Techniques for solving systems of linear equations lie at the heart of linear algebra. In high school we learn to solve systems with 2 or 3 variables using "elimination" and "substitution" of variables. In order to solve systems with a large number of variables we need to be more organized. The process of *Gauss-Jordan Elimination* gives us a systematic way of solving linear systems.

To solve a system of equations we should first drop as many unnecessary symbols as possible. This is done by constructing an *augmented matrix*. [*Note:* The word *matrix* comes to us from Latin. In Middle English, *matrix* meant *womb*. It can be used more generally to mean an enclosure. In mathematics, a matrix is a rectangular array filled numbers, functions, or more general objects.]

## Example:

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ & & 5y & - & 6z & = & 0 \\ -x & & & + & 4z & = & 7 \end{array} \quad \Longrightarrow \quad \left[\begin{array}{rrrcr} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{array}\right]$$

To solve our system we need to manipulate our equations. We will see that standard manipulations correspond to changing the *rows* of our augmented matrix. In the end, it turns out that we need just 3 types of operations to solve any linear system. We call these *elementary (row) operations*.

## Definition: Elementary Row Operations

| | **Effect on the linear system:** | | **Effect on the matrix:** |
|---|---|---|---|
| **Type I** | Interchange equation $i$ and equation $j$ (List the equations in a different order.) | $\Longleftrightarrow$ | Swap Row $i$ and Row $j$ |
| **Type II** | Multiply both sides of equation $i$ by a non-zero scalar $c$ | $\Longleftrightarrow$ | Multiply Row $i$ by $c$ where $c \neq 0$ |
| **Type III** | Multiply both sides of equation $i$ by $c$ and add to equation $j$ | $\Longleftrightarrow$ | Add $c$ times Row $i$ to Row $j$ where $c$ is any scalar |

If we can get matrix $A$ from matrix $B$ by performing a series of elementary row operations, then $A$ and $B$ are called **row equivalent matrices**.

Of course, there are also corresponding *elementary column operations*. If we can get matrix $A$ from matrix $B$ by performing a series of elementary column operations, we call $A$ and $B$ **column equivalent matrices**. Both of these *equivalences* are in fact equivalence relations (proof?). While both row and column operations are important, we will (for now) focus on row operations since they correspond to steps used when solving linear systems.

## Example: Type I — swap rows 1 and 3

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ & & 5y & - & 6z & = & 0 \\ -x & & & + & 4z & = & 7 \end{array} \quad \Longrightarrow \quad \begin{array}{rcrcrcr} -x & & & + & 4z & = & 7 \\ & & 5y & - & 6z & = & 0 \\ 2x & - & y & + & 3z & = & -1 \end{array}$$

$$\left[\begin{array}{rrrcr} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{array}\right] \underset{\sim}{\overset{R1 \leftrightarrow R3}{}} \left[\begin{array}{rrrcr} -1 & 0 & 4 & : & 7 \\ 0 & 5 & -6 & : & 0 \\ 2 & -1 & 3 & : & -1 \end{array}\right]$$

**Example:** Type II — scale row 3 by -2

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ & & 5y & - & 6z & = & 0 \\ -x & & & + & 4z & = & 7 \end{array} \quad \Longrightarrow \quad \begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ & & 5y & - & 6z & = & 0 \\ 2x & & & + & -8z & = & -14 \end{array}$$

$$\begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{bmatrix} \underset{\sim}{-2\times R3} \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ 2 & 0 & -8 & : & -14 \end{bmatrix}$$

**Example:** Type III — add 3 times row 3 to row 2

$$\begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ & & 5y & - & 6z & = & 0 \\ -x & & & + & 4z & = & 7 \end{array} \quad \Longrightarrow \quad \begin{array}{rcrcrcr} 2x & - & y & + & 3z & = & -1 \\ -3x & & 5y & + & 6z & = & 21 \\ -x & & & + & 4z & = & 7 \end{array}$$

$$\begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{bmatrix} \underset{\sim}{3\times R3+R2} \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ -3 & 5 & 6 & : & 21 \\ -1 & 0 & 4 & : & 7 \end{bmatrix}$$

It is important to notice several things about these operations. First, they are all reversible (that's why we want $c \neq 0$ in type II operations) — in fact the inverse of a type X operation is another type X operation. Next, these operations don't effect the set of solutions for the system — that is — row equivalent matrices represent systems with the same set of solutions. Finally, these are *row* operations — columns *never* interact with each other. This last point is quite important as it will allow us to check our work and later allow us to find bases for subspaces associated with matrices (see "Linear Correspondence" between columns).

If we already know about matrix multiplication, it is interesting to note that an elementary row operation performed on $A$ can be accomplished by multiplying $A$ on the *left* by a square matrix (called an elementary matrix). Likewise, multiplying $A$ on the *right* by an elementary matrix performs a column operation.

Let $\mathbf{e}_i = [0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]^T \in \mathbb{R}^{n\times 1}$ where the non-zero entry is located in the $i$-position and the "$T$" means transpose (i.e. this is a column vector). For example, $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$ (in $\mathbb{R}^2$ we sometimes call $\mathbf{e}_1 = \mathbf{i}$ and $\mathbf{e}_2 = \mathbf{j}$ and in $\mathbb{R}^3$ sometimes we say $\mathbf{e}_1 = \mathbf{i}$, $\mathbf{e}_2 = \mathbf{j}$, $\mathbf{e}_3 = \mathbf{k}$). Recall that the identity matrix $I_n$ is defined to be $[\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n]$. We also define $E_{ij} = \mathbf{e}_i\mathbf{e}_j^T$. This matrix has a 1 in the $(i,j)$-position and 0's elsewhere.

**Example:** In $\mathbb{R}^{2\times 2}$, $E_{21} = \mathbf{e}_2\mathbf{e}_1^T = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Let $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ be the *Kronecker delta*. Then $\mathbf{e}_j = [v_i]^T$ where $v_i = \delta_{ij}$. Also, $I_n = [\delta_{ij}]$. For example: $E_{ij} = \mathbf{e}_i\mathbf{e}_j^T$ has entry $(k,\ell)$ given by the $k^{\text{th}}$ entry (i.e. row) of $\mathbf{e}_i$ multiplied (i.e. dotted) by the $\ell^{\text{th}}$ entry (i.e. column) if $\mathbf{e}_j^T$. So the $(k,\ell)$-entry is $\delta_{ik}\delta_{j\ell}$. Of course, this is only non-zero (in which case it equals $1 \cdot 1 = 1$) when $i = k$ and $j = \ell$. In other words, the only non-zero entry of $E_{ij}$ is the $(i,j)$-entry – which we already "knew".

Notice that $A\mathbf{e}_\ell$ picks off the $\ell^{\text{th}}$ column of $A$. If $A \in \mathbb{R}^{m\times n}$ and $\mathbf{e}_\ell \in \mathbb{R}^{n\times 1}$. Then $A\mathbf{e}_\ell \in \mathbb{R}^{m\times 1}$ (a column vector with $m$-entries). The $i^{\text{th}}$ entry would come from dotting the $i^{\text{th}}$ row of $A$ with $\mathbf{e}_\ell$. This amounts to the sum $a_{i1}\delta_{1\ell}+a_{i2}\delta_{2\ell}+\cdots+a_{in}\delta_{n\ell}$. Of course, the only non-zero term in this sum is $a_{i\ell}\delta_{\ell\ell} = a_{i\ell}$. Therefore, $A\mathbf{e}_\ell$ does indeed pick off the $\ell^{\text{th}}$ column of $A$. Similarly, $\mathbf{e}_k^T A$ picks off the $k^{\text{th}}$ row of $A$.

**Example:** Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. $\quad A\mathbf{e}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\mathbf{e}_2^T A = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \end{bmatrix}$.

Next, since matrix multiplication can be viewed as $A[\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_\ell] = [A\mathbf{v}_1\ A\mathbf{v}_2\ \cdots\ A\mathbf{v}_\ell]$ (i.e. done column-by-column) or $\begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_k \end{bmatrix} A = \begin{bmatrix} \mathbf{w}_1 A \\ \vdots \\ \mathbf{w}_k A \end{bmatrix}$ (i.e. done row-by-row), we get that $AE_{ij} = A[0\ \cdots\ 0\ \mathbf{e}_i\ 0\ \cdots\ 0] = [0\ \cdots\ 0\ A\mathbf{e}_i\ 0\ \cdots\ 0]$ is merely the $i^{\text{th}}$ column of $A$ slapped into the $j^{\text{th}}$ column of the zero matrix. Likewise, $E_{ij}A$ is the $j^{\text{th}}$ row of $A$ slapped into the $i^{\text{th}}$ row of the zero matrix.

**Example:** Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. $AE_{21} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 4 & 0 \end{bmatrix}$ and $E_{21}A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}$.

**Question:** If $E = I_m - E_{ii} - E_{jj} + E_{ij} + E_{ji} \in \mathbb{R}^{m \times m}$ and $A \in \mathbb{R}^{m \times n}$, what is $EA$?

To answer this question, let's analyze what $E$ actually is. $E_{ii}A$ would be the $i^{\text{th}}$ row of $A$ left in place with all other rows zeroed out. $E_{ij}A$ would be the $j^{\text{th}}$ row of $A$ moved to the $i^{\text{th}}$ row with all other rows zeroed out. Therefore, $(I_m - E_{ii} - E_{jj})A = A - E_{ii}A - E_{jj}A$ wipes out rows $i$ and $j$. Then by adding in $(E_{ij} + E_{ji})A$, we put rows $i$ and $j$ back but interchanging their locations. Therefore, $EA$ swaps rows $i$ and $j$. In particular, $E = EI_m$ is the identity matrix with rows $i$ and $j$ swapped!

First, $\mathbf{e}_i\mathbf{e}_i^T$ and $\mathbf{e}_j\mathbf{e}_j^T$ are matrices with 1's in the $(i,i)$ and $(j,j)$ entries respectively. Thus $I_m - \mathbf{e}_i\mathbf{e}_i^T - \mathbf{e}_j\mathbf{e}_j^T$ removes the 1's from the $i^{\text{th}}$ and $j^{\text{th}}$ rows of the identity matrix. Then we add in $\mathbf{e}_i\mathbf{e}_j^T$ and $\mathbf{e}_j\mathbf{e}_i^T$. This just adds $1's$ in the $(i,j)$ and $(j,i)$ positions. The net effect is that $E$ is the identity matrix with the $i^{\text{th}}$ and $j^{\text{th}}$ rows swapped!

It is interesting to note that $E^T = I_m^T - E_{ii}^T - E_{jj}^T + E_{ij}^T + E_{ji}^T = I_m - E_{ii} - E_{jj} + E_{ji} + E_{ij} = E$. Also, since swapping twice undoes the swap, $E^{-1} = E$ ($E$ is its own inverse). It is also worth noting that if $B \in \mathbb{R}^{k \times m}$ (i.e. compatibly sized), then $BE$ is $B$ with *columns* $i$ and $j$ swapped.

This takes care of type I operations. Likewise we can deal with types II and III. To scaled a row by $c$ simply use $E = I_m - E_{ii} + cE_{ii}$. Notice that $E^{-1} = I_m - E_{ii} + c^{-1}E_{ii}$ (to undo scaling row $i$ by $c$ we should scale row $i$ by $1/c$). So $E^{-1}$ correspond to a type II operation. Again, $EA$ scales row $i$ of $A$ by $c$ and $BE$ scales column $i$ of $B$ by $c$.

For type III, we have that $E_{ji}$ (note the subscripts) will copy row $i$ into row $j$'s place. So $E = I_m + sE_{ji}$ will add $s$ times row $i$ to $j$. To undo adding $s$ times row $j$ to row $i$ we should subtract $s$ times row $j$ from row $i$. Therefore, $E^{-1} = I_m - sE_{ji}$, so yet again the inverse of an elementary operation is an elementary operation of the same type. Also, $EA$ adds $s$ times row $i$ to row $j$ while $BE$ adds $s$ times column $j$ to column $i$. Let's re-examine our row operation examples from before.

**Example:** Type I — swap rows 1 and 3 (so $E = I_3 - E_{11} - E_{33} + E_{13} + E_{31}$).

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 4 & : & 7 \\ 0 & 5 & -6 & : & 0 \\ 2 & -1 & 3 & : & -1 \end{bmatrix}$$

**Example:** Type II — scale row 3 by -2 (so $E = I_3 - E_{33} + (-2)E_{33}$).

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ 2 & 0 & -8 & : & -14 \end{bmatrix}$$

**Example:** Type III — add 3 times row 3 to row 2 (so $E = I_3 + 3E_{23}$).

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ 0 & 5 & -6 & : & 0 \\ -1 & 0 & 4 & : & 7 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 3 & : & -1 \\ -3 & 5 & 6 & : & 21 \\ -1 & 0 & 4 & : & 7 \end{bmatrix}$$

**Example:** Column Type III operation — add $-2$ times column 1 to column 2 (so $E = I_2 - 2E_{12}$).

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & -2 \\ 5 & -4 \end{bmatrix}$$

Now let's come back to trying to solve a linear system. Doing operations blindly probably won't get us anywhere. Instead we will choose our operations carefully so that we head towards some shape of equations which will let us read off the set of solutions. Thus the next few defintions.

**Definition:** A matrix is in **Row Echelon Form** (or REF) if...

- Each non-zero row is above all zero rows – that is – zero rows are "pushed" to the bottom.
- The leading entry of a row is *strictly* to the right of the leading entries of the rows above. (The leftmost non-zero entry of a row is called the "leading entry".)

If in addition...

- Each leading entry is "1". (*Note:* Some textbooks say this is a requirement of REF.)
- Only zeros appear above (& below) a leading entry of a row.

then a matrix is in **reduced row echelon form** (or RREF).

**Example:** The matrix $A$ (below) is in REF but is not reduced. The matrix $B$ is in RREF.

$$A = \begin{bmatrix} 0 & 2 & 3 & 4 & 1 & 5 \\ 0 & 0 & 0 & -3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

**Gauss-Jordan Elimination** is an "algorithm" which given a matrix returns a row equivalent matrix in reduced row echelon form (RREF). We first perform a **forward pass**:

1. Determine the leftmost non-zero column. This is a **pivot column** and the topmost entry is a **pivot position**. If "0" is in this pivot position, swap (an unignored) row with the topmost row (use a Type I operation) so that there is a non-zero entry in the pivot position.
2. Add appropriate multiples of the topmost (unignored) row to the rows beneath it so that only "0" appears below the pivot (use several Type III operations).
3. Ignore the topmost (unignored) row. If any non-zero rows remain, go to step 1.

The forward pass is now complete. Now our matrix in row echelon form (in my sense not our textbook's sense). Sometimes the forward pass alone is referred to as "Gaussian Elimination". However, we should be careful since the term "Gaussian Elimination" more commonly refers to *both* the forward and backward passes. Now let's finish Gauss-Jordan Elimination by performing a **backward pass**:

4. If necessary, scale the rightmost unfinished pivot to 1 (use a Type II operation).
5. Add appropriate multiples of the current pivot's row to rows above it so that only 0 appears above the current pivot (using several Type III operations).
6. The current pivot is now "finished". If any unfinished pivots remain, go to step 4.

It should be fairly obvious that the entire Gauss-Jordan algorithm will terminate in finitely many steps. Also, only elementary row operations have been used. So we end up with a row equivalent matrix. A tedious, wordy, and unenlightening proof would show us that the resulting matrix is in reduced row echelon form (RREF).

**Example:** Let's solve the system $\begin{array}{rcrcr} x & + & 2y & = & 1 \\ 3x & + & 4y & = & -1 \end{array}$

$$\begin{bmatrix} 1 & 2 & : & 1 \\ 3 & 4 & : & -1 \end{bmatrix} \underset{\sim}{\overset{-3\times R1+R2}{}} \begin{bmatrix} 1 & 2 & : & 1 \\ 0 & -2 & : & -4 \end{bmatrix}$$

The first non-zero column is just the first column. So the upper left hand corner is a pivot position. This position already has a non-zero entry so no swap is needed. The type III operation "$-3$ times row 1 added to row 2" clears the only position below the pivot, so after one operation we have finished with this pivot and can ignore row 1. $\begin{bmatrix} 1 & 2 & : & 1 \\ 0 & -2 & : & -4 \end{bmatrix}$

Among the (unignored parts of) columns the leftmost non-zero column is the second column. So the "$-2$" sits in a pivot position. Since it's non-zero, no swap is needed. Also, there's nothing below it, so no type III operations are necessary. Thus we're done with this row and we can ignore it.

$$\begin{bmatrix} 1 & 2 & : & 1 \\ 0 & -2 & : & -4 \end{bmatrix}$$

Nothing's left so we're done with the forward pass. $\begin{bmatrix} 1 & 2 & : & 1 \\ 0 & -2 & : & -4 \end{bmatrix}$ is in row echelon form.

Next, we need to take the rightmost pivot (the "$-2$") and scale it to 1 then clear everything above it.

$$\begin{bmatrix} 1 & 2 & : & 1 \\ 0 & -2 & : & -4 \end{bmatrix} \underset{\sim}{\overset{-1/2\times R2}{}} \begin{bmatrix} 1 & 2 & : & 1 \\ 0 & 1 & : & 2 \end{bmatrix} \underset{\sim}{\overset{-2\times R2+R1}{}} \begin{bmatrix} 1 & 0 & : & -3 \\ 0 & 1 & : & 2 \end{bmatrix}$$

This "finishes" that pivot. The next rightmost pivot is the 1 in the upper left hand corner. But it's already scaled to 1 and has nothing above it, so it's finished as well. That takes care of all of the pivots so the backward pass is complete leaving our matrix in reduced row echelon form.

Finally, let translate the RREF matrix back into a system of equations. The (new equivalent) system is $\begin{array}{rcr} x & = & -3 \\ y & = & 2 \end{array}$. So the only solution for this system is $x = -3$ and $y = 2$.

*Note:* One can also solve a system quite easily once (just) the forward pass is complete. This is done using "back substitution". Notice that the system after the forward pass was $\begin{array}{rcrcr} x & + & 2y & = & 1 \\ & & -2y & = & -4 \end{array}$. So we have $-2y = -4$ thus $y = 2$. Substituting this back into the first equation we get $x + 2(2) = 1$ so $x = -3$.

**Example:** Let's solve the system $\begin{array}{rcrcrcr} 2x & - & y & + & z & = & 0 \\ 4x & - & 2y & + & 2z & = & 1 \\ x & & & + & z & = & -1 \end{array}$

$$\begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 4 & -2 & 2 & : & 1 \\ 1 & 0 & 1 & : & -1 \end{bmatrix} \underset{\sim}{\overset{-2\times R1+R2}{}} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \\ 1 & 0 & 1 & : & -1 \end{bmatrix} \underset{\sim}{\overset{-1/2\times R1+R3}{}} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \\ 0 & 1/2 & 1/2 & : & -1 \end{bmatrix} \underset{\sim}{\text{Ignore } R1}$$

$$\begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \\ 0 & 1/2 & 1/2 & : & -1 \end{bmatrix} \underset{\sim}{\overset{R2\leftrightarrow R3}{}} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1/2 & 1/2 & : & -1 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{\text{Ignore } R2} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1/2 & 1/2 & : & -1 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{\text{Ignore } R3}$$

which leaves us with nothing. So the forward pass is complete and $\begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1/2 & 1/2 & : & -1 \\ 0 & 0 & 0 & : & 1 \end{bmatrix}$ is in REF.

$$\begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1/2 & 1/2 & : & -1 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{1{\times}R3{+}R2} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1/2 & 1/2 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{2{\times}R2} \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{1{\times}R2{+}R1}$$

$$\begin{bmatrix} 2 & 0 & 2 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{1/2{\times}R1} \begin{bmatrix} 1 & 0 & 1 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix}$$

This finishes the backward pass and our matrix is now in RREF. Our new system of equations is $\begin{array}{rcrcl} x & & & + & z & = & 0 \\ & & y & + & z & = & 0 \\ & & & & 0 & = & 1 \end{array}$. Of course $0 \neq 1$, so this is an **inconsistent** system — it has **no solutions**.

*Note:* If our only goal was to solve this system, we could have stopped after the very first operation (row number 2 already said "$0 = 1$").

**Example:** Let's solve the system $\begin{array}{rcrcrcr} x & + & 2y & + & 3z & = & 3 \\ 4x & + & 5y & + & 6z & = & 9 \\ 7x & + & 8y & + & 9z & = & 15 \end{array}$

$$\begin{bmatrix} 1 & 2 & 3 & : & 3 \\ 4 & 5 & 6 & : & 9 \\ 7 & 8 & 9 & : & 15 \end{bmatrix} \underset{\sim}{-4{\times}R1{+}R2} \begin{bmatrix} 1 & 2 & 3 & : & 3 \\ 0 & -3 & -6 & : & -3 \\ 7 & 8 & 9 & : & 15 \end{bmatrix} \underset{\sim}{-7{\times}R1{+}R3} \begin{bmatrix} 1 & 2 & 3 & : & 3 \\ 0 & -3 & -6 & : & -3 \\ 0 & -6 & -12 & : & -6 \end{bmatrix} \underset{\sim}{-2{\times}R2{+}R3}$$

$$\begin{bmatrix} 1 & 2 & 3 & : & 3 \\ 0 & -3 & -6 & : & -3 \\ 0 & 0 & 0 & : & 0 \end{bmatrix} \underset{\sim}{-1/3{\times}R2} \begin{bmatrix} 1 & 2 & 3 & : & 3 \\ 0 & 1 & 2 & : & 1 \\ 0 & 0 & 0 & : & 0 \end{bmatrix} \underset{\sim}{-2{\times}R2{+}R1} \begin{bmatrix} 1 & 0 & -1 & : & 1 \\ 0 & 1 & 2 & : & 1 \\ 0 & 0 & 0 & : & 0 \end{bmatrix}$$

Now our matrix is in RREF. The new system of equations is $\begin{array}{rcrcl} x & & & - & z & = & 1 \\ & & y & + & 2z & = & 1 \\ & & & & 0 & = & 0 \end{array}$. This is new — we don't have an equation of the form "$z = ...$" This is because $z$ does not lie in a pivot column. So we can make $z$ a "free variable." Let's relabel it $z = t$. Then we have $x - t = 1$, $y + 2t = 1$, and $z = t$. So $x = 1 + t$, $y = 1 - 2t$, and $z = t$ is a solution for any choice of $t$. In particular, $x = y = 1$ and $z = 0$ is a solution. But so is $x = 2$, $y = -1$, $z = 1$. In fact, there are infinitely many solutions.

*Note:* A system of linear equations will always have either one solution, infinitely many solutions, or no solution at all.

**Multiple Systems:** Gauss-Jordan can handle solving multiple systems at once, if these systems share the same coefficient matrix (the part of the matrix before the :'s).

Suppose we wanted to solve both $\begin{array}{rcrcr} x & + & 2y & = & 3 \\ 4x & + & 5y & = & 6 \\ 7x & + & 8y & = & 9 \end{array}$ and also $\begin{array}{rcrcr} x & + & 2y & = & 3 \\ 4x & + & 5y & = & 9 \\ 7x & + & 8y & = & 15 \end{array}$. These lead to the following augmented matrices: $\begin{bmatrix} 1 & 2 & : & 3 \\ 4 & 5 & : & 6 \\ 7 & 8 & : & 9 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & : & 3 \\ 4 & 5 & : & 9 \\ 7 & 8 & : & 15 \end{bmatrix}$. We can combine them together and get $\begin{bmatrix} 1 & 2 & : & 3 & 3 \\ 4 & 5 & : & 6 & 9 \\ 7 & 8 & : & 9 & 15 \end{bmatrix}$ which we already know has the RREF of $\begin{bmatrix} 1 & 0 & : & -1 & 1 \\ 0 & 1 & : & 2 & 1 \\ 0 & 0 & : & 0 & 0 \end{bmatrix}$ (from the last example —

only the :'s have moved). This corresponds to the augmented matrices $\begin{bmatrix} 1 & 0 & : & -1 \\ 0 & 1 & : & 2 \\ 0 & 0 & : & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & : & 1 \\ 0 & 1 & : & 1 \\ 0 & 0 & : & 0 \end{bmatrix}$. These in turn tell us that the first system's solution is $x = -1$, $y = 2$ and the second system's solution is $x = 1$ and $y = 1$.

This works because we aren't mixing columns together (all operations are row operations). Also, notice that the same matrix can be interpreted in a number of ways. Before we had a single system in 3 variables and now we have 2 systems in 2 variables.

**Pivoting:** It turns out that Gaussian elimination is not numerically stable. Small changes in a matrix can yield radically differential RREFs. To combat the propagation of errors (like round off error) one can implement partial or full pivoting methods. These are general purpose techniques to tackle solving systems numerically in a stable (or at least more stable) way.

First, we will mention **partial pivoting**. Solving a system using partial pivoting just modifies step 1 in the forward pass of Gauss-Jordan elimination. Instead of allowing just any element to sit in the pivot position, we swap the element (in an unignored row) of largest magnitude.

This doesn't add much overhead or runtime but greatly improves stability. Why? Well, essentially we get into trouble when doing type III operations to clear out positions below the pivot position. If the pivot is small, we will have to add a *large* multiple of our row to wipe out the entry below. This large multiple can greatly magnify any existing error. On the other hand, with partial pivoting since the pivot position contains an element of larger magnitude than the elements below it, we will never add a multiple of our row that is bigger than 1.

It turns out that partial pivoting is not stable (in theory). To get a fully stable technique we must use **full pivoting**. Here we consider all elements below and to the right of our pivot position and use both row and column swaps to get the element of largest possible magnitude into the pivot position. This gives us a fully stable method. However, we incur a lot of overhead in that every column swap is a relabeling of our variables (and this needs to be accounted for).

In practice partial pivoting is often used but full pivoting not so much. It turns out that most systems that arise in most applications can be solved stably just doing partial pivoting.

**Example:** Let's rework a previous example using partial pivoting.

Our goal os to find the RREF of $A = \begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 4 & -2 & 2 & : & 1 \\ 1 & 0 & 1 & : & -1 \end{bmatrix}$. Now the $(1, 1)$-position is clearly our first pivot. However, there is an element of larger magnitude below 2 so we need to swap rows.

$$\begin{bmatrix} 2 & -1 & 1 & : & 0 \\ 4 & -2 & 2 & : & 1 \\ 1 & 0 & 1 & : & -1 \end{bmatrix} \underset{\sim}{R1 \leftrightarrow R2} \begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 2 & -1 & 1 & : & 0 \\ 1 & 0 & 1 & : & -1 \end{bmatrix} \underset{\sim}{-1/2 \times R1 + R2} \begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 0 & 0 & : & -1/2 \\ 1 & 0 & 1 & : & -1 \end{bmatrix} \underset{\sim}{-1/4 \times R1 + R3}$$

$$\begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 0 & 0 & : & -1/2 \\ 0 & 1/2 & 1/2 & : & -5/4 \end{bmatrix} \underset{\sim}{R2 \leftrightarrow R3} \begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 1/2 & 1/2 & : & -5/4 \\ 0 & 0 & 0 & : & -1/2 \end{bmatrix}$$

Notice that after we cleared out the entries below 4 we found that the $(2, 2)$-position was out next pivot position. Now even though $-2$ in column 2 is bigger (in magnitude) than $1/2$, it does not lie below the pivot position so it should be ignored. Thus we choose from 0 and $1/2$ and so swap $1/2$ into the pivot position. This second column is already cleared out below this pivot. We locate the final pivot position at $(3, 4)$. Again, there are bigger elements in column 4, but none lying below our pivot position so no swap is needed. This completes the forward pass. Notice that our type III operations never involve scalings of magnitude bigger than 1.

$$\begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 1/2 & 1/2 & : & -5/4 \\ 0 & 0 & 0 & : & -1/2 \end{bmatrix} \underset{\sim}{-2\times R3} \begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 1/2 & 1/2 & : & -5/4 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{5/4\times R3+R2} \begin{bmatrix} 4 & -2 & 2 & : & 1 \\ 0 & 1/2 & 1/2 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{-1\times R3+R1}$$

$$\begin{bmatrix} 4 & -2 & 2 & : & 0 \\ 0 & 1/2 & 1/2 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{2\times R2} \begin{bmatrix} 4 & -2 & 2 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{2\times R2+R1} \begin{bmatrix} 4 & 0 & 4 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix} \underset{\sim}{(1/4)\times R1} \begin{bmatrix} 1 & 0 & 1 & : & 0 \\ 0 & 1 & 1 & : & 0 \\ 0 & 0 & 0 & : & 1 \end{bmatrix}$$

**Homework Problems:** For each of the following matrices perform Gauss-Jordan elimination (carefully following the algorithm defined above). I have given the result of the performing the forward pass to help you verify that you are on the right track. [Note: Each matrix has a unique RREF. However, REF is *not* unique. So if you do not follow my algorithm and use your own random assortment of operations, you will almost certainly get different REFs along the way to the RREF.] Once you have completed row reduction, identify the pivots and pivot columns. Finally, interpret your matrices as a system or collection of systems of equations and note the corresponding solutions.

Try doing these again using partial pivoting (you should generally get different REFs along the way).

1. $\begin{bmatrix} 1 & 5 \\ 2 & 2 \end{bmatrix} \Longrightarrow \text{forward pass} \Longrightarrow \begin{bmatrix} 1 & 5 \\ 0 & -8 \end{bmatrix}$

2. $\begin{bmatrix} 2 & -4 & 0 & 4 \\ 1 & -2 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 4 & -8 & 1 & 6 \end{bmatrix} \Longrightarrow \text{forward pass} \Longrightarrow \begin{bmatrix} 2 & -4 & 0 & 4 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \Longrightarrow \text{forward pass} \Longrightarrow \begin{bmatrix} 2 & 1 & 1 \\ 0 & 3/2 & 1/2 \\ 0 & 0 & 4/3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -3 & 0 \\ 2 & -6 & -2 \\ 1 & -3 & -2 \end{bmatrix} \Longrightarrow \text{forward pass} \Longrightarrow \begin{bmatrix} 1 & -3 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

**LU and PLU Decomposition:** After the forward pass of Gaussian elimination, our matrix is left in a REF. This means that every entry below the main diagonal is 0. Square matrices whose entries are 0 below the main diagonal are called **upper triangular** matrices. If all entries above the diagonal are 0 we call the matrix **lower triangular**. It is interesting to note that if we perform a forward pass without needing to use any row swaps, our elementary matrices are all lower triangular type III matrices (adding a multiple of a row to a lower row). We already know that the inverse of a type III matrix is also a type III matrix (with the off diagonal entry negated). Also, it turns out that the product of lower triangular matrices is still lower triangular. Thus if $E_1, \dots, E_\ell$ correspond to these type III row operations needed to finish a forward pass on some matrix $A$, we have $E_\ell \cdots E_2 E_1 A = U$ is upper triangular. Thus $A = E_1^{-1} E_2^{-1} \cdots E_\ell^{-1} U = LU$ where $L$ is lower triangular and $U$ is upper triangular. In fact, $L$ is *unit* lower triangular (meaning it has 1's down its diagonal).

**Example:** Let's find an LU decomposition for the matrix $A$ given below:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 1 \end{bmatrix} \underset{\sim}{-1 \times R1 + R3} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{bmatrix} \underset{\sim}{-2 \times R2 + R3} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} = U$$

So $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ correspond to our first and second operations. Therefore, from our discussion above we find that $L = E_1^{-1} E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$.

Notice that $LU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 4 & 1 \end{bmatrix} = A$.

It is natural to ask when an LU decomposition actually exists. You can find a proof of most matrix analysis texts that an *invertible* square matrix has an LU decomposition exactly when its principal submatrices are all invertible. The $k^{\text{th}}$ principal submatrix of an $m \times m$ matrix $A$ is obtained by deleting the last $k - m$ rows and columns of $A$. In the example above $[1]$ is the first principal submatrix, $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ is the second principal submatrix, and $A$ itself is the third principal submatrix (all of these are invertible and so $A$ had an LU decomposition).

If you perform several LU decompositions, you will noticed a tie between the entries below the diagonal of $L$ and the type III operations being used. This leads to a very efficient means of storing and computing an LU decomposition. Namely, if $s$ times row $i$ is added to row $j$, then the $(i, j)$-entry of $L$ should be $-s$. As we compute, let's store these multipliers in the matrix itself, but we'll highlight them in red and wrap them in parentheses to remind ourselves that they are not really part of the REF we are computing.

**Example:** Let's find an LU decomposition for the matrix $A$ given below:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \underset{\sim}{-4 \times R1 + R2} \begin{bmatrix} 1 & 2 & 3 \\ (4) & -3 & -6 \\ 7 & 8 & 9 \end{bmatrix} \underset{\sim}{-7 \times R1 + R3} \begin{bmatrix} 1 & 2 & 3 \\ (4) & -3 & -6 \\ (7) & -6 & -12 \end{bmatrix} \underset{\sim}{-2 \times R2 + R3} \begin{bmatrix} 1 & 2 & 3 \\ (4) & -3 & -6 \\ (7) & (2) & 0 \end{bmatrix}$$

We can then check that $LU = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = A$.

One reason for computing an LU decomposition is that it allows us to quickly solve systems using forward and backward substitution. If $A = LU$ and we want to solve $A\mathbf{x} = \mathbf{b}$, then we can first solve $L\mathbf{y} = \mathbf{b}$ (using forward substitution) and then solve $U\mathbf{x} = \mathbf{y}$ using back substitution.

**Example:** Let's solve $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b} = \begin{bmatrix} 1 \\ 7 \\ 13 \end{bmatrix}$ and $A$ is the same as the example above.

First we solve $L\mathbf{y} = \mathbf{b}$. Written as equations this is $\begin{array}{ccccccc} x & & & & & = & 1 \\ 4x & + & y & & & = & 7 \\ 7x & + & 2y & + & z & = & 13 \end{array}$.

Immediately we have $x = 1$. Then $4(1) + y = 7$ so $y = 3$. Then $7(1) + 2(3) + z = 13$ so $z = 0$. This means that $\mathbf{y} = [1\ 3\ 0]^T$. Now we need to solve $U\mathbf{x} = \mathbf{y}$. Written as equations this amounts to solving

$$\begin{array}{rrrrrrl} x & + & 2y & + & 3z & = & 1 \\ & & -3y & - & 6z & = & 3 \\ & & & & 0 & = & 0 \end{array}$$

. Therefore, $z = t$ (is free) and so $-3y - 6t = 3$ so that $y = -1 + 2t$. Thus $x + 2(-1+2t) + 3t = 1$ and so $x = 3 - 7t$. Our general solution is $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} -7 \\ 2 \\ 1 \end{bmatrix} t$ (where $t \in \mathbb{R}$).

Our next natural question is to ask if anything can be done when no LU decomposition exists. The answer is yes. In this case we can get a PLU decomposition where $P$ stands for permutation. It is not terribly difficult to prove that if we need row swaps, then these swaps move our parenthesized entries of $L$ around in a compatible way. So computing a PLU decomposition isn't much more difficult. We just need to keep track of how we permuted the rows.

**Example:** Let's compute the PLU decomposition of a matrix and then use it to solve a linear system. I will use partial pivoting as well (since we aren't concerned with trying to avoid row swaps).

$$A = \begin{bmatrix} 0 & 1 & 3 & -1 \\ 2 & 6 & 0 & 4 \\ 1 & 3 & 2 & -5 \\ -1 & 0 & 4 & 0 \end{bmatrix} \underset{\sim}{R1 \leftrightarrow R2} \begin{bmatrix} 2 & 6 & 0 & 4 \\ 0 & 1 & 3 & -1 \\ 1 & 3 & 2 & -5 \\ -1 & 0 & 4 & 0 \end{bmatrix} \underset{\sim}{-1/2 \times R1 + R3} \begin{bmatrix} 2 & 6 & 0 & 4 \\ 0 & 1 & 3 & -1 \\ (1/2) & 0 & 2 & -7 \\ -1 & 0 & 4 & 0 \end{bmatrix} \underset{\sim}{1/2 \times R1 + R4}$$

$$\begin{bmatrix} 2 & 6 & 0 & 4 \\ 0 & 1 & 3 & -1 \\ (1/2) & 0 & 2 & -7 \\ (-1/2) & 3 & 4 & 2 \end{bmatrix} \underset{\sim}{R2 \leftrightarrow R4} \begin{bmatrix} 2 & 6 & 0 & 4 \\ (-1/2) & 3 & 4 & 2 \\ (1/2) & 0 & 2 & -7 \\ 0 & 1 & 3 & -1 \end{bmatrix} \underset{\sim}{-1/3 \times R2 + R4}$$

$$\begin{bmatrix} 2 & 6 & 0 & 4 \\ (-1/2) & 3 & 4 & 2 \\ (1/2) & 0 & 2 & -7 \\ 0 & (1/3) & 5/3 & -5/3 \end{bmatrix} \underset{\sim}{-5/6 \times R3 + R4} \begin{bmatrix} 2 & 6 & 0 & 4 \\ (-1/2) & 3 & 4 & 2 \\ (1/2) & 0 & 2 & -7 \\ 0 & (1/3) & (5/6) & 25/6 \end{bmatrix}$$

Along the way we did two row swaps: $E_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ corresponding to swapping rows 1 and 2 and then rows 2 and 4. Our permutation matrix is then $P = (E_2E_1)^{-1} = E_1^{-1}E_2^{-1} = E_1E_2$.

$$PLU = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ 1/2 & 0 & 1 & 0 \\ 0 & 1/3 & 5/6 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 0 & 4 \\ 0 & 3 & 4 & 2 \\ 0 & 0 & 2 & -7 \\ 0 & 0 & 0 & 25/6 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 3 & -1 \\ 2 & 6 & 0 & 4 \\ 1 & 3 & 2 & -5 \\ -1 & 0 & 4 & 0 \end{bmatrix} = A$$

If we want to solve a system like $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b} = [-2\ 2\ 0\ -1]^T$ using this technique, we can use the same forward and backward substitutions. However, we need to adjust for the permutations. The easiest adjustment is just to use $LU$ and $P^{-1}\mathbf{b}$. Notice that $P^{-1} = E_2E_1$ just swaps the first and second entries then the second and fourth entries: $\mathbf{b} = [-2\ 2\ 0\ -1]^T \to [2\ -2\ 0\ -1]^T \to [2\ -1\ 0\ -2]^T = P^{-1}\mathbf{b}$. Then we forward substitute to solve $L\mathbf{y} = P^{-1}\mathbf{b}$ and get $\mathbf{y} = [2\ 0\ -1\ -7/6]^T$. Finally, back substitute solving $U\mathbf{x} = \mathbf{y}$ and get $\mathbf{x} = \frac{1}{25}[-123\ \ 54\ -37\ -7]^T$.

This PLU decomposition using partial pivoting is an extremely effective relatively stable method for solving linear systems. Once the decomposition is computed, it can be used to solve multiple systems (with different $\mathbf{b}$'s). It can be efficiently stored in memory as well (just overwrite $A$ storing $L$ and $U$ together). We really don't need to store $P$ as a matrix. Just the corresponding permutation needs to be kept around.

But this is not a cure-all. This kind of computation destroys any structure that $A$ might have. For any particular application with $A$ structured in some special way there is usually a better (more stable and faster) method for solving $A\mathbf{x} = \mathbf{b}$.

**Computing Inverses:** Since row operations will take us from $A$ to its RREF (call it $R$) and since these operations correspond to left multiplications by elementary matrices, we can get $E_\ell \cdots E_2 E_1 A = R$ for some list of elementary matrices. Let $P = E_\ell \cdots E_2 E_1$. Then $PA = R$. Note that the product of invertible matrices is still invertible (i.e. $P^{-1}$ must exist because the $E_i^{-1}$'s exist). Therefore, *there exists an invertible matrix $P$ such that $PA$ is the RREF of $A$.*

Now keep in mind that $PI = P$, so $P[A\ I] = [PA\ PI] = [R\ P]$. This means that if we adjoin the identity matrix to $A$ and then row reduce, we can find a matrix $P$ which puts $A$ into RREF.

**Example:** $[A\ :\ I_3] = \begin{bmatrix} 1 & 2 & 3 & : & 1 & 0 & 0 \\ 4 & 5 & 6 & : & 0 & 1 & 0 \\ 7 & 8 & 9 & : & 0 & 0 & 1 \end{bmatrix} \underset{\text{row reduce}}{\sim} \begin{bmatrix} 1 & 0 & -1 & : & -5/3 & 2/3 & 0 \\ 0 & 1 & 2 & : & 4/3 & -1/3 & 0 \\ 0 & 0 & 0 & : & 1 & -2 & 1 \end{bmatrix} = [R\ :\ P].$

Notice that we didn't finish Gauss-Jordan elimination (we could but we didn't need to). We get that...

$$PA = \begin{bmatrix} -5/3 & 2/3 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = R$$

It is worth mentioning that since $A$ is not invertible, the matrix $P$ putting $A$ into RREF is *not unique.* For example, if we finished Gaussian elimination (clearing above the third pivot), we would end up with a different (but valid) $P$.

Of course, this idea is most familiar as the way we compute $A^{-1}$. If $A$'s RREF is $I$ then $P[A\ :\ I] = [I\ :\ P]$ so $PA = I$ and thus $P = A^{-1}$.

It will be extremely useful to notice that when we preform elementary row operations on a matrix $A$, the linear relationships between columns of $A$ do not change. Specifically... Suppose that **a**, **b**, and **c** are columns of $A$, and suppose that we perform some row operation and get $A'$ whose corresponding columns are $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$. Then it turns out that:

$$x\mathbf{a} + y\mathbf{b} + z\mathbf{c} = \mathbf{0} \qquad \text{if and only if} \qquad x\mathbf{a}' + y\mathbf{b}' + z\mathbf{c}' = \mathbf{0}$$

for some real numbers $x$, $y$, and $z$.

This also holds for bigger or smaller collections of columns. Why? Essentially, since we are preforming **row** operations, the relationships between **columns** are unaffected – we aren't "mixing" the columns together.

More precisely, if $A$ is row equivalent to $B$, then there exists an invertible matrix $P$ such that $PA = B$. A homogeneous relation among the columns can be written as $A\mathbf{c} = \mathbf{0}$ (multiplying by a vector gives a linear combination of columns). Notice that if $A\mathbf{c} = \mathbf{0}$ then $PA\mathbf{c} = P\mathbf{0}$ so that $B\mathbf{c} = \mathbf{0}$. Conversely if $B\mathbf{c} = \mathbf{0}$, then $P^{-1}B\mathbf{c} = P^{-1}\mathbf{0}$ so that $A\mathbf{c} = \mathbf{0}$. Therefore, the (homogeneous) relations among the columns of $A$ also hold among the columns of $B$ (and vice-versa).

## Example:

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & -2 \\ 2 & 1 & 3 \end{bmatrix} \underset{\sim}{R1 \leftrightarrow R2} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & -1 \\ 2 & 1 & 3 \end{bmatrix} \underset{\sim}{2R1+R3} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \underset{\sim}{-R2+R3} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \underset{\sim}{-1 \times R1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

In the RREF (on the far right) we have that the final column is 2 times the first column plus $-1$ times the second column.

$$2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

So this must be true for all of the other matrices as well. In particular, we have that the third column of the original matrix is 2 times the first column plus $-1$ times the second column:

$$2\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix}$$

**Why is this important?** Well, first, Gauss-Jordan elimination typically requires a lot of computation. This correspondence gives you a way to check that you have row-reduced correctly! In some cases, the linear relations among columns are *obvious* and we can just *write down the RREF without performing Gaussian elimination at all!* We will see other applications of this correspondence later in the course.

**Example:** Let $A$ be a 3x3 matrix whose RREF is $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. Suppose that we know the first column of $A$ is $\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}$ and the second is $\begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$. Then, by the linear correspondence, we know that the third column must be $-1 \cdot \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + 2 \cdot \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$. Therefore, the mystery matrix is $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

**Example:** Let $B$ be a matrix whose RREF is $\begin{bmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Suppose that we know the first **pivot** column (i.e. the second column) of $B$ is $\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ and the second pivot column (i.e. the fourth column) is $\begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}$. Then, by the linear correspondence, we know that the third and fifth columns must be...

$$2 \cdot \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix} \qquad \text{and} \qquad 3 \cdot \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} - 2 \cdot \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ -11 \\ 5 \end{bmatrix}$$

Therefore, the mystery matrix is $B = \begin{bmatrix} 0 & 2 & 4 & -1 & 8 \\ 0 & -1 & -2 & 4 & -11 \\ 0 & 3 & 6 & 2 & 5 \end{bmatrix}$.

## Homework Problems:

5. Verify that the linear correspondence holds between the each matrix, its REF, and its RREF in the previous homework problems and examples.

6. I just finished row reducing the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 1 & 5 & 0 \end{bmatrix}$ and got $\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. Something's wrong. Just using the linear correspondence explain how I know something's wrong and then find the real RREF (without row reducing).

7. A certain matrix has the RREF of $\begin{bmatrix} 1 & 2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. The first column of the matrix is $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ and the third column is $\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$. Find the matrix.

8. A certain matrix has the RREF of $\begin{bmatrix} 1 & 2 & 0 & 5 & -2 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. The second column of the matrix is $\begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix}$ and the last column is $\begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix}$. Find the matrix.

9. A certain matrix has the RREF of $\begin{bmatrix} 1 & -5 & 0 & 3 & 0 & -1 \\ 0 & 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$. The first pivot column of the matrix is $\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, the second pivot column is $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, and the third pivot column is $\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$. Find the matrix.

10. $2 \times 2$ matrices can row reduce to either $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, or $\begin{bmatrix} 1 & c \\ 0 & 0 \end{bmatrix}$ for some $c \in \mathbb{R}$. What can be said about the original matrices in each of these cases?

## A Few Answers:

5. If you'd like a particular example worked out, just ask me.

6. According to the linear correspondence, the final column of my matrix should be 2 times the first pivot column minus the second pivot column. However, $3 \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3-1 \\ 0-1 \\ 3-5 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$ which does not match the third column of my original matrix. So the RREF must be wrong.

   Let's find $a$ and $b$ so that $a$ times column 1 plus $b$ times column 2 is column 3. In the final column of the original matrix, the second entry is "$-1$". So we must have $b = -1$. Thus $a \cdot$(column 1) = (column 2) + (column 3). Adding those columns together, we get $\begin{bmatrix} 5 \\ 0 \\ 5 \end{bmatrix}$. This is 5 times the first column. So $a = 5$. Therefore, the correct RREF is $\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$.

7. According to the RREF, the second column is 2 times the first column, the fourth column is obtained by subtracting the first column from the third column (i.e. the second pivot column), and the last column is twice the first column plus 3 times the third column. This gives us $\begin{bmatrix} 2 & 4 & -1 & -3 & 1 \\ 0 & 0 & 1 & 1 & 3 \\ 1 & 2 & 2 & 1 & 8 \end{bmatrix}$.

8. According to the RREF, the second column is twice the first column. Thus the first column must be $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$. The last column is $-2$ times the first column plus the second pivot column (i.e. column 3).

   Thus the second pivot column must be $\begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} + 2 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 10 \\ 12 \end{bmatrix}$. And finally, the fourth column is 5 times the first column plus 3 times the third column. So we get $\begin{bmatrix} 1 & 2 & 6 & 23 & 4 \\ 2 & 4 & 8 & 34 & 4 \\ 3 & 6 & 10 & 45 & 4 \\ 4 & 8 & 12 & 56 & 4 \end{bmatrix}$.

9. The second column is $-5$ times the first, the fourth column is 3 times column 1 plus 2 times column 3 (the second pivot column), and the sixth column is $-1$ times the first column, 4 times the third column, and 2 times the fifth column (the third pivot column). Thus we get $\begin{bmatrix} 3 & -15 & 1 & 11 & 2 & 5 \\ -1 & 5 & -1 & -5 & 3 & 3 \\ 2 & -10 & 1 & 8 & 1 & 4 \end{bmatrix}$.

10. If your RREF is the zero matrix, you must have started with the zero matrix (the result of doing or undoing any row operation on a zero matrix yields a zero matrix). If your RREF is the identity matrix $\left( \text{i.e. } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$, then the linear correspondence guarantees that both columns of the original matrix are non-zero and they must not be multiples of each other. If your RREF is the third option, then the first column of your original matrix was zeros and the second column was not a column made entirely of zeros. Finally, in the last case, the second column of the original matrix must be $c$ times the first column (the second column is a multiple of the first column).