第三章 典型优化问题

修贤超

https://xianchaoxiu.github.io

目录

半定规划

- 半定规划（SDP）是线性规划在矩阵空间中的一种推广
- 半定规划问题的标准形式

$$\begin{aligned} \min \quad & CX \\ \text{s.t.} \quad & A_1X = b_1 \\ & \cdots \\ & A_mX = b_m \\ & X \succeq 0 \end{aligned}$$

- 对偶形式

$$\begin{aligned} \min \quad & -b^\top y \\ \text{s.t.} \quad & y_1A_1 + y_2A_2 + \cdots + y_nA_n \preceq C \end{aligned}$$

LP, SOCP 与 SDP 的比较

- LP 与 SDP

LP $\quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Ax \le b \end{array}$ $\qquad$ SDP $\quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & \mathrm{diag}(Ax - b) \preceq 0 \end{array}$

- SOCP 与 SDP

SOCP
$$\begin{array}{ll} \min & f^\top x \\ \text{s.t.} & \|A_i x + b_i\|_2 \le c^\top x + d_i, \quad i = 1, \cdots, m \end{array}$$

SDP
$$\begin{array}{ll} \min & f^\top x \\ \text{s.t.} & \left[\begin{array}{cc} (c_i^\top x + d_i) I & A_i x + b_i \\ (A_i x + b_i)^\top & c_i^\top x + d_i \end{array} \right] \succeq 0, \quad i = 1, \cdots, m \end{array}$$

应用举例: 二次约束二次规划问题的半定规划松弛

- 设A_i 为 $n\times n$ 对称矩阵, 考虑二次约束二次规划问题

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & x^\top A_0 x + 2b_0^\top x + c_0 \\
\text{s.t.} \quad & x^\top A_i x + 2b_i^\top x + c_i \le 0, \quad i = 1,2,\cdots,m
\end{aligned}
$$

- 对任意 $x\in\mathbb{R}^n$ 以及 $A\in\mathcal{S}^n$，有恒等式

$$
x^\top A x = \mathrm{Tr}(x^\top A x) = \mathrm{Tr}(A x x^\top) = \langle A, x x^\top\rangle
$$

- 原始问题等价于

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & \langle A_0, X\rangle + 2b_0^\top x + c_0 \\
\text{s.t.} \quad & \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0, \quad i = 1,2,\cdots,m \\
& X = x x^\top
\end{aligned}
$$

应用举例: 二次约束二次规划问题的半定规划松弛

- 进一步地

$$
\begin{aligned}
x^\top A_i x + 2b_i^\top x + c_i &= \left\langle \begin{pmatrix} A_i & b_i \\ b_i^\top & c_i \end{pmatrix}, \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix} \right\rangle \\
&= \langle \overline{A}_i, \overline{X} \rangle, \quad i = 0, 1, \cdots, m
\end{aligned}
$$

- 约束 $X = xx^\top$ 松弛成半正定约束 $X \succeq xx^\top$（等价于 $\overline{X} \succeq 0$）
- 因此这个问题的半定规划松弛可以写成

$$
\begin{aligned}
\min \quad & \langle \overline{A}_0, \overline{X} \rangle \\
\text{s.t.} \quad & \langle \overline{A}_i, \overline{X} \rangle \le 0, \quad i = 1, 2, \cdots, m \\
& \overline{X} \succeq 0 \\
& \overline{X}_{n+1,n+1} = 1
\end{aligned}
$$

应用举例: 最大割问题的半定规划松弛

- 最大割问题是找到节点集合 V 的一个子集 S 使得 S 与它的补集 $\overline{S}$ 之间相连边的权重之和最大化
- 令 $x_j = 1,\ j \in S$ 和 $x_j = -1,\ j \in \overline{S}$，则

$$
\begin{aligned}
\max \quad & \frac{1}{2}\sum_{i<j}(1 - x_i x_j) w_{ij} \\
\text{s.t.} \quad & x_j \in \{-1, 1\},\ j = 1, 2, \cdots, n
\end{aligned}
$$

- 只有当 x_i 与 x_j 不同时，目标函数中 w_{ij} 的系数非零
- 最大割问题是一个离散优化问题，很难在多项式时间内找到最优解

应用举例：最大割问题的半定规划松弛

- 令$W = (w_{ij}) \in \mathcal{S}^n$，并定义 $C = -\frac{1}{4}(\mathrm{Diag}(W\mathbf{1}) - W)$，得到

$$
\begin{aligned}
\min \quad & x^\top C x \\
\text{s.t.} \quad & x_i^2 = 1,\ i = 1, 2, \cdots, n
\end{aligned}
$$

- 令 $X = xx^\top$，则最大割问题可以转化为

$$
\begin{aligned}
\min \quad & \langle C, X \rangle \\
\text{s.t.} \quad & X_{ii} = 1,\ i = 1, 2, \cdots, n \\
& X \succeq 0,\ \mathrm{rank}(X) = 1
\end{aligned}
$$

  - $x_i^2 = 1$ 意味着矩阵 X 对角线元素 $X_{ii} = 1$
  - $X = xx^\top$ 可以用约束 $X \succeq 0$ 和 $\mathrm{rank}(X) = 1$ 等价刻画

应用举例: 极小化最大特征值

- **极小化最大特征值问题可表示为**

$$\min \quad \lambda_{\max}(A_0 + \sum_i x_i A_i)$$

- **由于 $\lambda_{\max}(A) \leq t \Leftrightarrow A \preceq tI$，则极小化最大特征值可以转化为**

SDP 形式

$$\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & zI - \sum_i x_i A_i \succeq A_0 \end{aligned}$$

对偶问题形式

$$\begin{aligned} \max \quad & \langle A_0, Y \rangle \\ \text{s.t.} \quad & \langle A_i, Y \rangle = 0 \\ & \langle I, Y \rangle = 1 \\ & Y \succeq 0 \end{aligned}$$

应用举例: 极小化二范数问题

- 令 $A_i \in \mathbb{R}^{m\times n}$. **极小化** $A(x) = A_0 + \sum_i x_i A_i$ **的二范数**

$$\min_x \quad \|A(x)\|_2$$

- SDP 形式

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

- 约束形式来源于

$$\begin{aligned} \|A\|_2 \le t \Leftrightarrow \quad & A^\top A \preceq t^2 I, \quad t \ge 0 \\ \Leftrightarrow \quad & \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

目录

矩阵优化的基本形式

- 矩阵优化是在近几十年发展起来的一类变量含有矩阵的优化问题，广泛地出现在组合数学、材料科学、机器学习和统计学等
- 矩阵优化问题的形式

$$\min_{X \in \mathcal{X}} \quad \psi(X)$$

  - $\mathcal{X}$ 为特定的矩阵空间
  - $\psi(X) : \mathcal{X} \to \mathbb{R}$ 为给定的函数，可能是非光滑的
- 和向量相比，矩阵有许多新的性质，如秩、特征值等

矩阵优化的基本形式

- 低秩矩阵恢复问题

$$\min_{X \in \mathbb{R}^{m \times n}} \quad \frac{1}{2}\|X - M\|_F^2 + \mu\|X\|_*$$

考虑函数 $h(X) = \|X\|_*$ 的次微分

$$\partial h(X) = \{UV^\top + W \mid \|W\|_2 \le 1,\ U^\top W = 0,\ WV = 0\}$$

- 主成分分析问题

$$\min_{X \in \mathbb{R}^{p \times d}} \quad \psi(X) = -\mathrm{Tr}(X^\top A A^\top) \quad \text{s.t.} \quad X^\top X = I_d$$

考虑目标函数的微分

$$\nabla \psi(X) = -2AA^\mathrm{T} X$$

应用举例：非负矩阵分解

- 给定矩阵 $A = [a_1, a_2, \cdots, a_n] \in \mathbb{R}^{d\times n}$，将其分解成非负基矩阵 $X \in \mathbb{R}^{d\times p}$ 和非负系数矩阵 $Y \in \mathbb{R}^{p\times n}$ 的乘积，即

$$A = XY$$

- 由于观测含有噪声，原始数据矩阵 A 和分解 XY 不会完全吻合, 在这种情况下应考虑

$$\min_{X\in\mathbb{R}^{d\times p},\, Y\in\mathbb{R}^{p\times n}} \quad \frac{1}{2}\|A - XY\|_F^2 \quad \text{s.t.} \quad X \geq 0,\ Y \geq 0$$

- 本质上是将高维空间中的数据在一个低维空间中表示
- 和主成分分析模型类似, 但会得到比主成分分析模型更有实际意义的解

应用举例：非负矩阵分解

- 根据具体应用的不同，还可以考虑带正则项的非负矩阵分解模型

$$\begin{aligned} \min_{X\in\mathbb{R}^{d\times p},\, Y\in\mathbb{R}^{p\times n}} \quad & \frac{1}{2}\|A - XY\|_F^2 + \alpha r_1(X) + \beta r_2(Y) \\ \text{s.t.} \quad & X \geq 0, \quad Y \geq 0 \end{aligned}$$

  - $r_1(X)$ 和 $r_2(Y)$ 是正则项
  - $\alpha, \beta > 0$ 是用来权衡拟合项和正则项的正则化参数
- 如果基向量的线性无关性，取 $r_1(X) = \|X^\top X - I\|_F^2$
- 如果每一个观测值都可以用少数几个基向量来表示，取 $r_2(Y) = \|Y\|_1$

应用举例：相关系数矩阵估计

- 给定对称矩阵 $C \in \mathcal{S}^n$ 和非负对称权重矩阵 $H \in \mathcal{S}^n$，求解一个秩小于等于 p 的相关系数矩阵 X，使得在结合了权重矩阵的某种度量下最小化

$$\begin{aligned} \min_{X \succeq 0} \quad & \frac{1}{2}\|H \odot (X - C)\|_F^2 \\ \text{s.t.} \quad & X_{ii} = 1,\ i = 1, 2, \cdots, n \\ & \operatorname{rank}(X) \le p \end{aligned}$$

- 将 $\operatorname{rank}(X) \le p$ 表示为 $X = V^{\mathrm{T}} V$，其中 $V = [V_1, V_2, \cdots, V_n] \in \mathbb{R}^{p \times n}$，得到

$$\begin{aligned} \min_{V \in \mathbb{R}^{p \times n}} \quad & \frac{1}{2}\|H \odot (V^\top V - C)\|_F^2 \\ \text{s.t.} \quad & \|V_i\|_2 = 1,\ i = 1, 2, \cdots, n \end{aligned}$$

目录

典型优化算法软件

- 主流的商业软件
  - Gurobi（美国） 求解线性规划、二次规划、混合整数线性规划等问题
  - CPLEX（美国） 求解整数规划、线性规划、凸和非凸二次规划等问题
  - MOSEK（丹麦） 求解二次规划、二阶锥规划和半正定规划等问题
  - Knitro（法国）求解大规模非线性规划问题
- 国产软件
  - 杉数 COPT
  - 阿里 MindOPT
  - 华为天筹 OPTV

优化模型语言

- 模型语言的发展开始于 19 世纪 70 年代后期
  - 将容易出错的转化步骤交给计算机完成，降低错误率
  - 在模型和算法之间建立了一个清晰的界限
  - 对于困难的问题，可以尝试不用的求解器，得到更好的结果
- CVX 以 MATLAB 为基础的优化模型语言，求解凸优化问题
  - 快速构造和识别凸性
  - 调用已有软件包求解变形后的凸优化问题
  - 包括免费软件 SDPT3 和 SeDuMi 以及商业软件 Gurobi 和 MOSEK 等

CVX

- 考虑如下优化问题

$$
\begin{aligned}
\min \quad & \|Ax - b\|_2 \\
\text{s.t.} \quad & Cx = d \\
& \|x\|_\infty \le e
\end{aligned}
$$

```
m = 20; n = 10; p = 4;
A = randn(m,n); b = randn(m,1);
C = randn(p,n); d = randn(p,1); e = rand;
cvx_begin
    variable x(n)
    minimize( norm( A * x - b, 2 ) )
    subject to
    C * x == d
    norm( x, Inf ) <= e
cvx_end
```

$Q\&A$

$Thank\ you!$

感谢您的聆听和反馈