# Partial Derivatives

## 1 Introduction

Many real-world functions depend on more than one variable, for example

$$f(x, y), \quad f(x, y, z), \quad f(x_1, x_2, \ldots, x_n).$$

When we ask: *"How does the function change if only one variable changes while all the others stay constant?"* we use a **partial derivative**.

Partial derivatives are fundamental in optimization, machine learning, deep learning (backpropagation), and multivariable calculus.

## 2 Definition of a Partial Derivative

Let $f(x, y)$ be a function of two variables.

The partial derivative of $f$ with respect to $x$ at $(x, y)$ is defined as

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

Similarly, the partial derivative with respect to $y$ is

$$\frac{\partial f}{\partial y}(x, y) = \lim_{k \to 0} \frac{f(x, y + k) - f(x, y)}{k}.$$

**Key idea:** When taking $\dfrac{\partial f}{\partial x}$, treat $y$ as a constant. When taking $\dfrac{\partial f}{\partial y}$, treat $x$ as a constant.

## 3 A Real-Life Intuition for Partial Derivatives

Consider the function

$$f(x, y) = 3x + 4y,$$

where

- $x$ = hours studied,
- $y$ = hours slept,
- $f(x, y)$ = performance score.

This means your performance depends on two factors: how long you study and how much you sleep.

## Interpretation of Partial Derivatives

The partial derivative with respect to $x$ is

$$\frac{\partial f}{\partial x} = 3.$$

This represents the change in your score when you *increase study hours*, while keeping sleep fixed.

Similarly,

$$\frac{\partial f}{\partial y} = 4,$$

which represents the change in your score when you *increase sleep hours*, while keeping study fixed.

## Meaning

- If you study **1 extra hour** (and sleep the same), your score increases by **3 points**.
- If you sleep **1 extra hour** (and study the same), your score increases by **4 points**.

## Why this helps understanding

This example shows that:

- Partial derivatives measure the sensitivity of the function to each variable.
- They tell us which variable influences the outcome more.
- In this example, sleep ($y$) has a slightly stronger effect on performance than study time ($x$).

# 4 Basic Rules

Let $c$ be a constant and $f, g$ be functions of $x, y$.

## Linearity

$$\frac{\partial}{\partial x}(cf + g) = c\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}.$$

## Power rule (in one variable)

$$\frac{\partial}{\partial x}(x^n) = nx^{n-1} \quad \text{(treating other variables as constants).}$$

## Product rule

If $h(x, y) = f(x, y)g(x, y)$, then

$$\frac{\partial h}{\partial x} = f\frac{\partial g}{\partial x} + g\frac{\partial f}{\partial x}.$$

### Chain rule

If $u = g(x, y)$ and $f = f(u)$, then

$$\frac{\partial f}{\partial x} = \frac{df}{du} \cdot \frac{\partial u}{\partial x}.$$

# 5 Worked Examples

## 5.1 Example 1: Simple Polynomial Function

Consider

$$f(x, y) = 3x^2 y + 5y.$$

#### Partial derivative with respect to $x$

Treat $y$ as a constant:

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(3x^2 y) + \frac{\partial}{\partial x}(5y) \\
&= 3y \cdot \frac{\partial}{\partial x}(x^2) + 5 \cdot \frac{\partial}{\partial x}(y) \\
&= 3y \cdot 2x + 5 \cdot 0 \\
&= 6xy.
\end{aligned}$$

So

$$\boxed{\frac{\partial f}{\partial x} = 6xy.}$$

#### Partial derivative with respect to $y$

Treat $x$ as a constant:

$$\begin{aligned}
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}(3x^2 y) + \frac{\partial}{\partial y}(5y) \\
&= 3x^2 \cdot \frac{\partial}{\partial y}(y) + 5 \cdot \frac{\partial}{\partial y}(y) \\
&= 3x^2 \cdot 1 + 5 \cdot 1 \\
&= 3x^2 + 5.
\end{aligned}$$

So

$$\boxed{\frac{\partial f}{\partial y} = 3x^2 + 5.}$$

## 5.2 Example 2: Function with a Square Root

Let

$$f(x, y) = \sqrt{x^2 + y}.$$

Set $u = x^2 + y$. Then $f = u^{1/2}$.

**Partial derivative with respect to $x$**

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}\left(u^{1/2}\right) = \frac{1}{2}u^{-1/2} \cdot \frac{\partial u}{\partial x} \\
u &= x^2 + y \Rightarrow \frac{\partial u}{\partial x} = 2x \\
\frac{\partial f}{\partial x} &= \frac{1}{2}(x^2 + y)^{-1/2} \cdot 2x \\
&= \frac{x}{\sqrt{x^2 + y}}.
\end{aligned}
$$

So

$$
\boxed{\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y}}.}
$$

**Partial derivative with respect to $y$**

$$
\begin{aligned}
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(u^{1/2}\right) = \frac{1}{2}u^{-1/2} \cdot \frac{\partial u}{\partial y} \\
u &= x^2 + y \Rightarrow \frac{\partial u}{\partial y} = 1 \\
\frac{\partial f}{\partial y} &= \frac{1}{2}(x^2 + y)^{-1/2} \cdot 1 \\
&= \frac{1}{2\sqrt{x^2 + y}}.
\end{aligned}
$$

So

$$
\boxed{\frac{\partial f}{\partial y} = \frac{1}{2\sqrt{x^2 + y}}.}
$$

## 5.3 Example 3: Exponential Function

Let

$$
f(x, y) = e^{xy}.
$$

Set $u = xy$, so $f = e^u$.

**Partial derivative with respect to $x$**

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= e^{xy} \cdot \frac{\partial}{\partial x}(xy) \\
&= e^{xy} \cdot y \\
&= ye^{xy}.
\end{aligned}
$$

So

$$
\boxed{\frac{\partial f}{\partial x} = ye^{xy}.}
$$

**Partial derivative with respect to $y$**

$$\begin{aligned}\frac{\partial f}{\partial y} &= e^{xy} \cdot \frac{\partial}{\partial y}(xy) \\ &= e^{xy} \cdot x \\ &= xe^{xy}.\end{aligned}$$

So

$$\boxed{\frac{\partial f}{\partial y} = xe^{xy}.}$$

# 6 Higher-Order Partial Derivatives

Once we have first partial derivatives, we can differentiate again.

## Second-order partial derivatives

For $f(x, y)$, we define

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

## Mixed partial derivatives

We also define

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

Under suitable smoothness conditions (Clairaut's theorem),

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

## Example: Higher-order derivatives

Let

$$f(x, y) = x^2 y + y^3.$$

**First derivatives**

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2xy, \\ \frac{\partial f}{\partial y} &= x^2 + 3y^2.\end{aligned}$$

**Second-order derivatives**

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}(2xy) = 2y, \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}(x^2 + 3y^2) = 6y.\end{aligned}$$

**Mixed derivatives**

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}(x^2 + 3y^2) = 2x,$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}(2xy) = 2x.$$

We see the mixed partials are equal:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2x.$$

# 7 Geometric Interpretation

A function $f(x, y)$ defines a surface $z = f(x, y)$ in 3D.

- $\dfrac{\partial f}{\partial x}$ is the slope of the surface in the $x$-direction (with $y$ fixed).
- $\dfrac{\partial f}{\partial y}$ is the slope of the surface in the $y$-direction (with $x$ fixed).

You can think of taking a slice of the surface along a line parallel to the $x$-axis or $y$-axis and then computing an ordinary derivative on that slice.

# 8 Gradient Vector

For a function of two variables $f(x, y)$, the **gradient** is

$$\nabla f(x, y) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

For a function of many variables $f(x_1, x_2, \ldots, x_n)$,

$$\nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right).$$

**Important fact:** The gradient points in the direction of the *steepest increase* of the function.

In machine learning, if $L(w)$ is a loss function depending on parameters $w$, the gradient descent update is

$$w := w - \eta \nabla L(w),$$

where $\eta$ is the learning rate.

# 9 Applications in Machine Learning

- Computing gradients of loss functions with respect to weights and biases.
- Backpropagation in neural networks (repeated use of partial derivatives and chain rule).
- Sensitivity analysis: understanding which features affect the output most.
- Optimization algorithms (gradient descent, Adam, RMSProp, etc.).

## 10 Practice Problems

1. For $f(x,y) = x^3y^2 + 4y$, compute $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$.

2. For $f(x,y) = e^{3x+y}$, compute $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$.

3. For $f(x,y) = \ln(x^2 + y^2)$, compute both partial derivatives.

4. For $f(x,y) = x^2y + y^3$, compute $\dfrac{\partial^2 f}{\partial x^2}$, $\dfrac{\partial^2 f}{\partial y^2}$, and the mixed partials.

5. For $f(x,y,z) = xyz + x^2z$, find the gradient $\nabla f$.