Solving Systems of Linear Equations

1) Consider the following system of equations:

$$x_1 + x_3 = 1$$
$$x_1 - x_2 = 0$$
$$2x_2 + x_3 = 2$$

or

$$x_1 + 0\,x_2 + x_3 = 1$$
$$x_1 - x_2 + 0x_3 = 0 \qquad \text{Why?}$$
$$0x_1 + 2x_2 + x_3 = 2$$

We will solve this system using a procedure, which will lend itself to a solution using matrices, which is called the Gauss-Jordan elimination method. But first, two systems of equations are called equivalent if they have the same (set of) solutions. We will see that the above system of equations is equivalent to, as the same solutions, as the system

$$1x_1 + 0\,x_2 + 0x_3 = 1$$
$$0x_1 + 1\,x_2 + 0x_3 = 1$$
$$0x_1 + 0\,x_2 + 1x_3 = 0$$

Therefore we can "read off" the solutions directly from the above system, namely

$$x_1 = 1$$
$$x_2 = 1$$
$$x_3 = 0$$

The reader should check by substitution into the original system that these are indeed the solutions.

The method of reducing any system of equations to a simpler system where we can more easily "read off" the solutions is based on three simple rules which apply to any system of equations. These rules we incorporate in to the following Theorem.

Theorem 1. (Elementary Operations on Equations) If any sequence of the following operations is performed on a system of equations, the resulting system is equivalent to (has the same solutions as) the original system:

a) Interchange any two equations in the system.
b) Multiply both sides of any equation by a nonzero constant.
c) Multiply both sides of any equation by a nonzero constant and add the result to a second equation in the system, with the sum replacing the latter equation.

We will now apply the above Theorem to the original system given above. The original system is:

Example 1.

$x_1 + x_3 = 1$
$x_1 - x_2 = 0$
$2x_2 + x_3 = 2$

In order to get a clearer idea how the procedure works we will insert the "missing terms" and number the equations to obtain:

(1) $x_1 + 0\,x_2 + x_3 = 1$
(2) $x_1 - x_2 + 0x_3 = 0$
(3) $0x_1 + 2x_2 + x_3 = 2$

Multiply both sides of equation (1) by –1 and add the result to equation (2) to obtain:

(1) $x_1 + 0\,x_2 + x_3 = 1$
(2) $0x_1 - x_2 - x_3 = -1$ **Note**: Equation (1) did not change.
(3) $0x_1 + 2x_2 + x_3 = 2$

Multiply both sides of equation (2) by –1 to obtain:

(1) $x_1 + 0\,x_2 + x_3 = 1$
(2) $0x_1 + x_2 + x_3 = 1$
(3) $0x_1 + 2x_2 + x_3 = 2$

Multiply both sides of equation (2) by –2 and add the result to equation (3) to obtain:

(1)$x_1 + 0\,x_2 + x_3 = 1$
(2) $0x_1 + x_2 + x_3 = 1$ **Note**: Equation (2) did not change.
(3) $0x_1 + 0x_2 - x_3 = 0$

Multiply both sides of equation (3) by –1 to obtain:

(1) $x_1 + 0\,x_2 + x_3 = 1$
(2) $0x_1 + x_2 + x_3 = 1$
(3) $0x_1 + 0x_2 + x_3 = 0$

Multiply both sides of equation (3) by –1 and add the result to equation (1) to obtain:

(1) $x_1 + 0\,x_2 + 0x_3 = 1$
(2) $0x_1 + x_2 + x_3 = 1$ **Note**: Equation (3) did not change.
(3) $0x_1 + 0x_2 + x_3 = 0$

Multiply both sides of equation (3) by –1 and add the result to equation (2) to obtain:

(1) $x_1 + 0\,x_2 + 0x_3 = 1$
(2) $0x_1 + x_2 + 0x_3 = 1$ **Note**: Equation (3) did not change.
(3) $0x_1 + 0x_2 + x_3 = 0$

Therefore the solution to the system is: $x_1 = 1$, $x_2 = 1$ and $x_3 = 0$.

If you think about the step-by-step changes in the above equivalent systems the changes from system to system is in the numbers involved, that is, in the coefficients of the x's and the constants. The only purpose that the variables serve is to ensure that is that of keeping the coefficients (and the constants) in the appropriate location. We can effect this using matrices. We will write the original system in matrix notation as:

$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 1 & 2 \end{bmatrix}$ or $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 1 & 2 \end{array}\right]$. The only purpose of the vertical line in the latter version of this matrix is to separate the coefficients of the system from the constants for easier readability. Both ways of writing the matrix of the system are used. Since the first **three** columns of this matrix are the coefficients of the given system of equations the matrix consisting of the first three columns is called the **coefficient matrix**. That is, the **coefficient matrix** is $\begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$. The "complete matrix" above, $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 1 & 2 \end{array}\right]$, is referred to as the **augmented matrix**. So one way of using the tool of matrices to solve systems of equations is to take Theorem 1 above and to replace the word equation by row and the word system by matrix, that is, another version of Theorem 1 is:

Theorem 1. (Elementary Row Operations) If any sequence of the following operations is performed on a matrix, the resulting matrix is equivalent to the original.

a) Interchange any two rows in the matrix.
b) Multiply any row of the matrix by a nonzero constant.
c) Multiply both sides of any row by a nonzero constant and add the result to a second row, with the sum replacing the latter row.

If we use the convention R_i to stand for row i of a matrix an the symbol $\longrightarrow$ to stand for row equivalent then $A \xrightarrow{cR_i+R_j} B$ means that the matrix B is obtained from the matrix A by multiplying the ith row of A by c and adding it to the jth row of A. Remember for our purposes here if two matrices are row equivalent then they represent equivalent systems of equations.

We now redo example 1 using matrices.

Example 1 revisited.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 2 & 1 & 2 \end{array}\right] \xrightarrow{(-1)R_1+R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & -1 & -1 & -1 \\ 0 & 2 & 1 & 2 \end{array}\right]$$ **Note**: Row 1 (R_1) did not change.

$$\xrightarrow{(-1)R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 2 \end{array}\right]$$

$$\xrightarrow{(-2)R_2+R_3} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right]$$ **Note**: Row 2 (R_2) did not change

$$\xrightarrow{(-1)R_3} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{(-1)R_3+R_1} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$ **Note**: Row 3 (R_3) does not change

$$\xrightarrow{(-1)R_3+R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Note, one may prefer to use a different sequence of steps in solving the above.

Example 2. Solve the system. Recall that this means that we want to find all real numbers x_1, x_2, and x_3 which will satisfy each equation in the system.

$$\begin{aligned} 4x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + \ \ x_2 + x_3 &= 4 \\ 2x_1 + 2x_2 + x_3 &= 3 \end{aligned}$$

The (augmented) matrix of the system is:

$\begin{bmatrix} 4 & 2 & 1 & 1 \\ 2 & 1 & 1 & 4 \\ 2 & 2 & 1 & 3 \end{bmatrix}$ or if you prefer inserting the vertical line $\left[\begin{array}{ccc|c} 4 & 2 & 1 & 1 \\ 2 & 1 & 1 & 4 \\ 2 & 2 & 1 & 3 \end{array}\right]$.

$$\begin{bmatrix} 4 & 2 & 1 & 1 \\ 2 & 1 & 1 & 4 \\ 2 & 2 & 1 & 3 \end{bmatrix} \xrightarrow{\frac{1}{4} R_1} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & 1 & 1 & 4 \\ 2 & 2 & 1 & 3 \end{bmatrix}$$

$$\xrightarrow{-2 R_1 + R_2} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & 7/2 \\ 2 & 2 & 1 & 3 \end{bmatrix}$$

$$\xrightarrow{-2 R_1 + R_3} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & 7/2 \\ 0 & 1 & \frac{1}{2} & 5/2 \end{bmatrix}$$

$$\xrightarrow[\text{}]{\substack{\text{Interchange} \\ R_2 \text{ and } R_3}} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & \frac{1}{2} & 5/2 \\ 0 & 0 & \frac{1}{2} & 7/2 \end{bmatrix}$$

$$\xrightarrow{-\frac{1}{2} R_2 + R_1} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & 5/2 \\ 0 & 0 & \frac{1}{2} & 7/2 \end{bmatrix}$$

$$\xrightarrow{2 R_3} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & 5/2 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\xrightarrow{-\frac{1}{2} R_3 + R_2} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$

We now write in the variables and the equality symbols to obtain the system:

$$x_1 + 0x_2 + 0x_3 = -1$$
$$0x_1 + x_2 + 0x_3 = -1$$
$$0x_1 + 0x_2 + x_3 = 7$$

and read off the solution to the original system as $x_1 = -1$, $x_2 = -1$ and $x_3 = 7$.

I encourage the reader to substitute these values in the system to verify that they are indeed the solutions to the given system of equations.

Remark

Each system of equations will (usually) have its own set of solutions. The purpose of exercises 3 & 4 of the notes is to show that since the coefficient matrix of exercise 3 and that of 4 are the same the same elementary row operations could be used to solve each system. So if we were given 2 systems with the same coefficient matrix instead of solving them separately we could save time by solving them together by augmenting the coefficient by not one but 2 columns. Then use the usual process to row reduce the matrix. If all goes well the numbers in the first (added) column become the solution of the first system and those in the second (added) column become the solutions of the second system. Exercise 5 is an example where you can do this. One intent of the discussion is to lead people to thinking about what exercise 6 means and eventually to why the method of finding the inverse of a matrix in the next set of notes works.

So the matrix of problem six is really the matrix for solving 3 systems of 3 equations and 3 unknowns.

A key part of the definition of the inverse of a matrix A is to find a matrix B such that AB = I. If A is the 3x3 coefficient matrix given in problem 6, and if B (of the definition of inverse) is a 3x3 matrix of variables (since we are looking for B). Then AB = I becomes 3 systems of 3 equations and 3 unknowns, all with the same coefficient matrix. So we can solve all 3 systems simultaneously. The matrix of the 3 systems is that of problem 6. So if you solve problem 6 you are really finding the inverse (of the coefficient matrix).