2 Vector Spaces, Affine Spaces, and Metric Spaces

This chapter is only meant to give a short overview of the most important concepts in linear algebra, affine spaces, and metric spaces and is not intended as a course; for that we refer to the vast literature, e.g., [1] for linear algebra and [2] for metric spaces. We will in particular skip most proofs.

In Sect. 2.1 on vector spaces we present the basic concepts of linear algebra: vector space, subspace, basis, dimension, linear map, matrix, determinant, eigenvalue, eigenvector, and inner product. This should all be familiar concepts from a first course on linear algebra. What might be less familiar is the abstract view where the basic concepts are vector spaces and linear maps, while coordinates and matrices become derived concepts. In Sect. 2.1.5 we state the singular value decomposition which is used for mesh simplification and in the ICP algorithm for registration.

In Sect. 2.2 on affine spaces we only give the basic definitions: affine space, affine combination, convex combination, and convex hull. The latter concept is used in Delauney triangulation.

Finally in Sect. 2.3 we introduce metric spaces which makes the concepts of open sets, neighborhoods, and continuity precise.

2.1 Vector Spaces and Linear Algebra

A vector space consists of elements, called vectors, that we can add together and multiply with scalars (real numbers), such that the normal rules hold. That is,

Definition 2.1 A real vector space is a set V together with two binary operations $V \times V \to V : (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v}$ and $\mathbb{R} \times V \to V : (\lambda, \mathbf{v}) \mapsto \lambda\mathbf{v}$, such that:

1. For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
2. For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. There exists a zero vector $\mathbf{0} \in V$, i.e., for any $\mathbf{u} \in V$, $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
4. All $\mathbf{u} \in V$ has a negative element, i.e., there exists $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
5. For all $\alpha, \beta \in \mathbb{R}$ and $\mathbf{u} \in V$, $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$.

J.A. Bærentzen et al., *Guide to Computational Geometry Processing*,
DOI 10.1007/978-1-4471-4075-7_2, © Springer-Verlag London 2012

6. For all $\alpha, \beta \in \mathbb{R}$ and $\mathbf{u} \in V$, $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$.
7. For all $\alpha \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$.
8. Multiplication by $1 \in \mathbb{R}$ is the identity, i.e., for all $\mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$.

Remark 2.1 In the definition above the set $\mathbb{R}$ of real numbers can be replaced with the set $\mathbb{C}$ of complex numbers and then we obtain the definition of a complex vector space. We can in fact replace $\mathbb{R}$ with any field, e.g., the set $\mathbb{Q}$ of rational numbers, the set of rational functions, or with finite fields such as $\mathbb{Z}_2 = \{0, 1\}$.

Remark 2.2 We often write the sum $\mathbf{u} + (-\mathbf{v})$ as $\mathbf{u} - \mathbf{v}$.

We leave the proof of the following proposition as an exercise.

Proposition 2.1 *Let V be a vector space and let $\mathbf{u} \in V$ be a vector.*
1. *The zero vector is unique, i.e., if $\mathbf{0}', \mathbf{u} \in V$ are vectors such that $\mathbf{0}' + \mathbf{u} = \mathbf{u}$, then $\mathbf{0}' = \mathbf{0}$.*
2. *If $\mathbf{v}, \mathbf{w} \in V$ are negative elements to $\mathbf{u}$, i.e., if $\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{w} = \mathbf{0}$, then $\mathbf{v} = \mathbf{w}$.*
3. *Multiplication with zero gives the zero vector, i.e., $0\mathbf{u} = \mathbf{0}$.*
4. *Multiplication with -1 gives the negative vector, i.e., $(-1)\mathbf{u} = -\mathbf{u}$.*

Example 2.1 The set of vectors in the plane or in space is a real vector space.

Example 2.2 The set $\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}, i = 1, \ldots, n\}$ is a real vector space, with addition and multiplication defined as

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n), \tag{2.1}$$

$$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n). \tag{2.2}$$

Example 2.3 The complex numbers $\mathbb{C}$ with usual definition of addition and multiplication is a real vector space.

Example 2.4 The set $\mathbb{C}^n$ with addition and multiplication defined by (2.1) and (2.2) is a real vector space.

Example 2.5 Let Ω be a domain in $\mathbb{R}^n$. A real function $f : \Omega \to \mathbb{R}$ is called a C^n function if all partial derivatives up to order n exist and are continuous, the set of these functions is denoted $C^n(\Omega)$, and it is a real vector space with addition and multiplication defined as

$$(f + g)(x) = f(x) + g(x),$$

$$(\alpha f)(x) = \alpha f(x).$$

Example 2.6 Let Ω be a domain in $\mathbb{R}^n$. A map $f : \Omega \to \mathbb{R}^k$ is called a C^n map if each coordinate function is a C^n function. The set of these functions is denoted $C^n(\Omega, \mathbb{R}^k)$ and it is a real vector space, with addition and multiplication defined as

$$(f+g)(x) = f(x) + g(x),$$
$$(\alpha f)(x) = \alpha f(x).$$

Example 2.7 The set of real polynomials is a real vector space.

Example 2.8 The set of solutions to a system of homogeneous linear equations is a vector space.

Example 2.9 The set of solutions to a system of homogeneous linear ordinary differential equations is a vector space.

Example 2.10 If U and V are real vector spaces, then $U \times V$ is a real vector space too, with addition and multiplication defined as

$$(\mathbf{u}_1, \mathbf{v}_1) + (\mathbf{u}_2, \mathbf{v}_2) = (\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}_1 + \mathbf{v}_2),$$
$$\alpha(\mathbf{u}, \mathbf{v}) = (\alpha\mathbf{u}, \alpha\mathbf{v}).$$

Example 2.11 Let $a = t_0 < t_1 < \cdots < t_k = b$ be real numbers and let $n, m \in \mathbb{Z}_0$ be non zero integers. The space

$$\left\{ f \in C^n\big([a,b]\big) \mid f|_{[t_{\ell-1}, t_\ell]} \text{ is a polynomial of degree at most } m,\ \ell = 1, \ldots, k \right\}$$

is a real vector space.

2.1.1 Subspaces, Bases, and Dimension

A subset $U \subseteq V$ of a vector space is called a subspace if it is a vector space itself. As it is contained in a vector space we do not need to check all the conditions in Definition 2.1. In fact, we only need to check that it is *stable* with respect to the operations. That is,

Definition 2.2 A subset $U \subseteq V$ of a vector space V is a subspace if
1. For all $\mathbf{u}, \mathbf{v} \in U$, $\mathbf{u} + \mathbf{v} \in U$.
2. For all $\alpha \in \mathbb{R}$ and $\mathbf{u} \in U$, $\alpha\mathbf{u} \in U$.

Example 2.12 The subset $\{(x, y, 0) \in \mathbb{R}^3 \mid (x, y) \in \mathbb{R}^3\}$ is a subspace of $\mathbb{R}^3$.

Example 2.13 The subsets $\{\mathbf{0}\}, V \subseteq V$ are subspaces of V called the *trivial* subspaces.

Example 2.14 If $U, V \subseteq W$ are subspaces of W the $U \cap V$ is a subspace too.

Example 2.15 If U and V are vector spaces, then $U \times \{\mathbf{0}\}$ and $\{\mathbf{0}\} \times V$ are subspaces of $U \times V$.

Example 2.16 The subsets $\mathbb{R}, \mathrm{i}\mathbb{R} \subseteq \mathbb{C}$ of real and purely imaginary numbers, respectively, are subspaces of $\mathbb{C}$.

Example 2.17 The set of solutions to k real homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

Example 2.18 If $m \leq n$ then $C^n([a, b])$ is a subspace of $C^m([a, b])$.

Example 2.19 The polynomial of degree at most n is a subspace of the space of all polynomials.

Definition 2.3 Let $X \subseteq V$ be a non empty subset of a vector space. The subspace spanned by X is the smallest subspace of V that contains X. It is not hard to see that it is the set consisting of all linear combinations of elements from X,

$$\operatorname{span} X = \{\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n \mid \alpha_i \in \mathbb{R}, \mathbf{v}_1, \ldots, \mathbf{v}_n \in X, n \in \mathbb{N}\}. \tag{2.3}$$

If $\operatorname{span} X = V$ then we say that X spans V and X is called a spanning set.

Example 2.20 A non zero vector in space spans all vectors on a line.

Example 2.21 Two non zero vectors in space that are not parallel span all vectors in a plane.

Example 2.22 The complex numbers 1 and i span the set of real and purely imaginary numbers, respectively, i.e., $\operatorname{span}\{1\} = \mathbb{R} \subseteq \mathbb{C}$ and $\operatorname{span}\{\mathrm{i}\} = \mathrm{i}\mathbb{R} \subseteq \mathbb{C}$.

Definition 2.4 The *sum of two subspaces* $U, V \subseteq W$ is the subspace

$$U + V = \operatorname{span}(U \cup V) = \{\mathbf{u} + \mathbf{v} \in W \mid \mathbf{u} \in U \wedge \mathbf{v} \in V\}. \tag{2.4}$$

If $U \cap V = \{\mathbf{0}\}$ then the sum is called the *direct sum* and is written as $U \oplus V$.

Example 2.23 The complex numbers are the direct sum of the real and purely imaginary numbers, i.e., $\mathbb{C} = \mathbb{R} \oplus \mathrm{i}\mathbb{R}$.

Definition 2.5 A finite subset $X = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ is called linearly independent if the only solution to the equation

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n = \mathbf{0}$$

is the trivial one, $\alpha_1 = \cdots = \alpha_n = 0$. That is, the only linear combination that gives the zero vector is the trivial one. Otherwise, the set is called linearly dependent.

An important property of vector spaces is the existence of a *basis*. This is secured by the following theorem, which we shall not prove.

Theorem 2.1 *For a finite subset* $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ *of a vector space the following three statements are equivalent.*

1. $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *is a minimal spanning set.*
2. $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *is a maximal linearly independent set.*
3. *Each vector* $\mathbf{v} \in V$ *can be written as a unique linear combination*

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n.$$

If $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ *and* $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *both satisfy these conditions then* $m = n$.

Definition 2.6 A finite set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ of a vector space is called a basis if it satisfies one, and hence all, of the conditions in Theorem 2.1. The unique number of elements in a basis is called the dimension of the vector space and is denoted $\dim V = n$.

Theorem 2.2 *Let* V *be a finite dimensional vector space and let* $X \subseteq V$ *be a subset. Then the following holds*:

1. *If* X *is linearly independent then we can find a set of vectors* $Y \subseteq V$ *such that* $X \cup Y$ *is a basis.*
2. *If* X *is a spanning set then we can find a basis* $Y \subseteq X$.

The theorem says that we always can supplement a linearly independent set to a basis and that we always can extract a basis from a spanning set.

Corollary 2.1 *If* $U, V \subseteq W$ *are finite dimensional subspaces of* W *then*

$$\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V). \tag{2.5}$$

Example 2.24 Two vectors not on the same line are a basis for all vectors in the plane.

Example 2.25 Three vectors not in the same plane are a basis for all vectors in space.

Example 2.26 The vectors

$$\mathbf{e}_k = (\underbrace{0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0}_{n-k}) \in \mathbb{R}^n, \quad k = 1, \dots, n, \tag{2.6}$$

are a basis for $\mathbb{R}^n$ called the *standard basis*, so $\dim(\mathbb{R}^n) = n$.

Example 2.27 The complex numbers 1 and i are a basis for $\mathbb{C}$.

Example 2.28 If $U \cap V = \{\mathbf{0}\}$ are subspaces of a vector space and $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $\{\mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ are bases for U and V, respectively, then $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_\ell\}$ is a basis for $U \oplus V$.

Example 2.29 The monomials $1, x, \dots, x^n$ are a basis for the polynomials of degree at most n.

Example 2.30 The Bernstein polynomials $B_k^n(x) = \binom{n}{k}(1-x)^{n-k}x^k$, $k = 0, \dots, n$ are a basis for the polynomials of degree at most n.

2.1.2 Linear Maps, Matrices, and Determinants

A map between vector spaces is linear if it preserves addition and multiplication with scalars. That is,

Definition 2.7 Let U and V be vector spaces. A map $L : U \to V$ is linear if:
1. For all $\mathbf{u}, \mathbf{v} \in U$, $L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$.
2. For all $\alpha \in \mathbb{R}$ and $\mathbf{u} \in U$, $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u})$.

Example 2.31 If V is a vector space and $\alpha \in \mathbb{R}$ is a real number then multiplication by α: $V \to V : \mathbf{v} \mapsto \alpha\mathbf{v}$ is a linear map.

Example 2.32 The map $\mathbb{R} \to \mathbb{R} : x \mapsto ax + b$ with $b \neq 0$ is *not* linear, cf., Exercise 2.7.

Example 2.33 Differentiation $C^n([a,b]) \to C^{n-1}([a,b]) : f \mapsto \frac{\mathrm{d}f}{\mathrm{d}x}$ is a linear map.

Example 2.34 If $L_1, L_2 : U \to V$ are two linear maps, then the sum $L_1 + L_2 : U \to V : \mathbf{u} \mapsto L_1(\mathbf{u}) + L_2(\mathbf{u})$ is a linear map too.

Example 2.35 If $\alpha \in \mathbb{R}$ and $L : U \to V$ is a linear map, then the scalar product $\alpha L : U \to V : \mathbf{u} \mapsto \alpha L(\mathbf{u})$ is a linear map too.

Example 2.36 If $L_1 : U \to V$ and $L_2 : V \to W$ are linear maps, then the composition $L_2 \circ L_1 : U \to W$ is a linear map too.

Example 2.37 If $L : U \to V$ is linear and bijective, then the inverse map $L^{-1} : V \to U$ is linear too.

Examples 2.34 and 2.35 show that the space of linear maps between two vector spaces is a vector space.

Recall the definition of an *injective*, *surjective*, and *bijective* map.

Definition 2.8 A map $f : A \to B$ between two sets is

- injective if for all $x, y \in A$ we have $f(x) = f(y) \implies x = y$;
- surjective if there for all $y \in B$ exists $x \in A$ such that $f(x) = y$;
- bijective if it is both injective and surjective.

A map is *invertible* if and only if it is bijective.

Definition 2.9 Let $L : U \to V$ be a linear map. The *kernel* of L is the set

$$\ker L = L^{-1}(\mathbf{0}) = \{\mathbf{u} \in U \mid L(\mathbf{u}) = \mathbf{0}\}, \tag{2.7}$$

and the *image* of L is the set

$$L(U) = \{f(\mathbf{u}) \in V \mid \mathbf{u} \in U\}. \tag{2.8}$$

We have the following.

Theorem 2.3 *Let $L : U \to V$ be a linear map between two vector spaces. Then the kernel* $\ker L$ *is a subspace of U and the image $L(U)$ is a subspace of V. If U and V are finite dimensional then*
1. $\dim U = \dim \ker L + \dim L(U)$;
2. *L is injective if and only if* $\ker(L) = \{\mathbf{0}\}$;
3. *if L is injective then* $\dim U \leq \dim V$;
4. *if L is surjective then* $\dim U \geq \dim V$;
5. *if* $\dim U = \dim V$ *then L is surjective if and only if L is injective.*

If $L : U \to V$ is linear and $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is a basis for U and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a basis for V, then we can write the image of a basis vector $\mathbf{u}_j$ as $L(\mathbf{u}_j) = \sum_{i=1}^n a_{ij}\mathbf{v}_i$. Then the image of an arbitrary vector $\mathbf{u} = \sum_{j=1}^m x_j \mathbf{u}_j \in U$ is

$$\begin{aligned} L\left(\sum_{j=1}^m x_j \mathbf{u}_j\right) &= \sum_{j=1}^m x_j L(\mathbf{u}_j) = \sum_{j=1}^m x_j \sum_{i=1}^n a_{ij}\mathbf{v}_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij} x_j\right)\mathbf{v}_i = \sum_{i=1}^n y_i \mathbf{v}_i. \end{aligned} \tag{2.9}$$

We see that the coordinates y_i of the image vector $L(\mathbf{u})$ is given by the coordinates x_j of $\mathbf{u}$ by the following matrix equation:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}. \tag{2.10}$$

The matrix with entries a_{ij} is called the *matrix for L with respect to the bases* $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *and* $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Observe that the columns consist of the coordinates of the image of the basis vectors. Also observe that the first index i in a_{ij} gives the row number while the second index j gives the column number.

We denote the ith row in $\mathbf{A}$ by $\mathbf{A}_{i\_}$ and the jth column by $\mathbf{A}_{|j}$. That is,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{1\_} \\ \vdots \\ \mathbf{A}_{n\_} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{|1} \dots \mathbf{A}_{|m} \end{pmatrix}. \tag{2.11}$$

Addition of linear maps now corresponds to addition of matrices,

$$\begin{aligned} &\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix} \\ &\quad = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \dots & a_{nm}+b_{nm} \end{pmatrix} \end{aligned} \tag{2.12}$$

and scalar multiplication of linear maps corresponds to multiplication of a matrix with a scalar

$$\alpha \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \dots & \alpha a_{1m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n1} & \dots & \alpha a_{nm} \end{pmatrix}. \tag{2.13}$$

Composition of linear maps corresponds to matrix multiplication, which is defined as follows. If $\mathbf{A}$ is a $k \times m$ matrix with entries a_{ij} and $\mathbf{B}$ is an $m \times n$ matrix with entries b_{ij} then the product is an $k \times n$ matrix $\mathbf{C} = \mathbf{AB}$ where the element c_{ij} is the sum of the products of the elements in the ith row from $\mathbf{A}$ and the jth column from $\mathbf{B}$, i.e.,

$$c_{ij} = \mathbf{A}_{i\_}\mathbf{B}_{|j} = \sum_{k=1}^{m} a_{ik} b_{kj}. \tag{2.14}$$

The identity matrix is the $n \times n$ matrix with ones in the diagonal and zeros elsewhere,

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}. \tag{2.15}$$

If $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ is an $m \times n$ matrix then

$$\mathbf{IA} = \mathbf{A} \quad \text{and} \quad \mathbf{BI} = \mathbf{B}. \tag{2.16}$$

Definition 2.10 We say an $n \times n$ matrix $\mathbf{A}$ is invertible if there exists a matrix $\mathbf{A}^{-1}$ such that

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \tag{2.17}$$

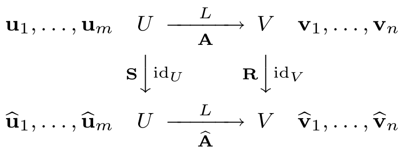

Fig. 2.1 The matrix for a linear map with respect to different bases

The matrix $\mathbf{A}^{-1}$ is then called the inverse of $\mathbf{A}$.

Theorem 2.4 *Let* $\mathbf{A}$ *be the matrix for a linear map* $L : U \to V$ *with respect to the bases* $\mathbf{u}_1, \dots, \mathbf{u}_m$ *and* $\mathbf{v}_1, \dots, \mathbf{v}_m$ *for* U *and* V*, respectively. Then* $\mathbf{A}$ *is invertible if and only if* L *is bijective. In that case* $\mathbf{A}^{-1}$ *is the matrix for* L^{-1} *with respect to the bases* $\mathbf{v}_1, \dots, \mathbf{v}_m$ *and* $\mathbf{u}_1, \dots, \mathbf{u}_m$.

An in some sense trivial, but still important special case is when $U = V$ and the map is the identity map $\mathrm{id} : \mathbf{u} \mapsto \mathbf{u}$. Let $\mathbf{S}$ be the matrix of id with respect to the bases $\mathbf{u}_1, \dots, \mathbf{u}_m$ and $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$. The jth column of $\mathbf{S}$ consists of the coordinates of $\mathrm{id}(\mathbf{u}_j) = \mathbf{u}_j$ with respect to the basis $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$. Equation (2.10) now reads

$$\underline{\underline{\widehat{\mathbf{u}}}} = \mathbf{S}\underline{\underline{\mathbf{u}}}, \tag{2.18}$$

and gives us the relation between the coordinates $\underline{\underline{\mathbf{u}}}$ and $\underline{\underline{\widehat{\mathbf{u}}}}$ of the same vector $\mathbf{u}$ with respect to the bases $\mathbf{u}_1, \dots, \mathbf{u}_m$ and $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$, respectively.

Now suppose we have a linear map $L : U \to V$ between two vector spaces, and two pairs of different bases, $\mathbf{u}_1, \dots, \mathbf{u}_m$ and $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$ for U and $\mathbf{v}_1, \dots, \mathbf{v}_n$ and $\widehat{\mathbf{v}}_1, \dots, \widehat{\mathbf{v}}_n$ for V. Let $\mathbf{A}$ be the matrix for L with respect to the bases $\mathbf{u}_1, \dots, \mathbf{u}_m$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ and let $\widehat{\mathbf{A}}$ be the matrix for L with respect to the bases $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$ and $\widehat{\mathbf{v}}_1, \dots, \widehat{\mathbf{v}}_n$. Let furthermore S be the matrix for the identity $U \to U$ with respect to the bases $\mathbf{u}_1, \dots, \mathbf{u}_m$ and $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_m$ and let R be the matrix for the identity $V \to V$ with respect to the bases $\mathbf{v}_1, \dots, \mathbf{v}_n$ and $\widehat{\mathbf{v}}_1, \dots, \widehat{\mathbf{v}}_n$; then

$$\widehat{\mathbf{A}} = \mathbf{RAS}^{-1}, \tag{2.19}$$

see Fig. 2.1. A special case is when $U = V$, $\mathbf{v}_i = \mathbf{u}_i$, and $\widehat{\mathbf{v}}_i = \widehat{\mathbf{u}}_i$. Then we have $\widehat{\mathbf{A}} = \mathbf{SAS}^{-1}$.

Definition 2.11 The transpose of a matrix $\mathbf{A}$ is the matrix $\mathbf{A}^T$ which is obtained by interchanging the rows and columns. That is, if $\mathbf{A}$ has entries a_{ij}, then $\mathbf{A}^T$ has entries α_{ij}, where $\alpha_{ij} = a_{ji}$.

Definition 2.12 An $n \times n$ matrix $\mathbf{A}$ is called *symmetric* if $\mathbf{A}^T = \mathbf{A}$.

Definition 2.13 An $n \times n$ matrix $\mathbf{U}$ is called *orthogonal* if $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, i.e., if $\mathbf{U}^{-1} = \mathbf{U}^T$.

Definition 2.14 An $n \times n$ matrix $\mathbf{A}$ is called *positive definite* if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all non zero column vectors $\mathbf{x}$.

Before we can define the determinant of a matrix we need the notion of permutations.

Definition 2.15 A permutation is a bijective map $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. If $i \neq j$, then σ_{ij} denotes the transposition that interchanges i and j, i.e., the permutation defined by

$$\sigma_{ij}(i) = j, \qquad \sigma_{ij}(j) = i, \qquad \sigma_{ij}(k) = k, \quad \text{if } k \neq i, j. \tag{2.20}$$

It is not hard to see that any permutation can be written as the composition of a number of transpositions $\sigma = \sigma_{i_k j_k} \circ \cdots \circ \sigma_{i_2 j_2} \circ \sigma_{i_1 j_1}$. This description is far from unique, but the number k of transpositions needed for a given permutation σ is either always even or always odd. If the number is even σ is called an *even permutation*, otherwise it is called an *odd permutation*. The sign of a sigma is now defined as

$$\operatorname{sign} \sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{if } \sigma \text{ is odd.} \end{cases} \tag{2.21}$$

Definition 2.16 The determinant of an $n \times n$ matrix $\mathbf{A}$ is the completely anti symmetric multilinear function of the columns of $\mathbf{A}$ that is 1 on the identity matrix. That is,

$$\det(\mathbf{A}_{|\sigma(1)}, \dots, \mathbf{A}_{|\sigma(n)}) = \operatorname{sign}(\sigma) \det(\mathbf{A}_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}) \tag{2.22}$$

$$\begin{aligned} \det\bigl(\mathbf{A}'_{|1} + \mathbf{A}''_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}\bigr) &= \det\bigl(\mathbf{A}'_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}\bigr) \\ &\quad + \det\bigl(\mathbf{A}''_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}\bigr), \end{aligned} \tag{2.23}$$

$$\det\bigl(\alpha \mathbf{A}_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}\bigr) = \alpha \det\bigl(\mathbf{A}_{|1}, \mathbf{A}_{|2}, \dots, \mathbf{A}_{|n}\bigr), \tag{2.24}$$

$$\det(\mathbf{I}) = 1, \tag{2.25}$$

where σ is a permutation. The determinant of $\mathbf{A}$ can be written

$$\det \mathbf{A} = \sum_{\sigma} \operatorname{sign} \sigma \prod_{i=1}^{n} a_{i\sigma(i)}, \tag{2.26}$$

where the sum is over all permutations σ of $\{1, \dots, n\}$.

The definition is not very practical, except in the case of 2×2 and 3×3 matrices. Here we have

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \tag{2.27}$$

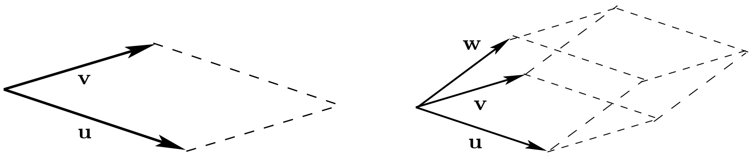

Fig. 2.2 The area and volume can be calculated as determinants: area = det(**u**, **v**) and volume = det(**u**, **v**, **w**)

The determinant of a 2×2 matrix $\mathbf{A}$ can be interpreted as the *signed area* of the parallelogram in $\mathbb{R}^2$ spanned by the vectors $\mathbf{A}_{1\_}$ and $\mathbf{A}_{2\_}$, see Fig. 2.2.

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \tag{2.28}$$

The determinant of a 3×3 matrix $\mathbf{A}$ can be interpreted as the *signed volume* of the parallelepiped spanned by the vectors $\mathbf{A}_{1\_}$, $\mathbf{A}_{2\_}$, and $\mathbf{A}_{3\_}$, see Fig. 2.2. The same is true in higher dimensions. The determinant of a $n \times n$ matrix A is the signed n-dimensional volume of the n-dimensional parallelepiped spanned by the columns of A.

For practical calculations one makes use of the following properties of the determinant.

Theorem 2.5 *Let* $\mathbf{A}$ *be an* $n \times n$ *matrix, then*

$$\det \mathbf{A}^T = \det \mathbf{A}. \tag{2.29}$$

The determinant changes sign if two rows or columns are interchanged, in particular

$$\det \mathbf{A} = 0, \quad \textit{if two rows or columns in } \mathbf{A} \textit{ are equal}, \tag{2.30}$$

$$\det \mathbf{A} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij}, \quad \textit{for } i = 1, \dots, n, \tag{2.31}$$

where $\mathbf{A}_{ij}$ *is the matrix obtained from* $\mathbf{A}$ *by deleting the* i*th row and* j*th column, i.e., the row and column where* a_{ij} *appears. If* $\mathbf{B}$ *is another* $n \times n$ *matrix then*

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}). \tag{2.32}$$

The matrix $\mathbf{A}$ *is invertible if and only if* $\det \mathbf{A} \neq 0$, *and in that case*

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det \mathbf{A}}. \tag{2.33}$$

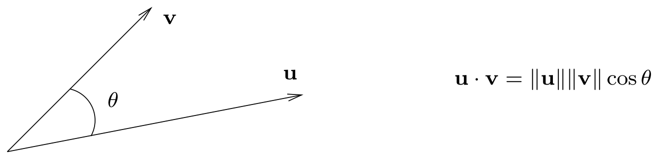

Fig. 2.3 The inner product between two vectors in the plane (or in space)

If $\mathbf{A}$ *is invertible then* $\mathbf{A}^{-1}$ *has entries* α_{ij}, *where*

$$\alpha_{ij} = \frac{(-1)^{i+j}\det \mathbf{A}_{ji}}{\det \mathbf{A}}. \tag{2.34}$$

Suppose $\mathbf{A}$ and $\widehat{\mathbf{A}}$ are matrices for a linear map $L : V \to V$ with respect to two different bases. Then we have $\widehat{\mathbf{A}} = \mathbf{SAS}^{-1}$ where $\mathbf{S}$ is an invertible matrix. We now have $\det \widehat{\mathbf{A}} = \det(\mathbf{SAS}^{-1}) = \det \mathbf{S} \det \mathbf{A} \det \mathbf{S}^{-1} = \det \mathbf{A}$. Thus, we can define the determinant of L as the determinant of any matrix representation and we clearly see that L is injective if and only if $\det L \neq 0$.

2.1.3 Euclidean Vector Spaces and Symmetric Maps

For vectors in the plane, or in space, we have the concepts of length and angles. This then leads to the definition of the inner product, see Fig. 2.3. For two vectors $\mathbf{u}$ and $\mathbf{v}$ it is given by

$$\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta, \tag{2.35}$$

where $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ is the length of a $\mathbf{u}$ and $\mathbf{v}$, respectively, and θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

A general vector space V does not have the a priori notions of length and angle and in order to be able to have the concepts of length and angle we introduce an abstract inner product.

Definition 2.17 An *Euclidean vector space* is a real vector space V equipped with a positive definite, symmetric, bilinear mapping $V \times V \to \mathbb{R} : (\mathbf{u}, \mathbf{v}) \mapsto \langle \mathbf{u}, \mathbf{v}\rangle$, called the *inner product*, i.e., we have the following:

1. For all $\mathbf{u}, \mathbf{v} \in V$, $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$.
2. For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\langle \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$.
3. For all $\alpha \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, $\langle \alpha\mathbf{u}, \mathbf{v}\rangle = \alpha\langle \mathbf{u}, \mathbf{v}\rangle$.
4. For all $\mathbf{u} \in V$, $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$.
5. For all $\mathbf{u} \in V$, $\langle \mathbf{u}, \mathbf{u}\rangle = 0 \iff \mathbf{u} = \mathbf{0}$.

Example 2.38 The set of vectors in the plane or in space equipped with the inner product (2.35) is an Euclidean vector space. The norm (2.41) becomes the usual length and the angle defined by (2.44) is the usual angle.

Example 2.39 The set $\mathbb{R}^n$ equipped with inner product

$$\big\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \big\rangle = x_1 y_1 + \cdots + x_n y_n, \tag{2.36}$$

is an Euclidean vector space.

Example 2.40 The space $C^n([a, b])$ of n times differentiable functions with continuous nth derivative equipped with the inner product

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, \mathrm{d}x, \tag{2.37}$$

is an Euclidean vector space. The corresponding norm is called the L^2-norm.

Example 2.41 If $(V_1, \langle \cdot, \cdot \rangle_1)$ and $(V_2, \langle \cdot, \cdot \rangle_2)$ are Euclidean vector spaces, then $V_1 \times V_2$ equipped with the inner product

$$\big\langle (u_1, u_2), (v_1, v_2) \big\rangle = \langle u_1, v_1 \rangle_1 + \langle u_2, v_2 \rangle_2, \tag{2.38}$$

is an Euclidean vector space.

Example 2.42 If $(V, \langle \cdot, \cdot \rangle)$ is an Euclidean vector space and $U \subseteq V$ is a subspace then U equipped with the restriction $\langle \cdot, \cdot \rangle|_{U \times U}$ of $\langle \cdot, \cdot \rangle$ to $U \times U$ is an Euclidean vector space too.

Example 2.43 The space $C_0^\infty([a, b]) = \{ f \in C^\infty([a, b]) \mid f(a) = f(b) = 0 \}$ of infinitely differentiable functions that are zero at the endpoints equipped with the restriction of the inner product (2.37) is an Euclidean vector space.

If $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis for V, $\mathbf{v} = \sum_{k=1}^n v_i \mathbf{u}_i$, and $\mathbf{w} = \sum_{k=1}^n w_i \mathbf{u}_i$ then the inner product of $\mathbf{v}$ and $\mathbf{w}$ can be written

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{k,\ell=1}^n v_k w_\ell \langle \mathbf{u}_k, \mathbf{u}_\ell \rangle = \underline{\underline{\mathbf{v}}}^T \mathbf{G} \underline{\underline{\mathbf{w}}}, \tag{2.39}$$

where $\underline{\underline{\mathbf{v}}}$ and $\underline{\underline{\mathbf{w}}}$ are the coordinates with respect to the basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ of $\mathbf{v}$ and $\mathbf{w}$, respectively, and $\mathbf{G}$ is the matrix

$$\mathbf{G} = \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \ldots & \langle \mathbf{u}_1, \mathbf{u}_n \rangle \\ \vdots & & \vdots \\ \langle \mathbf{u}_n, \mathbf{u}_1 \rangle & \ldots & \langle \mathbf{u}_n, \mathbf{u}_n \rangle \end{pmatrix}. \tag{2.40}$$

It is called the matrix for the inner product with respect to the basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and it is a positive definite symmetric matrix. Observe that we have the same kind of matrix representation of a symmetric bilinear map, i.e., a map that satisfies condition

(1), (2), and (3) in Definition 2.17. The matrix G is still symmetric but it need not be positive definite.

Let $\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_n$ be another basis, let $\widehat{\mathbf{G}}$ be the corresponding matrix for the inner product, and let $\mathbf{S}$ be the matrix for the identity on V with respect to the two bases. Then the coordinates of a vector $\mathbf{u}$ with respect to the bases satisfies (2.18) and we see that $\underline{\underline{\widehat{\mathbf{u}}}}^T \widehat{\mathbf{G}} \underline{\underline{\widehat{\mathbf{v}}}} = \underline{\underline{\mathbf{u}}}^T \mathbf{S}^T \widehat{\mathbf{G}} \mathbf{S} \underline{\underline{\mathbf{v}}}$. That is, $\mathbf{G} = \mathbf{S}^T \widehat{\mathbf{G}} \mathbf{S}$.

Definition 2.18 The norm of a vector $\mathbf{u} \in V$ in an Euclidean vector space $(V, \langle \cdot, \cdot \rangle)$ is defined as

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}. \tag{2.41}$$

A very important property of an arbitrary inner product is the Cauchy–Schwartz inequality.

Theorem 2.6 *If $(V, \langle \cdot, \cdot \rangle)$ is an Euclidean vector space then the inner product satisfies the Cauchy–Schwartz inequality*

$$\left| \langle \mathbf{u}, \mathbf{v} \rangle \right| \le \|\mathbf{u}\| \, \|\mathbf{v}\|, \tag{2.42}$$

with equality if and only if one of the vectors is a positive multiple of the other.

Corollary 2.2 *The norm satisfies the following conditions*:
1. *For all* $\alpha \in \mathbb{R}$ *and* $\mathbf{u} \in V$, $\|\alpha \mathbf{u}\| = |\alpha| \, \|\mathbf{u}\|$.
2. *For all* $\mathbf{u}, \mathbf{v} \in V$, $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$.
3. *For all* $\mathbf{u} \in V$, $\|\mathbf{u}\| \ge 0$.
4. *For all* $\mathbf{u} \in V$, $\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0}$.

This is the conditions for an abstract norm on a vector space and not all norms are induced by an inner product. But if a norm is induced by an inner product then this inner product is unique. Indeed, if $\mathbf{u}, \mathbf{v} \in V$ then symmetry and bilinearity imply that

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2 \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle.$$

That is, the inner product of two vectors $\mathbf{u}, \mathbf{v} \in V$ can be written as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{2} \bigl(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 \bigr). \tag{2.43}$$

The angle θ between two vectors $\mathbf{u}, \mathbf{v} \in V$ in an Euclidean vector space $(V, \langle \cdot, \cdot \rangle)$ can now be defined by the equation

$$\cos \theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}. \tag{2.44}$$

Two vectors $\mathbf{u}, \mathbf{v} \in V$ are called orthogonal if the angle between them is $\frac{\pi}{2}$, i.e., if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Example 2.44 If $(V, \langle\cdot,\cdot\rangle)$ is an Euclidean vector space and $U \subseteq V$ is a subspace then the *orthogonal complement*

$$U^{\perp} = \{\mathbf{v} \in V \mid \langle\mathbf{u}, \mathbf{v}\rangle = 0 \text{ for all } \mathbf{u} \in U\} \tag{2.45}$$

is a subspace of V, and $V = U \oplus U^{\perp}$.

Definition 2.19 A basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for an Euclidean vector space is called orthonormal if

$$\langle\mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{2.46}$$

That is, the elements of the basis are pairwise orthogonal and have norm 1.

If $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis for an Euclidean vector space V then we can construct an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ by Gram–Schmidt orthonormalization. The elements of that particular orthonormal basis is defined as follows:

$$\mathbf{v}_\ell = \mathbf{u}_\ell - \sum_{k=1}^{\ell-1} \langle\mathbf{u}_\ell, \mathbf{e}_k\rangle \mathbf{e}_k, \qquad \mathbf{e}_\ell = \frac{\mathbf{v}_\ell}{\|\mathbf{v}_\ell\|}, \quad \ell = 1, \ldots, n. \tag{2.47}$$

Definition 2.20 A linear map $L : U \to V$ between two Euclidean vector spaces is called an *isometry* if it is bijective and $\langle L(\mathbf{u}), L(\mathbf{v})\rangle_V = \langle\mathbf{u}, \mathbf{v}\rangle_U$ for all $\mathbf{u}, \mathbf{v} \in U$.

So an isometry preserves the inner product. As the inner product is determined by the norm it is enough to check that the map preserves the norm, i.e., if $\|L(\mathbf{u})\|_V = \|\mathbf{u}\|_U$ for all $\mathbf{u} \in U$ then L is an isometry.

Example 2.45 A rotation in the plane or in space is an isometry.

Example 2.46 A symmetry in space around the origin $\mathbf{0}$ or around a line through $\mathbf{0}$ is an isometry.

Theorem 2.7 *Let $L : U \to V$ be a linear map between two Euclidean vector spaces. Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be bases for U and V, respectively, and let $\mathbf{A}$ be the matrix for L with respect to these bases. Let furthermore G_U and G_V be the matrices for the inner product on U and V, respectively. Then L is an isometry if and only if*

$$\mathbf{A}^T \mathbf{G}_V \mathbf{A} = \mathbf{G}_U. \tag{2.48}$$

If $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ both are orthonormal then $\mathbf{G}_U = \mathbf{G}_V = \mathbf{I}$ and the equation reads

$$\mathbf{A}^T \mathbf{A} = \mathbf{I}, \tag{2.49}$$

i.e., $\mathbf{A}$ is orthogonal.

On a similar note, if $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and $\widehat{\mathbf{u}}_1, \ldots, \widehat{\mathbf{u}}_m$ are bases for an Euclidean vector space U and $\mathbf{u} \in U$ then the coordinates $\underline{\underline{\mathbf{u}}}$ and $\underline{\underline{\widehat{\mathbf{u}}}}$ for $\mathbf{u}$ with respect to the two bases are related by the equation $\underline{\underline{\widehat{\mathbf{u}}}} = \mathbf{S}\underline{\underline{\mathbf{u}}}$, cf. (2.18). If $\mathbf{G}$ and $\widehat{\mathbf{G}}$ are the matrices for the inner product with respect to the bases then we have

$$\underline{\underline{\mathbf{u}}}^T \mathbf{G} \underline{\underline{\mathbf{u}}} = \langle \mathbf{u}, \mathbf{u} \rangle = \underline{\underline{\widehat{\mathbf{u}}}}^T \widehat{\mathbf{G}} \underline{\underline{\widehat{\mathbf{u}}}} = (\mathbf{S}\underline{\underline{\mathbf{u}}})^T \widehat{\mathbf{G}} \mathbf{S} \underline{\underline{\mathbf{u}}} = \underline{\underline{\mathbf{u}}}^T \mathbf{S}^T \widehat{\mathbf{G}} \mathbf{S} \underline{\underline{\mathbf{u}}},$$

i.e., we have

$$\mathbf{G} = \mathbf{S}^T \widehat{\mathbf{G}} \mathbf{S}. \tag{2.50}$$

If the bases both are orthonormal then $\mathbf{G} = \widehat{\mathbf{G}} = \mathbf{I}$ and we see that $\mathbf{S}$ is orthogonal.

Definition 2.21 A linear map $L : V \to V$ from an Euclidean vector space to itself is called symmetric if

$$\big\langle L(\mathbf{u}), \mathbf{v} \big\rangle = \big\langle \mathbf{u}, L(\mathbf{v}) \big\rangle, \quad \text{for all } \mathbf{u}, \mathbf{v} \in V. \tag{2.51}$$

Example 2.47 The map $f \mapsto f''$ is a symmetric map of the space $(C_0^\infty([a, b]), \langle \cdot, \cdot \rangle)$ to itself, where the inner product $\langle \cdot, \cdot \rangle$ is given by (2.37).

If $\mathbf{A}$ is the matrix for a linear map L with respect to some basis and $\mathbf{G}$ is the matrix for the inner product then L is symmetric if and only if $\mathbf{A}^T \mathbf{G} = \mathbf{G}\mathbf{A}$. If the basis is orthonormal then $\mathbf{G} = \mathbf{I}$ and the condition reads $\mathbf{A}^T = \mathbf{A}$, i.e., $\mathbf{A}$ is a symmetric matrix.

2.1.4 Eigenvalues, Eigenvectors, and Diagonalization

Definition 2.22 Let $L : V \to V$ be a linear map. If there exist a non zero vector $\mathbf{v} \in V$ and a scalar $\lambda \in \mathbb{R}$ such that $L(\mathbf{v}) = \lambda \mathbf{v}$ then $\mathbf{v}$ is called an *eigenvector* with *eigenvalue* λ. If λ is an eigenvalue then the space

$$E_\lambda = \big\{ \mathbf{v} \in V \mid L(\mathbf{v}) = \lambda \mathbf{v} \big\} \tag{2.52}$$

is a subspace of V called the *eigenspace* of λ. The dimension of E_λ is called the *geometric multiplicity* of λ.

If $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is a basis for V, $\mathbf{A}$ is the matrix for L in this basis and a vector $\mathbf{v} \in V$ has coordinates $\underline{\underline{\mathbf{v}}}$ with respect to this basis then

$$L(v) = \lambda \mathbf{v} \iff \mathbf{A}\underline{\underline{\mathbf{v}}} = \lambda \underline{\underline{\mathbf{v}}} \tag{2.53}$$

We say that $\underline{\underline{\mathbf{v}}}$ is an eigenvector for the matrix $\mathbf{A}$ with eigenvalue λ.

Example 2.48 Consider the matrix $\left(\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right)$. The vector $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ is an eigenvector with eigenvalue 4 and $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ is an eigenvector with eigenvalue -2.

Example 2.49 The exponential map exp is an eigenvector with eigenvalue 1 for the linear map $C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}) : f \mapsto f'$.

Example 2.50 The trigonometric functions cos and sin are eigenvectors with eigenvalue -1 for the linear map $C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}) : f \mapsto f''$.

We see that λ is an eigenvalue for L if and only if the map $L - \lambda\,\mathrm{id}$ is not injective, i.e., if and only if $\det(L - \lambda\,\mathrm{id}) = 0$. In that case $E_\lambda = \ker(L - \lambda\,\mathrm{id})$. If $\mathbf{A}$ is the matrix for L with respect to some basis for V then we see that

$$\det(L - \lambda\,\mathrm{id}) = \det(\mathbf{A} - \lambda\mathbf{I}) = (-\lambda)^n + \operatorname{tr}\mathbf{A}(-\lambda)^{n-1} + \cdots + \det\mathbf{A} \tag{2.54}$$

is a polynomial of degree n in λ. It is called the *characteristic polynomial* of L (or $\mathbf{A}$). The eigenvalues are precisely the roots of the characteristic polynomial and the multiplicity of a root λ in the characteristic polynomial is called the *algebraic multiplicity* of the eigenvalue λ. The relation between the geometric and algebraic multiplicity is given in the following proposition.

Proposition 2.2 *Let $\nu_g(\lambda) = \dim(E_\lambda)$ be the geometric multiplicity of an eigenvalue λ and let $\nu_a(\lambda)$ be the algebraic multiplicity of λ. Then $1 \le \nu_g(\lambda) \le \nu_a(\lambda)$.*

The characteristic polynomial may have complex roots and even though they strictly speaking are not eigenvalues we will still call them *complex eigenvalues*. Once the eigenvalues are determined the eigenvectors belonging to a particular real eigenvalue λ can be found by determining a non zero solution to the linear equation $L(\mathbf{u}) - \lambda\mathbf{u} = \mathbf{0}$ or equivalently a non zero solution to the matrix equation

$$\begin{pmatrix} a_{1,1}-\lambda & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2}-\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,1} \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n}-\lambda \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.55}$$

If V has a basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ consisting of eigenvectors for L, i.e., $L(\mathbf{u}_k) = \lambda_k\mathbf{u}_k$ then the corresponding matrix is diagonal

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}, \tag{2.56}$$

and we say that L is *diagonalizable*. Not all linear maps (or matrices) can be diagonalized. The condition is that there is a basis consisting of eigenvectors and this is the same as demanding that $V = \bigoplus_\lambda E_\lambda$ or that all eigenvalues are real and the sum of the geometric multiplicities is the dimension of V. If there is a complex eigenvalue then this is impossible. The same is the case if $\nu_g(\lambda) < \nu_a(\lambda)$ for some real eigenvalue λ.

Example 2.51 The matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has no real eigenvalues.

Example 2.52 The matrix $\begin{pmatrix} \sqrt{2} & 1 \\ 0 & \sqrt{2} \end{pmatrix}$ has the eigenvalue $\sqrt{2}$ which has algebraic multiplicity 2 and geometric multiplicity 1.

In case of a symmetric map the situation is much nicer. Indeed, we have the following theorem, which we shall not prove.

Theorem 2.8 *Let $(V, \langle\cdot,\cdot\rangle)$ be an Euclidean vector space and let $L : V \to V$ be a symmetric linear map. Then all eigenvalues are real and V has an orthonormal basis consisting of eigenvectors for L.*

By choosing an orthonormal basis for V we obtain the following theorem for symmetric matrices.

Theorem 2.9 *A symmetric matrix $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{U}^T \boldsymbol{\Lambda} \mathbf{U}$, where $\boldsymbol{\Lambda}$ is diagonal and $\mathbf{U}$ is orthogonal.*

Let $(V, \langle\cdot,\cdot,\rangle)$ be an Euclidean vector space and let $h : V \times V \to \mathbb{R}$ be a symmetric bilinear map, i.e., it satisfies condition (1), (2), and (3) in Definition 2.17. Then there exists a unique symmetric linear map $L : V \to V$ such that $h(\mathbf{u}, \mathbf{v}) = \langle L(\mathbf{u}), \mathbf{v}\rangle$. Theorem 2.8 tells us that V has an orthonormal basis consisting of eigenvectors for L, and with respect to this basis the matrix representation for h is diagonal with the eigenvalues of L in the diagonal. Now suppose we have an arbitrary basis for V and let $\mathbf{G}$ and $\mathbf{H}$ be the matrices for the inner product $\langle\cdot,\cdot\rangle$ and the bilinear map h, respectively. Let furthermore $\mathbf{A}$ be the matrix for L. Then we have $\mathbf{H} = \mathbf{A}^T \mathbf{G}$, or as both $\mathbf{G}$ and $\mathbf{H}$ are symmetric $\mathbf{H} = \mathbf{G}\mathbf{A}$. That is, $\mathbf{A} = \mathbf{G}^{-1}\mathbf{H}$ and the eigenvalue problem $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ is equivalent to the *generalized eigenvalue problem* $\mathbf{H}\mathbf{v} = \lambda\mathbf{G}\mathbf{v}$. This gives us the following generalization of Theorem 2.9.

Theorem 2.10 *Let $\mathbf{G}, \mathbf{H}$ be symmetric $n \times n$ matrices with $\mathbf{G}$ positive definite. Then we can decompose $\mathbf{H}$ as $\mathbf{H} = \mathbf{S}^{-1} \boldsymbol{\Lambda} \mathbf{S}$, where $\boldsymbol{\Lambda}$ is diagonal and $\mathbf{S}$ is orthogonal with respect to $\mathbf{G}$, i.e., $\mathbf{S}^T \mathbf{G} \mathbf{S} = \mathbf{G}$.*

2.1.5 Singular Value Decomposition

Due to its numerically stability the singular value decomposition (SVD) is extensively used for practical calculations such as solving over- and under-determined systems and eigenvalue calculations. We will use it for mesh simplification and in the ICP algorithm for registration. The singular value decomposition can be formulated as

Theorem 2.11 *Let $L : V \to U$ be a linear map between two Euclidean vector spaces of dimension n and m, respectively, and let $k = \min\{m, n\}$. Then there exist an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for V, an orthonormal basis $\mathbf{f}_1, \ldots, \mathbf{f}_m$ for U, and non negative numbers $\sigma_1 \geq \sigma_1 \geq \cdots \geq \sigma_k \geq 0$, called the* singular values, *such that $L(\mathbf{u}_\ell) = \sigma_\ell \mathbf{v}_\ell$ for $\ell = 1, \ldots, k$ and $L(\mathbf{u}_\ell) = \mathbf{0}$ for $\ell = k+1, \ldots, n$.*

We see that $\sigma_1 = \max\{\|L(\mathbf{e})\| \mid \|\mathbf{e}\| = 1\}$ and that $\mathbf{e}_1$ realizes the maximum. We have in general that $\sigma_\ell = \max\{\|L(\mathbf{e})\| \mid \mathbf{e} \in \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_{\ell-1}\}^\perp \wedge \|\mathbf{e}\| = 1\}$ and that $\mathbf{e}_\ell$ realizes the maximum. The basis for V is simply given as $\mathbf{f}_\ell = \frac{L(\mathbf{e}_\ell)}{\|L(\mathbf{e}_\ell)\|}$ when $L(\mathbf{e}_\ell) \neq \mathbf{0}$. If this gives $\mathbf{f}_1, \ldots, \mathbf{f}_{k'}$ then the rest of the basis vectors are chosen as an orthonormal basis for $\operatorname{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_{k'}\}^\perp$. In terms of matrices it has the following formulation.

Theorem 2.12 *Let $\mathbf{A}$ be an $m \times n$ matrix and let $k = \min\{m, n\}$. Then $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, where $\mathbf{U}$ is an orthogonal $m \times m$ matrix, $\mathbf{V}$ is an orthogonal $n \times n$ matrix, and $\boldsymbol{\Sigma}$ is a diagonal matrix with non zero elements $\sigma_1 \geq \sigma_1 \geq \cdots \geq \sigma_k \geq 0$ in the diagonal.*

The singular values are the square root of the eigenvalues of $\mathbf{A}^T\mathbf{A}$, which is a positive semi definite symmetric matrix. The columns of $\mathbf{V}$, and hence the rows of $\mathbf{V}^T$, are the eigenvectors for $\mathbf{A}^T\mathbf{A}$.

Example 2.53

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 2.54

$$\begin{pmatrix} \sqrt{2} & 1 \\ 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \end{pmatrix}.$$

Example 2.55

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \end{pmatrix}.$$

Definition 2.23 *The Moore–Penrose pseudo inverse* of a matrix $\mathbf{A}$ is the matrix $\mathbf{A}^+ = \mathbf{V}\Sigma^+\mathbf{U}^T$ where $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ is the singular value decomposition of $\mathbf{A}$ and $\boldsymbol{\Sigma}^+$ is a diagonal matrix with $\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_k}$ in the diagonal. So $\mathbf{A}\mathbf{A}^+$ is a diagonal $m \times m$ matrix with $\underbrace{1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{m-k}$ in the diagonal and $\mathbf{A}^+\mathbf{A}$ is a diagonal $n \times n$ matrix with $\underbrace{1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{n-k}$ in the diagonal.

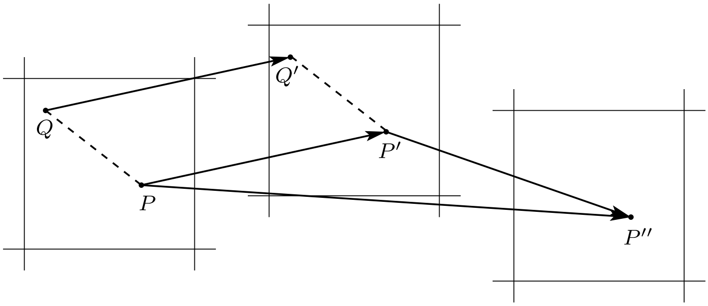

Fig. 2.4 There is a unique translation that maps a point P to another point P'. If it also maps Q to Q' then $\overrightarrow{PP'} = \overrightarrow{QQ'}$. Composition of translations corresponds to addition of vectors, $\overrightarrow{PP''} = \overrightarrow{PP'} + \overrightarrow{P'P''}$

Observe that the pseudo inverse of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma}^+$.

Example 2.56 If we have the equation $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ is the singular value decomposition of $\mathbf{A}$, then $\mathbf{U}$ and $\mathbf{V}^T$ are invertible, with inverse $\mathbf{U}^{-1} = \mathbf{U}^T$ and $\mathbf{V}^{T^{-1}} = \mathbf{V}$, respectively. We now have $\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{x} = \mathbf{U}^T\mathbf{b}$ and the best we can do is to let $\mathbf{V}^T\mathbf{x} = \boldsymbol{\Sigma}^+\mathbf{U}^T\mathbf{b}$ and hence $\mathbf{x} = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^T\mathbf{b} = \mathbf{A}^+\mathbf{b}$. If we have an overdetermined system we obtain the least square solution, i.e., the solution to the problem

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|^2. \tag{2.57}$$

If we have an underdetermined system we obtain the least norm solution, i.e., the solution to the problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|^2, \quad \text{such that } \mathbf{Ax} = \mathbf{b}. \tag{2.58}$$

2.2 Affine Spaces

We all know, at least intuitively, two affine spaces, namely the set of points in a plane and the set of points in space. If P and P' are two points in a plane then there is a *unique* translation of the plane that maps P to P', see Fig. 2.4. If the point Q is mapped to Q' then the vector from Q to Q' is the same as the vector from P to P', see Fig. 2.4. That is, we can identify the space of translation in the plane with the set of vectors in the plane. Under this identification addition of vectors corresponds to composition of translations, see Fig. 2.4. Even though we often identify our surrounding space with $\mathbb{R}^3$ and we can add elements of $\mathbb{R}^3$ it does obviously not make sense to add two points in space. The identification with $\mathbb{R}^3$ depends on the choice of coordinate system, and the result of adding the coordinates of two points depends on the choice of coordinate system, see Fig. 2.5.

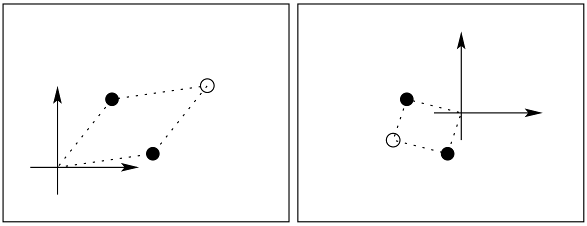

Fig. 2.5 If we add the coordinates of points in an affine space then the result depends on the choice of origin

What does make sense in the usual two dimensional plane and three dimensional space is the notion of translation along a vector $\mathbf{v}$. It is often written as adding a vector to a point, $\mathbf{x} \mapsto \mathbf{x} + \mathbf{v}$. An abstract affine space is a space where the notation of translation is defined and where this set of translations forms a vector space. Formally it can be defined as follows.

Definition 2.24 An affine space is a set X that admits a free transitive action of a vector space V. That is, there is a map $X \times V \to X : (\mathbf{x}, \mathbf{v}) \mapsto \mathbf{x} + \mathbf{v}$, called *translation* by the vector $\mathbf{v}$, such that

1. Addition of vectors corresponds to composition of translations, i.e., for all $\mathbf{x} \in X$ and $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{x} + (\mathbf{u} + \mathbf{v}) = (\mathbf{x} + \mathbf{u}) + \mathbf{v}$.
2. The zero vector acts as the identity, i.e., for all $\mathbf{x} \in X$, $\mathbf{x} + \mathbf{0} = \mathbf{x}$.
3. The action is free, i.e., if there for a given vector $\mathbf{v} \in V$ exists a point $\mathbf{x} \in X$ such that $\mathbf{x} + \mathbf{v} = \mathbf{x}$ then $\mathbf{v} = \mathbf{0}$.
4. The action is transitive, i.e., for all points $\mathbf{x}, \mathbf{y} \in X$ exists a vector $\mathbf{v} \in V$ such that $\mathbf{y} = \mathbf{x} + \mathbf{v}$.

The dimension of X is the dimension of the vector space of translations, V.

The vector $\mathbf{v}$ in Condition 4 that translates the point $\mathbf{x}$ to the point $\mathbf{y}$ is by Condition 3 unique, and is often written as $\mathbf{v} = \overrightarrow{\mathbf{xy}}$ or as $\mathbf{v} = \mathbf{y} - \mathbf{x}$. We have in fact a unique map $X \times X \to V : (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{y} - \mathbf{x}$ such that $\mathbf{y} = \mathbf{x} + (\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in X$. It furthermore satisfies

1. For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$, $\mathbf{z} - \mathbf{x} = (\mathbf{z} - \mathbf{y}) + (\mathbf{y} - \mathbf{x})$.
2. For all $\mathbf{x}, \mathbf{y} \in X$ and $\mathbf{u}, \mathbf{v} \in V$, $(\mathbf{y} + \mathbf{v}) - (\mathbf{x} + \mathbf{u}) = (\mathbf{y} - \mathbf{x}) + \mathbf{v} - \mathbf{u}$.
3. For all $\mathbf{x} \in X$, $\mathbf{x} - \mathbf{x} = \mathbf{0}$.
4. For all $\mathbf{x}, \mathbf{y} \in X$, $\mathbf{y} - \mathbf{x} = -(\mathbf{x} - \mathbf{y})$.

Example 2.57 The usual two dimensional plane and three dimensional space are affine spaces and the vector space of translations is the space of vectors in the plane or in space.

Example 2.58 If the set of solutions to k real inhomogeneous linear equations in n unknowns is non empty then it is an affine space and the vector space of translations is the space of solutions to the corresponding set of homogeneous equations.

Example 2.59 If (X, U) and (Y, V) are affine spaces then $(X \times Y, U \times V)$ is an affine space with translation defined by $(\mathbf{x}, \mathbf{y}) + (\mathbf{u}, \mathbf{v}) = (\mathbf{x} + \mathbf{u}, \mathbf{y} + \mathbf{v})$.

A *coordinate system* in an affine space (X, V) consists of a point $O \in X$, called the origin, and a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ for V. Any point $\mathbf{x} \in X$ can now be written as

$$\mathbf{x} = O + (\mathbf{x} - O) = O + \sum_{k=1}^{n} x_k \mathbf{v}_k, \tag{2.59}$$

where the numbers $x_1, \dots, x_n$ are the coordinates for the vector $\mathbf{x} - O$ with respect to the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$, they are now also called the coordinates for $\mathbf{x}$ with respect to the coordinate system $O, \mathbf{v}_1, \dots, \mathbf{v}_n$.

2.2.1 Affine and Convex Combinations

We have already noticed that it does not make sense to add points in an affine space, or more generally to take linear combination of points, see Fig. 2.5. So when a coordinate system is chosen it is important to be careful. It is of course possible to add the coordinates of two points and regard the result as the coordinates for a third point. But it is not meaningful. In fact, by changing the origin we can obtain any point by such a calculation.

But even though linear combination does not make sense, affine combination does.

Definition 2.25 A formal sum $\sum_{\ell=1}^{k} \alpha_\ell \mathbf{x}_\ell$ of k points $\mathbf{x}_1, \dots, \mathbf{x}_k$ is called an *affine combination* if the coefficients sum to 1, i.e., if $\sum_{\ell=1}^{k} \alpha_\ell = 1$. Then we have

$$\sum_{\ell=1}^{k} \alpha_\ell \mathbf{x}_\ell = O + \sum_{\ell=1}^{k} \alpha_\ell (\mathbf{x}_\ell - O), \tag{2.60}$$

where $O \in X$ is an arbitrary chosen point.

Observe that in the last sum we have a linear combination of vectors so the expression makes sense. If we choose an other point O' then the vector between the two results are

$$\begin{aligned}
&\left(O + \sum_{\ell=1}^{k} \alpha_\ell (\mathbf{x}_\ell - O)\right) - \left(O' + \sum_{\ell=1}^{k} \alpha_\ell \bigl(\mathbf{x}_\ell - O'\bigr)\right) \\
&= \bigl(O - O'\bigr) + \sum_{\ell=1}^{k} \alpha_\ell \bigl((\mathbf{x}_\ell - O) - \bigl(\mathbf{x}_\ell - O'\bigr)\bigr) \\
&= \bigl(O - O'\bigr) + \sum_{\ell=1}^{k} \alpha_\ell \bigl((\mathbf{x}_\ell - \mathbf{x}_\ell) + \bigl(O - O'\bigr)\bigr)
\end{aligned}$$

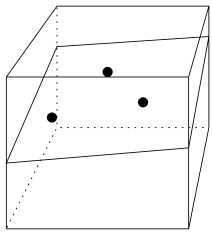

Fig. 2.6 The plane spanned by three points

$$= (O - O') - \left(\sum_{\ell=1}^{k} \alpha_\ell\right)(O - O')$$
$$= (O - O') - (O - O') = \mathbf{0}. \tag{2.61}$$

That is, the result does not depend on the auxiliary point O.

Example 2.60 The line spanned by two different points $\mathbf{x}$ and $\mathbf{y}$ in an affine space consists of affine combinations of the two points, that is, the points $(1-t)\mathbf{x} + t\mathbf{y} = \mathbf{x} + t(\mathbf{y} - \mathbf{x})$, $t \in \mathbb{R}$.

Example 2.61 The plane spanned by three points in space (not on the same line) consists of all affine combinations of the three points, see Fig. 2.6.

Unless the vector space of translations is equipped with an inner product there is no notion of lengths in an affine space. But for points on a line the *ratio of lengths* makes sense. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2$ be four points on a line and choose a non zero vector $\mathbf{v}$ on the line, e.g., the difference between two of the given points. Then there exist numbers $t_1, t_2 \in \mathbb{R}$ such that we have $\mathbf{y}_1 - \mathbf{x}_1 = t_1\mathbf{v}$ and $\mathbf{y}_2 - \mathbf{x}_2 = t_2\mathbf{v}$. The ratio between the line segments $\mathbf{x}_1\mathbf{y}_2$ and $\mathbf{x}_2\mathbf{y}_2$ is now defined as $\frac{t_1}{t_2}$. If we had chosen another vector $\mathbf{w}$ then $\mathbf{v} = \alpha\mathbf{w}$ and $\mathbf{y}_k - \mathbf{x}_k = t_k\alpha\mathbf{w}$ and the ratio $\frac{\alpha t_1}{\alpha t_2} = \frac{t_1}{t_2}$ is the same. Observe that we even have a well defined signed ratio.

Definition 2.26 A *convex combination* of points $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is an affine combination $\sum_{\ell=1}^{k} \alpha_\ell \mathbf{x}_\ell$ where all the coefficients are non negative, i.e., $\alpha_\ell \geq 0$ for all $\ell = 1, \ldots, k$.

Example 2.62 The line segment between two points consists of all convex combination of the two points.

Let X be an affine space of dimension n and let $\mathbf{x}_0, \ldots, \mathbf{x}_n$ be $n+1$ points that are affinely independent, i.e., none of the points can be written as an affine combination of the others. This is equivalent to the vectors $\mathbf{x}_1 - \mathbf{x}_0, \ldots, \mathbf{x}_n - \mathbf{x}_0$ being linearly independent. Then any point $\mathbf{y}$ in X can be written uniquely as an affine combination of the given points, $\mathbf{y} = \sum_{k=0}^{n} \alpha_k \mathbf{x}_k$. The numbers $\alpha_0, \ldots, \alpha_n$ are called *barycentric coordinates* for $\mathbf{y}$ with respect to the points $\mathbf{x}_0, \ldots, \mathbf{x}_n$. The case $n = 2$ is illustrated

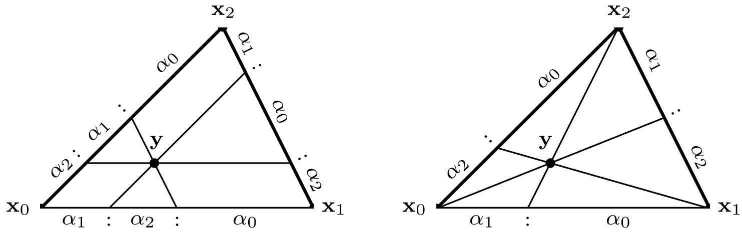

Fig. 2.7 The barycentric coordinates of a point $\mathbf{y} = \alpha_0 \mathbf{x}_0 + \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2$ can be given in terms of the ratio between different line segments

in Fig. 2.7 where two different geometric interpretations of the barycentric coordinates are shown.

2.2.2 Affine Maps

Definition 2.27 An affine map between two affine spaces X and Y is a map $f : X \to Y$ that preserves affine combinations, i.e.,

$$f\left(\sum_{\ell=1}^{k} \alpha_k \mathbf{x}_k\right) = \sum_{\ell=1}^{k} \alpha_k f(\mathbf{x}_k). \tag{2.62}$$

There is a close connection between affine maps between X and Y and linear maps between their vector spaces of translations U and V. More precisely we have the following proposition.

Proposition 2.3 *Let f be an affine map between two affine spaces (X, U) and (Y, V). Then there is a unique linear map $L : U \to V$ such that $f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + L(\mathbf{v})$ for all $\mathbf{x} \in X$ and $\mathbf{u} \in U$.*

We see that $L(\mathbf{v}) = f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x})$ and it turns out that this expression does not depend on $\mathbf{x}$ and that L is linear. If we now choose an origin $O \in X$ and $O' \in Y$ then

$$f(O + \mathbf{v}) = O' + \big(f(O) - O'\big) + L(\mathbf{v}) = O' + \mathbf{y} + L(\mathbf{v}). \tag{2.63}$$

and we see that f is the sum of the linear map L and the translation defined by $\mathbf{y} = f(O) - O' \in V$.

Definition 2.28 A hyperplane in an affine space X is a subset, H, of the form

$$H = f^{-1} = \big\{x \in X \mid f(x) = c\big\}, \tag{2.64}$$

where $f : X \to \mathbb{R}$ is affine and $c \in \mathbb{R}$.

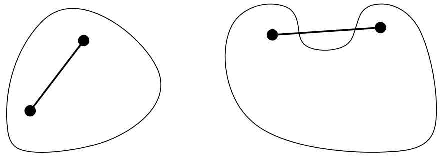

Fig. 2.8 To the left a convex set and to the right a non convex set

If $x_1, \ldots, x_n$ are coordinates for points in X with respect to some coordinate system, then a hyperplane is given by an equation of the form

$$a_1 x_1 + \cdots + a_n x_n = c. \tag{2.65}$$

A half space is defined in a similar manner.

Definition 2.29 A half space in an affine space X is a subset, H, of the form

$$H = f^{-1} = \{x \in X \mid f(x) \geq c\}, \tag{2.66}$$

where $f : X \to \mathbb{R}$ is affine and $c \in \mathbb{R}$.

If $x_1, \ldots, x_n$ are coordinates for points in X with respect to some coordinate system, then a half space is given by an equation of the form

$$a_1 x_1 + \cdots + a_n x_n \geq c. \tag{2.67}$$

2.2.3 Convex Sets

Definition 2.30 A subset $C \subseteq X$ of an affine space is called *convex* if for each pair of points in C the line segment between the points are in C, see Fig. 2.8.

Definition 2.31 Let $A \subseteq X$ be an arbitrary subset of an affine space X. The *convex hull* of A is the smallest convex set containing A and is denoted $CH(A)$.

There are alternative, equivalent, definitions of the convex hull.

1. The convex hull is the intersection of all convex sets containing A:

$$CH(A) = \bigcap_{\substack{A \subseteq C \\ C \text{ is convex}}} C. \tag{2.68}$$

2. The convex hull is the intersection of all half spaces containing A:

$$CH(A) = \bigcap_{\substack{A \subseteq H \\ H \text{ is a half space}}} H. \tag{2.69}$$

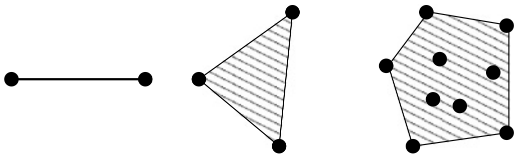

Fig. 2.9 All convex combinations of two, three, and nine points

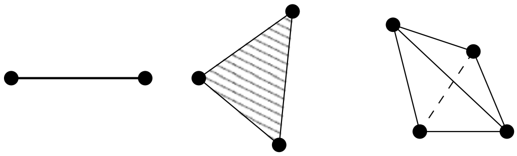

Fig. 2.10 From left to right, a simplex in dimension 1, 2, and 3

3. The convex hull is the set of all convex combinations of points in A:

$$CH(A) = \left\{ \sum_{\ell=1}^{k} \alpha_k \mathbf{x}_k \;\middle|\; \sum_{\ell=1}^{k} \alpha_k = 1 \wedge \alpha_1, \ldots, \alpha_k \geq 0 \wedge \mathbf{x}_1, \ldots, \mathbf{x}_k \in A \right\}, \quad (2.70)$$

see Fig. 2.9

Definition 2.32 A *simplex* is the convex hull of $n+1$ affinely independent points in a n dimensional affine space, see Fig. 2.10

Example 2.63 The convex hull of two points is the line segment between the two points, a two simplex.

Example 2.64 The convex hull of three points is a triangle, and its interior a three simplex.

Example 2.65 The convex hull of the unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ is the closed disk $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

2.3 Metric Spaces

A metric space is a space where an abstract notion of distance is defined. When we have such a notion we can define continuity of mappings between metric spaces, the notion of convergence and of open and closed sets, and the notion of neighborhoods of a point.

Definition 2.33 A metric space (X, d) is a set X equipped with a map $d : X \times X \to \mathbb{R}$ that satisfies the following three conditions:
1. Symmetry, for all $x, y \in X$: $d(x, y) = d(y, x)$.
2. The triangle inequality, for all $x, y, z \in X$: $d(x, z) \leq d(x, y) + d(y, z)$.

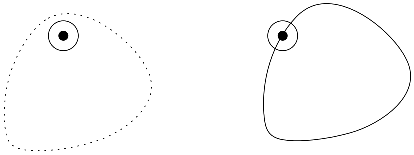

Fig. 2.11 To the *left* an open set, there is room for a ball around each point. To the *right* a non open set, any ball around a point on the boundary is not contained in the set

3. Positivity, for all $x, y \in X$: $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$.

Example 2.66 if $V, \langle\cdot,\cdot\rangle$ is an Euclidean vector space and $\|\cdot\|$ is the corresponding norm, then V equipped with the distance $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{v} - \mathbf{u}\|$ is a metric space.

Example 2.67 If (X, V) is an affine space and V is an Euclidean vector space with norm $\|\cdot\|$, then X equipped with the distance $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{y} - \mathbf{x}\|$ is a metric space.

Example 2.68 If $Y \subseteq X$ is a subset of a metric space (X, d), then Y equipped with the restriction of d to $Y \times Y$ is a metric space.

Example 2.69 If (X_1, d_1) and (X_2, d_2) are metric spaces then the Cartesian product $X_1 \times X_2$ equipped with the distance $d((\mathbf{x}_1, \mathbf{x}_2), (\mathbf{y}_1, \mathbf{y}_2)) = d_1(\mathbf{x}_1, \mathbf{y}_1) + d_2(\mathbf{x}_2, \mathbf{y}_2)$ is a metric space.

Example 2.70 If X is an arbitrary set and we define d by

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y, \end{cases} \tag{2.71}$$

then (X, d) is a metric space. This metric is called the *discrete metric*.

Definition 2.34 Let (X, d) be a metric space. The *open ball* with radius $r > 0$ and center $x \in X$ is the set $B(x, r) = \{y \in X \mid d(x, y) < r\}$.

Example 2.71 If X is equipped with the discrete metric and $x \in X$ is an arbitrary point then

$$B(x, r) = \begin{cases} \{x\} & \text{if } r \leq 1, \\ X & \text{if } r > 1. \end{cases}$$

Definition 2.35 Let X be a metric space. A subset $U \subseteq X$ is called an *open set* if there for all points $x \in U$ exists an open ball $B(x, r) \subseteq U$, see Fig. 2.11.

Example 2.72 If X is equipped with the discrete metric then all subsets are open.

Theorem 2.13 *If X is a metric space then the set of open sets has the following three properties*:

1. *The empty set* $\emptyset$ *and the whole space* X *are open sets.*
2. *If* U_i, $i \in I$ *is an arbitrary collection of open sets then their union* $\bigcup_{i\in I} U_i$ *is an open set.*
3. *If* $U_1, \dots, U_n$ *is a finite collection of open sets then their intersection* $U_1 \cap \cdots \cap U_n$ *is an open set.*

The three properties above are the defining properties of a topological space which is a more general concept. There exist many topological spaces that are not induced by a metric.

Definition 2.36 Let X be a metric space. A subset $F \subseteq X$ is called a *closed set* if the complement $X \setminus F$ is open.

Example 2.73 If X is equipped with the discrete metric then all subsets are closed.

Theorem 2.13 implies that the closed sets have the following three properties:

1. The empty set $\emptyset$ and the whole space X are closed sets.
2. If F_i, $i \in I$ is an arbitrary collection of closed sets then their intersection $\bigcap_{i\in I} F_i$ is a closed set.
3. If $F_1, \dots, F_n$ is a finite collection of closed sets then their union $F_1 \cup \cdots \cup F_n$ is an closed set.

Definition 2.37 Let $A \subseteq X$ be a subset of a metric space. The *interior* A° of A is the largest open set contained in A. The *closure* $\overline{A}$ is the smallest closed set containing A.

The interior and closure of a subset A of X can equivalently be defined as

$$A^\circ = \left\{x \in A \mid \exists r > 0 : B(x,r) \subseteq A\right\}, \tag{2.72}$$

$$\overline{A} = \left\{x \in X \mid \forall r > 0 : B(x,r) \cap A \neq \emptyset\right\}. \tag{2.73}$$

In other words all points in the interior has a surrounding ball contained in A, and all balls centered at points in the closure intersects A.

A subset A of a metric space X is a *neighborhood* of a set $B \subseteq X$ if $B \in A^\circ$, i.e., if and only if there for each $x \in Y$ exists an $r > 0$ such that $B(x,r) \subseteq A$.

Definition 2.38 A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) is called *convergent* with *limit* x if

$$\forall \epsilon > 0\ \exists n_0 \in \mathbb{N} : n > n_0 \implies d(x, x_n) < \epsilon. \tag{2.74}$$

We write $x_n \to x$ for $n \to \infty$ or $\lim_{n\to\infty} x_n = x$. The formal definition of continuity is as follows.

Definition 2.39 Let (X, d) and (Y, d') be metric spaces. A map $f : X \to Y$ is a *continuous* in a point $x \in X$ if

$$\forall \epsilon > 0\ \exists \delta > 0 : \forall y \in X : d(x, y) < \delta \implies d'\big(f(x), f(y)\big) < \epsilon.$$

A map $f : X \to Y$ is a *continuous map* if it is continuous at all points of X and it is a *homeomorphism* if it is bijective, continuous and the inverse $f^{-1} : Y \to X$ is continuous too.

There are alternative definitions of continuity.

Theorem 2.14 *Let (X, d) and (Y, d') be metric spaces. A map $f : X \to Y$ is a continuous map if and only if for all convergent sequences $(x_n)_{n\in\mathbb{N}}$ in X, the sequence $(f(x_n))_{n\in\mathbb{N}}$ is convergent in Y and* $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

Theorem 2.15 *Let (X, d) and (Y, d') be metric spaces. A map $f : X \to Y$ is a continuous map if and only if for all open set $U \subseteq Y$ the preimage $f^{-1}(U) = \{x \in X \mid f(x) \in U\}$ is an open set in X.*

The last concept we need is compactness.

Definition 2.40 A subset C of a metric space X is called compact if always when C is covered by a collection of open sets, i.e., $C \subseteq \bigcup_{i\in I} U_i$, where U_i is open for all $i \in I$, then there exists a finite number of $U_{i_1}, \dots, U_{i_n}$ of the given open sets that cover C, i.e., $C \subseteq U_{i_1} \cup \cdots \cup U_{i_n}$.

There is an alternative definition of compact sets.

Theorem 2.16 *A subset C of a metric space X is compact if and only if each sequence $(x_n)_{n\in\mathbb{N}}$ in C has a convergent sub sequence $(x_{n_k})_{k\in\mathbb{N}}$.*

Theorem 2.17 *If C is a compact subset of a metric space X then C is closed and bounded.*

If $X = \mathbb{R}^n$ then the converse is true.

Theorem 2.18 *A subset C of $\mathbb{R}^n$ is compact if and only if C is closed and bounded.*

One of the important properties of compact sets is the following result.

Theorem 2.19 *A continuous function $f : C \to \mathbb{R}$ on a compact set has a minimum and a maximum, i.e., there exist $x_0, x_1 \in C$ such that $f(x_0) \le f(x) \le f(x_1)$ for all $x \in C$.*

2.4 Exercises

Exercise 2.1 Prove Proposition 2.1.

Exercise 2.2 Prove that the examples in Examples 2.1–2.11 are vector spaces.

Exercise 2.3 Prove that if V is a vector space and $U \subseteq V$ satisfies the conditions in Definition 2.2 then U is a vector space.

Exercise 2.4 Prove that the examples in Examples 2.13–2.19 are subspaces.

Exercise 2.5 Prove Corollary 2.1.

Exercise 2.6 Show that the monomials as well as the Bernstein polynomials are a basis, cf. Examples 2.29 and 2.30.

Exercise 2.7 Show that the map $\mathbb{R} \to \mathbb{R} : x \mapsto ax + b$ is linear if and only if $b = 0$.

Exercise 2.8 Prove that the maps in Examples 2.31–2.37 are linear.

Exercise 2.9 Prove that the spaces in Examples 2.38–2.43 are Euclidean vector spaces.

Exercise 2.10 Prove the Cauchy–Schwartz inequality, Theorem 2.6. Hint: first note that the theorem is trivial if one of the vectors is the zero vector. Next, use $\langle \mathbf{u} - \alpha\mathbf{v}, \mathbf{u} - \alpha\mathbf{v} \rangle = \|\mathbf{u} - \alpha\mathbf{v}\|^2 \geq 0$ for all $\alpha \in \mathbb{R}$ and find the α that minimize the expression.

Exercise 2.11 Prove the statements in Example 2.44.

Exercise 2.12 Prove that the map in Example 2.47 is symmetric.

Exercise 2.13 Prove the statements in Examples 2.49 and 2.50.

Exercise 2.14 Prove Theorem 2.9. Hint: use Theorem 2.8.

Exercise 2.15 Prove Theorem 2.12. Hint: use Theorem 2.11.

Exercise 2.16 Prove the statements in Example 2.56.

Exercise 2.17 Prove that the spaces in Examples 2.57–2.59 are affine spaces.

Exercise 2.18 Prove Proposition 2.3.

Exercise 2.19 Determine the convex hull of the set

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid x = 0 \wedge y > 0\right\} \cup \left\{(x, y) \in \mathbb{R}^2 \mid x > 0 \wedge xy = 1\right\}.$$

Exercise 2.20 Prove that the spaces in Examples 2.66–2.70 are metric spaces.

Exercise 2.21 Prove the statements in Examples 2.71, 2.72, and 2.73.

Exercise 2.22 Let $\mathbf{A}$ be a real $n \times m$ matrix of rank $m \leq n$, let $\mathbf{x} \in \mathbb{R}^m$, and let $\mathbf{b} \in \mathbb{R}^n$. What is the solution to

$$\min_{\mathbf{x}} f(\mathbf{x})|_{\|\mathbf{x}\|=1},$$

where

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A}^T \mathbf{A}\mathbf{x}?$$

Hint: try first the case where $\mathbf{A}$ is a 2×2-matrix, e.g., $\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$.

Exercise 2.23 What are the solutions to

$$\max_{\mathbf{x}} f(\mathbf{x}),$$

where

$$f(\mathbf{x}) = \frac{\mathbf{b}^T\mathbf{x}}{\|\mathbf{b}\|\,\|\mathbf{x}\|}?$$

Exercise 2.24 What geometric object do the points $\mathbf{x} \in \mathbb{R}^3$, fulfilling the equation

$$\mathbf{n}^T\mathbf{x} = \alpha,$$

describe? Here $\mathbf{n} \in \mathbb{R}^3$ and $\alpha \in \mathbb{R}$. Please explain.

References

1. Strang, G.: Linear Algebra and Its Applications, 4th edn. Brooks Cole (2006)
2. Hansen, V.L.: Fundamental Concepts in Modern Analysis. World Scientific, River Edge (1999)

http://www.springer.com/978-1-4471-4074-0

Guide to Computational Geometry Processing
Foundations, Algorithms, and Methods
Bærentzen, J.A.; Gravesen, J.; Anton, F.; Aanæs, H.
2012, XVIII, 326 p., Hardcover
ISBN: 978-1-4471-4074-0