221900180 田永铭 凸优化作业6 (1)

3.8 凸性的二阶条件

证明：①先考虑 $f: \mathbb{R} \to \mathbb{R}$ 情形：

$\nabla^2 f(x) \succeq 0$ 条件化为 $f''(x) \geq 0$ (for $\forall x \in \text{dom} f$).

1° 若 f 为凸函数：

则定义域必为凸集. 有 $\forall x, y$, $f(y) \geq f'(x)(y-x) + f(x)$.

$f(x) \geq f'(y)(x-y) + f(y)$.

$\therefore f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x)$.

$\therefore [f'(y) - f'(x)](y-x) \geq 0$. 不妨设 $y > x$ ($y < x$ 同理易证),

$\therefore \dfrac{f'(y)-f'(x)}{y-x} \geq 0$. $\therefore \lim\limits_{y \to x} \dfrac{f'(y)-f'(x)}{y-x} \geq 0$, 即 $f''(x) \geq 0$ for $\forall x$.

$\therefore \Rightarrow$ 成立.

2° 若 f 定义域为凸集且 $f''(x) \geq 0$ (for $\forall x \in \text{dom} f$): 对 $\forall y, \forall x$, 可设 $y > x$, ($y < x$ 同理可证).

则有 $\int_x^y f''(z)(y-z)\,dz \geq 0$.

$\int_x^y f''(z)(y-z)\,dz = \int_x^y (y-z)\,d f'(z) = [f'(z)(y-z)]\Big|_{z=x}^{z=y} - \int_x^y f'(z)(-1)\,dz$

$\therefore \int_x^y f''(z)(y-z)\,dz$ $= f(y) - f(x) - f'(x)(y-x) \geq 0$.

即 $\forall x, \forall y$, $f(y) \geq f(x) + f'(x)(y-x)$. ($y < x$ 同理)

此即一阶条件. $\therefore f(x)$ 为凸函数.

$\therefore \Leftarrow$ 成立.

②再考虑高维：利用"切片法" $g(t) = f(\hat{x} + tv)$ 可将 f 降维. v 任取, 表 $\forall$ 方向.

有定理: $g(t)$ 凸 iff f 凸.

而 $g''(t) = v^T \nabla^2 f(\hat{x} + tv) v$

$\therefore f$ 凸 $\Leftrightarrow g(t)$ 凸 $\underset{①所证}{\Leftrightarrow} g''(t) \geq 0 \Leftrightarrow \underset{\text{for } \forall v}{v^T \nabla^2 f(\hat{x}+tv) v \geq 0} \Leftrightarrow \nabla^2 f(\hat{x}+tv) \succeq 0$.

$\therefore$ 证毕.

3.36 推导共轭函数

(a) max 函数

解: $f^*(y) = \sup\limits_{x \in \text{dom} f} (y^T x - f(x)) = \sup\limits_{x \in \mathbb{R}^n} (y^T x - x_1) = \sup\limits_{x \in \mathbb{R}^n} (y_1 x_1 + y_2 x_2 + \cdots + y_n x_n - x_1)$

(x_1 表示 x 最大的分量)

若 $\exists y_i < 0$, 则使其它变量固定, 仅改变 x_i, 使 x_i 无限小 $\to -\infty$, 则函数值 $\to +\infty$.

此时 $f^*(y) = \infty$. 若 $\exists y_i > 1$, 取 $x_i \to \infty$ 且 $y^T x - f(x) \to \infty$, $f^*(y) = \infty$.

$\therefore y \succeq 0$. 且 $y \preceq 1$.

同时: 若 $1^T y \neq 1$, 即 $y_1 + y_2 + \cdots + y_n \neq 1$. 1° 不妨先讨论 $y_1 + y_2 + \cdots + y_n > 1$ ($y_1, y_2 \cdots y_n \geq 0$).

则取 $\hat{x} = (x_1, x_1, x_1 \cdots)$ (全为最大分量), 有 $y^T x - x_1 = \lambda x_1$ 且 $\lambda > 0$,

当 $x_1 \to \infty$ 时, $\hat{x} \to \infty$, $y^T x - f(x) \to \infty$, $\therefore f^*(y) = \infty$.

2° 再讨论 $y_1 + y_2 + \cdots + y_n < 1$ ($y_1, y_2 \cdots y_n \geq 0$).

同理取 $\hat{x}$ 每个分量相同且 $\to -\infty$ 可使 $y^T x - f(x) = +\infty$. $\therefore f^*(y) = \infty$.

若 $y \succeq 0$ 且 $1^T y = 1$, 可得 $y^T x - x_1 \leq 0$ 且当 $x_i = x_1$ 且 $y_i = 1$ 时能取到等.

$\therefore f^*(y) = \begin{cases} 0, & y \succeq 0,\ 1^T y = 1, \\ +\infty, & \text{otherwise}. \end{cases}$

(b) 最大若干分量的和

$f(x)=\sum_{i=1}^{r} x_{[i]}$. $f^*(y)=\sup_x\left(y^Tx-\sum_{i=1}^{r}x_{[i]}\right)=\sup_x\left(y_1x_1+y_2x_2+\cdots+y_nx_n-x_{[1]}-x_{[2]}-\cdots-x_{[r]}\right)$.

① 若 $\exists y_i<0$：同(a)中理，$f^*(y)=\infty$.

若 $\exists y_i>1$：同(a)中理：$f^*(y)=\infty$.

② 若 $0\preceq y\preceq 1$：

1° 若 $1^Ty>r$：即 $y_1+y_2+\cdots+y_n>r$. 取 $\hat{x}=(x_{[1]},x_{[1]},\cdots x_{[1]})$（全为最大分量）.

则 $y^T\hat{x}-\cancel{f(\hat{x})}>r\,x_{[1]}-x_{[1]}-x_{[2]}-\cdots-x_{[r]}\ \cancel{\geq 0}>0$，

~~且固定 $x_{[1]}$，使 $x_{[2]}\ \ x_{[3]}$~~ 随 $x_{[1]}\to\infty$，$y^T\hat{x}-f(\hat{x})\to\infty$，$\therefore f^*(y)=\infty$.

$\therefore y^T\hat{x}-f(\hat{x})=\lambda x_{[1]}$ 且 $\lambda>0$. $\therefore$ 随 $x_{[1]}\to\infty$ $(\hat{x}=(x_{[1]},x_{[1]}\cdots x_{[1]}))$（全为最大分量），有 $y^T\hat{x}-f(\hat{x})\to+\infty$. $\therefore f^*(y)=\infty$.

2° 若 $1^Ty<r$，同理取 $\hat{x}=(x_{[1]},x_{[1]}\cdots x_{[1]})$，随 $x_{[1]}\to-\infty$，有 $y^T\hat{x}-f(\hat{x})\to+\infty$.

则 $y^T\hat{x}-f(\hat{x})=\lambda x_{[1]}$ 且 $\lambda<0$.

3° 若 $1^Ty=r$，则 $y^Tx-\sum_{i=1}^{r}x_{[i]}\leq 0$ 且等式能在 $x_{[1]},x_{[2]}\cdots x_{[r]}$ 对应的 $y_{[1]},y_{[2]}\cdots y_{[r]}$ 都取1的时候取到

注意限于 $y_i\leq 1$

综上：$f^*(y)=\begin{cases}0, & 0\preceq y\preceq 1,\ 1^Ty=r,\\ \infty, & \text{otherwise}.\end{cases}$

4.21 QCQP（easy的）问题

解(a) ①凸问题下的解：

令 $y=A^{\frac{1}{2}}x$，$\hat{c}=A^{-\frac{1}{2}}c$. (*)

则原问题转化为：

minimize $\hat{c}^Ty$

subject to $y^Ty\leq 1$.

这是一个unit ball的限制，$\therefore y^*=\frac{-\hat{c}}{\|\hat{c}\|_2}$.

$\therefore$ 最优值 $p^*=\hat{c}^Ty^*=\hat{c}^T\frac{-\hat{c}}{\|\hat{c}\|_2}=-\|\hat{c}\|_2$.

代入(*)式得：$x^*=\frac{-A^{-1}c}{\sqrt{c^TA^{-1}c}}$，$p^*=-\|A^{-\frac{1}{2}}c\|_2=-\sqrt{(A^{-\frac{1}{2}}c)^TA^{-\frac{1}{2}}c}=-\sqrt{c^TA^{-1}c}$.

②非凸问题下的解：

此时 $A\notin S^n_+$. A的逆不一定$\exists$. 因此不可和①一样地做.

将A进行特征值分解. 令 $A=P\Lambda(\lambda)P^T=\sum_{i=1}^{n}\lambda_ip_ip_i^T$.（$P$为正交阵）

~~令y~~ 原问题为 minimize c^Tx subject to $x^TAx\leq 1$.

令 $x=P^{-1}y$，$c=P^{-1}d$，则原问题 $\Leftrightarrow$ minimize $c^TP^{-1}y=d^T(P^{-1})^TP^{-1}y=d^Ty$（由于$P$为正交阵）

subject to $(P^{-1}y)^TP\Lambda(\lambda)P^TP^{-1}y=y^T\Lambda(\lambda)y=\sum_{i=1}^{n}\lambda_iy_i^2\leq 1$.

即 $\begin{cases}\text{minimize} & \sum_{i=1}^{n}d_iy_i\\ \text{subject to} & \sum_{i=1}^{n}\lambda_iy_i^2\leq 1\end{cases}$ →（向量不等式）

① 当 $\lambda\succ 0$ 时，①中已给出：$x^*=\frac{-A^{-1}c}{\sqrt{c^TA^{-1}c}}$，$p^*=-\sqrt{c^TA^{-1}c}$.

② 若$\exists i$，$\lambda_i<0$ 时，则可使 $y_i\to+\infty$ 或 $y_i\to-\infty$，则可使 $\sum_{i=1}^{n}d_iy_i\to-\infty$，$\therefore$ 问题p^*无下界.

$\because d$不为0矩阵

③ 若 ~~$\forall i,\lambda_i\geq 0$，$\exists i,\lambda_i=0$~~ λ_n（λ_n为最小特征值）$=0$，且$\exists i$，$\lambda_i=0$ 且 $d_i\neq 0$，则 p^*无下界

④ 若 λ_n（λ_n为最小特征值）$=0$，且$\forall i$，$\lambda_i=0\Rightarrow b_i=0$ 成立，则问题又可转化为①情形（找第一个非0 λ 就好取即可）.

解毕.

2221900180 邱永锐 凸优化作业6 (2)

4.21 (b) 解：与(a)类似地，我们希望 $(x-x_c)^TA(x-x_c)\le 1$ 条件能化为规范化的 $y^Ty\le 1$ 形式.

easy! 令 $y=A^{\frac{1}{2}}(x-x_c)$ 即 $x=A^{-\frac{1}{2}}y+x_c$, (*)

则原问题化为 minimize $c^TA^{-\frac{1}{2}}y+c^Tx_c$

subject to $y^Ty\le 1$.

$y^*=\dfrac{-c^TA^{-\frac{1}{2}}}{\|c^TA^{-\frac{1}{2}}\|_2}$, $p^*=c^TA^{-\frac{1}{2}}y+c^Tx_c$ → 定值.

代入(*)式，得：最优解为 $x^*=x_c-\dfrac{A^{-1}c}{\|A^{\frac{1}{2}}c\|_2}$.

(c) ①若 $B\in S_+^n$，则 $x^TBx\ge 0$，又 $x=0$ 时既满足 $x^TAx\le 1$，又使 x^TBx 取得最小值 0.

$\therefore x_*=0, p^*=0$.
(解)

②若 $B\notin S_+^n$：由题目提示和课本附录(B.1)得：

法1 该问题等价于 maximize γ

subject to $\lambda\ge 0$

$$\begin{bmatrix} B+\lambda A & 0 \\ 0 & -\lambda-\gamma \end{bmatrix}\succeq 0.$$

不等式约束可化为 $\lambda\ge 0$，$B+\lambda A\succeq 0$，$-\lambda-\gamma\ge 0$. 同时 $B+\lambda A\succeq 0$ 也能满足，

若 $\lambda\to 0$，则 γ 可 $\to 0$，此时 $\gamma\to 0$. 取 γ 达到0或大于0，λ 必 ≤ 0，则 $B\succeq 0$，与 $B\notin S_+^n$ 矛盾.

$\therefore$ 由上述讨论知：该问题最优值必 ≤ 0.

solve SDP半定规划即可解得：$p^*=\begin{cases}\lambda_{min}(A^{-\frac{1}{2}}BA^{\frac{1}{2}}), & \lambda_{min}(A^{\frac{1}{2}}BA^{\frac{1}{2}})\le 0,\\ 0, & \text{otherwise}.\end{cases}$

法(2)

可令 $y=A^{\frac{1}{2}}x$（基于 $A\succ 0$）. $\therefore$ 原问题化为 minimize $y^TA^{-\frac{1}{2}}BA^{\frac{1}{2}}y$

subject to $y^Ty\le 1$.

①若 $y^Ty=1$ 时达到最优，

则最优值即为 $\lambda_{min}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})$.

理由如下：若 $y^Ty=1$，则将 $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ 记为 K，并特征值分解为 $P\Lambda P^T$.

P 为正交阵. $\because y^Ty=1$ 且 $PP^T=E$，$\therefore (y^TP)(P^Ty)=y^T(PP^T)y=1$.

$\therefore$ minimize 的是 $y^TP\Lambda P^Ty$ 即 $U^T\Lambda U$ 且 $U^TU=1$. 即

minimize $\lambda_1u_1^2+\cdots+\lambda_nu_n^2$ subject to $u_1^2+\cdots+u_n^2=1$,

$\therefore p^*$ 为 λ_{min}.

②若 $y^Ty<1$ 时达到最优，则在梯度为0处取到. 有 $(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})y=0$. 又 $\because A\succ 0$，

$\therefore y^TA^{-\frac{1}{2}}BA^{-\frac{1}{2}}y=0$ $\therefore p^*=0$. 综上：$p^*=\begin{cases}\lambda_{min}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}), & \lambda_{min}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})\le 0,\\ 0, & \text{otherwise}.\end{cases}$ 证毕.

~~4.24 复 l_1-, l_2-, l_∞-范数逼近~~

4.24 复 l_1-, l_2- 和 l_∞- 范数逼近

(a) 复 l_1- 范数：本题中：记 $R(x)$ $(x\in C)$ 表示 x 的实部，$S(x)$ $(x\in C)$ 表示 x 的虚部。

(a_i^T $(i=1,2\cdots m)$ 为 A 的行向量)

该问题为 minimize $\sum_{i=1}^{m}|a_i^Tx-b|$.

引入变量 $t_1,t_2\cdots t_m$.

则问题 $\Leftrightarrow$ minimize $\sum_{i=1}^{m} t_i$

subject to $|a_i^Tx-b|<t_i$ $i=1,2,3\cdots m$.

实、虚部分解得 $|a_i^Tx-b|^2 = (R(a_i^T)R(x)-S(a_i^T)S(x)-R(b))^2+(R(a_i^T)S(x)+S(a_i^T)R(x)-S(b))^2$

$= R^2(a_i^T)R^2(x)+S^2(a_i^T)S^2(x)+R^2(b)-2R(a_i^T)R(x)S(a_i^T)S(x)$

$-2R(b)R(a_i^T)R(x)+2R(b)S(a_i^T)S(x)$

$+R^2(a_i^T)S^2(x)+S^2(a_i^T)R^2(x)+S^2(b)$

$+2R(a_i^T)S(x)S(a_i^T)R(x)-2R(a_i^T)S(x)S(b)-2R(x)S(a_i^T)S(b)$

$=R(a_i^Tx-b)^2+S(a_i^Tx-b)^2$

令 $z^T=[R(x)^T\ S(x)^T]$.

则 $|a_i^Tx-b|^2=\left\|\begin{bmatrix}R(a_i^T) & -S(a_i^T)\\ S(a_i^T) & R(a_i^T)\end{bmatrix}z-\begin{bmatrix}R(b)\\ S(b)\end{bmatrix}\right\|_2^2$.

$\therefore$ 原问题 $\Leftrightarrow$ minimize $\sum_{i=1}^{m} t_i$

subject to $\left\|\begin{bmatrix}R(a_i^T) & -S(a_i^T)\\ S(a_i^T) & R(a_i^T)\end{bmatrix}z-\begin{bmatrix}R(b)\\ S(b)\end{bmatrix}\right\|_2<t_i$, $i=1,2\cdots m$

即SOCP问题.

(b) 复 l_∞-范数：

发现：推导与上面的完全一致

而 l_∞-范数表示的是最大绝对值分量.

$\therefore$ 问题 $\Leftrightarrow$ minimize t

subject to $\left\|\begin{bmatrix}R(a_i^T) & S(a_i^T)\\ S(a_i^T) & R(a_i^T)\end{bmatrix}z-\begin{bmatrix}R(b)\\ S(b)\end{bmatrix}\right\|_2<t$ $i=1,2\cdots m$.

也为SOCP.

亦可将 t 和 $\|\cdots\|_2$ 均平方转化为QCQP.

(c) 复 l_2-范数：即 minimize $\|Ax-b\|_2^2$，即 $\left\|\begin{bmatrix}R(A) & -S(A)\\ S(A) & R(A)\end{bmatrix}z-\begin{bmatrix}R(b)\\ S(b)\end{bmatrix}\right\|_2^2$.

无约束问题.

(推导过程仍同上)

此为无约束 ~~QCQP问~~ QCQP(更特别地、QP)问题.

$\therefore$ 问题的二次表述形式如上给出. 解毕.

221900180 田永铭 凸优化作业6 (3)

5.1 (a) 简单例子

解：① 可行集：令 $(x-2)(x-4)\le 0$ 且 $x\in R$，得：$x\in[2,4]$．此即可行集．

② 令 $f_0(x)=x^2+1$，$\therefore f_0'(x)=2x+1$，当 $x\in[2,4]$ 时，$f_0'(x)>0$ 恒成立，$\therefore f_0(x)$ 在 $[2,4]$ 上增．

$\therefore f_0(x)_{min}=f_0(2)=2^2+1=5$．

$\therefore$ 最优解为 $x_0=2$，最优值为 $p^*=5$．

5.2 无界问题及不可行问题的弱对偶性

证明：(1) 若 $p^*=-\infty$：

$\because p^*=-\infty$，$\therefore$ 原问题无下界．即 $f_0(x)$ 在 $f_i(x)$ 和 $h_i(x)$ 的限制下有 $\tilde{x}$ 使 $f_0(\tilde{x})=-\infty$ $(\forall i)$．

而 $L(x,\lambda,\nu)=f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}\nu_i h_i(x)$．

$$g(\lambda,\nu)=\inf_x\left(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}\nu_i h_i(x)\right)\le f_0(\tilde{x})+\sum_{i=1}^{m}\lambda_i f_i(\tilde{x})+\sum_{i=1}^{p}\nu_i h_i(\tilde{x})\le -\infty+0+0=-\infty$$

$\therefore \sup\limits_{\lambda\ge 0,\nu} g(\lambda,\nu)$ 问题不可解．$\therefore d^*=-\infty$（注意：这是maximize问题，不可解时为 $-\infty$）．

(2) 若 $d^*=+\infty$，则 ~~dual~~ 对偶问题无上界．

即 $g(\lambda,\nu)$ 可取到 $+\infty$．而 $g(\lambda,\nu)=\inf\limits_x\left(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}\nu_i h_i(x)\right)$

反设 $p^*\ne+\infty$，即原问题可行，设最优解为 $\tilde{x}$，则 $g(\lambda,\nu)\le f_0(\tilde{x})+\sum_{i=1}^{m}\lambda_i f_i(\tilde{x})+\sum_{i=1}^{p}\nu_i h_i(\tilde{x})$

$\le p^*+0+0$；$\therefore g(\lambda,\nu)$ 不可能取到无穷．

$\therefore$ 假设不成立．$\therefore p^*=+\infty$．

综上：证毕．