Homework #2

Math 529 - Section 10

Spring 2013

Marco A. Montes de Oca

1. $\boxed{f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v)}$

Assuming f is convex:

a) $(\lambda u + (1-\lambda)v)^2 \leq \lambda u^2 + (1-\lambda)v^2$

$\lambda^2 u^2 + 2\lambda uv - 2\lambda^2 uv + v^2 - 2\lambda v^2 + \lambda^2 v^2 \leq \lambda u^2 + v^2 - \lambda v^2$

Subtracting right hand side from left hand side:

$(\lambda^2 - \lambda) u^2 + 2\lambda uv - 2\lambda^2 uv - \lambda v^2 + \lambda^2 v^2 \leq 0$

$(\lambda^2 - \lambda) u^2 - 2(\lambda^2 - \lambda) uv + (\lambda^2 - \lambda) v^2 \leq 0$

$(\lambda^2 - \lambda)(u^2 - 2uv + v^2) \leq 0$

$\lambda(\lambda - 1)(u - v)^2 \leq 0$, which occurs for all $\lambda \in [0,1]$. When $\lambda \in (0,1)$, $\lambda(\lambda-1)(u-v)^2 = 0$ only when $u = v$, which means that $f(x) = x^2$ is strictly convex.

b) Pick $\vec{u}, \vec{v} \in \mathbb{R}^2$. If $f(\vec{x})$ is convex, then

$$f(\lambda\vec{u} + (1-\lambda)\vec{v}) \leq \lambda f(\vec{u}) + (1-\lambda) f(\vec{v})$$

$$(\lambda u_1 + (1-\lambda)v_1)^2 + 2(\lambda u_2 + (1-\lambda)v_2)^2 \leq \lambda(u_1^2 + 2u_2^2) + (1-\lambda)(v_1^2 + 2v_2^2)$$

If we subtract the left hand side from the right hand side and simplify:

$$\lambda(1-\lambda)\left[(u_1 - v_1)^2 + 2(u_2 - v_2)^2\right] \geq 0$$

for all $\lambda \in [0,1]$, so $f(\vec{x})$ is indeed convex. Furthermore, when $\lambda \in (0,1)$, the expression above is zero only if $\vec{u} = \vec{v}$, which implies that $f(\vec{x})$ is strictly convex.

c) Using the same reasoning as in the two previous examples: If $f(x,y)$ is convex, then

$$2(\lambda u_1 + (1-\lambda)v_1)^2 - (\lambda u_1 + (1-\lambda)v_1)(\lambda u_2 + (1-\lambda)v_2) + (\lambda u_2 + (1-\lambda)v_2)^2 \leq$$
$$\lambda(2u_1^2 - u_1 u_2 + u_2^2) + (1-\lambda)(2v_1^2 - v_1 v_2 + v_2^2)$$

Simplifying, we get:
right hand side - left hand side =

$$\lambda(1-\lambda)\left[2(u_1-v_1)^2-(u_1-v_1)(u_2-v_2)+(u_2-v_2)^2\right]\geq 0$$

which is true for all $\lambda\in[0,1]$. Moreover, the expression is exactly zero for $\lambda\in(0,1)$ when $\vec{u}=\vec{v}$, so $f(x,y)$ is strictly convex.

2. $\boxed{f(\vec{u})+\nabla f(\vec{u})\cdot(\vec{v}-\vec{u})\leq f(\vec{v})}$

a) If $f(x)$ is convex, then

$$f(u)+f'(u)(v-u)\leq f(v)$$
$$-u^2+(-2u)(v-u)\leq -v^2$$
$$-u^2-2uv+2u^2\leq -v^2$$
$$u^2-2uv+v^2\leq 0$$
$$(u-v)^2\leq 0$$

← This is clearly not possible except when $u=v$. In fact, it is the case that $(u-v)^2\geq 0$, which implies that $f(x)$ is strictly concave.

b) Take $\vec{u}, \vec{v} \in \mathbb{R}^2$, then

$$(u_1+u_2)^2 + \left(2(u_1+u_2),\, 2(u_1+u_2)\right)\begin{pmatrix} v_1 - u_1 \\ v_2 - u_2 \end{pmatrix} = a$$ "left hand side"

$$(v_1+v_2)^2 = b$$ "right hand side"

If $f(\vec{x})$ is convex, then $a \leq b$.

$$b - a = \left(u_1 + u_2 - (v_1+v_2)\right)^2 \geq 0$$ in general,

and exactly equal to zero when $\vec{u} = \vec{v}$.

Therefore $f(\vec{x})$ is strictly convex.

c) Again, taking $\vec{u}, \vec{v} \in \mathbb{R}^2$, we have that

$$-u_1u_2 + (-u_2, -u_1)\begin{pmatrix} v_1 - u_1 \\ v_2 - u_2 \end{pmatrix} =$$

$$-u_1u_2 - u_2(v_1-u_1) - u_1(v_2-u_2) = a$$ "left hand side"

$$-v_1v_2 = b$$ "right hand side"

Then

$$b-a = -(u_1 - v_1)(u_2 - v_2)$$

$b-a \geq 0$ sometimes; for example, if

$\vec{u} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $b-a > 0$

However, $b-a \leq 0$ some other times. For instance, if $\vec{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $b-a < 0$.

We conclude therefore that $-xy$ is neither convex, nor concave.

3. $\boxed{\text{If } Hf(\vec{x}) \text{ is positive definite (or semidefinite) for all } \vec{x} \in \mathbb{R}^n, \text{ then } f(\vec{x}) \text{ is strictly convex (or just convex).}}$

a) $Hf(\vec{x}) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}$, whose eigenvalues are $\{0, 4, 4\}$. Therefore $Hf(\vec{x})$ is positive semidefinite and $f(\vec{x})$ is convex, but not strictly so.

b) $f(\vec{x}) = g(h(\vec{x}))$, where $h(\vec{x})$ is $x_1^2 + x_2^2 + x_3^2$, which is strictly convex because $Hh(\vec{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

which is positive definite for all $\vec{x} \in \mathbb{R}^3$.
The function $g(x) = e^x$ is strictly increasing and strictly convex because $g''(x) > 0$ for all $x \in \mathbb{R}$. Therefore, we may conclude that $g(h(\vec{x}))$ is strictly convex.

c) The Hessian of $f(x,y)$ is

$$Hf(x,y) = \begin{pmatrix} -400y + 1200x^2 + 2 & -400x \\ -400x & 200 \end{pmatrix}$$

and so

$$\Delta_1 = -400y + 1200x^2 + 2$$

$$\Delta_2 = 80000x^2 - 80000y + 400$$

If $(x,y) = (0,1)$, $\Delta_1 < 0$ & $\Delta_2 < 0$, so $Hf(0,1)$ is locally negative definite. However, if $(x,y) = (1,0)$ $Hf(1,0)$ is positive definite. Thus $f(x,y)$ is neither convex, nor concave.

4. The Hessian of $f(x,y)$ is

$$Hf(x,y) = \begin{pmatrix} 2 + \frac{1}{x^2} & -4 \\ -4 & 10 + \frac{1}{y^2} \end{pmatrix}$$

which is positive definite everywhere except $(0,0)$. Note, however, that $f(x,y)$ is not defined for cases where $xy < 0$. Therefore, $f(x,y)$ is strictly convex in $\{(x,y) \mid x>0, y>0\}$.

5. Let the set of global minimizers be X^*. Then, for any u and $v \in X^*$ and $\lambda \in [0,1]$

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v)$$

because f is convex. However, since u and v are global minimizers of f, we have $f(u) = f(v)$:

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda) f(v) = f(v)$$

Which means that

$$f(\lambda u + (1-\lambda)v) = f(v)$$

because by definition $f(v) \leq f(w)$ for all w in the domain of f.

We conclude therefore that the line segment joining u and $v \in X^*$ also belongs to X^*, and therefore X^* is a convex set.