The formula and its proof

Introduction

The formula

Consider a Jordan measurable set $E \subseteq \mathbb{R}^N$ and a function $f: \mathbb{R}^N \mapsto \mathbb{R}$ that is Riemann integrable on E. Now consider a transformation $T: \mathbb{R}^N \mapsto \mathbb{R}^N$:

$$T(\boldsymbol{u}) = T(u_1, \dots, u_N) = \begin{pmatrix} T_1(u_1, \dots, u_N) \\ \vdots \\ T_N(u_1, \dots, u_N) \end{pmatrix}. \tag{1}$$

We would like to have a formula of the form:

$$\int_E f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{T^{-1}(E)} \tilde{f}(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}. \tag{2}$$

It turns out that, under certain conditions on T, we have

$$\tilde{f}(\boldsymbol{u}) := f(T(\boldsymbol{u}))\,|\det(DT)(\boldsymbol{u})|\,\mathrm{d}\boldsymbol{u}. \tag{3}$$

Recall that DT is the Jacobian matrix

$$(DT)(\boldsymbol{u}) := \left(\frac{\partial T_i}{\partial u_j} \right) := \begin{pmatrix} \frac{\partial T_1}{\partial u_1} & \cdots & \frac{\partial T_1}{\partial u_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial T_N}{\partial u_1} & \cdots & \frac{\partial T_N}{\partial u_N} \end{pmatrix}. \tag{4}$$

Remark 1. Note that it is not at all obvious that such a formula should exist.

Remark 2. We only consider those T such that T^{-1} exists.

Plan of the proof

We will prove this formula through the following steps.

1. Measurability.

   We show that the set of Jordan measurable sets is invariant under invertible linear transformations.

2. Special linear transformations.

   We will consider two special cases: $L_1(\boldsymbol{u}) = D\boldsymbol{u}$ for a diagonal matrix D with non-negative entries; $L_2(\boldsymbol{u}) = O\,\boldsymbol{u}$ for an orthogonal matrix O.

3. Nonlinear transformations.

   We will prove the general case through approximating general nonlinear transformations locally by linear transformations.

Measurability under linear transformations

First we need to set up some results guaranteeing measurability of all sets under discussion.

Lemma 3. *Let $E \subseteq \mathbb{R}^N$. Let $L: \mathbb{R}^N \mapsto \mathbb{R}^N$ be linear and invertible. Then*

$$L(E^o) = (L(E))^o; \qquad L(\partial E) = \partial(L(E)); \qquad L(E^c) = (L(E))^c. \tag{5}$$

Proof. We prove the first one and leave the other two as exercises. Note that we only need to prove $L(E)^o \subseteq L(E^o)$. Once this is done, since L is invertible, we have

$$E^o = L^{-1}(L(E))^o \subseteq L^{-1}(L(E)^o) \Longrightarrow L(E^o) \subseteq L(L^{-1}(L(E)^o)) = L(E)^o \tag{6}$$

and equality follows.

Assume the contrary. That is there is $\boldsymbol{y} \in L(E)^o$, $\boldsymbol{y} = L(\boldsymbol{x})$ but $\boldsymbol{x} \notin E^o$. Thus we can find $\boldsymbol{x}_n \in E^c$ such that $\boldsymbol{x}_n \longrightarrow \boldsymbol{x}$. Since L is one-to-one, $\boldsymbol{x}_n \in E^c \Longrightarrow L(\boldsymbol{x}_n) \in L(E)^c$. But as L is continuous, we have $L(\boldsymbol{x}_n) \longrightarrow L(\boldsymbol{x}) = \boldsymbol{y}$ which means $\boldsymbol{y} \notin L(E)^o$. Contradiction. □

Exercise 1. Prove the other two equalities.

Lemma 4. *Let $E \subseteq \mathbb{R}^N$, $\mu(E) = 0$. Let $L: \mathbb{R}^N \mapsto \mathbb{R}^N$ be linear. Then $\mu(L(E)) = 0$.*

Exercise 2. Prove the above lemma.

Proposition 5. *Let $E \subseteq \mathbb{R}^N$ be Jordan measurable. Let $L: \mathbb{R}^N \mapsto \mathbb{R}^N$ be linear and invertible. Then $L(E)$ is Jordan measurable.*

Proof. All we need to prove is $\mu(\partial L(E)) = 0$. This follows immediately from Lemmas 3 and 4. □

Exercise 3. Find $E \subseteq \mathbb{R}^2$ measurable, and $L: \mathbb{R}^2 \mapsto \mathbb{R}$ linear, such that $L(E) \subseteq \mathbb{R}$ is not measurable.

Exercise 4. Let $E \subseteq \mathbb{R}^N$ be bounded (may not be measurable). Let $L: \mathbb{R}^N \mapsto \mathbb{R}^N$ be such that its matrix representation A is singular ($\det A = 0$). Prove that $\mu(L(E)) = 0$.

Diagonal transformations

Now we consider $L_1: \mathbb{R}^N \mapsto \mathbb{R}^N$ whose matrix representation is $D := \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_N \end{pmatrix}$ where $d_i > 0$ and all other entries are 0. The following lemmas are trivial to prove and left as exercises.

Lemma 6. *Let I be a compact interval. Then $\mu(L(I)) = d_1 \cdots d_N \, \mu(I)$.*

Lemma 7. *Let B be a simple graph. Then $\mu(L(B)) = d_1 \cdots d_N \, \mu(B)$.*

Lemma 8. *Let E be Jordan measurable. Then $\mu(L(E)) = d_1 \cdots d_N \, \mu(E)$.*

Lemma 9. *Let E be Jordan measurable and f be Riemann integrable on E. Then for any $\boldsymbol{x}_0 \in \mathbb{R}^N$,*

$$\int_E f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{L^{-1}(E - \boldsymbol{x}_0)} f(L\,\boldsymbol{u} + \boldsymbol{x}_0) \, d_1 \cdots d_N \, \mathrm{d}\boldsymbol{u}. \tag{7}$$

Orthogonal transformations

Next we consider $L_2: \mathbb{R}^N \mapsto \mathbb{R}^N$ whose matrix representation is an orthogonal matrix $O \in \mathbb{R}^{N \times N}$. We will prove the formula in this case via a different method.

Proposition 10. *For any Jordan measurable set $E \subseteq \mathbb{R}^N$, $\mu(L_2(E)) = \mu(E)$.*

Proof. Denote $\nu(E) := \mu(L_2(E))$. Set $c := \nu(I)$ where I is the unit interval $[0,1]^N$. It is clear that

$$\nu(I) = c\,\mu(I) \tag{8}$$

for any compact interval I.

Now by Lemma 4 we see that $\mu(E) = 0 \Longrightarrow \nu(E) = 0$. Furthermore it is obvious that $E_1 \cap E_2 = \varnothing \Longrightarrow \nu(E_1 \cup E_2) = \nu(E_1) + \nu(E_2)$. Thus for any simple graph B,

$$\nu(B) = c\,\mu(B). \tag{9}$$

From this we can prove that for any Jordan measurable E, $\nu(E) = c\,\mu(E)$.

Finally we determine c. Take the unit ball $B := \{\boldsymbol{x} \in \mathbb{R}^N | \, \|\boldsymbol{x}\| < 1\}$. Then $L_2(B) = B$. Consequently $c = 1$. □

Exercise 5. Prove that for any Jordan measurable E, $\nu(E) = c\,\mu(E)$.

Theorem 11. *Let E be Jordan measurable and f be Riemann integrable on E. Then*

$$\int_E f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{L_2^{-1}(E)} f(L_2(\boldsymbol{u}))\,\mathrm{d}\boldsymbol{u}. \tag{10}$$

Exercise 6. Prove the theorem.

Normal transformations

We consider invertible transformations T such that both T and T^{-1} are C^1, that is DT, DT^{-1} are both continuous.

We will prove that, for any $E \subseteq \mathbb{R}^N$ measurable, if T is C^1, then

$$\mu(T(E)) \leqslant \overline{\int_E} |\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{11}$$

Exercise 7. Explain why

$$\mu(E) \leqslant \overline{\int_{T^{-1}(E)}} |\det DT|\,\mathrm{d}\boldsymbol{u} \tag{12}$$

may not hold. (Hint: What if T is not invertible?)

This immediately leads to (we use $\overline{\int}$ be denote upper integral here), for invertible T with T, T^{-1} both C^1,

$$\overline{\int_E} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leqslant \overline{\int_{T^{-1}(E)}} f(T(\boldsymbol{u}))\,|\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{13}$$

Once this is done, applying the same formula to T^{-1} gives

$$\overline{\int_E} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geqslant \overline{\int_{T^{-1}(E)}} f(T(\boldsymbol{u}))\,|\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{14}$$

Thus

$$\overline{\int_E} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \overline{\int_{T^{-1}(E)}} f(T(\boldsymbol{u}))\,|\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{15}$$

Similar result can be proved for the lower integral and the conclusion follows from integrability of f.

Proposition 12. *Let $E \subseteq \mathbb{R}^N$ be measurable. Let T be C^1. Then*

$$\mu(T(E)) \leqslant \overline{\int_E} |\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{16}$$

Proof. Since T is C^1, DT is continuous and thus is uniformly continuous on E. Consequently there is a function $\delta: \mathbb{R}^+ \mapsto \mathbb{R}^+$ such that $\lim_{h\to 0+}\delta(h) = 0$ and

$$\forall i, j, \qquad \|\boldsymbol{u} - \boldsymbol{v}\| < h \Longrightarrow \|[(DT)_{ij}](\boldsymbol{u}) - [(DT)_{ij}](\boldsymbol{v})\| < \delta(h). \tag{17}$$

Denote

$$I_{i_1\ldots i_N,h} := [i_1\,h, (i_1+1)\,h] \times \cdots \times [i_N\,h, (i_N+1)\,h]. \tag{18}$$

Next define a piecewise constant function g_h through

$$g_h(\boldsymbol{u}) = \max_{I_{i_1\cdots i_N,h}} |\det DT| \qquad \forall \boldsymbol{u} \in I^o_{i_1\ldots i_N,h}. \tag{19}$$

and $g_h(\boldsymbol{u}) = \max_E |\det DT|$ everywhere else. Then we have

$$\lim_{h\to 0} \int_E g_h(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u} = \overline{\int_E} |\det DT|\,\mathrm{d}\boldsymbol{u}. \tag{20}$$

On the other hand, fix one cell $I_{i_1\ldots i_N,h}$. For simplicity of presentation we denote it by I_h. Take any $\boldsymbol{u}_0 \in I_h$. Denote by A the matrix representation of $(DT)(\boldsymbol{x}_0)$. By a theorem in linear algebra, we have

$$A = O_1\,D\,O_2 \tag{21}$$

where O_1, O_2 are orthogonal while D is diagonal with positive diagonal entries. Consider the transformation $\tilde{T}(\boldsymbol{u}) := O_1^{-1}\,T(\boldsymbol{u})\,O_2^{-1}$. Then we have

$$\mu\big(\tilde{T}(I_h)\big) = \mu(T(I_h)) \tag{22}$$

and

$$\big(D\,\tilde{T}\,\big)(\boldsymbol{u}_0) = D. \tag{23}$$

By MVT (note that here we have to apply MVT to each component of $\tilde{T}$) we have

$$\forall \boldsymbol{u} \in I_h, \qquad \big\|\tilde{T}(\boldsymbol{u}) - \tilde{T}(\boldsymbol{u}_0) - D\,\boldsymbol{u}_0\big\| \leqslant \delta(h)\,\|\boldsymbol{u} - \boldsymbol{u}_0\| \leqslant \delta(h)\,h. \tag{24}$$

Thus

$$\begin{aligned}\mu(T(I_h)) = \mu\big(\tilde{T}(I_h)\big) &\leqslant (d_1 + 2\,\delta(h)) \cdots (d_N + 2\,\delta(h))\,\mu(I_h) \\ &\leqslant (1 + C\,\delta(h))\,|\det D|\,\mu(I_h) \\ &\leqslant (1 + C\,\delta(h)) \max_{I_h} |\det(DT)|\,\mu(I_h) \\ &= (1 + C\,\delta(h)) \int_{I_h} g_h(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}. \end{aligned} \tag{25}$$

Here the constant C is independent of h. This leads to

$$\mu(T(E)) \leqslant (1 + C\,\delta(h)) \int_E g_h(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}. \tag{26}$$

Taking limit $h \longrightarrow 0$ we obtain the desired conclusion. □

Exercise 8. Complete the proof for the change of variables theorem.