Lecture Notes: Tensors for Beginners

Based on Youtube series by eigenchris

Contents

1 Motivation: Coordinate Independence in Physics

Physical laws must not depend on how we label space.

Why Does This Matter?

In physics, we describe phenomena using mathematical equations. However, these equations should describe reality itself, not our particular choice of coordinates. Consider:

- **Newton's Laws**: The force on an object doesn't change if you rotate your coordinate system
- **Maxwell's Equations**: Electromagnetic fields behave the same regardless of your reference frame
- **General Relativity**: The curvature of spacetime is independent of coordinate choice

This principle is fundamental: *physical reality is invariant, even though our descriptions of it may change with coordinates.*

Example: Velocity

A particle moving in space has a velocity that exists independently of coordinates. If we rotate our coordinate axes, the particle does not suddenly move differently — only the *numbers* describing its velocity change.

For instance, a ball moving northward at 5 m/s has the same physical motion whether you describe it in:

- Coordinates aligned north-south and east-west
- Coordinates rotated 45 degrees
- Polar coordinates centered at your location

The velocity vector itself (the geometric object) remains the same. Only its components (the numbers in your coordinate system) change.

The Core Distinction

This distinction between:

- the geometric object (velocity)
- its coordinate representation (components)

is the core motivation for tensors.

Key Principle

> If a quantity represents something physical or geometric, then changing coordinates must not change the object itself, only how we describe it.

Tensors are defined so this principle is automatically satisfied. They provide a mathematical framework where:

1. Objects exist independently of coordinates
2. We can compute how their components transform when we change coordinates
3. Physical laws written in tensor notation are automatically coordinate-independent

Thought Exercises

Exercise. *Consider the following physical quantities. For each, identify whether it is coordinate-independent (invariant) or coordinate-dependent:*

1. *The temperature at a point in space*
2. *The velocity components* (v_x, v_y, v_z) *of a moving particle*
3. *The speed (magnitude of velocity) of a moving particle*
4. *The coordinates* (x, y, z) *of a point in space*
5. *The distance between two points*
6. *Newton's second law:* $\vec{F} = m\vec{a}$

Hint: Ask yourself, "If I rotate my coordinate system, does this quantity change?"

Exercise. *Understanding the distinction:*

1. *A car travels east at 60 km/h. You set up coordinates with the x-axis pointing north. What are the velocity components in your coordinate system?*
2. *Your friend uses coordinates with the x-axis pointing east. What are the velocity components in their coordinate system?*
3. *Did the car's motion change between parts (a) and (b)? What changed?*
4. *Explain why we need a mathematical framework that distinguishes between "the velocity" and "the velocity components"*

Exercise. *Why tensors matter in physics:*

1. *If Maxwell's equations were not coordinate-independent, what would this imply about electromagnetic phenomena?*
2. *In general relativity, gravity is described by the curvature of spacetime. Why must this curvature be coordinate-independent?*
3. *Give an example of a physical law that would be problematic if it depended on coordinate choice*

2 Forward and Backward Transformations

Forward and backward transformations refer to the rules that make us move back and forward between different coordinate systems.

Tensors are invariant under a coordinate system change, so we need to understand how we move back and forward between different systems.

Let's assume we have 2 coordinate systems, an old basis $\vec{e_i}$ and a new basis $\tilde{\vec{e_i}}$, in $\mathbb{R}^2$:

- Old basis = $\{\vec{e_1}, \vec{e_2}\}$
- New basis = $\{\tilde{\vec{e_1}}, \tilde{\vec{e_2}}\}$

Moving from the old basis to the new basis, we want to express the new basis in terms of the old one, so:

$$\tilde{\vec{e_1}} = A\vec{e_1} + B\vec{e_2} \tag{1}$$

$$\tilde{\vec{e_2}} = C\vec{e_1} + D\vec{e_2} \tag{2}$$

The coefficients A, B, C, D can be inserted into a 2x2 matrix, writing down the basis vectors as *row* vectors:

$$[\vec{e_1}, \vec{e_2}] \begin{bmatrix} A & C \\ B & D \end{bmatrix} \tag{3}$$

You can see that performing the row-to-column multiplication, you obtain exactly the results in 1 and 2.

This might be confusing at first, because we're writing down vectors as row-vectors and not column-vectors, as you usually see in textbooks. The reason is that we're dealing with *basis vectors* and not *vector components* (which will be written in column form), and this is very important to make the multiplication we'll see later make sense and to enforce the fact that basis vectors transform differently (covariant) with respect to vector components (contravariant).

That being said, we can define the matrix F as forward matrix, to move from the old to the new basis vector coordinate systems:

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{4}$$

$$[\tilde{\vec{e_1}}, \tilde{\vec{e_2}}] = [\vec{e_1}, \vec{e_2}] \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{5}$$

And re-writing 1 and 2:

$$\tilde{\vec{e_1}} = F_{11}\vec{e_1} + F_{21}\vec{e_2} \tag{6}$$

$$\tilde{\vec{e_2}} = F_{12}\vec{e_1} + F_{22}\vec{e_2} \tag{7}$$

Now, the same can be done in the opposite direction, i.e. from the new basis to the old basis:

$$\vec{e_1} = E\tilde{\vec{e_1}} + F\tilde{\vec{e_2}} \tag{8}$$

$$\vec{e_2} = G\tilde{\vec{e_1}} + H\tilde{\vec{e_2}} \tag{9}$$

Obtaining with the same logic, a matrix that we'll call B for backward:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{10}$$

$$[\vec{e_1}, \vec{e_2}] = [\tilde{\vec{e_1}}, \tilde{\vec{e_2}}] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{11}$$

And re-writing 8 and 9:

$$\vec{e_1} = B_{11}\tilde{\vec{e_1}} + B_{21}\tilde{\vec{e_2}} \tag{12}$$

$$\vec{e_2} = B_{12}\tilde{\vec{e_1}} + B_{22}\tilde{\vec{e_2}} \tag{13}$$

All this can be extended to any $\mathbb{R}^n$ space dimension:

$$\begin{aligned} \tilde{\vec{e_1}} &= F_{11}\vec{e_1} + F_{21}\vec{e_2} + \cdots + F_{n1}\vec{e_n} \\ \tilde{\vec{e_2}} &= F_{12}\vec{e_1} + F_{22}\vec{e_2} + \cdots + F_{n2}\vec{e_n} \\ &\vdots \\ \tilde{\vec{e_n}} &= F_{1n}\vec{e_1} + F_{2n}\vec{e_2} + \cdots + F_{nn}\vec{e_n} \end{aligned}$$

With F being a $n \times n$ matrix:

$$\begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1n} \\ F_{21} & F_{22} & \cdots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nn} \end{bmatrix} \tag{14}$$

We can then write down in compact form, for both F and B:

$$\boxed{\begin{aligned} \tilde{\vec{e_i}} &= \sum_{j=1}^{n} F_{ji}\vec{e_j} \qquad (15a) \\ \vec{e_i} &= \sum_{j=1}^{n} B_{ji}\tilde{\vec{e_j}} \qquad (15b) \end{aligned}}$$

As you can already guess, the F and B matrices are one the inverse of the other: $B = F^{-1}$, and to simply prove it, we can easily re-arrange and substitute equations 15a and 15b:

$$\vec{e_i} = \sum_{j=1}^{n} B_{ji} \tilde{\vec{e_j}}$$

$$\vec{e_i} = \sum_{j} B_{ji} \left(\sum_{k} F_{kj} \tilde{\vec{e_k}} \right)$$

$$\vec{e_i} = \sum_{k} \left(\sum_{j} F_{kj} B_{ji} \right) \tilde{\vec{e_k}}$$

Now, to make sense of this, you see that we have the *blue* new basis vectors on both left and right side of the equation, so they obviously have to match when $i = k$, which means that:

$$\sum_{j} F_{kj} B_{ji} = \begin{cases} 1, & \text{if } i = k, \\ 0, & \text{if } i \neq k. \end{cases} \tag{16}$$

Which, if we expand in $\mathbb{R}^n$, represents the identity matrix. And this behavior is so common that we have a name for it, called the Kronecker delta:

$$\delta_{ik} = \begin{cases} 1, & \text{if } i = k, \\ 0, & \text{if } i \neq k. \end{cases} \tag{17}$$

Exercises

Exercise. *Consider the following basis transformation in $\mathbb{R}^2$:*

$$\tilde{\vec{e_1}} = 2\vec{e_1} + \vec{e_2}, \quad \tilde{\vec{e_2}} = -\vec{e_1} + \vec{e_2}$$

1. *Write down the forward transformation matrix F*
2. *Compute the backward transformation matrix $B = F^{-1}$*
3. *Verify that $FB = I$ (the identity matrix)*
4. *Express the old basis vectors in terms of the new basis using the matrix B*

Exercise. *Understanding the Kronecker delta:*

1. *Write out $\sum_{j=1}^{3} \delta_{ij}$ for $i = 1, 2, 3$. What pattern do you notice?*
2. *Evaluate $\sum_{j=1}^{n} F_{kj} B_{ji}$ and explain why it equals δ_{ki}*
3. *Show that for any vector components v^i, the expression $\sum_{i=1}^{n} \delta_{ij} v^i = v^j$*

Exercise. *Consider a rotation by angle θ counterclockwise. The forward transformation is:*

$$F = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

1. *Compute $B = F^{-1}$ (use the formula for 2×2 matrix inverse)*
2. *Compare B with F. What geometric transformation does B represent?*
3. *Verify that rotating by θ then by $-\theta$ returns you to the original basis*

3 Vectors

Now we'll move on definining what the first tensor we're seeing is: a vector.

A definition might be, that a vector is a list of numbers, that can be added together and multiplied by a number, but what we're actually describing like this, are the vector components, and not the vector itself.

We have to understand that **a vector is an invariant**, **vector components are not**, as they depend on the coordinate systems we use to compute them.

Another definition seen frequently is that a vector is like an arrow, having a direction and a magnitude. You can scale (up or down) vectors multiplying them by scalar numbers, and you can add them by using the tip-to-toe rule. **The problem with this definition, is that not all vectors can be visualized as arrows.** Indeed, the vectors that can be visualized as arrows are a special kind of vectors called "Euclidean vectors"

Moving to a different definition, we can say that a vector is a member of a vector space V. A Vector space V is defined as a collection of four things:

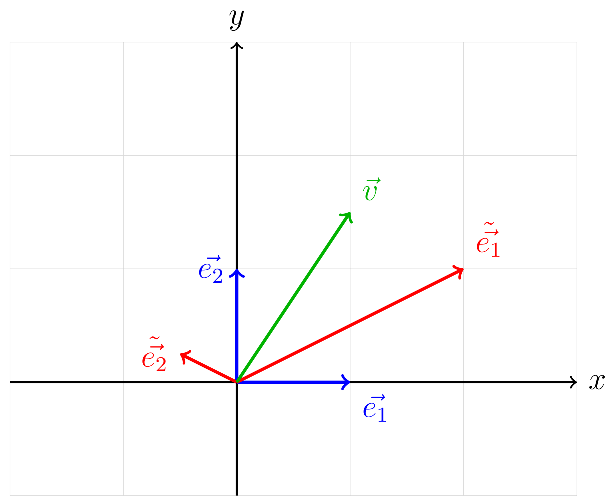
$$(V, S, +, \cdot) \tag{18}$$

- V is a set of vectors
- S is a set of scalars
- $+$ is a sort of addition rule by which we can add vectors together
- $\cdot$ is a sort of scaling rule by which we can scale vectors by acting with scalars

Now let's say that we have a vector $\vec{v}$ sitting in space and we want to find its components in our basis vectors defined before:

$$\{\vec{e_1}, \vec{e_2}\}$$
$$\{\tilde{\vec{e_1}}, \tilde{\vec{e_2}}\}$$

This means that we want to measure $\vec{v}$ in two different basis. Practically, we can do that in the following example, and try to understand how the vector components in different basis relate to each other.



In these basis, we have the forward and backward matrices as follows:

$$F = \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} \tag{19}$$

$$B = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ -1 & 2 \end{bmatrix} \tag{20}$$

And if you simply eyeball the diagram and calculate the components of the vector $\vec{v}$, you'll find that:

$$\begin{bmatrix} 1 \\ 1.5 \end{bmatrix}_{\vec{e_i}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\tilde{\vec{e_i}}}$$

We saw in the previous section, that, for basis vectors:

$$\{\vec{e_1}, \vec{e_2}\} \xrightarrow{\quad} \boxed{\mathbf{F}} \xrightarrow{\quad} \{\tilde{\vec{e_1}}, \tilde{\vec{e_2}}\}$$

$$\{\tilde{\vec{e_1}}, \tilde{\vec{e_2}}\} \xrightarrow{\quad} \boxed{\mathbf{B}} \xrightarrow{\quad} \{\vec{e_1}, \vec{e_2}\}$$

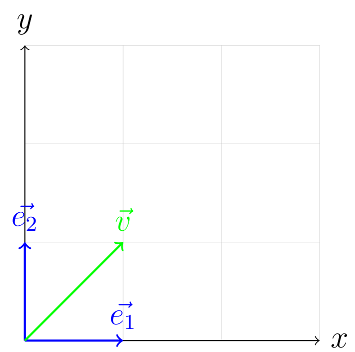
So let's try and check if following the same logic for the vector components, we can move from one coordinate systems to the other, applying the forward matrix, so:

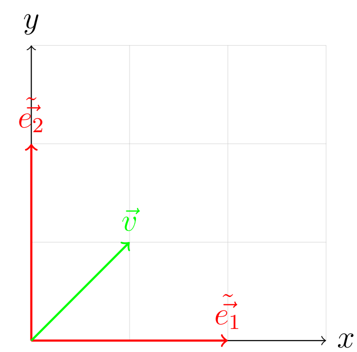
$$F \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}_{\vec{e_i}} = \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}_{\vec{e_i}} = \begin{bmatrix} 1.25 \\ 1.375 \end{bmatrix}$$

This does not seem right, doens't it? Why don't we try to apply the B matrix instead?

$$B\begin{bmatrix}1\\1.5\end{bmatrix}_{\vec{e_i}} = \begin{bmatrix}\frac{1}{4} & \frac{1}{2}\\-1 & 2\end{bmatrix}\begin{bmatrix}1\\1.5\end{bmatrix}_{\vec{e_i}} = \begin{bmatrix}1\\2\end{bmatrix}$$

So this tells us that, for basis vectors, forward matrix brings us from old to new, and backward from new to old, but for vector components, it's the opposite, backward matrix brings us from old to new, and forward from new to old. This might seem weird at a first look, but it actually makes sense. Let me give you a couple of examples. Let's start with a simple $\vec{v}$ in the basis $\vec{e_i}$ and imagine that you want to describe the same vector (invariant) in a different basis $\tilde{\vec{e_i}}$, which is only up-scaled by a factor of 2 wrt to the old basis:

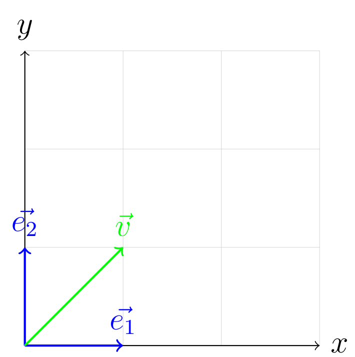


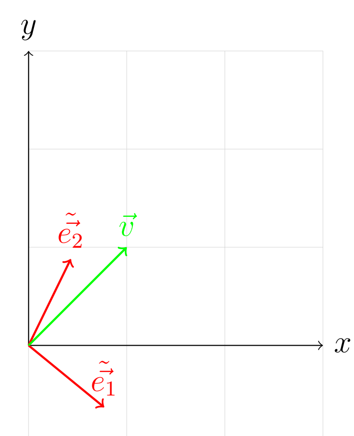


If you think about this, as basis vectors scaled up, the vector components have to scale down by the same amount, for the vector itself to be invariant wrt to change of coordinate system. $\vec{v}$ "looks" smaller as seen by the new "bigger" coordinate system basis vectors. Indeed:

$$\begin{bmatrix}1\\1\end{bmatrix}_{\vec{e_i}} \begin{bmatrix}1/2\\1/2\end{bmatrix}_{\tilde{\vec{e_i}}}$$

The same happens for a simple basis vector rotation:





In this case, intuitively and from the picture, it's clearly visible that a clockwise rotation of the basis vectors corresponds to a counter-clockwise rotation of the vector components.

Next step would be proving this in more dimentions. We can write the vector as a linear combination of its vector components in two different coordinate systems / basis vectors:

$$\vec{v} = v_1\vec{e_1} + v_2\vec{e_2} + \cdots + v_n\vec{e_n} = \sum_{j=1}^{n} v_j\vec{e_j}$$
$$\vec{v} = \tilde{v}_1\tilde{\vec{e_1}} + \tilde{v}_2\tilde{\vec{e_2}} + \cdots + \tilde{v}_n\tilde{\vec{e_n}} = \sum_{j=1}^{n} \tilde{v}_j\tilde{\vec{e_j}}$$

Bringing back our basis vector forward and backward transformations:

$$\boxed{\begin{aligned} \tilde{\vec{e_i}} &= \sum_{j=1}^{n} F_{ji}\vec{e_j} && (21a) \\ \vec{e_i} &= \sum_{j=1}^{n} B_{ji}\tilde{\vec{e_j}} && (21b) \end{aligned}}$$

We can write:

$$\vec{v} = \sum_{j=1}^{n} v_j\vec{e_j} = \sum_{j=1}^{n} v_j \left(\sum_{i=1}^{n} B_{ij}\tilde{\vec{e_i}} \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} B_{ij}v_j \right) \tilde{\vec{e_i}}$$

As you can see, **this actually proves that, to move from the old components to the new components, we use the backward transformation matrix**, and **to move from the new components to the old components, we use the forward transformation matrix**:

$$\boxed{\begin{aligned} \tilde{v}_i &= \sum_{j=1}^{n} B_{ij}v_j && (22a) \\ v_i &= \sum_{j=1}^{n} F_{ij}\tilde{v}_j && (22b) \end{aligned}}$$

So, summarizing what we've learned so far, we know the transformation rules that basis vectors and vector components obey:

$$\boxed{\begin{aligned} \tilde{\vec{e_i}} &= \sum_{j=1}^{n} F_{ji}\vec{e_j} && (23a) \\ \vec{e_i} &= \sum_{j=1}^{n} B_{ji}\tilde{\vec{e_j}} && (23b) \end{aligned}}$$

$$\boxed{\begin{aligned} \tilde{v}_i &= \sum_{j=1}^{n} B_{ij}v_j && (24a) \\ v_i &= \sum_{j=1}^{n} F_{ij}\tilde{v}_j && (24b) \end{aligned}}$$

Since the vector components behave contrary to the basis vectors, we say that they are contra-variant.
We'll see later indeed, that vectors are contra-variant tensors and from now on, we're going to make a small change in the way we write vector components due to this behavior, and we're writing them with the index on top and not on the bottom:

$$\vec{v} = \sum_{i=1}^{n} v^i \vec{e_i} = \sum_{i=1}^{n} \tilde{v}^i \tilde{\vec{e_i}}$$

Exercises

Exercise. *Given the forward transformation matrix*

$$F = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

1. *Compute the backward transformation matrix* $B = F^{-1}$
2. *If a vector has components* $v^1 = 2, v^2 = 3$ *in the old basis, find its components in the new basis*
3. *Verify that* $B \cdot \tilde{v} = v$ *gives you back the original components*

Exercise. *Consider a vector* $\vec{v} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ *in a standard orthonormal basis.*

1. *If the basis vectors are scaled by a factor of 2, what are the new vector components?*
2. *If the basis vectors are rotated counterclockwise by 45°, write down the forward transformation matrix*
3. *Explain in words why vector components transform opposite to basis vectors*

4 Covectors and covector components

You may find in some places, that covectors are defined to be "basically" row vectors, so you may think that's just it, if you have a vector written in column, you flip it and you have a covector, but that's not quite right and simple.

Column-vectors and row-vectors are fundamentally different types of objects. The reason you may think are basically the same but flipped, is that we normally deal with *orthonormal basis*, which is a basis where all vectors are one unit long and perpendicular to each other. But generally, this is not true in any coordinate system.

To realize this, we need to think at row vectors as functions acting on column vectors, so let's think about a general covector α acting on a general vector $\vec{v}$:

$$\alpha(\vec{v}) = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_n v^b = \sum_{i=1}^{n} \alpha_i v^i \tag{25}$$

Ultimately, a covector is a function that takes an input from a vector space and returns a scalar as output:

$$\alpha : V \to \mathbb{R} \tag{26}$$

They obey the linearity rule:

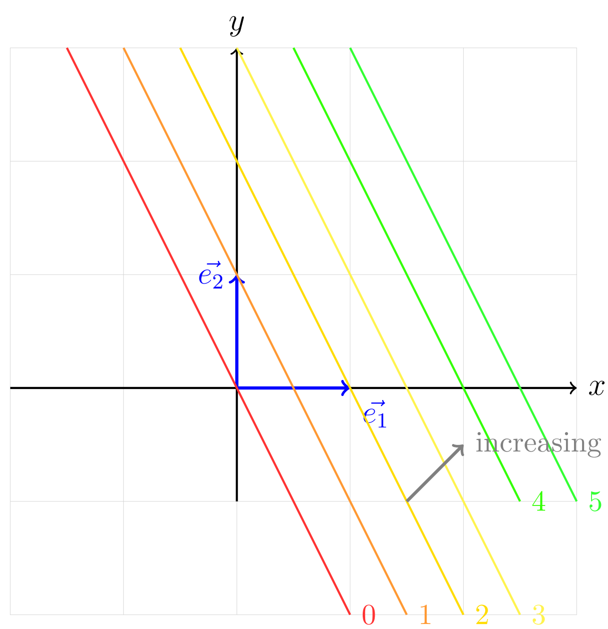
$$\alpha(n\vec{v} + m\vec{w}) = n\alpha(\vec{v}) + m\alpha(\vec{w}) \tag{27}$$

How can we visualized these covectors though? There's a nice way of doing it and we can start by thinking about a generic 2D covector as a function on two variables x and y:

$$\begin{bmatrix} 2 & 1 \end{bmatrix} \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = 2x + 1y$$

So how do we visualize a function of two variables that produces one output? This is very similar to what topographers do to visualize on a piece of 2D paper a topographic map of some mountains and valley. This is done by drawing curves of constant elevation value. And by looking at a map like this, we know that when we see these lines very close together they represent a place where the elevation changes very steeply, whereas where they are less dense, the elevation does not change so steeply.

Continuing with our example, we can start asking, where is this function equal to zero? $2x + 1y = 0$? This is the line $y = -2x$, same we can do for 1, 2, 3 and for negative as well.



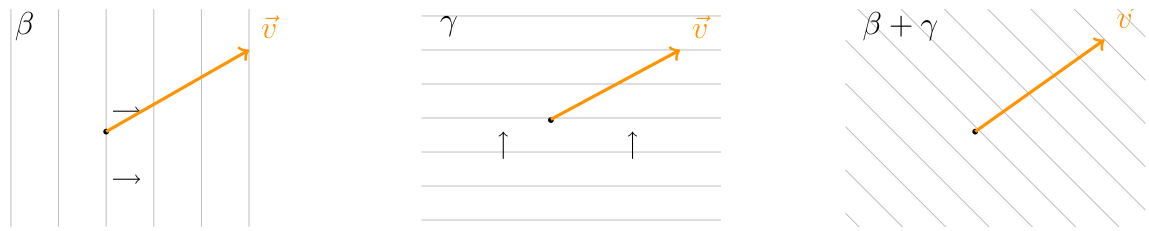
The stack is increasing towards the upper right so it has a direction towards north-east in our case.

You can actually think about a covector α acting on a vector $\vec{v}$ as giving a scalar output, equals to the number of times the vector pierces one of the covector's lines.

Now what happens if we scale up the covector, let's say by a factor of 2? We basically make is much denser, hence the vector will pierce the lines double the time it did before. And the result would be the same if we choose instead, of scaling up the vector by 2, as the vector will then pierce a double number of lines, being its magnitude longer.

Let's continue on how we can visualize two covector addition.

$$\beta(\vec{v}) = 3$$
$$\gamma(\vec{v}) = 2$$



The diagrams are not perfect, but a sum of covectors β and γ would show as a stack with the same density as β in the beta-direction, and same density as γ in the gamma-direction, which in this case would visualize as a NE-pointing stack, and since β has an horizontal density of 3 and γ a vertical density of 2, the sum will simply be the vector piercing the lines 5 times.

$$(\beta + \gamma)(\vec{v}) = 5$$
$$(\beta + \gamma)(\vec{v}) = \beta(\vec{v}) + \gamma(\vec{v})$$

Summing up, we've seen that for covectors, we also are able to scale them, and perform addition, and that gives us a hint about the fact that covectors are actually part of a Vector space.
We have that the set of covectors that act on vectors in V form a new vector space called the *dual space* V^*, with its own set of addition and scalar operations:

$$(V, S, +, \cdot)$$
$$(V^*, S, +, \cdot)$$

The elements of V^* are covectors, which are functions that go from V to the real numbers $\mathbb{R}$, with their own addition and scaling rules:

$$(n \cdot \alpha)(\vec{v}) = n\alpha(\vec{v})$$
$$(\beta + \gamma)(\vec{v}) = \beta(\vec{v}) + \gamma(\vec{v})$$

As per vectors, covectors are also invariant, they're purely geometric object, independend of the reference frame/coordinate system used to describe them. Their components though, exactly like vector components, are not invariant.

When we write a column vector, for example, as follows, we represent it by how much of each basis vector I need to make this vector, so as a linear combination of the "scaled" basis vectors (scaled by the vector components values):

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\vec{e_i}} \text{ we mean } 2\vec{e_1} + 1\vec{e_2} \tag{28}$$

But what does it mean to do this for covectors? Which are functions? Like if I write down

$$\begin{bmatrix} 2 & 1 \end{bmatrix} \tag{29}$$

This is not as intuitive because remember that covectors do not live in the same vector space, they live in the dual vector space, and they are functions from vectors to real numbers, so we can't use basis vectors in V to represent covectors of V^*.

What we can do is intruduce two special covectors, such that, considering the basis $\{\vec{e_1}, \vec{e_2}\}$ for V:

$$\begin{aligned} \epsilon^1 (\vec{e_1}) &= 1 & \qquad \epsilon^1 (\vec{e_2}) &= 0 \\ \epsilon^2 (\vec{e_1}) &= 0 & \qquad \epsilon^2 (\vec{e_2}) &= 1 \end{aligned}$$

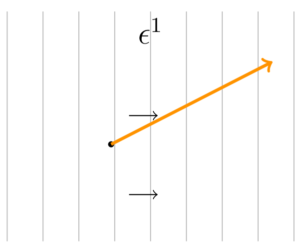
so basically:

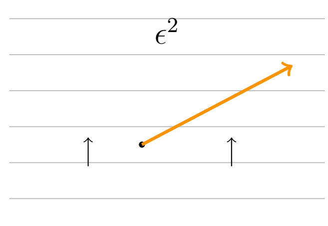
$$\epsilon^i (\vec{e_j}) = \delta^i{}_j = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{30}$$

What happens when we apply such covectors to a generic vector?

$$\begin{aligned} \epsilon^1 (\vec{v}) = \epsilon^1 \left(v^1 \vec{e_1} + v^2 \vec{e_2}\right) &= v^1 \\ \epsilon^2 (\vec{v}) = \epsilon^2 \left(v^1 \vec{e_1} + v^2 \vec{e_2}\right) &= v^2 \\ \epsilon^i (\vec{v}) &= v^i \end{aligned}$$

So what these covectors are doing, is projecting out vector components.





Let's now generalize and apply a general covector α to a vector $\vec{v}$:

$$\alpha (\vec{v}) = \alpha \left(v^1 \vec{e_1} + v^2 \vec{e_2}\right) = v^1 \alpha (\vec{e_1}) + v^2 (\vec{e_2})$$

We can write the components $v_i = \epsilon^i(\vec{v})$ so that:

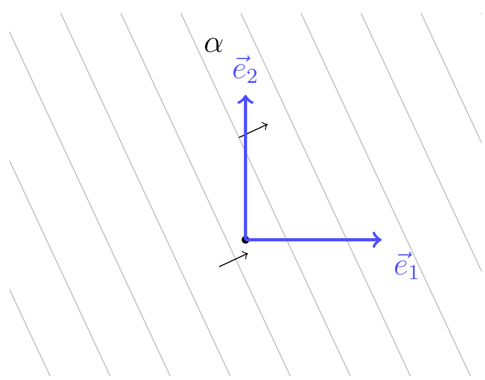
$$\alpha(\vec{v}) = \epsilon^1(\vec{v})\,\alpha(\vec{e_1}) + \epsilon^2(\vec{v})\,\alpha(\vec{e_2})$$

We define $\alpha(\vec{e_1}) = \alpha_1$ and $\alpha(\vec{e_2}) = \alpha_2$ so that:

$$\begin{aligned}
\alpha(\vec{v}) &= \alpha_1\epsilon^1(\vec{v}) + \alpha_2\epsilon^2(\vec{v}) \\
\alpha(\vec{v}) &= \left(\alpha_1\epsilon^1 + \alpha_2\epsilon^2\right)(\vec{v}) \\
\alpha &= \alpha_1\epsilon^1 + \alpha_2\epsilon^2
\end{aligned}$$

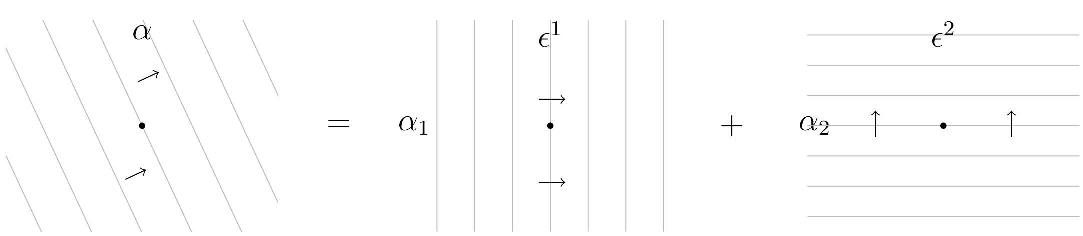
We've now written a covector α as linear combination of our epsilon covectors defined above. **What this means is that the ϵ covectors form a basis for the dual vector space V^*** and we call this ϵ the dual basis because they're a basis for the dual vector space V^*.

We may try to understand this visually and geometrically, since so far we derived this algebraically.



We can get the components of α by applying the covector to the basis vectors: $\alpha(\vec{e_1}) = \alpha_1$ and $\alpha(\vec{e_2}) = \alpha_2$. In terms of the dual basis $\{\epsilon^1, \epsilon^2\}$ we can visualize the decomposition

$$\alpha = \alpha_1\,\epsilon^1 + \alpha_2\,\epsilon^2.$$



The process is, we start with our vector basis $\vec{e_1}, \vec{e_2}$, then using this $\epsilon^i(\vec{e_j}) = \delta^i{}_j$, we get the dual covector basis, and then using those we can express any covector as a combination of the dual basis.

Remember though, that these ϵ covector basis is not the only one we can use to express α. We can start with a different vector basis, $\tilde{\vec{e_1}}, \tilde{\vec{e_2}}$ and then applying the rule $\tilde{\epsilon^i}\left(\tilde{\vec{e_j}}\right) = \delta^i{}_j$ we get another dual vector basis, that can be used to express the same α in a different covector basis.

Allright, so now, let's say we have a covector $\alpha = 2\epsilon^1 + 1\epsilon^2$ represented in the old covector basis ϵ^i:

$$\begin{bmatrix} 2 & 1 \end{bmatrix}_{\epsilon^i} \tag{31}$$

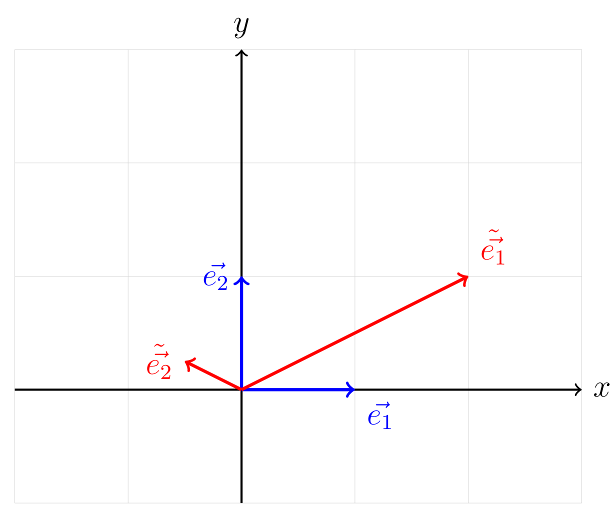
which means they have components:

$$\alpha\left(\vec{e_1}\right) = 2$$
$$\alpha\left(\vec{e_2}\right) = 1$$

What would these components look like in the new covector basis $\tilde{\epsilon^i}$?
For this we need to apply the covector α to the new basis vectors:

$$\alpha\left(\tilde{\vec{e_1}}\right) = \tilde{\alpha_1}$$
$$\alpha\left(\tilde{\vec{e_2}}\right) = \tilde{\alpha_2}$$

And taking this coordinate systems in example:



We see that $\tilde{\vec{e_1}} = (2\vec{e_1} + 1\vec{e_2})$ and $\tilde{\vec{e_2}} = (-1/2\vec{e_1} + 1/4\vec{e_2})$ so:

$$\tilde{\alpha_1} = \alpha\left(\tilde{\vec{e_1}}\right) = \alpha\left(2\vec{e_1} + 1\vec{e_2}\right) = 5$$
$$\tilde{\alpha_2} = \alpha\left(\tilde{\vec{e_2}}\right) = \alpha\left(-1/2\vec{e_1} + 1/4\vec{e_2}\right) = -3/4$$

$$\begin{bmatrix}2 & 1\end{bmatrix}_{\epsilon^i} \qquad\qquad \begin{bmatrix}5 & -3/4\end{bmatrix}_{\tilde{\epsilon^i}}$$

Remember what were the F and B matrices? **If you make your calculation, you will see that for covector components, forward brings from old to new, and backward brings from new to old**:

$$\begin{bmatrix}2 & 1\end{bmatrix}_{\epsilon^i} F = \begin{bmatrix}2 & 1\end{bmatrix}_{\epsilon^i} \begin{bmatrix}2 & -1/2 \\ 1 & 1/4\end{bmatrix} = \begin{bmatrix}5 & -3/4\end{bmatrix}_{\tilde{\epsilon^i}}$$
$$\begin{bmatrix}5 & -3/4\end{bmatrix}_{\tilde{\epsilon^i}} B = \begin{bmatrix}5 & -3/4\end{bmatrix}_{\tilde{\epsilon^i}} \begin{bmatrix}1/4 & 1/2 \\ -1 & 2\end{bmatrix} = \begin{bmatrix}2 & 1\end{bmatrix}_{\epsilon^i}$$

This is actually the opposite of what we've found for vector components under a change of basis. **This is why we can't just flip column vectors to row vectors to get covectors.** It works in an orthonormal basis, like the $\vec{e_i}$ and correspondent dual basis defined by ϵ^i. But it does not work in the new voctor basis $\tilde{\vec{e_i}}$ and correspondent covector basis $\tilde{\epsilon^i}$.

Like we did for vectors bafore, we've gone from old basis vectors to new basis vectors and we found that that requires the forward mateix F.

Now we want to do something similar for the covector basis, we want to go from an old covector basis ϵ^i to a new covector basis $\tilde{\epsilon^i}$.

$$\tilde{\epsilon^1} = Q_{11}\epsilon^1 + Q_{12}\epsilon^2 \tag{32}$$
$$\tilde{\epsilon^2} = Q_{21}\epsilon^1 + Q_{22}\epsilon^2 \tag{33}$$

To find the coefficients, we start by applying:

$$\tilde{\epsilon^1}\left(\vec{e_1}\right) = Q_{11}\epsilon^1\left(\vec{e_1}\right) + Q_{12}\epsilon^2\left(\vec{e_1}\right) = Q_{11}$$
$$\tilde{\epsilon^1}\left(\vec{e_2}\right) = Q_{11}\epsilon^1\left(\vec{e_2}\right) + Q_{12}\epsilon^2\left(\vec{e_2}\right) = Q_{12}$$

Given this, we can re-write:

$$\tilde{\epsilon^1} = \tilde{\epsilon^1}\left(\vec{e_1}\right)\epsilon^1 + \tilde{\epsilon^1}\left(\vec{e_2}\right)\epsilon^2$$

Now if we bring back our backward transformation, we know that:

$$\vec{e_1} = 1/4\tilde{\vec{e_1}} - 1\tilde{\vec{e_2}}$$
$$\vec{e_2} = 1/2\tilde{\vec{e_1}} + 2\tilde{\vec{e_2}}$$

We can type down:

$$\tilde{\epsilon^1} = \tilde{\epsilon^1}\left(1/4\tilde{\vec{e_1}} - 1\tilde{\vec{e_2}}\right)\epsilon^1 + \tilde{\epsilon^1}\left(1/2\tilde{\vec{e_1}} + 2\tilde{\vec{e_2}}\right)\epsilon^2$$
$$\tilde{\epsilon^1} = \left(1/4\tilde{\epsilon^1}\left(\tilde{\vec{e_1}}\right) - 1\tilde{\epsilon^1}\left(\tilde{\vec{e_2}}\right)\right)\epsilon^1 + \left(1/2\tilde{\epsilon^1}\left(\tilde{\vec{e_1}}\right) + 2\tilde{\epsilon^1}\left(\tilde{\vec{e_2}}\right)\right)\epsilon^2$$
$$\tilde{\epsilon^1} = 1/4\epsilon^1 + 1/2\epsilon^2$$

$$\tilde{\epsilon^2} = \tilde{\epsilon^2}\left(1/4\tilde{\vec{e_1}} - 1\tilde{\vec{e_2}}\right)\epsilon^1 + \tilde{\epsilon^2}\left(1/2\tilde{\vec{e_1}} + 2\tilde{\vec{e_2}}\right)\epsilon^2$$
$$\tilde{\epsilon^2} = \left(1/4\tilde{\epsilon^2}\left(\tilde{\vec{e_1}}\right) - 1\tilde{\epsilon^2}\left(\tilde{\vec{e_2}}\right)\right)\epsilon^1 + \left(1/2\tilde{\epsilon^2}\left(\tilde{\vec{e_1}}\right) + 2\tilde{\epsilon^2}\left(\tilde{\vec{e_2}}\right)\right)\epsilon^2$$
$$\tilde{\epsilon^2} = -1\epsilon^1 + 2\epsilon^2$$

If you notice, this is quite familiar to the backward transformation, that means that to go from the old dual basis to the new dual basis, we use the B matrix. This is also valid for every dimension, we'll leave this proof out, but this is the result, showing both the already seen vector basis transformation, and this new dual covector basis one:

$$\boxed{\begin{aligned} \tilde{\vec{e_j}} &= \sum_{j=1}^{n} F_{ij}\vec{e_i} \qquad (34a) \\ \vec{e_j} &= \sum_{j=1}^{n} B_{ij}\tilde{\vec{e_i}} \qquad (34b) \end{aligned}}$$

$$\boxed{\begin{aligned} \tilde{\epsilon^i} &= \sum_{j=1}^{n} B_{ij}\epsilon^j \qquad (35a) \\ \epsilon^i &= \sum_{j=1}^{n} F_{ij}\tilde{\epsilon^j} \qquad (35b) \end{aligned}}$$

That's why we write covector indeces on top, because the transorm like vector components, opposite to the basis vectors (i.e. contra-variantly)

With this now, we can also show how covector components transform:

$$\boxed{\begin{aligned} \tilde{\alpha_j} &= \sum_{i=1}^{n} F_{ij}\alpha_i \qquad (36a) \\ \alpha_j &= \sum_{i=1}^{n} B_{ij}\tilde{\alpha_j} \qquad (36b) \end{aligned}}$$

Covector components transform in the same way vector basis do.
To summarize all the transformation rules so far:

$$\tilde{\vec{e}}_j = \sum_{i=1}^{n} F_{ij}\,\vec{e}_i \qquad \vec{e}_j = \sum_{i=1}^{n} B_{ij}\,\tilde{\vec{e}}_i$$

$$\tilde{\epsilon}^{\,i} = \sum_{j=1}^{n} B_{ij}\,\epsilon^j \qquad \epsilon^i = \sum_{j=1}^{n} F_{ij}\,\tilde{\epsilon}^{\,j}$$

$$\tilde{v}^{\,i} = \sum_{j=1}^{n} B_{ij}\,v^j \qquad v^i = \sum_{j=1}^{n} F_{ij}\,\tilde{v}^{\,j}$$

$$\tilde{\alpha}_j = \sum_{i=1}^{n} F_{ij}\,\alpha_i \qquad \alpha_j = \sum_{i=1}^{n} B_{ij}\,\tilde{\alpha}_i$$

- Vector components are contravariant (high index), transform opposite to the basis vector transformation
- Covector components are covariant (low index), transform like the basis vector transform
- Dual vector basis are contravariant (high index), transform opposite to the basis vector transformation

Exercises

Exercise. *Consider a covector* α *with components* $[\alpha_1, \alpha_2] = [3, -1]$ *in the standard basis* $\{\vec{e_1}, \vec{e_2}\}$.

1. *Evaluate* $\alpha(\vec{v})$ *where* $\vec{v}$ *has components* $v^1 = 2, v^2 = 4$
2. *If we scale* α *by a factor of 2, what happens to the output* $\alpha(\vec{v})$*?*
3. *If instead we scale* $\vec{v}$ *by a factor of 2, what happens to the output?*
4. *Verify that* $\alpha(2\vec{u} + 3\vec{w}) = 2\alpha(\vec{u}) + 3\alpha(\vec{w})$ *(linearity)*

Exercise. *Understanding the dual basis:*

1. *Given basis vectors* $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, *verify that* $\epsilon^1(\vec{e_1}) = 1$ *and* $\epsilon^1(\vec{e_2}) = 0$
2. *For a covector* $\alpha = 2\epsilon^1 + 5\epsilon^2$, *compute* $\alpha(\vec{e_1})$ *and* $\alpha(\vec{e_2})$
3. *A vector* $\vec{v} = 3\vec{e_1} + 4\vec{e_2}$. *Compute* $\epsilon^1(\vec{v})$ *and* $\epsilon^2(\vec{v})$
4. *What is the relationship between* $\epsilon^i(\vec{v})$ *and the vector components* v^i*?*

Exercise. *Covector transformations: Using the same transformation from the vectors section with matrices:*

$$F = \begin{bmatrix} 2 & -1/2 \\ 1 & 1/4 \end{bmatrix}, \quad B = \begin{bmatrix} 1/4 & 1/2 \\ -1 & 2 \end{bmatrix}$$

1. *A covector has components* $[\alpha_1, \alpha_2] = [4, 2]$ *in the old basis. Find its components in the new basis*

2. *Compare this result with how vector components transform. Which matrix (F or B) is used?*

3. *Verify that* $\alpha(\vec{v})$ *gives the same scalar value in both coordinate systems*

4. *Explain in words why covector components are covariant (transform like basis vectors)*

Exercise. *Visualizing covectors:*

1. *Draw the level sets (lines where the function equals a constant) for the covector* $\alpha = [1, 1]$

2. *A vector* $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ *pierces these level sets. How many lines does it cross?*

3. *Verify this by computing* $\alpha(\vec{v})$

4. *If we double the covector to* $2\alpha = [2, 2]$*, how does the spacing of level sets change?*

5. *How does the answer to part (c) change when using* 2α *instead of* α*?*

Exercise. *Contravariant vs. Covariant:*

1. *Complete this table for the transformation rules:*

Quantity	*Index Position*	*Transforms with*
Basis vectors $\vec{e}_i$	*subscript*	*?*
Vector components v^i	*superscript*	*?*
Dual basis ϵ^i	*superscript*	*?*
Covector components α_i	*subscript*	*?*

2. *Why do we write vector components with upper indices and covector components with lower indices?*

3. *Give an intuitive explanation for why vector components transform opposite to basis vectors*

5 Linear Maps

We'll present linear maps with different definitions, one more practical, one geometric and finally the more abstract definition.

Practically, matrices are the coordinate version of linear maps. If you represent vectors as column, covectors as row vectors, linear maps are represented by matrices. How do matrices act on vectors? Let's see a simple example:

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \tag{37}$$

What is actually doing? If we test using a "copy" of the basis vectors $\vec{e_1}, \vec{e_2}$, we see that the outputs are respectively, the first and second column of the matrix:

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \tag{38}$$

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \tag{39}$$

Why do we stress about these being copies of basis vectors? Because Linear Maps tranform input vectors, they do not transform the basis. The i^{th} column of the matrix is the image of the i^{th} copy of the basis vector.

Geometrically, linear maps are spatial transforms that:

- Keep gridlines parallel
- Keep gridlines evenly spaced
- Keep the origin stationary

And abstractly, linear maps are maps from a vector space to another:

$$L : V \to W \tag{40}$$

And obey linearity rules:

$$\begin{aligned} L(\vec{v} + \vec{w}) &= L(\vec{v}) + L(\vec{w}) \\ L(n\vec{v}) &= nL(\vec{v}) \end{aligned}$$

Now let's see how the abstract and coordinate definitions are related to each other. Let's review the formula for matrix multiplication of a 2x2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \tag{41}$$

At a first glance, this rule may come out from nowhere, but this is actually a direct consequence of the abstract definition above. To show this, let's start by saying that we have a linear map $L : V \to V$ acting on a vector $\vec{v}$ to produce some output vector $\vec{w}$. **The assumption to stay in the same vector space is needed to ensure we can use the same basis vectors on both inputs and outputs**:

$$\begin{aligned} \vec{w} = L(\vec{v}) &= L\left(v^1\vec{e_1} + v^2\vec{e_2}\right) \\ &= v^1 L(\vec{e_1}) + v^2 L(\vec{e_2}) \end{aligned}$$

These $L(\vec{e_1})$ and $L(\vec{e_2})$ are just vectors in V so we can write them down in terms of the basis vectors:

$$\begin{aligned} L(\vec{e_1}) &= L_1^1\vec{e_1} + L_1^2\vec{e_2} \\ L(\vec{e_2}) &= L_2^1\vec{e_1} + L_2^2\vec{e_2} \end{aligned}$$

$$\begin{aligned} \vec{w} &= v^1\left(L_1^1\vec{e_1} + L_1^2\vec{e_2}\right) + v^2\left(L_2^1\vec{e_1} + L_2^2\vec{e_2}\right) \\ &= \left(L_1^1 v^1 + L_2^1 v^2\right)\vec{e_1} + \left(L_1^2 v^1 + L_2^2 v^2\right)\vec{e_2} \\ &= w^1\vec{e_1} + w^2\vec{e_2} \end{aligned}$$

So we've derived how to transform the v coefficients into the w coefficients.

$$\begin{aligned} w^1 &= L_1^1 v^1 + L_2^1 v^2 \\ w^2 &= L_1^2 v^1 + L_2^2 v^2 \end{aligned}$$

which are nothing more than the usual 2x2 matrix multiplication rule:

$$\begin{bmatrix} L_1^1 & L_2^1 \\ L_1^2 c & L_2^2 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = \begin{bmatrix} w^1 \\ w^2 \end{bmatrix} \tag{42}$$

For any number of dimensions, we can write:

$$\vec{w} = L(\vec{v}) = \sum_{i=1}^{n} w^i \vec{e_i} \tag{43}$$

And the linear maps coefficients:

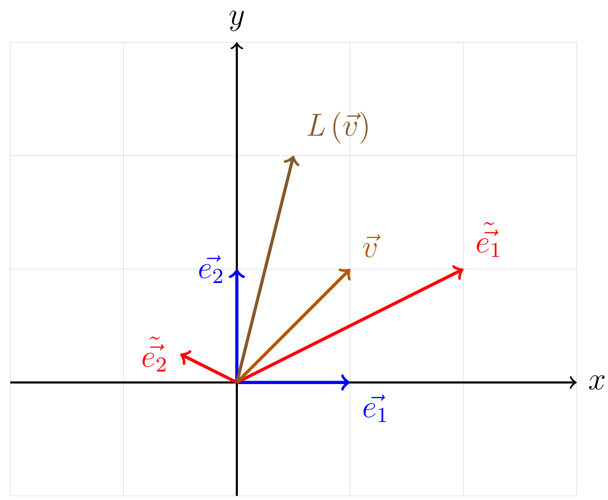
$$L(\vec{e_i}) = \sum_{j=1}^{n} L_i^j \vec{e_j} \tag{44}$$

And we can transform finally, the v components into w components:

$$w^i = \sum_{j=1}^{n} L_j^i v^j \tag{45}$$

which is indeed the matrix multiplication formula applied to n-dimension space.

We can now proceed checking how linear maps transform when moving from one basis to another.



Consider the diagram above and the linear map expressed by the matrix in the $\vec{e_i}$ basis:

$$\begin{bmatrix} 1/2 & 0 \\ 0 & 2 \end{bmatrix}_{\vec{e_i}} \tag{46}$$

As reminder, this means that:

$$L(\vec{e_1}) = 1/2\vec{e_1}$$
$$L(\vec{e_2}) = 2\vec{e_2}$$

Vector $\vec{v}$ has components $1, 1$ in the basis $\vec{e_i}$. So what would be the components of the vector $L(\vec{v})$?

$$L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\vec{e_i}}\right) = \begin{bmatrix} 1/2 \\ 2 \end{bmatrix}_{\vec{e_i}} \tag{47}$$

Remember that, if we apply the backward transformation to these vector components, we obtain the vector components in the new basis $\tilde{\vec{e_i}}$:

$$\begin{bmatrix} 1/4 & 1/2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\vec{e_i}} = \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}_{\tilde{\vec{e_i}}} \tag{48}$$

So, the next question might be, what are the components of the linear map output vector in the new basis?

$$L\left(\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}_{\tilde{\vec{e_i}}}\right) = \begin{bmatrix} ? \\ ? \end{bmatrix}_{\tilde{\vec{e_i}}} \tag{49}$$

We cannot use the matrix representation we used before, because that's tied to the old basis vectors, so what we need to do is finding a new matrix that represents the linear map in the new basis, which means we need to find these coefficients:

$$L\left(\tilde{\vec{e_1}}\right) = \tilde{L}_1^1\tilde{\vec{e_1}} + \tilde{L}_1^2\tilde{\vec{e_2}}$$
$$L\left(\tilde{\vec{e_2}}\right) = \tilde{L}_2^1\tilde{\vec{e_1}} + \tilde{L}_2^2\tilde{\vec{e_2}}$$

$$\sum_{q=1}^{n} \tilde{L}_i^q\tilde{\vec{e_q}} = L\left(\tilde{\vec{e_i}}\right) \tag{50}$$

We can use the forward transform to re-write the new basis vectors in terms of the old ones:

$$\sum_{q=1}^{n} \tilde{L}_i^q\tilde{\vec{e_q}} = L\left(\sum_{j=1}^{n} F_i^j\vec{e_j}\right)$$
$$= \sum_{j=1}^{n} F_i^j L\left(\vec{e_j}\right)$$

Now we can use the formula that expresses the linear map output as linear combinaton of the old basis vectors (44) and re-arranging:

$$\sum_{q=1}^{n} \tilde{L}_i^q\tilde{\vec{e_q}} = \sum_{j=1}^{n} F_i^j \sum_{k=1}^{n} L_j^k\vec{e_k}$$
$$= \sum_{j=1}^{n}\sum_{k=1}^{n} F_i^j L_j^k\vec{e_k}$$

We can proceed re-writing the old basis vector in terms of the new one, using the backward transformation 15b:

$$\sum_{q=1}^{n} \tilde{L}_i^q\tilde{\vec{e_q}} = \sum_{j=1}^{n}\sum_{k=1}^{n} F_i^j L_j^k \sum_{l=1}^{n} B_k^l\tilde{\vec{e_l}}$$
$$= \sum_{l=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} B_k^l L_j^k F_i^j\tilde{\vec{e_l}}$$

Since the letter we use for the summation does not really matter, we can just rename all the q to l and we get:

$$\sum_{l=1}^{n} \tilde{L}_i^l\tilde{\vec{e_l}} = \sum_{l=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} B_k^l L_j^k F_i^j\tilde{\vec{e_l}}$$

Finally, we see that the $\tilde{L}$ coefficients are obtained by multiplying the backward transform on the left, and forward transform on the right, of the old coefficients:

$$\boxed{\tilde{L}^l_i = \sum_{j=1}^{n}\sum_{k=1}^{n} B^l_k L^k_j F^j_i} \tag{51a}$$

You can think about this, the $\widetilde{L}$ transformation takes us from the input vector components to the output vector components in the new basis. Instead of "travelling" along this "arrow" (bottom of diagram below), we move from the new vector componens to the new vector components using the forward transformation. Then, to transform the components of the input vector to the components of the output vector, we apply the L transformation. And finally, to get from the old vector components to the new ones, we apply the backward transformation B

$$\begin{tikzcd}
L\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\vec{e}_i}\right) \arrow[r, bend left, "L"] & \begin{bmatrix} 1/2 \\ 2 \end{bmatrix}_{\vec{e}_i} \arrow[d, "B"] \\
L\left(\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}_{\widetilde{\vec{e}}_i}\right) \arrow[u, "F"] \arrow[r, bend right, "\widetilde{L}"'] & \begin{bmatrix} ? \\ ? \end{bmatrix}_{\widetilde{\vec{e}}_i}
\end{tikzcd}$$

Now you've noticed how heavy the derivation and notation, especially related to the summation involved, was above. We will, from now on, adopt the Einstein summation convention, which states that, everytime indeces appear both at the bottom and at the top, a summation is implicitly expected.

That means that in our case, of the linear map transformation:

$$\boxed{\tilde{L}^l_i = B^l_k L^k_j F^j_i} \tag{52a}$$

Because the repeating indeces, appearing at the top and bottom, are k and j, they're implicitly summed from 1 to n. Using this trick now, we can try deriving the backward transformation, from the $\widetilde{L}$ to the direct linear map L.

Understanding Index Position in Einstein Notation

Something you I was wondering is, when coming from conventional matrix multiplication, you clearly know that order or multiplication matters. When we derived the above equation, you notice that at the end of the day, we re-arrange coefficients to make it look ordered, but that really does not matter there, as all the components are scalar quantities, as such, they obey the standard commutative property.

How to we relate these concepts then? It's all a matter of index position and contraction. In tensor notation, the position of indices (upper vs. lower) carries crucial information about matrix multiplication:

Convention for Linear Maps: For a linear map L^i_j:

- **Lower index** j = column index (input)
- **Upper index** i = row index (output)

This comes directly from matrix-vector multiplication:

$$w^i = L^i_j v^j \tag{53}$$

The matrix "consumes" the j index (contraction) and "produces" the i index.
How do we read the transformation rule then?
For $\tilde{L}^l_i = B^l_k L^k_j F^j_i$, we trace the index flow:

$$\begin{array}{rll} F^j_i: & \text{input } i \text{ (lower)} & \to \text{output } j \text{ (upper)} \\ L^k_j: & \text{input } j \text{ (lower)} & \to \text{output } k \text{ (upper)} \\ B^l_k: & \text{input } k \text{ (lower)} & \to \text{output } l \text{ (upper)} \end{array}$$

The flow is $i \xrightarrow{F} j \xrightarrow{L} k \xrightarrow{B} l$, which tells us the matrix order is $\tilde{L} = B \cdot L \cdot F$.

Key Point: Repeated indices (one upper, one lower) indicate contractions (summations). The order of terms in Einstein notation ($B^l_k L^k_j F^j_i$ vs $L^k_j F^j_i B^l_k$) doesn't matter—it's the index connections that determine the matrix multiplication order.

Let's take this equation, and what we want, is isolating the L^k_j on the right side. To do that, we need to take out those B and F maps on the right side. How do we do this? We can consider that multiplying F and B together gives the identity matrix, so we can multiply by F and B on both sides of the equation.
But how do we choose the indeces? We need to follow a specific logic and rule: to contract with existing index, you need the opposit position (upper$\leftrightarrow$lower)

- We start with $\tilde{L}^l_i = B^l_k L^k_j F^j_i$;
- We want to obtain L with two new indeces, let's say s and t: L^s_t;
- We want to remove (i.e. contract) B and F, which have free indeces l and i;
- Let's start with B^l_k, we have l to contract on the upper, so we choose to multiply by F^s_l;
- Now for F^j_i, we want i to contract on the lower, so we choose to multiply by B^i_t;
- At the end, these new coefficients give: $F^j_i B^i_t = \delta^j_t$ which is the identity matrix.

$$
\begin{aligned}
\tilde{L}^l_i &= B^l_k L^k_j F^j_i \\
F^s_l \tilde{L}^l_i B^i_t &= F^s_l B^l_k L^k_j F^j_i B^i_t \\
F^s_l \tilde{L}^l_i B^i_t &= \delta^s_k L^k_j \delta^j_t \\
F^s_l \tilde{L}^l_i B^i_t &= \delta^s_k L^k_t \\
F^s_l \tilde{L}^l_i B^i_t &= L^s_t
\end{aligned}
$$

So we obtain:

$$
\boxed{F^s_l \tilde{L}^l_i B^i_t = L^s_t} \tag{54a}
$$

Now, summarizing this, we basically have found that linear maps are (1,1)-tensors, because they transform using both forward and backward transformations, so they're both contravariant and covariant once each.

Exercises

1. **Basic Matrix Action**

   Given the linear map L with matrix representation in basis $\{\vec{e_1}, \vec{e_2}\}$:

   $$L = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$$

   (a) What is $L(\vec{e_1})$ in terms of $\vec{e_1}$ and $\vec{e_2}$?

   (b) What is $L(\vec{e_2})$ in terms of $\vec{e_1}$ and $\vec{e_2}$?

   (c) Compute $L(\vec{v})$ where $\vec{v} = 3\vec{e_1} - 2\vec{e_2}$.

2. **Verifying Linearity**

   Which of the following maps are linear? For those that are not, explain why.

   (a) $T(\vec{v}) = 3\vec{v}$

   (b) $T(x, y) = (x + 1, y)$

   (c) $T(x, y) = (x + y, 2y)$

   (d) $T(x, y) = (x^2, y)$

3. **Constructing the Matrix**

   A linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ satisfies:

   $$
   \begin{aligned}
   L(\vec{e_1}) &= 2\vec{e_1} + 3\vec{e_2} \\
   L(\vec{e_2}) &= -\vec{e_1} + 5\vec{e_2}
   \end{aligned}
   $$

   Write down the matrix representation of L and compute $L(4\vec{e_1} + 7\vec{e_2})$.

4. **Einstein Notation Practice**

   Write out the following expressions explicitly without Einstein notation (showing all sums):

   (a) $v^i w_i$ (assuming 3 dimensions)

   (b) $A^i_j B^j_k$ (for 2×2 matrices)

   (c) $\delta^i_j v^j$

5. **Index Manipulation**

   Starting from $w^i = L^i_j v^j$, multiply both sides by a covector α_i and simplify. What type of mathematical object is the final result?

6. **Change of Basis - Forward and Backward**

   Consider two bases in $\mathbb{R}^2$: the standard basis $\{\vec{e_1}, \vec{e_2}\}$ and a new basis $\{\tilde{\vec{e_1}}, \tilde{\vec{e_2}}\}$ where:

$$\tilde{\vec{e_1}} = 2\vec{e_1} + \vec{e_2}$$
$$\tilde{\vec{e_2}} = \vec{e_1} + 3\vec{e_2}$$

   (a) Write down the forward transformation matrix F.

   (b) Compute the backward transformation matrix B by finding F^{-1}.

   (c) Verify that $FB = I$.

7. **Linear Map Transformation**

   Using the bases from Exercise 6, consider a linear map with matrix representation:

$$L = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}_{\vec{e_i}}$$

   in the original basis. Find the matrix representation $\tilde{L}$ in the new basis using the formula $\tilde{L}^l_i = B^l_k L^k_j F^j_i$.

8. **Index Flow Analysis**

   For the expression $T^a_b = C^a_c D^c_d E^d_b$:

   (a) Identify all free indices and all dummy (summed) indices.

   (b) Trace the index flow from input to output.

   (c) Write the equivalent matrix multiplication.

9. **Inverse Transformation**

   Given $\tilde{L}^l_i = B^l_k L^k_j F^j_i$, derive the expression for L^k_j in terms of $\tilde{L}^l_i$ by:

   (a) Multiplying by appropriate forward/backward transforms

   (b) Using the identity $F^s_l B^l_k = \delta^s_k$ and $B^i_t F^j_i = \delta^j_t$

(c) Simplifying using the Kronecker delta properties

10. **Composition of Linear Maps**

If ${L_1}^i_j$ and ${L_2}^j_k$ are two linear maps, their composition is $(L_1 \circ L_2)^i_k = {L_1}^i_j {L_2}^j_k$.

(a) Explain why the index structure makes sense for composition.

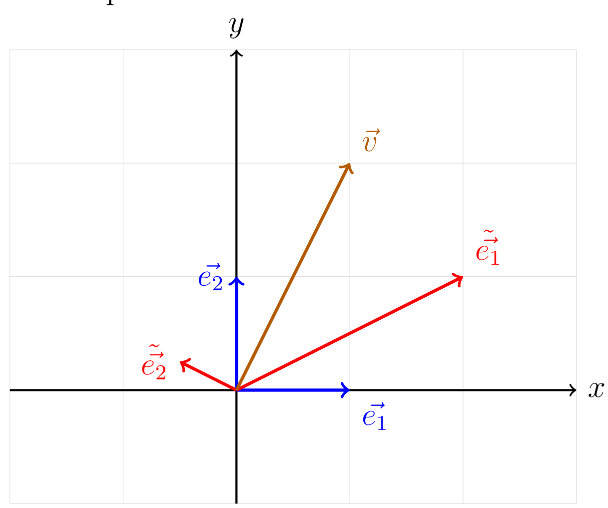
(b) Show that if both L_1 and L_2 transform according to $\tilde{L} = BLF$, then their composition transforms correctly as well.

(c) Is the order of composition important? Verify by computing ${L_1}^i_j {L_2}^j_k$ vs ${L_2}^i_j {L_1}^j_k$ for:

$$L_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$$

6 Metric Tensor

We start this chapter by asking a supposedly simple question, how do you get the length of a vector? So let's bring back our usual diagram with a sample vector living in our 2D space with a couple of basis vectors defined.



$$\begin{aligned} \vec{v} &= 1\vec{e_1} + 2\vec{e_2} \\ &= 5/4\tilde{\vec{e_1}} + 3\tilde{\vec{e_2}} \end{aligned}$$

Obviously, we can use Pythagora's theorem in the $\vec{e_i}$ basis, and compute the length of the vector as:

$$\|\vec{v}\|^2 = (1)^2 + (2)^2 = 5$$
$$\|\vec{v}\| = \sqrt{5} \approx 2.236$$

Well let's try doing the same in the new basis $\tilde{\vec{e_i}}$, why not?

$$
\begin{aligned}
\|\vec{v}\|^2 &= (5/4)^2 + (3)^2 = 169/16 \\
\|\vec{v}\| &= \sqrt{169/16} \approx 3.24
\end{aligned}
$$

Mmm, well that does not seems right, doesn't it? **Reason is that Pythagora's theorem only works in orthonormal basis like $\vec{e_i}$, and it does not yield the correct result when used in a non-orthonormal one like $\tilde{\vec{e_i}}$**

So what's the correct formula to use to get the length of a vector, independently of the basis used? This is the **dot product formula**, which in the two different coordinate systems (i.e. basis vectors):

$$
\begin{aligned}
\|\vec{v}\|^2 &= \vec{v} \cdot \vec{v} \\
&= \left(v^1 \vec{e_1} + v^2 \vec{e_2}\right) \cdot \left(v^1 \vec{e_1} + v^2 \vec{e_2}\right) \\
&= \left(v^1\right)^2 (\vec{e_1} \cdot \vec{e_1}) + 2 v^1 v^2 (\vec{e_1} \cdot \vec{e_2}) + \left(v^2\right)^2 (\vec{e_2} \cdot \vec{e_2}) \\
\\
&= \left(\tilde{v^1}\right)^2 \left(\tilde{\vec{e_1}} \cdot \tilde{\vec{e_1}}\right) + 2 \tilde{v^1} \tilde{v^2} \left(\tilde{\vec{e_1}} \cdot \tilde{\vec{e_2}}\right) + \left(\tilde{v^2}\right)^2 \left(\tilde{\vec{e_2}} \cdot \tilde{\vec{e_2}}\right)
\end{aligned}
$$

If we try applying this, in the orthonormal basis $\vec{e_i}$, it will simplify to the Pythagora's Theorem, whereas applying this to the new basis $\tilde{\vec{e_i}}$, you will see that, the coefficients obtained are not simply 1s and 0s, but are different numbers, that eventually allow the result to be consistent, and the length to be exactly the same computed in the old basis.

If you try it yourself, using the forward transform and getting the new basis in terms of the old, and computing the dot products, you will find that:

$$
\begin{aligned}
\tilde{\vec{e_1}} \cdot \tilde{\vec{e_1}} &= 5 \\
\tilde{\vec{e_1}} \cdot \tilde{\vec{e_2}} &= -3/4 \\
\tilde{\vec{e_2}} \cdot \tilde{\vec{e_2}} &= 5/16
\end{aligned}
$$

Now, the formula for the length using the dot product above, can be rewritten in matrix terms as:

$$
\begin{aligned}
\|\vec{v}\|^2 &= \vec{v} \cdot \vec{v} \\
&= \left(v^1\right)^2 + \left(v^2\right)^2 \\
&= 5\left(\tilde{v^1}\right)^2 + 2(-3/4) \tilde{v^1} \tilde{v^2} + (5/16)\left(\tilde{v^2}\right)^2 \\
\\
&= \begin{bmatrix} v^1 & v^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \\
&= \begin{bmatrix} \tilde{v^1} & \tilde{v^2} \end{bmatrix} \begin{bmatrix} 5 & -3/4 \\ -3/4 & 5/16 \end{bmatrix} \begin{bmatrix} \tilde{v^1} \\ \tilde{v^2} \end{bmatrix}
\end{aligned}
$$

So as you see, the key to getting the correct length, in any coordinate system, is that matrix in between you see there. **And that is what we call the metric tensor**

$$g_{\vec{e_i}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad\qquad g_{\tilde{\vec{e_i}}} = \begin{bmatrix} 5 & -3/4 \\ -3/4 & 5/16 \end{bmatrix}$$

This is a tensor, as it's invariant, but its components are different in different coordinate systems. And if you notice, at the end of the day, the components of the metric tensor in a specific basis, are the dot products between the basis vectors.

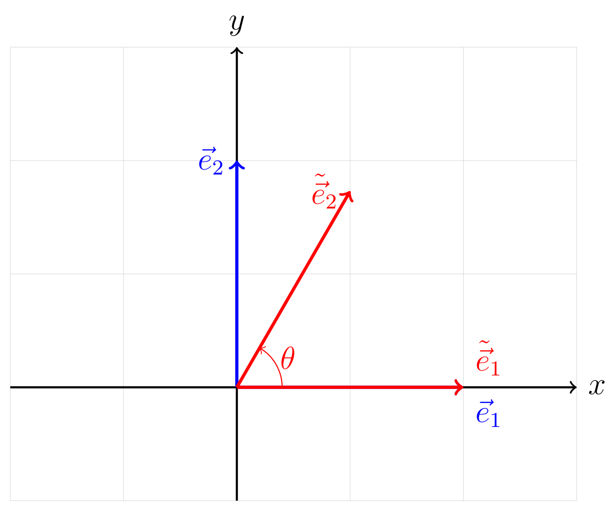
$$g_{\vec{e_i}} = \begin{bmatrix} \vec{e_1}\cdot\vec{e_1} & \vec{e_1}\cdot\vec{e_2} \\ \vec{e_2}\cdot\vec{e_1} & \vec{e_2}\cdot\vec{e_2} \end{bmatrix} \qquad\qquad g_{\tilde{\vec{e_i}}} = \begin{bmatrix} \tilde{\vec{e_1}}\cdot\tilde{\vec{e_1}} & \tilde{\vec{e_1}}\cdot\tilde{\vec{e_2}} \\ \tilde{\vec{e_2}}\cdot\tilde{\vec{e_1}} & \tilde{\vec{e_2}}\cdot\tilde{\vec{e_2}} \end{bmatrix}$$

So to sum up, and using the Einstein summation convention:

$$\begin{aligned} \|\vec{v}\|^2 &= v^i v^j \left(\vec{e_i}\cdot\vec{e_j}\right) = v^i v^j g_{ij} \\ &= \tilde{v^i}\tilde{v^j}\left(\tilde{\vec{e_i}}\cdot\tilde{\vec{e_j}}\right) = \tilde{v^i}\tilde{v^j}\tilde{g_{ij}} \end{aligned}$$

$$\boxed{\begin{aligned} g_{ij} &= \left(\vec{e_i}\cdot\vec{e_j}\right) && (55a) \\ \tilde{g_{ij}} &= \left(\tilde{\vec{e_i}}\cdot\tilde{\vec{e_j}}\right) && (55b) \end{aligned}}$$

The metric tensor is actually able to not only get us the length of vectors, but also measure angles. Let's see how that works, imagine a new basis wrt to an old one, with just one basis vector rotated by an angle θ:



Here, the forward transformation from old to new can be easily found:

$$\tilde{\vec{e_1}} = \vec{e_1}$$
$$\tilde{\vec{e_2}} = \cos(\theta)\vec{e_1} + \sin(\theta)\vec{e_2}$$

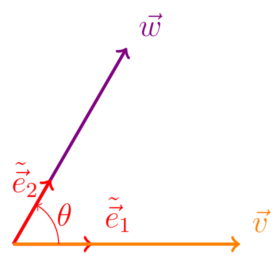
Let's compute the basis dot products in the new basis then:

$$\begin{aligned}
\tilde{\vec{e_1}} \cdot \tilde{\vec{e_1}} &= \vec{e_1} \cdot \vec{e_1} = 1 \\
\tilde{\vec{e_1}} \cdot \tilde{\vec{e_2}} &= \vec{e_1} \cdot (\cos(\theta)\vec{e_1} + \sin(\theta)\vec{e_2}) \\
&= \cos(\theta)\left(\vec{e_1} \cdot \vec{e_1}\right) + \sin(\theta)\left(\vec{e_1} \cdot \vec{e_2}\right) = \cos(\theta) \\
\tilde{\vec{e_2}} \cdot \tilde{\vec{e_2}} &= (\cos(\theta)\vec{e_1} + \sin(\theta)\vec{e_2}) \cdot (\cos(\theta)\vec{e_1} + \sin(\theta)\vec{e_2}) \\
&= \cos(\theta)^2\left(\vec{e_1} \cdot \vec{e_1}\right) + \sin(\theta)^2\left(\vec{e_2} \cdot \vec{e_2}\right) + \left(2\cos(\theta)\sin(\theta)\right)\left(\vec{e_1} \cdot \vec{e_2}\right) \\
&= \cos(\theta)^2 + \sin(\theta)^2 = 1
\end{aligned}$$

So the metric tensors in both basis are:

$$g_{\vec{e_i}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad\qquad g_{\tilde{\vec{e_i}}} = \begin{bmatrix} 1 & \cos(\theta) \\ \cos(\theta) & 1 \end{bmatrix}$$

Allright, now we can imagine having two vectors $\vec{w}$ and $\vec{v}$, for which we want to find the angle θ in between. We can define a new basis vector couple, where the basis vectors are pointing in the same direction as the two vectors, except for their length be unitary and then we can compute the dot product between two two vectors:



Since we chose such basis vectors, basically $\vec{w}$ is a scaled $\tilde{\vec{e_1}}$ and $\vec{v}$ is a scaled $\tilde{\vec{e_2}}$ so we can basically write something like:

$$\begin{aligned}
(\vec{v} \cdot \vec{w}) &= \left(a\tilde{\vec{e_1}}\right) \cdot \left(a\tilde{\vec{e_2}}\right) \\
&= ab\left(\tilde{\vec{e_1}} \cdot \tilde{\vec{e_2}}\right) \\
&= ab\cos(\theta)
\end{aligned}$$

Now, since a is just the length of $\vec{v}$ and b the length of $\vec{w}$:

$$\frac{(\vec{v} \cdot \vec{w})}{\|\vec{v}\|\|\vec{w}\|} = \cos(\theta) \tag{56}$$

References

[1] eigenchris, *Tensors for Beginners*, YouTube video series,
https://www.youtube.com/playlist?list=PLJHszsWbB6hrkmmq57lX8BV-o-YIOFsiG