Covariance and contravariance of vectors

In multilinear algebra and tensor analysis, **covariance** and **contravariance** describe how the quantitative description of certain geometric or physical entities changes with a change of basis. For holonomic bases, this is determined by a change from one coordinate system to another. When an orthogonal basis is rotated into another orthogonal basis, the distinction between co- and contravariance is invisible. However, when considering more general coordinate systems such as skew coordinates, curvilinear coordinates, and coordinate systems on differentiable manifolds, the distinction is significant.

- For a vector (such as a direction vector or velocity vector) to be basis-independent, the components of the vector must contra-vary with a change of basis to compensate. That is, the components must vary with the inverse transformation to that of the change of basis. The components of vectors (as opposed to those of dual vectors) are said to be **contravariant**. Examples of vectors with contravariant components include the position of an object relative to an observer, or any derivative of position with respect to time, including velocity, acceleration, and jerk. In Einstein notation, contravariant components are denoted with *upper indices* as in

\[ \mathbf{v} = v^i \mathbf{e}_i. \]

- For a dual vector (also called a covector) to be basis-independent, the components of the dual vector must co-vary with a change of basis to remain representing the same covector. That is, the components must vary by the same transformation as the change of basis. The components of dual vectors (as opposed to those of vectors) are said to be covariant. Examples of covariant vectors generally appear when taking a gradient of a function. In Einstein notation, covariant components are denoted with *lower indices* as in

\[ \mathbf{v} = v_\ell \mathbf{e}_\ell. \]

In physics, vectors often have units of distance or distance times some other unit (such as the velocity), whereas covectors have units the inverse of distance or the inverse of distance times some other unit. The distinction between covariant and contravariant vectors is particularly important for computations with tensors, which can have mixed variance. This means that they have components that are both covariant and contravariant. The valence or type of a tensor gives the number of covariant and contravariant component indices.

The terms covariant and contravariant were introduced by J.J. Sylvester in 1853 in order to study algebraic invariant theory. In this context, for instance, a system of simultaneous equations is contravariant in the variables. The use of
both terms in the modern context of multilinear algebra is a specific example of corresponding notions in category theory.

**Introduction**

In physics, a vector typically arises as the outcome of a measurement or series of measurements, and is represented as a list (or tuple) of numbers such as

\[ (v_1, v_2, v_3). \]

This list of numbers depends on the choice of coordinate system. For instance, if the vector represents position with respect to an observer (position vector), then the coordinate system may be obtained from a system of rigid rods, or reference axes, along which the components \( v_1, v_2, \) and \( v_3 \) are measured. For a vector to represent a geometric object, it must be possible to describe how it looks in any other coordinate system. That is to say, the components of the vectors will transform in a certain way in passing from one coordinate system to another.

A **contravariant vector** is required to have components that "transform in the same way as the coordinates" (the opposite way as the reference axes) under changes of coordinates such as rotation and dilation. The vector itself does not change under these operations; instead, the components of the vector make a change that cancels the change in the spatial axes, in the same way that co-ordinates change. In other words, if the reference axes were rotated in one direction, the component representation of the vector would rotate in exactly the opposite way. Similarly, if the reference axes were stretched in one direction, the components of the vector, like the co-ordinates, would reduce in an exactly compensating way. Mathematically, if the coordinate system undergoes a transformation described by an invertible matrix \( M \), so that a coordinate vector \( x \) is transformed to \( x' = Mx \), then a contravariant vector \( v \) must be similarly transformed via \( v' = Mv \). This important requirement is what distinguishes a contravariant vector from any other triple of physically meaningful quantities. For example, if \( v \) consists of the \( x, y, \) and \( z \)-components of velocity, then \( v \) is a contravariant vector: if the coordinates of space are stretched, rotated, or twisted, then the components of the velocity transform in the same way. On the other hand, for instance, a triple consisting of the length, width, and height of a rectangular box could make up the three components of an abstract vector, but this vector would not be contravariant, since rotating the box does not change the box’s length, width, and height. Examples of contravariant vectors include displacement, velocity and acceleration.

By contrast, a **covariant vector** has components that change oppositely to the coordinates or, equivalently, transform like the reference axes. For instance, the components of the gradient vector of a function

\[ \nabla f = \frac{\partial f}{\partial x_1} \hat{x}_1 + \frac{\partial f}{\partial x_2} \hat{x}_2 + \frac{\partial f}{\partial x_3} \hat{x}_3 \]

transform like the reference axes themselves. When only rotations of the axes are considered, the components of contravariant and covariant vectors behave in the same way. It is only when other transformations are allowed that the difference becomes apparent.

**Definition**

The general formulation of covariance and contravariance refers to how the components of a coordinate vector transform under a change of basis (passive transformation). Thus let \( V \) be a vector space of dimension \( n \) over the field of scalars \( S \), and let each of \( f = (X_1, \ldots, X_n) \) and \( f' = (Y_1, \ldots, Y_n) \) be a basis of \( V \). Also, let the change of basis from \( f \) to \( f' \) be given by

\[
\begin{align*}
\sum_i a_i^1 X_i, \ldots, \sum_i a_i^n X_i = f A \end{align*}\] 

(1)

\[
\begin{align*}
\sum_i a_i^1 X_i, \ldots, \sum_i a_i^n X_i = f A \end{align*}\] 

(1)
for some invertible $n \times n$ matrix $A$ with entries $a^i_j$. Here, each vector $Y_j$ of the $f$ basis is a linear combination of the vectors $X_i$ of the $f$ basis, so that

$$Y_j = \sum_i a^i_j X_i.$$  

**Contravariant transformation**

A vector $v$ in $V$ is expressed uniquely as a linear combination of the elements of the $f$ basis as

$$v = \sum_i v^i[f] X_i,$$

where $v^i[f]$ are scalars in $S$ known as the **components** of $v$ in the $f$ basis. Denote the column vector of components of $v$ by $v[f]$:

$$v[f] = \begin{bmatrix} v^1[f] \\ v^2[f] \\ \vdots \\ v^n[f] \end{bmatrix}$$

so that (2) can be rewritten as a matrix product

$$v = f v[f].$$

The vector $v$ may also be expressed in terms of the $f'$ basis, so that

$$v = f' v[f'].$$

However, since the vector $v$ itself is invariant under the choice of basis,

$$f v[f] = v = f' v[f'].$$

The invariance of $v$ combined with the relationship (1) between $f$ and $f'$ implies that

$$f v[f] = f A v[fA],$$

giving the transformation rule

$$v[fA] = A^{-1} v[f].$$

In terms of components,

$$v^i[fA] = \sum_j \tilde{a}^i_j v^j[f]$$

where the coefficients $\tilde{a}^i_j$ are the entries of the inverse matrix of $A$.

Because the components of the vector $v$ transform with the inverse of the matrix $A$, these components are said to **transform contravariantly** under a change of basis.

The way $A$ relates the two pairs is depicted in the following informal diagram using an arrow. The reversal of the arrow indicates a contravariant change:

$$f \rightarrow f'$$

$$v[f] \leftarrow v[f'].$$
Covariance and contravariance of vectors

Covariant transformation

A linear functional $\alpha$ on $V$ is expressed uniquely in terms of its components (scalars in $S$) in the $f$ basis as

$$\alpha(X_i) = \alpha_i[f], \quad i = 1, 2, \ldots, n.$$  

These components are the action of $\alpha$ on the basis vectors $X_i$ of the $f$ basis.

Under the change of basis from $f$ to $f'$ (1), the components transform so that

$$\alpha_i[fA] = \alpha(Y_i)$$

$$= \alpha \left( \sum_j a_i^j X_j \right)$$

$$= \sum_j a_i^j \alpha(X_j)$$

$$= \sum_j a_i^j \alpha_j[f].$$

Denote the row vector of components of $\alpha$ by $\alpha[f]$:

$$\alpha[f] = [\alpha_1[f], \alpha_2[f], \ldots, \alpha_n[f]]$$

so that (3) can be rewritten as the matrix product

$$\alpha[fA] = \alpha[f]A.$$

Because the components of the linear functional $\alpha$ transform with the matrix $A$, these components are said to transform covariantly under a change of basis.

The way $A$ relates the two pairs is depicted in the following informal diagram using an arrow. A covariant relationship is indicated since the arrows travel in the same direction:

$$f \longrightarrow f'$$

$$\alpha[f] \longrightarrow \alpha[f'].$$

Had a column vector representation been used instead, the transformation law would be the transpose

$$\alpha^T[fA] = A^T \alpha^T[f].$$

Coordinates

The choice of basis $f$ on the vector space $V$ defines uniquely a set of coordinate functions on $V$, by means of

$$x^i[f](v) = v^i[f].$$

The coordinates on $V$ are therefore contravariant in the sense that

$$x^i[fA] = \sum_{k=1}^n \tilde{a}_{ik} x^k[f].$$

Conversely, a system of $n$ quantities $v^i$ that transform like the coordinates $x^i$ on $V$ defines a contravariant vector. A system of $n$ quantities that transform oppositely to the coordinates is then a covariant vector.

This formulation of contravariance and covariance is often more natural in applications in which there is a coordinate space (a manifold) on which vectors live as tangent vectors or cotangent vectors. Given a local coordinate system $x^i$ on the manifold, the reference axes for the coordinate system are the vector fields

$$X_1 = \frac{\partial}{\partial x^1}, \ldots, X_n = \frac{\partial}{\partial x^n}.$$  

This gives rise to the frame $f = (X_1, \ldots, X_n)$ at every point of the coordinate patch.

If $y^i$ is a different coordinate system and

$$Y_1 = \frac{\partial}{\partial y^1}, \ldots, Y_n = \frac{\partial}{\partial y^n},$$
then the frame $f'$ is related to the frame $f$ by the inverse of the Jacobian matrix of the coordinate transition:

$$f' = fJ^{-1}, \quad J = \left( \frac{\partial y^i}{\partial x^j} \right)_{i,j=1}^n.$$ 

Or, in indices,

$$\frac{\partial}{\partial y^i} = \sum_{j=1}^n \frac{\partial x^j}{\partial y^i} \frac{\partial}{\partial x^j}.$$ 

A tangent vector is by definition a vector that is a linear combination of the coordinate partials $\partial/\partial x^i$. Thus a tangent vector is defined by

$$v = \sum_{i=1}^n v_i[f] X_i = f^i[v].$$ 

Such a vector is contravariant with respect to change of frame. Under changes in the coordinate system, one has

$$v[f'] = v[fJ^{-1}] = J v[f].$$ 

Therefore the components of a tangent vector transform via

$$v^i[f'] = \sum_{j=1}^n \frac{\partial y^j}{\partial x^i} v^j[f].$$ 

Accordingly, a system of $n$ quantities $v^i$ depending on the coordinates that transform in this way on passing from one coordinate system to another is called a contravariant vector.

Covariant and contravariant components of a vector with a metric

In a vector space $V$ over a field $K$ with a bilinear form $g : V \times V \to K$ (which may be referred to as the metric tensor), there is little distinction between covariant and contravariant vectors, because the bilinear form allows covectors to be identified with vectors. That is, a vector $v$ uniquely determines a covector $\alpha$ via

$$\alpha(w) = g(v, w)$$

for all vectors $w$. Conversely, each covector $\alpha$ determines a unique vector $v$ by this equation. Because of this identification of vectors with covectors, one may speak of the **covariant components** or **contravariant components** of a vector, that is, they are just representations of the same vector using reciprocal bases.

Given a basis $f = (X_1, \ldots, X_n)$ of $V$, there is a unique reciprocal basis $f^\# = (Y^1, \ldots, Y^n)$ of $V$ determined by requiring that

$$Y^i(X_j) = \delta^i_j,$$

de the Kronecker delta. In terms of these bases, any vector $v$ can be written in two ways:

$$v = \sum_i v_i[f] X_i = f^i[v]$$

$$= \sum_i v_i[f] Y^i = f^i[v^\#].$$

The components $v^i[f]$ are the **contravariant components** of the vector $v$ in the basis $f$, and the components $v_i[f]$ are the **covariant components** of $v$ in the basis $f$. The terminology is justified because under a change of basis,

$$v[fA] = A^{-1} v[f], \quad v^i[fA] = A^T v^i[f].$$
Euclidean plane

In the Euclidean plane, the dot product allows for vectors to be identified with covectors. If \( \mathbf{e}_1, \mathbf{e}_2 \) is a basis, then the dual basis \( \mathbf{e}^1, \mathbf{e}^2 \) satisfies

\[
\mathbf{e}^1 \cdot \mathbf{e}_1 = 1, \quad \mathbf{e}^1 \cdot \mathbf{e}_2 = 0 \\
\mathbf{e}^2 \cdot \mathbf{e}_1 = 0, \quad \mathbf{e}^2 \cdot \mathbf{e}_2 = 1.
\]

Thus, \( \mathbf{e}^1 \) and \( \mathbf{e}^2 \) are perpendicular to each other, as are \( \mathbf{e}^2 \) and \( \mathbf{e}_1 \), and the lengths of \( \mathbf{e}^1 \) and \( \mathbf{e}^2 \) normalized against \( \mathbf{e}_1 \) and \( \mathbf{e}_2 \), respectively.

Example

For example,\(^3\) suppose that we are given a basis \( \mathbf{e}_1, \mathbf{e}_2 \) consisting of a pair of vectors making a 45° angle with one another, such that \( \mathbf{e}_1 \) has length 2 and \( \mathbf{e}_2 \) has length 1. Then the dual basis vectors are given as follows:

- \( \mathbf{e}^2 \) is the result of rotating \( \mathbf{e}_1 \) through an angle of 90° (where the sense is measured by assuming the pair \( \mathbf{e}_1, \mathbf{e}_2 \) to be positively oriented), and then rescaling so that \( \mathbf{e}^2 \cdot \mathbf{e}_2 = 1 \) holds.
- \( \mathbf{e}^1 \) is the result of rotating \( \mathbf{e}_2 \) through an angle of 90°, and then rescaling so that \( \mathbf{e}^1 \cdot \mathbf{e}_1 = 1 \) holds.

Applying these rules, we find

\[
\mathbf{e}^1 = \frac{1}{2} \mathbf{e}_1 - \frac{1}{\sqrt{2}} \mathbf{e}_2
\]

and

\[
\mathbf{e}^2 = -\frac{1}{\sqrt{2}} \mathbf{e}_1 + 2 \mathbf{e}_2.
\]

Thus the change of basis matrix in going from the original basis to the reciprocal basis is

\[
R = \begin{bmatrix}
\frac{1}{2} & -1/\sqrt{2} \\
-1/\sqrt{2} & 2
\end{bmatrix},
\]

since
For instance, the vector
\[ v = \frac{3}{2}e_1 + 2e_2 \]
is a vector with contravariant components
\[ v^1 = \frac{3}{2}, \quad v^2 = 2. \]
The covariant components are obtained by equating the two expressions for the vector \( v \):
\[ v = v_1e^1 + v_2e^2 = v^1e_1 + v^2e_2 \]
so
\[
\begin{bmatrix}
  v_1 \\
  v_2
\end{bmatrix}
= R^{-1}
\begin{bmatrix}
  v^1 \\
  v^2
\end{bmatrix}
= \begin{bmatrix}
  4 & \sqrt{2} \\
  \sqrt{2} & 1
\end{bmatrix}
\begin{bmatrix}
  v^1 \\
  v^2
\end{bmatrix}
= \begin{bmatrix}
  6 + 2\sqrt{2} \\
  2 + 3\sqrt{2}
\end{bmatrix}.
\]

Three-dimensional Euclidean space
In the three-dimensional Euclidean space, one can also determine explicitly the dual basis to a given set of basis vectors \( e_1, e_2, e_3 \) of \( \mathbb{E}^3 \) that are not necessarily assumed to be orthogonal nor of unit norm. The contravariant (dual) basis vectors are:
\[
e^1 = \frac{e_2 \times e_3}{e_1 \cdot (e_2 \times e_3)}, \quad e^2 = \frac{e_3 \times e_1}{e_2 \cdot (e_3 \times e_1)}, \quad e^3 = \frac{e_1 \times e_2}{e_3 \cdot (e_1 \times e_2)}.
\]
Even when the \( e_i \) and \( e^i \) are not orthonormal, they are still mutually dual:
\[ e^i \cdot e^j = \delta^i_j. \]
Then the contravariant coordinates of any vector \( v \) can be obtained by the dot product of \( v \) with the contravariant basis vectors:
\[ q^1 = v \cdot e^1; \quad q^2 = v \cdot e^2; \quad q^3 = v \cdot e^3. \]
Likewise, the covariant components of \( v \) can be obtained from the dot product of \( v \) with covariant basis vectors, viz.
\[ q_1 = v \cdot e_1; \quad q_2 = v \cdot e_2; \quad q_3 = v \cdot e_3. \]
Then \( v \) can be expressed in two (reciprocal) ways, viz.
\[ v = q_3 e^3 = q_1 e^1 + q_2 e^2 + q_3 e^3 \]
or
\[ v = q^i e_i = q^1 e_1 + q^2 e_2 + q^3 e_3. \]
Combining the above relations, we have
\[ v = (v \cdot e_i) e^i = (v \cdot e^i) e_i \]
and we can convert from covariant to contravariant basis with
\[ q_i = v \cdot e_i = (q^j e_j) \cdot e_i = (e_j \cdot e_i) q^j \]
and
\[ q^i = v \cdot e^i = (q_j e^j) \cdot e^i = (e^j \cdot e^i) q_j. \]
The indices of covariant coordinates, vectors, and tensors are subscripts. If the contravariant basis vectors are orthonormal then they are equivalent to the covariant basis vectors, so there is no need to distinguish between the covariant and contravariant coordinates.
General Euclidean spaces

More generally, in an $n$-dimensional Euclidean space $V$, if a basis is
\[ \mathbf{e}_1, \ldots, \mathbf{e}_n, \]
the reciprocal basis is given by
\[ \mathbf{e}^i = e^{ij} \mathbf{e}_j \]
where the coefficients $e^{ij}$ are the entries of the inverse matrix of
\[ e_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j. \]
Indeed, we then have
\[ \mathbf{e}^i \cdot \mathbf{e}_k = e^{ij} \mathbf{e}_j \cdot \mathbf{e}_k = e^{ij} e_{jk} = \delta^i_k. \]
The covariant and contravariant components of any vector
\[ \mathbf{v} = q_i \mathbf{e}^i = q^i \mathbf{e}_i \]
are related as above by
\[ q_i = \mathbf{v} \cdot \mathbf{e}_i = (q^j \mathbf{e}_j) \cdot \mathbf{e}_i = q^j e_{ji} \]
and
\[ q^i = \mathbf{v} \cdot \mathbf{e}^i = (q_j \mathbf{e}^j) \cdot \mathbf{e}^i = q_j e^{ji}. \]

Informal usage

In the field of physics, the adjective covariant is often used informally as a synonym for invariant. For example, the Schrödinger equation does not keep its written form under the coordinate transformations of special relativity. Thus, a physicist might say that the Schrödinger equation is not covariant. In contrast, the Klein–Gordon equation and the Dirac equation do keep their written form under these coordinate transformations. Thus, a physicist might say that these equations are covariant.

Despite this usage of "covariant", it is more accurate to say that the Klein–Gordon and Dirac equations are invariant, and that the Schrödinger equation is not invariant. Additionally, to remove ambiguity, the transformation by which the invariance is evaluated should be indicated.

Because the components of vectors are contravariant and those of covectors are covariant, the vectors themselves are often referred to as being contravariant and the covectors as covariant. This usage may be misleading, however, since vectors push forward – are covariant under diffeomorphism – and covectors pull back – are contravariant under diffeomorphism. See Einstein notation for details.

Use in tensor analysis

The distinction between covariance and contravariance is particularly important for computations with tensors, which often have mixed variance. This means that they have both covariant and contravariant components, or both vector and dual vector components. The valence of a tensor is the number of variant and covariant terms, and in Einstein notation, covariant components have lower indices, while contravariant components have upper indices. The duality between covariance and contravariance intervenes whenever a vector or tensor quantity is represented by its components, although modern differential geometry uses more sophisticated index-free methods to represent tensors.

In tensor analysis, a covariant vector varies more or less reciprocally to a corresponding contravariant vector. Expressions for lengths, areas and volumes of objects in the vector space can then be given in terms of tensors with covariant and contravariant indices. Under simple expansions and contractions of the coordinates, the reciprocity is exact; under affine transformations the components of a vector intermingle on going between covariant and
Covariance and contravariance of vectors

On a manifold, a tensor field will typically have multiple indices, of two sorts. By a widely followed convention, covariant indices are written as lower indices, whereas contravariant indices are upper indices. When the manifold is equipped with a metric, covariant and contravariant indices become very closely related to one another. Contravariant indices can be turned into covariant indices by contracting with the metric tensor. The reverse is possible by contracting with the (matrix) inverse of the metric tensor. Note that in general, no such relation exists in spaces not endowed with a metric tensor. Furthermore, from a more abstract standpoint, a tensor is simply "there" and its components of either kind are only calculational artifacts whose values depend on the chosen coordinates.

The explanation in geometric terms is that a general tensor will have contravariant indices as well as covariant indices, because it has parts that live in the tangent bundle as well as the cotangent bundle.

A contravariant vector is one which transforms like $\frac{dx^\mu}{d\tau}$, where $x^\mu$ are the coordinates of a particle at its proper time $\tau$. A covariant vector is one which transforms like $\frac{\partial \phi}{\partial x^\mu}$, where $\phi$ is a scalar field.

Algebra and geometry

In category theory, there are covariant functors and contravariant functors. The dual space of a vector space is a standard example of a contravariant functor. Some constructions of multilinear algebra are of 'mixed' variance, which prevents them from being functors.

In geometry, the same map in/map out distinction is helpful in assessing the variance of constructions. A tangent vector to a smooth manifold $M$ is, to begin with, a curve mapping smoothly into $M$ and passing through a given point $P$. It is therefore covariant, with respect to smooth mappings of $M$. A contravariant vector, or 1-form, is in the same way constructed from a smooth mapping from $M$ to the real line, near $P$. It is in the cotangent bundle, built up from the dual spaces of the tangent spaces. Its components with respect to a local basis of one-forms $dx^i$ will be covariant; but one-forms and differential forms in general are contravariant, in the sense that they pull back under smooth mappings. This is crucial to how they are applied; for example a differential form can be restricted to any submanifold, while this does not make the same sense for a field of tangent vectors.

Covariant and contravariant components transform in different ways under coordinate transformations. By considering a coordinate transformation on a manifold as a map from the manifold to itself, the transformation of covariant indices of a tensor are given by a pullback, and the transformation properties of the contravariant indices is given by a pushforward.

Notes

[2] A basis $f$ may here profitably be viewed as a linear isomorphism from $\mathbb{R}^n$ to $V$. Regarding $f$ as a row vector whose entries are the elements of the basis, the associated linear isomorphism is then $x \mapsto fx$.

References


**External links**

• Invariance, Contravariance, and Covariance (http://www.mathpages.com/home/kmath398/kmath398.htm)

Image Sources, Licenses and Contributors


License

Creative Commons Attribution-Share Alike 3.0 Unported
creativecommons.org/licenses/by-sa/3.0/