最优化第一次作业

1. 设 $u,v\in V$，证明 $\langle u,v\rangle=0$ 当且仅当对所有 $a\in\mathbb{F}$ 均有 $\|u\|\le\|u+av\|$

证："$\Longrightarrow$" 若 $\langle u,v\rangle=0$

$$\|u+av\|^2=\|u\|^2+2a\langle u,v\rangle+a^2\|v\|^2=\|u\|^2+a^2\|v\|^2\ge\|u\|^2$$

所以对 $\forall a\in\mathbb{F}$ 均有 $\|a\|\le\|u+av\|$

"$\Longleftarrow$" 若对 $\forall a\in\mathbb{F}$，均有 $\|a\|\le\|u+av\|$

则 $\|u\|^2+2a\langle u,v\rangle+a^2\|v\|^2\ge\|u\|^2$

式① $=2a\langle u,v\rangle+a^2\|v\|^2\ge0$，对 $\forall a$

取 $a=-\dfrac{\langle u,v\rangle}{\|v\|^2}$

则式① $=-\dfrac{2\langle u,v\rangle}{\|v\|^2}\langle u,v\rangle+\dfrac{\langle u,v\rangle}{\|v\|^2}\langle u,v\rangle\ge0$

$\Longrightarrow -\dfrac{\langle u,v\rangle}{\|v\|^2}\ge0$

$\Longrightarrow -\dfrac{\langle u,v\rangle^2}{\|v\|^2}=0$

$\Longrightarrow \langle u,v\rangle=0$

2、证明如果 $C \subset \mathbb{R}^n$ 是凸集，那么 C 的闭包 $\bar{C}$ 也是凸集

证：只需证明 对 $\forall x, y \in \bar{C}$，$\lambda \in [0,1]$ 有

$$\lambda x + (1-\lambda) y \in \bar{C}$$

根据闭包的定义，选择 $\{x_k\}, \{y_k\} \in C$ 使得：

$$\lim_k \|x_k - x\| = 0, \quad \lim_k \|y_k - y\| = 0$$

则 $\|\lambda x_k + (1-\lambda) y_k - [\lambda x - (1-\lambda) y]\|$

$= \|\lambda (x_k - x) + (1-\lambda)(y_k - y)\|$

$\leq \lambda \|x_k - x\| + (1-\lambda) \|y_k - y\|$

两边取极限便有：

$$\lim_k \|\lambda x_k + (1-\lambda) y_k - [\lambda x + (1-\lambda) y]\| = 0$$

从而 $\lambda x + (1-\lambda) y \in \bar{C}$

3. 将下面的问题转化为线性规划：给定 $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$,

(a) $\min_{x\in\mathbb{R}^n}\|Ax-b\|_1 \quad s.t.\ \|x\|_\infty\le 1$

$$\min_{x\in\mathbb{R}^n}\|Ax-b\|_1 = \min_{x\in\mathbb{R}^n}\sum_{i=1}^n|Ax-b|$$

$$= \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^n}\sum_{i=1}^n z_i$$

$$s.t.\ -z\le Ax-b\le z$$

$$\|x\|_\infty\le 1 \Rightarrow -1\le x\le 1$$

$$\Rightarrow \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^n}\sum_{i=1}^n z_i$$

$$s.t.\ -z\le Ax-b\le z$$

$$-1\le x\le 1$$

(b) $\min_{x\in\mathbb{R}^n}\|x\|_1 \quad s.t.\ \|Ax-b\|_\infty\le 1$

$$\min_{x\in\mathbb{R}^n}\|x\|_1 = \min_{x\in\mathbb{R}^n}\sum_{i=1}^n|x_i| = \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^n}\sum_{i=1}^n z_i$$

$$s.t.\ -z\le x\le z$$

$$\|Ax-b\|_\infty\le 1 \Rightarrow -1\le Ax-b\le 1$$

$$\Rightarrow \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^n}\sum_{i=1}^n z_i$$

$$s.t.\ -z\le x\le z$$

$$-1\le Ax-b\le 1$$

(c) $\min_{x\in\mathbb{R}^n}\|Ax-b\|_1+\|x\|_\infty$

$$\min_{x\in\mathbb{R}^n}\|Ax-b\|_1+\|x\|_\infty = \min_{x\in\mathbb{R}^n}\sum_{i=1}^n|Ax-b| + \max_{\{x_1,\cdots,x_n\}}|x_i|$$

$$= \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^n, t\in\mathbb{R}_+}\sum_{i=1}^n z_i + t$$

$$s.t.\ -z\le Ax-b\le z$$

$$-t\mathbf{1}\le x\le t\mathbf{1}$$

$$\Rightarrow \min_{x\in\mathbb{R}^n, x\in\mathbb{R}_+^m, t\in\mathbb{R}_+}\sum_{i=1}^m z_i + t$$

$$s.t.\ -z\le Ax-b\le z$$

$$-t\mathbf{1}\le x\le t\mathbf{1}$$

(d) $\min_{x\in\mathbb{R}^n}\sum_{i=1}^m\max\{0, a_i^Tx+b_i\}$

$$\min_{x\in\mathbb{R}^n}\sum_{i=1}^m\max\{0, a_i^Tx+b_i\} = \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^m}\sum_{i=1}^m z_i$$

$$s.t.\ z_i\ge a_i^Tx+b_i$$

$$\Rightarrow \min_{x\in\mathbb{R}^n, z\in\mathbb{R}_+^m}\sum_{i=1}^m z_i$$

$$s.t.\ z\ge Ax+b$$

4. 设 $u, v \in \mathbb{R}^n$，对于 2-范数 $\|\ \|$，如果

$$\|u+v\| = \|u\| + \|v\|$$

证明对某个实数 λ，$u=0$ 或 $v = \lambda u$

证：$\|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2 + 2\langle u, v\rangle$

$$(\|u\|_2 + \|v\|_2)^2 = \|u\|_2^2 + \|v\|_2^2 + 2\|u\|_2\|v\|_2$$

$$\Rightarrow \langle u, v\rangle = \|u\|_2 \|v\|_2$$

又 $\langle u, v\rangle = \|u\|_2 \|v\|_2 \cos\theta$

$\Rightarrow$ $u=0$ 或 $v = \lambda u$

5、在线性空间中，证明：

1) $k\vec{0}=\vec{0}$

2) $k(\vec{\alpha}-\vec{\beta})=k\vec{\alpha}-k\vec{\beta}$

证：1) $k\vec{0}=k(\vec{\alpha}+(-\vec{\alpha}))=k\vec{\alpha}+k(-\vec{\alpha})=k\vec{\alpha}+k(-1)\vec{\alpha}$

$=k\vec{\alpha}+(-k)\vec{\alpha}+(k+(-k))\vec{\alpha}=0\vec{\alpha}=0$

2) $k(\vec{\alpha}-\vec{\beta})=k(\vec{\alpha}+(-\vec{\beta}))=k\vec{\alpha}+k(-\vec{\beta})=k\vec{\alpha}+k(-1)\vec{\beta}$

$=k\vec{\alpha}+(-1)k\vec{\beta}=k\vec{\alpha}+(-k\vec{\beta})=k\vec{\alpha}-k\vec{\beta}$

6. 二次不等式的解集。令 $C \subseteq \mathbb{R}^n$ 为下列二次不等式的解集

$$C = \{x \in \mathbb{R}^n \mid x^T A x + b^T x + c \leq 0\}$$

其中 $A \in S^n$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$

(a) 证明：如果 $A \succeq 0$，那么 C 是凸集

(b) 证明：如果对某些 $\lambda \in \mathbb{R}$，有 $A + \lambda g g^T \succeq 0$，那么 C 和由 $g^T x + h = 0$（这里 $g \neq 0$ 定义的超平面的交集是凸集）

以上命题的逆命题是否成立？

证明：$\forall\, x_1, x_2 \in C,\ \theta \in [0,1]$

有 $x_1^T A x_1 + b^T x_1 + c \leq 0$, $x_2^T A x_2 + b^T x_2 + c \leq 0$

只需证：$[\theta x_1 + (1-\theta) x_2]^T A [\theta x_1 + (1-\theta) x_2] + b^T[\theta x_1 + (1-\theta) x_2] + c \leq 0$

即证：$\theta^2 x_1^T A x_1 + \theta(1-\theta) x_1^T A x_2 + \theta(1-\theta) x_2^T A x_1 + (1-\theta)^2 x_2^T A x_2 +$

$\theta b^T x_1 + (1-\theta) b^T x_2 + c \leq 0$

即证：$\theta^2(x_1^T A x_1 + b^T x_1 + c) + (1-\theta)^2 (x_2^T A x_2 + b^T x_2 + c) + (\theta - \theta^2)(x_1^T A x_2$

$+ x_2^T A x_1 + b^T x_1 + b^T x_2 + 2c) \leq 0$

原式 $= \theta(x_1^T A x_1 + b^T x_1 + c) + (1-\theta)(x_1^T A x_2 + b^T x_2 + c) + \theta(1-\theta)$

$[x_1^T A (x_2 - x_1) - x_2^T A (x_2 - x_1)]$

$= \theta(x_1^T A x_1 + b^T x_1 + c) + (1-\theta)(x_2^T A x_2 + b^T x_2 + c)$

$\underbrace{-\theta[1-\theta]\underbrace{(x_2 - x_1)^T A [x_2 - x_1]}_{\geq 0}}_{\leq 0} \leq 0$

所以 C 为凸集.

(b) 不妨设超平面 $D=\{x\in\mathbb{R}^n \mid g^Tx+h=0\}$

$\forall x_1,x_2\in C\cap D,\ \theta\in[0,1]$ 有

$\begin{cases} x_1^TAx_1+b^Tx_1+c\leq 0 \\ x_2^TAx_2+b^Tx_2+c\leq 0 \end{cases}$ 且 $\begin{cases} g^Tx_1+h=0 \\ g^Tx_2+h=0 \end{cases}$

由(a)可知 将 $\theta x_1+(1-\theta)x_2$ 代入 $x_1^TAx_1+b^Tx_1+c$

可得：$\underbrace{\theta(x_1^TAx_1+b^Tx_1+c)+(1-\theta)(x_2^TAx_2+b^Tx_2+c)-(\theta-\theta^2)(x_2-x_1)^TA(x_2-x_1)}_{\text{记为式①}}$

$\begin{cases} g^Tx_1+h=0 \\ g^Tx_2+h=0 \end{cases} \Rightarrow g^T(x_2-x_1)=0$

又因为 $\exists\lambda$ s.t. $A+\lambda gg^T\succeq 0$

$\Rightarrow (x_2-x_1)^T(A+\lambda gg^T)(x_2-x_1)\geq 0$

即 $(x_2-x_1)^TA(x_2-x_1)+\lambda(x_2-x_1)^T\underbrace{g(x_2-x_1)}_{=0}\geq 0$

$\Rightarrow (x_2-x_1)^TA(x_2-x_1)\geq 0$

因此式①≤ 0

而 $g^T[\theta x_1+(1-\theta)x_2]+h=\theta(g^Tx_2+h)+(1-\theta)(g^Tx_2+h)=0$

所以 $\theta x_1+(1-\theta)x_2\in C\cap D$

因此 $C\cap D$ 是凸集.

7. 设D为$\mathbb{R}^n$中的闭集，$y\notin D$，证明存在唯一的点$\bar{x}\in D$，使得：

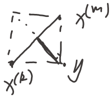
$$\|y-\bar{x}\| = \inf_{x\in D}\|y-x\|$$

证：令 $\inf\limits_{x\in D}\|y-x\| = r > 0$

由下确界的定义可知，存在序列$\{x^{(k)}\}$，$x^{(k)}\in D$，s.t. $\|y-x^{(k)}\|\to r$

先证$\{x^{(k)}\}$存在极限$\bar{x}\in D$，为此只需证明$\{x^{(k)}\}$为柯西序列. 根据平行四边形定律有（对角线的平方和等于一组邻边平方和的两倍）

$$\|x^{(k)}-x^{(m)}\|^2 = 2\|x^{(k)}-y\|^2+2\|x^{(m)}-y\|^2-4\left\|\frac{x^{(k)}+x^{(m)}}{2}-y\right\|^2$$



由于D是凸集，$\dfrac{x^{(k)}+x^{(m)}}{2}\in D$，由$r$的定义，有

$$\left\|\frac{x^{(k)}+x^{(m)}}{2}-y\right\|^2\geq r^2$$

因此 $\|x^{(k)}-x^{(m)}\|^2\leq 2\|x^{(k)}-y\|^2+2\|x^{(m)}-y\|^2-4r^2$

由此可知，当k和m充分大时，$\|x^{(k)}-x^{(m)}\|$充分接近于0，因此$\{x^{(k)}\}$为柯西序列，必存在极限$\bar{x}$，又因为D为闭集，所以$\bar{x}\in D$

再证唯一性. 设存在$\hat{x}\in D$，使

$$\|y-\bar{x}\| = \|y-\hat{x}\| = r$$

由于D为凸集，$\bar{x},\hat{x}\in D$，因此$\dfrac{(\bar{x}+\hat{x})}{2}\in D$，所以

$$\left\|y-\frac{\bar{x}+\hat{x}}{2}\right\| = \left\|\frac{1}{2}y-\frac{\bar{x}}{2}+\frac{1}{2}y-\frac{\hat{x}}{2}\right\|\leq\frac{1}{2}\|y-\bar{x}\|+\frac{1}{2}\|y-\hat{x}\| = r$$

等号成立当且仅当 $y-\bar{x}=\lambda(y-\hat{x})$

由r的定义可知

$$\left\|y-\frac{\bar{x}+\hat{x}}{2}\right\| = \frac{1}{2}\|y-\bar{x}\|+\frac{1}{2}\|y-\hat{x}\|$$

所以 $y-\bar{x}=\lambda(y-\hat{x})$

$\Rightarrow \|y-\bar{x}\| = |\lambda|\,\|y-\hat{x}\|$

由$\bar{x}$与$\hat{x}$的定义可知$|\lambda|=1$，若$\lambda=-1$，则 $y-\bar{x}=-(y-\hat{x})\Rightarrow$ $y=\dfrac{\hat{x}+\bar{x}}{2}\in D$ 与假设矛盾，所以$\lambda\neq-1$，故$\lambda=1$，因此$\bar{x}=\hat{x}$

2.5 两个平行的超平面 $\{x\in\mathbb{R}^n \mid a^Tx=b_1\}$ 和 $\{x\in\mathbb{R}^n \mid a^Tx=b_2\}$ 之间的距离是多少？

解：方法一：由于 a 是超平面的法向量

设 a 与超平面 $\{x\in\mathbb{R}^n \mid a^Tx=b_1\}$，$\{x\in\mathbb{R}^n \mid a^Tx=b_2\}$ 相交于点 x_1 与 x_2

则 $x_1=\left(\frac{b_1}{\|a\|^2}\right)a$

$x_2=\left(\frac{b_2}{\|a\|^2}\right)a$

因为 $a^Tx_1=a^T\frac{b_1}{\|a\|^2}a=\frac{\|a\|^2}{\|a\|^2}b_1=b_1$

因此 $x_1\in\{x\in\mathbb{R}^n \mid a^Tx=b_1\}$

x_1 是 a 的伸缩

因此 $\|x_1-x_2\|_2=\left\|\frac{b_1}{\|a\|^2}a-\frac{b_2}{\|a\|^2}a\right\|=\left\|\frac{b_1-b_2}{\|a\|^2}a\right\|_2=\frac{|b_1-b_2|}{\|a\|_2}$

方法2：设 x_1 为超平面 $\{x\in\mathbb{R}^n \mid a^Tx=b_1\}$ 上的一点

不妨设 x_1 在超平面 $\{x\in\mathbb{R}^n \mid a^Tx=b_2\}$ 上的投影为 x_2

则 $a^Tx_2=b_2$

那么 x_1-x_2 与法向量 a 平行

因此 $a(x_1-x_2)=\|a\|_2\|x_1-x_2\|_2\cos(0\text{或}\pi)$

$=\pm\|a\|_2\|x_1-x_2\|_2$

又因为 $a(x_1-x_2)=ax_1-ax_2=b_1-b_2$

因此 $\|x_1-x_2\|_2=\frac{|b_1-b_2|}{\|a\|_2}$