3. 二次规划

朱天宇

1 二次规划

二次规划是指等式、不等式约束都是二次的，所求的目标函数也是二次的情景。

$$
\begin{aligned}
\min \quad & Q(\mathsf{x}) = \frac{1}{2}\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x} \\
s.t. \quad & \mathsf{a}_i^\top \mathsf{x} = b_i,\ i \in \mathcal{E} = \{1, 2, ...m_e\} \\
& \mathsf{a}_i^\top \geq b_i,\ i \in \mathcal{I} = \{m_{e+1}, ...m\}
\end{aligned}
$$

这里要求 G 是对称阵、$\mathsf{a}_i (i \in \mathcal{E})$ 是线性无关的。

1.1 分类

1. 如果约束是不相容的，则可能没有有限的最小值，此时二次规划无解；
2. 如果矩阵 G 是半正定的，则问题是一个凸的二次规划问题，其任意局部解也是全局解；
3. 如果矩阵 G 是正定的，则问题是一个正定二次规划问题，只要存在解即是唯一解。
4. 如果矩阵 G 不定，则问题是一般的二次规划问题，有可能有非整体的局部解。

正定矩阵含义 如果一个 n 阶方阵 M，对于任意非零向量 z，都有 $\mathsf{z}^\top \mathsf{M}\mathsf{z} > 0$，则称 G 为正定矩阵。正定矩阵所有特征值都为正数 。

半正定矩阵含义 如果 $\mathsf{z}^\top \mathsf{M}\mathsf{z} \geq 0$，则称 G 为半正定矩阵。半正定矩阵所有特征值 非负 。

熟悉上方较为抽象的写法 设 $\mathsf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$，$\mathsf{c} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

则根据 G 的正定、半正定、不定的情况，分别举一个展开的例子。

$\mathsf{G} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ 是正定矩阵，因为 $\mathsf{x}^\top \mathsf{G}\mathsf{x} = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 > 0$。此时原函数 $\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x}$ 展开得到

$$Q = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 + x_1 + 2x_2 + 3x_3$$

对 x_1, x_2, x_3 分别求偏导，就可以解出唯一的值，因此只要存在解就是唯一解。

$\mathsf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 是半正定矩阵，此时原函数 $\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x}$ 展开得到

$$Q = x_1^2 + 2x_2^2 + x_1 + 2x_2 + 3x_3$$

对 x_1, x_2, x_3 分别求偏导，其中 x_1, x_2 是有唯一解的，而 x_3 无解。原函数没有全局最小值 ($x_3 \leftarrow -\infty$)

当矩阵不定的时候，整体解 (无约束的情况下) 不一定存在。

2 仅含等式约束的二次规划

$$
\begin{aligned}
\min \quad & Q(\mathsf{x}) = \frac{1}{2}\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x} \\
s.t. \quad & \mathsf{A}\mathsf{x} = \mathsf{b}
\end{aligned}
$$

在这里，A 是一个 $m \times n, m < n$ 的矩阵。不失一般性，设 $rank(\mathsf{A}) = m$，此时 A 是行满秩的，因此可行域非空。

2.1 基于消去法的两种方法

2.2 Lagrange 方法

拉格朗日方法基于求解可行域内的 (K-T) 点，即拉格朗日函数的稳定点。拉格朗日函数的稳定点是如下线性方程组的解：

$$\begin{cases} \mathsf{G}\mathsf{x} + \mathsf{c} = \mathsf{A}^\top \lambda \\ \mathsf{A}\mathsf{x} = \mathsf{b} \end{cases}$$

写成矩阵形式：

$$\begin{pmatrix} \mathsf{G} & -\mathsf{A}^\top \\ -\mathsf{A} & 0 \end{pmatrix} \begin{pmatrix} \mathsf{x} \\ \lambda \end{pmatrix} = - \begin{pmatrix} \mathsf{c} \\ \mathsf{b} \end{pmatrix}$$

这是一个更高维的线性方程组，可以强行求解但是没有必要。基于左侧矩阵的特殊结构，可以通过矩阵分解的方式来获得解。

设 $\begin{pmatrix} \mathsf{G} & -\mathsf{A}^\top \\ -\mathsf{A} & 0 \end{pmatrix}$ 的逆矩阵为 $\begin{pmatrix} \mathsf{U} & \mathsf{W}^\top \\ \mathsf{W} & \mathsf{V} \end{pmatrix}$

从而可以获得问题的唯一解为

$$\begin{cases} \mathsf{x}^\star = -\mathsf{U}\mathsf{c} - \mathsf{W}^\top \mathsf{b} \\ \lambda^\star = -\mathsf{W}\mathsf{c} - \mathsf{V}\mathsf{b} \end{cases}$$

从而可以计算出 $\mathsf{U}, \mathsf{V}, \mathsf{W}$ 的表示，如图。但这样的表示需要计算 G^{-1}，这个计算代价和直接计算增广的线性方程组没啥区别。因此要用 矩阵分解 的方法。

当G可逆，A行满秩，则$(AG^{-1}A^T)^{-1}$存在，不难验证

$$\begin{cases} U = G^{-1} - G^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1}, \\ V = -(AG^{-1}A^T)^{-1}, \\ W = -(AG^{-1}A^T)^{-1}AG^{-1}. \end{cases}$$

于是我们得到求解公式

$$\begin{cases} \mathbf{x}^* = -G^{-1}\mathbf{c} + G^{-1}A^T(AG^{-1}A^T)^{-1}(AG^{-1}\mathbf{c} + \mathbf{b}), \\ \lambda^* = (AG^{-1}A^T)^{-1}(AG^{-1}\mathbf{c} + \mathbf{b}). \end{cases}$$

Figure 1: 可以，但没必要。

取 Y, Z 满足 $(\mathsf{Y}, \mathsf{Z}) = \begin{pmatrix} \mathsf{A} \\ \mathsf{B} \end{pmatrix}^{-1}$，其中 Y 是矩阵 商空间 的一组基构成的矩阵；Z 是解空间的基构成的矩阵；即 $\mathsf{A}\mathsf{Y} = \mathsf{I}_{m\times m}, \mathsf{A}\mathsf{Z} = 0$。这里 A 是 $m \times n$ 的，Y 是 $n \times m$ 的。此时有

$$\begin{cases} \mathsf{U} = \mathsf{Z}(\mathsf{Z}^\top \mathsf{G}\mathsf{Z})^{-1}\mathsf{Z}^\top \\ \mathsf{V} = -\mathsf{Y}^\top \mathsf{G}\mathsf{P}^\top \mathsf{Y} \\ \mathsf{W} = -\mathsf{Y}^\top \mathsf{P} \end{cases}$$

这里 $\mathsf{P} = \mathsf{I} - \mathsf{G}\mathsf{Z}(\mathsf{Z}^\top \mathsf{G}\mathsf{Z})^{-1}\mathsf{Z}^\top$。

这里 B 取值我们不关心。所以就需要 根据矩阵 $\mathsf{A}_{m\times n}$，确定 $\mathsf{Y}_{n\times m}, \mathsf{Z}$ 使得 $\mathsf{A}\mathsf{Y} = \mathsf{I}$ 以及 $\mathsf{A}\mathsf{Z} = 0$ 。

通过对 A 做 QR 分解 可以获得一组特殊的 Y, Z 取值。

Q 是一个 $n \times n$ 的正交阵；R 是一个 $m \times m$ 的上三角阵；

设置 $\mathsf{A}^\top = \mathsf{Q} \begin{pmatrix} \mathsf{R} \\ 0 \end{pmatrix} = (\mathsf{Q}_1, \mathsf{Q}_2) \begin{pmatrix} \mathsf{R} \\ 0 \end{pmatrix}$，从而有

$$\mathsf{A} = (\mathsf{R}^\top, 0) \begin{pmatrix} \mathsf{Q}_1^\top \\ \mathsf{Q}_2^\top \end{pmatrix}$$

维度分别为 $\mathsf{Q}_{n\times n}, \mathsf{R}_{m\times m}$，维度不足则补 0。

可以令 $\mathsf{Y} = \mathsf{Q}_1 \mathsf{R}^{-\top}, \mathsf{Z} = \mathsf{Q}_2$。

从而求出 $\mathsf{U}, \mathsf{V}, \mathsf{W}$，进而求出 x。

3 包含不等式约束的二次规划

$$
\begin{aligned}
\min \quad & Q(\mathsf{x}) = \frac{1}{2}\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x} \\
s.t. \quad & \mathsf{a}_i^\top \mathsf{x} = b_i,\ i \in \mathcal{E} = \{1, 2, ...m_e\} \\
& \mathsf{a}_i^\top \geq b_i,\ i \in \mathcal{I} = \{m_{e+1}, ...m\}
\end{aligned}
$$

active set 根据函数的连续性，直观上，不积极的不等式约束在解的附近不起作用，可以去掉不予考虑；而积极的不等式约束，由于在解处等号成立，故而可以用等式约束来替代不等式约束。从而转化成等式约束。

需要设法判断积极约束和不积极约束。

3.1 积极集基本定理

定义记号 $\mathcal{I}(\mathsf{x}^k) = \{i \in \mathcal{I} | \mathsf{a}_i^\top \mathsf{x}^k = b_i\}$。这个记号是依赖于 x^k 的，表示在不等式约束构成的集合 $\mathcal{I}$ 中，在 x^k 点处，使得约束 $\mathsf{a}_i^\top \mathsf{x}^k = b_i$ 等号成立的这些约束。称这些约束在 x^k 点处是积极的。

定理 设 $\mathsf{x}^\star$ 是一般的二次规划问题 (包含所有等式和不等式约束) 的局部极小点。则 $\mathsf{x}^\star$ 也必然是等式约束问题

$$
(EQ) \begin{cases} \min \quad Q(\mathsf{x}) = \frac{1}{2}\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x} \\ s.t. \quad \mathsf{a}_i^\top \mathsf{x} = b_i,\ i \in \mathcal{E} \cup \mathcal{I}(\mathsf{x}^\star) \end{cases}
$$

的局部极小点。这个公式中 i 表示第 i 个不等式约束。反之，如果 $\mathsf{x}^\star$ 是一般问题的可行点，同时是问题 (EQ) 的 K-T 点，且相应的拉格朗日乘子满足 $\lambda_i^\star \geq 0, i \in \mathcal{I}^\star$，那么 $\mathsf{x}^\star$ 必然是原问题的 K-T 点。

也就是说问题 (EQ) 的 K-T 点同时还是原问题的可行点，且不等式约束的拉格朗日乘子都大于等于 0，那么 $\mathsf{x}^\star$ 就是原问题的 K-T 点。

因此，记 $\mathcal{E}_k = \mathcal{E} \cup \mathcal{I}(\mathsf{x}^k)$，尝试求解这个含等式约束的二次规划问题 $EQ1$：

$$
(EQ1) \begin{cases} \min \quad \frac{1}{2}\mathsf{s}^\top \mathsf{G}\mathsf{s} + (\mathsf{G}\mathsf{x}^k + \mathsf{c})^\top \mathsf{s} \\ s.t. \quad \mathsf{a}_i^\top \mathsf{s} = 0,\ i \in \mathcal{E}_k \end{cases}
$$

也就是使用**当前迭代点**x^k 对应的积极的不等式约束来替代局部极小点 $\mathsf{x}^\star$ 对应的积极的不等式约束。

在这里，$\mathsf{s} = \mathsf{x} - \mathsf{x}^k$。结果的 s 分为 3 种情况：

1. $\mathsf{s}^k \neq 0$，此时 x^k 不是原问题的 K-T 点；
2. $\mathsf{s}^k = 0$，则 x^k 是问题

$$
(EQ2) \begin{cases} \min \quad \frac{1}{2}\mathsf{x}^\top \mathsf{G}\mathsf{x} + \mathsf{c}^\top \mathsf{x} \\ s.t. \quad \mathsf{a}_i^\top \mathsf{x} = b_i,\ i \in \mathcal{E}_k \end{cases}
$$

   的 K-T 点。如果 $\lambda_i^k \geq 0, i \in \mathcal{I}(\mathsf{x}^k)$，则 x^k 也是原问题的 K-T 点。
3. 否则，选择不等式条件 i_q 使得 $\lambda_{i_q}^k = \min_{i \in \mathcal{I}(\mathsf{x}^k)} \lambda_i^k < 0$，此时，解要满足的约束太多，需要释放掉一些。求如下问题

$$
(EQ3) \begin{cases} \min \quad \frac{1}{2}\mathsf{s}^\top \mathsf{G}\mathsf{s} + (\mathsf{G}\mathsf{x}^k + \mathsf{c})^\top \mathsf{s} \\ s.t. \quad \mathsf{a}_i^\top \mathsf{s} = 0,\ i \in \mathcal{E}_k \backslash i_q \end{cases}
$$

   的解 $\hat{\mathsf{s}}$，作为当前点 x^k 处的可行方向。即 $\mathsf{a}_{i_q}^\top \hat{\mathsf{s}} \geq 0$。

3.2 二次规划的积极集方法

Algorithm 1 积极集方法

1: 初始化：给定可行点 x^0，令 $\mathcal{E}_0 \leftarrow \mathcal{E} \cup \mathcal{I}(\mathsf{x}^0), k \leftarrow 0$
2: **while** true **do**
3: 求解等式约束问题 (EQ1)，获得下降方向 s^k
4: **if** $\mathsf{s}^k = 0$ **then**
5: **if** $\lambda_i^k \geq 0, i \in \mathcal{I}(\mathsf{x}^k)$ **then**
6: 算法停止
7: **else**
8: 选择不等式约束 i_q，满足 $\lambda_{i_q}^k = \min_{i \in \mathcal{I}(\mathsf{x}^k)} \lambda_i^k < 0$
9: $\mathcal{E}_k \leftarrow \mathcal{E}_k \backslash i_q, \mathsf{x}^{k+1} \leftarrow \mathsf{x}^k$
10: **end if**
11: **else**
12: 确定步长 $\alpha_k \leftarrow \min\{1, \min_{i \in \mathcal{E}_k, \mathsf{a}_i^\top \mathsf{s}^k < 0} \frac{b_i - \mathsf{a}_i^\top \mathsf{x}^k}{\mathsf{a}_i^\top \mathsf{s}^k}\}$
13: $\mathsf{x}^{k+1} \leftarrow \mathsf{x}^k + \alpha_k \mathsf{s}^k$
14: **if** $\alpha_k \neq 1$ **then** {从不积极的约束中选择可以添加的积极约束}
15: 寻找 $p \notin \mathcal{E}_k$，满足 $\mathsf{a}_p^\top(\mathsf{x}^k + \alpha_k \mathsf{s}^k) = b_p$
16: $\mathcal{E}_k \leftarrow \mathcal{E}_k \cup \{p\}$
17: **end if**
18: **end if**
19: $\mathcal{E}_{k+1} \leftarrow \mathcal{E}_k, k \leftarrow k+1$
20: **end while**

第 14 行是考虑不在非积极约束，决定步长的时候要考虑这些约束。步长移动可能导致原本的不积极约束变成了积极约束，也有可能使得可行解变成不可行解。通过等式约束决定方向、通过不等式约束决定步长，这是在多个问题中会用到的思想。

3.3 二次规划的基于凸优化的方法

如果满足 G 是半正定的，则可以根据凸优化的相关方法来获取二次规划问题的解。