Chapter 5

Scalar product and orthogonality

5.1 Scalar product

Recall that the notion of a vector space has been introduced as an abstract version of the properties shared both by $\mathbb{R}^n$ and by $M_{mn}(\mathbb{R})$. Similarly, we have introduced the scalar product on $\mathbb{R}^n$ already in Chapter 1, let us now consider an abstract version of it. For simplicity, we introduce it on real vector spaces, but a slightly more general version will be considered once the complex numbers will be at our disposal.

Definition 5.1.1. A scalar product *on a real vector space V is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ such that for any $X, Y, Z \in V$ and $\lambda \in \mathbb{R}$ one has*

(i) $\langle X, Y\rangle = \langle Y, X\rangle$,

(ii) $\langle X + Y, Z\rangle = \langle X, Z\rangle + \langle Y, Z\rangle$,

(iii) $\langle \lambda X, Y\rangle = \lambda\langle X, Y\rangle$,

(iv) $\langle X, X\rangle \geq 0$ *and* $\langle X, X\rangle = 0$ *if and only if* $X = \mathbf{0}$.

Example 5.1.2. *For $V = \mathbb{R}^n$ and $X, Y \in V$ one sets $\langle X, Y\rangle := X \cdot Y$ and one can check that the four conditions above are satisfied.*

Example 5.1.3. *For $a, b \in \mathbb{R}$ with $a < b$ one considers $V = C\big([a, b]; \mathbb{R}\big)$ and for any $f, g \in V$ one defines*

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

It is easily checked that this defines a scalar product on V, see Exercise 5.5. For information, this scalar product extends to the set of L^2-functions (the set of square integrable functions).

Definition 5.1.4. *If V is a real vector space endowed with a scalar product, one says that $X, Y \in V$ are* orthogonal *if $\langle X, Y\rangle = 0$, and one writes $X \perp Y$. If S is a subset of V, one writes*

$$S^\perp := \{Y \in V \mid \langle X, Y\rangle = 0 \text{ for all } X \in S\}$$

and call it the orthogonal subspace of S.

One easily shows that $S^\perp$ is always a subspace of V.

Definition 5.1.5. *For any real vector space V endowed with a scalar product and for any $X \in V$ we set*

$$\|X\| := \sqrt{\langle X, X\rangle}$$

and call it the norm of X *(associated with the scalar product $\langle\cdot,\cdot\rangle$).*

Lemma 5.1.6. *For any real vector space V endowed with a scalar product, for any $X, Y \in V$ and for $\lambda \in \mathbb{R}$ one has*

(i) $\|\lambda X\| = |\lambda|\,\|X\|$,

(ii) $\|X+Y\|^2 = \|X\|^2 + \|Y\|^2$ *if and only if* $X \perp Y$ *(Pythagoras theorem)*

(iii) $\|X+Y\|^2 + \|X-Y\|^2 = 2\|X\|^2 + 2\|Y\|^2$,

(iv) $\|X+Y\| \leq \|X\| + \|Y\|$.

The proof will be provided in Exercise 5.1. The following statement is a generalization of a property already derived in the context of $\mathbb{R}^n$.

Lemma 5.1.7. *For any real vector space V endowed with a scalar product and for any $X, Y \in V$ one has*

$$|\langle X, Y\rangle| \leq \|X\|\,\|Y\|. \tag{5.1.1}$$

Proof. Let us first consider the trivial case $Y = \mathbf{0}$ for which (5.1.1) is an equality with 0 on both sides.

Now, assume that $Y \neq \mathbf{0}$ and set $c := \frac{\langle X,Y\rangle}{\|Y\|^2}$. Then let us observe that $(X - cY)\perp Y$, since

$$\langle X - cY, Y\rangle = \langle X, Y\rangle - \frac{\langle X, Y\rangle}{\|Y\|^2}\langle Y, Y\rangle = 0.$$

It follows by Pythagoras theorem that

$$\|X\|^2 = \|(X - cY) + cY\|^2 = \|X - cY\|^2 + \|cY\|^2 = \|X - cY\|^2 + c^2\|Y\|^2,$$

which implies that $\|X\|^2 \geq c^2\|Y\|$, or equivalently $\|X\| \geq |c|\ \|Y\|$. Note that this inequality can also be rewritten as $|c| \leq \frac{\|X\|}{\|Y\|}$.

By collecting these information one gets

$$|\langle X, Y\rangle| = |c|\,\|Y\|^2 \leq \frac{\|X\|}{\|Y\|}\|Y\|^2 = \|X\|\,\|Y\|,$$

which corresponds to the claim. □

5.2 Orthogonal bases

Definition 5.2.1. *Let V be a real vector space endowed with a scalar product, and let $\{V_1, \dots, V_n\}$ be a basis for V. The basis is called* orthogonal *if $\langle V_i, V_j\rangle = 0$ whenever $i, j \in \{1, \dots, n\}$ and $i \neq j$. If in addition $\langle V_i, V_i\rangle = 1$ for any $i \in \{1, \dots, n\}$ the basis is called* orthonormal.

Example 5.2.2. *The standard basis $\{E_1, \dots, E_n\}$ of $\mathbb{R}^n$ is an orthonormal basis.*

The following result is of conceptual importance, and rather well-known.

Theorem 5.2.3 (Graham-Schmidt)**.** *Let V be a real vector space of dimension n endowed with a scalar product. Then there exists an orthonormal basis for V.*

The proof consists in the explicit construction of an orthonormal basis.

Proof. Let $\{V_1, \dots, V_n\}$ be an arbitrary basis for V (such a basis exists since otherwise the dimension of V would not be defined), and let us set

$$\begin{aligned}
V_1' &:= \frac{1}{\|V_1\|} V_1 \\
V_2' &:= \frac{1}{\|V_2 - \langle V_2, V_1'\rangle V_1'\|}\big(V_2 - \langle V_2, V_1'\rangle V_1'\big) \\
&\vdots \\
V_n' &:= \frac{1}{\big\|V_n - \sum_{i=1}^{n-1}\langle V_n, V_i'\rangle V_i'\big\|}\Big(V_n - \sum_{i=1}^{n-1}\langle V_n, V_i'\rangle V_i'\Big),
\end{aligned}$$

where the prefactors are chosen such that $\|V_j'\| = 1$ (note that $V_j - \sum_{i=1}^{j-1}\langle V_j, V_i'\rangle V_i'$ is always different from $\mathbf{0}$ since otherwise V_j would be a linear combination of $V_1, \dots, V_{j-1}$ which is not possible by assumption). Then, it simply remains to observe that $V_j' \perp V_k'$ for any $j \neq k$. As a consequence, the elements V_j' generate an orthonormal basis for V, as expected. □

5.3 Bilinear maps

The notion of bilinear maps will be useful for calculus II.

Definition 5.3.1. *Let V, W, U be vector spaces over the same field $\mathbb{F}$. A map $\mathrm{T} : V \times W \to U$ is* bilinear *if it is linear in each argument, namely for any $X, X_1, X_2 \in V$, any $Y, Y_1, Y_2 \in W$ and $\lambda \in \mathbb{F}$ one has*

(i) $\mathrm{T}(X_1 + X_2, Y) = \mathrm{T}(X_1, Y) + \mathrm{T}(X_2, Y)$,

(ii) $\mathrm{T}(\lambda X, Y) = \lambda \mathrm{T}(X, Y)$,

(iii) $\mathrm{T}(X, Y_1 + Y_2) = \mathrm{T}(X, Y_1) + \mathrm{T}(X, Y_2)$,

(iv) $\mathrm{T}(X, \lambda Y) = \lambda \mathrm{T}(X, Y)$.

Example 5.3.2. *The scalar product* $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is a bilinear map on the Euclidean space* $\mathbb{R}^n$.

Example 5.3.3. *If* $\mathcal{A} \in M_{mn}(\mathbb{F})$ *one can define a bilinear map* $\mathrm{F}_{\mathcal{A}} : \mathbb{F}^m \times \mathbb{F}^n \to \mathbb{F}$ *for any* $X \in \mathbb{F}^m$ *and* $Y \in \mathbb{F}^n$ *by*

$$\mathrm{F}_{\mathcal{A}}(X, Y) = {}^tX\mathcal{A}Y \equiv \underbrace{{}^tX}_{\in M_{1m}(\mathbb{F})} \underbrace{\mathcal{A}}_{\in M_{mn}(\mathbb{F})} \underbrace{Y}_{\in M_{n1}(\mathbb{F})} \in \mathbb{F}. \tag{5.3.1}$$

Note that it is easily checked that $\mathrm{F}_{\mathcal{A}}$ *is indeed a bilinear map. For example, if* $\mathcal{A} = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, $X = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *and* $Y = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, *then*

$$\mathrm{F}_{\mathcal{A}}(X, Y) = (1\ 0) \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (1\ 0) \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2.$$

More generally, observe that if $\mathcal{A} = (a_{ij})$, $X = {}^t(x_1, \ldots, x_m)$ *and* $Y = {}^t(y_1, \ldots, y_n)$ *then*

$${}^tX\mathcal{A}Y = X \cdot (\mathcal{A}Y) = \sum_{i=1}^{m} x_i (\mathcal{A}Y)_i = \sum_{i=1}^{m} x_i \sum_{j=1}^{n} a_{ij}\, y_j = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, x_i\, y_j.$$

We shall now see that many bilinear maps are of the form presented in the previous example. For that purpose, recall from Section 4.5 that if $\mathcal{V} = \{V_1, \ldots, V_m\}$ is a basis for a vector space V over $\mathbb{F}$ and if $\mathcal{X} \in V$ then the coordinate vector of $\mathcal{X}$ is the element $X = {}^t(x_1, \ldots, x_m) \in \mathbb{F}^m$ such that $\mathcal{X} = x_1V_1 + \cdots + x_mV_m$. One has already introduced the notation $(\mathcal{X})_{\mathcal{V}} = X$. Similarly, for a basis $\mathcal{W} = \{W_1, \ldots, W_n\}$ of a vector space W over $\mathbb{F}$ and for any $\mathcal{Y} \in W$ one sets $(\mathcal{Y})_{\mathcal{W}} = Y = {}^t(y_1, \ldots, y_n) \in \mathbb{F}^n$ for its coordinate vector.

Lemma 5.3.4. *Let* V, W *be vector spaces over a field* $\mathbb{F}$ *and let* $\mathrm{F} : V \times \mathcal{W} \to \mathbb{F}$ *be a bilinear map. If* $\mathcal{V} = \{V_1, \ldots, V_m\}$ *is a basis for* V, *and if* $\mathcal{W} = \{W_1, \ldots, W_n\}$ *is a basis for* W *then there exists* $\mathcal{A} \in M_{mn}(\mathbb{F})$ *such that*

$$\mathrm{F}(\mathcal{X}, \mathcal{Y}) = {}^tX\mathcal{A}Y$$

for any $\mathcal{X} \in V$, *any* $\mathcal{Y} \in W$ *and with* $X = (\mathcal{X})_{\mathcal{V}}$ *and* $Y = (\mathcal{Y})_{\mathcal{W}}$.

Proof. By taking the bilinearity of F into account, one has

$$\mathrm{F}(\mathcal{X}, \mathcal{Y}) = \mathrm{F}\Big(\sum_{i=1}^{m} x_i\, V_i, \sum_{j=1}^{n} y_j\, W_j\Big) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i\, y_j \mathrm{F}(V_i, W_j).$$

Thus, by setting $a_{ij} = \mathrm{F}(V_i, W_j) \in \mathbb{F}$ one deduces that

$$\mathrm{F}(\mathcal{X}, \mathcal{Y}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, x_i\, y_j = {}^tX\mathcal{A}Y$$

with $\mathcal{A} = (a_{ij})$. ☐

Remark 5.3.5. *If V, W, U are vector spaces over the same field $\mathbb{F}$ and if $\mathrm{F}_i : V \times W \to U$ are bilinear maps for $i = 1, 2$, then $\mathrm{F}_1 + \mathrm{F}_2 : V \times W \to U$ is a bilinear map, and $\lambda \mathrm{F}_i$ is also a bilinear map. Thus, the set of bilinear maps from $V \times W$ to U is a vector space.*

Let us end this section with two questions:

Question: Let $V = W = \mathbb{R}^n$ and consider the map F defined by the usual scalar product

$$\mathrm{F}(X, Y) = X \cdot Y \qquad \text{for any } X, Y \in \mathbb{R}^n.$$

In view of Lemma 5.3.4, what is the matrix associated with this bilinear map with respect to the canonical basis of $\mathbb{R}^n$?

Question: How does a bilinear map change when one performs a change of bases for the vector spaces V and W ?

5.4 Exercises

Exercise 5.1. *Let V be a real vector space endowed with a scalar product. Prove the following relations for $X, Y \in V$ and $\lambda \in \mathbb{R}$:*

(i) $\|\lambda X\| = |\lambda|\,\|X\|$,

(ii) $\|X+Y\|^2 = \|X\|^2 + \|Y\|^2$ *if and only if* $X \perp Y$,

(iii) $\|X+Y\|^2 + \|X-Y\|^2 = 2\|X\|^2 + 2\|Y\|^2$,

(iv) $\|X+Y\| \leq \|X\| + \|Y\|$.

Exercise 5.2. *Let $\mathcal{A} = (a_{jk}) \in M_n(\mathbb{R})$ and define $\mathrm{Tr}(\mathcal{A}) = \sum_{j=1}^n a_{jj}$, where $\mathrm{Tr}(\mathcal{A})$ is called* the trace of $\mathcal{A}$. *Show the following properties:*

(i) $\mathrm{Tr} : M_n(\mathbb{R}) \to \mathbb{R}$ *is a linear map,*

(ii) $\mathrm{Tr}(\mathcal{A}\mathcal{B}) = \mathrm{Tr}(\mathcal{B}\mathcal{A})$, *for any* $\mathcal{A}, \mathcal{B} \in M_n(\mathbb{R})$,

(iii) If $\mathcal{C} \in M_n(\mathbb{R})$ *is an invertible matrix, then* $\mathrm{Tr}(\mathcal{C}^{-1}\mathcal{A}\mathcal{C}) = \mathrm{Tr}(\mathcal{A})$,

(iv) If $M_n^s(\mathbb{R})$ *denotes the vector space of all $n \times n$ symmetric matrices, then the map*

$$M_n^s(\mathbb{R}) \times M_n^s(\mathbb{R}) \ni (\mathcal{A}, \mathcal{B}) \mapsto \mathrm{Tr}(\mathcal{A}\mathcal{B}) \in \mathbb{R}$$

defines a scalar product on $M_n^s(\mathbb{R})$. We recall that a matrix $\mathcal{A}$ is symmetric if $\mathcal{A} = {}^t\mathcal{A}$.

Exercise 5.3. *Find an orthonormal basis for the subspace of $\mathbb{R}^4$ defined by the three vectors* $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}$ *and* $\begin{pmatrix} 1 \\ -2 \\ 0 \\ 0 \end{pmatrix}$.

Exercise 5.4. *Find an orthonormal basis for the space of solutions of the following systems:*

$$a)\ \begin{cases} 2x+y-z &= 0 \\ 2x+y+z &= 0 \end{cases} \quad b)\ \begin{cases} x-y+z &= 0 \end{cases} \quad c)\ \begin{cases} 4x+7y-\pi z &= 0 \\ 2x-y+z &= 0 \end{cases}$$

$$d)\ \begin{cases} x+y+z &= 0 \\ x-y &= 0 \\ y+z &= 0 \end{cases}$$

Exercise 5.5. *We consider the real vector space $V := C([0,1])$ made of continuous real functions on $[0,1]$ and endow it with the map*

$$V \times V \ni (f,g) \mapsto \langle f, g \rangle := \int_0^1 f(x)\,g(x)\,\mathrm{d}x \ \in \mathbb{R}.$$

Show that

(i) $\langle \cdot, \cdot \rangle$ *is a scalar product on* V,

(ii) *If* W *is the subspace of* V *generated by the three functions* $x \mapsto 1$ *(constant function),* $x \mapsto x$ *(identity function), and* $x \mapsto x^2$*, find an orthonormal basis for* W.

Exercise 5.6. *For any symmetric matrix* $\mathcal{A} = (a_{ij}) \in M_n(\mathbb{R})$*, we define the map*

$$\mathrm{F}_{\mathcal{A}} : \mathbb{R}^n \times \mathbb{R}^n \ni (X, Y) \mapsto \mathrm{F}_{\mathcal{A}}(X, Y) := {}^tX\mathcal{A}Y \in \mathbb{R}.$$

(i) *Show that* $\mathrm{F}_{\mathcal{A}}$ *is a bilinear map,*

(ii) *Show that* $\mathrm{F}_{\mathcal{A}}(X, Y) = \mathrm{F}_{\mathcal{A}}(Y, X)$ *for any* $X, Y \in \mathbb{R}^n$.

(iii) *When does* $\mathrm{F}_{\mathcal{A}}$ *define a scalar product ?*

(iv) *If* $\mathcal{A}$ *is one of the following matrices, does* $\mathrm{F}_{\mathcal{A}}$ *define a scalar product ?*

$$\mathcal{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \qquad \mathcal{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$