第 7 章：不等式约束

(2)：原对偶内点法

曹语

课程主页：`https://yucaoyc.github.io/math3806`

- **问题：**

$$
\begin{array}{ll}
\text{minimize} & f_0(\boldsymbol{x}) \\
\text{subject to} & f_i(\boldsymbol{x}) \leqslant 0, \quad i = 1, \cdots, m \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{array}
$$

- 障碍方法的思路是考虑一个原优化问题的近似问题

$$
\begin{array}{ll}
\text{minimize} & t f_0(\boldsymbol{x}) + \phi(\boldsymbol{x}) \\
\text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{array} \tag{1}
$$

函数 $\phi(\boldsymbol{x}) = -\sum_{i=1}^{m} \log\left(-f_i(\boldsymbol{x})\right)$ 是对数障碍函数。

该问题的 KKT 条件是:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}^\star(t) = \boldsymbol{b}, \qquad\qquad f_i\left(\boldsymbol{x}^\star(t)\right) < 0, \quad i = 1, \cdots, m, \\ t\nabla f_0\left(\boldsymbol{x}^\star(t)\right) + \sum_{i=1}^{m} \frac{1}{-f_i\left(\boldsymbol{x}^\star(t)\right)} \nabla f_i\left(\boldsymbol{x}^\star(t)\right) + \boldsymbol{A}^\top \widehat{\boldsymbol{\nu}} = \boldsymbol{0} \end{aligned} \tag{2}$$

通过线性化该方程组，在前一部分我们学习了牛顿法步径如下:

$$\begin{bmatrix} t\nabla^2 f_0(\boldsymbol{x}) + \nabla^2 \phi(\boldsymbol{x}) & \boldsymbol{A}^\top \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{x}_{\mathrm{nt}} \\ \boldsymbol{\omega} \end{bmatrix} = \begin{bmatrix} -t\nabla f_0(\boldsymbol{x}) - \nabla \phi(\boldsymbol{x}) \\ \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \end{bmatrix}$$

我们可以把公式(2)的 KKT 条件，修改成如下的等价问题

$$
\begin{aligned}
\nabla f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i \nabla f_i(\boldsymbol{x}) + \boldsymbol{A}^\top \boldsymbol{\nu} &= \mathbf{0} \\
-\lambda_i f_i(\boldsymbol{x}) &= \frac{1}{t}, \quad i = 1, \cdots, m \\
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\
\lambda_i &> 0, \quad i = 1, \cdots, m
\end{aligned}
\tag{3}
$$

我们可以对该新形式推导出对应的牛顿法，即原对偶内点法。

对此，我们将 $\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$，并且将方程组写成

$$\boldsymbol{r}_t(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{bmatrix} \nabla f_0(\boldsymbol{x}) + D\boldsymbol{f}(\boldsymbol{x})^\top \boldsymbol{\lambda} + \boldsymbol{A}^\top \boldsymbol{\nu} \\ -\operatorname{diag}(\boldsymbol{\lambda})\boldsymbol{f}(\boldsymbol{x}) - \frac{1}{t}\mathbf{1} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \end{bmatrix}$$

其中

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{bmatrix}, \qquad D\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \nabla f_1(\boldsymbol{x})^\top \\ \vdots \\ \nabla f_m(\boldsymbol{x})^\top \end{bmatrix}$$

类似之前的思路，将该方程组线性化，因此得到牛顿步径如下：

$$
\begin{bmatrix} \nabla^2 f_0(\boldsymbol{x}) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(\boldsymbol{x}) & D\boldsymbol{f}(\boldsymbol{x})^T & \boldsymbol{A}^\top \\ -\operatorname{diag}(\boldsymbol{\lambda}) D\boldsymbol{f}(\boldsymbol{x}) & -\operatorname{diag}(\boldsymbol{f}(\boldsymbol{x})) & 0 \\ \boldsymbol{A} & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{x} \\ \Delta \boldsymbol{\lambda} \\ \Delta \boldsymbol{\nu} \end{bmatrix} = - \begin{bmatrix} r_{\text{dual}} \\ r_{\text{cent}} \\ r_{\text{pri}} \end{bmatrix}.
$$

其中

$$
\begin{aligned}
r_{\text{dual}} &= \nabla f_0(\boldsymbol{x}) + D\boldsymbol{f}(\boldsymbol{x})^\top \boldsymbol{\lambda} + \boldsymbol{A}^\top \boldsymbol{\nu} && \text{对偶残差} \\
r_{\text{cent}} &= -\text{diag}(\boldsymbol{\lambda}) \boldsymbol{f}(\boldsymbol{x}) - \frac{1}{t}\mathbf{1} && \text{中心残差} \\
r_{\text{pri}} &= \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} && \text{原残差}
\end{aligned}
$$

Algorithm 1: 原对偶内点法

Input: 给定严格可行的 $\boldsymbol{x}$，$\boldsymbol{\lambda} \succ 0$，$\boldsymbol{\nu}$，阈值 ϵ，ϵ_{feas}，以及参数 $\mu > 1$

while *True* **do**
  令 $t = \mu m / \hat{\eta}$ $(\hat{\eta} := -\boldsymbol{f}(\boldsymbol{x})^\top \boldsymbol{\lambda})$；
  计算原对偶搜索方向 $\Delta \boldsymbol{y}$；
  线搜索并更新 $\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$；
  停止准则：如果 $\|r_{\text{pri}}\| \leq \epsilon_{\text{feas}}$，$\|r_{\text{dual}}\| \leq \epsilon_{\text{feas}}$，$\hat{\eta} \leq \epsilon$，则退出；
end

- μ 一般取 10 的数量级即可；
- 线搜索时需要验证 $\boldsymbol{x}$ 严格可行，$\boldsymbol{\lambda} \succ 0$，并且验证 $\|\boldsymbol{r}_t(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})\|^2$ 具有足够的下降。

总结

主要需要掌握的知识：

- 了解原对偶内点法形式的来源和思路

阅读作业 & 参考资料：

- 课本第 11.7 章