1.

(a) $y>0$ 时

$\frac{\partial g}{\partial x}=\frac{2x}{y}$ $\quad \frac{\partial g}{\partial y}=-\frac{x^2}{y^2}+1$

$$\text{Hessian}=\begin{bmatrix}\frac{2}{y} & -\frac{2x}{y^2}\\ -\frac{2x}{y^2} & \frac{3x^2}{y^3}\end{bmatrix}=\frac{6x^2}{y^4}-\frac{4x^2}{y^4}\geq 0$$

∵半正定 ∴ $g(x,y)$ 凸函数

(b) $h=\frac{x^2}{y}+y$

$\frac{\partial h}{\partial y}=-\frac{x^2}{y^2}+1=0 \quad y^2=x^2 \Rightarrow y=|x|$

代入 h 得 $\frac{x^2}{|x|}+|x|=2|x|$

(c) 由(b) $\min\limits_{\sigma_i\geq 0} g(\beta_i,\sigma_i)=2|\beta_i|$

∴ $\min\limits_{\beta} f(\beta)+\frac{\lambda}{2}\sum\limits_{i=1}^{n}2|\beta_i|=\min\limits_{\beta} f(\beta)+\lambda\|\beta\|_1$

2

a.

i→ii

$\|\nabla f(x)-\nabla f(y)\|_2 \le L\|x-y\|_2$

由柯西不等式得：$(\nabla f(x)-\nabla f(y))^T(x-y) \le \|\nabla f(x)-\nabla f(y)\|_2\|x-y\|_2$

$\therefore (\nabla f(x)-\nabla f(y))^T(x-y) \le L\|x-y\|_2^2$

ii→iii 考虑 f 在 x 处二阶近似 $y=x+sh$

则 $\nabla f(x+sh) \approx \nabla f(x)+s\nabla^2 f(x)h$

代入 x 和 $y=x+sh$ 得 $(\nabla f(x)-\nabla f(x+sh))^T(-sh) \le Ls^2$ 代入

得 $s^2 h^T\nabla^2 f(x)h \le Ls^2$ 得 $h^T\nabla^2 f(x)h \le L$

表明对 $\forall$ 单位向量 h，$\nabla^2 f(x)$ 在该方向上二次形式不超过 L

iii→iv $\forall$ 方向 d $d^T\nabla^2 f(x)d \le L\|d\|_2^2$ $\therefore (y-x)^T\nabla^2 f(s)(y-x) \le L\|y-x\|^2$

$f(y)=f(x)+\nabla f(x)^T(y-x)+\frac{1}{2}(y-x)^T\nabla^2 f(s)(y-x)$

$\therefore f(y) \le f(x)+\nabla f(x)^T(y-x)+\frac{L}{2}\|y-x\|_2^2$

iv→i $\forall x, y$

$f(y) \le f(x)+\nabla f(x)^T(y-x)+\frac{L}{2}\|y-x\|_2^2$

令 $g(t)=f(x+t(y-x))$ $t\in[0,1]$ 则 $g'(t)=\nabla f(x+t(y-x))^T(y-x)$

再对 $g'(t)$ 求导 得 $g''(t)=(y-x)^T\nabla^2 f(x+t(y-x))(y-x)$

$\because g''(t) \le L\|y-x\|^2$ $\therefore g'(1)-g'(0)=\int_0^1 g''(t)dt \le L\|y-x\|_2^2$

$\therefore (\nabla f(y)-\nabla f(x))^T(y-x) \le L\|y-x\|_2^2$ 由 Cauchy 得

$\therefore \|\nabla f(x)-\nabla f(y)\|_2 \le L\|x-y\|_2$

b

i→ii 若 f 是 m-强凸函数，对 $\forall x, y$，有 $f(y) \ge f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2$

令 $g(x) = f(x) - \frac{m}{2}\|x\|_2^2$，则 $g(x)$ 是一个凸函数

i) $g(y) \ge g(x) + \nabla g(x)^T(y-x)$ $g(x)$ 代入

$(\nabla f(x) - \nabla f(y))^T(x-y) \ge m\|x-y\|_2^2$

ii→iii 设 $y = x+sh$ $\nabla f(x+sh) \approx \nabla f(x) + s\nabla^2 f(x)h$

将 x 和 $y = x+sh$ 代入 ii 的不等式，我们得到

$(\nabla f(x) - \nabla f(x+sh))^T(-sh) \ge ms^2$

代入 $\nabla f(x+sh) \approx \nabla f(x) + s\nabla^2 f(x)h$

得 $s^2 h^T \nabla^2 f(x) h \ge ms^2$ 得 $h^T \nabla^2 f(x) h \ge m$

iii→iv

$d^T \nabla^2 f(x) d \ge m\|d\|_2^2$

$f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)$

$(y-x)^T \nabla^2 f(z)(y-x) \ge m\|y-x\|_2^2$

$\therefore f(y) \ge f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2$

iv→i iv 表明 f 是 m 强凸