http://shijuanfeng.blogbus.com/
来自我的同事 LPJ

内点法

凸优化第十一章
2014年1月20日

提纲

- 不等式约束的极小化问题
- 对数障碍函数和中心路径
- 障碍方法
- 可行性和阶段1方法
- 自和谐条件下的复杂性分析
- 广义不等式问题
- 原对偶内点法

不等式约束的极小化问题

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & Ax = b, \end{array} \qquad (1)$$

其中 $f_0, \dots, f_m : \mathbf{R}^n \to \mathbf{R}$ 是二次可微的凸函数；

$A \in \mathbf{R}^{p \times n}$ $\quad$ $\mathbf{rank}\, A = p < n$

$$\begin{array}{rcll} Ax^\star = b, \quad f_i(x^\star) & \le & 0, & i = 1, \dots, m \\ \lambda^\star & \succeq & 0 & \\ \nabla f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star \nabla f_i(x^\star) + A^T \nu^\star & = & 0 & \\ \lambda_i^\star f_i(x^\star) & = & 0, & i = 1, \dots, m. \end{array}$$

不等式约束的极小化问题

- 假设：目标函数、不等式约束函数是凸函数，并二次可微，等式约束是仿射函数。最优原变量及最优对偶变量满足KKT条件。
- Newton方法用来解决线性等式约束优化问题，目标函数二次可微，通过reducing为一系列线性等式约束二次问题来解决。
- 而内点法用来解决具有线性等式和不等式约束问题，通过reducing为一系列线性等式约束问题来解决。

对数障碍函数和中心路径

- 把不等式约束问题近似转换成等式约束问题

$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \\ \text{subject to} & Ax = b, \end{array}$$

$$I_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0. \end{cases}$$

- 障碍方法的基本思想是用以下函数近似示性函数

$$\widehat{I}_-(u) = -(1/t)\log(-u), \qquad \mathbf{dom}\,\widehat{I}_- = -\mathbf{R}_{++},$$

对数障碍函数和中心路径

$$\begin{array}{ll} \text{minimize} & f_0(x) + \sum_{i=1}^m -(1/t)\log(-f_i(x)) \\ \text{subject to} & Ax = b. \end{array}$$

- 对数障碍函数

$$\phi(x) = -\sum_{i=1}^m \log(-f_i(x)),$$

$$\begin{array}{ll} \text{minimize} & t f_0(x) + \phi(x) \\ \text{subject to} & Ax = b, \end{array}$$

对任意t>0,我们用$x^\star(t)$表示上式的解，称$x^\star(t)$,t>0为中心点，将这些点的集合定义为问题（1）的中心路径。

对数障碍函数和中心路径

- 中心路径的重要性质：每个中心点产生对偶可行解，因而给出最优值的一个下界。

$$\lambda_i^\star(t) = -\frac{1}{t f_i(x^\star(t))}, \quad i = 1, \ldots, m, \qquad \nu^\star(t) = \hat{\nu}/t.$$

- 无约束极小化方法:保证达到预定精度

$$\begin{array}{ll} \text{minimize} & (m/\epsilon) f_0(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

障碍方法

given strictly feasible x, $t := t^{(0)} > 0$, $\mu > 1$, tolerance $\epsilon > 0$.

repeat

1. *Centering step.*
   Compute $x^\star(t)$ by minimizing $tf_0 + \phi$, subject to $Ax = b$, starting at x.
2. *Update.* $x := x^\star(t)$.
3. *Stopping criterion.* **quit** if $m/t < \epsilon$.
4. *Increase* t. $t := \mu t$.

- 选择 $t^{(0)}$

$$\inf_{\nu} \left\| t\nabla f_0(x^{(0)}) + \nabla\phi(x^{(0)}) + A^T\nu \right\|_2$$

可行性和阶段1方法

- 定义：障碍方法需要一个严格可行的初始点，如果不知道这样一个可行点，在应用障碍方法之前需要一个预备阶段称为阶段1.
- 基本的阶段1方法
- 考虑 $x \in \mathbf{R}^n$, 的一组不等式和等式方程

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad Ax = b,$$

- $f_i : \mathbf{R}^n \to \mathbf{R}$ 是凸的，具有连续的二阶导数
- 阶段1优化问题：

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- 极小化不可行值的最大值
- 极小化不可行值的和

可行性和阶段1方法

- 采用不可行初始点Newton方法求解阶段1问题

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b. \end{array}$$

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \dots, m \\ & Ax = b, \quad s = 0, \end{array}$$

$$\begin{array}{ll} \text{minimize} & t^{(0)} f_0(x) - \sum_{i=1}^{m} \log(s - f_i(x)) \\ \text{subject to} & Ax = b, \quad s = 0. \end{array}$$

- 当问题不可行时没有好的停止准则，残差不能收敛到0

http://shijuanfeng.blogbus.com/

自和谐条件下的复杂性分析

自和谐假设

- 对于所有的 $t \geq t^{(0)}$， $tf_0 + \phi$是闭的自和谐性函数

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b, \end{array}$$

- 水平集有界。
- 如果 f_i是线性的或二次的，那么对于 $t \geq 0$， 下式是自和谐的。

$$t f_0 - \sum_{i=1}^{m} \log(-f_i)$$

自和谐条件下的复杂性分析

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \\ & Ax = b. \end{array}$$

$$tf_0(x) + \phi(x) = t\sum_{i=1}^n x_i \log x_i - \sum_{i=1}^m \log(g_i - f_i^T x),$$

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Fx \preceq g \\ & Ax = b \\ & x \succeq 0. \end{array}$$

$$tf_0(x) + \phi(x) = t\sum_{i=1}^n x_i \log x_i - \sum_{i=1}^n \log x_i - \sum_{i=1}^m \log(g_i - f_i^T x),$$

自和谐条件下的复杂性分析

- 中心点步骤的Newton迭代次数

$$\frac{m(\mu - 1 - \log \mu)}{\gamma} + c$$

- 总的Newton迭代次数

$$N = \left\lceil \frac{\log(m/(t^{(0)}\epsilon))}{\log \mu} \right\rceil \left(\frac{m(\mu - 1 - \log \mu)}{\gamma} + c \right)$$

$$\mu = 1 + 1/\sqrt{m}.$$

$$\begin{aligned} \mu - 1 - \log \mu &= 1/\sqrt{m} - \log(1 + 1/\sqrt{m}) \\ &\leq 1/\sqrt{m} - 1/\sqrt{m} + 1/(2m) \\ &= 1/(2m) \end{aligned}$$

http://shijuanfeng.blogbus.com/

自和谐条件下的复杂性分析

$$\log \mu = \log(1 + 1/\sqrt{m}) \geq (\log 2)/\sqrt{m}.$$

$$\begin{aligned}
N &\leq \left\lceil \frac{\log(m/(t^{(0)}\epsilon))}{\log \mu} \right\rceil \left(\frac{m(\mu - 1 - \log \mu)}{\gamma} + c \right) \\
&\leq \left\lceil \sqrt{m} \frac{\log(m/(t^{(0)}\epsilon))}{\log 2} \right\rceil \left(\frac{1}{2\gamma} + c \right) \\
&= \left\lceil \sqrt{m} \log_2(m/(t^{(0)}\epsilon)) \right\rceil \left(\frac{1}{2\gamma} + c \right) \\
&\leq c_1 + c_2 \sqrt{m},
\end{aligned}$$

自和谐条件下的复杂性分析

- 可行性问题

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \ldots, m \end{array}$$

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & f_i(x) \leq s, \quad i = 1, \ldots, m \\ & a^T x \leq 1. \end{array}$$

- 采用障碍方法求解，Newton迭代次数上界

$$\left\lceil \sqrt{m+1} \log_2 \frac{m(m+1)GR}{|\bar{p}^\star|} \right\rceil \left(\frac{1}{2\gamma} + c \right)$$

自和谐条件下的复杂性分析

- 阶段1和阶段2结合在一起的复杂性

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- 首先求解阶段1问题

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & f_i(x) \le s, \quad i = 1, \ldots, m \\ & f_0(x) \le M \\ & Ax = b \\ & a^T x \le 1, \end{array}$$

自和谐条件下的复杂性分析

- 阶段1和阶段2结合在一起的复杂性

$$N_{\mathrm{I}} = \left\lceil \sqrt{m+2}\log_2 \frac{(m+1)(m+2)GR}{|\bar{p}^\star|} \right\rceil \left(\frac{1}{2\gamma} + c \right)$$

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \\ & a^T x \leq 1. \end{array}$$

$$N_{\mathrm{II}} = \left\lceil \sqrt{m+1}\log_2 \frac{(m+1)(M-p^\star)}{\epsilon} \right\rceil \left(\frac{1}{2\gamma} + c \right)$$

$$N_{\mathrm{I}} + N_{\mathrm{II}}$$

广义不等式问题

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \ldots, m \\
& Ax = b,
\end{array}
$$

KKT条件：

$$
\begin{array}{rcll}
Ax^\star & = & b & \\
f_i(x^\star) & \preceq_{K_i} & 0, & i = 1, \ldots, m \\
\lambda_i^\star & \succeq_{K_i^*} & 0, & i = 1, \ldots, m \\
\nabla f_0(x^\star) + \sum_{i=1}^m Df_i(x^\star)^T \lambda_i^\star + A^T \nu^\star & = & 0 & \\
\lambda_i^{\star T} f_i(x^\star) & = & 0, & i = 1, \ldots, m,
\end{array}
$$

对数障碍和中心路径

广义不等式的对数障碍函数定义：

$$
\phi(x) = -\sum_{i=1}^m \psi_i(-f_i(x)), \qquad \mathbf{dom}\,\phi = \{x \mid f_i(x) \prec 0,\ i = 1, \ldots, m\}.
$$

广义不等式问题

- 中心路径定义

$$\begin{array}{ll} \text{minimize} & t f_0(x) - \sum_{i=1}^m \psi_i(-f_i(x)) \\ \text{subject to} & Ax = b \end{array}$$

- 中心点应该和某个 $\nu \in \mathbf{R}^p$ 一起满足以下最优性条件

$$\begin{aligned} & t\nabla f_0(x) + \nabla\phi(x) + A^T\nu \\ & \quad = \quad t\nabla f_0(x) + \sum_{i=1}^m Df_i(x)^T \nabla\psi_i(-f_i(x)) + A^T\nu = 0, \end{aligned}$$

- 中心路径上的对偶点
- 定义： $\lambda_i^\star(t) = \frac{1}{t}\nabla\psi_i(-f_i(x^\star(t))),$
- 令 $\nu^\star(t) = \nu/t,$ 那么 $\lambda_1^\star(t), \ldots, \lambda_m^\star(t),$ 和 $\nu^\star(t)$ 一起构成对偶可行解

广义不等式问题

- 中心路径的关键性质可以推广到广义不等式问题
- 阶段1方法也可以推广到广义不等式问题
- 为了确定等式和广义不等式的可行性

$$f_1(x) \preceq_{K_1} 0, \quad \ldots, \quad f_L(x) \preceq_{K_m} 0, \qquad Ax = b,$$

- 可以求解变量x和s的问题

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & f_i(x) \preceq_{K_i} s e_i, \quad i = 1, \ldots, m \\ & Ax = b, \end{array}$$

原对偶内点法

- 仅有一层迭代，即没有障碍方法的内部迭代和外部迭代的区分。每次迭代时同时更新原对偶变量。
- 通过将Newton方法应用于修改的KKT方程确定原对偶内点法的搜索方向。原对偶搜索方向和障碍方法导出的搜索方向相似，但不完全相同。
- 在原对偶内点法中，原对偶迭代值不需要是可行的

Algorithm 11.2 *Primal-dual interior-point method.*

given x that satisfies $f_1(x) < 0, \ldots, f_m(x) < 0$, $\lambda \succ 0$, $\mu > 1$, $\epsilon_{\text{feas}} > 0$, $\epsilon > 0$.

repeat

1. *Determine* t. Set $t := \mu m / \hat{\eta}$.
2. Compute primal-dual search direction Δy_{pd}.
3. *Line search and update.*
   Determine step length $s > 0$ and set $y := y + s\Delta y_{\text{pd}}$.

until $\|r_{\text{pri}}\|_2 \le \epsilon_{\text{feas}}$, $\|r_{\text{dual}}\|_2 \le \epsilon_{\text{feas}}$, and $\hat{\eta} \le \epsilon$.

谢谢