Notes on Geometry and Linear Algebra

Jacopo Tissino

October 18, 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:

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

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:

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

Co-hyper-planar 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

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

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

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 $\forall i : x_i = \alpha_i$. □

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

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

□

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 $\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| = (\|\mathbf{v}\| + \|\mathbf{w}\|)^2$ □

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

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

□

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

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

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 $\mathbf{v}$. □

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:

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

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

□

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 $\mathbf{v} \times \mathbf{w}$.

1.8 Perpendicular set

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

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

$S^\perp$ is always a subspace:

[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 $\alpha = \beta = 0$. □

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

□

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

1.9 Affine spaces

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

Definition 1.9.1. An affine space is the set of points on which the vector space $\mathbb{R}^3$ acts.

$$f : \mathbb{A}^3(\mathbb{R}) \times \mathbb{R} \to \mathbb{A}^3(\mathbb{R}) \tag{1.9.1}$$

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.

1.9.1 Linear varieties

They correspond to the subspaces of the vector space.

Definition 1.9.2. 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 $\mathbb{R}^3$.

$$P_0 + U = \{q \in \mathbb{A}^3(\mathbb{R}) : q - P_0 \in U\} \tag{1.9.2}$$

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 $\mathbf{R}^2$.

Equations of linear varieties

Describing a linear variety in $\mathbb{A}^3(\mathbb{R})$ as $P+ < \mathbf{x} >$ gives us parametric equations. We just apply the definition ***CHECK***.

$$r : P_1 + < P_2 - P_1 >= P_1 + \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix} \tag{1.9.3}$$

1.10 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

$$\mathrm{d}(S, T) := \inf \big\{ \mathrm{d}(P, Q), P \in S, Q \in T \big\} \subseteq [0, +\infty[\tag{1.10.1}$$

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

1.10.1 Distance between linear varieties

Theorem 1.10.1. *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

$$\mathrm{d}(p_0, q_0) = \| q_1 - \mathbf{w} - p_1 - \mathbf{u} \| = \mathbf{n} \tag{1.10.2}$$

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

$$d(p', q')^2 = \left\|q' - p'\right\|^2 = \left\|q_0 - p_0 + \mathbf{w}' - \mathbf{u}'\right\|^2 = \left\|\mathbf{n} + \mathbf{w}' - \mathbf{u}'\right\|^2 \tag{1.10.3}$$

but we also know that

$$\mathbf{n} \cdot (\mathbf{w} - \mathbf{u}) = 0 \tag{1.10.4}$$

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

$$\left\|\mathbf{n} + \mathbf{w}' - \mathbf{u}'\right\|^2 = \|\mathbf{n}\|^2 + \|\mathbf{w} - \mathbf{u}\|^2 \geq \|\mathbf{n}\| \tag{1.10.5}$$

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 $d(p', q') = d(p_0, q_0)$. □

Contents