LECTURE 10

ZIQUAN YANG

Let $g : U \to \mathbb{R}^m$ be a function on some open $U \subseteq \mathbb{R}^n$. We say that g is differentiable at a point $\mathbf{x} \in U$ if there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$g(\mathbf{x}) - g(\mathbf{x}') = T(\mathbf{x} - \mathbf{x}') + o(\|\mathbf{x} - \mathbf{x}'\|)$$

as $\mathbf{x}'$ approaches $\mathbf{x}$. Recall that the little-o notation means that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $0 < \|\mathbf{x}' - \mathbf{x}\| < \delta$,

$$|\big(g(\mathbf{x}) - g(\mathbf{x}')\big) - T(\mathbf{x} - \mathbf{x}')| < \varepsilon \|\mathbf{x} - \mathbf{x}'\|.$$

If such T exists, then we define T to be the Jacobian of g at $\mathbf{x}$, and denote it by $J_g(\mathbf{x})$. You should think of the Jacobian as the best approximation of g at $\mathbf{x}$ by a linear transformation.

Below we do not distinguish a linear transformation and its associated matrix, because we are always using the standard basis.

Theorem 1. *Suppose that $g : U \to \mathbb{R}^n$ is given by*

$$\begin{bmatrix} g_1(x_1, \cdots, x_n) & g_2(x_1, \cdots, x_n) & \cdots & g_m(x_1, \cdots, x_n) \end{bmatrix}^T.$$

If g is differentiable at $\mathbf{x}$, then all its partial derivatives $\partial g_i / \partial x_j$ exists at $\mathbf{x}$, and

$$J_g = \begin{bmatrix} \partial g_1/\partial x_1 & \partial g_1/\partial x_2 & \cdot & \partial g_1/\partial x_n \\ \partial g_2/\partial x_1 & \partial g_2/\partial x_2 & \cdots & \partial g_2/\partial x_n \\ \vdots & \vdots & \cdots & \vdots \\ \partial g_n/\partial x_1 & \partial g_n/\partial x_2 & \cdots & \partial g_n/\partial x_m \end{bmatrix}$$

at $\mathbf{x}$. Conversely, if for all $\mathbf{x}'$ in an open neighborhood of $\mathbf{x}$, $\partial g_i/\partial x_j(\mathbf{x}')$ in the above exists and is continous at $\mathbf{x}$, then f is differentiable.

Note that for each $\mathbf{x} \in U$, $J_g(\mathbf{x})$ is a $m \times n$ matrix. Therefore, $J_g : \mathbf{x} \mapsto J_g(\mathbf{x})$ itself is a function from U to $\mathrm{Mat}_{m\times n} \simeq \mathbb{R}^{mn}$. Let us call a function g *continously differentiable* on U if for g is differentiable at every $\mathbf{x} \in U$ (i.e., $J_g(\mathbf{x})$ exists) and $J_g : U \to \mathrm{Mat}_{m\times n}$ is continous. A special case of Theorem 1 is the following, which is usually what you need in practice.

Corollary 2. *A function $g : U \to \mathbb{R}^m$ with components g_i's is continously differentiable over U if and only if all its partial derivatives $\partial g_i/\partial x_j$ exists and is continous over U.*

Finally we are ready to introduce the general change-of-variable formula for integrals.

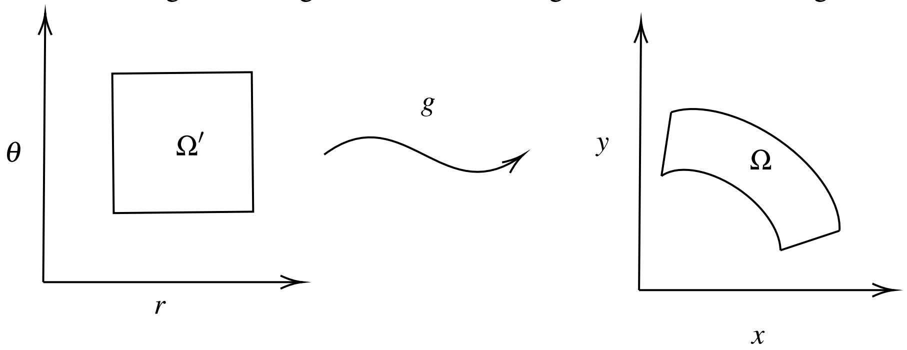
Theorem 3. *Suppose that $\Omega \subseteq \mathbb{R}^n$ is a closed bounded region and $f : \Omega \to \mathbb{R}$ is a continuous function. Suppose that for some closed $\Omega' \subseteq \mathbb{R}^n$ there exists a bijection function $g : \Omega' \to \Omega$ such that both g and g^{-1} are continously differentiable on the interior of Ω. Then f is integrable and*

$$\int_\Omega f = \int_{\Omega'} (f \circ g)|\det(J_g)|.$$

Let us consider the example of polar coordinates. We can think of it as applying an invertible function g from the (r, θ)-plane to the (x, y)-plane. More precisely, in order for g to be a bijection, we should only use the region $(0, \infty) \times (0, 2\pi)$ on the (r, θ)-plane, which corresponds to the open subset $\mathbb{R}^2 \setminus \{(x, 0) \mid$

Date: February 18, 2025.

$x \geq 0\}$ on the (x,y)-plane. When doing integrals, we often ignore this because oftentimes the deleting an area zero region from the region of integration does not change the value of the integral.



The function g is given by $(r,\theta) \mapsto (r\cos(\theta), r\sin(\theta))$.[1] Then the Jacobian is

$$J_g = \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix}$$

One quickly notice that this is the same as the rotation-by-θ matrix, except one of the columns is scaled by r, and $\det(J_g) = r$. This accounts for the "r" factor in the change-of-variable formula

$$dxdy = rdrd\theta.$$

Now let us try to account for the change-of-variable formula for the spherical coordinates. There the bijection g to be considered is $(0,\infty) \times (0,2\pi) \times (0,\pi) \to \mathbb{R}^3 \smallsetminus \big(\{(x,0,z) \mid x \geq 0\} \cup z\text{-axis}\big)$ given by

$$(\rho,\theta,\phi) \mapsto (\rho\sin(\phi)\cos(\theta), \rho\sin(\phi)\sin(\theta), \rho\cos(\phi)).$$

Let me leave it as an exercise to check that $\det(J_g) = \rho^2\sin(\phi)$. Since $\phi \in (0,\pi)$, $\sin(\phi) > 0$, so $|\rho^2\sin(\phi)| = \rho^2\sin(\phi)$.

Proof of Theorem 3 (Sketch). You do not need to know the full detail of the proof of theorem, which is beyond the scope of the course. However, let me offer the basic ideas. Assume for simplicity that g and g^{-1} extend to continuous functions over Ω—you can always reduce to this case.

First, we show the theorem when $f = 1$, in which case the theorem reduces to

$$\mathrm{vol}(\Omega) = \int_{\Omega'} |\det(J_g)| \tag{1}$$

Fix any $\varepsilon > 0$. We want to show that

$$\left|\mathrm{vol}(\Omega) - \int_{\Omega'} |\det(J_g)|\right| < M\varepsilon$$

for some bounded $M > 0$. Then since ε is arbitrary, we must have equality. Let Q be a very fine partition of Ω'. Then we have

$$\mathrm{vol}(\Omega) = \sum_{R\in Q} \mathrm{vol}(g(R)),$$

because now $\{g(R) : R \in Q\}$ becomes a partition for Ω. Choose a point x_R for each $R \in Q$. When Q is fine enough, we can make sure that for each $R \in Q$,

$$\Big|\mathrm{vol}(g(R)) - \mathrm{vol}(J_g(x_R)\cdot R)\Big| < \mathrm{vol}(R)\varepsilon.$$

Let me explain this: By definition, $J_g(x_R)$ is the linear transformation centered at x_R which best approximates the function g. Hence $g(R)$ and $J_g(x_R)\cdot R$ are "roughly the same", in the precise sense above. The reason that we have a factor "$\mathrm{vol}(R)$" on the right hand side is that in the definition of Jacobian we have a small o notation (as opposed to big O, for instance). As $J_g(x_R)$ is a linear transformation, we know from last time that

$$\mathrm{vol}(J_g(x_R)\cdot R) = |\det(J_g(x_R))|\mathrm{vol}(R).$$

[1] Last time I wrote $x \mapsto r\cos(\theta), y \mapsto r\sin(\theta)$. This is really a habit from algebraic geometry (which I normally do), and my apologies if you get confused. But hopefully it is clear what I mean.

On the other hand, by the continuity properties of J_g, we know that when R is very small, $J_g(x_R)$ is very close to $J_g(x)$ for $x \in R$. Therefore, after taking integrals we obtain

$$\Big| \int_R |\det(J_g(x_R))| - \int_R |\det(J_g)| \Big| < \mathrm{vol}(R)\varepsilon. \tag{2}$$

Putting these together, we obtain that

$$\begin{aligned}\Big|\mathrm{vol}(\Omega) - \int_{\Omega'} |\det(J_g)|\Big| &= \Big| \sum_{R\in Q} \mathrm{vol}(g(R)) - \int_R |\det(J_g)|\Big| \\ &\leq \sum_{R\in Q} \Big|\mathrm{vol}(R) - \int_R |\det(J_g)|\Big| \\ &\leq \sum_{R\in Q} 2\mathrm{vol}(R)\varepsilon = 2\mathrm{vol}(\Omega')\varepsilon.\end{aligned}$$

Hence we have completeted the proof when $f = 1$. Note that this generalizes to an arbitrary constant, not necessarily 1.

Now we treat the general f. Again take any $\varepsilon > 0$. Using the (uniform) continuity of f and g, when Q is fine enough, we can make sure that for every $R \in Q$ and $x \in R$, $|f(g(x_R)) - f(g(x))| < \varepsilon$. This implies that

$$\Big| \int_{g(R)} f - \int_{g(R)} f(g(x_R)) \Big| < \mathrm{vol}(g(R))\varepsilon.$$

As $f(x_R)$ is a constant over R, and we have already checked the theorem for constant functions, we have (letting R below play the role of Ω above)

$$\int_{g(R)} f(g(x_R)) = \int_R f(g(x_R))|\det(J_g)|.$$

Assume that $|f(x)| \leq N$ for all $x \in \Omega$, where N is some sufficiently large number. Such N exists again because of the uniform continuity of f.

$$\Big|\int_\Omega f - \int_{\Omega'} (f\circ g)|\det(J_g)|\Big| = \Big|\sum_{R\in Q}\Big(\int_{g(R)} f - \int_R (f\circ g)|\det(J_g)|\Big)\Big| \tag{3}$$

$$\leq \sum_{R\in Q}\Big|\int_{g(R)} f - \int_R (f\circ g)|\det(J_g)|\Big| \tag{4}$$

$$\leq \sum_{R\in Q}\Big(\Big|\int_{g(R)} f(g(x_R)) - \int_R (f\circ g)|\det(J_g)|\Big| + \mathrm{vol}(g(R))\varepsilon\Big) \tag{5}$$

$$\leq \mathrm{vol}(\Omega)\varepsilon + \sum_{R\in Q}\Big|\int_R \big(f(g(x_R)) - (f\circ g)\big)|\det(J_g)|\Big| \tag{6}$$

$$\leq \mathrm{vol}(\Omega)\varepsilon + \sum_{R\in Q} \varepsilon \int_R |\det(J_g)|\Big| \tag{7}$$

$$\leq \mathrm{vol}(\Omega)\varepsilon + \varepsilon\Big(\int_R |\det(J_g)|\Big) \leq 2\mathrm{vol}(\Omega)\varepsilon. \tag{8}$$

Hence we are done. □

Remark 4. When we were defining the integral of a function f over a rectangular region Ω, we used the rectangular partitions. In fact, if f is integrable, then a foritiori there is nothing special about using rectangular partitions. We can consider partitions P of Ω into any union of closed bounded subsets, i.e., P is a set of closed and bounded subsets of Ω such that for $\Omega = \cup_{R\in P} R$ and for any $R \neq R' \in P$, $\mathrm{vol}(R \cap R') = 0$. We define $\|P\|$ to be the maximum value of $\{\mathrm{vol}(R) \mid R \in P\}$. Then we still have

$$\int_\Omega f = \lim_{\|P\|\to 0} L(P, f).$$