最优化方法

东南大学
计算机&人工智能学院
宋沫飞
songmf@seu.edu.cn

内点法

- 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
- 自和谐条件的复杂性分析

不等式约束的极小化

不等式约束的极小化

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

不等式约束的极小化

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

函数f为凸函数，二次连续可微

不等式约束的极小化

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

函数f为凸函数，二次连续可微

$$A \in \mathbf{R}^{p \times n} \text{ with } \mathbf{rank}\, A = p$$

不等式约束的极小化

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

函数f为凸函数，二次连续可微

$$A \in \mathbf{R}^{p \times n} \text{ with } \mathbf{rank}\, A = p$$

假设最优值 $p^\star$ 为有限且存在

不等式约束的极小化

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

函数f为凸函数，二次连续可微

$$A \in \mathbf{R}^{p \times n} \text{ with } \mathbf{rank}\, A = p$$

假设最优值 $p^\star$ 为有限且存在

假设问题为严格可行：存在 $\tilde{x}$ 满足

不等式约束的极小化

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

函数f为凸函数，二次连续可微

$$A \in \mathbf{R}^{p \times n} \text{ with } \mathbf{rank}\, A = p$$

假设最优值 $p^\star$ 为有限且存在

假设问题为严格可行：存在 $\tilde{x}$ 满足

$$\tilde{x} \in \mathbf{dom}\, f_0, \qquad f_i(\tilde{x}) < 0, \quad i = 1, \ldots, m, \qquad A\tilde{x} = b$$

不等式约束的极小化

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

函数f为凸函数，二次连续可微

$$A \in \mathbf{R}^{p \times n} \text{ with } \mathbf{rank}\, A = p$$

假设最优值 $p^\star$ 为有限且存在

假设问题为严格可行：存在 $\tilde{x}$ 满足

$$\tilde{x} \in \mathbf{dom}\, f_0, \qquad f_i(\tilde{x}) < 0, \quad i = 1, \ldots, m, \qquad A\tilde{x} = b$$

因此，具有强对偶且对偶优化存在

例子

例子

- 线性规划、二次规划、二次约束的二次规划

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

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

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

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

- 其中，$\mathbf{dom}\, f_0 = \mathbf{R}_{++}^n$

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

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

- 其中，$\mathbf{dom}\, f_0 = \mathbf{R}^n_{++}$
- 可微性需要重形式化问题

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

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

- 其中，$\mathbf{dom}\, f_0 = \mathbf{R}_{++}^{n}$
- 可微性需要重形式化问题
    - 分段线性极小化

例子

- 线性规划、二次规划、二次约束的二次规划
- 基于线性不等式约束的熵最大化

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

- 其中，$\mathbf{dom}\ f_0 = \mathbf{R}_{++}^n$
- 可微性需要重形式化问题
  - 分段线性极小化
  - 基于线性规划的范数逼近

对数障碍

对数障碍

- 通过示性函数重形式化不等式约束问题

对数障碍

- 通过示性函数重形式化不等式约束问题

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

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \le 0$, $I_-(u) = \infty$ otherwise

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \le 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \leq 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

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

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \leq 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

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

- 为等式约束问题

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \le 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

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

- 为等式约束问题
- $t > 0$, $-(1/t)\log(-u)$为 I_-的平滑近似

对数障碍

- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \le 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

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

- 为等式约束问题
- $t > 0$, $-(1/t)\log(-u)$为 I_-的平滑近似
- 近似程度随 $t \to \infty$ 提高

对数障碍

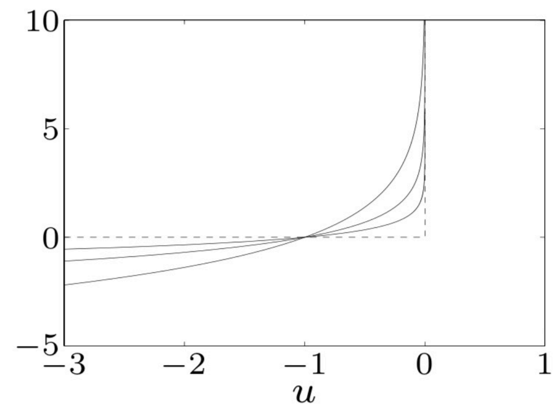
- 通过示性函数重形式化不等式约束问题

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

- 其中，$I_-(u) = 0$ if $u \le 0$, $I_-(u) = \infty$ otherwise
- 通过对数障碍近似

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

- 为等式约束问题
- $t > 0$, $-(1/t)\log(-u)$为 I_-的平滑近似
- 近似程度随 $t \to \infty$ 提高



对数障碍函数

对数障碍函数

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \mathbf{dom}\,\phi = \{x \mid f_1(x) < 0, \ldots, f_m(x) < 0\}$$

对数障碍函数

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \qquad \mathbf{dom}\,\phi = \{x \mid f_1(x) < 0, \ldots, f_m(x) < 0\}$$

- 根据组合规则为凸函数

对数障碍函数

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \mathbf{dom}\,\phi = \{x \mid f_1(x) < 0, \ldots, f_m(x) < 0\}$$

- 根据组合规则为凸函数
- 二阶可微连续函数，导数为

对数障碍函数

$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)), \quad \mathbf{dom}\,\phi = \{x \mid f_1(x) < 0, \ldots, f_m(x) < 0\}$$

- 根据组合规则为凸函数
- 二阶可微连续函数，导数为

$$\begin{aligned}
\nabla\phi(x) &= \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) \\
\nabla^2\phi(x) &= \sum_{i=1}^{m} \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla^2 f_i(x)
\end{aligned}$$

中心路径

中心路径

- 对$t > 0$，定义 $x^*(t)$为如下问题的解

中心路径

- 对$t > 0$，定义 $x^*(t)$为如下问题的解

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

中心路径

- 对$t>0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）

中心路径

- 对$t>0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）
- 中心路径为 $\{x^\star(t) \mid t>0\}$

中心路径

- 对$t > 0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）

- 中心路径为 $\{x^\star(t) \mid t > 0\}$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, 6 \end{array}$$

中心路径

- 对$t > 0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）
- 中心路径为 $\{x^\star(t) \mid t > 0\}$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, 6 \end{array}$$

- 超平面 $c^T x = c^T x^\star(t)$ 与 ϕ 的等值线

中心路径

- 对$t > 0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）
- 中心路径为 $\{x^\star(t) \mid t > 0\}$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, 6 \end{array}$$

- 超平面 $c^T x = c^T x^\star(t)$ 与 ϕ 的等值线
- 相切于$x^*(t)$

中心路径

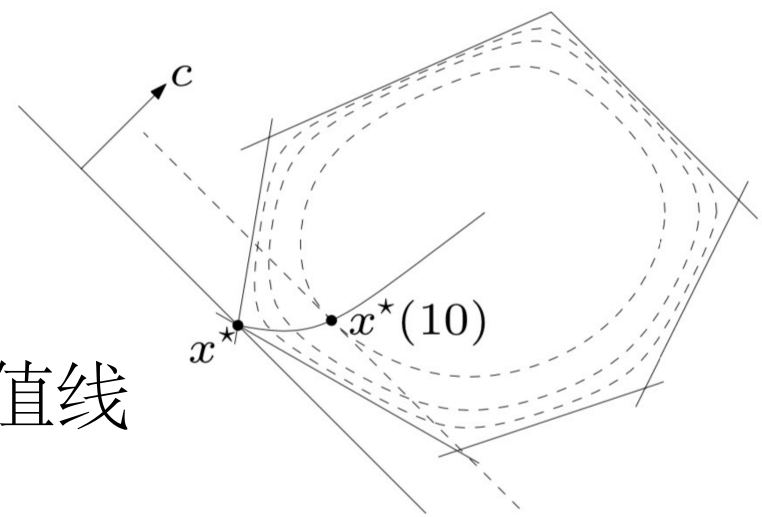
- 对$t > 0$，定义 $x^*(t)$为如下问题的解

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

  - （此时，假设$x^*(t)$存在且唯一）

- 中心路径为 $\{x^\star(t) \mid t > 0\}$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i, \quad i = 1, \ldots, 6 \end{array}$$



- 超平面 $c^T x = c^T x^\star(t)$ 与 ϕ 的等值线
- 相切于$x^*(t)$

中心路径的对偶点

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

- 因此，$x^*(t)$最小化**Lagrangian**函数

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

- 因此，$x^*(t)$最小化**Lagrangian**函数

$$L(x, \lambda^\star(t), \nu^\star(t)) = f_0(x) + \sum_{i=1}^{m} \lambda_i^\star(t) f_i(x) + \nu^\star(t)^T (Ax - b)$$

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

- 因此，$x^*(t)$最小化**Lagrangian**函数

$$L(x, \lambda^\star(t), \nu^\star(t)) = f_0(x) + \sum_{i=1}^{m} \lambda_i^\star(t) f_i(x) + \nu^\star(t)^T (Ax - b)$$

- 其中，定义 $\lambda_i^\star(t) = 1/(-t f_i(x^\star(t)))$ and $\nu^\star(t) = w/t$

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

- 因此，$x^*(t)$最小化**Lagrangian**函数

$$L(x, \lambda^\star(t), \nu^\star(t)) = f_0(x) + \sum_{i=1}^{m} \lambda_i^\star(t) f_i(x) + \nu^\star(t)^T(Ax - b)$$

- 其中，定义 $\lambda_i^\star(t) = 1/(-t f_i(x^\star(t)))$ and $\nu^\star(t) = w/t$
- 这确认了直观的想法：$f_0(x^\star(t)) \to p^\star$

中心路径的对偶点

- $x = x^*(t)$，若存在w满足

$$t\nabla f_0(x) + \sum_{i=1}^{m} \frac{1}{-f_i(x)} \nabla f_i(x) + A^T w = 0, \qquad Ax = b$$

- 因此，$x^*(t)$最小化**Lagrangian**函数

$$L(x, \lambda^\star(t), \nu^\star(t)) = f_0(x) + \sum_{i=1}^{m} \lambda_i^\star(t) f_i(x) + \nu^\star(t)^T (Ax - b)$$

- 其中，定义 $\lambda_i^\star(t) = 1/(-t f_i(x^\star(t)))$ and $\nu^\star(t) = w/t$
- 这确认了直观的想法：$f_0(x^\star(t)) \to p^\star$

$$\begin{aligned} p^\star &\geq g(\lambda^\star(t), \nu^\star(t)) \\ &= L(x^\star(t), \lambda^\star(t), \nu^\star(t)) \\ &= f_0(x^\star(t)) - m/t \end{aligned}$$

基于KKT条件的解释

基于KKT条件的解释

- $x = x^\star(t)$, $\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$ 满足

基于KKT条件的解释

- $x = x^\star(t),\ \lambda = \lambda^\star(t),\ \nu = \nu^\star(t)$ 满足
- 原约束：$f_i(x) \le 0,\ i = 1, \ldots, m,\ Ax = b$

基于KKT条件的解释

- $x = x^\star(t)$, $\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$ 满足
- 原约束：$f_i(x) \le 0,\ i = 1, \ldots, m,\ Ax = b$
- 对偶约束： $\lambda \succeq 0$

基于KKT条件的解释

- $x = x^\star(t)$, $\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$　满足
- 原约束：$f_i(x) \leq 0,\ i = 1, \ldots, m,\ Ax = b$
- 对偶约束：$\lambda \succeq 0$
- 互补松弛条件：$-\lambda_i f_i(x) = 1/t,\ i = 1, \ldots, m$

基于KKT条件的解释

- $x = x^\star(t),\ \lambda = \lambda^\star(t),\ \nu = \nu^\star(t)$　满足
- 原约束：$f_i(x) \le 0,\ i = 1, \ldots, m,\ Ax = b$
- 对偶约束：$\lambda \succeq 0$
- 互补松弛条件：$-\lambda_i f_i(x) = 1/t,\ i = 1, \ldots, m$
- **Lagrangian**函数的梯度

基于KKT条件的解释

- $x = x^\star(t)$, $\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$ 满足
- 原约束：$f_i(x) \le 0,\ i = 1, \ldots, m,\ Ax = b$
- 对偶约束： $\lambda \succeq 0$
- 互补松弛条件： $-\lambda_i f_i(x) = 1/t,\ i = 1, \ldots, m$
- **Lagrangian**函数的梯度

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T \nu = 0$$

基于KKT条件的解释

- $x = x^\star(t),\ \lambda = \lambda^\star(t),\ \nu = \nu^\star(t)$ 满足
- 原约束：$f_i(x) \le 0,\ i = 1, \ldots, m,\ Ax = b$
- 对偶约束： $\lambda \succeq 0$
- 互补松弛条件： $-\lambda_i f_i(x) = 1/t,\ i = 1, \ldots, m$
- **Lagrangian**函数的梯度

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T \nu = 0$$

- 和**KKT**的不同是条件**3**替换了 $\lambda_i f_i(x) = 0$

障碍法

障碍法

算法 11.1　障碍方法。

给定严格可行点 x，$t := t^{(0)} > 0$，$\mu > 1$，误差阈值 $\epsilon > 0$。

重复进行

1. **中心点步骤。**

   从 x 开始，在 $Ax = b$ 的约束下极小化 $tf_0 + \phi$，最终确定 $x^\star(t)$。

2. **改进。** $x := x^\star(t)$。
3. **停止准则。** 如果 $m/t < \epsilon$ 则退出。
4. **增加 t。** $t := \mu t$。

障碍法

算法 11.1 障碍方法。

给定严格可行点 x，$t := t^{(0)} > 0$，$\mu > 1$，误差阈值 $\epsilon > 0$。

重复进行

1. **中心点步骤。**
   从 x 开始，在 $Ax = b$ 的约束下极小化 $tf_0 + \phi$，最终确定 $x^\star(t)$。
2. **改进。** $x := x^\star(t)$。
3. **停止准则。** 如果 $m/t < \epsilon$ 则退出。
4. **增加 t。** $t := \mu t$。

- 终结时，$f_0(x) - p^\star \le \epsilon$ 停止条件来自 $f_0(x^\star(t)) - p^\star \le m/t$

障碍法

算法 11.1　障碍方法。

给定严格可行点 x，$t := t^{(0)} > 0$，$\mu > 1$，误差阈值 $\epsilon > 0$。

重复进行

1. **中心点步骤。**
   从 x 开始，在 $Ax = b$ 的约束下极小化 $tf_0 + \phi$，最终确定 $x^\star(t)$。
2. **改进。** $x := x^\star(t)$。
3. **停止准则。** 如果 $m/t < \epsilon$ 则退出。
4. **增加 t。** $t := \mu t$。

- 终结时，$f_0(x) - p^\star \le \epsilon$ 停止条件来自 $f_0(x^\star(t)) - p^\star \le m/t$
- 使用**Newton**方法完成中心点步骤，从当前点*x*开始

障碍法

算法 11.1　障碍方法。

给定严格可行点 x，$t := t^{(0)} > 0$，$\mu > 1$，误差阈值 $\epsilon > 0$。

重复进行

1. **中心点步骤。**
   从 x 开始，在 $Ax = b$ 的约束下极小化 $tf_0 + \phi$，最终确定 $x^\star(t)$。
2. **改进。** $x := x^\star(t)$。
3. **停止准则。** 如果 $m/t < \epsilon$ 则退出。
4. **增加 t。** $t := \mu t$。

- 终结时，$f_0(x) - p^\star \leq \epsilon$ 停止条件来自 $f_0(x^\star(t)) - p^\star \leq m/t$
- 使用**Newton**方法完成中心点步骤，从当前点*x*开始
- μ的选择需要权衡：取大表示较少的外部循环，更多的内部训练；通常设置为 $\mu = 10\text{–}20$

障碍法

算法 11.1　障碍方法。

给定严格可行点 x，$t := t^{(0)} > 0$，$\mu > 1$，误差阈值 $\epsilon > 0$。

重复进行

1. **中心点步骤。**
   从 x 开始，在 $Ax = b$ 的约束下极小化 $tf_0 + \phi$，最终确定 $x^\star(t)$。
2. **改进。** $x := x^\star(t)$。
3. **停止准则。** 如果 $m/t < \epsilon$ 则退出。
4. **增加 t。** $t := \mu t$。

- 终结时，$f_0(x) - p^\star \le \epsilon$ 停止条件来自 $f_0(x^\star(t)) - p^\star \le m/t$
- 使用**Newton**方法完成中心点步骤，从当前点*x*开始
- μ的选择需要权衡：取大表示较少的外部循环，更多的内部训练；通常设置为 $\mu = 10$–20
- 初始点选择有很多启发式策略

收敛性分析

收敛性分析

- 外部训练的次数：精确的

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤

$$\text{minimize} \quad t f_0(x) + \phi(x)$$

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤
- 中心点步骤 minimize $tf_0(x) + \phi(x)$

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤
- 中心点步骤 minimize $t f_0(x) + \phi(x)$
- 见**Newton**法的收敛性分析

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤
- 中心点步骤 minimize $t f_0(x) + \phi(x)$
- 见**Newton**法的收敛性分析
- 对 $t \geq t^{(0)}$ ， $t f_0 + \phi$ 必须有闭的下水平集

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤
- 中心点步骤 minimize $\quad t f_0(x) + \phi(x)$
- 见**Newton**法的收敛性分析
- 对 $t \geq t^{(0)}$ ， $t f_0 + \phi$ 必须有闭的下水平集
- 经典分析需要强凸性、**Lipschitz**条件

收敛性分析

- 外部训练的次数：精确的

$$\left\lceil \frac{\log(m/(\epsilon t^{(0)}))}{\log \mu} \right\rceil$$

- 加上初始化中心点步骤
- 中心点步骤 minimize $\quad t f_0(x) + \phi(x)$
- 见**Newton**法的收敛性分析
- 对 $t \geq t^{(0)}$ ， $t f_0 + \phi$ 必须有闭的下水平集
- 经典分析需要强凸性、**Lipschitz**条件
- 基于自适应的分析需要 $t f_0 + \phi$ 为自适应函数

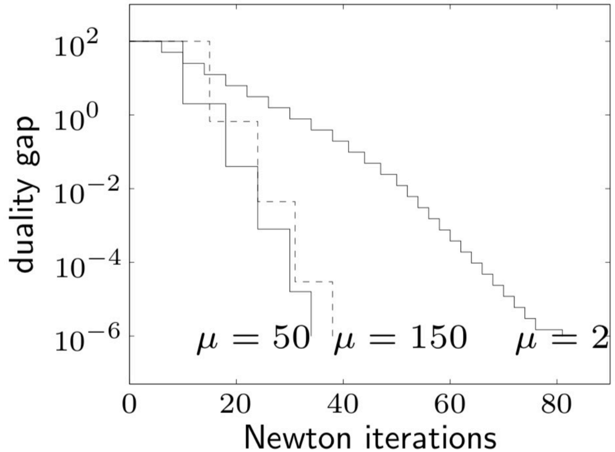
例

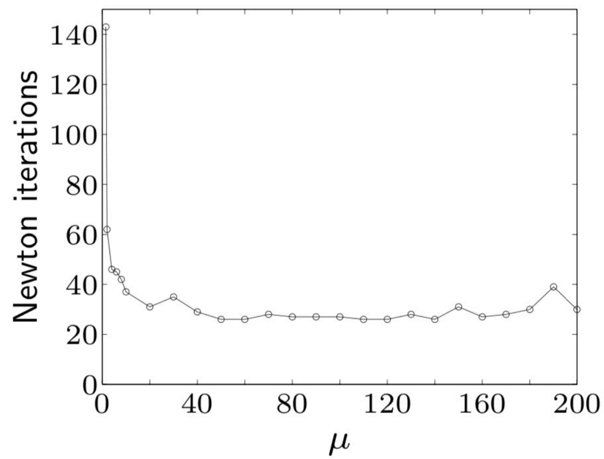
例

- 不等式形式的线性规划（m=**100**个不等式，n=**50**个优化变量）

例

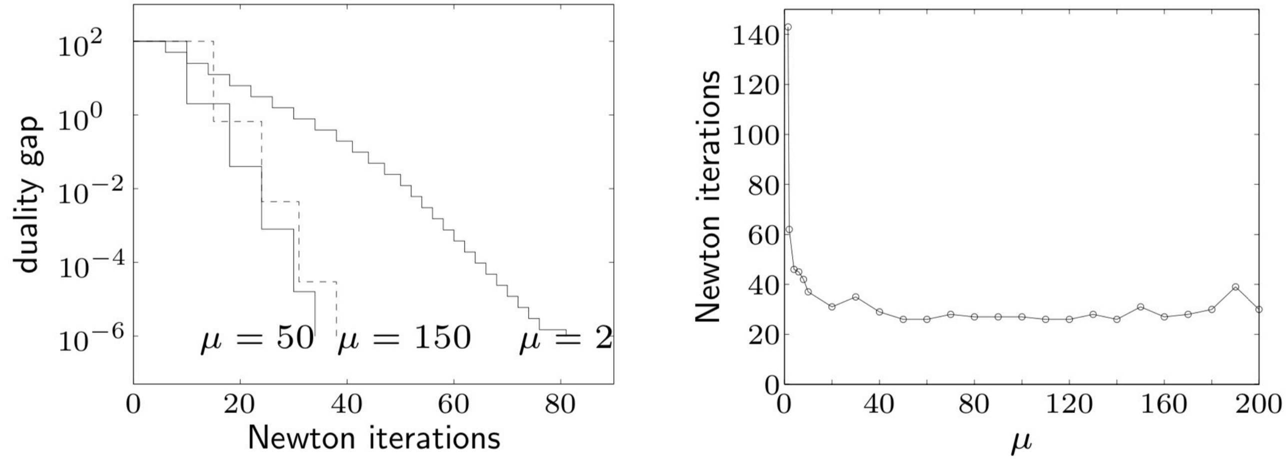
- 不等式形式的线性规划（m=**100**个不等式，n=**50**个优化变量）





例

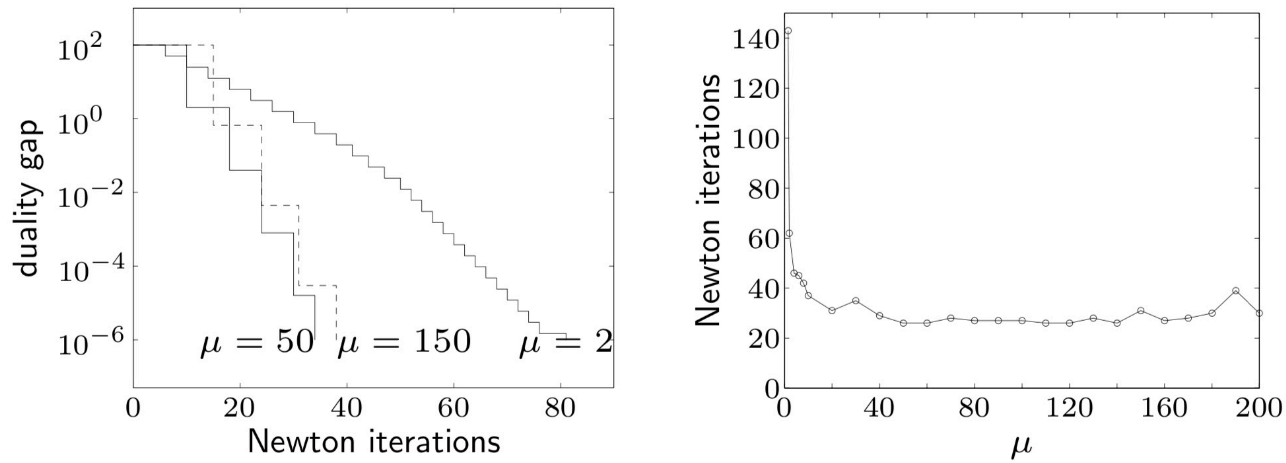
- 不等式形式的线性规划（m=**100**个不等式，n=**50**个优化变量）



- 从中心路径的点x开始（ $t^{(0)} = 1$，对偶间隙为**100**）

例

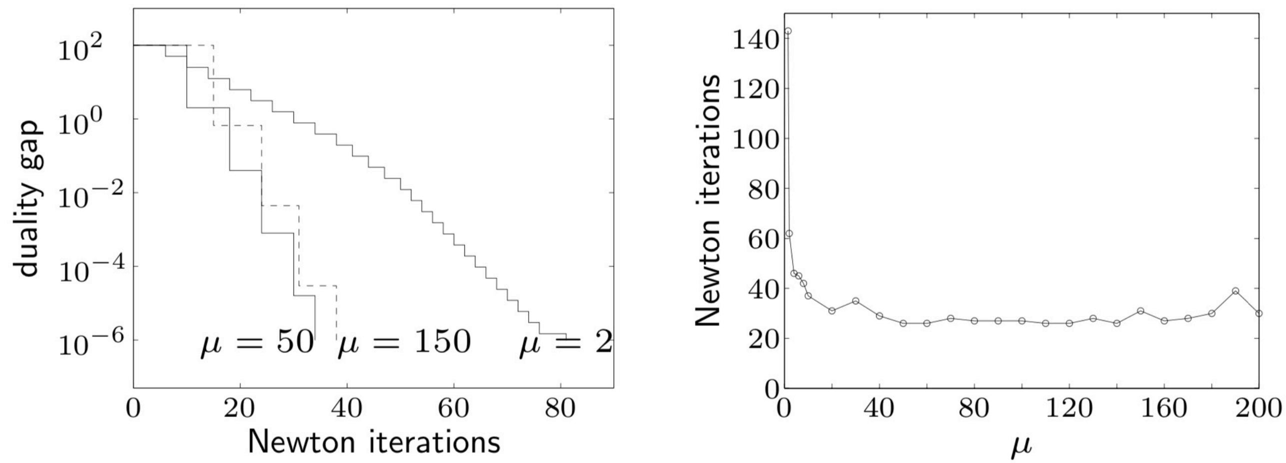
- 不等式形式的线性规划（***m*=100**个不等式，***n*=50**个优化变量）



- 从中心路径的点x开始（ $t^{(0)} = 1$ ，对偶间隙为**100**）
- 当t=**10**^{8}时，停止

例

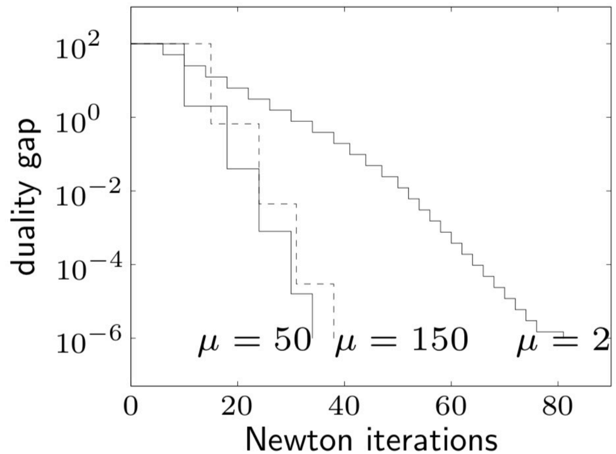
- 不等式形式的线性规划（m=**100**个不等式，n=**50**个优化变量）

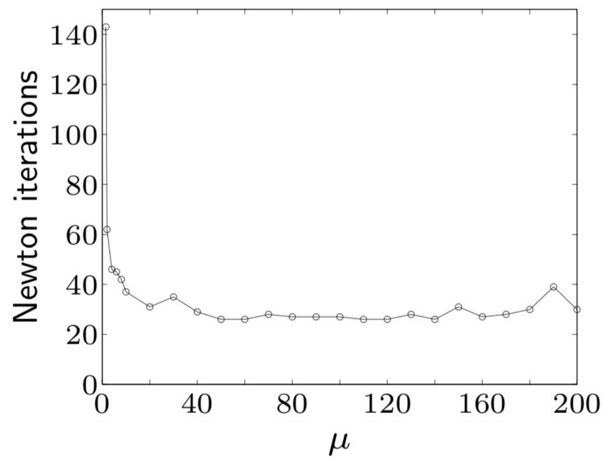


- 从中心路径的点x开始（ $t^{(0)} = 1$，对偶间隙为**100**）
- 当t=**10**^{8}时，停止
- 中心点步骤使用带回溯的**Newton**法

例

- 不等式形式的线性规划（m=**100**个不等式，n=**50**个优化变量）





- 从中心路径的点x开始（ $t^{(0)} = 1$，对偶间隙为**100**）
- 当t=**10**^{8}时，停止
- 中心点步骤使用带回溯的**Newton**法
- 若 $\mu \geq 10$，**Newton**法的整体迭代次数不会太敏感

一簇标准的线性规划

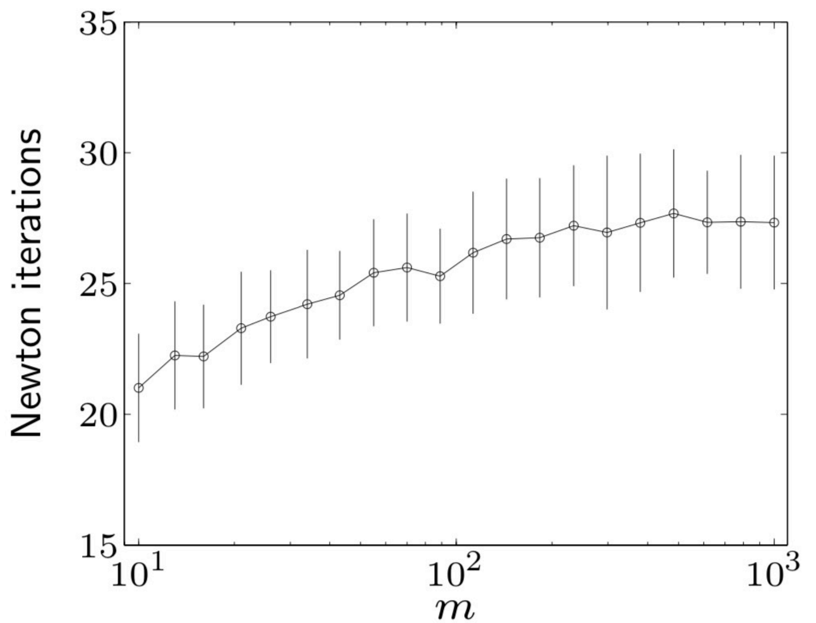
一簇标准的线性规划

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0\end{array}$$

一簇标准的线性规划

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$

- m=10，…，1000；对每一个m，求解100个随机生成的实例



- 当问题维数变化比值为**100:1**时，迭代次数增长的较为缓慢

可行性和阶段1方法

可行性和阶段1方法

- 可行性问题：寻找*x*满足 $f_i(x) \le 0, \quad i = 1, \ldots, m, \qquad Ax = b$

可行性和阶段1方法

- 可行性问题：寻找x满足 $f_i(x) \leq 0, \quad i=1,\ldots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点

可行性和阶段1方法

- 可行性问题：寻找*x*满足 $f_i(x) \le 0, \quad i=1,\dots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

可行性和阶段1方法

- 可行性问题：寻找*x*满足 $f_i(x) \le 0, \quad i=1,\dots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

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

可行性和阶段1方法

- 可行性问题：寻找*x*满足 $f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad Ax = b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

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

- 若*x*，*s*可行，且*s*<**0**，则*x*是严格可行的

可行性和阶段1方法

- 可行性问题：寻找x满足 $f_i(x) \le 0, \quad i=1,\ldots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

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

- 若x，s可行，且s**<0**，则x是严格可行的
- 若最优值为正，则问题为非可行性的

可行性和阶段1方法

- 可行性问题：寻找x满足 $f_i(x) \leq 0, \quad i=1,\ldots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

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

- 若x，s可行，且s<**0**，则x是严格可行的
- 若最优值为正，则问题为非可行性的
- 若最优值为**0**且存在，则问题为可行的（但不是严格可行的）

可行性和阶段1方法

- 可行性问题：寻找x满足 $f_i(x) \le 0, \quad i=1,\ldots,m, \qquad Ax=b$
- 阶段**1:**为障碍法计算研究可行的起始点
- 基本阶段**1**方法

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

- 若x，s可行，且s<**0**，则x是严格可行的
- 若最优值为正，则问题为非可行性的
- 若最优值为**0**且存在，则问题为可行的（但不是严格可行的）
- 若最优值为**0**且不可达，则问题为非可行的

不可行值之和

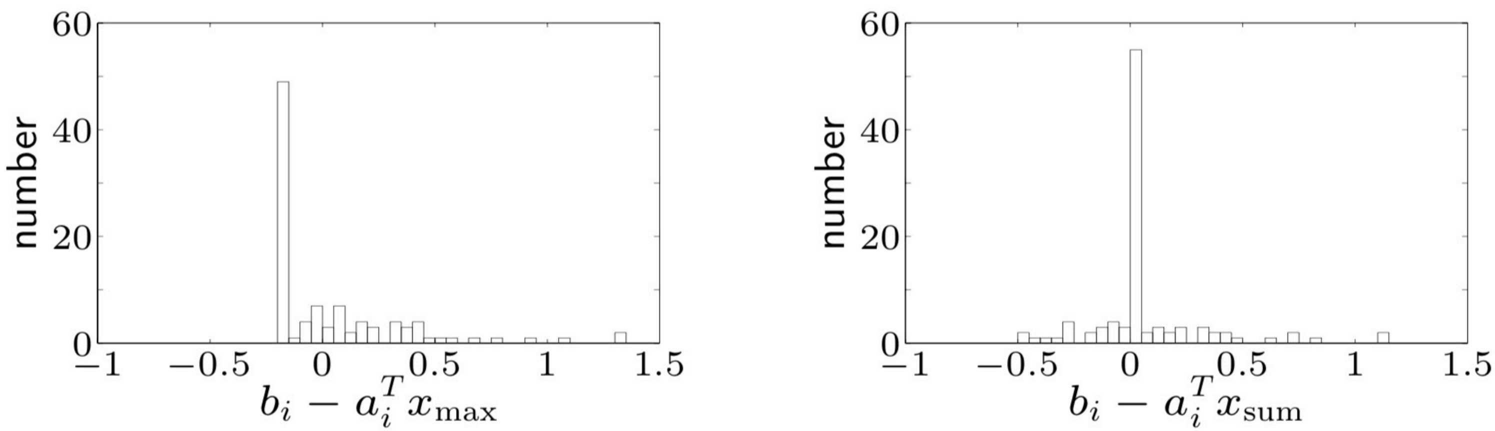
不可行值之和

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

不可行值之和

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

- 对不可行点，比基本阶段**1**方法产生满足更多不等式约束的解
- 例（由**50**个变量和**100**个线性不等式构成的不可行集）



- 左图：基本阶段**1**方法，满足**39**个不等式
- 右图：不可行值之和，满足**79**个不等式

例：一族线性不等式

例：一族线性不等式

$$Ax \preceq b + \gamma \Delta b$$

例：一族线性不等式

$$Ax \preceq b + \gamma \Delta b$$

- 所需的数据对 $\gamma > 0$ 是严格可行的，对 $\gamma \leq 0$ 是不可行的

例：一族线性不等式

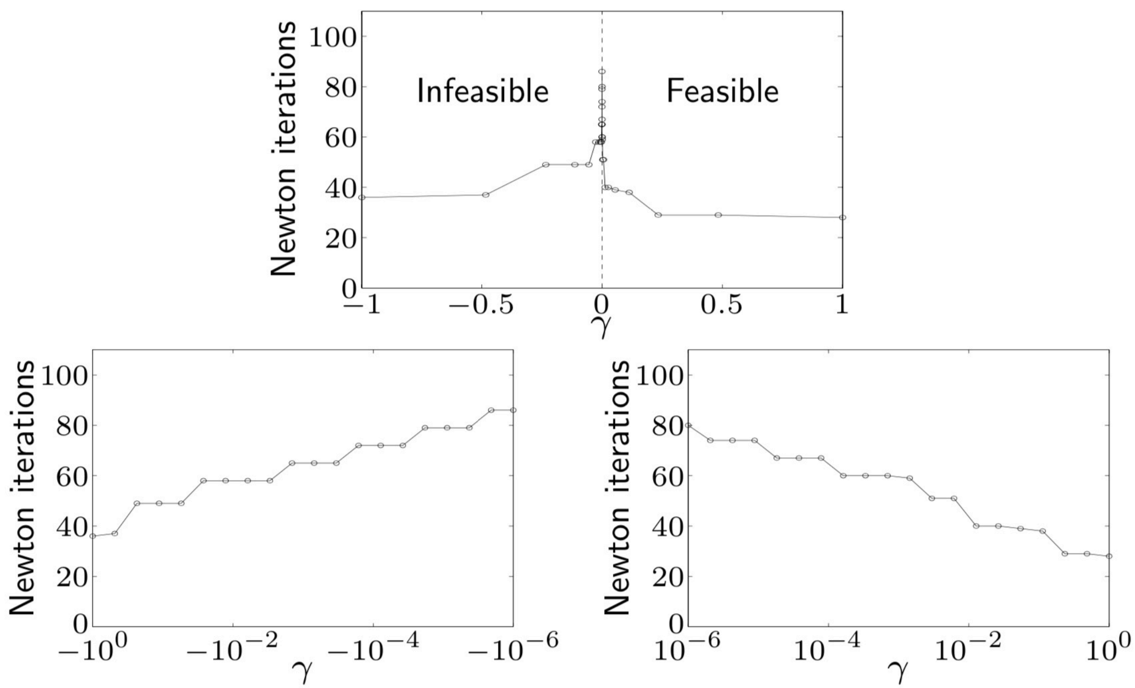
$$Ax \preceq b + \gamma \Delta b$$

- 所需的数据对 $\gamma > 0$ 是严格可行的，对 $\gamma \leq 0$ 是不可行的
- 使用基本阶段**1**方法，当 $s < 0$ 或对偶目标为正时，结束

例：一族线性不等式

$$Ax \preceq b + \gamma \Delta b$$

- 所需的数据对 $\gamma > 0$ 是严格可行的，对 $\gamma \leq 0$ 是不可行的
- 使用基本阶段**1**方法，当 $s < 0$ 或对偶目标为正时，结束



$$\log(1/|\gamma|)$$

基于自和谐条件的复杂性分析

基于自和谐条件的复杂性分析

- 假设：

基于自和谐条件的复杂性分析

- 假设:
- 水平子集有界

基于自和谐条件的复杂性分析

- 假设：
- 水平子集有界
- 所有的 $tf_0+\phi$ 为自和谐函数，且是闭的

基于自和谐条件的复杂性分析

- 假设:
- 水平子集有界
- 所有的 $tf_0 + \phi$ 为自和谐函数，且是闭的
  - 对线性规划、二次规划等很多问题均成立

基于自和谐条件的复杂性分析

- 假设:
- 水平子集有界
- 所有的 $tf_0+\phi$ 为自和谐函数，且是闭的
  - 对线性规划、二次规划等很多问题均成立
  - 或许需要重形式化问题，

基于自和谐条件的复杂性分析

- 假设:
- 水平子集有界
- 所有的 $tf_0 + \phi$ 为自和谐函数，且是闭的
  - 对线性规划、二次规划等很多问题均成立
  - 或许需要重形式化问题,

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

基于自和谐条件的复杂性分析

- 假设：
- 水平子集有界
- 所有的 $tf_0 + \phi$ 为自和谐函数，且是闭的
  - 对线性规划、二次规划等很多问题均成立
  - 或许需要重形式化问题，

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

- 可用于复杂性分析；当不采用自和谐条件时，障碍法仍然有效

中心点步骤的Newton迭代次数

中心点步骤的Newton迭代次数

- 根据自和谐理论

中心点步骤的Newton迭代次数

- 根据自和谐理论

$$\#\text{Newton iterations} \le \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

中心点步骤的Newton迭代次数

- 根据自和谐理论

$$\#\text{Newton iterations} \leq \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

- 以 $x = x^\star(t)$ 为起始点计算 $x^+ = x^\star(\mu t)$ 的代价上界

中心点步骤的Newton迭代次数

- 根据自和谐理论

$$\#\text{Newton iterations} \le \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

- 以 $x = x^\star(t)$ 为起始点计算 $x^+ = x^\star(\mu t)$ 的代价上界
- γ, c 为常数（取决于很多**Newton**算法参数）

中心点步骤的Newton迭代次数

- 根据自和谐理论

$$\#\text{Newton iterations} \le \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

- 以 $x = x^\star(t)$ 为起始点计算 $x^+ = x^\star(\mu t)$ 的代价上界
- γ, c 为常数（取决于很多**Newton**算法参数）
- 根据对偶性（其中，$\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$）

中心点步骤的Newton迭代次数

- 根据自和谐理论

$$\#\text{Newton iterations} \le \frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c$$

- 以 $x = x^\star(t)$ 为起始点计算 $x^+ = x^\star(\mu t)$ 的代价上界
- γ, c 为常数（取决于很多**Newton**算法参数）
- 根据对偶性（其中，$\lambda = \lambda^\star(t)$, $\nu = \nu^\star(t)$）

$$\begin{aligned}
& \mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+) \\
&= \mu t f_0(x) - \mu t f_0(x^+) + \sum_{i=1}^m \log(-\mu t \lambda_i f_i(x^+)) - m \log \mu \\
&\le \mu t f_0(x) - \mu t f_0(x^+) - \mu t \sum_{i=1}^m \lambda_i f_i(x^+) - m - m \log \mu \\
&\le \mu t f_0(x) - \mu t g(\lambda, \nu) - m - m \log \mu \\
&= m(\mu - 1 - \log \mu)
\end{aligned}$$

总的Newton迭代次数

总的Newton迭代次数

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

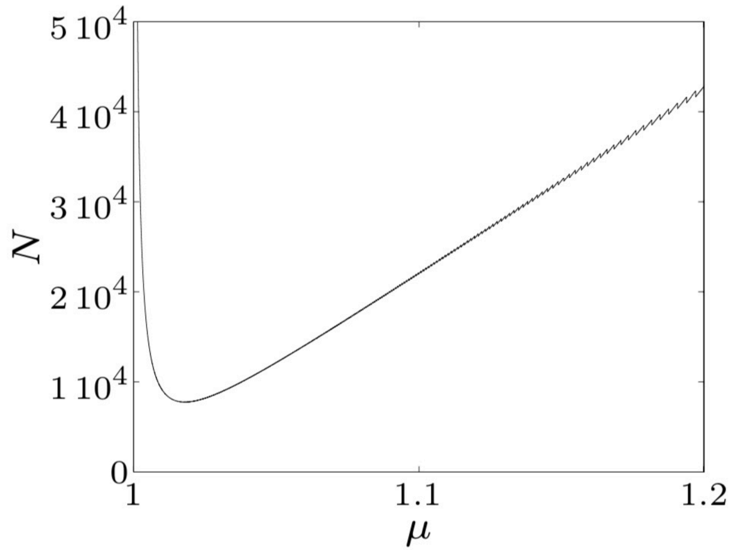


figure shows N for typical values of γ, c,

$$m = 100, \qquad \frac{m}{t^{(0)}\epsilon} = 10^5$$

总的Newton迭代次数

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

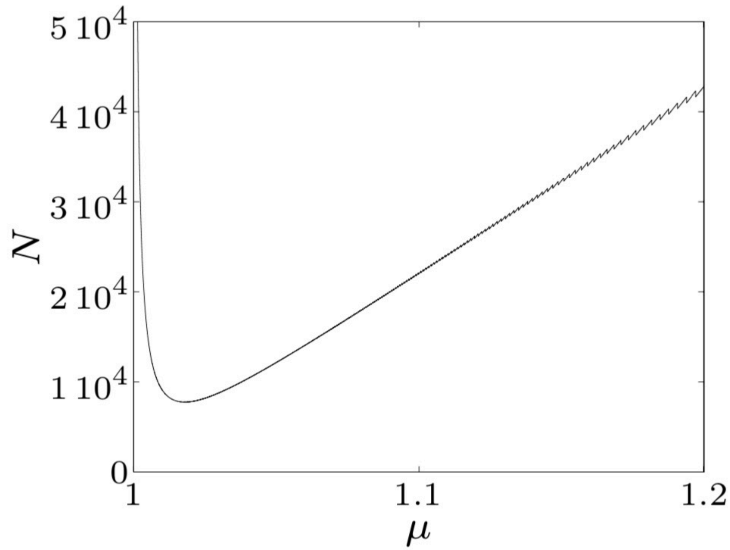


figure shows N for typical values of γ, c,

$$m = 100, \qquad \frac{m}{t^{(0)}\epsilon} = 10^5$$

- 确认关于参数μ 的权衡

总的Newton迭代次数

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

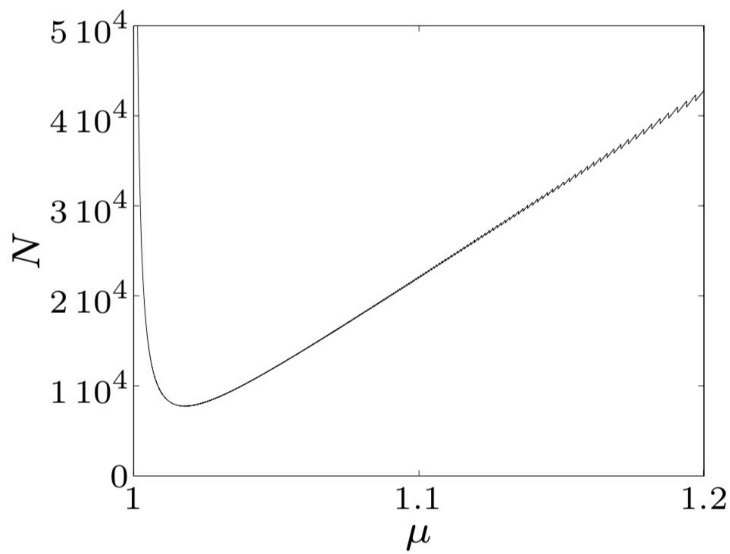


figure shows N for typical values of γ, c,

$$m = 100, \qquad \frac{m}{t^{(0)}\epsilon} = 10^5$$

- 确认关于参数μ 的权衡
- 实际上，迭代参数数十次的量, 当 $\mu \ge 10$ 不敏感

障碍法的多项式复杂度

障碍法的多项式复杂度

- 根据 $\mu = 1 + 1/\sqrt{m}$:

障碍法的多项式复杂度

- 根据 $\mu = 1 + 1/\sqrt{m}$:

$$N = O\left(\sqrt{m}\log\left(\frac{m/t^{(0)}}{\epsilon}\right)\right)$$

障碍法的多项式复杂度

- 根据 $\mu = 1 + 1/\sqrt{m}$:

$$N = O\left(\sqrt{m}\log\left(\frac{m/t^{(0)}}{\epsilon}\right)\right)$$

- 固定间隙压缩的**Newton**法的迭代次数为 $O(\sqrt{m})$

障碍法的多项式复杂度

- 根据 $\mu = 1 + 1/\sqrt{m}$:

$$N = O\left(\sqrt{m}\log\left(\frac{m/t^{(0)}}{\epsilon}\right)\right)$$

- 固定间隙压缩的**Newton**法的迭代次数为 $O(\sqrt{m})$
- 乘以一次**Newton**迭代的代价，得到浮点运算次数的上界

障碍法的多项式复杂度

- 根据 $\mu = 1 + 1/\sqrt{m}$:

$$N = O\left(\sqrt{m}\log\left(\frac{m/t^{(0)}}{\epsilon}\right)\right)$$

- 固定间隙压缩的**Newton**法的迭代次数为 $O(\sqrt{m})$
- 乘以一次**Newton**迭代的代价，得到浮点运算次数的上界
- μ 的选择优化了最坏情况的复杂性；事实上，设置 μ 为固定值

原-对偶内点法

原-对偶内点法

- 当需要高性能时，比障碍法更为有效

原-对偶内点法

- 当需要高性能时，比障碍法更为有效
- 每次迭代更新原问题和对偶问题的变量；不区分内部和外部迭代

原-对偶内点法

- 当需要高性能时，比障碍法更为有效
- 每次迭代更新原问题和对偶问题的变量；不区分内部和外部迭代
- 通常展现出超线性收敛性质

原-对偶内点法

- 当需要高性能时，比障碍法更为有效
- 每次迭代更新原问题和对偶问题的变量；不区分内部和外部迭代
- 通常展现出超线性收敛性质
- 搜索方向可以理解为**Newton**方向，来处理修改的**KKT**条件

原-对偶内点法

- 当需要高性能时，比障碍法更为有效
- 每次迭代更新原问题和对偶问题的变量；不区分内部和外部迭代
- 通常展现出超线性收敛性质
- 搜索方向可以理解为**Newton**方向，来处理修改的**KKT**条件
- 可以从非可行点开始

原-对偶内点法

- 当需要高性能时，比障碍法更为有效
- 每次迭代更新原问题和对偶问题的变量；不区分内部和外部迭代
- 通常展现出超线性收敛性质
- 搜索方向可以理解为**Newton**方向，来处理修改的**KKT**条件
- 可以从非可行点开始
- 每次迭代的计算代价和障碍法相同