Course 7: 12.11.2020

Chapter 3 MATRICES AND LINEAR SYSTEMS

In this chapter we present the essential connection between linear maps and matrices, developing a more computational apparatus. Using elementary operations, we will be able to give practical methods for computing the rank or the inverse of a matrix as well as for solving linear systems of equations.

Throughout the present chapter K will always denote a field.

3.1 Elementary operations

Definition 3.1.1 Let V be a vector space over K. Then an *elementary operation* is one of the functions $\varepsilon_{ij}, \varepsilon_{i\alpha}, \varepsilon_{ij\alpha} : V^n \to V^n$ defined for every $(v_1, \ldots, v_n) \in V^n$ by:

$$\varepsilon_{ij}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = (v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n), \quad (1)$$

$$\varepsilon_{i\alpha}(v_1, \ldots, v_i, \ldots, v_n) = (v_1, \ldots, \alpha v_i, \ldots, v_n), \quad \alpha \in K^*, \quad (2)$$

$$\varepsilon_{ij\alpha}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = (v_1, \ldots, v_i + \alpha v_j, \ldots, v_j, \ldots, v_n), \quad \alpha \in K. \quad (3)$$

Theorem 3.1.2 *Using the previous notation, we have $\varepsilon_{ij}, \varepsilon_{i\alpha}, \varepsilon_{ij\alpha} \in Aut_K(V^n)$.*

Proof. It is easy to show that V^n has a structure of a vector space over K, where the operations are defined by

$$(v_1, \ldots, v_n) + (v'_1, \ldots, v'_n) = (v_1 + v'_1, \ldots, v_n + v'_n),$$

$$k(v_1, \ldots, v_n) = (kv_1, \ldots, kv_n),$$

$\forall k \in K$ and $\forall (v_1, \ldots, v_n), (v'_1, \ldots, v'_n) \in V^n$. Also, it is easy to check that $\varepsilon_{ij}, \varepsilon_{i\alpha}, \varepsilon_{ij\alpha}$ are K-linear maps. They are also bijections, having the inverses

$$(\varepsilon_{ij})^{-1} = \varepsilon_{ji}, \quad (\varepsilon_{i\alpha})^{-1} = \varepsilon_{i\alpha^{-1}}, \quad (\varepsilon_{ij\alpha})^{-1} = \varepsilon_{ij(-\alpha)}.$$

Definition 3.1.3 Let V be a vector space over K. Then two lists X and X' of vectors in the vector space V^n over K are called *equivalent* if one of them can be obtained from the other by applying a finite number of elementary operations, that is, there exists a finite composition $\varphi : V^n \to V^n$ of elementary operations such that $\varphi(X) = X'$ or $\varphi(X') = X$.

Remark 3.1.4 (1) The composition $\varphi : V^n \to V^n$ of elementary operations is obviously bijective, hence by Theorem 3.1.2, if one of the lists can be obtained from the other by applying a finite number of elementary operations, then the other one can.

(2) The name "equivalent" is justified by the fact that the previously defined relation is an equivalence relation (reflexive, transitive and symmetric).

Theorem 3.1.5 *Let V be a vector space over K and let X and X' be equivalent lists of vectors in the vector space V^n over K. Then:*

(i) X is linearly independent in $V^n \iff X'$ is linearly independent in V^n;

(ii) X is a system of generators for $V^n \iff X'$ is a system of generators for V^n;

(iii) X is a basis of $V^n \iff X'$ is a basis of V^n.

Proof. Since the lists X and X' are equivalent, there exists a finite composition $\varphi : V^n \to V^n$ of elementary operations such that $\varphi(X) = X'$. Since by Theorem 3.1.2, each elementary operation is an isomorphism, it follows that φ is an isomorphism. But we know that isomorphisms preserve linearly independent lists, systems of generators, and bases. □

Remark 3.1.6 In what follows, let us apply the previously presented theory of elementary operations in the case of the vector space $M_{mn}(K)$ of $m \times n$-matrices over K. In order to do that, we will see a matrix $A = (a_{ij}) \in M_{mn}(K)$ as a list of vectors $(a^1, \ldots, a^n)$, each of them being a column $a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$.

Then we get the following well-known elementary operations for a matrix:

(1) Interchanging any two columns of the matrix.

(2) Multiplying a column of the matrix by a non-zero scalar.

(3) Multiplying a column of the matrix by a scalar and add the result to another column of the matrix.

Applying Definition 3.1.3 in this case, we say that two matrices are *equivalent* if one of them can be obtained from the other by a finite number of the previous elementary operations on columns.

Theorem 3.1.7 *The value of an elementary operation applied on a matrix $A = (a_{ij}) \in M_{mn}(K)$, seen as a list of column-vectors $(a^1, \ldots, a^n)$, is equal to A multiplied on the right hand side by the matrix obtained from the identity matrix I_n, also seen as a list of columns, by applying the same elementary operation.*

Proof. For simplicity of writing we are going to prove the theorem for the first two columns involved in the elementary operations. We have

$$\varepsilon_{12}(A) = \varepsilon_{12}(a^1, a^2, a^3, \ldots, a^n) = (a^2, a^1, a^3, \ldots, a^n) = \begin{pmatrix} a_{12} & a_{11} & a_{13} & \ldots & a_{1n} \\ a_{22} & a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m2} & a_{m1} & a_{m3} & \ldots & a_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} = A \cdot E_{12},$$

where E_{12} is the matrix obtained from the identity matrix I_n by interchanging the first two columns.

Furthermore, $\forall \alpha \in K^*$,

$$\varepsilon_{1\alpha}(A) = \varepsilon_{1\alpha}(a^1, a^2, \ldots, a^n) = (\alpha a^1, a^2, \ldots, a^n) = \begin{pmatrix} \alpha a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \alpha & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \vdots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = A \cdot E_{1\alpha},$$

where $E_{1\alpha}$ is the matrix obtained from the identity matrix I_n by multiplying the first column by α.

Finally, $\forall \alpha \in K$,

$$\varepsilon_{12\alpha}(A) = \varepsilon_{12\alpha}(a^1, a^2, a^3, \ldots, a^n) = (a^1 + \alpha a^2, a^2, a^3, \ldots, a^n) = \begin{pmatrix} a_{11} + \alpha a_{12} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} + \alpha a_{22} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} + \alpha a_{m2} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \alpha & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} = A \cdot E_{12\alpha},$$

where $E_{12\alpha}$ is the matrix obtained from the identity matrix I_n by multiplying the second column by α and adding it to the first column. □

Definition 3.1.8 The matrices E_{ij}, $E_{i\alpha}$, $E_{ij\alpha}$, obtained by applying elementary operations on the identity matrix I_n, are called *elementary matrices*.

Theorem 3.1.9 *The elementary matrices are invertible.*

Proof. We have $\det I_n = 1 \neq 0$. But the determinant remains non-zero by applying elementary operations on I_n. Hence the elementary matrices are invertible. □

Remark 3.1.10 (1) It is also possible to see a matrix $A = (a_{ij}) \in M_{mn}(K)$ as a list of vectors $(a_1, \ldots, a_m)$, each of them being a row $a_i = (a_{i1} \ldots a_{in})$. In this case, the elementary operations are made on rows and the value of an elementary operation applied on A is equal to A multiplied on the left hand side by the matrix obtained from the identity matrix I_m by applying the same elementary operation.

(2) By the reason of methods of computing the rank of a matrix and solving linear systems of equations, from now on we will apply the elementary operations on the rows of a matrix.

Now we would like to find for a given matrix a "better" equivalent matrix. This better form is described in the following definition.

Definition 3.1.11 We say that a matrix $A \in M_{mn}(K)$ is in *echelon form* with $r \geq 1$ non-zero rows if:
(1) the rows $1, \ldots, r$ are non-zero and the rows $r+1, \ldots, m$ are zero;
(2)

$$0 \leq N(1) < N(2) < \cdots < N(r),$$

where $N(i)$ denotes the number of zero elements from the beginning of the row i, $\forall i \in \{1, \ldots, r\}$.

An echelon form is called:
- *trapezoidal* if $N(i) = i - 1$, $\forall i \in \{1, \ldots, r\}$.
- *triangular* if it is trapezoidal and $r = n$.
- *diagonal* if it has non-zero elements only on the diagonal.

Theorem 3.1.12 *Every non-zero matrix is equivalent to a matrix in echelon form.*

Proof. As we have mentioned, we are going to use elementary operations on the rows of a matrix.

Let $A = (a_{ij}) \in M_{mn}(K)$. Look for a row i with the least $N(i)$, that is, the least number of zero elements from the beginning of the row. Then interchange it with the first row. Let $j_1 = N(1) + 1$, that is, a_{1j_1} is the first non-zero element of the first row. Then make zeros on the column j_1 below the element a_{1j_1} by applying elementary operations on rows. In order to do that, multiply the first row by $-a_{kj_1}a_{1j_1}^{-1}$ and add it to the row k, $\forall k \in \{2, \ldots, m\}$.

Now look for a row i with the least $N(i)$, where $i \in \{2, \ldots, n\}$. Then interchange it with the second row. Let $j_2 = N(2) + 1$, that is, a_{2j_2} is the first non-zero element of the second row. Then make zeros on the column j_2 below the element a_{1j_2} by applying elementary operations on rows.

Repeating this procedure, we get a matrix in echelon form. □

Example 3.1.13 Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 3 & 2 & -2 & 6 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Then by applying elementary operations only on rows, we have the following succession of equivalent matrices (we denote by $\sim$ their equivalence):

$$A \sim \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix}.$$

In the first step, we have multiplied the first row by -3 and then by 1 and add it to the second row and to the third row respectively. In the second step, we have multiplied the second row by 2 and add it to the third row. The last matrix is in (trapezoidal) echelon form and it is equivalent to A.

3.2 Applications of elementary operations

Let us now see how to use elementary operations to compute easier the rank of a matrix and the inverse of a square matrix. Recall that we are going to apply elementary operations on the rows of a matrix.

We have the following theorem, whose proof will be omitted.

Theorem 3.2.1 *Let $0 \neq A = (a_{ij}) \in M_{mn}(K)$, seen as a list of row-vectors $(a_1, \ldots, a_m)$ or as a list of column-vectors $(a^1, \ldots, a^n)$. Then:*

(i) $\operatorname{rank}(A) = \dim < a_1, \ldots, a_m >= \dim < a^1, \ldots, a^n >$.

(ii) $\operatorname{rank}(A) = \operatorname{rank}(C) = r$, *where C is a matrix in echelon form with r non-zero rows equivalent to A.*

Example 3.2.2 Consider the matrix

$$A = \begin{pmatrix} -3 & 5 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -3 \end{pmatrix}.$$

Then

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

hence $\operatorname{rank}(A) = 2$.

Theorem 3.2.3 *Let $A \in M_n(K)$ with $\det A \neq 0$. Then A is equivalent to the identity matrix I_n and the inverse matrix A^{-1} of A is obtained from the identity matrix I_n by applying the same elementary operations as one does to obtain I_n from A.*

Proof. Since $\det(A) \neq 0$, $A \in M_n(K)$ is invertible, hence we have $\operatorname{rank}(A) = n$. Then by applying elementary operations we get to a matrix C in echelon form, having n non-zero rows. The matrix C has all the elements below the principal diagonal zero. Then one can make zeros above the principal diagonal, by applying elementary operations on rows starting with the last row. Thus, A is equivalent to a matrix in diagonal form. But since its rank is n, all the elements on the diagonal are non-zero, hence we may multiply by their inverses to get I_n. Therefore, A is equivalent to I_n.

But by Theorem 3.1.7 this means that there exist some elementary matrices $E_1, \ldots, E_k$ such that $E_k \ldots E_1 A = I_n$. It follows that $A^{-1} = E_k \ldots E_1 I_n$, that is, A^{-1} is obtained from the identity matrix I_n by applying the same elementary operations as one does to obtain I_n from A. □

Example 3.2.4 Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Then $\det(A) = 1 \neq 0$, hence A is invertible. Let us determine its inverse with the above described method. For simplicity of writing, we will put in the same matrix both A and the identity matrix I_3 and we will apply in parallel elementary operations on rows. We have

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right).$$

We read the inverse matrix A^{-1} from the right hand half of the last matrix, hence we have

$$A^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

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

Example 3.2.6 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:
$$[p]_E^T = [c]_E^T \cdot K^{-1} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}.$$