Course 9: 29.11.2021

3.5 Change of bases

We would like now to establish a connection between the matrices of a vector or of a linear map in different bases of a vector space.

Definition 3.5.1 Let V be a vector space over K, and let $B = (v_1, \dots, v_n)$ and $B' = (v'_1, \dots, v'_n)$ be bases of V. Then we can uniquely write

$$\begin{cases} v'_1 = t_{11}v_1 + t_{21}v_2 + \dots + t_{n1}v_n \\ v'_2 = t_{12}v_1 + t_{22}v_2 + \dots + t_{n2}v_n \\ \dots\dots\dots\dots\dots\dots \\ v'_n = t_{1n}v_1 + t_{2n}v_2 + \dots + t_{nn}v_n \end{cases}$$

for some $t_{ij} \in K$. Then the matrix $(t_{ij}) \in M_n(K)$, having as its columns the coordinates of the vectors of the basis B' in the basis B, is called the *change matrix* (or *transition matrix*) *from the basis* B *to the basis* B' and is denoted by $T_{BB'}$.

Remark 3.5.2 B is referred to as the "old" basis, while B' is referred to as the "new" basis.

Theorem 3.5.3 *Let V be a vector space over K, and let $B = (v_1, \dots, v_n)$ and $B' = (v'_1, \dots, v'_n)$ be bases of V. Then the change matrix $T_{BB'}$ is invertible and its inverse is the change matrix $T_{B'B}$.*

Proof. Denote by $T = (t_{ij}) \in M_n(K)$ the change matrix from the basis B to the basis B'. Then $v'_j = \sum_{i=1}^n t_{ij}v_i$, $\forall j \in \{1, \dots, n\}$. Denote by $S = (s_{ij}) \in M_n(K)$ the change matrix from the basis B' to the basis B. Then $v_i = \sum_{k=1}^n s_{ki}v'_k$, $\forall i \in \{1, \dots, n\}$. It follows that

$$v'_j = \sum_{i=1}^n t_{ij} \left(\sum_{k=1}^n s_{ki} v'_k \right) = \sum_{k=1}^n \left(\sum_{i=1}^n s_{ki} t_{ij} \right) v'_k .$$

By the uniqueness of writing of each v'_j as linear combination of the vectors of the basis B', We have

$$\sum_{i=1}^n s_{ki}t_{ij} = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases},$$

that is, $S \cdot T = I_n$.

Similarly, one shows that $T \cdot S = I_n$. Hence T is invertible and its inverse is S. $\square$

Let us now see how one can use the change matrix from one basis to another in order to compute the coordinates of a vector in different bases or the matrix of a linear map in different bases.

Theorem 3.5.4 *Let V be a vector space over K, let $B = (v_1, \dots, v_n)$ and $B' = (v'_1, \dots, v'_n)$ be bases of V and let $v \in V$. Then*

$$[v]_B = T_{BB'} \cdot [v]_{B'} .$$

Proof. Consider the writings of the vector $v \in V$ in the two bases B and B', say $v = \sum_{i=1}^n k_i v_i$ and $v = \sum_{j=1}^n k'_j v'_j$ for some $k_i, k'_j \in K$. Since $T_{BB'} = (t_{ij}) \in M_n(K)$, we have $v'_j = \sum_{i=1}^n t_{ij}v_i$, $\forall j \in \{1, \dots, n\}$. It follows that

$$v = \sum_{j=1}^n k'_j \left(\sum_{i=1}^n t_{ij} v_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n t_{ij} k'_j \right) v_i .$$

By the uniqueness of writing of v as a linear combination of the vectors of the basis B, it follows that $k_i = \sum_{j=1}^n t_{ij}k'_j$, whence $[v]_B = T_{BB'} \cdot [v]_{B'}$. $\square$

Example 3.5.5 Consider the bases $E = (e_1, e_2, e_3)$ and $B = (v_1, v_2, v_3)$ of the canonical real vector space $\mathbb{R}^3$, where E is the canonical basis and $v_1 = (0, 1, 1)$, $v_2 = (1, 1, 2)$, $v_3 = (1, 1, 1)$. Let us determine the change matrices from E to B and viceversa. We have

$$\begin{cases} v_1 = \phantom{e_1 +{}} e_2 + e_3 \\ v_2 = e_1 + e_2 + 2e_3 \\ v_3 = e_1 + e_2 + e_3 \end{cases} \Longrightarrow \begin{cases} e_1 = -v_1 \phantom{{}-v_2} + v_3 \\ e_2 = v_1 - v_2 + v_3 \\ e_3 = \phantom{v_1 -{}} v_2 - v_3 \end{cases} .$$

Hence we get

$$T_{EB} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}, \quad T_{BE} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

We must have $T_{BE} = T_{EB}^{-1}$, so that we could have obtained T_{BE} by computing the inverse of T_{EB}.

Now consider the vector $u = (1, 2, 3)$. Clearly, its coordinates in the canonical basis E are 1, 2 and 3. By Theorem 3.5.4, it follows that

$$[u]_B = T_{BE} \cdot [u]_E = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Hence the coordinates of u in the basis B are 1, 1 and 0.

Next we give a theorem relating the matrices of a linear map in two different bases, whose proof will be omitted.

Theorem 3.5.6 *Let $f \in End_K(V)$, and let B and B' be bases of V. Then*

$$[f]_{B'} = T_{BB'}^{-1} \cdot [f]_B \cdot T_{BB'} .$$

Example 3.5.7 Consider the bases $E = (e_1, e_2, e_3)$ and $B = (v_1, v_2, v_3)$ of the canonical real vector space $\mathbb{R}^3$, where E is the canonical basis and $v_1 = (0, 1, 1)$, $v_2 = (1, 1, 2)$, $v_3 = (1, 1, 1)$. Also let $f \in End_{\mathbb{R}}(\mathbb{R}^3)$ be defined by

$$f(x, y, z) = (x + y, y - z, z + x), \ \forall (x, y, z) \in \mathbb{R}^3 .$$

Let us determine the matrix of f in the basis E and in the basis B. We have

$$\begin{cases} f(e_1) = (1, 0, 1) = e_1 + e_3 \\ f(e_2) = (1, 1, 0) = e_1 + e_2 \\ f(e_3) = (0, -1, 1) = -e_2 + e_3 \end{cases} \Longrightarrow [f]_E = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Using Theorem 3.5.6 and the change matrices T_{EB} and T_{BE}, that we have determined in Example 3.5.5, we have

$$[f]_B = T_{EB}^{-1} \cdot [f]_E \cdot T_{EB} = T_{BE} \cdot [f]_E \cdot T_{EB} =$$
$$= \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -3 & -2 \\ 1 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix}.$$

It is worth to be mentioned that we could have reached the same result using the definition of the matrix of a linear map and expressing the vectors $f(v_1)$, $f(v_2)$ and $f(v_3)$ as linear combinations of the vectors v_1, v_2 and v_3 of the basis B.

3.6 Eigenvectors and eigenvalues

The study of endomorphisms of vector spaces also makes use of vectors whose images are just scalar multiples of themselves. They are the subject of the present section.

Definition 3.6.1 Let V be a vector space over K and $f \in End_K(V)$. A non-zero vector $v \in V$ is called an *eigenvector of f* if there exists $\lambda \in K$ such that $f(v) = \lambda \cdot v$. Here λ is called an *eigenvalue of f*.

Remark 3.6.2 Clearly, each eigenvector has a unique corresponding eigenvalue. But different eigenvectors may have the same corresponding eigenvalue.

For $f \in End_K(V)$, denote

$$V(\lambda) = \{v \in V \mid f(v) = \lambda v\}\,,$$

that is, the set consisting of the zero vector and the eigenvectors of f with eigenvalue λ.

Theorem 3.6.3 *Let V be a vector space over K, $f \in End_K(V)$ and λ an eigenvalue of f. Then $V(\lambda)$ is a subspace of V.*

Proof. Clearly, $0 \in V(\lambda)$, hence $V(\lambda) \neq \emptyset$. Now let $k_1, k_2 \in K$ and $v_1, v_2 \in V(\lambda)$. Then we have $f(v_1) = \lambda v_1$ and $f(v_2) = \lambda v_2$. It follows that

$$f(k_1v_1 + k_2v_2) = k_1f(v_1) + k_2f(v_2) = k_1(\lambda v_1) + k_2(\lambda v_2) = (k_1\lambda)v_1 + (k_2\lambda)v_2 = \lambda(k_1v_1 + k_2v_2)\,.$$

Hence, $k_1v_1 + k_2v_2 \in V(\lambda)$ and consequently, $V(\lambda)$ is a subspace of V. □

Definition 3.6.4 Let V be a vector space over K, $f \in End_K(V)$ and λ an eigenvalue of f. Then $V(\lambda)$ is called the *eigenspace* (or the *characteristic subspace*) of λ with respect to f.

The next theorem offers the essence of the practical method to determine eigenvalues and eigenvectors.

Theorem 3.6.5 *Let V be a vector space over K, B a basis of V and $f \in End_K(V)$ with the matrix $[f]_B = A = (a_{ij}) \in M_n(K)$. Then $\lambda \in K$ is an eigenvalue of f if and only if*

$$\det(A - \lambda \cdot I_n) = 0 \tag{1}$$

Proof. The element $\lambda \in K$ is an eigenvalue of f if and only if there exists a non-zero $v \in V$ such that $f(v) = \lambda v$. Consider $[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Then it follows that

$$\begin{aligned}
&f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff (f - \lambda \cdot 1_V)(v) = 0 \iff [(f - \lambda \cdot 1_V)(v)]_B = [0]_B \\
&\iff [f - \lambda \cdot 1_V]_B \cdot [v]_B = [0]_B \iff ([f]_B - \lambda \cdot [1_V]_B) \cdot [v]_B = [0]_B \iff (A - \lambda \cdot I_n) \cdot [v]_B = [0]_B \\
&\iff \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \iff \begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \dots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn} - \lambda)x_n = 0. \end{cases}
\end{aligned}$$

Then λ is an eigenvalue of f if and only if the final system (S) of linear equations has a non-zero solution if and only if its determinant $\det(A - \lambda \cdot I_n)$ is zero. □

Definition 3.6.6 The equality (1) is called the *characteristic equation* and the system (S) is called the *characteristic system*. The determinant $\det(A - \lambda I_n)$ may be seen as a polynomial $p_A(\lambda)$ in λ and it is called the *characteristic polynomial of f* with respect to A (or the *characteristic polynomial* of A).

Now a question arises naturally: if we take another basis B' of V and use the matrix $[f]_{B'}$, do we get the same eigenvalues and eigenvectors of f? The answer is positive, as we may see in the following result.

Theorem 3.6.7 *Let V be a vector space over K, B and B' bases of V and $f \in End_K(V)$ with the matrices $[f]_B = A \in M_n(K)$ and $[f]_{B'} = A' \in M_n(K)$. Then $p_A(\lambda) = p_{A'}(\lambda)$.*

Proof. By Theorem 3.5.6, we have $[f]_{B'} = T_{BB'}^{-1} \cdot [f]_B \cdot T_{BB'}$. Denote $T = T_{BB'}$. Hence we have $A' = T^{-1} \cdot A \cdot T$. Then

$$\begin{aligned}
p_{A'}(\lambda) &= \det(A' - \lambda I_n) = \det(T^{-1}AT - \lambda I_n T^{-1}T) = \det(T^{-1}(A - \lambda I_n)T) \\
&= \det(T^{-1}) \cdot \det(A - \lambda I_n) \cdot \det(T) = \det(A - \lambda I_n) = p_A(\lambda)\,.
\end{aligned}$$

□

Remark 3.6.8 (1) Therefore, the eigenvalues and the eigenvectors *do not depend* on the basis chosen for writing the matrix of the endomorphism. Of course, the matrices might be different, but in the end we get the same characteristic polynomial. Consequently, we can say that the eigenvalues of an endomorphism (or simply, of a matrix) are just the roots in K of its unique characteristic polynomial.

(2) A non-zero vector $v \in K^n$ is an eigenvector of a matrix $A \in M_n(K)$ if there exists $\lambda \in K$ such that $A[v]_E = \lambda[v]_E$, where E is the canonical basis of the canonical vector space K^n over K. In this case, λ is an eigenvalue of A.

Example 3.6.9 Let $f \in End_{\mathbb{R}}(\mathbb{R}^3)$ be defined by

$$f(x, y, z) = (2x, y + 2z, -y + 4z)\,,\ \forall (x, y, z) \in \mathbb{R}^3\,.$$

We write its matrix in the simplest basis, namely in the canonical basis of $\mathbb{R}^3$. Then

$$[f] = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -1 & 4 \end{pmatrix}.$$

The characteristic polynomial is $p(\lambda) = -(\lambda - 2)^2(\lambda - 3)$, so the eigenvalues are $\lambda_1 = \lambda_2 = 2$ and $\lambda_3 = 3$.

Let us take first $\lambda_1 = \lambda_2 = 2$. An eigenvector (x_1, x_2, x_3) is a non-zero solution of the characteristic system

$$\begin{pmatrix} 2-\lambda_1 & 0 & 0 \\ 0 & 1-\lambda_1 & 2 \\ 0 & -1 & 4-\lambda_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow \begin{cases} -x_2 + 2x_3 = 0 \\ -x_2 + 2x_3 = 0 \end{cases}.$$

Then $x_2 = 2x_3$ and $x_1, x_3 \in \mathbb{R}$, whence

$$V(2) = \{(x_1, 2x_3, x_3) \mid x_1, x_3 \in \mathbb{R}\} = \langle (1, 0, 0), (0, 2, 1) \rangle.$$

Any non-zero vector in $V(2)$ is an eigenvector of f with the associated eigenvalue $\lambda_1 = \lambda_2 = 2$.

Consider now $\lambda_3 = 3$. The corresponding characteristic system is

$$\begin{pmatrix} 2-\lambda_3 & 0 & 0 \\ 0 & 1-\lambda_3 & 2 \\ 0 & -1 & 4-\lambda_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow \begin{cases} -x_1 = 0 \\ -2x_2 + 2x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}.$$

We get the solution $x_1 = 0$, $x_2 = x_3$ and $x_3 \in \mathbb{R}$. Then

$$V(3) = \{(0, x_3, x_3) \mid x_3 \in \mathbb{R}\} = \langle (0, 1, 1) \rangle.$$

Any non-zero vector in $V(3)$ is an eigenvector of f with the associated eigenvalue $\lambda_3 = 3$.

The following famous theorem, which we give without proof, involves the characteristic polynomial.

Theorem 3.6.10 (Hamilton-Cayley) *Every matrix $A \in M_n(K)$ is a root of its characteristic polynomial.*

Corollary 3.6.11 *Let $A \in M_2(K)$. Then:*

(i) The characteristic polynomial of A is $p_A(\lambda) = \lambda^2 - Tr(A)\lambda + det(A)$, where $Tr(A)$ denotes the trace *of A (that is, the sum of the elements of the principal diagonal of A).*

(ii) $A^2 - Tr(A) \cdot A + det(A) \cdot I_2 = 0_2$.

The theory of eigenvectors and eigenvalues of an endomorphism is important for finding a basis in which the matrix of the endomorphism has a "nicer" form, namely a diagonal one if possible. As a sample result in this sense, we give the following theorem, whose proof will be omitted.

Theorem 3.6.12 *Let V be a vector space over K with $\dim V = n$ and $f \in End_K(V)$. If f has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, then*

$$[f]_B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

where $B = (v_1, \dots, v_n)$ is the basis of the corresponding eigenvectors.

Extra: PageRank

PageRank is a number assigned by Google to each web page. Pages with higher rank come higher in search results. We will describe a simplified version.

- Consider pages $S_1, \dots, S_n$, with some links between them. A link from S_j to S_i is a vote by S_j that S_i is important.
- Links from important pages should count for more (because the probability of visiting S_i will clearly increase); links from pages with many links should count for less (because that will decrease the probability that we click the one that leads to S_i).
- We want rankings $r_1, \dots, r_n \geq 0$, normalized so that $\sum_{i=1}^{n} r_i = 1$.
- Say S_j links to N_j different pages, and assume $N_j > 0$. We use the rule: a link from S_j to S_i contributes $\frac{r_j}{N_j}$ to r_i.
- Thus, for every $i \in \{1, \dots, n\}$, the following consistency condition should be satisfied:

$$r_i = \sum_{j \in J_i} \frac{r_j}{N_j},$$

  where $J_i = \{j \in \{1, \dots, n\} \mid \text{page } S_j \text{ links to page } S_i\}$.
- Define the matrix $P = (p_{ij}) \in M_n(\mathbb{R})$ by

$$p_{ij} = \begin{cases} \frac{1}{N_j} & \text{if there is a link from } S_j \text{ to } S_i \\ 0 & \text{otherwise.} \end{cases}$$

- Hence, for every $i \in \{1, \dots, n\}$, the consistency condition becomes:

$$r_i = \sum_{j \in J_i} p_{ij} r_j.$$

- But this is equivalent to the matrix equation $Pr = r$, and so r is an eigenvector of the matrix P with eigenvalue 1.

Reference: N. Strickland, Linear Mathematics for Applications.
`https://neil-strickland.staff.shef.ac.uk/courses/MAS201/MAS201.pdf`.