67731 | Convex Optimization and Applications | Ex 2

Guy Lutsker 207029448

1.

1.Not convex.
Let $x = (2, -4), y = (-2, 4), \theta = 0.5$
We see that $\theta x + (1-\theta)y = 0 \to f(\theta x + (1-\theta)y) = 0$
Yet, $\theta f(x) + (1-\theta)f(y) = -4 + -4 = -8$
And so $\theta f(x) + (1-\theta)f(y) < f(\theta x + (1-\theta)y)$

2.Not convex
$x = (-1, 0.5), y = (12, 6), \theta = 0.5$
$0.5f(x) + 0.5f(y) = 0.5 \cdot -2 + 0.5 \cdot 2 = 0 \leq f(0.5x + 0.5y) = f(5.5, 3.25)$ which is larger than 0.

3.
It is convex, proof:

$$\nabla f = [\frac{2x_1}{x_2}, \frac{-x_1^2}{x_2^2}]$$

$$\nabla f^2 = \begin{bmatrix} \frac{2}{x_2} & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix} = \frac{2}{x_2} \begin{bmatrix} 1 & \frac{-x_1}{x_2} \\ \frac{-x_1}{x_2} & \frac{x_1^2}{x_2^2} \end{bmatrix}$$

And since $x_2 \geq 0$ what is left to show is that $\begin{bmatrix} 1 & \frac{-x_1}{x_2} \\ \frac{-x_1}{x_2} & \frac{x_1^2}{x_2^2} \end{bmatrix}$ is PSD.

$$|\begin{bmatrix} 1 & \frac{-x_1}{x_2} \\ \frac{-x_1}{x_2} & \frac{x_1^2}{x_2^2} \end{bmatrix}| = \frac{x_1^2}{x_2^2} - \frac{x_1^2}{x_2^2} = 0 \geq 0$$

$$Tr(\begin{bmatrix} 1 & \frac{-x_1}{x_2} \\ \frac{-x_1}{x_2} & \frac{x_1^2}{x_2^2} \end{bmatrix}) = 1 + \frac{x_1^2}{x_2^2} \geq 0$$

And so both the trace and the determinant are non negative, meaning ∇f^2 is PSD

4.

2

25. Prove that if f is convex then ∇f is monotonic increasing.

$$Solution:$$

A criteria for convexity is that $\nabla^2 f$ is PSD which means that ∇f is monotonic increasing... But I guess we can prove it from definition:

Let f be convex, and so lets look at the expression we want to prove:

$$(\nabla f(x)^\intercal - \nabla f(y)^\intercal)(x-y) \overset{?}{\geq} 0 \Leftrightarrow$$

$$\nabla f(x)^\intercal (x-y) - \nabla f(y)^\intercal (x-y) \overset{?}{\geq} 0 \Leftrightarrow$$

And if we remember the definition of high dimensional convexity, with regard to the first order derivative we get:

$$f(x) \geq f(y) + \nabla f(x)^\intercal (y-x)$$

$$f(y) \geq f(x) + \nabla f(x)^\intercal (x-y)$$

Summing up these equations yields:

$$\nabla f(x)^\intercal (x-y) - \nabla f(y)^\intercal (x-y) + f(x) + f(y) \leq f(x) + f(y) \Leftrightarrow$$

$$\nabla f(x)^\intercal (x-y) - \nabla f(y)^\intercal (x-y) \leq 0$$

Which is awfully close to what we want to prove.

But if we remember that the statement in the question is regarding any $x, y \in dom(f)$ we realize that the order isn't important and we can flip it to get:

$$\nabla f(x)^\intercal (x-y) - \nabla f(y)^\intercal (x-y) \leq 0 \Leftrightarrow$$

$$\nabla f(y)^\intercal (y-x) - \nabla f(x)^\intercal (y-x) \geq 0$$

Which is what we wanted.

3.

4.

1.Define $h(x) = \sum a_i x_i$ we see that h is affine and so convex.

We see that $\frac{\partial h}{\partial x_i} = a_i$ which are non-negative and so we can write it as $f = h(f_1..f_n)$ and from Q3 first clain we get that f is convex.

2.f is affine at each coordinate in x with each coordinate being $\langle a_i, x\rangle + b_i$ where a_i is the i row of A .

f is convex and from Q3 claim 3 we get that $g(x)$ is also convex.

3. max is a convex function and if we have the freedom to look at each coordinate and choose the maximum we get a convex function as well.

from Q3 first claim we get that $max(f_1...f_n)$ is convex.

5.

1.

$$f(x) = log(\frac{1+e^x}{e^x}) = log(1+e^x) - log(e^x) = log(1+e^x) - x.$$

$$\frac{\partial^2 f(x)}{\partial x^2} = \frac{\partial(\frac{e^x}{1+e^x} - 1)}{\partial x} = \frac{e^x}{(1+e^x)^2}$$

which is non negative for all x. And so f is convex.

2.$f(x) = -|x|^{\frac{1}{n}}$
We seen in class that f is convex iff $\forall x \in Dom(f), \forall v\ g(t) = f(x+tv)$ is convex.
define $g(t) = f(x+tvv)$ with $dom(g) = \{t \in \mathbb{R} | x+tv \in dom(f)\}$.
Notice that:

$$g(t) = -|x+tv|^{\frac{1}{n}} = -|x|^{\frac{1}{n}} |I + tx^{\frac{1}{2}} vz^{\frac{1}{2}}|^{\frac{1}{n}} = -|x|^{\frac{1}{n}} \prod (1+t\lambda_i)^{\frac{1}{n}}$$

We have seen in class that $-\prod (x_i)^{\frac{1}{n}}$ is convex and since $1+t\lambda_i$ is affine in x and so $\prod(1+t\lambda_i)^{\frac{1}{n}}$ is convex too. $|x|^{\frac{1}{n}}$ is always positive and so all together $g(t)$ is convex which means that $f(x)$ is convex as well.

3.$f(x) = \sqrt{\sum (x^{\intercal} P_i x)^2}$
define $g(x) = ||x||$ and $h_i = x^{\intercal} p_i x$.
g is convex as we saw in class, and we have also seen in class that h_i is convex and so their combination - $f = (g(h_i(x)...)$ is also convex.

4.$f(x) = \frac{x^{\intercal} x}{\prod x_i^{\frac{1}{n}}}$
We can use Q3. define $g(x) = ||x||, h(x) = \prod x_i^{\frac{1}{n}}$ both of then are convex, and if we shall define $r(x,y) = \frac{x^2}{y}$ which we also know from class is convex, we get that :

$$f(x) = r(g(x), h(x))$$

is convex as well.

6.

1.

Let $y_1, y_2 \in dom(f^*)$.

$$f^*(\theta y_1 + (1-\theta) y_2) = sup\{\langle x, \theta y_1 + (1-\theta) y_2 \rangle - f(x)\}$$

$$= sup\{\theta(\langle x, y_1 \rangle - f(x) + (1-\theta)(\langle x, y_2 \rangle - f(x))\}$$

$$\leq sup\{(\langle x, y_1 \rangle - f(x))\} + (1-\theta) sup\{(\langle x, y_2 \rangle - f(x))\}$$

$$= \theta f^*(y_1) + (1-\theta) f^*(y_2)$$

Which means f^* is convex.

2.

Let $Q > 0$ which means Q is ivertible and so:

$$f^*(y) = sup\{\langle x, y \rangle - 0.5 x^{\intercal} Q x\}$$

$$\nabla \langle x, y \rangle - 0.5 x^{\intercal} Q x = 0$$

and so $x = Q^{-1} y$
which means that:

$$f^*(y) = \langle Q^{-1} y, y \rangle - 0.5 (Q^{-1} y)^{\intercal} P P^{-1} y = 0.5 y^{\intercal} Q^{-1} y$$

3.a not for submission
3.b
since $\|y\|^* > 1$ we get tat there exists w with $\|w\| \leq 1$ and $\langle w, y \rangle > 1$ and $\langle w, y \rangle - \|w\| > 1 - \|w\| \geq 0$.
$f^*(y) = sup\{\langle x, y \rangle - \|x\|\} \geq \langle tw, y \rangle - \|tw\| = t(\langle w, y \rangle - \|w\|)$
And for $t \to \infty$ it clearly goes off to infinity.