Integration by Substitution with multiple variables

Branden Xia

Jun 28, 2026

Linear Algebra 2025-2026

Contents

1. Introduction

Integration by substitution is frequently used in integrations to simplify the integrand and make the integral easier to evaluate. Typical examples are trig-substitution and the Gaussian integral. In this project, I'll first explore an example in 2D, and then generalize it to n-dimensions. After that, I'll try to explore some applications of this formula when working on integrals.

2. Notations and Definitions

Before diving into the main content, let's clarify some notations and definitions that will be used throughout this project.

2.1. Partial Derivative

Partial derivative is defined as the derivative of a multi-variable function with respect to one variable while keeping the other variables constant.

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = \lim_{h \to 0} \frac{f(\boldsymbol{x}_0 + h \cdot \boldsymbol{e}_i) - f(\boldsymbol{x}_0)}{h}$$

2.2. Jacobian

The Jacobian matrix is a matrix of all first-order partial derivatives of a vector-valued function.

Given a function $\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n$, its Jacobian an $n \times m$ matrix:

$$\mathbf{J}_{\boldsymbol{f}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

2.3. Integral over n-dimensional Region

An integral over an n-dimensional region can be defined as the limit of Riemann sums:

$$\int_D f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{\max V_i \to 0} \sum_i f(\boldsymbol{x}_i) \cdot V_i$$

2.4. C^1 Transformation

A C^1 transformation is defined by a function $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ where U is an open set in $\mathbb{R}^n$ with the following properties:

1. The function $\boldsymbol{\varphi}$ is continuously differentiable on U, which means all first-order partial derivatives of $\boldsymbol{\varphi}$ exist and are continuous on U.
2. The function $\boldsymbol{\varphi}$ is injective on U, which means for any two distinct points $\boldsymbol{u}_1, \boldsymbol{u}_2$ in U, we have $\boldsymbol{\varphi}(\boldsymbol{u}_1) \neq \boldsymbol{\varphi}(\boldsymbol{u}_2)$.

3. Deriving the Formula

Before working on the general case, let's consider some cases in lower dimensions.

3.1. Example: 2D Coordinate Transformation

Consider three points: (u_0, v_0), $(u_0 + \Delta u, v_0)$, and $(u_0, v_0 + \Delta v)$. They can define a rectangle on the uv-plane. Their volume can be represented by two vectors:

$$\boldsymbol{u} = \begin{pmatrix} \Delta u \\ 0 \end{pmatrix} \text{ and } \boldsymbol{v} = \begin{pmatrix} 0 \\ \Delta v \end{pmatrix}$$

Their volume is product of the lengths of the two vectors:

$$V_{uv} = \Delta u \cdot \Delta v$$

Consider a C^1 transformation from the uv-plane to the xy-plane:

$$\boldsymbol{T}(u, v) = \langle x(u, v), y(u, v) \rangle$$

The two vectors on the xy-plane corresponding to the two vectors on the uv-plane are given by linear approximations:

$$\boldsymbol{T}(u_0 + \Delta u, v_0) - \boldsymbol{T}(u_0, v_0) \approx \frac{\partial \boldsymbol{T}}{\partial u}(u_0, v_0) \cdot \Delta u$$
$$\boldsymbol{T}(u_0, v_0 + \Delta v) - \boldsymbol{T}(u_0, v_0) \approx \frac{\partial \boldsymbol{T}}{\partial v}(u_0, v_0) \cdot \Delta v$$

Therefore, the volume on the xy-plane is given by the determinant of the matrix formed by these two vectors:

$$V_{xy} \approx \left| \det \begin{bmatrix} \frac{\partial x}{\partial u}(u_0, v_0) \cdot \Delta u & \frac{\partial x}{\partial v}(u_0, v_0) \cdot \Delta v \\ \frac{\partial y}{\partial u}(u_0, v_0) \cdot \Delta u & \frac{\partial y}{\partial v}(u_0, v_0) \cdot \Delta v \end{bmatrix} \right| \approx |\det \mathbf{J}_{\boldsymbol{T}}| \cdot \Delta u \, \Delta v$$

where $\mathbf{J}_{\boldsymbol{T}}$ is the Jacobian matrix of the transformation $\boldsymbol{T}$.

Now, consider a double integral on the xy-plane:

$$\begin{aligned} \int_D f(x, y) \, \mathrm{d}A &= \lim_{\max V_{xy} \to 0} \sum_i f(x_i, y_i) \cdot V_{(xy)_i} \\ &= \lim_{\Delta u \to 0} \lim_{\Delta v \to 0} \sum_i f(x(u_i, v_i), y(u_i, v_i)) \cdot |\det \mathbf{J}_{\boldsymbol{T}}(u_i, v_i)| \cdot \Delta u \, \Delta v \\ &= \int_{D'} f(x(u, v), y(u, v)) \cdot |\det \mathbf{J}_{\boldsymbol{T}}(u, v)| \, \mathrm{d}u \mathrm{d}v \end{aligned}$$

where D' is the region on the uv-plane that corresponds to the region D on the xy-plane under the transformation $\boldsymbol{T}$.

3.2. General Case

Theorem. (Change of variable) Let $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ be a C^1 transformation where U is an open set in $\mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function, we have:

$$\int_{\boldsymbol{\varphi}(U)} f(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = \int_U f(\varphi(\boldsymbol{u})) \cdot \left| \det \mathbf{J}_{\boldsymbol{\varphi}}(\boldsymbol{u}) \right| \mathrm{d}\boldsymbol{u}$$

The proof of this theorem follows the same idea as the 2D case, and can be proven similarly.

4. Application: Simplifying Integral

Now let's apply the formula derived above to derive some interesting results.

Given a function $f : \mathbb{R} \to \mathbb{R}$, whose anti-derivative F is known, and another function $g : \mathbb{R}^n \to \mathbb{R}$. Let D be an open set in $\mathbb{R}^n$. Consider the following integral:

$$\int_D (f \circ g)(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v}$$

Note that we can safely assume g is dependent on all dimensions, since, if there exists a dimension j that g is independent of, we can define $g' : \mathbb{R}^{n-1} \to \mathbb{R}$ as following:

$$g'(\boldsymbol{x}) = g\big(\boldsymbol{x} + 0 \cdot \boldsymbol{e_j}\big) \text{ where } \boldsymbol{x} \in \mathbb{R}^{n-1}$$

Similarly, the integral can be rewritten as:

$$\begin{aligned}\int_D (f \circ g)(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} &= \int_{u_k} \mathrm{d}u_k \int_{u_i, i \neq k} (f \circ g)\left(\sum_{i \neq k} v_i \boldsymbol{e_i}\right) \mathrm{d}u_i \\ &= \int_{u_k} \mathrm{d}u_k \int_{D'} (f \circ g)(\boldsymbol{v}') \, \mathrm{d}\boldsymbol{v}'\end{aligned}$$

It might or might not be easy to evaluate this integral directly. However, let's consider a C^1 transformation $\boldsymbol{\varphi} : U \to D$ where U is an open set in $\mathbb{R}^n$, such that $\boldsymbol{\varphi}(U) = D$. By the change of variable formula, we have:

$$\int_D (f \circ g)(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = \int_U (f \circ g \circ \boldsymbol{\varphi})(\boldsymbol{u}) \cdot \left| \det \mathbf{J}_{\boldsymbol{\varphi}}(\boldsymbol{u}) \right| \mathrm{d}\boldsymbol{u}$$

Notice if we have a good choice of $\boldsymbol{\varphi}$ with the following properties for a fixed $i \in \mathbb{N}$:

$$\begin{cases} (g \circ \boldsymbol{\varphi})(\boldsymbol{u}) = h(u_i) \\ \left|\det \mathbf{J}_{\boldsymbol{\varphi}}(\boldsymbol{u})\right| = h'(u_i) \end{cases} \quad \text{where } h : \mathbb{R} \to \mathbb{R}$$

We can simplify the integral to:

$$\begin{aligned} \int_U (f \circ g \circ \boldsymbol{\varphi})(\boldsymbol{u}) \cdot \left| \det \mathbf{J}_{\boldsymbol{\varphi}}(\boldsymbol{u}) \right| \mathrm{d}\boldsymbol{u} &= \int_U (f \circ h)(u_i) \cdot h'(u_i) \, \mathrm{d}\boldsymbol{u} \\ &= \int_{u_k, k \neq i} \mathrm{d}u_k \int_{u_i} (f \circ h)(u_i) \cdot h'(u_i) \, \mathrm{d}u_i \\ &= \int_{u_k, k \neq i} F(h(u_i))\big|_{\text{boundary of } u_i} \, \mathrm{d}u_k \end{aligned}$$

Now, we have reduced the problem to evaluating a single-variable integral. The only challenge left is to find a good $\boldsymbol{\varphi}$ that satisfies the properties mentioned above.

Now, let's look at the properties of $\boldsymbol{\varphi}$ carefully. Since $h(x_i)$ and $h'(x_i)$ is on the right hand side of the equations, we can rewrite the equations into a differential equation:

$$\left| \det \mathbf{J}_{\boldsymbol{\varphi}} \right| = \frac{\partial h}{\partial u_i} = \frac{\partial}{\partial u_i}(g \circ \boldsymbol{\varphi})$$

At the first glance, this seems to be a complicated PDE. Let's fix $j \in \mathbb{N}$:

$$\boldsymbol{\varphi}(\boldsymbol{u}) = \sum_{k \neq j} u_k \boldsymbol{e_k} + \psi(\boldsymbol{u}) \boldsymbol{e_j} \text{ where } \psi : U \to \mathbb{R}$$

This means that the transformation only affects the jth dimension, while keeping other dimensions unchanged.

With this constraint, the Jacobian matrix becomes:

$$\mathbf{J}_{\boldsymbol{\varphi}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \psi}{\partial u_1} & \frac{\partial \psi}{\partial u_2} & \cdots & \frac{\partial \psi}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

According to the properties of determinants, we have:

$$\det \mathbf{J}_{\boldsymbol{\varphi}} = \frac{\partial \psi}{\partial u_j}$$

Therefore, the PDE becomes:

$$\frac{\partial \psi}{\partial u_j} = \frac{\partial}{\partial u_i}(g \circ \boldsymbol{\varphi})$$

4.1. 2D Case

Let's consider a specific example when $n = 2$, $i = 1$, and $j = 2$ to intuitively understand the solution.

Now, the PDE becomes:

$$\begin{aligned} \frac{\partial \psi}{\partial u_2} &= \frac{\partial}{\partial u_1}(g(u_1, \psi(u_1, u_2))) \\ \frac{\partial \psi}{\partial u_2} &= \frac{\partial g}{\partial x}(u_1, \psi(u_1, u_2)) + \frac{\partial g}{\partial y}(u_1, \psi(u_1, u_2)) \cdot \frac{\partial \psi}{\partial u_1} \end{aligned}$$

Its characteristic equations are given by:

$$\frac{\mathrm{d}u_2}{1} = \frac{\mathrm{d}u_1}{-\partial_y g(u_1, \psi(u_1, u_2))} = \frac{\mathrm{d}\psi}{\partial_x g(u_1, \psi(u_1, u_2))}$$

Now we need a specific example of g. Consider the function:

$$g(x, y) = x^2 + y^2$$

Now, the characteristic equations become:

$$\frac{\mathrm{d}u_2}{1} = \frac{\mathrm{d}u_1}{-2\psi(u_1, u_2)} = \frac{\mathrm{d}\psi}{2u_1}$$

The second and third terms give us the following ODE:

$$2u_1 \,\mathrm{d}u_1 = -2\psi \,\mathrm{d}\psi$$

Integrating both sides, we have:

$$\int 2u_1 \,\mathrm{d}u_1 = \int 2\psi \,\mathrm{d}\psi$$
$$u_1^2 + \psi^2 = C_1$$

Similarly, we can obtain:

$$2u_2 + \arctan\left(\frac{u_1}{\psi}\right) = C_2$$

Thus, the solution of the PDE is given by:

$$G\left(u_1^2 + \psi^2, 2u_2 + \arctan\left(\frac{u_1}{\psi}\right)\right) = 0$$

where G is an arbitrary function with continuous first-order partial derivatives.

Now, let $G(x, y) = xy$, this gives us:

$$(u_1^2 + \psi^2) \cdot \left(2u_2 + \arctan\left(\frac{u_1}{\psi}\right)\right) = 0$$

Notice that the first part in the product is always non-negative, and equals to zero only when both u_1 and ψ equal to zero. Therefore, assuming it's non-zero, we have:

$$\begin{aligned}
2u_2 + \arctan\left(\frac{u_1}{\psi}\right) &= 0 \\
\arctan\left(\frac{u_1}{\psi}\right) &= -2u_2 \\
\frac{u_1}{\psi} &= -\tan(2u_2) \\
\psi &= -u_1 \cot(2u_2) \\
\boldsymbol{\varphi}(u_1, u_2) &= \langle u_1, -u_1 \cot(2u_2) \rangle
\end{aligned}$$

To verify the solution, we can compute the determinant of its Jacobian:

$$\det \mathbf{J}_{\boldsymbol{\varphi}}(u_1, u_2) = \begin{vmatrix} 1 & 0 \\ -\cot(2u_2) & 2u_1 \csc^2(2u_2) \end{vmatrix} = 2u_1 \csc^2(2u_2)$$

And compute $g \circ \boldsymbol{\varphi}$:

$$g \circ \boldsymbol{\varphi}(u_1, u_2) = u_1^2 + (-u_1 \cot(2u_2))^2 = u_1^2 \csc^2(2u_2)$$

Plugging them into the PDE, we have:

$$\frac{\partial}{\partial u_1}(g \circ \boldsymbol{\varphi})(u_1, u_2) = 2u_1 \csc^2(2u_2) = \det \mathbf{J}_{\boldsymbol{\varphi}}(u_1, u_2)$$

which confirms our solution.

4.2. Conclusion

Given a function $f : \mathbb{R} \to \mathbb{R}$, whose anti-derivative F is known, and another function $g : \mathbb{R}^n \to \mathbb{R}$ where $n \geq 2$. Let D be an open set in $\mathbb{R}^n$. For the following integral:

$$\int_D (f \circ g)(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v}$$

We can choose $i, j \in \mathbb{N}$ and define a C^1 transformation $\boldsymbol{\varphi} : U \to D$ where U is an open set in $\mathbb{R}^n$:

$$\boldsymbol{\varphi}(\boldsymbol{u}) = \sum_{k \neq j} u_k \boldsymbol{e_k} + \psi(\boldsymbol{u}) \boldsymbol{e_j}$$

Through solving the following PDE:

$$\frac{\partial \psi}{\partial u_j} = \frac{\partial}{\partial u_i}(g \circ \boldsymbol{\varphi})$$

We can find the function ψ that simplifies the integral:

$$\begin{aligned} \int_D (f \circ g)(\boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} &= \int_U (f \circ g \circ \boldsymbol{\varphi})(\boldsymbol{u}) \cdot \left| \det \mathbf{J}_{\boldsymbol{\varphi}}(\boldsymbol{u}) \right| \mathrm{d}\boldsymbol{u} \\ &= \int_{u_k, k \neq i} F(h(u_i))\big|_{\text{boundary of } u_i} \, \mathrm{d}u_k \end{aligned}$$