Linear Algebra Short Course

Lecture 1

Matthew J. Holland

matthew-h@is.naist.jp

Mathematical Informatics Lab
Graduate School of Information Science, NAIST

Some useful references

- Properties of vector spaces: Axler (1997, Ch. 1–2)
- Various spaces, more analysis than LA: Luenberger (1968, Ch. 2–3)
- Basic topology of metric spaces: Spivak (1965, Ch. 1), Mendelson (1990, Ch. 2), Rudin (1976, Ch. 1–2)

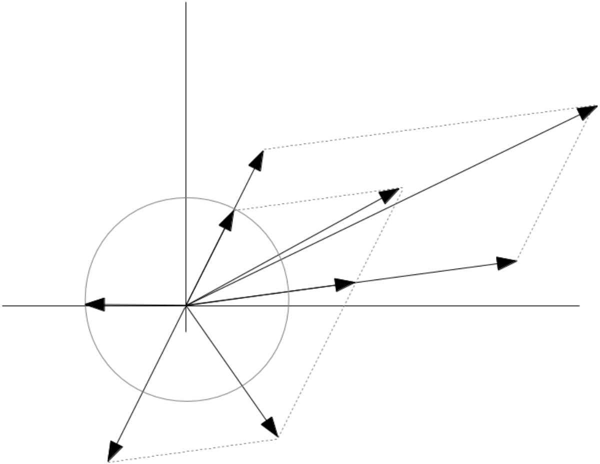
Lecture contents

1. **Basic framework of linear algebra**
2. **Properties and structure of linear spaces**
3. **Analysis on general vector spaces**
4. **Some important spaces**

Lecture contents

1. **Basic framework of linear algebra**
2. **Properties and structure of linear spaces**
3. **Analysis on general vector spaces**
4. **Some important spaces**

General motivations 1



In high school physics, we build intuition with objects called *vectors* and *scalars*, their properties of *length* and *angle*, and operations such as *rotations* and *translations*.

General motivations 2

Even at the elementary level, there is a natural progression:

Geometric (drawing arrows, etc.)

$\Downarrow$

Algebraic (using symbols, <u>defining operations</u>)

In basic "vector analysis," what sort of operations do we define?

V + V, S * V, projection, scalar product, vector product, . . .

Defining these for "vectors" in $\mathbb{R}^3$ and "scalars" in $\mathbb{R}$ is very fruitful!

But in mathematics, $\mathbb{R}^3$ is not the only space we're interested in.
What about $\mathbb{R}^n$ or even $\mathbb{R}^\infty$? Spaces of functions?

General motivations 3

What we do in linear algebra:

Define analogous operations on more general spaces

$\Downarrow$

Investigate their properties (i.e., prove interesting theorems)

We take an axiomatic approach to this.
Why is this a productive endeavour?

We only need to prove things once!

If $\mathbb{R}^3$ and $\mathbb{R}^\infty$ and $\mathcal{C}[a,b]$ all satisfy our axioms, and our proofs only use those axioms, proving it for one implies the others.

Enough heuristics, let's get started.

Basic framework: axioms for scalars 1

Our pool of scalars will be a “field” $\mathbb{F}$, defined below.

Defn. If non-empty set $\mathbb{F}$ with binary addition/multiplication operations defined such that (FA), (FM), and (FD) hold, we call $\mathbb{F}$ a **field**. Taking arbitrary $x, y, z \in \mathbb{F}$,

Addition axioms:

FA.1 $x, y \in \mathbb{F} \implies (x + y) \in \mathbb{F}$

FA.2 $x + y = y + x$

FA.3 $(x + y) + z = x + (y + z)$

FA.4 $\exists\, x' \in \mathbb{F}, x' + x = x, \forall\, x \in \mathbb{F}$. Denote 0.

FA.5 $\exists\, x' \in x', x' + x = 0$. Denote $-x$.

Basic framework: axioms for scalars 2

Multiplication and distribution axioms:

FM.1 $x, y \in \mathbb{F} \implies xy \in \mathbb{F}$

FM.2 $xy = yx$

FM.3 $(xy)z = x(yz)$

FM.4 $\exists\, x' \in \mathbb{F}, x'x = x, \forall\, x \in \mathbb{F}$. Denote 1.

FM.5 $\exists\, x' \in x', x'x = 1$. Denote $1/x$.

FD.1 $x(y+z) = xy + xz$

(*) Note $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ with usual operations are fields.

Basic framework: axioms for vectors 1

With underlying field $\mathbb{F}$ used as our source of scalars, now we discuss axioms for V, a set from which we get our vectors.

Defn. The non-empty set V, equipped with scalar multiplication and vector addition operations, is called a **(linear) vector space** on $\mathbb{F}$ when (VA), (VM), and (VD) below hold. Take $u, v, w \in V, \alpha, \beta, 0, 1 \in \mathbb{F}$.

Vector addition axioms:

VA.1 $u + v = v + u \in V$

VA.2 $(u + v) + w = u + (v + w)$

VA.3 $\exists\, \theta \in V, \theta + u = u, \forall\, u \in V$

VA.4 $\exists\, u' \in V, u + u' = \theta, \forall\, u \in V$. Denote $-u$.

Basic framework: axioms for vectors 2

Scalar multiplication and distribution axioms:

VM.1 $(\alpha\beta)u = \alpha(\beta u) \in V$

VM.2 $0u = \theta, \forall\, u \in V$

VM.3 $1u = u, \forall\, u \in V$

VD.1 $\alpha(u + v) = \alpha u + \alpha v$

VD.2 $(\alpha + \beta)u = \alpha u + \beta u$

Typically we just denote all additive identities by 0, so let $\theta = 0$.

That's all the groundwork we'll need to build our framework.

Basic properties

(*) Let $U_1, \ldots, U_n$ be vector spaces on common field $\mathbb{F}$. Then with the usual definition of the Cartesian product, verify $U_1 \times \cdots \times U_n$ is a vector space on $\mathbb{F}$.

(*) The following properties follow from our axioms on V:

VM.4 $\alpha 0 = 0$

VD.3 $(\alpha - \beta)x = \alpha x - \beta x$

VD.4 $\alpha(x - y) = \alpha x - \alpha y$

VC.1 $x + y = y + z \implies x = z$

VC.2 $\alpha \neq 0, \alpha x = \alpha y \implies x = y$

VC.3 $x \neq 0, \alpha x = \beta x \implies \alpha = \beta$

VM.5 $(-\alpha)x = \alpha(-x) = -(\alpha x)$

VM.6 $xy = 0 \implies x = 0$ or $y = 0$

(*) For vector space V, additive identity is always unique. Also, for each $v \in V$, additive inverse always unique.

Subspaces

Certain subsets of vector spaces will be of particular interest:

Defn. Let V be a vector space on field $\mathbb{F}$. Taking $X \subseteq V$, if

$$u + v \in X$$
$$\alpha u \in X$$

$\forall\, u, v \in X, \alpha \in \mathbb{F}$, then we call X a **(linear) subspace** of V.

Subspaces are thus the subsets closed under vector sums and scalar products.

(*) Note $X \subseteq V$ a subspace $\iff$ X is a vector space.

Basic framework of linear spaces

Example. (*) Given the "usual" algebraic operations, the following are linear spaces. Consider the operations and the vector space and field upon which they live.

- $\mathbb{F}^n$, given field $\mathbb{F}$.
- $\mathbb{F}^\infty = \{(x_1, x_2, \ldots) : x_i \in \mathbb{F}, i = 1, 2, \ldots\}$, given field $\mathbb{F}$.
- $\mathcal{P}(\mathbb{F}) = \{p : p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_m x^m\}$, given field $\mathbb{F}$, and coefficients $\alpha \in \mathbb{F}^m, m \geq 1$.
- $\{(x_1, x_2, \ldots) \in \mathbb{R}^\infty : x_n \to 0\}$
- $\{f : [a, b] \to \mathbb{R}; f$ continuous on $[a, b]\}$.

Lecture contents

1. **Basic framework of linear algebra**
2. **Properties and structure of linear spaces**
3. **Analysis on general vector spaces**
4. **Some important spaces**

Lecture contents

Sums and decompositions into direct sums

Defn. Let V be a vector space, and $U_1, \ldots, U_n \subset X$ be subsets. We define the **sum** of these sets respectively as

$$U_1 + \cdots + U_n := \{u_1 + \cdots + u_n : u_i \in U_i, 1 \leq i \leq n\}.$$

If for every $z \in V$, we have that each representation

$$z = u_1 + \cdots + u_n, \text{ where } u_i \in U_i, 1 \leq i \leq n$$

is unique, then we write $V = U_1 \oplus \cdots \oplus U_n$, the **direct sum** of the U_i.

What operations preserve linearity?

Given a sum, under what conditions is it a direct sum decomposition?

Basic properties of sums

Let $U, W, U_1, \ldots, U_n \subset V$ be subspaces of vector space V.
(*) Then,

$$V = U_1 \oplus \cdots \oplus U_n \iff U_1 + \cdots + U_n = V$$
$$\text{and } 0 = u_1 + \cdots + u_n \text{ uniquely } u_i = 0.$$

(*) This leads to a nice corollary,

$$V = U \oplus W \iff V = U + W \text{ and } U \cap W = \{0\}.$$

(**) Sums and intersections (happily) preserve linearity:

$$U + W \text{ and } U \cap W \text{ are subspaces.}$$

(*) Note this extends to arbitrary sums/intersections of subspaces.
(*) Unions need not preserve linearity.
(*) For subspaces $U, W \subset V$, have that $[U \cup W] = U + W$.

Linear combinations

Defn. Given vector space V on $\mathbb{F}$, for any $m \geq 1$ elements $x_1, \ldots, x_m \in V$ and $\alpha \in \mathbb{F}^m$, we call

$$\alpha_1 x_1 + \cdots + \alpha_m x_m$$

a **linear combination** of these elements.

(*) Note we only defined pairwise sums, but the axioms imply this is notation is unambiguous.

(*) If $S \subset V$ is a subspace then S closed under linear combinations.

Defn. For subset $T \subset V$, define

$$[T] := \{\text{all linear combinations of elements in } T\}$$

called the **subspace generated by** T, or the “span” of T.

(*) Validate this defn; $[T]$ a subspace of V, the “smallest” containing T.

Translations of linear spaces

Example. Consider the hyperplane $H \subset \mathbb{R}^n$ given by

$$H = \{x : \alpha^T x = b\}, b \neq 0.$$

(*) While defined by a linear relation, this is *not* a subspace.

Defn. Any $W \subset V$ containing all lines through any two points we call an **affine** set. That is, for $u, v \in W$, have

$$\lambda u + (1 - \lambda)v \in W, \ \forall \lambda \in \mathbb{R}.$$

The **affine hull** of a set $T \subset V$ is defined

$$\operatorname{aff} T := \cap W_i$$

intersection over all affine sets $W_i \subset V$ s.t. $T \subset W_i$.

(*) Validate this definition; $\operatorname{aff} T$ is well-defined, is affine.
(*) Every affine set is a translation of a subspace.

Linear dependence, dimension, basis

Foundational concepts for analysis of linear spaces.
Assume V a vector space on $\mathbb{F}$.

Defn. Take non-empty $S \subset V$. We say $x \in V$ is **linearly dependent** on S if x is a linear combination of elements of S. Equivalently,

$$x \text{ is linearly dependent on } S \iff x \in [S].$$

If this doesn't hold, say x is **linearly independent** of S. Analogously, say $S \subset V$ is a **linearly independent set** of vectors when

$$u \text{ lin indep of } S \setminus \{u\}, \ \forall\, u \in S.$$

(*) For any finite set $\{x_1, \ldots, x_n\} \subset V$, the following is key:

$$\{x_1, \ldots, x_n\} \text{ is linearly indep.} \iff \sum_{i=1}^{n} \alpha_i x_i = 0 \text{ implies all } \alpha_i = 0.$$

Linear dependence, dimension, basis

Defn. We call a vector space V **finite dimensional** if there exists a *finite* subset $B \subset V$ such that $[B] = V$. If no such finite subset exists, call V **infinite-dimensional**.

If B is linearly independent, call B a **basis** of V.

(*) If $\{v_1, \ldots, v_n\}$ a basis of V, then every $v \in V$ may be uniquely represented in the form

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

(*) In fact, the uniqueness of this representation *characterizes* $\{v_1, \ldots, v_n\}$ as a basis.

Does a basis B always exist?
Can we make any statements about its length $|B|$?
Is dimensionality monotonic in some sense?

Linear dependence, dimension, basis

(**) The following foundational results are valid:

Let V be a finite-dim vector space. Then,

- $\exists\, B \subset V$ s.t. B a basis of V
- if B, C bases of V, then $|B| = |C|$.
- $S \subset V$ a subspace $\implies$ S finite-dim.
- $S \subset V$ a subspace $\implies$ $|B_S| \leq |B_V|$, each respective bases.

These results completely motivate a (now well-definable) dimension notion.

Linear dependence, dimension, basis

Defn. Let V be a finite-dimensional vector space. Define the **dimension** of V by $\dim V := |B|$, where B is any basis of V. If V is infinite-dim, let $\dim V := \infty$, if $V = \{0\}$, let $\dim V = 0$.

(*) Clearly we have our monotonicity, where subspace $S \subset V$ satisfies

$$\dim S \leq \dim V.$$

As well, one may show that if we know $\dim V = n$, we only need one more piece of information to validate a given $\{v_1, \ldots, v_n\} \subset V$ as a basis, since

$$[\{v_1, \ldots, v_n\}] = V \implies \{v_1, \ldots, v_n\} \text{ a basis of } V$$
$$\{v_1, \ldots, v_n\} \text{ lin indep} \implies \{v_1, \ldots, v_n\} \text{ a basis of } V.$$

Linear dependence, dimension, basis

Let $S, T, U, U_1, \dots, U_n \subset V$ be subspaces, $\dim V < \infty$.

(**) Handily, it is possible to verify

$$\dim(S+T) = \dim S + \dim T - \dim(S \cap T)$$
$$\dim(U_1 + \cdots + U_n) \leq \dim U_1 + \cdots + \dim U_n.$$

(*) Unfortunately, the following *does not* hold in general:

$$\begin{aligned}\dim(S+T+U) = {} & \dim S + \dim T + \dim U - \dim(S \cap T) \\ & - \dim(S \cap U) - \dim(T \cap U) + \dim(S \cap T \cap U).\end{aligned}$$

Linear dependence, dimension, basis

Interestingly, the "structure" of vector space V, $\dim V < \infty$ is captured well by its dimension.

(*) For any vector space V, $\dim V = n$, there exist one-dim $V_1, \ldots, V_n$ s.t.

$$V = V_1 \oplus \cdots \oplus V_n.$$

(*) Also, if $S \subset V$ is a subspace, then

$$\dim S = \dim V \implies S = V.$$

(**) Taking subspaces $U_1, \ldots, U_n \subset V$,

$$V = U_1 + \cdots + U_n \text{ and } \dim V = \sum_{i=1}^{n} \dim U_i \iff V = U_1 \oplus \cdots \oplus U_n.$$

This proof is another straightforward exercise.

Lecture contents

1. Basic framework of linear algebra
2. Properties and structure of linear spaces
3. Analysis on general vector spaces
4. Some important spaces

Lecture contents

Brief analysis review

Linear spaces are *ubiquitous* in mathematics. To introduce some important examples, let's review a few basic concepts from analysis.

Defn. Given set X, a function $d : X \times X \to \mathbb{R}_+$ is called a **metric** if $\forall\, x, y, z \in X$,

M.1 $d(x, y) \geq 0,$ with equality iff $x = y$

M.2 $d(x, y) \leq d(x, y) + d(y, z)$

M.3 $d(x, y) = d(y, x)$

We call X equipped with a metric d a **metric space**.

(**) The following are metric spaces:

- $\mathbb{R}^n$ with $d(x, y) := (\sum_{i=1}^n (x_i - y_i)^2)^{1/2}$
- $\mathbb{R}^n$ with $d(x, y) := \max_i |x_i - y_i|$
- $\mathcal{C}[a, b]$ with $d(f, g) := \int_a^b |f(t) - g(t)|\, dt$
- $d(f, g) := \sup_{a \leq x \leq b} |f(x) - g(x)|, f, g$ bounded on $[a, b] \subset \mathbb{R}$.

Brief analysis review

Defn. Denoting the ε-radius ball at x_0 in metric space X by

$$\varepsilon B(x_0) := \{x \in V : d(x, x_0) < \epsilon\},$$

for any $S \subset X$, call $u \in S$ an **interior point** if $\exists\, \varepsilon > 0$ s.t.

$$\varepsilon B(u) \subset S.$$

Denote all such points by $\operatorname{int} S$, the **interior** of S. If $S = \operatorname{int} S$, call S an **open** subset of X.

Call $p_0 \in V$ a **limit point** of $S \subset V$ if $\forall\, \delta > 0$,

$$\exists\, x \in S, x \neq p_0, \text{ s.t. } x \in \delta B(p_0).$$

If S^* is all limit points of S, call $\bar{S} := S \cup S^*$ the **closure** of S. Call S a **closed** subset of X when $S = \bar{S}$.

Vector "magnitude" in linear spaces

Defn. If V a vector space on field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, we call a map $x \mapsto \|x\| \in \mathbb{R}_+$ a **norm** if $\forall\, u, v \in V, \alpha \in \mathbb{F}$,

N.1 $\|u\| > 0$ for $u \neq 0$, and $\|0\| = 0$.

N.2 $\|u + v\| \leq \|u\| + \|v\|$

N.3 $\|\alpha u\| = |\alpha| \|u\|$

We call V equipped with a norm a **normed linear space**.

(*) Note any norm on V induces a valid metric on V.

(*) What about the converse? Consider a "reverse indicator" metric.

(*) $\mathcal{C}[a, b]$ with $\|f\| := \sup_{a \leq x \leq b} |f(x)|$ is normed vec space.

Convergence in normed linear spaces

Let $(X, \|\cdot\|)$ be a normed vector space, and (x_n) a sequence of vectors $x_1, x_2, \ldots \in X$.

Defn. We say a sequence (x_n) **converges** to $x \in X$ (in the norm $\|\cdot\|$), denoted $x_n \to x$,

$$\text{whenever } \lim_{n\to\infty} \|x_n - x\| = 0,$$

noting $(\|x_n - x\|)$ is a sequence of real numbers.

(*) The limits of convergent sequences are unique.
(*) $S \subset X$ is closed $\iff$ Every sequence (x_n) in S converges in S.

Continuous maps in normed linear spaces

Once again say $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$, are normed vector spaces.

Defn. Continuity of $f : X \to Y$ extends in the natural way, of course. Namely, f is **continuous** at $x_0 \in X$ if $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ s.t.

$$\|x - x_0\|_X < \delta \implies \|f(x) - f(x_0)\|_Y < \varepsilon.$$

Clearly this depends on both norms.

(*) f is continuous at $x_0 \iff x_n \to x_0$ implies $f(x_n) \to f(x_0)$.

Lecture contents

Lecture contents

Banach spaces

Defn. We call (x_n) a **Cauchy sequence** if $\forall\, \varepsilon > 0$, $\exists\, N_0 < \infty$ s.t.

$$m, n \geq N_0 \implies \|x_n - x_m\| < \varepsilon.$$

If all Cauchy sequences on X converge, we say X is **complete** (in norm $\|\cdot\|$). We call a complete normed linear space a **Banach space**.

The "Cauchy condition" is precisely why Banach spaces are nice.

(*) All Cauchy sequences are bounded in norm $\|\cdot\|$.
(*) All convergent sequences are Cauchy.
(*) $\mathcal{C}[a, b]$ with $\|f\| := \sup_{a \leq x \leq b} |f(x)|$ is Banach.
(**) $\mathcal{C}[a, b]$ with $\|f\| := \int_a^b |f(x)|\, dx$ is *not* Banach.

More on Banach spaces

(*) If X, Y are Banach, the “usual” product space $(X \times Y, \|\cdot\|)$ with $\|\cdot\| := \|\cdot\|_X + \|\cdot\|_Y$ is Banach.

(*) Let X be Banach; subset $S \subset X$ is complete $\iff$ S is closed.

(**) Another key result: if X is a normed linear space, for $S \subseteq X$,

$$\dim S < \infty \implies S \text{ is complete.}$$

Example: ℓ_p space, $1 \le p \le \infty$

Here we introduce “the” classical Banach space.

Defn. Define **ℓ_p-space** for $1 \le p < \infty$ by

$$\ell_p := \left\{ (x_1, x_2, \ldots) : \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}.$$

The norm of interest is of course $\|x\|_p := \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$.

For the $p = \infty$ case, we consider *bounded* sequences, and intuitively we define $\|x\|_\infty := \sup |x_i|$.

Example: $L_p(\Omega, \mathcal{A}, \mathbf{P})$ space, $1 \leq p \leq \infty$

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space.
Defn. We define L_p**-space** on $(\Omega, \mathcal{A}, \mathbf{P})$ for $1 \leq p < \infty$ by

$$L_p := \left\{ h : \mathbf{E}\,|h|^p = \int_\Omega |h(\omega)|^p \ \mathbf{P}(d\omega) < \infty \right\},$$

and the usual norm is $\|h\|_p := (\mathbf{E}\,|h|^p)^{1/p}$.

For $p = \infty$ case, consider bounded functions $\sup_\omega |h(\omega)| < \infty$.

Minor complication:
Even if $g, h \in L_p$ are $g \neq h$, we might have $g = h$ a.e. $[\mathbf{P}]$.

(*) For those familiar with Lebesgue integration, why is defining $\|h\|_\infty = \sup_\omega |h(\omega)|$ inadvisable? Any ideas for an alternative?

Both ℓ_p and L_p are Banach

Very important, classic results, proofs are critical for any serious student of analysis (out of scope here).

The basic flow (ℓ case) is:

(1) For $x \in \ell_p, y \in \ell_q$ where $1/p + 1/q = 1$, prove Hölder's inequality,

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_q.$$

(2) Using Hölder, for $x, y \in \ell_p$ prove Minkowski's inequality:

$$x + y \in \ell_p \text{ and } \|x + y\|_p \leq \|x\|_p + \|y\|_p,$$

namely the triangle inequality. Definiteness easy on ℓ_p, but requires thought on L_p.

(3) Then just need completeness. ℓ_p is basic analysis, L_p requires some Lebesgue theory.

Inner product and Hilbert space

Consider vector space V on field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Defn. Call $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ an **inner product** on V if $\forall\, u, v, w \in V, \alpha \in \mathbb{F}$,

IP.1 $\langle u, v \rangle = \overline{\langle v, u \rangle}$

IP.2 $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$

IP.3 $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$

IP.4 $\langle u, u \rangle \geq 0$, and $\langle u, u \rangle = 0 \iff u = 0$.

(*) Additivity holds in both arguments. Also, $\langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle$.

(**) We'll later see that an IP on V induces a norm on V (Lec 4).

Defn. Call $(V, \langle \cdot, \cdot \rangle)$ an **inner product space**. A complete IP space is called **Hilbert space**. Much more later.

Lecture contents

Lecture contents

References

Axler, S. (1997). *Linear Algebra Done Right*. Springer, 2nd edition.

Luenberger, D. G. (1968). *Optimization by Vector Space Methods*. Wiley.

Mendelson, B. (1990). *Introduction to Topology*. Dover, 3rd edition.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition.

Spivak, M. (1965). *Calculus on Manifolds*. Addison-Wesley.