MATH 223: Gaussian Elimination.

Consider the following system of equations solved by elementary row operations.

$$\left.\left\{\begin{array}{cccccc} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ -4x_1 & +5x_3 & +9x_3 & = & -9 \end{array}\right.\right\}(1)$$

$$\Rightarrow \left.\left\{\begin{array}{cccccc} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ & -3x_2 & +13x_3 & = & -9 \end{array}\right.\right\}(2)$$

$$\Rightarrow \left.\left\{\begin{array}{cccccc} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ & & x_3 & = & 3 \end{array}\right.\right\}(3)$$

Because this is a triangular system, the system can (always) be solved by back substitution:

$$x_3 = 3, \quad x_2 = 16, \quad x_1 = 29$$

In matrix terms the system of equations is

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

The elementary row operation of adding 4 times row 1 to row 3 can be seen to be a linear transformation and so can be represented by a matrix.

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z+4x \end{bmatrix} = \begin{bmatrix} x & & \\ & y & \\ 4x & & +z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Note that this linear transformation has an inverse, namely subtracting 4 times row 1 from row 3 which has the matrix

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

We apply our elementary row operations to the original matrix and the right hand side.

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 3 \end{bmatrix}$$

Theorem 1 *Let A and $\mathbf{b}$ be given. Let B be an invertible matrix. Then*

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} = \{\mathbf{x} : BA\mathbf{x} = B\mathbf{b}\}$$

Proof: If $\mathbf{x}$ satisfies $A\mathbf{x} = \mathbf{b}$ then $\mathbf{x}$ satisfies $BA\mathbf{x} = B\mathbf{b}$ and hence

$$\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} \subseteq \{\mathbf{x} : BA\mathbf{x} = B\mathbf{b}\}$$

where $\subseteq$ denotes *subset*. Now if $\mathbf{x}$ satisfies $BA\mathbf{x} = B\mathbf{b}$ then we may multiply on the left by B^{-1} to obtain $A\mathbf{x} = B^{-1}BA\mathbf{x} = B^{-1}B\mathbf{b} = \mathbf{b}$. Thus

$$\{\mathbf{x} : BA\mathbf{x} = B\mathbf{b}\} \subseteq \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}.$$

This completes the proof.

Our elementary row operations can be seen to be invertible by considering them as linear transformations.

Let $D(i, \lambda)$ denote the operation of multiplying the ith row by the non zero constant λ. Then the inverse is multiplying row i by $1/\lambda$ namely $D(i, 1/\lambda)$.

$$D(2,17) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (D(2,17))^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/17 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Let $E(i, j)$ denote the matrix corresponding to interchanging row i and row j. Then its inverse is seen to be $E(i, j)$.

$$E(1,2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (E(1,2))^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = E(1,2)$$

Let $F(i, j, \lambda)$ denote the operation of adding λ times the ith row to row j. Then the inverse will be adding $-\lambda$ times the ith row to row j and hence $F(i, j, -\lambda)$.

$$F(2,3,5) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}, \quad (D(2,3,5))^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix} = F(2,3,-5)$$

Sample problem

$$\left\{ \begin{array}{cccccccc} 2x_1 & -2x_2 & & +2x_4 & +x_5 & & = & 0 \\ 4x_1 & -4x_2 & & +4x_4 & +3x_5 & +2x_6 & = & 2 \\ 2x_1 & -x_2 & +3x_3 & +4x_4 & +x_5 & +x_6 & = & 3 \\ 2x_1 & & +6x_3 & +6x_4 & +3x_5 & +6x_6 & = & 10 \end{array} \right\}$$

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 4 & -4 & 0 & 4 & 3 & 2 & 2 \\ 2 & -1 & 3 & 4 & 1 & 1 & 3 \\ 2 & 0 & 6 & 6 & 3 & 6 & 10 \end{array} \right]$$

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 1 & 3 & 2 & 0 & 1 & 3 \\ 0 & 2 & 6 & 4 & 2 & 6 & 10 \end{array} \right]$$

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 2 & 6 & 4 & 2 & 6 & 10 \end{array}\right]$$

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 2 & 4 & 4 \end{array}\right]$$

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We are now at echelon form or staircase pattern. You may but I do not encourage you to go on to reduced echelon form. Back in our original for as a set of equations we have:

$$\left\{\begin{array}{ccccccc} 2x_1 & -2x_2 & & +2x_4 & +x_5 & & = & 0 \\ & x_2 & +x_3 & +2x_4 & & +x_6 & = & 3 \\ & & & & x_5 & +2x_6 & = & 2 \end{array}\right\}$$

where we have discarded the equation $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 = 0$.

We have the *pivot variables* to be $\{x_1, x_2, x_5\}$ where the remaining variables $\{x_3, x_4, x_6\}$ are called *free variables* since if we choose any values at all for the three free variables , then because of the triangular form of the remaining system we can solve (uniquely) for the pivot variables.

$$\left\{\begin{array}{cccccccc} 2x_1 & -2x_2 & +x_5 & = & & & -2x_4 & \\ & x_2 & & = & 3 & -3x_3 & -2x_4 & -x_6 \\ & & x_5 & = & 2 & & & -2x_6 \end{array}\right\}$$

To avoid confusion (or perhaps create it) we introduce parameters s, t, u to be the values chosen for the free variables $\{x_3, x_4, x_6\}$.

$$x_3 = s, \quad x_4 = t, \quad x_6 = u$$

(this is the usual problem of distinguishing the variable from the value it takes)

$$\left\{\begin{array}{cccccccc} 2x_1 & -2x_2 & +x_5 & = & & & -2t & \\ & x_2 & & = & 3 & -3s & -2t & -u \\ & & x_5 & = & 2 & & & -2u \end{array}\right\}$$

Back substitute to solve:

$$\begin{array}{cccccc} x_5 & = & 2 & & & -2u \\ x_2 & = & 3 & -3s & -2t & -u \\ x_1 & = & 2 & -3s & -3t & \end{array}$$

We may write this as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2 & -3s & -3t & \\ 3 & -3s & -2t & -u \\ & s & & \\ & & t & \\ 2 & & & -2u \\ & & & u \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} + r \begin{bmatrix} -3 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

thus the set of solutions is

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ \boxed{0} \\ \boxed{0} \\ 2 \\ \boxed{0} \end{bmatrix} + r \begin{bmatrix} -3 \\ -3 \\ \boxed{1} \\ \boxed{0} \\ 0 \\ \boxed{0} \end{bmatrix} + s \begin{bmatrix} -3 \\ -2 \\ \boxed{0} \\ \boxed{1} \\ 0 \\ \boxed{0} \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ \boxed{0} \\ \boxed{0} \\ -2 \\ \boxed{1} \end{bmatrix} \;:\; r, s, t \in \mathbf{R} \right\}$$

This called vector parametric form (or parametric vector form) and gives a different way of looking at the solutions to $A\mathbf{x} = \mathbf{b}$. If we follow this procedure then we end up with the boxed entries as shown, 0's in the rows of the free variables in the first vector then an identity matrix in the rows of the free variables in the remaining vectors.

When we have written the set of solutions

$$\{\mathbf{x} \,:\, A\mathbf{x} = \mathbf{b}\} = \{\mathbf{w} + r\mathbf{v}_1 + s\mathbf{v}_2 + t\mathbf{v}_3, \quad r, s, t \in \mathbf{R}\}$$

we deduce that

$$A\mathbf{w} = \mathbf{b}$$

(using $r = s = t = 0$) and then we may deduce that

$$A\mathbf{v}_1 = \mathbf{0}, \quad A\mathbf{v}_2 = \mathbf{0}, \quad A\mathbf{v}_3 = \mathbf{0}$$

by noting that for example with $r = 1, s = t = 0$, we have $A(\mathbf{w}+\mathbf{v}_1) = \mathbf{b}$ and then $A(\mathbf{w}+\mathbf{v}_1) - A\mathbf{w} = \mathbf{b} - \mathbf{b} = \mathbf{0}$ but $A(\mathbf{w} + \mathbf{v}_1) - A\mathbf{w} = A\mathbf{w} + A\mathbf{v}_1 - A\mathbf{w} = A\mathbf{v}_1$.

This observation will provide a way to check your vector parametric form but you can also use it to solve for the vectors by substituting the entries as given by our boxes and solving for the rest of the vector. We solve for $\mathbf{w}$ by setting the free variables to 0 and the solving the triangular system. We solve for $\mathbf{v}_1$ by setting $x_3 = 1$, $x_4 = 0$, $x_6 = 0$ and then solving the triangular system WHERE YOU HAVE SET THE RIGHT HAND SIDE TO THE ZERO VECTOR. You do the same for v_2 and $\mathbf{v}_3$, again BY SETTING THE RIGHT HAND SIDE TO THE ZERO VECTOR. You might wish to try this alternate strategy for reading of the vector parametric form from the echelon form.

If we had changed the last equation to $2x_1 + 6x_3 + 6x_4 + 3x_5 + 6x_6 = 11$, then we would obtain the system

$$\left[\begin{array}{cccccc|c} 2 & -2 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

The last equation is now $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 = 1$ for which there is no solution. In this case we say the system of equations is *inconsistent.*

Theorem 2 *Let A be an $n \times n$ matrix. The following are equivalent.*
i) A^{-1} exists (we say A is invertible *or* nonsingular.
ii) $A\mathbf{x} = \mathbf{0}$ has only one solution, namely $\mathbf{x} = \mathbf{0}$.
iii) A can be transformed to a triangular matrix, with nonzeros on the main diagonal, by elementary row operations.
iv) A can be transformed to I by elementary row operations.

Proof: We can verify by Gaussian elimination that iii)⇒ iv)⇒ i)⇒ ii)⇒ iii), the last implication following because there can be no free variables (the system $A\mathbf{x} = \mathbf{0}$ is always consistent) and so elementary row operations must result in every variable being a pivot variable.

Finding A^{-1}

$$\text{Let } A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

To solve for A^{-1} we can solve

$$A\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad A\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad A\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and it makes sense to solve for all three vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ at the same time:

$$\begin{array}{c} \begin{array}{cc} \quad A & \quad I \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{array} \quad \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \end{array}$$

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad \begin{array}{c} \begin{array}{cc} \quad E_1A & \quad E_1 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{array} \quad \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{array}\right] \end{array}$$

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \begin{array}{c} \begin{array}{cc} \quad E_2E_1A & \quad E_2E_1 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{array} \quad \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{array}\right] \end{array}$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1/2 \end{bmatrix} \qquad \begin{array}{c} \begin{array}{cc} \quad E_3E_2E_1A & \quad E_3E_2E_1 \end{array} \\ \left[\begin{array}{ccc} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array} \quad \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & -1/2 \end{array}\right] \end{array}$$

$$E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}, E_5 = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{c} \begin{array}{cc} E_5E_4E_3E_2E_1A \qquad & E_5E_4E_3E_2E_1 \end{array} \\ \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1/2 & 1/2 & 1/2 \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{array}\right] \end{array}$$

Thus we have $(E_5E_4E_3E_2E_1)A = I$ and so $A^{-1} = E_5E_4E_3E_2E_1$ and $A = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}$. By the way, always check your work:

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