

FINDING SOLUTIONS TO SYSTEMS OF LINEAR EQUATIONS

In First Year students learn to first represent a given system as an augmented matrix, reduce that matrix to RE (Row Echelon) form and then after that if there are solutions most students write the equivalent system of equations for the RE matrix and continue in the way they did before they knew about matrices. The trouble with this is, most students fail to appreciate the benefits of the matrix approach and abandon it once they have finished the course, reverting to the traditional somewhat *ad hoc* approach.

So ... here we describe how we can continue further with the matrix approach to a point where we can essentially write the solution down. The *algorithm* is described and applied to the immediately following problem, below (description – on the left; application to problem – on the right).

Problem. Find the solutions to the following linear system:

$$\begin{aligned}x_1 + 3x_2 - 3x_3 + 2x_4 &= -13 \\2x_1 + 6x_2 + 2x_3 + 4x_4 &= 14 \\3x_1 + 9x_2 - x_3 + 6x_4 &= 1\end{aligned}$$

Algorithm.

1. Represent the system of linear equations as an augmented matrix.

Let the augmented matrix of the system be A . Then A is

$$\left[\begin{array}{cccc|c} 1 & 3 & -3 & 2 & -13 \\ 2 & 6 & 2 & 4 & 14 \\ 3 & 9 & -1 & 6 & 1 \end{array} \right]$$

2. Reduce the augmented matrix to RE form.

$$\begin{aligned} &\sim \left[\begin{array}{cccc|c} 1 & 3 & -3 & 2 & -13 \\ 0 & 0 & 8 & 0 & 40 \\ 0 & 0 & 8 & 0 & 40 \end{array} \right] \\ &\sim \left[\begin{array}{cccc|c} 1 & 3 & -3 & 2 & -13 \\ 0 & 0 & 8 & 0 & 40 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

3. Decide whether the system has solutions. That is, does the RE matrix have a row of form

The system has no rows of form

$$[00 \dots 0 | \neq 0]$$

$$[00 \dots 0 | \neq 0]$$

So the system is *consistent* and we continue to the *next* step.

(all zeros except for the right-hand side entry)? ... if so, we conclude: *the system is inconsistent* (i.e. has no solutions) and we STOP. Otherwise we continue to the *next* step.

4. Reduce the RE matrix to RRE (Reduced Row Echelon) form. That is, the leading entry of each row should be 1; and above each leading 1 should be all zeros.

$$\begin{aligned} A &\sim \left[\begin{array}{cccc|c} 1 & 3 & -3 & 2 & -13 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \\ &\sim \left[\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

5. Rearrange the RRE matrix (by inserting all-zero rows, if necessary) in such a way that all the leading 1s lie on the diagonal of the coefficient matrix.

$$A \sim \left[\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

6. Change all the 0s (if there are any) that appear on the diagonal of the coefficient matrix to a * (just to signify these are in some way special).

$$\sim \left[\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 2 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & * & 0 \end{array} \right]$$

7. Now we are ready to write the general solution down. For this we prefer to write our vectors as *columns*. Identify the columns of the final matrix that have a *. Let's call these vectors: *-vectors, and label them: $\mathbf{u}_1^*, \dots, \mathbf{u}_k^*$. Let's call the right-most column of the final matrix \mathbf{v} . Then the final solution is:

$$\mathbf{x} = \mathbf{v} + \lambda_1(-\mathbf{u}_1^*) + \dots + \lambda_k(-\mathbf{u}_k^*)$$

except that each $-*$ we replace by a 1. (In practice, in writing this solution down we negate each entry of a *-vector, except the *, and replace the * with a 1 in a single step ... rather than negate the * and replace $-*$ with 1. If that confuses you ... look at the example.)

So the general solution of the system is:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 5 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

for $\lambda, \mu \in \mathbb{R}$.

FINAL COMMENTS.

In the final solution of our example we chose parameters λ, μ rather than λ_1, λ_2 as would have been our choice if we had followed Step 7. to the letter.

Why does the algorithm work? If we follow the traditional approach we identify that x_2, x_4 are non-step variables (there is no leading 1 in the 2nd and 4th columns of the RRE matrix at Step 4.) ... and so we should choose these as free parameters λ, μ , say. Translating the non-zero rows of the RRE matrix at Step 4. to equations we get

$$\begin{aligned} x_1 + 3x_2 + 2x_4 &= 2 \\ x_3 &= 5 \end{aligned}$$

Now we replace x_2, x_4 by λ, μ , respectively, and rearrange the equations so that x_1, x_3 are the subject. If you look carefully at the *magic* operations done in Step 7. this is precisely what is accomplished; in particular, the * replacement operation recovers (reading the 2nd and 4th rows of the final vector solution)

$$\begin{aligned} x_2 &= 0 + \lambda.1 + \mu.0 \\ x_4 &= 0 + \lambda.0 + \mu.1 \end{aligned}$$

i.e. $x_2 = \lambda$ and $x_4 = \mu$. The sign changes occur because terms are being swapped to the right-hand side in order to make x_1, x_3 the subject.

Of course, if the system has a *unique* solution then we would not need to introduce *s and Steps 5. to 7. become trivial operations, and if the system is *inconsistent* we stop at Step 3.