数据科学与工程数学基础 作业7

2021年6月30日 上午

5.1k 字 43 分钟

一

> 下面的集合哪些是凸集？
>
> 1. 平板， 即形如 $\left\{x \in \mathbb{R}^n | \alpha \leq a^T x \leq \beta\right\}$ 的集合。
> 2. 矩形， 即形如 $\{x \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i, i = 1, \cdots, n\}$ 的集合。当 $n > 2$ 使，矩形有时也称为超矩形。
> 3. 楔形，即 $\left\{x \in \mathbb{R}^n | a_1^T x \leq b_1, a_2^T x \leq b_2\right\}$。
> 4. 距离给定点比距离给定集合近的点构成的集合，即
>
> $$\{x|||x - x_0||_2 \leq ||x - y||_2, \forall y \in S\}$$
>
> ，其中 $S \subseteq \mathbb{R}^n$

(a) 该集合可写为 $\{x \in \mathbb{R}^n | a^T x \leq \beta\} \cap \{x \in \mathbb{R}^n | a^T x > \alpha\}$，故为一个凸集

(b) 该集合可写为一组半空间的交集，故为一个凸集

(c) 该集合可写为 $\{x\in\mathbb{R}^n|a_1^Tx\leq b_1\}\cap\{x\in\mathbb{R}^n|a_2^Tx\leq b_2\}$，故为一个凸集

(d) 该集合可写为

$$\bigcap_{y\in S}\{x|||x-x_0||_2\leq||x-y||_2\}$$

故为一个凸集

二

> 下面的函数哪些是凸函数？请说明理由：
>
> 1. $f(x)=e^x+1,x\in\mathbb{R}$
> 2. $f(x)=\max\left(\|Ax+b\|_2,\left\|x^TAx\right\|_1\right),A\in\mathbb{R}^{m\times n}x\in\mathbb{R}^nb\in\mathbb{R}^m$
> 3. $f(x)=-\cos x,x\in[-\pi/2,\pi/2]$

(a)

$$f'(x)=e^x,f''(x)=e^x>0,x\in\mathbb{R}$$

故 $f(x)$ 为凸函数

(b) 由于仿射映射、仿射函数取范数、取最大值为保凸运算，故 $f(x)$ 为凸函数

(c)

$$f'(x)=\sin x,f''(x)=\cos x\geq 0,x\in[-\frac{\pi}{2},\frac{\pi}{2}]$$

故 $f(x)$ 为凸函数

三

> 证明 $x^*=(1,0.5,-1)$ 是如下优化问题的最优解：
>
> $$\begin{array}{ll}\min & \frac{1}{2}x^TPx+q^Tx+r\\ \text{s.t} & -1\leq x_i\leq 1,i=1,2,3\end{array}$$
>
> 其中
>
> $$P=\begin{pmatrix}13&12&-2\\12&17&6\\-2&6&12\end{pmatrix},q=\begin{pmatrix}-22\\-14.5\\13\end{pmatrix},r=1$$

由于 $\nabla f_0=\frac{1}{2}(P+P^T)x+q=px+q$

故 $\nabla f_0(x^*)=(-1,0,2)^T$

因此 $\forall y\in[-1,1]^n$，$\nabla f_0(x^*)^T(y-x)\geq 0$

即 x^* 满足最优性条件，也即目标函数的最优点

四

> 计算函数 $f(x)$ 的共轭函数，以及共轭函数的定义域：
>
> 1. $f(x)=-\log x$
> 2. $f(x)=e^x$

(a) 由 $f(x)=-\log x$ 可知 $dom f=\{x|x>0\}$

令 $g(x,y)=xy+\log x$

当 $y\geq 0$ 时，$\sup g(x,y)=+\infty$

当 $y<0$ 时，$\sup g(x,y)=-1-\log(-y)$ 当且仅当 $x=-\frac{1}{y}$

故 $f^*(y)=-1-\log(-y), y<0$

(b) $dom f=\mathbb{R}$

令 $g(x,y)=xy-e^x$

当 $y>0$ 时，$\sup g(x,y)=y\log y-y$ 当且仅当 $x=\log y$

当 $y=0$ 时，$\sup g(x,y)=\sup(-e^x)=0$

当 $y<0$ 时，$\sup g(x,y)=+\infty$

故

$$f^*(y)=\begin{cases} y\log y-y, y>0 \\ 0, y=0 \end{cases}$$

五

求解线性规划

$$\begin{array}{cr} \min & e^Tx \\ \text{s.t} & Gx\leq h \\ & Ax=b \end{array}$$

的对偶函数，给出对偶问题。

拉格朗日函数

$$\begin{aligned} L(x,\lambda,\mu) &= e^Tx+\lambda^T(Gx-h)+\mu^T(Ax-b) \\ &= (e^T+\lambda G+\mu^TA)x-\lambda^Th-\mu^Tb \end{aligned}$$

故

$$g(\lambda,\mu)=\begin{cases} -\lambda^Th-\mu^Tb, & e+G^T\lambda+A^T\mu=0 \\ -\infty, & otherwise \end{cases}$$

因此其对偶问题为

$$\max_{\lambda,\mu}\left(-\lambda^Th-\mu^Tb\right)$$
$$s.t.\ \ e+G^T\lambda+A^T\mu=0, \lambda\geq 0$$

六

证明：Gauss概率密度函数的累积分布函数

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du$$

是对数-凹函数。即 $\log(\Phi(x))$ 是凹函数。

由于

$$\Phi'(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$
$$\Phi''(x)=-\frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

故

$$
\begin{aligned}
(\Phi'(x))^2 &= \frac{1}{2\pi}e^{-\frac{x^2}{2}} \\
\Phi(x)\Phi''(x) &= -\frac{x}{2\pi}e^{-\frac{x^2}{2}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}du
\end{aligned}
$$

当 $x \geq 0$ 时，易见 $\Phi(x)\Phi''(x) \leq (\Phi'(x))^2$

当 $x < 0$ 时，由 $\frac{u^2}{2}$ 的凸性可知

$$
\begin{aligned}
\Phi(x)\Phi''(x) &= -\frac{x}{2\pi}e^{-\frac{x^2}{2}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}du \\
&\leq -\frac{x}{2\pi}e^{-\frac{x^2}{2}}\int_{-\infty}^{x} e^{-\frac{x^2}{2}-(u-x)x}du \\
&= \frac{1}{2\pi}e^{-\frac{x^2}{2}} \\
&= (\Phi'(x))^2
\end{aligned}
$$

由此可知 $\Phi(x)$ 时对数凸函数

七

> 求优化问题 $\arg\min_{x_1,x_2,x_3} x_1x_2x_3$ 当 x_1,x_2,x_3 满足 $x_1^2+x_2^2+x_3^2=1$ 的解

拉格朗日函数 $L = x_1x_2x_3 + \lambda(x_1^2+x_2^2+x_3^2-1)$

令 $\nabla L = 0$，即

$$
\begin{cases}
\frac{\partial L}{\partial x_1} = x_2x_3 + 2\lambda x_1 = 0 \\
\frac{\partial L}{\partial x_2} = x_1x_3 + 2\lambda x_2 = 0 \\
\frac{\partial L}{\partial x_3} = x_1x_2 + 2\lambda x_3 = 0 \\
\frac{\partial L}{\partial \lambda} = x_1^2+x_2^2+x_3^2-1 = 0
\end{cases}
$$

解得 $|x_1| = |x_2| = |x_3| = \frac{1}{\sqrt{3}}$

故 $|x_1x_2x_3| = \frac{\sqrt{3}}{9}$

代入原方程可知 $x_1x_2x_3 = -\frac{\sqrt{3}}{9}$

八

> 已知矩阵 $A \in \mathbb{R}^{p\times q}, B \in \mathbb{R}^{p\times r}, \mathrm{rank}(A) = \min(p,q)$，未知矩阵 $X \in \mathbb{R}^{q\times r}$，求以下优化问题：
>
> 若 $p < q$，求Frobenius范数最小的矩阵 X，使得 $AX = B$，也即优化问题为
>
> $$
> \begin{aligned}
> \min \qquad & f(X) = \frac{1}{2}||X||_F^2 \\
> \text{s.t} \qquad & AX = B
> \end{aligned}
> $$

拉格朗日函数 $L = Tr(\frac{1}{2}X^TX) - Tr(\Lambda^T(AX-B))$

令 $\nabla L = 0$，即

$$
\begin{cases}
\frac{\partial L}{\partial X} = X - A^T\Lambda = 0 \\
\frac{\partial L}{\partial \Lambda} = AX - B = 0
\end{cases}
$$

由于 A 行满秩，故 AA^T 可逆

故 $AX = B = AA^T(AA^T)^{-1}B$

即 $X=A^T(AA^T)^{-1}B$

九

> 给出优化问题 $\min_x(x^3-ax)$ 使用牛顿法时的迭代格式。

$$f'(x)=3x^2-a,f''(x)=6x$$

故

$$x_n=x_{n-1}-\frac{f'(x_n)}{f''(x_n)}=x_{n-1}-\frac{3x_n^2-a}{6x_n}$$

十

> 梯度下降法是最常用的优化方法之一。考虑优化问题
>
> $$\min f(x)=x_1^2+x_2^2+2x_3^2$$
>
> 证明:在点 $x_0=(x_1,x_2,x_3)$ 处沿负梯度方向迭代的最佳步长为
>
> $$\lambda=\frac{x_1^2+x_2^2+4x_3^2}{2x_1^2+2x_2^2+16x_3^2}$$

令 $x'=x-\lambda\nabla f(x)$

故

$$\begin{aligned}g(\lambda)&=f(x')\\&=f(x-\lambda\nabla f(x))\\&=(1-2\lambda)^2x_1^2+(1-2\lambda)^2x_2^2+2(1-4\lambda)^2x_3^2\end{aligned}$$

于是

$$g'(\lambda)=-4(1-2\lambda)^2x_1^2-4(1-2\lambda)^2x_2^2-16(1-4\lambda)x_3^2$$

令 $g'(\lambda)=0$

解得

$$\lambda=\frac{x_1^2+x_2^2+4x_3^2}{2x_1^2+2x_2^2+16x_3^2}$$

数据科学数学基础 Mathematics DataScience

寻宝游戏(MongoDB)
数据科学与工程数学基础 作业6