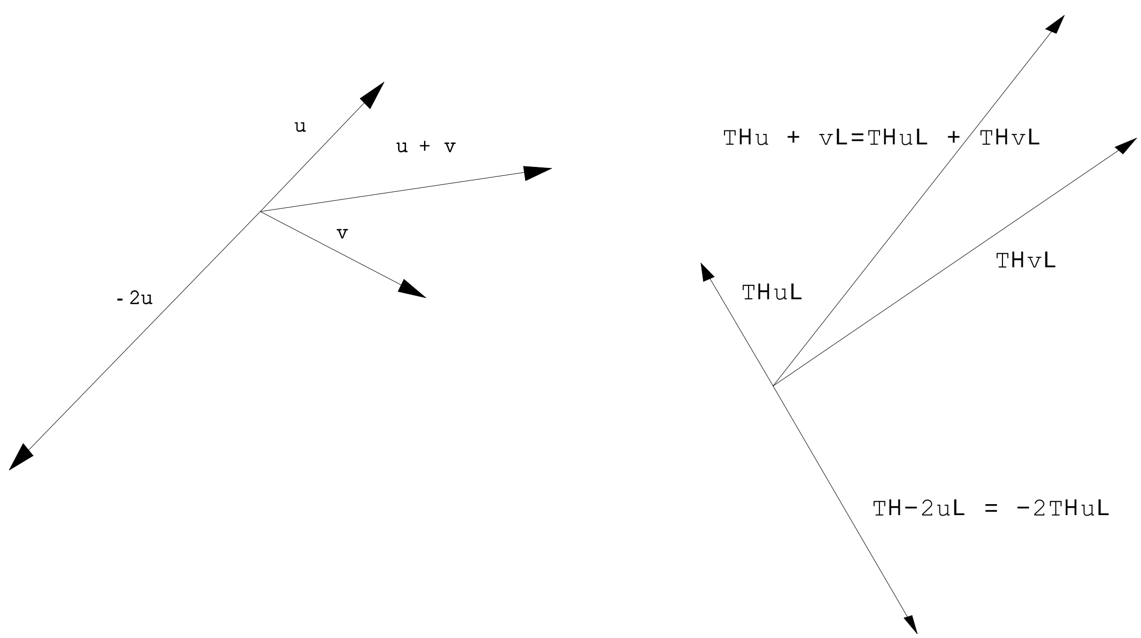
7 - Linear Transformations

Mathematics has as its objects of study sets with various structures. These sets include sets of numbers (such as the integers, rationals, reals, and complexes) whose structure (at least from an algebraic point of view) arise from the operations of addition and multiplication with their relevant properties. Metric spaces consist of sets of points whose structure comes from a distance function. Various sets of functions with certain properties make up other objects with their structure coming from the operation of composition. In linear algebra the objects are sets of vectors with the operations of addition and scalar multiplication providing the structure. In every instance the most interesting and useful functions between these sets are those that preserve the structure-whether that is preserving distance, closeness, sums or products. In linear algebra we call these functions or maps *linear transformations.*

Definition

Let V and W be vector spaces over the real numbers. Suppose that T is a function from V to W, T:V 6 W. T is **linear** (or a **linear transformation**) provided that T preserves vector addition and scalar multiplication, i.e. for all vectors **u** and **v** in V, T(**u + v**) = T(**u**) + T(**v**) and for any scalar c we have T(c**v**) = cT(**v**).



The picture above illustrates a linear transformation T:$\mathbf{R}^n$ 6 $\mathbf{R}^n$. If we assume that T is defined on the vectors **u** and **v** (where it could be chosen without restriction), then its linearity forces T(-2**u**) and T(**u** + **v**) to be as shown.

<u>Examples</u>

1. Define T: $\mathbf{R}^4 \to \mathbf{R}^2$ by T(a, b, c, d) = (a + 3c, b - c + 2d). Then T is a linear transformation for let $\mathbf{v_1} = (a_1, b_1, c_1, d_1)$ and $\mathbf{v_2} = (a_2, b_2, c_2, d_2)$ be vectors in $\mathbf{R}^4$.

$$\begin{aligned} T(\mathbf{v_1} + \mathbf{v_2}) &= T(a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2) \\ &= ((a_1 + a_2) + 3(c_1 + c_2), (b_1 + b_2) - (c_1 + c_2) + 2(d_1 + d_2)) \\ &= (a_1 + 3c_1, b_1 - c_1 + 2d_1) + (a_2 + 3c_2, b_2 - c_2 + 2d_2) \\ &= T(\mathbf{v_1}) + T(\mathbf{v_2}) \end{aligned}$$

Thus, T preserves addition.

Also, if **v** = (a, b, c, d) and k is any scalar, then

T(k**v**) = T(ka, kb, kc, kd) = (ka + 3(kc), kb - kc + 2(kd)) = k (a + 3c, b - c + 2d) = kT(**v**)

Hence, T preserves scalar multiplication and so is linear.

2. Define S: $\mathbf{R}^2 \to \mathbf{R}^4$ by S(x, y) = (0, x, x, y). For any vectors (x, y) and (s, t) and any scalar c:

$$\begin{aligned} S((x, y) + (s, t)) &= S(x + s, y + t) \\ &= (0, x + s, x + s, y + t) \\ &= (0, x, x, y) + (0, s, s, t) \\ &= S(x, y) + S(s, t) \end{aligned}$$

and S(c(x, y)) = S(cx, cy) = (0, cx, cx, cy) = c (0, x, x, y) = c S(x, y).

Therefore, S is a linear transformation.

3. Set $T(x, y, z) = (x^2, y + 3)$. Show that T is <u>not</u> linear. T does not preserve addition since

T((1, 0, 0) + (1, 0, 0)) = T(2, 0, 0) = (4, 3)

yet T(1, 0, 0) + T(1, 0, 0) = (1, 3) + (1, 3) = (2, 6).

This is sufficient to show that T is not linear, but we also show that T does not preserve scalar multiplication. Indeed, $T(2(0, 0, 0)) = T(0, 0, 0) = (0, 3) \neq (0, 6) = 2T(0, 0, 0)$.

4. An example of a linear transformation between polynomial vector spaces is $D: P_4 \to P_3$ given

by $D(ax^4 + bx^3 + cx^2 + dx + e) = 4ax^3 + 3bx^2 + 2cx + d$.

5. The map $T: M_{2\times 2} \to P_3$ defined by $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ax^3 + bx^2 + cx + d$ is a linear transformation.

6. Consider $S: P_3 \to \mathbf{R}^2$ given by $S(p(x)) = (p(1), p(2))$ where $p(x)$ is any vector in P_3 (and so any polynomial of degree 3 or less.) Is S linear?

7. Let $\mathbf{v}$ be some fixed vector in $\mathbf{R}^n$, say for example $\mathbf{v} = (1, 2, 3, 1, 2, 3, ...)$. Define a map, $T: \mathbf{R}^n \to \mathbf{R}$ by using the dot product setting $T(\mathbf{x}) = \mathbf{x} \cdot \mathbf{v}$. T is a linear transformation.

Linear transformations are defined as functions between vector spaces which preserve addition and multiplication. This is sufficient to insure that they preserve additional aspects of the spaces as well as the result below shows.

Theorem

Suppose that $T: V \to W$ is a linear transformation and denote the zeros of V and W by 0_v and 0_w, respectively. Then $T(0_v) = 0_w$.

Proof

Since $0_w + T(0_v) = T(0_v) = T(0_v + 0_v) = T(0_v) + T(0_v)$, the result follows by cancellation.

This property can be used to prove that a function is not a linear transformation. Note that in example 3 above $T(\mathbf{0}) = (0, 3) \neq \mathbf{0}$ which is sufficient to prove that T is not linear. The fact that a function may send $\mathbf{0}$ to $\mathbf{0}$ is not enough to guarantee that it is linear. Defining $S(x, y) = (xy, 0)$ we get that $S(\mathbf{0}) = \mathbf{0}$, yet S is not linear.

Definitions

Suppose that $T: V \to W$ is a linear transformation.

1. The **kernel** of T is defined by ker T = {**v** | T(**v**) = **0**}.
2. The **range** of T = {T(**v**) | **v** is in V}.

Theorem

Let T: V 6 W be a linear transformation. Then

1. Ker T is a subspace of V and
2. Range T is a subspace of W.

Proof

1. The kernel of T is not empty since **0** is in ker T by the previous theorem. Suppose that **u** and **v** are in ker T so that T(**u**) = **0** and T(**v**) = **0**. Then T(**u** + **v**) = T(**u**) + T(**v**) = **0** + **0** = **0**. Thus, **u** + **v** is in ker T. If, in addition, c is any scalar, we have T(c**v**) = cT(**v**) = c**0** = **0**. Hence, c**v** is in ker T which, therefore, is a subspace of V.
2. Since **0** is in V, T(**0**) is in the range of T which is not empty. Suppose that $\mathbf{w}_1$ and $\mathbf{w}_2$ are in range T. Then there exist vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V with $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. We then have $\mathbf{v}_1 + \mathbf{v}_2$ is a vector in V and $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$, i.e. $\mathbf{w}_1 + \mathbf{w}_2$ is in range T. Also if **w** is in range T, say with **v** in V and T(**v**) = **w**, and c is any scalar, then c**v** is in V and T(c**v**) = cT(**v**) = c**w** which shows that c**w** is in range T. Consequently, range T is a subspace of W.

Examples

1. Consider the linear transformation T(x, y, z) = (x - 3y + 5z, -4x + 12y, 2x - 6y + 8z). To compute the kernel of T we solve T(x, y, z) = **0**. This corresponds to the homogeneous system of linear equations

$$\begin{aligned} x - 3y + 5z &= 0 \\ -4x + 12y \phantom{{}+8z} &= 0 \\ 2x - 6y + 8z &= 0 \end{aligned}$$

So we reduce the coefficient matrix $\begin{pmatrix} 1 & -3 & 5 \\ -4 & 12 & 0 \\ 2 & -6 & 8 \end{pmatrix}$ to get $\begin{pmatrix} 1 & -3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Hence ker T = {(x, y, z) | x = 3y and z = 0} = < (3, 1, 0) >.

Range T = {T(x, y, z) | (x, y, z) is in $\mathbf{R}^3$}

= {(x - 3y + 5z, -4x + 12y, 2x - 6y + 8z)| x, y, and z are real numbers}

= {x(1, -4, 2) + y(-3, 12, -6) + z(5, 0, 8)| x, y, and z are real numbers}

= < (1, -4, 2), (5, 0, 8) >.

2. Let S: $\mathbf{R}^3$ 6 $M_{2\times 2}$ be given by S(x, y, z) = $\begin{pmatrix} x & z \\ y & 0 \end{pmatrix}$. Then S is linear and it is easy o see that

ker S = {(0, 0, 0)} and range S = $\left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\rangle$.

For any function f: X 6 Y, f is said to be *one-to-one* if f(a) = f(b) implies that a = b (no two elements of the domain of f map to the same element of Y.) For any linear transformation there is a straightforward method of determining whether or not it is one-to-one. It is an important reason why we are interested in kernels.

<u>**Theorem**</u>

Suppose that T: V 6 W is a linear transformation. T is one-to-one if and only if ker T = {**0**}.

Proof

Suppose that ker T = {**0**}. Let **a** and **b** be vectors in V with T(**a**) = T(**b**). Then T(**a - b**) = T(**a**) - T(**b**) = **0**. Thus, **a - b** is in the kernel of T, so **a - b** = **0**. Hence, **a** = **b** which shows that T is one-to-one.

Conversely, suppose that ker T … {**0**} say **v** is in ker T and **v** … **0**. We then have T(**v**) = **0** = T(**0**) yet **v** … **0**. Hence, T is not one-to-one. So if T is one-to-one, ker T = { **0**}.

We have seen that matrix multiplication distributes over addition (so, when the addends are column vectors, A(**x** + **y**) = A**x** + A**y**) and that scalars can be factored out of products (A(c**x**) = c(A**x**)). We have also defined the null space of A as the set of vectors **x** for which A**x** = **0** and seen that the column space of A is the set of all vectors for which there is a solution to A**x** = **b**, i.e. all vectors **b** such that there exists a vector **x** with A**x** = **b**. Thus we have the following

Theorem.

Let A be an m×n matrix. Define $T:\mathbf{R}^n \to \mathbf{R}^m$ by, for any **x** in $\mathbf{R}^n$, T(**x**) = A**x**. Then T is a linear transformation. Furthermore, the kernel of T is the null space of A and the range of T is the column space of A.

Thus matrix multiplication provides a wealth of examples of linear transformations between real vector spaces. In fact, every linear transformation (between finite dimensional vector spaces) can be thought of as matrix multiplication. We will see this shortly, but first a little ground work.

Suppose that B = $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ is a basis for the vector space V and **v** is any vector in V. By the unique representation theorem, $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$ where the scalars $c_1, c_2, ..., c_n$ are uniquely determined.

Definition

The **coordinate vector** of **v** with respect to the basis B, $\mathbf{v}_B$, is defined by setting

$$\mathbf{v}_B = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

Example

Let **v** = (50, -36, 134) and consider the bases B = {(1, 0, 0), (0, 1, 0), (0, 0, 1)} and B′ = {(2, -3, 5), (1, 4, -7), (6, 0, 9)} of $\mathbf{R}^3$. We have

$$\mathbf{v_B} = \begin{pmatrix} 50 \\ -36 \\ 134 \end{pmatrix} \quad \text{and } \mathbf{v_{BN}} = \begin{pmatrix} 4 \\ -6 \\ 8 \end{pmatrix}$$

(The latter since (50, -36, 134) = 4(2, -3, 5) - 6(1, 4, -7) + 8(6, 0, 9).)

Definition

Let V and W be vector spaces. V and W are **isomorphic** if there is a linear transformation $T: V \to W$ which is one-to-one and onto (i.e. range T = W) in which case T is called an **isomorphism.**

Vector spaces which are isomorphic have the same number of vectors and corresponding vectors in the two vector spaces behave in precisely the same way. The spaces have the same structure; they differ only in notation or name (as vector spaces.) The following theorem justifies the special attention accorded $\mathbf{R}^n$.

Theorem

Suppose that V is a vector space with dim V = n. Then V is isomorphic to $\mathbf{R}^n$.

Proof

Let $B = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ be a basis for V. Define $I_B: V \to \mathbf{R}^n$ as follows. For any $\mathbf{v}$ in V let $c_1, c_2, ..., c_n$ be scalars so that $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n}$. Set $I_B(\mathbf{v}) = (c_1, c_2, ..., c_n)$. Then I_B is an isomorphism. Notice that, in essence, $I_B(\mathbf{v}) = \mathbf{v_B}$.

Examples

1. P_n is isomorphic to $\mathbf{R}^{n+1}$ by, e.g. $T(a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_0) = (a_n, a_{n-1}, ..., a_1, a_0)$.
2. $M_{m\times n}$ is isomorphic to $\mathbf{R}^n$ by defining, for $A = (a_{ij})$, $S(A) = (a_{11}, a_{12}, ..., a_{1n}, a_{21}, a_{22}, ..., a_{nn})$.

Any linear transformation is determined by its effect on any basis of its domain. For suppose that $T: V \to W$ is linear and that $B = \{\mathbf{v_1}, \mathbf{v_2}, . . . , \mathbf{v_n}\}$ is a basis for V. For any vector $\mathbf{v}$ in V there are scalars c_i so that $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + . . . + c_n\mathbf{v_n}$. But then $T(\mathbf{v}) = c_1T(\mathbf{v_1}) + c_2T(\mathbf{v_2}) + . . . + c_nT(\mathbf{v_n})$.

So once the images of the basis elements $\mathbf{v_i}$ are fixed, the vectors T($\mathbf{v_i}$), the image of an arbitrary vector **v**, T(**v**), is forced. Now let BN = {$\mathbf{w_1}, \mathbf{w_2}, ..., \mathbf{w_n}$} be a basis for W. For each j = 1, 2, ..., n write T($\mathbf{v_j}$) = $\sum_{k=1}^{m} a_{kj} w_k$

Theorem

Suppose that V and W are vector spaces with bases B and BN, respectively. Given any linear transformation T: V 6 W set matrix A = (a_{ij}) where each a_{ij} is defined as the scalar above. Then for any **v** in V, T(**v**)$_{BN}$ = A$\mathbf{v_B}$, i.e. the coordinate vector of T(**v**) with respect to the basis BN is just the product of the coordinate vector of **v** with respect to the basis B times A. The matrix A is called the *matrix representation of T with respect to (w.r.t.) the bases B and BN.*

Example

Define T:P_2 6 $M_{2\times 2}$ by T($ax^2 + bx + c$) = $\begin{pmatrix} a+2b & b-c \\ 3a+4c & 5a \end{pmatrix}$. Using the bases

B = {$x^2 + 1$, $x + 2$, $x^2 + x$} and BN = $\left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}$

we compute a, the matrix of T w.r.t. B and BN. First we calculate the image of each element of B under T.

T($x^2 + 1$) = $\begin{pmatrix} 1 & -1 \\ 7 & 5 \end{pmatrix}$ T($x + 2$) = $\begin{pmatrix} 2 & -1 \\ 8 & 0 \end{pmatrix}$ T($x^2 + x$) = $\begin{pmatrix} 3 & 1 \\ 3 & 5 \end{pmatrix}$

Next the coordinate vectors of each of these w.r.t. BN are calculated.

T($x^2 + 1$)$_{BN}$ = $\begin{pmatrix} 5 \\ -6 \\ 8 \\ -6 \end{pmatrix}$ T($x + 2$)$_{BN}$ = $\begin{pmatrix} 0 \\ -1 \\ 9 \\ -6 \end{pmatrix}$ T($x^2 + x$)$_{BN}$ = $\begin{pmatrix} 5 \\ -4 \\ 2 \\ 0 \end{pmatrix}$

The first vector is correct since $\begin{pmatrix} 1 & -1 \\ 7 & 5 \end{pmatrix} = 5\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - 6\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + 8\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - 6\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Then we have A = $\begin{pmatrix} 5 & 0 & 5 \\ -6 & -1 & -4 \\ 8 & 9 & 2 \\ -6 & -6 & 0 \end{pmatrix}$.

We illustrate the results of the theorem with $\mathbf{v} = 4x^2 - 2x + 10$.

$\mathbf{v}_B = \begin{pmatrix} 22/3 \\ 4/3 \\ -10/3 \end{pmatrix}$ since $\mathbf{v} = 4x^2 - 2x + 10.= \frac{22}{3}(x^2+1) + \frac{4}{3}(x+2) - \frac{10}{3}(x^2+x)$.

Then $A\mathbf{v}_\mathbf{B} = \begin{pmatrix} 20 \\ -32 \\ 64 \\ -52 \end{pmatrix}$. Comparing we have $T(\mathbf{v}) = \begin{pmatrix} 0 & -12 \\ 52 & 20 \end{pmatrix}$ Then, since

$20\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} - 32\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + 64\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} - 52\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, we have $T(\mathbf{v})_{BN} = \begin{pmatrix} 20 \\ -32 \\ 64 \\ -52 \end{pmatrix} = A\mathbf{v}_\mathbf{B}$ as claimed.

Special Case.

Consider bases B_1 and B_2 of $\mathbf{R^n}$. Let I : $\mathbf{R^n}$ 6 $\mathbf{R^n}$ be the identity map, i.e. I(**v**) = **v** for every vector **v**, The matrix representation of I w.r.t. B_1 and B_2 (call it A) provides a change of basis. If a vector

v is expressed in terms of B_1 as v_{B_1}, then A v_{B_1} is **v** in terms of. the basis B_2, the coordinate vector of **v** w.r.t. B_2.

Using the isomorphisms introduced in the previous theorem we could express the results of the theorem above as $I_{B'} \circ T = A \circ I_B$ or $T = I_{B'}^{-1} \circ A \circ I_B$. Not only do linear transformation correspond to multiplying vectors by matrices, but the composition of linear transformations amounts to matrix multiplication.

Theorem

Suppose that U, V, and W are vector spaces with bases B, B′, and B″, respectively. Suppose that $S:U \to V$ and $T:V \to W$ are linear transformations with the matrix representation of S w.r.t. B and B′ being C and the matrix representation of T w.r.t. B′ and B″ the matrix A. Then the composition $S \circ T$ has as its matrix representation w.r.t. B and B″, the matrix product CA.

Now consider a linear transformation $T:\mathbf{R}^n \to \mathbf{R}^n$. Let A be the matrix representation of T w.r.t the standard basis of $\mathbf{R}^n$. Suppose that λ is an eigenvalue of A with eigenvector **x**. Then $A\mathbf{x} = \lambda\mathbf{x} = T(\mathbf{x})$ so T sends **x** to a scalar multiple of itself In fact, the space generated by **x** (a subspace of the eigenspace of λ) is mapped into itself by T.