Solving Equations: REF and RREF example

Our text, and my lecture the other day, discuss how to apply elementary row operations, one after another, to take any matrix and change it into a matrix in either Row Echelon Form or Reduced Row Echelon Form, and those in turn are supposed to let us find the solutions to a set of equations corresponding to the matrix. (Really to any of the matrices along the way, since the elementary row operations are chosen so that they don't change the set of solutions.) It is possible to write all of the steps out in general, with phrases like "if $a_{ij} \neq 0$ and $a_{kl} = 0$ for all $l < j$" to say a_{ij} is in the first non-zero column. But the description in that form gets to be so complicated that it is hard to see exactly what it saying. Most people find all of this much easier to follow by reading examples. The text has many examples. Here is another, with my comments.

(If you are at all interested in computing, you might want to think about the two processes (a) converting a matrix to REF or RREF and (b) reading off equation solutions from the converted matrix and see if you could write code to carry them out. Almost any computational language works pretty well, although older ones like Fortran and C are actually easier to use for this than some object-oriented ones like Java. Don't try it with a language designed for AI, like LISP! I once had to read code for doing this in LISP and I still have not recovered...)

For our example we start with the matrix

$$A = \left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ -1 & 1 & 0 & 3 & 0 & -4 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ -1 & 2 & 1 & 3 & 0 & -3 \end{array}\right].$$

(This is the same matrix I used in lecture on Monday 1/31, except that I have written in the vertical bar we would use for an augmented matrix. That does not change any of the arithmetic, and sets us up to consider an associated set of equations after we get to Row Echelon Form.)

We want to get this to Row Echelon Form (REF) and eventually to Reduced Row Echelon Form (RREF). I note that the upper left entry is already a 1, so that can be the leading entry in the first row. But property (b) of REF, that the leading entries move to the right as we go down the rows, implies that all the entries below a leading entry must be 0. So we proceed to "clear out" the numbers below that 1. We can add the first row to the second, a type III elementary row operation, and that will make $a_{21} = 0$.

Adding 1 times the first row to the second row we get

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 2 & 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ -1 & 2 & 1 & 3 & 0 & -3 \end{array}\right].$$

We want to clear out the rest of the first column, making the -2 in the lower left corner into a 0. We can do that by adding 2 times the first row to the last row, and we get

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 2 & 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 3 & 2 & 1 & 1 & 5 \end{array}\right].$$

Now one of the nice things about proceeding in a somewhat orderly fashion is that we have the first column set up so that the row operations we need to perform to get to REF (or RREF) won't mess up that leading 1: We won't need to swap the first row with any other, we won't need to multiply the first row by a

constant, and adding any multiple of some other row to the first would only add 0 to the 1 because of all those 0's below the 1, so it would not change it.

So we move over and work to get the second column "right". If we were going to try to get to RREF directly we would want to get a 0 in position a_{12} at the top of the column, but I am going to go first to REF and later to RREF so we don't need that. What we do want is to get a 1 as the leading entry in the second row, then make all the entries below it into 0's. We see that the 2 in a_{22} is right now in the Leading Entry position, the first non-zero entry in the second row, and we need to get a 1 there. We could multiply the second row by $\frac{1}{2}$, which is non-zero so that is a legal elementary row operation. I choose instead to swap the second and third rows, just to avoid working with fractions when possible. Swapping those rows we get

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 1 & 1 & 4 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 3 & 2 & 1 & 1 & 5 \end{array}\right].$$

To make the entries a_{32}, a_{42}, and a_{52} zero, we subtract multiples of the second row from the third, fourth, and fifth rows: "Officially" we do one of those, then the next, then the last, so that what we do fits the definition of an <u>elementary</u> row operation. But we frequently don't rewrite the result after each one, instead showing the result after several operations are performed. For the first one we want to subtract twice the second row from the third row: Again this is not the official way an elementary row operation is defined, but it is easier to say (or type) "subtract" than to say we add -1 times the second row to the third. No matter how you say it, when we subtract twice the second row from the third, subtract the second row from the fourth, and subtract 3 times the second row from the fifth row, we get

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 1 & 1 & 2 \end{array}\right].$$

Now isn't that surprising! Those three rows coming out the same is not something you would likely have guessed from looking at the original matrix! But it shows there was a lot of interrelationship between the rows, or equations, and will turn out to say something about the solutions to corresponding equations.

Now we move to the third column. I want a leading entry 1 in the row, at the first non-zero place, i.e. a_{33}. Now it is -1, so we can just multiply the third row by -1 and get

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & -1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 1 & 1 & 2 \end{array}\right].$$

We proceed to "clear out" the lower elements in the third column, by adding the third row to each of the fourth and fifth rows, getting

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

That is in Row Echelon Form. But it is not in Reduced Row Echelon Form, since in each of the second and third columns, we don't have <u>all</u> of the column zero except for the leading entries. We could find the

2

solutions to the corresponding equations from this matrix: Solving this way is called Gaussian Elimination. But I will go on to get to Reduced Row Echelon Form before writing out the solutions, a process called Gauss-Jordan Elimination, so that we can compare what has to be done to read out the solutions. So that we can come back to this one later, let's call it A_R.

As noted, to get to RREF we need to clear out the upper parts of columns two and three. I usually do that "from the bottom up". I will first subtract the third row from the second, getting

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

and then subtract the third row from the first, and the second row from the first, which I will combine here to get

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -2 & 1 & 7 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

which is Reduced Row Echelon Form. I will call it A_{RR}.

Now we want to read off solutions to associated equations. Since at each stage our matrices changed by (combinations of) elementary row operations, the solutions should be the same whether I used equations corresponding to the first matrix, the second, or any other. In a typical application the equations we care about came first, so they would be what goes with the original matrix A. Those are

$$\begin{aligned} x_1 + x_2 + x_3 - 2x_4 + x_5 &= 8 \\ -x_1 + x_2 + 0x_3 + 0x_4 + 0x_5 &= -4 \\ 0x_1 + x_2 + x_3 + 0x_4 + 0x_5 &= 1 \\ 0x_1 + x_2 + 0x_3 + x_4 + x_5 &= 3 \\ -x_1 + 2x_2 + x_3 + 3x_4 + 0x_5 &= -3. \end{aligned}$$

So what we would want are the solutions to those equations. But the whole point to this process is that we can read off the solutions more easily from the version corresponding to A_R or A_{RR}. I'll use $A_{R}R$ first, it is easier to read off the solutions from Reduced Row Echelon Form (i.e. to apply the Gauss-Jordan process). That matrix was

$$A_{RR} = \left[\begin{array}{ccccc|c} 1 & 0 & 0 & -2 & 1 & 7 \\ 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

and the corresponding equations would be $x_1 - 2x_4 + x_5 = 7$, $x_2 + x_4 + x_5 = 3$, and $x_3 - x_4 - x_5 = -2$. (We ignore equations that say multiplying all the variables by zero and adding gives zero: Those equations contribute nothing to finding solutions, since any numbers at all could be substituted for the variables without changing the truth of the equations.) We see that each of those three equations is easily transformed into an equation that represents one of x_1, x_2, or x_3 as a combination of x_4 and x_5 and a constant: $x_1 = 7 + 2x_4 - x_5$, $x_2 = 3 - x_4 - x_5$, and $x_3 = -2 + x_4 + x_5$. We can give arbitrary values to x_4 and x_5 and use those three equations to determine values of x_1, x_2, and x_3 which go with the numbers we chose for x_4 and x_5 to make a solution! Also, any solution must be of this form: If a, b, c, d, e are five numbers

3

that fit the equations (with $x_1 = a$, etc.), then d and e are some numbers being used for x_4 and x_5: x_1 has to be $7 + 2d - e = 7 + 2x_4 - x_5$ since the numbers had to fit the first equation, and similarly for x_2 and x_3.

Note that the variables x_4 and x_5 which were given arbitrary values matched up with the columns of the matrix (to the left of the dividing bar) that don't contain leading entries, while the variables determined from those arbitrary values are the ones going with columns that do have leading entries. We could formally write all solutions in vector form as

$$\begin{pmatrix} 7 + 2\alpha - \beta \\ 3 - \alpha - \beta \\ -2 + \alpha + \beta \\ \alpha \\ \beta \end{pmatrix} \quad \text{for any numbers } \alpha \text{ and } \beta.$$

Now let's see how we could have got the same solutions out of the REF matrix,

$$A_R = \begin{bmatrix} 1 & 1 & 1 & -2 & 1 & 8 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The process is called "back substitution", and we start from the bottom (skipping the zero rows) with the row

$$0 \quad 0 \quad 1 \quad -1 \quad -1 \mid -2.$$

Interpreted as an equation, and moving things around a bit, that says $x_3 = -2 + x_4 + x_5$. That is the same information we read out of A_{RR}: The third rows of A_R and A_{RR} are the same, since the difference between Reduced REF and just REF only specifies what happens <u>above</u> a leading entry. Now we move up a row, to

$$0 \quad 1 \quad 1 \quad 0 \quad 0 \mid 1$$

Writing that as an equation, and rearranging the terms, that says $x_2 = 1 - x_3$. That is certainly not the same as what we had before, but using $x_3 = -2 + x_4 + x_5$ that we got above, we have $x_2 = 1 - (-2 + x_4 + x_5) = 3 - x_4 - x_5$, which is exactly what we got before. Continuing upward we get to row 1,

$$1 \quad 1 \quad 1 \quad -2 \quad 1 \mid 8$$

Again we interpret it as an equation, $x_1 = 8 - x_2 - x_3 + 2x_4 - x_5$. Substituting in $x_2 = 1 - (-2 + x_4 + x_5) = 3 - x_4 - x_5$ and $x_3 = -2 + x_4 + x_5$ we get $x_1 = 8 - (3 - x_4 - x_5) - (-2 + x_4 + x_5) + 2x_4 - x_5$, and simplifying we have $x_1 = 7 + 2x_4 - x_5$, which is the same as we got from A_{RR}. Again we can let x_4 and x_5 take on arbitrary values and determine x_1, x_2, and x_3 from them, and this process, working from the (non Reduced) Row Echelon Form A_R, got the same answers as before. There was more work to do after getting the matrix, but less work getting the matrix. The "clearing out" above leading entries gets us to the point where the variables we did not give arbitrary values (in this case x_1, x_2, and x_3) do not interact.

<u>NOTE:</u> For this example, the leading entries were in the first three columns, and the variables given arbitrary values were the last two. The definition of "leading entry" tends to make them appear in early columns, and since arbitrary values are given to variables not going with leading entries they go with the remaining, hence typically later, columns. But the separation does not have to be this complete! See more examples in the textbook.