# Chapter 0

# 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 $d$ can be described as sets that can be put into a 1-1 correspondence with regular rectangular grids of points in $\mathbb{R}^d$ whose coordinates are integers, used as index labels:

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

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, $2 \times 3$ 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 $d$-dimensional arrays for $d > 2$, 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. (We will sometimes use an over arrow symbol to signal a vector in $\mathbb{R}^n$ for emphasis, but not always.) 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 $\mathbb{R}^3$, 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, which in this example, streamlines the notation for taking a "linear combination of variables," namely the sum of the matched products of corresponding coefficients and variables.

This index positioning notation encodes the distinction between rows and columns in the matrix notation for a linear transformation. We will represent a matrix $(a_{ij})$ representing a linear transformation instead as $(a^i{}_j)$ 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

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

where the entries of a single row matrix are labeled by the column index (down), and the entries of a single column matrix are labeled by the row index (up).

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

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

If we agree to use an underlined kernel symbol $\underline{x}$ for a column vector, and the transpose $\underline{a}^T$ for a row vector, where the transpose simply interchanges rows and columns of a matrix, this can be represented as $\underline{a}^T \underline{x} = \vec{a} \cdot \vec{x}$. Since many geometric objects also have component matrices, it will be useful to link them together by using the same kernel symbol and underlining the matrix symbol to distinguish it from the object from which the components are taken.

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

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

Thinking of the coefficient matrix as a 1-dimensional vertical array of row vectors (the first right hand side of this sequence of equations), 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. Denoting the left matrix factor by $\underline{A}$, then the product column matrix has entries

$$[\underline{A}\,\underline{x}]^i = \sum_{k=1}^{3} a^i{}_k x^k = a^i{}_k x^k , \qquad 1 \le i \le 2 .$$

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

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

Denoting the new left matrix factor again by $\underline{A}$ and the right matrix factor by $\underline{X}$, then the product matrix has entries (row index left up, column index right down)

$$[\underline{A}\,\underline{X}]^i{}_j = \sum_{k=1}^{3} a^i{}_k x^k{}_j = a^i{}_k x^k{}_j , \qquad 1 \le i \le 2 ,\ 1 \le j \le 2 ,$$

where the sum over three entries (representing the dot product) is implied by our summation convention in the second equality, and the row and column indices here go from 1 to 2 to label the entries of the 2 rows and 2 columns of the product matrix. Thus matrix multiplication in this example is just an organized way of displaying all such dot products of two ordered sets of vectors in an array where the rows of the left factor in the matrix product correspond to the coefficient vectors in the left set and the columns in the right factor in the matrix product correspond to the variable vectors in the right set. The dot product itself in this context of matrix multiplication is representing the natural evaluation of linear functions (left row) on vectors (right column). No geometry (lengths and angles in Euclidean geometry) is implied in this context, only linearity and the process of linear combination.

The matrix product of a matrix with a single 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

$$\begin{aligned} \begin{pmatrix} \underline{v}_1 & \underline{v}_2 & \underline{v}_3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} &= \begin{pmatrix} v^1{}_1 & v^1{}_2 & v^1_3 \\ v^2{}_1 & v^2{}_2 & v^2_3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} v^1{}_1 x^1 + v^1{}_2 x^2 + v^1{}_3 x^3 \\ v^2{}_1 x^1 + v^2{}_2 x^2 + v^2{}_3 x^3 \end{pmatrix} \\ &= x^1 \begin{pmatrix} v^1{}_1 \\ v^2{}_1 \end{pmatrix} + x^2 \begin{pmatrix} v^1{}_2 \\ v^2{}_2 \end{pmatrix} + x^3 \begin{pmatrix} v^1{}_3 \\ v^2{}_3 \end{pmatrix} = x^1 \underline{v}_1 + x^2 \underline{v}_2 + x^3 \underline{v}_3 = x^i \underline{v}_i \,. \end{aligned}$$

Thus in this case the summed-over index pair performs a linear combination of the columns of the left factor of the matrix product, whose coefficients are the entries of the right column matrix factor. This interpretation extends to more columns in the right matrix factor, leading to a matrix product consisting of the same number of columns, each of which represents a linear combination of the column vectors of the left factor matrix. 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 $2 \times 3$ matrix multiplied by a $3 \times 1$ matrix), one is dealing with coefficient vectors $\underline{v}_i$ and vectors $\underline{x}$ of different dimensions, in this example combining three 2-component vectors by linear combination.

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 $\mathbb{R}^3$ is our basic vector space consisting of all possible ordered triplets of real numbers, and the space of all linear functions on it is equivalent to another copy of $\mathbb{R}^3$, 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 their matrix of coefficients 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 $d$-dimensional arrays for any positive integer $d$. Index positioning together with the Einstein summation convention is essential in letting the notation itself directly carry the information about its role in this scheme of linear mathematics extended beyond the elementary level.

Combining this linear algebra structure with multivariable calculus leads to differential geometry. Consider $\mathbb{R}^3$ 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 $\partial_i = \partial/\partial x^i$ for the partial derivative operator and then the abbreviation $f_{,i}$ for the corresponding partial derivatives of the function $f$. At each point of $\mathbb{R}^3$, the differentials $df$ and $dx^i$ 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 a tangent vector $\vec{v}$ at a given point are the values of the partial derivative functions there, and hence have indices down compared to the up indices of the tangent vector itself, which belongs 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 $\partial f/\partial x^i$ 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 $\mathbb{R}^3$ 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 $\mathbb{R}^3$ discussion where $\vec{x} = \langle x^1, x^2, x^3\rangle = x^i e_i$ is expressed as a linear combination of the standard basis vectors $\{e_i\} = \{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ usually denoted by $\mathbf{i}, \mathbf{j}, \mathbf{k}$? This question will be answered in the following notes, making the link between old fashioned tensor analysis and modern differential geometry.

One last remark about matrix notation is needed. We adopt here the notational conventions of the computer algebra system Maple for matrices and vectors. A vector $\langle u^1, u^2\rangle$ will be interpreted as a column matrix in matrix expressions

$$\underline{u} = \langle u^1, u^2\rangle = \begin{pmatrix} u^1 \\ u^2 \end{pmatrix}$$

while its transpose will be denoted by

$$\underline{u}^T = \langle u^1 | u^2\rangle = \begin{pmatrix} u^1 & u^2 \end{pmatrix} \,.$$

In other words within triangle bracket delimiters, a comma will represent a vertical separator in a list, while a vertical line will represent a horizontal separator in a list. A matrix can then be represented as a vertical list of rows or as a horizontal list of columns, as in

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \langle\langle a|b\rangle, \langle c|d\rangle\rangle = \langle\langle a,c\rangle|\langle b,d\rangle\rangle \,.$$

Finally if $\underline{A}$ is a matrix, we will not use a lowercase letter $a^i{}_j$ for its entries but retain the same symbol: $\underline{A} = (A^i{}_j)$.

Since the matrix notation and matrix multiplication which suppresses all indices and the summation is so efficient, it is important to be able to translate between the summed indexed notation to the corresponding index-free matrix symbols. In the usual language, matrix multiplication the $i$th row and $j$th column entry of the product matrix is

$$[\underline{A}\,\underline{B}]_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}\,,$$

while in our streamlined notation when these represent linear transformations of the vectors space into itself, this becomes

$$[\underline{A}\,\underline{B}]^{i}{}_{j} = A^{i}{}_{k} B^{k}{}_{j}\,.$$

However, as we will see all other index position combinations are possible with corresponding different meanings. In our application of the matrix product to matrices with indices in various up/down positions, the left index will always be the row index and the right index the column index and to translate from indexed notation to symbolic matrices we always have to use the above correspondence independent of the index up or down position: only left-right position counts. Thus to translate an expression like $M_{ij} B^{i}{}_{m} B^{j}{}_{n}$ we need to first rearrange the factors to $B^{i}{}_{m} M_{ij} B^{j}{}_{n}$ and then recognize that the second summed index $j$ is in the right adjacent pair of positions for interpretation of matrix multiplication, but the first summed index $i$ is in the row instead of column position so the transpose is required to place it adjacent to the middle matrix factor

$$(B^{i}{}_{m} M_{ij} B^{j}{}_{n}) = ([\underline{B}^{T}\,\underline{M}\,\underline{B}]_{mn}) = \underline{B}^{T}\,\underline{M}\,\underline{B}\,.$$

## Geometry?

Finally it is important to give some sense of what all this index business is needed for, connecting up with what has already been encountered in undergraduate multivariable calculus. In particular, what does all of this have to do with geometry? First, the term differential geometry refers to the study of the differential structure of "manifolds" which encompass not only the familiar straight line, flat plane and flat space of multivariable calculus but more complicated "curved spaces" like circles, spheres and cylinders or other conic-section related surfaces or more general surfaces in space as well as higher dimensional examples like the 3-sphere within 4-dimensional Euclidean space, a space which is often encountered in popular discussions of closed universe models. Second, geometry we first encounter in the context of lengths of line segments or angles between them (high school geometry class!). This idea of lengths and angles underlies the foundations of Riemannian (or pseudo-Riemannian) geometry in which the differential structure of differential geometry is given an additional "metric structure" that we associate with the distance formula and the dot product in multivariable calculus. To explain in more detail, we need more preliminary tools, but we can begin with an example that gives us a preview of what a metric is.

When we study arclengths of curves, we easily accept the expression for the square of the differential of arclength for a differential displacement in the plane from a point $(x, y)$ to $(x + dx, y + dy)$ as the Pythagorean relation for right triangles

$$ds^2 = dx^2 + dy^2 = \begin{pmatrix} dx & dy \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}\,, \tag{1}$$

which is very useful for integrating up finite arclengths along curves once the curve is parametrized. We also learn to re-express this formula passing from the standard Cartesian coordinates to polar coordinates in the plane by way of the coordinate transformation

$$(x,y) = (r\sin\theta, r\cos\theta) \tag{2}$$

whose differential

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \cos\theta\, dr - r\sin\theta\, d\theta \\ \sin\theta\, dr + r\cos\theta\, d\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} dr \\ d\theta \end{pmatrix} \equiv \underline{J} \begin{pmatrix} dr \\ d\theta \end{pmatrix}, \tag{3}$$

which defines the so-called Jacobian matrix $\underline{J}$ of partial derivatives of the old coordinates with respect to the new ones. It is then a simple matter to substitute these relations into the square of the differential arclength, and expand and simplify the result using the fundamental trigonometric identity

$$ds^2 = (\cos\theta\, dr - r\sin\theta\, d\theta)^2 + (\sin\theta\, dr + r\cos\theta\, d\theta)^2 = \ldots = dr^2 + r^2 d\theta^2 . \tag{4}$$

In terms of the matrix representation of this same calculation, we have instead

$$ds^2 = \begin{pmatrix} dx \\ dy \end{pmatrix}^T \begin{pmatrix} dx \\ dy \end{pmatrix} = \left(\underline{J} \begin{pmatrix} dr \\ d\theta \end{pmatrix}\right)^T \left(\underline{J} \begin{pmatrix} dr \\ d\theta \end{pmatrix}\right) = \begin{pmatrix} dr \\ d\theta \end{pmatrix}^T \underline{J}^T \underline{J} \begin{pmatrix} dr \\ d\theta \end{pmatrix}, \tag{5}$$

where

$$\underline{J}^T \underline{J} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \tag{6}$$

and we have used the well known property $(\underline{A}\,\underline{B})^T = \underline{B}^T \underline{A}^T$ of the matrix transpose which converts columns into rows and vice versa.

The quantity $ds^2$ is a quadratic function of the coordinate differentials, or a "quadratic form," usually called the "line element" in differential geometry. The diagonal matrix $g = \underline{J}^T \underline{J}$ consists of the coefficients of this quadratic form, or the components of the "metric."

But so far no indices! For that we have to introduce superscripted coordinate variables. The old and new coordinates are

$$(x^1, x^2) = (x,y)\,,\ \ (y^1, y^2) = (r,\theta) \tag{7}$$

and we can let $i, j, k, \ldots = 1, 2$. Then their differential relationship is

$$dx^i = \frac{\partial x^i}{\partial y^j} dy^j \equiv J^i{}_j dy^j \tag{8}$$

and

$$ds^2 = g_{ij} dy^i dy^j\,, \quad g_{ij} = \left(\underline{J}^T \underline{J}\right)_{ij} = \sum_{m=1}^{2} J^m{}_i J^m{}_j \tag{9}$$

defines the metric expressed in the new coordinates. More precisely, the components of the metric $g_{ij}$ are the entries of the matrix of the quadratic form represented by the "line element"

$ds^2$. The only way to get rid of the summation symbol here is to introduce the unit matrix with both indices down to be able to use the summation convention

$$\underline{I} = (\delta_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

Then we can write

$$g_{ij} = J^m{}_i J^n{}_j \delta_{mn} = J^m{}_i \delta_{mn} J^n{}_j = (\underline{J}^T \underline{I}\, \underline{J})_{ij} .$$

In fact since lowered indices are associated with coefficients of linear forms, it makes sense that the coefficients of a bilinear quadratic form $g_{ij}X^iY^j$ also have lowered indices.

In my institution the section of the textbook on changes of variables and the Jacobian matrix is omitted from the multivariable calculus syllabus, but as you can see, it easily determines the differential arclength in non-Cartesian coordinates as in this toy calculation. It also determines the correction factor for the differential of area in the plane in polar coordinates $dA = r\, dr d\theta$ through its determinant $|\underline{J}| = r$, which we instead alternatively derive using the formula for the area of a sector of a circle. In fact while the dot product and its generalization to a metric determine lengths of line segments and arclengths of curves, the 1-dimensional measure associated with geometry, it is the determinant and its generalizations that allow this 1-dimensional measure to be extended to higher dimensional structures like parallelograms and parallelopipeds and differentials of surface area and volume in non-Cartesian coordinates in $\mathbb{R}^3$. In $\mathbb{R}^n$ with $n > 3$, there are $p = 1, 2, \ldots, n-1, n$ dimensional structures and measures, all governed by the mathematics of determinants and Jacobians.

Old fashioned tensor calculus deals with understanding differential properties of spaces described by different coordinate systems, giving preference to those quantities which do not depend on the choice of coordinates and therefore have to "transform" in a certain way to guarantee that coordinate independence. The Jacobian is the matrix of the linear transformation of derivatives between different coordinate systems. Calculations with it can become tedious, as we will see when we later study the corresponding derivation for spherical coordinates in space. Fortunately computer algebra systems can now save us from a lot of the hand algebra that becomes quite cumbersome in this subject. For this reason it is important to use either Maple or Mathematica as a tool in learning the ropes of this area of mathematics as one works problems to gain familiarity with the concepts. These software products have slightly different conventions from each other and the classical notation of the discipline, so we have to be a bit flexible in dealing with notation. Modern differential geometry organizes old fashioned tensor calculus in such a way that we deal with invariant objects instead of collections of components which change with the choice of coordinate system, like the notion of a vector as an abstract arrow rather than as a list of components, but those components are still implicit in everything we do, even if we do it with a modern flair.

**Exercise 0.0.1.**
**arclength in the plane**

a) Evaluate $ds^2$ for polar coordinates, filling in the lower dots in the above derivation.

b) Evaluate $g_{ij}$ for polar coordinates from the matrix product $\underline{J}^T \underline{J}$.

c) If you want to try something unfamiliar, evaluate $ds^2$ for $x = uv, y = \frac{1}{2}(u^2 - v^2)$. This too results in a sum of squares, which characterizes what are called "orthogonal coordinates" as we will learn about later. These particular orthogonal coordinates in the plane are called parabolic coordinates since the coordinate lines for both coordinates are parabolas. We will return to these much later.

■

**Exercise 0.0.2.**
**matrix multiplication and the trace**

The matrix equation defining the inverse of a matrix $\underline{A}^{-1}\underline{A} = \underline{I}$ can be written with our index conventions as ${A^{-1}}^i{}_j A^j{}_k = \delta^i{}_k$, where here we need one index up and one index down on the identity matrix component representation in this context for consistency with our index conventions (the identity matrix plays a different role here than above!). The identity matrix property $\underline{I}\,\underline{A} = \underline{A}$ translates to $\delta^i{}_j A^j{}_k = A^i{}_k$. Given that the trace is defined by $\mathrm{Tr}\,\underline{A} = A^i{}_i$, write out the matrix identity $\mathrm{Tr}\left(\underline{A}^{-1}\underline{B}\,\underline{A}\right) = \mathrm{Tr}\,\underline{B}$ in our index notation and use the inverse property to show why the left hand side reduces to the right hand side, thus proving the identity easily. Similarly show the more general property of the trace holds: $\mathrm{Tr}\,(\underline{A}\,\underline{B}\,\underline{C}) = \mathrm{Tr}\,(\underline{C}\,\underline{A}\,\underline{B})$, etc., for cyclic permutations of the factor matrices. These are simple examples of how this streamlined index notation for linear behavior easily allows one to prove identities that otherwise are not obvious.

■