Section 1: Vector spaces

Matthew Macauley

Department of Mathematical Sciences Clemson University http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Algebraic structures

Definition

A group is a set G and associative binary operation * with:

- closure: $a, b \in G$ implies $a * b \in G$;
- **identity**: there exists $e \in G$ such that a * e = e * a = a for all $a \in G$;
- inverses: for all $a \in G$, there is $b \in G$ such that a * b = e.

A group is abelian if a * b = b * a for all $a, b \in G$.

Definition

A field is a set \mathbb{F} (or K) containing $1 \neq 0$ with two binary operations: + (addition) and \cdot (multiplication) such that:

(i) F is an abelian group under addition;

(ii) $\mathbb{F} \setminus \{0\}$ is an abelian group under multiplication;

(iii) The distributive law holds: a(b+c) = ab + ac for all $a, b, c \in \mathbb{F}$.

Remarks

- \mathbb{Q} , \mathbb{R} , \mathbb{C} , \mathbb{Z}_p (prime p), $\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ are all fields.
- **\blacksquare** \mathbb{Z} is not a field. Nor is \mathbb{Z}_n (composite *n*).
- the additive identity is 0, and the inverse of a is -a.
- the multiplicative identity is 1, and the inverse of a is a^{-1} , or $\frac{1}{a}$.

Vector spaces

Definition

A vector space is a set X ("vectors") over a field \mathbb{F} ("scalars") such that:

(i) X is an abelian group under addition;

(ii) + and \cdot are "compatible" via natural associative and distributive laws relating the two:

• $a(bv) = (ab)v$,	for all $a, b \in \mathbb{F}$, $v \in X$;
a(v+w) = av + aw,	for all $a \in \mathbb{F}$, $v, w \in X$;
(a+b)v = av + bv,	for all $a \in \mathbb{F}$, $v, w \in X$;
• $1v = v$,	for all $v \in X$.

Intuition

Think of a vector space as a set of vectors that is:

- (i) Closed under addition and subtraction;
- (ii) Closed under scalar multiplication;

(iii) Equipped with the "natural" associative and distributive laws.

Proposition (exercise)

In any vector space X,

- (i) The zero vector 0 is unique;
- (ii) 0x = 0 for all $x \in X$;

(iii)
$$(-1)x = -x$$
 for all $x \in X$.

Linear maps

Definition

A linear map between vector spaces X and Y over \mathbb{F} is a function $\varphi \colon X \to Y$ satisfying:

- $\varphi(v+w) = \varphi(v) + \varphi(w),$ for all $v, w \in X;$
- $\varphi(av) = a \varphi(v)$, for all $a \in \mathbb{F}$, $v \in X$.

An isomorphism is a linear map that is bijective (1-1 and onto).

Proposition

The two conditions for linearity above can be replaced by the single condition:

$$\varphi(av + bw) = a\varphi(v) + b\varphi(w),$$
 for all $v, w \in X$ and $a, b \in \mathbb{F}$

Examples of vector spaces

- (i) $K^n = \{(a_1, \ldots, a_n) : a_i \in K\}$. Addition and multiplication are defined componentwise.
- (ii) Set of functions $\mathbb{R} \longrightarrow \mathbb{R}$ (with $K = \mathbb{R}$).
- (iii) Set of functions $S \longrightarrow K$ for an abitrary set S.
- (iv) Set of polynomials of degree < n, with coefficients from K.

Exercise

In the list of vector spaces above, (i) is isomorphic to (iv), and to (iii) if |S| = n.

Subspaces

Definition

A subset Y of a vector space X is a subspace if it too is a vector space. We'll write $Y \leq X$.

Examples

- (i) $Y = \{(0, a_2, \dots, a_{n-1}, 0) : a_i \in K\} \subseteq K^n$.
- (ii) $Y = \{ \text{functions with period } T | \pi \} \subseteq \{ \text{functions } \mathbb{R} \to \mathbb{R} \}.$
- (iii) $Y = \{ \text{constant functions } S \to K \} \subseteq \{ \text{functions } S \to K \}.$

(iv) $Y = \{a_0 + a_2x^2 + a_4x^4 + \dots + a_{n-1}x^{n-1} : a_i \in K\} \subseteq \{\text{polynomials of degree} < n\}.$

Definition

If Y and Z are subsets of a vector space X, then their:

sum is
$$Y + Z = \{y + z \mid y \in Y, z \in Z\};$$

• intersection is $Y \cap Z = \{x \mid x \in Y, x \in Z\}.$

Exercise

If Y and Z are subspaces of X, then Y + Z and $Y \cap Z$ are also subspaces.

Spanning and Independence

Definition

A linear combination of vectors x_1, \ldots, x_k is a vector of the form $a_1x_1 + \cdots + a_kx_k$, where each $a_i \in K$.

Definition

Given a subset $S \subseteq X$, the subspace spanned by S is the set of all linear combinations of vectors in S, and denoted Span(S).

Exercise

For any subset $S \subseteq X$,

$$\operatorname{Span}(S) = igcap_{S \subseteq Y_{lpha} \leq X} Y_{lpha},$$

where the intersection is taken over all subspaces of X that contain S.

Definition

The vectors x_1, \ldots, x_k are linearly dependent if we can write $a_1x_1 + \cdots + a_kx_k = 0$, where not all $a_i = 0$. Otherwise, the vectors are linearly independent.

Spanning and linear independence

Lemma 1.1

If $X = \text{Span}(x_1, \dots, x_n)$, and the vectors $y_1, \dots, y_k \in X$ are linearly independent, then $k \leq n$.

Proof outline (details to be done on the board)

Write $y_1 = a_1 x_1 + \cdots + a_n x_n$, and assume WLOG that $a_1 \neq 0$.

Now, "solve" for x_1 and eliminate it, and conclude that

 $\operatorname{Span}(x_1, x_2, \ldots, x_n) = \operatorname{Span}(y_1, x_2, \ldots, x_n) = X$

Repeat this process: eliminating each x_2, x_3, \ldots

Note that k > n is impossible. (Why?)

Basis of a vector space

Definition

- A set $B \subseteq X$ is a **basis** for X if:
 - *B* spans *X*. (is "big enough");
 - *B* is linearly independent. (isn't "too big").

Exercise

The following are equivalent for a subset $B \subseteq X$:

- (i) B is a basis of X;
- (ii) B is a minimal spanning set;
- (iii) B is a maximal linearly independent set.

Examples

Let's find bases for some familiar vector spaces.

- 1. $K^n = \{(a_1, \ldots, a_n) : a_i \in K\}$. Addition and multiplication are defined componentwise.
- 2. Set of functions $S \longrightarrow K$ from a finite set S.
- 3. Set of polynomials of degree < n, with coefficients from K.

Bases

Lemma 1.2

If Span $(x_1, \ldots, x_n) = X$, then some subset of $\{x_1, \ldots, x_n\}$ is a basis for X.

Proof

If x_1, \ldots, x_n are linearly dependent, then we can write (WLOG; renumber of necessary)

$$x_n=a_1x_1+\cdots+a_{n-1}x_{n-1}.$$

Now, $\text{Span}(x_1, \ldots, x_{n-1}) = X$, and we can repeat this process until the remaining set is linearly independent.

Definition

A vector space X is finite dimensional (f.d.) if it has a finite basis.

Examples

- (i) In \mathbb{R}^n , any two vectors that don't lie on the same line (i.e., aren't scalar multiples) are linearly independent.
- (ii) In \mathbb{R}^3 , any three vectors are linearly independent iff they do not lie on the same plane.
- (iii) Any two vectors in \mathbb{R}^2 that aren't scalar multiples form a basis.

Dimension

Theorem / Definition 1.3

All bases for a f.d. vector space have the same cardinality, called the dimension of X.

Proof

Let x_1, \ldots, x_n and y_1, \ldots, y_m be two bases for X. By Lemma 1.1, $m \le n$ and $n \le m$.

Theorem 1.4

Every linear independent set of vectors y_1, \ldots, y_j in a finite-dimensional vector space X can be extended to a basis of X.

Proof

If $\text{Span}(y_1, \ldots, y_j) \neq X$, then find $y_{j+1} \in X$ not in $\text{Span}(y_1, \ldots, y_j)$, add it to the set and repeat the process.

This will terminate in less than $n = \dim X$ steps because otherwise, X would contain more than n linearly independent vectors.

An example from ODEs

Let X be the set of all smooth functions x(t) that satisfy the second order differential equation $\frac{d^2}{dt^2}x + x = 0$.

If $x_1(t)$, $x_2(t)$ are solutions, then so are $x_1(t) + x_2(t)$ and $cx_1(t)$. Thus X is a vector space.

Solutions describe the motion of a mass-spring system (simple harmonic motion). A particular solution is determined completely by specifying:

 $x(0) = x_0$ (initial position) $x'(0) = v_0$ (initial velocity).

Thus, we can describe an element $x(t) \in X$ by a pair (x_0, v_0) , where $x_0, v_0 \in \mathbb{R}$ (or in \mathbb{C}).

This defines an isomorphism $X \longrightarrow \mathbb{C}^2$, by $x(t) \longmapsto (x(0), x'(0))$.

Note that $\cos x$ and $\sin x$ are two linearly independent solutions, so the general solution to this ODE is $a \cos x + b \sin x$; $a, b \in \mathbb{C}$.

Said differently, $\{\cos x, \sin x\}$ is a basis for the solution space of x'' + x = 0.

Note that $\cos x + i \sin x = e^{ix}$ and $\cos x - i \sin x = e^{-ix}$ are linearly independent, and so $\{e^{ix}, e^{-ix}\}$ is another basis! Thus, the general solution can be written as $C_1e^{ix} + C_2e^{-ix}$ instead!

Complements and direct sums

Theorem 1.5

- (a) Every subspace Y of a finite-dimensional vector space X is finite-dimensional.
- (b) Every subspace Y has a complement in X: another subspace Z such that every vector x ∈ X can be written uniquely as

x = y + z, $y \in Y$, $z \in Z$, $\dim X = \dim Y + \dim Z$.

Proof

Pick $y_1 \in Y$ and extend this to a basis y_1, \ldots, y_j of Y. By Lemma 1.1, $j \leq \dim X < \infty$. Extend this to a basis $y_1, \ldots, y_j, z_{j+1}, \ldots, z_n$ of X [and define $Z := \text{Span}(z_{j+1}, \ldots, z_n)$].

Clearly, Y and Z are complements, and dim $X = n = j + (n - j) = \dim Y + \dim Z$.

Definition

X is the direct sum of subspaces Y and Z that are complements of each other.

More generally, X is the direct sum of subspaces Y_1, \ldots, Y_m if every $x \in X$ can be expressed uniquely as

$$x = y_1 + \cdots + y_m, \qquad y_i \in Y_i.$$

We denote this as $X = Y_1 \oplus \cdots \oplus Y_m$.

Direct products

Definition

The direct product of X_1 and X_2 is the vector space

$$X_1 \times X_2 := \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\},\$$

with addition and multiplication defined componentwise.

Proposition

$$\dim(Y_1 \oplus \cdots \oplus Y_m) = \sum_{i=1}^m \dim Y_i; \qquad \qquad \ \ \, \dim(X_1 \times \cdots \times X_m) = \sum_{i=1}^m \dim X_i.$$

Example

Let $X = \mathbb{R}^4$, $Y_1 = \{(a, b, 0, 0) : a, b \in \mathbb{R}\}$, $Y_2 = \{(0, 0, c, d) : c, d \in \mathbb{R}\}$, $X_1 = X_2 = \mathbb{R}^2$.

Clearly,
$$X = Y_1 \oplus Y_2$$
, since $(a, b, c, d) = (a, b, 0, 0) + (0, 0, c, d)$ [uniquely].

$$X_1 \times X_2 = \left\{ \left((a,b), (c,d) \right) : (a,b) \in \mathbb{R}^2, (c,d) \in \mathbb{R}^2 \right\} \cong \left\{ (a,b,c,d) : a,b,c,d \in \mathbb{R} \right\} = X.$$

Direct sums vs. direct products

In the finite-dimensional cases, there is no difference (up to isomorphism) of direct sums vs. direct products.

Not the case when dim $X = \infty$. Consider the vector space:

$$X = \mathbb{R}^{\infty} := \{(a_1, a_2, a_3, \dots) : a_i \in \mathbb{R}\} \cong \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots$$

and the following subspaces:

 $X_1 = \{(a_1, 0, 0, 0, \dots,) : a_1 \in \mathbb{R}\}, \qquad X_2 = \{(0, a_2, 0, 0, \dots,) : a_2 \in \mathbb{R}\}, \qquad \text{and so on}.$

Elements in the subspace $X_1 \oplus X_2 \oplus X_3 \oplus \cdots$ of X are finite sums

$$x = x_{i_1} + x_{i_2} + \dots + x_{i_k}, \quad x_{i_i} \in X_{i_i}.$$

Thus, we can write the direct sum as follows:

$$X_1 \oplus X_2 \oplus X_3 \oplus \cdots = \{(a_1, \ldots, a_k, 0, 0, \ldots) : a_i \in \mathbb{R}, \ k \in \mathbb{Z}\} \subsetneq \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots$$

- Elements in the direct product are sequences, e.g., x = (1, 1, 1, ...).
- Elements in the direct sum are finite sums, e.g., $x = 3e_1 5.25e_4 + 78e_{11}$.

Congruence of subspaces

Sums and products "multiply" vector spaces. We can also "divide" by a subspace.

Definition

If Y is a subspace of X, then two vectors $x_1, x_2 \in X$ are congruent modulo Y, denoted $x_1 \equiv x_2 \pmod{Y}$, if $x_1 - x_2 \in Y$.

Proposition (exercise)

Congruence modulo Y is an equivalence relation, i.e., it is:

- (i) symmetric: $x \equiv y$ imples $y \equiv x$;
- (ii) **reflexive**: $x \equiv x$ for all $x \in X$;

(iii) **transitive**: $x \equiv y$ and $y \equiv z$ implies $x \equiv z$.

The equivalence classes are called congruence classes mod Y, or cosets. Denote the class containing x by $\{x\}$. [Sometimes written \overline{x} or $x + Y := \{x + y : y \in Y\}$.]

Example

Let
$$X = \mathbb{R}^3$$
, $Y = \{(x, y, 0) : x, y \in \mathbb{R}\}$ = xy-plane, $Z = \{(0, 0, z) : z \in \mathbb{R}\}$ = z-axis

• $v \equiv w \mod Y$ if they lie on the same horizontal plane.

• $v \equiv w \mod Z$ if they lie on the same vertical line.

Quotient spaces

Let X/Y denote the set of equivalence classes in X, modulo Y.

This can be made into a vector space by defining addition and scalar multiplication as

$$\{x\} + \{z\} := \{x + z\}, \quad a\{x\} := \{ax\}.$$

Need to check that this is well-defined, i.e., that it is *independent of the choice of representative* from the classes.

This means showing (HW exercise) that if $x_1 \equiv x_2 \mod Y$ and $z_1 \equiv z_2 \mod Y$, then

$$\{x_1\} + \{z_1\} = \{x_2\} + \{z_2\}, \qquad a\{x_1\} = a\{x_2\}.$$

Definition

The vector space X/Y is called the quotient space of X modulo Y.

Alternate notations

Since $\{x\}$ is sometimes written \overline{x} , or $x + Y := \{x + y : y \in Y\}$, then addition and multiplication becomes:

•
$$\overline{x} + \overline{z} = \overline{x + z}$$
, and $a\overline{x} = \overline{ax}$;

$$(x + Y) + (z + Y) = x + z + Y, \text{ and } a(x + Y) = ax + Y.$$

Dimension of quotient spaces

Theorem 1.6

If Y is a subspace of a finite-dimensional vector space X, then dim $Y + \dim X/Y = \dim X$.

Proof

Let y_1, \ldots, y_k be a basis for Y. Extend this to a basis $y_1, \ldots, y_k, x_{k+1}, \ldots, x_n$ of X.

Claim: $\{x_{k+1}\}, \ldots, \{x_n\}$ is a basis of X/Y.

• Show this spans X/Y:

Pick
$$\{x\}$$
 in X/Y and write $x = \sum_{i=1}^{k} a_i y_i + \sum_{j=k+1}^{n} b_j x_j$. By definition,

$$\{x\} = \left\{\sum a_i y_i + \sum b_j x_j\right\} = \sum a_i \{y_i\} + \sum b_j \{x_j\} = \sum b_j \{x_j\}.$$

Show this is linearly independent: Suppose $\sum_{j=k+1}^{n} c_j \{x_j\} = \{0\}$, which means $\sum c_j x_j = y$ for some $y \in Y$. Write $y = \sum_{i=1}^{k} d_i y_i$, and so $\sum c_k x_k - \sum d_i y_i = 0$, and hence all $c_k, d_i = 0$ (Why?).

Corollary

If a subspace Y of a finite-dimensional space X has dim $Y = \dim X$, then Y = X.

M. Macauley (Clemson)

Dimension of sums

Theorem 1.7

Let U, V be subspaces of a finite-dimensional space X with U + V = X. Then

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

Proof

Let $W = U \cap V$. The result trivially holds when $W = \{0\}$ (Theorem 1.5).

Define
$$\overline{U} = U/W$$
, $\overline{V} = V/W$ and $\overline{X} = X/W$.

Note that $\overline{U} \cap \overline{V} = \{0\}$ (why?), and $\overline{X} = \overline{U} + \overline{V}$, so dim $\overline{X} = \dim \overline{U} + \dim \overline{V}$ (Theorem 1.5).

By Theorem 1.6: $\dim \overline{X} = \dim X - \dim W$ $\dim \overline{U} = \dim U - \dim W$ $\dim \overline{V} = \dim V - \dim W$

Therefore, $(\dim X - \dim W) = (\dim U - \dim W) + (\dim V - \dim W)$.

From which it easily follows that dim $X = \dim U + \dim V - \dim W$.

Scalar functions

Let X be a vector space over a field K. A scalar function is any function from X to K.

A scalar function $\ell \colon X \to K$ is linear if

•
$$\ell(x+y) = \ell(x) + \ell(y)$$
, for all $x, y \in X$;

•
$$\ell(cx) = c\ell(x)$$
, for all $x \in X$, $c \in K$.

Or equivalently, if

$$\ell(c_1x_1 + \dots + c_nx_n) = c_1\ell(x_1) + \dots + c_n\ell(x_n), \quad \text{for all } c_i \in K, \ x_i \in X$$

Definition

The set of linear scalar functions $\ell: X \to K$ is a vector space called the dual of X, and denoted X'.

Addition and scalar multiplication is defined naturally:

- Addition: $(\ell + m)(x) := \ell(x) + m(x)$,
- Scalar multiplication: $(c\ell)(x) := c\ell(x)$.

Examples of scalar functions

Example 1

Let $X = C([0,1],\mathbb{R})$, the continuous functions $[0,1] \to \mathbb{R}$, and fix $t_1, \ldots, t_n \in [0,1]$. The following are linear scalar functions:

•
$$\ell(f) = f(t_1);$$

• $\ell(f) = \sum_{i=1}^{n} a_i f(t_i), \quad a_i \in \mathbb{R};$
• $\ell(f) = \int_0^1 f(t) dt.$

Example 2

Let $X = \mathcal{C}^{\infty}(\mathbb{R})$ be the set of smooth functions $\mathbb{R} \to \mathbb{R}$. For a fixed $t_0 \in \mathbb{R}$,

$$\ell := \sum_{i=1}^{n} \mathsf{a}_i \left. \frac{d^i}{dt^i} \right|_{t=t_0}, \qquad \ell \colon f \longmapsto \sum_{i=1}^{n} \mathsf{a}_i \left. \frac{d^i f}{dt^i} \right|_{t=t_0}$$

is a linear scalar function (i.e., an element of X').

The dual space

If dim X = n, then $X \cong K^n$. Thus, we can associate a vector $x \in X$ with an *n*-tuple $x = (c_1, \ldots, c_n)$ of scalars.

For any fixed $a_1, \ldots, a_n \in K$, the function

$$\ell: X \longrightarrow K, \qquad \ell(x) = a_1 c_1 + \dots + a_n c_n$$
 (1)

is linear, i.e., $\ell \in X'$.

Theorem 1.8

If dim $X = n < \infty$, then every $\ell \in X'$ can be written as in Eq. (1).

Proof

The dual space

Corollary 1.9

If dim $X < \infty$, then $X \cong X'$.

One way to think of this is to:

- 1. associate a vector $x \in X$ with a column vector,
- 2. associate a scalar function $\ell \in X'$ with a row vector.

Notation

A linear function $\ell \in X'$ applied to a vector $x \in X$ depends on the *n*-tuples (c_1, \ldots, c_n) for x and (a_1, \ldots, a_n) for ℓ . We can use scalar product notation

$$(\ell, x) := \ell(x).$$

Sometimes, elements $\ell \in X'$ are called co-vectors, or dual vectors.

Definition

Let x_1, \ldots, x_n be a basis for X. The dual basis in X' is ℓ_1, \ldots, ℓ_n , where

$$(\ell_i, x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Think of ℓ_i as the function that "picks off" the coefficient of x_i .

M. Macauley (Clemson)

Duality in infinite dimensional spaces

Consider the vector space

$$X = \ell^1(\mathbb{R}) := \Big\{ (x_1, x_2, \dots) \mid x_i \in \mathbb{R}, \sum_{i=1}^{\infty} |x_i| < \infty \Big\}.$$

Given vectors $y = (a_1, a_2, \dots)$ and $x = (c_1, c_2, \dots)$,

$$(y,x)=\sum_{i=1}^{\infty}a_ic_i<\infty,$$

so every $y \in X$ defines a co-vector in X'.

But there are others! If z = (1, 1, 1, ...),

$$(z,x)=\sum_{i=1}^{\infty}c_i<\infty,$$

but $z \notin X$.

The double dual

The scalar product (ℓ, x) is a bilinear function of ℓ and x. That is, if we fix one argument, it is linear in the other. Equivalently,

$$\underbrace{(a\ell, x)}_{=a\ell(x)} = a(\ell, x) = \underbrace{(\ell, ax)}_{\ell(ax)}$$
 for all $x \in X, \ \ell \in X', \ a \in K$

If dim $X = n < \infty$, then every linear scalar function $X \to K$ is of the form

$$(\ell, x)$$
, for some fixed $\ell = (a_1, \ldots, a_n) \in K^n$.

Since X' is a vector space, it has a dual, called the double dual of X, and denoted X'' := (X')'. Every linear scalar function $X' \to K$ is of the form

$$(\ell, x)$$
, for some fixed $x = (c_1, \ldots, c_n) \in K^n$.

Key points

Let x_1, \ldots, x_n be a basis of X

- Think of the dual basis ℓ_1, \ldots, ℓ_n as "pick-off functions"
- Think of elements in the double dual as "evaluation functions"

The bilinear function (ℓ, x) naturally identifies X'' with X.

Annihilators

Definition

Let $Y \leq X$. The set of linear functions that vanish on Y is its annihilator, denoted

$$Y^{\perp} = \big\{ \ell \in X' \mid \ell(y) = 0, \ \forall y \in Y \big\}.$$

Theorem 1.10

Let $Y \leq X$ with dim $X < \infty$. Then

 $\dim Y + \dim Y^{\perp} = \dim X.$

Proof

The annihilator of the annihilator

Definition

The dimension of Y^{\perp} is called the codimension of Y in X, denoted codim Y.

By Theorem 1.10,

```
\dim Y + \operatorname{codim} Y = \dim X.
```

Since Y^{\perp} is a subspace of X', its annihilator $Y^{\perp \perp}$ is a subspace of X''.

Theorem 1.11

Assume dim $X < \infty$ and identify X'' with X. Then $Y^{\perp \perp} = Y$.

Proof

The annihilator of a subset

We can define the annihilator of an arbitrary subset $S \subseteq X$, as

$$S^{\perp} := \big\{ \ell \in X' \mid \ell(s) = 0, \ \forall s \in S \big\}.$$

Recall that the smallest subspace containing S is

$$\operatorname{Span}(S) = igcap_{S \subseteq Y_{lpha} \leq X} Y_{lpha}.$$

Exercises

Let
$$S, T \subseteq X$$
.
If $S \subseteq T$, then $T^{\perp} \subseteq S^{\perp}$,
 $S^{\perp} = \text{Span}(S)^{\perp}$.