# Contents

# 1 Matrices and Systems of Linear Equations

In this chapter, we will discuss the problem of solving systems of linear equations, reformulate the problem using matrices, and then give the general procedure for solving such systems. We will then study basic matrix operations (addition and multiplication) and discuss the inverse of a matrix and how it relates to another quantity known as the determinant. We will then revisit systems of linear equations after reformulating them in the language of matrices.

## 1.1 Systems of Linear Equations

- Our motivating problem is to study the solutions to a system of linear equations, such as the system

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & 5 \\ 3x_1 & + & x_2 & = & -1 \end{array}.$$

  ○ Recall that a linear equation is an equation of the form $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, for some constants $a_1, \ldots a_n, b$, and variables $x_1, \ldots, x_n$.

  ○ When we seek to find the solutions to a system of linear equations, this means finding all possible values of the variables such that all equations are satisfied simultaneously.

- Example: One can check that the system

$$\begin{array}{rcrcr} x_1 & + & 3x_2 & = & 5 \\ 3x_1 & + & x_2 & = & -1 \end{array}$$

  in the two variables $x_1, x_2$ has a solution $(x_1, x_2) = (-1, 2)$. (Indeed, it turns out that this is the only solution.)

- Example: One can check that the system

$$\begin{array}{rcrcrcr} 2x_1 & - & x_2 & + & 3x_3 & = & 6 \\ x_1 & + & x_2 & - & 2x_3 & = & 1 \\ 4x_1 & + & x_2 & - & x_3 & = & 8 \end{array}$$

  in the three variables $x_1, x_2, x_3$ has as solutions $(x_1, x_2, x_3) = (2, 1, 1)$, $(1, 8, 4)$, and, more generally, any 3-tuple of the form $(2 - t, 1 + 7t, 1 + 3t)$ for any real number $t$. (Indeed, it turns out that these are all the solutions.)

### 1.1.1 Elimination, Matrix Formulation

- The traditional method for solving a system of linear equations (likely familiar from basic algebra) is by elimination: we solve the first equation for one variable $x_1$ in terms of the others, and then plug in the result to all the other equations to obtain a reduced system involving one fewer variable. Eventually, the system will simplify either to a contradiction (e.g., $1 = 0$), a unique solution, or an infinite family of solutions.

- Example: Solve the system of equations

$$\begin{array}{ccccc} x & + & y & = & 7 \\ 2x & - & 2y & = & -2 \end{array} .$$

  ◦ We can solve the first equation for $x$ to obtain $x = 7 - y$.

  ◦ Plugging in this relation to the second equation gives $2(7 - y) - 2y = -2$, or $14 - 4y = -2$, so that $\boxed{y = 4}$. Then since $x = 7 - y$ we obtain $\boxed{x = 3}$.

- Another way to perform elimination is to add and subtract multiples of the equations, so to eliminate variables (and remove the need to solve for each individual variable before eliminating it).

  ◦ In the example above, instead of solving the first equation for $x$, we could multiply the first equation by $-2$ and then add it to the second equation, so as to eliminate $x$ from the second equation.

  ◦ This yields the same overall result, but is less computationally difficult.

- Example: Solve the system of equations

$$\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ 2x & + & 3y & - & z & = & 1 \\ -x & + & 2y & + & 2z & = & 1 \end{array} .$$

  ◦ If we label the equations #1, #2, #3, then we can eliminate $x$ by taking $[\#2] - 2[\#1]$ and $[\#3] + [\#1]$. This gives the new system
$$\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ & & y & - & 7z & = & -7 \\ & & 3y & + & 5z & = & 5 \end{array} .$$

  ◦ Now we can eliminate $y$ by taking $[\#3] - 3[\#2]$. This yields

$$\begin{array}{ccccccc} x & - & 2y & + & 3z & = & 4 \\ & & y & - & 7z & = & -7 \\ & & & & 26z & = & 26 \end{array} .$$

  ◦ Now the third equation immediately gives $z = 1$. Then the second equation requires $y = 0$, and the first equation gives $x = 1$, so the solution is $(x, y, z) = \boxed{(1, 0, 1)}$.

- This procedure of elimination can be simplified even more, because we don't really need to write the variable labels down every time. We only need to keep track of the coefficients, which we can do by putting them into an array.

  ◦ Example: The system
$$\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ 2x & + & 3y & - & z & = & 1 \\ -x & + & 2y & + & 2z & = & 1 \end{array}$$
  can be written in simplified form using the array

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

- ○ We can then do operations on the entries in the array that correspond to manipulations of the associated system of equations.

- Since we will frequently work with arrays of different sizes, it is useful to have a way to refer to them:

- <u>Definition</u>: An $m \times n$ <u>matrix</u> is an array of numbers with $m$ rows and $n$ columns. A <u>square matrix</u> is one with the same number of rows and columns: that is, an $n \times n$ matrix for some $n$.

  - ○ <u>Examples</u>: $\begin{bmatrix} 4 & 1 & -1 \\ 3 & 2 & 0 \end{bmatrix}$ is a $2 \times 3$ matrix, and $\begin{bmatrix} \pi & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & 9 \end{bmatrix}$ is a $3 \times 3$ square matrix.

  - ○ When working with a coefficient matrix, we will draw a line to separate the coefficients of the variables from the constant terms. This type of matrix is often called an <u>augmented matrix</u>.

- <u>Definition</u>: If $A$ is a matrix, the entry in the $i$th row and $j$th column of $A$ is called the <u>$(i, j)$-entry</u> of $A$, and will be denoted $a_{i,j}$.

  - ○ <u>Warning</u>: It is easy to mix up the coordinates. Remember that the first coordinate specifies the row, and the second coordinate specifies the column.
  - ○ <u>Example</u>: If $A = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 0 & 5 \end{bmatrix}$, then the $(2, 2)$-entry is $a_{2,2} = 0$ and the $(1, 3)$-entry is $a_{1,3} = 4$.

- <u>Definition</u>: In an $n \times n$ square matrix, the $(i, i)$-entries for $1 \leq i \leq n$ form the <u>diagonal</u> of the matrix. The <u>trace</u> of a square matrix is the sum of its diagonal entries.

  - ○ <u>Example</u>: The diagonal entries of $\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$ are 4 and 2, so the trace of this matrix is 6.

- When doing elimination, each step involves one of the three <u>elementary row operations</u> on the rows of the coefficient matrix:

  1. Interchange two rows.
  2. Multiply all entries in a row by a nonzero constant.
  3. Add a constant multiple of one row to another row.

- Each of these elementary row operations leaves unchanged the solutions to the associated system of linear equations. The idea of elimination is to apply these elementary row operations to the coefficient matrix until it is in a simple enough form that we can simply read off the solutions to the original system of equations.

- <u>Example</u>: Use elementary row operations to solve the system $\begin{array}{rcrcr} x & + & y & = & 2 \\ 2x & - & 3y & = & 9 \end{array}$.

  - ○ The associated augmented coefficient matrix is $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 2 & -3 & 9 \end{array}\right]$.

  - ○ Subtracting twice the first row from the second row produces $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -5 & 5 \end{array}\right]$.

  - ○ Scaling the second row by $-\dfrac{1}{5}$ then produces $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \end{array}\right]$.

  - ○ Subtracting the second row from the first row yields $\left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -1 \end{array}\right]$.

  - ○ Now we can easily read off the solutions: the first equation says $x = 3$ and the second says $y = -1$.

**1.1.2 Row-Echelon Form and Reduced Row-Echelon Form**

- The procedure of applying elementary row operations to a matrix until it is in a simpler form, is called row-reduction.

- We will now take a brief digression to discuss some standard forms of coefficient matrices that arise frequently. The most basic simple form is called row-echelon form:

- Definition: A matrix is in row-echelon form if (i) all rows with at least one nonzero element are above any row of all zero entries, and (ii) the first nonzero term in each row is always to the right of the first nonzero term in the row above it. (The first nonzero term in each row is called the pivot.)

  ○ A shorter way of writing the two conditions is (i) all rows without a pivot (the rows of all zeroes) are at the bottom, and (ii) any row's pivot, if it has one, lies to the right of the pivot of the row directly above it.

- Here are some examples of matrices in row-echelon form, where the pivot elements have been boxed:

  ○ $\begin{bmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & \boxed{1} & 2 & 3 & 4 \\ 0 & 0 & \boxed{1} & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}$, $\begin{bmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0 & 0 & \boxed{3} & 4 & 5 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

- Here are some examples of matrices not in row-echelon form:

  ○ $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$ : the pivot in the second row is not strictly to the right of the pivot element above it.

  ○ $\begin{bmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$ : the pivot in the third row is not strictly to the right of the pivot element above it.

  ○ $\begin{bmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ : the row of all zeroes is not at the bottom of the matrix.

- If the coefficient matrix is in row-echelon form, it is easy to read off the solutions to the corresponding system of linear equations by working from the bottom up.

  ○ Example: The augmented matrix $\begin{bmatrix} 1 & 1 & 3 & | & 4 \\ 0 & 1 & -1 & | & 1 \\ 0 & 0 & 2 & | & 4 \end{bmatrix}$ corresponding to the system

  $$\begin{array}{rrrrcl} x & + & y & + & 3z & = & 4 \\ & & y & - & z & = & 1 \\ & & & & 2z & = & 4 \end{array}.$$

  is in row-echelon form. The bottom equation immediately gives $z = 2$. Then the middle equation gives $y = 1 + z = 3$, and the top equation gives $x = 4 - y - 3z = -5$.

- Example: Use row operations to put the matrix $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}$ into row-echelon form.

  ○ We apply elementary row operations to clear out the first column, and then we can clear out the third row.

  $$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow{R_3-R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow{R_4-2R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_3-R_2} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

○ The matrix is now in row-echelon form.

○ Notice that there are other possible combinations of row operations we could have performed to put the matrix into a row-echelon form.

○ For example, we could have started by swapping rows 1 and 3, and then cleared out the first column and the lower rows:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow{R_3 - R_1} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & -1 & -2 & -3 \\ 2 & 4 & 6 & 8 \end{bmatrix} \xrightarrow{R_4 - 2R_1} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & -1 & -2 & -3 \\ 0 & -2 & -4 & -6 \end{bmatrix}$$

$$\xrightarrow{R_3 + R_2} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -4 & -6 \end{bmatrix} \xrightarrow{R_4 + 2R_2} \begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

○ Notice that we obtain a different row-echelon form in this case.

• An even simpler form is called reduced row-echelon form:

• Definition: A matrix is in reduced row-echelon form if it is in row-echelon form, all pivot elements are equal to 1, and each pivot element is the only nonzero term in its column.

○ Here are some matrices in reduced row-echelon form (with pivot elements boxed): $\begin{bmatrix} \boxed{1} & 0 & 0 & 4 & 5 \\ 0 & \boxed{1} & 0 & 3 & 4 \\ 0 & 0 & \boxed{1} & 0 & 1 \end{bmatrix}$,

$$\begin{bmatrix} \boxed{1} & 2 & 3 & 0 & 5 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \boxed{1} & 2 & 0 & 4 & 0 \\ 0 & 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}.$$

○ Here are some examples of matrices not in reduced row-echelon form:

∗ $\begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$ : the pivot in the second row has a nonzero entry in its column.

∗ $\begin{bmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ : the pivot in the first row is not equal to 1.

• Example: Use row operations to put the matrix $\begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 1 \\ 5 & 2 & 0 \end{bmatrix}$ into reduced row-echelon form.

○ Applying elementary row operations to clear out the first column, and then the bottom row, yields

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 1 \\ 5 & 2 & 0 \end{bmatrix} \xrightarrow{R_2 - 3R_1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -5 \\ 5 & 2 & 0 \end{bmatrix} \xrightarrow{R_3 - 5R_1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -5 \\ 0 & 2 & -10 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 0 \end{bmatrix}.$$

• There are numerous ways to row-reduce a given matrix until it is in row-echelon form, and many different row-echelon forms are possible. However, it turns out that the reduced row-echelon form is always unique:

• Theorem: Every matrix has a unique reduced row-echelon form.

○ We will not prove this theorem. However, it is useful from a theoretical standpoint to know that, regardless of the way we perform row-reductions, we will always obtain the same reduced row-echelon form when we are finished.

• Definition: The rank of a matrix is the number of nonzero rows in its (reduced) row-echelon form. Equivalently, it is the number of pivots that appear when the matrix is in (reduced) row-echelon form.

○ Examples: The rank of $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$ is 3, while the rank of $\begin{bmatrix} 1 & 2 & 3 & 0 & 5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is 2.

### 1.1.3 Gaussian Elimination

- By using row operations on a coefficient matrix, we can find the solutions to the associated system of equations, since as we noted before each of the elementary row operations does not change the solutions to the system.

- The only remaining difficulty is interpreting the results we obtain once the coefficient matrix is in row-echelon form. There are three different possible cases, which we will illustrate by using an augmented $3 \times 3$ coefficient matrix in row-echelon form:

  1. The system has a unique solution.
     ○ Example: $\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 4 \end{array}\right]$ whose corresponding system is $\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ & & y & - & z & = & 1 \\ & & & & 2z & = & 4 \end{array}$ . The unique
     solution is $z = 2$, $y = 3$, $x = -5$ as we can see by reading the system from the bottom up.
     ○ Note that in this case, each of the columns on the left side of the matrix is a pivotal column.

  2. The system is inconsistent ("overdetermined") and has no solutions.
     ○ Example: $\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{array}\right]$ whose corresponding system is $\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ & & y & - & z & = & 1 \\ & & & & 0 & = & 4 \end{array}$ . The bot-
     tom equation is contradictory.
     ○ Note that in this case, the column on the right side of the matrix is a pivotal column.

  3. The system has infinitely many solutions ("underdetermined").
     ○ Example: $\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$ whose corresponding system is $\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ & & y & - & z & = & 1 \\ & & & & 0 & = & 0 \end{array}$ . The bot-
     tom equation is always true so there are really only two relations between the three variables $x, y, z$.
     ○ If we take $z = t$ to be an arbitrary parameter, then the equations require $y = 1 + z = 1 + t$ and $x = 4 - y - 3z = 3 - 4t$, and it is equally easy to see that every triple of the form $(x, y, z) = (3 - 4t, 1 + t, t)$ satisfies the equations.
     ○ Note that in this case, there are nonpivotal columns on the left side of the matrix: specifically, the column corresponding to the third variable $z$, which was also the variable we assigned to have an arbitrary parameter value.

- Our observations about the pivotal columns will hold in general, and gives us a simple way to determine the structure of the solution set:

- Definition: If a variable is associated to a nonpivotal column, it is called a free variable. If a variable is associated to a pivotal column, it is called a bound variable.

  ○ Example: For the system $\begin{array}{ccccccc} x & + & y & + & 3z & = & 4 \\ & & y & - & z & = & 1 \\ & & & & 0 & = & 0 \end{array}$ with matrix $\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$ , $x$ and $y$ are
  bound variables and $z$ is a free variable.

- Theorem (Solutions to Linear Systems): Suppose the augmented matrix for a system of linear equations is in row-echelon form. If there is a pivot in the column on the right side (the column of constants) then the system is inconsistent, and otherwise the system is consistent. In the latter case, if each column on the left side has a pivot, the system will have a unique solution. Otherwise, the system has infinitely many solutions and, more specifically, the variable associated to each nonpivotal column is a free variable that can be assigned an arbitrary value, and the bound variables associated to the pivotal columns can be written uniquely in terms of these free variables.

◦ The statement of the theorem is rather lengthy but it is really just an encapsulation of what we saw in the three cases above.

◦ <u>Proof</u>: Suppose first that there is a pivot in the column of constants on the right side of the augmented matrix, which we assume to be in reduced row-echelon form (nothing in the statement changes if we use only the row-echelon form, but it is easier to see what is going on with the reduced row-echelon form).

◦ The only way there can be a pivot in the right-side column is if the coefficient matrix has a row of the form $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & | & a \end{bmatrix}$ for some nonzero constant $a$. The corresponding equation reads $0 = a$, which is impossible since $a$ is not zero: this is a contradiction, so the system cannot have any solutions.

◦ Now suppose that the column of constants does not have a pivot. We claim that, if we assign each of the free variables (the variables in the nonpivotal columns) an arbitrary parameter value, the system of equations will uniquely specify each of the bound variables in terms of the free variables.

◦ To see this we simply read the equations from the bottom up: there will be some number of rows reading $0 = 0$, and then, by the assumption that the matrix is in row-echelon form, each bound variable will occur as the leading term of exactly one equation (namely, the equation corresponding to the row where the pivot element for that variable's column is located).

◦ Each row that does not have all zero entries has exactly one pivot, and the corresponding equation will define the associated bound variable in terms of the free variables. So every bound variable is defined in terms of the free variables, and every equation is satisfied, so the solutions have been completely found.

◦ Finally, the system will have infinitely many solutions if there are any free variables (since a free variable can take an infinite number of possible values), and will have exactly one solution if there are no free variables. Having no free variables is equivalent to each variable being a bound variable, which happens precisely when every column on the left side of the matrix has a pivot.

• Using the theorem above we can give an algorithm for solving a system of linear equations by putting the coefficient matrix in row-echelon form. This procedure is known as <u>Gaussian elimination</u>:

◦ <u>Step 1</u>: Convert the system to its augmented coefficient matrix.

◦ <u>Step 2</u>: If all entries in the first column are zero, remove this column from consideration and repeat this step until an entry in the first column is nonzero. Swap rows, if necessary, so that the upper-left entry in the first column is nonzero.

◦ <u>Step 3</u>: Use row operations to clear out all entries in the first column below the first row.

◦ <u>Step 4a</u>: Repeat steps 2 and 3 on the submatrix obtained by ignoring the first row and first column, until all remaining rows have all entries equal to zero.

   ∗ After following these steps, the matrix will now be in row-echelon form, and the system can be fairly easily solved. To put the matrix in reduced row-echelon form, we have an optional extra step:

◦ <u>Step 4b</u>: Identify the "pivotal columns" (columns containing a leading row-term), and then perform row operations to clear out all non-leading entries in each pivotal column.

◦ <u>Step 5</u>: Convert the matrix back to its corresponding system of equations and interpret the results:

   ∗ If there is a pivot in the column on the right side (the column of constants) then the system is inconsistent. Otherwise, the system is consistent.

   ∗ If each column on the left side has a pivot, the system will have a unique solution. Read the system from the bottom up to find it.

   ∗ If there are columns without a pivot, the corresponding variables are free variables. Introduce appropriate parameters for the free variables, and use the equations from the bottom up to write the other variables in terms of the free variables.

• <u>Example</u>: Solve the system $x + y + 3z = 4$, $2x + 3y - z = 1$, $-x + 2y + 2z = 1$ using Gaussian elimination.

◦ Given the system
$$\begin{array}{rrrrrrl} x & + & y & + & 3z & = & 4 \\ 2x & + & 3y & - & z & = & 1 \\ -x & + & 2y & + & 2z & = & 1 \end{array}$$

7

we write it in matrix form to get

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

○ We apply elementary row operations to clear out the first column:

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 2 & 3 & -1 & 1 \\ -1 & 2 & 2 & 1 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ -1 & 2 & 2 & 1 \end{array}\right] \xrightarrow{R_3 + R_1} \left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ 0 & 3 & 5 & 5 \end{array}\right].$$

○ Now we are done with the first column and can focus on the other columns:

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ 0 & 3 & 5 & 5 \end{array}\right] \xrightarrow{R_3 - 3R_2} \left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ 0 & 0 & 26 & 26 \end{array}\right] \xrightarrow{\frac{1}{26}R_3} \left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

○ The system is in row-echelon form. To put it in reduced row-echelon form, we can work from the bottom up:

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & -7 & -7 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_2 + 7R_3} \left[\begin{array}{ccc|c} 1 & 1 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_1 - R_2} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_2 - 3R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

○ From here we see that the right-hand column is not pivotal, so the system has a solution. Furthermore, every column on the left side has a pivot, so there are no free variables and so the system has a unique solution.

○ If we convert back we immediately see that the unique solution is $\boxed{x = 1, \ y = 0, \ z = 1}$. (Note of course that this is the same answer we got when we did elimination without using a matrix.)

• <u>Example</u>: Solve the system $a - b + c - d + e = 5$, $-a + b - c + d = 4$, $2d + 3e = 3$ using Gaussian elimination.

○ The coefficient matrix is

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

○ We put it in row-echelon form:

$$\left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ -1 & 1 & -1 & 1 & 0 & 4 \\ 0 & 0 & 0 & 2 & 3 & 3 \end{array}\right] \xrightarrow{R_2 + R_1} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 2 & 3 & 3 \end{array}\right] \xrightarrow{R_2 \Leftrightarrow R_3} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 2 & 3 & 3 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right].$$

○ Now we put it in reduced row-echelon form:

$$\left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 2 & 3 & 3 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right] \xrightarrow{R_2 - 3R_3} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 2 & 0 & -24 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right] \xrightarrow{\frac{1}{2}R_2} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 1 & 0 & -12 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right]$$

$$\xrightarrow{R_1 - R_3} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & -1 & 0 & -4 \\ 0 & 0 & 0 & 1 & 0 & -12 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right] \xrightarrow{R_1 + R_2} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & 0 & 0 & -16 \\ 0 & 0 & 0 & 1 & 0 & -12 \\ 0 & 0 & 0 & 0 & 1 & 9 \end{array}\right].$$

○ From here we see that the right-hand column is not pivotal, so the system has a solution. Furthermore, columns 1, 4, and 5 are pivotal and columns 2 and 3 are nonpivotal, so $a, d, e$ are bound variables and $b, c$ are free variables.

○ To write down the general solution we set $b = t_1$ and $c = t_2$ for parameters $t_1$ and $t_2$, and then solve for $a, d, e$ in terms of these parameters: the third equation gives $e = 9$, the second gives $d = -12$, and the first gives $a - b + c = -16$, so that $a = t_1 - t_2 - 16$.

8

○ Thus the general solution to the system is $\boxed{(a, b, c, d, e) = (t_1 - t_2 - 16,\ t_1,\ t_2,\ -12,\ -16)}$, where $t_1$ and $t_2$ are arbitrary parameters.

● <u>Example</u>: Solve the system $x + y - 2z = 3$, $-x + 3y - 5z = 1$, $3x - y + z = 2$ using Gaussian elimination.

○ We row-reduce the coefficient matrix:

$$\left[\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ -1 & 3 & -5 & 1 \\ 3 & -1 & 1 & 2 \end{array}\right] \xrightarrow{R_2 + R_1} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 0 & 4 & -7 & 4 \\ 3 & -1 & 1 & 2 \end{array}\right] \xrightarrow{R_3 - 3R_1} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 0 & 4 & -7 & 4 \\ 0 & -4 & 7 & -7 \end{array}\right] \xrightarrow{R_2 + R_3} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 3 \\ 0 & 4 & -7 & 4 \\ 0 & 0 & 0 & -3 \end{array}\right].$$

○ From here we see that the right-hand column is pivotal, so the system has $\boxed{\text{no solution}}$. (The corresponding row is the contradictory equation $0 = -3$.)

## 1.2 Matrix Operations: Addition and Multiplication

● We will now discuss some algebraic operations we can do with matrices.

● Like with vectors, we can add and subtract matrices of the same dimension, and we can also multiply a matrix by a scalar. Each of these operations is done "componentwise": to add or subtract, we just add or subtract the corresponding entries of the two matrices. To multiply by a scalar, we just multiply each entry by that scalar.

○ <u>Example</u>: If $A = \begin{bmatrix} 1 & 6 \\ 2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, then $A + B = \begin{bmatrix} 1+3 & 6+0 \\ 2+0 & 2+2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 2 & 4 \end{bmatrix}$, $2A = \begin{bmatrix} 2 \cdot 1 & 2 \cdot 6 \\ 2 \cdot 2 & 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 4 & 4 \end{bmatrix}$, and $A - \frac{1}{3}B = \begin{bmatrix} 0 & 6 \\ 2 & \frac{4}{3} \end{bmatrix}$.

● We also have a transposition operation, where we interchange the rows and columns of the matrix:

● <u>Definition</u> (Matrix Transpose): If $A$ is an $n \times m$ matrix, then the <u>transpose</u> of $A$, denoted $A^T$, is the $m \times n$ matrix whose $(i, j)$-entry is equal to the $(j, i)$-entry of $A$.

○ <u>Example</u>: If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

● Matrix multiplication, however, is NOT performed componentwise. Instead, the product of two matrices is the "row-column product".

● <u>Definition</u> (Matrix Product): If $A$ is an $m \times n$ matrix and $B$ is an $n \times q$ matrix, then the matrix <u>product</u> $A \cdot B$, often written simply as $AB$, is the $m \times q$ matrix whose $(i, j)$-entry is the sum $(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$, the sum of products of corresponding entries from from the $i$th row of $A$ with the $j$th column of $B$.

○ This product is sometimes called the <u>row-column product</u> to emphasize the fact that it is a product involving the rows of $A$ with the columns of $B$.

○ <u>Important Note</u>: In order for the matrix product to exist, the number of columns of $A$ must equal the number of rows of $B$. In particular, if $A$ and $B$ are the same size, their product exists only if they are square matrices. Also, if $AB$ exists, then $BA$ may not necessarily exist.

○ A shorter way to summarize the definition of the matrix product is to use the associated idea of a dot product of two vectors: if $\mathbf{v} = \langle a_1, a_2, \cdots, a_n \rangle$ and $\mathbf{w} = \langle b_1, b_2, \cdots, b_n \rangle$ are two vectors, then their dot product is defined to be the sum $a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$. Then the $(i, j)$ entry of the matrix product $AB$ is the dot product of the $i$th row of $A$ with the $j$th column of $B$ (each thought of as vectors).

● <u>Example</u>: If $A = \begin{bmatrix} -1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 3 & 3 \end{bmatrix}$, find $AB$ and $BA$.

○ Since $A$ is a $2 \times 3$ matrix and $B$ is a $3 \times 2$ matrix, we see that $AB$ is defined and will be a $2 \times 2$ matrix.

○ The $(1, 1)$ entry of $AB$ is $(-1)(1) + (1)(2) + (2)(3) = 7$.

○ The $(1, 2)$ entry of $AB$ is $(-1)(2) + (1)(0) + (2)(3) = 4$.

○ The $(2, 1)$ entry of $AB$ is $(0)(1) + (1)(2) + (1)(3) = 5$.

○ The $(2, 2)$ entry of $AB$ is $(0)(2) + (1)(0) + (1)(3) = 3$.

○ Putting all of this together gives $AB = \boxed{\begin{bmatrix} 7 & 4 \\ 5 & 3 \end{bmatrix}}$.

○ We see that $BA = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ is also defined and will be a $3 \times 3$ matrix.

○ The $(1, 1)$ entry of $BA$ is $(1)(-1) + (2)(0) = -1$, the $(1, 2)$ entry is $(1)(1) + (2)(1) = 3$, and the $(1, 3)$ entry is $(1)(2) + (2)(1) = 4$.

○ In a similar way we can compute the other six entries: the result is $BA = \boxed{\begin{bmatrix} -1 & 3 & 4 \\ -2 & 2 & 4 \\ -3 & 6 & 9 \end{bmatrix}}$.

• The definition of matrix multiplication seems very peculiar at first. Ultimately, it is defined the way it is in order to make changes of variables in a system of equations work correctly. It also allows us to rewrite a system of linear equations as a single "matrix equation". (We will return to these ideas later.)

○ <u>Example</u> (Matrix Equations): The system of equations

$$\begin{array}{rcrcr} x & + & y & = & 7 \\ 2x & - & 2y & = & -2 \end{array}$$

can be rewritten as a matrix equation

$$\begin{bmatrix} 1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$$

since the product on the left-hand side is the column matrix $\begin{bmatrix} x + y \\ 2x - 2y \end{bmatrix}$.

○ <u>Example</u> (Substitution): Consider what happens if we are given the equations $x_1 = y_1 + y_2$, $x_2 = 2y_1 - y_2$ and the equations $y_1 = 3z_1 - z_2$, $y_2 = z_1 - z_2$, and want to express $x_1$ and $x_2$ in terms of $z_1$ and $z_2$. It is straightforward to plug in and check that $x_1 = 4z_1 - 2z_2$ and $x_2 = 5z_1 - z_2$.

* In terms of matrices this says $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$.

* So we would want to be able to say $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$.

* Indeed, we have the matrix product $\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 5 & -1 \end{bmatrix}$. So the definition of matrix multiplication makes everything consistent with what we'd want to happen.

• If we restrict our attention to square matrices, then matrices under addition and multiplication obey some, but not all, of the algebraic properties that real numbers do.

○ In general, matrix multiplication is NOT commutative: $AB$ typically isn't equal to $BA$, even if $A$ and $B$ are both square matrices.

* <u>Example</u>: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$. Then $AB = \begin{bmatrix} 10 & 13 \\ 22 & 29 \end{bmatrix}$ while $BA = \begin{bmatrix} 11 & 16 \\ 19 & 28 \end{bmatrix}$.

○ Matrix multiplication distributes over addition, on both sides: $(A + B)C = AC + BC$ and $A(B + C) = AB + AC$.

* This property can be derived from the definition of matrix multiplication, along with some arithmetic.

○ Matrix multiplication is associative: $(AB)C = A(BC)$, if $A, B, C$ are of the proper dimensions.

  * In particular, taking the $n$th power of a square matrix is well-defined for every positive integer $n$.
  * This property can also be derived from the definition, but the arithmetic is very cumbersome.

○ The transpose of the product of two matrices is the product of their transposes in reverse order: $(AB)^T = B^T A^T$.

  * This property can likewise be derived from the definition of matrix multiplication.

- <u>Definition</u>: If $A$ is an $n \times n$ matrix, then there is a zero matrix $Z_n$ which has the properties $Z_n + A = A$ and $Z_n A = A Z_n = Z_n$. This matrix $Z_n$ is the matrix whose entries are all zeroes.

  ○ <u>Example</u>: The $2 \times 2$ zero matrix is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

  ○ <u>Remark</u>: In contrast to real (or complex) numbers, where $x^2 = 0$ implies $x = 0$, there exist nonzero matrices whose square is nonetheless the zero matrix. One such matrix is $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$: it is easy to check that $A^2$ is the zero matrix, but of course $A$ itself is nonzero.

- <u>Definition</u>: If $A$ is an $n \times n$ matrix, then there is an $n \times n$ <u>identity matrix</u> $I_n$ which has the property that $I_n A = A I_n = A$. This matrix $I_n$ is the matrix whose diagonal entries are 1s and whose other entries are 0s.

  ○ <u>Example</u>: The $2 \times 2$ identity matrix is $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and the $3 \times 3$ identity matrix is $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

  ○ Observe that $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ for any $2 \times 2$ matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

- Each of the elementary row operations on an $n \times n$ matrix corresponds to left-multiplication by the matrix obtained by applying the corresponding row operation to the identity matrix:

  1. Interchanging the $i$th and $j$th rows is equivalent to multiplying by the matrix which is the identity matrix, except that the $i$th and $j$th rows have been interchanged.

     ○ <u>Example</u>: The $3 \times 3$ matrix with rows 2 and 3 swapped is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, and indeed we can compute $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$.

  2. Multiplying all entries in the $i$th row by the constant $\alpha$ is equivalent to multiplying by the matrix which is the identity matrix, except that the $(i, i)$ entry is $\alpha$ rather than 1.

     ○ <u>Example</u>: The $2 \times 2$ matrix with row 2 doubled is $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, and indeed $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix}$.

  3. Adding $\alpha$ times the $j$th row to the $i$th row is equivalent to multiplying by the matrix which is the identity matrix, except the $(i, j)$ entry is $\alpha$ rather than 0.

     ○ <u>Example</u>: The $2 \times 2$ matrix with twice row 2 added to row 1 is $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, and indeed $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a + 2c & b + 2d \\ c & d \end{bmatrix}$.

- The matrices appearing above have a special name:

- <u>Definition</u>: An <u>elementary row matrix</u> is a matrix obtained by performing a single elementary row operation on the identity matrix.

○ Examples: The matrices $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, and $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$ are elementary matrices associated to the operations "add twice row 2 to row 1", "double row 1", "swap rows 2 and 3", and "add 3 times row 2 to row 3", respectively.

## 1.3 Determinants and Inverses

- In this section we will discuss some important quantities relevant to matrix multiplication: the inverse of a matrix and the determinant of a matrix.

### 1.3.1 The Inverse of a Matrix

- Given a square $n \times n$ matrix $A$, we might like to know whether it has a multiplicative inverse.

- Definition: If $A$ is an $n \times n$ square matrix, then we say $A$ is invertible (or nonsingular) if there exists an $n \times n$ matrix $A^{-1}$, the inverse matrix, such that $AA^{-1} = A^{-1}A = I_n$, where $I_n$ is the $n \times n$ identity matrix. If no such matrix $A^{-1}$ exists, we say $A$ is not invertible (or singular).

  ○ Example: The matrix $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ has inverse matrix $A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$, since we can compute
  $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$.

- Not every matrix has a multiplicative inverse.

  ○ Obviously, $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ does not have a multiplicative inverse, since $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ times any matrix is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, and cannot be the identity matrix.

  ○ Similarly observe that $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}$, which is never equal to the identity matrix for any choice of $a, b, c, d$ since the top and bottom rows are always equal. Thus, $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ does not have an inverse either.

- Here are a few basic properties of inverse matrices:

- Proposition (Uniqueness of Inverse): If a matrix is invertible, then it has only one inverse matrix.

  ○ Proof: Suppose $B_1$ and $B_2$ both had $AB_1 = I_n = B_1A$ and $AB_2 = I_n = B_2A$.
  ○ Then $B_1 = B_1I_n = B_1(AB_2) = (B_1A)B_2 = I_nB_2 = B_2$ and so $B_1 = B_2$.

- Proposition (Inverse of Product): If $A$ and $B$ are invertible $n \times n$ matrices with inverses $A^{-1}$ and $B^{-1}$, then $AB$ is also an invertible matrix with inverse $B^{-1}A^{-1}$.

  ○ Proof: We simply compute $(AB)(B^{-1}A^{-1}) = A(B\,B^{-1})A^{-1} = A(I_n)A^{-1} = A\,A^{-1} = I_n$. Similarly, the product in the other order will also come out to be the identity matrix.
  ○ We can extend this result to show that $(A_1A_2 \cdots A_n)^{-1} = (A_n)^{-1} \cdots (A_2)^{-1}(A_1)^{-1}$, provided that each of $A_1, A_2, \ldots, A_n$ is invertible.

- Proposition: Any square matrix with a row or column of all zeroes cannot be invertible.

  ○ Proof: Suppose the $n \times n$ matrix $A$ has all entries in its $i$th row equal to zero. Then for any $n \times n$ matrix $B$, the product $AB$ will have all entries in its $i$th row equal to zero, so it cannot be the identity matrix.
  ○ Similarly, if the $n \times n$ matrix $A$ has all entries in its $i$th column equal to zero, then for any $n \times n$ matrix $B$, the product $BA$ will have all entries in its $i$th column equal to zero.

- <u>Proposition</u> (Inverses of $2 \times 2$ Matrices): The $2 \times 2$ the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$, and if so, the inverse is given by $\dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

  - <u>Proof</u>: This follows simply from solving the system of equations for $e, f, g, h$ in terms of $a, b, c, d$ that arises from comparing entries in the product $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$: one obtains precisely the solution given above. If $ad = bc$ then the system is inconsistent and there is no solution; otherwise there is exactly one solution as given.

- Using row operations we can give a simple criterion for deciding whether a matrix is invertible:

- <u>Theorem</u> (Invertible Matrices): An $n \times n$ matrix $A$ is invertible if and only if it is row-equivalent to the identity matrix $I_n$.

  - <u>Proof</u>: Consider the reduced row-echelon form of the matrix $A$. Because $A$ is a square matrix, the reduced row-echelon form is either the identity matrix, or a matrix with a row of all zeroes.
  - Suppose $A$ is row-equivalent to the identity matrix. Each elementary row operation corresponds to left-multiplication by an invertible matrix, so there are elementary matrices $E_i$ with $1 \leq i \leq k$ such that $E_k E_{k-1} \cdots E_1 A = I_n$.
  - So if we let $B = E_k E_{k-1} \cdots E_1$, then $B$ is invertible (its inverse is $B^{-1} = E_1^{-1} \cdots E_{k-1}^{-1} E_k^{-1}$) and $BA = I_n$.
  - Multiplying the expression $BA = I_n$ on the left by $B^{-1}$ and on the right by $B$ produces $AB = B^{-1} B = I_n$, so we see $AB = BA = I_n$. Thus $B$ is the inverse of $A$, as claimed.
  - Now suppose that $A$ is not row-equivalent to the identity matrix. Then its reduced row-echelon form $A_{red}$ must contain a row of all zero entries. From our results above we see that $A_{red}$ cannot be invertible, and since $A = E_1 E_2 \ldots E_k A_{red}$, then if $A$ had an inverse $B$ then $A_{red}$ would have an inverse, namely $B E_1 E_2 \cdots E_k$.

- From the proof of this theorem we see that if $A$ has an inverse, we can compute it as the composition of the appropriate row operations that convert $A$ into the identity matrix. Explicitly, in order to compute the inverse of an $n \times n$ matrix $A$ using row reduction (or to see if it is non-invertible), we can perform the following procedure, called <u>Gauss-Jordan elimination</u>:

  - <u>Step 1</u>: Set up a "double" matrix $[A \,|\, I_n]$ where $I_n$ is the identity matrix.
  - <u>Step 2</u>: Perform row operations to put $A$ in reduced row-echelon form. (Carry the computations through on the entire matrix, but only pay attention to the left side when deciding what operations to do.)
  - <u>Step 3</u>: If $A$ can be row-reduced to the $n \times n$ identity matrix, then row-reducing $A$ will produce the double matrix $[I_n \,|\, A^{-1}]$. If $A$ cannot be row-reduced to the $n \times n$ identity matrix, then $A$ is not invertible.

- <u>Example</u>: Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -1 & 1 \\ 0 & 2 & -5 \end{bmatrix}$.

  - First, we set up the starting matrix $\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ 0 & 2 & -5 & 0 & 0 & 1 \end{array}\right]$.
  - Now we perform row operations to row-reduce the matrix on the left:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ 0 & 2 & -5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 3 & -2 & 1 & 0 \\ 0 & 2 & -5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 + 2R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 3 & -2 & 1 & 0 \\ 0 & 0 & 1 & -4 & 2 & 1 \end{array}\right]$$

$$\xrightarrow[R_1 + R_3]{R_2 - 3R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & -1 & 0 & 10 & -5 & -3 \\ 0 & 0 & 1 & -4 & 2 & 1 \end{array}\right] \xrightarrow{(-1)R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & 1 & 0 & -10 & 5 & 3 \\ 0 & 0 & 1 & -4 & 2 & 1 \end{array}\right].$$

  - We have row-reduced $A$ to the identity matrix, so $A$ is invertible and $\boxed{A^{-1} = \begin{bmatrix} -3 & 2 & 1 \\ -10 & 5 & 3 \\ -4 & 2 & 1 \end{bmatrix}}$.

### 1.3.2 The Determinant of a Matrix

- We might like to know, without performing all of the row-reductions, if a given large matrix is invertible. This motivates the idea of the determinant, which will tell us precisely when a matrix is invertible.

- Definition: The determinant of a square matrix $A$, denoted $\det(A)$ or $|A|$, is defined inductively. For a $1 \times 1$ matrix $[a]$ it is just the constant $a$. For an $n \times n$ matrix we compute the determinant via "cofactor expansion": define $A^{(1,k)}$ to be the matrix obtained from $A$ by deleting the 1st row and $k$th column. Then $\det(A) = \sum_{k=1}^{n} (-1)^{k+1} a_{1,k} \det(A^{(1,k)})$.

  $\circ$ The best way to understand determinants is to work out some examples.

- Example: The determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ is given by $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

  $\circ$ So, as particular cases, $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = (1)(4) - (2)(3) = -2$ and $\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = (1)(2) - (1)(2) = \boxed{0}$.

- Example: The determinant $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ is given by $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} +$ $a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$.

  $\circ$ As a particular case, $\begin{vmatrix} 1 & 2 & 4 \\ -1 & 1 & 0 \\ -2 & 1 & 3 \end{vmatrix} = 1 \begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} - 2 \begin{vmatrix} -1 & 0 \\ -2 & 3 \end{vmatrix} + 4 \begin{vmatrix} -1 & 1 \\ -2 & 1 \end{vmatrix} = 1(3) - 2(-3) + 4(1) = \boxed{13}$.

- There is a nice way to interpret the determinant geometrically:

- Proposition: If $\mathbf{v}_1, \mathbf{v}_2$ are vectors in $\mathbb{R}^2$, then the determinant of the matrix whose rows are $\mathbf{v}_1, \mathbf{v}_2$ is the signed area of the parallelogram formed by $\mathbf{v}_1, \mathbf{v}_2$. Furthermore, if $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ are vectors in $\mathbb{R}^3$, then the determinant of the matrix whose rows are $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ is the signed volume of the parallelepiped (skew box) formed by $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$.

  $\circ$ Note that a signed area (and a signed volume) can be negative: the sign indicates the relative orientation of the vectors. For two vectors, the signed area is positive if the second vector is counterclockwise from the first one and negative otherwise. For three vectors, the signed volume is positive if the vectors are arranged per the right-hand rule: align the fingers of the right hand along the first vector and then curl them into the direction of the second vector; the orientation is positive if the thumb is pointing in the direction of the third vector, and negative if it is pointing in the opposite direction.

  $\circ$ Here are pictures of the appropriate regions:
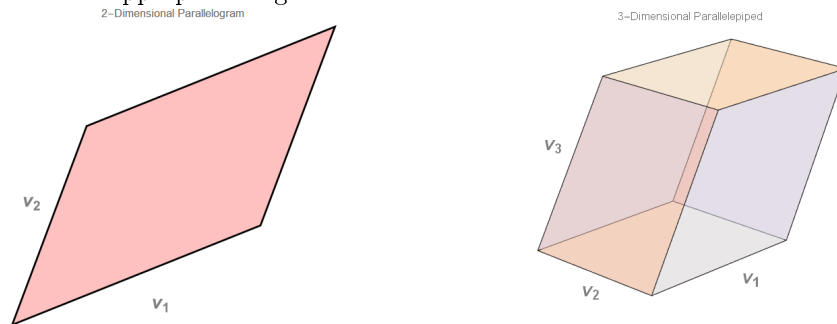


  $\circ$ The proof is a straightforward geometric calculation, which we omit.

- Example: Find the volume of the skew box formed by the vectors $\mathbf{w}_1 = \langle 1, 0, 1 \rangle$, $\mathbf{w}_2 = \langle 2, 1, 1 \rangle$, and $\mathbf{w}_3 = \langle 0, 2, 2 \rangle$.

- ○ We simply compute the appropriate determinant: $\begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 2 & 2 \end{vmatrix} = 1 \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} - 0 \begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} + 1 \begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} =$ $0 - 0 + 4 = \boxed{4}$.

- The determinant behaves in a very predictable way under the elementary row operations (showing these results requires a more careful technical analysis of the determinant and so we will omit the details):

  - ○ Interchanging two rows multiplies the determinant by $-1$.
    - ∗ <u>Example</u>: $\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} = 5$ while $\begin{vmatrix} -1 & 1 \\ 3 & 2 \end{vmatrix} = -5$.
  - ○ Multiplying all entries in one row by a constant scales the determinant by the same constant.
    - ∗ <u>Example</u>: $\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} = 5$ while $\begin{vmatrix} 9 & 6 \\ -1 & 1 \end{vmatrix} = 3 \cdot 5 = 15$.
  - ○ Adding or subtracting a scalar multiple of one row to another leaves the determinant unchanged.
    - ∗ <u>Example</u>: $\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} = 5$ while $\begin{vmatrix} 3+2(-1) & 2+2(1) \\ -1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ -1 & 1 \end{vmatrix} = 5$.

- From the above analysis we can deduce a number of other properties of determinants:

- <u>Property</u>: If a matrix has a row or column of all zeroes, its determinant is zero.

  - ○ <u>Example</u>: $\begin{vmatrix} 3 & 2 \\ 0 & 0 \end{vmatrix} = 0$ and $\begin{vmatrix} 3 & 0 \\ -1 & 0 \end{vmatrix} = 0$.
  - ○ For rows, this follows by observing that the matrix is unchanged upon multiplying the row of zeroes by the scalar 0. For columns, applying the definition will yield $n-1$ smaller determinants, each of which will have first column containing all zeroes.

- <u>Property</u>: If one row or column is a scalar multiple of another, then its determinant is zero. (In particular, if two rows are equal, the determinant is zero.) More generally, if the matrix is row-equivalent to a matrix with a row or column of all zeroes, then the determinant is zero.

  - ○ <u>Example</u>: $\begin{vmatrix} 3 & 2 & 1 \\ 6 & 4 & 2 \\ 1 & -1 & 1 \end{vmatrix} = 0$ (rows 1,2) and $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 1 & 1 & 3 \end{vmatrix} = 0$ (columns 1,3).
  - ○ The first statement follows from the second one, since if one row is a scalar multiple of another, we can use a row operation to form a matrix with a row of all zeroes. The second statement follows by the observation that applying a row operation will scale the determinant by a nonzero constant, and thus cannot change a nonzero determinant into a zero determinant or vice versa.

- <u>Property</u>: The determinant is multiplicative: $\det(AB) = \det(A)\det(B)$.

  - ○ <u>Example</u>: If $A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$ with $\det(A) = 2$ and $B = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$ with $\det(B) = 3$, then $AB = \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$ with $\det(AB) = 6$.
  - ○ <u>Example</u>: If $A = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$ with $\det(A) = 5$ and $B = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$ with $\det(B) = 10$, then $AB = \begin{bmatrix} 9 & 2 \\ 2 & 6 \end{bmatrix}$ with $\det(AB) = 50$.
  - ○ This can be shown by row-reducing the matrix $A$ and using the facts about how the elementary row operations affect the determinant.

- <u>Property</u>: The determinant of any upper-triangular matrix (a matrix whose entries below the diagonal are all zeroes) is equal to the product of the diagonal entries. In particular, the determinant of the identity matrix is 1.

- ○ Example: $\begin{vmatrix} 6 & -1 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 36.$

  - ○ Using the definition of the determinant will produce $n$ smaller determinants, and each determinant except the first one has first column all zeroes (and is therefore zero).

- **Property**: A matrix $A$ is invertible precisely when $\det(A) \neq 0$.

  - ○ We can see this by applying the elementary row operations to put a matrix in reduced row-echelon form: if the matrix is invertible then the reduced row-echelon form will be the identity matrix, and the determinant calculations will yield a nonzero result.
  - ○ If the matrix is not invertible, then its reduced row-echelon form will have a row of all zeroes, and the determinant of such a matrix is zero.

- **Property**: If $A$ is invertible, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

  - ○ Apply $\det(AB) = \det(A)\det(B)$ when $B = A^{-1}$: we get $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$.

- **Property**: The determinant of the transpose matrix is the same as the original determinant: $\det(A^T) = \det(A)$.

  - ○ Example: $\begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix} = 5$ and $\begin{vmatrix} 3 & -1 \\ 2 & 1 \end{vmatrix} = 5$ also.

- **Property**: The determinant is "linear" in columns: if $A = \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{a}_i & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix}$, $B = \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b}_i & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix}$, and $C = \begin{bmatrix} | & & | & | & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{c}_i & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \\ | & & | & | & | & & | \end{bmatrix}$, where the $\mathbf{a}_j$ are column vectors and $\mathbf{a}_i = \mathbf{b}_i + \mathbf{c}_i$, then $\det(A) = \det(B) + \det(C)$. The same property also holds for rows.

  - ○ Example: If $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$, where we decompose the second column as $\begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}$: then $B = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. We compute $\det(A) = 1$, $\det(B) = 2$, $\det(C) = -1$, and indeed $\det(A) = \det(B) + \det(C)$.

### 1.3.3 Cofactor Expansions and the Adjugate

- There are a few other results about determinants that require an additional definition to state:

- **Definition**: If $A$ is a square $n \times n$ matrix, define $A^{(j,k)}$ to be the matrix obtained from $A$ by deleting the $j$th row and $k$th column. The $(j,k)$ cofactor of $A$, $C^{(j,k)}$, is defined to be $(-1)^{j+k}\det(A^{(j,k)})$.

  - ○ Example: The $(1,1)$ cofactor of $\begin{bmatrix} 6 & -1 & 3 \\ 0 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$ is $(-1)^{1+1}\begin{vmatrix} 2 & 0 \\ 0 & 3 \end{vmatrix} = (-1)^2 6 = 6$, and the $(2,3)$ cofactor is $(-1)^{2+3}\begin{vmatrix} 6 & -1 \\ 3 & 0 \end{vmatrix} = (-1)^5 3 = -3$.

- **Theorem** (Expansion by Minors): If $A$ is a square $n \times n$ matrix, then for any fixed $j$, $\det(A) = \displaystyle\sum_{k=1}^{n} a_{j,k} C^{(j,k)}$ and also $\det(A) = \displaystyle\sum_{i=1}^{n} a_{i,j} C^{(i,j)}$.

- The statement of this theorem requires some unpacking. Essentially, the idea is that we can compute the determinant by expanding along any row, rather than just the first row (as in the original definition), or along any column.

- The only difficulty is remembering which terms have which sign (plus or minus). Each term has a particular sign based on its location in the matrix, as follows: the $(1, 1)$ entry has a plus sign, and the remaining elements are filled in in an alternating "checkerboard" pattern: $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$.

- The calculation of the determinant this way is called "expansion by minors".

  - Expanding along the second row: $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = -b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}$.

  - Expanding down the third column: $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$.

- When choosing a row or column to expand along, it is best to choose one with many zeroes, as this will reduce the number of smaller determinants that need to be evaluated.

- <u>Example</u>: Find the determinant $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 \\ 0 & 3 & 4 & 0 \end{vmatrix}$.

  - We will start by expanding along the second row: this will give us a single $3 \times 3$ determinant, which we can then evaluate by expanding along its bottom row:

  $$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 1 \\ 0 & 3 & 4 & 0 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 3 & 4 \\ 2 & 1 & 1 \\ 0 & 4 & 0 \end{vmatrix} = (-4) \begin{vmatrix} 1 & 4 \\ 2 & 1 \end{vmatrix} = (-4)(-7) = \boxed{28}.$$

- Although expansion by minors (or even just the definition of the determinant) gives a recursive method for computing any $n \times n$ determinant, these methods are quite slow unless the matrix has many zero entries.

  - Evaluating an $n \times n$ determinant using the definition requires computing $n$ total $(n - 1) \times (n - 1)$ determinants, each of which requires evaluating $(n - 1)$ total $(n - 2) \times (n - 2)$ determinants.

  - Continuing in this way, we see that evaluating an $n \times n$ determinant from the definition requires $n!$ total computations (where we say a $1 \times 1$ determinant is one computation).

  - Since $5! = 120$, it is already quite unreasonable to compute a $5 \times 5$ determinant by hand using this method, while a $10 \times 10$ determinant (note $10! = 3628800$) is entirely out of reach, and even a computer would have trouble with a $30 \times 30$ determinant ($30! = 2.65 \cdot 10^{32}$).

- Row-reduction is a far more efficient method for computing large determinants.

  - It is sufficient to row-reduce a matrix to put it into row-echelon form, since any row-echelon matrix is upper-triangular, and the determinant of an upper triangular matrix is simply the product of the diagonal entries.

  - Row-reducing an $n \times n$ matrix requires approximately $n^2$ individual multiplications, although due to the fact that the sizes of the entries in the matrix can grow quite large (if one is trying to avoid introducing denominators), the total number of calculations is on the order of $n^3$.

  - For $n = 5$, one can typically row-reduce a matrix by hand to compute a determinant, and even a $10 \times 10$ determinant (approximately a few hundred computations) would not be impossible by hand. A computer can easily deal with a $1000 \times 1000$ determinant using row-reductions.

- Example: Find the determinant $\begin{vmatrix} 1 & 2 & -1 & 3 \\ 3 & 7 & 0 & 4 \\ -2 & 1 & 1 & 2 \\ -1 & 3 & 16 & 5 \end{vmatrix}$.

  ○ By row-reducing, $\begin{vmatrix} 1 & 2 & -1 & 3 \\ 3 & 7 & 0 & 4 \\ -2 & 1 & 1 & 2 \\ -1 & 3 & 16 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 5 & -1 & 8 \\ 0 & 5 & 15 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 3 & -7 \\ 0 & 0 & -16 & 43 \\ 0 & 0 & 0 & 36 \end{vmatrix} = \boxed{-576}$.

- There is also a formula for the inverse of a matrix in terms of its cofactors:

- Theorem (Matrix Inverse Formula): If the matrix $A$ is invertible, so that $\det(A) \neq 0$, we have the formula $A^{-1} = \dfrac{1}{\det(A)} [\mathrm{adj}(A)]$, where $\mathrm{adj}(A)$ is the matrix whose $(i,j)$-entry is given by the $(j,i)$ cofactor $C^{(j,i)}$ of $A$.

  ○ The name $\mathrm{adj}(A)$ is short for adjugate. (Some authors refer to this matrix as the "adjoint", but that term is usually reserved to mean something else in modern usage.)

  ○ Proof: Since $A$ is invertible and the inverse of a matrix is unique, it is sufficient to show that the product $A \cdot \mathrm{adj}(A)$ is equal to $\det(A)$ times the identity matrix.

  ○ First consider the $(k,k)$ entry in the product $A \cdot \mathrm{adj}(A)$: it is the sum $\displaystyle\sum_{l=1}^{n} a_{k,l} C^{(k,l)}$, which is the expansion of the determinant of the matrix $A$ along the $k$th row. So the $(k,k)$ entry is equal to $\det(A)$ for each $k$.

  ○ Now consider the $(i,j)$ entry of the product, for $i \neq j$: it is the sum $\displaystyle\sum_{i=1}^{n} a_{i,l} C^{(j,l)}$, which is the expansion of the determinant of the matrix obtained by replacing the $j$th row of $A$ with the $i$th one, along its $i$th row. This determinant is zero since the matrix has two equal rows.

- Example: Use the adjugate formula to compute the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -1 & 1 \\ 0 & 2 & -5 \end{bmatrix}$.

  ○ First, we have $\det(A) = 1 \begin{vmatrix} -1 & 1 \\ 2 & -5 \end{vmatrix} - 0 \begin{vmatrix} 2 & 1 \\ 0 & -5 \end{vmatrix} - 1 \begin{vmatrix} 2 & -1 \\ 0 & 2 \end{vmatrix} = 3 - 4 = -1$.

  ○ Now we compute all of the entries of the adjugate matrix:

  $$\mathrm{adj}(A) = \begin{bmatrix} +\begin{vmatrix} -1 & 1 \\ 2 & -5 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 0 & -5 \end{vmatrix} & +\begin{vmatrix} 2 & -1 \\ 0 & 2 \end{vmatrix} \\ -\begin{vmatrix} 0 & -1 \\ 2 & -5 \end{vmatrix} & +\begin{vmatrix} 1 & -1 \\ 0 & -5 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} \\ +\begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} & +\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 3 & -2 & -1 \\ 10 & -5 & -3 \\ 4 & -2 & -1 \end{bmatrix}.$$

  ○ Thus, we have $\mathrm{adj}(A) = \begin{bmatrix} 3 & -2 & -1 \\ 10 & -5 & -3 \\ 4 & -2 & -1 \end{bmatrix}$. Since $\det(A) = -1$, we thus obtain $\boxed{A^{-1} = \begin{bmatrix} -3 & 2 & 1 \\ -10 & 5 & 3 \\ -4 & 2 & 1 \end{bmatrix}}$.

  ○ We can of course check this answer by computing $AA^{-1}$ and verifying it is the identity matrix (which it is).

- Although the adjugate formula does give an explicit formula for the inverse, it is not computationally useful: it is much faster to compute $A^{-1}$ using row reductions.

  ○ Using the adjugate formula requires finding an $n \times n$ determinant and $n^2$ total $(n-1) \times (n-1)$ determinants, so even for a $3 \times 3$ matrix, the adjugate formula is far less efficient than row reduction.

  ○ However, the adjugate formula does yield one important piece of information: if the entries of the matrix are integers, the entries of the inverse will be rational numbers whose denominators are (at largest) given by the determinant of the original matrix.

18

## 1.4  Matrices and Systems of Linear Equations, Revisited

- As an application of the utility of the matrix approach, let us revisit systems of linear equations.

- First suppose we have a system of $k$ equations in $n$ variables:

$$
\begin{aligned}
a_{1,1}x_1 + \cdots + a_{n,1}x_n &= c_1 \\
&\vdots \qquad \vdots \\
a_{1,k}x_1 + \cdots + a_{n,k}x_n &= c_k
\end{aligned}
$$

  - Let $A = \begin{bmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & \ddots & \vdots \\ a_{1,k} & \cdots & a_{n,k} \end{bmatrix}$ be the matrix of coefficients, $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}$ the column vector of variables,

    and $\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$ the coefficient vector.

  - Then we can rewrite the system as the much simpler matrix equation $A\mathbf{x} = \mathbf{c}$. We can use our results about matrices to say things about the solutions to such systems.

- Definition: A matrix equation $A\mathbf{x} = \mathbf{c}$ is homogeneous if $\mathbf{c}$ is the zero vector, and it is nonhomogeneous otherwise.

- Proposition: Suppose $\mathbf{x}_p$ is one solution to the matrix equation $A\mathbf{x} = \mathbf{c}$. Then the general solution to this equation may be written as $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_{hom}$, where $\mathbf{x}_{hom}$ is a solution to the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

  - This proposition says that if we can find *one* solution to the original system, then we can find *all* of them just by solving the homogeneous system.

  - Proof: To see this, first observe that if $A\mathbf{x}_p = \mathbf{c}$ and $A\mathbf{x}_{hom} = \mathbf{0}$, then

$$
A(\mathbf{x}_p + \mathbf{x}_{hom}) = A\mathbf{x}_p + A\mathbf{x}_{hom} = \mathbf{c} + \mathbf{0} = \mathbf{c}
$$

    and so $\mathbf{x}_p + \mathbf{x}_{hom}$ is also a solution to the original system.

  - Conversely, if $\mathbf{x}_1$ and $\mathbf{x}_2$ are two solutions to the original system, then $A(\mathbf{x}_2 - \mathbf{x}_1) = A\mathbf{x}_2 - A\mathbf{x}_1 = \mathbf{c} - \mathbf{c} = \mathbf{0}$, so that $\mathbf{x}_2 - \mathbf{x}_1$ is a solution to the homogeneous system.

- We now mention a few properties of the solutions to a homogeneous equation $A\mathbf{x} = \mathbf{0}$:

- Proposition: The zero vector $\mathbf{0}$ is always a solution to any homogeneous equation $A\mathbf{x} = \mathbf{0}$. Also, if $\mathbf{x}_1$ and $\mathbf{x}_2$ are two solutions, then $\mathbf{x}_1 + \mathbf{x}_2$ is also a solution, as is $r\mathbf{x}_1$ for any real number scalar $r$.

  - Proof: We clearly have $A\mathbf{0} = \mathbf{0}$ so $\mathbf{0}$ is a solution.

  - Next, if $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions, then $A\mathbf{x}_1 = A\mathbf{x}_2 = \mathbf{0}$: then $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$, so $\mathbf{x}_1 + \mathbf{x}_2$ is also a solution,

  - Finally, we also have $A(r\mathbf{x}_1) = r(A\mathbf{x}_1) = r\mathbf{0} = \mathbf{0}$, so $r\mathbf{x}_1$ is also a solution.

- We can use some of our results about inverses and determinants when the coefficient matrix is square.

- If $A$ is not invertible, then the homogeneous system has infinitely many solutions (as the reduced row-echelon form of $A$ must have a row of all zeroes, and hence at least one free variable). The original system can then either have no solutions or infinitely many solutions, as we saw.

  - In this case we cannot really hope to write down a formula for the solutions, although of course we can still compute them using Gauss-Jordan elimination.

- In the event that the coefficient matrix is invertible we can write down the solutions to the system using determinants:

- <u>Theorem</u> (Cramer's Rule): If $A$ is an invertible $n \times n$ matrix, then the matrix equation $A\mathbf{x} = \mathbf{c}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{c}$. Specifically, the $i$th element of $\mathbf{x}$ is given by $x_i = \dfrac{\det(C_i)}{\det(A)}$, where $C_i$ is the matrix obtained by replacing the $i$th column of $A$ with the column vector $\mathbf{c}$.

    - <u>Proof</u>: Suppose $A$ is invertible: then we can multiply both sides of the equation $A\mathbf{x} = \mathbf{c}$ on the left by $A^{-1}$ to see that $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{c}$.

    - We can then write $\mathbf{x} = I_n\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{c}$. In particular, we see that the solution is unique, since $A^{-1}$ is unique.

    - For the other part, we use the adjugate formula for $A^{-1}$: recall that we showed $A^{-1} = \dfrac{1}{\det(A)}\mathrm{adj}(A)$, so we see $\mathbf{x} = \dfrac{\mathrm{adj}(A)\mathbf{c}}{\det(A)}$.

    - Upon writing out the product in the numerator we see that the $i$th element is $x_i = \displaystyle\sum_{k=1}^{n} c_i(-1)^{k+i}\det(A^{(i,k)})$, and this is the expansion by minors along the $i$th column for the determinant of the matrix $C_i$.

    - Therefore, $x_i = \dfrac{\det(C_i)}{\det(A)}$ as claimed.

- <u>Example</u>: Solve the system of equations $3x + z = 0$, $x + 2y - 3z = 1$, $2x - 2y - z = 2$ using Cramer's rule.

    - The coefficient matrix is $C = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -3 \\ 2 & -2 & -1 \end{bmatrix}$ whose determinant is $\det(C) = 3\begin{vmatrix} 2 & -3 \\ -2 & -1 \end{vmatrix} + 1\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = -30$.

    - Since this matrix is invertible the system will have a unique solution.

    - We have $C_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & -3 \\ 2 & -2 & -1 \end{bmatrix}$, $C_2 = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 1 & -3 \\ 2 & 2 & -1 \end{bmatrix}$, and $C_3 = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & -2 & 2 \end{bmatrix}$, and the respective determinants are $\det(C_1) = 1\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = -6$, $\det(C_2) = 3\begin{vmatrix} 1 & -3 \\ 2 & -1 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 15$, and $\det(C_3) = 3\begin{vmatrix} 2 & 1 \\ -2 & 2 \end{vmatrix} = 18$.

    - Thus, by Cramer's rule, the solution is $(x, y, z) = \left( \dfrac{-6}{-30}, \dfrac{15}{-30}, \dfrac{18}{-30} \right) = \boxed{\left( \dfrac{1}{5}, -\dfrac{1}{2}, -\dfrac{3}{5} \right)}$.

- As with the other formulas involving determinants, Cramer's rule is not particularly useful for practical computation, at least when the coefficients of the system are real numbers.

    - In addition to requiring the coefficient matrix to be square and invertible, the total amount of computation is much larger: solving an $n \times n$ system with Cramer's rule requires computing $n + 1$ total $n \times n$ determinants, while (in comparison) solving the system via row-reduction directly requires only row-reducing one $n \times n$ matrix.

    - We will remark that if the coefficients of the system are functions (as can arise, for example, in solving certain kinds of differential equations), then Cramer's rule can end up being more useful.

- Cramer's rule does have one useful theoretical consequence: if the entries of the matrix $A$ are integers, the solution vector to the system will have rational number entries whose denominators divide $\det(A)$.

---

Well, you're at the end of my handout. Hope it was helpful.