(c)

HomeWork : 6

UID:

Name: Aishwarya Hiremath

7.1] Consider the univariate function

$$f(x) = x^3 + 6x^2 - 3x - 5$$

Find its stationary points and indicate whether they are maximum, minimum, or saddle points.

soln

To find the stationary points of the function $f(x) = x^3 + 6x^2 - 3x - 5$

We need to find the stationary points by setting the gradient to zero, and solving for x

OR

we need to find the values of x where the derivative $f'(x)$ is equal to zero. After finding these stationary points, we can analyze the second derivative $f''(x)$ to determine whether they are maximam, minimum OR saddle points.

Find the derivative OR obtain gradient.

$$\frac{df}{dx} = 3x^2 + 12x - 3$$

to find the stationary point set $\frac{df}{dx} = 0$ and solve for x

$$3x^2+12x-3=0$$

$$x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

$$= \frac{-12 \pm \sqrt{(12)^2-4(3)(-3)}}{2(3)}$$

$$= \frac{-12 \pm \sqrt{144+36}}{6} = \frac{-12 \pm \sqrt{180}}{6}$$

$$= \frac{-12 \pm 6\sqrt{5}}{6} = \cancel{6}\left(\frac{-2 \pm \sqrt{5}}{\cancel{6}}\right)$$

$$x = -2 \pm \sqrt{5}$$

$\therefore$ The stationary points are

$x = -2+\sqrt{5}$ and $x = -2-\sqrt{5}$

Analyze the second derivative

$$\frac{d^2f}{dx^2} = 6x+12 \longrightarrow \text{Hessian}$$

substituting the sol^n of $\frac{df}{dx}=0$ into Hessian

$$\frac{d^2f}{dx^2}(-2-\sqrt{5}) = 6(-2-\sqrt{5})+12$$

$$= -12 - 6\sqrt{5}+12 = -13.4164$$

$$\frac{d^2f}{dx^2}(-2+\sqrt{5}) = 6(-2+\sqrt{5})+12$$

$$= -12+6\sqrt{5}+12$$

$$= 13.4164$$

Based on the second derivative test, we conclude that.

The point $x = -2-\sqrt{5}$ is a maximum

The point $x = -2+\sqrt{5}$ is a minimum

7.2] Consider the update equation for stochastic gradient descent. Write down the update when we use a mini-batch size of one.

Sol^n Equation 7.15 is given by

$$\theta_{i+1} = \theta_i - \gamma_i \left(\nabla L(\theta_i)\right)^T$$

$$= \theta_i - \gamma_i \sum_{n=1}^{N} \left(\nabla L_n(\theta_i)\right)^T \longrightarrow (1)$$

write the eq^n (1) as a update when we use a mini-batch size of one

$$\theta_{i+1} = \theta_i - \gamma_i \left(\nabla L(\theta_i)\right)^T$$

$$= \theta_i - \gamma_i \left[\nabla L_k(\theta_i)\right]^T$$

where K is the index of the example that is randomly chosen.

7.3] Consider whether the following statements are true or false:

a] The intersection of any two convex sets is convex.

sol^n True:

The intersection of any two convex sets is convex. This property is a fundamental property of convex sets. If you take two convex sets and find their intersection, the resulting set will also be convex. Intuitively, this property holds because any line segment connecting two points within the intersection will also lie within the intersection

b] ~~False~~: The union of any two convex sets is ~~is~~ convex

sol^n False:

The union of two convex sets is not necessarily convex. Consider two convex sets in a two-dimensional space that are disjoint (have no overlap). Their union would not form a convex set since you can find points within the union that do not belong to either of the original sets.

e] The difference of a convex set A from another convex set B is convex.

sol^n False:

The difference of a convex set A from another convex set B is not necessarily convex. If you take the difference A - B, it may result in a set that is not convex. Consider a scenario where A is a convex set and B is a subset of A. The difference A - B may contain non-convex regions, leading to a set that is not convex.

7.4] Consider whether the following statements are true or false:

a] The sum of any two convex functions is convex.

sol^n True:

The sum of any two convex functions is convex. This is a property of convex functions. If you take two convex functions and add them together, the resulting function will also be convex.

b] The difference of any two convex function is convex sol^n False:

The difference of any two convex functions is not necessarily convex. While the sum of convex functions is convex, the difference might not retain convexity. Counterexamples can be constructed to demonstrate this

c] The product of any two convex function is convex.

sol^n False:

The product of any two convex functions is not necessarily convex. Convexity does not generally hold for the product of two convex function

d] The maximum of any two convex function is convex.

sol^n True:

The maximum of any two convex function is convex. because the function is already convex and it is choosen from set which is convex.

7.5] Express the following optimization problem as a standard linear program in matrix notation

$$\max_{x \in \mathbb{R}^2, \xi \in \mathbb{R}} \quad p^T x + \xi$$

subject to the constraints that $\xi \geq 0$, $x_0 \leq 0$ and $x_1 \leq 3$.

Sol^n The optimization target and constraints must all be linear. To make the optimization target linear,

we combine x and ξ into one matrix equation

$$\max_{x \in \mathbb{R}^2, \xi \in \mathbb{R}} \begin{bmatrix} p_0 \\ p_1 \\ 1 \end{bmatrix}^T \begin{bmatrix} x_0 \\ x_1 \\ \xi \end{bmatrix}$$

We can then combine all of the inequalities into one matrix inequality.

$$\text{subject to} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \xi \end{bmatrix} - \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} \leq 0$$