A review of tensors in non-Cartesian coordinate systems

Kris Innanen

ABSTRACT

This is a brief review of some of the ideas of tensor mathematics in non-Cartesian coordinate systems. There is no new material here, only a particular selection and arrangement of theory available in many good textbooks. The mathematics is needed in order to discuss model space re-parameterizations as transformations between coordinate systems, and to develop the new re-parameterizations we attempt elsewhere. The document ends with a listing of transformation rules we will use repeatedly elsewhere.

INTRODUCTION

In this report are a set of papers on the problem of re-parameterization of model space for geophysical inverse problems, how many types of re-parameterization currently in use can be formulated as transform problems between appropriately general coordinate systems, and how useful new coordinate systems (and thereby re-parameterizations) can be designed. Transformations back and forth between Cartesian coordinate systems rotated relative to one another are not sufficient for this purpose, however, and we must allow transformations to be between systems with non-orthogonal coordinate axes, and length scales which vary from one coordinate direction to the next. This in turn requires us to use more general geometrical considerations than is common for geophysicists, including use of and distinction between covariant and contravariant components of tensors and vectors. This review paper is geared towards readers familiar with Cartesian vector and tensor quantities and their transformation rules, but not the more general covariant notation. Of course, all of this information can be found in many good texts on the subject, and at most the current paper is a useful collection of particular features of the problem needed for our purposes. I myself followed the concise and logically perfect textbook on general relativity by Dirac (1975), occasionally making use of Arfken and Weber (2001).

TENSORS AND VECTORS IN NON-CARTESIAN COORDINATE SYSTEMS

Let us start with a tutorial type introduction, exemplified with some simple cases and calculations. In this section we will use variables x and y to make it look like familiar algebra and geometry; presently these will be switched to s and r, which look a little odd but are more appropriate for the geophysical applications in other companion reports.

Vectors and tensors in indicial notation

When dealing with rectilinear but non-Cartesian coordinate systems (we will often use the word *oblique*), the indicial notation we are familiar with in continuum mechanics remains useful, but it includes one additional wrinkle, which is that the system requires both upper (i.e., superscript) and lower (subscript) indices. Vectors have a single free index which may be in either upper or lower positions, and tensors can be constructed through products of vectors with any combination of upper and lower indices. For instance,

$$s^{\alpha\beta}_{\gamma} = p^{\alpha} q^{\beta} r_{\gamma} \tag{1}$$

is a tensor of rank 3. As in continuum mechanical indicial notation, repeated indices imply a sum, however, in this more general system repeated indices must be in upper and lower pairs. For instance,

$$p^{\alpha\beta}_{\gamma} q^{\gamma}_{\beta} \tag{2}$$

is a vector with a single free (upper) index α, with γ and β being dummy indices which are summed out. The scalar (or inner) product of two vectors $\mathbf{a}$ and $\mathbf{b}$ in indicial notation is

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = a_{\mu} b^{\mu}. \tag{3}$$

This is referred to as a scalar product, because after it is complete there are no free indices remaining, and any quantity with no free indices is referred to as a *scalar*. Another scalar is the squared length of a vector, e.g., for $\mathbf{a}$,

$$|\mathbf{a}|^2 = a_{\mu} a^{\mu}. \tag{4}$$

Scalars are numbers which are invariant under coordinate transformations. They are therefore distinct from the components of a vector, which, though also numbers, tend to change under coordinate transformations. Derivatives of scalars, vectors, or tensors with respect to spatial coordinates can be expressed using commas or ∂ symbols in indicial notation. For instance, the derivative of a scalar ϕ with respect to x^{μ} is

$$\frac{\partial \phi}{\partial x^{\mu}} = \partial_{\mu} \phi = \phi_{,\mu}, \tag{5}$$

and the divergence of the components of a vector with upper indices is

$$\frac{\partial a^{\mu}}{\partial x^{\mu}} = a^{\mu}_{,\mu}, \tag{6}$$

etc. Repeated indices both in the upper or the lower spots have no meaning, and do not feature in any properly formed mathematical statements.

An example coordinate system with oblique axes

Consider two coordinate systems within which to describe points on the plane: a reference Cartesian system x, and an oblique system y. In the oblique system the vertical axis is tilted clockwise; the tilt can be described by the angle θ away from the vertical (see Figure 2a). The axes are characterized by unit vectors $\mathbf{e}_0(x)$ and $\mathbf{e}_1(x)$ in the x system, and $\mathbf{e}_0(y)$ and $\mathbf{e}_1(y)$ in the y system.

Let us show that in such a system, a point with x coordinates (x^0, x^1) has y coordinates (y^0, y^1), and vice versa, where

$$\begin{bmatrix} y^0 \\ y^1 \end{bmatrix} = \begin{bmatrix} \sec\theta & 0 \\ -\tan\theta & 1 \end{bmatrix} \begin{bmatrix} x^0 \\ x^1 \end{bmatrix}, \text{ and } \begin{bmatrix} x^0 \\ x^1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} y^0 \\ y^1 \end{bmatrix}. \tag{7}$$

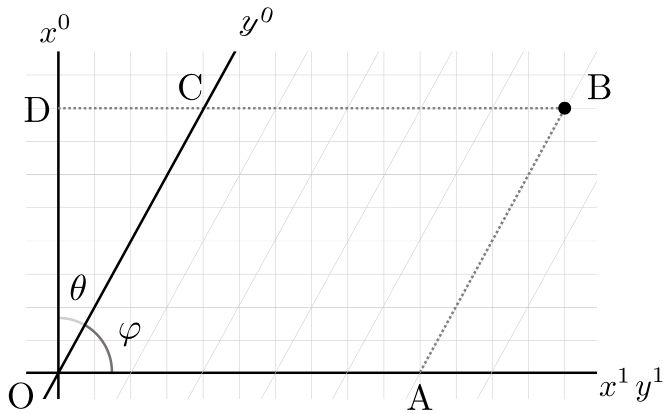

FIG. 1. A diagram to support determining coordinate values of a point B in the orthogonal x system and the oblique y system. It could be by turns convenient to frame results in terms of the angle θ between y^0 and x^1 or φ between y^0 and x^0.

The relevant systems x and y of interest are illustrated in Figure 1. The values of the x coordinates of some point, say B, are given respectively by the lengths OD along the vertical axis and DB along the horizontal axis. Let lines parallel to OD and DB, forced to go through O, define the x^0 and x^1 coordinate axes respectively. Meanwhile, the y^1 coordinate axis is set equal to the x^1 coordinate axis, and the y^0 axis is the oblique line containing the points O and C. The oblique axis makes an angle θ with the vertical axis and an angle φ with the horizontal axis. Formulae for determining the y coordinates (y^0, y^1) of a point in terms of the x coordinates (x^0, x^1) and vice versa are determined as follows. The y^1 value of the point B is determined by measuring its distance from the y^0 axis along a direction parallel to the y^1 axis, which is CB. The y^0 value of the point B is likewise determined by measuring its distance from the y^1 axis along a direction parallel to the y^0 axis, which is AB. That is,

$$\begin{aligned} y^0(x^0, x^1) &= \text{AB} = x^0 \csc\varphi = x^0 \sec\theta, \\ y^1(x^0, x^1) &= \text{DB} - \text{DC} = x^1 - x^0 \cot\varphi = x^1 - x^0 \tan\theta. \end{aligned} \tag{8}$$

Solving instead for the x coordinates, we obtain

$$\begin{aligned} x^0(y^0, y^1) &= y^0 \sin\varphi = y^0 \cos\theta \\ x^1(y^0, y^1) &= y^0 \sin\varphi \cot\varphi + y^1 = y^0 \cos\varphi + y^1 = y^0 \sin\theta + y^1. \end{aligned} \tag{9}$$

Assembling these into matrix form, we reproduce (7) as desired.

Given this, consider a small displacement of a point in the plane. In the y system such a displacement can be written as

$$dy^\mu = \begin{bmatrix} dy^0 \\ dy^1 \end{bmatrix}. \tag{10}$$

The components in (10) can be related to their counterparts in the x system by taking differentials. For instance,

$$dy^0 = \frac{\partial y^0(x^0, x^1)}{\partial x^0} dx^0 + \frac{\partial y^0(x^0, x^1)}{\partial x^1} dx^1; \tag{11}$$

and similarly for dy^1; the x differentials can also be determined from the y derivatives, by

$$dx^0 = \frac{\partial x^0(y^0,y^1)}{\partial y^0}dy^0 + \frac{\partial x^0(y^0,y^1)}{\partial y^1}dy^1, \tag{12}$$

etc. The elements of the matrices in equation (7) provide these derivatives directly, and so the rules for the vector displacements are

$$\begin{bmatrix} dy^0 \\ dy^1 \end{bmatrix} = \begin{bmatrix} \sec\theta & 0 \\ -\tan\theta & 1 \end{bmatrix} \begin{bmatrix} dx^0 \\ dx^1 \end{bmatrix}, \text{ and } \begin{bmatrix} dx^0 \\ dx^1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} dy^0 \\ dy^1 \end{bmatrix}. \tag{13}$$

Vectors with components referred to oblique axes

We labelled the displacement vectors with superscript indices. Vector components which transform in the same way as small displacements are called *contravariant*, and are all similarly labelled. For the arbitrary vector "**a**" in Figure 2b, this means we go back and forth between the x and y systems using (13) as a template:

$$\begin{aligned} a^\mu(y) &= \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix} = \begin{bmatrix} \sec\theta & 0 \\ -\tan\theta & 1 \end{bmatrix} \begin{bmatrix} a^0(x) \\ a^1(x) \end{bmatrix}, \\ a^\mu(x) &= \begin{bmatrix} a^0(x) \\ a^1(x) \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix}. \end{aligned} \tag{14}$$

This is a special case of the more general formula

$$\begin{aligned} a^\mu(y) &= \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix} = \begin{bmatrix} y^0_{,0} & y^0_{,1} \\ y^1_{,0} & y^1_{,1} \end{bmatrix} \begin{bmatrix} a^0(x) \\ a^1(x) \end{bmatrix} = y^\mu_{,\nu} a^\nu(x) \\ a^\mu(x) &= \begin{bmatrix} a^0(x) \\ a^1(x) \end{bmatrix} = \begin{bmatrix} x^0_{,0} & x^0_{,1} \\ x^1_{,0} & x^1_{,1} \end{bmatrix} \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix} = x^\mu_{,\nu} a^\nu(y), \end{aligned} \tag{15}$$

where in the second case the comma notation implies derivatives of $x^\mu(y^0, y^1)$ with respect to the y variables.

Vector components with lower indices are called *covariant*. They transform according to different rules. To develop them, consider again the scalar product in (3). The value of this product must be independent of coordinate system, so

$$a^\nu(y)b_\nu(y) = a^\mu(x)b_\mu(x). \tag{16}$$

But, since $a^\mu(x) = x^\mu_{,\nu}a^\nu(y)$, per equation (15), it must also be that

$$a^\nu(y)b_\nu(y) = a^\nu(y)\, x^\mu_{,\nu}\, b_\mu(x). \tag{17}$$

Comparing the right and left sides of this relation, we see that

$$b_\nu(y) = x^\mu_{,\nu}(y)\, b_\mu(x), \tag{18}$$

which gives the transformation rule for the covariant components of the vectors. Assembling this into something that looks more like a counterpart to the contravariant rule in equation (15), we have then:

$$\begin{aligned} a_\mu(y) &= \begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix} = \begin{bmatrix} x^0_{,0} & x^1_{,0} \\ x^0_{,1} & x^1_{,1} \end{bmatrix} \begin{bmatrix} a_0(x) \\ a_1(x) \end{bmatrix} = x^\mu_{,\nu} a^\nu(x) \\ a_\mu(x) &= \begin{bmatrix} a_0(x) \\ a_1(x) \end{bmatrix} = \begin{bmatrix} y^0_{,0} & y^1_{,0} \\ y^0_{,1} & y^1_{,1} \end{bmatrix} \begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix} = y^\mu_{,\nu} a^\nu(y). \end{aligned} \tag{19}$$

In the tilted vertical axis example we have been working with, this amounts to

$$\begin{aligned} a_\mu(y) &= \begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix} = \begin{bmatrix} \sec\theta & -\tan\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_0(x) \\ a_1(x) \end{bmatrix}, \\ a_\mu(x) &= \begin{bmatrix} a_0(x) \\ a_1(x) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix}. \end{aligned} \tag{20}$$

Geometrical meaning of contravariant and covariant components

The contravariant and covariant components of a vector are distinct from one another in situations where synthesis and analysis of the vector are not the same. The contravariant components of a vector tell how much of each basis vector is needed for its synthesis, in either the x or y systems. They are in our 2D example the a^0 and a^1 in the construction

$$\mathbf{a} = a^0(x)\,\mathbf{e}_0(x) + a^1(x)\,\mathbf{e}_1(x) = a^0(y)\,\mathbf{e}_0(y) + a^1(y)\,\mathbf{e}_1(y). \tag{21}$$

Graphically, these components correspond to the lengths of the black and gray bars in Figure 2c. The two basis vectors in either system, when weighted by the contravariant components in this way, can be observed to form "vector parallelograms" which correctly sum to form $\mathbf{a}$. In the x case, the parallelogram is a rectangle; in the y case, it is not. So, when contravariant components are under discussion, one is talking about the process of building or synthesizing a vector.

In the analysis of a vector, we are pulling $\mathbf{a}$ apart, rather than building it. We do this with projections (i.e., inner products) of $\mathbf{a}$ onto basis vectors. Cartesian coordinate systems like x have the special property that these projections are in fact equal to the weights needed to construct $\mathbf{a}$. In Cartesian systems, the synthesis and analysis problems are, in other words, the same. This is not true of the y system. In Figure 2d, the lengths of the orthogonal projections of $\mathbf{a}$ onto the four basis vectors are illustrated, again as black and grey bars. In the x system (black bars), the lengths match those in Figure 2c, but in the y case (grey bars) these lengths are not the ones needed in a synthesis like that in equation (21). Nevertheless, the two sets of numbers based on the projections, namely

$$a_0(y) = \mathbf{a}\cdot\mathbf{e}_0(y),\ a_1(y) = \mathbf{a}\cdot\mathbf{e}_1(y),\ \text{and}\ a_0(x) = \mathbf{a}\cdot\mathbf{e}_0(x),\ a_1(x) = \mathbf{a}\cdot\mathbf{e}_1(x), \tag{22}$$

each fully characterize the vector, and in fact form valid sets of vector components. These of course are the *covariant* components of $\mathbf{a}$.

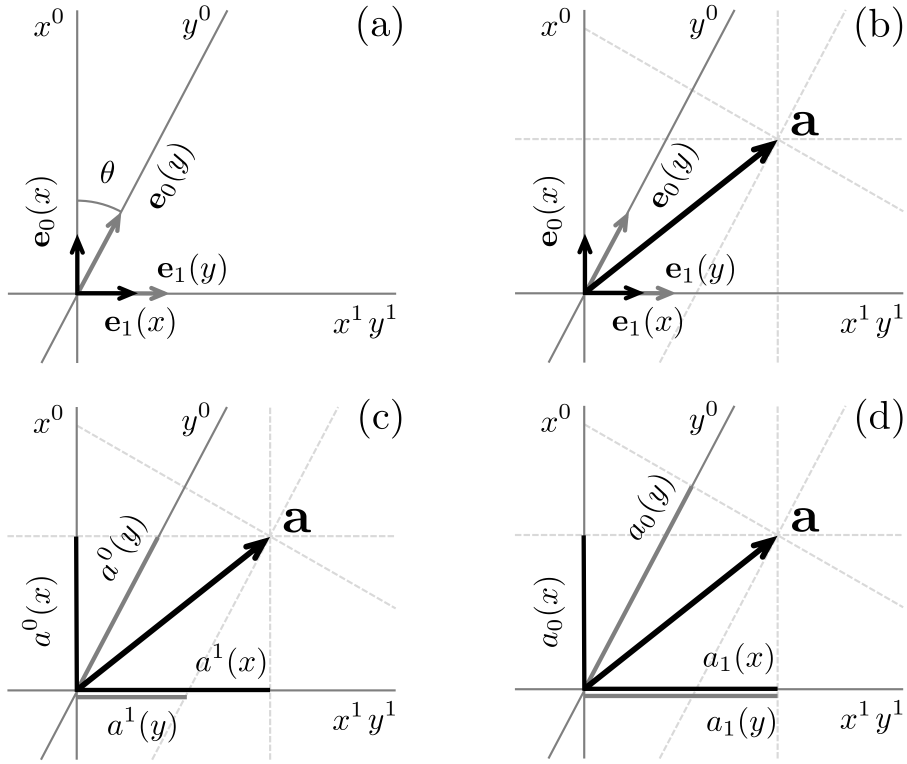

FIG. 2. Positions and vector components referred to orthogonal (x) and oblique (y) coordinate systems. (a) The coordinate axes and their basis vectors; note, the $\mathbf{e}_0(x)$, $\mathbf{e}_1(x)$, $\mathbf{e}_0(y)$ and $\mathbf{e}_1(y)$ are all unit vectors, but the arrows in the y system (grey) have been lengthened for illustration purposes. (b) An arbitrary vector $\mathbf{a}$ can be resolved into components in either system. These components can be either of (c) contravariant type, with grey and blue bars indicating their size, or (d) covariant type, again with black and grey bars indicating their size. Notice the black bars are unchanged from (c) to (d), but the grey bars are. This reflects the different numerical values taken on by covariant and contravariant components when axes are oblique.

Raising and lowering indices

A vector expressed in terms of its contravariant components can be easily re-expressed in terms of its covariant components. Let us develop this algebraically. The contravariant components of $\mathbf{a}$ are the weights involved in its construction, namely:

$$\mathbf{a} = a^0(y)\mathbf{e}_0(y) + a^1(y)\mathbf{e}_1(y). \tag{23}$$

Whereas, the covariant components are the projections of $\mathbf{a}$ onto the basis vectors:

$$a_0(y) = \mathbf{a}\cdot\mathbf{e}_0(y),\ a_1(y) = \mathbf{a}\cdot\mathbf{e}_1(y). \tag{24}$$

Substituting equation (23) into (24), we obtain

$$\begin{aligned} a_0(y) &= a^0(y)\mathbf{e}_0(y)\cdot\mathbf{e}_0(y) + a^1(y)\mathbf{e}_1(y)\cdot\mathbf{e}_0(y) = a^0(y) + \sin\theta\, a^1(y) \\ a_1(y) &= a^0(y)\mathbf{e}_0(y)\cdot\mathbf{e}_1(y) + a^1(y)\mathbf{e}_1(y)\cdot\mathbf{e}_1(y) = \sin\theta\, a^0(y) + a^1(y), \end{aligned} \tag{25}$$

using $\mathbf{e}_1(y)\cdot\mathbf{e}_0(y) = \mathbf{e}_0(y)\cdot\mathbf{e}_1(y) = \cos(90-\theta) = \sin\theta$, per Figure 2a. Equations (25) in matrix form are

$$\begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix} = \begin{bmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix}. \tag{26}$$

Comparing the vectors on either side of this relationship we surmise that the matrix / tensor

$$g_{\mu\nu}(y) = \begin{bmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{bmatrix} \tag{27}$$

in equation (25) is a machine for lowering the index of a vector. In indicial notation, we thus have the formal index-lowering procedure:

$$a_\mu(y) = g_{\mu\nu}(y) a^\nu(y). \tag{28}$$

The inverse of $g_{\mu\nu}(y)$ is also of interest. Calculating it explicitly we obtain

$$(g_{\mu\nu}(y))^{-1} = \frac{1}{\cos^2\theta} \begin{bmatrix} 1 & -\sin\theta \\ -\sin\theta & 1 \end{bmatrix}. \tag{29}$$

To be a valid mathematical quantity, this inverse must be able to act on both sides of equation (28), which means it must act on a vector with a lowered index, and produce a result with a single upper free index. This is suggestive that it itself must have two upper indices; in fact it is conventionally expressed as, simply:

$$(g_{\mu\nu}(y))^{-1} = g^{\mu\nu}(y). \tag{30}$$

This produces the formal index-raising procedure:

$$a^\mu(y) = g^{\mu\nu}(y) a_\nu(y). \tag{31}$$

In our example oblique system, then, the processes of index-lowering (i.e., transformation from contravariant to covariant components) and the process of index-raising (i.e., transformation from covariant to contravariant components), occur via

$$g_{\mu\nu}(y) = \begin{bmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{bmatrix}, \quad g^{\mu\nu}(y) = \frac{1}{\cos^2\theta} \begin{bmatrix} 1 & -\sin\theta \\ -\sin\theta & 1 \end{bmatrix}, \tag{32}$$

respectively. Since contravariant and covariant components are equal in the orthogonal x system, we can add to this the trivial case

$$g_{\mu\nu}(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad g^{\mu\nu}(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{33}$$

The metric tensor

The squared length of the vector $\mathbf{a}$:

$$|\mathbf{a}|^2 = a^\mu a_\mu, \tag{34}$$

appears to call for both the contravariant and covariant components of $\mathbf{a}$ to be used in combination. To calculate $|\mathbf{a}|^2$ in terms of contravariant components only, we can use the "index lowering" tensor $g_{\mu\nu}$ on the second of the two quantities in equation (34):

$$|\mathbf{a}|^2 = a^\mu (g_{\mu\nu} a^\nu) = \left[a^0(y)\ a^1(y)\right] \begin{bmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} a^0(y) \\ a^1(y) \end{bmatrix}. \tag{35}$$

To do so in terms of covariant components only, we can instead use the "index raising" tensor $g^{\mu\nu}$ on the first of the two quantities:

$$|\mathbf{a}|^2 = (g^{\mu\nu}a_\nu)a_\mu = \frac{1}{\cos^2\theta}\left[a_0(y)\ a_1(y)\right]\begin{bmatrix} 1 & -\sin\theta \\ -\sin\theta & 1 \end{bmatrix}\begin{bmatrix} a_0(y) \\ a_1(y) \end{bmatrix}. \tag{36}$$

In the x system all this is trivial, because both $g_{\mu\nu}(x)$ and $g^{\mu\nu}(x)$ are identity operators. But, in the y system, where the contravariant and covariant components are numerically different, the detailed forms of these g tensors are a critical part of the calculation, responsible for ensuring that $|\mathbf{a}|^2$ itself is preserved. In fact, everything we need to know in order to assign measures, or metrics, in a coordinate system y is included in $g^{\mu\nu}(y)$ and $g_{\mu\nu}(y)$. They are referred to for this reason as *metric tensors*.

Individual indices of tensors can be raised and lowered in the same way as those of as well as vectors. For instance,

$$s^{\mu\nu}_{\alpha} = g^{\nu\eta}\, s^{\mu}_{\alpha\eta}. \tag{37}$$

One might observe that $g^{\mu\nu}$ and $g_{\mu\nu}$ are themselves tensors, and then ask if it were possible to use $g^{\mu\nu}$ to raise one of the indices of $g_{\mu\nu}$. The answer is yes, in fact

$$g^{\mu\alpha}g_{\alpha\nu} = g^{\mu}_{\nu}. \tag{38}$$

But recall that the two-upper-index and two-lower-index metric tensors are just inverses of one another, which means that a g tensor with one lower and one upper index is the identity operator, or Kronecker delta:

$$g^{\mu}_{\nu} = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \delta^{\mu}_{\nu}. \tag{39}$$

Scalar functions

Consider a scalar field ϕ, which is a number assigned to every possible vector in some space. For instance, in a 2D environment like the one we have been considering, we can define a $\phi = \phi(x^\mu) = \phi(x^0, x^1)$, which assigns a number ϕ to every possible position vector x^μ in the plane. The vector x^μ points to a unique position on the plane, and if we point to the same position with a vector in a different coordinate system, say y^μ, we must wind up with the same number ϕ. So, although it itself is not a vector, ϕ, and in particular the way ϕ varies with position, is nevertheless subject to transformation rules.

Gradients

We might ask how ϕ varies when we take the small but finite step δx^μ. From a Taylor's series expansion in the 2D environment, we have, to first order,

$$\phi(x^\mu + \delta x^\mu) = \phi(x^\mu) + \frac{\partial\phi}{\partial x^0}\delta x^0 + \frac{\partial\phi}{\partial x^1}\delta x^1. \tag{40}$$

The variation of ϕ is defined as $\delta\phi(x^\mu) = \phi(x^\mu + \delta x^\mu) - \phi(x^\mu)$, so from equation (40) and using the comma and summation notation we have

$$\delta\phi(x^\mu) = \phi_{,\mu}\delta x^\mu, \tag{41}$$

with $\phi_{,\mu} = \partial\phi/\partial x^\mu$ being the gradient of ϕ. Notice that, if displacements like δx^μ are the prototypes for contravariant vectors, then the need for $\delta\phi$ to be a scalar forces the gradient to be expressed as a covariant vector. Gradients are in fact the prototypes for covariant vectors. This means in moving from one oblique coordinate system to another, the gradient transforms according to the rule in equation (20):

$$\phi_{,\mu}(y) = \frac{\partial x^\nu}{\partial y^\mu}\frac{\partial\phi}{\partial x^\nu} = x^\nu_{,\mu}\phi_{,\nu}(x),\ \phi_{,\mu}(x) = \frac{\partial y^\nu}{\partial x^\mu}\frac{\partial\phi}{\partial y^\nu} = y^\nu_{,\mu}\phi_{,\nu}(y). \tag{42}$$

CONCLUSIONS

Let us finalize this paper by summarizing the transformation rules. The above x,y-centric terminology needs to be translated to vector/tensor rules in the terms we will use in the companion papers in this report. First, as a reference coordinate system, rather than x we adopt the label s. Reference sets of parameters representing geophysical unknowns will be organized into position vectors s^μ in such systems ("model vectors"). These reference systems can be considered to be Cartesian or non-Cartesian as needed. Although in the earlier review, and in many tensor texts, indices start from 0, we will also now move to a numbering system starting at 1. We will consider transformations from objects in s to counterparts in a second system, which we will assume is non-Cartesian. This system will be labelled r, rather than y. Under transformation model vectors will be denoted r^μ. Similarly to the rule set out in (15), displacements in the s system will be related to displacements in the r system through

$$\delta s^\mu = s^\mu_{,\nu}\delta r^\nu = t^\mu_\nu \delta r^\nu. \tag{43}$$

The coefficients of the transformation are, as before, the derivatives $s^\mu_{,\nu} = \partial s^\mu/\partial r^\nu$, but for convenience we will write them as components of the tensor t^μ_ν. All subsequent rules for transformation of contravariant and covariant components from either system to its counterpart then follow. We will include both general expressions and explicit 2D examples.

To transform the contravariant components of a vector $\mathbf{a}$ from r to s, use:

$$a^\mu(s) = t^\mu_\nu a^\nu(r), \quad \text{or} \quad \begin{bmatrix} a^1(s) \\ a^2(s) \end{bmatrix} = \begin{bmatrix} t^1_1 & t^1_2 \\ t^2_1 & t^2_2 \end{bmatrix} \begin{bmatrix} a^1(r) \\ a^2(r) \end{bmatrix}. \tag{44}$$

To transform the contravariant components of a vector $\mathbf{a}$ from s to r, use:

$$a^\mu(r) = (t^{-1})^\mu_\nu a^\nu(s), \quad \text{or} \quad \begin{bmatrix} a^1(r) \\ a^2(r) \end{bmatrix} = \begin{bmatrix} t^1_1 & t^1_2 \\ t^2_1 & t^2_2 \end{bmatrix}^{-1} \begin{bmatrix} a^1(s) \\ a^2(s) \end{bmatrix}. \tag{45}$$

To transform the covariant components of a vector $\mathbf{a}$ from r to s, use:

$$a_\mu(s) = (t^{-1})^\nu_\mu a_\nu(r), \quad \text{or} \quad \begin{bmatrix} a_1(s) \\ a_2(s) \end{bmatrix} = \left(\begin{bmatrix} t^1_1 & t^1_2 \\ t^2_1 & t^2_2 \end{bmatrix}^{-1} \right)^T \begin{bmatrix} a_1(r) \\ a_2(r) \end{bmatrix}. \tag{46}$$

To transform the covariant components of a vector $\mathbf{a}$ from s to r, use:

$$a_\mu(r) = t^\nu_\mu a_\nu(s), \quad \text{or} \quad \begin{bmatrix} a_1(r) \\ a_2(r) \end{bmatrix} = \begin{bmatrix} t^1_1 & t^1_2 \\ t^2_1 & t^2_2 \end{bmatrix}^T \begin{bmatrix} a_1(s) \\ a_2(s) \end{bmatrix}. \tag{47}$$

The transformation matrix itself can be derived for particular systems s and r based on its definition $t^\mu_\nu = s^\mu_{,\nu}$, but it may be more helpful to design it geometrically. If one supposes that the reference s system is Cartesian and the r system is not, and one has a sense of the desired basis vectors of the r system, as resolved in the s system, then organizing these as columns of an $N \times N$ matrix we correctly generate t^μ_ν.

ACKNOWLEDGEMENTS

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grants CRDPJ 461179-13 and CRDPJ 543578-19, and in part by an NSERC-DG.

REFERENCES

Arfken, and Weber, 2001, Mathematical Methods for Physicists: Harcourt/Academic Press, San Diego, 5th edn.

Dirac, P. A. M., 1975, General theory of relativity: Princeton University Press, 1st edn.