CHAPTER I

Duality of finite-dimensional vector spaces

1. Dual space

Let E be a finite-dimensional vector space over a field K. The vector space of linear maps $E \to K$ is denoted by E^*, so

$$E^* = L(E, K).$$

This vector space is called the *dual space* of E. Its elements are called *linear forms* on E. For $\varphi \in E^*$ and $u \in E^*$, let

$$\langle \varphi, u \rangle = \varphi(u) \in K.$$

1.1. Proposition. *For φ, $\psi \in E^*$, u, $v \in E$ and α, $\beta \in K$,*

1. $\langle \alpha\varphi + \beta\psi, u \rangle = \alpha\langle \varphi, u \rangle + \beta\langle \psi, u \rangle$.
2. $\langle \varphi, u\alpha + v\beta \rangle = \langle \varphi, u \rangle\alpha + \langle \varphi, v \rangle\beta$.
3. *If $\langle \varphi, u \rangle = 0$ for all $u \in E$, then $\varphi = 0$.*
4. *If $\langle \varphi, u \rangle = 0$ for all $\varphi \in E^*$, then $u = 0$.*

Proof. 1. This follows from the definition of the sum and the scalar multiplication in E^*.
2. This follows from the fact that φ is a linear map.
3. This is the definition of the zero map.
4. If $u \neq 0$, then u is the first element in some basis of E. Let $(u, e_2, \dots, e_n)$ be such a basis. From the construction principle for linear maps, it follows that there is a linear form $\varphi \in E^*$ which maps u to 1 and the other basis vectors to 0. This proves the contrapositive of implication 4.

$\square$

1.1. Dual basis

Let $e = (e_1, \dots, e_n)$ be a basis of E. By the construction principle for linear maps, there exists for all $i = 1, \dots, n$ a linear form $e_i^* \in E^*$ which maps e_i to 1 and the other basis vectors to 0, i.e.,

$$\langle e_i^*, e_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus, for $u = \sum_{j=1}^n e_j u_j \in E$ we have

$$\langle e_i^*, u \rangle = \sum_{j=1}^n \langle e_i^*, e_j \rangle u_j = \sum_{j=1}^n \delta_{ij} u_j = u_i,$$

which shows that the linear form e_i^* maps every vector of E to its i-th coordinate with respect to the basis e. Therefore, for all $u \in E$,

$$u = \sum_{i=1}^n e_i \langle e_i^*, u \rangle. \tag{1}$$

1.2. Theorem. *The sequence $e^* = (e_1^*, \dots, e_n^*)$ is a basis of E^*. In particular,* $\dim E^* = \dim E$.

Proof. *The elements e_1^*, ..., e_n^* are linearly independent:* Let $\sum_{i=1}^n \alpha_i e_i^* = 0$. For $j = 1, \dots, n$,

$$0 = \Big\langle \sum_{i=1}^n \alpha_i e_i^*, e_j \Big\rangle = \sum_{i=1}^n \alpha_i \langle e_i^*, e_j \rangle = \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j.$$

The sequence e^* spans E^*: Let us show that for all $\varphi \in E^*$,

$$\varphi = \sum_{i=1}^{n} \langle \varphi, e_i \rangle e_i^*. \tag{2}$$

It suffices to show that both sides take the same value on every vector $u \in E$. From (1), it follows that

$$\langle \varphi, u \rangle = \Big\langle \varphi, \sum_{j=1}^{n} e_j \langle e_j^*, u \rangle \Big\rangle = \sum_{j=1}^{n} \langle \varphi, e_j \rangle \langle e_j^*, u \rangle.$$

On the other hand,

$$\Big\langle \sum_{i=1}^{n} \langle \varphi, e_i \rangle e_i^*, u \Big\rangle = \sum_{i=1}^{n} \langle \varphi, e_i \rangle \langle e_i^*, u \rangle.$$

□

1.2. Bidual space

Every vector $u \in E$ defines a linear form $\mathrm{ev}_u \colon E^* \to K$ by

$$\mathrm{ev}_u(\varphi) = \varphi(u).$$

Thus, $\mathrm{ev}_u \in (E^*)^*$. Moreover, the map

$$\mathrm{ev} \colon E \to E^{**}; \qquad u \mapsto \mathrm{ev}_u$$

is linear since for u, $v \in E$, α, $\beta \in K$ and $\varphi \in E^*$,

$$\mathrm{ev}_{u\alpha+v\beta}(\varphi) = \varphi(u\alpha + v\beta) = \varphi(u)\alpha + \varphi(v)\beta = \mathrm{ev}_u(\varphi)\alpha + \mathrm{ev}_v(\varphi)\beta.$$

1.3. Theorem. *The map $\mathrm{ev} \colon E \to E^{**}$ is a vector space isomorphism.*

Proof. First, we show that ev is injective: if $u \in E$ is such that $\mathrm{ev}_u = 0$, then $\langle \varphi, u \rangle = 0$ for all $\varphi \in E^*$, hence $u = 0$ by Proposition 1.1. Therefore, $\operatorname{Ker} \mathrm{ev} = 0$ and ev is injective.

On the other hand, applying Theorem 1.2 twice, we find

$$\dim E^{**} = \dim E^* = \dim E.$$

Since ev is injective, it is therefore also surjective. □

The isomorphism ev is *natural*, inasmuch as it does not depend on the choice of bases in E and E^{**}. It is used to *identify* these two vector spaces. Thus, for $u \in E$, it is agreed that

$$u = \mathrm{ev}_u \in E^{**}.$$

Therefore, $\langle u, \varphi \rangle$ is defined for $u \in E$ and $\varphi \in E^*$, and we have

$$\langle u, \varphi \rangle = \langle \mathrm{ev}_u, \varphi \rangle = \varphi(u) = \langle \varphi, u \rangle.$$

2. Orthogonality

For every subspace $V \subset E$, the *orthogonal subspace* $V^0 \subset E^*$ is defined by

$$V^0 = \{\varphi \in E^* \mid \langle \varphi, u \rangle = 0 \text{ for all } u \in V\}.$$

It is easily checked that V^0 is a subspace of E^*.

2.1. Proposition. $\dim V^0 = \dim E - \dim V$.

Proof. Let $(e_1, \dots, e_r)$ be a basis of V, which we extend into a basis $e = (e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ of E. Consider the last $n - r$ elements of the dual basis $e^* = (e_1^*, \dots, e_r^*, \dots, e_n^*)$. We shall show:

$$V^0 = \operatorname{span}\langle e_{r+1}^*, \dots, e_n^* \rangle.$$

The proposition follows, since the right side is a subspace of dimension $n - r$, as e_{r+1}^*, …, e_n^* are linearly independent.

Equation (2) shows that every $\varphi \in V^0$ is a linear combination of e_{r+1}^*, …, e_n^*. Therefore,

$$V^0 \subset \operatorname{span}\langle e_{r+1}^*, \dots, e_n^* \rangle.$$

On the other hand, every $v \in V$ is a linear combination of $(e_1, \dots, e_r)$, hence equation (1) shows that

$$\langle e_i^*, v\rangle = 0 \qquad \text{for } i = r+1, \dots, n.$$

Therefore, $e_{r+1}^*, \dots, e_n^* \in V^0$, hence

$$\operatorname{span}\langle e_{r+1}^*, \dots, e_n^*\rangle \subset V^0.$$

$\square$

2.2. Corollary. *For every subspace $V \subset E$,*

$$V^{00} = V$$

(under the identification of E and E^{**}, since $V^{00} \subset E^{**}$).

Proof. By definition,

$$V^{00} = \{u \in E \mid \langle u, \varphi\rangle = 0 \text{ for all } \varphi \in V^0\}.$$

Since $\langle \varphi, u\rangle = 0$ for $u \in V$ and $\varphi \in V^0$, we have $V \subset V^{00}$. Now, the preceding proposition and Theorem 1.2 yield

$$\dim V^{00} = \dim E^* - \dim V^0 = \dim E - \dim V^0 = \dim V.$$

Therefore, the inclusion $V \subset V^{00}$ is an equality. $\square$

3. Transposition

Let $A\colon E \to F$ be a linear map between finite-dimensional vector spaces. The transpose map

$$A^t\colon F^* \to E^*$$

is defined by

$$A^t(\varphi) = \varphi \circ A \qquad \text{for } \varphi \in F^*.$$

3.1. Proposition. *A^t is the only linear map from F^* to E^* such that*

$$\big\langle A^t(\varphi), u\big\rangle = \big\langle \varphi, A(u)\big\rangle \qquad \textit{for all } \varphi \in F^* \textit{ and all } u \in E.$$

Proof. For $\varphi \in F^*$ and $u \in E$ we have

$$\big\langle A^t(\varphi), u\big\rangle = \langle \varphi \circ A, u\rangle = \varphi \circ A(u) = \big\langle \varphi, A(u)\big\rangle,$$

hence A^t satisfies the property.

Suppose $B\colon F^* \to E^*$ is another linear map with the same property. Then for all $\varphi \in F^*$ and all $u \in E$,

$$\big\langle B(\varphi), u\big\rangle = \big\langle \varphi, A(u)\big\rangle = \big\langle A^t(\varphi), u\big\rangle,$$

hence $B(\varphi) = A^t(\varphi)$ and therefore $B = A^t$. $\square$

Now, let e and f be arbitrary bases of E and F respectively, and let e^*, f^* be the dual bases.

3.2. Proposition. *The matrices of A^t and A are related as follows:*

$${}_{e^*}(A^t)_{f^*} = {}_f(A)_e^t.$$

Proof. Let ${}_f(A)_e = (a_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$, i.e.,

$$A(e_j) = \sum_{i=1}^m f_i a_{ij} \qquad \text{for } j = 1, \dots, n \tag{1}$$

and ${}_{e^*}(A^t)_{f^*} = (a'_{ij})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$, i.e.,

$$A^t(f_j^*) = \sum_{i=1}^n e_i^* a'_{ij} \qquad \text{for } j = 1, \dots, m.$$

We have to prove $a'_{ij} = a_{ji}$ for $i = 1, \dots, n$ and $j = 1, \dots, m$. Substituting $A^t(f_j^*)$ for φ in equation (2) of section 1.1, we obtain

$$A^t(f_j^*) = \sum_{i=1}^{n} \langle A^t(f_j^*), e_i \rangle e_i^* \qquad \text{for } j = 1, \dots, m,$$

hence $a'_{ij} = \langle A^t(f_j^*), e_i \rangle$. By Proposition 3.1, it follows that

$$a'_{ij} = \langle f_j^*, A(e_i) \rangle.$$

On the other hand, we may compute the right side of this last equation by using equation (1), which yields

$$\langle f_j^*, A(e_i) \rangle = \Big\langle f_j^*, \sum_{k=1}^{n} f_k a_{ki} \Big\rangle = \sum_{k=1}^{n} \langle f_j^*, f_k \rangle a_{ki} = a_{ji}.$$

Thus, $a'_{ij} = a_{ji}$ for $i = 1, \dots, n$ and $j = 1, \dots, m$. $\square$

3.3. Proposition. 1. *For A, $B\colon E \to F$ and α, $\beta \in K$,*

$$(\alpha A + \beta B)^t = \alpha A^t + \beta B^t.$$

2. *For $A\colon E \to F$ and $B\colon F \to G$,*

$$(B \circ A)^t = A^t \circ B^t.$$

3. *For $A\colon E \to F$,*

$$(A^t)^t = A$$

(under the identifications $E^{**} = E$ and $F^{**} = F$, since $(A^t)^t\colon E^{**} \to F^{**}$).

Proof. These properties can be proved either by reduction to the corresponding properties of matrices (by using arbitrary bases of the relevant spaces), or by using the characterization of transpose maps in Proposition 3.1. $\square$

CHAPTER II

Quotient spaces

1. Definition

Let E be an arbitrary vector space over a field K and let $V \subset E$ be an arbitrary subspace. For $x \in E$ let

$$x + V = \{x + v \mid v \in V\} \subset E.$$

1.1. Lemma. *For x, $y \in E$, the equation $x + V = y + V$ is equivalent to $x - y \in V$.*

Proof. Suppose first $x+V = y+V$. Since $x = x+0 \in x+V$, we have $x \in y+V$, hence $x = y+v$ for some $v \in V$ and therefore $x - y = v \in V$.

Conversely, suppose $x - y = v \in V$. For all $w \in V$ we then have $x + w = y + (v + w) \in y + V$, hence $x + V \subset y + V$. We also have

$$y + w = x + (w - v) \in x + V,$$

hence $y + V \subset x + V$, and therefore $x + V = y + V$. $\square$

Let

$$E/V = \{x + V \mid x \in E\}.$$

An addition and a scalar multiplication are defined on this set by

$$(x + V) + (y + V) = (x + y) + V \qquad \text{for } x,\ y \in E,$$
$$(x + V)\alpha = x\alpha + V \qquad \text{for } x \in E \text{ and } \alpha \in K.$$

It is easily checked that these operations are well-defined and endow the set E/V with a K-vector space structure.

In view of the definitions above, the map

$$\varepsilon_V \colon E \to E/V$$

defined by

$$\varepsilon_V(x) = x + V$$

is a surjective linear map. It is called the *canonical epimorphism* of E onto E/V. Its kernel is the set of $x \in E$ such that $x + V = 0 + V$. The lemma shows that this condition holds if and only if $x \in V$; therefore

$$\operatorname{Ker} \varepsilon_V = V.$$

1.2. Proposition. *Suppose E is finite-dimensional; then*

$$\dim E/V = \dim E - \dim V.$$

Moreover,

(a) *if $(e_1, \ldots, e_n)$ is a basis of E which extends a basis $(e_1, \ldots, e_r)$ of V, then $(e_{r+1}+V, \ldots, e_n+V)$ is a basis of E/V;*

(b) *if $(e_1, \ldots, e_r)$ is a basis of V and $u_{r+1}, \ldots, u_n$ are vectors of E such that $(u_{r+1}+V, \ldots, u_n+V)$ is a basis of E/V, then $(e_1, \ldots, e_r, u_{r+1}, \ldots, u_n)$ is a basis of E.*

PROOF. Since $\operatorname{Ker} \varepsilon_V = V$ and $\operatorname{Im} \varepsilon_V = E/V$, the relation between the dimensions of the kernel and the image of a linear map yields

$$\dim E = \dim V + \dim E/V.$$

(a) Since $\dim E/V = \dim E - \dim V = n - r$, it suffices to prove that the sequence $(e_{r+1}+V, \dots, e_n+V)$ spans E/V. Every $x \in E$ is a linear combination of $e_1, \dots, e_n$. If

$$x = e_1x_1 + \cdots + e_nx_n,$$

then

$$x + V = \varepsilon_V(x) = \varepsilon_V(e_1)x_1 + \cdots + \varepsilon_V(e_n)x_n.$$

Now, $\varepsilon_V(e_i) = 0$ for $i = 1, \dots, r$ since $e_i \in V$, hence

$$x + V = (e_{r+1} + V)x_{r+1} + \cdots + (e_n + V)x_n.$$

(b) Since $\dim E = \dim V + \dim E/V$, it suffices to prove that $e_1, \dots, e_r, u_{r+1}, \dots, u_n$ are linearly independent. Suppose

$$e_1\alpha_1 + \cdots + e_r\alpha_r + u_{r+1}\beta_{r+1} + \cdots + u_n\beta_n = 0. \tag{1}$$

Taking the image of each side under ε_V, we obtain $(u_{r+1} + V)\beta_{r+1} + \cdots + (u_n + V)\beta_n = 0$, hence $\beta_{r+1} = \cdots = \beta_n = 0$. Therefore, by (1) it follows that $\alpha_1 = \cdots = \alpha_r = 0$, since $(e_1, \dots, e_r)$ is a basis of V. □

2. Induced linear maps

Let $A\colon E \to F$ be a linear map and let $V \subset E$ be a subspace. A linear map $\overline{A}\colon E/V \to F$ is said to be *induced* by A if

$$\overline{A}(x + V) = A(x) \qquad \text{for all } x \in E.$$

In other words,

$$\overline{A} \circ \varepsilon_V = A.$$

This condition is also expressed as follows: the diagram

$$\begin{tikzcd} E \arrow[rr, "\varepsilon_V"] \arrow[dr, "A"'] & & E/V \arrow[dl, "\overline{A}"] \\ & F & \end{tikzcd}$$

is *commutative* or *commutes*.

2.1. PROPOSITION. *A linear map $A\colon E \to F$ induces a linear map*

$$\overline{A}\colon E/V \to F$$

if and only if $V \subset \operatorname{Ker} A$.

PROOF. If $\overline{A} \circ \varepsilon_V = A$, then $A(v) = 0$ for all $v \in V$ since $V = \operatorname{Ker} \varepsilon_V$. For the converse, suppose $V \subset \operatorname{Ker} A$ and consider $x, x' \in E$. If $x + V = x' + V$, then $x - x' \in \operatorname{Ker} A$, hence $A(x - x') = 0$ and therefore

$$A(x) = A(x').$$

We may then define a map $\overline{A}\colon E/V \to F$ by

$$\overline{A}(x + V) = A(x) \qquad \text{for all } x \in E.$$

It is easily checked that the map $\overline{A}$ thus defined is linear. □

Now, consider a linear map $A\colon E \to F$ and subspaces $V \subset E$ and $W \subset F$. A linear map

$$\overline{A}\colon E/V \to F/W$$

is said to be *induced* by A if

$$\overline{A}(x+V) = A(x) + W \qquad \text{for all } x \in E$$

or, alternatively, if $\overline{A} \circ \varepsilon_V = \varepsilon_W \circ A$. This condition can also be expressed by saying that the following diagram is *commutative*:

$$\begin{tikzcd}
E \arrow[r, "A"] \arrow[d, "\varepsilon_V"'] & F \arrow[d, "\varepsilon_W"] \\
E/V \arrow[r, "\overline{A}"] & F/W
\end{tikzcd}$$

or that $\overline{A}$ makes the above diagram *commute*.

2.2. Corollary. *The linear map $A\colon E \to F$ induces a linear map $\overline{A}\colon E/V \to F/W$ if and only if $A(V) \subset W$.*

Proof. Consider the linear map $A' = \varepsilon_W \circ A\colon E \to F/W$. The preceding proposition shows that there exists a linear map $\overline{A}$ such that $A' = \overline{A} \circ \varepsilon_V$ if and only if $V \subset \operatorname{Ker} A'$. In view of the definition of A', this condition is equivalent to

$$A(V) \subset \operatorname{Ker} \varepsilon_W = W.$$

□

If $A(V) \subset W$, we may restrict the linear map A to the subspace V and get a linear map

$$A|_V\colon V \to W.$$

We aim to compare the matrices of A, of $A|_V$ and of $\overline{A}\colon E/V \to F/W$.

Consider a basis $e' = (e_1, \dots, e_r)$ of V, a basis $\overline{e} = (e_{r+1} + V, \dots, e_n + V)$ of E/V and the sequence $e = (e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ which, according to Proposition 1.2, is a basis of E. Similarly, let $f' = (f_1, \dots, f_s)$ be a basis of W, $\overline{f} = (f_{s+1} + W, \dots, f_m + W)$ a basis of F/W and consider the basis $f = (f_1, \dots, f_s, f_{s+1}, \dots, f_m)$ of F.

2.3. Proposition.

$${}_f(A)_e = \left(\begin{array}{c|c} {}_{f'}(A|_V)_{e'} & * \\ \hline 0 & {}_{\overline{f}}(\overline{A})_{\overline{e}} \end{array}\right).$$

Proof. Let $A(e_j) = \sum_{i=1}^m f_i a_{ij}$ for $j = 1, \dots, n$, so that

$${}_f(A)_e = (a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}.$$

For $j = 1, \dots, r$ we have $e_j \in V$, hence $A(e_j) \in W$. Therefore, $A(e_j)$ is a linear combination of f_1, $\dots$, f_s, hence

$$a_{ij} = 0 \qquad \text{for } j = 1, \dots, r \text{ and } i = s+1, \dots, m.$$

Moreover, $A(e_j) = A|_V(e_j)$ for $j = 1, \dots, n$, hence $A|_V(e_j) = \sum_{i=1}^s f_i a_{ij}$ and therefore

$${}_{f'}(A|_V)_{e'} = (a_{ij})_{\substack{1 \leqslant i \leqslant s \\ 1 \leqslant j \leqslant r}}.$$

On the other hand, for $j = r+1, \dots, n$,

$$\overline{A}(e_j + V) = A(e_j) + W = \sum_{i=1}^m (f_i + W) a_{ij}.$$

For $i = 1, \dots, s$, we have $f_i \in W$, hence $f_i + W = 0$. The preceding equation thus yields $\overline{A}(e_j + V) = \sum_{i=s+1}^m (f_i + W) a_{ij}$ for $j = r+1, \dots, n$, hence

$${}_{\overline{f}}(\overline{A})_{\overline{e}} = (a_{ij})_{\substack{s+1 \leqslant i \leqslant m \\ r+1 \leqslant j \leqslant n}}.$$

The proof is thus complete. $\square$

In particular, if $F = E$ and $W = V$,

$$ {}_e(A)_e = \left(\begin{array}{c|c} {}_{e'}(A|_V)_{e'} & * \\ \hline 0 & {}_{\overline{e}}(\overline{A})_{\overline{e}} \end{array}\right) $$

and since the determinant of a block-triangular matrix is the product of the determinants of the blocks the following corollary follows:

2.4. Corollary. $\det A = \det A|_V \cdot \det \overline{A}$ *and* $\mathrm{Pc}_A = \mathrm{Pc}_{A|_V} \cdot \mathrm{Pc}_{\overline{A}}$.

3. Triangulation

3.1. Proposition. *Let* $A\colon E \to E$ *be a linear operator on a finite-dimensional vector space* E*. There exists a basis* e *of* E *such that the matrix* ${}_e(A)_e$ *is upper triangular if and only if the characteristic polynomial* Pc_A *decomposes into a product of factors of degree* 1*.*

Proof. If the matrix ${}_e(A)_e$ is upper triangular, let

$$ {}_e(A)_e = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, $$

then $\mathrm{Pc}_A(X) = (X - \lambda_1)\dots(X - \lambda_n)$.

Conversely, suppose Pc_A decomposes into a product of factors of degree 1. We argue by induction on the dimension of E. If $\dim E = 1$, there is nothing to prove since every matrix of order 1 is upper triangular. Let $\dim E = n$. By hypothesis, the characteristic polynomial of A has a root λ. Let $e_1 \in E$ be an eigenvector of A with eigenvalue λ and let $V = e_1 K$ be the vector space spanned by e_1. We have

$$ A(V) = A(e_1)K = e_1 \lambda K \subset V, $$

hence A induces a linear operator

$$ \overline{A}\colon E/V \to E/V. $$

By Corollary 2.4, we have $\mathrm{Pc}_A(X) = (X - \lambda)\,\mathrm{Pc}_{\overline{A}}(X)$, hence the characteristic polynomial of $\overline{A}$ decomposes into a product of factors of degree 1. By induction, there exists a basis $\overline{e} = (e_2 + V, \dots, e_n + V)$ of E/V such that the matrix ${}_{\overline{e}}(\overline{A})_{\overline{e}}$ is upper triangular. Proposition 1.2 shows that the sequence $e = (e_1, e_2, \dots, e_n)$ is a basis of E. Moreover, by Proposition 2.3,

$$ {}_e(A)_e = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & {}_{\overline{e}}(\overline{A})_{\overline{e}} \end{array}\right). $$

Since the matrix ${}_{\overline{e}}(\overline{A})_{\overline{e}}$ is upper triangular, so is ${}_e(A)_e$. $\square$

CHAPTER III

Tensor product

1. Tensor product of vector spaces

Let $E_1, \dots, E_n, F$ be vector spaces over a field K. Recall that a map

$$f\colon\ E_1 \times \cdots \times E_n \to F$$

is called *n-linear* or *multilinear* if

$$f(x_1, \dots, x_{i-1}, x_i\alpha + x_i'\alpha', x_{i+1}, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)\alpha + f(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n)\alpha'$$

for each $i = 1, \dots, n$ and for all $x_1 \in E_1, \dots, x_{i-1} \in E_{i-1}$, $x_i, x_i' \in E_i$, $x_{i+1} \in E_{i+1}, \dots, x_n \in E_n$ and $\alpha, \alpha' \in K$. In other words, for each $i = 1, \dots, n$ and for all $x_1 \in E_1, \dots, x_{i-1} \in E_{i-1}$, $x_{i+1} \in E_{i+1}, \dots, x_n \in E_n$ fixed, the map

$$f(x_1, \dots, x_{i-1}, \bullet, x_{i+1}, \dots, x_n)\colon\ E_i \to F$$

which carries $x \in E_i$ to $f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$ is linear. For example, all the products for which the distributive law holds are bilinear maps.

Clearly, if $f\colon\ E_1 \times \cdots \times E_n \to F$ is multilinear and $A\colon\ F \to G$ is linear, then the composite map $A \circ f\colon\ E_1 \times \cdots \times E_n \to G$ is multilinear. It turns out that, for given $E_1, \dots, E_n$, there exists a "universal" linear map $\otimes\colon\ E_1 \times \cdots \times E_n \to E_1 \otimes \cdots \otimes E_n$, called the *tensor product* of $E_1, \dots, E_n$, from which all the multilinear maps originating in $E_1 \times \cdots \times E_n$ can be derived. The goal of this section is to define this universal map.

1.1. Universal property

Let $E_1, \dots, E_n$ be vector spaces over a field K. A pair (f, F) consisting of a K-vector space F and a multilinear map

$$f\colon\ E_1 \times \cdots \times E_n \to F$$

is called a *universal product* of $E_1, \dots, E_n$ if the following property (called the *universal property*) holds: for every K-vector space G and every multilinear map

$$g\colon\ E_1 \times \cdots \times E_n \to G$$

there exists a unique linear map $A\colon\ F \to G$ such that $g = A \circ f$ or, in other words, which makes the following diagram commute:

$$\begin{tikzcd} E_1 \times \cdots \times E_n \arrow[rr, "f"] \arrow[dr, "g"'] & & F \arrow[dl, "A"] \\ & G. & \end{tikzcd}$$

The existence of a universal product is far from obvious. By contrast, the uniqueness (up to isomorphism) easily follows from the definition by abstract nonsense.

1.1. Proposition. *Suppose (f, F) and (f', F') are two universal products of $E_1, \dots, E_n$. Then there is one and only one vector space isomorphism $\varphi\colon\ F \to F'$ such that the following diagram* *commutes:*

$$\begin{tikzcd} E_1 \times \cdots \times E_n \arrow[rr, "f"] \arrow[dr, "f'"'] & & F \arrow[dl, "\varphi"] \\ & F'. & \end{tikzcd}$$

PROOF. The existence of a unique linear map $\varphi\colon F \to F'$ such that $f' = \varphi \circ f$ follows from the universal property of (f, F). We have to show that φ is bijective. To achieve this goal, observe that the universal property of (f', F') yields a linear map $\varphi'\colon F' \to F$ such that $f = \varphi' \circ f'$. Now, the following diagrams commute:

$$\begin{tikzcd} E_1 \times \cdots \times E_n \arrow[rr, "f"] \arrow[dr, "f"'] & & F \arrow[dl, "\mathrm{Id}_F"] \\ & F & \end{tikzcd} \qquad \begin{tikzcd} E_1 \times \cdots \times E_n \arrow[rr, "f"] \arrow[dr, "f"'] & & F \arrow[dl, "\varphi' \circ \varphi"] \\ & F. & \end{tikzcd}$$

In view of the uniqueness condition in the universal property of (f, F), it follows that $\varphi' \circ \varphi = \mathrm{Id}_F$. Similarly, the universal property of (f', F') shows that $\varphi \circ \varphi' = \mathrm{Id}_{F'}$. Therefore, φ and φ' are reciprocal bijections. □

This proposition shows that a universal product of $E_1, \ldots, E_n$, if it exists, can be considered as unique. We shall refer to this universal product as *the tensor product* of $E_1, \ldots, E_n$ and denote it as $(\otimes, E_1 \otimes \cdots \otimes E_n)$. The image of an n-tuple $(x_1, \ldots, x_n) \in E_1 \times \cdots \times E_n$ under $\otimes$ is denoted $x_1 \otimes \cdots \otimes x_n$.

Our next goal is to prove the existence of the tensor product. We consider first the case where $n = 2$ and then derive the general case by associativity.

1.2. The tensor product of two vector spaces

1.2. PROPOSITION. *Let $\mu\colon K^m \times K^n \to K^{m\times n}$ be defined by*

$$\mu\big((x_i)_{1\leqslant i\leqslant m}, (y_j)_{1\leqslant j\leqslant n}\big) = (x_i y_j)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}.$$

The pair $(\mu, K^{m\times n})$ is a universal product of K^m and K^n.

PROOF. Denoting the elements of K^m and K^n as column vectors, we have

$$\mu(x, y) = x \cdot y^t \qquad \text{for } x \in K^m \text{ and } y \in K^n,$$

where $\cdot$ denotes the matrix multiplication $K^{m\times 1} \times K^{1\times n} \to K^{m\times n}$. Therefore, the bilinearity of μ follows from the properties of the matrix multiplication.

Let $(c_i)_{1\leqslant i\leqslant m}$ (resp. $(c'_j)_{1\leqslant j\leqslant n}$) be the canonical basis of K^m (resp. K^n). Let also

$$e_{ij} = \mu(c_i, c'_j) \qquad \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n.$$

Thus, $e_{ij} \in K^{m\times n}$ is the matrix whose only non-zero entry is a 1 at the i-th row and j-th column. The family $(e_{ij})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ is a basis of $K^{m\times n}$. Therefore, given any bilinear map

$$B\colon K^m \times K^n \to F,$$

we may use the construction principle to obtain a linear map $\varphi\colon K^{m\times n} \to F$ such that $\varphi(e_{ij}) = B(c_i, c'_j)$. For all $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in K^m$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$, we have $x = \sum_{i=1}^m c_i x_i$, $y = \sum_{j=1}^n c'_j y_j$, hence

$$\mu(x, y) = \sum_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} x_i y_j \mu(c_i, c'_j) = \sum_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} x_i y_j e_{ij}$$

and

$$B(x,y) = \sum_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} x_i y_j B(c_i, c'_j) = \sum_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} x_i y_j \varphi(e_{ij}) = \varphi\Big(\sum_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} x_i y_j e_{ij}\Big).$$

Therefore, $B = \varphi \circ \mu$.

Uniqueness of the linear map $\varphi\colon K^{m\times n} \to F$ such that $B = \varphi\circ\mu$ is clear, since this equation yields

$$B(c_i, c'_j) = \varphi\big(\mu(c_i, c'_j)\big) = \varphi(e_{ij}),$$

and a linear map is uniquely determined by the image of a basis. Thus, $(\mu, K^{m\times n})$ is a universal product of K^m and K^n. □

1.3. Theorem. *Any two vector spaces over a field K have a tensor product.*

Proof. We give the proof for finite-dimensional vector spaces only. If $\dim E = m$ and $\dim F = n$, we may identify E to K^m and F to K^n by the choice of arbitrary bases of E and of F (by mapping every vector in E (resp. F) to the m-tuple (resp. n-tuple) of its coordinates with respect to the chosen basis). The preceding proposition then shows that $(\mu, K^{m\times n})$ is a universal tensor product of E and F. □

Note that, even though bases of E and F were chosen in the proof of the theorem, the tensor product itself does not depend on the choice of bases (but its identification with $(\mu, K^{m\times n})$ does). It is a *canonical construction*. To simplify the notation, we suppress the bilinear map $\otimes$ in the tensor product notation $(\otimes, E\otimes F)$, and refer to the tensor product of E and F simply as $E\otimes F$.

1.4. Corollary. *If $e = (e_i)_{1\leqslant i\leqslant m}$ is a basis of E and $f = (f_j)_{1\leqslant j\leqslant n}$ is a basis of F, then $(e_i\otimes f_j)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}$ is a basis of $E\otimes F$. In particular,* $\dim(E\otimes F) = \dim E \dim F$.

Proof. It was observed in the proof of Proposition 1.2 that $\big(\mu(c_i, c'_j)\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}$ is a basis of $K^{m\times n}$. If E is identified with K^m and F with K^n by means of the bases e and f, then e and f correspond to the canonical bases of K^m and K^n respectively, hence $\mu(c_i, c'_j) = e_i\otimes f_j$. □

The corollary shows that every element in $E\otimes F$ has the form $\sum_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} \alpha_{ij} e_i\otimes f_j$ for some $\alpha_{ij}\in K$. Grouping terms, we obtain

$$\sum_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} \alpha_{ij} e_i\otimes f_j = \sum_{j=1}^{n}\Big(\sum_{i=1}^{m}\alpha_{ij}e_i\Big)\otimes f_j$$

hence every element in $E\otimes F$ can be written in the form

$$x_1\otimes y_1 + \cdots + x_r\otimes y_r$$

for suitable $x_1, \ldots, x_r \in E$ and $y_1, \ldots, y_r\in F$ (and $r\in\mathbb{N}$). Note that this expression is *not* unique: for instance,

$$(e_1+e_2)\otimes f_1 + (e_1-e_2)\otimes f_2 = e_1\otimes(f_1+f_2) + e_2\otimes(f_1-f_2).$$

1.3. The tensor product of three or more vector spaces

Let E_1, E_2, E_3 be vector spaces over K. We may consider the tensor products $E_1\otimes E_2$ and $(E_1\otimes E_2)\otimes E_3$. The map

$$\otimes'\colon E_1\times E_2\times E_3 \to (E_1\otimes E_2)\otimes E_3$$

defined by $\otimes'(x_1,x_2,x_3) = (x_1\otimes x_2)\otimes x_3$ is 3-linear.

1.5. Proposition. *The pair $\big(\otimes', (E_1\otimes E_2)\otimes E_3\big)$ is a universal product of E_1, E_2 and E_3.*

PROOF. Let $T\colon E_1 \times E_2 \times E_3 \to F$ be a 3-linear map. For all $x_3 \in E_3$, the map

$$T(\bullet, \bullet, x_3)\colon E_1 \times E_2 \to F$$

is bilinear, hence there is a unique linear map $\varphi_{x_3}\colon E_1 \otimes E_2 \to F$ such that

$$\varphi_{x_3}(x_1 \otimes x_2) = T(x_1, x_2, x_3) \qquad \text{for all } x_1 \in E_1,\ x_2 \in E_2.$$

For $x_1 \in E_1$, $x_2 \in E_2$, x_3, $x_3' \in E_3$ and α, $\alpha' \in K$, we have

$$T(x_1, x_2, \alpha x_3 + \alpha' x_3') = \alpha T(x_1, x_2, x_3) + \alpha' T(x_1, x_2, x_3'),$$

hence

$$\varphi_{\alpha x_3 + \alpha' x_3'}(x_1 \otimes x_2) = \alpha \varphi_{x_3}(x_1 \otimes x_2) + \alpha' \varphi_{x_3'}(x_1 \otimes x_2).$$

Since $E_1 \otimes E_2$ has a basis consisting of products $x_1 \otimes x_2$, by Corollary 1.4, and since a linear map is uniquely determined by the image of a basis, it follows that

$$\varphi_{\alpha x_3 + \alpha' x_3'} = \alpha \varphi_{x_3} + \alpha' \varphi_{x_3'}. \tag{1}$$

Consider the map $\Phi\colon (E_1 \otimes E_2) \times E_3 \to F$ defined by

$$\Phi(\xi, x_3) = \varphi_{x_3}(\xi) \qquad \text{for } \xi \in E_1 \otimes E_2 \text{ and } x_3 \in E_3.$$

By (1) and since each map φ_{x_3} is linear, the map Φ is bilinear. Therefore, the universal property of tensor products yields a linear map

$$\varphi\colon (E_1 \otimes E_2) \otimes E_3 \to F$$

such that

$$\varphi(\xi \otimes x_3) = \Phi(\xi, x_3) \qquad \text{for all } \xi \in E_1 \otimes E_2,\ x_3 \in E_3.$$

In particular,

$$\varphi\big((x_1 \otimes x_2) \otimes x_3\big) = \varphi_{x_3}(x_1 \otimes x_2) = T(x_1, x_2, x_3),$$

hence $T = \varphi \circ \otimes'$.

Uniqueness of the map φ satisfying this equation is clear, since every element in $(E_1 \otimes E_2) \otimes E_3$ can be written as $\sum_i (x_i \otimes y_i) \otimes z_i = \sum_i \otimes'(x_i, y_i, z_i)$ for some $x_i \in E_1$, $y_i \in E_2$ and $z_i \in E_3$. If φ and φ' satisfy $T = \varphi \circ \otimes' = \varphi' \circ \otimes'$, we must have

$$\varphi\Big(\sum_i (x_i \otimes y_i) \otimes z_i\Big) = \sum_i T(x_i, y_i, z_i) = \varphi'\Big(\sum_i (x_i \otimes y_i) \otimes z_i\Big),$$

hence $\varphi = \varphi'$. □

The proposition above proves the existence of the tensor product $E_1 \otimes E_2 \otimes E_3$, and shows that there is a canonical isomorphism

$$E_1 \otimes E_2 \otimes E_3 = (E_1 \otimes E_2) \otimes E_3.$$

Obviously, we also have $E_1 \otimes E_2 \otimes E_3 = E_1 \otimes (E_2 \otimes E_3)$. By induction, it follows that the tensor product of arbitrarily many vector spaces exists, and may be defined by associativity.

The following corollary also follows by induction on the number of factors:

1.6. COROLLARY. *Let $(e_{ij_i})_{1 \leqslant j_i \leqslant n_i}$ be a basis of E_i, for $i = 1, \ldots, r$. Then*

$$(e_{1j_1} \otimes e_{2j_2} \otimes \cdots \otimes e_{rj_r})_{\substack{1 \leqslant j_1 \leqslant n_1 \\ \cdots \\ 1 \leqslant j_r \leqslant n_r}}$$

is a basis of $E_1 \otimes \cdots \otimes E_r$. In particular,

$$\dim E_1 \otimes \cdots \otimes E_r = \dim E_1 \ldots \dim E_r.$$

1.4. Example

Let $E_1, \dots, E_n$ be finite-dimensional K-vector spaces and let F be an arbitrary K-vector space. Denote by $M(E_1 \times \cdots \times E_n, F)$ the set of all multilinear maps

$$f\colon\ E_1 \times \cdots \times E_n \to F.$$

It is easily checked that $M(E_1 \times \cdots \times E_n, F)$ is a subspace of the space $\mathcal{F}(E_1 \times \cdots \times E_n, F)$ of all mappings $E_1 \times \cdots \times E_n \to F$.

Let E_i^* be the dual space of E_i for $i = 1, \dots, n$. For $y \in F$ and $\varphi_1 \in E_1^*, \dots, \varphi_n \in E_n^*$, define

$$\theta_{y,\varphi_1,\dots,\varphi_n}\colon\ E_1 \times \cdots \times E_n \to F$$

by

$$\theta_{y,\varphi_1,\dots,\varphi_n}(x_1, \dots, x_n) = y\langle\varphi_1, x_1\rangle \cdots \langle\varphi_n, x_n\rangle.$$

Clearly, the map $\theta_{y,\varphi_1,\dots,\varphi_n}$ is multilinear, hence

$$\theta_{y,\varphi_1,\dots,\varphi_n} \in M(E_1 \times \cdots \times E_n, F).$$

Moreover, it is also clear that

$$(y, \varphi_1, \dots, \varphi_n) \mapsto \theta_{y,\varphi_1,\dots,\varphi_n}$$

is a multilinear map $F \times E_1^* \times \cdots \times E_n^* \to M(E_1 \times \cdots \times E_n, F)$. We denote it by θ. By the universal property of tensor products, θ induces a linear map

$$\Theta\colon\ F \otimes E_1^* \otimes \cdots \otimes E_n^* \to M(E_1 \times \cdots \times E_n, F).$$

1.7. Theorem. *The map Θ is an isomorphism of vector spaces.*

Proof. We have to show that Θ is one-to-one and onto. To achieve this, we construct a reciprocal map. For $i = 1, \dots, n$, let $(e_{ij_i})_{1\leqslant j_i\leqslant m_i}$ be a basis of E_i, and let $(e_{ij_i}^*)_{1\leqslant j_i\leqslant m_i}$ be the dual basis. For every multilinear map $f \in M(E_1 \times \cdots \times E_n, F)$, define

$$\Psi(f) = \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} f(e_{1j_1}, \dots, e_{nj_n}) \otimes e_{1j_1}^* \otimes \cdots \otimes e_{nj_n}^* \in F \otimes E_1^* \otimes \cdots \otimes E_n^*.$$

The map $\Psi\colon\ M(E_1 \times \cdots \times E_n, F) \to F \otimes E_1^* \otimes \cdots \otimes E_n^*$ is linear. To prove that Ψ and Θ are reciprocal bijections, we shall show[1] that $\Psi \circ \Theta$ is the identity on $F \otimes E_1^* \otimes \cdots \otimes E_n^*$ and $\Theta \circ \Psi$ is the identity on $M(E_1 \times \cdots \times E_n, F)$.

Every element in $F \otimes E_1^* \otimes \cdots \otimes E_n^*$ can be written in the form

$$\xi = \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} y_{j_1\cdots j_n} \otimes e_{1j_1}^* \otimes \cdots \otimes e_{nj_n}^*.$$

For this ξ, we have

$$\Theta(\xi) = \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} \theta_{y_{j_1\cdots j_n}, e_{1j_1}^*, \dots, e_{nj_n}^*}$$

hence

$$\begin{aligned}\Theta(\xi)(e_{1k_1}, \dots, e_{nk_n}) &= \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} y_{j_1\cdots j_n}\langle e_{1j_1}^*, e_{1k_1}\rangle \cdots \langle e_{nj_n}^*, e_{nk_n}\rangle\\ &= y_{k_1\cdots k_n}\end{aligned}$$

for all $k_1 = 1, \dots, m_1, \dots, k_n = 1, \dots, m_n$.

[1] The notation makes the following proof hardly readable. The reader is warmly encouraged to check the statement by him/herself!

Now, by definition of Ψ we have

$$\begin{aligned}\Psi\circ\Theta(\xi) &= \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} \Theta(\xi)(e_{1k_1},\dots,e_{nk_n})\otimes e^*_{1k_1}\otimes\cdots\otimes e^*_{nk_n}\\ &= \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} y_{k_1\cdots k_n}\otimes e^*_{1k_1}\otimes\cdots\otimes e^*_{nk_n}\\ &= \xi.\end{aligned}$$

Therefore, $\Psi\circ\Theta$ is the identity on $F\otimes E_1^*\otimes\cdots\otimes E_n^*$.

Now, let $f\in M(E_1\times\cdots\times E_n,F)$. We have

$$\Theta\circ\Psi(f)=\sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} \theta_{f(e_{1j_1},\dots,e_{nj_n}),e^*_{1j_1},\dots,e^*_{nj_n}},$$

hence

$$\begin{aligned}\Theta\circ\Psi(f)(e_{1k_1},\dots,e_{nk_n}) &= \sum_{\substack{1\leqslant j_1\leqslant m_1\\ \cdots\\ 1\leqslant j_n\leqslant m_n}} f(e_{1j_1},\dots,e_{nj_n})\langle e^*_{1j_1},e_{1k_1}\rangle\cdots\langle e^*_{nj_n},e_{nk_n}\rangle\\ &= f(e_{1k_1},\dots,e_{nk_n})\end{aligned}$$

for all $k_1=1,\dots,n_1,\dots,k_n=1,\dots,m_n$. Since multilinear maps are uniquely determined by their values on basis elements, it follows that $\Theta\circ\Psi(f)=f$, hence $\Theta\circ\Psi$ is the identity on $M(E_1\times\cdots\times E_n,F)$. □

Note that the isomorphism Θ is canonical, even though its reciprocal isomorphism Ψ is defined above through the choice of bases of $E_1,\dots,E_n$. We may therefore use Θ to *identify*

$$F\otimes E_1^*\times\cdots\times E_n^*=M(E_1\times\cdots\times E_n,F).$$

This provides an alternative definition for the tensor product of vector spaces, at least when all but one of the factors are finite-dimensional.

2. Formalism

2.1. Covariant and contravariant behaviour

Let E be a finite-dimensional vector space over a field K, and let $(e_i)_{1\leqslant i\leqslant n}$ be a basis of E. It is sometimes convenient to write the coordinates of vectors $x\in E$ with *upper* indices instead of lower indices.[2] We thus write

$$x=\sum_{i=1}^n e_ix^i.$$

This notation is simplified by the *Einstein convention*: on every index appearing twice in a formula, the summation is understood:

$$x=e_ix^i\qquad(=x^ie_i).$$

For notational coherence, the entries in any change-of-basis matrix receive an upper index and a lower index: if $(e'_\alpha)_{1\leqslant\alpha\leqslant n}$ is another basis of E we write

$$e'_\alpha=e_ia^i_\alpha \tag{1}$$

(instead of $e'_\alpha=\sum_{i=1}^n e_ia_{i\alpha}$). So, $(a^i_\alpha)_{1\leqslant i\leqslant n}$ are the coordinates of e'_α with respect to $(e_i)_{1\leqslant i\leqslant n}$. The upper index is the row index and the lower index is the column index. If $(b^\alpha_i)_{1\leqslant i,\alpha\leqslant n}$ is the inverse matrix, we have (by definition of the inverse matrix)

$$a^i_\alpha b^\beta_i=\delta^\beta_\alpha\qquad\text{and}\qquad a^i_\alpha b^\alpha_j=\delta^i_j$$

[2] Of course, upper indices should not be confused with exponents. There should not be any confusion, since in (multi)linear algebra exponents appear very exceptionally.

where δ^i_j (resp. δ^β_α) is the Kronecker symbol, i.e., $\delta^i_j = 1$ if $i = j$ and $\delta^i_j = 0$ if $i \neq j$. We also have

$$e_i = e'_\alpha b^\alpha_i. \tag{2}$$

Now, consider a vector $x = e_i x^i = e'_\alpha x'^\alpha$. Substituting $e_i a^i_\alpha$ for e'_α, we obtain $x = e_i a^i_\alpha x'^\alpha$, hence

$$x^i = a^i_\alpha x'^\alpha. \tag{3}$$

Similarly,

$$x'^\alpha = b^\alpha_i x^i. \tag{4}$$

It is a crucial observation that, although (3) and (4) look like (1) and (2), these equations are essentially different: multiplying by the matrix a^i_α transforms e_i into e'_α while it has the reverse effect on coordinates, transforming x'^α into x^i.

Now, consider the dual bases $(e^*_i)_{1\leqslant i\leqslant n}$ and $(e'^*_\alpha)_{1\leqslant \alpha\leqslant n}$. (These are bases of the dual space E^*.) Equation (2) in section 1.1 of Chapter I yields

$$e'^*_\alpha = e^*_i \langle e'^*_\alpha, e_i\rangle \qquad \text{and} \qquad e^*_i = e'^*_\alpha \langle e^*_i, e'_\alpha\rangle$$

hence, taking into account (1) and (2),

$$e'^*_\alpha = e^*_i b^\alpha_i \qquad \text{and} \qquad e^*_i = e'^*_\alpha a^i_\alpha. \tag{5}$$

In view of these equations, notational coherence dictates to denote e^i for e^*_i and e'^α for e'^*_α; then (5) can be rewritten as

$$e'^\alpha = e^i b^\alpha_i \qquad \text{and} \qquad e^i = e'^\alpha a^i_\alpha. \tag{6}$$

These equations are exactly like (4) and (3).

To complete the picture, consider a linear form $\varphi \in E^*$. Equation (2) of section 1.1 of Chapter I yields

$$\varphi = e^i \langle \varphi, e_i\rangle = e'^\alpha \langle \varphi, e'_\alpha\rangle.$$

Therefore, the coordinates $(\varphi_i)_{1\leqslant i\leqslant n}$ (resp. $(\varphi'_\alpha)_{1\leqslant \alpha\leqslant n}$) of φ with respect to the basis $(e^i)_{1\leqslant i\leqslant n}$ (resp. $(e'^\alpha)_{1\leqslant \alpha\leqslant n}$) of E^* are given by

$$\varphi_i = \langle \varphi, e_i\rangle, \qquad \varphi'_\alpha = \langle \varphi, e'_\alpha\rangle.$$

Using (1) and (2), we get

$$\varphi'_\alpha = \langle \varphi, e_i\rangle a^i_\alpha = \varphi_i a^i_\alpha \qquad \text{and} \qquad \varphi_i = \langle \varphi, e'_\alpha\rangle b^\alpha_i = \varphi'_\alpha b^\alpha_i, \tag{7}$$

which should be compared to (3) and (4).

Thus, coordinates of linear forms behave under a change of basis exactly like basis vectors: this is the *covariant* behaviour. By contrast, coordinates of vectors and dual basis vectors behave in the opposite way to basis vectors: they display the so-called *contravariant* behaviour.

2.2. Covariant and contravariant tensors

For p and q any natural numbers, let

$$T^q_p(E) = \underbrace{E^* \otimes \cdots \otimes E^*}_{p} \otimes \underbrace{E \otimes \cdots \otimes E}_{q} \qquad \text{if } p,\ q \text{ are not both zero,}$$

and $T^0_0(E) = K$. The elements in $T^q_p(E)$ are called *p-covariant and q-contravariant tensors on E*. This is a shameless abuse of terminology, since the elements in $T^q_p(E)$ do not vary at all. It is motivated by the fact that the *coordinates* of elements in $T^q_p(E)$ have p covariant (i.e., lower) indices and q contravariant (i.e., upper) indices, as we proceed to show.

Let $(e_i)_{1\leqslant i\leqslant n}$ be a basis of E, and let $(e^i)_{1\leqslant i\leqslant n}$ be the dual basis, which is a basis of E^*. Corollary 1.6 shows that $(e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q})_{1\leqslant i_1,\dots,j_q\leqslant n}$ is a basis of $T^q_p(E)$, hence every element in $T^q_p(E)$ can be written as

$$t = t^{j_1\cdots j_q}_{i_1\cdots i_p} e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}. \tag{8}$$

Let $(e'_\alpha)_{1\leqslant\alpha\leqslant n}$ be another basis of E and $(e'^\alpha)_{1\leqslant\alpha\leqslant n}$ be its dual basis, and suppose the bases are related by

$$e'_\alpha = e_i a^i_\alpha \qquad \text{and} \qquad e_i = e'_\alpha b^\alpha_i \tag{9}$$

hence

$$e'^\alpha = e^i b^\alpha_i \qquad \text{and} \qquad e^i = e'^\alpha a^i_\alpha,$$

as in the preceding section. Substituting in (8), we get

$$\begin{aligned} t &= t^{j_1\cdots j_q}_{i_1\cdots i_p}(e'^{\alpha_1}a^{i_1}_{\alpha_1})\otimes\cdots\otimes(e'^{\alpha_p}a^{i_p}_{\alpha_p})\otimes(e'_{\beta_1}b^{\beta_1}_{j_1})\otimes\cdots\otimes(e'_{\beta_q}b^{\beta_q}_{j_q}) \\ &= \big(t^{j_1\cdots j_q}_{i_1\cdots i_p}a^{i_1}_{\alpha_1}\cdots a^{i_p}_{\alpha_p}b^{\beta_1}_{j_1}\cdots b^{\beta_q}_{j_q}\big)e'^{\alpha_1}\otimes\cdots\otimes e'^{\alpha_p}\otimes e'_{\beta_1}\otimes\cdots\otimes e'_{\beta_q}. \end{aligned}$$

Therefore, the coordinates $t'^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p}$ of $t\in T^q_p(E)$ with respect to the basis $(e'^{\alpha_1}\otimes\cdots\otimes e'^{\alpha_p}\otimes e'_{\beta_1}\otimes\cdots\otimes e'_{\beta_q})_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}$ are given by

$$t'^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p} = t^{j_1\cdots j_q}_{i_1\cdots i_p}a^{i_1}_{\alpha_1}\cdots a^{i_p}_{\alpha_p}b^{\beta_1}_{j_1}\cdots b^{\beta_q}_{j_q}. \tag{10}$$

This shows that the lower indices are indeed covariant while the upper indices are contravariant (as they should).

Similarly, we have

$$t^{j_1\cdots j_q}_{i_1\cdots i_p} = t'^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p}b^{\alpha_1}_{i_1}\cdots b^{\alpha_p}_{i_p}a^{j_1}_{\beta_1}\cdots a^{j_q}_{\beta_q}.$$

From a formal point of view, the change of basis formula (10) is the only condition for an indexed array to be the coordinates of a p-covariant and q-contravariant tensor, as the following tensoriality criterion shows.

2.1. Theorem. *Suppose that to each basis $e=(e_i)_{1\leqslant i\leqslant n}$ of E is attached an array of scalars $\big(u(e)^{j_1\cdots j_q}_{i_1\cdots i_p}\big)_{1\leqslant i_1,\dots,j_q\leqslant n}\in K^{n^{p+q}}$. There exists a tensor in $T^q_p(E)$ whose coordinates with respect to each basis e are $u(e)^{j_1\cdots j_q}_{i_1\cdots i_p}$ if and only if the arrays attached to any two bases e, e' related by (9) satisfy*

$$u(e')^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p} = u(e)^{j_1\cdots j_q}_{i_1\cdots i_p}a^{i_1}_{\alpha_1}\cdots a^{i_p}_{\alpha_p}b^{\beta_1}_{j_1}\cdots b^{\beta_q}_{j_q}.$$

Proof. The "if" part readily follows from the computation leading to equation (10). To prove the "only if" part, pick an arbitrary basis $\tilde{e}$ of E and define a tensor $u\in T^q_p(E)$ by

$$u = u(\tilde{e})^{j_1\cdots j_q}_{i_1\cdots i_p}\tilde{e}^{i_1}\otimes\cdots\otimes\tilde{e}^{i_p}\otimes\tilde{e}_{j_1}\otimes\cdots\otimes\tilde{e}_{j_q}.$$

By definition, $\big(u(\tilde{e})^{j_1\cdots j_q}_{i_1\cdots i_p}\big)_{1\leqslant i_1,\dots,j_q\leqslant n}$ is the array of coordinates of u with respect to the basis $(\tilde{e}^{i_1}\otimes\cdots\otimes\tilde{e}^{i_p}\otimes\tilde{e}_{j_1}\otimes\cdots\otimes\tilde{e}_{j_q})_{1\leqslant i_1,\dots,j_q\leqslant n}$. We have to show that for *any* basis $\hat{e}=(\hat{e}_\alpha)_{1\leqslant\alpha\leqslant n}$, the coordinates of u with respect to $(\hat{e}^{\alpha_1}\otimes\cdots\otimes\hat{e}^{\alpha_p}\otimes\hat{e}_{\beta_1}\otimes\cdots\otimes\hat{e}_{\beta_q})_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}$ are $\big(u(\hat{e})^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p}\big)_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}$. Suppose $\hat{e}_\alpha=\tilde{e}_i a^i_\alpha$ and $\tilde{e}_i=\hat{e}_\alpha b^\alpha_i$, hence $\hat{e}^\alpha=\tilde{e}^i b^\alpha_i$ and $\tilde{e}^i=\hat{e}^\alpha a^i_\alpha$. Then by (10) the coordinates of u with respect to $(\hat{e}^{\alpha_1}\otimes\cdots\otimes\hat{e}^{\alpha_p}\otimes\hat{e}_{\beta_1}\otimes\cdots\otimes\hat{e}_{\beta_q})_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}$ are

$$\Big(u(\tilde{e})^{j_1\cdots j_q}_{i_1\cdots i_p}a^{i_1}_{\alpha_1}\cdots a^{i_p}_{\alpha_p}b^{\beta_1}_{j_1}\cdots b^{\beta_q}_{j_q}\Big)_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}.$$

By hypothesis, this array is $\big(u(\hat{e})^{\beta_1\cdots\beta_q}_{\alpha_1\cdots\alpha_p}\big)_{1\leqslant\alpha_1,\dots,\beta_q\leqslant n}$. $\square$

2.3. Tensors on euclidean spaces

In this section, we assume that the vector space E is finite-dimensional and endowed with a scalar product

$$(\cdot|\cdot)\colon\ E\times E\to K,$$

i.e., a symmetric bilinear form which is non-degenerate in the sense that $(x|y)=0$ for all $y\in E$ implies $x=0$. Using the scalar product, we define a linear map

$$\Phi\colon\ E\to E^*$$

by mapping every vector $x \in E$ to the linear form $y \mapsto (x|y)$, so

$$\langle \Phi(x), y \rangle = (x|y) \quad (= (y|x)) \qquad \text{for } x,\, y \in E.$$

The non-degeneracy condition on the scalar product means that Φ is injective, hence it is bijective since $\dim E = \dim E^*$, by Theorem 1.2 of Chapter I. We may thus use Φ to identify E and E^*, hence also $T_p^q(E)$ with $T_{p+q}(E)$ or $T^{p+q}(E)$.

In the rest of this section, we show how this identification works at the level of coordinates.

Let $(e_i)_{1\leqslant i\leqslant n}$ be an arbitrary basis of E. For $i,\, j = 1, \dots, n$, let

$$g_{ij} = (e_i|e_j) \in K.$$

The matrix $(g_{ij})_{1\leqslant i,j\leqslant n}$ is the Gram matrix of the scalar product with respect to the basis $(e_i)_{1\leqslant i\leqslant n}$. Its entries are the coordinates of the 2-covariant tensor

$$g = (e_i|e_j)e^i \otimes e^j \in E^* \otimes E^*$$

which corresponds to the scalar product (or to the map Φ) under the identification of Theorem 1.7:

$$M(E \times E, K) = E^* \otimes E^* \quad (= M(E, E^*)).$$

The tensor g is called the *metric tensor* on the euclidean space E.

For $i = 1, \dots, n$, the basis vector $e_i \in E$ is identified with $\Phi(e_i) \in E^*$. By equation (2) of section 1.1 of Chapter I, we have

$$\Phi(e_i) = e^j \langle \Phi(e_i), e_j \rangle = e^j (e_i|e_j) = g_{ij} e^j,$$

hence

$$e_i = g_{ij} e^j \qquad \text{under the identification } E = E^*. \tag{11}$$

Consequently, any p-covariant and q-contravariant tensor

$$t = t_{i_1\cdots i_p}^{j_1\cdots j_q} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q} \in T_p^q(E)$$

is identified to

$$\begin{aligned} t_{i_1\cdots i_p}^{j_1\cdots j_q} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes (g_{j_1k_1} e^{k_1}) \otimes \cdots \otimes (g_{j_qk_q} e^{k_q}) \\ = \big(t_{i_1\cdots i_p}^{j_1\cdots j_q} g_{j_1k_1} \cdots g_{j_qk_q}\big) e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e^{k_1} \otimes \cdots \otimes e^{k_q}. \end{aligned}$$

We may therefore define the $(p+q)$-covariant coordinates of t as follows:

$$t_{i_1\cdots i_p j_1\cdots j_q} = t_{i_1\cdots i_p}^{k_1\cdots k_q} g_{k_1j_1} \cdots g_{k_qj_q}.$$

To define the $(p+q)$-contravariant coordinates of t, we use the inverse of the Gram matrix $(g_{ij})_{1\leqslant i,j\leqslant n}$. (The Gram matrix is invertible because the scalar product is non-degenerate.) Let $(g^{ij})_{1\leqslant i,j\leqslant n}$ be the inverse matrix. From equation (11), we derive

$$e^i = g^{ij} e_j,$$

hence

$$\begin{aligned} t = t_{i_1\cdots i_p}^{j_1\cdots j_q} (g^{i_1k_1} e_{k_1}) \otimes \cdots \otimes (g^{i_pk_p} e_{k_p}) \otimes e_{j_1} \otimes \cdots \otimes e_{j_q} \\ = \big(t_{i_1\cdots i_p}^{j_1\cdots j_q} g^{i_1k_1} \cdots g^{i_pk_p}\big) e_{k_1} \otimes \cdots \otimes e_{k_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q}. \end{aligned}$$

Therefore, the $(p+q)$-contravariant coordinates of t are

$$t^{i_1\cdots i_p j_1\cdots j_q} = t_{k_1\cdots k_p}^{j_1\cdots j_q} g^{k_1i_1} \cdots g^{k_pi_p}.$$

3. Tensor product of linear maps

3.1. THEOREM. *For $i = 1, \dots, n$, let $A_i\colon E_i \to F_i$ be a linear map. There is a linear map*

$$A_1 \otimes \cdots \otimes A_n\colon E_1 \otimes \cdots \otimes E_n \to F_1 \otimes \cdots \otimes F_n$$

such that

$$(A_1 \otimes \cdots \otimes A_n)(x_1 \otimes \cdots \otimes x_n) = A_1(x_1) \otimes \cdots \otimes A_n(x_n) \tag{1}$$

for all $x_1 \in E_1, \dots, x_n \in E_n$.

PROOF. This follows from the universal property of tensor products, applied to the multilinear map

$$E_1 \times \cdots \times E_n \to F_1 \otimes \cdots \otimes F_n$$

which carries $(x_1, \dots, x_n) \in E_1 \times \cdots \times E_n$ to $A_1(x_1) \otimes \cdots \otimes A_n(x_n)$. □

Clearly, the linear map $A_1 \otimes \cdots \otimes A_n$ is uniquely determined by condition (1), since every element in $E_1 \otimes \cdots \otimes E_n$ is a sum of products $x_1 \otimes \cdots \otimes x_n$.

3.2. PROPOSITION. 1. *If $A_i\colon E_i \to F_i$ and $B_i\colon F_i \to G_i$ are linear maps, for $i = 1, \dots, n$, then*

$$(B_1 \circ A_1) \otimes \cdots \otimes (B_n \otimes A_n) = (B_1 \otimes \cdots \otimes B_n) \circ (A_1 \otimes \cdots \otimes A_n).$$

2. $\mathrm{Id}_{E_1} \otimes \cdots \otimes \mathrm{Id}_{E_n} = \mathrm{Id}_{E_1 \otimes \cdots \otimes E_n}$.
3. *If A_i is injective (resp. surjective) for all $i = 1, \dots, n$, then $A_1 \otimes \cdots \otimes A_n$ is injective (resp. surjective).*

PROOF. 1. For $x_1 \in E_1, \dots, x_n \in E_n$, we have

$$(B_1 \circ A_1) \otimes \cdots \otimes (B_n \circ A_n)(x_1 \otimes \cdots \otimes x_n) = B_1 \circ A_1(x_1) \otimes \cdots \otimes B_n \circ A_n(x_n)$$

and

$$\begin{aligned}(B_1 \otimes \cdots \otimes B_n) \circ (A_1 \otimes \cdots \otimes A_n)(x_1 \otimes \cdots \otimes x_n) &= (B_1 \otimes \cdots \otimes B_n)\big(A_1(x_1) \otimes \cdots \otimes A_n(x_n)\big) \\ &= B_1 \circ A_1(x_1) \otimes \cdots \otimes B_n \circ A_n(x_n).\end{aligned}$$

2. For $x_1 \in E_1, \dots, x_n \in E_n$ we have

$$\begin{aligned}(\mathrm{Id}_{E_1} \otimes \cdots \otimes \mathrm{Id}_{E_n})(x_1 \otimes \cdots \otimes x_n) &= \mathrm{Id}_{E_1}(x_1) \otimes \cdots \otimes \mathrm{Id}_{E_n}(x_n) \\ &= x_1 \otimes \cdots \otimes x_n \\ &= \mathrm{Id}_{E_1 \otimes \cdots \otimes E_n}(x_1 \otimes \cdots \otimes x_n).\end{aligned}$$

3. If A_i is injective (resp. surjective), then there is a linear map $B_i\colon F_i \to E_i$ such that $B_i \circ A_i = \mathrm{Id}_{E_i}$ (resp. $A_i \circ B_i = \mathrm{Id}_{F_i}$), hence

$$(B_1 \otimes \cdots \otimes B_n) \circ (A_1 \otimes \cdots \otimes A_n) = \mathrm{Id}_{E_1 \otimes \cdots \otimes E_n}$$
$$(\text{resp. } (A_1 \otimes \cdots \otimes A_n) \circ (B_1 \otimes \cdots \otimes B_n) = \mathrm{Id}_{F_1 \otimes \cdots \otimes F_n})$$

by the first two parts of the proposition. It follows that $A_1 \otimes \cdots \otimes A_n$ is injective (resp. surjective).

□