Partial Derivatives

Raiqa Rasool

November 13, 2023

1 Purpose of Partial Derivatives

Partial derivatives are used in multivariable calculus to understand how a function changes concerning one specific independent variable while keeping the other independent variables constant. In other words, they allow us to study the rate of change of a function concerning one variable at a time, holding all other variables fixed.

Partial derivatives are particularly useful when dealing with functions of multiple variables, which often arise in fields like physics, engineering, economics, and more. These derivatives help us analyze how a change in one variable affects the output of the function, while ignoring the effects of other variables.

Mathematically, if you have a function with multiple variables, say z = f(x, y), then the partial derivative of z with respect to x, denoted as dz/dx, represents the rate of change of z concerning x while treating y as a constant. Similarly, dz/dy represents the rate of change of z concerning y while treating x as a constant.

Partial derivatives are essential in optimization problems, such as finding local extrema (maxima and minima) of functions with multiple variables, as well as in understanding the behavior of systems with interrelated variables. They are a fundamental tool in the study of multivariable calculus and are crucial in various scientific and engineering applications.

2 Why Partial Derivatives were invented?

Partial derivatives were not "invented" in the traditional sense, but rather they emerged as a natural mathematical concept to address the need for analyzing functions of multiple variables. They were developed as a tool to make sense of and work with functions that depend on more than one independent variable.

The concept of a derivative, which measures the rate of change of a function with respect to a single variable, dates back to the early development of calculus in the 17th century, primarily associated with the work of Isaac Newton and Gottfried Wilhelm Leibniz. These early developments focused on functions of a single variable, which were sufficient for many problems in physics and mathematics at the time.

However, as scientific and engineering disciplines advanced, the need to work with functions that depended on multiple independent variables became apparent. For example, problems in physics, like describing the motion of objects in three-dimensional space or analyzing heat conduction in materials, naturally led to the development of functions with multiple variables.

Partial derivatives were introduced to address this new mathematical challenge. They allow mathematicians and scientists to study how a function changes concerning one specific independent variable while keeping the others constant. This concept was further developed and formalized by mathematicians like Augustin-Louis Cauchy and Simon Denis Poisson in the 19th century.

In summary, partial derivatives were not "invented" but rather developed as a mathematical tool to extend the principles of calculus to functions with multiple independent variables, enabling a deeper understanding of the behavior of systems in various fields of science and engineering. They emerged as a natural extension of the fundamental concept of derivatives when dealing with more complex mathematical and real-world problems.

3 Partial Derivative in Deep Learning

Partial derivatives are important in deep learning for optimization

3.1 Gradient Descent and Backpropogation

Deep learning models are typically trained using optimization algorithms like gradient descent. These algorithms rely on partial derivatives to find the optimal model parameters that minimize a loss function and are calculated by backpropogation. Refer to this link https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/ for detailed understanding with example solution.

Backpropagation is the process of computing gradients efficiently within a neural network, while gradient descent is the optimization algorithm that uses these gradients to adjust the network's parameters to minimize the loss function. These two concepts work together to train neural networks effectively, but they serve different purposes.