Linear Algebra 4

University of Ostrava

Version of May 7, 2024

Literature

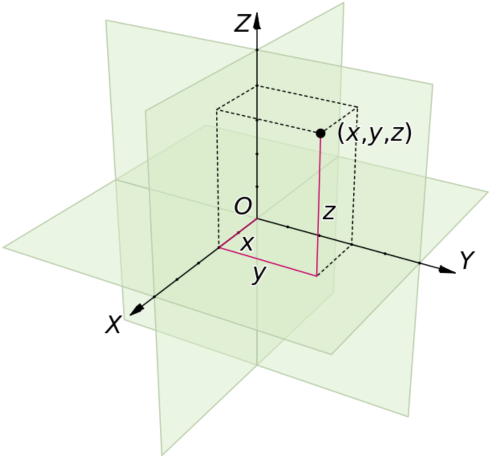
- A.I. Kostrikin and Yu.I. Manin, *Linear Algebra and Geometry*, Gordon and Breach, 1997
  (referred in what follows as KOSTRIKIN–MANIN)
- S. Mac Lane and G. Birkhoff, *Algebra*, 3rd ed., AMS Chelsea, 1999
  (referred as MAC LANE–BIRKHOFF)
- A. Onishchik and R. Sulanke, *Projective and Cayley–Klein Geometries*, Springer, 2006
  (referred as ONISHCHIK–SULANKE)
- M. Postnikov, *Lectures in Geometry. Semester I. Analytic Geometry*, Mir, Moscow, 1982
  (referred as POSTNIKOV)

(All images, unless specified otherwise, are courtesy of Wikipedia)

Affine spaces

Refresher: Vector space

From the previous linear algebra courses, you (suppose to) know already what a vector (aka linear) space is.



The vector space $\mathbb{R}^3$

Refresher: Vector space (cont.)

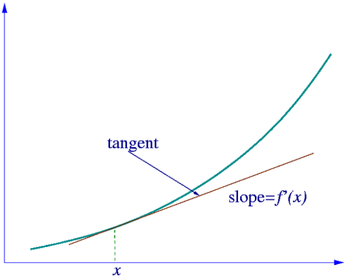
Abstractly, the vector space is a set equipped with operation of addition and multiplication on scalars (elements of the base field) subject to the usual condition satisfied by addition of vectors in $\mathbb{R}^n$ and multiplication of a vector by a scalar:

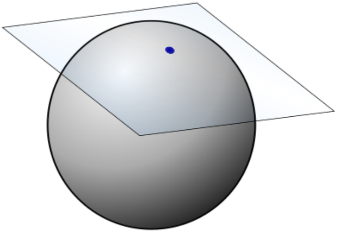
$$a + b = b + a$$
$$(a + b) + c = a + (b + c)$$
$$\lambda(a + b) = \lambda a + \lambda b$$
$$\text{etc} \ldots$$

Vector space is one of the central notions of mathematics and embodies the ideas of *linearity* and *linearization*.

Refresher: Importance of vector spaces

Most of the phenomena in the nature are described by nonlinear structures, but they can be approximated by linear ones; the paradigmatic examples are the notion of derivative in analysis (which is nothing but a slope of a tangent line – i.e., an one dimensional space – at a point to a curve), or, more generally, the notion of the tangent plane at a point to the manifold:





Affine space

Now meet affine spaces, a specialization of the notion of vector spaces.

Definition

An *affine space* over a field K is a triple $(A, V, +)$, where A is a set, V is a vector space over K, and $+$ is a binary operation $A \times V \to A$, $(a, v) \mapsto a + v$, satisfying the following axioms:

1. $(a + u) + v = a + (u + v)$ for any $a \in A$, $u, v \in V$
2. $a + 0 = a$ for any $a \in A$
3. for any $a, b \in A$ there exists a unique $v \in V$ such that $b = a + v$

In other words: the additive group of V acts on A (conditions 1 and 2) transitively and freely (condition 3).

Affine space (cont.)

Elements of A are called *points*, and elements of V are called *vectors*. The unique v from axiom 3 is called the *difference* of the corresponding points and is denoted by $b - a$.

The maps $t_v : A \to A$, $a \mapsto a + v$ are called *translations*.

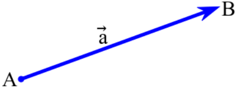
The *dimension* of the affine space is the dimension of its underlying vector space, V.

Often we will omit the symbol of operation "$+$", and denote an affine space as merely a pair (A, V).

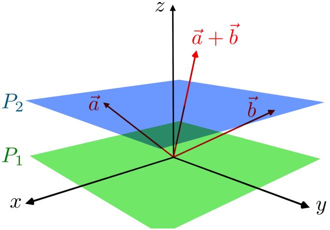
Intuitively, affine space may be thought as a vector space whose origin 0 is "forgotten": only operations of translation by a vector, summation of translations, and multiplication of a translation by a scalar are retained.

Examples of affine spaces

- A paradigmatic example: the "points" are points in $\mathbb{R}^3$, and "vectors" are vectors in $\mathbb{R}^3$, with vectors acting at a point.



- More generally: $(V, V, +)$, where V is a vector space, i.e. the sets of vectors and points are the same, with the action being the usual addition of vectors.
- $(P_2, P_1, +)$, where P_2 is the set of solutions of the inhomogeneous linear system $Ax = b$, P_1 is the set of solutions of the corresponding homogeneous system $Ax = 0$.



Examples of affine spaces (cont.)

Exercise

Try to give more examples of affine spaces. What is the general pattern for constructing them?

Hint: see, for example, Wikipedia.

Properties of affine spaces

Lemma

Let $(A, V, +)$ be an affine space.

1. For any $v \in V$, the map $A \to A$, $a \mapsto a + v$ is a bijection.
2. For any $a \in A$, the map $V \to A$, $v \mapsto a + v$ is a bijection.
3. For any $a \in A$, $v \in V$, there is a unique $b \in A$ such that $b - a = v$.
4. For any $a, b, c \in A$, it holds $(c - b) + (b - a) = c - a$.
5. For any $a \in A$, it holds $a - a = 0$.
6. For any $a, b \in A$, $v, u \in V$, it holds $(a + v) - (b + u) = (a - b) + (v - u)$.

Warning

Be careful! While $a + v$ is a legitimate expression, $v + a$ is not! For example, in heading 4 from the previous Lemma we cannot merely "open the brackets" and write something like $(c - b) + b$, etc.

For the general rules and examples of manipulating with expressions containing multiple points and vectors with plus and minus signs, see Kostrikin–Manin, pp. 196–197, and Mac Lane–Birkhoff, pp. 566–567.

Affine maps

Affine maps

Definition

Let (A_1, V_1) and (A_2, V_2) be affine spaces. The pair (f, Df) of maps $f : A_1 \to A_2$, $Df : V_1 \to V_2$ such that Df is a linear map, and for any $a, b \in A_1$,

$$f(a) - f(b) = Df(a - b),$$

is called an *affine map* of the first affine space into the second one.

Note

Since a, b run through all A, Df is defined by f uniquely, so often (by abuse of language) we will call an affine map just f.

Examples of affine maps

- A linear map $f : V_1 \to V_2$ from a vector spaces V_1 to V_2 is an affine map from (V_1, V_1) to (V_2, V_2).
- Any translation is an affine map of the affine space to itself.
- An identity map id_A is an affine map.

Exercise

What will be the corresponding Df in all these cases?

Composition of affine maps

Theorem 1

The composition of affine maps is an affine map.

Note

For those who knows (and appreciates) category theory: this Theorem, together with the fact that the identity map is an affine map, implies that the affine spaces form a category in which morphisms are affine maps.

Theorem 2

Let $f, g : (A, V) \to (B, W)$ be two affine maps. Then $Df = Dg$ if and only if one of them is a composition of another one with a translation by some (unique) vector (i.e., say, $g = t_w \circ f$ for some $w \in W$).

Affine coordinates

Affine coordinates

A *system of affine coordinates* of an n-dimensional affine space (A, V) over a field K consists of a point $a_0 \in A$, and a basis $\{e_1, \dots, e_n\}$ of V.
The *coordinates* of the point $a \in A$ in this system is an n-tuple $(x_1, \dots, x_n) \in K^n$ such that

$$a = a_0 + x_1 e_1 + \cdots + x_n e_n.$$

Theorem

For any system of affine coordinates, every point has a unique coordinates in it.

Hint: the proof is not difficult, and follows from the previous theorems and facts.

Barycentric combination

Definition

Let (A, V) be an affine space over a field K, and $a_1, \ldots, a_n \in A$. For any $x_1, \ldots, x_n \in K$ such that $x_1 + \cdots + x_n = 1$, define the formal sum $\sum_{i=1}^{n} x_i a_i$ as

$$\sum_{i=1}^{n} x_i a_i = a + \sum_{i=1}^{n} x_i (a_i - a)$$

for some $a \in A$. It is called the *barycentric combination* of the points $\{a_i\}$ with coefficients $\{x_i\}$.

Theorem

This definition does not depend on a.

Warning

For example, for two points a, b, the expression $\frac{1}{2}a + \frac{1}{2}b$ makes sense by above, while the summands, $\frac{1}{2}a$ and $\frac{1}{2}b$, are meaningless!

A physical meaning of the barycentric combination

In the affine space $(\mathbb{R}^3, \mathbb{R}^3)$, the barycentric combination represents the "center of gravity" of a system of masses located at the points $\{a_i\}$, weighted accordingly by coefficients $\{x_i\}$.

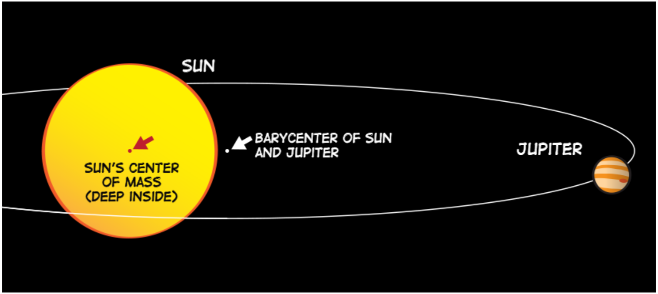


Image courtesy of NASA

Barycentric combination and affine maps

Theorem

Let (A, V), (B, W) be two affine spaces. A function $f : A \to B$ is an affine map if and only if for any finite sets $\{a_i\}_{i=1}^n$ and $\{x_i\}_{i=1}^n$ of elements of A and of the ground field respectively, such that $\sum_{i=1}^n x_i = 1$,

$$f(\sum_{i=1}^n x_i a_i) = \sum_{i=1}^n x_i f(a_i).$$

(I.e., a map is affine if and only if it "preserves" the barycentric combinations).

Characterization of affine spaces

Theorem

Any affine space is affine isomorphic to the affine space of the form (V, V) for some vector space V.

The proof uses the barycentric coordinates and the previous theorem.

Corollary

Any n-dimensional affine space over a field K is affine isomorphic to (K^n, K^n).

Affine group

Affine group

Theorem 1

The set of bijective affine maps of an affine space (A, V) forms a group with respect to composition.

Definition

This group is called the *affine group* of the affine space (A, V), and is denoted by $Aff(A, V)$.

Theorem 2

For any affine space (A, V), it holds $Aff(A, V) \simeq V \rtimes GL(V)$.

For an explicit isomorphism, and rules of computation in $V \rtimes GL(V)$, see Kostrikin–Manin, pp. 203–204.

Affine subspaces

Affine subspace

Definition

An *affine subspace* of an affine space (A, V) is an affine space (B, W) satisfying one of the following conditions:

1. $B = \varnothing$ and $W = 0$;
2. $B \subseteq A$, $W = \{b_1 - b_2 \mid b_1, b_2 \in B\}$ is a vector subspace of V, and $t_w(B) \subseteq B$ for any $w \in W$.

Note

As W is determined by B uniquely, often (and again by abuse of terminology) we will call an affine subspace just B.

Parallel affine subspaces

Definition

Two affine subspaces of the same dimension of the same affine space are called *parallel* if they have the common space W from the previous definition.

Theorem

Parallel affine subspaces either do not have common points, or do coincide.

Exercise

What happens in the situation where W's do not coincide, but one of them is contained in another?

See Kostrikin–Manin, p. 208.

Affine span

Definition

Let S be a subset of the set of points of an affine space (A, V). The smallest affine subspace of (A, V) whose set of points contains S, is called an *affine span* of S.

Affine span always exists. (Why?)

Theorem

The set of points of the affine span of a set S is equal to the set of barycentric combinations of elements of S.

Affine subspace (cont.)

Theorem 1

Let $f_1, \dots, f_n$ be affine maps from affine space A to the one-dimensional affine space K. Then the set of points $\{a \in A \mid f_1(a) = \dots = f_n(a) = 0\}$ is an affine subspace of A. If A is finite-dimensional, then any its affine subspace has this form.

Theorem 2

A subset of an affine space is an affine subspace if and only if it contains the affine span of any its two points (or, in other words, the straight line passing though those two points).

Euclidean spaces

Euclidean space

Euclidean space = affine space + metric space

Definition

An *(affine) Euclidean space* is an affine space (A, V) over $\mathbb{R}$ such that the set of points A is equipped with the metric d satisfying the following property: for any $a, b \in A$, $d(a, b) = |a - b|$, where $|\cdot|$ is the standard Euclidean metric in the real vector space V. $|a - b|$ is called the *distance* between two points a and b.

Motions of the Euclidean space

Definition 1

A *motion* of an Euclidean space (A, V) is an affine map (f, Df) such that $f : A \to A$ preserves distances: $f(d(a,b)) = d(a,b)$ for any $a, b \in A$.

Definition 2

Let (A, V) be an affine space, and G a subgroup of $GL(V)$. The inverse image of G with respect to the canonical homomorphism $Aff(A, V) \to GL(V)$ is called the *affine extension* of G.

Theorem

The motions of an Euclidean space forms a group with respect to composition, which coincides with affine extension of the group of orthogonal isometries $O(V)$ of the Euclidean real space V.

Examples of motions

Motions of an n-dimensional Euclidean space are:

- $n = 1$: translations and reflections relative to a point.
- $n = 2$: translations, rotations relative to a point, and composition of a reflection relative to a straight line, and translation along this line.

For a detailed study of these examples, as well as (much more complicated) 3-dimensional and the general n-dimensional cases, see KOSTRIKIN–MANIN, pp. 206–207.

Distance and angles in Euclidean space

Definition

An *angle* φ between two vectors v, w in an Euclidean vector space is defined by the formula

$$cos\,\varphi = \frac{|v+w|^2 - |v|^2 - |w|^2}{2\,|v|\cdot|w|}.$$

This conforms our usual geometric intuition about angles.
Exactly the same way as in the standard 2-dimensional and 3-dimensional Euclidean geometry, one can define the distance and angle between two lines, between a line and a plane, the notion of orthogonality, etc., in an abstract Euclidean space.
See POSTNIKOV, p. 129–132,156–159,161–165 for details.

Basics of linear programming

An example

A farmer has a piece of land with the area of 10 km^2, to be planted with either wheat, or barley, or some combination of the two. The farmer has a limited amount of fertilizer, 50 kilograms, and pesticide, 15 kilograms. Every square kilometer of wheat requires 4 kilograms of fertilizer and 1 kilogram of pesticide, while every square kilometer of barley requires 3 kilograms of fertilizer and 2 kilograms of pesticide. Let 1000 EUR be the price of wheat per square kilometer, and 1500 EUR be the price of barley. How to maximize the farmer's profit?

(adapted from Wikipedia)

An example (cont.)

Denoting the area of land planted with wheat and barley by x_1 and x_2 respectively, the problem can be expressed as follows:

Maximize

$$1000x_1 + 1500x_2$$

subject to constraints:

$$\begin{aligned} x_1 + x_2 &\leq 10 \\ 4x_1 + 3x_2 &\leq 50 \\ x_1 + 2x_2 &\leq 15 \\ x_1 \geq 0, x_2 &\geq 0. \end{aligned}$$

Linear programming problem

Given row-vectors $b, c \in \mathbb{R}^n$, and $n \times n$ real matrix M, maximize (or minimize) $c^\top x$ over $x \in \mathbb{R}^n$, subject to $Ax \leq b$ and $x \geq 0$.

In what follows, fix an affine space (A, V) over $\mathbb{R}$. All affine maps are assumed to act from A to itself.

Definition

A *polyhedron* is an intersection of a finite number of sets of points the form $\{a \in A \mid f(a) \geq 0\}$, where f is a non-constant affine map.

Theorem

Assume that and affine map f is bounded above on the polyhedron S. Then f assumes its maximal value at all points of some faces of S, that is also a polyhedron. If S is bounded, then f assumes its maximal value at some vertex of S.

Affine quadrics

Quadratic functions

The notion of quadratic function is an “affinization” of the notion of quadratic form.

Definition

A *quadratic function* on an affine space (A, V) over a field K is a map $Q : A \to K$ of the form

$$Q(a) = q(a - a_0) + \ell(a - a_0) + c, \qquad (*)$$

where $a_0 \in A$ is a point, $q : V \to K$ a quadratic form, $\ell : V \to K$ is a linear form, and $c \in K$ is an element of the ground field.

Lemma

The “quadratic part” q of Q is independent on the choice of a_0. Namely, if (*) holds, then for any $a_0' \in A$

$$Q(a) = q(a - a_0') + \ell'(a - a_0) + c',$$

for some linear map $\ell' : V \to K$ and $c' \in K$.

Center

Definition

The point a_0 is called a *central point* of the quadratic function Q if the linear part ℓ in the equality (*) is equal to zero.
The set of all central points is called the *center* of Q.

The geometric meaning of the center: identifying A with V such that a_0 identifies with 0, the function Q becomes symmetric with respect to reflection $v \mapsto -v$.

Canonical forms of quadratic functions

Theorem

Let Q be a quadratic function on an n-dimensional affine space (A, V) over a field K. Then there exists an affine coordinate system in (A, V) in which Q has one of the following forms:

1. If q is non-degenerate, then $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \lambda_i x_i^2 + c$ for some $\lambda_i, c \in K$;
2. If q is degenerate of rank r, and the center of Q is not empty, then $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \lambda_i x_i^2 + c$ for some $\lambda_i, c \in K$;
3. If q is degenerate of rank r, and the center of Q is empty, then $Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \lambda_i x_i^2 + x_{r+1}$ for some $\lambda_i \in K$.

Affine quadrics

Definition

Let (A, V) be an affine space. An *affine quadric* is a set $\{a \in A \mid Q(a) = 0\}$, where Q is a quadratic function on A.

Theorem

Let the affine quadric X is given by equation $Q_1 = 0$ and $Q_2 = 0$ for some quadratic functions Q_1, Q_2. Then either X is an affine subspace, or $Q_1 = \lambda Q_2$ for some element of the ground field λ.

(In other words, the quadratic function defining a quadric is unique up to multiplication by a scalar).

Conic sections

In the 2-dimensional real case, the quadrics are called *conic sections*. Using the theorem on canonical forms above, all conic sections are reduced to either parabola, or hyperbola, or circle, or a pair of intersecting lines, or a pair of parallel lines, with the center having the usual geometric meaning.
See MAC LANE–BIRKHOFF, p. 585 and
POSTNIKOV, pp. 166–186,193–202.

Projective spaces

Projective spaces

Definition

Let V be a vector space. The set $P(V)$ of straight lines in V (i.e., one-dimensional subspaces) is called the *projective space* of V. The *dimension* of $P(V)$ is equal to $\dim V - 1$.

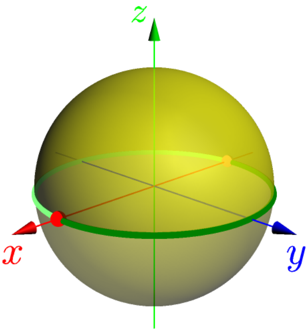
Definition

Select a basis $\{e_0, \ldots, e_n\}$ in V. Every point $p \in P(V)$ is uniquely determined by any nonzero vector on p. The coordinates $\{x_0, \ldots, x_n\}$ of this vector are called *homogeneous coordinates* of p. They are defined up to nonzero scalar factor and denoted by $(x_0 : x_1 : \cdots : x_n)$.

In other words: the n-dimensional projective coordinate space $P(K^{n+1})$ is the set of orbits of the multiplicative group $K^* = K \setminus \{0\}$ acting on the set of nonzero vectors $K^{n+1} \setminus \{0\}$.

Visualization of a projective space

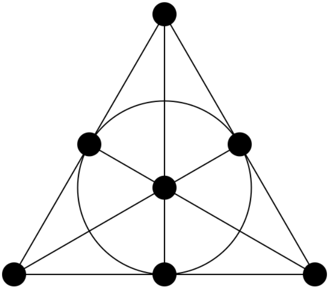
Let $K = \mathbb{R}$ and $V = \mathbb{R}^3$. Any line in $\mathbb{R}^3$ meets the sphere of radius 1 centered at 0 in two antipodal points. Identifying antipodal points, we get a realization of the projective space $P(\mathbb{R}^3)$.



For more details, including higher-dimensional spaces and other ways of visualization, see KOSTRIKIN–MANIN, pp.224–225 and ONISHCHIK–SULANKE, pp. 3–4.

The Fano plane

An example of a 2-dimensional projective space over $GF(2)$:



Exercise

Prove that this is the smallest 2-dimensional projective space.

Hint: $2^3 - 1 = 7$.

Projective subspaces

Projective subspaces

Definition

A *projective subspace* of a projective space $P(V)$ is a set of the form $P(W)$, where W is a vector subspace of V.

Lemma

For two vector spaces V, W, we have $P(V \cap W) = P(V) \cap P(W)$.

Hence the family of the projective subspaces of a given projective space is closed with respect to intersection.

Definition

Let S be a subset of the projective space $P(V)$. The smallest projective subspace of $P(V)$ containing S is called a *projective span* of S, and is denoted by $\overline{S}$.

Theorem

Let P_1, P_2 be finite-dimensional projective subspaces of the same projective space. Then

$$\dim P_1 \cap P_2 + \dim \overline{P_1 \cup P_2} = \dim P_1 + \dim P_2.$$

Projective duality

The dual projective space

Definition

Let V be a vector space, and V^* its dual. Then the projective space $P(V^*)$ is called the *dual space* of the projective space $P(V)$.

Geometrical meaning: points in $P(V^*)$ are hyperplanes in $P(V)$.

When we pass to the projective dual:

- The intersection of projective subspaces turned into the projective span of their dual spaces, and vice versa.
- The inclusion of projective subspaces turned to containment, and vice versa.

The principle of projective duality

Suppose that there is a theorem about projective subspaces of projective spaces, whose formulation involves the notions of dimension, inclusion, intersection, and projective span. Then there is a dual theorem, in which all terms are replaced by their duals according to the rules at the preceding slide.

Note

As stated, this is not a rigorous mathematical theorem. The exact formulation of this principle will require some mathematical logic!

Example

The theorem "Two different planes in a 3-dimensional projective space intersect along one straight line" is the dual of the theorem "One straight line passes through any two points in a 3-dimensional projective space".

Projective group

The projective group

Lemma 1

An injective linear map $f : V \to W$ between vector spaces V, W induces a map $P(f) : P(V) \to P(W)$.

Definition

$P(f)$ is called the *projectivization* of f.

Lemma 2

The set of all projectivizations of invertible endomorphisms of V forms a group with respect to composition.

Definition

This group is called the *projective group* of V and is denoted by $PGL(V)$.

Three groups meet together (projective, general linear, affine)

Lemma

For any vector space V, $GL(V) \to PGL(V)$, $f \mapsto P(f)$ is a surjective homomorphism of groups.

Exercise

What is the kernel of this homomorphism?

See Kostrikin–Manin, p. 234, or Onishchik–Sulanke, p. 20, or Mac Lane–Birkhoff, p. 595.

Theorem

Let (A, W) be an affine space, and W a subspace of codimension 1 in the vector space V. Then the subgroup of $PGL(V)$ consisting of projectivizations which map $P(W)$ to itself, is isomorphic to $Aff(A, W)$.

Projective quadrics

Definition

A *projective quadric* in the projective space $P(V)$ over a field K, is the set $\{p \in P(V) \mid q(p) = 0\}$ for some quadratic form $q : V \to K$. If $\dim P(V) = 2$, the projective quadrics are called *projective conics*. If the quadratic form q has rank r, then the corresponding quadric is said to have rank $r - 1$.

Theorem

In the real projective plane (i.e., in the 2-dimensional projective space), any two nonempty projective conics of rank 2 are equivalent under the projective group.

The End