Linear Algebra Study Guide

Systems of Linear Equations

A system of linear equations in *n* variables is a set of *m* equations, each of which is linear in the same n variables. The double subscript notation indicates a_{ij} is the coefficient of x_j in the *i*th equation.

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = b_1 \\
&a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n = b_2 \\
&a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + ... + a_{3n}x_n = b_3 \\
&\qquad\qquad\qquad\qquad\quad . \\
&\qquad\qquad\qquad\qquad\quad . \\
&\qquad\qquad\qquad\qquad\quad . \\
&a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ... + a_{mn}x_n = b_n
\end{aligned}
$$

Systems of two Equations in two Variables

If there is one solution the equations are a consistent system and the lines of the graphs will intersect (intersecting lines, one solution). If all the values of one equation satisfy the other equation, then the two consistent equations are called dependent equations and the lines of their graphs will coincide (coincident lines, infinite solutions). To **coincide** is to correspond exactly or occupy the same place in space. If there is no set of values that will satisfy both equations they are called inconsistent equations and the lines of the graph s will be parallel (parallel lines, no solutions).

Consistency of systems of equations is the property possessed by a system of equations when there is at least one set of values of the variables that satisfies each equation, for example the solution sets have one or more common points. If they are not satisfied by any one set of values of the variables they are inconsistent.

There are three methods for solving systems of two equations in two variables. You can solve by graphing to determine whether the lines of the graph intersect, are coincident, or are parallel. You can also add or subtract the equations to solve for one variable which can then be substituted back into the original equation to solve for the other variable. For solving larger systems Gauss Jordan elimination is used.

Operations that lead to equivalent systems of equations:

1. Interchange two equations
2. Multiply an equation by a nonzero constant
3. Add a multiple of an equation to another equation

Using the three basic operations of Gaussian elimination, work from the upper left corner of the system saving the *x* variable in the upper left position (making the *x* variable equal to one) and eliminating the other x variables (making the other x variables equal to zero) from the first column.

Properties of a Matrix in Row-echelon Form

1. All rows consisting entirely of zeros occur at the bottom of the matrix.
2. For each row that does not consist entirely of zeros, the first nonzero entry is one, called a leading one.
3. For two successive nonzero rows, the leading one in the higher row is farther to the left than the leading one in the lower row.

A system of equations is **homogeneous** when each of the constant terms is zero. A homogeneous equation in n variables has the form $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n = 0$. Every homogeneous system of linear equations is consistent and must have at least one solution. If the system has fewer equations than variables, it must have an infinite number of solutions. Solutions in which at least one of the variables has a value different from zero are nontrivial solutions. If all the variables in a homogeneous system have the value zero they are referred to as trivial solutions of a set of homogeneous linear equations.

Matrices

A **matrix** is a rectangular array of numbers written using brackets. Matrices are written in row and column format ($r \times c$, or $m \times n$). The individual elements of a matrix are labeled using a_{ij} notation, where i is called the row index and j the column index. Element a_{ij} is in the ith row and the jth column of the matrix.

$$A = \left[a_{ij}\right] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & ... & a_{1n} \\ a_{21} & a_{22} & a_{23} & ... & a_{2n} \\ a_{31} & a_{32} & a_{33} & ... & a_{3n} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & a_{m2} & a_{m3} & ... & a_{mn} \end{bmatrix}$$

Add example matrix showing a_{ij} with numeric subscripts (e.g. a_{11})

Matrix Operations

Addition and Subtraction of Matrices

Given two m x n matrices $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$, their sum is $A + B = \left[a_{ij} + b_{ij}\right]$ and their difference is $A - B = \left[a_{ij} - b_{ij}\right]$. To add and subtract matrices they must be the same order or size. To calculate the sum, add the elements in the corresponding positions of the matrix. To calculate the difference, subtract the elements in the corresponding positions of the matrix. Addition of matrices is both commutative and associative.

Matrix Multiplication

For an $m \times n$ matrix $\mathrm{A} = \left[a_{ij}\right]$ and an $n \times p$ matrix $\mathrm{B} = \left[b_{ij}\right]$, the product $\mathrm{AB} = \left[c_{ij}\right]$ is an $m \times p$ matrix, where $c_{ij} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + a_{i3} \cdot b_{3j} + ... + a_{in} \cdot b_{nj}$ The entry c_{ij} in **AB** is obtained by multiplying the entries in row *i* of **A** by the corresponding entries in column *j* of **B** and adding the results. In other words, multiply the rows times the columns and add. To multiply matrices, the number of columns in the first matrix must be equal to the number of rows in the second matrix.

Scalar Multiplication

The **scalar product** of a number *k* and a matrix **A** is the matrix denoted *k***A**, obtained by multiplying each entry of **A** by the number *k*. The number *k* is called a **scalar**. A **scalar** as contrasted with a vector, which represents both magnitude and direction, is a number that measures magnitude, positive or negative, but not direction.

Example: Find the scalar product of 3**A** where $\mathrm{A} = \begin{bmatrix} -3 & 0 \\ 4 & 5 \end{bmatrix}$.

$$3\mathrm{A} = 3\begin{bmatrix} -3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 3(-3) & 3(0) \\ 3(4) & 3(5) \end{bmatrix} = \begin{bmatrix} -9 & 0 \\ 12 & 15 \end{bmatrix}$$

Inverse Matrices

For an $m \times n$ matrix **A**, if there is a matrix $\mathbf{A}^{-1}$ for which $\mathrm{A}^{-1} \cdot \mathrm{A} = \mathrm{I} = \mathrm{A} \cdot \mathrm{A}^{-1}$ then $\mathbf{A}^{-1}$ is the inverse of **A**. To find the inverse matrix, form an augmented matrix consisting of an identity matrix of the same order and perform row equivalent operations as in the Gauss-Jordan elimination method. If the resulting matrix has an inverse it is called an inverted or nonsingular matrix. If a row consisting entirely of zeroes occurs in either of the two matrices in the augmented matrix the matrix has no inverse and is called a singular matrix. Not all matrices have an inverse. An **identity matrix** is a square matrix with *n* rows and *n* columns that contains zeroes for each of its elements except those along the diagonal that begins with the element in the first row and the first column. The symbol *I* is used to represent an identity matrix when its dimensions are not necessary and when the dimensions can be determined from the context. The symbol I_n represents the identity matrix of dimension *n* by *n*. For example:

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 2 x 2 identity matrix I_2 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 3 x 3 identity matrix I_3

LU Factorization

LU factorization also known as LU decomposition is defined as follows:

If a $n \times n$ matrix *A* can be written as the product of a lower triangular matrix *L* and upper triangular matrix *U*, then $A = LU$ is an LU factorization of *A*.

Given a random square matrix, *A*, the only way we can guarantee that *A* will have an LU decomposition is if we can reduce it to row-echelon form without interchanging any rows. If we do have to interchange rows then there is a good chance that the matrix will not have an LU Decomposition. There is no single unique LU-Decomposition for *A*.

Stochastic Matrices

Stochastic means regarding conjecture. Conjecture is an opinion, or judgment, formed on defective or presumptive evidence; to infer on slight evidence.

Matrix Solutions of Systems of Equations

For a system of *n* linear equations in *n* variables, **AX** = **B**, if **A** is an invertible matrix, then the unique solution of the system is given by $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$.

$\mathbf{AX} = \mathbf{B}$	
$\mathbf{A}^{-1}(\mathbf{AX}) = \mathbf{A}^{-1}\mathbf{B}$	Multiply by $\mathbf{A}^{-1}$ on the left on both sides
$(\mathbf{A}^{-1}\mathbf{A})\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$	Associative property of matrices
$\mathbf{IX} = \mathbf{A}^{-1}\mathbf{B}$	$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$
$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$	$\mathbf{IX} = \mathbf{X}$

Example: Solve the following system of equations using an inverse matrix:

$$-2x + 3y = 4$$
$$-3x + 4y = 5$$

$$\begin{bmatrix} -2 & 3 \\ -3 & 4 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{B}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\mathbf{A}^{-1} \cdot \mathbf{B} = \mathbf{X}$$

The solution to the system of equations is (1, 2).

Determinants

History: Determinants were originally considered without reference to matrices. A determinant was defined as a property of a system of linear equations. The determinant "determines" whether the system has a unique solution (which occurs precisely if the determinant is non-zero). In this sense, determinants were first used in the 3rd century BC Chinese mathematics textbook *The Nine Chapters on the Mathematical Art*.

In Europe, 2×2 determinants were considered by Cardano (1501 - 1576) at the end of the 16th century and larger ones by Leibniz (1646 - 1716). Vandermonde (1735 - 1796) was the first to recognize determinants as independent functions. Laplace (1749 - 1827) gave the general method of expanding a determinant in terms of its complementary minors. Following the work of Laplace, Lagrange (1736 - 1813) developed determinants of the second and third order. Cauchy (1789 – 1857) was the first to use the word determinant in its present sense. He also summarized and simplified what was then known on the subject, improved the notation, and provided a proof of the multiplication theorem.

Definition: A **determinant** is a square array of quantities, called elements, symbolizing the sum of certain products of these elements. The number of rows or columns is the order of the determinant. The determinant of a square matrix is a scalar calculated by multiplying, adding, and subtracting various elements of the matrix.

If a determinant has two rows and two columns, it is called a determinant of the second order. The value of this determinant is the difference of the products of its diagonals:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

The value of this determinant is the difference of the products of its diagonals: $ad - bc$

A third order determinant is an array of numbers that has three rows and three columns. To find the determinant of a system with an order (size) greater than three, you need to use minors and cofactors.

There are two ways to determine the determinant of a 3 x 3 matrix. The first method is the diagonal method and the second is cofactor expansion.

To apply the diagonal method, copy the first and second columns of A to form fourth and fifth columns. The determinant is obtained by adding and subtracting the products of the six diagonals.

$$\mathrm{A} = \begin{bmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & -4 & 1 \end{bmatrix} \qquad \begin{matrix} 0 & 2 & 1 & 0 & 2 \\ 3 & -1 & 2 & 3 & -1 \\ 4 & -4 & 1 & 4 & -4 \end{matrix}$$

$$\begin{array}{ccccc} 0 & 2 & 1 & 0 & 2 \\ 3 & -1 & 2 & 3 & -1 \\ 4 & -4 & 1 & 4 & -4 \end{array}$$

By adding the lower three products and subtracting the upper three products you can find the determinant.

$$|A| = 0 + 16 + (-12) - (-4) - 0 - 6 = 2$$

Minor is defined as something that is less or smaller. The **minor** of an element in a determinant is the determinant formed by striking out the row and column in which the element occurs. For a square matrix $\mathrm{A} = \left[a_{ij}\right]$, the minor M_{ij} of an element a_{ij} is the determinant of the matrix formed by deleting the *i*th row and *j*th column of **A**. The real number M_{ij} is the determinant of a sub-matrix of dimension $n \times 1$ by $n \times 1$ which contains everything except row *i* and column *j* of the original matrix. The number M_{ij} is called the minor for element *ij* of the matrix. For example: given the following matrix find the following minors; $\mathrm{M}_{31}, \mathrm{M}_{32}, \mathrm{M}_{33}$.

$$\mathrm{A} = \left[a_{ij}\right] = \begin{bmatrix} -8 & 0 & 6 \\ 4 & -6 & 7 \\ -1 & -3 & 5 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{-8} & 0 & 6 \\ \mathbf{4} & -6 & 7 \\ \mathbf{-1} & \mathbf{-3} & \mathbf{5} \end{bmatrix} \qquad \mathrm{M}_{31} = \begin{bmatrix} 0 & 6 \\ -6 & 7 \end{bmatrix} \qquad = 0 \cdot 7 - (-6)6 = 0 - (-36) = 36$$

$$\begin{bmatrix} -8 & \mathbf{0} & 6 \\ 4 & \mathbf{-6} & 7 \\ \mathbf{-1} & \mathbf{-3} & \mathbf{5} \end{bmatrix} \qquad \mathrm{M}_{32} = \begin{bmatrix} -8 & 6 \\ 4 & 7 \end{bmatrix} \qquad = -8 \cdot 7 - 4 \cdot 6 = -56 - 24 = -80$$

$$\begin{bmatrix} -8 & 0 & \mathbf{6} \\ 4 & -6 & \mathbf{7} \\ \mathbf{-1} & \mathbf{-3} & \mathbf{5} \end{bmatrix} \qquad \mathrm{M}_{33} = \begin{bmatrix} -8 & 6 \\ 4 & -6 \end{bmatrix} \qquad = -8 \cdot (-6) - 4 \cdot 0 = 48 - 0 = 48$$

The **cofactor** of an element in the *i*th row and *j*th column of a determinant is the value of the minor of that element if *i* + *j* is even, or the negative of the value of the minor of the element if $i + j$ is odd. For a square matrix $\mathrm{A} = \left[a_{ij}\right]$, the cofactor a_{ij} of an element a_{ij} is given by $A_{ij} = (-1)^{i+j} M_{ij}$, where M_{ij} is the minor of a_{ij}.

Example 2: Find the cofactors of the minors from example 1 where $\mathrm{M}_{31} = 36$, $\mathrm{M}_{32} = 80$, and $\mathrm{M}_{33} = 48$.

$$\mathrm{A}_{ij} = (-1)^{i+j}\mathrm{M}_{ij}$$

$$\begin{aligned} \mathrm{A}_{31} &= (-1)^{3+1}36 \\ &= (-1)^4 36 \\ &= 1\cdot 36 \\ &= 36 \end{aligned} \qquad \begin{aligned} \mathrm{A}_{32} &= (-1)^{3+2}(-80) \\ &= (-1)^5(-80) \\ &= (-1)(-80) \\ &= 80 \end{aligned} \qquad \begin{aligned} \mathrm{A}_{33} &= (-1)^{3+3}48 \\ &= (-1)^6 48 \\ &= 1\cdot 48 \\ &= 48 \end{aligned}$$

$$\begin{aligned} \left|\mathrm{A}\right| &= (-1)(36) + (-3)(80) + 5(48) \\ &= -36 - 240 + 240 \\ &= -36 \end{aligned}$$

The determinant of a matrix, denoted $\left|\mathrm{A}\right|$ can be determined by choosing any row or column, multiplying each element in that row or column by its cofactor and adding.

Shortcuts and Special Cases

- The determinant of a triangular matrix (one in which either all the entries above the main diagonal are zero or all the entries below it are) is the product of the entries on the diagonal.
- The determinant of a matrix is the same as the determinant of its transpose $\det(\mathrm{A}) = \det(\mathrm{A}^T)$
- If a matrix has a row or column of zeros, or if one row or column is a multiple of another, then its determinant is zero.

The Adjoint of a Matrix

The adjoint of a matrix is the transpose of the matrix obtained by replacing each element by its cofactor; the matrix obtained by replacing each element a_{ij} by the cofactor of the element a_{ji}. The transpose of a matrix is formed by writing its column as rows. If $\mathrm{A} = m \times n$ then $\mathrm{A}^T = n \times m$. The Hermitian conjugate matrix is frequently called the adjoint matrix by writers on quantum mechanics.

If **A** is an $n \times m$ invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)}\,\mathrm{adj}(A)$$

Cramer's Rule

Cramer's rule, named after Gabriel Cramer (1704 -1752), provides a method for solving systems of linear equations in two or more unknowns by expressing the unknown variables as the ratio of two determinants. The formal definition is as follows:

For a system of n linear equations in n unknowns, let A be the coefficient matrix of the system. If $\det A \neq 0$, then the solution of the system is given as

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad ..., \qquad x_n = \frac{\det A_n}{\det A}$$

Where the *i*th column of A_i is the column of constants in the system of equations.

Cramer's rule can be extended to a system of *n* linear equations in *n* variables. The value of each variable is the quotient of two determinants. The denominator is the determinant of the coefficient matrix, and the numerator is the determinant formed by replacing the column corresponding to the variable being solved for with the column representing the constants. Cramer's rule is derived as follows:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Multiplying the first equation by $-a_{21}$ and the second equation by a_{11} and adding the results produces the following:

$$\begin{array}{r} -a_{21}a_{11}x_1 - a_{21}a_{12}x_2 = -a_{21}b_1 \\ \underline{a_{11}a_{21}x_1 + a_{11}a_{22}x_2 = a_{11}b_2} \\ (a_{11}a_{22} - a_{21}a_{12})\,x_2 = a_{11}b_2 - a_{21}b_1 \end{array}$$

In order to add or subtract the constants the subscripts must be the same but do not have to be in the same order. In this example, when you add $-a_{21}a_{11}x_1$ and $a_{11}a_{21}x_1$ the result is zero.

$$\frac{(a_{11}a_{22} - a_{21}a_{12})\,x_2}{a_{11}a_{22} - a_{21}a_{12}} = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}$$ divide both sides by $(a_{11}a_{22} - a_{21}a_{12})$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}$$

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}$$

Since the numerators and denominators of both x_1 and x_2 can be represented as determinants (e.g. $a_{11}a_{22} - a_{21}a_{12}$), you have the following:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \qquad a_{11}a_{22} - a_{21}a_{12} \neq 0$$

The denominator for both x_1 and x_2 is the determinant of the coefficient matrix **A**. The determinant forming the numerator of x_1 can be obtained from **A** by replacing its first column by the column representing the constants of the system. The determinant forming the numerator of x_2 can be obtained by replacing its second column by the column representing the constants of the system. The two determinants are denoted $|\mathrm{A}_1|$ and $|\mathrm{A}_2|$.

$$|\mathrm{A}_1| = \begin{vmatrix} \boldsymbol{b_1} & a_{12} \\ \boldsymbol{b_2} & a_{22} \end{vmatrix} \qquad |\mathrm{A}_2| = \begin{vmatrix} a_{11} & \boldsymbol{b_1} \\ a_{21} & \boldsymbol{b_2} \end{vmatrix}$$

Which results in $x_1 = \frac{|\mathrm{A}_1|}{|\mathrm{A}|}$ and $x_2 = \frac{|\mathrm{A}_2|}{|\mathrm{A}|}$.

The syntax for systems of equations and determinants can vary from one textbook to another, but the operations are the same. The syntax used most often for linear equations is as follows:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned} \qquad\qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

Cramer’s rule using both types of syntax is given below. The solution of the systems of equations where $|\mathrm{A}| \neq 0$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

$$x_1 = \frac{|A_1|}{|A|} = \frac{\begin{vmatrix} \boldsymbol{b_1} & a_{12} & a_{13} \\ \boldsymbol{b_2} & a_{22} & a_{23} \\ \boldsymbol{b_3} & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}} \qquad x_2 = \frac{|A_2|}{|A|} = \frac{\begin{vmatrix} a_{11} & \boldsymbol{b_1} & a_{13} \\ a_{21} & \boldsymbol{b_2} & a_{23} \\ a_{31} & \boldsymbol{b_3} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

$$x_3 = \frac{|A_3|}{|A|} = \frac{\begin{vmatrix} a_{11} & a_{12} & \boldsymbol{b_1} \\ a_{21} & a_{22} & \boldsymbol{b_2} \\ a_{31} & a_{32} & \boldsymbol{b_3} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}$$

The solution of the systems of equations where $D \neq 0$:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}$$

is given by

$$x = \frac{D_x}{D}, \qquad y = \frac{D_y}{D}, \qquad z = \frac{D_z}{D},$$

where

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad D_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

$$D_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix} \qquad D_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

Example 1: Solve the following system using Cramer's rule.

$$x - 3y + 7z = 13$$
$$x + y + z = 1$$
$$x - 2y + 3z = 4$$

$$\mathrm{D} = \begin{vmatrix} 1 & -3 & 7 \\ 1 & 1 & 1 \\ 1 & -2 & 3 \end{vmatrix} = -10 \qquad \mathrm{D}_x = \begin{vmatrix} \mathbf{13} & -3 & 7 \\ \mathbf{1} & 1 & 1 \\ \mathbf{4} & -2 & 3 \end{vmatrix} = 20$$

$$\mathrm{D}_y = \begin{vmatrix} 1 & \mathbf{13} & 7 \\ 1 & \mathbf{1} & 1 \\ 1 & \mathbf{4} & 3 \end{vmatrix} = -6 \qquad \mathrm{D}_z = \begin{vmatrix} 1 & -3 & \mathbf{13} \\ 1 & 1 & \mathbf{1} \\ 1 & -2 & \mathbf{4} \end{vmatrix} = -24$$

$$x = \frac{\mathrm{D}_x}{\mathrm{D}} = \frac{20}{-10} = -2$$
$$y = \frac{\mathrm{D}_y}{\mathrm{D}} = \frac{-6}{-10} = \frac{3}{5}$$

$$z = \frac{\mathrm{D}_z}{\mathrm{D}} = \frac{-24}{-10} = \frac{12}{5}$$

Triangular Matrix

For an upper triangular matrix the matrix must be square and all the entries below the main diagonal are zero and the main diagonal entries and the entries above it may or may not be zero. A lower triangular matrix is just the opposite. The matrix is still a square matrix and all the entries of a lower triangular matrix above the main diagonal are zero and the main diagonal entries and those below it may or may not be zero.

If A is a triangular matrix of order n, then its determinant is the product of the entries on the main diagonal:

$$\det(\mathrm{A}) = |\mathrm{A}| = a_{11}a_{22}a_{33}\ldots a_{nn}$$

Rules for finding determinants using triangular matrices:

Let A be a square matrix:

1. If B is the matrix that results from multiplying a row or column of A by a scalar c, then $\det(B) = c\det(A)$.
2. If B is the matrix that results from interchanging two rows or two columns of A, then $\det(B) = -\det(A)$. If you interchange two rows or columns you have to place a minus sign in front of the determinant.

3. If B is the matrix that results from adding a multiple of one row of A onto another row of A, or adding a multiple of one column of A onto another column of A, then $\det(B) = \det(A)$.
4. When you multiply a row by a fraction it changes the value of the determinant so you have to multiply the resulting matrix by the reciprocal of the scalar that was used in the row operation.

Example:

$$A = \begin{bmatrix} -2 & 10 & 2 \\ 1 & 0 & 7 \\ 0 & -3 & 5 \end{bmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 7 \\ -2 & 10 & 2 \\ 0 & -3 & 5 \end{vmatrix} \quad R_1 \leftrightarrow R_2$$

$$= -\begin{vmatrix} 1 & 0 & 7 \\ 0 & 10 & 16 \\ 0 & -3 & 5 \end{vmatrix} \quad R_2 + 2R_1$$

$$= -(10)\begin{vmatrix} 1 & 0 & 7 \\ 0 & 1 & \frac{8}{5} \\ 0 & -3 & 5 \end{vmatrix} \quad \frac{1}{10}R_2$$

$$= -(10)\begin{vmatrix} 1 & 0 & 7 \\ 0 & 1 & \frac{8}{5} \\ 0 & 0 & \frac{49}{5} \end{vmatrix} \quad R_3 + 3R_2$$

$$\det(A) = -10(1)(1)\left(\frac{49}{5}\right) = -98$$

Parametric Equations

Parametric equations are equations in which coordinates are each expressed in terms of quantities called parameters. A parametrically defined curve uses individual equations to represent the coordinates of each point. In 2-space the x and y coordinates of each point are defined in terms of a third variable t called the parameter. You then substitute a value for t into both parametric equations to identify the coordinates of one point on the graph of the parametric curve. For example, the linear equation $y = 2x - 7$ can be represented by the parametric equations $x = t + 3$ and $y = 2t - 1$. To eliminate the parameter and express the parametric curve in rectangular form you eliminate the parameter leaving behind only the variables x and y. To do this you solve one of the parametric equations for t and substitute it into the other parametric equation.

Example: Eliminate the parameter from the equations $x = t + 3$ and $y = 2t - 1$, and express the equation in rectangular form.

$x = t + 3$
$x - 3 = t + 3 - 3$ subtract 3 from both sides of the equation
$t = x - 3$

$y = 2t - 1$
$y = 2(x - 3) - 1$ replace the parameter t with its equivalent
$y = 2x - 6 - 1$
$y = 2x - 7$

In selecting the variable for linear algebra problems it is recommended that choose the variables that occur last in a given equation to be the free variables. For example,

$$x_1 = 4 - 2x_2$$

In this form, the variable x_2 is free, which means it can take on any real value. The variable x_1 is not free because its value depends on the value assigned to x_2. To represent the infinite number of solutions of this equation, it is convenient to introduce a third variable called the parameter. By letting $x_2 = t$ you can represent the solution set as $x_1 = 4 - 2t, \quad x_2 = t$

Eigenvalues and Eigenvectors

Eigen is a German word which means proper or characteristic. Eigenvectors are a special set of vectors associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic vectors, proper vectors, or latent vectors.

The determination of the eigenvectors and eigenvalues of a system is extremely important in physics and engineering, where it is equivalent to matrix diagonalization and arises in such common applications as stability analysis, the physics of rotating bodies, and small oscillations of vibrating systems, to name only a few. Each eigenvector is paired with a corresponding so-called eigenvalue. Mathematically, two different kinds of eigenvectors need to be distinguished: left eigenvectors and right eigenvectors. However, for many problems in physics and engineering, it is sufficient to consider only right eigenvectors. The term "eigenvector" used without qualification in such applications can therefore be understood to refer to a right eigenvector. An eigenvalue of a square matrix is a scalar that is usually represented by the Greek letter λ (pronounced Lambda). An eigenvector is a vector. Moreover, we require that an eigenvector be a non-zero vector, in other words, an eigenvector can not be the zero vector. We will denote an eigenvector by the small letter ***x***. All eigenvalues and eigenvectors satisfy the equation $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for a given square matrix, *A*. Eigenvalues are values of λ for which $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Eigenvectors are solutions of ***x*** corresponding to particular values of λ. An eigenvector cannot be zero. A matrix can have more than one eigenvalue.

Let A be an $n \times n$ matrix. Then,

1. An eigenvalue of *A* is a scalar λ such that $\det(\lambda I - A) = 0$
2. The eigenvalues of *A* corresponding to λ are the nonzero solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$

The equation $\det(\lambda I - A) = 0$ is called the characteristic equation of *A*. When expanded to polynomial form the polynomial $|(\lambda I - A)| = \lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda + c_0$ is called the characteristic polynomial. From this definition it follows that the eigenvalues of an $n \times n$ matrix *A* correspond to the roots of the characteristic polynomial of *A*.

Algorithm: Let A be an $n \times n$ matrix, then:

1. Form the characteristic equation $|(\lambda I - A)| = 0$
2. Find the roots of the characteristic equation. These are the eigenvalues of A
3. For each eigenvalue λ_i, find the eigenvectors corresponding to λ_i by solving the homogeneous system $(\lambda_i I - A)\mathbf{x} = \mathbf{0}$. Use reduced row echelon form on the $n \times n$ matrix. The resulting matrix must have at least one row of zeros (the system must have nontrivial solutions).

Example: Find the eigenvalues and eigenvectors corresponding to the following matrix:

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

First, solve the characteristic equation to find the eigenvalues.

$$|(\lambda I - A)| = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \overset{I}{\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}} - \overset{A}{\begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}}$$

$$= \begin{bmatrix} \lambda - 2 & 12 \\ 1 & \lambda + 5 \end{bmatrix} \qquad \text{result of } I - A$$

$$= (\lambda - 2)(\lambda + 5) - (-1)(12) \qquad \text{find the determinant}$$

$$= \lambda^2 + 5\lambda - 2\lambda - 10 + 12$$
$$= \lambda^2 + 3\lambda + 2 \qquad \text{factor and solve for } \lambda$$
$$= (\lambda + 1)(\lambda + 2)$$

$$\lambda_1 + 1 = 0 \qquad \lambda_2 + 2 = 0$$
$$\lambda_1 + 1 - 1 = 0 - 1 \qquad \lambda_2 + 2 - 2 = 0 - 2$$
$$\lambda_1 = -1 \qquad \lambda_2 = -2$$

Place eigenvalues back into the characteristic formula to form a new matrix. Next, use Gauss-Jordon elimination to reduce the matrix to row echelon form with at least one row of zeroes.

For eigenvalue $\lambda_1 = -1$:

$$(-1)I - A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix} = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \quad R2 \leftrightarrow R1 \quad \begin{bmatrix} -1 & 4 \\ -3 & 12 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 4 \\ -3 & 12 \end{bmatrix} \quad -3R1 + R2 \quad \begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix}$$

$$-1\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix}$$ multiply by -1 to get a positive coefficient in the first row

$$\begin{bmatrix} 1 & -4 \\ 0 & 0 \end{bmatrix}$$

This gives the following equation:

$x_1 - 4x_2 = 0$ where $x_2 = t$ and $x_1 - 4t = 0$

This gives the eigenvector for λ_1 as follows:

$$\mathbf{x}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix}, t \neq 0$$

For eigenvalue $\lambda_2 = -2$:

$$(-2)I - A = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} - \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \quad R2 \leftrightarrow R1 \quad \begin{bmatrix} -1 & 3 \\ -4 & 12 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 3 \\ -4 & 12 \end{bmatrix} \quad -4R1 + R2 \quad \begin{bmatrix} -1 & 3 \\ 0 & 0 \end{bmatrix}$$

$$-1\begin{bmatrix} -1 & 3 \\ 0 & 0 \end{bmatrix}$$ multiply by -1 to get a positive coefficient in the first row

$$\begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

This gives the following equation:

$x_1 - 3x_2 = 0$ where $x_2 = t$ and $x_1 - 3t = 0$

This gives the eigenvector for λ_2 as follows:

$$\mathbf{x}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t\begin{bmatrix} 3 \\ 1 \end{bmatrix}, t \neq 0$$

For a linear transformation *T* on a vector space *V*, an eigenvalue is a scalar λ for which there is a nonzero member ***v*** of *V* for which $T\boldsymbol{v} = \lambda\boldsymbol{v}$. The vector ***v*** is an eigenvector (or characteristic vector). For a matrix *A*, the eigenvalues are the roots (characteristic roots) of the characteristic equation of the matrix. The number λ being an eigenvalue means there is a nonzero vector $\boldsymbol{x} = (x_1, x_2, ... x_n)$ for which $A\boldsymbol{x} = \lambda\boldsymbol{x}$, where multiplication is matrix multiplication and ***x*** is considered to be a one column matrix. If *A* is an $n \times n$ triangular matrix, then its eigenvalues are the entries on its main diagonal.

Eigenspace is the set of all eigenvectors of a given eigenvalue λ together with the zero vector and is a subspace of R^n If A is an $n \times n$ matrix with an eigenvalue λ, the the set of all eigenvectors of λ, together with the zero vector $\{\mathbf{0}\} \cup \{\mathbf{x} : \mathbf{x} \text{ is an eigenvalue of } \lambda\}$ is a subspace of R^n . This subspace is called the eigenspace of λ. Wolfram definition: If *A* is an $n \times n$ square matrix and λ is an eigenvalue of *A*, then the union of the zero vector **0** and the set of all eigenvectors corresponding to eigenvalues λ is a subspace of R^n known as the eigenspace of λ.

Diagonalization

A **diagonal** matrix is a square matrix all of whose nonzero elements are in the principal diagonal. If, in addition, all the diagonal elements are equal, the matrix is a scalar matrix. An identity (or unit) matrix is a diagonal matrix whose elements in the principal diagonal are all unity.

Matrix **diagonalization** is the process of taking a square matrix and converting it into a special type of matrix--a so-called diagonal matrix--that shares the same fundamental properties of the underlying matrix. Matrix diagonalization is equivalent to transforming the underlying system of equations into a special set of coordinate axes in which the matrix takes this canonical form. Diagonalizing a matrix is also equivalent to finding the matrix's eigenvalues, which turn out to be precisely the entries of the diagonalized matrix. Similarly, the eigenvectors make up the new set of axes corresponding to the diagonal matrix.

The remarkable relationship between a diagonalized matrix, eigenvalues, and eigenvectors follows from the beautiful mathematical identity (the eigen decomposition) that a square matrix A can be decomposed into the very special form

Suppose that *A* is a square matrix and further suppose that there exists an invertible matrix *P* (of the same size as *A*) such that $P^{-1}AP$ is a diagonal matrix. In such a case we call *A* diagonalizable and say that *P* diagonalizes *A*. An $n \times n$ matrix *A* is diagonalizable if *A* is similar to a diagonal matrix. That is, *A* is diagonalizable if there exists and invertible matrix *P* such that $P^{-1}AP$. Two square matrices *A* and *B* are called similar if there exists and invertible matrix *P* such that $B = P^{-1}AP$. An $n \times n$ matrix *A* is diagonalizable if and only if it has n linearly independent eigenvectors.

Algorithm: To find *P*, provided it exists of course. First find the eigenvalues for the matrix *A* and then for each eigenvalue find a basis for the eigenspace corresponding to that eigenvalue. The set of basis vectors will then serve as a set of linearly independent eigenvectors for the eigenvalue. If for all the eigenvalues we have a set of *n* eigenvectors then *A* is diagonalizable and we use the eigenvectors to form *P*. If we don't have a set of *n* eigenvectors then *A* is not diagonalizable. The steps are as follows:

Let a be an $n \times n$ matrix:

1. Find *n* linearly independent eigenvectors $\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_n$ for *A* with corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. If *n* linearly independent eigenvectors do not exist, then *A* is not diagonalizable.
2. If *A* has *n* linearly independent eigenvectors, let *P* be the $n \times n$ matrix whose columns consist of these eigenvectors. That is

   $$P = [\mathbf{p}_1 \vdots \mathbf{p}_2 \vdots ... \vdots \mathbf{p}_n]$$

3. The diagonal matrix $D = P^{-1}AP$ will have eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ on its main diagonal and zeros elsewhere. The order of the eigenvectors used to form *P* will determine the order in which the eigenvalues appear on the main diagonal D.

Example: Show that the matrix A is diagonalizable. Then find a matrix P such that $P^{-1}AP$ is diagonal.

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

The characteristic polynomial of A is as follows:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 1 & -1 \\ -1 & \lambda - 3 & -1 \\ 3 & -1 & \lambda + 1 \end{vmatrix} = (\lambda - 2)(\lambda + 2)(\lambda - 3)$$

This gives the eigenvalues of $\lambda_1 = 2$, $\lambda_2 = -2$, and $\lambda_3 = 3$. From these eigenvalues obtain the reduced row echelon forms and corresponding eigenvectors.

$$2I - A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \quad \mathbf{x}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$-2I - A = \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{bmatrix} \to \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

$$3I - A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \to \quad \mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Form the matrix P from the eigenvectors above.

$$P = \begin{bmatrix} -1 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & 4 & -1 \end{bmatrix}$$

The matrix is nonsingular, which implies that the eigenvectors are linearly independent and A is diagonalizable. P inverse is

$$P^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ \frac{1}{5} & 0 & \frac{1}{5} \\ \frac{1}{5} & 1 & \frac{1}{5} \end{bmatrix}$$

So it follows that:

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Symmetric Matrices

A square matrix *A* is symmetric if it is equal to its transpose $A = A^T$. If *A* is an $n \times n$ symmetric matrix, then the following properties are true:

1. *A* is diagonalizable
2. All eigenvalues of *A* are real
3. If λ is an eigenvalue of *A* with multiplicity k, the λ has *k* linearly independent eigenvectors. That is, the eigenspace of λ has dimension *k*.

If the above properties are true, this is called the Real Spectral Theorem, and the set of eigenvalues of *A* is called the spectrum of *A*.

Orthogonal Matrix

An orthogonal matrix is a matrix that is equal to the inverse of its transpose. A square matrix *P* is called orthogonal if it is invertible and if $P^{-1} = P^T$. For example:

The matrix $P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is orthogonal because $P^{-1} = P^T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

An $n \times n$ matrix *P* is orthogonal if and only if its column vectors form and orthonormal set.

Orthogonal Diagonalization of a Symmetric Matrix

Let *A* be an $n \times n$ symmetric matrix:

1. Find all eigenvalues of *A* and determine the multiplicity of each
2. For each eigenvalue of multiplicity 1, choose a unit eigenvector. (Choose any eigenvector and normalize it). The word multiplicity is a general term meaning "the number of values for which a given condition holds." For example, the term

is used to refer to the value of the totient valence function or the number of times a given polynomial equation has a root at a given point.

3. For each eigenvalue of multiplicity $k \geq 2$, find a set of *k* linearly independent eigenvectors. If this set is not orthonormal apply the Gram-Schmidt orthonormalization process.
4. The composite of steps 2 and 3 produces an orthonormal set of *n* eigenvectors. Use these eigenvectors to form the columns of *P*. The matrix $P^{-1}AP = P^T AP = D$ will be diagonal. (The main diagonal entries of *D* are the eigenvalues of *A*)

Example: Find an orthogonal matrix P that orthogonally diagonalizes the following matrix:

$$A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$$

The characteristic polynomial of *A* is:

$$|\lambda I - A| = \begin{vmatrix} \lambda + 2 & -2 \\ -2 & \lambda - 1 \end{vmatrix} = (\lambda + 3)(\lambda - 2)$$

Which gives the eigenvalues of $\lambda_1 = -3$ and $\lambda_2 = 2$.
For each eigenvalue, find an eigenvector by converting the matrix $\lambda I - A$ to reduced row echelon form.

$$-3I - A = \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \rightarrow \quad \mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$2I - A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} \rightarrow \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The eigenvectors (-2, 1) and (1, 2) form an orthogonal basis for R^2. Each of these eigenvectors is normalized to produce and orthonormal basis. The normalized vector of *X* is a vector in the same direction but with norm (length) 1. It is denoted $\hat{\mathbf{X}}$ and given by

$$\hat{\mathbf{X}} = \frac{\mathbf{X}}{|\mathbf{X}|}$$ where $|\mathbf{X}|$ is the norm of $\mathbf{X}$, also called the unit vector.

$$\mathbf{p}_1 = \frac{(-2,1)}{\|(-2,1)\|} = \frac{1}{\sqrt{5}}(-2,1) = \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$$

$$\mathbf{p}_2 = \frac{(1,2)}{\|(1,2)\|} = \frac{1}{\sqrt{5}}(1,2) = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

Because each eigenvalue has a multiplicity of 1, construct the matrix P

$$P = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

Verify P is correct by computing $P^{-1}AP = P^T AP$.

$$P^T AP = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

Rank of a Matrix

The dimension of the row or column space of a matrix A is called the **rank** of A and is denoted by rank(A). The rank is the order of the nonzero determinant of greatest order that can be selected from the matrix by taking out rows and columns. The concept rank facilitates the statement of the condition for consistency of simultaneous linear equations: m linear equations in n unknowns are consistent when, and only when, the rank of the matrix of the coefficients is equal to the rank of the augmented matrix. For example, in the system of equations

$$\begin{aligned} x + y + z + 3 = 0 \\ 2x + y + z + 4 = 0 \end{aligned}$$

The matrix of the coefficients is

$$\begin{Vmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{Vmatrix}$$

And the augmented matrix is

$$\begin{Vmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 4 \end{Vmatrix}$$

The rank of both is two, because the determinant $\begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix}$ is not zero.

The dimension of the null space of *A* is called the **nullity** of *A* and is denoted by nullity(*A*).

Let *A* be an $n \times m$ matrix:

1. The subspace of R^m that is spanned by the row vectors of *A* is called the row space of *A*.
2. The subspace of R^n that is spanned by the column vectors of *A* is called the column space of *A*.
3. The set of all *x* in R^m such that $A\mathbf{x} = \mathbf{0}$ is called the null space of A.

Example: Find a basis for the null space, row space and column space of the following matrix. Determine the rank and nullity of the matrix.

$$A = \begin{bmatrix} -1 & 2 & -1 & 5 & 6 \\ 4 & -4 & -4 & -12 & -8 \\ 2 & 0 & -6 & -2 & 4 \\ -3 & 1 & 7 & -2 & 12 \end{bmatrix}$$

We'll find the nullspace first since that was the first thing asked for. The nullspace is the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$. To do this we'll need to solve the following system of equations. To solve this system write the augmented matrix $[A \quad \vdots \quad \mathbf{0}]$ in reduced row-echelon form.

$$\begin{aligned} -x_1 + 2x_2 - x_3 + 5x_4 + 6x_5 &= 0 \\ 4x_1 - 4x_2 - 4x_3 - 12x_4 - 8x_5 &= 0 \\ 2x_1 + 0x_2 - 6x_3 - 2x_4 + 4x_5 &= 0 \\ -3x_1 + x_2 + 7x_3 - 2x_4 + 12x_5 &= 0 \end{aligned}$$

The solutions for this system are as follows:

$x_1 = 3t \qquad x_2 = 2t - 8s \qquad x_3 = t \qquad x_4 = 2s \qquad x_5 = s$

The nullspace is as follows:

$$x = \begin{bmatrix} 3t \\ 2t - 8s \\ t \\ 2s \\ s \end{bmatrix} = t \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ -8 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

The basis for the nullspace is:

$$x_1 \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad x_2 \begin{bmatrix} 0 \\ -8 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

So the nullity is $\text{nullity}(A) = 2$ and the rank is equal to 3.

Rank = n(# of columns) - nullity(A)
= 5 - 2
= 3

The rank provides a check when you find a basis for the row and column space. In this case each should contain three vectors.

Next, the first, second and fourth columns of U contain leading 1's so they will form a basis for the column space of U and this tells us that the first, second and fourth columns of A will form a basis for the column space of A.

Matrix A in row echelon form is given below:

$$U = \begin{bmatrix} -1 & -2 & 1 & -5 & -6 \\ 0 & 1 & -2 & 2 & 4 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The rows containing the leading 1's will form a basis for the row space of A:

$$r_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \\ -5 \\ -6 \end{bmatrix} \qquad r_2 = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 2 \\ 4 \end{bmatrix} \qquad r_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -2 \end{bmatrix}$$

The first, second and fourth columns of U contain leading 1's so they will form a basis for the column space of U. This also suggests that the first, second and fourth columns of A will form a basis for the column space of A.

$$c_1 = \begin{bmatrix} -1 \\ 4 \\ 2 \\ -3 \end{bmatrix} \qquad c_2 = \begin{bmatrix} 2 \\ -4 \\ 0 \\ 1 \end{bmatrix} \qquad c_3 = \begin{bmatrix} 5 \\ -12 \\ -2 \\ -2 \end{bmatrix}$$

Summary of Equivalent Conditions for Square Matrices - $A = n \times n$ matrix

1. A is invertible
2. $A\mathbf{x} = \mathbf{b}$ has a unique solution for any $n \times 1$ matrix $\mathbf{b}$
3. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
4. A is row equivalent to I_n
5. $|A| \neq 0$
6. $\text{Rank}(A) = n$
7. The n row vectors of A are linearly independent
8. The n column vectors of A are linearly independent

Finding a Basis for a Row Space

Find a basis for the row space:

$$A = \begin{bmatrix} 1 & 3 & -1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & -2 \end{bmatrix}$$

Rewrite A in row-echelon form:

$$B = \begin{bmatrix} 1 & 3 & -1 & 3 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The nonzero row vectors of B form a basis for the row space of A as shown below:

$$\mathbf{w}_1 = (1,3,1,3) \qquad \mathbf{w}_2 = (0,1,1,0) \qquad \mathbf{w}_3 = (0,0,0,1)$$

Finding a Basis for a Column Space

Find a basis for the column space of matrix *A*:

$$A = \begin{bmatrix} 1 & 3 & -1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & -2 \end{bmatrix}$$

Take the transpose of *A* and rewrite A^T in row-echelon form:

$$A^T = \begin{bmatrix} 1 & 0 & -3 & 3 & 2 \\ 3 & 1 & 0 & 4 & 0 \\ 1 & 1 & 6 & -2 & -4 \\ 3 & 0 & -1 & 1 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3 & 3 & 2 \\ 0 & 1 & 9 & -5 & -6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The following form a basis for the row space of A^T :

$$\mathbf{w}_1 = (1,0,-3,3,2) \qquad \mathbf{w}_2 = (0,1,9,-5,-6) \qquad \mathbf{w}_3 = (0,0,1,-1,-1)$$

This is equivalent to stating that the following column vectors form a basis for the column space of *A*:

$$\begin{bmatrix} 1 \\ 0 \\ -3 \\ 3 \\ 2 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 9 \\ -5 \\ -6 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

Finding a Basis for a Subspace

Find a basis for the subspace of R^3 spanned by:

$$S = \{\underset{\mathbf{v}_1}{(-1,2,5)}, \underset{\mathbf{v}_2}{(3,0,3)}, \underset{\mathbf{v}_3}{(5,1,8)}\}$$

Use $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ to form the rows of matrix *A*. Then write *A* in row-echelon form.

$$A = \begin{bmatrix} -1 & 2 & 5 \\ 3 & 0 & 3 \\ 5 & 1 & 8 \end{bmatrix} \rightarrow \quad B \begin{bmatrix} -1 & 2 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

The nonzero row vectors of *B* that form a basis for the row space of *A* are as follows:

$\mathbf{w}_1 = (1, -2, -5) \qquad \mathbf{w}_2 = (0, 1, 3)$

These nonzero row vectors form a basis for the subspace spanned by $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Change of Basis

Generally, you are provided with the coordinates of a vector relative to one basis *B* and are asked to find the coordinates relative to another basis B'. If you let B be the standard basis, then the problem of finding the coordinate matrix relative to the basis B', becomes one of solving for $c_1, c_2, ..., c_n$.

Coordinate representation relative to a basis

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ be an ordered basis for a vector space *V* and let **x** be a vector in *V* such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$$

The scalars $c_1, c_2, ..., c_n$ are called the coordinates of **x** relative to the basis B. The coordinate matrix (or coordinate vector) of x relative to *B* is the column matrix in R^n whose components are the coordinates of **x**.

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_3 \end{bmatrix}$$

Example 1: Find the coordinate matrix of $\mathbf{x} = (-2, 1, 3)$ in R^3 relative to the standard basis $S = \{(1,0,0), (0,1,0), (0,0,1)\}$.

Because x can be written as:

$$\mathbf{x} = (-2, 1, 3) = -2(1,0,0) + 1(0,1,0) + 3(0,0,1)$$

The coordinate matrix of x relative to the standard basis is:

$$[\mathbf{x}]_S = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$$

The components of x are the same as its coordinates relative to the standard basis.

Example 2: Find the coordinate matrix relative to a standard basis. The coordinate matrix of $\mathbf{x}$ in R^2 relative to the ordered basis $B = \{\mathbf{v}_1, \mathbf{v}_2\} = \{(1,0),(1,2)\}$ is

$$[\mathbf{x}]_B = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Find the coordinates of $\mathbf{x}$ relative to the standard basis $B^{'} = \{\mathbf{u}_1, \mathbf{u}_2\} = \{(1,0),(0,1)\}$.

Because $[\mathbf{x}]_B = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ then $\mathbf{x} = 3\mathbf{v}_1 + 2\mathbf{v}_2 = 3(1,0) + 2(1,2) = (5,4)$ and since $(5,4) = 5(1,0) + 4(01)$, it follows that the coordinates of x relative to $B^{'}$ are $[\mathbf{x}]_{B^{'}} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$

Example 3: Find the coordinate matrix of $\mathbf{x} = (1,2,-1)$ in R^3 relative to the nonstandard basis $B^{'} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2\} = \{(1,0,1),(0,-1,2),(2,3,-5)\}$.

Begin by writing x as a linear combination of $\mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{u}_3$:

$$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3$$
$$(1,2,-1) = c_1(1,0,1) + c_2(0,-1,2) + c_3(2,3,-5)$$

Equating corresponding components produces the following system of linear equations:

$$\begin{aligned} c_1 \qquad + 2c_3 &= 1 \\ -c_2 + 3c_3 &= 2 \\ c_1 + 2c_2 - 5c_3 &= 1 \end{aligned}$$

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

The solution of this system is $c_1 = 5, \qquad c_2 = -8, \qquad c_3 = -2$ so

$$\mathbf{x} = 5(1,0,1) + (-8)(0,1,2) + (-2)(2,3,-5)$$

And the coordinate matrix of **x** relative to $B^{'}$ is

$$[\mathbf{x}]_{B^{'}} = \begin{bmatrix} 5 \\ -8 \\ -2 \end{bmatrix}$$

The Inverse of a Transition Matrix

Inverse is defined as opposite in nature and effect; -- said with reference to any two operations, which, when both are performed in succession upon any quantity, reproduce that quantity; as, multiplication is the inverse operation to division. The symbol of an inverse operation is the symbol of the direct operation with -1 as an index; thus $\sin^{-1} x$ means the arc whose sine is *x*.

A **transition** is a change from one form to another or the movement from one place or state to another.

If *P* is the transition matrix from a basis $B^{'}$ to a basis *B* in R^n, then *P* is invertible and the transition matrix from *B* to $B^{'}$ is given by P^{-1}.

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_n\}$ and $B^{'} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, ..., \mathbf{u}_n\}$ be two bases for R^n. Then the transition matrix P^{-1} from *B* to $B^{'}$ can be found by using Gauss-Jordan elimination on the $n \times 2n$ matrix $[B^{'} \quad \vdots \quad B]$, as follows:

$$[B^{'} \quad \vdots \quad B] \to [I_n \quad \vdots \quad P^{-1}]$$

Example:

Gram-Schmidt Process

The **Gram-Schmidt Process** is the process of forming an orthogonal sequence from a linearly independent sequence.

Linear Transformations

The Geometry of Linear Transformations in the Plane

Reflection in y-axis

$$A=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

$$T(x,y)=(-x,y)$$
$$\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}-x\\ y\end{bmatrix}$$

Reflection in x-axis

$$A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

$$T(x,y)=(x,-y)$$
$$\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x\\ -y\end{bmatrix}$$

Reflections in Line *y = x*

$$A=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

$$T(x,y)=(y,x)$$
$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}y\\ x\end{bmatrix}$$

Horizontal expansion $(k>1)$ **or contraction** $(0<k<1)$ **where *k* is a positive scalar**

$$A=\begin{bmatrix}k & 0\\ 0 & 1\end{bmatrix}$$

$$T(x,y)=(kx,y)$$
$$\begin{bmatrix}k & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}kx\\ y\end{bmatrix}$$

Vertical expansion $(k>1)$ **or contraction** $(0<k<1)$

$$A=\begin{bmatrix}1 & 0\\ 0 & k\end{bmatrix}$$

$$T(x,y)=(x,ky)$$
$$\begin{bmatrix}1 & 0\\ 0 & k\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x\\ ky\end{bmatrix}$$

Horizontal Shear

$$A=\begin{bmatrix}1 & 0\\ 0 & k\end{bmatrix}$$

$$T(x,y)=(x+ky,y)$$
$$\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x+ky\\ y\end{bmatrix}$$

Vertical Shear

$$A=\begin{bmatrix}1 & 0\\ k & 1\end{bmatrix}$$

$$T(x,y)=(x,y+kx)$$
$$\begin{bmatrix}1 & 0\\ k & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x\\ kx+y\end{bmatrix}$$

Shear is a transformation in which all points along a given line *L* remain fixed while other points are shifted parallel to *L* by a distance proportional to their perpendicular distance from *L*. Shearing a plane figure does not change its area. The shear can also be generalized to three dimensions, in which planes are translated instead of lines.

Rotations in 3-space

Rotation about the x-axis ***Rotation about the y-axis*** ***Rotation about the z-axis***

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \quad \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In each case the rotation is oriented counterclockwise relative to a person facing the negative direction of the indicated axis.

Example: Given the eight vertices of a rectangular box having sides of length 1, 2, and 3, find the coordinates of the box when it is rotated counterclockwise 60 degrees about the *z*-axis.

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Original Vertex	Rotated Vertex
$V_1 = (0,0,0)$	$(0,0,0)$
$V_2 = (1,0,0)$	$(0.5,0.87,0)$
$V_3 = (1,2,0)$	$(-1.23,1.87,0)$
$V_4 = (0,2,0)$	$(-1.73,1,0)$
$V_5 = (0,0,3)$	$(0,0,3)$
$V_6 = (1,0,3)$	$(0.5,0.87,3)$
$V_7 = (1,2,3)$	$(-1.23,1.87,3)$
$V_8 = (0,2,3)$	$(-1.73,1,3)$

Groups and Fields

A **group** is a set G with a binary operation whose domain is the set of all ordered pairs of members of G, whose range is contained in G, and which satisfies the following conditions:

1. There is a member of G (called the identity or unit element) such that its product with any member, in either order, is that same member
2. For each member of G there is a member (called the inverse) such that the product of the two, in either order, is the identity
3. The associative law holds $a(b\cdot c) = (a\cdot b)c$

The positive and negative integers and zero form a group under ordinary addition, the identity being zero and the inverse of an integer its negative. A group is Abelian (or commutative) if (in addition to the three assumption listed above it satisfies the commutative law. A group for which all members are powers of one member is cyclic. The number of members of a finite group is the order of the group.

A **field** is a set for which two operations call addition and multiplication are defined and have the following properties:

1. The set F is a commutative group with addition as the group operation
2. Multiplication is commutative and the set, with the identity zero of the additive group omitted is a group with multiplication as the group operation
3. The distributive law holds $a(b+c)=ab+ac$ for all a, b, and c in the set
4. Each nonzero element, a in F has an inverse element a^{-1} in F relative to operation of multiplication.

Hermitian (Hermitian conjugate of a matrix) **–** The transpose of the complex conjugate of the matrix; called the adjoint of the matrix by some writers on quantum mechanics.

Linear Programming

Linear programming is the mathematical theory of the minimization or maximization of a linear function subject to linear constraints. A solution of a linear programming problem is any set of values x_i that satisfy the m linear constraints. A solution consisting of nonnegative numbers is a **feasible solution**. A solution consisting of m x's for which the matrix of coefficients in the constraints is not singular, and otherwise consisting of zeros, is a **basic solution**. A feasible solution that minimizes the linear form is an **optimal solution**.

A two-dimensional linear programming problem consists of a linear objective function and a system of linear inequalities called **constraints**. The objective function gives the quantity that is to be maximized (or minimized), and the constraints determine the set of feasible solutions.

If a linear programming problem has an optimal solution, it must occur at a vertex of the set of feasible solutions. If the problem has more than one optimal solution, then at least one of them must occur at a vertex of the set of feasible solutions. In either case, the value of the objective function is unique.

To solve a linear programming problem involving two variables by the graphical method, use the following steps.

1. Sketch the region corresponding to the system of constraints. (The points inside or on the boundary of the region are called feasible solutions.)
2. Find the vertices of the region.

3. Test the objective function at each of the vertices and select the values of the variables that optimize the objective function. For a bounded region, both a minimum and maximum value will exist. (For an unbounded region, if an optimal solution exists, it will occur at a vertex.)

Example: Find the maximum values of:

$z = 3x + 2y$ *Objective function*

Subject to the following constraints:

$$\left.\begin{aligned} &x \geq 0 \\ &y \geq 0 \\ &x + 2y \leq 4 \\ &x - y \leq 1 \end{aligned}\right. \qquad \textit{Constraints}$$

Set the equations equal to the constant and graph. The constraints form an area or region. At the vertices of this region the values of the objective function are as follows:

At (0, 0) z = 3(0) + 2(0) = 0
At (1, 0) z = 3(1) + 2(0) = 3
At (2, 1) z = 3(2) + 2(1) = 8
At (0, 2) z = 3(0) + 2(2) = 4

The maximum value occurs at z = 8 when x = 2 and y = 1.

When solving a linear programming problem, it is possible that the maximum (or minimum) value occurs at two different vertices. Some linear programming problems have no optimal solution. This can occur if the region determined by the constraints is unbounded.

Simplex Method

The simplex method is carried out by performing elementary row operations on a matrix called the simplex tableau. This tableau consists of the augmented matrix corresponding to the constraint equations together with the coefficients of the objective function written in the form:

$$-c_1x_1 - c_2x_2 - ... - c_nx_n + (0)s_1 + (0)s_2 + ... + (0)s_m + z = 0$$

Because the left-hand side of each inequality is less than or equal to the right-hand side, there must exist nonnegative numbers $s_1, s_2, ..., s_n$ that can be added to the left side of each equation to produce the system of linear equations. The numbers $s_1, s_2, ..., s_n$ are called slack variables because they represent the "slack" in each inequality. Occasionally, the constraints in a linear programming problem will include an equation. In such cases, you can still add a “slack variable” called an artificial variable to form the initial simplex tableau. Technically, this new variable is not a slack variable (because there is no slack to be taken). Once you have determined an optimal solution in such a problem, you should check to see that any equations in the original constraints are satisfied.

Once you have set up the initial simplex tableau for a linear programming problem, the simplex method consists of checking for optimality and then, if the current solution is not optimal, improving the current solution. (An improved solution is one that has a larger -value than the current solution.) To improve the current solution, bring a new basic variable into the solution, the entering variable. This implies that one of the current basic variables (the departing variable) must leave, otherwise you would have too many variables for a basic solution. You choose the entering and departing variables as follows.

1. The entering variable corresponds to the smallest (the most negative) entry in the bottom row of the tableau.
2. The departing variable corresponds to the smallest nonnegative ratio of in the column determined by the entering variable.
3. The entry in the simplex tableau in the entering variable’s column and the departing variable’s row is called the pivot.

Finally, to form the improved solution, apply Gauss-Jordan elimination to the column that contains the pivot.

A basic solution of a linear programming problem in standard form is a solution $\left(x_1, x_2, \ldots, x_n, s_1, s_2, \ldots s_m\right)$ of the constraint equations in which at most *m* variables are nonzero, and the variables that are nonzero are called basic variables. A basic solution for which all variables are nonnegative is called a basic feasible solution.

Algorithm for the Simplex Method

1. Determine the objective function
2. Determine the constraints
3. Convert each inequality in the set of constraints to an equation by adding slack variables.
4. Create the initial simplex tableau.
5. Locate the most negative entry in the bottom row. The column for this entry is called the entering column. (If ties occur, any of the tied entries can be used to determine the entering column.)
6. Form the ratios of the entries in the “*b*-column” with their corresponding positive entries in the entering column. The departing row corresponds to the smallest nonnegative ratio $\frac{b_i}{a_{ij}}$. (If all entries in the entering column are 0 or negative, then there is no maximum solution. For ties, choose either entry.) The entry in the departing row and the entering column is called the pivot.

7. Use elementary row operations so that the pivot is 1, and all other entries in the entering column are 0. This process is called pivoting.
8. If all entries in the bottom row are zero or positive, this is the final tableau. If not, go back to step 3.
9. If you obtain a final tableau, then the linear programming problem has a maximum solution, which is given by the entry in the lower right corner of the tableau.

Example: Use the simplex method to find the maximum value of the objective function:

$$z = 2x_1 - x_2 + 2x_3$$

Constraints:

$$\begin{aligned} 2x_1 + x_2 \qquad &\le 10 \\ x_1 + 2x_2 - 2x_3 &\le 20 \\ x_2 + 2x_3 &\le 5 \end{aligned}$$

Where $x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0$ and the basic feasible solution is $(x_1, x_2, x_3, s_1, s_2, s_3) = (0, 0, 0, 10, 20, 5)$

Add slack values and create the initial simplex tableau:

x_1	x_2	x_3	s_1	s_2	s_3	b	
2	1	0	1	0	0	10	s_1
1	2	-2	0	1	0	20	s_2
0	1	2	0	0	1	5	s_3
-2	1	-2	0	0	0	0	

Locate the most negative entry in the bottom row and find the smallest nonnegative ratio $\frac{b_i}{a_{ij}}$.

$$\frac{10}{0} = ND, \qquad \frac{20}{-2} = -10 \qquad \frac{5}{2} = 2\frac{1}{2}$$

The first ratio is division by zero so it is not defined, the second ratio is negative, and the third ration is the smallest nonnegative value so 2 will be the pivot, shown in bold below.

x_1	x_2	x_3	s_1	s_2	s_3	b	
2	1	0	1	0	0	10	s_1
1	2	-2	0	1	0	20	s_2
0	1	**2**	0	0	1	5	x_3
-2	1	-2	0	0	0	0	

Use Gauss-Jordan elimination so that the pivot is 1, and all other entries in the entering column are 0. This process is called pivoting. Notice that the basic variables change as a result of the pivot process.

Use the following row operations:

$R3 + R2$

$\frac{1}{2}R3$

x_1	x_2	x_3	s_1	s_2	s_3	b	
2	1	0	1	0	0	10	s_1
1	3	0	0	1	1	25	s_2
0	$\frac{1}{2}$	1	0	0	$\frac{1}{2}$	$\frac{5}{2}$	x_3
-2	2	0	0	0	1	5	

Since there is still a negative value in the bottom row of the tableau you need to repeat the pivot process.

$\frac{10}{2} = 5,\qquad \frac{25}{1} = 25\qquad \frac{5}{2}/0 = ND$

x_1	x_2	x_3	s_1	s_2	s_3	b	
1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	5	x_1
0	$\frac{5}{2}$	0	$-\frac{1}{2}$	1	1	20	s_2
0	$\frac{1}{2}$	1	0	0	$\frac{1}{2}$	$\frac{5}{2}$	x_3
0	3	0	1	0	1	15	

This implies that the optimal solution is $(x_1, x_2, x_3, s_1, s_2, s_3) = \left(5, 0, \frac{5}{2}, 0, 20, 0\right)$ and the maximum value of z is 15.

$$z = 2x_1 - x_2 + 2x_3$$

$$z = 2(5) - 0 + 2\left(\frac{5}{2}\right)$$

$$z = 10 + 5$$

$$z = 15$$