

MAT237 Multivariable Calculus

Lecture Notes

Yuchen Wang, Tingfeng Xia

February 3, 2019

Contents

1	Taylor's Theorem	3
1.1	Review of Taylor's Theorem in 1 Dimension	3
1.2	Taylor's Theorem in higher dimensions	3
2	Critical Points	4
3	Lagrange Multipliers	6
3.1	Two constraints	6
3.2	Inequality Constraints	7
4	The Implicit Function Theorem	7
5	The Inverse Function Theorem	8
6	Some Important Coordinate Systems	8
6.1	Polar Coordinates in \mathbb{R}^2	8
6.2	Spherical Coordinates in \mathbb{R}^3	9
6.3	Cylindrical Coordinates in \mathbb{R}^3	9
7	k-Dimensional Manifolds in \mathbb{R}^n	9
7.1	The General Case	9
7.2	The Specific Cases	10
7.3	Remarks on smoothness of a parametric smooth curve	11
8	Zero content	11
9	Theorems of 1-D Integral Calculus	12

10 Generalized Integral Calculus	14
11 Change of Variables	15
11.1 Some important determinants	16

1 Taylor's Theorem

1.1 Review of Taylor's Theorem in 1 Dimension

Definition of Taylor polynomials Assume $I \subset \mathbb{R}$ is an open interval and that $f : I \rightarrow \mathbb{R}$ is a function of class C^k on I .

For a point $a \in I$, the k th order Taylor polynomial of f at a is the unique polynomial of order at most k , denoted $P_{a,k}(h)$ such that

$$\begin{aligned} f(a) &= P_{a,k}(0) \\ f'(a) &= P'_{a,k}(0) \\ &\vdots \\ f^{(k)}(a) &= P^{(k)}_{a,k}(0) \end{aligned}$$

$$\begin{aligned} P_{a,k}(h) &= f(a) + hf'(a) + \frac{h^2}{2}f''(a) + \dots + \frac{h^k}{k!}f^{(k)}(a) \\ &= \sum_{j=0}^k \frac{h^j}{j!}f^{(j)}(a) \end{aligned}$$

Remark Taylor's Theorem guarantees that $P_{a,k}(h)$ is a very good approximation of $f(a+h)$, and that the quality of the approximation increases as k increases.

Taylor's Theorem in 1D Assume $I \subset \mathbb{R}$ is an open interval and that $f : I \rightarrow \mathbb{R}$ is a function of class C^k on I . For $a \in I$ and $h \in \mathbb{R}$ such that $a+h \in I$, let $P_{a,k}(h)$ denote the k th-order Taylor polynomial at a and define the remainder

$$R_{a,k}(h) := f(a+h) - P_{a,k}(h)$$

Then

$$\lim_{h \rightarrow 0} \frac{R_{a,k}(h)}{h^k} = 0$$

1.2 Taylor's Theorem in higher dimensions

Assume $S \subset \mathbb{R}^n$ is an open set and that $f : S \rightarrow \mathbb{R}$ is a function of class C^k on S . For a point $a \in S$, the k th order Taylor polynomial of f at a is the

unique polynomial of order at most k , denoted $P_{a,k}(\mathbf{h})$ such that

$$\begin{aligned} f(\mathbf{a}) &= P_{\mathbf{a},k}(\mathbf{0}) \\ \partial^\alpha f(\mathbf{a}) &= \partial^\alpha P_{\mathbf{a},k}(\mathbf{0}) \end{aligned}$$

for all partial derivatives of order up to k .

Taylor's Theorem in nD Assume $S \subset \mathbb{R}^n$ is an open interval and that $f : S \rightarrow \mathbb{R}$ is a function of class C^k on I . For $a \in S$ and $h \in \mathbb{R}^n$ such that $a + h \in S$, let $P_{a,k}(h)$ denote the k th-order Taylor polynomial at a and define the remainder

$$R_{a,k}(h) := f(a + h) - P_{a,k}(h)$$

Then

$$\lim_{h \rightarrow 0} \frac{R_{a,k}(h)}{|h|^k} = 0$$

A Taylor polynomial formula for $k = 2$

$$P_{\mathbf{a},2}(\mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2} (H(\mathbf{a})\mathbf{h}) \cdot \mathbf{h}$$

where we remember that $\mathbf{h} = \mathbf{x} - \mathbf{a}$ if we want the result in terms of x, y .

2 Critical Points

Definition A symmetric $n \times n$ matrix A is

1. **positive definite** if $\mathbf{x}^T A \mathbf{x} > 0$ for all $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$
2. **nonnegative definite** if $\mathbf{x}^T A \mathbf{x} \geq 0$ for all $x \in \mathbb{R}^n$

In addition, we say that A is

1. **negative definite** if $-A$ is positive definite
2. **nonpositive definite** if $-A$ is nonnegative definite

A matrix A is **indefinite** if none of the above holds. Equivalently, A is indefinite if there exist $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{x}^T A \mathbf{x} < 0 < \mathbf{y}^T A \mathbf{y}$

Theorem 1 Assume that A is a symmetric matrix. Then

1. A is positive definite \iff all its eigenvalues are positive
 $\iff \exists \lambda_1 > 0$ such that $\mathbf{x}^T A \mathbf{x} \geq \lambda_1 |\mathbf{x}|^2$ for all $\mathbf{x} \in \mathbb{R}^n$
2. A is nonnegative definite \iff all its eigenvalues are nonnegative
3. A is indefinite \iff A has both positive and negative eigenvalues

Remark If A is a symmetric matrix then

The smallest eigenvalue of $A = \min_{\{\mathbf{u} \in \mathbb{R}^n: |\mathbf{u}|=1\}} \mathbf{u}^T A \mathbf{u}$

Theorem 2 For the matrix $A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$,

1. if $\det A < 0$, then A is indefinite
2. if $\det A > 0$, then
 - if $\alpha > 0$ then A is positive definite
 - if $\alpha < 0$ then A is negative definite
3. if $\det A = 0$ then at least one eigenvalue equals zero.

Definition A critical point \mathbf{a} of C^2 function \mathbf{f} is degenerate if $\det(D_{\mathbf{H}}(\mathbf{a})) = 0$

Theorem 3 - first derivative test If $\mathbf{f} : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable, then every local extremum is a critical point.

Theorem 4 - second derivative test

1. If $f : S \rightarrow \mathbb{R}$ is C^2 and \mathbf{a} is a local minimum point for f , then \mathbf{a} is a critical point of f and $H(\mathbf{a})$ is nonnegative definite.
2. If \mathbf{a} is a critical point and $H(\mathbf{a})$ is positive definite, then \mathbf{a} is a local minimum point.

Corollary Assume that f is C^2 and $\nabla f(\mathbf{a}) = \mathbf{0}$

1. If $H(\mathbf{a})$ is positive definite, then \mathbf{a} is a local min;
2. If $H(\mathbf{a})$ is negative definite, then \mathbf{a} is a local max;
3. If $H(\mathbf{a})$ is indefinite, then \mathbf{a} is a saddle point;
4. If none of the above hold, then we cannot determine the character of the critical point without further thought.

E.Knight's approach to critical points. In solving a question of $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ we could use the following "quick check" approach:

1. Calculate the gradient of F , equating it to zero to find the critical points
2. Calculate the Hessian of F , find the corresponding matrices for each critical points, where the Hessian is defined as

$$H(f) = \begin{bmatrix} \partial_{xx}f & \partial_{xy}f = \partial_{yx}f \\ \partial_{xy}f = \partial_{yx}f & \partial_{yy}f \end{bmatrix}$$

3. Calculate the determinant of the hessian, and there are the following cases to consider
 - (a) $\det H < 0$, then $\text{sig}(H) = (1, 1)$ and the point is a saddle point
 - (b) $\det H > 0$, then
 - i. $\text{tr}(H) < 0 \implies \text{sig}(H) = (2, 0)$ and the point is a local minimum
 - ii. $\text{tr}(H) > 0 \implies \text{sig}(H) = (0, 2)$ and the point is a local maximum
 - (c) $\det H = 0$, then the test is inconclusive. We have to do this case by starring at it.

3 Lagrange Multipliers

3.1 Two constraints

Assume that f, g_1 and g_2 are functions $\mathbb{R}^n \rightarrow \mathbb{R}$ of class C^1 . Assume also that $\{\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x})\}$ are linearly independent at all \mathbf{x} where $g_1(\mathbf{x}) = g_2(\mathbf{x}) = 0$

Then if x is any solution to the optimization problem, there exists $\lambda_1, \lambda_2 \in \mathbb{R}$ such that the following system of equations is satisfied by \mathbf{x}, λ_1 and λ_2 :

$$\begin{cases} \nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x}) &= \mathbf{0} \\ g_1(\mathbf{x}) &= 0 \\ g_2(\mathbf{x}) &= 0 \end{cases}$$

3.2 Inequality Constraints

Consider the problem

$$\begin{cases} \text{minimize/maximize } f(\mathbf{x}) \\ \text{subject to the constraint: } g(\mathbf{x}) \leq 0 \end{cases}$$

where we assume that g is C^2 , say, and that $\nabla g(\mathbf{x}) \neq 0$ on the set $\{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$

We can reduce this to problems we already know how to solve. EVT guarantees that the problem has a solution.

Case 1 The max or min occurs in the set $\{x \in \mathbb{R}^n : g(\mathbf{x}) < 0\}$
Then it is a critical point, which we know how to find and classify

Case 2 The max or min occurs in the set $\{x \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$
Then we can find it by the Lagrange Multiplier technique.

Finally we can choose the smallest/largest value of f and the point where that value is attained from among all the candidates found in steps 1 and 2 above.

4 The Implicit Function Theorem

Assume that S is an open subset of \mathbb{R}^{n+k} and that $F : S \rightarrow \mathbb{R}^k$ is a function of class C^1 . Assume also that (\mathbf{a}, \mathbf{b}) is a point in S such that $\mathbf{F}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ and $\det D_{\mathbf{y}}\mathbf{F}(\mathbf{a}, \mathbf{b}) \neq 0$

1. Then there exists $r_0, r_1 > 0$ such that for every $\mathbf{x} \in \mathbb{R}^n$ such that $|\mathbf{x} - \mathbf{a}| < r_0$, there exists a unique $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y} - \mathbf{b}| < r_1$

$$\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}(1)$$

In other words, equation (1) implicitly defines a function $\mathbf{y} = \mathbf{f}(\mathbf{x})$ for $x \in \mathbb{R}^n$ near \mathbf{a} , with $\mathbf{y} = \mathbf{f}(\mathbf{x})$ close to \mathbf{b} . Note in particular that $\mathbf{b} = \mathbf{f}(\mathbf{a})$.

2. Moreover, the function $\mathbf{f} : B(r_0, \mathbf{a}) \rightarrow B(r_1, \mathbf{b}) \subset \mathbb{R}^k$ from part (1) above is of class C^1 , and its derivatives may be determined by differentiating the identity

$$\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{0}$$

and solving to find the partial derivatives of \mathbf{f} .

Remark

$$D\mathbf{f}(\mathbf{a}) = -[D_{\mathbf{y}}\mathbf{F}(\mathbf{a}, \mathbf{b})]^{-1}D_{\mathbf{x}}\mathbf{F}(\mathbf{a}, \mathbf{b})$$

5 The Inverse Function Theorem

Let U and V be open sets in \mathbb{R}^n , and assume that $\mathbf{f} : U \rightarrow V$ is a mapping of class C^1 .

Assume that $\mathbf{a} \in U$ is a point such that $D\mathbf{f}(\mathbf{a})$ is invertible.

and let $\mathbf{b} := \mathbf{f}(\mathbf{a})$. Then there exist open sets $M \subset U$ and $N \subset V$ such that

1. $\mathbf{a} \in M$ and $\mathbf{b} \in N$
2. \mathbf{f} is one-to-one from M onto N (hence invertible), and
3. the inverse function $f^{-1} : N \rightarrow M$ is of class C^1

Moreover, if $x \in M$ and $y = \mathbf{f}(\mathbf{x}) \in N$, then

$$D(\mathbf{f}^{-1})(\mathbf{y}) = [D\mathbf{f}(\mathbf{x})]^{-1}$$

In particular,

$$D(\mathbf{f}^{-1})(\mathbf{b}) = [D\mathbf{f}(\mathbf{a})]^{-1}$$

6 Some Important Coordinate Systems

6.1 Polar Coordinates in \mathbb{R}^2

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix} = \mathbf{f}(r, \theta)$$

For \mathbf{f} to be a bijection between open sets, we have to restrict its domain and range. A common choice is to specify that \mathbf{f} is a function $U \rightarrow V$ where

$$U := \{(r, \theta) : r > 0, |\theta| < \pi\}, V := \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$$

(Note that there is a half of the x-axis missing)

6.2 Spherical Coordinates in \mathbb{R}^3

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \theta \sin \varphi \\ r \sin \theta \sin \varphi \\ r \cos \varphi \end{pmatrix} = \mathbf{f}(r, \theta, \varphi)$$

If we want \mathbf{f} to be a bijection between open sets U and V , it is necessary to restrict the domain and range in some appropriate way.

6.3 Cylindrical Coordinates in \mathbb{R}^3

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ z \end{pmatrix} = \mathbf{f}(r, \theta, z)$$

7 k-Dimensional Manifolds in \mathbb{R}^n

7.1 The General Case

Fix $k < n$. For a k -dimensional manifold M in \mathbb{R}^n , we say that M has "degrees of freedom" k . There are 3 natural ways to represent M (be careful with the dimensions!!!):

1. As a graph:

$$\mathbf{f} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$$

where U is open.

$$S = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) \in \mathbb{R}^n : \mathbf{x} \in U\}$$

2. As a level set:

$$\mathbf{F} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^{n-k}$$

where U is open.

$$S = \{\mathbf{x} \in U : \mathbf{F}(\mathbf{x}) = \mathbf{c}\}$$

for some $\mathbf{c} \in \mathbb{R}^{n-k}$.

This is also called the "zero locus" of \mathbf{F} when $\mathbf{c} = \mathbf{0}$

Remark 1 The regularity conditions that guarantees that S is smooth is that

1. $\nabla F_1(\mathbf{x}), \dots, \nabla F_{n-k}(\mathbf{x})$ are linearly independent at each $\mathbf{x} \in S$. Or equivalently,
2. the matrix $D\mathbf{F}(\mathbf{x})$ has rank $n - k$ at every $\mathbf{x} \in S$.

Remark 2 It can happen that the above conditions are satisfied but S is not smooth. Example: The square of a smooth F , $c = 0$

3. Parametrically

$$\mathbf{f} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$$

where U is open.

$$S = \{\mathbf{f}(\mathbf{u}) : \mathbf{u} \in U\}$$

Remark The regularity conditions that guarantees that S is smooth is that

1. $\partial_{u_1}\mathbf{f}(\mathbf{u}), \dots, \partial_{u_k}\mathbf{f}(\mathbf{u})$ are linearly independent at each $\mathbf{u} \in U$. Or equivalently,
2. the matrix $D\mathbf{f}(\mathbf{u})$ has rank k at every $\mathbf{u} \in U$.

Notes We can prove that if the above conditions are satisfied, then S is smooth. Construct $\mathbf{F} : \mathbb{R}^{2k} \rightarrow \mathbb{R}^k$, then use IFT (the proof is hard but worthwhile to think about since the general case implies every specific case).

7.2 The Specific Cases

Theorem 1 - When is a curve regular? Assume that $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}$ is C^1 , and let

$$S := \{\mathbf{x} \in \mathbb{R}^2 : F(\mathbf{x}) = 0\}$$

If $\mathbf{a} \in S$ and $\nabla F(\mathbf{a}) \neq 0$, then there exists some $r > 0$ such that $B(r, \mathbf{a}) \cap S$ is a C^1 graph.

(Prove directly using IFT)

Theorem 2 - When is the parametrization regular? Assume that $\mathbf{f} : (a, b) \rightarrow \mathbb{R}^2$ is C^1 , and let

$$S := \{\mathbf{f}(t) : t \in (a, b)\}$$

If $\mathbf{f}'(c) \neq 0$ for some $c \in (a, b)$, then there exists some $r > 0$ such that $\{\mathbf{f}(t) : |t - c| < r\}$ is a C^1 graph.

Remark It says only that the parametrization is regular near $t = c$, it does not say that S is regular near $\mathbf{f}(c)$. What it means is that when increasing/decreasing t , we have no control over the path of $\mathbf{f}(t)$.

Theorem 3- When is a surface regular? conditions: $\mathbf{a} \in S$ and $\nabla F(\mathbf{a}) \neq 0$

Theorem 4 - When is the parametrization regular? conditions: $Df(\mathbf{c})$ has rank 2 at some c

7.3 Remarks on smoothness of a parametric smooth curve

Definition. If $I \subseteq \mathbb{R}$ is an interval, a \mathcal{C}^1 map $\gamma : I \rightarrow \mathbb{R}^2$ is said to be

1. A *regular curve* if $\gamma'(t) \neq 0, \forall t \in I$
2. A *simple curve* if γ is injective on the interior of I .

Hence if γ is regular, then there is a neighbourhood of each point whose image looks like a graph of a \mathcal{C}^1 -function. Simplicity guarantees that no funny overlaps can happen, and this is what is needed for the curve to be smooth. **As a conclusion, we say a parametric curve is smooth if and only if it is *regular* and *simple*.** Equivalently we could also convert such parametrization into a zero locus of a F to use the good old method of gradient directly.

8 Zero content

Zero content in 1-D A set $S \subset \mathbb{R}$ is said to have zero content if

$$\forall \epsilon > 0, \exists \text{ intervals } I_1, \dots, I_n \text{ s.t. } S \subseteq \bigcup_{i=1}^n I_i \wedge \sum_{i=1}^n \text{Len}(I_i) < \epsilon$$

Multidimensional zero content. A set $S \subset \mathbb{R}^n$ is said to have zero content if

$$\forall \epsilon > 0, \exists \text{ boxes } B_1, \dots, B_n \text{ s.t. } S \subseteq \bigcup_{i=1}^n B_i \wedge \sum_{i=1}^n \text{Area}(B_i) < \epsilon$$

Consequence of zero content. If a set Z has zero content, then

$$\forall \epsilon > 0, \exists \text{ boxes } B_1, \dots, B_n \text{ s.t. } S \subseteq \bigcup_{i=1}^n B_i^{\text{int}} \wedge \sum_{i=1}^n \text{Area}(B_i) < \epsilon$$

Notice the extra *int*.

Proposition on zero content

1. If $Z \subset \mathbb{R}^2$ has zero content and $U \subset Z$, then U has zero content.
2. If Z_1, \dots, Z_k have zero content, then so does $\bigcup_1^k Z_j$
3. $\mathbf{f} : (a_0, b_0) \rightarrow \mathbb{R}^2$ is of class C_1 , then $\mathbf{f}([a, b])$ has zero content whenever $a_0 < a < b < b_0$

9 Theorems of 1-D Integral Calculus

Lemma: Refined partitions give better approximations Let P be some partition over an interval and let P' be a refinement of P , then

$$LS_{P'}f \geq LS_Pf \wedge US_{P'} \leq US_Pf$$

Where LS and US stands for lower sum and upper sum respectively.

Lemma: Lower sum is always less then or equal to upper sum If P and Q are any partitions of $[a, b]$, then $LS_Pf \leq US_Qf$. The essence of this proof is to consider the common refinement of these two partitions.

Lemma. $\epsilon - \delta$ definition of integrability If f is a bounded function on $[a, b]$, the following conditions are equivalent:

1. f is integrable on $[a, b]$
2. $\forall \epsilon > 0, \exists P$ of $[a, b]$ such that $US_Pf - LS_Pf < \epsilon$

Theorem: Integration is “Linear”

1. Suppose $a < b < c$. If f is integrable on $[a, b]$ and on $[b, c]$, then f is integrable on $[a, c]$, further more

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$$

2. If f and g are integrable on $[a, b]$, then so is $f + g$, further more

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

Theorem. Suppose f is integrable on $[a, b]$.

1. If $c \in \mathbb{R}$, the cf is integrable on $[a, b]$, and $\int_a^b cf(x) = c \int_a^b f(x)dx$
2. Of $[c, d] \subset [a, b]$, then f is integrable on $[c, d]$.
3. If g is integrable on $[a, b]$ and $f(x) \leq g(x), \forall x \in [a, b]$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$
4. $|f|$ is integrable on $[a, b]$, and $\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx$

Theorem: Bounded + monotone \implies integrable If f is bounded and monotone on $[a, b]$, then f is integrable on $[a, b]$. The proof of this uses the $\epsilon - \delta$ definition of integrability

Theorem: Continuous \implies integrable If f is continuous on $[a, b]$, then f is integrable on $[a, b]$. Note that continuous is a sufficient but not necessary condition of integrability

Theorem: discontinuous at only finite pts \implies integrable If f is bounded on $[a, b]$ and continuous at all except finitely many points in $[a, b]$, then f is integrable on $[a, b]$. A easy example of this would be any \mathbb{R} function that has a hole in it.

Theorem: Discontinuous at only zero content \implies integrable If f is bounded on $[a, b]$ and the set of points in $[a, b]$ at which f is discontinuous has zero content, then f is integrable on $[a, b]$.

Proposition. Suppose f and g are integrable on $[a, b]$ and $f(x) = g(x)$ for all except finitely many points $x \in [a, b]$. Then $\int_a^b f(x)dx = \int_a^b g(x)dx$.

The Fundamental Theorem Of Calculus

1. Let f be an integrable function on $[a, b]$. For $x \in [a, b]$, let $F(x) = \int_a^x f(t)dt$. Then F is continuous on $[a, b]$; more-over, $F'(x)$ exists and equals $f(x)$ at every x at which f is continuous,
2. Let F be a continuous function on $[a, b]$ that is differentiable except perhaps at finitely many points in $[a, b]$, and let f be a function on $[a, b]$ that agrees with F' at all points where the latter is defined. If f is integrable on $[a, b]$, then $\int_a^b f(t)dt = F(b) - F(a)$

Proposition. Suppose f is integrable on $[a, b]$. Given $\epsilon > 0, \exists \delta > 0$ such that if $P = \{x_0, \dots, x_J\}$ is any partition of $[a, b]$ satisfying

$$\max\{x_j - x_{j-1} | 1 \leq j \leq J\} < \delta$$

the sums $LS_P f$ and $US_P f$ differ from $\int_a^b f(x)dx$ by at most ϵ .

10 Generalized Integral Calculus

Theorems of double integrals

1. If f_1 and f_2 are integrable on the bounded set S and $c_1, c_2 \in \mathbb{R}$, then $c_1 f_1 + c_2 f_2$ is integrable on S , and

$$\iint_S [c_1 f_1 + c_2 f_2] dA = c_1 \iint_S f_1 dA + c_2 \iint_S f_2 dA$$

2. Let S_1 and S_2 be bounded sets with no points in common (intersection $= \emptyset$), and let f be a bounded function. If f is integrable on S_1 and on S_2 , then f is integrable on $S_1 \cup S_2$, in which case

$$\iint_{S_1 \cup S_2} f dA = \iint_{S_1} f dA + \iint_{S_2} f dA$$

3. If f and g are integrable on S and $f(\mathbf{x}) \leq g(\mathbf{x})$ for $\mathbf{x} \in S$, then $\iint_S f dA \leq \iint_S g dA$
4. If f is integrable on S , then so is $|f|$, and

$$\left| \iint_S f dA \right| \leq \iint_S |f| dA$$

Theorem. Suppose f is a bounded function on the rectangle R . If the set of points in R at which f is discontinuous has zero content, then f is integrable on R .

Discontinuity of characteristic function The function χ_S is discontinuous at \mathbf{x} if and only if \mathbf{x} is in the boundary of S .

Theorem. Let S be a measurable subset of \mathbb{R}^2 . Suppose $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is bounded and the set of points in S at which f is discontinuous has zero content. Then f is integrable on S .

Remark on this theorem: The only points where f_{χ_S} can be discontinuous are those points in the closure of S where either f or χ_S is discontinuous. Both of these cases are discontinuity on a set of zero content. And we can definitely fix S inside of a rectangle, then by the previously stated theorem (The theorem directly above), such function is integrable.

Proposition: Integration on a set of zero content evaluates to zero.

Suppose $Z \subset \mathbb{R}^2$ has zero content. If $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is bounded, then f is integrable on Z and $\int_Z f dA = 0$

Corollary

1. Suppose that f is integrable on the set $S \subset \mathbb{R}^2$. If $g(\mathbf{x}) = f(\mathbf{x})$ except for \mathbf{x} in a set of zero content, then g is integrable on S and $\int_S g dA = \int_S f dA$
2. Suppose that f is integrable on S and on T , and $S \cap T$ has zero content. Then f is integrable on $S \cup T$, and $\int_{S \cup T} f dA = \int_S f dA + \int_T f dA$

Fubini's Theorem Let $R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$, and let f be an integrable function on R . Suppose that, for each $y \in [c, d]$, the function f_y defined by $f_y(x) = f(x, y)$ is integrable on $[a, b]$, and the function $g(y) = \int_a^b f(x, y) dx$ is integrable on $[c, d]$. Then

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

Likewise, if $f^x(y) = f(x, y)$ is integrable on $[c, d]$ for each $x \in [a, b]$, and $h(x) = \int_c^d f(x, y) dy$ is integrable on $[a, b]$, then

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

11 Change of Variables

Change of Variable formula 1D If g is a one-to-one function of class C^1 on the interval $[a, b]$, then for any continuous function f ,

$$\int_{[a, b]} f(g(u)) |g'(u)| du = \int_{g([a, b])} f(x) dx$$

In practice it is often more convenient to have all the g 's on one side of the equation. If we set $I = g([a, b])$, we have $[a, b] = g^{-1}(I)$, and

$$\int_I f(x) dx = \int_{g^{-1}(I)} f(g(u)) |g'(u)| du$$

goal: find the analogous formula for multiple integrals. The question is: How does the volume of a tiny piece of n -space change when one applies the transformation \mathbf{G} ?

Theorem - Change of Variable for linear mappings Let A be an invertible $n \times n$ matrix, and let $\mathbf{G}(\mathbf{u}) = A\mathbf{u}$ be the corresponding linear transformation of \mathbb{R}^n . Suppose S is a measurable region in \mathbb{R}^n and f is an integrable function on S . Then $G^{-1}(S) = \{A^{-1}\mathbf{x} : \mathbf{x} \in S\}$ is measurable and $f \circ \mathbf{G}$ is integrable on $\mathbf{G}^{-1}(S)$, and

$$\int \dots \int_S f(\mathbf{x}) d^n \mathbf{x} = |\det A| \int \dots \int_{\mathbf{G}^{-1}(S)} f(A\mathbf{u}) d^n \mathbf{u}$$

Theorem - Change of Variable for general functions Given open sets U and V in \mathbb{R}^n , let $\mathbf{G} : U \rightarrow V$ be a one-to-one transformation of class C^1 whose derivative $D\mathbf{G}(\mathbf{u})$ is invertible for all $\mathbf{u} \in U$. Suppose that $T \subset U$ and $S \subset V$ are measurable sets such that $\mathbf{G}(T) = S$. If f is an integrable function on S , then $f \circ \mathbf{G}$ is integrable on T , and

$$\int \dots \int_S f(\mathbf{x}) d^n \mathbf{x} = \int \dots \int_T f(\mathbf{G}(\mathbf{u})) |\det D\mathbf{G}(\mathbf{u})| d^n \mathbf{u}$$

11.1 Some important determinants

1. polar coordinates: factor = r
2. cylindrical coordinates: factor = $r^2 \sin \varphi$