Chain Rule in Calculus

The chain rule is a fundamental concept in calculus used to compute the derivative of a composite function. When you have a function that is composed of other functions, the chain rule allows you to differentiate it by taking into account the derivatives of the inner functions.

Chain Rule for Single-Variable Functions

If you have two functions f and g, where $y = f(u)$ and $u = g(x)$, the composite function y can be written as $y = f(g(x))$. To find the derivative of y with respect to x, you apply the chain rule:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

Example

Consider the functions $f(u) = u^2$ and $g(x) = \sin(x)$. The composite function is $y = f(g(x)) = (\sin(x))^2$. To find $\frac{dy}{dx}$:

1. Compute $\frac{dy}{du}$:

$$\frac{dy}{du} = 2u$$

Since $u = \sin(x)$, this becomes:

$$\frac{dy}{du} = 2\sin(x)$$

2. Compute $\frac{du}{dx}$:

$$\frac{du}{dx} = \frac{d}{dx}\sin(x) = \cos(x)$$

3. Apply the chain rule:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 2\sin(x) \cdot \cos(x)$$

Chain Rule for Multi-Variable Functions

The chain rule also extends to functions of multiple variables. If $z = f(u, v)$ where $u = g(x, y)$ and $v = h(x, y)$, then the partial derivatives of z with respect to x and y are given by:

$$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial x}$$

$$\frac{\partial z}{\partial y} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial y}$$

Example

Consider $z = f(u, v) = u^2 + v^2$ with $u = x + y$ and $v = x - y$. To find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$:

1. Compute the partial derivatives of f:

$$\frac{\partial f}{\partial u} = 2u, \quad \frac{\partial f}{\partial v} = 2v$$

2. Compute the partial derivatives of u and v:

$$\frac{\partial u}{\partial x} = 1, \quad \frac{\partial u}{\partial y} = 1$$

$$\frac{\partial v}{\partial x} = 1, \quad \frac{\partial v}{\partial y} = -1$$

3. Apply the chain rule:

$$\frac{\partial z}{\partial x} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial x} = 2u \cdot 1 + 2v \cdot 1 = 2(x + y) + 2(x - y) = 4x$$

$$\frac{\partial z}{\partial y} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial y} = 2u \cdot 1 + 2v \cdot (-1) = 2(x + y) - 2(x - y) = 4y$$

Summary

The chain rule is an essential tool in calculus that allows for the differentiation of composite functions. For single-variable functions, it connects the derivative of the outer function with the derivative of the inner function. For multi-variable functions, it relates the partial derivatives of the composite function to the partial derivatives of the inner functions. This rule is widely used in various fields, including physics, engineering, and especially in training neural networks in machine learning.