CHAPTER II
DIFFERENTATION IN SEVERAL VARIABLES

Department of Foundation Year,

Institute of Technology of Cambodia

2014–2015

Contents

Contents

1. Introduction to Functions of Several Variables
2. Limits and Continuity
3. Partial Derivatives
4. Differentials
5. Chain rules for functions of several variables
6. Directional Derivatives and Gradients
7. Tangent planes and normal lines
8. Extrema of Functions of Several Variables
9. Lagrange Multipliers

Introduction to Functions of Several Variables

Definition 1

Let D be a nonempty set and that for each element x in D there corresponds a unique value $y = f(x)$ in R, then f is called a **function of** x. The set D is the **domain** of f, and R is the **range** of f. We write

$$f : D \longrightarrow R; x \longmapsto f(x).$$

In this case x is called **independent variable** and y is called **dependent variable**.

Extrema of Functions of Several Variables

Definition 52

Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$ be a function.

- The set D is said to be **convex** if every two points $x, y \in D$, the line segment $L(x, y) \subset D$.
- The function f is said to be **convex** on a convex set D if for any $x, y \in D$ and for any $t \in [0,1], f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$.
- The function f is said to be **concave** on a convex set D if for any $x, y \in D$ and for any $t \in [0,1], f((1-t)x + ty) \geq (1-t)f(x) + tf(y)$.

Theorem 53

Let $D \subset \mathbb{R}^n$ be a convex set and $f : D \to \mathbb{R}$ be a C^2 function. The function f is a convex function on D if and only if for any $x \in D$, for any $k = 1, 2, \ldots, n$, the kth principal minor of the Hessian $H_k(x) \geq 0$.

Extrema of Functions of Several Variables

Theorem 54

If $f : D \to \mathbb{R}$ is a convex function on a convex set D and $f(a)$ is a local minimum then $f(a)$ is the global minimum on D.
If $f : D \to \mathbb{R}$ is a concave function on a convex set D and $f(a)$ is a local maximum then $f(a)$ is the global maximum on D.

Contents

Lagrange Multipliers

Theorem 55

Let $f, g_1, \ldots, g_k : D \to \mathbb{R}$ be C^1 functions where $D \subset \mathbb{R}^n$ is open and $k < n$. Suppose there is an $a \in D$ such that

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_k}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial x_1}(a) & \cdots & \frac{\partial g_k}{\partial x_k}(a) \end{pmatrix} \neq 0.$$

If $f(a)$ is local extremum of f subject to the constraints $g_i(x) = c_i$ for $i = 1, 2, \ldots, k$, then there exist scalars $\lambda_1, \ldots, \lambda_k$ (called **Lagrange Multipliers***) such that*

$$\nabla f(a) + \sum_{i=1}^{k} \lambda_i \nabla g_i(a) = 0.$$

Extrema of Functions of Several Variables

Theorem 46 (A necessary condition for extremum)

If f has a local extremum at a on an open region D, then a is a critical point of f.

Note that the converse of the theorem above is not true in general. That is, a critical point does not yield a local extremum.

Definition 47 (Quadratic form)

Let $b_{ij} \in \mathbb{R}$ such that $b_{ij} = b_{ji}$ and $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$.

1. A **quadratic form** in $h_1, \ldots, h_n$ is a function defined by

$$Q(h_1, \ldots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} h_i h_j.$$

This quadratic form can be written in term of matrices as

$$Q(h) = (h_1 \ldots h_n) \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = h^t B h$$

where $h^t = \begin{pmatrix} h_1 & \ldots & h_n \end{pmatrix}$ and $B = (b_{ij})_{n \times n}$ is a symmetric matrix.

2. The quadratic form Q (and also symmetric matrix B) is said to be **positive definite** if $Q(h) > 0$ for all $h \neq 0$ and **negative definite** if $Q(h) < 0$ for all $h \neq 0$.

Extrema of Functions of Several Variables

Theorem 48

Let $A = (a_{ij})_{n \times n}$ be a symmetric matrix. Then the matrix A is positive definite if and only if all kth principal minors $A_k > 0$ for $k = 1, 2, \ldots, n$.

Theorem 49

Let$A = (a_{ij})_{n \times n}$ be a symmetric matrix. Then A is positive definite if and only if it can be reduced to upper triangular form using only elementary row operations $E_{i,j}(\lambda)$ and the diagonal elements of resulting matrix are greater than zero.

Extrema of Functions of Several Variables

Theorem 50

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$ be a function. If all second-order partial derivatives of f exist at $a \in D$ and $d^{(2)} f_a(h) > 0$ for all $h \neq 0$, then there is a $\lambda > 0$ such that $d^{(2)} f_a(x) \geq \lambda \|x\|^2$ for all $x \in \mathbb{R}^n$.

Theorem 51 (The Second Derivative Test)

Let $D \subset \mathbb{R}^n$ be open containing a and $f : D \to \mathbb{R}$ satisfy $\nabla f(a) = 0$. Suppose further that all second-order partial derivatives of f exist on D and continuous at a.

- *If the quadratic form $Q = d^{(2)} f_a(h)$ is positive definite, then $f(a)$ is a local minimum of f.*
- *If the quadratic form $Q = d^{(2)} f_a(h)$ is negative definite, then $f(a)$ is a local maximum of f.*
- *If the quadratic form $Q = d^{(2)} f_a(h)$ takes on both positive and negative values for $h \in \mathbb{R}^n$, then a is a saddle point of f.*

Contents

Extrema of Functions of Several Variables

Definition 41

Let $D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ be a function and $a \in D$.

- $f(a)$ is called a **local minimum** of f if there is an $r > 0$ such that $f(a) \leq f(x)$ for all $x \in B_r(a) \cap D$.
- $f(a)$ is called a **local maximum** of f if there is an $r > 0$ such that $f(a) \geq f(x)$ for all $x \in B_r(a) \cap D$.
- $f(a)$ is called a **local extremum** of f if there $f(a)$ is a local minimum or a local maximum of f.
- $f(a)$ is called a **global minimum** of f on D if $f(a) \leq f(x)$ for all $x \in D$.
- $f(a)$ is called a **global maximum** of f on D if $f(a) \geq f(x)$ for all $x \in D$.
- $f(a)$ is called a **global extremum** of f on D if $f(a)$ is a global minimum or a global maximum of f.

Extrema of Functions of Several Variables

Theorem 42

If $D \subset \mathbb{R}^n$ is closed and bounded and $f : D \to \mathbb{R}$ is a continuous function, then f must have both a global maximum and a global minimum somewhere on D.

Theorem 43

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$ be a function. If the first-order partial derivatives of f exist at $a \in D$ and $f(a)$ is a local extremum of f, then $\nabla f(a) = 0$.

Extrema of Functions of Several Variables

Definition 44

Let f be defined on an open region D containing a. The point a is a **critical point** of f if one of the following is true.

1. $f_{x_i}(a) = 0$, for all $i = 1, 2, \ldots, n$.
2. $f_{x_i}(a)$ does not exist for some $i \in \{1, 2, \ldots, n\}$.

Definition 45

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$ be differentiable at $a \in D$. Then a is called a **saddle point** of f if $\nabla f(a) = 0$ and there is an $r_0 > 0$ such that given any $0 < \rho < r_0$ there are points $x, y \in B_\rho(a)$ which satisfy $f(x) < f(a) < f(y)$.

Tangent planes and normal lines

So far, you have represented surfaces in space primarily by equations of the form

$$z = f(x, y)$$

In the development to follow, however, it is convenient to use the more general representation $F(x, y, z) = 0$. For a surface S given by $z = f(x, y)$, you can convert to the general form by defining F as

$$F(x, y, z) = f(x, y) - z.$$

Because $f(x, y) - z = 0$, you can consider S to be the level surface of F given by

$$F(x, y, z) = 0.$$

Tangent planes and normal lines

Let S be a surface given by $F(x, y, z) = 0$ and let $P(x_0, y_0, z_0)$ be a point on S. Let C be a curve S on through P that is defined by the vector-valued function

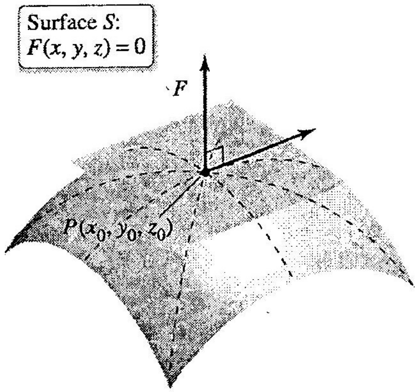
$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$$

Then, for all t

$$F(x(t), y(t), z(t)) = 0.$$

If F is differentiable and $x'(t), y'(t)$ and $z'(t)$ all exist, then we have

$$F'(t) = F_x(x, y, z)x'(t) + F_y(x, y, z)y'(t) + F_z(x, y, z)z'(t) = 0$$



Tangent planes and normal lines

At point $P(x_0, y_0, z_0)$, the equivalent vector form is

$$\nabla F(x_0, y_0, z_0).\mathbf{r}'(t_0) = 0.$$

This result means that the gradient at P is orthogonal to the tangent vector of every curve on S through P. So, all tangent lines on S lie in a plane that is normal to $\nabla F(x_{0,0}, z_0)$ and contains P, as shown in the Figure.

Definition 38

Let F be differentiable at the point $P(x_0, y_0, z_0)$ on the surface given by $F(x, y, z) = 0$ such that $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$.

- The plane through P that is normal to $\nabla F(x_0, y_0, z_0)$ is called the **tangent plane to S at P.**
- The line through P having the direction of $\nabla F(x_0, y_0, z_0)$ is called the **normal line to S at P.**

Tangent planes and normal lines

Theorem 39 (Equation of Tangent plane)

If F is differentiable at (x_0, y_0, z_0) then an equation of the tangent plane to the surface given by $F(x, y, z) = 0$ at (x_0, y_0, z_0) is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

Theorem 40

If the surface given by equation $z = f(x, y)$, then an equation of tangent line to the the surface at the point (x_0, y_0, z_0) is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0.$$

Hessian Matrix

Definition 34 (Principal Minor)

Let $A = (a_{ij})_{n\times n}$ be a square matrix. The determinant $A_k = \det(a_{ij})_{k\times k}$ is called the kth principal minor of the $n\times n$ matrix.

Definition 35 (The Hessian of a Function)

Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$ be a function having second-order partial derivatives $\frac{\partial^2 f}{\partial x_j \partial x_i}$. The **Hessian** of f is the matrix whose (i,j) entry is $f_{x_i x_j}$. That is,

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)_{n\times n}$$

We call $H_s = \det\left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)_{s\times s}$, sth principal minor of H_f for $s = 1, 2, \ldots, n$.

Higher Order Differential

Definition 36 (Higher Order Differential)

Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$ be a function. Suppose that the partial derivatives of f order $k-1$ exist on D. If each $(k-1)$th order partial derivative of f is differentiable at $a \in D$. Let $h = (h_1, \ldots, h_n)$. We call the kth **differential of** f **at** a is the expression

$$d^{(k)} f_a(h) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a) h_{i_1} \ldots h_{i_k}.$$

Theorem 37 (Taylor's formula)

Let $D \subset \mathbb{R}^n$ be open, $a, x \in D$, and $f : D \to \mathbb{R}$ be a function, and suppose that the partial derivatives of f order $k-1$ exist on D. If each $(k-1)$th order partial derivative of f is differentiable on D and the line segment $L(a,x) = \{(1-t)a + tx, 0 \le t \le 1\} \subset D$, then there is a point $c \in L(a,x)$ such that

$$f(x) = f(a) + \sum_{j=1}^{k-1} \frac{1}{j!} d^{(j)} f_a(h) + \frac{1}{k!} d^{(k)} f_c(h)$$

$$f(x) = f(a) + df_a(h) + \frac{1}{2} h^t H_f(a) h + \sum_{j=3}^{k-1} \frac{1}{j!} d^{(j)} f_a(h) + R_k$$

where $d^{(j)} f_a(h) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(a) h_{i_1} \ldots h_{i_j}$,
$R_k = d^{(k)} f_c(h) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_k}}(c) h_{i_1} \ldots h_{i_k}$ *and*
$h = (h_1, \ldots, h_n) = x - a$.

Contents

Differential fo Vector-Valued Functions

Definition 29

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function of n variables. Let $x = (x_1, \ldots, x_n)$ denote a point of $\mathbb{R}^n$ and $f = (f_1, \ldots, f_m)$. We define the **matrix of partial derivatives of** f, denoted Df, to be the $m \times n$ matrix whose (i, j) entry is $\frac{\partial f_i}{\partial x_j}$. That is, $Df = \left(\frac{\partial f_i}{\partial x_j}\right)$. The matrix $Df(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)$ is also called **Jacobian matrix of** f **at** a.

Differential fo Vector-Valued Functions

Theorem 30

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^m$ be a vector-valued function of $x = (x_1, \ldots, x_n)$. If $f = (f_1, \ldots, f_m)$ is differentiable at a then the first-order partial derivatives of f exist at a and the differential of f at a is

$$df\,(dx) = Df(a)(dx) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}$$

For short,

$$df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}.$$

Differential fo Vector-Valued Functions

Theorem 31

If $f, g : D \to \mathbb{R}^m$ are differentiable at a then

$$d(f + \alpha g)(a) = df(a) + \alpha dg(a)$$

where α is a constant.

Theorem 32

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^m$ be a function. If all first-order partial derivatives of f exist at a and are continuous at a, then the function f is differentiable at a.

Differential fo Vector-Valued Functions

Theorem 33 (Chain rule)

Let $D_1 \subset \mathbb{R}^n$ and $D_2 \subset \mathbb{R}^m$ be open. If $f : D_1 \to \mathbb{R}^m$ is differentiable at a and $g : D_2 \to \mathbb{R}^p$ is differentiable at $f(a) \in D_2$, then $k = g \circ f$ is differentiable at a and

$$d(g \circ f)(a) = dg\,(f(a))\,df(a).$$

If $M_{g \circ f}(a)$ is the Jacobian matrix of $g \circ f$ at a, $M_g\,(f(a))$ the Jacobian matrix of g at $f(a)$, and $M_f(a)$ is the Jacobian matrix of f at a, then

$$M_{g \circ f}(a) = M_g\,(f(a))\,M_f(a).$$

Gradient

Definition 23

For a real-valued function $f(x_1, x_2, \ldots, x_n)$, the **gradient of f at a point** a, denoted by $\nabla f(a)$, is the vector

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

Theorem 24

Let $D \subset \mathbb{R}^n$ be open, and suppose $f : D \to \mathbb{R}$ is differentiable at $a \in D$. Then the directional derivative of f at a exists for all directions (unit vectors) $\mathbf{u}$ and, moreover, we have

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

Gradient

Theorem 25

Let f be a continuously differentiable real-valued function, with $\nabla f \neq \mathbf{0}$. Then:

- *The value of $f(x)$ increases the fastest in the direction of ∇f. The maximum value of $D_{\mathbf{u}} f$ is $\|\nabla f\|$.*
- *The value of $f(x)$ decreases the fastest in the direction of $-\nabla f$. The minimum value of $D_{\mathbf{u}} f$ is $-\|\nabla f\|$.*

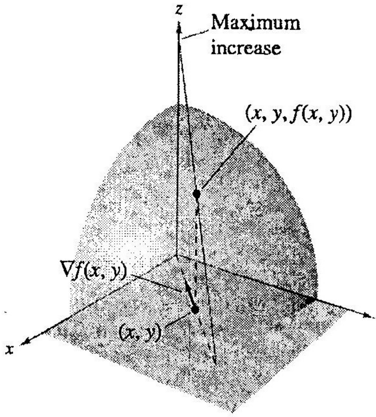


Figure: The maximum increase of f in the direction of $\nabla f(x, y)$ in the xy-plane

Gradient

Theorem 26

Let $D \subset \mathbb{R}^n$ be open, and $f : D \to \mathbb{R}$ be a function of class C^1. If a is a point on the level hypersurface $S = \{x \in D : f(x) = c\}$, then the vector $\nabla f(a)$ is perpendicular to S.

Differential fo Vector-Valued Functions

Definition 27

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}^m$ be a vector-valued function of n variables. Then f is said to be differentiable at a if there is a mapping $L_a : \mathbb{R}^n \to \mathbb{R}^m$ (called **the differential of f at** a denoted by $df_a = L_a$ or $df = L$ for short, that is $df(h) = L(h)$) such that

1. $L_a(\alpha x + y) = \alpha L_a(x) + L_a(y)$ for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.
2. $f(a + h) = f(a) + L_a(h) + o(\|h\|)$.

Theorem 28

If the function $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a then there is only one mapping L_a and f is continuous at a.

Chain rules for functions of several variables

Theorem 19 (Chain rules for function of several variables)

Let $y = f(x_1, x_2, \ldots, x_n)$, where f is differentiable function of $x_i, i = 1, 2, \ldots, n$. If each $x_i, i = 1, 2, \ldots, n$ is a differentiable function of m variables $t_1, t_2, \ldots, t_m$, then y is a differentiable function of $t_1, t_2, \ldots, t_m$ and

$$\frac{\partial y}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \frac{\partial x_i}{\partial t_j}.$$

for $j = 1, 2, \ldots, m$.
In particular, if $x_i, i = 1, 2, \ldots, n$ is a function of a single variable t, then we have

$$\frac{dy}{dt} = \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \frac{dx_i}{dt}.$$

Chain rules for functions of several variables

Theorem 20 (Chain rule: Implicit Differentiation)

If the equation $F(x_1, x_2, \ldots, x_n, y) = 0$ defines y implicitly as a differentiable function of $x_i, i = 1, 2, \ldots, n$, then

$$\frac{\partial y}{\partial x_i} = -\frac{F_{x_i}(x_1, x_2, \ldots, x_n, y)}{F_y(x_1, x_2, \ldots, x_n, y)}, \quad F_y(x_1, x_2, \ldots, x_n, y) \neq 0$$

for $i = 1, 2, \ldots, n$.

Contents

Directional Derivatives

Definition 21

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function and a be a point in $\mathbb{R}^n$. If $\mathbf{u} \in \mathbb{R}^n$ is any unit vector. Then, **the directional derivative of f at a in the direction of $\mathbf{u}$** is defined by

$$D_{\mathbf{u}} f = \lim_{t \to 0} \frac{f(a + t\mathbf{u}) - f(a)}{t}, \quad t \in \mathbb{R},$$

provided that the limit exists.

Theorem 22

Let $f : D \to \mathbb{R}$ and $a = (a_1, a_2, \ldots, a_n) \in D$. Suppose that the first partial derivatives of f exist and continue at a. Then the **directional derivative of** *f in the direction of a* **unit vector** *$\mathbf{u} = (u_1, u_2, \ldots, u_n)$ is given by*

$$D_{\mathbf{u}} f(a) = \sum_{i=1}^{n} u_i f_{x_i}(a).$$

Differentials

Definition 14

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$. The function f is said to be **differentiable at** a if there is a mapping $L_a : \mathbb{R}^n \to \mathbb{R}$ (called **differential of** f **at** a denoted by $df_a = L_a$ or $df = L$ for short, that is $df(h) = L(h)$), such that

1. $L_a(\alpha x + y) = \alpha L_a(x) + L_a(y)$ for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.
2. $f(a+h) = f(a) + L_a(h) + o(\|h\|)$.

Theorem 15

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$ be a function. If the function f is differentiable at $a \in D$ then there is only one mapping L_a and f is continuous at a.

Theorem 16

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$. If the function f is differentiable at a then all partial derivatives of function f exist at a and the differential of f at a is

$$df = L(h) = \frac{\partial f}{\partial x_1}(a)h_1 + \frac{\partial f}{\partial x_2}(a)h_2 + \cdots + \frac{\partial f}{\partial x_n}(a)h_n$$

where $h = (h_1, h_2, \ldots, h_n)$.

Note that if we denote the increments

$$h_i = \Delta x_i = dx_i, \quad i = 1, 2, \ldots, n$$

(called **the differential of** the independent variable $x_i, i = 1, 2, \ldots, n$, respectively), then

$$df = \frac{\partial f}{\partial x_1}(a)dx_1 + \frac{\partial f}{\partial x_2}(a)dx_2 + \cdots + \frac{\partial f}{\partial x_n}(a)dx_n$$

Differentials

and the increment $\Delta f = f(a + \Delta x) - f(a)$ of dependent variable f is

$$\Delta f = \frac{\partial f}{\partial x_1}(a)\Delta x_1 + \cdots + \frac{\partial f}{\partial x_n}(a)\Delta x_n + o(\|\Delta x\|).$$

If $\Delta x_i = dx_i, i = 1, 2, \ldots, n$ are small enough tending to zero, then Δf can be approximated by df.

Theorem 17

Let $D \subset \mathbb{R}^n$ be open and $f : D \to \mathbb{R}$. If all first-order partial derivatives of function f exist and are continuous at a, then the function f is differentiable at a.

Theorem 18

If a function is differentiable at a, then it is continuous at a.

Contents

Partial Derivatives

Definition 12

- A function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is said to be of **class** C^k if all its partial derivatives of order $\leq k$ are continuous.
- A function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is said to be of **class** C^∞ if f has continuous partial derivatives of all orders.
- A function $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be of class C^k if each of component functions is of class C^k.
- A function $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is said to be of class C^∞ if each of component functions is of class C^∞.

Partial Derivatives

- Note that if the function if $\dfrac{\partial f}{\partial x_i}$ has a partial derivative with respect to x_j, we denote the partial derivative by

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i} = f_{x_i x_j}.$$

- The function obtained by differentiating f successively with respect to $x_{i_1}, x_{i_2}, \dots x_{i_r}$ at x is denoted by

$$\frac{\partial^k f}{\partial x_{i_r} \partial x_{i_{r-1}} \dots \partial x_{i_1}} = f_{x_{i_1} x_{i_2} \dots x_{i_r}} \quad \text{where} \quad i_1 + \cdots + i_r = k.$$

It is called a k**th-order partial derivative of** f.

Partial Derivatives

Theorem 13

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be a C^k function. Then

$$f_{x_{i_1} x_{i_2} \dots x_{i_r}}(x) = f_{x_{j_1} x_{j_2} \dots x_{j_r}}(x)$$

where $i_1 + i_2 + \cdots + i_r = j_1 + j_2 + \cdots + j_r = k$.

Contents

Continuity

Theorem 9

Let $f : D \subset \mathbb{R}^n \longmapsto \mathbb{R}^m$, where $f = (f_1, \ldots, f_m)$. Then f is continuous at $a \in D$ (respectively f is continous on D) if and only if its component functions $f_i : D \to \mathbb{R}$ are all continuous at a.

Contents

Partial derivatives of a function of several variables

Definition 10

Let $f : D \subset \mathbb{R}^n \longmapsto \mathbb{R}$. If $y = f(x) = f(x_1, x_2, \ldots, x_n)$, then the **first partial derivative** of f with respect to $x_i, i \in \{1, 2, \ldots, n\}$, is defined by $f_{x_i}(x)$ or $\frac{\partial f}{\partial x_i}(x)$ and

$$f_{x_i}(x) = \frac{\partial}{\partial x_i} f(x) = \lim_{\Delta x_i \to 0} \frac{f(x_1, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, \ldots, x_n)}{\Delta x_i}$$

provided the limit exists.

Partial Derivatives

Theorem 11

Let f, g be two scalar-valued functions of n variables and let $x = (x_1, \ldots, x_n)$. If $f_{x_i}(x)$ and $g_{x_i}(x)$ exist, then

- $\dfrac{\partial(f + \lambda g)}{\partial x_i} x) = f_{x_i}(x) + \lambda g_{x_i}(x)$ *where λ is some constant*
- $\dfrac{\partial(fg)}{\partial x_i}(x) = f_{x_i}(x) g(x) + f(x) g_{x_i}(x)$
- $\dfrac{\partial(f/g)}{\partial x_i}(x) = \dfrac{f_{x_i}(x) g(x) - f(x) g_{x_i}(x)}{g^2(x)}$ *if $g(x) \neq 0$.*

Limit of a Function of Several Variables

Definition 3

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and $a \in \bar{D}$. Then we say that the limit of $f(x)$ equals L as x approaches a, written as

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

if given any $\epsilon > 0$, there exists a $\delta > 0$ such that

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

Theorem 4

If a limit exists, it is unique.

Limit of a Function of Several Variables

Theorem 5

Suppose that $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ both exist and that k is a scalar. Then

- $\lim_{x \to a}[f(x) \pm g(x)] = [\lim_{x \to a} f(x)] \pm [\lim_{x \to a} g(x)]$
- $\lim_{x \to a}[kf(x)] = k[\lim_{x \to a} f(x)]$
- $\lim_{x \to a}[f(x)g(x)] = [\lim_{x \to a} f(x)][\lim_{x \to a} g(x)]$
- $\lim_{x \to a}[f(x)/g(x)] = [\lim_{x \to a} f(x)]/[\lim_{x \to a} g(x)]$ *provided* $\lim_{x \to a} g(x) \neq 0$ *and both f and g are real-valued functions.*
- *If $f(x) \leq g(x)$ for all x, then $\lim_{x \to a} f(x) \leq \lim_{x \to a} g(x)$, where f and g are real-valued functions.*
- *If $\| \lim_{x \to a} f(x) - L\| \leq g(x)$ for all x and if $\lim_{x \to a} g(x) = 0$, then $\lim_{x \to a} f(x) = L$.*

Limit of a Function of Several Variables

Note

To show that the limit does not exist, we need to show that the function approaches different values as x approaches a along different paths in $\mathbb{R}^n$.

Theorem 6

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be a vector-value functions, $f = (f_1, f_2, \ldots, f_m)$ and $L = (L_1, L_2, \ldots, L_m)$. Then $\lim_{x \to a} f(x) = L$ if and only if $\lim_{x \to a} f_i(x) = L_i$ for $i = 1, 2, \ldots, m$.

Continuity

Definition 7

Let $f : D \subset \mathbb{R}^n \longmapsto \mathbb{R}^m$. We say that the function f is continuous at a point a in D if

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

We say that f is a continuous function on D if it is continuous at every point in its domain D.

Theorem 8 (Algebraic properties)

Let $f, g : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be continuous vector-value function and let $\alpha \in \mathbb{R}$ be a scalar. Then

- *$f + \alpha g$ and fg are continuous.*
- *If both f and g are real-valued functions and if $g(x) \neq 0$, then f/g is continuous.*

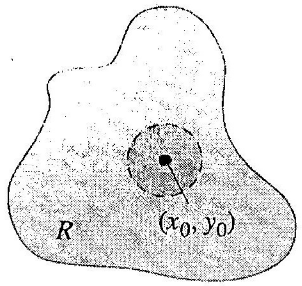


Figure: Interior point

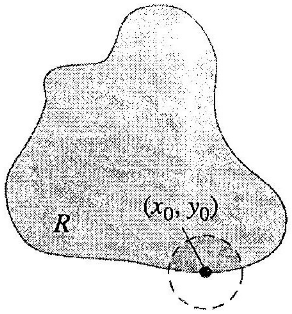


Figure: Boundary point

- A region R is **open** if it is a subset of its interior. That is, $R \subset \overset{\circ}{R}$
- A region R is **closed** if it contains its entire boundary. That is, $\partial R \subset R$.
- The closure of R is denoted by $\overline{R}$ and defined by

$$\overline{R} = \overset{\circ}{R} \cup \partial R$$

Level curve, level surface and level hypersurfce

Definition 2 (Level curve, level surface and level hypersurfce)

The set of points $(x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$ where a function of n independent variables has a constant value $f(x_1, x_2, \ldots, x_n) = c$ is called a **level hypersurface** of f.
In particular,

- the set of points in the plane where a function $f(x, y)$ has a constant value $f(x, y) = c$ is called a **level curve** of f.
- the set of points in the space where a function $f(x, y, z)$ has a constant value $f(x, y, z) = c$ is called a **level surface** of f.
- Note that if $n \geq 4$, the set of points satisfying the equation $f(x_1, x_2, \ldots, x_n) = c$ is called **level hypersurface**.

Example of level curves

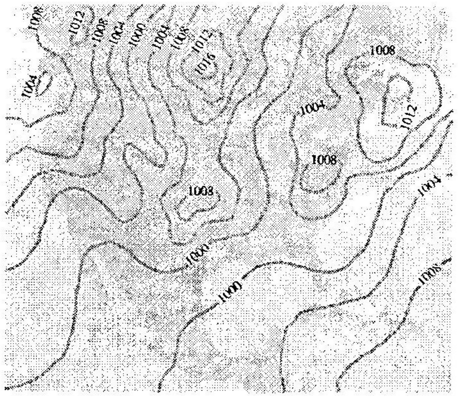


Figure: Level curves show the lines of equal pressure (isobars) measured in millibars

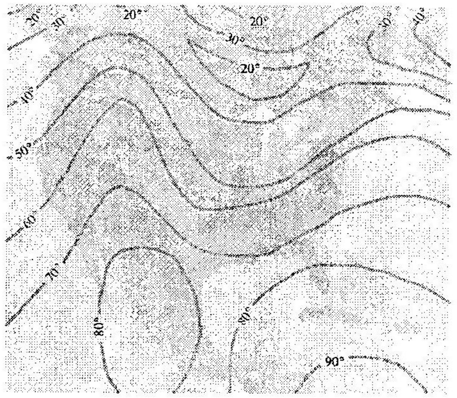


Figure: Level curves show the lines of equal temperature (isotherms) measured in degree Fahrenheit.

Contents

4

Note

A **real-valued function** on subset D of $\mathbb{R}^n$ is a function whose range is $\mathbb{R}$. That is,

$$f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}; (x_1, x_2, \dots, x_n) \longmapsto y = f(x_1, x_2, \dots, x_n)$$

Special cases for $n = 2$ and $n = 3$ will be mainly concerned since they help to visualise their geometrical meaning.

Some Operations on

Let $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. We define

- Addition:
$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$
- Scalar multiplication:
$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$$
- Inner product:
$$x \cdot y = \langle x, y \rangle = xy = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

- In this study we use only **Euclidean Norm**, that is if $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, then norm of x is defined by
$$\|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$
- If $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$, then **norm of the difference** x **and** y (or **Euclidean distance between** x **and** y) is defined by
$$\|x - y\| = \left[\sum_{i=1}^{n} (x_i - y_i)^2\right]^{1/2}.$$

- Let $x_0 \in \mathbb{R}^n$ and let $\epsilon > 0$. A **neighbourhood** or **ϵ-neighbourhood** about x_0 is denoted and defined by
$$N_\epsilon(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \epsilon\}.$$
- A point x_0 in R is called an **interior point** of R if there exists an ϵ-neighbourhood about x_0 that lies entirely in R. That is,
$$x_0 \in N_\epsilon(x_0) \subset R.$$
The **interior** of R, denoted by $\overset{\circ}{R}$ or $\text{int}(R)$, is the set of all interior points of R.
- A point x_0 is a **boundary point** of R if every neighbourhood about x_0 contains points inside R and points outside R.
$$\forall \epsilon > 0 : N_\epsilon(x_0) \cap R \neq \emptyset \quad \text{and} \quad N_\epsilon(x_0) \cap R^c \neq \emptyset$$
The **boundary** of R, denoted by ∂R or $b(R)$, is the set of all boundary points of R.

Lagrange Multipliers

Theorem 56 (Second derivative test for constrained local extremum)

$f, g_1, \ldots, g_k : D \to \mathbb{R}$ be C^2 functions where $D \subset \mathbb{R}^n$ is open and $k < n$. Denote $(\lambda; x)$ and so-called ***Lagrange function*** *$L(\lambda; x) = f(x) - \sum_i^k \lambda_i (g_i(x) - c_i)$. Suppose there is an $x \in D$ such that*

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_k}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial x_1}(a) & \cdots & \frac{\partial g_k}{\partial x_k}(a) \end{pmatrix} \neq 0$$

where $h_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) - \sum_{b=1}^k \lambda_b \frac{\partial^2 g_b}{\partial x_j \partial x_i}(a)$. Let d_j be jth principal minor of $HL(\lambda; a)$.

Lagrange Multipliers

We calculate the sequence of $n - k$ numbers

$$(s) : (-1)^k d_{2k+1}, (-1)^k d_{2k+2}, \ldots, (-1)^k d_{k+n}$$

1. If the sequence in (s) consists entirely of positive numbers, then $f(a)$ is a local minimum subject to the constraints $g_i x) = c_i$ for all $i = 1, 2, \ldots, k$.
2. If the sequence in (s) begin with a negative number and thereafter alternates in sign, then $f(a)$ is a local maximum subject to the constraints $g_i(x) = c_i$ for all $i = 1, 2, \ldots, k$.
3. If neither case 1 nor case 2 holds, then f has a constrained saddle point at a.

References

1. R. Larson and B. Edwards, *Multivariable Calculus*, Ninth Edition, Brooks/Cole, Cengage Learning, 2010.
2. S. T. TAN, *Multivariable Calculus*, Brooks/Cole, Cengage Learning, 2010.