221900180 田永铭 凸优化作业3

3.3 增的凸函数的反函数

解：g是凹函数，证明如下：$y=f(x)$ 与 $x=g(y)$ 互为反函数，由题：$f'(x)>0$ 且 $f''(x)>0$.

$$\cancel{g'(y)=g'(f(x))=\frac{1}{f'(x)}\left(\frac{dy}{dx}=\frac{dx}{dy}\right)}$$

$$\cancel{\therefore g''(y)=g''(f(x))=-\frac{f''(x)}{f'(x)^2}\quad\left(\frac{1}{y'}\right)'=-\frac{1}{y'^2}}$$

$$g'(y)=\frac{dx}{dy}=\frac{1}{\frac{dy}{dx}}=\frac{1}{f'(x)}.$$

$$g''(y)=\frac{d(g'(y))}{dy}=\frac{d\left(\frac{1}{f'(x)}\right)}{dx}\cdot\frac{dx}{dy}=-\frac{f''(x)}{[f'(x)]^2}\cdot\frac{1}{f'(x)}=-\frac{f''(x)}{f'(x)^3}<0.$$

∴ g为凹函数.

3.4 凸函数等价定义 $f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$.

证明：① ⇒：若f为凸函数，则 $\forall\theta\in[0,1]$, $f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$.

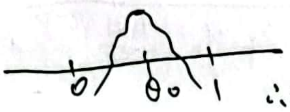
⇔ $\forall\lambda\in[0,1]$, $f(x+\lambda(y-x))\le f(x)+\lambda(f(y)-f(x))$.

∴ $\int_0^1 f(x+\lambda(y-x))d\lambda\le\int_0^1[f(x)+\lambda(f(y)-f(x))]d\lambda=\frac{f(x)+f(y)}{2}$ ∴充分性得证

② ⇐：若 $\forall x,y\in R^n$, $\int_0^1 f(x+\lambda(y-x))d\lambda\le\frac{f(x)+f(y)}{2}$，则：

反假设f不是凸函数，即 $\exists\theta_0\in(0,1)$, s.t. $f(\theta_0x+(1-\theta_0)y)>\theta_0f(x)+(1-\theta_0)f(y)$

令 $g(\theta)=f(\theta x+(1-\theta)y)-\theta f(x)-(1-\theta)f(y)$. 则 $g(0)=g(1)=0$，且 $g(\theta_0)>0$.



∴ $\exists\delta$, s.t. 在 $\Delta=(\theta_0-\delta,\theta_0+\delta)$ 内，$g(\theta)>0$ 恒成立.

∴ 对于 θ_1、$\theta_2\in\Delta$ 邻域，$g(\theta_1)>0$, $g(\theta_2)>0$. $g(\lambda\theta_1+(1-\lambda)\theta_2)>0$ $\lambda\in[0,1]$.

∴ $\int_0^1 f(\theta_1+\theta(\theta_2-\theta_1))d\theta>\int_0^1(\theta f(\theta_2)+(1-\theta)f(\theta_1))d\theta=\frac{f(\theta_1)+f(\theta_2)}{2}$

与 $\int_0^1 f(x+\lambda(y-x))d\lambda\le\frac{f(x)+f(y)}{2}$ 对 $\forall x,y$ 成立相矛盾. ∴假设不成立.

∴原结论得证.

3.13 Kullback-Leibler散度和信息不等式.

证明：① $D_{KL}(u,v)=f(u)-f(v)-\nabla f(v)^T(u-v)$，要证 $D_{KL}\ge0$，只需证明 $f(x)$ 是凸函数.

而 $f(x)=\sum_{i=1}^n x_i\log x_i$，有 $x_i>0$，令 $g(x)=x\log x$，∴ $g'(x)=\log x+1$，$g''(x)=\frac{1}{x}>0$.

∴ $g(x)=x\log x$ 为凸函数. 由结论"凸函数求和仍为凸函数"知，$f(x)$ 为凸函数，

即有 $f(u)-f(v)-\nabla f(v)^T(u-v)\ge0$ 恒成立.

② 当 $u=v$ 时，等式显然成立；

若 $f(u)-f(v)-\nabla f(v)^T(u-v)=0$，即 u,v 两点连线与u处f切线(梯度)相同，

又∵ $f(x)$ 处处不线性，∴仅可能 $u=v$. 得证.

3.17 $f(x)=(\sum_{i=1}^{n}x_i^p)^{\frac{1}{p}}$. $(p<1, p\neq 0)$. 凹. $\operatorname{dom} f = R_{++}^n$

(a) 证明: $\frac{\partial f(x)}{\partial x_i} = \frac{1}{p}(\sum_{i=1}^{n}x_i^p)^{\frac{1-p}{p}} \cdot p x_i^{p-1} = (\frac{f(x)}{x_i})^{1-p}$.

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{(1-p)}{x_i}(\frac{f(x)}{x_i})^{1-2p} - \frac{(1-p)}{x_i^2}(\frac{f(x)}{x_i})^{-p}f(x) = \frac{1-p}{f(x)}(\frac{f(x)^2}{x_i^2})^{1-p} - \frac{1-p}{x_i}(\frac{f(x)}{x_i})^{1-p}$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{f(x)}(\frac{f(x)^2}{x_i x_j})^{1-p} \quad (i\neq j).$$

(b) $f(x)$凹 $\Leftrightarrow y^T\nabla^2 f(x) y \le 0 \Leftrightarrow (y_1, y_2 \cdots y_n)\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots \\ \vdots & \ddots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n\end{pmatrix} \le 0.$

$\Leftrightarrow \frac{1-p}{f(x)}\left((\sum_{i=1}^{n}\frac{y_i f(x)^{1-p}}{x_i^{1-p}})^2 - \sum_{i=1}^{n}\frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) \le 0$

(这步实在化不出来，参考了一下答案)

又 $\frac{1-p}{f(x)}>0$, ∴只需证 $(\sum_{i=1}^{n}\frac{y_i f(x)^{1-p}}{x_i^{1-p}})^2 - \sum_{i=1}^{n}\frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} \le 0$.

∵ $a^Tb \le \|a\|_2\|b\|_2$ (柯西不等式), ∴令 $a_i = (\frac{f(x)}{x_i})^{-\frac{p}{2}}$, $b_i = y_i(\frac{f(x)}{x_i})^{1-\frac{p}{2}}$,

则 $a^Tb \le \|a\|_2\|b\|_2$

$$\Leftrightarrow \left((\frac{f(x)}{x_1})^{-\frac{p}{2}} \quad (\frac{f(x)}{x_2})^{-\frac{p}{2}} \quad \cdots \quad (\frac{f(x)}{x_n})^{-\frac{p}{2}}\right)\begin{pmatrix} y_1(\frac{f(x)}{x_1})^{1-\frac{p}{2}} \\ y_2(\frac{f(x)}{x_2})^{1-\frac{p}{2}} \\ \vdots \\ y_n(\frac{f(x)}{x_n})^{1-\frac{p}{2}}\end{pmatrix}$$

$$\le \underbrace{\sqrt{(\frac{f(x)}{x_1})^{-p} + (\frac{f(x)}{x_2})^{-p} + \cdots + (\frac{f(x)}{x_n})^{-p}}}_{=\frac{x_1^p}{x_1^p+x_2^p+\cdots+x_n^p}+\cdots+\frac{x_n^p}{x_1^p+x_2^p+\cdots+x_n^p}} \sqrt{y_1^2(\frac{f(x)}{x_1})^{2-p} + \cdots + y_n^2(\frac{f(x)}{x_n})^{2-p}},$$

$= 1$

∴ $(\sum_{i=1}^{n}\frac{y_i f(x)^{1-p}}{x_i^{1-p}})^2 - \sum_{i=1}^{n}\frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} \le 0$. ∴得证.

3.19 非负加权和以及积分

(a) 证明: 由题: $f(x)=\sum_{i=1}^{r}x_{[i]}$ 是凸函数,

且 $f(x)=\sum_{i=1}^{r}\alpha_i x_{[i]} = \alpha_r(x_{[1]}+x_{[2]}+\cdots+x_{[r]}) + (\alpha_{r-1}-\alpha_r)(x_{[1]}+x_{[2]}+\cdots+x_{[r-1]})$
$+\cdots+(\alpha_1-\alpha_2)x_{[1]}$

又 $\alpha_1, \alpha_2 \cdots \alpha_r \ge 0$, 且非负加权求和后仍为凸函数, ∴得证.
凸函数

(b) 证明: $f(x) = -\int_0^{2\pi}\log T(x,\omega)d\omega = \int_0^{2\pi} -\log T(x,\omega)d\omega$ 是对 $-\log T(x,\omega)$ 的求和(相当于).

而固定ω, 令 $g(x,\omega) = -\log(x_1+x_2\cos\omega+\cdots+x_n\cos(n-1)\omega)$.
$\triangleq -\log(a^Tx+b)$, 是$-\log$(仿射), 是凸函数.

而凸函数求和(积分)仍为凸函数.

∴ $f(x)$为凸函数. 得证.

221900180 周永铭 凸优化作业3 (2)

3.23 透视函数

(a) 证明: $\because t>0$, ~~且 g(x) 为 g(x)~~ 且 $g(x)=|x_1|^p+\cdots+|x_n|^p$ 的透视函数

为 $h(x,t)=tg(\frac{x}{t})=t\left[\frac{|x_1|^p}{t^p}+\cdots+\frac{|x_n|^p}{t^p}\right]=\frac{\|x\|_p^p}{t^{p-1}}=f(x,t)$,

$\therefore f(x,t)$ 为 $g(x)$ 透视函数. 又 $\because g(x)$ 为凸函数范数之和, $\therefore g(x)$ 为凸函数.

又 $\because$ 凸函数的透视函数仍为凸函数, $\therefore f(x,t)$ 为凸函数.

(b) 证明: $f(x)=\frac{\|Ax+b\|_2^2}{c^Tx+d}$. 令 $y=Ax+b$, $t=c^Tx+d$.

则 $f(x)=g(y,t)=\frac{\|y\|_2^2}{t}=\frac{y^Ty}{t}$.

令 $h(x)=x^Tx=\|x\|_2^2$, 则 $h(x)$ 的透视函数为 $l(x,t)=th(\frac{x}{t})$

$=t(\frac{x}{t})^T(\frac{x}{t})=\frac{x^Tx}{t}$

由范数是凸的, 知 $h(x)$ 凸; 由凸函数的透视函数亦凸, 知 $l(x,t)$ 或说 $g(y,t)$ 凸.

又 $\because y=Ax+b$ 和 $t=c^Tx+d$ 均为仿射变换, $\therefore$ 最终 $f(x)$ 为凸函数.

3.32 凸函数的积或比

(a) 证明: 不妨以非减为例 (非增同理).

$\because f,g$ 凸, 设 $\forall\theta\in[0,1]$, $\therefore f(\theta x+(1-\theta)y)g(\theta x+(1-\theta)y)\le[\theta f(x)+(1-\theta)f(y)][\theta g(x)+(1-\theta)g(y)]$

$=\theta f(x)g(x)+(1-\theta)f(y)g(y)+\underbrace{\theta(1-\theta)(f(y)-f(x))(g(x)-g(y))}_{(*)}$.

由于 $f、g>0$ 且 $f、g$ 非减, $\therefore(*)$式 ≤0, $\therefore f(\theta x+(1-\theta)y)g(\theta x+(1-\theta)y)\le\theta f(x)g(x)+(1-\theta)f(y)g(y)$.

$\therefore fg$ 在此区间上是凸函数.

(b) $\because f、g$ 凹, $\therefore$ 设 $\forall\theta\in[0,1]$, 有 $f(\theta x+(1-\theta)y)g(\theta x+(1-\theta)y)$

$\ge[\theta f(x)+(1-\theta)f(y)][\theta g(x)+(1-\theta)g(y)]$

$=\theta f(x)g(x)+(1-\theta)f(y)g(y)+\underbrace{\theta(1-\theta)(f(y)-f(x))(g(x)-g(y))}_{(*)}$.

$\because f、g$ 一个非减, 一个非增, $\therefore(*)$式 ≥0,

$\therefore f(\theta x+(1-\theta)y)g(\theta x+(1-\theta)y)\ge\theta f(x)g(x)+(1-\theta)f(y)g(y)$.

$\therefore fg$ 是凹函数.

(c) $f/g=f\cdot\frac{1}{g}$. $\because g>0$, g 非增, g 凹, $\therefore\frac{1}{g}>0$, $\frac{1}{g}$ 非减, $\frac{1}{g}$ 凸.

$\therefore f、\frac{1}{g}$ 满足 (a) 中的条件, $\therefore\frac{f}{g}$ 凸.

3.47 证明: f 是对数凹函数

$\Leftrightarrow$ $\log f$ 是凹函数

$\Leftrightarrow$ for $\forall\theta\in[0,1]$, $\log f(y)\le\log f(x)+\nabla(\log f(x))^T(y-x)$.

$\Leftrightarrow$ for $\forall\theta\in[0,1]$, $\log f(y)-\log f(x)\le\frac{\nabla f(x)^T}{f(x)}(y-x)$

取 exp. $\Leftrightarrow$ $\frac{f(y)}{f(x)}\le\exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right)$.

$\therefore$ 得证.

3.57 证明 $f(X)=X^{-1}$ 在 S^n_{++} 上是矩阵凸的.

证明：要证 $f(X)=X^{-1}$ 在 S^n_{++} 上矩阵凸，

只要说明对 $\forall$ 向量 z，有 $z^T f(x) z$ 是凸函数（标量函数）.

即下证 ~~$f(x,y)=f(z)$~~ $g(x,Y)=x^TY^{-1}x$ 为凸函数.

证明：$\mathrm{epi}\, g=\{(x,Y,t)\mid Y\succ 0,\ x^TY^{-1}x\le t\}$

~~$=\left\{(x,Y,t)\ \middle|\ \begin{bmatrix} Y & x \\ x^T & t\end{bmatrix}\succeq 0,\ Y\succ 0\right\}$~~

由 Schur 补条件：$\begin{bmatrix} Y & x \\ x^T & t\end{bmatrix}$ 半正定 iff $\begin{cases} Y\succeq 0 \ \checkmark \\ t-x^TY^{-1}x\ge 0 \ \checkmark \\ (I-YY^{-1})x=0. \end{cases}$

~~[illegible]~~

~~$\therefore \mathrm{epi}\, g$ 为凸集.~~

$\therefore \begin{bmatrix} Y & x \\ x^T & t\end{bmatrix}$ 半正定 $\therefore \mathrm{epi}\, g$ 为凸锥，是凸集

$\therefore g$ 为凸函数，

$\therefore f(X)=X^{-1}$ 在 S^n_{++} 上矩阵凸.