3

General Vector Spaces

Much of what we did with traditional vectors can also be done with other sets of "things". We will develop the appropriate theory here, and extend our notation appropriately.

One change we will make is that we will no longer restrict ourselves to real scalars. Henceforth, the term "scalar" will refer to either a real number or a complex number, depending on the context. Unless otherwise indicated, the default will be that scalars are complex numbers.

Since we will be using complex numbers as scalars, we will begin by briefly reviewing some of the basic ideas and notation of complex analysis. We will also briefly review the complex exponential since complex exponentials will be used in some of the examples.

3.1 Elementary Complex Analysis*

Basic Ideas, Notation and Terminology

Recall that z is a *complex number* if and only if it can be written as

$$z = x + iy$$

where x and y are real numbers and i is a "complex constant" satisfying $i^2 = -1$. The *real part* of z , denoted by $\text{Re}[z]$, is the real number x , while the *imaginary part* of z , denoted by $\text{Im}[z]$, is the real number y .[1] If $\text{Im}[z] = 0$ (equivalently, $z = \text{Re}[z]$), then z is said to be real. Conversely, if $\text{Re}[z] = 0$ (equivalently, $z = i\,\text{Im}[z]$), then z is said to be imaginary.

The *complex conjugate* of $z = x + iy$, which we will denote by z^* , is the complex number $z^* = x - iy$.

In the future, given any statement like "the complex number $z = x + iy$ ", it should automatically be assumed (unless otherwise indicated) that x and y are real numbers.

The algebra of complex numbers can be viewed as simply being the algebra of real numbers with the addition of a number i whose square is negative one. Thus, choosing some computations that will be of particular interest,

$$zz^* = z^*z = (x - iy)(x + iy) = x^2 - (iy)^2 = x^2 + y^2$$

* Much of this section has been stolen, slightly modified, from *Principles of Fourier Analysis* by Howell, with the permission of the author. It may contain a little more than we really need for now.

[1] Our text uses $\Re$ and $\Im$ instead of Re and Im .

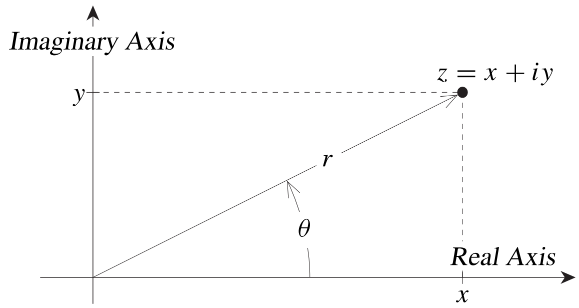

Figure 3.1: Coordinates in the complex plane for $z = x + iy$, where $x > 0$ and $y > 0$.

and

$$\frac{1}{z} = \frac{1}{x+iy} = \frac{1}{x+iy} \cdot \frac{x-iy}{x-iy} = \frac{x-iy}{x^2+y^2} = \frac{z^*}{zz^*} \quad .$$

We will often use the easily verified facts that, for any pair of complex numbers z and w,

$$(z+w)^* = z^* + w^* \qquad \text{and} \qquad (zw)^* = (z^*)(w^*) \quad .$$

The set of all complex numbers is denoted by $\mathbb{C}$. By associating the real and imaginary parts of the complex numbers with the coordinates of a two-dimensional Cartesian system, we can identify $\mathbb{C}$ with a plane (called, unsurprisingly, the complex plane). This is illustrated in figure 3.1. Also indicated in this figure are the corresponding polar coordinates r and θ for $z = x + iy$. The value r, which we will also denote by $|z|$, is commonly referred to as either the *magnitude*, the *absolute value*, or the *modulus* of z, while θ is commonly called either the *argument*, the *polar angle*, or the *phase* of z. It is easily verified that

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{z^* z} \quad ,$$

$$x = r\cos(\theta) \qquad \text{and} \qquad y = r\sin(\theta) \quad .$$

From this it follows that the complex number $z = x + iy$ can be written in *polar form*,

$$z = x + iy = r\,[\cos(\theta) + i\sin(\theta)] \quad .$$

It should also be pretty obvious that

$$|x| \le |z| \qquad \text{and} \qquad |y| \le |z| \quad .$$

The Complex Exponential

The basic complex exponential, denoted either by e^z or, especially when z is given by a formula that is hard to read as a superscript, by $\exp(z)$, will play a role in some of our examples, and later may play a bigger role in some applications

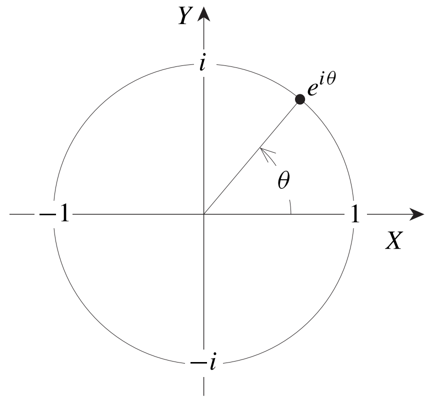

Figure 3.2: Plot of $e^{i\theta}$.

Basic Formulas

Recall that, if $z = x + iy$ then,

$$e^{x+iy} = e^x [\cos(y) + i \sin(y)]$$

and

$$e^{x-iy} = e^x [\cos(y) - i \sin(y)] \quad .$$

(Those who wonder why these are the formulas for $e^{x \pm iy}$ may want to look at the derivation that follows in a page or so.)

Letting $x = 0$ and $y = \theta$, these identities become the pair

$$e^{i\theta} = \cos(\theta) + i \sin(\theta) \tag{3.2a}$$

and

$$e^{-i\theta} = \cos(\theta) - i \sin(\theta) \quad . \tag{3.2b}$$

We can then solve for $\cos(\theta)$ and $\sin(\theta)$, obtaining

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

and

$$\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad .$$

We may occasionally need to compute the value of $e^{i\theta}$ and $\left|e^{\pm i\theta}\right|$ for specific real values of θ . Computing $\left|e^{\pm i\theta}\right|$ is easy. For any *real value* θ ,

$$\left|e^{\pm i\theta}\right| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = 1 \quad .$$

This also tells us that $e^{i\theta}$ is a point on the unit circle in the complex plane. In fact, comparing formula (3.2a) with the polar form for the complex number $e^{i\theta}$, we find that θ is, in fact, a polar angle for $e^{i\theta}$. The point $e^{i\theta}$ has been plotted in figure 3.2 for some unspecified θ between 0 and $\pi/2$. The real and imaginary parts of $e^{i\theta}$ can be computed either by using formula (3.2a) (or (3.2b)) or, at least for some values of θ , by inspection of figure 3.2. Clearly, for example,

$$e^{i\frac{1}{2}\pi} = i \quad , \quad e^{i\pi} = -1$$

and

$$e^{i\frac{3}{2}\pi} = e^{-i\frac{1}{2}\pi} = -i \quad .$$

?►*Exercise 3.1:* *Verify that*

$$e^{i2\pi n} = 1 \qquad \text{and} \qquad e^{i\pi n} = (-1)^n \qquad \text{for} \quad n = 0, \pm 1, \pm 2, \pm 3, \ldots \quad .$$

We should also note that, given any two complex numbers A and B,

1. $e^{A+B} = e^A e^B$,

2. e^A is the real exponential of A whenever A is real,

and

3. $\frac{d}{dt}e^{At} = Ae^{At}$.

Derivation†

Our goal is to derive a meaningful formula for e^z that extends our notion of the exponential to the case where z is complex. We will derive this formula by requiring that the complex exponential satisfies some of the same basic properties as the well-known real exponential function, and that it reduces to the real exponential function when z is real.

First, let us insist that the law of exponents (i.e., that $e^{A+B} = e^A e^B$) holds. Thus,

$$e^z = e^{x+iy} = e^x e^{iy} \quad . \tag{3.4}$$

We know the first factor, e^x. It's the real exponential from elementary calculus (a function you should be able to graph in your sleep).

To determine the second factor, consider the yet undefined function

$$f(t) = e^{it} \quad .$$

Since we insist that the complex exponential reduces to the real exponential when the exponent is real, we must have

$$f(0) = e^{i0} = e^0 = 1 \quad .$$

Recall, also, that $\frac{d}{dt}e^{at} = ae^{at}$ whenever a is a real constant. Requiring that this formula be true for imaginary constants gives

$$f'(t) = \frac{d}{dt}e^{it} = ie^{it} \quad . \tag{3.5}$$

Differentiating again gives

$$f''(t) = \frac{d}{dt}f'(t) = \frac{d}{dt}ie^{it} = i^2 e^{it} = -f(t) \quad ,$$

which can be rewritten as

$$f''(t) + f(t) = 0 \quad .$$

† For the interested reader only.

But this is a simple differential equation, the general solution of which is easily verified to be $A\cos(t) + B\sin(t)$ where A and B are arbitrary constants. So

$$e^{it} = f(t) = A\cos(t) + B\sin(t) \quad .$$

The constant A is easily determined from the requirement that $e^{i0} = 1$:

$$1 = e^{i0} = A\cos(0) + B\sin(0) = A\cdot 1 + B\cdot 0 = A \quad .$$

From equation (3.5) and the observation that

$$f'(t) = \frac{d}{dt}[A\cos(t) + B\sin(t)] = -A\sin(t) + B\cos(t) \quad ,$$

we see that

$$i = ie^{i0} = f'(0) = -A\sin(0) + B\cos(0) = -A\cdot 0 + B\cdot 1 = B \quad .$$

Thus $A = 1$, $B = i$, and

$$e^{it} = \cos(t) + i\sin(t) \quad . \tag{3.6}$$

Formula (3.6) is Euler's (famous) formula for e^{it}. It and equation (3.4) yield the formula

$$e^{x+iy} = e^x e^{iy} = e^x[\cos(y) + i\sin(y)]$$

for all real values of x and y. We will take this formula as the *definition* of the complex exponential. It then follows that

$$e^{x-iy} = e^x e^{i(-y)} = e^x[\cos(-y) + i\sin(-y)] = e^x[\cos(y) - i\sin(y)] \quad .$$

3.2 General Vector Spaces

Definition

A *vector* (or *linear*) *space* $\mathcal{V}$ is a collection of things (which we will usually call vectors and often denote using the same notation as for traditional vectors) that can be added together and multiplied by scalars. One of the elements of $\mathcal{V}$ must be the *zero vector* $\mathbf{0}$, which is the vector such that

$$\mathbf{0} + \mathbf{v} = \mathbf{v} \qquad \text{for each} \quad \mathbf{v} \in \mathcal{V} \quad .$$

In addition, we will require the following to hold whenever $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathcal{V}$, and α and β are scalars:

$$\alpha\mathbf{v} + \beta\mathbf{w} \in \mathcal{V}$$

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

$$(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$$

$$\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$$

$$1\mathbf{v} = \mathbf{v}$$

$$0\mathbf{v} = \mathbf{0}$$

The space $\mathcal{V}$ will be called a *real vector space* if only real numbers are allowed as scalars, and will be called a *complex vector space* if scalars can be complex. As already noted, the default is for our vector spaces to be complex.

Linear Combinations, Basis and Components
Linear Combinations and Linear (In)Dependence

Just about everything defined and discussed in section 2.2 regarding linear combinations also applies when the traditional vectors are replaced with more general entities.

A *linear combination* of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ... and $\mathbf{v}_N$ is any expression of the form

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_N\mathbf{v}_N$$

where α_1 , α_2 , ... and α_N are scalars. We will continue to limit ourselves to linear combinations with finitely many terms (to avoid having to deal with convergence issues involving infinite series).

The *span* of a set of vectors is the set of all linear combinations of those vectors.

A set of vectors is *linearly dependent* if one vector can be written as a linear combination of the others, and is *linearly independent* otherwise.

?►*Exercise 3.2 (Tests for linear (in)dependence):*

a: *Show that a finite set of vectors* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ *is linearly independent if and only if the only choice for scalars* α_1 , α_2 , α_1 , ...*and* α_N *such that*

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_N\mathbf{v}_N = \mathbf{0}$$

is $\alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$.[2]

b: *Convince yourself that an infinite set of vectors* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ *is linearly independent if and only if every finite subset of this set is linearly independent.*

Bases and Dimension

A *basis* for a vector space is a linearly independent set of vectors whose span equals the entire space. That is

$$\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots\}$$

is a basis for a vector space $\mathcal{V}$ if and only if

1. $\mathcal{B}$ is a linearly independent set, and
2. each vector in $\mathcal{V}$ can be written as a linear combination of the $\mathbf{b}_k$'s .

[2] This is basically the same as exercise 2.7 on page 2–7.

The *(scalar) components* of a vector $\mathbf{v}$ with respect to $\mathcal{B}$ are those scalars v_1, v_2, ... such that

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots \quad .$$

If the vector space has a basis with finitely many vectors, say N vectors, then the space is said to be "*finite dimensional* with *dimension* N" (or "N*-dimensional*"). If there is no basis with finitely many elements, the space is said to be *infinite dimensional*.

Here are some easily verified facts concerning any vector space $\mathcal{V}$:

1. The components of a vector with respect to a given basis are unique. That is, if

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots \qquad \text{and} \qquad \mathbf{w} = w_1\mathbf{b}_1 + w_2\mathbf{b}_2 + \cdots$$

where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots\}$ is some basis for $\mathcal{V}$, then

$$\mathbf{v} = \mathbf{w} \iff v_k = w_k \qquad \text{for each} \quad k \quad .$$

2. The dimension is unique. That is, if $\mathcal{V}$ has one basis with exactly N vectors (where N is some finite integer), then every basis of $\mathcal{V}$ has exactly N vectors.

3. Suppose $\mathcal{V}$ has finite dimension N, and let $\mathcal{B}$ be some set of vectors from $\mathcal{V}$. Then

$$\begin{aligned} \mathcal{B} \text{ is a basis for } \mathcal{V} &\iff \mathcal{B} \text{ is a linearly independent set of } N \text{ vectors} \\ &\iff \mathcal{B} \text{ is a set of } N \text{ vectors that spans } \mathcal{V} \quad . \end{aligned}$$

Convenient Notation

Assume $\mathcal{V}$ is an N-dimensional vector space with basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \} \quad .$$

Given any vector $\mathbf{v}$, we will use $|\mathbf{v}\rangle_{\mathcal{B}}$ to denote the $N \times 1$ matrix of components of $\mathbf{v}$ with respect to $\mathcal{B}$; that is,

$$|\mathbf{v}\rangle_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad \text{where} \quad \mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_N\mathbf{b}_N \quad .$$

Note that, using the "row vector" of the basis vectors and basic matrix multiplication,

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_N\mathbf{b}_N \\ &= \mathbf{b}_1 v_1 + \mathbf{b}_2 v_2 + \cdots + \mathbf{b}_N v_N \\ &= \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_N \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} |\mathbf{v}\rangle_{\mathcal{B}} \quad . \end{aligned}$$

In short,

$$\mathbf{v} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_N \end{bmatrix} |\mathbf{v}\rangle_{\mathcal{B}} \qquad \text{where} \qquad \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N \} = \mathcal{B} \quad .$$

If there is little likelihood of confusion, the "subscript $\mathcal{B}$" will be dropped, and we will just write $|\mathbf{v}\rangle$.

The "$|\ \rangle$" is sometimes called a *ket*, and is one element of the "bra-ket" notation we will develop more fully later. This notation gives us a convenient way to represent arbitrary vectors. One advantage of this notation is that it naturally converts vector computations to matrix computations.

?►Exercise 3.3: *Let* $\mathbf{v}$ *and* $\mathbf{w}$ *be two vectors, and let* α *and* β *be two scalars. Verify that*

$$\underbrace{|\alpha\mathbf{v} + \beta\mathbf{w}\rangle}_{\text{vector operations}} = \underbrace{\alpha\,|\mathbf{v}\rangle + \beta\,|\mathbf{w}\rangle}_{\text{matrix operations}} \quad .$$

Examples of Vector Spaces

Where ever mentioned below, N denotes some finite positive integer.

1. Any traditional vector space (as defined in the chapter on traditional vectors) is easily verified to be a real vector space as defined in this set of notes.

?►Exercise 3.4: *Go back over our development of traditional vectors and convince yourself that we showed that any traditional vector space satisfies the requirements for being a general vector space.*

By the way, a traditional vector space can be made into a complex vector space by simply defining complex scalar multiplication by

$$(\alpha + i\beta)\mathbf{v} = \alpha\mathbf{v} + i\beta\mathbf{v} \quad .$$

Mathematically, this is perfectly acceptable. How to physically interpret $i\mathbf{v}$, however, may be another issue.[3] Be aware that the "dot product" has to be re-defined when we allow complex scalars (see the next section).

2. $\mathbb{R}^3$, the set of all ordered triples of real numbers, with scalar multiplication and vector addition defined in the usual way,

$$\alpha(x, y, z) = (\alpha x, \alpha y, \alpha z)$$

and

$$(a, b, c) + (x, y, z) = (a + x, b + y, c + z) \quad ,$$

is a three-dimensional vector space. Most people's favorite basis for $\mathbb{R}^3$ is

$$\{ (1, 0, 0), (0, 1, 0), (0, 0, 1) \} \quad .$$

[3] But this is often done in E&M and other areas involving things that vary sinusoidally with time.

Note that, using this basis,

$$|(3,-2,5)\rangle = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad .$$

Often, people identify any real three-dimensional vector space with $\mathbb{R}^3$. Indeed, the ket notation essentially does this. Though this is convenient, don't take it too seriously. $\mathbb{R}^3$ is just a bunch of triple numbers.

3. In the same sense, $\mathbb{R}^N$ is an N-dimensional real vector space, and $\mathbb{C}^N$ is an N-dimensional complex vector space. Using the ket notation, any N-dimensional real vector space can be identified with $\mathbb{R}^N$ and any N-dimensional complex vector space can be identified with $\mathbb{C}^N$.

4. The set of all 2×3 matrices with the usual notions of scalar multiplication and matrix addition is a vector space. It is real or complex depending on whether we are allowing only real entries or not.

?►*Exercise 3.5* **a:** *What is a basis for this space?*

b: *What is the dimension of this space?*

5. Let $\mathcal{P}^3$ be the set of all (complex) third-degree polynomials, that is, $\mathcal{P}^3$ is the set of all functions of the form

$$p(z) = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \alpha_3 z^3$$

where the coefficients α_0, α_1, α_2 and α_3 are complex numbers. Using the usual notions of scalar multiplication of polynomials and addition of two polynomials, $\mathcal{P}^3$ is a vector space.

?►*Exercise 3.6* **a:** *Verify that $\mathcal{P}^3$ is a vector space.*

b: *What is a natural basis for $\mathcal{P}^3$?*

c: *What is the dimension of $\mathcal{P}^3$?*

6. Let $\mathcal{P}$ be the set of all (complex) polynomials; that is, $\mathcal{P}$ is the set of all functions of the form

$$p(z) = \sum_{k=0}^{M} \alpha_k z^k$$

where M can be any nonnegative integer and the coefficients are complex numbers. Using the usual notions of scalar multiplication of polynomials and addition of two polynomials, $\mathcal{P}$ is a complex vector space.

?►*Exercise 3.7* **a:** *Verify that $\mathcal{P}$ is a vector space.*

b: *What is a natural basis for $\mathcal{P}$?*

c: *What is the dimension of $\mathcal{P}$?*

7. The set of all functions of the form

$$f(x) = \sum_{k=-2}^{2} \alpha_k e^{i2\pi kx}$$

where the α_k's are complex numbers, is a complex vector space. An obvious basis for this space is

$$\mathcal{B}_E = \left\{ e^{i2\pi kx} : k = -2, -1, 0, 1, 2 \right\} \quad .$$

So this is a five-dimensional space.

?►*Exercise 3.8:* *Show that*

$$\mathcal{B}_T = \{ 1,\ \cos(2\pi x)\ ,\ \sin(2\pi x)\ ,\ \cos(4\pi x)\ ,\ \sin(4\pi x) \}$$

is also a basis for this space.

8. The set of all continuous functions on $[0, 1]$ is a vector space.

Vector Subspaces

Let $\mathcal{V}$ be a vector space. A *vector subspace* (or *linear subspace*) $\mathcal{U}$ of $\mathcal{V}$ is simply a vector space contained in $\mathcal{V}$; that is, $\mathcal{U}$ is some subset of vectors from $\mathcal{V}$ that, themselves, form a vector space.

Strictly speaking, $\mathcal{V}$, itself, and the set containing only the zero vector $\{\mathbf{0}\}$ are subspaces of $\mathcal{V}$, but we are rarely interested in these as subspaces. More often, our interest is in *proper subspaces* of $\mathcal{V}$, which are subspaces containing some, but not all nonzero vectors from $\mathcal{V}$.

Do note that the span of any set of vectors in $\mathcal{V}$ is a subspace of $\mathcal{V}$.

Here are a few examples

1. Let $\mathcal{V}$ be any three-dimensional traditional vector space, and let $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be a standard basis for this space.

   (a) The span of just $\{\mathbf{i}, \mathbf{j}\}$ is a two-dimensional subspace of $\mathcal{V}$.

   (b) So is the set of all linear combinations of $\mathbf{b}_1$ and $\mathbf{b}_2$ where

$$\mathbf{b}_1 = 2\mathbf{i} + \mathbf{j} \qquad \text{and} \qquad \mathbf{b}_2 = \mathbf{i} - 3\mathbf{j} \quad .$$

   Each of the above examples is typically viewed as a "plane in three-space".

2. Let $\mathcal{P}$ and $\mathcal{P}^3$ be, respectively, the set of all polynomials and the set of all third-degree polynomials (as discussed above). $\mathcal{P}^3$ is a finite-dimensional subspace of the infinite-dimensional vector space $\mathcal{P}$.

3. Along a similar line, $\mathcal{P}$, the set of all polynomials, is a subspace of the vector space of all continuous functions on $(0, 1)$.[4]

[4] Later in this course, we will spend a great deal of effort learning how to find the "best" "simple" infinite subspace of the vector space of all continuous functions on some interval, and use that subspace in finding the solutions to some partial differential equation problem. Among other things, this will lead to "Fourier series".

3.3 Inner Products and Norms

Definitions

An *inner product* for a vector space $\mathcal{V}$ is an operation that converts a pair of vectors $\mathbf{v}$ and $\mathbf{w}$ into a scalar — denoted by $\langle \mathbf{v} \mid \mathbf{w} \rangle$ — and which satisfies the following three rules:

1. (Linearity on the Right) For any three vectors $\mathbf{a}$, $\mathbf{v}$ and $\mathbf{w}$, and any scalar γ,

$$\langle \mathbf{a} \mid \mathbf{v}+\mathbf{w} \rangle = \langle \mathbf{a} \mid \mathbf{v} \rangle + \langle \mathbf{a} \mid \mathbf{w} \rangle$$

and

$$\langle \mathbf{a} \mid \gamma\mathbf{v} \rangle = \gamma \langle \mathbf{a} \mid \mathbf{v} \rangle \quad .$$

Thus,

$$\langle \mathbf{a} \mid \alpha\mathbf{v}+\beta\mathbf{w} \rangle = \alpha \langle \mathbf{a} \mid \mathbf{v} \rangle + \beta \langle \mathbf{a} \mid \mathbf{w} \rangle \quad .$$

2. (Positive Definiteness) For any vector $\mathbf{v}$,

$$\langle \mathbf{v} \mid \mathbf{v} \rangle \geq 0$$

with

$$\langle \mathbf{v} \mid \mathbf{v} \rangle = 0 \quad \iff \quad \mathbf{v} = \mathbf{0} \quad .$$

3. ((Conjugate) Symmetry) For any two vectors $\mathbf{a}$ and $\mathbf{v}$,

$$\langle \mathbf{a} \mid \mathbf{v} \rangle = \langle \mathbf{v} \mid \mathbf{a} \rangle^* \quad .$$

Remember, the * denotes complex conjugation. If $\mathcal{V}$ is a real vector space, then this requirement reduces to

$$\langle \mathbf{a} \mid \mathbf{v} \rangle = \langle \mathbf{v} \mid \mathbf{a} \rangle \quad .$$

Do observe that, unless the vector space is real, we do not quite have "linearity on the left". Instead, by combining the first two requirements above, we have

$$\begin{aligned}\langle \alpha\mathbf{v}+\beta\mathbf{w} \mid \mathbf{a} \rangle &= \langle \mathbf{a} \mid \alpha\mathbf{v}+\beta\mathbf{w} \rangle^* \\ &= [\alpha \langle \mathbf{a} \mid \mathbf{v} \rangle + \beta \langle \mathbf{a} \mid \mathbf{w} \rangle]^* = \alpha^* \langle \mathbf{v} \mid \mathbf{a} \rangle + \beta^* \langle \mathbf{w} \mid \mathbf{a} \rangle \quad .\end{aligned}$$

The corresponding *norm* $\|\cdot\|$ on $\mathcal{V}$ is then defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} \mid \mathbf{v} \rangle} \quad .$$

Much of what was done with dot products of traditional vectors can be mimicked in general using an inner product.

?►*Exercise 3.9 (inner product test for equality):*

a: *Suppose* $\mathbf{D}$ *is a vector such that*

$$\langle \mathbf{D} \mid \mathbf{c} \rangle = 0 \qquad \text{for every} \quad \mathbf{c} \quad .$$

Show that $\mathbf{D} = \mathbf{0}$. *(Hint: What if we choose* $\mathbf{c} = \mathbf{D}$ *?)*

b: *Suppose* $\mathbf{A}$ *and* $\mathbf{B}$ *are two vectors such that*

$$\langle \mathbf{A} \mid \mathbf{c} \rangle = \langle \mathbf{B} \mid \mathbf{c} \rangle \qquad \text{for every} \quad \mathbf{c} \quad .$$

Show that $\mathbf{A} = \mathbf{B}$. *(Hint: What about* $\mathbf{D} = \mathbf{A} - \mathbf{B}$ *?)*

Examples

Several notable examples of inner products are listed below. For each example, *you* should verify that the given example really is an inner product by verifying that it satisfies the defining requirements just given.

1. The dot product
$$\langle \mathbf{v} \mid \mathbf{w} \rangle = \|\mathbf{v}\| \, \|\mathbf{w}\| \cos(\theta)$$
for any traditional traditional vector space is an inner product.

2. The traditional inner product and corresponding norm on $\mathbb{R}$ are simply the ordinary product and the absolute value,
$$\langle a \mid x \rangle = ax \quad \text{and} \quad \|a\| = \sqrt{a^2} = |a| \quad .$$
On $\mathbb{R}^3$, the traditional inner product is, as you well know,
$$\langle (a,b,c) \mid (x,y,z) \rangle = ax + by + cz \quad .$$
More generally, the traditional inner product on $\mathbb{R}^N$ is
$$\Big\langle (a_1, a_2, \ldots, a_N) \,\Big|\, (x_1, x_2, \ldots, x_N) \Big\rangle = \sum_{k=1}^{N} a_k x_k \quad .$$
The corresponding norm is then
$$\|(a_1, a_2, \ldots, a_N)\| = \sqrt{\sum_{k=1}^{N} {a_k}^2} \quad .$$

3. On $\mathbb{C}$, the traditional inner product is
$$\langle a \mid z \rangle = a^* z \quad .$$
So
$$\|a\| = \sqrt{\langle a \mid a \rangle} = \sqrt{a^* a} = |a| \quad .$$
On $\mathbb{C}^3$, the traditional inner product is the natural extension of the inner product on $\mathbb{C}$,
$$\langle (a,b,c) \mid (x,y,z) \rangle = a^* x + b^* y + c^* z \quad .$$
More generally, the traditional inner product on $\mathbb{C}^N$ is
$$\Big\langle (a_1, a_2, \ldots, a_N) \,\Big|\, (x_1, x_2, \ldots, x_N) \Big\rangle = \sum_{k=1}^{N} {a_k}^* x_k \quad .$$
The corresponding norm is then
$$\|(a_1, a_2, \ldots, a_N)\| = \sqrt{\sum_{k=1}^{N} {a_k}^* a_k} = \sqrt{\sum_{k=1}^{N} |a_k|^2} \quad .$$

4. Consider the set of all continuous (complex-valued) functions on the interval $[0, 1]$. A standard inner product for this space is

$$\langle\, f \mid g \,\rangle \;=\; \int_0^1 f^*(x) g(x)\, dx \quad .$$

The corresponding norm,

$$\|f\| \;=\; \sqrt{\int_0^1 f^*(x) f(x)\, dx} \;=\; \sqrt{\int_0^1 |f(x)|^2\, dx} \quad ,$$

is called either the *energy norm* or the L^2 norm.

5. More generally, an inner product on the set of all continuous (complex-valued) functions on the interval $[0, 1]$ can be defined by

$$\langle\, f \mid g \,\rangle \;=\; \int_0^1 f^*(x) g(x) w(x)\, dx$$

where w is any "pre-chosen" continuous function on $[1, 0]$ which is positive on $(0, 1)$ (say, $w(x) = \sqrt{x}$). The $w(x)$ is called the *weight function* for this inner product. (This sort of inner product will become extremely important later when we study Sturm-Liouville problems and develop methods for solving partial differential equations.)

?►Exercise 3.10: *Verify that each of the above described inner products does satisfy the three rules stated on page 3–11 for inner products.*

Orthogonality

Generally, there is not a natural notion of "angles between vectors" when our vector space is not a space of traditional vectors. In fact, with complex vector spaces the notion becomes very problematic! Still, we can say that two (nonzero) vectors $\mathbf{v}$ and $\mathbf{w}$ are parallel if one is a scalar multiple of the other. (In practice, though, the term "parallel" is rarely used with nontraditional vectors.)

Moreover, using inner products, we can generalize the notion of "orthogonality" in a useful way, namely:

$$\text{A set } \{\mathbf{v}_1, \mathbf{v}_2, \ldots\} \text{ is } \textit{orthogonal} \quad \Longleftrightarrow \quad \left\langle\, \mathbf{v}_j \,\middle|\, \mathbf{v}_k \,\right\rangle = 0 \quad \text{whenever} \quad j \neq k \quad .$$

$$\text{A set } \{\mathbf{v}_1, \mathbf{v}_2, \ldots\} \text{ is } \textit{orthonormal} \quad \Longleftrightarrow \quad \left\langle\, \mathbf{v}_j \,\middle|\, \mathbf{v}_k \,\right\rangle = \delta_{jk} \quad .$$

This, of course, assumes that an inner product is defined for the space from whence we took the $\mathbf{v}_k$'s .

Strictly speaking, we should say that sets of vectors are orthogonal or orthonormal *with respect to a given inner product*. Many inner products can be defined on any vector space (look at all the inner products mentioned for the space of all continuous functions on $[0, 1]$). It is quite possible for a pair of vectors to be orthogonal with respect to one inner product and not orthogonal with respect to another.

Still, despite the arbitrariness of the choice of the inner product, much of what could be derived using the dot product is re-derivable using any inner product. For example, it is easy to show that any orthogonal set of nonzero vectors is linearly independent (just trust me on that for now). And, in the homework, you will show that, given any two nonzero vectors $\mathbf{a}$ and $\mathbf{v}$, $\mathbf{v}$ can be decomposed into the sum of an orthogonal pair of vectors, with one being the generalized projection of $\mathbf{v}$ onto $\mathbf{a}$.

Be aware that choosing the "right" inner product and basis for a given problem is an important element of mathematical physics. Later (much later!), we will see that one reason the Sturm-Liouville theory is so important is that it tells us how to pick the right inner products and bases for solving various partial differential equation problems.

3.4 Inner Products and Vector Components

Finding the components of a vector $\mathbf{v}$ with respect to an orthonormal basis is (in theory) quite easy if we can use the inner product, less easy if the basis is merely orthogonal, and somewhat painful if the basis is not even orthogonal. We (i.e., *you*) will derive these formulas here. They will be used again and again later on.

In the following, $\mathbf{a}$ and $\mathbf{v}$ are arbitrary vectors in a vector space $\mathcal{V}$ having a basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \dots \}$$

and inner product $\langle \cdot | \cdot \rangle$. We will not insist that $\mathcal{V}$ be finite dimensional, but we will continue to require that vectors in $\mathcal{V}$ can be expressed as finite linear combinations of basis vectors (i.e., linear combinations with finitely many terms) to avoid dealing with convergence issues regarding infinite series.

Because $\mathcal{B}$ is a basis, we can express $\mathbf{a}$ and $\mathbf{v}$ in terms of their components with respect to $\mathcal{B}$,

$$\mathbf{a} = \sum_j a_j \mathbf{b}_j \qquad \text{and} \qquad \mathbf{v} = \sum_k v_k \mathbf{b}_k \quad .$$

Plugging these into the inner product and using the properties of inner products, we will first derive the component formulas for inner products and norms. Then we will discuss how to find the components from the inner products.

By the way, in all of this, it is assumed we know the values of the inner products of the basis vectors with themselves,

$$\left\langle \mathbf{b}_j \;\middle|\; \mathbf{b}_k \right\rangle \quad .$$

(This will later correspond to knowing the "metric" for a space.)

When the Basis is Arbitrary

Using the basic properties of the inner product, you can easily verify the following:

1. $\langle \mathbf{a} \mid \mathbf{v} \rangle = \sum_j \sum_k {a_j}^* v_k \gamma_{jk}$ where $\gamma_{jk} = \left\langle \mathbf{b}_j \;\middle|\; \mathbf{b}_k \right\rangle$.

2. $\|\mathbf{v}\|^2 = \sum_j \sum_k {v_j}^* v_k \gamma_{jk}$.

?► ***Exercise 3.11:*** *Verify both of the above statements. (It should still only take one to three lines for each.)*

Note that, if $\mathbf{a} = \mathbf{b}_1$, then

$$\mathbf{a} = \mathbf{b}_1 = 1\mathbf{b}_1 + 0\mathbf{b}_2 + 0\mathbf{b}_3 + \cdots = \sum_j \delta_{1j}\mathbf{b}_j \quad .$$

Thus, using the above component formula for the inner products,

$$\langle \mathbf{b}_1 \mid \mathbf{v} \rangle = \sum_j \sum_k \delta_{1j}^* v_k \gamma_{jk} = \sum_j \sum_k \begin{Bmatrix} 1 & \text{if} & j = 1 \\ 0 & \text{if} & j \neq 1 \end{Bmatrix} v_k \gamma_{jk} = \sum_k v_k \gamma_{1k} \quad .$$

Rewriting this slightly and repeating with all the basis vectors yields the linear system

$$\begin{aligned}
\langle \mathbf{b}_1 \mid \mathbf{v} \rangle &= \gamma_{11}v_1 + \gamma_{12}v_2 + \gamma_{13}v_3 + \cdots \quad , \\
\langle \mathbf{b}_2 \mid \mathbf{v} \rangle &= \gamma_{21}v_1 + \gamma_{22}v_2 + \gamma_{23}v_3 + \cdots \quad , \\
\langle \mathbf{b}_3 \mid \mathbf{v} \rangle &= \gamma_{31}v_1 + \gamma_{32}v_2 + \gamma_{33}v_3 + \cdots \quad , \\
&\vdots \quad .
\end{aligned}$$

Assuming we know each $\langle \mathbf{b}_j \mid \mathbf{v} \rangle$, we can then find the components of $\mathbf{v}$ by solving the above system for the v_k's . Keep in mind that $\mathbf{v}$ will have only a finite number of (nonzero) components, so the above system really does consist of just a finite number of equations, each with a finite number of nonzero terms on the right.

When the Basis is Orthogonal, but Not Necessarily Orthonormal

Suppose our basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \ldots \}$$

is orthogonal with respect to the given inner product, but not necessarily orthonormal. Then

$$\left\langle \mathbf{b}_j \mid \mathbf{b}_k \right\rangle = \begin{cases} \left\| \mathbf{b}_j \right\|^2 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases} \quad ,$$

and you can easily verify that

1. $\langle \mathbf{a} \mid \mathbf{v} \rangle = \sum_j a_j{}^* v_j \left\| \mathbf{b}_j \right\|^2$.

2. $\|\mathbf{v}\|^2 = \sum_j \left| v_j \right|^2 \left\| \mathbf{b}_j \right\|^2$.

3. For each j ,

$$\left\langle \mathbf{b}_j \mid \mathbf{v} \right\rangle = v_j \left\| \mathbf{b}_j \right\|^2 \qquad \text{and} \qquad \left\langle \mathbf{v} \mid \mathbf{b}_j \right\rangle = v_j^* \left\| \mathbf{b}_j \right\|^2 \quad .$$

?►*Exercise 3.12:* *Verify each of the three statements above. (It should still only take one to three lines for each.)*

From the above formula for $\langle \mathbf{v} \mid \mathbf{b}_j \rangle$ we see that the components of $\mathbf{v}$ with respect to this basis can be computed from the appropriate inner products by

$$v_j = \frac{\langle \mathbf{b}_j \mid \mathbf{v} \rangle}{\|\mathbf{b}_j\|^2} \quad . \tag{3.7}$$

Hence

$$\mathbf{v} = \sum_j \frac{\langle \mathbf{b}_j \mid \mathbf{v} \rangle}{\|\mathbf{b}_j\|^2} \mathbf{b}_j \quad . \tag{3.8}$$

(By the way: Formulas (3.7) and (3.8) turn out to be the ones actually used in constructing Fourier series and other “eigenfunction” expansions for solving partial differential equations.)

When the Basis is Orthonormal

If our basis

$$\mathcal{B} = \{ \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \ldots \}$$

is orthonormal with respect to the given inner product, then

$$\langle \mathbf{b}_j \mid \mathbf{b}_k \rangle = \begin{cases} 1 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases} \quad ,$$

and you can easily verify that the above formulas reduce to our favorite formulas for inner products, norms, and components:

1. $\langle \mathbf{a} \mid \mathbf{v} \rangle = \sum_j {a_j}^* v_j$.
2. $\|\mathbf{v}\|^2 = \sum_j |v_j|^2$.
3. For each j ,
   $$v_j = \langle \mathbf{b}_j \mid \mathbf{v} \rangle \qquad \text{and} \qquad v_j^* = \langle \mathbf{v} \mid \mathbf{b}_j \rangle \quad .$$
   Hence
   $$\mathbf{v} = \sum_j \langle \mathbf{b}_j \mid \mathbf{v} \rangle \, \mathbf{b}_j \quad .$$

?►*Exercise 3.13:* *Verify each of the statements above. (It should still only take one to three lines for each.)*

3.5 Constructing Orthonormal Sets from Arbitrary Sets

Obviously, we would prefer to do most of our vector computations using an orthonormal basis. In theory, this is easy to arrange.

Normalizing

Given any orthogonal set of nonzero vectors

$$\{\, \mathbf{b}_1 ,\ \mathbf{b}_2 ,\ \dots \} \quad ,$$

we can construct a corresponding orthonormal set

$$\{\, \mathbf{u}_1 ,\ \mathbf{u}_2 ,\ \dots \}$$

by *normalizing* each $\mathbf{b}_k$, that is, defining $\mathbf{u}_k$ by

$$\mathbf{u}_k = \frac{\mathbf{b}_k}{\|\mathbf{b}_k\|} \quad .$$

Components Orthogonal to a Set

Earlier, we discussed decomposing a traditional vector $\mathbf{v}$ into vectors parallel and orthogonal to any nonzero traditional vector $\mathbf{u}$,

$$\mathbf{v} = \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) + \overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v})$$

Recall that $\overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v})$, the vector component of $\mathbf{v}$ parallel to $\mathbf{u}$, is given by

$$\overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\,\mathbf{u} \quad .$$

Thus, $\overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v})$, the vector component of $\mathbf{v}$ orthogonal to $\mathbf{u}$, is given by

$$\overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|^2}\,\mathbf{u} \quad .$$

You can easily show (and may have done so in the homework) that the above holds for more general vector spaces with inner products, provided you replace the $\mathbf{v}\cdot\mathbf{u}$ with $\langle\, \mathbf{u} \,|\, \mathbf{v} \,\rangle$ (remember, the order in inner products can be important). That is, if $\mathbf{v}$ and $\mathbf{u}$ are two vectors in a vector space with an inner product, and $\mathbf{u} \neq \mathbf{0}$, then

$$\mathbf{v} = \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) + \overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v})$$

where

$$\overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle\, \mathbf{u} \mid \mathbf{v} \,\rangle}{\|\mathbf{u}\|^2}\,\mathbf{u}$$

is "parallel" to $\mathbf{u}$ and

$$\overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \frac{\langle\, \mathbf{u} \mid \mathbf{v} \,\rangle}{\|\mathbf{u}\|^2}\,\mathbf{u} \quad .$$

is orthogonal to $\mathbf{u}$.

Of particular interest is the case where $\mathbf{u}$ is a unit vector. Then $\|\mathbf{u}\| = 1$ and the above reduces to

$$\mathbf{v} = \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) + \overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v})$$

where

$$\overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \langle\, \mathbf{u} \mid \mathbf{v} \,\rangle \mathbf{u}$$

is "parallel" to $\mathbf{u}$ and

$$\overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \overrightarrow{\mathrm{pr}}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \langle\, \mathbf{u} \mid \mathbf{v} \,\rangle \mathbf{u} \quad .$$

is orthogonal to $\mathbf{u}$.

Observe that all we are doing to obtain $\overrightarrow{\mathrm{or}}_{\mathbf{u}}(\mathbf{v})$ is to subtract from $\mathbf{v}$ the vector component of $\mathbf{v}$ in the direction of $\mathbf{u}$.

Now suppose we have any finite orthonormal set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots \mathbf{u}_M\}$. Then subtracting from $\mathbf{v}$ the vector components of $\mathbf{v}$ with respect to each of these $\mathbf{u}_j$'s leaves us with a vector

$$\mathbf{w} = \mathbf{v} - \sum_{j=1}^{M} \langle\, \mathbf{u}_j \mid \mathbf{v} \,\rangle \mathbf{u}_j$$

having no (nonzero) components with respect to each of these $\mathbf{u}_j$'s. With a little thought, you'll realize that this means $\mathbf{w}$ is othogonal to each each of these $\mathbf{u}_j$'s. And with a little computation, you can rigorously derive this.

?►*Exercise 3.14:* *Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots \mathbf{u}_M\}$ be an orthonormal set of vectors in some vector space with an inner product, and let $\mathbf{v}$ be any vector in that space. Set*

$$\mathbf{w} = \mathbf{v} - \sum_{j=1}^{M} \langle\, \mathbf{u}_j \mid \mathbf{v} \,\rangle \mathbf{u}_j$$

and show that

$$\langle\, \mathbf{u}_k \mid \mathbf{w} \,\rangle = 0 \qquad \text{for} \quad k = 1, 2, \dots, M \quad .$$

All this leads to a natural extension of previous notation and concepts: Let $\mathcal{S}$ be some set of vectors from $\mathcal{V}$ (possibly a subspace of $\mathcal{V}$), and let $\mathbf{v}$ be any nonzero vector in $\mathcal{V}$. Then we define the *projection of* $\mathbf{v}$ *onto the space* $\mathcal{S}$ (or the *component of* $\mathbf{v}$ *in* $\mathcal{S}$) and the *(vector) component of* $\mathbf{v}$ *orthogonal to* $\mathcal{S}$ — denoted, respectively, by $\overrightarrow{\mathrm{pr}}_{\mathcal{S}}(\mathbf{v})$ and $\overrightarrow{\mathrm{or}}_{\mathcal{S}}(\mathbf{v})$ — to be those vectors such that all the following holds

1. $\mathbf{v} = \overrightarrow{\mathrm{pr}}_{\mathcal{S}}(\mathbf{v}) + \overrightarrow{\mathrm{or}}_{\mathcal{S}}(\mathbf{v})$,
2. $\overrightarrow{\mathrm{pr}}_{\mathcal{S}}(\mathbf{v})$ is in the span of $\mathcal{S}$, and
3. $\overrightarrow{\mathrm{or}}_{\mathcal{S}}(\mathbf{v})$ is a vector orthogonal to every vector in $\mathcal{S}$.

It should be clear (using the ideas in this subsection), that if the span of $\mathcal{S}$ is an M-dimensional subspace with an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_M\}$, then these two components exist and are "easily" computed by

$$\overrightarrow{\mathrm{pr}}_{\mathcal{S}}(\mathbf{v}) = \overrightarrow{\mathrm{pr}}_{\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\dots,\mathbf{u}_M\}}(\mathbf{v}) = \sum_{j=1}^{M} \langle\, \mathbf{u}_j \mid \mathbf{v} \,\rangle \mathbf{u}_j \tag{3.9a}$$

and

$$\overrightarrow{\text{or}}_{\mathcal{S}}(\mathbf{v}) = \overrightarrow{\text{or}}_{\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\ldots,\mathbf{u}_M\}}(\mathbf{v}) = \mathbf{v} - \overrightarrow{\text{pr}}_{\mathcal{S}}(\mathbf{v}) = \mathbf{v} - \sum_{j=1}^{M} \langle\, \mathbf{u}_j \mid \mathbf{v} \,\rangle \mathbf{u}_j \quad . \tag{3.9b}$$

In practice, the above general formula for $\overrightarrow{\text{pr}}_{\mathcal{S}}(\mathbf{v})$ is used to find the "best finite-dimensional approximation" to things in infinite-dimensional spaces — but that discussion would lead us away from the main computational topic of this section, which is the Gram-Schmidt procedure.

It should also be noted that we will get the same projections no matter what basis we use for $\mathcal{S}$. That is, we have:

Lemma 3.1

Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M\}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_M\}$ be two sets of vectors having the same span, and let $\mathcal{S}$ be this span. Then for any vector $\mathbf{v}$,

$$\overrightarrow{\text{pr}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_M\}}(\mathbf{v}) = \overrightarrow{\text{pr}}_{\mathcal{S}}(\mathbf{v}) = \overrightarrow{\text{pr}}_{\{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_M\}}(\mathbf{v})$$

and

$$\overrightarrow{\text{or}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_M\}}(\mathbf{v}) = \overrightarrow{\text{or}}_{\mathcal{S}}(\mathbf{v}) = \overrightarrow{\text{or}}_{\{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_M\}}(\mathbf{v}) \quad .$$

The Gram-Schmidt Procedure

Basic Idea

Using the above results, we can create a procedure, known as the "Gram-Schmidt procedure", that will take any linearly independent set of vectors[5]

$$\{\, \mathbf{v}_1,\ \mathbf{v}_2,\ \ldots \}$$

and construct a corresponding orthonormal set

$$\{\, \mathbf{u}_1,\ \mathbf{u}_2,\ \ldots \}$$

having the same span as the first set. The basic idea is that, once you have

$$\mathcal{B}_{K-1} = \{\, \mathbf{u}_1,\ \mathbf{u}_2,\ \ldots,\ \mathbf{u}_{K-1} \}$$

constructed from the first $K-1$ $\mathbf{v}_j$'s, you find the vector $\mathbf{w}_K$ which is the component of $\mathbf{v}_K$ orthogonal to all the vectors in $\mathcal{B}_{K-1}$,

$$\begin{aligned} \mathbf{w}_K &= \overrightarrow{\text{or}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{K-1}\}}(\mathbf{v}_K) \\ &= \mathbf{v}_K - \overrightarrow{\text{pr}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{K-1}\}}(\mathbf{v}_K) = \mathbf{v}_K - \sum_{j=1}^{K-1} \langle\, \mathbf{u}_j \mid \mathbf{v}_K \,\rangle \mathbf{u}_j \quad . \end{aligned}$$

Then normalize this vector, call the result $\mathbf{u}_K$, and add it to the set $\mathcal{B}_{K-1}$ to get the slightly bigger orthonormal set

$$\mathcal{B}_K = \{\, \mathbf{u}_1,\ \mathbf{u}_2,\ \ldots,\ \mathbf{u}_K \,\} \quad .$$

Continue until you run out of $\mathbf{v}_k$'s.

That is the basic idea behind the *Gram-Schmidt procedure*. Next are the details:

[5] The method also works if the set is not linearly independent, provided you throw out the zero vectors that are generated.

The Gram-Schmidt Procedure (Steps)

We start with some linearly independent set

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots\} \quad .$$

To construct the corresponding orthonormal set

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots\} \quad ,$$

do the following:

1. Normalize $\mathbf{v}_1$ to get $\mathbf{u}_1$

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \quad .$$

2. Set $\mathbf{w}_2$ equal to $\mathbf{v}_2$ with the vector component in the direction of $\mathbf{u}_1$ subtracted off,

$$\mathbf{w}_2 = \mathbf{v}_2 - \langle \mathbf{u}_1 \mid \mathbf{v}_2 \rangle \mathbf{u}_1 \quad ,$$

   and normalize $\mathbf{w}_2$ to get $\mathbf{u}_2$,

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \quad .$$

3. Set $\mathbf{w}_3$ equal to $\mathbf{v}_3$ with the vector components in the direction of $\mathbf{u}_1$ and $\mathbf{u}_2$ subtracted off,

$$\mathbf{w}_3 = \mathbf{v}_3 - \langle \mathbf{u}_1 \mid \mathbf{v}_3 \rangle \mathbf{u}_1 - \langle \mathbf{u}_2 \mid \mathbf{v}_3 \rangle \mathbf{u}_2 \quad ,$$

   and normalize $\mathbf{w}_3$ to get $\mathbf{u}_3$,

$$\mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \quad .$$

4. ...

5. In general, once you have found

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{K-1}\} \quad ,$$

   you find $\mathbf{u}_K$ by setting $\mathbf{w}_K$ equal to $\mathbf{v}_K$ with the vector components in the directions of $\mathbf{u}_1$, $\mathbf{u}_2$, ... and $\mathbf{u}_{K-1}$ all subtracted off,

$$\mathbf{w}_K = \mathbf{v}_K - \sum_{j=1}^{K-1} \langle \mathbf{u}_j \mid \mathbf{v}_K \rangle \mathbf{u}_j \quad ,$$

   and then normalizing $\mathbf{w}_K$ to get $\mathbf{u}_K$,

$$\mathbf{u}_K = \frac{\mathbf{w}_K}{\|\mathbf{w}_K\|} \quad .$$

6. ...

7. (Continue until you run out of $\mathbf{v}_j$'s.)

Observations

The main purpose of the Gram-Schmidt procedure is to generate an orthonormal basis from a basis that is not orthonormal. However, we will also find it useful in a different task in a few pages, and for that, we'll find some of the following observations useful.

First of all, it is fairly easy to see that, at each stage,

$$\mathcal{N} = \{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_K \}$$

is an orthonormal set with

$$\text{span of } \{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K \} = \text{span of } \{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_K \} \quad .$$

So we can view $\mathcal{N}$ as an orthonormal basis for the vector space spanned by these $\mathbf{v}_j$'s. In a few pages, we will find it useful to know the components of these $\mathbf{v}_j$'s with respect to basis $\mathcal{N}$.

To find these components, we start with the formulas for finding the vectors in $\mathcal{N}$. For the first, we have

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1$$

Solving this for $\mathbf{v}_1$ we get

$$\mathbf{v}_1 = \|\mathbf{v}_1\| \, \mathbf{u}_1 = \|\mathbf{v}_1\| \, \mathbf{u}_1 + 0\mathbf{u}_2 + 0\mathbf{u}_3 + \cdots + 0\mathbf{u}_K \quad .$$

So,

$$|\mathbf{v}_1\rangle_{\mathcal{N}} = \begin{bmatrix} \|\mathbf{v}_1\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad .$$

For $k > 1$,

$$\mathbf{u}_k = \frac{1}{\|\mathbf{w}_k\|}\mathbf{w}_k = \frac{1}{\|\mathbf{w}_k\|}\left[\mathbf{v}_k - \sum_{j=1}^{k-1} \langle \, \mathbf{u}_j \mid \mathbf{v}_k \, \rangle \mathbf{u}_j \right] \quad .$$

Do remember that

$$\mathbf{w}_k = \overrightarrow{\text{or}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}\}}(\mathbf{v}_k) \quad ,$$

and since the corresponding sets of $\mathbf{u}_j$'s and $\mathbf{v}_j$'s span the same space, we also have

$$\mathbf{w}_k = \overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k) \quad .$$

Solving the above for the $\mathbf{v}_k$'s in terms of the $\mathbf{u}_j$'s, we get

$$\mathbf{v}_k = \|\mathbf{w}_k\| \, \mathbf{u}_k + \sum_{j=1}^{k-1} \langle \, \mathbf{u}_j \mid \mathbf{v}_k \, \rangle \mathbf{u}_j$$

where

$$\mathbf{w}_k = \overrightarrow{\text{or}}_{\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_{k-1}\}}(\mathbf{v}_k) = \overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k) \quad .$$

In particular, for $\mathbf{v}_2$ and $\mathbf{v}_2$, we have

$$\begin{aligned}\mathbf{v}_2 &= \left\|\overrightarrow{\text{or}}_{\mathbf{v}_1}(\mathbf{v}_2)\right\| \mathbf{u}_2 + \langle\, \mathbf{u}_1 \mid \mathbf{v}_2 \,\rangle\, \mathbf{u}_1 \\ &= \langle\, \mathbf{u}_1 \mid \mathbf{v}_2 \,\rangle\, \mathbf{u}_1 + \left\|\overrightarrow{\text{or}}_{\mathbf{v}_1}(\mathbf{v}_2)\right\| \mathbf{u}_2 + 0\mathbf{u}_3 + \cdots + 0\mathbf{v}_K \quad ,\end{aligned}$$

and

$$\begin{aligned}\mathbf{v}_3 &= \left\|\overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2\}}(\mathbf{v}_3)\right\| \mathbf{u}_3 + \langle\, \mathbf{u}_1 \mid \mathbf{v}_3 \,\rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2 \mid \mathbf{v}_3 \,\rangle\, \mathbf{u}_2 \\ &= \langle\, \mathbf{u}_1 \mid \mathbf{v}_3 \,\rangle\, \mathbf{u}_1 + \langle\, \mathbf{u}_2 \mid \mathbf{v}_3 \,\rangle\, \mathbf{u}_2 + \left\|\overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2\}}(\mathbf{v}_3)\right\| \mathbf{u}_3 + 0\mathbf{u}_4 + \cdots + 0\mathbf{v}_K \quad .\end{aligned}$$

So,

$$|\mathbf{v}_2\rangle_{\mathcal{N}} = \begin{bmatrix} \langle\, \mathbf{u}_1 \mid \mathbf{v}_2 \,\rangle \\ \left\|\overrightarrow{\text{or}}_{\mathbf{v}_1}(\mathbf{v}_2)\right\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad |\mathbf{v}_3\rangle_{\mathcal{N}} = \begin{bmatrix} \langle\, \mathbf{u}_1 \mid \mathbf{v}_3 \,\rangle \\ \langle\, \mathbf{u}_2 \mid \mathbf{v}_3 \,\rangle \\ \left\|\overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2\}}(\mathbf{v}_3)\right\| \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad .$$

Continuing these computations leads to:

Lemma 3.2
Let

$$\mathcal{N} = \{\, \mathbf{u}_1,\, \mathbf{u}_2,\, \dots,\, \mathbf{u}_K \,\}$$

be the orthonormal set of vectors generated by the Gram-Schmidt procedure from a linearly independent set

$$\{\, \mathbf{v}_1,\, \mathbf{v}_2,\, \dots,\, \mathbf{v}_K \,\} \quad .$$

Then, treating $\mathcal{N}$ *as a basis for the space spanned by these* $\mathbf{v}_k$*'s,*

$$|\mathbf{v}_1\rangle_{\mathcal{N}} = \begin{bmatrix} \|\mathbf{v}_1\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad ,$$

and, for $1 < k \le K$ *,*

$$|\mathbf{v}_k\rangle_{\mathcal{N}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_K \end{bmatrix} \quad \textit{with} \quad \alpha_j = \begin{cases} \left\langle\, \mathbf{u}_j \;\middle|\; \mathbf{v}_k \,\right\rangle & \textit{if} \quad j < k \\ \left\|\overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k)\right\| & \textit{if} \quad j = k \\ 0 & \textit{if} \quad j > k \end{cases} \quad .$$

3.6 "Volumes" of N-Dimensional Hyper-Parallelepipeds (Part I)

Basics and Notation

Here is the situation: In a N-dimensional Euclidean space we have a "hyper-parallelepiped" generated by a linearly independent set of N vectors

$$\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N \} \quad .$$

Our goal in this section is to find a way to compute the N-dimensional volume of this object. (We will later expand on this subject.)

For convenience, for $K = 1, 2, \ldots, N$, let

$$\mathcal{P}_K = \text{" K-dimensional parallelepiped" generated by vectors } \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K\} \quad .$$

So, in fact $\mathcal{P}_1$ is a straight line segment, $\mathcal{P}_2$ is a parallelogram, $\mathcal{P}_3$ is a traditional 3-dimensional parallelepiped, etc.

We'll also let

$$V_K(\mathcal{P}_K) = \text{" K-dimensional volume" of } \mathcal{P}_K \quad .$$

In particular,

$$\begin{aligned} V_1(\mathcal{P}_1) &= \text{the one-dimensional "volume" of line segment } \mathcal{P}_1 \\ &= \text{length of } \mathcal{P}_1 \\ &= \|\mathbf{v}_1\| \quad , \end{aligned}$$

$$\begin{aligned} V_2(\mathcal{P}_2) &= \text{the two-dimensional "volume" of parallelogram } \mathcal{P}_2 \\ &= \text{the area of parallelogram } \mathcal{P}_2 \\ &= \|\mathbf{v}_2 \times \mathbf{v}_1\| \qquad \text{if} \quad N = 3 \end{aligned}$$

and

$$\begin{aligned} V_3(\mathcal{P}_3) &= \text{the three-dimensional "volume" of parallelogram } \mathcal{P}_3 \\ &= \text{the traditional volume of } \mathcal{P}_3 \\ &= |\mathbf{v}_3 \cdot (\mathbf{v}_2 \times \mathbf{v}_1)| \qquad \text{if} \quad N = 3 \quad . \end{aligned}$$

The above classic vector formulas were derived from the classical "base × height" formula. That "base × height" formula applies in higher dimensions, as well. In terms of our notation, this formula is

$$V_K(\mathcal{P}_K) = V_{K-1}(\mathcal{P}_{K-1}) \times H_K \qquad \text{for} \quad K = 2, 3, 4, \ldots$$

where H_K is the length of the vector component of $\mathbf{v}_K$ orthogonal to $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{K-1}\}$; that is,

$$H_k = \left\| \overrightarrow{\text{or}}_{\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{K-1}\}}(\mathbf{v}_K) \right\| \quad .$$

Applying the above formula repeatedly, we have

$$V_2(\mathcal{P}_2) = V_1(\mathcal{P}_1) \times H_2 = \|\mathbf{v}_1\| H_2 \quad ,$$

$$V_3(\mathcal{P}_3) = V_2(\mathcal{P}_2) \times H_3 = \|\mathbf{v}_1\| H_2 H_3 \quad ,$$

$$V_4(\mathcal{P}_4) \;=\; V_3(\mathcal{P}_3) \times H_4 \;=\; \|\mathbf{v}_1\|\, H_2 H_3 H_4 \quad ,$$

$$\vdots$$

In general,

$$V_N(\mathcal{P}_N) \;=\; \|\mathbf{v}_1\|\, H_2 H_3 \cdots H_N \tag{3.10a}$$

where

$$H_k \;=\; \left\| \overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k) \right\| \quad . \tag{3.10b}$$

And how do we find these H_k's ? Well, recall that they are just the lengths of the "$\mathbf{w}_k$'s" computed in the Gram-Schmidt process. Consequently, a major part of using the above is to simply go through the Gram-Schmidt process to find the orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ generated from $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$, and, at the point where we set

$$\mathbf{u}_k \;=\; \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|} \qquad \text{with} \quad \mathbf{w}_k \;=\; \overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k) \;=\; \mathbf{v}_k \;-\; \sum_{j=1}^{k-1} \langle\, \mathbf{u}_j \mid \mathbf{v}_k \,\rangle \mathbf{u}_j \quad ,$$

we also set

$$H_k \;=\; \|\mathbf{w}_k\| \quad .$$

Try it yourself.

?►*Exercise 3.15:* *Let $\mathcal{V}$ be a four-dimensional space of traditional vectors with orthonormal basis*

$$\{\,\mathbf{i}\,,\,\mathbf{j}\,,\,\mathbf{k}\,,\,\mathbf{l}\,\}$$

and let

$$\mathbf{v}_1 \;=\; 3\mathbf{i} \quad ,$$

$$\mathbf{v}_2 \;=\; 2\mathbf{i} \;+\; 4\mathbf{j} \quad ,$$

$$\mathbf{v}_3 \;=\; 8\mathbf{i} \;-\; 3\mathbf{j} \;+\; 5\mathbf{k}$$

and

$$\mathbf{v}_4 \;=\; 4\mathbf{i} \;+\; 7\mathbf{j} \;-\; 2\mathbf{k} \;+\; 2\mathbf{l} \quad .$$

Using formula (3.10) (not the vector formulas), compute the following:

a: *The area of the parallelogram generated by* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *.*

b: *The volume of the parallelepiped generated by* $\mathbf{v}_1$ *,* $\mathbf{v}_2$ *and* $\mathbf{v}_3$ *.*

c: *The "four-dimensional volume" of the hyper-parallelepiped generated by* $\mathbf{v}_1$ *,* $\mathbf{v}_2$ *,* $\mathbf{v}_3$ *and* $\mathbf{v}_4$ *.*

d: *The area of the parallelogram generated by* $\mathbf{v}_3$ *and* $\mathbf{v}_4$ *.*

Linear Dependence and Volumes of Hyper-Parallelepiped

In the above, we assumed the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ is linearly independent. That insures $\mathbf{v}_1$ is nonzero, and that no $\mathbf{v}_k$ is a linear combination of the other $\mathbf{v}_j$'s . And since the span of

$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}\}$ equals the span of $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-1}\}$, we also cannot have $\mathbf{v}_k$ being a linear combination of vectors from this last set. Thus,

$$H_k = \|\overrightarrow{\text{or}}_{\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_{k-1}\}}(\mathbf{v}_k)\| = \left\| \mathbf{v}_k - \sum_{j=1}^{k-1} \langle \mathbf{u}_j \mid \mathbf{v}_k \rangle \mathbf{u}_j \right\| \neq 0$$

And that, in turn, insures that the hyper-volume $V_N(\mathcal{P}_N)$ (which is given by $\|\mathbf{v}_1\| H_2 H_3 \cdots H_N$) is not zero.

However, with minor adjustments in our procedure[6], the computation of $V_N(\mathcal{P}_N)$ can still be done if the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ is NOT linearly independent. But then, either $\mathbf{v}_1 = \mathbf{0}$ or one of the other $\mathbf{v}_k$'s is a linear combination of the other $\mathbf{v}_j$'s. With a little thought, you'll then realize that this then means one of the H_k's must be zero, and, hence,

$$V_N(\mathcal{P}_N) = \|\mathbf{v}_1\| H_2 H_3 \cdots H_N = 0 \quad .$$

In summary,

Lemma 3.3
Let $V_k(\mathcal{P}_K)$ be the K-dimensional volume" of the hyper-parallelepiped $\mathcal{P}_K$ generated by the set of vectors

$$\{ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N \} \quad .$$

Then this set of $\mathbf{v}_j$'s is linearly independent if and only if $V_k(\mathcal{P}_K) \neq 0$

This may later be used to develop a more practical "test for linear independence" using components.

Component Formulas

Basics

It should be noted that the computations for exercise 3.15 were considerably simplified by the $\mathbf{v}_k$'s being "nicely oriented" relative to the given orthonormal basis. Obviously, it would be nice to have the simplest possible formulas for these "hyper-volumes" in terms of the components of the $\mathbf{v}_k$'s using whatever basis is convenient. That, ultimately, will require material we will develop later. But we can at least do a little preliminary work in anticipation.

First of all, when given any basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ and any set of N vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$, we will let $\mathbf{V}_\mathcal{B}$ be the "matrix of components of the $\mathbf{v}_j$'s with respect to basis $\mathcal{B}$" given by

$$\mathbf{V}_\mathcal{B} = \text{“} \left[|\mathbf{v}_1\rangle_\mathcal{B} \quad |\mathbf{v}_2\rangle_\mathcal{B} \quad |\mathbf{v}_3\rangle_\mathcal{B} \quad \cdots \quad |\mathbf{v}_N\rangle_\mathcal{B} \right] \text{”} = \begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} & \cdots & v_{N,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & \cdots & v_{N,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & \cdots & v_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{1,N} & v_{2,N} & v_{3,N} & \cdots & v_{N,N} \end{bmatrix}$$

[6] e.g., setting $\mathbf{u}_k = \mathbf{0}$ if $\mathbf{w}_k = \mathbf{0}$. Of course, this means that the resulting set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$ will not be a basis. We'll have to throw away all the zero vectors, first.

where, for $j = 1, 2, \ldots, N$,

$$\mathbf{v}_j = \sum_{k=1}^{N} v_{j,k} \mathbf{b}_k \quad .$$

(Warning: The indexing in this matrix is "nonstandard" in that the $(j,k)^{\text{th}}$ entry is actually $v_{k,j}$.)

Using the Gram-Schmidt Generated Basis

In particular, let's consider $\mathbf{V}_{\mathcal{N}}$ where

$$\mathcal{N} = \{ \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N \}$$

is the orthonormal basis generated by the Gram-Schmidt procedure from the set of $\mathbf{v}_j$'s. Since this will later serve as an intermediate basis, let us denote the coresponding components of the $\mathbf{v}_j$'s with primes,

$$\mathbf{v}_j = \sum_{k=1}^{N} v'_{j,k} \mathbf{u}_k \qquad \text{for} \quad j = 1, 2, \ldots, N \quad .$$

Fortunately, we got a good idea of what these $v'_{j,k}$'s are a few pages ago, and summarized our findings in lemma 3.2 on page 3–22. Checking that corollary (and keeping in mind what H_k denotes), we see that

$$|\mathbf{v}_1\rangle_{\mathcal{N}} = \begin{bmatrix} v'_{1,1} \\ v'_{1,2} \\ v'_{1,3} \\ \vdots \\ v'_{1,N} \end{bmatrix} = \begin{bmatrix} \|\mathbf{v}_1\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

And, for $k > 1$,

$$|\mathbf{v}_k\rangle_{\mathcal{N}} = \begin{bmatrix} v'_{k,1} \\ v'_{k,2} \\ v'_{k,3} \\ \vdots \\ v'_{k,N} \end{bmatrix} \qquad \text{with} \quad v'_{k,j} = \begin{cases} \langle \mathbf{u}_j \mid \mathbf{v}_k \rangle & \text{if} \quad j < k \\ H_k & \text{if} \quad j = k \\ 0 & \text{if} \quad j > k \end{cases} \quad .$$

Consequently,

$$\mathbf{V}_{\mathcal{N}} = \text{“} \big[|\mathbf{v}_1\rangle_{\mathcal{N}} \quad |\mathbf{v}_2\rangle_{\mathcal{N}} \quad |\mathbf{v}_3\rangle_{\mathcal{N}} \quad \cdots \quad |\mathbf{v}_N\rangle_{\mathcal{N}} \big] \text{”} = \begin{bmatrix} \|\mathbf{v}_1\| & v'_{2,1} & v'_{3,1} & \cdots & v'_{N,1} \\ 0 & H_2 & v'_{3,2} & \cdots & v'_{N,2} \\ 0 & 0 & H_3 & \cdots & v'_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & H_N \end{bmatrix}$$

with

$$v'_{k,j} = \langle \mathbf{u}_j \mid \mathbf{v}_k \rangle \qquad \text{for} \quad k > j \quad .$$

Presumably, you already know enough about determinants of matrices to realize that the determinant of this matrix is very easily computed,

$$\det(\mathbf{V}_{\mathcal{N}}) = \|\mathbf{v}_1\| \, H_2 H_3 \cdots H_N \quad .$$

Recall seeing the right side of this equation before? Yes! It's the formula we derived for the volume of the hyper-parallelepiped generated by $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$. Amazing! That gives us the following theorem.

Theorem 3.4 (a component formula for the volume of a hyper-parallelepiped)
Let $\mathcal{P}_N$ be the hyper-parallelepiped generated by the linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ of traditional vectors, and let $V_N(\mathcal{P}_N)$ denote its N-dimensional volume. Then

$$V_N(\mathcal{P}_N) = \det(\mathbf{V}_{\mathcal{N}})$$

where $\mathcal{N}$ is the orthonormal set generated from $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ via the Gram-Schmidt procedure and $\mathbf{V}_{\mathcal{N}}$ is the matrix of components of the $\mathbf{v}_j$'s with respect to basis $\mathcal{N}$

$$\mathbf{V}_{\mathcal{N}} = \begin{bmatrix} v'_{1,1} & v'_{2,1} & v'_{3,1} & \cdots & v'_{N,1} \\ v'_{1,2} & v'_{2,2} & v'_{3,2} & \cdots & v'_{N,2} \\ v'_{1,3} & v'_{2,3} & v'_{3,3} & \cdots & v'_{N,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v'_{1,N} & v'_{2,N} & v'_{3,N} & \cdots & v'_{N,N} \end{bmatrix} \quad \textit{with} \quad \begin{bmatrix} v'_{k,1} \\ v'_{k,2} \\ v'_{k,3} \\ \vdots \\ v'_{k,N} \end{bmatrix} = |\mathbf{v}_k\rangle_{\mathcal{N}} \quad .$$

We will later build on this theorem to obtain component formulas for volume using more natural bases.

Geometric Formulas

Using the above, we can also derive "geometric" formulas for volume based directly on the lengths and inner products of the vectors in the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ generating the parallelepided $\mathcal{P}_N$. These formulas may be useful later when discussing differential elements of length, area and volume when using arbitrary coordinate systems.

The Formulas

Here are those formulas for $N = 1$, $N = 2$ and $N = 3$.

Theorem 3.5 (geometric formulas for the volumes of parallelepipeds)
Let $\mathcal{P}_N$ be the hyper-parallelepiped generated by the linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N\}$ of traditional vectors, and let $V_N(\mathcal{P}_N)$ denote its N-dimensional volume. Then

$$V_1(\mathcal{P}_1) = \|\mathbf{v}_1\| \quad ,$$

$$V_2(\mathcal{P}_2) = \sqrt{\|\mathbf{v}_1\|^2 \, \|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)^2}$$

and

$$V_3(\mathcal{P}_3) = \sqrt{\alpha - \beta - 2\gamma}$$

where

$$\alpha = \|\mathbf{v}_1\|^2 \, \|\mathbf{v}_2\|^2 \, \|\mathbf{v}_3\|^2 \quad ,$$

$$\beta = \|\mathbf{v}_1\|^2 (\mathbf{v}_2 \cdot \mathbf{v}_3)^2 + \|\mathbf{v}_2\|^2 (\mathbf{v}_1 \cdot \mathbf{v}_2)^2 + \|\mathbf{v}_3\|^2 (\mathbf{v}_1 \cdot \mathbf{v}_2)^2$$

and

$$\gamma = \frac{(\mathbf{v}_1 \cdot \mathbf{v}_2)^2 (\mathbf{v}_1 \cdot \mathbf{v}_3)^2}{\|\mathbf{v}_1\|^2} = \frac{(\mathbf{v}_2 \cdot \mathbf{v}_1)^2 (\mathbf{v}_2 \cdot \mathbf{v}_3)^2}{\|\mathbf{v}_2\|^2} = \frac{(\mathbf{v}_3 \cdot \mathbf{v}_2)^2 (\mathbf{v}_3 \cdot \mathbf{v}_1)^2}{\|\mathbf{v}_3\|^2} \quad .$$

(Since these vectors are assumed to be "traditional" vectors, we're using the traditional "dot" notation.) The derivations of these formulas will be given in a few pages.

It should be noted that these geometric formulas apply no matter what the dimension of the vector space is. That can make these formulas a little more applicable than the vector formulas we have or the component formulas that we'll later derive.

?►*Exercise 3.16:* *Let $\mathcal{V}$ be a four-dimensional space of traditional vectors with orthonormal basis*

$$\{ \mathbf{i} , \mathbf{j} , \mathbf{k} , \mathbf{l} \}$$

and let

$$\mathbf{a} = 3\mathbf{i} \quad ,$$

$$\mathbf{b} = 2\mathbf{i} + 4\mathbf{j} \quad ,$$

$$\mathbf{c} = 8\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$$

and

$$\mathbf{d} = 4\mathbf{i} + 7\mathbf{j} - 2\mathbf{k} + 2\mathbf{l} \quad .$$

Using the above geometric formulas (not the vector formulas or formula (3.10)), compute the following:

a: *The area of the parallelogram generated by* $\mathbf{a}$ *and* $\mathbf{b}$*.*

b: *The area of the parallelogram generated by* $\mathbf{c}$ *and* $\mathbf{d}$*.*

c: *The volume of the parallelepiped generated by* $\mathbf{a}$*,* $\mathbf{b}$ *and* $\mathbf{c}$*.*

d: *The volume of the parallelepiped generated by* $\mathbf{b}$*,* $\mathbf{c}$ *and* $\mathbf{d}$*.*

On occasion, each $\mathbf{v}_k$ will be a positive scalar multiple Δv_k of some other vector $\mathbf{b}_k$,

$$\mathbf{v}_k = \Delta v_k \, \mathbf{b}_k \quad .$$

On such occasions, the formulas in the above theorem can be rewritten with the Δv_k's factored out:

$$V_1(\mathcal{P}_1) = \|\mathbf{b}_1\| \, \Delta v_1 \quad , \tag{3.11}$$

$$V_2(\mathcal{P}_2) = \sqrt{\|\mathbf{b}_1\|^2 \, \|\mathbf{b}_2\|^2 - (\mathbf{b}_2 \cdot \mathbf{b}_1)^2} \, \Delta v_1 \Delta v_2 \tag{3.12}$$

and

$$V_3(\mathcal{P}_3) = \sqrt{A - B - 2C} \, \Delta v_1 \Delta v_2 \Delta v_3 \tag{3.13}$$

where

$$A = \|\mathbf{b}_1\|^2 \, \|\mathbf{b}_2\|^2 \, \|\mathbf{b}_3\|^2 \quad , \tag{3.14a}$$

$$B = \|\mathbf{b}_1\|^2 \, (\mathbf{b}_2 \cdot \mathbf{b}_3)^2 + \|\mathbf{b}_2\|^2 \, (\mathbf{b}_1 \cdot \mathbf{b}_2)^2 + \|\mathbf{b}_3\|^2 \, (\mathbf{b}_1 \cdot \mathbf{b}_2)^2 \tag{3.14b}$$

and

$$C = \frac{(\mathbf{b}_1 \cdot \mathbf{b}_2)^2 (\mathbf{b}_1 \cdot \mathbf{b}_3)^2}{\|\mathbf{b}_1\|^2} = \frac{(\mathbf{b}_2 \cdot \mathbf{b}_1)^2 (\mathbf{b}_2 \cdot \mathbf{b}_3)^2}{\|\mathbf{b}_2\|^2} = \frac{(\mathbf{b}_3 \cdot \mathbf{b}_2)^2 (\mathbf{b}_3 \cdot \mathbf{b}_1)^2}{\|\mathbf{b}_3\|^2} \quad . \tag{3.14c}$$

These formulas reduce even further if the $\mathbf{b}_k$'s are unit vectors, and even further if the set of $\mathbf{b}_k$'s is orthonormal.

?►Exercise 3.17: *To what do the above formulas for $V_1(\mathcal{P}_1)$, $V_2(\mathcal{P}_2)$ and $V_3(\mathcal{P}_3)$ reduce*

a: *when the $\mathbf{b}_k$'s are unit vectors.*

b: *when the set of $\mathbf{b}_k$ is orthonormal.*

Derivations for Theorem 3.5

Here are the derivations for the formulas in theorem 3.5. Frankly, I don't expect you to spend much (or any) time on them. They are included just to assure you that I did not make up those formulas.[7]

The orthonormal set $\mathcal{N} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ generated from the $\mathbf{v}_s$'s via the Gram-Schmidt procedure are given by

$$\mathbf{u}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \quad , \quad \mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \quad , \quad \mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} \quad , \quad \cdots$$

where

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1 \\
\mathbf{w}_2 &= \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 \\
\mathbf{w}_3 &= \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2 \\
&\ \vdots
\end{aligned}$$

Computing the norm squared of these:

$$\|\mathbf{w}_1\|^2 = \|\mathbf{v}_1\|^2 \quad ,$$

$$\begin{aligned}
\|\mathbf{w}_2\|^2 &= [\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1] \cdot [\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1] \\
&= \mathbf{v}_2 \cdot \mathbf{v}_2 - 2(\mathbf{v}_2 \cdot \mathbf{u}_1)^2 + (\mathbf{v}_2 \cdot \mathbf{u}_1)^2 \\
&= \|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2 \quad ,
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{w}_3\|^2 &= [\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2] \cdot [\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2] \\
&= \mathbf{v}_3 \cdot \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 \\
&\qquad - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 + (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 + (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 \\
&= \mathbf{v}_3 \cdot \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 \quad .
\end{aligned}$$

Also,

$$\mathbf{v}_1 \cdot \mathbf{w}_1 = \|\mathbf{v}_1\|^2 = \|\mathbf{w}_1\|^2 \quad ,$$

$$\mathbf{v}_2 \cdot \mathbf{w}_2 = \mathbf{v}_2 \cdot [\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1]$$

[7] I wrote these derivations before writing up some of the earlier sections in these notes. It is very likely that the following derivations can be 'cleaned up' and shortened using some of the material already given in these notes.

$$= \mathbf{v}_2 \cdot \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2$$
$$= \|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2 = \|\mathbf{w}_2\|^2 \quad ,$$

and

$$\begin{aligned} \mathbf{v}_3 \cdot \mathbf{w}_3 &= \mathbf{v}_3 \cdot [\mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{v}_3 \cdot \mathbf{u}_2)\mathbf{u}_2] \\ &= \mathbf{v}_3 \cdot \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 \\ &= \|\mathbf{v}_3\|^2 - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2 = \|\mathbf{w}_3\|^2 \quad . \end{aligned}$$

Thus,

$$v'_{1,1} = \cdot = \|\mathbf{v}_1\| \quad ,$$

$$v'_{2,2} = \mathbf{v}_2 \cdot \mathbf{u}_2 = \mathbf{v}_2 \cdot \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{\mathbf{v}_2 \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|} = \|\mathbf{w}_2\| \quad ,$$

and

$$v'_{3,3} = \mathbf{v}_3 \cdot \mathbf{u}_3 = \mathbf{v}_3 \cdot \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{\mathbf{v}_3 \cdot \mathbf{w}_3}{\|\mathbf{w}_3\|} = \|\mathbf{w}_3\| \quad .$$

For $N = 1$,

$$V_1(\mathcal{P}_1) = \left|v'_{1,1}\right| = \cdots = \|\mathbf{v}_1\| \quad .$$

For $N = 2$, (since $\mathbf{v}_1 = \|\mathbf{v}_1\|\,\mathbf{u}_1$),

$$\begin{aligned} V_2(\mathcal{P}_2) &= \left|v'_{1,1}v'_{2,2}\right| \\ &= \|\mathbf{v}_1\|\,\|\mathbf{w}_2\| \\ &= \|\mathbf{v}_1\| \sqrt{\|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2} \\ &= \sqrt{\|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2 - \|\mathbf{v}_1\|^2\,(\mathbf{v}_2 \cdot \mathbf{u}_1)^2} \\ &= \sqrt{\|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)^2} \quad . \end{aligned}$$

For $N = 3$,

$$\begin{aligned} [V_3(\mathcal{P}_3)]^2 &= (v'_{1,1})^2(v'_{2,2})^2(v'_{3,3})^2 \\ &= \|\mathbf{v}_1\|^2\,\|\mathbf{w}_2\|^2\,\|\mathbf{w}_3\|^2 \\ &= \|\mathbf{v}_1\|^2\,\|\mathbf{w}_2\|^2 \left(\|\mathbf{v}_3\|^2 - (\mathbf{v}_3 \cdot \mathbf{u}_1)^2 - (\mathbf{v}_3 \cdot \mathbf{u}_2)^2\right) \\ &= \|\mathbf{v}_1\|^2\,\|\mathbf{w}_2\|^2 \left(\|\mathbf{v}_3\|^2 - \frac{(\mathbf{v}_3 \cdot \mathbf{v}_1)^2}{\|\mathbf{v}_1\|^2} - \frac{(\mathbf{v}_3 \cdot \mathbf{w}_2)^2}{\|\mathbf{w}_2\|^2}\right) \\ &= \|\mathbf{v}_1\|^2 \left(\|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2\right) \|\mathbf{v}_3\|^2 \\ &\quad - (\mathbf{v}_3 \cdot \mathbf{v}_1)^2\,\|\mathbf{w}_2\|^2 - \|\mathbf{v}_1\|^2\,(\mathbf{v}_3 \cdot \mathbf{w}_2)^2 \\ &= \|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2\,\|\mathbf{v}_3\|^2 - \|\mathbf{v}_1\|^2\,(\mathbf{v}_2 \cdot \mathbf{u}_1)^2\,\|\mathbf{v}_3\|^2 \\ &\quad - (\mathbf{v}_3 \cdot \mathbf{v}_1)^2 \left[\|\mathbf{v}_2\|^2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)^2\right] \\ &\quad - \|\mathbf{v}_1\|^2\,(\mathbf{v}_3 \cdot [\mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{u}_1)\mathbf{u}_1])^2 \\ &= \|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2\,\|\mathbf{v}_3\|^2 - (\mathbf{v}_2 \cdot \mathbf{v}_1)^2\,\|\mathbf{v}_3\|^2 \\ &\quad - \|\mathbf{v}_2\|^2\,(\mathbf{v}_3 \cdot \mathbf{v}_1)^2 + (\mathbf{v}_2 \cdot \mathbf{u}_1)^2(\mathbf{v}_3 \cdot \mathbf{v}_1)^2 \end{aligned}$$

$$
\begin{aligned}
&\quad - \ \|\mathbf{v}_1\|^2\,(\mathbf{v}_3\cdot\mathbf{v}_2) \ + \ \|\mathbf{v}_1\|^2\,(\mathbf{v}_3\cdot\mathbf{u}_1)^2(\mathbf{v}_2\cdot\mathbf{u}_1)^2 \\
&= \ \|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2\,\|\mathbf{v}_3\|^2 \\
&\quad - \ \|\mathbf{v}_1\|^2\,(\mathbf{v}_2\cdot\mathbf{v}_3) \ - \ \|\mathbf{v}_2\|^2\,(\mathbf{v}_1\cdot\mathbf{v}_2)^2 \ - \ \|\mathbf{v}_3\|^2\,(\mathbf{v}_1\cdot\mathbf{v}_2)^2 \\
&\quad + \ (\mathbf{v}_2\cdot\mathbf{u}_1)^2(\mathbf{v}_3\cdot\mathbf{v}_1)^2 \ + \ \|\mathbf{v}_1\|^2\,(\mathbf{v}_3\cdot\mathbf{u}_1)^2(\mathbf{v}_2\cdot\mathbf{u}_1)^2 \quad .
\end{aligned}
$$

Keeping in mind that $\mathbf{v}_1 = \|\mathbf{v}_1\|\,\mathbf{u}_1$, this becomes

$$
\begin{aligned}
[V_3(\mathcal{P}_3)]^2 \ = \ & \|\mathbf{v}_1\|^2\,\|\mathbf{v}_2\|^2\,\|\mathbf{v}_3\|^2 \\
& - \ \|\mathbf{v}_1\|^2\,(\mathbf{v}_2\cdot\mathbf{v}_3)^2 \ - \ \|\mathbf{v}_2\|^2\,(\mathbf{v}_1\cdot\mathbf{v}_2)^2 \ - \ \|\mathbf{v}_3\|^2\,(\mathbf{v}_1\cdot\mathbf{v}_2)^2 \\
& + \ 2\frac{(\mathbf{v}_1\cdot\mathbf{v}_2)^2(\mathbf{v}_1\cdot\mathbf{v}_3)^2}{\|\mathbf{v}_1\|^2} \quad .
\end{aligned}
$$

Note: by “symmetry” it should be clear that the last term could be replaced with

$$
2\frac{(\mathbf{v}_2\cdot\mathbf{v}_1)^2(\mathbf{v}_2\cdot\mathbf{v}_3)^2}{\|\mathbf{v}_2\|^2} \qquad \text{or} \qquad 2\frac{(\mathbf{v}_3\cdot\mathbf{v}_2)^2(\mathbf{v}_3\cdot\mathbf{v}_1)^2}{\|\mathbf{v}_3\|^2} \quad .
$$