

Contravariant and Covariant as Transforms

There is a lot more “behind” the concepts of contravariant and covariant tensors (of any rank) than the fact that their basis vectors are mutually orthogonal to each other. The terms contravariant and covariant apply based on *how they are derived* :

- Contravariant tensors (of any rank) are derived from the derivative of the coordinate axes
- Covariant tensors (of any rank) are derived from the gradient

which are also related to *how the tensor's components transform* between different coordinate systems.

Transforms are used to change from one Lorentz frame to another (boosts, rotations, etc.), from one local coordinate chart to another, or between bases. This can include bases from different points in curved space, from a holonomic basis to a non-holonomic basis, or equivalently, from a tangent vector basis to a tetrad basis (described in GR3x). Keep in mind that the tensor itself does not change under these transformations, only the way we represent it.

Contravariant Transforms

For a contravariant vector to be coordinate system invariant, the components of the vector must transform *oppositely* (“contra” as in “against”, like “contrary”) to the change of the basis vectors (the axes) to *cancel the change*. For example, **if the axes were rotated in one direction, the component representation of the vector would rotate in the opposite way**.

Let (x^1, x^2, x^3) be the coordinates in a surface in one basis, and (y^1, y^2, y^3) be the coordinates in the same surface in a different (transformed) basis, where each y^i is a function of (x^1, x^2, x^3) .

Then the derivative of each *coordinate* y^i can be expressed in terms of its partials with respect to the x^i 's by using the chain rule :

$$\begin{aligned} dy^1 &= (\partial y^1 / \partial x^1) dx^1 + (\partial y^1 / \partial x^2) dx^2 + (\partial y^1 / \partial x^3) dx^3 \\ dy^2 &= (\partial y^2 / \partial x^1) dx^1 + (\partial y^2 / \partial x^2) dx^2 + (\partial y^2 / \partial x^3) dx^3 \\ dy^3 &= (\partial y^3 / \partial x^1) dx^1 + (\partial y^3 / \partial x^2) dx^2 + (\partial y^3 / \partial x^3) dx^3 \end{aligned}$$

Or in index notation,

$$dy^i = (\partial y^i / \partial x^j) dx^j$$

where each $(\partial y^i / \partial x^j)$ is *calculated from the relationship between the bases*, and becomes a “constant”.

For example, consider a 2-D Euclidean space with Cartesian coordinates. → The first basis $[x^1, x^2]$ is the usual $[x, y]$. The new basis has been rotated about the origin by θ degrees, so

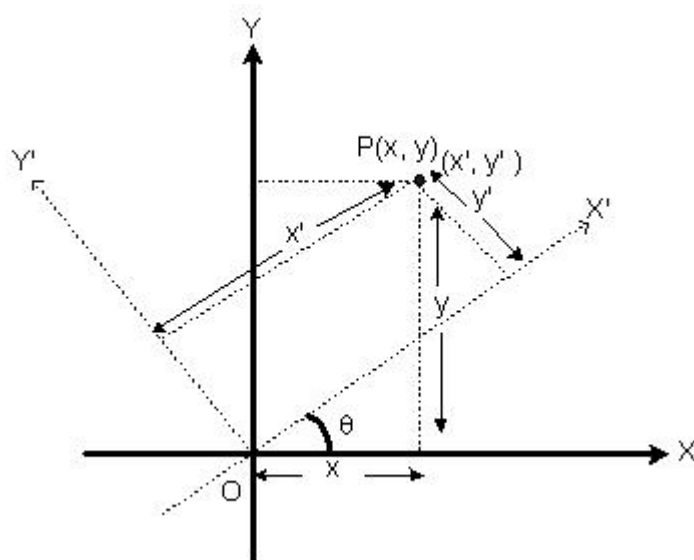
$$\begin{aligned} y^1 &= x^1 \cos\theta + x^2 \sin\theta \\ y^2 &= -x^1 \sin\theta + x^2 \cos\theta \end{aligned}$$

Then

$$\begin{aligned} dy^1 &= (\partial y^1 / \partial x^1) dx^1 + (\partial y^1 / \partial x^2) dx^2 \\ &= (\cos\theta) dx^1 + (\sin\theta) dx^2 \\ dy^2 &= (\partial y^2 / \partial x^1) dx^1 + (\partial y^2 / \partial x^2) dx^2 \\ &= (-\sin\theta) dx^1 + (\cos\theta) dx^2 \end{aligned}$$

When the *components* of a contravariant vector V^i in the first basis are transformed into *components* W^i in the new basis (remember, the vector itself is not changing), they transform like the coordinate differentials :

$$W^i = (\partial y^i / \partial x^j) V^j$$



In the example above, the point P is given by $V=(0.75, 1.25)$ in $[x^1, x^2]$, and $\theta=35^\circ$. So

$$W^1 = (\partial y^1 / \partial x^1) V^1 + (\partial y^1 / \partial x^2) V^2 = \cos(35) * 0.75 + \sin(35) * 1.25 = 1.33$$

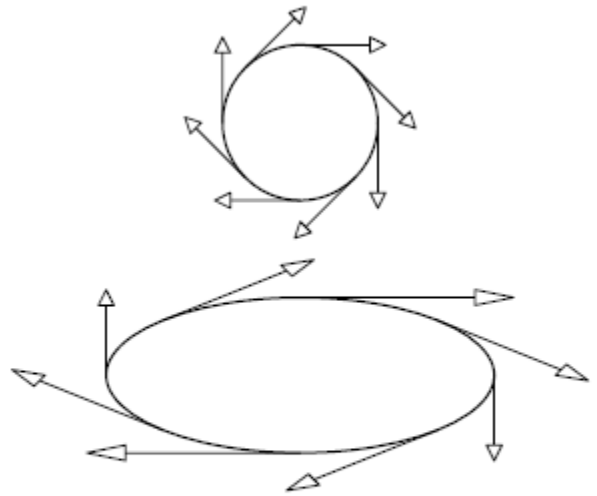
$$W^2 = (\partial y^2 / \partial x^1) V^1 + (\partial y^2 / \partial x^2) V^2 = -\sin(35) * 0.75 + \cos(35) * 1.25 = 0.59$$

which matches (x', y') in the figure.

Looking at the transformation again in Cartesian coordinates,

$$W^i = (\partial y^i / \partial x^j) V^j \\ = (\partial y^i / \partial x^1) V^1 + (\partial y^i / \partial x^2) V^2 + (\partial y^i / \partial x^3) V^3$$

the basis vectors $(\partial / \partial x^j)$ are the same as in the definition of a *tangent basis vector* (from GR1b). It follows that **contravariant vectors are tangent vectors**. When the circle to the left and its tangent vectors undergo the same contraction transformation, the new vectors are still tangent to the new surface. Note that for a spacetime 4-vector, the length $(ct)^2 - x^2 - y^2 - z^2$ would remain invariant.



Covariant Transforms

For a covariant vector to be coordinate system invariant, the components of the vector must transform in the *same way* (“co-” as in “together”, like “comoving”) as the change of the basis vectors (the axes) to *maintain the same meaning* (like “normal”).

As before, let (x^1, x^2, x^3) be the coordinates in a surface in one basis, and (y^1, y^2, y^3) be the coordinates in the same surface in a different (transformed) basis, where each y^i is a function of (x^1, x^2, x^3) .

To find the partial derivatives of a *scalar function* $f(x^1, x^2, x^3)$ in the new basis, we use the chain rule :

$$\partial f / \partial y^1 = (\partial f / \partial x^1) (\partial x^1 / \partial y^1) + (\partial f / \partial x^2) (\partial x^2 / \partial y^1) + (\partial f / \partial x^3) (\partial x^3 / \partial y^1)$$

$$\partial f / \partial y^2 = (\partial f / \partial x^1) (\partial x^1 / \partial y^2) + (\partial f / \partial x^2) (\partial x^2 / \partial y^2) + (\partial f / \partial x^3) (\partial x^3 / \partial y^2)$$

$$\partial f / \partial y^3 = (\partial f / \partial x^1) (\partial x^1 / \partial y^3) + (\partial f / \partial x^2) (\partial x^2 / \partial y^3) + (\partial f / \partial x^3) (\partial x^3 / \partial y^3)$$

Or in index notation,

$$\partial f / \partial y^i = (\partial f / \partial x^j) (\partial x^j / \partial y^i)$$

where each $(\partial x^j / \partial y^i)$ is *calculated from the relationship between the bases*, and becomes a “constant”.

For example, if a scalar field in Cartesian coordinates $(x^1, x^2, x^3) = (x, y, z)$ is : $f = x + 5y^2z$ then its partial derivatives are

$$\partial f / \partial x = 1 \qquad \partial f / \partial y = 10yz \qquad \partial f / \partial z = 5y^2$$

To find the partial derivatives in cylindrical coordinates $(y^1, y^2, y^3) = (r, \phi, z)$ where

$$x = r \cos \phi \qquad y = r \sin \phi \qquad z = z$$

We have

$$\partial f / \partial r = (\partial f / \partial x^1) (\partial x^1 / \partial y^1) + (\partial f / \partial x^2) (\partial x^2 / \partial y^1) + (\partial f / \partial x^3) (\partial x^3 / \partial y^1)$$

$$\partial f / \partial r = (\partial f / \partial x) (\partial x / \partial r) + (\partial f / \partial y) (\partial y / \partial r) + (\partial f / \partial z) (\partial z / \partial r)$$

$$\partial f / \partial r = (1) (\cos \phi) + (10yz) (\sin \phi) + (5y^2) (0)$$

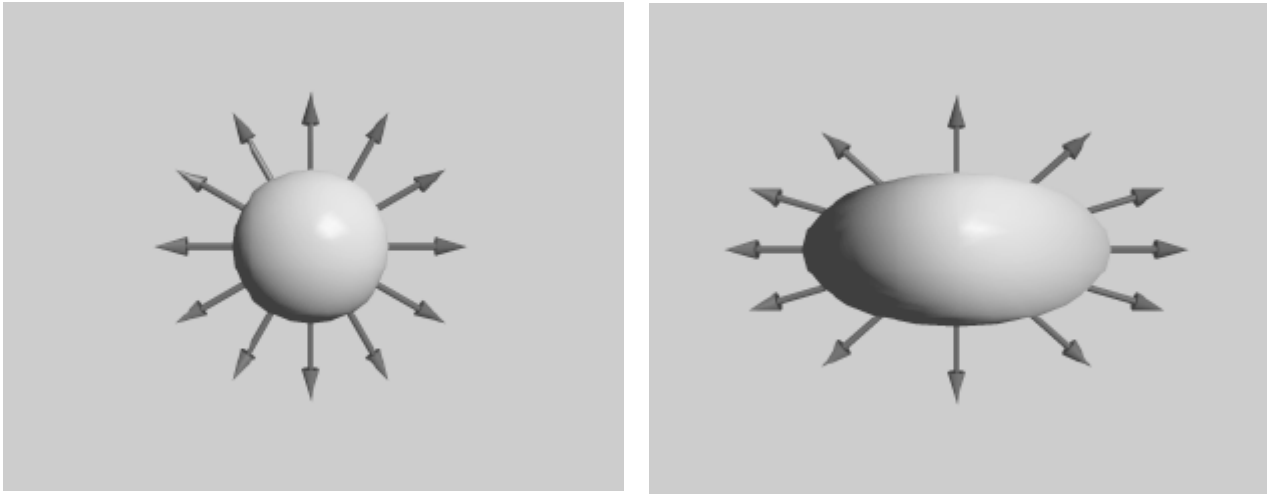
$$\partial f / \partial r = \cos \phi + 10yz \sin \phi$$

When the *components* of a covariant vector V_i in the first basis are transformed into *components* W_i in the new basis, they transform like the derivative of a function :

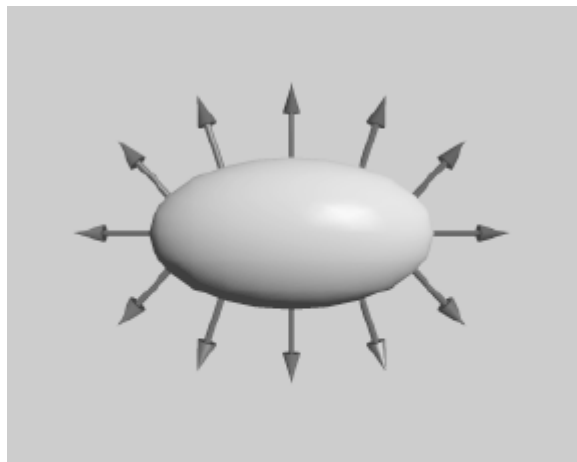
$$W_i = (\partial x^j / \partial y^i) V_j$$

$$= (\partial x^1 / \partial y^i) V_1 + (\partial x^2 / \partial y^i) V_2 + (\partial x^3 / \partial y^i) V_3$$

Consider the gradient of a scalar function $f = \nabla f = (\partial f / \partial x^1, \partial f / \partial x^2, \partial f / \partial x^3) = (\partial f / \partial x^j)$. Thus the components of a covariant vector are components of a normal vector. When the circle below and its normal vectors undergo the same transformation, the new vectors are not normal to the new surface :



With the proper covariant transformation, the normal vectors stay normal :



The Jacobian

When transforming from coordinates x^i to coordinates y^i and remembering that each y^i is a function of (x^1, x^2, x^3) , we have :

$$\partial y^1 / \partial x = (\partial y^1 / \partial x^1) + (\partial y^1 / \partial x^2) + (\partial y^1 / \partial x^3)$$

$$\partial y^2 / \partial x = (\partial y^2 / \partial x^1) + (\partial y^2 / \partial x^2) + (\partial y^2 / \partial x^3)$$

$$\partial y^3 / \partial x = (\partial y^3 / \partial x^1) + (\partial y^3 / \partial x^2) + (\partial y^3 / \partial x^3)$$

Which can be rearranged and written in matrix form as

$$\begin{array}{l} | \partial y^1 | \quad | \partial y^1 / \partial x^1 \quad \partial y^1 / \partial x^2 \quad \partial y^1 / \partial x^3 | \quad | \partial x^1 | \\ | \partial y^2 | = | \partial y^2 / \partial x^1 \quad \partial y^2 / \partial x^2 \quad \partial y^2 / \partial x^3 | \quad | \partial x^2 | \\ | \partial y^3 | \quad | \partial y^3 / \partial x^1 \quad \partial y^3 / \partial x^2 \quad \partial y^3 / \partial x^3 | \quad | \partial x^3 | \end{array}$$

The matrix is a tensor called the **Jacobian** (J^i_j), which in general can convert between coordinate systems, do Lorentz transformations, etc. Then contravariant vector components transform as

$$W^i = J^i_j V^j$$

While covariant vector components transform as

$$W_i = J_j^{i-1} V_j$$

Where the definition of the Jacobian is $J_j^i = \partial y^i / \partial x^j$ and $J_j^{i-1} = \partial x^i / \partial y^j$ is the inverse transform :

$$(J_j^i)^{-1} = \begin{vmatrix} \partial x^1 / \partial y^1 & \partial x^1 / \partial y^2 & \partial x^1 / \partial y^3 \\ \partial x^2 / \partial y^1 & \partial x^2 / \partial y^2 & \partial x^2 / \partial y^3 \\ \partial x^3 / \partial y^1 & \partial x^3 / \partial y^2 & \partial x^3 / \partial y^3 \end{vmatrix}$$

Note that $(J_j^i)^{-1}$ is **not** the matrix inverse of J_{ij} . Some authors refer to the determinant of the matrix as “the Jacobian” [in which case $\det(J_j^{i-1})$ does = $1/\det(J_{ij})$], so be careful!

For example, the transformation *from* polar coordinates $[x^1, x^2] = [r, \theta]$ to Cartesian coordinates $[y^1, y^2] = [x, y]$ is given by the functions :

$$\begin{aligned} x &= r \cos \theta \\ y &= r \sin \theta \end{aligned}$$

Then the Jacobian for this transform is (ϕ 's should be θ) :

$$J(r, \phi) = \begin{bmatrix} \frac{dx}{dr} & \frac{dx}{d\phi} \\ \frac{dy}{dr} & \frac{dy}{d\phi} \end{bmatrix} = \begin{bmatrix} \frac{d(r \cos \phi)}{dr} & \frac{d(r \cos \phi)}{d\phi} \\ \frac{d(r \sin \phi)}{dr} & \frac{d(r \sin \phi)}{d\phi} \end{bmatrix} = \begin{bmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{bmatrix}$$

And the inverse is :

$$\frac{\partial P^b}{\partial C^a} = \begin{pmatrix} \partial r / \partial x & \partial r / \partial y \\ \partial \phi / \partial x & \partial \phi / \partial y \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\frac{\sin \phi}{r} & \frac{\cos \phi}{r} \end{pmatrix}$$

While the ordinary derivative is the first derivative of a *scalar function* of one variable [$df(x)/dx$], the gradient is the first derivative of a *scalar function* of several variables [$\nabla f(x,y,z)$], and the Jacobian is the first derivative of a *vector function* of several variables [for example, $\mathbf{f}(r,\theta) = (r \cos \theta, r \sin \theta)$].

The Jacobian can also be thought of as describing the amount of "stretching" that a transformation imposes. For example, if $(x_2, y_2) = f(x_1, y_1)$ is used to transform an image at each pixel (x_1, y_1) , the Jacobian of f describes how much the image is stretched in the x , y , and xy (diagonal) directions.

One-Forms

One-forms go by many alternate names : “1-form”, “form”, “linear form”, “outer form”, “differential form”, “cotangent vector”, “covector”, “dual vector”, and a name we already know : **covariant vector!**

The figure to the right shows a series of contour lines for air pressure over a surface: lines along which the pressure is the same value. The values of air pressure are a *scalar field*, and a **one-form is the geometrical object represented by the contour lines of the scalar field**. There are plenty of ways in which contours can be useful. Which way will a boulder roll at a certain point on a complicated hilly terrain? The height of the earth is a scalar field; draw contour lines,

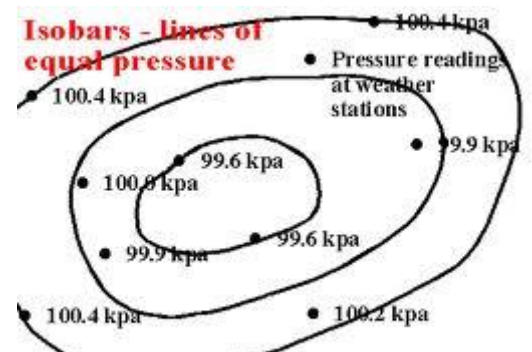
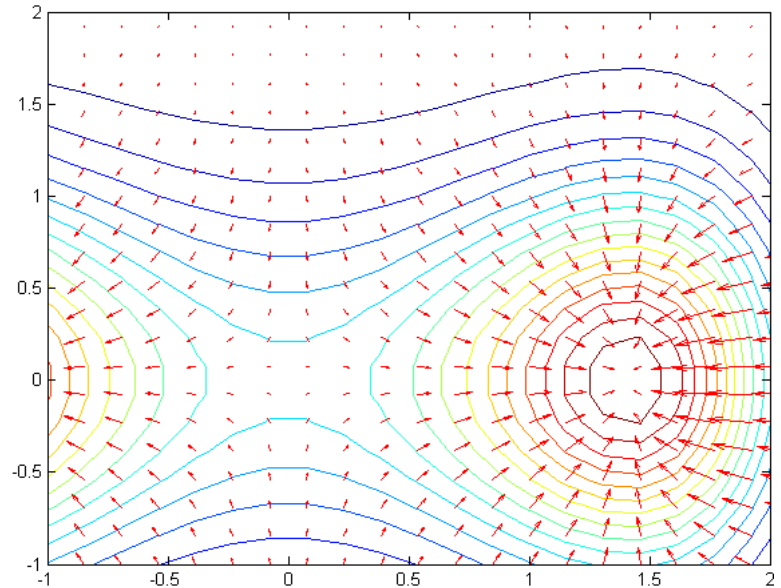


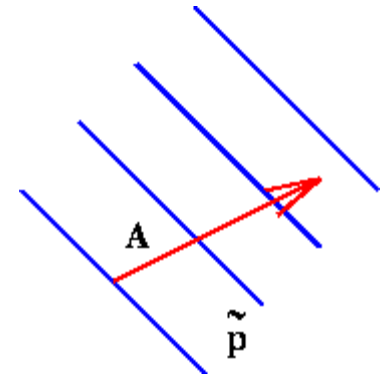
Fig 11

then draw an arrow **perpendicular** to the contour line passing through the boulder in the direction of decreasing height, and that's the way the boulder will fall (we're assuming Euclidean space for the moment, where "perpendicular" actually means something). You can even give the arrow a length proportional to the steepness of the ground; that is, the rate of change of height as you move around. → That is the **gradient**, which is a vector representing the rate at which a scalar field is changing at each point in a given direction. *The closer together the contour lines, the longer the vector will be.* So the gradient is related to the contour lines, and thus to one-forms.



A more mathematical definition is that a one-form is anything that we can combine with a (non-gradient) contravariant vector to yield a scalar number at any point in the surface, such that the number we get is *linear* with respect to the vector (doubling the length of the vector doubles the value of the scalar, etc.).

There's no reason why we can't have a one-form that satisfies this definition, but doesn't happen to be the gradient of *any* scalar field. We could draw a set of lines at any point that yielded the same scalar result, → but the lines at this point wouldn't be perpendicular to the vector, nor would they connect with the lines at a point next to it, like actual contour lines would. In this case, it makes more sense to think of the lines being drawn in the tangent plane for each point, since they don't really belong in the surface.



The magnitude of the one-form is related to the spacing between the lines. A one-form with twice the magnitude would have twice as many lines in the same space.

In 2-D pictures, the one-form is represented as a set of lines. If you have two points in 3-D Euclidean space, then after drawing a vector between them, the one-form would be a series of planes along the vector. In 4-D spacetime, the one-form is a series of three-dimensional hyperplanes (really 3-D spaces, but "hyperplane" is the standard generic term for any straight/flat surface which has one less dimension than the space it's in).

Briefly, a **vector space** is any collection of vectors which satisfy certain mathematical rules, such as association, commutation, distribution, etc. (ordinary 2-D and 3-D Cartesian spaces are vector spaces). In a vector space, any *linear* function which when combined with a vector results in a scalar is called a **covector** (short for "**cotangent vector**", just like "tangent vector"), and covectors also form a vector space. When the vector space of vectors and the vector space of covectors have the same number of dimensions, they can be put in a linear one-to-one correspondence with each other. This means that for each vector, there is a covector that can in some sense represent or replace it (and vice-versa). The distinction between vectors and covectors is needed when spacetime or the coordinates are curved. **Technically, a one-form is a covector field** (just like a "scalar field" or "vector field"), which associates a covector with each point on the surface (spacetime). **Many authors make no distinction between a one-form and the covector associated with it, and the term "covector" is sometimes never used.**

One-forms are denoted by Latin letters with tildes ($\tilde{\mathbf{u}}$) or overbars ($\bar{\mathbf{u}}$, which can be confused with vectors), by letters with lines under them as the complement to regular vectors ($\underline{\mathbf{u}}$, which is also used for tensors or 4-vectors), and also by bold Greek letters ($\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\sigma}$), which will be used here when possible.

The following figure gives several examples of vectors and their one-forms in a spacetime diagram :

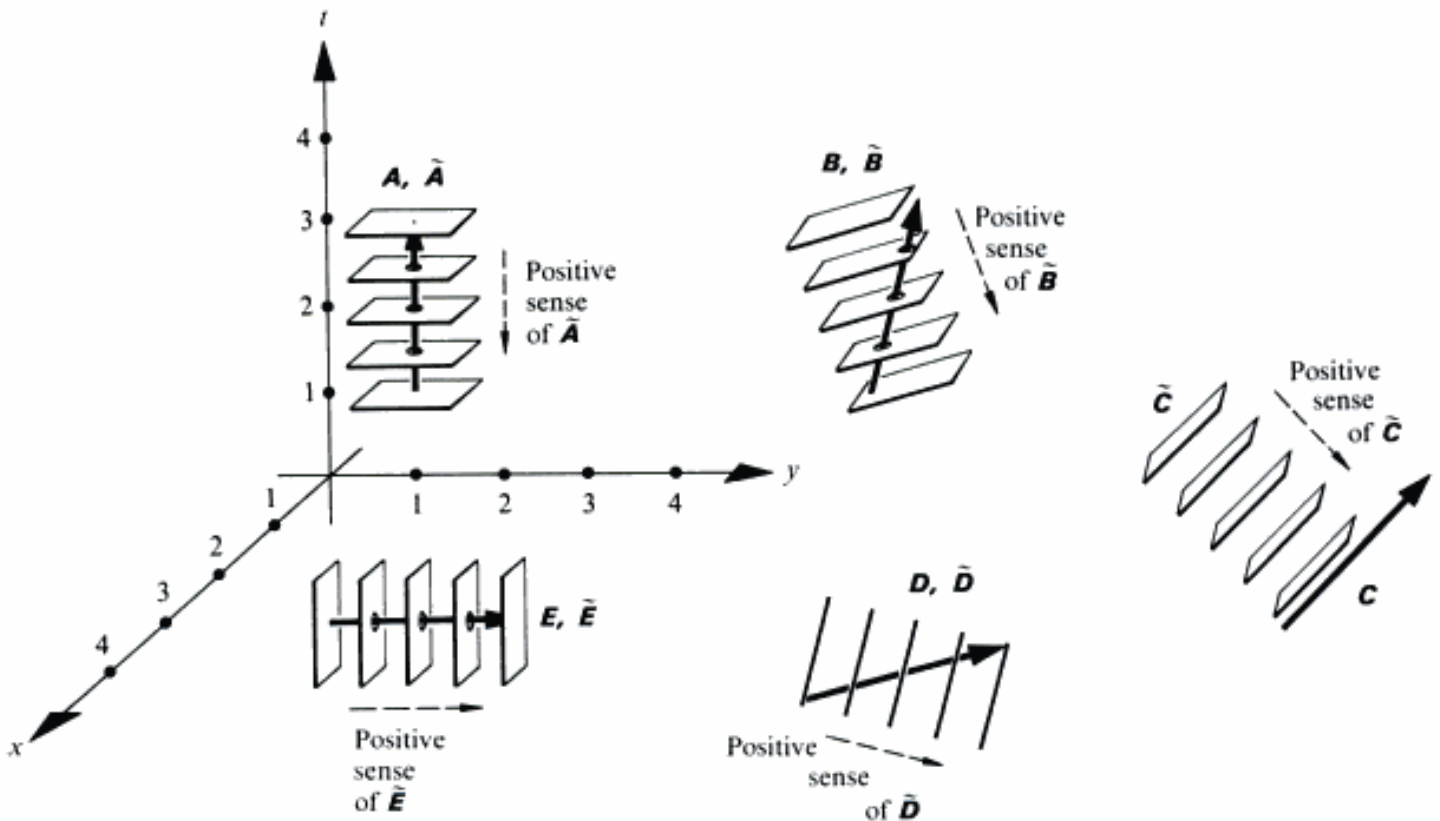
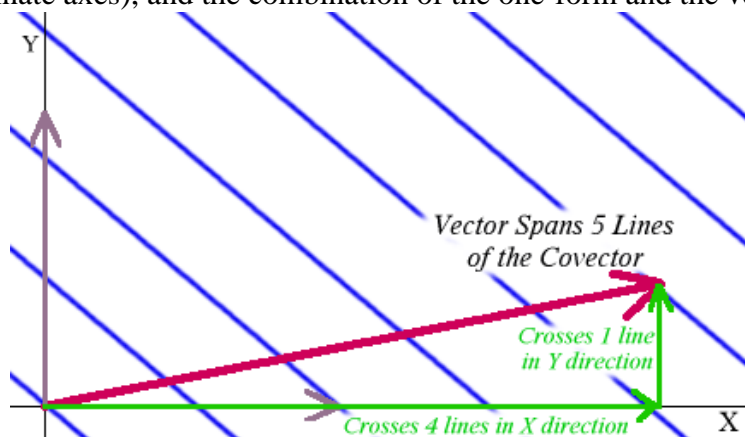


Figure 2.7.

Several vectors, \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} , \mathbf{E} , and corresponding 1-forms $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\tilde{\mathbf{C}}$, $\tilde{\mathbf{D}}$, $\tilde{\mathbf{E}}$. The process of drawing $\tilde{\mathbf{U}}$ corresponding to a given vector \mathbf{U} is quite simple. (1) Orient the surfaces of $\tilde{\mathbf{U}}$ orthogonal to the vector \mathbf{U} . (Why? Because any vector \mathbf{V} that is perpendicular to \mathbf{U} must pierce no surfaces of $\tilde{\mathbf{U}}$ ($0 = \mathbf{U} \cdot \mathbf{V} = \langle \tilde{\mathbf{U}}, \mathbf{V} \rangle$) and must therefore lie in a surface of $\tilde{\mathbf{U}}$.) (2) Space the surfaces of $\tilde{\mathbf{U}}$ so that the number of surfaces pierced by some arbitrary vector \mathbf{Y} (e.g., $\mathbf{Y} = \mathbf{U}$) is equal to $\mathbf{Y} \cdot \mathbf{U}$.

Note that in the figure the surfaces of $\tilde{\mathbf{B}}$ are, indeed, orthogonal to \mathbf{B} ; those of $\tilde{\mathbf{C}}$ are, indeed, orthogonal to \mathbf{C} , etc. If they do not look so, that is because the reader is attributing Euclidean geometry, not Lorentz geometry, to the spacetime diagram. He should recall, for example, that because \mathbf{C} is a null vector, it is orthogonal to itself ($\mathbf{C} \cdot \mathbf{C} = 0$), so it must itself lie in a surface of the 1-form $\tilde{\mathbf{C}}$. Confused readers may review spacetime diagrams in a more elementary text, e.g., Taylor and Wheeler (1966).

A one-form also provides a way to *measure* a vector. If a one-form is visualized as a collection of parallel hyperplanes, then the “length” of a vector with respect to that one-form corresponds to the *number of hyperplanes crossed by the vector*. The figure below shows a one-form being “applied” to a vector (for the moment, ignore the coordinate axes), and the combination of the one-form and the vector is 5.



If we call the vector \mathbf{v} and the one-form σ , this is called the “dot product of \mathbf{v} and σ ”, “applying σ to \mathbf{v} ”, “the value of σ on \mathbf{v} ” or “the contraction of σ with \mathbf{v} ” (or just “tensor contraction” in general). The notation is just as varied : $\mathbf{v} \cdot \sigma = \mathbf{v}(\sigma) = \sigma(\mathbf{v}) = \langle \mathbf{v}, \sigma \rangle = \langle \mathbf{v} | \sigma \rangle = v^i \sigma_i$ (which should look familiar!)

The notation indicates that a one-form σ_i (which is a covariant vector) can be thought of as an operator $\sigma(\mathbf{v})$ with one slot in which we insert a contravariant vector v^i and get a scalar. Similarly, a contravariant vector v^i can be thought of as a function $\mathbf{v}(\sigma)$ which we give a one-form σ_i and get a scalar. This is just the directional derivative : if we take the dot product of a vector v^a and a one-form f_a ($v^a f_a$) we get a scalar representing the magnitude of the rate of change of the contours in the direction of the vector.

If we now look at the one-form and the vector in the figure above with respect to the coordinate axes, we see that the total number of lines the vector crosses is just the number of lines it crosses in the direction of each basis vector (green lines) in turn. It doesn't matter what path you follow – you need to cross the same number of lines! In other words, the count along each axis is the product of each component of the vector with the corresponding component of the one-form ($= v^i \sigma_i$). So in the x direction $v^1 \sigma_1 = 4$ and in the y direction $v^2 \sigma_2 = 1$. Note that we are not calculating a *distance*, since $(1^2+4^2)^{1/2} \neq 5$. The figure below shows another example :

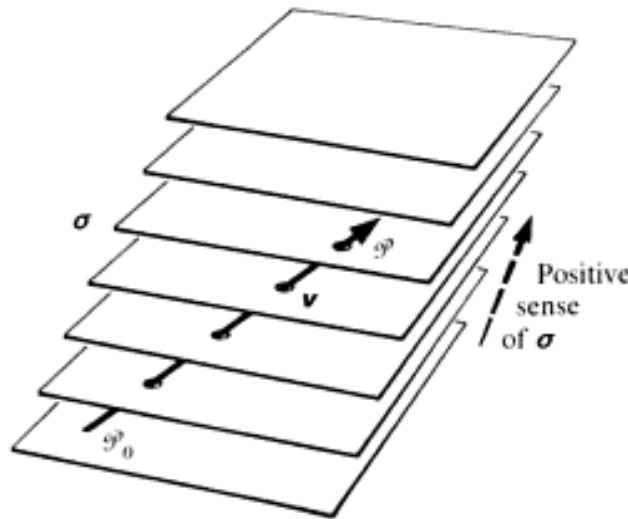


Figure 2.4.

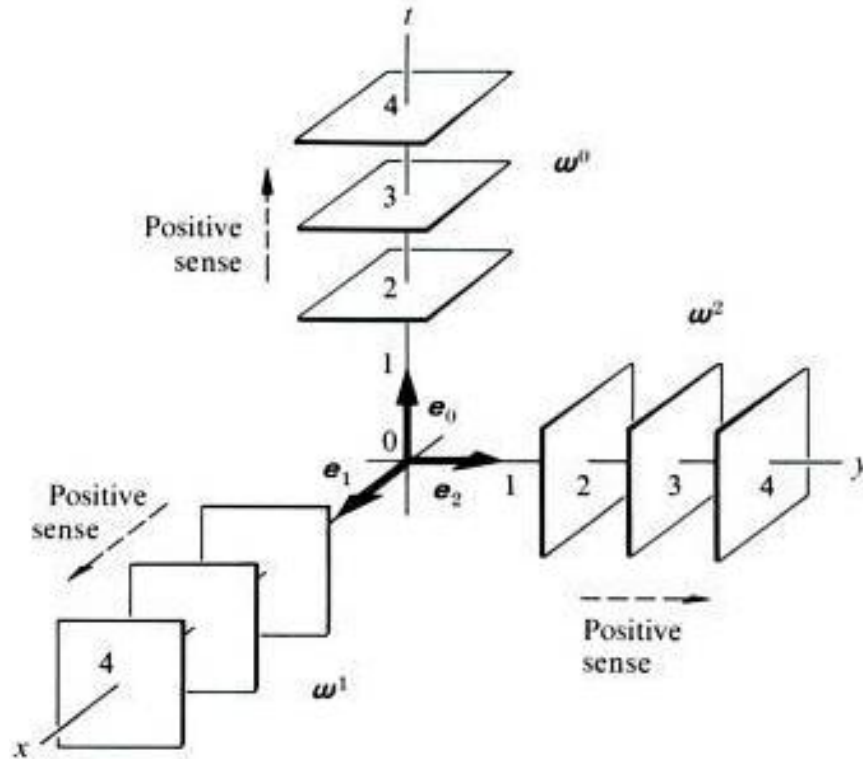
The vector separation $\mathbf{v} = \mathcal{P} - \mathcal{P}_0$ between two neighboring events \mathcal{P}_0 and \mathcal{P} ; a **l-form** σ ; and the piercing of σ by \mathbf{v} to give the number

$$\langle \sigma, \mathbf{v} \rangle = (\text{number of surfaces pierced}) = 4.4$$

In matrix notation, we can also think of a spacetime vector A^i as a 1x4 column and a one-form p_i as a 4x1 row :

$$\tilde{\mathbf{p}}(\mathbf{A}) = (p_0 \ p_1 \ p_2 \ p_3) \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix}$$

In flat spacetime, the coordinates $[x^0, x^1, x^2, x^3]$ can be thought of as functions, so their gradients can be calculated. The resulting “**basis one-forms**” (just like “basis vectors”) are described by $\omega^a = dx^a$, and have hyperplanes whose surfaces are defined by $x^a = \text{constant}$:



This means that each basis vector \mathbf{e}_a pierces exactly one surface of the corresponding basis one-form ω^a (plane #1 of each axis is not shown for clarity, but would be at the tip of each bold arrow), while the other basis vectors (axes) lie parallel to the surface of ω^a and thus pierce none : $\langle \omega^a, \mathbf{e}_\beta \rangle = \delta^a_\beta$.

Just as any vector can be described in terms of its components of the basis vectors ($\mathbf{v} = v^a \mathbf{e}_a$), one-forms can be described in terms of the basis one-forms : $\sigma = \sigma_a \omega^a$.

The idea of a one-form can be extended to tensors of all ranks, and is called a “differential form”. Specifically, a “**differential p-form**” or just “**p-form**” is a type (0,p) tensor. Thus the stress-energy-momentum tensor $T_{\alpha\beta}$ is a 2-form. *2-forms are always anti-symmetric, and are created by the wedge product of basis one-forms* (in the same way that tensors are created by the *dyadic product of basis vectors*).

Any two objects are called “**duals**” of each other if they combine to give a single number. In matrix algebra, row vectors and column vectors multiply to give a number, so they are duals. Likewise, vectors and one-forms (or equivalently contravariant vs. covariant components) are duals. Because of this, when dealing with one-forms the term “dual” often appears :

covariant vector \leftrightarrow dual vector

vector space of one-forms \leftrightarrow dual vector space, dual basis, dual space (last two most common)

Note : a * is often used to indicate dual basis spaces and their cotangent vector spaces

basis one-form \leftrightarrow dual basis, one-form basis, dual basis one-form, dual basis vector, basis dual vector

covector field \leftrightarrow dual field, dual vector field, form field, p-form field

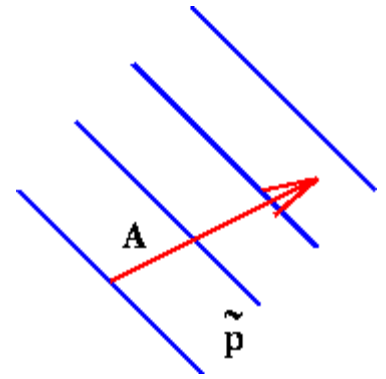
Some authors make a strong distinction between the covariant and contravariant forms of certain 4-vectors, claiming that (for example) velocity must be contravariant and momentum must be covariant, while other authors do not. Regardless, keep in mind that the contravariant and covariant forms can be easily transformed using the metric tensor. Any physical quantity can be described equally well by a vector or a one-form; the version used often depends on which is easiest in the given situation.

One-Forms as Frequency

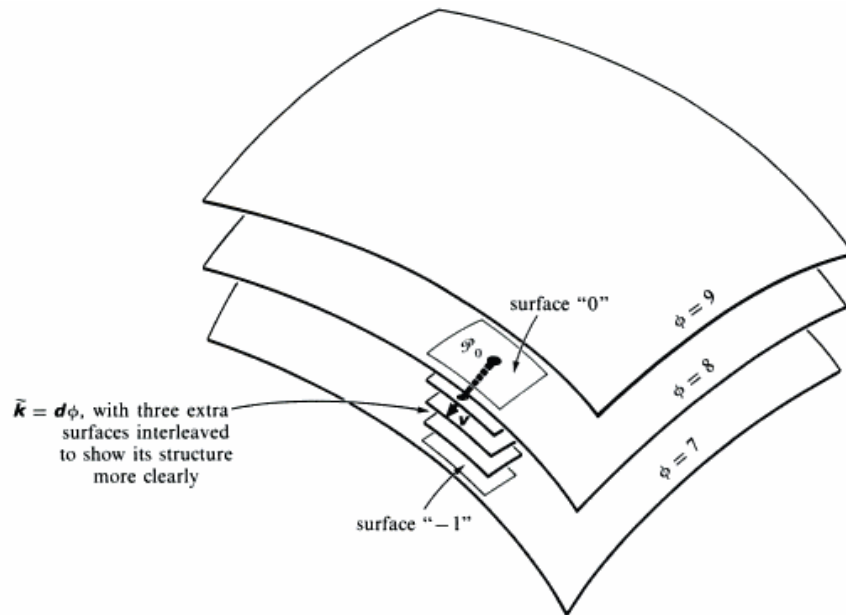
Another way to look at one-forms is as a kind of “frequency”: start at the beginning of the vector and draw cosine waves towards the end of the vector whose *frequency* is related to the magnitude of the one-form. Think of the lines representing the one-form as the crests of the waves. →

When calculating the dot product, count the number of cycles between the beginning and end of the vector. The dot product is the total (probably fractional) number of cycles.

For this figure, the length of the vector *with respect to the one-form* is 2.6 times the “wavelength” of the one-form. If we were to multiply the one-form by 2, we would *double* the number of blue lines, meaning the distance between the lines would *decrease* if the magnitude of the one-form *increased*. So the one-form does indeed act like a frequency: double it, and the number of crests increases, but the wavelength *decreases*.



Quantum physics associates a de Broglie wavelength with each particle. Take 2 points in spacetime, A and B, and run a vector between them, $\mathbf{v} = (B-A)$. It will pierce a certain number of one-form surfaces σ of integer phase. In this case, that number $\langle \sigma, \mathbf{v} \rangle$ represents the phase difference between the “waves” at A and B. Think of σ as a local pattern of waves near A, with fractional values uniformly interpolated between the integer surfaces (the figure below shows additional surfaces at $\phi=7.25, 7.50, 7.75$). **Just as the tangent vector $\mathbf{u} = dx/dt$ represents the local approximation to a worldline, the one-form represents the local form of the de Broglie wave’s surfaces.**



Quantum physics also deals with Hilbert-space state vectors “kets” $|\psi\rangle$ and their conjugate transposes “bras” $\langle\phi|$, which when combined with each other $\langle\phi|\psi\rangle$ produce a (complex) scalar that represents a probability amplitude. So bras and kets are also duals of each other, with bras equivalent to covariant one-forms and kets equivalent to contravariant vectors.



Exterior Derivatives

The exterior derivative is the generalization of the regular derivative to covariant “forms”. The exterior derivative of a “k-form” is a “(k+1)-form”. If you think of a scalar field as a “0-form”, then the exterior derivative of zero-, one- and two-forms corresponds to the operators grad, curl, and div (respectively).

The mathematical notations for the exterior derivative of a k-form ω include $\nabla \wedge \omega$ or $d\omega$

Specifically, the exterior derivative of a scalar field f is the one-form df :

$$df = \omega = (\partial f / \partial x_i) dx^i$$

Where df/dx_i is calculated for the particular function, and their formulas prepended to “dxⁱ”. See the “Covariant Transforms” subsection at the top of this document for an example.

A general k-form ω with covariant tensor components b_1, b_2, \dots has the format

$$\omega^1 = b_1 dx_1 + b_2 dx_2 + \dots$$

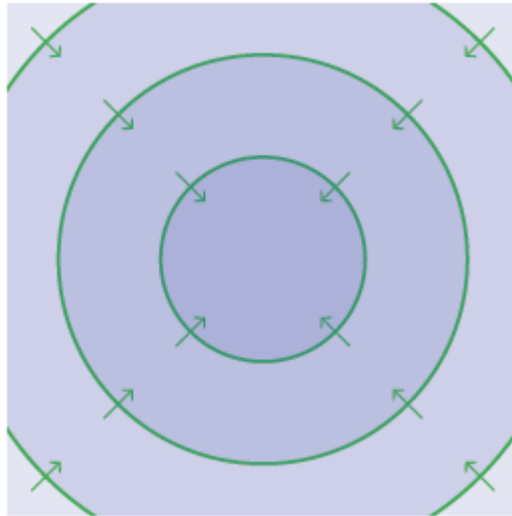
and its exterior derivative is

$$d\omega^1 = db_1 \wedge dx_1 + db_2 \wedge dx_2 + \dots$$

The external derivative can also be described in terms of covariant derivatives :

$$\begin{aligned} \nabla_{\mathbf{a}} \wedge \phi &= \nabla_{\mathbf{a}} \phi \\ \nabla_{\mathbf{a}} \wedge \omega_{\mathbf{b}} &= \nabla_{\mathbf{a}} \omega_{\mathbf{b}} - \nabla_{\mathbf{b}} \omega_{\mathbf{a}} \\ \nabla_{\mathbf{a}} \wedge \omega_{[\mathbf{bc}]} &= \nabla_{\mathbf{a}} \omega_{[\mathbf{bc}]} + \nabla_{\mathbf{b}} \omega_{[\mathbf{ca}]} + \nabla_{\mathbf{c}} \omega_{[\mathbf{ab}]} \end{aligned}$$

Visually, the exterior derivative is related to the “boundaries” of the k-form. In the picture below, a scalar field (blue) is a zero-form, and its exterior derivative is the boundaries between different values (green lines), which is the field’s contours (the green arrows indicate the “direction” of the one-form).



Because there is no “boundary of a boundary”, $d(d\omega) = \nabla \wedge \nabla \omega = 0$



Recap : Contravariant Vectors and Covariant (Dual) One-Forms

Basis vector : $\mathbf{e}_a = \partial_a = \partial/\partial x^a$
(contravariant)

Dual basis one-form : $\mathbf{e}^a = dx^a$
(covariant)

$$\mathbf{V} = V^a \mathbf{e}_a$$

$$V^a = \mathbf{e}^a \cdot \mathbf{V}$$

$$\mathbf{e}_a \cdot \mathbf{e}^b = \delta_a^b$$

$$\mathbf{V} = V_a \mathbf{e}^a$$

$$V_a = \mathbf{e}_a \cdot \mathbf{V}$$

$$g_{ac} \cdot g^{cb} = \delta_a^b$$

Remember, the values of the coefficients V^a and V_a depend on the particular basis, but the actual vector \mathbf{V} they represent does not!

It may seem odd that the components are defined/calculated as $V^a = \mathbf{e}^a \cdot \mathbf{V}$ and then the vector is defined as $\mathbf{V} = V^a \mathbf{e}_a$ but note : $\mathbf{e}_a V^a = \mathbf{e}_a (\mathbf{e}^a \cdot \mathbf{V}) = 1 \cdot \mathbf{V} = \mathbf{V}$

For a coordinate (holonomic) basis :

$$\mathbf{e}_a \cdot \mathbf{e}_b = g_{ab}$$

$$\mathbf{e}_a = g_{ab} \mathbf{e}^b$$

$$V_a = g_{ab} V^b$$

$$T_{mn} = g_{mu} g_{vn} T^{uv}$$

$$\mathbf{e}^a \cdot \mathbf{e}^b = g^{ab}$$

$$\mathbf{e}^a = g^{ab} \mathbf{e}_b$$

$$V^a = g^{ab} V_b$$

$$T^{mn} = g^{mu} g^{vn} T_{uv}$$

For an orthonormal (non-holonomic, tetrad) basis :

$$\mathbf{e}_a \cdot \mathbf{e}_b = \eta_{ab}$$

$$\mathbf{e}_a = \eta_{ab} \mathbf{e}^b$$

$$V_a = \eta_{ab} V^b$$

$$T_{mn} = \eta_{mu} \eta_{vn} T^{uv}$$

$$\mathbf{e}^a \cdot \mathbf{e}^b = \eta^{ab}$$

$$\mathbf{e}^a = \eta^{ab} \mathbf{e}_b$$

$$V^a = \eta^{ab} V_b$$

$$T^{mn} = \eta^{mu} \eta^{vn} T_{uv}$$

The “dot product” between two (contravariant) vectors \mathbf{A} and \mathbf{B} or two one-forms $\underline{\mathbf{P}}$ and $\underline{\mathbf{Q}}$ is :

$$\mathbf{A} \cdot \mathbf{B} = g_{uv} A^u B^v \quad \langle \underline{\mathbf{A}}, \underline{\mathbf{P}} \rangle = A^u P_u \quad \underline{\mathbf{P}} \cdot \underline{\mathbf{Q}} = g^{uv} P_u Q_v$$

Put another way, the “dot product” between two generic “vectors” (ignoring type) \mathbf{V} and \mathbf{W} can be written :

$$\mathbf{V} \cdot \mathbf{W} = g_{ab} V^a W^b = V^a W_b = V_a W^b = g^{ab} V_a W_b$$

Multiplying a tensor by two vectors can be done many ways :

$$x_i A^i_j y^j = g_{ir} x^r A_{ij} y^j = A_{ij} x^r y^j = x_i A^i_j g^{jr} y_r = A^{ir} x_i y_r$$

but only the first form can be correctly written in matrix form : $\mathbf{x}^T \mathbf{A} \mathbf{y}$