Chapter 12
Euclidean and Unitary Vector Spaces

In this chapter we study vector spaces over the fields $\mathbb{R}$ and $\mathbb{C}$. Using the definition of bilinear and sesquilinear forms, we introduce scalar products on such vector spaces. Scalar products allow the extension of well-known concepts from elementary geometry, such as length and angles, to abstract real and complex vector spaces. This, in particular, leads to the idea of orthogonality and to orthonormal bases of vector spaces. As an example for the importance of these concepts in many applications we study least-squares approximations.

12.1 Scalar Products and Norms

We start with the definition of a scalar product and the Euclidean or unitary vector spaces.

Definition 12.1 Let $\mathcal{V}$ be a K-vector space, where either $K = \mathbb{R}$ or $K = \mathbb{C}$. A map

$$\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to K, \quad (v, w) \mapsto \langle v, w \rangle,$$

is called a *scalar product* on $\mathcal{V}$, when the following properties hold:

(1) If $K = \mathbb{R}$, then $\langle \cdot, \cdot \rangle$ is a symmetric bilinear form.
If $K = \mathbb{C}$, then $\langle \cdot, \cdot \rangle$ is an Hermitian sesquilinear form.
(2) $\langle \cdot, \cdot \rangle$ is *positive definite*, i.e., $\langle v, v \rangle \geq 0$ holds for all $v \in \mathcal{V}$, with equality if and only if $v = 0$.

An $\mathbb{R}$-vector space with a scalar product is called a *Euclidean vector space*[1], and a $\mathbb{C}$-vector space with a scalar product is called a *unitary vector space*.

Scalar products are sometimes called *inner products*. Note that $\langle v, v \rangle$ is nonnegative and *real* also when $\mathcal{V}$ is a $\mathbb{C}$-vector space. It is easy to see that a subspace $\mathcal{U}$ of

[1]Euclid of Alexandria (approx. 300 BC).

J. Liesen and V. Mehrmann, *Linear Algebra*, Springer Undergraduate Mathematics Series, DOI 10.1007/978-3-319-24346-7_12

a Euclidean or unitary vector space $\mathcal{V}$ is again a Euclidean or unitary vector space, respectively, when the scalar product on the space $\mathcal{V}$ is restricted to the subspace $\mathcal{U}$.

Example 12.2

(1) A scalar product on $\mathbb{R}^{n,1}$ is given by

$$\langle v, w \rangle := w^T v.$$

It is called the *standard scalar product of* $\mathbb{R}^{n,1}$.

(2) A scalar product on $\mathbb{C}^{n,1}$ is given by

$$\langle v, w \rangle := w^H v.$$

It is called the *standard scalar product of* $\mathbb{C}^{n,1}$.

(3) For both $K = \mathbb{R}$ and $K = \mathbb{C}$,

$$\langle A, B \rangle := \mathrm{Spur}(B^H A)$$

is a scalar product on $K^{n,m}$.

(4) A scalar product on the vector space of the continuous and real valued functions on the real interval $[\alpha, \beta]$ is given by

$$\langle f, g \rangle := \int_\alpha^\beta f(x)g(x)dx.$$

We will now show how to use the Euclidean or unitary structure of a vector space in order to introduce geometric concepts such as the length of a vector or the angle between vectors.

As a motivation of a general concept of length we have the *absolute value* of real numbers, i.e., the map $|\cdot| : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$. This map has the following properties:

(1) $|\lambda x| = |\lambda| \cdot |x|$ for all $\lambda, x \in \mathbb{R}$.
(2) $|x| \geq 0$ for all $x \in \mathbb{R}$, with equality if and only if $x = 0$.
(3) $|x + y| \leq |x| + |y|$ for all $x, y \in \mathbb{R}$.

These properties are generalized to real or complex vector spaces as follows.

Definition 12.3 Let $\mathcal{V}$ be a K-vector space, where either $K = \mathbb{R}$ or $K = \mathbb{C}$. A map

$$\|\cdot\| : \mathcal{V} \to \mathbb{R}, \quad v \mapsto \|v\|,$$

is called a *norm* on $\mathcal{V}$, when for all $v, w \in \mathcal{V}$ and $\lambda \in K$ the following properties hold:

(1) $\|\lambda v\| = |\lambda| \cdot \|v\|$.
(2) $\|v\| \geq 0$, with equality if and only if $v = 0$.
(3) $\|v + w\| \leq \|v\| + \|w\|$ (triangle inequality).

A K-vector space on which a norm is defined is called a *normed space*.

Example 12.4

(1) If $\langle \cdot, \cdot \rangle$ is the standard scalar product on $\mathbb{R}^{n,1}$, then

$$\|v\| := \langle v, v \rangle^{1/2} = (v^T v)^{1/2}$$

defines a norm that is called the *Euclidean norm of* $\mathbb{R}^{n,1}$.

(2) If $\langle \cdot, \cdot \rangle$ is the standard scalar product on $\mathbb{C}^{n,1}$, then

$$\|v\| := \langle v, v \rangle^{1/2} = (v^H v)^{1/2}$$

defines a norm that is called the *Euclidean norm of* $\mathbb{C}^{n,1}$. (This is common terminology, although the space itself is unitary and not Euclidean.)

(3) For both $K = \mathbb{R}$ and $K = \mathbb{C}$,

$$\|A\|_F := (\text{trace}(A^H A))^{1/2} = \Big(\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2 \Big)^{1/2}$$

is a norm on $K^{n,m}$ that is called the *Frobenius norm*[2] *of* $K^{n,m}$. For $m = 1$ the Frobenius norm is equal to the Euclidean norm of $K^{n,1}$. Moreover, the Frobenius norm of $K^{n,m}$ is equal to the Euclidean norm of $K^{nm,1}$ (or K^{nm}), if we identify these vector spaces via an isomorphism.
Obviously, we have $\|A\|_F = \|A^T\|_F = \|A^H\|_F$ for all $A \in K^{n,m}$.

(4) If $\mathcal{V}$ is the vector space of the continuous and real valued functions on the real interval $[\alpha, \beta]$, then

$$\|f\| := \langle f, f \rangle^{1/2} = \Big(\int_{\alpha}^{\beta} (f(x))^2 dx \Big)^{1/2}$$

is a norm on $\mathcal{V}$ that is called the L^2-*norm*.

(5) Let $K = \mathbb{R}$ or $K = \mathbb{C}$, and let $p \in \mathbb{R}$, $p \geq 1$ be given. Then for $v = [\nu_1, \ldots, \nu_n]^T \in K^{n,1}$ the *p-norm of* $K^{n,1}$ is defined by

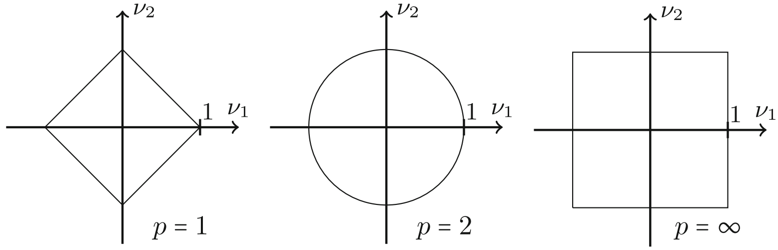
$$\|v\|_p := \Big(\sum_{i=1}^{n} |\nu_i|^p \Big)^{1/p}. \tag{12.1}$$

[2]Ferdinand Georg Frobenius (1849–1917).

For $p = 2$ this is the Euclidean norm on $K^{n,1}$. For this norm we typically omit the index 2 and write $\|\cdot\|$ instead of $\|\cdot\|_2$ (as in (1) and (2) above). Taking the limit $p \to \infty$ in (12.1), we obtain the *∞-norm of $K^{n,1}$*, given by

$$\|v\|_\infty = \max_{1 \le i \le n} |\nu_i|.$$

The following figures illustrate the unit circle in $\mathbb{R}^{2,1}$ with respect to the p-norm, i.e., the set of all $v \in \mathbb{R}^{2,1}$ with $\|v\|_p = 1$, for $p = 1$, $p = 2$ and $p = \infty$:



(6) For $K = \mathbb{R}$ or $K = \mathbb{C}$ the *p-norm of $K^{n,m}$* is defined by

$$\|A\|_p := \sup_{v \in K^{m,1} \setminus \{0\}} \frac{\|Av\|_p}{\|v\|_p}.$$

Here we use the p-norm of $K^{m,1}$ in the denominator and the p-norm of $K^{n,1}$ in the numerator. The notation sup means *supremum*, i.e., the least upper bound that is known from Analysis. One can show that the supremum is attained by a vector v, and thus we may write max instead of sup in the definition above. In particular, for $A = [a_{ij}] \in K^{n,m}$ we have

$$\|A\|_1 = \max_{1 \le j \le m} \sum_{i=1}^{n} |a_{ij}|,$$

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{m} |a_{ij}|.$$

These norms are called *maximum column sum* and *maximum row sum norm of $K^{n,m}$*, respectively. We easily see that $\|A\|_1 = \|A^T\|_\infty = \|A^H\|_\infty$ and $\|A\|_\infty = \|A^T\|_1 = \|A^H\|_1$. However, for the matrix

$$A = \begin{bmatrix} 1/2 & -1/4 \\ -1/2 & 2/3 \end{bmatrix} \in \mathbb{R}^{2,2}$$

we have $\|A\|_1 = 1$ and $\|A\|_\infty = 7/6$. Thus, this matrix A satisfies $\|A\|_1 < \|A\|_\infty$ and $\|A^T\|_\infty < \|A^T\|_1$. The 2-norm of matrices will be considered further in Chap. 19.

The norms in the above examples (1)–(4) have the form $\|v\| = \langle v, v\rangle^{1/2}$, where $\langle\cdot,\cdot\rangle$ is a given scalar product. We will show now that the map $v \mapsto \langle v, v\rangle^{1/2}$ always defines a norm. Our proof is based on the following theorem.

Theorem 12.5 *If $\mathcal{V}$ is a Euclidean or unitary vector space with the scalar product $\langle\cdot,\cdot\rangle$, then*

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle \cdot \langle w, w\rangle \quad \textit{for all } v, w \in \mathcal{V}, \tag{12.2}$$

with equality if and only if v, w are linearly dependent.

Proof The inequality is trivial for $w = 0$. Thus, let $w \neq 0$ and let

$$\lambda := \frac{\langle v, w\rangle}{\langle w, w\rangle}.$$

Then

$$\begin{aligned}
0 \leq \langle v - \lambda w, v - \lambda w\rangle &= \langle v, v\rangle - \overline{\lambda}\langle v, w\rangle - \lambda\langle w, v\rangle - \lambda(-\overline{\lambda})\langle w, w\rangle \\
&= \langle v, v\rangle - \frac{\overline{\langle v, w\rangle}}{\langle w, w\rangle}\langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\overline{\langle v, w\rangle} + \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle^2}\langle w, w\rangle \\
&= \langle v, v\rangle - \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle},
\end{aligned}$$

which implies (12.2).

If v, w are linearly dependent, then $v = \lambda w$ for a scalar λ, and hence

$$\begin{aligned}
|\langle v, w\rangle|^2 &= |\langle \lambda w, w\rangle|^2 = |\lambda\langle w, w\rangle|^2 = |\lambda|^2|\langle w, w\rangle|^2 = \lambda\overline{\lambda}\,\langle w, w\rangle\,\langle w, w\rangle \\
&= \langle \lambda w, \lambda w\rangle\,\langle w, w\rangle = \langle v, v\rangle\,\langle w, w\rangle.
\end{aligned}$$

On the other hand, let $|\langle v, w\rangle|^2 = \langle v, v\rangle\langle w, w\rangle$. If $w = 0$, then v, w are linearly dependent. If $w \neq 0$, then we define λ as above and get

$$\langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle - \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle} = 0.$$

Since the scalar product is positive definite, we have $v - \lambda w = 0$, and thus v, w are linearly dependent. □

The inequality (12.2) is called *Cauchy-Schwarz inequality*.[3] It is an important tool in Analysis, in particular in the estimation of approximation and interpolation errors.

[3] Augustin Louis Cauchy (1789–1857) and Hermann Amandus Schwarz (1843–1921).

Corollary 12.6 *If $\mathcal{V}$ is a Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$, then the map*

$$\| \cdot \| : \mathcal{V} \to \mathbb{R}, \quad v \mapsto \|v\| := \langle v, v \rangle^{1/2},$$

is a norm on $\mathcal{V}$ that is called the norm induced by the scalar product.

Proof We have to prove the three defining properties of the norm. Since $\langle \cdot, \cdot \rangle$ is positive definite, we have $\|v\| \geq 0$, with equality if and only if $v = 0$. If $v \in \mathcal{V}$ and $\lambda \in K$ (where in the Euclidean case $K = \mathbb{R}$ and in the unitary case $K = \mathbb{C}$), then

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \overline{\lambda} \langle v, v \rangle = |\lambda|^2 \langle v, v \rangle,$$

and hence $\|\lambda v\| = |\lambda| \, \|v\|$. In order to show the triangle inequality, we use the Cauchy-Schwarz inequality and the fact that $\mathrm{Re}(z) \leq |z|$ for every complex number z. For all $v, w \in \mathcal{V}$ we have

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\
&= \langle v, v \rangle + \langle v, w \rangle + \overline{\langle v, w \rangle} + \langle w, w \rangle \\
&= \|v\|^2 + 2\,\mathrm{Re}(\langle v, w \rangle) + \|w\|^2 \\
&\leq \|v\|^2 + 2\,|\langle v, w \rangle| + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\| \, \|w\| + \|w\|^2 \\
&= (\|v\| + \|w\|)^2,
\end{aligned}$$

and thus $\|v + w\| \leq \|v\| + \|w\|$. □

12.2 Orthogonality

We will now use the scalar product to introduce angles between vectors. As motivation we consider the Euclidean vector space $\mathbb{R}^{2,1}$ with the standard scalar product and the induced Euclidean norm $\|v\| = \langle v, v \rangle^{1/2}$. The Cauchy-Schwarz inequality shows that

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \, \|w\|} \leq 1 \quad \text{for all } v, w \in \mathbb{R}^{2,1} \setminus \{0\}.$$

If $v, w \in \mathbb{R}^{2,1} \setminus \{0\}$, then the angle between v and w is the uniquely determined real number $\varphi \in [0, \pi]$ with

$$\cos(\varphi) = \frac{\langle v, w \rangle}{\|v\| \, \|w\|}.$$

The vectors v, w are orthogonal if $\varphi = \pi/2$, so that $\cos(\varphi) = 0$. Thus, v, w are orthogonal if and only if $\langle v, w\rangle = 0$.

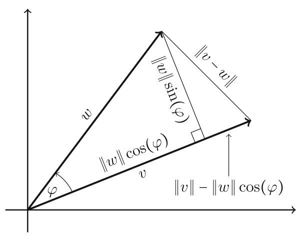
An elementary calculation now leads to the *cosine theorem for triangles*:

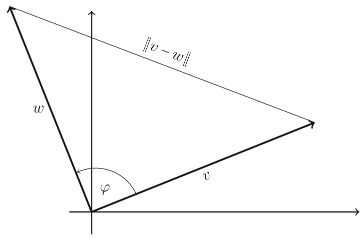
$$\|v - w\|^2 = \langle v - w, v - w\rangle = \langle v, v\rangle - 2\langle v, w\rangle + \langle w, w\rangle$$
$$= \|v\|^2 + \|w\|^2 - 2\|v\|\,\|w\|\cos(\varphi).$$

If v, w are orthogonal, i.e., $\langle v, w\rangle = 0$, then the cosine theorem implies the *Pythagorean theorem*[4]:

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2.$$

The following figures illustrate the cosine theorem and the Pythagorean theorem for vectors in $\mathbb{R}^{2,1}$:





In the following definition we generalize the ideas of angles and orthogonality.

Definition 12.7 Let $\mathcal{V}$ be a Euclidean or unitary vector space with the scalar product $\langle\cdot,\cdot\rangle$.

(1) In the Euclidean case, the *angle* between two vectors $v, w \in \mathcal{V} \setminus \{0\}$ is the uniquely determined real number $\varphi \in [0, \pi]$ with

$$\cos(\varphi) = \frac{\langle v, w\rangle}{\|v\|\,\|w\|}.$$

(2) Two vectors $v, w \in \mathcal{V}$ are called *orthogonal*, if $\langle v, w\rangle = 0$.
(3) A basis $\{v_1, \ldots, v_n\}$ of $\mathcal{V}$ is called an *orthogonal basis*, if

$$\langle v_i, v_j\rangle = 0, \quad i, j = 1, \ldots, n \ \text{ and } \ i \neq j.$$

If, furthermore,

$$\|v_i\| = 1, \quad i = 1, \ldots, n,$$

where $\|v\| = \langle v, v\rangle^{1/2}$ is the norm induced by the scalar product, then $\{v_1, \ldots, v_n\}$ is called an *orthonormal basis* of $\mathcal{V}$. (For an orthonormal basis we therefore have $\langle v_i, v_j\rangle = \delta_{ij}$.)

[4]Pythagoras of Samos (approx. 570–500 BC).

Note that the terms in (1)−(3) are defined with respect to the given scalar product. Different scalar products yield different angles between vectors. In particular, the orthogonality of two given vectors may be lost when we consider a different scalar product.

Example 12.8 The standard basis vectors $e_1, e_2 \in \mathbb{R}^{2,1}$ are orthogonal and $\{e_1, e_2\}$ is an orthonormal basis of $\mathbb{R}^{2,1}$ with respect to the standard scalar product (cp. (1) in Example 12.2). Consider the symmetric and invertible matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \in \mathbb{R}^{2,2},$$

which defines a symmetric and non-degenerate bilinear form on $\mathbb{R}^{2,1}$ by

$$(v, w) \mapsto w^T A v$$

(cp. (1) in Example 11.10). This bilinear form is positive definite, since for all $v = [\nu_1, \nu_2]^T \in \mathbb{R}^{2,1}$ we have

$$v^T A v = \nu_1^2 + \nu_2^2 + (\nu_1 + \nu_2)^2.$$

The bilinear form therefore is a scalar product on $\mathbb{R}^{2,1}$, which we denote by $\langle \cdot, \cdot \rangle_A$. We denote the induced norm by $\| \cdot \|_A$.

With respect to the scalar product $\langle \cdot, \cdot \rangle_A$ the vectors e_1, e_2 satisfy

$$\langle e_1, e_1 \rangle_A = e_1^T A e_1 = 2, \quad \langle e_2, e_2 \rangle_A = e_2^T A e_2 = 2, \quad \langle e_1, e_2 \rangle_A = e_2^T A e_1 = 1.$$

Clearly, $\{e_1, e_2\}$ is not an orthonormal basis of $\mathbb{R}^{2,1}$ with respect to $\langle \cdot, \cdot \rangle_A$. Also note that $\|e_1\|_A = \|e_2\|_A = \sqrt{2}$.

On the other hand, the vectors $v_1 = [1,\ 1]^T$ and $v_2 = [-1,\ 1]^T$ satisfy

$$\langle v_1, v_1 \rangle_A = v_1^T A v_1 = 6, \quad \langle v_2, v_2 \rangle_A = v_2^T A v_2 = 2, \quad \langle v_1, v_2 \rangle_A = v_2^T A v_1 = 0.$$

Hence $\|v_1\|_A = \sqrt{6}$ and $\|v_2\|_A = \sqrt{2}$, so that $\{6^{-1/2} v_1,\ 2^{-1/2} v_2\}$ is an orthonormal basis of $\mathbb{R}^{2,1}$ with respect to the scalar product $\langle \cdot, \cdot \rangle_A$

We now show that every finite dimensional Euclidean or unitary vector space has an orthonormal basis.

Theorem 12.9 *Let $\mathcal{V}$ be a Euclidean or unitary vector space with the basis $\{v_1, \ldots, v_n\}$. Then there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathcal{V}$ with*

$$\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{v_1, \ldots, v_k\}, \quad k = 1, \ldots, n.$$

Proof We give the proof by induction on $\dim(\mathcal{V}) = n$. If $n = 1$, then we set $u_1 := \|v_1\|^{-1}v_1$. Then $\|u_1\| = 1$, and $\{u_1\}$ is an orthonormal basis of $\mathcal{V}$ with $\text{span}\{u_1\} = \text{span}\{v_1\}$.

Let the assertion hold for an $n \geq 1$. Let $\dim(\mathcal{V}) = n + 1$ and let $\{v_1, \ldots, v_{n+1}\}$ be a basis of $\mathcal{V}$. Then $\mathcal{V}_n := \text{span}\{v_1, \ldots, v_n\}$ is an n-dimensional subspace of $\mathcal{V}$. By the induction hypothesis there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathcal{V}_n$ with $\text{span}\{u_1, \ldots, u_k\} = \text{span}\{v_1, \ldots, v_k\}$ for $k = 1, \ldots, n$. We define

$$\widehat{u}_{n+1} := v_{n+1} - \sum_{k=1}^{n} \langle v_{n+1}, u_k \rangle u_k, \quad u_{n+1} := ||\widehat{u}_{n+1}||^{-1}\widehat{u}_{n+1}.$$

Since $v_{n+1} \notin \mathcal{V}_n = \text{span}\{u_1, \ldots, u_n\}$, we must have $\widehat{u}_{n+1} \neq 0$, and Lemma 9.16 yields $\text{span}\{u_1, \ldots, u_{n+1}\} = \text{span}\{v_1, \ldots, v_{n+1}\}$.

For $j = 1, \ldots, n$ we have

$$\begin{aligned}
\langle u_{n+1}, u_j \rangle &= \langle \|\widehat{u}_{n+1}\|^{-1}\widehat{u}_{n+1}, u_j \rangle \\
&= \|\widehat{u}_{n+1}\|^{-1} \left(\langle v_{n+1}, u_j \rangle - \sum_{k=1}^{n} \langle v_{n+1}, u_k \rangle \, \langle u_k, u_j \rangle \right) \\
&= \|\widehat{u}_{n+1}\|^{-1} \big(\langle v_{n+1}, u_j \rangle - \langle v_{n+1}, u_j \rangle \big) \\
&= 0.
\end{aligned}$$

Finally, $\langle u_{n+1}, u_{n+1} \rangle = \|\widehat{u}_{n+1}\|^{-2} \langle \widehat{u}_{n+1}, \widehat{u}_{n+1} \rangle = 1$ which completes the proof. □

The proof of Theorem 12.9 shows how a given basis $\{v_1, \ldots, v_n\}$ can be *orthonormalized*, i.e., transformed into an orthonormal basis $\{u_1, \ldots, u_n\}$ with

$$\text{span}\{u_1, \ldots, u_k\} = \text{span}\{v_1, \ldots, v_k\}, \quad k = 1, \ldots, n.$$

The resulting algorithm is called the *Gram-Schmidt method*[5]:

Algorithm 12.10 Given a basis $\{v_1, \ldots, v_n\}$ of $\mathcal{V}$.

(1) Set $u_1 := \|v_1\|^{-1}v_1$.
(2) For $j = 1, \ldots, n - 1$ set

$$\begin{aligned}
\widehat{u}_{j+1} &:= v_{j+1} - \sum_{k=1}^{j} \langle v_{j+1}, u_k \rangle u_k, \\
u_{j+1} &:= \|\widehat{u}_{j+1}\|^{-1}\widehat{u}_{j+1}.
\end{aligned}$$

[5] Jørgen Pedersen Gram (1850–1916) and Erhard Schmidt (1876–1959).

A slight reordering and combination of steps in the Gram-Schmidt method yields

$$\underbrace{(v_1, v_2, \ldots, v_n)}_{\in \mathcal{V}^n} = \underbrace{(u_1, u_2, \ldots, u_n)}_{\in \mathcal{V}^n} \begin{pmatrix} \|v_1\| & \langle v_2, u_1\rangle & \ldots & \langle v_n, u_1\rangle \\ & \|\widehat{u}_2\| & \ddots & \vdots \\ & & \ddots & \langle v_n, u_{n-1}\rangle \\ & & & \|\widehat{u}_n\| \end{pmatrix}.$$

The upper triangular matrix on the right hand side is the coordinate transformation matrix from the basis $\{v_1, \ldots, v_n\}$ to the basis $\{u_1, \ldots, u_n\}$ of $\mathcal{V}$ (cp. Theorem 9.25 or 10.2). Thus, we have shown the following result.

Theorem 12.11 *If $\mathcal{V}$ is a finite dimensional Euclidean or unitary vector space with a given basis B_1, then the Gram-Schmidt method applied to B_1 yields an orthonormal basis B_2 of $\mathcal{V}$, such that $[\mathrm{Id}_{\mathcal{V}}]_{B_1,B_2}$ is an invertible upper triangular matrix.*

Consider an m-dimensional subspace of $\mathbb{R}^{n,1}$ or $\mathbb{C}^{n,1}$ with the standard scalar product $\langle \cdot, \cdot \rangle$, and write the m vectors of an orthonormal basis $\{q_1, \ldots, q_m\}$ as columns of a matrix, $Q := [q_1, \ldots, q_m]$. Then we obtain in the real case

$$Q^T Q = [q_i^T q_j] = [\langle q_j, q_i\rangle] = [\delta_{ji}] = I_m,$$

and analogously in the complex case

$$Q^H Q = [q_i^H q_j] = [\langle q_j, q_i\rangle] = [\delta_{ji}] = I_m.$$

If, on the other hand, $Q^T Q = I_m$ or $Q^H Q = I_m$ for a matrix $Q \in \mathbb{R}^{n,m}$ or $Q \in \mathbb{C}^{n,m}$, respectively, then the m columns of Q form an orthonormal basis (with respect to the standard scalar product) of an m-dimensional subspace of $\mathbb{R}^{n,1}$ or $\mathbb{C}^{n,1}$, respectively. A "matrix version" of Theorem 12.11 can therefore be formulated as follows.

Corollary 12.12 *Let $K = \mathbb{R}$ or $K = \mathbb{C}$ and let $v_1, \ldots, v_m \in K^{n,1}$ be linearly independent. Then there exists a matrix $Q \in K^{n,m}$ with its m columns being orthonormal with respect to the standard scalar product of $K^{n,1}$, i.e., $Q^T Q = I_m$ for $K = \mathbb{R}$ or $Q^H Q = I_m$ for $K = \mathbb{C}$, and an upper triangular matrix $R \in GL_m(K)$, such that*

$$[v_1, \ldots, v_m] = QR. \tag{12.3}$$

The factorization (12.3) is called a *QR-decomposition* of the matrix $[v_1, \ldots, v_m]$. The QR-decomposition has many applications in Numerical Mathematics (cp. Example 12.16 below).

Lemma 12.13 *Let $K = \mathbb{R}$ or $K = \mathbb{C}$ and let $Q \in K^{n,m}$ be a matrix with orthonormal columns with respect to the standard scalar product of $K^{n,1}$. Then $\|v\| = \|Qv\|$ holds for all $v \in K^{m,1}$. (Here $\|\cdot\|$ is the Euclidean norm of $K^{m,1}$ and of $K^{n,1}$.)*

Proof For $K = \mathbb{C}$ we have

$$\|v\|^2 = \langle v, v\rangle = v^H v = v^H (Q^H Q)v = \langle Qv, Qv\rangle = \|Qv\|^2,$$

and the proof for $K = \mathbb{R}$ is analogous. □

We now introduce two important classes of matrices.

Definition 12.14

(1) A matrix $Q \in \mathbb{R}^{n,n}$ whose columns form an orthonormal basis with respect to the standard scalar product of $\mathbb{R}^{n,1}$ is called *orthogonal*.
(2) A matrix $Q \in \mathbb{C}^{n,n}$ whose columns form an orthonormal basis with respect to the standard scalar product of $\mathbb{C}^{n,1}$ is called *unitary*.

A matrix $Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n,n}$ is therefore orthogonal if and only if

$$Q^T Q = [q_i^T q_j] = [\langle q_j, q_i\rangle] = [\delta_{ji}] = I_n.$$

In particular, an orthogonal matrix Q is invertible with $Q^{-1} = Q^T$ (cp. Corollary 7.20). The equation $QQ^T = I_n$ means that the n rows of Q form an orthonormal basis of $\mathbb{R}^{1,n}$ (with respect to the scalar product $\langle v, w\rangle := wv^T$).

Analogously, a unitary matrix $Q \in \mathbb{C}^{n,n}$ is invertible with $Q^{-1} = Q^H$ and $Q^H Q = I_n = QQ^H$. The n columns of Q form an orthonormal basis of $\mathbb{C}^{1,n}$.

Lemma 12.15 *The sets $\mathcal{O}(n)$ of orthogonal and $\mathcal{U}(n)$ of unitary $n \times n$ matrices form subgroups of $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$, respectively.*

Proof We consider only $\mathcal{O}(n)$; the proof for $\mathcal{U}(n)$ is analogous.

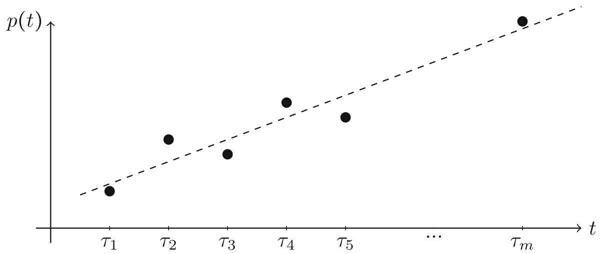
Since every orthogonal matrix is invertible, we have that $\mathcal{O}(n) \subset GL_n(\mathbb{R})$. The identity matrix I_n is orthogonal, and hence $I_n \in \mathcal{O}(n) \neq \emptyset$. If $Q \in \mathcal{O}(n)$, then also $Q^T = Q^{-1} \in \mathcal{O}(n)$, since $(Q^T)^T Q^T = QQ^T = I_n$. Finally, if $Q_1, Q_2 \in \mathcal{O}(n)$, then

$$(Q_1 Q_2)^T (Q_1 Q_2) = Q_2^T (Q_1^T Q_1) Q_2 = Q_2^T Q_2 = I_n,$$

and thus $Q_1 Q_2 \in \mathcal{O}(n)$. □

Example 12.16 In many applications measurements or samples lead to a data set that is represented by tuples $(\tau_i, \mu_i) \in \mathbb{R}^2$, $i = 1, \ldots, m$. Here $\tau_1 < \cdots < \tau_m$, are the pairwise distinct measurement points and $\mu_1, \ldots, \mu_m$ are the corresponding measurements. In order to approximate the given data set by a simple model, one can try to construct a polynomial p of small degree so that the values $p(\tau_1), \ldots, p(\tau_m)$ are as close as possible to the measurements $\mu_1, \ldots, \mu_m$.

The simplest case is a real polynomial of degree (at most) 1. Geometrically, this corresponds to the construction of a straight line in $\mathbb{R}^2$ that has a minimal distance to the given points, as shown in the figure below (cp. Sect. 1.4). There are many possibilities to measure the distance. In the following we will describe one of them in more detail and use the Gram-Schmidt method, or the QR-decomposition, for the construction of the straight line. In Statistics this method is called *linear regression*.



A real polynomial of degree (at most) 1 has the form $p = \alpha t + \beta$ and we are looking for coefficients $\alpha, \beta \in \mathbb{R}$ with

$$p(\tau_i) = \alpha\tau_i + \beta \approx \mu_i, \quad i = 1, \dots, m.$$

Using matrices we can write this problem as

$$\begin{bmatrix} \tau_1 & 1 \\ \vdots & \vdots \\ \tau_m & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \approx \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix} \quad \text{or} \quad [v_1, v_2] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \approx y.$$

As mentioned above, there are different possibilities for interpreting the symbol "$\approx$". In particular, there are different norms in which we can measure the distance between the given values $\mu_1, \dots, \mu_m$ and the polynomial values $p(\tau_1), \dots, p(\tau_m)$. Here we will use the Euclidean norm $\|\cdot\|$ and consider the minimization problem

$$\min_{\alpha,\beta\in\mathbb{R}} \left\| [v_1, v_2] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} - y \right\|.$$

The vectors $v_1, v_2 \in \mathbb{R}^{m,1}$ are linearly independent, since the entries of v_1 are pairwise distinct, while all entries of v_2 are equal. Let

$$[v_1, v_2] = [q_1, q_2]R$$

be a QR-decomposition. We extend the vectors $q_1, q_2 \in \mathbb{R}^{m,1}$ to an orthonormal basis $\{q_1, q_2, q_3, \dots, q_m\}$ of $\mathbb{R}^{m,1}$. Then $Q = [q_1, \dots, q_m] \in \mathbb{R}^{m,m}$ is an orthogonal matrix and

$$\begin{aligned}
\min_{\alpha,\beta\in\mathbb{R}} \left\| [v_1, v_2] \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - y \right\| &= \min_{\alpha,\beta\in\mathbb{R}} \left\| [q_1, q_2] R \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - y \right\| \\
&= \min_{\alpha,\beta\in\mathbb{R}} \left\| Q \begin{bmatrix} R \\ 0_{m-2,2}\end{bmatrix} \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - y \right\| \\
&= \min_{\alpha,\beta\in\mathbb{R}} \left\| Q \left(\begin{bmatrix} R \\ 0_{m-2,2}\end{bmatrix} \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - Q^T y \right) \right\| \\
&= \min_{\alpha,\beta\in\mathbb{R}} \left\| \begin{bmatrix} R \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} q_1^T y \\ q_2^T y \\ q_3^T y \\ \vdots \\ q_m^T y \end{bmatrix} \right\| .
\end{aligned}$$

Here we have used that $QQ^T = I_m$ and $\|Qv\| = \|v\|$ for all $v \in \mathbb{R}^{m,1}$. The upper triangular matrix R is invertible and thus the minimization problem is solved by

$$\begin{bmatrix}\widetilde{\alpha}\\ \widetilde{\beta}\end{bmatrix} = R^{-1} \begin{bmatrix} q_1^T y \\ q_2^T y \end{bmatrix}.$$

Using the definition of the Euclidean norm, we can write the minimizing property of the polynomial $\widetilde{p} := \widetilde{\alpha} t + \widetilde{\beta}$ as

$$\begin{aligned}
\left\| [v_1, v_2] \begin{bmatrix}\widetilde{\alpha}\\ \widetilde{\beta}\end{bmatrix} - y \right\|^2 &= \sum_{i=1}^{m} (\widetilde{p}(\tau_i) - \mu_i)^2 \\
&= \min_{\alpha,\beta\in\mathbb{R}} \left(\sum_{i=1}^{m} ((\alpha\tau_i + \beta) - \mu_i)^2 \right).
\end{aligned}$$

Since the polynomial $\widetilde{p}$ minimizes the sum of squares of the distances between the measurements μ_i and the polynomial values $\widetilde{p}(\tau_i)$, this polynomial yields a *least squares approximation* of the measurement values.

Consider the example from Sect. 1.4. In the four quarters of a year, a company has profits of 10, 8, 9, 11 million Euros. Under the assumption that the profits grows linearly, i.e., like a straight line, the goal is to estimate the profit in the last quarter of the following year. The given data leads to the approximation problem

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \approx \begin{bmatrix} 10 \\ 8 \\ 9 \\ 11 \end{bmatrix} \quad \text{or} \quad [v_1, v_2] \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \approx y.$$

The numerical computation of a QR-decomposition of $[v_1, v_2]$ yields

$$\begin{bmatrix} \widetilde{\alpha} \\ \widetilde{\beta} \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{30} & \frac{1}{3}\sqrt{30} \\ 0 & \frac{1}{3}\sqrt{6} \end{bmatrix}^{-1}}_{=R^{-1}} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & \frac{3}{\sqrt{30}} & \frac{4}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} \end{bmatrix}}_{=[q_1,q_2]^T} \begin{bmatrix} 10 \\ 8 \\ 9 \\ 11 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 8.5 \end{bmatrix},$$

and the resulting profit estimate for the last quarter of the following year is $\widetilde{p}(8) = 11.7$, i.e., 11.7 million Euros.

MATLAB-Minute.
In Example 12.16 one could imagine that the profit grows quadratically instead of linearly. Determine, analogously to the procedure in Example 12.16, a polynomial $\widetilde{p} = \widetilde{\alpha}t^2 + \widetilde{\beta}t + \widetilde{\gamma}$ that solves the least squares problem

$$\sum_{i=1}^{4} (\widetilde{p}(\tau_i) - \mu_i)^2 = \min_{\alpha,\beta,\gamma\in\mathbb{R}} \left(\sum_{i=1}^{4} \left((\alpha\tau_i^2 + \beta\tau_i + \gamma) - \mu_i \right)^2 \right).$$

Use the MATLAB command `qr` for computing a QR-decomposition, and determine the estimated profit in the last quarter of the following year.

We will now analyze the properties of orthonormal bases in more detail.

Lemma 12.17 *If $\mathcal{V}$ is a Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$ and the orthonormal basis $\{u_1, \ldots, u_n\}$, then*

$$v = \sum_{i=1}^{n} \langle v, u_i \rangle u_i$$

for all $v \in \mathcal{V}$.

Proof For every $v \in \mathcal{V}$ there exist uniquely determined coordinates $\lambda_1, \ldots, \lambda_n$ with $v = \sum_{i=1}^n \lambda_i u_i$. For every $j = 1, \ldots, n$ we then have $\langle v, u_j \rangle = \sum_{i=1}^n \lambda_i \langle u_i, u_j \rangle = \lambda_j$. □

The coordinates $\langle v, u_i \rangle$, $i = 1, \ldots, n$, of v with respect to an orthonormal basis $\{u_1, \ldots, u_n\}$ are often called the *Fourier coefficients*[6] of v with respect to this basis. The representation $v = \sum_{i=1}^n \langle v, u_i \rangle u_i$ is called the (abstract) *Fourier expansion* of v in the given orthonormal basis.

[6]Jean Baptiste Joseph Fourier (1768–1830).

Corollary 12.18 *If $\mathcal{V}$ is a Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$ and the orthonormal basis $\{u_1, \ldots, u_n\}$, then the following assertions hold:*

(1) $\langle v, w\rangle = \sum_{i=1}^{n} \langle v, u_i\rangle \langle u_i, w\rangle = \sum_{i=1}^{n} \langle v, u_i\rangle \overline{\langle w, u_i\rangle}$ *for all* $v, w \in \mathcal{V}$ *(Parseval's identity*[7]*).*

(2) $\langle v, v\rangle = \sum_{i=1}^{n} |\langle v, u_i\rangle|^2$ *for all* $v \in \mathcal{V}$ *(Bessel's identity*[8]*).*

Proof

(1) We have $v = \sum_{i=1}^{n} \langle v, u_i\rangle u_i$, and thus

$$\langle v, w\rangle = \left\langle \sum_{i=1}^{n} \langle v, u_i\rangle u_i, w \right\rangle = \sum_{i=1}^{n} \langle v, u_i\rangle \langle u_i, w\rangle = \sum_{i=1}^{n} \langle v, u_i\rangle \overline{\langle w, u_i\rangle}.$$

(2) is a special case of (1) for $v = w$. □

By Bessel's identity, every vector $v \in \mathcal{V}$ satisfies

$$\|v\|^2 = \langle v, v\rangle = \sum_{i=1}^{n} |\langle v, u_i\rangle|^2 \geq \max_{1 \leq i \leq n} |\langle v, u_i\rangle|^2,$$

where $\| \cdot \|$ is the norm induced by the scalar product. The absolute value of each coordinate of v with respect to an orthonormal basis of $\mathcal{V}$ is therefore bounded by the norm of v. This property does not hold for a general basis of $\mathcal{V}$.

Example 12.19 Consider $\mathcal{V} = \mathbb{R}^{2,1}$ with the standard scalar product and the Euclidean norm, then for every real $\varepsilon \neq 0$ the set

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \right\}$$

is a basis of $\mathcal{V}$. For every vector $v = [\nu_1, \nu_2]^T$ we then have

$$v = \left(\nu_1 - \frac{\nu_2}{\varepsilon}\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{\nu_2}{\varepsilon} \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}.$$

If $|\nu_1|, |\nu_2|$ are moderate numbers and if $|\varepsilon|$ is (very) small, then $|\nu_1 - \nu_2/\varepsilon|$ and $|\nu_2/\varepsilon|$ are (very) large. In numerical algorithms such a situation can lead to significant problems (e.g. due to roundoff errors) that are avoided when orthonormal bases are used.

[7] Marc-Antoine Parseval (1755–1836).

[8] Friedrich Wilhelm Bessel (1784–1846).

Definition 12.20 Let $\mathcal{V}$ be a Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$, and let $\mathcal{U} \subseteq \mathcal{V}$ be a subspace. Then

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

is called the *orthogonal complement of* $\mathcal{U}$ (in $\mathcal{V}$).

Lemma 12.21 *The orthogonal complement* $\mathcal{U}^{\perp}$ *is a subspace of* $\mathcal{V}$.

Proof Exercise. □

Lemma 12.22 *If* $\mathcal{V}$ *is an* n*-dimensional Euclidean or unitary vector space, and if* $\mathcal{U} \subseteq \mathcal{V}$ *is an* m*-dimensional subspace, then* $\dim(\mathcal{U}^{\perp}) = n - m$ *and* $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^{\perp}$.

Proof We know that $m \leq n$ (cp. Lemma 9.27). If $m = n$, then $\mathcal{U} = \mathcal{V}$, and thus

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

so that the assertion is trivial.

Thus let $m < n$ and let $\{u_1, \ldots, u_m\}$ be an orthonormal basis of $\mathcal{U}$. We extend this basis to a basis of $\mathcal{V}$ and apply the Gram-Schmidt method in order to obtain an orthonormal basis $\{u_1, \ldots, u_m, u_{m+1}, \ldots, u_n\}$ of $\mathcal{V}$. Then $\text{span}\{u_{m+1}, \ldots, u_n\} \subseteq \mathcal{U}^{\perp}$ and therefore $\mathcal{V} = \mathcal{U} + \mathcal{U}^{\perp}$. If $w \in \mathcal{U} \cap \mathcal{U}^{\perp}$, then $\langle w, w \rangle = 0$, and hence $w = 0$, since the scalar product is positive definite. Thus, $\mathcal{U} \cap \mathcal{U}^{\perp} = \{0\}$, which implies that $\mathcal{V} = \mathcal{U} \oplus \mathcal{U}^{\perp}$ and $\dim(\mathcal{U}^{\perp}) = n - m$ (cp. Theorem 9.29). In particular, we have $\mathcal{U}^{\perp} = \text{span}\{u_{m+1}, \ldots, u_n\}$. □

12.3 The Vector Product in $\mathbb{R}^{3,1}$

In this section we consider a further product on the vector space $\mathbb{R}^{3,1}$ that is frequently used in Physics and Electrical Engineering.

Definition 12.23 The *vector product* or *cross product* in $\mathbb{R}^{3,1}$ is the map

$$\mathbb{R}^{3,1} \times \mathbb{R}^{3,1} \to \mathbb{R}^{3,1}, \quad (v, w) \mapsto v \times w := [\nu_2\omega_3 - \nu_3\omega_2,\ \nu_3\omega_1 - \nu_1\omega_3,\ \nu_1\omega_2 - \nu_2\omega_1]^T,$$

where $v = [\nu_1, \nu_2, \nu_3]^T$ and $w = [\omega_1, \omega_2, \omega_3]^T$.

In contrast to the scalar product, the vector product of two elements of the vector space $\mathbb{R}^{3,1}$ is not a scalar but again a vector in $\mathbb{R}^{3,1}$. Using the canonical basis vectors of $\mathbb{R}^{3,1}$,

$$e_1 = [1, 0, 0]^T, \quad e_2 = [0, 1, 0]^T, \quad e_3 = [0, 0, 1]^T,$$

we can write the vector product as

$$v \times w = \det\left(\begin{bmatrix} \nu_2 & \omega_2 \\ \nu_3 & \omega_3 \end{bmatrix}\right) e_1 - \det\left(\begin{bmatrix} \nu_1 & \omega_1 \\ \nu_3 & \omega_3 \end{bmatrix}\right) e_2 + \det\left(\begin{bmatrix} \nu_1 & \omega_1 \\ \nu_2 & \omega_2 \end{bmatrix}\right) e_3.$$

Lemma 12.24 *The vector product is linear in both components and for all $v, w \in \mathbb{R}^{3,1}$ the following properties hold:*

(1) $v \times w = -w \times v$, *i.e., the vector product is* anti commutative *or* alternating.
(2) $\|v \times w\|^2 = \|v\|^2\,\|w\|^2 - \langle v, w\rangle^2$, *where* $\langle\cdot,\cdot\rangle$ *is the standard scalar product and* $\|\cdot\|$ *the Euclidean norm of* $\mathbb{R}^{3,1}$.
(3) $\langle v, v \times w\rangle = \langle w, v \times w\rangle = 0$, *where* $\langle\cdot,\cdot\rangle$ *is the standard scalar product of* $\mathbb{R}^{3,1}$.

Proof Exercise. □

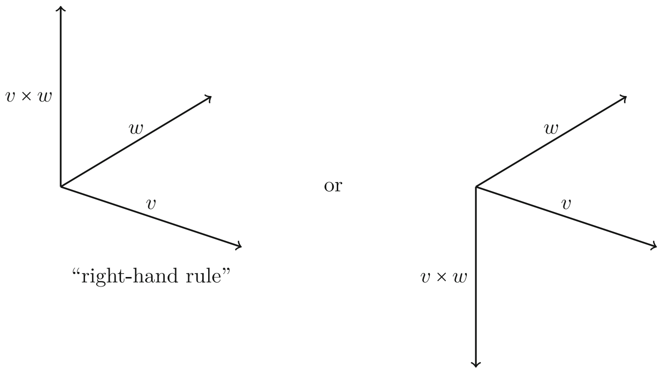
By (2) and the Cauchy-Schwarz inequality (12.2), it follows that $v \times w = 0$ holds if and only if v, w are linearly dependent. From (3) we obtain

$$\langle \lambda v + \mu w, v \times w\rangle = \lambda\langle v, v \times w\rangle + \mu\langle w, v \times w\rangle = 0,$$

for arbitrary $\lambda, \mu \in \mathbb{R}$. If v, w are linearly independent, then the product $v \times w$ is orthogonal to the plane through the origin spanned by v and w in $\mathbb{R}^{3,1}$, i.e.,

$$v \times w \ \in \ \{\lambda v + \mu w \mid \lambda, \mu \in \mathbb{R}\}^{\perp}.$$

Geometrically, there are two possibilities:



The positions of the three vectors $v, w, v \times w$ on the left side of this figure correspond to the "right-handed orientation" of the usual coordinate system of $\mathbb{R}^{3,1}$, where the canonical basis vectors e_1, e_2, e_3 are associated with thumb, index finger and middle finger of the right hand. This motivates the name *right-hand rule*. In order to explain this in detail, one needs to introduce the concept of *orientation*, which we omit here.

If $\varphi \in [0, \pi]$ is the angle between the vectors v, w, then

$$\langle v, w\rangle = \|v\|\,\|w\|\,\cos(\varphi)$$

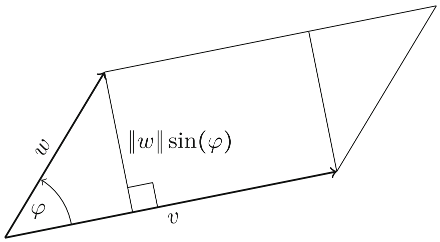
(cp. Definition 12.7) and we can write (2) in Lemma 12.24 as

$$\|v \times w\|^2 = \|v\|^2 \, \|w\|^2 - \|v\|^2 \, \|w\|^2 \, \cos^2(\varphi) = \|v\|^2 \, \|w\|^2 \, \sin^2(\varphi),$$

so that

$$\|v \times w\| = \|v\| \, \|w\| \, \sin(\varphi).$$

A geometric interpretation of this equation is the following: *The norm of the vector product of v and w is equal to the area of the parallelogram spanned by v and w.* This interpretation is illustrated in the following figure:



Exercises

12.1 Let $\mathcal{V}$ be a finite dimensional real or complex vector space. Show that there exists a scalar product on $\mathcal{V}$.

12.2 Show that the maps defined in Example 12.2 are scalar products on the corresponding vector spaces.

12.3 Let $\langle \cdot, \cdot \rangle$ be an arbitrary scalar product on $\mathbb{R}^{n,1}$. Show that there exists a matrix $A \in \mathbb{R}^{n,n}$ with $\langle v, w \rangle = w^T A v$ for all $v, w \in \mathbb{R}^{n,1}$.

12.4 Let $\mathcal{V}$ be a finite dimensional $\mathbb{R}$- or $\mathbb{C}$-vector space. Let s_1 and s_2 be scalar products on $\mathcal{V}$ with the following property: If $v, w \in \mathcal{V}$ satisfy $s_1(v, w) = 0$, then also $s_2(v, w) = 0$. Prove or disprove: There exists a real scalar $\lambda > 0$ with $s_1(v, w) = \lambda s_2(v, w)$ for all $v, w \in \mathcal{V}$.

12.5 Show that the maps defined in Example 12.4 are norms on the corresponding vector spaces.

12.6 Show that

$$\|A\|_1 = \max_{1 \le j \le m} \sum_{i=1}^{n} |a_{ij}| \quad \text{and} \quad \|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{m} |a_{ij}|$$

for all $A = [a_{ij}] \in K^{n,m}$, where $K = \mathbb{R}$ or $K = \mathbb{C}$ (cp. (6) in Example 12.4).

12.7 Sketch for the matrix A from (6) in Example 12.4 and $p \in \{1, 2, \infty\}$, the sets $\{Av \mid v \in \mathbb{R}^{2,1}, \ \|v\|_p = 1\} \subset \mathbb{R}^{2,1}$.

12.8 Let $\mathcal{V}$ be a Euclidean or unitary vector space and let $\| \cdot \|$ be the norm induced by a scalar product on $\mathcal{V}$. Show that $\| \cdot \|$ satisfies the *parallelogram identity*

$$\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$$

for all $v, w \in \mathcal{V}$.

12.9 Let $\mathcal{V}$ be a K-vector space ($K = \mathbb{R}$ or $K = \mathbb{C}$) with the scalar product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. Show that $v, w \in \mathcal{V}$ are orthogonal with respect to $\langle \cdot, \cdot \rangle$ if and only if $\|v + \lambda w\| = \|v - \lambda w\|$ for all $\lambda \in K$.

12.10 Does there exist a scalar product $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^{n,1}$, such that the 1-norm of $\mathbb{C}^{n,1}$ (cp. (5) in Example 12.4) is the induced norm by this scalar product?

12.11 Show that the inequality

$$\Big(\sum_{i=1}^{n} \alpha_i \beta_i\Big)^2 \leq \sum_{i=1}^{n} (\gamma_i \alpha_i)^2 \;\cdot\; \sum_{i=1}^{n} \Big(\frac{\beta_i}{\gamma_i}\Big)^2$$

holds for arbitrary real numbers $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ and positive real numbers $\gamma_1, \ldots, \gamma_n$.

12.12 Let $\mathcal{V}$ be a finite dimensional Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$. Let $f : \mathcal{V} \to \mathcal{V}$ be a map with $\langle f(v), f(w) \rangle = \langle v, w \rangle$ for all $v, w \in \mathcal{V}$. Show that f is an isomorphism.

12.13 Let $\mathcal{V}$ be a unitary vector space and suppose that $f \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ satisfies $\langle f(v), v \rangle = 0$ for all $v \in \mathcal{V}$. Prove or disprove that $f = 0$.
Does the same statement also hold for Euclidean vector spaces?

12.14 Let $D = \mathrm{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n,n}$ with $d_1, \ldots, d_n > 0$. Show that $\langle v, w \rangle = w^T D v$ is a scalar product on $\mathbb{R}^{n,1}$. Analyze which properties of a scalar product are violated if at least one of the d_i is zero, or when all d_i are nonzero but have different signs.

12.15 Orthonormalize the following basis of the vector space $\mathbb{C}^{2,2}$ with respect to the scalar product $\langle A, B \rangle = \mathrm{trace}(B^H A)$:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}.$$

12.16 Let $Q \in \mathbb{R}^{n,n}$ be an orthogonal or let $Q \in \mathbb{C}^{n,n}$ be a unitary matrix. What are the possible values of $\det(Q)$?

12.17 Let $u \in \mathbb{R}^{n,1} \setminus \{0\}$ and let

$$H(u) = I_n - 2\,\frac{1}{u^T u} u u^T \;\in\; \mathbb{R}^{n,n}.$$

Show that the n columns of $H(u)$ form an orthonormal basis of $\mathbb{R}^{n,1}$ with respect to the standard scalar product. (Matrices of this form are called *Householder matrices*.[9] We will study them in more detail in Example 18.15.)

12.18 Prove Lemma 12.21.

[9] Alston Scott Householder (1904–1993), pioneer of Numerical Linear Algebra.

12.19 Let

$$[v_1, v_2, v_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{3,3}.$$

Analyze whether the vectors v_1, v_2, v_3 are orthonormal with respect to the standard scalar product and compute the orthogonal complement of $\text{span}\{v_1, v_2, v_3\}$.

12.20 Let $\mathcal{V}$ be a Euclidean or unitary vector space with the scalar product $\langle \cdot, \cdot \rangle$, let $u_1, \dots, u_k \in \mathcal{V}$ and let $\mathcal{U} = \text{span}\{u_1, \dots, u_k\}$. Show that for $v \in \mathcal{V}$ we have $v \in \mathcal{U}^\perp$ if and only if $\langle v, u_j \rangle = 0$ for $j = 1, \dots, k$.

12.21 In the unitary vector space $\mathbb{C}^{4,1}$ with the standard scalar product let $v_1 = [-1, \mathbf{i}, 0, 1]^T$ and $v_2 = [\mathbf{i}, 0, 2, 0]^T$ be given. Determine an orthonormal basis of $\text{span}\{v_1, v_2\}^\perp$.

12.22 Prove Lemma 12.24.