Machine Learning WS 2018

Solution to Assignment 5

Submitted By : Vindhya Singh **Enrollment Number :** 03693296
Colaborated By : Wasiq Rumaney **Enrollment Number :** 03694978

Solution 1 :

i. $f(x,y)=x^2+2y+\cos(\sin(\sqrt{\pi})-min(-x^2,\log(y)))$ D = (-100,100) x (1,50)

We know that,
$-min\{-x^2,\log(y)\}=max\{x^2,-\log(y)\}$

$\frac{\partial^2(x^2)}{\partial x^2}=2\geq 0$ thus, convex (A)

$\frac{\partial^2(-\log(y))}{\partial y^2}=\frac{1}{y^2}\geq 0$ over the given domain D. Thus, convex.

$\Rightarrow max\{x^2,-\log(y)\}$ is convex. [Lec 5, Slide 17, Property c] (1)

x^2 is a convex function (from (A)) (2)

2y is a linear function and thus convex too. (3)

$\cos(\sin(\sqrt{\pi}))$ has twice differential equal to 0 because it is constant in x and y (4)

From (1),(2),(3),(4) we can see that **f(x,y) is convex** too as the sum of convex functions is also convex **in the given domain D.**

ii. $f(x)=\log(x)-x^3$ D = (1,∞)

For convexity, we know that,

$$\frac{d^2f(x)}{dx^2}\geq 0$$

$$f'(x)=x^{-1}-3x^2$$

$\frac{d^2f(x)}{d(x)^2}=\frac{-1}{x^2}-6x\geq 0$ for the function to be convex.

However, in the given domain D, the above condition does not hold because any value from the given domain will be positive and on substitution in the above inequality, it will make the function negative. Thus, the function $f(x)=\log(x)-x^3$ is **not convex**.

iii. $f(x) = -min\{\log(3x+1), -x^4 - 3x^2 + 8x - 42\}$ **(D = R⁺)**

We know that,

$$-min\{\log(3x+1), -x^4 - 3x^2 + 8x - 42\} = max\{-\log(3x+1), x^4 + 3x^2 - 8x + 42\}$$

Let, $g(x) = -\log(3x+1)$ and $h(x) = x^4 + 3x^2 - 8x + 42$

$$\frac{d\,g(x)}{dx} = \frac{-3}{(3x+1)}$$

$$\frac{d^2 g(x)}{dx^2} = \frac{9}{(3x+1)^2} \geq 0$$ on D $\rightarrow$ R^+ Therefore, function g(x) is convex. (1)

$$\frac{dh(x)}{dx} = 4x^3 + 6x - 8$$

$$\frac{d^2 h(x)}{dx^2} = 12x^2 + 6 \geq 0$$ on D. Therefore, function h(x) is convex. (2)

Since both the given functions are convex (from (1) and (2)). Therefore, from Lec 5, slide 17, Property 2, the given function, $f(x) = -min\{\log(3x+1), -x^4 - 3x^2 + 8x - 42\}$ is **convex in the given domain.**

iv. $f(x, y) = yx^3 - yx^2 + y^2 + y + 4$ D = (-10,10) x (-10,10)

For a function to be convex, the second derivative must not be negative. (1)

Taking the second derivative of the given function,

$$f''(x, y) = 6x^2 - 4x + 6xy - 2y + 2$$

In the given domain, if we take the values of $x = 2, y = -9$, then the value of above equation becomes -72 i.e. negative. This contradicts the statement (1). Hence, the given function is **not convex in the given domain**

Solution 2 :

Given : $f_1 : \mathbb{R}^D \rightarrow \mathbb{R}$ and $f_2 : \mathbb{R}^D \rightarrow \mathbb{R}$ are two convex functions

To Prove : $h(x) = max\{f_1(x), f_2(x)\}$ is also convex

Proof : We know that for convexity,

$$f_1(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f_1(x_1) + (1-\lambda) f_1(x_2) \tag{1}$$

$$f_2(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f_2(x_1) + (1-\lambda) f_2(x_2) \tag{2}$$

(Since given that both the functions are convex)

Thus for,

$h(x) = max\{f_1(x), f_2(x)\}$ if we prove that,

$$h(\lambda x_1 + (1-\lambda)x_2) \leq \lambda h(x_1) + (1-\lambda) h(x_2)$$

or,

$$h(t) \leq \lambda h(x_1) + (1-\lambda) h(x_2)$$

because $h(x) = max\{f_1(x), f_2(x)\}$ it is safe to say that,

$$f_1(t) \leq \lambda h(x_1) + (1-\lambda) h(x_2) \tag{3}$$

$$f_2(t) \leq \lambda h(x_1) + (1-\lambda) h(x_2) \tag{4}$$

$h(t) = max\{f_1(t), f_2(t)\}$ Thus, either f_1(t) or f_2(t) will be the maximum. (5)

From (1), (2), (3), (4) and (5), the following holds :

$$h(t) \leq \lambda h(x_1) + (1-\lambda) h(x_2)$$

Thus, h is a convex function.

Hence, Proved.

Solution 3 :

Given : $f_2 : \mathbb{R} \rightarrow \mathbb{R}$ and $f_1 : \mathbb{R} \rightarrow \mathbb{R}$ are two convex functions

To Prove/Disprove : $g(x) = f_1(f_2(x))$ is also convex

Proof : For a function f(x) to be convex, we need to prove that,

$$\frac{d^2 f(x)}{dx^2} \geq 0 \qquad (1)$$

From (1), for g(x) to be convex,

$$\frac{d^2 g(x)}{dx^2} \geq 0$$

$$\frac{d\,g(x)}{dx} = f_1'(f_2(x)).f_2'(x)$$

$$\frac{d^2(g(x))}{d\,x^2} = \frac{d^2 f_1(f_2(x))}{dx^2} \frac{d f_2(x)}{dx} . \frac{d f_2(x)}{dx} + \frac{d^2 f_2(x)}{dx^2} . \frac{d f_1(f_2(x))}{dx} \geq 0$$

$$\frac{d^2(g(x))}{d\,x^2} = \frac{d^2 f_1(f_2(x))}{dx^2} \left(\frac{d f_2(x)}{dx^2}\right)^2 + \frac{d^2 f_2(x)}{dx^2} . \frac{d f_1(f_2(x))}{dx} \geq 0 \qquad (2)$$

Since it is given that the functions f_1 and f_2 are convex functions (given).
Therefore, second derivative of both f_1 and f_2 are greater than or equal to zero. The first part of (2) is thus greater than or equal to zero.

This implies that, for (2) to hold, $\frac{d f(x)}{dx} \geq 0$ for g(x) to be convex.

Hence, g(x) is convex if f_1 is convex and non-decreasing.

Proved.

<u>Solution 4 :</u>

<u>Given :</u> $f: \mathbb{R}^N \rightarrow \mathbb{R}$ is a convex function

<u>To Prove/Disprove :</u> In a convex function, every local minima is a global minima, i.e. $\nabla f(\theta^*) = 0$ then θ^* is a global minimum.

<u>Proof :</u> We know that for a convex function f,

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \text{where} \quad \lambda = [0,1] \tag{1}$$

Adopting a proof by contradiction approach,

Let there be a point z such that $f(z) < f(\theta^*)$ where θ^* is a local minima and U is the neighborhood region such that $\exists z \in U$

Thus,

$$f(\theta^*) = f(\lambda \theta^* + (1-\lambda) z) \leq \lambda f(\theta^*) + (1-\lambda) f(z) \qquad \text{[from (1)]}$$

$$< \lambda f(\theta^*) + (1-\lambda) f(\theta^*) \tag{2}$$

Thus (2) becomes $f(\theta^*)$

Because of the convexity of U,

$$\lambda \theta^* + (1-\lambda) z \in U \quad \text{where} \quad \lambda = [0,1] \tag{3}$$

If $\lambda \rightarrow 1$ (3) tends to $\lambda \theta^* = \theta^*$

This contradicts our assumption.

Hence, if θ^* is a local minima of f, f being a convex function, then θ^* is a global minima.

So in a convex function, every local minima is a global minima.

Proved.