Matrices and Linear Systems

A *matrix* is a rectangular array of numbers, usually real or complex numbers, aligned along its rows and columns. Examples include

$$(1)\quad \begin{pmatrix} 1 & 2 \\ 4 & 3 \\ 7 & 6 \end{pmatrix} \qquad (2)\quad \begin{pmatrix} 9 \\ 8 \\ 1 \\ 2 \end{pmatrix} \qquad (3)\quad \begin{pmatrix} 2 & 1 & 3 \\ 7 & 2 & 4 \end{pmatrix}.$$

I'll use round brackets, i.e. $\big($ and $\big)$, to delineate matrix elements. Some use square brackets, i.e. $\big[$ and $\big]$, but I'll use square brackets for other purposes.

The *size* of a matrix is its number of rows by its number of columns. The matrix in (1) has 3 rows by 2 columns, or simply 3×2 (rows $\times$ columns). The matrix in (2) is a 4×1 matrix, and the matrix in (3) is 2×3. It's very common to refer to the class of matrices of size $m \times n$ which have real numbers as elements by $\mathbb{R}^{m\times n}$. For example, the matrix (1) is in $\mathbb{R}^{3\times 2}$. The class of $m \times n$ matrices with complex elements will be denoted by $\mathbb{C}^{m\times n}$.

I'll use Roman letters (often upper case but sometimes lower) to signify a given matrix and the equal sign for matrix assignment,

$$A = \begin{pmatrix} -1 & 3 & 2 \\ 5 & -4 & 7 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

The double subscript $A_{i,j}$ will be used to denote the particular element (i.e. a number) of matrix A located at row i, column j. For example, for the matrix A just defined, $A_{2,1} = 5$ and $A_{2,2} = -4$.

Matrix addition is only defined between two matrices which have the same size,

$$\begin{aligned} & A \in \mathbb{R}^{m\times n},\ B \in \mathbb{R}^{m\times n} \quad (\text{or } \mathbb{C}^{m\times n}) \\ \Rightarrow\quad & C = A + B \in \mathbb{R}^{m\times n} \quad (\text{or } \mathbb{C}^{m\times n}) \\ & C_{i,j} = A_{i,j} + B_{i,j} \quad \text{for each } 1 \le i \le m \text{ and } 1 \le j \le n. \end{aligned}$$

For example

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \\ 5 & 1 \end{pmatrix},\ B = \begin{pmatrix} 2 & 1 \\ 0 & 4 \\ 2 & 3 \end{pmatrix} \quad \Rightarrow \quad C = A + B = \begin{pmatrix} 1+2 & 2+1 \\ 4+0 & 3+4 \\ 5+2 & 1+3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 4 & 7 \\ 7 & 4 \end{pmatrix}.$$

Scalar matrix multiplication is defined between any scalar and matrix,

$$\begin{aligned} & \alpha \in \mathbb{R} \quad (\text{or } \mathbb{C}),\ A \in \mathbb{R}^{m\times n} \quad (\text{or } \mathbb{C}^{m\times n}) \\ \Rightarrow\quad & C = \alpha\, A \in \mathbb{R}^{m\times n} \quad (\text{or } \mathbb{C}^{m\times n}) \\ & C_{i,j} = \alpha\, A_{i,j} \quad \text{for each } 1 \le i \le m \text{ and } 1 \le j \le n. \end{aligned}$$

For example

$$\alpha = 2,\ A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 4 & 3 \\ 2 & 1 \end{pmatrix} \quad \Rightarrow \quad C = \alpha\, A = \begin{pmatrix} 2\cdot 1 & 2\cdot 2 \\ 2\cdot 3 & 2\cdot 4 \\ 2\cdot 4 & 2\cdot 3 \\ 2\cdot 2 & 2\cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 8 & 6 \\ 4 & 2 \end{pmatrix}.$$

From now on I'm going to drop saying the $\mathbb{C}$ stuff. Everything generalizes in the obvious way from $\mathbb{R}$ to $\mathbb{C}$.

Matrix multiplication is only defined between two matrices which have complimentary size,

$$A \in \mathbb{R}^{m\times l},\ B \in \mathbb{R}^{l\times n}$$

!!! (note that A has the same number of columns as B has rows)

$$\Rightarrow \quad C = A\,B \in \mathbb{R}^{m\times n} \ \text{ where } C\text{'s elements are } \ C_{i,j} = \sum_{k=1}^{l} A_{i,k} B_{k,j}.$$

The Σ notation defining matrix multiplication above might be confusing to some. All it says is the i,jth element of the product C is found by *dotting* the ith row of A by the jth column of B. I'll say more about this in class. For example, consider

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 3},\ B = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \in \mathbb{R}^{3\times 1}.$$

For these, the product AB is defined (since A's columns $=$ B's rows), and the result has size 2×1 (A's rows by B's columns), and

$$A\,B = \begin{pmatrix} 1\cdot 1 + 2\cdot 0 + 1\cdot 3 \\ 2\cdot 1 + 1\cdot 0 + 2\cdot 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}.$$

For these two particular matrices, the product BA is not defined (since B's columns, 1, is not equal to A's rows, 2).

Here are some useful facts you are expected to know.

When the following matrix addition is defined, we always have

$$\begin{array}{ll} (A+B)+C = A+(B+C) & \text{addition is } \textit{associative}, \\ A+B = B+A & \text{addition is } \textit{communtative}. \end{array}$$

When the following matrix multiplication is defined, we always have

$$(AB)C = A(BC) \quad \text{multiplication is } \textit{associative}.$$

Please note however, even when both AB and BA are defined, matrix multiplication does not always commute, i.e. generally $AB \neq BA$. For example

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix}.$$

When either side below is defined, we also have

$$\begin{array}{ll} (A+B)C = AC + BC & \text{addition/multiplication is } \textit{distributive}, \\ A(B+C) = AB + AC & \text{multiplication/addition is } \textit{distributive}. \end{array}$$

Let's prove the second distributive rule above is true. Suppose A has size $m_A \times n_A$, B has size $m_B \times n_B$ and C has size $m_C \times n_C$. In one direction, suppose $A(B+C)$ is defined. Then since $B+C$ is defined $\Rightarrow$ $m_B = m_C$ and $n_B = n_C$. Let $l = m_B = m_C$ and $n = n_B = n_C$ denote the common values. Since the product A times $B+C$ is defined we must also have $n_A = l$. Therefore, conclude that

$$A \in \mathbb{R}^{m_A \times l}, \; B \in \mathbb{R}^{l \times n}, \; C \in \mathbb{R}^{l \times n}.$$

This implies AB is defined and has size $m_A \times n$, AC is defined and has size $m_A \times n$ and $AB + AC$ is defined and has size $m_A \times n$. Finally, according to the definitions of matrix addition and matrix multiplication, we have for every $1 \le i \le m_A$ and $1 \le j \le n$

$$\begin{aligned} (A(B+C))_{i,j} &= \sum_{k=1}^{l} A_{i,k}(B+C)_{k,j} \\ &= \sum_{k=1}^{l} A_{i,k}(B_{k,j} + C_{k,j}) = \sum_{k=1}^{l} A_{i,k}B_{k,j} + \sum_{k=1}^{l} A_{i,k}C_{k,j} \\ &= (AB)_{i,j} + (AC)_{i,j} = (AB+AC)_{i,j}. \end{aligned}$$

Thus, given $A(B+C)$ is defined, so is $AB+AC$, and the two are equal. The other direction is shown similarly.

---

1. Consider the matrices.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad D = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Compute the following matrix sums when defined.

(a) $A+B$ (b) $B+C$ (c) $C+D$ (d) $A+D$

2. Consider the matrices.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

Compute the following matrix products when defined.

(a) AB (b) AC (c) AD (d) CD

(e) CB (f) BC (g) DA (h) BD

3. Prove the addition/multiplication distributive rule, i.e. show $(A+B)C = AC + BC$ given that either side is defined.

4. Consider the following matrices.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 3 & 4 \\ 1 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \qquad E = \begin{pmatrix} 2 & 0 & 1 \\ 3 & 1 & 1 \end{pmatrix} \qquad F = \begin{pmatrix} 1 & 2 \end{pmatrix}$$

Verify by example the following are true.

(a) $(A+B)C = AC + BC$ (b) $F(D+E) = FD + FE$

Here's an example of a coupled linear system of three equations in three unknowns, x_1, x_2 and x_3.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_2 + 2x_3 &= 3 \\ 5x_3 &= 6 \end{aligned}$$

Geometrically, each equation defines a plane in three dimensional space. As long as none are parallel, we expect these three planes to intersect at a single point. This is an example of what's called a *triangular system* for obvious reasons. Watch how easy it is to find the intersection point.

$$\begin{aligned} 5x_3 &= 6 & \Rightarrow \quad x_3 &= 6/5 \\ 2x_2 &= 3 - 2x_3 & \Rightarrow \quad x_2 &= 3/10 \\ x_1 &= 1 - 2x_2 - 3x_3 & \Rightarrow \quad x_1 &= -16/5 \end{aligned}$$

So the point where the three planes intersect is $(x_1, x_2, x_3) = (-16/5, 3/10, 6/5)$. You've just seen what's called *back substitution* applied to this triangular system.

Here's another example of three equations in three unknowns.

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

The intersection point here is not as easy to compute because this system is not triangular. However, if we subtract the second equation from the third and make this the new third equation we get

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_2 + 2x_3 &= 3 \\ -5x_3 &= -6 \end{aligned}$$

Look familiar? Multiply the third equation by -1 to get exactly the same triangular system we solved earlier.

Here's one more example.

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

All I've done here is switch around equations one and two from the previous example. Casting this into to triangular form therefore requires two steps and in the proper order. They are:

(1) Switch around equations one and two.

(2) Subtract the second equation from the third and make the result the new third.

While the current nontriangular system and the earlier triangular system appear different, and they are, they are equivalent in the following sense. Both share the same solution.

The three equations in three unknowns given in the previous example can be rewritten as a matrix equation

$$\begin{pmatrix} 0 & 2 & 2 \\ 1 & 2 & 3 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -3 \end{pmatrix},$$

or symbolically $AX = B$, where here $A \in \mathbb{R}^{3\times 3}$ (called the coefficient matrix), $X \in \mathbb{R}^{3\times 1}$ (called the matrix of unknowns) and $B \in \mathbb{R}^{3\times 1}$ (called the right hand side matrix).

I'll use this to introduce you to a systematic procedure, called *Gaussian elimination*, which will convert a nontriangular system to an equivalent triangular one.

Form what's called the *augmented matrix* from matrices A and B

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

I'll use square brackets, $\big[$ and $\big]$, to delineate the augmented matrix and a vertical bar, $\big|$, to separate the coefficient matrix from the right hand side matrix. Gaussian elimination is comprised of a sequence of *elementary row operations* applied to the augmented matrix. An elementary row operation is one of the following.

(E1) Interchange two rows.

(E2) Multiply each element in a row by a nonzero number.

(E3) Replace a row by subtracting from it a multiple of a different row.

Such a sequence of elementary row operations could look like this.

$$\begin{array}{r} R_1 \leftrightarrow R_2 \\ \\ \\ \end{array} \left[\begin{array}{ccc|c} 0 & 2 & 2 & 3 \\ 1 & 2 & 3 & 1 \\ 0 & 2 & -3 & -3 \end{array}\right] \sim$$

$$\begin{array}{r} \\ \\ R_3 \to R_3 - R_2 \end{array} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 2 & 2 & 3 \\ 0 & 2 & -3 & -3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & -5 & -6 \end{array}\right].$$

I've used $R_1 \leftrightarrow R_2$ above to signify interchanging rows 1 and 2, and $R_3 \to R_3 - R_2$ to signify replacing row 3 with row 3 minus row 2.

Gaussian elimination employs the three elementary row operations listed above, E1, E2 and E3. However, what you really want to pay attention to is their <u>systematic</u> application. The word systematic is at the heart of Gaussian elimination.

Eliminate columnwise starting from the left most column and work right. When up to the jth column, use its jth element (the element on the current matrix diagonal) and row operations E3 to eliminate to zero column entries below it. Of course this assumes the jth element (called the *pivot* element) is nonzero. If it is zero, interchange a row below with the current row so that (if possible) the new pivot element is not zero. (This was necessary for column 1 in the previous example.) Once done with column j, move to column $j+1$.

Here's another example.

$$\begin{pmatrix} 3 & 5 & -4 \\ -3 & -5 & 5 \\ 6 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 11 \end{pmatrix} \quad \Rightarrow \quad \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ -3 & -5 & 5 & 2 \\ 6 & 1 & 1 & 11 \end{array}\right].$$

Eliminate down the first column below the pivot element.

$$\begin{array}{r} \\ R_2 \to R_2 + R_1 \\ R_3 \to R_3 - 2R_1 \end{array} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ -3 & -5 & 5 & 2 \\ 6 & 1 & 1 & 11 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & -9 & 9 & 9 \end{array}\right].$$

Next, let's eliminate down the second column. But here the pivot is zero. Interchange rows 2 and 3 to obtain a nonzero pivot element.

$$\begin{array}{r} \\ R_2 \leftrightarrow R_2 \\ \\ \end{array} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & -9 & 9 & 9 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ 0 & -9 & 9 & 9 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Now column 2 is done and we've finished the elimination procedure. But it might simplify later work if we scale row 2 by $-1/9$.

$$\begin{array}{r} \\ R_2 \to -\frac{1}{9}R_2 \\ \\ \end{array} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ 0 & -9 & 9 & 9 \\ 0 & 0 & 1 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & -4 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

There you go. Now, we can solve for the unknowns, x_1, x_2 and x_3, by back substitution.

The augmented matrix above right corresponds to the triangular system

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 1 \\ x_2 - x_3 &= -1 \\ x_3 &= 3 \end{aligned} \quad \text{which yields} \quad \begin{aligned} x_3 &= 3, \\ x_2 &= -1 + x_3 \quad \Rightarrow \quad x_2 = 2, \\ 3x_1 &= 1 - 5x_2 + 4x_3 \quad \Rightarrow \quad x_1 = 1. \end{aligned}$$

There's one more issue I want to address before giving exercises. We've used the geometric interpretation of intersecting planes to motivate the solution of three linear equations in three unknowns. But do three planes always intersect at a point? Of course they don't. Two or more of the planes can be parallel but not coplanar, i.e. two planes never intersect. There is one other possibility. There may be an infinite number of solutions when planes are parallel but coplanar.

Here's an example which exemplifies such possibilities.

$$\begin{pmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \\ -4 \end{pmatrix} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{array}\right].$$

Let's start the elimination phase.

$$\begin{array}{l} \\ R_2 \to R_2 + R_1 \\ R_3 \to R_3 - 2R_1 \end{array} \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{array}\right]$$

$$\begin{array}{l} \\ \sim \\ R_3 \to R_3 + 3R_2 \end{array} \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & -9 & 0 & -18 \end{array}\right] \sim \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The third row in augmented matrix above right says $0x_3 = 0$. So x_3 can have any value, say $x_3 = \alpha$. The second row says $3x_2 + 0x_3 = 6$ or $x_2 = 2$. The first row says $3x_1 + 5x_2 - 4x_3 = 7$ or $x_1 = -1 + \frac{4}{3}\alpha$. Here we have an infinite number of solutions,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 + \frac{4}{3}\alpha \\ 2 \\ \alpha \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} \frac{4}{3} \\ 0 \\ 1 \end{pmatrix} \quad \text{is a solution for any parameter } \alpha.$$

If the right hand side in previous example is slightly tweaked,

$$\begin{pmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \\ -3 \end{pmatrix} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -3 \end{array}\right],$$

compute that Gaussian elimination gives

$$\left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -3 \end{array}\right] \sim \left[\begin{array}{rrr|r} 3 & 5 & -4 & 7 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Here the third row in the reduced augmented matrix says $0x_3 = 1$. But this is impossible. Therefore, this system has no solution.

All of the examples given so far deal with small and square systems, i.e. the number of equations is the same as the number of unknowns. We'll generalize the elimination procedure after giving some exercises.

Do the following for the next four linear systems.

(a) Write it in the form of a matrix equation, $AX = B$.

(b) Reduce the augmented matrix to triangular form if possible.

(c) Solve for x_1, x_2 and x_3.

5. $\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_1 + x_2 - 3x_3 &= 5 \\ 4x_1 - 7x_2 + x_3 &= -1 \end{aligned}$

6. $\begin{aligned} 2x_1 - x_2 + 3x_3 &= 5 \\ 2x_1 + 2x_2 + 3x_3 &= 7 \\ -2x_1 + 3x_2 &= -3 \end{aligned}$

7. $\begin{aligned} 3x_1 - 4x_2 + 5x_3 &= 7 \\ -3x_1 + 4x_2 - 2x_3 &= -1 \\ 6x_1 - 8x_2 + x_3 &= -4 \end{aligned}$

8. $\begin{aligned} x_1 + x_2 - 3x_3 &= 4 \\ 2x_1 + x_2 - x_3 &= 2 \\ 3x_1 + 2x_2 - 4x_3 &= 7 \end{aligned}$

The point of elimination is, by means of the three elementary row operations, to get the augmented matrix into a form where solving by back substitution is possible. Up to now I've focused on small square systems. But not all applications are small and/or square.

Suppose after eliminating down the first column of the augmented matrix for some 4×4 system, the augmented matrix looked like

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

where the $*$'s represent arbitrary values. What's your next step? It's clear. Swap rows 4 and 2, and then move on to eliminating down column 3.

Now, suppose our hypothetical augmented matrix looked like this

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

(Note the 2 has been changed to 0.) What do we do next here? There's no row swapping that can help. Remember, elimination is a systematic application of elementary row operations that, when completed, should facilitate back substitution to solve the resulting problem. With that in mind, leave column 2 alone and move on to eliminating down column 3 to get for example

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

The goal of Gaussian elimination isn't to convert the coefficient matrix part of the augmented matrix to diagonal form with nonzero diagonal elements. As we've seen, this isn't always possible. Rather, Gaussian elimination should convert the augmented matrix to what is called *row echelon form.*

Borrowed from `https://wikipedia.org/wiki/Row_echelon_form`:

A matrix is in row echelon form if

- all nonzero rows (rows with at least one nonzero element) are above any rows of all zeroes (all zero rows, if any, belong at the bottom of the matrix), and
- the leading coefficient (the first nonzero number from the left, also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.

These two conditions imply that all entries in a column below a leading coefficient are zeros.

By means of elementary row operations, E1, E2 and E3, it's **always possible to convert** a matrix of any size, square or not square, **to row echelon form**.

To help clarify what matrices are in echelon form, recall what Justice Potter Stewart famously said, "I know it when I see it." So here are three examples of augmented matrices in row echelon form.

$$\left[\begin{array}{cccc|c} 0 & 2 & * & * & * \\ 0 & 0 & 3 & * & * \end{array}\right], \quad \left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 0 & 3 & * & * \\ 0 & 0 & 0 & 4 & * \end{array}\right], \quad \left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 2 & * & * & * \\ 0 & 0 & 0 & 4 & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & * \end{array}\right].$$

There's nothing special about the particular numbers 1, 2, 3 and 4 used above. Any <u>nonzero</u> number could replace them. The $*$ entries on the other hand can contain any zero or nonzero numbers.

Let's finish the discussion with examples of how back substitution works in such nondiagonal cases.

First, suppose we had a linear system of two equations in four unknowns, say x_1, x_2, x_3 and x_4, whose reduced augmented matrix is

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

Here, x_1 and x_4 are what are called *free variables.* Call $x_1 = \alpha$ and $x_4 = \beta$ where α and β can be any real number. Then read off (backwards) from the augmented matrix to see

$$\begin{aligned} 3x_3 + x_4 &= 1 \\ 2x_2 + x_3 + x_4 &= 1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_3 &= \tfrac{1}{3}(1-\beta) = \tfrac{1}{3} - \tfrac{1}{3}\beta, \\ x_2 &= \tfrac{1}{2}(1 - \tfrac{1}{3}(1-\beta) - \beta) = \tfrac{1}{3} - \tfrac{1}{3}\beta. \end{aligned}$$

So this example has a two parameter family of solutions

$$x_1 = \alpha,\ x_2 = \tfrac{1}{3} - \tfrac{1}{3}\beta,\ x_3 = \tfrac{1}{3} - \tfrac{1}{3}\beta,\ x_4 = \beta.$$

Second, suppose we had a linear system of three equations in four unknowns whose reduced augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 4 & 1 \end{array}\right].$$

Here, x_2 is the only free variable, say $x_2 = \alpha$. Again read off backwards to see

$$\begin{aligned} 4x_4 &= 1 \\ 3x_3 + x_4 &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_4 &= \tfrac{1}{4}, \\ x_3 &= \tfrac{1}{4}, \\ x_1 &= \tfrac{1}{2} - \alpha. \end{aligned}$$

Third, suppose we had a linear system of five equations in four unknowns whose reduced augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & b_4 \\ 0 & 0 & 0 & 0 & b_5 \end{array}\right].$$

Observe that the fourth and fifth equations say

$$\begin{aligned} 0x_1 + 0x_2 + 0x_3 + 0x_4 &= b_4, \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= b_5, \end{aligned}$$

so there can be <u>no</u> solution <u>unless</u> $b_4 = b_5 = 0$. If these are zero, the free variable is $x_3 = \alpha$ and read off backwards from the augmented matrix to obtain

$$\begin{aligned} 4x_4 &= 1 \\ 2x_2 + x_3 + x_4 &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 1 \end{aligned} \quad \Rightarrow \quad \begin{aligned} x_4 &= \tfrac{1}{4}, \\ x_2 &= \tfrac{3}{8} - \tfrac{1}{2}\alpha, \\ x_1 &= \tfrac{3}{8} - \tfrac{1}{2}\alpha. \end{aligned}$$

Here is your last group of exercises for this assignment.

For the following systems, written as $AX = B$, reduce the augmented matrix to row echelon form, and then find all solutions X if one exists.

9. $\begin{pmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 4 & 3 \\ 2 & 4 & -4 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ 10. $\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & -1 & 1 & -1 \\ 3 & -1 & 4 & -5 \\ 2 & 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 7 \\ 5 \end{pmatrix}$

I got for 9: $x_1 = -1 - 2\alpha,\ x_2 = \alpha,\ x_3 = 0,\ x_4 = 1$ for any $\alpha \in \mathbb{R}$.

I got for 10: $x_1 = (5 - 3\alpha)/2,\ x_2 = (1 - \alpha)/2,\ x_3 = \alpha,\ x_4 = 0$.