优化作业

1. ① $\exists x>0 \quad Ax=b \Rightarrow \exists \lambda \quad A^T\lambda \geq 0 \quad A^T\lambda \neq 0 \quad b^T\lambda \leq 0$

先证引理：若 $cx=d$ 对 $\forall x$ 满足 $Ax=b$ 成立，则 $\exists \lambda$, s.t $c=A^T\lambda$, $d=b^T\lambda$

$\because b\in Range(A)$ $\therefore \{x \mid Ax=b\}$ 非空 由线性代数知 可表示为 $x_0+Null(A)$，其中 x_0 为 $Ax_0=b$.

$c^Tx=d \Rightarrow c^T(x_0+x_p)=d \quad \forall x_p\in Null(A)$

$\Rightarrow c^Tx_p=0$，$c^Tx_0=d$. $\forall x_p\in Null(A)$. $\Rightarrow c\in (Null(A))^{\perp}$

$\Rightarrow c\in Range(A^T) \Rightarrow \exists\lambda \quad c=A^T\lambda$.

$\Rightarrow d=c^Tx_0=(A^T\lambda)^Tx_0=\lambda^TAx_0=\lambda^Tb$ 引理得证

考虑 $S_1=\{x \mid Ax=b\}$ $S_2=\{x \mid x>0\}$ 则 $S_1\cap S_2=\emptyset$. S_1, S_2 凸. $\therefore$ 根据分离超平面

$\Rightarrow \exists c\in R^n$. $c\neq 0$. $\alpha\in R \quad \forall x\in S_1 \quad c^Tx\leq\alpha$. 且 $\forall x\in S_2 \quad c^Tx\geq\alpha$.

$\because \forall x\in S_2 \quad c^Tx>\alpha \Rightarrow \alpha=0$ 且 $c\geq 0$.

又 $S_1=x_0+N(A)$ $\therefore \forall x\in S_1 \quad x=x_0+x_p \quad x_p\in N(A)$

$c^Tx=c^Tx_0+c^Tx_p$ $\because N(A)$ 是线性空间 $c^Tx\leq\alpha$ 恒成立 对 $\forall x\in S_1$

$\Rightarrow c^Tx_p=0$

$\therefore \forall x\in S_1 \quad c^Tx=c^Tx_0\leq\alpha$.

由引理 $\exists\lambda \quad c^T=A^T\lambda$. $c^Tx_0=b^T\lambda$

$\Rightarrow A^T\lambda=c^T\geq 0$ 且 $c\neq 0 \quad b^T\lambda=c^Tx_0\leq 0$ 证毕！

② $\exists x>0 \quad Ax=b \Rightarrow \nexists\lambda \quad A^T\lambda\geq 0$. $A^T\lambda\neq 0$. $b^T\lambda\leq 0$

设 $Ax_1=b$, $x_1>0$ $\forall\lambda$ 若 $A^T\lambda\geq 0 \Rightarrow 0\leq x_1^TA^T\lambda=(Ax_1)^T\lambda=b^T\lambda\leq 0$ 矛盾.

2. $\partial C=\{x\in\mathbb{R}^n \mid \|x\|_\infty=1\}$

• $H=\{y\in\mathbb{R}^n: a^Ty=\alpha\}$ 是支持超平面 则

① $x\in H \Rightarrow \alpha=a^Tx$

② $\max\limits_{y\in C} a^Ty\le\alpha$. $\quad y\in C \Leftrightarrow |y_i|\le 1 \ \forall i$. $\quad \max\limits_{y\in C} a^Ty=\sum_{i=1}^n |a_i|$. (取 $y_i=\mathrm{sgn}(a_i)$)

$\therefore \sum_{i=1}^n |a_i|\le\alpha=a^Tx \quad \Rightarrow a_i\begin{cases}=0 & |x_i|<1\\ \ge 0 & x_i=1\\ \le 0 & x_i=-1\end{cases}$

综上. 对 $x\in\partial C$ $\quad H=\{y\in\mathbb{R}^n: a^Ty=a^Tx,\ a_i\begin{cases}=0 & |x_i|<1\\ \ge 0 & x_i=1\\ \le 0 & x_i=-1\end{cases}$

3. (a) 凸 非强凸 $\quad f''(x)=e^x>0$ 且 $\mathbb{R}$ 凸 $\quad \nabla^2 f(x)=e^x$ 无正下界

(b) 非凸 非凹 $\quad \nabla^2 f(x)=\begin{bmatrix}0&1\\1&0\end{bmatrix}$ $\quad \lambda_{1,2}=\pm 1$ 不定

(c) 凸. 非强凸. $\quad \nabla^2 f(x)=\begin{bmatrix}\frac{2}{x_1^3x_2} & \frac{1}{x_1^2x_2^2}\\ \frac{1}{x_1^2x_2^2} & \frac{2}{x_1x_2^3}\end{bmatrix}$ $\quad \det(\nabla^2 f(x))>0$ 对角线>0. ∴正定 ⇒ 凸

无正下界.

(d) ~~凸. 非强凸~~ 非凸 非凹 $\quad \nabla^2 f(x)=\begin{bmatrix}0 & -\frac{1}{x_2^2}\\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3}\end{bmatrix}$ $\quad \det(\nabla^2 f(x))<0$ 对角线>0 ⇒ $\nabla^2 f(x)$ 非正定. ⇒ ~~凸~~ 非凹

(e) 凸. 非强凸 $\quad \nabla^2 f(x)=\begin{bmatrix}\frac{2}{x_2} & -\frac{2x_1}{x_2^2}\\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3}\end{bmatrix}$ $\quad \det(\nabla^2 f(x))=0$ $\quad$ 半正定

(f) 凹. 非强凹. $\quad \nabla^2 f(x)=\begin{bmatrix}\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha}\\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)\alpha x_1^{\alpha}x_2^{-\alpha-1}\end{bmatrix}$ = ~~α(α-1)~~

$$=(1-\alpha)\alpha\begin{bmatrix}-x_1^{\alpha-2}x_2^{1-\alpha} & x_1^{\alpha-1}x_2^{-\alpha}\\ x_1^{\alpha-1}x_2^{-\alpha} & -x_1^{\alpha}x_2^{-\alpha-1}\end{bmatrix}$$ $\det\nabla^2 f(x)=0$. 半负定

4. $f'(t)=\ln\frac{t}{t+1}+(t-1)\frac{1}{t(t+1)}$

$f''(t)=\frac{1}{t(t+1)}+\frac{1}{t(t+1)}+(t-1)\frac{-2t-1}{(t(t+1))^2}=\frac{2t^2+2t-2t^2+2t-t+1}{t^2(t+1)^2}=\frac{3t+1}{t^2(t+1)^2}>0$, 对 $t>0$ 凸

∴ $f(t)$ 强凸.

5. 对 $p\in\mathbb{R}^n$ $\sum p_i=1$ $p_i>0$ $P=\{p\in\mathbb{R}^n \mid \sum p_i=1,\ p_i>0\}$ 是凸集.

$f(p)=\sum p_i \ln p_i$ $\quad (\nabla^2 f(p))_{(i,j)} = \begin{cases} \cancel{\ln p_i+1}\ \frac{1}{p_i} & i=j \\ 0 & \end{cases}$

∴ $\nabla^2 f(p) \succeq I$ $\quad$ (∵ $p_i\le 1$ ∴ $\frac{1}{p_i}\ge 1$).

∴ $f(p)$ 是 1-强凸函数

6. 不妨 $a_3=1$ (否则两边同除以 a_2+a_3 即可) $\Rightarrow a_1+a_2=1$.

由1阶条件 $\quad \langle \nabla f(x_3), x_i-x_3\rangle < f(x_i)-f(x_3) \quad i=1,2.$

∴ $LHS = a_1\langle \nabla f(x_3), x_1-x_3\rangle + a_2\langle \nabla f(x_3), x_2-x_3\rangle$

$< a_1(f(x_1)-f(x_3)) + a_2(f(x_2)-f(x_3)) = a_1 f(x_1) + a_2 f(x_2) - f(x_3) = RHS.$

证毕