范数、可微性、展开

• 向量范数

若实值函数 $\|\cdot\|:\mathbb{R}^n \to \mathbb{R}$ 满足

$$ullet$$
 $\forall oldsymbol{x} \in \mathbb{R}^n, \|oldsymbol{x}\| \geq 0 oldsymbol{eta} \|oldsymbol{x}\| = 0 \Longleftrightarrow oldsymbol{x} = oldsymbol{0}$

$$\bullet \ \forall \alpha \in \mathbb{R}, \forall \boldsymbol{x} \in \mathbb{R}^n, \|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|$$

$$ullet \ \ \ orall oldsymbol{x}, oldsymbol{y} \in \mathbb{R}^n, \|oldsymbol{x} + oldsymbol{y}\| \leq \|oldsymbol{x}\| + \|oldsymbol{y}\|$$

则称其为向量范数, 其中 \mathbb{R}^n 表示 n 维向量空间

设
$$\boldsymbol{x}=\left(x_{1},x_{2},\cdots,x_{n}\right)^{T}\in\mathbb{R}^{n}$$

$$\circ$$
 L_1 范数: $\|oldsymbol{x}\|_1 = \sum\limits_{i=1}^m |x_i|$

$$ullet$$
 $ullet$ L_2 范数: $\|oldsymbol{x}\|_2 = \left(\sum\limits_{i=1}^n x_i^2
ight)^{rac{1}{2}}$

。
$$L_{\infty}$$
 范数: $\|oldsymbol{x}\|_{\infty} = \max_{i} |x_{i}|$

。
$$L_p$$
 范数: $\left\|oldsymbol{x}
ight\|_p = \left(\sum\limits_{j=1}^n \left|x_i
ight|^p
ight)^{rac{1}{p}}$

• 矩阵范数

Frobenius 范数:
$$\left\|oldsymbol{A}
ight\|_F = \left(\sum\limits_{i=1}^n\sum\limits_{j=1}^n\left|a_{ij}
ight|^2
ight)^{rac{1}{2}} = \left(\operatorname{tr}\left(oldsymbol{A}^Toldsymbol{A}
ight)
ight)^{rac{1}{2}}$$

$$m{A} \in \mathbb{R}^{m imes n}$$
,诱导的矩阵范数为 $\|m{A}\| = \max_{x
eq 0} rac{\|m{A}m{x}\|}{\|m{x}\|}$

设
$$oldsymbol{A} = (a_{ij})_{m imes n}, \; \lambda_{oldsymbol{A}^Toldsymbol{A}}$$
 表示 $oldsymbol{A}^Toldsymbol{A}$ 的最大特征值

$$egin{aligned} & \left\|oldsymbol{A}
ight\|_1 = \max_j \sum_{i=1}^m \left|a_{ij}
ight| \ & \left\|oldsymbol{A}
ight\|_2 = \sqrt{\lambda_{oldsymbol{A}^Toldsymbol{A}}} \end{aligned}$$
(谱范数)

。
$$\|oldsymbol{A}\|_2 = \sqrt{\lambda_{oldsymbol{A}^Toldsymbol{A}}}$$
 (谱范数)

若矩阵 $m{A}$ 正定,则条件数 $\mathrm{cond}(m{A}) = \|A\| \|A^{-1}\|$

$$\circ$$
 在 L_2 范数定义下, $\operatorname{cond}(oldsymbol{A}) = rac{\lambda_{\max}(oldsymbol{A})}{\lambda_{\min}(oldsymbol{A})}$

○ 条件数是判断矩阵病态与否的一种度量,条件数越大矩阵越病态

梯度

$$abla f(oldsymbol{x}) = \left[rac{\partial f(oldsymbol{x})}{\partial x_1}, rac{\partial f(oldsymbol{x})}{\partial x_2}, \cdots, rac{\partial f(oldsymbol{x})}{\partial x_n}
ight]^T$$

f 在点 x^0 关于 p 的方向导数 $rac{\partial f(x^0)}{\partial p} = \lim_{t o 0^+} rac{f(x^0+tp)-f(x^0)}{t}$

$$\circ \nabla \boldsymbol{b}^T \boldsymbol{x} = \boldsymbol{b}$$

。
$$abla oldsymbol{x}^T oldsymbol{A} oldsymbol{x} = ig(oldsymbol{A} + oldsymbol{A}^Tig) oldsymbol{x}$$

$$ullet \ ullet \ egin{pmatrix} igl(rac{1}{2}oldsymbol{x}^Toldsymbol{A}oldsymbol{x} + oldsymbol{b}^Toldsymbol{x} + oldsymbol{c} \end{pmatrix} = oldsymbol{A}oldsymbol{x} + oldsymbol{b}$$

• Jacobi 矩阵

$$m{h'}(m{x}) =
abla m{h}(m{x})^T = egin{pmatrix} rac{\partial h_1(m{x})}{\partial x_1} & rac{\partial h_1(m{x})}{\partial x_2} & \cdots & rac{\partial h_1(m{x})}{\partial x_n} \ rac{\partial h_2(m{x})}{\partial x_1} & rac{\partial h_2(m{x})}{\partial x_2} & \cdots & rac{\partial h_2(m{x})}{\partial x_n} \ dots & dots & dots & dots \ rac{\partial h_m(m{x})}{\partial x_1} & rac{\partial h_m(m{x})}{\partial x_2} & \cdots & rac{\partial h_m(m{x})}{\partial x_n} \end{pmatrix}$$

• Hessian 矩阵

$$abla^2 f(oldsymbol{x}) = egin{pmatrix} rac{\partial^2 f(oldsymbol{x})}{\partial x_1^2} & rac{\partial^2 f(oldsymbol{x})}{\partial x_1 x_2} & \cdots & rac{\partial^2 f(oldsymbol{x})}{\partial x_1 x_n} \ rac{\partial^2 f(oldsymbol{x})}{\partial x_2 x_1} & rac{\partial^2 f(oldsymbol{x})}{\partial x_2^2} & \cdots & rac{\partial^2 f(oldsymbol{x})}{\partial x_2 x_n} \ dots & dots & dots & dots \ rac{\partial^2 f(oldsymbol{x})}{\partial x_n x_1} & rac{\partial^2 f(oldsymbol{x})}{\partial x_n x_2} & \cdots & rac{\partial^2 f(oldsymbol{x})}{\partial x_n^2} \end{pmatrix}$$

• Taylor 展开

。 一阶展开

 $f:\mathbb{R}^n o \mathbb{R}$ 连续可微,向量 $oldsymbol{p} \in \mathbb{R}^n$

$$f(\boldsymbol{x} + \boldsymbol{p}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{p} + o(\|\boldsymbol{p}\|)$$

。 二阶展开

 $f:\mathbb{R}^n o \mathbb{R}$ 二阶连续可微

$$f(oldsymbol{x} + oldsymbol{p}) = f(oldsymbol{x}) +
abla f(oldsymbol{x}) +
abla f(oldsymbol{x})^T oldsymbol{p} + rac{1}{2} oldsymbol{p}^T
abla^2 f(oldsymbol{x}) oldsymbol{p} + o\left(\|oldsymbol{p}\|
ight)$$

凸集

• 凸集的定义

。 仿射集

设 S 为 n 维欧氏空间 \mathbb{R}^n 中的一个集合。若对任意两点 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S$ 及每个实数 $\lambda \in \mathbb{R}$,有 $\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)} \in S$,则称 S 为仿射集。 $\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}$ 称为 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ 的仿射组合

- 泛化定义: $\forall x_1, x_2, \cdots, x_k \in S, \lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{R}, \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$, 有 $\lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k \in S$
- 。 仿射包

$$ext{aff} S = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k \left| egin{array}{c} orall x_1, x_2, \cdots, x_k \in S \ \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1 \end{array}
ight\}$$

。 凸集

设 S 为 n 维欧氏空间 \mathbb{R}^n 中的一个集合。若对任意两点 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S$ 及每个实数 $\lambda \in [0,1]$,有 $\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)} \in S$,则称 S 为凸集。 $\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}$ 称为 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ 的凸组合

。 凸包

$$\mathrm{conv} C = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k \left| egin{array}{l} orall x_1, x_2, \cdots, x_k \in S \ \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1 \ \lambda_1, \lambda_2, \cdots, \lambda_k \in [0, 1] \end{array}
ight.
ight.$$

○ 锥

设有集合 $C\subset\mathbb{R}^n$, 若对没一点 $x\in C$, 当 $\lambda>0$ 时, 都有 $\lambda x\in C$, 称 C 为锥

。 凸锥

任意两点 $m{x}^{(1)},m{x}^{(2)}\in S$ 及实数 $\lambda_1,\lambda_2\geq 0$ 都有 $\lambda_1m{x}^{(1)}+\lambda_2m{x}^{(2)}\in C$,称 C 为凸锥

• 典型凸集

。 超平面

 $H = \{ \boldsymbol{x} \mid \boldsymbol{p}^T \boldsymbol{x} = \alpha \}$,其中 \boldsymbol{p} 为 n 维列向量, α 为实数

■ 半空间

正的闭半空间
$$H^+ = \left\{ m{x} \mid m{p}^T m{x} \geq lpha
ight\}$$

负的闭半空间 $H^- = \left\{ m{x} \mid m{p}^T m{x} \leq lpha
ight\}$
正的开半空间 $\dot{H}^+ = \left\{ m{x} \mid m{p}^T m{x} > lpha
ight\}$

负的开半空间 $\dot{H}^- = \left\{ oldsymbol{x} \mid oldsymbol{p}^T oldsymbol{x} < lpha
ight\}$

。 多面体

有限个线性等式和不等式的解集

有限个半空间与超平面的交集

。 多胞形

有限点集 $\left\{oldsymbol{x}^0,oldsymbol{x}^1,\cdots,oldsymbol{x}^m
ight\}\subset\mathbb{R}^n$ 的凸包称为多胞形

o 单纯形

 $\left\{m{x}^0,m{x}^1,\cdots,m{x}^m
ight\}$ 仿射无关, $m{x}^1-m{x}^0,\cdots,m{x}^m-m{x}^0$ 线性无关,对应的凸包称为 m 维单纯形,向量 $m{x}^i$ 称为该单纯形的顶点

。 极点

非空凸集中的点 $m{x}$ 称为极点,若 $m{x}=\lambdam{x}_1+(1-\lambda)m{x}_2, \lambda\in(0,1), m{x}_1, m{x}_2\in S$,则 $m{x}=m{x}_1=m{x}_2$

• 集合的保凸运算

。 保凸运算

设 S_1, S_2 为 \mathbb{R}^n 中的两个凸集, β 为实数, 则

- $\beta S_1 = \{ \beta \boldsymbol{x} \mid \boldsymbol{x} \in S_1 \}$ 为凸集
- $S_1 + \beta = \{ \boldsymbol{x} + \beta \mid \boldsymbol{x} \in S_1 \}$ 为凸集
- $S_1 \cap S_2$ 为凸集
- ullet $S_1+S_2=\{oldsymbol{x}^{(1)}+oldsymbol{x}^{(2)}\midoldsymbol{x}^{(1)}\in S_1,oldsymbol{x}^{(2)}\in S_2\}$ 为凸集
- ullet $S_1-S_2=\{oldsymbol{x}^{(1)}-oldsymbol{x}^{(2)}\midoldsymbol{x}^{(1)}\in S_1,oldsymbol{x}^{(2)}\in S_2\}$ 为凸集
- 。 集合的保凸运算

函数 $f:\mathbb{R}^n \to \mathbb{R}^m$ 是仿射的,当 $f(x)=Ax+b, A\in\mathbb{R}^{m\times n}, b\in\mathbb{R}^m$,若 $S\in\mathbb{R}^n$ 为凸集,则 $f(S)=\{f(x)\mid x\in S\}$ 为凸

凸函数

• 凸函数定义

设 $S \subset \mathbb{R}^n$ 是非空凸集,函数 $f: S \to \mathbb{R}$,若对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$,每一 $\lambda \in (0,1)$ 都有 $f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda)f(\boldsymbol{x}_2)$ 则称 $f \in S$ 上的凸函数

若对于 $x_1 \neq x_2$ 严格成立,则称 $f \in S$ 上的严格凸函数

。 一阶条件定理

设 $S\subset\mathbb{R}^n$ 是非空开凸集,函数 $f:S\to\mathbb{R}$ 是可微函数,则 $f(\boldsymbol{x})$ 是凸函数当且仅当对任意的 $\boldsymbol{x}^*\in S$,有 $\forall \boldsymbol{x}\in S, f(\boldsymbol{x})\geq f(\boldsymbol{x}^*)+\nabla f(\boldsymbol{x}^*)^T(\boldsymbol{x}-\boldsymbol{x}^*)$

 $f(oldsymbol{x})$ 严格凸当且仅当 $orall oldsymbol{x} \in S, f(oldsymbol{x}) > f(oldsymbol{x}^*) +
abla f(oldsymbol{x}^*)^T (oldsymbol{x} - oldsymbol{x}^*)$

■ 推广

设 $S\subset\mathbb{R}^n$ 是非空开凸集,函数 $f:S\to\mathbb{R}$ 是可微函数,则 $f({m x})$ 是凸函数当且仅当对任意的 ${m x}_1,{m x}_2\in S$,有 $(\nabla f({m x}_1)-\nabla f({m x}_2))^T({m x}_2-{m x}_1)\geq 0$

$$f(m{x})$$
 严格凸当且仅当对任意的 $m{x}_1,m{x}_2\in S$,有 $\left(
abla f(m{x}_1)-
abla f(m{x}_2)
ight)^T(m{x}_2-m{x}_1)>0$

。 二阶条件定理

设 $S\subset\mathbb{R}^n$ 是非空凸集,函数 $f:S\to\mathbb{R}$ 的二次可微函数,则 $f(\boldsymbol{x})$ 是凸函数当且仅当 S 上每一点的 Hessian 矩阵是半正定的

f(x) 严格凸当且仅当 S 上每一点的 Hessian 矩阵是正定的

。 切西瓜定理

设 $S\subset\mathbb{R}^n$ 是非空凸集,函数 $f:S\to\mathbb{R}$ 的可微函数,则 $f(\boldsymbol{x})$ 是凸函数当且仅当对任意 $\boldsymbol{x}\in S$,任意的方向 $\boldsymbol{v}\in\mathbb{R}^n$,关于 t 的一元函数 $g(t)=f(\boldsymbol{x}+t\boldsymbol{v})$ 为凸函数,其中 $\mathrm{dom}g=\{t\mid \boldsymbol{x}+t\boldsymbol{v}\in\mathrm{dom}f\}$

• 常见函数的凹凸性

- 函数的保凸运算
 - \circ 非负加权和: $f = w_1 f_1 + \cdots + w_m f_m$
 - 仿射映射: g(x) = f(Ax + b)
 - 逐点最大: $f(x) = \max\{f_1(x), \dots, f_m(x)\}$
 - \circ 函数透视: g(x,t)=tf(x/t)
 - 函数共轭: $f^*(y) = \sup_{x \in \text{dom} f} (y^T x f(x))$
 - 复合函数: f(x) = h(g(x)), 若 h 非减,则凹凸性一致; 若 h 非增,则凹凸性相反
- 凸函数与凸集的关系
 - 。 凸函数 f 的水平集 $L(f,\alpha)=\{m{x}\mid m{x}\in S, f(m{x})\leq \alpha\}, \alpha\in\mathbb{R}$ 和上镜图 $\mathrm{epi}(f)=\{(x,y)\mid x\in S, y\in\mathbb{R}, y\geq f(x)\} \text{ 都是凸集}$
 - 。 拟凸函数:若函数 $f: \mathbb{R}^n \to \mathbb{R}$ 所有的 α -sublevel sets $S_{\alpha} \{ \boldsymbol{x} \mid f(\boldsymbol{x}) \leq \alpha, \boldsymbol{x} \in \mathrm{dom} f \}$ 都为凸集合,则 f 为拟凸函数

拟凹函数:若函数 $f:\mathbb{R}^n \to \mathbb{R}$ 所有的 lpha-superlevel sets

 $S_{lpha}\left\{m{x}\mid f(m{x})\geq lpha,m{x}\in\mathrm{dom}f
ight\}$ 都为凸集合,则 f 为拟凹函数

若 $\mathrm{dom}f$ 为凸集且 $f(heta oldsymbol{x} + (1- heta)oldsymbol{y}) \leq \max{\{f(oldsymbol{x}),f(oldsymbol{y})\}}$,则 f 为拟凸函数

凸优化问题

• 凸规划

○ 目标函数: 凸函数

可行域: 凸集

。 对凸问题,任何局部最优解都是全局最优解

• 等价形式

。 变量替换

$$m{x} = \phi(m{z}), ilde{f}_i(m{z}) = f_i(\phi(m{z}))$$
 等价于 minimize $ilde{f}_0(m{z})$ subject to $ilde{f}_i(m{z}) \leq 0$ $i=1,\cdots,m$ $ilde{h}_i(m{z}) = 0$ $i=1,\cdots,p$

。 消除/引入等式约束

$$egin{aligned} & \min_{m{x}} & f_0(m{x}) \ & ext{subject to} & f_i(m{x}) \leq 0 \quad i=1,\cdots,m \ & m{A}m{x} = m{b} \ & m{A}m{x} = m{b} \leftrightarrow m{x} = m{F}m{z} + m{x}_0 \end{aligned}$$

$$egin{array}{ll} ext{minimize} & f_0(oldsymbol{F}oldsymbol{z}+oldsymbol{x}_0) \ ext{subject to} & f_i(oldsymbol{F}oldsymbol{z}+oldsymbol{x}_0) \leq 0 & i=1,\cdots,m \end{array}$$

。 引入松弛变量

$$egin{aligned} & \min _{oldsymbol{x}} & f_0(oldsymbol{x}) \ & ext{subject to} & oldsymbol{a}_i^T oldsymbol{x} \leq b_i \quad i=1,\cdots,m \ & rac{oldsymbol{x}}{s} & f_0(oldsymbol{x}) \ & ext{subject to} & oldsymbol{a}_i^T oldsymbol{x} + s_i \leq b_i \quad i=1,\cdots,m \ & s_i \geq 0 & i=1,\cdots,m \end{aligned}$$

。 目标函数放到约束条件

$$egin{array}{ll} ext{minimize} & f_0(oldsymbol{x}) \ ext{subject to} & f_i(oldsymbol{x}) \leq 0 & i=1,\cdots,m \ & oldsymbol{A}oldsymbol{x} = oldsymbol{b} \end{array}$$

等价于

$$egin{array}{ll} ext{minimize} & t \ ext{subject to} & f_0(oldsymbol{x}) - t \leq 0 \ & f_i(oldsymbol{x}) \leq 0 & i = 1, \cdots, m \ & oldsymbol{A} oldsymbol{x} = oldsymbol{b} \end{array}$$

。 先优化某些变量

$$egin{aligned} & \min_{m{x},m{y}} & f_0(m{x},m{y}) \ & ext{subject to} & f_i(m{x}) \leq 0 \quad i=1,\cdots,m \ & \inf_{m{x},m{y}} f(m{x},m{y}) = \inf_{m{x}} ilde{f}(m{x}) \ & ilde{f}(m{x}) = \inf_{m{y}} f(m{x},m{y}) \end{aligned}$$

$$egin{aligned} & \min_{m{x}} & ilde{f}_0(m{x}) \ & ext{subject to} & f_i(m{x}) \leq 0 & i = 1, \cdots, m \end{aligned}$$

• 线性规划 (LP)

。 标准形式:

$$egin{array}{ll} ext{minimize} & oldsymbol{c}^T oldsymbol{x} + d \ ext{subject to} & oldsymbol{G} oldsymbol{x} \leq oldsymbol{h} \ oldsymbol{A} oldsymbol{x} = oldsymbol{b} \end{array}$$

。 可行域是一个多面体

二次规划

• 二次规划 (QP)

$$egin{array}{ll} ext{minimize} & rac{1}{2}oldsymbol{x}^Toldsymbol{P}oldsymbol{x}+oldsymbol{q}^Toldsymbol{x}+r \ ext{subject to} & oldsymbol{G}oldsymbol{x} \leq oldsymbol{h} \ & oldsymbol{A}oldsymbol{x} = oldsymbol{b} \end{array}$$

• 二次约束二次规划 (QCQP)

$$\begin{aligned} & \underset{m{x}}{\text{minimize}} & & \frac{1}{2} m{x}^T m{P}_0 m{x} + m{q}_0^T m{x} + r_0 \\ & \text{subject to} & & \frac{1}{2} m{x}^T m{P}_i m{x} + m{q}_i^T m{x} + r_i \leq 0 \quad i = 1, \cdots, m \\ & & & \bm{A} m{x} = m{b} \end{aligned}$$

• 半正定规划 (SDP)

$$egin{array}{ll} & \min _{m{x}} & \operatorname{tr}\left(m{C}m{X}
ight) \ & \operatorname{subject} \ \operatorname{to} & \operatorname{tr}\left(m{A}_{i}m{X}
ight) = b_{i} \ & m{X}\succeq 0 \ & \min _{m{x}} & m{c}^{T}m{x} \ & \operatorname{subject} \ \operatorname{to} & x_{1}m{F}_{1} + \cdots x_{n}m{F}_{n} \preceq m{G} \ & m{A}m{x} = m{b} \end{array}$$

对偶问题

• 拉格朗日

标准形式最优化问题

minimize
$$f_0(x)$$
 subject to $f_i(x) \leq 0$ $i=1,\cdots,m$ $h_i(x)=0$ $i=1,\cdots,p$

$$D = \bigcap_{i=0}^m \mathrm{dom} f_i \cap \bigcap_{i=0}^p \mathrm{dom} h_i$$

拉格朗日:目标函数和约束函数的加权和

$$L(x,\lambda,v)=f_0(x)+\sum\limits_{i=1}^m\lambda_if_i(x)+\sum\limits_{i=1}^pv_ih_i(x)$$

• 对偶函数

$$g:x\in R^m\times R^p\to R$$

$$egin{array}{lll} g(\lambda,v) &=& \inf_{x\in D}L(x,\lambda,v) \ &=& \inf_{x\in D}\left(f_0(x)+\sum\limits_{i=1}^m\lambda_if_i(x)+\sum\limits_{i=1}^pv_ih_i(x)
ight) \end{array}$$

对偶函数总是凹函数

• 对偶问题

 $\underset{x}{\text{minimize}} \quad g(\lambda, v)$

subject to $\lambda \geq 0$

设原问题最优化目标函数值为 p^* , 对偶问题为 d^* , 则 $d^* \leq p^*$

对偶问题总是凸问题

• 弱对偶和强对偶

。 弱对偶

$$d^* \leq p^*$$

对偶间隙 $d^* - p^*$

。 强对偶

$$d^* = p^*$$

Slater 约束规范:对凸优化问题,只要有一个严格可行点,强对偶即成立

最优性条件

• KKT 条件

。 原问题可行

$$f_i(oldsymbol{x}^*) \leq 0, i=1,\cdots,m, h_i(oldsymbol{x}^*) = 0, i=1,\cdots,p$$

。 对偶问题可行

$$\boldsymbol{\lambda}^* \geq 0$$

。 互补松弛条件

$$\forall i, \lambda_i^* f_i(\boldsymbol{x}^*) = 0$$

 \circ 拉格朗日对 x 求导为零

$$abla f_0(oldsymbol{x}^*) + \sum\limits_{i=1}^m \lambda_i
abla f_i(oldsymbol{x}^*) + \sum\limits_{i=1}^p v_i
abla h_i(oldsymbol{x}^*) = 0$$

算法

0.618 法

- 单峰函数: 若 $\exists \alpha^* \in [a,b]$,使得 f(x) 在 $[a,\alpha^*]$ 上严格递减,在 $[\alpha^*,b]$ 上严格递增,则 f(x) 是 [a,b] 上的单峰函数
- 设 $\varphi(\alpha)$ 是搜索区间 $[a_1,b_1]$ 上的单峰函数,设在第 k 次迭代时搜索区间为 $[a_k,b_k]$ 。取两个试探点 $\lambda_k,\mu_k\in[a_k,b_k],\lambda_k<\mu_k$,计算 $\varphi(\lambda_k),\varphi(\mu_k)$

。 若
$$arphi(\lambda_k) \leq arphi(\mu_k)$$
,令 $a_{k+1} = a_k, b_{k+1} = \mu_k$

。 若
$$\varphi(\lambda_k) > \varphi(\mu_k)$$
,令 $a_{k+1} = \lambda_k, b_{k+1} = b_k$

牛顿法

$$x^{(k+1)} = x^{(k)} - rac{f'(x^{(k)})}{f''(x^{(k)})}$$

割线法

$$x^{(k+1)} = x^{(k)} - rac{x^{(k)} - x^{(k-1)}}{f'(x^{(k)}) - f'(x^{(k-1)})} f'\left(x^{(k)}
ight)$$

梯度下降法

- 给定初始点 $x^{(1)}$
- 计算负梯度方向 $d^{(k)} = -\nabla f\left(x^{(k)}\right)$
- 计算步长 λ ,满足 $arphi(\lambda_k)=f\left(x^{(1)}+\lambda_k d^{(k)}
 ight)$, $arphi'(\lambda_k)=0$
- 更新 $x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)}$
- 中止准则: $\|\nabla f(x)\|_2 \leq \epsilon$

牛顿法

$$x^{(k+1)} = x^{(k)} - \left(
abla^2 f\left(x^{(k)}
ight)
ight)^{-1}
abla f\left(x^{(k)}
ight)$$