cdots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j=x^{\mathrm{T}}Ax$$

在单位球面$S=\left\{(x_1,x_2,\cdots,x_n)\middle|\sum_{i=1}^{n}x_i^2=1\right\}$上的最大

值和最小值.

解: 构造辅助函数

$$L(x_1,x_2,\cdots,x_n,\lambda)=\sum_{i,j=1}^{n}a_{ij}x_ix_j-\lambda\left(\sum_{i=1}^{n}x_i^2-1\right).$$

$L(x,y,z,\lambda)$的驻点满足方程组:

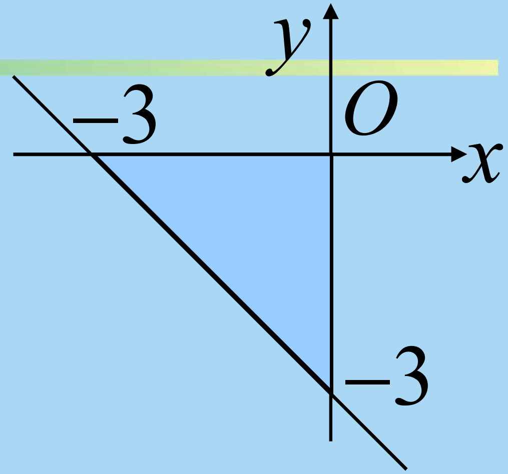
$$\begin{cases} L'_{x_i}=2\left[a_{i1}x_1+\cdots+(a_{ii}-\lambda)x_i+\cdots+a_{in}x_n\right]=0, \\ \qquad\qquad\qquad\qquad\qquad\qquad i=1,2,\cdots,n \\ L'_{\lambda}=\sum_{i=1}^{n}x_i^2-1=0 \end{cases}$$

即 $\begin{cases} \mathrm{A}x-\lambda x=0 \\ x^{\mathrm{T}}x=1. \end{cases}$ 即λ为A的特征值, x为与之对应的单位长度的特征向量.

此时 $$f(x)=x^{\mathrm{T}}\mathrm{A}x=\lambda x^{\mathrm{T}}x=\lambda.$$

于是, f在单位球面S上的最大(小)值分别是矩阵A的最大(小)特征值.□

例: 求$z = x^2 + y^2 - xy + x + y$在闭区域$\{(x, y) \mid x \le 0, y \le 0, x + y \ge -3\}$中的最大值与最小值.



分析:最值能在内部达到,也可能在边界上达到.

解: (1)研究函数在区域内部的情况.

$$\text{由}\begin{cases} z'_x = 2x - y + 1 = 0 \\ z'_y = 2y - x + 1 = 0 \end{cases}\text{得驻点}x = y = -1,$$

此时$z(-1, -1) = -1$.

(2)研究函数在边界上的情况.

•当$x=0$时，$z=y^2+y(-3\le y\le 0)$，此时

$$z_{\max}\big|_{x=0}=z(0,-3)=6,\ z_{\min}\big|_{x=0}=z(0,-\frac{1}{2})=-\frac{1}{4}.$$

•当$y=0$时，$z=x^2+x(-3\le x\le 0)$，

$$z_{\max}\big|_{y=0}=z(-3,0)=6,\ z_{\min}\big|_{y=0}=z(-\frac{1}{2},0)=-\frac{1}{4}.$$

•当$x+y=-3$时，$z=3(x^2+3x+2)(-3\le x\le 0)$，

$$z_{\max}\big|_{x+y=-3}=z(0,-3)=z(-3,0)=6.$$

$$z_{\min}\big|_{x+y=-3}=z(-\frac{3}{2},-\frac{3}{2})=-\frac{3}{4}.$$

综上所述，在点$(0,-3)$和$(-3,0)$处函数取最大值6，在点$(-1,-1)$处函数取最小值-1. □

作业:习题1.9

No.7(3),8,9(3),10(1)