

# 一范数规范化最小二乘问题

## 一、作业要求

1141374791正在观看直播

腾讯课堂

8

01:41:59

作业 一范数规范化最小二乘问题  $x \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, \lambda \in \mathbb{R}$

$\min \frac{1}{2} \|Ax - b\|^2 + \lambda \|x\|_1$

当  $\lambda \geq \lambda_{\max}$  时, 最优解必为  $x=0$ . 计算  $\lambda_{\max}$  的取值

## 二、作业原理

1. 凸问题的重要性质:

腾讯课堂

7

凸问题的重要性质: 局部最优  $\Leftrightarrow$  全局最优

回顾:  $x$  局部最优  $\Leftrightarrow \exists R > 0, f(x) = \inf \{ f(z) \mid z \in X, \|x - z\|_2 \leq R \}$

证明. 设  $x$  不是全局最优 则  $\exists y \in X, f(y) < f(x)$

因  $x$  局部最优, 故  $\|y - x\|_2 > R$

考虑  $z = (1-\theta)x + \theta y$   $\theta = \frac{R}{2\|y-x\|_2} \in (0, 1)$

有  $\|z - x\|_2 = \frac{R\|y-x\|_2}{2\|y-x\|_2} = \frac{R}{2}$

因可行域凸,  $z$  必可行 (凸组合)

因  $x$  局部最优,  $\|z - x\|_2 \leq R$ , 则  $f(z) \geq f(x)$

因  $f_0$  为凸,  $f(z) \leq (1-\theta)f_0(x) + \theta f_0(y) < f_0(x)$ , 矛盾

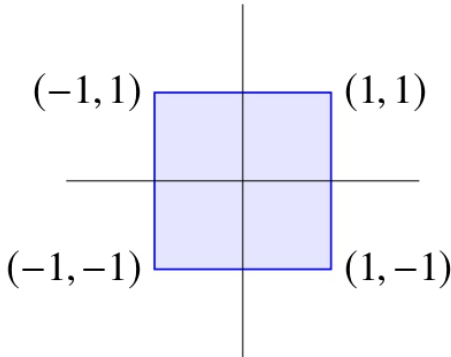
2. 1-范数的次梯度:

## Example: $\ell_1$ -norm

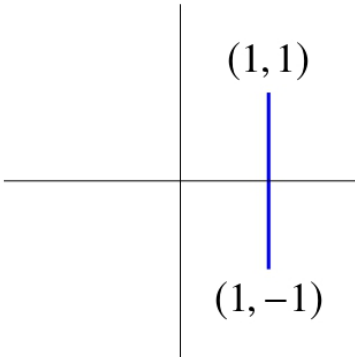
$$f(x) = \|x\|_1 = \max_{s \in \{-1,1\}^n} s^T x$$

the subdifferential is a product of intervals

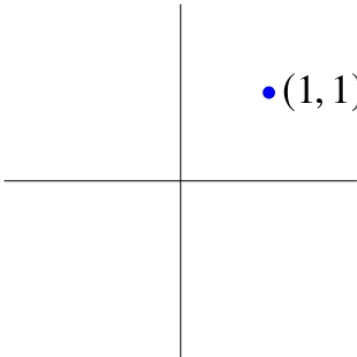
$$\partial f(x) = J_1 \times \cdots \times J_n, \qquad J_k = \begin{cases} [-1, 1] & x_k = 0 \\ \{1\} & x_k > 0 \\ \{-1\} & x_k < 0 \end{cases}$$



$$\partial f(0,0) = [-1, 1] \times [-1, 1]$$



$$\partial f(1,0) = \{1\} \times [-1, 1]$$



$$\partial f(1,1) = \{(1,1)\}$$

Subgradients

2.16

### 三、作业内容

因为1-范数， 2-范数是凸函数

所以 $f$ 是凸函数

所以该问题是无约束的图问题， 局部最优解即全局最优解

所以 $\mathbf{0}$ 是该问题最优解当且仅当 $\exists R > 0, \forall z \in \mathbb{R}^n$ , 若 $\| z - \mathbf{0} \| \leq R$ ,  $f(z) > f(\mathbf{0})$ 即 $f(z) - f(\mathbf{0}) > \mathbf{0} \cdot (z - \mathbf{0})$ 即 $\mathbf{0} \in \partial f(\mathbf{0})$

因为 $\nabla^{\frac{1}{2}} \| Ax - b \|^2 = A^T(Ax - b)$ （分子布局）。

又因为 $\partial \| x \|_1 = J_1 \times \cdots \times J_n$ , 其中 $J_k = \begin{cases} [-1, 1] & x_k = 0 \\ \{1\} & x_k > 0 \\ \{-1\} & x_k < 0 \end{cases}$

所以 $\partial f(\mathbf{0}) = \{A^T(-b) + \lambda C | \forall i \in \mathbb{N}^+ \wedge i \leq n, C_i \in [-1, 1]\}$

所以 $\mathbf{0} \in \partial f(\mathbf{0})$ 当且仅当 $\forall i \in \mathbb{N}^+ \wedge i \leq n, 0 \in [(A^T(-b))_i - \lambda, (A^T(-b))_i + \lambda]$

当且仅当 $\lambda \geq \max_{1 \leq i \leq n} \{|(A^T b)_i|\}$

所以 $\lambda_{max} = \max_{1 \leq i \leq n} \{|(A^T b)_i|\}$