In the preceding section we arrived, after a lengthy example, at the idea that we might be able to single out a “preferred” transformed version of the augmented matrix of a linear system. In this section we investigate the type of matrix we have in mind, but we emphasize that this discussion does not have anything to do with linear systems per se, rather it is about a particular type of matrix!

\( A \in M_{m \times n} \) is in reduced row echelon form if there exist natural numbers 
\[ 1 \leq j_1 < j_2 < \cdots < j_r \leq n \]

such that

1. for each \( i = 1, \ldots, r \), column \( j_i \) of \( A \) is the \( i \)th standard basis vector \( e_i \) of \( \mathbb{R}^m \), and
2. every other column of \( A \) is in the span if the columns preceding it.

**Example.** Let’s verify that the \( 4 \times 7 \) matrix

\[
A = \begin{pmatrix}
1 & 2 & 0 & 3 & 5 & 0 & 1 \\
0 & 0 & 1 & 4 & 6 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

is in reduced row echelon form. Columns 1, 3, and 6 of \( A \) are

\[ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \]

in that order. For the other columns, we have:

- Column 2 is 2 times column 1.
- Column 4 is 3 times column 1 plus 4 times column 2.
- Column 5 is 5 times column 1 plus 6 times column 2.
- Column 7 is equal to column 1.

Thus every other column is in the span of the columns preceding it.

Note that in this case it happens that column 7 is also a standard basis vector, namely \( e_1 \), but we ignore this, because it is irrelevant to the properties of reduced row echelon form. Also, in the notation of the definition of reduced row echelon form, for this example we have \( r = 3 \), and

\[ j_1 = 1, \quad j_2 = 3, \quad \text{and} \quad j_3 = 6. \]
**Example.** Here are all the $2 \times 2$ matrices in reduced row echelon form:

$$
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}, \quad \begin{pmatrix}
1 & * \\
0 & 0
\end{pmatrix}, \quad \begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix}, \quad \begin{pmatrix}
0 & 0 \\
0 & 0
\end{pmatrix},
$$

where "*" means the entry can be any real number.

How do we recognize that a matrix is in reduced row echelon form? As we examine the columns from left to right, there might be a string of zero columns at the beginning. If the matrix is nonzero, the first nonzero column (from left to right) will be the first standard basis vector $e_1$. Continuing to examine columns from left to right, there might be a string of columns which are multiples of $e_1$. If there exists a column which is not a multiple of $e_1$, the first such column will be the second standard basis vector $e_2$. Continuing to the right, there might be a string of columns which are in the span of $\{e_1, e_2\}$, but if there exists a column which is not in this span then the first such column will be $e_3$. This pattern will continue until the end of the columns is reached.

We will not prove the following theorem, because we do not really need it for our work, and it would require a rather fussy and lengthy argument. The main idea of the proof uses the theory of matrix representations of linear functions. However, it is interesting enough to mention:

**Theorem.** For all $A \in M_{m \times n}$ there exists $D \in GL_m$ such that $DA$ is in reduced row echelon form. Moreover, this reduced row echelon form is uniquely determined by $A$.

The uniqueness of the reduced row echelon form allows us to speak of "the" reduced row echelon form, and also to attach certain aspects of the reduced row echelon form to the matrix $A$. For example, in the definition of reduced row echelon form, we call $j_1, \ldots, j_r$ the *leading column numbers* of the reduced row echelon form (because the nonzero entries in these columns are the left-most nonzero entries in their rows). The number $r$ itself is something we already know:

**Theorem.** Let $A$ be in reduced row echelon form. In the notation of the definition of reduced row echelon form, the number $r$ is the rank of $A$.

**Proof.** The column space of $A$ contains the independent vectors $e_1, \ldots, e_r$, and every column is in the span of these, so we have

$$
\text{rank } A = \text{dim col } A = r.
$$

QED

Regarding the existence of the reduced row echelon form, we do not usually care about the invertible matrix $D$ per se (in fact, it is far away from being unique); our interest lies in the algorithm which allows us to find the reduced row echelon form itself. In fact, our algorithm will not explicitly produce a single invertible matrix $D$ for which $DA$ is in reduced
row echelon form, rather it will use a finite sequence of such matrices. We will develop this algorithm, called row reduction, or Gauss-Jordan elimination, in this section.

We will use particularly simple invertible matrices, which are in fact called “elementary matrices”. To describe these, it is convenient to first introduce “row operations” on $M_{m \times n}$, which will be functions $R : M_{m \times n} \to M_{m \times n}$ defined in terms of rows. In the following definition we’ll write a matrix $A \in M_{m \times n}$ in terms of its rows:

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

where $a_1, \ldots, a_m \in M_{1 \times n}$.

**Elementary row operation of type 1:** Let $i \in \{1, \ldots, m\}$ and $c \in \mathbb{R}\{0\}$, and define $R$ by

$$R \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} ca_i \\ \vdots \\ a_m \end{pmatrix},$$

i.e., $R$ multiplies row $i$ by $c$.

**Elementary row operation of type 2:** Let $i, j \in \{1, \ldots, m\}$ with $i \neq j$, let $c \in \mathbb{R}$, and define $R$ by

$$R \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_i + ca_j \\ \vdots \\ a_m \end{pmatrix},$$

i.e., $R$ replaces row $i$ by the sum of row $i$ and $c$ times row $j$.

**Elementary row operation of type 3:** Let $i, j \in \{1, \ldots, m\}$, with $i < j$, and define $R$ by

$$R \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_m \end{pmatrix},$$

i.e., $R$ switches rows $i$ and $j$. \[\]
Now we can define the “elementary matrices”: for \( t = 1, 2, 3 \), if \( R \) is an elementary row operation of type \( t \), then \( R(I) \) is an \textit{elementary matrix of type} \( t \).

**Example.**
1. If \( R \) multiplies row 2 by 8, then
   \[
   R(I_5) = \begin{pmatrix}
   1 & 0 & 0 & 0 & 0 \\
   0 & 8 & 0 & 0 & 0 \\
   0 & 0 & 1 & 0 & 0 \\
   0 & 0 & 0 & 1 & 0 \\
   0 & 0 & 0 & 0 & 1
   \end{pmatrix}
   \]
   is an elementary matrix of type 1.
2. If \( R \) replaces row 3 by the sum of row 3 and 4 times row 1, then
   \[
   R(I_5) = \begin{pmatrix}
   1 & 0 & 0 & 0 & 0 \\
   0 & 1 & 0 & 0 & 0 \\
   4 & 0 & 1 & 0 & 0 \\
   0 & 0 & 0 & 1 & 0 \\
   0 & 0 & 0 & 0 & 1
   \end{pmatrix}
   \]
   is an elementary matrix of type 2.
3. If \( R \) switches rows 2 and 4, then
   \[
   R(I_5) = \begin{pmatrix}
   1 & 0 & 0 & 0 & 0 \\
   0 & 0 & 0 & 1 & 0 \\
   0 & 1 & 0 & 0 & 0 \\
   0 & 0 & 0 & 0 & 1
   \end{pmatrix}
   \]
   is an elementary matrix of type 3.

**Proposition.** Let \( R \) be an elementary row operation, and let \( E \) be the corresponding elementary matrix. Then for all \( A \in M_{m \times n} \) we have
\[
R(A) = EA.
\]

**Proof.** In each case this follows easily from the properties of matrix multiplication. \( \text{QED} \)

**Corollary.** Every elementary matrix is invertible. In fact, if \( E \) is an elementary matrix of type \( t \), then \( E^{-1} \) is also an elementary matrix of type \( t \).

**Proof.** The corresponding properties of elementary row operations are obvious:

1. If \( R \) multiplies row \( i \) by \( c \), then \( R^{-1} \) multiplies row \( i \) by \( c^{-1} \);
2. If \( R \) replaces row \( i \) by the sum of row \( i \) and \( c \) times row \( j \), then \( R^{-1} \) replaces row \( i \) by the sum of row \( i \) and \( -c \) times row \( j \);
3. If \( R \) switches rows \( i \) and \( j \), then \( R^{-1} = R \).
The result now follows from the correspondence between elementary matrices and elementary row operations, because in each case, if $F$ is the elementary matrix corresponding to the inverse operation $R^{-1}$, we have

$$FE = R^{-1}(R(I)) = I,$$

so $E$ is invertible and $F = E^{-1}$. QED

It should be fairly clear (although, as we mentioned above, we will not prove it) that every $A \in M_{m \times n}$ can be transformed into reduced row echelon form by a finite sequence of elementary row operations. If the associated elementary matrices are $E_1, \ldots, E_k$, then the matrix

$$D := E_k \ldots E_1$$

is invertible, and $DA$ is the reduced row echelon form of $A$. However, we reiterate that we usually won’t care about the matrix $D$ itself.

The following rather lengthy example illustrates the row reduction algorithm. The strategy is to proceed by columns from left to right: begin by looking for the first (from left to right) nonzero column. Use elementary row operations to transform this column into $e_1$. Continuing to the right, look for the first column which is not a scalar multiple of $e_1$. Use elementary row operations to transform this column into $e_2$. Continuing to the right, look for the first column which is not in the span of $\{e_1, e_2\}$. Use elementary row operations to transform this column into $e_3$. Continue this process until all the columns have been processed.

**Example.** Let’s find the reduced row echelon form of

$$A = \begin{pmatrix}
0 & 0 & -1 & 2 & 0 & 1 \\
-2 & -4 & -3 & 4 & -4 & -2 \\
2 & 4 & 4 & -6 & 4 & 2 \\
1 & 2 & 2 & -3 & 2 & 0
\end{pmatrix}.$$

Switch rows 1 and 4:

$$\begin{pmatrix}
1 & 2 & 2 & -3 & 2 & 0 \\
-2 & -4 & -3 & 4 & -4 & -2 \\
2 & 4 & 4 & -6 & 4 & 2 \\
0 & 0 & -1 & 2 & 0 & 1
\end{pmatrix}.$$

Add 2 times row 1 to row 2:

$$\begin{pmatrix}
1 & 2 & 2 & -3 & 2 & 0 \\
0 & 0 & 1 & -2 & 0 & -2 \\
2 & 4 & 4 & -6 & 4 & 2 \\
0 & 0 & -1 & 2 & 0 & 1
\end{pmatrix}.$$
Add $-2$ times row 1 to row 3:
\[
\begin{pmatrix}
1 & 2 & 2 & -3 & 2 & 0 \\
0 & 0 & 1 & -2 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & -1 & 2 & 0 & 1
\end{pmatrix}.
\]

Add $-2$ times row 2 to row 1:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 4 \\
0 & 0 & 1 & -2 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & -1 & 2 & 0 & 1
\end{pmatrix}.
\]

Add row 2 to row 4:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 4 \\
0 & 0 & 1 & -2 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & -1
\end{pmatrix}.
\]

Multiply row 3 by $1/2$:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 4 \\
0 & 0 & 1 & -2 & 0 & -2 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & -1
\end{pmatrix}.
\]

Add $-4$ times row 3 to row 1:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 0 \\
0 & 0 & 1 & -2 & 0 & -2 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & -1
\end{pmatrix}.
\]

Add 2 times row 3 to row 2:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 0 \\
0 & 0 & 1 & -2 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & -1
\end{pmatrix}.
\]

Add row 3 to row 4:
\[
\begin{pmatrix}
1 & 2 & 0 & 1 & 2 & 0 \\
0 & 0 & 1 & -2 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

Thus, the leading column numbers are 1, 2, and 6, and the rank of $A$ is 3.