

求解一类非凸约束的 QCQP 问题

2018 年 11 月 5 日

1 问题形式

在此, 我们关注一类特殊问题: 非凸二次约束的二次规划. 具体形式表示如下

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, m, \end{aligned} \quad (1)$$

其中, 变量 $\mathbf{x} \in \mathbf{R}^n$, 参数 $\mathbf{P}_i \in \mathbf{S}^n, \mathbf{q}_i \in \mathbf{R}^n$ 以及 $r_i \in \mathbf{R}$. 注意, 在这里我们关注 \mathbf{P}_i 不是正半定矩阵 (至少有一个不是) 情况下的优化问题. 如果含有等式约束, 可以转化为不等式约束来求解. 显然, 非凸 QCQP 问题是 NP-难的.

2 半定松弛

我们先考虑简单的问题, 齐次 (homogeneous) 的 QCQP 问题:

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{P}_0 \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{P}_i \mathbf{x} \leq 0, \quad i = 1, \dots, m, \end{aligned} \quad (2)$$

首先, 也是最重要的, 我们观察到

$$\mathbf{x}^T \mathbf{P}_i \mathbf{x} = \text{Tr}(\mathbf{x}^T \mathbf{P}_i \mathbf{x}) = \text{Tr}(\mathbf{P}_i^T \mathbf{x} \mathbf{x}^T) = \langle \mathbf{P}_i, \mathbf{x} \mathbf{x}^T \rangle, \quad i = 0, 1, \dots, m, \quad (3)$$

这里的 \mathbf{P}_i 是对称矩阵. 令 $\mathbf{X} = \mathbf{x} \mathbf{x}^T$, 那么 \mathbf{X} 显然是一个秩为 1 的对称正半定矩阵. 因此原问题可以转化为如下的等价形式:

$$\begin{aligned} \min \quad & \langle \mathbf{P}_0, \mathbf{X} \rangle \\ \text{s.t.} \quad & \langle \mathbf{P}_i, \mathbf{X} \rangle \leq 0, \\ & \mathbf{X} \succeq 0, \text{rank}(\mathbf{X}) = 1. \end{aligned} \quad (4)$$

最后一个约束是根据 $\mathbf{X} = \mathbf{x} \mathbf{x}^T$ 得到的, 这里我们用 $\mathbf{X} \succeq 0$ 表示 \mathbf{X} 是一个正半定矩阵 (PSD).

实际上, 到此为止, 问题并没有变简单. 不过求解这个问题似乎可以用投影梯度法, 因为 $\text{rank}(X) = 1$ 的投影问题比之前的非凸投影问题略为简单.

显然, 问题的难点在于这个秩约束 ($\text{rank}(\mathbf{X}) = 1$) 是非凸的. 我们直接去掉 $\text{rank}(\mathbf{X}) = 1$, 显然原始问题被转变为了半定规划 (SDP) 问题

$$\begin{aligned} \min \quad & \langle \mathbf{P}_0, \mathbf{X} \rangle \\ \text{s.t.} \quad & \langle \mathbf{P}_i, \mathbf{X} \rangle \leq 0, \\ & \mathbf{X} \succeq 0. \end{aligned} \quad (5)$$

该问题被称为半定松弛问题 (SDR), 这是一个凸问题. 一般来说, 这个半定松弛问题的复杂度下界为

$$\mathcal{O}(\max m, n^4 n^{1/2} \log(1/\epsilon)), \quad (6)$$

而 $\epsilon > 0$ 是要求的求解精度. 这里并没有对 $P_i, i = 0, \dots, m$ 做出任何稀疏或者别的特别的假设.

进一步的, 就是如何将问题 (5) 全局最优解 \mathbf{X}^* 转化为原始非凸 QCQP 问题 (1) 的可行解 $\tilde{\mathbf{x}}$. 首先, 如果 \mathbf{X}^* 本身就是秩为 1 的矩阵, 那么显然就有 $\mathbf{X}^* = \mathbf{x}^* \mathbf{x}^{*T}$ 并且 \mathbf{x}^* 即为全局最优解. 其次, 如果 \mathbf{X}^* 的秩大于 1, 那么就变成了秩-1 近似 (rank-one approximation) 问题, 这里我们选择 Frobenius 范数, 有

$$\min \|\mathbf{X}^* - \mathbf{x} \mathbf{x}^T\|_F, \quad (7)$$

将 \mathbf{X}^* 进行奇异值分解可以得到

$$\mathbf{X}^* = \sum_{i=1}^r \lambda_i \mathbf{q}_i \mathbf{q}_i^T, \quad (8)$$

其中 r 是矩阵 \mathbf{X}^* 的秩, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ 是奇异值. 而 $\mathbf{q}_1, \dots, \mathbf{q}_r \in R^n$ 是对应的奇异向量. 则最优解为

$$\lambda_1 \mathbf{q}_1 \mathbf{q}_1^T = \arg \min_{\mathbf{x}} \|\mathbf{X}^* - \mathbf{x} \mathbf{x}^T\|_F, \quad (9)$$

实际上 $\tilde{\mathbf{x}}$ 有可能不满足约束条件, 如果 $\tilde{\mathbf{x}}$ 是可行解, 我们定义 $\tilde{\mathbf{x}} = \sqrt{\lambda_1} \mathbf{q}_1$ 为原始问题的候选解. 否则我们考虑将其映射为约束中最邻近的可行解.

除此以外, 还有一种概率方法值得考虑, 我们在求得半定矩阵 \mathbf{X}^* 之后, 根据正态分布随机生成 L 个向量 $\boldsymbol{\xi}_\ell, \ell = 1, \dots, L$, 即 $\boldsymbol{\xi}_\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{X})$. 然后选择其中的最优解 $\ell^* = \arg \min_{\ell=1, \dots, L} \boldsymbol{\xi}_\ell^T \mathbf{P}_0 \boldsymbol{\xi}_\ell$. 注意需要检查 $\boldsymbol{\xi}_\ell$ 是否满足约束条件.

最终, 我们从 \mathbf{X}^* 中提取出了可行解 $\tilde{\mathbf{x}}$, 值得一提的是, 用 SDR 的方法, 得出的解的质量完全取决于我们的提取方法.

而对于非齐次的形式, 只需将齐次化即可, 考虑如下形式

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{P}_0 \mathbf{x} + 2 \mathbf{q}_0^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{P}_i \mathbf{x} + 2 \mathbf{q}_i^T \mathbf{x} \leq 0, \quad i = 1, \dots, m, \end{aligned} \quad (10)$$

其对应的齐次形式可写为

$$\begin{aligned}
\min \quad & \begin{bmatrix} \mathbf{x}^T & t \end{bmatrix} \begin{bmatrix} \mathbf{P}_0 & \mathbf{q}_0 \\ \mathbf{q}_0^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \\
\text{s.t.} \quad & t^2 = 1 \\
& \begin{bmatrix} \mathbf{x}^T & t \end{bmatrix} \begin{bmatrix} \mathbf{P}_i & \mathbf{q}_i \\ \mathbf{q}_i^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \leq 0, \quad i = 1, \dots, m,
\end{aligned} \tag{11}$$

同样的方法解得最优解 \mathbf{x}^*, t^* , 若 $t^* = 1$, 则原问题最优解为 \mathbf{x}^* ; 若 $t^* = -1$, 则原问题最优解为 $-\mathbf{x}^*$.

实际上, 再回到最开始, 我们将非凸约束 $\mathbf{X} \in S_+^n, \text{rank}(\mathbf{X}) = 1$, 写成 $\mathbf{X} = \mathbf{x}\mathbf{x}^T$, 有

$$\begin{aligned}
\min \quad & \langle \mathbf{P}_0, \mathbf{X} \rangle + 2\langle \mathbf{q}_0, \mathbf{x} \rangle \\
\text{s.t.} \quad & \langle \mathbf{P}_i, \mathbf{X} \rangle + 2\langle \mathbf{q}_i, \mathbf{x} \rangle \leq 0, \quad i = 1, \dots, m, \\
& \mathbf{X} = \mathbf{x}\mathbf{x}^T,
\end{aligned} \tag{12}$$

通过凸松弛, 将 $\mathbf{X} = \mathbf{x}\mathbf{x}^T$ 转化为了 $\mathbf{X} \succeq \mathbf{x}\mathbf{x}^T$, 即

$$\begin{aligned}
\min \quad & \langle \mathbf{P}_0, \mathbf{X} \rangle + 2\langle \mathbf{q}_0, \mathbf{x} \rangle \\
\text{s.t.} \quad & \langle \mathbf{P}_i, \mathbf{X} \rangle + 2\langle \mathbf{q}_i, \mathbf{x} \rangle \leq 0, \quad i = 1, \dots, m, \\
& \mathbf{X} \succeq \mathbf{x}\mathbf{x}^T,
\end{aligned} \tag{13}$$

写成 Schur 补的形式为

$$\begin{aligned}
\min \quad & \langle \mathbf{P}_0, \mathbf{X} \rangle + 2\langle \mathbf{q}_0, \mathbf{x} \rangle \\
\text{s.t.} \quad & \langle \mathbf{P}_i, \mathbf{X} \rangle + 2\langle \mathbf{q}_i, \mathbf{x} \rangle \leq 0, \quad i = 1, \dots, m, \\
& \begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^T & 1 \end{bmatrix} \succeq 0,
\end{aligned} \tag{14}$$

可以看到和前面的直接去除 $\text{rank}(\mathbf{X}) = 1$ 是一样的结果 ($t = 1$).

以上即为原始非凸 QCQP 问题的 SDP 松弛, 其对应的 SDR 的最优值为原始问题最优值的下界. 此外, 已经证明当非凸约束为 1 个时, 能够得到全局最优解. 实际求解过程中, 我们可能会将该问题变成两阶段的方法, 先用 SDR 求得一个候选解. 再用非线性优化的方法 (往往选择内点法) 使得这个解变得更好.

3 拉格朗日松弛

原问题如下

$$\begin{aligned}
\min \quad & \mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \\
\text{s.t.} \quad & \mathbf{x}^T \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^T \mathbf{x} + r_i \leq 0, \quad i = 1, \dots, m,
\end{aligned} \tag{15}$$

应用拉格朗日乘子法

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T \left(\mathbf{P}_0 + \sum_{i=1}^m \lambda_i \mathbf{P}_i \right) \mathbf{x} + \left(\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i \right)^T \mathbf{x} + r_0 + \sum_{i=1}^m \lambda_i r_i, \quad (16)$$

我们先关于 \mathbf{x} 最小化拉格朗日函数, 由于

$$\inf_{\mathbf{x} \in \mathbf{R}^n} \mathbf{x}^T P \mathbf{x} + \mathbf{q}^T \mathbf{x} + r = \begin{cases} r - \frac{1}{4} \mathbf{q}^T P^\dagger \mathbf{q}, & \text{if } P \succeq 0 \text{ and } \mathbf{q} \in \mathcal{R}(P) \\ -\infty, & \text{otherwise} \end{cases} \quad (17)$$

因此其对偶函数为

$$\begin{aligned} g(\lambda) &= \inf_{\mathbf{x} \in \mathbf{R}^n} L(\mathbf{x}, \lambda) \\ &= -\frac{1}{4} \left(\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i \right)^T \left(\mathbf{P}_0 + \sum_{i=1}^m \lambda_i \mathbf{P}_i \right)^\dagger \left(\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i \right) + \sum_{i=1}^m \lambda_i r_i + r_0 \end{aligned} \quad (18)$$

取 γ 为第一项的下界, 有

$$-\gamma - \frac{1}{4} \left(\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i \right)^T \left(\mathbf{P}_0 + \sum_{i=1}^m \lambda_i \mathbf{P}_i \right)^\dagger \left(\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i \right) \geq 0, \quad (19)$$

我们再次使用 Schur 补, 得到其对偶问题

$$\begin{aligned} \min \quad & \gamma + \sum_{i=1}^m \lambda_i r_i + r_0 \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{P}_0 + \sum_{i=1}^m \lambda_i \mathbf{P}_i & (\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i)/2 \\ (\mathbf{q}_0 + \sum_{i=1}^m \lambda_i \mathbf{q}_i)^T/2 & -\gamma \end{bmatrix} \succeq 0 \\ & \lambda_i \geq 0, \quad i = 1, \dots, m, \end{aligned} \quad (20)$$

该方法被称为拉格朗日松弛. 显然这这也是一个半定规划问题, 其最优值也是原问题的下界. 并且, 实际上, 半定松弛和拉格朗日松弛互为对偶问题. 因此两者的下界实际上是一致的.

4 单个约束

非凸问题的弱对偶性表明, 松弛之后的最优解仅仅只是原问题最优值的下界. 但在某些特殊情况下, 即使原问题是非凸的, 这个对偶界限仍然是 0. 那么凸松弛的情况下我们就能够得到原非凸问题的全局最优解.

单约束的 QCQP 问题就是一个典型的例子

$$\begin{aligned} \min \quad & \mathbf{x}^T \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{x}^T \mathbf{P}_1 \mathbf{x} + \mathbf{q}_1^T \mathbf{x} + r_1 \leq 0, \end{aligned} \quad (21)$$

其对偶问题为

$$\begin{aligned} \min \quad & \gamma + \sum_{i=1}^m \lambda_i r_i + r_0 \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{P}_0 + \lambda_1 \mathbf{P}_1 & (\mathbf{q}_0 + \lambda_1 \mathbf{q}_1)/2 \\ (\mathbf{q}_0 + \lambda_1 \mathbf{q}_1)^T/2 & -\gamma \end{bmatrix} \succeq 0 \\ & \lambda \geq 0, \end{aligned} \quad (22)$$

这二者可以产生相同的最优解, 即使第一个问题是非凸问题. 这个结果也是控制理论里面著名的 S-procedure.

5 线性化以及凸逼近

实际上, 松弛技术仅仅给出了最优值的下界, 而对应解很可能并不是可行解. 在此, 我们尝试去寻找一个好的局部最优解. 首先, 我们令 $\mathbf{x}^{(0)}$ 为一个初始可行点.

5.1 线性化

我们保持所有的凸约束不变. 局部线性化原始可行解 $\mathbf{x}^{(0)}$ 点附近的非凸约束. 例如, 考虑如下约束

$$\mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{q}^T \mathbf{x} + r \leq 0, \quad (23)$$

我们将 \mathbf{P} 分解为正定和负定矩阵两部分

$$\mathbf{P} = \mathbf{P}_+ - \mathbf{P}_-, \quad \mathbf{P}_+, \mathbf{P}_- \succeq 0. \quad (24)$$

那么原始约束可以写成

$$\mathbf{x}^T \mathbf{P}_+ \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \leq \mathbf{x}^T \mathbf{P}_- \mathbf{x}, \quad (25)$$

这样等式两边都是凸二次函数, 我们局部线性化 $\mathbf{x}^{(0)}$ 附近的右式, 有

$$\mathbf{x}^T \mathbf{P}_+ \mathbf{x} + \mathbf{q}_0^T \mathbf{x} + r_0 \leq \mathbf{x}^{(0)T} \mathbf{P}_- \mathbf{x}^{(0)} + 2\mathbf{x}^{(0)T} \mathbf{P}_- (\mathbf{x} - \mathbf{x}^{(0)}). \quad (26)$$

不等式右边就变成了一个仿射界. 这个约束是凸的但是比原始的约束更为保守. 换句话说, 我们用一个凸子集代替了原先的非凸集合. 这就是原问题的凸逼近形式.

5.2 迭代法

迭代法实际上即为 **concave-convex procedure** 方法, 这是典型的序列凸逼近方法, 可以得到一个局部最优解 (stationary point).

我们直接考虑 DC 规划问题 (Difference of convex programming), 具体形式如下

$$\begin{aligned} \min \quad & f_0(\mathbf{x}) - g_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) - g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \end{aligned} \quad (27)$$

其中 $\mathbf{x} \in \mathbf{R}^n$ 是优化变量, $f_i: \mathbf{R}^n \rightarrow \mathbf{R}$ 和 $g_i: \mathbf{R}^n \rightarrow \mathbf{R}, i = 1, \dots, m$ 是凸函数.

我们将函数 g_i 线性化. 但与之前方法不同的是, 在这里, 每一步线性化的时候用上一步的点而非最初的可行点.

Algorithm 1 Basic CCP algorithm

given an initial feasible point $\mathbf{x}^{(0)}$.

$k := 0$.

repeat

1. Convexify.

$$\hat{g}_i(\mathbf{x}; \mathbf{x}^{(k)}) \simeq g_i(\mathbf{x}^{(k)}) + \nabla g_i(\mathbf{x}^{(k)})^T (\mathbf{x} - \mathbf{x}^{(k)}), \quad \text{for } i = 1, \dots, m.$$

2. Solve. Set the value of $\mathbf{x}^{(k+1)}$ to a solution of the convex problem

$$\begin{aligned} \min \quad & f_0(\mathbf{x}) - \hat{g}_0(\mathbf{x}; \mathbf{x}^{(k)}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) - \hat{g}_i(\mathbf{x}; \mathbf{x}^{(k)}) \leq 0, \quad i = 1, \dots, m. \end{aligned} \tag{28}$$

3. Update iteration. $k := k + 1$.

until stopping criterion is satisfied.

这里的停止条件建议选择为前后两步目标函数的差值小于给定误差值. 此外, 对于子问题

$$\begin{aligned} \min \quad & f_0(\mathbf{x}) - (g_0(\mathbf{x}^{(k)}) + \nabla g_0(\mathbf{x}^{(k)})^T (\mathbf{x} - \mathbf{x}^{(k)})) \\ \text{s.t.} \quad & f_i(\mathbf{x}) - (g_i(\mathbf{x}^{(k)}) + \nabla g_i(\mathbf{x}^{(k)})^T (\mathbf{x} - \mathbf{x}^{(k)})) \leq 0, \quad i = 1, \dots, m, \end{aligned} \tag{29}$$

也许选择线性搜索方向能够加快子问题的收敛速度

Algorithm 2 Linear search for CCP

given a solution $\mathbf{x}^{(k+1)}$ for subproblem and $\alpha > 1$.

$t := 1$.

$\Delta \mathbf{x} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$.

while The following condition unsatisfied

$$\begin{aligned} f_0(\mathbf{x}^{(k)} + \alpha t \Delta \mathbf{x}) - g_0(\mathbf{x}^{(k)} + \alpha t \Delta \mathbf{x}) &\leq f_0(\mathbf{x}^{(k)} + t \Delta \mathbf{x}) - g_0(\mathbf{x}^{(k)} + t \Delta \mathbf{x}) \\ f_i(\mathbf{x}^{(k)} + \alpha t \Delta \mathbf{x}) - g_i(\mathbf{x}^{(k)} + \alpha t \Delta \mathbf{x}) &\leq 0, \quad i = 1, \dots, m. \end{aligned}$$

do

$t := \alpha t$.

end while
