NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: MTH 311

COURSE TITLE: CALCULUS OF SEVERAL VARIABLES

MTH 311 **CALCULUS OF SEVERAL VARIABLES**

COURSE MATERIAL

NATIONAL OPEN UNIVERSITY OF NIGERIA

Course Code MTH 311

Course Title **CALCULUS OF SEVERAL VARIABLES**

Course Writers Mr. Toyin Olorunnishola
And
Dr. Ajibola Oluwatoyin
School of Science and Technology
National Open University of Nigeria

Course Editing Team

NATIONAL OPEN UNIVERSITY OF NIGERIA

COURSE TITLE: CALCULUS OF SEVERAL VARIABLES

COURSE CODE: **M T H 311**

TABLE OF CONTENT

MODULE 1 Limit and Continuity of Functions of Several Variables

- Unit 1: Real Functions
- Unit 2: Limit of Function of Several Variables.
- Unit 3: Continuity of Function of Several Variables.

MODULE 2 PARTIAL DERIVATIVES OF FUNCTION OF SEVERAL VARIABLES

- Unit 1: Derivative
- Unit 2: Partial derivative.
- Unit 3: Application of Partial derivative.

MODULE 3 TOTAL DERIVATIVE OF A FUNCTION.

- Unit 1: Derivation of a function.
- Unit 2: Total derivative of a function.
- Unit 3: Application of total derivative of a function.

MODULE 4 PARTIAL DIFFERENTIABILITY AND TOTAL DIFFERENTIABILITY OF FUNCTION OF SEVERAL VARIABLE

- Unit 1: Partial differentials of function of several variables.
- Unit 2: Total differentials of function of several variables.
- Unit 3:Application of partial and total differentials of function of several variables.

MODULE 5 COMPOSITE DIFFERENTIATION, FULLER'S THEOREM, IMPLICIT DIFFERENTIATION.

- Unit 1: Composite differentiation
- Unit 2: Fuller's Theorem
- Unit 3: Implicit differentiation.

MODULE 6 TAYLOR'S SERIES EXPANSION

- Unit 1: Function of two variables
- Unit 2: Taylor's series expansion for functions of two variables.
- Unit 3: Application of Taylor's series.

MODULE 7 MAXIMISATION AND MINIMISATION OF FUNCTIONS OF SEVERAL VARIABLES

- Unit 1Maximisation and Minimisation Of Functions Of Several Variables.
- Unit 2: Lagrange’s Multipliers.
- Unit 3: Application of Lagrange’s Multipliers.

MODULE 8 THE JACOBIANS

- Unit 1: Jacobians
- Unit 2: Jacobian determinants
- Unit 3: Applications of Jacobian

MODULE 1 Limit and Continuity of Functions of Several Variables

Unit 1: Real Functions

Unit 2: Limit of Function of Several Variables.

Unit 3: Continuity of Function of Several Variables.

UNIT 1 REAL FUNCTION

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 domain
  3.2 real function
  3.3 value of functions
  3.4 types of graph
  3.5 types of function
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

INTRODUCTION

A real-valued function, f, of x, y, z, ... is a rule for manufacturing a new number, written f(x, y, z, ...), from the values of a sequence of independent variables (x, y, z, ...).

The function f is called a real-valued function of two variables if there are two independent variables, a real-valued function of three variables if there are three independent variables, and so on.

As with functions of one variable, functions of several variables can be represented numerically (using a table of values), algebraically (using a formula), and sometimes graphically (using a graph).

Examples

1. $f(x, y) = x - y$ — Function of two variables

$f(1, 2) = 1 - 2 = -1$ — Substitute 1 for x and 2 for y f(2,˜ 1) = 2˜ ˜(1) = 3 — Substitute 2 for x and˜ 1 for y

f(y, x) = y˜ x — Substitute y for x and x for y

2. h(x, y, z) = x + y + xz — Function of three variables

h(2, 2,˜ 2) = 2 + 2 + 2˜(2) = — Substitute 2 for x, 2 for y, and˜ 2 for z.

OBJECTIVES

At the end of this unit, you should be able to know:

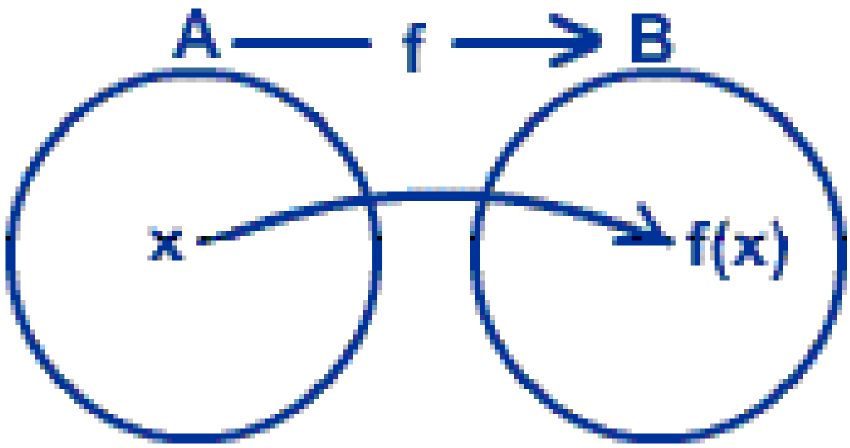
- domain
- real function
- value of functions
- types of graph
- types of function

MAIN CONTENT

f is a function from set A to a set B if each element x in A can be associated with a unique element in B.

Usually written as $f: A \to B$

The unique element B which f associates with x in A denoted by f (x).



Domain

In the above definition of the function, set A is called domain.

Co-domain

In the above definition of the function, set B is called co-domain.

Real Functions

A real valued function f : A to B or simply a real function 'f ' is a rule which associates to each possible real number $x \in A$, a unique real number $f(x) \in B$, when A and B are subsets of R, the set of real numbers.

In other words, functions whose domain and co-domain are subsets of R, the set of real numbers, are called real valued functions.

Value of a Function

If 'f ' is a function and x is an element in the domain of f, then image

f(x) of x under f is called the value of 'f ' at x.

Types of Functions and their Graphs

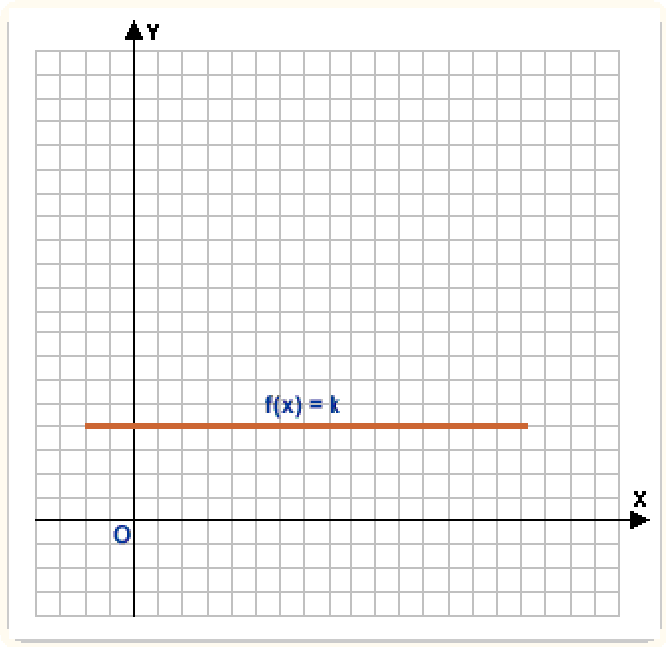
Constant Function

A function f : A ® B Such that A, B Ì R, is said to be a constant function if there exist K Î B such that f(x) = k.

Domain = A

Range = {k}

The graph of this function is a line or line segment parallel to x-axis. Note that, if k>0, the graph B is above X-axis. If k<0, the graph is below the x-axis. If k = 0, the graph is x-axis itself.

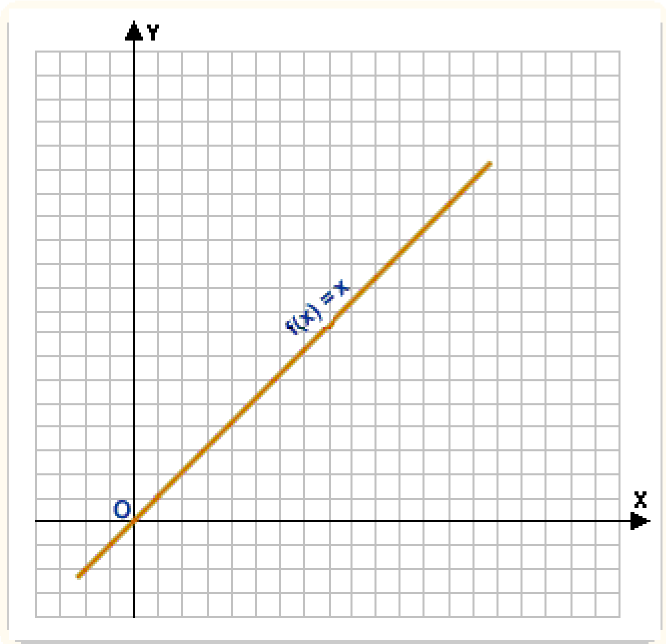


Identity Function

A function f : R® R is said to be an identity function if for all x Î R, f(x) = x.

Domain = R

Range = R



Polynomial Function

A function f : R® R is said to be a polynomial function if for each x Î R, f(x) is a polynomial in x.

$f(x) = x^3 + x^2 + x$

$g(x) = x^4 + 3x^2 + 2\sqrt{3}x + \sqrt{5}$ are examples of polynomial functions.

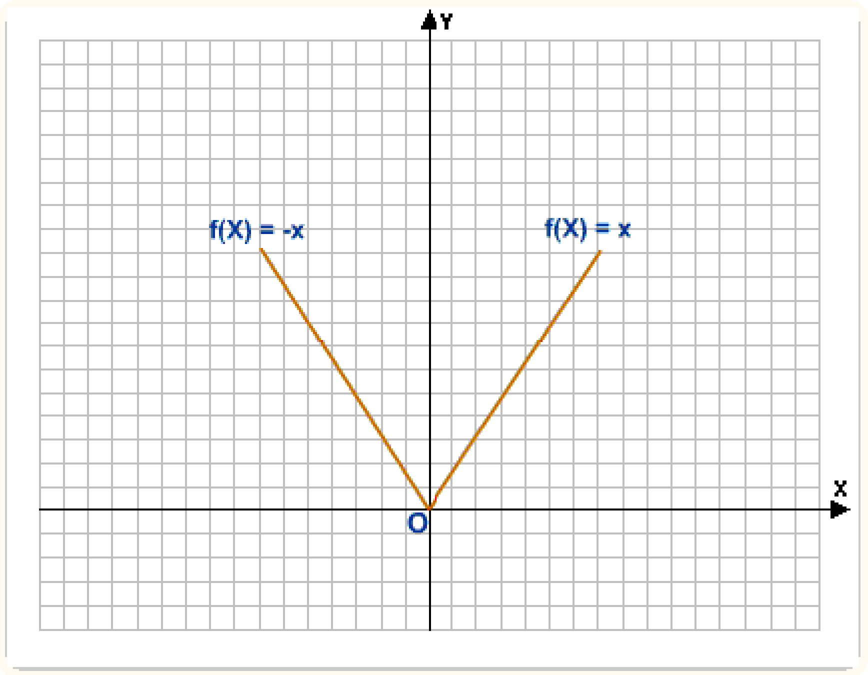
$h(x) = 3x^2 + \frac{2}{x}$ is not a polynomial function.

Modulus Function

f : R ® R such that f(x) = |x|, $f(x) = \begin{cases} x & x \geq 0 \\ -x & x < 0 \end{cases}$ is called the modulus function or absolute value function.

Domain = R

$\text{Range} = \begin{Bmatrix} x : x \geq 0 \\ x \in R \end{Bmatrix}$



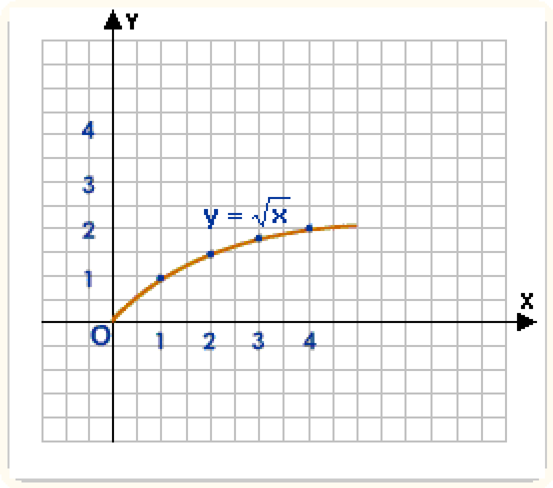
Square Root Function

Since square root of a negative number is not real, we define a function f : R^+ ® R such that

$f(x) = \sqrt{x}$

Domain of $f = R^+$ (set of all non-negative real numbers)

Range = R^+ (set of all non – negative real numbers)



Greatest Integer Function or Step Function (floor Function)

f (x) = [x] = greatest integer less than or equal to x

[x] = n, where n is an integer such that $n \leq x < n + 1$

Smallest Integer Function (ceiling Function)

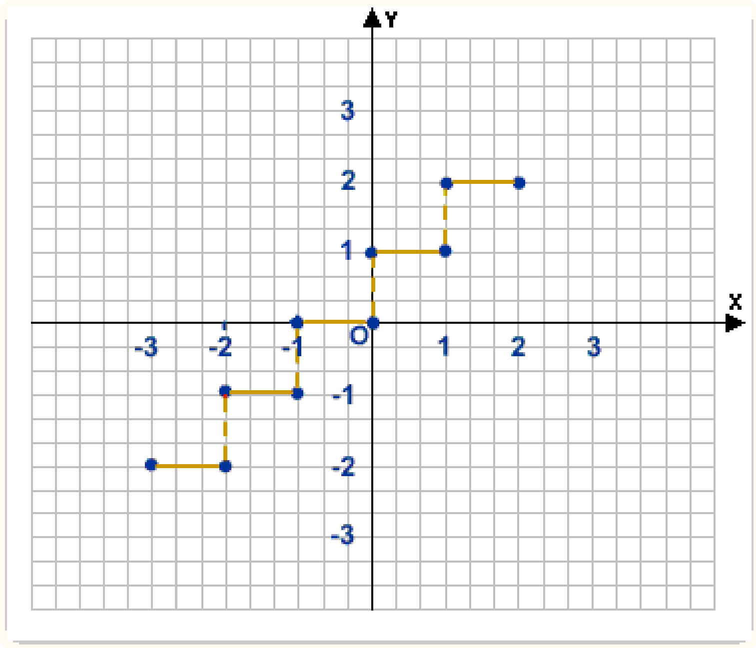
For a real number x, we denote by [x], the smallest integer greater than or equal to x. For example, [5 . 2] = 6, [-5 . 2] = -5, etc. The function f:R $\rightarrow$R defined by

f(x) = [x], x $\in$R

is called the smallest integer function or the ceiling function.

Domain: R

Range : Z

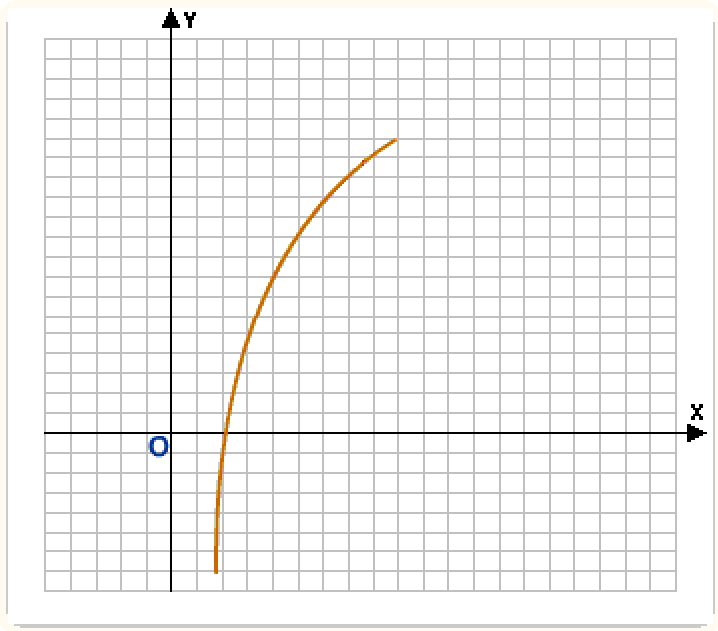


Exponential Function

The exponential function is defined as $f(x) = e^x$. Its graph is

Logarithmic Function

Logarithmic function is f (x) = log x. Its graph is



Trigonometric Functions

Trigonometric functions are sinx, cosx, tanx, etc. The graph of these functions have been done in class XI.

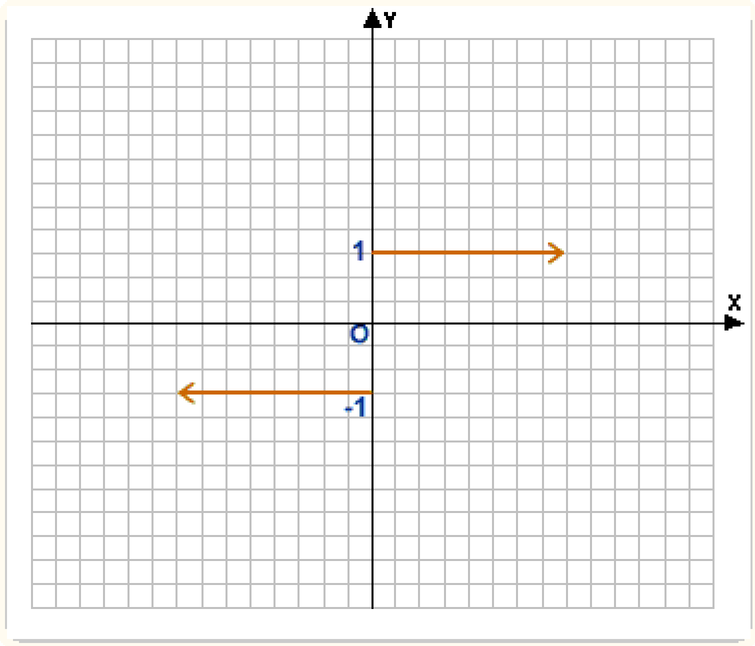
Inverse Functions

Inverse functions are $\sin^{-1}x$, $\cos^{-1}x$, $\tan^{-1}x$ etc. The graph of these functions have been done in class XI.

Signum Functions

$$f(x) = \begin{cases} \dfrac{|x|}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

$$\text{i.e., } f(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$



Odd Function

A function f : A →B is said to be an odd function if

f(x) = - f(-x) for all x ∈ A

The domain and range of f depends on the definition of the function.

Examples of odd function are

y = sinx, y = x^3, y = tanx

Even Function

A function f : A →B is said to be an even function if

f(x) = f(-x) for all x ∈ A.

The domain and range of f depends on the definition of the function.

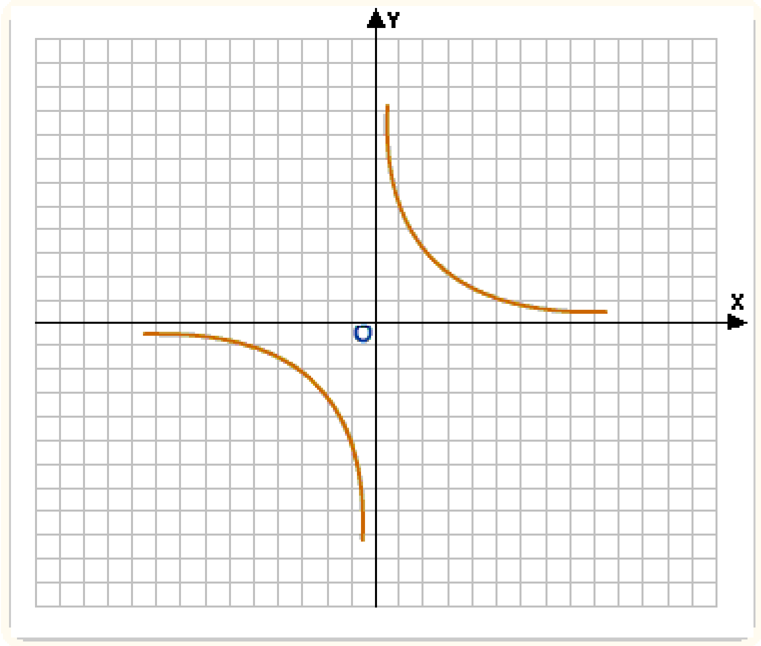
Examples of even function are

y = cosx, y = x^2, y = secx

A polynomial with only even powers of x is an even function.

Reciprocal Function

$$f(x) = \frac{1}{x} \quad x \neq 0$$



CONCLUSION

In this unit, you have defined domain and types of domain. You have known real functions and have also learnt value of functions. You have also known types of graph and type of function.

SUMMARY

In this unit, you have studied :

- domain
- real function
- value of functions
- types of graph
- types of function

TUTOR – MARKED ASSIGNMENT

1. Function f is defined by $f(x) = - 2 x^2 + 6 x - 3$. find f(- 2).
2. Function h is defined by $h(x) = 3 x^2 - 7 x - 5$. find h(x - 2).
3. Functions f and g are defined by $f(x) = - 7 x - 5$ and $g(x) = 10 x - 12$. find (f + g)(x)

4. Functions f and g are defined by f(x) = 1/x + 3x and g(x) = -1/x + 6x - 4 . find (f + g)(x) and its domain

5. Functions f and g are defined by f(x) = x^2 -2 x + 1 and g(x) = (x - 1)(x + 3) . find (f / g)(x) and its domain.

REFERENCES

Boas, Ralph P., Jr.: "A primer of real functions", The Carus Mathematical Monographs, No. 13; Published by The Mathematical Association of America, and distributed by John Wiley and Sons, Inc.; New York 1960 189 pp. MR22#9550

Smith, Kennan T.: "Primer of modern analysis", Second edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1983. 446 pp. ISBN 0-387-90797-1 MR84m:26002

Krantz, Steven G.; Parks, Harold R.: "A primer of real analytic functions", Basler Lehrbücher [Basel Textbooks], 4; Birkhäuser Verlag, Basel, 1992. 184 pp. ISBN 3-7643-2768-5 MR93j:26013

UNIT 2: Limit of Function of Several Variables

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
    3.1: Definition
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

1.0: INTRODUCTION

Let f be a function of two variables defined on a disk with center (a,b), except possibly at (a,b). Then we say that the **limit of f(x,y) as (x,y) approaches (a,b)** is L and we write

$$\lim_{(x,y)\to(a,b)} f(x,y) = L$$

If for every number ε > 0 there is a corresponding number δ > 0 such that

$$|f(x,y) - L| < \varepsilon \quad \text{whenever} \quad 0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$$

Other notations for the limit are

$\lim\limits_{\substack{x \to a \\ y \to b}} f(x,y) = L$ and $f(x,y) \to L$ as $(x,y) \to (a,b)$

Since $|f(x,y) - L|$ is the distance between the numbers f(x,y) and L, and $\sqrt{(x-a)^2 + (y-b)^2}$ is the distance between the point (x,y) and the point (a,b), Definition 12.5 says that the distance between f(x,y) and L can be made arbitrarily small by making the distance from (x,y) to (a,b) sufficiently small (but not 0). Figure 12.15 illustrates Definition 12.5 by means of an arrow diagram. If any small interval (L - ε, L +ε) is given around L, then we can find a disk D_δ with center (a,b) and radius δ > 0 such that f maps all the points in D_δ [except possibly (a,b)] into the interval (L - ε, L + ε).

2.0: OBJECTIVES

At this unit, you should be able to know the definition of terms.

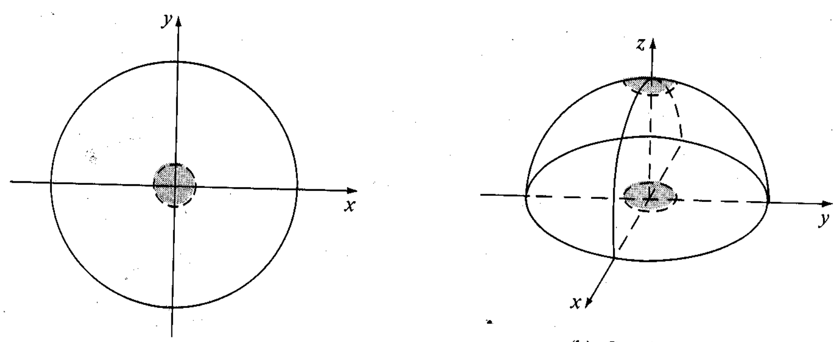
3.0: MAIN CONTENTS

Consider the function f(x,y) = $\sqrt{9 - x^2 - y^2}$ whose domain is the closed disk D = $\{(x,y) \mid x^2 + y^2 \le 9\}$ shown in Figure 12.14(a) and whose graph is the hemisphere shown in Figure 12.14(b)

If the point (x,y) is close to the origin, then x and y are both close to 0, and so f(x,y) is close to 3. In fact, if (x,y) lies in a small open disk $x^2 + y^2 < \delta^2$, then

f(x,y) = $\sqrt{9 - (x^2 + y^2)} > \sqrt{9 - \delta^2}$

Figure 12.14



(a) Domain of *f* (b) Graph of *f*

Thus we can make the values of f(x,y) as close to 3 as we like by taking (x,y) in a small enough disk with centre (0,0). We describe this situation by using the notation

$$\lim_{(x,y) \to (a,b)} \sqrt{9 - (x^2 + y^2)} = 3$$

In general, the notation

$\lim_{(x,y)\to(a,b)}$ f(x,y) = L

Means that the values of f(x,y) can be made as close as we wish to the number L by taking the point (x,y) close enough to the point (a,b). A more precise definition follows.

12.5 Definition

Let f be a function of two variables defined on a disk with centre (a,b), except possibly at (a,b). Then we say that the **limit of f(x,y) as (x,y) approaches (a,b)** is L and we write

$\lim_{(x,y)\to(a,b)}$ f(x,y) = L

If for every number ε > 0 there is a corresponding number δ > 0 such that

$|f(x,y) - L| < \varepsilon$ whenever $0 < \sqrt{(x-a)^2+(y-b)^2} < \delta$

Other notations for the limit are

$\lim_{\substack{x\to a\\ y\to b}} f(x,y) = L$ and $f(x,y) \to L$ as $(x,y) \to (a,b)$

Since $|f(x,y) - L|$ is the distance between the numbers f(x,y) and L, and $\sqrt{(x-a)^2+(y-b)^2}$ is the distance between the point (x,y) and the point (a,b), Definition 12.5 says that the distance between f(x,y) and L can be made arbitrarily small by making the distance from (x,y) to (a,b) sufficiently small (but not 0). Figure 12.15 illustrates Definition 12.5 by means of an arrow diagram. If any small interval (L - ε, L +ε) is given around L, then we can find a disk D_δ with center (a,b) and radius δ > 0 such that f maps all the points in D_δ [except possibly (a,b)] into the interval (L - ε, L + ε).

Another illustration of Definition 12.5 is given in Figure 12.16 where the surface S is the graph of f. If ε > 0 is given, we can find δ > 0 such that if (x,y) is restricted to lie in the disk D_δ and (x,y) ≠ (a,b), then the corresponding part of S lies between the horizontal planes z = L - ε and z = L + ε. For functions of a single variable, when we let x approach a, there are only two possible directions of approach, from the left or right. Recall from Chapter 2 that if $\lim_{x\to a^-} f(x) \neq \lim_{x\to a^+} f(x)$, then $\lim_{x\to a} f(x)$ does not exist.

For functions of two variables the situation is not as simple because we can let (x,y) approach (a,b) from an infinite number of directions in any manner whatsoever (see Figure 12.7).

Definition 12.5 refers only to the *distance* between (x,y) and (a,b). It does not refer to the direction of approach. Therefore if the limit exists, then f(x,y) must approach the same limit no matter how (x,y) approaches (a,b). Thus if we can find two different paths of approach along which f(x,y) has different limits, then it follows that $\lim_{(x,y)\to(a,b)} f(x,y)$ does not exist.

Figure 12.15

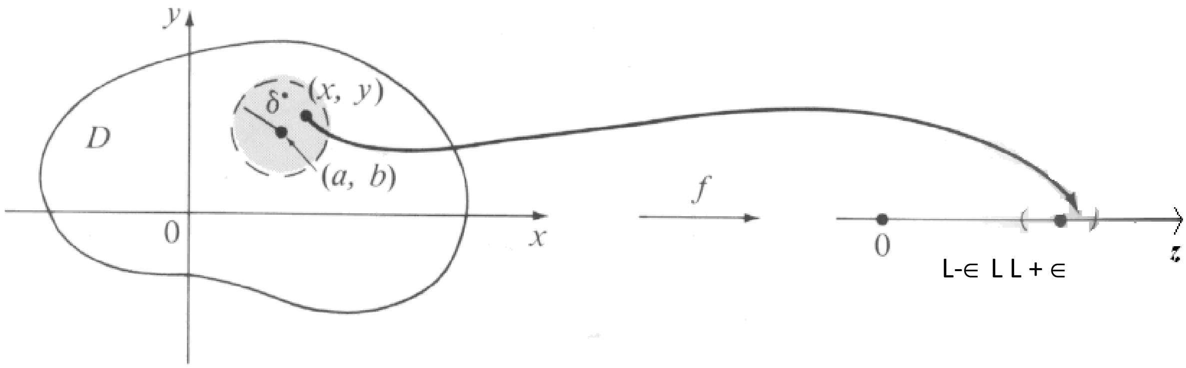


Figure 12.16

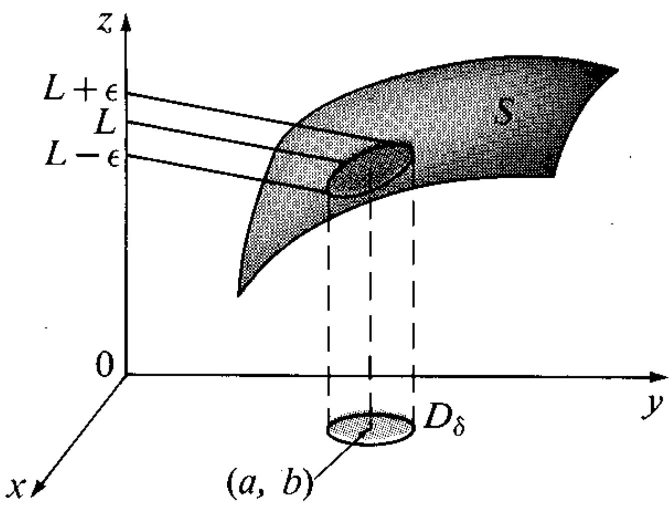
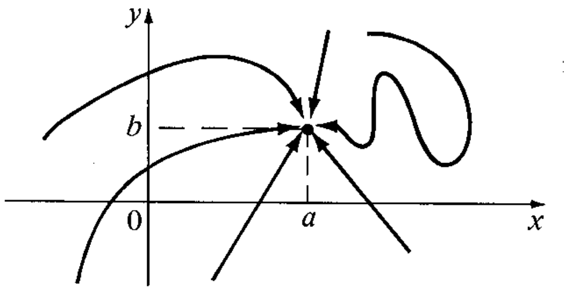


Figure 12.17



If $f(x,y) \to L_1$ as $(x,y) \to (a,b)$ along a path C_1, and $f(x,y) \to L_2$ as $(x,y) \to (a,b)$ along a path C_2, where $L_1 \neq L_2$, then $\lim_{(x,y)\to(a,b)} f(x,y)$

Example 1

Find $\lim_{(x,y)\to(0,0)} \dfrac{x^2 - y^2}{x^2 + y^2}$ if it exists.

Solution

Let $f(x,y) = (x^2 - y^2)/(x^2 + y^2)$. First let us approach (0,0) along the x-axis. Then $y = 0$ gives $f(x,0) = x^2/x^2 = 1$ for all $x \neq 0$, so

$f(x,y) \to 1$ as $(x,y) \to (0,0)$ along the x-axis

We now approach along the y-axis by putting $x = 0$. Then $f(0,y) = -y^2/y^2 = -1$ for all $y \neq 0$, so

$f(x,y) \to 1$ as $(x,y) \to (0,0)$ along the y-axis (see Figure 12.18.) Since f has two different limits along two different lines, the given limit does not exist.

Figure 12.18

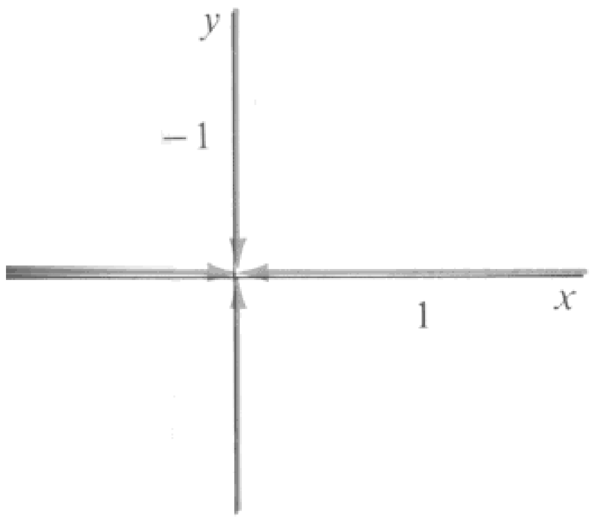
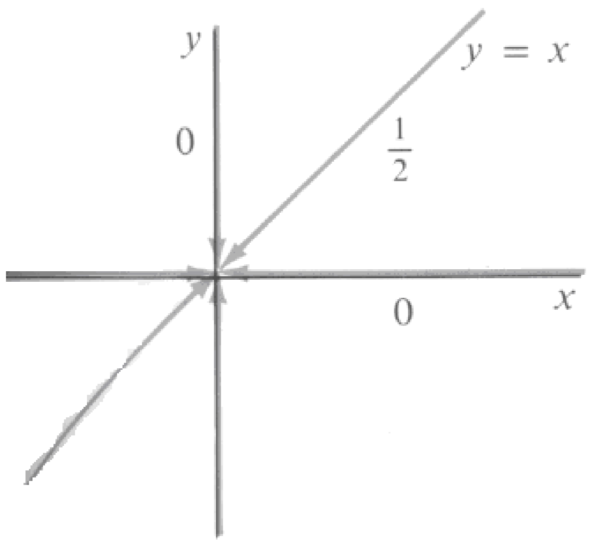


Figure 12.19



Example 2

If $f(x,y) = xy/(x^2 + y^2)$, does $\lim_{(x,y)\to(0,0)} f(x,y)$ exist?

Solution

If y = 0, then f(x,0) = $0/x^2 = 0$. Therefore

f(x,y) $\to$ 0 as (x,y) $\to$ (0,0) along the x-axis

If x = 0, then f(0,y) = $0/y^2 = 0$, so

f(x,y) $\to$ 0 as (x,y) $\to$ (0,0) along the y-axis

Although we have obtained identical limits along the axes, that does not show that the given limit is 0. Let us now approach (0,0) along another line, say y = x. For all x $\neq$ 0.

$$f(x,y) = \frac{x^2}{x^2 + x^2} = \frac{1}{2}$$

Therefore f(x,y) $\to \frac{1}{2}$ as (x,y) $\to$ (0,0) along y = x

(See Figure 12.19.) Since we obtained different limits along different paths, the given limit does not exist.

Example 3

If f(x,y) = $\frac{xy^2}{x^2 + y^4}$, does $\lim_{(x,y)\to(0,0)}$ f(x,y) exist?

Solution

With the solution of Example 2 in mind, let us try to save time by letting (x,y) $\to$ (0,0) along any line through the origin. Then y = mx, where m is the slope, and if m $\neq$ 0,

$$f(x,y) = f(x,mx) = \frac{x(mx)^2}{x^2 + (mx)^4} = \frac{m^2x^3}{x^2 + m^4x^4} = \frac{m^2x}{1 + m^4x^2}$$

So f(x,y) $\to$ 0 as (x,y) $\to$ (0,0) along y = mx

Thus f has the same limiting value along every line through the origin. But that does not show that the given limit is 0, for if we now let (x,y) $\to$ (0,0) along the parabola $x = y^2$ we have

$$f(x,y) = f(y^2,y) = \frac{y^2.y^2}{(y^2)^2 + y^4} = \frac{y^4}{2y^4} = \frac{1}{2}$$

so f(x,y) $\to \frac{1}{2}$ as (x,y) $\to$ (0,0) along $x = y^2$

Since different paths lead to different limiting values, the given limit does not exist.

Example 4

Find $\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^2 + y^2}$ if it exists.

Solution

As in Example 3, one can show that the limit along any line through the origin is 0. This does not prove that the given limit is 0, but the limits along the parabolas $y = x^2$ and $x = y^2$ also turn out to be 0, so we begin to suspect that the limit does exist.

Let ε > 0. We want to find δ > 0 such that

$$\left| \frac{3x^2 y}{x^2 + y^2} - 0 \right| < \varepsilon \text{ whenever } 0 < \sqrt{x^2 + y^2} < \delta$$

That is, $\frac{3x^2|y|}{x^2 + y^2} < \varepsilon$ whenever $0 < \sqrt{x^2 + y^2} < \delta$

But $x^2 \le x^2 + y^2$ since $y^2 \ge 0$, so

$$\frac{3x^2|y|}{x^2 + y^2} \le 3|y| = 3\sqrt{y^2} \le 3\sqrt{x^2 + y^2}$$

Thus if we choose δ = ε/3 and let $0 < \sqrt{x^2 + y^2} < \delta$, then

$$\left| \frac{3x^2 y}{x^2 + y^2} - 0 \right| \le 3\sqrt{x^2 + y^2} < 3\delta = 3\left(\frac{\varepsilon}{3}\right) = \varepsilon$$

Hence, by Definition 12.5.

$$\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^2 + y^2} = 0$$

4.0: CONCLUSION

In this unit, you have known several definitions and have worked various examples.

5.0: SUMMARY

In this unit, you have studied the definition of terms and have solved various examples .

6.0: TUTOR-MARKED-ASSIGNMENT

1. Find the limit

$$\lim_{x\to 1^-} \frac{x^2 + 2x - 3}{|x - 1|}$$

2. Find the limit

$$\lim_{x\to 5} \frac{x^2-25}{x^2+x-30}$$

3. Calculate the limit

$$\lim_{x\to 2} \frac{x^2+4x-12}{|x-2|}$$

4.Calculate the limit

$$\lim_{x\to -1^+} \sqrt[3]{x+1}\,\ln(x+1)$$

5. Find the limit

$$\lim_{x\to 9} \frac{\sqrt{x}-3}{x-9}$$

6.Find the limit

$$\lim_{t\to 0} \frac{\sin t - t}{\tan t}$$

7.Find the limit

$$\lim_{x\to\infty} \frac{3x}{\sqrt{16x^2+1}}$$

7.0: REFERENCES / FURTHER READING

1. G. B. Thomas, Jr. and R. L. Finney, *Calculus and Analytic Geometry*, 9th ed., Addison-Wesley, 1996.

2. S. Wolfram, *The Mathematica Book*, 3rd ed., Wolfram Media, 1996

3.Bartle, R. G. and Sherbert, D. *Introduction to Real Analysis.* New York: Wiley, p. 141, 1991.

4.Kaplan, W. "Limits and Continuity." §2.4 in *Advanced Calculus, 4th ed.* Reading, MA: Addison-Wesley, pp. 82-86, 1992.

Unit 3: Continuity of Function of Several Variables.

CONTENT

1.0 Introduction
2.0 Objectives
3.0 Main Content
3.1 Definitions and examples

4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

1.0: INTRODUCTION

Just as for functions of one variable, the calculation of limits can be greatly simplified by the use of properties of limits and by the use of continuity.

The properties of limits listed in Tables 2.14 and 2.15 can be extended to functions of two variables. The limit of a sum is the sum of the limits, and so on.

Recall that evaluating limits of *continuous* functions of a single variable is easy. It can be accomplished by direct substitution because the defining property of a continuous function is $\lim_{x\to a} - f(x) = f(a)$. Continuous functions of two variables are also defined by the direct substitution property.

Definition

Let f be a function of two variables defined on a disk with center (a,b). Then f is called **continuous at** (a,b) if $\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$

2.0: OBJECTIVE

At this unit, you should be able to know the definition of terms

3.0 : MAIN CONTENTS

Let f be a function of two variables defined on a disk with center (a,b). Then f is called **continuous at** (a,b) if $\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$

If the domain of f is a set $D \subset R^2$, then Definition 12.6 defines the continuity of f at an **interior point** (a,b) of D, that is, a point that is contained in a disk $D_\delta \subset D$ [seek Figure 12.20(a)]. But D may also contain a **boundary point**, that is, a point (a,b) such that every disk with center (a,b) contains points in D and also points not in D [see Figure 12.20(b)].

If (a,b) is a boundary of D, then Definition 12.5 is modified so that the last line reads

$$|f(x,y) - L| < \varepsilon \text{ whenever } (x,y) \in D \text{ and } 0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$$

With this convention, Definition 12.6 also applies when f is defined at a boundary point (a,b) of D.

Finally, we say f is **continuous on** D if f is continuous at every point (a,b) in D.

The intuitive meaning of continuity is that if the point (x,y) changes by a small amount, then the value of f(x,y) changes by a small amount. This means that a surface that is the graph of a continuous function has no holes or breaks.

Using the properties of limits, you can see that sums, differences, products, and -quotients of continuous functions are continuous on their domains. Let us use this fact to give examples of continuous functions.

A polynomial function of two variables (or polynomial, for short) is a sum of terms of the form $cx^m y^n$, where c is a constant and m and n are non-negative integers. A **rational function** is a ratio of polynomials. For instance,

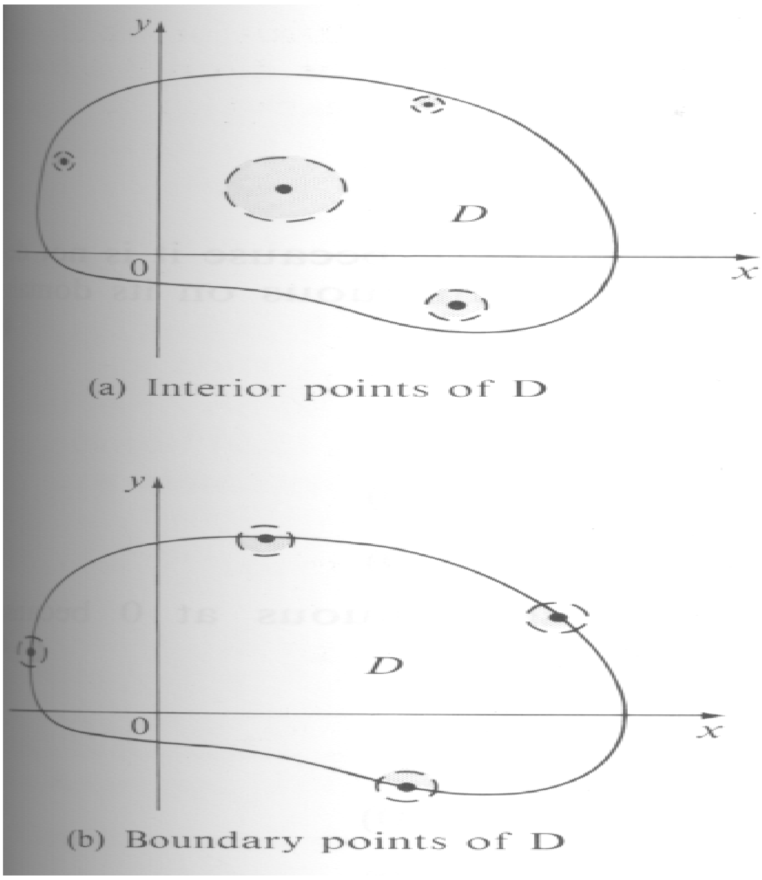
$f(x,y) = x^4 + 5x^3y^2 + 6xy^4 - 7y + 6$

is a polynomial, whereas

$$g(x,y) = \frac{2xy+1}{x^2+y^2}$$

is a rational function.

Figure 12.20



(a) Interior points of D

(b) Boundary points of D

From Definition it can be shown that

$$\lim_{(x,y)\to(a,b)} x = a \qquad \lim_{(x,y)\to(a,b)} y = b \quad \lim_{(x,y)\to(a,b)} c = c$$

These limits show that the functions f(x,y) = x, g(x,y) = y, and h(x,y) = c are continuous. Since any polynomial can be built up out of the simple functions f, g and h by multiplication and addition, it follows that all polynomials are continuous on R2. Likewise, any rational function is continuous on its domain since it is a quotient of continuous functions.

Example 5

Evaluate $\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y)$.

Solution

Since $f(x,y) = x^2y^3 - x^3y^2 + 3x + 2y$ is a polynomial, it is continuous everywhere, so the limit can be found by direct substitution:

$$\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y) = 1^2.2^3 - 1^3.2^2 + 3.1 + 2.2 = 11$$

Example 6

Where is the function

$$f(x,y) = \frac{x^2 + y^2}{x^2 + y^2} \quad \text{Continuous?}$$

Solution

The function f is discontinuous at (0,0) because it is not defined there. Since f is a rational function it is continuous on its domain $D = \{(x,y) \mid (x,y) \neq (0,0\}$.

Example 7

Let

$$g(x,y) = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2} & \text{If}(x,y) \neq (0,0) \\ 0 \end{cases}$$

Here g is defined at (0,0) but g is still discontinuous at 0 because

$\text{Lim}_{(x,y)\to(0,0)}$ g(x,y) does not exist (see Example 1).

Example 8

Let

$$f(x,y) = \begin{cases} \dfrac{3x^2 y}{x^2 + y^2} & \text{If}(x,y) \neq (0,0) \\ 0 \end{cases}$$

We know f is continuous for (x,y) ≠ (0,0) since it is equal to a rational function there. Also, from Example 4, we have

$$\lim_{(x,y)\to(a,b)} f(x,y) = \lim_{(x,y)\to(a,b)} \frac{3x^2 y}{x^2 + y^2} = 0 = f(0,0)$$

Therefore f is continuous at (0,0), and so it is continuous on R^2.

Example 9

Let

$$h(x,y) = \begin{cases} \dfrac{3x^2 y}{x^2 + y^2} & \text{If}(x,y) \neq (0,0) \\ 17 \end{cases}$$

Again from Example 4, we have

$$\lim_{(x,y)\to(a,b)} g(x,y) = \lim_{(x,y)\to(a,b)} \frac{3x^2 y}{x^2 + y^2} = 0 \neq 17 = g(0,0)$$

And so g is discontinuous at (0,0). But g is continuous on the set S = {(x,y) | (x,y) ≠ (0,0} since it is equal to a rational function on S.

Composition is another way of combining two continuous functions to get at third. The proof of the following theorem is similar to that of Theorem 2.27.

Theorem

If f is continuous at (a,b) and g is a function of a single variable that is continuous at f(a,), then the composite function $h = g \circ f$ defined by h(x,y) = g(f(x,y)) is continuous at (a,b).

Example 10

On what set is the function $h(x,y) = \ln(x^2 + y^2 - 1)$ continuous?

Solution

Let $f(x,y) = x^2 + y^2 - 1$ and $g(t) = \ln t$. Then

$g(f(x,y)) = \ln(x^2 + y^2 - 1) = h(x,y)$

So $h = g \circ f$. Now f is continuous everywhere since it is a polynomial and g is continuous on its domain $\{t \mid t > 0\}$. Thus, by Theorem 12.7, h is continuous on its domain

$D = \{(x,y) \mid x^2 + y^2 - 1 > 0\} = \{(x,y) \mid x^2 + y^2 > 1\}$

Which consists of all points outside the circle $x^2 + y^2 = 1$.

Everything in this section can be extended to functions of three or more variables. The distance between two points (x,y,z) and (a,b,c) in R^3 is $\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}$, so the definitions of limit and continuity of a function of three variables are as follows.

Definition

Let $f: D \subset R^3 \to R$.

(a) $\lim_{(x,y,z)\to(a,b,c)} f(x,y,z) = L$

Means that for every number $\varepsilon > 0$ there is a corresponding number $\delta > 0$ such that

$|f(x,y,z) - L| < \varepsilon$ whenever $(x,y,z) \in D$ and

$0 < \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} < \delta$

(b) f is **continuous** at (a,b,c) if

$$\lim_{(x,y,z)\to(a,b,c)} f(x,y,z) = f(a,b,c)$$

If we use the vector notation introduced at the end of Section 12.1, then the definitions of a limit for functions of two or three variables can be written in a single compact form as follows.

If f: $D \subset R^n \to R$, then $\lim_{x\to a} f(x) = L$ means that for every number ε > 0 there is a corresponding number δ > 0 such that

$$|f(x) - L| < \varepsilon \text{ whenever } 0 < |x - a| < \delta$$

Notice that if n = 1, then **x** = x and **a** = a, and (12.9) is just the definition of a limit for functions of a single variable. If n = 2, then x = (x,y), **a** = (a,b), and $|\mathbf{x} - \mathbf{a}| = \sqrt{(x-a)^2+(y-b)^2}$, so (12.9) becomes Definition 12.5. If n = 3, then **x** = (x,y,z), **a** = (a,b,c), and (12.9) becomes part (a) of Definition 12.8. In each case the definition of continuity can be written as

$$\lim_{x\to a} f(\mathbf{x}) = f(\mathbf{a})$$

4.0: CONCLUSION

In this unit, you have known several definitions and have worked various examples.

5.0: SUMMARY

In this unit, you have studied the definition of terms and have solved various examples . T**he following limits** $\lim_{(x,y)\to(a,b)} x = a$, $\lim_{(x,y)\to(a,b)} y = b$ and $\lim_{(x,y)\to(a,b)} c = c$

Show that the functions f(x,y) = x, g(x,y) = y, and h(x,y) = c are continuous. Obviously any polynomial can be built up out of the simple functions f, g and h by multiplication and addition, it follows that all polynomials are continuous on R2. Likewise, any rational function is continuous on its domain since it is a quotient of continuous functions.

6.0: TMA

In Exercises 1 – 3 determine the largest set on which the given function is continuous

1. F(x,y) = $\dfrac{x^2+y^2+1}{x^2+y^2-1}$

2. F(x,y) = $\dfrac{x^6 + x^3y^3 + y^6}{x^3 + y^3}$

3. G(x,y) = $\sqrt{x+y} - \sqrt{x-y}$

4. For what values of the number r is the function

$$f(x,y,z) = \begin{cases} \dfrac{(x+y+z)^r}{x^2+y^2+z^2} & \text{If}(x,y,z) \neq (0,0,0) \\ 0 & \end{cases}$$

continuous on R^3?

5. If **c** $\in V_n$, show that the function f: $R^n \to R$ given by f(**x**) = **c**.**x** is continuous on R^n.

6.0 TUTOR – MARKED ASSIGNMENT

1.Show that function f defined below is not continuous at x = - 2.

f(x) = 1 / (x + 2)

2. Show that function f is continuous for all values of x in R.

f(x) = 1 / (x^4 + 6)

3. Show that function f is continuous for all values of x in R.

f(x) = | x - 5 |

4. Find the values of x at which function f is discontinuous.

f(x) = (x - 2) / [(2 x^2 + 2x - 4)(x^4 + 5)]

5. Evaluate the limit

$\lim_{xx\to a} \sin(2x + 5)$

6.Show that any function of the form $e^{ax + b}$ is continous everywhere, a and b real numbers.

7.0: REFERENCES / FURTHER READING

1.Bartle, R. G. and Sherbert, D. *Introduction to Real Analysis.* New York: Wiley, p. 141, 1991.

2.Kaplan, W. "Limits and Continuity." §2.4 in *Advanced Calculus, 4th ed.* Reading, MA: Addison-Wesley, pp. 82-86, 1992

3. Richard Gill. Associate Professor of Mathematics. Tidewater Community

MODULE 2 PARTIAL DERIVATIVES OF FUNCTION OF SEVERAL VARIABLES

-Unit 1: Derivative

-Unit 2: Partial derivative.

-Unit 3: Application of Partial derivative.

UNIT 1: DERIVATIVE

CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
3.1 The derivative of a function
3.2 Higher derivative
3.3 Computing derivative
3.4 Derivative of higher dimension
3.0 Conclusion
4.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

1.0 INTRODUCTION

In calculus, a derivative is a measure of how the function changes as the input changes. Loosely speaking, a derivative can be thought of how much one quantity is changing in response to changes in some other quantity. For example, the derivative of the position of a moving object with respect to time, is the object instantaneous velocity.

The derivative of a function at a given chosen input value describe the best linear approximation of the function near that input value. For a real valued function of a single real variable. The derivative at a point equals the slope of the tangent line to the graph of the function at that point. In higher dimension, the derivative of a function at a point is linear transformation called the linearization. A closely related notion is the differential of a function. The process of finding a derivative is differentiation. The reverse is Integration.

The derivative of a function represents an infinitesimal change in the function with respect to one of its variables,

The "simple" derivative of a function f with respect to a variable x is denoted either $f'(x)$ or

df/dx

2.0 OBJECTIVE

In this Unit, you should be able to:

- Know the derivative of a function
- Identify higher derivative
- solve problems by Computing derivative
- identify derivative of higher dimension

3.**0 MAIN CONTENT**

3.1 The Derivative of a Function

Let f be a function that has a derivative at every point a in the domain of f. Because every point a has a derivative, there is a function that sends the point a to the derivative of f at a. This function is written $f'(x)$ and is called the *derivative function* or the *derivative* of f. The derivative of f collects all the derivatives of f at all the points in the domain of f.

Sometimes f has a derivative at most, but not all, points of its domain. The function whose value at a equals $f'(a)$ whenever $f'(a)$ is defined and elsewhere is undefined is also called the derivative of f. It is still a function, but its domain is strictly smaller than the domain of f.

Using this idea, differentiation becomes a function of functions: The derivative is an operator whose domain is the set of all functions that have derivatives at every point of their domain and whose range is a set of functions. If we denote this operator by D, then $D(f)$ is the function $f'(x)$. Since $D(f)$ is a function, it can be evaluated at a point a. By the definition of the derivative function, $D(f)(a) = f'(a)$.

For comparison, consider the doubling function $f(x) = 2x$; f is a real-valued function of a real number, meaning that it takes numbers as inputs and has numbers as outputs:

$$1 \mapsto 2,$$
$$2 \mapsto 4,$$
$$3 \mapsto 6.$$

The operator *D*, however, is not defined on individual numbers. It is only defined on functions:

$$D(x \mapsto 1) = (x \mapsto 0),$$
$$D(x \mapsto x) = (x \mapsto 1),$$
$$D(x \mapsto x^2) = (x \mapsto 2 \cdot x).$$

Because the output of *D* is a function, the output of *D* can be evaluated at a point. For instance, when *D* is applied to the squaring function,

$$x \mapsto x^2,$$

D outputs the doubling function,

$$x \mapsto 2x,$$

which we named $f(x)$. This output function can then be evaluated to get $f(1) = 2$, $f(2) = 4$, and so on.

3.2 Higher derivative

Let f be a differentiable function, and let $f'(x)$ be its derivative. The derivative of $f'(x)$ (if it has one) is written $f''(x)$ and is called the second derivative **of f**. Similarly, the derivative of a second derivative, if it exists, is written $f'''(x)$ and is called the **third derivative of f**. These repeated derivatives are called *higher-order derivatives*.

If $x(t)$ represents the position of an object at time t, then the higher-order derivatives of x have physical interpretations. The second derivative of x is the derivative of $x'(t)$, the velocity, and by definition this is the object's acceleration. The third derivative of x is defined to be the jerk, and the fourth derivative is defined to be the jounce.

A function f need not have a derivative, for example, if it is not continuous. Similarly, even if f does have a derivative, it may not have a second derivative. For example, let

$$f(x) = \begin{cases} +x^2, & \text{if } x \geq 0 \\ -x^2, & \text{if } x \leq 0. \end{cases}$$

Calculation shows that f is a differentiable function whose derivative is

$$f'(x) = \begin{cases} +2x, & \text{if } x \geq 0 \\ -2x, & \text{if } x \leq 0. \end{cases}$$

f′(x) is twice the absolute value function, and it does not have a derivative at zero. Similar examples show that a function can have *k* derivatives for any non-negative integer *k* but no *(k + 1)*-order derivative. A function that has *k* successive derivatives is called ***k* times differentiable**. If in addition the *k*th derivative is continuous, then the function is said to be of differentiability class C^k. (This is a stronger condition than having *k* derivatives.) A function that has infinitely many derivatives is called **infinitely differentiable**.

On the real line, every polynomial function is infinitely differentiable. By standard differentiation rules, if a polynomial of degree *n* is differentiated *n* times, then it becomes a constant function. All of its subsequent derivatives are identically zero. In particular, they exist, so polynomials are smooth functions.

The derivatives of a function *f* at a point *x* provide polynomial approximations to that function near *x*. For example, if *f* is twice differentiable, then

$$f(x+h) \approx f(x) + f'(x)h + \tfrac{1}{2}f''(x)h^2$$

in the sense that

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2}{h^2} = 0.$$

If *f* is infinitely differentiable, then this is the beginning of the Taylor series for *f*.

Inflection Point

A point where the second derivative of a function changes sign is called an **inflection point**.At an inflection point, the second derivative may be zero, as in the case of the inflection point $x=0$ of the function $y=x^3$, or it may fail to exist, as in the case of the inflection point $x=0$ of the function $y=x^{1/3}$. At an inflection point, a function switches from being a convex function to being a concave function or vice versa.

3.3 Computing the derivative

The derivative of a function can, in principle, be computed from the definition by considering the difference quotient, and computing its limit. In practice, once the derivatives of a few simple functions are known, the derivatives of other functions are more easily computed using *rules* for obtaining derivatives of more complicated functions from simpler ones.

Derivative of Elementary Function

Most derivative computations eventually require taking the derivative of some common functions. The following incomplete list gives some of the most frequently used functions of a single real variable and their derivatives.

- Derivative power: if

$$f(x) = x^r,$$

where *r* is any real number, then

$$f'(x) = rx^{r-1},$$

wherever this function is defined. For example, if $f(x) = x^{1/4}$, then

$$f'(x) = (1/4)x^{-3/4},$$

and the derivative function is defined only for positive *x*, not for *x* = 0. When *r* = 0, this rule implies that *f*′(*x*) is zero for $x \neq 0$, which is almost the constant rule (stated below).

Exponential *and* logarithm *functions*:

$$\frac{d}{dx}e^x = e^x$$
$$\frac{d}{dx}a^x = \ln(a)a^x$$
$$\frac{d}{dx}\ln(x) = \frac{1}{x}, \qquad x > 0$$
$$\frac{d}{dx}\log_a(x) = \frac{1}{x\ln(a)}$$

Trigonometric Functions :

$$\frac{d}{dx}\sin(x) = \cos(x).$$
$$\frac{d}{dx}\cos(x) = -\sin(x).$$
$$\frac{d}{dx}\tan(x) = \sec^2(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x).$$

Inverse Trigonometric Function :

$$\frac{d}{dx}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}.$$
$$\frac{d}{dx}\arccos(x) = -\frac{1}{\sqrt{1-x^2}}.$$
$$\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}.$$

Rules for finding the derivative

In many cases, complicated limit calculations by direct application of Newton's difference quotient can be avoided using differentiation rules. Some of the most basic rules are the following.

Constant rule: if $f(x)$ is constant, then

$$f' = 0$$

Sine rule :

$$(af + bg)' = af' + bg'$$ for all functions f and g and all real numbers a and b.

Product rule :

$$(fg)' = f'g + fg'$$ for all functions f and g.

Quotient rule :

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$ for all functions f and g where $g \neq 0$.

Chain rule : If $f(x) = h(g(x))$, then

$$f'(x) = h'(g(x)) \cdot g'(x).$$

Example computation

The derivative of

$$f(x) = x^4 + \sin(x^2) - \ln(x)e^x + 7$$

is

$$f'(x) = 4x^{(4-1)} + \frac{d\left(x^2\right)}{dx}\cos(x^2) - \frac{d\left(\ln x\right)}{dx}e^x - \ln x\frac{d\left(e^x\right)}{dx} + 0$$
$$= 4x^3 + 2x\cos(x^2) - \frac{1}{x}e^x - \ln(x)e^x.$$

Here the second term was computed using the chain rule and third using the product rule. The known derivatives of the elementary functions x^2, x^4, $\sin(x)$, $\ln(x)$ and $\exp(x) = e^x$, as well as the constant 7, were also used.

3.4 Derivatives in higher dimensions

Derivative of vector valued function

A vector valued function $\mathbf{y}(t)$ of a real variable sends real numbers to vectors in some vector space $\mathbf{R}^n$. A vector-valued function can be split up into its coordinate functions $y_1(t)$, $y_2(t)$, …, $y_n(t)$, meaning that $\mathbf{y}(t) = (y_1(t), ..., y_n(t))$. This includes, for example, parametric curve in $\mathbf{R}^2$

or $\mathbf{R}^3$. The coordinate functions are real valued functions, so the above definition of derivative applies to them. The derivative of $\mathbf{y}(t)$ is defined to be the vector, called the tangent vector, whose coordinates are the derivatives of the coordinate functions. That is,

$$\mathbf{y}'(t) = (y_1'(t), \ldots, y_n'(t)).$$

Equivalently,

$$\mathbf{y}'(t) = \lim_{h\to 0} \frac{\mathbf{y}(t+h) - \mathbf{y}(t)}{h},$$

if the limit exists. The subtraction in the numerator is subtraction of vectors, not scalars. If the derivative of $\mathbf{y}$ exists for every value of t, then $\mathbf{y}'$ is another vector valued function.

If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is the standard basis for $\mathbf{R}^n$, then $\mathbf{y}(t)$ can also be written as $y_1(t)\mathbf{e}_1 + \ldots + y_n(t)\mathbf{e}_n$. If we assume that the derivative of a vector-valued function retains the linearity property, then the derivative of $\mathbf{y}(t)$ must be

$$y_1'(t)\mathbf{e}_1 + \cdots + y_n'(t)\mathbf{e}_n$$

because each of the basis vectors is a constant.

This generalization is useful, for example, if $\mathbf{y}(t)$ is the position vector of a particle at time t; then the derivative $\mathbf{y}'(t)$ is the velocity vector of the particle at time t.

Partial derivative

Suppose that f is a function that depends on more than one variable. For instance,

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

f can be reinterpreted as a family of functions of one variable indexed by the other variables:

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

In other words, every value of x chooses a function, denoted f_x, which is a function of one real number. That is,

$$x \mapsto f_x,$$

$$f_x(y) = x^2 + xy + y^2.$$

Once a value of x is chosen, say a, then $f(x,y)$ determines a function f_a that sends y to $a^2 + ay + y^2$:

$$f_a(y) = a^2 + ay + y^2.$$

In this expression, a is a *constant*, not a *variable*, so f_a is a function of only one real variable. Consequently the definition of the derivative for a function of one variable applies:

$$f_a'(y) = a + 2y.$$

The above procedure can be performed for any choice of a. Assembling the derivatives together into a function gives a function that describes the variation of f in the y direction:

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

This is the partial derivative of f with respect to y. Here ∂ is a rounded d called the **partial derivative symbol**. To distinguish it from the letter d, ∂ is sometimes pronounced "der", "del", or "partial" instead of "dee".

In general, the **partial derivative** of a function $f(x_1, \ldots, x_n)$ in the direction x_i at the point $(a_1 \ldots, a_n)$ is defined to be:

$$\frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_i, \ldots, a_n)}{h}.$$

In the above difference quotient, all the variables except x_i are held fixed. That choice of fixed values determines a function of one variable

$$f_{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n}(x_i) = f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$

and, by definition,

$$\frac{df_{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n}}{dx_i}(a_i) = \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n).$$

In other words, the different choices of a index a family of one-variable functions just as in the example above. This expression also shows that the computation of partial derivatives reduces to the computation of one-variable derivatives.

An important example of a function of several variables is the case of a scalar valued function $f(x_1, \ldots x_n)$ on a domain in Euclidean space $\mathbf{R}^n$ (e.g., on $\mathbf{R}^2$ or $\mathbf{R}^3$). In this case f has a partial derivative $\partial f/\partial x_j$ with respect to each variable x_j. At the point a, these partial derivatives define the vector

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

This vector is called the gradient of f at a. If f is differentiable at every point in some domain, then the gradient is a vector-valued function ∇f that takes the point a to the vector $\nabla f(a)$. Consequently the gradient determines a vector field.

Generalizations

The concept of a derivative can be extended to many other settings. The common thread is that the derivative of a function at a point serves as a linear approximation of the function at that point.

4.0 CONCLUSION

In this unit, you have known the derivative of a function .Through the derivative of functions, you have identified higher derivative, and you have solved problems by computing derivative through the use of this functions. You have also identified derivative of higher dimension.

5.0 SUMMARY

In this unit, you have studied the following:

- the derivative of a function
- identify higher derivative
- solve problems by Computing derivative
- identify derivative of higher dimension

6.0 TUTOR MARKED ASSIGNMENT

Find the derivative of F(x,y) = 3sin(3xy)

Find the derivative of F(x,y)= $(x^3 + \ln 6)(\sqrt{y})$

Evaluate the derivative F(x,y) = $x^2 + 3xy - 2\tan(y)$

Find the derivative of F(x,y) = $\dfrac{y\sin x}{e^{\cos x}}$

REFERENCES

Anton, Howard; Bivens, Irl; Davis, Stephen (February 2, 2005), *Calculus: Early Transcendentals Single and Multivariable* (8th ed.), New York: Wiley, ISBN 978-0-471-47244-5

Apostol, Tom M. (June 1967), *Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00005-1

- Apostol, Tom M. (June 1969), *Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00007-5
- Courant, Richard; John, Fritz (December 22, 1998), *Introduction to Calculus and Analysis, Vol. 1*, Springer-Verlag, ISBN 978-3-540-65058-4
- Eves, Howard (January 2, 1990), *An Introduction to the History of Mathematics* (6th ed.), Brooks Cole, ISBN 978-0-03-029558-4

- Larson, Ron; Hostetler, Robert P.; Edwards, Bruce H. (February 28, 2006), *Calculus: Early Transcendental Functions* (4th ed.), Houghton Mifflin Company, ISBN 978-0-618-60624-5
- Spivak, Michael (September 1994), *Calculus* (3rd ed.), Publish or Perish, ISBN 978-0-914098-89-8
- Stewart, James (December 24, 2002), *Calculus* (5th ed.), Brooks Cole, ISBN 978-0-534-39339-7
- Thompson, Silvanus P. (September 8, 1998), *Calculus Made Easy* (Revised, Updated, Expanded ed.), New York: St. Martin's Press, ISBN 978-0-312-18548-0

- Crowell, Benjamin (2003), *Calculus*, http://www.lightandmatter.com/calc/
- Garrett, Paul (2004), *Notes on First-Year Calculus*, University of Minnesota, http://www.math.umn.edu/~garrett/calculus/
- Hussain, Faraz (2006), *Understanding Calculus*, http://www.understandingcalculus.com/
- Keisler, H. Jerome (2000), *Elementary Calculus: An Approach Using Infinitesimals*, http://www.math.wisc.edu/~keisler/calc.html
- Mauch, Sean (2004), *Unabridged Version of Sean's Applied Math Book*, http://www.its.caltech.edu/~sean/book/unabridged.html
- Sloughter, Dan (2000), *Difference Equations to Differential Equations*, http://synechism.org/drupal/de2de/
- Strang, Gilbert (1991), *Calculus*, http://ocw.mit.edu/ans7870/resources/Strang/strangtext.htm
- Stroyan, Keith D. (1997), *A Brief Introduction to Infinitesimal Calculus*, http://www.math.uiowa.edu/~stroyan/InfsmlCalculus/InfsmlCalc.htm

Unit 2 : Partial derivatives

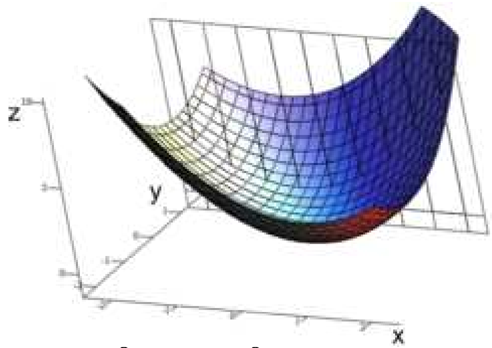
CONTENTS

1.0 Introduction
2.0 Objectives
3.0 Main Content
  3.1 Define Partial derivative
  3.2 Know the geometric interpretation
  3.3 Identify anti derivative analogue
  3.4 Solve problems on partial derivative for function of several variables
  3.5 Identify higher order derivatives
4.0 Conclusion
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

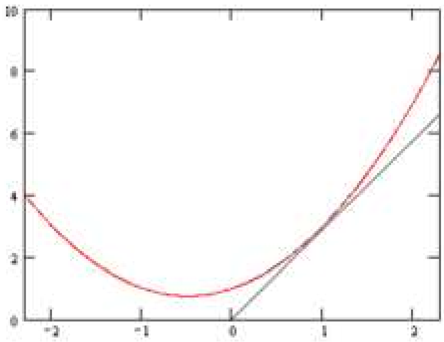
1.0 INTRODUCTION

Suppose that f is a function of more than one variable. For instance,

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



A graph of $z = x^2 + xy + y^2$. For the partial derivative at (1, 1, 3) that leaves y constant, the corresponding tangent line is parallel to the xz-plane.



A slice of the graph above at y= 1

The graph of this function defines a surface in Euclidean space. To every point on this surface, there are an infinite number of tangent lines. Partial differentiation is the act of choosing one of these lines and finding its slope. Usually, the lines of most interest are those that are parallel to the xz-plane, and those that are parallel to the yz-plane.

To find the slope of the line tangent to the function at (1, 1, 3) that is parallel to the xz-plane, the y variable is treated as constant. The graph and this plane are shown on the right. On the graph below it, we see the way the function looks on the plane $y = 1$. By finding the derivative of the equation while assuming that y is a constant, the slope of f at the point (x, y, z) is found to be:

$$\frac{\partial z}{\partial x} = 2x + y$$

So at (1, 1, 3), by substitution, the slope is 3. Therefore

$$\frac{\partial z}{\partial x} = 3$$

at the point. (1, 1, 3). That is, the partial derivative of z with respect to x at (1, 1, 3) is 3

2.0: OBJECTIVES

After studying this, you should be able to :

- define Partial derivative

- know the geometric interpretation
- identify anti derivative analogue
- solve problems on partial derivative for function of several variables
- identify higher order derivatives

3.0 MAIN CONTENT

Let us consider a function

1) $u = f(x, y, z, p, q, \ldots)$

of several variables. Such a function can be studied by holding all variables except one constant and observing its variation with respect to one single selected variable. If we consider all the variables except x to be constant, then

$$\frac{du}{dx} = \frac{d\,f(x, \hat{y}, \hat{z}, \hat{p}, \hat{q}, \ldots)}{dx}$$

represents the partial derivative of f(x, y, z, p, q, ...) with respect to x (the hats indicating variables held fixed). The variables held fixed are viewed as parameters.

Definition of Partial derivative.

The partial derivative of a function of two or more variables with respect to one of its variables is the ordinary derivative of the function with respect to that variable, considering the other variables as constants.

Example 1. The partial derivative of $3x^2y + 2y^2$ with respect to x is 6xy. Its partial derivative with respect to y is $3x^2 + 4y$.

The partial derivative of a function z = f(x, y, ...) with respect to the variable x is commonly written in any of the following ways:

$$\frac{\partial z}{\partial x}, \quad \frac{\partial f}{\partial x}, \quad \frac{\partial f(x,y,\ldots)}{\partial x}, \quad D_x f(x,y,\ldots), \quad D_x f, \quad f_x(x,y,\ldots), \quad f_x, \quad f_1(x,y,\ldots)$$

Its derivative with respect to any other variable is written in a similar fashion.

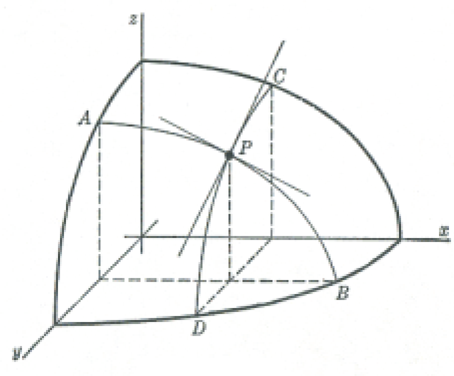


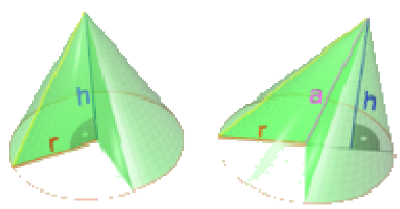
Fig. 1

Geometric interpretation. The geometric interpretation of a partial derivative is the same as that for an ordinary derivative. It represents the slope of the tangent to that curve represented by the function at a particular point P. In the case of a function of two variables

z = f(x, y)

Fig. 1 shows the interpretation of $\partial f / \partial x$ and $\partial f / \partial y$. $\partial f / \partial x$ corresponds to the slope of the tangent to the curve APB at point P (where curve APB is the intersection of the surface with a plane through P perpendicular to the y axis). Similarly, $\partial f / \partial y$ corresponds to the slope of the tangent to the curve CPD at point P (where curve CPD is the intersection of the surface with a plane through P perpendicular to the x axis).

Examples 2



The volume of a cone depends on height and radius

The volume *V* of a cone depends on the cone's height *h* and its radius *r* according to the formula

$$V(r,h) = \frac{\pi r^2 h}{3}.$$

The partial derivative of *V* with respect to *r* is

$$\frac{\partial V}{\partial r} = \frac{2\pi r h}{3},$$

which represents the rate with which a cone's volume changes if its radius is varied and its height is kept constant. The partial derivative with respect to *h* is

$$\frac{\partial V}{\partial h} = \frac{\pi r^2}{3},$$

which represents the rate with which the volume changes if its height is varied and its radius is kept constant.

By contrast, the *total* derivative of *V* with respect to *r* and *h* are respectively

$$\frac{\mathrm{d}V}{\mathrm{d}r} = \overbrace{\frac{2\pi rh}{3}}^{\frac{\partial V}{\partial r}} + \overbrace{\frac{\pi r^2}{3}}^{\frac{\partial V}{\partial h}} \frac{\mathrm{d}h}{\mathrm{d}r}$$

and

$$\frac{\mathrm{d}V}{\mathrm{d}h} = \overbrace{\frac{\pi r^2}{3}}^{\frac{\partial V}{\partial h}} + \overbrace{\frac{2\pi rh}{3}}^{\frac{\partial V}{\partial r}} \frac{\mathrm{d}r}{\mathrm{d}h}$$

The difference between the total and partial derivative is the elimination of indirect dependencies between variables in partial derivatives.

If (for some arbitrary reason) the cone's proportions have to stay the same, and the height and radius are in a fixed ratio *k*,

$$k = \frac{h}{r} = \frac{\mathrm{d}h}{\mathrm{d}r}.$$

This gives the total derivative with respect to *r*:

$$\frac{\mathrm{d}V}{\mathrm{d}r} = \frac{2\pi rh}{3} + k\frac{\pi r^2}{3}$$

Equations involving an unknown function's partial derivatives are called partial differential equations and are common in physics, engineering, and other sciences and applied disciplines.

Notation

For the following examples, let *f* be a function in *x*, *y* and *z*.

First-order partial derivatives:

$$\frac{\partial f}{\partial x} = f_x = \partial_x f.$$

Second-order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = \partial_{xx} f.$$

Second-order mixed derivatives:

$$\frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{xy} = \partial_{yx} f.$$

Higher-order partial and mixed derivatives:

$$\frac{\partial^{i+j+k} f}{\partial x^i\, \partial y^j\, \partial z^k} = f^{(i,j,k)}.$$

When dealing with functions of multiple variables, some of these variables may be related to each other, and it may be necessary to specify explicitly which variables are being held constant. In fields such as statistical mechanics, the partial derivative of *f* with respect to *x*, holding *y* and *z* constant, is often expressed as

$$\left(\frac{\partial f}{\partial x}\right)_{y,z}.$$

Anti derivative analogue

There is a concept for partial derivatives that is analogous to anti derivatives for regular derivatives. Given a partial derivative, it allows for the partial recovery of the original function.

Consider the example of $\frac{\partial z}{\partial x} = 2x + y$. The "partial" integral can be taken with respect to *x* (treating *y* as constant, in a similar manner to partial derivation):

$$z = \int \frac{\partial z}{\partial x}\, dx = x^2 + xy + g(y)$$

Here, the "constant" of integration is no longer a constant, but instead a function of all the variables of the original function except *x*. The reason for this is that all the other variables are treated as constant when taking the partial derivative, so any function which does not involve *x* will disappear when taking the partial derivative, and we have to account for this when we take the antiderivative. The most general way to represent this is to have the "constant" represent an unknown function of all the other variables. Thus the set of functions

$x^2 + xy + g(y)$, where g is any one-argument function, represents the entire set of functions in variables x,y that could have produced the x-partial derivative $2x+y$.

If all the partial derivatives of a function are known (for example, with the gradient), then the antiderivatives can be matched via the above process to reconstruct the original function up to a constant

Example 3

For the function

$$f(x, y) = x^2 + x^3y^2 + y^4$$

find the partial derivatives of f with respect to x and y and compute the rates of change of the function in the x and y directions at the point (-1,2).

Initially we will not specify the values of x and y when we take the derivatives; we will just remember which one we are going to hold constant while taking the derivative. First, hold y fixed and find the partial derivative of f with respect to x:

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

Second, hold x fixed and find the partial derivative of f with respect to y:

Now, plug in the values x=-1 and y=2 into the equations. We obtain f_x(-1,2)=10 and f_y(-1,2)=28.

Partial Derivatives for Functions of Several Variables

We can of course take partial derivatives of functions of more than two variables. If f is a function of n variables x_1, x_2, ..., x_n, then to take the partial derivative of f with respect to x_i we hold all variables besides x_i constant and take the derivative.

Example 4

To find the partial derivative of f with respect to t for the function

$$f(x, y, z, t) = x^2 + y^2 + z^2 + t^2 + xyzt^{-3}$$

we hold x, y, and z constant and take the derivative with respect to the remaining variable t. The result is

$$\frac{\partial f}{\partial t}(x, y, z, t) = 0 + 0 + 0 + 2t - 3xyzt^{-4}$$

Interpretation

$\frac{\partial f}{\partial x}$ Is the rate at which f changes as x changes, for a fixed (constant) y.

$\frac{\partial f}{\partial y}$ Is the rate at which f changes as y changes, for a fixed (constant) x.

Higher Order Partial Derivatives

If f is a function of x, y, and possibly other variables, then

$\frac{\partial^2 f}{\partial x^2}$ is defined to be $\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x}\right]$

Similarly,

$\frac{\partial^2 f}{\partial y^2}$ is defined to be $\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial y}\right]$

$\frac{\partial^2 f}{\partial y \partial x}$ is defined to be $\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right]$

$\frac{\partial^2 f}{\partial x \partial y}$ is defined to be $\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right]$

The above second order partial derivatives can also be denoted by f_{xx}, f_{yy}, f_{xy}, and f_{yx} respectively.

The last two are called **mixed derivatives** and will always be equal to each other when all the first order partial derivatives are continuous.

Some examples of partial derivatives of functions of several variables are shown below, variable as we did in Calculus I.

Example 1 Find all of the first order partial derivatives for the following functions.

(a) $f(x, y) = x^4 + 6\sqrt{y} - 10$

(b) $w = x^2 y - 10 y^2 z^3 + 43x - 7\tan(4y)$

(c) $h(s, t) = t^7 \ln(s^2) + \dfrac{9}{t^3} - \sqrt[7]{s^4}$

(d) $f(x, y) = \cos\left(\dfrac{4}{x}\right) \mathrm{e}^{x^2 y - 5y^3}$

Solution

(a) $f(x, y) = x^4 + 6\sqrt{y} - 10$

Let's first take the derivative with respect to *x* and remember that as we do so all the *y*'s will be treated as constants. The partial derivative with respect to *x* is,

$$f_x(x, y) = 4x^3$$

Notice that the second and the third term differentiate to zero in this case. It should be clear why the third term differentiated to zero. It's a constant and we know that constants always differentiate to zero. This is also the reason that the second term differentiated to zero. Remember that since we are differentiating with respect to *x* here we are going to treat all *y*'s as constants. That means that terms that only involve *y*'s will be treated as constants and hence will differentiate to zero.

Now, let's take the derivative with respect to *y*. In this case we treat all *x*'s as constants and so the first term involves only *x*'s and so will differentiate to zero, just as the third term will. Here is the partial derivative with respect to *y*.

$$f_y(x, y) = \frac{3}{\sqrt{y}}$$

(b) $w = x^2 y - 10 y^2 z^3 + 43x - 7\tan(4y)$

With this function we've got three first order derivatives to compute. Let's do the partial derivative with respect to *x* first. Since we are differentiating with respect to *x* we will treat all *y*'s and all *z*'s as constants. This means that the second and fourth terms will differentiate to zero since they only involve *y*'s and *z*'s.

This first term contains both *x*'s and *y*'s and so when we differentiate with respect to *x* the *y* will be thought of as a multiplicative constant and so the first term will be differentiated just as the third term will be differentiated.

Here is the partial derivative with respect to *x*.

$$\frac{\partial w}{\partial x} = 2xy + 43$$

Let's now differentiate with respect to *y*. In this case all *x*'s and *z*'s will be treated as constants. This means the third term will differentiate to zero since it contains only *x*'s while the *x*'s in the first term and the *z*'s in the second term will be treated as multiplicative constants. Here is the derivative with respect to *y*.

$$\frac{\partial w}{\partial y} = x^2 - 20yz^3 - 28\sec^2(4y)$$

Finally, let's get the derivative with respect to *z*. Since only one of the terms involve *z*'s this will be the only non-zero term in the derivative. Also, the *y*'s in that term will be treated as multiplicative constants. Here is the derivative with respect to *z*.

$$\frac{\partial w}{\partial z} = -30y^2z^2$$

(c) $h(s,t) = t^7 \ln(s^2) + \frac{9}{t^3} - \sqrt[7]{s^4}$

With this one we'll not put in the detail of the first two. Before taking the derivative let's rewrite the function a little to help us with the differentiation process.

$$h(s,t) = t^7 \ln(s^2) + 9t^{-3} - s^{\frac{4}{7}}$$

Now, the fact that we're using *s* and *t* here instead of the "standard" *x* and *y* shouldn't be a problem. It will work the same way. Here are the two derivatives for this function.

$$h_s(s,t) = \frac{\partial h}{\partial s} = t^7\left(\frac{2s}{s^2}\right) - \frac{4}{7}s^{-\frac{3}{7}} = \frac{2t^7}{s} - \frac{4}{7}s^{-\frac{3}{7}}$$

$$h_t(s,t) = \frac{\partial h}{\partial t} = 7t^6\ln\left(s^2\right) - 27t^{-4}$$

Remember how to differentiate natural logarithms.

$$\frac{d}{dx}\left(\ln g(x)\right) = \frac{g'(x)}{g(x)}$$

(d) $f(x,y) = \cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}$

Now, we can't forget the product rule with derivatives. The product rule will work the same way here as it does with functions of one variable. We will just need to be careful to remember which variable we are differentiating with respect to.

Let's start out by differentiating with respect to *x*. In this case both the cosine and the exponential contain *x*'s and so we've really got a product of two functions involving *x*'s and so we'll need to product rule this up. Here is the derivative with respect to *x*.

$$f_x(x,y) = -\sin\left(\frac{4}{x}\right)\left(-\frac{4}{x^2}\right)\mathbf{e}^{x^2y-5y^3} + \cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}\left(2xy\right)$$
$$= \frac{4}{x^2}\sin\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3} + 2xy\cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2y-5y^3}$$

Do not forget the chain rule for functions of one variable. We will be looking at the chain rule for some more complicated expressions for multivariable functions in a latter section. However, at this point we're treating all the *y*'s as constants and so the chain rule will continue to work as it did back in Calculus I.

Also, don't forget how to differentiate exponential functions,

$$\frac{d}{dx}\left(\mathbf{e}^{f(x)}\right) = f'(x)\,\mathbf{e}^{f(x)}$$

Now, let's differentiate with respect to *y*. In this case we don't have a product rule to worry about since the only place that the *y* shows up is in the exponential. Therefore, since *x*'s are considered to be constants for this derivative, the cosine in the front will also be thought of as a multiplicative constant. Here is the derivative with respect to *y*.

$$f_y(x, y) = \left(x^2 - 15y^2\right)\cos\left(\frac{4}{x}\right)\mathbf{e}^{x^2 y - 5y^3}$$

Example 2 Find all of the first order partial derivatives for the following functions.

(a) $z = \dfrac{9u}{u^2 + 5v}$

(b) $g(x, y, z) = \dfrac{x\sin(y)}{z^2}$

(c) $z = \sqrt{x^2 + \ln\left(5x - 3y^2\right)}$

Solution

(a) $z = \dfrac{9u}{u^2 + 5v}$

We also can't forget about the quotient rule. Since there isn't too much to this one, we will simply give the derivatives.

$$z_u = \frac{9\left(u^2 + 5v\right) - 9u\left(2u\right)}{\left(u^2 + 5v\right)^2} = \frac{-9u^2 + 45v}{\left(u^2 + 5v\right)^2}$$

$$z_v = \frac{\left(0\right)\left(u^2 + 5v\right) - 9u\left(5\right)}{\left(u^2 + 5v\right)^2} = \frac{-45u}{\left(u^2 + 5v\right)^2}$$

In the case of the derivative with respect to *v* recall that *u*'s are constant and so when we differentiate the numerator we will get zero!

(b) $g\left(x, y, z\right) = \dfrac{x\sin\left(y\right)}{z^2}$

Now, we do need to be careful however to not use the quotient rule when it doesn't need to be used. In this case we do have a quotient, however, since the *x*'s and *y*'s only appear in the numerator and the *z*'s only appear in the denominator this really isn't a quotient rule problem.

Let's do the derivatives with respect to *x* and *y* first. In both these cases the *z*'s are constants and so the denominator in this is a constant and so we don't really need to worry too much about it. Here are the derivatives for these two cases.

$$g_x\left(x, y, z\right) = \frac{\sin\left(y\right)}{z^2} \qquad\qquad g_y\left(x, y, z\right) = \frac{x\cos\left(y\right)}{z^2}$$

Now, in the case of differentiation with respect to *z* we can avoid the quotient rule with a quick rewrite of the function. Here is the rewrite as well as the derivative with respect to *z*.

$$g(x,y,z) = x\sin(y)z^{-2}$$

$$g_z(x,y,z) = -2x\sin(y)z^{-3} = -\frac{2x\sin(y)}{z^3}$$

We went ahead and put the derivative back into the "original" form just so we could say that we did. In practice you probably don't really need to do that.

(c) $z = \sqrt{x^2 + \ln\left(5x - 3y^2\right)}$

In this last part we are just going to do a somewhat messy chain rule problem. However, if you had a good background in Calculus I chain rule this shouldn't be all that difficult of a problem. Here are the two derivatives,

$$\begin{aligned} z_x &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \frac{\partial}{\partial x}\left(x^2 + \ln\left(5x - 3y^2\right)\right) \\ &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}}\left(2x + \frac{5}{5x - 3y^2}\right) \\ &= \left(x + \frac{5}{2\left(5x - 3y^2\right)}\right)\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \end{aligned}$$

$$\begin{aligned} z_y &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \frac{\partial}{\partial y}\left(x^2 + \ln\left(5x - 3y^2\right)\right) \\ &= \frac{1}{2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \left(\frac{-6y}{5x - 3y^2}\right) \\ &= -\frac{3y}{5x - 3y^2}\left(x^2 + \ln\left(5x - 3y^2\right)\right)^{-\frac{1}{2}} \end{aligned}$$

So, there are some examples of partial derivatives. Hopefully you will agree that as long as we can remember to treat the other variables as constants these work in exactly the same manner that derivatives of functions of one variable do. So, if you can do Calculus I derivative you shouldn't have too much difficulty in doing basic partial derivatives.

There is one final topic that we need to take a quick look at in this section, implicit differentiation. Before getting into implicit differentiation for multiple variable functions let's first remember how implicit differentiation works for functions of one variable.

Example 3 Find $\frac{dy}{dx}$ for $3y^4 + x^7 = 5x$.

Solution

Remember that the key to this is to always think of *y* as a function of *x*, or $y = y(x)$ and so whenever we differentiate a term involving *y*'s with respect to *x* we will really need to use the chain rule which will mean that we will add on a $\frac{dy}{dx}$ to that term.

The first step is to differentiate both sides with respect to *x*.

$$12y^3 \frac{dy}{dx} + 7x^6 = 5$$

The final step is to solve for $\frac{dy}{dx}$

$$\frac{dy}{dx} = \frac{5 - 7x^6}{12y^3}$$

Now, we did this problem because implicit differentiation works in exactly the same manner with functions of multiple variables. If we have a function in terms of three variables *x*, *y*, and *z* we will assume that *z* is in fact a function of *x* and *y*. In other words, $z = z(x, y)$. Then whenever we differentiate *z*'s with respect to *x* we will use the chain rule and add on a $\frac{\partial z}{\partial x}$. Likewise, whenever we differentiate *z*'s with respect to *y* we will add on a $\frac{\partial z}{\partial y}$.

Let's take a quick look at a couple of implicit differentiation problems.

Example 4 Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ for each of the following functions.

(a) $x^3 z^2 - 5xy^5 z = x^2 + y^3$

(b) $x^2 \sin(2y - 5z) = 1 + y\cos(6zx)$

Solution

(a) $x^3 z^2 - 5xy^5 z = x^2 + y^3$

Let's start with finding $\frac{\partial z}{\partial x}$ We first will differentiate both sides with respect to *x* and remember to add on a $\frac{\partial z}{\partial x}$ whenever we differentiate a *z*.

$3x^2 z^2 + 2x^3 z\frac{\partial z}{\partial x} - 5y^5 z - 5xy^5\frac{\partial z}{\partial x} = 2x$ Remember that since we are assuming $z = z(x, y)$ then any product of *x*'s and *z*'s will be a product and so will need the product rule! Now, solve for

$$3x^2z^2 + 2x^3z\frac{\partial z}{\partial x} - 5y^5z - 5xy^5\frac{\partial z}{\partial x} = 2x \quad \frac{\partial z}{\partial x}.$$

$$\left(2x^3z - 5xy^5\right)\frac{\partial z}{\partial x} = 2x - 3x^2z^2 + 5y^5z$$

$$\frac{\partial z}{\partial x} = \frac{2x - 3x^2z^2 + 5y^5z}{2x^3z - 5xy^5}$$

Now we'll do the same thing for $\frac{\partial z}{\partial y}$ except this time we'll need to remember to add on a $\frac{\partial z}{\partial y}$ whenever we differentiate a *z*.

$$2x^3z\frac{\partial z}{\partial y} - 25xy^4z - 5xy^5\frac{\partial z}{\partial y} = 3y^2$$

$$\left(2x^3z - 5xy^5\right)\frac{\partial z}{\partial y} = 3y^2 + 25xy^4z$$

$$\frac{\partial z}{\partial y} = \frac{3y^2 + 25xy^4z}{2x^3z - 5xy^5}$$

(b) $x^2\sin\left(2y - 5z\right) = 1 + y\cos\left(6zx\right)$

We'll do the same thing for this function as we did in the previous part. First let's find $\frac{\partial z}{\partial x}$.

$$2x\sin\left(2y - 5z\right) + x^2\cos\left(2y - 5z\right)\left(-5\frac{\partial z}{\partial x}\right) = -y\sin\left(6zx\right)\left(6z + 6x\frac{\partial z}{\partial x}\right)$$

Don't forget to do the chain rule on each of the trig functions and when we are differentiating the inside function on the cosine we will need to also use the product rule. Now let's solve for $\frac{\partial z}{\partial x}$

$$2x\sin(2y-5z)-5\frac{\partial z}{\partial x}x^2\cos(2y-5z)=-6zy\sin(6zx)-6yx\sin(6zx)\frac{\partial z}{\partial x}$$

$$2x\sin(2y-5z)+6zy\sin(6zx)=\left(5x^2\cos(2y-5z)-6yx\sin(6zx)\right)\frac{\partial z}{\partial x}$$

$$\frac{\partial z}{\partial x}=\frac{2x\sin(2y-5z)+6zy\sin(6zx)}{5x^2\cos(2y-5z)-6yx\sin(6zx)}$$

Now let's take care of $\frac{\partial z}{\partial y}$. This one will be slightly easier than the first one.

$$x^2\cos(2y-5z)\left(2-5\frac{\partial z}{\partial y}\right)=\cos(6zx)-y\sin(6zx)\left(6x\frac{\partial z}{\partial y}\right)$$

$$2x^2\cos(2y-5z)-5x^2\cos(2y-5z)\frac{\partial z}{\partial y}=\cos(6zx)-6xy\sin(6zx)\frac{\partial z}{\partial y}$$

$$\left(6xy\sin(6zx)-5x^2\cos(2y-5z)\right)\frac{\partial z}{\partial y}=\cos(6zx)-2x^2\cos(2y-5z)$$

$$\frac{\partial z}{\partial y}=\frac{\cos(6zx)-2x^2\cos(2y-5z)}{6xy\sin(6zx)-5x^2\cos(2y-5z)}$$

4.0 CONCLUSION

In this unit, you have defined a Partial derivative of a function of several variables. You have used the partial derivative of a function of several variable to know the geometric interpretation of a function and anti derivative analogue has been identified. You have Solved problems on partial derivative for function of several variables and identified higher order derivatives.

5.0 SUMMARY

In this unit, you have studied the following:

the definition of Partial derivative of functions of several variable

the geometric interpretation of partial derivative of functions of several variables the identification of antiderivative analogue of partial derivative of functions of several variable

Solve problems on partial derivative for function of several variables

The identification of higher order derivatives of functions of several variables

TUTOR MARKED ASSIGNMENT

1.Find the partial derivatives f_x and f_y if f(x , y) is given by

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

2: Find f_x and f_y if f(x , y) is given by

$$f(x , y) = \sin(x y) + \cos x$$

3.Find f_x and f_y if f(x , y) is given by

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

4.Find f_x and f_y if f(x , y) is given by

$$f(x , y) = \ln (x^2 + 2 y)$$

5.Find $f_x(2 , 3)$ and $f_y(2 , 3)$ if f(x , y) is given by

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

6.Find partial derivatives f_x and f_y of the following functions

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

B. $f(x , y) = \ln (2 x + y x)$

C. $f(x , y) = x \sin(x - y)$

7.0 REFERENCE

Jeff Miller (2009-06-14). "Earliest Uses of Symbols of Calculus". *Earliest Uses of Various Mathematical Symbols*. http://jeff560.tripod.com/calculus.html. Retrieved 2010-02-20.

Abramowitz, M. and Stegun, I. A. (Eds.). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing.* New York: Dover, pp. 883-885, 1972.

Fischer, G. (Ed.). Plate 121 in *Mathematische Modelle aus den Sammlungen von Universitäten und Museen, Bildband.* Braunschweig, Germany: Vieweg, p. 118, 1986.

Thomas, G. B. and Finney, R. L. §16.8 in *Calculus and Analytic Geometry, 9th ed.* Reading, MA: Addison-Wesley, 1996.

Wagon, S. *Mathematica in Action.* New York: W. H. Freeman, pp. 83-85, 1991.

Unit 3 APPLICATION OF PARTIAL DERIVATIVE

CONTENT

1.0 INTRODUCTION

2.0 **OBJECTIVES**

3.0 MAIN CONTENT

3.1 Apply partial derivative of functions of several variable in Chain rule.
3.2 Apply partial derivative of functions of several variable in Curl (Mathematics)
3.3 Apply partial derivative of functions of several variable in Derivatives
3.4 Apply partial derivative of functions of several variable in D' Alamber operator
3.5 Apply partial derivative of functions of several variable in Double integral
3.6 Apply partial derivative of functions of several variable in Exterior derivative
3.7 Apply partial derivative of function of several variable in Jacobian matrix and determinant

4.0 CONCLUSION

5.0 SUMMARY

6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

The **partial derivative of f with respect to x** is the derivative of f with respect to x, treating all other variables as constant.

Similarly, the **partial derivative of f with respect to y** is the derivative of f with respect to y, treating all other variables as constant, and so on for other variables. The partial derivatives are written as ∂f/∂x, ∂f/∂y, and so on. The symbol "∂" is used (instead of "d") to remind us that there is more than one variable, and that we are holding the other variables fixed.

OBJECTIVES

In this Unit, you should be able to:

Apply partial derivative of functions of several variable in Chain rule.

Apply partial derivative of functions of several variable in Curl (Mathematics)

Apply partial derivative of functions of several variable in Derivatives

Apply partial derivative of functions of several variable in D' Alamber operator

Apply partial derivative of functions of several variable in Double integral

Apply partial derivative of functions of several variable in Exterior derivative

Apply partial derivative of function of several variable in Jacobian matrix and determinant

MAIN CONTENT

APPLICATIONS OF PARTIAL DERIVATIVE OF FUNCTIONS IN SEVERAL VARIABLE.

Chain rule

Composites of more than two functions

The chain rule can be applied to composites of more than two functions. To take the derivative of a composite of more than two functions, notice that the composite of *f*, *g*, and *h* (in that order) is the composite of *f* with $g \circ h$. The chain rule says that to compute the derivative of $f \circ g \circ h$, it is sufficient to compute the derivative of *f* and the derivative of $g \circ h$. The derivative of *f* can be calculated directly, and the derivative of $g \circ h$ can be calculated by applying the chain rule again.

For concreteness, consider the function

$$y = e^{\sin x^2}.$$

This can be decomposed as the composite of three functions:

$$\begin{aligned} y &= f(u) = e^u, \\ u &= g(v) = \sin v, \\ v &= h(x) = x^2. \end{aligned}$$

Their derivatives are:

$$\frac{dy}{du} = f'(u) = e^u,$$
$$\frac{du}{dv} = g'(v) = \cos v,$$
$$\frac{dv}{dx} = h'(x) = 2x.$$

The chain rule says that the derivative of their composite at the point $x = a$ is:

$$(f \circ g \circ h)'(a) = f'((g \circ h)(a))(g \circ h)'(a) = f'((g \circ h)(a))g'(h(a))h'(a).$$

In Leibniz notation, this is:

$$\frac{dy}{dx} = \left.\frac{dy}{du}\right|_{u=g(h(a))} \cdot \left.\frac{du}{dv}\right|_{v=h(a)} \cdot \left.\frac{dv}{dx}\right|_{x=a},$$

or for short,

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

The derivative function is therefore:

$$\frac{dy}{dx} = e^{\sin x^2} \cdot \cos x^2 \cdot 2x.$$

Another way of computing this derivative is to view the composite function $f \circ g \circ h$ as the composite of $f \circ g$ and h. Applying the chain rule to this situation gives:

$$(f \circ g \circ h)'(a) = (f \circ g)'(h(a))h'(a) = f'(g(h(a))g'(h(a))h'(a).$$

This is the same as what was computed above. This should be expected because $(f \circ g) \circ h = f \circ (g \circ h)$.

The quotient rule

The chain rule can be used to derive some well-known differentiation rules. For example, the quotient rule is a consequence of the chain rule and the product rule. To see this, write the function $f(x)/g(x)$ as the product $f(x) \cdot 1/g(x)$. First apply the product rule:

$$\begin{aligned}\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) &= \frac{d}{dx}\left(f(x) \cdot \frac{1}{g(x)}\right) \\ &= f'(x) \cdot \frac{1}{g(x)} + f(x) \cdot \frac{d}{dx}\left(\frac{1}{g(x)}\right).\end{aligned}$$

To compute the derivative of $1/g(x)$, notice that it is the composite of g with the reciprocal function, that is, the function that sends x to $1/x$. The derivative of the reciprocal function is $-1/x^2$. By applying the chain rule, the last expression becomes:

$$f'(x)\cdot\frac{1}{g(x)}+f(x)\cdot\left(-\frac{1}{g(x)^2}\cdot g'(x)\right)=\frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2},$$

which is the usual formula for the quotient rule.

Derivatives of inverse functions

inverse functions and differentiation

Suppose that $y = g(x)$ has an inverse function. Call its inverse function f so that we have $x = f(y)$. There is a formula for the derivative of f in terms of the derivative of g. To see this, note that f and g satisfy the formula

$$f(g(x)) = x.$$

Because the functions $f(g(x))$ and x are equal, their derivatives must be equal. The derivative of x is the constant function with value 1, and the derivative of $f(g(x))$ is determined by the chain rule. Therefore we have:

$$f'(g(x))g'(x) = 1.$$

To express f' as a function of an independent variable y, we substitute $f(y)$ for x wherever it appears. Then we can solve for f'.

$$\begin{aligned} f'(g(f(y)))g'(f(y)) &= 1 \\ f'(y)g'(f(y)) &= 1 \\ f'(y) &= \frac{1}{g'(f(y))}. \end{aligned}$$

For example, consider the function $g(x) = e^x$. It has an inverse which is denoted $f(y) = \ln y$. Because $g'(x) = e^x$, the above formula says that

$$\frac{d}{dy}\ln y=\frac{1}{e^{\ln y}}=\frac{1}{y}.$$

This formula is true whenever g is differentiable and its inverse f is also differentiable. This formula can fail when one of these conditions is not true. For example, consider $g(x) = x^3$. Its inverse is $f(y) = y^{1/3}$, which is not differentiable at zero. If we attempt to use the above formula to compute the derivative of f at zero, then we must evaluate $1/g'(f(0))$. $f(0) = 0$ and $g'(0) = 0$, so we must evaluate $1/0$, which is undefined. Therefore the formula fails in this case. This is not surprising because f is not differentiable at zero.

Higher derivatives

Faà di Bruno's formula generalizes the chain rule to higher derivatives. The first few derivatives are

$$\frac{d(f\circ g)}{dx}=\frac{df}{dg}\frac{dg}{dx}$$

$$\frac{d^2(f\circ g)}{dx^2}=\frac{d^2f}{dg^2}\left(\frac{dg}{dx}\right)^2+\frac{df}{dg}\frac{d^2g}{dx^2}$$

$$\frac{d^3(f\circ g)}{dx^3}=\frac{d^3f}{dg^3}\left(\frac{dg}{dx}\right)^3+3\frac{d^2f}{dg^2}\frac{dg}{dx}\frac{d^2g}{dx^2}+\frac{df}{dg}\frac{d^3g}{dx^3}$$

$$\frac{d^4(f\circ g)}{dx^4}=\frac{d^4f}{dg^4}\left(\frac{dg}{dx}\right)^4+6\frac{d^3f}{dg^3}\left(\frac{dg}{dx}\right)^2\frac{d^2g}{dx^2}+\frac{d^2f}{dg^2}\left\{4\frac{dg}{dx}\frac{d^3g}{dx^3}+3\left(\frac{d^2g}{dx^2}\right)^2\right\}+\frac{df}{dg}$$

Example

Given $u=x^2+2y$ where $x=r\sin(t)$ and $y=\sin^2(t)$, determine the value of $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial t}$ using the chain rule.

$$\frac{\partial u}{\partial r}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}=(2x)(\sin(t))+(2)(0)=2r\sin^2(t)$$

and

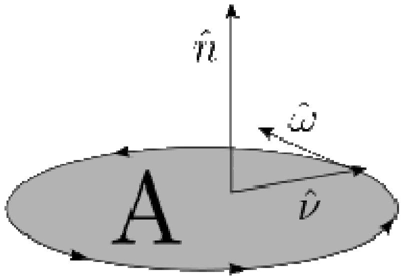
$$\frac{\partial u}{\partial t}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial t}=(2x)(r\cos(t))+(2)(2\sin(t)\cos(t))$$
$$=2(r\sin(t))r\cos(t)+4\sin(t)\cos(t)=2\left(r^2+2\right)\sin(t)\cos(t).$$

Curl (mathematics)

In vector calculus, the **curl** (or **rotor**) is a vector operator that describes the infinitesimal rotation of a 3-dimensional vector field. At every point in the field, the curl is represented by a vector. The attributes of this vector (length and direction) characterize the rotation at that point.

The curl of a vector field **F**, denoted curl **F** or ∇×**F**, at a point is defined in terms of its projection onto various lines through the point. If $\hat{\mathbf{n}}$ is any unit vector, the projection of the curl of **F** onto $\hat{\mathbf{n}}$ is defined to be the limiting value of a closed line integral in a plane orthogonal to $\hat{\mathbf{n}}$ as the path used in the integral becomes infinitesimally close to the point, divided by the area enclosed.

As such, the curl operator maps C^1 functions from $\mathbf{R}^3$ to $\mathbf{R}^3$ to C^0 functions from $\mathbf{R}^3$ to $\mathbf{R}^3$.



Convention for vector orientation of the line integral

Implicitly, curl is defined by:[2]

$$(\nabla \times \mathbf{F}) \cdot \hat{\mathbf{n}} \stackrel{\text{def}}{=} \lim_{A \to \mathbf{0}} \frac{\oint_C \mathbf{F} \cdot d\mathbf{r}}{|A|}$$

The above formula means that the curl of a vector field is defined as the infinitesimal area density of the *circulation* of that field. To this definition fit naturally (i) the Kelvin-Stokes theorem, as a global formula corresponding to the definition, and (ii) the following "easy to memorize" definition of the curl in orthogonal curvilinear coordinates, e.g. in cartesian coordinates, spherical, or cylindrical, or even elliptical or parabolical coordinates:

$$(\text{curl}\ \mathbf{F})_3 = \frac{1}{a_1 a_2} \cdot \left(\frac{\partial (a_2 F_2)}{\partial u_1} - \frac{\partial (a_1 F_1)}{\partial u_2} \right)$$

If (x_1,x_2,x_3) are the Cartesian coordinates and (u_1,u_2,u_3) are the curvilinear coordinates, then $a_i = \sqrt{\sum_{j=1}^{3} \left(\frac{\partial x_j}{\partial u_i} \right)^2}$ is the length of the coordinate vector corresponding to u_i. The remaining two components of curl result from cyclic index-permutation: 3,1,2 → 1,2,3 → 2,3,1.

Usage

In practice, the above definition is rarely used because in virtually all cases, the curl operator can be applied using some set of curvilinear coordinates, for which simpler representations have been derived.

The notation $\nabla\times\mathbf{F}$ has its origins in the similarities to the 3 dimensional cross product, and it is useful as a mnemonic in Cartesian coordinates if we take ∇ as a vector differential operator del. Such notation involving operators is common in physics and algebra. If certain coordinate systems are used, for instance, polar-toroidal coordinates (common in plasma physics) using the notation $\nabla\times\mathbf{F}$ will yield an incorrect result.

Expanded in Cartesian coordinates (see: Del in cylindrical and spherical coordinates for spherical and cylindrical coordinate representations), $\nabla\times\mathbf{F}$ is, for $\mathbf{F}$ composed of $[F_x, F_y, F_z]$:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}$$

where **i**, **j**, and **k** are the unit vectors for the *x*-, *y*-, and *z*-axes, respectively. This expands as follows:[4]

$$\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\mathbf{k}$$

Although expressed in terms of coordinates, the result is invariant under proper rotations of the coordinate axes but the result inverts under reflection.

In a general coordinate system, the curl is given by[2]

$$(\nabla \times \mathbf{F})^k = \epsilon^{k\ell m}\partial_\ell F_m$$

where ε denotes the Levi-Civita symbol, the metric tensor is used to lower the index on **F**, and the Einstein summation convention implies that repeated indices are summed over. Equivalently,

$$(\nabla \times \mathbf{F}) = \boldsymbol{e}_k \epsilon^{k\ell m}\partial_\ell F_m$$

where $\mathbf{e}_k$ are the coordinate vector fields. Equivalently, using the exterior derivative, the curl can be expressed as:

$$\nabla \times \mathbf{F} = \left[\star\left(\mathbf{d}F^\flat\right)\right]^\sharp$$

Here $\flat$ and $\sharp$ are the musical isomorphisms, and $\star$ is the Hodge dual. This formula shows how to calculate the curl of F in any coordinate system, and how to extend the curl to any oriented three dimensional Riemannian manifold. Since this depends on a choice of orientation, curl is a chiral operation. In other words, if the orientation is reversed, then the direction of the curl is also reversed.

Directional derivative

The directional derivative of a scalar function

$$f(\vec{x}) = f(x_1, x_2, \ldots, x_n)$$

along a unit vector

$$\vec{u} = (u_1, \ldots, u_n)$$

is the function defined by the limit

$$\nabla_{\vec{u}} f(\vec{x}) = \lim_{h \to 0^+} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h}.$$

(See other notations below.) If the function *f* is differentiable at $\vec{x}$, then the directional derivative exists along any unit vector $\vec{u}$, and one has

$$\nabla_{\vec{u}} f(\vec{x}) = \nabla f(\vec{x}) \cdot \vec{u}$$

where the ∇ on the right denotes the gradient and $\cdot$ is the Euclidean inner product. At any point $\vec{x}$, the directional derivative of *f* intuitively represents the rate of change in *f* along $\vec{u}$ at the point $\vec{x}$.

One sometimes permits non-unit vectors, allowing the directional derivative to be taken in the direction of $\vec{v}$, where $\vec{v}$ is any nonzero vector. In this case, one must modify the definitions to account for the fact that $\vec{v}$ may not be normalized, so one has

$$\nabla_{\vec{v}} f(\vec{x}) = \lim_{h \to 0^+} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h|\vec{v}|},$$

or in case *f* is differentiable at $\vec{x}$,

$$\nabla_{\vec{v}} f(\vec{x}) = \nabla f(\vec{x}) \cdot \frac{\vec{v}}{|\vec{v}|}$$

Such notation for non-unit vectors (undefined for the zero vector), however, is incompatible with notation used elsewhere in mathematics, where the space of derivations in a derivation algebra is expected to be a vector space.

Notation

Directional derivatives can be also denoted by:

$$\nabla_{\vec{u}} f(\vec{x}) \sim \frac{\partial f(\vec{x})}{\partial u} \sim f'_{\mathbf{u}}(\mathbf{x}) \sim D_{\mathbf{u}} f(\mathbf{x}) \sim \mathbf{u} \cdot \nabla f(\mathbf{x})$$

In the continuum mechanics of solids

Several important results in continuum mechanics require the derivatives of vectors with respect to vectors and of tensors with respect to vectors and tensors.[1] The **directional directive** provides a systematic way of finding these derivatives.

The definitions of directional derivatives for various situations are given below. It is assumed that the functions are sufficiently smooth that derivatives can be taken.

Derivatives of scalar valued functions of vectors

Let $f(\mathbf{v})$ be a real valued function of the vector $\mathbf{v}$. Then the derivative of $f(\mathbf{v})$ with respect to $\mathbf{v}$ (or at $\mathbf{v}$) in the direction $\mathbf{u}$ is the **vector** defined as

$$\frac{\partial f}{\partial \mathbf{v}} \cdot \mathbf{u} = Df(\mathbf{v})[\mathbf{u}] = \left[\frac{d}{d\alpha}\ f(\mathbf{v} + \alpha\ \mathbf{u})\right]_{\alpha=0}$$

for all vectors $\mathbf{u}$.

Properties:

1) If $f(\mathbf{v}) = f_1(\mathbf{v}) + f_2(\mathbf{v})$ then $\frac{\partial f}{\partial \mathbf{v}} \cdot \mathbf{u} = \left(\frac{\partial f_1}{\partial \mathbf{v}} + \frac{\partial f_2}{\partial \mathbf{v}}\right) \cdot \mathbf{u}$

2) If $f(\mathbf{v}) = f_1(\mathbf{v})\ f_2(\mathbf{v})$ then $\frac{\partial f}{\partial \mathbf{v}} \cdot \mathbf{u} = \left(\frac{\partial f_1}{\partial \mathbf{v}} \cdot \mathbf{u}\right)\ f_2(\mathbf{v}) + f_1(\mathbf{v})\ \left(\frac{\partial f_2}{\partial \mathbf{v}} \cdot \mathbf{u}\right)$

3) If $f(\mathbf{v}) = f_1(f_2(\mathbf{v}))$ then $\frac{\partial f}{\partial \mathbf{v}} \cdot \mathbf{u} = \frac{\partial f_1}{\partial f_2}\ \frac{\partial f_2}{\partial \mathbf{v}} \cdot \mathbf{u}$

Derivatives of vector valued functions of vectors

Let $\mathbf{f}(\mathbf{v})$ be a vector valued function of the vector $\mathbf{v}$. Then the derivative of $\mathbf{f}(\mathbf{v})$ with respect to $\mathbf{v}$ (or at $\mathbf{v}$) in the direction $\mathbf{u}$ is the **second order tensor** defined as

$$\frac{\partial \mathbf{f}}{\partial \mathbf{v}} \cdot \mathbf{u} = D\mathbf{f}(\mathbf{v})[\mathbf{u}] = \left[\frac{d}{d\alpha}\ \mathbf{f}(\mathbf{v} + \alpha\ \mathbf{u})\right]_{\alpha=0}$$

for all vectors $\mathbf{u}$.

Properties:

1) If $\mathbf{f}(\mathbf{v}) = \mathbf{f}_1(\mathbf{v}) + \mathbf{f}_2(\mathbf{v})$ then $\frac{\partial \mathbf{f}}{\partial \mathbf{v}} \cdot \mathbf{u} = \left(\frac{\partial \mathbf{f}_1}{\partial \mathbf{v}} + \frac{\partial \mathbf{f}_2}{\partial \mathbf{v}}\right) \cdot \mathbf{u}$

2) If $\mathbf{f}(\mathbf{v}) = \mathbf{f}_1(\mathbf{v}) \times \mathbf{f}_2(\mathbf{v})$ then $\frac{\partial \mathbf{f}}{\partial \mathbf{v}} \cdot \mathbf{u} = \left(\frac{\partial \mathbf{f}_1}{\partial \mathbf{v}} \cdot \mathbf{u}\right) \times \mathbf{f}_2(\mathbf{v}) + \mathbf{f}_1(\mathbf{v}) \times \left(\frac{\partial \mathbf{f}_2}{\partial \mathbf{v}} \cdot \mathbf{u}\right)$

3) If $\mathbf{f}(\mathbf{v}) = \mathbf{f}_1(\mathbf{f}_2(\mathbf{v}))$ then $\frac{\partial \mathbf{f}}{\partial \mathbf{v}} \cdot \mathbf{u} = \frac{\partial \mathbf{f}_1}{\partial \mathbf{f}_2} \cdot \left(\frac{\partial \mathbf{f}_2}{\partial \mathbf{v}} \cdot \mathbf{u}\right)$

Derivatives of scalar valued functions of second-order tensors

Let $f(\boldsymbol{S})$ be a real valued function of the second order tensor $\boldsymbol{S}$. Then the derivative of $f(\boldsymbol{S})$ with respect to $\boldsymbol{S}$ (or at $\boldsymbol{S}$) in the direction $\boldsymbol{T}$ is the **second order tensor** defined as

$$\frac{\partial f}{\partial \boldsymbol{S}} : \boldsymbol{T} = Df(\boldsymbol{S})[\boldsymbol{T}] = \left[\frac{d}{d\alpha}\ f(\boldsymbol{S} + \alpha\ \boldsymbol{T})\right]_{\alpha=0}$$

for all second order tensors $\boldsymbol{T}$.

Properties:

1) If $f(\boldsymbol{S}) = f_1(\boldsymbol{S}) + f_2(\boldsymbol{S})$ then $\frac{\partial f}{\partial \boldsymbol{S}} : \boldsymbol{T} = \left(\frac{\partial f_1}{\partial \boldsymbol{S}} + \frac{\partial f_2}{\partial \boldsymbol{S}}\right) : \boldsymbol{T}$

2) If $f(\boldsymbol{S}) = f_1(\boldsymbol{S})\ f_2(\boldsymbol{S})$ then $\frac{\partial f}{\partial \boldsymbol{S}} : \boldsymbol{T} = \left(\frac{\partial f_1}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)\ f_2(\boldsymbol{S}) + f_1(\boldsymbol{S})\ \left(\frac{\partial f_2}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)$

3) If $f(\boldsymbol{S}) = f_1(f_2(\boldsymbol{S}))$ then $\frac{\partial f}{\partial \boldsymbol{S}} : \boldsymbol{T} = \frac{\partial f_1}{\partial f_2}\ \left(\frac{\partial f_2}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)$

Derivatives of tensor valued functions of second-order tensors

Let $\boldsymbol{F}(\boldsymbol{S})$ be a second order tensor valued function of the second order tensor $\boldsymbol{S}$. Then the derivative of $\boldsymbol{F}(\boldsymbol{S})$ with respect to $\boldsymbol{S}$ (or at $\boldsymbol{S}$) in the direction $\boldsymbol{T}$ is the **fourth order tensor** defined as

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{S}} : \boldsymbol{T} = D\boldsymbol{F}(\boldsymbol{S})[\boldsymbol{T}] = \left[\frac{d}{d\alpha}\ \boldsymbol{F}(\boldsymbol{S} + \alpha\ \boldsymbol{T})\right]_{\alpha=0}$$

for all second order tensors $\boldsymbol{T}$.

Properties:

1) If $\boldsymbol{F}(\boldsymbol{S}) = \boldsymbol{F}_1(\boldsymbol{S}) + \boldsymbol{F}_2(\boldsymbol{S})$ then $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{S}} : \boldsymbol{T} = \left(\frac{\partial \boldsymbol{F}_1}{\partial \boldsymbol{S}} + \frac{\partial \boldsymbol{F}_2}{\partial \boldsymbol{S}}\right) : \boldsymbol{T}$

2) If $\boldsymbol{F}(\boldsymbol{S}) = \boldsymbol{F}_1(\boldsymbol{S}) \cdot \boldsymbol{F}_2(\boldsymbol{S})$ then $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{S}} : \boldsymbol{T} = \left(\frac{\partial \boldsymbol{F}_1}{\partial \boldsymbol{S}} : \boldsymbol{T}\right) \cdot \boldsymbol{F}_2(\boldsymbol{S}) + \boldsymbol{F}_1(\boldsymbol{S}) \cdot \left(\frac{\partial \boldsymbol{F}_2}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)$

3) If $\boldsymbol{F}(\boldsymbol{S}) = \boldsymbol{F}_1(\boldsymbol{F}_2(\boldsymbol{S}))$ then $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{S}} : \boldsymbol{T} = \frac{\partial \boldsymbol{F}_1}{\partial \boldsymbol{F}_2} : \left(\frac{\partial \boldsymbol{F}_2}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)$

4) If $f(\boldsymbol{S}) = f_1(\boldsymbol{F}_2(\boldsymbol{S}))$ then $\frac{\partial f}{\partial \boldsymbol{S}} : \boldsymbol{T} = \frac{\partial f_1}{\partial \boldsymbol{F}_2} : \left(\frac{\partial \boldsymbol{F}_2}{\partial \boldsymbol{S}} : \boldsymbol{T}\right)$

Exterior derivative

The exterior derivative of a differential form of degree k is a differential form of degree $k + 1$. There are a variety of equivalent definitions of the exterior derivative.

Exterior derivative of a function

If f is a smooth function, then the exterior derivative of f is the differential of f. That is, $\mathrm{d}f$ is the unique one-form such that for every smooth vector field X, $\mathrm{d}f(X) = Xf$, where Xf is the directional derivative of f in the direction of X. Thus the exterior derivative of a function (or 0-form) is a one-form.

Exterior derivative of a *k*-form

The exterior derivative is defined to be the unique **R**-linear mapping from k-forms to (k+1)-forms satisfying the following properties:

1. $\mathrm{d}f$ is the differential of f for smooth functions f.
2. $\mathrm{d}(\mathrm{d}f) = 0$ for any smooth function f.
3. $\mathrm{d}(\alpha\wedge\beta) = \mathrm{d}\alpha\wedge\beta + (-1)^p(\alpha\wedge\mathrm{d}\beta)$ where α is a p-form. That is to say, d is an antiderivation of degree 1 on the exterior algebra of differential forms.

The second defining property holds in more generality: in fact, $\mathrm{d}(\mathrm{d}\alpha) = 0$ for any k-form α. This is part of the Poincaré lemma. The third defining property implies as a special case that if f is a function and α a k-form, then $\mathrm{d}(f\alpha) = \mathrm{d}f\wedge\alpha + f\wedge\mathrm{d}\alpha$ *because functions are forms of degree 0.*

Exterior derivative in local coordinates

Alternatively, one can work entirely in a local coordinate system $(x^1,...,x^n)$. First, the coordinate differentials $\mathrm{d}x^1,...,\mathrm{d}x^n$ form a basic set of one-forms within the coordinate chart. Given a multi-index $I = (i_1,...,i_k)$ with $1 \le i_p \le n$ for $1 \le p \le k$, the exterior derivative of a k-form

$$\omega = f_I \mathrm{d}x^{\boldsymbol{I}} = f_{i_1,i_2\cdots i_k}\mathrm{d}x^{i_1} \wedge \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

over $\mathbf{R}^n$ is defined as

$$\mathrm{d}\omega = \sum_{i=1}^{n} \frac{\partial f_I}{\partial x^i}\mathrm{d}x^i \wedge \mathrm{d}x^{\boldsymbol{I}}.$$

For general k-forms $\omega = \Sigma_I f_I \, dx_I$ (where the components of the multi-index I run over all the values in {1, ..., n}), the definition of the exterior derivative is extended linearly. Note that whenever i is one of the components of the multi-index I then $dx_i \wedge dx_I = 0$ (see wedge product).

The definition of the exterior derivative in local coordinates follows from the preceding definition. Indeed, if $\omega = f_I \, dx_{i1} \wedge ... \wedge dx_{ik}$, then

$$\mathrm{d}\omega = \mathrm{d}(f_I \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k})$$

$$= \mathrm{d}f_I \wedge (\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}) + f_I \mathrm{d}(\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k})$$

$$= \mathrm{d}f_I \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} + \sum_{p=1}^{k} (-1)^{(p-1)} f_I \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_{p-1}} \wedge \mathrm{d}^2 x^{i_p} \wedge \mathrm{d}x^{i_{p+1}} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

$$= \mathrm{d}f_I \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

$$= \sum_{i=1}^{n} \frac{\partial f_I}{\partial x^i} \mathrm{d}x^i \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

Here, we have here interpreted f_I as a zero-form, and then applied the properties of the exterior derivative.

Invariant formula

Alternatively, an explicit formula can be given for the exterior derivative of a k-form ω, when paired with k+1 arbitrary smooth vector fields $V_1, V_2, ..., V_k$:

$$\mathrm{d}\omega(V_1, ..., V_k) = \sum_i (-1)^{i-1} V_i \left(\omega(V_1, \ldots, \hat{V_i}, \ldots, V_k) \right)$$

$$+ \sum_{i<j} (-1)^{i+j} \omega([V_i, V_j], V_1, \ldots, \hat{V_i}, \ldots, \hat{V_j}, \ldots, V_k)$$

where $[V_i, V_j]$ denotes Lie bracket and the hat denotes the omission of that element:

$$\omega(V_1, \ldots, \hat{V_i}, \ldots, V_k) = \omega(V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_k).$$

In particular, for 1-forms we have: $d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y])$, where X and Y are vector fields.

Examples

1. Consider $\sigma = u \, dx^1 \wedge dx^2$ over a 1-form basis $dx^1,...,dx^n$. The exterior derivative is:

$$d\sigma = d(u) \wedge dx^{\mathbf{1}} \wedge dx^{\mathbf{2}}$$

$$= \left(\sum_{i=1}^{n} \frac{\partial u}{\partial x^i} dx^i \right) \wedge dx^1 \wedge dx^2$$

$$= \sum_{i=3}^{n} \left(\frac{\partial u}{\partial x^i} dx^i \wedge dx^1 \wedge dx^2 \right)$$

The last formula follows easily from the properties of the wedge product. Namely, $dx^i \wedge dx^i = 0$.

2. For a 1-form $\sigma = u\,dx + v\,dy$ defined over $\mathbf{R}^2$. We have, by applying the above formula to each term (consider $x^1 = x$ and $x^2 = y$) the following sum,

$$d\sigma = \left(\sum_{i=1}^{2} \frac{\partial u}{\partial x^i} dx^i \wedge dx \right) + \left(\sum_{i=1}^{2} \frac{\partial v}{\partial x^i} dx^i \wedge dy \right)$$

$$= \left(\frac{\partial u}{\partial x} dx \wedge dx + \frac{\partial u}{\partial y} dy \wedge dx \right) + \left(\frac{\partial v}{\partial x} dx \wedge dy + \frac{\partial v}{\partial y} dy \wedge dy \right)$$

$$= 0 - \frac{\partial u}{\partial y} dx \wedge dy + \frac{\partial v}{\partial x} dx \wedge dy + 0$$

$$= \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx \wedge dy.$$

D'Alembert operator

In special relativity, electromagnetism and wave theory, the **d'Alembert operator** (represented by a box: $\sqcup$), also called the **d'Alembertian** or the **wave operator**, is the Laplace operator of Minkowski space. The operator is named for French mathematician and physicist Jean le Rond d'Alembert. In Minkowski space in standard coordinates (t, x, y, z) it has the form:

$$\sqsupset = \partial_\mu \partial^\mu = g_{\mu\nu} \partial^\nu \partial^\mu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}$$

$$= \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta$$

Applications

he Klein–Gordon equation has the form

$$(\Box + m^2)\psi = 0.$$

The wave equation for the electromagnetic field in vacuum is

$$\Box A^{\mu} = 0$$

where A^{μ} is the electromagnetic four-potential.

The wave equation for small vibrations is of the form

$$\Box_c u\,(x,t) \equiv u_{tt} - c^2 u_{xx} = 0,$$

where $u(x,t)$ is the displacement.

Green's function

The Green's function $G(x-x')$ for the d'Alembertian is defined by the equation

$$\Box G(x-x') = \delta(x-x')$$

where $\delta(x-x')$ is the Dirac delta function and x and x' are two points in Minkowski space.

Explicitly we have

$$G(t,x,y,z) = \frac{1}{2\pi}\Theta(t)\delta(t^2 - x^2 - y^2 - z^2)$$

where Θ is the Heaviside step function

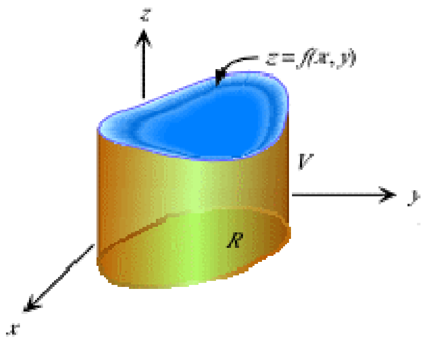
Double Integral

The **double integral of f(x, y) over the region R in the xy-plane** is defined as

$$\iint_R f(x,y)\,dx\,dy$$

= (volume above R and under the graph of f)
- (volume below R and above the graph of f).

- The following figure illustrates this volume (in the case that the graph of f is above the region R).

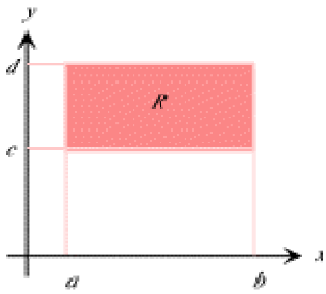
- 

- ---

- **Computing Double Integrals**
  If R is the rectangle $a \le x \le b$ and $c \le y \le d$ (see figure below) then

$$\iint_R f(x, y)\,dx\,dy = \int_c^d \left[\int_a^b f(x, y)\,dx \right] dy$$

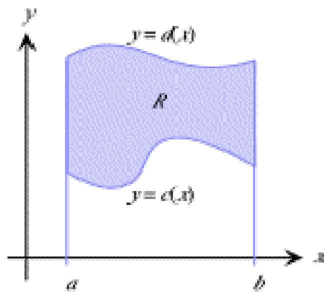
$$= \int_a^b \left[\int_c^d f(x, y)\,dy \right] dx$$

- 

- ---

- If R is the region $a \le x \le b$ and $c(x) \le y \le d(x)$ (see figure below) then we integrate over R according to the following equation.

$$\iint_R f(x, y)\,dx\,dy = \int_a^b \left[\int_{c(x)}^{d(x)} f(x, y)\,dy \right] dx$$

- 

JACOBIAN MATRIX

The Jacobian of a function describes the orientation of a tangent plane to the function at a given point. In this way, the Jacobian generalizes the gradient of a scalar valued function of multiple variables which itself generalizes the derivative of a scalar-valued function of a scalar. Likewise, the Jacobian can also be thought of as describing the amount of "stretching" that a transformation imposes. For example, if $(x_2,y_2) = f(x_1,y_1)$ is used to transform an image, the Jacobian of f, $J(x_1,y_1)$ describes how much the image in the neighborhood of (x_1,y_1) is stretched in the x and y directions.

If a function is differentiable at a point, its derivative is given in coordinates by the Jacobian, but a function doesn't need to be differentiable for the Jacobian to be defined, since only the partial derivatives are required to exist.

The importance of the Jacobian lies in the fact that it represents the best linear approximation to a differentiable function near a given point. In this sense, the Jacobian is the derivative of a multivariate function.

If $\mathbf{p}$ is a point in $\mathbf{R}^n$ and F is differentiable at $\mathbf{p}$, then its derivative is given by $J_F(\mathbf{p})$. In this case, the linear map described by $J_F(\mathbf{p})$ is the best linear approximation of F near the point $\mathbf{p}$, in the sense that

$$F(\mathbf{x}) = F(\mathbf{p}) + J_F(\mathbf{p})(\mathbf{x}-\mathbf{p}) + o(\|\mathbf{x}-\mathbf{p}\|)$$

for $\mathbf{x}$ close to $\mathbf{p}$ and where o is the little o-notation (for $x \to p$) and $\|\mathbf{x}-\mathbf{p}\|$ is the distance between $\mathbf{x}$ and $\mathbf{p}$.

In a sense, both the gradient and Jacobian are "first derivatives" — the former the first derivative of a *scalar function* of several variables, the latter the first derivative of a *vector function* of several variables. In general, the gradient can be regarded as a special version of the Jacobian: it is the Jacobian of a scalar function of several variables.

The Jacobian of the gradient has a special name: the Hessian matrix, which in a sense is the "second derivative" of the scalar function of several variables in question.

Inverse

According to the inverse function theorem, the matrix inverse of the Jacobian matrix of an invertible function is the Jacobian matrix of the *inverse* function. That is, for some function $F : \mathbf{R}^n \to \mathbf{R}^n$ and a point p in $\mathbf{R}^n$,

$$J_{F^{-1}}(F(p)) = [J_F(p)]^{-1}.$$

It follows that the (scalar) inverse of the Jacobian determinant of a transformation is the Jacobian determinant of the inverse transformation.

Uses

Dynamical systems

Consider a dynamical system of the form $x' = F(x)$, where x' is the (component-wise) time derivative of x, and $F : \mathbf{R}^n \to \mathbf{R}^n$ is continuous and differentiable. If $F(x_0) = 0$, then x_0 is a stationary point (also called a fixed point). The behavior of the system near a stationary point is related to the eigenvalues of $J_F(x_0)$, the Jacobian of F at the stationary point.[1] Specifically, if the eigenvalues all have a negative real part, then the system is stable in the operating point, if any eigenvalue has a positive real part, then the point is unstable.

Newton's method

A system of coupled nonlinear equations can be solved iteratively by Newton's method. This method uses the Jacobian matrix of the system of equations.

The following is the dettail code in MATLAB

```
function s = jacobian(f, x, tol) % f is a multivariable function handle, x is a starting point

  if nargin == 2
     tol = 10^-5;
  end
  while 1
     % if x and f(x) are row vectors, we need transpose operations here
     y = x' - jacob(f, x)\f(x)';            % get the next point
     if norm(f(y))<tol                      % check error tolerate
        s = y';
        return;
     end
     x = y';
  end

function j = jacob(f, x) % approximately calculate Jacobian matrix

  k = length(x);
  j = zeros(k, k);
  for m = 1: k
     x2 = x;
     x2(m) =x(m)+0.001;
     j(m, :) = 1000*(f(x2)-f(x));       % partial derivatives in m-th row
  end
```

Jacobian determinant

If $m = n$, then F is a function from n-space to n-space and the Jacobian matrix is a square matrix. We can then form its determinant, known as the **Jacobian determinant**. The Jacobian determinant is sometimes simply called "the Jacobian."

The Jacobian determinant at a given point gives important information about the behavior of F near that point. For instance, the continuously differentiable function F is invertible near a point $\mathbf{p} \in \mathbf{R}^n$ if the Jacobian determinant at **p** is non-zero. This is the inverse function theorem. Furthermore, if the Jacobian determinant at **p** is positive, then F preserves orientation near **p**; if it is negative, F reverses orientation. The absolute value of the Jacobian determinant at **p** gives us the factor by which the function F expands or shrinks volumes near **p**; this is why it occurs in the general substitution rule.

Uses

The Jacobian determinant is used when making a change of variables when evaluating a multiple integral of a function over a region within its domain. To accommodate for the change of coordinates the magnitude of the Jacobian determinant arises as a multiplicative factor within the integral. Normally it is required that the change of coordinates be done in a manner which maintains an injectivity between the coordinates that determine the domain. The Jacobian determinant, as a result, is usually well defined.

Examples

Example 1. The transformation from spherical coordinates (r, θ, φ) to Cartesian coordinates (x_1, x_2, x_3) , is given by the function $F : \mathbf{R}^+ \times [0,\pi] \times [0,2\pi) \to \mathbf{R}^3$ with components:

$$x_1 = r\,\sin\theta\,\cos\phi$$

$$x_2 = r\,\sin\theta\,\sin\phi$$

$$x_3 = r\,\cos\theta.$$

The Jacobian matrix for this coordinate change is

$$J_F(r,\theta,\phi) = \begin{bmatrix} \dfrac{\partial x_1}{\partial r} & \dfrac{\partial x_1}{\partial \theta} & \dfrac{\partial x_1}{\partial \phi} \\ \dfrac{\partial x_2}{\partial r} & \dfrac{\partial x_2}{\partial \theta} & \dfrac{\partial x_2}{\partial \phi} \\ \dfrac{\partial x_3}{\partial r} & \dfrac{\partial x_3}{\partial \theta} & \dfrac{\partial x_3}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \sin\theta\,\cos\phi & r\,\cos\theta\,\cos\phi & -r\,\sin\theta\,\sin\phi \\ \sin\theta\,\sin\phi & r\,\cos\theta\,\sin\phi & r\,\sin\theta\,\cos\phi \\ \cos\theta & -r\,\sin\theta & 0 \end{bmatrix}.$$

The determinant is $r^2 \sin\theta$. As an example, since $dV = dx_1\, dx_2\, dx_3$ this determinant implies that the differential volume element $dV = r^2 \sin\theta\, dr\, d\theta\, d$□. Nevertheless this determinant varies with coordinates. To avoid any variation the new coordinates can be defined as $w_1 = \dfrac{r^3}{3},\ w_2 = -\cos\theta,\ w_3 = \phi.$[2] Now the determinant equals to 1 and volume element becomes $r^2 dr\ \sin\theta\ d\theta\ d\phi = dw_1 dw_2 dw_3.$

Example 2. The Jacobian matrix of the function $F : \mathbf{R}^3 \to \mathbf{R}^4$ with components

$$y_1 = x_1$$

$$y_2 = 5x_3$$

$$y_3 = 4x_2^2 - 2x_3$$

$$y_4 = x_3 \sin(x_1)$$

is

$$J_F(x_1,x_2,x_3) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \\ \frac{\partial y_4}{\partial x_1} & \frac{\partial y_4}{\partial x_2} & \frac{\partial y_4}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 8x_2 & -2 \\ x_3\cos(x_1) & 0 & \sin(x_1) \end{bmatrix}.$$

This example shows that the Jacobian need not be a square matrix.

Example 3.

$$x = r\cos\phi;$$
$$y = r\sin\phi.$$

$$J(r,\phi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \frac{\partial(r\cos\phi)}{\partial r} & \frac{\partial(r\cos\phi)}{\partial \phi} \\ \frac{\partial(r\sin\phi)}{\partial r} & \frac{\partial(r\sin\phi)}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{bmatrix}$$

The Jacobian determinant is equal to *r*. This shows how an integral in the Cartesian coordinate system is transformed into an integral in the polar coordinate system:

$$\iint_A dx\,dy = \iint_B r\,dr\,d\phi.$$

Example 4. The Jacobian determinant of the function $F : \mathbf{R}^3 \to \mathbf{R}^3$ with components

$$y_1 = 5x_2$$
$$y_2 = 4x_1^2 - 2\sin(x_2x_3)$$
$$y_3 = x_2x_3$$

is

$$\begin{vmatrix} 0 & 5 & 0 \\ 8x_1 & -2x_3\cos(x_2x_3) & -2x_2\cos(x_2x_3) \\ 0 & x_3 & x_2 \end{vmatrix} = -8x_1 \cdot \begin{vmatrix} 5 & 0 \\ x_3 & x_2 \end{vmatrix} = -40x_1x_2.$$

From this we see that *F* reverses orientation near those points where x_1 and x_2 have the same sign; the function is locally invertible everywhere except near points where $x_1 = 0$ or $x_2 = 0$. Intuitively, if you start with a tiny object around the point (1,1,1) and apply *F* to that object, you will get an object set with approximately 40 times the volume of the original one

CONCLUSION

In this unit you have applied partial derivative of functions of several variable to solve chain rule and curl (mathematics) . You have also applied partial derivative of functions of several variable solve derivatives and D' Alamber operator. You have applied partial derivative of functions of several variable in Double integral and Exterior derivative. You also used partial derivative of function of several variable in Jacobian matrix and determinant.

SUMMARY

In this unit, you have studied the :

Application of partial derivative of functions of several variable in Chain rule.

Application of partial derivative of functions of several variable in Curl (Mathematics)

Application of partial derivative of functions of several variable in Derivatives

Application of partial derivative of functions of several variable in D' Alamber operator

Application of partial derivative of functions of several variable in Double integral

Application of partial derivative of functions of several variable in Exterior derivative

Application of partial derivative of function of several variable in Jacobian matrix and determinant

TUTOR – MARKED ASSIGNMENT

1. Find the equation of the tangent plane to $z = \ln(2x+y) z = \ln(2x+y)$ at $(-1,3)$

2.Find the linear approximation to $z = 3 + \frac{x^2}{16} + \frac{y^2}{9}$ ATA a at $(-4,3)$

2.Find the absolute minimum and absolute maximum of $f(x,y) = x^2 + 4y^2 - 2x^2y + 4$ on the rectangle given by $-1 \le x \le 1$ and $-1 \le y \le 1$

4. Find the absolute minimum and absolute maximum of $f(x,y) = 2x^2 - y^2 + 6y$ on the disk of radius 4, $x^2 + y^2 \le 16$
5. **Find the partial derivatives of the following in the second order :**

a. F(x,y) = $\chi^2 - 2xy + 6x - 2y + 1$

b. F(x,y) = e^{xy}

References

J. E. Marsden and T. J. R. Hughes, 2000, *Mathematical Foundations of Elasticity*, Dover.

Hildebrand, F. B. (1976). *Advanced Calculus for Applications*. Prentice Hall. ISBN 0130111899.

Hernandez Rodriguez and Lopez Fernandez, *A Semiotic Reflection on the Didactics of the Chain Rule*, The Montana Mathematics Enthusiast, ISSN 1551-3440, Vol. 7, nos.2&3, pp.321–332.

Apostol, Tom (1974). *Mathematical analysis* (2nd ed. ed.). Addison Wesley. Theorem 5.5.

Flanders, Harley (1989). *Differential forms with applications to the physical sciences*. New York: Dover Publications. pp. 20. ISBN 0-486-66169-5.

Ramanan, S. (2005). *Global calculus*. Providence, Rhode Island: American Mathematical Society. pp. 54. ISBN 0-8218-3702-8.

Conlon, Lawrence (2001). *Differentiable manifolds*. Basel, Switzerland: Birkhäuser. pp. 239. ISBN 0-8176-4134-3.

Darling, R. W. R. (1994). *Differential forms and connections*. Cambridge, UK: Cambridge University Press. pp. 35. ISBN 0-521-46800-0.

Retrieved from "http://en.wikipedia.org/w/index.php?title=Exterior_derivative&oldid=472668495"
D.K. Arrowsmith and C.M. Place, *Dynamical Systems*, Section 3.3, Chapman & Hall, London, 1992. ISBN 0-412-39080-9.

Taken from http://www.sjcrothers.plasmaresources.com/schwarzschild.pdf - On the Gravitational Field of a Mass Point according to Einstein’s Theory by K. Schwarzschild - arXiv:physics/9905030 v1 (text of the original paper, in Wikisource

MODULE 3 TOTAL DERIVATIVES OF FUNCTION OF SEVERAL VARIABLES

-Unit 1:Derivative

-Unit 2: Total derivative.

-Unit 3:Application of Total derivative.

UNIT 1 : DERIVATIVE

CONTENT

1.0 INTRODUCTION

2.0 OBJECTIVES

3.0 MAIN CONTENT

Solve directional derivatives

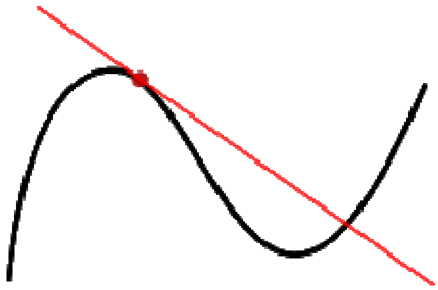
Use derivative to solve Total derivative, total differential and Jacobian matrix

4.0 Conclusion

5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

Introduction

This article is an overview of the term as used in calculus. For a less technical overview of the subject, see Differential calculus. For other uses, see Derivative (disambiguation).



The graph of a function, drawn in black, and a tangent line to that function, drawn in red. The slope of the tangent line is equal to the derivative of the function at the marked point.

In calculus, a branch of mathematics, the **derivative** is a measure of how a function changes as its input changes. Loosely speaking, a derivative can be thought of as how much one quantity is changing in response to changes in some other quantity; for example, the derivative of the position of a moving object with respect to time is the object's instantaneous velocity.

The derivative of a function at a chosen input value describes the best linear approximation of the function near that input value. For a real-valued function of a single real variable, the derivative at a point equals the slope of the tangent line to the graph of the function at that point. In higher dimensions, the derivative of a function at a point is a linear transformation called the linearization.[1] A closely related notion is the differential of a function.

The process of finding a derivative is called **differentiation**. The reverse process is called **antidifferentiation**. The fundamental theorem of calculus states that antidifferentiation is the same as integration. Differentiation and integration constitute the two fundamental operations in single-variable calculus.

OBJECTIVES

In this unit, you should be able to :

Solve directional derivatives

Use derivative to solve Total derivative, total differential and Jacobian matrix

Main content

Directional derivatives

If f is a real-valued function on $\mathbf{R}^n$, then the partial derivatives of f measure its variation in the direction of the coordinate axes. For example, if f is a function of x and y, then its partial derivatives measure the variation in f in the x direction and the y direction. They do not, however, directly measure the variation of f in any other direction, such as along the diagonal line $y = x$. These are measured using directional derivatives. Choose a vector

$$\mathbf{v} = (v_1, \ldots, v_n).$$

The **directional derivative** of f in the direction of $\mathbf{v}$ at the point $\mathbf{x}$ is the limit

$$D_{\mathbf{v}} f(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}.$$

In some cases it may be easier to compute or estimate the directional derivative after changing the length of the vector. Often this is done to turn the problem into the computation of a directional derivative in the direction of a unit vector. To see how this works, suppose that $\mathbf{v} = \lambda\mathbf{u}$. Substitute $h = k/\lambda$ into the difference quotient. The difference quotient becomes:

$$\frac{f(\mathbf{x} + (k/\lambda)(\lambda\mathbf{u})) - f(\mathbf{x})}{k/\lambda} = \lambda \cdot \frac{f(\mathbf{x} + k\mathbf{u}) - f(\mathbf{x})}{k}.$$

This is λ times the difference quotient for the directional derivative of f with respect to $\mathbf{u}$. Furthermore, taking the limit as h tends to zero is the same as taking the limit as k tends to zero because h and k are multiples of each other. Therefore $D_{\mathbf{v}}(f) = \lambda D_{\mathbf{u}}(f)$. Because of this rescaling property, directional derivatives are frequently considered only for unit vectors.

If all the partial derivatives of f exist and are continuous at $\mathbf{x}$, then they determine the directional derivative of f in the direction $\mathbf{v}$ by the formula:

$$D_{\mathbf{v}} f(\boldsymbol{x}) = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}.$$

This is a consequence of the definition of the total derivative. It follows that the directional derivative is linear in $\mathbf{v}$, meaning that $D_{\mathbf{v} + \mathbf{w}}(f) = D_{\mathbf{v}}(f) + D_{\mathbf{w}}(f)$.

The same definition also works when f is a function with values in $\mathbf{R}^m$. The above definition is applied to each component of the vectors. In this case, the directional derivative is a vector in $\mathbf{R}^m$.

Total derivative, total differential and Jacobian matrix

When f is a function from an open subset of $\mathbf{R}^n$ to $\mathbf{R}^m$, then the directional derivative of f in a chosen direction is the best linear approximation to f at that point and in that direction. But when $n > 1$, no single directional derivative can give a complete picture of the behavior of f. The total derivative, also called the (**total**) **differential**, gives a complete picture by considering all directions at once. That is, for any vector **v** starting at **a**, the linear approximation formula holds:

$$f(\mathbf{a}+\mathbf{v}) \approx f(\mathbf{a}) + f'(\mathbf{a})\mathbf{v}.$$

Just like the single-variable derivative, $f'(\mathbf{a})$ is chosen so that the error in this approximation is as small as possible.

If n and m are both one, then the derivative $f'(a)$ is a number and the expression $f'(a)v$ is the product of two numbers. But in higher dimensions, it is impossible for $f'(\mathbf{a})$ to be a number. If it were a number, then $f'(\mathbf{a})\mathbf{v}$ would be a vector in $\mathbf{R}^n$ while the other terms would be vectors in $\mathbf{R}^m$, and therefore the formula would not make sense. For the linear approximation formula to make sense, $f'(\mathbf{a})$ must be a function that sends vectors in $\mathbf{R}^n$ to vectors in $\mathbf{R}^m$, and $f'(\mathbf{a})\mathbf{v}$ must denote this function evaluated at **v**.

To determine what kind of function it is, notice that the linear approximation formula can be rewritten as

$$f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}) \approx f'(\mathbf{a})\mathbf{v}.$$

Notice that if we choose another vector **w**, then this approximate equation determines another approximate equation by substituting **w** for **v**. It determines a third approximate equation by substituting both **w** for **v** and **a** + **v** for **a**. By subtracting these two new equations, we get

$$f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}+\mathbf{w}) + f(\mathbf{a}) \approx f'(\mathbf{a}+\mathbf{v})\mathbf{w} - f'(\mathbf{a})\mathbf{w}.$$

If we assume that **v** is small and that the derivative varies continuously in **a**, then $f'(\mathbf{a} + \mathbf{v})$ is approximately equal to $f'(\mathbf{a})$, and therefore the right-hand side is approximately zero. The left-hand side can be rewritten in a different way using the linear approximation formula with **v** + **w** substituted for **v**. The linear approximation formula implies:

$$\begin{aligned} 0 &\approx f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a}+\mathbf{w}) + f(\mathbf{a}) \\ &= (f(\mathbf{a}+\mathbf{v}+\mathbf{w}) - f(\mathbf{a})) - (f(\mathbf{a}+\mathbf{v}) - f(\mathbf{a})) - (f(\mathbf{a}+\mathbf{w}) - f(\mathbf{a})) \\ &\approx f'(\mathbf{a})(\mathbf{v}+\mathbf{w}) - f'(\mathbf{a})\mathbf{v} - f'(\mathbf{a})\mathbf{w}. \end{aligned}$$

This suggests that $f'(\mathbf{a})$ is a linear transformation from the vector space $\mathbf{R}^n$ to the vector space $\mathbf{R}^m$. In fact, it is possible to make this a precise derivation by measuring the error in the approximations. Assume that the error in these linear approximation formula is bounded by a constant times $\|\mathbf{v}\|$, where the constant is independent of **v** but depends continuously on **a**. Then, after adding an appropriate error term, all of the above approximate equalities can be rephrased as inequalities. In particular, $f'(\mathbf{a})$ is a linear transformation up to a small error term. In the limit as **v** and **w** tend to zero, it must therefore be a linear transformation. Since we define the total derivative by taking a limit as **v** goes to zero, $f'(\mathbf{a})$ must be a linear transformation.

In one variable, the fact that the derivative is the best linear approximation is expressed by the fact that it is the limit of difference quotients. However, the usual difference quotient does not make sense in higher dimensions because it is not usually possible to divide vectors. In particular, the numerator and denominator of the difference quotient are not even in the same vector space: The numerator lies in the codomain $\mathbf{R}^m$ while the denominator lies in the domain $\mathbf{R}^n$. Furthermore, the derivative is a linear transformation, a different type of object from both the numerator and denominator. To make precise the idea that $f'(\mathbf{a})$ is the best linear approximation, it is necessary to adapt a different formula for the one-variable derivative in which these problems disappear. If $f : \mathbf{R} \to \mathbf{R}$, then the usual definition of the derivative may be manipulated to show that the derivative of f at a is the unique number $f'(a)$ such that

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0.$$

This is equivalent to

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - f'(a)h|}{|h|} = 0$$

because the limit of a function tends to zero if and only if the limit of the absolute value of the function tends to zero. This last formula can be adapted to the many-variable situation by replacing the absolute values with norms.

The definition of the **total derivative** of f at **a**, therefore, is that it is the unique linear transformation $f'(\mathbf{a}) : \mathbf{R}^n \to \mathbf{R}^m$ such that

$$\lim_{\|\mathbf{h}\|\to 0} \frac{\|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - f'(\mathbf{a})\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Here **h** is a vector in $\mathbf{R}^n$, so the norm in the denominator is the standard length on $\mathbf{R}^n$. However, $f'(\mathbf{a})\mathbf{h}$ is a vector in $\mathbf{R}^m$, and the norm in the numerator is the standard length on $\mathbf{R}^m$. If v is a vector starting at a, then $f'(\mathbf{a})\mathbf{v}$ is called the pushforward of **v** by f and is sometimes written $f_*\mathbf{v}$.

If the total derivative exists at **a**, then all the partial derivatives and directional derivatives of f exist at **a**, and for all **v**, $f'(\mathbf{a})\mathbf{v}$ is the directional derivative of f in the direction **v**. If we write f using coordinate functions, so that $f = (f_1, f_2, ..., f_m)$, then the total derivative can be expressed using the partial derivatives as a matrix. This matrix is called the **Jacobian matrix** of f at **a**:

$$f'(\mathbf{a}) = \mathrm{Jac}_{\mathbf{a}} = \left(\frac{\partial f_i}{\partial x_j}\right)_{ij}.$$

The existence of the total derivative $f'(\mathbf{a})$ is strictly stronger than the existence of all the partial derivatives, but if the partial derivatives exist and are continuous, then the total derivative exists, is given by the Jacobian, and depends continuously on **a**.

The definition of the total derivative subsumes the definition of the derivative in one variable. That is, if f is a real-valued function of a real variable, then the total derivative exists if and only if the usual derivative exists. The Jacobian matrix reduces to a 1×1 matrix whose only entry is the derivative $f'(x)$. This 1×1 matrix satisfies the property that $f(a + h) - f(a) - f'(a)h$ is approximately zero, in other words that

$$f(a+h) \approx f(a) + f'(a)h.$$

Up to changing variables, this is the statement that the function $x \mapsto f(a) + f'(a)(x-a)$ is the best linear approximation to f at a.

The total derivative of a function does not give another function in the same way as the one-variable case. This is because the total derivative of a multivariable function has to record much more information than the derivative of a single-variable function. Instead, the total derivative gives a function from the tangent bundle of the source to the tangent bundle of the target.

The natural analog of second, third, and higher-order total derivatives is not a linear transformation, is not a function on the tangent bundle, and is not built by repeatedly taking the total derivative. The analog of a higher-order derivative, called a jet, cannot be a linear transformation because higher-order derivatives reflect subtle geometric information, such as concavity, which cannot be described in terms of linear data such as vectors. It cannot be a function on the tangent bundle because the tangent bundle only has room for the base space and the directional derivatives. Because jets capture higher-order information, they take as arguments additional coordinates representing higher-order changes in direction. The space determined by these additional coordinates is called the jet bundle. The relation between the total derivative and the partial derivatives of a function is paralleled in the relation between the kth order jet of a function and its partial derivatives of order less than or equal to k.

Conclusion

In this unit, you have used derivative to solve problems on directional derivatives and have also solve problems on total derivative,total differentiation and Jacobian matrix.

Summary

In this unit you have studied :

Solve directional derivatives

Use derivative to solve problems on total derivative, total differentiation and Jacobian matrix.

Tutor-Marked Assignment

1.Evaluate the derivative of F(x,y,z) = $3(x^2 + y)\sin(z^2)$

2.Find the derivative of F(x,y,z) = $xy^3 + z^4$

3.Let F(x,y,z) = $x^5 + y^4 z^3 + \sin z^2$,find the derivative.

4.Evaluate the derivatives of F(x,y,z) = $x^2 - xy + z^4$

5.Find the derivative of F(x,y,z) = $\dfrac{\sin x + \cos^2 x}{\tan^{-1} x}$

References/Further Readings

- Anton, Howard; Bivens, Irl; Davis, Stephen (February 2, 2005), *Calculus: Early Transcendentals Single and Multivariable* (8th ed.), New York: Wiley, ISBN 978-0-471-47244-5

- Apostol, Tom M. (June 1967), *Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00005-1

- Apostol, Tom M. (June 1969), *Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00007-5

- Courant, Richard; John, Fritz (December 22, 1998), *Introduction to Calculus and Analysis, Vol. 1*, Springer-Verlag, ISBN 978-3-540-65058-4

- Eves, Howard (January 2, 1990), *An Introduction to the History of Mathematics* (6th ed.), Brooks Cole, ISBN 978-0-03-029558-4

- Larson, Ron; Hostetler, Robert P.; Edwards, Bruce H. (February 28, 2006), *Calculus: Early Transcendental Functions* (4th ed.), Houghton Mifflin Company, ISBN 978-0-618-60624-5

- Spivak, Michael (September 1994), *Calculus* (3rd ed.), Publish or Perish, ISBN 978-0-914098-89-8

- Stewart, James (December 24, 2002), *Calculus* (5th ed.), Brooks Cole, ISBN 978-0-534-39339-7

- Thompson, Silvanus P. (September 8, 1998), *Calculus Made Easy* (Revised, Updated, Expanded ed.), New York: St. Martin's Press, ISBN 978-0-312-18548-0

UNIT 2: TOTAL DERIVATIVE

CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT

3.1 Differentiate with indirect dependent
3.2 The total derivative via differentials
3.3 The total derivative as a linear map
3.4 sTotal differential equation.

4.**0 Conclusion**
5.0 Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

1.0 INTRODUCTION

In the mathematical field of differential calculus, the term **total derivative** has a number of closely related meanings.

The total derivative (full derivative) of a function *f*, of several variables, e.g., *t*, *x*, *y*, etc., with respect to one of its input variables, e.g., *t*, is different from the partial derivative (∂). Calculation of the total derivative of *f* with respect to *t* does not assume that the other arguments are constant while *t* varies; instead, it allows the other arguments to depend on *t*. The total derivative adds in these *indirect dependencies* to find the overall dependency of *f* on *t*. For example, the total derivative of *f*(*t*,*x*,*y*) with respect to *t* is

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

Consider multiplying both sides of the equation by the differential $\mathrm{d}t$:

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

The result will be the differential change $\mathrm{d}f$ in the function *f*. Because *f* depends on *t*, some of that change will be due to the partial derivative of *f* with respect to *t*. However, some of that change will

also be due to the partial derivatives of *f* with respect to the variables *x* and *y*. So, the differential $\mathrm{d}t$ is applied to the total derivatives of *x* and *y* to find differentials $\mathrm{d}x$ and $\mathrm{d}y$, which can then be used to find the contribution to $\mathrm{d}f$.

- It refers to a differential operator such as

$$\frac{\mathrm{d}}{\mathrm{d}x} = \frac{\partial}{\partial x} + \sum_{j=1}^{k} \frac{\mathrm{d}y_j}{\mathrm{d}x}\frac{\partial}{\partial y_j},$$

  which computes the total derivative of a function (with respect to *x* in this case).

- It refers to the (total) differential d*f* of a function, either in the traditional language of infinitesimals or the modern language of differential forms.

- A differential of the form

$$\sum_{j=1}^{k} f_j(x_1, \ldots, x_k)\,\mathrm{d}x_j$$

  is called a **total differential** or an **exact differential** if it is the differential of a function. Again this can be interpreted infinitesimally, or by using differential forms and the exterior derivative.

- It is another name for the derivative as a linear map, i.e., if *f* is a differentiable function from $\mathbf{R}^n$ to $\mathbf{R}^m$, then the (total) derivative (or differential) of *f* at $x \in \mathbf{R}^n$ is the linear map from $\mathbf{R}^n$ to $\mathbf{R}^m$ whose matrix is the Jacobian matrix of *f* at *x*.

- It is a synonym for the gradient, which is essentially the derivative of a function from $\mathbf{R}^n$ to **R**.

- It is sometimes used as a synonym for the material derivative, $\frac{D\mathbf{u}}{Dt}$, in fluid mechanics

2.0 OBJECTIVE

After studying this unit, you should be able to

differentiate with indirect dependent

find the derivative via differentials

solve total derivative as a linear map

know total differential equation.

3.0 MAIN CONTENT

Differentiation with indirect dependencies

Suppose that *f* is a function of two variables, *x* and *y*. Normally these variables are assumed to be independent. However, in some situations they may be dependent on each other. For example *y* could be a function of *x*, constraining the domain of *f* to a curve in R^2. In this case the partial derivative of *f* with respect to *x* does not give the true rate of change of *f* with respect to changing *x* because changing *x* necessarily changes *y*. The **total derivative** takes such dependencies into account.

For example, suppose

$$f(x,y) = xy.$$

The rate of change of *f* with respect to *x* is usually the partial derivative of *f* with respect to *x*; in this case,

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

However, if *y* depends on *x*, the partial derivative does not give the true rate of change of *f* as *x* changes because it holds *y* fixed.

Suppose we are constrained to the line

$$y = x$$

then

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

In that case, the total derivative of *f* with respect to *x* is

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 2x.$$

Notice that this is not equal to the partial derivative:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 2x \neq \frac{\partial f}{\partial x} = y = x.$$

While one can often perform substitutions to eliminate indirect dependencies, the chain rule provides for a more efficient and general technique. Suppose $M(t, p_1, ..., p_n)$ is a function of time t and n variables p_i which themselves depend on time. Then, the total time derivative of M is

$$\frac{\mathrm{d}\,M}{\mathrm{d}\,t} = \frac{\mathrm{d}}{\mathrm{d}\,t} M\big(t, p_1(t), \ldots, p_n(t)\big).$$

This expression is often used in physics for a gauge transformation of the Lagrangian, as two Lagrangians that differ only by the total time derivative of a function of time and t generalized coordinates lead to the same equations of motion. The operator in brackets (in the final expression) is also called the total derivative operator (with respect to t).

For example, the total derivative of $f(x(t), y(t))$ is

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

Here there is no $\partial f / \partial t$ term since f itself does not depend on the independent variable t

directly

The total derivative via differentials

Differentials provide a simple way to understand the total derivative. For instance, suppose $M(t, p_1, \ldots, p_n)$ is a function of time t and n variables p_i as in the previous section. Then, the differential of M is

$$\mathrm{d}\,M = \frac{\partial M}{\partial t}\mathrm{d}\,t + \sum_{i=1}^{n} \frac{\partial M}{\partial p_i}\mathrm{d}\,p_i.$$

This expression is often interpreted *heuristically* as a relation between infinitesimals. However, if the variables t and p_j are interpreted as functions, and $M(t, p_1, \ldots, p_n)$ is interpreted to mean the composite of M with these functions, then the above expression makes perfect sense as an equality of differential 1-forms, and is immediate from the chain rule for the exterior derivative. The advantage of this point of view is that it takes into account arbitrary dependencies between the variables. For example, if $p_1^2 = p_2 p_3$ then $2p_1\,\mathrm{d}\,p_1 = p_3\,\mathrm{d}\,p_2 + p_2\,\mathrm{d}\,p_3$. In particular, if the variables p_j are all functions of t, as in the previous section, then

$$\mathrm{d}\,M = \frac{\partial M}{\partial t}\mathrm{d}\,t + \sum_{i=1}^{n} \frac{\partial M}{\partial p_i}\frac{\partial p_i}{\partial t}\;\mathrm{d}\,t.$$

The total derivative as a linear map

Let $U \subseteq \mathbb{R}^n$ be an open subset. Then a function $f : U \to \mathbb{R}^m$ is said to be (**totally**) **differentiable** at a point $p \in U$, if there exists a linear map $\mathrm{d} f_p : \mathbb{R}^n \to \mathbb{R}^m$ (also denoted $D_p f$ or $Df(p)$) such that

$$\lim_{x \to p} \frac{\|f(x) - f(p) - \mathrm{d} f_p(x-p)\|}{\|x-p\|} = 0.$$

The linear map $\mathrm{d} f_p$ is called the (**total**) **derivative** or (**total**) **differential** of f at p. A function is (**totally**) **differentiable** if its total derivative exists at every point in its domain.

Note that f is differentiable if and only if each of its components $f_i : U \to \mathbb{R}$ is differentiable. For this it is necessary, but not sufficient, that the partial derivatives of each function f_j exist. However, if these partial derivatives exist and are continuous, then f is differentiable and its differential at any point is the linear map determined by the Jacobian matrix of partial derivatives at that point.

Total differential equation

A *total differential equation* is a differential equation expressed in terms of total derivatives. Since the exterior derivative is a natural operator, in a sense that can be given a technical meaning, such equations are intrinsic and *geometric*.

CONCLUSION

In this unit, you have known how to differentiate with indirect dependent. You have used total derivative via differentials and have known the total derivative as a linear map. You have

SUMMARY

In this unit, you have studied the following :

Differentiation with indirect dependent

The total derivative via differentials

The total derivative as a linear map

The total differential equation

TUTOR – MARK ASSIGNMENT

1.Find the total derivative for the second – order of the function

F(x,y,z)= $x^3 + y^4 - z^3$

2.Find the total derivative for the function

F(x,y,z)= $x^2 y^3 + z^3$

3.Solve the total derivative to the third - order of the function

F(x,y,z)= $x^3 y^4 + x^2 y + y^3 x^4 z^4$

REFFERENCE

A. D. Polyanin and V. F. Zaitsev, *Handbook of Exact Solutions for Ordinary Differential Equations (2nd edition)*, Chapman & Hall/CRC Press, Boca Raton, 2003. ISBN 1-58488-297-2

nton, Howard; Bivens, Irl; Davis, Stephen (February 2, 2005), *Calculus: Early Transcendentals Single and Multivariable* (8th ed.), New York: Wiley, ISBN 978-0-471-47244-5

Apostol, Tom M. (June 1967), *Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00005-1

Apostol, Tom M. (June 1969), *Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications*, **1** (2nd ed.), Wiley, ISBN 978-0-471-00007-5

Courant, Richard; John, Fritz (December 22, 1998), *Introduction to Calculus and Analysis, Vol. 1*, Springer-Verlag, ISBN 978-3-540-65058-4

Eves, Howard (January 2, 1990), *An Introduction to the History of Mathematics* (6th ed.), Brooks Cole, ISBN 978-0-03-029558-4

Larson, Ron; Hostetler, Robert P.; Edwards, Bruce H. (February 28, 2006), *Calculus: Early Transcendental Functions* (4th ed.), Houghton Mifflin Company, ISBN 978-0-618-60624-5

Spivak, Michael (September 1994), *Calculus* (3rd ed.), Publish or Perish, ISBN 978-0-914098-89-8

Stewart, James (December 24, 2002), *Calculus* (5th ed.), Brooks Cole, ISBN 978-0-534-39339-7

Thompson, Silvanus P. (September 8, 1998), *Calculus Made Easy* (Revised, Updated, Expanded ed.), New York: St. Martin's Press, ISBN 978-0-312-18548-0

UNIT 3: APPLICATION OF TOTAL DERIVATIVE OF A FUNCTION.

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT
3.1 chain rule
3.2 directional derivative
3.3 differentiation under integral sign
3.4 lebnitz rule
4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

Let us consider a function

1) $u = f(x, y, z, p, q, \ldots)$

of several variables. Such a function can be studied by holding all variables except one constant and observing its variation with respect to one single selected variable. If we consider all the variables except x to be constant, then

$$\frac{du}{dx} = \frac{d\, f(x, \hat{y}, \hat{z}, \hat{p}, \hat{q}, \ldots)}{dx}$$

represents the partial derivative of f(x, y, z, p, q, ...) with respect to x (the hats indicating variables held fixed). The variables held fixed are viewed as parameters.

OBJECTIVES

After studying this unit, you should be able to correctly :

apply total derivative on chain rule for functions of functions

apply total derivative to find directional derivative
apply total derivative to solve differentiation under integral sign
apply total derivative on lebnitz rule

APPLICATION OF TOTAL DERIVATIVES.

Chain rule for functions of functions.

If w = f(x, y, z, ...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives $\partial w / \partial x, \partial w / \partial y, \partial w / \partial z$, ... and if x, y, z, ... are differentiable functions x = x(t), y = y(t) , z = z(t), etc. of a variable t, then the **total derivative** of w with respect to t is given by

2) $$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} + \cdots$$

This rule is called the **chain rule** for the partial derivatives of functions of functions.

Similarly, if w = f(x, y, z, ...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives $\partial w / \partial x, \partial w / \partial y, \partial w / \partial z$, ... and if x, y, z, ... are differentiable functions of m independent variables r, s, t ... , then

$$\frac{dw}{dr} = \frac{\partial w}{\partial x}\frac{dx}{dr} + \frac{\partial w}{\partial y}\frac{dy}{dr} + \frac{\partial w}{\partial z}\frac{dz}{dr} + \cdots$$

$$\frac{dw}{ds} = \frac{\partial w}{\partial x}\frac{dx}{ds} + \frac{\partial w}{\partial y}\frac{dy}{ds} + \frac{\partial w}{\partial z}\frac{dz}{ds} + \cdots \qquad etc.$$

This rule is called the **chain rule** for the partial derivatives of functions of functions.

Similarly, if w = f(x, y, z, ...) is a continuous function of n variables x, y, z, ..., with continuous partial derivatives $\partial w / \partial x, \partial w / \partial y, \partial w / \partial z$, ... and if x, y, z, ... are differentiable functions of m independent variables r, s, t ... , then

$$\frac{dw}{dr} = \frac{\partial w}{\partial x}\frac{dx}{dr} + \frac{\partial w}{\partial y}\frac{dy}{dr} + \frac{\partial w}{\partial z}\frac{dz}{dr} + \cdots$$

$$\frac{dw}{ds} = \frac{\partial w}{\partial x}\frac{dx}{ds} + \frac{\partial w}{\partial y}\frac{dy}{ds} + \frac{\partial w}{\partial z}\frac{dz}{ds} + \cdots \qquad etc.$$

Note the similarity between total differentials and total derivatives. The total derivative above can be obtained by dividing the total differential

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz + \cdots$$ by dt.

As a special application of the chain rule let us consider the relation defined by the two equations

z = f(x, y); y = g(x)

Here, z is a function of x and y while y in turn is a function of x. Thus z is really a function of the single variable x. If we apply the chain rule we get

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}$$

which is the total derivative of z with respect to x.

Defination of Scalar point function. A scalar point function is a function that assigns a real number (i.e. a scalar) to each point of some region of space. If to each point (x, y, z) of a region R in space there is assigned a real number $u = \Phi(x, y, z)$, then Φ is called a scalar point function.

Examples. 1. The temperature distribution within some body at a particular point in time. 2. The density distribution within some fluid at a particular point in time.

Directional derivatives. Let $\Phi(x, y, z)$ be a scalar point function defined over some region R of space. The function $\Phi(x, y, z)$ could, for example, represent the temperature distribution within some body. At some specified point P(x, y, z) of R we wish to know the rate of change of Φ in a particular direction. The rate of change of a function Φ at a particular point P, in a specified direction, is called the **directional derivative** of Φ at P in that direction. We specify the direction by supplying the direction angles or direction cosines of a unit vector **e** pointing in the desired direction.

Theorem. The rate of change of a function $\Phi(x, y, z)$ in the direction of a vector with direction angles (α, β, γ) is given by

3) $$\frac{d\Phi}{ds} = \frac{\partial \Phi}{\partial x}\cos\alpha + \frac{\partial \Phi}{\partial y}\cos\beta + \frac{\partial \Phi}{\partial z}\cos\gamma$$

where s corresponds to distance in the metric of the coordinate system. That direction for which the function Φ at point P has its maximum value is called the **gradient** of Φ at P.

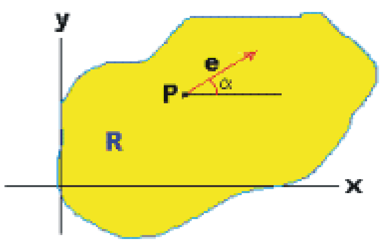


Fig. 4

We shall prove the theorem shortly. First let us consider the same problem for two dimensional space.

Let Φ(x, y) be a scalar point function defined over some region R of the plane. At some specified point P(x, y) of R we wish to know the rate of change of Φ in a particular direction. We specify the direction by supplying the angle α that a unit vector **e** pointing in the desired direction makes with the positive x direction. See Fig. 4. The rate of change of function Φ at point P in the direction of **e** corresponding to angle α is given by

$$4)\quad \frac{d\Phi}{ds} = \frac{\partial \Phi}{\partial x}\cos\alpha + \frac{\partial \Phi}{\partial y}\sin\alpha$$

where s corresponds to distance in the metric of the coordinate system. We show this as follows:

Let

T = f(x, y)

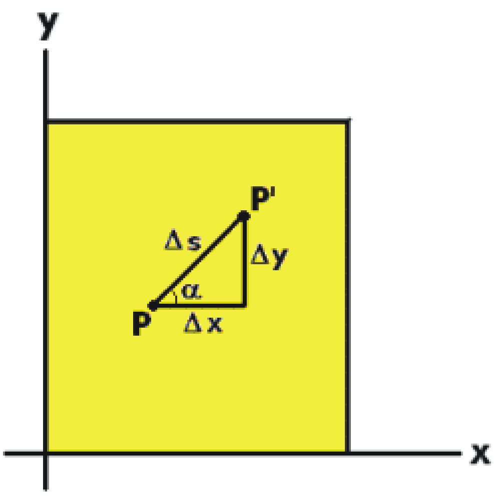


Fig. 5

where T is the temperature at any point of the plate shown in Fig. 5. We wish to derive expression 4) above. In other words, we wish to derive the expression for the rate of change of T with respect to the distance moved in any selected direction. Suppose we move from point P to point P'. This represents a displacement Δx in the x-direction and Δy in the y-direction. The distance moved along the plate is

$$PP' = \Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2}\ .$$

The direction is given by the angle α that PP' makes with the positive x-direction. The change in the value of T corresponding to the displacement from P to P' is

$$\Delta T = \frac{\partial T}{\partial x}\Delta x + \frac{\partial T}{\partial y}\Delta y + \varepsilon\sqrt{(\Delta x)^2 + (\Delta y)^2}\ .$$

where ε is a quantity that approaches 0 when Δx and Δy approach 0.

If we divide ΔT by the distance moved along the plate, we have

$$\frac{\Delta T}{\Delta s} = \frac{\partial T}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial T}{\partial y}\frac{\Delta y}{\Delta s} + \varepsilon\sqrt{\left(\frac{\Delta x}{\Delta s}\right)^2 + \left(\frac{\Delta y}{\Delta s}\right)^2}\ .$$

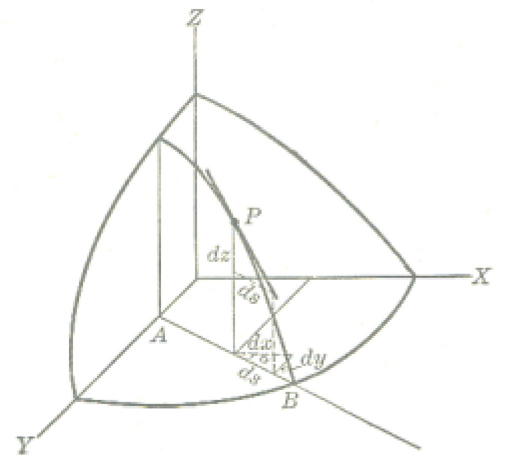


Fig. 6

From Fig. 5 we observe that Δx/Δs = cos α and Δy/Δs = sin α . Making these substitutions and letting P' approach P along line PP', we have

$$\frac{dT}{ds} = \frac{\partial T}{\partial x}\cos\alpha + \frac{\partial T}{\partial y}\sin\alpha$$

This is the directional derivative of T in the direction α.

A geometric interpretation of a directional derivative in the case of a function z = f(x, y) is that of a tangent to the surface at point P as shown in Fig. 6.

Def. Directional derivative. The directional derivative of a scalar point function Φ(x, y, z) is the rate of change of the function Φ(x, y, z) at a particular point P(x, y, z) as measured in a specified direction.

Tech. Let Φ(x, y, z) be a scalar point function possessing first partial derivatives throughout some region R of space. Let P(x_0, y_0, z_0) be some point in R at which we wish to compute the directional derivative and let P'(x_1, y_1, z_1) be a neighboring point. Let the distance from P to P' be Δs. Then the directional derivative of Φ in the direction PP' is given by

5) $$\frac{d\Phi}{ds} = \lim_{P'\to P}\frac{\Phi[P'(x_1, y_1, z_1)] - \Phi[P(x_0 y_0, z_0)]}{\Delta s}$$

where P' approaches P along the line PP' and Δs approaches 0.

Using this definition, let us now derive 3) above. In moving from P to P' the function Φ will change by an amount

$$\Delta\Phi = \frac{\partial\Phi}{\partial x}\Delta x + \frac{\partial\Phi}{\partial y}\Delta y + \frac{\partial\Phi}{\partial z}\Delta z + \varepsilon_1\Delta x + \varepsilon_2\Delta y + \varepsilon_3\Delta z$$

where ε_1, ε_2, ε_3 are higher order infinitesimals which approach zero as P' approaches P i.e. as Δx, Δy and Δz approach zero. If we divide the change ΔΦ by the distance Δs we obtain a measure of the rate at which Φ changes as we move from P to P':

$$\frac{\Delta\Phi}{\Delta s} = \frac{\partial\Phi}{\partial x}\frac{\Delta x}{\Delta s} + \frac{\partial\Phi}{\partial y}\frac{\Delta y}{\Delta s} + \frac{\partial\Phi}{\partial z}\frac{\Delta z}{\Delta s} + \varepsilon_1\frac{\Delta x}{\Delta s} + \varepsilon_2\frac{\Delta y}{\Delta s} + \varepsilon_3\frac{\Delta z}{\Delta s} \tag{6)}$$

We now observe that Δx/Δs, Δy/Δs, Δz/Δs are the direction cosines of the line segment PP'. They are also the direction cosines of a unit vector **e** located at P pointing in the direction of '. If the direction angles of **e** are α, β, γ, then Δx/Δs, Δy/Δs, Δz/Δs are equal to cos α, cos β, and cos γ, respectively. Thus 6) becomes

$$\frac{\Delta\Phi}{\Delta s} = \frac{\partial\Phi}{\partial x}\cos\alpha + \frac{\partial\Phi}{\partial y}\cos\beta + \frac{\partial\Phi}{\partial z}\cos\gamma + \varepsilon_1\frac{\Delta x}{\Delta s} + \varepsilon_2\frac{\Delta y}{\Delta s} + \varepsilon_3\frac{\Delta z}{\Delta s}$$

and

$$\frac{d\Phi}{ds} = \lim_{\Delta s\to 0}\frac{\Delta\Phi}{\Delta s} = \frac{\partial\Phi}{\partial x}\cos\alpha + \frac{\partial\Phi}{\partial y}\cos\beta + \frac{\partial\Phi}{\partial z}\cos\gamma \tag{7)}$$

Let us note that 7) can be written in vector form as the following dot product:

$$\frac{d\Phi}{ds} = \begin{bmatrix}\frac{\partial\Phi}{\partial x} & \frac{\partial\Phi}{\partial y} & \frac{\partial\Phi}{\partial z}\end{bmatrix} \cdot \begin{bmatrix}\cos\alpha & \cos\beta & \cos\gamma\end{bmatrix} = \begin{bmatrix}\frac{\partial\Phi}{\partial x} & \frac{\partial\Phi}{\partial y} & \frac{\partial\Phi}{\partial z}\end{bmatrix} \cdot e \tag{8)}$$

The vector

$$\begin{bmatrix}\frac{\partial\Phi}{\partial x} & \frac{\partial\Phi}{\partial y} & \frac{\partial\Phi}{\partial z}\end{bmatrix}$$

is called the gradient of Φ. Thus the directional derivative of Φ is equal to the dot product of the gradient of Φ and the vector **e**. In other words,

$$\left(\frac{d\Phi}{ds}\right)_e = grad\,\Phi\cdot e$$

where

$$\left(\frac{d\Phi}{ds}\right)_e$$

is the directional derivative of Φ in the direction of unit vector e.

If the vector **e** is pointed in the same direction as the gradient of Φ then the directional derivative of Φ is equal to the gradient of Φ.

Differentiation under the integral sign. Leibnitz's rule. We now consider differentiation with respect to a parameter that occurs under an integral sign, or in the limits of integration, or in both places.

Theorem 1. Let

$$F(x) = \int_a^x f(t)dt$$

where a $\leq$x $\leq$b and f is assumed to be integrable on [a, b]. Then the function *F*(x) is continuous and $F'(x)$ = f(x) at each point where f(x) is continuous.

Theorem 2. Let f(x, α) and ∂f/∂α be continuous in some region **R**: (a $\leq$x $\leq$b, c $\leq$α $\leq$d) of the x-α plane. Let

$$9) \quad G(\alpha) = \int_a^b f(x,\alpha)\,dx \qquad c \leq \alpha \leq d$$

Then

$$10) \quad \frac{dG}{d\alpha} = \int_a^b \frac{\partial f(x,\alpha)}{\partial \alpha}\, dx$$

Theorem 3. Leibnitz's rule. Let

$$11) \quad G(\alpha) = \int_{u_1}^{u_2} f(x,\alpha)\, dx \qquad c \le \alpha \le d$$

where u_1 and u_2 are functions of the parameter α i.e.

$u_1 = u_1(\alpha)$

$u_2 = u_2(\alpha)$.

Let f(x, α) and ∂f/∂α be continuous in both x and α in a region **R** of the x-α plane that includes the region $u_1 \le x \le u_2$, $c \le \alpha \le d$. Let u_1 and u_2 be continuous and have continuous derivatives for $c \le \alpha \le d$. Then

$$12) \quad \frac{dG}{d\alpha} = \int_{u_1}^{u_2} \frac{\partial f(x,\alpha)}{\partial \alpha}\, dx + f(u_2,\alpha)\frac{du_2}{d\alpha} - f(u_1,\alpha)\frac{du_1}{d\alpha}$$

where $f(u_1, \alpha)$ is the expression obtained by substituting the expression $u_1(\alpha)$ for x in f(x, α). Similarly for $f(u_2, \alpha)$. The quantities $f(u_1, \alpha)$ and $f(u_2, \alpha)$ correspond to $\partial G/u_1$ and $\partial G/u_2$ respectively and 12) represents the chain rule.

Order of differentiation. For most functions that one meets

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

However, in some cases it is not true. Under what circumstances is it true? It is true if both functions f_{yx} and f_{xy} are continuous at the point where the partials are being taken.

Theorem. Let the function f(x, y) be defined in some neighborhood of the point (a, b). Let the partial derivatives f_x, f_y, f_{xy}, and f_{yx} also be defined in this neighborhood. Then if f_{xy} and f_{yx} are both continuous at (a, b), $f_{xy}(a, b) = f_{yx}(a, b)$.

EXAMPLE

Given $u = x^2 + 2y$ where $x = r\sin(t)$ and $y = \sin^2(t)$, determine the value of $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial t}$ using the chain rule.

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = (2x)(\sin(t)) + (2)(0) = 2r\sin^2(t)$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} = (2x)(r\cos(t)) + (2)(2\sin(t)\cos(t))$$
$$= 2(r\sin(t))r\cos(t) + 4\sin(t)\cos(t) = 2\left(r^2+2\right)\sin(t)\cos(t).$$

4.0 CONCLUSION

In this unit, you have applied total derivative on chain rule. You have solved problems on directional derivatives using total derivative. You have used total derivative to solve differentiation under integral sign and leibnitz rule.

5,0 SUMMARY

In this unit, you have studied the following:

The application of total derivative on chain rule

The application of total derivative on directional derivative

The application of total derivative on differentiation under integral sign

The application of total derivative on leibnitz rule.

6.0 TUTOR - MARKED ASSIGNMENT

1.Find all directional derivatives of the function

$$F(x,y) = \left(3x^2+y^4\right)^{\frac{1}{4}}$$

(where $(x;y)\in\mathbb{R}^2$), in the point $(0;0)$

2. Find the integral of the function

F(x,y,z) = $3x^2$ + 2xyz

In the point (0,1)

3.Find the total derivative of the function

F(xy) = 3xy + 4 y^2

REFERENCES

James & James. Mathematics Dictionary

Middlemiss. Differential and Integral Calculus

Spiegel. Advanced Calculus

Taylor. Advanced Calculus

Spiegel. Vector Analysis

Bernard R. Gelbaum and John M. H. Olmsted. Counterexamples in Analysis.

Dover, 2003.

MODULE 4

PARTIAL DIFFERENTIABILITY AND TOTAL DIFFERENTIABILITY OF FUNCTION OF SEVERAL VARIABLE

-Unit 1: Partial differentials of function of several variables.

-Unit 2: Total differentials of function of several variables.

-Unit 3: Application of partial and total differentials of function of several variables.

UNIT 1 PARTIAL DIFFERENTIABILITY OF FUNCTION OF SEVERAL VARIABLE

CONTENT

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT
3.1 Partial derivatives
3.2 Second partial derivatives

4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

Differentiation is a method to compute the rate at which a dependent output *y* changes with respect to the change in the independent input *x*. This rate of change is called the **derivative** of *y* with respect to *x*. In more precise language, the dependence of *y* upon *x* means that *y* is a function of *x*. This functional relationship is often denoted $y = f(x)$, where f denotes the function. If *x* and *y* are real numbers, and if the graph of *y* is plotted against *x*, the derivative measures the slope of this graph at each point.

The simplest case is when *y* is a linear function of *x*, meaning that the graph of *y* against *x* is a straight line. In this case, $y = f(x) = m\,x + b$, for real numbers *m* and *b*, and the slope *m* is given by

$$m = \frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x}$$

where the symbol Δ (the uppercase form of the Greek letter Delta) is an abbreviation for "change in." This formula is true because

$y + \Delta y = f(x+ \Delta x) = m\,(x + \Delta x) + b = m\,x + b + m\,\Delta x = y + m\Delta x.$

It follows that $\Delta y = m\,\Delta x$.

This gives an exact value for the slope of a straight line. If the function f is not linear (i.e. its graph is not a straight line), however, then the change in y divided by the change in x varies: differentiation is a method to find an exact value for this rate of change at any given value of x.

Rate of change as a limiting value

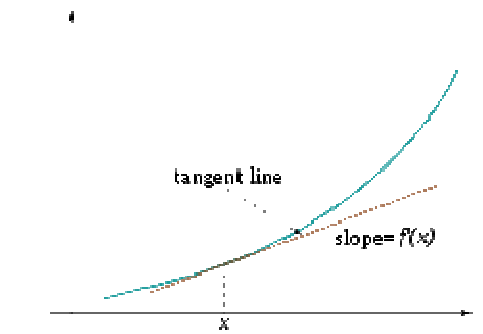


Figure 1. The tangent line at $(x, f(x))$

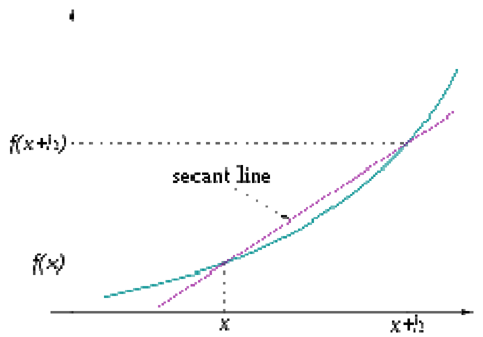


Figure 2. The secant to curve $y= f(x)$ determined by points $(x, f(x))$ and $(x+h, f(x+h))$

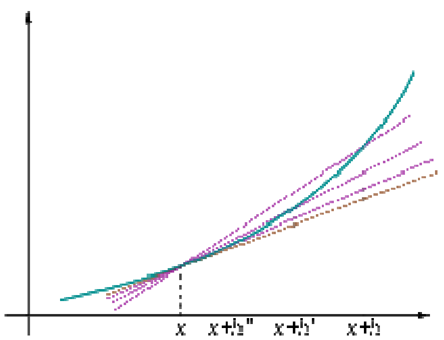


Figure 3. The tangent line as limit of secants

The idea, illustrated by Figures 1-3, is to compute the rate of change as the limiting value of the ratio of the differences $\Delta y / \Delta x$ as Δx becomes infinitely small.

In Leibniz's notation, such an infinitesimal change in x is denoted by dx, and the derivative of y with respect to x is written

$$\frac{dy}{dx}$$

suggesting the ratio of two infinitesimal quantities. (The above expression is read as "the derivative of y with respect to x", "d y by d x", or "d y over d x". The oral form "d y d x" is often used conversationally, although it may lead to confusion.)

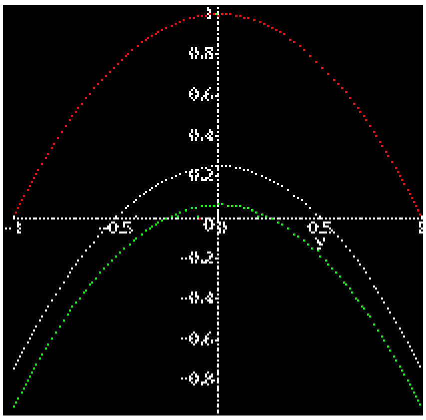
The most common approach[2] to turn this intuitive idea into a precise definition uses limits, but there are other methods, such as non-standard analysis.[3]

Derivatives

Bound as we humans are to three spacial dimensions, multi-variable functions can be very difficult to get a good feel for. (Try picturing a function in the 17th dimension and see how far you get!) We can at least make three-dimensional models of two-variable functions, but even then at a stretch to our intuition. What is needed is a way to cheat and look at multi-variable functions as if they were one-variable functions.

We can do this by using **partial functions**. A partial function is a one-variable function obtained from a function of several variables by assigning constant values to all but one of the independent variables. What we are doing is taking two-dimensional "slices" of the surface represented by the equation.

For Example: $z=x^2-y^2$ can be modeled in three dimensional space, but personally I find it difficult to sketch! In the section on critical points a picture of a plot of this function can be found as an example of a saddle point. But by alternately setting x=1 (red), x=0.5 (white), and x=0.25 (green), we can take slices of $z=x^2-y^2$ (each one a plane parallel to the z-y plane) and see different partial functions. We can get a further idea of the behavior of the function by considering that the same curves are obtained for x=-1, -0.5 and -0.25.



Food For Thought: How do partial functions compare to level curves and level surfaces? If the function f is a continuous function, does the level set or surface have to be continuous? What about partial functions?

All of this helps us to get to our main topic, that is, partial differentiation. We know how to take the derivative of a single-variable function. What about the derivative of a multi-variable function? What does that even mean? Partial Derivatives are the beginning of an answer to that question.

OBJECTIVES

In this unit, you should be able to :

Identify and solve partial derivatives

Solve second partial derivatives.

MAIN CONTENT

A **partial derivative** is the rate of change of a multi-variable function when we allow only one of the variables to change. Specifically, we differentiate with respect to only one variable, regarding all others as constants (now we see the relation to partial functions!). Which essentially means if you know how to take a derivative, you know how to take a partial derivative.

A partial derivative of a function *f* with respect to a variable *x*, say $z=f(x,y_1,y_2,...y_n)$ (where the y_i's are other independent variables) is commonly denoted in the following ways:

$$\frac{\partial z}{\partial x}$$ (*referred to as ``partial z, partial x''*)

$$\frac{\partial f}{\partial x}$$ (*referred to as ``partial f, partial x''*)

Note that this is not the usual derivative ``*d*'' . The funny ``d'' symbol in the notation is called ``roundback d'', ``curly d'' or ``del d'' (to distinguish from ``delta d''; the symbol is actually a ``lowercase Greek `delta' '').

The next set of notations for partial derivatives is much more compact and especially used when you are writing down something that uses lots of partial derivatives, especially if they are all different kinds:

z_x (*referred to as ``partial z, partial x''*)

f_x (*referred to as ``partial f, partial x''*)

$f_x(x,y)$ *(referred to as ``partial f, partial x'')*

Any of the above is equivalent to the limit

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

To get an intuitive grasp of partial derivatives, suppose you were an ant crawling over some rugged terrain (a two-variable function) where the x-axis is north-south with positive x to the north, the y-axis is east-west and the z-axis is up-down. You stop at a point $P=(x_0, y_0, z_0)$ on a hill and wonder what sort of slope you will encounter if you walk in a straight line north. Since our longitude won't be changing as we go north, the y in our function is constant. The slope to the north is the value of $f_x(x_0, y_0)$.

The actual calculations of partial derivatives for most functions is very easy! Treat every indpendent variable except the one we are interested in as if it were a constant and apply the familiar rules!

Example:

Let's find f_x and f_y of the function $z=f=x^2 -3x^2y+y^3$. To find f_x, we will treat y as a constant and differentiate. So, $f_x=2x-6xy$. By treating x as a constant, we find $f_y=-3x^2+3y^2$.

Second Partial Derivatives

Observe carefully that the expression f_{xy} implies that the function f is differentiated first with respect to x and then with respect to y, which is a natural inference since f_{xy} is really $(f_x)_y$.

For the same reasons, in the case of the expression,

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

it is implied that we differentiate first with respect to y and then with respect to x.

Below are examples of **pure second partial derivatives**:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} \quad \frac{\partial^2 f}{\partial y^2} = f_{yy}$$

Example:

Lets find f_{xy} and f_{yx} of $f=e^{xy} + y(\sin x)$.

- $f_x=ye^{xy} + y\cos x$

- $f_{xy}=xye^{xy} + \cos x$
- $f_y=xe^{xy} + \sin x$
- $f_{yx}=xye^{xy} + \cos x$

In this example $f_{xy}=f_{yx}$. Is this true in general? Most of the time and in most examples that you will probably ever see, yes. More precisely, if

- both f_{xy} and f_{yx} exist for all points near (x_0,y_0)
- and are continuous at (x_0,y_0),

then $f_{xy}=f_{yx}$.

Partial Derivatives of higher order are defined in the obvious way. And as long as suitable continuity exists, it is immaterial in what order a sequence of partial differentiation is carried out.

total differential (Definition)

There is the generalisation of the theorem in the parent entry concerning the real functions of several variables; here we formulate it for three variables:

Theorem. Suppose that S is a ball in $\mathbb{R}^3$, the function $f:S \longrightarrow \mathbb{R}$ is continuous and has partial derivatives f_x', f_y', f_z' in S and the partial derivatives are continuous in a point (x, y, z) of S . Then the increment

$\Delta f:=f(x+\Delta x, y+\Delta y, z+\Delta z)-f(x, y, z),$

which f gets when one moves from (x, y, z) to another point $(x+\Delta x, y+\Delta y, z+\Delta z)$ of S , can be split into two parts as follows:

$$\Delta f = [f_x'(x, y, z)\Delta x + f_y'(x, y, z)\Delta y + f_z'(x, y, z)\Delta z] + \langle \varrho \rangle \varrho. \qquad (1)$$

Here, $\varrho:=\sqrt{\Delta x^2+\Delta y^2+\Delta z^2}$ and $\langle \varrho \rangle$ is a quantity tending to 0 along with ϱ.

The former part of Δx is called the (total) differential or the exact differential of the function f in the point (x, y, z) and it is denoted by $df(x, y, z)$ of briefly df . In the special case $f(x, y, z) \equiv x$, we see that $df=\Delta x$ and thus $\Delta x=dx$; similarly $\Delta y=dy$ and $\Delta z=dz$. Accordingly, we obtain for the general case the more consistent notation

$$df = f_x'(x, y, z)dx + f_y'(x, y, z)dy + f_z'(x, y, z)dz, \qquad (2)$$

where dx, dy, dz may be thought as independent variables.

We now assume conversely that the increment of a function f in $\mathbb{R}^3$ can be split into two parts as follows:

$$f(x+\Delta x,\ y+\Delta y,\ z+\Delta z) - f(x,\ y,\ z) \ = \ [A\Delta x + B\Delta y + C\Delta z] + \langle \varrho \rangle \varrho \quad (3)$$

where the coefficients A ,B ,C are independent on the quantities Δx , Δy , Δz and ϱ , $\langle \varrho \rangle$ are as in the above theorem. Then one can infer that the partial derivatives fx′ ,fy′ ,fz′ exist in the point (x ,y ,z) and have the values A ,B ,C , respectively. In fact, if we choose $\Delta y = \Delta z = 0$, then $\varrho = |\Delta x|$ whence (3) attains the form

$f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y, z) = A\Delta x + \langle \Delta x \rangle \Delta x$

and therefore

$A = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y, z)}{\Delta x} = f_x'(x, y, z)$.

Similarly we see the values of fy′ and fz′ .

The last consideration showed the uniqueness of the total differential.

Definition. A function f in $\mathbb{R}^3$, satisfying the conditions of the above theorem is said to be differentiable in the point (x ,y ,z) .

CONCLUSION

In this unit, you have identified and solved problem on partial differential of function of several variables. You have also used partial differential of function of several variables to solve problems on second partial derivatives.

SUMMARY

In this unit, you have studied :

Partial derivatives

Second partial derivatives

TUTOR – MARK ASSIGNMENTS

1.Find the first order derivative of the following function

F(x,y,z) = $x^2 y z^4$

2.Find $f_{xx} f_{yy} f_{zz}$,given that F(x,y,z) = sin(xyz)

3.Evaluate the second order derivative of

$f_{xx} f_{yy} f_{zz} = x^3 y^4 + 2xy + z^4$

4.Evaluate the second order derivative of

F(x,y,z) = $x^3 + y^2 z^3$

REFFERENCE

Jacques, I. 1999. *Mathematics for Economics and Business.* 3*rd* Edition. Prentice Hall.

UNIT 2 TOTAL DIFFERENTIABILITY OF FUNCTION OF SEVERAL VARIABLE

CONTENT

1.0 **INTRODUCTION**
2.0 OBJECTIVES
3.0 MAIN CONTENT

Identify and solve problems on total differentials of functions of several variables

4.0 Conclusion
5.0- Summary
6.0 Tutor-Marked Assignment
7.0 References/Further Readings

INTRODUCTION

In the case of a function of a single variable the **differential** of the function y = f(x) is the quantity

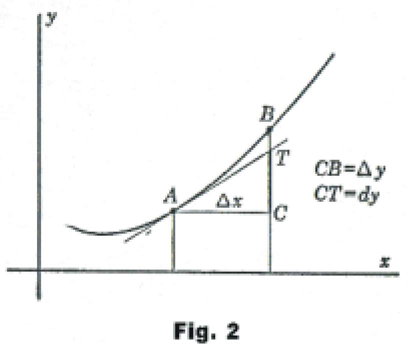


Fig. 2

dy = f '(x) Δx .

This quantity is used to compute the approximate change in the value of f(x) due to a change

Δx in x. As is shown in Fig. 2,

$$\Delta y = CB = f(x + \Delta x) - f(x)$$

while $dy = CT = f'(x)\Delta x$.

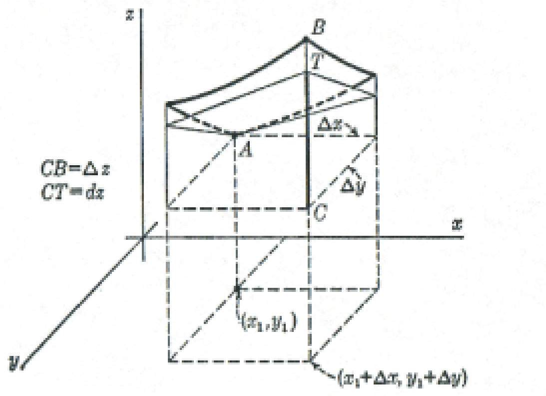


Fig. 3

When Δx is small the approximation is close. Line AT represents the tangent to the curve at point A.

OBJECTIVES

At the end of this unit, you should be able to identify and solve problems on total differentials of functions of several variables

MAIN CONTENT

In the case of a function of two variables the situation is analogous. Let us start at point $A(x_1, y_1, z_1)$ on the surface

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

shown in Fig. 3 and let x and y change by small amounts Δx and Δy, respectively. The change produced in the value of the function z is

$$\Delta z = CB = f(x_1 + \Delta x, y_1 + \Delta y) - f(x_1, y_1) .$$

An approximation to Δz is given by

$$CT = \left(\frac{\partial z}{\partial x}\right)_A \Delta x + \left(\frac{\partial z}{\partial y}\right)_A \Delta y .$$

When Δx and Δy are small the approximation is close. Point T lies in that plane tangent to the surface at point A.

The quantity

$$dz = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$$

is called the **total differential** of the function z = f(x, y). Because is customary to denote increments Δx and Δy by dx and dy, the total differential of a function z = f(x, y) is defined as

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\ .$$

The total differential of three or more variables is defined similarly. **For a function z = f(x, y, .. , u) the total differential is defined as**

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy + \cdots + \frac{\partial z}{\partial u}du\ .$$

Each of the terms represents a **partial differential**. For example, the term

$$\frac{\partial z}{\partial x}dx$$

is the partial differential of z with respect to x. The total differential is the sum of the partial differentials.

4.0 CONCLUSION

In this unit, you have identified and solved problems on total differentials of functions of several variables

5.0 SUMMARY

In this unit, you have studied total differentials of functions of several variables.

6.0TUTOR – MARKED ASSIGNMENT

Find the total differentiability of the following :

a. F(x,y) = x + 2xy + y^2

b. F(x,y,z) = $x^4 + 2y^3 + z^2$

c. F(x,y,z) = $x^3 y^2 z^3$

d. F(x,y,z) = $4x^2 y^3 + z^2$

e. F(x,y,z) = $\sqrt{x^2 + y^3 - 2xyz}$

7.0 REFERENCES

James & James. Mathematics Dictionary

Middlemiss. Differential and Integral Calculus

Spiegel. Advanced Calculus

Taylor. Advanced Calculus

MODULE 5 COMPOSITE DIFFERENTIATION, EULER'S THEOREM, IMPLICIT DIFFERENTIATION.

Unit 1: Composite differentiation

Unit 2: Euler's Theorem

Unit 3: Implicit differentiation.

UNIT 1: COMPOSITE DIFFERENTIATION

1.0 INTRODUCTION

2.0 OBJECTIVES

3.0 MAIN CONTENT

3.1 The chain rule

3.2 Composites of more than two functions

3.4 The quotient rule

3.5 Higher derivative

3.6 Proof of the chain rule

3.7 The rule in higher dimension

4.0 **CONCLUSION**

5.0 SUMMARY

6.0 TUTOR-MARKED ASSIGNMENT

7.0 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

In calculus, the chain rule is a formula for computing the derivative of the composition of two or more functions. That is, if f is a function and g is a function, then the chain rule expresses the derivative of the composite function $f \circ g$ in terms of the derivatives of f and g.

Calculate the derivatives of each function. Write in fraction form, if needed, so that all exponents are positive in your final answer. Use the "modified power rule" for each.

2.0 OBJECTIVES

At the end of this unit, you should be able to use chain rule to solve mathematical problems

solve composites of more than two functions

use the quotient rule to solve composite functions

identify problems in composite function which could be solve by the use of higher derivative.

Proof the chain rule

Know the rule in higher dimension

3.0 MAIN CONTENT

Statement of the Rule

The simplest form of the chain rule is for real-valued functions of one real variable. It says that if g is a function that is differentiable at a point c (i.e. the derivative g′(c) exists) and f is a function that is differentiable at g(c), then the composite function f ∘ g is differentiable at *c*, and the derivative is

$$(f \circ g)'(c) = f'(g(c)) \cdot g'(c).$$

The rule is sometimes abbreviated as

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

If *y* = *f*(*u*) and *u* = *g*(*x*), then this abbreviated form is written in Leibniz notation as:

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

The points where the derivatives are evaluated may also be stated explicitly:

$$\left.\frac{dy}{dx}\right|_{x=c} = \left.\frac{dy}{du}\right|_{u=g(c)} \cdot \left.\frac{du}{dx}\right|_{x=c}.$$

Further examples

The chain rule in the absence of formulas

It may be possible to apply the chain rule even when there are no formulas for the functions which are being differentiated. This can happen when the derivatives are measured directly. Suppose that a car is driving up a tall mountain. The car's speedometer measures its speed directly. If the grade is known, then the rate of ascent can be calculated using trigonometry. Suppose that the car is ascending at 2.5 km/h. Standard models for the Earth's atmosphere imply that the temperature drops about 6.5 °C per kilometer ascended (see lapse rate). To find the temperature drop per hour, we apply the chain rule. Let the function *g*(*t*) be the altitude of the car at time *t*, and let the function *f*(*h*) be the temperature *h* kilometers above sea level. *f* and *g* are not known exactly: For example, the altitude where the car starts is not known and the temperature on the mountain is not known. However, their derivatives are known: *f′* is −6.5 °C/km, and *g′* is 2.5 km/h. The chain rule says that the derivative of the composite function is the product of the derivative of *f* and the derivative of *g*. This is −6.5 °C/km · 2.5 km/h = −16.25 °C/h.

One of the reasons why this computation is possible is because *f′* is a constant function. This is because the above model is very simple. A more accurate description of how the temperature near the car varies over time would require an accurate model of how the temperature varies at different altitudes. This model may not have a constant derivative. To compute the temperature change in such a model, it would be necessary to know *g* and not just *g′*, because without knowing *g* it is not possible to know where to evaluate *f′*.

Composites of more than two functions

The chain rule can be applied to composites of more than two functions. To take the derivative of a composite of more than two functions, notice that the composite of *f*, *g*, and *h* (in that order) is the composite of *f* with *g* ∘ *h*. The chain rule says that to compute the derivative of *f* ∘ *g* ∘ *h*, it is sufficient to compute the derivative of *f* and the derivative of *g* ∘ *h*. The derivative of *f* can be calculated directly, and the derivative of *g* ∘ *h* can be calculated by applying the chain rule again.

For concreteness, consider the function

$$y = e^{\sin x^2}.$$

This can be decomposed as the composite of three functions:

$$\begin{aligned} y &= f(u) = e^u, \\ u &= g(v) = \sin v, \\ v &= h(x) = x^2. \end{aligned}$$

Their derivatives are:

$$\begin{aligned} \frac{dy}{du} &= f'(u) = e^u, \\ \frac{du}{dv} &= g'(v) = \cos v, \\ \frac{dv}{dx} &= h'(x) = 2x. \end{aligned}$$

The chain rule says that the derivative of their composite at the point *x* = *a* is:

$$(f \circ g \circ h)'(a) = f'((g \circ h)(a))(g \circ h)'(a) = f'((g \circ h)(a))g'(h(a))h'(a)$$

In Leibniz notation, this is:

$$\frac{dy}{dx} = \left.\frac{dy}{du}\right|_{u=g(h(a))} \cdot \left.\frac{du}{dv}\right|_{v=h(a)} \cdot \left.\frac{dv}{dx}\right|_{x=a},$$

or for short,

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

The derivative function is therefore:

$$\frac{dy}{dx} = e^{\sin x^2} \cdot \cos x^2 \cdot 2x.$$

Another way of computing this derivative is to view the composite function $f \circ g \circ h$ as the composite of $f \circ g$ and h. Applying the chain rule to this situation gives:

$$(f \circ g \circ h)'(a) = (f \circ g)'(h(a))h'(a) = f'(g(h(a))g'(h(a))h'(a).$$

This is the same as what was computed above. This should be expected because $(f \circ g) \circ h = f \circ (g \circ h)$.

The quotient rule

The chain rule can be used to derive some well-known differentiation rules. For example, the quotient rule is a consequence of the chain rule and the product rule. To see this, write the function $f(x)/g(x)$ as the product $f(x) \cdot 1/g(x)$. First apply the product rule:

$$\begin{aligned} \frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) &= \frac{d}{dx}\left(f(x) \cdot \frac{1}{g(x)}\right) \\ &= f'(x) \cdot \frac{1}{g(x)} + f(x) \cdot \frac{d}{dx}\left(\frac{1}{g(x)}\right). \end{aligned}$$

To compute the derivative of $1/g(x)$, notice that it is the composite of g with the reciprocal function, that is, the function that sends x to $1/x$. The derivative of the reciprocal function is $-1/x^2$. By applying the chain rule, the last expression becomes:

$$f'(x) \cdot \frac{1}{g(x)} + f(x) \cdot \left(-\frac{1}{g(x)^2} \cdot g'(x)\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2},$$

which is the usual formula for the quotient rule.

Derivatives of inverse functions

Suppose that $y = g(x)$ has an inverse function. Call its inverse function f so that we have x = f(y). There is a formula for the derivative of f in terms of the derivative of g. To see this, note that f and g satisfy the formula

$f(g(x)) = x$.

Because the functions $f(g(x))$ and x are equal, their derivatives must be equal. The derivative of *x* is the constant function with value 1, and the derivative of $f(g(x))$ is determined by the chain rule. Therefore we have:

$f'(g(x))g'(x) = 1.$

To express f' as a function of an independent variable *y*, we substitute $f(y)$ for *x* wherever it appears. Then we can solve for f'.

$$\begin{aligned} f'(g(f(y)))g'(f(y)) &= 1 \\ f'(y)g'(f(y)) &= 1 \\ f'(y) &= \frac{1}{g'(f(y))}. \end{aligned}$$

For example, consider the function $g(x) = e^x$. It has an inverse which is denoted $f(y) = \ln y$. Because $g'(x) = e^x$, the above formula says that

$$\frac{d}{dy}\ln y = \frac{1}{e^{\ln y}} = \frac{1}{y}.$$

This formula is true whenever *g* is differentiable and its inverse *f* is also differentiable. This formula can fail when one of these conditions is not true. For example, consider $g(x) = x^3$. Its inverse is $f(y) = y^{1/3}$, which is not differentiable at zero. If we attempt to use the above formula to compute the derivative of *f* at zero, then we must evaluate $1/g'(f(0))$. $f(0) = 0$ and $g'(0) = 0$, so we must evaluate 1/0, which is undefined. Therefore the formula fails in this case. This is not surprising because *f* is not differentiable at zero.

Higher derivatives

Faà di Bruno's formula generalizes the chain rule to higher derivatives. The first few derivatives are

$$\frac{d(f\circ g)}{dx} = \frac{df}{dg}\frac{dg}{dx}$$

$$\frac{d^2(f\circ g)}{dx^2} = \frac{d^2 f}{dg^2}\left(\frac{dg}{dx}\right)^2 + \frac{df}{dg}\frac{d^2 g}{dx^2}$$

$$\frac{d^3(f\circ g)}{dx^3} = \frac{d^3 f}{dg^3}\left(\frac{dg}{dx}\right)^3 + 3\frac{d^2 f}{dg^2}\frac{dg}{dx}\frac{d^2 g}{dx^2} + \frac{df}{dg}\frac{d^3 g}{dx^3}$$

$$\frac{d^4(f\circ g)}{dx^4} = \frac{d^4 f}{dg^4}\left(\frac{dg}{dx}\right)^4 + 6\frac{d^3 f}{dg^3}\left(\frac{dg}{dx}\right)^2\frac{d^2 g}{dx^2} + \frac{d^2 f}{dg^2}\left\{4\frac{dg}{dx}\frac{d^3 g}{dx^3} + 3\left(\frac{d^2 g}{dx^2}\right)\right.$$

Proofs of the chain rule

First proof

One proof of the chain rule begins with the definition of the derivative:

$$(f \circ g)'(a) = \lim_{x \to a} \frac{f(g(x)) - f(g(a))}{x - a}.$$

Assume for the moment that $g(x)$ does not equal $g(a)$ for any x near a. Then the previous expression is equal to the product of two factors:

$$\lim_{x \to a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \frac{g(x) - g(a)}{x - a}.$$

When g oscillates near a, then it might happen that no matter how close one gets to a, there is always an even closer x such that $g(x)$ equals $g(a)$. For example, this happens for $g(x) = x^2 \sin(1 / x)$ near the point $a = 0$. Whenever this happens, the above expression is undefined because it involves division by zero. To work around this, introduce a function Q as follows:

$$Q(y) = \begin{cases} \frac{f(y) - f(g(a))}{y - g(a)}, & y \neq g(a), \\ f'(g(a)), & y = g(a). \end{cases}$$

We will show that the difference quotient for $f \circ g$ is always equal to:

$$Q(g(x)) \cdot \frac{g(x) - g(a)}{x - a}.$$

Whenever $g(x)$ is not equal to $g(a)$, this is clear because the factors of $g(x) - g(a)$ cancel. When $g(x)$ equals $g(a)$, then the difference quotient for $f \circ g$ is zero because $f(g(x))$ equals $f(g(a))$, and the above product is zero because it equals $f'(g(a))$ times zero. So the above product is always equal to the difference quotient, and to show that the derivative of $f \circ g$ at a exists and to determine its value, we need only show that the limit as x goes to a of the above product exists and determine its value.

To do this, recall that the limit of a product exists if the limits of its factors exist. When this happens, the limit of the product of these two factors will equal the product of the limits of the factors. The two factors are $Q(g(x))$ and $(g(x) - g(a)) / (x - a)$. The latter is the difference quotient for g at a, and because g is differentiable at a by assumption, its limit as x tends to a exists and equals $g'(a)$.

It remains to study $Q(g(x))$. Q is defined wherever f is. Furthermore, because f is differentiable at $g(a)$ by assumption, Q is continuous at $g(a)$. g is continuous at a because it is differentiable at a, and therefore $Q \circ g$ is continuous at a. So its limit as x goes to a exists and equals $Q(g(a))$, which is $f'(g(a))$.

This shows that the limits of both factors exist and that they equal $f'(g(a))$ and $g'(a)$, respectively. Therefore the derivative of $f \circ g$ at a exists and equals $f'(g(a))g'(a)$.

Second proof

Another way of proving the chain rule is to measure the error in the linear approximation determined by the derivative. This proof has the advantage that it generalizes to several variables. It relies on the following equivalent definition of differentiability at a point: A function *g* is differentiable at *a* if there exists a real number $g'(a)$ and a function $\varepsilon(h)$ that tends to zero as *h* tends to zero, and furthermore

$$g(a+h) - g(a) = g'(a)h + \varepsilon(h)h.$$

Here the left-hand side represents the true difference between the value of *g* at *a* and at $a + h$, whereas the right-hand side represents the approximation determined by the derivative plus an error term.

In the situation of the chain rule, such a function ε exists because *g* is assumed to be differentiable at *a*. Again by assumption, a similar function also exists for *f* at $g(a)$. Calling this function η, we have

$$f(g(a)+k) - f(g(a)) = f'(g(a))k + \eta(k)k.$$

The above definition imposes no constraints on $\eta(0)$, even though it is assumed that $\eta(k)$ tends to zero as *k* tends to zero. If we set $\eta(0) = 0$, then η is continuous at 0.

Proving the theorem requires studying the difference $f(g(a + h)) - f(g(a))$ as *h* tends to zero. The first step is to substitute for $g(a + h)$ using the definition of differentiability of *g* at *a*:

$$f(g(a + h)) - f(g(a)) = f(g(a) + g'(a)h + \varepsilon(h)h) - f(g(a)).$$

The next step is to use the definition of differentiability of *f* at $g(a)$. This requires a term of the form $f(g(a) + k)$ for some *k*. In the above equation, the correct *k* varies with *h*. Set $k_h = g'(a)h + \varepsilon(h)h$ and the right hand side becomes $f(g(a) + k_h) - f(g(a))$. Applying the definition of the derivative gives:

$$f(g(a)+k_h) - f(g(a)) = f'(g(a))k_h + \eta(k_h)k_h.$$

To study the behavior of this expression as *h* tends to zero, expand k_h. After regrouping the terms, the right-hand side becomes:

$$f'(g(a))g'(a)h + [f'(g(a))\varepsilon(h) + \eta(k_h)g'(a) + \eta(k_h)\varepsilon(h)]h.$$

Because $\varepsilon(h)$ and $\eta(k_h)$ tend to zero as *h* tends to zero, the bracketed terms tend to zero as *h* tends to zero. Because the above expression is equal to the difference $f(g(a + h)) - f(g(a))$, by the definition of the derivative $f \circ g$ is differentiable at *a* and its derivative is $f'(g(a))g'(a)$.

The role of *Q* in the first proof is played by η in this proof. They are related by the equation:

$$Q(y) = f'(g(a)) + \eta(y - g(a)).$$

The need to define *Q* at $g(a)$ is analogous to the need to define η at zero. However, the proofs are not exactly equivalent. The first proof relies on a theorem about products of limits to show that the derivative exists. The second proof does not need this because showing that the error term vanishes proves the existence of the limit directly.

The chain rule in higher dimensions

The simplest generalization of the chain rule to higher dimensions uses the total derivative. The total derivative is a linear transformation that captures how the function changes in all directions. Let $f : \mathbf{R}^m \to \mathbf{R}^k$ and $g : \mathbf{R}^n \to \mathbf{R}^m$ be differentiable functions, and let D be the total derivative operator. If **a** is a point in $\mathbf{R}^n$, then the higher dimensional chain rule says that:

$$D_{\mathbf{a}}(f \circ g) = D_{g(\mathbf{a})} f \circ D_{\mathbf{a}} g,$$

or for short,

$$D(f \circ g) = Df \circ Dg.$$

In terms of Jacobian matrices, the rule says

$$J_{\mathbf{a}}(f \circ g) = J_{g(\mathbf{a})}(f) J_{\mathbf{a}}(g),$$

That is, the Jacobian of the composite function is the product of the Jacobians of the composed functions. The higher-dimensional chain rule can be proved using a technique similar to the second proof given above.

The higher-dimensional chain rule is a generalization of the one-dimensional chain rule. If k, m, and n are 1, so that $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$, then the Jacobian matrices of f and g are 1×1. Specifically, they are:

$$\begin{aligned} J_a(g) &= \left(g'(a)\right), \\ J_{g(a)}(f) &= \left(f'(g(a))\right). \end{aligned}$$

The Jacobian of $f \circ g$ is the product of these 1×1 matrices, so it is $f'(g(a))g'(a)$, as expected from the one-dimensional chain rule. In the language of linear transformations, $D_a(g)$ is the function which scales a vector by a factor of $g'(a)$ and $D_{g(a)}(f)$ is the function which scales a vector by a factor of $f'(g(a))$. The chain rule says that the composite of these two linear transformations is the linear transformation $D_a(f \circ g)$, and therefore it is the function that scales a vector by $f'(g(a))g'(a)$.

Another way of writing the chain rule is used when f and g are expressed in terms of their components as $\mathbf{y} = f(\mathbf{u}) = (f_1(\mathbf{u}), \ldots, f_k(\mathbf{u}))$ and $\mathbf{u} = g(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$. In this case, the above rule for Jacobian matrices is usually written as:

$$\frac{\partial(f_1, \ldots, f_k)}{\partial(x_1, \ldots, x_n)} = \frac{\partial(f_1, \ldots, f_k)}{\partial(u_1, \ldots, u_m)} \frac{\partial(g_1, \ldots, g_m)}{\partial(x_1, \ldots, x_n)}.$$

The chain rule for total derivatives implies a chain rule for partial derivatives. Recall that when the total derivative exists, the partial derivative in the ith coordinate direction is found by multiplying the Jacobian matrix by the ith basis vector. By doing this to the formula above, we find:

$$\frac{\partial(f_1,\ldots,f_k)}{\partial x_i} = \frac{\partial(f_1,\ldots,f_k)}{\partial(u_1,\ldots,u_m)}\frac{\partial(g_1,\ldots,g_m)}{\partial x_i}.$$

Since the entries of the Jacobian matrix are partial derivatives, we may simplify the above formula to get:

$$\frac{\partial(f_1,\ldots,f_k)}{\partial x_i} = \sum_{\ell=1}^{m}\frac{\partial(f_1,\ldots,f_k)}{\partial u_\ell}\frac{\partial g_\ell}{\partial x_i}.$$

More conceptually, this rule expresses the fact that a change in the x_i direction may change all of g_1 through g_k, and any of these changes may affect f.

In the special case where $k = 1$, so that f is a real-valued function, then this formula simplifies even further:

$$\frac{\partial f}{\partial x_i} = \sum_{\ell=1}^{m}\frac{\partial f}{\partial u_\ell}\frac{\partial g_\ell}{\partial x_i}.$$

Example

Given $u = x^2 + 2y$ where $x = r\sin(t)$ and $y = \sin^2(t)$, determine the value of $\frac{\partial u}{\partial r}$ and $\frac{\partial u}{\partial t}$ using the chain rule.

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = (2x)(\sin(t)) + (2)(0) = 2r\sin^2(t)$$

and

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} = (2x)(r\cos(t)) + (2)(2\sin(t)\cos(t))$$

$$= 2(r\sin(t))\,r\cos(t) + 4\sin(t)\cos(t) = 2\left(r^2+2\right)\sin(t)\cos(t).$$

Higher derivatives of multivariable functions

Faà di Bruno's formula for higher-order derivatives of single-variable functions generalizes to the multivariable case. If f is a function of $u = g(x)$ as above, then the second derivative of $f \circ g$ is:

$$\frac{\partial^2(f\circ g)}{\partial x_i\partial x_j} = \sum_k \frac{\partial f}{\partial u_k}\frac{\partial^2 g_k}{\partial x_i\partial x_j} + \sum_{k,\ell}\frac{\partial^2 f}{\partial u_k\partial u_\ell}\frac{\partial g_k}{\partial x_i}\frac{\partial g_\ell}{\partial x_j}.$$

The composite function chain rule notation can also be adjusted for the multivariate case:

$$\text{Given } z = f(u) \text{ and } u = g(x,y)$$
$$\text{such that } z = f\left[g(x,y)\right]$$

Then the partial derivatives of z with respect to its two independent variables are defined as:

$$\frac{\partial z}{\partial x} = \frac{dy}{du} \cdot \frac{\partial u}{\partial x}$$
$$\frac{\partial z}{\partial y} = \frac{dy}{du} \cdot \frac{\partial u}{\partial y}$$

Let's do the same example as above, this time using the composite function notation where functions within the z function are renamed. Note that either rule could be used for this problem, so when is it necessary to go to the trouble of presenting the more formal composite function notation? As problems become more complicated, renaming parts of a composite function is a better way to keep track of all parts of the problem. It is slightly more time consuming, but mistakes within the problem are less likely.

$$\text{Given } z = \left(2x + y^2\right)^3$$
$$\text{let } z = f(u) = u^3 \text{ and } u = g(x,y) = 2x + y^2$$
$$\text{then } \frac{\partial z}{\partial x} = \frac{dy}{du} \cdot \frac{\partial u}{\partial x} = \left(3u^2\right)(2) = 6u^2$$
$$\text{and } \frac{\partial z}{\partial y} = \frac{dy}{du} \cdot \frac{\partial u}{\partial y} = \left(3u^2\right)(2y) = (6y)u^2$$

The final step is the same, replace u with function g:

$$\frac{\partial z}{\partial x} = 6u^2 = 6\left(2x + y^2\right)^2$$
$$\frac{\partial z}{\partial y} = 6y(u)^2 = (6y)\left(2x + y^2\right)^2$$

Multivariate function

The rule for differentiating multivariate natural logarithmic functions, with appropriate notation changes is as follows:

$$\text{Given } z = f(u) = \ln(u) \text{ and } u = g(x,y)$$
$$\text{such that } z = \ln g(x,y)$$

Then the partial derivatives of z with respect to its independent variables are defined as:

$$\frac{\partial z}{\partial x} = \frac{1}{u} \cdot \frac{\partial g}{\partial x} = \frac{1}{g(x,y)} \cdot \frac{\partial g}{\partial x}$$
$$\frac{\partial z}{\partial y} = \frac{1}{u} \cdot \frac{\partial g}{\partial y} = \frac{1}{g(x,y)} \cdot \frac{\partial g}{\partial y}$$

Let's do an example. Find the partial derivatives of the following function:

$$z = \ln\left(2x^2 + 4y^2\right)$$
$$\frac{\partial z}{\partial x} = \frac{1}{g(x,y)} \cdot \frac{\partial g}{\partial x} = \frac{1}{\left(2x^2 + 4y^2\right)} \cdot (4x) = \frac{4x}{\left(2x^2 + 4y^2\right)}$$
$$\frac{\partial z}{\partial y} = \frac{1}{g(x,y)} \cdot \frac{\partial g}{\partial y} = \frac{1}{\left(2x^2 + 4y^2\right)} \cdot (8y) = \frac{4y}{\left(x^2 + 2y^2\right)}$$

The rule for taking partials of exponential functions can be written as:

$$\text{Given } z = f(u) = e^u \text{ and } u = g(x, y)$$
$$\text{such that } z = e^{g(x,y)}$$

Then the partial derivatives of z with respect to its independent variables are defined as:

$$\frac{\partial z}{\partial x} = e^{g(x,y)} \cdot \frac{\partial g}{\partial x}$$
$$\frac{\partial z}{\partial y} = e^{g(x,y)} \cdot \frac{\partial g}{\partial y}$$

One last time, we look for partial derivatives of the following function using the exponential rule:

$$z = e^{\left(3xy^2\right)}$$
$$\frac{\partial z}{\partial x} = e^{g(x,y)} \cdot \frac{\partial g}{\partial x} = e^{\left(3xy^2\right)} \cdot \left(3y^2\right) = \left(3y^2\right)e^{\left(3xy^2\right)}$$
$$\frac{\partial z}{\partial y} = e^{g(x,y)} \cdot \frac{\partial g}{\partial y} = e^{\left(3xy^2\right)} \cdot \left(6xy\right) = \left(6xy\right)e^{\left(3xy^2\right)}$$

Higher order partial and cross partial derivatives

The story becomes more complicated when we take higher order derivatives of multivariate functions. The interpretation of the first derivative remains the same, but there are now two second order derivatives to consider.

First, there is the direct second-order derivative. In this case, the multivariate function is differentiated once, with respect to an independent variable, holding all other variables constant. Then the result is differentiated a second time, again with respect to the same independent variable. In a function such as the following:

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

There are 2 direct second-order partial derivatives, as indicated by the following examples of notation:

$$f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2} = Z_{xx}$$
$$f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial y^2} = Z_{yy}$$

These second derivatives can be interpreted as the rates of change of the two slopes of the function z.

Now the story gets a little more complicated. The cross-partials, f_{xy} and f_{yx} are defined in the following way. First, take the partial derivative of z with respect to x. Then take the derivative again, but this time, take it with respect to y, and hold the x constant. Spatially, think of the cross partial as a measure of how the slope (change in z with respect to x) changes, when the y variable changes. The following are examples of notation for cross-partials:

$$f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial y \partial x}$$
$$f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial x \partial y}$$

We'll discuss economic meaning further in the next section, but for now, we'll just show an example, and note that in a function where the cross-partials are continuous, they will be identical. For the following function:

$$z = 2x^3 + 3xy + 2y^2$$

Take the first and second partial derivatives.

$$z = 2x^3 + 3xy + 2y^2$$
$$z_x = 6x^2 + 3y \qquad z_y = 3x + 4y$$
$$z_{xx} = 12x \qquad z_{yy} = 4$$

Now, starting with the first partials, find the cross partial derivatives:

$$z_x = 6x^2 + 3y \qquad z_y = 3x + 4y$$
$$z_{xy} = 3 \qquad z_{yx} = 3$$

4.0 **CONCLUSION**

In this unit, you have been introduced to the composite differentiation also called the chain rule. You have known the Composites of more than two functions. You have also known the quotient rule. You have solved problems on higher derivative with the use of composite differentiation. You have proof the chain rule and known the rule in higher dimension.

5.0 **SUMMARY**

In this unit, you have studied :

The chain rule

Composites of more than two functions

The quotient rule

Higher derivative

Proof of the chain rule

The rule in higher dimension

6.0 **TUTOR-MARKED ASSIGNMENT**

1.0 What are the second – order derivatives of the function F(x,y)= $xy^2 + x^3 y^5$

2.0 Express x- and y- derivatives of W($x^3 y^3$) in terms of x,y.

3.0 What are the second - order derivatives of the function F(x,y) = $x^4 y^6$.

4.0 What are the second – order derivatives of the function K(x,y) = ln (2x-3y).

5.0 What are the second – order derivatives of the function R(x,y) = $x^{\frac{1}{2}} y^{\frac{1}{3}}$.

6.0 What are the second – order derivatives of the function N(x,y) = $\tan^{-1}(x, y)$.

REFERENCES

Hernandez Rodriguez and Lopez Fernandez, *A Semiotic Reflection on the Didactics of the Chain Rule*, The Montana Mathematics Enthusiast, ISSN 1551-3440, Vol. 7, nos.2&3, pp.321–332.

Apostol, Tom (1974). *Mathematical analysis* (2nd ed. ed.). Addison Wesley. Theorem 5.

UNIT 2: EULER'S THEOREM

CONTENT

1.**0 INTRODUCTION**

2.0 OBJECTIVES

3.0 MAIN CONTENT

Statement and prove of Euler's theorem

4.0 CONCLUSION

5.0 SUMMARY

6.0 TUTOR-MARKED ASSIGNMENT

7.0 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

In number theory, **Euler's theorem** (also known as the **Fermat–Euler theorem** or **Euler's totient theorem**) states that if *n* and *a* are coprime positive integers, then

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

where φ(*n*) is Euler's totient function and "... ≡ ... (mod *n*)" denotes congruence modulo n.

2.0 OBJECTIVES

In this unit, the student should able to state and prove the Euler's theorem.

3.0 MAIN CONTENT

The converse of Euler's theorem is also true: if the above congruence holds for positive integers a and n, then a and n are coprime.

The theorem is a generalization of Fermat's little theorem, and is further generalized by Carmichael's theorem.

The theorem may be used to easily reduce large powers modulo *n*. For example, consider finding the ones place decimal digit of 7^{222}, i.e. 7^{222} (mod 10). Note that 7 and 10 are coprime, and φ(10) = 4. So Euler's theorem yields $7^4 \equiv 1$ (mod 10), and we get $7^{222} \equiv 7^{4\times55 + 2} \equiv (7^4)^{55}\times7^2 \equiv 1^{55}\times7^2 \equiv 49 \equiv 9$ (mod 10).

In general, when reducing a power of *a* modulo *n* (where *a* and *n* are coprime), one needs to work modulo φ(*n*) in the exponent of *a*:

if $x \equiv y \pmod{\varphi(n)}$, then $a^x \equiv a^y \pmod{n}$.

Euler's theorem also forms the basis of the RSA encryption system: encryption and decryption in this system together amount to exponentiating the original text by $k\varphi(n)+1$ for some positive integer k, so Euler's theorem shows that the decrypted result is the same as the original.

Proofs

1. Leonhard Euler published a proof in 1789. Using modern terminology, one may prove the theorem as follows: the numbers b which are relatively prime to n form a group under multiplication mod n, the group G of (multiplicative) units of the ring $\mathbf{Z}/n\mathbf{Z}$. This group has $\varphi(n)$ elements. The element $\mathbf{a} := a$ (mod n) is a member of the group G, and the order o(**a**) of **a** (the least k > 0 such that $\mathbf{a}^k = 1$) must have a multiple equal to the size of G. (The order of **a** is the size of the subgroup of G generated by **a**, and Lagrange's theorem states that the size of any subgroup of G divides the size of *G*.)

Thus for some integer $M > 0$, $M\cdot o(\mathbf{a}) = \varphi(n)$. Therefore $\mathbf{a}^{\varphi(n)} = a^{o(\mathbf{a})\cdot M} = (\mathbf{a}^{o(\mathbf{a})})^M = 1^M = 1$. This means that $a^{\varphi(n)} = 1 \pmod{n}$.

2. Another direct proof: if a is coprime to n, then multiplication by a permutes the residue classes mod n that are coprime to n; in other words (writing R for the set consisting of the $\varphi(n)$ different such classes) the sets $\{ x : x \text{ in } R \}$ and $\{ ax : x \text{ in } R \}$ are equal; therefore, the two products over all of the elements in each set are equal. Hence, $P \equiv a^{\varphi(n)}P \pmod{n}$ where P is the product over all of the elements in the first set. Since P is coprime to n, it follows that $a^{\varphi(n)} \equiv 1$ (mod

4.0 CONCLUSION

In this unit, you have stated and proved the Euler's theorem

5.0 SUMMARY

In this unit, you have known the statement of euler's theorem and proved euler's theorem.

6.0 TUTOR-MARKED ASSIGNMENT

State and prove euler's theorem.

References

Hernandez Rodriguez and Lopez Fernandez, *A Semiotic Reflection on the Didactics of the Chain Rule*, The Montana Mathematics Enthusiast, ISSN 1551-3440, Vol. 7, nos.2&3, pp.321–332.

Apostol, Tom (1974). *Mathematical analysis* (2nd ed. ed.). Addison Wesley. Theorem 5.5.

UNIT 3 :IMPLICIT DIFFERENTIATION

CONTENTS

1.0 **INTRODUCTION**

2.0 OBJECTIVES

3.0 MAIN CONTENT

3.1 Know the derivatives of Inverse Trigonometric Functions

3.2 Define and identify Implicit differentiation

3.3 Know formula for two variables

3.4 Know applications in economics

3.5 Solve Implicit differentiation problems

4.0 **CONCLUSION**

5.0 SUMMARY

6.0 TUTOR-MARKED ASSIGNMENT

7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

Most of our math work thus far has always allowed us to solve an equation for *y* in terms of *x*. When an equation can be solved for y we call it an *explicit* function. But not all equations can be solved for *y*. An example is:

$$x^3 + y^3 = 6xy$$

This equation cannot be solved for *y*. When an equation cannot be solved for *y*, we call it an *implicit* function. The good news is that we can still differentiate such a function. The technique is called *implicit differentiation*.

When we implicitly differentiate, we must treat *y* as a composite function and therefore we must use the chain rule with *y* terms. The reason for this can be seen in Leibnitz notation: $\frac{d}{dx}$. This notation tells us that we are differentiating with respect to *x*. Because y is not native to what are differentiating with respect to, we need to regard it as a composite function. As you know, when we differentiate a composite function we must use the chain rule.

Let's now try to differentiate the implicit function, $x^3 + y^3 = 6xy$.

$x^3 + y^3 = 6xy$

This is a "folium of Descartes" curve. This would be very difficulty to solve for y, so we will want to use implicit differentiation.

$\frac{d}{dx}\left(x^3 + y^3 = 6xy\right)$

Here we show with Leibnitz notation that we are implicitly differentiating both sides of the equation.

$\frac{d}{dx}\left(x^3\right) + \frac{d}{dx}\left(y^3\right) = \frac{d}{dx}(6$

On the left side we need to individually take the derivative of each term. On the right side we will have to use the product rule. ($6x \cdot y$)

$3x^2 + 3y^2 y' - 6xy' + 6y$

Here we take the individual derivatives. Note: Where did the *y'* come from? Because we are differentiating with respect to *x*, we need to use the chain rule on the *y*. Notice that we did use the product rule on the right side.

$3y^2 y' - 6xy' = 6y - 3x^2$

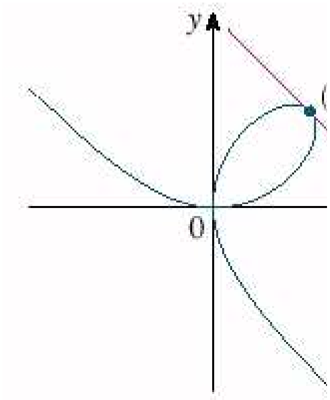
Now we get the *y'* terms on the same side of the equation.

$y'\left(3y^2 - 6x\right) = 6y - 3x^2$

Now we factor *y'* out of the expression on the left side.

$y' = \frac{6y - 3x^2}{3y^2 - 6x} = \frac{3(2y - x^2}{3(y^2 - 2x}$

Now we divide both sides by the $3y^2 - 6x$ factor and simplify.



We can see in a plot of the implicit function that the slope of the tangent line at the point (3,3) does appear to be -1.

Another example: Differentiate: $x^2 - 2xy + y^3 = c$

$$x^2 - 2xy + y^3 = c$$

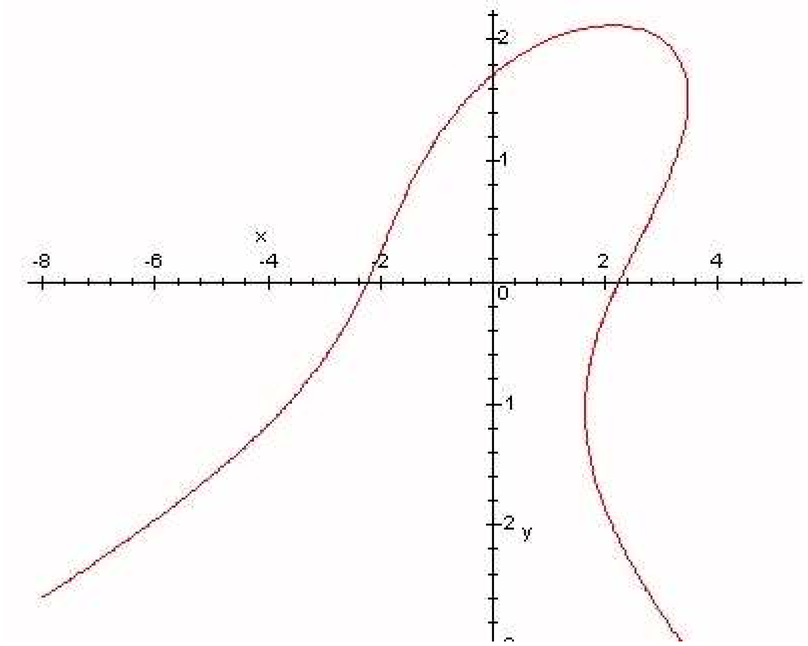
Given implicit function

$$2x - (2xy' + 2y) + 3y^2y' = 0$$

Doing implicit differentiation on the function. Note the use of the product rule on the second term

$$2x - 2xy' - 2y + 3y^2y' = 0$$
$$-2xy' + 3y^2y' = 2y - 2x$$
$$y'(3y^2 - 2x) = 2y - 2x$$
$$y' = \frac{2y - 2x}{3y^2 - 2x}$$

We do the algebra to solve for y'.



Here we see a portion of plot of the implicit equation with *c* set equal to 5.. When does it appear that the slope of the tangent line will be zero? It appears to be at about (2.2,2.2).

$$0 = \frac{2y - 2x}{3y^2 - 2x}$$
$$0 = 2y - 2x$$
$$y = x$$

We take our derivative, set it equal to zero, and solve.

$$x^2 - 2xy + y^3 = 5$$
$$x^2 - 2x^2 + x^3 = 5$$
$$x^3 - x^2 - 5 = 0$$

Now putting $x = y$ in the original implicit equation, we find that...

$x = y = 2.116343299$

We still must use a computer algebra system to solve this cubic equation. The one real answer is shown at the left. This answer does seem consistent with our visual estimate.

This can be done in Maple with the following

2.0 OBJECTIVES

At the end of this unit, you should be able to :

Know the derivatives of Inverse Trigonometric Functions

Define and identify Implicit differentiation

Know formula for two variables

Know applications in economics

Solve Implicit differentiation problems

3.0 MAIN CONTENT

Links to other explanations of Implicit Differentiation

Derivatives of Inverse Trigonometric Functions

Thanks to implicit differentiation, we can develop important derivatives that we could not have developed otherwise. The inverse trigonometric functions fall under this category. We will develop and remember the derivatives of the inverse sine and inverse tangent.

$y = \sin^{-1} x$	Inverse sine function.
$\sin y = x$	This is what inverse sine means.
$\cos y \frac{dy}{dx} = 1$	We implicitly differentiate both sides of the equation with respect to *x*. Because we are differentiating with respect to *x*, we need to use the chain rule on the left side.
$\frac{dy}{dx} = \frac{1}{\cos y}$	We solve the equation for $\frac{dy}{dx}$.
$\frac{dy}{dx} = \frac{1}{\sqrt{1-\sin^2 y}}$	This is because of the trigonometric identity, $\sin^2 y + \cos^2 y = 1$.
$\frac{dy}{dx} = \frac{1}{\sqrt{1-x^2}}$	Refer back to the equation in step two above. We have our derivative.

$$\frac{d}{dx}(\sin^{-1}x) = \frac{1}{\sqrt{1 - x^2}}$$

$$\frac{d}{dx}(\tan^{-1}x) = \frac{1}{1 + x^2}$$

Implicit differentiation

In

$y = \tan^{-1} x$	The inverse tangent function.
$\tan y = x$	This is what inverse tangent means.
$\sec^2 y \frac{dy}{dx} = 1$	We implicitly differentiate both sides of the equation with respect to *x*. Because we are differentiating with respect to *x*, we need to use the chain rule on the left side.
$\frac{dy}{dx} = \frac{1}{\sec^2 y}$	We solve the equation for $\frac{dy}{dx}$.
$\frac{dy}{dx} = \frac{1}{1+\tan^2 x}$	This is because of the trigonometric identity, $\tan^2 y + 1 = \sec^2 y$.
$\frac{dy}{dx}\left(\tan^{-1} x\right) = \frac{1}{1+x^2}$	Refer back to the equation in step two above. We have our derivative.

calculus, a method called **implicit differentiation** makes use of the chain rule to differentiate implicitly defined functions.

As explained in the introduction, *y* can be given as a function of *x* implicitly rather than explicitly. When we have an equation *R*(*x*, *y*) = 0, we may be able to solve it for *y* and then differentiate. However, sometimes it is simpler to differentiate *R*(*x*, *y*) with respect to *x* and *y* and then solve for *dy*/*dx*.

Examples

1. Consider for example

$$y + x + 5 = 0$$

This function normally can be manipulated by using algebra to change this equation to one expressing *y* in terms of an explicit function:

$$y = -x - 5\,,$$

where the right side is the explicit function whose output value is *y*. Differentiation then gives $\frac{dy}{dx} = -1$. Alternatively, one can totally differentiate the original equation:

$$\frac{dy}{dx} + \frac{dx}{dx} + \frac{d}{dx}(5) = 0;$$

$$\frac{dy}{dx} + 1 = 0.$$

Solving for $\frac{dy}{dx}$ gives:

$$\frac{dy}{dx} = -1,$$

the same answer as obtained previously.

2. An example of an implicit function, for which implicit differentiation might be easier than attempting to use explicit differentiation, is

$$x^4 + 2y^2 = 8$$

In order to differentiate this explicitly with respect to *x*, one would have to obtain (via algebra)

$$y = f(x) = \pm\sqrt{\frac{8 - x^4}{2}},$$

and then differentiate this function. This creates two derivatives: one for $y > 0$ and another for $y < 0$.

One might find it substantially easier to implicitly differentiate the original function:

$$4x^3 + 4y\frac{dy}{dx} = 0,$$

giving,

$$\frac{dy}{dx} = \frac{-4x^3}{4y} = \frac{-x^3}{y}$$

3. Sometimes standard explicit differentiation cannot be used and, in order to obtain the derivative, implicit differentiation must be employed. An example of such a case is the equation $y^5 - y = x$. It is impossible to express *y* explicitly as a function of *x* and therefore *dy*/*dx* cannot be found by explicit differentiation. Using the implicit method, *dy*/*dx* can be expressed:

$$5y^4\frac{dy}{dx} - \frac{dy}{dx} = \frac{dx}{dx}$$

where $\frac{dx}{dx} = 1$. Factoring out $\frac{dy}{dx}$ shows that

$$\frac{dy}{dx}(5y^4 - 1) = 1$$

which yields the final answer

$$\frac{dy}{dx} = \frac{1}{5y^4 - 1},$$

which is defined for $y \neq \pm\frac{1}{\sqrt[4]{5}}.$

Formula for two variables

"The Implicit Function Theorem states that if *F* is defined on an open disk containing (*a*,*b*), where *F*(*a*,*b*) = 0, $F_y(a,b) \neq 0$, and F_x and F_y are continuous on the disk, then the equation *F*(*x*,*y*) = 0 defines *y* as a function of *x* near the point (*a*,*b*) and the derivative of this function is given by..."[1]:§ 11.5

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{F_x}{F_y},$$

where F_x and F_y indicate the derivatives of *F* with respect to *x* and *y*.

The above formula comes from using the generalized chain rule to obtain the total derivative—with respect to *x*—of both sides of *F*(*x*, *y*) = 0:

$$\frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0,$$

and hence

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0.$$

Implicit function theorem

It can be shown that if *R*(*x*,*y*) is given by a smooth submanifold *M* in $\mathbb{R}^2$, and (*a*,*b*) is a point of this submanifold such that the tangent space there is not vertical (that is $\frac{\partial R}{\partial y} \neq 0$), then *M* in some small enough neighbourhood of (*a*,*b*) is given by a parametrization (*x*,*f*(*x*)) where *f* is a smooth function. In less technical language, implicit functions exist and can be differentiated, unless the tangent to the supposed graph would be vertical. In the standard case where we are given an equation

R(*x*,*y*) = 0

the condition on *R* can be checked by means of partial derivatives .

Applications in economics

Marginal rate of substitution

In economics, when the level set $R(x,y) = 0$ is an indifference curve for the quantities x and y consumed of two goods, the absolute value of the implicit derivative is interpreted as the marginal rate of substitution of the two goods: how much more of y one must receive in order to be indifferent to a loss of 1 unit of x.

IMPLICIT DIFFERENTIATION PROBLEMS

The following problems require the use of implicit differentiation. Implicit differentiation is nothing more than a special case of the well-known chain rule for derivatives. The majority of differentiation problems in first-year calculus involve functions y written EXPLICITLY as functions of x . For example, if

$$y = 3x^2 - \sin(7x + 5)$$

,

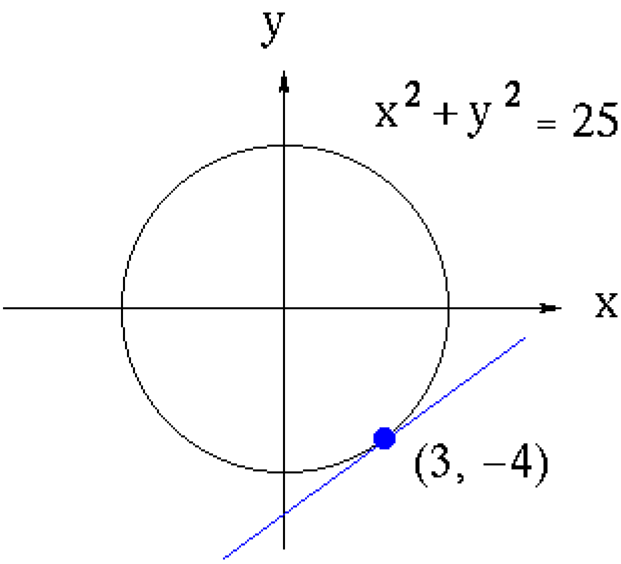
then the derivative of y is

.

However, some functions y are written IMPLICITLY as functions of x . A familiar example of this is the equation

$x^2 + y^2 = 25$,

which represents a circle of radius five centered at the origin. Suppose that we wish to find the slope of the line tangent to the graph of this equation at the point (3, -4) .



How could we find the derivative of y in this instance ? One way is to first write y explicitly as a function of x . Thus,

$x^2 + y^2 = 25$,

$y^2 = 25 - x^2$,

and

$$y = \pm\sqrt{25 - x^2}$$,

where the positive square root represents the top semi-circle and the negative square root represents the bottom semi-circle. Since the point (3, -4) lies on the bottom semi-circle given by

$$y = -\sqrt{25 - x^2}$$,

the derivative of *y* is

$$y' = -(1/2)\left(25 - x^2\right)^{-1/2}(-2x) = \frac{x}{\sqrt{25 - x^2}}$$,

i.e.,

$$y' = \frac{x}{\sqrt{25 - x^2}}$$.

Thus, the slope of the line tangent to the graph at the point (3, -4) is

$$m = y' = \frac{(3)}{\sqrt{25 - (3)^2}} = \frac{3}{4}$$.

Unfortunately, not every equation involving *x* and *y* can be solved explicitly for *y* . For the sake of illustration we will find the derivative of *y* WITHOUT writing *y* explicitly as a function of *x* . Recall that the derivative (D) of a function of *x* squared, $(f(x))^2$, can be found using the chain rule :

$$D\{(f(x))^2\} = 2f(x)\ D\{f(x)\} = 2f(x)f'(x)$$.

Since *y* symbolically represents a function of *x*, the derivative of y^2 can be found in the same fashion :

$$D\{y^2\} = 2y\ D\{y\} = 2yy'$$.

Now begin with

$x^2 + y^2 = 25$.

Differentiate both sides of the equation, getting

$D\ (\ x^2 + y^2\ ) = D\ (\ 25\ )$,

$D (x^2) + D (y^2) = D (25)$,

and

$2x + 2 y y' = 0$,

so that

$2 y y' = - 2x$,

and

$$y' = \frac{-2x}{2y} = \frac{-x}{y} ,$$

i.e.,

$$y' = \frac{-x}{y} .$$

Thus, the slope of the line tangent to the graph at the point (3, -4) is

$$m = y' = \frac{-(3)}{(-4)} = \frac{3}{4} .$$

This second method illustrates the process of implicit differentiation. It is important to note that the derivative expression for explicit differentiation involves *x* only, while the derivative expression for implicit differentiation may involve BOTH *x* AND *y* .

The following problems range in difficulty from average to challenging.

PROBLEM 1 : Assume that *y* is a function of *x* . Find $y' = dy/dx$ for $x^3 + y^3 = 4$.

SOLUTION 1 : Begin with $x^3 + y^3 = 4$. Differentiate both sides of the equation, getting

$D (x^3 + y^3) = D (4)$,

$D (x^3) + D (y^3) = D (4)$,

(Remember to use the chain rule on $D (y^3)$.)

$3x^2 + 3y^2 y' = 0$,

so that (Now solve for *y'* .)

$3y^2 y' = - 3x^2$,

and

$$y' = \frac{-3x^2}{3y^2} = \frac{-x^2}{y^2}$$

SOLUTION 2 : Begin with $(x-y)^2 = x + y - 1$. Differentiate both sides of the equation, getting

$D(x-y)^2 = D(x + y - 1)$,

$D(x-y)^2 = D(x) + D(y) - D(1)$,

(Remember to use the chain rule on $D(x-y)^2$.)

$$2(x-y)\ D(x-y) = 1 + y' - 0 ,$$

$2(x-y)(1-y') = 1 + y'$,

so that (Now solve for y' .)

$2(x-y) - 2(x-y)y' = 1 + y'$,

$-2(x-y)y' - y' = 1 - 2(x-y)$,

(Factor out y' .)

$y'[-2(x-y) - 1] = 1 - 2(x-y)$,

and

$$y' = \frac{1-2(x-y)}{-2(x-y)-1} = \frac{2y-2x+1}{2y-2x-1} .$$

SOLUTION 3 : Begin with $y = \sin(3x + 4y)$. Differentiate both sides of the equation, getting

$$D(y) = D(\sin(3x+4y)) ,$$

(Remember to use the chain rule on $D(\sin(3x+4y))$.)

$$y' = \cos(3x+4y)\ D(3x+4y) ,$$

$$y' = \cos(3x+4y)(3+4y') ,$$

so that (Now solve for y' .)

$$y' = 3\cos(3x+4y) + 4y'\cos(3x+4y)$$,

$$y' - 4y'\cos(3x+4y) = 3\cos(3x+4y)$$,

(Factor out *y*' .)

$$y'[1 - 4\cos(3x+4y)] = 3\cos(3x+4y)$$,

and

$$y' = \frac{3\cos(3x+4y)}{1 - 4\cos(3x+4y)}$$.

SOLUTION 4 : Begin with $y = x^2 y^3 + x^3 y^2$. Differentiate both sides of the equation, getting

$D(y) = D(x^2 y^3 + x^3 y^2)$,

$D(y) = D(x^2 y^3) + D(x^3 y^2)$,

(Use the product rule twice.)

$$y' = \{x^2 D(y^3) + D(x^2) y^3\} + \{x^3 D(y^2) + D(x^3) y^2\}$$,

(Remember to use the chain rule on $D(y^3)$ and $D(y^2)$.)

$$y' = \{x^2 (3y^2 y') + (2x) y^3\} + \{x^3 (2yy') + (3x^2) y^2\}$$,

$y' = 3x^2 y^2 y' + 2x y^3 + 2x^3 y y' + 3x^2 y^2$,

so that (Now solve for *y*' .)

$y' - 3x^2 y^2 y' - 2x^3 y y' = 2x y^3 + 3x^2 y^2$,

(Factor out *y*' .)

$y' [1 - 3x^2 y^2 - 2x^3 y] = 2x y^3 + 3x^2 y^2$,

and

$$y' = \frac{2xy^3 + 3x^2 y^2}{1 - 3x^2 y^2 - 2x^3 y}$$.

SOLUTION 5 : Begin with $e^{xy} = e^{4x} - e^{5y}$. Differentiate both sides of the equation, getting

,

$$D(e^{xy}) = D(e^{4x}) - D(e^{5y})$$

,

$$e^{xy} D(xy) = e^{4x} D(4x) - e^{5y} D(5y)$$

,

$$e^{xy}(xy' + (1)y) = e^{4x}(4) - e^{5y}(5y')$$

,

so that (Now solve for y'.)

$$xe^{xy}y' + ye^{xy} = 4e^{4x} - 5e^{5y}y'$$

,

$$xe^{xy}y' + 5e^{5y}y' = 4e^{4x} - ye^{xy}$$

,

(Factor out y'.)

$$y'[xe^{xy} + 5e^{5y}] = 4e^{4x} - ye^{xy}$$

,

and

$$y' = \frac{4e^{4x} - ye^{xy}}{xe^{xy} + 5e^{5y}}$$

.

SOLUTION 6 : Begin with $\cos^2 x + \cos^2 y = \cos(2x + 2y)$. Differentiate both sides of the equation, getting

$$D(\cos^2 x + \cos^2 y) = D(\cos(2x + 2y))$$

,

$$D(\cos^2 x) + D(\cos^2 y) = D(\cos(2x + 2y))$$

,

$$(2\cos x)D(\cos x) + (2\cos y)D(\cos y) = -\sin(2x + 2y)\, D(2x + 2y)$$

,

$$2\cos x(-\sin x) + 2\cos y(-\sin y)(y') = -\sin(2x+2y)(2+2y')$$,

so that (Now solve for *y'* .)

$$-2\cos x\sin x - 2y'\cos y\sin y = -2\sin(2x+2y) - 2y'\sin(2x+2y)$$,

$$2y'\sin(2x+2y) - 2y'\cos y\sin y = -2\sin(2x+2y) + 2\cos x\sin x$$,

(Factor out *y'* .)

$$y'[2\sin(2x+2y) - 2\cos y\sin y] = 2\cos x\sin x - 2\sin(2x+2y)$$,

$$y' = \frac{2\cos x\sin x - 2\sin(2x+2y)}{2\sin(2x+2y) - 2\cos y\sin y}$$,

$$y' = \frac{2[\cos x\sin x - \sin(2x+2y)]}{2[\sin(2x+2y) - \cos y\sin y]}$$,

and

$$y' = \frac{\cos x\sin x - \sin(2x+2y)}{\sin(2x+2y) - \cos y\sin y}$$.

SOLUTION 7 : Begin with $x = 3 + \sqrt{x^2+y^2}$. Differentiate both sides of the equation, getting

$$D(x) = D(3 + \sqrt{x^2+y^2})$$,

$1 = (1/2)(x^2 + y^2)^{-1/2} D(x^2 + y^2)$,

$1 = (1/2)(x^2 + y^2)^{-1/2} (2x + 2y y')$,

so that (Now solve for *y'* .)

$$1 = \frac{(1/2)(2)(x+yy')}{\sqrt{x^2+y^2}}$$,

$$1 = \frac{x+yy'}{\sqrt{x^2+y^2}}$$,

$$\sqrt{x^2+y^2} = x + yy'$$
,

$$\sqrt{x^2+y^2} - x = yy'$$
,

and

$$y' = \frac{\sqrt{x^2+y^2} - x}{y}$$
.

SOLUTION 8 : Begin with $\frac{x-y^3}{y+x^2} = x+2$. Clear the fraction by multiplying both sides of the equation by $y + x^2$, getting

$$\frac{x-y^3}{y+x^2}(y+x^2) = (x+2)(y+x^2)$$
,

or

$x - y^3 = xy + 2y + x^3 + 2x^2$.

Now differentiate both sides of the equation, getting

$D (x - y^3) = D (xy + 2y + x^3 + 2x^2)$,

$D (x) - D (y^3) = D (xy) + D (2y) + D (x^3) + D (2x^2)$,

(Remember to use the chain rule on $D (y^3)$.)

$1 - 3 y^2 y' = (xy' + (1)y) + 2 y' + 3x^2 + 4x$,

so that (Now solve for y' .)

$1 - y - 3x^2 - 4x = 3 y^2 y' + xy' + 2 y'$,

(Factor out y' .)

$1 - y - 3x^2 - 4x = (3y^2 + x + 2) y'$,

and

$$y' = \frac{1-y-3x^2-4x}{3y^2+x+2}$$
.

SOLUTION 9 : Begin with $\frac{y}{x^3}+\frac{x}{y^3}=x^2y^4$. Clear the fractions by multiplying both sides of the equation by $x^3 y^3$, getting

$$\left\{\frac{y}{x^3}+\frac{x}{y^3}\right\}(x^3y^3)=x^2y^4(x^3y^3)$$,

$$\frac{yx^3y^3}{x^3}+\frac{xx^3y^3}{y^3}=x^2x^3y^4y^3$$,

$y^4 + x^4 = x^5 y^7$.

Now differentiate both sides of the equation, getting

$D (y^4 + x^4) = D (x^5 y^7)$,

$D (y^4) + D (x^4) = x^5 D (y^7) + D (x^5) y^7$,

(Remember to use the chain rule on $D (y^4)$ and $D (y^7)$.)

$4 y^3 y' + 4 x^3 = x^5 (7 y^6 y') + (5 x^4) y^7$,

so that (Now solve for y' .)

$4 y^3 y' - 7 x^5 y^6 y' = 5 x^4 y^7 - 4 x^3$,

(Factor out y' .)

$y' [4 y^3 - 7 x^5 y^6] = 5 x^4 y^7 - 4 x^3$,

and

$$y'=\frac{5x^4y^7-4x^3}{4y^3-7x^5y^6}$$.

SOLUTION 10 : Begin with $(x^2+y^2)^3 = 8x^2y^2$. Now differentiate both sides of the equation, getting

$D (x^2+y^2)^3 = D (8x^2y^2)$,

$3 (x^2+y^2)^2 D (x^2+y^2) = 8x^2 D (y^2) + D (8x^2) y^2$,

(Remember to use the chain rule on $D (y^2)$.)

$3 (x^2+y^2)^2 (2x + 2 y y') = 8x^2 (2 y y') + (16 x) y^2$, so that (Now solve for y' .)

$6x (x^2+y^2)^2 + 6 y (x^2+y^2)^2 y' = 16 x^2 y y' + 16 x y^2$,

$6 y (x^2+y^2)^2 y' - 16 x^2 y y' = 16 x y^2 - 6x (x^2+y^2)^2$,

(Factor out y' .)

$y' [6 y (x^2+y^2)^2 - 16 x^2 y] = 16 x y^2 - 6x (x^2+y^2)^2$,

and

$$y' = \frac{16x y^2 - 6x (x^2 + y^2)^2}{6y(x^2 + y^2)^2 - 16x^2 y}$$
.

Thus, the slope of the line tangent to the graph at the point (-1, 1) is

$$m = y' = \frac{16(-1)(1)^2 - 6(-1)((-1)^2 + (1)^2)^2}{6(1)((-1)^2 + (1)^2)^2 - 16(-1)^2(1)} = \frac{8}{8} = 1$$
,

and the equation of the tangent line is

$y - (1) = (1) (x - (-1))$

or

$y = x + 2$

SOLUTION 11 : Begin with $x^2 + (y-x)^3 = 9$. If $x=1$, then

$(1)^2 + (y-1)^3 = 9$

so that

$(y-1)^3 = 8$,

$y-1 = 2$,

$y = 3$,

and the tangent line passes through the point (1, 3) . Now differentiate both sides of the original equation, getting

$D (x^2 + (y-x)^3) = D (9)$,

$D (x^2) + D (y-x)^3 = D (9)$,

$2x + 3 (y-x)^2 D (y-x) = 0$,

$2x + 3 (y-x)^2 (y'-1) = 0$,

so that (Now solve for y' .)

$2x + 3\,(y-x)^2\,y' - 3\,(y-x)^2 = 0$,

$3\,(y-x)^2\,y' = 3\,(y-x)^2 - 2x$,

and

$$y' = \frac{3(y-x)^2 - 2x}{3(y-x)^2}$$.

Thus, the slope of the line tangent to the graph at (1, 3) is

$$m = y' = \frac{3(3-1)^2 - 2(1)}{3(3-1)^2} = \frac{10}{12} = \frac{5}{6}$$,

and the equation of the tangent line is

$y - (3) = (5/6)(x - (1))$,

or

$y = (7/6)\,x + (13/6)$.

SOLUTION 12 : Begin with $x^2y + y^4 = 4 + 2x$. Now differentiate both sides of the original equation, getting

$D(x^2 y + y^4) = D(4 + 2x)$,

$D(x^2 y) + D(y^4) = D(4) + D(2x)$,

$(x^2 y' + (2x)\,y) + 4\,y^3\,y' = 0 + 2$,

so that (Now solve for y' .)

$x^2 y' + 4\,y^3\,y' = 2 - 2x\,y$,

(Factor out y' .)

$y'\,[\,x^2 + 4\,y^3\,] = 2 - 2x\,y$,

and

(Equation 1)

$$y' = \frac{2 - 2xy}{x^2 + 4y^3}$$.

Thus, the slope of the graph (the slope of the line tangent to the graph) at (-1, 1) is

$$y' = \frac{2-2(-1)(1)}{(-1)^2+4(1)^3} = \frac{4}{5}$$

.

Since *y*'= 4/5 , the slope of the graph is 4/5 and the graph is increasing at the point (-1, 1) . Now determine the concavity of the graph at (-1, 1) . Differentiate Equation 1, getting

$$y'' = \frac{(x^2+4y^3)D(2-2xy)-(2-2xy)D(x^2+4y^3)}{(x^2+4y^3)^2}$$

$$= \frac{(x^2+4y^3)((-2x)y'+(-2)y)-(2-2xy)(2x+12y^2y')}{(x^2+4y^3)^2}$$

.

Now let *x*=-1 , *y*=1 , and *y*'=4/5 so that the second derivative is

$$y'' = \frac{[(-1)^2+4(1)^3][(-2(-1))(4/5)+(-2)(1)]-[2-2(-1)(1)][2(-1)+12(1)^2(4/5)]}{((-1)^2+4(1)^3)^2}$$

$$= \frac{(5)(8/5-2)-(4)(-2+48/5)}{25}$$

$$= \frac{-2-(152/5)}{25}$$

$$= \frac{-162}{125}$$

.

Since *y*'' < 0 , the graph is concave down at the point (-1, 1)

4.0 CONCLUSION

In this unit you have studied the derivative of inverse of trigonometric functions. You have known the definition of implicit differentiation and have identified problems on implicit differentiation. You have also studied the formular for two variables and implicit differentiation applications in economics. You have solved various examples on implicit differentiation.

5.0 SUMMARY

In this course you have studied

The derivatives of Inverse Trigonometric Functions

Definition and identification of Implicit differentiation

The formula for two variables

The applications in economics

Implicit differentiation problems

5.0 TUTOR-MARKED ASSIGNMENT

Find the equation of the tangent line to the ellipse 25 $x^2 + y^2 = 109$

if Find y' if $y^4 + 4y - 3x^3 \sin(y) = 2x + 1$.

Find y' if $xy^3 + x^2y^2 + 3x^2 - 6 = 1$.

. Show that if a normal line to each point on an ellipse passes through the center of an ellipse, then the ellipse is a circle.

7.0 REFERENCES

^ *a b* Stewart, James (1998). *Calculus Concepts And Contexts*. Brooks/Cole Publishing Company. ISBN 0-534-34330-9.

Rudin, Walter (1976). *Principles of Mathematical Analysis*. McGraw-Hill. ISBN 0-07-054235-X.

Spivak, Michael (1965). *Calculus on Manifolds*. HarperCollins. ISBN 0-8053-9021-9.

Warner, Frank (1983). *Foundations of Differentiable Manifolds and Lie Groups*. Springer. ISBN 0-387-90894

MODULE 6 TAYLOR'S SERIES EXPANSION

-Unit 1: Function of two variables

-Unit 2: Taylor's series expansion for functions of two variables.

-Unit 3: Application of Taylor's series.

UNIT 1 : FUNCTIONS OF TWO VARIABLES

CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT

3.1 Solve problems on partial derivatives in calculus

3.2 Solve problems on higher order partial derivative

3.3 State and apply clairauts theorem

3.4 Solve problem on maxima and manima

3.5 Identify Taylor series of function of two variable

3.6 Know analytical function

4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

Functions of Two Variables
Definition of a function of two variables
Until now, we have only considered functions of a single variable.
However, many real-world functions consist of two (or more) variables. E.g., the area function of a rectangular shape depends on both its width and its height. And, the pressure of a given quantity of gas varies with respect to the temperature of the gas and its volume. We define a function of two variables as follows:
A function *f* of two variables is a relation that assigns to every ordered pair of input values x, y **in a set called the *domain of* a unique output value denoted by,** $f(x, y)$**. The set of output values is called the *range*.**
Since the domain consists of ordered pairs, we may consider the domain to be all (or part) of the *x*-*y* plane.

Unless otherwise stated, we will assume that the variables *x* and *y* and the output Value *f* (*x*, *y*).

2.0 OBJECTIVE

At this unit, you should be able to :

- Solve problems on partial derivatives in calculus
- Solve problems on higher order partial derivative
- State and apply clairauts theorem
- Solve problem on maxima and manima
- Identify Taylor series of function of two variable
- Know analytical function

3.0 **MAIN CONTENT**

Partial Derivatives in Calculus

Let f(x,y) be a function with two variables. If we keep y constant and differentiate f (assuming f is differentiable) with respect to the variable x, we obtain what is called the **partial derivative** of f with respect to x which is denoted by

$$\frac{\partial f}{\partial x} \text{ or } f_x$$

We might also define partial derivatives of function f as follows:

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

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

We now present several examples with detailed solution on how to calculate partial derivatives.

Example 1: Find the partial derivatives f_x and f_y if f(x , y) is given by

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

Solution to Example 1:

Assume y is constant and differentiate with respect to x to obtain

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}[\, x^2 y + 2x + y \,]$$

$$= \frac{\partial}{\partial x}[\, x^2 y] + \frac{\partial}{\partial x}[\, 2x \,] + \frac{\partial}{\partial x}[\, y \,] = [2xy] + [\, 2 \,] + [\, 0 \,] = 2xy + 2$$

Now assume x is constant and differentiate with respect to y to obtain

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}[\, x^2 y + 2x + y \,]$$

$$= \frac{\partial}{\partial y}[\, x^2 y] + \frac{\partial}{\partial y}[\, 2x \,] + \frac{\partial}{\partial y}[\, y \,] = [\, x^2 \,] + [\, 0 \,] + [\, 1 \,] = x^2 + 1$$

Example 2: Find f_x and f_y if f(x , y) is given by

f(x , y) = sin(x y) + cos x

Solution to Example 2:

Differentiate with respect to x assuming y is constant

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}[\, \sin(xy) + \cos x \,] = y\cos(xy) - \sin x$$

Differentiate with respect to y assuming x is constant

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}[\, \sin(xy) + \cos x \,] = x\cos(xy)$$

Example 3: Find f_x and f_y if f(x , y) is given by

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

Solution 3:

Differentiate with respect to x assuming y is constant

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}[\, x e^{xy} \,] = e^{xy} + xy\, e^{xy} = (xy + 1)e^{xy}$$

Differentiate with respect to y

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} [\, x\, e^{x\,y} \,] = (x)\,(x\, e^{x\,y}) = x^2\, e^{x\,y}$$

Example 4: Find f_x and f_y if f(x , y) is given by

$f(x , y) = \ln (x^2 + 2 y)$

Solution

Differentiate with respect to x to obtain

$$f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} [\, \ln (x^2 + 2 y) \,] = \frac{2x}{x^2 + 2 y}$$

Differentiate with respect to y

$$f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} [\, \ln (x^2 + 2 y) \,] = \frac{2}{x^2 + 2 y}$$

Example 5: Find $f_x(2 , 3)$ and $f_y(2 , 3)$ if f(x , y) is given by

$f(x , y) = y\, x^2 + 2 y$

Solution to Example 5:

We first find f_x and f_y

$f_x(x,y) = 2x\, y$

$f_y(x,y) = x^2 + 2$

We now calculate $f_x(2 , 3)$ and $f_y(2 , 3)$ by substituting x and y by their given values

$f_x(2,3) = 2\,(2)(3) = 12$

$f_y(2,3) = 2^2 + 2 = 6$

Exercise: Find partial derivatives f_x and f_y of the following functions

1. $f(x , y) = x\, e^{x + y}$

2. f(x , y) = ln (2 x + y x)

3. f(x , y) = x sin(x - y)

Answer to Above Exercise:

1. f_x =(x + 1)$e^{x + y}$, f_y = x $e^{x + y}$

2. f_x = 1 / x , f_y = 1 / (y + 2)

3. f_x = x cos (x - y) + sin (x - y) , f_y = -x cos (x - y)

More on partial derivatives and mutlivariable functions. Multivariable Functions

Higher Order Partial Derivatives

Just as we had higher order derivatives with functions of one variable we will also have higher order derivatives of functions of more than one variable. However, this time we will have more options since we do have more than one variable..Consider the case of a function of two variables, $f(x,y) f(x,y)$ since both of the first order partial derivatives are also functions of *x* and *y* we could in turn differentiate each with respect to *x* or *y*. This means that for the case of a function of two variables there will be a total of four possible second order derivatives. Here they are and the notations that we'll use to denote them.

$$(f_x)_x = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$$
$$(f_x)_y = f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$
$$(f_y)_x = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$$
$$(f_y)_y = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$$

The second and third second order partial derivatives are often called mixed partial derivatives since we are taking derivatives with respect to more than one variable. Note as well that the order that we take the derivatives in is given by the notation for each these. If we are using the subscripting notation, *e.g.* f_{xy}, then we will differentiate from left to right. In other words, in this case, we will differentiate first with respect to *x* and then with respect to *y*. With the fractional notation *e.g.* $\frac{\partial^2 f}{\partial y \partial x}$, it is the opposite. In these cases we differentiate moving along the denominator from right to left. So, again, in this case we differentiate with respect to *x* first and then

Let's take a quick look at an example.

Example 1 Find all the second order derivatives for

$$f(x,y) = \cos(2x) - x^2 e^{5y} + 3y^2$$

Solution
We'll first need the first order derivatives so here they are.

$$f_x(x,y) = -2\sin(2x) - 2x e^{5y}$$
$$f_y(x,y) = -5x^2 e^{5y} + 6y$$

Now, let's get the second order derivatives.

$$f_{xx} = -4\cos(2x) - 2e^{5y}$$
$$f_{xy} = -10x e^{5y}$$
$$f_{yx} = -10x e^{5y}$$
$$f_{yy} = -25x^2 e^{5y} + 6$$

Notice that we dropped the (x,y) from the derivatives. This is fairly standard and we will be doing it most of the time from this point on. We will also be dropping it for the first order derivatives in most cases.

Now let's also notice that, in this case, $f_{xy} = f_{yx}$. This is not by coincidence. If the function is "nice enough" this will always be the case. So, what's "nice enough"? The following theorem tells us.

Clairaut's Theorem

Suppose that *f* is defined on a disk *D* that contains the point (a,b). If the functions f_{xy} and f_{yx} are continuous on this disk then,

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

Now, do not get too excited about the disk business and the fact that we gave the theorem is for a specific point. In pretty much every example in this class if the two mixed second order partial derivatives are continuous then they will be equal.

Example 2 Verify Clairaut's Theorem for $f(x,y) = x e^{-x^2 y^2}$.

Solution
We'll first need the two first order derivatives.

$$f_x(x,y) = \mathbf{e}^{-x^2y^2} - 2x^2y^2\mathbf{e}^{-x^2y^2}$$
$$f_y(x,y) = -2yx^3\mathbf{e}^{-x^2y^2}$$

Now, compute the two fixed second order partial derivatives.

$$f_{xy}(x,y) = -2yx^2\mathbf{e}^{-x^2y^2} - 4x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2} = -6x^2y\mathbf{e}^{-x^2y^2} + 4x^4y^3\mathbf{e}^{-x^2y^2}$$
$$f_{yx}(x,y) = -6yx^2\mathbf{e}^{-x^2y^2} + 4y^3x^4\mathbf{e}^{-x^2y^2}$$

Sure enough they are the same.

So far we have only looked at second order derivatives. There are, of course, higher order derivatives as well. Here are a couple of the third order partial derivatives of function of two variables.

$$f_{xyx} = (f_{xy})_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial^3 f}{\partial x \partial y \partial x}$$
$$f_{yxx} = (f_{yx})_x = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) = \frac{\partial^3 f}{\partial x^2 \partial y}$$

Notice as well that for both of these we differentiate once with respect to *y* and twice with respect to *x*. There is also another third order partial derivative in which we can do this, f_{xxy}. There is an extension to Clairaut's Theorem that says if all three of these are continuous then they should all be equal,

$$f_{xxy} = f_{xyx} = f_{yxx}$$

To this point we've only looked at functions of two variables, but everything that we've done to this point will work regardless of the number of variables that we've got in the function and there are natural extensions to Clairaut's theorem to all of these cases as well. For instance,

$$f_{xz}(x,y,z) = f_{zx}(x,y,z)$$

provided both of the derivatives are continuous.

In general, we can extend Clairaut's theorem to any function and mixed partial derivatives. The only requirement is that in each derivative we differentiate with respect to each variable the same number of times. In other words, provided we meet the continuity condition, the following will be equal

$$f_{ssrtsrr} = f_{trsrssr}$$

because in each case we differentiate with respect to *t* once, *s* three times and *r* three times.

Let’s do a couple of examples with higher (well higher order than two anyway) order derivatives and functions of more than two variables.

Example 3 Find the indicated derivative for each of the following functions.

(a) Find f_{xxyzz} for $f(x,y,z) = z^3 y^2 \ln(x)$

(b) Find $\dfrac{\partial^3 f}{\partial y \partial x^2}$ for $f(x,y) = \mathbf{e}^{xy}$

Solution

(a)Find f_{xxyzz} for $f(x,y,z) = z^3 y^2 \ln(x)$

In this case remember that we differentiate from left to right. Here are the derivatives for this part.

$$f_x = \frac{z^3 y^2}{x}$$

$$f_{xx} = -\frac{z^3 y^2}{x^2}$$

$$f_{xxy} = -\frac{2z^3 y}{x^2}$$

$$f_{xxyz} = -\frac{6z^2 y}{x^2}$$

$$f_{xxyzz} = -\frac{12zy}{x^2}$$

(b) Find $\dfrac{\partial^3 f}{\partial y \partial x^2}$ for $f(x,y) = \mathbf{e}^{xy}$

Here we differentiate from right to left. Here are the derivatives for this function.

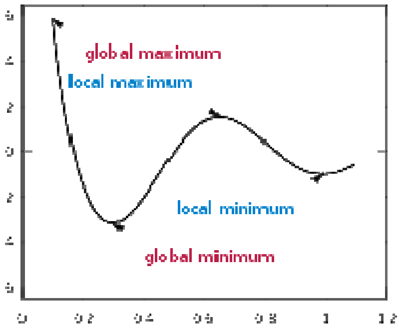
$$\frac{\partial f}{\partial x} = y\mathbf{e}^{xy}$$

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

$$\frac{\partial^3 f}{\partial y \partial x^2} = 2y e^{xy} + xy^2 e^{xy}$$

Maxima and minima

For other uses, see Maxima (disambiguation) and Maximum (disambiguation). For use in statistics, see Maximum (statistics).



Local and global maxima and minima for cos(3*πx*)/*x*, 0.1≤*x*≤1.1

In mathematics, the **maximum** and **minimum** (plural: maxima and minima) of a function, known collectively as **extrema** (singular: extremum), are the largest and smallest value that the function takes at a point either within a given neighborhood (*local* or *relative* extremum) or on the function domain in its entirety (*global* or *absolute* extremum). More generally, the maximum and minimum of a set (as defined in set theory) are the greatest and least element in the set. Unbounded infinite sets such as the set of real numbers have no minimum and maximum.

To locate extreme values is the basic objective of optimization

real-valued function *f* defined on a real line is said to have a **local (or relative) maximum point** at the point x^*, if there exists some $\varepsilon > 0$ such that $f(x^*) \geq f(x)$ when $|x - x^*| < \varepsilon$. The value of the function at this point is called **maximum** of the function. Similarly, a function has a **local minimum point** at x^*, if $f(x^*) \leq f(x)$ when $|x - x^*| < \varepsilon$. The value of the function at this point is called **minimum** of the function. A function has a **global** (or **absolute**) **maximum point** at x^* if $f(x^*) \geq f(x)$ for all x. Similarly, a function has a **global (or absolute) minimum point** at x^* if $f(x^*) \leq f(x)$ for all x. The global maximum and global minimum points are also known as the arg max and arg min: the argument (input) at which the maximum (respectively, minimum) occurs.

Restricted domains: There may be maxima and minima for a function whose domain does not include all real numbers. A real-valued function, whose domain is any set, can have a global maximum and minimum. There may also be local maxima and local minima points, but only at points of the domain set where the concept of neighborhood is defined. A neighborhood plays the role of the set of x such that $|x - x^*| < \varepsilon$.

A continuous (real-valued) function on a compact set always takes maximum and minimum values on that set. An important example is a function whose domain is a closed (and bounded) interval of real numbers (see the graph above). The neighborhood requirement precludes a *local* maximum or minimum at an endpoint of an interval. However, an endpoint may still be a *global* maximum or minimum. Thus it is *not always true*, for finite domains, that a global maximum (minimum) must also be a local maximum (minimum).

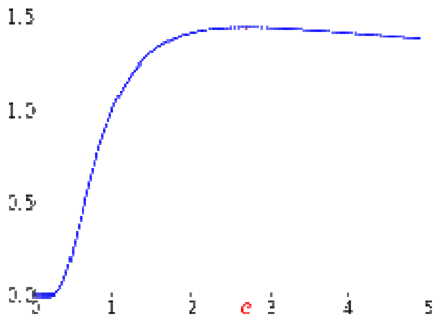
Finding functional maxima and minima

Finding global maxima and minima is the goal of mathematical optimization. If a function is continuous on a closed interval, then by the extreme value theorem global maxima and minima exist. Furthermore, a global maximum (or minimum) either must be a local maximum (or minimum) in the interior of the domain, or must lie on the boundary of the domain. So a method of finding a global maximum (or minimum) is to look at all the local maxima (or minima) in the interior, and also look at the maxima (or minima) of the points on the boundary; and take the biggest (or smallest) one.

Local extrema can be found by Fermat's theorem, which states that they must occur at critical points. One can distinguish whether a critical point is a local maximum or local minimum by using the first derivative test or second derivative test.

For any function that is defined piecewise, one finds a maxima (or minima) by finding the maximum (or minimum) of each piece separately; and then seeing which one is biggest (or smallest).

Examples



The global maximum of $\sqrt[x]{x}$ occurs at $x = e$.

- The function x^2 has a unique global minimum at $x = 0$.
- The function x^3 has no global minima or maxima. Although the first derivative ($3x^2$) is 0 at $x = 0$, this is an inflection point.
- The function $\sqrt[x]{x}$ has a unique global maximum at $x = e$. (See figure at right)
- The function x^{-x} has a unique global maximum over the positive real numbers at $x = 1/e$.
- The function $x^3/3 - x$ has first derivative $x^2 - 1$ and second derivative $2x$. Setting the first derivative to 0 and solving for x gives stationary points at −1 and +1. From the sign of the second derivative we can see that −1 is a local maximum and +1 is a local minimum. Note that this function has no global maximum or minimum.

- The function $|x|$ has a global minimum at $x = 0$ that cannot be found by taking derivatives, because the derivative does not exist at $x = 0$.
- The function $\cos(x)$ has infinitely many global maxima at 0, $\pm 2\pi$, $\pm 4\pi$, …, and infinitely many global minima at $\pm\pi$, $\pm 3\pi$, ….
- The function $2\cos(x) - x$ has infinitely many local maxima and minima, but no global maximum or minimum.
- The function $\cos(3\pi x)/x$ with $0.1 \le x \le 1.1$ has a global maximum at $x = 0.1$ (a boundary), a global minimum near $x = 0.3$, a local maximum near $x = 0.6$, and a local minimum near $x = 1.0$. (See figure at top of page.)
- The function $x^3 + 3x^2 - 2x + 1$ defined over the closed interval (segment) $[-4,2]$ has two extrema: one local maximum at $x = -1-\sqrt{15}/3$, one local minimum at $x = -1+\sqrt{15}/3$, a global maximum at $x = 2$ and a global minimum at $x = -4$.

Functions of more than one variable

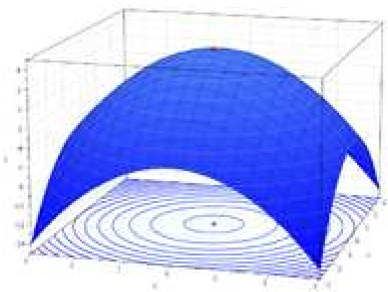
Second partial derivative test

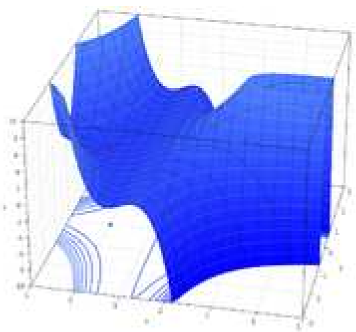
For functions of more than one variable, similar conditions apply. For example, in the (enlargeable) figure at the right, the necessary conditions for a *local* maximum are similar to those of a function with only one variable. The first partial derivatives as to z (the variable to be maximized) are zero at the maximum (the glowing dot on top in the figure). The second partial derivatives are negative. These are only necessary, not sufficient, conditions for a local maximum because of the possibility of a saddle point. For use of these conditions to solve for a maximum, the function z must also be differentiable throughout. The second partial derivative test can help classify the point as a relative maximum or relative minimum.

In contrast, there are substantial differences between functions of one variable and functions of more than one variable in the identification of global extrema. For example, if a bounded differentiable function f defined on a closed interval in the real line has a single critical point, which is a local minimum, then it is also a global minimum (use the intermediate value theorem and Rolle's theorem to prove this by reductio ad absurdum). In two and more dimensions, this argument fails, as the function

$$f(x,y) = x^2 + y^2(1-x)^3, \qquad x,y \in \mathbb{R},$$

shows. Its only critical point is at (0,0), which is a local minimum with $f(0,0) = 0$. However, it cannot be a global one, because $f(4,1) = -11$.





The global maximum is the point at the top Counterexample

In relation to sets

Maxima and minima are more generally defined for sets. In general, if an ordered set *S* has a greatest element *m*, *m* is a maximal element. Furthermore, if *S* is a subset of an ordered set *T* and *m* is the greatest element of *S* with respect to order induced by *T*, *m* is a least upper bound of *S* in *T*. The similar result holds for least element, minimal element and greatest lower bound.

In the case of a general partial order, the **least element** (smaller than all other) should not be confused with a **minimal element** (nothing is smaller). Likewise, a greatest element of a partially ordered set (poset) is an upper bound of the set which is contained within the set, whereas a **maximal element** *m* of a poset *A* is an element of *A* such that if $m \le b$ (for any *b* in *A*) then $m = b$. Any least element or greatest element of a poset is unique, but a poset can have several minimal or maximal elements. If a poset has more than one maximal element, then these elements will not be mutually comparable.

In a totally ordered set, or *chain*, all elements are mutually comparable, so such a set can have at most one minimal element and at most one maximal element. Then, due to mutual comparability, the minimal element will also be the least element and the maximal element will also be the greatest element. Thus in a totally ordered set we can simply use the terms ***minimum*** and ***maximum***. If a chain is finite then it will always have a maximum and a minimum. If a chain is infinite then it need not have a maximum or a minimum. For example, the set of natural numbers has no maximum, though it has a minimum. If an infinite chain *S* is bounded, then the closure *Cl(S)* of the set occasionally has a minimum and a maximum, in such case they are called the **greatest lower bound** and the **least upper bound** of the set *S*, respectively.

TAYLOR SERIES

The Maclaurin series for any polynomial is the polynomial itself.

The Maclaurin series for $(1 - x)^{-1}$ for |x| < 1 is the geometric series

$$1 + x + x^2 + x^3 + \cdots$$

so the Taylor series for x^{-1} at $a = 1$ is

$$1 - (x-1) + (x-1)^2 - (x-1)^3 + \cdots.$$

By integrating the above Maclaurin series we find the Maclaurin series for $\log(1 - x)$, where log denotes the natural logarithm:

$$-x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \frac{1}{4}x^4 - \cdots$$

and the corresponding Taylor series for $\log(x)$ at $a = 1$ is

$$(x-1)-\frac{1}{2}(x-1)^2+\frac{1}{3}(x-1)^3-\frac{1}{4}(x-1)^4+\cdots.$$

The Taylor series for the exponential function e^x at a = 0 is

$$1+\frac{x^1}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots=1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\frac{x^5}{120}+\cdots=\sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

The above expansion holds because the derivative of e^x with respect to x is also e^x and e^0 equals 1. This leaves the terms $(x - 0)^n$ in the numerator and $n!$ in the denominator for each term in the infinite sum.

History

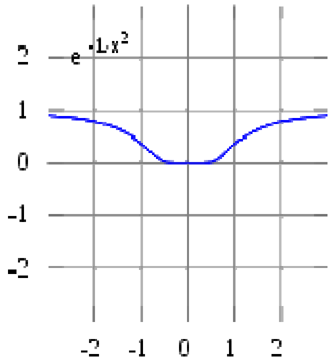
The Greek philosopher Zeno considered the problem of summing an infinite series to achieve a finite result, but rejected it as an impossibility: the result was Zeno's paradox. Later, Aristotle proposed a philosophical resolution of the paradox, but the mathematical content was apparently unresolved until taken up by Democritus and then Archimedes. It was through Archimedes's method of exhaustion that an infinite number of progressive subdivisions could be performed to achieve a finite result. Liu Hui independently employed a similar method a few centuries later

In the 14th century, the earliest examples of the use of Taylor series and closely related methods were given by Madhava of Sangamagrama Though no record of his work survives, writings of later Indian mathematicians suggest that he found a number of special cases of the Taylor series, including those for the trigonometric functions of sine, cosine, tangent, and arctangent. The Kerala school of astronomy and mathematics further expanded his works with various series expansions and rational approximations until the 16th century.

In the 17th century, James Gregory also worked in this area and published several Maclaurin series. It was not until 1715 however that a general method for constructing these series for all functions for which they exist was finally provided by Brook Taylor, after whom the series are now named.

The Maclaurin series was named after Colin Maclaurin, a professor in Edinburgh, who published the special case of the Taylor result in the 18th century.

Analytic functions



The function $\mathbf{e}^{-1/x^2}$ is not analytic at $x = 0$: the Taylor series is identically 0, although the function is not.

If $f(x)$ is given by a convergent power series in an open disc (or interval in the real line) centered at b, it is said to be analytic in this disc. Thus for x in this disc, f is given by a convergent power series

$$f(x) = \sum_{n=0}^{\infty} a_n (x - b)^n.$$

Differentiating by x the above formula n times, then setting $x=b$ gives:

$$\frac{f^{(n)}(b)}{n!} = a_n$$

and so the power series expansion agrees with the Taylor series. Thus a function is analytic in an open disc centered at b if and only if its Taylor series converges to the value of the function at each point of the disc.

If $f(x)$ is equal to its Taylor series everywhere it is called entire. The polynomials and the exponential function e^x and the trigonometric functions sine and cosine are examples of entire functions. Examples of functions that are not entire include the logarithm, the trigonometric function tangent, and its inverse arctan. For these functions the Taylor series do not converge if x is far from a. Taylor series can be used to calculate the value of an entire function in every point, if the value of the function, and of all of its derivatives, are known at a single point.

4.0 **CONCLUSION**

In this unit, you have been introduced to partial derivative in calculus and some higher order partial derivative. Clairauts theorem was stated and applied.You have been introduced to

Maxima and minima, functions of more than one variable and the relation of maxima and minima to set.

5.0 **SUMMARY**

In this unit you have studied :

Partial derivatives in calculus

Higher order partial derivative

Clairauts theorem

Maxima and manima

Taylor series of function of two variable

Analytical function

6.0 TUTOR-MARKED ASSIGNMENT

7.0 REFERENCES

1. Stewart, James (2008). *Calculus: Early Transcendentals* (6th ed.). Brooks/Cole. ISBN 0-495-01166-5.
2. Larson, Ron; Edwards, Bruce H. (2009). *Calculus* (9th ed.). Brooks/Cole. ISBN 0-547-16702-4.

   Thomas, George B.; Weir, Maurice D.; Hass, Joel (2010). *Thomas' Calculus: Early Transcendentals* (12th ed.). Addison-Wesley. ISBN 0-321-58876-2.

- Maxima and Minima From MathWorld--A Wolfram Web Resource.
- Thomas Simpson's work on Maxima and Minima at Convergence

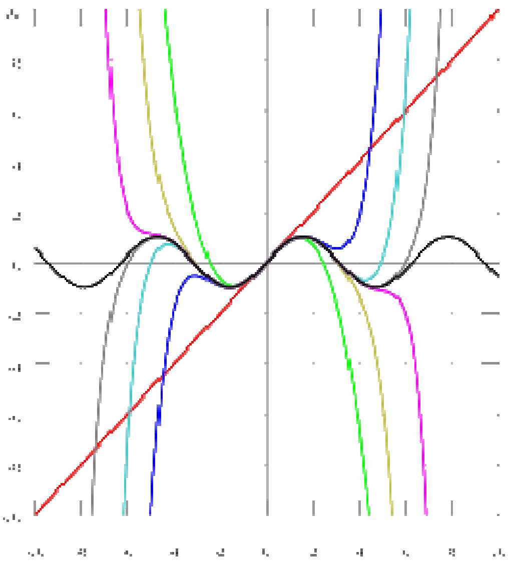
- Apostol, Tom (1967), *Calculus*, Jon Wiley & Sons, Inc., ISBN 0-471-00005-1.
- Bartle; Sherbert (2000), *Introduction to Real Analysis* (3rd ed.), John Wiley & Sons, Inc., ISBN 0-471-32148-6.
- Hörmander, L. (1976), *Linear Partial Differential Operators, Volume 1*, Springer-Verlag, ISBN 978-3540006626.
- Klein, Morris (1998), *Calculus: An Intuitive and Physical Approach*, Dover, ISBN 0-486-40453-6.
- Pedrick, George (1994), *A First Course in Analysis*, Springer-Verlag, ISBN 0-387-94108-8.
- Stromberg, Karl (1981), *Introduction to classical real analysis*, Wadsworth, Inc., ISBN 978-0534980122.
- Rudin, Walter (1987), *Real and complex analysis, 3rd ed.*, McGraw-Hill Book Company, ISBN 0-07-054234

UNIT 2 :TAYLOR SERIES OF EXPANSION FOR FUNCTIONS OF TWO VARIABLES

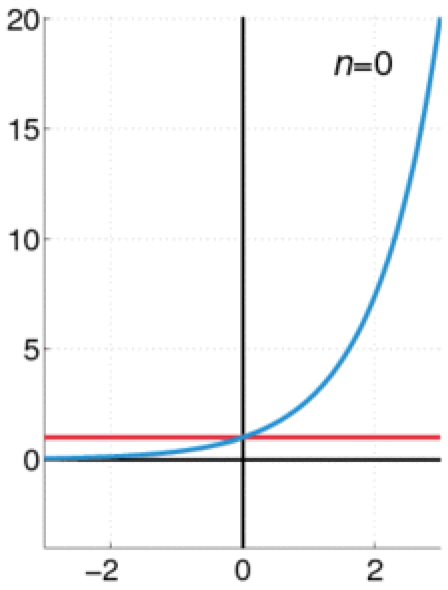
CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT
3.1 Definition of tailors series of expansion
3.2 Analytical function
3.3 Uses of taylor series for analytical functions
3.4 Approximation and convergence
3.5 List of maclaurine series of some common function
3.6 Calculation of tailors series
3.7 Taylors series in several variable
3.8 Fractional taylor series
4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

Introduction



As the degree of the Taylor polynomial rises, it approaches the correct function. This image shows sin *x* (in black) and Taylor approximations, polynomials of degree 1, 3, 5, 7, 9, 11 and



The exponential function (in blue), and the sum of the first n+1 terms of its Taylor series at 0 (in red).

In mathematics, a **Taylor series** is a representation of a function as an infinite sum of terms that are calculated from the values of the function's derivatives at a single point.

The concept of a Taylor series was formally introduced by the English mathematician Brook Taylor in 1715. If the Taylor series is centered at zero, then that series is also called a **Maclaurin series**, named after the Scottish mathematician Colin Maclaurin, who made extensive use of this special case of Taylor series in the 18th century.

It is common practice to approximate a function by using a finite number of terms of its Taylor series. Taylor's theorem gives quantitative estimates on the error in this approximation. Any finite number of initial terms of the Taylor series of a function is called a Taylor polynomial. The Taylor series of a function is the limit of that function's Taylor polynomials, provided that the limit exists. A function may not be equal to its Taylor series, even if its Taylor series converges at every point. A function that is equal to its Taylor series in an open interval (or a disc in the complex plane) is known as an analytic function.

OBJECTIVE

At the end of this unit, you should be able to :

Definition taylor series of functions of two varables

Solve problems on analytical problem

Use the taylor series to solve analytic function

Solve problems that involve approximation and convergence

The list of maclaurine series of some common functions

Calculation of taylor series

Taylors series in several variables

Fractional taylor series

3.0 MAIN CONTENT

Definition

The Taylor series of a real or complex function $f(x)$ that is infinitely differentiable in a neighborhood of a real or complex number a is the power series

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots.$$

which can be written in the more compact sigma notation as

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

where $n!$ denotes the factorial of n and $f^{(n)}(a)$ denotes the nth derivative of f evaluated at the point a. The zeroth derivative of f is defined to be f itself and $(x-a)^0$ and 0! are both defined to be 1. In the case that $a = 0$, the series is also called a Maclaurin series.

Examples

The Maclaurin series for any polynomial is the polynomial itself.

The Maclaurin series for $(1-x)^{-1}$ for $|x| < 1$ is the geometric series

$$1 + x + x^2 + x^3 + \cdots$$

so the Taylor series for x^{-1} at $a = 1$ is

$$1 - (x-1) + (x-1)^2 - (x-1)^3 + \cdots.$$

By integrating the above Maclaurin series we find the Maclaurin series for $\log(1-x)$, where log denotes the natural logarithm:

$$-x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \frac{1}{4}x^4 - \cdots$$

and the corresponding Taylor series for $\log(x)$ at $a = 1$ is

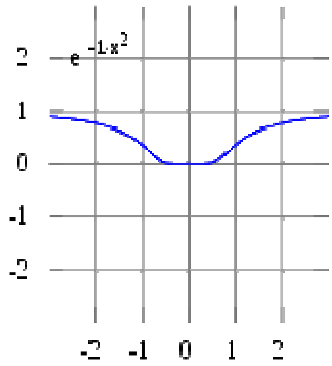
$$(x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4 + \cdots.$$

The Taylor series for the exponential function e^x at $a = 0$ is

$$1+\frac{x^1}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots = 1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}+\frac{x^5}{120}+\cdots = \sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

The above expansion holds because the derivative of e^x with respect to x is also e^x and e^0 equals 1. This leaves the terms $(x - 0)^n$ in the numerator and $n!$ in the denominator for each term in the infinite sum.

Analytic functions



The function $\mathbf{e}^{-1/x^2}$ is not analytic at $x = 0$: the Taylor series is identically 0, although the function is not.

If $f(x)$ is given by a convergent power series in an open disc (or interval in the real line) centered at b, it is said to be analytic in this disc. Thus for x in this disc, f is given by a convergent power series

$$f(x) = \sum_{n=0}^{\infty} a_n (x-b)^n.$$

Differentiating by x the above formula n times, then setting $x=b$ gives:

$$\frac{f^{(n)}(b)}{n!} = a_n$$

and so the power series expansion agrees with the Taylor series. Thus a function is analytic in an open disc centered at b if and only if its Taylor series converges to the value of the function at each point of the disc.

If $f(x)$ is equal to its Taylor series everywhere it is called entire. The polynomials and the exponential function e^x and the trigonometric functions sine and cosine are examples of entire functions. Examples of functions that are not entire include the logarithm, the trigonometric function tangent, and its inverse arctan. For these functions the Taylor series do not converge if x is far from a. Taylor series can be used to calculate the value of an entire function in every point, if the value of the function, and of all of its derivatives, are known at a single point.

Uses of the Taylor series for analytic functions include:

The partial sums (the Taylor polynomials) of the series can be used as approximations of the entire function. These approximations are good if sufficiently many terms are included.

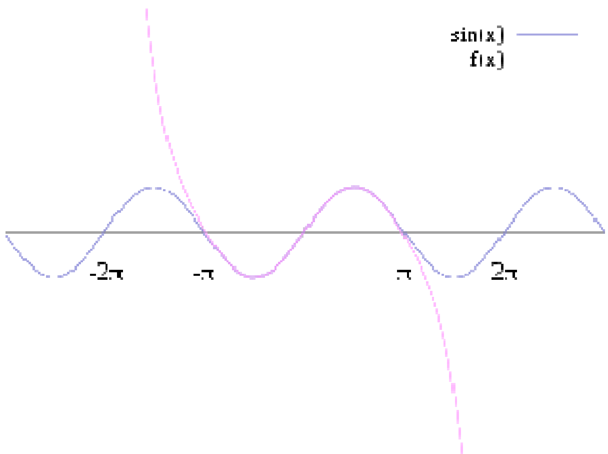
Differentiation and integration of power series can be performed term by term and is hence particularly easy.

An analytic function is uniquely extended to a holomorphic function on an open disk in the complex plane. This makes the machinery of complex analysis available.

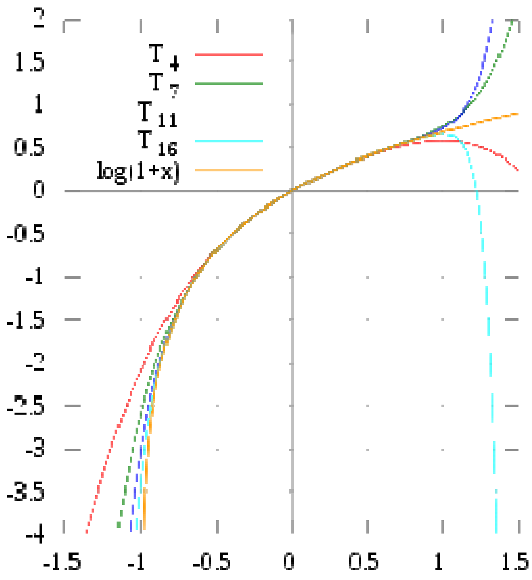
The (truncated) series can be used to compute function values numerically, (often by recasting the polynomial into the Chebyshev form and evaluating it with the Clenshaw algorithm).

Algebraic operations can be done readily on the power series representation; for instance the Euler's formula follows from Taylor series expansions for trigonometric and exponential functions. This result is of fundamental importance in such fields as harmonic analysis.

Approximation and convergence



The sine function (blue) is closely approximated by its Taylor polynomial of degree 7 (pink) for a full period centered at the origin.



The Taylor polynomials for log(*1*+*x*) only provide accurate approximations in the range $-1 < x \leq 1$. Note that, for $x > 1$, the Taylor polynomials of higher degree are ***worse*** approximations.

Pictured on the right is an accurate approximation of sin(*x*) around the point *x* = 0. The pink curve is a polynomial of degree seven:

$$\sin(x) \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}.$$

The error in this approximation is no more than $|x|^9/9!$. In particular, for $-1 < x < 1$, the error is less than 0.000003.

In contrast, also shown is a picture of the natural logarithm function log(1 + *x*) and some of its Taylor polynomials around *a* = 0. These approximations converge to the function only in the region $-1 < x \leq 1$; outside of this region the higher-degree Taylor polynomials are ***worse*** approximations for the function. This is similar to Runge's phenomenon.

The **error** incurred in approximating a function by its *n*th-degree Taylor polynomial is called the **remainder** or *residual* and is denoted by the function $R_n(x)$. Taylor's theorem can be used to obtain a bound on the size of the remainder.

In general, Taylor series need not be convergent at all. And in fact the set of functions with a convergent Taylor series is a meager set in the Fréchet space of smooth functions. Even if the Taylor series of a function *f* does converge, its limit need not in general be equal to the value of the function *f*(*x*). For example, the function

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is infinitely differentiable at $x = 0$, and has all derivatives zero there. Consequently, the Taylor series of $f(x)$ about $x = 0$ is identically zero. However, $f(x)$ is not equal to the zero function, and so it is not equal to its Taylor series around the origin.

In real analysis, this example shows that there are infinitely differentiable functions $f(x)$ whose Taylor series are *not* equal to $f(x)$ even if they converge. By contrast in complex analysis there are *no* holomorphic functions $f(z)$ whose Taylor series converges to a value different from $f(z)$. The complex function $e^{-z^{-2}}$ does not approach 0 as z approaches 0 along the imaginary axis, and its Taylor series is thus not defined there.

More generally, every sequence of real or complex numbers can appear as coefficients in the Taylor series of an infinitely differentiable function defined on the real line, a consequence of Borel's lemma (see also Non-analytic smooth function#Application to Taylor series). As a result, the radius of convergence of a Taylor series can be zero. There are even infinitely differentiable functions defined on the real line whose Taylor series have a radius of convergence 0 everywhere.[5]

Some functions cannot be written as Taylor series because they have a singularity; in these cases, one can often still achieve a series expansion if one allows also negative powers of the variable x; see Laurent series. For example, $f(x) = e^{-x^{-2}}$ can be written as a Laurent series.

There is, however, a generalization[6][7] of the Taylor series that does converge to the value of the function itself for any bounded continuous function on $(0,\infty)$, using the calculus of finite differences. Specifically, one has the following theorem, due to Einar Hille, that for any $t > 0$,

$$\lim_{h\to 0^+}\sum_{n=0}^{\infty}\frac{t^n}{n!}\frac{\Delta_h^n f(a)}{h^n} = f(a+t).$$

Here Δ_h^n is the n-th finite difference operator with step size h. The series is precisely the Taylor series, except that divided differences appear in place of differentiation: the series is formally similar to the Newton series. When the function f is analytic at a, the terms in the series converge to the terms of the Taylor series, and in this sense generalizes the usual Taylor series.

In general, for any infinite sequence a_i, the following power series identity holds:

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\Delta^n a_i = e^{-u}\sum_{j=0}^{\infty}\frac{u^j}{j!}a_{i+j}.$$

So in particular,

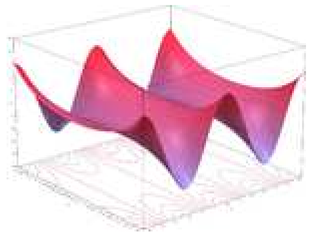
$$f(a+t) = \lim_{h\to 0^+} e^{-t/h}\sum_{j=0}^{\infty} f(a+jh)\frac{(t/h)^j}{j!}.$$

The series on the right is the expectation value of $f(a + X)$, where X is a Poisson distributed random variable that takes the value jh with probability $e^{-t/h}(t/h)^j/j!$. Hence,

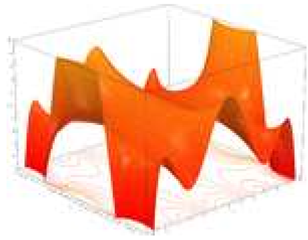
$$f(a+t) = \lim_{h\to 0+} \int_{-\infty}^{\infty} f(a+x)\,dP_{t/h,h}(x).$$

The law of large numbers implies that the identity holds.

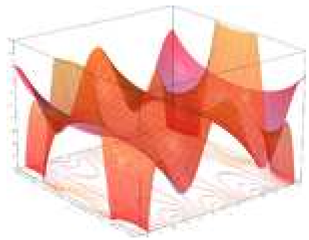
List of Maclaurin series of some common functions



The real part of the cosine function in the complex plane.



An 8th degree approximation of the cosine function in the complex plane.



The two above curves put together.

Several important Maclaurin series expansions follow.All these expansions are valid for complex arguments *x*.

Exponential function:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad \text{for all } x$$

Natural logarithm:

$$\log(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n} \text{ for } -1 \le x < 1$$

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \text{ for } -1 < x \le 1$$

Finite geometric series:

$$\frac{1-x^{m+1}}{1-x} = \sum_{n=0}^{m} x^n \quad \text{for } x \ne 1 \text{ and } m \in \mathbb{N}_0$$

Infinite geometric series:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \text{ for } |x| < 1$$

Variants of the infinite geometric series:

$$\frac{x^m}{1-x} = \sum_{n=m}^{\infty} x^n \quad \text{for } |x| < 1 \text{ and } m \in \mathbb{N}_0$$

$$\frac{x}{(1-x)^2} = \sum_{n=1}^{\infty} n x^n \quad \text{for } |x| < 1$$

$$\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1} \quad \text{for } |x| < 1$$

Square root:

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{(1-2n)(n!)^2 (4^n)} x^n = 1 + \tfrac{1}{2}x - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 - \tfrac{5}{128}x^4 + \dots \text{ for } |x| \le$$

Binomial series (includes the square root for $\alpha = 1/2$ and the infinite geometric series for $\alpha = -1$):

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \quad \text{for all } |x| < 1 \text{ and all complex } \alpha$$

with generalized binomial coefficients

$$\binom{\alpha}{n} = \prod_{k=1}^{n} \frac{\alpha - k + 1}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}$$

Trigonometric functions:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \text{ for all } x$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \text{ for all } x$$

$$\tan x = \sum_{n=1}^{\infty} \frac{B_{2n}(-4)^n(1-4^n)}{(2n)!} x^{2n-1} = x + \frac{x^3}{3} + \frac{2x^5}{15} + \cdots \text{ for } |x| < \frac{\pi}{2}$$

where the B_s are Bernoulli numbers.

$$\sec x = \sum_{n=0}^{\infty} \frac{(-1)^n E_{2n}}{(2n)!} x^{2n} \text{ for } |x| < \frac{\pi}{2}$$

$$\arcsin x = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)^2 (2n+1)} x^{2n+1} \text{ for } |x| \le 1$$

$$\arccos x = \frac{\pi}{2} - \arcsin x = \frac{\pi}{2} - \sum_{n=0}^{\infty} \frac{(2n)!}{4^n (n!)^2 (2n+1)} x^{2n+1} \text{ for } |x| \le 1$$

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \text{ for } |x| \le 1$$

Hyperbolic functions:

$$\sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \text{ for all } x$$

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \text{ for all } x$$

$$\tanh x = \sum_{n=1}^{\infty} \frac{B_{2n} 4^n (4^n - 1)}{(2n)!} x^{2n-1} = x - \frac{1}{3}x^3 + \frac{2}{15}x^5 - \frac{17}{315}x^7 + \cdots \text{ for } |x| < \frac{\pi}{2}$$

$$\operatorname{arsinh}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{4^n (n!)^2 (2n+1)} x^{2n+1} \text{ for } |x| \le 1$$

$$\operatorname{artanh}(x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1} \text{ for } |x| < 1$$

Lambert's W function:

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n \text{ for } |x| < \frac{1}{e}$$

The numbers B_k appearing in the *summation* expansions of tan(x) and tanh(x) are the Bernoulli numbers. The E_k in te expansion of sec(x) are Euler numbers.

Calculation of Taylor series

Several methods exist for the calculation of Taylor series of a large number of functions. One can attempt to use the Taylor series as-is and generalize the form of the coefficients, or one can use manipulations such as substitution, multiplication or division, addition or subtraction of standard Taylor series to construct the Taylor series of a function, by virtue of Taylor series being power series. In some cases, one can also derive the Taylor series by repeatedly applying integration by parts. Particularly convenient is the use of computer algebra systems to calculate Taylor series.

First example

Compute the 7th degree Maclaurin polynomial for the function

$$f(x) = \log \cos x, \quad x \in (-\pi/2, \pi/2).$$

First, rewrite the function as

$$f(x) = \log(1 + (\cos x - 1)).$$

We have for the natural logarithm (by using the big O notation)

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4)$$

and for the cosine function

$$\cos x - 1 = -\frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + O(x^8)$$

The latter series expansion has a zero constant term, which enables us to substitute the second series into the first one and to easily omit terms of higher order than the 7th degree by using the big O notation

$$
\begin{aligned}
f(x) &= \log(1+(\cos x-1)) \\
&= (\cos x-1)-\frac{1}{2}(\cos x-1)^2+\frac{1}{3}(\cos x-1)^3+O((\cos x-1)^4) \\
&= \left(-\frac{x^2}{2}+\frac{x^4}{24}-\frac{x^6}{720}+O(x^8)\right)-\frac{1}{2}\left(-\frac{x^2}{2}+\frac{x^4}{24}+O(x^6)\right)^2+\frac{1}{3}\left(-\frac{x^2}{2}+O
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{x^2}{2}+\frac{x^4}{24}-\frac{x^6}{720}-\frac{x^4}{8}+\frac{x^6}{48}-\frac{x^6}{24}+O(x^8) \\
&= -\frac{x^2}{2}-\frac{x^4}{12}-\frac{x^6}{45}+O(x^8).
\end{aligned}
$$

Since the cosine is an even function, the coefficients for all the odd powers x, x^3, x^5, x^7, ... have to be zero.

Second example

Suppose we want the Taylor series at 0 of the function

$$g(x)=\frac{e^x}{\cos x}.$$

We have for the exponential function

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots$$

and, as in the first example,

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots$$

Assume the power series is

$$\frac{e^x}{\cos x}=c_0+c_1x+c_2x^2+c_3x^3+\cdots$$

Then multiplication with the denominator and substitution of the series of the cosine yields

$$
\begin{aligned}
e^x &= (c_0+c_1x+c_2x^2+c_3x^3+\cdots)\cos x \\
&= \left(c_0+c_1x+c_2x^2+c_3x^3+c_4x^4+\cdots\right)\left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots\right) \\
&= c_0-\frac{c_0}{2}x^2+\frac{c_0}{4!}x^4+c_1x-\frac{c_1}{2}x^3+\frac{c_1}{4!}x^5+c_2x^2-\frac{c_2}{2}x^4+\frac{c_2}{4!}x^6+c_3x^3-\frac{c_3}{2}x^5+
\end{aligned}
$$

Collecting the terms up to fourth order yields

$$= c_0 + c_1 x + \left(c_2 - \frac{c_0}{2}\right) x^2 + \left(c_3 - \frac{c_1}{2}\right) x^3 + \left(c_4 + \frac{c_0}{4!} - \frac{c_2}{2}\right) x^4 + \cdots$$

Comparing coefficients with the above series of the exponential function yields the desired Taylor series

$$\frac{e^x}{\cos x} = 1 + x + x^2 + \frac{2x^3}{3} + \frac{x^4}{2} + \cdots.$$

Comparing coefficients with the above series of the exponential function yields the desired Taylor series

$$\frac{e^x}{\cos x} = 1 + x + x^2 + \frac{2x^3}{3} + \frac{x^4}{2} + \cdots.$$

Third example

Here we use a method called "Indirect Expansion" to expand the given function. This method uses the known function of Taylor series for expansion.

Q: Expand the following function as a power series of x

$(1 + x)e^x$.

We know the Taylor series of function e^x is:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots, -\infty < x < +\infty$$

Thus,

$$\begin{aligned}(1+x)e^x = e^x + xe^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n!} = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n!} \\ &= 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} + \sum_{n=1}^{\infty} \frac{x^n}{(n-1)!} = 1 + \sum_{n=1}^{\infty} \left(\frac{1}{n!} + \frac{1}{(n-1)!}\right) x^n \\ &= 1 + \sum_{n=1}^{\infty} \frac{n+1}{n!} x^n, -\infty < x < +\infty \\ &= \sum_{n=0}^{\infty} \frac{n+1}{n!} x^n\end{aligned}$$

.

Taylor series in several variables

The Taylor series may also be generalized to functions of more than one variable with

$$T(x_1,\ldots,x_d) = \\ = \sum_{n_1=0}^{\infty}\cdots\sum_{n_d=0}^{\infty}\frac{(x_1-a_1)^{n_1}\cdots(x_d-a_d)^{n_d}}{n_1!\cdots n_d!}\left(\frac{\partial^{n_1+\cdots+n_d}f}{\partial x_1^{n_1}\cdots\partial x_d^{n_d}}\right)(a_1,\ldots,a_d).$$

For example, for a function that depends on two variables, *x* and *y*, the Taylor series to second order about the point (*a*, *b*) is:

$$f(x,y)\approx f(a,b)+(x-a)\,f_x(a,b)+(y-b)\,f_y(a,b) \\ +\frac{1}{2!}\left[(x-a)^2\,f_{xx}(a,b)+2(x-a)(y-b)\,f_{xy}(a,b)+(y-b)^2\,f_{yy}(a,b)\right],$$

where the subscripts denote the respective partial derivatives.

A second-order Taylor series expansion of a scalar-valued function of more than one variable can be written compactly as

$$T(\mathbf{x})=f(\mathbf{a})+(\mathbf{x}-\mathbf{a})^T\mathrm{D}f(\mathbf{a})+\frac{1}{2!}(\mathbf{x}-\mathbf{a})^T\left\{\mathrm{D}^2f(\mathbf{a})\right\}(\mathbf{x}-\mathbf{a})+\cdots,$$

ere $Df(\mathbf{a})$is the gradient of $f$evaluated at $\mathbf{x}=\mathbf{a}$and $D^2f(\mathbf{a})$is the Hessian matrix. Applying the multi-index notation the Taylor series for several variables becomes

$$T(\mathbf{x})=\sum_{|\alpha|\ge 0}\frac{(\mathbf{x}-\mathbf{a})^{\alpha}}{\alpha!}\,(\partial^{\alpha}f)(\mathbf{a}),$$

which is to be understood as a still more abbreviated multi-index version of the first equation of this paragraph, again in full analogy to the single variable case.

Example

Second-order Taylor series approximation (in gray) of a function $f(x,y) = e^x \log(1 + y)$ around origin.

Compute a second-order Taylor series expansion around point $(a,b) = (0,0)$ of a function

$$f(x,y) = e^x \log(1+y).$$

Firstly, we compute all partial derivatives we need

$$f_x(a,b) = e^x \log(1+y)\Big|_{(x,y)=(0,0)} = 0\,,$$

$$f_y(a,b) = \frac{e^x}{1+y}\Big|_{(x,y)=(0,0)} = 1\,,$$

$$f_{xx}(a,b) = e^x \log(1+y)\Big|_{(x,y)=(0,0)} = 0\,,$$

$$f_{yy}(a,b) = -\frac{e^x}{(1+y)^2}\Big|_{(x,y)=(0,0)} = -1\,,$$

$$f_{xy}(a,b) = f_{yx}(a,b) = \frac{e^x}{1+y}\Big|_{(x,y)=(0,0)} = 1.$$

The Taylor series is

$$T(x,y) = f(a,b) + (x-a)\, f_x(a,b) + (y-b)\, f_y(a,b) + \frac{1}{2!}\left[(x-a)^2\, f_{xx}(a,b) + 2(x-a)(y-b)\, f_{xy}(a,b) + (y-b)^2\, f_{yy}(a$$

which in this case becomes

$$T(x,y) = 0 + 0(x-0) + 1(y-0) + \frac{1}{2}\Big[0(x-0)^2 + 2(x-0)(y-0) + (-1)(y$$
$$= y + xy - \frac{y^2}{2} + \cdots\,.$$

Since $\log(1 + y)$ is analytic in $|y| < 1$, we have

$$e^x \log(1+y) = y + xy - \frac{y^2}{2} + \cdots$$

for $|y| < 1$.

Fractional Taylor series

With the emergence of fractional calculus, a natural question arises about what the Taylor Series expansion would be. Odibat and Shawagfeh answered this in 2007. By using the

Caputo fractional derivative, $0 < \alpha < 1$, and $x+$ indicating the limit as we approach x from the right, the fractional Taylor series can be written as

$$f(x+\Delta x) = f(x) + D_x^\alpha f(x+)\frac{(\Delta x)^\alpha}{\Gamma(\alpha+1)} + D_x^\alpha D_x^\alpha f(x+)\frac{(\Delta x)^{2\alpha}}{\Gamma(2\alpha+1)} + \cdots .$$

4.0 **CONCLUSION**

In this unit, you have defined tailors series of function of two variable. You have studied analytical function and have used tailors series to solve problem s that involve analytical functions. You have studied approximation and convergence. You have also studied the list of maclaurine series of some common functions and have done some calculation of tailors series. You have also studied tailors in several variables and the fractional taylor series.

5.0 SUMMARY

In this unit, you have studied the following :

Definition taylor series of functions of two varables

Solve problems on analytical problem

Use the taylor series to solve analytic function

Solve problems that involve approximation and convergence

The list of maclaurine series of some common functions

Calculation of taylor series

Taylors series in several variables

Fractional taylor series

TUTOR – MARKED ASSIGNMENT

1.Use the tailor series to expand F(z) = $\frac{1}{z+1}$ about the point z = 1 ,and find the values of z for which the expansion is valid.

2.Use the tailor series to expand F(x) = $\frac{1}{x+2}$ about the point x = 1 ,and find the values of z for which the expansion is valid.

3.Use the tailor series to expand F(x) = $\frac{1}{(x-2)^2}$ about the point x = 2 ,and find the values of z for which the expansion is valid.

4.Use the tailor series to expand F(x) = $\frac{1}{(x+4)^2}$ about the point x = 2 ,and find the values of z for which the expansion is valid.

5.Use the tailor series to expand F(b) = $\frac{2}{(b+2)^3}$ about the point b = 1 ,and find the values of z for which the expansion is valid.

REFERENCES

Abramowitz, Milton; Stegun, Irene A. (1970), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* New York: Dover Publications, Ninth printing

Thomas, George B. Jr.; Finney, Ross L. (1996), *Calculus and Analytic Geometry (9th ed.),* Addison Wesley, ISBN 0-201-53174-7

Greenberg, Michael (1998), *Advanced Engineering Mathematics (2nd ed.),* Prentice Hall, ISBN 0-13-321431

Abramowitz, M. and Stegun, I. A. (Eds.). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing.* New York: Dover, p. 880, 1972.

Arfken, G. "Taylor's Expansion." §5.6 in *Mathematical Methods for Physicists, 3rd ed.* Orlando, FL: Academic Press, pp. 303-313, 1985.

Askey, R. and Haimo, D. T. "Similarities between Fourier and Power Series." *Amer. Math. Monthly* **103**, 297-304, 1996.

Comtet, L. "Calcul pratique des coefficients de Taylor d'une fonction algébrique." *Enseign. Math.* **10**, 267-270, 1964.

Morse, P. M. and Feshbach, H. "Derivatives of Analytic Functions, Taylor and Laurent Series." §4.3 in *Methods of Theoretical Physics, Part I.* New York: McGraw-Hill, pp. 374-398, 1953.

Whittaker, E. T. and Watson, G. N. "Forms of the Remainder in Taylor's Series." §5.41 in *A Course in Modern Analysis, 4th ed.* Cambridge, England: Cambridge University Press, pp. 95-96, 1990.

UNIT 3 : APPLICATIONS OF TAYLOR SERIES

CONTENT

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT

3.1 Evaluating definite integrals

3.2 Understanding the asymptotic behaviour
3.3 Understanding the growth of functions
3.4 Solving differential equations

4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

We started studying Taylor Series because we said that polynomial functions are easy and that if we could find a way of representing complicated functions as series ("infinite polynomials") then maybe some properties of functions would be easy to study too. In this section, we'll show you a few ways in Taylor series can make life easy.

2.0 OBJECTIVES

At the end of this unit, you should be able to :

Evaluate definite integrals with taylors series

Understand the asymptotic behaviour with taylors series

Understand the growth of functions with taylors series

Solve differential equations with taylors series

3.0 MAIN CONTENT

Evaluating definite integrals

Remember that we've said that some functions have no antiderivative which can be expressed in terms of familiar functions. This makes evaluating definite integrals of these functions difficult because the Fundamental Theorem of Calculus cannot be used. However, if we have a series representation of a function, we can often times use that to evaluate a definite integral.

Here is an example. Suppose we want to evaluate the definite integral

$$\int_0^1 \sin(x^2)\, dx$$

The integrand has no antiderivative expressible in terms of familiar functions. However, we know how to find its Taylor series: we know that

$$\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \ldots$$

Now if we substitute $t = x^2$, we have

$$\sin(x^2) = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{10!} - \frac{x^{14}}{14!} + \ldots$$

In spite of the fact that we cannot antidifferentiate the function, we can antidifferentiate the Taylor series:

$$\begin{aligned}\int_0^1 \sin(x^2)\, dx &= \int_0^1 (x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \frac{x^{14}}{7!} + \ldots)\, dx \\ &= (\frac{x^3}{3} - \frac{x^7}{7 \cdot 3!} + \frac{x^{11}}{11 \cdot 5!} - \frac{x^{15}}{15 \cdot 7!} + \ldots)|_0^1 \\ &= \frac{1}{3} - \frac{1}{7 \cdot 3!} + \frac{1}{11 \cdot 5!} - \frac{1}{15 \cdot 7!} + \ldots\end{aligned}$$

Notice that this is an alternating series so we know that it converges. If we add up the first four terms, the pattern becomes clear: the series converges to **0.31026.**

Understanding asymptotic behaviour

Sometimes, a Taylor series can tell us useful information about how a function behaves in an important part of its domain. Here is an example which will demonstrate.

A famous fact from electricity and magnetism says that a charge **q** generates an electric field whose strength is inversely proportional to the square of the distance from the charge. That is, at a distance **r** away from the charge, the electric field is

$$E = \frac{kq}{r^2}$$

where **k** is some constant of proportionality.

Often times an electric charge is accompanied by an equal and opposite charge nearby. Such an object is called an electric dipole. To describe this, we will put a charge **q** at the point $x = d$ and a charge **-q** at $x = -d$.

Along the **x** axis, the strength of the electric fields is the sum of the electric fields from each of the two charges. In particular,

$$E = \frac{kq}{(x-d)^2} - \frac{kq}{(x+d)^2}$$

If we are interested in the electric field far away from the dipole, we can consider what happens for values of **x** much larger than **d.** We will use a Taylor series to study the behaviour in this region.

$$E = \frac{kq}{(x-d)^2} - \frac{kq}{(x+d)^2} = \frac{kq}{x^2(1-\frac{d}{x})^2} - \frac{kq}{x^2(1+\frac{d}{x})^2}$$

Remember that the geometric series has the form

$$\frac{1}{1-u} = 1 + u + u^2 + u^3 + u^4 + \dots$$

If we differentiate this series, we obtain

$$\frac{1}{(1-u)^2} = 1 + 2u + 3u^2 + 4u^3 + \dots$$

Into this expression, we can substitute $u = \frac{d}{x}$ to obtain

In the same way, if we substitute $u = -\frac{d}{x}$, we have

$$\frac{1}{(1+\frac{d}{x})^2} = 1 - \frac{2d}{x} + \frac{3d^2}{x^2} - \frac{4d^3}{x^3} + \dots$$

Now putting this together gives

$$\begin{aligned} E &= \frac{kq}{x^2(1-\frac{d}{x})^2} - \frac{kq}{x^2(1+\frac{d}{x})^2} \\ &= \frac{kq}{x^2}[(1 + \frac{2d}{x} + \frac{3d^2}{x^2} + \frac{4d^3}{x^3} + \dots) - (1 - \frac{2d}{x} + \frac{3d^2}{x^2} - \frac{4d^3}{x^3} + \dots)] \\ &= \frac{kq}{x^2}[\frac{4d}{x} + \frac{8d^3}{x^3} + \dots] \\ &\approx \frac{4dq}{x^3} \end{aligned}$$

In other words, far away from the dipole where **x** is very large, we see that the electric field strength is proportional to the inverse *cube* of the distance. The two charges partially cancel one another out to produce a weaker electric field at a distance.

Understanding the growth of functions

This example is similar is spirit to the previous one. Several times in this course, we have used the fact that exponentials grow much more rapidly than polynomials. We recorded this by saying that

$$\lim_{n\to\infty} \frac{e^x}{x^n} = \infty$$

for any exponent **n** . Let's think about this for a minute because it is an important property of exponentials. The ratio $\frac{e^x}{x^n}$ is measuring how large the exponential is compared to the polynomial. If this ratio was very small, we would conclude that the polynomial is larger than the exponential. But if the ratio is large, we would conclude that the exponential is much larger than the polynomial. The fact that this ratio becomes arbitrarily large means that the exponential becomes larger than the polynomial by a factor which is as large as we would like. This is what we mean when we say "an exponential grows faster than a polynomial."

To see why this relationship holds, we can write down the Taylor series for e^x.

To see why this relationship holds, we can write down the Taylor series for e^x.

$$\begin{aligned} \frac{e^x}{x^n} &= \frac{1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \frac{x^n}{n!} + \frac{x^{n+1}}{(n+1)!} + \ldots}{x^n} \\ &= \frac{1}{x^n} + \frac{1}{x^{n-1}} + \ldots + \frac{1}{n!} + \frac{x}{(n+1)!} + \ldots \\ &> \frac{x}{(n+1)!} \end{aligned}$$

Notice that this last term becomes arbitrarily large as $x \to \infty$. That implies that the ratio we are interested in does as well:

$$\lim_{x\to\infty} \frac{e^x}{x^n} = \infty$$

Basically, the exponential e^x grows faster than any polynomial because it behaves like an infinite polynomial whose coefficients are all positive.

Solving differential equations

Some differential equations cannot be solved in terms of familiar functions (just as some functions do not have antiderivatives which can be expressed in terms of familiar functions).

However, Taylor series can come to the rescue again. Here we will present two examples to give you the idea.

Example 1: We will solve the initial value problem

$$\begin{aligned} \frac{dy}{dx} &= y \\ y(0) &= 1 \end{aligned}$$

Of course, we know that the solution is $y(x) = e^x$, but we will see how to discover this in a different way. First, we will write out the solution in terms of its Taylor series:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots$$

Since this function satisfies the condition $y(0) = 1$, we must have $y(0) = a_0 = 1$.

We also have

$$\frac{dy}{dx} = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \dots$$

Since the differential equation says that $\frac{dy}{dx} = y$, we can equate these two Taylor series:

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots = a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \dots$$

If we now equate the coefficients, we obtain:

$$\begin{aligned} &a_0 = a_1 = 1, \quad && a_1 = 1 \\ &a_1 = 2a_2, && a_2 = \tfrac{a_1}{2} = \tfrac{1}{2} \\ &a_2 = 3a_3, && a_3 = \tfrac{a_2}{3} = \tfrac{1}{2\cdot 3} \\ &a_3 = 4a_4, && a_4 = \tfrac{a_3}{4} = \tfrac{1}{2\cdot 3\cdot 4} \\ &a_{n-1} = na_n, && a_n = \tfrac{a_{n-1}}{n} = \tfrac{1}{1\cdot 2\cdot 3 \dots n} = \tfrac{1}{n!} \end{aligned}$$

This means that $y = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots = e^x$ as we expect.

Of course, this is an intial value problem we know how to solve. The real value of this method is in studying initial value problems that we do not know how to solve.

Example 2: Here we will study *Airy's equation* with initial conditions:

$$\begin{aligned} y'' &= xy \\ y(0) &= 1 \\ y'(0) &= 0 \end{aligned}$$

This equation is important in optics. In fact, it explains why a rainbow appears the way in which it does! As before, we will write the solution as a series:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + \ldots$$

Since we have the initial conditions, $y(0) = a_0 = 1$ and $y'(0) = a_1 = 0$.

Now we can write down the derivatives:

$$\begin{aligned} y' &= a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + 5a_5 x^4 + \ldots \\ y'' &= 2a_2 + 2 \cdot 3x + 3 \cdot 4x^2 + 4 \cdot 5x^3 + \ldots \end{aligned}$$

The equation then gives

$$\begin{aligned} y'' &= xy \\ 2a_2 + 2 \cdot 3a_3 x + 3 \cdot 4a_4 x^2 + 4 \cdot 5a_5 x^3 + \ldots &= x(a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots) \\ 2a_2 + 2 \cdot 3a_3 x + 3 \cdot 4a_4 x^2 + 4 \cdot 5a_5 x^3 + \ldots &= a_0 x + a_1 x^2 + a_2 x^3 + a_3 x^4 + \ldots \end{aligned}$$

Again, we can equate the coefficients of **x** to obtain

$$\begin{array}{ll} 2a_2 = 0 & a_2 = 0 \\ 2 \cdot 3a_3 = a_0 & a_3 = \frac{1}{2 \cdot 3} \\ 3 \cdot 4a_4 = a_1 & a_4 = 0 \\ 4 \cdot 5a_5 = a_2 & a_5 = 0 \\ 5 \cdot 6a_6 = a_3 & a_6 = \frac{1}{2 \cdot 3 \cdot 5 \cdot 6} \end{array}$$

This gives us the first few terms of the solution:

$$y = 1 + \frac{x^3}{2 \cdot 3} + \frac{x^6}{2 \cdot 3 \cdot 5 \cdot 6} + \ldots$$

If we continue in this way, we can write down many terms of the series (perhaps you see the pattern already?) and then draw a graph of the solution. This looks like this:

Notice that the solution oscillates to the left of the origin and grows like an exponential to the right of the origin. Can you explain this by looking at the differential equation

4.0 CONCLUSION

In this unit, you have been introduced to the application of taylors series and
some basic ways of using taylors series such as the evaluating of definite integrals, understanding the asymptotic behaviour, understanding the growth of functions and solving differential equations. Some examples where used to illustrate the applications.

5 SUMMARY

Having gone through this unit, you now know that;
In this section, we show you ways in which Taylor series can make life easy :

i. In evaluating definite integrals , we used series representation of a function to evaluate some functions that have no antiderivative .

Suppose we want to evaluate the definite integral

$$\int_0^1 \sin(x^2)\ dx$$

The integrand has no antiderivative expressible in terms of familiar functions. However, we know how to find its Taylor series: we know that

$$\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \ldots$$

Now if we substitute $t = x^2$, we have

$$\sin(x^2) = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{10!} - \frac{x^{14}}{14!} + \ldots$$

In spite of the fact that we cannot antidifferentiate the function, we can antidifferentiate the Taylor series:

(ii) We used taylors series to understand asymptotic behaviour of functions that behave in the important part of the domain . And some examples are shown to demonstrate ,

(iii) Taylors series is used to understand the growth of functions. Because we know the fact that exponentials grow much more rapidly than polynomials. We recorded this by saying that

$$\lim_{n\to\infty} \frac{e^x}{x^n} = \infty$$

for any exponent **n** .
(iv) We used taylors series to solve problems which could not be solved ordinarily through differential equations.

Tutor-Marked Assignment

1. Compute a second-order Taylor series expansion around point $(a,b) = (0,0)$ of a function

F(x,y)= e^x log(2+y)

2. Show that the taylor series expansion of f(x,y) = e^{xy} about the point (2,3) .

7.0 REFERENCE

Arfken and Weber, Mathematical Methods for Physicists, 6th Edition, 352-354, Academic press 200

[2] http://www.ugrad.math.ubc.ca/coursedoc/math101 /notes/series/appsTaylor.html September,29 2008

[3] Broadhead MK (Broadhead, Michael K.) , geophysical prospecting 56, 5, 729-735 SEP 2008

[4] Guyenne P (Guyenne, Philippe), Nicholls DP (Nicholls, David P.), siam journal on scientific computing 30, 1, 81-101, 2007

[5] Popa C (Popa, Cosmin), IEEE transactions on very large scale integration (VLSI) systems,16, 3,318-321, MAR 2008

[6] Janssen AJEM (Janssen, A. J. E. M.), van Leeuwaarden JSH (van Leeuwaarden, J. S. H.) stochastic processes and their applications 117 ,12, 1928-1959, DEC 2007

[7] Sahu S (Sahu, Sunil), Baker AJ (Baker, A. J.) Source: international journal for numerical methods in fluids 55, 8, 737-783, NOV 20, 2007

MODULE 7 MAXIMA AND MINIMA OF FUNCTIONS OF SEVERAL VARIABLES, STATIONARY POINT, LAGRANGE'S METHOD OF MULTIPLIERS

Unit 1:MAXIMISATION AND MINIMISATION OF FUNCTIONS OF SEVERAL VARIABLES

CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT
3.1 Recognise problems on maximum and minimum functions of several variables
3.2 Necessary condition for a maxima or minima function of several variable
3.3 Sufficient condition for a maxima or minima function of several variable
3.4 Maxima and minima of functions subject to constraints
3.5 Method of finding maxima and minima of functions subject to constraints
3.6Identify the different types of examples of maxima and minima functions of several variables
3.7Solve problems on maxima and minima functions of several variables
4.0CONCLUSION
5.0SUMMARY
6.0TUTOR-MARKED ASSIGNMENT
7.0REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

Def. Stationary (or critical) point. For a function y = f(x) of a single variable, a stationary (or critical) point is a point at which dy/dx = 0; for a function $u = f(x_1, x_2, \ldots, x_n)$ of n variables it is a point at which

$$1) \qquad \frac{\partial u}{\partial x_1} = 0 \qquad \frac{\partial u}{\partial x_2} = 0 \quad \cdots \quad \frac{\partial u}{\partial x_n} = 0 .$$

In the case of a function y = f(x) of a single variable, a stationary point corresponds to a point on the curve at which the tangent to the curve is horizontal. In the case of a function y = f(x, y) of two variables a stationary point corresponds to a point on the surface at which the tangent plane to the surface is horizontal.

In the case of a function y = f(x) of a single variable, a stationary point can be any of the following three: a maximum point, a minimum point or an inflection point. For a function y = f(x, y) of two variables, a stationary point can be a maximum point, a minimum point or a saddle point. For a function of n variables it can be a maximum point, a minimum point or a point that is analogous to an inflection or saddle point.

2.0 OBJECTIVE

At the end of this unit, you should be able to :

- recognise problems on maximum and minimum functions of several variables
- know the necessary condition for a maxima or minima function of several variable
- know the Sufficient condition for a maxima or minima function of several variable
- identify the maxima and minima of functions subject to constraints
- know the method of finding maxima and minima of functions subject to constraints
- identify the different types of examples of maxima and minima functions of several variables
- solve problems on maxima and minima functions of several variables

Maxima and minima of functions of several variables.
A function f(x, y) of two independent variables has a **maximum** at a point (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0). Such a function has a **minimum** at a point (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0).

A function $f(x_1, x_2, \dots, x_n)$ of n independent variables has a **maximum** at a point $(x_1', x_2', \dots, x_n')$ if $f(x_1', x_2', \dots, x_n') \geq f(x_1, x_2, \dots, x_n)$ at all points in the neighborhood of $(x_1', x_2', \dots, x_n')$. Such a function has a **minimum** at a point $(x_1', x_2', \dots, x_n')$ if $f(x_1', x_2', \dots, x_n') \leq f(x_1, x_2, \dots, x_n)$ at all points in the neighborhood of $(x_1', x_2', \dots, x_n')$.

Necessary condition for a maxima or minima. A necessary condition for a function f(x, y) of two variables to have a maxima or minima at point (x_0, y_0) is that

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0$$

at the point (i.e. that the point be a **stationary point**).

In the case of a function $f(x_1, x_2, \dots, x_n)$ of n variables, the condition for the function to have a maximum or minimum at point $(x_1', x_2', \dots, x_n')$ is that

$$\frac{\partial f}{\partial x_1} = 0, \qquad \frac{\partial f}{\partial x_2} = 0, \quad \cdots, \frac{\partial f}{\partial x_n} = 0$$

at that point (i.e. that the point be a **stationary point**).

To find the maximum or minimum points of a function we first locate the stationary points using 1) above. After locating the stationary points we then examine each stationary point to determine if it is a maximum or minimum. To determine if a point is a maximum or minimum we may consider values of the function in the neighborhood of the point as well as the values of its first and second partial derivatives. We also may be able to establish what it is by arguments of one kind or other. The following theorem may be useful in establishing maximums and minimums for the case of functions of two variables.

Sufficient condition for a maximum or minimum of a function z = f(x, y). Let z = f(x, y) have continuous first and second partial derivatives in the neighborhood of point (x_0, y_0). If at the point (x_0, y_0)

$$\frac{\partial f}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} = 0$$

and

$$\Delta = \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} < 0$$

then there is a maximum at (x_0, y_0) if

$$\frac{\partial^2 f}{\partial x^2} < 0$$

and a minimum if

$$\frac{\partial^2 f}{\partial x^2} > 0.$$

If $\Delta > 0$, point (x_0, y_0) is a saddle point (neither maximum nor minimum). If $\Delta = 0$, the nature of point (x_0, y_0) is undecided. More investigation is necessary.

Example. Find the maxima and minima of function $z = x^2 + xy + y^2 - y$.

Solution..

$$\frac{\partial z}{\partial x} = 2x + y, \qquad \frac{\partial z}{\partial y} = x + 2y - 1$$

$$\frac{\partial^2 z}{\partial x^2} = 2, \qquad \frac{\partial^2 z}{\partial x \partial y} = 1, \qquad \frac{\partial^2 z}{\partial y^2} = 2$$

$2x + y = 0$
$x + 2y = 1$

$x = -1/3$, $y = 2/3$

This is the stationary point. At this point $\Delta > 0$ and

$$\frac{\partial^2 z}{\partial x^2} > 0$$

and the point is a minimum. The minimum value of the function is - 1/3.

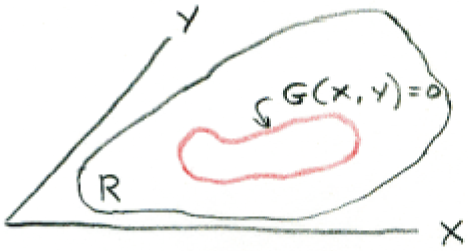


Fig. 1

Maxima and minima of functions subject to constraints. Let us set ourselves the following problem: Let F(x, y) and G(x, y) be functions defined over some region R of the x-y plane. Find the points at which the function F(x, y) has maximums subject to the side condition G(x, y) = 0. Basically we are asking the question: At what points on the solution set of G(x, y) = 0 does F(x, y) have maximums? The solution set of G(x, y) = 0 corresponds to some curve in the plane. See Figure 1. The solution set (i.e. locus) of G(x, y) = 0 is shown in red. Figure 2 shows the situation in three dimensions where function z = F(x, y) is shown rising up above the x-y plane along the curve G(x, y) = 0. The problem is to find the maximums of z = F(x, y) along the curve G(x, y) = 0.

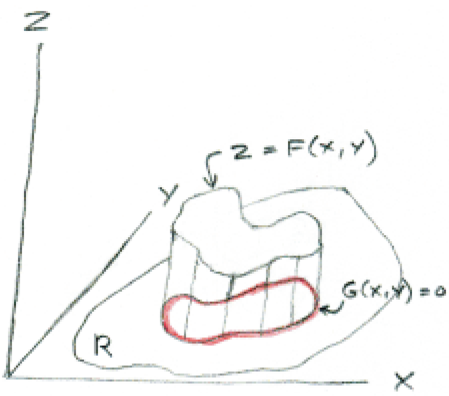


Fig. 2

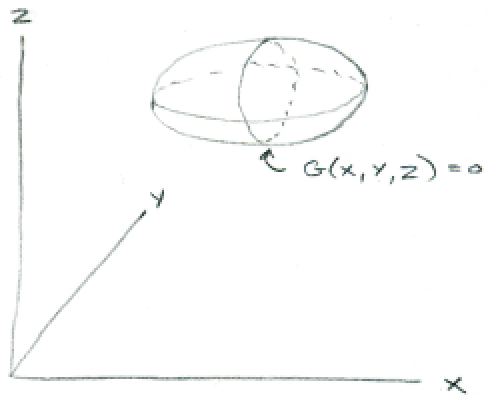


Fig. 3

Let us now consider the same problem in three variables. Let F(x, y, z) and G(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums subject to the side condition G(x, y, z) = 0. Basically we are asking the question: At what points on the solution set of G(x, y, z) = 0 does F(x, y, z) have maximums? G(x, y, z) = 0 represents some surface in space. In Figure 3, G(x, y, z) = 0 is depicted as a spheroid in space. The problem then is to find the maximums of the function F(x, y, z) as evaluated on this spheroidal surface.

Let us now consider another problem. Suppose instead of one side condition we have two. Let F(x, y, z), G(x, y, z) and H(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums subject to the side conditions

2) G(x, y, z) = 0

3) H(x, y, z) = 0.

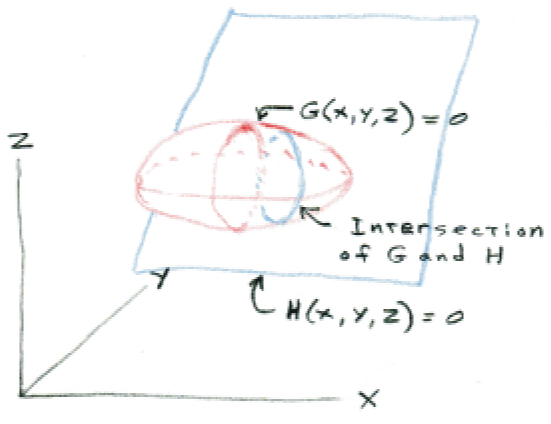


Fig. 4

Here we wish to find the maximum values of F(x, y, z) on that set of points that satisfy both equations 2) and 3). Thus if D represents the solution set of G(x, y, z) = 0 and E represents the solution set of H(x, y, z) = 0 we wish to find the maximum points of F(x, y, z) as evaluated on set F = D $\cap$ E (i.e. the intersection of sets D and E). In Fig. 4 G(x, y, z) = 0 is depicted as an ellipsoid and H(x, y, z) = 0 as a plane. The intersection of the ellipsoid and the plane is the set F on which F(x, y, z) is to be evaluated.

The above can be generalized to functions of n variables $F(x_1, x_2, \dots, x_n)$, $G(x_1, x_2, \dots, x_n)$, etc. and m side conditions.

Methods for finding maxima and minima of functions subject to constraints.

1. Method of direct elimination. Suppose we wish to find the maxima or minima of a function F(x, y) with the constraint Φ(x, y) = 0. Suppose we are so lucky that Φ(x, y) = 0 can be solved explicitly for y, giving y = g(x). We can then substitute g(x) for y in F(x, y) and then find the maximums and minimums of F(x, g(x)) by standard methods. In some cases, it may be possible to do this kind of thing. We express some of the variables in the equations of constraint in terms of other variables and then substitute into the function whose extrema are sought, and find the extrema by standard methods.

2. Method of implicit functions. Suppose we wish to find the maxima or minima of a function u = F(x, y, z) with the constraint Φ(x, y, z) = 0. We note that Φ(x, y, z) = 0 defines z implicitly as a function of x and y i.e. z = f(x, y). We thus seek the extrema of the quantity

$$u = F(x, y, f(x, y)) .$$

The necessary condition for a **stationary point**, as given by 1) above, becomes

4) $$\frac{\partial u}{\partial x} = F_1 + F_3 \frac{\partial z}{\partial x} = 0 \qquad \frac{\partial u}{\partial y} = F_2 + F_3 \frac{\partial z}{\partial y} = 0$$

(where F_1 represents the partial of F with respect to x, etc.)

Taking partials of Φ with respect to x and y it follows that

5) $$\Phi_1 + \Phi_3 \frac{\partial z}{\partial x} = 0 \qquad \Phi_2 + \Phi_3 \frac{\partial z}{\partial y} = 0 .$$

(since the partial derivative of a function that is constant is zero).
From the pair of equations consisting of the first equation in 4) and 5) we can eliminate $\partial z / \partial x$ giving

6) $F_1\Phi_3 - F_3\Phi_1 = 0$

From the pair of equations consisting of the second equation in 4) and 5) we can eliminate $\partial z / \partial y$ giving

7) $F_2\Phi_3 - F_3\Phi_2 = 0$

Equations 6) and 7) can be written in determinant form as

$$8) \quad \begin{vmatrix} F_1 & F_3 \\ \Phi_1 & \Phi_3 \end{vmatrix} = 0 \qquad \begin{vmatrix} F_2 & F_3 \\ \Phi_2 & \Phi_3 \end{vmatrix} = 0$$

Equations 8) combined with the equation Φ(x, y, z) = 0 give us three equations which we can solve simultaneously for x, y, z to obtain the stationary points of function F(x, y, z). The maxima and minima will be among the stationary points.

This same method can be used for functions of an arbitrary number of variables and an arbitrary number of side conditions (smaller than the number of variables).

Extrema for a function of four variables with two auxiliary equations. Suppose we wish to find the maxima or minima of a function

u = F(x, y, z, t)

with the side conditions

9) Φ(x, y, z, t) = 0 ψ(x, y, z, t) = 0.

Equations 9) define variables z and t implicitly as functions of x and y i.e.

10) $z = f_1(x,y)$ $t = f_2(x, y)$.

We thus seek the extrema of the quantity

$u = F(x, y, f_1(x, y), f_2(x, y))$.

The necessary condition for a **stationary point**, as given by 1) above, becomes

$$11) \quad \frac{\partial u}{\partial x} = F_1 + F_3 \frac{\partial z}{\partial x} + F_4 \frac{\partial t}{\partial x} = 0 \qquad \frac{\partial u}{\partial y} = F_2 + F_3 \frac{\partial z}{\partial y} + F_4 \frac{\partial t}{\partial x} = 0$$

Taking partials of Φ with respect to x and y it follows that

$$12) \quad \Phi_1 + \Phi_3 \frac{\partial z}{\partial x} + \Phi_4 \frac{\partial t}{\partial x} = 0 \qquad \Phi_2 + \Phi_3 \frac{\partial f}{\partial y} + \Phi_4 \frac{\partial t}{\partial y} = 0 .$$

Taking partials of ψ with respect to x and y it follows that

$$13) \quad \psi_1 + \psi_3 \frac{\partial z}{\partial x} + \psi_4 \frac{\partial t}{\partial x} = 0 \qquad \psi_2 + \psi_3 \frac{\partial f}{\partial y} + \psi_4 \frac{\partial t}{\partial y} = 0 .$$

From 12) and 13) we can derive the conditions

$$14)\quad \begin{vmatrix} F_1 & F_3 & F_4 \\ \Phi_1 & \Phi_3 & \Phi_4 \\ \Psi_1 & \Psi_3 & \Psi_4 \end{vmatrix} = 0 \qquad \begin{vmatrix} F_2 & F_3 & F_4 \\ \Phi_2 & \Phi_3 & \Phi_4 \\ \Psi_2 & \Psi_3 & \Psi_4 \end{vmatrix} = 0$$

Equations 14) combined with the auxiliary equations Φ(x, y, z, t) = 0 and ψ(x, y, z, t) = 0 give us four equations which we can solve simultaneously for x, y, z, t to obtain the stationary points of function F(x, y, z, t). The maxima and minima will be among the stationary points.

Extrema for a function of n variables with p auxiliary equations.

The p equations corresponding to equation 14) above for the case of a function of n variables

$$u = F(x_1, x_2, \ldots .x_n)$$

and p auxiliary equations (i.e. side conditions)

$\Phi(x_1, x_2, \ldots , x_n) = 0$
$\Psi(x_1, x_2, \ldots , x_n) = 0$
....

$\Omega(x_1, x_2, \ldots , x_n) = 0$

are

$$15)\quad \begin{vmatrix} F_k & F_{n-p+1} & \cdots & F_n \\ \Phi_k & \Phi_{n-p+1} & \cdots & \Phi_n \\ \cdots & \cdots & \cdots & \cdots \\ \Omega_k & \Omega_{n-p+1} & \cdots & \Omega_n \end{vmatrix} = 0 \qquad k = 1, 2, \cdots , p$$

These p equations along with the p auxiliary equations

$\Phi(x_1, x_2, \ldots , x_n) = 0$
$\Psi(x_1, x_2, \ldots , x_n) = 0$
....

$\Omega(x_1, x_2, \ldots , x_n) = 0$

can be solved simultaneously for the n variables x_1, x_2, ...x_n to obtain the stationary points of F(x_1, x_2, ...x_n). The maxima and minima will be among the stationary points.

****

Geometrical interpretation for extrema of function F(x, y, z) with a constraint. We shall now present a theorem that gives a geometrical interpretation for the case of extremal values of functions of type F(x, y, z) with a constraint.

Theorem 1. Suppose the functions F(x, y, z) and Φ(x, y, z) have continuous first partial derivatives throughout a certain region R of space. Let the equation Φ(x, y, z) = 0 define a surface S, every point of which is in the interior of R, and suppose that the three partial derivatives Φ_1, Φ_2, Φ_3 are never simultaneously zero at a point of S. Then a necessary condition for the values of F(x, y, z) on S to attain an extreme value (either relative or absolute) at a point of S is that F_1, F_2, F_3 be proportional to Φ_1, Φ_2, Φ_3 at that point. If C is the value of F at the point, and if the constant of proportionality is not zero, the geometric meaning of the proportionality is that the surface S and the surface F(x, y, z) = C are tangent at the point in question.

Rationale behind theorem. From 8) above, a necessary condition for F(x, y, z) to attain a maxima or minima (i.e. a condition for a stationary point) at a point P is that

$$F_1\Phi_3 - F_3\Phi_1 = 0 \qquad\qquad F_2\Phi_3 - F_3\Phi_2 = 0$$

or

$$16)\qquad \frac{F_1}{\Phi_1} = \frac{F_3}{\Phi_3} \qquad\qquad \frac{F_2}{\Phi_2} = \frac{F_3}{\Phi_3}\,.$$

Thus at a stationary point the partial derivatives F_1, F_2, F_3 and Φ_1, Φ_2, Φ_3 are proportional. Now the partial derivatives F_1, F_2, F_3 and Φ_1, Φ_2, Φ_3 represent the gradients of the functions F and Φ; and the gradient, at any point P, of a scalar point function ψ(x, y, z) is a vector that is normal to that level surface of ψ(x, y, z) that passes through point P. If C is the value of F at the stationary point P, then the vector (F_1, F_2, F_3) at point P is normal to the surface F(x, y, z) = C at P. Similarly, the vector (Φ_1, Φ_2, Φ_3) at point P is normal to the surface Φ(x, y, z) = 0 at P. Since the partial derivatives F_1, F_2, F_3 and Φ_1, Φ_2, Φ_3 are proportional, the normals to the two surfaces point in the same direction at P and the surfaces must be tangent at point P.

Example. Consider the maximum and minimum values of F(x, y, z) = $x^2 + y^2 + z^2$ on the surface of the ellipsoid

$$G(x, y, z) = \frac{x^2}{64} + \frac{y^2}{36} + \frac{z^2}{25} = 1\,.$$

Since F(x, y, z) is the square of the distance from (x, y, z) to the origin, it is clear that we are looking for the points at maximum and minimum distances from the center of the ellipsoid. The maximum occurs at the ends of the longest principal axis, namely at ($\pm$8, 0, 0). The minimum occurs at the ends of the shortest principal axis, namely at (0, 0, $\pm$5). Consider the maximum point (8, 0, 0). The value of F at this point is 64, and the surface F(x, y, z) = 64 is a sphere. The sphere and the ellipsoid are tangent at (8, 0, 0) as asserted by the theorem. In this case the ratios G_1:G_2:G_3 and F_1:F_2:F_3 at (8, 0, 0) are 1/4 : 0 : 0 and 16 : 0 : 0 respectively.

This example brings out the fact that the tangency of the surfaces (or the proportionality of the two sets of ratios), is a necessary but not a sufficient condition for a maximum or minimum value of F, for we note that the condition of proportionality exists at the points (0, $\pm$6, 0), which are the ends of the principal axis of intermediate length. But the value of F in neither a maximum nor a minimum at this point.

Case of extrema of function F(x, y) with a constraint. A similar geometrical interpretation can be given to the problem of extremal values for F(x, y) subject to the constraint Φ(x, y) = 0. Here we have a curve defined by the constraint, and a one-parameter family of curves F(x, y) = C. At a point of extremal value of F the curve F(x, y) = C through the point will be tangent to the curve defined by the constraint.

3. Lagrange's Method of Multipiers. Let F(x, y, z) and Φ(x, y, z) be functions defined over some region R of space. Find the points at which the function F(x, y, z) has maximums and minimums subject to the side condition Φ(x, y, z) = 0. Lagrange's method for solving this problem consists of forming a third function G(x, y, z) given by

17) $G(x, y, z) = F(x, y, z) + \lambda\Phi(x, y, z)$,

where λ is a constant (i.e. a parameter) to which we will later assign a value, and then finding the maxima and minima of the function G(x, y, z). A reader might quickly ask, "Of what interest are the maxima and minima of the function G(x, y, z)? How does this help us solve the problem of finding the maxima and minima of F(x, y, z)?" The answer is that examination of 17) shows that for those points corresponding to the solution set of Φ(x, y, z) = 0 the function G(x, y, z) is equal to the function F(x, y, z) since at those points equation 17) becomes

$$G(x, y, z) = F(x, y, z) + \lambda\cdot 0 .$$

Thus, for the points on the surface Φ(x, y, z) = 0, functions F and G are equal so the maxima and minima of G are also the maxima and minima of F. The procedure for finding the maxima and minima of G(x, y, z) is as follows: We regard G(x, y, z) as a function of three independent variables and write down the necessary conditions for a stationary point using 1) above:

18) $F_1 + \lambda\Phi_1 = 0 \qquad F_2 + \lambda\Phi_2 = 0 \qquad F_3 + \lambda\Phi_3 = 0$

We then solve these three equations along with the equation of constraint Φ(x, y, z) = 0 to find the values of the four quantities x, y, z, λ. More than one point can be found in this way and this will give us the locations of the stationary points. The maxima and minima will be among the stationary points thus found.

Let us now observe something. If equations 18) are to hold simultaneously, then it follows from the third of them that λ must have the value

$$\lambda = -\frac{F_3}{\Phi_3} .$$

If we substitute this value of λ into the first two equations of 18) we obtain

$F_1\Phi_3 - F_3\Phi_1 = 0$ $\quad\quad\quad$ $F_2\Phi_3 - F_3\Phi_2 = 0$

or

19) $$\begin{vmatrix} F_1 & F_3 \\ \Phi_1 & \Phi_3 \end{vmatrix} = 0 \qquad \begin{vmatrix} F_2 & F_3 \\ \Phi_2 & \Phi_3 \end{vmatrix} = 0$$

We note that the two equations of 19) are identically the same conditions as 8) above for the previous method. Thus using equations 19) along with the equation of constraint $\Phi(x, y, z) = 0$ is exactly the same procedure as the previous method in which we used equations 8) and the same constraint.

One of the great advantages of Lagrange's method over the method of implicit functions or the method of direct elimination is that it enables us to avoid making a choice of independent variables. This is sometimes very important; it permits the retention of symmetry in a problem where the variables enter symmetrically at the outset.

Lagrange's method can be used with functions of any number of variables and any number of constraints (smaller than the number of variables). In general, given a function $F(x_1, x_2, \ldots, x_n)$ of n variables and h side conditions $\Phi_1 = 0$, $\Phi_2 = 0$, ... , $\Phi_h = 0$, for which this function may have a maximum or minimum, equate to zero the partial derivatives of the auxiliary function $F + \lambda_1\Phi_1 + \lambda_2\Phi_2 + \ldots\ldots + \lambda_h\Phi_h$ with respect to $x_1, x_2, \ldots, x_n$, regarding $\lambda_1, \lambda_2, \ldots\ldots, \lambda_h$ as constants, and solve these n equations simultaneously with the given h side conditions, treating the λ's as unknowns to be eliminated.

The parameter λ in Lagrange's method is called Lagrange's multiplier.
Further examples

Example 1.

Let us find the critical points of

$$z = f(x,y) = \exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

The partial derivatives are

$$f_x(x,y) = (-x^2+1)\exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

$$f_y(x,y) = -2y\exp\left(-\frac{1}{3}x^3 + x - y^2\right)$$

f_x=0 if 1-x^2=0 or the exponential term is 0. f_y=0 if -2y=0 or the exponential term is 0.
The exponential term is not 0 except in the degenerate case. Hence we require 1-x^2=0 and -2y=0, implying x=1 or x=-1 and y=0. There are two critical points (-1,0) and (1,0)

The Second Derivative Test for Functions of Two Variables

How can we determine if the critical points found above are relative maxima or minima? We apply a second derivative test for functions of two variables.

Let (x_c,y_c) be a critical point and define

$$D(x_c, y_c) = f_{xx}(x_c, y_c) f_{yy}(x_c, y_c) - [f_{xy}(x_c, y_c)]^2.$$

We have the following cases:

- If D>0 and $f_{xx}(x_c, y_c) \langle 0,$ then f(x,y) has a relative maximum at x_c, y_c.
- If D>0 and $f_{xx}(x_c, y_c) \langle 0,$ then f(x,y) has a relative minimum at x_c, y_c.
- If D<0, then f(x,y) has a saddle point at x_c, y_c.
- If D=0, the second derivative test is inconclusive.

An example of a saddle point is shown in the example below.

Example: Continued

For the example above, we have

$$f_{xx}(x, y) = (-2x + (1 - x^2)^2) exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

$$f_{yy}(x, y) = (-2 + 4y^2) exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

$$f_{xy}(x, y) = -2y(1 - x^2) exp\left(-\frac{1}{3}x^3 + x - y^2\right),$$

For x=1 and y=0, we have D(1,0)=4exp(4/3)>0 with f_xx(1,0)=-2exp(2/3)<0. Hence, (1,0) is a relative maximum. For x=-1 and y=0, we have D(-1,0)=-4exp(-4/3)<0. Hence, (-1,0) is a saddle point.

Example 2: Maxima and Minima in a Disk

Another example of a bounded region is the disk of radius 2 centered at the origin. We proceed as in the previous example, determining in the 3 classes above. (1,0) and (-1,0) lie in the interior of the disk.

The boundary of the disk is the circle x^2+y^2=4. To find extreme points on the disk we parameterize the circle. A natural parameterization is x=2cos(t) and y=2sin(t) for 0<=t<=2*pi. We substitute these expressions into z=f(x,y) and obtain

$$z = f(x, y) = f(\cos(t), \sin(t)) = exp\left(-\frac{8}{3}\cos^3 t + 2\cos t - 4\sin^2 t\right) = g(t)$$

On the circle, the original functions of 2 variables is reduced to a function of 1 variable. We can determine the extrema on the circle using techniques from calculus of on variable.

In this problem there are not any corners. Hence, we determine the global max and min by considering points in the interior of the disk and on the circle. An alternative method for finding the maximum and minimum on the circle is the method of Lagrange multipliers.

4.0 CONCLUSION

You have been introduced to maximum and minimum functions of several variables, necessary condition for a maxima or minima function of several variables, problems on maximum and minimum functions of several variables e.t.c

5.0 SUMMARY

A summary of maximum and minimum functions of several variables are as follows :

A function f(x, y) of two independent variables has a **maximum** at a point (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0). Such a function has a **minimum** at a point (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all points (x, y) in the neighborhood of (x_0, y_0).

Solve the following problem, Find the maxima and minima of function $z = x^2 + xy + y^2 - y$.

Solution..

$$\frac{\partial z}{\partial x} = 2x + y\ , \qquad \frac{\partial z}{\partial y} = x + 2y - 1 \qquad \frac{\partial^2 z}{\partial x^2} = 2\ , \qquad \frac{\partial^2 z}{\partial x \partial y} = 1\ , \qquad \frac{\partial^2 z}{\partial y^2} = 2$$

2x + y = 0 , x + 2y = 1

x = -1/3 , y = 2/3

This is the stationary point. At this point Δ > 0 and

$$\frac{\partial^2 z}{\partial x^2} > 0$$

and the point is a minimum. The minimum value of the function is - 1/3.

6.0 TUTOR-MARKED ASSIGNMENT

1.Determine the critical points and locate any relative minimum, maxima and saddle points of functions f defined by

$F(x,y) = 2x^2 - 2xy + 2y^4 - 6x$

2.Determine the critical points and locate any relative minimum, maxima and saddle points of functions f defined by

$F(x,y) = 2x^4 - 4xy + y^3 + 4$

3.Determine the critical points and locate any relative minimum, maxima and saddle points of functions f defined by

F(x,y)= $x^4 - y^4 + 4xy$

Determine the critical points of the functions below and find out whether each point corresponds to a relative minimum, maximum and saddle point, or no conclusion can be made

4.F(x,y)= $x^2 + 3y^2 - 2xy - 8x$

5. F(x,y)= $x^3 + 12x + y^3 + 3y^2 - 9y$

7.0 REFERENCES.

Taylor. Advanced Calculus

Osgood. Advanced Calculus.

James and James. Mathematics Dictionary.

Mathematics, Its Content, Methods and Meaning. Vol.

UNIT 2 LAGRANGE MULTIPLIERS

CONTENTS

1.0 **INTRODUCTION**

2.0 OBJECTIVES

3.0 MAIN CONTENT

3.1 Handling multiple constraints

3.2 interpretation of the langrange multiplie

3.3 example

3.4 applications

4.0 CONCLUSION

5.0 SUMMARY

6.0 TUTOR-MARKED ASSIGNMENT

7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

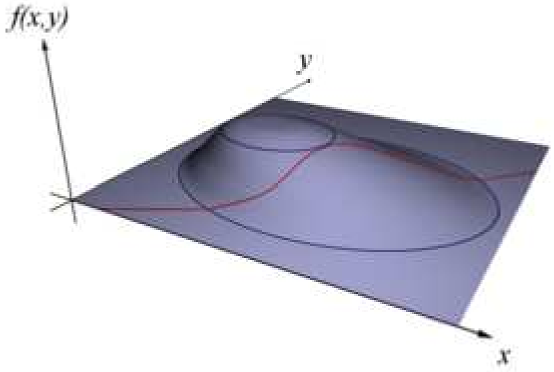


Figure 1: Find *x* and *y* to maximize *f*(*x*,*y*) subject to a constraint (shown in red) *g*(*x*,*y*) = *c*.

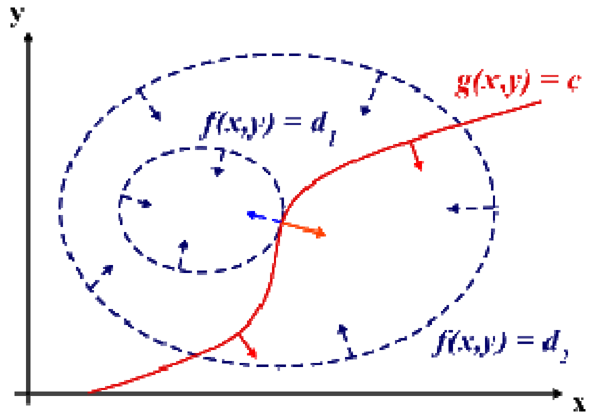


Figure 2: Contour map of Figure 1. The red line shows the constraint *g*(*x*,*y*) = *c*. The blue lines are contours of *f*(*x*,*y*). The point where the red line tangentially touches a blue contour is our solution.

In mathematical optimization, the method of **Lagrange multipliers** (named after Joseph Louis Lagrange) provides a strategy for finding the maxima and minima of a function subject to constraints.

For instance (see Figure 1), consider the optimization problem

maximize $f(x, y)$
subject to $g(x, y) = c.$

We introduce a new variable (λ) called a Lagrange multiplier, and study the Lagrange function defined by

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda \cdot \Big(g(x, y) - c\Big),$$

where the λ term may be either added or subtracted. If $f(x,y)$ is a maximum for the original constrained problem, then there exists λ such that (x,y,λ) is a stationary point for the Lagrange function (stationary points are those points where the partial derivatives of Λ are zero). However, not all stationary points yield a solution of the original problem. Thus, the method of Lagrange multipliers yields a necessary condition for optimality in constrained problems

2.0 OBJECTIVES

After studying this unit, you should be to correctly:

i. Identify problem which could be solve by langranges multiplier
ii. Know single and multiple constraints
iii. Know the interpretation of lagrange multiplier
iv. Solve problems with the use of langranges multiplier

3.0 MAIN CONTENT

One of the most common problems in calculus is that of finding maxima or minima (in general, "extrema") of a function, but it is often difficult to find a closed form for the function being extremized. Such difficulties often arise when one wishes to maximize or minimize a function subject to fixed outside conditions or constraints. The method of Lagrange multipliers is a powerful tool for solving this class of problems without the need to explicitly solve the conditions and use them to eliminate extra variables.

Consider the two-dimensional problem introduced above:

maximize $f(x, y)$
subject to $g(x, y) = c.$

We can visualize contours of *f* given by

$$f(x,y) = d$$

for various values of *d*, and the contour of *g* given by $g(x,y) = c$.

Suppose we walk along the contour line with $g = c$. In general the contour lines of *f* and *g* may be distinct, so following the contour line for $g = c$ one could intersect with or cross the contour lines of *f*. This is equivalent to saying that while moving along the contour line for $g = c$ the value of *f* can vary. Only when the contour line for $g = c$ meets contour lines of *f* tangentially, do we not increase or decrease the value of *f* — that is, when the contour lines touch but do not cross

The contour lines of *f* and *g* touch when the tangent vectors of the contour lines are parallel. Since the gradient of a function is perpendicular to the contour lines, this is the same as saying that the gradients of *f* and *g* are parallel. Thus we want points (x,y) where $g(x,y) = c$ and

$$\nabla_{x,y} f = -\lambda \nabla_{x,y} g,$$

where

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

and

$$\nabla_{x,y} g = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y} \right)$$

are the respective gradients. The constant λ is required because although the two gradient vectors are parallel, the magnitudes of the gradient vectors are generally not equal.

To incorporate these conditions into one equation, we introduce an auxiliary function

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda \cdot \Big(g(x, y) - c \Big),$$

and solve

$$\nabla_{x,y,\lambda} \Lambda(x, y, \lambda) = 0.$$

This is the method of Lagrange multipliers. Note that $\nabla_\lambda \Lambda(x, y, \lambda) = 0$ implies $g(x,y) = c$.

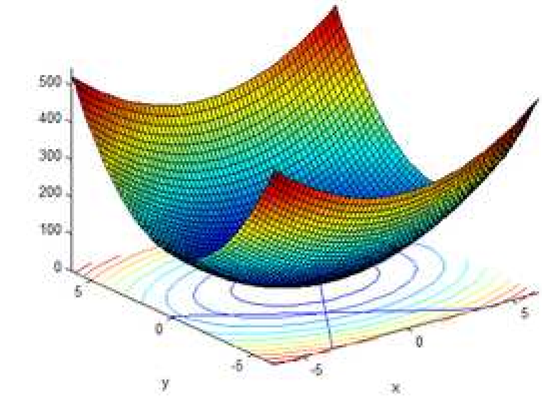
Not necessarily extrema

The constrained extrema of *f* are *critical points* of the Lagrangian Λ, but they are not *local extrema* of Λ (see Example 2 below).

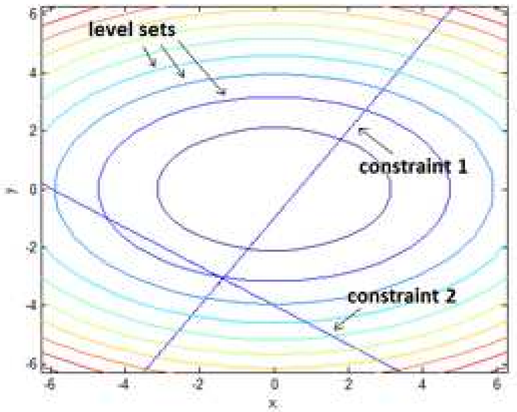
One may reformulate the Lagrangian as a Hamiltonian, in which case the solutions are local minima for the Hamiltonian. This is done in optimal control theory, in the form of Pontryagin's minimum principle.

The fact that solutions of the Lagrangian are not necessarily extrema also poses difficulties for numerical optimization. This can be addressed by computing the *magnitude* of the gradient, as the zeros of the magnitude are necessarily local minima, as illustrated in the numerical optimization example.

Handling multiple constraints



A paraboloid, some of its level sets (aka contour lines) and 2 line constraints.



Zooming in on the levels sets and constraints, we see that the two constraint lines intersect to form a "joint" constraint that is a point. Since there is only one point to analyze, the corresponding point on the paraboloid is automatically a minimum and maximum. Yet the simplified reasoning presented in sections above seems to fail because the level set definitely appears to "cross" the point and at the same time its gradient is not parallel to the gradients of either constraint. This shows we must refine our explanation of the method to handle the kinds of constraints that are formed when we have more than one constraint acting at once.

The method of *Lagrange multipliers* can also accommodate multiple constraints. To see how this is done, we need to reexamine the problem in a slightly different manner because the concept of "crossing" discussed above becomes rapidly unclear when we consider the types of constraints that are created when we have more than one constraint acting together.

As an example, consider a paraboloid with a constraint that is a single point (as might be created if we had 2 line constraints that intersect). The level set (i.e., contour line) clearly appears to "cross" that point and its gradient is clearly not parallel to the gradients of either of the two line constraints. Yet, it is obviously a maximum *and* a minimum because there is only one point on the paraboloid that meets the constraint.

While this example seems a bit odd, it is easy to understand and is representative of the sort of "effective" constraint that appears quite often when we deal with multiple constraints intersecting. Thus, we take a slightly different approach below to explain and derive the Lagrange Multipliers method with any number of constraints.

Throughout this section, the independent variables will be denoted by $x_1, x_2, \ldots, x_N$ and, as a group, we will denote them as $p = (x_1, x_2, \ldots, x_N)$. Also, the function being analyzed will be denoted by and the constraints will be represented by the equations $g_1(p) = 0, g_2(p) = 0, \ldots, g_M(p) = 0$.

The basic idea remains essentially the same: if we consider only the points that satisfy the constraints (i.e. are *in* the constraints), then a point $(p, f(p))$ is a stationary point (i.e. a point in a "flat" region) of *f* if and only if the constraints at that point do not allow movement in a direction where *f* changes value.

Once we have located the stationary points, we need to do further tests to see if we have found a minimum, a maximum or just a stationary point that is neither.

We start by considering the level set of *f* at $(p, f(p))$. The set of vectors $\{v_L\}$ containing the directions in which we can move and still remain in the same level set are the directions where the value of *f* does not change (i.e. the change equals zero). Thus, for every vector *v* in $\{v_L\}$, the following relation must hold:

$$\Delta f = \frac{df}{dx_1} v_{x1} + \frac{df}{dx_2} v_{x2} + \cdots + \frac{df}{dx_N} v_{xN} = 0$$

where the notation v_{x_K} above means the x_K-component of the vector *v*. The equation above can be rewritten in a more compact geometric form that helps our intuition:

$$\underbrace{\begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \\ \vdots \\ \frac{df}{dx_N} \end{bmatrix}}_{\nabla f} \cdot \underbrace{\begin{bmatrix} v_{x_1} \\ v_{x_2} \\ \vdots \\ v_{x_N} \end{bmatrix}}_{v} = 0 \qquad \Rightarrow \qquad \nabla f \cdot v = 0$$

This makes it clear that if we are at *p*, then *all* directions from this point that do *not* change the value of *f must be perpendicular* to $\nabla f(p)$ (the gradient of *f* at *p*).

Now let us consider the effect of the constraints. Each constraint limits the directions that we can move from a particular point and still satisfy the constraint. We can use the same procedure, to look for the set of vectors $\{v_C\}$ containing the directions in which we can move and still satisfy the constraint. As above, for every vector *v* in $\{v_C\}$, the following relation must hold:

$$\Delta g = \frac{dg}{dx_1}v_{x_1} + \frac{dg}{dx_2}v_{x_2} + \cdots + \frac{dg}{dx_N}v_{x_N} = 0 \qquad \Rightarrow \qquad \nabla g \cdot v = 0$$

From this, we see that at point *p*, all directions from this point that will still satisfy this constraint must be perpendicular to $\nabla g\,(p)$.

Now we are ready to refine our idea further and complete the method: *a point on f is a constrained stationary point if and only if the direction that changes f violates at least one of the constraints*. (We can see that this is true because if a direction that changes *f* did *not* violate any constraints, then there would a “legal” point nearby with a higher or lower value for *f* and the current point would then not be a stationary point.)

Single constraint revisited

For a single constraint, we use the statement above to say that at stationary points the direction that changes *f* is in the same direction that violates the constraint. To determine if two vectors are in the same direction, we note that if two vectors start from the same point and are “in the same direction”, then one vector can always “reach” the other by changing its length and/or flipping to point the opposite way along the same direction line. In this way, we can succinctly state that two vectors point in the same direction if and only if one of them can be multiplied by some real number such that they become equal to the other. So, for our purposes, we require that:

$$\nabla f\,(p) = \lambda \nabla g\,(p) \qquad \Rightarrow \qquad \nabla f\,(p) - \lambda\, \nabla g\,(p) = 0$$

If we now add another simultaneous equation to guarantee that we only perform this test when we are at a point that satisfies the constraint, we end up with 2 simultaneous equations that when solved, identify all constrained stationary points:

$$\begin{cases} g\,(p) = 0 & \text{means point satisfies constraint} \\ \nabla f\,(p) - \lambda\, \nabla g\,(p) = 0 & \text{means point is a stationary point} \end{cases}$$

Note that the above is a succinct way of writing the equations. Fully expanded, there are N + 1 simultaneous equations that need to be solved for the N + 1 variables which are λ and $x_1,\ x_2, \ldots, x_N$:

$$\begin{aligned}
g\left(x_1, x_2, \ldots, x_N\right) &= 0 \\
\frac{df}{dx_1}\left(x_1, x_2, \ldots, x_N\right) - \lambda \frac{dg}{dx_1}\left(x_1, x_2, \ldots, x_N\right) &= 0 \\
\frac{df}{dx_2}\left(x_1, x_2, \ldots, x_N\right) - \lambda \frac{dg}{dx_2}\left(x_1, x_2, \ldots, x_N\right) &= 0 \\
&\vdots \\
\frac{df}{dx_N}\left(x_1, x_2, \ldots x_N\right) - \lambda \frac{dg}{dx_N}\left(x_1, x_2, \ldots, x_N\right) &= 0
\end{aligned}$$

Multiple constraints

For more than one constraint, the same reasoning applies. If there is more than one constraint active together, each constraint contributes a direction that will violate it. Together, these "violation directions" form a "violation space", where infinitesimal movement in any direction within the space will violate one or more constraints. Thus, to satisfy multiple constraints we can state (using this new terminology) that at the stationary points, the direction that changes *f* is in the "violation space" created by the constraints acting jointly.

The *violation space* created by the constraints consists of all points that can be reached by adding any combination of scaled and/or flipped versions of the individual violation direction vectors. In other words, all the points that are "reachable" when we use the individual violation directions as the basis of the space. Thus, we can succinctly state that *v* is in the space defined by $b_1, b_2, \ldots, b_M$ if and only if there exists a set of "multipliers" $\lambda_1, \lambda_2, \ldots, \lambda_M$ such that:

$$\sum_{k=1}^{M} \lambda_k b_k = v$$

which for our purposes, translates to stating that the direction that changes *f* at *p* is in the "violation space" defined by the constraints $g_1, g_2, \ldots, g_M$ if and only if:

$$\sum_{k=1}^{M} \lambda_k \nabla g_k(p) = \nabla f(p) \quad \Rightarrow \quad \nabla f(p) - \sum_{k=1}^{M} \lambda_k \nabla g_k(p) = 0$$

As before, we now add simultaneous equation to guarantee that we only perform this test when we are at a point that satisfies every constraint, we end up with simultaneous equations that when solved, identify all constrained stationary points:

$$\begin{aligned}
g_1(p) &= 0 \\
g_2(p) &= 0 \qquad \text{these mean the point satisfies all constraints} \\
&\vdots \\
g_M(p) &= 0 \\
\nabla f(p) - \sum_{k=1}^{M} \lambda_k \nabla g_k(p) &= 0 \qquad \text{this means the point is a stationary point}
\end{aligned}$$

The method is complete now (from the standpoint of solving the problem of finding stationary points) but as mathematicians delight in doing, these equations can be further condensed into an even more elegant and succinct form. Lagrange must have cleverly noticed that the equations above look like partial derivatives of some larger scalar function *L* that takes all the $x_1, x_2, \ldots, x_N$ and all the $\lambda_1, \lambda_2, \ldots, \lambda_M$ as inputs. Next, he might then have noticed that setting every equation equal to zero is exactly what one would have to do to solve for the *unconstrained* stationary points of that larger function. Finally, he showed that a larger function *L* with partial derivatives that are exactly the ones we require can be constructed very simply as below:

$$L(x_1, x_2, \ldots, x_N, \lambda_1, \lambda_2, \ldots, \lambda_M)$$
$$= f(x_1, x_2, \ldots, x_N) - \sum_{k=1}^{M} \lambda_k g_k (x_1, x_2, \ldots, x_N)$$

Solving the equation above for its *unconstrained* stationary points generates exactly the same stationary points as solving for the *constrained* stationary points of *f* under the constraints $g_1, g_2, \ldots, g_M$.

In Lagrange's honor, the function above is called a *Lagrangian*, the scalars $\lambda_1, \lambda_2, \ldots, \lambda_M$ are called *Lagrange Multipliers* and this optimization method itself is called *The Method of Lagrange Multipliers*.

The method of Lagrange multipliers is generalized by the Karush–Kuhn–Tucker conditions, which can also take into account inequality constraints of the form $h(\mathbf{x}) \leq c$.

Interpretation of the Lagrange multipliers

Often the Lagrange multipliers have an interpretation as some quantity of interest. To see why this might be the case, observe that:

$$\frac{\partial L}{\partial g_k} = \lambda_k.$$

So, λ_k is the rate of change of the quantity being optimized as a function of the constraint variable. As examples, in Lagrangian mechanics the equations of motion are derived by finding stationary points of the action, the time integral of the difference between kinetic and potential energy. Thus, the force on a particle due to a scalar potential, $F = -\nabla V$, can be interpreted as a Lagrange multiplier determining the change in action (transfer of potential to kinetic energy) following a variation in the particle's constrained trajectory. In economics, the optimal profit to a player is calculated subject to a constrained space of actions, where a Lagrange multiplier is the increase in the value of the objective function due to the relaxation of a given constraint (e.g. through an increase in income or bribery or other means) – the marginal cost of a constraint, called the shadow price.

In control theory this is formulated instead as costate equations.

Examples

Example 1

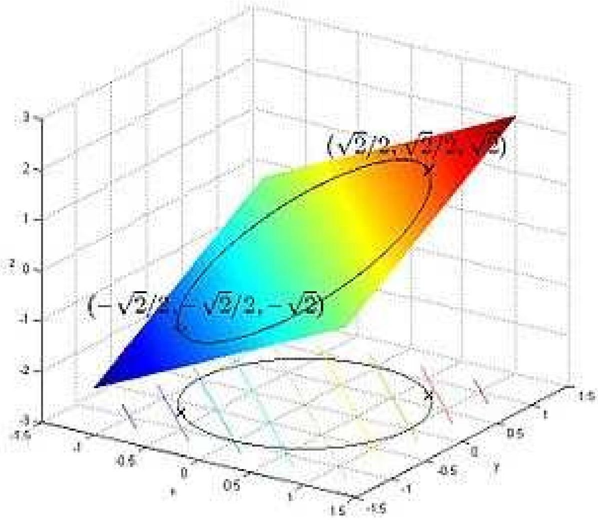


Fig. 3. Illustration of the constrained optimization problem

Suppose one wishes to maximize $f(x,y) = x + y$ subject to the constraint $x^2 + y^2 = 1$. The feasible set is the unit circle, and the level sets of f are diagonal lines (with slope -1), so one can see graphically that the maximum occurs at $(\sqrt{2}/2, \sqrt{2}/2)$, and the minimum occurs at $(-\sqrt{2}/2, -\sqrt{2}/2)$.

Formally, set $g(x,y) - c = x^2 + y^2 - 1$, and

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda(g(x,y) - c) = x + y + \lambda(x^2 + y^2 - 1)$$

Set the derivative $d\Lambda = 0$, which yields the system of equations:

$$\begin{aligned} \frac{\partial \Lambda}{\partial x} &= 1 + 2\lambda x &&= 0, \qquad \text{(i)} \\ \frac{\partial \Lambda}{\partial y} &= 1 + 2\lambda y &&= 0, \qquad \text{(ii)} \\ \frac{\partial \Lambda}{\partial \lambda} &= x^2 + y^2 - 1 &&= 0, \qquad \text{(iii)} \end{aligned}$$

As always, the $\partial \lambda$equation ((iii) here) is the original constraint.

Combining the first two equations yields $x = y$ (explicitly, $\lambda \neq 0$, otherwise (i) yields $1 = 0$, so one has $x = -1 / (2\lambda) = y$).

Substituting into (iii) yields $2x^2 = 1$, so $x = y = \pm\sqrt{2}/2$ and $\lambda = \mp\sqrt{2}/2$, showing the stationary points are $(\sqrt{2}/2, \sqrt{2}/2)$ and $(-\sqrt{2}/2, -\sqrt{2}/2)$. Evaluating the objective function *f* on these yields
$f(\sqrt{2}/2, \sqrt{2}/2) = \sqrt{2}$ and $f(-\sqrt{2}/2, -\sqrt{2}/2) = -\sqrt{2},$

thus the maximum is $\sqrt{2}$, which is attained at $(\sqrt{2}/2, \sqrt{2}/2)$, and the minimum is $-\sqrt{2}$, which is attained at $(-\sqrt{2}/2, -\sqrt{2}/2)$.

Example 2

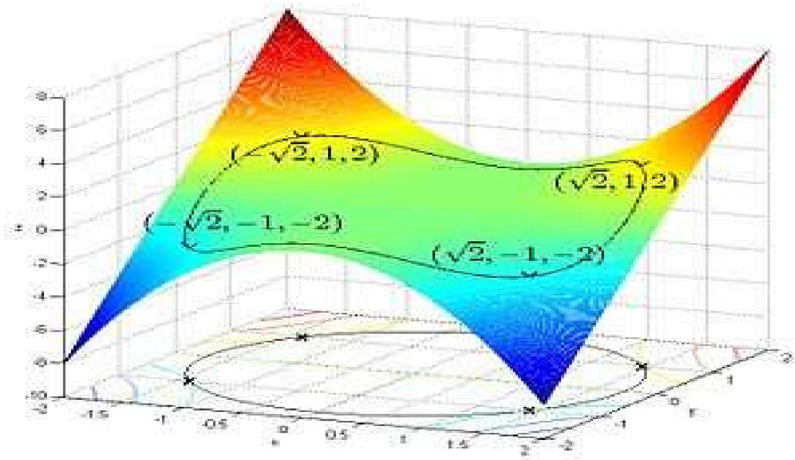


Fig. 4. Illustration of the constrained optimization problem

Suppose one wants to find the maximum values of

$$f(x, y) = x^2 y$$

with the condition that the *x* and *y* coordinates lie on the circle around the origin with radius √3, that is, subject to the constraint

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

As there is just a single constraint, we will use only one multiplier, say λ.

The constraint *g*(*x*, *y*)-3 is identically zero on the circle of radius √3. So any multiple of *g*(*x*, *y*)-3 may be added to *f*(*x*, *y*) leaving *f*(*x*, *y*) unchanged in the region of interest (above the circle where our original constraint is satisfied). Let

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda(g(x, y) - 3) = x^2 y + \lambda(x^2 + y^2 - 3).$$

The critical values of Λ occur where its gradient is zero. The partial derivatives are

$$\frac{\partial \Lambda}{\partial x} = 2xy + 2\lambda x \quad = 0, \qquad \text{(i)}$$
$$\frac{\partial \Lambda}{\partial y} = x^2 + 2\lambda y \quad = 0, \qquad \text{(ii)}$$
$$\frac{\partial \Lambda}{\partial \lambda} = x^2 + y^2 - 3 \quad = 0. \qquad \text{(iii)}$$

Equation (iii) is just the original constraint. Equation (i) implies $x = 0$ *or* $\lambda = -y$. In the first case, if $x = 0$ then we must have $y = \pm\sqrt{3}$ by (iii) and then by (ii) $\lambda = 0$. In the second case, if $\lambda = -y$ and substituting into equation (ii) we have that,

$$x^2 - 2y^2 = 0.$$

Then $x^2 = 2y^2$. Substituting into equation (iii) and solving for y gives this value of y:

$$y = \pm 1.$$

Thus there are six critical points:

$$(\sqrt{2}, 1); \quad (-\sqrt{2}, 1); \quad (\sqrt{2}, -1); \quad (-\sqrt{2}, -1); \quad (0, \sqrt{3}); \quad (0, -\sqrt{3}).$$

Evaluating the objective at these points, we find

$$f(\pm\sqrt{2}, 1) = 2; \quad f(\pm\sqrt{2}, -1) = -2; \quad f(0, \pm\sqrt{3}) = 0.$$

Therefore, the objective function attains the global maximum (subject to the constraints) at $(\pm\sqrt{2}, 1)$ and the global minimum at $(\pm\sqrt{2}, -1)$. The point $(0, \sqrt{3})$ is a local minimum and $(0, -\sqrt{3})$ is a local maximum, as may be determined by consideration of the Hessian matrix of Λ.

Note that while $(\sqrt{2}, 1, -1)$ is a critical point of Λ, it is not a local extremum. We have $\Lambda(\sqrt{2} + \epsilon, 1, -1 + \delta) = 2 + \delta(\epsilon^2 + (2\sqrt{2})\epsilon)$. Given any neighborhood of $(\sqrt{2}, 1, -1)$, we can choose a small positive ϵ and a small δ of either sign to get Λ values both greater and less than 2.

Example: entropy

Suppose we wish to find the discrete probability distribution on the points $\{x_1, x_2, \ldots, x_n\}$ with maximal information entropy. This is the same as saying that we wish to find the least biased probability distribution on the points $\{x_1, x_2, \ldots, x_n\}$. In other words, we wish to maximize the Shannon entropy equation:

$$f(p_1, p_2, \ldots, p_n) = -\sum_{j=1}^{n} p_j \log_2 p_j.$$

For this to be a probability distribution the sum of the probabilities p_i at each point x_i must equal 1, so our constraint is $g(\vec{p})$= 1:

$$g(p_1, p_2, \ldots, p_n) = \sum_{j=1}^{n} p_j.$$

We use Lagrange multipliers to find the point of maximum entropy, $\vec{p}^{\,*}$, across all discrete probability distributions $\vec{p}$ on $\{x_1, x_2, \ldots, x_n\}$. We require that:

$$\left.\frac{\partial}{\partial \vec{p}}(f + \lambda(g-1))\right|_{\vec{p}=\vec{p}^{\,*}} = 0,$$

which gives a system of *n* equations, $k = 1, \ldots, n$, such that:

$$\left.\frac{\partial}{\partial p_k}\left(-\sum_{j=1}^{n} p_j \log_2 p_j + \lambda\left(\sum_{j=1}^{n} p_j - 1\right)\right)\right|_{p_k=p_k^*} = 0.$$

Carrying out the differentiation of these *n* equations, we get

$$-\left(\frac{1}{\ln 2} + \log_2 p_k^*\right) + \lambda = 0.$$

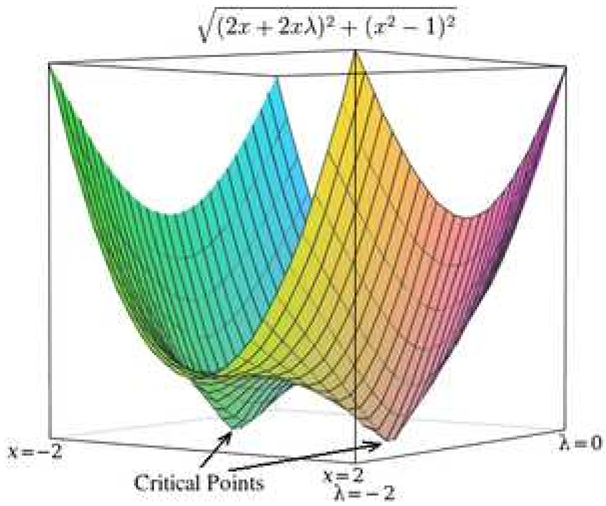
This shows that all $p_k^*$are equal (because they depend on λ only). By using the constraint $\sum_j p_j = 1$, we find

$$p_k^* = \frac{1}{n}.$$

Hence, the uniform distribution is the distribution with the greatest entropy, among distributions on *n* points.

Example: numerical optimization

Lagrange multipliers cause the critical points to occur at saddle points.



The magnitude of the gradient can be used to force the critical points to occur at local minima.

With Lagrange multipliers, the critical points occur at saddle points, rather than at local maxima (or minima). Unfortunately, many numerical optimization techniques, such as hill climbing, gradient descent, some of the quasi-Newton methods, among others, are designed to find local maxima (or minima) and not saddle points. For this reason, one must either modify the formulation to ensure that it's a minimization problem (for example, by extremizing the square of the gradient of the Lagrangian as below), or else use an optimization technique that finds stationary points (such as Newton's method without an extremum seeking line search) and not necessarily extrema.

As a simple example, consider the problem of finding the value of x that minimizes $f(x) = x^2$, constrained such that $x^2 = 1$. (This problem is somewhat pathological because there are only two values that satisfy this constraint, but it is useful for illustration purposes because the corresponding unconstrained function can be visualized in three dimensions.)

Using Lagrange multipliers, this problem can be converted into an unconstrained optimization problem:

$$\Lambda(x,\lambda) = x^2 + \lambda(x^2 - 1)$$

The two critical points occur at saddle points where $x = 1$ and $x = -1$.

In order to solve this problem with a numerical optimization technique, we must first transform this problem such that the critical points occur at local minima. This is done by computing the magnitude of the gradient of the unconstrained optimization problem.

First, we compute the partial derivative of the unconstrained problem with respect to each variable:

$$\frac{\partial \Lambda}{\partial x} = 2x + 2x\lambda$$

$$\frac{\partial \Lambda}{\partial \lambda} = x^2 - 1$$

If the target function is not easily differentiable, the differential with respect to each variable can be approximated as

$$\frac{\partial \Lambda}{\partial x} \approx \frac{\Lambda(x+\epsilon, \lambda) - \Lambda(x, \lambda)}{\epsilon},$$

$$\frac{\partial \Lambda}{\partial \lambda} \approx \frac{\Lambda(x, \lambda+\epsilon) - \Lambda(x, \lambda)}{\epsilon},$$

where ϵ is a small value.

Next, we compute the magnitude of the gradient, which is the square root of the sum of the squares of the partial derivatives:

$$h(x,\lambda) = \sqrt{(2x+2x\lambda)^2 + (x^2-1)^2} \approx \sqrt{\left(\frac{\Lambda(x+\epsilon,\lambda) - \Lambda(x,\lambda)}{\epsilon}\right)^2 + \left(\frac{\Lambda(x,\lambda+\epsilon) - \Lambda(x,\lambda)}{\epsilon}\right)^2}$$

(Since magnitude is always non-negative, optimizing over the squared-magnitude is equivalent to optimizing over the magnitude. Thus, the ``square root" may be omitted from these equations with no expected difference in the results of optimization.)

The critical points of h occur at $x = 1$ and $x = -1$, just as in Λ. Unlike the critical points in Λ, however, the critical points in h occur at local minima, so numerical optimization techniques can be used to find them.

CONCLUSION

In this unit, you have studied how to identify problem which could be solve by langranges multiplier. You studied single and multiple constraints. You have studied the interpretation of lagranges multiplier.You couls solve problems with the use of langranges multiplier.

Summary

In this unit, you have :

i. identified problem which could be solved by langranges multiplier
ii. known single and multiple constraints
iii. known the interpretation of lagrange multiplier
iv. solved problems with the use of langranges multiplier

Problems

Problem 1. Let $h(x, y) = x^2 + .3y^2 + .4y + 1$ be our objective function. (Note that the coefficients are decimals 0 .3 and 0 .4 and not 3 and 4.) Let and the ellipse $g(x, y) = 1$ be our constraint. Find the maximum and the minimum values of $h(x, y)$ subject to $g(x, y) = 1$ following the steps below.

(a) Plot the 3d graph of the function $h(x, y)$, the ellipse $g(x, y) = 1$ in the xy-plane, and the curve on the graph $z = h(x, y)$ corresponding to the values of $h(x, y)$ along the ellipse in one coordinate system. Use a parametric representation of the ellipse that you should know from last semester. How many solutions you will expect the Lagrangian system of equations to have. Explain your reasoning.

(b) Define the Lagrangian function for the optimization problem and set up the corresponding system of equations.

(c) Find solutions to the system using the solve command. Check that you didn't obtain any extraneous solutions. Is the number of solutions what you expected?

(d) Using results of (c), find the minimum and the maximum values of $h(x, y)$ subject to the constraint $g(x, y) = 1$.

TUTOR-MARKED ASSIGNMENT

3. Find the maximum and minimum of $f(x, y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$
4. Find the maximum and minimum values of $f(x, y, z) = xyz$ subject to the constraint $x + y + z = 1$ Assume that $x, y, z \geq 0$
5. Find the maximum and minimum values of $f(x, y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$
6. Find the maximum and minimum of $f(x, y, z) = 4y - 2z$ $f(x, y, z) = 4y - 2z$ subject to the constraints $2x - y - z = 2$ and $x^2 + y^2 = 1$.

REFERENCES/FURTHER READINGS

Bertsekas, Dimitri P. (1999). *Nonlinear Programming* (Second ed.). Cambridge, MA.: Athena Scientific. ISBN 1-886529-00-0.

Vapnyarskii, I.B. (2001), "Lagrange multipliers", in Hazewinkel, Michiel, *Encyclopedia of Mathematics*, Springer, ISBN 978-1556080104, http://www.encyclopediaofmath.org/index.php?title=Lagrange_multipliers.

Lasdon, Leon S. (1970). *Optimization theory for large systems*. Macmillan series in operations research. New York: The Macmillan Company. pp. xi+523. MR337317.

Lasdon, Leon S. (2002). *Optimization theory for large systems* (reprint of the 1970 Macmillan ed.). Mineola, New York: Dover Publications, Inc.. pp. xiii+523. MR1888251.

Hiriart-Urruty, Jean-Baptiste; Lemaréchal, Claude (1993). "XII Abstract duality for practitioners". *Convex analysis and minimization algorithms, Volume II: Advanced theory and bundle methods*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. **306**. Berlin: Springer-Verlag. pp. 136–193 (and Bibliographical comments on pp. 334–335). ISBN 3-540-56852-2.

Lemaréchal, Claude (2001). "Lagrangian relaxation". In Michael Jünger and Denis Naddef. *Computational combinatorial optimization: Papers from the Spring School held in Schloß Dagstuhl, May 15–19, 2000*. Lecture Notes in Computer Science. **2241**. Berlin: Springer-Verlag. pp. 112–156. doi:10.1007/3-540-45586-8_4. ISBN 3-540-42877-1. MRdoi:[http://dx.doi.org/10.1007%2F3-540-45586-8_4 10.1007/3-540-45586-8_4 1900016.[[Digital object identifier|doi]]:[http://dx.doi.org/10.1007%2F3-540-45586-8_4 10.1007/3-540-45586-8_4]].

UNIT 3.0 APPLICATIONS OF LANGRANGES MULTIPLIER

CONTENTS

8.0 INTRODUCTION
9.0 OBJECTIVES
10.0 MAIN CONTENT
  3.1: DEFINITION
11.0 CONCLUSION
12.0 SUMMARY
13.0 TUTOR-MARKED ASSIGNMENT
14.0 REFERENCES/FURTHER READINGS

1.0 Introduction

Optimization problems, which seek to minimize or maximize a real function, play an important role in the real world. It can be classified into unconstrained op timization problems and constrained optimization problems. Many practical uses in science, engineering, economics, or even in our everyday life can be formulated as constrained optimization problems, such as the minimization of the energy of a particle in physics;[1] how to maximize the profit of the investments in economics.[2]In unconstrained problems, the stationary points theory gives the necessary condition to find the extreme points of the objective function *f* (*x*1*; ¢ ¢ ¢ ; xn*). The stationary points are the points where the gradient *rf* is zero, that is each of the partial derivatives is zero. All the variables in *f* (*x*1*; ¢ ¢ ¢ ; xn*)are independent, so they can be arbitrarily set to seek the extreme of *f*. However when it comes to the constrained optimization problems, the arbitration of the variables does not exist. The constrained optimization problems can be formulated into the standard form.

2.0 Objectives

At the end of this unit, you should be able to :

i. Apply the lagranges multiplier on a pringle surface
ii. Apply lagranges multiplier on Economics
iii. Apply lagranges multiplier on control theory
iv. Solve problems with the application of lagrange multiplier

3.0 Main content

There are many cool applications for the Lagrange multiplier method. For example, we will show you how to find the extrema on the world famous Pringle surface. The Pringle surface can be given by the equation

$$f(x,y) = x^2 - y^2$$

Let us bound this surface by the unit circle, giving us a very happy pringle. :) In this case, the boundary would be

$$G(x,y) = x^2 + y^2 - 1$$

The first step is to find the extrema on an unbounded f.

. Economics

Constrained optimization plays a central role in economics. For example, the choice problem for a consumer is represented as one of maximizing a utility function subject to a budget constraint. The Lagrange multiplier has an economic interpretation as the shadow price associated with the constraint, in this example the marginal utility of income.

Control theory

In optimal control theory, the Lagrange multipliers are interpreted as costate variables, and Lagrange multipliers are reformulated as the minimization of the Hamiltonian, in Pontryagin's minimum principle.

Example 1 Find the dimensions of the box with largest volume if the total surface area is 64 cm^2.

We first need to identify the function that we're going to optimize as well as the constraint. Let's set the length of the box to be *x*, the width of the box to be *y* and the height of the box to be *z*. Let's also note that because we're dealing with the dimensions of a box it is safe to assume that *x*, *y*, and *z* are all positive quantities.

We want to find the largest volume and so the function that we want to optimize is given by,

$f(x, y, z) = xyz$ Next we know that the surface area of the box must be a constant 64. So this is the constraint. The surface area of a box is simply the sum of the areas of each of the sides so the constraint is given by,

$$2xy + 2xz + 2yz = 64 \qquad \Rightarrow \qquad xy + xz + yz = 32$$

Note that we divided the constraint by 2 to simplify the equation a little. Also, we get the function $g(x, y, z)$ from this.

$$g(x, y, z) = xy + xz + yz$$

Here are the four equations that we need to solve.

(1) $yz = \lambda(y+z)$ $\qquad (f_x = \lambda g_x)$

(2) $xz = \lambda(x+z)$ $\qquad (f_y = \lambda g_y)$

$xy + xz + yz = 32$ $\qquad (g(x, y, z) = 32)$

(3) $xy = \lambda(x+y)$ $\qquad (f_z = \lambda g_z)$

(4)There are many ways to solve this system. We'll solve it in the following way. Let's multiply equation (1) by *x*, equation (2) by *y* and equation (3) by *z*. This gives,

$xyz = \lambda x(y+z)$

(5) $xyz = \lambda y(x+z) xyz = \lambda y(x+z)$

(6) $xyz = \lambda z(x+y)$

(7)Now notice that we can set equations (5) and (6) equal. Doing this gives,

$$\lambda x(y+z) = \lambda y(x+z)$$

$$\lambda(xy+xz) - \lambda(yx+yz) = 0$$

$$\lambda(xz-yz) = 0 \qquad \Rightarrow \qquad \lambda = 0 \quad \text{or} \quad xz = yz$$

This gave two possibilities. The first, $\lambda = 0$ $\lambda = 0 \lambda = 0$ is not possible since if this was the case equation (1) would reduce to

$$yz = 0 \qquad \Rightarrow \qquad y = 0 \ \text{or} \ z = 0$$

Since we are talking about the dimensions of a box neither of these are possible so we can discount $\lambda = 0$ This leaves the second possibility.

$$xz = yz$$

Since we know that $z \neq 0$ (again since we are talking about the dimensions of a box) we can cancel the *z* from both sides. This gives, $x = y$ (8)

Next, let's set equations (6) and (7) equal. Doing this gives,

$$\lambda y(x+z) = \lambda z(x+y)$$

$$\lambda(yx+yz-zx-zy) = 0$$

$$\lambda(yx-zx) = 0 \qquad \Rightarrow \qquad \lambda = 0 \ \text{or} \ yx = zx$$

As already discussed we know that $\lambda = 0$ won't work and so this leaves,

$$yx = zx$$

We can also say that $x \neq 0$ since we are dealing with the dimensions of a box so we must have,

$$z = y$$

(9)

Plugging equations (8) and (9) into equation (4) we get,

$$y^2 + y^2 + y^2 = 3y^2 = 32 \qquad\qquad y = \pm\sqrt{\frac{32}{3}} = \pm 3.266$$

However, we know that *y* must be positive since we are talking about the dimensions of a box. Therefore the only solution that makes physical sense here is

$$x = y = z = 3.266$$

So, it looks like we've got a cube here.

We should be a little careful here. Since we've only got one solution we might be tempted to assume that these are the dimensions that will give the largest volume. The method of Lagrange Multipliers will give a set of points that will either maximize or minimize a given function subject to the constraint, provided there actually are minimums or maximums.

The function itself, $f(x, y, z) = xyz$ will clearly have neither minimums or maximums unless we put some restrictions on the variables. The only real restriction that we've got is that all the variables must be positive. This, of course, instantly means that the function does have a minimum, zero.

The function will not have a maximum if all the variables are allowed to increase without bound. That however, can't happen because of the constraint,

$$xy + xz + yz = 32$$

Here we've got the sum of three positive numbers (because *x*, *y*, and *z* are positive) and the sum must equal 32. So, if one of the variables gets very large, say *x*, then because each of the products must be less than 32 both *y* and *z* must be very small to make sure the first two terms are less than 32. So, there is no way for all the variables to increase without bound and so it should make some sense that the function, $f(x, y, z) = xyz$, will have a maximum.

This isn't a rigorous proof that the function will have a maximum, but it should help to visualize that in fact it should have a maximum and so we can say that we will get a maximum volume if the dimensions are : $x = y = z = 3.266$.

Notice that we never actually found values for λ in the above example. This is fairly standard for these kinds of problems. The value of λ isn't really important to determining if the point is a maximum or a minimum so often we will not bother with finding a value for it. On occasion we will need its value to help solv

Example 2

Find the maximum and minimum of $f(x, y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$

Solution

This one is going to be a little easier than the previous one since it only has two variables. Also, note that it's clear from the constraint that region of possible solutions lies on a disk of radius $\sqrt{136}$ which is a closed and bounded region and hence by the Extreme Value Theorem we know that a minimum and maximum value must exist.

Here is the system that we need to solve.

$$5 = 2\lambda x$$
$$-3 = 2\lambda y$$
$$x^2 + y^2 = 136$$

Notice that, as with the last example, we can't have $\lambda = 0$ since that would not satisfy the first two equations. So, since we know that $\lambda \neq 0$ we can solve the first two equations for *x* and *y* respectively. This gives,

$$x = \frac{5}{2\lambda} \qquad\qquad y = -\frac{3}{2\lambda}$$

Plugging these into the constraint gives,

$$\frac{25}{4\lambda^2} + \frac{9}{4\lambda^2} = \frac{17}{2\lambda^2} = 136$$

We can solve this for λ

$$\lambda^2 = \frac{1}{16} \qquad \Rightarrow \qquad \lambda = \pm\frac{1}{4}$$

Now, that we know λ we can find the points that will be potential maximums and/or minimums.

If $\lambda = -\frac{1}{4}$ we get, $x = -10$ $\qquad$ $y = 6$

and if

$$x = 10 \qquad y = -6$$

$$x = 10 \qquad y = -6$$

To determine if we have maximums or minimums we just need to plug these into the function. Also recall from the discussion at the start of this solution that we know these will be the minimum and maximums because the Extreme Value Theorem tells us that minimums and maximums will exist for this problem.

Here are the minimum and maximum values of the function.

$$f(-10,6) = -68 \qquad \textbf{Minimum at } (-10,6)$$
$$f(10,-6) = 68 \qquad \textbf{Maximum at } (10,-6)$$

Example 3

- Set up equations for the volume and the cost of building the silo.
- Using the Lagrange multiplier method, find the cheapest way to buld the silo.
- Do these dimensions seem reasonable? Why?

Next, we will look at the cost of building a silo of volume 1000 cubic meters. The curved surface on top of the silo costs $3 per square meter to build, while the walls cost $1 per square meter.

Of course, if all situations were this simple, there would be no need for the Lagrange multiplier method, since there are other methods for solving 2 variable functions that are much nicer. However, with a greater number of variables, the Lagrange multiplier method is much more fun.

For the next example, imagine you are working at the State Fair (since you're so desperate for money that you can't even buy a bagel anymore). You find yourself at the snowcone booth, and your boss, upon hearing that you are good at math, offers you a bonus if you can design the most efficient snowcone. You assume the snowcone will be modelled by a half-ellipsoid perched upon a cone.

Your boss only wants to use 84 square centimeters of paper per cone, and wants to have it hold the maximum amount of snow. This can be represented in 3 variables: r (the radius of the cone), h (the height of the cone), and s (the height of the half-ellipsoid). In order to keep the snow from tumbling off the cone, s cannot be greater than 1.5*r. We have provided hints for the equations if you need them.

CONCLUSION:

In this unit, you should be able to apply the lagranges multiplier on a pringle surface, apply lagranges multiplier on Economics, apply lagranges multiplier on control theory and solve problems with the application of lagrange multiplier

SUMMARY

The Lagrange multipliers method is a very sufficient tool for the nonlinear optimization problems,which is capable of dealing with both equality constrained and inequality constrained nonlinear optimization problems.Many computational programming methods, such as the barrier and interior point method, penalizing and augmented Lagrange method,The lagrange multipliers method and its extended methods are widely applied in science, engineering, economics and our everyday life.

TUTOR-MARKED ASSIGNMENT

1. Find the dimensions of the box with largest volume if the total surface area is 64 cm^2.

2.Consider two curves on the xy-plane: y = e^x and y = $-(x-2)^2$. Find two points (x,y), (X,Y) on each of the two curves, respectively, whose distance apart is as small as possible. Use the method of Lagrange multipliers. Make a graph that illustrates your solution

3. Find the maximum and minimum values of $f(x,y,z) = xyz$ subject to the constraint $x + y + z = 1$ Assume that $x, y, z \geq 0$

4.Find the maximum and minimum values of $f(x,y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$

5.Find the maximum and minimum of $f(x,y,z) = 4y - 2z$ subject to the constraints $2x - y - z = 2$ and $x^2 + y^2 = 1$

REFERENCE

[1] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Elsevier academic, 2005), 5th ed.
[2] N. Scho¯eld,*Mathematical Methods in Economics and Social Choice* (Springer, 2003), 1st ed.
[3] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, 1982), 1st ed.
[4] R. Courant, *Di®erential and Integral Calculus* (Interscience Publishers, 1937), 1st ed.
[5] D. P. Bertsekas, *Nonlinear Programming* (Athena Scient¯c, 1999), 2nd ed.
[6] M. Crow, *Computational Methods for Electric Power Systems* (CRC, 2003), 1st ed.

MODULE 8 THE JACOBIANS

UNIT 1: JACOBIANS

UNIT 2: JACOBIAN DETERMINANTS

UNIT 3: APPLICATIONS OF JACOBIAN

UNIT 1 JACOBIAN

CONTENTS

1.0 INTRODUCTION
4 OBJECTIVES
3.0 MAIN CONTENT

3.1 Recognise the Jacobian rule

3.2 How to use the Jacobian

5 CONCLUSION
6 SUMMARY
7 TUTOR-MARKED ASSIGNMENT
8 REFERENCES/FURTHER READINGS

1.0 INTRODUCTION

Jacobian

The Jacobian of functions $f_i(x_1,x_2, \ldots,x_n), i= 1, 2, \ldots,n$, of real variables x_i is the determinant of the matrix whose ith row lists all the first-order partial derivatives of the function $f_i(x_1,x_2, \ldots,x_n)$. Also known as Jacobian determinant.

(or functional determinant), a determinant $|a_{ik}|_1^n$ with elements $a_{ik} = \partial y_i/\partial x_k$ where $y_i = f_i(x_1, \ldots, x_n)$, $1 \le i \le n$, are functions that have continuous partial derivatives in some region Δ. It is denoted by

$$\frac{D(y_1, \ldots, y_n)}{D(x_1, \ldots, x_n)}$$

The Jacobian was introduced by K. Jacobi in 1833 and 1841. If, for example, $n = 2$, then the system of functions

(1) $y_1 = f_1(x_1, x_2)$ $y_2 = f_2(x_1, x_2)$

defines a mapping of a region Δ, which lies in the plane x_1x_2, onto a region of the plane y_1y_2. The role of the Jacobian for the mapping is largely analogous to that of the derivative for a function of a single variable. Thus, the absolute value of the Jacobian at some point M is equal to the local factor by which areas at the point are altered by the mapping; that is, it is equal to the limit of the ratio of the area of the image of the neighborhood of M to the area of the neighborhood as the dimensions of the neighborhood approach zero. The Jacobian at M is positive if mapping (1) does not change the orientation in the neighborhood of M, and negative otherwise.

OBJECTIVE

At the end of this unit, you should be able to :

recognise the Jacobian rule

how to use rhe Jacobian

MAIN CONTENT

If the Jacobian does not vanish in the region Δ and if $\varphi(y_1, y_2)$ is a function defined in the region Δ_1 (the image of Δ), then

$$\iint_{\Delta_1} \Phi(y_1, y_2)\, dy_1 dy_2 =$$
$$\iint_{\Delta} \Phi[f_1(x_1, x_2), f_2(x_1, x_2)] \left| \frac{D(y_1, y_2)}{D(x_1, x_2)} \right| dx_1 dx_2$$

(the formula for change of variables in a double integral). An analogous formula obtains for multiple integrals. If the Jacobian of mapping (1) does not vanish in region Δ, then there exists the inverse mapping

$x_1 = \psi(y_1, y_2)$ $x_2 = \psi_2(y_1, y_2)$

and

$$\frac{D(x_1, x_2)}{D(y_1, y_2)} = 1 : \frac{D(y_1, y_2)}{D(x_1, x_2)}$$

(an analogue of the formula for differentiation of an inverse function). This assertion finds numerous applications in the theory of implicit functions.

In order for the explicit expression, in the neighborhood of the point $M(x_1^{(0)}, \ldots, x_n^{(0)}, y_1, \ldots, y_m^{(0)})$, of the functions $y_1, \ldots y_m$ that are implicitly defined by the equations

(2) $F_k(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$ $I \le k \le m$

to be possible, it is sufficient that the coordinates of M satisfy equations (2), that the functions F_k have continuous partial derivatives, and that the Jacobian

$$\frac{D(F_1, \ldots, F_m)}{D(y_1, \ldots, y_m)}$$

be nonzero at M. The Jacobian is been classified into two :

The Jacobian matrix and the Jacobian determinant.

Examples

1.Let F: $R^2 \rightarrow R^2$ be the mapping defined by

F(x,y) = $\begin{pmatrix} x^2 + y^2 \\ e^{xy} \end{pmatrix} \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$

Find the Jacobian matrix $J_f(p)$ for p = (1,1)

The Jacobian matrix at an arbitrary point (x,y) is

$$\begin{pmatrix} \frac{df}{dx} & \frac{df}{dy} \\ \frac{dy}{dx} & \frac{dg}{dy} \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ ye^{xy} & xe^{xy} \end{pmatrix}$$

Hence when x=1 ,y=1 ,we find J_f (1, 1) = $\begin{pmatrix} 2 & 2 \\ e & e \end{pmatrix}$

2.Let F : $R^2 \rightarrow R^3$ be the mapping defined by

$$F(x,y) = \begin{pmatrix} xy \\ \sin x \\ x^2 y \end{pmatrix}$$

Find J_F (P) at the point P = $(\Pi, \frac{\Pi}{2})$.

The Jacobian matrix at an arbitrary point(x,y) is

$$J_F(x, y) = \begin{pmatrix} y & x \\ \cos x & 0 \\ 2xy & x^2 \end{pmatrix}$$

Hence, $J_F(\Pi, \frac{\Pi}{2}) = \begin{pmatrix} \frac{\Pi}{2} & \Pi \\ -1 & 0 \\ \Pi^2 & \Pi^2 \end{pmatrix}$

CONCLUSION

In this unit, you have been able to recognise the Jacobian rule and how to use the formular.

SUMMARY

In this unit, you have studied the basic concept of Jacobian with the identification of the formular below as :

$$\frac{D(F_1, \ldots, F_m)}{D(y_1, \ldots, y_m)}$$

be nonzero at *M*.

Tutor – Marked Assignment

1.Define the Jacobian matrix and the Jacobian determinant.

2.Compute the Jacobian matrix of the following cases below :

a. F(x,y) = (x+y, $\chi^2 y$)
b. F(x,y) = (sinx,cosxy)
c. F(x,y,z) = (xyz, χ^2 z)

REFERENCE

D.K. Arrowsmith and C.M. Place, *Dynamical Systems*, Section 3.3, Chapman & Hall, London, 1992. ISBN 0-412-39080-9.

Taken from http://www.sjcrothers.plasmaresources.com/schwarzschild.pdf - On the Gravitational Field of a Mass Point according to Einstein's Theory by K. Schwarzschild - arXiv:physics/9905030 v1 (text of the original paper, in Wikisource).

UNIT 2 : JACOBIAN DETERMINANT

CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT

3.1 apply the Jacobean concept
3.2 know the Jacobean matrix
3.3 apply the inverse transformation
3.4 solve problems on Jacobean determinant

4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

The Jacobian of functions $f_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$, of real variables x_i is the determinant of the matrix whose ith row lists all the first-order partial derivatives of the function $f_i(x_1, x_2, \ldots, x_n)$. Also known as Jacobian determinant.

In vector calculus, the **Jacobian matrix** : is the matrix of all first-order partial derivatives of a vector- or scalar-valued function with respect to another vector. Suppose $F : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is a function from Euclidean n-space to Euclidean m-space. Such a function is given by m real-valued component functions, $y_1(x_1,...,x_n)$, ..., $y_m(x_1,...,x_n)$. The partial derivatives of all these functions (if they exist) can be organized in an m-by-n matrix, the Jacobian matrix J of F, as follows:

$$J = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

This matrix is also denoted by $J_F(x_1, \ldots, x_n)$ and $\dfrac{\partial(y_1, \ldots, y_m)}{\partial(x_1, \ldots, x_n)}$. If $(x_1,...,x_n)$ are the usual orthogonal Cartesian coordinates, the i th row ($i = 1, ..., n$) of this matrix corresponds to the gradient of the i^{th} component function y_i: (∇y_i). Note that some books define the

Jacobian as the transpose of the matrix given above.

The **Jacobian determinant** (often simply called the **Jacobian**) is the determinant of the Jacobian matrix (if $m = n$).

These concepts are named after the mathematician Carl Gustav Jacob Jacobi.

OBJECTIVE

After reading through this unit, you should be able to :

i. apply the jacobian concept
ii. know the Jacobian matrix
iii. apply the inverse transformation
iv. solve problems on Jacobian determinant

MAIN CONTENT

Jacobian matrix

The Jacobian of a function describes the orientation of a tangent plane to the function at a given point. In this way, the Jacobian generalizes the gradient of a scalar valued function of multiple variables which itself generalizes the derivative of a scalar-valued function of a scalar. Likewise, the Jacobian can also be thought of as describing the amount of "stretching" that a transformation imposes. For example, if $(x_2,y_2) = f(x_1,y_1)$ is used to transform an image, the Jacobian of f, $J(x_1,y_1)$ describes how much the image in the neighborhood of (x_1,y_1) is stretched in the x and y directions.

If a function is differentiable at a point, its derivative is given in coordinates by the Jacobian, but a function doesn't need to be differentiable for the Jacobian to be defined, since only the partial derivatives are required to exist.

The importance of the Jacobian lies in the fact that it represents the best linear approximation to a differentiable function near a given point. In this sense, the Jacobian is the derivative of a multivariate function.

If $\mathbf{p}$ is a point in $\mathbf{R}^n$ and F is differentiable at $\mathbf{p}$, then its derivative is given by $J_F(\mathbf{p})$. In this case, the linear map described by $J_F(\mathbf{p})$ is the best linear approximation of F near the point $\mathbf{p}$, in the sense that

$$F(\mathbf{x}) = F(\mathbf{p}) + J_F(\mathbf{p})(\mathbf{x} - \mathbf{p}) + o(\|\mathbf{x} - \mathbf{p}\|)$$

for $\mathbf{x}$ close to $\mathbf{p}$ and where o is the little o-notation (for $x \to p$) and $\|\mathbf{x} - \mathbf{p}\|$ is the distance between $\mathbf{x}$ and $\mathbf{p}$.

In a sense, both the gradient and Jacobian are "first derivatives" — the former the first derivative of a *scalar function* of several variables, the latter the first derivative of a *vector function* of several variables. In general, the gradient can be regarded as a special version of the Jacobian: it is the Jacobian of a scalar function of several variables.

The Jacobian of the gradient has a special name: the Hessian matrix, which in a sense is the "second derivative" of the scalar function of several variables in question.

Inverse

According to the inverse function theorem, the matrix inverse of the Jacobian matrix of an invertible function is the Jacobian matrix of the *inverse* function. That is, for some function $F : \mathbf{R}^n \to \mathbf{R}^n$ and a point p in $\mathbf{R}^n$,

$$J_{F^{-1}}(F(p)) = [J_F(p)]^{-1}.$$

It follows that the (scalar) inverse of the Jacobian determinant of a transformation is the Jacobian determinant of the inverse transformation.

Examples

Example 1. The transformation from spherical coordinates (r, θ, φ) to Cartesian coordinates (x_1, x_2, x_3) , is given by the function $F : \mathbf{R}^+ \times [0,\pi] \times [0,2\pi) \to \mathbf{R}^3$ with components:

$$\begin{aligned} x_1 &= r\ \sin\theta\ \cos\phi \\ x_2 &= r\ \sin\theta\ \sin\phi \\ x_3 &= r\ \cos\theta. \end{aligned}$$

The Jacobian matrix for this coordinate change is

$$J_F(r,\theta,\phi) = \begin{bmatrix} \dfrac{\partial x_1}{\partial r} & \dfrac{\partial x_1}{\partial \theta} & \dfrac{\partial x_1}{\partial \phi} \\ \dfrac{\partial x_2}{\partial r} & \dfrac{\partial x_2}{\partial \theta} & \dfrac{\partial x_2}{\partial \phi} \\ \dfrac{\partial x_3}{\partial r} & \dfrac{\partial x_3}{\partial \theta} & \dfrac{\partial x_3}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \sin\theta\ \cos\phi & r\ \cos\theta\ \cos\phi & -r\ \sin\theta\ \sin\phi \\ \sin\theta\ \sin\phi & r\ \cos\theta\ \sin\phi & r\ \sin\theta\ \cos\phi \\ \cos\theta & -r\ \sin\theta & 0 \end{bmatrix}.$$

The determinant is $r^2 \sin\theta$. As an example, since $dV = dx_1\ dx_2\ dx_3$ this determinant implies that the differential volume element $dV = r^2 \sin\theta\ dr\ d\theta\ d\phi$. Nevertheless this determinant varies with coordinates. To avoid any variation the new coordinates can be defined as $w_1 = \dfrac{r^3}{3},\ w_2 = -\cos\theta,\ w_3 = \phi.$[2] Now the determinant equals to 1 and volume element becomes $r^2 dr\ \sin\theta\ d\theta\ d\phi = dw_1 dw_2 dw_3$.

Example 2. The Jacobian matrix of the function $F : \mathbf{R}^3 \to \mathbf{R}^4$ with components

$$y_1 = x_1$$

$$y_2 = 5x_3$$
$$y_3 = 4x_2^2 - 2x_3$$
$$y_4 = x_3 \sin(x_1)$$

is

$$J_F(x_1, x_2, x_3) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \\ \frac{\partial y_4}{\partial x_1} & \frac{\partial y_4}{\partial x_2} & \frac{\partial y_4}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 8x_2 & -2 \\ x_3 \cos(x_1) & 0 & \sin(x_1) \end{bmatrix}.$$

This example shows that the Jacobian need not be a square matrix.

Example 3.

$$x = r \cos \phi;$$
$$y = r \sin \phi.$$

$$J(r, \phi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \frac{\partial (r \cos \phi)}{\partial r} & \frac{\partial (r \cos \phi)}{\partial \phi} \\ \frac{\partial (r \sin \phi)}{\partial r} & \frac{\partial (r \sin \phi)}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{bmatrix}$$

The Jacobian determinant is equal to *r*. This shows how an integral in the Cartesian coordinate system is transformed into an integral in the polar coordinate system:

$$\iint_A dx\, dy = \iint_B r\, dr\, d\phi .$$

Example 4. The Jacobian determinant of the function $F : \mathbf{R}^3 \to \mathbf{R}^3$ with components

$$y_1 = 5x_2$$
$$y_2 = 4x_1^2 - 2\sin(x_2 x_3)$$
$$y_3 = x_2 x_3$$

is

$$\begin{vmatrix} 0 & 5 & 0 \\ 8x_1 & -2x_3 \cos(x_2 x_3) & -2x_2 \cos(x_2 x_3) \\ 0 & x_3 & x_2 \end{vmatrix} = -8x_1 \cdot \begin{vmatrix} 5 & 0 \\ x_3 & x_2 \end{vmatrix} = -40 x_1 x_2 .$$

From this we see that F reverses orientation near those points where x_1 and x_2 have the same sign; the function is locally invertible everywhere except near points where $x_1 = 0$ or $x_2 = 0$. Intuitively, if you start with a tiny object around the point (1,1,1) and apply F to that object, you will get an object set with approximately 40 times the volume of the original one.

CONCLUSION

In this unit, you have studied the application of the Jacobian concept. You have known the Jacobian matrix and the application of the inverse transformation of Jacobian determinants. You have solved problems on Jacobian determinant.

SUMMARY

In this unit ;

i you have studied application of the Jacobian concept

ii you have known the Jacobian matrix

iii you have known the inverse transformation of Jacobian determinant

iv you have solve problems on Jacobian determinant such as ;

. The Jacobian matrix of the function $F : \mathbf{R}^3 \to \mathbf{R}^4$ with components

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= 5x_3 \\ y_3 &= 4x_2^2 - 2x_3 \\ y_4 &= x_3 \sin(x_1) \end{aligned}$$

is

$$J_F(x_1, x_2, x_3) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \\ \frac{\partial y_4}{\partial x_1} & \frac{\partial y_4}{\partial x_2} & \frac{\partial y_4}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 8x_2 & -2 \\ x_3 \cos(x_1) & 0 & \sin(x_1) \end{bmatrix}.$$

This example shows that the Jacobian need not be a square matrix.

Tutor-Marked Assignment

1.In each of the following cases, compute the Jacobian matrix of F, and evaluate at the following points;

F(x,y) = (sinx,cosxy) at points (1,2)

F(x,y,z) = (sinxyz,xz) at points (2,-1,-1)

F(x,y,z) =(xz,xy,yz) at points (1,1,-1)

2. Transform the following from spherical coordinates (r,θ,ϕ) to Cartesian coordinate (x_1, x_2, x_3) by the function F: $R^+ \times (0,\Pi) \times (0,2\Pi) \to R^3$ with components :

$r_1 = r \tan\theta \cos\theta$

$r_2 = r \sin\theta \tan\theta$

$r = r \sin\theta_1$

3.The Jacobian matrix of the function F: $R^3 \to R^4$ with components

$y_1 = x_2$

$y_2 = 4x_1$

$y_3 = 5x_2^2 - 4x_3$

$y_4 = x_1 \sin x_3$

4.The Jacobian matrix of the function F: $R^3 \to R^3$ with components

$y_1 = 4x_1^2 - 3\sin x_2 x_3$

$y_2 = 3x_2$

$y_3 = x_2 3 x_3$

The Jacobian matrix of the function F: $R^3 \to R^3$ with components

x=rtanϕ

y=rcosϕ

REFERENCES

Kudriavtsev, L. D. *Matematicheskii analiz*, 2nd ed., vol. 2. Moscow, 1973.
Il'in, V. A., and E. G. Pozniak. *Osnovy matematicheskogo analiza*, 3rd ed.. part I. Moscow, 1971.

D.K. Arrowsmith and C.M. Place, *Dynamical Systems*, Section 3.3, Chapman & Hall, London, 1992.

UNIT 3 APPLICATION OF JACOBIAN

CONTENTS

1.0 INTRODUCTION
2.0 OBJECTIVES
3.0 MAIN CONTENT

3.1 apply the jacobian concept
3.2 know the Jacobian matrix
3.3 apply the inverse transformation
3.4 solve problems on Jacobian determinant

4.0 CONCLUSION
5.0 SUMMARY
6.0 TUTOR-MARKED ASSIGNMENT
7.0 REFERENCES/FURTHER READINGS

INTRODUCTION

If $m = n$, then F is a function from m-space to n-space and the Jacobian matrix is a square matrix. We can then form its determinant, known as the **Jacobian determinant**. The Jacobian determinant is sometimes simply called "the Jacobian."

OBJECTIVE

MAIN CONTENT

Dynamical systems

Consider a dynamical system of the form $x' = F(x)$, where x' is the (component-wise) time derivative of x, and $F : \mathbf{R}^n \to \mathbf{R}^n$ is continuous and differentiable. If $F(x_0) = 0$, then x_0 is a stationary point (also called a fixed point). The behavior of the system near a stationary point is related to the eigenvalues of $J_F(x_0)$, the Jacobian of F at the stationary point.Specifically, if the eigenvalues all have a negative real part, then the system is stable in the operating point, if any eigenvalue has a positive real part, then the point is unstable.

.

The Jacobian determinant at a given point gives important information about the behavior of F near that point. For instance, the continuously differentiable function F is invertible near a point $\mathbf{p} \in \mathbf{R}^n$ if the Jacobian determinant at $\mathbf{p}$ is non-zero. This is the inverse function theorem. Furthermore

if the Jacobian determinant at $\mathbf{p}$ is positive, then F preserves orientation near $\mathbf{p}$; if it is negative, F reverses orientation. The absolute value of the Jacobian determinant at $\mathbf{p}$ gives us the factor by which the function F expands or shrinks volumes near $\mathbf{p}$; this is why it occurs in the general substitution rule.

Uses

The Jacobian determinant is used when making a change of variables when evaluating a multiple integral of a function over a region within its domain. To accommodate for the change of coordinates the magnitude of the Jacobian determinant arises as a multiplicative factor within the integral. Normally it is required that the change of coordinates be done in a manner which maintains an injectivity between the coordinates that determine the domain. The Jacobian determinant, as a result, is usually well defined.

Newton's method

A system of coupled nonlinear equations can be solved iteratively by Newton's method. This method uses the Jacobian matrix of the system of equations

CONCLUSION

In this unit, you have known the application of Jacobian concept. You have studied the application of Jacobian matrix. You have used Jacobian in in the application of inverse transformation and have also solved problems on Jacobian determinant.

SUMMARY

In this unit, you have studied the following :

Application of the Jacobian concept

Application of Jacobian on the Jacobian matrix

Application of the Jacobian on the inverse concept

Application of the Jacobian to solve problems on Jacobian determinant

TUTOR – MARK ASSIGNMENTS

1.Find the Jacobian determinant of the map below, and determine all the points where the Jacobian determinants is equal to zero(0).

a.F(x,y,z) = (xy,y,xz)

b.F(x,y) = (e^{xy} ,x)

c.F(x,y) = (xy, χ^2)

2.The transformation from spherical coordinates (r, θ, φ) to Cartesian coordinates (x_1, x_2, x_3) , is given by the function $F : \mathbf{R}^+ \times [0,\pi] \times [0,2\pi) \to \mathbf{R}^3$ with components:

$$x_1 = r \sin\theta \cos\phi$$
$$x_2 = r \sin\theta \sin\phi$$
$$x_3 = r \cos\theta.$$

3.The Jacobian determinant of the function $F : \mathbf{R}^3 \to \mathbf{R}^4$ with components

$$y_1 = x_1$$
$$y_2 = 5x_3$$

$$y_3 = 4x_2^2 - 2x_3$$
$$y_4 = x_3 \sin(x_1)$$

4.The Jacobian determinant of the function $F : \mathbf{R}^3 \to \mathbf{R}^4$ with components

$$x = r \cos \phi;$$
$$y = r \sin \phi.$$

5.The Jacobian determinant of the function $F : \mathbf{R}^3 \to \mathbf{R}^3$ with components

$$y_1 = 5x_2$$
$$y_2 = 4x_1^2 - 2\sin(x_2 x_3)$$
$$y_3 = x_2 x_3$$

REFFERENCE

Gradshteyn, I. S. and Ryzhik, I. M. "Jacobian Determinant." §14.313 in *Tables of Integrals, Series, and Products, 6th ed.* San Diego, CA: Academic Press, pp. 1068-1069, 2000.

Kaplan, W. *Advanced Calculus, 3rd ed.* Reading, MA: Addison-Wesley, pp. 98-99, 123, and 238-245, 1984.

Simon, C. P. and Blume, L. E. *Mathematics for Economists.* New York: W. W. Norton, 1994.

D.K. Arrowsmith and C.M. Place, *Dynamical Systems*, Section 3.3, Chapman & Hall, London, 1992. ISBN 0-412-39080-9