Jacobians Overview

P. Reany

July 5, 2021

Abstract

This short note is not really a detailed introduction to Jacobians, because I assume that the reader has already been exposed to the concept. My goal here is to define the Jacobian (matrix) of a transformation as the total derivative of the transformation, and then use my own Structured Differentiation (SD) to show what I mean by this. After that, I want to relate the SD version to some of the major versions in use, and in doing so, maybe to clear up some confusion surrounding the subject.

1 Main Body

My first comment is that if you are not confused about what a Jacobian is, maybe you should skip this paper. But if you are determined to continue reading, but you are not sure what I mean by "structured differentiation," you can read any number of companion papers that describe it in full.

So, let me immediately state my most controversial claim: The **Jacobian matrix** is the result of the total derivative of a transformation of variables from R^n to R^n and it reveals itself as an $n \times n$ matrix of n^2 total derivatives. Thus, the Jacobian is always the derivative of the 'old' variables with respect to the 'new' variables.

Now we'll continue without much controversy:

This derivative matrix can be put into the form of n^2 partial derivatives, which, because it is a square matrix, has a determinant, called the **Jacobian**. And it amazes me just how prevalent and useful this scalar function can be.

Let's begin with an example. Say we wish to perform an indefinite integration of $f(x)$,

$$I = \int f(x)\, dx \,. \tag{1}$$

Say that, upon reflection, we decide to make a change of variables to make the integration simpler, in which we'll go from the old variable x to the new variable

u, say. So, how does this affect the integral? Well,

$$f(x) \to \bar{f}(u) = f(x(u)) \quad \text{and} \quad dx \to \frac{dx}{du}\, du\,. \tag{2}$$

With these substitutions in hand, (1) becomes

$$I = \int \bar{f}(u)\, \frac{dx}{du}\, du\,. \tag{3}$$

So what appears as an 'extra' factor of dx/du in the integrand is really vital to getting it right. It's called the Jacobian of the transformation of variables from the 'old' variable x to the 'new' variable u.

But let's be a bit formal about what just happened. We needed to know x as a function of u, so we could write formally that

$$x = g(u)\,, \tag{4}$$

and thus

$$\frac{dx}{du} = g'(u)\,. \tag{5}$$

Indeed, by this we can rewrite (3) as

$$I = \int \bar{f}(u)\, g'(u)\, du\,. \tag{6}$$

Now, this is how mathematicians would probably prefer things, but I have a background in physics and engineering, and thus I find the perfunctory duplication of each variable by some formal avatar (such as g is to x) to be prohibited by the fact that practically all letters in physics are already used up! Thus, we are likely to see such unmathematical (but concise) things in physics as

$$x = x(t)\,, \tag{7}$$

in which x stands for both a variable and its avatar function. (To my knowledge, there are very few formal functions used in physics.)

Next, we go to a double integral, say,

$$I = \iint f(x,y)\, dxdy\,. \tag{8}$$

So, by analogy to the last integral, say we wish to integrate by two other variables like u and v, thus we have

$$x = x(u,v) \quad \text{and} \quad y = y(u,v)\,, \tag{9}$$

and

$$f(x,y) \to \bar{f}(u,v) = f(x(u,v), y(u,v)) \quad \text{and} \quad dxdy \to \frac{\delta(x,y)}{\delta(u,v)}\, dudv\,. \tag{10}$$

I show in my SD papers how to interpret this total derivative

$$\frac{\delta(x,y)}{\delta(u,v)} \tag{11}$$

and also how it can be replaced by 'partial' derivatives to give us

$$dxdy \to \frac{\partial(x,y)}{\partial(u,v)}\,dudv\,. \tag{12}$$

But wait! We're not quite there yet! There is a discrepancy between how convention treats $\frac{\partial(x,y)}{\partial(u,v)}$ and how SD treats it. In no way do I consider this a serious discrepancy, but it should be noted and kept in mind. In SD, but not in standard mathematics, the expressions

$$\frac{\delta(x,y)}{\delta(u,v)} \quad \text{and} \quad \frac{\partial(x,y)}{\partial(u,v)} \tag{13}$$

are treated as matrices, in this case as 2×2 matrices. Therefore, to properly state the correct transformed integral in SD, I should write

$$dxdy \to \left|\,\frac{\partial(x,y)}{\partial(u,v)}\,\right|\,dudv\,, \tag{14}$$

where

$$\left|\,\frac{\partial(x,y)}{\partial(u,v)}\,\right| \tag{15}$$

is the determinant of

$$\frac{\partial(x,y)}{\partial(u,v)}\,. \tag{16}$$

The matrices in (13) are called **Jacobian matrices** and their determinants are call Jacobians of the transformation of the old variables into the new. And, therefore, the double integral in (8) becomes

$$I = \iint \bar{f}\,(u,v)\,\left|\,\frac{\partial(x,y)}{\partial(u,v)}\,\right|\,dudv\,. \tag{17}$$

So, before I move on to more general considerations, I'd like to give a short defense of the SD convention to denote the Jacobian by use of explicit determinant signs. It's really quite simple. Jacobian matrices are of great computational value in themselves. If I am then to represent the Jacobian of a transformation, say, by

$$\frac{\partial(x,y)}{\partial(u,v)}\,, \tag{18}$$

then how do I represent the Jacobian matrix of the transformation? You see, something has to give somewhere.

In SD we can do differentiation of transformations of R^m to R^n, but for our immediate purposes, that is more general than is needed, so I'll continue to restrict our transformations from R^n to R^n.

Okay, so we have a set of n variables $\{y_1, y_2, \dots y_n\}$, each of which is a function of n new variables $\{x_1, x_2, \dots x_n\}$, which we can write as

$$y_i = f_i(x_1, x_2, \dots x_n)\,, \tag{19}$$

where f_i is the formal functional avatar for y_i for all $i \in [1, 2, \dots, n]$. So here I'm regarding the y_i as the 'old' variables, which are being written in terms of the 'new' variables x_i, yet, the transformation f_i takes the new variables into the old variables, which it must do.

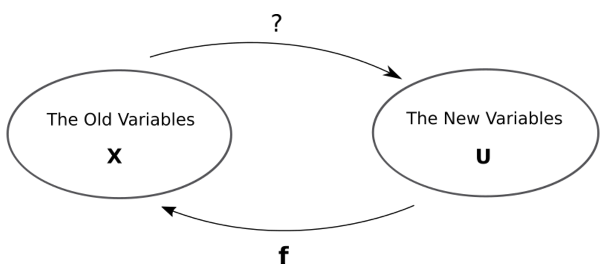

Figure 1. Although we can think of going from the 'old' variables to the 'new' variables, the formal transformation of variables f_i goes the other way. The derivative of **f** is the Jacobian matrix. The question mark is a placeholder for the inverse transformation, if it exists. It also represents my questionable intuition that felt that the mapping 'should' go the other way.

The SD equivalent of (19) is

$$\mathbf{y} = \mathbf{f}(\mathbf{x})\,, \tag{20}$$

where **f** is mapping the new variables into the old variables. On taking the differential on both sides, we get

$$\delta\mathbf{y} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\delta\mathbf{x}\,. \tag{21}$$

Dividing this by $\delta\mathbf{x}$ we have

$$\frac{\delta\mathbf{y}}{\delta\mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\delta\mathbf{x}}{\delta\mathbf{x}}\,. \tag{22}$$

Now, if all the variables x_i are independent of each other then $\dfrac{\delta\mathbf{x}}{\delta\mathbf{x}} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. And then (22) becomes

$$\frac{\delta\mathbf{y}}{\delta\mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\,. \tag{23}$$

We call $\delta\mathbf{f}/\delta\mathbf{x}$ the *total derivative* (or just *derivative*) of $\mathbf{f}$ with respect to $\mathbf{x}$; its matrix form is

$$\frac{\delta\mathbf{f}}{\delta\mathbf{x}} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \dots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \dots & \partial f_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \partial f_n/\partial x_2 & \dots & \partial f_n/\partial x_n \end{bmatrix}. \tag{24}$$

Definition: The ordered list of all variables on which a function is explicitly dependent is called the function's **variant** (vector). Each of these variables is also referred to as a *variant*, and context seems to keep the distinction clear.

If all the variants of a function are mutually independent then the total derivatives of that function reduce to partial derivatives, and, in the case of (24), we get

$$\frac{\partial\mathbf{f}}{\partial\mathbf{x}} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \dots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \dots & \partial f_2/\partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \partial f_n/\partial x_2 & \dots & \partial f_n/\partial x_n \end{bmatrix}. \tag{25}$$

2 Comparison to Convention

Some conventional views of this subject share with me the conviction that we should think of the Jacobian matrix as a total derivative. In any case, when conventional mathematicians long ago decided to remove the 'differential' (infinitesimal) from real mathematics, that left a brand new usage available for the term, which has been taken up by some mathematicians as the Jacobian matrix being the 'Differential' of a transformation. However, in SD, the Jacobian matrix is never called a 'differential', which is a term reserved for infinitesimals.

I am quite unrepentant over this viewpoint, since I doubt that any engineer has ever had bridge fail for his or her employing a classical differential in their analysis, nor is it likely that any rocket has failed its mission because a physicist thought in terms of classical differentials. And I cannot imagine classical thermodynamics taught without the heuristic of infinitesimal differentials.[1]

3 Conclusion

SD is not really a novel system of differentiation. It has its bits and pieces spread all over the literature for many decades before I came along in the 1980s

[1] I'm not denying that thermodynamics can be presented from the viewpoint of differential forms, but that is perhaps a lot to ask of an already-mathematically-burdened undergraduate engineering student.

and tried to synthesize the best parts of it into a seamless organic whole. In particular, I was influenced much by Buck's *Advanced Calculus* [1].

As for how the general literature treats Jacobian matrices and their determinants, I can only say that, by contrast, I sought for consistency and clarity. The reader will find many variations in the literature regarding the nomenclature and symbolism of these primary objects, some similar to mine, some very different from mine.

References

[1] R.C. Buck. *Advanced Calculus*, 3ed. McGraw-Hill Book Co. (1984).