Part I

Background Topics

1 Linear Algebra

1.1 *Preface*

The goal of this chapter is to review a few of the more advanced topics in linear algebra, which are important both for learning the other background topics, as well as the actual topic of spinors.

For example, the topic of dual vectors can help with understanding the what the Dira/bra-ket notation actually means, and cement a deeper understanding of the basic ideas of quantum physics. The topic of absrtact vector spaces provides a good foundation for topics such as geometric- and abstract algebra.

More to be written...

1.2 *Change of Coordinates*

In introductory linear algebra courses you should have learned about change of coordinate systems: a coordinate system is just another name for a basis set of whatever vector space is used (in this section it's $\mathbb{R}^n$). A change of coordinate system is the transformation of vectors from being represented in one basis set $B = \{e_1, e_2, \ldots, e_n\}$ to being represented in another basis set $\tilde{B} = \{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\}$. Since such transformations are linear they are commonly represented in a matrix form.

In this section we will discuss *how* vectors and their components transform under change of basis sets. These transformations can be quite confusing, so I color coded the equations consistently as a visual guide. In addition, the $\mathbb{R}^2$ case is intoduced first, before giving the generalized form for $\mathbb{R}^n$. [1]

[1] This section is heavily based, with permission, on Eigenchris' video series Tensors for Beginners.

Before we do this, there is one important idea to understand: vectors themselves do not change under change of basis - their magnitude and orientation stay the same no matter how *we* measure them. The component of a vector are a way to describe (or "measure") the vector using a certain basis set, so it should not be surprising that they change under change of basis.

With that in mind, let us look at an example of a basis set change in $\mathbb{R}^2$.

1.2.1 *Change of basis set in* $\mathbb{R}^2$

Suppose we use the standard basis set to represent $\mathbb{R}^2$:

$$B = \{e_1, e_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \tag{1.1}$$

and we want to change our coordinate system to use the following basis set:

$$\tilde{B} = \{\tilde{e}_1, \tilde{e}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{4} \end{bmatrix} \right\}. \tag{1.2}$$

(the vectors composing the two basis sets are shown in Figure 1.1)

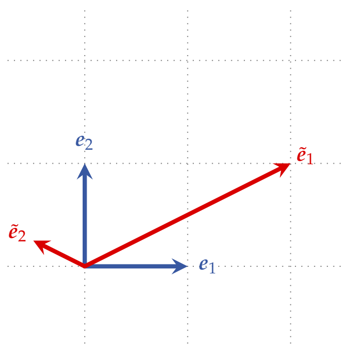

Figure 1.1: The standard basis set B and a new basis set $\tilde{B}$ shown together.

The transformation between the basis sets can be represented in matrix form:

$$\begin{array}{cc} & \begin{array}{cc} \tilde{e}_1 & \tilde{e}_2 \\ \downarrow & \downarrow \end{array} \\ F = & \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} \end{array}. \tag{1.3}$$

Note 1.1 The components of F

The columns of $\boldsymbol{F}$ are exactly the two vectors of $\tilde{B}$ because B is the standard basis. The transformation matrix is not that obvious between any two basis sets.

!

Now, if we want to transform e_1 and e_2 into $\tilde{e}_1$ and $\tilde{e}_2$ using $\boldsymbol{F}$, we simply multiply each of the vectors by $\boldsymbol{F}$:

$$\tilde{e}_1 = \boldsymbol{F}e_1,$$
$$\tilde{e}_2 = \boldsymbol{F}e_2.$$

To make Equation 1.4 more concise, we can collect the two vectors into a matrix disguised as a row vecor:

$$e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e_1, e_2 \end{bmatrix}, \tag{1.4}$$

and find that Equation 1.4 can be written in vector-matrix notation as

$$\tilde{e} = \begin{bmatrix} \tilde{e}_1, & \tilde{e}_2 \end{bmatrix} = \begin{bmatrix} e_1, & e_2 \end{bmatrix} \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} 2e_1 + e_2, & -\frac{1}{2}e_1 + \frac{1}{4}e_2 \end{bmatrix}. \tag{1.5}$$

The reverse transformation can be calculated by applying the inverse transformation $\boldsymbol{F^{-1}}$ on Equation 1.5:

$$\begin{aligned} \begin{bmatrix} \tilde{e}_1, & \tilde{e}_2 \end{bmatrix} \boldsymbol{F^{-1}} &= \left(\begin{bmatrix} e_1, & e_2 \end{bmatrix} \boldsymbol{F} \right) \boldsymbol{F^{-1}} \\ &= \begin{bmatrix} e_1, & e_2 \end{bmatrix} \boldsymbol{F}\boldsymbol{F^{-1}} \\ &= \begin{bmatrix} e_1, & e_2 \end{bmatrix} \boldsymbol{I}_2 \\ &= \begin{bmatrix} e_1, & e_2 \end{bmatrix}. \end{aligned}$$

(where in our case $\boldsymbol{F^{-1}} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ -1 & 2 \end{bmatrix}$)

To summarize: given a set B of basis vectors in $\mathbb{R}^2$, we can transform it into the basis vectors set $\tilde{B}$ by using the forward transformation $\boldsymbol{F} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$. To get back the original basis vectors from the transformed vectors we use the inverse transfomation $\boldsymbol{F^{-1}} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}^{-1}$ (if it exists - i.e. if $\boldsymbol{F}$ is invertible).

1.2.2 *The more general case: $\mathbb{R}^n$*

In $\mathbb{R}^n$ the transformations behave in a similar way: given the transformation rule that each new basis vector $\tilde{e}_i \in \tilde{B}$ is a linear combination of the old basis vector set B, i.e.

$$\begin{aligned}
\tilde{e}_1 &= F_{11}e_1 + F_{12}e_2 + \cdots + F_{1n}e_n, \\
\tilde{e}_2 &= F_{21}e_1 + F_{22}e_2 + \cdots + F_{2n}e_n, \\
\vdots &= \vdots \\
\tilde{e}_n &= F_{n1}e_1 + F_{n2}e_2 + \cdots + F_{nn}e_n,
\end{aligned}$$

we can write the transformation in matrix form as

$$\left[\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\right] = \left[e_1, e_2, \ldots, e_n\right] \begin{bmatrix} F_{11} & F_{12} & \ldots & F_{1n} \\ F_{21} & F_{22} & \ldots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \ldots & F_{nn} \end{bmatrix}. \tag{1.6}$$

Equation 1.6 can be written in index notation, which shows us how each vector is transformed:

$$\tilde{e}_j = \sum_{k=1}^{n} F_{kj} e_k. \tag{1.7}$$

Similarly, the inverse operation is given by

$$\begin{aligned}
e_1 &= F_{11}^{-1}\tilde{e}_1 + F_{12}^{-1}\tilde{e}_2 + \cdots + F_{1n}^{-1}\tilde{e}_n, \\
e_2 &= F_{21}^{-1}\tilde{e}_1 + F_{22}^{-1}\tilde{e}_2 + \cdots + F_{2n}^{-1}\tilde{e}_n, \\
\vdots &= \vdots \\
e_n &= F_{n1}^{-1}\tilde{e}_1 + F_{n2}^{-1}\tilde{e}_2 + \cdots + F_{nn}^{-1}\tilde{e}_n,
\end{aligned}$$

with the matrix form being

$$\left[e_1, e_2, \ldots, e_n\right] = \left[\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\right] \begin{bmatrix} F_{11}^{-1} & F_{12}^{-1} & \ldots & F_{1n}^{-1} \\ F_{21}^{-1} & F_{22}^{-1} & \ldots & F_{2n}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1}^{-1} & F_{n2}^{-1} & \ldots & F_{nn}^{-1} \end{bmatrix}, \tag{1.8}$$

and the index notation being

$$e_i = \sum_{j=1}^{n} F_{ji}^{-1} \tilde{e}_j. \tag{1.9}$$

By subtituting Equation 1.7 into Equation 1.9, we get

$$\begin{aligned}
e_i &= \sum_{j=1}^{n} F_{ji}^{-1} \tilde{e}_j \\
&= \sum_{j=1}^{n} F_{ji}^{-1} \left(\sum_{k=1}^{n} F_{kj} e_k \right) \\
&= \sum_{k=1}^{n} \underbrace{\left(\sum_{j=1}^{n} F_{kj} F_{ji}^{-1} \right)}_{\text{this is just a number!}} e_k.
\end{aligned} \tag{1.10}$$

Now, Equation 1.10 simply tells us something we already know: each basis vector e_i is equal to a linear combination of the *same set* of basis vectors. This must mean that for $k = i$ the number in paranthesis is one, and for any other value of k it is zero - i.e. it equals δ_{ik}:

$$e_i = \sum_{k=1}^{n} \delta_{ik} e_k. \tag{1.11}$$

In turn, Equation 1.11 means that the matrix $\boldsymbol{F}\boldsymbol{F}^{-1} = \boldsymbol{I}_n$, the identity matrix in $\mathbb{R}^n$ - and thus the matrices $\boldsymbol{F}$ and $\boldsymbol{F}^{-1}$ are eachother's inverses, exactly as we expect.

1.2.3 *Components transformation*

Now let us discuss how the *components* of a vector transform under change of basis: suppose we have the vector v as depicted in Figure 1.2. Using the standard basis set (depicted in blue) we can write $v = e_1 + e_2$, or simply $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. However, by using the basis set $\tilde{B} = \{2e_1; e_2\}$ (i.e. e_1 is scaled by 2), v is now $v = \frac{1}{2}\tilde{e}_1 + \tilde{e}_2 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$.

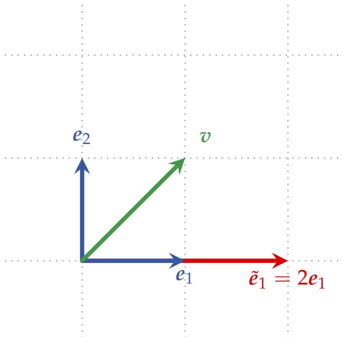

Figure 1.2: A vector v and two different sets of basis vectors: in blue the standard basis vectors and in red the basis set $\tilde{B} = \{2e_1, e_2\}$.

We see that by scaling a basis vector by 2, its repspective component is scaled by $\frac{1}{2}$. In general, when we scale the basis vector i by a non-zero scalar α, the respective i-th component is scaled by $\frac{1}{\alpha}$.

What about rotating the basis set? Again, we take the vector from Figure 1.2, but this time we rotate the standard basis set by some angle θ (Figure 1.3). The result is that the vector v is rotated by $-\theta$ with respect to the new basis set. If you don't see this, look at the angle between v and e_1: it is exactly 45° (since $v = e_1 + e_2$). The angle between v and $\tilde{e}_1$, on the other hand, is smaller than 45° - which means that v got relatively rotated *towards* it. This is the opposite direction of the rotation from e_1 to $\tilde{e}_1$.

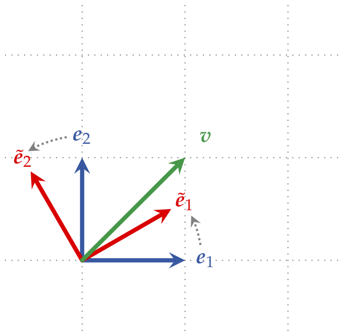

Figure 1.3: A vector v and two different sets of basis vectors: in blue the standard basis vectors and in red the basis set $\tilde{B} = \{2e_1, e_2\}$.

It seems that vector coordinates transform *counter* to the change in basis. To be sure, let us take a look at the most general case. Given any vector v we can express it in a basis set $B = \{e_1, e_2, \ldots, e_n\}$ using the respective components v_i:

$$v = \sum_{i=1}^{n} v_i e_i. \tag{1.12}$$

In some other basis $\tilde{B} = \{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\}$ it has other components which we denote as $\tilde{v}_i$:

$$v = \sum_{j=1}^{n} \tilde{v}_j \tilde{e}_j. \tag{1.13}$$

We can subtitute into Equation 1.12 the explicit form of the vectors in B (Equation 1.9):

$$\begin{aligned} v &= \sum_{i=1}^{n} v_i e_i \\ &= \sum_{i=1}^{n} v_i \left(\sum_{j=1}^{n} F^{-1}_{ji} \tilde{e}_j \right) \\ &= \sum_{i=1}^{n} \underbrace{\sum_{j=1}^{n} \left(F^{-1}_{ji} v_i \right)}_{\tilde{v}_i} \tilde{e}_i . \end{aligned}$$

Comparing the last equality in Equation 1.14 to Equation 1.12, we see that under a basis change the components of v go from v_i to $\tilde{v}_i$ via the *inverse transformation* F^{-1}. This is in agreement with what we saw in the specific example: the inverse transformation of scaling is scaling by the inverse, and the inverse transformation of rotation by an angle is rotation by the opposite angle.

Due to this behaviour we say that vectors are *contravarient*, and sometimes even refer to them as *contravarient vectors*[2]

[2] this is a hint that there is at least another type of vectors: those that transform together with a change in basis, or *covarient vectors*. Indeed, we will meet these "co-vectors" soon.

1.2.4 *Index position and the Einstein summation convention*

To always remember that vector components are contravarient in regards to change of basis, we will from now on denote them using a superscript instead of a subscript:

$$v = \sum_{i=1}^{n} v^i e_i . \tag{1.14}$$

In the case of a generic vector in $\mathbb{R}^n$ the upper-index notation translates into writing its explicit column-component form as follows:

$$v = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} \tag{1.15}$$

It is important to always keep in mind that using the upper-index notation for vector components means that powers have to be more explicitely written, as to not confuse the reader. For example, the standard L_2 norm is written as

$$\|v\| = \sqrt{\left(v^1\right)^2 + \left(v^2\right)^2 + \cdots + \left(v^n\right)^2} . \tag{1.16}$$

While this can seem as unnessecary complication now, it will soon become clear why it is needed - especially when we use multiple components of different types (some contravarient and some covariant).

In addition to the upper index notation for vector components, this book (and almost all relevant texts) use the *Einstein summation convention*: where an index appears twice in a single term - once as an upper index and once as a lower index (the order does not matter), we should view the term as having a sum infront of it. This sum is done over all relevant values of the index in question (unless otherwise stated). Below are some examples of using this convention.

Example 1.1 Einstein summation convention #1

The standard way of writing a vector explicitely in the basis set $B = \{e_1, e_2, \dots, e_n\}$ is

$$v = \sum_{i=1}^{n} v^i e_i. \tag{1.17}$$

Using Einstein's convention, this becomes simply

$$v = v^i e_i, \tag{1.18}$$

where from context we know that $i \in \{1, 2, \dots, n\}$.

Example 1.2 Einstein summation convention #2

The inner product of two vectors v, w can be written explicitely as

$$\langle v, w \rangle = \sum_{i=1}^{n} v^i w^i. \tag{1.19}$$

(note that the upper-index notation is used here)
However, if we consider the vectors as being matrices, the first being an $1 \times n$ matrix and the second an $n \times 1$ matrix, we can transpose the first vector from standard column form

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

into the row form:

$$v^\top = \begin{bmatrix} v_1, & v_2, & \dots, & v_n \end{bmatrix}.$$

(note that the components of v as a row-vector are in lower-index notation. The reason for this will become clearer in the next section) Indeed, the matrix product of a $1 \times n$ matrix with an $n \times 1$ matrix has dimension of $1 \times 1 = 1$, just like a scalar.
Using Einstein's convention, we can write the inner product

continues in the next page

continues from the previous page

as

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \begin{bmatrix} v_1, & v_2, & \ldots, & v_n \end{bmatrix} \begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{bmatrix} = v_i w^i. \tag{1.20}$$

where, again, we know that $i \in \{1, 2, \ldots, n\}$ from the context.

Example 1.3 Einstein summation convention #3

The product of an $m \times n$ matrix A and a vector $\boldsymbol{v}$ can be written in component form as

$$(A\boldsymbol{v})^i = \sum_{j=1}^{n} A_{ij} v^j. \tag{1.21}$$

. . . FINISH WRITING THIS CONFUSING PART

1.3 Dual Vectors and Dual Spaces

1.3.1 Linear measurements and rulers (or: why do we care about dual vectors?)

Frequently in linear algebra we want to measure vectors. A measure in this context is a way to assign each vector a real number which somehow reflects its properties (i.e. direction and/or magnitude). Mathematically, we are looking for a function $\phi : \mathbb{R}^n \to \mathbb{R}$ - we feed it a vector, and it returns some measurement.

One common way to measure vectors is using the norm: the norm of a vector in $\mathbb{R}^n$ using the standard basis set $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ is given by

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2 + \cdots + v_n^2}, \tag{1.22}$$

where the numbers $v_1, v_2, \ldots, v_n$ are the components in each of the respective basis vectors.

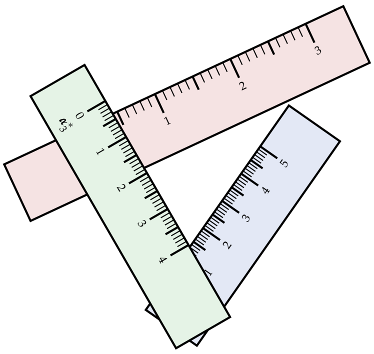

Figure 1.4: Three rulers, $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*, \boldsymbol{\alpha}_3^*$, each with its own orientation and frequency of graduation lines.

While the norm does tell us something useful about a vector, it has two drawbacks: it doesn't tell us anything about the vector's orientation in space, and even worse - it is not linear. We can try and simplify the calculation of the norm by dropping the square root and calculating the *square norm*, but that too isn't linear, nor does it tell us anything about the vector's orientation.

Another way of measuring vectors is by using *rulers*. Rulers are nothing more than a set of graduation lines with an orientation in space (Figure 1.4). We can therefore represent a ruler using a vector: the magnitude of the vector is the frequency of the graduation lines, while its direction is the direction of the ruler (which is orthogonal to the graduation lines). To not confuse ruler vectors with regular vectors, I will denote ruler vectors using a greek letter and an asterisk, e.g. $\boldsymbol{\alpha}^*$.

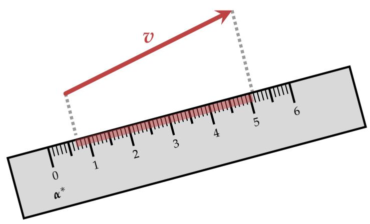

Figure 1.5: Measuring a vector using a ruler by projecting the vector onto the ruler. The result of the projection is drawn as a red line on the ruler.

In real life we usually rotate rulers to align with the orientation of the magnitude we wish to measure. However, in our case we want the rulers to also measure orientation (otherwise why not just use the norm?). So instead of rotating the ruler to align with a vector we wish to measure, we *project* the vector on the ruler and then take our measurement (Figure 1.5).

Since projection in linear algebra is calculated via the inner product, the measurement m of a vector $\boldsymbol{v}$ using a ruler $\boldsymbol{\alpha}^*$ is given by

$$m = \langle \boldsymbol{\alpha}^*, \boldsymbol{v} \rangle. \tag{1.23}$$

However, since we can represent a ruler using a vector, we can use the vector representation in Equation 1.23. And in order to distinguish ruler vectors from regular vectors written in an excplicit form, we write the ruler vectors as row vectors:

$$\begin{aligned}\langle \boldsymbol{\alpha}^*, \boldsymbol{v}\rangle &= \begin{bmatrix}\alpha_1, & \alpha_2, & \dots, & \alpha_n\end{bmatrix}\begin{bmatrix}v^1 \\ v^2 \\ \vdots \\ v^n\end{bmatrix} \\ &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n.\end{aligned} \tag{1.24}$$

Another reason for representing rulers as row vectors is that it corresponds well to matrix multiplication: if we regard a row vector as a $1 \times n$ matrix, and a column vector as an $n \times 1$ matrix, the resulting product is a 1×1 matrix, which is equivalent to a scalar. This will become handy when we discuss higher-order rulers (which in turn can handle higher-order "vectors").

A convinient visualization for rulers in $\mathbb{R}^2$ is their representation as a set of parallel lines which are the extensions of the ruler's graduation marks. We draw the first line going through the origin and orthogonal to the orientation of the vector representation of the ruler. Then, we draw a line for each integer multiple of the frequency of the ruler's graduation marks. This results in a set of evenly spaced and parllel lines, which fit precisely with the main graduation mark of the ruler they represent (Figure 1.6).

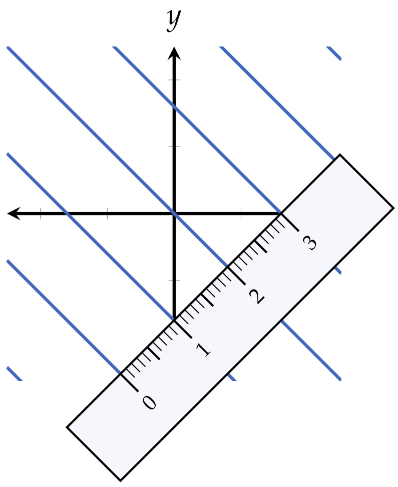

Figure 1.6: The graphical representation of a ruler shown next to the ruler, such that the infinite set of lines drawn match the ruler's graduation marks.

In $\mathbb{R}^3$ a ruler is represented in a similar way, but with *planes* instead of lines (Figure 1.7). And in general, the rulers for a space $\mathbb{R}^n$ can be represented using an infinite set of $(n-1)$-dimensional *hyperplanes*.

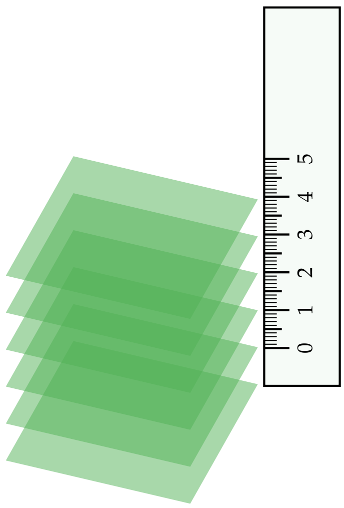

Figure 1.7: Representation of rulers in $\mathbb{R}^3$ as planes.

From Equation 1.25 we can see that any ruler $\boldsymbol{\alpha}^*$ can be written as the left half of an inner product, i.e. $\langle \boldsymbol{\alpha}^*, \cdot \rangle$. This hints that the ruler is *acting* on vectors, yielding a scalar. Looking at the explicit form of a ruler acting on a vector, we can see that not only do rulers represent linear functions, but they in-fact represent *all* the possible *linear* functions of the type $\phi : \mathbb{R}^n \to \mathbb{R}$. Such functions are called *linear forms*, or more specifically in this case a *1-form*.

Since rulers can be represented as vectors but aren't exactly the same as regular vectors, they are also called *dual vectors*, which is the name I will use for them from now on. For any vector space $\mathbb{R}^n$, the dual vectors form a vector space of their own, called the *dual space to* $\mathbb{R}^n$.

The linearity of the dual vectors space simplifies many calculations - foremost, it allows us to construct dual vectors from other dual vectors using linear combinations. See 1.4 below.

Example 1.4 Linear combination of dual vectors

Suppose we have two dual vectors

$$\begin{aligned} \boldsymbol{\alpha}^* &= \begin{bmatrix} 1, & -4, & 0 \end{bmatrix}, \\ \boldsymbol{\beta}^* &= \begin{bmatrix} 2, & 1, & 3 \end{bmatrix}. \end{aligned} \tag{1.25}$$

We can use $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ to create a third dual vector γ^*:

$$\begin{aligned} \gamma^* &= 5\begin{bmatrix} 1, & -4, & 0 \end{bmatrix} + 2\begin{bmatrix} 2, & 1, & 3 \end{bmatrix} \\ &= \begin{bmatrix} 5+4, & -20+2, & 0+6 \end{bmatrix} \\ &= \begin{bmatrix} 9, & -18, & 6 \end{bmatrix}. \end{aligned} \tag{1.26}$$

The application of γ^* on any vector $\boldsymbol{v} \in \mathbb{R}^3$ is then given by

$$\langle \gamma^*, \boldsymbol{v} \rangle = 9v_1 - 18v_2 + 6v_3, \tag{1.27}$$

which is exactly the application of $\boldsymbol{\alpha}^*$ on the vector five times followed by the application of $\boldsymbol{\beta}^*$ twice (try it for yourself!).

1.3.2 *Basis sets and coordinate transformations*

In 1.4 I wrote dual vectors as coordinates in a row form. Since the dual space $\mathbb{R}^{n*}$ is an n-dimensional real vector space, it has infinitely many basis sets, each composed of n vectors. We can use any basis set $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ in $\mathbb{R}^n$ to create a dual basis set $B^* = \left\{\boldsymbol{\epsilon}^1, \boldsymbol{\epsilon}^2, \dots, \boldsymbol{\epsilon}^n\right\}$ in $\mathbb{R}^{n*}$: since dual vectors are linear maps, we simply define the vectors in B^* as follows:

$$\langle \boldsymbol{\epsilon}^i, \boldsymbol{e}_j \rangle = \delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{1.28}$$

Note 1.2 Dual basis vectors notation

To distinguish dual basis vectors from regular basis vectors, I denote them using a greek letter (since they represent a linear maps) white also using upper-index notation. The latter will be consistent with their propeties as we shall see soon.

!

Some examples of calculating dual basis vectors:

Example 1.5 The standard dual basis set

Given the standard basis set in $\mathbb{R}^3$,

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$$

continues in the next page

continues from the previous page

its corresponding dual is then

$$B^* = \left\{\epsilon^1, \epsilon^2, \epsilon^3\right\}.$$

For ϵ^1 the following is true:

$$\langle \epsilon^1, e_1 \rangle = 1, \quad \langle \epsilon^1, e_2 \rangle = 0, \quad \langle \epsilon^1, e_3 \rangle = 0.$$

If we write $\epsilon^1 = \begin{bmatrix} \alpha, & \beta, & \gamma \end{bmatrix}$, then we get that the first equality gives

$$1 = \langle \epsilon^1, e_1 \rangle = \begin{bmatrix} \alpha, & \beta, & \gamma \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \alpha,$$

i.e. $\alpha = 1$. Similarly, the second equation gives

$$0 = \langle \epsilon^1, e_2 \rangle = \begin{bmatrix} \alpha, & \beta, & \gamma \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \beta,$$

i.e. $\beta = 0$. The third equation then gives

$$0 = \langle \epsilon^1, e_3 \rangle = \begin{bmatrix} \alpha, & \beta, & \gamma \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \gamma,$$

i.e. $\gamma = 0$. Altogether we get $\epsilon^1 = \begin{bmatrix} 1, & 0, & 0 \end{bmatrix}$.
Similar calculations for ϵ^2 and ϵ^3 show that $\epsilon^2 = \begin{bmatrix} 0, & 1, & 0 \end{bmatrix}$ and $\epsilon^3 = \begin{bmatrix} 0, & 0, & 1 \end{bmatrix}$.
In $\mathbb{R}^n$ we get a similar result: the i-th dual basis vector is simply the row form of the i-th basis vector.
This is however <u>not</u> true for all dual basis sets, as we will see in the next example.

Example 1.6 Dual basis set in $\mathbb{R}^2$

Consider the following basis set in $\mathbb{R}^2$:

$$B = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix} \right\}.$$

To find its dual,

$$B^* = \left\{ \begin{bmatrix} \alpha, & \beta \end{bmatrix}, \begin{bmatrix} c, & d \end{bmatrix} \right\},$$

we simply solve the following four equations:

$$\langle \epsilon^1, e_1 \rangle = \begin{bmatrix} \alpha, & \beta \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2\alpha - \beta = 1,$$

continues in the next page

continues from the previous page

$$\langle \epsilon^1, e_2 \rangle = \begin{bmatrix} \alpha, & \beta \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = -\alpha + 3\beta = 0,$$

$$\langle \epsilon^2, e_1 \rangle = \begin{bmatrix} c, & d \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2c - d = 0,$$

$$\langle \epsilon^2, e_2 \rangle = \begin{bmatrix} c, & d \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = -c + 3d = 1.$$

which gives $\alpha = 0.6, \beta = 0.2, c = 0.2, d = 0.4$, and thus

$$\epsilon^1 = \begin{bmatrix} 0.6, & 0.2 \end{bmatrix},$$
$$\epsilon^2 = \begin{bmatrix} 0.2, & 0.4 \end{bmatrix}.$$

Note that $\epsilon^1 \neq e_1^\top$ and $\epsilon^2 \neq e_2^\top$! This is important, and we will discuss this thoroughly in the following paragraphs.

In the most general case (i.e. in $\mathbb{R}^n$), a basis set has n vectors, and Equation 1.28 has the following form:

$$\begin{aligned} &\langle \epsilon^1, e_1 \rangle = 1,\ \langle \epsilon^1, e_2 \rangle = 0,\ \dots,\ \langle \epsilon^1, e_n \rangle = 0. \\ &\langle \epsilon^1, e_1 \rangle = 0,\ \langle \epsilon^2, e_2 \rangle = 1,\ \dots,\ \langle \epsilon^2, e_n \rangle = 0. \\ &\quad\vdots \\ &\langle \epsilon^n, e_1 \rangle = 0,\ \langle \epsilon^n, e_2 \rangle = 0,\ \dots,\ \langle \epsilon^n, e_n \rangle = 1. \end{aligned} \tag{1.29}$$

This above equation is begging to be represented as a matrix equation: we collect the *dual* basis vectors $\left\{\epsilon^i\right\}$ as **rows** in an $n \times n$ matrix E^*. The *regular* basis vectors $\left\{e_j\right\}$ are collected as **columns** of another $n \times n$ matrix E. Then, Equation 1.29 can be written as

$$E^* E = \mathbb{I}_n, \tag{1.30}$$

From the matrix equation it is clear that $E^* = E^{-1}$, and since E must have a non-zero determinant[3], we are guaranteed to find E^*. This means that for any basis set in $\mathbb{R}^n$ the dual basis exists and is unique.

[3] if you're not sure why, I strongly recommend reviewing basic linear algebra.

Example 1.7 Dual set in $\mathbb{R}^3$

The following basis set

$$B = \left\{ \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \right\}$$

can be represented as the matrix

$$E = \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & -2 \\ -1 & 1 & 2 \end{bmatrix}.$$

continues in the next page

continues from the previous page

The dual vectors can then be calculated from the inverse of E:

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

$$\Rightarrow B^* = \left\{ \begin{bmatrix} -2 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 4 & -1 & -3 \end{bmatrix}, \begin{bmatrix} -3 & 1 & 3 \end{bmatrix} \right\}.$$

(I found the inverse using python, but feel free to do the calculation for yourself as an exercise!)

Indeed, for example, $\langle \epsilon^1, e_1 \rangle = -2 \cdot 0 + 1 \cdot 3 - 1 \cdot 2 = 1$, and $\langle \epsilon^3, e_2 \rangle = -3 \cdot 1 + 1 \cdot 0 + 3 \cdot 1 = 0$ as expected.

In section 1.2 we discussed how vectors, basis vectors and vector components behave under change of basis. Now let us discuss how dual vectors, dual basis vectors and dualvector components behave under the same change. When we transform our representation of vectors from one basis set B to a new basis set $\tilde{B}$ we also generate a new basis for the dual space, given by Equation 1.30. I will denote the dual basis of B as B^* and the dual of $\tilde{B}$ as $\tilde{B}^*$. Let's define B and $\tilde{B}$ with actual numbers in $\mathbb{R}^3$ so it is easier to follow the operations:

$$\begin{aligned} B &= \mathbb{I}_3, \\ \tilde{B} &= \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & -2 \\ -1 & 1 & 2 \end{bmatrix}. \end{aligned} \tag{1.31}$$

(i.e. B is the standard basis in for $\mathbb{R}^3$ and $\tilde{B}$ is the basis set described in 1.7)

In this case, the transformation matrix $\boldsymbol{F}$ is simply $\tilde{B}$, since $\mathbb{I}_3\tilde{B} = \tilde{B}$. Similarly, the dual basis is represented by $\mathbb{I}_3^{-1} = \mathbb{I}_3$, and thus the inverse transformation matrix is simply $\boldsymbol{F}^{-1} = \tilde{B}^*$, since then $\mathbb{I}_3\tilde{B}^* = \tilde{B}^*$. This already shows us the first transformation property of dual vectors in a change of basis: the dual basis vectors themselves transform via the inverse transformation $\boldsymbol{F}^{-1}$ (recall that the regular basis vectors transform with the "forward" transformation matrix $\boldsymbol{F}$).

Now we understand why the upper-index notation is used for dual basis vectors: this is used since they transform in a *contravariant* fashion to the change in regular basis vectors. Accordingly, we say that the regualr basis vectors transform in a *contravariant* fashion, and indeed they are denoted using lower-index notation. For this reason, dual vectors are also sometimes called *covectors*.

At this point, I will introduce back Einstein's index convention. Using the convention, all *covariant* components of any object we are dealing with are denoted using lower-indices, while all *contravariant* components of any such object are denoted using upper-indices. When we see an index repeating twice in a single term - once as an upper-index and once as a lower-index (the order does not matter) - we sum over that index.

Table ?? shows the transformation rules for regualr basis vectors vs. dual basis vectors - and their respective components. In each equation you should find the repeating index, and imagine that there is a summation sign over than index before the transformation matrix. For example, the full regular basis vector transformation is $\tilde{\boldsymbol{e}}_j = \sum_{i=1}^{n} F^i{}_j \boldsymbol{e}_i$. Note: the inverse transformation $\boldsymbol{F}^{-1}$ is denoted in component form as $B^i{}_j$ to avoid the confusion arising from the use of the inverse notation, which is also upper-index.

Table 1.1: Transformation rules for regular basis vectors, dual basis vectors and their respective components. Note: $\boldsymbol{B} = \boldsymbol{F}^{-1}$.

	Regular		Dual	
	Form	Equation	Form	Equation
Basis vectors	Co	$\tilde{\boldsymbol{e}}_j = F^i{}_j \boldsymbol{e}_i$	Contra	$\boldsymbol{\epsilon}^i = B^i{}_j \boldsymbol{\epsilon}^j$
Components	Contra	$v^i = B^i{}_j v^j$	Co	$\alpha_j = F^i{}_j \alpha_i$

This kind of implicit summation is called a *contraction*. The transformation matrices $F^i{}_j$ and $B^i{}_j$ are in fact *tensors* of the type $(1, 1)$: they each have a single contravariant component (upper-index) and a single covariant component (lower-index).

1.3.3 *Dual vectors in* $\mathbb{C}^n$ *and the Dirac notation*

In physics, and especially quantum physics, complex vector spaces are often used. For example, consider the space in which quantum states exist: $\mathbb{C}^n$, where n can be any natural number and even infinity. The standard inner product in $\mathbb{C}^n$ (which we will discuss momentarily) gives rise to a natural norm and thus also a distance function:

$$\mathrm{D}\left(\boldsymbol{x}, \boldsymbol{y}\right) = \left\|\boldsymbol{x} - \boldsymbol{y}\right\| = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle. \tag{1.32}$$

Therefore, $\mathbb{C}^n$ together with the standard inner product is a metric space, and in fact a complete one - i.e. it is a *Hilbert space*. Naturally, we would like $\mathrm{D}\left(\boldsymbol{x}\right)$ to be always *real* and non-negative - otherwise it is quite strange to consider[4].

[4] what is the meaning of a negative, or worse - complex, distance?

To that end, and to stay consistent with the form of the inner product in $\mathbb{R}^n$, we define the standard inner product in $\mathbb{C}^n$ to use the *complex conjugate* in one of its arguments, in our case the first argument[5]:

[5] this is the so-called "physicist convention". The other, "mathematician convention" conjugates the second argument.

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} \overline{x}_i y_i. \tag{1.33}$$

This definition makes the induced norm to be always real and non-negative:

$$\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \sum_{i=1}^{n} \overline{x}_i x_i = \sum_{i=1}^{n} |x|^2 \geq 0. \tag{1.34}$$

We can rewrite Equation 1.33 in terms of a dual-vector/vector product:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \begin{bmatrix} \overline{x}_1, & \overline{x}_2, & \dots, & \overline{x}_n \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}. \tag{1.35}$$

If we then use a vertical line to separate the arguments of the inner product, i.e. writing it as $\langle \boldsymbol{x} | \boldsymbol{y} \rangle$, we can "separate" it into two parts:

$$\langle \boldsymbol{x} | = \begin{bmatrix} \overline{x}_1, & \overline{x}_2, & \dots, & \overline{x}_n \end{bmatrix} \quad \text{and} \quad | \boldsymbol{y} \rangle = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{bmatrix}. \tag{1.36}$$

Since the notation for the inner product uses *brackets*, we call $\langle \boldsymbol{x} |$ a *bra*, and $| \boldsymbol{y} \rangle$ a *ket* - and generally identify bras with dual vectors and kets with regular vectors. This notation is called the *Dirac notation*[6], named for Paul Dirac who introduced it. It is a surprisingly useful notation for vectors and dualvectors.

[6] and also, unsurprisngly, the *bra-ket notation*.

Since it is clear that bras and kets represent covectors and vectors, respectively, I will from now on drop the vector notation (bold letters) *within them*: $\langle a|, \langle x_2|, |u_i\rangle$, etc. Also, depending on context I will sometimes drop the letters all together in favour of just the indices: for example, if we are discussing some vectors $|v_1\rangle, |v_2\rangle, \dots, |v_i\rangle$, etc. - I might just write $|1\rangle, |2\rangle, \dots, |i\rangle$, etc.

Vector addition looks exactly as it does with the standard notation:

$$\begin{aligned} \boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c} &\Leftrightarrow |a\rangle + |b\rangle = |c\rangle, \\ \boldsymbol{\alpha}^* + \boldsymbol{\beta}^* = \boldsymbol{\gamma}^* &\Leftrightarrow \langle \alpha| + \langle \beta| = \langle \gamma|. \end{aligned} \tag{1.37}$$

And much like the fact that adding a vector and a dual vector does not make sense, adding a bra and a ket is undefined: $\xcancel{\langle \alpha| + |b\rangle}$. Sometimes, in order to declutter an equation, the addition can be written inside a bra or a ket:

$$\langle \alpha + \beta| \equiv \langle \alpha| + \langle \beta|. \tag{1.38}$$

Similarly, scaling dual vectors and vectors by a scalar looks the same as the usual notation, and sometimes this is also inserted into the respective bra/ket:

$$\begin{aligned} |2x\rangle &\equiv 2\,|x\rangle, \\ \langle -3\alpha| &\equiv -3\,\langle \alpha|. \end{aligned} \tag{1.39}$$

A nice example of using the notation is writing the definition of a dual basis set: ...

1.4 *Vector Spaces Beyond* $\mathbb{R}^n$

This section will cover the formal definition of vectors, so that the reader learns how to apply linear algebra ideas from $\mathbb{R}^n$ into any abstract vector space.

1.4.1 *Dual vector spaces*

Text text text.

Definition 1.1 Dual space

Given a vector space V over a field $\mathbb{F}$, its *dual space*, denoted V^*, is the set of all the linear functions $\phi : V \to \mathbb{F}$. The elements of V^* are called *dual vectors*.

π

A dual space equipped with a closed addition operation between any two of its elements and a closed product between its elements and the elements of the field $\mathbb{F}$ is itself a *vector space*, with similar structure to the space V.

Challenge 1.1 Dual space as a vector space

Show that a dual space euipped with addition and scalar product as defined above is indeed a vector space (Use the definition XXX).

?

To write:

1. Examples of dual vectors of functions?..

1.5 Further Reading

2 *Geometric Algebra*

2.1 Preface

This is a temp text.

3 Abstract Algebra

3.1 Preface

This is a temp text.

4 Lie Groups and Algebras

4.1 Preface

This is a temp text.

Part II

Spinors