Homework 8

3.5.1 a) $f(x) = \ln(1+x^2)$

$$\frac{d}{dx} f(x) = \frac{1}{1+x^2}(2x) \quad \therefore \quad \frac{d^2 f}{dx^2} = \frac{(1+x^2)2 - 2x(2x)}{(1+x^2)^2} = \frac{2-2x^2}{(1+x^2)^2}$$

for $x = 2$ for instance $\frac{d^2f}{dx^2} = \frac{2-8}{(1+5)^2} < 0 \quad \therefore \; Hf(x)$ not psd when $x=2$

$\Rightarrow f$ not convex

c) $f(x) = |x^3| = \begin{cases} x^3 & \text{if } x \geq 0 \\ -x^3 & \text{if } x < 0 \end{cases} \qquad \frac{d^2f}{dx^2} = \begin{cases} 6x & x \geq 0 \\ -6x & x < 0 \end{cases}$

If $x \geq 0$ $6x \geq 0$, If $x < 0$ $-6x > 0$

$\therefore Hf(x)$ psd $\forall x \in \mathbb{R}$ $\Rightarrow f$ convex

e) $f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{3}{2}x_2^2 + \sqrt{3}x_1x_2 \qquad \frac{\partial f}{\partial x_1} = x_1 + \sqrt{3}x_2$

$\Rightarrow \quad Hf = \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$ psd $\therefore f$ convex $\qquad \frac{\partial f}{\partial x_2} = 3x_2 + \sqrt{3}x_1$

Note x_1x_2 is not convex $\therefore$ you cannot use the fact that a non-negative linear comb^n of convex functions is convex

g) $f(x) = \|Bx - d\|^2 \qquad B \in \mathbb{R}^{m\times n} \qquad d \in \mathbb{R}^m$

$= \sum_{i=1}^{m} (B_i x - d_i)^2$ where B_i = i^{th} row of B and d_i = i^{th} component of d.

Let $h(x) = \sum_{i=1}^{n} x_i^2$. Then $f(x) = h(Bx - d)$

$\therefore$ By property ③ its enough to show h is convex.

$Hh(x) = \begin{bmatrix} 2 & & 0 \\ & 2 & \\ & \ddots & \\ 0 & & 2 \end{bmatrix}$ which is psd $\forall x \in \mathbb{R}^n$. $\therefore h$ is convex

$\Rightarrow f$ is convex.

3.5.3 $f: \mathbb{R}^n \to \mathbb{R}$ convex function $\frac{\partial f}{\partial x_i}$ conts on $\mathbb{R}^n$.

Prove $\nabla f(x^*) = 0 \iff f(x^*) \le f(x) \quad \forall\, x \in \mathbb{R}^n$.

Proof "$\Rightarrow$" Suppose $\nabla f(x^*) = 0$. Since f is convex by Thm 3.5.1 (a) we have that $\forall\, x, y \in \mathbb{R}^n$,

$$f(x) + \nabla f(x)^T (y - x) \le f(y)$$

Taking $x = x^*$ and $y = x$ we get

$$f(x^*) + \nabla f(x^*)^T (x - x^*) \le f(x) \quad \forall\, x \in \mathbb{R}^n$$

$$\Rightarrow f(x^*) \le f(x) \quad \forall\, x \in \mathbb{R}^n \quad \text{since } \nabla f(x^*) = 0.$$

"$\Leftarrow$" Suppose $f(x^*) \le f(x) \quad \forall\, x \in \mathbb{R}^n$. Then x^* is a local opt of f and by Thm 3.3.2 $\nabla f(x^*) = 0$.
(Note: you don't need f to be convex for this)

3.5.5 $g_1, \dots, g_\ell$ convex functions, $h_1, \dots, h_m$ affine functions.
$D = \{x \in \mathbb{R}^n : g_1(x) \le 0, \dots, g_\ell(x) \le 0, h_1(x) \le 0, \dots, h_m(x) \le 0\}$

a) Show D is convex: Let p, q be 2 pts in D and $0 \le t \le 1$.
Need to show that $tp + (1-t)q \in D$.

$p, q \in D \Rightarrow g_i(p) \le 0 \quad g_i(q) \le 0 \quad \forall\, i = 1, \dots, \ell$
$h_i(p) = 0 \quad h_i(q) = 0 \quad \forall\, i = 1, \dots, m$

$\therefore g_i(tp + (1-t)q) \le t g_i(p) + (1-t) g_i(q)$ since g_i convex
≤ 0 since $t, 1-t \ge 0$, $g_i(p), g_i(q) \le 0$

$h_i(tp + (1-t)q) = t h_i(p) + (1-t) h_i(q) = 0$
since h_i affine

$\therefore tp + (1-t)q \in D$.

$$D_1 = \left\{ (x_1, x_2, x_3) : \begin{array}{l} x_1^2 + x_2^2 - 9 \le 0 \\ x_2 - 2x_3 - 3 \le 0 \end{array} \right\} \qquad \begin{array}{l} g_1 = x_1^2 + x_2^2 - 9 \\ g_2 = x_2 - 2x_3 - 3 \end{array}$$

g_1 convex since $Hg_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ psd

g_2 convex since it's affine

$\therefore D_1$ is convex.

b) f convex. Suppose $\exists\ x^*, x^{**} \in D$ s.t $f(x^*) = f(x^{**}) \le f(x)\ \forall x \in D$.
Then since D is convex by a) all pts on the line segment $[x^*, x^{**}]$ is in D.
A pt on this segment has the form $t x^* + (1-t) x^{**}$ $0 \le t \le 1$.
Since f is convex,
$f(tx^* + (1-t)x^{**}) \le t f(x^*) + (1-t) f(x^{**}) = f(x^*)$ (since $f(x^*) = f(x^{**})$)
$\Rightarrow f(tx^* + (1-t)x^{**}) = f(x^*)$ since x^* is a global opt.
$\therefore$ Every pt on $[x^*, x^{**}]$ is a global opt of f.

3.5.6 $(3.4.3) = (2.4.1)$

$$\begin{array}{llll} \min & x_1 + x_2^2 & & \\ \text{s.t} & 3x_1 + 2x_2 = 3 & & \\ & x_1 \ge 0 & & \\ & x_2 \ge 0 & & \end{array} \quad \overset{\text{std form}}{\rightsquigarrow} \quad \begin{array}{ll} \min & x_1 + x_2^2 \\ \text{s.t} & -x_1 \le 0 \\ & -x_2 \le 0 \\ & 3x_1 + 2x_2 = 3 \end{array}$$

$g_1 = -x_1$ $g_2 = -x_2$ both convex since they are linear
$h_1 = 3x_1 + 2x_2 - 3$ affine.
$f = x_1 + x_2^2$ is convex since $Hf = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$ psd.
$\therefore$ This is a convex program.
By KKT calculation we got $(x_1 = 7/9,\ x_2 = 1/3,\ \mu_1 = -\frac{1}{3},\ \lambda_1 = \lambda_2 = 0)$ as a sol^n of KKT cond^{n}s + feasibility.
$\therefore (7/9, 1/3)$ is global opt.

3 more problems to do.

3.5.7 $\min f(\underline{x}) = \sum_{i=1}^{n} c_i x_i$ f convex

s.t $x_1^2 + x_2^2 + \dots + x_n^2 \le b$ $g_1 = \sum x_i^2 - b$ also convex
$\therefore$ This is a convex program so Thm 3.5.2 applies.

$\nabla f = c$ $\nabla g_1 = (2x_1, 2x_2, \dots, 2x_n)^t$

a) $\therefore$ KKT condition $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} + \lambda^* \begin{pmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{pmatrix} = 0$ $\lambda^* \ge 0$
$\lambda^* g(x^*) = 0$

b) KKT $\Rightarrow$ $c_i + 2\lambda^* x_i = 0$ $\Rightarrow$ $x_i = \frac{-c_i}{2\lambda^*}$ $\forall\, i = 1, \ldots, n$

feasibility $\Rightarrow$ $\sum x_i^2 \le b$

$\Leftrightarrow$ $\sum \frac{c_i^2}{4(\lambda^*)^2} \le b$ $\Leftrightarrow$ $\sum c_i^2 \le 4(\lambda^*)^2 b$

Case 1: $\lambda^* = 0$ gives $c = 0$. If $c = 0$ all feasible sol$^{\underline{n}}$s are optimal.

Case 2: $g(x^*) = 0$ $\Leftrightarrow$ $\sum x_i^2 = b$ $\therefore$ $\sum c_i^2 = 4(\lambda^*)^2 b$

$$\Rightarrow (\lambda^*)^2 = \frac{\sum c_i^2}{4b} \Rightarrow \lambda^* = \frac{\sqrt{\sum c_i^2}}{2\sqrt{b}}$$

$$\therefore \quad x_i = \frac{-c_i}{\cancel{2}\sqrt{\sum c_i^2}} \cdot \cancel{2}\sqrt{b} = \frac{-\sqrt{b}\, c_i}{\sqrt{\sum c_i^2}}$$

$\therefore$ Global opt $x^* = \left(\frac{-\sqrt{b}\, c_i}{\sqrt{\sum c_i^2}} \right)$ //