Chain Rule for Multivariable Functions

Investigation

How does the chain rule work when we have more than one independent variable?
We'll start by using parametric equations.

Elementary Definition of Chain Rule

For a single-variable function, the chain rule is given by:

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

Here, y is a function of x, and x is a function of t.

Demonstration

Consider the function:

$$y = (3x)^4$$

Let $t = 3x$. Now, t is a function of x.

Thus, we can express y in terms of t:

$$y = t^4$$

Here, t is the only independent variable, even though we originally had three variables.

Conceptual Understanding

In general:

- Suppose $y = f(x)$ and $x = g(t)$.
- Thus, $y = f(g(t))$.

The chain rule is used for compositions of functions. We need an "intermediate variable" to connect y to t and t to x. In this example:

- x is the intermediate variable for t.
- By substituting x with $g(t)$, x effectively becomes an independent variable.

Multivariable Example

Consider the function:

$$w = x^2 - y^2$$

where:

$$x = t^2 + 1$$

$$y = t^3 + t$$

Substituting x and y into the function for w:

$$w = (t^2 + 1)^2 - (t^3 + t)^2$$

Analysis

In this example:

- x and y are independent variables in the function w.
- However, both x and y are dependent on t.

This means that while x and y are independent in the expression for w, their actual values are functions of t.

Concept

As t changes, both x and y depend on t, so x and y change. Consequently, w changes as well. To understand how w changes with respect to t, we need to consider how much x and y change. By adding these changes together, we get the total change in w.

If the slope is defined as $\mathrm{rise}/\mathrm{run}$, then:

$$\text{rise} = \text{slope} \times \text{run}$$

In other words, the change in height (in one dimension) is related to the distance traveled in some slope direction.

Thus, the change in height (Δw) is given by:

$$\Delta w = \left(\frac{\partial w}{\partial x} \cdot \Delta x\right) + \left(\frac{\partial w}{\partial y} \cdot \Delta y\right)$$

Concise Formula

The total change in w can be written as:

$$\Delta w = \left(\frac{\partial w}{\partial x} \cdot \Delta x\right) + \left(\frac{\partial w}{\partial y} \cdot \Delta y\right)$$

To find the rate of change with respect to t:

$$\frac{\Delta w}{\Delta t} = \frac{\left(\frac{\partial w}{\partial x} \cdot \Delta x\right) + \left(\frac{\partial w}{\partial y} \cdot \Delta y\right)}{\Delta t}$$

This can be broken down as:

$$\frac{\Delta w}{\Delta t} = \frac{\partial w}{\partial x} \cdot \frac{\Delta x}{\Delta t} + \frac{\partial w}{\partial y} \cdot \frac{\Delta y}{\Delta t}$$

We need to take the limit as $\Delta t \to 0$:

$$\lim_{\Delta t \to 0} \left(\frac{\Delta w}{\Delta t}\right) = \lim_{\Delta t \to 0} \left(\frac{\partial w}{\partial x} \cdot \frac{\Delta x}{\Delta t}\right) + \lim_{\Delta t \to 0} \left(\frac{\partial w}{\partial y} \cdot \frac{\Delta y}{\Delta t}\right)$$

As $\Delta t \to 0$, Δx and Δy also approach 0, representing infinitesimally small changes. Therefore:

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial w}{\partial y} \cdot \frac{dy}{dt}$$

In plain words, the change in w with respect to t is given by the partial derivative of w with respect to x multiplied by the derivative of x with respect to t, plus the partial derivative of w with respect to y multiplied by the derivative of y with respect to t.

Important: Always keep in mind that these concepts are based on compositions of functions.

misc

lecture video: https://youtu.be/tXryaM-mTpY