TENSORS

Contents

Number
Vector
Linear map
Operator
Scalar
Linear function
Linear form
Multilinear map
Inner product

Being linear within a set defines a vector space. Being linear between sets defines a linear map. A vector itself is a linear map. A linear map can be an element of a new vector space. We map the new vector space to some other vector spaces—that is, we map the maps. So goes the never-ending story of linear algebra.

These notes are incomplete, and have not been tested in classroom.

Number

Set. A collection of objects is called a set. Each object in the collection is called an element of the set.

In linear algebra, we will study many sets. We form each set by listing its attributes. We study elements within each set, and relate elements in one set to elements in another set.

Here we go with our first set. A set *F* is called a number field if the following conditions hold.

Adding two elements in *F* gives an element in *F*. To any two elements α and β in *F* there corresponds an element in *F*, written as $\alpha+\beta$, called the addition of α and β. The addition obeys the following rules.

1) Addition is commutative: $\alpha+\beta=\beta+\alpha$ for every α and β in *F*.

2) Addition is associative: $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ for every α, β and γ in *F*.

3) There exists an element in *F*, called 0 (zero), such that $0+\alpha=\alpha$ for every α in *F*.

4) For every α in *F*, there exists an element γ in *F*, such that $\alpha+\gamma=0$. We also write $\gamma=-\alpha$.

Multiplying two elements in *F* gives an element in *F*. To any two elements α and β in *F* there corresponds an element in *F*, written as $\alpha\beta$, called the multiplication of α and β. The multiplication obeys the following rules.

5) Multiplication is commutative: $\alpha\beta=\beta\alpha$ for every α and β in *F*.

6) Multiplication is associative: $(\alpha\beta)\gamma=\alpha(\beta\gamma)$ for every α, β and γ in *F*.

7) There exists an element in *F*, called 1, such that $1\cdot\alpha=\alpha$ for every α in *F*.

8) For every $\alpha\neq 0$ in *F*, there exists an element γ in *F*, such that $\alpha\gamma=1$. We also write $\gamma=1/\alpha$.

Multiplication is distributive over addition. A final rule involves both operations of addition and multiplication:

9) $\gamma(\alpha+\beta)=\gamma\alpha+\gamma\beta$ for every α, β and γ in *F*.

Remark. We call each element in the set *F* a number. Addition and multiplication are two binary maps. Each binary map turns two elements in *F* into an element in *F*, namely, $(F,F)\to F$. That is, the set *F* is closed under the two binary maps. We need to memorize nothing new: the two binary maps follow the usual arithmetic rules.

Example. The set of rational numbers is a field. The set of real numbers is a field. The set of complex numbers is a field. The field of complex numbers contains the field of real numbers. The field of real numbers contains the field of rational numbers.

Counter example. The set of integers is not a field, because the set violates Axiom 8.

Example. A set consists of all numbers of the form $\alpha + \beta\sqrt{2}$, where α and β are rational numbers. We adopt the usual definition of addition and multiplication for real numbers. This set is a number field. This field is contained in the field of real numbers.

Example. A set consists of all numbers of the form $\alpha + \beta i$, where α and β are rational numbers, and $i^2 = -1$. We adopt the usual definition of addition and multiplication for complex numbers. This set is a number field. This field is contained in the field of complex numbers.

Counter example. A set consists of all numbers of the form $\alpha\sqrt{2} + \beta\sqrt{3} + \gamma\sqrt{5}$, where α, β and γ are rational numbers. We adopt the usual definition of addition and multiplication for real numbers. The multiplication of two elements in the set does not always give another element in the set. This set is not a number field.

Counter example. Each element in a set is a piece of gold of some amount. We define the addition of two pieces in the set by melting them together, resulting in a piece in the set. However, we do not have a sensible definition of the multiplication that makes the multiplication of two pieces of mass into another piece of mass. This set is not a number field.

Vector

A set V is a vector space over a number field F if the following conditions hold.

Adding two elements in *V* gives an element in *V*. To any two elements x and y in V there corresponds an element in V, written as $x+y$, called the addition of x and y. Addition obeys the following rules:

1) Addition is commutative: $x+y=y+x$ for every x and y in V.

2) Addition is associative: $(x+y)+z=x+(y+z)$ for every x, y and z in V.

3) There exists an element 0 (the zero element) in V such that $0+x=x$ 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 $z=-x$.

Multiplying an element in *F* and an element in *V* gives an element in *V*. To every element α in F and every element x in V there corresponds an element in V, written as αx, called the multiplication of α and x. The multiplication obeys the following rules:

5) $1 \cdot x = x$ for every x in V.

6) $\alpha(\beta x)=(\alpha\beta)x$ for every x in V and for every α, β in F.

Multiplication is distributive over addition. The multiplication of an element in F and an element in V is distributive over two types of addition: addition of elements in F and addition of elements in V.

7) $(\alpha+\beta)x=\alpha x+\beta x$ for every x in V and for every α, β in F.

8) $\alpha(x+y)=\alpha x+\alpha y$ for every α in F and every x and y in V.

Remark. We call each element in F a number, and each element in V a vector. A vector space over a number field consists of two sets and four binary maps. The set F is closed under two binary maps: adding two numbers gives a number, $(F,F)\to F$; multiplying two numbers give a number, $(F,F)\to F$. The set V is closed under two other binary maps: adding two vectors gives a vector, $(V,V)\to V$; multiplying a number and a vector gives a vector, $(F,V)\to V$.

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, but we can tell them apart when we see them in any context. Likewise, 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.

A number field F is a vector space over itself. The definitions of number field and vector space have many similarities. Indeed, the number field F is a vector space over itself. For any two elements x and y in F, the addition $x+y$ is an element in F. The rules for the addition of two numbers are identical to the rules for the addition of two vectors. For any two elements α and x in F, the multiplication αx is an element in F. The rules of this multiplication are identical to the rules of the multiplication of a number and a vector, if we regard α as a number and x as a vector.

The number field is a peculiar vector space. We now turn to more representative vector spaces.

Example. An arrow drawn on a piece of paper is a directed segment. Two directed segments are regarded as the same element in a set if one of them can be translated onto the other. We follow the rules in plane geometry. For two directed segments x and y, we define $x+y$ to be the diagonal of the parallelogram of sides x and y. For a positive real number α and a directed segment x, we define αx to be a directed segment of the same direction as x, and of length α times that of x. For a negative number α, we define αx to be a directed segment parallel to x, in the opposite direction, of length $|\alpha|$ times that of x. To complete the set, we include the element zero, which has zero length and unspecified direction. The set of directed segments in a plane is a vector space over the field of real numbers.

Example. The set of directed segments in a solid is a vector space over the field of real numbers. This example is of great importance: it is a model of our physical space.

Example. The set of directed segments in a straight line is a vector space over the field of real numbers. This example is of great importance: it is a model of physical quantities like amount of time, volume, mass, energy, matter, and electric charge. It also models population, and amount of money.

Remark. The set of directed segments in a plane appeals to our intuition. Like any example of an abstract concept, however, this example has distracting features that have nothing to do with the definition of the vector space. The definition of vector space does not limit a vector to a geometric object.

Example. 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. As we have noted before, this set is not a number field. This set, however, is a vector space over the field of rational numbers.

Example. A polynomial of order up to a positive integer n takes the form

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

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*.

Linear combination. Given two sets *V* and *F*, we can think of many possible binary maps. Most binary maps, however, do not appear in 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 two binary maps that we do use define the notion of being linear. In particular, they lead to the idea of linear combination of elements in the set *V*.

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

Example. As noted before, the set of directed segments in a solid is a vector field over the field real numbers. For any segment *x* and any real number α, their multiplication αx is a segment of the same direction as *x*, of length α times that of *x*. Similarly, for any segment y and any real number β, their multiplication βy is a segment 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.

Example. 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 independence. Vectors $u, v, ..., z$ in *V* are said to be linearly dependent if there exist numbers $\alpha, \beta, ..., \zeta$ in *F*, not all of which are zero, such that

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

The vectors $u, v, ..., z$ are said to be linearly independent if the above equation holds only when all the numbers are zero.

Dimension of a vector space. A vector space V over a number field F is said to be n-dimensional if there exist n linearly independent vectors in V, but any $n+1$ vectors in V are linearly dependent.

The set containing only the zero vector is a zero-dimensional vector space over any number field. This is the only type of zero-dimensional vector space.

Example. The set of directed segments in a plane is a two-dimensional vector space over the field of real numbers. Any two segments in the same direction are linearly dependent. Any two segments, not parallel to each other, are linearly independent. Any three segments regardless their directions are linearly dependent.

The set of directed segments in a solid is a three-dimensional vector space over the field of real numbers.

The set of directed segments in a straight line is a one-dimensional vector space over the field of real numbers.

The set of directed intervals of time is a one-dimensional vector space.

A number field F is a one-dimensional vector space over F.

Example. The set of numbers of the form $\alpha+\beta\sqrt{2}$, where α and β are rational numbers, is a two-dimensional vector space over the field of rational numbers. The two elements of the set, 1 and $\sqrt{2}$, are linearly independent. Any three elements of the set are linearly dependent.

Example. 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.

A set of gold. Each element in a set is a piece of gold of some amount. Adding two pieces in the set means putting them together, resulting in a piece in the set. We have noted before that this set is not a number field, because we do not have a sensible definition of the multiplication of two pieces in the set. However, we can readily define the multiplication of a piece and a number: multiplying a piece by a real number r means finding in the set a piece r times the amount.

We model this set as a one-dimensional vector space over the field of real numbers. We do so with caution. The definition of the vector space requires that a piece times any real number be still a piece in the set. If the number is too large, we may not have that much gold. If the number is too small, we may reach subatomic dimension, in which case the "piece" is no longer gold. Gold is made of atoms; they come in discrete lumps. This physical fact requires us to use integers, but integers do not form a number field. Also, the definition of the vector space will require that negative amounts gold be in the set.

In representing a physical phenomenon with a mathematical model, we may choose to ignore these inconvenient truths initially. Once we obtain a prediction, we can check what it means. For example, if the model gives a negative amount of gold, it means that we are in deficit.

Apples and oranges. Each element in a set is a pile containing some number of apples and some number of oranges. Adding two piles means putting them together, resulting in a pile in the set. Multiplying a pile and a real number *r* means finding in the set a pile *r* times the amount. We model each pile as a vector, and model the set of piles as a two-dimensional vector space over the field of real numbers.

A vector lumps different objects together as a single entity. A pile containing apples and oranges is a vector. The addition of two vectors does not require us to add apples and oranges. Rather, in adding two piles, we add apples to apples, and oranges to oranges. The addition of vectors generalizes the addition of numbers: adding two vectors corresponds to adding two lists of numbers in parallel.

Mass and volume. We can also list different physical quantities together as a single object. Consider a set, each element of which is a piece of some mass and some volume. Adding two pieces means putting them together, resulting in a piece in the set. Multiplying a piece and a real number *r* means finding in the set a piece *r* times the amount. This set is a two-dimensional vector space over the field of real numbers.

Basis. In an n-dimensional vector space *V* over a number field *F*, there exist n linearly independent vectors. We call any list of n linearly independent vectors a basis of *V*.

Components of a vector relative to a basis. We now establish an important theorem: any vector *x* is a unique linear combination of the vectors in a basis. Let $e_1,...,e_n$ 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, 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 superscripts indicate n distinct numbers. 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.

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 vectors in a basis.

To prove the uniqueness, let $a^1,...,a^n$ be a list of numbers in F that satisfy $x = a^1e_1 + ... + a^ne_n$, and let $b^1,...,b^n$ be a list of numbers that satisfy $x = b^1e_1 + ... + b^ne_n$. 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 vectors $e_1,...,e_n$ are linearly independent.

Thus, given a basis $e_1,...,e_n$, a vector x is a unique linear combination of the vectors in the basis:

$$x = x^1e_1 + ... + x^ne_n .$$

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

Example. The set of directed segments in a plane is a two-dimensional vector space over the field of real numbers. Any two segments not parallel to each other are linearly independent, and constitute a basis of this vector space. We designate two such segments as e_1 and e_2. The linear combination $x = x^1e_1 + x^2e_2$ means that the segment 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 segment e_2.

Example. The set of directed segments in a solid is a three-dimensional vector space over the field of real numbers. Any three segments not in the same plane are linearly independent, and constitute a basis of this vector space. We designate three such segments as e_1, e_2 and e_3. The linear combination $x = x^1e_1 + x^2e_2 + x^3e_3$ means that the segment 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.

Example. 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,

$x^2, x, 1$, are linearly independent. Consequently, the three elements can serve as a basis of this vector space—that is, 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 $x^2, x, 1$.

Gold and silver. Each element in a set is a piece containing some amount of gold and some amount of silver. Adding two pieces means putting them together, resulting a piece in the set. Multiplying a piece by a real number *r* means finding in the set a piece *r* times the amount. This set is a two-dimensional vector space over the field of real numbers.

A basis of this vector space consists of two pieces in the set. To ensure that the vectors in the basis are linearly independent, we pick two pieces with disproportional amounts of gold and silver. Here is a basis: e_1 is a piece containing 1 gold atom and no silver atom, and e_2 is a piece containing no gold atom and 1 silver atom. Here is another basis: $\tilde{e}_1$ is a piece containing 2 gold atoms and 3 silver atoms, and $\tilde{e}_2$ is a piece containing 5 gold atoms and 7 silver atoms.

The two bases are related as

$$\tilde{e}_1 = 2e_1 + 3e_2$$
$$\tilde{e}_2 = 5e_1 + 7e_2$$

We can also invert this relation and write

$$e_1 = -7\tilde{e}_1 + 3\tilde{e}_2$$
$$e_2 = 5\tilde{e}_1 - 2\tilde{e}_2$$

Now consider a particular piece *x* that contains 11 gold atoms and 13 silver atoms. This piece is a linear combination of the base vectors e_1 and e_2:

$$x = 11e_1 + 13e_2 .$$

The numbers 11 and 13 are components of the piece *x* relative the basis e_1 and e_2. The same piece *x* is also a linear combination of the base vectors $\tilde{e}_1$ and $\tilde{e}_2$:

$$x = -12\tilde{e}_1 + 7\tilde{e}_2 .$$

The numbers -12 and 7 are components of the piece *x* relative to the basis $\tilde{e}_1$ and $\tilde{e}_2$.

Change of basis. The above example illustrates two general facts: a vector space has different bases, and the components of a vector change when the basis changes. Let $e_1,...,e_n$ and $\tilde{e}_1,...,\tilde{e}_n$ be two bases of an *n*-dimensional vector space V on 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 a vector relative to different bases. A vector x in V is a linear combination of either basis:

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

These equations define numbers $x^1,...,x^n$ and $\tilde{x}^1,...,\tilde{x}^n$ in F as the components of the vector x relative to the two bases.

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 $\left(x^2, x, 1\right)$ as one basis, and $\left(x^2+2, 5x^2+7x, 4x+3\right)$ 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 $\left(x^2+2,5x^2+7x,4x+3\right)$:

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

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

Matrix notation. The above expressions suggest a commonly used 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

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

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.

Summation convention. We write this sum $x=x^1e_1+..+x^ne_n$ as

$$x=x^je_j.$$

The index for the sum repeats as a superscript and as a subscript. The repeated index implies sum from 1 to *n*. The value of the sum is independent of the choice of the symbol for the repeated index, *j*. Thus, the repeated index is known as the dummy index. Similarly write the sum $x=\tilde{x}^1\tilde{e}_1+...+\tilde{x}^n\tilde{e}_n$ as

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

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

$$\tilde{e}_i=p_i^je_j.$$

Here the index *j* is dummy, but the index *i* is not. This expression represents three distinct equations. A combination of the above three equations gives

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

We have written the linear map in three equivalent ways: in longhand, using a matrix, and using the summation convention.

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

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

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

Example. The vector space *V* is a subspace of *V*.

Example. The set containing only the zero vector is a subspace of any vector space.

Example. 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.

Example. The set of directed segments in a solid is a three-dimensional vector space over the field of real numbers. Let *x* and *y* be two directed segments 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*.

Example. 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.

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 $x+y$ with x in U and y in V. A sum of subspaces of a vector space W is also a subspace of W.

Example. 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.

Example. 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.

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 if each element in $U+V$ is the addition of a unique element in U and a unique element in V. The direct sum of subspaces U and V of a vector space W is also a subspace of W, and is written as $U \oplus V$.

The sum of two subspaces U and V is a direct sum if the two subspaces share only the zero vector, $U \cap V = \{0\}$.

Example. In the two examples above, 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.

An *n*-dimensional vector space is a direct sum of *n* one-dimensional vector spaces. Let V be an n-dimensional vector space over a number field F, and let $e_1,...,e_n$ be a basis. The set U_1 of all the vectors of the form αe_1 with α in F is a one-dimensional subspace of V. Similarly, the set U_2

of all the vectors of the form βe_2 with β in F is also a one-dimensional subspace of V. The sum of all n such one-dimensional subspaces is a direct sum, and the result of the direct sum is the vector space V: $U_1 \oplus U_2 \oplus \ldots \oplus U_n = V$.

For a vector x in V, the set of all the vectors of the form αx with α in F is a one-dimensional subspace of V. Any n-dimensional vector space is a direct sum of n one-dimensional vector spaces.

We can form the vector space of gold-silver pieces by the direct sum of the one-dimensional vector space of gold pieces and the one-dimensional vector space of silver pieces. The result of the direct sum is a two-dimensional vector space. As noted before, a basis of this two dimensional vector space consists of two pieces: e_1 is a piece containing 1 gold atom and no silver atom, and e_2 is a piece containing no gold atom and 1 silver atom. This basis happens to be obtained directly from the bases of the two one-dimensional vector spaces.

Once the two-dimensional vector space is formed by the direct sum, however, we are not obliged to use the bases of the subspaces. As noted before, here is another basis of the two-dimensional vector space: $\tilde{e}_1$ is a piece containing 2 gold atoms and 3 silver atoms, and $\tilde{e}_2$ is a piece containing 5 gold atoms and 7 silver atoms.

The Minkowski spacetime is a direct sum of space and time, which may be regarded as a three-dimensional subspace and a one-dimensional subspace. Once the four-dimensional space is formed, we can use any four linearly independent vectors as a basis. Each choice of a basis is called an inertial frame of reference.

Linear Map

Map. Let V and W be two sets. With each element v in V, a map A associates an element w in W. We write this relation as

$$w = A(v).$$

We also write this relation as

$$A : V \to W.$$

We call the input v an argument, and the output w a value. We call the set V the domain, and the set W the codomain.

Let V' be a subset of the domain V. The elements in W associated with all elements in V' constitute a subset in the codomain W, and the subset is called the image of V' under A, written as $A(V')$.

Let W' be a subset of the codomain W. The elements in V associated with all elements in W' constitute a subset in the codomain V, and the subset is called the preimage of W' under A, written as $A^{-1}(W')$.

Injection, surjection and bijection. A map is injective (one-to-one) if it sends each element in the domain to a distinct element in the codomain. That is, for every two distinct elements x and y in V, $A(x)$ and $A(y)$ are distinct elements in W.

A map is surjective (onto) if it sends at least one element in the domain to every element in the codomain. That is, for every element w in W, there is at least one element v in V to satisfy $w = A(v)$.

A map is bijective (one-to-one correspondent) if it is both injective and surjective.

Linear map. A map $A : V \to W$ is a linear map if V and W are vector spaces over a number field F, and if the following conditions hold.

1. $A(x+y) = A(x) + A(y)$ for any x and y in V
2. $A(\alpha x) = \alpha A(x)$ for any α in F and any x in V

We write a linear map $A(x)$ as Ax if it does not cause confusion.

Remark. To define a linear map, we require that the two sets V and W be vector spaces. Thus, a linear combination of elements in V is an element in V, and a linear combination of elements in W is an element in W. Without these requirements, the conditions 1 and 2 are meaningless.

The two conditions are equivalent to a single statement

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$$

for any α and β in *F*, as well as any x and y in *V*. With a linear combination of two elements in *V*, the linear map *A* associates a linear combination of two elements in *W*.

Example. The set of directed segments in a plane is a vector space over real numbers. Now we paint the segments on a sheet of rubber. A homogeneous stretch of the sheet transforms every segment into a segment. The homogeneous stretch is a linear map that maps a segment in the upstretched sheet to a segment in the stretched sheet.

We can demonstrate a similar map by painting segments on a rubber band. When the rubber band is stretched, the segments become longer.

We can also demonstrate the linear map of segments in a sold by using a transparent block of elastomer. When we deform the block, the segments stretch and rotate.

Example. A beam of light casts shadows of directed segments in a solid onto a plane. This shadow casting is a linear map that maps segments in the solid to segments in the plane.

Example. The set *T* of directed intervals of time is a one-dimensional vector space. The set *D* of directed segments in a solid is a three-dimensional vector space. The linear map $v:T \to D$ defines velocity. Velocity is a linear map that maps a directed internal of time to a directed segments. This linear map is also a vector.

Example. The set *V* of all velocities is a three-dimensional vector space. The linear map $a:T \to V$ defines acceleration. Thus, acceleration is a linear map that maps a directed interval of time to a velocity.

Number as a linear map. Let ξ be a number in *F*. The definition of the vector space says that

1. $\xi(x+y) = \xi x + \xi y$ for any x and y in V
2. $\xi(\alpha x) = \alpha \xi x$ for any α in *F* and any x in V

Thus, any number ξ in *F* is a linear map that maps a vector in *V* to a vector in *V*, $\xi : V \to V$. This map scales every vector by the number ξ.

The set of all such linear maps is *F*. As we noted before, the set *F* is a one-dimensional vector space over the number field *F*.

We can also confirm that a number in *F* is a linear map that maps every element in *F* to an element in *F*.

Vector as a linear map. Let *V* be a vector space over a number field *F*. For a vector v in *V*, the definition of the multiplication says that

1 $v(x+y) = vx + vy$ for any x and y in F

2 $v(\alpha x) = \alpha v x$ for any α and x in F

Thus, the vector v is a linear map that maps the one-dimensional vector space F to the vector space V, $v : F \rightarrow V$. This map turns each number in F into a vector in V that scales with v.

The set of all vectors, which is V, is of course a vector space over the number field F.

Graphic representation. In the example of gold and silver, pieces in the set have two independent variations: the amount of gold and the amount of silver. This set is a two-dimensional vector space over the field of real numbers. We can represent the vector space graphically on a piece of paper. We draw one coordinate to represent the amount of gold, and another coordinate to represent the amount of silver. Along each coordinate we mark specific quantities proportional to the length. The origin of the coordinates represents zero amounts of gold and silver. Each point in the plane represents certain amounts of gold and silver.

The definition of the vector space does not contain the notion of the angle between two vectors. In the example of gold and silver, we gain nothing by drawing two coordinates perpendicular to each other. We may as well leave the angle between the two coordinates arbitrary. In practice, we are so used to drawing perpendicular coordinates that drawing coordinates with any other angle will require explanation.

Isomorphism. The set of gold-silver pieces contains different objects from the set of directed segments. Yet one can represent the other. Indeed, we can use the vector space of directed segments on a plane to represent any two-dimensional vector space. We now formalize this notion of equivalence between different vector spaces.

Two vector spaces V and W over a number field F are isomorphic if a one-to-one correspondence A exists to associate a vector v in V with a vector w in W, written as $w = A(v)$, such that

1) $A(x+y) = A(x) + A(y)$ for every x and y in V

2) $A(\alpha x) = \alpha A(x)$ for every x in V and every α in F

Such a relation between two vector spaces is a bijective linear map.

One-to-one correspondence of linear combination. If two vector spaces V and W over a number field F are isomorphic through a one-to-one correspondence A, we observe that

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$$

for every x and y in V and every α and β are is in F. The linear map preserves the linear combination in associating vectors in one space to vectors in the other space.

The vector zero in V corresponds to the vector zero in W, namely, $A(0)=0$. Because A is one-to-one correspondence, $A(x)\neq 0$ for every non-zero x in V.

If x and y are linearly independent vectors in V, then $A(x)$ and $A(y)$ are linearly independent vectors in W. If $A(x)$ and $A(y)$ were linearly dependent, we would find two numbers α and β in F, not both being zero, such that $\alpha A(x)+\beta A(y)=0$. This would imply that $\alpha x+\beta y=0$, namely, x and y would be linearly dependent. The contradiction supplies a proof.

The above conclusion applies to any finite numbers of linearly independent vectors. Thus, if two finite-dimensional vector spaces V and W over a number field F are isomorphic, the two vector spaces have the same dimension.

All *n*-dimensional vector spaces over a number field *F* are isomorphic. Here is a way to find a bijective linear map between any two n-dimensional vector spaces V and W over a number field F. Let $e_1,...,e_n$ be a basis of V, and $f_1,...,f_n$ be a basis of W. For any n numbers $\alpha^1,...,\alpha^n$ in F, the linear combination $\alpha^1 e_1+...+\alpha^n e_n$ is a vector in V, and the linear combination $\alpha^1 f_1+...+\alpha^n f_n$ is a vector in W. This procedure establishes a bijective linear map A between the two vector spaces V and W. The linear map A sends each vector in the basis of V to a vector in the basis of W: $f_1=A(e_1),...,f_n=A(e_n)$.

Graphic representation as isomorphism. In representing the vector space of gold-silver pieces by the vector space of directed segments in a plane, we establish a bijective linear map. We choose two pieces e_1 and e_2 of disproportional amounts of gold and silver as a basis of the vector space of gold-silver pieces. We choose two directed segments f_1 and f_2 not in the same direction as a basis of the vector space of directed segments in a plane. The linear map A associates one basis to the other, $f_1=A(e_1)$ and $f_2=A(e_2)$. For any two numbers α^1 and α^2 in F, the map A associates the gold-silver piece $\alpha^1 e_1+\alpha^2 e_2$ to the directed segment $\alpha^1 f_1+\alpha^2 f_2$. Being isomorphic does not mean being the same. We do not confuse a gold-silver piece with a directed segment.

Components of a linear map. Let V be an m-dimensional vector space over a field F, and $e_1,...,e_m$ be a basis of V. A vector v in V is a linear combination of the basis:

$$v=v^K e_K.$$

We use a capital letter to indicate an index associated the vector space V, and the dummy index implies a sum from 1 to m, the dimension of V. This expression defines the numbers $v^1,...,v^m$ in F as the components of the vector v in V relative to the basis $e_1,...,e_m$.

Let W be an n-dimensional vector space over a field F, and $f_1,...,f_n$ be a basis of W. A vector w in W is a linear combination

$$w = w^i f_i .$$

We use a small letter to indicate an index associated the vector space W, and the dummy index implies a sum from 1 to n, the dimension of W. This expression defines the numbers $w^1,...,w^n$ in F as the components of the vector w in W relative to the basis $f_1,...,f_n$.

Note that $A(e_K)$ is a vector in W, and is a linear combination of the basis of W:

$$A(e_K) = A^i_K f_i .$$

This equation defines the mn numbers A^i_K in F as the components of the linear map A relative to the two sets of basis.

Because A is a linear map, we write that

$$Av = A(v^K e_K) = v^K A(e_K) = v^K A^i_K f_i .$$

Comparing this equation with $w = Av$ and $w = w^i f_i$, we obtain that

$$w^i = A^i_K v^K .$$

This expression relates the components of the vector w in W to the components of the vector v in V, through the components of the linear map A. These components are relative to the bases of the two vector spaces, $e_1,...,e_m$ and $f_1,...,f_n$.

Matrix notation. The above expression can be written in the notation of columns and a matrix. For example, in the case $m = 2$ and $n = 3$, a vector in V is a linear combination $v = v^1 e_1 + v^2 e_2$, a vector in W is a linear combination $w = w^1 f_1 + w^2 f_2 + w^3 f_3$. We write the linear map as

$$\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \begin{bmatrix} A^1_1 & A^1_2 \\ A^2_1 & A^2_2 \\ A^3_1 & A^3_2 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$$

Example. The set V of numbers of the form $\alpha\sqrt{2}+\beta\sqrt{3}$ is a two-dimensional vector space over the field of rational numbers. The set W of numbers of the form $\xi\sqrt{5}+\zeta\sqrt{7}+\gamma\sqrt{11}$ is a three-dimensional vector space over the field of rational numbers. Here α , β, ξ, ζ and γ are arbitrary rational numbers. A particular linear map turns every number $\alpha\sqrt{2}+\beta\sqrt{3}$ in V into a number $(\alpha+2\beta)\sqrt{5}+(3\alpha+4\beta)\sqrt{7}+(5\alpha+6\beta)\sqrt{11}$ in W. We have just specified a linear map without identifying any basis.

Now we choose a basis of V as $e_1=\sqrt{2}, e_2=\sqrt{3}$, and a basis of W as $f_1=\sqrt{5}, f_2=\sqrt{7}, f_3=\sqrt{11}$. An element in V is a linear combination $v=v^1e_1+v^2e_2$, a vector in W is a linear combination $w=w^1f_1+w^2f_2+w^3f_3$. We write the linear map as

$$\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}$$

Chickens and rabbits. Each element in a set V is a cage that contains a number c of chickens and a number r of rabbits. Each element in a set W is a column of two entries: a number h of heads and a number f of feet. A linear map sends an element in V to an element in W:

$$\begin{bmatrix} h \\ f \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} c \\ r \end{bmatrix}$$

In the above, we have used the following vectors as a basis of V: : e_1 is a cage containing 1 chicken and no rabbit, and e_2 is a cage containing no chicken and 1 rabbit. Now consider another basis of the vector space V: $\tilde{e}_1$ is a cage containing 2 chickens and 3 rabbits, and $\tilde{e}_2$ is a piece containing 5 chickens and 7 rabbits. The two bases are related as

$$\tilde{e}_1 = 2e_1 + 3e_2$$
$$\tilde{e}_2 = 5e_1 + 7e_2$$

We can also invert this relation and write

$$e_1 = -7\tilde{e}_1 + 3\tilde{e}_2$$
$$e_2 = 5\tilde{e}_1 - 2\tilde{e}_2$$

Now consider an element v in V: a cage that contains c chickens and r rabbits. This element is a linear combination of the base vectors e_1 and e_2:

$$v = ce_1 + re_2 .$$

The element v is also a linear combination of the base vectors $\tilde{e}_1$ and $\tilde{e}_2$:

$$v = a\tilde{e}_1 + b\tilde{e}_2 .$$

This expression defines the numbers a and b as components of the vector x relative to the basis $\tilde{e}_1$ and $\tilde{e}_2$. Express the above expression in terms of the basis e_1 and e_2

$$\begin{aligned} v &= a\left(2e_1 + 3e_2\right) + b\left(5e_1 + 7e_2\right) \\ &= \left(2a + 5b\right)e_1 + \left(3a + 7b\right)e_2 \end{aligned}$$

Comparing this expression with $v = ce_1 + re_2$, we relate two sets of components by

$$\begin{bmatrix} c \\ r \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Relative to the basis $\tilde{e}_1$ and $\tilde{e}_2$, the linear map from V to W takes the form

$$\begin{bmatrix} h \\ f \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Change of bases of the two vector spaces. We now formalize this change of basis. Let $e_1,...,e_m$ and $\tilde{e}_1,...,\tilde{e}_m$ be two bases of V. The two bases are related as

$$\tilde{e}_L = p_L^K e_K , \quad e_K = q_K^L \tilde{e}_L .$$

These equations define the numbers p_L^K and q_K^L in F.

Let $f_1,...,f_n$ and $\tilde{f}_1,...,\tilde{f}_n$ be two bases of W. The two bases are related as

$$\tilde{f}_j = r_j^i f_i , \quad f_i = s_i^j \tilde{f}_j .$$

These equations define the numbers r_j^i and s_i^j in F.

Recall that $A\left(e_K\right)$ is a vector in W, and is a linear combination of the basis of W:

$$A\left(e_K\right) = A_K^i f_i .$$

This expression defines the number A_K^i in F as the components of the linear map A relative to the two bases $e_1,...,e_m$ and $f_1,...,f_n$.

Similarly, $A\left(\tilde{e}_K\right)$ is a vector in W, and is a linear combination of the basis of W:

$$A\left(\tilde{e}_L\right) = \tilde{A}_L^j \tilde{f}_j .$$

This expression defines the number $\tilde{A}_L^j$ in F as the components of the linear map A relative to the two bases $\tilde{e}_1,...,\tilde{e}_m$ and $\tilde{f}_1,...,\tilde{f}_m$.

Because A is a linear map, we write

$$A\left(\tilde{e}_L\right) = A\left(p_L^K e_K\right) = p_L^K A\left(e_K\right) = p_L^K A_K^i f_i = p_L^K A_K^i s_i^j \tilde{f}_j$$

A comparison of this expression with $A\left(\tilde{e}_L\right) = \tilde{A}_L^j \tilde{f}_j$ gives that

$$\tilde{A}_L^j = p_L^K s_i^j A_K^i .$$

The components of the linear map are covariant with the basis of V, but contravariant with the basis of W.

Kernel. Let A be a linear map from a vector space V to a vector space W over the number field F. The kernel of the linear map A, written as ker A, is the set whose elements are vectors in V that satisfy $Av = 0$.

The kernel of a linear map is a subspace of V. The kernel is also called the nullspace, written as Null A. The dimension of the nullspace of A is called the nullity of A.

Range. Let $A : V \to W$ be a linear map. The range of the linear map A, written as range A, is the set whose elements are vectors in W of the form Av for some v in V. The range of a linear map is a subspace of W.

$$\dim \operatorname{null} A + \dim \operatorname{range} A = \dim V .$$

Example. Let α , β, ξ, ζ and γ be arbitrary rational numbers. The set V of numbers of the form $\alpha\sqrt{2} + \beta\sqrt{3}$ is a two-dimensional vector space over the field of rational numbers. The set W of numbers of the form $\xi\sqrt{5} + \zeta\sqrt{7} + \gamma\sqrt{11}$ is a three-dimensional vector space over the field of rational numbers. A particular linear map turns every number $\alpha\sqrt{2} + \beta\sqrt{3}$ in V into a number $\left(\alpha + 2\beta\right)\sqrt{5} + \left(3\alpha + 4\beta\right)\sqrt{7} + \left(5\alpha + 6\beta\right)\sqrt{11}$ in W.

Here we have used the numbers 1,2,3,4,6 to specify a linear map, but you can use any other six rational numbers to specify other linear maps from V to W. The collection of all linear maps from V to W is a six-dimensional vector space over the field of rational numbers.

Linear map as an element of a vector space. We now paraphrase the above example in general terms. Let V and W be vector spaces on a number field F. A linear map $A : V \to W$ can be an isolated object. However, If there exist multiple linear maps from V to W, we can examine their collective properties. Let

A and B be two linear maps from V to W. The addition of the two linear maps, written as $A+B$, is defined by

$$(A+B)v = Av + Bv$$

for every element v in V. The multiplication of an number α in F and a linear map A, written as αA is defined by

$$(\alpha A)v = \alpha(Av)$$

for every element v in V.

Given two vector spaces V and W over a number field F, so long as we can find multiple linear maps from V and W, the collection of all linear maps from V to W, denoted by $L(V,W)$, is a vector space over the number field F.

For an m-dimensional vector space V and an n-dimensional vector space W, the collection of all linear maps $L(V,W)$ is an mn-dimensional vector space.

On being linear. A vector space defines being linear within a set. A linear map defines being linear in relating different sets. A vector is itself a linear map. A linear map can be an element in a new vector space. We map the new vector space to some other vector space. We map the maps. That is how we get so many vector spaces and linear maps. So goes the never-ending story of linear algebra.

Successive linear maps. Let U, V and W be three vector spaces on a number field F. A linear map $B:U\to V$ maps an element u in U to an element v in V:

$$v = Bu$$

A linear map $A:V\to W$ maps an element v in V to an element w in W:

$$w = Av$$

We write the two successive maps together as

$$w = A(Bu).$$

With every vector u in U, the successive maps associate a vector w in W.

The multiplication of the two linear maps, written as AB, is defined by

$$(AB)(u) = A(Bu)$$

for every u in U.

We can write the above operation in terms of components once we choose a basis of each of the three vector spaces. Let $d_1, d_2, ..., d_l$ be a basis of an l-dimensional vector space U, $e_1, e_2, ..., e_m$ be a basis of an m-dimensional vector space U, and $f_1, f_2, ..., f_n$ be a basis of n-dimensional space W. A vector in each of the three vector spaces is a linear combination of a basis:

$$u = u^\alpha d_\alpha, \quad v = v^K e_K, \quad w = w^i f_i.$$

These equations define the numbers u^α, v^K and w^i in F as components of the vectors relative to the bases. We use a Greek letter to indicate an index associated the vector space U, and the dummy index implies a sum from 1 to l, the dimension of U. We use a capital letter to indicate an index associated the vector space V, and the dummy index implies a sum from 1 to m, the dimension of V. We use a small letter to indicate an index associated the vector space W, and the dummy index implies a sum from 1 to n, the dimension of W.

The components of the two linear maps are defined by

$$A\left(e_K\right) = A_K^i f_i ,$$

$$B\left(d_\alpha\right) = B_\alpha^K e_K .$$

The successive maps become that

$$w = A\left(Bu\right) = A\left(B\left(u^\alpha d_\alpha\right)\right) = u^\alpha A\left(B_\alpha^K\left(e_K\right)\right) = A_K^i B_\alpha^K u^\alpha f_i .$$

Comparing this expression with $w = w^i f_i$, we obtain that

$$w^i = A_K^i B_\alpha^K u^\alpha .$$

The linear map B corresponds to a matrix B_α^K of l columns and m rows, whereas the linear map A corresponds to a matrix A_K^i of m columns and n rows. The successive maps correspond to the multiplication of matrix:

$$\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \begin{bmatrix} A_1^1 & A_2^1 \\ A_1^2 & A_2^2 \\ A_1^3 & A_2^3 \end{bmatrix} \begin{bmatrix} B_1^1 & B_2^1 & B_3^1 & B_4^1 \\ B_1^2 & B_2^2 & B_3^2 & B_4^2 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \end{bmatrix}$$

Here we write the equation for the case $l = 4$, $m = 2$, and $n = 3$.

Operator

Operator. Let V be a vector space over a number field F. With every vector x in V, a map A associates a vector z in V,

$$z = A(x).$$

The map is linear if the following conditions hold.

1. $A(x+y) = A(x) + A(y)$, for any x and y in V.
2. $A(\alpha x) = \alpha A(x)$, for any α in F and any x in V.

We call such a linear map an operator on the vector space V over the number field F. We write

$$z = Ax,$$

if it does not cause confusion.

Components of an operator. Let V be an n-dimensional vector space, and $e_1, e_2, ..., e_n$ be a basis of V. The vector x in V is a linear combination of the base vectors:

$$x = x^i e_i.$$

The dummy index i implies the summation from 1 to n. This expression defines the numbers $x^1, x^2, ..., x^n$ in F as the components of the vector x relative to the basis $e_1, e_2, ..., e_n$.

Similarly, the vector z in V is also a linear combination of the basis:

$$z = z^j e_j.$$

The dummy index j implies the summation from 1 to n. This expression defines the numbers $z^1, z^2, ..., z^n$ in F as the components of the vector z relative to the basis $e_1, e_2, ..., e_n$.

Observe that $A(e_i)$ is a vector in V, and is a linear combination of the vectors in the basis:

$$A(e_i) = A_i^j e_j.$$

This expression defines the n^2 numbers A_i^j in F as the components of the operator A relative to the basis $e_1, e_2, ..., e_n$.

Because A is a linear map, we write

$$Ax = A(x^i e_i) = x^i A(e_i) = A_i^j x^i e_j.$$

A comparison of the this expression with $z = Ax$ and $z = z^j e_j$ gives that

$$z^j = A_i^j x^i.$$

This expression calculates the components of the vector z once we know the components of the vector x and those of the operator A. The repeated index i

implies a summation. The result of the summation gives a vector—an object of a single index. Such an operation reduces the number of indices, and is known as *contraction*.

Matrix notation. These expressions are also written in the matrix notation:

$$\begin{bmatrix} z^1 \\ z^2 \\ z^3 \end{bmatrix} = \begin{bmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}$$

As a convention, we use the superscript of A_i^j to indicate the row, and the subscript to indicate the column.

Change of basis. The components of a vector in *V* depend on a basis of *V*. So are the components of an operator on *V*. Let e_1, e_2, e_3 and $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$ be two bases of the space *V*. The two bases are related by

$$\tilde{e}_i = p_i^j e_j, \quad e_i = q_i^j \tilde{e}_j .$$

We have already defined the components of the operator *A* with respect to the basis e_1, e_2, e_3 by using the expression

$$A(e_i) = A_i^j e_j .$$

Similarly, observe that $A(\tilde{e}_i)$ is a vector V, and is a linear combination of the $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$:

$$A(\tilde{e}_i) = \tilde{A}_i^j \tilde{e}_j .$$

This expression defines the nine numbers $\tilde{A}_i^j$ in *F* as the components of the operator *A* relative to the basis $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$.

Because *A* is a linear map, we write

$$A(\tilde{e}_i) = A(p_i^j e_j) = p_i^k A(e_k) = p_i^k A_k^l e_l = p_i^k A_k^l q_l^j \tilde{e}_j .$$

Comparing this expression with $A(\tilde{e}_i) = \tilde{A}_i^j \tilde{e}_j$, we obtain that

$$\tilde{A}_i^j = p_i^k q_l^j A_k^l .$$

This expression relates both the matrix of the forward change from e_1, e_2, e_3 to $\tilde{e}_1, \tilde{e}_2, \tilde{e}_3$, and the matrix of the inverse change the other way around. The operator *A* on *V* is partly contravariant and partly covariant with respect to the basis of *V*.

Eigenvectors and eigenvalues. Let A be an operator on a vector space V over a number field F. For a vector v in V, Av is another vector in V. In general, the two vectors, v and Av, are in different directions. For a certain vector v, the two vectors v and Av are in the same direction, namely,

$$Av = \lambda v ,$$

where λ is a number in F. Such a vector is known as an eigenvector, and the number λ the eigenvalue of A associated with the eigenvector v.
Write the above equation as

$$\left(A - \lambda I\right)v = 0 .$$

This is a linear algebraic equation for v. A non-trivial solution exists if and only if the matrix is singular, namely,

$$\det\left(A - \lambda I\right) = 0 .$$

For an operator on an n-dimensional vector space, this equation involves an nth order polynomial of λ. If the number field F is the field of complex numbers, according to the fundamental theorem of algebra, the polynomial has n roots. In general, the roots are complex numbers. The n roots are n eigenvalues.

Successive maps. Let A and B be two operators on the vector space V over the number field F. Let u be a vector in a vector space V, the object Bu is also a vector in V. Write $v = Bu$. The object Av is still a vector in V and we write $w = Av$. We write

$$w = A\left(v\right) = A\left(B\left(u\right)\right) .$$

The successive mapping associates every vector in V with a vector in V. We call the successive mapping the multiplication of the two operators. We write the successive mapping as AB, and write

$$w = \left(AB\right)\left(u\right) .$$

We can write the above operation in terms of components. Let a basis of the vector space V be e_1, e_2, e_3. The vector u in V is a linear combination of the base vectors:

$$u = u^j e_j .$$

This expression defines the numbers u^j in F as the components of the vector u relative to the basis e_1, e_2, e_3. Similarly, the vector w in V is a linear combination of the base vectors:

$$w = w^i e_i .$$

This expression defines the numbers w_i in F as the components of the vector w relative to the basis e_1, e_2, e_3. The components of the two operators are defined by

$$A(e_k) = A_k^i e_i ,$$

$$B(e_j) = B_j^k e_k .$$

The composite operation becomes that

$$w = A(B(u)) = A(B(u^j e_j)) = A(B_j^k (u^j e_k)) = A_k^i B_j^k u^j e_i .$$

A comparison of this expression with $w = w^i e_i$ gives that

$$w^i = A_k^i B_j^k u^j .$$

This expression is also written in the matrix form:

$$\begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \begin{bmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{bmatrix} \begin{bmatrix} B_1^1 & B_2^1 & B_3^1 \\ B_1^2 & B_2^2 & B_3^2 \\ B_1^3 & B_2^3 & B_3^3 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix}$$

The successive maps correspond to the multiplication of matrix.

The set of all operators on a vector space. Let $L(V)$ be the set of all operators on the vector *V* over the number field *F*. For every vector *x* in *V*, and two operators *A* and *B* on *V*, $Ax + Bx$ is a vector in *V*. The addition of the two operators *A* and *B*, written as $A + B$, is defined by

$$(A + B)x = Ax + Bx .$$

For every vector *x* in *V*, every number α in *F*, and every operator *A* on *V*, $\alpha(Ax)$ is a vector in *V*. The multiplication of α and *A*, written as αA, is defined by

$$(\alpha A)x = \alpha(Ax).$$

Both $A + B$ and αA are operators on *V*. That is, the set $L(V)$ is closed under the following two binary maps.

The set $L(V)$ of all operators on a vector space *V* over a number field *F* is closed under the two binary maps. Thus, this set is also a vector space over the number field *F*. For an *n*-dimensional vector space *V*, the dimension of the vector space $L(V)$ is n^2.

Scalar

Various usages of the word scalar. In linear algebra, the word *scalar* is used in two ways. First, the word scalar is used as a synonym to the word *number*, an element in the field F. Second, the word scalar is used in defining a linear form, an object to be studied in a later section. In this second usage, the scalars behave in a way different from numbers. Adding two scalars results in a scalar, and multiplying a scalar with a number resulting in a scalar; however, multiplying two scalars does not result in a scalar in the same set. The situation is the same in defining bilinear form, quadratic form, and inner product.

In physics, the word scalar is used to indicate properties like mass, volume and energy. The usage is consistent with the second usage in linear algebra, but is inconsistent with the first one in several ways:

- A physical property like mass is more than just a number; it has a unit.
- The multiplication defined on a field makes no sense to a physical quantity like mass: the multiplication of two elements in F gives yet another element in F, but the multiplication of two masses does not give another mass.
- If we regard both mass and volume as elements in the field F, then we need to assign a meaning to the addition of mass and volume. What does that even mean?

We will call each element in the field F a number, and will reserve the word scalar for another object in linear algebra, defined as follows.

Scalar set. We call a collection S of objects a scalar set if it is a one-dimensional vector space over a number field F. We call each element in S a scalar, and each element in F a number.

Remark. This definition relies on the definitions of number field and vector space. The definition invokes two sets F and S, as well as four binary maps. The set F is closed under two binary maps: adding two elements in F gives an element in F, $(F,F) \to F$; multiplying two elements in F gives an element in F, $(F,F) \to F$. The set S is closed under two other binary maps: adding two elements in S gives an element in S, $(S,S) \to S$; multiplying an element in F and an element in S gives an element in S, $(F,S) \to S$.

The four binary maps follow the familiar rules, as listed in the definitions of number field and vector space.

Example. The set of all numbers of form $q\sqrt{2}$, where q is a rational number, is a scalar set over the field of rational numbers.

Example. The set of all numbers of form bi, where b is a real number and $i = \sqrt{-1}$, is a scalar set over the field of real numbers.

Example. The set of all different amounts of money is a scalar field.

Example. The set of points on a straight line is not a scalar set, because it is unclear how we define the addition of two points, or the multiplication of a point and a number. However, we can form a scalar set by the following procedure. Mark a particular point on the line as the origin. The position of any point on the line relative to the origin defines a directed segment. The set of all directed segments is a scalar set over the field of real numbers. The addition of two directed segments x and y is a directed segment, formed by placing the tail of x at the origin, and placing the tail of y at the tip of x. The multiplication of a real number r and a segment x is a segment of length r times that of x.

Example. Similarly, the set of all times is not a scalar set, because it is unclear how we define the addition of two times, or the multiplication of a time and a number. However, we can form a scalar set by following a similar procedure. Mark a particular time as the reference. The difference of any other time relative to this reference defines a directed interval. The set of all directed intervals is a scalar set over the field of real numbers. The addition of two directed intervals x and y is a directed interval, formed by placing the tail of x at the reference, and placing the tail of y at the tip of x. The multiplication of a real number r and a n interval x is a segment of length r times that of x. We can use the same procedure to construct a scalar set by using differences in energy with respect to a reference point.

A set of gold. Consider a set of all different amounts of gold. We can define the addition of the pieces, but we do not have a sensible definition of the multiplication of the pieces. Thus, this set is not a number field. This set, however, is a one-dimensional vector space over the field of real numbers. That is, the set of pieces of all different amounts of gold is a scalar set. We stipulate the two binary maps in a natural way. Adding two pieces of gold corresponds to a piece of gold. Multiplying a real number r and a piece of gold corresponds to a piece of gold r times the amount.

Unit and magnitude. Of all vector spaces, one-dimensional vector spaces have a unique attribute: elements in a one-dimensional vector space *scale* with one another by numbers in F. Let u be a nonzero element in S. An element s in S scales with u, namely,

$$s = s_M u,$$

where s_M is a number in F. We call u a unit of the scalar set S, and s_M the magnitude of the scalar s relative to the unit u.

Remark. A scalar set is, by definition, a one-dimensional vector space. Nonetheless it is helpful to replace generic terms for a vector space with specific terms for a scalar set. We replace the phrase "one-dimensional vector space" with the phase "scalar set". We call an element in the vector space a vector, and call an element in the scalar set a scalar. A basis of an *n*-dimensional vector space consists of *n* linearly independent elements in the vector space, whereas a unit of a scalar set is a nonzero element in the scalar set. We replace the phrase "components of a vector with respect to a basis" with the phrase "magnitude of a scalar relative to a unit".

Do not confuse a scalar with its magnitude. For a three-dimensional vector space *V* over a number field *F*, given a basis e_1, e_2, e_3, a vector x in *V* is a linear combination, $x = x^1 e_1 + x^2 e_2 + x^3 e_3$, where the components x^1, x^2, x^3 are numbers in *F*. We should not confuse a vector with its components. For the set of directed segments in a plane, a vector is an object, a directed segment, and its components are real numbers.

Similarly, for a scalar set *S* over a number field *F*, given a unit *u*, a scalar *s* in *S* scales with the unit, $s = s_M u$, where the magnitude s_M is a number in *F*. For the set of gold, the scalar is an object, a piece of gold, and the magnitude is a real number. We do not confuse a piece of gold with a number.

Extensive property. A piece of a substance, such as gold, has many physical properties, including volume, shape, color, temperature, mass, energy. A physical property is extensive if it is proportional to the amount of the substance. Volume, mass and energy are extensive properties. Shape, color, temperature are not extensive properties.

We can use a one-dimensional vector space *S* to model an extensive physical property such as mass. In this model, *F* is the field of real numbers. We define the addition of two masses by lumping them together. We define the multiplication of a mass and a real number *r* by another mass *r* times the amount.

We use a particular element in *S* as the unit for this quantity. For example, for the scalar set *S* of all masses, the unit mass, kilogram, is the mass of a block metal, called the International Prototype Kilogram (IPK), preserved in a vault located in Sevres, France. Any other mass equals this unit times a real number. For example, 1.7 kg means a mass, which is 1.7 times the mass of the IPK. The mass of the IPK is an element in the scalar set of masses, and by an international convention we agree to call it a unit of mass. The mass 1.7 kg is another element in the set.

Like mass, each of extensive physical properties, including energy, entropy, and electric charge, forms a scalar set.

Counter example. The set of all temperatures is not a scalar set.

Change of unit. Let u and $\tilde{u}$ be two non-zero scalars in a scalar set S over a number field F. The two scalars are proportional to each other:

$$\tilde{u} = pu ,$$

where p is a number in F, and is the magnitude of the scalar $\tilde{u}$ relative to the scalar u.

A scalar s in S scales with either unit:

$$s = s_M u = \tilde{s}_M \tilde{u} ,$$

where the number s_M in F and is the magnitude of the scalar s relative to the unit u, and the number $\tilde{s}_M$ is the magnitude of the scalar s relative to the unit $\tilde{u}$.

A combination of the above expressions gives that

$$s_M = p\tilde{s}_M .$$

This expression relates three numbers in F. The magnitude of a scalar converts in a way opposite to the way in which the unit of the scalar set converts. Thus, the scalar set is contravariant.

Example. Kilogram and pound are two units of mass. They convert to each other by 1 kilogram = 2.20462 pounds. Thus, a 10-pound turkey is 4.5 kilograms.

Example. Money in the United States comes in many units: cent, nickel, dime, quarter, and dollar. These units are represented by distinct physical objects.

Linear Function

Let S and T be two scalar sets over a number field F. Write a map $g:S \rightarrow T$ as

$$t = g(s),$$

where s is an element in S, and t is and element T. The map is a linear function if

1) $g(x+y) = g(x)+g(y)$ for any x and y in S
2) $g(\alpha x) = \alpha g(x)$ for any α in F and any x in S

We write the linear map $g(s)$ as gs if it does not cause any confusion.

Remark. We have just written conditions 1 and 2 out of habit. Because S is a scalar set, (2) implies (1). Let u be a non-zero element in S. The two scalars x and y in S scale as $x = \tilde{x}u$ and $y = \tilde{y}u$, where $\tilde{x}$ and $\tilde{y}$ are numbers in F. Consequently,

$$g(x+y) = g((\tilde{x}+\tilde{y})u) = (\tilde{x}+\tilde{y})g(u) = \tilde{x}g(u)+\tilde{y}g(u) = g(x)+g(y).$$

Example. Let U be the set of various amounts of US dollars, and E be the set of various amounts of Euros. A linear map $r:U \rightarrow E$ is a rate of exchange.

Example. Let G be the set of various amounts of gold, and C be the set of various amounts of money. A linear map $p:G \rightarrow C$ is a price of gold.

Example. Let M be the set of different masses of gold, and V be the set of different volumes. The linear map $\rho:V \rightarrow M$ is the density of gold.

Example. Led T be the set of directed intervals of time, and D be the set of directed segments on a geometric line. The linear map $c:T \rightarrow D$ is the speed.

Example. Led T be the set of directed intervals of time, and E be the set of differences in energy. The linear map $P:T \rightarrow E$ is the power.

Example. Led T be the set of directed intervals of time, and C be the set of various amounts of electric charge. The linear map $I:T \rightarrow C$ is the electric current.

Example. Let C be the set of various amounts of electric charge, and E be the set of various amounts of energy. The linear map $\phi:C \rightarrow E$ is the electric potential.

Example. Let N be the set of various amounts of a species of molecules, and E be the set of various amounts of energy. The linear map $\mu : N \to E$ is the chemical potential.

Magnitude of a linear map. The magnitude of a linear map is relative to the units of the two scalar sets. Let S and T be two scalar sets over a number field F, and $g : S \to T$ be a linear map. The linear map associates an element s in S to an element t in T:

$$t = gs .$$

Let u be a unit of S. An element s in S scales with u, namely,

$$s = s_M u .$$

This expression defines the number s_M in F as the magnitude of the scalar s relative to the unit u. Similarly, let v be a unit of T. An element t in T scales with v, namely,

$$t = t_M v .$$

This expression defines the number t_M in F as the magnitude of the scalar t relative to the unit v.

Because gs is a linear map, we write

$$gs = g\left(s_M u\right) = s_M g u .$$

Note that u is an element in S, so that gu is an element in T, and must scale with v. We write

$$gu = g_M v .$$

This expression maps the unit u of S to the unit v of T. We call the number g_M in F as the magnitude of the linear map with respect to the units u and v.

A combination of the above expressions, we obtain that

$$t_M v = g_M s_M v .$$

The components must equal:

$$t_M = g_M s_M$$

This expression relates the three magnitudes, the numbers in F.

Example. Let S be the set of all numbers of form $q\sqrt{2}$, and T be the set of all numbers of form bi, where q and b are rational numbers, and $i = \sqrt{-1}$. A particular a linear map $g : S \to T$ associates an element $q\sqrt{2}$ in S to an element $1.6qi$ in T; that is, $g\left(q\sqrt{2}\right) = 1.6qi$. This description is independent of the choice of units of the two scalar sets.

We can also choose units for the two scalar sets, say $u = 2\sqrt{2}$ as a unit of *S*, and $v = 3i$ as a unit of *T*. Write an element *s* in *S* as $s = \hat{s}2\sqrt{2}$, and an element *t* in *T* as $t = \hat{t}3i$, where $\hat{s}$ and $\hat{t}$ are rational numbers, and are the magnitudes of the two scalars relative the two units.

Recall that $g(u) = \hat{g}v$ defines the rational number $\hat{g}$ as the magnitude of the linear map *g* relative to the two units *u* and *v*. Write

$$g(2\sqrt{2}) = (1.6)2i = (1.6)(2/3)(3i),$$

so that the magnitude of the linear map *g* is $\hat{g} = (1.6)(2/3)$.

Change of units of the two scalar sets. Let *u* and $\tilde{u}$ be two non-zero scalars in *S*. The two scalars are proportional to each other:

$$\tilde{u} = pu,$$

where *p* is a number in *F*, and is the magnitude of the scalar $\tilde{u}$ relative to the scalar *u*. Let *v* and $\tilde{v}$ be two non-zero scalars in *S*. The two scalars are proportional to each other:

$$\tilde{v} = rv,$$

where *r* is a number in *F*, and is the magnitude of the scalar $\tilde{v}$ relative to the scalar *v*.

Recall that $gu = g_M v$ defines the number g_M in *F* as the magnitude of the linear map *g* relative to the two units *u* and *v*. Similarly, $g(\tilde{u}) = \tilde{g}_M \tilde{v}$ defines the number $\tilde{g}_M$ in *F* as the magnitude of the linear map *g* relative to the two units $\tilde{u}$ and $\tilde{v}$. The two magnitudes g_M and $\tilde{g}_M$ are related as

$$\tilde{g}_M = rg_M / p.$$

Thus the magnitude of the linear map is covariant with respect to the unit of *T*, but contravariant with respect to the unit of *S*.

The collection of all linear maps from one scalar set to the other scalar. Let $L(S,T)$ be the collection of all linear maps from a scalar set *S* to a scalar set *T*, both over the number field *F*. For every element *s* in *S*, every number α in *F*, and every linear map $g: S \to T$, $\alpha(gs)$ is an element in *T*. The multiplication of α and *g*, written as αg, is defined by

$$(\alpha g)s = \alpha(gs).$$

That is, the collection $L(S,T)$ is also a scalar set space over the number field *F*.

Example. Let S be the set of all numbers of form $q\sqrt{2}$, and T be the set of all numbers of form bi, where q and b are rational numbers, and $i = \sqrt{-1}$. Each linear from S to T associates an element $q\sqrt{2}$ in S to an element rqi in T, where r is a particular rational number. That is, each rational number r corresponds to a distinct linear map from S to T. All these linear maps constitute $L(S,T)$.

Successive linear maps. Let S, T and U be three scalar sets over a number field F. Consider two successive linear maps, $g: S \to T$ and $h: T \to U$. We can use the successive maps to define a map $hg: S \to U$.

Linear Form

Linear form. Let V be a vector space and S be a scalar set, both over a number field F. Let a be a map that associates a scalar s in S with every vector x in V:

$$s = a(x).$$

We also write this map as

$$a : V \to S .$$

The map is linear if

1. $a(x+y) = a(x) + a(y)$ for any x and y in V
2. $a(\lambda x) = \lambda a(x)$ for any λ in F and any x in V

We call such a linear map a linear form. The linear form is a special case of a linear map from one vector space to another vector space.

Example. Each element in a set M is a piece containing some amount of gold and some amount of silver. We model this set as a two-dimensional vector space over the field of real numbers. The set C of various amounts of money is a scalar set over the field of real numbers. We will call this set the set of cost. The linear map $P : M \to C$ is a list of prices of the two metals. Another list of prices is another linear map.

Example. The set X of directed segments in a solid is a three-dimensional vector space over the field of real numbers. The set E of various amounts of energy is a scalar set over the field of real numbers. The linear map $F : X \to E$ is the force.

Example. The set X of directed segments in a solid is a three-dimensional vector space over the field of real numbers. The set U of various levels of electric potential is a scalar set over the field of real numbers. The linear map $E : X \to U$ is the electric field.

Components of a linear form. Let V be an n-dimensional vector space over a number field F, and $e_1,...,e_n$ be a basis of V. A vector x in V is a linear combination the vectors in the basis:

$$x = x^i e_i .$$

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

Let S be a scalar set over the number field F, and u be a unit of S. A scalar s in S scales with the unit u:

$$s = s_M u ,$$

where the number s_M in *F* is the magnitude of the scalar *s* in *S* relative to the unit *u*.

Observe that e_i is a vector in *V*, and $a(e_i)$ is a scalar in *S*. Consequently, $a(e_i)$ scales with the unit *u* of *S*:

$$a(e_i) = a_i u .$$

This expression defines the numbers $a_1,...,a_n$ in *F* as the components of the linear form *f* relative to the basis $e_1,...,e_n$ of *V* and the unit *u* of *S*.

Linear form as a sum of products. Because $a(x)$ is a linear map, we write

$$a(x) = a(x^i e_i) = x^i a(e_i) = x^i a_i .$$

The repeated index implies a sum from 1 to *n*. Comparing this expression with $s = a(x)$ and $s = s_M u$, we obtain that

$$s_M = a_i x^i .$$

Thus, once the components $a_1,...,a_n$ of a linear form are known, the above expression calculates the magnitude s_M of the scalar for any vector *x*.

Example. For the vector space of silver and gold, each vector is a piece that contains some amount of gold and some amount of silver. The cost of each piece is a scalar. Suppose we determine the cost of each piece in a simple way:

$$\left(\text{cost of a piece}\right) = \text{(price of gold)(amount of gold)} + \text{(price of silver)(amount of silver)} .$$

The linear form turns each vector (piece) into a scalar (cost).

Change of basis and unit. For the linear map $f: V \to S$, the components of the linear map depend on the basis of *V* and the unit of *S*. Let $e_1,...,e_n$ be a basis of the vector space *V*, and $\tilde{e}_1,...,\tilde{e}_n$ be another basis of the same vector space. The two bases are related as

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

where the numbers p_j^i in *F* relate the two bases. A vector *x* in *V* is a linear combination of either basis:

$$x = x^i e_i = \tilde{x}^i \tilde{e}_i ,$$

where the numbers $x^1,...,x^n$ in *F* are the components of *x* relative to the basis $e_1,...,e_n$, and the numbers $\tilde{x}^1,...,\tilde{x}^n$ in *F* are the components of *x* relative to the basis $\tilde{e}_1,...,\tilde{e}_n$. The two sets of components are related as

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

Thus, the components of the vector in *V* transform in the opposite way as the basis. We say that the vector is *contravariant*.

Let *u* be a unit of the scalar set *S*, and $\tilde{u}$ be another unit of the same scalar set. The two units are related as

$$\tilde{u} = ru ,$$

where the number *r* in *F* converts the two units. A scalar *s* in *S* scales with either unit:

$$s = s_M u = \tilde{s}_M u ,$$

where the number s_M in *F* is the magnitude of *s* relative to the unit *u*, and the number $\tilde{s}_M$ in *F* is the magnitude of *s* relative to the unit $\tilde{u}$. The two magnitudes are related as

$$s_M = r\tilde{s}_M .$$

Thus, the scalar is *contravariant*.

Associated with a change of the basis of *V* and a change with the unit of *S*, the components of the linear form also change. Because $\tilde{e}_i$ is a vector in *V*, the linear form $a(\tilde{e}_i)$ maps the vector $\tilde{e}_i$ to an element in *S*. Write

$$a(\tilde{e}_i) = \tilde{a}_i \tilde{u} .$$

The numbers $\tilde{a}_1,...,\tilde{a}_n$ are the components of the linear form *a* relative to the basis $\tilde{e}_1,...,\tilde{e}_n$ and the unit $\tilde{u}$. Because $a(\tilde{e}_i)$ is a linear function, we write

$$a(\tilde{e}_i) = a(p_i^j e_j) = p_i^j a(e_j) = p_i^j a_j u .$$

A combination of the above two expressions gives

$$\tilde{a}_i = p_i^j f_j / r .$$

The components of the linear form transform in the same way as the basis of *V*, but in the opposite way as the unit of *S*. We say that the linear form is *covariant* with respect to the vector, but *contravariant* with respect to the scalar.

Dual space. Given a vector space *V* and a scalar set *S* over a number field *F*, a linear map from *V* to *S* can be an individual object. If there exist many linear maps from *V* to *S*, however, we can study their collective properties. Let *a* and *b* be two linear maps from *V* to *S*. The addition of the two linear maps, written as $a+b$, is defined by

$$(a+b)(x) = a(x) + b(x)$$

for every element *x* in *V*. The multiplication of a number λ in *F* and a linear map *a*, written as λa is defined by

$$(\lambda a)(x) = \lambda(a(x))$$

for every element x in V.

Given a vector space V and a scalar set S over a number field F, the collection of all linear maps from V to S, denoted by $L(V,S)$, is a vector space over the number field F. This space is called the dual space of V with respect to the scalar set S.

For an n-dimensional vector space V, the dual space with respect to any scalar set is an n-dimensional vector space.

Example. The set of all lists of prices is the dual space of the vector space of gold and silver with respect to the set of cost. Adding two lists of prices means forming a new list of prices. This operation is useful, for example, when we set a list of prices for the metals sold in one location, and then set another list of prices associated with transportation of the metals to a different location. Multiplying a list of prices by a real number means changing the price of every metal by a factor r.

Example. The set of forces is the dual space of the vector space of directed segments with respect to energy.

The set of electric fields is the dual space of the vector space of directed segments with respect to electric potential.

Basis of the dual space. Why do we talk about this? Given an n-dimensional vector space V and a scalar set S over a number field F, the set of all linear maps from V to S is also an n-dimensional vector space, the dual space, $L(V,S)$. Let $f^1,...,f^n$ be a basis of $L(V,S)$. An element a in $L(V,S)$ is a linear combination in the basis:

$$a = a_1 f^1 + ... + a_n f^n .$$

This expression defines numbers $a_1,...,a_n$ as the components of the vector a in $L(V,S)$ relative the basis $f^1,...,f^n$. Contrast this expression with $a(e_i) = a_i$. In general, the two usages of the numbers $a_1,...,a_n$ are different.

Let $e_1,...,e_n$ be a basis of the vector space V, and u be a unit of the scalar set S. Note that f^γ is a vector in $L(V,S)$, e_i is a vector in V, and $f^\gamma(e_i)$ is a scalar in S. The scalar scales with the unit of S:

$$f^\gamma(e_i) = D_i^\gamma u .$$

This expression defines the n^2 numbers D_i^γ in F that relate to the basis $f^1,...,f^n$ of the dual space $L(V,S)$ to the basis $e_1,...,e_n$ of V.

For a vector x in V and a vector a in $L(V,S)$, $a(x)$ is a scalar in S. We write

$$a(x) = a_\gamma f^\gamma (x^i e_i) = a_\gamma x^i f^\gamma (e_i) = D_i^\gamma a_\gamma x^i u .$$

A comparison with $s = a(x)$ and $s = s_M u$, we write

$$s_M = D_i^\gamma a_\gamma x^i .$$

Dual basis. Because $e_1,...,e_n$ are linearly independent vectors in V, and $f^1,...,f^n$ are linearly independent vectors in $L(V,S)$, the n by n matrix D_i^γ is invertible. Conversely, given a basis $e_1,...,e_n$ of V, any invertible n by n matrix D_i^γ defines a basis $f^1,...,f^n$ of the dual space $L(V,S)$ through

$$f^\gamma (e_i) = D_i^\gamma u .$$

Given a basis $e_1,...,e_n$ in V, A basis $f^1,...,f^n$ of $L(V,S)$ is called a dual basis if

$$f^\gamma (e_i) = \begin{cases} u & \text{for}\, \gamma = i \\ 0 & \text{for}\, \gamma \neq i \end{cases}$$

The dual basis corresponds to the matrix

$$D_i^\gamma = \begin{cases} 1 & \text{for}\, \gamma = i \\ 0 & \text{for}\, \gamma \neq i \end{cases}$$

When $f^1,...,f^n$ is the dual basis of $e_1,...,e_n$, the two expressions $a(e_i) = a_i$ and $a = a_\gamma f^\gamma$ define the same list of numbers $a_1,...,a_n$.

A vector x in V is a linear combination $x = x^i e_i$, and a vector a in $L(V,S)$ is a linear combination $a = a_i f^i$. If the basis $f^1,...,f^n$ is the dual basis, we write

$$a(x) = a_\gamma f^\gamma (x^i e_i) = a_\gamma x^i f^\gamma (e_i) = a_i x^i u .$$

By choosing the basis of the dual space to be the dual basis, we “tie” the basis $f^1,...,f^n$ to the basis $e_1,...,e_n$. To indicate this tie, we write $f^1,...,f^n$ as $e^1,...,e^n$.

$$e^\gamma (e_i) = u \delta_i^\gamma .$$

Once we change the basis $e_1,...,e_n$ of V, we also change the dual basis $e^1,...,e^n$ of $L(V,S)$ accordingly. Let the new basis $\tilde{e}_1,...,\tilde{e}_n$ be a new basis of V, which relate to the old basis as

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

$$\tilde{e}^{\xi}\left(\tilde{e}_j\right) = \delta_j^{\xi} u$$

$$e^{\gamma}\left(\tilde{e}_j\right) = e^{\gamma}\left(p_j^i e_i\right) = p_j^i e^{\gamma}\left(e_i\right) = p_j^i u \delta_i^{\gamma}$$

The new dual basis relate to the old dual basis by

$$e^{\gamma} = p_{\xi}^{\gamma} \tilde{e}^{\xi}$$

Consequently, the components of a vector in the dual space $L(V,S)$ are covariant with the basis of the space V.

$$e^i$$

Tensor. Perhaps move this section to the end.

$$c : V \times V \ldots \times L(V,S) \times L(V,S) \to T$$

$$t = c(x, y \ldots, a, b \ldots)$$

Write a linear form as a bilinear form
Basis
Vector space
Associate withe each scalar set is a dual space, many dual spaces

Force linear in displacement. Stiffness
Energy well

Taylor expansion of a function of a vector. A nonlinear map from v to s.multiliear form

Multilinear Map

Bilinear map. Let B be a map from two sets U and V to a set W. To each element u in U and each element v in V, there corresponds an element in W, written as

$$w = B(u,v).$$

We also write the map as

$$B:(U,V) \to W .$$

We call the sets U and V the domain of the map, and the set W the codomain of the map. The map is a by linear map if U, V and W are vector spaces on a number field F, and if the map is linear in each of its independent variables, namely,

$$B(\alpha_1 u_1 + \alpha_2 u_2, v) = \alpha_1 B(u_1, v) + \alpha_2 B(u_2, v),$$

$$B(u, \alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 B(u, v_1) + \alpha_2 B(u, v_2),$$

for every u, u_1, u_2 in U, every v, v_1, v_2 in V, and every α_1 and α_2 in F. One can similarly define multilinear maps.

Example. The set V of directed segments in a solid is a three-dimensional vector space on the field of real numbers. Three directed segments are edges of a parallelepiped. The volume of the parallelepiped is a trilinear map on the vector space.

Components of a multilinear map. Let U, V and W be three vector spaces on a number field F. A bilinear map $B:(U,V) \to W$ maps an element u in U and an element v in V to an element w in W:

$$w = B(u,v).$$

Let $d_1, d_2, ..., d_l$ be a basis of an l-dimensional vector space U, let $e_1, e_2, ..., e_m$ be a basis of an m-dimensional vector space U, and let $f_1, f_2, ..., f_n$ be a basis of n-dimensional space W. A vector in each of the three vector spaces is a linear combination of a basis:

$$u = u^\alpha d_\alpha , \quad v = v^K e_K , \quad w = w^i f_i .$$

These equations define the numbers u^α, v^K and w^i in F as components of the vectors relative to the bases.

Because d_α is an element in U, e_K is an element in V, and $B(d_\alpha, e_K)$ is an element in W, we can write $B(d_\alpha, e_K)$ as a linear combination of the basis of W:

$$B(d_\alpha, e_K) = B^i_{\alpha K} f_i .$$

This expression defines a total of lmn numbers $B^i_{\alpha K}$ in F as the components of the bilinear map. We write

$$B(u,v) = B(u^\alpha d_\alpha, v^K e_K) = u^\alpha v^K B(d_\alpha, e_K) = B^i_{\alpha K} u^\alpha v^K f_i .$$

A comparison of this expression with $w = B(u,v)$ and $w = w^i f_i$ gives that

$$w^i = B^i_{\alpha K} u^\alpha v^K .$$

The components of the bilinear map $B^i_{\alpha K}$ have three indices, and cannot be listed as a matrix.

The components of the bilinear map $B^i_{\alpha K}$ are covariant with the bases of *U* and *V*, and contravariant with the basis of *W*.

Example. Let a, b, α , β , ξ , ζ and γ be arbitrary rational numbers. The set *U* of numbers of the form $a+bi$, the set *V* of numbers of the form $\alpha\sqrt{2}+\beta\sqrt{3}$, and the set *W* of numbers of the form $\xi\sqrt{5}+\zeta\sqrt{7}+\gamma\sqrt{11}$ are two-, two- and three-dimensional vector spaces over the field of rational numbers. A particular linear map turns every element $a+bi$ in *U* and every element $\alpha\sqrt{2}+\beta\sqrt{3}$ in *V* into a element in *W*:

$$(a+2b)(\alpha+2\beta)\sqrt{5}+(3a+4b)(3\alpha+4\beta)\sqrt{7}+(5a+6b)(5\alpha+6\beta)\sqrt{11}$$

You can use any other rational numbers to specify other bilinear maps from *U* and *V* to *W*. The collection of all linear maps from *U* and *V* to *W* is a 12-dimensional vector space over the field of rational numbers. We next paraphrase this example in general terms.

Multilinear map as an element of a vector space. Let *U*, *V* and *W* be vector spaces over a number field *F*. Let $A:(U,V)\to W$ and $B:(U,V)\to W$ be two bilinear maps. The addition of the two linear maps, written as $A+B$, is defined by

$$(A+B)(u,v) = A(u,v)+B(u,v)$$

for every *u* in *U* and every *v* in *V*. The multiplication of an number α in *F* and a bilinear map *A*, written as αA is defined by

$$(\alpha A)(u,v) = \alpha(A(u,v))$$

for every *u* in *U* and every *v* in *V*.

Given three vector spaces *U*, *V* and *W* over a number field *F*, so long as we can find multiple bilinear maps from *U* and *V* to *W*, the collection of all bilinear is a vector space over the number field *F*.

For and *l*-dimensional vector space *U*, an *m*-dimensional vector space *V* and an *n*-dimensional vector space *W*, the collection of all bilinear maps is an *lmn*-dimensional vector space.

Inner Product

This section considers an n-dimensional vector space V and a scalar set S. All numbers are real numbers.

Bilinear form. With any two vectors x and y in V, a binary map $g:(V,V)\to S$ associates an element s in S, written as

$$s = g(x,y).$$

The map is a bilinear form if $g(x,y)$ is linear in each of its arguments, namely,

$$g(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 g(x_1,y) + \alpha_2 g(x_2,y),$$
$$g(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 g(x,y_1) + \alpha_2 g(x,y_2),$$

for every x, x_1, x_2, y, y_1, y_2 in V, and every numbers α_1 and α_2.

Components of a bilinear form. Let $e_1,...,e_n$ be a basis of the vector space V. A vector x in V is a linear combinations of the base vectors:

$$x = x^i e_i .$$

The numbers $x^1,...,x^n$ are the components of the vector x relative to the basis $e_1,...,e_n$. Similarly, a vector y is a linear combination of the base vectors:

$$y = y^j e_j .$$

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

Let u be a unit of the scalar set S. Any scalar s in S scales with the unit:

$$s = s_M u ,$$

The number s_M is the magnitude of the scalar s relative to the unit u.

Because $g(e_i,e_j)$ is a scalar in S, we write

$$g(e_i,e_j) = g_{ij} u .$$

This expression defines n^2 numbers g_{ij} as the components of the bilinear form relative to the basis $e_1,...,e_n$ and the unit u.

We can list the components of the bilinear from as a matrix:

$$\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

Here we list the matrix for the case n = 2. The first subscript indicates the row, and the second subscript indicates the column.

Bilinear form as a sum of products. Because $g(x,y)$ is a bilinear map, we write

$$g(x,y) = g\left(x^i e_i, y^i e_i\right) = g\left(e_i, e_i\right) x^i y^j = g_{ij} x^i y^j u .$$

Comparing this expression with $s = g(x,y)$ and $s = s_M u$, we obtain that

$$s_M = g_{ij} x^i y^j .$$

The magnitude s_M of a scalar is calculated from the components of the bilinear form and the components of the two vectors.

We can write the sum term by term:

$$s_M = g_{11} x^1 y^1 + g_{12} x^1 y^2 + g_{21} x^2 y^1 + g_{22} x^2 y^2 .$$

We have used the case n = 2. We can also write the above expression in the matrix notation:

$$s_M = \begin{bmatrix} x^1 & x^2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix}$$

Change of basis and unit. Let $e_1,...,e_n$ be a basis of the vector space *V*, and $\tilde{e}_1,...,\tilde{e}_n$ be another basis of the same vector space. The two bases are related as

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

where the real numbers p_j^i relate the two bases. A vector *x* in *V* is a linear combination of either basis:

$$x = x^i e_i = \tilde{x}^i \tilde{e}_i ,$$

where the numbers $x^1,...,x^n$ are the components of *x* relative to the basis $e_1,...,e_n$, and the numbers $\tilde{x}^1,...,\tilde{x}^n$ are the components of *x* relative to the basis $\tilde{e}_1,...,\tilde{e}_n$. The two sets of components are related as

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

The vector is *contravariant*.

Let *u* be a unit of the scalar set *S*, and $\tilde{u}$ be another unit of the same scalar set. The two units are related as

$$\tilde{u} = ru ,$$

where the real number *r* converts the two units. A scalar *s* in *S* scales with either unit:

$$s = s_M u = \tilde{s}_M u ,$$

where the real number s_M is the magnitude of s relative to the unit u, and the number $\tilde{s}_M$ in F is the magnitude of s relative to the unit $\tilde{u}$. The two magnitudes are related as

$$s_M = r\tilde{s}_M .$$

The scalar is *contravariant*.

Because $\tilde{e}_i$ and $\tilde{e}_j$ are two vectors in V, the bilinear linear form $g\left(\tilde{e}_i,\tilde{e}_j\right)$ maps the two vectors to a scalar in S. Write

$$g\left(\tilde{e}_i,\tilde{e}_j\right) = \tilde{g}_{ij}\tilde{u} .$$

The numbers $\tilde{g}_{ij}$ are the components of the bilinear form g relative to the basis $\tilde{e}_1,...,\tilde{e}_n$ and the unit $\tilde{u}$. Because $g\left(\tilde{e}_i,\tilde{e}_j\right)$ is a linear function, we write

$$g\left(\tilde{e}_i,\tilde{e}_j\right) = g\left(p_i^k e_k, p_j^l e_l\right) = p_i^k p_j^l g\left(e_k,e_l\right) = p_i^k p_j^l g_{kl} u .$$

A comparison of the above two expressions gives that

$$\tilde{g}_{ij} = p_i^k p_j^l g_{kl} / r .$$

The components of the bilinear form transform in the same way as the basis of V, but in the opposite way as the unit of S. We say that the bilinear form is *covariant* with respect to the vector, but *contravariant* with respect to the scalar.

Symmetric bilinear form. A bilinear form is symmetric if

$$g\left(x,y\right) = g\left(y,x\right)$$

for every x and y in V.

Let $e_1,...,e_n$ be a basis of the vector space V, and u be a unit of the scalar set S. Relative to the basis and the unit, the bilinear form is

$$g\left(x,y\right) = g_{ij}x^i y^j u .$$

The sum contains terms like $g_{12}x^1 y^2 u$ and $g_{21}x^2 y^1 u$. The bilinear form is symmetric if and only if

$$g_{ij} = g_{ji}$$

for every i and j. That is, g_{ij} is a symmetric matrix.

Quadratic form. Let $s = g\left(x,y\right)$ be a bilinear map from a vector space V to a scalar set S. The bilinear form defines a function

$$Q\left(x\right) = g\left(x,x\right)$$

Reducing a quadratic form into a sum of squares.
Positive-definite inner product.

Non-degenerative.
Isomorphism.
One vector field, many scalar sets, generate a lot but liner maps
Invariants
Vector space of 1 d, nsquqred
Inner product
Cross product

Want a scalar to represent the size of a vector.
Scalar we can rank one d all elements
Zero vector rank a university. Rank a reseracher. Number of papers, times citation s of paper, years after PhD. Flawed.
Desired attribute

Zero vector has zero size
Positive and negative vectors have the same size
The size depends on vector smoothly
Size is scalar. A map from a vector space to a scalar.
Taylor expansion. Leading order term is quadratic. We decide go no further. Flat space.
Still need to determine the quadratic form.
There exists a basis, quadratic form is a sum
The number of positive and negative are fixed. Exclude zero.
Alpha times x gives alpha size

References

iMechanica users, Textbooks on linear algebra, http://imechanica.org/node/15843. This iMechanica thread goes much beyond recommendations of textbook. Topics include tensors and ways in which mechanician use old tensors and create new ones.

iMechanica users, Scalar done wrong, http://imechanica.org/node/15857. This iMechanica thread initiated the discussion on scalar sets.

S. Axler, Linear Algebra Done Right, 2nd ed. Springer, 1997.

W.C. Brown, A Second Course in Linear Algebra, Wiley, 1988.

P.R. Halmos, Finite-Dimensional Vector Spaces, 1958.

K. Hoffman and R. Kunze, Linear Algebra, 2nd ed., 1971.

I.M. Gelfand, Lectures on Linear Algebra, Dover Publications, New York, 1961.

N. Jeevanjee, An Introduction to Tensors and Group Theory for Physicists. Birkhauser, 2010.

G.E. Shilov, Linear Algebra, Dover Publications, New York, 1977.

S. Treil, Linear Algebra Done Wrong, free on line: http://www.math.brown.edu/~treil/papers/LADW/LADW.html

李开泰,黄艾香, 张量分析极其应用, 西安交通大学出版社, 1984.