hw6_51215903008_陈诺

作业链接

https://www.wolai.com/v2SV4T7sSoBoEAJDKoNCTu

习题部分

$$\begin{gathered}\text{习题1.下面的集合哪些是凸集?}\\(a)\text{平板，即形如}\left\{x\in\mathbf{R}^n\mid\alpha\leqslant a^Tx\leqslant\beta\right\}\text{的集合。}\\(b)\text{矩形,即形如}\left\{x\in\mathbf{R}^n\mid\alpha_i\leqslant x_i\leqslant\beta_i,i=1,\cdots,n\right\}\text{的集合。}\\\text{当}n>2\text{时,矩形有时也称为超矩形。}\\(c)\text{楔形,即}\left\{x\in\mathbf{R}^n\mid a_1^Tx\leqslant b_1,a_2^Tx\leqslant b_2\right\}\text{。}\\(d)\text{到定点的距离比到定集合的距离更近的点构成的集合,即}\\\left\{x\mid\|x-x_0\|_2\leqslant\|x-y\|_2,\forall y\in S\right\}\\\text{其中}S\subseteq\mathbf{R}^n\text{。}\end{gathered}$$

$$(a)\quad\begin{aligned}&a^\top\left(\theta x_1+(1-\theta)x_2\right)\\&=\theta a^\top x_1+(1-\theta)a^\top x_2\ (1)\\&(1)\ \geqslant\theta\alpha+(1-\theta)\alpha=\alpha\\&(1)\ \geqslant\theta\beta+(1-\theta)\beta=\beta(\theta\in[0,1])\end{aligned}\quad\text{是凸集}$$

$$\text{或者由于平板是两个半空间的交集所以是凸集。}$$

$$(b)\quad\begin{aligned}&\text{对}(x_1,..,x_n)^\top\text{的其中一维}x_i,\\&\text{取该维上的两个坐标}x_1^{'},x_2^{'}\\&\theta x_1^{'}+(1-\theta)x_2^{'}\ (1)\\&(1)\ \geqslant\theta\alpha_i+(1-\theta)\alpha_i=\alpha_i\\&(1)\ \geqslant\theta\beta_i+(1-\theta)\beta_i=\beta_i(\theta\in[0,1])\end{aligned}\quad\text{是凸集}$$

$$\text{或者由于矩形是有限个半空间的交集所以是凸集。}$$

$$(c)\quad\begin{aligned}&\text{有}x_1,x_2\text{都满足}\left\{\begin{array}{l}a_1^\top x\leq b_1\\a_2^\top x\leq b_2\end{array}\right.\\&\text{则}a_1{}^\top\left(\theta x_1+(1-\theta)x_2\right)\leqslant\theta b_1+(1-\theta)b_1=b_1\\&\text{同理}a_2^\top\left(\theta x_1+(1-\theta)x_2\right)\leq\theta b_2+(1-\theta)b_2=b_2\end{aligned}\quad\text{是凸集}$$

$$\text{或者由于楔形是两个半空间的交集所以是凸集。}$$

$$\begin{gathered}(d)\\\forall y,\|x-x_0\|_2^2\leq\|x-y\|_2^2\\(x-x_0)^\top(x-x_0)\leq(x-y)^\top(x-y)\\-2x^\top x_0+x_0^\top x_0\leq-2x^\top y+y^\top y\\-2x^\top(x_0-y)\leq y^\top y-x_0^\top x_0\\\text{是一个半空间}\\\text{而该集合可记为}\bigcap_{y\in S}\left\{x\mid\|x-x_0\|_2\leq\|x-y\|_2\right\}\\\text{由于是半空间的交集所以是凸集。}\end{gathered}$$

$$\text{习题2.下面的函数哪些是凸函数?请说明理由。}P66P73$$

$$(a) f(x)=e^{x}+1, x\in\mathbf{R}$$
$$(b) f(x)=\max\left(\|Ax+b\|_2,\left\|x^T x\right\|_1\right), A\in\mathbf{R}^{m\times n}, x\in\mathbf{R}^n, b\in\mathbf{R}^m$$
$$(c) f(x)=-\cos x, x\in[-\pi/2,\pi/2]$$

$$(a)\frac{\partial^2 f}{\partial x^2}=e^x\ge 0\quad f\text{是凸函数}$$

$$(b)\begin{array}{l}\text{由于范数是凸函数，且复合仿射映射是保凸运算，}\\ \text{故若记}f_1(x)=||Ax+b||_2, f_2(x)=||x^Tx||_2,\\ \text{则}f_1(x), f_2(x)\text{是凸函数，}\\ \text{而逐点最大}\max\{f_1(x), f_2(x)\}\text{是保凸运算(证明如下)}\end{array}\qquad f\text{是凸函数}$$

下证逐点最大是保凸运算：

即证如果函数f_1和f_2均为凸函数，则逐点最大函数$f(x)=\max\{f_1(x), f_2(x)\}$

定义域为$\operatorname{dom} f=\operatorname{dom} f_1\cap\operatorname{dom} f_2$,仍然是凸函数。

证：$\forall 0\leqslant\theta\leqslant 1$以及$x, y\in\operatorname{dom} f$,有

$$\begin{aligned}f(\theta x+(1-\theta)y)&=\max\left\{f_1(\theta x+(1-\theta)y), f_2(\theta x+(1-\theta)y)\right\}\\&\leqslant\max\left\{\theta f_1(x)+(1-\theta)f_1(y),\theta f_2(x)+(1-\theta)f_2(y)\right\}\\&\leqslant\theta\max\left\{f_1(x), f_2(x)\right\}+(1-\theta)\max\left\{f_1(y), f_2(y)\right\}\\&=\theta f(x)+(1-\theta)f(y),\end{aligned}$$

从而说明了函数f的凸性。

同样很容易证明,如果函数$f_1,\cdots,f_m$为凸函数,则它们的逐点最大函数

$$f(x)=\max\left\{f_1(x),\cdots,f_m(x)\right\}$$

仍然是凸函数。

补充：最大值函数是凸函数：

由于对任意$0\leqslant\theta\leqslant 1$,函数$f(x)=\max_i x_i$满足

$$\begin{aligned}&f(\theta x+(1-\theta)y)=\max_i\left(\theta x_i+(1-\theta)y_i\right)\\&\leqslant\theta\max_i x_i+(1-\theta)\max_i y_i\\&=\theta f(x)+(1-\theta)f(y)\end{aligned}$$

$$(c)\frac{\partial^2 f}{\partial x^2}=cos(x)\ge 0, x\in[-\pi/2,\pi/2]\quad f\text{是凸函数}$$

习题3.考虑极小化二次函数

$$f_0(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}+\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+r$$

其中,$\boldsymbol{P}\in\mathbb{S}_+^n$ (n 阶半正定矩阵)。给出$\boldsymbol{x}$为f_0最小解的充要条件,

并说明$\boldsymbol{x}$何时无解,有唯一解,有多个解。

$$\frac{\partial f}{\partial x}=\frac{1}{2}\left(P^{\top}+P\right)x+q\overset{P\in S_+^n}{=}Px+q$$

$$\frac{\partial^2 f}{\partial x\partial x^{\top}}=P\succcurlyeq 0\text{，故}f_0\text{是凸函数，有最小解}$$

$$f_0\text{最小解时，}\frac{\partial f}{\partial x}=Px+q=0\text{，}$$

该线性方程的解有以下几种情况：

当$r(P)=r(P,q)=n$时，$P\succ 0$，存在唯一x^*使f最小解,$x^*=-P^{-1}q$

(如果P正定,即P是满秩矩阵,则存在唯一最小解$x^*=-P^{-1}q$)

当$r(P)=r(P,q)<n$时，存在无穷x^*使f最小解,$x^*=-P^{-1}q+\operatorname{Null}(P)$

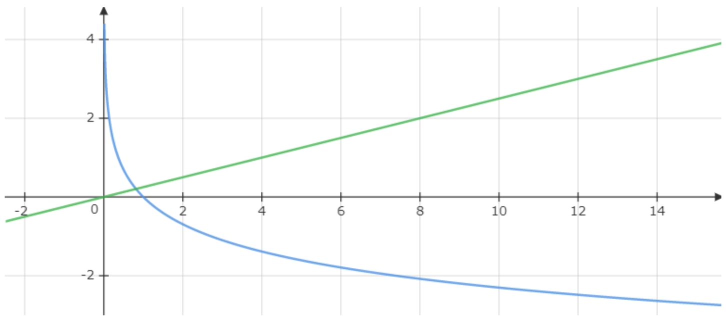
(如果P奇异,最优解集合为$x^*=-P^{\dagger}q+\operatorname{Null}(P)$,其中$P^{\dagger}$为$P$的伪逆。)

当$r(P)<r(P,q)$时，无解,f_0无下界。

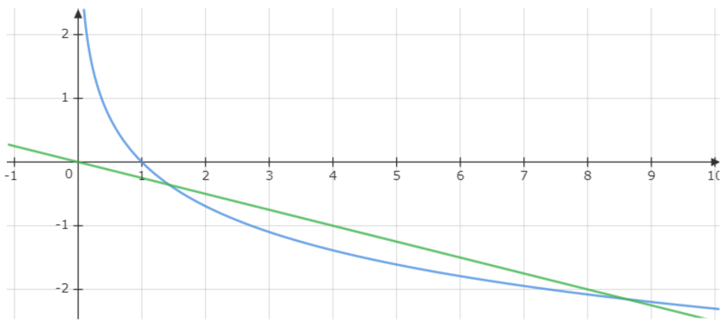
(如果$q\notin\operatorname{Col}(P)$,则无解。此类情况$f_0$无下界。)

$$\text{习题4.计基}f(x)\text{的共轭函数,以及共轭函数的定义域。}P85$$
$$(a)f(x)=-\log x$$
$$(b)f(x)=e^x$$

$$(a)f^*(y)=\sup_x\{yx+\log(x)\}$$
$$if\ \ y\geqslant 0, f^*(y)=\infty$$
$$if\ \ y<0, \frac{\partial(yx+\log(x))}{\partial x}=y+\frac{1}{x}=0$$
$$x=-\frac{1}{y}>0$$
$$f^*(y)=\begin{cases} yx+\log(x)|_{x=-\frac{1}{y}}=-1+\log\left(-\frac{1}{y}\right), y<0 \\ \infty, y\geqslant 0\end{cases}$$
$$D=y|y<0$$

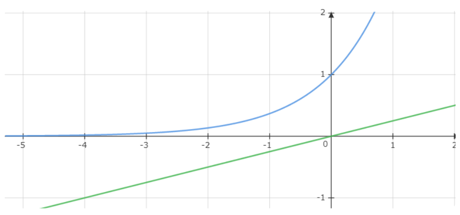


$y\geq 0$

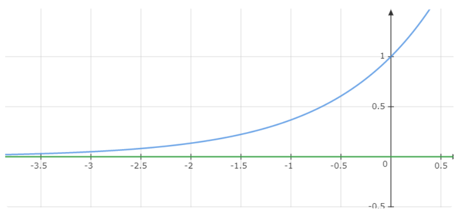


$y<0$

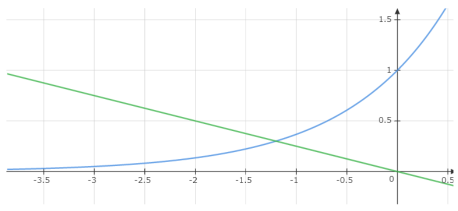
$$(b)f^*(y)=\sup_x\{yx-e^x\}$$
$$=\begin{cases}\sup -e^x=0, & y=0\\ \infty, & y<0\\ y\ln y-y, & y>0\end{cases}$$
$$=\begin{cases} y\ln y-y, & y\geqslant 0\\ \infty, & y<0\end{cases}$$
$$D=\{y|y\geq 0\}$$



$y>0$



$y=0$



$y<0$

$$\text{习题5.证明：}Gauss\text{概率密度函数的累积分布函数}$$
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-u^2/2}du$$
$$\text{是对数}-\text{凹函数.即}\log(\Phi(x))\text{是凹函数。}P101P118$$

$$\frac{\partial\log f}{\partial x}=\frac{f'}{f}\qquad\qquad \text{记}f'=\frac{\partial f}{\partial x}$$
$$\frac{\partial\left(\frac{f'}{f}\right)}{\partial x}=\frac{f''f-f'^2}{f^2}\qquad\qquad f''=\frac{\partial^2 f}{\partial x^2}$$
$$\text{由于}\frac{\partial\left(\frac{f'}{f}\right)}{\partial x}\leq 0\text{为凹}$$
$$\text{对数凹}\Leftrightarrow f''f\leq f'^2$$
$$\text{故函数}f\text{是对数}-\text{凹函数的充要条件是,对任意}x\text{有}f''(x)f(x)\leqslant f'(x)^2\text{。}$$

法1

$$\text{根据两个对数}-\text{凹函数的卷积}(P101\ (f*g)\ (x)=\int f(x-y)g(y)dy)\text{仍然是对数}-\text{凹函数,}$$
$$\text{易得出}f\text{是对数}-\text{凹函数的结论。}$$

下为证明

$$(f*g)(x)=\int f(x-u)g(u)du$$

$$u\in(-\infty,x]$$

$$x-u\in[0,+\infty)$$

此例中，$f(x-u)=\begin{cases}\frac{1}{\sqrt{2\pi}},x-u\in[0,+\infty)\\ 0,\text{others}\end{cases}$

由于$[0,+\infty)$是凸集，故f是对数$-$凹函数

$$g(u)=e^{-\frac{u^2}{2}}$$

$$g'=g\times(-u)$$

$$g''=g'\times(-u)-g=gu^2-g$$

$$gg''=g^2\times\left(u^2-1\right)$$

$$\leqslant g^2u^2=g'^2$$

故g是对数$-$凹函数

由f和g是对数$-$凹函数，故他们的卷积$(f*g)(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}e^{-u^2/2}du$仍为对数凹函数.

法2

证明对数$-$凹函数充要条件：对任意x有$\Phi''(x)\Phi(x)\leqslant\Phi'(x)^2$。

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

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

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

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

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

(a)证明对$x\geqslant 0$有$f''(x)f(x)\leqslant f'(x)^2$。证明如下：

当 $x\geq 0$ 时, $\left(\Phi'(x)\right)^2\geq 0\geq\Phi(x)\Phi''(x)$.

当$x<0$时在后续步骤说明。

(b)证明对任意t和x有$t^2/2\geqslant -x^2/2+xt$。证明如下：

当$x<0$时,由于$\dfrac{u^2}{2}$是凸函数,由

$f(u)\geq f(x)+f'(x)(u-x)$得

$$\frac{u^2}{2}\geq\frac{x^2}{2}+(u-x)x=xu-\frac{x^2}{2}$$

(c)利用(b)证明$e^{-t^2/2}\leqslant e^{x^2/2-xt}$。因此,对$x<0$有

$\displaystyle\int_{-\infty}^{x}e^{-t^2/2}dt\leqslant e^{x^2/2}\int_{-\infty}^{x}e^{-xt}dt$证明如下：

$$\int_{-\infty}^{x}e^{-u^2/2}du\leq\int_{-\infty}^{x}e^{\frac{x^2}{2}-xu}du$$

$$=e^{\frac{x^2}{2}}\int_{-\infty}^{x}e^{-xu}du=e^{\frac{x^2}{2}}\frac{e^{-xu}}{-x}\bigg|_{u=-\infty}^{x}$$

$$=e^{\frac{x^2}{2}}\cdot\frac{e^{-x^2}}{-x}$$

(d)利用(c)证明当$x\leqslant 0$时,$f''(x)f(x)\leqslant f'(x)^2$。证明如下：

$\Phi(x)\Phi''(x)\leq\dfrac{1}{2\pi}e^{-x^2}=\left(\Phi'(x)\right)^2$, $\Phi(x)$ 是对数凹函数.