Course 8

3.3 The matrix of a list of vectors

In the previous section, we have seen a matrix as a list of row-vectors. Now we discuss a converse, namely we define the matrix associated to a list of vectors, with respect to a basis.

Definition 3.3.1 Let V be a vector space over K, $B = (v_1, \dots, v_n)$ a basis of V and $X = (u_1, \dots, u_m)$ a list of vectors in V. Let

$$\begin{cases} u_1 = a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ u_2 = a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \dots\dots\dots\dots\dots\dots \\ u_m = a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{cases}$$

be the unique writings of the vectors in X as linear combinations of vectors of the basis B, for some $a_{ij} \in K$. The *matrix of the list of vectors* X in the basis B is the matrix having as its rows the coordinates of the vectors in X in the basis B, that is,

$$[X]_B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Example 3.3.2 Consider the canonical basis $B = (e_1, e_2, e_3, e_4)$ and the list $X = (u_1, u_2, u_3)$ in the canonical real vector space $\mathbb{R}^4$, where

$$\begin{cases} u_1 & = (1, 2, 3, 4) \\ u_2 & = (5, 6, 7, 8) \\ u_3 & = (9, 10, 11, 12) \end{cases}.$$

Since the coordinates of a vector in the canonical basis are just its components, we get

$$[X]_B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}.$$

Now we give a theorem which allows one to determine the dimension of the subspace generated by a list of vectors.

Theorem 3.3.3 *Let V be a vector space over K, $B = (v_1, \dots, v_n)$ a basis of V and $X = (u_1, \dots, u_m)$ a list of vectors in V having the matrix A in the basis B. Then:*
(i) $\dim\langle X\rangle = \operatorname{rank}(A)$.
(ii) A basis of $\langle X\rangle$ is the list of non-zero row-vectors $(c_1, \dots, c_r)$ of an echelon form C equivalent to A.

Example 3.3.4 Let us determine the dimensions of the subspaces S, T, $S+T$ and $S\cap T$ of the canonical real vector space $\mathbb{R}^4$, where

$$S = \langle(-3, 5, -1, 1), (-1, 1, 0, 1), (1, 1, -1, -3)\rangle,$$
$$T = \langle(1, 0, 2, 0), (2, 1, -1, 2)\rangle.$$

One can easily show that the ranks of the matrices in the canonical basis corresponding to the vectors from S and from T respectively are both 2. Hence $\dim S = \dim T = 2$.

Furthermore, $S+T=\langle S\cup T\rangle$. We write the matrix of $S\cup T$ in the canonical basis and we have

$$\begin{pmatrix}-3&5&-1&1\\-1&1&0&1\\1&1&-1&-3\\1&0&2&0\\2&1&-1&2\end{pmatrix}\sim\begin{pmatrix}1&1&-1&-3\\-1&1&0&1\\-3&5&-1&1\\1&0&2&0\\2&1&-1&2\end{pmatrix}\sim\begin{pmatrix}1&1&-1&-3\\0&2&-1&-2\\0&8&-4&-8\\0&-1&3&3\\0&-1&1&8\end{pmatrix}\sim\begin{pmatrix}1&1&-1&-3\\0&-1&3&3\\0&2&-1&-2\\0&2&-1&-2\\0&-1&1&8\end{pmatrix}$$

$$\sim\begin{pmatrix}1&1&-1&-3\\0&-1&3&3\\0&0&5&4\\0&0&5&4\\0&0&-2&5\end{pmatrix}\sim\begin{pmatrix}1&1&-1&-3\\0&-1&3&3\\0&0&5&4\\0&0&0&\frac{33}{5}\\0&0&0&0\end{pmatrix}\sim\begin{pmatrix}1&1&-1&-3\\0&-1&3&3\\0&0&5&4\\0&0&0&33\\0&0&0&0\end{pmatrix}.$$

Then by Theorem 3.3.3, $\dim(S+T)=4$ and a basis of $S+T$ consists of the non-zero row-vectors from the echelon form, that is, $((1,1,-1,-3),(0,-1,3,3),(0,0,5,4),(0,0,0,33))$. Now by the Second Dimension Theorem, it follows that $\dim(S\cap T)=\dim S+\dim T-\dim(S+T)=2+2-4=0$.

Now we are going to define the matrix of a vector in a basis of a vector space. Even if one might expect to define it as a row-matrix, by considering a single vector list, it is more convenient to define it as a column-matrix for our purposes concerning linear maps in order to avoid formulas involving transposes.

Definition 3.3.5 Let V be a vector space over K, $v\in V$ and $B=(v_1,\dots,v_n)$ a basis of V. If $v=k_1v_1+\dots+k_nv_n$ $(k_1,\dots,k_n\in K)$ is the unique writing of v as a linear combination of the vectors of the basis B, then the *matrix of the vector* v in the basis B is

$$[v]_B=\begin{pmatrix}k_1\\\vdots\\k_n\end{pmatrix}.$$

Example 3.3.6 Consider the vector $v=(1,2,3)$ in the canonical real vector space $\mathbb{R}^3$, and let E be the canonical basis. Then $[v]_E=\begin{pmatrix}1\\2\\3\end{pmatrix}$.

3.4 The matrix of a linear map

Definition 3.4.1 Let $f:V\to V'$ be a K-linear map, $B=(v_1,\dots,v_n)$ a basis of V and $B'=(v'_1,\dots,v'_m)$ a basis of V'. Then we can uniquely write the vectors in $f(B)$ as linear combinations of the vectors of the basis B', say

$$\begin{cases}f(v_1)=a_{11}v'_1+a_{21}v'_2+\dots+a_{m1}v'_m\\f(v_2)=a_{12}v'_1+a_{22}v'_2+\dots+a_{m2}v'_m\\\dots\dots\dots\dots\dots\dots\\f(v_n)=a_{1n}v'_1+a_{2n}v'_2+\dots+a_{mn}v'_m\end{cases}$$

for some $a_{ij}\in K$.

Then the *matrix of the K-linear map* f in the bases B and B' is the matrix having as its columns the coordinates of the vectors of $f(B)$ in the basis B', that is,

$$[f]_{BB'}=\begin{pmatrix}a_{11}&a_{12}&\dots&a_{1n}\\a_{21}&a_{22}&\dots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\dots&a_{mn}\end{pmatrix}.$$

If $V=V'$ and $B=B'$, then we simply denote $[f]_B=[f]_{BB'}$.

Remark 3.4.2 We have to emphasize that we put the coordinates on the columns of the matrix of a linear map and not on the rows as we did for the matrix of a list of vectors.

Example 3.4.3 Consider the $\mathbb{R}$-linear map $f : \mathbb{R}^4 \to \mathbb{R}^3$ defined by

$$f(x,y,z,t) = (x+y+z, y+z+t, z+t+x)\,,\ \forall (x,y,z,t) \in \mathbb{R}^4\,.$$

Let $E = (e_1,e_2,e_3,e_4)$ and $E' = (e_1',e_2',e_3')$ be the canonical bases in $\mathbb{R}^4$ and $\mathbb{R}^3$ respectively. Since

$$\begin{cases} f(e_1) = f(1,0,0,0) = (1,0,1) = e_1' + e_3' \\ f(e_2) = f(0,1,0,0) = (1,1,0) = e_1' + e_2' \\ f(e_3) = f(0,0,1,0) = (1,1,1) = e_1' + e_2' + e_3' \\ f(e_4) = f(0,0,0,1) = (0,1,1) = e_2' + e_3' \end{cases}$$

it follows that the matrix of f in the bases E and E' is

$$[f]_{EE'} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Theorem 3.4.4 *Let $f : V \to V'$ be a K-linear map, $B = (v_1,\dots,v_n)$ a basis of V, $B' = (v_1',\dots,v_m')$ a basis of V' and $v \in V$. Then*

$$[f(v)]_{B'} = [f]_{BB'} \cdot [v]_B\,.$$

Proof. Let $[f]_{BB'} = (a_{ij}) \in M_{m,n}(K)$. Let $v = \sum_{j=1}^n k_j v_j$ and

$$f(v) = \sum_{i=1}^m k_i' v_i'$$

for some $k_i, k_i' \in K$. On the other hand, using the definition of the matrix of f in the bases B and B', we have

$$f(v) = f\left(\sum_{j=1}^n k_j v_j\right) = \sum_{j=1}^n k_j f(v_j) = \sum_{j=1}^n k_j \left(\sum_{i=1}^m a_{ij} v_i'\right) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} k_j\right) v_i'\,.$$

But the writing of $f(v)$ as a linear combination of the vectors of the basis B' is unique, hence we must have $k_i' = \sum_{j=1}^n a_{ij} k_j$ for every $i \in \{1,\dots,m\}$. Therefore, $[f(v)]_{B'} = [f]_{BB'} \cdot [v]_B$. □

Now we give a connection between the ranks of a linear map and of its matrix in a pair of bases.

Theorem 3.4.5 *Let $f : V \to V'$ be a K-linear map. Then*

$$\operatorname{rank}(f) = \operatorname{rank}([f]_{BB'})\,,$$

where B and B' are any bases of V and V' respectively.

Proof. Let $B = (v_1,\dots,v_n)$ and $[f]_{BB'} = A$. Using our results relating ranks and dimensions, we have

$$\begin{aligned} \operatorname{rank}(f) &= \dim(\operatorname{Im} f) = \dim f(V) = \dim f(\langle v_1,\dots,v_n\rangle) \\ &= \dim\langle f(v_1),\dots,f(v_n)\rangle = \operatorname{rank}(\ A^T) = \operatorname{rank}(A) = \operatorname{rank}([f]_{BB'})\,. \end{aligned}$$

Now take some other bases $B_1 = (u_1,\dots,u_n)$ of V and B_1' of V' and denote $[f]_{B_1B_1'} = A_1$. Then

$$\begin{aligned} \operatorname{rank}([f]_{B_1B_1'}) &= \operatorname{rank}(A_1) = \operatorname{rank}(\ A_1^T) = \dim\langle f(u_1),\dots,f(u_n)\rangle \\ &= \dim(\operatorname{Im} f) = \dim\langle f(v_1),\dots,f(v_n)\rangle = \operatorname{rank}([f]_{BB'})\,. \end{aligned}$$

This shows the result. □

Remark 3.4.6 Notice that the rank of a linear map does not depend on the pair of bases in which we write its matrix. Also notice that, considering matrices of a linear map in different pairs of bases, their ranks are the same. Some other connection between matrices of a linear map in different pairs of bases will be discussed in the next section.

Example 3.4.7 Consider the $\mathbb{R}$-linear map $f:\mathbb{R}^4\to\mathbb{R}^3$ defined by

$$f(x,y,z,t)=(x+y+z,y+z+t,z+t+x)\,,\ \forall (x,y,z,t)\in\mathbb{R}^4\,.$$

Let $E=(e_1,e_2,e_3,e_4)$ and $E'=(e'_1,e'_2,e'_3)$ be the canonical bases in $\mathbb{R}^4$ and $\mathbb{R}^3$ respectively. Using Example 3.4.3 it follows that

$$[f]_{EE'}=\begin{pmatrix}1&1&1&0\\0&1&1&1\\1&0&1&1\end{pmatrix}\sim\begin{pmatrix}1&1&1&0\\0&1&1&1\\0&-1&0&1\end{pmatrix}\sim\begin{pmatrix}1&1&1&0\\0&1&1&1\\0&0&1&2\end{pmatrix}.$$

Now by Theorem 3.4.5 it follows that $\operatorname{rank}(f)=\operatorname{rank}([f]_{EE'})=3$.

We end this section with a key result in Linear Algebra, connecting linear maps and matrices.

Theorem 3.4.8 *Let V, V' and V'' be vector spaces over K with $\dim V=n$, $\dim V'=m$ and $\dim V''=p$ and let $B=(v_1,\dots,v_n)$, $B'=(v'_1,\dots,v'_m)$ and $B''=(v''_1,\dots,v''_p)$ be bases of V, V' and V'' respectively. Then $\forall f,g\in\operatorname{Hom}_K(V,V')$, $\forall h\in\operatorname{Hom}_K(V',V'')$ and $\forall k\in K$, we have*

$$\begin{aligned}[f+g]_{BB'}&=[f]_{BB'}+[g]_{BB'}\,,\\ [kf]_{BB'}&=k\cdot[f]_{BB'}\,,\\ [h\circ f]_{BB''}&=[h]_{B'B''}\cdot[f]_{BB'}\,.\end{aligned}$$

Proof. Let $[f]_{BB'}=(a_{ij})\in M_{m,n}(K)$, $[g]_{BB'}=(b_{ij})\in M_{m,n}(K)$ and $[h]_{B'B''}=(c_{ki})\in M_{pm}(K)$. Then

$$f(v_j)=\sum_{i=1}^m a_{ij}v'_i\,,\quad g(v_j)=\sum_{i=1}^m b_{ij}v'_i\,,\quad h(v'_i)=\sum_{k=1}^p c_{ki}v''_k$$

$\forall j\in\{1,\dots,n\}$ and $\forall i\in\{1,\dots,m\}$.

Then $\forall k\in K$ and $\forall j\in\{1,\dots,n\}$ we have

$$\begin{aligned}(f+g)(v_j)&=f(v_j)+g(v_j)=\sum_{i=1}^m a_{ij}v'_i+\sum_{i=1}^m b_{ij}v'_i=\sum_{i=1}^m(a_{ij}+b_{ij})v'_i\,,\\ (kf)(v_j)&=kf(v_j)=k\cdot\left(\sum_{i=1}^m a_{ij}v'_i\right)=\sum_{i=1}^m(ka_{ij})v'_i\,,\end{aligned}$$

hence $[f+g]_{BB'}=[f]_{BB'}+[g]_{BB'}$ and $[kf]_{BB'}=k\cdot[f]_{BB'}$.

Finally, $\forall j\in\{1,\dots,n\}$ we have

$$\begin{aligned}(h\circ f)(v_j)&=h(f(v_j))=h\left(\sum_{i=1}^m a_{ij}v'_i\right)=\sum_{i=1}^m a_{ij}h(v'_i)\\ &=\sum_{i=1}^m a_{ij}\left(\sum_{k=1}^p c_{ki}v''_k\right)=\sum_{k=1}^p\sum_{i=1}^m(c_{ki}a_{ij})v''_k\,,\end{aligned}$$

hence $[h\circ f]_{BB''}=[h]_{B'B''}\cdot[f]_{BB'}$. □

Theorem 3.4.9 *Let V and V' be vector spaces over K with $\dim V=n$ and $\dim V'=m$, and let B and B' be bases of V and V' respectively. Then the map*

$$\varphi:\operatorname{Hom}_K(V,V')\to M_{m,n}(K),\quad \varphi(f)=[f]_{BB'}\,,\ \forall f\in\operatorname{Hom}_K(V,V')$$

is an isomorphism of vector spaces.

Proof. We have seen that $\mathrm{Hom}_K(V,V')$ is a vector space over K with respect to the following addition and scalar multiplication: $\forall f,g \in \mathrm{Hom}_K(V,V')$ and $\forall k \in K$, $f+g, k\cdot f \in \mathrm{Hom}_K(V,V')$, where $\forall x \in V$,

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

Also, $M_{m,n}(K)$ is a vector space over K. By Theorem 3.4.8 it follows that φ is a K-linear map.

Finally, let us prove that φ is bijective. Consider $B = (v_1,\dots,v_n)$ and $B' = (v'_1,\dots,v'_m)$. Let $f,g \in \mathrm{Hom}_K(V,V')$ be such that $\varphi(f) = \varphi(g)$. Then $[f]_{BB'} = [g]_{BB'} = (a_{ij}) \in M_{m,n}(K)$, hence

$$f(v_j) = a_{1j}v'_1 + a_{2j}v'_2 + \cdots + a_{mj}v'_m = g(v_j)\,,$$

$\forall j \in \{1,\dots,n\}$. We have seen that two K-linear maps are equal if and only if they have the same values at all vectors of a basis. Hence $f = g$, which shows that φ is injective. Now let $A = (a_{ij}) \in M_{m,n}(K)$, seen as a list of column-vectors $(a^1,\dots,a^n)$, where $a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$. Define a K-linear map $f : V \to V'$ on the basis of the domain by

$$f(v_j) = a_{1j}v'_1 + \cdots + a_{mj}v'_m\,,$$

$\forall j \in \{1,\dots,n\}$. Then $\varphi(f) = [f]_{BB'} = (a_{ij}) = A$. Thus, φ is surjective. $\square$

Remark 3.4.10 The extremely important isomorphism given in Theorem 3.4.9 allows us to work with matrices instead of linear maps, which is much simpler from a computational point of view. Under this isomorphism, the kernel and the image of a linear map $f : V \to V'$, where V and V' are vector spaces over K with $\dim(V) = n$ and $\dim(V') = m$, and bases B and B' respectively, correspond to the *null space* and to the *column space* of its associated matrix $A = [f]_{BB'} \in M_{m,n}(K)$ respectively. Thus, the *null space* of A consists of vectors $x \in K^n$ such that $Ax = 0$, while the *column space* of A consists of all linear combinations of the columns of A. A vector $b \in K^m$ belongs to the column space of A if and only if the system $Ax = b$ has a solution. By the First Dimension Theorem it follows that the sum of the dimensions of the null space and the column space of A equals n.

Theorem 3.4.11 *Let V be a vector space over K with $\dim V = n$, and let B be a basis of V. Then the map*

$$\varphi : \mathrm{End}_K(V) \to M_n(K), \quad \varphi(f) = [f]_B\,,\ \forall f \in \mathrm{End}_K(V)$$

is an isomorphism of vector spaces and of rings.

Proof. Note that $(\mathrm{End}_K(V),+,\circ)$ and $(M_n(K),+,\cdot)$ are rings. The required isomorphisms follow by Theorem 3.4.9. $\square$

Corollary 3.4.12 *Let $f \in \mathrm{End}_K(V)$. Then $f \in \mathrm{Aut}_K(V) \iff \det([f]_B) \neq 0$, where B is any basis of V.*

Proof. Let B a basis of V. By Theorem 3.4.11, $f \in \mathrm{Aut}_K(V) \iff f$ is invertible in the ring $(\mathrm{End}_K(V),+,\circ) \iff [f]_B$ is invertible in the ring $(M_n(K),+,\cdot) \iff \det([f]_B) \neq 0$. $\square$

Extra: Hill cipher

Let $n \in \mathbb{N}^*$ and consider the canonical vector space $V = \mathbb{Z}_2^n$ over $\mathbb{Z}_2$ with canonical basis E. The vectors of V may be identified with n-bit binary strings. Suppose that Alice needs to send an n-bit plaintext $p \in \mathbb{Z}_2^n$ to Bob.

Hill cipher:

1. (*Key establishment*) Alice and Bob randomly choose an invertible matrix $K \in M_n(\mathbb{Z}_2)$ as a key, and compute its inverse.
2. (*Encryption*) Alice computes the ciphertext c according to the formula $[c]_E^T = [p]_E^T \cdot K$.

3. (*Decryption*) Bob computes the plaintext p according to the formula $[p]_E^T = [c]_E^T \cdot K^{-1}$.

Remark 3.4.13 The Hill cipher, which is nowadays insecure, was the first application of linear algebra to cryptography.

Example 3.4.14 Alice wants to send the message $p = (1, 0, 1) \in \mathbb{Z}_2^3$ to Bob.
Alice and Bob agree on the matrix

$$K = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{Z}_2)$$

as a key, and compute its inverse

$$K^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{Z}_2).$$

Alice encrypts the message by computing the ciphertext c as:

$$[c]_E^T = [p]_E^T \cdot K = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}.$$

Bob decrypts the message by computing the plaintext p as:

$$[p]_E^T = [c]_E^T \cdot K^{-1} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}.$$

Extra: Image transformations

Suppose that we have a 2D-image that we want to rotate counterclockwise with θ degrees around the origin. By such a rotation, the point of coordinates $(1, 0)$ becomes the point of coordinates $(\cos\theta, \sin\theta)$, while the point of coordinates $(0, 1)$ becomes the point of coordinates $(-\sin\theta, \cos\theta)$.
We look for an $\mathbb{R}$-linear map $f : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying the following conditions:

$$f(1, 0) = (\cos\theta, \sin\theta),$$
$$f(0, 1) = (-\sin\theta, \cos\theta).$$

Recall that every linear map is determined by its values at the elements of a basis (the canonical basis in our case). Hence the matrix of the linear map f in the canonical basis E of the canonical real vector space $\mathbb{R}^2$ is:

$$[f]_E = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

For any point $v = (x, y) \in \mathbb{R}^2$ of a 2D-image, its corresponding point in the rotated image is computed as $f(v) = (x', y') \in \mathbb{R}^2$, where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = [f(v)]_E = [f]_E \cdot [v]_E = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

For instance, for a counterclockwise rotation of 90^0 around the origin one has the matrix:

$$[f]_E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Extra: Graphs and networks (see [Crivei])