

Course 7

Elementary operations



Prof. dr. Septimiu Crivei

Chapter 3. Matrices and Linear Systems

- 1 Elementary operations
- 2 Applications of elementary operations

Application: LU decomposition

We present a matrix decomposition which offers more efficient ways for computing the inverse or the determinant of a matrix or for solving square linear systems of equations.

$$A = LU = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \cdot \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix} \in M_n(K).$$

Throughout the present chapter K will always denote a field, and $m, n \in \mathbb{N}^*$.

Definition

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

$$\varepsilon_{ij}(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = (v_1, \dots, v_j, \dots, v_i, \dots, v_n),$$

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

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

Property of elementary operations

Theorem

Using the previous notation, we have $\varepsilon_{ij}, \varepsilon_{i\alpha}, \varepsilon_{ij\alpha} \in \text{Aut}_K(V^n)$.

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

$$\begin{aligned}(v_1, \dots, v_n) + (v'_1, \dots, v'_n) &= (v_1 + v'_1, \dots, v_n + v'_n), \\ k(v_1, \dots, v_n) &= (kv_1, \dots, kv_n),\end{aligned}$$

$\forall k \in K$ and $\forall (v_1, \dots, v_n), (v'_1, \dots, v'_n) \in V^n$. 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

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

which imply that they are automorphisms. □

Definition

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 $\varphi : V^n \rightarrow V^n$ of elementary operations such that

$$\varphi(X) = X' \quad \text{or} \quad \varphi(X') = X.$$

The composition $\varphi : V^n \rightarrow V^n$ of elementary operations is obviously bijective, hence if one of the lists can be obtained from the other by applying a finite number of elementary operations, then the other one also can.

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

Theorem

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 $\varphi : V^n \rightarrow V^n$ of elementary operations such that $\varphi(X) = X'$. Since 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. □

Elementary operations for matrices

In what follows, let us apply the previously presented theory of elementary operations in the case of the vector space $M_{m,n}(K)$ of $m \times n$ -matrices over K . We will see a matrix $A = (a_{ij}) \in M_{m,n}(K)$ as a list of vectors (a^1, \dots, a^n) , each of them being a column

$$a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Then we get the 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.*

Two matrices are *equivalent* if one of them can be obtained from the other by a finite number of elementary operations.

Theorem

The value of an elementary operation applied on a matrix $A = (a_{ij}) \in M_{m,n}(K)$, seen as a list of column-vectors (a^1, \dots, 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.

$$\begin{aligned}
 \varepsilon_{12}(A) &= \varepsilon_{12}(a^1, a^2, a^3, \dots, a^n) = (a^2, a^1, a^3, \dots, a^n) \\
 &= \begin{pmatrix} a_{12} & a_{11} & a_{13} & \dots & a_{1n} \\ a_{22} & a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m2} & a_{m1} & a_{m3} & \dots & a_{mn} \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = A \cdot E_{12},
 \end{aligned}$$

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

Elementary matrices III

$$\begin{aligned}\varepsilon_{1\alpha}(A) &= \varepsilon_{1\alpha}(a^1, a^2, \dots, a^n) = (\alpha a^1, a^2, \dots, a^n) \\&= \begin{pmatrix} \alpha a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \\&= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = A \cdot E_{1\alpha},\end{aligned}$$

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

Elementary matrices IV

$$\begin{aligned}\varepsilon_{12\alpha}(A) &= \varepsilon_{12\alpha}(a^1, a^2, a^3, \dots, a^n) = (a^1 + \alpha a^2, a^2, a^3, \dots, a^n) \\&= \begin{pmatrix} a_{11} + \alpha a_{12} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} + \alpha a_{22} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} + \alpha a_{m2} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \\&= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ \alpha & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \\&= A \cdot E_{12\alpha},\end{aligned}$$

Elementary matrices V

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

Definition

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

Theorem

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. □

(1) It is also possible to see a matrix $A = (a_{ij}) \in M_{m,n}(K)$ as a list of vectors (a_1, \dots, a_m) , each of them being a row

$$a_i = (a_{i1} \dots a_{in}).$$

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.

Echelon form of a matrix I

For a given matrix we look for a “better” equivalent matrix.

Definition

We say that a matrix $A \in M_{m,n}(K)$ is in (row) *echelon form* with $r \geq 1$ non-zero rows if:

(1) the rows $1, \dots, r$ are non-zero and the rows $r + 1, \dots, m$ are zero;

(2)

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

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

Theorem

Every non-zero matrix $A \in M_{m,n}(K)$ is equivalent to a matrix in echelon form.

Echelon form of a matrix II

Proof. Let $A = (a_{ij}) \in M_{m,n}(K)$. 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. Such an element is sometimes called a *pivot*. 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 , $\forall k \in \{2, \dots, m\}$. Now look for a row i with the least $N(i)$, where $i \in \{2, \dots, n\}$. 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_{2j_2} by applying elementary operations on rows. Repeating this, one gets an echelon form. □

Example

Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 3 & 2 & -2 & 6 \\ -1 & 1 & 1 & 0 \end{pmatrix} \in M_{3,4}(\mathbb{R}).$$

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

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

We have pointed out the pivots by putting them in a box. 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 .

Applications

Lemma

Let $A = (a_{ij}) \in M_{m,n}(K)$, seen as a list of row-vectors (a_1, \dots, a_m) . Then a sublist of (a_1, \dots, a_m) consisting of r vectors is linearly independent in K^n if and only if there exists a non-zero minor of order r of the matrix A .

Proof. For simplicity, let us consider the sublist $X = (a_1, \dots, a_r)$ of (a_1, \dots, a_m) . Then X is linearly independent if and only if

$$k_1 a_1 + \cdots + k_r a_r = 0 \implies k_1 = \cdots = k_r = 0 \quad (k_1, \dots, k_r \in K)$$

if and only if the next system has only the zero solution:

[illegible]

if and only if there is a non-zero minor of order r of A .

Theorem

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

- (i) $\text{rank}(A) = \dim \langle a_1, \dots, a_m \rangle = \dim \langle a^1, \dots, a^n \rangle$.
- (ii) $\text{rank}(A) = \text{rank}(C) = r$, where C is a matrix in echelon form with r non-zero rows equivalent to A .

Proof. (i) Denote $r = \dim \langle a_1, \dots, a_m \rangle$. Then the maximum number of linearly independent vectors in the list (a_1, \dots, a_m) is r . Then the maximum order of the non-zero minors of A is r .

Therefore, $\text{rank}(A) = r$.

Similarly, $\text{rank}(A) = \dim \langle a^1, \dots, a^n \rangle$.

(ii) Let us see the matrices A and C as lists of row-vectors (a_1, \dots, a_m) and $(c_1, \dots, c_r, c_{r+1}, \dots, c_m)$. Clearly, $c_{r+1} = \dots = c_m = 0$ and the list (c_1, \dots, c_r) is linearly

Computation of the rank II

independent. But the elementary operations preserve the rank of a matrix. It follows that

$$\begin{aligned}\operatorname{rank}(A) &= \dim \langle a_1, \dots, a_m \rangle \\ &= \dim \langle c_1, \dots, c_m \rangle \\ &= \dim \langle c_1, \dots, c_r \rangle = r.\end{aligned}$$

Clearly, $\operatorname{rank}(C) = r$. □

Example. By the above method, we obtain that the following matrix

$$A = \begin{pmatrix} -3 & 5 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -3 \end{pmatrix} \in M_{3,4}(\mathbb{R})$$

has rank 2 [...].

Theorem

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 $\text{rank}(A) = n$. 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 .

Computation of the inverse II

But this means that there exist some elementary matrices E_1, \dots, E_k such that $E_k \dots E_1 A = I_n$. It follows that $A^{-1} = E_k \dots E_1 I_n$, 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 . \square

Example. Consider the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \in M_3(\mathbb{R}).$$

Then $\det(A) = 1 \neq 0$, hence A is invertible.

By the above method we obtain its inverse as

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

Definition

A matrix $A \in M_n(K)$ has an *LU decomposition* if it may be written as $A = L \cdot U$ for some lower triangular matrix L (that is, a matrix all of whose elements above its principal diagonal are zero) and upper triangular matrix U (that is, a matrix all of whose elements under its principal diagonal are zero).

Theorem

If $A \in M_n(K)$ can be reduced to an echelon form without interchanging any rows, then A has an LU decomposition, not necessarily unique. More generally, for every $A \in M_n(K)$ there is a permutation matrix P (that is, a matrix obtained by repeatedly interchanging the rows and columns of an identity matrix) such that $P \cdot A$ has an LU decomposition.

Extra: LU decomposition II

If A has an LU decomposition, then $\det(A) = \det(L) \cdot \det(U)$. Moreover, if A is also invertible, then $A^{-1} = U^{-1} \cdot L^{-1}$. The determinants and the inverses of L and U are computed much easier than the determinant and the inverse of A .

We have already seen that the matrix

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \in M_3(\mathbb{R})$$

is invertible and $\det(A) = 1$.

Alternatively, note that there is the permutation matrix

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in M_3(\mathbb{R})$$

Extra: LU decomposition III

(corresponding to interchanging the first two rows) such that

$$P \cdot A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = L \cdot U$$

as a product of a lower triangular matrix L and an upper triangular matrix U .

Then $\det(A) = \det(P)^{-1} \cdot \det(L) \cdot \det(U) = 1$, and

$$\begin{aligned} A^{-1} &= U^{-1} \cdot L^{-1} \cdot P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}. \end{aligned}$$