2022秋季 学院路校区(凸优化与近似)
最优化方法　　第 5 次作业
提交日期：2022 年11 月24日周四课前

2022 年 11 月 9 日

1. (将点到仿射集的距离建模为凸二次规划问题. 特别地，当所给仿射集是$n-1$维的，即为超平面时，就是中学所学知识“点到直线的距离”和“点到平面距离”公式的推广.) 考虑点 $\boldsymbol{x}^{(0)}$ 到多面集 $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{b}\}$ (其中 $\boldsymbol{A}$ 是列满秩的)的最短距离问题. 它可以表述为二次规划

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^{(0)})^T(\boldsymbol{x}-\boldsymbol{x}^{(0)}) \\ \text{subject to} & \boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

   这个问题是凸规划吗？利用几何最优性条件证明该问题的最优解

$$\boldsymbol{x}^* = \boldsymbol{x}^{(0)} - \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{x}^{(0)} - \boldsymbol{b}).$$

   说明当 $\boldsymbol{A} = \boldsymbol{a} \neq \boldsymbol{0}$ 是一个列向量，$b \in \mathbb{R}$时, 从 $\boldsymbol{x}^{(0)}$ 到超平面 $\boldsymbol{a}^T\boldsymbol{x} = b$ 的最短距离是 $\frac{|\boldsymbol{a}^T\boldsymbol{x}^{(0)}-b|}{\|\boldsymbol{a}\|_2}$.

2. [**最优性条件**] 给出以下问题最优解的充分必要条件.

   (a) 已知$y \in \mathbb{R}^n$. 考虑问题

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & \frac{1}{2}\|y-x\|_2^2 \\ \text{subject to} & x \geq 0, \sum_{i=1}^n x_i = 1. \end{array}$$

   (b) 设$f : \mathbb{R}^n \to \mathbb{R}$是凸的可微函数，已知$A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. 考虑问题

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & Ax \leq b. \end{array}$$

3. [体会写对偶问题时其中集合约束选取的灵活性] 考虑问题

$$
\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\
& \boldsymbol{x} \geq \mathbf{0}.
\end{aligned}
$$

分别基于集合约束 $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \geq \mathbf{0}\}$ 和 $\Omega = \mathbb{R}^n$ 写出该问题的对偶问题.

4. **支撑向量机中相关优化问题的若干分析.**

给定训练数据：

$$(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_l, y_l),$$

其中$\boldsymbol{x}_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$. 假设数据几乎线性可分离，则支撑向量机中求解最优分离超平面的优化问题是

$$
\begin{aligned}
\underset{\boldsymbol{w}\in\mathbb{R}^d, b\in\mathbb{R}}{\text{minimize}} \quad & \tfrac{1}{2}\|\boldsymbol{w}\|_2^2 + C\textstyle\sum_{i=1}^l \xi_i \\
\text{subject to} \quad & y_i(\boldsymbol{w}^T\boldsymbol{x}_i - b) \geq 1 - \xi_i, \quad i = 1, 2, \cdots, l \\
& \xi_i \geq 0, \quad i = 1, 2, \cdots, l,
\end{aligned}
\tag{1}
$$

其中C是折中参数. 设该优化问题的最优解是$\boldsymbol{w}^*, b^*$，则最终的分类超平面是

$$(\boldsymbol{w}^*)^T\boldsymbol{x} = b^*,$$

最终的分类器是

$$g(\boldsymbol{x}) = \operatorname{sgn}((\boldsymbol{w}^*)^T\boldsymbol{x} - b^*).$$

该模型的直观含义是使得两类几乎线性可分数据之间的Margin $= \frac{1}{\|\boldsymbol{w}\|_2}$尽可能地大，从而使分类器具备更强的推广能力，同时使经验误差$\sum_{i=1}^l \xi_i$尽可能地小.

针对约束优化问题(1)进行以下分析：

(a) 写出(1)的Lagrange函数.

(b) 求出(1)的对偶函数.

(c) 写出(1)的Lagrange对偶问题，并指出如何由对偶问题的最优解确定(1)的最优解中的$\boldsymbol{w}$.

(d) 写出优化问题(1)的KKT条件.

(e) 比较(c)和(d)的结果. 并给出所观察到的现象的一个可能的解释.

5. [**凸共轭**] 对任意函数$f : C \subset \mathbb{R}^n \to \mathbb{R}$，定义其凸共轭$f^*$为

$$f^*(y) = \sup_{x\in C}(y^T x - f(x)). \tag{2}$$

f^*的定义域取为集合$D = \{y \in \mathbb{R}^d : f^*(y) < \infty\}$.

(1) 求f^*和D，若

(a) $f(x) = 1/x, C = (0, \infty)$,

(b) $f(x) = \frac{1}{2}\|x\|_2^2, C = \mathbb{R}^n$,

(c) $f(x) = \log \sum_{j=1}^n \exp(x_j), x = (x_1, \cdots, x_n), C = \mathbb{R}^n$,

下面假设f为严格凸可微函数，且$C = \mathbb{R}^d$.

(2) 证明$\forall x \in C, f(x) \geq f^{**}(x)$成立.

(3) 证明(2)中的上确界在满足$\nabla f(x^*) = y$的x^* 处取到.

(4) 回想$D_f(\cdot, \cdot)$表示关于f的Bregman散度. 证明

$$D_f(x, y) = D_{f^*}(\nabla f(y), \nabla f(x))$$

6. 考虑问题

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 + x_2 - 1 = 0. \end{array}$$

(a) 这个问题是凸规划吗？

(b) 写这个问题的几何最优性条件，求出最优解和对应的Lagrange乘子.

(c) 求出这个问题的Lagrange对偶问题，并求解.

(d) 设$\eta = 2, \lambda_1 = 0$. 分别给出用对偶上升法和乘子法求解所给问题得到的迭代点λ_2.

(e) 比较(d)中两种方法得到的λ_2与(c)中得到的对偶最优解的远近.