习题

2024年3月11日 9:36

4. Proof. $\forall a_1, a_2 \in S$, $\forall \alpha \in [0,1]$. wlog $|a_1| \geq |a_2|$

$$[\alpha a_1 + (1-\alpha) a_2]^2 = \alpha^2 a_1^2 + 2\alpha(1-\alpha) a_1 a_2 + (1-\alpha)^2 a_2^2$$
$$\leq a_1^2 + (1-\alpha)[-(1+\alpha) a_1^2 + 2\alpha a_1 a_2 + (1-\alpha) a_2^2]$$
$$= a_1^2 + (1-\alpha)[-\alpha(a_1 - a_2)^2 - a_1^2 + a_2^2]$$
$$\leq a_1^2 + (1-\alpha)[-\alpha(a_1 - a_2)^2] \leq a_1^2 \leq 1$$

so $\alpha a_1 + (1-\alpha) a_2 \in S$. i.e. S is convex.

5. $d^{1T} A d^2 = (1 \ \ 0)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = (2 \ \ 1)\begin{pmatrix} 1 \\ -2 \end{pmatrix} = 0$

so d^1 & d^2 are conjugate, wrt. A.

1. $\lambda^{-1} x^*$ $\quad k\lambda^{-1} y^*$

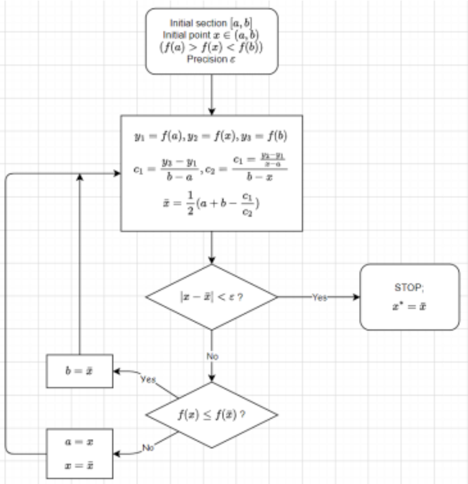
2. $-\nabla f(x^k)$ $\quad -(\nabla^2 f(x^k))^{-1} \nabla f(x^k)$ $\quad -\nabla f(x^k) + \dfrac{\|\nabla f(x^k)\|^2}{\|\nabla f(x^{k-1})\|^2} p^{k-1}$

3. $y_k = B_{k+1} S_k$

4. $f_1' = 10x_1 - 6x_2$, $f_2' = -6x_1 + 10x_2$

$f_{11}'' = 10$, $f_{12}'' = -6$

$f_{21}'' = -6$, $f_{22}'' = 10$ $\quad H = \begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$



Quadratic Convergence/Terminality（二次收敛/终止性）: Starting from any initial point, the positive definite quadratic function can always be achieved the minimum point within finite iterations.

Second-Order Convergence（二阶收敛性）: Refers to the convergent speed. $x^{n+1} - x^* \sim (x^n - x^*)^2, n \to \infty$

Steepest Descent Method （最速下降法）

- The search direction of two adjacent iterations is orthogonal. (theoretically, if linear search is accurate) $d^k = -\nabla f(x^k), \varphi'(\lambda^k) = \nabla f(x^{k+1})^T d^k = 0$
  The approach to the minimun point is zigzagged forward **(Sawtooth phenomenon)**.

1. $Q = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$ $x^1 = (1,1)^T$, $\varepsilon = 10^{-4}$

$Df(x) = Qx$

$Df(x^1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\|Df(x^1)\| = 2 > \varepsilon$.

$p^1 = -H_1 Df(x^1) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$

$\lambda = \arg\min f(x^1 + \lambda p^1) = \frac{1}{2}(2 - 8\lambda + 12\lambda^2) \Rightarrow \lambda = \frac{1}{3}$

$\Rightarrow x^2 = x^1 + \lambda p^1 = \begin{pmatrix} \frac{1}{3} \\ 1 \end{pmatrix}$.

$Df(x^2) = \begin{pmatrix} 0 \\ \frac{2}{3} \end{pmatrix}$, $\|Df(x^2)\| = \frac{2}{3} > \varepsilon$.

$\Delta x_1 = \begin{pmatrix} -\frac{2}{3} \\ 0 \end{pmatrix}$ $\quad \Delta g_1 = \begin{pmatrix} -2 \\ \frac{2}{3} \end{pmatrix}$, $r_1 = H_1 \Delta g_1 = \begin{pmatrix} -2 \\ \frac{2}{3} \end{pmatrix}$

$$H_2 = H_1 + \frac{\Delta x_1 \Delta x_1^T}{\Delta x_1^T \Delta g_1} - \frac{r_1 r_1^T}{\Delta g_1^T H_1 \Delta g_1}$$
$$= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \frac{3}{4}\begin{pmatrix} \frac{4}{9} & 0 \\ 0 & 0 \end{pmatrix} - \frac{9}{40}\begin{pmatrix} 4 & -\frac{4}{3} \\ -\frac{4}{3} & \frac{4}{9} \end{pmatrix}$$
$$= \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \frac{9}{10} & -\frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{13}{30} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{pmatrix}$$

$p^2 = -H_2 g_2 = \begin{pmatrix} -\frac{1}{5} \\ -\frac{3}{5} \end{pmatrix}$

$\lambda = \arg\min f(x^2 + \lambda p^2) = \frac{1}{15}(3\lambda - 5)^2 \Rightarrow \lambda = \frac{5}{3}$

$\Rightarrow x_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$Df(x^3) = 0$ $\quad \|Df(x^3)\| < \varepsilon$.

第二次测试

1、 ≤ ≥ ≤

2、基本 顶

3. $\begin{pmatrix}2\\1\end{pmatrix}$ 是

内点法 $f(x)=x_1^2+2x_2$, $Df=\begin{pmatrix}2x_1\\2\end{pmatrix}$

$g(x)=x_2-1$, $Dg=\begin{pmatrix}0\\1\end{pmatrix}$

取对数罚函数 $B(x)=-\ln(x_2-1)$

$T(x,r_1)=f(x)+r_1B(x)=x_1^2+2x_2-r\ln(x_2-1)$

$DT=\begin{pmatrix}2x_1\\2-\frac{r}{x_2-1}\end{pmatrix}$ 令 $r\to 0$, 得零点 $\begin{pmatrix}0\\1\end{pmatrix}$

此时 $D^2T>0$, 是极小值点、即 $x^*=\begin{pmatrix}0\\1\end{pmatrix}$, $f_{min}=2$.

线性规划

原: $\min\ c^Tx$ $c=\begin{pmatrix}1\\-2\\1\\0\end{pmatrix}$

s.t. $Ax=b$
$x\geq 0$

$$\begin{aligned}
x_1+x_2-2x_3+x_4 &\geq 10\\
-x_1-x_2+2x_3-x_4 &\geq -10\\
-2x_1+x_2-4x_3 &\geq -8\\
x_1-2x_2+4x_3 &\geq -4\\
x_1,x_2,x_3,x_4 &\geq 0
\end{aligned}$$

$A=\begin{pmatrix}1&1&-2&1\\-1&-1&2&-1\\-2&1&-4&0\\1&-2&4&0\end{pmatrix}$

$b=\begin{pmatrix}10\\-10\\-8\\-4\end{pmatrix}$

1. 对偶: $\max\ W=10y_1-10y_2-8y_3-4y_4$

$$\begin{aligned}
\text{s.t.}\quad y_1-y_2-2y_3+y_4 &\leq 1\\
y_1-y_2+y_3-2y_4 &\leq -2\\
-2y_1+2y_2-4y_3+4y_4 &\leq 1\\
y_1-y_2 &\leq 0\\
y_{1,2,3,4} &\geq 0
\end{aligned}$$

2、标准形式:

$\min\ x_1-2x_2+x_3$

$$\begin{aligned}
\text{s.t.}\quad x_1+x_2-2x_3+x_4 &=10\\
2x_1-x_2+4x_3+x_5 &=8\\
-x_1+2x_2-4x_3+x_6 &=4\\
x_{1\sim 6} &\geq 0
\end{aligned}$$

	c	1	-2	1	0	0	0		
C_B	x_B	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{b}$	θ
0	x_4	1	1	-2	1	0	0	10	10
0	x_5	2	-1	4	0	1	0	8	—
0	x_6	-1	2	-4	0	0	1	4	2
	σ_j	-1	2	-4	0	0	0		

	c	1	-2	1	0	0	0		
C_B	x_B	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{b}$	θ
0	x_4	$\frac{3}{2}$	0	0	1	0	$-\frac{1}{2}$	8	
0	x_5	$\frac{3}{2}$	0	2	0	1	$\frac{1}{2}$	10	
-2	x_2	$-\frac{1}{2}$	1	-2	0	0	$\frac{1}{2}$	2	
	σ_j	0	0	3	0	0	-1		

	c	1	-2	1	0	0	0		
C_B	x_B	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{b}$	θ
0	x_4	$\frac{3}{2}$	0	0	1	0	$-\frac{1}{2}$	8	
1	x_3	$\frac{3}{4}$	0	1	0	$\frac{1}{2}$	$\frac{1}{4}$	5	
-2	x_2	1	1	0	0	1	1	12	
	σ_j	$-\frac{9}{4}$	0	0	0	$-\frac{3}{2}$	$-\frac{7}{4}$		

≤ 0

故最优解 $x^*=\begin{pmatrix}0\\12\\5\\8\end{pmatrix}$ $f_{min}=-19$

非线性规划 $g_3(x)=-x_1+3x_2-1$?

(1) $g_1(x^0)=0$, $g_2(x^0)=0$, $g_3(x^0)\neq 0$

$g_1(x)\geq 0$, $g_2(x)\geq 0$ 是紧约束

(2) $Df=\begin{pmatrix}4(x_1-2)\\2(x_2-1)\end{pmatrix}$ $Dg_1=\begin{pmatrix}-2x_1\\1\end{pmatrix}$ $Dg_2=\begin{pmatrix}-1\\-1\end{pmatrix}$

$Df(x^0)-\left(w_1Dg_1(x^0)+w_2Dg_2(x^0)\right)=0$

$\Leftrightarrow\begin{cases}4w_1+w_2=4\\w_1-w_2=0\end{cases}$ 解得 $\begin{cases}w_1=\frac{4}{5}\\w_2=\frac{4}{5}\end{cases}$, $\begin{matrix}w_1\geq 0\\w_2\geq 0\end{matrix}$

故 x^0 是 K-T 点.

2024年2月27日 13:58

定理2 (1) 设 A_i 是 g_i 的解集，记 D_i 为定义域

i) 如果 $g_i \geq 0$, $\forall x \in D_i$，显然 A_i 闭

ii) 如果存在使 $g_i < 0$ 的点，$\forall x^* \in D_i$：$g_i(x^*) < 0$

根据连续函数的性质，存在 x^* 的邻域 U_{x^*} s.t. $g_i(x) < 0$，$x \in U_{x^*}$.

$B_i = \{U_{x^*} \mid x^* \in D_i, g_i(x^*) < 0\}$ 是 A_i^c 的开覆盖，并且每个 $U_{x^*} \subset A_i^c$，故 $A_i^c = B_i$

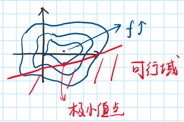
于是 $A_i^c = \bigcup \{U_{x^*} \mid x \in D_i, g_i(x) < 0\}$ 是开集，故 A_i 闭.

综上，可行域 $R = \bigcup_i A_i$ 闭.

(2) 设 $\{x_n\}$ 是 R^* 内的各项互异点列（如果存在的话），并且 $x_n \to x_0 (n \to \infty)$

$f(x) = b^T x$	$\nabla f(x) = b$	$\nabla^2 f(x) = O_n$
$f(x) = \frac{1}{2} x^T x$	$\nabla f(x) = x$	$\nabla^2 f(x) = I_n$
$f(x) = \frac{1}{2} x^T Q x$	$\nabla f(x) = Qx$	$\nabla^2 f(x) = Q$

等值面法求极小值



分派问题

$$\max \sum_{i,j} d_{ij} x_{ij}$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^{n} x_{ij} = 1, & j = 1, \cdots, n \\ \sum_{j=1}^{n} x_{ij} = 1, & i = 1, \cdots, n \\ x_{ij} \in \{0, 1\}, & i, j = 1, \cdots, n \end{cases}$$

凸集

Def. $\forall x_1, x_2 \in D \subset \mathbb{R}^n$，$x_1 x_2$ 连线 $\subset D$.

即 $\alpha x_1 + (1-\alpha) x_2 \in D$, $\forall \alpha \in [0,1]$

Pro1. A, B 是凸集，则 $A \cap B$, $A+B$, $A-B$ 也是凸集 ($A \cup B$ 未必)

Pro2. S 是凸集 $\Leftrightarrow$ S 中任意有限多个点的凸组合属于 S.

即 $\forall x_1, \cdots, x_n \in S$，$\lambda_1 x_1 + \cdots + \lambda_n x_n \in S$，$\forall \lambda_1 + \cdots + \lambda_n = 1$, $\lambda_1, \cdots, \lambda_n \in [0,1]$.

Proof. ($\Rightarrow$)

(凸集分离定理) 两不相交凸集必有分离超平面

设 $S_1, S_2 \subset \mathbb{R}^n$ 非空，如果超平面 $H = \{x \in \mathbb{R}^n \mid p^T x = \alpha\}$ 满足：

$\forall x \in S_1$, $y \in S_2$，都有 $p^T x \geq \alpha \geq p^T y$，则称 H 分离 $S_{1,2}$

(Gordan定理) 设 $A_{m \times n} \in \mathbb{R}^{m \times n}$，则

i) $\exists x \in \mathbb{R}^n$, s.t. $Ax < 0$

ii) $\exists y \in \mathbb{R}^n$, $y \neq 0$, $y \geq 0$, s.t. $A^T y = 0$

恰有一个成立. y的每个分量≥0，但不全为0.

Proof. i) $\Rightarrow \neg$ ii) 设 $\bar{x} \in \mathbb{R}^n$, $A\bar{x} < 0$.

假设ii)成立，设 $\bar{y} \in \mathbb{R}^m$, $A^T \bar{y} = 0$, $\bar{y} \geq 0$

(Farkas定理) 设 $A_{m \times n} \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$，则

i) $\exists x \in \mathbb{R}^n$, s.t. $Ax \leq 0$, $c^T x = 0$

ii) $\exists y \in \mathbb{R}^m$, s.t.

恰有一个成立.

三个定理关系？

(择一性定理)

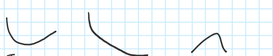
(凸包) 包含集合D的所有凸集交集 $co(D)$

(凸函数) 任意A, B，弦AB在弧AB之上.

$\xrightarrow{\mathbb{R}^n}$ 凸集上的凸函数

设 $D \subseteq \mathbb{R}^n$ 为凸集，$f: D \to \mathbb{R}$，如果 $\forall x_1, x_2 \in D$

$f(x) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$, $\forall x \in [x_1, x_2]$, $\alpha \in [0,1]$



严格凸　　凸　　凹（严格）

凹函数取负即为凸函数.

$f(x) = x^T A x$

若A半正定，$f(x)$ 凸

正定　　严格凸

凸函数的判定

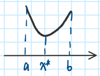
(一阶条件)

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$$

(二阶条件) $\forall x \in D$, $\nabla^2 f(x)$ 半正定

$\exists x^* \in [a,b]$, f 在 $[a,x^*]$ 上严格递减, 在 $[x^*,b]$ 上严格递增

f 单峰函数 $[a,b]$ 极小化问题的搜索区间

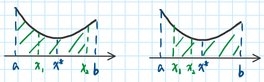


Th. 若 f 是 $[a,b]$ 上的单峰函数, $a \leq x_1 < x_2 \leq b$.

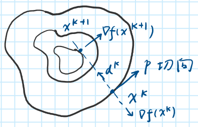
(1) 若 $f(x_1) \leq f(x_2)$, $x^* \in [a, x_2]$;

(2) $\geq$ $[x_1, b]$.

(缩小搜索区间)



取 x_1, x_2 中更高的点代替一个原端点. (x_1 左 x_2 右)



f 在切向上方向导数为 0,

即 $\frac{df}{dp} = 0 \Leftrightarrow \nabla f(x^k)^T \underbrace{p_0}_{\text{单位切向量}} = 0$

2024年3月12日 9:09

Levenberg-Marquardt (L-M)

找到最小的μ使得$\nabla^2 f(x)+\mu I$正定

整理 无约束优化中一维搜索的公式！
(代入梯度时，不直接求 argmin)

共轭方向法

A_n对称正定，$p, q \in \mathbb{R}^n$.

若$p^T A q = 0$，称方向p, q A共轭(正交)

若$\lambda = \arg\min f(x+\lambda d^{(2)})$

即$\nabla f(x+\lambda d^{(2)}) = A(x+\lambda d^{(2)}) + b = 0$

则 $0 = Ax + \lambda A d^{(2)} + b = 0$

$$[d^{(1)}]^T A x + \lambda [d^{(1)}]^T A d^{(2)} + [d^{(1)}]^T b = 0$$

$$[d^{(1)}]^T (Ax+b) = 0$$

即$d^{(1)}$与$\nabla f(x)$方向正交

作业

2024年6月23日 20:23

11. (1) $\forall z_1, z_2 \in \Omega_1 + \Omega_2$，设 $z_1 = x_1 + y_1$，$z_2 = x_2 + y_2$.

$\forall \alpha \in [0,1]$，$\alpha z_1 + (1-\alpha) z_2 = \alpha x_1 + (1-\alpha) x_2 + \alpha y_1 + (1-\alpha) y_2$

因为 Ω_1, Ω_2 凸，$\alpha x_1 + (1-\alpha) x_2 \in \Omega_1$，$\cdots \in \Omega_2$.

于是 $z_1 + z_2 \in \Omega_1 + \Omega_2$. 故…

(2) 同理

13. (1) $D^2 f = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$AC - B^2 = 3 > 0$，$A > 0$，严格凸

(2) $D^2 f = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 4 & 4 \\ 0 & 4 & 8 \end{pmatrix}$

各顺序主子式 > 0，正定的，严格凸

(3) $D^2 f = \begin{pmatrix} 2 & -2 & 2 \\ -2 & -6 & -6 \\ 2 & -6 & 0 \end{pmatrix}$ 不是凸函数

16. (1) $D^2 f = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$ 正定，严格凸

(2) $D^2 f = \begin{pmatrix} -2 & 2 \\ 2 & -10 \end{pmatrix}$ 负定，严格凹

(3) $D^2 f = \begin{pmatrix} 4 & 1 & -6 \\ 1 & 2 & 0 \\ -6 & 0 & 4 \end{pmatrix}$ 不定，非凹非凸.

18. 因为 $g(x)$ 凹，$\forall x_1, x_2 \in \{x \mid g(x) > 0\}$，$\forall \lambda \in [0,1]$，

$g(\lambda x_1 + (1-\lambda) x_2) \geq \lambda g(x_1) + (1-\lambda) g(x_2) > 0$.

从而 $f(\lambda x_1 + (1-\lambda) x_2) \leq \dfrac{1}{\dfrac{\lambda}{f(x_1)} + \dfrac{1-\lambda}{f(x_2)}}$

令 $f(x_1) = a$，$f(x_2) = b$，则 $a, b > 0$.

下证 $\dfrac{ab}{\lambda b + (1-\lambda) a} \leq \lambda a + (1-\lambda) b$

即证 $ab \leq [\lambda b + (1-\lambda) a][\lambda a + (1-\lambda) b]$

$LHS = ab + (-\lambda^2 + \lambda)(a-b)^2$

由 $\lambda \in [0,1]$，$-\lambda^2 + \lambda \geq 0$，故 $ab \leq LHS$.

从而 $f(\lambda x_1 + (1-\lambda x_2)) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$. 证

19. $D^2 f = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$ 正定，严格凸

显然约束都是凸函数. 故是为凸规划.

1.(1) 若 $x_k \le 1$，$\left|\frac{x_k}{2}\right| \le |x_k|$，下降.

若 $x_k > 1$，$\left|\frac{x_k-1}{2}+1\right| = \left|\frac{x_k+1}{2}\right| \le |x_k|$，下降.

(2) 若 $a \le 1$，显然 $\lim_{k\to\infty} x_k = 0$.

若 $a>1$，$x_{k+1} = \frac{x_k+1}{2}$，即 $2(x_{k+1}-1) = x_k - 1$

从而 $x_k = \frac{a-1}{2^k}+1$，$\lim_{k\to\infty} x_k = 1$.

16. $Df = \begin{pmatrix} 6x_1-4 \\ 4x_2-6 \end{pmatrix}$.

$-Df(x^0) = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$,

argmin $f(0+4\lambda, 0+6\lambda) = 120\lambda^2 - 52\lambda \Rightarrow \lambda = \frac{7}{30}$

得 $x_1 = \begin{pmatrix} \frac{14}{15} \\ \frac{7}{5} \end{pmatrix}$，$-Df(x^1) = \begin{pmatrix} -\frac{8}{5} \\ \frac{2}{5} \end{pmatrix}$

得 $x_2 = \begin{pmatrix} \frac{11}{15} \\ \frac{1}{15} \end{pmatrix}$

18. $Df = \begin{pmatrix} 8x_1-8 \\ 2x_2-4 \end{pmatrix}$ $D^2 f = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$

Newton: $Df(x^0) = \begin{pmatrix} -8 \\ -4 \end{pmatrix}$

$x' = x^0 - (D^2 f(x^0))^{-1} Df(x^0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

阻尼Newton: $S^0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

argmin $f = 8\lambda^2 - 16\lambda \Rightarrow \lambda = 1$

得 $x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$Df\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 0$

故极小点为 $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

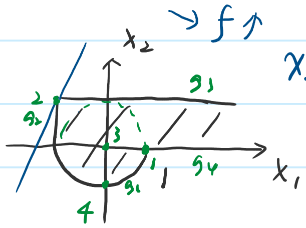
23. $Df = \begin{pmatrix} 8x_1 - 4x_2 \\ -4x_1 + 8x_2 - 12 \end{pmatrix}$

得 $p_0 = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$

argmin $f \Rightarrow \lambda = \frac{17}{104}$.

$\Rightarrow x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

1.



$x_2 = 2x_1 - f$

$\min f = -3$

$x = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

t-紧 l-松	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
$g_1(x) = -\sqrt{1-x_1^2} - x_2 \le 0$	t	l	l	t
$g_2(x) = -x_1 \le 1$	l	t	l	l
$g_3(x) = x_2 \le 1$	l	t	l	l
$g_4(x) = -x_2 \le 0$	t	l	l	l

3. $Df = \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix}$ $Dg_1 = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ $Dg = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} - w_1 \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} - w_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0, \ w_1 \ge 0, \ w_2 \ge 0$

$\Rightarrow w_1 = 0, \ w_2 = 0, \ x = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$

2. $Df = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ $Dg_1 = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$ $Dg_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $Dg_3 = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ $Dg_4 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ $Dg_5 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

$Df - \sum_{i=1}^{5} w_i Dg_i = 0, \ w_i \ge 0$, 在 x^*

$\Rightarrow w = (1, 1, 1, 0, 0)$

7. (1) $DF = \begin{pmatrix} 2x_1 - 2M\max(0, 1-x_1) \\ 2x_2 \end{pmatrix} = 0 \xRightarrow{M \to +\infty} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

(2) $DF = \begin{pmatrix} -x_2 + 2M\cdots \\ -x_1 + 2M\cdots \end{pmatrix} = 0 \xRightarrow{M \to +\infty} x = \begin{pmatrix} \frac{2}{3} \\ \frac{13}{3} \end{pmatrix}$

8. (2) $F = x_1^2 + x_2^2 - r\ln(2 - x_1 - x_2) + r\ln(x_2 - 1)$

$DF = \begin{pmatrix} 2x_1 + \frac{r}{2-x_1-x_2} \\ 2x_2 + \frac{r}{2-x_1-x_2} - \frac{r}{x_2-1} \end{pmatrix} \xRightarrow{r \to 0} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

(3) $DF = \begin{pmatrix} 1 - r\frac{-3x_1^2 + x_2}{-x_1^3 + x_1x_2} \\ 1 - r\frac{x_1}{-x_1^3 + x_1x_2} \end{pmatrix} \xRightarrow{r \to 0} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\left(\cdots \ -x_1^3 + x_1 x_2 \quad 1 - r \frac{x_1}{-x_2^3 + x_1 x_2} \right) \Rightarrow x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

一维搜索

2024年9月17日 15:36

$$\lim_{k\to\infty}\frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|^{\alpha}}=q$$ Q - α 阶收敛速度

$\alpha=1$ Q - 线性收敛速度

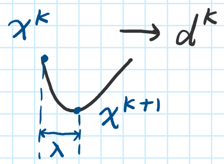
$1<\alpha<2,\ q>0$ 或 $\alpha=1,\ q=0$ Q - 超线性

$\alpha=2$ Q - 二阶

Th. Q - 超线性 $\Rightarrow \lim\limits_{k\to\infty}\frac{\|x^{k+1}-x^k\|}{\|x^k-x^*\|}=1$

可用 $\|x^{k+1}-x^k\|<\varepsilon$ 来代替 $\|x^k-x^*\|<\varepsilon$ 来给出终止判断.

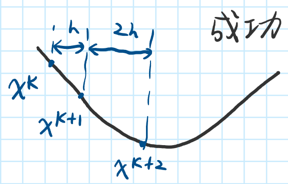
$\nabla f(x^{k+1})^T d^k=0$

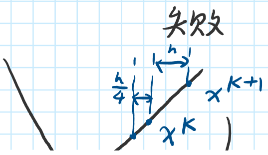


$$\frac{d f(x^k+\lambda d^k)}{d\lambda}=0,\quad x^{k+1}=x^k+\lambda d^k$$

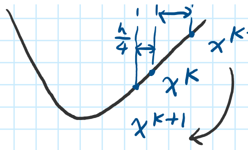
$\Rightarrow \nabla f(x^{k+1})^T d^k=0$

成功 - 失败法



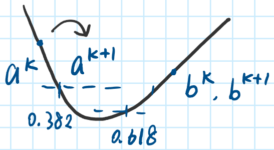


$|h|<\varepsilon$ 则终止



$|h|<\varepsilon$ 则终止

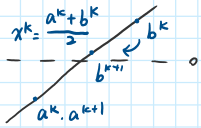
0.618法



(变成$[a^{k+1}, b^{k+1}]$的0.382分位点)

若$|a^k-b^k|<\varepsilon$,
则终止, 取 $x^*=\dfrac{a^k+b^k}{2}$

二分法(导数)



若$|a^k-b^k|<\varepsilon$或$f'(x)<\varepsilon$则终止
取$x^*=x^k$.

Newton法

(近似到二阶才有极值)

$$f(x^{k+1})\approx f(x^k)+f'(x^k)(x^{k+1}-x^k)+\frac{1}{2}f''(x^k)(x^{k+1}-x^k)^2$$

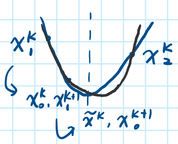
$\Rightarrow \underset{x^{k+1}}{\operatorname{argmin}}\ f(x^{k+1})$ 即 $f'(x^k)+f''(x^k)(x^{k+1}-x^k)$ 的零点

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

当$|f'(x^k)|<\varepsilon$时终止,
取$x^*=x^k$.

要求对充分大的k, $f''(x^k)>0$.

抛物线法

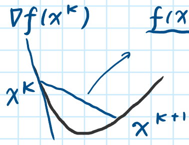


若$|f(\tilde{x}^k)-f(x_0^k)|<\varepsilon$, 终止
取$x^*=\tilde{x}^k$

非精确

Wolfe-Powell准则

$0<c_1<c_2<1$

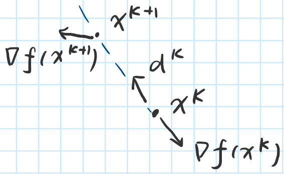


$\dfrac{f(x^{k+1})-f(x^k)}{\lambda^k d^k}$ (仅作示意)

1° $f(x^k)-f(x^{k+1}) \geq -c_1\lambda^k \nabla f(x^k)^T d^k$ (控制 λ 上界)

2° $\nabla f(x^{k+1}) d^k \geq c_2 \nabla f(x^k) d^k$ (控制 λ 下界)

一般 $c_1=0.1$, $c_2=0.5$



$\nabla f(x^{k+1}) d^k$ 随 λ↑

$a=0$, $b=+\infty$, $\lambda=1$

满足1°, 2°终止, $\lambda^k=\lambda$

不满足1°, 令 $b=\lambda$, $\lambda=\dfrac{a+b}{2}$.

满足1°不满足2°, 令 $a=\lambda$, $\lambda=\min\{2a, \dfrac{a+b}{2}\}$

可根据实际需要变化

无约束优化

2024年9月17日 18:26

Lem. f 在 x^k 处可微，

若对于方向 d^k，$\nabla f(x^k)^T d^k < 0$，

则 $\exists \delta > 0$，$\forall \lambda \in (0, \delta)$，$f(x^k + \lambda d^k) < f(x^k)$

Th.（一阶必要条件）

设 f 在 x^* 处可微，则 x^* 为局部最优解的

必要条件为 $\nabla f(x^*) = 0$.

proof. 假设 $\nabla f(x^*) \neq 0$，则令 $d = -\nabla f(x^*)$，

$\nabla f(x^*)^T d = -\|\nabla f(x^*)\| < 0$

与引理矛盾（x^* 为局部最优解）

Th.（二阶必要条件）

设 f 在 x^* 二次可微，则 x^* 为局部最优解的

必要条件为 $\nabla f(x^*) = 0$ 且 $|\nabla^2 f(x^*)| \geq 0$

proof. 因为二次可微，$\nabla f(x^*) = 0$，

$f(x^* + \lambda d) - f(x^*) = \frac{1}{2}\lambda^2 d^T \nabla^2 f(x^*) d + o(\lambda^2 \|d\|^2) \leq 0$

当 $\lambda \to 0$ 时，$o(\lambda^2 \|d\|^2) \to 0$

于是 $d^T \nabla^2 f(x^*) d \geq 0$，则 $|\nabla^2 f(x^*)| \geq 0$.

Th.（充分条件）

设 f 在 x^* 二次可微，则 x^* 为局部最优解的

充分条件为 $\nabla f(x^*) = 0$ 且 $|\nabla^2 f(x^*)| > 0$

最速下降法（Steepest Descend Method）

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

若 $\|\nabla f(x^k)\| < \varepsilon$ 停止，$x^* = x^k$

Th.（SDM 的收敛性定理）

在极值点附近收敛速度慢

对小扰动不稳定（初值敏感）

改进：加速梯度法

阻尼牛顿法（Damped Newton）

牛顿方向：$d^k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$ （不一定是下降方向！）

若 $\|\nabla f(x^k)\|$ 停止，$x^* = x^k$.

对二次凸函数一次迭代即达到最优解.

（二次终止性） ... 有限次

Th.（收敛性定理）

当 Hesse 阵奇异时，难以计算 d^k

Hesse 阵计算量大

改进

1° 每 m 次迭代使用同一个 Hesse 阵

2° 取正定阵 $G^k = \nabla^2 f(x^k) + \varepsilon^k I$.（$I$ 是单位阵）

$d^k = -(G^k)^{-1} \nabla f(x^k)$ （扰动方程）

鞍点处可取 d^k 满足 $(d^k)^T \nabla^2 f(x^k) d^k < 0$

共轭梯度法 (Conjugate Gradient Methods)

A-共轭 $\underbrace{(d^1)^T A d^2}_{\text{也是一种内积}} = 0$

一组A-共轭向量线性无关.

$x \in \mathbb{R}^n$ 可表示为 $x = \sum_{i=1}^{n} \underbrace{\frac{(d^i)^T A x}{(d^i)^T A d^i}}_{A\text{-投影}} d^i$ $\qquad x = \sum_{i=1}^{n} \frac{(\alpha_i, x)}{\|\alpha_i\|} \alpha_i$

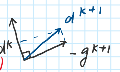
(二次终止性) $f(x) = \frac{1}{2} x^T A x + b^T x + c$

$d^1, \cdots, d^n$ A-共轭

依次沿 $d^1, \cdots, d^n$ 精确一维搜索, 至多n次到达最优解.

令 $g^1 = \nabla f(x^1)$, $d^1 = -g^1$. 初始方向必需是负梯度方向! (使 g^2 与 g^1 正交)

$d^{k+1} = -g^{k+1} + \alpha_k d^k$, $\alpha_k = \frac{(g^{k+1})^T A d^k}{(d^k)^T A d^k}$ (该式不要求 d^k 与 $-g^{k+1}$ 正交)



$g^{k+1} = A x^{k+1} + b = A(x^k + \lambda_k d^k) + b = g^k + \lambda_k A d^k$

对于二次函数, $\lambda_k = -\frac{(g^k)^T d^k}{(d^k)^T A d^k}$

$d^{k+1} = -g^{k+1} + \alpha_k d^k$

$\Rightarrow g^{k+1} = \alpha_k d^k - d^{k+1}$

由 d^k, d^{k+1} A-共轭, α_k 即 g^{k+1} 在 d^k 上的A-投影

取 $x^* = x^{n+1}$

反之, $(d^{k+1})^T A d^k = -g^{k+1} A d^k + \alpha_k (d^k)^T A d^k = 0$

故若 $d^{k+1} = -g^{k+1} + \alpha_k d^k$, d^{k+1} 与 d^k A-共轭 $\Leftrightarrow \alpha_k = \frac{(g^{k+1})^T A d^k}{(d^k)^T A d^k}$

归纳法, 假设 $d^1, \cdots, d^{k-1}$ A-共轭, $g^1, \cdots, g^{k-1}$ 正交,
则由 g^k 与 d^{k-1} 正交, $(g^k)^T d^{k-2} = \alpha_{k-2}^{-1} (g^k)^T (d^{k-1} + g^{k-1}) = 0$
于是 g^k 与 d^{k-2} 正交. 归纳可知 g^k 与 $d^1, \cdots, d^{k-1}$ 正交.

$g^{k+1} = A x^{k+1} + b = A(x^k + \lambda_k d^k) + b$

$(g^{k+1})^T d^k = 0 \Rightarrow$

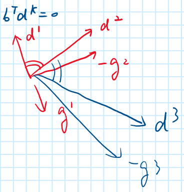
$i \le k$, $(g^k)^T g^i = (g^{k-1} + \lambda_{k-1} A d^{k-1}) g^i$

$(x^k + \lambda_k d^k)^T A^T d^k + b^T d^k$

$(x^k)^T A d^k + \lambda_k (d^k)^T A d^k + b^T d^k = 0$

$\lambda_k = -\frac{(x^k A + b)^T d^k}{(d^k)^T A d^k}$

$= -\frac{(g^k)^T d^k}{(d^k)^T A d^k}$



$(g^k)^T g^{k-1}$

$(g^k)^T d^{k-1}$

$d^k = -g^k + \alpha_{k-1}(-g^{k-1} + \alpha_{k-2} d^{k-2})$

$\Rightarrow (g^k)^T d^k = -(g^k)^T g^k + \alpha_{k-1} (g^k)^T d^{k-1}$

$\Rightarrow (g^k)^T (d^k + g^k) = 0$

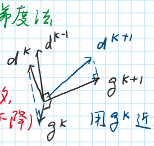
改进: 计算 α_k 时不使用A.

$(d^{k-1})^T g^k = 0$, $(d^k)^T g^{k+1} = 0$, $(d^{k-1})^T A d^k = 0$

FR (Fletcher-Reeves) 共轭梯度法

重新开始策略

$\alpha_k = \frac{(g^{k+1})^T g^{k+1}}{(g^k)^T g^k}$ (对二次函数, 等价最速下降)



用 g^k 近似 d^k

若 $\|g^k\| < \varepsilon$ 则停,

取 $x^* = x^k$

若 $k = n+1$, 置 $x^1 = x^{n+1}$, $g^1 = \nabla f(x^1)$, $d^1 = -g^1$, $k = 1$.

PRP 共轭梯度法

$\alpha_k = \frac{(g^{k+1})^T (g^{k+1} - g^k)}{(g^k)^T g^k}$

DFP 变尺度法 (一种拟Newton法)

在 x^{k+1} 附近, $\nabla f(x) \approx g^{k+1} + H_{k+1}(x - x^{k+1})$

取 $x = x^k$, 则 $\Delta x^k \approx H_{k+1}^{-1} \Delta g^k$

$(\Delta x^k = x^{k+1} - x^k$, $\Delta g^k = g^{k+1} - g^k)$

要构造 $\mathbb{H}_k$, s.t. $\Delta x^k = \mathbb{H}_{k+1} \Delta g^k$ 拟Newton方程

$\mathbb{H}_1 = I$, $d^k = -\mathbb{H}_k g^k$,

$\mathbb{H}_{k+1} = \mathbb{H}_k + \frac{\Delta x^k (\Delta x^k)^T}{(\Delta x^k)^T \Delta g^k} - \frac{r^k (r^k)^T}{(r^k)^T \Delta g^k}$, $r^k = \mathbb{H}_k \Delta g^k$ DFP 修正公式

若 $\|g^k\| < \varepsilon$ 则停, $x^* = x^k$.

$k = n+1$ 则置 $x^1 = x^{n+1}$, $g^1 = \nabla f(x^1)$, $\mathbb{H}_1 = I$, $d^1 = -\mathbb{H}_1 g^1$, $k = 1$

$x^{k+1} = x^k + \lambda_k d^k$, $\Delta x^k = \mathbb{H}_{k+1} \Delta g^k$

$\Rightarrow \lambda_k d^k = \mathbb{H}_{k+1} \Delta g^k$

$d^k = \lambda_k^{-1} (\mathbb{H}_{k+1} g^{k+1} - \mathbb{H}_{k+1} g^k)$

$\mathbb{H}_k$ 对称正定, 故 d^k 为下降方向.

$\Rightarrow \lambda_{k-1} d^{k-1} = -\mathbb{H}_k g^{k-1} - d^k$

$\Rightarrow x^k = x^{k-1} - (\mathbb{H}_k g^{k-1} + d^k)$

$\Rightarrow d^k + \mathbb{H}_k g^{k-1} = -\mathbb{H}_k \Delta g^{k-1}$

$\Rightarrow d^k = -\mathbb{H}_k g^k$

相当于每次迭代用不同的A定义内积和范数的共轭梯度法.

改进: BFGS方法

$\mathbb{B}_{k+1} \Delta x^k = \Delta g^k$

$\mathbb{B}_{k+1} = \mathbb{B}_k + \frac{\Delta g^k (\Delta g^k)^T}{\cdots} - \frac{\Delta s^k (\Delta s^k)^T}{\cdots}$, $\Delta s^k = \mathbb{B}_k \Delta x^k$

改进：BFGS方法

$$\mathbb{B}_{k+1}\Delta x^k=\Delta g^k$$

$$\mathbb{B}_{k+1}=\mathbb{B}_k+\frac{\Delta g^k(\Delta g^k)^T}{(\Delta g^k)^T\Delta x^k}-\frac{\Delta s^k(\Delta s^k)^T}{(\Delta s^k)^T\Delta x^k},\quad \Delta s^k=\mathbb{B}_k\Delta x^k$$

$$d^k=-\mathbb{B}_k^{-1}g^k$$

检验后的变化

考虑到 √，x^1还是x_1^1？

方向加速法（Powell）

检验

精确一维搜索

$$x^1=x_1^1\xrightarrow{e_1}x_2^1\xrightarrow{e_2}\cdots\xrightarrow{e_{n-1}}x_n^1\xrightarrow{e_n}x_{n+1}^1 \quad \text{令 } d^1=x_{n+1}^1-x^1,\ x_{n+1}^1\xrightarrow{d^1}x^2$$

$$x^2=x_1^2\xrightarrow{e_2}x_2^2\xrightarrow{e_2}\cdots\xrightarrow{e_n}x_n^2\xrightarrow{d_1}x_{n+1}^2 \quad \text{令 } d^2=x_{n+1}^2-x^2,\ x_{n+1}^2\xrightarrow{d^2}x^3$$

$$\vdots$$

$$x^n=x_1^n\xrightarrow{e_n}x_2^n\xrightarrow{d^1}\cdots\xrightarrow{d^{n-2}}x_n^n\xrightarrow{d^{n-1}}x_{n+1}^n \quad \text{令 } d^n=x_{n+1}^n-x^n,\ x_{n+1}^n\xrightarrow{d^n}x^{n+1}$$

$x_{n+1}^1\xrightarrow{d^1}x^2$，$x_n^2\xrightarrow{d^1}x_{n+1}^2$

分别从不同点沿d^1进行精确一维搜索，

则d^1与$d^2=x_{n+1}^2-x^2$ A-共轭.

如此找到一组共轭方向，并依次沿$d^1,\cdots d^n$进行了精确一维搜索.

若$\|x^{k+1}-x^k\|<\varepsilon$则停，取$x^*=x^{k+1}$.

若$f(x_1^k)-2f(x_{n+1}^k)+f(2x_{n+1}^k-x_1^k)\geq 2\max\limits_{1\le i\le n}[f(x_i^k)-f(x_{i+1}^k)]$，则令$x_1^k=x_{n+1}^k$，重新寻找$d^k$.

Th. $(d^i)^TAd^i=1$

$B=(d^1,\cdots,d^n)$

$\det B$ 取到最大值 $\Longleftrightarrow$ $d^1,\cdots,d^n$ A-共轭.

Hooke-Jeeves 步长加速法

约束优化方法

2024年9月18日 17:42

Theorem

(Kuhn-Tucker必要条件) 设在问题(1)中，$\bar{x} \in S$，函数$f(x), g_i(x)(i \in I)$在点$\bar{x}$处可微，$g_i(x)(i \notin I)$ 在点$\bar{x}$处连续，$h_j(x)(j = 1, \cdots, n)$ 在点$\bar{x}$处连续可微，且

(CQ)向量集$\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}) | i \in I, j = 1, \cdots, n\}$线性无关，

若$\bar{x}$是问题(1)的局部最优解，则存在数$w_i(i \in I)$和$v_j, j = 1, \cdots, n$使得

$$\nabla f(\bar{x}) = \sum_{i \in I} w_i \nabla g_i(\bar{x}) + \sum_{j=1}^{n} v_j \nabla h_j(\bar{x}), \; w_i \geq 0, i \in I. \qquad (2)$$

★ 在上述K-T定理中，若进一步假设$g_i(x)(i \notin I)$ 在$\bar{x}$ 处可微，则K-T条件(2)可写成如下形式

$$\begin{aligned}
\nabla f(\bar{x}) &= \sum_{i=1}^{m} w_i \nabla g_i(\bar{x}) + \sum_{j=1}^{n} v_j \nabla h_j(\bar{x}), \\
w_i g_i(\bar{x}) &= 0, i = 1, \cdots, m, \longleftarrow\text{-----}\longrightarrow \text{互补松弛条件} \\
w_i &\geq 0, i = 1, \cdots, m.
\end{aligned}$$

★ 若存在$\bar{x} \in S$使得K-T条件成立，则$\bar{x}$ 称为问题(1)的K-T点.

K-T点

CG： $g_i(\bar{x})$ 与 $h_i(\bar{x})$ 线性无关
$i \in I$ $\quad i=1,\cdots,m_2$

$F(x, \sigma) = f(x) + \sigma p(x)$

外点法

$$F(x, M_k) = f(x) + M_k\left(\sum_{i=1}^{m_1} \min\{0, [g_i(x)]^2\} + \sum_{i=1}^{m_2} [h_i(x)]^2\right)$$

若 $|\sigma p(x^k)| < \varepsilon$，则停，否则增大$M$，如$M_{k+1} = 10 M_k$

$$x^k = \underset{x \in \mathbb{R}^n}{\arg\min}\, F(x, M_k)$$

1° $p(x^{k+1}) \leq p(x^k)$ $\quad M\uparrow$ 对$p(x)$更敏感

2° $f(x^{k+1}) \geq f(x^k)$ $\quad$ 约束的影响更大，无约束目标更劣

3° $F(x^{k+1}, M_{k+1}) \geq F(x^k, M_k)$

4° $F(x^k, M_{k+1}) \leq f(x), \forall x \in S$

$$\begin{cases} f(x^k) + M_k p(x^k) \leq f(x^{k+1}) + M_k p(x^{k+1}) & ① \\ f(x^{k+1}) + M_{k+1} p(x^{k+1}) \leq f(x^k) + M_{k+1} p(x^k) & ② \end{cases}$$

①+② $\Rightarrow (M_{k+1} - M_k)[p(x^{k+1}) - p(x^k)] \leq 0$

$\qquad >0$ ③ $\Rightarrow$ $\qquad \leq 0$ ④

①.④ $\Rightarrow f(x^{k+1}) \geq f(x^k)$

①.③ $\Rightarrow F(x^k, M_k) \leq F(x^{k+1}, M_{k+1})$

内点法 $\qquad$ 对等式约束问题不适用

$$-\sum_{i=1}^{m} \ln g_i(x) \qquad \sum_{i=1}^{m} \frac{1}{g_i(x)} \qquad \sum_{i=1}^{m} \frac{1}{[g_i(x)]^2}$$

初始内点

$g_i(x) \le 0 \qquad g_i(x) > 0$

$\min P(x, r_k) = -\sum_{i\in S_k} g_i(x) + r_k \sum_{i\in T_k} \frac{1}{g_i(x)}$ 保持 $i\in T_k$ 的约束仍有效

减小 r_k. 让 $i\in S_k$ 的 $g_i(x)\uparrow$

1° $B(x^{k+1}) \ge B(x^k)$

2° $f(x^{k+1}) \le f(x^k)$

3° $F(x^{k+1}, M_{k+1}) \le F(x^k, M_k)$

4° $F(x^k, M_{k+1}) > f(x^*)$

乘子法

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

$H(x,\lambda,M) = f(x) - \sum_{i=1}^{m} \lambda_i h_i(x) + \frac{M}{2}\sum_{i=1}^{m}[h_i(x)]^2$ 增广 Lagrange 函数

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

$h_i(x)=0$ 在 x 处的法线方向为 $\nabla h_i(x)$

当 $\nabla L(x,\lambda)=0$，则 f 在 $\nabla h_i(x)$ 上的方向导数

$\nabla f(x)^T \nabla h_i(x) = 0$

$x^{k+1} = \arg\min_x H(x, \lambda^k, M_k)$

$\lambda_j^{k+1} = \lambda_j^k - M h_j(x^{k+1})$

（则 $\nabla f(x^{k+1}) = \sum_{j=1}^{m} \lambda_j^{k+1} \nabla h_j(x^{k+1})$，满足 K-T 条件）

若 $\sum_{j=1}^{m}[h_j(x^k)]^2 < \varepsilon$，则停，取 $x^* = x^k$.

不等式约束

引进人工变量，则 $g_i(x) - z_i^2 = 0$

$$\begin{cases} f(x) - \sum_{i=1}^{m} \mu_i (g_i(x) - z_i^2) + \frac{M}{2}\sum_{i=1}^{m}(g_i(x)-z_i^2)^2 \\ \frac{\partial \cdots}{\partial z_i} = 0 \end{cases}$$

$\Rightarrow H(x,\lambda,M) = f(x) + \frac{1}{2M}\sum_{i=1}^{m}\left(\max\{0, (\mu_i - M g_i(x))^2\} - \mu_i^2\right)$

可行方向法

Zoutendijk

起作用约束 A_1, b_1

松弛 A_2, b_2

解 $\min\limits_{d} \nabla f(x^k)^T d$

$$\text{s.t.} \begin{cases} A_1 d \geq 0 \\ |d_j| \leq 1 \end{cases}$$

得到可行方向 d^k

指第i个分量

步长入应满足 $\lambda \leq \min\left\{\frac{\bar{b}_i}{\bar{d}_i} \middle| \bar{d}_i < 0\right\}$, if $\exists\, \bar{d}_i < 0$

$\bar{b} = b_2 - A_2 x^k$

$\bar{d} = A_2 d^k$

Frank-Wolfe

求 $\min \nabla f(x^k)^T y$
s.t. $y \in D$ $\Rightarrow y^k$ 若 $|\nabla f(x^k)^T d^k| < \varepsilon$ 则停

$d^k = y^k - x^k$

$0 \leq \lambda \leq 1$

梯度投影法

$M_{n\times n} = \begin{pmatrix} A_1 \\ I \end{pmatrix}$, 若M空, 令 $P = I$, 否则 $P = I - M^T(MM^T)^{-1}M$

$d^k = -P\nabla f(x^k)$, 若 $\|d^k\| < \varepsilon$, 令 $w = (MM^T)^{-1}M\nabla f(x) = \begin{pmatrix} u \\ v \end{pmatrix}$, 若 $u \geq 0$ 则停,

否则令 $u_j < 0$, 去掉 A_i 中 j 行重求

一维搜索同 Zoutendijk

多目标优化

2024年9月20日 11:04