Notes on Geometry and Linear Algebra

Jacopo Tissino

November 23, 2016

Chapter 1

Vectors

1.1 Euclidean space

We will use the Euclidean n-dimentional space as a model for the physical space. Space is a set, whose elements are points, and whose subsets are lines, planes and hyperplanes, which we will treat algebraically. We will need notions of parallelism, measure and orthogonality.

1.2 Segments

An oriented segment (or "applied vector") is a subset of a space S_n characterized by an ordered pair of points. It is denoted as such: P_1P_2 is the segment (P_1, P_2), where $\forall k \in \mathbb{N} : P_k \in S_n$.

There exist "trivial" segments P_1P_1.

We define an equivalence relation on $S_n \times S_n$: $\sim$, where $P_1P_2 \sim Q_1Q_2 \iff P_1P_2Q_2Q_1$ is a parallelogram.[1] Notably, we do not need a notion of distance for this.

We denote the representative as such: $[(P_1, P_2)] := \overrightarrow{P_1P_2} = \mathbf{v}$. These are "free vectors", or "geometrical vectors".

We have the following operations between them:

- addition: $\mathbf{a} + \mathbf{b}$

- multiplication by a scalar: $\alpha\mathbf{a}, \quad \alpha \in \mathbb{R}$

Addition If we want to add together two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$, first we must represent both as starting from the same point: so we change $\overrightarrow{CD}$ to $\overrightarrow{AE} = \overrightarrow{CD}$. Then, there exists a K such that $\overrightarrow{EK} = \overrightarrow{AB}$, and

$$\overrightarrow{AK} := \overrightarrow{AB} + \overrightarrow{CD} \tag{1.2.1}$$

---

[1]It is easy enough to check the three conditions.

Multiplication by a scalar   We use a real number as a scalar because of the continuum hypothesis.[2]

For any $\alpha \in \mathbb{R}$ and any vector , we define $\alpha\overrightarrow{AB}$ as a vector $\overrightarrow{AC}$ such that $\left|\overrightarrow{AC}\right| = |\alpha|\left|\overrightarrow{AB}\right|$ and C is on the same side of A as B if $\alpha > 0$ and on the opposite side if $\alpha < 0$.

Defining the zero vector $\mathbf{0} = \overrightarrow{AA}$ we immediately see that:

- $0\mathbf{v} = \mathbf{0}$

- $\alpha\mathbf{0} = \mathbf{0}$

Properties of vector operations   The following hold:

1. addition is commutative: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$

2. addition is associative: $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$, since they define a parallelepiped

3. there exists a zero vector: $\exists \mathbf{0} : \mathbf{a} + \mathbf{0} = \mathbf{a}$

4. there exists an opposite for every vector: $\exists(-\mathbf{a}) : \mathbf{a} + (-\mathbf{a}) = \mathbf{0}$, also, we notice that $(-\mathbf{a}) = -1\mathbf{a}$

5. scalar multiplication is distributive: $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$

6. scalar addition distributes as vector addition over scalar multiplication: $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{b}$

7. scalar multiplication is associative: $(\alpha\beta)\mathbf{a} = \alpha(\beta\mathbf{a})$

8. $1\mathbf{a} = \mathbf{a}$

9. $0\mathbf{a} = \mathbf{0}$

1.3   Reference frames

For an n-dimensional reference frame we need n vectors (the "basis vectors") which are neither parallel to one another nor lying in the same (hyper)plane, so that we can express any vector we want through a linear combination of them; they, however, need not be orthogonal. We will call our three-dimensional basis vectors $\hat{x}$, $\hat{y}$ and $\hat{z}$. Any vector $\mathbf{p}$ can then be seen as:

$$\mathbf{p} = \alpha\hat{x} + \beta\hat{y} + \gamma\hat{z} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \tag{1.3.1}$$

---

[2]CH states that there is no set whose cardinality is between that of the integers and that of the reals.

We can verify through the identities that, as long as we work in a consistent reference system, we can express vector addition and scalar mutiplication through the coordinates as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}, \qquad c \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \end{pmatrix} \qquad (1.3.2)$$

1.4   Linear dependence

Parallelism   To write the fact that $\mathbf{a} \parallel \mathbf{b}$, we could say that $\exists \lambda \in \mathbb{R} : \mathbf{a} = \lambda \mathbf{b}$, but this fails to account for one of the vectors being $\mathbf{0}$. A better formula is this one:

$$\exists \alpha, \beta \in \mathbb{R} : \neg(\alpha = \beta = 0) : \alpha \mathbf{a} + \beta \mathbf{b} = \mathbf{0} \qquad (1.4.1)$$

Linearly dependent vectors   With a similar argument, we say that n vectors $\mathbf{x}_i$, where $0 < i \leq n$ are in the same $n - 1$-plane if

$$\forall i \in \mathbb{N} : 0 < i \leq n : \exists \alpha_i \in \mathbb{R} : \exists \alpha_i \neq 0 : \sum_{i=1}^{n} \alpha_i \mathbf{x}_i = \mathbf{0} \qquad (1.4.2)$$

and we call this a *linear combination* of the vectors $\mathbf{x}_i$.

One vector is linearly dependend only if it is $\mathbf{0}$.

If there is no such linear combination (with at least one nonzero index) yielding $\mathbf{0}$, the vectors are said to be *linearly independent*; otherwise they are *linearly dependent*. If one of the vectors is the zero vector, all of them are automatically linearly dependent.

Definition 1.4.1. A *subspace* is a subset S of the geometrical vector space such that:

1. $\mathbf{0} \in S$;

2. $\forall \mathbf{a}, \mathbf{b} \in S : \mathbf{a} + \mathbf{b} \in S$;

3. $\forall \mathbf{a} \in S : \forall \lambda \in \mathbb{R} : \lambda \mathbf{a} \in S$;

In n-dimensional space, the linear combinations of up to $n - 1$ vectors always generate subspaces.

To represent the subspaces generated by the vectors $\mathbf{v}_i$, we write $\langle v_1; v_2; \ldots; v_n \rangle$. To check whether the number of generators is minimal, we just need to see if they are linearly independent. If they are, they form a *basis* of S. The number of basis vectors is called the dimension of S.

The canonical basis vectors for $\mathbb{R}^3$ are

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.4.3}$$

The representation of any $\mathbf{v}$ as a combination of the basis vectors is unique.

Proof. Suppose there were two sequences of coefficients x_i and α_i, both of which represent $\mathbf{v}$. Then,

$$\mathbf{v} = \sum_i x_i \mathbf{e}_i = \sum_i \alpha_i \mathbf{e}_i \implies \mathbf{0} = \sum_i (x_i - \alpha_i)\mathbf{e}_i \tag{1.4.4}$$

but the basis vectors are linearly independent, so $\forall i : x_i = \alpha_i$. $\qquad\square$

To find a basis for a given subspace (expressed with an equation or set of equations) we write the vectors in the subspace with as few coordinates as we can, and then separate the variables.

1.5 Scalar product

Given two vectors $\mathbf{v}$ and $\mathbf{w}$, their scalar product is defined as:

$$\mathbf{v} \cdot \mathbf{w} = \sum_i v_i w_i \tag{1.5.1}$$

Its properties are:

1. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$;

2. $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$;

3. $(\alpha\mathbf{v}) \cdot \mathbf{w} = \alpha(\mathbf{v} \cdot \mathbf{w})$;

4. $\mathbf{v} \cdot \mathbf{v} \geq 0$, and $\mathbf{v} \cdot \mathbf{v} \iff \mathbf{v} = \mathbf{0}$

Definition 1.5.1. We define the *norm* of $\mathbf{v}$ as $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Properties:

1. By definition $\|\mathbf{v}\| \geq 0$, and $\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0}$.

2. $\|\alpha\mathbf{v}\| = |\alpha|\|\mathbf{v}\|$;

3. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$

Property 3 can be proven using the Cauchy-Schwarz inequality:

Theorem 1.5.1 (Cauchy-Schwarz). $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \cdot \|\mathbf{w}\|$

Proof. For a $t \in \mathbb{R}$, we take $(\mathbf{v} + t\mathbf{w})^2 = (\mathbf{v} \cdot \mathbf{v}) + 2t(\mathbf{v} \cdot \mathbf{w}) + t^2(\mathbf{w} \cdot \mathbf{w})$. By definition, $(\mathbf{v} + t\mathbf{w})^2 \geq 0$.

Then, in the variable t, it must be that $\Delta < 0$, or

$$4(\mathbf{v} \cdot \mathbf{w})^2 - 4(\mathbf{w} \cdot \mathbf{w})(\mathbf{v} \cdot \mathbf{v}) < 0$$
$$\mathbf{v} \cdot \mathbf{w} < \|\mathbf{w}\| \cdot \|\mathbf{v}\| \tag{1.5.2}$$

$\square$

Proof of property 3. Take $(\|\mathbf{v} + \mathbf{w}\|)^2 = (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w})$. This is equal to $\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\,\mathbf{v} \cdot \mathbf{w}$ which, by (1.5.2), is less than or equal to $\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$

$\square$

Since

$$\forall \mathbf{v}, \mathbf{w} \neq \mathbf{0} : \quad -1 \leq \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1 \tag{1.5.3}$$

there exists a $\theta \in [0; \pi]$ such that $\cos\theta = (\mathbf{v} \cdot \mathbf{w})/(\|\mathbf{v}\|\|\mathbf{w}\|)$. We will show that θ is the angle between the vectors.

Proof. Take h to be the height (relative to the side $\mathbf{v}$) of the triangle defined by $\mathbf{v}$ and $\mathbf{w}$, and l to be $\|\mathbf{v}\|$ minus the projection of $\mathbf{w}$ on $\mathbf{v}$.

Then $h = \|\mathbf{w}\| \sin\theta$ and $\|\mathbf{v}\| - l = \|\mathbf{w}\| \cos\theta$. So

$$h^2 + l^2 = (\|\mathbf{v} - \mathbf{w}\|)^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\| \cos\theta \tag{1.5.4}$$

but it is also the case that

$$(\|\mathbf{v} - \mathbf{w}\|)^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2(\mathbf{v} \cdot \mathbf{w}) \tag{1.5.5}$$

$\square$

Angle between vectors

$$\frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} = \cos(\theta) \tag{1.5.6}$$

where θ is the angle between the vectors, always taken to be positive.

Orthogonality   From this follows that $\mathbf{v} \cdot \mathbf{w} = 0 \iff \mathbf{v} \perp \mathbf{w}$.

1.6 Vector product

It is defined as an operation on two vectors in $\mathbb{R}^3$, with the informal determinant

$$\mathbf{v} \times \mathbf{w} = \det \begin{pmatrix} \hat{i} & \hat{j} & \hat{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \tag{1.6.1}$$

This weird way of writing comes from the fact that, if we define the following function:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \det \begin{pmatrix} x & v_1 & w_1 \\ y & v_2 & w_2 \\ z & v_3 & w_3 \end{pmatrix} \tag{1.6.2}$$

we can check that it is linear, thus there is some vector $\mathbf{p}$ for which

$$\mathbf{p} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \det \begin{pmatrix} x & v_1 & w_1 \\ y & v_2 & w_2 \\ z & v_3 & w_3 \end{pmatrix} \tag{1.6.3}$$

and so it becomes clear that, in fact, $\mathbf{p} = \mathbf{v} \times \mathbf{w}$ is the vector we can scalar-multiply $\mathbf{x}$ by to get the volume of the parallelepiped defined by $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{x}$.

The properties of the vector product are:

1. $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$;

2. $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$;

3. $(\alpha \mathbf{v}) \times \mathbf{w} = \alpha(\mathbf{v} \times \mathbf{w}) = \mathbf{v} \times (\alpha \mathbf{w})$;

4. $\mathbf{v} \times \mathbf{w} \iff \mathbf{v} \parallel \mathbf{w}$;

5. $\mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{v} \times \mathbf{w})$;

6. $\|\mathbf{v} \times \mathbf{w}\|^2 = \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 - (\mathbf{v} \cdot \mathbf{w})^2$ (Lagrange's identity);

The first three are trivial, although the calculations get cumbersome.

Proof of point 4. Suppose that $\mathbf{v} \neq \mathbf{0}$ (otherwise the proposition would be automatically true). So WLOG suppose that $v_1 \neq 0$. Then $x_1 y_3 = x_3 y_1$, and we can divide by x_1:

$$w_3 = \left(\frac{w_1}{v_1}\right) w_3 \tag{1.6.4}$$

And also

$$w_2 = \left(\frac{w_1}{v_1}\right) v_2 \quad \text{and} \quad w_1 = \left(\frac{w_1}{v_1}\right) v_1 \tag{1.6.5}$$

So $\mathbf{w}$ is a multiple of $\mathbf{v}$. $\qquad \square$

The proofs of points 5 and 6 follow immediately from expanding the calculations.

A neat consequence of point 6 is the fact that, dividing everything by the square norms, and remembering equation (1.5.6) we get:

$$\frac{\|\mathbf{v} \times \mathbf{w}\|^2}{\|\mathbf{v}\|^2\|\mathbf{w}\|^2} + \cos^2(\theta) = 1 \implies \frac{\|\mathbf{v} \times \mathbf{w}\|}{\|\mathbf{v}\|\|\mathbf{w}\|} = \sin(\theta) \tag{1.6.6}$$

where $\theta \in [0; \pi[$. So the norm of the vector product is the area of the parallelogram.

1.7 Mixed product

Given three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, the mixed product is defined as $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) \in \mathbb{R}$.

We can prove that $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = 0 \iff$ the three vectors are linearly dependent, i. e. $\mathbf{u} \in \langle \mathbf{v}; \mathbf{w} \rangle$.

Proof. First we show the leftward implication: $(\alpha \mathbf{v} + \beta \mathbf{w}) \cdot (\mathbf{v} \times \mathbf{w}) = 0$.

Then, for the rightward implication: we use the fact that if $\mathbf{v}$ and $\mathbf{w}$ are linearly independent then $\forall \mathbf{a} \in \mathbb{R}^3$, we can write $\mathbf{a}$ as a linear combination of $\mathbf{v}, \mathbf{w}$ and $\mathbf{v} \times \mathbf{w}$,[3] which are linearly independent. [4]

Then $\mathbf{u} = a\mathbf{v} + b\mathbf{w} + c(\mathbf{v} \times \mathbf{w})$.

If we scalar-multiply everything by $\mathbf{v} \times \mathbf{w}$ we get:

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = c \|\mathbf{v} \times \mathbf{w}\|^2 \implies c = 0 \tag{1.7.2}$$

So $\mathbf{u} = a\mathbf{v} + b\mathbf{w}$.

$\square$

The mixed product of three vectors is the volume of the parallelepiped they define:

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \|\mathbf{u}\|\|\mathbf{v} \times \mathbf{w}\| \cos \alpha = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \tag{1.7.3}$$

where α is the angle between $\mathbf{u}$ and $\mathbf{v} \times \mathbf{w}$.

---

[3]This will be proved later.

[4]

Proof. We will prove this by showing that the only linear combination of these three vectors yielding $\mathbf{0}$ is one where the three coefficients are all zero: take

$$\alpha \mathbf{v} + \beta \mathbf{w} + \gamma(\mathbf{v} \times \mathbf{w}) = 0 \tag{1.7.1}$$

and scalar-multiply everything by $\mathbf{v} \times \mathbf{w}$. We get $\gamma \|\mathbf{v} \times \mathbf{w}\|^2 = 0$, so $\gamma = 0$.

Then it must be that $\alpha = \beta = 0$.

$\square$

1.8 Perpendicular set

Given some vectors $\mathbf{v}_i$, we define the perpendicular set as

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}^\perp := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \forall i : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \mathbf{v}_i = 0 \right\} \tag{1.8.1}$$

$S^\perp$ is always a subspace:

Proof.   1. $\mathbf{0} \in S^\perp$;

2. $\mathbf{w}_1 \in S^\perp, \mathbf{w}_2 \in S^\perp \implies \forall i : (\mathbf{w}_1 + \mathbf{w}_2) \cdot \mathbf{v}_i = 0$ because of the distributive property;

3. $\forall i : (\alpha \mathbf{w}_1) \cdot \mathbf{v}_i = \alpha(\mathbf{w}_1 \cdot \mathbf{v}_i) = 0$.

$\square$

In $\mathbb{R}^n$, the perpendicular set of a subspace of dimension d has dimension $n - d$.

Chapter 2

Spaces

2.1 Affine spaces

They resolve the ambiguity between vectors and points. $\mathbb{A}^3(\mathbb{R})$ is affine to $\mathbb{R}^3$.

Definition 2.1.1. An affine space is the set of points on which the vector space $\mathbb{R}^3$ acts.

$$f : \mathbb{A}^3(\mathbb{R}) \times \mathbb{R} \to \mathbb{A}^3(\mathbb{R}) \tag{2.1.1}$$

So to a couple (point, vector) we associate a vector.

Properties

1. Transitive: $\forall (p, q) \in \mathbb{A}^3(\mathbb{R}) : \exists! \mathbf{v} \in \mathbb{R}^3$ such that $q = p + \mathbf{v}$ (or, $\mathbf{v}$ is uniquely determined by (p, q);

2. Compatibility with the algebraic operations in $\mathbf{R}^3$ (sum, neutral element...);

3. If we choose an $O \in \mathbb{A}^3(\mathbb{R})$ we get $\mathbb{A}^3(\mathbb{R}) \leftrightarrow \mathbb{R}^3$ (a reference frame in the affine space corresponds to a basis in the vector space);

An affine space has more freedom than a vector space, since we can change the reference point.

2.1.1 Linear varieties

They correspond to the subspaces of the vector space.

Definition 2.1.2. They are subsets of $\mathbb{A}^3(\mathbb{R})$ in the form $P_0 + U = \{Q \in \mathbb{A}^3(\mathbb{R}) : Q - P_0 \in U\}$, where P_0 is a point in space and U is a subspace of $\mathbb{R}^3$.

$$P_0 + U = \{q \in \mathbb{A}^3(\mathbb{R}) : q - P_0 \in U\} \tag{2.1.2}$$

If Q_1 and Q_2 belong to $P_0 + U : Q_i = P_0 + U_i$, where $U_i \in U$, then $Q_2 - Q_1 = U_2 - U_1 \in U$ is the vector connecting the points.

Potential linear varieties of $\mathbb{A}^3(\mathbb{R})$

- $P_0 + \{\mathbf{0}\}$ is just P_0;

- $P_0 + \mathbf{R}^3 = \mathbb{A}^3(\mathbb{R})$, is all of space;

- if $U = \langle \mathbf{u} \rangle$, with $\mathbf{u} \neq 0$, the dimension of U is 1, and then $P_0 + U = \{P_0 + \lambda\mathbf{u}\}$;

- if $U = \langle \mathbf{u}_1; \mathbf{u}_2 \rangle$, the dimension of U is 2, and $P_0 + U = \{P_0 + \lambda\mathbf{u}_1 + \mu\mathbf{u}_2\}$.

The subspace U cannot be changed for another: it is called "giacitura".
We can do an analogous operation in $\mathbf{R}^2$.

Parametric equations of linear varieties

Describing a linear variety in $\mathbb{A}^3(\mathbb{R})$ as $P + \langle \mathbf{x} \rangle$ gives us parametric equations. We just apply the definition; for example, in the case of a linear variety of dimension 1:

$$r : P_1 + \langle P_2 - P_1 \rangle = P_1 + \alpha \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix} \tag{2.1.3}$$

Cartesian equations of linear varieties

We can find them by taking the orthogonal of the orthogonal subspace, and substituting $(x_i - P_i)$ for each coordinate.

2.2 Euclidean spaces

If we consider the structure "vector subspace" equipped with the notion of scalar product, we get an Euclidean space, denoted $\mathbb{E}^3(\mathbb{R})$. This also means we have a notion of distance: $\mathrm{d}(P,Q) = \left\| \overrightarrow{PQ} \right\|$, where $\overrightarrow{PQ}$ is $Q - P \in \mathbb{R}^3$.

Given $S, T \subseteq \mathbb{E}^3(\mathbb{R})$, we define

$$\mathrm{d}(S,T) := \inf \left\{ \mathrm{d}(P,Q), P \in S, Q \in T \right\} \subseteq [0, +\infty[\tag{2.2.1}$$

Note that this does not abide the usual axioms for a distance function: the distance between two different sets can be 0.

2.2.1 Distance between linear varieties

Theorem 2.2.1. If S and T are linear varieties, $\exists p_0 \in S, q_0 \in T : \mathrm{d}(p_0, q_0) = \mathrm{d}(S,T)$.

Proof. The linear varieties are $S = p_1 + U$ and $T = q_1 + W$, where $U, W \subseteq \mathbb{R}^3$.

It will be shown later that there exist $\mathbf{u} \in U, \mathbf{w} \in W, \mathbf{n} \in \{U \cup W\}^\perp$ such that $p_1 q_1 = \mathbf{u} + \mathbf{w} + \mathbf{n}$, but we will use the result straight away.

Take the points $p_0 = p_1 + \mathbf{u}$, $q_0 = q_1 + \mathbf{w}$. Then

$$\mathrm{d}(p_0, q_0) = \|q_1 - \mathbf{w} - p_1 - \mathbf{u}\| = \mathbf{n} \tag{2.2.2}$$

Now to show that (p_0, q_0) is minimal: take $(p', q') = (p_0 + \mathbf{u}', q_0 + \mathbf{w}')$.

$$\mathrm{d}(p', q')^2 = \|q' - p'\|^2 = \|q_0 - p_0 + \mathbf{w}' - \mathbf{u}'\|^2 = \|\mathbf{n} + \mathbf{w}' - \mathbf{u}'\|^2 \tag{2.2.3}$$

but we also know that

$$\mathbf{n} \cdot (\mathbf{w} - \mathbf{u}) = 0 \tag{2.2.4}$$

since $\mathbf{n} \in \{U \cup W\}^\perp$, so

$$\|\mathbf{n} + \mathbf{w}' - \mathbf{u}'\|^2 = \|\mathbf{n}\|^2 + \|\mathbf{w} - \mathbf{u}\|^2 \geq \|\mathbf{n}\| \tag{2.2.5}$$

as desired. We have equality only if $\|\mathbf{w} - \mathbf{u}\|^2 = 0$; if U and W have nonzero vectors in commons (i. e. they are parallel), there are (p', q') such that $\mathrm{d}(p', q') = \mathrm{d}(p_0, q_0)$. $\qquad\square$

We can find intersections between linear varieties using cartesian or parametric equations, but it is often best to mix: use the cartesian equation for one variety and the parametric for the other.

2.3   Plane bundles

In $\mathbb{A}^3(\mathbb{R})$, plane bundles are a useful construct. We start with two planes:

$$\pi : ax + by + cz + d = 0 \tag{2.3.1}$$
$$\pi' : a'x + b'y + c'z + d' = 0 \tag{2.3.2}$$

Their plane bundle is made as such:

$$\lambda(ax + by + cz + d) + \mu(a'x + b'y + c'z + d') = 0 \tag{2.3.3}$$

for some λ and $\mu \in \mathbb{R}$, such that $(\lambda, \mu) \neq (0, 0)$. We immediately notice that for any $\alpha \in \mathbb{R}^*$, (λ, μ) and $(\alpha\lambda, \alpha\mu)$ correspond to the same plane.

Intersecting planes   If $\pi \cup \pi' = r$, a line, then $P \in r \implies P \in \zeta$ for any ζ in the bundle, because the coordinates of P solve both of the planes' equations. So this is the *proper bundle* of axis r.

Parallel planes If $\pi : P + \langle \mathbf{v}, \mathbf{u} \rangle$ and $\pi' : P' + \langle \mathbf{v}, \mathbf{u} \rangle$, then it must be that $a = a'$, $b = b'$ and $c = c'$, since the perpendicular vector is the same. So the plane equation looks like:

$$(\lambda + \mu)(ax + by + cz) + \lambda d + \mu d' = 0 \tag{2.3.4}$$

Which is parallel to the two starting planes, so the bundle consists of all the parallel planes to the starting ones.

2.3.1 Plane-point distance

We want to find the distance between the point P and the plane $T = Q + V$. We call the projection of P onto the plane Q_0.

The cartesian equation of T is $ax + by + cz + d = 0$. Then the perpendicular subspace $V^\perp$ is

$$V^\perp = \langle \begin{pmatrix} a \\ b \\ c \end{pmatrix} \rangle = \langle \mathbf{n} \rangle \tag{2.3.5}$$

$P_0 Q_0$ is proportional to $\mathbf{n}$, $Q Q_0$ is perpendicular to it. Now, interpret QP as $QQ_0 + Q_0 P$ and take $QP \cdot \mathbf{n} = QQ_0 \cdot \mathbf{n} + Q_0 P \cdot \mathbf{n} = \alpha \|\mathbf{n}\|^2$. Since $\mathbf{n}$ is arbitrary, we choose $\|\mathbf{n}\| = 1$, so $\|Q_0 P\| = |\alpha|$, so we can calculate α with the scalar product $QP \cdot \mathbf{n}$, without knowing anything about Q_0.

With our choice of $\mathbf{n}$, the perpendicular subspace became

$$V^\perp = \left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix} \frac{1}{\sqrt{a^2 + b^2 + c^2}} \right\rangle \tag{2.3.6}$$

Now we calculate

$$\mathrm{d}(P, T) = |\alpha| = \left| QP \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} \frac{1}{\sqrt{a^2 + b^2 + c^2}} \right| \tag{2.3.7}$$

The vector QP is

$$\begin{pmatrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{pmatrix} \tag{2.3.8}$$

So

$$\mathrm{d}(P, T) = \left| \frac{ax_0 + by_0 + cz_0 - ax_1 - bx_1 - cx_1}{\sqrt{a^2 + b^2 + c^2}} \right| \tag{2.3.9}$$

But $Q \in T$, so $ax_1 + bx_1 + cx_1 + d = 0$, so $d = -ax_1 - bx_1 - cx_1$.

$$d(P,T) = \left| \frac{ax_0 + by_0 + cz_0 + d}{\sqrt{a^2 + b^2 + c^2}} \right| \tag{2.3.10}$$

which is the final formula.

2.3.2 Plane-plane or plane-line distance

If they intersect, the distance is 0. Otherwise, we just pick a point from one plane, and then calculate its distance from the other, since they are parallel. The same holds for a plane and a line.

2.3.3 Point-line distance

If $P \in r$, the distance is 0. Otherwise, $d(P,r) = d(P,Q_0)$, where Q_0 is the projection of P onto the line. If we take a plane σ such that $\sigma \perp r$ and $P \in \sigma$, then $Q_0 = \sigma \cap r$.

If we do not care about the position of Q_0, we can just take any point $Q \in r$ and $\mathbf{v} \in r$, and calculate:

$$d(P,r) = \frac{\left\| \overrightarrow{QP} \times \mathbf{v} \right\|}{\|v\|} \tag{2.3.11}$$

which is the height of the parallelogram spanned by $\mathbf{v}$ and $\overrightarrow{QP}$.

2.3.4 Line-line distance

If $r \cap s \neq \emptyset$, the distance is 0. If $r \parallel s$, we use the point-line formula.

Mixed product method Otherwise, we can choose $P \in r$ and $Q \in s$, $\mathbf{w}$ in the direction of r and $\mathbf{v}$ in the direction of s. Now, we take the volume of the parallelepiped spanned by $\overrightarrow{QP}$, $\mathbf{w}$ and $\mathbf{v}$ and divide it by the area of its base:

$$h = \frac{\left| \overrightarrow{QP} \cdot (\mathbf{v} \times \mathbf{w}) \right|}{\|\mathbf{v} \times \mathbf{w}\|} \tag{2.3.12}$$

Plane method Take a plane σ such that $s \in \sigma$ and $\sigma \parallel r$. Then, $d(r,s) = d(r,\sigma) = d(P \in r, \sigma)$

Linear variety method We can write $r = P + \langle \mathbf{w} \rangle$ and $s = Q + \langle \mathbf{v} \rangle$. We must choose our P and Q such that $\overrightarrow{PQ} \cdot \mathbf{v} = \overrightarrow{PQ} \cdot \mathbf{w} = 0$. Then $\left\| \overrightarrow{PQ} \right\|$ is the distance.

2.4 Abstract vector spaces

Definition 2.4.1. A vector space over a field $\mathbb{F}$ is a set $V \neq \emptyset$ equipped with:

- addition: $v_1, v_2 \in V \implies v_1 + v_2 \in V$;

- scalar multiplication: $\alpha \in \mathbb{F}, v \in V \implies \alpha v \in V$;

which must satisfy, $\forall a, b, c \in V$ and $\forall \alpha, \beta \in \mathbb{F}$:

- commutativity for addition: $a + b = b + a$;

- associativity for addition: $(a + b) + c = a + (b + c)$;

- existence of a neutral element for addition $\exists \mathbf{0} \in V : a + \mathbf{0} = a$;

- existence of additive inverses: $\exists v \in V : v + a = \mathbf{0}$;

- distributivity of scalar multiplication with respect to addition: $\alpha(a + b) = \alpha a + \alpha b$;

- distrivutivity of scalar addition with respect to scalar multiplication: $(\alpha + \beta)a = \alpha a + \beta a$;

- associativity of scalar multiplication: $\alpha(\beta a) = (\alpha \beta)a$;

- identity element of scalar multiplication: $\exists 1 \in \mathbb{F} : 1a = a$.

2.4.1 Examples

$\mathbb{R}^n$ and the geometric vectors are real vector spaces. $\mathbb{R}[x]$ are the polynomials with real coefficients; addition and scalar multiplication are defined as usual.

If we take $I \subseteq \mathbb{R}$, the set of all $f : I \to \mathbb{R}$ is a vector space. Addition and scalar multiplication can be defined point by point on the image. Subspaces of this vector space are the continuous functions and the differentiable functions.

Real or complex matrices with n rows and m columns are a vector space (denoted by $M_{n \times m}(\mathbb{R})$). Addition and scalar multiplication are defined as: $a_{ij} + b_{ij} = c_{ij}$, $\alpha a_{ij} = c_{ij}$.

2.5 Subspaces

$U \subseteq V$, $U \neq \emptyset$ is a vector subspace if the operations of V, restricted to U, make it a vector space. U is subspace iff it contains all the linear combinations of its vectors.

Generators and bases

Definition 2.5.1. Given the subspace U, the vectors $\mathbf{u}_i$ are *generators* if every vector in U can be written as a linear combination of them.

2.6 Bases

Definition 2.6.1. A basis of a subspace U is a set of linearly independent generators.

For example, the canonical basis of $M_{n \times m}(\mathbb{R})$ is $\mathbf{e}_{ij} = \delta_{ij}$.[1]

Theorem 2.6.1. Given a vector subspace $U \subseteq V$, the vectors $\mathbf{u}_i \in U$ are a basis iff every vector in U can be written in a unique way as a linear combination of $\mathbf{u}_i$.

Proof. Leftward implication: if every vector can be written in a unique way, $\mathbf{u}_i$ are generators, and $\mathbf{0} = \sum 0\mathbf{u}_i$: by the fact that the representation is unique $\mathbf{u}_i$ are linearly independent.

Rightward implication: if $\mathbf{u}_i$ is a basis, take $\mathbf{v} = \sum \alpha_i \mathbf{u}_i = \sum \beta_i \mathbf{u}_i$. Then $\mathbf{v} - \mathbf{v} = \mathbf{0} = \sum (\alpha_i - \beta_i)\mathbf{u}_i$, which implies that $\forall i \in \mathbb{N} \, \alpha_i = \beta_i$, that is, the representation is unique. $\qquad \square$

A set of vectors S can be a set of generators even if it is infinite, as long as any vector is written as a combination of just a finite subset of them. For example, the set $S = \{x^n | \forall n \in \mathbb{N}\}$ is a set of generators for the polynomials.

Elimination

Theorem 2.6.2. If $\mathbf{u}_i, i \in [1, n]$ are linearly dependent generators of U, then $\exists k \in [1, n] : \mathbf{u}_i, i \in [1, n] \smallsetminus \{k\}$ are generators of U.

Proof. In the combination $\mathbf{0} = \sum \alpha_i \mathbf{u}_i$ at least one α_i must be nonzero. Take one such coefficient and divide by it:

$$\frac{\alpha_i \mathbf{u}_i}{\alpha_i} = -\frac{1}{\alpha_i} \sum_{\substack{k=1 \\ k \neq i}}^{n} \alpha_k \mathbf{u}_k \qquad (2.6.1)$$

So we have written $\mathbf{u}_i$ as a linear combination of the other basis vectors. Now, we can express any vector in U without it. $\qquad \square$

In a finitely generated space (a space with finitely many generators), we can always find a basis applying this theorem repeatedly.

Theorem 2.6.3. If $\mathbf{u}_i$ are linearly independent vectors in a subspace U, given a vector $\mathbf{v} \in U$, the vectors $\{\mathbf{u}_i, \mathbf{v}\}$ are linearly independent iff $\mathbf{v} \notin \langle \mathbf{u}_i \rangle$.

Proof. Leftward implication: if $\mathbf{v} \in \langle \mathbf{u}_i \rangle$, then $\mathbf{v} - \sum \alpha_i \mathbf{u}_i = \mathbf{0}$, so they are linearly dependent.

Rightward implication: if there exist coefficients $\{\beta, \alpha_i\}$ such that $\beta \mathbf{v} + \sum \alpha_i \mathbf{u}_i = 0$, then $\beta \neq 0$ since $\mathbf{u}_i$ are linearly independent, then we can divide by β and:

$$\mathbf{v} = -\frac{1}{\beta} \sum_i \alpha_i \in \langle \mathbf{u}_i \rangle \qquad (2.6.2)$$

$\square$

---

[1] δ_{ij} is the Dirac delta, which is 1 only if the two indices are equal, and 0 otherwise.

Constructing a basis Given a finitely generated subspace $U \neq \mathbf{0}$, we can construct a basis for it: choose a vector $\mathbf{u}_1$, which is independent. If $U = \langle \mathbf{u}_1 \rangle$, we are done. If not, $\exists \mathbf{u}_2 \in U, \mathbf{u}_2 \notin \langle \mathbf{u}_1 \rangle$, so they are independent. We can go on as long as the generated subspace does not equal U.

Theorem 2.6.4 (Steinitz exchange lemma). Given the linearly independent vectors $\{\mathbf{u}_i\}_{i \in [1,k]} \in V$ and the vectors $\langle \mathbf{v}_i \rangle_{i \in [1,p]} = V$, then $k \leq p$.

Proof. $\mathbf{u}_1 = \sum \alpha_i \mathbf{v}_i$. $\forall i : \mathbf{u}_i \neq \mathbf{0}$, so at least one α_i is not 0. WLOG, we can take $\alpha_1 \neq 0$. Now,

$$\mathbf{v}_1 = \frac{1}{\alpha_1} \mathbf{u}_1 - \sum_{i=2}^{p} \frac{\alpha_i \mathbf{v}_i}{\alpha_1} \tag{2.6.3}$$

so $\langle \mathbf{u}_1, \mathbf{v}_i \rangle = V$. We have substituted one of the generators: we wish to substitute k generators, but we must watch out for an issue: in

$$\mathbf{u}_2 = \beta_1 \mathbf{u}_1 \sum_{i=2}^{p} \beta_i \mathbf{v}_i \tag{2.6.4}$$

one of the β_i must be $\neq 0$, but the proof would not work if it were just β_1.

This is easily fixed: if $\forall i \in [2,p] : \beta_i = 0$, then we would have $\mathbf{u}_2 = \beta_1 \mathbf{u}_1$, which is excluded since they are independent. So we suppose, again WLOG, that $\beta_2 \neq 0$, and we can continue:

$$\mathbf{v}_2 = -\frac{1}{\beta_2} \left(\beta_1 \mathbf{u}_1 - \mathbf{u}_2 + \sum_{i=3}^{p} \beta_i \mathbf{v}_i \right) \tag{2.6.5}$$

This way, we can substitute k of the p generators and we will still have generators. So, $k \leq p$. $\qquad\square$

Corollary 2.6.4.1. All the bases of a subspace have the same number of elements, we call this the *dimension* of the subspace.

Proof. Given n vectors in a basis and k in another basis, we use alternatively the fact that they are generators and independent to see that $k \leq n$ and $n \leq k$. $\quad\square$

Corollary 2.6.4.2. If $U \subseteq W \subseteq V$ and U, W are subspaces, then if $\dim U = \dim W$ then $U = W$.

Proof. Given the basis vectors of U, $\{\mathbf{u}_i\}_{i \in [1,n]}$, $\{\mathbf{u}_i, \mathbf{w}\}$ are linearly independent iff $\mathbf{w} \notin U$, so if $\exists \mathbf{w} \in W : \mathbf{w} \notin U$, then there would be $n+1$ linearly independent vectors in W, which is impossible, so $\forall \mathbf{w} \in W, \mathbf{w} \in U$. $\qquad\square$

2.7 Operations on subspaces

2.7.1 Intersection

Theorem 2.7.1. Given a vector space V and the subspaces $W \subseteq V$ and $U \subseteq V$, then $W \cap U$ is a subspace.

Proof. $\mathbf{0} \in W \cap U$, and given $\mathbf{v}_1, \mathbf{v}_2 \in W \cap U$, $\mathbf{v}_1 + \mathbf{v}_2 \in U$ and $\mathbf{v}_1 + \mathbf{v}_2 \in W$, so $\mathbf{v}_1 + \mathbf{v}_2 \in W \cap U$. The same holds for scalar multiplication. $\square$

2.7.2 Sum

The union of subspaces is generally not a subspace, and we only need one example to show this. So, we define:

Definition 2.7.1. Given the subspaces U and W, their *sum space* is defined as $U + W := \langle U \cup W \rangle$.[2] This is the smallest subspace containing both U and W.

Definition 2.7.2. We say that the sum of two subspaces is *direct* iff a vector in it is written uniquely as a linear combination of vectors of each starting subspace; this is denoted like: $S \oplus T$.

For example, $\mathbb{R}^2 = \langle \mathbf{e}_1 \rangle \oplus \langle \mathbf{e}_2 \rangle$.

Theorem 2.7.2. The sum of two subspaces is direct iff their intersection is exactly $\mathbf{0}$.

Proof. Leftward implication: given $\mathbf{v} \in S \oplus T$, take $\mathbf{v} \in S \cap T$. Then $\mathbf{v} = (\mathbf{v}_S + \mathbf{0}_T) = (\mathbf{0}_S + \mathbf{v}_T)$. By the uniqueness of the representation, $\mathbf{v} = \mathbf{0}$.

Rightward implication: given that $S \cap T = \{\mathbf{0}\}$, take $\mathbf{v} \in S + T$, and suppose that $\mathbf{v} = \mathbf{s}_1 + \mathbf{t}_1 = \mathbf{s}_2 + \mathbf{t}_2$.

Now, $\mathbf{s}_1 - \mathbf{s}_2 \in S$ and $\mathbf{t}_1 - \mathbf{t}_2 \in T$, but $\mathbf{s}_1 - \mathbf{s}_2 = \mathbf{t}_1 - \mathbf{t}_2$ so they belong to $S \cap T$, which implies $\mathbf{s}_1 - \mathbf{s}_2 = \mathbf{t}_1 - \mathbf{t}_2 = \mathbf{0}$. $\square$

It is clear that

$$\dim U + \dim W \geq \dim(U + W) \geq \max\{\dim U, \dim W\} \tag{2.7.1}$$

Theorem 2.7.3 (Grassmann's formula).

$$\dim(S + T) = \dim S + \dim T - \dim(S \cap T) \tag{2.7.2}$$

A sketch of the proof is as follows:

We construct a basis for $S \cap T$, then we add vectors until we get to a basis of S, and we do the same for T. All these vectors must then be a basis for $S + T$ (it must be proven that they are independent generators), but we counted the ones from $S \cap T$ twice.

---

[2]Obviously the set of genenerators given for this subspace are not a basis.

Chapter 3

Matrices

3.1 Elementary operations

Given the set of generators $U = \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \rangle$, we can perform the following operations without changing U:

1. switching places: $U = \langle \mathbf{u}_2, \mathbf{u}_1, \ldots, \mathbf{u}_2 \rangle$;

2. adding to any generator any other generator scaled by some amount: $U = \langle (\mathbf{u}_1 + \alpha \mathbf{u}_2), \mathbf{u}_2, \ldots, \mathbf{u}_n \rangle$

3. scaling any generator: $U = \langle \alpha \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \rangle$

We can perform these to the rows of a matrix.

Definition 3.1.1. A matrix is said to be in *row echelon form* if:

1. all the rows composed by just zeroes are on the bottom of the matrix;

2. the pivot[1] of any row is strictly to the right of the one of the row above it.

Definition 3.1.2. The *rank by rows* of a matrix is the dimension of the space spanned by its rows.

If a matrix is in row echelon form, this is the number of nonzero rows. The rank of a matrix is invariant under elementary operations.

3.2 Matrix multiplication

This is the first operation we introduce that distinguishes matrices from any other vector space.

If we multiply $\mathbf{A} \in M_{m \times n}(\mathbb{F})$ and $\mathbf{B} \in M_{p \times n}(\mathbb{F})$ we get $\mathbf{AB} = \mathbf{C}$ such that:

---

[1]The first nonzero coefficient in a row of a matrix is called *pivot*.

$$\mathbf{C}_{ij} = \sum_{k=1}^{m} \mathbf{A}_{ik}\mathbf{B}_{kj} \tag{3.2.1}$$

It has all the properties we would expect, with respect to both addition and scalar multiplication: associativity, distributivity, but it is not commutative. The identity matrix $I_n \in M_{n \times n}(\mathbb{F})$ is defined so that $I_{ij} = \delta_{ij}$, and its columns (or rows) are the canonical basis of $\mathbb{R}^n$.

3.2.1 Elementary operations through matrix multiplication

We can express all the elementary operations through multiplying to the left by some matrix, which can be obtained by performing the operations on the identity matrix.

3.2.2 Linear applications

An application is linear if it preserves the operations on the spaces:

$$f : V \to W \text{ linear } \iff \forall \mathbf{v}_1, \mathbf{v}_2 \in V : f(\mathbf{v}_1) + f(\mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2) \tag{3.2.2}$$
$$\forall \alpha \in \mathbb{R} : f(\alpha\mathbf{v}_1) = \alpha f(\mathbf{v}_1) \tag{3.2.3}$$

A linear function $f : \mathbb{R} \to \mathbb{R}$ must satisfy $f(x) = f(1x) = x f(1)$, so it is necessarily of the form $f(x) = \alpha x$ for some $\alpha \in \mathbb{R}$.

By the same reasoning, a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is uniquely determined by the vectors onto which the basis vectors land, and is described by a matrix $\mathbf{A} \in M_{m \times n}(\mathbb{R})$: $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$. The columns of $\mathbf{A}$ are the vectors onto which the basis vectors land.

Theorem 3.2.1. Given a linear function $f : V \to W$, it preserves subspaces:

1. if $S \subseteq V$ is a subspace, $f(S)$ is a subspace in W;

2. if $T \subseteq W$ is a subspace, $f^{-1}(T) = \{\mathbf{v} \in V : f(\mathbf{v}) \in T\}$ is a subspace in V.

Proof. The proof follows trivially from the properties of linear functions. $\square$

We find that the image of the function $f(V)$ is a subspace of W.

Definition 3.2.1.

$$\ker f = f^{-1}(\mathbf{0}) = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\} \tag{3.2.4}$$

Theorem 3.2.2. A linear function $f : V \to W$ is surjective iff $f(V) = W$ and injective iff $\ker f = \{\mathbf{0}\}$. [2]

---

[2]Note that this is only the case for *linear* functions.

Proof. The surjective part is trivial.

Leftward implication: if f is injective, $\mathbf{v} \in \ker f$, $\mathbf{v} = f(\mathbf{0}) \implies \mathbf{v} = \mathbf{v} = \mathbf{0}$ since $f(\mathbf{0}) = \mathbf{0}$.

Rightward implication: if $\ker f = \{\mathbf{0}\}$, take $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Then

$$f(\mathbf{v}_1) - f(\mathbf{v}_2) = f(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0} \implies \mathbf{v}_1 - \mathbf{v}_2 \in \ker f \implies \mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0} \quad (3.2.5)$$

$\square$

If $V = \langle \mathbf{v}_i \rangle$, and $f : V \to W$ is linear, then $f(V) = \langle f(\mathbf{v}_i) \rangle$.
So $\dim(f(V)) \leq \dim V$, since there will be $\dim V$ generators.

3.2.3 Dimensions

Theorem 3.2.3 (Dimension theorem). Given a linear function $f : V \to W$, if V is finitely generated, then

$$\dim V = \dim f(V) + \dim \ker f \quad (3.2.6)$$

Proof. We construct a basis for V from a basis of $\ker f$ and one of $f(V)$.

If $\{\mathbf{u}_i\}$ is a basis for $\ker f$ and $\{\mathbf{w}_i\}$ is a basis for $f(V)$, we take the vectors $\{\mathbf{v}_i\}$ such that $\forall i : f(\mathbf{v}_i) = \mathbf{w}_i$. Then $\{\mathbf{v}_i, \mathbf{u}_i\}$ is a basis for V:

Generators: take any $\mathbf{x} \in V$, then $f(\mathbf{x}) \in f(V)$, so $f(\mathbf{x}) = \sum \beta_i \mathbf{w}_i$. Therefore, $f(\mathbf{x} - \sum \beta_i \mathbf{v}) = f(\mathbf{x}) - \sum \beta_i f(\mathbf{v}_i) = \mathbf{0}$.

So, $\mathbf{x} - \sum \beta_i \mathbf{v}_i \in \ker f \implies \mathbf{x} - \sum \beta_i \mathbf{v}_i = \sum \alpha_i \mathbf{u}_i$, which gives us the final formula:

$$\mathbf{x} = \sum_i \beta_i \mathbf{v}_i + \sum_i \alpha_i \mathbf{u}_i \quad (3.2.7)$$

Independent: we take

$$\sum_i \beta_i \mathbf{v}_i + \sum_i \alpha_i \mathbf{u}_i = \mathbf{0} \quad (3.2.8)$$

and apply f: the αs must go to 0 since their combination is in the kernel, and the βs must go to 0 since the basis vectors for V are independent. $\square$

3.3 Linear applications through matrices

Given a matrix $\mathbf{A} \in M_{m \times n}(\mathbb{R})$, we define $L_A : \mathbb{R}^m \to \mathbb{R}^n$ as: $\mathbf{x} \to \mathbf{A}\mathbf{x}$. The image of L_A is the subspace generatd by the columns of $\mathbf{A}$.

Definition 3.3.1. The rank by columns of $\mathbf{A}$ is the dimension of the subspace generated by its columns, and the same goes for the rows: $\operatorname{rank}_C A = \dim \operatorname{im}(L_A)$.

The kernel of L_A is just the set of solutions to $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Theorem 3.3.1.

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^T \tag{3.3.1}$$

Proof. ??? Abate-DF pag. 101 □

The columns containing the pivots also form a basis for the column space of **A**.

Inverse images What does the set $f^{-1}(\mathbf{w})$ look like? If it is not the kernel it is not a subspace, but, if $f^{-1}(\mathbf{w}) \neq \emptyset$, we can take two vectors $\mathbf{v}_1, \mathbf{v}_2 \in f^{-1}(\mathbf{w})$, and write:

$$\mathbf{0} = \mathbf{w} - \mathbf{w} = f(\mathbf{v}_1) - f(\mathbf{v}_2) \implies \mathbf{v}_1 - \mathbf{v}_2 \in \ker f \tag{3.3.2}$$

So, all the sets of counterimages of vectors can be written as $f^{-1}(\mathbf{w}) = \{\mathbf{u} + \mathbf{v}, \forall \mathbf{v} \in \ker f\}$ and are thus linear varieties of direction.

3.4 Linear systems

Given the matrix $\mathbf{A}_{m \times n}(\mathbb{R})$, we define the function L_A as usual. Now, for some $\mathbf{b} \in \mathbb{R}^n$, $L_A^{-1}(\mathbf{b}) \neq \emptyset \iff \mathbf{b} \in \text{im } L_A \iff \text{im } L_A = \langle \mathbf{A}_i, \mathbf{b} \rangle$ where $\mathbf{A}_i$ are the columns of the matrix.

So, we can write the matrix $(\mathbf{A}|\mathbf{b}) \in M_{m \times (n+1)}(\mathbb{R})$ which is **A** with an additional column containing **b**.

Theorem 3.4.1 (Ronché — Capelli). The linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ has solution iff $\text{rank } \mathbf{A} = \text{rank}(\mathbf{A}|\mathbf{b})$. If it does, the solutions form a linear variety with $\ker L_A$ as its directing space; therefore it has dimension $n - \text{rank } A$.

Contents