DIFFERENTIAL GEOMETRY NOTES BASED ON UNDERGRADUATE LINEAR ALGEBRA AND MULTIVARIABLE CALCULUS

bob jantzen
Department of Mathematical Sciences
Villanova University

in progress: version: January 15, 2008

Contents

Preface

These handwritten notes from a course given at Villanova University in the spring semester of 1991 were scanned and posted on the web in 2006 at

`http://www34.homepage.villanova.edu/robert.jantzen/notes/dg1991/`

and were converted to a LaTeX compuscript and revised in 2007 with the help of Hans Kuo of Taiwan through a serendipitous internet collaboration.

Most undergraduate courses on differential geometry are leftovers from the early part of the last century, focusing on curves and surfaces in space, which is not very useful for the most important application of the twentieth century: general relativity and field theory in theoretical physics. Most mathematicians who teach such courses are not well versed in physics, so perhaps this is a natural consequence of the distancing of mathematics from physics, two fields which developed together in creating these ideas from Newton to Einstein and beyond. The idea of these notes is to develop the essential tools of modern differential geometry while bypassing more abstract notions like manifolds, which although important for global questions, are not essential for local differential geometry and therefore need not steal precious time from a first course aimed at undergraduates.

Part 1 (Algebra) develops the vector space structure of $\mathbb{R}^n$ and its dual space of real-valued linear functions, and builds the tools of tensor algebra on that structure, getting the index manipulation part of tensor analysis out of the way first. Part 2 (Calculus) then develops $\mathbb{R}^n$ as a manifold first analyzed in Cartesian coordinates, beginning by redefining the tangent space of multivariable calculus to be the space of directional derivatives at a point, so that all of the tools of Part 1 then can be applied pointwise in the space. Non-Cartesian coordinates and the Euclidean metric are then used as a shortcut to what would be the consideration of more general manifolds with Riemannian metrics in a more ambitious course, followed by the covariant derivative and parallel transport, leading naturally into curvature. The exterior derivative and integration of differential forms is the final topic, showing how conventional vector analysis fits into a more elegant unified framework.

The theme of Part 1 is that one needs to distinguish the linearity properties from the inner product ("metric") properties of linear algebra. The inner product geometry governs lengths and angles, and the determinant then enables one to extend the linear measure of length to area and volume in the plane or 3-dimensional space, and to p-dimensional objects in $\mathbb{R}^n$. The determinant also tests linear independence of a set of vectors and hence is key to characterizing subspaces independent of the particular set of vectors we use to describe them while assigning an actual measure to the p-parallelepipeds formed by a particular set, once an inner product sets the length scale for orthogonal directions. By appreciating the details of these basic notions in the setting of $\mathbb{R}^n$, one is ready for the tools needed point by point in the tangent spaces to $\mathbb{R}^n$, once one understands the relationship between each tangent space and the simpler enveloping space.

Introduction: motivating index algebra

Elementary linear algebra is the mathematics of linearity, whose basic objects are 1- and 2-dimensional arrays of numbers, which can be visualized as at most 2-dimensional rectangular arrangements of those numbers on sheets of paper or computer screens. Arrays of numbers of dimension m can be described as sets that can be put into a 1-1 correspondence with regular rectangular grids of points in R^m:

$$\begin{aligned} \{a_i \mid i = 1, \ldots, n\} &\qquad \text{1-}d \text{ array} : n \text{ entries} \\ \{a_{ij} \mid i = 1, \ldots, n_1, j = 1, \ldots, n_2\} &\qquad \text{2-}d \text{ array} : n_1 n_2 \text{ entries} \\ \{a_{ijk} \mid i = 1, \ldots, n_1, j = 1, \ldots, n_2, k = 1, \ldots, n_3\} &\qquad \text{3-}d \text{ array} : n_1 n_2 n_3 \text{ entries} \end{aligned}$$

1-dimensional arrays (vectors) and 2-dimensional arrays (matrices), coupled with the basic operation of matrix multiplication, itself an organized way of performing dot products of two sets of vectors, combine into a powerful machine for linear computation. When working with arrays of specific dimensions (3 component vectors, 2×3 matrices, etc.), one can avoid index notation and the sigma summation symbol $\sum_{i=1}^{n}$ after using it perhaps to define the basic operation of dot products for vectors of arbitrary dimension, but to discuss theory for indeterminant dimensions (n-component vectors, $m \times n$ matrices), index notation is necessary. However, index "positioning" (distinguishing subscript and superscript indices) is not essential and rarely used, especially by mathematicians. Going beyond 2-dimensional arrays to m-dimensional arrays for $m > 2$, the arena of "tensors", index notation and index positioning are instead both essential to an efficient computational language.

Suppose we start with 3-vectors to illustrate the basic idea. The dot product between two vectors is symmetric in the two factors

$$\begin{aligned} \vec{a} &= \langle a_1, a_2, a_3 \rangle \,, \ \vec{b} = \langle b_1, b_2, b_3 \rangle \\ \vec{a} \cdot \vec{b} &= a_1 b_1 + a_2 b_2 + a_3 b_3 = \sum_{i=1}^{3} a_i b_i = \vec{b} \cdot \vec{a} \,, \end{aligned}$$

but using it to describe a linear function in R^3, a basic asymmetry is introduced

$$f_{\vec{a}}(\vec{x}) = \vec{a} \cdot \vec{x} = a_1 x_1 + a_2 x_2 + a_3 x_3 = \sum_{i=1}^{3} a_i x_i \,.$$

The left factor is a constant vector of "coefficients", while the right factor is the vector of "variables" and this choice of left and right is arbitrary but convenient, although some mathematicians like to reverse it for some reason. To reflect this distinction, we introduce superscripts (up position) to denote the variable indices and subscripts (down position) to denote the coefficient indices, and then agree to sum over the understood 3 values of the index range for any repeated such pair of indices (one up, one down)

$$f_{\vec{a}}(\vec{x}) = a_1 x^1 + a_2 x^2 + a_3 x^3 = \sum_{i=1}^{3} a_i x^i = a_i x^i \,.$$

The last convention, called the Einstein summation convention, turns out to be an extremely convenient and powerful shorthand.

This index positioning notation encodes the distinction between rows and columns in matrix notation, with row indices (left) associated with superscripts, and column indices (right) with subscripts. A single row matrix or column matrix is used to denote respectively a "coefficient" vector and a "variable" vector

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, \quad \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}.$$

The matrix product re-interprets the dot product between two vectors as the way to combine a row vector (left factor) with a column vector (right factor) to produce a single number, the value of a linear function of the variables

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = a_1 x^1 + a_2 x^2 + a_3 x^3 = \vec{a} \cdot \vec{x}\,.$$

If we agree to use a boldface kernel symbol $\mathbf{x}$ for a column vector, and the transpose $\boldsymbol{a}^T$ for a row vector, where the transpose simply interchanges rows and columns of a matrix, this can be represented as $\boldsymbol{a}^T \mathbf{x} = \vec{a} \cdot \vec{x}$.

Extending the matrix product to more than one row in the left factor is the second step in defining a general matrix product, leading to a column vector result

$$\begin{bmatrix} a^1{}_1 & a^1{}_2 & a^1_3 \\ a^2{}_1 & a^2{}_2 & a^2_3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}^{1T} \\ \boldsymbol{a}^{2T} \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} \vec{a}^1 \cdot \vec{x} \\ \vec{a}^2 \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} a^1{}_i x^i \\ a^2{}_i x^i \end{bmatrix}.$$

Thinking of the matrix as a 1-dimensional vertical array of row vectors (second equation of previous equation array), one gets a corresponding array of numbers (a column) as the result, consisting of the corresponding dot products of the rows with the single column.

Finally, adding more columns to the right factor in the matrix product, we generate corresponding columns in the matrix product, with the resulting array of numbers representing all possible dot products between the row vectors on the left and the column vectors on the right, labeled by the same row and column indices as the factor vectors from which they come.

$$\begin{bmatrix} a^1{}_1 & a^1{}_2 & a^1_3 \\ a^2{}_1 & a^2{}_2 & a^2_3 \end{bmatrix} \begin{bmatrix} x^1{}_1 & x^1{}_2 \\ x^2{}_1 & x^2{}_2 \\ x^3{}_1 & x^3{}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}^{1T} \\ \boldsymbol{a}^{2T} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \vec{a}^1 \cdot \vec{x}_1 & \vec{a}^1 \cdot \vec{x}_2 \\ \vec{a}^2 \cdot \vec{x}_1 & \vec{a}^2 \cdot \vec{x}_2 \end{bmatrix}.$$

Denoting these two matrix factors by $\mathbf{A}$ and $\mathbf{X}$, then the product matrix has entries (row index left up, column index right down)

$$[AX]^i{}_j = a^i{}_k x^k{}_j \quad i = 1..2\,, \; j = i..2$$

where the sum $\sum_{k=1}^{3}$ is implied. Thus matrix multiplication is just an organized way of displaying all such dot products in an array where the rows correspond to the coefficient vectors in the left set and the columns correspond to the variable vectors in the right set. The dot product itself in this context is representing the natural evaluation of linear functions (left) on vectors (right). No geometry (lengths and angles in Euclidean geometry) is implied in this context, only linearity.

The matrix product of a matrix with a column vector can be reinterpreted in terms of the more general concept of a vector-valued linear function of vectors, namely a linear combination of vectors, in which case the right factor column vector entries play the role of coefficients. In this case the left factor matrix must be thought of as a horizontal array of column vectors

$$\begin{bmatrix} v^1{}_1 & v^1{}_2 & v^1_3 \\ v^2{}_1 & v^2{}_2 & v^2_3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} v^1{}_1 x^1 + v^1{}_2 x^2 + v^1{}_3 x^3 \\ v^2{}_1 x^1 + v^2{}_2 x^2 + v^2{}_3 x^3 \end{bmatrix}$$
$$= \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = x^1 \vec{v}_1 + x^2 \vec{v}_2 + x^3 \vec{v}_3 \,.$$

This interpretation extends to more columns on the right, leading to a matrix of columns, each of which represents a linear combination of the column vectors of the left factor. In this case the coefficient indices are superscripts since the labels of the vectors being combined linearly are subscripts, but the one up, one down repeated index summation is still consistent. Note that when the left factor matrix is not square (in this example, a 2×3 matrix), one is dealing with coefficient vectors and vectors of different dimensions, here combining three 2-component vectors for example.

If we call our basic column vectors just vectors (contravariant vectors, indices up) and call row vectors "covectors" (covariant vectors, indices down), then combining them with the matrix product represents the evaluation operation for linear functions, and implies no geometry in the sense of lengths and angles usually associated with the dot product, although one can easily carry over this interpretation. In this example R^3 is our basic vector space, and the space of all linear functions on it is equivalent to another copy of R^3, the space of all coefficient vectors. The space of linear functions on a vector space is called the dual space, and given a basis of the original vector space, expressing linear functions with respect to this basis leads to a component representation in terms of matrices as above.

It is this basic foundation of a vector space and its dual, together with the natural evaluation represented by matrix multiplication in component language, reflected in superscript and subscript index positioning respectively associated with column vectors and row vectors, that is used to go beyond elementary linear algebra to the algebra of tensors, or m-dimensional arrays for any positive integer m. Index positioning together with the Einstein summation convention is essential in letting the notation itself directly carry the information about its function in this scheme of linear mathematics extended beyond the elementary level.

Combining this linear algebra structure with multivariable calculus leads to differential geometry. Consider R^3 with the usual Cartesian coordinates x^1, x^2, x^3 thought of as functions on this space. The differential of any function on this space can be expressed in terms of partial derivatives by the formula

$$df = \frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + \frac{\partial f}{\partial x^3}dx^3 = \partial_i f dx^i = f_{,i}dx^i$$

using first the abbreviation $\partial_i = \frac{\partial}{\partial x^i}$ for the partial derivative operator and then the abbreviation $f_{,i}$ for the corresponding partial derivatives of the function f. At each point of R^3, the differentials dx^i play the role of linear functions on the tangent space. The differential of f acts on a tangent vector $\vec{v}$ at a given point by evaluation to form the directional derivative along the vector

$$D_{\vec{v}}f = \frac{\partial f}{\partial x^1}v^1 + \frac{\partial f}{\partial x^2}v^2 + \frac{\partial f}{\partial x^3}v^3 = \frac{\partial f}{\partial x^i}v^i \,,$$

so that the coefficients of this linear function of tangent vectors at a given point are the values of the partial derivative functions there, and hence have indices down compared to the indices up of the tangent vectors themselves, which belong to the tangent space, the fundamental vector space describing the diffential geometry near each point of the whole space. In the linear function notation, the application of the linear function df to the vector $\vec{v}$ gives the same result

$$df(\vec{v}) = \frac{\partial f}{\partial x^i}v^i \,.$$

If $\frac{\partial f}{\partial x^i}$ are therefore the components of a covector, and v^i the components of a vector in the tangent space, what is the basis of the tangent space, analogous to the natural (ordered) basis $\{e_1, e_2, e_3\} = \{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ of R^3 thought of as a vector space in our previous discussion? In other words how do we express a tangent vector in the abstract form like in the naive R^3 discussion where $\vec{x} = \langle x^1, x^2, x^3\rangle = x^i e_i$? This question will be answered in the following notes, making the link between old fashioned tensor analysis and modern differential geometry.