Linear algebra fundamentals

Matthew Macauley

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

Math 8530, Fall 2020

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) $\mathbb{F}$ 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.
- $\mathbb{Z}$ 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 $\mathbf{0}$ is unique;

(ii) $0x = \mathbf{0}$ for all $x \in X$;

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

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:

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

Examples of vector spaces

(i) $K^n = \{(a_1, \dots, a_n) : a_i \in K\}$. 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 $|S| = n$. $\square$

Subspaces

Definition

A subset Y of a vector space X is a subspace if it too is a vector space.

Examples

(i) $Y = \{(0, a_2, \dots, a_{n-1}, 0) : a_i \in K\} \subseteq K^n$.

(ii) $Y = \{\text{functions with period } T\big|\pi\} \subseteq \{\text{functions } \mathbb{R} \to \mathbb{R}\}$.

(iii) $Y = \{\text{constant functions } S \to K\} \subseteq \{\text{functions } S \to K\}$.

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

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. $\square$

Spanning and Independence

Definition

A linear combination of vectors $x_1, \dots, x_j$ is a vector of the form $a_1 x_1 + \cdots + a_j x_j$, 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 $\mathrm{Span}(S)$.

Exercise

For any subset $S \subseteq X$,

$$\mathrm{Span}(S) = \bigcap_{S \subseteq Y_\alpha \leq X} Y_\alpha \,,$$

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

Definition

The vectors $x_1, \dots, x_j$ are linearly dependent if we can write $a_1 x_1 + \cdots a_j x_j = 0$, where not all $a_i = 0$. Otherwise, the vectors are linearly independent.

Spanning, linear independence, and bases

Lemma 1.1

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

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

$$\text{Span}(x_1, x_2 \dots, x_n) = \text{Span}(y_1, x_2 \dots, x_n) = X$$

Repeat this process: eliminating each $x_2, x_3, \dots$. Note that $j > n$ is impossible. (Why?) $\square$

Definition

A set $B \subset X$ is a basis for X if:

- B spans X. (is "big enough");
- B is linearly independent. (isn't "too big").

Bases

Lemma 1.2

If $\mathrm{Span}(x_1, \ldots, x_n) = X$, 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)

$$x_n = a_1 x_1 + \cdots + a_{n-1} x_{n-1}\,.$$

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

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, $m \leq n$ and $n \leq m$. $\square$

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 $\mathrm{Span}(y_1, \ldots, y_j) \neq X$, then find $y_{j+1} \in X$ not in $\mathrm{Span}(y_1, \ldots, y_j)$, 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. $\square$

An example from ODEs

Let X be the set of all (twice-differential) 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:

$$x(0) = x_0 \quad \text{(initial position)} \qquad\qquad x'(0) = v_0 \quad \text{(initial velocity)}.$$

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 $C_1 e^{ix} + C_2 e^{-ix}$ 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 $x \in X$ can be written uniquely as

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

Proof

Pick $y_1 \in Y$ and extend this to a basis $y_1, \dots, y_j$ of Y. By Lemma 1.1, $j \leq \dim X < \infty$.

Extend this to a basis $y_1, \dots, y_j, z_{j+1}, \dots, z_n$ of X [and define $Z := \operatorname{Span}(z_{j+1}, \dots, z_n)$].

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

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 $Y_1, \dots, Y_m$ 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

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

Example

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

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

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

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:

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

and the following subspaces:

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

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

$$x = x_{i_1} + x_{i_2} + \cdots + x_{i_k}, \qquad x_{i_j} \in X_{i_j}.$$

Thus, we can write the direct sum as follows:

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

- Elements in the direct product are sequences, e.g., $x = (1, 1, 1, \dots)$.
- 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**: $x \equiv y$ implies $y \equiv x$;

(ii) **reflexive**: $x \equiv x$ for all $x \in X$;

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

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

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

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 $a\overline{x}=\overline{ax}$;
- $(x+Y)+(z+Y)=x+z+Y$, 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 $y_1, \dots, y_j$ be a basis for Y. Extend this to a basis $y_1, \dots, y_j, x_{j+1}, \dots, x_n$ of X.

Claim: $\{x_{j+1}\}, \dots, \{x_n\}$ *is a basis of* X/Y.

- Show this spans X/Y:

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

  $$\{x\} = \Big\{\sum a_i y_i + \sum b_k x_k\Big\} = \sum a_i\{y_i\} + \sum b_k\{x_k\} = \sum b_k\{x_k\}\,.$$

- Show this is linearly independent:

  Suppose $\sum_{k=j+1}^{n} c_k\{x_k\} = \{0\}$, which means $\sum c_k x_k = y$ for some $y \in Y$.

  Write $y = \sum_{i=1}^{j} d_i y_i$, and so $\sum c_k x_k - \sum d_i y_i = 0$, and hence all $c_k, d_i = 0$ (Why?). $\square$

Corollary

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

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 $\overline{X} = \overline{U} + \overline{V}$, and so $\dim \overline{X} = \dim \overline{U} + \dim \overline{V}$ (Theorem 1.5).

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

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

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