Tensor Analysis I [gmd5-A]

Introduction:

In the module entitled *Matrices* [gmd6], we begin with lists of elements, where the elements are (real or complex) numbers or functions. Matrices are lists organized into arrays of n columns and m rows.

We investigate matrix specifications, attributes, and operations. We identify $n \times 1$ matrices as column vectors and $1 \times m$ matrices as row vectors. However, there is more to vectors than being matrices of one column or row.

In the module entitled *Vector Analysis* [gmd1], the focus is on vectors and vector fields in 3D Euclidean space. It goes without saying that these quantities feature prominently in the formulation of many laws of physics.

– Newton's second law, $d\mathbf{p}/dt = \mathbf{F}$, is a relation between the derivative of the vectors $\mathbf{p}$ (momentum) with respect to the scalar t (time) and the vector $\mathbf{F}$ (force) in a coordinate-independent equation.

– Maxwell's equations are four (coordinate-independent) relations between the vector fields $\mathbf{E}$ (electric field), $\mathbf{B}$ (magnetic field), $\mathbf{J}$ (current density), and the scalar field ρ (charge density).

When we solve a physics problem, we apply laws of physics by using coordinates adapted to the symmetry of the situation. The module entitled *Coordinate Systems* [gmd2] is designed to help find the optimal choice.

Coordinate systems are not a part of nature. They are mathematical tools employed for working out predictions extractable from laws of nature.

Therefore, laws of physics, understood to be laws of nature, must not depend on the choice of coordinate system.

Not all coordinate-independent relations between physical quantities (including laws of physics) can be expressed by scalars and vectors. Tensors are natural extensions of scalars and vectors.

– Scalars are rank-zero tensors, expressible as numbers or functions in any coordinate system.

– Vectors are rank-one tensors (columns/rows of components in any system of coordinates). Each component is a number or a function.

– Rank-two tensors are expressible as square matrices of elements, where each element is a number or a function.

– Rank-n tensors are arrays of elements with n indices.

Tensors of any rank must satisfy specific conditions under coordinate transformations to guarantee coordinate-independent relations between physical quantities (including laws of physics).

- Quantities which do not change under a coordinate transformation are named *invariants* or *scalars*. They are rank-zero tensors. Within the context of Newtonian mechanics, time t and mass m are invariant.
- A coordinate-free visualization of vectors (rank-1 tensors) in 3D Euclidean space employs arrows with length and direction. The components of vectors in different (Cartesian) coordinate systems amount to projections of the same arrow onto different sets of coordinate axes.
- Tensors of higher rank are harder to visualize geometrically. The need for them in the coordinate-free formulation of relations between physical quantities can best be demonstrated for the concept of inertia.

One source of confusion in tensor analysis is the word *covariant*, which has different meanings.

▷ The coordinate-independent formulation of a law of physics or any relation between physical quantities is called *covariant*.

  The equations of electrostatics in vector form are covariant in a limited sense, i.e. independent under coordinate transformations which include translations and rotations in 3D Euclidean space.

  Maxwell's equations of electrodynamic are covariant in a more general sense, also including transformations between coordinate systems in relative motion. The more general covariance is made explicit in 4D Minkowski spacetime.

▷ A tangent vector is generated as the derivative of a vector with respect to a scalar (e.g. velocity from position). This type of vector is named *contravariant*.

  A gradient vector is generated as the derivative of a scalar with respect to a vector (e.g. gravitational field from gravitational potential). This type of vector is named *covariant*.

  Tensors of rank two or higher can be contravariant, covariant, or mixed. The status may vary between indices.

  Superscript indices signal contravariance and subscript indices covariance. The distinction between contravariance and covariance i.e. (upper and lower indices) does not matter for all metrics.

Inertia tensor:

Consider a rigid body of mass m in purely translational motion. A coordinate-independent formulation of the relation between momentum $\mathbf{p}$ and velocity $\mathbf{v}$ (two vectors) involves the mass m (a scalar):

$$\mathbf{p} = m\mathbf{v}, \quad p_i = mv_i. \tag{1}$$

The two vectors have the same directions at all times, irrespective of the forces in operation. The second version uses tensor notation.[1]

A purely rotational motion of the same rigid body is described by two different vectors: the angular momentum $\mathbf{L}$ and the angular velocity $\boldsymbol{\omega}$.

The relation between these two vectors is more complex. They are not parallel, in general. A coordinate-independent formulation is possible and involves the (rank-two) inertia tensor $\mathbf{I} = I_{ij}$ [gex96], [gex97]:

$$\mathbf{L} = \mathbf{I} \cdot \boldsymbol{\omega}, \quad L_i = \sum_j I_{ij}\omega_j. \tag{2}$$

The first version relates two column vectors via matrix multiplication. The second version employs tensor notation.

The inner product in the first expression becomes a tensor contraction (sum over the repeated index j) in the second expression.

The inertia tensor turns out to be symmetric, $I_{ij} = I_{ji}$. Hence there exists a coordinate system for which it is diagonal: $\bar{I}_{ij} = \bar{I}_i\delta_{ij}$, where the $\bar{I}_i$ are the principal moments of inertia.

For this coordinate system, Eq. (2) reduces to $L_i = \bar{I}_i\omega_i$. The vectors $\mathbf{L}$ and $\boldsymbol{\omega}$ are parallel if the rotation is about a principal axis.

Kinetic energy K is a scalar quantity. The coordinate-independent expressions for translational motion and rotational motion read as follows:

$$K_{\text{trans}} = \frac{1}{2}m\mathbf{v} \cdot \mathbf{v} = \frac{1}{2}\sum_i mv_iv_i, \tag{3}$$

$$K_{\text{trans}} = \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{I} \cdot \boldsymbol{\omega} = \frac{1}{2}\sum_{ij} I_{ij}\omega_i\omega_j. \tag{4}$$

[1]The distinction between superscript indices and subscript indices is without significance in this application. Also, the summation convention is set aside here.

From matrices to tensors:

From [gmd6] we know that matrices (including row or column vectors) are elements organized into arrays, where each element is a number or a function.

Tensors of rank one or two can be represented as matrices in any particular coordinate system, which means that all matrix attributes and matrix operations discussed in [gmd6] are applicable to tensors.

Here we summarize the linear algebra of rank-1 and rank-2 tensors with emphasis on tensor notation including the summation convention. Keep in mind that there is more to tensors than their linear algebra.

– Rank-1 tensor (vector):

$$\mathbf{v} \doteq v_i = (v_1, \ldots, v_n), \quad \mathbf{w} \doteq w^i = (w^1, \ldots, w^n). \tag{5}$$

– Rank-2 tensor (matrix):

$$\mathbf{a} \doteq a_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad \mathbf{b} \doteq b^{ij} = \begin{pmatrix} b^{11} & \cdots & b^{1n} \\ \vdots & & \vdots \\ b^{n1} & \cdots & b^{nn} \end{pmatrix},$$
$$\mathbf{c} \doteq c^i{}_j = \begin{pmatrix} c^1{}_1 & \cdots & c^1{}_n \\ \vdots & & \vdots \\ c^n{}_1 & \cdots & c^n{}_n \end{pmatrix}. \tag{6}$$

– Matrix multiplication:

$$\mathbf{a} \doteq a_{ij}, \quad \mathbf{b} \doteq b^{ij}, \quad \mathbf{p} = \mathbf{a} \cdot \mathbf{b} \quad \Leftrightarrow \quad p_i{}^k = a_{ij} b^{jk}. \tag{7}$$

– Identity matrix:

$$\mathbf{I} = \delta_{ij} = \delta^{ij} = \delta_i{}^j = \delta^i{}_j. \tag{8}$$

– Inverse matrix (Kronecker symbol):

$$\mathbf{a} \doteq a_{ij}, \quad \mathbf{b} \doteq b^{ij}, \quad \mathbf{b} = \mathbf{a}^{-1}$$
$$\Rightarrow \ \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} = \mathbf{I} \quad \Leftrightarrow \quad a_{ij} b^{jk} = \delta_i{}^k, \quad b^{ij} a_{jk} = \delta^i{}_k. \tag{9}$$

– Transpose:

$$\mathbf{a} \doteq a_{ij} \quad \Rightarrow \ \mathbf{a}^T = a_{ji}, \qquad \mathbf{b} \doteq b^{ij} \quad \Rightarrow \ \mathbf{b}^T = b^{ji} \tag{10}$$

– Symmetric [antisymmetric] matrix:

$$\mathbf{a} = \mathbf{a}^T \quad \Leftrightarrow \quad a_{ij} = a_{ji} \qquad \Big[\mathbf{b} = -\mathbf{b}^T \quad \Leftrightarrow \quad b_{ij} = -b_{ji}\Big]. \tag{11}$$

– Permutation symbol:

$$\epsilon_{ijk\cdots} = \begin{cases} (-1)^P & : \text{ no repeated indices,} \\ 0 & : \text{ at least one index repeated,} \end{cases} \tag{12}$$

where permutation $\{123\cdots\} \to \{ijk\cdots\}$ involves P transpositions.

– Determinant of matrix $\mathbf{a} \doteq a_{ij}$:

$$\mathrm{Det}[\mathbf{a}] = \epsilon_{i_1\cdots i_n} a_{1i_1} \cdots a_{ni_n}. \tag{13}$$

– Products between vectors $\mathbf{u} \doteq u_i,\ \mathbf{v} \doteq v_i$ (with real elements):

$$\mathbf{u} \cdot \mathbf{v} = u_i v_i = w \quad : \text{ scalar product,} \tag{14}$$

$$\mathbf{u} \times \mathbf{v} = \epsilon_{ijk} u_j v_k = w_i \quad : \text{ vector product.} \tag{15}$$

– Norm of a vector $\mathbf{u} \doteq u_i$:

$$||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_i u_i}. \tag{16}$$

– Linear equations with $\mathbf{a} \doteq a_{ij},\ \mathbf{b} \doteq b_i$ for vector $\mathbf{x} \doteq x_i$:

$$\mathbf{a} \cdot \mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad a_{ij} x_j = b_i. \tag{17}$$

– Inverse matrix [gmd6]:

$$\mathbf{x} = \mathbf{a}^{-1} \cdot \mathbf{b}. \tag{18}$$

– Quadratic form of vector $\mathbf{x} \doteq x_i$ with matrix $\mathbf{a} \doteq a_{ij}$:

$$q = \mathbf{x}^T \cdot \mathbf{a} \cdot \mathbf{x} = a_{ij} x_i x_j = a_{ij}^{(s)} x_i x_j, \quad a_{ij}^{(s)} = \frac{1}{2}(a_{ij} + a_{ji}). \tag{19}$$

– Linear coordinate transformation:[2]

$$\bar{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x} \quad \Rightarrow \ \mathbf{x} = \mathbf{A}^{-1} \cdot \bar{\mathbf{x}} \doteq \mathbf{B} \cdot \bar{\mathbf{x}}. \tag{20}$$

– Transformation of quadratic form $C_{ij} x_i x_j$:

$$\mathbf{x}^T \cdot \mathbf{C} \cdot \mathbf{x} = (\mathbf{B} \cdot \bar{\mathbf{x}})^T \cdot \mathbf{C} \cdot (\mathbf{B} \cdot \mathbf{x}) = \bar{\mathbf{x}}^T \cdot \underbrace{\mathbf{B}^T \cdot \mathbf{C} \cdot \mathbf{B}}_{\bar{\mathbf{C}}} \cdot \bar{\mathbf{x}}. \tag{21}$$

[2]Tensor analysis is mostly concerned with the (passive) *alias* interpretation of coordinate transformation, where $\bar{\mathbf{x}}$ and $\mathbf{x}$ represent the same point in different coordinate systems. In the (active) *alibi* interpretation, $\bar{\mathbf{x}}$ and $\mathbf{x}$ are different points in one coordinate system.

– Linear transformation of (invariant) distance:[3]

$$\begin{aligned} d(\mathbf{x},\mathbf{y}) \doteq ||\mathbf{x}-\mathbf{y}|| &= \sqrt{(\mathbf{x}-\mathbf{y})^T \cdot (\mathbf{x}-\mathbf{y})} \\ &= \sqrt{(\mathbf{B}\cdot\bar{\mathbf{x}}-\mathbf{B}\cdot\bar{\mathbf{y}})^T \cdot (\mathbf{B}\cdot\bar{\mathbf{x}}-\mathbf{B}\cdot\bar{\mathbf{y}})} \\ &= \sqrt{(\bar{\mathbf{x}}-\bar{\mathbf{y}})^T \cdot \mathbf{B}^T \cdot \mathbf{B} \cdot (\bar{\mathbf{x}}-\bar{\mathbf{y}})} \\ &= \sqrt{(\bar{\mathbf{x}}-\bar{\mathbf{y}})^T \cdot \mathbf{G} \cdot (\bar{\mathbf{x}}-\bar{\mathbf{y}})} \doteq d(\bar{\mathbf{x}},\bar{\mathbf{y}}), \end{aligned} \tag{22}$$

with $\mathbf{G} = \mathbf{B}^T \cdot \mathbf{B} = (\mathbf{A}\cdot\mathbf{A}^T)^{-1}$.

Special case of orthogonal transformation:[4] $\mathbf{A}^T = \mathbf{A}^{-1} \quad \Rightarrow \; \mathbf{G} = \mathbf{I}$.

– General coordinate transformation:

$$\bar{\mathbf{x}} = \mathbf{T}(\mathbf{x}) \quad \Leftrightarrow \quad \bar{x}_i = T_i(x_1,\ldots,x_n). \tag{23}$$

▷ Passive *alias* interpretation: $\mathbf{T}$ establishes a one-on-one correspondence between coordinates $\bar{\mathbf{x}}$ and $\mathbf{x}$ of a point P.

▷ Active *alibi* interpretation: Any point P with coordinates $\mathbf{x}$ in the domain of $\mathbf{T}$ has its unique image Q with coordinates $\bar{\mathbf{x}}$ in the range of $\mathbf{T}$ and vice versa.

[3] Linear coordinate transformations are called *affine*.

[4] Transformations between Cartesian, cylindrical, and spherical coordinates are orthogonal, but not linear (affine).

Tensors in real coordinate space:

The n-dimensional coordinate space is denoted $\mathbb{R}^n$.

Cartesian (rectangular) coordinates $x_i,\ i = 1, \ldots, n$ are defined by a specific form of the distance relation:

$$d = \sqrt{\delta_{ij}\Delta x^i \Delta y^j} = \sqrt{(x^1 - y^1)^2 + \cdots (x^n - y^n)^2}, \quad \Delta x^i \doteq x^i - y^i. \tag{24}$$

General coordinate transformation $\mathcal{T}:\ \bar{x}^i = \bar{x}^i(x^1, \ldots, x^n)$.

Inverse transformation $\mathcal{T}^{-1}:\ x^i = x^i(\bar{x}^1, \ldots, \bar{x}^n)$.

– Cylindrical coordinates $\bar{x}^1 \doteq \rho,\ \bar{x}^2 \doteq \phi,\ \bar{x}^3 \doteq z$ in $\mathbb{R}^3$:

$$\mathcal{T}: \begin{cases} \rho = \sqrt{(x^1)^2 + (x^2)^2} \\ \phi = \arctan\left(\dfrac{x^2}{x^1}\right) \\ z = x^3 \end{cases} \qquad \mathcal{T}^{-1}: \begin{cases} x^1 = \rho\cos\phi \\ x^2 = \rho\sin\phi \\ x^3 = z \end{cases} \tag{25}$$

– Spherical coordinates $\bar{x}^1 = r,\ \bar{x}^2 = \theta,\ \bar{x}^3 = \phi$ in $\mathbb{R}^3$:

$$\mathcal{T}: \begin{cases} r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} \\ \theta = \arccos\left(\dfrac{x^3}{\sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}}\right) \\ \phi = \arctan\left(\dfrac{x^2}{x^1}\right) \end{cases}$$

$$\mathcal{T}^{-1}: \begin{cases} x^1 = r\sin\theta\cos\phi \\ x^2 = r\sin\theta\sin\phi \\ x^3 = r\cos\theta \end{cases} \tag{26}$$

The one-on-one mapping required for a coordinate transformation $\mathcal{T}$ is guaranteed in any region of vanishing Jacobian determinant: $\mathrm{Det}[J] \neq 0$.

Jacobian matrix: $J \doteq \begin{pmatrix} \partial\bar{x}^1/\partial x^1 & \cdots & \partial\bar{x}^1/\partial x^n \\ \vdots & & \vdots \\ \partial\bar{x}^n/\partial x^1 & \cdots & \partial\bar{x}^n/\partial x^n \end{pmatrix}$.

Coordinate transformations, in general, involve two kinds of metric which are independent of each other: the metric of the space and the metric of the transformation (more about this later).

The metric of the transformation is a generalization of the distance function introduced earlier for linear transformations.

The transformations between rectangular and cylindrical or spherical coordinates are nonlinear and orthogonal. They are quite artificial in a tensor context [gex113], but important for considerations of symmetry [gmd2].

In physics applications of tensor analysis, the focus is on symmetry transformations.

Contravariance versus covariance:

Vector fields, in general, are lists of components (one row or one column), where each component is a scalar function of the space coordinates:

$$\mathbf{V} = \begin{pmatrix} V_1(x^1, \ldots, x^n) \\ \vdots \\ V_n(x^1, \ldots, x^n) \end{pmatrix}. \tag{27}$$

Such vector fields are candidates for rank-1 tensors under a specific coordinate transformation and its inverse:

$$\mathcal{T} : \ \bar{x}^i = \bar{x}^i(x^1, \ldots, x^n), \qquad \mathcal{T}^{-1} : \ x^i = x^i(\bar{x}^1, \ldots, \bar{x}^n). \tag{28}$$

Tensors of rank 1 must themselves transform in one or the other of two distinct ways under the coordinate transformation (28):

– Contravariant vector:

$$\bar{V}^i = V^j \frac{\partial \bar{x}^i}{\partial x^j} \doteq V^j \partial_j \bar{x}^i, \tag{29}$$

– Covariant vector:

$$\bar{U}_i = U_j \frac{\partial x^j}{\partial \bar{x}^i} \doteq U_j \partial_i \bar{x}^j. \tag{30}$$

▷ The last expression in (29) and (30) is a short-hand notation for partial derivatives often used when tensors are ubiquitous.

▷ Contravariant tensors are identified by superscript indices and covariant vectors by subscript indices.

▷ Tangent vectors to curves (e.g. velocity) are contravariant: $x^i(t)$

$$\Rightarrow \ v^i(t) = \frac{dx^i}{dt} \quad \Rightarrow \ \bar{v}^i = \frac{d\bar{x}^i}{dt} = \frac{\partial \bar{x}^i}{\partial x^j}\frac{dx^j}{dt} = v^j \frac{\partial \bar{x}^i}{\partial x^j}. \tag{31}$$

▷ Gradient vectors (of scalar field) are covariant: $\Phi(x^1, \ldots, x^n)$

$$
\begin{aligned}
&\Rightarrow\ \bar{\Phi}(\bar{x}^1, \ldots, \bar{x}^n) = \Phi\big(x^1(\bar{x}^1, \ldots, \bar{x}^n), \ldots, x^n(\bar{x}^1, \ldots, \bar{x}^n)\big) \\
&\quad\ \mathbf{F} \doteq \nabla\Phi \quad \Rightarrow\ F_i = \frac{\partial \Phi}{\partial x^i}, \quad \bar{F}_i = \frac{\partial \bar{\Phi}}{\partial \bar{x}^i} \\
&\Rightarrow\ \frac{\partial \bar{\Phi}}{\partial \bar{x}^i} = \frac{\partial \Phi}{\partial x^j}\frac{\partial x^j}{\partial \bar{x}^i} \quad \Rightarrow\ \bar{F}_i = F_j\,\frac{\partial x^j}{\partial \bar{x}^i}.
\end{aligned} \tag{32}
$$

Invariance:

Scalar quantities which do ot change under a corrdinate transformation are called invariants. They are rank-0 tensors. The scalar field field Φ used in (32) is an invariant by construction.

The inner product[5] of a contravariant vector V^i and a covariant vector U_i (two rank-1 tensors) can be shown to be an invariant (rank-0 tensor):

$$
\begin{aligned}
&\bar{V}^i = V^j\,\frac{\partial \bar{x}^i}{\partial x^j}, \qquad \bar{U}_i = U_j\,\frac{\partial x^j}{\partial \bar{x}^i} \\
\Rightarrow\ &\bar{E} = \bar{V}^i\bar{U}_i = V^j\,\frac{\partial \bar{x}^i}{\partial x^j}U_k\,\frac{\partial x^k}{\partial \bar{x}^i} = V^jU_k\,\underbrace{\frac{\partial \bar{x}^i}{\partial x^j}\frac{\partial x^k}{\partial \bar{x}^i}}_{\delta_j{}^k} = V^iU_j = E.
\end{aligned} \tag{33}
$$

Mixed variance:

Tensors of rank 2 exist in three types:

– contravariant: $\bar{T}^{ij} = T^{kl}\,\dfrac{\partial \bar{x}^i}{\partial x^k}\dfrac{\partial \bar{x}^j}{\partial x^l}$.

– covariant: $\bar{T}_{ij} = T_{kl}\,\dfrac{\partial x^k}{\partial \bar{x}^i}\dfrac{\partial x^l}{\partial \bar{x}^j}$.

– mixed: $\bar{T}^i{}_j = T^k{}_l\,\dfrac{\partial \bar{x}^i}{\partial x^k}\dfrac{\partial x^l}{\partial \bar{x}^j}$.

Tensors of rank 2 or higher or often characterized differently, namely of order $m = p + q$ when they have p superscript indices and q subscript indices.

The tensor $T^i{}_{jkl}$, for example, has order $1 + 3$.

[5]Tensor operations are a later topic. Inner products are like scalar products of vectors.

Affine tensors:

Physical quantities which transform as tensors under linear coordinate transformations are called *affine* tensors.

- Linear coordinate transformation: $\mathcal{T}: \; \bar{x}^i = a^i{}_j x^j$ with constant $a^i{}_j$.
- Jacobian and inverse: $J = \dfrac{\partial \bar{x}^i}{\partial x^j} = a^i{}_j, \quad J^{-1} = \dfrac{\partial x^i}{\partial \bar{x}^j} = b^i{}_j.$
- Contravariant tensors: $\bar{T}^i = a^i{}_j T^j, \quad \bar{T}^{ij} = a^i{}_k a^j{}_l T^{kl}, \ldots$
- Covariant tensors: $\bar{T}_i = b^j{}_i T_j, \quad \bar{T}_{ij} = b^k{}_i b^l{}_j T_{kl}, \ldots$
- Mixed tensors: $\bar{T}^i{}_j = a^i{}_k b^l{}_j T^k_l, \ldots.$

The position vector itself is a contravariant affine rank-1 tensor.

Cartesian tensors:

If we restrict the linear coordinate transformations to orthogonal ones, then more more physical quantities qualify as tensors. They are Cartesian tensors.

- Jacobian and inverse: $J = \dfrac{\partial \bar{x}^i}{\partial x^j} = a^i{}_j, \quad J^{-1} = J^T = \dfrac{\partial x^i}{\partial \bar{x}^j} = b^i{}_j = a^j{}_i.$
- Contravariant tensors: $\bar{T}^i = a^i{}_j T^j, \quad \bar{T}^{ij} = a^i{}_k a^j{}_l T^{kl}, \ldots$
- Covariant tensors: $\bar{T}_i = b^j{}_i T_j = a^i{}_j T_j, \quad \bar{T}_{ij} = b^k{}_i b^l{}_j T_{kl} = a^i{}_k a^j{}_l T_{kl}, \ldots$
- Mixed tensors: $\bar{T}^i{}_j = a^i{}_k b^l{}_j T^k_l = a^i{}_k a^j{}_l T^k_l, \ldots.$

The distinction between contravariance and covariance becomes artificial for Cartesian tensors. The notation is often simplified to subscripts only.