Question 1: First Order Derivative

Consider the function $f(x) = 3x\ 2 + 5x + 2$ a) Find the first-order derivative $f'(x)$ b) Determine the critical points of $f(x)$ by setting f and solving for $f'(x) = 0$. c) Use the first-order derivative test to classify each critical point as a local minimum, local maximum, or neither.

a) Given function $f(x) = 3x^2 + 5x + 2$, let's find its derivative $f'(x)$:

$f'(x)$=d/dx $(3x^2+5x+2)$

$f'(x)=6x+5$

So, the first-order derivative of $f(x)$ is $\boldsymbol{f'(x)=6x+5}$

b) Determine the critical points of $f(x)$ by setting $f'(x) = 0$

$6x+5= 0$

Solving for x:

$6x= -5$

$x= -5/6$

So, **the critical point is $x= -5/6$**

c) To determine if the critical point $x= -5/6$ is a local minimum, local maximum, or neither, we can use the first-order derivative test. The test involves checking the sign of $f'(x)$ around the critical point.

Evaluate $f'(-1)$:

$f'(-1)=6(-1)+5= -1$

Since $f'(-1)$ is negative, it indicates that the function is decreasing to the left of the critical point.

Evaluate $f'(-5/6)$):

$f'(-5/6)=6(-5/6)+5= -1$

Since $f'(-5/6)$ is also negative, it indicates that the function is decreasing to the right of the critical point.

Therefore, by the first-order derivative test**, the critical point $x= -5/6$ is a local maximum**.

Question 2: Second Order Derivative

Continuing from Question 1, let $f(x) = 3x^2 + 5x + 2$ a) Find the second-order derivative $f''(x)$ b) Evaluate $f''(x)$ at the critical points found in Question 1. c) Use the second-order derivative test to determine whether each critical point is a local minimum, local maximum, or neither.

a) Given $f(x)$=$3x^2$+5x+2, let's find its second-order derivative $f''(x)$:

$f''(x)$= d^2/dx^2 (3x2+5x+2)

$f''(x)$= d/dx(6x+5)

$f''(x)$= 6

So, the second-order derivative of $f(x)$ is **$f''(x)$=6**.

b) Evaluate $f''(x)$ at the critical point x= −5/6,

f'' (−5/6) = 6.

c) The second-order derivative test involves checking the sign of $f''(x)$ at the critical point.

Since f'' (−5/6) = 6 (positive), it indicates that the function is concave up at x=−5/6.

Therefore, by the second-order derivative test, **the critical point x= −5/6 is a local minimum**.

Question 3: Chain Rule

Consider the functions $g(u) = u^3$and $h(x) = 2x - 1$, where u is a function of x, i.e., $u = h(x)$.
a) Find $g'(u)$ and $h'(x)$
b) Apply the chain rule to find dg/dx.

a) Given functions:

$g(u)$=u

$h(x)$=2x−1

u=$h(x)$

Let's find $g'(u)$ and $h'(x)$:

$g'(u)$=dg/du (u) = 1

$h'(x)$=dh/dx (2x−1) = 2

b) The chain rule states that:

$dg/dx = dg/du$ x du/dx

= $g'(u)$ x $d[h(x)]/dx$

= $g'(u)$ x $h'(x)$

=1·2

=2

Therefore**, *dg/dx* =2**.

Question 4: Gradient Descent

A machine learning model has a cost function $J(\theta) = \theta^2 - 4\theta + 5$, where θ is the model parameter.
a) Find the first-order partial derivative $\partial J/\partial\theta$
b) Apply the gradient descent update rule: $\theta_{new} = \theta_{old} - \alpha\, \partial J/\partial\theta$, where α is the learning rate (assume $\alpha = 0.1$)
c) Explain how the gradient descent process helps in finding the minimum of the cost function.

a) Given the cost function $J(\theta)=\theta^2-4\theta+5$,

let's find its first-order partial derivative with respect to θ:

$\partial J/\partial\theta = \partial/\partial\theta\ [\theta^2-4\theta+5]$

$=2\theta-4$

So, the first-order partial derivative is $\boldsymbol{\partial J/\partial\theta = 2\theta-4}$**.**

b) The gradient descent update rule is given by:

$\theta_{new}=\theta_{old}-\alpha\ \partial J/\partial\theta$

Substituting the given values:

$\boldsymbol{\theta_{new}=\theta_{old}-0.1(2\theta_{old}-4)}$

c) Gradient descent is an optimization algorithm that aims to minimize the cost function by iteratively moving towards the minimum.

1. **Initialization:** Start with an initial guess for the model parameter θ, denoted as θ_{old}.

Calculate Gradient: Compute the gradient of the cost function with respect to θ, i.e., $\partial J/\partial\theta$.

2. **Update $\boldsymbol{\theta}$:** Use the gradient to update the parameter θ in the direction that reduces the cost. The update is done using the gradient descent update rule.
3. **Iterate:** Repeat steps 2 and 3 until convergence or a predefined number of iterations.

The process helps in finding the minimum of the cost function because the algorithm adjusts the model parameter in the direction opposite to the gradient, effectively "descending" down the slope of the cost function. As it iteratively updates θ, it approaches the optimal value where the cost function is minimized.