Convex optimization HW3.

1、a. $\nabla \|x\|_p = \|x\|_p^{1-p} \cdot \begin{bmatrix} x_1^{p-1} \\ \vdots \\ x_n^{p-1} \end{bmatrix}$. 记 $\vec{x}^p = \begin{bmatrix} x_1^p \\ \vdots \\ x_n^p \end{bmatrix}$.

记 $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \odot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 \\ \vdots \\ a_n b_n \end{bmatrix}$. $\odot$ 为逐元素相乘.

在一阶近似下，有:

$$\begin{bmatrix} (x_1+\Delta x_1)^{p-1} \\ \vdots \\ (x_n+\Delta x_n)^{p-1} \end{bmatrix} - \begin{bmatrix} x_1^{p-1} \\ \vdots \\ x_n^{p-1} \end{bmatrix} \approx (p-1)\vec{x}^{p-2} \odot \Delta\vec{x}, \qquad \nabla \|x\|_p^{1-p} = (1-p)\|x\|_p^{1-2p} \cdot \vec{x}^{p-1}$$

$$= \mathrm{diag}\left((p-1)x^{p-2}\right) \cdot \Delta\vec{x}$$

故: $\Delta(\nabla\|x\|_p) = \nabla\|x+\Delta x\|_p - \nabla\|x\|_p$

Hessian.

$$\approx \boxed{\left(\|x_p\|^{1-p} \mathrm{diag}\left((p-1)\vec{x}^{p-2}\right) + (1-p)\|x\|_p^{1-2p} \cdot (\vec{x}^{p-1})(\vec{x}^{p-1})^T \right)} \Delta x$$

$= \nabla^2 \|x\|_p \cdot \Delta x$ 即得 Hessian.

b. $f(x+\Delta x) = (a^T(x+\Delta x))(b^T(x+\Delta x))$

$$= (a^T x + a^T \Delta x)(b^T x + b^T \Delta x)$$

$$= (a^T x)(b^T x) + ((a^T x) b^T + (b^T x) a^T) \Delta x + (a^T \Delta x)(b^T \Delta x)$$

$$= f(x) + [(a^T x) b + (b^T x) a] \Delta x + \frac{1}{2} \Delta x^T (ab^T + ba^T) \Delta x$$

故: $\nabla^2 f = ab^T + ba^T$

c. $\nabla f = A^T(Ax - b)$

而 $\nabla f(x+\Delta x) = A^T(Ax + A\Delta x - b)$

$$= \nabla f(x) + A^T A \Delta x$$

故: $\nabla^2 f = A^T A$.

d. $\nabla f = R^T \begin{bmatrix} u_1 g'(y_1) \\ \vdots \\ u_n g'(y_n) \end{bmatrix}$，其中 $\vec{y} = R\vec{x}$

$= R^T(\vec{u} \odot g'(\vec{y}))$，其中 $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \odot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \equiv \begin{bmatrix} a_1 b_1 \\ \vdots \\ a_n b_n \end{bmatrix}$，为 element-wise multiplication

则 $\Delta(\nabla f) = \nabla f(\vec{x} + \Delta\vec{x}) - \nabla f(\vec{x})$

$\cancel{= R^T(\vec{u} \odot g(R\Delta\vec{x}))}$ （一阶近似）

~~而~~ $\cancel{g(R\Delta\vec{x}) \simeq R\, g(\Delta\vec{x})}$

$= R^T\left(\vec{u} \odot \left(g'(R(x+\Delta\vec{x})) - g'(Rx)\right)\right)$

而 $g'(R(x+\Delta\vec{x})) - g'(Rx) \simeq g''(Rx) \odot R\Delta\vec{x}$.

故 $\Delta(\nabla f) \simeq R^T(u \odot g'(Rx) \odot R\Delta\vec{x})$

$= R^T \operatorname{diag}(u \odot g''(Rx)) R \cdot \Delta\vec{x}$

$\therefore \nabla^2 f = R^T \operatorname{diag}(u \odot g''(Rx)) R$