HW3 M529 - Spring 14

Solutions.

1. a) Yes, $f''(x) = e^x > 0$, so e^x is convex.
   b) Yes, the epigraph of $|x^3|$ is a convex set.
   c) No, the epigraph of $\ln x$ is not a convex set.

2. See attachment.

3. $f(x) = x^2 + x$

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b) \ ?$$

$$(\lambda a + (1-\lambda)b)^2 + (\lambda a + (1-\lambda)b) \leq \lambda(a^2 + a) + (1-\lambda)(b^2 + b) \ ?$$

$$\underbrace{\lambda^2 a^2 + 2\lambda a(1-\lambda)b + (1-\lambda)^2 b^2 + \lambda a + (1-\lambda)b}_{A} \leq \underbrace{\lambda a^2 + \lambda a + (1-\lambda)b^2 + (1-\lambda)b}_{B} \ ?$$

$$B - A \geq 0 \ ?$$

$$\lambda a^2 + \lambda a + (1-\lambda)b^2 + (1-\lambda)b - \lambda^2 a^2 - 2\lambda(1-\lambda)ab - (1-\lambda)^2 b^2 - \lambda a - (1-\lambda)b$$

$$\lambda a^2 + \lambda a + b^2 - \lambda b^2 + b - \lambda b - \lambda^2 a^2 - 2\lambda ab + 2\lambda^2 ab - (1 - 2\lambda + \lambda^2)b^2 - \lambda a - b + \lambda b$$

Grouping in terms of λ:

$$\lambda(a^2+a-b^2-b-2ab+2b^2-a+b) - \lambda^2(a^2-2ab+b^2) =$$

$$\lambda(a^2-2ab+b^2) - \lambda^2(a^2-2ab+b^2) =$$

$$\lambda(a-b)^2 - \lambda^2(a-b)^2 =$$

$\lambda(1-\lambda)(a-b)^2$ which is clearly greater than or equal to zero if $0 \leq \lambda \leq 1$.

4. Using the results of question 2: We see that e^x is convex, $3x_1 - x_2$ (a plane) is also convex and $x_1^2 + x_2^2$ (a paraboloid) is convex as well. Thus by composition $e^{3x_1-x_2}$ is convex and $e^{x_1^2+x_2^2}$ is convex. Finally since the addition of convex functions is convex, the function

$$e^{3x_1-x_1} + e^{x_1^2+x_2^2} \text{ is convex.}$$

5. } See code on my website.
6. }

7. $\min\ x_1 + x_2$ subject $x_1^2 + x_2^2 = 1$

From the constraint: $x_2^2 = 1 - x_1^2$

$$x_2 = \pm\sqrt{1 - x_1^2}$$

For: $x_2 = +\sqrt{1-x_1^2}$

$$\min\ x_1 + \sqrt{1-x_1^2} = f(x_1)$$

$$f'(x_1) = 1 + \frac{1}{2}(1-x_1^2)^{-\frac{1}{2}}(-2x_1) = 0$$

$$= 1 - \frac{x_1}{\sqrt{1-x_1^2}} = 0$$

$$\frac{x_1}{\sqrt{1-x_1^2}} = 1 \Rightarrow x_1 = \sqrt{1-x_1^2}$$

$$x_1^2 = 1 - x_1^2$$

$$2x_1^2 = 1$$

$$x_1^2 = \frac{1}{2} \Rightarrow x_1 = \pm\frac{1}{\sqrt{2}}$$

Solutions:

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

From these solutions, the point

$\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ is the minimizer.

For $x_2 = -\sqrt{1 - x_1^2}$

$$\min \; x_1 - \sqrt{1-x_1^2} = f(x_1)$$

$$f'(x_1) = 1 - \frac{1}{2}(1-x_1^2)^{-\frac{1}{2}}(-2x_1) = 0$$

$$1 + \frac{x_1}{\sqrt{1-x_1^2}} = 0 \Rightarrow \frac{x_1}{\sqrt{1-x_1^2}} = -1$$

$$x_1 = -\sqrt{1-x_1^2}$$

$$x_1^2 = 1 - x_1^2 \Rightarrow x_1 = \pm\frac{1}{\sqrt{2}}$$

Solutions:

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

From these solutions, the point $\left(\frac{-1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ is the minimizer.

Clearly, the actual minimizer is $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ so, if we use $x_2 = +\sqrt{1-x_1^2}$, we miss the actual solution!!

8. $L = x + 2y + \lambda_1(1 - x - y - z) + \lambda_2(4 - y^2 - z^2)$

$$\nabla L = \begin{pmatrix} 1 - \lambda_1 \\ 2 - \lambda_1 - 2\lambda_2 y \\ -\lambda_1 - 2\lambda_2 z \\ 1 - x - y - z \\ 4 - y^2 - z^2 \end{pmatrix} = 0$$

From (1): $\lambda_1 = 1$

In (2) & (3): $2 - 1 - 2\lambda_2 y = 0$

$$1 - 2\lambda_2 y = 0 \Rightarrow \lambda_2 = \frac{1}{2y}$$

$$-1 - 2\lambda_2 z = 0 \Rightarrow \lambda_2 = \frac{-1}{2z}$$

$$\therefore \frac{1}{2y} = -\frac{1}{2z} \Rightarrow 2z = -2y$$

$$z = -y \quad (*)$$

(*) in (5):

$$4 - y^2 - (-y)^2 = 0$$

$$4 - 2y^2 = 0 \Rightarrow y^2 = 2 \Rightarrow y = \pm\sqrt{2}$$

$$\therefore z = \mp\sqrt{2}$$

In (4):

$$x = 1 - y - z = 1 - y - (-y) = \underline{1}$$

Points: $\underbrace{(1, \sqrt{2}, -\sqrt{2})}_{\text{maximizer}}$ and $\underbrace{(1, -\sqrt{2}, \sqrt{2})}_{\text{minimizer}}$

$f(1, \sqrt{2}, -\sqrt{2}) = 1 + 2\sqrt{2}$ $\qquad$ $f(1, -\sqrt{2}, \sqrt{2}) = 1 - 2\sqrt{2}$

9. $L = yz + xy + \lambda_1(1 - xy) + \lambda_2(1 - x^2 - z^2)$

$$\nabla L = \begin{pmatrix} y - \lambda_1 y - 2\lambda_2 x \\ z + x - \lambda_1 x \\ y - 2\lambda_2 z \\ 1 - xy \\ 1 - x^2 - z^2 \end{pmatrix} = 0$$

This system of equations does not have solutions, so there are no maxima or minima for this problem

10. $\min \sqrt{x^2 + y^2 + z^2}$ (distance to the origin)

s.t. $x + y + z = 1$

$2x - y - 3z = 3$

Since the minimum of $\sqrt{x^2+y^2+z^2}$ and $x^2+y^2+z^2$, we can use $f(x,y,z) = x^2 + y^2 + z^2$ as objective function.

$L = x^2 + y^2 + z^2 + \lambda_1(1 - x - y - z) + \lambda_2(3 - 2x + y + 3z)$

$$\nabla L = \begin{pmatrix} 2x - \lambda_1 - 2\lambda_2 \\ 2y - \lambda_1 + \lambda_2 \\ 2z - \lambda_1 + 3\lambda_2 \\ 1 - x - y - z \\ 3 - 2x + y + 3z \end{pmatrix} = 0$$

$$x = \frac{\lambda_1 + 2\lambda_2}{2} \; ; \; y = \frac{\lambda_1 - \lambda_2}{2} \; ; \; z = \frac{\lambda_1 - 3\lambda_2}{2}$$

So

$$1 - \left(\frac{\lambda_1 + 2\lambda_2}{2}\right) - \left(\frac{\lambda_1 - \lambda_2}{2}\right) - \left(\frac{\lambda_1 - 3\lambda_2}{2}\right) = 0$$

$$2 - \lambda_1 \quad 2\lambda_2 - \lambda_1 + \lambda_2 - \lambda_1 + 3\lambda_2 = 0$$

$$2 - 3\lambda_1 + 6\lambda_2 = 0 \Rightarrow \lambda_1 = \frac{2 + 6\lambda_2}{3}$$

$$3 - 2\left(\frac{\lambda_1 + 2\lambda_2}{2}\right) + \left(\frac{\lambda_1 - \lambda_2}{2}\right) + 3\left(\frac{\lambda_1 - 3\lambda_2}{2}\right) = 0$$

$$6 - 2\lambda_1 - 4\lambda_2 + \lambda_1 - \lambda_2 + 3\lambda_1 - 9\lambda_2 = 0$$

$$6 + 2\lambda_1 - 14\lambda_2 = 0$$

$$6 + 2\left(\frac{2 + 6\lambda_2}{3}\right) - 14\lambda_2 = 0$$

$$6 + \frac{4}{3} + 4\lambda_2 - 14\lambda_2 = 0$$

$$-10\lambda_2 + \frac{18 + 4}{3} = 0$$

$$10\lambda_2 = \frac{22}{3} \Rightarrow \lambda_2 = \frac{22}{30}$$

$$\lambda_1 = \frac{2 + 6\left(\frac{22}{30}\right)}{3} = \frac{\frac{60 + 132}{30}}{3} = \frac{\overset{64}{\cancel{192}}}{\underset{30}{\cancel{90}}}$$

$$x = \frac{\lambda_1 + 2\lambda_2}{2} \; ; \quad y = \frac{\lambda_1 - \lambda_2}{2} \; ; \quad z = \frac{\lambda_1 - 3\lambda_2}{2}$$

So

$$1 - \left(\frac{\lambda_1 + 2\lambda_2}{2}\right) - \left(\frac{\lambda_1 - \lambda_2}{2}\right) - \left(\frac{\lambda_1 - 3\lambda_2}{2}\right) = 0$$

$$2 - \lambda_1 - 2\lambda_2 - \lambda_1 + \lambda_2 - \lambda_1 + 3\lambda_2 = 0$$

$$2 - 3\lambda_1 + 2\lambda_2 = 0 \implies \lambda_1 = \frac{2 + 2\lambda_2}{3}$$

$$3 - 2\left(\frac{\lambda_1 + 2\lambda_2}{2}\right) + \left(\frac{\lambda_1 - \lambda_2}{2}\right) + 3\left(\frac{\lambda_1 - 3\lambda_2}{2}\right) = 0$$

$$6 - 2\lambda_1 - 4\lambda_2 + \lambda_1 - \lambda_2 + 3\lambda_1 - 9\lambda_2 = 0$$

$$6 + 2\lambda_1 - 14\lambda_2 = 0$$

$$6 + 2\left(\frac{2 + 2\lambda_2}{3}\right) - 14\lambda_2 = 0$$

$$6 + \frac{4}{3} + \frac{4}{3}\lambda_2 - 14\lambda_2 = 0$$

$$\frac{18 + 4 + 4\lambda_2 - 42\lambda_2}{3} = 0$$

$$22 - 38\lambda_2 \implies \lambda_2 = \frac{\cancel{22}^{\,11}}{\cancel{38}_{\,19}}$$

$$\lambda_1 = \frac{2 + 2\left(\frac{11}{19}\right)}{3} = \frac{\frac{38 + 22}{19}}{3} = \frac{\frac{60}{19}}{3} = \frac{20}{19}$$

$$x = \frac{21}{19} \; ; \quad y = \frac{9}{38} \; ; \quad z = -\frac{13}{38}$$

So the distance is:

$$\sqrt{\left(\frac{21}{19}\right)^2 + \left(\frac{9}{38}\right)^2 + \left(-\frac{13}{38}\right)^2} = \sqrt{\frac{53}{38}} \simeq \underline{1.18}$$

of results points in this direction. The following theorem shows that convex functions can be combined in a variety of ways to produce new convex functions.

(2.3.10) Theorem.
(a) *If $f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})$ are convex functions on a convex set C in R^n, then*

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}) + \cdots + f_k(\mathbf{x})$$

is convex. Moreover, if at least one $f_i(\mathbf{x})$ is strictly convex on C, then the sum $f(\mathbf{x})$ is strictly convex.
(b) *If $f(\mathbf{x})$ is convex (resp. strictly convex) on a convex set C in R^n and if α is a positive number, then $\alpha f(\mathbf{x})$ is convex (resp. strictly convex) on C.*
(c) *If $f(\mathbf{x})$ is a convex (resp. strictly convex) function defined on a convex set C in R^n, and if $g(y)$ is an increasing (resp. strictly increasing) convex function defined on the range of $f(\mathbf{x})$ in R, then the composite function $g(f(\mathbf{x}))$ is convex (resp. strictly convex) on C.*

PROOF. (a) To show that any finite sum of convex functions on C is convex on C, it suffices to show that the sum $(f_1 + f_2)(\mathbf{x})$ of two convex functions $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ on C is again convex on C. If $\mathbf{y}, \mathbf{z}$ belong to C and $0 \le \lambda \le 1$, then

$$\begin{aligned}(f_1 + f_2)(\lambda\mathbf{y} + [1-\lambda]\mathbf{z}) &= f_1(\lambda\mathbf{y} + [1-\lambda]\mathbf{z}) + f_2(\lambda\mathbf{y} + [1-\lambda]\mathbf{z}) \\ &\le \lambda f_1(\mathbf{y}) + [1-\lambda]f_1(\mathbf{z}) + \lambda f_2(\mathbf{y}) + [1-\lambda]f_2(\mathbf{z}) \\ &= \lambda(f_1 + f_2)(\mathbf{y}) + [1-\lambda](f_1 + f_2)(\mathbf{z}).\end{aligned}$$

Hence, $(f_1 + f_2)(\mathbf{x})$ is convex on C. Moreover, it is clear from this computation that if either $f_1(\mathbf{x})$ or $f_2(\mathbf{x})$ is strictly convex, then $(f_1 + f_2)(\mathbf{x})$ is strictly convex because strict convexity of either function introduces a strict inequality at the right place.
(b) This result follows by an argument similar to that used in (a).
(c) If $\mathbf{y}, \mathbf{z}$ belong to C and if $0 \le \lambda \le 1$, then

$$f(\lambda\mathbf{y} + [1-\lambda]\mathbf{z}) \le \lambda f(\mathbf{y}) + [1-\lambda]f(\mathbf{z})$$

since $f(\mathbf{x})$ is convex on C. Consequently, since g is an increasing, convex function on the range of $f(\mathbf{x})$, it follows that

$$\begin{aligned}g(f(\lambda\mathbf{y} + [1-\lambda]\mathbf{z})) &\le g(\lambda f(\mathbf{y}) + [1-\lambda]f(\mathbf{z})) \\ &\le \lambda g(f(\mathbf{y})) + [1-\lambda]g(f(\mathbf{z})).\end{aligned}$$

Thus, the composite function $g(f(\mathbf{x}))$ is convex on C. If $f(\mathbf{x})$ is strictly convex and g is strictly increasing, the first inequality in the preceding computation is strict for $\mathbf{y} \ne \mathbf{z}$ and $0 < \lambda < 1$, so $g(f(\mathbf{x}))$ is strictly convex on C.

(2.3.11) Examples
(a) The function $f(\mathbf{x})$ defined on R^3 by

At first glance, it might seem that the most direct path to verify that f is strictly convex on R^3 would be to show that the Hessian $Hf(\mathbf{x})$ of $f(\mathbf{x}$ positive definite on R^3. However, the Hessian turns out to be

$$Hf(\mathbf{x}) = \begin{pmatrix} (2+4x_1^2)e^{x_1^2+x_2^2+x_3^2} & (4x_1x_2)e^{x_1^2+x_2^2+x_3^2} & (4x_1x_3)e^{x_1^2+x_2^2+x_3^2} \\ (4x_1x_2)e^{x_1^2+x_2^2+x_3^2} & (2+4x_2^2)e^{x_1^2+x_2^2+x_3^2} & (4x_2x_3)e^{x_1^2+x_2^2+x_3^2} \\ (4x_1x_3)e^{x_1^2+x_2^2+x_3^2} & (4x_2x_3)e^{x_1^2+x_2^2+x_3^2} & (2+4x_3^2)e^{x_1^2+x_2^2+x_3^2} \end{pmatrix}$$

Obviously, proving that the Hessian is positive definite for all $x \in R^3$ v involve quite tedious algebra. No matter! There is a much simpler way handle the problem.
First, note that

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

is strictly convex since its Hessian

$$Hh(x_1, x_2, x_3) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

is obviously positive definite. Also, $g(t) = e^t$ is strictly increasing (since $g'(t)$ $e^t > 0$ for all $t \in R$) and (strictly) convex (since $g''(t) = e^t > 0$ for all $t \in$ Therefore, by (2.3.10)(c), $f(\mathbf{x}) = g(h(\mathbf{x}))$ is strictly convex on R^3.
(b) Suppose that $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(k)}$ are fixed vectors in R^n and that c_1, $\ldots, c_k$ are positive real numbers. Then the function $f(\mathbf{x})$ defined on R^n by

$$f(\mathbf{x}) = \sum_{i=1}^{k} c_i e^{\mathbf{a}^{(i)}\cdot\mathbf{x}}$$

is convex.
To prove this statement, first observe that the functions $g_i(\mathbf{x})$ on R^n defin by

$$g_i(\mathbf{x}) = \mathbf{a}^{(i)} \cdot \mathbf{x}, \qquad i = 1, 2, \ldots, k$$

are linear and therefore convex on R^n. Since $h(t) = e^t$ is increasing and conv on R, it follows from (2.3.10)(c) that the functions

$$h(g_i(\mathbf{x})) = e^{\mathbf{a}^{(i)}\cdot\mathbf{x}}, \qquad i = 1, 2, \ldots, k$$

are all convex on R^n. Since $c_1, c_2, \ldots, c_k$ are positive real numbers, we c apply (2.3.10)(a), (b) to conclude that

$$f(\mathbf{x}) = \sum_{i=1}^{k} c_i e^{\mathbf{a}^{(i)}\cdot\mathbf{x}}$$

is convex on R^n.
(c) The function $f(\mathbf{x})$ defined on R^2 by