Chain Rule: Step-by-Step Example

From Calculus to Neural Networks

Praveen Kumar Chandaliya

Department of Artificial Intelligence
SVNIT Surat

January 9, 2026

Problem Statement

Given the function:

$$f(x, y, z) = (x + y)z$$

with input values:

$$x = -2, \quad y = 5, \quad z = -4$$

Goal:

- Apply the chain rule step by step
- Compute partial derivatives
- Interpret the result using a neural network view

Step 1: Introduce Intermediate Variable

Define an intermediate variable:

$$u = x + y$$

Then the function becomes:

$$f = u \cdot z$$

Dependency structure:

$$x, y \to u \to f, \quad z \to f$$

Step 2: Partial Derivatives

Derivatives of intermediate variables:

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

Derivative of output:

$$\frac{\partial f}{\partial u} = z, \quad \frac{\partial f}{\partial z} = u$$

Step 3: Apply the Chain Rule

Partial derivatives:

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

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

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

Step 4: Numerical Evaluation

Compute intermediate value:

$$u = x + y = -2 + 5 = 3$$

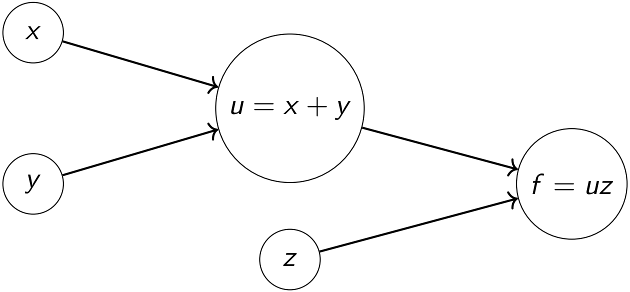
Final gradients:

$$\frac{\partial f}{\partial x} = -4, \quad \frac{\partial f}{\partial y} = -4, \quad \frac{\partial f}{\partial z} = 3$$

Gradient vector:

$$\nabla f = \begin{bmatrix} -4 \\ -4 \\ 3 \end{bmatrix}$$

Computational Graph (Forward & Backward)



Backpropagation: gradients flow backward using chain rule

Neural Network Interpretation

This computation is equivalent to a small neural network:

- $u = x + y \rightarrow$ hidden neuron (linear layer)
- $f = uz \rightarrow$ output neuron
- Gradients computed using chain rule = backpropagation

$$\text{Forward pass} \rightarrow \text{Loss} \rightarrow \text{Backward pass}$$

Key Insight:

Backpropagation is repeated application of the chain rule

Key Takeaways

- Introduce intermediate variables
- Apply chain rule step by step
- Use computational graphs to simplify derivatives
- Same principle powers all deep neural networks

Without the chain rule, deep learning cannot train.