Lecture Notes 9: Measure and Probability
Part B: Measure and Multiple Integration

Peter Hammond

2024 September 20th,
typeset from `measProbB24.tex`

Outline

Measurable Rectangles

Let (X, Σ_X) and (Y, Σ_Y) be two measurable spaces, with their respective σ-algebras Σ_X and Σ_Y.

The Cartesian product of X and Y is

$$X \times Y = \{(x, y) \mid x \in X \; y \in Y\}$$

Let $\Sigma_X \times \Sigma_Y = \{A \times B \mid A \in \Sigma_X, \; B \in \Sigma_Y\}$ denote the set of measurable rectangles that are the Cartesian product of two measurable sets

Example

Suppose that $X = \{a, b\}$ and $Y = \{c, d\}$, with $\Sigma_X = 2^X$ and $\Sigma_Y = 2^Y$.

Then $\#\Sigma_X = \#\Sigma_Y = 4$ and $\#(\Sigma_X \times \Sigma_Y) = 10$ after identifying $E \times \emptyset = \emptyset \times F = \emptyset$ for all $E \subseteq X$ and all $F \subseteq Y$.

But then $(X \times Y) \setminus \{a, c\} = (X \times \{d\}) \cup (\{b\} \times Y) \notin \Sigma_X \times \Sigma_Y$.

This implies that $\Sigma_X \times \Sigma_Y$ is not a σ-algebra.

The Product of Two Measurable Spaces

So we define the product σ-algebra, denoted by $\Sigma_X \otimes \Sigma_Y$, as $\sigma(\Sigma_X \times \Sigma_Y)$, the σ-algebra generated by $\Sigma_X \times \Sigma_Y$.

It is the smallest σ-algebra that contains all measurable rectangles $A \times B$ with $A \in \Sigma_X$ and $B \in \Sigma_Y$.

And we define the product of the two measurable spaces (X, Σ_X) and (Y, Σ_Y) as the measurable space $(X \times Y, \Sigma_X \otimes \Sigma_Y)$.

The function $X \times Y \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ of two variables (x, y) is product measurable just in case, for each Borel set $E \in \mathcal{B}(\mathbb{R})$, the inverse $f^{-1}(B)$ is $\Sigma_X \otimes \Sigma_Y$-measurable.

The Product of Two Measure Spaces

Let (X, Σ_X, μ_X) and (Y, Σ_Y, μ_Y) be two measure spaces, and $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ the product measurable space.

Say that μ on $(X \times Y, \Sigma_X \otimes \Sigma_Y)$ is a product measure just in case it is a measure that satisfies $\mu(E \times F) = \mu_X(E) \times \mu_Y(F)$ for all measurable rectangles $E \times F \in \Sigma_X \times \Sigma_Y$.

Typically there is a unique product measure with this property, which we denote by $\mu_X \otimes \mu_Y$.

Then $(X \times Y, \Sigma_X \otimes \Sigma_Y, \mu_X \otimes \mu_Y)$ is the product of the two measure spaces.

The Fubini Theorem

Theorem (Fubini)

Provided that $X \times Y \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$
is measurable w.r.t. the product σ*-algebra* $\Sigma_X \otimes \Sigma_Y$*,*
its integral w.r.t. the product measure $\mu_X \otimes \mu_Y$ *satisfies*

$$\begin{aligned} &\int_{X \times Y} f(x, y)(\mu_X \otimes \mu_Y)(\mathrm{d}\, x \times \mathrm{d}\, y) \\ &= \int_X \left[\int_Y f(x, y) \mu_Y(\mathrm{d}\, y) \right] \mu_X(\mathrm{d}\, x) \\ &= \int_Y \left[\int_X f(x, y) \mu_X(\mathrm{d}\, x) \right] \mu_Y(\mathrm{d}\, y) \end{aligned}$$

That is, for any product measurable function,
the order of integration is irrelevant.

Product Measure as a Double Integral

Corollary

For every $E \in \Sigma_X \otimes \Sigma_Y$, its product measure satisfies

$$
\begin{aligned}
(\mu_X \otimes \mu_Y)(E) &= \int_E 1_E(x,y)(\mu_X \otimes \mu_Y)(\mathrm{d}\, x \times \mathrm{d}\, y) \\
&= \int_X \left[\int_Y 1_E(x,y) \mu_Y(\mathrm{d}\, y) \right] \mu_X(\mathrm{d}\, x) \\
&= \int_Y \left[\int_X 1_E(x,y) \mu_X(\mathrm{d}\, x) \right] \mu_Y(\mathrm{d}\, y)
\end{aligned}
$$

The Lebesgue Plane

Example

Suppose the two measure spaces (X, Σ_X, μ_X) and (Y, Σ_Y, μ_Y) are both copies of the Lebesgue real line $(\mathbb{R}, \mathcal{L}, \lambda)$ where:

1. $\mathcal{L}$ is the Lebesgue completion of the Borel σ-algebra on $\mathbb{R}$;
2. λ is the Lebesgue measure which satisfies $\lambda(I) = b - a$ for any interval $I \subset \mathbb{R}$ with endpoints a and b satisfying $a \leq b$.

Then the measure product $(\mathbb{R}, \mathcal{L}, \lambda)^2$ is the Lebesgue plane in the form of the measure space $(\mathbb{R}^2, \mathcal{A}, \alpha)$, where:

1. $\mathcal{A} = \mathcal{L} \otimes \mathcal{L}$ is the product of the Lebesgue σ-algebra on $\mathbb{R}$ with itself;
2. $\alpha = \lambda \otimes \lambda$ has the property that, for each $E \in \mathcal{A}$, the measure $\alpha(E)$ is its area.

In particular, the measure α on the measurable space $(\mathbb{R}^2, \mathcal{A})$ is the unique measure that satisfies $\alpha(I_X \times I_Y) = \lambda(I_X)\lambda(I_Y)$ for every product measurable rectangle $I_X \times I_Y$.

Outline

Recalling the Definition of an Antiderivative in $\mathbb{R}$

The following definition is taken (with some changes of notation) from the review set out in FMEA, Section 4.1.

Definition

Let $I \ni x \mapsto f(x) \in \mathbb{R}$ be a continuous function defined on an interval $I \subset \mathbb{R}$.

An indefinite integral of f is a function $I \ni x \mapsto F(x) \in \mathbb{R}$ whose derivative, for all x in I, exists and is equal to $f(x)$ — in symbols $\int f(\xi)\, \mathrm{d}\,\xi = F(x) + C$ where $F'(x) = f(x)$.

In effect, this defines an equivalence class of functions, where $F \sim G \iff \exists C \in \mathbb{R}; \forall x \in I : F(x) - G(x) = C$.

An indefinite integral is often described as an antiderivative, or an N–L integral where "N–L" stands for "Newton–Leibniz".

Relating Definite to Indefinite Integrals

The following definition is taken (with some changes of notation) from EMEA6, Section 10.2, (10.2.3).

Definition

Let $I \ni x \mapsto f(x) \in \mathbb{R}$ be a continuous function defined on an interval $I \subset \mathbb{R}$.

The definite integral of f over any interval $[a, b] \subset I$ is

$$\int_a^b f(\xi)\, \mathrm{d}\,\xi = F(b) - F(a)$$

where F is any indefinite integral of f.

Existence of an Antiderivative

Definition

Let $I \ni x \mapsto f(x) \in \mathbb{R}$ be any Lebesgue integrable function which is defined on an interval $I \subset \mathbb{R}$.

For each fixed $a \in \operatorname{int} I$, define the N–L integral function

$$(a, +\infty) \cap \operatorname{int} I \ni x \mapsto F(x) := \int_a^x f(\xi)\, d\,\xi = \int_a^x f(\xi)\, \lambda(d\,\xi)$$

where λ denotes Lebesgue measure on $\mathbb{R}$.

Theorem

Let $I \ni x \mapsto f(x) \in \mathbb{R}$ be any integrable function defined on an interval $I \subset \mathbb{R}$.

Then at any point $x_0 \in I$ where f is continuous, the N–L integral function F is differentiable with $F'(x_0) = f(x_0)$.

Proof.

The proof using upper and lower integrals is left as an exercise. □

Outline

A Definition of Antiderivative in Two Dimensions

Definition

Let $D \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^2$.

An indefinite integral of f is a function $D \ni (x, y) \mapsto F(x, y) \in \mathbb{R}$ whose mixed partial derivative, for all $(x, y) \in D$, exists and is equal to $f(x, y)$ — in symbols

$$\int f(\xi, \eta) \, \mathrm{d}\, \xi \, \mathrm{d}\, \eta = F(x, y) + C$$

$$\text{where} \quad F''_{12}(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y) = f(x, y)$$

Definition of an Integral Function

Given any point $(a, b) \in \mathbb{R}^2$, let

$$(a, b)_{\geqq} := \{(x, y) \in \mathbb{R}^2 \mid x \geqq a \text{ and } y \geqq b\}$$

denote the set $\{(a, b)\} + \mathbb{R}^2_+$ that results
when the non-negative quadrant $\mathbb{R}^2_+$ is shifted
so that its the bottom left-hand corner $(0, 0)$ is moved to (a, b).

Definition

Let $D \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ be a continuous function
defined on an open and convex domain $D \subset \mathbb{R}^2$.

For each fixed $(a, b) \in D$, define the definite integral function

$$\begin{aligned} (a, b)_{\geqq} \cap D \ni (x, y) \mapsto I_f(x, y) &:= \int_a^x \int_b^y f(\xi, \eta) \, \mathrm{d}\, \xi \, \mathrm{d}\, \eta \\ &= \int_a^x \int_b^y f(\xi, \eta) \, \lambda^2(\mathrm{d}\, \xi \times \mathrm{d}\, \eta) \end{aligned}$$

where λ^2 denotes Lebesgue measure on $\mathbb{R}^2$.

Existence of an Antiderivative: Statement of Theorem

Theorem

Let $D \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^2$.

Then given any fixed $(a, b) \in D$, the function

$$(a, b)_{\geqq} \cap D \ni (x, y) \mapsto F(x, y) := \int_a^x \int_b^y f(\xi, \eta) \, \mathrm{d}\,\xi \, \mathrm{d}\,\eta \in \mathbb{R}$$

has a mixed second derivative $F''_{12}(x, y) = F''_{21}(x, y)$ that equals $f(x, y)$ at (x, y).

Existence of an Antiderivative: Proof of Theorem

Proof.

Recall the definition

$$(a,b)_{\geqq} \cap D \ni (x,y) \mapsto F(x,y) := \int_a^x \int_b^y f(\xi,\eta)\,\mathrm{d}\,\xi\,\mathrm{d}\,\eta \in \mathbb{R}$$

Differentiating this definition once partially w.r.t. x gives $F_1'(x,y) = \int_b^y f(x,\eta)\,\mathrm{d}\,\eta$.

Differentiating this equation partially w.r.t. y gives $F_{21}''(x,y) = f(x,y)$.

Because $F_{21}''(x,y) = f(x,y)$ is continuous, Young's theorem on the symmetry of second-order partial derivatives implies that $F_{12}''(x,y) = F_{21}''(x,y)$. □

Useful Lemma in Two Dimensions

Lemma

Let $D \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$ *be a continuous function defined on an open and convex domain* $D \subset \mathbb{R}^2$.

For every fixed $(a, b) \in D$, *as well as* $d, e > 0$, *one has*

$$\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^2} \int_a^{a+\epsilon d} \int_b^{b+\epsilon e} f(\xi, \eta) \, \mathrm{d}\,\xi \, \mathrm{d}\,\eta = d \cdot e \cdot f(a, b)$$

Proof of Lemma

Proof.

Let $\langle \epsilon_k \rangle_{k \in \mathbb{N}}$ be any sequence of positive numbers such that $\epsilon_k \to 0$ as $k \to \infty$.

By the mean value theorem for double integrals, for each $k \in \mathbb{N}$ there exists a point (x_k, y_k) in the rectangle $[a, a + \epsilon_k d] \times [b, b + \epsilon_k e] \subset \mathbb{R}^2$, whose area is $\epsilon_k^2 \cdot d \cdot e$, such that

$$\frac{1}{\epsilon_k^2} \int_a^{a+\epsilon_k d} \int_b^{b+\epsilon_k e} f(\xi, \eta) \, \mathrm{d}\,\xi \, \mathrm{d}\,\eta = d \cdot e \cdot f(x_k, y_k)$$

Because $a \leq x_k \leq a + \epsilon_k d$ and $b \leq y_k \leq b + \epsilon_k e$, taking limits as $k \to \infty$ and so $\epsilon_k \downarrow 0$ implies that $x_k \to a$ and $y_k \to b$.

Then continuity of f implies that $f(x_k, y_k)$ converges to $f(a, b)$, so the result follows. □

Outline

A Definition of Antiderivative in n Dimensions

Given a function $\mathbb{R}^n \supset S \ni \mathbf{x} \mapsto F(\mathbf{x}) \in \mathbb{R}$,
we introduce the notation $\partial^n F(\mathbf{x})$ as an abbreviation
for the nth order partial derivative $\frac{\partial^n}{\partial x_1 \partial x_2 \ldots, \partial x_n} F(\mathbf{x})$, when it exists.

Definition

Let $D \ni \mathbf{x} \mapsto f(\mathbf{x}) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^n$.

An indefinite integral of f is a function $D \ni \mathbf{x} \mapsto F(\mathbf{x}) \in \mathbb{R}$ whose mixed partial derivative $\partial^n F(\mathbf{x})$, for all $\mathbf{x} \in D$, exists and is equal to $f(\mathbf{x})$ — in symbols

$$\int\int \cdots \int f(\mathbf{x})\, \mathrm{d}\,\mathbf{x} = F(\mathbf{x}) + C \quad \text{where} \quad \partial^n F(\mathbf{x}) = f(\mathbf{x})$$

Orthants and Cuboids in $\mathbb{R}^n$

Given any two points $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$,
define the following three subsets of $\mathbb{R}^n$:

1. $\mathbf{a}_{\geqq} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \geqq \mathbf{a}\} = \{\mathbf{a}\} + \mathbb{R}^n_+$, the set that results when the non-negative orthant $\mathbb{R}^n_+$ of $\mathbb{R}^n$ is shifted so that the corner or extreme point at $\mathbf{0}$ is moved to $\mathbf{a}$;
2. $\mathbf{b}_{\leqq} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \leqq \mathbf{b}\} = \{\mathbf{b}\} - \mathbb{R}^n_+$, the set that results when the non-positive orthant $\mathbb{R}^n_- = -\mathbb{R}^n_+$ of $\mathbb{R}^n$ is shifted so that the corner or extreme point at $\mathbf{0}$ is moved to $\mathbf{b}$;
3. $[\mathbf{a}, \mathbf{b}] := \mathbf{a}_{\geqq} \cap \mathbf{b}_{\leqq}$ denote the (possibly empty) n-dimensional cuboid $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \leqq \mathbf{x} \leqq \mathbf{b}\}$.

Definition of an Integral Function

For each $E \subseteq \mathbb{R}^n$, recall the definition $\mathbb{R}^n \ni \mathbf{x} \mapsto 1_E(\mathbf{x}) \in \{0, 1\}$ of the indicator function for the set E that satisfies $1_E(\mathbf{x}) = 1 \iff x \in E$.

Definition

Let $D \ni \mathbf{x} \mapsto f(\mathbf{x}) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^n$.

For each fixed $\mathbf{a} \in D$, define the definite integral function

$$\begin{aligned} \mathbf{a}_{\geqq} \cap D \ni \mathbf{b} \mapsto F(\mathbf{b}) &:= \int_{\mathbf{a}}^{\mathbf{b}} 1_D(\mathbf{x})\, f(\mathbf{x})\, \lambda^n(\mathrm{d}\,\mathbf{x}) \\ &= \int_D 1_{[\mathbf{a},\mathbf{b}]}(\mathbf{x}) f(\mathbf{x})\, \lambda^n(\mathrm{d}\,\mathbf{x}) \end{aligned}$$

where λ^n denotes Lebesgue measure on $\mathbb{R}^n$.

Existence of an Antiderivative

Theorem

Let $D \ni \mathbf{x} \mapsto f(\mathbf{x}) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^n$.

Then given any fixed $\mathbf{a} \in D$, for each $\mathbf{b} \in \mathbf{a}_{\geqq} \cap D$,

the function $\mathbf{b} \mapsto F(\mathbf{b}) := \int_{\mathbf{a}}^{\mathbf{b}} f(\mathbf{x})\, \mathrm{d}\,\mathbf{x}$

has a mixed nth derivative $\partial^n F(\mathbf{x})$ that equals $f(\mathbf{x})$ at $\mathbf{x}$.

Proof.

The proof, based on integrating n times the function $\mathbf{x} \mapsto f(\mathbf{x})$, is a straightforward extension of the proof given for $\mathbb{R}^2$. □

Useful Lemma in n Dimensions

Lemma

Let $D \ni \mathbf{x} \mapsto f(\mathbf{x}) \in \mathbb{R}$ be a continuous function defined on an open and convex domain $D \subset \mathbb{R}^n$.

For every fixed $\mathbf{a} \in D$ and $\mathbf{e} = \langle e_i \rangle_{i=1}^n \in \mathbb{R}^n_{++}$, one has

$$\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^n} \int_{\mathbf{a}}^{\mathbf{a}+\epsilon\mathbf{e}} f(\mathbf{x})\, \mathrm{d}\,\mathbf{x} = \prod\nolimits_{i=1}^n e_i \cdot f(\mathbf{a})$$

Proof.

The proof is similar to that we gave when $n = 2$. □

Remark

Recall that,
given the diagonal matrix $\mathbf{diag}\,\mathbf{e} = \mathbf{diag}(e_1, e_2, \ldots, e_n)$,
the product $\prod_{i=1}^n e_i$ equals the volume $\mathrm{vol}_n(\mathbf{diag}\,\mathbf{e})$
of the n-dimensional cuboid $\sum_{i=1}^n [\mathbf{0}, e_i \mathbf{e}_i]$
where each $\mathbf{e}_i = (\delta_{ij})_{j=1}^n$ is the ith column of the identity matrix $\mathbf{I}$.

Outline

Integration by Substitution in One Variable

Suppose that, in looking for an antiderivative function

$$\mathbb{R} \ni x \mapsto F(x) = \int f(x) \, \mathrm{d}\, x \in R$$

such that $F'(x) = f(x)$, we try the substitution $x = g(u)$.

This implies that $\mathrm{d}\, x = g'(u) \, \mathrm{d}\, u$.

So the original antiderivative $F(x) = \int f(x) \, \mathrm{d}\, x$ becomes the transformed antiderivative $G(u) = \int f(g(u))g'(u) \, \mathrm{d}\, u$, which may be easier to find.

Outline

Change of Variables (FMEA, Theorem 4.7.2)

Theorem

Suppose that $A' \ni \mathbf{u} \mapsto \mathbf{g}(\mathbf{u}) = (g_1(\mathbf{u}), \ldots, g_n(\mathbf{u})) \in \mathbb{R}^n$
is used to specify the transformation $\mathbf{x} = \mathbf{g}(\mathbf{u})$
from an open and bounded set $A' \subset \mathbb{R}^n$ *in "*$\mathbf{u}$*-space"*
onto an open and bounded set $A \subset \mathbb{R}^n$ *in "*$\mathbf{x}$*-space".*

Suppose that the Jacobian matrix function

$$A' \ni \mathbf{u} \mapsto \mathbf{J}(\mathbf{u}) = \frac{\partial(g_1, \ldots, g_n)}{\partial(u_1, \ldots, u_n)}(\mathbf{u}) = \frac{\partial \mathbf{g}}{\partial \mathbf{u}}(\mathbf{u}) \in \mathbb{R}^{n \times n}$$

is bounded.

Let f *be a bounded, continuous function defined on* A*. Then*

$$\int \ldots \int_A f\left(x_1, \ldots x_n\right) \mathrm{d}\, x_1 \ldots \mathrm{d}\, x_n$$
$$= \int \ldots \int_{A'} f\left(g_1(\mathbf{u}), \ldots, g_n(\mathbf{u})\right) \left|\det \mathbf{J}(\mathbf{u})\right| \mathrm{d}\, u_1 \ldots \mathrm{d}\, u_n$$

Outline

An Instructive Example, I

In one dimension, integration by substitution leads to the formula $\int f(g(u))g'(u) \, \mathrm{d}\, u$.

By contrast, in n dimensions, one has

$$\int \ldots \int_A f\left(x_1, \ldots x_n\right) \, \mathrm{d}\, x_1 \ldots \mathrm{d}\, x_n$$
$$= \int \ldots \int_{A'} f\left(g_1(\mathbf{u}), \ldots, g_n(\mathbf{u})\right) |\det \mathbf{J}(\mathbf{u})| \, \mathrm{d}\, u_1 \ldots \mathrm{d}\, u_n$$

with the absolute value of the Jacobian determinant.

Why is there this contrast?

An Instructive Example, II

Consider the definite integral

$$J = \int_0^1 (1 - x)\, \mathrm{d}\, x = |_0^1 (x - \tfrac{1}{2}x^2) = 1 - \tfrac{1}{2} = \tfrac{1}{2}$$

Suppose we try to make things even simpler by using the substitution $u = 1 - x$.

Then $u = 1$ when $x = 0$ and $u = 0$ when $x = 1$.

Also $\mathrm{d}\, x = -\mathrm{d}\, u$, so the integral becomes

$$J = \int_1^0 u(-\mathrm{d}\, u) = |_1^0 (-\tfrac{1}{2}u^2) = \tfrac{1}{2}$$

An Instructive Example, III

We are integrating over the interval $I = [0, 1]$, so $J = \int_I (1 - x)\,\mathrm{d}\,x$.

When we make the substitution $u = 1 - x$, where $\mathrm{d}\,x = (-1)\,\mathrm{d}\,u$, the integration by substitution formula seems to suggest the transformation

$$\tilde{J} = \int_I u(-1)\,\mathrm{d}\,u = \int_0^1 u(-1)\,\mathrm{d}\,u = |_0^1(-\tfrac{1}{2}u^2) = -\tfrac{1}{2}$$

But then $\tilde{J} = -J$, so we evidently have a wrong answer!

To get the right answer, we need to consider the absolute value $+1$ of the Jacobian scalar -1.

This gives $J^* = \int_I u(+1)\,\mathrm{d}\,u = \int_0^1 u\,\mathrm{d}\,u = |_0^1(\tfrac{1}{2}u^2) = \tfrac{1}{2}$ which is the right answer.

Outline

Outline of a Justification in a Special Case, I

Let $D \ni \mathbf{x} \mapsto f(\mathbf{x}) \in \mathbb{R}$ be a C^1 function
defined on an open and convex domain $D \subset \mathbb{R}^n$.

Consider the special case when the mapping $D' \ni \mathbf{u} \mapsto \mathbf{g}(\mathbf{u}) \in \mathbb{R}^n$
determines a C^1 diffeomorphism
between a cuboid $[\mathbf{a}, \mathbf{b}] \subset D'$ and its image $\mathbf{g}([\mathbf{a}, \mathbf{b}]) \subset D$.

That is, suppose there exists
a continuously differentiable bijection $[\mathbf{a}, \mathbf{b}] \ni \mathbf{u} \mapsto \mathbf{g}(\mathbf{u}) \in \mathbf{g}([\mathbf{a}, \mathbf{b}])$
whose inverse $\mathbf{g}([\mathbf{a}, \mathbf{b}]) \ni \mathbf{v} \mapsto \mathbf{g}^{-1}(\mathbf{v}) \in [\mathbf{a}, \mathbf{b}]$
is also continuously differentiable.

Suppose too that at each $\mathbf{u} \in [\mathbf{a}, \mathbf{b}]$, each partial derivative $\partial g_i / \partial x_j$
of the Jacobian matrix $\mathbf{J}(\mathbf{u})$ is positive.

Outline of a Justification in a Special Case, II

Now, given any $\mathbf{e} \gg \mathbf{0}$, the "useful lemma" can be applied, together with the fact that, with $\mathbf{c} = \mathbf{g}(\mathbf{a})$ and so $\mathbf{g}(\mathbf{a} + \epsilon\mathbf{e}) \approx \mathbf{c} + \epsilon\mathbf{J}(\mathbf{a})\mathbf{e}$, one has

$$\begin{aligned}\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^n} \int_{\mathbf{g}([\mathbf{a},\mathbf{a}+\epsilon\mathbf{e}])} f(\mathbf{x})\,\mathrm{d}\,\mathbf{x} &= \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^n} \int_{\mathbf{c}}^{\mathbf{c}+\epsilon\mathbf{J}(\mathbf{a})\mathbf{e}} f(\mathbf{x})\,\mathrm{d}\,\mathbf{x} \\ &= \mathrm{vol}_n(\mathbf{J}(\mathbf{a})\,\mathbf{diag}(\mathbf{e})) \cdot f(\mathbf{c}) \\ \text{and } \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon^n} \int_{\mathbf{a}}^{\mathbf{a}+\epsilon\mathbf{e}} f(\mathbf{g}(\mathbf{u}))\,\mathrm{d}\,\mathbf{u} &= \mathrm{vol}_n(\mathbf{diag}(\mathbf{e})) \cdot f(\mathbf{g}(\mathbf{a}))\end{aligned}$$

Outline of a Justification in a Special Case, II

Recall that, for any $n \times n$ matrix $\mathbf{A}$,
the volume $\mathrm{vol}_n(\mathbf{A})$ of the paralleliped $\sum_{j=1}^n [\mathbf{0}, \mathbf{a}^j]$
spanned by its columns $\mathbf{a}^j$ ($j \in \mathbb{N}_n$) equals $|\det \mathbf{A}|$.

It follows that

$$\begin{aligned}
\mathrm{vol}_n(\mathbf{J}(\mathbf{a})\,\mathbf{diag}(\mathbf{e})) &= |\det(\mathbf{J}(\mathbf{a})\,\mathbf{diag}(\mathbf{e}))| \\
&= |\det(\mathbf{J}(\mathbf{a})| \cdot |\det(\mathbf{diag}(\mathbf{e}))| \\
&= |\det(\mathbf{J}(\mathbf{a})| \cdot \mathrm{vol}_n(\mathbf{diag}(\mathbf{e}))
\end{aligned}$$

For this special case when each element of $\mathbf{J}(\mathbf{u})$ is positive,
this allows us to conclude that when the variables of integration
are transformed from $\mathbf{x} = \mathbf{g}(\mathbf{u})$ to $\mathbf{u}$,
the integrand $f(\mathbf{x})$ should be replaced,
not by $f(\mathbf{g}(\mathbf{u}))$, but by $f(\mathbf{g}(\mathbf{u})) \cdot |\det(\mathbf{J}(\mathbf{a})|$.

Outline

Carl-Friedrich Gauss (1777–1855) on a German Banknote

Portrait with (i) the graph of the "bell curve"; (ii) part of the University of Göttingen (where Gauss was a professor); (iii) the formula $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

The Gaussian Integral, I

For each $b \in \mathbb{R}_+$, let $S(b) := [-b, b]^2$
denote the Cartesian product
of the line interval $[-b, b]$ with itself.

That is, $S(b)$ is the solid square subset of $\mathbb{R}^2$
which is centred at the origin and has sides of length $2b$.

For each $b \in \mathbb{R}$ define $I(b) := \int_{-b}^{+b} e^{-x^2} \, \mathrm{d}\, x$.

Then the Fubini theorem implies that

$$\begin{aligned}
[I(b)]^2 &= \left(\int_{-b}^{+b} e^{-x^2} \, \mathrm{d}\, x\right) \left(\int_{-b}^{+b} e^{-y^2} \, \mathrm{d}\, y\right) \\
&= \int_{-b}^{+b} \left(\int_{-b}^{+b} e^{-y^2} \, \mathrm{d}\, y\right) e^{-x^2} \, \mathrm{d}\, x \\
&= \int_{-b}^{+b} \int_{-b}^{+b} e^{-x^2} e^{-y^2} \, \mathrm{d}\, x \, \mathrm{d}\, y \\
&= \int_{S(b)} e^{-x^2 - y^2} \, \mathrm{d}\, x \, \mathrm{d}\, y
\end{aligned}$$

The Gaussian Integral, II

Next, let $D(b) := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq b^2\}$
denote the disk of radius b centred at the origin.

Consider the transformation $(r, \theta) \mapsto (x, y) = (r \cos\theta, r \sin\theta)$
from polar to Cartesian coordinates.

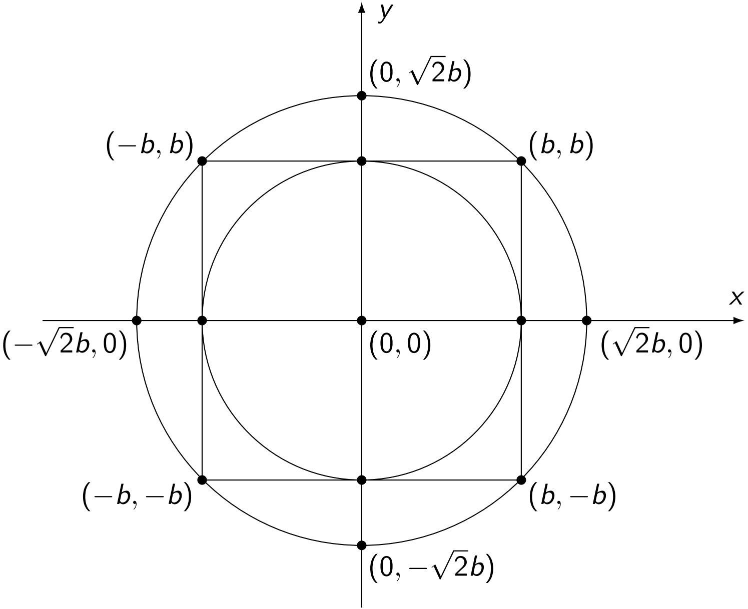
The Jacobian determinant of this transformation is

$$\begin{vmatrix} \partial x / \partial r & \partial x / \partial \theta \\ \partial y / \partial r & \partial y / \partial \theta \end{vmatrix} = \begin{vmatrix} \cos\theta & -r \sin\theta \\ \sin\theta & r \cos\theta \end{vmatrix} = r(\cos^2\theta + \sin^2\theta) = r$$

It follows that changing to polar coordinates
in the double integral $J(b) = \int_{D(b)} e^{-(x^2+y^2)} \, \mathrm{d}\, x \, \mathrm{d}\, y$
transforms it to

$$\begin{aligned} J(b) &= \int_0^b \int_0^{2\pi} r e^{-r^2} \, \mathrm{d}\, r \, \mathrm{d}\, \theta &&= \left(\int_0^b r e^{-r^2} \, \mathrm{d}\, r \right) \left(\int_0^{2\pi} 1 \, \mathrm{d}\, \theta \right) \\ &= \left[\, \Big|_0^b \left(-\tfrac{1}{2} e^{-r^2} \right) \right] 2\pi &&= \pi(1 - e^{-b^2}) \end{aligned}$$

Square with Inscribed and Circumscribed Circles



The Gaussian Integral, III

In the previous slide:

1. $S(b)$ is the square whose four corners are $(\pm b, \pm b)$;
2. $D(b)$ is the circular disk that is inscribed in $S(b)$;
3. $D(b\sqrt{2})$ is the circular disk that circumscribes $S(b)$.

It follows that $D(b) \subset S(b) \subset D(b\sqrt{2})$.

But the integrand $e^{-x^2-y^2}$ is non-negative, so

$$\begin{array}{rrcl} & J(b) & = & \int_{D(b)} e^{-(x^2+y^2)} \, \mathrm{d}\, x \, \mathrm{d}\, y \\ \leq & [I(b)]^2 & = & \int_{S(b)} e^{-(x^2+y^2)} \, \mathrm{d}\, x \, \mathrm{d}\, y \\ \leq & J(b\sqrt{2}) & = & \int_{D(b\sqrt{2})} e^{-(x^2+y^2)} \, \mathrm{d}\, x \, \mathrm{d}\, y \end{array}$$

From the previous definitions and calculations, it follows that

$$\pi(1 - e^{-b^2}) = J(b) \leq [I(b)]^2 \leq J(b\sqrt{2}) = \pi(1 - e^{-2b^2})$$

The Gaussian Integral, IV

Given $I(b) = \int_{-b}^{+b} e^{-x^2} \, \mathrm{d}\, x$,
we have shown that $\pi(1 - e^{-b^2}) \leq [I(b)]^2 \leq \pi(1 - e^{-2b^2})$.

As $b \to \infty$, both the lower bound $\pi(1 - e^{-b^2})$
and upper bound $\pi(1 - e^{-2b^2})$ converge to π.

From the squeezing principle, it follows that $[I(b)]^2 \to \pi$.

Because $I(b)$ is evidently positive, one has $I(b) \to \sqrt{\pi}$.

This proves that:

Theorem
The Gaussian integral $\int_{-\infty}^{+\infty} e^{-x^2} \, \mathrm{d}\, x$ *equals* $\sqrt{\pi}$*.*