DIRECTORATE OF DISTANCE EDUCATION

UNIVERSITY OF NORTH BENGAL

MASTERS OF SCIENCE-MATHEMATICS

SEMESTER -I

ANALYSIS OF SEVERAL VARIABLES

DEMATH1SCORE3

BLOCK-1

UNIVERSITY OF NORTH BENGAL

Postal Address:

The Registrar,

University of North Bengal,

Raja Rammohunpur,

P.O.-N.B.U.,Dist-Darjeeling,

West Bengal, Pin-734013,

India.

Phone: (O) +91 0353-2776331/2699008

Fax:(0353) 2776313, 2699001

Email: regnbu@sancharnet.in ; regnbu@nbu.ac.in

Wesbsite: www.nbu.ac.in

First Published in 2019

All rights reserved. No Part of this book may be reproduced or transmitted,in any form or by any means, without permission in writing from University of North Bengal. Any person who does any unauthorised act in relation to this book may be liable to criminal prosecution and civil claims for damages. This book is meant for educational and learning purpose. The authors of the book has/have taken all reasonable care to ensure that the contents of the book do not violate any existing copyright or other intellectual property rights of any person in any manner whatsoever. In the even the Authors has/ have been unable to track any source and if any copyright has been inadvertently infringed, please notify the publisher in writing for corrective action.

FOREWORD

The Self-Learning Material (SLM) is written with the aim of providing simple and organized study content to all the learners. The SLMs are prepared on the framework of being mutually cohesive, internally consistent and structured as per the university’s syllabi. It is a humble attempt to give glimpses of the various approaches and dimensions to the topic of study and to kindle the learner’s interest to the subject

We have tried to put together information from various sources into this book that has been written in an engaging style with interesting and relevant examples. It introduces you to the insights of subject concepts and theories and presents them in a way that is easy to understand and comprehend.

We always believe in continuous improvement and would periodically update the content in the very interest of the learners. It may be added that despite enormous efforts and coordination, there is every possibility for some omission or inadequacy in few areas or topics, which would definitely be rectified in future.

We hope you enjoy learning from this book and the experience truly enrich your learning and help you to advance in your career and future endeavours.

ANALYSIS OF SEVERAL VARIABLES

BLOCK 1

BLOCK 2

BLOCK 1 ANALYSIS OF SEVERAL VARIABLES

Introduction to the Block

In this block we will go through In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables.

Unit I Deals with Topology

Unit II Deals with The Space Of Linear Transformations From R^m To R^n

Unit III Deals with Local Properties Of Continuous Functions

Unit IV Deals with The Basic Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables

Unit V Deals with Transition To The Case Of A Relation

Unit VI Deals withTaylor's Theorem

Unit VII Deals with A Sufficient Condition For A Constrained Extremum

UNIT - 1: TOPOLOGY

STRUCTURE

1.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about Topology

Learn, Understand about Topological spaces

Learn, Understand about Constructions with topological spaces

Learn, Understand about Properties of topological spaces

Learn, Understand about One variable calculus

Learn, Understand about The differential calculus of functions of several variables

Notes

1.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Topology, Metric Spaces, Topological spaces, Constructions with topological spaces, Properties of topological spaces, One variable calculus, The differential calculus of functions of several variables

1.2 TOPOLOGY

METRIC SPACES

A Map f R^M M R between Euclidean Spaces is Continuous iff

Vx G X Ve > 34 > 0 Vy G X d(x, y)< 4 ^ d(f(x), f(y))< e

Where d(x, y)=||x - y\ \=V(xi - yi)^2 4….4(x_{ra} - y,,)^2 G R>o

is the Euclidean distance between two points x, y in R^n.

Example(Examples of continuous maps.)

1. The addition map a: R^2 M R, x=(x_i, x_2)M $x_i + x_2$;
2. The multiplication map m: R^2 M R, x=(x_i, x_2)M $x_i x_2$;

The proofs that these maps are continuous are simple estimates that you probably remember from calculus. Since the continuity of all the maps we'll look at in these notes is proved by expressing them in terms of the maps a and m, we include the proofs of continuity of a and m for completeness.

Proof. To prove that the addition map a is continuous, suppose x=(x^ x_2)E R^2 and e > 0 are given. We claim that for 8 := e /2 and y=(yi, y_2)E R^2 with d(x, y)<8 we have d(a(x), a(y)) < e and hence a is a continuous function. To prove the claim, we note that d(x, y)=V| x i - yi | 2 + | x 2 - y2 | 2

and hence $|x_1 — y_1| < d(x, y)$, $|x_1 — y_1| < d(x, y)$. It follows that d(a(x), a(y))=|a(x) — a(y)|=

|xi + x_2 — yi — y2 1 < |xi — yi | + |x2 — y2 1 < 2d(x, y)<

28=e.

To prove that the multiplication map m is continuous, we claim that for 8 := min{ 1, e /(|xi| + |x21 + 1)} and y=(y_i, y_2)E R^2 with d(x, y)<8 we have d(m(x), m(y)) < e and hence m is a continuous function. The claim follows from the following estimates:

d(m(y), m(x))=|yi y2 — x^|=|yi y2 — xi y2 + xi y2 — xi x2 |

< |yi y2 — xi y2 | + |xi y2 — xi x2 | =|yi — xi ||y2 | + |xi ||y2 —

x2 |

< d(x, y)(|y2| + |xi|)< d(x, y)(|x2| + |y2 — x2| + |xi|)

< d(x, y)(|xi | + |x2 1 + 1)< 8(|xi | + |x2 1 + 1)< e

Theorem. The function d: R^n x R^n — $R>_0$ has the following properties:

1. d(x, y)=0 if and only if x=y;
2. d(x, y)=d(y, x)(symmetry);
3. d(x, y)< d(x, z)+ d(z, y)(triangle inequality)

Definition. A metric space is a set X equipped with a map

d: X x X —y R> o with properties(1)-(3)above. A map f: X — Y between metric spaces X, Y is continuous if condition is satisfied. an isometry if d(f(x), f(y))=d(x, y)for all x, y E X;

Two metric spaces X, Y are homeomorphic(resp. isometric)if there are continuous maps(resp. isometries)f: X — Y and g: Y — X which are inverses of each other.

Example . An important class of examples of metric spaces are subsets of R^n.

1. The n-disk D^n := { x E R^n||x| < 1 } C R^n, and Dn := { x E R^n||x| < r }, the n-disk of radius r > 0. The dilation map D^n —► D^n x ^ rx

is a homeomorphism between D^n and Df with inverse given by multiplication by 1 /r. However, these two metric spaces are not isometric for r =1. To observe this, define the diameter diam(X)of a metric space X by diam(X):= sup{ d(x, y)| x, y E X } E $R>_0$ U { ^ }.

For example, diam(Df)=2r. It is easy to observe that if two metric spaces X, Y are isometric, then their diameters agree. In particular, the disks Df and Df are not isometric unless r=r'.

2. The n-sphere S^n := { x E R^{n+1}||x|=1 } C R^{n+1}.

3. The torus T={ v E R^3|d(v, C)=r } for 0 < r < 1. Here

C={(x, y, 0)| $x^2 + y^2$=1 } C R^3 is the unit circle in xy-plane, and d(v, C)=$\inf_{weC}$ d(v, w)is the distance between v and C.

4. The general linear group

GL_n(R)={ vector space isomorphisms f: R^n ^ R^n }

^^ {(vi, ..., vn)| vi E R^n, det(vi, ..., vn)=0 }

= { invertible n x n-matrices } C R^n x — ■ x R"=R^n

Here we think of(v_1, ..., v_n)as an n x n-matrix with column vectors v_i, and the bijection is the usual one in linear algebra that sends a linear map f: R^n ^ R^n to the matrix(f(e_1), ..., f(e_n)) whose column vectors are the images of the standard basis elements e_i E R^n.

5. The special linear group

SLn(R)={(v1, . .., vn)| vi E R^n, det(v1, . .., v, $_n$)=1 } C R^n

6. The orthogonal group

O(n)={ linear isometries f: R^n ^ R^n }

= {(v_1, ..., v_n)| v_i E R^n, v_i's are orthonormal } C R^n

Collection of vectors v_i E R^n is orthonormal if $|v_i|$=1 for all i, and v_i is perpendicular to vj for i=j.

The special orthogonal group

SO(n)=((v_f, ..., v_n)e O(n)| det(v_f, ..., v_n)=1 } C R^n

7. The Stiefel manifold

$V_k(R^n)$=(linear isometries f: R^k m R^n }

=((v_1, ..., v_k)| v_i e R^n, Vj's are orthonormal } C R^{kn}

Example . The following maps between metric spaces are continuous. While it is possible to prove their continuity using the definition of continuity, it will be much simpler to prove their continuity by 'building' these maps using compositions and products from the continuous maps a and Every polynomial function f: R^n m R is continuous is of the form f(xi, ...,Xn)=Z)$_{il}$, ..., $_{i}$, aq, ..., i_n xf x^ for a^, ..., ^ e R.

Let $M_{nxn}(R)=R^{n2}$ be the set of n x n matrices. Then the map

Mnxn(R)X Mnxn(R) -- M M, $_n$xn(R)(A, B)M AB

given by matrix multiplication is continuous. Here we use the fact that a map to the product $M_{nxn}(R)=R^n$=R x • • • x R is continuous if and only if each component map is continuous and each matrix entry of AB is a polynomial and hence a continuous function of the matrix entries of A and B. Restricting to the invertible matrices GL_n(R)C M_{nxn}(R), we observe that the multiplication map GL_n(R)x GL_n(R) —m GL_n(R)is continuous. The same holds for the subgroups SO(n)C O(n)C GL_n(R).

The map GL_n(R)m GL_n(R), A A^{-1} is continuous for the subgroups of GL_n(R). The Euclidean metric on

R^n given by d(x, y)=y /(x_f — $y_f)^2$ + • • • +(x_n — $y_n)^2$ for x, y e R^n

is not the only reasonable metric on R^n. Another metric on R^n is given by

$$d_i(x, y) = 5\wedge |x_i - y_i|.$$

The question arises whether it can happen that a map f: R^n m R^n is continuous with respect to one of these metrics, but not with respect to the other. To observe that this doesn’t happen, it is useful to characterize continuity of a map f: X m Y between metric spaces X, Y in a way that involves the metrics on X and Y.
This alternative characterization will be based on the following notion of "open subsets" of a metric space.

Definition. Let X be a metric space. A subset U C X is open if for every point x E U there is some e > 0 such that $B_e(x)$C U. Here $B_e(x)$={ y E X|d(y, x)< e } is the ball of radius e around x.

To illustrate this, Lets look at examples of subsets of R^n equipped with the Euclidean metric. The subset D™={ v E R^n |||v|| < r } C R^n is not open, since for a point v E Df with ||v||=r any open ball $B_e(v)$with center v will contain points not in Df. By contrast, the subset $B_r(0)$C R^n is open, since for any x E $B_r(0)$the ball B$(x)of radius 8=r — ||x||
is contained in $B_r(0)$, since for y E B$(x)by the triangle inequality we have

d(y, 0)< d(y, x)+ d(x, 0)<8 + ||x||=(r — ||x||)+ ||x||=r.

Theorem. A map f: X ^ Y between metric spaces is continuous if and only if $f^{-1}(V)$
is an open subset of X for every open subset V C Y.

Corollary. If f: X ^ Y and g: Y ^ Z are continuous maps, then so it their composition g o f: X ^ Z.

Exercise . Let that d, d' are two metrics on a set X which are equivalent in the sense that there are constants C, C' > 0 such that d(x, y)< Cd_1(x, y)and $d_1(x, y)< C'd(x, y)$for all x, y E X. Show that a subset U C X is open with respect to d if and only if it is open with respect to d'.

Show that the Euclidean metric d and the metric on R^n are equivalent. This shows in particular that a map f: R^n ^ R^n is continuous w.r.t. d if and only if it is continuous w.r.t. d1 .

1.3 TOPOLOGICAL SPACES

To define continuity of maps between metric spaces in terms of the open subsets of these metric space instead of the original. In fact, we can go one step further, forget about the metric on a set X altogether, and just consider a collection T of subsets of X that we declare to be "open". The next result summarizes the basic properties of open subsets of a metric space X, which then motivates the restrictions that we wish to put on such collections T.

Theorem. Open subsets of a metric space X have the following properties.

X and 0 are open.
Any union of open sets is open.
The intersection of any finite number of open sets is open.

Definition. A topological space is a set X together with a collection T of subsets of X, known open sets which are required to satisfy conditions(i), (ii)and(iii)of the Theorem above. The collection T is known a topology on X. The sets in T are known the open sets, and their complements in X are known closed sets. A subset of X can be neither closed nor open, either closed or open, or both.

A map f: X ^ Y between topological spaces X, Y is continuous if the inverse image f^{-1}(V)of every open subset V C Y is an open subset of X.

It is easy to observe that the composition of continuous maps is again continuous.

Examples of topological spaces.

Let X be a metric space, and T the collection of those subsets of X that are unions of balls $B_e(x)$in X(i.e., the subsets which are open in the sense of Definition. Then T is a topology on X, the metric topology.

Let X be a set. Then T={ all subsets of X } is a topology, the discrete topology. We note that any map f: X ^ Y to a topological space Y is continuous. We will observe later that the only continuous maps R^n ^ X are the constant maps.

Let X be a set. Then T={ 0, X } is a topology, the indiscrete topology.

Sometimes it is convenient to define a topology U on a set X by first describing a smaller collection B of subsets of X, and then defining U to be those subsets of X that can be written as unions of subsets belonging to B. We've done this already when defining the metric topology: Let X be a metric space and Let B be the collection of subsets of X of the form $B_e(x)$:= { y E X|d(y, x)< e }(the balls in X). Then the metric topology U on X consists of those subsets U which are unions of subsets belonging to B.

Theorem. Let B be a collection of subsets of a set X satisfying the following conditions

1. Every point x E X belongs to some subset B E B.

2. If Bi, B_2 E B, then for every x E B_1 f B_2 there is some B E B with x E B and
B c Bi f B2.

Then T := { unions of subsets belonging to B } is a topology on X.

Definition. If the above conditions are satisfied, we call the collection B is known a basis for the topology T or we say that B generates the topology T.

It is easy to check that the collection of balls in a metric space satisfies the above conditions and hence the collection of open subsets is a topology.

1.4 CONSTRUCTIONS WITH TOPOLOGICAL SPACES

Subspace topology

Definition . Let X be a topological space, and A c X a subset. Then

T={ A f U|U c X } open is a topology on A known the subspace topology.

Theorem. Let X be a metric space and A c X. Then the metric topology on A agrees with the subspace topology on A(as a subset of X equipped with the metric topology).

Theorem. Let X, Y be topological spaces and Let A be a subset of X equipped with the subspace topology. Then the inclusion map i: A ^ X is continuous and a map f: Y ^ A is continuous if and only if the composition i o f: Y ^ X is continuous.

Product topology

Definition . The product topology on the Cartesian product X x Y={(x, y)| x E X, y E Y } of topological spaces X, Y is the topology with basis

B={ U x V|U c X, V c Y } open

The collection B obviously satisfies property of a basis; property holds since(U x V)C(U' x V')=(U C U')x(V n V'). We note that the collection B is not a topology since the union of U x V and U' x V' is typically not a Cartesian product(e.g., draw a picture for the case where X=Y=R and U, U', V, V' are open intervals).

Theorem. The product topology on R^{-} x R^{n}(with each factor equipped with the metrictopology)agrees with the metric

Notes

topology on $R^{-}+^{n}=R^{-} \times R^{n}$.

Theorem. Let X, Y_i, Y_2 be topological spaces. Then the projection maps pi: Yi xY_2 4 Y is continuous and a map f: X 4 Y_i x Y_2 is continuous if and only if the component maps X^f Yi x Y2 ^WYi are continuous for i=1, 2.

Theorem. . Let X be a topological space and Let f, g: X 4 R be continuous maps.
Then f + g and f • g continuous maps from X to R. If g(x)=0 for all x E X, then also f /g is continuous.

Any polynomial function f: R^n 4 R is continuous.

The multiplication map p: $GL_n(R)$x $GL_n(R)$4 $GL_n(R)$is continuous.

Proof. we note that the map f + g: X 4 R can be factored in the form X —4 R x R -4 R

The map f x g is continuous since its component maps f, g are continuous; the map a is continuous, and hence the composition f + g is continuous. The argument for f • g is the same, with a replaced by m.

To prove that f /g is continuous, we factor it in the form

X ------------------------------ ^{fx}^-R x R^{x} $^{PlX(Jopa)}$; R x R^x --------------- -— ^R,

where R^x={ t E R|t=0 }, p_i(resp. p_2)is the projection to the first(resp. second)factor of R x R^x, and I: R^x 4 R^x is the inversion map t 4 t^{-i}.

we note that the constant map R^n M R, x=(x \,...,x_n)M a is obviously continuous, and that the projection map p_i: R^n M R, x=(xi, ..., x_n)M x_i is continuous monomial function x M axl^1 ■ ■ ■

$x^{\%}n$ is continuous. Any polynomial function is a sum of monomial functions and hence continuous. Let $M_{nxn}(R)=R^n$ be the set of n x n matrices and Let L : Mnxn(R)x Mnxn(R) M M, $_n$xn(R)(A, B)M AB be the map given by matrix multiplication. The map p is continuous if and only if the composition

Mnxn(R)X Mnxn(R) —M M, $_n$xn(R)~ ~M R

is continuous for all 1 < i, j < n, where p_ij is the projection map that sends a matrix A to its entry A_ij G R. Since the pj(p(A, B))=(A ■ B)$_i$j is a polynomial in the entries of the matrices A and B, this is a continuous map and hence p is continuous. Restricting p to invertible matrices, we obtain the multiplication map

L|: GLn(R)x GLn(R) —m GLn(R)

that we want to show is continuous. We will argue that in general if f: X M Y is a continuous map with f(A)C B for subsets A C X, B C Y, then the restriction $f|_A$: A M B is continuous.

AB

i j

X Y

where i, j are obvious inclusion maps. These inclusion maps are continuous w.r.t. the subspace topology on A, B by Theorem. The continuity of f and i implies the continuity of f o i=j o f|A which again by Theorem implies the continuity of $f|_A$.

Quotient topology.

Definition. Let X be a topological space and Let ~ be an equivalence relation on X. We denote by X/~ be the set of equivalence classes and by

p: X M X/~ x M [x]

be the projection map that sends a point x G X to its equivalence class [x]. The quotient topology on X/~ is given by the collection of subsets

U={ U C X/~| p^{-1}(U)is an open subset of X }.

Notes

The set X/~ equipped with the quotient topology is known the quotient space.

The quotient topology is often used to construct a topology on a set Y which is not a subset of some Euclidean space R^n, or for which it is not clear how to construct a metric.

If there is a surjective map p: X —y Y from a topological space X, then Y can be identified with quotient space X/ —, where the equivalence relation is given by x — x' if and only if p(x)=p(x'). In particular, Y=X/ —can be equipped with the quotient topology. Here are important examples.

Example. The real projective space of dimension n is the set

RP^n := { 1-dimensional subspaces of R^{n+1} }.

The map S^n —y RP^n R^{n+1} 9 v M subspace generated by v is surjective, leading to the identification $RP^n=S^n$ /(v - ±v), and the quotient topology on RP^n.

Similarly, working with complex vector spaces, we obtain a quotient topology on the complex projective space

CP^n := { 1-dimensional subspaces of C^{n+1} }=S^{2n+1} /(v — zv), z E S^1

Generalizing, we can consider the Grassmann manifold $G_k(R^{n+fc})$:= { k-dimensional subspaces of R^{n+fc} }. There is a surjective map

Vk(R^{n+fc})={(v1, ..., v_{fc})| v E R^{n+fc}, vj's are orthonormal } ^ $G_k(R^{n+fc})$

given by sending(v_1, ..., v_k)E $V_k(R^{n+fc})$to the k-dimensional subspace of R^{n+fc} spanned by the v^s. Hence the subspace topology on the Stiefel manifold $V_k(R^{n+fc})$C $R^{(n+fc)fc}$ gives a quotient topology on the Grassmann manifold $G_k(R^{n+fc})=V_k(R^{n+fc})$/ —. The same construction works for the complex Grassmann manifold $G_k(C^{n+fc})$.

As the examples below will show, sometimes a quotient space X/ — is homeomorphic to a topological space Z constructed in a different way. To establish the homeomorphism between X/ — and Z, we need to construct continuous maps

f: X /---------y Z g: Z M X/ —

that are inverse to each other. The next Theorem shows that it is easy to check continuity of the map f, the map out of the quotient space.

Theorem. The projection map p: X y X/ — is continuous and a map f: X/ — y Z to a topological space Z is continuous if and only if the composition f op: X y Z is continuous.

As we will observe in the next section, there are many situations where the continuity of the inverse map for a continuous bijection f is automatic. So in the examples below, and for the exercises in this section, we will defer checking the continuity of f^{-1} to that section.

Notation. Let A be a subset of a topological space X. Define a equivalence relation ~ on X by x ~ y if x=y or x, y E A. We use the notation X /A for the quotient space X/~.

Example. We claim that the quotient space [— 1, +1] /{ ±1 } is homeomorphic to

S^1 via the map f: [-1, +1] /{ ±1 } M S^1 given by [t] M e^{nit}. Geometrically

speaking, the map f wraps the interval [-1, +1] once around the circle. It is

easy to check that the map f is a bijection.

To observe that f is continuous, consider the composition [-1, +1] [-1, +1] /{ ±1 } S^1 C=R^2, where p is the projection map and i the inclusion map. This composition sends t E [-1, +1] to e^{nit}=(sin nt, cos nt)E R^2.

It is a continuous function, since its component functions sin nt and cos nt are continuous functions. The continuity of i o f o p implies the continuity of i o f, which implies the continuity of f. As mentioned above, we'll postpone the proof of the continuity of the inverse map f^{-1} to the next section. homeomorphic to a subspace of R^3.

1.5 PROPERTIES OF TOPOLOGICAL SPACES

In the previous subsection we described a number of examples of topological spaces X, Y that we claimed to be homeomorphic. We typically constructed a bijection f: X ^ Y and argued that f is continuous. However, we did not finish the proof that f is a homeomorphism, since we defered the argument that the inverse map f^{-1}: Y ^ X is continuous. We note that not every continuous bijection is a homeomorphism. For example if X is a set, X$(resp. X_{ind})
is the topological space given by equipping the set X with the discrete(resp. indiscrete)topology, then the identity map is a continuous bijection from X$ to X_{ind}. However its inverse, the identity map X_{ind} ^ X$ is not continuous if X contains at least two points.

Fortunately, there are situations where the continuity of the inverse map is automatic as the following proposition shows.

Proposition. Let f: X ^ Y be a continuous bijection. Then f is a homeomorphism provided X is compact and Y is Hausdorff.

The goal of this section is to define these notions, prove the proposition above, and to give a tools to recognize that a topological space is compact and /or Hausdorff.

Hausdorff spaces

Definition. Let X be a topological space, x E X, i = 1, 2, ... a sequence in X and x E X. Then x is the limit of the x's if for any open subset U C X containing x there is some N such that Xj E U for all i > N.

Caveat: If X is a topological space with the indiscrete topology, every point is the limit of every sequence. The limit is unique if the topological space has the following property:

Definition. A topological space X is Hausdorff if for every x, y E X, x=y, there are disjoint open subsets U, V C X with x E U, y E V.

Note: if X is a metric space, then the metric topology on X is Hausdorff(since for x=y and e=d(x, y)/2, the balls $B_e(x)$, $B_e(y)$are disjoint open subsets). In particular, any subset of R^n, equipped with the subspace topology, is Hausdorff. The notion of Cauchy sequences can be defined in metric spaces, but not in general for topological spaces(even when they are Hausdorff).

Theorem. Let X be a topological space and A a closed subspace of X. If x_n E A is a sequence with limit x, then x E A.

Proof. Let x /A. Then x is a point in the open subset X \ A and hence by the definition of limit, all but finitely many elements xn must belong to X \ A, contradicting our assumptions.

Compact spaces

Definition. An open cover of a topological space X is a collection of open subsets of X whose union is X. If for every open cover of X there is a finite sub collection which also covers X, then X is known compact. Some books(like Munkres' Topology)refer to open covers as open coverings, while newer books(and wikipedia) observem to prefer to above terminology, probably for the same reasons as me: to avoid confusions with covering spaces, a notion we'll introduce soon.

Theorem. If f: X ^ Y is a continuous map and X is compact, then the image f(X)is compact.

In particular, if X is compact, then any quotient space X/~ is compact, since the projection map X ^ X/~ is continuous with image X/~.

Proof. To show that f(X)is compact let that { U_a }, a E A is an open cover of the subspace f(X). Then each U_a is of the form U_a=V_a if f(X)for some open subset V_a E Y. Then { f^{-1}(Va)}, a E A is an open cover of X. Since X is compact, there is a finite subset
A' of A such that { f^{-1}(K)}, a E A' is a cover of X. This implies that { U_a }, a E A' is a finite cover of f(X), and hence f(X)is compact.

Theorem. 1. If K is a closed subspace of a compact space X, then K is compact.

If K is compact subspace of a Hausdorff space X, then K is closed.

Proof. Let that { U_a }, a E A is an open covering of K. Since the U_a's are open w.r.t. the subspace topology of K, there are open subsets Va of X such that U_a=K if K. Then the Va's together with the open subset X \ K form an open covering of X. The compactness of X implies that there is a finite subset A' C A such that the subsets Va for a E A', together with X \ K still cover X. It follows that U_a, a E A' is a finite cover of K, showing that K is compact.

Corollary. If f: X ^ Y is a continuous bijection with X compact and Y Hausdorff, then f is a homeomorphism.

Proof. We need to show that the map g: Y ^ X inverse to f is continuous, i.e., that g^{-1}(U)=f(U)is an open subset of Y for any open subset U of X. Equivalently(by passing to complements), it suffices to show that g^{-1}(C)=f(C)is a closed subset of Y for any closed subset C of C.

Now the assumption that X is compact implies that the closed subset C C X is compact that Y is Hausdorff then implies by part of Theorem that f(C)is closed.

Theorem. Let K be a compact subset of R^n. Then K is bounded, meaning that there is some r > 0 such that K is contained in the open ball $B_r(0)$:= { x E R^n|d(x, 0)< r }.

Proof. The collection $B_r(0)$if K, r E(0, ro), is an open cover of K. By compactness, K is covered by a finite number of these balls; if R is the maximum of the radii of these finitely many balls, this implies K C $B_R(0)$as desired.

Corollary. If f: X ^ R is a continuous function on a compact space X, then f has a maximum and a minimum.

Proof. K=f(X)is a compact subset of R. Hence K is bounded, and thus K has an infimum a := inf K E R and a supremum b := sup K E R. The infimum(resp. supremum)of K is the limit of a sequence of elements in K; since K is closed, the limit points a and b belong to K. In other words, there are elements x_{min}, x_{max} E X with $f(x_{min})=a < f(x)$for all x E X and $f(x_{max})=b > f(x)$for all x E X.spaces we are interested in, are in fact compact. Note that this is quite difficult just working from the definition of compactness: you need to ensure that every open cover has a finite subcover. That sounds like a lot of work...

Fortunately, there is a very simple classical characterization of compact subspaces of Euclidean spaces:

Theorem.(Heine-Borel Theorem)A subspace X C R^n is compact if and only if X is closed and bounded.

We note that we've already proved that if K C R^n is compact, then K is a closed subset of R^n and K is bounded

There two important ingredients to the proof of the converse, namely the following two results:

Theorem. A closed interval [a, b] is compact.

Theorem. If Xi, ...,X_n are compact topological spaces, then their product XiX- • -xX_n is compact.

The statement is true more generally for a product of infinitely many compact, the correct definition of the product topology for infinite products this result is known Tychonoff's Theorem, Proof of the Heine-Borel Theorem. Let K C R^n be closed and bounded, say K C $B_r(0)$. We note that $B_r(0)$is contained in the n-fold product

P := [—r, r] x • • • x [—r, r] C R^n

which is compact. So K is a closed subset of P and hence compact

Connected spaces

Definition. A topological space X is connected if it can't be written as decomposed in the form X=U U V, where U, V are two non-empty disjoint open subsets of X.

For example, if a, b, c, d are real numbers with a < b < c < d, consider the subspace X=(a, b)H(c, d)C R. The topological space X is not connected, since U=(a, b), V=(c, d)are open disjoint subsets of X whose union is X. This remains true if we replace the open intervals by closed intervals. The space X'=[a, b] H [c, d] is not connected, since it is the disjoint union of the subsets U'=[a, b], V'=[c, d]. We want to emphasize that while U' and V' are not open as subsets of R, they are open subsets of X', since they can be
written as

U'=(-ro, c)n X' V'=(b, ro)n X',
showing that they are open subsets for the subspace
topology of X' C R.

Theorem. Any interval I in R(open, closed, half-open, bounded or not)is connected.

Proof. Using proof by contradiction, Let us let that I has a decomposition I=U U V as the union of two non-empty disjoint open subsets. Pick points u E U and v E V, and Let us let u < v without loss of generality. Then

[u, v]=U' U V' with U' := U n [u, v] V' := U n [u, v] is a decomposition of [u, v] as the disjoint union of non-empty disjoint open subsets U', V' of [u, v]. We claim that the supremum c := sup U' belongs to both, U' and

V', thus leading
to the desired contradiction. Here is the argument.

Assuming that c doesn't belong to U', for any e > 0, there must be some element of U' belonging to the interval(c — e, c), allowing us to construct a sequence of elements u E U' converging to c. This implies c E U' since U' is a closed subspace of [u, v](its complement V' is open).

By construction, every x E [u, v] with x > c=sup U' belongs to V'. So we can construct a sequence v E V' converging to c. Since V' is a closed subset of [u, v], we conclude c E V'.

•

Theorem.(Intermediate Value Theorem)Let X be a connected topological space and f: X ^ R a continuous map. If elements a, b E R belong to the image of f, then also any real number c between a and b belongs to the image of f.

Proof. Let that c is not in the image of f. Then X=f^{-1}(—ro, c) U f^{-1}(c, ro)is a decomposion of X as a union of non-empty disjoint open subsets.

There is another notion, closely related to the notion of connected topological space which might be easier to think of geometrically.

Definition. A topological space X is path connected if for any points x, y E X there is a path connecting them. In other words, there is a continuous map 7: [a, b] ^ X from some interval to X with 7(a)=x, 7(b)=y.

Theorem. Any path connected topological space is connected.

Proof. Using proof by contradiction, Let us let that the topological space X is path connected, but not connected. So there is a decomposition X=U U V of X as the union of non-empty open subsets U, V C X. The assumption that X is path connected allows us to find a path 7: [a, b] ^ X with 7(a)E U and 7(b)E V. Then we obtain the decomposition [a, b]=f^{-1}(U)U f^{-1}(V)of the interval [a, b] as the disjoint union of open subsets.

Notes

These are non-empty since a E

f^{-1}(U)and b E f^{-1}(V). This implies that [a, b] is not connected, the desired contradiction.

For typical topological spaces we will consider, the properties "connected" and "path connected" are equivalent. But here is an example known as the topologist's sine curve which is connected, but not path connected. It is the following subspace of R^2:

$X=\{(x, \sin -^{1})E\ R^2|0 < x < 1\}\ U\ \{(0, y)E\ R^2|\ —1 < y < 1\}$

Check your Progress-1

Discuss Topology,

___

___

___

Discuss Topological spaces

___

___

___

1.6 ONE VARIABLE CALCULUS

In this brief discussion of one variable calculus, we introduce the Riemann

integral, and relate it to the derivative. We will define the Riemann integral

of a bounded function over an interval I=[a, b] on the real line. For now, we let f is real valued. To start, we partition I into smaller intervals.

A partition P of I is a finite collection of subintervals $\{ J_k : 0 < k < N \}$, disjoint except for their endpoints, whose union is I. We can order the J_k so that $J_k=[x_k,$ %k**+i],** where

$xo < xi < \cdots < x_N < x_N+_1, x_0=a, x_N+_1=b.$

We call the points x_k the endpoints of P. We set

$l(J_k)= x_{k+1} — x_k$, maxsize(P)= max^ $l(J_k)$

We then set

IP(f)=**J**^sup f(x)l(Jk), k Jk

Lp(f)=^ **J** f(x)l(Jk). Jk k

the definition of sup and inf. We call Ip(f)and Lp(f)respectively the upper sum and lower sum of f, associated to the partition P. Note that Lp(f)< Ip(f). These quantities should approximate the Riemann integral of f, if the partition P is sufficiently "fine."

To be more precise, if P and Q are two partitions of I, we say P refines

Q, and write P y Q, if P is formed by partitioning each interval in Q. Equivalently, P y Q if and only if all the endpoints of Q are also endpoints

of P. It is easy to observe that any two partitions have a common refinement;

just take the union of their endpoints, to form a new partition. Note also that refining a partition lowers the upper sum of f and raises its lower sum:

P y Q =^ I p(f)< I $_Q$ (f), and Ip(f)> I_c(f).

Consequently, if Pj are any two partitions and Q is a common refinement,

we have

P(f)< I Q(f)< I Q(f)< Ip_2(f).

Now, whenever f : I ^ R is bounded, the following quantities are well

defined:

(1.b6) I(f)= f 1 p(f), I(f)=sup Lp(f ^Pen(/) pen(/)

Notes

where n(I)is the set of all partitions of I. We call Iff)the lower integral of f and I(f)its upper integral. Clearly Iff)< I(f). We then say that f is Riemann integrable provided I(f)=I(f), and in such a case, we set

J f(x)dx=J f(x)dx=I(f)=Iff).

We will denote the set of Riemann integrable functions on I by R(I).

We derive some basic properties of the Riemann integral.

Proposition If f, g € R(I), then f + g € R(I), and

J(f + g)dx=j f dx + J gdx.I ii

Proof. If Jk is any subinterval of I, then

sup(f + g)< sup f + sup g, and inf(f + g)> inf f + inf g,

jk Jk Jk Jk Jk Jk

so, for any partition P, we have Ip(f + g)< Ip(f)+ Ip(g). Also, using common refinements, we can simultaneously approximate I(f)and I(g)by Ip(f)and Ip(g), and ditto for I(f + g). Thus the characterization implies I(f + g)< I(f)+ I(g). A parallel argument implies I(f + g)> I(f)+ I(g), and the proposition follows.

Next, there is a fair supply of Riemann integrable functions.

Proposition If f is continuous on I, then f is Riemann integrable.

Proof. Any continuous function on a compact interval is bounded and uni-
formly continuous. Let u(S)be a modulus of continuity for f, so

\x - y\< 5 =^\f(x)- f(y)\ < u(S), u(S)^ 0 as 5 ^ 0. Then

maxsize(P)< 5 =^ Ip(f) — Lp(f)< w(5)• l(I),

which yields the proposition.

We denote the set of continuous functions on I by C(I).

C(I)c R(I).

The criterion on a partition guaranteeing that Ip(f)and I_p(f)are close to J_I f dx when f is continuous.

We produce an extension, giving a condition under which Ip(f)and I(f)

are close, and Ip(f)and I(f)are close, given f bounded on I. Given a partition P_0 of I, set minsize(P_0)=min{ l(J_k): J_k € P_0 }.

Theorem Let P and Q be two partitions of I. Let

maxsize(P)< — minsize(Q). k

Let \f \ < M on I. Then

$^I P^{(f)} < ^I Q^{(f)} + \sim T^{l(I)}>$

Ip(f)> IQ(f)- -^HI).

Proof. Let P_1 denote the minimal common refinement of P and Q. Consider on the one hand those intervals in P that are contained in intervals in Q and on the other hand those intervals in P that are not contained in intervals in Q. Each interval of the first type is also an interval in Pi. Each interval of the second type gets partitioned, to yield two intervals in P_1. Denote by P**i** the collection of such divided intervals. The lengths of the intervals in P^1 sum to < l(I)/k. It follows that

\ I P (f)- I P - I (f)\ <E 2Ml(J)< $2M^l k^)$, J p

and similarly \Ip(f) — Ip_1(f)\ < 2Ml(I)/k. Therefore

IP(f)< I P I (f)+ ~^l(I), IP(f)> IPi(f) — ~^l(I).

Since also Ip_1(f)< IQ (J)and Ip_1(f)> I Q (J)

The following consequence is sometimes known Darboux's Theorem.

Theorem Let P_v be a sequence of partitions of I into v intervals J_{vk}, 1 < k < v, such that maxsize(P_v) —> 0.

If f : I ^ R is bounded, then(EE$^{14)}$ $^I P_V^{(f)}$^$^{I(f)and}$

LP„$^{(f)}$^ $^{I(f)}$. Consequently,

f € R(I)^ I(f)=lim E f(ivk)l(Jvk), is—Yoo ^{LJ} **k**=1 for arbitrary f_{vk} € J_{vk}, in which case the limit is f dx.

Proof. As before, let \f \ < M. Pick e=1 /k > 0. Let Q be a partition such that

I(f)< IQ(f)< I(f)+ e,

I(f)> LQ(f)> I(f)- e.

Now pick N such that

v > N =^ maxsize P_v < e minsize Q.

yields, for v > N,

LPv(f)< LQ(f)+ 2Ml(I)e,

L_Pv(f)> L_Q(f)- 2Ml(I)e.

Hence, for v > N,

L(f)< LPv(f)< L(f)+ [2Ml(L)+ 1]e,

L(f)> LPv(f)> L(f)- [2Ml(L)+ 1]e.f '(a)=0.

pr

Then there exists a' G Z

1.7 THE DIFFERENTIAL CALCULUS OF FUNCTIONS OF SEVERAL VARIABLES

The Linear Structure on Rm

Rm as a Vector Space

The concept of a vector space is already familiar to you from your study of algebra.

If we introduce the operation of addition of elements X\=(x \,...,0^) and X2=(#2, • • • > XT)*n by the formula

Xi +X2=(x } + x \,..., X™ +x™),

and multiplication of an element x=(x1, ..., xm)by a number A G R via the relation

Ax=(Ax1, ..., \xm),

then Rm becomes a vector space over the field of real numbers. Its points can
now be known vectors. .

The vectors

d=(0, ..., 0, 1, 0, ..., 0)(i=1, ...,m)

(where the 1 stands only in the ith place)form a maximal linearly independent set of vectors in this space, as a result of which it turns out to be an m-dimensional vector space.

Any vector x G Rm can be expanded with respect to the basis, that is, represented in the form

x=xxe\ H V xrnern

When vectors are indexed, we shall write the index as a subscript, while denoting its coordinates, as we have been doing, by superscripts. This is convenient for many reasons can make the convention of writing expressions like briefly in the form taking the simultaneous presence of subscript and superscript with the same Letter to indicate summation with respect to that Letter over its range of variation.

Check your Progress-2

Discuss One variable calculus

____

____

____

____

Discuss The differential calculus of functions of several variables
____

1.8 LET US SUM UP

In this unit we have discussed the definition and example of Topology, Topological spaces, Constructions with topological spaces, Properties of topological spaces, One variable calculus, The differential calculus of functions of several variables.

1.9 KEYWORDS

Topology: In mathematics, topology is concerned with the properties of a geometric object that are preserved under continuous deformations, such as stretching, twisting, crumpling and bending, but not tearing or gluing.

Topological spaces: A topological space (X, T) is a set X together with a topology T on X.

1.10 QUESTIONS FOR REVIEW

Explain Topology, Topological spaces

Explain One variable calculus

The differential calculus of functions of several variables

1.11 ANSWERS TO CHECK YOUR PROGRESS

Topology, Topological spaces (answer for Check your Progress-1 Q)

One variable calculus

The differential calculus of functions of several variables (answer for Check your Progress-2

Q)

1.12 REFERENCES

- Analysis of Several Variables
- Application of Several Variables
- Function of Several Variables
- Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables

UNIT - 2: FUNCTIONS OF SEVERAL VARIABLES

STRUCTURE

2.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about Functions Of Several Variables

Learn, Understand about Space Of Linear Transformations From R^m To R^n

Learn, Understand about The Space Rm And Its Subsets Open And Closed Sets In Mm

Learn, Understand about Limits And Continuity Of Functions Of Several Variables

Learn, Understand about Continuity Of A Function Of Several Variables And Properties Of Continuous Functions

2.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Functions Of Several Variables, The Space Of Linear Transformations From R^m To R^n, The Space Rm And Its Subsets, Open And Closed Sets In Mm, Limits And Continuity Of Functions Of Several Variables, Continuity Of A Function Of Several Variables, And Properties Of Continuous Functions

2.2 FUNCTIONS OF SEVERAL VARIABLES

Up to now we have considered almost exclusively numerical-valued functions x \f(x)in which the number f(x)was determined by giving a single number x from the domain of definition of the function.

However, many quantities of interest depend on not just one, but many factors, and if the quantity itself and each of the factors that determine it can be characterized by some number, then this dependence reduces to the fact that a value y= /(x1, ...,xn)of the quantity in question is made to correspond to an ordered set(x1, ..., xn)of numbers, each of which describes the state of the corresponding factor. The quantity lets this value when the factors determining this quantity are in these states.

For example, the area of a rectangle is the product of the lengths of its sides. The volume of a given quantity of gas is computed by the formula

V=R

Where R is a constant, m is the mass, T is the absolute temperature, and p is the pressure of the gas. Thus the value of V depends on a variable ordered triple of numbers (m, T, p), or, as we say, V is a function of the three variables m, T, and p.

Notes

Our goal is to learn how to study functions of several variables just as we learned how to study functions of one variable.

As in the case of functions of one variable, the study of functions of several numerical variables begins by describing their domains of definition.

The Space Mm and the Most Important Classes of its Subsets

The Set Mm and the Distance in it

We make the convention that Mm denotes the set of ordered m-tuples (x1, ..., xm)of real numbers xlGl, (i=1, ..., m).

Each such m-tuple will be denoted by a single Letter x=(x1, ...,xm) and, in accordance with convenient geometric terminology, will be known a point of Mm. The number x% in the set(x1, ...,xm)will be known the zth coordinate of the point x=(a:1, ..., xm).

The geometric analogies can be extended by introducing a distance on Mm between the points x\=(x \,..., x™)and X2=(x2, ..., x™)according to the formula

^(x\ - 4)2 .

2=1

The function

d : Mm x Mm -)> R
defined by the formula obviously has the following properties:

d(xi, x2)> 0;

(d(xi, x2)=0)(xi=x2);

d(xux2)=d(x2, x 1);

d(xi, x3)< d(xi, x2)+ d(x2, x3).

This last inequality(known, again because of geometric analogies, the triangle inequality)is a special case of Minkowski's inequality.

A function defined on pairs of points { x \,x2)of a set X and possessing the properties is known a metric or distance on X.

A set X together with a fixed metric on it is known a metric space.

Thus we have turned Mm into a metric space by endowing it with the metric given by relation.

The reader can get information on arbitrary metric spaces. Here we do not wish to become distracted from the particular metric space Mm that we need at the moment.

Since the space Mm with metric will be our only metric space, forming our object of study, we have no need for the general definition of a metric space at the moment. It is given only to explain the term "space" used in relation to Mm and the term "metric" in relation to the function.

It follows that for i e { 1, ..., m }

\x\ — xl2\ < d(x \,x2)< y /m max \x\ — xl21,

l<2<m

that is, the distance between the points xi, x2 <= Mm is small if and only if the corresponding coordinates of these points are close together.

It is clear that for m=1, the set M1 is the same as the set of real numbers, between whose points the distance is measured in the standard way by the absolute value of the difference of the numbers.

2.3 THE SPACE OF LINEAR TRANSFORMATIONS FROM RM TO RN

Linear transformations mapping R^m To R^n We can add such linear transformations in the usual way:$(L_1 + L_2)(x) = L_1(x) + L_2(x)$. Similarly we can multiply such a linear transformation by a scalar. In this way, the set

Notes

$L(R^n, R^m)$= { linear transformations from R^m To R^n becomes a vector space. If we choose bases for R^n and R^m, say the standard bases, then each element of $L(R^n, R^m)$has an m X n matrix with respect to these bases.

Since there are mn entries in such a matrix, and they all can be chosen independently of each other, $L(R^n, R^m)$has dimension mn. A basis is the set of m X n matrices which are all zero except for a 1 in one entry.

Second derivative

Recall that if f : R^n ! R^m is differentiate at a point x 2 R^n, then D f(x)is a linear transformation from R^m To R^n . Hence, for each x, Df(x)2 $L(R^m, R^n)$. From this we observe that Df is a function from R^n to $L(R^n, R^m)$.

We can then discuss D(Df) , or D^2f, the second derivative of f. For each x 2 R^n, D2f

(x)is a linear transformation from R^n to $L(R^n,R^m)$. Hence, for any v 2 R^n, $D^2f(x)(v)$2 $L(R^n, R^m)$. Therefore, for any w 2 R^n, $D^2f(x)(v)(w)$2 R^m.

Recall that R^n X R^n= {(v, w)| v and w are in R^ng . We can therefore consider $D^2f(x)$as a linear transformation from R^n X R^n to R^m. So instead of writing D2f

(x)(v)(w), we write $D^2f(x)(v, w)$.

This linear transformation from R^n to R^m is known "bilinear", because it is linear as a function of v for each fixed w, and also as a function of w for each fixed v. In other words,

$D^2f(x)(av^1 + ,dv^2, yw^1 + hw^2)= ayD^2f(x)(v^1, w^1)+ \hat{}yD^2f(x)(v^2, w^1)$

$+ ahD^2f(x)(v^1, w^2)+ \hat{}hD^2f(x)(v^2, w^2)$

Now we will only consider the case $m = 1$. Thus,

We wish to consider the nature of a general bilinear function L from R^n x R^n to R.

f : R^n ! R

For each x, Df(x): R^n ! R Df

For each x. $D^2f(x)$Equivalently, $D^2f(x)$

R^n ! $L(R^n, R)$

R^n x R^n ! R

Let ei, ..., e_n be the standard basis vectors of R^n. Then for each i and j, L(ej, ej) G R. Let L(ej, ej) = ajj.

It will be simplest now to consider the case n= 2 . Suppose that $v=c_1e_1 + c_2e_2$, and $w=d_1e_1 + d_2e_2$. The bilinearity implies that

L(v, w)= L(c1e1 + C2e2, d1e1 + d2e2)

= C1d1 an + C1d2a12 + C2d1a21 + C2d2a22.

It turns out that this equals an a12 \(d1

(cl, C2) \ a a I i d

a21 a22 d2

(Check by multiplying this out.)In this way, each L is associated with an n x n

matrix A. In the case where $L= D^2f(x)$, it is shown in the text that

f @ ^{2}f

A= (@x_i^x_j(x)

If you recall that for most functions, the order in which you take partial derivatives doesn't matter, you observe that under some assumptions on f, A is a symmetric matrix.

All of the second partial derivatives of f are continuous.

Example: Let $f(x, y)= x^2y+xy^3$. We will find the standard matrix for $D^2f(1, 1)$, and check that the limit formula for derivative works. For this function we have

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

By this we mean that

$(2xy + y^3)u$

$(x^2 + 3xy^2)v$

Notes

Also,

$2 \ y \ 2x + 3y^2$

$2x + 3y^2 \ 6 \ xy$

(Notice that= _@LL.)Recall that $D^2 f(x, y)$: R^n ! $L(R^n, R)$. Therefore, must be a map from R^n to R. We saw such a map before: Df(x) maps R^n to R. Any element of $L(R^n, R)$can be written in the form

x —► bx

where b is a 1 X n matrix; that is, a row vector. And any linear map L from R^n to fall n-dimensional row vectors } can be written as y ! $y^T A$ for some n X n matrix A. It is shown in the text that if L= $D^2 f(x)$partial derivatives of f, known the "Hessian".

This leads us to the equation We now check this last formula using the definition of derivative. However, it is a bit complicated to describe just what is meant by the norm of a linear operator. It turns out to be equivalent to discuss the corresponding matrices. Once again the sup norm will be convenient. We wish to check that(Notice that in the numerator we are dealing with row vectors.)We obtain

$||(2xy + y^3, \ x^2 + 3xy^2) - (3, 4) - (2(x - 1) + 5(y - 1), \ 5(x - 1) + 6(y - 1))$^

It is sufficient to show that the ratio of the absolute value of each component of the vector in this expression to the norm in the denominator tends to zero as(x, y)!

(1, 1)*

The first component is $y^3 + 2xy - 2x - 5y + 4$. A little algebra is necessary: Since

$2x = 2(x - 1) + 2$ and $5y = 5(y - 1) + 5$, we have

$y^3 + (2(x - 1) + 2)y - 2x - 5(y - 1) - 5 + 4 = y^3 - 3y + 2 + (x - 1)(2y - 2)$

Further, it turns out that $y^3 - 3y + 2 = (y - 1)^2 (y + 2)$. Hence if$(^{x,} y) = {}^{(1,1)}$

then the ratio of the absolute value of the first component of the numerator to the denominator is

|(y - 1)^2(y + 2)+ 2(x - 1)(s - 1)^1((* ')╫"j'$^{2|x}$~" i^f 1^x - ^{1} > |y -
1

max f|x - 1|, |y - 1| }fa-Vfeyl+%-"^{2} if |x - 1| < |y - 1| '

Both alternatives on the right tend to zero as as(x, y)!(1, 1). The second component can be handled similarly. It would be a nice algebra exercise to do this.

Third derivative

Notice the pattern: f : R^n ! R, and for each x 2 R^n(where f is differentiable),

Df(x): R^n ! R. In other words, Df(x)2 L(R^n, R). The linear transformation

Df(x)has the standard matrix(1 X n)given by the gradient, which is in R^n. Thus, Df : R^n ! R^n. Df is not usually a linear transformation.

As we explained, D^2f(x)is a linear transformation from R^n X R^n to R, and this linear transformation has the standard n X n matrix given above.

Therefore, D^2f : R^n ! L(R^n X R^n, R). Hence, we expect that for each x,

D^3f(x): R^n ! L(R^n X R^n, R). This will involve the third derivatives $_{dx}$.@ /.$_{dxk}$.

We will consider this further below. First, we have a review of Taylor series in one variable.

Taylor series for f : R ! R

First recall the general formula for a Taylor series in one variable. Suppose that

f : R ! R, and all derivatives of f exist at every x 2 R. If x_0 2 R, then the Taylor series for f at x_0 is

m_0^f<$^{n)}$(xo)(x - Xo)^n .
n!

Here f <$^{n)}$is the n-th derivative of f. We have the usual conventions that 0 ! = 1 and f <$^{0)}$=f.

This series can converge for all x, or only for x in some interval containing x_0.
(It obviously converges if $x = x_0$.)And if it converges for some $x = x_0$, it might not converge to f(x). Examples of these possibilities will be given in class.

Definition If the series converges to $f(x)$ in some neighborhood of x_0, then f is known "analytic" at x_0.

Perhaps of even more importance is using a finite sum of the terms in the Taylor series to approximate f on some interval containing x_0. This can sometimes be done even if f is not analytic at x_0, perhaps because not all of the derivatives of f at x_0 are defined. The theorem which allows us to give such approximations is known Taylor's theorem. To state Taylor's theorem we first need a definition.

Definition Suppose that $f : I \; C \; R \; ! \; R$, where I is an open interval containing a point x_0. Suppose that r is a nonnegative integer. We say that f is of class C^r on I if the first r derivatives, f, f^0, f^{00}, ..., $f <^{r)}$exist and are continuous on I.

Theorem Suppose that $f : I \; C \; R \; ! \; R$ where I is an open interval containing a point x_0, and f is of class C^r on I. Suppose that x and y are in I. Then there is a c between x and y such that

$$^{f(}y^{)}-^{f(x)} = s\, n=\backslash \; n^{f} <^{n)(x)(}y \; -^{x)\,n} + \; r, \; ^{f} <^{r)(c)(}y \; -^{x)\,r}$$

If $r = 1$, then this is the mean value theorem.

As an example, Let $f(x) = |x|^{5}=^{2}$, and consider $f(2) - f(-1)$. Notice that $f^0(0) = f^{00}(0) = 0$, but $f^{000}(0)$ doesn't exist. Also, $f(x) = (-x)^{5}=^{2}$ if $x < 0$. We wish to find $c \; 2(-1, \; 2)$ such that

$f(2)-f(-1)=f(-1)(2-(-1))+1\ f^{00}(c)(2-(-1))^2$

$$2^5=^2-1=_2(-1)(3)+-f^{00}(c)(9).$$

If c > 0, then f"(c)=||c$^{1/2}$, while if d= —c, then f"(d)= (—|)(—|)(—d)^{1} $^{/2}$= f"(c), so f" is an even function. Therefore we can let c > 0, and we want

$$o5\ /_2\ ^{13\ 9\ 53\ 1}$$
$$2^{5/2}+----=\ ------c\ 2,$$
$$2\ 2\ 22\ '$$

or

$$r$$
$$^8(n_5$$
$$_{/2}\ 13$$
$$\backslash f c\ =\ ---\ 2^{5/2}+-------$$

^{V} 135 V 2

Since we letd that c > 0, we have to check that c < 2. That is easily done.

$$P\ <\ —(^8+7)=T$$

Taylor's theorem of order **2** and quadratic forms.

As pointed out earlier, the mean value theorem is a special case of Taylor's Theorem.
if f is differentiable at every x, then for any x_o and x there is a c between x_o and x such that

$$f(x)=f(xo)+\ Df(c)(x-xo).$$

Recall that Df(c)is a row vector, (the gradient).

Extending by one more term

$$f(x)=f(xo)+\ Df(xo)(x-xo)+\ \ 2\,D^2f(c)(x-xo,\ x-xo).$$

We need to explain the last term. From the theory above, we observe that if we write
x — x_o as a column vector, then the last term is of the form

Notes

$$(x - x_o)^T A(x - x_o)$$

and A is the n X n matrix . Writing this out, we have the expression

$$D^2f(c)(x - xo, x - xo) = ((x - xo)_1, (x - xo)_2, \cdots, (x - xo)_n) A$$

Let's look again at n= 2 . Let

$$x - x \text{»} = (v)$$

for scalars u and v. Then the expression in becomes

$$a_n u + a_{12}uv + a_{21}uv + a_{22}v .$$

But A is symmetric, so we get

$$a_{11}u^2 + 2a_{12} uv + a_{22}v^2.$$

Such an expression is known a "quadratic form". In the n dimensional case with $(x - x_0) = u$, we get

$$a n u 1 + a22u2 + \cdots + a_{nn} u^2{}_n + 2a12u1u2 + 2a13 u1u3 + \cdots + 2a(,,\_1)_n u_{n}\_1u,,,$$

which is again known a quadratic form. One of the chief questions one asks about a quadratic form is whether it is positive whenever u= 0 . In that case it is known a "positive definite" quadratic form. One major reason that question is important is its application in the next section of the text to maxima and minima of functions

$f : R^n - R.$

2.4 THE SPACE RM AND ITS SUBSETS OPEN AND CLOSED SETS IN MM

Definition. For 8 > 0 the set

B(a; 8)={ x <= Mm|d(a, x)< 8 }

is known the ball with center a <= Mm of radius 6 or the 8-neighborhood of the point a <= Mm.

Definition. A set G C Mm is open in Mm if for every point x <= G there is a ball B(x; 8)such that B(x\ 8)C G.

Example. Mm is an open set in Mm.

Example. The empty set 0 contains no points at all and hence can be regarded as satisfying Definition 2, that is, 0 is an open set in Mm.

Example. A ball F(a;r)is an open set in Mm. Indeed, if x <= B(a;r), that is, d(a, x)< r, then for 0 < 8 < r — d(a, x), we have B(x\ 8)C B(a; r), since

(<= <= B(x;8)) =>(d(x, <=)< 8)=>

=>(d(a, <=)< d(a, x)+ d(x, <=)< d(a, x)+ r — d(a, x)=r).

Example. A set G={ x <= Mm| d(a, x)> r }, that is, the set of points whose distance from a fixed point a <= Mm is larger than r, is open. This fact is easy to verify using the triangle inequality for the metric.

Definition. The set F C Mm is closed in Mm if its complement G=Mm\F is open in Mm.

Example. The set B{ a\r)={ x <= Mm| d(a, x)< r }, r > 0, that is, the set of points whose distance from a fixed point a <= Mm is at most r, is closed, as follows from Definition. The set B(a\r)is known the closed ball with center a of radius r.

Proposition The union(J Ga of the sets of any system { Ga, a <= A }

ot<=A

of open sets in Mm is an open set in Mm.

n

The intersection p| Gi of a finite number of open sets in Mm is an

Notes

2=1

open set in Mm.

The intersection p| Fa of the sets of any system { Fq,, a <= A } of
aeA

closed sets Fa in Mm is a closed set in Mm.

n

The union(J Fi of a finite number of closed sets in Mm is a closed

2=1

set in Mm.

Proof. If x e(J Ga, then there exists ao <= A such that x <= Gao, and

ot<=A consequently there is a ^-neighborhood B(x\ 8)of x such that B(x\ 8)C Gao.
But then B(x\ 8)C(J G^.

a€A

n

Let x <= P| G{ . Then x <= G*, (i=1, ..., n). Let <5i, ..., 8n be positive

i= 1

numbers such that B{ x\8i)C G*, (i=1, ..., n). Setting <5=min{ 5i, • • •>^n },

n

we obviously find that 8 > 0 and B(x\ 8)C H

2=1

Let us show that the set G f f] Faj complementary to f] Fa in Mm

' oteA

is an open set in Mm.

Indeed,

c(n f«)=u(cp<«)=u -

acEA acEA a€;A

where the sets Ga=CFa are open in Mm. Part a')now follows from a), b')Similarly, from b)we obtain

n n n

c(u^=n(^)=n^. a

2=1 2=1 2=1

Example. The set 5(a; r)={ x <= Mm| d(a, x)=r }, r > 0, is known the sphere of radius r with center a <= Mm. The complement of S{ a\ r)in Mm, is the union of open sets. Hence by the proposition just proved it is open, and the sphere S{ a\r)is closed in Rm.

Definition. An open set in Mm containing a given point is known a neighborhood of that point in Mm.

In particular, the ^-neighborhood of a point is a neighborhood of it.

Definition. In relation to a set E c Mm a point x <= Mm is
an interior point if some neighborhood of it is contained in E\
an exterior point if it is an interior point of the complement of E in Mm;
a boundary point if it is neither an interior point nor an exterior point.

It follows from this definition that the characteristic property of a boundary point of a set is that every neighborhood of it contains both points of the set and points not in the set.

Example. The sphere 5(a;r), r > 0 is the set of boundary points of both the open ball B(a\r)and the closed ball B{ a\r).

Example. A point a <= Mm is a boundary point of the set Mm \ a, which has no exterior points.

Notes

Example. All points of the sphere 5(a;r)are boundary points of it; re- garded as a subset of Mm, the sphere S(a;r)has no interior points.

Definition. A point a E Mm is a limit point of the set E C Mm if for any neighborhood 0(a)of a the intersection E D 0(a)is an infinite set.

Definition. The union of a set E and all its limit points in Rm is the closure of E in Mm.

The closure of the set E is usually denoted E.

Example. The set B(a\r)=B(a\r)U S(a\r)is the set of limit points of the open ball B(a\r)\ that is why H(a;r), in contrast to B(a\r), is known a closed ball.

Example. S(a\r)=S(a;r).

Proposition.(F is closed in Mm)<=>(F=F in Mm).

In other words, F is closed in Mm if and only if it contains all its limit points.

Proof. Let F be closed in Mm, x E Mm, and x <= F. Then the open set G=Mm \F is a neighborhood of x that contains no points of F. Thus we have shown that if x <= F, then x is not a limit point of F.

Let F=F. We shall verify that the set G=Mm \F is open in Mm. If x E G, then x <= F, and therefore x is not a limit point of F. Hence there is a neighborhood of x containing only a finite number of points xi, ..., xn of F.
Since x <= F, one can construct, for example, balls about x, Oi(ar), ..., On(x)

n

such that Xi <= Oi(x). Then 0(x)=p| Oi(x)is an open neighborhood of x

2=1

containing no points of F at all, that is, 0(x)C Mm \ F and hence the set Mm \ F=Mm \F is open. Therefore F is closed in Mm. □

Compact Sets in Mm

Definition. A set K C Mm is compact if from every covering of K by sets that are open in Mm one can extract a finite covering.

Example. A closed interval [a, b] C M1 is compact by the finite covering Theorem(Heine-Borel theorem).

Example A generalization to Mm of the concept of a closed interval is the set

I={ x e Mm| a1 < x% < bl, i=1, ..., m },

which is known an m-dimensional interval, or an m-dimensional block or an m-dimensional parallelepiped.

We shall show that I is compact in Mm.

Proof. Let that from some open covering of I one cannot extract a finite covering. Bisecting each of the coordinate closed intervals P={ x1 E M : a% <
x1 < F }, (i=1, ...,m), we break the interval I into 2m intervals, at least one of which does not admit a covering by a finite number of sets from the open system we started with. We proceed with this interval exactly as with the original interval. Continuing this division process, we obtain a sequence of nested intervals I=I\ D I2 ID • • • ID In ID • • •, none of which admits a finite covering. If In={ x E Mm| a%n < x1 < bln, i, ...,m }, then for each i E { 1, ..., m } the coordinate closed intervals ai< xi < b\(n =1, 2, ...)
form, by construction, a system of nested closed intervals whose lengths tend to zero. By finding the point <=z E [«n?^n] common to all of these intervals for each i E { 1, ..., m }, we obtain a point <==(<=x, ...,

<=m)belonging to all the intervals I=Ji, J2, ..., /n, Since <= E /, there is an open set G in the system of covering sets such that <= E G. Then for some 6 > 0 we also have B(<=;<5)C G. But by construction and the relation there exists N such that In C B(<=; S)C G for n > N. We have now reached a contradiction with the fact that the intervals In do not admit a finite covering by sets of the given system.

Proposition. If K is a compact set in Mm, then K is closed in Mm;

any closed subset of Mm contained in K is itself compact.

Proof. We shall show that any point a E Mm that is a limit point of K must belong to K. Suppose a <= K. For each point x <= K we

construct a neighborhood G(x)such that a has a neighborhood disjoint from

G(x). The set(G(x)}, x E K, consisting of all such neighborhoods forms an open covering of the compact set if, from which we can select a finite covering G(xi), ..., G(xn). If now Oi(a)is a neigborhood of a such that n

G(xi)fl Oi(a)=0, then the set 0(a) — f)Oi(a)is also a neighborhood of

i=1

a, and obviously K fl 0(a)=0. Thus a cannot be a limit point of K.

Suppose F is a closed subset of Mm and F C K. Let { Ga }, ol E A, be a covering of F by sets that are open in Mm. Adjoining to this collection the

open set G=Mm \ F, we obtain an open covering of Mm, and in particular an open covering of if, from which we select a finite covering of if. This finite covering of if will also cover the set F. Observing that G fl F=0, one can say that if G belongs to this finite covering, we will still have a finite covering of F by sets of the original system { Gq, }, a E A, if we remove G.

Definition. The diameter of a set E C Mm is the quantity d(E):= sup d(xi, X2).

xi,X2 <=E

Definition. A set E C Mm is bounded if its diameter is finite.

Proposition. If K is a compact set in Mm, then K is a bounded subset of Mm.

Proof Take an arbitrary point a <= Mm and consider the sequence of open balls(7?(a;n)}, (n=1, 2, ...,). They form an open covering of Mm and consequently also of K. If K were not bounded, it would be impossible to select a finite covering of K from this sequence.

Proposition. The set K C Mm is compact if and only if K is closed and bounded in Mm.

Proof The necessity of these conditions .

Let us verify that the conditions are sufficient. Since K is a bounded set, there exists an m-dimensional interval I containing K.

Example. I is compact in Mm. But if AT is a closed set contained in the compact set 7, then it is itself compact.

The distance d(E \,E2)between the sets E \,E2 C Rm is the quantity

d(Ei, E2):= inf d(xi, x2).

xiEEi, X2 G-E /2

Give an example of closed sets E\ and E2 in Rm having no points in common for which d(Ei, E2)=0.

Show that the closure E in Rm of any set E C Rm is a closed set in Rm; the set dE of boundary points of any set E C Rm is a closed set;

Notes

if G is an open set in Rm and F is closed in Rm, then G\F is open in Rm.

Show that if K\ D K2 D • • • D Kn D • • • is a sequence of nested nonempty

00

compact sets, then p| Ki 7^ 0.

i=1

In the space Rk a two-dimensional sphere S2 and a circle S1 are situated so that the distance from any point of the sphere to any point of the circle is the same.
Is this possible?

Consider problem for spheres 5m, Sn of arbitrary dimension in Rk. Under what relation on m, n, and k is this situation possible

2.5 LIMITS AND CONTINUITY OF FUNCTIONS OF SEVERAL VARIABLES

The Limit of a Function

In the operation of passing to the limit for a
real-valued function /: X —)> R defined on a set in which a base B was fixed.

In the next few sections we shall be considering functions /: X —)> Rn defined on subsets of Rm with values in R=R1 or more generally in Rn, n G N. We shall now make a number of additions to the theory of limits connected with the specifics of this class of functions.

However, we begin with the basic general definition.

Definition. A point A e Rn is the limit of the mapping f : X —)> Rn over a base B in X if for every neighborhood V(A)of the point there exists an element B e B of the base whose image f(B)is contained in V(A).

In brief,

(lim f(x)=A):=(W(A)3B € B(f(B)c V{ A))).

We observe that the definition of the limit of a function /: X —»> Rn is exactly
the same as the definition of the limit of a function f : X R if we keep in mind what a neighborhood V(A)of a point A <= Rn is for every n G N.

Definition . A mapping /: X —)> Rn is bounded if the set f(X)C Rn is bounded in Rn.

Definition. Let B be a base in A. A mapping f : X Rn is ultimately bounded over the base B if there exists an element B of B on which /is bounded. Taking these definitions into account and using the same reasoning that one can verify without difficulty that

a function f : X Rn can have at most one limit over a given base B in

X;

a function f : X Rn having a limit over a base B is ultimately bounded over that base.

Definition can be rewritten in another form making explicit use of the metric in Rn

Definition

(lim /(x)=A € Rn):=(Ve > 0 3B e B Vx € B(d(f(x), A)< e))

Or

Definition

(lim /(a:)=ietnj :=(lim(d(f(x), A)=o).

The specific property of a mapping /: X —>> Mn is that, since a point y E Mn is an ordered n-tuple(y1, ..., yn)of real numbers, defining a function
/: X —)> Mn is equivalent to defining n real-valued functions f% : X —)> M
(z=1, ..., n), where /*(x)=yi(i=1, ..., n).

If A=(A1, ..., An)and y=(y1, ..., yn), we have the inequalities

|yl - Al\ < d(y, A)< s /n max |y* - A%\,

1 <i<n

from which one can observe that

lim /(x)=A<& lim /1(a;)=A%(i=l, ..., n),

that is, convergence in Mn is coordinatewise.

Now Let X=N be the set of natural numbers and B the base k oo, k e N, in X. A function /: N —)> Mn in this case is a sequence { ? /& }, k e N,
of points of Mn.

Definition. A sequence { yk }, k e N, of points G Mn is fundamental(a Cauchy sequence)if for every e > 0 there exists a number N G N such that
d(yk^yk2)< <= for all kuk2> N.

One can conclude from the inequalities that a sequence of points Vk —(vli • • •»Vk)C Mn is a Cauchy sequence if and only if each sequence of coordinates having the same labels { ylk }, k G N, i=1, ..., n, is a Cauchy sequence.

Taking account of relation and the Cauchy criterion for numerical sequences, one can now assert that a sequence of points Mn converges if and only if it is a Cauchy sequence.

In other words, the Cauchy criterion is also valid in Mn.

Later on we shall call metric spaces in which every Cauchy sequence has a limit compLete metric spaces. Thus we have now established that Mn is a compLete metric space for every n G N.

Definition. The oscillation of a function /: X —>> Mn on a set E C X is the quantity

w(/;E):= d(f(E)),
where d(f(E)) is the diameter of f(E).

As one can observe, this is a direct generalization of the definition of the oscillation of a real-valued function, which Definition becomes
when n=1.

The validity of the following Cauchy criterion for the existence of a limit for functions /: X —)> Mn with values in Mn results from the completeness of Mn.

Theorem. Let X be a set and B a base in X. A function f : X —)> Mn has a limit over the base B if and only if for every e > 0 there exists an element B <= B of the base on which the oscillation of the function is less than e.

Thus,

lim f(x)We >03 BeB(w(/;B)< s).

This is a verbatim repetition of the proof of the Cauchy
criterion for numerical functions, except for one minor change: \f(xi) — f(x2)| must be replaced throughout by d(f(x\), f(x2)).

Notes

One can also verify another way, regarding the Cauchy criterion as known for real-valued functions and using relations.

The important theorem on the limit of a composite function also remains valid for functions with values in Mn.

Theorem. Let Y be a set, By a base in Y, and g : Y —)> Mn a mapping having a limit over the base By.

Let X be a set, Bx a base in X, and f : X Y a mapping of X into
Y such that for each By <= By there exists Bx € Bx such that the image f(Bx)is contained in By.

Under these conditions the composition g o f : X —)> Mn of the mappings f and g is defined and has a limit over the base Bx, and

lim(g o f)(x)=lim g(y).

ISx Oy

The proof can be carried out either by repeating the replacing M by Mn, or by invoking that theorem and using relation

Up to now we have considered functions /: X —>> Mn with values in Mn, without specifying their domains of definition X in any way. From now on we shall primarily be interested in the case when X is a subset of Mm.

As before, we make the following conventions.

U(a)is a neighborhood of the point a <= Mm;

o o

U(a)is a deleted neighborhood of a <= Mm, that is, U{ a):= U{ a)\ a;

Ue(cl)is a neighborhood of a in the set E C Mm, that is, Ue(cl):= EnU(a);

o 00

Ue(o)is a deLeted neighborhood of a in E, that is, Ue(o)'-= E D U(a)', x —)> a is the base of deLeted neighborhoods of a in Mm;

x —)> 00 is the base of neighborhoods of infinity, that is, the base consisting of the sets Mm \ B(a; r);

x —)> a, x <= E or(E 3 x —)> a)is the base of deLeted neighborhoods of a in E if a is a limit point of E;

x 00, x e E or(E 3 x 00)is the base of neighborhoods of infinity in E consisting of the sets E\B(a;r), if E is an unbounded set.

In accordance with these definitions, one can, for example, give the following specific form of the limit of a function when speaking of a function /: E —)> Rn mapping a set E C Rm into Rn:

(lim f(x)=:=(Ve > 0 3 Ue(o)Vx € Ue{ o)(d(f(x), A)< e)) .

\ EBx—hi /

The same thing can be written another way:

(lim f(x)=a):=

V EBx—ta))

=(Vs > 0 36 > 0 Vx G E(0 < d(x, a)< 6 => d(f(x), A)< e)) .

Here it is understood that the distances d(x, a)and d(f(x), A)are measured in the spaces(Mm and Mn)in which these points lie.

Finally,

(lim f(x)=A):=(Ve > 03H(a;r)Vx G Mm \B(a;r)(d(f(x), A)< e).

Let us also agree to say that, in the case of a mapping /: X —)> Mn, the phrase " /(x) —)> oo in the base means that for any ball B(A\r)C En there exists B e B of the base B such that f(B)Cln\ B(A\ r).

Example. Let x i-)» 7rz(x)be the mapping n7, : Rm R assigning to each x=(x1, ..., <=m)in Rm its zth coordinate xl. Thus

-Kl(x)=.

If a=(a1, ..., am), then obviously

7rz(x)a1 as x a .

The function x i->> 7rz(x)does not tend to any finite value nor to infinity as x oo if m > 1.

On the other hand,

m

f(x)=J2(7riw)oo as x —)> oo .

2=1

One should not think that the limit of a function of several variables can be found by computing successively the limits with respect to each of its coordinates. The following examples show why this is not the case.

Example Let the function /: R2 —)> R be defined at the point(x, y)G R2 as follows:

if*2 + rV0,

f(x, y)=i

[0, if x2 + y2=0 .

Then /(0, y)=f(x, 0)=0, while f(x, x)=\ for x ^ 0.

Hence this function has no limit as(x, y)(0, 0).
On the other hand,

lim(lim f(x, y))=lim(0)=0,
y^0lx->0M 'y)) 0V ' '

lim(lim f(x, y))=lim(0)=0 .
x_>0 v y_>0J v 'yj)x—>o y

Example. For the function

grg.ifsa + yVo,

. x2+y

/(®, y).

0, if x2 + y2=0,

we have

lim(lim f(x, y))=lim(^)=1,
x^0ly^0M 'y)) x^0\x2)2

lim(lim f(x, y))=lim(—=-
oVx->(TV ' y)) 2 /—>0 V y2)

Example For the function

x + ysin^, if x ^ 0,

f(x, y)"

0, if x=0,

we have

lim f(x, y)=0,
(*, j/)->(0, 0)

lim,,(lini f(x, y)) =

x—>0 y—> 0

yet at the same time the iterated limit

lim(lim f(x, y))
y->ovx-xr v yjJ

does not exist at all.

Example The function

. _ >if x2 + y2 ^0 >

f(x, y)=_

0, if x2 + y2=0,

Notes

has a limit of zero upon approach to the origin along any ray

x=at, y=fit.

At the same time, the function equals \ at any point of the form(a, a2), where a ^ 0, and so the function has no limit as(x, y)(0, 0).

2.6 CONTINUITY OF A FUNCTION OF SEVERAL VARIABLES AND PROPERTIES OF CONTINUOUS FUNCTIONS

Let E be a subset of Mm and /: E —>> Mn a function defined on E with values in Mn.

Definition. The function /: E —)> Mn is continuous at a E E if for every neighborhood V of the value f(a)that the function lets at a, there exists a neighborhood Ue(o)of a in E whose image f(ue(a)) is contained i nV{ f(a)).

Thus

(/: E —)> Mn is continuous at a e E):=

=(W(/(a)) 3UE(a)(f(UE(a)) C V(f(a)))) .

We observe that has the same form as Definition for continuity of a real-valued function, which we are familiar with from. As was the case there, we can give the following alternate expression for this definition:

(/: E —)> Mn is continuous at a G E):=

=(ye > 0 35 > 0 Vx e E(d(x, a)< S => d(f(x), f(a)) < <=)),
or, if a is a limit point of E,

(/: E —)> Mn is continuous at a G E):=(lim /(x)= /(a)) .

EBx—hi

As noted, the concept of continuity is of interest precisely in connection with a point a G E that is a limit point of the set E on which the function /is defined.

It follows from Definition and relation that the mapping /: E —)> Mn defined by the relation

(x1, ...,xm)=xt-Uy=(y\..., i /n)=is continuous at a point if and only if each of the functions y% =,..., xm)is continuous at that point.

In particular, we recall that we defined a path in Mn to be a mapping /: I —)> Mn of an interval /Cl defined by continuous functions /1(x), ..., /n(x)
in the form

x^y= { yi, ..., yn)={ f1{ x), ..., fn{ x)) .

Thus we can now say that a path in Mn is a continuous mapping of an interval I C M of the real line into Mn.

By analogy with the definition of oscillation of a real-valued function at a point, we introduce the concept of oscillation at a point for a function with values in Mn.

Let E be a subset of Mm, a € E, and E^(a; r)=E D E(a; r).

Definition. The oscillation of the function f : E —>> Rn at the point a G E is the quantity

w(/;a):= lim w(/; BE(a;r)) .

i—>-+0 v '

From Definition of continuity of a function, taking account of the properties of a limit and the Cauchy criterion, we obtain a set of frequently used local properties of continuous functions. We now list them.

Check your Progress-1

Discuss Functions Of Several Variables

____

____

____

____

Discuss Limits And Continuity of functions of several variables

____

____

2.7 LET US SUM UP

In this unit we have discussed the definition and example of Functions Of Several Variables, The Space Of Linear Transformations From To R^m To R^n

The Space Rm And Its Subsets, Open And Closed Sets In Mm, Limits And Continuity Of Functions Of Several Variables, Continuity Of A Function Of Several Variables, And Properties Of Continuous Functions

2.8 KEYWORDS

Functions Of Several Variables: Here numerical-valued functions x \f(x)in which the number f(x)was determined by giving a single number x from the domain of definition of the function.

The Space Of Linear Transformations From To R^m To R^n ….. Linear transformations mapping R^m To R^n We can add such linear transformations in the usual way:$(L_1 + L_2)(x)= L_1(x)+ L_2(x)$

The Space Rm And Its Subsets: Continuity Of A Function Of Several Variables And Properties Of Continuous Functions

Open And Closed Sets In Mm For 8 > 0 the set B(a; 8)={ x <= Mm|d(a, x)< 8 } is known the ball with center a <= Mm of radius 6 or the 8-neighborhood of the point a <= Mm.

Limits And Continuity Of Functions Of Several Variables: The Limit of a Function In the operation of passing to the limit for a real-valued function /: X —)> R defined on a set in which a base B was fixed

Continuity Of A Function Of Several Variables And Properties Of Continuous Functions: Let E be a subset of Mm and /: E —>> Mn a function defined on E with values in Mn.

2.9 QUESTIONS FOR REVIEW

Explain Functions Of Several Variables

Explain Limits And Continuity Of Functions Of Several Variables

2.10 ANSWERS TO CHECK YOUR PROGRESS

Functions Of Several Variables (answer for Check your Progress-1 Q)

Limits And Continuity Of Functions Of Several Variables

(answer for Check your Progress-1 Q)

2.11 REFERENCES

- System of Equation
- Function of Real Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables

UNIT - 3: LOCAL PROPERTIES OF CONTINUOUS FUNCTIONS

STRUCTURE

3.0 OBJECTIVES

After studying this unit, you should be able to:

Local Properties Of Continuous Functions

Linear Transformations L : Rm —> R™

The Norm In Rm

The Euclidean Structure On Rm

The Differential Of A Function Of Several Variables

The Differential And Partial Derivatives Of A Real-Valued Function

Coordinate Representation Of The Differential Of A Mapping. The Jacobi Matrix

Continuity, Partial Derivatives And Differentiability Of A Function At A Point

The Basic Laws Of Differentiation Linearity Of The Operation Of Differentiation

Differentiation Of A Composition Of Mappings (Chain Rule)

The Main Theorem

The Differential And Partial Derivatives Of A Composite Real Valued Function

3.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Local Properties Of Continuous Functions

Linear Transformations L : Rm —> R™

The Norm In Rm

The Euclidean Structure On Rm

The Differential Of A Function Of Several Variables

The Differential And Partial Derivatives Of A Real-Valued Function

Coordinate Representation Of The Differential Of A Mapping. The Jacobi Matrix

Continuity, Partial Derivatives And Differentiability Of A Function At A Point

The Basic Laws Of Differentiation Linearity Of The Operation Of Differentiation

Differentiation Of A Composition Of Mappings (Chain Rule)

The Main Theorem

The Differential And Partial Derivatives Of A Composite Real Valued Function

3.2 LOCAL PROPERTIES OF CONTINUOUS FUNCTIONS

A mapping f : E —)> Rn of a set E C Rm is continuous at a point a G E if and only if uo{ f\ a)=0.

A mapping f : E —)> Rn that is continuous at ae E is bounded in some neighborhood Ue(a)of that point.

If the mapping g :Y of the set Y C Rn is continuous at a point yo G Y and the mapping f : X Y of the set X C Rm is continuous at a point xq G X and f(xo)=yo, then the mapping g o f : X —)> Rfc is defined, and it is continuous at xq G X.

Real-valued functions possess, in addition, the following properties.

If the function f : E —)> M is continuous at the point a <= E and f(a)> 0(or f(a)< 0), there exists a neighborhood Ue(cl)of a in E such that f(x)> 0(resp. f(x)<0)for all x eUe(q>).

If the functions f : E —)> M and g : E —)> M are continuous at a G E, then any linear combination of them(af+(3g): E —)> R, where a, (3 <= R, t /iezr

product(f - g): E —)> R, and, if g(x)^ 0 on E, their quotient(^): E —)> M are defined on E and continuous at a.

Let us agree to say that the function /: E —> Rn is continuous on the set E if it is continuous at each point of the set.

The set of functions /: E —> Rn that are continuous on E will be denoted C(E';Rn)or simply C(E), if the range of values of the functions is unambiguously determined from the context. As a rule, this abbreviation will be

used when Rn=R.

Example. The functions(a;1, ...,<=m)x%(i=1,..., m), mapping Rm onto R(projections)are obviously continuous at each point a = (a1, ...,am)G Rm, since lim 7rl(x)=a1=7r%{ a).

x—ta

Example . Any function x f(x)defined on R, for example x i-> sinx, can also be regarded as a function(x, y)f(x)defined, say, on R2. In that case,

if /was continuous as a function on R, then the new function(x, y)\—> f(x)will be continuous as a function on R2. This can be verified either directly from the definition of continuity or by remarking that the function F is the composition(/o 7r1)(x, 2 /)of continuous functions.

In particular, it follows from this, when we take account of c)and e), that the functions

f(x, y)=sinx + exy, f(x, y)=arctan(ln(|x| + \y\ + 1)),

for example, are continuous on M2.

We remark that the reasoning just used is essentially local, and the fact that the functions /and F studied in Example were defined on the entire real line R or the plane R2 respectively was purely an accidental circumstance.

Notes

Example. The function f(x, y)of Example 2 is continuous at any point of the space R2 except(0, 0). We remark that, despite the discontinuity of f(x, y)at this point, the function is continuous in either of its two variables for each fixed value of the other variable.

Example . If a function /: E —> Rn is continuous on the set E and E is a subset of E, then the restriction f\g of /to this subset is continuous on E, as follows immediately from the definition of continuity of a function at a point.

We now turn to the global properties of continuous functions. To state them for functions /: E —> Rn, we first give two definitions.

Definition . A mapping /: E —> Rn of a set E C Rm into Rn is uniformly continuous on E if for every e > 0 there is a number 5 > 0 such that

d(f(xi), f(x2)) < <= for any points X \,X2 € E such that d(x 1, ^2)< S.

As before, the distances d(x 1, ^2)and d(/(xi), f(x2)) are let d to be measured in Rm and Rn respectively.

When m=n=1, this definition is the definition of uniform continuity of numerical-valued functions.

Definition . A set E C Rm is path wise connected if for any pair of its points <=0, #i, there exists a path r : I —> E with support in E and endpoints at these points.

In other words, it is possible to go from any point x$ G E to any other point x\ G E without leaving E.

Since we shall not be considering any other concept of connectedness for a set except path wise connectedness for the time being, for the sake of brevity we shall temporarily call path wise connected sets simply connected.

Definition. A domain in Rm is an open connected set.

Example. An open ball B(a\r), r > 0, in Rm is a domain. We already know that B(a; r)is open in Rm. Let us verify that the ball is connected.

Let

xo=(xl, x™)and x\=(x \,..., x™)be two points of the ball. The path defined by the functions xl{ t)=tx\ -f(1 — t)xl0, (i=1, ...,ra), defined on the closed interval 0 < t < 1, has x$ and x\ as its endpoints. In addition, its support lies in the ball B(a;r), since, by Minkowski's inequality, for any t <= [0, 1],

Example. The circle(one-dimensional sphere)of radius r > 0 is the subset of R2 given by the equation { x1)2 +(x2)2=r2. Setting x1=rcost, x2=rsin<=, we observe that any two points of the circle can be joined by a path that goes along the circle. Hence a circle is a connected set. However, this set is not a domain in R2, since it is not open in R2.

We now state the basic facts about continuous functions in the large. Global properties of continuous functions

If a mapping f : K Rn is continuous on a compact set K C Rm, then it is uniformly continuous on K.

If a mapping f : K Rn is continuous on a compact set K C Rm, then it is bounded on K.

If a function f : K —> R is continuous on a compact set K C Rm, then it lets its maximal and minimal values at some points of K.

If a function f : E —> R is continuous on a connected set E and lets the values f(a)=A and f(b)=B at points a, 6 G E, then for any C between A and B, there is a point c G E at which f(c)=C.

Earlier when we were studying the local and global properties of functions of one variable, we gave proofs of these properties that extend to the more general case considered here. The only change that must be made in the earlier proofs is that expressions of the type \x\ — X2\ or \f{ x\) — f{ x2)\
must be replaced by d(x 1, 2:2)and d(/(xi), /(a^)), where d is the metric

Notes

in the space where the points in question are located. This remark applies fully to everything except the last statement d).

Proof, d)Let r : I —> E be a path that is a continuous mapping of an interval
[a, /?]=I C R such that r(a)=a, r(/3)=b. By the connectedness of E there exists such a path. The function /of : I —> R, being the composition of continuous functions, is continuous; therefore there is a point 7 <= [a, /?] on the closed interval [a, ft] at which /o r(j)=C. Set c=^(7). Then c <= E
and /(c)=C.

Example The sphere 5(0; r) defined in Rm by the equation

(*1)2 + ... +(zm)2=r2,

is a compact set.

Indeed, it follows from the continuity of the function

(z1, ..., zm)^(z1)2 + ... +(zrn)2

that the sphere is closed, and from the fact that \x%\ < r(i=1, ...,ra)on the sphere that it is bounded.

The function

(x\ ..., xm)^(x1)2 + . . • +(xk)2 -(x^1)2 (xm)2

is continuous on all of Rm, so that its restriction to the sphere is also continuous, and by the global property c)of continuous functions lets its minimal and maximal values on the sphere. At the points(1, 0, ...,0)and(0, ...,0, 1)this function lets the values 1 and —1 respectively. By the connectedness of the sphere(observe Problem 3 at the end of this section), global property d)of continuous functions now enables us to assert that there is a point on the sphere where this function lets the value 0.

Example. The open set Rm \ 5(0; r)for r > 0 is not a domain, since it is not connected.

Indeed, if r : I —> Rm is a path one end of which is at the point x$ =

(0, ..., 0)and the other at some point x\=(x \,..., x™)such that(x\)2-\

(x™)2 > r2, then the composition of the continuous functions r : I —> Rm and /: Rm —> R, where

(x1, ..., xm)J->(x1)2 + --- +(xm)2,

is a continuous function on I assuming values less than r2 at one endpoint and greater than r2 at the other. Hence there is a point 7 on I at which (/o r)(7)=r2. Then the point x1=r(7)in the support of the path turns out to lie on the sphere 5(0; r). We have thus shown that it is impossible to get out of the ball B(0;r)C Rm without intersecting its boundary sphere 5(0; r).

3.3 LINEAR TRANSFORMATIONS L : RM —> R™

We recall that a mapping L : X —» Y from a vector space X into a vector space Y is known linear if

L{ X\xi + X2x2)=X\L(xi)+ X2L(x2)

for any xi, x2 G X, and Ai, A2 G R. We shall be interested in linear mappings
L : Rm Rn.

If { ei, ..., em } and { ei, ..., en } are fixed bases of Rm and Rn respectively, then, knowing the expansion

L(e{)=a\ei H b afen=a{ ej(i=1, ..., m)

of the images of the basis vectors under the linear mapping L : Rm —> Rn,
we can use the linearity of L to find the expansion of the image L(h)of any vector h=hle\ H b /imem=hlei in the basis { ei, ..., en }. To be specific,

. L(h)=L(tiei)== h{ a{ ej=a{ h% .

Hence, in coordinate notation:

L(h)=(a1ihi, ..., a?hi).

For a fixed basis in Rn the mapping L : Rm —> Rn can thus be regarded as a set

L=(L\..., Ln)

of n(coordinate)mappings U : Rm —> R.

Taking account easily conclude that a mapping L : Rm —> Rn
is linear if and only if each mapping U in the set is linear.

If as a column, taking account of relation we have

(L\h)\ (a\ al\(V\

L(h)= •••= •• •

\Ln(h)J \a1 ■■■ a^J \hm)

Thus, fixing bases in Rm and Rn enables us to establish a one-to-one correspondence between linear transformations L : Rm —> Rn and m x n-matrices
(a^), (i=1, ..., m, j=1, ..., n). When this is done, the ith column of the matrix(a^)corresponding to the transformation L consists of the coordinates of L(ei), the image of the vector e$ G { ei, ...,em }. The coordinates of the image of an arbitrary vector h=h%ei G Rm can be obtained from by multiplying the matrix of the linear transformation by the column of coordinates of h. Since Rn has the structure of a vector space, one can speak of linear combinations X\f\ -f A2 /2 of mappings fi : X —> Rn and : X —> Rn,

setting

(A1 /1 + A2 /2X2O := Ai /i(x)+ A2 /2ipt)•

In particular, a linear combination of linear transformations L\ : Rm -» Rn and L2 : Rm —> Rn is, according to the definition a mapping h^\iL1(h)+ \2L2(h)=L(h),

which is obviously linear. The matrix of this transformation is the corresponding linear combination of the matrices of the transformations L\ and L2.

The composition C=B o A of linear transformations A : Rm —> Rn and B : Rn —> Rfc is obviously also a linear transformation, whose matrix, as follows from, is the product of the matrix of A and the matrix of B(which is multiplied on the left). Actually, the law of multiplication for matrices was defined in the way you are familiar with precisely so that the product of matrices would correspond to the composition of the transformations.

3.4 THE NORM IN RM

The quantity

IMI=V(xl)2 H 1-(xm)2

is known the norm of the vector x=(x1, ..., xm)G Rm.

It follows from this definition, taking account of Minkowski's inequality, that

i° IMI > 0,

2°(||x||=0)(x=0),

3° ||Ax||=|A| • ||x||, where A G R,

4° ||x1+x2|| <|M + 11*211.

In general, any function || || : X —> R on a vector space X satisfying conditions l°-4° is known a norm on the vector space. Sometimes, to be precise as to which norm is being discussed, the norm sign has a symbol attached to it to denote the space in which it is being considered. For example, we can write |M|jRm or IMk71- As a rule, however, we shall not do that, since it will always be clear from the context which space and which norm are meant.

Notes

We remark that

\\X2-X\W=d(x \,x2)

where d(x 1, <=2)is the distance in Rm between the vectors x\ and X2, regarded as points of Rm.

It is clear from that the following conditions are equivalent:

x —> xo, d(x, xo) —> 0, \\x — xo|| —> 0 .

In view we have, in particular,

\\x\\=d(0, x).

Property 4° of a norm is known the triangle inequality, and it is now clear why.

The triangle inequality extends by induction to the sum of any finite number of terms. To be specific, the following inequality holds:

||xi + l-Zfcll < ||»l|| + --- + ||«fc|| •

The presence of the norm of a vector enables us to compare the size of values of functions /: X —> Rm and g : X —> Rn.

Let us agree to write f(x)=o(g(x)) or /=o(g)over a base B in X if
|| /(x)||Rm =

°(lb(x)lk71)over the base B.

If f(x)=(/1(x), ..., fm(x)) is the coordinate representation of the map-ping /: X —> Rm, then in view of the inequalities

m

\f(x)\ <\\f(x)\ \<^2\f(x)\

2=1

one can make the following observation, which will be useful below:

= o(g)over the base B)(/z=o{ g)over the base B\ i=1, ..., m).

We also make the convention that the statement f=0(g)over the base B in X will mean that || /(x)||]Rm=0(||^(x)||]Rn)over the base B.

We then obtain from

(/=0(g)over the base B)(fl=0(g)over the base B\ i=1, ..., m).

Example. Consider a linear transformation L : Rm —> Rn. Let h=

be an arbitrary vector in Rm. Let us estimate ||L(ft)||Rn:

m m

Y, hiL^)<E 11^)11 E 11^)" ii^ii-

2=1 2=1 2=1 '

Thus one can assert that

L(h)=0(h)as h —> 0 .

In particular, it follows from this that L(x — xo)=L(x) — L(xo) —> 0 as x —> Xo, that is, a linear transformation L : Rm —> Rn is continuous at every point x$ G Rm. Prom estimate it is even clear that a linear transformation is uniformly continuous.

3.5 THE EUCLIDEAN STRUCTURE ON RM

The concept of the inner product in a real vector space is known from algebra as a numerical function(x, y)defined on pairs of vectors of the space and possessing the properties

(x, x)> 0,

(x, x)=0 x=0,

(xux2)=(x2, Xi),

(Xxi, x2)= \(x \, x2), where A G R,

(xi + x2, x3)=(xi, x3)+(x2, x3).

It follows in particular from these properties that if a basis { ei, ..., em } is fixed in the space, then the inner product(x, y)of two vectors x and y can be expressed in terms of their coordinates(x1, ..., xm)and(y1, ..., ym)as the bilinear form

(x, y)=9ijXlyJ

(where summation over i and j is understood), in which gij=(e^e^). Vectors are said to be orthogonal if their inner product equals 0.

A basis { ei, ..., em } is orthonormal if gij=5^ -, where

[0, if i + j,

Sij=< [1, if i=j.

In an orthonormal basis the inner product has the very simple form

(x, y)=5ijXly3,

or

(x, y)=x1 • y1 H h xm • y™ .

Coordinates in which the inner product has this form are known Cartesian coordinates.

We recall that the space Rm with an inner product defined in it is known Euclidean space.

Between the inner product and the norm of a vector there is an obvious connection

(x, x)=\\x\\2 .

The following inequality is known from algebra:

{ x, yf <(x, x)(y, y).

It shows in particular that for any pair of vectors there is an angle ip G [0, 7r] such that

(x, y)=\\x\\ ||y|| cosy .

This angle is known the angle between the vectors x and y. That is the reason we regard vectors whose inner product is zero as orthogonal.

We shall also find useful the following simple, but very important fact, known from algebra:

any linear function L : Rm —> R in Euclidean space has the form

L(x)={ <=, x },

where L G Rm is a fixed vector determined uniquely by the function L.

3.6 THE DIFFERENTIAL OF A FUNCTION OF SEVERAL VARIABLES

Differentiability and the Differential of a Function at a Point

Definition. A function /: E —> Rn defined on a set E C Rm is differentiate at the point x G E, which is a limit point of E, if

f(x 4- ft) — f(x)=L(x)h 4- a{ x\ ft),

where L(x): Rm —> Rn is a function2 that is linear in ft and a(x; ft)=o(h)as ft —y 0, x -f- ft G E.

The vectors

Ax(h):=(x 4- ft) — x=ft,

Af(x; h):= /(x + h)~ f(x)

are known respectively the increment of the argument and the increment of the function(corresponding to this increment of the argument). These vectors are traditionally denoted by the symbols of the functions of ft themselves Ax and Af(x). The linear function L(x): Rm -> Rn in is

Notes

known the differential, tangent mapping, or derivative mapping of the function /: E —» Rn at the point x G E.

The differential of the function /: E —> Rn at a point x G E is denoted by the symbols d f(x), Df(x), or f'(x).

In accordance with the notation just introduced, we can rewrite relation as

f(x + h)- f(x)=f'(x)h + a(x; h)

or

Af(x; h)=d f(x)h + a(x\ h).

We remark that the differential is defined on the displacements h from the point x G Rm.

To emphasize this, we attach a copy of the vector space Rm to the point x G Rm and denote it TxRm, TRm(x), or TR™. The space TR™ can be interpreted as a set of vectors attached at the point x G Rm. The vector space TR™ is known the tangent space to Rm at x G Rm. The origin of this terminology will be explained below.

The value of the differential on a vector h G TR™ is the vector f'(x)h G TW }{ x)attached to the point f(x)and approximating the increment f(x + h) — f(x)of the function caused by the increment h of the argument x.

Thus df(x)or f(x)is a linear transformation f(x): TR™ —> TR^xy

We observe that, in complete agreement with the one-dimensional case that

we studied, a vector-valued function of several variables is differentiable at a point if its increment Af(x; h)at that point is linear as a function of h up to the correction term a(x;h), which is infinitesimal as h —> 0 compared to the increment of the argument.

3.7 THE DIFFERENTIAL AND PARTIAL DERIVATIVES OF A REAL-VALUED FUNCTION

If the vectors /(x + /i), f(x), L(x)h, a{ x\ h)in Rn are written in coordinates becomes equivalent to the n equalities

f(x + h)~ f(x)=L\x)h + <*'(*; h)(i=1, ..., n) between real-valued functions, in which, as follows from relations, Ll(x): Rm —> R are linear functions and a%{ x\ h)=o(h)
as h —> 0, x 4- h G E, for every i=1, ..., n.

Thus we have the following proposition.

Proposition. A mapping f : E —> Rn of a set E C Rm is differentiate at a point x G E that is a limit point of E if and only if the functions f% : E —> R
(i=1, ...,n)that define the coordinate representation of the mappping are differentiate at that point.

Since relations are equivalent, to find the differential
L(x)of a mapping f : E —> Rn it suffices to learn how to find the differentials Ll(x)of its coordinate functions f1 : E —> R.

Thus, Let us consider a real-valued function f : E —> R, defined on a set E C Rm and differentiate at an interior point x <= E of that set. We remark that in the future we shall mostly be dealing with the case when E is a domain in Rm. If x is an interior point of E, then for any sufficiently small displacement ft from x the point x + h will also belong to E, and consequently will also be in the domain of definition of the function /: E —> R.

If we pass to the coordinate notation for the point x —(x1, ..., xm), the vector ft=(ft1, ...,ftm), and the linear function L(x)h=a\{ x)h } -f • • • am(x)ftm, then the condition

f(x 4- ft) — f(x)=L(x)h 4- o(h)as ft —> 0

can be rewritten as

f(xl+ h\..., xrn + hm)-f(x\..., xm)= ,

= ai{ x)hl H h arn(x)hm 4- o(h)as ft —> 0

Notes

where ai(x), ..., am(x)are real numbers connected with the point x.

We wish to find these numbers. To do this, instead of an arbitrary displacement ft we consider the special displacement

hi=Wei=0 • ei 4- • • • 4- 0 • e^_i 4- Wei 4~ 0 • e^\ 4- • • • + 0 • em

by a vector ft* collinear with the vector of the basis { ei, ..., em } in Rm.

When ft=fti, it is obvious that ||ft||=|ftz|, and so by(8.24), for ft=hi we obtain

f(x1, ..., xi-\xi + hi, a?+\,xm)-f(x\..., x\..., xm)=

= di(x)W 4- o(W)as W —> 0 .

This means that if we fix all the variables in the function /(x1, ...,xm) except the zth one, the resulting function of the zth variable alone is differentiate at the point x%.

In that way, we find that mix)=

— lim ->x% + WiX1*1, ...,xm)- /(x1, ...,xl, ...,xm)

h*->0 W

Definition. The limit is known the partial derivative of the function f(x)at the point x=(x1, ..., xm)with respect to the variable x%. We denote it by one of the following symbols:

, dif(x), Dif(x), f'xi(x)

3.8 COORDINATE REPRESENTATION OF THE DIFFERENTIAL OF A MAPPING. THE JACOBI MATRIX

Thus we have found formula for the differential of a real-valued function f : E R. But then, by the equivalence of relations for any mapping /: E R^n of a set E C R^m that is differentiate at an interior point x G E, we can write the coordinate representation of the differential

df(x)as(d f^l(x)h\ df(x)h= \dr(x)h / \$<=(x)■■■ §<=(x)j $\h^m$J

Definition. The matrix(difi(x))(i=1, ..., m, j=1, ..., n)of partial derivatives of the coordinate functions of a given mapping at the point x G E is known the Jacobi matrix[3] or the Jacobian[4] of the mapping at the point. In the case when n=1, we are simply brought back to formula and when n=1 and m=1, we arrive at the differential of a real-valued function of one real variable. The equivalence of relations and the uniqueness of the differential of a real-valued function implies the following result.

Proposition . If a mapping f : E —>• R^n of a set E C R^m is ifferentiable at an interior point x G E, then it has a unique differential df{ x)at that point, and the coordinate representation of the mapping df(x): TRJ[1] —>TR^($_x$)is given by relation.

3. 9 CONTINUITY, PARTIAL DERIVATIVES, AND DIFFERENTIABILITY OF A FUNCTION AT A POINT

We complete our discussion of the concept of differentiability of a function at a point by pointing out some connections among the continuity of a function at a

point, the existence of partial derivatives of the function at that point, and differentiability at the point.

we established that if L : R^m —> R^n is a linear transformation, then Lh 0 as ft 0. Therefore, one can conclude from relation that a function that is differentiate at a point is continuous at that point, since

f(x 4- ft) — f(x)=L(x)h + o(ft)as ft —0, x + h € E .

The converse, of course, is not true because, as we know, it fails even in the one-dimensional case.

.

Notes

. Thus the relation between continuity and differentiability of a function at a point in the multidimensional case is the same as in the one-dimensional case.

The situation is completely different in regard to the relations between partial derivatives and the differential. In the one-dimensional case, that is, in the case of a real-valued function of one real variable, the existence of the differential and the existence of the derivative for a function at a point are equivalent conditions. For functions of several variables, we have shown that differentiability of a function at an interior point of its domain of definition guarantees the existence of a partial derivative with respect to each variable at that point. However, the converse is not true.

Example. The function

$$f(x^1, x^2) = \begin{cases} 0, & \text{if } x^1x^2 = 0, \\ 1, & \text{if } x^1x^2 \neq 0, \end{cases}$$

equals 0 on the coordinate axes and therefore has both partial derivatives at the point(0, 0):

' fci-y 0 h^1 fci-y 0 h^1

ft /(0, 0)=,jm W')- /(0.0)= !^o=0.

h^2—*0 ft^2 h^2—*0 ft^2

At the same time, this function is not differentiate at(0, 0), since it is obviously discontinuous at that point.

The function given in fails to have one of its partial derivatives at points of the coordinate axes different from(0, 0). However, the function

f(x, y)=

has partial derivatives at all points of the plane, but it also is discontinuous at the origin and hence not differentiate there.

Thus the possibility of writing the right-hand side still does not guarantee that this expression will represent the differential of the function we are considering, since the function can be non differentiable.

This circumstance might have been a serious hindrance to the entire differential calculus of functions of several variables, if it had not been determined(as will be proved below)that continuity of the partial derivatives at a point is a sufficient condition for differentiability of the function at that point.

3.10 THE BASIC LAWS OF DIFFERENTIATION LINEARITY OF THE OPERATION OF DIFFERENTIATION

Theorem If the mappings R^n and /2 : E R^n, defined on a

set E C R^m, are differentiate at a point x <= E, then a linear combination of them(Ai /i -f A2 /2): E R^n is also differentiate at that point, and the following equality holds:

(Ai /i + A2 /2)'(x) —(Aif[+ A2 /2XX).

Equality shows that the operation of differentiation, that is, forming the differential of a mapping at a point, is a linear transformation on the vector space of mappings f : E R^n that are differentiate at a given point of the set E.

The left-hand side of contains by definition the linear transformation(A1 /1+A2 /2)'(x), while the right-hand side contains the linear combination(Ai /{ -f A2 /2)^)of linear transformations f[(x): R^m R^n, and f^{ x): R^m —)• R^n, which, as we know from Sect. 8.1, is also a linear transformation.

Theorem asserts that these mappings are the same.

Proof

(A1 /1 + X_2f_2)(x + h)-(Ai $/_2$ + Xih){ x)=

=(Ai /i(a; + h)4- \2f2(% + h)) ~ { Xifi(x)+ \2f2{ x)) =

= Ai(/i(x + h)~ fi(x)) + A_2($/_2$(x + h)~ f_2(x)) =

= Ai(f{(x)h + o(h)) + \2(f2(x)h + o(h)) =

=(Xif[(x)+ X_2f2(x))h + o{ h). □

If the functions in question are real-valued, the operations of multiplication and division(when the denominator is not zero)can also be performed.

We have then the following theorem.

Theorem:. If the functions f : E R and g : E ^ R, defined on a set E c R^m, are differentiate at the point x <= E, then

a) their product is differentiate at x and

(/• 9)'(x)=9(x)f(x)+ f(x)g'(x);

b) their quotient is differentiate at x if g(x)^ 0, and

(<=)'(*)=p<=jGK*)/'(s)- f(x)9'(x)) •

The proof of this theorem is the same as the proof of the corresponding parts of Theorem, so that we shall omit the details.

Relations can be rewritten in the other notations for the differential. To be specific:

d(Ai /i(a:)+ A_2 $/_2$)(x)=(Aid /i + A_2d $/_2$)(x),

d(/• g)(x)=g(x)df(x)+ f(x)dg(x),

^{d}(f)^=^)(^(®)^d /(®)- /(®)^{d}5(®)) •

Let us observe what these equalities mean in the coordinate representation of the mappings. We know that if a mapping : E —>• R^n that is differentiable at an interior point x of the set E C R^m is written in the coordinate form

<p(x)=

then the Jacobi matrix

v'(x)=(ft"')(x)=(a, vJ)(x)

will correspond to its differential d<^(x): R^m R^n at this point.

For fixed bases in R^m and R^n the correspondence between linear transformations L : R^m —»• R^n and m x n matrices is one-to-one, and hence the linear transformation L can be identified with the matrix that defines it.

Even so, we shall as a rule use the symbol f'{ x)rather than d /(x)to denote the Jacobi matrix, since it corresponds better to the traditional distinction between the concepts of derivative and differential that holds in the one-dimensional case.

Thus, by the uniqueness of the differential, at an interior point x of E we obtain the following coordinate notation for denoting the equality of the corresponding Jacobi matrices:

(<9i(Ai /i +Mfi)){ x)=(Ai<9i /i + A 2difi)(x)

(i = j=1, ..., n),

(di(f ■ 9))(x)=g(x)dif(x)+ f(x)dig(x)=

= <72^ ~ f{ x)d%g{ x))(i=1, ..., m).

It follows from the elementwise equality of these matrices, for example, that the partial derivative with respect to the variable x^1 of the product of real-valued functions /(x^1, ..., x^m)and ^(x^1, ..., x^m)should be taken as follows:

d^'s)(xi xm)_

dxi (® >•••>*

fl(®1, ..., ®m)^(®1, ..., ®m)+ /(s1, ..., ®TO)^(®1, ..., ®m).

We note that both this equality and the matrix equalities are obvious consequences of the definition of a partial derivative

Notes

and the usual rules for differentiating real-valued functions of one real variable. However, we know that the existence of partial derivatives can still turn out to be insufficient for a function of several variables to be differentiate. For that reason, along with the important and completely obvious equalities, the assertions about the existence of a differential for the corresponding mapping in Theorems acquire a particular importance.

We remark finally that by induction using one can obtain the relation

d(/i, • • •, fk){ x)=($/_2$ • • • fk)(x)dfi(x)+ 1-(/i • • • fk-iWk(x)

for the differential of a product(fi' — fk)of differentiate real-valued functions.

3.11 DIFFERENTIATION OF A COMPOSITION OF MAPPINGS (CHAIN RULE)

THE MAIN THEOREM

Theorem. If the mapping f : X —>• Y of a set X C R^m into a set Y C R^n is differentiate at a point x G X, and the mapping f : Y R^k is differentiate at the pointy=f(x)G Y, then their composition go f : X R^{fc} is differentiate at x and the differential d(g o /): TR™ —)• TR^^ of the composition equals the composition dg(y)o d /(x)of the differentials

df(x): TR™ TW $\}_{ix)=y}$, dg(y): TR™ $TR^k_{g\{ y)}$

The proof of this theorem repeats almost completely the proof of Theorem In order to call attention to a new detail that arises in this case, we shall nevertheless carry out the proof again, without going into technical details that have already been discussed, however.

Proof. Using the differentiability of the mappings /and g at the points x and y= /(#), and also the linearity of the differential g'(x), we can write

(9 ^{0} f)(x + h)-(go f)(x)=g(f(x + h)) - g{ f(x)) =

= 9'{ f(x))(f(x + h)~ f(x)) + o(f(x + h)~ f(x)) =

= 9'(y)(f'{ x)h + o{ h)) + o(f(x + h)~ f(x)) =

= g'{ y)(f(x)h)+ g'(y)(o(h)) + o(f(x + h)- /(x)) =

=(:g'{ y)° f'(x))h + a(x; h),

where g'{ y)o f'{ x)is a linear mapping(being a composition of linear mappings), and

at(x; h)=g'(y)(o(h)) + o(f(x + h)- f(xj).

g'(y)(o(h))=o(h)ash-t 0,

f(x + h) — /(x)=f'(x)h + o(h)=0(h)+ o(h)=0(h)as h -¥ 0,

and

°(f(^{x} + h) — f(x))=o(0(h))=o(h)as h —>• 0 .

Consequently,

a(x\ h)=o(h)+ o(/i)=o(h)as h —)► 0,

and the theorem is proved. □

When rewritten in coordinate form, Theorem means that if x is an interior point of the set X and

/dif^1(x)••• dmf^1(x)\

=(dif^j)(x),

\dif^n(x)dmf^n(x)J

and y=f(x)is an interior point of the set Y and

/dig^l(y)••• d_ng^1(y)\

= { djg^k)(y),

\dig^k(y)■■■ d_ng^k(y)\ di9^k (y) ••• d_ng^k(y) /

In the equality

{ di(g^l o f))(x)=djg^l{ f{ x)) ■ dif^j(x))

summation is understood on the right-hand side with respect to the index j over its interval of variation, that is, from 1 to n even in the sense of

Notes

elementwise equality of the matrices occurring in it.

Let us now consider some important cases of the theorem just proved.

3.12 THE DIFFERENTIAL AND PARTIAL DERIVATIVES OF A COMPOSITE REAL VALUED FUNCTION

Let $z=g(y^1, ..., y^n)$be a real-valued function of the real variables $y^1, ..., y^n$, each of which in turn is a function yi= /^($x^1, ..., x^m$) (j=1, ..., n)of the variables $x^1, ..., x^m$. Assuming that the functions g and are differentiate(j=1, ..., n), Let us find the partial derivative ^{d}^°P(x)of the composition of the mappings /: X —^ Y and g : Y —> R.

According to formula, in which I=1 under the present conditions, we find

$$di(g \circ f)(x)=d_jg(f(x)) \blacksquare dif^3(x),$$

or, in notation that shows more detail= di g{ f(x)) ■ difix)H+ d_ng(f{ x)) ■ $dif^n(x)$.

The Derivative with Respect to a Vector and the Gradient of a Function at a Point Consider the stationary flow of a liquid or gas in some domain G of R^3. The term "stationary" means that the velocity of the flow at each point of G does not vary with time, although of course it can vary from one point of G to another. Suppose, for example, /(x)= /(x^1, x^2, x^3)is the pressure in the flow at the point x=(x^x^x^3)G G. If we move about in the flow according to the law x=x(<=), where t is time, we shall record a pressure(/o x)(t)=f(x(t)) at time t. The rate of variation of pressure over time along our trajectory is obviously the derivative of function(f ox)(t)with respect to time. Let us find this derivative, assuming that /(x^1, x^2, x^3)is a differentiate function in the domain G. By the rule for differentiating composite functions, we find

$$(t)=+ \wedge(®(*))®^2W + \wedge(®(0)®^3w »(8\text{-}36)$$

where $x^l(t)$=^(t)(i=1, 2, 3).

Since the vector(x^x, x^2, x^3)=v(t)is the velocity of our displacement at time t and(dif, d2f, dsf)(x)is the coordinate notation for the differential d /(x)of the function /at the point x, can also be rewritten as that is, the required quantity is the value of the differential d /(x(t)) of the function /(x)at the point x(t)evaluated at the velocity vector v(t)of the motion.

In particular, if we were at the point xo=x(0)at time t=0, then

^l<=^.(0)=df(xo)v,

where v=t>(0)is the velocity vector at time t=0.

The right-hand side depends only on the point Xo G G and the velocity vector v that we have at that point; it is independent of the specific form of the trajectory x=x(<=), provided the condition x(0)=v holds. That means that the value of the left-hand side is the same on any trajectory of the form

x(t)=xo 4- vt 4- ct(t),

where a(t)=o(t)as t 0, since this value is completely determined by giving the point xo and the vector v G TR^3_o attached at that point. In particular, if we wished to compute the value of the left-hand side of directly(and hence also the right-hand side), we could choose the law of motion to be

x(t)=xo 4- vt,

corresponding to a uniform motion at velocity v under which we are at the point x(0)=xo at time t=0.

We now give the following

Definition. If the function /(x)is defined in a neighborhood of the point Xo G R^m and the vector v G TR™ is attached at the point xo, then the quantity

fix0 + vt)- fixo)

Notes

$D_v f(x_0)$:= lim ---------- ---------

(if the indicated limit exists)is known the derivative of f at the point Xo with

respect to the vector v or the derivative along the vector v at the point xq.

It follows from these considerations that if the function /is differentiate at the point #o, then the following equality holds for any function x(t)of the nd in particular, for any function of the form

$$D_v f(x_0)=<V(0)=df(x_0)v .$$

In coordinate notation, this equality says

$$D_v f(x_0)=\text{^}(xoW + \bullet\bullet\bullet + T\sim(xo)v^m \blacksquare$$

In particular, for the basis vectors e\=(1, 0, ..., 0), ..., e_m=(0, ..., 0, 1) this formula implies

D_eJ{ xo)=^i(xo)(i=l, ..., m).

By virtue of the linearity of the differential d /(#o), we deduce from that if /is differentiate at the point #o, then for any vectors v±, V2 G TR™ and any Ai, A2 G R the function has a derivative at the point xq with respect to the vector(Aitq -f X2V2)G TR™, and that

$$D\backslash_{lVj+}\backslash_{2V2}f(x_0)=XiD_vJ(x_0)+ \backslash_2 D_{V2}f(x_0)\bullet$$

If R^m is regarded as a Euclidean space, that is, as a vector space with an inner product, then it is possible to write any linear functional L(v)as the inner product(<=, v)of a fixed vector <==<=(L)and the variable vector v.

In particular, there exists a vector <= such that

$$df(x_0)v=(<=, v).$$

Definition. The vector <= G TR!^ corresponding to the differential df(xo)of the function /at the point xo in the sense is known the gradient of the function at that point and is denoted grad /(#o).

Thus, by definition

$df(x_0)v=(\text{grad } /(x_0), v)$.

If a Cartesian coordinate system has been chosen in R^m, then, by comparing relations we conclude that the gradient has the following representation in such a coordinate system:

We shall now explain the geometric meaning of the vector grad /(#o)-

Let e G TR™$_0$ be a unit vector. Then

$$D_e f(xo)=|\text{grad } f(xo)\backslash \text{ cosip},$$

where tp is the angle between the vectors e and grad f(xo).

Thus if grad f(xo)^ 0 and e=||grad /(xo)||$^{-1}$grad /(#o), the derivative D_ef{ xo)lets a maximum value. That is, the rate of increase of the func- tion /(expressed in the units of /relative to a unit length in M^m)is maximal
and equal to ||grad f(xo)|| for motion from the point Xo precisely when the
displacement is in the direction of the vector grad f(xo). The value of the function decreases most sharply under displacement in the opposite direction,
and the rate of variation of the function is zero in a direction perpendicular
to the vector grad f(xo).

The derivative with respect to a unit vector in a given direction is usually
known the directional derivative in that direction.

Since a unit vector in Euclidean space is determined by its direction cosines

e —(cos oc cos Ocfft)i

where ai is the angle between the vector e and the basis vector e* in a Cartesian coordinate system, it follows that

$D_e f(xo)=(\text{grad } f(x_0), e)=$ —(a;$_0$)coso;i + ------b ^^-(xo)cosa$_m$.

The vector grad f(xo)is encountered very frequently and has numerous applications. For example the so-known gradient methods for finding extrema of functions of several variables numerically(using a computer)are based on the geometric property of the gradient just noted.

Many important vector fields, such as, for example, a Newtonian gravi- tational field or the electric field due to charge, are the gradients of certain scalar-valued functions, known as the potentials of the fields

Many physical laws use the vector grad /in their very statement. For example, in the mechanics of continuous media the equivalent of Newton's basic law of dynamics ma=F is the relation

pa= —grad p,

which connects the acceleration a=a(x, t)in the flow of an ideal liquid or gas free of external forces at the point x and time t with the density of the medium p=p(x, t)and the gradient of the pressure p=p(x, t)at the same point and time

We shall discuss the vector grad /again later, when we study vector analysis and the elements of field theory.

Check your Progress-1

Discuss Local Properties Of Continuous Functions

___

___

___

Discuss Functions of variable

____

____

____

3.13 LET US SUM UP

In this unit we have discussed the definition and example of Local Properties Of Continuous Functions, Linear Transformations L : Rm —> R™, The Norm In Rm, The Euclidean Structure On Rm, The Differential Of A Function Of Several Variables, The Differential And Partial Derivatives Of A Real-Valued Function, Coordinate Representation Of The Differential Of A Mapping. The Jacobi Matrix, Continuity, Partial Derivatives And Differentiability Of A Function At A Point, The Basic Laws Of Differentiation Linearity Of The Operation Of Differentiation, Differentiation Of A Composition Of Mappings (Chain Rule) The Main Theorem, The Differential And Partial Derivatives Of A Composite Real Valued Function

3.14 KEYWORDS

1. Local Properties Of Continuous Functions: A mapping f : E —)> Rn of a set E C Rm is continuous at a point a G E if and only if uo{ f\ a)=0.

2.Linear Transformations: L : Rm —> R™ We recall that a mapping L : X —» Y from a vector space X into a vector space Y is known linear if

L{ X\xi + X2x2)=X\L(xi)+ X2L(x2)

3.The Norm In Rm The quantity IMI=V(xl)2 H 1-(xm)2 is known the norm of the vector x=(x1, ..., xm)G Rm.

4. The Euclidean Structure On Rm: The concept of the inner product in a real vector space is known from algebra as a numerical function(x, y) defined on pairs of vectors of the space and possessing the properties

5.The Differential Of A Function Of Several Variables: Differentiability and the Differential of a Function at a Point

6. The Basic Laws Of Differentiation Linearity Of The Operation Of Differentiation Theorem If the mappings R^n and /2 : E R^n, defined on a set E C R^m, are differentiate at a point x <= E

7. Differentiation Of A Composition Of Mappings (Chain Rule) The Main Theorem : If the mapping f : X —>• Y of a set X C R^m into a set Y C R^n is differentiate at a point x G X

8.The Differential And Partial Derivatives Of A Composite Real Valued Function Let $z=g(y^1, ..., y^n)$ be a real-valued function of the real variables $y^1, ..., y^n$, each of which in turn is a function

3.15 QUESTIONS FOR REVIEW

Explain Local Properties Of Continuous Functions

Explain Function of Variables

3.16 ANSWERS TO CHECK YOUR PROGRESS

Local Properties Of Continuous Functions

(answer for Check your Progress-1 Q)

Function of Variables

3.17 REFERENCES

- Analysis of Several Variables
- Application of Several Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables

UNIT - IV: DIFFERENTIAL CALCULUS OF REAL-VALUED FUNCTIONS OF SEVERAL VARIABLES

STRUCTURE

4.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables, The Mean-Value Theorem

Learn, Understand about A Sufficient Condition For Differentiability Of A Function Of Several Variables

Learn, Understand about Higher-Order Partial Derivatives

Learn, Understand about Real-Valued Functions Of Several Variables

Learn, Understand about Taylor's Formula

Learn, Understand about Extrema Of Functions Of Several Variables

Learn, Understand about The Implicit Function Theorem

Learn, Understand about Elementary Version Of The Implicit Function Theorem

4.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

The Basic Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables, The Mean-Value Theorem, A Sufficient Condition For Differentiability Of A Function Of Several Variables, Higher-Order Partial Derivatives, Real-Valued Functions Of Several Variables, Taylor's Formula, Extrema Of Functions Of Several Variables, The Implicit Function Theorem, Elementary Version Of The Implicit Function Theorem

4.2 THE BASIC FACTS OF DIFFERENTIAL CALCULUS OF REAL-VALUED FUNCTIONS OF SEVERAL VARIABLES

THE MEAN-VALUE THEOREM

Theorem. Let f : G -» R be a real-valued function defined in a region G C M.m, and Let the closed line segment [x, x + h\ with endpoints x and

x + h be contained in G. If the function f is continuous at the points of the closed line segment [x, x + h] and differentiate at points of the open interval]x, x+ft[, then there exists a point <= E]x, x+ft[such that the following equality holds:

f(x + h)~ f(x)=f'{ <=)h ■

Proof Consider the auxiliary function

F(t)=f(x + th)

defined on the closed interval 0 < t < 1. This function satisfies all the hypotheses of Lagrange's theorem: it is continuous on [0, 1], being the composition of continuous mappings, and differentiable on the open interval]0, 1[,
being the composition of differentiable mappings. Consequently, there exists a point 9 E]0, 1[such that

F(l)- F(0)=F\9)• 1 .

But F(l)=f(x + /i), F(0)= /(x), F'{ 6)=f'{ x + 0 /i)/i, and hence the equality just written is the same as the assertion of the theorem.

We now give the coordinate form of relation.

If x=(x1, ..., xm), h=(ft1, ..., ftm), and <==(x1 + Oh1, ..., xm + 6hm),
means that f(x + h)- f(x)= /(x1 + h1, .. ., xm + hm)- /(x1, ..., xm)=

=   ^«>)(•",,)=

= dx fiOh1 + ---+dmf(Ohm

m   .

= Y, 9if{ x1+6h1, ..., xm + ehm)hi .

2=1

Using the convention of summation on an index that appears as both superscript and subscript, we can finally write

fix1 + h1, ..., xm + hm)- fix1, ..., xm)=

= difix1 + eh1, ..., xm + ehm)ti, (8.54)

where 0 < 6 < 1 and 6 depends on both x and ft.

Remark. Theorem is known the mean-value theorem because there exists
a certain "average" point <= g]x, x + h[at which holds. We have already noted in our discussion of Lagrange's theorem that the mean-value theorem is specific to real-valued functions. A general finite increment theorem for mappings will be proved.

Corollary. If the function f : G -» R is differentiate in the domain G c Mm and its differential equals zero at every point x g G, then f is constant in the domain G.

Proof The vanishing of a linear transformation is equivalent to the vanishing of all the elements of the matrix corresponding to it. In the present case

df(x)h=(5i /, ..., dmf)(x)h,

and therefore d\ f(x)=• • •=dmf(x)=0 at every point x G G.

By definition, a domain is an open connected set. We shall make use of this fact.

We first show that if x g G, then the function /is constant in a ball B(x;r)c G. Indeed, if(x + h)g B(x;r), then [x, x + ft] c B(x;r)c G. Applying relation, we obtain

f(x + ft)- f(x)=f'(Qh=0 • ft=0,

that is, f(x + ft)= /(x), and the values of /in the ball B{ x\ r)are all equal to the value at the center of the ball.

Now Let Xo, Xi g G be arbitrary points of the domain G. By the connectedness of G, there exists a path t h* x(t)g G such that x(0)=xo and x(l)=X\. We let that the continuous mapping t x(t)is defined on the closed interval 0 < t < 1. Let B{ x$\r)be a ball with center at xo contained in G. Since x(0)=Xo and the mapping t h* x(t)is continuous, there is a positive number 8 such that x(t)G B(xo; r)C G for 0 < t < 8. Then, by what has been proved, (/o x)(t)=f(xo)on the interval [0, 8].

Let I=sup 5, where the upper bound is taken over all numbers 8 G [0, 1] such that(/o x)(t)=f(xo)on the interval [0, 5]. By the continuity of the function /(x(<=)) we have /(x(Z))=f(xo). But then 1=1. Indeed, if that were not so, we could take a ball B(x(Z); r)C G, in which f(x)=f(x(l))=f(x0), and then by the continuity of the mapping t x(t)find A > 0 such that x(t)G B(x{ l)\ r)for I < t < I + A. But then(/o x)(t)=f(x(l))=f(xo)for
0 <t<l + A, and so I ^ sup 8.

Thus we have shown that(fox)(t)=f(xo)for any t G [0, 1]. In particular (f°x)(1)=f(xi)= /(x0), and we have verified that the values of the function /: G -> R are the same at any two points xo, x\ G G.

4.3 A SUFFICIENT CONDITION FOR DIFFERENTIABILITY OF A FUNCTION OF SEVERAL VARIABLES

Theorem. Let f : U(x)-» R be a function defined in a neighborhood U(x)C Mm of the point x=(x1, ..., xm).

If the function f has all partial derivatives,..., at each point of the neighborhood U(x)and they are continuous at x, then f is differentiate at x.

Proof Without loss of generality we shall let that U(x)is a ball B{ x\r). Then, together with the points x=(x1, ...^™)and x + h=(x1 + ft1, ...,xm + /im), the points(x\x2 + /i2, ..., xm + /im), ..., (x1, x2, ..., xm_1, xm + hm)and the lines connecting them must also belong to the domain U(x). We shall use this fact, applying the Lagrange theorem for functions of one variable in the following computation:

f(x + h)- f(x)=fix1 +h1, ..., xm + hm)- fix1, ...,xm)=
= fix1 +H1, ..., xm + hm) — fix1, x2 +h2, ..., xm + hm)+

Notes

+ fixx, x2 + h2, ...,xm + hm) — fix1,x2,x2 + h3, ...,xm + hm)H b

+ fix1,x2, ..., xm-\xm + hm)- fix1,..., Xm)=

= di fix1 +01h1, x2 + h2, ..., xm + hm)h } +

+ d2fixx, x2 4- d2h2, x3 + h, 3, ...,xm + hm)h2 H b

+ Omfix^x2, x™'1, xm + emhm)hm .

So far we have used only the fact that the function /has partial derivatives with respect to each of its variables in the domain U(x). We now use the fact that these partial derivatives are continuous at x.
Continuing the preceding computation, we obtain

f(x + h)- f(x)=di /(x1, .. ., xm)h } + a1^1 +

+d2f(x1, .. ., xm)h2 + a2h2 H h

+dmf(x1, ..., xm)hm+amhm,

where the quantities a \,..., am tend to zero as h —» 0 by virtue of the continuity of the partial derivatives at the point x.

But this means that

f(x + h) — f(x)=L(x)h + o(h)as h -» 0,

where L(x)h=<9i /(x\..., xm)/i1 H h dmf(x\

It follows from Theorem that if the partial derivatives of a function /: G -» R are continuous in the domain G C Rm, then the function is differentiable at that point of the domain.

Let us agree from now on to use the symbol C^^(G;R), or, more simply, C^(G)to denote the set of functions having continuous partial derivatives in the domain G.

4.4 HIGHER-ORDER PARTIAL DERIVATIVES

If a function /: G R defined in a domain G C Rm has a partial derivative with respect to one of the variables x1, ..., xm, this partial derivative is a function dif : G -» R, which in turn can have a partial derivative dj(dif)(x)with respect to a variable xK

The function dj(dif): G -» R is known the second partial derivative of f with respect to the variables x% and and is denoted by one of the following
symbols:

d2f

dxi dx1

The order of the indices indicates the order in which the differentiation is carried out with respect to the corresponding variables.

We have now defined partial derivatives of second order.

If a partial derivative of order k

= dJ.'.dxi.(*>

has been defined, we define by induction the partial derivative of order k + 1
by the relation

dih-ij(x):=ai(ail...ifc/)(x).

At this point a question arises that is specific for functions of several variables: Does the order of differentiation affect the partial derivative computed?

4.5 REAL-VALUED FUNCTIONS OF SEVERAL VARIABLES

Theorem. If the function f : G -» R has partial derivatives

i W 5

dxidx^ dxidx1

Notes

in a domain G, then at every point x G G at which both partial derivatives are continuous, their values are the same.

Proof Let x <= G be a point at which both functions dijf : G -» R and djif G —» R are continuous. From this point on all of our arguments are carred out in the context of a ball B(x;r)C G, r > 0, which is a convex neighborhood of the point x. We wish to verify that

^ ^ -(x1 Xm)=d f /gl

dxidx^ ""   dx^dx1

Since only the variables x* and xJ will be changing in the computations to follow, we shall let for the sake of brevity that /is a function of two variables /(x1, #2), and we need to verify that

df(x\x2)=   { x\x2),

dx1dx2 '   dx2dx1

if the two functions are both continuous at the point(x1, x2).

Consider the auxiliary function

F^ft1, ft2)= /(x1 + ft1, x2 + ft2) — /(x1 + ft1, x2) — /(x1, x2 + ft2)+ /(x1, x2),

where the displacement ft=(ft1, ft2)is letd to be sufficiently small, namely so small that x + ft G B(x; r).

If we regard F(ftx, ft2)as the difference

Fih1, ^)=ip{ 1)- <p(0),

where y>(<=)= /(x1 + 2ft1, x2 + ft2) — /(x1 + 2ft1, x2), we find by Lagrange's theorem that

F(ftx, ft2)=^'(^i)=(d\f(x1 + 6\hl,x2 + ft2) — ^ /(x1 + flift^x2)) /^1 .
Again applying Lagrange's theorem to this last difference, we find that

F(hl, h2)=^ /(x1 + #ift\x2 + 62h2)h2hl .

If we now represent F(ftx, ft2)as the difference
F(/i\ /I2)=^(1)-^(0),

where(p{ t)= /(x1 + ftx, x2 + th2) — /(xx, x2 + 2ft2), we find similarly that

F(hl, h2)=d\2f(xl + 6ihl, x2 + 62h2)hlh2 .

Comparing we conclude that

021 /(x1 + 6ihl, x2 + 02h2)=d\2f{ xl + 6\hl, x2 + 02h2),

where 6i, 62, 6 \,62 G]0, 1[. Using the continuity of the partial derivatives at

the point(x1, x2), as ft —» 0, we get the equality we need as a consequence of d2if(x1, x2)=di2f{ xl, x2). □

We remark that without additional assumptions we cannot say in general that dijf(x)=djif(x)if both of the partial derivatives are defined at the point x(observe Problem 2 at the end of this section).

Let us agree to denote the set of functions /: G —» R all of whose partial derivatives up to order k inclusive are defined and continuous in the domain G C Rm by the symbol C«(G; R)or C^(G).

Proposition . Iff G C^k\G\ R), the value di1...ikf(x)of the partial derivative is independent of the order i \,..., i^ of differentiation, that is, remains the same for any permutation of the indices i \,..., zV

Proof. In the case k=2 this proposition is contained

Let us let that the proposition holds up to order n inclusive. We shall show that then it also holds for order n + 1.

But dili2...in+1f(x)=dil(di2...in+1f)(x). By the induction assumption the indices i2, ...,in+i can be permuted without changing the function di2."in+if(x)i and hence without changing di1...in+1f(x). For that reason it

suffices to verify that one can also permute, for example, the indices i\ and i2 without changing the value of the derivative di1i2...in+1f(x).

Since

0»i<2-i«+i /(s)=9ili2{ di3."in+1f)(x),

the possibility of this permutation follows immediately By the induction principle

Example. Let /(x)= /(xx, x2)be a function of class C^(G;R).

Let h=(ft1, ft2)be such that the closed interval [x, x + ft] is contained in the domain G. We shall show that the function

<p(t)=f(x + th),

which is defined on the closed interval [0, 1], belongs to class C^[0, 1] and find its derivative of order k with respect to t.

We have

cp'(t)=d\f(xl + tft1, x2 + ^ft2)ft1 +d2f(x1 + ^ft1, x2 + <=ft2)ft2,
ip"(t)=dn f(x + th)hlhl + d2\f(x + ^ft)ft2ftx +

+ <9i2 f(x + ^ft)ftxft2 + d22 f(x + th)h2h2 =
= <9n /(x + ^ft)(ftx)2 + 2<9i2 /(x + ^ft)ftxft2 + d22f(x + ^ft)(ft2)2 .

These relations can be written as the action of the operator(hld\ + h2):

ip'(t)=(hld 1 + h2d2)f(x + th)=hldif(x + th),

<p"{ t)=(h }d\ + h2d2)2f(x + th)=hllhl2dili2f(x + th).

By induction we obtain

cp^k\t)=(h }d\ + h2d2)kf(x + th)=h11 - " h%kdil...ikf{ x + th)

(summation over all sets i \,..., «<= of k indices, each assuming the values 1 and 2, is meant).

Example. If f(x)= /(a;1, ..., #™)and /G C^(G;R), then, under the assumption that [x, x + h] C G, for the function(p(t)=f(x + th)defined on the closed interval [0, 1] we obtain

ip^k\t)=h11 - • • h'lkdi1...ikf(x + th), where summation over all sets of indices z'i, ..., each assuming all values from 1 to m inclusive, is meant on the right.

We can also write formula as

= { rtdi + • • • + hmdm)kf(x + th)

4.6 TAYLOR'S FORMULA

Theorem:. If the function f : U(x)-» R is defined and belongs to class C<n>(Cf(s); R)in a neighborhood U(x)C Rm of the point x G Rm, and the closed interval [x, x + h] is completely contained in U{ x), then the following equality holds:

/(x1 + h1, ..., xm + hm)- f(xx, ..., xm)=

n—1

= ^2 T \(hldi "l•" hmdm)kf(x)+ rn-i(x; h),

k—1 *

where

1 n 1

rn_1(»;ft)=J(|(fe1^ + • • • + hmdm)nf(x + th)dt.

Taylor's formula with integral form of the remainder.
Taylor's formula follows immediately from the corresponding Taylor formula for a function of one variable. In fact, consider the auxiliary function

(p(t)=f(x + th),

which, by the hypotheses is defined on the closed interval
0 < t < 1 and(as we have verified above)belongs to the class C^[0, 1].

Then for r G [0, 1], by Taylor's formula for functions of one variable, we can write that

<P(r)=<p(0)+ ^'(0)r + • • • + +

Notes

+ }i10Wv*, { tT }T'"it'

0

Setting r=1 here, we obtain

V(1)=v(0)+ ^'(0)+ • • • +(^yn-"(0)+

+ /(1(„"-"i)!ly""(')'it-(8'62)

0

Substituting the values

v?(fc)(0)=(h'd, + • • • + hmdm)kf{ x)(k=0, ..., n - 1),

^n\t)=(h }d 1 + • • • + hmdm)nf(X + th),

into this equality in accordance with formula

Remark. If we write the remainder term in relation in the Lagrange form rather than the integral form, then the equality

<p(i)=m + jj<p'(o)+ ■■■ + + ±r(n)(0),

where 0 < 9 < 1, implies Taylor's formula with remainder term

rn_i(x; h)=^(hldi H b hmdrn)nf(x + 6h).

This form of the remainder term, as in the case of functions of one variable is known the Lagrange form of the remainder term in Taylor's formula.

Since /G(U(x); R), it follows from that

rwfo h)=^{ h }d 1 + • • • + hmdm)nf{ x)+ o(\\h\r)as h -> 0,
and so we have the equality

fix1 + h\..., xm + hm)-f(x1, ..., xm)=

= it(^1 + • • • + hmdm)kf(x)+ o(||*P)as h -> 0

k=1

known Taylor's formula with the remainder term in Peano form.

4.7 EXTREMA OF FUNCTIONS OF SEVERAL VARIABLES

One of the most important applications of differential calculus is its use in finding extrema of functions.

Definition. A function /: E -» R defined on a set E C Rm has a local maximum(resp. local minimum)at an interior point Xo of E if there exists a neighborhood U(xo)C E of the point xo such that f(x)< f(xo)(resp.f{ x)> f(xo)) for all x G U(x0).

If the strict inequality f(x)< f(xo)holds for x G U(xo)\ xo(or, respectively, f(x)> f(xo)), the function has a strict local maximum(resp. strict local minimum)at xo.

Definition. The local minima and maxima of a function are known its local extrema.

Theorem. Suppose a function f : U(xo) —» R defined in a neighborhood U(xq)C Rm of the point xo=(xj, ..., x^)^as partial derivatives with respect to each of the variables x1, ..., xm at the point xo-

Then a necessary condition for the function to have a local extremum at xo is that the following equalities hold at that point:

= =

Proof. Consider the function(p(xx)= /(x1, Xq, ..., x™)of one variable defined, according to the hypotheses of the theorem, in some neighborhood of the point xj on the real line. At xj the function(fix1)has a local extremum, and since

<p'(xo)=§^{ xl4, ..., xV),

it follows that Jjt(#o)=0.

The other equalities are proved similarly.

Notes

We call attention to the fact that relations give only necessary but not sufficient conditions for an extremum of a function of several variables.

An example that confirms this is any example constructed for this purpose for functions of one variable. Thus, where previously we spoke of the function x h* x3, whose derivative is zero at zero, but has no extremum there

f(x\..., xm)=(x1)3,

all of whose partial derivatives are zero at Xo=(0, ..., 0), while the function obviously has no extremum at that point.

Theorem shows that if the function /: G —> R is defined on an open set G C Rm, its local extrema are found either among the points at which /

is not differentiate or at the points where the differential d /(#o)or, what is the same, the tangent mapping f'(xo), vanishes.

We know that if a mapping /: U(xo)-» Rn defined in a neighborhood U(xo)C Rm of the point Xo G Rm is differentiate at #o, then the matrix of the tangent mapping f'(xo): Rm -» Rn has the form

d1f1(x0)■■■ dmf1(xo)\

difn(x0)■ • ■ dmfn(x0)J

Definition. The point Xo is a critical point of the mapping f : U(xo) —» Rn if the rank of the Jacobi matrix of the mapping at that point is less than min{ m, n }, that is, smaller than the maximum possible value it can have.

In particular, if n=1, the point Xo is critical if condition holds, that is, all the partial derivatives of the function /: U(xo)-» R vanish.

The critical points of real-valued functions are also known the stationary points of these functions.

After the critical points of a function have been found by solving the system, the subsequent analysis to determine whether they are extrema or not can often be carried out using Taylor's formula and the following sufficient conditions for the presence or absence of an extremum provided by that formula.

Theorem. Let f : U(xo)-» R be a function of class C^(U(xo);R)de- fined in a neighborhood U(xo)C Rm of the point xo=(xj, ...,x™)G Rm, and Let xo be a critical point of the function f.

If in the Taylor expansion of the function at the point xq

f(xo + h\ ...,x^ + hm)

1 m <=p. f

= /(®o."-.®o*)+ 2! S datdx* +

$$\sum^{m} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)h^i h^j \equiv \partial_{ij} f(x_0)h^i h^j$$

is positive-definite or negative-definite, then the point xo has a local extremum at xo, which is a strict local minimum if the quadratic form is positive-definite and a strict local maximum if it is negative definite lets both positive and negative values, then the function does not have an extremum at xq .

Proof. Let h 0 and xq + h G U(xq). Let us represent in the form where o(l)is infinitesimal as h —» 0. It is clear from that the sign of the difference f(xo + h) — f(xo)is compLetely determined by the sign of the quantity in brackets. We now undertake to study this quantity.

The vector e=(h1/\ \h\ \,...,hm/\ \h\\)obviously has norm 1. The quadratic form is continuous as a function h G Rm, and therefore its restriction to the unit sphere 5(0; 1)={ x G Rm| ||x||=1 } is also continuous on 5(0; 1). But the sphere 5 is a closed bounded subset in Rm, that is, it is compact. Consequently, the form has both a minimum point and a maximum point on 5, at which it lets respectively the values m and M.

Notes

If the form is positive-definite, then $0 < m < M$, and there is a number $5 > 0$ such that $|o(l)| < m$ for \\h\\ < 5. Then for \\h\\ < 5 the bracket on the right-hand side of is positive, and consequently $f(xo + h) — f(xo) > 0$ for $0 <$ \\h\\ < 6. Thus, in this case the point xo is a strict local minimum of the function.

One can verify similarly that when the form is negative-definite, the function has a strict local maximum at the point xq.Thus Let em and eM be points of the unit sphere at which the form lets the values m and M respectively, and Let $m < 0 < M$. Setting h=tem, where t is a sufficiently small positive number(so small that Xo + tem <= U(xo)), we find that

/(x0 + tem)- f(xo)=^t2(m + o(l)),

where o(l) —> 0 as t —> 0. Starting at some time(that is, for all sufficiently small values of t), the quantity m+o(l)on the right-hand side of this equality will have the sign of m, that is, it will be negative. Consequently, the left-hand side will also be negative, Similarly, setting h=teM, we obtain

f(xo + teM)- f(xo)=7y t2(M + o(l))

and consequently for all sufficiently small t the difference f(xo + ^m) — /(#o)is positive.

Thus, if the quadratic form lets both positive and negative values on the unit sphere, or, what is obviously equivalent, in Rm, then in any neighborhood of the point xo there are both points where the value of the function is larger than f(xo)and points where the value is smaller than f(xo). Hence, in that case xo is not a local extremum of the function.

We now make a number of remarks in connection with this theorem.

Remark . Theorem says nothing about the case when the form is semi-definite, that is, nonpositive or nonnegative. It turns out that in this case the point can be an extremum, or it can not. This can be observen, in particular from the following example.

Example. Let us find the extrema of the function /(#, y)=x4 + y4 — 2x2 which is defined in R2.

In accordance with the necessary conditions we write the system of equations

= 4x3 - 4x=0,

<

^(x, y)=4y3=0,

from which we find three critical points:(—1, 0), (0, 0), (1, 0).

Since

0(x, »)=12**-4, 0(x, »)=12!, 2,

at the three critical points the quadratic form has respectively the form

8(^)2, -A(h })2, 8(^)2 .

That is, in all cases it is positive semi-definite or negative semi-definite. Theorem is not applicable, but since /(#, y)=(x2 — l)2 + y4 — 1, it is obvious that the function f(x, y)has a strict minimum —1(even a global minimum)at the points(—1, 0), and(1, 0), while there is no extremum at(0, 0), since for x=0 and y ^ 0, we have /(0, y)=y4 > 0, and for y=0 and sufficiently small x ^ 0 we have /(#, 0)=x4 — 2x2 < 0.

Remark. After the quadratic form has been obtained, the study of its definiteness can be carried out using the Sylvester5 criterion. We recall

m

that by the Sylvester criterion, a quadratic form with symmetric

ijj=i

matrix

(flu • * • aim ^

\ ®ml ' ' ' ^mm J

is positive-definite if and only if all its principal minors are positive; the form is negative-definite if and only if an <0 and the sign of the principal minor reverses each time its order increases by one.

Example• Let us find the extrema of the function

f(x, y)=xy\n(x2 + y2),

except at the origin.

which is defined everywhere in the plane
Solving the system of equation

~(x, y)=a: In(x2 + y2)+ ^ =0,
,, ay xz + yz

we find all the critical points of the function

<0, ±1);(±1, 0);(± '±');(± -L T')Since the function is odd with respect to each of its arguments individually, the points(0, d=l)and(±1, 0)are obviously not extrema of the function.

It is also clear that this function does not change its value when the signs of both variables x and y are changed. Thus by studying only one of the remaining critical points, for example, we be able t° draw

conclusions on the nature of the others.

Since local maxima of the function. This, however, could have been verified
directly, by checking the definiteness of the corresponding quadratic form.
For example, at the point(— ^==, ^==)the matrix of the quadratic form has the form from which it is clear that it is negative-definite.

Remark . It should be kept in mind that we have given necessary conditions and sufficient conditions for an extremum of a function only at an interior point of its domain of definition. Thus in observeking the absolute maximum or minimum of a function, it is necessary to examine the boundary points of the domain of definition along with the critical interior points, since the function can let its maximal or minimal value at one of these boundary points.

The general principles of studying noninterior extrema will be considered in more detail later(observe the section devoted to extrema with constraint). It is useful to keep in mind that in searching for minima and maxima one can use certain simple considerations connected with the nature of the problem along with the formal techniques, and sometimes even instead of them. For example, if a differentiate function being studied in Rm must have a minimum because of the nature of the problem and turns out to be unbounded above, then if the function has only one critical point, one can assert without further investigation that that point is the minimum.

Example. Huygens problem. On the basis of the laws of conservation of energy and momentum of a closed mechanical system one can show by a simple computation that when two perfectly elastic balls having mass m\ and m2 and initial velocities v\ and V2 collide, their velocities after a central collision(when the velocities are directed along the line joining the centers) are determined by the relations

„(mi - m2)ui + 2rri2V2

vi= —

mi + m2

(m2 - mi)v2 + 2m\V\

mi + m2

In particular, if a ball of mass M moving with velocity V strikes a motionless ball of mass m, then the velocity v acquired by the latter can be found from the formula

V=2M V,

m + M v 1

from which one can observe that if 0 < m < M, then V < v < 2V.

How can a significant part of the kinetic energy of a larger mass be communicated to a body of small mass? To do this, for example, one can insert
balls with intermediate masses between the balls of small and large mass: m < mi < m2 < • • • < mn < M. Let us compute(after Huygens)how the masses mi, m2, ..., mn should be chosen to that the body m will acquire maximum velocity after successive central collisions.

In accordance with formula we obtain the following expression for the required velocity as a function of the variables mi, m2, ..., mn:

v= —— ^-2-+1V.

m + mi mi + m2 mn_i + mn mn + M

Thus Huygens' problem reduces to finding the maximum of the function

mi mn M

f{ m \,...,mn)=

m + mi mn_i+mn mn + M

The system of equations, which gives the necessary conditions for an interior extremum, reduces to the following system in the present case:

m • m2 — ml=0,
mi • ms — m%=0,

mn-1 ■ M -17%=0,

from which it follows that the numbers m, mi, ..., mn, M form a geometric
progression with ratio q equal to n+y /M /m.

The value of the velocity that results from this choice of masses is given by which agrees with if n=0.

It is clear from physical considerations that formula gives the maximal value of the function. However, this can also be verified formally (without invoking the cumbersome second derivatives).

We remark that it is clear from that if m —> 0, then v —>> 2n+1V. Thus the intermediate masses do indeed significantly increase the portion of
the kinetic energy of the mass M that is transmitted to the small mass m.

4.8 THE IMPLICIT FUNCTION THEOREM

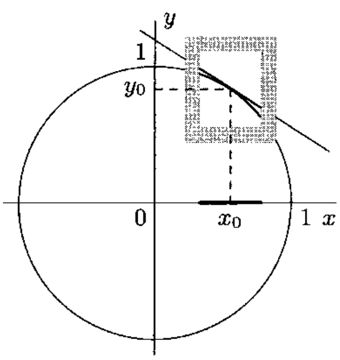
Statement of the Problem and Preliminary Considerations

In this section we shall prove the implicit function theorem, which is important both intrinsically and because of its numerous applications.

Let us begin by explaining the problem. Suppose, for example, we have the relation

x2+y2-1=0

between the coordinates x, y of points in the plane M2. The set of all points
of M2 satisfying this condition is the unit circle



The presence of the relation shows that after fixing one of the coordinates, for example, x, we can no longer choose the second coordinate
arbitrarily. Thus relation determines the dependence of y on x. We are interested in the question of the conditions under which the implicit relation
can be solved as an explicit functional dependence y=y{ x).

we find that

y=±yjl -x2

that is, to each value of x such that |x| < 1, there are actually two admissible
values of y. In forming a functional relation y=y{ x)satisfying relation one cannot give preference to either of the values without invoking additional requirements. For example, the function y(x)that lets the value +a /1 — x2 at rational points of the closed interval [—1, 1] and the value
—a /1 — x2 at irrational points obviously satisfies

It is clear that one can create infinitely many functional relations satisfying by varying this example.

The question whether the set defined in M2 by is the graph of a function y=y(pc)obviously has a negative answer, since from the geometric point of view it is equivalent to the question whether it is possible to establish
a one-to-one direct projection of a circle into a line.

But observation suggests that nevertheless, in a neighbor-
hood of a particular point(#o, 2 /o)the arc projects in a one-to-one manner
into the x-axis, and that it can be represented uniquely as y=y(x), where x y(x)is a continuous function defined in a neighborhood of the point xo and assuming the value yo at xo- In this aspect, the only bad points are
(—1, 0)and(1, 0), since no arc of the circle having them as interior points projects in a one-to-one manner into the x-axis. Even so, neighborhoods

of

these points on the circle are well situated relative to the y-axis, and can be represented as the graph of a function x=x(y)that is continuous in a neighborhood of the point 0 and lets the value —1 or 1 according as the arc in question contains the point(—1, 0)or(1, 0).

How is it possible to find out analytically when a geometric locus of points

defined by a relation of the type can be represented in the form of an explicit function y=y(x)or x=x{ y)in a neighborhood of a point

on the locus?

We shall discuss this question using the following, now familiar, method. We have a function F(x, y)=x2 + y2 — 1. The local behavior of this function

in a neighborhood of a point is well described by its differential

K(x0, yo)(x - ®o)+ Fy(XO, yo)(y - 2 /0),

since

F(x, y)=F(x0, 2 /o)+ K(xo, 2 /o)(« - x0)+

+ Fy(xo, 2 /0)(2/- 2 /o)+ o(\x - x0\ + \y~ 2 /o|)

as(x, y)->•(x0, y0).

If F(x0, yo)=0 and we are interested in the behavior of the level curve

F(x, y)=0

of the function in a neighborhood of the point we can judge that behavior from the position of the(tangent)line

K(xo, Vo)(x - x0)+ Fy(x0, 2 /0X2/- 2 /o)=0 .

If this line is situated so that its equation can be solved with respect to 2 /, then, since the curve F{ x, y)=0 differs very little from this line in a neighborhood of the point we can hope that it also can be written

in the form y=y(x)in some neighborhood of the point-

Notes

The same can be said about local solvability of F(x, y)=0 with respect to x.

Writing for the specific relation, we obtain the following equation for the tangent line:

x0(x - #0)+ 2 /0(2/ — 2 /o)=0 .

This equation can always be solved for y when 2 /0 7^ 0, that is, at all points
of the circle except(—1, 0)and(1, 0). It is solvable with respect to x at all points of the circle except(0, -1)and(0, 1).

4.9 ELEMENTARY VERSION OF THE IMPLICIT FUNCTION THEOREM

In this section we shall obtain the implicit function theorem by a very intuitive, but not very constructive method, one that is adapted only to the case of real-valued functions of real variables. The reader can become familiar with another method of obtaining this theorem, one that is in many ways
preferable, and with a more detailed analysis is an elementary version of the implicit function theorem.

Proposition. If the function F : U(xo, yo) —> K defined in a neighborhood
U(xo, yo)of the point(xo, yo)<= R2 ^ siac /i that
1° F e C^\U;R), where p> I,

2° F(aj0, 2 /0)=0,

3° Fy{ x0, yo)7^ 0,

<= /&en <= /&ere exist a two-dimensional interval I=Ix x Iy where

Ix={ x<= R| I a; - Zol < a },   = { y e K| |y - y0| < /? }, that is a neighborhood of the point(xo, yo)contained in U{ xo, yo), and a function f G C^p\lx\Iy)such that

F(x, y)=0&y= /(x)

for any point(x, y)G Ix x Iy and the derivative of the function y=f(x)at the points x G Ix can be computed from the formula

f'(x)=-[F^xJix^-^F^xJix))] .

Before taking up the proof, we shall give several possible reformulations of the conclusion, which should bring out the meaning of the relation itself.

Proposition says that under hypotheses 1°, 2°, and 3° the portion of the set defined by the relation F{ x, y)=0 that belongs to the neighborhood
Ix x Iy of the point(xo, 2 /o)1S the graph of a function /: Ix —> Iy of class
C<*\lx;Iy).

In other words, one can say that inside the neighborhood I of the point (x0, yo)the equation F(x, y)=0 has a unique solution for y, and the function
y= /(x)is that solution, that is, F(x, /(«))=0 on Ix.

It follows in turn from this that if y=f{ x)is a function defined on Ix that is known to satisfy the relation F(x, /(x))=0 on /x, /(xo)=yo, and this function is continuous at the point xo G Ix, then there exists a neighborhood
Aclx of xo such that f(A)C Iy, and then /(x)= /(x)for x G A.

Without the assumption that the function /is continuous at the point xo and the condition f(xo)=yo, this last conclusion could turn out to be incorrect, as can be observen from the example of the circle already studied.
Let us now prove Proposition.

Proof Suppose for definiteness that Fy(xo, yo)> 0. Since F G C^l\U\ R), it follows that Fy(x, y)> 0 also in some neighborhood of(xo, 2 /o)- In order

to avoid introducing new notation, we can let without loss of generality that $F_y(x, y) > 0$ at every point of the original neighborhood $U(x_o, y_o)$-

Moreover, shrinking the neighborhood t /(xo, 2 /o)if necessary, we can as-

sume that it is a disk of radius r=2 /3 > 0 with center at(xq, yo)•

Since $F_y(x, y) > 0$ in [/, the function $F(x_o, y)$ is defined and monotonically

increasing as a function of y on the closed interval yo —(3 < y < yo + /?.

Consequently,

F(xo, y0 - 0)< F(x0, yo)=0 < F(x0, yo + P)

By the continuity of the function F in [/, there exists a positive number a < p such that the relations

F(x, y0 —(3)<0 < F(x, y0 + 0)

hold for \x — xo\ < ot>

We shall now show that the rectangle $I = I_x \times I_y$, where

Ix={ x € E| I a: - x0| < a }, Iy={ y G R| \y - y0)| < 0 },

is the required two-dimensional interval in which relation holds.

For each x G Ix we fix the vertical closed interval with endpoints(x, yo —

/?), (x, yo +(3). Regarding F{ x, y)as a function of y on that closed interval,

we obtain a strictly increasing continuous function that lets values of opposite sign at the endpoints of the interval. Consequently, for each x G ix,

there is a unique point y{ x)G Iy such that F(x, ? /(x))=0. Setting y(x)=/(x) .

We now establish that /G C^p\lx\Iy).

We begin by showing that the function /is continuous at xo and that f(xo)=yo- This last equality obviously follows from the fact that for x=Xo
there is a unique point y(xo)G Iy such that F(xo, y(xo))=0. At the same time, F(x0, 2 /o)=0, and so /(x0)=yo-

Given a number e, 0 < e <(3, we can repeat the proof of the existence of the function /(x)and find a number 5, 0 < S < a such that in the two-dimensional interval I=Ix x Iy, where

Ix={ xe R| |x - x0| < S }, iy={ y G R| \y - y0\ < e },

the relation

(F(x, y)=0 in I)(y=f(x), x € Ix)

holds with a new function /: Ix Iy.

But Ix C /x, Iy C Iy, and I C /, and therefore it follows from that /(x)= /(x)for x G Ix C Ix. We have thus verified that

1 /0*0 - f(xo)l=1 /0*0 -yo\<<= for \x - Xol < <5-

We have now established that the function /is continuous at the point xo- But any point(x, y)g I at which F(x, y)=0 can also be taken as the initial point of the construction, since conditions 2° and 3° hold at that
point. Carrying out that construction inside the interval /, we would once again arrive via at the corresponding part of the function /considered in a neighborhood of x. Hence the function /is continuous at x. Thus we have established that /G C(Ix;Iy).

We shall now show that /G C^(Ix;Iy)
Let the number Ax be such that
x + Ax G Ix. Let y=f(x)and y + Ay =
f(x + Ax). Applying the mean-value theorem to the function F{ x, y)inside
the interval I, we find that

0=F{ x + Ax, f(x + Ax)) - F{ x, f(x)) =

= F(x + Ax, y + Ay)- F{ x, y)=

= Fx(x + 6Ax, y + 6Ay)Ax + Fy(x + 6Ax, y + 6Ay)Ay(0 < 0 < 1), from which, taking account of the relation Fy(x, y)^ 0 in I, we obtain Ay F'x(x + 0Ax, y + OAy)

Ax Fy(x + OAx, y + OAy)

Since /G C(Ix;Iy), it follows that Ay —> 0 as Ax —> 0, and, taking account of the relation F G C^(U; R), as Ax —> 0

f'(x)=Fx(*, ?/)

nx } F'{ x, y)'

where y=f(x).

By the theorem on continuity of composite functions that /G C^\lx;Iy).

If F G C^2\U', R), the right-hand side can be differentiated with respect to x, and we find that

r(x)=IK'X + Ky • f(x)]Fj - F'x[F^y + F^y ■ f'(x)} ^(g 84/)

(■^3 /)2

where Fx, Fy, Fxx, Fxy, and Fyy are all computed at the point(x, f(x)).

Thus /G C{ Ix, Iy)if F G C^(U; R). Since the order of the derivatives of /on the right-hand side and so forth, is one less than the order on the left-hand side of the equality, we find by induction that /G C^(IX; Iy)if F G C(p\U; R)

Example. Let us return to relation studied above, which defines a circle in R2, and verify Proposition on this example.

In this case

F(x, y)=x2 + y2 - 1,
and it is obvious that F G C^°°^(R2;R). Next,

Fx{ x, y)=2x, F'y{ x, y)=2y,

so that Fy(x, y)^ 0 if y ^ 0. Thus, by Proposition 1, for any point(#o, Vo)of
this circle different from the points(—1, 0)and(1, 0)there is a neighborhood such that the arc of the circle contained in that neighborhood can be written
in the form y= /(#). Direct computation confirms this, and f{ x)=\ /l — x2
or f(x)=-\ /l - x2.

Next, by Proposition

f(x0)=-F^V0\=-- •

Fy(xo, yo)2 /o

f'(x)="t^t=--,

f(x)y

and computation with it leads to the same result,

/•// \

/(%o)= ,

2 /o

as computation from formula obtained from Proposition

It is important to note that compute ff(x)without even having an explicit expression for the relation y= /(#), if only we know that f(xo)=yo•
The condition y$=f(xo)must be prescribed, however, in order
to distinguish the portion of the level curve
F(x, y)=0 that we intend to describe in the form y=f(x).

It is clear from the example of the circle that giving only the coordinate does not determine an arc of the circle, and only after fixing yo have we distinguished one of the two possible arcs in this case.

Notes

Check your Progress-1

Discuss Differential Calculus Of Real-Valued Functions Of Several Variables

____

____

____

Discuss Higher-Order Partial Derivatives

____

____

____

4.10 LET US SUM UP

In this unit we have discussed the definition and example of The Basic Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables, The Mean-Value Theorem, A Sufficient Condition For Differentiability Of A Function Of Several Variables, Higher-Order Partial Derivatives, Real-Valued Functions Of Several Variables, Taylor's Formula, Extrema Of Functions Of Several Variables, The Implicit Function Theorem, Elementary Version Of The Implicit Function Theorem

4.11 KEYWORDS

1. The Basic Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables …The Mean-Value Theorem: Let f : G -» R be a real-valued function defined in a region G C M.m

2. A Sufficient Condition For Differentiability Of A Function Of Several Variables…. Let f : U(x)-» R be a function defined in a neighborhood U(x)C Mm of the point x=(x1, ..., xm).

3. Higher-Order Partial Derivatives If a function /: G R defined in a domain G C Rm has a partial derivative with respect to one of the

variables

4. Real-Valued Functions Of Several Variables…. The function f : G -» R has partial derivatives.

5. Taylor's Formula: If the function f : U(x)-» R is defined and belongs to class C<n>(Cf(s).

6. Extrema Of Functions Of Several Variables: One of the most important applications of differential calculus is its use in finding extrema of functions.

4.12 QUESTIONS FOR REVIEW

Explain Differential Calculus Of Real-Valued Functions Of Several Variables

Explain Higher-Order Partial Derivatives

4.13 ANSWERS TO CHECK YOUR PROGRESS

Differential Calculus Of Real-Valued Functions Of Several Variables

(answer for Check your Progress-1 Q)

Higher-Order Partial Derivatives

(answer for Check your Progress-1 Q)

4.14 REFERENCES

- Function of Variables
- System of Equation
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables

UNIT - 5: TRANSITION TO THE CASE OF A RELATION

STRUCTURE

5.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about Transition To The Case Of A Relation

Learn, Understand about The Implicit Function Theorem

Learn, Understand about Some Corollaries Of The Implicit Function Theorem

The Inverse Function Theorem

Learn, Understand about Differential Calculus In Several Variables

Learn, Understand about Differential Calculus In Several Variables Functional Dependence

Learn, Understand about Local Resolution Of A Diffeomorphism Into A Composition Of Elementary Ones

Learn, Understand about Morse's Theorem

5.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Transition To The Case Of A Relation, The Implicit Function Theorem, Some Corollaries Of The Implicit Function Theorem, The Inverse Function Theorem, Differential Calculus In Several Variables, Differential Calculus In Several Variables Functional Dependence, Local Resolution Of A Diffeomorphism Into A Composition Of Elementary Ones, Morse's Theorem

5.2 TRANSITION TO THE CASE OF A RELATION

F(x1, ..., a?m, y)=0

The following proposition is a a simple generalization of Proposition to the case of a relation F(xx, ..., xm, y)=0.

Proposition. If a function F : U —> R defined in a neighborhod U C Rm+1 of the point(x0, yo)=(xb ■ • ■ > ^o1.2 /o)€ ®m+1 is such that
1° FeC^(U -,R), p> 1,

2° F(xo, j /o)=F{ x \,..., a#\ 2 /0)=0,

3° Fy(xo, yo)=Fy(xl, ..., zg1, 2 /0)+ 0,

then there exists an(m + 1)-dimensional interval I=I™ x l }j; where

I™={ x=(x1, ..., xm)<= Rm| \xl - Xq| < ot% i=1, ...,m },

l\={ y <= R| \y-yo\ < /? },

which is a neighborhood of the point(#o, 2 /o)contained in U, and a function

f <= Iy)such that for any point(x, y)<= I™ x I1

F(x1, ..., xm, y)=0= /(x1, ..., xm)

and the partial derivatives of the function y <= fix1, ..., xm)at the points of Ix can be computed from the formula

f<=(®)=-[^(®. /(®))]_1[^(®. /(®))] •

Proo /. The proof of the existence of the interval /m+1=I™ x Iy and the existence of the function y=f{ x)=fix1, ..., xm)and its continuity in I™ is a verbatim repetition of the corresponding part of the proof of Proposition with only a single change, which reduces to the fact that the symbol x must now be interpreted as(x1, ..., xm)and a as(a1, ..., am).

If we now fix all the variables in the functions Fix1, ...,xm, y)and /(x1, ..., xm)except xl and y, we have the hypotheses of Proposition where now the role of x is played by the variable x\ Formula follows from this. It is clear from this formula that <= C(I'™; I1)(i=1, ..., m), that is, /<= C^il™', !1). Reasoning as in the proof of Proposition we establish by induction that /<= C(p\l™', Iy)when F <= C^P\U; R). □

Example. Let that the function F : G —> R is defined in a domain G C Rm and belongs to the class C^(G;R); xo=(xq, ..., x^)<= G and jF(xo)=F(xJ, ... jXq1)=0. If xo is not a critical point of F, then at least one of the partial derivatives of F at xo is nonzero. Suppose, for example, that ^r(xo)^ 0.

Then, by Proposition in some neighborhood of Xo the subset of Rm defined by the equation Fix1, ...,xm)=0 can be defined as the graph of a function xm= /(x1, ..., xm_1), defined in a neighborhood of the point (xq, ...,x^l_1)<= Rm_1 that is continuously differentiate in this neighborhood and such that /(xj, ..., x™-1)=x^.

Thus, in a neighborhood of a noncritical point xo of F the equation

F(x1, ..., xm)=0

defines an(m — l)-dimensional surface.

In particular, in the case of R3 the equation

Fix, y, z)=0

defines a two-dimensional surface in a neighborhood of a noncritical point ixo, yo, zo)satisfying the equation, which, when the condition §f(xo? 2 /o? ^o)7^ 0 holds, can be locally written in the form

*=ffay).

As we know, the equation of the plane tangent to the graph of this function at the point(xo, 2 /o?^o)has the form

Z- Z0=x0, y0)(x - x0)+ x0, yo)(y ~ 2 /o)■

But by formula

df(\- Fx(xo, yo, z0)df ^_ Fy(x0, yo, z0)
8x [X0>V0)~ F'z{ x0, 2 /0)zo)' dy(*0'Vo)~ F'z{ x0, y0, z0)'

and therefore the equation of the tangent plane can be rewritten as
F'x{ xo, t /o, zo)(x - x0)+ Fy{ x0, 2 /0, z0)(y - 2 /0)+ Fz(x0, 2 /0, z0)(z - z0)=0,

which is symmetric in the variables x, y, z.

Similarly, in the general case we obtain the equation

m

Y.F'Ax o)(xi-4)=0

i= 1

of the hyperplane in Rm tangent at the point=(xj, ..., x™)to the surface given by the equation F(xx, ..., xm)=0(naturally, under the assumptions that F(xo)=0 and that xo is a noncritical point of F).

It can be observen from these equations that, given the Euclidean structure
on Rm, one can assert that the vector

dF dF

sradF(l°)=(&r

is orthogonal to the r-level surface F(x)=r of the function F at a corresponding point xo <= Rm.

For example, for the function

x^ z^

Fi*, v, *)=j + yw + f,

defined in R3, the r-level is the empty set if r < 0, a single point if r=0, and the ellipsoid

x2y2z2

~2 + To + ~2 — T

az bz cz

if r > 0. If(xo, 2 /o?^o)is a point on this ellipsoid, then by what has been proved, the vector

a t?(^(2x° 2y° 2z°\

grad F(xo, 2 /o, ^o)= -5-)

is orthogonal to this ellipsoid at the point(xo, 2 /o?^o)? and the tangent plane
to it at this point has the equation

5.3 THE IMPLICIT FUNCTION THEOREM

We now turn to the general case of a system of equations
xo(x-xo), 2 /0(2 /-2 /0). Zo(z-Zo)

= 0,

a' b2 c*

which, when we take account of the fact that the point(#o, Vo, zo)lies on the
ellipsoid, can be rewritten as

xox . VoV . z0z

—2~ + ~U2~ + T=r '
az bz cz F\x\...,
xm, y\...yn)=0 which
we shall solve with respect to 2 /1, -- -,2 /n, that is, find a system of functional relations

y1= /1(a:1, ...,xm)

^yn=fn(x\..., xm),

locally equivalent to the system

For the sake of brevity, convenience in writing, and clarity of statement,
Let us agree that x=(x1, ..., xm), y=(y1, ..., yn). We shall write the left hand side of the system as F(x, y), the system of equations as
F(x, y)=0, and the mapping as y= /(x).

If

Notes

x0 = , J /0=(2 /o»---»2 /o)»

a=(a1, ..., am),(3=(/J1, ..., /T),

the notation \x — xo\ < a or \y — yo\ <(3 will mean that \x% — Xq| < a1
(i=1, ..., m)or \y] — y30\ <(3j(j=1, ..., n)respectively.

We next set

dFn

V dy1 dyn J

We remark that the matrix Fy(x, y)is square and hence invertible if and only if its determinant is nonzero. In the case n=1, it reduces to a single element, and in that case the invertibility of Fy(x, y)is equivalent to the condition that that single element is nonzero. As usual, we shall denote the
matrix inverse to Fy(x, y)by [Fy(x, y)] .

We now state the main result of the present section.

Theorem.(Implicit function theorem). If the mapping F : U Rn defined in a neighborhood U of the point(#o, 2 /o)^ Mm+n is such that
1° Fe C^(U', Rn), p> 1,

2° F(xo, yo)=0,

3° Fy(xo, yo)is an invertible matrix,
then there exists an(m + n)-dimensional interval I=I™ x I™ C U, where

I™={ x g Mm| \x - x0\ < a }, = { y e Rn| |y - y0\ < /? },

and a mapping f e(/™; I™)such that

F(x, y)=0^y=f(x)

for any point(x, y)G I™ x Iy and

f'{ x)=-[F^xJix^y'lF^xJix))]

Proof The proof of the theorem will rely on Proposition and the elementary properties of determinants. We shall break it into stages, reasoning by

induction. For n=l, the theorem is the same as Proposition and is

therefore true.

Suppose the theorem is true for dimension n — 1. We shall show that it

is

then valid for dimension n.

a)By hypothesis 3°, the determinant of the matrix is nonzero at

the point(xo, t /o)^ Mm+n and hence in some neighborhood of the point

(#o?2 /o)- Consequently at least one element of the last row of this

matrix is

nonzero. Up to a change in the notation, we can let that the element

dFn .

—— is nonzero.

dyn

Then applying Proposition to the relation

Fn(x1, ...,xm, y1, ..., yn)=0,

we find an interval Im+n=(J™ x /™-1)x J* c U and a function /e

CW(7™ x Iy-^Jy)such that

(Fn(a:\..., xm, y1, ..., yn)= 0 in /m+n)

^(yn=fix1, ..., xm, yx, ...,yra_1),

Substituting the resulting expression yn=f(x, y1, ..., yn_1)for the

variable yn in the first n — 1 equations of(8.90), we obtain n — 1

relations

^(x1, ...,xm, y1, ..., j /n_1):= _

= F1(x1, ...,xm, y1, ...,yn~1, /(x1, ...,xm, y1, ...,yn~1))=0,

<7n 1(x1, ..., xm, y1, ..., yn *):= _

= Fn-1(x1, ..., xm, y1, ..., yn-1J(x1, ..., xm, y1, ..., yn-1))=0 .

It is clear that € C^(I™ x 7™_1;1R)(t=1, ..., n - 1), and

Notes

^(4, -- -,x^ -,yo, ..., yo~1)=0   1),

since /(®J, ..., xg1, yj, ..., y%_1)=y% and Fi(x0, yo)=0(i=1, ..., n).
By definition of the functions $k(k=1, ..., n — 1),

Qpk Qpk Qf

= +   (i, k=l, ..., n-l)

Qyl   Qyl   QyU Qyl

Further setting

*n(x1, ..., *m, y1, ..., yn-1):=

= Fn(x1, ..., xm, y\..., yn-\f(x1, ..., xm, y1, ..., yn-1)),

we find by that <Pn=0 in its domain of definition, and therefore

d$n dFn dFn df

= +   = o(i=1, ..., n - 1).

dyl   Qyl   QyU Qyl

Taking account of relations the properties of determinants
we can now observe that the determinant of the matrix
equals the determinant of the matrixF1 dF1 df dF1 \+ 3 *

dy1 dyn dy1 dyn_1 dyn dyn_1 dy

Since I™ 1 c I™ \ and J™ c J™, substituting Z1, ..., /72 1 from
in place of the corresponding variables in the function

Vn=f(x1, ..., xm, y1, ..., yn~1)

we obtain a relation

yn=fn(x1, ..., xm)

between yn and(x1, ..., xm).

We now show that the system

' y1=f1(x1, ...xm),

< xel? .

, yn= /n(xV.., xm),

which defines a mapping /<= C^(/™; J™), where I™=Iy~x x /*, is equivalent to the system of equations in the neighborhood /m+n=I™ x I™.

In fact, inside /m+n=(J™ x /^_1)x J* we began by replacing the last equation of the original system with the equality yn = f(x, y1, ..., yn_1), which is equivalent to it by virtue From the second system so obtained, we passed to a third system equivalent to it by replacing the variable yn in the first n — 1 equations with /(#, y1, ..., yn_1). We then replaced the first n — 1 equations of the third system inside I™ x Iy~l C /-X Iy~l with relations which are equivalent to them. In that way, we obtained a fourth system, after which we passed to the final system which is equivalent to it inside I™ x J™-1 x I*= /m+n? by replacing the variables y1, ..., yn_1 with their expressions in the last equation yn= /(x1, ..., xm, y1, ..., yn_1)of the fourth system, obtaining as the last equation.

To compLete the proof of the theorem it remains only to verify formula Since the systems are equivalent in the neighborhood I™ x Iy of the point(xo, yo), it follows that

F(x, f(x))=o, iixel™.

In coordinates this means that in the domain I™

Fk(x\..., xm, f\x\..., xm), ..., fn(x\ ...,xm))=0

(k = .(8.104)

Since /€ C(p'(/™; /™)and F 6 C^V\U;R"), wherep > 1, it follows that G <=(p)(/™; Rn)an(j)differentiating the identity we obtain

+ (*-I

J=1

Relations are obviously equivalent to the single matrix equality

K(x, y)+ Fy(x> v)• f'(x)=0 >

in which y= /(x).

Taking account of the invertibility of the matrix F'y{ x, y)in a neighborhood of the point(xq, 2 /0)> we find by this equality that

f'{ x)= ,

and the theorem is compLetely proved.

The solution of this problem gives another proof of the fundamental theorem of this section, perhaps less intuitive and constructive than the one given above, but shorter. Suppose the hypotheses of the implicit function theorem are satisfied, and

Let

^,(dFl dFi \,,

Fy^v)=W, -", dr){ x, v)

be the ith row of the matrix Fy(x, y).

Show that the determinant of the matrix formed from the vectors Fy(xi, yi)is nonzero if all the points(xi, yi)(i=l, ..., n)lie in some sufficiently small neighborhood U=I™ x Iy of(xo, yo). Show that, if for x <= I™ there are points 2 /1, 2 /2 <= Iy such that F(x, yi)=0
and F(x, 2 /2)=0, then for each i G { 1, ..., n)there is a point(x, y{)lying on the closed interval with endpoints(x, 2 /1)and(x, 2 /2)such that Fi(x>y*)(y2 -2 /i)=0(«=i, ..., n). Show that this implies that 2 /1 =2 /2? that is, if the implicit function /: I™ -» Iy exists, it is unique. Show that if the open ball B(yo; r)is contained in J™, then F(xo, y)^ 0 for ||2/ — 2 /o11= r > 0.

The function ||F(xo, 2/)lliri continuous and has a positive minimum value p on the sphere ||2/ — 2 /o||r^=r.

There exists S > 0 such that for ||x — xo||r™ < S we have

||F(a:, j/)||Rn > i /i, if \\y - yolk"=r,

\\F(x, y)\ \ln <, if y=y0 .

For any fixed x such that ||x — xo|| < S the function ||F(x, 2/)||r^ attains a minimum at some interior point y=f(x)of the open ball \\y — 2 /0 Hr™ < r, and since the matrix F^x, /(x)^ is invertible, it follows that F^x, /(x)^=0. This establishes the existence of the implicit function /: B(xo;S)-> B(yo;r).

If Ay=f(x + Ax)- /(x), then

Ay =■ [F^]Ax,

where F'y is the matrix whose rows are the vectors Fy(xi, yi), (i=1, ..., n), (xj, y{)being a point on the closed interval with endpoints(x, y)and(x -f Ax, y -f Ay). The symbol F'x has a similar meaning.

Show that this relation implies that the function y=f(x)is continuous.

Show that

f'(x)=- [.Fy(x, /(a:))] •(x, /(a:))]

"If f(x, y, z)=0, then ff - - ff=-1."

Give a precise meaning to this statement.

Verify that it holds in the example of Clapeyron's ideal gas equation

P'V ♦

———=const

T

and in the general case of a function of three variables.

Write the analogous statement for the relation /(a?1, ...,#m)=0 among m variables. Verify that it is correct.

Show that the roots of the equation

zn + c\zn~1 H b Cn=0

are smooth functions of the coefficients, at least when they are all distinct.

5.4 SOME COROLLARIES OF THE IMPLICIT FUNCTION THEOREM

THE INVERSE FUNCTION THEOREM

Definition. A mapping /: U -» V, where U and V are open subsets of Mm, is a -diffeomorphism or a diffeomorphism of smoothness p (P=0, 1, ...), if

i)fecv)(u -,vy,

/is a bijection;

f-1 eC<*\V;U).

A -diffeomorphism is known a homeomorphism.

As a rule, in this book we shall consider only the smooth case, that is, the case p G N or p=oo.

The basic idea of the following frequently used theorem is that if the differential of a mapping is invertible at a point, then the mapping itself is invertible in some neighborhood of the point.

Theorem.(Inverse function theorem). If a mapping f : G -» Mm of a domain G C Mm is such that

1° /€C, W(G;Rm), p> l,

2° 2 /o=fix0)at x0 e G,

3° fix0)is invertible,

then there exists a neighborhood U(xo)C G of xq and a neighborhood V(yo)of yo such that f : U(xo)-» V(yo)is a -diffeomorphism. Moreover, if x G U(xo)and y=f(x)G V(yo), then

(/1)'(2/)=(fix)) 1

Proof. We rewrite the relation y=f(x)in the form

F(x, y)=f(x)-y=0.

The function F(x, y)=f(x) — y is defined for x G G and y G Mm, that is it is defined in the neighborhood G x Mm of the point(xq, yo)G Mm x Mm.

We wish to solve with respect to x in some neighborhood of (xoiVo)- By hypotheses 1°, 2°, 3° of the theorem the mapping F(x, y)has the property that

F€Cw(Gxr;r), P> i,

F(x0, yo)=0,

K(xo, Vo)=f'(xo)is invertible.

By the implicit function theorem there exist a neighborhood Ix x Iy of (#Oj2 /o)and a mapping g G C(p\ly; Ix)such that

f(x)- y=0 ^ x=g(y)

for any point(x, y)G Ix x Iy and

g'iv)=- [K(x>y)]_1 y)] ■

In the present case

F'x{ x, y)=f'{ x), Fy(x, y)= —E,
where E is the identity matrix; therefore

g\y)=(/'(s))"1 •

If we set V — Iy and U=^(V), relation shows that the mappings /: U —V and g : V -» U are mutually inverse, that is, g= /-1 on V.

Since V=Iy, it follows that V is a neighborhood of yo- This means that under hypotheses 1°, 2°, and 3° the image yo=f(xo)of xo G G, which is an interior point of G, is an interior point of the image /(G)of G. By formula the matrix g'{ yo)is invertible. Therefore the mapping g : V U has properties 1°, 2°, and 3° relative to the domain V and the point yo G V.

Hence by what has already been proved xo=g(yo)is an interior point of U=g(V).

Since by hypotheses 1°, 2°, and 3° obviously hold at any point y G V, any point x=g(y)is an interior point of U. Thus U is an open(and obviously even connected)neighborhood of xo G Mm.

We have now verified that the mapping /: U -» V satisfies all the conditions. The inverse function theorem is very often used in converting from one coordinate system to another. The simplest version of such a change of coordinates was studied in analytic geometry and linear algebra and has the form

$$\begin{pmatrix} y^1 \\ \cdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_1^1 & \cdots & a_m^1 \\ \cdots & \cdots & \cdots \\ a_1^m & \cdots & a_m^m \end{pmatrix} \begin{pmatrix} x^1 \\ \cdots \\ x^m \end{pmatrix}$$

or, in compact notation, yi=a\xl.

This linear transformation A : R™ —R™
has an inverse A~x : R™ -» R™ defined on the entire space R™ if and only if the matrix(a^)is invertible, that is, det(a^)^ 0.

The inverse function theorem is a local version of this proposition, based on the fact that in a neighborhood of a point a smooth mapping behaves approximately like its differential at the point.

Example. Polar coordinates. The mapping /: R+ -» R2 of the half-plane R%={(p, (p)G R2\ p > 0 } onto the plane R2 defined by the formula

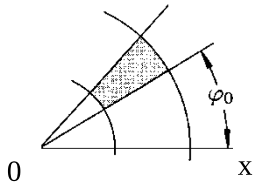
x=p cos p,
y=psmp,

The Jacobian of this mapping, as can be easily computed, is p, that is, it is nonzero in a neighborhood of any point(p, p), where p > 0. Therefore formulas are locally invertible and hence locally the numbers p and p can be taken as new coordinates of the point previously determined by the Cartesian coordinates x and y.

The coordinates(p, p)are a well known system of curvilinear coordinates on the plane - polar coordinates. Their geometric interpretation is shown

in Figure. We note that by the periodicity of the functions cos<^ and sin<^ the mapping is only locally a diffeomorphism when p > 0; it is not bijective on the entire plane. That is the reason that the change from Cartesian to polar coordinates always involves a choice of a branch of the argument p
(that is, an indication of its range of variation).

Y 7r 2



Polar coordinates(p, ip)in three-dimensional space M3 are known spherical coordinates. They are connected with Cartesian coordinates by the formulas

z=pcosip,

y=psmipsirup,

x=p sin ^ cos(p .

The geometric meaning of the parameters p.

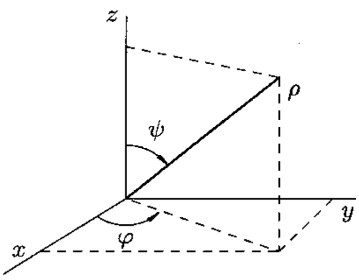


Fig. 8.6.

The Jacobian of the mapping is p2sinip, and so by Theorem
the mapping is invertible in a neighborhood of each point(p, ip, p)at which
p > 0 and sin ip ^ 0.

The sets where p=const, (p=const, or ip=const in(x, y, z)-space obviously correspond to a spherical surface(a sphere of radius p), a

Notes

halfplane passing through the z-axis, and the surface of a cone whose axis is the z-axis respectively.

Thus in passing from coordinates(x, y, z)to coordinates(p, i^,(p), for example, the spherical surface and the conical surface are flattened; they correspond to pieces of the planes p=const and ip=const respectively. We observed a similar phenomenon in the two dimensional case, where an arc of a circle in the(x, y)-plane corresponded to a closed interval on the line in the plane with coordinates(p,(p). Please note that this is a local straightening.

In the m-dimensional case polar coordinates are introduced by the relations

x1=pcos<<=i,
x2=p sin ipi cos ip2 ?

xm"i=psin^i sin^?2 • • • sin(^m_2 cos(^m_i,
x171=p sin <pi sin(p2 • • • sin(pm-2 sin(pm_ i .

and by Theorem it is also locally invertible everywhere where this Jacobian is nonzero.

Example. The general idea of local rectification of curves. New coordinates are usually introduced for the purpose of simplifying the analytic expression for the objects that occur in a problem and making them easier to visualize in the new notation.

Suppose for example, a curve in the plane R2 is defined by the equation

F(x, y)=0 .

Let that F is a smooth function, that the point(xo, yo)lies on the curve, that is, F(xo, yo) — 0, and that this point is not a critical point of F. For example, suppose Fy(x, y)^ 0.

Let us try to choose coordinates <= & 77, so that in these coordinates a closed interval of a coordinate line, for example, the line rj=0, corresponds to an arc of this curve.

We set

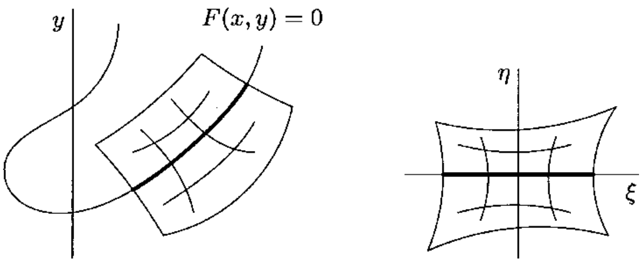
<==X - x0, T]=F(x, y).

The Jacobi matrix

$$\begin{pmatrix} 1 & 0 \\ F'_x & F'_y \end{pmatrix}(x, y)$$

of this transformation has as its determinant the number F'y(x, y), which by assumption is nonzero at(xo, yo)- Then by Theorem this mapping is a diffeomorphism of a neighborhood of(xo, yo)onto a neighborhood of the point(<=, ??)=(0, 0). Hence, inside this neighborhood, the numbers <= & 77 can be taken as new coordinates of points lying in a neighborhood of(xq, t /o)-

In the new coordinates, the curve obviously has the equation 77=0, and in this sense we have indeed achieved a local rectification.



Local Reduction of a Smooth Mapping to Canonical Form

In this subsection we shall consider only one question of this type. To be specific, we shall exhibit a canonical form to which one can locally reduce any smooth mapping of constant rank by means of a suitable choice of coordinates.

We recall that the rank of a mapping /: U -» W1 of a domain U C Mm at a point x G U is the rank of the linear transformation tangent to it at the point, that is, the rank of the matrix f'(x). The rank of a mapping at a point is usually denoted rank /(x).

Theorem.(The rank theorem). Let f : U -» Mn be a mapping defined in a neighborhood U C Mm of a point xq G Mm. If f G C^p\U;Rn), p > 1, and the mapping f has the same rank k at every point x G U, then there

Notes

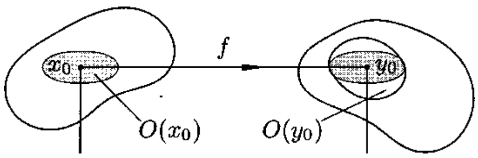
exist neighborhoods 0(xo)of xq and O(yo)of yo=f(xo)and diffeomorphisms

u=(p(x), v=tp(y)of those neighborhoods, of class C^p\ such that the mapping v=o f o has the coordinate representation

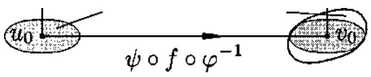
(u1, ...,uk, ...,um) — u i—y v —(v1, ..., vn)=(u1, ...,^, 0, ...,0)

in the neighborhood 0(uf)=tp(0(xo)) of uq=^>{ xf).

In other words, the theorem asserts(observe Fig. 8.8)that one can choose coordinates(u1, ..., um)in place of(x1, ..., xm)and(v1, ..., vn)in place of(y1, ...,yn)in such a way that locally the mapping has the form in the new coordinates, that is, the canonical form for a linear transformation of rank k.



O(uo) 0(v 0)_



5.5 DIFFERENTIAL CALCULUS IN SEVERAL VARIABLES

We write the coordinate representation

y6=f1(x1, ..., xm),

yk=fk(x\..., xm),

yk+1= /fe+i(xi;,

yn=fn(x\..., xm)

of the mapping /: U -» M™, which is defined in a neighborhood of the point
xq G M™. In order to avoid relabeling the coordinates and the neighborhood we shall let that at every point x € U, the principal minor of order k the upper left corner of the matrix f'(x)is nonzero.

Let us consider the mapping defined in a neighborhood U of xq by the equalities

u1=(f1(x1, ..., xm)=f1(x1, ..., xm)

<pk(x\..., xm)=fk(x1, ..., xm),
y>k+1{ x1, ..., xm)=xk+1,

um=y>m(x1, ..., xm)=xm .
The Jacobi matrix of this mapping has the form

df1 df1 \

fdfl

dx1

dfk dfk

dx1 dxk

dxk+l

dx71

and by assumption its determinant is nonzero in U.

By the inverse function theorem, the mapping u=<p(x)is a diffeomorphism of smoothness p of some neighborhood 0(xo)C U of xq onto a neighborhood O(uo)=(p(0(xo)) of uo=<p(xo).

y*=fnQ(p-l^ul^ ^um>)= m,um).

Since the mapping(p_1 : O(i^o)-> 0(xo)has maximal rank m at each point u G O(uo), and the mapping /: O(xo) has rank k at every point x G O(xo)? h follows, as is known from linear algebra, that the matrix g'(u)=f'{(p~l(u)){ p~l)(u)has rank k at every point u G 0{ uq).

Notes

Direct computation of the Jacobi matrix of the mapping(8.122)yields

/

duk+1

dum /

dgj

Hence at each point u G 0(u$)we obtain -r—="(u)=0 for i=k +1, ..., m;

ou% _

j -h 1, ..., n. Assuming that the neighborhood O(uo)is convex(which can be achieved by shrinking 0{ uq)to a ball with center at uq, for example)we can conclude from this that the functions gi, j=k + 1, ..., n, really are independent of the variables ukJrl, ..., um.

After this decisive observation, we can rewrite the mapping as

yk=uk,
yk+1=gk+-i-{ ra1, ..., uk),

yn=gn{ u1, ..., uk).

At this point we can exhibit the mapping ip. We set

v1=y1 =■■ vHy)>

= yk =■ ipk(y),
vk+1=yk+1 -gk+1(y1, -- -,yk)=■ ipk+1(y),

vn=yn-gn(y1, ..., yk)=: ipn(y)■

It is clear from the construction of the functions gi(j=k +1, ..., n)that the mapping is defined in a neighborhood of yo and belongs to class C^ in that neighborhood.

The Jacobi matrix of the mapping has the form

dy1

dyk

Its determinant equals 1, and so by Theorem 1 the mapping i /i is a diffeomorphism of smoothness p of some neighborhood O(yo)of yo G R™ onto a neighborhood 0(vq)=i)(0{ yo)) of vo G R™.

Comparing relations, we observe that in a neighborhood O(uo)C O(uo)ofuo so small that g(0(uo)) C O(yo)> the mapping t /jo /o^-1 :

O(uo)-> M.y is a mapping of smoothness p from this neighborhood onto some neighborhood O(vo)C 0(vq)of vq G and that it has the canonical form

vn=0.

Setting p~l(O(uo))=0(xo)and o))=O(yo)> we obtain the

neighborhoods of xq and yo whose existence is asserted in the theorem. The proof is now compLete. Theorem is obviously a local version of the corresponding theorem from linear algebra.

In connection with the proof just given of Theorem we make the following remarks, which will be useful in what follows.

Remark. If the rank of the mapping /: U -» Mn is n at every point of the original neighborhood U C Mm, then the point yo=f(xo), where Xo G U, is an interior point of f(U), that is, f(U)contains a neighborhood of this point.

Proof. Indeed, from what was just proved, the mapping t /jo /o^-1 :
O(uo)-»
0(vo)has the form

(u1, ...,un, ..., um)=u v=(v1, ..., vn)=(u1, ..., /un),

in this case, and so the image of a neighborhood of uo=(f(xo)contains some neighborhood of vo=° /° y>~l(uo)-

But the mappings(p : 0(xo)-> O(uo)and ip : O(yo)-> 0(vo)are diffeomorphisms, and therefore they map interior points to interior points. Writing the original mapping /as /=V7"1 °(V7 ° /°(P~1)° we conclude that yo=f(xo)is an interior point of the image of a neighborhood of xo

Remark. If the rank of the mapping f : U -» W1 is k at every point of a neighborhood U and k < n, then in some neighborhood of xo G U C Mm the following n — k relations hold:

f(x\..., xm)=gi(f1(x\...>xm)>..., fk

(i=k + l, ..., n)

These relations are written under the assumption we have made that the principal minor of order k of the matrix f'(xo)is nonzero, that is, the rank k is realized on the set of functions Z1, ..., fk. Otherwise one can relabel the functions f1, ..., fn and again have this situation.

5.6 DIFFERENTIAL CALCULUS IN SEVERAL VARIABLES FUNCTIONAL DEPENDENCE

Definition. A system of continuous functions f%(x)= ..., xm)

(i — l, ..., n)is functionally independent in a neighborhood of a point xq=(xj, .. •, Xq1)if for any continuous function F(y)=F(y1, ..., yn)defined in a neighborhood of y0=(j /J, ..., y%)=(f1(x0), fn(x 0))=f(x0),
the relation

F(f1(x\..., xm), ..., fn(x\..., xm))=0

is possible at all points of a neighborhood of xq only when F(y1, ..., yn)=0
in a neighborhood of y®.

The linear independence studied in algebra is independence with respect to linear relations

F(y1, ..., yn)=X1y1 + --- + Xnyn .

If a system is not functionally independent, it is said to be functionally dependent When vectors are linearly dependent, one of them obviously is a linear combination of the others. A similar situation holds in the relation of functional dependence of a system of smooth functions.

Proposition. If a system fl(xl, ..., xm)(i=1, ..., n)of smooth

functions defined on a neighborhood U(xo)of the point xq g Mm is such that the rank of the matrix

(df1 Of1 \

dx1 dxm

dfn dfn

\ dx1 dxm)

is equal to the same number k at every point x g U, then

when k=n, the system is functionally independent in a neighborhood ofx0;

when k < n, there exist a neighborhood of x$ and k functions of the system, say f1, ..., fk such that the other n — k functions can be represented
as

f(x\. .., xm)=gi(f1(x1, .. ., xm), fk(x\.. ., xm))

in this neighborhood, where gt(y1, ..., yk), (i=k + 1, ..., n)are smooth functions defined in a neighborhood of yo=(f1(xo), ..., fn(xo)) and depending only on k coordinates of the variable point y=(y1, ..., yn).

Proof In fact, if k=n, then by Remark 1 after the rank theorem, the image of a neighborhood of the point xq under the mapping

can hold in a neighborhood of xq only if

F(y\..., yn)= 0

in a neighborhood of yo> This proves assertion.

If k < n and the rank k of the mapping is realized on the functions Z1, ..., /*^ then by Remark after the rank theorem, there exists a neighborhood of yo=f(xo)and n — k functions gl(y)=gt(y1, ..., yk) (i=k -b 1, ...,n), defined on that neighborhood, having the same order of smoothness as the functions of the original system, and such that relations hold in some neighborhood of xq.

We have now shown that if k < n there exist n — k special functions Fl(y) — y%— 9l(yl, • • •> yk)(® — k + 1, ..., n)that establish the relations between the functions of the system f1, ..., fk, ..., fn in a neighborhood of the point xq.

5.7 LOCAL RESOLUTION OF A DIFFEOMORPHISM INTO A COMPOSITION OF ELEMENTARY ONES

In this subsection we shall show how, using the inverse function theorem, one can represent a diffeomorphic mapping locally as a composition of diffeomorphisms, each of which changes only one coordinate.

Definition. A diffeomorphism g : U —» Mm of an open set U C Mm will be known elementary if its coordinate representation is

yi=x*, i e , i¥=j,

yj=gj(,

that is, under the diffeomorphism g : U —» Rm only one coordinate of the point being mapped is changed.

Proposition. If f : G —> Mm is a diffeomorphism of an open set G C Mm, then for any point xq G G there is a neighborhood of the point in which the representation f=g\ o • • • o gn holds, where g \,...,gn are elementary diffeomorphisms.

Proof We shall verify this by induction.

If the original mapping /is itself elementary, the proposition holds trivially for it.

Let that the proposition holds for diffeomorphisms that alter at most (k — 1)coordinates, where k — l<n. Now consider a diffeomorphism /: G -»
Mm that alters k coordinates:

1=p(x1, ..., xm)

ym=x71

We have letd that it is the first k coordinates that are changed, which can be achieved by linear changes of variable. Hence this assumption causes no loss in generality.

Since /is a diffeomorphism, its Jacobi matrix f'(x)is nondegenerate at each point, for

(/_1)'(/(^))=[/'(*)]'

Let us fix xq G G and compute the determinant of f\xo):

minor of order A: — 1 is nonzero. Now consider the auxiliary mapping g : G —»
Mm defined by the equalities

u1= /1(a;1, ..., a;m),

, k-1

uk=xk

u"v=x"

Since the Jacobian
df1 df1 ; df1

&e

the mapping g : G -> Mm is nonzero at xq G G, the mapping g is a diffeomorphism in some neighborhood of xq.

Then, in some neighborhood of u$=g{ xo)the mapping inverse to g, x=g~l{ u), is defined, making it possible to introduce new coordinates (u1, ..., um)in a neighborhood of x$.

Let h=f o g~l. In other words, the mapping y=h(u)is the mapping y=f(x)written in ^-coordinates. The mapping /i, being the composition of diffeomorphisms, is a diffeomorphism of some neighborhood of uq. Its coordinate expression obviously has the form

V =f °9(u)=u,

yk 1=fk 1og l(u)=uk x,

yk=fk°,

yk+l=Uk+l,

y171=Um,

that is, h is an elementary diffeomorphism.

But /=h o g, and by the induction hypothesis the mapping g defined can be resolved into a composition of elementary diffeomorphisms. Thus, the diffeomorphism /, which alters k coordinates, can also be resolved into a composition of elementary diffeomorphisms in a neighborhood of x$, which compLetes the induction.

5.8 MORSE'S THEOREM

This same circle of ideas contains an intrinsically beautiful Theorem of Morseon the local reduction of smooth real-valued functions to canonical form in a neighborhood of a nondegenerate critical point. This Theorem is also important in applications.

Definition. Let xq be a critical point of the function /E C^2\U\ R) defined in a neighborhood U of this point.

The critical point xq is a nondegenerate critical point of f if the Hessian

d2 f

of the function at that point(that is, the matrix _ .0 .(x0)formed from

UX%Ox3

the second-order partial derivatives)has a nonzero determinant.

If xq is a critical point of the function, that is, ff(xo)=0, then by Taylor's formula

f(x)-f(x0)=1 dx*dxi(a?o)(gi-4)(^-^)+ o(lk-»o||2)•

' id

Morse's Theorem asserts that one can make a local change of coordinates x=g(y)such that the function will have the form

(/o g){ y)- /(so)=~(y1)2 (yk)2 +(yk+1)2 + • • • +(ym)2

when expressed in y-coordinates.

If the remainder term o(\\x — xo\\2)were not present on the right-hand side of Eq. that is, the difference f(x) — f(xo)were a simple quadratic form, then, as is known from algebra, it could be brought into the indicated canonical form by a linear transformation. Thus the assertion we are about to prove is a local version of the theorem on reduction of a quadratic form to canonical form. The proof will use the idea of the proof of this algebraic theorem. We shall also rely on the inverse function theorem and the following proposition.

Hadamard's Theorem. Let /:[/—> R be a function of class C^P\U;M), p > 1, defined in a convex neighborhood U of the point 0=(0, ..., 0)E Rm and such that /(0)=0. Then there exist functions gi G R)

(i=1, ..., m)such that the equality

m

fix1, ..., xm)=x'giix1, ..., xm)

2=1

holds in U, and gi{ 0)=J^r(0).

Proof. Equality is essentially another useful expression for Taylor's formula with the integral form of the remainder term. It follows from the equalities

/(*', . .., *»)= /a «=<=></gy*',

0 i=1 0

if we set

l

<=i(x\..., xm)=J ^{ txl, ... ^tx™)&t(i=l, ..., ra).

o

The fact that <fa(0)=J^r(0)(i=1, ..., m)is obvious, and it is also not difficult to verify that gi G C^p~^(U\ R). However, we shall not undertake the verification just now, since we shall later give a general rule for differentiating an integral depending on a parameter, from which the property we need for the functions gi will follow immediately.

Thus, up to this verification, Hadamard's formula is proved.

Morse's Theorem. If f : G —> R is a function of class(G; R)defined on an open set G C Rm and Xo G G is a nondegenerate critical point of that function, then there exists a diffeomorphism g : V —> U of some neighborhood of the origin 0 in Rm onto a neighborhood U of xo such that if ° 9)iv)=fixo)- [(y1)2 + • • • +(yfc)2] + [ivk+1)2 + • • • +(ym)2] for all y G V.

Proof By linear changes of variable we can reduce the problem to the case when xo=0 and f(xo)=0, and from now on we shall let that these conditions hold.

Since xo=0 is a critical point of /, we have ^(0)=0 in formula (i=1, ..., m). Then, also by Hadamard's Theorem,

m

Qiix1, . . ■,Xm)=

3=1

where h^ are smooth functions in a neighborhood of 0 and consequently m

f{ xl, ...,xm)=xtxjhij(x1, ...,xm).

i, j=1

By making the substitution hij=\{ hij-\-hji)if necessary, we can let that=hji. We remark also that, by the uniqueness of the Taylor expansion, the continuity of the functions h^ implies that hij(0)= .(0)and

ux uxJ

hence the matrix(/^(0)) is nondegenerate.

The function /has now been written in a manner that resembles a quadratic form, and we wish, so to speak, to reduce it to diagonal form. As in the classical case, we proceed by induction.

Let that there exist coordinates ix1, ...,xxm in a neighborhood U\ of 0 G Mm, that is, a diffeomorphism x=(p(u), such that

m

if ° vOM — i^1)2 ± • • • =t K"1)2 -f ^2 tfv? Hijfa1, ...,ixm)

i, j=r

in the coordinates ix1, ..., um, where r > 1 and H\j=Hji.

We observe that relation holds for r=1, as one can observe from, where=hij.

m

By the hypothesis of the Theorem the quadratic form J2 xlx^hij(0)is

i, 3=1

nondegenerate, that is, det(fcij(O)) 7^ 0. The change of variable x=(p(u)is carried out by a diffeomorphism, so that det^'(O)7^ 0. But then the matrix

m

Notes

of the quadratic form ±(u1)2 ± • • • ±(ur~1)2 + J2 utu^Hij(0)obtained from

i, j=r

the matrix { hij{ 0)) through right-multiplication by the matrix(p'(0)and left-
multiplication by the transpose of y /(0)is also nondegenerate. Consequently,
at least one of the numbers Hij(0)(hj=r, ...,ra)is nonzero. By a linear

m

change of variable we can bring the form ^2 u%vPHij{ Q)to diagonal form,

i, j=r

and so we can let that Hrr(0)7^ 0 in Eq.(8.133). By the continuity of the functions Hij(u)the inequality Hrr(u)7^ 0 will also hold in some

neighborhood of u=0.

Let us set ipiu1, ...,um)=yJ\Hrr{ u)\ . Then the function ip belongs to the class in some neighborhood U2 C U\ of u=0. We now change to coordinates(i;1, ..., vm)by the formulas

<8-134>

The Jacobian of the transformation at u=0 is obviously equal
to V>(0)> that is, it is nonzero. Then by the inverse function theorem we can assert that in some neighborhood Us C U2 of u=0 the mapping v=ip(u)defined is a diffeomorphism of class C^(Us; Mm)and therefore the variables(i;1, ..., vm)can indeed serve as coordinates of points in Us-

We now separate off in all terms

m

ururHrr(v }, ...,um)+ 2 ^ urv?Hrj{ v } -,... iUm),

j=r+1

containing ur. In the expression for the sum of these terms we have used the fact that Hij=Hji.

Comparing we observe that we can rewrite in the form

±VTVr -(<= u'Hiriu1, . . ., Um))2 •

rr i>r

The ambiguous sign ± appears in front of vrvr because Hrr=±(V02> the positive sign being taken if Hrr > 0 and the negative sign if Hrr < 0. Thus, after the substitution v=ip(u), the expression becomes the equality

r

(/o   [±(w1)2] + v'v7Hijiv1, ...,vm),

2=1   i, j>r

where Hij are new smooth functions that are symmetric with respect to the indices i and j. The mapping(poif;"1 is a diffeomorphism. Thus the induction from r — 1 to r is now compLete, and Morse's Theorem is proved.

Check your Progress-1

Discuss Transition To The Case Of A Relation

Discuss Morse’s Theorem

5.9 LET US SUM UP

In this unit we have discussed the definition and example of Transition To The Case Of A Relation, The Implicit Function Theorem, Some Corollaries Of The Implicit Function Theorem, The Inverse Function Theorem, Differential Calculus In Several Variables, Differential Calculus In Several Variables Functional Dependence, Local Resolution Of A Diffeomorphism Into A Composition Of Elementary Ones, Morse's Theorem

5.10 KEYWORDS

1. Transition To The Case Of A Relation: The following proposition is a a simple generalization of Proposition to the case of a relation F(xx, ..., xm, y)=0

2. Some Corollaries Of The Implicit Function Theorem Definition. A mapping /: U -» V, where U and V are open subsets of Mm, is a -diffeomorphism or a diffeomorphism of smoothness p (P=0, 1, ...)

3.The Inverse Function Theorem:We write the coordinate representation y8=f1(x1, ..., xm)

4. Differential Calculus In Several Variables Functional Dependence Definition. A system of continuous functions f%(x)= ..., xm)

5. Morse's Theorem This same circle of ideas contains an intrinsically beautiful Theorem of Morseon the local reduction of smooth real-valued functions to canonical form in a neighborhood of a nondegenerate critical point

5.11 QUESTIONS FOR REVIEW

Explain Transition To The Case Of A Relation

Explain Morse's Theorem

5.12 ANSWERS TO CHECK YOUR PROGRESS

Transition To The Case Of A Relation

(answer for Check your Progress-1 Q)

Morse's Theorem

(answer for Check your Progress-1 Q)

5.13 REFERENCES

- Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables

UNIT - 6: TAYLOR'S THEOREM, MAXIMA AND MINIMA

STRUCTURE

6.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about Taylor's Theorem, Maxima And Minima

Learn, Understand about Application To Maxima And Minima

Learn, Understand about Surfaces In Rn And The Theory Of Extrema With Constraint

Learn, Understand about The Tangent Space

Learn, Understand about Differential Calculus In Several Variables

Learn, Understand about Extrema With Constraint

Learn, Understand about Surfaces In Rn And Constrained Extrema

Learn, Understand about Some Geometric Images Connected With Functions Of Several Variables

6.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Taylor's Theorem, Maxima And Minima, Application To Maxima And Minima, Surfaces In Rn And The Theory Of Extrema With Constraint, The Tangent Space, Differential Calculus In Several Variables, Extrema With Constraint, Surfaces In Rn And Constrained Extrema, Some Geometric Images Connected With Functions Of Several Variables

6.2 TAYLOR'S THEOREM, MAXIMA & MINIMA

TAYLOR'S THEOREM

This is somewhat complicated, and the longest proof in either content here to carry the expansion out one more term than thus adding a third derivative and discussing the resulting expression. It is

$$f(x)=f(xo)+Df(xo)(x - Xo)+2D^2f(xo)(x - xo, x - xo)$$

$$+3D^3f(c)(x - xo, x - xo, x - Xo)$$

Notes

The last term with $D^3f(c)$, is known the "remainder term". Here c is a point on the line segment between xo and x. From the last term that $D^3f(c)$: R^n X R^n X R^n ! R. There are 3^n

third derivatives, ^—(c). Thus, they won't fit into a square matrix. We will denote these derivatives by $f_{abc}(c)$where a, b, and c are each one of x.or y. The third derivative term when n= 2 , and "x" is(^{X} |, turns out to be

$3(f_{xxx}(c)(x - xo)^3 + 3f_{xxy}(c)(x - xo)^2(y - yo)$

$+ 3f_{xyy}(c)(x - Xo)(y - yo)^2 + f_{,,,}(c)(y - yo)^3)$.

'

Can you observe what the third order derivative would be when n= 3 ?

What about the fourth derivative term for n= 2 and n= 3 ?

MAXIMA AND MINIMA

Positive definite quadratic forms.

A quadratic form is a function Q : R^n ! R of the form

$$Q(u)=u^TAu$$

for some symmetric n X n matrix A. The relevance of this to Taylor's Theorem is observen by looking at equations.

Definition A symmetric matrix A is known "positive definite" if Q(u)> 0 for every u=0 in R^n.

There are two particularly useful criteria for determining of A is positive definite. These are from linear algebra, and won't be proved here.

Theorem A symmetric n X n matrix A is positive definite if either of the following conditions holds:

All eigenvalues of A are positive

All n "upper left" sub determinants of A are positive.

An upper left sub determinant is one formed by deleting between zero and n of the last rows and columns of A/This will be illustrated in class.

If A is a symmetric matrix and —A is positive definite, then A is known negative definite.

6.3 APPLICATION TO MAXIMA AND MINIMA

The equation above allows us to determine criteria guaranteeing that a point x_o is a local maximum or local minimum for the function f. To apply it, we must let that f 2 C^2. There cannot be a local maximum at x_o unless Df(xo)= 0 , for otherwise there is a nonzero directional derivative in some direction e, which means that

d~r, f(xo + te)|t=o=0; dt

and so there are larger values of f either for t positive or t negative, and |t| small.

Definition x_0 is known a "critical point" of f if $Df(x_0)=0$.

We then repeat :

$f(x)= f(xo)+ Df(xo)(x - xo) + D^2f(c)(x - xo, x - xo).$

Assuming that $Df(x_o)= 0$, we get

$$f(x)= f(xo)+ (x - xo)^T D^2f(c)(x - xo).$$

Theorem If x_o is a critical point of f and the matrix corresponding to

$D^2f(x_o)$

is positive definite, then x_o is a local minimum for f. If $D^2f(x_o)$is negative definite, then x_o is a local maximum.

Proof. Suppose that x_o is a critical point of f and $A := D^2f(x_o)$is positive definite. Then $e^TAe > 0$ for every unit vector e 2 R^n. If u is a nonzero vector in R^n, then e= is a unit vector. It follows that A is positive definite if and only if $e^TAe > 0$ for every unit vector e. But the set of all unit vectors in A is a compact set. Hence,

$$n = \min e^TAe > 0.$$

I|e|=1|

Because $D^2f(x)$is continuous, it follows that there is a 5 such that if $\|c — x\| < 5$, then $e^TD^2f(c)e > \wedge > 0$ for every unit vector e. Hence D^2f is also positive definite.(i.e. the symmetric matrix corresponding to $D^2f(c)$is positive definite.)

If $\|x — x_o\| <5$ then $\|c — x_o\| < 5$, because c is on the line segment between x and x_o. Equation then implies that if $0 < \|x — x_o\| < 5$, then $f(x)> f(x_o)$. Hence x_o is a local minimum for f. ■

6.4 SURFACES IN RN AND THE THEORY OF EXTREMA WITH CONSTRAINT

To acquire an informal understanding of the theory of extrema with con-straint, which is important in applications, it is useful to have some elemen-
tary information on surfaces(manifolds)in Rn.

Fc-DIMENSIONAL SURFACES IN Rn

Generalizing the concept of a law of motion of a point mass x=x(<=), we have
previously introduced the concept of a path in Rn as a continuous

mapping
r : I —> Rn of an interval I C R. The degree of smoothness of the path was
defined as the degree of smoothness of this mapping. The support f(/)C Rn
of a path can be a rather peculiar set in Rn, which it would be a great stretch
to call a curve in some instances. For example, the support of a path might
be a single point.

Similarly, a continuous or smooth mapping /: Ik —> Rn of a ^-dimensional
interval Ik cRfc, known a singular k-cell in Rn, can have as its image f(Ik)
not at all what one would like to call a ^-dimensional surface in Rn. For
example, it might again be simply a point.

In order for a smooth mapping /: G —> Rn of a domain G C Rfc to
define a ^-dimensional geometric figure in Rn whose points are described by
k independent parameters(t1, ...,tk)G G, it suffices, as we know from the
preceding section, to require that the rank of the mapping /: G —> Rn be k at
each point t G G(naturally, k < n). In that case the mapping /: G —> f(G)
is locally one-to-one(that is, in a neighborhood of each point t G G).

Indeed, suppose rank /(to)=k and this rank is realized, for example, on
the first k of the n functions

' X1 = ,

<

= fn{ t\..., tk)

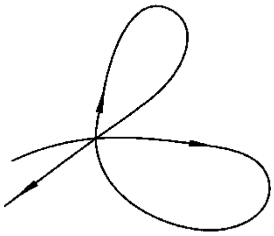
that define the coordinate expressions for the mapping /: G —> Rn.
Then, by the inverse function theorem the variables t1, ..., tk can be ex-

pressed in terms of x1, ... in some neighborhood U(to)of to- It follows that the set f(U(to)) can be written as

=<pk+1(x1, ..., xk),...,xn=ipn{ x\..., xk)

(that is, it projects in a one-to-one manner onto the coordinate plane of x1, ...,xfc), and therefore the mapping /: U(to) —> f(U(to)) is indeed one-to-one.

However, even the simple example of a smooth one-dimensional path makes it clear that the local injectivity of the mapping /: G —> Rn



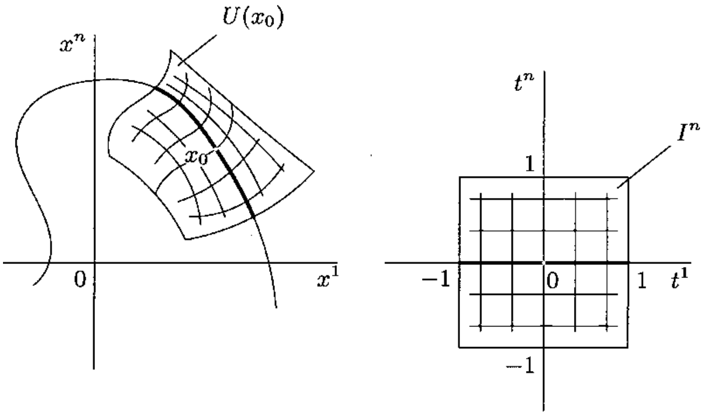
from the parameter domain G into Rn is by no means necessarily a global injectivity. The trajectory can have multiple self-intersections, so that if we
wish to define a smooth ^-dimensional surface in Rn and picture it as a set
that has the structure of a slightly deformed piece of a ^-dimensional plane
(a ^-dimensional subspace of Rn)near each of its points, it is not enough merely to map a canonical piece G C Rfc of a ^-dimensional plane in a reg-
ular manner into Rn. It is also necessary to be sure that it happens to be globally imbedded in this space.

Definition . We shall call a set S C Rn a k-dimensional smooth surface in Rn(or a k-dimensional submanifold of Rn)if for every point xo G S there exist a neighborhood U(xo)in Rn and a diffeomorphism(p : U(xo) —> In of
this neighborhood onto the standard n-dimensional cube In={ t G Rn| \tl\ <

1, i=1, ..., n } of the space Rn under which the image of the set SnU(xo)is the portion of the ^-dimensional plane in Rn defined by the relations

tk+1 =

0, ..., tn=0 lying



We shall measure the degree of smoothness of the surface S by the degree

of smoothness of the diffeomorphism(p.

If we regard the variables t1, ..., tn as new coordinates in a neighborhood of U(xo), Definition 1 can be rewritten briefly as follows: the set S C Rn is

a ^-dimensional surface(^-dimensional submanifold)in Rn if for every point

xo G S there is a neighborhood U(xo)and coordinates in U(xo)

such that in these coordinates the set S D U(xo)is defined by the relations

tk+1=• • •=tn=0 .

The role of the standard n-dimensional cube in Definition 1 is rather artificial and approximately the same as the role of the standard size and shape of a page in a geographical atlas. The canonical location of the interval

in the coordinate system t1, ..., tn is also a matter of convention and nothing more, since any cube in Rn can always be transformed into the standard n-dimensional cube by an additional linear diffeomorphism.

We shall often use this remark when abbreviating the verification that a set S C Rn is a surface in Rn.

Let us consider some examples.

Example. The space Rn itself is an n-dimensional surface of class C^°°\ As the mapping(p : Rn —> In here, one can take, for example, the mapping

2

C=~ arctan x%(i=1, ..., n)

7T

Example . The mapping constructed in Example 1 also establishes that the subspace of the vector space Rn defined by the conditions xfc+1=• • •=xn = 0 is a ^-dimensional surface in Rn(or a ^-dimensional submanifold of Rn).

Example . The set in Rn defined by the system of relations

' ajx1 H b a\xh + <4+1xfc+1 Hh a^x71=0,

k V + • • • + al~kxk + al~kxk+l + • • • + a<=-'kxn=0,

provided this system has rank n — &, is a ^-dimensional submanifold of Rn.
Indeed, suppose for example that the determinant is nonzero. Then the linear transformation

tn=a?" V + • • • + aZ~kxn,

is obviously nondegenerate. In the coordinates t1, ..., tn the set is defined by the conditions tk+1=• • •=tn=0, already considered in Example 2.

Example The graph of a smooth function xn= /(x1, ..., xn_1)defined in a domain G C Rn-1 is a smooth(n — l)-dimensional surface in Rn.

Indeed, setting

tl=xl(i=1, ..., n — 1),

$t^n = x^n - f(x^1, \ldots, x^{n-1})$,

we obtain a coordinate system in which the graph of the function has the equation tn=0.

Example . The circle x2 + y2=1 in R2 is a one-dimensional submanifold of
R2, as is established by the locally invertible conversion to polar coordinates
(p, <p)studied in the preceding section. In these coordinates the circle has equation p=1.

Example . This example is a generalization of Example and at the same time, as can be observen from Definition, gives a general form for the coordinate
expression of submanifolds of Rn.

Let Ft(x11..., xn)(i=1, ..., n — k)be a system of smooth functions of rank n — k. We shall show that the relations

r F1(x1, ...,xfc, xfc+1, ...,xn)=0,

(8.137)

k Fn~k{ xl, ..., xk, xk+1, ..., xn)=0

define a ^-dimensional submanifold S in Rn.
Suppose the condition

dF1 dF1

dxkJtl dxn

(x0)^ 0

Notes

Qjpn-k Qpn-k

dxkJtl dxn

holds at a point xo G S. Then by the inverse function theorem the transfor-
mation

(ti=xi (i=1, ..., k),

(tl=Fl~k(x1, ...,xn)(i=k + 1, ..., n)

is a diffeomorphism of a neighborhood of this point.

In the new coordinates t1, ..., tn the original system will have the form tk+1=- —=tn=0; thus, S is a ^-dimensional smooth surface in Rn.

Example. The set E of points of the plane R2 satisfying the equation x2 —
y2=0 consists of two lines that intersect at the origin. This set is not a one-
dimensional sub manifold of R2(verify this!)precisely because of this point
of intersection.

If the origin 0 G R2 is removed from E, then the set E \ 0 will now obviously satisfy Definition 1. We remark that the set <='\0 is not connected.
It consists of four pairwise disjoint rays.

Thus a ^-dimensional surface in Rn satisfying Definition 1 can happen to be a disconnected subset consisting of several connected components(and
these components are connected ^-dimensional surfaces). A surface in Rn is
often taken to mean a connected ^-dimensional surface. Just now we shall be
interested in the problem of finding extrema of functions defined on surfaces.

These are local problems, and therefore connectivity will not manifest itself
in them.

Example. If a smooth mapping /: G —> Rn of the domain G C Rn defined
in coordinate form has rank k at the point to G G, then there exists a
neighborhood U(to)C G of this point whose image f(U(to)) C Rn
is a smooth surface in Rn.

Indeed, as already noted above, in this case relations can be re-
placed by the equivalent system

' xk+1 _ pk+1^1^ ^ ^ ^ ^ xk^ ^

< (8.139)

^ xn=^(x1, ...,xk)

in some neighborhood U(to)of to <= G.(For simplicity of notation, we let
that the system /x, ..., fk has rank k.)Setting

Fi(x\ ..., xn)=xk+i - /+1(x\ ..., xk)(i=1, ..., n - k),

the set f(U(to)) is indeed a ^-dimensional smooth surface in Rn

6.5 THE TANGENT SPACE

In studying the law of motion x=x(t)of a point mass in R3, starting from
the relation

x(t)=x(0)+ x'(0)<= + o(t)as t —> 0

and assuming that the point t=0 is not a critical point of the mapping
R 3 t x(t)G R3, that is, x'(0)7^ 0, we defined the line tangent to the
trajectory at the point x(0)as the linear subset of R3 given in parametric
form by the equation

x-x0=x'(0)t •

or the equation

x - xo=<= • t,

where xq=x(0)and <==x'(0)is a direction vector of the line.

In essence, we did a similar thing in defining the tangent plane to the graph of a function z=f(x, y)in R3. Indeed, supplementing the relation z= /(x, y)with the trivial equalities x=x and y=y, we obtain a mapping R2 3(x, y)i->(x, 2 /, /(x, y)) G R3 to which the tangent at the point(xo, 2 /o)
is the linear mapping

=(0 1)(®:®°).

\z-zo)\ f'x(xo, yo)fy(x0, 2 /0)/ ^ '

where z0=f(x0, y0).

Setting t=(x — xo, y — yo)and x=(x — xo, y — yo>z ~ ^0)here, and denoting the Jacobi matrix in for this transformation by x'(0), we remark that its rank is two and that in this notation relation has the form

The peculiarity of relation is that only the last equality in the set of three equalities

{

X — Xo=X — Xo,

y - yo=y - yo, (8.144)

^-^0=fx(xo, yo)(x-x0)+ ffao, y-0)(y-yo),

to which it is equivalent is a nontrivial relation. That is precisely the reason
it is retained as the equation defining the plane tangent to the graph of s= /(x, y)at(x0, yo, zo).

This observation can now be used to give the definition of the k-dimensional plane tangent to a ^-dimensional smooth surface S C Rn.

It can be observen from Definition 1 of a surface that in a neighborhood of

each of its points Xo G S a ^-dimensional surface S can be defined para- metrically, that is, using mappings Ik 3(t1, ..., ^)(x1, ...^71)G S.
Such a parametrization can be taken to be the restriction of the mapping(p 1 : In —> U(xo)to the ^-dimensional plane tk+1=• • •=tn=0

Since(p~x is a diffeomorphism, the Jacobian of the mapping <p~x : In —>
U(xo)is nonzero at each point of the cube In. But then the mapping Ik 3 (t1, ..., tk)(x1, ..., xn)G S obtained by restricting <^-1 to this plane must also have rank k at each point of Ik.

Now setting(t1, ...,tk)=t G Ik and denoting the mapping Ik 3 t »-> x G S by x=x(t), we obtain a local parametric representation of the surface S possessing the property expressed by(8.140), on the basis of which we take as the equation of the tangent space or tangent plane to the surface S C En at xo G S.

Thus we adopt the following definition.

Definition. If a ^-dimensional surface S C Rn, 1 < A; < n, is defined parametrically in a neighborhood of Xo G 5 by means of a smooth mapping
(t1, ..., tk)=t x=(x1, ..., xn)such that Xo=x(0)and the matrix x'(0) has rank k, then the ^-dimensional surface in Rn defined parametrically by
the matrix equality is known the tangent plane or tangent space to the surface S at xo G S.

In coordinate form the following system of equations corresponds to Eq.
xl~x°=+ ■ • •+ %^tk >

dxn, ^., . dxn k

dtl-(o)t1 + --- + -^ir(o)tk.

We shall denote9 the tangent space to the surface S at x G 5, as before, by
TSX.

Notes

An important and useful exercise, which the reader can do independently,
is to prove the invariance of the definition of the tangent space and the verification that the linear mapping t x'(0)t tangent to the mapping t x(t), which defines the surface S locally, provides a mapping of the space

= TRq onto the plane T5x(o)(observe Problem 3 at the end of this section).

Let us now determine the form of the equation of the tangent plane to the ^-dimensional surface S defined in Rn by the system. For definiteness we shall let that condition holds in a neighborhood of the point Xo G S.

Setting(x\..., xfc)=u, (xfc+1, ..., xn)=v, (F1, ...,Fn~k)=F, we write the system in the form

F(u, v)=0,

6.6 DIFFERENTIAL CALCULUS IN SEVERAL VARIABLES

F'v(u, v)7^ 0

Using the implicit function theorem, in a neighborhood of the point (i<=o, ^o)=(#0' • • • >xo>xo+1> • • • >xo)we Pass from relation to the equivalent relation

V=f(u),

which, when we supplement it with the identity u — u, yields the parametric
representation of the surface S in a neighborhood of Xq G S:

On the basis of Definition we obtain from the parametric equation .

u — Uo=E -1,

v-v0=f'{ u0)■ t

of the tangent plane; here E is the identity matrix and t=u — Uq.

Just as was done in the case of the system we retain in the system only the nontrivial equation

v-vo=f'(uo)(u - uo)

which contains the connection of the variables x1, ... with the variables <=fc+1, ..., xn that determine the tangent space.

Using the relation

f'(u0)=-[^(tio, «o)]_1[Fi(«o, «o)],

which follows from the implicit function theorem, we rewrite as

K(uo, v0)(u - Uo)+ F'v(uo -,vo){ v - v0)=0,

from which, after returning to the variables(x1, ...,xn)=x, we obtain the equation we are observeking for the tangent space TSXq C Mn, namely

F'x(xo)(x-x o)=0.

In coordinate representation Eq. is equivalent to the system of equations

iO /?!   iO /?!

^OroX*1 - xl)+ • • • + ^(*o)(»n - Off)=o,
fpn—k rpn—k

dx!(xo){ xl - Xq)+ • • • + dxn(x0){ xn - x$)=0 .

By hypothesis the rank of this system is n — /c, and hence it defines a /c-dimensional plane in Rn.

The affine equation is equivalent(given the point xo)to the vector equation

F'x(x oK=0,

in which <==x — xo.

Hence the vector <= lies in the plane TSXo tangent at xo G S to the surface
S C Rn defined by the equation F{ x)=0 if and only if it satisfies condition

Thus TSXo can be regarded as the vector space consisting of the vectors <= that satisfy

It is this fact that motivates the use of the term tangent space.

Let us now prove the following proposition, which we have already en-countered

Proposition. The space TSXo tangent to a smooth surface S C Mn at a point
xo G S consists of the vectors tangent to smooth curves lying on the surface
S and passing through the point xo.

Proof Let the surface S be defined in a neighborhood of the point xo G S by a system of equations, which we write briefly as

F{ x)=0,

where F=(F1, ..., Fn~k), x=(x1, ..., xn). Let JH : I —> S be an arbitrary smooth path with support on S. Taking J={ tGR||t|<l }, we shall let that x(0)=xo. Since x(t)G S for t G /, after substituting x{ t)into Eq.

, we obtain

F(x(t))=0

for tel. Differentiating this identity with respect to <=, we find that

F'x(x(t))-x'{ t)=0.

In particular, when t=0, setting <==^'(O), we obtain

K(x o)<==0,

that is, the vector <= tangent to the trajectory at xo(at time t=0)satisfies Eq. of the tangent space TSXo.

We now show that for every vector <= satisfying Eq. there exists a smooth path T : I —> S that defines a curve on S passing through xo at t=0 and having the velocity vector <= at time t — 0.

This will simultaneously establish the existence of smooth curves on S passing through xo, which we letd implicitly in the proof of the first part of the proposition.

Suppose for definiteness that condition holds. Then, knowing the first k coordinates <=x, ..., of the vector <==(<=x, ..., <=fc, <=fc+1, ..., <=n), we determine the other coordinates <=fc+1, ..., <=n uniquely from Eq(which is equivalent to the system). Thus, if we establish that a vector <==(f1, ..., <=n)satisfies Eq. we can conclude that <= — <=• We shall make use of this fact.

Again, as was done above, we introduce for convenience the notation u=(x1, ...,xk), v=(xk+1, ...,xn), x=(x1, ...,xn)=(u, v), and F(x)= F(u, v). In the subspace Rk C Rn of the variables x1, ...,xk we choose a parametrically defined line

X1 -xl =,

f G]R,

xk - »*=zkt,

having direction vector(C1, • • •, <=fc), which we denote In more abbreviated
notation this line can be written as

u=uo + <=ut.

Solving Eq. for v, by the implicit function theorem we obtain a smooth function, which, when the right-hand side of Eq. is substituted as its argument is taken account of, yields a smooth curve in Rn defined as follows:

u=uo -f <=ut,

t e 17(0)G R .

v=f(uo 4- <=ut),

Since F(u, f(u))=0, this curve obviously lies on the surface S. Moreover, it is clear from Eqs. that at t=0 the curve passes through the point (uo, v0)=(zj, ..., a:§, a;<=+1, ..., a#)=r0<=S.

Notes

Differentiating the identity

F(u(t), v(t))=F(u0 + <=ut, f(u0 + <=ut))=0

with respect to t, we obtain for t=0

Fu(uOivo)€u 4- Fv(uo, vo)<=v=0,

where <=u=i /(0)=(Cfc+1, ...,^n). This equality shows that the vector <= =

(tu, tv) —(C1 ? • • • ? €k, • • • ? <=n)satisfies Eq. Thus by the remark made above, we conclude that <==<=. But the vector <= is the velocity vector

at t=0 for the trajectory. The proposition is now proved.

6.7 EXTREMA WITH CONSTRAINT

Statement of the Problem One of the most brilliant and well-known achievements of differential calculus is the collection of recipes it provides
for finding the extrema of functions. The necessary conditions and sufficient
differential tests for an extremum that we obtained from Taylor's theorem apply, as we have noted, to interior extrema.

In other words, these results are applicable only to the study of the be- havior of functions Rn 3 x f{ x)G R in a neighborhood of a point G Rn, when the argument x can let any value in some neighborhood of Xq in

Rn. Frequently a situation that is more complicated and from the practical
point of view even more interesting arises, in which one observeks an extremum of
a function under certain constraints that limit the domain of variation of the
argument. A typical example is the isoperimetric problem, in which we observek
a body of maximal volume subject to the condition that its boundary surface
has a fixed area. To obtain a mathematical expression for such a problem that will be accessible to us, we shall simplify the statement and let that the problem is to choose from the set of rectangles having a fixed perimeter
2p the one having the largest area a. Denoting the lengths of the sides of the
rectangle by x and y, we write

a(x, h)=x-y,
x + y=p.

Thus we need to find an extremum of the function a(x, y)under the condition that the variables x and y are connected by the equation x +

$y=p$.

Therefore, the extremum is being sought only on the set of points of R^2 satisfying this relation. This particular problem, of course, can be solved without difficulty: it suffices to write $y=p — x$ and substitute this expression
into the formula for $a(x, y)$, then find the maximum of the function $x(p — x)$
by the usual methods. We needed this example only to explain the statement
of the problem itself.

In general the problem of an extremum with constraint usually amounts to finding an extremum for a real-valued function

$y=f(x_1, ..., x_n)$

of n variables under the condition that these variables must satisfy a system
of equations

' $Fl(xl, ..., xn)=0,$

<

w $F_m(x_1, ...,x_n)=0.$

Since we are planning to obtain differential conditions for an extremum, we shall let that all these functions are differentiate and even continuously differentiate. If the rank of the system of functions $F_1, ..., F_m$ is n — /c, conditions define a /c-dimensional smooth surface S in R^n, and
from the geometric point of view the problem of extremum with constraint
amounts to finding an extremum of the function /on the surface S. More precisely, we are considering the restriction f\s of the function /to the surface
S and observeking an extremum of that function.

The meaning of the concept of a local extremum itself here, of course, remains the same as before, that is, a point xo G S is a local extremum of

/on,S, or, more briefly f\s, if there is a neighborhood10 Us(xo)of Xq in S C Rn such that f{ x)> f(xo)for any point x G Us(xo)(in which case xo is
a local minimum)or f(x)< f{ xo)(and then Xo is a local maximum). If these
inequalities are strict for x G Us(xo)\#o, then the extremum, as before, will
be known strict.

A Necessary Condition for an Extremum with Constraint

Theorem . Let f : D —> R be a function defined on an open set D CW1 and belonging to C^(D; R). Let S be a smooth surface in D.

A necessary condition for a point x$ G S that is noncritical for f to be a local extremum of f\s is that

TSXn C TNXl

where TSXo is the tangent space to the surface S at x$ and TNXo is the tangent space to the level surface N={ x G D\ f(x)=f(xo)} of f to which xo belongs.

We begin by remarking that the requirement that the point xo be non-critical for /is not an essential restriction in the context of the problem of finding an extremum with constraint, which we are discussing. Indeed, even
if the point Xo G D were a critical point of the function /: D -¥ R or an extremum of the function, it is clear that it would still be a possible or

actual extremum respectively for the function f\s. Thus, the new element in
this problem is precisely that the function f\s can have critical points and extrema that are different from those of /.

Proof We choose an arbitrary vector <= G TSXo and a smooth path x=x(t)

Notes

on S that passes through this point at t=0 and for which the vector <= is the velocity at t=0, that is,

§«>)=<=.

If xo is an extremum of the function f\s, the smooth function f[x(t)) must have an extremum at t=0. By the necessary condition for an extremum, its derivative must vanish at t=0, that is, we must have

/'(«<>)- ^

Since Xo is a noncritical point of /, condition is equivalent to the condition that <= g TNXo, for relation is precisely the equation of the tangent space TNXo.

Thus we have proved that TSXq C TNXq . □

If the surface S is defined by the system of equations in a neighborhood of <=o, then the space T5Xo, as we know, is defined by the system of linear equations and, since every solution of is a solution of the latter equation
is a consequence

It follows from these considerations that the relation TSXo C TNXo is equivalent to the analytic statement that the vector grad /(xo)is a linear combination of the vectors gradF2(xo), (i=1, ...,m), that is,

grad /(a:o)=XigradFl(x0)

2=1

Taking account of this way of writing the necessary condition for an extremum of a function whose variables are connected by Lagrange proposed using the following auxiliary function when observeking a constrained extremum:

L{ x, \)=fix)-J2Xipi(x)

2=1

in n -f m variables(x, A)=(x1, ..., xm, Ai, ..., An).

This function is known the Lagrange function and the method of using it is the method of Lagrange multipliers.

The function is convenient because the necessary conditions for an extremum of it, regarded as a function of(x, A)=(x1, ..., xm, Ai, ..., An), are precisely

Indeed,

/)f m f)F*

1-1

—Fl{ x)=0 (z=1, ..., m).

Thus, in observeking an extremum of a function whose variables are subject to the constraints one can write the Lagrange function with undetermined multipliers and look for its critical points. If it is possible to find Xo=(xj, ...,Xq)from the system without finding A=(Ai, ..., Am), then, as far as the original problem is concerned, that is what should be done.

As can be observen from the multipliers A^(i=1, ..., m)are uniquely determined if the vectors grad Fl(xo)(i=1, ..., m)are linearly independent. The independence of these vectors is equivalent to the statement that the rank of the system is m, that is, that all the equations in this system are essential(none of them is a consequence of the others).

This is usually the case, since it is letd that all the relations are independent, and the rank of the system of functions F1, ..., Fm is m at every point x G X.

The Lagrange function is often written as

m

L(x, \)=f(x)+ J2*iFi(x),

2=1

which differs from the preceding expression only in the inessential replace-
ment of Ai by — Ai.11

Example . Let us find the extrema of a symmetric quadratic form

n

f(x)=Y, a, ijXlxj (ctij=aji)

on the sphere

n

F(x)= -1=0-(8-170)

2=1

6.8 SURFACES IN RN AND CONSTRAINED EXTREMA

Let us write the Lagrange function for this problem

n /n \

L(X> X)= Y2 aijX%X1 ~ A ~ 1)'

i, j=1 ^ i=l '

and the necessary conditions for an extremum of L(x, A), taking account of
the relation aij=a^:

dL f n \

q^{ x, A)=2(J2 aiix3 ~ Xxt)=0(i=I, -- -,"),

<

JlfeA)=(E(,T-i)=o.

Multiplying the first equation by x% and summing the first relation over i, we find, taking account of the second relation, that the equality

n

^2 - a=0

i, j=1

must hold at an extremum.

The system minus the last equation can be rewritten as

n

dijxi=\xl(i=1, ..., n),

2=1

from which it follows that A is an eigenvalue of the linear operator A defined
by the matrix(a^-), and x=(x1, ...,xn)is an eigenvector of this operator corresponding to this eigenvalue.

Since the function which is continuous on the compact set S=ix G Mn|
=must let its maximal value at some point, the 2= 1 J
system and hence also must have a solution. Thus we have established along the way that every real symmetric matrix(a^)has at least one real eigenvalue. This is a result well-known from linear algebra, and is fundamental in the proof of the existence of a basis of eigenvectors for a symmetric operator.

To show the geometric meaning of the eigenvalue A, we remark that if A > 0, then, passing to the coordinates t%=xl /V\ we find, instead of
n

Y2=1,

M=1

and, instead of

fy?=\ .

2=1

n

But <=(t*)2 is the square of the distance from origin to the point

2=1

t=(t1, ...,tn), which lies on the quadric surface. Thus if
represents, for example, an ellipsoid, then the reciprocal 1 /A of the eigen-
value A is the square of the length of one of its semi-axes.

This is a useful observation. It shows in particular that relations, which are necessary conditions for an extremum with constraint, are still not sufficient. After all, an ellipsoid in R3 has, besides its largest and smallest semi-axes, a third semi-axis whose length is intermediate between the two, in any neighborhood of whose endpoint there are both points nearer to the origin and points farther away from the origin than the endpoint. This last becomes completely obvious if we consider the ellipses obtained by taking a section of the original ellipsoid by two planes passing through the intermediate-length semi-axis, one passing through the smallest semi-axis and the other through the largest. In one of these cases the intermediate axis will be the major semi-axis of the ellipse of intersection. In the other it will be the minor semi-axis.

To what has just been said we should add that if 1/is the length of this intermediate semi-axis, then, as can be observen from the canonical equation of

an ellipsoid, A will be an eigenvalue of the operator A. Therefore the system which expresses the necessary conditions for an extremum of the function f\s will indeed have a solution that does not give an extremum of the function.

The result obtained in Theorem(the necessary condition for an extremum with constraint)is illustrated The first of these figures explains why the point of the surface S cannot

be an extremum of f\s if S is not tangent to the surface N={ x G Rn| f(x)= f(xo)=Co } at Xq. It is letd here that grad /(xo)7^ 0. This last condition guarantees that in a neighborhood of Xo there are points of a higher, C2-level

surface of the function /and also points of a lower, c\-level surface of the function.

Since the smooth surface S intersects the surface AT, that is, the co-level surface of the smooth function /, it follows that S will intersect both higher and lower level surfaces of /in a neighborhood of Xq. But this means that

xo cannot be an extremum of f\s.

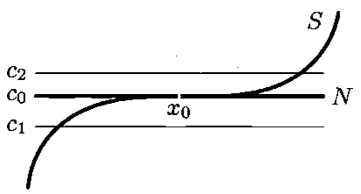
The second figure shows why, when N is tangent to S at Xo, this point can turn out to be an extremum. In the figure Xo is a local maximum of f\s.

These same considerations make it possible to sketch a picture whose analytic expression can show that the necessary criterion for an extremum is

not sufficient.

Indeed, in accordance with we set, for example,

f(x, y)=y, F(x, y)=x3 -y=0 .



It is then obvious that y has no extremum at the point(0, 0)on the curve S C M2 defined by the equation y=x3, even though this curve is tangent to the level line f(x, y)= 0 of the function f at that point. We remark that grad /(0, 0)=(0, 1)^0.

It is obvious that this is essentially the same example that served earlier to illustrate the difference between the necessary and sufficient conditions for a classical interior extremum of a function.

6.9 SOME GEOMETRIC IMAGES CONNECTED WITH FUNCTIONS OF SEVERAL VARIABLES

The Graph of a Function and Curvilinear Coordinates Let x, y, and z be Cartesian coordinates of a point in R3 and Let z= /(x, y)be a continuous function defined in some domain G of the plane R2 of the variables x and y.

By the general definition of the graph of a function, the graph of the function /: G —» R in our case is the set S={(x, y, z)G R3|(x, y)G G, z= /(x, y)} in the space R3.

It is obvious that the mapping G —> S defined by the relation(x, y)i-» (x, 2 /, /(x, y)) is a continuous one-to-one mapping of G onto S, by which one
can determine every point of S by exhibiting the point of G corresponding
to it, or, what is the same, giving the coordinates(x, y)of this point of G.

Thus the pairs of numbers(x, y)G G can be regarded as certain coordinates of the points of a set S - the graph of the function z= /(x, y). Since the points of S are given by pairs of numbers, we shall conditionally agree to
call S a two-dimensional surface in R3.(The general definition of a surface
will be given later.)

If we define a path T : I —» G in G, then a path F o T : I —» S automatically appears on the surface S. If x=x(t)and y=y(t)is a parametric definition of the path T, then the path F o T on S is given by the three functions x=x(<=), y=y(t), z=z(t)=f(x(t), y(t)). In particular, if we set x=xo +1, y=2 /0? we obtain a curve x=xo +1, y=yo, z=f(xo +1, yo)on the surface S along which the coordinate y=yo of the points of S does not change. Similarly one can exhibit a curve x=xo, y=yo + <=, z — f(xo? 2 /o + <=)on S along which the first coordinate xo of the points of S does not change.

By analogy with the planar case these curves on S are naturally known coordinate lines on the surface S. However, in contrast to the coordinate lines in G C R2, which are pieces of straight lines, the coordinate lines on S are in general curves in R3. For that reason, the coordinates(x, y)of points of the
surface S are often known curvilinear coordinates on S.

Thus the graph of a continuous function z= /(x, y), defined in a domain G C R2 is a two-dimensional surface S in R3 whose points can be defined by
curvilinear coordinates(x, y)G G.

At this point we shall not go into detail on the general definition of a surface, since we are interested only in a special case of a surface - the graph
of a function. However, we let that from the course in analytic geometry the reader is well acquainted with some important particular surfaces in R3
(such as a plane, an ellipsoid, paraboloids, and hyperboloids).

The Tangent Plane to the Graph of a Function Differentiability of a function z= /(x, y)at the point(xo, yo)G G means that

/(®> v)=f(xo, yo)+ A(x ~ xo)+ B(y - y0)+

+o(y /(x - x0)2 +(y - yo)2)as(x, y)->•(x0, yo)

where A and B are certain constants.

In R3 Let us consider the plane

Notes

$z=z0 + A(x - x0)+ B(y - y0)$,

where zo= /(xo, yo)• Comparing equalities, we observe that the graph of the function is well approximated by the plane in a neighborhood of the point(xo, yo, zo). More precisely, the point(x, y, /(x, y)) of
the graph of the function differs from the point(x, y, z(x, y)) of the plane by an amount that is infinitesimal in comparison with the magnitude yj(x — xo)2 +(y — yo)2 of the displacement of its curvilinear coordinates
(x, y)from the coordinates(xo^yo)of the point(xo, 2 /o?^o)-

By the uniqueness of the differential of a function, the plane possessing this property is unique and has the form

2=f(xo, yo)+ yo)(x - x0)+ yo)(y - yo)•

This plane is known the tangent plane to the graph of the function z= /(x, y)
at the point(x0, yo, /(xo, yo)) •

Thus, the differentiability of a function z=f(x, y)at the point(xo^yo) and the existence of a tangent plane to the graph of this function at the point
(#0? 2 /o, /(^o? Vo)) are equivalent conditions.

The Normal Vector for the tangent plane in the canonical form

|^(zo, yo)(x - xo)+ yo)(y - yo)-(2 - fix0, y0))=0.

Check your Progress-1

Discuss Taylor's Theorem, Maxima And Minima

____

____

____

Discuss Differential Calculus In Several Variables

____

6.10 LET US SUM UP

In this unit we have discussed the definition and example of Taylor's Theorem, Maxima And Minima, Application To Maxima And Minima, Surfaces In Rn And The Theory Of Extrema With Constraint, The Tangent Space, Differential Calculus In Several Variables, Extrema With Constraint, Surfaces In Rn And Constrained Extrema, Some Geometric Images Connected With Functions Of Several Variables

6.11 KEYWORDS

1. Taylor's Theorem, Maxima And Minima: This is somewhat complicated, and the longest proof in either content here to carry the expansion out one more term than thus adding a third derivative and discussing the resulting expression.

2.Application To Maxima And Minima The equation above allows us to determine criteria guaranteeing that a point x_o is a local maximum or local minimum for the function f

3. Surfaces In Rn And The Theory Of Extrema With Constraint To acquire an informal understanding of the theory of extrema with

constraint, which is important in applications, it is useful to have some elemen-
tary information on surfaces(manifolds)in Rn.

4.The Tangent Space In studying the law of motion x=x(t)of a point mass in R3, starting from the relation x(t)=x(0)+ x'(0)<= + o(t)as t —> 0

5. Differential Calculus In Several Variables: Using the implicit function theorem, in a neighborhood of the point (i<=o, ^o)=(#0' • • • >xo>xo+1> • • • >xo)we Pass from relation to the equivalent relation.

6.Surfaces In Rn And Constrained Extrema Lagrange function

7. Some Geometric Images Connected With Functions Of Several Variables

8.The Graph of a Function and Curvilinear Coordinates Let x, y, and z be Cartesian coordinates of a point in R3 and Let z= /(x, y)be a continuous function defined in some domain G of the plane R2 of the variables x and y.

6.12 QUESTIONS FOR REVIEW

Explain Taylor's Theorem, Maxima And Minima

Explain Differential Calculus In Several Variables

6.13 ANSWERS TO CHECK YOUR PROGRESS

Taylor's Theorem, Maxima And Minima

(answer for Check your Progress-1 Q)

Differential Calculus In Several Variable

(answer for Check your Progress-1 Q)

6.14 REFERENCES

- Calculus of Several Variables
- Advance Calculus of Several Variables
- Analysis of Several Variables
- Application of Several Variables
- Function of Several Variables

UNIT - 7: A SUFFICIENT CONDITION FOR A CONSTRAINED EXTREMUM

STRUCTURE

7.0 OBJECTIVES

After studying this unit, you should be able to:

Learn, Understand about A Sufficient Condition For A Constrained Extremum

Learn, Understand about Differential Calculus Differentiable Functions

Learn, Understand about Functions Differentiable At A Point

Notes

Learn, Understand about The Tangent Line; Geometric Meaning Of The Derivative And Differential

Learn, Understand about The Role Of The Coordinate System

7.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

A Sufficient Condition For A Constrained Extremum, Differential Calculus Differentiable Functions, Functions Differentiable At A Point, The Tangent Line; Geometric Meaning Of The Derivative And Differential, The Role Of The Coordinate System

7.2 A SUFFICIENT CONDITION FOR A CONSTRAINED EXTREMUM

We now prove the following sufficient condition for the presence or absence of a constrained extremum.

Theorem. Let f : D —> R be a function defined on an open set DcRn and belonging to the class C^(D;R); Let S be the surface in D defined by Eqs. where F% G C^2^(D;R)(i=1, ...,m)and the rank of the system of functions { F1, ..., Fm } at each point of D is m.

Suppose that the parameters Ai, ..., Am in the Lagrange function

m

L(x)=L(x;X)= /(x1, ..., xn)- ^ \iF\x1, ...,xn)

2=1

have been chosen in accordance with the necessary criterion for a constrained extremum of the function f\s at xo G

A sufficient condition for the point xo to be an extremum of the function f\s is that the quadratic form

f)2 r

~(xo)ee

dxidxi

be either positive-definite or negative-definite for vectors <= G TSXo.

If the quadratic form is positive-definite on TSXo, then x$ is a strict local minimum of f\s; if it is negative-definite, then Xo is a strict local maximum.

A sufficient condition for the point Xo not to be an extremum of f\s is that the form let both positive and negative values on TSXo.

Proof We first note that L(x)=f{ x)for x G <S, so that if we show that x$ G
S is an extremum of the function L\s, we shall have shown simultaneously that it is an extremum of f\s.

By hypothesis, the necessary criterion for an extremum of f\s at xo is fulfilled, so that gradI /(xo)=0 at this point. Hence the Taylor expansion of L{ x)in a neighborhood of Xo=(xj, ..., Xq)has the form

1 rP T ....

L{ x)- L(x0)=2idxidxj(x o)(g* - 4)(xj ~A)+ °{ \\x ~ xof)

as x —> xo.

We now recall that, in motivating Definition, we noted the possibility of a local(for example, in a neighborhood of Xo G S)parametric definition of a smooth /c-dimensional surface S(in the present case, k=n — m).
In other words, there exists a smooth mapping

R 3(t1, ..., tk)=t\-+ x=(x1, ..., xn)G Rn

(as before, we shall write it in the form x=x(t)), under which a neighborhood of the point 0=(0, ..., 0)G Rfc maps bijectively to some neighborhood of Xo
on 5, and xq=x(0).

We remark that the relation

x(t) — x(0)=^'(O)^ 4- o(||t||)as t —> 0,

which expresses the differentiability of the mapping t x(t)at t=0, is equivalent to the n coordinate equalities

s*(i)- x'(0)=^(0)iQ + o(\\t\\)(i=1, ..., n),

in which the index a ranges over the integers from 1 to k and the summation is over this index.

It follows from these numerical equalities that

|xl(t) — x2(0)|=0(||<=||)as t -> 0

and hence

IIx(t)- a;(0)||Rn=0(||i||Rfc)as t -> 0 .

Using relations, we find that as t —> 0

L(x{ t)) — L(a:(0))=^dijL(xo)dax'l(0)di3x(0)tat<3 + o(p||2)

Hence under the assumption of positive- or negative-definiteness of the form

dijL(x0)daxi(0)0^x^0)^

it follows that the function L(x(t)) has an extremum at t=0. If the form lets both positive and negative values, then L(x(t)) has no extremum at t=0. But, since some neighborhood of the point 0 G Rk maps to a neighborhood of x(0)=Xo G S on the surface S under the mapping t i-» #(<=), we can conclude that the function L\s also will either have an extremum at xo of the same nature as the function L(#(<=)) or, like L(#(<=)), will not have an extremum.

Thus, it remains to verify that for vectors <= G TSXo the expressions and are merely different notations for the same object.

Indeed, setting

<==x'(0)t,

we obtain a vector <= tangent to S at xo, and if <==(<=x, ..., <=n), x{ t)=(x1, ..., xn)(t), and t=(t1, ...,tk), then

<==dpx3(0)^(j=l, ..., n),

from which it follows that the quantities are the same

We note that the practical use of Theorem is hindered by the fact that only k — n — m of the coordinates of the vector <==(<=x, ...,<=n)G TSXo are independent, since the coordinates of <= must satisfy the system defining the space TSXo. Thus a direct application of the Sylvester criterion to the quadratic form generally yields nothing in the present case the form can not be positive- or negative-definite on TR<=o and yet be definite on TSXo. But if we express m coordinates of the vector <= in terms of the other k coordinates by relations and then substitute the resulting linear forms into, we arrive at a quadratic form in k variables whose positive- or negative-definiteness can be investigated using the Sylvester criterion.

Let us clarify what has just been said by some elementary examples.

Example. Suppose we are given the function

f(x, y, z)=x2 -y2 + z2

in the space R3 with coordinates x, y, z. We observe an extremum of this function on the plane S defined by the equation

F(x, y1z)=2x — y — 3=0.

Writing the Lagrange function

L(x, y, z)=(x2 - y2 + z2)- \(2x - y - 3)

and the necessary conditions for an extremum

2x — 2A=0,

dL n

tt~ — —2 y + A — 0,

dy

dL

dz

^= —{ 2x — y — 3)=0,

we find the possible extremum p=(2, 1, 0).

Next we find the form

fatfe=(?)*-(?)* +(?)*

We note that in this case the parameter A did not occur in this quadratic form, and so we did not compute it.

We now write the condition <= E TSP:

1 — <=2=0 .

From this equality we find <=2=2<=x and substitute it into the form after which it lets the form

where this time and <=3 are independent variables.

This last form can obviously let both positive and negative values, and therefore the function f\s has no extremum at p E S.

Example. Under the hypotheses of Example we replace R3 by R2 and the function /by

f(x, y)=x2 - y2,

retaining the condition

2x — y - 3=0,

which now defines a line S in the plane R2.

We find p=(2, 1)as a possible extremum.

Instead of the form we obtain the form

(e1)2 -(<=2)2

with the previous relation between and <=2.

Thus the form now has the form

-3(^)2

on TSP, that is, it is negative-definite. We conclude from this that the point p=(2, 1)is a local maximum of f\s.

The following simple examples are instructive in many respects. On them one can distinctly trace the working of both the necessary and the sufficient conditions for constrained extrema, including the role of the parameter and the informal role of the Lagrange function itself.

Example. On the plane R2 with Cartesian coordinates(x, y)we are given the function

f(x, y)=x2 + y2 .

Let us find the extremum of this function on the ellipse given by the canonical relation

x2 t /2

y)= ~2 + =

where 0 < a < b.

It is obvious from geometric considerations that min /|^=a2 and max /|^=b2. Let us obtain this result on the basis of the procedures recommended

By writing the Lagrange function

L(x, y, a)=(x2 + y2)-a(^ + ~ *)

and solving the equation dL=0, that is, the system — —
find the solutions

(x, y, A)=(±a, 0, a2), (0, ±b, b2).

Now in accordance with Theorem 2 we write and study the quadratic form
\d2Z /<=2, the second term of the Taylor expansion of the Lagrange function in a neighborhood of the corresponding points:

J**+(!-<=)«•>■.

At the points(±a, 0)of the ellipse S the tangent vector <==(<=\<=2)has the form(0, <=2), and for A=a2 the quadratic form lets the form

0-<=)<«2>2-

Taking account of the condition 0 < a < 6, we conclude that this form is positive-definite and hence at the points(±a, 0)G S there is a strict local(and in this case obviously also global)minimum of the function f\s, that is,
min /|g=a2.

Similarly we find the form

('-?)«■>•■

which corresponds to the points(0, ±b)G 5, and we find max /|^=b2.

Remark. Note the role of the Lagrange function here compared with the role of the function /. At the corresponding points on these tangent vectors the differential of /(like the differential of L)vanishes, and the quadratic form
|d 2 /€2=(C1)2 +(<=2)2 is positive definite at whichever of these points it is computed. Nevertheless, the function f\s has a strict minimum at the points
(±a, 0)and a strict maximum at the points(0, ±b).

To understand what is going on here, look again at the proof of Theorem. and try to obtain relation by substituting /for L in. Note that an additional term containing ^"(O)arises here. The reason it does not vanish is that, in contrast to dL the differential df of /is not identically zero at the corresponding points, even though its values are indeed zero on the tangent vectors(of the form ^'(O)).

Example. Let us find the extrema of the function f(x, y, z)=x2+y2 + z2

on the ellipsoid S defined by the relation

F(s, y, *)=^ + g + 4- 1=0,

az bz c

where 0 < a < b < c.

By writing the Lagrange function

L{ x, y, z, \)=(x2 4-y2 -f z2) — ^(^2 + ^2 ^2 ~~ '

in accordance with the necessary criterion for an extremum, we find the solutions of the equation dL=0, that is, the system = §X=0:

(x, y, z, X)=(±a, 0, 0, a2),(0, ±b, 0, b2),(0, 0, ±c, c2),

On each respective tangent plane the quadratic form

in each of these cases has the form

(l-<=)(??+(l-<=)((>?, (a)

+(>"?)«■>•■ «

0-?)(«■)'+(i-p)<«2>'- (c)

Since 0 < a < b < c, it follows from Theorem which gives a sufficient criterion for the presence or absence of a constrained extremum, that one can conclude that in cases(a)and(c), we have found respectively min /|^=a2 and max /|^=c2, while at the points(0, ±6, 0)G S corresponding to case

(b)the function f\s has no extremum. This is in compLete agreement with the obvious geometric considerations stated in the discussion of the necessary criterion for a constrained extremum.

Certain other aspects of the concepts of analysis and geometry encountered in this section, which are sometimes quite useful, including the physical interpretation of the problem of a constrained extremum itself, as well as the necessary criterion for it as the resolution of forces at an equilibrium point and the interpretation of the Lagrange multipliers as the magnitude of the reaction of ideal constraints.

7.3 DIFFERENTIAL CALCULUS DIFFERENTIABLE FUNCTIONS

Statement of the Problem and Introductory Considerations

Suppose, following Newton, 12 we wish to solve the Kepler problem of two bodies, that is, we wish to explain the law of motion of one celestial body m(a planet)relative to another body M(a star). We take a Cartesian coordinate system in the plane of motion with origin at M. Then the position of m at time t can be characterized numerically by the coordinates(x(t), y(t)) of the point in that coordinate system. We wish to find the functions x(t)and y(t).

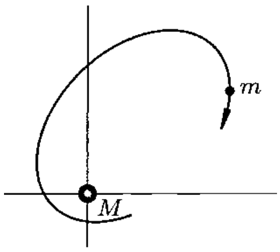


Fig. 5.1.

The motion of m relative to M is governed by Newton's two famous laws: the general law of motion

ma=F,

connecting the force vector with the acceleration vector that it produces via the coefficient of proportionality m - the inertial mass of the body and the law of universal gravitation, which makes it possible to find the gravitational action of the bodies m and M on each other according to the formula

_ „mM . .

= "jrj^r'

where r is a vector with its initial point in the body to which the force is applied and its terminal point in the other body and |r| is the length of the vector r, that is, the distance between m and M.

Knowing the masses m and M, we can easily use to express the right-hand side in terms of the coordinates x(t)and y(t)of the body m at time <=, and thereby take account of all the data for the given motion.

To obtain the relations on x(t)and y(t)contained in we must learn how to express the left-hand side of in terms of x(t)and y(t). Acceleration is a characteristic of a change in velocity v(<=). More precisely, it is simply the rate at which the velocity changes. Therefore, to solve the problem we must first of all learn how to compute the velocity v(<=)at time t possessed by a body whose motion is described by the radius-vector r(t)=(x(t), y(t)).

Thus we wish to define and learn how to compute the instantaneous velocity of a body that is implicit in the law of motion

To measure a thing is to compare it to a standard. In the present case, what can serve as a standard for determining the instantaneous velocity of motion

The simplest kind of motion is that of a free body moving under inertia. This is a motion under which equal displacements of the body in space(as vectors)occur in equal intervals of time. It is the so-known uniform(rectilinear)motion. If a point is moving uniformly, and r(0)and r(l)are its radius-vectors relative to an inertial coordinate system at times t=0 and t=1 respectively, then at any time t we shall have

$$r(t) — r(0)=v \cdot t,$$

where v=r(l) — r(0). Thus the displacement r(t) — r(0)turns out to be a linear function of time in this simplest case, where the role of the constant of proportionality between the displacement r(t) — r(0)and the time t is played by the vector v that is the displacement in unit time. It is this vector that we call the velocity of uniform motion. The fact that the motion is rectilinear can be observen from the parametric representation of the trajectory: r(<=)=r(0)+ v • t, which is the equation of a straight line, as you will recall from analytic geometry.

We thus know the velocity v of uniform rectilinear motion given. By the law of inertia, if no external forces are acting on a body, it moves uniformly in a straight line. Hence if the action of M on m were to cease at time t, the latter would continue its motion, in a straight line at a certain velocity from that time on. It is natural to regard that velocity as the instantaneous velocity of the body at time t.

However, such a definition of instantaneous velocity would remain a pure abstraction, giving us no guidance for explicit computation of the quantity, if not for the circumstance of primary importance that we are about to discuss.

While remaining within the circle we have entered(logicians would call it a "vicious" circle)when we wrote down the equation of motion and then undertook to determine what is meant by instantaneous velocity and acceleration, we nevertheless remark that, even with the most general ideas about these concepts, one can draw the following heuristic conclusions. If there is no force, that is, F=0, then the acceleration is also zero. But if the rate of change a(<=)of the velocity v(<=)is zero, then the velocity v(<=)itself must not vary over time. In that way, we arrive at the law of inertia, according to which the body indeed moves in space with a velocity that is constant in time.

From this same Equation we can observe that forces of bounded magnitude are capable of creating only accelerations of bounded magnitude. But if the absolute magnitude of the rate of change of a quantity P(t)over a time interval [0, t] does not exceed some constant c, then, in our picture of the situation, the change |P(t) — P(0)| in the quantity P over time t cannot exceed c • t, that is, in this situation, the quantity changes by very little in a small interval of time.(In any case, the function P(t)turns out to be continuous.)

Thus, in a real mechanical system the parameters change by small amounts over a small time interval.

In particular, at all times t close to some time to the velocity v(t)of the body m must be close to the value v(to)that we wish to determine. But in that case, in a small neighborhood of the time to the motion itself must differ by only a small amount from uniform motion at velocity v(<=q), and the closer to to, the less it differs.

If we photographed the trajectory of the body m through a telescope, depending on the power of the telescope The portion of the trajectory corresponds to a time interval so small that it is difficult to distinguish the actual trajectory from a straight line, since this portion of the trajectory really does resemble a straight line, and the motion resembles uniform rectilinear motion. From this observation, as it happens, we can conclude that by solving the problem of
determining the instantaneous velocity(velocity being a vector quantity)

will at the same time solve the purely geometric problem of defining and finding the tangent to a curve(in the present case the curve is the trajectory of motion).

Thus we have observed that in this problem we must have v(t)« v(<=o) for t close to to, that is, v(t)v(<=o)as t to, or, what is the same, v(t)=v(to)+ o(l)as t to- Then we must also have

r(t)- r(t0)« v(*0)•(t - t0)

for t close to to. More precisely, the value of the displacement r(t) — r(to)is equivalent to v(to)(t — to)as t to, or

r(t)- r(t0)=v(t0)(t - t0)+ o(v(t0)(t - t0)),

where o(v(to)(t — to)) is a correction vector whose magnitude tends to zero faster than the magnitude of the vector v(to)(t—to)as t to. Here, naturally, we must except the case when v(to)=0. So as not to exclude this case
from consideration in general, it is useful to observe that13 |v(<=o)(<= — ^o)|=|v(to)| 11 — to\. Thus, if |v(to)| 7^ 0, then the quantity \v(to)(t

Notes

— <=o)| is of the same order as \t — to \, and therefore o(v(to)(t — to))=o(t — to). Hence, instead of we can write the relation

$r(t) - r(t_0) = v(t_0)(t - t_0) + o(t - t_0),$

which does not exclude the case v(to)=0. Thus, starting from the most general, and perhaps vague ideas about
velocity, we have arrived at which the velocity must satisfy. But
the quantity v(to)can be found unambiguously

v('«»=<=?„ r('llt?o)•

Therefore both the fundamental relation and the relation equivalent to it can now be taken as the definition of the quantity v(to), the instantaneous velocity of the body at time

At this point we shall not allow ourselves to be distracted into a detailed discussion of the problem of the limit of a vector-valued function. Instead, we shall confine ourselves to reducing it to the case of the limit of a real valued function, which has already been discussed in compLete detail. Since the vector r(<=) — r(to)has coordinates(x(t) — x(to), y(t) — y(to)), we have in a position to answer the question whether a pair of functions(x(t), y(t)) can describe the motion of the body m about the body M. To answer this question, one must find x(t)and y(t)and check whether hold. The system is an example of a system of so-known differential equations. At this point we can only check whether a set of functions is a solution of the system. How to find the solution or, better expressed, how to investigate the
properties of solutions of differential equations is studied in a special and, as one can now appreciate, critical area of analysis - the theory of differential equations.

The operation of finding the rate of change of a vector quantity, as has been shown, reduces to finding the rates of change of several numerical-valued
functions - the coordinates of the vector. Thus we must first of all learn how to carry out this operation easily in the simplest case of real-valued functions of a real-valued argument, which we now take up.

7.4 FUNCTIONS DIFFERENTIABLE AT A POINT

We begin with two preliminary definitions that we shall shortly make precise.

Definition Oi. A function /: E R defined on a set E C R is differentiate at a point a <= E that is a limit point of E if there exists a linear function A -(x — a)of the increment x — a of the argument such that f(x) — f(a)can be represented as

f(x) — f(a)=A •(x — a)+ o(x — a)as x a, x <= E .

In other words, a function is differentiable at a point a if the change in its values in a neighborhood of the point in question is linear up to a correction that is infinitesimal compared with the magnitude of the displacement x — a from the point a.

Remark. As a rule we have to deal with functions defined in an entire neighborhood of the point in question, not merely on a subset of the neighborhood.

Definition. The linear function A •(x — a)in is known the differential of the function /at a. The differential of a function at a point is uniquely determined; for it follows from that

f(x) — f(a)..(. o(x — a)\ A

lim = lim A + }- J=A,

EBx-±a X — a E3x-±a \ X — a /

so that the number A is unambiguously determined due to the uniqueness of the limit.

Definition. The number

/(«)= Um MiiM

e^x—Yd x — a

is known the derivative of the function /at a.

Notes

Relation can be rewritten in the equivalent form

f(x)-f(a)

———=f(a)+a(x),

where a(x)0 as x a, x <= E, which in turn is equivalent to

f(x) — f(a)=f(a)(x — a)+ o(x — a)as x a, x <= E .

Thus, differentiability of a function at a point is equivalent to the existence of its derivative at the same point.

If we compare these definitions with what was said in Subsect., we can conclude that the derivative characterizes the rate of change of a function at the point under consideration, while the differential provides the best linear approximation to the increment of the function in a neighborhood of the same point.

If a function /: E R is differentiable at different points of the set E, then in passing from one point to another both the quantity A and the function o(x — a)in Eq.(5.9)can change(a result at which we have already arrived explicitly. This circumstance should be noted in the very definition of a differentiable function, and we now write out this fundamental definition in full.

Definition . A function /: E R defined on a set E C R is differentiable at a point x <= E that is a limit point of E if

f(x + h) — f(x)=A(x)h + ol{ x\ h),

where h h-* A(x)h is a linear function in h and a(x;h)=o(h)as h 0, x -j- h <= E.

The quantities

Ax{ h):=(x + h) — x=h

and

Af(x; h):= f(x + h)- f(x)

are known respectively the increment of the argument and the increment of the function(corresponding to this increment in the argument).

They are often denoted(not quite legitimately, to be sure)by the symbols Ax and Af(x)representing functions of h.

Thus, a function is differentiable at a point if its increment at that point, regarded as a function of the increment h in its argument, is linear up to a correction that is infinitesimal compared to h as h 0.

Definition. The function h h-* A(x)h of Definition which is linear in /i, is known the differential of the function f : E at the point x G E and is denoted d f(x)or Df(x).

Thus, df{ x){ h)=A{ x)h.

Prom Definitions we have

Af(x; h)- df(x)(h)=a(x; h),

and a(x;h)=o(h)as h 0, x + h G E; that is, the difference between the increment of the function due to the increment h in its argument and the value of the function df(x), which is linear in /i, at the same /i, is an infinitesimal of higher order than the first in h.

For that reason, we say that the differential is the(principal)linear part of the increment of the function.

As follows from relation and Definition,

A(*)= /'(*)=lim +

/i->o h

x+h, x<=E

and so the differential can be written as

d f(x)(h)=f(x)h .

In particular, if f(x)=x, we obviously have f'(x)=1 and

dx{ h)=1 • h=h,

Notes

so that it is sometimes said that "the differential of an independent variable equals its increment".

Taking this equality into account, we deduce from that

df(x)(h)=f'(x)dx(h),

that is,

d /Oz)=f(x)dx .

The equality should be understood as the equality of two functions of h. we obtain

that is, the function(the ratio of the functions d f(x)and dx)is constant and equals f(x). For this reason, following Leibniz, we frequently denote the derivative by the symbol, alongside the notation f'(x)proposed by Lagrange.

In mechanics, in addition to these symbols, the symbol ip(t)(read "phi-dot of t")is also used to denote the derivative of the function ip(t)with respect to time t.

7.5 THE TANGENT LINE; GEOMETRIC MEANING OF THE DERIVATIVE AND DIFFERENTIAL

Let /: E R be a function defined on a set E C R and xq a given limit point of E. We wish to choose the constant cq so as to give the best possible description of the behavior of the function in a neighborhood of the point xq among constant functions. More precisely, we want the difference f(x) — Co to be infinitesimal compared with any nonzero constant as x xo, x <= E, that is

f(x)=Co + o(l)as x -> xo, x <= E .

This last relation is equivalent to saying lim f(x)=cq. If, in particular the function is continuous at xo, then lim f(x)=f(xo), and naturally

EBx-±x o

Co=f(x o).

Now Let us try to choose the function Co + c\{ x — xq)so as to have

f(x)=Co + ci(x — xo)+ o(x — xo)as x —> xo, x e E .

This is obviously a generalization of the preceding problem, since the formula can be rewritten as

f(x)=Co + o((x — xo)°)as x -> xo, x e E .

It follows immediately from that Co=lim f(x), and if the

EBx—tx o

function is continuous at this point, then Co=f(xo).

If cq has been found, it then follows from that

oi — iim .

E3x—^xo X — Xq

And, in general, if we were observeking a polynomial

Pn(xo; x)=Co + c \(x — xo)+

V cn{ x — xq)71 such that

f(x)=Co + ci(x - xo)H

h cn(x - x0)n + o((x - x0)n)

as x x0, x <= E,

we would find successively, with no ambiguity that

Co=lim f(x),

E3x—>x o

ci=*im

EBx-±x o x x°

^ /(x)-[coH hcn_i(x-x0)n x]

Notes

Cn — ^ lim (rr_T\n >

<=9x-»x0 ^

assuming that all these limits exist. Otherwise condition cannot be fulfilled, and the problem has no solution.

If the function /is continuous at xo, it follows from, as already pointed out, that Co=f(xo), and we then arrive at the relation

f(x) — f(x0)=ci(x — x0)+ o(x — xo)as x x0, x G E,

which is equivalent to the condition that f(x)be differentiable at xq.

Prom this we find

C1= lim

E3x->x o X — Xq

We have thus proved the following proposition.

Proposition . A function f : E —>]R that is continuous at a point xq G E that is a limit point of E Cl admits a linear approximation if and only if it is differentiable at the point

The function

(p(x)=Co + ci(x - x0)

with Co=f(xo)and c\=f'{ xo)is the only function of the form.

Thus the function

ip(x)=f(xo)+ f'(xo)(x - x0)

provides the best linear approximation to the function /in a neighborhood of xo in the sense that for any other function <p(x)of the form we have f(x) — if(x)7^ o(x — xo)as x xq, x E E.

The graph of the function is the straight line

y-f(xo)=f'(x0)(x-x0),

passing through the point(xo, f(xo)) and having slope f'(xo).

Since the line provides the optimal linear approximation of the graph of the function y=f(x)in a neighborhood of the point(xo, f(xo)), it is natural to make the following definition.

Definition . If a function f : E M is defined on a set E C R and differentiable at a point xo G the line defined is known the tangent to the graph of this function at the point(xo, f(xo)) illustrates all the basic concepts we have so far introduced in connection with differentiability of a function at a point: the increment of the argument, the increment of the function corresponding to it, and the value of the differential. The figure shows the graph of the function, the tangent to the graph at the point Po=(xo, f(xo))> and for comparison, an arbitrary line(usually known a secant)passing through Po and some point P ^ Po of the graph of the function.

f{ x o + h)

f(x o)

Xo Xo + h X

Definition . If the mappings f : E R and g : E —>■ R are continuous at a point xo € E that is a limit point of E and f(x) — g(x)=o((x — xo)n) as x —>■ Xo, x <= P, we say that /and g have nth order contact at xo(more precisely, contact of order at least n).

For n=1 we say that the mappings /and g are tangent to each other at

xo.

According to Definition the mapping is tangent at Xo to a mapping f : E R that is differentiable at that point.

We can now also say that the polynomial Pn(xo; x)=Co + ci(x — xo)+ - + cn(x — x0)n of relation has contact of order at least n with the function /.

The number h=x — xo, that is, the increment of the argument, can be regarded as a vector attached to the point Xq and defining the transition from #o to x=#o + h. We denote the set of all such vectors by TR(#o)or TTHXo Similarly, we denote by TR(? /o)or TRyo the set of all displacement vectors from the point yo along the y-axis. It can then be observen from the definition of the differential that the mapping

d /(*o): TR(xo)TR(/(®o)),

defined by the differential h i—>• f'(xo)h=df(xo)(h)is tangent to the mapping

h i-> f(xo + h)- f(x0)=Af(x0; h),

defined by the increment of a differentiable function.

ordinate of the graph of the function y=f(x)as the argument passes from #o to #o + then the differential gives the increment in the ordinate of the tangent to the graph of the function for the same increment h in the argument.

7.6 THE ROLE OF THE COORDINATE SYSTEM

The analytic definition of a tangent can be the cause of some vague uneasiness. We shall try to state what it is exactly that makes one uneasy. However, we shall first point out a more geometric construction of the tangent to a curve at one of its points Po.

Take an arbitrary point P of the curve different from Po. The line determined by the pair of points Po and P, as already noted, is known a secant in relation to the curve. We now force the point P to approach Po along the curve. If the secant tends to some limiting position as we do so, that limiting position of the secant is the tangent to the curve at Po- Despite its intuitive nature, such a definition of the tangent is not available to us at the moment, since we do not know what a curve is, what it means to say that "a point tends to another point along a curve", and finally, in what sense we are to interpret the phrase "limiting position of the secant".

Rather than make all these concepts precise, we point out a fundamental difference between the two definitions of tangent that we have introduced. The second was purely geometric, unconnected(at least until it is made more precise)with any coordinate system. In the first case, however, we have defined the tangent to a curve that is the graph of a differentiate function in some coordinate system. The question naturally arises whether, if the curve is written in a different coordinate system, it might not cease to be different iable, or might be different iable but yield a different line as tangent when the computations are carried out in the new coordinates.

This question of invariance, that is, independence of the coordinate system, always arises when a concept is introduced using a coordinate system.

The question applies in equal measure to the concept of velocity, which we discussed in Subsect. and which as we have mentioned already, includes the concept of a tangent.

Points, vectors, lines, and so forth have different numerical characteris tics in different coordinate systems(coordinates of a point, coordinates of a vector, equation of a line). However, knowing the formulas that connect two coordinate systems, one can always determine from two numerical representations of the same type whether or not they are expressions for the same geometric object in different coordinate systems. Intuition suggests that the
procedure for defining velocity described in Subsect. leads to the same vector independently of the coordinate system in which the computations are carried out. At the appropriate time in the study of functions of several variables we shall give a detailed discussion of questions of this sort. The invariance of the definition of velocity with respect to different coordinate systems will be verified in the next section.

Before passing to the study of specific examples, we now summarize some of the results.

We have encountered the problem of the describing mathematically the instantaneous velocity of a moving body.

This problem led us to the problem of approximating a given function in the neighborhood of a given point by a linear function, which on the geometric level led to the concept of the tangent Functions describing the motion of a real mechanical system are letd to admit such a linear approximation.

In this way we have distinguished the class of differentiable functions in the class of all functions.

The concept of the differential of a function at a point has been introduced. The differential is a linear mapping defined on displacements from the point under consideration that describes the behavior of the increment of a differentiable function in a neighborhood of the point, up to a quantity that is infinitesimal in comparison with the displacement.

The differential df(xo)h=f(xo)h is compLetely determined by the number f'{ xo), the derivative of the function /at Xo, which can be found by taking the limit

rM=Bm m-fM

E3x-±xq X — Xq

The physical meaning of the derivative is the rate of change of the quantity /(x)at time xo; its geometrical meaning is the slope of the tangent to the graph of the function y=f(x)at the point(xo, /(#o))-

Some Examples

Example. Let f(x)=sinx. We shall show that f'{ x)=cosx.

Proof.

s\n(x + h)-smx 2 sin(4)cos(x + 4)
lim — /=lim— =

h—>-0 h h—>-0 h

1.

f h\ .. Sin(I)

= lim cos [x -f - • lim —, =cos x . □

h—*0 V 2/h—*0(|)

Here we have used the theorem on the limit of a product, the continuity of the function cosx, the equivalence sint ~ t as t —>■ 0, and the theorem on
the limit of a composite function.

Example. We shall show that cos' x= — sin x.

Proof.

cos(x + h)-cosx —2 sin(^)sin(x -t- j)

lim —-=lim

h—>-0

ft

h—>-0

h

.. . /h\ .. sin(|).
= — lim sin [x -f - • lim —, ,\= — sin x . □
h—>0 V 2/ /i—>o(|)

Example. We shall show that if f{ t)=r cos art, then /'(<=)= — no sin out.
Proof. .. r cosouit -f /i) — r cosout .. —2sin(^)sinu;(<=-f §)

lim = r lim— =

h—>-0

Notes

h

h—>-0 h

. f h \ sin(^) .

= — ruo lim sinw[t -f — • lim —/\ f= —ruosmuot. □

h^o \ 2 J h->o(^)

Example. The instantaneous velocity and instantaneous acceleration of a point mass. Suppose a point mass is moving in a plane and that in some given coordinate system its motion is described by differentiate functions of

time

X=X(t), y=y(t)

or, what is the same, by a vector

r(t)=(x(t), y(t)) .

As we have explained in Subsect., the velocity of the point at time t is the vector

v(t)=r(i)=(x(t), y(t)),

where x(t)and y(t)are the derivatives of x(t)and y(t)with respect to time t.

The acceleration a(t)is the rate of change of the vector v(<=), so that

a(<)=v(t)=r(i)=(x(t), y(t)),

where x(t)and y(t)are the derivatives of the functions x(t)and y(t)with respect to time, the so-known second derivatives of x(t)and y(t).

Thus, in the sense of the physical problem, functions x(t)and y(t)that describe the motion of a point mass must have both first and second derivatives.

In particular, Let us consider the uniform motion of a point along a circle of radius r. Let u be the angular velocity of the point, that is, the magnitude

of the central angle over which the point moves in unit time.

In Cartesian coordinates(by the definitions of the functions cos# and sinx)this motion is written in the form

r(t)=(rcos(ut + a), rsin(ut + a)),

and if r(0)=(r, 0), it lets the form

r(t)=(r cos ut, r sin ut).

Without loss of generality in our subsequent deductions, for the sake of brevity, we shall let that r(0)=(r, 0).

v(t)=r(<=)=(—ru sin ut, ru cos ut).

From the computation of the inner product

(v(t), r(t))= —r2u sinu)t cosut + r2cu cos ut sinut=0,

as one should expect in this case, we find that the velocity vector v(t)is or-

thogonal to the radius-vector r(<=)and is therefore directed along the tangent

to the circle.

Next, for the acceleration, we have

a(<=)=v(t)=r(t)=(—rev2 cos ut, —rev2 sinu;<=),

that is, a(t)= — u;2r(<=), and the accleration is thus indeed centripetal, since

it has the direction opposite to that of the radius-vector r(<=).

Moreover,

|a(*)|=w2|r(i)|=w2r =   = —,

r   r

where v=|v(t)|.

Starting from these formulas, Let us compute, for example, the speed of a low-altitude satellite of the Earth. In this case r equals the radius of the earth,

that is, r=6400 km, while |a(<=)|=#, where g « 10m /s2 is the acceleration of free fall at the surface of the earth.

Thus, v2=|a(<=)|r « 10m /s2 x 64 • 105m=64 • 106(m /s)2, and so

v « 8 • 103 m /s.

Example. The optic property of a parabolic mirror. Let us consider the parabola y=^x2(p > 0, observe Fig. 5.4), and construct the tangent to it at

the point(x0, yo)=(xo, ^«o)-

Since f(x)=^x2, we have

f'(x0)=lim ^ lim(x + x0)=-x0 .

x^-xo X — Xq Zp x^xq p

Hence the required tangent has the equation

V~Y*o=~* o(x-xo)

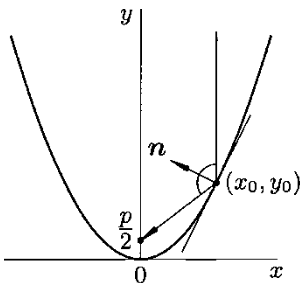
2p p

or

-x0(x - Xo)-(y - yo)=0,

P

where y0=^x%.



The vector n = as can be observen from this last equation, is

orthogonal to the line whose equation. We shall show that the vectors ey=(0, 1)and e /=(— #o, § ~~ 2 /o)form equal angles with n. The vector ey
is a unit vector directed along the ? /-axis, while e/is directed from the point
of tangency(#o, 2 /o)=(xo? ^xo)to the point(0j f)' which is the focus of the parabola. Thus we have shown that a wave source located at the point(0, |), the
focus of the parabola, will emit a ray parallel to the axis of the mirror(the y-axis), and that a wave arriving parallel to the axis of the mirror will pass
through the focus.

Example. With this example we shall show that the tangent is merely the best linear approximation to the graph of a function in a neighborhood of the
point of tangency and does not necessarily have only one point in common
with the curve, as was the case with a circle, or in general, with convex curves.
(For convex curves we shall give a separate discussion.)

Let the function be given by

(x2 sin i, if X ^ 0,

/(*)={

[0, if x=0 .

The graph of this function is shown by the thick line

Notes

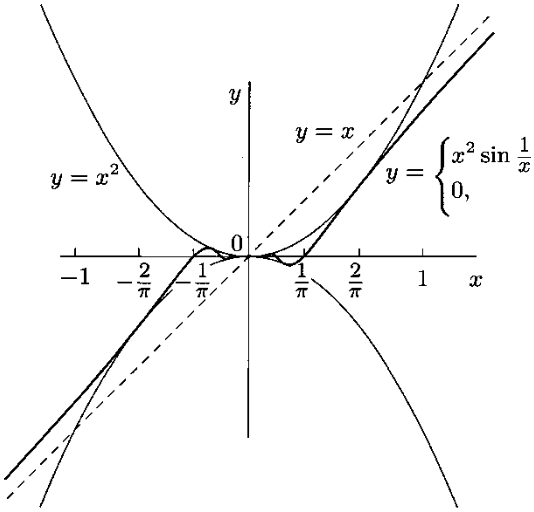


Fig. 5.5.

Let us find the tangent to the graph at the point(0, 0). Since

»/f^.\ x2sin<=—0 .. .1 _

/(0)=lim = lim x sm —=0,

a—>o x - 0 z—>o x

the tangent has the equation y — 0=0 •(x — 0), or simply y=0.

Thus, in this example the tangent is the x-axis, which the graph intersects infinitely many times in any neighborhood of the point of tangency.

By the definition of differentiability of a function /: E —>■ R at a point xo <= E, we have

/(x)- /(x0)=A(x0)(x - x0)+ o(x - x0)as x x0, x <= E .

Since the right-hand side of this equality tends to zero as x —xq, x e E, it follows that lim f{ x)=f{ xo), so that a function that is differentiate

E3x^xq

at a point is necessarily continuous at that point.

We shall show that the converse, of course, is not always true.

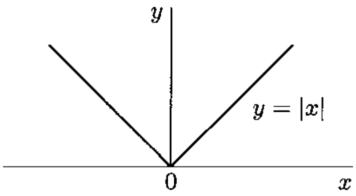
Example. Let f{ x)=|x|, Then at the point xo=0 we have
|im <=HW=lim !f^o= lim =

x—^xo—0 X — Xo x->-0 x — 0 x—>—0 x

llm /w - fM=lim = 1|m x =!

x-txo+0 X — Xo a:->+0 x — 0 z->+0 x

Consequently, at this point the function has no derivative and hence is not differentiate at the point.



Example. We shall show that ex+h — ex=exh + o(h)as h —>■ 0.

Thus, the function exp(x)=e* is differentiate and dexp(x)h=exp(x)/i, or de*=exdx, and therefore exp'x=expx, or=ex.

Proof.

ex+h _ex= ex(eh _l- }=ex(h + = exh +

Here we have used the formula eh — 1=h + o(h)

Example. If a > 0, then ax+h — ax=ah(\na)h + o(h)as h —>■ 0. Thus dax=ax(\na)dx and=ax In a.

Proof.

ax+h _ax=ax(ah _ J)=ax(ehlna _y =

= ax(/ilna + o(hlna))=ax(lna)/i + o(h)as h —>• 0 . □

Example. If x ^ 0, then In |a: + h\ — In |m|=+ o(h)as h —> 0. Thus dln|ar|=±dx and ±

Proof.

h

In \x + h\ — In b|=In |l H— .

x

Notes

For |ft| < \x\ we have |l -f=1 -f and so for sufficiently small values of h we can write

In \x + h\ — In |x|=In(l +=- + o(-)= —h + o{ h)
\ xJ x \xJ x

as h —>■ 0. Here we have used the relation ln(l -f1)=t -f o(t)as <= —>■ 0,

Example. If x /0 and 0 < a /1, then loga \x+h\ — loga |x|=^^h+o(h) as ft -»• 0. Thus, d loga |x|=^dz and dl°f<=N =

Proof.

loga |« + ^| - loga M=^ga |l +=loga(* + =

= r— ln(i + -)=7~■("+°(~))= —f-h + o(h).
ma V x/ma\x \xJJ xma

Here we have used the formula for transition from one base of logarithms

Check your Progress-1

Discuss A Sufficient Condition For A Constrained Extremum

Discuss Functions Differentiable At A Poin

7.7 LET US SUM UP

In this unit we have discussed the definition and example of A Sufficient Condition For A Constrained Extremum, Differential Calculus, Differentiable Functions, Functions Differentiable At A Point, The Tangent Line; Geometric Meaning Of The Derivative And Differential, The Role Of The Coordinate System

7.8 KEYWORDS

1. Differential Calculus, Differentiable Functions   Statement of the Problem and Introductory Considerations.

2. Functions Differentiable At A Point: We begin with two preliminary definitions that we shall shortly make precise.

3.The Tangent Line; Geometric Meaning Of The Derivative And Differential Let /: E R be a function defined on a set E C R and xq a given limit point of E.

4. The Role Of The Coordinate System : The analytic definition of a tangent can be the cause of some vague uneasiness. We shall try to state what it is exactly that makes one uneasy. However, we shall first point out a more geometric construction of the tangent to a curve at one of its points Po.

7.9 QUESTIONS FOR REVIEW

Explain A Sufficient Condition For A Constrained Extremum

Explain Functions Differentiable At A Point

7.10 ANSWERS TO CHECK YOUR PROGRESS

A Sufficient Condition For A Constrained Extremum

(answer for Check your Progress-1 Q)

Functions Differentiable At A Point

(answer for Check your Progress-1 Q)

7.11 REFERENCES

- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables