Linear Mappings

Daoping Zhang

The Chinese University of Hong Kong

dpzhang@math.cuhk.edu.hk

April 22, 2020

Outline

Mappings

Let S, S' be two sets. A mapping from S to S' is an association which to every element of S associates an element of S'. Instead of saying that F is a mapping from S into S', we shall often write the symbols $F : S \to S'$. A mapping will also be called a map, for the sake of brevity.

A function is a special type of mapping, namely it is a mapping from a set into the set of numbers, i.e. into $\mathbb{R}$.

If $T : S \to S'$ is a mapping, and if u is an element of S, then we denote by $T(u)$, or Tu, the element of S' associated to u by T. We call $T(u)$ the value of T at u, or also the image of u under T. The symbols $T(u)$ are read "T of u". The set of all elements $T(u)$, when u ranges over all elements of S, is called the image of T. If W is a subset of S, then the set of elements $T(w)$, when w ranges over all elements of W, is called the image of W under T, and is denoted by $T(w)$.

Let $F : S \to S'$ be a map from a set S into a set S'. If x is an element of S, we often write

$$x \mapsto F(x)$$

with a special arrow $\mapsto$ to denote the image of x under F. Thus, for instance, we would speak of the map F such that $F(x) = x^2$ as the map $x \mapsto x^2$.

For any set S we have the identity mapping $I : S \to S$. It is defined by $I(x) = x$ for all x.

Let S and S' be both equal to $\mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^2$ (i.e. the function whose value at a number x is x^2). Then f is a mapping from $\mathbb{R}$ into $\mathbb{R}$. Its image is the set of numbers ≥ 0.

Let S be the set of numbers ≥ 0, and let $S' = \mathbb{R}$. Let $g : S \to S'$ be the function such that $g(x) = x^{1/2}$. Then g is a mapping from S into $\mathbb{R}$.

Let S be the set of functions having derivatives of all orders on the interval $0 < t < 1$, and let $S' = S$. Then the derivative $D = d/dt$ is a mapping from S into S. Indeed, our mapping D associates the function $df/dt = Df$ to the function f. According to our terminology, Df is the value of the mapping D at f.

Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be the mapping given by

$$\text{F(x,y)} = \text{(2x,2y)}.$$

Describe the image under F of the points lying on the circle $x^2 + y^2 = 1$. Let (x, y) be a point on the circle of radius 1. Let $u = 2x$ and $v = 2y$. Then u, v satisfy the relation

$$(u/2)^2 + (v/2)^2 = 1$$

or in other words,

$$\tfrac{u^2}{4} + \tfrac{v^2}{4} = 1.$$

Hence (u, v) is a point on the circle of radius 2. Therefore the image under F of the circle of radius 1 is a subset of the circle of radius 2. Conversely, given a point (u, v) such that

$$u^2 + v^2 = 4,$$

let $x = u/2$ and $y = v/2$. Then the point (x, y) satisfies the equation

$$x^2 + y^2 = 1,$$

and hence is a point on the circle of radius 1. Furthermore,

$$F(x, y) = (u, v).$$

Hence every point on the circle of radius 2 is the image of some point on the circle of radius 1. We conclude finally that the image of the circle of radius 1 under F is precisely the circle of radius 2.

In general, let S, S' be two sets. To prove that $S = S'$, one frequently proves that S is a subset of S' and that S' is a subset of S.

Let V be a vector space, and let u be a fixed element of V. We let

$$T_u : V \to V$$

be the map such that $T_u(v) = v + u$. We call T_u the translation by u. If S is any subset of V, then $T_u(S)$ is called the translation of S by u, and consists of all vectors $v + u$, with $v \in S$. We often denote it by $S + u$. Rotation counterclockwise around the origin by an angle θ is a mapping, which we may denote by $\mathbb{R}_\theta$. Let $\theta = \pi/2$. The image of the point $(1, 0)$ under the rotation $\mathbb{R}_{\pi/2}$ is the point $(0, 1)$. We may write this as

$$\mathbb{R}_{\pi/2}(1, 0) = (0, 1).$$

Linear Mappings

Let V, W be two vector spaces. A linear mapping

$$L : V \to W$$

is a mapping which satisfies the following two properties. First, for any elements u, v in V, and any scalar c, we have:
LM1 $L(u + v) = L(u) + L(v)$.
LM2 $L(cu) = cL(u)$.

The most important linear mapping of this course is described as follows. Let A be a given $m \times n$ matrix. Define

$$L_A : \mathbb{R}^n \to \mathbb{R}^m$$

by the formula

$$L_A(X) = AX.$$

Then L_A is linear. Indeed, this is nothing but a summary way of expressing the properties

$$A(X+Y) = AX + AY \quad \text{and} \quad A(cX) = cAX$$

for any vertical X, Y in $\mathbb{R}^n$ and any number c.

The dot product is essentially a special case. Let $A = (a_1, ..., a_n)$ be a fixed vector, and define

$$L_A(X) = A \cdot X.$$

Then L_A is a linear map from $\mathbb{R}^n$ into $\mathbb{R}$, because

$$A \cdot (X + Y) = A \cdot X + A \cdot Y \quad \text{and} \quad A \cdot (cX) = c(A \cdot X).$$

Note that the dot product can also be viewed as multiplication of matrices if we view A as a row vector, and X as a column vector.

Let $V = \mathbb{R}^3$ be the vector space of vectors in 3-space. Let $V' = \mathbb{R}^2$ be the vector space of vectors in 2-space. We can define a mapping.

$$F : \mathbb{R}^3 \to \mathbb{R}^2$$

by the projection, namely $F(x, y, z) = (x, y)$, which is a linear mapping. More generally, suppose $n = r + s$ is expressed as a sum of two positive integers. We can separate the coordinates $(x_1, ..., x_n)$ into two bunches $(x_1, ..., x_r, x_{r+1}, ..., x_{r+s})$, namely the first r coordinates, and the last s coordinates. Let

$$F : \mathbb{R}^n \to \mathbb{R}^r$$

be the map such that $F(x_1, ..., x_n) = (x_1, ..., x_r)$. Then you can verify easily that F is linear. We call F the projection on the first r coordinates. Similarly, we would have a projection on the last s coordinates, by means of the liner map L such that

$$L(x_1, ..., x_n) = (x_{r+1}, ..., x_n).$$

Let $L : V \to W$ be a linear mapping. Let u, v, w be elements of V. Then

$$L(u + v + w) = L(u) + L(v) + L(w).$$

This can be seen stepwise, using the definition of linear mappings. Thus

$$L(u + v + w) = L(u + v) + L(w) = L(u) + L(v) + L(w).$$

Similarly, given a sum of more than three elements, an analogous property is satisfied. For instance, let $u_1, ..., u_n$ be elements of V. Then

$$L(u_1 + \cdots + u_n) = L(u_1) + \cdots + L(u_n).$$

The sum on the right can be taken in any order.

If $a_1, ..., a_n$ are numbers, then

$$L(a_1 u_1 + \cdots + a_n u_n) = a_1 L(u_1) + \cdots + a_n L(u_n).$$

We show this for three elements.

$$\begin{aligned} L(a_1 u + a_2 v + a_3 w) &= L(a_1 u) + L(a_2 v) + L(a_3 w) \\ & a_1 L(u) + a_2 L(v) + a_3 L(w). \end{aligned}$$

With the notation of summation signs, we would write

$$L(\textstyle\sum_{i=1}^n a_i u_i) = \sum_{i=1}^n a_i L(u_i).$$

Property

Let $L : V \to W$ be a linear map. Then $L(O) = O$.

Proof.

We have

$$L(O) = L(O + O) = L(O) + L(O).$$

Subtracting $L(O)$ from both sides yields $O = L(O)$, as desired. $\square$

Property

Let $L : V \to W$ be a linear map. Then $L(-v) = -L(v)$.

Proof.

We have

$$O = L(O) = L(v - v) = L(v) + L(-v).$$

Add $-L(v)$ to both sides to get the desired assertion. $\square$

We observe that the values of a linear map are determined by knowing the values on the elements of a basis.
Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be a liner map. Suppose that

$$L(1,1) = (1,4) \quad \text{and} \quad L(2,-1) = (-2,3).$$

Find $L(3,-1)$.
To do this, we write $(3,-1)$ as a linear combination of $(1,1)$ and $(2,-1)$.
Thus we have to solve

$$(3,-1) = x(1,1) + y(2,-1).$$

The solution is $x = \frac{1}{3}, y = \frac{4}{3}$. Hence,

$$L(3,-1) = xL(1,1) + yL(2,-1) = \frac{1}{3}(1,4) + \frac{4}{3}(-2,3) = (\frac{-7}{3}, \frac{16}{3}).$$

Let first

$$F : V \to \mathbb{R}^n$$

be any mapping. Then each value $F(v)$ is an element of $\mathbb{R}^n$, and so has coordinates. Thus we can write

$$F(v) = (F_1(v), ..., F_n(v)), \quad \text{or} \quad F = (F_1, ..., F_n).$$

Each F_i is a function of V into $\mathbb{R}$, which we write

$$F_i : V \to \mathbb{R}.$$

Let $F : \mathbb{R}^2 \to \mathbb{R}^3$ be the mapping

$$F(x, y) = (2x - y, 3x + 4y, x - 5y).$$

Then

$$F_1(x, y) = 2x - y, \quad F_2(x, y) = 3x + 4y, \quad F_3(x, y) = x - 5y.$$

Observe that each coordinate function can be expressed in terms of a dot product. For instance, let

$$A_1 = (2, -1), \quad A_2 = (3, 4), \quad A_3 = (1, -5).$$

Then

$$F_i(x, y) = A_i \cdot (x, y) \quad \text{for} \quad i = 1, 2, 3.$$

Each function

$$X \mapsto A_i \cdot X$$

is linear.

Proposition

Let $F : V \to \mathbb{R}^n$ be a mapping of a vector space V into $\mathbb{R}^n$. Then F is linear if and only if each coordinate function $F_i : V \to \mathbb{R}$ is linear, for $i = 1, ..., n$.

Proof.

For $v, w \in V$, we have

$$F(v + w) = (F_1(v + w), ..., F_n(v + w)),$$
$$F(v) = (F_1(v), ..., F_n(v)),$$
$$F(w) = (F_1(w), ..., F_n(w)).$$

Thus $F(v + w) = F(v) + F(w)$ if and only if $F_i(v + w) = F_i(v) + F_i(w)$ for all $i = 1, ..., n$ by the definition of n-tuples. The same argument shows that if $c \in \mathbb{R}$, then $F(cv) = cF(v)$ if and only if

$$F_i(cv) = cF_i(v) \quad \text{for all } i = 1, ..., n.$$

This proves the propostion. □

Let V, W be two vectors spaces. We consider the set of all linear mappings from V into W, and denote this set by $\mathcal{L}(V, W)$, or simply $\mathcal{L}$ if the reference to V and W is clear. We shall define the addition of linear mappings and their multiplication by numbers in such a way as to make $\mathcal{L}$ into a vector space.

Let $L : V \to W$ and $F : V \to W$ be two linear mappings. We define their sum $L + F$ to be the map whose value at an element u of V is $L(u) + F(u)$. Thus we may write

$$(L + F)(u) = L(u) + F(u).$$

The map $L + F$ is then a linear map. Indeed, it is easy to verify that the two conditions which define a liner map are satisfied. For any elements u, v of V, we have

$$\begin{aligned}(L + F)(u + v) &= L(u + v) + F(u + v)\\ &= L(u) + L(v) + F(u) + F(v)\\ &= (L + F)(u) + (L + F)(v).\end{aligned}$$

Furthermore, if c is a number, then

$$\begin{aligned}(L+F)(cu) &= L(cu)+F(cu)\\ &= cL(u)+cF(u)\\ &= c[L(u)+F(u)]\\ &= c[(L+F)(u)].\end{aligned}$$

Hence $L+F$ is a linear map.
If a is a number, and $L : V \to W$ is a linear map, we define a map aL from V into W by giving its value at an element u of V, namely $(aL)(u) = aL(u)$. Then it is easily verified that aL is a linear map.

We have just defined operations of addition and multiplication by numbers in our set $\mathcal{L}$. Furthermore, if $L : V \to W$ is a linear map, i.e. an element of $\mathcal{L}$, then we define $-L$ to be $(-1)L$, i.e. the product of the number -1 by L. Finally, we have the zero-map, which is to every element of V associates the element O of W. Then $\mathcal{L}$ is a vector space. In other words, the set of linear maps from V into W is itself a vector space. The verification that the rules VS 1 through VS 8 for a vector space are satisfied is easy.

Let $V = W$ be the vector space of functions which have derivatives of all orders. Let D be the derivative, and let I be the identity. If f is in V, then

$$(D + I)f = Df + f.$$

Thus, when $f(x) = e^x$, then $(D + I)f$ is the function whose value at x is $e^x + e^x = 2e^x$.
If $f(x) = sinx$, then $(D + 3I)f$ is the function such that

$$((D + 3I)f)(x) = (Df)(x) + 3If(x) = cosx + 3sinx.$$

We note that $3 \cdot I$ is a linear map, whose value at f is $3f$. Thus $(D + 3 \cdot I)f = Df + 3f$. At any number x, the value of $(D + 3 \cdot I)f$ is $Df(x) + 3f(x)$. We can also write $(D + 3I)f = Df + 3f$.

The Kernel and Image of a Linear Map

Let $F : V \to W$ be a linear map. The image of F is the set of elements w in W such that there exists an element v of V such that $F(v) = w$.

Property

The image of F is a subspace of W.

Proof.

Observe first that $F(O) = O$, and hence O is in the image. Next, suppose that w_1, w_2 are in the image. Then there exist elements v_1, v_2 of V such that $F(v_1) = w_1$ and $F(v_2) = w_2$. Hence

$$F(v_1 + v_2) = F(v_1) + F(v_2) = w_1 + w_2,$$

thereby proving that $w_1 + w_2$ is in the image. If c is a number, then

$$f(cv_1) = cF(v_1) = cw_1.$$

Hence cw_1 is in the image. This proves that the assertion. $\square$

Let V, W be vectors spaces, and let $F : V \to W$ be a linear map. The set of elements $v \in V$ such that $F(v) = O$ is called the kernel of F.

Property

The kernel of F is a subspace of V.

Proof.

Since $F(O) = O$, we see that O is in the kernel. Let v, w be in the kernel. Then $F(v + w) = F(v) + F(w) = O + O = O$, so that $v + w$ is in the kernel. If c is a number, then $F(cv) = cF(v) = O$ so that cv is also in the kernel. Hence the kernel is a subspace. $\square$

Let $L : \mathbb{R}^3 \to \mathbb{R}$ be the map such that

$$L(x, y, z) = 3x - 2y + z.$$

Thus if $A = (3, -2, 1)$, then we can write

$$L(x) = X \cdot A = A \cdot X.$$

Then the kernel of L is the set of solutions of the equation.

$$3x - 2y + z = 0.$$

Of course, this generalizes to n-space. If A is an arbitrary vector in $\mathbb{R}^n$, we can define the linear map

$$L_A : \mathbb{R}^n \to \mathbb{R}$$

such that $L_A(X) = A \cdot X$. Its kernel can be interpreted as the set of all X which are perpendicular to A.

Theorem

Let $F : V \to W$ be a linear map whose kernel is $\{O\}$. If $v_1, ..., v_n$ are linearly independent elements of V, then $F(v_1), ..., F(v_n)$ are linearly independent elements of W.

Proof.

Let $x_1, ..., x_n$ be numbers such that

$$x_1 F(v_1) + \cdots + x_n F(v_n) = O.$$

By linearity, we get

$$F(x_1 v_1 + \cdots + x_n v_n) = O.$$

Hence $x_1 v_1 + \cdots + x_n v_n = O$. Since $v_1, ..., v_n$ are linearly independent it follows that $x_i = 0$ for $i = 1, ..., n$. This proves the theorem. □

We often abbreviate kernel and image by writing Ker and Im respectively. The next theorem relates the dimensions of the kernel and image of a linear map, with the dimension of the space on which the map is defined.

Theorem

Let V be a vector space. Let $L : V \to W$ be a linear map of V into another space W. Let n be the dimension of V, q the dimension of the kernel of L, and s the dimension of the image of L. Then $n = q + s$. In other words,

$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L.$$

Proof

If the image of L consists of O only, then our assertion is trivial. We may therefore assume that $s > 0$. Let $\{w_1, ..., w_s\}$ be a basis of the image of L. Let $v_1, ..., v_s$ be elements of V such that $L(v_i) = w_i$ for $i = 1, ..., s$. If the kernel is not $\{O\}$, it is understood that all reference to $\{u_1, ..., u_q\}$ is to be omitted in what follows. We contend that

$$\{v_1, ..., v_s, u_1, ..., u_q\}$$

is a basis of V. This will suffice to prove our assertion. Let v be any element of V. Then there exist numbers $x_1, ..., x_s$ such that

$$L(v) = x_1 w_1 + \cdots + x_s w_s,$$

because $\{w_1, ..., w_s\}$ is a basis of the image of L. By linearity,

$$L(v) = L(x_1 v_1 + \cdots + x_s v_s),$$

and again by linearity, subtracting the right-hand side from the left-hand side, it follows that

Proof

$$L(v - x_1 v_1 - \cdots - x_s v_s) = O.$$

Hence $v - x_1 v_1 - \cdots - x_s v_s$ lies in the kernel of L, and there exist numbers $y_1, ..., y_q$ such that

$$v - x_1 v_1 - \cdots - x_s v_s = y_1 u_1 + \cdots + y_q u_q.$$

Hence

$$v = x_1 v_1 + \cdots + x_s v_s = y_1 u_1 + \cdots + y_q u_q.$$

is a linear combination of $v_1, ... v_s, u_1, ..., u_q$. This proves that these $s + q$ elements of V generate V.
We now show that they are linearly independent, and hence that they constitute a basis. Suppose that there exists a linear relation:

$$x_1 v_1 + \cdots + x_s v_s + y_1 u_1 + \cdots + y_q u_q = O.$$

Proof.

Applying L to this relation, and using the fact that $L(u_j) = O$ for $j = 1, ..., q$, we obtain

$$x_1 L(v_1) + \cdots + x_s L(v_s) = O.$$

But $L(v_1), ..., L(v_s)$ are none other that $w_1, ..., w_s$, which have been assumed linearly independent. Hence $x_i = 0$ for $i = 1, ..., s$. Hence

$$y_1 u_1 + \cdots + y_q u_q = O.$$

But $u_1, ..., u_q$ constitue a basis of the kernel of L, and in particular, are linearly independent. Hence all $y_j = 0$ for $j = 1, ..., q$. This concludes the proof of our assertion. □

The linear map $L : \mathbb{R}^3 \to \mathbb{R}$ is given by the formula

$$L(x, y, z) = 3x - 2y + z.$$

Its kernel consists of all solutions of the equation

$$3x - y + z = 0.$$

Its image is a subspace of $\mathbb{R}$, is not $\{O\}$, and hence consists of all of $\mathbb{R}$. Thus its image has dimension 1. Hence its kernel has dimension 2.
The image of the projection

$$P : \mathbb{R}^3 \to \mathbb{R}^2$$

is all of $\mathbb{R}^2$, and the kernel has dimension 1.

The Rank and Linear Equations Again

Let A be an $m \times n$ matrix,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & & a_{mn} \end{pmatrix}.$$

Let $L_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear map which has been defined previously, namely

$$L_A(X) = AX.$$

As we have mentioned, the kernel of L_A is the space of solutions of the system of linear equations written briefly as

$$AX = O.$$

Let us now analyze its image.

Let $E^1, \ldots, E^n$ be the standard unit vectors of $\mathbb{R}^n$, written as column vectors, so

$$E^1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad E^n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Then ordinary matrix multiplication shows that

$$AE^j = A^j$$

is the j-th column of A. Consequently for any vector

$$X = x_1 E^1 + \cdots + x_n E^n,$$

we find that

$$AX = L_A(X) = x_1 A^1 + \cdots + x_n A^n.$$

Thus we see:

Theorem

The image of L_A is the subspace generated by the columns of A.

Previously, we gave a name to the dimension of that space, namely the column rank, which we have already seen is equal to the row rank, and is simply called the rank of A. Now we can interpret this rank also in the following way:

The rank of A is the dimension of the image of L_A.

Theorem

Let r be the rank of A. Then the dimension of the space of solutions of $AX = O$ is equal to $n - r$.

Find the dimension of the space of solutions of the system of equations

$$\begin{array}{ccccccccc} 2x & - & y & + & z & + & 2w & = & 0; \\ x & + & y & - & 2z & - & w & = & 0. \end{array}$$

We recall that the system of linear equations could also be written in the form

$$X \cdot A_i = 0 \quad \text{for } i = 1, \ldots m,$$

where A_i are the rows of the matrix A. This means that X is perpendicular to each row of A. Then X is also perpendicular to the row space of A, i.e. to the space generated by the rows. It is now convenient to introduce some terminology.
Let U be a subspace of $\mathbb{R}^n$. We let

$$U^\perp = \text{set of all elements } X \text{ in } \mathbb{R}^n \text{ such that } X \cdot Y = 0 \text{ fro all } Y \text{ in } U.$$

We call $U^\perp$ the orthogonal complement of U. It is the set of vectors which are perpendicular to all elements of U, or as we shall also say, perpendicular to U itself. Then it is easily verified that $U^\perp$ is a subspace.

Let U be the subspace generated by the row vectors of the matrix $A = (a_{ij})$. Then its orthogonal complement $U^\perp$ is precisely the set of solutions of the homogeneous equations

$$X \cdot A_i = 0 \quad \text{for all } i.$$

In other words, we have

$$(\text{row space of } A)^\perp = \operatorname{Ker} L_A = \text{space of solutions of } AX = O.$$

Theorem

Let U be a subspace of $\mathbb{R}^n$. Then

$$\dim U + \dim U^\perp = n.$$

In 3-dimensional space, for instance, this theorem proves the fact that the orthogonal complement of a line is a plane, and vice versa.

Let us now discuss briefly non-homogeneous equations, i.e. a system of the form

$$AX = B,$$

where B is a given vector (m-tuple). Such a system may not have a solution, in other words, the equations may be what is called "inconsistent".

Consider the system

$$\begin{array}{ccccccc} 3x & - & y & + & z & = & 1 \\ 2x & + & y & - & z & = & 2, \\ x & - & 2y & + & 2z & = & 5. \end{array}$$

There cannot be a solution to the above system of equations.

Theorem

Consider a non-homogeneous system of linear equations

$$AX = B.$$

Suppose that there exists at least one solution X_0. Then the set of solutions is precisely

$$X_0 + Ker\ L_A.$$

In other words, all the solutions are of the form

$$X_0 + Y, \text{ where } Y \text{ is a solution of } AY = O.$$

When there exists one solution at least to the system $AX = B$, then dim Ker L_A is called the dimension of the set of solutions. It is the dimension of the homogeneous system.

Proof.

Let $Y \in \text{Ker } L_A$. This means $AY = O$. Then

$$A(X_0 + Y) = AX_0 + AY = B + O = B.$$

so $X_0 + \text{Ker } L_A$ is contained in the set of solutions. Conversely, let X be any solution of $AX = B$. Then

$$A(X - X_0) = AX - AX_0 = B - B = O.$$

Hence $X = X_0 + (X - X_0)$, where $X - X_0 = Y$ and $AY = O$. This proves the theorem. $\square$

Find the dimension of the set of solutions of the following system of equations, and determine this set in $\mathbb{R}^3$.

$$\begin{array}{ccccccc} 2x & + & y & + & z & = & 1 \\ & & y & - & z & = & 0. \end{array}$$

We see by inspection that there is at least one solution, namely $x = 1/2, y = z = 0$. The rank of the matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

is 2. Hence the dimension of the set of solutions is 1. The vector space of solutions of the homogeneous system has dimension 1, and one solution is easily found to be

$$y = z = 1,\ x = -1.$$

Hence the set of solutions of the inhomogeneous system is the set of all vectors

$$(1/2, 0, 0) + t(-1, 1, 1),$$

where t ranges over all real numbers. We see that our set of solutions is a straight line.

The Matrix Associated with a Linear Map

To every matrix A we have associated a linear map L_A. Conversely, given a linear map

$$L : \mathbb{R}^n \to \mathbb{R}^m,$$

we shall now prove that there is some associated matrix A such that $L = L_A$.

Let $E^1, ..., E^n$ be the unit column vectors of $\mathbb{R}^n$. For each $j = 1, ..., n$ let $L(E^j) = A^j$, where A^j is a column vector of $\mathbb{R}^m$. Then for every element X in $\mathbb{R}^n$ we can write

$$X = x_1 E^1 + \cdots + x_n E^n$$

and therefore

$$\begin{aligned} L(X) &= x_1 L(E^1) + \cdots + x_n L(E^n) \\ &= x_1 A^1 + \cdots + x_n A^n \\ &= AX \end{aligned}$$

where A is the matrix whose column vectors are $A^1, ..., A^n$. Hence $L = L_A$, which proves the theorem.

This matrix A will be called the matrix associated with the linear map L.

Let $L : \mathbb{R}^4 \to \mathbb{R}^2$ be the linear map such that

$$L(E^1) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad L(E^2) = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad L(E^3) = \begin{pmatrix} -5 \\ 4 \end{pmatrix}, \quad L(E^4) = \begin{pmatrix} 1 \\ 7 \end{pmatrix}.$$

We see that the matrix associated with L is the matrix

$$\begin{pmatrix} 2 & 3 & -5 & 1 \\ 1 & -1 & 4 & 7 \end{pmatrix}.$$

Let V be an n-dimensional vector space. If we pick some basis $\{v_1, ..., v_n\}$ of V, then every element of V can be written in terms of coordinates

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Thus to each element v of V we can associate the coordinate vector X. If

$$w = y_1 v_1 + \cdots + y_n v_n$$

so Y is the coordinate vector of w, then

$$v + w = (x_1 + y_1)v_1 + \cdots + (x_n + y_n)v_n,$$

so $X + Y$ is the coordinate vector of $v + w$. Let c be a number. Then

$$cv = cx_1 v_1 + \cdots + cx_n v_n,$$

so cX is the coordinate vector of cv. Thus after choosing a basis, we can identify V with $\mathbb{R}^n$ via the coordinate vectors.

Let $L : V \to V$ be a linear map. Then after choosing a basis which gives us an identification of V with $\mathbb{R}^n$, we can then represent L by a matrix. Different choices of bases will give rise to different associated matrices. Some choices of bases will ofter give rise to especially simple matrices. Suppose that there exists a basis $\{v_1, ..., v_n\}$ and numbers $c_1, ..., c_n$ such that

$$Lv_i = c_i v_i \quad \text{for} \quad i = 1, ..., n.$$

Then with respect to this basis, the matrix of L is the diagonal matrix

$$\begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \vdots & c_n \end{pmatrix}.$$

If we picked another basis, the matrix of L might not be so simple.

Let $\{v_1, ..., v_n\}$ be the given basis of V. Then there exist numbers c_{ij} such that

$$\begin{array}{ccccc} Lv_1 & = & c_{11}v_1 & +\cdots+ & c_{1n}v_n \\ \vdots & & \vdots & & \vdots \\ Lv_n & = & c_{n1}v_1 & +\cdots+ & c_{nn}v_n. \end{array}$$

What is the effect of L on the coordinate vector X of an element $v \in V$? Such an element is of the form

$$v = x_1v_1 + \cdots + x_nv_n.$$

Then

$$Lv = \sum_{i=1}^n x_i L(v_i) = \sum_{i=1}^n x_i \sum_{j=1}^n c_{ij}v_j = \sum_{j=1}^n (\sum_{i=1}^n x_i c_{ij})v_j.$$

Hence we find:
If $C = (c_{ij})$ is the matrix such that $L(v_i) = \sum_{j=1}^{n} c_{ij} v_j$, and X is the coordinate vector of v, then the coordinate vector of Lv is $C^t X$. In other words, on coordinate vectors, L is represented by the matrix C^t (transpose of C).
We note the transpose C rather than C itself. This is because when writing Lv_i as linear combination of $v_1, ..., v_n$ we have written it horizontally, whereas before we wrote it vertically in terms of the vertical unit vectors $E^1, ..., E^n$. We call C^t the matrix associated with L with respect to the given basis.

Let $L : V \to V$ be a linear map. Let $\{v_1, v_2, v_3\}$ be a basis of V such that

$$\begin{aligned} L(v_1) &= 2v_1 - v_2, \\ L(v_2) &= v_1 + v_2 - 4v_3, \\ L(v_3) &= 5v_1 + 4v_2 + 2v_3. \end{aligned}$$

Then the matrix associated with L on the coordinate vectors is the matrix

$$\begin{pmatrix} 2 & 1 & 5 \\ -1 & 1 & 4 \\ 0 & -4 & 2 \end{pmatrix}.$$

It is the transpose of the matrix

$$\begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & -4 \\ 5 & 4 & 2 \end{pmatrix}.$$