Math 173

Alex Iosevich

1. Linear spaces: basic definitions

Let V be a non-empty set of objects, called elements. We say that V is a linear space if the following axioms hold:

1) For every $x, y \in V$, $x + y \in V$.

2) For every $x \in V$ and $a \in \mathbb{R}$, $ax \in V$.

3) For all $x, y \in V$, $x + y = y + x$.

4) For all $x, y, z \in V$, $(x + y) + z = x + (y + z)$.

5) There exists an element O such that $x + O = x$ for all $x \in V$.

6) For every $x \in V$, $x + (-1)x = O$.

7) For all $x \in V$ and all $a, b \in \mathbb{R}$, $a(bx) = (ab)x$.

8) For all $x, y \in V$ and $a \in \mathbb{R}$, $a(x + y) = ax + ay$.

9) For all $x \in V$ and $a, b \in \mathbb{R}$, $(a + b)x = ax + by$.

10) For every $x \in V$, $1 \cdot x = x$.

Example 1.1. Perhaps the simplest example of a linear space is $V = \mathbb{R}$, the field of real numbers. All the axioms above are satisfied instantly due to the familiar properties of $\mathbb{R}$.

Example 1.2. A slightly more interesting example is $\mathbb{R}^n$, the set of n-tuples of elements of $\mathbb{R}$. If $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ then $x+y$ is defined to be $(x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$. Once again the axioms are satisfied because matters can be reduced to the properties of $\mathbb{R}$. For example, if $x = (x_1, x_2, \ldots, x_n)$, then $-1 \cdot x = (-x_1, -x_2, \ldots, -x_n)$, so $x + (-1)x = (0, \ldots, 0)$, which is the O elements in this linear space.

Example 1.3. The previous example already leads to much interesting mathematics, but let's consider something slightly more exotic. Let V be the space of functions continuous on $[0, 1]$. Here addition is just the usual addition on $\mathbb{R}$ and the O elements is the function identically equal to 0.

The only properties that require checking are the closure under addition and multiplication. But we know from calculus that the sum of two continuous function is continuous and a scalar multiple of any continuous function is also continuous.

EXAMPLE 1.4. Yet another example that gives a taste of things to come is the linear space of all polynomials of degree n, where n is a positive integer. Any element of this space is of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

Any such polynomial is uniquely determined by $(n+1)$-tuples $(a_0, a_1, \ldots, a_n)$. This brings up an immediate question of how this space is *different* from $\mathbb{R}^{n+1}$. In particular, this example suggests that at some point we may want to introduce a notion of what it means for two linear space to be *the same* or *different*.

PROBLEM 1.5. Do problems 1-28 on page 7. Don't be scared, it is not as much as it looks...

We shall now prove some basic consequences of the axioms. We being by demonstrating that the zero element is unique.

THEOREM 1.6. *In any linear space there is one and only one zero element.*

We know that there is at least one zero element. Suppose that there are two of them, O_1 and O_2. Taking $x = O_1$ and $O = O_2$, we see that

$$O_1 + O_2 = O_1.$$

Similarly, taking $x = O_2$ and $O = O_1$, we see that

$$O_2 + O_1 = O_2.$$

By commutativity (axiom 3), $O_1 = O_2$ and we are done. This completes the proof.

THEOREM 1.7. *For every $x \in V$ there is one and only one element y such that $x + y = O$.*

Axiom 6 tells that we can take y to be $(-1)x$. Suppose that there are two elements y_1, y_2 such that $x + y_1 = x + y_2 = O$. It follows that

$$y_2 + (x + y_1) = y_2 + O = y_2$$

and

$$(y_2 + x) + y_1 = O + y_1 = y_1.$$

It follows that $y_1 = y_2$, so the only negative element of x is $(-1)x$. This completes the proof. And now comes a barrage of elementary properties that we shall use over and over.

THEOREM 1.8. *In a given vector space, let x, y denote arbitrary elements and a, b denote arbitrary scalars. Then the following properties hold:*

a) $0x = O$.

b) $aO = O$.

c) $(-a)x = (-ax) = a(-x)$.

d) If $ax = O$, *then either* $x = O$ *or* $a = 0$.

e) If $ax = ay$ *and* $a \neq 0$, *then* $x = y$.

f) If $ax = bx$ and $x \neq O$, then $a = b$.

g) $-(x + y) = (-x) + (-y) = -x - y$.

h) $x + x = 2x$, $x + x + x = 3x$, and in general $\sum_{i=1}^{n} x = nx$.

To prove a), observe that by axiom 9,

$$0x + 0x = (0 + 0)x = 0x.$$

Adding $(-1)0x$ to both sides yields $0x = O$, as desired.
To prove b) do the same thing using axiom 8.
To prove c), observe that

$$(-a)x + ax = (-a + a)x = 0x = O.$$

Similarly,

$$a(-x) + ax = a(-x + x) = a \cdot O = O.$$

This completes the proof of a), b), c).

PROBLEM 1.9. Prove d), e), f), g) and h).

2. Subspaces

We say that S is a subspace of a linear space V if S is linear subspace with the same operations of addition and multiplication by scalars.

EXAMPLE 2.1. Let V be the linear space of continuous function on $[0, 1]$ and let S be the linear space of continuous function on $[0, 1]$ that are equal to 0 at 0.

EXAMPLE 2.2. Let

$$V = \{x = (x_1, x_2) : x_j \in \mathbb{R}\}$$

and let

$$S = \{(t, 0) : t \in \mathbb{R}\}.$$

THEOREM 2.3. *Let S be a non-empty subset of a linear space V. Then S is a subspace if and only if S satisfies the closure axioms.*

One direction is obvious. Conversely, we must show that if S is a subset of V and satisfies the closure axioms, then it satisfies the other axioms as well. The only axioms that are not automatically satisfied are axioms 5 and 6, the existence of the 0 element and the existence of the additive inverse.

By multiplicative closure, $ax \in S$ whenever $x \in S$ for every $a \in \mathbb{R}$. It follows that $0x \in S$. But we proved in Theorem 1.8 above that $0x = O$, so $O \in S$, satisfying the axiom 5. Similarly, $-1 \cdot x \in S$ whenever $x \in S$. But $x + (-1)x = O$ since both x and $(-1)x$ are elements of V. Therefore the axiom 6 is also satisfied. This completes the proof.

DEFINITION 2.4. Let S be a non-empty subset of V. An element $x \in V$ of the form

$$x = \sum_{i=1}^{k} c_i x_i,$$

where c_is are scalars and x_is are elements of S, is called a finite linear combination of elements of S.

Note that the set of all linear combinations of elements of S satisfies the closure axioms and is thus a subspace of V, denoted by $L(S)$. The following example makes a connection between this notion and things we are very well familiar with.

EXAMPLE 2.5. Let $V = \mathbb{R}^2$, i.e $V = \{x = (x_1, x_2) : x_j \in \mathbb{R}\}$. Let $S = \{(1,0), (0,1)\}$, a two-element set. Observe that given any $x = (x_1, x_2) \in V$, $x = x_1(1,0) + x_2(0,1)$, so $L(S) = V$!!

EXAMPLE 2.6. Let $V = \mathbb{R}^2$ as above and let $S = \{(1,1), (2,3)\}$. Observe that

$$x = (x_1, x_2) = (3x_1 - 2x_2)(1,1) + (x_2 - x_1)(2,3),$$

so $L(S) = V$ once again!

PROBLEM 2.7. Let $V = \mathbb{R}^2$. Suppose that $S = \{(v_1, v_2), (w_1, w_2)\}$ and there does not exist $t \in \mathbb{R}$ so that $v_1 = tw_1$ and $v_2 = tw_2$. Prove that $L(S) = V$.

PROBLEM 2.8. Generalize the problem above to $V = \mathbb{R}^3$.

DEFINITION 2.9. We say that a set S in a linear space V is dependent if there is a finite collection of elements of S, say, $x_1, \dots, x_n$ and scalars $c_1, \dots, c_n$ such that

$$\sum_{i=1}^{n} c_i x_i = O.$$

Such an equation is called a non-trivial representation of O. The set is called independent if it is not dependent. Note that in this case, for any collection of vectors $x_1, \dots, x_n$ and scalars $c_1, \dots, c_n$, the condition

$$\sum_{i=1}^{n} c_i x_i = O \text{ implies that } c_1 = c_2 = \cdots = c_n = 0.$$

PROBLEM 2.10. Let $V = \mathbb{R}^2$. Prove that any three element set is dependent. Prove that a two element set is independent if and only if the two vectors are not scalars multiples of each other.

EXAMPLE 2.11. Let $u_k(t) = t^k$ for $k = 0, 1, \dots$ and $t \in \mathbb{R}$. Is the set $S = \{u_0, u_1, \dots\}$ independent? Suppose that

$$\sum_{i=0}^{n} c_k u_k = \sum_{i=0}^{n} c_k t^k = 0.$$

Setting $t = 0$ we see that $c_0 = 0$. Differentiating the expression and setting $t = 0$ yields $c_1 = 0$ and so on.

PROBLEM 2.12. In the example above we proved that $u_0, \dots, u_n$ are independent for each n. But by definition of independence we must show that ANY finite collection of u_js is independent. Extend the argument above to cover the general case.

THEOREM 2.13. *Let $S = \{x_1, \dots, x_k\}$ be an independent set consisting of k elements in a linear space V and let $L(S)$ be the space spanned by S. Then every set of $k+1$ elements in $L(S)$ is dependent.*

Before proving this theorem, let us make sure that we understand what it is we are trying to establish. Let $V = \mathbb{R}^2$ and let S be a two-element set consisting of two-vectors that are not multiples of each others. Then by Problem 2.10 this set is independent. The space $L(S)$ is just $\mathbb{R}^2$ (check!) and any three element set in $\mathbb{R}^2$ is dependent, once again by Problem 2.10.

Let us now generalize this reasoning. Suppose that $k = 1$. Then S consists of one element, x_1 and $L(S) = \{cx_1 : c \in \mathbb{R}\}$. Choose any two distinct elements in this set, say, $y_1 = c_1x_1$ and $y_2 = c_2x_1$. It follows that

$$c_2y_1 - c_1y_2 = 0, \text{ so } y_1 \text{ and } y_2 \text{ are dependent.}$$

Now assume that theorem is true for $k-1$ and prove it for k. Let $S = \{x_1, \dots, x_k\}$ and choose a $k+1$ element subset of $L(S)$, say, $T = \{y_1, \dots, y_{k+1}\}$. We must prove that T is dependent. By definition,

$$y_i = \sum_{j=1}^{k} a_{ij}x_j. \tag{2.1}$$

Let us first consider the case when $a_{i1} = 0$ for every $i = 1, \dots, k$. Then the sum in (2.1) does not involve x_1, so each y_i in T is a linear combination of elements in $S' = \{x_2, \dots, x_k\}$. The set S' is independent and consists of $k-1$ elements. By the induction hypothesis, the set T is dependent.

Let us now consider the case where not all scalars a_{i1} are zero. Without loss of generality, suppose that $a_{11} \neq 0$. Take $i = 1$ in (2.1) and multiply both sides by $c_i = \frac{a_{i1}}{a_{11}}$. We obtain

$$c_iy_1 = a_{i1}x_1 + \sum_{j=2}^{k} c_ia_{ij}x_j.$$

Subtract the original equation and obtain

$$c_iy_1 - y_i = \sum_{j=2}^{k} (c_ia_{1j} - a_{ij})x_j$$

for $i = 2, \dots, k+1$.

This expresses k elements $c_iy_1 - y_i$ as a linear combination of the $k-1$ elements $x_2, \dots, x_k$. By the induction hypothesis the collection $c_iy_1 - y_i$ must be dependent. Therefore there exist $t_2, \dots, t_{k+1}$, not all zero, such that

$$\sum_{i=2}^{k+1} t_i(c_iy_1 - y_i) = O.$$

This implies that

$$y_1 \cdot \sum_{i=2}^{k+1} t_ic_i - \sum_{i=2}^{k+1} t_iy_i = O,$$

which is a non-trivial linear combination of elements T yielding O. This means that T is a dependent set and the proof is complete.

3. Bases and dimension

DEFINITION 3.1. A finite set of elements of a linear subspace V is called a finite basis for V if S is independent and spans V . The space V is called finite dimensional if it has a finite basis, or if V consists of O alone. Otherwise V is called infinite dimensional.

THEOREM 3.2. *Let V be a finite dimensional linear space. Then every finite basis for V has the same number of elements.*

The proof follows almost instantly from Theorem 1.5. Suppose that S and T are finite bases for V with k and m elements, respectively. Theorem 1.5 tells us that $k+1$ elements of V are dependent, therefore $m \leq k$. Running this argument with respect to T finishes the proof.

DEFINITION 3.3. If a linear space has a basis with n elements, then we say that $dimV = n$. If $V = \{O\}$, then we say that V has dimension 0.

EXAMPLE 3.4. The space of all polynomials $p(t)$ is infinite dimensional, though you should be prepared to produce a rigorous proof if asked. The idea is that $\{1, t, t^2, \dots, t^n, \dots\}$ spans the space, but how do you proof that no finite subset of this collection does the job?

Here is some more food for thought. Consider the space of all infinitely differentiable functions on $[0, 1]$. It must be infinite dimensional because it contains the space of all the polynomials. Now consider the space of all twice differentiable functions on $[0, 1]$. Is it infinite dimensional? Do we have the means to prove this fact?

THEOREM 3.5. *Let V be a finite dimensional linear space with $dim(V) = n$. Then we have the following:*

a) Any set of independent elements of V is a subset of some basis for V.
b) Any set of n independent elements is a basis for V.

Let $S = \{x_1, x_2, \dots, x_k\}$ be an independent set with n elements in V. If $L(S) = V$, then S is a basis. If not, then there is an element $y \in V$ such that $y \notin L(S)$. Add y to the collection S and consider $S' = \{x_1, \dots, x_k, y\}$. If this set is were dependent, then

$$\sum_{i=1}^{k} c_i x_i + c_{k+1} y = 0 \text{ with } c_{k+1} \neq 0.$$

But this means that we can solve for y contradicting the assumption that $y \notin L(S)$. Therefore S' must be independent. If $L(S') = V$, then S' is a basis and since $S \subset S'$, we are done. If not, keep adding elements until we obtain an independent set with $n+1$ elements, which is impossible by the assumption that $dimV = n$.

Part b) is automatic because by part a) a set S with n independent elements is a subset of some basis. But every basis has exactly n elements, so S must be a basis.

Another important notion is that of components. Let's consider basis elements $e_1, e_2, \dots, e_n$ taken in a given order. We denote such an ordered basis as an n-tuple $(e_1, e_2, \dots, e_n)$. If $x \in V$ we can x as a linear combination of the e_i:

$$x = \sum_{i=1}^{n} c_i e_i.$$

The coefficients c_is are uniquely determined because if there is another representation in the form

$$x = \sum_{i=1}^{n} d_i e_i,$$

then by subtracting we get

$$\sum_{i=1}^{n} (c_i - d_i) e_i = O$$

and we get $c_i \equiv d_i$ by the independence assumption.

PROBLEM 3.6. Do problems 1-10 on page 13. (This is just one problem, really)

PROBLEM 3.7. Do problems 11-20 on page 13.

PROBLEM 3.8. Do problems 23 and 24 on page 14.

4. Inner products and norms

DEFINITION 4.1. We say that $(\cdot, \cdot)$ is an inner product on a linear space V, if given $x, y \in V$ there exists a unique (x, y) satisfying the following axioms:
(1) $(x, y) = (y, x)$ (Symmetry)
(2) $(x, y + z) = (x, y) + (x, z)$ (Linearity)
(3) $c(x, y) = (cx, y)$ (Homogeneity)
(4) $(x, x) > 0$ if $x \neq O$ (Positivity)

A real linear space with an inner product is a called a real Euclidean space. If the scalars in a linear space come from complex numbers instead of real numbers, we replace (1) by (1'), which says

$$(x, y) = \overline{(y, x)}.$$

We also replace (3) by (3') which says

$$(x, cy) = \overline{(cy, x)} = \overline{c}\overline{(y, x)} = \overline{c}(x, y).$$

A complex space with an inner product is called a complex linear inner product space.

EXAMPLE 4.2. In $\mathbb{R}^n$, let $(x, y) = x \cdot y$, the usual dot product.

EXAMPLE 4.3. Let $C([a, b])$ denote the linear space of all the continuous functions on $[a, b]$. Let

$$(f, g) = \int_a^b f(t) g(t) dt.$$

EXAMPLE 4.4. If we consider the space of all polynomials on $\mathbb{R}$, then the inner product we used for $C([a, b])$ does not quite work because the integral does not in general converge. So we go with

$$(p, q) = \int_0^\infty e^{-t} p(t) q(t) dt.$$

THEOREM 4.5. *In real Euclidean space, every inner product satisfies*

$$|(x, y)|^2 \leq (x, x) \cdot (y, y) \text{ for all } x \text{ and } y \text{ in } V.$$

To prove this, observe that if $z \in V$,

$$0 \leq (z,z) = (ax+by, ax+by) = (ax,ax) + (ax,by) + (by,ax) + (by,by)$$
$$a\overline{a}(x,x) + a\overline{b}(x,y) + b\overline{a}(y,x) + b\overline{b}(y,y).$$

Let $a = (y,y)$ and divide by it. We obtain

$$(y,y)(x,x) + \overline{b}(x,y) + b(y,x) + b\overline{b} \geq 0.$$

Now take $b = -(x,y)$. The last inequality becomes

$$(y,y)(x,x) \geq (x,y)(y,x) = |(x,y)|^2,$$

as desired.

PROBLEM 4.6. Do problems 1,2,8, 9, 10, 14 and 16 on pages 20-22.

5. Orthogonality in Euclidean space

DEFINITION 5.1. We say that vectors x and y in Euclidean space are orthogonal if

$$(x,y) = 0.$$

The next result is one of the most important in linear algebra.

THEOREM 5.2. *In Euclidean space V every set of orthogonal vectors is independent. In particular, every set consisting of n orthogonal vectors in an n-dimensional space V is a basis.*

To prove this, suppose that S is an orthogonal set, $x_i \in S$ and

$$\sum_{i=1}^{k} c_i x_i = O.$$

Take the inner product of this quantity with each x_i. By assumption we obtain a relation

$$c_i(x_i, x_i) = O,$$

which implies that $c_i = 0$. This proves that S is independent. If S is independent and has n elements, then Theorem 1.7 (b) shows that S is a basis.

The next result allows us to compute the components of a vector with respect to an orthogonal basis.

THEOREM 5.3. *Let V be a finite dimensional Euclidean space with dimension n and assume that $S = \{e_1, \ldots, e_n\}$ is an orthogonal basis for V. If an element x is a linear combination of the basis elements:*

$$x = \sum_{i=1}^{n} c_i e_i,$$

then its component relative to the ordered basis $(e_1, \ldots, e_n)$ is given by

$$c_j = \frac{(x, e_j)}{(e_j, e_j)}, \ j = 1, 2, \ldots, n.$$

The proof is the same as the one above. We just take inner products and use orthogonality. The next result, known as the Parseval formula, allows us to express the inner product in terms of an orthonormal basis.

THEOREM 5.4. *Let V be a finite dimensional Euclidean space of dimension n and assume that $\{e_1, \dots, e_n\}$ is an orthonormal basis for V. Then for every pair of elements x, y in V we have*

$$(x, y) = \sum_{i=1}^{n} (x, e_i)\overline{(y, e_i)}.$$

In particular, if $x = y$,

$$||x||^2 = \sum_{i=1}^{n} |(x, e_i)|^2.$$

To prove this, note that the previous theorem allows us to write

$$x = \sum_{i=1}^{n} (x, e_i)e_i.$$

Now just take an inner product of both sides with y.

PROBLEM 5.5. Write a paragraph explaining a cryptic remark following the proof of this theorem in the book. Namely, explain why Theorem 1.12 can be viewed as a generalization of the Pythagorean theorem. Please give this some thought.

6. The Gram-Schmidt process

We now learn how to construct orthogonal bases out of regular ones.

THEOREM 6.1. *Let $x_1, x_2, \dots$ be a finite or infinite sequence of elements in Euclidean space V, and let $L(x_1, \dots, x_k)$ be the subspace spanned by the first k elements of this collection. Then there is a sequence $y_1, y_2, \dots$ in V which for each k has the following properties:*

a) The element y_k is orthogonal to every element in the subspace $L(y_1, \dots, y_{k-1})$.

b) The subspace $L(y_1, \dots, y_k)$ is the same as the one spanned by $x_1, \dots, x_k$.

c) The sequence $y_1, \dots,$ is unique up to scalar factors.

To prove this we construct y_js by induction. Let $y_1 = x_1$. Now assume that we have constructed $y_1, \dots, y_r$ so that a) and b) are satisfied when $k = r$. Then we define y_{r+1} by the equation

$$y_{r+1} = x_{r+1} - \sum_{i=1}^{r} a_i y_i,$$

where $a_1 \dots, a_r$ is to be determined.

For $j \leq r$, the inner product of y_{r+1} with y_i is given by

$$(y_{r+1}, y_j) = (x_{r+1}, y_j) - \sum_{i=1}^{r} a_i (y_i, y_j)$$

$$= (x_{r+1}, y_j) - a_j (y_j, y_j)$$

by orthogonality.

If $y_j \neq O$, then we can make y_{r+1} orthogonal go y_j by setting

$$a_j = \frac{(x_{r+1}, y_j)}{(y_j, y_j)}.$$

If $y_j = O$, then y_{r+1} is already orthogonal to y_j and we choose $a_j = 0$. In this way, y_{r+1} is well-defined and is orthogonal to all the previous elements. Therefore, it is orthogonal to every element in the subspace $L(y_1, \dots, y_r)$. This proves a) when $k = r + 1$.

To prove b) when $k = r + 1$, we must show that $L(x_1, \dots, x_{r+1}) = L(y_1, \dots, y_{r+1})$. Let us first show that $L(y_1, \dots, y_{r+1}) \subset L(x_1, \dots, x_{r+1})$. The first r elements are in $L(x_1, \dots, x_{r+1})$ by the induction hypothesis. The remanning element y_{r+1} is a linear combination of x_{r+1} and $y_1, \dots, y_r$, so we are done. We must now show that $L(x_1, \dots, x_{r+1}) \subset L(y_1, \dots, y_{r+1})$. But this argument is the same since x_{r+1} is a linear combination of $y_1, \dots y_{r+1}$.

We prove c) by induction on k. The case $k = 1$ is trivial. So assume true for $k = r$ and consider y'_{r+1}. This element belongs to $L(y_1, \dots, y_{r+1})$ by part b), so

$$y'_{r+1} = \sum_{i=1}^{r+1} c_i y_i = z_r + c_{r+1} y_{r+1},$$

where $z_r \in L(y_1, \dots, y_r)$.

By part a), both $c_{r+1}y_{r+1}$ and y'_{r+1} are orthogonal to z_r. Therefore their difference is also orthogonal to z_r. This means that z_r is orthogonal to itself, which means that $z_r = O$. This completes the proof.

The proof above is quite far-reaching. In particular, it gives us a method (called the Gramm-Schmidt process) for constructing a non-trivial orthogonal set out of an independent set. The formula is

$$y_1 = x_1, \ y_{r+1} = x_{r+1} - \sum_{i=1}^{r} \frac{(x_{r+1}, y_i)}{(y_i, y_i)} y_i \text{ for } r = 1, 2, \dots, k-1.$$

Formulas always remain abstractions until you get your hands dirty.

Example 6.2. Let $V = \mathbb{R}^2$ and consider vectors $(1, 0)$ and $(1, 1)$. They form an independent set, but this set is not orthogonal. Let's apply the GS process. Then

$$y_1 = (1, 0) \text{ and } y_2 = (1, 1) - \frac{((1, 1), (1, 0))}{((1, 0), (1, 0))}(1, 0) = (0, 1).$$

It is hard to argue with the fact that $(1, 0)$ and $(0, 1)$ for an orthogonal set.

We have the following immediate yet important corollary of the previous result.

Theorem 6.3. *Every finite dimensional Euclidean space has an orthonormal basis.*

Problem 6.4. (due Monday, September 28) Do problems 2,3,4 on page 30.

7. Orthogonal complements and projections

Let V be a Euclidean space and S a finite-dimensional subspace. The question we ask is, given $x \in V$, find $y \in S$, such that $||x - y||$ is as small as possible.

Definition 7.1. Let S be a subset of the Euclidean space V. An element in V is said to be orthogonal to S if it is orthogonal to every element of S. The set of all elements orthogonal to S is denoted by $S^{\perp}$. If S is a subspace, then $S^{\perp}$ is called the orthogonal complement of S.

The following result allows us to express any vector in a linear space in terms of a subspace and its orthogonal complement.

THEOREM 7.2. *Let V be a Euclidean space and let S be a finite-dimensional subspace of V. Then every element $x \in V$ can be represented uniquely as a sum of two elements, one in S and one in $S^\perp$. More precisely,*

$$x = s + s^\perp,\ s \in S,\ s^\perp \in S^\perp.$$

Moreover,

$$||x||^2 = ||s||^2 + ||s^\perp||^2.$$

To prove this, we must first show that the decomposition actually exists. Let $\{e_1, \ldots, e_n\}$ be an orthonormal decomposition of S. Given x, define

$$s = \sum_{i=1}^{n} (x, e_i)e_i,\ s^\perp = x - s.$$

By construction, $s \in S$. We must now show that $s^\perp \in S^\perp$. We have

$$(s^\perp, e_j) = (x - s, e_j) = (x, e_j) - (s, e_j).$$

However,

$$(s, e_j) = \sum_{i=1}^{n} (x, e_i)(e_i, e_j) = (x, e_j)$$

and we are done, except for uniqueness. Suppose that

$$x = s + s^\perp \text{ and } x = t + t^\perp.$$

We wish to prove that $s = t$ and $s^\perp = t^\perp$. We have

$$s - t = t^\perp - s^\perp,$$

so it is enough to show that $s - t = O$. But $s - t \in S$ and $t^\perp - s^\perp \in S^\perp$, so $s - t$ is both equal to and orthogonal to $t^\perp - s^\perp$, which implies that $s = t$.

To prove the Pythagorean formula, we observe that

$$||x||^2 = (x, x) = (s + s^\perp, s + s^\perp) = (s, s) + (s^\perp, s^\perp) = ||s||^2 + ||s^\perp||^2,$$

as claimed

PROBLEM 7.3. Do problems 4-10 on page 30.

We now define the notion of a projection.

DEFINITION 7.4. Let S be a finite dimensional subspace of a Euclidean space V, and let $e_1, \ldots, e_n$ be an orthonormal basis for S. If $x \in V$, the element s defined by

$$s = \sum_{i=1}^{n} (x, e_i)e_i$$

is called a projection of x onto S.

EXAMPLE 7.5. Let $V = R^2$ and $S = \{(t, 0) : t \in \mathbb{R}\}$. Let $e_1 = (1, 0)$ be an orthonormal basis for $\mathbb{R}^2$. The notion of orthonormal is vacuous here since S is one-dimensional. Let $x = (x_1, x_2)$. Then the projection of x onto S is

$$s = x \cdot (1, 0)(1, 0) = (x_1, 0).$$

PROBLEM 7.6. Let $V = \mathbb{R}^3$ and $S = \{(u, v, u - v) : u \in \mathbb{R}, v \in \mathbb{R}\}$. Construct an orthonormal basis for S and given an element $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, compute its projection onto S.

8. Best approximation

THEOREM 8.1. *Let S be a finite dimensional subspace of a Euclidean space V, and let x be any element of V. Then the projection of x onto S is nearer to x than any other element of S. In other words, if s is the projection of x onto S and t is any other element of S, then*

$$||x - s|| \leq ||x - t||,$$

and the equality holds only if $s = t$.

To prove this, use the previous result to write $x = s + s^{\perp}$, where $s \in S$ and $s^{\perp} \in S^{\perp}$. Then for any $t \in S$, we have

$$x - t = (x - s) + (s - t).$$

Since $s - t \in S$ and $x - s = s^{\perp} \in S^{\perp}$, this is an orthogonal decomposition of $x - t$. Therefore,

$$||x - t||^2 = ||x - s||^2 + ||s - t||^2.$$

Since $||s - t||^2 \geq 0$, we are done and the equality only holds if $s = t$. We are done.

9. Linear transformations and matrices

DEFINITION 9.1. If V and W are linear spaces, a function $T : V \to W$ is called a linear transformation of V onto W if it has the following two properties: a) $T(x + y) = T(x) + T(y)$ for all $x, y \in V$.

b) $T(cx) = cT(x)$ for all $x \in V$ and all scalars c.

These properties combine by induction into the statement

$$T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$$

for any collection $x_i \in V$ and scalars a_i.

EXAMPLE 9.2. The identity transformation and the zero transformation are linear transformations, as is multiplication by a scalar. Among the non-trivial linear transformations, it is especially important to note the case $V = V_n$ and the transformation that takes $(x_1, \ldots, x_n)$ to $(y_1, \ldots, y_n)$ via

$$y_i = \sum_{k=1}^{n} a_{ik} x_k \text{ for } i = 1, 2, \ldots, m.$$

EXAMPLE 9.3. Another incredibly important example is when V is the space of real valued continuous function on $[a, b]$. For $f \in V$, define

$$g(x) = \int_a^b f(t)dt \text{ if } a \leq x \leq b.$$

Such a transformation T is called an integral operator.

10. Null space and range

THEOREM 10.1. *Suppose that $T : V \to W$, a linear transformation. Then $T(V)$ is a subspace of W. Moreover the zero element of V gets mapped to the zero element of W.*

The proof follows instantly from the definition.

DEFINITION 10.2. Let $T : V \to W$. Then

$$N(T) = \{x \in V : T(x) = O\}.$$

THEOREM 10.3. *The null space $N(T)$ is a subspace of V.*

The proof is, once again, immediate.

EXAMPLE 10.4. Beyond the identity and the zero transformation, a very important example is that of a dot product, i.e

$$T(x) = x \cdot y,$$

where y is a fixed element of V_n. In this case the null space is the space of vectors perpendicular to y.

EXAMPLE 10.5. Another critical example is the projection onto a subspace S. If $x \in V$, $x = s + s^{\perp}$ with $s \in S$ and $s^{\perp} \in S^{\perp}$. Since $T(x) = s$, we have $T(x) = O$ if and only if $x = s^{\perp}$. It follows that $N(T) = S^{\perp}$.

11. Rank/Nullity Theorem

There is no mathematics without the following result.

THEOREM 11.1. *If V is finite dimensional, then $T(V)$ is also finite dimensional and*

$$dim(N(T)) + dim(T(V)) = dim(V).$$

To prove this result, let $n = dim(V)$ and let $e_1, \ldots, e_k$ be a basis for $N(T)$, where $k = dim(N(T)) \leq n$. We know that this set is a part of a basis of V, say,

$$e_1, \ldots, e_k, e_{k+1}, \ldots, e_{k+r},$$

where $k + r = n$.

We need to establish the fact that $T(e_{k+1}), \ldots, T(e_{k+r})$ form a basis for $T(V)$, which instantly implies that $dim(V) = r$.

We begin by showing that these vectors span $T(V)$. If $y \in T(V)$, then $y = T(x)$ for some $x \in V$. Express x in the form

$$c_1 e_1 + \cdots + c_{k+r} e_{k+r}.$$

Hence

$$Y = T(x) = \sum_{i=1}^{k+r} c_i T(e_i) = \sum_{i=1}^{k} c_i T(e_i) + \sum_{i=k+1}^{k+r} c_i T(e_i) = \sum_{i=k+1}^{k+r} c_i T(e_i)$$

since $T(e_j) = O$ for $1 \leq j \leq k$, which implies that $T(e_{k+1}), \ldots, T(e_{k+r})$ span $T(V)$.

We must now establish the independence of these vectors. Suppose that there exist scalars such that

$$\sum_{i=k+1}^{k+r} c_i T(e_i) = O.$$

This implies that

$$T(\sum_{i=k+1}^{k+r} c_i e_i) = O,$$

which implies that

$$x = c_{k+1}e_{k+1} + \cdots + c_{k+r}e_{k+r} \in N(T).$$

This means that there exist $c_1, \ldots, c_k$ such that

$$x - x = \sum_{i=1}^{k} c_i e_i - \sum_{i=k+1}^{k+r} c_i e_i = O,$$

but this implies that all the c_js are zero since $e_1, \ldots, e_k, \ldots, e_{k+r}$ is a basis for V. This completes the proof.

PROBLEM 11.2. Do the problems 1,3,5,7,9,12,14,16,18,21,23 on page 35. Show your work.

PROBLEM 11.3. Do problem 29 on page 36.

12. Basic manipulations with linear transformations

DEFINITION 12.1. Let $S : V \to W$ and $T : V \to W$ be two function with a common domain V and values in a linear space W. If c is any scalar in W, let

$$(S + T)(x) = S(x) + T(x), \ (cT)(x) = cT(x)$$

for all $x \in V$.

We shall mostly use this concept in the context when V is also a linear space with the same scalars as W. In this case let $\mathcal{L}(V, W)$ denote the set of all linear transformations from V to W. We have the following basic result.

DEFINITION 12.2. The set $\mathcal{L}(V, W)$ is a linear space with operations given by the definition above.

The proof is straightforward. Beyond addition and multiplication, we also have a composition operation.

DEFINITION 12.3. Let U, V, W be sets. Let $T : U \to V$ be a function with domain U and values in V and let $S : V \to W$ be defined similarly with respect to V and W. Then ST is the function from U to W defined by

$$(ST)(x) = S(T(x)) \text{ for every } x \in U.$$

It is not in general true that $ST = TS$ (example?), but associativity does hold.

THEOREM 12.4. *If $T : U \to V, S : V \to W, R : W \to X$ are functions, then*

$$R(ST) = (RS)(T).$$

The proof is immediate from the definition.

It is useful to define powers of a function that maps $V \to V$ by the relations

$$T^0 = I, T^n = TT^{n-1}.$$

THEOREM 12.5. *Let U, V, W be linear spaces with the same scalars, and $T : U \to V, S : V \to W$ are linear transformations, then $ST : U \to W$ is a linear transformation.*

The proof once again follows directly from the definition as does the following observation.

THEOREM 12.6. *Let U, V, W be linear spaces with the same scalars, assume S and T are in $\mathcal{L}(V, W)$, and let c be any scalar.*

a) For any function R with values in V, we have

$$(S + T)(R) = SR + TR \text{ and } (cS)R = c(SR).$$

b) For any linear transformation $R : W \to U$, we have

$$R(S + T) = RS + RT \text{ and } R(cS) = c(RS).$$

We are now rapidly driving towards cool objects of study. Given a function T, we wish to find a function S such that ST and TS are both the identity function. This is not always possible and indeed the composition is not typically commutative. This forces us to define left and right inverses, at least in the beginning.

DEFINITION 12.7. Given two sets V, W and a function $T : V \to W$, a function $S : T(V) \to V$ is called a left inverse of T if $S[T(x)] = x$ for all $x \in V$, that is, if

$$ST = I_V,$$

the identity transformation on V. The right inverse is defined analogously, i.e.

$$TR = I_{T(V)}.$$

THEOREM 12.8 (LEFT INVERSE). *A function $T : V \to W$ can have at most one left inverse. If T has a left inverse S, then S is also a right inverse.*

To prove this, suppose that T has two left inverses, S and S'. Choose any $y \in T(V)$. Then $y = T(x)$ for some $x \in V$. We must show that $S(y) = S'(y)$. We have

$$S(T(x)) = x \text{ and } S'(T(x)) = x,$$

since both S and S' are left inverses. This means that $S(y) = S'(y)$, so $S = S'$ since y is arbitrary.

Let's now prove that if the left inverse exists, then the right inverse exists also. We have

$$x = S(T(x)) = S(y).$$

Applying T we get

$$T(x) = T(S(y)).$$

But since $y = T(x)$,

$$y = T(S(y))$$

and the proof is complete.

A companion result is the following.

THEOREM 12.9 (EXISTENCE OF LEFT INVERSE). *A function $T : V \to W$ has a left inverse if and only if T maps distinct elements of V onto distinct elements of W (i.e. T is one-to-one) In other words for all $x, y \in V$,*

$$x \neq y \text{ implies } T(x) \neq T(y).$$

First assume that T has a left inverse. Suppose that $T(x) = T(y)$. We must show that $x = y$. Applying the left inverse S we see that $S(T(x)) = S(T(y))$, so $x = y$, as desired.

Now assume that T is one-to-one. If $y \in T(V)$, then $y = T(x)$ for some $x \in V$ and by one-to-one property there is exactly one such x. Define $S(y) = x$. This gives us the left inverse and we are done.

DEFINITION 12.10. Let $T : V \to W$ be one-to-one on V. The unique left inverse of T, which we know is also the right inverse of T is called the inverse of T and denoted by T^{-1}.

13. One-to-one linear transformations

We are now firmly in the realm of linear spaces for the foreseeable future.

THEOREM 13.1. *Let $T : V \to W$ be a linear transformation in $\mathcal{L}(V, W)$. Then the following statements are equivalent:*

a) T is 1-1 on V.

b) T is invertible and its inverse $T^{-1} : T(V) \to V$ is linear.

c) For all $x \in V$, $T(x) = 0$ implies $x = O$.

We shall prove that a) implies b), b) implies c) and c) implies a).

Suppose that a) holds. We know that T is invertible with the inverse T^{-1}. We must prove that T^{-1} is linear. Consider $y, y' \in T(V)$. Then $y = T(x)$, $y' = T(x')$ or some $x, x' \in V$. Then

$$T^{-1}(ay + by') = T^{-1}(aT(x) + bT(x')) = T^{-1}(T(ax) + T(bx'))$$
$$= T^{-1}(T(ax + bx')) = ax + bx' = aT^{-1}(y) + bT^{-1}(y')$$

as desired. This proves that a) implies b).

Let's assume b) and prove c). Suppose that c) does not hold. Then there exists $x \neq O$ such that $Tx = 0$. Then T is not 1-1 and b) is violated. Alternatively, we can argue direction that if $T(x) = O$, then $x = T^{-1}(T(x)) = T^{-1}(O) = O$ since T^{-1} is linear.

Let's assume c) and prove a). Suppose that $T(u) = T(v)$. Then $T(u - v) = O$ by linearity. This means that $u - v \in N(T)$ which implies, by assumption, that $u - v = O$, hence $u = v$.

We need another set of equivalences in terms of the basis elements of V.

THEOREM 13.2. *Let $T : V \to W$ be a linear transformation in $\mathcal{L}(V, W)$ and assume that $dim(V) = n < \infty$. Then the following are equivalent.*

a) T is 1-1 on V.

b) If $e_1, \ldots, e_k$ are independent elements in V, then $T(e_1), \ldots, T(e_k)$ are independent elements of $T(V)$.

c) $dim(T(V)) = n$.

d) If $\{e_1, \ldots, e_n\}$ is a basis for V, then $\{T(e_1), \ldots, T(e_n)\}$ is a basis for $T(V)$.

The proof is routine and left to the reader. Please go through it carefully, however...

PROBLEM 13.3. Do problems 1-12 on page 42.

PROBLEM 13.4. Do problems 16-20 on page 43.

PROBLEM 13.5. Do problems 25, 29, and 30 on page 43.

14. Concrete linear transformations

The purpose of the next theorem is to construct a linear transformation that hits assigned values in the ambient space.

THEOREM 14.1. *Let $e_1, \ldots, e_n$ be a basis for an n-dimensional linear space V. Let $u_1, \ldots, u_n$ be arbitrary elements in a linear space W. Then there is a unique linear transformation T such that $T(e_k) = u_k$, $1 \leq k \leq n$. More precisely, if $x = \sum_{k=1}^{n} x_k e_k$, then*

$$T(x) = \sum_{k=1}^{n} x_k u_k.$$

The proof is once again straightforward.

15. Matrices!!!