Elementary Matrices and Toolkit Sequences

- Elementary Matrices
  - Definition of elementary matrix
  - Computer algebra systems and elementary matrices
  - Constructing elementary matrices $\boldsymbol{E}$ and their inverses $\boldsymbol{E}^{-1}$
- Fundamental Theorem on Elementary Matrices
- A certain $\mathbf{6}$-frame toolkit sequence.
- Toolkit Sequence Details
- Fundamental Theorem Illustrated
- The RREF Inverse Method

Elementary Matrices

Definition. An elementary matrix $\boldsymbol{E}$ is the result of applying a combination, multiply or swap rule to the identity matrix.

An elementary matrix is then the **second frame** after a `combo`, `swap` or `mult` toolkit operation which has been applied to a **first frame** equal to the identity matrix.

Example:

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

First frame = identity matrix.

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{-5} & \mathbf{0} & \mathbf{1} \end{pmatrix}$$

Second frame
Elementary combo matrix
`combo(1,3,-5)`

Computer algebra systems and elementary matrices

The computer algebra system `maple` displays typical 4×4 elementary matrices (C=Combination, M=Multiply, S=Swap) as follows.

`with(linalg):`	`with(LinearAlgebra):`
`Id:=diag(1,1,1,1);`	`Id:=IdentityMatrix(4);`
`C:=addrow(Id,2,3,c);`	`C:=RowOperation(Id,[3,2],c);`
`M:=mulrow(Id,3,m);`	`M:=RowOperation(Id,3,m);`
`S:=swaprow(Id,1,4);`	`S:=RowOperation(Id,[4,1]);`

The answers:

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

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

Constructing elementary matrices E and their inverses E^{-1}

Mult	Change a one in the identity matrix to symbol $m \neq 0$.
Combo	Change a zero in the identity matrix to symbol c.
Swap	Interchange two rows of the identity matrix.

Constructing E^{-1} from elementary matrix E

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

Fundamental Theorem on Elementary Matrices

Theorem 1 (Toolkit sequences and elementary matrices)
In a Toolkit sequence, let the second frame $\boldsymbol{A_2}$ be obtained from the first frame $\boldsymbol{A_1}$ by a `combo`, `swap` or `mult` toolkit operation. Let $\boldsymbol{n}$ equal the row dimenson of $\boldsymbol{A_1}$.Then there is correspondingly an $\boldsymbol{n \times n}$ `combo`, `swap` or `mult` elementary matrix $\boldsymbol{E}$ such that

$$\boldsymbol{A_2 = EA_1.}$$

Theorem 2 (The $\mathbf{rref}$ and elementary matrices)
Let $\boldsymbol{A}$ be a given matrix of row dimension $\boldsymbol{n}$. Then there exist $\boldsymbol{n \times n}$ elementary matrices $\boldsymbol{E_1}$, $\boldsymbol{E_2}$, ..., $\boldsymbol{E_k}$ such that

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

Proof of Theorem 1

The first result is the observation that left multiplication of matrix $\boldsymbol{A}_1$ by elementary matrix $\boldsymbol{E}$ gives the answer $\boldsymbol{A_2} = \boldsymbol{E}\boldsymbol{A}_1$ which is obtained by applying the corresponding `combo`, `swap` or `mult` toolkit operation. This fact is discovered by doing examples, then a formal proof can be constructed (not presented here).

Proof of Theorem 2

The second result applies the first result multiple times to obtain elementary matrices $\boldsymbol{E}_1$, $\boldsymbol{E_2}, \ldots$ which represent the multiply, combination and swap operations performed in the toolkit sequence which take the First Frame $\boldsymbol{A}_1 = \boldsymbol{A}$ into the Last Frame $\boldsymbol{A}_{k+1} = \mathbf{rref}(\boldsymbol{A}_1)$. Combining the identities

$$\boldsymbol{A}_2 = \boldsymbol{E}_1\boldsymbol{A}_1, \quad \boldsymbol{A}_3 = \boldsymbol{E}_2\boldsymbol{A}_2, \quad \ldots, \quad \boldsymbol{A}_{k+1} = \boldsymbol{E}_k\boldsymbol{A}_k$$

gives the matrix multiply equation

$$\boldsymbol{A}_{k+1} = \boldsymbol{E}_k\boldsymbol{E}_{k-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}_1$$

or equivalently the theorem's result, because $\boldsymbol{A}_{k+1} = \mathbf{rref}(\boldsymbol{A})$ and $\boldsymbol{A}_1 = \boldsymbol{A}$.

A certain 6-frame toolkit sequence ______________________________.

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 1, original matrix.

$$A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -6 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 2, combo(1,2,-2).

$$A_3 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 3 & 6 & 3 \end{pmatrix}$$ Frame 3, mult(2,-1/6).

$$A_4 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & -6 \end{pmatrix}$$ Frame 4, combo(1,3,-3).

$$A_5 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ Frame 5, combo(2,3,-6).

$$A_6 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$ Frame 6, combo(2,1,-3). Found $\mathbf{rref}(A_1)$.

Continued

The corresponding 3×3 elementary matrices are

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Frame 2, combo(1,2,-2) applied to I.

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/6 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$ Frame 3, mult(2,-1/6) applied to I.

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$$ Frame 4, combo(1,3,-3) applied to I.

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix}$$ Frame 5, combo(2,3,-6) applied to I.

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

Frame Sequence Details

$\boldsymbol{A_2 = E_1 A_1}$	Frame 2, $\boldsymbol{E_1}$ equals combo(1,2,-2) on $\boldsymbol{I}$.
$\boldsymbol{A_3 = E_2 A_2}$	Frame 3, $\boldsymbol{E_2}$ equals mult(2,-1/6) on $\boldsymbol{I}$.
$\boldsymbol{A_4 = E_3 A_3}$	Frame 4, $\boldsymbol{E_3}$ equals combo(1,3,-3) on $\boldsymbol{I}$.
$\boldsymbol{A_5 = E_4 A_4}$	Frame 5, $\boldsymbol{E_4}$ equals combo(2,3,-6) on $\boldsymbol{I}$.
$\boldsymbol{A_6 = E_5 A_5}$	Frame 6, $\boldsymbol{E_5}$ equals combo(2,1,-3) on $\boldsymbol{I}$.
$\boldsymbol{A_6 = E_5 E_4 E_3 E_2 E_1 A_1}$	Summary frames 1-6.

Then

$$\mathbf{rref}(\boldsymbol{A_1}) = \boldsymbol{E_5 E_4 E_3 E_2 E_1 A_1},$$

which is the result of the Theorem.

Fundamental Theorem Illustrated

The summary:

$$A_6 = \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{6} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A_1$$

Because $A_6 = \mathbf{rref}(A_1)$, the above equation gives the inverse relationship

$$A_1 = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} E_5^{-1}\, \mathbf{rref}(A_1).$$

Each inverse matrix is simplified by the rules for constructing E^{-1} from elementary matrix E, the result being

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

Theorem 3 (RREF Inverse Method)

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

Proof: For *any* matrix $\boldsymbol{E}$ there is the matrix multiply identity

$$\boldsymbol{E}\langle \boldsymbol{C}|\boldsymbol{D}\rangle = \langle \boldsymbol{EC}|\boldsymbol{ED}\rangle.$$

This identity is proved by arguing that each side has identical columns. For example, $\mathbf{col}(\mathrm{LHS}, 1) = \boldsymbol{E}\,\mathbf{col}(\boldsymbol{C}, 1) = \mathbf{col}(\mathrm{RHS}, 1)$.

Assume $\boldsymbol{C} = \langle \boldsymbol{A}|\boldsymbol{I}\rangle$ satisfies $\mathbf{rref}(\boldsymbol{C}) = \langle \boldsymbol{I}|\boldsymbol{B}\rangle$. The fundamental theorem of elementary matrices implies $\boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{C} = \mathbf{rref}(\boldsymbol{C})$. Then

$$\mathbf{rref}(\boldsymbol{C}) = \langle \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{A} \,|\, \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{I}\rangle = \langle \boldsymbol{I}|\boldsymbol{B}\rangle$$

implies that $\boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{A} = \boldsymbol{I}$ and $\boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{I} = \boldsymbol{B}$. Together, $\boldsymbol{BA} = \boldsymbol{I}$ and then $\boldsymbol{B}$ is the inverse of $\boldsymbol{A}$.

Conversely, assume that $\boldsymbol{AB} = \boldsymbol{I}$. Then $\boldsymbol{A}$ has inverse $\boldsymbol{B}$. The fundamental theorem of elementary matrices implies the identity $\boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{A} = \mathbf{rref}(\boldsymbol{A}) = \boldsymbol{I}$. It follows that $\boldsymbol{B} = \boldsymbol{E}_k \cdots \boldsymbol{E}_1$. Then $\mathbf{rref}(\boldsymbol{C}) = \boldsymbol{E}_k \cdots \boldsymbol{E}_1 < \boldsymbol{A}|\boldsymbol{I} >=< \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{A} \,|\, \boldsymbol{E}_k \cdots \boldsymbol{E}_1 \boldsymbol{I} >=< \boldsymbol{I}|\boldsymbol{B} >$.