最优化理论与方法

第2讲 微分

无约束问题：

$$\min\ f(x), \qquad x \in R^n \tag{1.4}$$

约束问题：

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \geq 0, i \in I = \{1,2,\cdots,m_1\} \\ & h_j(x) = 0, j \in E = \{m_1+1,\cdots,\mathrm{m}\} \end{aligned} \tag{1.5}$$

约束函数：g_i, h_j

不等式约束条件：$g_i(x) \geq 0$

等式约束条件：$h_j(x) = 0$

可行域：$D = \{x : g_i(x) \geq 0, i \in I; h_j(x) = 0, j \in E\}$

满足(1.4)或(1.5)的x称为最优解或解。最优化就是研究关于解的理论及如何求解

最优解的概念

设 $x^* \in D$

定义5.1.1

x^*是问题(1.1)的局部最优解：若存在邻域$U_\delta(x^*)$,使得

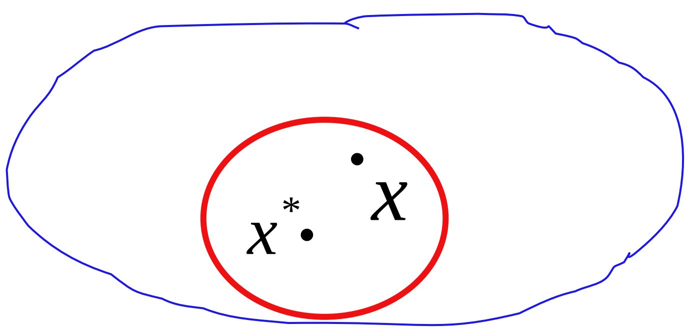
$$f(x^*) \le f(x), \qquad \forall x \in U_\delta(x^*) \bigcap D \qquad (1.6)$$

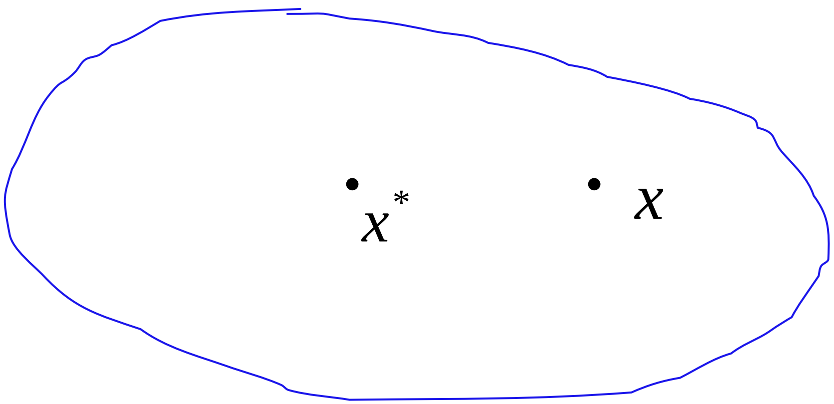
如果(1.6)成立严格不等式，则称x^*是问题(1.1)的严格局部最优解

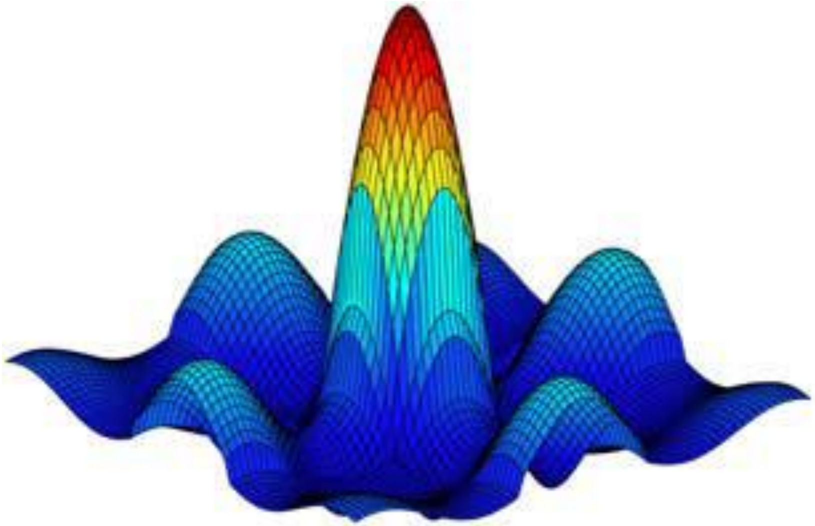
x^*是问题(1.1)的全局(整体)最优解：若

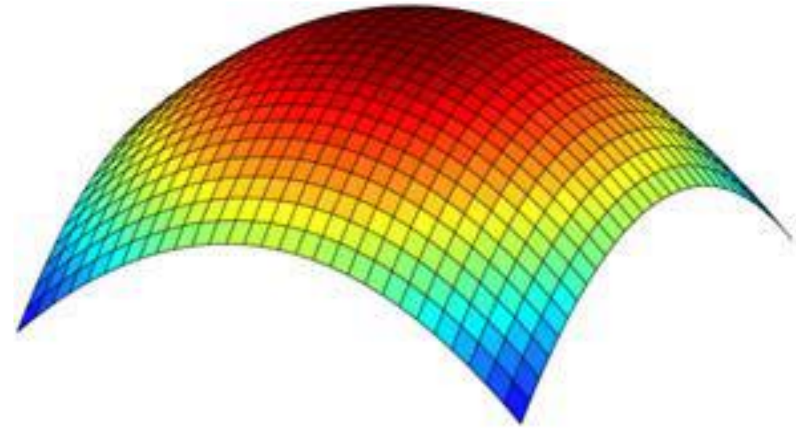
$$f(x^*) \le f(x), \qquad \forall x \in D \qquad (1.7)$$

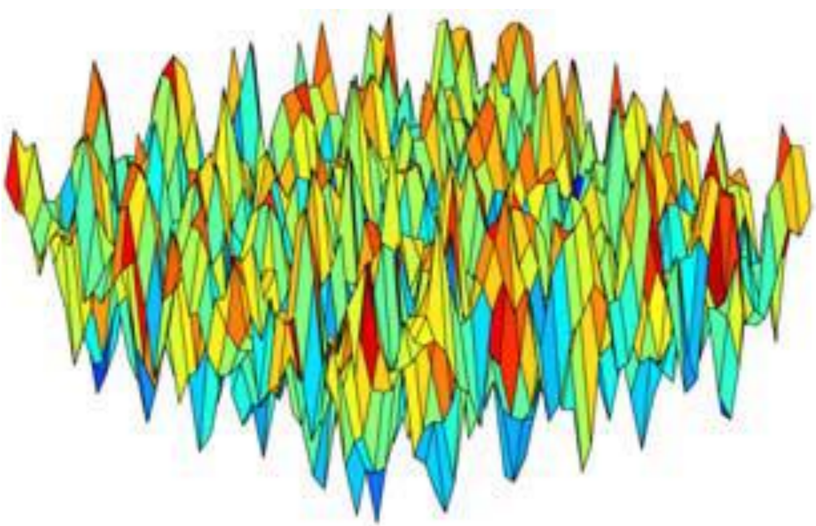
如果(1.7)成立严格不等式$(x \ne x^*)$，则称x^*是问题(1.1)的严格全局最优解

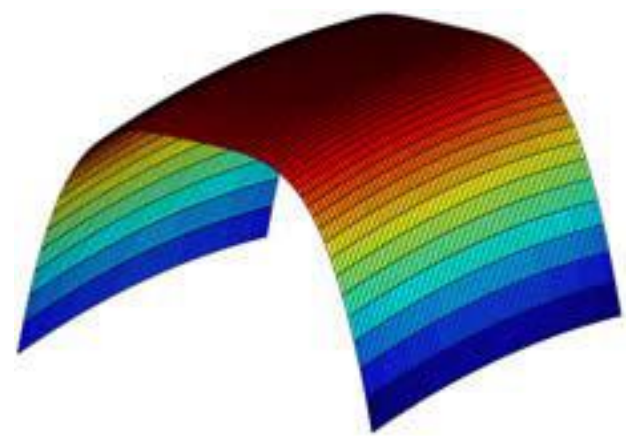












三类特殊的优化问题：

线性规划：g_i, h_j线性函数，f线性函数

二次规划：g_i, h_j线性函数，f二次函数

特别地：

凸规划：f凸函数，D凸集（有关概念后面介绍）

关于多元函数的微分

设$f:R^n\to R$是二次连续可微函数，在x处

梯度：

$$\nabla f(x)=\begin{pmatrix}\dfrac{\partial f(x)}{\partial x_1}\\ \vdots\\ \dfrac{\partial f(x)}{\partial x_n}\end{pmatrix}$$

Hessian矩阵：

$$\nabla^2 f(x)=\begin{pmatrix}\dfrac{\partial^2 f(x)}{\partial x_1{}^2} & \dfrac{\partial^2 f(x)}{\partial x_1\partial x_2}\cdots & \dfrac{\partial^2 f(x)}{\partial x_1\partial x_n}\\ \vdots & & \vdots\\ \dfrac{\partial^2 f(x)}{\partial x_n\partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_n\partial x_2}\cdots & \dfrac{\partial^2 f(x)}{\partial x_n{}^2}\end{pmatrix}$$

思考

(1) $\nabla(b^T x) = b$, 其中b是n维常向量

(2) $\nabla c = O$, 其中c是n维常向量，O为n阶零矩阵

(3) $\nabla^2(x^T Qx) = Q + Q^T$, 其中$Q$为$n$阶矩阵

(4) 设$\phi(t) = f(x+tp)$, 其中$f: R^n \to R$, $\phi: R \to R$, 则

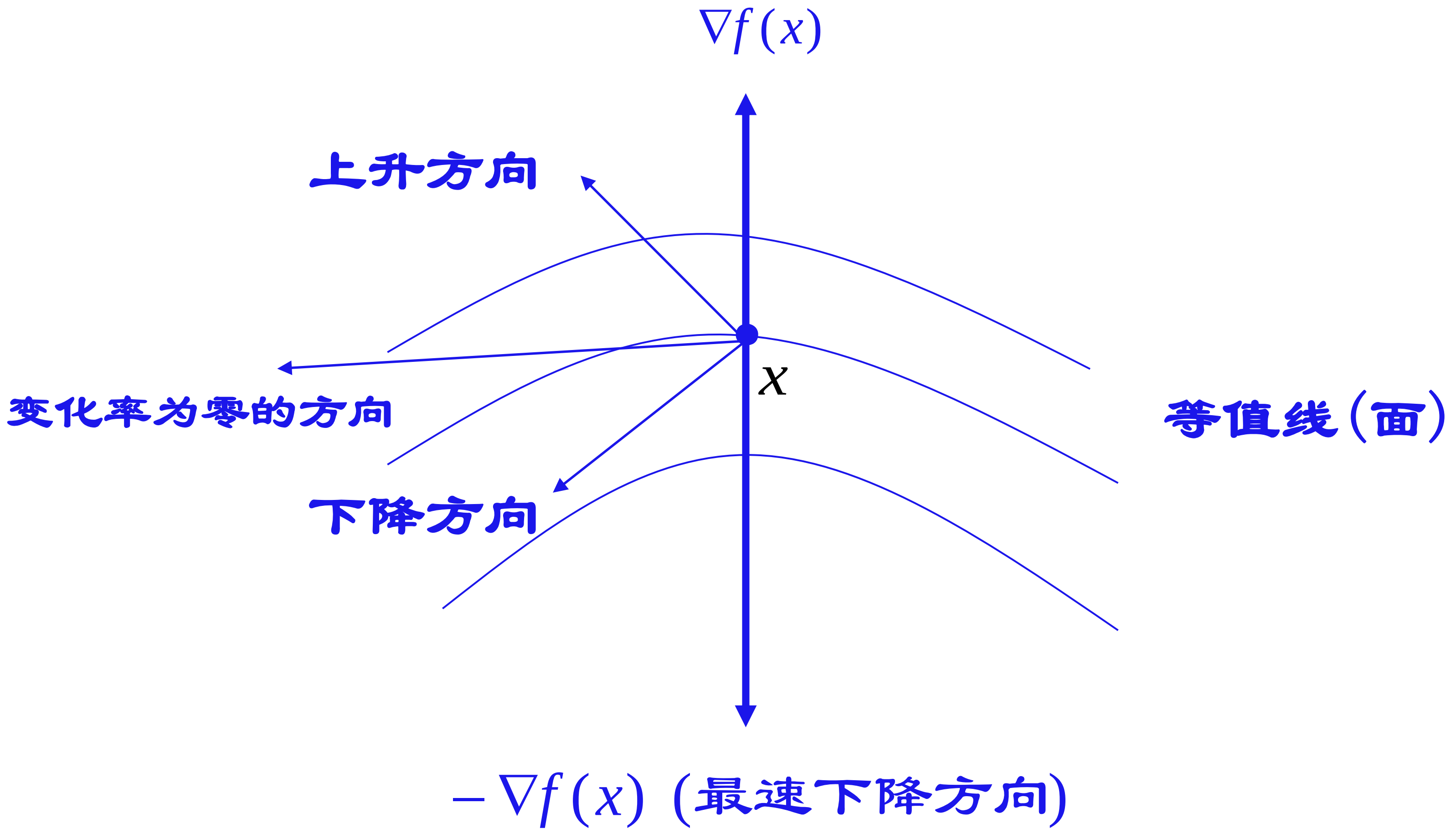
$$\phi'(t) = \nabla f(x+tp)^{\mathrm{T}} p$$

$$\phi''(t) = p^T \nabla^2 f(x+tp) p$$

如何证明？

利用：多元函数微分、一元函数导数

梯度的意义



多元函数的一阶Taylor展开式：

$$f(y) = f(x) + \nabla f(x)^T (y-x) + o(\| y-x \|)$$
$$= f(x) + \nabla f[x+\theta(y-x)]^T (y-x)$$
$$= f(x) + \int_0^1 \nabla f[x+\tau(y-x)]^T (y-x) d\tau$$

如何证明？　　利用一元函数中值定理

$$\phi(t) = f(x+t(y-x))$$

$$\phi'(t) = \nabla f(x+tp)^{\mathrm{T}} p$$

多元函数的二阶Taylor展开式

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) + o(\| y-x \|^2)$$

$$= f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f[x+\theta(y-x)](y-x)$$

$$= f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \int_0^1 \nabla^2 f[x+\tau(y-x)](1-\tau) d\tau (y-x)$$

令 $y = x + tp, p \in R^n, t \in R$, 则有

$$f(x+tp) = f(x) + t\nabla f(x)^T p + \frac{t^2}{2} p^T \nabla^2 f(x) p + o(\| tp \|^2)$$

向量函数的微分

设$F=(F_1,F_2,\cdots,F_m)^T:R^n\to R^m$连续可微

Jacobi矩阵：$F'(x)=\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \vdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}=JF(x)$

梯度：$\nabla F(x):=(\nabla F_1(x),\nabla F_2(x),\cdots,\nabla F_m(x))=F'(x)^{\mathrm{T}}$

向量值函数的一阶Taylor展开式：

设$F=(F_1,F_2,\cdots,F_m)^T:R^n\to R^m$连续可微，其Jacobi矩阵

$$F'(x)=(\nabla F_1(x),\nabla F_2(x),\cdots,\nabla F_m(x))^T$$

$$F(y)=F(x)+F'(x)(y-x)+o(\|y-x\|)$$

$$=F(x)+\int_0^1 F'[x+\tau(y-x)](y-x)d\tau$$

$$\neq F(x)+F'[x+\theta(y-x)](y-x)$$

思考

(1) $\nabla x = I$,其中x是n维向量，I为n阶单位矩阵

(2) $\nabla (\mathrm{Q}x) = Q^T$,其中$Q$为$n$阶矩阵

常用的梯度或Hessian矩阵公式：

(1) $\nabla c = O$, 其中c是n维常向量，O为n阶零矩阵

(2) $\nabla x = I$, 其中x是n维向量，I为n阶单位矩阵

(3) $\nabla(b^T x) = b$, 其中b是n维常向量

(4) $\nabla(Qx) = Q^T$, 其中Q为n阶矩阵

(5) 设$\phi(t) = f(x+tp)$, 其中$f: R^n \to R, \phi: R \to R$, 则

$$\phi'(t) = \nabla f(x+tp)^{\mathrm{T}} p$$

$$\phi''(t) = p^T \nabla^2 f(x+tp) p$$

(4) $\nabla^2(x^T Qx) = Q + Q^T$, 其中$Q$为$n$阶矩阵

链式法则

复合函数 $h(x) = f(g(x))$,其中向量函数f(g)和$g(x)$可微,则

$$h'(x) = f'(g(x))g'(x)$$

由于 $f' = \nabla f^T$,因此

$$\nabla h(x) = \nabla g(x) \nabla f(g(x))$$

练习：

设有复合函数$h(x)=f(u(x))$,其中

$$f(u)=\begin{bmatrix} f_1(u) \\ f_2(u) \end{bmatrix}=\begin{bmatrix} u_1^2-u_2 \\ u_1+u_2^2 \end{bmatrix}, \quad u(x)=\begin{bmatrix} u_1(x) \\ u_2(x) \end{bmatrix}=\begin{bmatrix} x_1+x_3 \\ x_2^2-x_3 \end{bmatrix}.$$

试求复合函数$h(x)=f(u(x))$的导数.

例1 已知函数 $f(x)=\frac{1}{2}x_1^2+x_2^2-x_1x_2-x_1$，求$\nabla f(x),\nabla^2 f(x)$

例2. 已知 $f(x)=\frac{1}{2}x^TQx+q^Tx$，求$\nabla f(x),\nabla^2 f(x)$.

例3 函数 $f(x)=\frac{1}{2}\sum_{i=1}^{m}F_i(x)^2=\frac{1}{2}\|F(x)\|^2$，求$\nabla f(x),\nabla^2 f(x)$

例4 函数 $f(d)=\frac{1}{2}\|F(x^k)+F'(x^k)d\|^2$，求$\nabla f(d),\nabla^2 f(d)$

作 业

- 习题1(李董辉等《最优化理论和方法》)

  2，3（i）