14.3: Partial Derivatives,
14.4: The Chain Rule, and
14.5: Directional Derivatives and Gradient Vectors

TA: Sam Fleischer

November 10

Section 14.3: Partial Derivatives

Definition: Partial Derivative with respect to x, y

The **partial derivative of $f(x,y)$ with respect to x at the point** (x_0, y_0) is

$$f_x = \left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \left.\frac{\mathrm{d}}{\mathrm{d}x} f(x, y_0)\right|_{(x_0,y_0)} = \lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

The **partial derivative of $f(x,y)$ with respect to y at the point** (x_0, y_0) is

$$f_y = \left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = \left.\frac{\mathrm{d}}{\mathrm{d}y} f(x_0, x)\right|_{(x_0,y_0)} = \lim_{h\to 0} \frac{f(x_0, y_0+h) - f(x_0, y_0)}{h}$$

Example: Taking Partial Derivatives

$$f(x,y) = y\sin(xy)$$

To find f_x (the partial derivative of f with respect to x), treat y as a *constant*.

$$f_x = y\cos(xy)y = y^2\cos(xy)$$

To find f_y (the partial derivative of f with respect to y), treat x as a *constant*.

$$f_y = (1)\sin(xy) + y\cos(xy)x = \sin(xy) + xy\cos(xy)$$

Example: Implicit Differentiation

Find $\dfrac{\partial z}{\partial x}$ if the equation

$$yz - \ln z = x + y$$

defines z as a function of two independent variables x and y and the partial derivative exists.

$$\frac{\mathrm{d}}{\mathrm{d}x}(yz - \ln z) = \frac{\mathrm{d}}{\mathrm{d}x}(x+y)$$

$$
\begin{aligned}
y\frac{\partial z}{\partial x} - \frac{1}{z}\frac{\partial z}{\partial x} &= 1 + 0 \\
\left(y - \frac{1}{z}\right)\frac{\partial z}{\partial x} &= 1 \\
\frac{\partial z}{\partial x} &= \frac{x}{yz - 1}
\end{aligned}
$$

Definition: Second-Order Partial Derivatives

When we differentiate a function $f(x, y)$ twice, we produce its **second-order derivatives**. Notation:

$$
\frac{\partial^2 f}{\partial x^2} \text{ or } f_{xx}, \qquad \frac{\partial^2 f}{\partial y^2} \text{ or } f_{yy}
$$
$$
\frac{\partial^2 f}{\partial x \partial y} \text{ or } f_{yx}, \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial x} = f_{xy}
$$

The defining equations are

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

Note that $f_{yx} = \dfrac{\partial^2 f}{\partial x \partial y}$ means you *first* take the derivative with respect to y, and *then* take the derivate with respect to x.

In general, $f_{xy} \neq f_{yx}$.

Theorem 2: Mixed Derivative Theorem

If $f(x, y)$ and its partial derivatives f_x, f_y, f_{xy}, and f_{yx} are defined throughout an open region containing a point (a, b) and are all continuous at (a, b), then

$$
f_{xy}(a, b) = f_{yx}(a, b)
$$

Example: Partial Derivatives of Higher Order

Find f_{yxyz} if $f(x, y, z) = 1 - 2xy^2z + x^2y$

$$
\begin{aligned}
f(x, y, z) &= 1 - 2xy^2z + x^2y \\
f_y &= -4xyz + x^2 \\
f_{yx} &= -4yz + 2x \\
f_{yxy} &= -4z \\
f_{yxyz} &= -4
\end{aligned}
$$

Definition: Differentiability

A function $z = f(x, y)$ is **differentiable at** $(x_0, y_)$ if $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exists and $\mathbf{\Delta} z$ satisfies an equation of the form

$$\mathbf{\Delta} z = f_x(x_0, y_0)\mathbf{\Delta} x + f_y(x_0, y_0)\mathbf{\Delta} y + \epsilon_1 \mathbf{\Delta} x + \epsilon_2 \mathbf{\Delta} y$$

in which each of $\epsilon_1, \epsilon_2 \to 0$ as both $\mathbf{\Delta} x, \mathbf{\Delta} y \to 0$. We call f **differentiable** if it is differentiable at every point in its domain, and say that its graph is a **smooth surface**.

Theorem 3: The Increment Theorem for Functions of Two Variables

Suppose that the first partial derivatives of $f(x, y)$ are defined throughout an open region R containing the point (x_0, y_0) and that f_x and f_y are continuous at (x_0, y_0). Then the change

$$\mathbf{\Delta} z = f(x_0 + \mathbf{\Delta} x, y_0 + \mathbf{\Delta} y) - f(x_0, y_0)$$

in the value of f that results from moving from (x_0, y_0) to another point $(x_0 + \mathbf{\Delta} x, y_0 + \mathbf{\Delta} y)$ in R satisfies an equation of the form

$$\mathbf{\Delta} z = f_x(x_0, y_0)\mathbf{\Delta} x + f_y(x_0, y_0)\mathbf{\Delta} y + \epsilon_1 \mathbf{\Delta} x + \epsilon_2 \mathbf{\Delta} y$$

in which each of $\epsilon_1, \epsilon_2 \to 0$ as both $\mathbf{\Delta} x, \mathbf{\Delta} y \to 0$.

Corollary of Theorem 3

If the partial derivatives f_x and f_y of a function $f(xy)$ are continuous throughout an open region R, then f is differentiable at every point of R.

Theorem 4 - DIfferentiability Implies Continuity

If a function $f(x, y)$ is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

Section 14.4: The Chain Rule

Theorem 5: Chain Rule for Functions of One Independent Variable and Two Intermediate Variables

If $w = f(x, y)$ is differentiable and if $x = x(t)$, $y = y(t)$ are differentiable functions of t, then the composite $w = f(x(t), y(t))$ is a differentiable function of t and

$$\frac{\mathrm{d}w}{\mathrm{d}t} = f_x(x(t), y(t)) \cdot x'(t) + f_y(x(t), y(t)) \cdot y'(t)$$

or

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

Theorem 6: Chain Rule for Functions of One Independent Variable and Three Intermediate Variables

If $w = f(x, y, z)$ is differentiable and x, y, and z are differentiable functions of t, then w is a differentiable function of t and

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t}$$

Example

Find $\dfrac{\mathrm{d}w}{\mathrm{d}t}$ if

$$w = xy + z, \qquad x = \cos t, \qquad y = \sin t, \qquad z = t$$

Use Theorem 6:

$$\begin{aligned}
\frac{\mathrm{d}w}{\mathrm{d}t} &= \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} \\
&= (y)(-\sin t) + (x)(\cos t) + (1)(1) \\
&= (\sin t)(-\sin t) + (\cos t)(\cos t) + 1 \\
&= -\sin^2 t + \cos^2 t + 1 \\
&= \cos 2t + 1
\end{aligned}$$

What is the derivative at $t = 0$?

$$\left.\left(\frac{\mathrm{d}w}{\mathrm{d}t}\right)\right|_{t=0} = 1 + \cos 0 = 2$$

Theorem 7: Chain Rule for Two Independent Variables and Three Intermediate Variables

Suppose that $w = f(x, y, z)$, $x = g(r, s)$, $y = h(r, s)$, and $z = k(r, s)$. If all four functions are differentiable, then w has partial derivatives with respect to r and s given by the formulas

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}$$

Theorem 8: A Formula for Implicit Differntiation

Suppose that $F(x, y)$ is differentiable and that the equation $F(x, y) = 0$ defines y as a differentiable function of x. Then at any point where $F_y \neq 0$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{F_x}{F_y}$$

Example

Find $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ if $y^2 - x^2 - \sin xy = 0$.

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{F_x}{F_y} \\
&= -\frac{-2x - y\cos xy}{2y - x\cos xy} \\
&= \frac{2x + y\cos xy}{2y - x\cos xy}
\end{aligned}$$

Expansion of Theorem 8 to Three Variables

Suppose $F(x, y, z) = 0$ and $z = f(x, y)$. Assuming F and f are differentiable functions, then

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} \qquad \text{and} \qquad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

Expansion of Chain Rule to Functions of n variables

In general, suppose $z = f(x_1, x_2, \ldots, x_n)$ is a differential function of the intermediate variables $x_1, x_2, \ldots, x_n$ where n is a positive integer ($n \in \mathbb{N}$). Also suppose each x_i is a differentiable function of the independent variables $t_1, t_2, \ldots, t_m$, with $m \in \mathbb{N}$. In equation form,

$$\begin{aligned}
x_1 &= g_1(t_1, t_2, \ldots, t_m) \\
x_2 &= g_2(t_1, t_2, \ldots, t_m) \\
&\vdots \\
x_n &= g_n(t_1, t_2, \ldots, t_m)
\end{aligned}$$

Then w is a differential function of each of the independent variables $t_1, t_2, \ldots, t_m$, and the partial derivatives of w with respect to each t_i are

$$\begin{aligned}
\frac{\partial w}{\partial t_1} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_1} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i}\frac{\partial x_i}{\partial t_1} \\
\frac{\partial w}{\partial t_2} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_2} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i}\frac{\partial x_i}{\partial t_2} \\
&\vdots \\
\frac{\partial w}{\partial t_m} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_m} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_m} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_m} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i}\frac{\partial x_i}{\partial t_m}
\end{aligned}$$

More compactly,

$$\frac{\partial w}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i}\frac{\partial x_i}{\partial t_j} \qquad \text{for } j = 1, 2, \ldots, m$$

Directional Derivatives and Gradient Vectors

Definition: Directional Derivative

The **derivative of f at $P_0(x_0, y_0)$ in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} = u_2\mathbf{j}$** is the number

$$\left(\frac{\mathrm{d}f}{\mathrm{d}s}\right)_{\mathbf{u},P_0} = \lim_{s\to 0} \frac{f(x_0 + su_1, y_0 + su_2) - f(x_0, y_0)}{s}$$

provided the limit exists. The directional derivative is also denoted

$$\left(\frac{\mathrm{d}f}{\mathrm{d}s}\right)_{\mathbf{u},P_0} = (D_{\mathbf{u}}f)_{P_0}$$

and is read "The derivative of f at P_0 in the direction of $\mathbf{u}$".

Definition: Gradient Vector

The **gradient vector (gradient)** of $f(x, y)$ at a point P is the vector

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

Theorem 9: The Directional Derivative is a Dot Product

If $f(x, y)$ is differentiable in an open region containing $P_0(x_0, y_0)$, then

$$(D_{\mathbf{u}}f)_{P_0} = (\boldsymbol{\nabla} f)_{P_0} \cdot \mathbf{u}$$

In words, the derivative of f at P_0 in the direction of $\mathbf{u}$ is the dot product of the gradient $\boldsymbol{\nabla} f$ at P_0 and $\mathbf{u}$. In brief,

$$(D_{\mathbf{u}})f = \boldsymbol{\nabla} f \cdot \mathbf{u}$$

Example

Find the derivative of $f(x, y) = xe^y + \cos xy$ at the point $(2, 0)$ in the direction of $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$.

First, find the unit direction vector

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{5} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}$$

Then we need to find the partial derivatives of f at $(2, 0)$ because together, they make up the gradient, $\boldsymbol{\nabla} f$.

$$\begin{aligned} f_x &= e^y - y\sin xy \\ f_x(2,0) &= e^0 - 0\sin(2\cdot 0) = 1 - 0 = 1 \\ f_y &= xe^y - x\sin xy \\ f_y(2,0) &= 2e^0 - 2\sin(2\cdot 0) = 2 - 0 = 2 \end{aligned}$$

Plug these values into the definition of gradient.

$$\begin{aligned} \nabla f|_{(2,0)} &= f_x(2,0)\mathbf{i} + f_y(2,0)\mathbf{j} \\ &= \mathbf{i} + 2\mathbf{j} \end{aligned}$$

Then the directional derivative of f at $(2,0)$ in the direction of u is

$$\begin{aligned} (D_{\mathbf{u}} f)_{(2,0)} &= \nabla f|_{(2,0)} \cdot \mathbf{u} \\ &= \left(\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\right) \cdot (\mathbf{i} + 2\mathbf{j}) \\ &= \frac{3}{5} - 2 \cdot \frac{4}{5} \\ &= -1 \end{aligned}$$

Properties of the Directional Derivative

1. The function f increases most rapidly when $\cos\theta = 1$, i.e. when $\theta = 0$, i.e. when $\mathbf{u}$ is the direction of ∇f. That is, *at each point P in its domain, f increases most rapidly in the direction of the gradient vector ∇f at P.* The derivative in this direction is

$$D_{\mathbf{u}} f = \|\nabla f\| \cos 0 = \|\nabla f\|$$

2. Similarly, f decreases most rapidly in the direction of $-\nabla f$. The derivative in this direction is

$$D_{\mathbf{u}} f = \|\nabla f\| \cos \pi = -\|\nabla f\|$$

3. Any direction $\mathbf{u}$ orthogonal to a gradient $\nabla f \neq 0$ is adirection of zero change in f because $\theta = \dfrac{\pi}{2}$ and

$$D_{\mathbf{u}} f = \|\nabla f\| \cos \frac{\pi}{2} = \|\nabla f\| \cdot 0 = 0$$

Example

Let $f(x,y) = \dfrac{x^2}{2} + \dfrac{y^2}{2}$, and consider the point $(1,1)$.

The function increases most rapidly in the direction of ∇f.

$$(\nabla f) = x\mathbf{i} + y\mathbf{j} \implies (\nabla f)_{(1,1)} = \mathbf{i} + \mathbf{j}$$

The unit vector of $(\nabla f)_{(1,1)}$ is

$$\mathbf{u} = \frac{\mathbf{i} + \mathbf{j}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}$$

The function *decreases* most rapidly in the direction $-(\nabla f)_{(1,1)}$

$$-\frac{1}{\sqrt{2}}\mathbf{i} - \frac{1}{\sqrt{2}}\mathbf{j}$$

The directions of zero change at $(1,1)$ are the directions ofthogonal to ∇f:

$$\mathbf{n} = -\frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j} \qquad \text{and} \qquad -\mathbf{n} = \frac{1}{\sqrt{2}}\mathbf{i} - \frac{1}{\sqrt{2}}\mathbf{j}$$

Important Concept

At every point (x_0, y_0) in the domain of a differentiable function $f(x, y)$, the gradient of f is normal to the level curve through (x_0, y_0).

Tangent Line to a Level Curve

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$$

Notice this is the same as point-slope form from elementary algebra.

$$y - y_0 = m(x - x_0)$$

where

$$m = -\frac{f_x}{f_y} = \frac{\mathrm{d}y}{\mathrm{d}x}$$

by Theorem 8.

Algebra Rules for Gradients

1.	*Sum Rule:*	$\boldsymbol{\nabla}(f + g) = \boldsymbol{\nabla} f + \boldsymbol{\nabla} g$
2.	*Difference Rule:*	$\boldsymbol{\nabla}(f - g) = \boldsymbol{\nabla} f - \boldsymbol{\nabla} g$
3.	*Constant Multiple Rule:*	$\boldsymbol{\nabla}(kf) = k\boldsymbol{\nabla} f$
4.	*Product Rule:*	$\boldsymbol{\nabla}(fg) = f\boldsymbol{\nabla} g + g\boldsymbol{\nabla} f$
5.	*Quotient Rule:*	$\boldsymbol{\nabla}\left(\frac{f}{g}\right) = \frac{g\boldsymbol{\nabla} f - f\boldsymbol{\nabla} g}{g^2}$

Gradients of Functions of n variables

For a differential function $f(x_1, x_2, \dots x_n)$ and a unit vector $\mathbf{u} = \langle u_1, u_2, \dots, u_n\rangle$ in space, we have

$$\boldsymbol{\nabla} f = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right\rangle$$

and

$$D_{\mathbf{u}} f = \boldsymbol{\nabla} f \cdot \mathbf{u} = \frac{\partial f}{\partial x_1} u_1 + \frac{\partial f}{\partial x_2} u_2 + \dots \frac{\partial f}{\partial x_n} u_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} u_i$$

The Derivative Along a Path

Let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ be a smooth path C and $w = f(\mathbf{r}(t))$ a scalar function along C. Then

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t}$$

or in vector notation,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\mathbf{u}(t)) = \boldsymbol{\nabla} f(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$$