MATH 304 - Linear Algebra

We have seen that matrices form a complex structure with addition and multiplication which, on one hand, have many of the familiar properties of addition and multiplication of numbers, but, on the other hand, multiplication shows some significantly different behavior (like it is not commutative, two non-zero matrices can multiply to a zero matrix, etc.). Nevertheless, we can compute with matrices and, as in the case with numbers, this often leads to equations with unknown matrix (or matrices). Perhaps one of the simplest such equations (which involves multiplication) is an equation of the form $AX = B$. Here A, B are given matrices of size $k \times m$ and $k \times n$ respectively and X is an unknown matrix of size $m \times n$. When A is a square invertible matrix then we can solve such equations easily by just multiplying both sides on the left by A^{-1}: $A^{-1}(AX) = A^{-1}B$. Since multiplication is associative, we have $A^{-1}(AX) = (A^{-1}A)X = X$, so $X = A^{-1}B$ is the unique solution. But when A is not invertible, we need a different method.

Let us look at an example. Let

$$A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & 2 & -1 \\ 1 & 2 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 3 & 0 & -1 & 4 \\ 6 & 5 & 3 & 8 \\ 6 & 3 & 1 & 8 \end{bmatrix}, \ X = \begin{bmatrix} x_{1,1} & x_{1.2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2.2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3.2} & x_{3,3} & x_{3,4} \end{bmatrix}.$$

We want to solve the equation $AX = B$. Even though A in this example is a square matrix, it is not invertible. The most naive way of approaching this problem would be to perform the multiplication AX and then make each entry equal to the corresponding entry in B. For example, the $1,3$-entry of AX is $-x_{1,3} + x_{2,3} + 2x_{3,3}$ so we get equation $-x_{1,2} + x_{2,2} + 2x_{3,2} = -1$. Doing this for all 12 entries will result is a system of 12 equations with 12 unknowns. The good news is that this is a system of linear equations, so we have tools to solve it. The bad news is that we would need to work with 12×12 matrices. We can simplify this approach by observing that this system splits into 4 systems, each having three equations and 3 unknowns (the unknowns in each system being independent of each other). This follows from the fact that the entries in the first column of AX only involve the first column of X, and the same applies to other columns of X. In our example, the 4 systems are:

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

$$\begin{bmatrix} -1 & 1 & 2 \\ 3 & 2 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_{1,3} \\ x_{2,3} \\ x_{3,3} \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} -1 & 1 & 2 \\ 3 & 2 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_{1,4} \\ x_{2,4} \\ x_{3,4} \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 8 \end{bmatrix}.$$

At this point it looks like even though we reduced the size of the systems, but now we have to solve four such systems, which may still require quite a bit of labor. Final simplification comes from the observation that all 4 systems have the same coefficient matrix. Recall that our method of solving systems of linear equations is to form the augmented matrix and then perform elementary row operations to transform the coefficient part into reduced row-echelon form. Since all 4 systems have the same coefficient parts, the elementary row operations involved will be the same for each system. We can organize the computations through a single 3×7 matrix as follows.

We form a "combined augmented matrix": $\left[\begin{array}{rrr|rrrr} -1 & 1 & 2 & 3 & 0 & -1 & 4 \\ 3 & 2 & -1 & 6 & 5 & 3 & 8 \\ 1 & 2 & 1 & 6 & 3 & 1 & 8 \end{array}\right]$. If we focus only on the i-th column to the right of the dividing line, we are working with the i-th system of equations, i.e. solving for the i-th column of X. We perform $E_{2,1}(3)$, $E_{3,1}(1)$ and get $\left[\begin{array}{rrr|rrrr} -1 & 1 & 2 & 3 & 0 & -1 & 4 \\ 0 & 5 & 5 & 15 & 5 & 0 & 20 \\ 0 & 3 & 3 & 9 & 3 & 0 & 12 \end{array}\right]$. Next we do $D_1(-1), D_2(1/5), D_3(1/3)$ and get $\left[\begin{array}{rrr|rrrr} 1 & -1 & -2 & -3 & 0 & 1 & -4 \\ 0 & 1 & 1 & 3 & 1 & 0 & 4 \\ 0 & 1 & 1 & 3 & 1 & 0 & 4 \end{array}\right]$. Then we do $E_{3,2}(-1)$, $E_{1,2}(1)$ and get $\left[\begin{array}{rrr|rrrr} 1 & 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 3 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$. The part to the left of the dividing line is now in the reduced row-echelon form.

If we had pivot columns to the right of the dividing line, one of the four system would be inconsistent and we would conclude that there are no solutions to our original matrix equation. In our example though there are no pivot columns to the right of the dividing line. This means that all 4 systems are consistent and now we need to read the solutions for each of them. Note that the last column of the left part is not a pivot column and the other two columns are pivot columns. This means that the third unknown in each system will be a free variable and the first two will be dependent variables expressed in terms of the free variable. Working with the first column to the right of the dividing line we see that

$$x_{3,1} = a \text{ is a free variable and } x_{2,1} = 3 - a,\ x_{1,1} = a.$$

Similarly, working with the second column to the right of the dividing line we get

$$x_{3,2} = b \text{ is a free variable and } x_{2,2} = 1 - b,\ x_{1,2} = 1 + b$$

From the third column we get

$$x_{3,3} = c \text{ is a free variable and } x_{2,3} = -c,\ x_{1,3} = 1 + c,$$

and from the fourth column we get

$$x_{3,4} = d \text{ is a free variable and } x_{2,4} = 4 - d\ x_{1,4} = d.$$

Summarizing the above computations, the equation $AX = B$ in the example has infinitely many solutions, depending on 4 parameters a, b, c, d, as follows:

$$X = \begin{bmatrix} a & 1+b & 1+c & d \\ 3-a & 1-b & -c & 4-d \\ a & b & c & d \end{bmatrix}.$$

We can choose any values for the parameters a, b, c, d and get a concrete solution. For example, when $a = 1,\ b = 0,\ c = 0,\ d = 1$, we get $X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 3 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ as one of the possible solutions.

The method outlined in the above example works in general. To solve $AX = B$, we form the "combined augmented matrix" $[A|B]$ and we perform elementary row operations to transform the part to the left of the dividing line into reduced row-echelon form. If in the resulting matrix there is a pivot column to the right of the dividing line, there are no solutions. Otherwise, we read the solutions column by column.

Note that our method of finding the inverse of a matrix is a special case of the above method. It should not come as a surprise, as inverting the matrix A is the same as solving the equation $AX = I$.

A natural questions which may come to mind at this point is: what about equations of the form $XA = B$, where A, B are given matrices of size $m \times n$ and $k \times n$ respectively and X is unknown matrix of the size $k \times m$? We can easily reduce it to the case discussed above by using a tool we learned in the previous note. Namely, we can transpose both side of the equation to get a new equation $(XA)^T = B^T$, i.e. $A^T X^T = B^T$. Now we can use the method we just learned to solve for X^T and then transpose the solution to get X.

Let us consider an example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & -1 \end{bmatrix},\ B = \begin{bmatrix} 7 & 18 & 5 \\ 12 & 31 & 8 \\ 5 & 13 & 3 \end{bmatrix},\ X = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{bmatrix}.$$

In order to solve the equation $XA = B$, we first solve the equation $A^T X^T = B^T$. Here

$$A^T = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & -1 \end{bmatrix},\ B^T = \begin{bmatrix} 7 & 12 & 5 \\ 18 & 31 & 13 \\ 5 & 8 & 3 \end{bmatrix}$$

and the "combined augmented matrix" is $\left[\begin{array}{cc:ccc} 1 & 1 & 7 & 12 & 5 \\ 2 & 3 & 18 & 31 & 13 \\ 3 & -1 & 5 & 8 & 3 \end{array}\right]$. We do $E_{2,1}(-2)$, $E_{3,1}(-3)$ and get $\left[\begin{array}{cc:ccc} 1 & 1 & 7 & 12 & 5 \\ 0 & 1 & 4 & 7 & 3 \\ 0 & -4 & -16 & -28 & -12 \end{array}\right]$. Next do $E_{3,2}(4)$, $E_{1,2}(-1)$ to get $\left[\begin{array}{cc:ccc} 1 & 0 & 3 & 5 & 2 \\ 0 & 1 & 4 & 7 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$. This matrix is in reduced row echelon form and no column to the right of the dividing line is a pivot column, so there are solutions. Moreover, we have no free variables (all columns to the left of the dividing line are pivot), so the solution is actually unique: the first column of X^T (written as a row) is $3, 4$, the second column is $5, 7$ and the third is $2, 3$. In other words, $X^T = \begin{bmatrix} 3 & 5 & 2 \\ 4 & 7 & 3 \end{bmatrix}$. It follows that

$$X = (X^T)^T = \begin{bmatrix} 3 & 4 \\ 5 & 7 \\ 2 & 3 \end{bmatrix}.$$

Using the correspondence between linear transformations and matrices we can restate the two matrix equations considered above into problems about linear transformations.
The equation $AX = B$ is equivalent to the following question:

Given linear transformations $L_B : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and $L_A : \mathbb{R}^m \longrightarrow \mathbb{R}^k$, is there a linear transformation $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that $L_A \circ L = L_B$?
Any L satisfying this condition will be of the form L_X where $AX = B$.

Similarly, The equation $XA = B$ is equivalent to the following question:

Given linear transformations $L_B : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, is there a linear transformation $L : \mathbb{R}^m \longrightarrow \mathbb{R}^k$ such that $L \circ L_A = L_B$?
Any L satisfying this condition will be of the form L_X where $XA = B$.