Theory and method of optimization

郭英明

2020/10/1

目录

1 The first homework

1.1

Question: the intersection of convex sets is convex.

Answer:

假设凸集为 $S_1, S_2 \cdots S_n$，若证明 $S_1 \cap S_2$ 为凸集，则 $S_1 \cap S_2 \cdots \cap S_n$ 为凸集。

下证 $S_1 \cap S_2$ 为凸集：

设 $x_1, x_2 \in S_1 \cap S_2$, $x_0 = \theta * x_1 + (1-\theta) * x_2, (0 \leq \theta \leq 1)$，

因为 $x_1 \in S_1$，$x_2 \in S_1$, 且 S_1 为凸集，

所以 $x_0 = \theta * x_1 + (1-\theta) * x_2 \in S_1$，

同理，$x_0 \in S_2$，

所以，, $x_0 \in S_1 \cap S_2$。

所以，对 $\forall x_1, x_2 \in S_1 \cap S_2, x_0 = \theta * x_1 + (1-\theta) * x_2 \in S_1 \cap S_2, (0 \leq \theta \leq 1)$。

所以 $S_1 \cap S_2$ 为凸集，

得证。

1.2

Question: if C is convex,then

$$aC + b = ax + b : x \in C$$

is convex for any a,b

Answer:

对 $\forall x_1, x_2 \in C$，

$ax_1 + b \in aC + b, ax_2 + b \in aC + b$

设 $x_0 = \theta * x_1 + (1-\theta) * x_2, (0 \leq \theta \leq 1)$，

则

$$\begin{aligned} ax_0 + b &= a * (\theta * x_1 + (1-\theta) * x_2) + b \\ &= \theta * (ax_1 + b) + (1-\theta) * (ax_2 + b) \end{aligned}$$

所以对 $\forall ax_1 + b \in aC + b, ax_2 + b \in aC + b$，

$ax_0 + b = \theta * (ax_1 + b) + (1-\theta) * (ax_2 + b) \in aC + b$

所以 $aC + b$ 为凸集。

1.3

Question: if $f(x) = Ax + b$ and C is convex, and if D is convex then

$$f^{-1}(D) = x : f(x) \in D$$

is convex

Answer:

因为 D 为凸集，b 为常数，

所以，$D - b$ 为凸集。

对 $\forall D_1, D_2 \in D - b$，

设

$$\begin{aligned} Ax_1 &= D_1 \\ Ax_2 &= D_2 \\ x_0 &= \theta * x_1 + (1-\theta) * x_2), (0 \leq \theta \leq 1) \end{aligned}$$

所以 $Ax_0 = \theta * D_1 + (1-\theta) * D_2 \in D - b$

所以 $x_0 \in$ 集合 $f^{-1}(D)$

所以集合 $f^{-1}(D)$ 为凸集。

1.4

Question: log-sum-exp:$f(x) = \log \sum_{i=1}^{n} \exp x_k$ is convex

Answer:

$$\nabla f(x) = \begin{bmatrix} \frac{\exp x_1}{\sum_{i=1}^{n} \exp x_i} & \frac{\exp x_2}{\sum_{i=1}^{n} \exp x_i} & \cdots & \frac{\exp x_n}{\sum_{i=1}^{n} \exp x_i} \end{bmatrix} * x$$

二阶导分母相同且为正，省去后：$\nabla^2 f(x) =$

$$\begin{bmatrix} \exp x_1 \sum_{i=1}^{n} \exp x_i - \exp x_1 \exp x_1 & -\exp x_1 \exp x_2 & \cdots & -\exp x_1 \exp x_n \\ -\exp x_1 \exp x_2 & \exp x_2 \sum_{i=1}^{n} \exp x_i - \exp x_2 \exp x_2 & \cdots & -\exp x_2 \exp x_n \\ \vdots & \vdots & \ddots & \vdots \\ -\exp x_1 \exp x_n & \cdots & \cdots & \exp x_n \sum_{i=1}^{n} \exp x_i - \exp x_n \exp x_n \end{bmatrix}$$

$n = 1$ 时矩阵为半正定，

假设 $n = k - 1$ 时矩阵为半正定阵，记为矩阵 A，

则 $n = k$ 时矩阵的 $k - 1$ 阶主子式可写为：

$$A+\begin{bmatrix} \exp x_1 exp x_k & 0 & \cdots & 0 \\ 0 & \exp x_2 \exp x_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \exp x_{k-1} \exp x_k \end{bmatrix}$$

因为半正定具有可加性，
所以 $n=k$ 时的 $k-1$ 阶顺序主子式为半正定矩阵，
所以 $n=k$ 时的前 $k-1$ 阶顺序主子式均大于等于 0，
对 $\nabla^2 f(x)$ 做初等行变换，

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ -\exp x_1 \exp x_2 & \exp x_2 \sum_{i=1}^n \exp x_i - \exp x_2 \exp x_2 & \cdots & -\exp x_2 \exp x_n \\ \vdots & \vdots & \ddots & \vdots \\ -\exp x_1 \exp x_n & \cdots & \cdots & \exp x_n \sum_{i=1}^n \exp x_i - \exp x_n \exp x_n \end{bmatrix}$$

按第一行展开，行列式值为 0，
所以 $n=k$ 时的前 k 阶顺序主子式均大于等于 0，
所以 $\nabla^2 f(x)$ 半正定。
所以 $f(x)$ 为凸函数。

2 The second homework

2.1

Question: 证明无穷范数 $||X||_\infty := max|x_i|$ 满足范数的三个性质。
Answer:
性质一：$||X||_\infty := \max |x_i| \geq 0$，当且仅当 $X=0$ 时等号成立。
性质二：

$$\begin{aligned} ||tX||_\infty &= \sqrt[n]{(tx_1)^n + (tx_2)^n + \cdots + (tx_n)^n} \\ &= t * \sqrt[n]{(x_1)^n + (x_2)^n + \cdots + (x_n)^n} \\ &= t \max_{0 \leq i \leq n} |x_i| \\ &= t||X||_\infty \end{aligned}$$

性质三：

$$
\begin{aligned}
||X+Y||_\infty &= \sqrt[n]{(x_1+y_1)^n+(x_2+y_2)^n+\cdots+(x_n+y_n)^n} \\
&= \max_{0\le i\le n}|x_i+y_i| \\
&\le \max_{0\le i\le n}|x_i|+\max_{0\le i\le n}|y_i| \\
&\le \max_{0\le i\le n}|x_i|+\max_{0\le j\le n}|y_j| \\
&= ||X||_\infty + ||Y||_\infty
\end{aligned}
$$

2.2

Question: 凸函数的局部最优为全局最优

Answer:

设 x_0 为全局最优点，若存在局部最优点 x_1,

若 $f(x_0)=f(x_1)$,

则局部最优为全局最优。

若 $f(x_0)<f(x_1)$, 不妨设 $x_0<x_1$。

因为 x_1 为局部最优，

所以存在 x_2 使 $x_0<x_2<x_1$,$f(x_2)>f(x_1)>f(x_0)$,

设 $x_2=\theta x_0+\theta x_1,(0\le\theta\le1)$，

$f(x_2)>\theta f(x_0)+(1-\theta)(f(x_1)$

与 $f(x)$ 为凸函数矛盾。

所以 $f(x_0)=f(x_1)$，x_1 为全局最优。

3 The third homework

3.1

Question: 无穷范数的对偶范数是一范数

Answer:

$$
\begin{aligned}
||u||_* &= \sup_{||v||_\infty \le 1} \mathbf{u}^{\mathrm{T}} v \\
&= \frac{\sum_{i=1}^n u_i v_i}{\max\limits_{0 \le j \le n} |v_j|} \\
&\le \frac{\sum_{i=1}^n |u_i||v_i|}{\max\limits_{0 \le j \le n} |v_j|} \\
&\le \frac{\sum_{i=1}^n |u_i| \max\limits_{0 \le j \le n} |v_j|}{\max\limits_{0 \le j \le n} |v_j|} \\
&= \sum_{i=1}^n |u_i| \\
&= ||u||_1
\end{aligned}
$$

所以，无穷范数的对偶范数是一范数。

4 The fourth homework

4.1

Question: 求一范数的次梯度

Answer:

$f(x) = ||X||_1 = \sum_{i=1}^n |x_i|$

当 $x_i > 0$ 时，

$\frac{\partial f}{\partial x_i} = \frac{\partial x_i}{\partial x_i} = 1$

当 $x_i < 0$ 时，

$\frac{\partial f}{\partial x_i} = \frac{-\partial x_i}{\partial x_i} = -1$

当 $x_i = 0$ 时，

$\partial f(x) = \left\{ g | \mathbf{g}^{\mathrm{T}}(y - x) \le f(y) - f(x) \right\}$

当 $y_i > x_i$ 时，其它项相等，

$\mathbf{g}_i^{\mathrm{T}} \le \frac{f(y)-f(x)}{y_i - x_i} = \frac{y_i - 0}{y_i - 0} = 1$

同理，当 $y_i < x_i$ 时，

$\mathbf{g}_i^{\mathrm{T}} \ge \frac{f(y)-f(x)}{y_i - x_i} = \frac{-y_i - 0}{y_i - 0} = -1$

所以 $-1 \le \frac{\partial f}{\partial x_i} \le 1$

综上：

$$\frac{\partial f}{\partial x_i} = \begin{cases} 1 & \text{if} \quad x_i > 0 \\ -1 & \text{if} \quad x_i < 0 \\ [-1, 1] & \text{if} \quad x_i = 0 \end{cases}$$