Homework 1
Shengchao Liu

1. (a) Is polyhedra.

$$\{x| \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} x \geq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & \dots 1 \\ a_1 & a_2 & \dots a_n \\ a_1^2 & a_2^2 & \dots a_n^2 \end{bmatrix} x = \begin{bmatrix} 1 \\ b \\ c \end{bmatrix} \}$$

   (b) Not polyhedra.

   (c) Is polyhedra.

   $x^T y \leq 1$ for all y with $\|y\|^2 = 1$, this implies that $\max_i |x_i| \leq 1$

   Proof:

   If $\max_i |x_i| > 1$ when $i = k$, then we can find y with $\|y\|^2 = 1$ and $y_k = 1$ if $x_k = 1$, $y_k = -1$ if $x_k = -1$, and $\underset{i \neq k}{y_i} = 0$.

   Then $x^T = \sum x_i y_i = x_k y_k > 1$, contradicts with assumption.

   So $\underset{i}{|x_i|} \leq 1 \Longrightarrow 1 \succeq x \succeq -1$

$$\{x| \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{bmatrix} x \geq \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} \}$$

2. According to definition, x^* is isolated local minima if there is a neighbourhood N of x^* such that x^* is the only local minima in N.

   Assume x^* is not a strict local minima, which means we can find x^k such that both x^* and x^k are local minima, and $x^k \in N$. And this contradicts with the definition of x^* being the only local minima in N.

   $\therefore$ all isolated local minima are strict.

3. (a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$, the eigenvalues are -3, and 1. So it is not positive definite.

   (b) Yes.

   If $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ is positive semi definite, according to definition, $x^T A x \geq 0, \forall x = \begin{bmatrix} x_1, x_2, \dots x_n \end{bmatrix}$.

We can get $x^T Ax = a_{11}x_1^2+a_{22}x_2^2+\dots a_{nn}x_n^2+a_{12}x_1x_2+a_{13}x_1x_3+\dots a_{(n-1)n}x_{n-1}x_n$. Assume $\exists i \in [1,n], a_{ii} < 0$. Then we can set $x = \begin{bmatrix}0,0,\dots,x_i,\dots,0\end{bmatrix}, x_i \neq 0$. Then $x^T Ax = a_{ii}x_i^2 < 0$, which contradicts $x^T Ax \geq 0$.

4. $f(x) = 100(x_2 - x_1^2)^2 + (1-x_1)^2$

$$\triangledown f(x) = \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{bmatrix} = \begin{bmatrix} 400x_1^3 - 400x_2x_1 + 2x_1 - 2 \\ 200x_2 - 200x_1^2 \end{bmatrix}$$

$$\triangledown^2 f(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

Suppose x^* is local minimizer. Then we can get $\triangledown f(x^*) = 0$ and $\triangledown^2 f(x^*) \succeq 0$.

$$\begin{bmatrix} 400x_1^{*3} - 400x_2^*x_1^* + 2x_1^* - 2 \\ 200x_2^* - 200x_1^{*2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\implies x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\therefore\ \triangledown^2 f(x) = \begin{bmatrix} 1200 - 400 + 2 & -400 \\ -400 & 200 \end{bmatrix} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$, and eigenvalues are $501 \pm \sqrt{501^2 - 20^2} > 0$, so Hessian at x^* is positive definite.

5. $f(x) = (x_1 + x_2)^2$

$$-\triangledown f(x) = -\begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

Let θ be the angle between p and $-\triangledown f(x)$. So $\cos\theta = \frac{\sqrt{2}}{2} > 0$, which means p has an angle strictly less than $\frac{\pi}{2}$ with $-\triangledown f(x)$, and p is a descent direction.

$$\min_{\alpha>0} f(x+\alpha p) = \min_{\alpha>0} f(\begin{bmatrix} 1-\alpha \\ \alpha \end{bmatrix}) = \min_{\alpha>0}(1-\alpha+\alpha^2)^2$$

So when $\alpha = \frac{1}{2}$, we can get minima $\frac{9}{16}$.

6. $\tilde{f}(z) = f(x) = f(Sz + s)$

$$\triangledown \tilde{f}(z) = \frac{d\tilde{f}}{dz} = \begin{bmatrix} \frac{d\tilde{f}}{dz_1} \\ \frac{d\tilde{f}}{dz_2} \\ \vdots \end{bmatrix}$$

$$\frac{d\tilde{f}}{dz_j} = \sum_i \frac{df}{dx_i}\frac{dx_i}{dz_j} = \sum_i S_{ij}\frac{df}{dx_i}$$

$$\triangledown \tilde{f}(z) = \frac{d\tilde{f}}{dz} = \begin{bmatrix} \frac{d\tilde{f}}{dz_1} \\ \frac{d\tilde{f}}{dz_2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{df}{dx_1}S_{11} + \frac{df}{dx_2}S_{21} + \dots \\ \frac{df}{dx_1}S_{12} + \frac{df}{dx_2}S_{22} + \dots \\ \vdots \end{bmatrix} = \begin{bmatrix} S_{11} & S_{21} & \dots \\ S_{12} & S_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \\ \vdots \end{bmatrix} = S^T \triangledown f(x)$$

Similarly, we can get:

$$\frac{d^2}{z_j} = \sum_i \frac{d}{dx_i}(\sum_k \frac{df}{dx_k}\frac{dx_k}{dz_j})\frac{dx_i}{dz_j} = \sum_i\sum_k S_{ij}S_{kj}\frac{d^2 f}{dx_i dx_k} = \begin{bmatrix} S_{1j} & S_{2j} & \dots \end{bmatrix} \begin{bmatrix} \sum_k S_{kj}\frac{d^2}{dx_1 dx_k} \\ \sum_k S_{kj}\frac{d^2}{dx_2 dx_k} \\ \vdots \end{bmatrix} =$$

$$\begin{bmatrix} S_{1j} & S_{2j} & \dots \end{bmatrix} \begin{bmatrix} \frac{d^2}{dx_1 dx_1} & \frac{d^2}{dx_1 dx_2} & \dots \\ \frac{d^2}{dx_2 dx_1} & \frac{d^2}{dx_2 dx_2} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} S_{1j} \\ S_{2j} \\ \vdots \end{bmatrix}$$

So $\nabla^2 \tilde{f}(z) = S^T \nabla^2 f(x) S$

7. $f(x) = x_1^2 + x_2^2 + \beta x_1 x_2 + x_1 + 2x_2$

   $\nabla f(x) = \begin{bmatrix} 2x_1 + \beta x_2 + 1 \\ 2x_2 + \beta x_1 + 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

   $f(x) = x^T \begin{bmatrix} 2 & \beta \\ \beta & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} x$

   $\begin{cases} \begin{bmatrix} \frac{4-4\beta}{\beta^2-4} \\ \frac{4-\beta}{\beta^2-4} \end{bmatrix}, & \beta \neq \pm 2 \\ \text{no solution}, & \beta = \pm 2 \end{cases}$

   For local minimizer, $\nabla^2 f(x) = \begin{bmatrix} 2 & \beta \\ \beta & 2 \end{bmatrix} \succeq 0$, and eigenvalues are $2 \pm \beta \geq 0$. That is $-2 \leq \beta \leq 2$. Combined with condition $\nabla f(x) = 0$, we have $-2 < \beta < 2$.

   And since Q is positive definite, then this function is convex, so these points are global minimizer. That is $\begin{bmatrix} \frac{4-4\beta}{\beta^2-4} \\ \frac{4-\beta}{\beta^2-4} \end{bmatrix}, -2 < \beta < 2$.

8. Sufficient: f is convex $\Longrightarrow \Omega$ is convex and $\nabla^2 f(x) \succeq 0, \forall x \in \Omega$.

   Prove Ω is convex:

   Because f is convex, $f((1-\alpha)x + \alpha y) = (1-\alpha)f(x) + \alpha f(y)$, for $\forall x, y \in \Omega$, which implies that $(1-\alpha)x + \alpha y \in \Omega$, so Ω is convex.

   Prove $\nabla^2 f(x) \succeq 0, \forall x \in \Omega$:

   Because f is convex, we can get $f(y) \geq f(x) + \nabla f(x)^T(y - x), \forall x, y \in \Omega$.

   And let $y = x + p$, we can get $f(x + p) \geq f(x) + \nabla f(x)^T p, \forall x, p \in \Omega$.

   According to Taylor's Thrm,

   $f(x + p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x + tp)p, \exists t \in (0, 1)$.

   So $\forall x, p \in \Omega$, $\exists t \in (0, 1)$, $\frac{1}{2} p^T \nabla^2 f(x + tp)p \geq 0 \Longrightarrow \nabla^2 f(x) \succeq 0$.

Necessary: Ω is convex and $\nabla^2 f(x) \succeq 0, \forall x \in \Omega \implies f$ is convex.

According to Taylor's Thrm:

$f(x) = f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}(x-y)^T \nabla^2 f(y + t(x-y))^T (x-y)$, $\exists t \in (0,1)$

And $\nabla^2 f(x)$ is positive semi-definite, then $f(x) \geq f(y) + \nabla f(y)^T(x-y)$.

Then we set $z = \lambda y + (1-\lambda)x$, $\forall \lambda \in [0,1]$.

$f(y) \geq f(z) + \nabla f(z)^T(y-z) = f(z) + \nabla f(z)^T(1-\lambda)(y-x)$

$f(x) \geq f(z) + \nabla f(z)^T(x-z) = f(z) + \nabla f(z)^T(-\lambda)(y-x)$

$\implies$

$\lambda f(y) \geq \lambda f(z) + \lambda(1-\lambda)\nabla f(z)^T(y-x)$

$(1-\lambda) f(x) \geq (1-\lambda) f(z) - \lambda(1-\lambda)\nabla f(z)^T(y-x)$

$\implies$

$\lambda f(y) + (1-\lambda) f(x) \geq f(\lambda y + (1-\lambda)x)$

So f is convex.

9. (a) $\nabla f(x) = \begin{bmatrix} 4x_1^3 - 16x_1 \\ 2x_2 \end{bmatrix} = 0$

$\implies x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}$

$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 - 16 & 0 \\ 0 & 2 \end{bmatrix}$

When we take $x = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 32 & 0 \\ 0 & 2 \end{bmatrix} \succeq 0$, so both are local minima.

And $x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} -12 & 0 \\ 0 & 2 \end{bmatrix} \not\succeq 0$, so this is a saddle point.

(b) $\nabla f(x) = \begin{bmatrix} x_1 + \cos x_2 \\ -x_1 \sin x_2 \end{bmatrix} = 0$

$\implies x = \begin{bmatrix} 0 \\ \frac{\pi}{2} + k\pi \end{bmatrix}, \begin{bmatrix} 1 \\ \pi + 2k\pi \end{bmatrix}, \begin{bmatrix} -1 \\ 2\pi + 2k\pi \end{bmatrix}$

$\nabla^2 f(x) = \begin{bmatrix} 1 & -\sin x_2 \\ -\sin x_2 & -x_1 \cos x_2 \end{bmatrix} \succeq 0$

$x = \begin{bmatrix} 0 \\ \frac{\pi}{2} + 2k\pi \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix} \not\succeq 0$, not local minima.

$x = \begin{bmatrix} 1 \\ \pi + 2k\pi \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succeq 0$, is local minima.

$x = \begin{bmatrix} 0 \\ \frac{3\pi}{2} + 2k\pi \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \not\succeq 0$, not local minima.

$x = \begin{bmatrix} -1 \\ 2\pi + 2k\pi \end{bmatrix}$, $\nabla^2 f(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succeq 0$, is local minima.

(c) $\nabla f(x) = \begin{bmatrix} -4x_1(x_2 - x_1^2) - 2x_1 \\ 2(x_2 - x_1^2) \end{bmatrix} = 0$

$\implies x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 - 4x_2 - 2 & -4x_1 \\ -4x_1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \not\succeq 0$. So this is a saddle point.