Math 346 Lecture #18
8.7 Change of Variables

Change of Variables for integrals of real-valued functions of a real variable is the familiar substitution formula or u-substitution: for an continuous function $f : [a, b] \to \mathbb{R}$, there holds

$$\int_c^d f(g(s))g'(s)\ ds = \int_a^b f(\tau)\ d\tau$$

where $g : [c, d] \to [a, b]$ is a bijection with continuous derivative that doesn't change sign, and $g(c) = a$ and $g(d) = b$.

An analogous substitution formula or change of variables holds in higher dimensions.

8.7.1 Diffeomorphisms

Recall that a homeomorphism is a continuous bijection $h : U \to V$ for U and V open sets in Banach spaces where the inverse map $h^{-1} : V \to U$ is continuous on V.

A homeomorphism is a diffeomorphism when h and h^{-1} are continuously differentiable. We state the formal definition in the Banach space $\mathbb{R}^n$.

Definition 8.7.1. For open sets U and V in $\mathbb{R}^n$, a function $\Psi : U \to V$ is called a C^1-diffeomorphism if Ψ is a C^1 bijection whose inverse Ψ^{-1} is C^1.

Note. Recall that the Inverse Function Theorem shows that a C^1 function $F : U \to V$ is a "local" diffeomorphism when $DF(\mathrm{u}_0) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$ has bounded inverse, i.e., there exist open U_0 containing u_0 and an open V_0 containing $F(\mathrm{u}_0)$ such that F restricted to a function from U_0 to V_0 is a C^1-diffeomorphism.

Note. A C^k-diffeomorphism $\Psi : U \to V$ is a C^k bijection whose inverse $\Psi^{-1} : V \to U$ is C^k. A smooth or C^∞-diffeomorphism is a bijection $\Psi : U \to V$ that is C^k for all $k \in \mathbb{N}$ and whose inverse Ψ^{-1} is C^k for all $k \in \mathbb{N}$. By a diffeomorphism $\Psi : U \to V$ we mean a C^1-diffeomorphism.

Examples (in lieu of 8.7.3 and 8.7.4). (i) The function $f : (1, 2) \to (1, \sqrt{2})$ defined by

$$f(x) = \sqrt{x}$$

is a diffeomorphism because f is a continuously differentiable bijection whose inverse map $f^{-1}(y) = y^2$ from $(1, \sqrt{2})$ to $(1, 2)$ is continuously differentiable. In fact, f is a C^∞-diffeomorphism.

(ii) The function

$$g(x) = \arctan(x)$$

from $\mathbb{R}$ to $(-\pi/2, \pi/2)$ is a diffeomorphism because it is a continuously differentiable bijection whose inverse $g^{-1}(y) = \tan(y)$ is continuously differentiable. Here again, g is a C^∞-diffeomorphism.

(iii) The function

$$f(x) = x^3 - 3x$$

is not a diffeomorphism from $\mathbb{R}$ to $\mathbb{R}$ because it is not injective: $f(-\sqrt{3}) = 0 = f(\sqrt{3})$.

However, if we restrict f to the open interval $U = (-1, 1)$ (the interval between the two critical points of f) and, by abuse of notation, refer to the restriction of f to U also by f, then f is a continuously differentiable bijection from U to $V = (-2, 2)$.

To see that f^{-1} is a continuously differentiable, we have $f'(u) = 3u^2 - 3 \neq 0$ for each $u \in (-1, 1)$, so by the Inverse Function Theorem, f is a "local" diffeomorphism on a neighbourhood U_0 of each $u \in (-1, 1)$.

Thus we can differentiate $(f^{-1} \circ f)(x) = x$ with $x \in U_0$ to get

$$Df^{-1}(f(x)) = \frac{1}{Df(x)}$$

which since $Df(x) \neq 0$ for all $x \in U_0$ gives the continuous differentiability of f^{-1}.

Since $u \in U$ is arbitrary, we get that f^{-1} is continuously differentiable, and hence f is diffeomorphism.

(iv) The function $h : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$h(x, y) = (x^3 - y^2, xy)$$

is a "local" diffeomorphism from an open set U containing the point $(1, 2)$ to an open set V containing $h(1, 2) = (-3, 2)$ because

$$Dh(1, 2) = \begin{bmatrix} 3x^2 & -2y \\ y & x \end{bmatrix}_{x=1,y=2} = \begin{bmatrix} 3 & -4 \\ 2 & 1 \end{bmatrix}$$

has a nonzero determinant of 11.

But h is not a diffeomorphism from $\mathbb{R}^2$ to $\mathbb{R}^2$ because h is not injective, i.e.,

$$h(0, 2) = (-4, 0) = h(0, -2).$$

8.7.2 Change of Variables

To anticipate the form of change of variables for integration in higher dimensions we rewrite the substitution formula.

Suppose that g is a diffeomorphism on an open set containing $[c, d]$ for which the continuous g' doesn't change sign.

This gives two cases: g' is positive and g' is negative.

When g' is positive then $g(c) < g(d)$ and we have

$$\int_{g([c,d])} f = \int_{g(c)}^{g(d)} f(\tau)\, d\tau = \int_c^d f(g(s))g'(s)\, ds = \int_{[c,d]} (f \circ g)|g'|.$$

When g' is negative then $g(d) < g(c)$ and we have

$$\begin{aligned} \int_{g([c,d])} f &= \int_{g(d)}^{g(c)} f(\tau)\, d\tau = -\int_{g(c)}^{g(d)} f(\tau)\, d\tau = -\int_d^c f(g(s))g'(s)\, ds \\ &= -\int_{[c,d]} (f \circ g)g' = \int_{[c,d]} (f \circ g)|g'| \end{aligned}$$

Theorem 8.7.5 (Change of Variables Theorem). For open sets U and V in $\mathbb{R}^n$ suppose $\Psi : U \to V$ is a diffeomorphism. If $X \subset U$ is measurable, then $Y = \Psi(X)$ is measurable. If $f \in L^1(Y, \mathbb{R})$, then

1. $(f \circ \Psi)|\det(D\Psi)| \in L^1(X, \mathbb{R})$ and
2. $\displaystyle\int_Y f = \int_X (f \circ \Psi)|\det(D\Psi)|$.

The proof of this is in Chapter 9.

Remark 8.7.6. To see what is happening here in the Change of Variables Theorem, take Ψ to be an invertible linear operator on $\mathbb{R}^n$, the measurable set $X = [\mathrm{a}, \mathrm{b}]$ a compact n-interval in $\mathbb{R}^n$, and $f = 1$. Then $D\Psi(\mathrm{x}) = \Psi$ for all $x \in \mathbb{R}^n$ so that

$$\lambda(\Psi([\mathrm{a},\mathrm{b}])) = \int_{\Psi([\mathrm{a},\mathrm{b}])} 1 = \int_{[\mathrm{a},\mathrm{b}]} (1 \circ \Psi)|\det(\Psi)| = |\det(\Psi)| \int_{[\mathrm{a},\mathrm{b}]} 1 = |\det(\Psi)|\lambda([\mathrm{a},\mathrm{b}]).$$

This says the measure of $\Psi([\mathrm{a},\mathrm{b}])$ is precisely the product of $|\det(\Psi)|$ and the measure of $[\mathrm{a},\mathrm{b}]$. Using a SVD $\Psi = U\Sigma V^{\mathrm{H}}$, the value of $|\det(\Psi)|$ is the product of the diagonal entries of Σ (the singular values of Ψ) because the determinants of the orthornormal matrices have modulus 1. Each singular value of Ψ scales the i^{th} standard basis vector by σ_i, so that the overall volume scales by the product of the singular values. The value of $|\det(\Psi)|$ gives the change of volume.

Corollary 8.7.8. For a diffeomorphism $\Psi : U \to V$, if $E \subset U$ is a set of measure zero, then $\Psi(E)$ is a set of measure zero.

Proof. Suppose E has measure zero. Then by the Change of Variables Theorem we have that $\Psi(E)$ is measurable and

$$\int_{\Psi(E)} 1 = \int_E |\det(D\Psi)| = \int_U \chi_E |\det(D\Psi)| = 0$$

because $\chi_E = 0$ a.e. on U.

The measurability of $\Psi(E)$ implies that the function $\chi_{\Psi(E)}$ is measurable.

By Exercise 8.20 (and 8.14), the set $\Psi(E)$ has measure zero. □

Nota Bene. If $\Psi : U \to V$ is only a homeomorphism, then it is possible for $\Psi(E)$ to have positive measure when E has measure zero, and for $\Psi(E)$ to have measure zero when E has positive measure. A homeomorphism $\Psi : [0,1] \to [0,1]$ is constructed in Math 541 for which $\Psi(E)$ has measure 1 for a certain set E of measure 0, and passing to the complement of E in $[0,1]$, that $\Psi([0,1] - E)$ has measure zero with $[0,1] - E$ having measure 1. Diffeomorphisms cannot do these strange things to sets of measure zero.

8.7.3 Polar Coordinates

Polar coordinates on $\mathbb{R}^2$ are determined by the diffeomorphism $\Psi : (0,\infty) \times (0, 2\pi) \to V$ given by

$$\Psi(r,\theta) = (r\cos\theta, r\sin\theta)$$

where

$$V = \mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2 : x \geq 0\}.$$

To see that Ψ is indeed a diffeomorphism, we recognize that it is bijective with inverse given by

$$\Psi^{-1}(x,y) = (\sqrt{x^2+y^2}, \theta)$$

where

$$\theta = \begin{cases} \arctan(y/x) & \text{if } x > 0 \text{ and } y > 0, \\ \pi/2 & \text{if } \ x = 0 \text{ and } y > 0, \\ \pi + \arctan(y/x) & \text{if } x < 0, \\ 3\pi/2 & \text{if } x = 0 \text{ and } y < 0, \\ 2\pi - \arctan(y/x) & \text{if } x > 0 \text{ and } y < 0. \end{cases}$$

One can check (and you should) that θ is at least a continuous function of (x,y).

The function Ψ is differentiable with continuous derivative

$$D\Psi(r,\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix}.$$

Since $\det(D\Psi(r,\theta)) = r > 0$ on $(0,\infty)\times(0,2\pi)$, the Inverse Function Theorem guarantees that the inverse Ψ^{-1} is also continuously differentiable.

Thus Ψ is a diffeomorphism.

For any measurable $A \subset (0,\infty) \times (0,2\pi)$, the image $B = \Psi(A) \subset V$ is measurable and for any measurable $f : B \to \mathbb{R}$ we have

$$\iint_B f(x,y)\ dxdy = \int_B f = \int_A (f \circ \Psi)r = \iint_A f(r\cos\theta, \sin\theta)r\ drd\theta.$$

We can extend the equality to $[0,\infty) \times [0,2\pi]$ because the rays defined by $\theta = 0$ and $\theta = 2\pi$ have measure zero and contribute zero to the integrals.

8.7.4 Spherical and Hyperspherical Coordinates

We recall spherical coordinates on $\mathbb{R}^3$, and then extend this to hyperspherical coordinates in $\mathbb{R}^n$ for $n \geq 4$.

Definition 8.7.12. Let $U = (0,2\pi) \times (0,\pi) \times (0,\infty)$. Spherical coordinates (θ,ϕ,r) on $\mathbb{R}^3$ are defined by

$$S(\theta,\phi,r) = (r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\psi).$$

One verifies that S is a diffeomorphism by computing

$$DS(\theta,\phi,r) = \begin{bmatrix} -r\sin\phi\sin\theta & r\cos\phi\cos\theta & \sin\phi\cos\theta \\ r\sin\phi\cos\theta & r\cos\phi\sin\theta & \sin\phi\cos\theta \\ 0 & -r\sin\phi & \cos\phi \end{bmatrix}$$

and

$$|\det(DS(\theta,\phi,r))| = r^2\sin\phi > 0.$$

One can check that S is C^1 and bijective with range $V = S(U)$.

Since $DS(\theta, \phi, r) \neq 0$ on U, the Inverse Function Theorem guarantees that the inverse function is C^1.

Thus S is a diffeomorphism.

Applying the Change of Variables Theorem to S gives for any measurable subset X of U and integrable $f : X \to \mathbb{R}$ the formula

$$\begin{aligned}\int_{S(X)} f = \iiint_{S(X)} f(x,y,z)\ dxdydz &= \int_X (f \circ S)\, r^2 \sin\phi \\ &= \int_X f(r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi) r^2 \sin\phi\ drd\phi d\theta.\end{aligned}$$

We can extend this formula to closure of U because the $\overline{U} \setminus U$ is a set of measure zero.

Definition 8.7.14. Hyperspherical coordinates $(\phi_1, \ldots, \phi_{n-2}, \phi_{n-1}, r)$ on $\mathbb{R}^n$ are given the function

$$\begin{aligned}\Psi &: (0,\pi) \times \cdots \times (0,\pi) \times (0, 2\pi) \times (0,\infty) \\ &\to r \begin{bmatrix} \cos(\phi_1) \\ \sin(\phi_1)\cos(\phi_2) \\ \sin(\phi_1)\sin(\phi_2)\cos(\phi_3) \\ \vdots \\ \sin(\phi_1)\sin(\phi_2)\cdots\sin(\phi_{n-2})\cos(\phi_{n-1}) \\ \sin(\phi_1)\sin(\phi_2)\cdots\sin(\phi_{n-2})\sin(\phi_{n-1}) \end{bmatrix} \in \mathbb{R}^n,\end{aligned}$$

where there are $n-2$ copies of $(0,\pi)$ in the domain U of Ψ.

A straightforward but tedious computation shows that

$$\det\big(D\Psi(\phi_1, \ldots, \phi_{n-2}, \phi_{n-1}, r)\big) = r^{n-1} \sin^{n-2}(\phi_1) \sin^{n-3}(\phi_2) \cdots \sin(\phi_{n-2}).$$

One shows that Ψ is a C^1 bijection from U to $V = \Psi(U)$.

Since $D\Psi \neq 0$, the Inverse Function Theorem shows that Ψ^{-1} is C^1 on V.

Thus Ψ is a diffeomorphism from U to V.

Applying the Change of Variables Theorem to Ψ gives for any measurable subset X of U and integrable $f : X \to \mathbb{R}$ the formula

$$\int_V f = \int_U (f \circ \Psi)\, r^{n-1} \sin^{n-2}(\phi_1) \sin^{n-3}(\phi_2) \cdots \sin(\phi_{n-2}).$$

You will use hyperspherical coordinates to find the volume of the unit ball (in the 2-norm) in $\mathbb{R}^n$ (Exercise 8.36).