A note on the factorisation of non-singular matrices

Afonso Carvalho
Instituto Superior Técnico
Universidade de Lisboa
e-mail: afonso.s.c.carvalho@tecnico.ulisboa.pt

Resumo: Mostra-se que toda a matriz invertível é um produto de matrizes elementares de apenas dois tipos, sendo que nenhuma delas é a matriz de permutação.

Abstract It is shown that any non-singular matrix is a product of only two types of elementary matrices none of which is a permutation matrix.

palavras-chave: Matriz elementar; matriz invertível; matriz de permutação.

keywords: Elementary matrix; non-singular matrix; permutation matrix.

1 The elementary matrices and the permutation matrix

Let $\mathbb{M}_{k\times n}(\mathbb{K})$ be the algebra of the $n \times n$ matrices over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Given a matrix A in $\mathbb{M}_{k\times n}(\mathbb{K})$, there are classically three types of *elementary (row) operations* performed on the rows of A, namely:

(i) replacing row L_i by αL_i, where α is a non-zero scalar (Type I);

(ii) replacing row L_i by $L_i + \alpha L_j$, where α is a scalar and $i \neq j$ (Type II);

(iii) exchanging two rows, i.e., exchanging row L_i with row L_j, with $i \neq j$ (Type III).

Here the scalars α lie in the same field as the entries in the matrix.

Alternatively, we can see these operations as matrix multiplications. In fact, these operations correspond to multiplying A on the left, respectively, by the following three types of $k \times k$ *elementary matrices*:

- Type I. $D_i(\alpha)$ (with $\alpha \neq 0$): the matrix that is obtained from the identity matrix multiplying row i by α;

- Type II. $E_{ij}(\alpha)$ (with $i \neq j$): the matrix that is obtained from identity matrix by adding to row i row j multiplied by $\alpha \in \mathbb{K}$;
- Type III. P_{ij} (with $i < j$): the matrix that is obtained from the identity matrix by exchanging rows i and j.

It is well known that A is a non-singular matrix if, and only if, A is the product of elementary matrices of Type I, II, or III. Equivalently rephrasing it, the set of matrices of Type I, II, or III generates the group of invertible matrices (cf. [1], Proposition 2.18, [2], (3.9.3)).

It is the purpose of this note to show that the use of Type III matrices can be avoided. More precisely,

Proposition 1. *Let A be a $n \times n$ real or complex matrix. Then A is non–singular if, and only if, A is a product of elementary matrices of Type I or II.*

Proof. Since A is non-singular if, and only if, A is a product of elementary matrices of Type I, II or III, it suffices to show that any elementary matrix of Type III is a product of elementary matrices of the remaining types.

Let P_{ij} be a $n \times n$ matrix of Type III, that is,

$$P_{ij} = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix},$$

where it is supposed that $i < j$, without loss of generality. Then

$$P_{ij} = D_j(-1)E_{ij}(1)E_{ji}(-1)E_{ij}(1).$$

In fact,

$$D_j(-1)E_{ij}(1)E_{ji}(-1)E_{ij}(1) = D_j(-1)E_{ij}(1)$$

$$\times \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

$$= D_j(-1) \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} = P_{ij},$$

as required. □

The next example shows how the method of this proof can be applied to exchange two rows i, j of a given matrix A. We shall suppose, without loss of generality, that $i < j$.

Example. Let A be a $k \times n$ matrix and let i and j be two given rows with $i < j$. Similarly to the proof above, these two rows can be exchanged using only Type I and II elementary operations yielding a new matrix A'. The permutation of rows i and j correspond to the following:

1. Adding to row i row j multiplied by 1;
2. Adding to row j row i multiplied by -1;
3. Adding to row i row j multiplied by 1;
4. Multiplying row j by $\alpha = -1$.

Starting with matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix},$$

the corresponding matrices at each step of the process are:

1. Adding to row i row j multiplied by 1,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + a_{j1} & a_{i2} + a_{j2} & \cdots & a_{in} + a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix};$$

2. Adding to row j row i multiplied by -1,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + a_{j1} & a_{i2} + a_{j2} & \cdots & a_{in} + a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} - (a_{i1} + a_{j1}) & a_{j2} - (a_{i2} + a_{j2}) & \cdots & a_{jn} - (a_{in} + a_{jn}) \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}+a_{j1} & a_{i2}+a_{j2} & \cdots & a_{in}+a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{i1} & -a_{i2} & \cdots & -a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix};$$

3. Adding to row i row j multiplied by 1,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}+a_{j1}-a_{i1} & a_{i2}+a_{j2}-a_{i2} & \cdots & a_{in}+a_{jn}-a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{i1} & -a_{i2} & \cdots & -a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{i1} & -a_{i2} & \cdots & -a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix}.$$

4. Multiplying a row j by $\alpha =$ -1.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} = A'.$$

Referências

[1] M. Artin, *Algebra*, Prentice-Hall, New Jersey, 1991.

[2] Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.