

Vanilla Neural Networks

Shamel Bhimani

August 2025

Contents

1	Precursors to Backpropagation	2
1.1	Chain Rule	2
1.1.1	Core Principle of the Chain Rule	2
1.1.2	The Single-Variable Chain Rule	2

1 Precursors to Backpropagation

test

1.1 Chain Rule

1.1.1 Core Principle of the Chain Rule

The chain rule is a formula for finding the *derivative* of a **composite function**, that is, a function that is formed by the composition of two or more other functions. It is an indispensable tool in science, engineering, and statistics, particularly for optimization problems where functions are dependent on a chain of intermediate variables. In the context of neural networks, the backpropagation algorithm is, at its core, a highly efficient and systematic application of the chain rule.

1.1.2 The Single-Variable Chain Rule

Let y be a function of u , denoted as $y = f(u)$, and let u in turn be a function of x , denoted as $u = g(x)$. The composition of these two functions forms a new function, $y = F(x) = f(g(x))$.

The **Chain Rule Theorem** states that if both f and g are differentiable functions, then the derivative of the composition function $F(x)$ with respect to x is the product of the derivative of the outer function f with respect to its input u , and the derivative of the inner function g with respect to its input x .

Theorem 1.1 (Chain Rule). *If $y = f(u)$ and $u = g(x)$ are differentiable functions, then the derivative of y with respect to x is given by:*

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

An equivalent notation, often used for simplicity, is the prime notation:

$$F'(x) = f'(g(x)) \cdot g'(x)$$

Example 1.1.1.

Let $y = (x^2 + 1)^3$. We can identify this as a composite function. Let the inner function be $u = g(x) = x^2 + 1$. Let the outer function be $y = f(u) = u^3$.

First, we find the individual derivatives:

$$\begin{aligned} \frac{dy}{du} &= \frac{d}{du}(u^3) = 3u^2 \\ \frac{du}{dx} &= \frac{d}{dx}(x^2 + 1) = 2x \end{aligned}$$

Now, we apply chain rule:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = (3u^2) \cdot (2x)$$

Finally, we substitute the expression for u back into the result:

$$\frac{dy}{dx} = 3(x^2 + 1)^2 \cdot (2x) = 6x(x^2 + 1)^2$$