

5.2 Matrix Equations

Linear Equations. An $m \times n$ system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

can be written as a matrix multiply equation $A\vec{X} = \vec{b}$. Let A be the matrix of coefficients a_{ij} , let \vec{X} be the column vector of variable names x_1, \dots, x_n and let \vec{b} be the column vector with components b_1, \dots, b_n . Then, because equal vectors are defined by equal components,

$$\begin{aligned} A\vec{X} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \\ &= \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \end{aligned}$$

Therefore, $A\vec{X} = \vec{b}$. Conversely, a matrix equation $A\vec{X} = \vec{b}$ corresponds to a set of linear algebraic equations, because of reversible steps above and equality of vectors.

A system of linear equations can be represented by its **variable list** x_1, x_2, \dots, x_n and its **augmented matrix**

$$(1) \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{pmatrix}.$$

The augmented matrix of system $A\vec{X} = \vec{b}$ is denoted $\langle A|\vec{b} \rangle$. Given an augmented matrix $C = \langle A|\vec{b} \rangle$ and a variable list x_1, \dots, x_n , the conversion back to a linear system of equations is made by expanding $C\vec{Y} = \mathbf{0}$, where \vec{Y} has components $x_1, \dots, x_n, -1$. This expansion involves just dot products, therefore rapid display is possible for the linear system

corresponding to an augmented matrix. Hand work often contains the following kind of exposition:

$$(2) \quad \begin{array}{cccc|c} & \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n & \\ \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array} \right) \end{array}$$

In (2), a dot product is applied to the first n elements of each row, using the variable list written above the columns. The symbolic answer is set equal to the rightmost column's entry, in order to recover the equations.

It is usual in homogeneous systems to *not write* the column of zeros, but to deal directly with A instead of $\langle A|\mathbf{0} \rangle$. This convention is justified by arguing that the rightmost column of zeros is unchanged by swap, multiply and combination rules, which are defined for matrix equations the next paragraph.

Elementary Row Operations. The three operations on equations which produce equivalent systems can be translated directly to row operations on the augmented matrix for the system. The rules produce **equivalent systems**, that is, the three rules neither create nor destroy solutions.

- Swap** Two rows can be interchanged.
- Multiply** A row can be multiplied by multiplier $m \neq 0$.
- Combination** A multiple of one row can be added to a different row.

Documentation of Row Operations. Throughout the display below, symbol **s** stands for **source**, symbol **t** for **target**, symbol **m** for **multiplier** and symbol **c** for **constant**.

- Swap** **swap(s,t)** \equiv swap rows **s** and **t**.
- Multiply** **mult(t,m)** \equiv multiply row **t** by $m \neq 0$.
- Combination** **combo(s,t,c)** \equiv add **c** times row **s** to row **t** \neq **s**.

The standard for documentation is to write the notation next to the target row, which is the row to be changed. For swap operations, the notation is written next to the first row that was swapped, and optionally next to both rows. The notation was developed from early `maple` notation for the corresponding operations `swaprow`, `mulrow` and `addrow`, appearing in the `maple` package `linalg`. In early versions of `maple`, an

additional argument A is used to reference the matrix for which the elementary row operation is to be applied. For instance, `addrow(A,1,3,-5)` selects matrix A as the target of the combination⁶ rule, which is documented in written work as `combo(1,3,-5)`. In written work on paper, symbol A is omitted, because A is the matrix appearing on the previous line of the sequence of steps.

Maple Remarks. Versions of `maple` use packages to perform toolkit operations. A short conversion table appears below.

On paper	Maple with(<code>linalg</code>)	Maple with(<code>LinearAlgebra</code>)
<code>swap(s,t)</code>	<code>swaprow(A,s,t)</code>	<code>RowOperation(A,[t,s])</code>
<code>mult(t,c)</code>	<code>mulrow(A,t,c)</code>	<code>RowOperation(A,t,c)</code>
<code>combo(s,t,c)</code>	<code>addrow(A,s,t,c)</code>	<code>RowOperation(A,[t,s],c)</code>

Conversion between packages can be controlled by the following function definitions, which causes the `maple` code to be the same regardless of which linear algebra package is used.⁷

Maple `linalg`

```

combo:=(a,s,t,c)->addrow(a,s,t,c);
swap:=(a,s,t)->swaprow(a,s,t);
mult:=(a,t,c)->mulrow(a,t,c);

```

Maple `LinearAlgebra`

```

combo:=(a,s,t,c)->RowOperation(a,[t,s],c);
swap:=(a,s,t)->RowOperation(a,[t,s]);
mult:=(a,t,c)->RowOperation(a,t,c);
macro(matrix=Matrix);

```

RREF Test. The reduced row-echelon form of a matrix, or `rref`, is defined by the following requirements.

1. Zero rows appear last. Each nonzero row has first element 1, called a **leading one**. The column in which the leading one appears, called a **pivot column**, has all other entries zero.
2. The pivot columns appear as consecutive initial columns of the identity matrix I . Trailing columns of I might be absent.

⁶Textbooks on linear algebra also call this operation **replace**, with re-ordered argument list. The logic is $\mathbf{t} = \mathbf{t} + c\mathbf{s}$, a replacement of row \mathbf{t} .

⁷The acronym ASTC is used for the signs of the trigonometric functions in quadrants I through IV. The argument lists for `combo`, `swap`, `mult` use the same order, ASTC, memorized in trigonometry as **All Students Take Calculus**.

The matrix (3) below is a typical **rref** which satisfies the preceding properties. Displayed secondly is the reduced echelon system (4) in the variables x_1, \dots, x_8 represented by the augmented matrix (3).

$$(3) \quad \begin{pmatrix} 1 & 2 & 0 & 3 & 4 & 0 & 5 & 0 & 6 \\ 0 & 0 & 1 & 7 & 8 & 0 & 9 & 0 & 10 \\ 0 & 0 & 0 & 0 & 0 & 1 & 11 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 13 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(4) \quad \begin{array}{rcccccc} x_1 & +2x_2 & & +3x_4 & +4x_5 & & +5x_7 & & =6 \\ & & x_3 & +7x_4 & +8x_5 & & +9x_7 & & =10 \\ & & & & & x_6 & +11x_7 & & =12 \\ & & & & & & & x_8 & =13 \end{array}$$

Matrix (3) is an **rref** and system (4) is a reduced echelon system. The initial 4 columns of the 7×7 identity matrix I appear in natural order in matrix (3); the trailing 3 columns of I are absent.

If the **rref** of the augmented matrix has a leading one in the last column, then the corresponding system of equations then has an equation “ $0 = 1$ ” displayed, which signals an **inconsistent** system. It must be emphasized that an **rref** always exists, even if the corresponding equations are inconsistent.

Elimination Method. The elimination algorithm for equations (see page 197) has an implementation for matrices. A row is marked **processed** if either (1) the row is all zeros, or else (2) the row contains a leading one and all other entries in that column are zero. Otherwise, the row is called **unprocessed**.

1. Move each unprocessed row of zeros to the last row using **swap** and mark it *processed*.
2. Identify an unprocessed nonzero row having the least number of leading zeros. Apply the **swap** rule to make this row the very first unprocessed row. Apply the **multiply** rule to insure a leading one. Apply the **combination** rule to change to zero all other entries in that column. The number of leading ones (lead variables) has been increased by one and the current column is a column of the identity matrix. Mark the row as *processed*, e.g., box the leading one: $\boxed{1}$.
3. Repeat steps 1–2, until all rows have been processed. Then all leading ones have been defined and the resulting matrix is in reduced row-echelon form.

Computer algebra systems and computer numerical laboratories automate computation of the **reduced row-echelon form** of a matrix A .

Literature calls the algorithm **Gauss-Jordan elimination**. Two examples:

$\mathbf{rref}(\mathbf{0}) = \mathbf{0}$ In **step 2**, all rows of the zero matrix $\mathbf{0}$ are zero. No changes are made to the zero matrix.

$\mathbf{rref}(I) = I$ In **step 2**, each row has a leading one. No changes are made to the identity matrix I .

Visual RREF Test. The habit to mark pivots with a box leads to a visual test for a RREF. An illustration:

$\left(\begin{array}{cccc c} \boxed{1} & 0 & 0 & 0 & 1/2 \\ 0 & \boxed{1} & 0 & 0 & 1/2 \\ 0 & 0 & \boxed{1} & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>Each boxed leading one $\boxed{1}$ appears in a column of the identity matrix. The boxes trail downward, ordered by columns 1, 2, 3 of the identity. No 4th pivot, therefore trailing identity column 4 is not used.</p>
---	--

Toolkit Sequence. A sequence of swap, multiply and combination steps applied to a system of equations is called a **toolkit sequence**. The viewpoint is that a camera is pointed over the shoulder of an expert who writes the mathematics, and after the completion of each toolkit step, a photo is taken. The ordered sequence of cropped photo frames is a **filmstrip** or toolkit sequence. The **First Frame** displays the original system and the **Last Frame** displays the reduced row echelon system.

The terminology applies to systems $A\vec{x} = \vec{b}$ represented by an augmented matrix $C = \langle A|\vec{b} \rangle$. The First Frame is C and the Last Frame is $\mathbf{rref}(C)$.

Documentation of toolkit sequence steps will use this textbook's notation, page 317:

$$\text{swap}(s,t), \text{mult}(t,m), \text{combo}(s,t,c),$$

each written next to the target row t . During the sequence, consecutive initial columns of the identity, called **pivot columns**, are created as steps toward the **rref**. Trailing columns of the identity might not appear. An illustration:

$\text{Frame 1: } \left(\begin{array}{cccc c} 1 & 2 & -1 & 0 & 1 \\ 1 & 4 & -1 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>Original augmented matrix.</p>
--	-----------------------------------

Frame 2:	$\left(\begin{array}{cccc c} 1 & 2 & -1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>combo(1,2,-1) Pivot column 1 completed.</p>
Frame 3:	$\left(\begin{array}{cccc c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>swap(2,3)</p>
Frame 4:	$\left(\begin{array}{cccc c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>combo(2,3,-2)</p>
Frame 5:	$\left(\begin{array}{cccc c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>Pivot column 2 completed by operation combo(2,1,-2). Back-substitution postpones this step.</p>
Frame 6:	$\left(\begin{array}{cccc c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>All leading ones found. mult(3,-1/2)</p>
Frame 7:	$\left(\begin{array}{cccc c} 1 & 0 & -3 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>combo(3,2,-1) Zero other column 3 entries. Next, finish pivot column 3.</p>
Last Frame:	$\left(\begin{array}{cccc c} 1 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$	<p>combo(3,1,3) rref found. Column 4 of the identity does not appear! There is no 4th pivot column.</p>

Avoiding fractions. A matrix A with only integer entries can often be put into reduced row-echelon form without introducing fractions. The **multiply** rule introduces fractions, so its use should be limited. It is advised that leading ones be introduced only when convenient, otherwise make the leading coefficient nonzero and positive. Divisions at the end of the computation will produce the **rref**.

Clever use of the **combination** rule can sometimes create a leading one without introducing fractions. Consider the two rows

$$\begin{array}{cccc} 25 & 0 & 1 & 0 & 5 \\ 7 & 0 & 2 & 0 & 2 \end{array}$$

The second row multiplied by -4 and added to the first row effectively replaces the 25 by -3 , whereupon adding the first row twice to the

second gives a leading one in the second row. The resulting rows are fraction-free.

$$\begin{array}{ccccc} -3 & 0 & -7 & 0 & -3 \\ 1 & 0 & -12 & 0 & -4 \end{array}$$

Rank and Nullity. What does it mean, if the first column of a **rref** is the zero vector? It means that the corresponding variable x_1 is a **free variable**. In fact, every column that does not contain a leading one corresponds to a free variable in the standard general solution of the system of equations. Symmetrically, each leading one identifies a pivot column and corresponds to a **leading variable**.

The number of leading ones is the **rank** of the matrix, denoted $\mathbf{rank}(A)$. The rank cannot exceed the row dimension nor the column dimension. The column count less the number of leading ones is the **nullity** of the matrix, denoted $\mathbf{nullity}(A)$. It equals the number of free variables.

Regardless of how matrix B arises, augmented or not, we have the relation

$$\text{variable count} = \mathbf{rank}(B) + \mathbf{nullity}(B).$$

If $B = \langle A|\vec{b} \rangle$ for $A\vec{X} = \vec{b}$, then the variable count n comes from \vec{X} and the column count of B is *one more*, or $n + 1$. Replacing the *variable count* by the *column count* can therefore lead to fundamental errors.

Back-substitution and efficiency. The algorithm implemented in the preceding toolkit sequence is *easy to learn*, because the actual work is organized by creating pivot columns, via swap, combination and multiply. The created pivot columns are initial columns of the identity. You are advised to learn the algorithm in this form, but *please change the algorithm* as you become more efficient at doing the steps. See the examples for illustrations.

Back Substitution. Computer implementations and also hand computation can be made more efficient by changing steps 2 and 3, then adding **step 4**, as outlined below.

1. Move each unprocessed row of zeros to the last row using **swap** and mark it *processed*.
- 2a. Identify an unprocessed nonzero row having the least number of leading zeros. Apply the **swap** rule to make this row the very first unprocessed row. Apply the **multiply** rule to insure a leading one. Apply the **combination** rule to change to zero all other entries in that column which are **below** the leading one.
- 3a. Repeat steps 1–2a, until all rows have been processed. The matrix has all leading ones identified, a triangular shape, but it is not generally a RREF.

4. **Back-Substitution.** Identify a row with a leading one. Apply the **combination** rule to change to zero all other entries in that column which are **above** the leading one. Repeat until all rows have been processed. The resulting matrix is a RREF.

Literature refers to **step 4** as **back-substitution**, a process which is exactly the original elimination algorithm applied to the system created by step 3a with *reversed variable list*.

Inverse Matrix. An efficient method to find the inverse B of a square matrix A , should it happen to exist, is to form the augmented matrix $C = \langle A|I \rangle$ and then read off B as the package of the last n columns of $\mathbf{rref}(C)$. This method is based upon the equivalence

$$\mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle \quad \text{if and only if} \quad AB = I.$$

The next theorem aids not only in establishing this equivalence but also in the practical matter of testing a candidate solution for the inverse matrix. The proof is delayed to page 332.

Theorem 8 (Inverse Test)

If A and B are square matrices such that $AB = I$, then also $BA = I$. Therefore, only one of the equalities $AB = I$ or $BA = I$ is required to check an inverse.

Theorem 9 (The Matrix Inverse and the rref)

Let A and B denote square matrices. Then

- (a) If $\mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle$, then $AB = BA = I$ and B is the inverse of A .
- (b) If $AB = BA = I$, then $\mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle$.
- (c) If $\mathbf{rref}(\langle A|I \rangle) = \langle C|B \rangle$, then $C = \mathbf{rref}(A)$. If $C \neq I$, then A is not invertible. If $C = I$, then B is the inverse of A .
- (d) Identity $\mathbf{rref}(A) = I$ holds if and only if A has an inverse.

The proof is delayed to page 333.

Finding Inverses. The method will be illustrated for the matrix

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

Define the first frame of the toolkit sequence to be $C_1 = \langle A|I \rangle$, then compute the toolkit sequence to $\mathbf{rref}(C_1)$ as follows.

$$\begin{array}{l}
C_1 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) & \text{First Frame} \\
C_2 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & -1 & 1 \end{array} \right) & \text{combo}(3,2,-1) \\
C_3 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array} \right) & \text{mult}(3,1/2) \\
C_4 = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array} \right) & \text{combo}(3,2,1) \\
C_5 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & 1 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 & -1/2 & 1/2 \end{array} \right) & \begin{array}{l} \text{combo}(3,1,-1) \\ \text{Last Frame} \end{array}
\end{array}$$

The theory implies that the inverse of A is the matrix in the right half of the last frame:

$$A^{-1} = \begin{pmatrix} 1 & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix}$$

Answer Check. Let B equal the matrix of the last display, claimed to be A^{-1} . The **Inverse Test**, Theorem 8, says that we do not need to check both $AB = I$ and $BA = I$. It is enough to check one of them. Details:

$$\begin{aligned}
AB &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1/2 - 1/2 & -1/2 + 1/2 \\ 0 & 1/2 + 1/2 & 1/2 - 1/2 \\ 0 & 1/2 - 1/2 & 1/2 + 1/2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

Elementary Matrices. The purpose of elementary matrices is to express toolkit operations of swap, combination and multiply in terms of matrix multiply.

Typically, toolkit operations produce a finite sequence of k linear algebraic equations, the first is the original system and the last is the reduced

row echelon form of the system. We are going to re-write a typical toolkit sequence as matrix multiply equations. Each step is obtained from the previous by left-multiplication by a square matrix E :

$$(5) \quad \begin{aligned} A\vec{X} &= \vec{b} && \text{Original system} \\ E_1 A\vec{X} &= E_1 \vec{b} && \text{After one toolkit step} \\ E_2 E_1 A\vec{X} &= E_2 E_1 \vec{b} && \text{After two toolkit steps} \\ E_3 E_2 E_1 A\vec{X} &= E_3 E_2 E_1 \vec{b} && \text{After three toolkit steps} \end{aligned}$$

Definition 5 (Elementary Matrix)

An elementary matrix E is created from the identity matrix by applying a single toolkit operation, that is, exactly one of the operations combination, multiply or swap.

Elementary Combination Matrix. Create square matrix E by applying `combo(s, t, c)` to the identity matrix. The result equals the identity matrix except for the zero in row t and column s which is replaced by c .

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Identity matrix.}$$

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{pmatrix} \quad \begin{array}{l} \text{Elementary combination matrix,} \\ \text{combo}(2, 3, c). \end{array}$$

Elementary Multiply Matrix. Create square matrix E by applying `mult(t, m)` to the identity matrix. The result equals the identity matrix except the one in row t is replaced by m .

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Identity matrix.}$$

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & m \end{pmatrix} \quad \begin{array}{l} \text{Elementary multiply matrix,} \\ \text{mult}(3, m). \end{array}$$

Elementary Swap Matrix. Create square matrix E by applying `swap(s, t)` to the identity matrix.

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Identity matrix.}$$

$$E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \begin{array}{l} \text{Elementary swap matrix,} \\ \text{swap}(1, 3). \end{array}$$

If square matrix E represents a combination, multiply or swap rule, then the definition of matrix multiply applied to matrix EB gives the same matrix as obtained by applying the toolkit rule directly to matrix B . The statement is justified by experiment. See the exercises and Theorem 10.

Typical 3×3 elementary matrices (C=Combination, M=Multiply, S=Swap) can be displayed in computer algebra system `maple` as follows.

On Paper	Maple with(linalg)	Maple with(LinearAlgebra)
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	<code>B:=diag(1,1,1);</code>	<code>B:=IdentityMatrix(3);</code>
<code>combo(2,3,c)</code>	<code>C:=addrow(B,2,3,c);</code>	<code>C:=RowOperation(B,[3,2],c);</code>
<code>mult(3,m)</code>	<code>M:=mulrow(B,3,m);</code>	<code>M:=RowOperation(B,3,m);</code>
<code>swap(1,3)</code>	<code>S:=swaprow(B,1,3);</code>	<code>S:=RowOperation(B,[3,1]);</code>

The reader is encouraged to write out several examples of elementary matrices by hand or machine. Such experiments lead to the following observations and theorems, proofs delayed to the section end.

Constructing an Elementary Matrix E .

- Combination** Change a zero in the identity matrix to symbol c .
- Multiply** Change a one in the identity matrix to symbol $m \neq 0$.
- Swap** Interchange two rows of the identity matrix.

Constructing E^{-1} from an Elementary Matrix E .

- Combination** Change multiplier c in E to $-c$.
- Multiply** Change diagonal multiplier $m \neq 0$ in E to $1/m$.
- Swap** The inverse of E is E itself.

Theorem 10 (Matrix Multiply by an Elementary Matrix)

Let B_1 be a given matrix of row dimension n . Select a toolkit operation combination, multiply or swap, then apply it to matrix B_1 to obtain matrix B_2 . Apply the identical toolkit operation to the $n \times n$ identity I to obtain elementary matrix E . Then

$$B_2 = EB_1.$$

Theorem 11 (Toolkit Sequence Identity)

If C and D are any two frames in a toolkit sequence, then corresponding toolkit operations are represented by square elementary matrices E_1, E_2, \dots, E_k and the two frames C, D satisfy the matrix multiply equation

$$D = E_k \cdots E_2 E_1 C.$$

Theorem 12 (The rref and Elementary Matrices)

Let A be a given matrix of row dimension n . Then there exist $n \times n$ elementary matrices E_1, E_2, \dots, E_k representing certain toolkit operations such that

$$\mathbf{rref}(A) = E_k \cdots E_2 E_1 A.$$

Illustration. Consider the following 6-frame sequence.

$$\begin{aligned} A_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 6 & 3 \end{pmatrix} && \text{Frame 1, original matrix.} \\ A_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 3 & 6 & 3 \end{pmatrix} && \text{Frame 2, combo(1,2,-2).} \\ A_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 3 & 6 & 3 \end{pmatrix} && \text{Frame 3, mult(2,-1/6).} \\ A_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -6 \end{pmatrix} && \text{Frame 4, combo(1,3,-3).} \\ A_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} && \text{Frame 5, combo(2,3,-6).} \\ A_6 &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} && \text{Frame 6, combo(2,1,-3). Found rref.} \end{aligned}$$

The corresponding 3×3 elementary matrices are

$$\begin{aligned} E_1 &= \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{Frame 2, combo(1,2,-2) applied to } I. \\ E_2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/6 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{Frame 3, mult(2,-1/6) applied to } I. \\ E_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} && \text{Frame 4, combo(1,3,-3) applied to } I. \\ E_4 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix} && \text{Frame 5, combo(2,3,-6) applied to } I. \end{aligned}$$

$$E_5 = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{Frame 6, combo}(2,1,-3) \text{ applied to } I.$$

Because each frame of the toolkit sequence has the succinct form EB , where E is an elementary matrix and B is the previous frame, the complete toolkit sequence can be written as follows.

$$\begin{array}{ll} A_2 = E_1 A_1 & \text{Frame 2, } E_1 \text{ equals } \text{combo}(1,2,-2) \text{ on } I. \\ A_3 = E_2 A_2 & \text{Frame 3, } E_2 \text{ equals } \text{mult}(2,-1/6) \text{ on } I. \\ A_4 = E_3 A_3 & \text{Frame 4, } E_3 \text{ equals } \text{combo}(1,3,-3) \text{ on } I. \\ A_5 = E_4 A_4 & \text{Frame 5, } E_4 \text{ equals } \text{combo}(2,3,-6) \text{ on } I. \\ A_6 = E_5 A_5 & \text{Frame 6, } E_5 \text{ equals } \text{combo}(2,1,-3) \text{ on } I. \\ A_6 = E_5 E_4 E_3 E_2 E_1 A_1 & \text{Summary, frames 1-6. This relation is} \\ & \mathbf{rref}(A_1) = E_5 E_4 E_3 E_2 E_1 A_1, \text{ which is the} \\ & \text{result of Theorem 12.} \end{array}$$

The summary is the equation

$$\mathbf{rref}(A_1) = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A_1$$

The inverse relationship $A_1 = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} E_5^{-1} \mathbf{rref}(A_1)$ is formed by the rules for constructing E^{-1} from elementary matrix E , page 326, the result being

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{rref}(A_1)$$

Examples and Methods

1 Example (Identify a Reduced Row–Echelon Form) Identify the matrices in reduced row–echelon form using the RREF Test page 318.

$$\begin{array}{ll} A = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & B = \begin{pmatrix} 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ C = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & D = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Solution:

Matrix A. There are two nonzero rows, each with a leading one. The pivot columns are 2, 4 and they are consecutive initial columns of the 4×4 identity matrix. Yes, it is a RREF.

Matrix B. Same as *A* but with pivot columns 1, 4. Yes, it is a RREF. Column 2 is not a pivot column. The example shows that a scan for columns of the identity is not enough: you must skip over non-pivot columns.

Matrix C. Immediately not a RREF, because the leading nonzero entry in row 1 is not a one.

Matrix D. Not a RREF. Swap row 3 twice to bring it to row 1. Then it will be a RREF. This example has pivots in columns 1, 4. But columns 1, 4 fail to be the first two initial columns of the 4×4 identity matrix.

Visual RREF Test. More experience is needed to use the visual test for RREF, but it is worth the effort. Details are very brief. The ability to use the visual test is learned by working examples that use the basic RREF test.

Leading ones are boxed:

$$A = \begin{pmatrix} 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} \boxed{1} & 1 & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Matrices *A*, *B* pass the visual test. Matrices *C*, *D* fail the test. Visually, we look for a boxed one starting on row 1. Boxes occupy consecutive rows, marching down and right, to make a triangular diagram.

2 Example (Reduced Row–Echelon Form) Find the reduced row–echelon form of the coefficient matrix *A* using the **elimination method**, page 319. Then solve the system.

$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= 0, \\ x_1 + 3x_2 - x_3 + 2x_4 &= 0, \\ x_2 + x_4 &= 0. \end{aligned}$$

Solution: The coefficient matrix *A* and its **rref** are given by (details below)

$$A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{rref}(A) = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Using variable list x_1, x_2, x_3, x_4 , the equivalent reduced echelon system is

$$\begin{aligned} x_1 - x_3 - x_4 &= 0, \\ x_2 + x_4 &= 0, \\ 0 &= 0. \end{aligned}$$

which has lead variables x_1, x_2 and free variables x_3, x_4 .

The *last frame algorithm* applies to write the standard general solution. This algorithm assigns invented symbols t_1, t_2 to the free variables, then back-substitution is applied to the lead variables. The solution to the system is

$$\begin{aligned} x_1 &= t_1 + t_2, \\ x_2 &= -t_2, \\ x_3 &= t_1, \\ x_4 &= t_2, \quad -\infty < t_1, t_2 < \infty. \end{aligned}$$

Details of the Elimination Method.

$$\begin{pmatrix} 1^* & 2 & -1 & 1 \\ 1 & 3 & -1 & 2 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

The coefficient matrix A . Leading one identified and marked as 1^* .

$$\begin{pmatrix} \boxed{1} & 2 & -1 & 1 \\ 0 & 1^* & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Apply the **combination** rule to zero the other entries in column 1. Mark the row processed. Identify the next leading one, marked 1^* .

$$\begin{pmatrix} \boxed{1} & 0 & -1 & -1 \\ 0 & \boxed{1} & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Apply the **combination** rule to zero the other entries in column 2. Mark the row processed. The matrix passes the **Visual RREF Test**.

3 Example (Back-Substitution) Display a toolkit sequence which uses numerical efficiency ideas of back substitution, page 322, in order to find the RREF of the matrix

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

Solution: The answer for the reduced row-echelon form of matrix A is

$$\mathbf{rref}(A) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Back-substitution details appear below.

Meaning of the computation. Finding a RREF is part of solving the homogeneous system $A\vec{X} = \vec{0}$. The Last Frame Algorithm is used to write the general solution. The algorithm requires a toolkit sequence applied to the augmented matrix $\langle A|\vec{0} \rangle$, ending in the Last Frame, which is the RREF with an added column of zeros.

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

The given matrix A . Identify row 1 for the first pivot.

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

combo(1,2,-1) applied to introduce zeros below the leading one in row 1.

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{combo}(2,3,-1) \text{ applied to introduce zeros below the leading one in row 2. The RREF has not yet been found. The matrix is triangular.}$$

$$\begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{Begin back-substitution: } \text{combo}(2,1,-2) \text{ applied to introduce zeros above the leading one in row 2.}$$

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{Continue back-substitution: } \text{combo}(3,2,-1) \text{ and } \text{combo}(3,1,1) \text{ applied to introduce zeros above the leading one in row 3.}$$

$$\begin{pmatrix} \boxed{1} & 0 & -1 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix} \quad \begin{array}{l} \text{RREF Visual Test passed.} \\ \text{This matrix is the answer.} \end{array}$$

4 Example (Answer Check a Matrix Inverse) Display the answer check details for the given matrix A and its proposed inverse B .

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

Solution:

Details. We apply the **Inverse Test**, Theorem 8, which requires one matrix multiply:

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} && \text{Expect } AB = I. \\ &= \begin{pmatrix} 1 & -3+2+1 & 1-2+1 & 1-1 \\ 0 & 1 & -1+1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1-1 & -1+1 & 1 \end{pmatrix} && \text{Multiply.} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} && \text{Simplify. Then } AB = I. \\ &&& \text{Because of Theorem 8, we don't check } BA = I. \end{aligned}$$

5 Example (Find the Inverse of a Matrix) Compute the inverse matrix of

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

Solution: The answer:

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

Details. Form the augmented matrix $C = \langle A|I \rangle$ and compute its reduced row-echelon form by toolkit steps.

$$\begin{pmatrix} 1 & 2 & -1 & 1 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & | & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{Augment } I \text{ onto } A.$$

$$\begin{pmatrix} 1 & 2 & -1 & 1 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & | & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{swap}(3,4).$$

$$\begin{pmatrix} 1 & 2 & -1 & 1 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & | & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{combo}(2,3,-1). \text{ Triangular matrix.}$$

$$\begin{pmatrix} 1 & 2 & -1 & 1 & | & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{Back-substitution: combo}(4,2,-1).$$

$$\begin{pmatrix} 1 & 2 & -1 & 0 & | & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{combo}(4,1,-1).$$

$$\begin{pmatrix} 1 & 0 & -1 & 0 & | & 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{combo}(2,1,-2).$$

$$\begin{pmatrix} 1 & 0 & -1 & 0 & | & 1 & -3 & 1 & 1 \\ 0 & 1 & 0 & 0 & | & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & | & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & | & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{combo}(3,1,1). \text{ Identity left, inverse right.}$$

Details and Proofs

Proof of Theorem 8:

Assume $AB = I$. Let $C = BA - I$. We intend to show $C = \mathbf{0}$, then $BA = C + I = I$, as claimed.

Compute $AC = ABA - A = AI - A = \mathbf{0}$. It follows that the columns \vec{y} of C are solutions of the homogeneous equation $A\vec{y} = \vec{0}$. To complete the proof, we

show that the only solution of $A\vec{y} = \vec{0}$ is $\vec{y} = \vec{0}$, because then C has all zero columns, which means C is the zero matrix.

First, $B\vec{u} = \vec{0}$ implies $\vec{u} = I\vec{u} = AB\vec{u} = A\vec{0} = \vec{0}$, hence B has an inverse, and then $B\vec{x} = \vec{y}$ has a unique solution $\vec{x} = B^{-1}\vec{y}$.

Suppose $A\vec{y} = \vec{0}$. Write $\vec{y} = B\vec{x}$. Then $\vec{x} = I\vec{x} = AB\vec{x} = A\vec{y} = \vec{0}$. This implies $\vec{y} = B\vec{x} = B\vec{0} = \vec{0}$. The proof is complete.

Proof of Theorem 9:

Details for (a). Let $C = \langle A|I \rangle$ and assume $\mathbf{rref}(C) = \langle I|B \rangle$. Solving the $n \times 2n$ system $C\vec{X} = \vec{0}$ is equivalent to solving the system $A\vec{Y} + I\vec{Z} = \vec{0}$ with n -vector unknowns \vec{Y} and \vec{Z} . This system has exactly the same solutions as $I\vec{Y} + B\vec{Z} = \vec{0}$, by the equation $\mathbf{rref}(C) = \langle I|B \rangle$. The latter is a reduced echelon system with lead variables equal to the components of \vec{Y} and free variables equal to the components of \vec{Z} . Multiplying by A gives $A\vec{Y} + AB\vec{Z} = \vec{0}$, hence $-\vec{Z} + AB\vec{Z} = \vec{0}$, or equivalently $AB\vec{Z} = \vec{Z}$ for every vector \vec{Z} (because its components are free variables). Letting \vec{Z} be a column of I shows that $AB = I$. Then $AB = BA = I$ by Theorem 8, and B is the inverse of A .

Details for (b). Assume $AB = I$. We prove the identity $\mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle$. Let the system $A\vec{Y} + I\vec{Z} = \vec{0}$ have a solution \vec{Y}, \vec{Z} . Multiply by B to obtain $BA\vec{Y} + B\vec{Z} = \vec{0}$. Use $BA = I$ to give $\vec{Y} + B\vec{Z} = \vec{0}$. The latter system therefore has \vec{Y}, \vec{Z} as a solution. Conversely, a solution \vec{Y}, \vec{Z} of $\vec{Y} + B\vec{Z} = \vec{0}$ is a solution of the system $A\vec{Y} + I\vec{Z} = \vec{0}$, because of multiplication by A . Therefore, $A\vec{Y} + I\vec{Z} = \vec{0}$ and $\vec{Y} + B\vec{Z} = \vec{0}$ are *equivalent* systems. The latter is in reduced row-echelon form, and therefore $\mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle$.

Details for (c). Toolkit steps that compute $\mathbf{rref}(\langle A|I \rangle)$ also compute $\mathbf{rref}(A)$. Readers learn this fact first by working examples. Elementary matrix formulas can make the proof more transparent: see the Miscellany exercises. We conclude that $\mathbf{rref}(\langle A|I \rangle) = \langle C|B \rangle$ implies $C = \mathbf{rref}(A)$.

We prove $C \neq I$ implies A is not invertible. Suppose not, then $C \neq I$ and A is invertible. Then (b) implies $\langle C|B \rangle = \mathbf{rref}(\langle A|I \rangle) = \langle I|B \rangle$. Comparing columns, this equation implies $C = I$, a contradiction.

To prove $C = I$ implies B is the inverse of A , apply (a).

Details for (d). Assume A is invertible. To be proved: $\mathbf{rref}(A) = I$. Part (b) says $F = \langle A|I \rangle$ satisfies $\mathbf{rref}(F) = \langle I|B \rangle$ where B is the inverse of A . Part (c) says $\mathbf{rref}(F) = \mathbf{rref}(A, I) = \langle C|B \rangle$ where $C = \mathbf{rref}(A)$. Comparing matrix columns in $\langle C|B \rangle = \langle I|B \rangle$ gives $\mathbf{rref}(A) = C = I$.

Converse: assume $\mathbf{rref}(A) = I$, to prove A invertible. Let $F = \langle A|I \rangle$, then $\mathbf{rref}(F) = \langle C|B \rangle$ for some C, B . Part (c) says $C = \mathbf{rref}(A) = I$. Part (a) says B is the inverse of A . This proves A is invertible and completes (d).

Proof of Theorem 10: It is possible to organize the proof into three cases, by considering the three possible toolkit operations. We don't do the tedious details. Instead, we refer to the *Elementary Matrix Multiply* exercises page 335, for suitable experiments that provide the intuition needed to develop formal proof details.

Proof of Theorem 11: The idea of the proof begins with writing Frame 1 as $C_1 = E_1C$, using Theorem 10. Repeat to write Frame 2 as $C_2 = E_2C_1 =$

E_2E_1C . By induction, Frame k is $C_k = E_kC_{k-1} = E_k \cdots E_2E_1C$. But Frame k is matrix D in the sequence. The proof is complete.

Proof of Theorem 12: The reduced row-echelon matrix $D = \mathbf{rref}(A)$ paired with $C = A$ imply by Theorem 11 that $\mathbf{rref}(A) = D = E_k \cdots E_2E_1C = E_k \cdots E_2E_1A$. The proof is complete.

Exercises 5.2

Identify RREF. Mark the matrices which pass the RREF Test, page 318. Explain the failures.

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

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

3.
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

4.
$$\begin{pmatrix} 1 & 1 & 4 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Lead and Free Variables. For each matrix A , assume a homogeneous system $A\vec{x} = \vec{0}$ with variable list x_1, \dots, x_n . List the lead and free variables. Then report the rank and nullity of matrix A .

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

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

7.
$$\begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

8.
$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

9.
$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

10.
$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

11.
$$\begin{pmatrix} 1 & 1 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

12.
$$\begin{pmatrix} 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

13.
$$\begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

14.
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

15.
$$\begin{pmatrix} 0 & 1 & 0 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

16.
$$\begin{pmatrix} 1 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Elementary Matrices. Write the 3×3 elementary matrix E and its inverse E^{-1} for each of the following operations, defined on page 317.

17. $\text{combo}(1, 3, -1)$

18. $\text{combo}(2, 3, -5)$

19. $\text{combo}(3, 2, 4)$

20. `combo(2,1,4)`
 21. `combo(1,2,-1)`
 22. `combo(1,2,-e2)`
 23. `mult(1,5)`
 24. `mult(1,-3)`
 25. `mult(2,5)`
 26. `mult(2,-2)`
 27. `mult(3,4)`
 28. `mult(3,5)`
 29. `mult(2,- π)`
 30. `mult(1,e2)`
 31. `swap(1,3)`
 32. `swap(1,2)`
 33. `swap(2,3)`
 34. `swap(2,1)`
 35. `swap(3,2)`
 36. `swap(3,1)`

Elementary Matrix Multiply. For each given matrix B_1 , perform the toolkit operation (`combo`, `swap`, `mult`) to obtain the result B_2 . Then compute the elementary matrix E for the identical toolkit operation. Finally, verify the matrix multiply equation $B_2 = EB_1$.

37. $\begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$, `mult(2,1/3)`.
 38. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, `mult(1,3)`.
 39. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, `combo(3,2,-1)`.
 40. $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, `combo(2,1,-3)`.

41. $\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, `swap(2,3)`.

42. $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, `swap(1,2)`.

Inverse Row Operations. Given the final frame B of a sequence starting with matrix A , and the given operations, find matrix A . Do not use matrix multiply.

43. $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations
`combo(1,2,-1)`, `combo(2,3,-3)`,
`mult(1,-2)`, `swap(2,3)`.

44. $B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations
`combo(1,2,-1)`, `combo(2,3,3)`,
`mult(1,2)`, `swap(3,2)`.

45. $B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations
`combo(1,2,-1)`, `combo(2,3,3)`,
`mult(1,4)`, `swap(1,3)`.

46. $B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations
`combo(1,2,-1)`, `combo(2,3,4)`,
`mult(1,3)`, `swap(3,2)`.

Elementary Matrix Products.

Given the first frame B_1 of a sequence and elementary matrix operations E_1, E_2, E_3 , find matrix $F = E_3E_2E_1$ and $B_4 = FB_1$. Hint: Compute $\langle B_4|F \rangle$ from toolkit operations on $\langle B_1|I \rangle$.

47. $B_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations
`combo(1,2,-1)`, `combo(2,3,-3)`,
`mult(1,-2)`.

48. $B_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, operations
 combo(1,2,-1), combo(2,3,3),
 swap(3,2).

49. $B_1 = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations
 combo(1,2,-1), mult(1,4),
 swap(1,3).

50. $B_1 = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$, operations
 combo(1,2,-1), combo(2,3,4),
 mult(1,3).

Miscellany.

51. Justify with English sentences why all possible 2×2 matrices in reduced row-echelon form must look like

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

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

where $*$ denotes an arbitrary number.

52. Display all possible 3×3 matrices in reduced row-echelon form. Besides the zero matrix and the identity matrix, please report five other forms, most containing symbol $*$ representing an arbitrary number.
53. Determine all possible 4×4 matrices in reduced row-echelon form.
54. Display a 6×6 matrix in reduced row-echelon form with rank 4 and only entries of zero and one.
55. Display a 5×5 matrix in reduced row-echelon form with nullity 2 having entries of zero, one and two, but no other entries.

56. Display the rank and nullity of any $n \times n$ elementary matrix.

57. Let $F = \langle C|D \rangle$ and let E be a square matrix with row dimension matching F . Display the details for the equality

$$EF = \langle EC|ED \rangle.$$

58. Let $F = \langle C|D \rangle$ and let E_1, E_2 be $n \times n$ matrices with n equal to the row dimension of F . Display the details for the equality

$$E_2 E_1 F = \langle E_2 E_1 C | E_2 E_1 D \rangle.$$

59. Display details explaining why $\mathbf{rref}(\langle A|I \rangle)$ equals the matrix $\langle \mathbf{rref}(A)|B \rangle$, where matrix $B = E_k \cdots E_1$. Symbols E_i are elementary matrices in toolkit steps taking $\langle A|I \rangle$ into reduced row-echelon form. Suggestion: Use the preceding exercises.

60. Assume E_1, E_2 are elementary matrices in toolkit steps taking A into reduced row-echelon form. Prove that $A^{-1} = E_2 E_1$. In words, A^{-1} is found by doing the same toolkit steps to the identity matrix.

61. Assume E_1, \dots, E_k are elementary matrices in toolkit steps taking $\langle A|I \rangle$ into reduced row-echelon form. Prove that $A^{-1} = E_k \cdots E_1$.

62. Assume A, B are 2×2 matrices. Assume $\mathbf{rref}(\langle A|B \rangle) = \langle I|D \rangle$. Explain why the first column \vec{x} of D is the unique solution of $A\vec{x} = \vec{b}$, where \vec{b} is the first column of B .

63. Assume A, B are $n \times n$ matrices. Explain how to solve the matrix equation $AX = B$ for matrix X using the augmented matrix of A, B .