· $\min f(x)$ 牛顿法

· x^k $\min\limits_x f(x^k)+\underbrace{\nabla f(x^k)^T(x-x^k)+\frac{1}{2}(x-x^k)^T\nabla^2 f(x^k)}_{\text{二次函数}}(x-x^k)$

$x^{k+1}=x^k-[\nabla^2 f(x^k)]^{-1}\nabla f(x^k)$

$\min f(x)$
$g_1(x)=0$
$g_2(x)\le 0$

x^k
$$\begin{cases}\min & f(x^k)+\nabla f(x^k)^T(x-x^k)+\frac{1}{2}(x-x^k)^T\nabla^2 f(x^k)(x-x^k)\\ \text{s.t.} & g_1(x^k)+\nabla g_1(x^k)^T(x-x^k)=0\\ & g_2(x^k)+\nabla g_2(x^k)^T(x-x^k)\le 0\end{cases}$$

$\downarrow$

x^{k+1}
$$\begin{cases}\min & \frac{1}{2}x^TGx+c^Tx\\ \text{s.t.} & a_1^Tx=b_1\\ & a_2^Tx\le b_2\end{cases}$$ 二次规划

h.w.

$$\begin{cases}\min\limits_x & x_1^2+x_2^2-2x_1-4x_2\\ \text{s.t.} & x_1+x_2\le 1\\ & -x_1\le 0\\ & -x_2\le 0\end{cases}$$

$G=\begin{bmatrix}2&0\\0&2\end{bmatrix}$ $c=\begin{bmatrix}-2\\-4\end{bmatrix}$

$a_1=\begin{pmatrix}1\\1\end{pmatrix}$ $b_1=1$

$a_2=\begin{pmatrix}-1\\0\end{pmatrix}$ $b_2=0$

$a_3=\begin{pmatrix}0\\-1\end{pmatrix}$ $b_3=0$

第一次迭代

$x^1=\begin{pmatrix}0\\0\end{pmatrix}$ $A_1=\{2,3\}$

$\bar{x}_1=\begin{pmatrix}0\\0\end{pmatrix}$ $d^1=\begin{pmatrix}0\\0\end{pmatrix}$

$$\begin{cases}\min & x_1^2+x_2^2-2x_1-4x_2\\ \text{s.t.} & -x_1=0\\ & -x_2=0\end{cases}$$

$$\begin{bmatrix}\bar{x}_1\\ \lambda^1\end{bmatrix}=\begin{bmatrix}G&A^T\\A&0\end{bmatrix}\begin{bmatrix}-c\\b\end{bmatrix}=\begin{bmatrix}2&0&-1&0\\0&2&0&-1\\-1&0&0&0\\0&-1&0&0\end{bmatrix}^{-1}\begin{bmatrix}2\\4\\0\\0\end{bmatrix}$$

$$=\begin{bmatrix}0\\0\\-2\\-4\end{bmatrix}$$ $\lambda^1=\begin{pmatrix}-2\\-4\end{pmatrix}$

第二次迭代 $x^2=\begin{pmatrix}0\\0\end{pmatrix}$ $A_2=\{2\}$

$$\begin{cases}\min & x_1^2+x_2^2-2x_1-4x_2\\ \text{s.t.} & -x_1=0\end{cases}$$

$\bar{x}^2=\begin{pmatrix}0\\2\end{pmatrix}$ $d^2=\begin{pmatrix}0\\2\end{pmatrix}$

$\bar{x}_2\notin\Phi$ $i\in I\setminus A_2=\{1,3\}$

$a_1^Td^2=2>0$

$a_3^Td^2=-2$

$\alpha_2=\dfrac{-(a_1^Tx^2-b_1)}{a_1^Td^2}=\dfrac{-(-1)}{2}=\dfrac{1}{2}$

第三次迭代

$x^3=x^2+\alpha_2d^2=\begin{pmatrix}0\\0\end{pmatrix}+\frac{1}{2}\begin{pmatrix}0\\2\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$

$A_3=\{1,2\}$ $\bar{x}_3=\begin{pmatrix}0\\1\end{pmatrix}$

$$\begin{cases} \min \ x_1^2+x_2^2-2x_1-4x_2 \\ s.t. \quad x_1+x_2=1 \\ \qquad -x_1=0 \end{cases}$$

$$\begin{bmatrix} x_3 \\ \lambda^3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 & -1 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 4 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

Lecture 7
序列二次规划方法

Print version of the lecture in *优化方法 (Optimization Methods)*

presented on 2020, 秋

by 肖现涛 from 数学科学学院 at 大连理工大学

7.1

Contents

7.2

1 二次规划

等式约束二次规划

- 给定 $G \succeq 0$, 行满秩矩阵 $A \in \mathbb{R}^{m\times n}$
- 考虑等式约束二次规划

$$\begin{array}{ll} \min_x & \frac{1}{2}x^TGx + c^Tx \\ \text{subject to} & Ax = b \end{array}$$

- 令 $Z \in \mathbb{R}^{n\times(n-m)}$ 由 A 的零空间的一组基构成
- 若 Z^TGZ 正定, 令

$$\begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} G & -A^T \\ -A & 0 \end{bmatrix}^{-1} \begin{bmatrix} -c \\ -b \end{bmatrix}$$

则 x^* 是全局最优解, λ^* 是对应的 Lagrange 乘子

$$L(x,\lambda) = \frac{1}{2}x^TGx + c^Tx + \lambda^T(Ax-b)$$

KKT条件 $\nabla_x L(x,\lambda) = Gx + c + A^T\lambda = 0$

$Ax = b$

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

$$\begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -c \\ b \end{bmatrix}$$

等式约束两者皆同
积极集法

7.3

凸二次规划

- 考虑凸二次规划

$$\begin{array}{ll} \min_x & \frac{1}{2}x^TGx + c^Tx \\ \text{subject to} & a_i^Tx = b_i, \ i \in \mathcal{E} \\ & a_i^Tx \le b_i, \ i \in \mathcal{I} \end{array}$$

- $Q(x) = \frac{1}{2}x^TGx + c^Tx$
- 可行集 $\Phi = \{x | a_i^Tx = b_i, i \in \mathcal{E}, \ a_i^Tx \le b_i, i \in \mathcal{I}\}$
- 积极集 (active set): $\mathcal{A}(x) = \{i \in \mathcal{E} \cup \mathcal{I} | a_i^Tx = b_i\}$

定理
设 x^* 是凸二次规划的最优解, LICQ 在 x^* 成立, 则 x^* 是下述问题的最优解:

$$\begin{array}{ll} \min_x & \frac{1}{2}x^TGx + c^Tx \\ \text{subject to} & a_i^Tx = b_i, \ i \in \mathcal{A}(x^*) \end{array}$$

凸二次规划的KKT条件

$$L(x,\lambda) = \frac{1}{2}x^TGx + c^Tx + \sum_{i\in\mathcal{E}}\lambda_i(a_i^Tx - b_i) + \sum_{i\in\mathcal{I}}\lambda_i(a_i^Tx - b_i)$$

$$\nabla_x L(x^*,\lambda^*) = Gx^* + c + \sum_{i\in\mathcal{E}}\lambda_i^*a_i + \sum_{i\in\mathcal{I}}\lambda_i^*a_i = 0$$

$a_i^Tx^* = b_i \quad i \in \mathcal{E} \qquad a_i^Tx^* \le 1 \quad i \in \mathcal{I}$

$\lambda_i^* \ge 0 \quad i \in \mathcal{I}$

$\lambda_i^*(a_i^Tx^* - b_i) = 0 \quad i \in \mathcal{I}$

$$\Longrightarrow \begin{cases} Gx^* + c + \sum_{i\in\mathcal{A}(x^*)}\lambda_i^*a_i = 0 \\ a_i^Tx^* = b_i \quad i \in \mathcal{A}(x^*) \end{cases}$$

等式约束问题的KKT条件

$L(x,\lambda)=\frac{1}{2}x^TGx+c^Tx+\sum_{i\in A(x)}\lambda_i(a_i^Tx-b_i)$

$$\begin{cases}\nabla_x L(x^*,\lambda^*)=Gx^*+c+\sum_{i\in A(x)}\lambda_i^* a_i=0\\ a_i^Tx=b_i, \quad i\in A(x^*)\end{cases}$$

Proof.

7.4

积极集方法 (active-set method)

基本思想

只需求出最优积极集 $\mathcal{A}(x^*)$. 从可行点 x^1 及其积极集 $\mathcal{A}_1$ 出发, 按目标函数减少的原则, 逐步调整积极集, 直到 $\mathcal{A}(x^*)$

$X^1\in\Phi$. $\mathcal{A}^1 \longrightarrow \bar{x}^1, \lambda^1$

① $x_1\neq\bar{x}_1$

(1) 若 $\bar{x}_1\in\Phi$ 令 $x_2=\bar{x}_1, \mathcal{A}_2$

(2) 若 $\bar{x}_1\notin\Phi$

② $x_1=\bar{x}_1$ λ^1

(1) 若 ______ x_1是KKT点

初始步

选取可行点 x^1, 其积极集为 $\mathcal{A}_1$, 设 $a_i, i\in\mathcal{A}_1$ 线性无关, 求解

$$\begin{array}{ll}\min_x & \frac{1}{2}x^TGx+c^Tx\\ \text{subject to} & a_i^Tx=b_i,\ i\in\mathcal{A}_1\end{array}$$

记其最优解为 $\bar{x}^1$, 乘子 λ^1

7.5

---

case: $\bar{x}^1\neq x^1$

若 $\bar{x}^1\neq x^1$, 则 $\frac{1}{2}(\bar{x}^1)^TG\bar{x}^1+c^T\bar{x}^1<\frac{1}{2}(x^1)^TGx^1+c^Tx^1$

1. 若 $\bar{x}^1\in\Phi$, 则令 $x^2=\bar{x}^1$, 其积极集为 $\mathcal{A}_2$

2. 若 $\bar{x}^1\notin\Phi$, 令 $d^1=\bar{x}^1-x^1$. 从 x^1 沿 d^1 向 $\bar{x}^1$ 走, 目标函数 $Q(x)$ 减少; 走 α_1 的距离, 直到某个约束 $a_i^Tx<b_i$ 先变为 $a_i^Tx=b_i$. 记
$$x^2=x^1+\alpha_1 d^1,$$
$\mathcal{A}_2$ 为 x^2 的积极集

7.6

---

case: $\bar{x}^1=x^1$

若 $\bar{x}^1=x^1$,

1. 若 $\lambda_i^1\geq 0, i\in\mathcal{I}\cap\mathcal{A}_1$, 则 $\bar{x}^1$ 为凸二次规划的 KKT 点, 即最优解 (exercise: 证明 $\bar{x}^1$ 为 KKT 点)

2. 否则, 令
$$k=\arg\min_i\{\lambda_i^1|i\in\mathcal{I}\cap\mathcal{A}_1\}$$
则 $\lambda_k^1<0$. 取 $x^2=\bar{x}^1$, 令 $\mathcal{A}_2=\mathcal{A}_1/\{k\}$, 求解
$$\begin{array}{ll}\min_x & \frac{1}{2}x^TGx+c^Tx\\ \text{subject to} & a_i^Tx=b_i,\ i\in\mathcal{A}_2\end{array}$$
令其最优解为 $\bar{x}^2$, 这时 $\bar{x}^2\neq x^2$

7.7

积极集法求解凸二次规划

Step 0 取 $x^1 \in \Phi$, 求积极集 $\mathcal{A}_1$, 令 $k=1$

Step 1 求解等式约束二次规划

$$\begin{array}{ll} \min_x & \frac{1}{2}x^TGx + c^Tx \\ \text{subject to} & a_i^Tx = b_i,\ i \in \mathcal{A}_k \end{array}$$

令其解为 $\bar{x}^k, \lambda^k$, 记 $d^k = \bar{x}^k - x^k$

Step 2 若 $d^k = 0$, 去 Step 3; 否则, 去 Step 4

Step 3 若 $\lambda_i^k \geq 0, i \in \mathcal{A}_k \cap \mathcal{I}$, 则 x^k 为最优解, 停止; 否则, 求 $\lambda_l^k = \min\{\lambda_i^k | i \in \mathcal{A}_k \cap \mathcal{I}\}$, 令 $x^{k+1} = x^k$, $\mathcal{A}_{k+1} = \mathcal{A}_k / \{l\}$

Step 4 若 $\bar{x}^k \in \Phi$, 令 $x^{k+1} = \bar{x}^k$, $\mathcal{A}_{k+1}$ 为 $\bar{x}^k$ 处积极集; 否则, 计算

$$\alpha_k = \min\{\frac{-(a_i^Tx^k - b_i)}{a_i^Td^k} | i \in \mathcal{I}/\mathcal{A}_k, a_i^Td^k > 0\}$$

令 $x^{k+1} = x^k + \alpha_k d^k$, $\mathcal{A}_{k+1}$ 为 x^{k+1} 处积极集

Step 5 令 $k = k+1$, 回 Step 2

7.8

作业 1. 利用积极集法求解

$$\begin{array}{ll} \min_x & \frac{1}{2}x_1^2 + x_2^2 - 2x_1 - 4x_2 \\ \text{subject to} & x_1 + x_2 - 1 \leq 0,\ x_1 \geq 0,\ x_2 \geq 0 \end{array}$$

7.9

2 局部 SQP 方法

等式约束问题

- 考虑等式约束问题

$$\min_x f(x) \quad \text{subject to } g(x) = 0$$

其中 $f: \mathbb{R}^n \to \mathbb{R}$ 和 $g: \mathbb{R}^n \to \mathbb{R}^m$ 是光滑函数

$g = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$

- 令 $A(x) = [\nabla g_1(x), \ldots, \nabla g_m(x)]^T \in \mathbb{R}^{m\times n}$
- KKT 条件为

$$F(x,\lambda) := \begin{bmatrix} \nabla f(x) - A(x)^T\lambda \\ g(x) \end{bmatrix} = 0$$

- $F(x,\lambda)$ 的雅可比矩阵为

$$\nabla F(x,\lambda) = \begin{bmatrix} \nabla_{xx}^2 L(x,\lambda) & -A(x)^T \\ A(x) & 0 \end{bmatrix}$$

- 若 LICQ 在 x^k 成立, 则 $A(x^k)$ 行满秩; 若二阶充分性条件成立, 则 $\nabla_{xx}^2 L(x^k,\lambda^k)$ 在 $A(x^k)$ 的零空间上正定. 则 $\nabla F(x^k,\lambda^k)$ 可逆

7.10

$\cdot x^k$ $\begin{cases} \min & \nabla f(x^k)^T(x-x^k) + \frac{1}{2}(x-x^k)^T\nabla^2 f(x^k)(x-x^k) \\ \text{s.t.} & g_1(x^k) + \nabla g_1(x^k)^T(x-x^k) = 0 \\ & \vdots \\ & g_m(x^k) + \nabla g_m(x^k)^T(x-x^k) = 0 \end{cases}$

令 $d = x - x^k$ $\begin{cases} \min_d & \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d \\ \text{s.t.} & A(x^k) d = -g(x^k) \end{cases}$

KKT条件 $L(x,\lambda) = f(x) - \lambda^T g(x)$

$\nabla f(x) - A(x)^T\lambda = 0 \qquad g(x) = 0$

$g(x): \mathbb{R}^n \to \mathbb{R}^m$ 雅可比阵

$A(x) = \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x)^T \end{bmatrix} \in \mathbb{R}^{m\times n}$

$\begin{bmatrix} \nabla f(x) - A(x)^T\lambda \\ g(x) \end{bmatrix} = 0$ 非线性方程组

牛顿法

relation between Newton method and SQP method

- 令 $W_k = \nabla^2_{xx} L(x^k, \lambda^k)$, $A_k = A(x^k)$, $g^k = g(x^k)$
- Newton 法求解 $F(x, \lambda) = 0$

$$\begin{bmatrix} x^{k+1} \\ \lambda^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ \lambda^k \end{bmatrix} + \begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla f(x^k) + A_k^T \lambda^k \\ -g^k \end{bmatrix}$$

- 令 $d^k = x^{k+1} - x^k$, 则

$$\begin{bmatrix} d^k \\ \lambda^{k+1} \end{bmatrix} = \begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla f(x^k) \\ -g^k \end{bmatrix}$$

- 考虑二次规划问题

$$\begin{aligned} \min_p \quad & \tfrac{1}{2} p^T W_k p + \nabla f(x^k)^T p \\ \text{subject to} \quad & A_k p + g^k = 0 \end{aligned}$$

- 它的 KKT 点 (p^k, μ^k) 满足

$$\begin{bmatrix} p^k \\ \mu^k \end{bmatrix} = \begin{bmatrix} W_k & -A_k^T \\ A_k & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla f(x^k) \\ -g^k \end{bmatrix}$$

7.11

local SQP method

局部 SQP 方法

Step 0 取 x^0, λ^0, 令 $k = 0$

Step 1 求解二次规划问题

$$\begin{aligned} \min_p \quad & \tfrac{1}{2} p^T W_k p + \nabla f(x^k)^T p \\ \text{subject to} \quad & A_k p + g^k = 0 \end{aligned}$$

令其解为 p^k, μ^k

Step 2 若 x^k 满足最优性条件, 停止; 否则, 令 $x^{k+1} = x^k + p^k$, $\lambda^{k+1} = \mu^k$, $k = k+1$, 回 Step 1

7.12

SQP methods for (NLP)

- 考虑 (NLP)

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & g_i(x) = 0, \ i \in \mathcal{E} \\ & g_i(x) \le 0, \ i \in \mathcal{I} \end{aligned}$$

- SQP 中对应的二次规划问题

$$\begin{aligned} \min_x \quad & \tfrac{1}{2} p^T W_k p + \nabla f(x^k)^T p \\ \text{subject to} \quad & \nabla g_i(x^k)^T p + g_i(x^k) = 0, \ i \in \mathcal{E} \\ & \nabla g_i(x^k)^T p + g_i(x^k) \le 0, \ i \in \mathcal{I} \end{aligned}$$

7.13

收敛性

定理

设 x^* 是 (NLP) 的局部最优解, λ^* 是其对应的乘子, 且

1. LICQ 在 x^* 处成立
2. 严格互补条件成立,$\{i | g_i(x^*) = \lambda_i^* = 0,\ i \in \mathcal{I}\} = \emptyset$
3. 二阶充分性条件在 (x^*, λ^*) 成立

则当 (x^k, λ^k) 充分接近 (x^*, λ^*) 时, 存在二次规划子问题的局部最优解 $\bar{x}^k$, 它的积极集与 x^* 处的积极集相同

Remark

上述定理表明, 当 (x^k, λ^k) 充分接近 (x^*, λ^*) 时, $\mathcal{A}_k = \mathcal{A}_*$, 二次规划子问题变为等式约束二次规划, SQP 方法可视为牛顿法

7.14

not covered

- line search SQP method
- the Maratos effect
- interior point method for (NLP)

7.15