ON THE EXPLICIT FORMULA FOR GAUSS-JORDAN ELIMINATION

NAM VAN TRAN, JULIA JUSTINO AND IMME VAN DEN BERG

Abstract. The elements of the successive intermediate matrices of the Gauss-Jordan elimination procedure have the form of quotients of minors. Instead of the proof using identities of determinants of [5], a direct proof by induction is given.

1. Introduction

Gantmacher’s book [1] on linear algebra contains an explicit formula for all elements $a_{ij}^{(k)}$ of the intermediate matrix obtained after $k$ Gaussian operations applied to a matrix $A = [a_{ij}]$, below and to the right of the $k$th pivot. The formula is given in terms of quotients of minors, and follows when applying Gaussian elimination to the two minors, which happen to have common factors all wiping themselves out, except for $a_{ij}^{(k)}$. Up to changing indices, the formula holds also for Gauss-Jordan elimination, and then a similar formula, with alternating sign, holds at the upper part of the intermediate matrices. A proof is given by Y. Li in [5], using identities of determinants. To our opinion the notation used for minors in different parts of the matrix is somewhat confusing. We considered it worthwhile to present the result in the notation of [1], with a direct proof by induction. Indeed, up to some changes, it is possible to extend the method of simultaneous Gaussian elimination of minors to all elements of the $k$th intermediate matrix.

The status of the explicit formula for the Gauss-Jordan elimination procedure seems to be uncertain. As regards to the formula for the lower part of the intermediate matrices Gantmacher refers to [2]; [3] contains some historical observations and presents a proof using identities of determinants.

In [3] the explicit formula for Gaussian elimination was applied to error analysis. We came across the explicit formula for Gauss-Jordan elimination also in relation to error analysis [4]. Here we apply the formula to prove that principal minors satisfying a maximality property are non-zero. This has also some numerical relevance, for a consequence is that maximal pivots are automatically non-zero.

2. The explicit formula for Gauss-Jordan elimination

We start with some definitions and notations related to the Gauss-Jordan operations, where we use the common representation by matrix multiplications. It is convenient when the matrix is diagonally eliminable, i.e. all pivots lie on the principal diagonal, and we verify that this can be assumed without restriction of generality through a condition on minors, and that then the pivots are non-zero.

2020 Mathematics Subject Classification. 15A09.

Key words and phrases. Gauss-Jordan elimination, minors.
Indeed, Theorem 2.6 is the main theorem and gives explicit expressions for the intermediate matrices of the Gauss-Jordan procedure. Explicit formulas for the matrices representing the Gauss-Jordan operations follow directly and are given in Theorem 2.7.

Theorem 2.11 is a straightforward consequence of Theorem 2.6 and states, together with Proposition 2.10, that without restriction of generality we may impose a maximal condition on the principal minors, and then the pivots are also maximal, and nonzero indeed.

We will always consider \( m \times n \) matrices with \( m, n \in \mathbb{N}, m, n \geq 1 \). We denote by \( M_{m,n}(\mathbb{R}) \) the set of all \( m \times n \) matrices over the field \( \mathbb{R} \).

**Definition 2.1.** Let \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) be of rank \( r \geq 1 \). Assume that \( a_{11} \neq 0 \). We let \( G_1 = [g_{ij}^{(1)}]_{m \times m} \) be the matrix which corresponds to the multiplication of the entries of the first line of \( A \) by \( 1/a_{11} \), such that the first pivot of \( A^{(1)} \equiv G_1A = [a_{ij}^{(1)}]_{m \times n} \) becomes \( a_{11}^{(1)} = 1 \). This means that

\[
G_1 = [g_{ij}^{(1)}]_{m \times m} = \begin{bmatrix}
\frac{1}{a_{11}} & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}.
\]

Assume \( a_{22}^{(1)} \neq 0 \). Let \( G_2 \) be the matrix which corresponds to the creation of zero’s in the first column of \( A^{(1)} \), except for \( a_{11}^{(1)} \), and let \( A^{(2)} = G_2A^{(1)} = [a_{ij}^{(2)}]_{m \times n} \), i.e.

\[
G_2 = [g_{ij}^{(2)}]_{m \times m} = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
-a_{21} & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
-a_{m1} & 0 & \cdots & 1
\end{bmatrix},
\]

and

\[
A^{(2)} = [a_{ij}^{(2)}]_{m \times n} = \begin{bmatrix}
1 & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\
0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\
\vdots & \vdots & \ddots & \vdots \\
0 & a_{m2}^{(2)} & \cdots & a_{mn}^{(2)}
\end{bmatrix}.
\]

Assume that \( G_{2k} = [g^{(2k)}]_{m \times m} \) and \( A^{(2k)} = [a^{(2k)}]_{m \times n} \) are defined for \( k < r \), and \( a_{k+1k+1}^{(2k)} \neq 0 \). The matrix \( G_{2k+1} \) corresponds to the multiplication of row \( k + 1 \) of \( A^{(2k)} \) by \( 1/a_{k+1k+1}^{(2k)} \), leading to \( A^{(2k+1)} = G_{2k+1}A^{(2k)} \), and the matrix \( G_{2k+2} \) corresponds to transforming the entries of column \( k \) of \( A^{(2k+1)} \) into zero, except for the entry \( a_{k+1k+1}^{(2k+1)} = 1 \), resulting in \( A^{(2k+2)} = G_{2k+2}A^{(2k+1)} \). So we have \( G_{2k+1} = [g_{ij}^{(2k+1)}]_{m \times m} \), where

\[
g_{ij}^{(2k+1)} = \begin{cases}
1 & \text{if } i = j \neq k + 1 \\
0 & \text{if } i \neq j \\
1/a_{k+1k+1}^{(2k)} & \text{if } i = j = k + 1
\end{cases}
\]

(1)
and \( G_{2k+2} = \left[ g^{(2k+2)}_{ij} \right]_{m \times m} \), where

\[
g^{(2k+2)}_{ij} = \begin{cases} 
0 & \text{if } j \notin \{i, k+1\} \\
1 & \text{if } i = j \\
-a^{(2k+1)}_{ik+1} & \text{if } i \neq k+1, j = k+1
\end{cases}
\]

For \( 1 \leq q \leq 2r \) we call the matrix \( A^{(q)} \) the \( q^{th} \) Gauss-Jordan intermediate matrix, and the matrix \( G_q \) is called the \( q^{th} \) Gauss-Jordan operation matrix. We write \( G = G_{2r}G_{2r-1} \cdots G_1 \).

The product \( G_{2k+1}A^{(2k)} \) corresponds to the Gaussian operation of multiplying the \((k+1)^{th}\) row of the matrix \( A^{(2k)} \) by the non-zero scalar \( \frac{1}{a^{(2k)}_{k+1k+1}} \) and the product \( G_{2k}A^{(2k-1)} \) corresponds to the repeated Gauss-Jordan operation of adding a scalar multiple of a row to some other row of \( A^{(2k-1)} \).

**Definition 2.2.** Assume \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) has \( \text{rank} r \geq 1 \). Then \( A \) is called diagonally eliminable up to \( r \) if \( a^{(2k-2)}_{kk} \neq 0 \) for \( 1 \leq k \leq r \); if \( r = m \leq n \) we say that \( A \) is diagonally eliminable.

**Notation 2.3.** Let \( A \in M_{m,n}(\mathbb{R}) \). For each \( k \in \mathbb{N} \) such that \( 1 \leq k \leq \min\{m, n\} \), let \( 1 \leq i_1 < \cdots < i_k \leq m \) and \( 1 \leq j_1 < \cdots < j_k \leq n \).

1. We denote the \( k \times k \) submatrix of \( A \) consisting of the rows with indices \( \{i_1, \ldots, i_k\} \) and columns with indices \( \{j_1, \ldots, j_k\} \) by \( A^{i_1 \cdots i_k}_{j_1 \cdots j_k} \).
2. We denote the corresponding \( k \times k \) minor by \( m^{i_1 \cdots i_k}_{j_1 \cdots j_k} = \det(A^{i_1 \cdots i_k}_{j_1 \cdots j_k}) \).
3. For \( 1 \leq k \leq \min\{m, n\} \) we may denote the principal minor of order \( k \) by \( m_k = m_{1 \cdots k}^{1 \cdots k} \). We define formally \( m_0 = 1 \).

Let \( A \) be a matrix of \( \text{rank} r \geq 1 \). It is well-known and not difficult to see that up to changing rows and columns one may always assume that the principal minors up to \( r \) are all non-zero. Proposition 2.5 shows that this condition is equivalent to being diagonally eliminable, and gives also a formula for the pivots. The proposition is a consequence of the following lemma; its proof uses the idea found in [1], on how the value of certain elements of the intermediate matrices can be related to minors by simplifying determinants.

**Lemma 2.4.** Let \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) be of \( \text{rank} r > 1 \). Assume that \( a_{11}, a_{22}^{(2)}, \ldots, a_{kk}^{(2k-2)} \) are non-zero for \( 1 \leq k < r \). Then for \( 1 \leq k < r \) it holds that

\[
m_{k+1} = a_{11}^{(2)}a_{22}^{(2)} \cdots a_{kk}^{(2k-2)}a_{k+1k+1}^{(2k-1)}.
\]

As a consequence \( a_{11}, a_{22}^{(2)}, \ldots, a_{kk}^{(2k-2)}, a_{k+1k+1}^{(2k-1)} \neq 0 \) if and only if \( m_1, \ldots, m_k \neq 0 \).
Theorem 2.6 (Explicit expressions for the Gauss-Jordan intermediate matrices). Let $A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R})$ be of rank $r > 1$ and diagonally eliminable up to $r$. Let $k < r$. Then

$$A^{(2k)} = \begin{bmatrix}
1 & \cdots & 0 & a^{(2k)}_{1k+1} & \cdots & a^{(2k)}_{1n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & a^{(2k)}_{kk+1} & \cdots & a^{(2k)}_{kn} \\
0 & \cdots & 0 & a^{(2k)}_{k+1k+1} & \cdots & a^{(2k)}_{k+1n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & a^{(2k)}_{m(k+1)} & \cdots & a^{(2k)}_{m(n)}
\end{bmatrix},$$

where

$$m_{k+1} = a_{11} \cdots a_{kk}^{(2k-1)} a_{k+1k+1}^{(2k)} \det \begin{bmatrix}
1 & \cdots & 0 & a^{(2k-1)}_{1k+1} \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 1 & a^{(2k-1)}_{kk+1} \\
0 & \cdots & 0 & a^{(2k)}_{k+1k+1}
\end{bmatrix}.$$
where

\[
 a_{ij}^{(2k)} = \begin{cases} 
(-1)^{k+i} \frac{m_{1\ldots i-1i+1\ldots kj}}{m_k} & \text{if } 1 \leq i \leq k, k+1 \leq j \leq n \\
\frac{m_{1\ldots kj}}{m_k} & \text{if } k+1 \leq i \leq m, k+1 \leq j \leq n 
\end{cases}
\]

**Proof.** Firstly, let \( k+1 \leq i \leq m \) and \( k+1 \leq j \leq n \). Let

\[
 U_{i,j} = \begin{bmatrix} a_{11} & \cdots & a_{1k} & a_{1j} \\
 \vdots & \ddots & \vdots & \vdots \\
 a_{i1} & \cdots & a_{ik} & a_{kj} \\
 a_{i1} & \cdots & a_{ik} & a_{ij} \end{bmatrix}.
\]

Then \( \det(U_{i,j}) = m_{1\ldots kj}^{1\ldots kj} \). By applying the first \( 2k \) Gauss-Jordan operations to \( U_{i,j} \), we obtain

\[
\det(U_{i,j}) = a_{11} \cdots a_{kk}^{(2k-2)} \det \begin{bmatrix} 1 & \cdots & 0 & a_{ij}^{(2k-1)} \\
 \vdots & \ddots & \vdots & \vdots \\
 0 & \cdots & 1 & a_{kj}^{(2k-1)} \\
 0 & \cdots & 0 & a_{ij}^{(2k)} \end{bmatrix} = m_k a_{ij}^{(2k)}.
\]

Hence \( a_{ij}^{(2k)} = \frac{m_{1\ldots kj}}{m_k} a_{ij}^{(2k)} \).

Secondly, we let \( 1 \leq i < k+1 \) and \( k+1 \leq j \leq n \). Let

\[
 V_{i,j} = \begin{bmatrix} a_{11} & \cdots & a_{1i-1} & a_{1i+1} & \cdots & a_{1k} & a_{1j} \\
 \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 a_{i-11} & \cdots & a_{i-1i-1} & a_{i-1i+1} & \cdots & a_{i-1k} & a_{i-1j} \\
 a_{i1} & \cdots & a_{ii-1} & a_{ii+1} & \cdots & a_{ik} & a_{ij} \\
 a_{i+11} & \cdots & a_{i+1i-1} & a_{i+1i+1} & \cdots & a_{i+1k} & a_{i+1j} \\
 \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 a_{k1} & \cdots & a_{ki-1} & a_{ki+1} & \cdots & a_{kk} & a_{kj} \end{bmatrix}.
\]

Then

\[
\det(V_{i,j}) = m_{1\ldots i-1i+1\ldots kj}.
\]

Let \( V'_{i,j} \) be the matrix obtained by applying the first \( 2k \) Gauss-Jordan operations to \( V_{i,j} \). Then, using (3),

\[
\det(V'_{i,j}) = a_{11} \cdots a_{kk}^{(2k-2)} \det \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & a_{ij}^{(2k)} \\
 \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
 0 & \cdots & 1 & \cdots & 0 & a_{ij}^{(2k)} \\
 0 & \cdots & 0 & 1 & \cdots & 0 \\
 \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} = m_k \det(V'_{i,j}).
\]

Expanding \( \det(V'_{i,j}) \) along the \( i^{th} \) row, we derive that

\[
\det(V'_{i,j}) = (-1)^{i+k} a_{ij}^{(2k)}.
\]
Combining, we conclude that \( a_{ij}^{(2k)} = (-1)^{i+j} \frac{m_{1\ldots i}^{1\ldots k+1} m_{i+1\ldots j}^{1\ldots k}}{m_k}. \) \( \square \)

The next theorem gives explicit formulas for the matrices \( G_q \) associated to the Gauss-Jordan operations. At odd order \( q = 2k + 1 \) we have to divide the row \( k + 1 \) by \( a_{k+1\ldots k+1}^{(2k)} \) as given by (4), and at even order \( q = 2k + 2 \), in the column \( j = k + 1 \) we have to subtract by \( a_{k+1\ldots k}^{(2k)} \) as given by (5).

**Theorem 2.7 (Explicit expressions for the Gauss-Jordan operation matrices).** Let \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) be of rank \( r > 1 \) and diagonally eliminable up to \( r \). For \( k < r \) the Gauss-Jordan operation matrix of odd order \( G_{2k+1} = [g_{ij}^{(2k+1)}]_{m \times m} \) satisfies

\[
g_{ij}^{(2k+1)} = \begin{cases} 1 & \text{if } i = j \neq k + 1 \\ 0 & \text{if } i \neq j \\ \frac{m_k}{m_{k+1}} & \text{if } i = j = k + 1 \\ \end{cases}.
\]

and the Gauss-Jordan operation matrix of even order \( G_{2k+2} = [g_{ij}^{(2k+2)}]_{m \times m} \) satisfies

\[
g_{ij}^{(2k+2)} = \begin{cases} 0 & \text{if } j \notin \{i, k + 1\} \\ (-1)^{k+i+1} \frac{m_{1\ldots i}^{1\ldots k+1} m_{j+1\ldots k+1}^{1\ldots k}}{m_k} & \text{if } 1 \leq i \leq k, j = k + 1 \\ \frac{m_{1\ldots i}^{1\ldots k+1} m_{j+1\ldots k+1}^{1\ldots k}}{m_k} & \text{if } k + 1 < i \leq m, j = k + 1 \\ \end{cases}.
\]

**Proof.** The theorem follows from formulas (1), (2) and Theorem 2.6. \( \square \)

Applying Theorem 2.6 to a diagonally eliminable \( n \times n \) matrix we find at the end the identity matrix \( I_n \). As a result the product of the Gauss-Jordan operation matrices is equal to the inverse matrix.

**Corollary 2.8.** Let \( A = [a_{ij}]_{n \times n} \) be a diagonally eliminable matrix. Then \( GA = I_n \) and \( G = A^{-1} \).

**Proof.** By Theorem 2.7, the matrices \( G_q \) are well-defined for all \( 1 \leq q \leq 2n \). Then \( GA = A^{(2n)} = I_n \) by Theorem 2.6. Hence \( G = A^{-1} \). \( \square \)

Theorem 2.6 holds under the condition that the matrix is diagonally eliminable, which is equivalent to asking that the principal minors are non-zero. Alternatively we may ask that the absolute values of the principal minors \( m_{k+1} \) are maximal with respect to minors of the same size which share the first \( k \) rows and columns. It well-known that by appropriately changing rows and columns this may always be achieved, and then we speak of properly arranged matrices. We will use Theorem 2.6 to show that the principal minors of properly arranged matrices are non-zero, which implies that they are diagonally eliminable. From a numerical point-of-view we are better off, the pivots being maximal.
Definition 2.9. Assume \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) has rank \( r \geq 1 \). Then \( A \) is called properly arranged, if
\[
|a_{ij}| \leq |a_{11}| \quad \text{for all } 1 \leq i \leq m, 1 \leq j \leq n, \\
\text{and for every } k \in \mathbb{N} \text{ such that } 1 \leq k < \min\{m, n\} \\
|m_{i \ldots k\ldots j}| \leq |m_{k+1}| \quad \text{for all } k + 1 \leq i \leq m, k + 1 \leq j \leq n. 
\]

Proposition 2.10. Assume \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) has rank \( r \geq 1 \). By changing rows and columns of \( A \), if necessary, we may obtain that \( A \) is properly arranged.

Theorem 2.11. Let \( A = [a_{ij}]_{m \times n} \in M_{m,n}(\mathbb{R}) \) be of rank \( r \geq 1 \) and properly arranged. Then \( A \) is diagonally eliminable up to \( r \). Moreover, \( |a_{k+1k+1}^{(2k)}| \geq |a_{ij}^{(2k)}| \) for all \( k \) with \( 0 \leq k < r \) and \( i, j \) with \( k + 1 \leq i \leq m \) and \( k + 1 \leq j \leq n \).

Proof. Suppose there exists \( k \) with \( 0 \leq k < r \) such that \( m_{k+1} = 0 \). We may also assume that \( k \) is the smallest index satisfying this condition. If \( m_1 = 0 \), also \( a_{11} = 0 \), and then by (7) all entries of \( A \) are zero. Hence \( \text{rank}(A) = 0 \), a contradiction. From now on we suppose that \( k \geq 1 \). By Theorem 2.7 the matrices \( G_1, \ldots, G_{2k} \) are well-defined. By Theorem 2.9 one has
\[
A^{(2k)} = \begin{bmatrix}
1 & \cdots & 0 & a_{1k+1}^{(2k)} & \cdots & a_{1n}^{(2k)} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 1 & a_{kk+1}^{(2k)} & \cdots & a_{kn}^{(2k)} \\
0 & \cdots & 0 & a_{k+1k+1}^{(2k)} & \cdots & a_{k+1n}^{(2k)} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & a_{mk+1}^{(2k)} & \cdots & a_{mn}^{(2k)} 
\end{bmatrix},
\]
where
\[
a_{ij}^{(2k)} = \begin{cases}
(-1)^{k+i} \frac{m_{1 \ldots i-i+1 \ldots k\ldots j}}{m_k} & \text{if } 1 \leq i \leq k, k + 1 \leq j \leq n \\
\frac{m_{1 \ldots k\ldots j}}{m_k} & \text{if } k + 1 \leq i \leq m, k + 1 \leq j \leq n
\end{cases}.
\]
Because \( A \) is properly arranged
\[
|a_{ij}^{(2k)}| = \frac{m_{1 \ldots k\ldots j}}{m_k} \leq \frac{m_{k+1}}{m_k} = |a_{k+1k+1}^{(2k)}|.
\]
for \( k + 1 \leq i \leq m, k + 1 \leq j \leq n \). Because \( m_{k+1} = 0 \), it holds that \( a_{k+1k+1}^{(2k)} = \frac{m_{k+1}}{m_k} = 0 \). Then \( a_{ij}^{(2k)} = 0 \) for \( k + 1 \leq i \leq m, k + 1 \leq j \leq n \). Hence \( \text{rank}(A) = k < r \). Then also \( \text{rank}(A) = \text{rank}(A^{(2k)}) = k < r \), a contradiction. Hence for \( 0 \leq k < r \) one has \( m_{k+1} \neq 0 \), and also (9).

Corollary 2.12. Let \( A = [a_{ij}]_{m \times n} \in M_{n}(\mathbb{R}) \) be non-singular and properly arranged. Then \( A \) is diagonally eliminable.

Proof. Because \( A = [a_{ij}]_{m \times n} \) is non-singular, it follows that \( \text{rank}(A) = n \). By Theorem 2.11 the matrix \( A \) is diagonally eliminable.
References

[1] F.R. Gantmacher. *The theory of matrices*, vols I, and II. Chelsea Publishing Co., New York, 1960.

[2] D.P. Grossman. On the problem of the numerical solution of systems of simultaneous linear algebraic equations. *Uspekhi Mat. Nauk*, 5(3):87–103, 1950.

[3] D.S. Parker. Explicit Formulas for the results of Gaussian Elimination. Available from: [http://web.cs.ucla.edu](http://web.cs.ucla.edu) 1995.

[4] N.V. Tran, J. Justino and I.P. van den Berg. Gauss-Jordan elimination applied to flexible systems of linear equations (in preparation).

[5] Y. Li. An explicit construction of Gauss-Jordan elimination matrix. Available from: [http://arxiv.org/pdf/0907.5038.pdf](http://arxiv.org/pdf/0907.5038.pdf) 2009.

namtv@hcmute.edu.vn; julia.justino@estsetubal.ips.pt; ivdb@uevora.pt