VECTOR

Algebra is an art to structure sets. Set is a permissive idea: any collection of any things is a set. Given a set, someone will discern some structure on the set. There must be more structures than sets. Algebra keeps us disciplined. We choose to regard most sets as unstructured, and focus on structured sets of a few types, such as groups, number fields, and scalar sets. These structured sets permeate much of science and everyday activities.

The object of algebra—and much of science and everyday activities—is to see structures in sets. Algebra structures sets by axioms, deduces theorems and algorithms, and relates them to phenomena in the world.

Group, number field, and scalar set. A group G is a set that satisfies four axioms: G is closed under an operation, the operation is associative, G has an identity element, and every element in G is invertible. Furthermore, if the operation is commutative, we say that G is a commutative group.

A number field F is a commutative group under one operation called addition and, with the identity element of addition excluded, is a commutative group under another operation called multiplication. Furthermore, multiplication distributes over addition. We will mostly use the field of real numbers, R, and the field of complex numbers, C. We will occasionally use the field of rational numbers, Q.

A scalar set S over a number field F is defined by twelve axioms. Axioms 1-5 say that S is a commutative group under addition—that is, elements in S are additive. Axioms 6-8 define an operation on S and F: multiplying an element in S and an element in F gives an element in S—that is, elements in S are scalable by numbers in F. Axioms 9 and 10 say that the multiplication of elements in S and elements in F distribute over two types of addition: adding elements in F and adding elements in S. Axioms 11 and 12 say that the scalar set is one-dimensional. Familiar scalar sets include piles of apples, pieces of gold, amounts of money, lumps of mass, quantities of energy, durations between moments, and distances between places.

Vector Space

We now generalize scalar sets to vector spaces by modifying Axioms 11 and 12 only. Thus, a scalar set is a one-dimensional vector space, and an n-dimensional vector space is the Cartesian product of n scalar sets. Like scalars, vectors are additive, and vectors are scalable by numbers.

This section lists axioms that define vector spaces. Take a look at them, but do not be intimidated. Later we will replace all these axioms by a single statement: an n-dimensional vector space over a number field is the Cartesian product of n scalar sets over the same number field.

Six items. A vector space involves six items:

$$\left(V,+,0,F,\cdot,n\right).$$

- *V* is a collection of things, each of which is called a *vector*.
- + is a map from the Cartesian product $V\times V$ to *V*. That is, to every two elements *x* and *y* in *V*, there corresponds a unique element in *V*, written as $x+y$, called the *addition* of *x* and *y*. We say that the elements in *V* are *additive*, and that the set *V* is *closed* under addition.
- 0 is an *identity element* in *V* with respect to the addition. That is, there exists an element in *V*, written as 0 and called the *zero vector*, such that $x+0=x$ for every *x* in *V*.
- *F* is a number field. Elements in *F* are called *numbers*.
- $\cdot$ is a map from the Cartesian product $V\times F$ to *V*. That is, to every element α in *F* and every element *x* in *V*, there corresponds a unique element in *V*, written as $\alpha\cdot x$ (or just αx), called the *multiplication* of α and *x*. The vector αx results from *scaling* vector *x* by number α. We say that the elements in *V* are *scalable* by numbers, and that the set *V* is *closed* under the multiplication by numbers.
- *n* is a nonnegative integer called the *dimension* of the vector space. We also designate the dimension of the vector space by $\dim V$. For an *n*-dimensional vector space *V*, write $\dim V=n$. When *n* is finite, *V* is called a finite-dimensional vector space; otherwise, *V* is called an infinite-dimensional vector space.

Twelve axioms. A set *V* is called an *n*-dimensional vector space over a number field *F* if the following axioms hold.

1. $x+y\in V$ for every *x* and *y* in V.
2. $\left(x+y\right)+z=x+\left(y+z\right)$ for every *x*, *y* and *z* in *V*.
3. $x+0=x$ for every *x* in *V*.
4. For every *x* in *V* there exists an element (the negative element) *z* in *V* such that $x+z=0$. We write $z=-x$.
5. $x+y=y+x$ for every *x* and *y* in *V*.
6. $\alpha x\in V$ for every element α in *F* and every element *x* in *V*.
7. $1x=x$ for every *x* in *V*, where 1 is the identity element in *F* with respect to the multiplication of elements in F.
8. $\alpha\left(\beta x\right)=\left(\alpha\beta\right)x$ for every *x* in *V* and for every α, β in *F*.
9. $\left(\alpha+\beta\right)x=\alpha x+\beta x$ for every *x* in *V* and for every α, β in *F*.
10. $\alpha\left(x+y\right)=\alpha x+\alpha y$ for every α in *F* and every *x* and *y* in *V*.
11. There exist *n* elements $x_1,...,x_n$ in *V* such that, for any *n* elements $\alpha_1,...,\alpha_n$ in *F*, the equality $\alpha_1x_1+,...,+\alpha_nx_n=0$ implies that $\alpha_1=...=\alpha_n=0$.

12. For any $n+1$ elements $x_1,...,x_{n+1}$ in *V*, there exist elements $\alpha_1,...,\alpha_{n+1}$ in *F*, not all zero, such that $\alpha_1 x_1+,...,+\alpha_{n+1}x_{n+1}=0$.

Remarks

Two sets. The definition of a vector space mentions two sets, *V* and *F*. In most applications, *F* stands for either the set of real numbers *R*, or the set of complex numbers *C*. The vector space is called a *real vector space* if *F* = *R*, and is called a *complex vector space* if *F* = *C*.

The definition of vector space does not say what *V* is, or what vectors are. Whenever we find things in the world that fulfill the twelve axioms, we call these things vectors. Much of linear algebra is to discover and invent useful vector spaces. These notes will give many examples of vector spaces.

Four binary maps. A vector space *V* over a number field *F* involves four binary maps.

By the definition of number field, the set *F* is closed under two binary maps. Adding two elements in *F* gives a unique element in *F*:

$$F\times F\xrightarrow{\text{add}}F.$$

The number field *F* is a commutative group under addition. Multiplying two elements in *F* gives a unique element in *F*:

$$F\times F\xrightarrow{\text{multiply}}F.$$

The number field *F* with element zero removed is a commutative group under multiplication.

The definition of the vector space introduces two other binary maps. Axioms 1-5 do not mention the set *F*, and are devoted entirely to the set *V*. Adding two elements in *V* gives a unique element in *V*:

$$V\times V\xrightarrow{\text{add}}V.$$

Axioms 1-5 define *V* as a commutative group, with the vectors as the elements, addition as the operation, and the vector zero as the identity element. Axioms 6-8 are devoted to another binary map. Multiplying an element in *F* and an element in *V* gives a unique element in *V*:

$$F\times V\xrightarrow{\text{multiply}}V.$$

Axioms 9 and 10 say that the number-vector multiplication distributes over two both the number-number addition and the vector-vector addition.

Axioms 1-10 say how the two binary maps, vector-vector addition and number-vector multiplication should behave, by do not say what the two maps are. Much of linear algebra is to find maps in the world that fulfill Axioms 1-10. These notes will give many examples of vector-vector addition and number-vector multiplication.

Given two sets *V* and *F*, we can think of many possible binary maps to combine elements in the two sets. Most binary maps, however, do not enter the definition of the vector space. In particular, we exclude from the definition of the vector space any binary map that might represent the addition of an element in F and an element in V, or represent the multiplication of two elements in V.

The same words mean different things. We distinguish the addition of two numbers from the addition of two vectors, and distinguish the multiplication of two numbers from the multiplication a number and a vector. We call the four binary maps number-number addition, number-number multiplication, vector-vector addition, and vector-number multiplication.

The element zero in the set F is an object different from the element zero in the set V. The two objects have the same notation, 0. We tell them apart by seeing them in context.

Incidentally, "1" is the identity element for the multiplication on F. The definition of vector space does *not* introduce any identity element in V.

Zero-dimensional vector space. When $n = 0$, Axioms 11 and 12 define the zero-dimensional vector space V that contains a single element:

$$V = \{0\}.$$

This set also fulfills Axioms 1-10.

One-dimensional vector space is a scalar set. When $n = 1$, Axioms 11 and 12 are identical to the corresponding axioms defining a scalar set. That is, a one-dimensional vector space is a scalar set.

Any nonzero element in a vector space of any dimension is a unit of a scalar set. Let V be an n-dimensional vector space over a number field F. For any nonzero element x in V, define the set

$$S = \{ s | s = \alpha x, \alpha \in F \}.$$

We next confirm that the set S is a scalar set over F. The set S clearly fulfills Axioms 1-10 that define a scalar set. Because x is a nonzero element in S, the set S also fulfills Axiom 11. All elements in S are generated by multiplying x by numbers, so that S also fulfills Axiom 12. Thus, every nonzero element of a vector space of any dimension is a unit of a scalar set.

Tuples of Numbers

Let F be a number field and n be a positive integer. We now confirm that the Cartesian product F^n is an n-dimensional vector space over F. We identify the six items in $(V,+,0,F,\cdot,n)$ that fulfill the twelve axioms.

The set F^n plays the role of the set V. Recall that an n-tuple is an ordered list of n items, $(a_1,...,a_n)$. When each item is a number in F, the collection of all such n-tuples constitutes the Cartesian product F^n. Examples include the collection of all n-tuples of real numbers, R^n, and the collection of all n-tuples of complex numbers, C^n.

Tuples are additive. We define the addition of the two elements $(a_1,...,a_n)$ and $(b_1,...,b_n)$ in F^n as

$$(a_1,...,a_n)+(b_1,...,b_n)=(a_1+b_1,...,a_n+b_n).$$

The sign "+" appears on both sides of the equation with different meanings. On the right side, "+" means the addition of numbers, which comes from the definition of the number field *F*. On the left side, "+" means the addition of tuples. The equation thus defines the addition of the elements in F^n by using the addition of the elements in *F*. We further define the element zero in F^n by $(0,...,0)$, and define the negative element of $(a_1,...,a_n)$ by $(-a_1,...,-a_n)$. The addition of tuples defined this way satisfies Axioms 1-5.

Tuples are scalable by numbers. We define the multiplication of an element λ in *F* and an element $(a_1,...,a_n)$ in F^n as

$$\lambda(a_1,...,a_n)=(\lambda a_1,...,\lambda a_n).$$

This equation uses the multiplication of the elements in *F* to define the multiplication of elements in *F* and elements in F^n. The multiplication defined this way satisfies Axioms 6-8.

Multiplication distributes over additions of two types. The multiplication satisfies Axiom 9, which requires that the multiplication distribute over the addition in *F*. For $(a_1,...,a_n)$ in F^n and for every λ, β in *F*, we confirm the following equation:

$$(\lambda+\mu)(a_1,...,a_n)=\lambda(a_1,...,a_n)+\mu(a_1,...,a_n).$$

The multiplication satisfies Axiom 10, which requires that the multiplication distribute over the addition in F^n. For every λ in *F* and every $(a_1,...,a_n)$ and $(b_1,...,b_n)$ in F^n, we confirm the following equation:

$$\lambda\left[(a_1,...,a_n)+(b_1,...,b_n)\right]=\lambda(a_1,...,a_n)+\lambda(b_1,...,b_n).$$

F^n is an *n*-dimensional vector space. We finally confirm the dimension of F^n is n. To satisfy Axiom 11, consider n elements in F^n:

$$(1,0,...,0),(0,1,...,0),...,(0,0,...,1).$$

We confirm that the n elements are linearly independent. Any n elements $\alpha_1,\alpha_2,...,\alpha_n$ in *F* satisfy that

$$\alpha_1(1,0,...,0)+\alpha_2(0,1,...,0)+...+\alpha_n(0,0,...,1)=(\alpha_1,\alpha_2,...,\alpha_n).$$

So long as not all $\alpha_1,\alpha_2,...,\alpha_n$ are zero, $(\alpha_1,\alpha_2,...,\alpha_n)\neq(0,0,...,0)$.

To satisfy Axiom 12, consider any m elements in F^n:

$$(a_{11},a_{12},...,a_{1n}),$$
$$(a_{21},a_{22},...,a_{2n}),$$
$$......................,$$
$$(a_{m1},a_{m2},...,a_{mn}).$$

If $m>n$, we can confirm that the m elements are linearly dependent. This statement is evident by row reduction.

Tuples of Scalars

We next generalize the above example. The Cartesian product of scalar sets is a vector space. In this sense, scalar sets are the building blocks of linear algebra.

Cartesian product serves as *V*. Let $S_1,...,S_n$ be n scalar sets over a number field *F*. Recall that the Cartesian product of the scalar sets is defined as

$$S_1\times...\times S_n=\{(a_1,...,a_n)|a_1\in S_1,...,a_n\in S_n\}.$$

Write $V=S_1\times...\times S_n$. Thus, V is a set, each element of which is an n-tuple of scalars, $(a_1,...,a_n)$. We next confirm that *V* is an n-dimensional vector space over *F*.

Tuples are additive. Define the addition of two elements $(a_1,...,a_n)$ and $(b_2,...,b_n)$ in *V* as

$$(a_1,...,a_n)+(b_1,...,b_n)=(a_1+b_1,...,a_n+b_n).$$

The item a_1+b_1 is the addition of two elements in S_1, ..., and the item a_n+b_n is the addition of two elements in S_n. This equation uses the additions of elements in individual scalar sets to define the addition of elements in *V*. In general, the elements in each scalar set have no relation to the elements in another scalar set. In defining the addition of elements in *V*, we only need the additions for the individual scalar sets, and do not need to add an element in one scalar set to an element in another scalar set. The addition of elements in *V* so defined satisfies Axioms 1-5.

Tuples are scalable. Define the multiplication of a number λ in *F* and an element $(a_1,...,a_n)$ in *V* as

$$\lambda(a_1,...,a_n) = (\lambda a_1,...,\lambda a_n).$$

The item λa_1 means multiplying an element a_1 in S_1 by an element λ in *F*, ..., and the item λa_n means multiplying an element a_n in S_n by an element λ in *F*. The multiplication of elements in *V* and *F* so defined satisfies Axioms 6-10.

Cartesian product of *n* scalar sets is an *n*-dimensional vector space. To satisfy Axiom 11, we must find *n* linearly independent elements in *V*. Let a_1 be a nonzero element in S_1, a_2 be a nonzero element in S_2,..., and a_n be a nonzero element in S_n. Consider *n* elements in *V*:

$$(a_1,0,...,0)$$
$$(0,a_2,...,0)$$
$$......$$
$$(0,0,...,a_n)$$

Any numbers $\lambda_1,\lambda_2,...,\lambda_n$, not all zero, satisfy the inequality:

$$\lambda_1(a_1,0,...,0)+\lambda_2(0,a_2,...,0)+...+\lambda_n(0,0,...,a_n)\neq(0,0,...,0).$$

To satisfy Axiom 12, consider any *m* elements in *V*:

$$(a_{11},a_{12},...,a_{1n}),$$
$$(a_{21},a_{22},...,a_{2n}),$$
$$......................,$$
$$(a_{m1},a_{m2},...,a_{mn}).$$

If $m>n$, we can confirm that the *m* elements are linearly dependent. This statement is evident by row reduction.

Apple, Orange, Appleorange

To see how scalars build vectors, consider piles and piles of apples, piles and piles of oranges, and piles and piles of apples and oranges.

Apple. We have piles and piles of apples. We perform operations of two types. The addition of any two piles gives another pile having the same quantity of apples as the two piles put together. The multiplication of any pile of apples by any real number α corresponds to another pile α times the quantity of applies. The multiplication requires us to multiply apples by numbers, but does not require us to multiply apples by apples.

The piles and piles of apples form a set, denoted as APPLE. Each element in the set APPLE is a pile containing a distinct quantity of apples. We model the set APPLE as a scalar set over the field of real numbers.

Orange. We have piles and piles of oranges. We can similarly perform the two operations: add two piles of oranges, and multiple a pile of orange by a number. The piles and piles of oranges form a set, denoted as ORANGE. Each element in the set ORANGE is a pile containing a distinct quantity of oranges. We model the set ORANGE as a scalar set over the field of real numbers.

Appleorange. Any pile of apples and any pile of oranges can form an ordered pair, a pile of apples *and* oranges. All such ordered pairs form a new set, the *Cartesian product* of APPLE and ORANGE. We call the new set APPLEORANGE. Thus,

$$(\text{APPLEORANGE}) = (\text{APPLE}) \times (\text{ORANGE}).$$

Each element in APPLEORANGE is a pile containing some quantities of apples and some quantities of oranges. Any two piles differ in either the quantity of apples, or the quantity of oranges, or both.

We model the set APPLEORANGE as a two-dimensional vector space. We perform operations of two types. Adding two piles in APPLEORANGE means putting two piles together, resulting in yet another pile in the set APPLEORANGE. Multiplying a pile in APPLEORANGE and a real number α means finding in the set a pile α times the quantity of apples and α times the quantity of oranges.

The addition of two piles does not require us to add apples and oranges. Rather, in adding two piles, we add apples to apples, and oranges to oranges. Adding two piles in APPLEORANGE corresponds to adding the two types of things in parallel.

Space, Time, Spacetime*

Space. The physical space is a set, each element of which is called a *place*. The arrow from one place to another place is called a *displacement*. We model the collection of all displacements as a three-dimensional vector space over the field of real numbers. Denote this vector space by $V_{\text{displacement}}$.

Time. Time is a set, each element of which is called a *moment*. The arrow from one moment to another moment is called a *duration*. We model the collection of all durations as a scalar set over the field of real numbers. Denote this scalar set by V_{duration}.

Spacetime. Nothing stops us from lumping a place and a moment into a single object, which we call *event*. Thus, an event is an ordered pair of a moment and a place:

$$\text{event}=\left(\text{place, moment}\right).$$

An event happens at a place in space and a moment in time. The place tells us where the event happens, and the moment tells us when the event happens. The collection of all events constitutes a set, called *spacetime*. The set spacetime is the Cartesian product of two sets, space and time:

$$\left(\text{spacetime}\right)=\left(\text{space}\right)\times\left(\text{time}\right).$$

Minkowski (1908) called the Cartesian product of space and time the *world*. He called a ordered pair of a place and moment a *world-point*. When a bird flies in space and time, it goes through a sequence of world points—a subset of the world, which he called a *world-line*.

Four-dimensional vector space. Now consider two events. One event happens at a place and a moment, and the other event happens at another place and another moment. We now have two distinct objects: the displacement from one place to another place, and the duration from one moment to another moment. We lump the two objects together to form a single object, which we call *displacement-duration*, or *4-displacement*. The 4-displacement is an arrow from one event to the other event. The set of 4-displacements between all pairs of events constitute a four-dimensional vector space.

This four-dimensional vector space is the Cartesian product of two vector spaces: the three-dimensional vector space of displacements the one-dimensional vector space of durations:

$$\left(V_{\text{SPACETIME}}\right)=\left(V_{\text{displacement}}\right)\times\left(V_{\text{duration}}\right).$$

This Cartesian product defines the collection of all 4-displacements, $V_{\text{SPACETIME}}$, as a four-dimensional vector space.

Appleorange and Spacetime*

Cartesian product of appleorange and spacetime. We lump space and time into a single object, spacetime. Should we keep lumping? How about we lump spacetime and appleorange? Actually, such a lumping happens everyday. For example, when we ship apples and oranges to different places at different moments. We even have a name for such a lumped object. We call it a shipment: a pile of some quantity of apples and some quantity of oranges, shipped over some displacement and some duration. All shipments form a six-dimensional vector space.

Appleorange vs. spacetime. When we lump quantities of apples and quantities of oranges, the result does not surprise us. But when Minkowski lumped places and times, the result was shocking. Minkowski said, "Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."

What makes spacetime, but not appleorange, so shocking? The answer to this question is the key to the theory of relativity. The answer brings out another unity: the speed of light in vacuum is the same in all inertial frames. This experimental fact leads to the Minkowski metric in spacetime.

The Euclidean space differs from appleorange in two ways. First, the Euclidean space is an affine space, and appleorange is a vector space. Second, the Euclidean space has a matric, but appleorange does not. Spacetime is also an affine space with a metric, but its metric—the Minkowski metric—differs from the Euclidean metric. This difference will be explained later.

Cartesian Product of Vector Spaces

The above idea can be readily generalized. Let U and V be two vector spaces over a number field F. The Cartesian product of the two sets, $U \times V$, is also a vector space over F. Note that

$$\dim(U \times V) = \dim U + \dim V .$$

The idea works for any number of vector spaces over a number field.

In particular, if U is a vector space over a number field F, so is the Cartesian product U^m. Note that

$$\dim(U^m) = m \dim U .$$

We next look at several more examples.

Arrows in a solid. The set of arrows in a solid is a three-dimensional vector space over the field of real numbers. This example is of great importance: it is a model of our physical space. We may regard the set of arrows in a solid as the Cartesian product of arrows in three lines.

Complex numbers. In the set-building notation, we write the set of complex numbers as

$$C = \left\{ z \middle| z = a + bi, a \in R, b \in R \right\},$$

where R is the field of real numbers, and i is the unit of the imaginary numbers. The set C is a two-dimensional vector space over R.

Rational combinations of irrational numbers. A set A consists of numbers of the form $\alpha\sqrt{2} + \beta\sqrt{3} + \gamma\sqrt{5}$, where α, β and γ are rational numbers. In the set-building notation, the set is defined as

$$A = \left\{ x \middle| x = \alpha\sqrt{2} + \beta\sqrt{3} + \gamma\sqrt{5}, \alpha, \beta, \gamma \in Q \right\}.$$

This set is a three-dimensional vector space over the field of rational numbers, Q. We can confirm that the set A fulfills Axioms 1-10. Furthermore, the three elements of the set, $\sqrt{2}$, $\sqrt{3}$ and $\sqrt{5}$, are linearly independent, and any four elements of the set are linearly dependent.

Polynomials. Consider a polynomial

$$\alpha_0 + \alpha_1\xi + \alpha_2\xi^2 + ... + \alpha_n\xi^n .$$

The coefficients of the polynomial $\alpha_0, \alpha_1, \alpha_2, ..., \alpha_n$, as well as the variable ξ, are elements in a number field *F*. The set of all such polynomials is a vector space over the number field *F*.

In particular, consider the set of all quadratic polynomials, $ax^2 + bx + c$, where the variable is in the field of complex numbers, and the coefficients *a*, *b* and *c* are in the field of rational numbers. This set is a three-dimensional vector space over the field of rational numbers. The three elements in the vector space, $x^2 + 1$, $x + 3$ and 2, are linearly independent. However, the two elements in the vector space, $x^2 + 3$ and $2x^2 + 6$, are linearly dependent.

Linear Combination

Linear combination of vectors. Let *V* be a vector space over a number field *F*. For any vectors $u, v, ..., z$ in *V* and any numbers $\alpha, \beta, ..., \zeta$ in *F*, Axioms 1 and 6 ensure that the expression

$$\alpha u + \beta v + ... + \zeta z$$

is a vector in *V*. We call this vector a *linear combination* of the vectors $u, v, ..., z$ with coefficients $\alpha, \beta, ..., \zeta$.

Examples. *Arrows*. As noted before, the set of arrows in a solid is a vector field over the field real numbers. For any arrow *x* and any real number α, their multiplication αx is an arrow of the same direction as *x*, of length α times that of *x*. Similarly, for any arrow *y* and any real number β, their multiplication βy is an arrow of the same direction as *y*, of length β times that of *y*. The linear combination $\alpha x + \beta y$ is the diagonal of the parallelogram of sides αx and βy.

Polynomials. Consider the set of all quadratic polynomials, $ax^2 + bx + c$, where the variable is in the field of complex numbers, and the coefficients *a*, *b* and *c* are in the field of rational numbers. This set is a vector space over the field of rational numbers. $x^2 + 1$ is an element in the vector space, and $x + 3$ is another element of the vector space. A linear combination of the two elements, $6\left(x^2 + 1\right) + 7\left(x + 3\right)$, is an element in the vector space.

Linear Dependence and Linear Independence

Linear dependence. This notion appears in Axioms 11 and 12. Vectors $u, v, ..., z$ in *V* are said to be *linearly dependent* if there exist numbers $\alpha, \beta, ..., \zeta$ in *F*, not all zero, such that

$$\alpha u + \beta v + ... + \zeta z = 0 .$$

Linear independence. The vectors $u,v,...,z$ are said to be *linearly independent* if the equation

$$\alpha u + \beta v + ... + \zeta z = 0$$

holds only when all the numbers $\alpha,\beta,...,\zeta$ are zero.

Examples. *Vector zero*. Any list of vectors containing the vector zero are linearly dependent.

Arrows in a plane. Any arrows in parallel are linearly dependent. Any two arrows not in parallel are linearly independent. Any three arrows in a plane are linearly dependent.

Arrows in a solid. Any three arrows not in a plane are linearly independent. Any four arrows in a solid are linearly dependent.

Polynomials. The polynomials $1+x+x^2$ and $1+x+x^2+x^3$ are linearly independent.

Tuples. Any three ordered pairs are linearly dependent. In general, any $n+1$ number of n-tuples are linearly dependent.

Basis

Basis of a vector space. In an n-dimensional vector space V over a number field F, Axiom 11 says that there exist n linearly independent vectors. We call any list of n linearly independent vectors a *basis*, and call each vector in the list a *base vector*.

Thus, a basis is an n-tuple of vectors in V. Let $\left(e_1,...,e_n\right)$ be a basis—that is, a list of n linearly independent vectors in V. For $\beta^1,...,\beta^n$ in F, the equation

$$\beta^1 e_1 + ... + \beta^n e_n = 0$$

implies that $\beta^1 =,...,= \beta^n = 0$. The superscripts indicate n distinct numbers.

Any vector is a unique linear combination of the base vectors. To prove this statement, let $\left(e_1,...,e_n\right)$ be a basis of an n-dimensional vector space V over a number field F. Let x be a vector in V. Because V is an n-dimensional vector space, Axiom 12 says that the $n+1$ vectors x, $e_1,...,e_n$ are linearly dependent—that is, there exist numbers $\alpha,\beta^1,...,\beta^n$ in F, not all of which are zero, such that

$$\alpha x + \beta^1 e_1 + ... + \beta^n e_n = 0 .$$

The number α cannot be zero; otherwise, if α were zero, the above expression would become $\beta^1 e_1 + ... + \beta^n e_n = 0$, not all of $\beta^1,...,\beta^n$ being zero, implying that

$e_1,...,e_n$ were linearly independent and therefore were not a basis. This contradiction affirms that $\alpha \neq 0$. Dividing the above expression by α and rearranging, we obtain that

$$x = -\frac{\beta^1}{\alpha} e_1 - ... - \frac{\beta^n}{\alpha} e_n .$$

Consequently, any vector x is a linear combination of the base vectors.

We next prove that any vector x is a *unique* linear combination of the base vectors. Let $a^1,...,a^n$ be a list of numbers in F that satisfy $x = a^1 e_1 + ... + a^n e_n$, and let $b^1,...,b^n$ be a list of numbers that satisfy $x = b^1 e_1 + ... + b^n e_n$. The difference between the two linear combinations is $0 = \left(a^1 - b^1\right) e_1 + ... + \left(a^3 - b^3\right) e_n$. This equation implies that $a^1 = b^1,...,a^n = b^n$ because the base vectors $e_1,...,e_n$ are linearly independent.

Components of a vector relative to a basis. Thus, given a basis $e_1,...,e_n$ of an n-dimensional vector space V over a number field F, any vector x in V is a unique linear combination of the base vectors:

$$x = x^1 e_1 + ... + x^n e_n .$$

We call the n numbers $x^1,...,x^n$ in F the *components* of the vector x relative to the basis $\left(e_1,...,e_n\right)$.

The components of a vector constitute an n-tuple $\left(x^1,...,x^n\right)$ in F^n. The above equation defines a bijection between two sets, V and F^n. That is, once we choose a basis $\left(e_1,...,e_n\right)$, the equation $x = x^1 e_1 + ... + x^n e_n$ establishes a one-one correspondence between a vector x in V and a tuple $\left(x^1,...,x^n\right)$ in F^n:

$$x \leftrightarrow \left(x^1,...,x^n\right).$$

Examples. *Arrows in a plane*. The set of arrows in a plane is a two-dimensional vector space over the field of real numbers. Any two arrows not parallel to each other are linearly independent, and constitute a basis of this vector space. We designate two such arrows as e_1 and e_2. The linear combination $x = x^1 e_1 + x^2 e_2$ means that the arrow x is a diagonal of a parallelogram, one edge being in the direction of e_1 and of length x^1 times that of e_2, and the other edge being in the direction of e_1 and of length x^2 times that of the arrow e_2.

Arrows in a solid. The set of arrows in a solid is a three-dimensional vector space over the field of real numbers. Any three arrows not in the same plane are linearly independent, and constitute a basis of this vector space. We designate three such arrows as e_1, e_2 and e_3. The linear combination $x = x^1 e_1 + x^2 e_2 + x^3 e_3$ means that the arrow x is a diagonal of a parallelepiped, each of the three edges being in the direction of one base vector e_i, and of length x^i times that of e_i.

Polynomials. Consider the set of all quadratic polynomials, $ax^2 + bx + c$, where the variable x is in the field of complex numbers, and the coefficients a, b and c are in the field of rational numbers. This set is a three-dimensional vector space over the field of rational numbers. The three elements in this vector space, $\left(x^2, x, 1\right)$, are linearly independent. Consequently, the three elements can serve as a basis of this vector space. Any quadratic polynomial is a linear combination of this basis. The polynomial $2x^2 + 6x + 3$ is an element in the vector space, and has components 2, 6, 3 relative to the basis $\left(x^2, x, 1\right)$.

Vector space is Cartesian product of scalar sets. We have seen that the Cartesian product of n scalar sets is an n-dimensional vector space. We now show that the converse is also true: Any n-dimensional vector space is a Cartesian product of n scalar sets.

Let $\left(e_1, ..., e_n\right)$ be a basis of an n-dimensional vector space V over a field F. Each base vector is a unit of a scalar set. In the set-building notation, we define n scalar sets as

$$S_1 = \left\{ s \middle| s = \alpha e_1, \alpha \in F \right\},$$

......

$$S_n = \left\{ s \middle| s = \alpha e_n, \alpha \in F \right\}.$$

The Cartesian product of these scalar sets gives the vector space:

$$V = S_1 \times ... \times S_n .$$

That is, each element of V is an n-tuple of scalars:

$$V = \left\{ x \middle| x = \left(x^1 e_1, ..., x^n e_n\right), x^1, ..., x^n \in F \right\}.$$

Change of Basis

Two bases. A vector space has different bases, and the components of a vector change when the basis changes. Let $\left(e_1, ..., e_n\right)$ and $\left(\tilde{e}_1, ..., \tilde{e}_n\right)$ be two bases of an n-dimensional vector space V over a number field F. Each of the vectors $\tilde{e}_1, ..., \tilde{e}_n$ is an element in V, and is a linear combination of the basis $e_1, ..., e_n$, namely,

$$\tilde{e}_1 = p_1^1 e_1 + ... + p_1^n e_n$$
$$......$$
$$\tilde{e}_n = p_n^1 e_1 + .. + p_n^n e_n$$

These equations define n^2 numbers p_j^i in F that relate the two bases.

Components of one vector relative to two bases. A vector x in V is a linear combination of the vectors in the basis $e_1,...,e_n$:

$$x = x^1 e_1 + ... + x^n e_n .$$

The numbers $x^1,...,x^n$ in F are the components of the vector x relative to the basis $e_1,...,e_n$.

The same vector x is also a linear combination of the vectors in the basis $\tilde{e}_1,...,\tilde{e}_n$:

$$x = \tilde{x}^1 \tilde{e}_1 + ... + \tilde{x}^n \tilde{e}_n .$$

The numbers $\tilde{x}^1,...,\tilde{x}^n$ in F are the components of the vector x relative to the basis $\tilde{e}_1,...,\tilde{e}_n$.

Given two bases of V, $(e_1,...,e_n)$ and $(\tilde{e}_1,...,\tilde{e}_n)$, the vector x is a linear combination of base vectors in either basis. In the second equation above, replacing the basis $\tilde{e}_1,...,\tilde{e}_n$ with the basis $e_1,...,e_n$, we obtain that

$$x = \left(\tilde{x}^1 p_1^1 + ... + \tilde{x}^n p_n^1\right) e_1$$
$$+ ....$$
$$+ \left(\tilde{x}^1 p_1^n + ... + \tilde{x}^n p_n^n\right) e_n$$

Compare this equation to $x = x^1 e_1 + ... + x^n e_n$, and we obtain that

$$x^1 = \tilde{x}^1 p_1^1 + ... + \tilde{x}^n p_n^1$$
$$......$$
$$x^n = \tilde{x}^1 p_1^n + ... + \tilde{x}^n p_n^n$$

These equations relate the two sets of components of the same vector relative to two bases.

The change of the components of a vector is in a way opposition to the change of the basis of the vector space. We say that the vector is *contravariant*.

Example. The set of all quadratic polynomials is a three-dimensional vector space. We designate $(x^2, x, 1)$ as one basis, and $(x^2 + 2, 5x^2 + 7x, 4x + 3)$ as another basis. The second basis relate to the first basis through the coefficients

$$p_1^1 = 1, \quad p_1^2 = 0, \quad p_1^3 = 2$$
$$p_2^1 = 5, \quad p_2^2 = 7, \quad p_2^3 = 0$$
$$p_3^1 = 0, \quad p_3^2 = 4, \quad p_3^3 = 3$$

The polynomial $2x^2 + 6x + 3$ is an element in the vector space formed by all quadratic polynomials. The components of this element with respect to the basis $\left(x^2, x, 1\right)$ are (2, 6, 3). The same vector is a linear combination of the vectors in the other basis $\left(x^2 + 2, 5x^2 + 7x, 4x + 3\right)$:

$$2x^2 + 6x + 3 = (\quad)\left(x^2 + 2\right) + (\quad)\left(5x^2 + 7x\right) + (\quad)\left(4x + 3\right).$$

The three components of this element relative to the basis are left blank as an exercise.

Inverse Change of Basis

A change of basis is a bijection. Let *V* be an *n*-dimensional vector space over a number field *F*. The collection of all bases of *V* constitutes a set *B*, each member of which is a basis of *V* (i.e., a list of *n* linearly independent vectors in *V*).

We have used n^2 numbers p_j^i in *F* to change a basis $\left(e_1, ..., e_n\right)$ to a basis $\left(\tilde{e}_1, ..., \tilde{e}_n\right)$. The same n^2 numbers p_j^i in *F* can, of course, change any basis of *V* to another basis. Consequently, the n^2 numbers p_j^i define a map on *B*. The map sends one basis to another.

The map must be a bijection. Any basis can be changed to any other basis. In the above, we have changed from basis $\left(e_1, ..., e_n\right)$ to the basis $\left(\tilde{e}_1, ..., \tilde{e}_n\right)$. Because this change is a bijection, we can reverse the change.

Inverse change of basis. Each of the vectors $e_1, ..., e_n$ is an element in *V*, and is a linear combination of the basis $\tilde{e}_1, ..., \tilde{e}_n$. Write

$$e_1 = q_1^1 \tilde{e}_1 + ... + q_1^n \tilde{e}_n$$
$$......$$
$$e_n = q_n^1 \tilde{e}_1 + .. + q_n^n \tilde{e}_n$$

These equations define n^2 numbers q_j^i in *F* that relate the two bases.

The two sets of numbers p_j^i and q_j^i must be related. Combing the two changes, we obtain that

$$e_1 = q_1^1\left(p_1^1 e_1 + p_1^2 e_2 + ... + p_1^n e_n\right) + q_1^2\left(p_2^1 e_1 + p_2^2 e_2 + ... + p_2^n e_n\right) ... + q_1^n\left(p_n^1 e_1 + p_n^2 e_2 + ... + p_n^n e_n\right)$$
$$e_2 = q_2^1\left(p_1^1 e_1 + p_1^2 e_2 + ... + p_1^n e_n\right) + q_2^2\left(p_2^1 e_1 + p_2^2 e_2 + ... + p_2^n e_n\right) ... + q_2^n\left(p_n^1 e_1 + p_n^2 e_2 + ... + p_n^n e_n\right)$$
......
$$e_n = q_n^1\left(p_1^1 e_1 + p_1^2 e_2 + ... + p_1^n e_n\right) + q_n^2\left(p_2^1 e_1 + p_2^2 e_2 + ... + p_2^n e_n\right) ... + q_n^n\left(p_n^1 e_1 + p_n^2 e_2 + ... + p_n^n e_n\right)$$

Because each vector is a unique linear combination of the base vectors, the above equation must reduce to $e_1 = e_1, e_2 = e_2, ..., e_n = e_n$. We obtain a total of n^2 relations between the two sets of numbers p_j^i and q_j^i:

$$q_1^1 p_1^1 + q_1^2 p_2^1 + ... + q_1^n p_n^1 = 1, \quad q_1^1 p_1^2 + q_1^2 p_2^2 + ... + q_1^n p_n^2 = 0, \quad q_1^1 p_1^n + q_1^2 p_2^n + ... + q_1^n p_n^n = 0$$
$$q_2^1 p_1^1 + q_2^2 p_2^1 + ... + q_2^n p_n^1 = 0, \quad q_2^1 p_1^2 + q_2^2 p_2^2 + ... + q_2^n p_n^2 = 1, \quad q_2^1 p_1^n + q_2^2 p_2^n + ... + q_2^n p_n^n = 0$$
....
$$q_n^1 p_1^1 + q_n^2 p_2^1 + ... + q_n^n p_n^1 = 0, \quad q_n^1 p_1^2 + q_n^2 p_2^2 + ... + q_n^n p_n^2 = 0, \quad q_n^1 p_1^n + q_n^2 p_2^n + ... + q_n^n p_n^n = 1$$

It is hard to keep track of these equations. Later we will describe more convenient notation.

Summation Convention

Shorthand. Let *V* be an *n*-dimensional vector space over a number field *F*. Let $\left(e_1, ..., e_n\right)$ be a basis of *V*. Any vector *x* in *V* is a unique linear combination of the base vectors:

$$x = x^1 e_1 + ... + x^n e_n .$$

The components of the vector *x* relative to the basis $\left(e_1, ..., e_n\right)$ form an *n*-tuple $\left(x^1, ..., x^n\right)$ in F^n.

We can write the equation $x = x^1 e_1 + ... + x^n e_n$ in shorthand:

$$x = x^j e_j .$$

This way of writing follows a convention: The repeated index implies sum from 1 to *n*. The index for the sum repeats as a superscript and as a subscript. The value of the sum is independent of the choice of the symbol for the repeated index, *j*. Thus, the repeated index is called a dummy index.

Change of basis. Let $\left(\tilde{e}_1, ..., \tilde{e}_n\right)$ be another basis of *V*. The vector x is also a unique linear combination of the three new base vectors:

$$x = \tilde{x}^1 \tilde{e}_1 + ... + \tilde{x}^n \tilde{e}_n .$$

Similarly write the sum in shorthand:

$$x = \tilde{x}^i \tilde{e}_i .$$

The two bases are related as

$$\tilde{e}_1 = p_1^1 e_1 + ... + p_1^n e_n$$

......

$$\tilde{e}_n = p_n^1 e_1 + .. + p_n^n e_n$$

Using the summation convention, we write the relation between the two bases as

$$\tilde{e}_i = p_i^j e_j .$$

Here the index *j* is dummy, but the index *i* is not. This expression represents *n* distinct equations.

A combination of the above three equations gives

$$x^j = p_i^j \tilde{x}^i .$$

Inverse change of basis. We can also express each of the vectors $e_1,...,e_n$ as a linear combination of $\tilde{e}_1,...,\tilde{e}_n$:

$$e_i = q_i^k \tilde{e}_k .$$

This expression defines n^2 numbers q_i^k in *F* that relate the two bases.

This change of basis inverts the previous change, $\tilde{e}_k = p_k^j e_j$. A combination of two expressions gives that

$$e_j = q_j^k p_k^i e_i .$$

Consequently, the two sets of coefficients are related as

$$q_i^k p_k^j = \delta_i^j$$

where $\delta_i^j = 0$ when $i \neq j$, and $\delta_i^j = 1$ when $i = j$. Similarly, one can confirm that

$$p_i^k q_k^j = \delta_i^j .$$

Matrix

Change of basis. Recall the change of components associated with a change of basis:

$$x^1 = \tilde{x}^1 p_1^1 + ... + \tilde{x}^n p_n^1$$

......

$$x^n = \tilde{x}^1 p_1^n + ... + \tilde{x}^n p_n^n$$

The above expressions suggest yet another notation. List the components of the vector *x* relative to a basis as a column, list the components of the same vector *x* relative to the other basis as another column, and list the coefficients that transforming one basis to the other basis as a matrix. In this notation, we write the relation that transforms the two sets of components of the same vector as

$$\begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} p_1^1 & p_2^1 & p_3^1 \\ p_1^2 & p_2^2 & p_3^2 \\ p_1^3 & p_2^3 & p_3^3 \end{pmatrix} \begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix}.$$

Here we write the equation for a three-dimensional vector space. This expression means the same thing as the previous equation, a requirement that suggests the usual rule of matrix operation. A comparison of the two expressions also requires that the superscript indicate row, and the subscript indicate column.

Inverse change of basis. The components of the vector x relative to the two bases are also related by

$$\begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix} = \begin{pmatrix} q_1^1 & q_2^1 & q_3^1 \\ q_1^2 & q_2^2 & q_3^2 \\ q_1^3 & q_2^3 & q_3^3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

A combination of the above two relations gives that

$$\begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix} = \begin{pmatrix} q_1^1 & q_2^1 & q_3^1 \\ q_1^2 & q_2^2 & q_3^2 \\ q_1^3 & q_2^3 & q_3^3 \end{pmatrix} \begin{pmatrix} p_1^1 & p_2^1 & p_3^1 \\ p_1^2 & p_2^2 & p_3^2 \\ p_1^3 & p_2^3 & p_3^3 \end{pmatrix} \begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix}.$$

We obtain that

$$\begin{pmatrix} q_1^1 & q_2^1 & q_3^1 \\ q_1^2 & q_2^2 & q_3^2 \\ q_1^3 & q_2^3 & q_3^3 \end{pmatrix} \begin{pmatrix} p_1^1 & p_2^1 & p_3^1 \\ p_1^2 & p_2^2 & p_3^2 \\ p_1^3 & p_2^3 & p_3^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Two matrices p_j^i and q_j^i satisfying the above equation are said to be the *inverse* matrix of each other, and each matrix is said to be *invertible*. The matrix on the right-hand side is called the *identity matrix*.

Invertible matrices form a group. By definition, a number field F, with the element "0" excluded, is a commutative group, with the multiplication as the operation, and with the element "1" as the identity element.

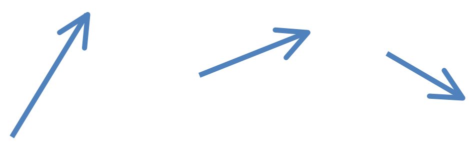
This statement can be generalized as follows. The collection of all n-by-n invertible matrices over a number field F forms a group, with the matrix multiplication as the operation, and with the identity matrix as the identity element. When $n = 1$, this statement recovers the previous statement, and the group is commutative. When $n > 1$, the matrix multiplication in general does not commutate, and the group is not commutative.

Arrows in a Plane

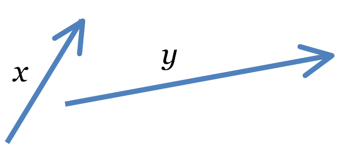
Arrows drawn in a plane form a two-dimensional vector space over the field of real numbers. This example is a special case of the Cartesian product scalar sets: the set of arrows in a plane is the Cartesian product of arrows in two lines.

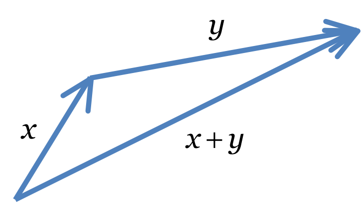
On the other hand, this example is so important that we examine it with care. We use the rules in plane geometry to specify addition and multiplication, and confirm that the arrows in the plane constitute a two-dimensional vector space over the field of real numbers.

Arrows in a plane constitute a set *V*. An arrow drawn on a piece of paper has a direction and a length. Two arrows are said to be the same element in a set *V* if one arrow can be translated onto the other.

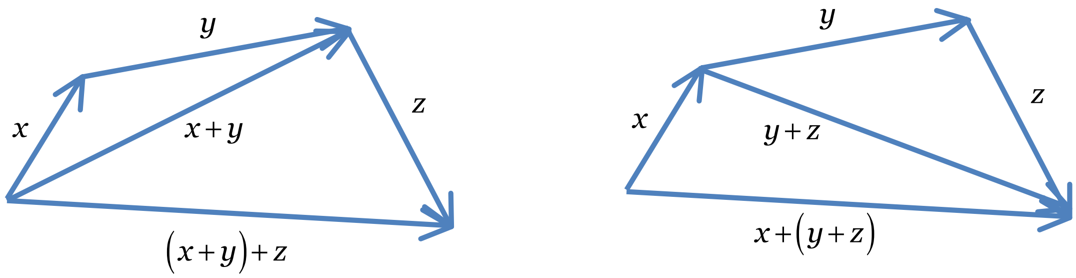


Arrows are additive. For two arrows x and y, plane geometry defines the addition $x+y$ as follows. Translate y to make the tail of y coincide with the head of x, and call the arrow from the tail of x to the head of y the addition $x+y$. The addition so defined clearly fulfills Axiom 1. That is, the procedure sends every two arrows in *V* to a unique arrow in *V*.

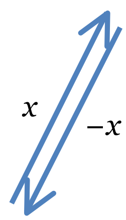




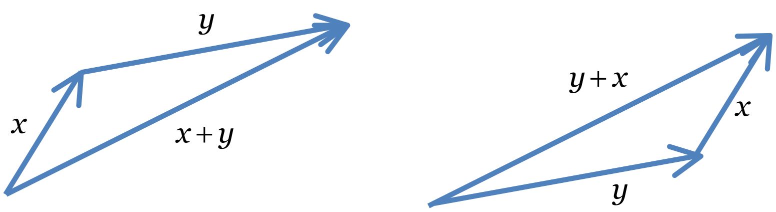
The procedure also fulfills Axiom 2. For every three arrows x, y, z, the procedure ensures that $(x+y)+z=x+(y+z)$. We translate the vectors to join the tail of one vector to the head of another. The arrow from the tail of the first vector to the head of the last vector gives the addition of all vectors.



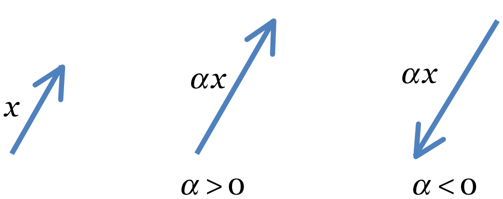
Plane geometry defines the "arrow 0" as the thing of zero length and unspecified direction. We can confirm Axioms 3 and 4. For an arrow x in V, $-x$ is an arrow parallel to x, of the opposite direction and the same length.



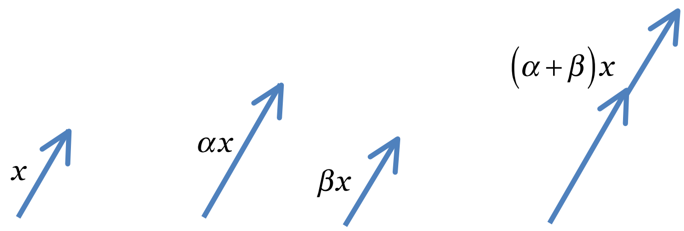
The addition so defined fulfills Axiom 5. That is, the procedure of addition is commutative.



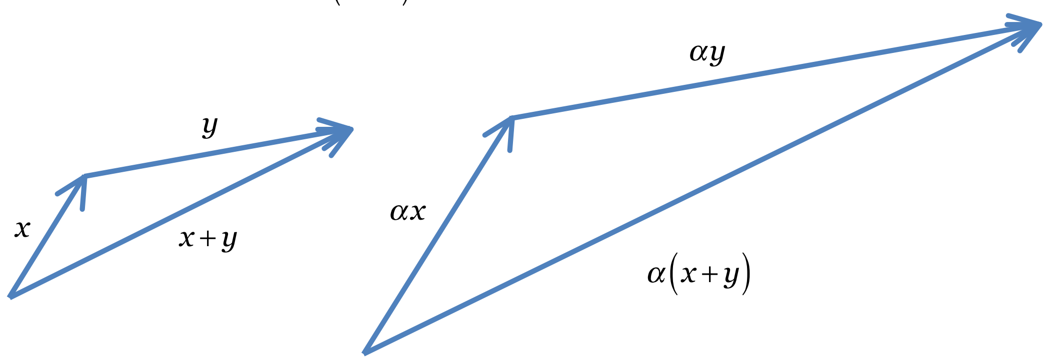
Arrows are scalable by real numbers. For a positive real number α and an arrow x, we define αx to be an arrow of the same direction as x, and of length α times that of x. For a negative number α, we define αx to be an arrow parallel to x, in the opposite direction, of length $|\alpha|$ times that of x. The scaling procedure so defined clearly fulfills Axioms 6-8.



Multiplication distributes over the two types of additions. To fulfill Axiom 9, we confirm that the multiplication distributes over the number-number addition, $(\alpha+\beta)x=\alpha x+\beta x$.

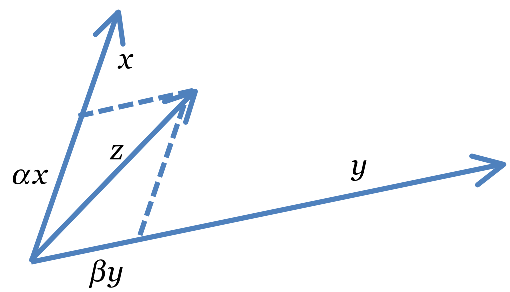


To fulfill Axiom 10, we confirm that the multiplication distributes over the vector-vector addition, $\alpha(x+y)=\alpha x+\alpha y$.



Arrows in a plane constitute a two-dimensional vector space. That is, arrows in a plane fulfill Axioms 11 and 12 for n = 2. Let us look at Axiom 11 first. Any two arrows x and y not parallel to each other are linearly independent. Otherwise, two real numbers would exist, not both zero, such that $\alpha x+\beta y=0$. Say $\alpha\neq 0$, so that $x=-(\beta/\alpha)y$, and x would be parallel to y.

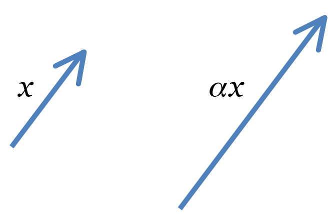
To fulfill Axiom 12, we need to confirm that any three arrows in a plane are linearly dependent. That is, for every three arrows in a plane, x, y and z, there exist three real numbers, α,β and γ, not all zero, such that $\alpha x+\beta y+\gamma z=0$. To see this, assume all the three vectors are nonzero; otherwise, if any one of the vectors is zero, say z = 0, we could choose $\alpha=0,\beta=0,\gamma\neq 0$ to make $\alpha x+\beta y+\gamma z=0$. Further assume that x and y are not parallel, otherwise, we could choose $\alpha=-1/\beta,\gamma=0$ to make $\alpha x+\beta y+\gamma z=0$. For any three nonzero vectors, we translate them such that their tails coincide. Because x and y are not parallel, from the head of z, we can draw a line parallel to y and intersecting the line containing x. Let αx be the arrow from the common tail to the intersection. Similarly, from the head of z, we can draw another line parallel to x and intersecting the line containing y. Let βx be the arrow from the common tail to the intersection. Recall the rule for addition, and this procedure gives that $\alpha x + \beta y - z = 0$.



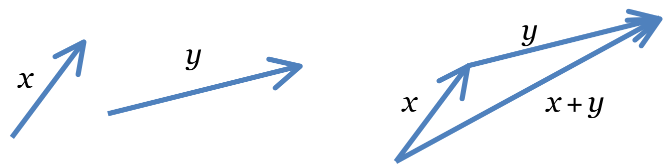
Ignore distractions. The set of arrows in a plane appeals to our intuition. Like any concrete example of an abstract concept, however, this example has distracting features (i.e., structures) that do not belong to the definition of vector space. For example, the definition of vector space does not contain the notion of the length and direction of a vector. But whenever we draw an arrow, the arrow has a length and a direction.

Representing Appleorange by Arrows

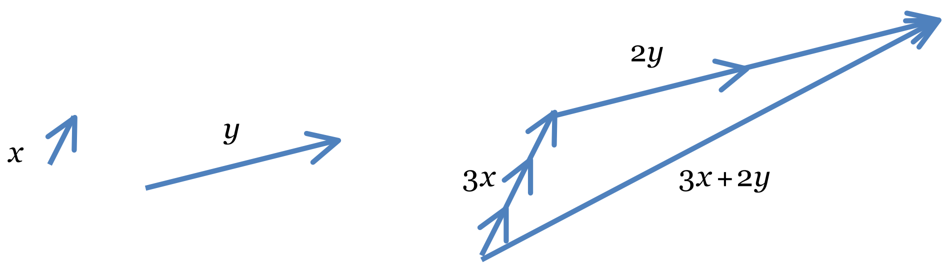
Represent piles in APPLEORANGE as arrows on a piece of paper. We can represent a pile x in APPLEORANGE as an arrow on a piece of paper. Any length and any orientation will work. To multiply pile x and a number α, we draw a new arrow in parallel to arrow x, but make the length of the new arrow a factor α times the length of arrow x. If $\alpha > 0$, the arrow αx is in the same direction as the arrow x. If $\alpha < 0$, the arrow αx is in the opposite direction as the arrow x. The locations of the two arrows do not concern us. Once we give a length and a direction to a pile x, all piles in APPLEORANGE proportional to x are parallel to x, and have the length proportional to the length of x.



We draw two piles x and y of disproportional quantities of apples and oranges as two arrows in different orientations. We have not defined metric in APPLEORANGE. Thus, we cannot relate the lengths of two arrows in different directions, or assign any meaning to the angle between the two arrows. The relative length and the angle between the two arrows have no significance. To add the two piles, we translate the two arrows to put the head of one arrow to the tail of the other arrow. The sum of the two piles is the arrow from the tail of one arrow to the head of the other.



Given two piles of disproportional quantities of apples and oranges, x and y, any other pile in APPLEORANGE z is a linear combination of the two given piles, $z = \alpha x + \beta y$, where α and β are numbers. Thus, once we represent two piles of disproportional quantities of apples and oranges, x and y, as two arrows on a piece of paper, we can represent any other pile by an arrow using the two given piles. We have the freedom to choose the lengths and orientations of the two arrows that represent piles x and y, but do not have freedom to choose the length and orientation of other piles.



Basis. The APPLEORANGE is a two-dimensional vector space. We can choose any two piles that contain disproportional quantities of apples and oranges. Every other pile is a *linear combination* of these two piles.

The two chosen piles are called *base vectors*. The two base vectors together form a *basis* in the vector space. Here is a basis: e_1 is a pile containing 1 apple and no orange, and e_2 is a pile containing no apple and 1 orange. Now consider a pile x that contains 9 apples and 5 oranges. This pile is a *linear combination* of the base vectors: $x = 9e_1 + 5e_2$. The numbers 9 and 5 are the *components* of the pile x relative to the basis e_1 and e_2.

Once two base vectors e_1 and e_2 are chosen, any vector x in the vector space is a linear combination of the base vectors:

$$x = x^1 e_1 + x^2 e_2 ,$$

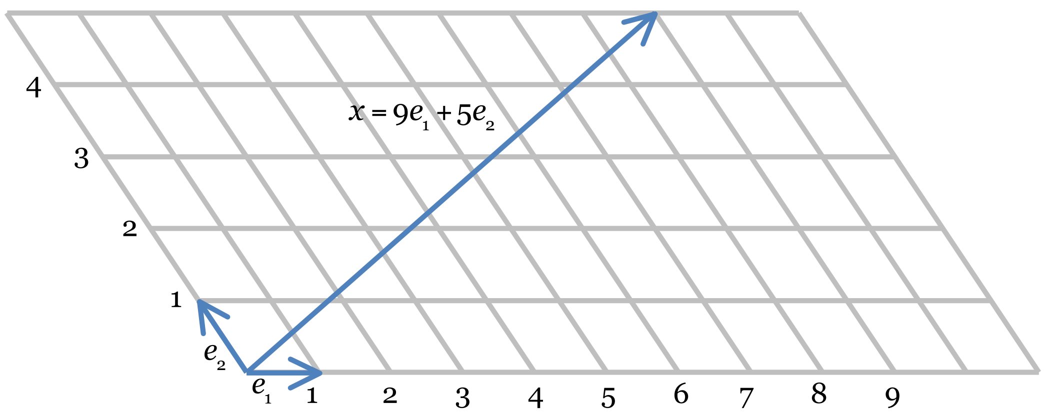
where the numbers x^1 and x^2 are the components of the pile relative to the basis e_1 and e_2. A quicker way of writing the above equation is

$$x = x^i e_i ,$$

We adopt the convention that a repeated index means a sum. We can also list the components of a vector as a column:

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}.$$

This column of two numbers represents a pile of apples and oranges. How many apples and oranges are in the pile? To answer this question, you need to know the two base piles e_1 and e_2, and return to the equation $x = x^1 e_1 + x^2 e_2$.



We represent the two base vectors e_1 and e_2 as two arrows on the piece of paper. The two arrows need not be of the same length, or be normal to each other. Indeed, we have not defined an inner product to relate the lengths of two vectors in different directions, or assign any meaning to the angle between the two vectors. In the direction of each base vector we draw a coordinate. The origin of the two coordinates represents zero amounts of apples and oranges. Along each coordinate we mark quantities proportional to the base vector. We also add a grid of parallel lines to guide the eye. Each arrow from the origin to a point in the plane represents a pile of apples and oranges, namely, a vector in APPLEORANGE. We draw the vector

$$x = 9e_1 + 5e_2,$$

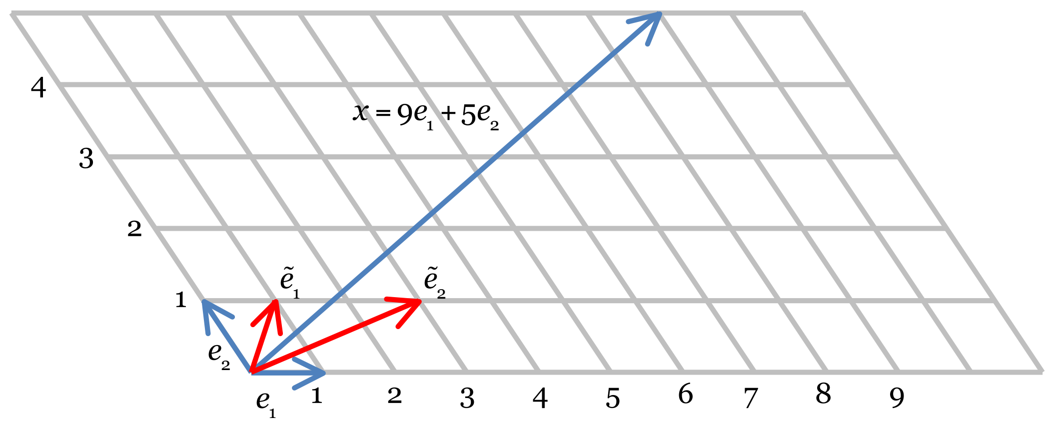
a pile of 9 apples and 5 oranges.

Change of basis. Here is another basis: $\tilde{e}_1$ is a pile containing 1 apple and 1 orange, and $\tilde{e}_2$ is a pile containing 3 apples and 1 orange. The same pile x is also a linear combination of the base vectors $\tilde{e}_1$ and $\tilde{e}_2$:

$$x = 3\tilde{e}_1 + 2\tilde{e}_2,$$

where the numbers 3 and 2 are components of the pile x relative to the basis $\tilde{e}_1$ and $\tilde{e}_2$. You can verify that x does contain 9 apples and 5 oranges. A vector is independent of the choice of basis, but the components of the vector are relative to the basis.

We add the two new base vectors $\tilde{e}_1$ and $\tilde{e}_2$ to the drawing, but decide not to add the coordinates and the grid associated with the new base vectors. To do so would make the drawing confusing. The graphic representation gets complicated and useless quickly.



Let's return to algebra. Each of the vectors $\tilde{e}_1$ is a pile of apples and oranges, and is a linear combination of the basis e_1 and e_2:

$$\tilde{e}_1 = p_1^1 e_1 + p_1^2 e_2 .$$

Similarly write

$$\tilde{e}_2 = p_2^1 e_1 + p_2^2 e_2 .$$

In our numerical example, the two equations are $\tilde{e}_1 = e_1 + e_2$ and $\tilde{e}_2 = 3e_1 + e_2$.

We can list the four numbers p_a^i as a matrix:

$$P = \begin{pmatrix} p_1^1 & p_1^2 \\ p_2^1 & p_2^2 \end{pmatrix}.$$

We call this matrix the *matrix of change of basis*.

Also write the transformation of the components of a vector $x^i = p_a^i \tilde{x}^a$ as two equations:

$$x^1 = p_1^1 \tilde{x}^1 + p_2^1 \tilde{x}^2 ,$$
$$x^2 = p_1^2 \tilde{x}^1 + p_2^2 \tilde{x}^2 .$$

Summary of our numerical example. In one basis, e_1 is a pile containing 1 apple and no orange, and e_2 is a pile containing no apple and 1 orange. In the other basis, $\tilde{e}_1$ is a pile containing 1 apple and 1 orange, and $\tilde{e}_2$ is a pile containing 3 apples and 1 orange. The change of basis is

$$\tilde{e}_1 = e_1 + e_2$$

$$\tilde{e}_2 = 3e_1 + e_2 .$$

Consider a pile x containing 9 apples and 5 oranges. This pile is a linear combination of set of base piles, $x = 9e_1 + 5e_2$. The same pile is also a linear combination of the other set of base piles, $x = 3\tilde{e}_1 + 2\tilde{e}_2$.

The transformation of the components of the vector—the pile of 9 apples and 5 oranges—is

$$\begin{pmatrix} 9 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

The matrix of the change of basis and the matrix of the transformation of the components of a vector are transpose to each other.

Subspace

As mentioned before, much of linear algebra is to discover and invent vector spaces. So far we have described one basic method to create vector space: any Cartesian product of scalar sets is a vector space. We now describe methods that rely on the idea of subspace.

Subspace. Let *V* be a vector space over a number field *F*. A subset *U* of *V* is called a subspace if the elements in *U* form a vector space under the same two operations—that is, the addition of two elements in *U* is an element in *U*, and the multiplication of an element in *F* and an element in *U* is an element in *U*.

Examples. *Trivial subspaces*. The vector space *V* is a subspace of *V*. The set containing only the zero vector is a subspace of any vector space.

Scalar set. Let *x* be a vector in a vector space *V* over a number field *F*. The set that contains all vectors of the form αx, where α is in *F*, is a subspace of *V*. This subspace is one-dimensional.

Arrows. The set of arrows in a solid is a three-dimensional vector space over the field of real numbers. Let *x* and *y* be two arrows that have different directions. The set of all linear combinations $\alpha x + \beta y$, where α and β are in *F*, is a two-dimensional subspace of the three-dimensional vector space. The subspace corresponds to a plane spanned by the two vectors *x* and *y*.

Mixed roots. A set consists of numbers of the form $\alpha\sqrt{2} + \beta\sqrt{3} + \gamma\sqrt{5}$, where α, β and γ are rational numbers. This set is a three-dimensional vector space over the field of rational numbers. The set consists all numbers of the form

$\alpha\sqrt{2}+\beta\sqrt{3}$ is a two-dimensional subspace. The set consists all numbers of the form $\alpha\left(\sqrt{2}+\sqrt{3}\right)$ is a one-dimensional subspace.

Span of vectors. Let *V* be a vector space over a number field *F*. The collection of all linear combinations of a list of vectors $u,v,...,z$ in *V* is called the span of the list of vectors, and is written as $\text{span}(u,v,...,z)$.

Thus, a span of a collection of vectors $u,v,...,z$ is a set, which can be defined using the set-building notation:

$$\text{span}(u,v,...,z)=\left\{\alpha u+\beta v...+\zeta z \middle| \alpha,\beta,...,\zeta\in F\right\}.$$

Examples. *Span of a single vector*. Let *u* be an arrow in a three-dimensional space. $\text{span}(u)$ is the collection of all vectors of the form αu, where α is any number in *F*. That is, $\text{span}(u)$ corresponds to all arrows in a line that contains the arrow *u*. The span of a single vector is a one-dimensional space, i.e., a scalar set.

Span of two vectors. Let *u* and *v* be two arrows in a three-dimensional space. $\text{span}(u,v)$ is the collection of all vectors of the form $\alpha u+\beta v$, where α and β are any numbers in *F*. So long as the two arrows *u* and *v* are not parallel, $\text{span}(u,v)$ corresponds to all arrows in a plane that contains the arrows *u* and *v*.

Span of powers. Note that 1, *x*, and x^2 are elements of the vector space of all polynomials. $\text{span}(1,x,x^2)$ is the collection of expressions of the form ax^2+bx+c, where *a*, *b*, and *c* are any numbers in *F*. That is, $\text{span}(1,x,x^2)$ is the set of all second order polynomials.

The span of vectors in *V* is a subspace of *V*. This statement means two things. First, the span of a list of vectors in *V* is subset of *V*. This is true because any linear combination of a list of vectors in *V* is a vector in *V*.

Second, the span is a vector space over the same number field and under the same addition and multiplication. This is true once we go over Axioms 1-10.

Inspecting Axiom 11 and 12, we can readily confirm that that a span of *k* linearly independent vectors in *V* is a *k*-dimensional vector space.

Intersection of subspaces is a subspace. Let *U* and *V* be subspaces of a vector space *W* over a number field *F*. Recall the definition of the union of two sets:

$$U\cap V=\left\{x \middle| x \text{ is in both } U \text{ and } V\right\}.$$

Because U and V are subsets of W, the intersection $U \cap V$ is a subset of U, V, and W. To confirm that $U \cap V$ is a vector space over F, we must show that elements in $U \cap V$ are closed under vector-vector addition and number-vector multiplication.

If $x \in (U \cap V)$ and $y \in (U \cap V)$, of course $x \in U$, $x \in V$, $y \in U$, and $y \in V$. Because U and V are vector spaces, $(x+y) \in U$ and $(x+y) \in V$. Consequently, $(x+y) \in (U \cap V)$.

If $x \in (U \cap V)$, of course $x \in U$ and $x \in V$. Because U and V are vector spaces, $\alpha x \in U$ and $\alpha x \in V$ for every α in F. Consequently, $\alpha x \in (U \cap V)$.

For example, arrows in a solid form a three-dimensional vector space W over the field of real numbers. Arrows in a plane form a subspace of W. If two planes U and V are not parallel, their intersection $U \cap V$ defines a line. Arrows in the line form a subspace of U, V, and W.

Union of subspaces may not be a vector space. Let U and V be subspaces of a vector space W over a number field F. Recall the definition of the union of two sets:

$$U \cup V = \left\{ x \middle| x \text{ is in either } U \text{ or } V \right\}.$$

The union is closed under the number-vector multiplication, but may not be closed under the vector-vector addition.

For example arrows in a solid form a three-dimensional vector space W over the field of real numbers. Arrows in a plane form a subspace of W. If two planes U and V are not parallel, their union $U \cup V$ defines two planes of arrows. The sum of an arrow u in U and an arrow v in V is an arrow in W, but not in $U \cup V$.

Sum of subspaces. Let U and V be subspaces of a vector space W over a number field F. The sum of the two subspaces, denoted by $U+V$, is the set whose elements are all the vectors $u+v$ with u in U and v in V. We can write this definition in the set-building notation:

$$U+V = \left\{ x \middle| x = u+v, u \in U, v \in V \right\}.$$

Examples. *Gold, silver, and platinum*. The set whose elements are all amounts of gold is a one-dimensional vector space, and the set whose elements are all amounts of silver is another one-dimensional vector space. Both vector spaces can be regarded as subspaces of a vector space whose elements are pieces containing gold, silver and platinum. The sum of the two subspaces is a two-dimensional vector space, which is a subspace of the vector space of gold, silver, and platinum.

Mixed roots. A set whose elements are all numbers of the form $\alpha\sqrt{2}+\beta\sqrt{3}$, where α,β are rational numbers, is a two-dimensional vector space U over the field of rational numbers. A set whose elements are all numbers of the form $\gamma\sqrt{3}+\xi\sqrt{5}$, where γ,ξ are rational numbers, is a two-dimensional vector space V over the field of rational numbers. Both vector spaces are subspaces of a vector space W whose elements are all numbers of the form $a\sqrt{2}+b\sqrt{3}+c\sqrt{5}+d\sqrt{7}$, where a, b, c, d are rational numbers. The sum $U+V$ is the set whose elements are all numbers of the form $a\sqrt{2}+b\sqrt{3}+c\sqrt{5}$, where a, b, c are rational numbers. This set is a three-dimensional subspace of W.

A sum of subspaces is a subspace. By their definitions, it is evident that the union of two subspaces is a subset of the sum of the two subspaces:

$$(U\cup V)\subset(U+V).$$

We have shown that $U\cup V$ is in general not a vector space. We now show that $U+V$ is a vector space.

If $x\in(U+V)$ and $y\in(U+V)$, by definition we can write $x=u+v$ where $u\in U$ and $v\in V$, and $y=a+b$ where $a\in U$ and $b\in V$. Because U and V are vector spaces, $(u+a)\in U$ and $(v+b)\in V$. By the definition of the sum of subspaces, $(u+a+v+b)\in(U+V)$. Note that $x+y=(u+a)+(v+b)$, so that $(x+y)\in(U+V)$.

If $x\in(U+V)$, by definition we can write $x=u+v$ where $u\in U$ and $v\in V$. Because U and V are vector spaces, $\alpha x\in U$ and $\alpha x\in V$ for every α in F. By the definition of the sum of subspaces, $(\alpha u+\alpha v)\in(U+V)$. Note that $\alpha x=\alpha u+\alpha v$, so that $\alpha x\in(U+V)$.

An identity of the dimensions of subspaces. Let U and V be subspaces of a vector space W over a number field F. We have just learned that $U\cap V$ and $U+V$ are also subspaces of W. The four subspaces obey an identity:

$$\dim(U+V)+\dim(U\cap V)=\dim U+\dim V.$$

For example, arrows in a solid form a three-dimensional vector space W over the field of real numbers. Arrows in a plane form a subspace of W. If two planes U and V are not parallel, their intersection $U\cap V$ defines a line of arrows. In this example,

$$U\times V=\left\{x\middle|x=(u,v),u\in U,v\in V\right\}.$$

They satisfy the above identity.

To prove the identity in general, let $\dim(U \cap V) = k$ and $(e_1,...,e_k)$ be k linearly independent vectors that span $U \cap V$. Then let $\dim U = m$ and $(e_1,...,e_k,u_{k+1},...,u_m)$ be m linearly independent vectors that span U, and let $\dim V = n$ and $(e_1,...,e_k,v_{k+1},...,v_n)$ be n linearly independent vectors that span V. By the definition of the sum of the vector spaces, if $x \in (U+V)$, we can write $x = u + v$ where $u \in U$ and $v \in V$. Thus, u is a linear combination of the vectors in the list $(e_1,...,e_k,u_{k+1},...,u_m)$, v is a linear combination of the vectors in the list $(e_1,...,e_k,v_{k+1},...,v_n)$, and $u+v$ is a linear combination of the vectors in the list $(e_1,...,e_k,u_{k+1},...,u_m,v_{k+1},...,v_n)$. Consequently, the list of vectors

$$(e_1,...,e_k,u_{k+1},...,u_m,v_{k+1},...,v_n)$$

spans the vector space $U+V$. These vectors are linearly independent. Counting the number of vectors in the list, we obtain that

$$\dim(U+V) = m+n-k,$$

which concludes the proof.

Direct sum of subspaces. Let U and V be subspaces of a vector space W over a number field F. The sum $U+V$ is called a direct sum, written as $U \oplus V$, if each element in $U+V$ is the addition of a unique element in U and a unique element in V.

For example, the sum of the space of gold and the space of silver is a direct sum. The sum of the space of $\alpha\sqrt{2}+\beta\sqrt{3}$ and the space $\gamma\sqrt{3}+\xi\sqrt{5}$ is not a direct sum. The sum of two planes of arrows is not a direct sum. The sum of a plane of arrows and a line of arrows, when the line is not in the plane, is a direct sum.

Direct sum and Cartesian product. The direct sum of subspaces U and V of a vector space W is also a subspace of W.

The sum of two subspaces U and V is a direct sum if and only if the two subspaces share only the zero vector, $U \cap V = \{0\}$. To see this, let $\dim U = m$ and $(u_1,...,u_m)$ be m linearly independent vectors that span U, and let $\dim V = n$ and $(v_1,...,v_n)$ be n linearly independent vectors that span V. The combined vectors

$$(u_1,...,u_m,v_1,...,v_n)$$

are linearly independent and span $U \oplus V$. The identity of dimensions reduces to

$$\dim(U \oplus V) = \dim U + \dim V.$$

Recall the definition of Cartesian product of two sets:

$$U \times V = \left\{ x \middle| x = (u,v), u \in U, v \in V \right\}.$$

If U and V are vector spaces over a number field F, the Cartesian product $U \times V$ is also a vector space. The dimensions of the three vector spaces obey an identity

$$\dim(U \times V) = \dim U + \dim V .$$

So far as vector spaces are concerned, the two concepts $U \oplus V$ and $U \times V$ are identical.

Map and Vector Space

Yet another method to create new vector spaces is through maps. Indeed, the collection of all maps from an arbitrary set to a vector space is a new vector space.

Maps from one vector space to another vector space. In single-variable calculus, we study maps that send a real number to another real number:

$$f : R \to R .$$

The collection of all such maps is a vector space over the field of real numbers. This vector space has infinite dimensions.

We can generalize this idea as follows. Let V and W be two vector spaces over a number field F. All maps from V to W constitute a vector space. When F is the field of real numbers, such maps are studied in multivariable calculus.

Maps from an arbitrary set to a number field. Let X be an arbitrary set, and F be a number field. All maps from X to F constitute a vector space over F. The number of elements in X is the dimension of the vector space. If X is a set of infinitely many elements, all maps from X to F constitute a vector space of infinite dimensions.

This example generalizes F^n, where $X = \{1,...,n\}$.

Maps from an arbitrary set to a vector space. Let X be an arbitrary set, and V be a vector space over a number field F. All maps from X to V constitute a vector space over F. We designate this vector space by the triple (X,V,F).

To substantiate the claim that the set of all maps from X to V is a vector space over F, we need to specify an addition of every two elements in the set (X,V,F), a multiplication of an element in (X,V,F) and an element in F, and then confirm that the two operations satisfy Axioms (1)-(10).

Recall that we write a map f from X and V in several ways; for example,

$$f : X \to V ,$$

$$x \mapsto f(x) .$$

Because X is an arbitrary set, the elements in X need not be additive to scalable. Because V is a vector space over F, the elements in V are additive and scalable.

Thus, for every two maps f and g in (X,V,F), for any element x in X, $f(x)$ and $g(x)$ are elements in V, so we will use the addition on V for the addition $f(x)+g(x)$. Similarly, for every map f in (X,V,F) and every number α in F, we will use the multiplication of an element in F and an element in V for the multiplication $\alpha f(x)$. The addition and multiplication so specified will satisfy Axioms (1)-(10).

The dimension of the vector space is the product of the number of elements in X and the dimension of the vector space V:

$$\dim(X,V,F)=|X|\dim(V).$$

Maps from the Cartesian product of arbitrary sets to the Cartesian product of vector spaces. This idea can be generalized as follows. Let X and Y be two arbitrary sets, and V be a vector space over a number field F. All maps from the Cartesian product $X\times Y$ to V constitute a vector space over F. We designate this vector space by the quadruple (X,Y,V,F), and note that

$$\dim(X,Y,V,F)=|X||Y|\dim(V).$$

The idea can be further generalized to any number of arbitrary sets.

Another generalization goes as follows. Let X be an arbitrary sets, and U and V be two vector spaces over a number field F. All maps from the Cartesian product X to $U\times V$ constitute a vector space over F. We designate this vector space by the quadruple (X,U,V,F), and note that

$$\dim(X,U,V,F)=|X|\left[\dim(U)+\dim(V)\right].$$