The Jacobian

Adarsh Kishore

Contents

1 Introduction

The Jacobian is a generalization of the substitution method used in single variable calculus.
The formula for the single variable case is,

$$\int_a^b f(x)dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(u))g'(u)du$$

Here, $g'(u)$ is the Jacobian for the transformation $x = g(u)$. To understand the geometric significance of this, it is necessary to view the derivative from a new perspective.

2 A New Interpretation of Derivative

It is convenient to think of the derivative as the slope of the tangent line to the graph of a function at a point. However, we also know that the derivative is a rate measurer, $\frac{dy}{dx}$ measures the rate at which y changes w.r.t. x.

As an example to clear up things, consider $y = x^2$. Then $y' = 2x$ is the derivative of this function. The statement reads that $\frac{dy}{dx} = 2x$, or $dy = 2x\,\mathrm{d}x$, which in plain english reads that "y changes twice as fast as x at any point". We can take a point, say $(2, 4)$ and zoom in to a very close negbourhood. Let's say we are only concerned about the interval $(2 - 10^{-37}, 2 + 10^{-37})$ and nothing beyond it.

In such a small interval, the function value remains almost constant, indeed the difference is atmost about $2\,\mathrm{x}\,10^{-37}$. In this case, the function can be treated almost linear and can be approximated by a linear function. Then the derivative at this point is just a linear coefficient of x, in this case 2. It means that when x changes by some value, say 10^{-38}, y changes by $2\,\mathrm{x}\,10^{-38}$. This coefficient 2 is actually the Jacobian of this function. The Jacobian is like the local "stretching" factor for a function. In the 1D case, it turns out that the Jacobian is numerically equal to the derivative.

Now we need to figure out how this Jacobian fits into the substiitution formula. For this (you guessed it), we need a better interpretation of the integral of a function.

3 A New Interpretation of the Integral

The most common interpretation of $\int_a^b f(x)\,\mathrm{d}x$ is the area under the graph of $f(x)$ from a to b. However, we'll change that now. Consider a very thin wire, with only length and no thickness, almost like a straight line. To find the mass of the wire, we need two things: the length of the wire l and the linear density λ. Let both of these be given. If λ is constant, then we have

$$m = \lambda \cdot l$$

However, if λ is not constant, and is instead a function of the position x on the wire, then how to find its mass? Suppose the end points of the wire are at a and b. We can divide the entire wire into a large number of strips of very small length. Let's zoom in on a point x and, say we are only concerned about finding the mass of the wire in the interval $(x - 10^{-37}, x + 10^{-37})$. In such a small length, the density is almost like constant, and we can calculate its mass as,

$$\Delta m = \lambda(x) \cdot \Delta x$$

We do the same thing for all the small strips, and add them up to get the total mass. To be accurate, we should let the length of each strip tend to 0, which makes the number of strips tend to ∞. Thus,

$$m = \lim_{n\to\infty} \sum_{i=1}^{n} \lambda(x_i) \cdot \Delta x_i := \int_a^b \lambda(x)\,\mathrm{d}x$$

Thus the integral of a function can be interpreted physically as the mass of a wire of length $(b - a)$ whose linear density function is $f(x)$.

4 The Jacobian in One Dimension

Now, suppose we have a wire whose end points are themselves functions, like $g(a)$ and $g(b)$. In this case it is quite natural to desire some way to make the end points just a and b. In this case, the substitution suggests itself. We make $x = g(u)$ where u is our new variable. So now the integral becomes,

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x \to \int_a^b f(g(u))\,\mathrm{d}u$$

But they're not quite equal, yet. Due to our substitution the length of each small strip has changed. Earlier it was $\mathrm{d}x$, now it has become $\mathrm{d}u$. How are they related. Well, we just discussed this interpretation of the derivative. Since both x and u are changing by very small values in a small strip, they are approximately linearly related. And that "stretching" factor is the Jacobian, which is just the derivative in the 1D case. Thus $\mathrm{d}x = g'(u)\,\mathrm{d}u$. And so we get

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(u))g'(u)\,\mathrm{d}u$$

This was the 1D case. With our new interpretations, let's jump into the 2D case.

5 The Jacobian in Two Dimensions

As the number of variables increase, it is convenient to take help from vectors, because vectors package two or more components into a single entity and have well defined operations like the dot and cross product. In other words, they'll do the technical hardwork for us and we concentrate only on the intuition purposes.

Let us consider a vector function $\mathbf{T}(u, v) = g(u, v)\mathbf{i} + h(u, v)\mathbf{j}$. What this is really saying is that give me a plane, the uv-plane, and I'll transform it into another plane, the xy-plane according to the rule $x = g(u, v)$ and $y = h(u, v)$. After all, what $y = f(x)$ meant that give me a line, the X-axis and I'll transform it into the curve $f(X)$ (notice the capital X). Thus it makes sense that as we push the dimension up, the line becomes a plane, then space, then hyperspace and beyond.

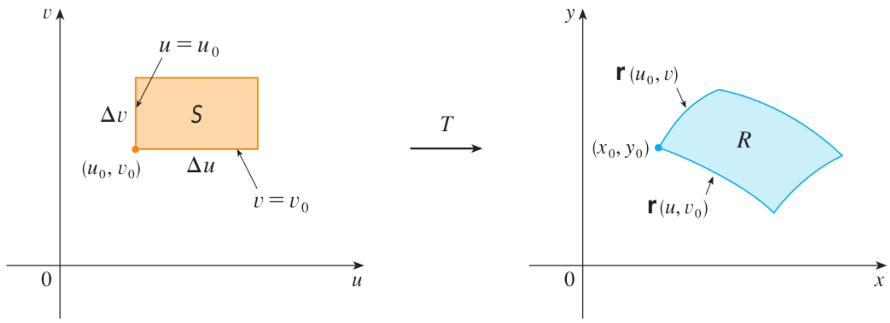

Figure 1: The Transformation $\mathbf{T}(u, v) = g(u, v)\mathbf{i} + h(u, v)\mathbf{j}$

Let's see what is going on in Figure 1. To simplify things, we're basically reducing things to one dimension only. This is possible if we examine the square with sides Δu and Δv beginning at (u_0, v_0) in the figure on the left. We we travel along the u-axis only, then $v = v_0$ is a constant. Thus the x and y coordinates become single variable functions of u only as v is fixed. Upon transformation this will yield in general a curve, given by $\mathbf{T}(u, v_0) = x(u, v_0)\mathbf{i} + y(u, v_0)\mathbf{j}$. Since these are x and y coordinates basically, this vector is basically the radius vector. Thus $\mathbf{T}$ is also sometimes referred by $\mathbf{r}$ as in the figure. Nothing new just a change of notation.

Likewise if travel along the line $u = u_0$ in the left figure, we'll be tracing out another curve $\mathbf{r}(u_0, v) = x(u_0, v)\mathbf{i} + y(u_0, v)\mathbf{j}$, a single variable function only. Now the thing is, if our increments Δu and Δv are small, then these curves are approximately linear, a concept we just explored in the new interpretation of the derivative. Thus the figure on the right is approximately a parallelogram, as the sides are straight lines.

We know what the area of the parallelogram is. It is the magnitude of the two adjacent vectors forming its sides. And the Jacobian in 1D was the stretching factor for lengths. It makes sense if it is now defined as the stretching factor for areas, specifically ratio of area of $\mathbf{R}$ to that of $\mathbf{S}$ as the increments tend to 0.

So what remains is to determine those adjacent vectors. But that's easy, it's just the difference of the radius vectors at the end points. Let these vectors be $\mathbf{v}_1$ (the horizontal-like one corresponding to Δu) and $\mathbf{v}_2$ (the vertical-like one corresponding to Δv).

We can write them as -

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{r}(u_0 + \Delta u, v_0) - \mathbf{r}(u_0, v_0) \\
&= \frac{\mathbf{r}(u_0 + \Delta u, v_0) - \mathbf{r}(u_0, v_0)}{\Delta u} \cdot \Delta u \\
&\approx \frac{\partial \mathbf{r}}{\partial u}(u_0, v_0) \cdot \Delta u \\
&\approx \mathbf{r}_u(u_0, v_0) \cdot \Delta u
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}_2 &= \mathbf{r}(u_0, v_0 + \Delta v) - \mathbf{r}(u_0, v_0) \\
&\approx \mathbf{r}_v(u_0, v_0) \cdot \Delta v
\end{aligned}$$

Thus, area of parallelogram **R** is -

$$\begin{aligned}
\Delta A &\approx \mid \mathbf{r}_u \times \mathbf{r}_v \mid \Delta u \cdot \Delta v \\
\implies \Delta A(S) &\approx \mid \mathbf{r}_u \times \mathbf{r}_v \mid \Delta A(R) \\
\implies \mathrm{d}A(S) &= \mid \mathbf{r}_u \times \mathbf{r}_v \mid \mathrm{d}A(R) \\
\implies \frac{\mathrm{d}A(S)}{\mathrm{d}A(R)} &= \mid \mathbf{r}_u \times \mathbf{r}_v \mid
\end{aligned}$$

This is the stretching factor of the areas, in other words this *is* the Jacobian in two dimensions. Computing it is very easy now.

$$\begin{aligned}
J := \frac{\partial(x, y)}{\partial(u, v)} &= \mid \mathbf{r}_u \times \mathbf{r}_v \mid \\
&= \left\| \begin{matrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_u & y_u & 0 \\ x_v & y_v & 0 \end{matrix} \right\| \\
&= \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}
\end{aligned}$$

6 Putting Everything in Place

For the 1D integration case, we had interpreted our function as the linear mass density of a very thin wire. Now we can think of it as the surface mass density of a thin plate. Suppose the area of the plate is A. If the surface mass density σ is a constant throughout the plate then we can calculate the mass of the plate as $m = \sigma \cdot A$.

However, this time let σ vary throughout the plate itself, i.e., it is a function of both x and y coordinates. Then just as before we can divide the plate into small rectangles of length Δx and width Δy. Then the density is nearly constant inside the rectangle with these sides and we can find this small rectangle's mass as $\Delta m = \sigma(x, y) \cdot \Delta A$. The total mass of the wire is-

$$m = \lim_{n\to\infty} \sum_{i=1}^{n} \sigma(x_i, y_i) \cdot \Delta A_i := \iint\limits_{R} \sigma(x, y)\, \mathrm{d}A$$

Thus the double integral of a function $f(x, y)$ on a domain R can be interpreted as the mass of a wire of the shape of domain R whose surface mass density is $f(x, y)$.

Now let's say the domain is R, but the integral is easier to evaluate on a domain S which is defined by $x = g(u, v)$ and $y = h(u, v)$. Then we have to see how much the area $\mathrm{d}A$(R) is changed into $\mathrm{d}A$(S). This "stretching factor" is the Jacobian. Since we only want the magnitude of scaling of areas (since area is always positive physically speaking), we take the absolute value of the Jacobian. Thus,

$$\iint\limits_{R} f(x, y)\, \mathrm{d}A = \iint\limits_{S} f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \mathrm{d}A$$

7 The Integral in Three Dimensions

As before, we'll see integration from the point of view of density. Imagine a 3D surface D whose volumetric density at each point is $\rho(x, y, z)$. Then the total mass of the surface is equal to the sum of all the little masses, which are just its density at each point times its volume. We'll divide the surface into a number of small cubes this time and add up the masses of each cube. Then total mass is

$$m = \lim_{n\to\infty} \sum_{i=1}^{n} \rho(x, y, z) \cdot \Delta V := \iiint\limits_{D} \rho(x, y, z)\, \mathrm{d}V$$

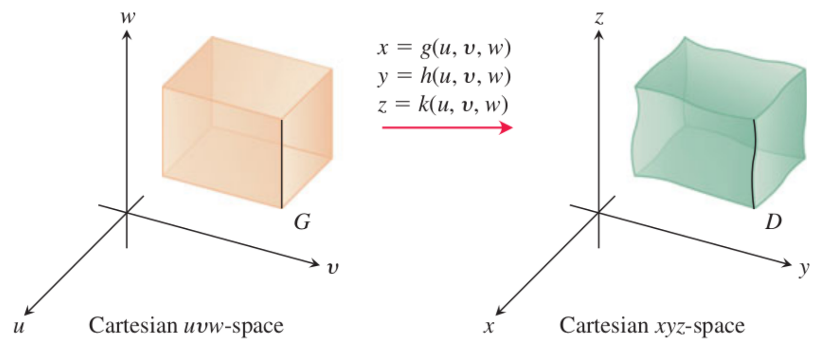

Figure 2: A 3D mapping

8 The Jacobian in Three Dimensions

As you can see in Figure 2, we have a 3D mapping, where the radius vector $\mathbf{r}$ is defined as $\mathbf{r}(u,v,w) = g(u,v,w)\mathbf{i} + h(u,v,w)\mathbf{j} + k(u,v,w)\mathbf{k}$. Now I don't have a neat figure to show the transformation but I hope you can imagine that a cube with sides $\Delta u, \Delta v, \Delta w$ will get mapped to a region whose end points are the radius vectors $\mathbf{r}(u_0, v_0, w_0), \mathbf{r}(u_0 + \Delta u, v_0, w_0), \mathbf{r}(u_0, v_0 + \Delta v)$ and $\mathbf{r}(u_0, v_0, w_0 + \Delta w)$.

These are approximately straight lines, and thus the entire thing is approximately a parallelopiped. The volume of this parallelopiped is the **scalar triple product** of its adjacent vectors on the sides.

Let these be $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ then by arguments similar to what we discussed in the 2D case,

$$\begin{aligned}
\mathbf{v}_1 &\approx \mathbf{r}_u(u_0, v_0, w_0) \cdot \Delta u \\
\mathbf{v}_2 &\approx \mathbf{r}_v(u_0, v_0, w_0) \cdot \Delta v \\
\mathbf{v}_3 &\approx \mathbf{r}_w(u_0, v_0, w_0) \cdot \Delta w
\end{aligned}$$

Therefore volume of parallelopiped D is,

$$\begin{aligned}
\mathrm{d}V(D) &= \mathbf{r}_u \cdot (\mathbf{r}_v \times \mathbf{r}_w)\,\mathrm{d}V(G) \\
&= \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\[2ex] \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \\[2ex] \dfrac{\partial x}{\partial w} & \dfrac{\partial y}{\partial w} & \dfrac{\partial z}{\partial w} \end{vmatrix} \mathrm{d}V(G)
\end{aligned}$$

The Jacobian was the stretching factor for length in 1D, area in 2D. It makes sense to define it as the stretching factor for volume in 3D.

$$J := \frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} & \dfrac{\partial z}{\partial u} \\[2ex] \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} & \dfrac{\partial z}{\partial v} \\[2ex] \dfrac{\partial x}{\partial w} & \dfrac{\partial y}{\partial w} & \dfrac{\partial z}{\partial w} \end{vmatrix}$$

9 Change of Variables in Three Dimensions

Suppose we have to measure the mass of the region D in Figure 2 but for some reason we find it easier to transform to region G. Then we have to make x, y, z as functions of u, v, w. Moreover, we have to see how each small cube is transforming. Since for small regions any well-behaved (differentiable) transform is approximately linear, we need to multiply the volume by its stretching factor, the Jacobian. As we want just the scale factor and not its sign, we take absolute value of the Jacobian.

$$\iiint\limits_D f(x,y,z)\,\mathrm{d}V = \iiint\limits_G f(x(u,v,w), y(u,v,w), z(u,v,w)) \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| \mathrm{d}V$$

10 Conclusion

This concludes our discussion. Hope you enjoyed it!