§lec49~50.

无约束优化问题及算法

$\min f(x)$

gradient descent $d^k = -\nabla f(x^k)$

steepest descent $d^k = \underset{\|v\|=1}{\arg\min} \{\nabla f^T(x^k) v\}$.

coordinate descent

subgradient descent $d^k = -\frac{\partial f(x^k)}{\partial x}$

Newton's Method $d^k = \underset{v}{\arg\min} \{\nabla f^T(x) v + \frac{1}{2} v^T \nabla^2 f(x) v\}$

$= -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$.

Quasi-Newton Method $d^k = -B^{-1} \nabla f(x^k)$

有约束优化问题.

$$\begin{cases} \min & f(x) \\ \text{s.t} & Ax=b \end{cases}$$

KKT条件: $Ax^* = b$

$\nabla f(x^*) + A^T v^* = 0$.

1) 线性方程组.

$\min \ \frac{1}{2} x^T P x + q^T x + r \quad P \in S_+^n$

$\text{s.t} \quad Ax = b$.

KKT条件: $Ax^* = b$

$Px^* + q + A^T v^* = 0$.

$$\Leftrightarrow \begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

2) 非线性方程组.

$\underset{d}{\min} \ f(\underbrace{x^k+d}) \quad x^{k+1}$

$\text{s.t.} \ \underbrace{A(x^k+d)=b}$. $\quad A \cdot d = 0$

$$\Leftrightarrow \underset{d}{\min} \ f(x^k) + \nabla f^T(x^k) d + \frac{1}{2} d^T \nabla^2 f(x^k) d.$$
$$\text{s.t.} \quad Ad = 0.$$

$$\Leftrightarrow \text{KKT条件:} \begin{bmatrix} \nabla^2 f(x^k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} d^* \\ v^* \end{bmatrix} = \begin{bmatrix} -\nabla f(x^k) \\ 0 \end{bmatrix}$$

$$\begin{cases} \alpha^k = \underset{\alpha \geq 0}{\arg\min}\ f(x^k + \alpha d^k) \quad (\text{步长}). \\ x^{k+1} = x^k + \alpha^k d^k \end{cases}$$

拉格朗日法. Lagrangian Method.

$$\begin{cases} x^{k+1} = x^k - \alpha^k (\nabla f(x^k) + A^T v^k). \\ v^{k+1} = v^k + \alpha^k (Ax^k - b) \end{cases}$$

$$L(x,v) = f(x) + v^T (Ax - b)$$

$$(x^*, v^*) = \arg\max_{v} \min_{x} L(x,v)$$

$$(x^*, v^*) = \arg\min_{x} \max_{v} L(x,v).$$

$$\begin{cases} x^* = \arg\min_{x} L(x, v^*). \\ v^* = \arg\max_{v} L(x^*, v). \end{cases}$$

$$-\nabla f(x^k) - A^T v^*$$

$$x^{k+1} = x^k + \alpha^k (-\nabla f(x^k) - A^T \underset{v^k.}{\underline{v^*}}).$$

$$-Ax^* + b$$

$$v^{k+1} = v^k + \alpha^k (A\underset{x^k.}{\underline{x^*}} - b)$$

$$\boxed{\text{KKT条件}: \quad \begin{aligned} &Ax^* = b \\ &\nabla f(x^*) + A^T v^* = 0. \end{aligned}}$$

$$\min P(x,v) = \frac{1}{2}\|Ax - b\|_2^2 + \frac{1}{2}\|\nabla f(x) + A^T v\|_2^2$$

负梯度方向:

$$-\nabla P(x,v)\Big|_{(x^k, v^k)} = -\begin{pmatrix} A^T(Ax^k - b) - \nabla^2 f(x^k)(\nabla f(x^k) + A^T x^k) \\ A(\nabla f(x^k) + A^T v^k \end{pmatrix}$$

拉格朗日法方向:

$$d^k = \begin{pmatrix} -(f(x^k) + A^T v^k) \\ Ax^k - b \end{pmatrix}$$

拉格朗日法与负梯度方向夹角<90°, 是一个下降的方向.

$$\begin{aligned} &(d^k)^T(-\nabla P(x^k, v^k)) \\ &= (\nabla f(x^k) + A^T v^k)^T A^T (Ax^k - b) + (\nabla f(x^k) + A^T v^k)^T \nabla^2 f(x^k)(\nabla f(x^k) + A^T v^k) \\ &\qquad - (A^T x^k - b)^T A (\nabla f(x^k) + A^T v^k) \\ &= (\nabla f(x^k) + A^T v^k)^T \nabla^2 f(x^k)(\nabla f(x^k) + A^T v^k) \end{aligned}$$

$$\begin{cases} \nabla f(x^k) + A^T v^k \neq 0 \\ \nabla^2 f(x^k) \succ 0 \end{cases}$$

增广拉格朗日法. Augmented Lagrangian Method.

$$L_c(x, v) = f(x) + v^T(Ax-b) + \frac{c}{2}\|Ax-b\|_2^2$$ 惩罚项.

$$\begin{cases} \min\ f(x) \\ \text{s.t.}\ \ Ax=b \end{cases} \overset{x^*}{\underset{v^*}{\Longleftrightarrow}} \begin{cases} \min\ f(x) + \frac{c}{2}\|Ax-b\|_2^2 \\ \text{s.t.}\ \ Ax=b \end{cases}$$

最优解相同，当满足约束时，惩罚项为0.