

Part I

Background Topics

1 Linear Algebra

1.1 Preface

THE GOAL OF THIS CHAPTER is not to teach you, the reader, linear algebra from scratch - nor to be a thorough source of information on the topic. Rather, my aim is to introduce important “advanced” concepts for those who took a basic linear algebra course as part of an undergraduate university program. These concepts should help you gain a basic knowledge of the topics needed for understanding the rest of the background material, as well as the topic of spinors itself.

My approach to teaching topics in linear algebra - and in mathematics as a whole - is to first build an intuition and only then formalize and generalize the ideas as needed. In my personal experiences, when I was studying linear algebra I completely failed to understand it (and indeed, failed the course) until it “clicked” for me in regards to 2- and 3-dimensional real spaces, i.e. - visible geometry. After that I didn’t even have to study for exams anymore, as everything became clear enough to grasp and develop on the spot even during an exam (except for later, more advanced concepts). That is why, for example, I absolutely adore courses and study materials of the topic¹ which use animation, such as *3Blue1Brown* great video essay series [Essence of linear algebra](#)².

¹ And other mathematical topics as well.

² Temporary sidenote which should become a citation for the mentioned 3B1B video series

There are very few proofs in this chapter, and those that are shown are not completely rigorous. For more in-depth materials, see the last section (further read). With that out of the way - let’s begin!

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 $B = \{e_1, e_2, \dots, e_n\}$ to being represented in another basis set $\tilde{B} = \{\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_n\}$. 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 introduced first, before giving the generalized form for \mathbb{R}^n .³

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 :

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

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

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

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

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

$$F = \begin{matrix} \begin{matrix} \tilde{e}_1 & \tilde{e}_2 \\ \downarrow & \downarrow \end{matrix} \\ \begin{bmatrix} 2 & -\frac{1}{2} \\ 1 & \frac{1}{4} \end{bmatrix} \end{matrix}. \quad (1.3)$$

³ This section is heavily based, with permission, on Eigenchris' video series [Tensors for Beginners](#).

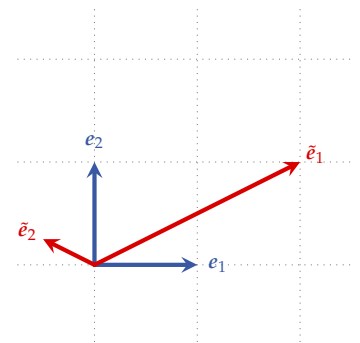


Figure 1.1: The standard basis set B and a new basis set \tilde{B} shown together.

Note 1.1 The components of F

The columns of 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 F , we simply multiply each of the vectors by F :

$$\tilde{e}_1 = F e_1,$$

$$\tilde{e}_2 = F e_2.$$

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

$$e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [e_1, e_2], \quad (1.4)$$

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

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

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

$$\begin{aligned} [\tilde{e}_1, \tilde{e}_2] F^{-1} &= ([e_1, e_2] F) F^{-1} \\ &= [e_1, e_2] F F^{-1} \\ &= [e_1, e_2] I_2 \\ &= [e_1, e_2]. \end{aligned}$$

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

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

$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$. To get back the original basis vectors

from the transformed vectors we use the inverse transformation

$$F^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}^{-1} \quad (\text{if it exists - i.e. if } F \text{ is invertible}).$$

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 combi-

nation of the old basis vector set B , i.e.

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

we can write the transformation in matrix form as

$$\begin{bmatrix} \tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1n} \\ F_{21} & F_{22} & \cdots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nn} \end{bmatrix}. \quad (1.6)$$

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

$$\tilde{\mathbf{e}}_j = \sum_{k=1}^n F_{kj} \mathbf{e}_k. \quad (1.7)$$

Similarly, the inverse operation is given by

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

with the matrix form being

$$\begin{bmatrix} \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \dots, \tilde{\mathbf{e}}_n \end{bmatrix} \begin{bmatrix} F_{11}^{-1} & F_{12}^{-1} & \cdots & F_{1n}^{-1} \\ F_{21}^{-1} & F_{22}^{-1} & \cdots & F_{2n}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1}^{-1} & F_{n2}^{-1} & \cdots & F_{nn}^{-1} \end{bmatrix}, \quad (1.8)$$

and the index notation being

$$\mathbf{e}_i = \sum_{j=1}^n F_{ji}^{-1} \tilde{\mathbf{e}}_j. \quad (1.9)$$

By substituting Equation 1.7 into Equation 1.9, we get

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

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} :

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

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

1.2.3 Components transformation

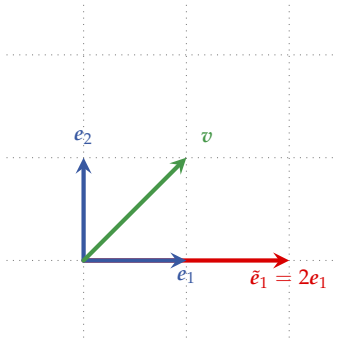


Figure 1.2: A vector v and two different sets of basis vectors: in blue the standard basis vectors and in red the basis set $\tilde{B} = \{2e_1, e_2\}$.

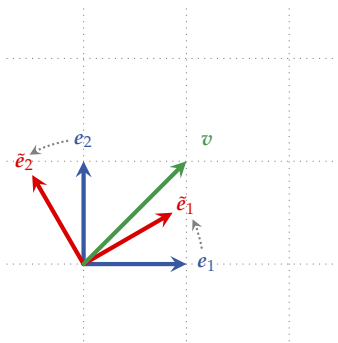


Figure 1.3: A vector v and two different sets of basis vectors: in blue the standard basis vectors and in red the basis set $\tilde{B} = \{2e_1, e_2\}$.

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 $v = \frac{1}{2}\tilde{e}_1 + \tilde{e}_2 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$.

We see that by scaling a basis vector by 2, its respective 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 .

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 $B = \{e_1, e_2, \dots, e_n\}$ using the respective components v_i :

$$v = \sum_{i=1}^n v_i e_i. \quad (1.12)$$

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

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

We can substitute into Equation 1.12 the explicit form of the vectors

in B (Equation 1.9):

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

Comparing the last equality in Equation 1.14 to Equation 1.12, we see that under a basis change the components of \mathbf{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 *contravariant*, and sometimes even refer to them as *contravariant vectors*⁴

⁴ this is a hint that there is at least another type of vectors: those that transform together with a change in basis, or *covariant 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 contravariant in regards to change of basis, we will from now on denote them using a superscript instead of a subscript:

$$\mathbf{v} = \sum_{i=1}^n v^i \mathbf{e}_i. \quad (1.14)$$

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:

$$\mathbf{v} = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} \quad (1.15)$$

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

$$\|\mathbf{v}\| = \sqrt{(v^1)^2 + (v^2)^2 + \cdots + (v^n)^2}. \quad (1.16)$$

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 contravariant 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 in front 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 explicitly in the basis set $B = \{e_1, e_2, \dots, e_n\}$ is

$$v = \sum_{i=1}^n v^i e_i. \quad (1.17)$$

Using Einstein's convention, this becomes simply

$$v = v^i e_i, \quad (1.18)$$

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 explicitly as

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

(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:

$$v^\top = [v_1, v_2, \dots, v_n].$$

(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

$$\langle v, w \rangle = [v_1, v_2, \dots, v_n] \begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{bmatrix} = v_i w^i. \quad (1.20)$$

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



Example 1.3 Einstein summation convention #3

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

$$(Av)^i = \sum_{j=1}^n A_{ij} v^j. \quad (1.21)$$

... 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 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ - 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 e_1, e_2, \dots, e_n is given by

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}, \quad (1.22)$$

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

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. α^* .

Now, we usually rotate rulers to align with the orientation of the magnitude we wish to measure - however here we want our rulers to also measure orientation. Therefore, 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 v using a ruler α^* is given by

$$m = \langle \alpha^*, v \rangle. \quad (1.23)$$

However, since we can represent a ruler using a vector, we can use the vector representation in Equation 1.23. And again, to distinguish ruler vectors from "regular" vectors written in an explicit

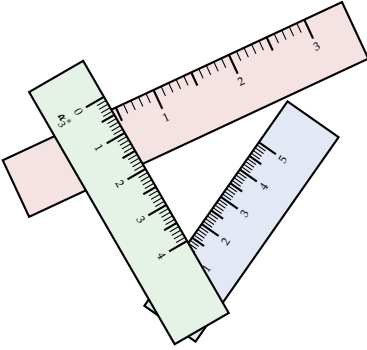


Figure 1.4: Three rulers, $\alpha_1^*, \alpha_2^*, \alpha_3^*$, each with its own orientation and frequency of graduation lines.

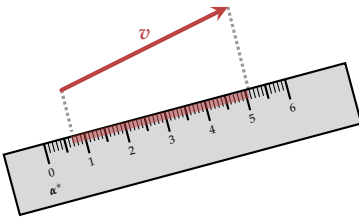


Figure 1.5: Measuring a vector using a ruler by projecting the vector onto the ruler. The result of the projection is drawn as a red line on the ruler.

form, we write the ruler vectors as row vectors:

$$\begin{aligned} \langle \alpha^*, v \rangle &= [\alpha_1, \alpha_2, \dots, \alpha_n] \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{bmatrix} \\ &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n. \end{aligned} \quad (1.24)$$

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 parallel lines, which fit precisely with the main graduation mark of the ruler they represent (Figure 1.6).

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*.

From Equation 1.25 we can see that any ruler α^* can be written as the left half of an inner product, i.e. $\langle \alpha^*, \cdot \rangle$. 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 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$. 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.

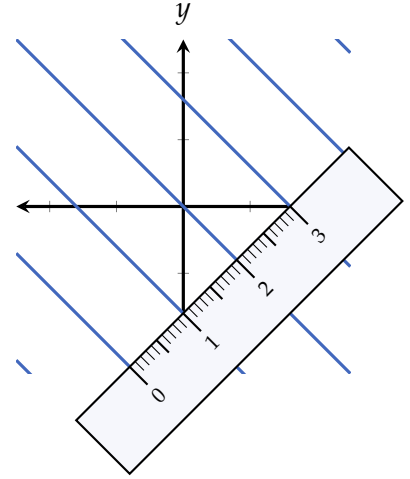


Figure 1.6: The graphical representation of a ruler shown next to the ruler, such that the infinite set of lines drawn match the ruler’s graduation marks.

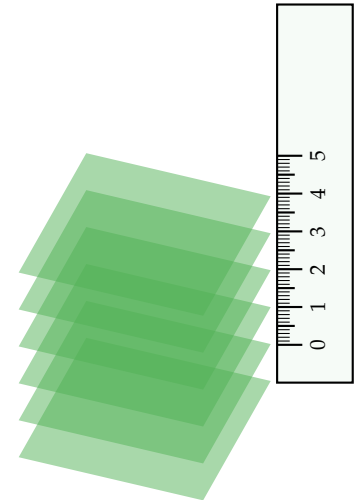


Figure 1.7: Representation of rulers in \mathbb{R}^3 as planes.

Example 1.4 Linear combination of dual vectors

Suppose we have two dual vectors

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

We can use α^* and β^* to create a third dual vector γ^* :

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

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

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

which is exactly the application of α^* on the vector five times followed by the application of β^* twice (try it for yourself!).



1.3.2 Basis sets and coordinate transformations

1. Dual basis: converting from a basis set in V to its dual in V^* .
2. Covariance of dual vectors basis change vs. contra-variance of vectors.

1.3.3 The complex dual (dual vectors in \mathbb{C}^n)

text

1.4 *The Bracket Notation and Einstein's Summation*

1.5 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.

1.5.1 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 $\phi : V \rightarrow \mathbb{F}$. 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 equipped 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?..

1.6 *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

