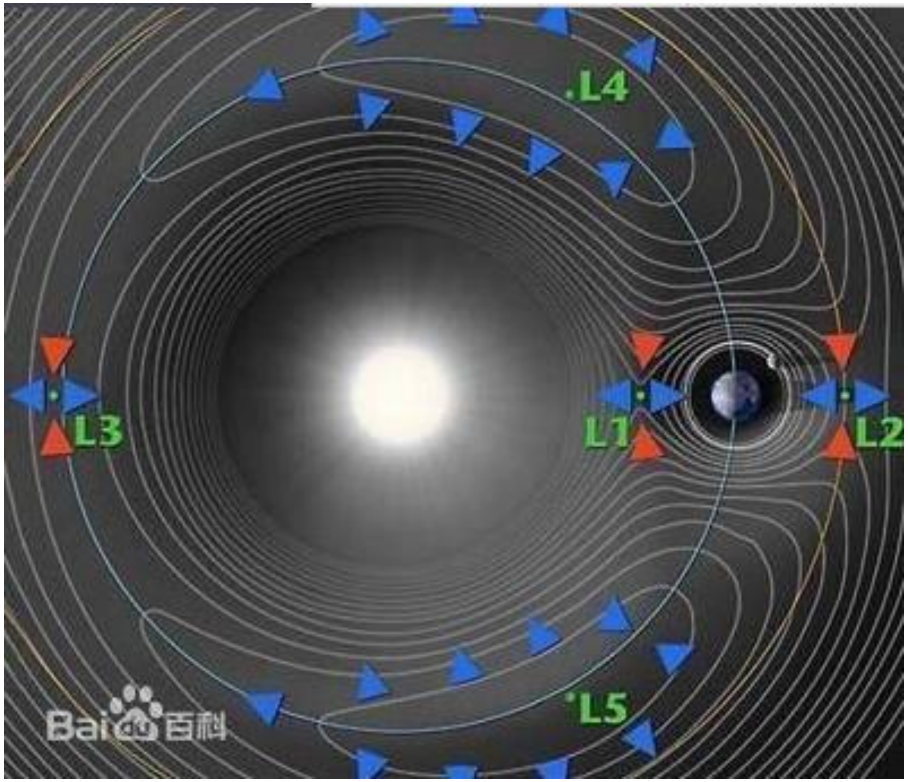
最优化理论与方法

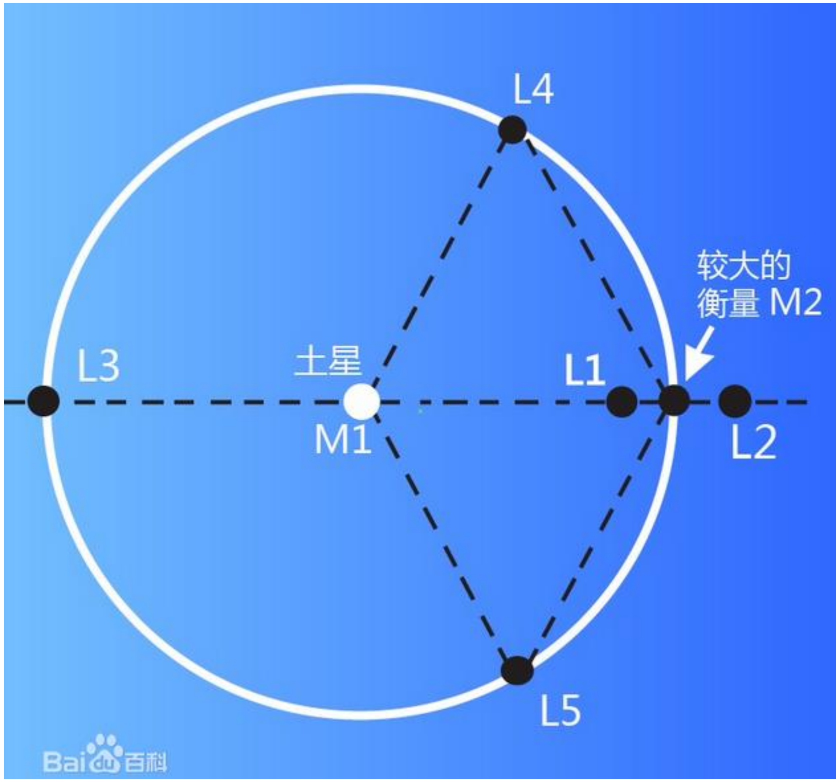
研究生学位课

陈军华（副教授、主任）

- **等式约束的非线性规划**
  - $\min f(x)$
  - $s.t. \quad g_i(x) = 0 \qquad i = 1,2,\dots,m < n$
  - 基本思路：将约束问题转化为无约束问题。
  - 使用拉格朗日乘子(Lagangian)
  - $L(x, \lambda)$

- **拉格朗日点**





- **原问题模型**
  - $minf(x)$
  - $s.t. \quad g_i(x) = 0 \qquad \forall i = 1,2, \dots, m < n$

- **拉格朗日函数**
  - $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$

- **例：6_0**

$$minf(x, y) = x^2 + y^2$$

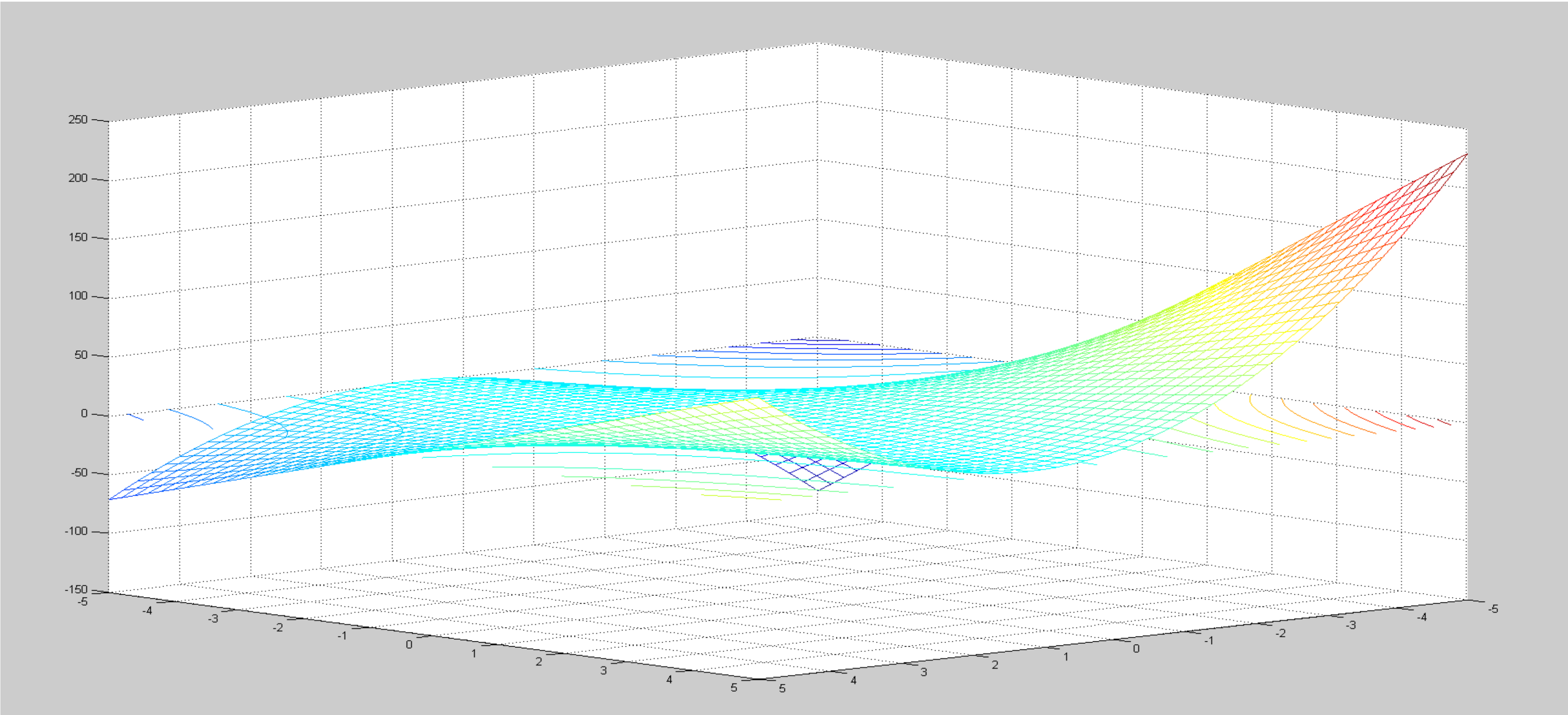
$$s.t. \quad x + y = 1$$

- **例：6_1**

$$\min\ f(x) = (x-2)^2$$

$$s.t.\quad x = 1$$

  - $L(x,\lambda) = (x-2)^2 + \lambda(x-1)$ $\qquad x^* = 1, \lambda^* = 2$



- **等式约束问题解存在的必要性条件**
  - $L(x,\lambda)=f(x)+\sum_{i=1}^{m}\lambda_i g_i(x)$
  - $\frac{\partial L}{\partial x_j}=\frac{\partial L}{\partial x_j}+\sum_{i=1}^{m}\lambda_i\frac{\partial g_i}{\partial x_j}=0 \qquad j=1,2,\dots,n$
  - $\frac{\partial L}{\partial \lambda_i}=g_i(x)=0 \qquad i=1,2,\dots,m$
  - 以上条件成立基础是 无约束问题存在最优解。

- **例6_2**
  - $\boldsymbol{max} \quad \boldsymbol{f}(\boldsymbol{x}) = x_1^2 + \mathbf{4}x_2^2$
  - $\boldsymbol{s.t} \quad \boldsymbol{x_1} + \mathbf{2}\boldsymbol{x_2} = \mathbf{6}$

- **例6_2**
  - $\boldsymbol{max}\ \ \boldsymbol{f(x)} = x_1^2 + 4x_2^2$
  - $\boldsymbol{s.t}\quad \boldsymbol{x_1 + 2x_2 = 6}$

  - $\boldsymbol{L(x,\lambda)} = x_1^2 + 4x_2^2 + \boldsymbol{\lambda(x_1 + 2x_2 - 6)}$

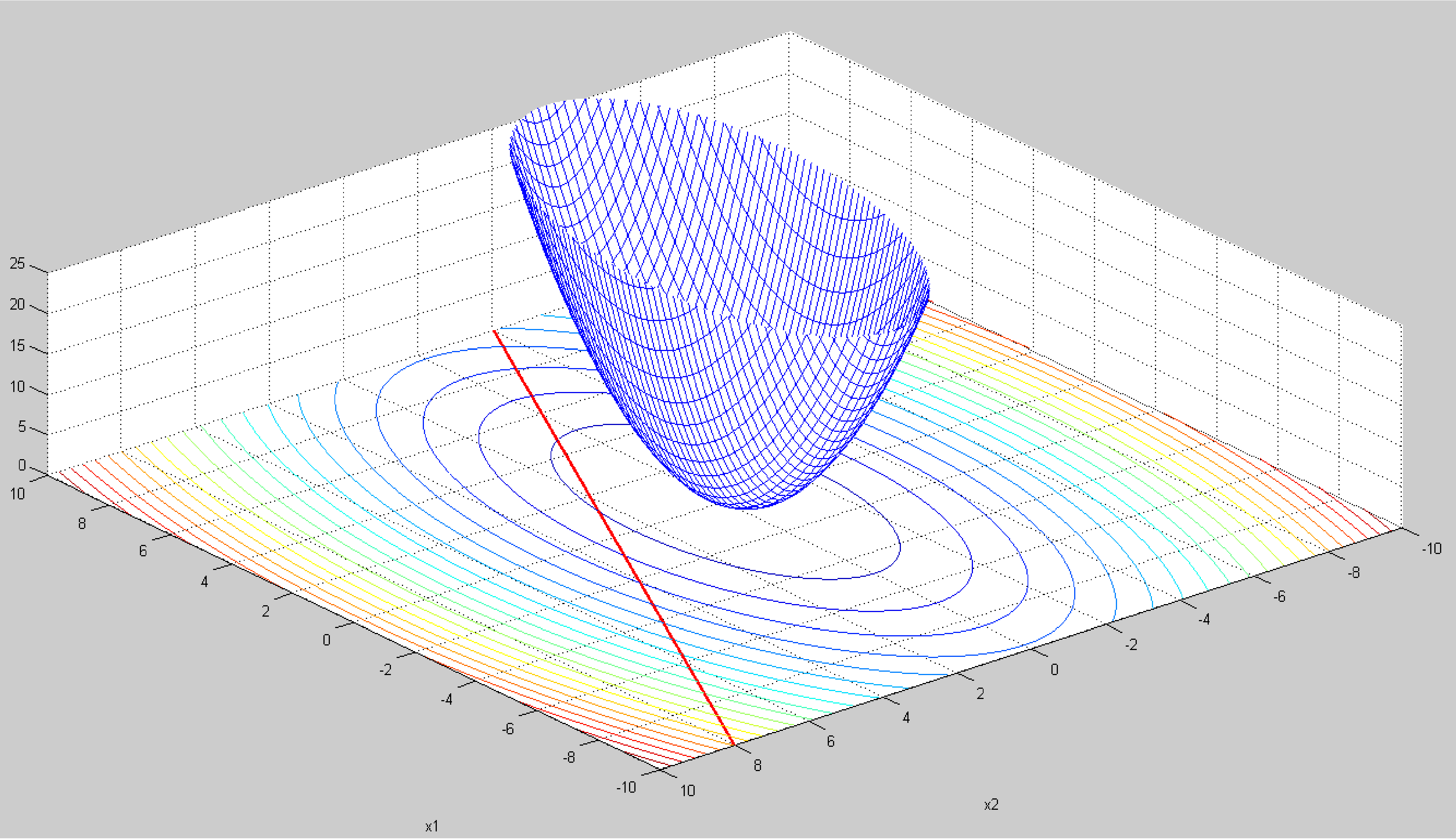
$$\frac{\partial L}{\partial x_1} = 2x_1 + \lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 8x_2 + 2\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = x_1 + 2x_2 - 6 = 0$$

  - 求得 $x_1^* = 3$， $x_2^* = \frac{3}{2}$， $\lambda^* = -6$　　正确否？

- 例6_2



- **关于约束**
  - 如果约束条件可以约减，则先约减。
  - 加上约束条件后，其最优解一定不优于无约束问题。

$$\min f(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2$$
$$s.t. \quad x_1^2 + x_1 x_2 = 1$$
$$x_1 + x_3 = 4$$

• λ存在性讨论

- 对于Lagrangian 函数

$$L(x,\lambda)=f(x)+\sum_{i=1}^{m}\lambda_i g_i(x)$$

及必要性条件：

$$\nabla f(x)+\left[\frac{\partial g}{\partial x}\right]\lambda=0,\qquad g(x)=0$$

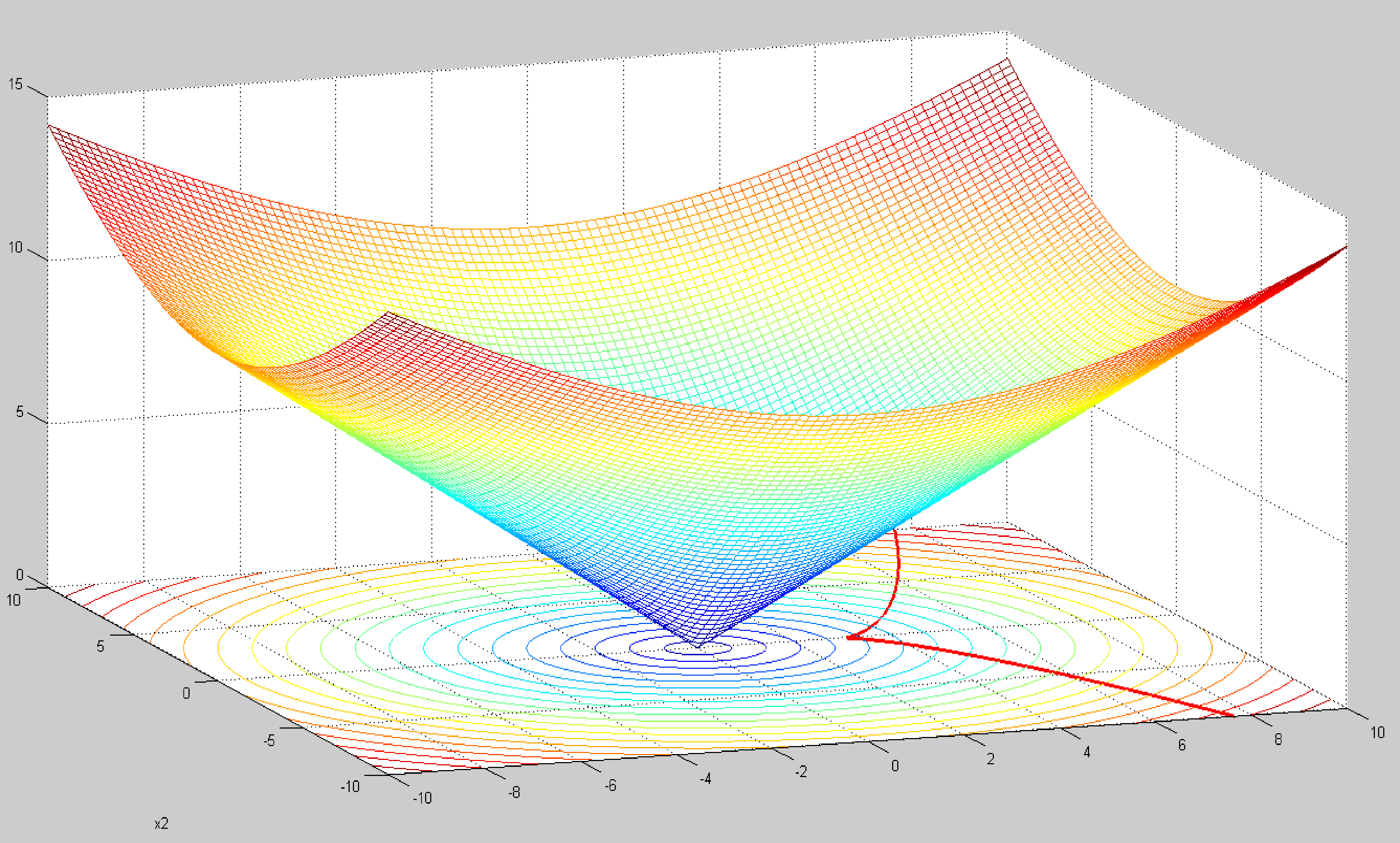
$\left[\frac{\partial g}{\partial x}\right]$为$n\times m$矩阵。若找到一个解$x^0$，满足$g(x^0)=0$,如果此解为最优解，则 λ应满足：

$\left[\frac{\partial g}{\partial x}\right]_{x^0}\lambda=-\nabla f(x^0)$,此解存在的条件是：$\left[\frac{\partial g}{\partial x}\right]$的秩为m

- **例6_3**
  - $\min \quad f(x) = \sqrt{x^2 + y^2}$

$s.t \quad y^2 - (x-1)^3 = 0$

- **例6_3**



- **例6_3**

  - $\min\ f(x)=\sqrt{x^2+y^2}$

    $s.t\quad y^2-(x-1)^3=0$

$$L=\sqrt{x^2+y^2}+\lambda(y^2-(x-1)^3)$$

$$\frac{\partial L}{\partial x}=(x^2+y^2)^{-1/2}\cdot\frac{1}{2}\cdot 2x+(-3)(x-1)^2\lambda=0$$

$$\frac{\partial L}{\partial y}=(x^2+y^2)^{-1/2}\cdot\frac{1}{2}\cdot 2y+2y\lambda=0$$

$$\frac{\partial L}{\partial \lambda}=y^2-(x-1)^3=0$$

若将最优点（1，0）代入上式存在问题。Why?

$$\frac{\partial g}{\partial X}=\begin{bmatrix}-3(x-1)^2\\ 2y\end{bmatrix}\Big|_{(1,0)}=\begin{bmatrix}0\\0\end{bmatrix},\ \text{rank}\left(\frac{\partial g}{\partial X}\right)=1$$

- 不等式约束的非线性规划
  - $\min f(x)$
  - $s.t. \quad h_j(x) \geq 0 \qquad j = 1,2,\ldots,m < n$

  - 基本思路：将不等式约束问题转化为等式约束问题。
  - 使用松驰变量$\theta_j^2 = h_j(x) \geq 0$

拉格朗日函数(不等式约束)

- $L(x,\lambda,\theta) = f(x) + \sum_{j=1}^{m} \lambda_j (h_j(x) - \theta_j^2)$

- $\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial h_j(x)}{\partial x_i} = 0 \qquad i = 1, \dots, n$

- $\frac{\partial L}{\partial \lambda_j} = h_j(x) - \theta_j^2 = 0 \qquad j = 1, \dots, m$

- $\frac{\partial L}{\partial \theta_j} = -2\lambda_j \theta_j = 0 \qquad j = 1, \dots, m$

- 由最后一式可知：$\lambda_j^* = 0$或 $\theta_j^* = 0$或两者都为0

- **拉格朗日函数(不等式约束)**
  - 情形I：$\lambda_j^* = 0$且 $\theta_j^* \neq 0$，故$h_j(x^*) = (\theta_j^*)^2 > 0$,故最优解在可行域内部。若所有 $\lambda_j^* = 0$，等价于无约束情形。
  - 情形II：$\lambda_j^* \neq 0$且 $\theta_j^* = 0$，故 $h_j(x^*)$=0， 故最优解在可行域边界。由于 $\lambda_j^* \neq 0$，故不满足$\nabla f(x^*) = 0$.
  - 情形III：$\lambda_j^* = 0$且 $\theta_j^* = 0$, 故 $h_j(x^*)$=0,且 $\nabla f(x^*) = 0$，故约束条件的边界穿过无约束问题的最优解。

- **例6_4**
  - $\boldsymbol{min \quad f(x) = (x-a)^2 + b}$

$\boldsymbol{s.t \quad x \geq c}$

- **例6_4**
  - $\min \quad f(x) = (x-a)^2 + b$

$s.t \quad x \geq c$

令 $\theta^2 = x - c \geq 0$,得到:

$$L = (x-a)^2 + b + \lambda(x - c - \theta^2)$$

  - $\dfrac{\partial L}{\partial x} = 2(x-a) + \lambda = 0$
  - $\dfrac{\partial L}{\partial \lambda} = x - c - \theta^2 = 0$
  - $\dfrac{\partial L}{\partial \theta} = -2\lambda\theta = 0$

- **例6_4**
  - $2(x-a)+\lambda=0$
  - $x-c-\theta^2=0$
  - 情形I：$\lambda=0, x^*=a, \theta^2=a-c$.若 $c<a$, θ为实值.
  - 情形II： $\theta=0, x^*=c, \lambda^*=-2(c-a)$
  - 情形III：$\lambda=0$， $\theta=0$， $x^*=a=c$

- **作业(例6-5)**
  - 求极大和极小
  - $f(x) = 2x^2 - 3y^2 - 2x$

$s.t \quad x^2 + y^2 \leq 1$

- **作业(6_5)**
  - 求极大和极小
  - $f(x) = 2x^2 - 3y^2 - 2x$

$s.t \quad x^2 + y^2 \leq 1$

令：$\theta^2 = 1 - x^2 - y^2$

$$L = 2x^2 - 3y^2 - 2x + \lambda(\theta^2 - 1 + x^2 + y^2)$$

  - $\frac{\partial L}{\partial x} = 4x - 2 + 2\lambda x = 0$
  - $\frac{\partial L}{\partial y} = -6y + 2\lambda y = 0$
  - $\frac{\partial L}{\partial \lambda} = \theta^2 - 1 + x^2 + y^2 = 0$
  - $\frac{\partial L}{\partial \theta} = 2\lambda\theta = 0$

- **作业1**
  - $1)\ \theta = 0$
  - $4x - 2 + 2\lambda x = 0$
  - $-6y + 2\lambda y = 0$
  - $-1 + x^2 + y^2 = 0$

解得：

$$\begin{pmatrix} 0.2 \\ \pm\sqrt{0.96} \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ -3 \end{pmatrix}$$

-3.2，0，4

- **作业1**
  - 2) $\lambda = 0$
  - $4x - 2 = 0$
  - $-6y = 0$
  - $-1 + x^2 + y^2 = \theta^2$

解得:

$$\begin{pmatrix} 0.5 \\ 0 \\ \pm\sqrt{1.25} \end{pmatrix}$$

-0.5

- **作业1**
  - $3)\ \lambda = 0, \theta = 0$
  - $4x - 2 = 0$
  - $-6y = 0$
  - $-1 + x^2 + y^2 = 0$

无解。

故：极大值为4，极小值为-3.2