3.3

(法1：利用定理) 函数是凹函数 $\Leftrightarrow$ 其亚图 $\mathrm{hypo}\, f=\{(x,t)\mid t\le f(x)\}$ 是凸集

$\mathrm{hypo}\, g=\{(y,t)\mid t\le g(y)\}=\{(y,t)\mid f(t)\le f(g(y))=y\}=\begin{bmatrix}0&1\\1&0\end{bmatrix}\mathrm{epi}\, f$，故 g 是凹函数

(法2：利用定义)

$f(x)$ 在 (a,b) 上是凸函数 $\Rightarrow$ 对 $x_1,x_2\in(a,b)$，有 $f(\theta x_1+(1-\theta)x_2)\le\theta f(x_1)+(1-\theta)f(x_2)$

其中 $0\le\theta\le 1$

$f(x)$ 在 (a,b) 上递增，g 是其反函数 $\Rightarrow$ g 在 $(f(a),f(b))$ 上递增

$g(\theta f(x_1)+(1-\theta)f(x_2))\ge g(f(\theta x_1+(1-\theta)x_2))=\theta x_1+(1-\theta)x_2=\theta g(f(x_1))+(1-\theta)g(f(x_2))$

故 g 是凹函数

注：凸函数几乎处处二阶可微，进一步[强化条件 g,f 可微]，~~可得~~可得

$g'(f(x))=\dfrac{1}{f'(x)}$，故 g 递增

$g''(f(x))=-\dfrac{f''(x)}{f'(x)^3}$，故 g 是凹函数

3.4

①假设f是凸函数

$$\int_0^1 f(x+\lambda(y-x))d\lambda \leq \int_0^1 (f(x)+\lambda(f(y)-f(x)))d\lambda = \frac{f(x)+f(y)}{2}$$

②假设f不是凸函数

存在x,y和$\theta_0\in(0,1)$, s.t. $f(\theta_0 x+(1-\theta_0)y) > \theta_0 f(x)+(1-\theta_0)f(y)$

~~即~~ $\cancel{f(y+\theta_0(x-y)) > \theta_0 f(x)+(1-\theta_0)f(y)}$

$\cancel{\Rightarrow \int_0^1 f(y+\theta_0(x-y)}$

设$F(\theta)=f(\theta x+(1-\theta)y)-\theta f(x)-(1-\theta)f(y)$, F连续, $\cancel{F(\theta)}$ $F(0)=F(1)=0$, $F(\theta_0)>0$

设α,β为F在$(0,\theta_0)$零最大的零点交叉和在$(\theta_0,1)$最小的零点交叉,则

$F(\theta)=f(\cancel{\theta x}\ \theta x+(1-\theta)y)-\theta f(x)-(1-\theta)f(y)>0$, $\theta\in(\alpha,\beta)$

取$u=\alpha x+(1-\alpha)y$, $v=\beta x+(1-\beta)y$, 对$\theta\in(0,1)$, $f(\theta u+(1-\theta)v)>\theta f(u)+(1-\theta)f(v)$

$\int_0^1 f(u+\theta(u-v))d\theta > \int_0^1 (f(u)+\theta(f(u)-f(v)))d\theta = \frac{f(u)+f(v)}{2}$, 矛盾!

注:不可对θ_0直接积分

3.13

负熵是严格凸且可微的(见3.1.5)，故 $f(u) > f(v) + \nabla f(v)^T(u-v),\ u \neq v$

当 $u=v$ 时，$f(u) = f(v) + \nabla f(v)^T(u-v)$

故 $\forall u,v \in R^n_{++},\ D_{kl}(u,v) \geq 0$

当 $u \neq v$ 时，$D_{kl}(u,v) = \sum_{i=1}^{n}\left(u_i \log\left(\frac{u_i}{v_i}\right) - u_i + v_i\right) > 0$

故 $u=v \Leftrightarrow D_{KL}(u,v) = 0$

3.17

$$\frac{\partial f}{\partial x_i} = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1-p}{p}} x_i^{p-1} = \left(\frac{f(x)}{x_i}\right)^{1-p}$$

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

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

$$y^T \nabla^2 f(x) y = \frac{1-p}{f(x)}\left(\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right)$$

取 $a_i = (f(x)/x_i)^{-\frac{p}{2}},\ b_i = y_i(f(x)/x_i)^{1-\frac{p}{2}},\ \sum_i a_i^2 = 1$,

则 $y^T \nabla^2 f(x) y \leq 0 \Rightarrow \nabla^2 f(x) \preceq 0 \Rightarrow f$ 为凹函数

3.19

(a) $f(x)=\alpha_r(x_{[1]}+x_{[2]}+\cdots+x_{[r]})+(\alpha_{r-1}-\alpha_r)(x_{[1]}+x_{[2]}+\cdots+x_{[r-1]})$

$+(\alpha_{r-2}-\alpha_{r-1})(x_{[1]}+x_{[2]}+\cdots+x_{[r-2]})+\cdots+(\alpha_1-\alpha_2)x_{[1]}$

$f(x)$是凸函数的非负加权求和，故$f(x)$是凸函数

(b) ~~$T(x,w)$是凸的~~ $T(x,w)$是x的线性函数 $\Rightarrow g(x,w)=-\log T(x,w)$是凸的$\Rightarrow f(x)=\int_0^{2\pi}g(x,w)$是凸的

注：保凸运算，见3.2.1

3.23

(a) $f(x,t)=t\cdot\frac{\|x\|_p^p}{t^p}$，设$g(x)=\|x\|_p^p$，则$f(x,t)=t\cdot g(\frac{x}{t})$，故$f(x,t)$是凸的

(b) 令$(y,t)=(Ax+b, c^Tx+d)$，$g(y,t)=\frac{y^Ty}{t}=t\cdot\varphi(\frac{y}{t})$，其中$\varphi(\frac{y}{t})=\frac{\|y\|_2^2}{t^2}$

$\varphi(y)=\|y\|_2^2$是凸的$\Rightarrow g(y,t)$是凸的$\Rightarrow f(x)$是凸的

3.32

(a) 对于 $0\le\theta\le1$,

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

$$\le(\theta f(x)+(1-\theta)f(y))(\theta g(x)+(1-\theta)g(y))$$

$$=\theta f(x)g(x)+(1-\theta)f(y)g(y)$$

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

f与g同时递增或递减

$$\Rightarrow(f(y)-f(x))(g(x)-g(y))\le0$$

则 $f(\theta x+(1-\theta)y)g(\theta x+(1-\theta)y)$

$$\le\theta f(x)g(x)+(1-\theta)f(y)g(y)$$

得证

(b) 与(a)同理，利用Jensen不等式

(c) $\frac{1}{g}$是凸的，非减且大于0，由(a)可知，结论成立

3.47

~~必要性~~: $\cancel{\log f(y)-\log f(x)\le\frac{\nabla f^T(x)(y-x)}{f(x)}}$ (取对数)

$$\cancel{\log f(y)\le\log f(x)+\frac{1}{f(x)}\nabla f^T(x)(y-x)}$$

~~则 log f 为凹函数~~

~~又 f(x)>0，故 f 为对数凹~~

~~充分性~~: $\log f(y)\le\log f(x)+\frac{1}{f(x)}\nabla f^T(x)(y-x)$ (一阶条件

$$\log\frac{f(y)}{f(x)}\le\frac{1}{f(x)}\nabla f^T(x)(y-x)$$

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

3.57

$\forall v\in R^n$，只需证 $g(X)=v^TX^{-1}v$ 在 $X\in S^n_{++}$ 上是凸的

由例3.4可知，结论成立

见例3.48