Chapter 3 Linear Algebra Supplement

Properties of the Dot Product

Note: this and the following two sections should logically appear at the end of these notes, but I have moved them to the front because they are the most important for our understanding of support vector machines.

Let $\overline{u}, \overline{v} \in \mathcal{R}^n$ be given. Then the **dot product** between $\overline{u}$ and $\overline{v}$, denoted $\overline{u} \cdot \overline{v}$, is defined as

$$\overline{u} \cdot \overline{v} = u_1 v_1 + \cdots + u_n v_n.$$

Properties of the dot product. Let $\overline{u}, \overline{v}, \overline{w} \in \mathcal{R}^n$ and $k \in \mathcal{R}$ be arbitrary.

Symmetry $\overline{u} \cdot \overline{v} = \overline{v} \cdot \overline{u}$

Additivity $\overline{u} \cdot (\overline{v} + \overline{w}) = \overline{u} \cdot \overline{v} + \overline{u} \cdot \overline{w}$

Scalar Associativity $(k\overline{u}) \cdot \overline{v} = k(\overline{u} \cdot \overline{v})$

Positivity $\overline{u} \cdot \overline{u} \geq 0$ and $\overline{u} \cdot \overline{u} = 0$ iff $\overline{u} = \overline{0}$

Measuring length and angle with the dot product.

$$|\overline{v}| = \sqrt{\overline{v} \cdot \overline{v}},$$

while

$$\theta(\overline{u}, \overline{v}) = \cos^{-1}\left(\frac{\overline{u} \cdot \overline{v}}{|\overline{u}||\overline{v}|}\right).$$

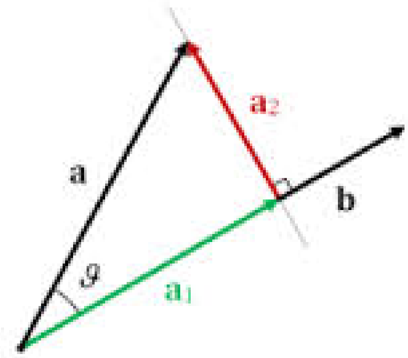

Figure 1: Projection of vector a on to a vector b

Projections onto Lines and Planes

In the theory of support vector machines, it is important to be able to compute the distance between a point and plane. This is because lines and planes may act as (linear) classifiers, and we want the classifier to be a maximum distance away from the closest training point. By definition, the **projection of a point/vector $\overline{a}$ onto a surface** S is denoted as $\text{proj}(\overline{a}, S)$, and is defined as the point on S that is closest in Euclidean distance to $\overline{a}$. For example, in Figure below, S is the line that passes through vector $\overline{b}$, which is denoted by $\overline{b}$, and $\text{proj}(\overline{a}, \overline{b})$ is equal to the vector $\overline{a}_1$. Notice also that the distance from the head of $\overline{a}_1$ to $\overline{a}$ is $|\overline{a}_2|$, and that, since this distance is minimum, $\overline{a}_2$ is orthogonal to $\overline{b}$.

The following theorem porvides a formula for computing $\overline{a}_1$.

Theorem 1. Let $\overline{a}$ and $\overline{b}$ be vectors. Then

$$\text{proj}(\overline{a}, \overline{b}) = (\frac{\overline{a} \cdot \overline{b}}{|\overline{b}|^2})\overline{b}.$$

Moreover,

$$|\text{proj}(\overline{a}, \overline{b})| = \frac{|\overline{a} \cdot \overline{b}|}{|\overline{b}|}.$$

Proof of Theorem 1. We prove it for the case of two-dimensional vectors. Let θ be the angle made between $\overline{a}$ and $\overline{b}$, and that the tails of $\overline{a}$ and $\overline{b}$ are both positioned at the origin. Let $\overline{a}$ have coordinates $(a\cos\alpha, a\sin\alpha)$, and $\overline{b}$ have coordinates $(b\cos\beta, b\sin\beta)$, where $a, b > 0$ and α and β are between $0°$ and $90°$. Then,

$$|\overline{a}_1| = a\cos\theta = a\cos(\alpha - \beta) = a\cos\alpha\cos\beta + a\sin\alpha\sin\beta =$$

$$\frac{1}{b}((a\cos\alpha)(b\cos\beta) + (a\sin\alpha)(b\sin\beta)) = \frac{|\overline{a} \cdot \overline{b}|}{|\overline{b}|}.$$

Finally, $\overline{a}_1$ can be expressed as $|\overline{a}_1|\overline{u}$, where $\overline{u}$ is a unit vector that lies in the direction of $\overline{b}$. In other words, $\overline{u} = \frac{\overline{b}}{|\overline{b}|}$. Therefore,

$$\text{proj}(\overline{a}, \overline{b}) = (\frac{\overline{a} \cdot \overline{b}}{|\overline{b}|})\frac{\overline{b}}{|\overline{b}|} = (\frac{\overline{a} \cdot \overline{b}}{|\overline{b}|^2})\overline{b}.$$

Example 1. Compute $\text{proj}((1, -5, 6), (-3, 4, 5))$.

Distance from Point to Plane

We may use the concept of projection to compute the distance from a point $\overline{a}$ to a plane having equation $\overline{w} \cdot \overline{x} = b$, where $\overline{w}, \overline{x} \in \mathcal{R}^n$, $\overline{w} \neq \overline{0}$, and $b \geq 0$ is a constant. Recall that $\overline{w}$ is orthogonal to the plane.

Theorem 2. Let $\overline{w} \cdot \overline{x} = b$ be the equation of a plane P, where $\overline{w}, \overline{x} \in \mathcal{R}^n$, $\overline{w} \neq \overline{0}$, and $b \geq 0$ is a constant. Then the distance from point $\overline{a}$ to P is

$$d(\overline{a}, P) = \frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|}.$$

Proof of Theorem 2. First notice that, if $\overline{x}$ is a point on the plane, then, by Theorem 1,

$$|\text{proj}(\overline{x}, \overline{w})| = \frac{\overline{w} \cdot \overline{x}}{|\overline{w}|} = \frac{b}{|\overline{w}|}$$

is a constant, and is equal to the distance from the origin to the plane. We refer to this distance as the **translation distance** of the plane, since it is the distance needed to translate the plane to the origin, in a direction parallel to $\overline{w}$.

We also use the following terminology to describe the spatial location of a point $\overline{a}$'s in relation to a plane P. We say that

1. $\overline{a}$ lies in **positive space** with respect to P iff $\overline{w} \cdot \overline{a} > b$,
2. $\overline{a}$ **lies on** P iff $\overline{w} \cdot \overline{a} = b$
3. $\overline{a}$ lies in **semi-negative space** with repsect to P iff $0 \leq \overline{w} \cdot \overline{a} < b$, and
4. $\overline{a}$ lies in **negative space** with respec to P iff $\overline{w} \cdot \overline{a} < 0$.

Notice that, if $\overline{a}$ is in positive space, then $\overline{w} \cdot \overline{a} - b > 0$, while $\overline{a}$ is in (semi-) negative space when $\overline{w} \cdot \overline{a} - b < 0$.

Next, notice that the distance from $\overline{a}$ to P is always measured along a direction that is parallel with $\overline{w}$, since $\overline{w}$ is orthogonal to P. Moreover, since $|\text{proj}(\overline{a}, \overline{w})|$ measures the distance from $\overline{a}$ to the origin along $\overline{w}$, it follows that the the distance from $\overline{a}$ to P may be obtained by adding/subtracting $\text{proj}(\overline{a}, \overline{w})$ to P's translation distance. The details of this are now worked out in the following cases of where $\overline{a}$ lies in relation to P.

Case 1: $\overline{a}$ lies in positive space with respect to P. Then $d(\overline{a}, P)$ is the difference between the distance from $\overline{a}$ to the origin along $\overline{w}$ and the distance from P to the origin along $\overline{w}$. In other words,

$$d(\overline{a}, P) = |\text{proj}(\overline{a}, \overline{w})| - \frac{b}{|\overline{w}|} = \frac{\overline{w} \cdot \overline{a}}{|\overline{w}|} - \frac{b}{|\overline{w}|} =$$

$$\frac{\overline{w} \cdot \overline{a} - b}{|\overline{w}|} = \frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|},$$

since $\overline{w} \cdot \overline{a} > b$.

Case 2: $\overline{a}$ lies on P. In this case,

$$d(\overline{a}, P) = 0 = \frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|},$$

since $\overline{w} \cdot \overline{a} = b$.

Case 3: $\overline{a}$ lies in semi-negative space with respect to P. Then $\text{proj}(\overline{a}, \overline{w})$ lies between P and the origin, and thus has the same direction as $\overline{w}$. Hence, $d(\overline{a}, P)$ is the translation distance of P *minus* $|\text{proj}(\overline{a}, \overline{w})|$; i.e.,

$$d(\overline{a}, P) = \frac{b}{|\overline{w}|} - \frac{|\overline{w} \cdot \overline{a}|}{|\overline{w}|} = \frac{b}{|\overline{w}|} - \frac{\overline{w} \cdot \overline{a}}{|\overline{w}|} =$$

$$\frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|},$$

where the second-to-last equality is due to the fact that $\overline{w} \cdot \overline{a} \geq 0$, and the last equality is due to $\overline{w} \cdot \overline{a} < b$.

Case 4: $\overline{a}$ lies in negative space with respect to P. Then $\text{proj}(\overline{a}, \overline{w})$ has a direction that is opposite of $\overline{w}$. Hence, $d(\overline{a}, P)$ is the translation distance of P *plus* $|\text{proj}(\overline{a}, \overline{w})|$; i.e.,

$$d(\overline{a}, P) = \frac{b}{|\overline{w}|} + \frac{|\overline{w} \cdot \overline{a}|}{|\overline{w}|} = \frac{b}{|\overline{w}|} - \frac{\overline{w} \cdot \overline{a}}{|\overline{w}|} =$$

$$\frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|},$$

where the last two equalities are due to the fact that $\overline{w} \cdot \overline{a} < 0 \leq b$.

Therefore, in all four cases,

$$d(\overline{a}, P) = \frac{|\overline{w} \cdot \overline{a} - b|}{|\overline{w}|}.$$

Example 2. Compute the distance from point $\overline{a} = (1, 2, 4)$ to the plane $x - y + 2z = 4$. Determine the point on the plane that is closest to $\overline{a}$.

Linear Equations

To begin, a **linear expression** over real-valued variables $x_1, \ldots, x_n$ is one of the form

$$a_1x_1 + \cdots + a_nx_n,$$

where a_i is a real-valued constant called the **coefficient** of x_i. For example, $3x_1 - 2.2x_2 + 7x_3$ and $-5x + 2y - 3z + 4w$ are linear expressions, but $2\sqrt{x} - 3y$ is not beccause the square-root function has been applied to variable x, which is not allowed. Linear expressions are also referred to as **linear combinations**.

A **linear inequality** is one of the form $a_1x_1 + \cdots + a_nx_n \leq b$, where the left side is a linear expression over variables $x_1, \ldots, x_n$, and b is a real-valued constant. Note that $\leq$ may be replaced with one of the other three inequality relations: $<$, $\geq$, and $>$. Examples of linear inequalities include $2x_1 - 3x_3 + 0.51x_4 \leq 7$, $3x - 5y > 2$, and $-4a + b - 9c - 10d \geq 20$. On the other hand, $3x^2 + 4y^2 \leq 16$ is not because $3x^2 + 4y^2$ is not a linear expression (why ?).

A **linear equation** is one of the form $a_1x_1 + \cdots + a_nx_n = b$. Examples of linear equations include $2x_1 - 3x_3 + 0.51x_4 = 7$, $3x - 5y = 2$, and $-4a + b - 9c - 10d = 20$. On the other hand, $y - \sin x = 0$ is not a linear equation (why ?).

A linear equation or inequality is an example of a **constraint**, which restricts the combination of values that a set of variables may assume. For example, in the equation $2x + 3y = 7$, if x is assigned the value 2, then y is forced to be assigned the value 1. Any other (x, y)-combination with $x = 2$ is invalid. Constraints arise naturally in problem modeling because every realm of human life is governed by constraints. For example, if an individual has a total of D dollars to invest in three different companies, and if x_i, $i = 1, 2, 3$, is the amount she will invest in company i, then necessarily $x_1 + x_2 + x_3 \leq D$.

Given a linear equation $a_1x_1 + \cdots + a_nx_n = b$, a **solution** to the equation is a sequence of numbers $s_1, \ldots, s_n$ which, when respectively substituted for the variables, $x_1, \ldots, x_n$, satisfies the equation in the sense that

$$a_1s_1 + \cdots + a_ns_n = b$$

is a true statement. For example, $x = 2, y = 1$ is a solution to equation $2x + 3y = 7$.

The **solution set** for an equation is the set of all possible solutions to the equation. For an equation with n variables, its solution set can be described with the help of $n - 1$ paramters.

Example 3. Find the solution set for the equation $2x + 3y = 7$.

Example 3 Solution. Equation $2x + 3y = 7$ has two variables and its solution set can be described using a single parameter t. To do this, set $y = t$. Then, solving for x in the equation gives

$$x = (7 - 3t)/2 = 7/2 + (-3/2)t.$$

Thus, the solution set for $2x + 3y = 7$ is $x = 7/2 - 3t/2$ and $y = t$, where t is any real number. Here we say that y is an **independent variable**, since it can be assigned any real value, while x is a **dependent variable** whose value depends on the value assigned to y.

Example 4. Represent the solution set $-2x - 5y + 4z = 10$ using two parameters s and t. What values of s and t will generate the solution $x = 4, y = 2, z = 2$?

Systems of Linear Equations

A set of linear equations over the same set of variables is called a **system of linear equations**, or **linear system**. Systems of linear equations arise in practice when each constraint has been modeled with an linear equation. For example,

$$5x_1 - 7x_2 + 3x_3 = 38$$

$$2x_1 + 4x_2 + 8x_3 = 22$$

represents a system of two equations over the variables x_1, x_2, x_3.

Given a system of linear equations over variables $x_1, \ldots, x_n$, a **system solution** is a sequence of numbers $s_1, \ldots, s_n$ which, when respectively substituted for the variables, $x_1, \ldots, x_n$, satisfies each of the equations in the system. For example $x_1 = 1, x_2 = -3, x_3 = 4$ is a solution to the above system. Moreover, a **system solution set** is the set of all system solutions.

We now review the Guass-Jordan (Variable Elimination) algorithm for finding the set of solutions to a system of linear equations. Since this algorithm only acts on the variable coefficients and constant of each equation, it serves well to convert the system into a single matrix, and then perform operations on the matrix rows.

The first step is to assume a variable ordering, such as $x_1, \ldots, x_n$. The next step is to transfer the coefficients into a **matrix**, which is a two-dimensional array of numbers. Moreover, assuming m equations and n variables, the matrix will have m rows and $n+1$ columns (the extra column is needed for the right-side constants). The resulting $m \times n$ matrix is called the **augmented matrix** for the system of equations. The first n columns are called the **variable columns**, while the last column is called the **constant** column. Moreover, the matrix that is formed by removing the constant column from the augmented matrix is called the **coefficient matrix**.

Example 5. Provide the augmented matrix for the following system of equations.

$$e_1 : \ 2x + 5y + 7z = 9$$

$$e_2 : \ -x + 2y + 4z = -3$$

$$e_3 : \ 3x - y - 6z = 6$$

Example 5 Solution. The following is the augmented matrix for the above system.

$$\begin{pmatrix} 2 & 5 & 7 & 9 \\ -1 & 2 & 4 & -3 \\ 3 & -1 & -6 & 6 \end{pmatrix}$$

Example 6. Provide the augmented matrix for the following system of equations. Assume the variable ordering x, y, z, w.

$$x - 3z = -2$$

$$5y + 8w = -2$$

$$2x + 7w = -4$$

$$6y - 5z = 0$$

Example 6 Solution.

The Gauss-Jordan algorithm takes as input an agumented matrix, and transforms it to a matrix that is in *reduced row-echelon form*. An augmented matrix M for a system of equations is in **reduced row-echelon form** iff the following statements hold.

REF1. Each row of M either consists of all zeros (tautology), all zeros except a 1 in the constant column (fallacy), or has a leading 1 in a variable column thar is preceded by all zeros (conditional).

REF2. If a column of M has a leading 1 from some conditional row, then all othe entries in that column are zero.

REF3. The rows of M are ordered top to bottom as follows: conditional rows, followed by fallacies and tautologies. Moreover, if r_1 and r_2 are two conditional rows, then r_1 occurs before r_2 iff its leading 1 occurs in an earlier column than r_2's leading 1.

Example 7. Explain why, of the following matrices, only M_4 is in reduced row-echelon form.

$$M_1 = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 2 & 0 & -3 \\ 0 & 0 & 1 & 7 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 1 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$M_3 = \begin{pmatrix} 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 6 \end{pmatrix}$$

$$M_4 = \begin{pmatrix} 0 & 1 & 0 & 2 & 7 \\ 0 & 0 & 1 & -4 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Notice that, for an augmented matrix in reduced row-echelon form, a variable column that has a leading 1 coincides with a dependent variable v. This is because v only occurs in one row (i.e. equation), and it is the first variable in this row with a nonzero coefficient. Moreover, once the (independent) variaibles in the equation have been assigned, then v's value can be determined.

Also, a variable column with no leading 1 corresponds with an independent variable. Such a variable can be assigned any real value, since, in any row (i.e. equation) that has a nonzero coefficient in that column, to the left of that column will be a dependent-variable column. This leads to the following theorem which is stated without proof.

Theorem 3. Let S be a system of linear equations whose augmented matrix M is in reduced row-echelon form. Then one of the following is true about S.

1. M has a fallacy row and S is inconsistent.
2. Every variable column of M has a leading 1, and so S has exactly one solution.
3. M has at least one variable column with no leading 1, and so S has infinitely many solutions.

We now describe a set of **elementary row operations** that are used by the Gauss-Jordan algorithm for converting an augmented matrix into one that is in reduced row-echelon form.

Constant Times Row If r is a row and k is a nonzero constant, then kr is the row whose i th entry is equal to ka_i, where a_i is the i th entry of r, for $i = 1, \ldots, n+1$. This is analogous to multiplying both sides of an equation by k.

Row Plus Row If r_1 and r_2 are rows, then $r_1 + r_2$ is the row whose i th entry is $a_i + b_i$, where a_i and b_i are the respective i th entries of r_1 and r_2, for $i = 1, \ldots, n+1$. This is analogous to adding to equations to form a new equation.

Swap Two Rows If r_1 and r_2 are two rows of a matrix, then this operation swaps the positions of r_1 and r_2.

The first two operations are often combined into a single operation of the form $kr_1 + r_2$; i.e. add a multiple of r_1 to r_2. So from here on we asssume that $kr_1 + r_2$ is one of the elementary operations.

When transforming an augmented matrix M by applying one of the above elementary operations, we assume that i) kr is replacing r in M, ii) $r_1 + r_2$ is replacing r_2 in M, iii) $kr_1 + r_2$ is replacing r_2 in M, and iv) $r_1 \leftrightarrow r_2$ is swapping rows r_1 and r_2 of M. Thus, if M has m rows, then it will still have m rows after applying one of the elementary operations. The proof of the following proposition is left as an exercise.

Proposition 1. Let M_0 be the augmented matrix for a system of linear equations S. If M_k is the matrix that results from applying k consecutive elementary row operations to M_0, then M_k is the augmented matrix of a linear system that has the same set of solutions as S does.

Gauss-Jordan Variable Elimination algorithm. Let A be the augmented matrix for a linear system of equations.

Mark each row and column of A as "independent".

Initialize the number d of dependent rows: $d \leftarrow 0$.

For each variable column c of A,

1. Let r be the first independent row of A that has a nonzero entry a in c (if no such r exists, then continue to the next variable column).
2. Divide r by a: $r \leftarrow \frac{1}{a} \cdot r$.
3. For every r' with nonzero entry a' in column c.

   Add $-a$ times r to r': $r' \leftarrow -ar + r'$.
4. Swap r with row $d+1$: $r \leftrightarrow r_{d+1}$.
5. Mark r and c as "dependent".
6. Increment d: $d \leftarrow d + 1$.

Theorem 4. The above Gauss-Jordan Variable Elimination algorithm transforms an augmented matrix A into one that is in reduced row-echelon form.

Proof of Theorem 4. Let A' be the matrix that results from applying the Gauss-Jordan algorithm on A. Let r be a row of A'. If r is marked as "independent", then its variable-column entries must all be zero. For if c is any variable column, if it is marked "indpendent", then all the independent rows, including r, have zero entries in this column. Also, if c is marked as dependent, then a multiple of a dependent row would have been added to r in order to eliminate its nonzero entry in column c (see Step 3 within the for loop). Thus, r could only possibly have a nonzero entry in its constant column. On the other hand, if r is marked as "dependent", then it must have a leading one (see Step 2 within the for loop). Moreover, the entries before this leading one must all be zero, since r was independent before being marked dependent, and, by the above reasoning, the previous entries of r must all be zero. Hence, REF1 is satisfied.

Also, REF2 is satisfied due to Step 3 within the for loop.

Finally REF3 is satisfied since i) the for loop examines the columns in order from left to right, and ii) Step 4 within the for loop ensures the dependent rows are ordered before independent rows.

Example 8. Use the Gauss-Jordan elimination algorithm to transform the following matrix into reduced row-echelon form.

$$\begin{pmatrix} 2 & 3 & 1 & -9 \\ -5 & 1 & 4 & 10 \\ 1 & -2 & 5 & 15 \end{pmatrix}$$

Find the solution set for the corresponding system.

Example 8 Solution.

$r_1 \leftrightarrow r_3$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ -5 & 1 & 4 & 10 \\ 2 & 3 & 1 & -9 \end{pmatrix}$$

$5r_1 + r_2$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & -9 & 29 & 85 \\ 2 & 3 & 1 & -9 \end{pmatrix}$$

$-2r_1 + r_3$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & -9 & 29 & 85 \\ 0 & 7 & -9 & -39 \end{pmatrix}$$

$(-1/9)r_2$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & 1 & -29/9 & -85/9 \\ 0 & 7 & -9 & -39 \end{pmatrix}$$

$-7r_2 + r_3$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & 1 & -29/9 & -85/9 \\ 0 & 0 & 122/9 & 244/9 \end{pmatrix}$$

$9/122r_3$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & 1 & -29/9 & -85/9 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$29/9r_3 + r_2$:

$$\begin{pmatrix} 1 & -2 & 5 & 15 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$2r_2 + r_1$:

$$\begin{pmatrix} 1 & 0 & 5 & 9 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$-5r_3 + r_1$:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Therefore, $x_1 = -1$, $x_2 = -3$, and $x_3 = 2$ is the unique solution.

A **homogeneous linear equation** is one of the form $a_1x_1 + \cdots + a_nx_n = 0$. In other words, a homogeneous equation has a constant term $b = 0$. A **homogeneous system** of linear equations is one in which each equation is homogeneous. Notice that a homogeneous system can never possess a fallacy, because the right side of each equation is always zero. Moreover, a homogeneous system will always possess the **trivial solution** $x_1 = x_2 = \cdots = 0$, since this assignment makes the left side of each equation evaluate to zero. From this we may conclude that a homogeneous system of equations will either have one (trivial) solution, or infinitely many solutions, depending on whether or not its reduced form has independent variables.

Axioms of a Vector Space

A **vector space** is a triple $(\mathcal{V}, +, \cdot)$, where

1. $\mathcal{V}$ is a set of elements, called **vectors**,
2. $+$ is a binary addition operation that assigns two vectors $\overline{u}, \overline{v} \in \mathcal{V}$, to the vector $\overline{u} + \overline{v} \in \mathcal{V}$, and
3. $\cdot$ is a binary scalar operation that assigns a real number r and a vector $\overline{v} \in \mathcal{V}$ to the vector $r \cdot \overline{v} = r\overline{v} \in \mathcal{V}$.

Moreover, $(\mathcal{V}, +, \cdot)$ satisfies the following addition axioms.

Associative $\overline{u} + (\overline{v} + \overline{w}) = (\overline{u} + \overline{v}) + \overline{w}$

Commutative $\overline{u} + \overline{v} = \overline{v} + \overline{u}$

Zero Element $\overline{v} + \overline{0} = \overline{0} + \overline{v} = \overline{v}$

Additive Inverse there exists an integer $-\overline{v}$ for which $\overline{v} + (-\overline{v}) = \overline{0}$

Finally, $(\mathcal{V}, +, \cdot)$ satisfies the following scalar axioms.

Unit Scalar $1\overline{v} = \overline{v}$

Associative $r(s\overline{v}) = (rs)\overline{v}$

Distributive over $\mathcal{V}$ $r(\overline{u} + \overline{v}) = r\overline{u} + r\overline{v}$

Distributive over $\mathcal{R}$ $(r + s)\overline{v} = r\overline{v} + s\overline{v}$

Vector Subspaces

Theorem 4. Let $\mathcal{V}$ be a vector space with addition and scalar operations $+$ and $\cdot$, and $\mathcal{W} \subseteq \mathcal{V}$ a nonempty subset of $\mathcal{V}$. Then $\mathcal{W}$ is a **subspace** of $\mathcal{V}$, meaning $\mathcal{W}$ is also a vector space under the same addition and scalar operations, iff $\mathcal{W}$ is closed under $+$ and $\cdot$. In other words, $\mathcal{W}$ is a subspace provided

1. For all $\overline{u}, \overline{v} \in \mathcal{W}$, $\overline{u} + \overline{v} \in \mathcal{W}$.
2. If $\overline{v} \in \mathcal{W}$ and r is a scalar, then $r \cdot \overline{v} \in \mathcal{W}$.

Proof of Theorem 4. Since addition is associative and commutative in $\mathcal{V}$, it must also be associative and commutative in $\mathcal{W}$. Also, since each of the scalar axioms hold in $\mathcal{V}$, they must also hold in $\mathcal{W}$.

Furthermore, given $\overline{v} \in \mathcal{W}$, $0\overline{v} = \overline{0} \in \mathcal{W}$ by 2. So $\mathcal{W}$ contains the identity element. Also by 2., for any $\overline{v} \in \mathcal{W}$, $-1\overline{v} = -\overline{v} \in \mathcal{W}$. So every vector in $\mathcal{W}$ also has its additive inverse in $\mathcal{W}$. Therefore, addition over $\mathcal{W}$ satisfies all the addition and scalar axioms.

Example 9. Let $\overline{w} \in \mathcal{R}^n$ be a vector, and consider the plane equation $\overline{w} \cdot \overline{x} = 0$. Show that the set of points on the plane form a vector subspace of $\mathcal{R}^n$.

Linear Spans

Let S be a subset of vector space $\mathcal{V}$. The **linear span** of S, denoted span(S), is the set of all possible linear combinations of the form $c_1\overline{v}_1 + \cdots + c_n\overline{v}_n$, where $\overline{v}_1, \ldots, \overline{v}_n \in S$, and $c_1, \ldots, c_n$ are scalars. Note that a linear combination in span(S) need not use all vectors in S. In fact, S could be an infinite set.

Theorem 5. Let S be a nonempty subset of vector space $\mathcal{V}$. Then span(S) is a vector subspace of $\mathcal{V}$. Moreover, any other subspace $\mathcal{W}$ of $\mathcal{V}$ that contains S, must also contain span(S). In other words, span(S) is the smallest subspace that contains S.

Proof of Theorem 5. To prove that span(S) is a subspace, it must be shown that span(S) is closed under both addition and scalar multiplication. Let $\overline{u}, \overline{v} \in$ span(S) be arbitrary. Then there are vectors $\overline{v}_1, \ldots, \overline{v}_n \in S$ for which $\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n$ and $\overline{v} = d_1\overline{v}_1 + \cdots + d_n\overline{v}_n$, where $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$ are scalars (why may we assume that $\overline{u}$ and $\overline{v}$ are linear combinations of the exact same vectors $\overline{v}_1, \ldots, \overline{v}_n$ from S?). Then

$$\overline{u} + \overline{v} = (c_1\overline{v}_1 + \cdots + c_n\overline{v}_n) + (d_1\overline{v}_1 + \cdots + d_n\overline{v}_n) = (c_1\overline{v}_1 + d_1\overline{v}_1) + \cdots + (c_n\overline{v}_n + d_n\overline{v}_n) =$$

$$(c_1 + d_1)\overline{v}_1 + \cdots + (c_n + d_n)\overline{v}_n$$

is a linear combination of vectors from S, and hence $\overline{u} + \overline{v} \in$ span(S). Similarly, if r is a scalar, then

$$r\overline{u} = r(c_1\overline{v}_1 + \cdots + c_n\overline{v}_n) = r(c_1\overline{v}_1) + \cdots + r(c_n\overline{v}_n) = (rc_1)\overline{v}_1 + \cdots + (rc_n)\overline{v}_n$$

is a linear combination of vectors from S, and hence $r\overline{u} \in$ span(S).

Finally, suppose $\mathcal{W}$ is a subspace of $\mathcal{V}$ and contains S. Let $\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n$ be an arbitrary element in span(S), where $\overline{v}_1, \ldots, \overline{v}_n \in S$ and $c_1, \ldots, c_n$ are scalars. Then, since $S \subseteq \mathcal{W}$, it follows that $\overline{v}_1, \ldots, \overline{v}_n \in \mathcal{W}$. Moreover, since $\mathcal{W}$ is a subspace, it is closed under both additiion and scalar multiplication. In other words, any linear combination of $\overline{v}_1, \ldots, \overline{v}_n$ must also be in $\mathcal{W}$. In particular $\overline{u} \in \mathcal{W}$. And since $\overline{u} \in$ span(S) was arbitrary, it follows that span(S) is a subset of $\mathcal{W}$, and hence span(S) is the smallest subspace that contains S.

Example 10. Express $(2, 0, 6)$ as a linear combination of $\overline{u} = (2, 1, 4)$, $\overline{v} = (1, -1, 3)$, and $\overline{w} = (3, 2, 5)$. Determine if these three vectors span $\mathcal{R}^3$.

Linear Independence

Another property of a set of vectors that proves very useful is that of linear independence. Given set $S = \{v_1, \ldots, v_n\}$ of vectors, S is said to be **linearly independent** iff the equation

$$c_1\overline{v}_1 + \cdots + c_n\overline{v}_n = \overline{0}$$

is only satisfied when $c_1 = \cdots = c_n = 0$. Otherwise S is said to be **linearly dependent**.

Example 11. Consider the vectors $(1,0,0)$, $(1,2,0)$, and $(1,2,3)$. We show that these vectors are linearly independent. Consider the equation.

$$c_1(1,0,0) + c_2(1,2,0) + c_3(1,2,3) = \overline{0}.$$

This implies the system of equations

$$c_1 + c_2 + c_3 = 0$$

$$2c_2 + 2c_3 = 0$$

$$c_3 = 0$$

where the i th equation $(i = 1,2,3)$ is obtained by equating the i th components on both sides of the linear-independence equation. From these equations it is clear that $c_1 = c_2 = c_3 = 0$ is the only solution. Therefore, the vectors are linearly independent.

Theorem 6. If $S = \{\overline{v}_1, \ldots, \overline{v}_n\}$ is a linearly independent set, and $\overline{u} \in \text{span}(S)$, then there is unique set of coefficients $c_1, \ldots, c_n$ for which

$$\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n.$$

Proof of Theorem 6. Suppose that $S = \{\overline{v}_1, \ldots, \overline{v}_n\}$ is a linearly independent set, and $\overline{u} \in \text{span}(S)$ has two sets of coefficients $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$ for which

$$\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n$$

and

$$\overline{u} = d_1\overline{v}_1 + \cdots + d_n\overline{v}_n.$$

Then subtracting the second equation from the first yields

$$\overline{0} = (c_1 - d_1)\overline{v}_1 + \cdots + (c_n - d_n)\overline{v}_n.$$

But since $\overline{v}_1, \ldots, \overline{v}_n$ are linearly independent, it follows that $c_1 - d_1 = \cdots = c_n - d_n = 0$, which implies $c_1 = d_1, \ldots, c_n = d_n$. Therefore, the two sets of coefficients are identical, and hence there is a unique set of coefficients for expressing $\overline{u}$ as a linear combination of $\overline{v}_1, \ldots, \overline{v}_n$.

Basis and Dimension

Let $\mathcal{V}$ be a vector space and suppose the linearly-indenpendent vectors $\overline{v}_1, \ldots, \overline{v}_n$ span $\mathcal{V}$. It then follows from Theorem 1 that every vector $\overline{u} \in \mathcal{V}$ can be uniquely represented as

$$\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n.$$

This is very convenient since, regardless of the kind of vectors that are in $\mathcal{V}$, each vector $\overline{u}$ can be represented by its unique set of coefficients $(c_1, \ldots, c_n)$. Thus, in some sense $\mathcal{V}$ is structurally the same as the vector space $\mathcal{R}^n$, since each vector $u \in \mathcal{V}$ can be uniquely represented as an n-tuple. Moreover, these n-tuples can be added and (scalar) multiplied. For example, the unique coefficients for $\overline{u} + \overline{v}$, can be obtained by adding the coefficients of $\overline{u}$ to the coefficients of $\overline{v}$ as n-tuples in $\mathcal{R}^n$.

From the above paragraph, we see that a linearly-independent set of vectors $\overline{v}_1, \ldots, \overline{v}_n$ that span $\mathcal{V}$ seems quite useful and fundamental for representing each vector in $\mathcal{V}$ as an n-tuple. For this reason, when $\overline{v}_1, \ldots, \overline{v}_n$ are linearly independent and span $\mathcal{V}$, then $\overline{v}_1, \ldots, \overline{v}_n$ is called a **basis** for $\mathcal{V}$, and n is called the **dimension** of $\mathcal{V}$.

Example 12. $\overline{e}_1 = (1, 0, 0)$, $\overline{e}_2 = (0, 1, 0)$, and $\overline{e}_3 = (0, 0, 1)$ is a basis for $\mathcal{R}^3$, since they are linearly independent, and any vector (x, y, z) in $\mathcal{R}^3$ may be written as $x\overline{e}_1 + y\overline{e}_2 + z\overline{e}_3$.

One aspect of the definitiion of vector-space dimension that warrants scrutiny is the question of whether or not the dimension n of a vector space is a unique number. In other words, can a vector space have two different finite bases that are of different sizes? The following theorem answers this in the negative. Before stating the theorem, we first prove a needed proposition about homogenous systems of equations.

Proposition 2. A homogeneous system of equations that has more variables than equations has an infinite number of solutions.

Proof of Proposition 2. Suppose a homogenous system of equations has m equations and $n > m$ variables. Consider the result of converting the system's augmented matrix to reduced row-echelon form. This form can have at most m rows with leading ones, and hence at most m dependent variables. Therefore, the system has at least one independent variable, and thus an infinite number of solutions.

Theorem 7. Suppose vector space $\mathcal{V}$ has a basis $\overline{v}_1, \ldots, \overline{v}_n$. Then any vectors $\overline{w}_1, \ldots, \overline{w}_n, \overline{w}_{n+1}$ in $\mathcal{V}$ are linearly dependent. Conclusion: if B_1 and B_2 are two finite bases for $\mathcal{V}$, then $|B_1| = |B_2|$.

Proof of Theorem 7. Consider the linear-independence equation

$$c_1\overline{w}_1 + \cdots + c_n\overline{w}_n + c_{n+1}\overline{w}_{n+1} = \overline{0}.$$

We must show that this can be satisfied by $c_1, \ldots, c_n, c_{n+1}$, where $c_i \neq 0$ for some $i = 1, \ldots, n, n+1$. Let $\overline{v}_1, \ldots, \overline{v}_n$ be a basis for $\mathcal{V}$. Then each w_i can be written as

$$\overline{w}_i = a_{1i}\overline{v}_1 + \cdots + a_{ni}\overline{v}_n,$$

for some coefficients $a_{1i}, \ldots, a_{ni}$. Moreover, substituting for each w_i in the linear-indenpendence equation yields,

$$c_1(a_{11}\overline{v}_1 + \cdots + a_{n1}\overline{v}_n) + \cdots + c_n(a_{1n}\overline{v}_1 + \cdots + a_{nn}\overline{v}_n) + c_{n+1}(a_{1,n+1}\overline{v}_1 + \cdots + a_{n,n+1}\overline{v}_n) = \overline{0}.$$

Next, grouping the coefficients in accordance with each $\overline{v}_i$ yields the equivalent equation

$$(c_1a_{11} + \cdots + c_na_{1n} + c_{n+1}a_{1,n+1})\overline{v}_1 + \cdots + (c_1a_{n1} + \cdots + c_na_{nn} + c_{n+1}a_{n,n+1})\overline{v}_n = \overline{0}.$$

But since $v_1, \ldots, v_n$ are linearly independent, this equation implies the following homogeneous system of equations in which each of the above $\overline{v}_i$ coefficients are set equal to zero.

$$c_1a_{11} + \cdots + c_na_{1n} + c_{n+1}a_{1,n+1} = 0$$

$$\vdots$$

$$c_1a_{n1} + \cdots + c_na_{nn} + c_{n+1}a_{n,n+1} = 0.$$

This is a homogeneous linear system with n equations and $n+1$ variables. Hence, by Proposition 2, it has infinitely many solutions for $(c_1, \ldots, c_n, c_{n+1})$, which means that the equation

$$c_1\overline{w}_1 + \cdots + c_n\overline{w}_n + c_{n+1}\overline{w}_{n+1} = \overline{0}$$

can be satisfied with a non-trivial solution (i.e. where some $c_i \neq 0$). Therefore, $\overline{w}_1, \ldots, \overline{w}_n, \overline{w}_{n+1}$ are linearly dependent.

Theorem 8. Let $\mathcal{V}$ be an n-dimensional vector space, and $S = \{\overline{v}_1, \ldots, \overline{v}_n\}$ be a set of $n \geq 1$ distinct vectors. Then S is a basis for $\mathcal{V}$ iff either S spans $\mathcal{V}$ or S is a linearly independent set. In other words, so long as S has n vectors, having one of the two properties (linear independence or spanning $\mathcal{V}$) will imply also having the second property.

Proof of Theorem 8. First assume S spans $\mathcal{V}$. To show that S is a basis for $\mathcal{V}$ we must show that it is a linearly independent set. If $n = 1$, this is certainly true. If $n > 1$ and S is a linearly dependent set, then by Exercise 23 of this lecture there is a vector $\overline{v}$ in S that can be written as a linear combination of other vectors in S. Moreover, by Exercise 24 $\overline{v}$ can be removed from S, and yet S still spans $\mathcal{V}$. This process can be continued until S has been reduced to a linearly-independent subset S', which spans $\mathcal{V}$. Then, by definition, S' is a basis for $\mathcal{V}$, but has fewer than n vectors, which contradicts Theorem 7. Therefore, S must be a linearly independent set.

Now assume that S is a linearly independent set. Let $\overline{u}_1, \ldots, \overline{u}_n$ be a basis for $\mathcal{V}$. If S does not span $\mathcal{V}$, then one or more of the vectors $\overline{u}_1, \ldots, \overline{u}_n$ are not in span(S) (why?). If all of these missing basis vectors are added to S to form the set S', then S' is still linearly independent (why?), and now spans $\mathcal{V}$. Then, by definition, S' is a basis for $\mathcal{V}$, but has more than n vectors, which contradicts Theorem 7. Therefore, S must in fact span $\mathcal{V}$.

Exercises

1. Find $\text{proj}(\overline{a}, \overline{b})$ for the following vectors.

   a. $\overline{a} = (2, 1)$, $\overline{b} = (-3, 2)$

   b. $\overline{a} = (2, 6)$, $\overline{b} = (-9, 3)$

   c. $\overline{a} = (-7, 1, 3)$, $\overline{b} = (5, 0, 1)$

   d. $\overline{a} = (0, 0, 1)$, $\overline{b} = (8, 3, 4)$

2. Given the plane equation $-2x+5y-3z = 7$, determine its translation distance; i.e. the distance from the origin to plane.

3. Given the plane equation $-2x+5y-3z = 7$ for plane P, determine where each of the following points lie with respect to P (i.e. positive space, on P, negative space, or semi-negative space).

   a. $(1, 2, 1/3)$

   b. $(2, 1, 3)$

   c. $(2, 3, 0)$

   d. $(0, 1, 0)$

4. In each part, find the distance between the point and the plane.

   a. $(1, -2, 3)$, $2x - 2y + z = 4$

   b. $(0, 1, 5)$, $3x + 6y - 2z - 5 = 0$

   c. $(7, 2, -1)$, $20x + 4y - 5z = 0$

5. Find the solution set for $6x - 7y + z = 3$

6. Provide the augmented matrix for the following system of linear equations.

$$x_1 - 2x_2 = 0$$

$$3x_1 + 4x_2 = -1$$

$$2x_1 - x_2 = 3$$

7. Find a linear system of equations that corresponds with the following augmented matrix.

$$\left(\begin{array}{ccc|c} 3 & 0 & -1 & 1 \\ -2 & 1 & 4 & -3 \\ 0 & 1 & 3 & 3 \end{array}\right)$$

8. Solve the following linear system using Gauss-Jordan elimination.

$$x_1 + x_2 + 2x_3 = 8$$

$$-x_1 - 2x_2 + x_3 = 1$$

$$3x_1 - 7x_2 + 4x_3 = 10$$

9. Solve the following linear system using Gauss-Jordan elimination.

$$x - y + 2z - w = -1$$

$$2x + y - 2z - 2w = -2$$

$$-x + 2y - 4z + w = 1$$

$$3x - 3w = -3$$

10. Solve the following linear system using Gauss-Jordan elimination.

$$2x + 2y + 2z = 0$$

$$-2x + 5y + 2z = 0$$

$$-7x + 7y + z = 0$$

11. Solve the following linear system using Gauss-Jordan elimination.

$$2x_1 - 3x_2 = -2$$

$$2x_1 + x_2 = 1$$

$$3x_1 + 2x_2 = 1$$

12. Find a linear equation over constants a, b, and c in order for the system to have at least one solution.

$$x_1 + x_2 + 2x_3 = a$$

$$x_1 + x_3 = b$$

$$2x_1 + x_2 + 3x_3 = c$$

13. Notice that a linear equation $a_1x_1 + \cdots + a_nx_n = b$ may be written as the vector equation $\overline{a} \cdot \overline{x} = b$, where $\overline{a} = (a_1, \ldots, a_n$, and $\overline{x} = (x_1, \ldots, x_n$. Use this notation and properties of the dot product to prove the following.

   a. If $\overline{x}$ satisfies the linear equation $\overline{a} \cdot \overline{x} = b$, then it satisfies the equation $(k\overline{a}) \cdot \overline{x} = kb$, for any real number k.

   b. If $\overline{x}$ satisfies the linear equations $\overline{a} \cdot \overline{x} = b_1$ and $\overline{d} \cdot \overline{x} = b_2$, then it satisfies the equation $(\overline{a} + \overline{d}) \cdot \overline{x} = b_1 + b_2$.

   c. If $\overline{x}$ satisfies the linear equations $\overline{a} \cdot \overline{x} = b_1$ and $(\overline{a} + \overline{d}) \cdot \overline{x} = b_1 + b_2$, then it satisfies the equation $\overline{d} \cdot \overline{x} = b_2$, then it satisfies the equation.

   Conclude that elementary row operations performed on an augmented matrix do not change the set of solutions for the corresponding system of equations.

14. Prove that $\mathcal{R}^2$ satisfies the four vector-space addition axioms.

15. Prove that $\mathcal{R}^2$ satisfies the four vector-space scalar axioms and the four vector-space scalar axioms.

16. Prove the following statements are true in a vector space, where r is an arbitrary scalar, and $\overline{v}$ is an arbitrary vector.

    a. $0\overline{v} = \overline{0}$

    b. $r\overline{0} = \overline{0}$

    c. $(-1)\overline{v} = -\overline{v}$

    d. if $r\overline{v} = \overline{0}$, then either $r = 0$ or $\overline{v} = \overline{0}$

17. Which of the following are linear combinations of $\overline{u} = (1, -1, 3)$, and $\overline{v} = (2, 4, 0)$? a) $(3, 3, 3)$, b) $(4, 2, 6)$, c) $(1, 5, 6)$, d) $(0, 0, 0)$

18. Express $(2, 2, 4)$ as a linear combination of $\overline{u} = (2, 1, 4)$, $\overline{v} = (1, -1, 3)$, and $\overline{w} = (3, 2, 5)$.

19. Do the vectors $\overline{v}_1 = (1, 3, 3)$, $\overline{v}_2 = (1, 3, 4)$, $\overline{v}_3 = (1, 4, 3)$, $\overline{v}_4 = (6, 2, 1)$ span $\mathcal{R}^3$? Show work.

20. Do the vectors $\overline{v}_1 = (3, 1, 4)$, $\overline{v}_2 = (2, -3, 5)$, $\overline{v}_3 = (5, -2, 9)$, $\overline{v}_4 = (1, 4, -1)$ span $\mathcal{R}^3$? Show work.

21. Determine if vectors $(2, -1, 4)$, $(3, 6, 2)$, and $(2, 10, -4)$ are linearly independent. If yes, explain how we may conclude that the vectors form a basis for $\mathcal{R}^3$.

22. Prove that if $\overline{u}$ can be expressed as a linear combination of $\overline{v}_1, \ldots, \overline{v}_n$, then $\overline{u}, \overline{v}_1, \ldots, \overline{v}_n$ are linearly dependent.

23. Prove the converse of Exercise 22: if S is a linearly dependent set of two or more vectors, then at least one vector in S can be written as a linear combination of one or more other vectors in S.

24. Prove that if S has a vector $\overline{v}$ that is a linear combination of other vectors in S, then removing $\overline{v}$ from S does not change the span of S.

25. Recall that the set of solutions to a homogeneous system is a subspace of $\mathcal{R}^n$. For this reason it is called the **solution space**. Provide a basis for the solution space of the homogeneous system

$$x_1 + x_2 - x_3 = 0$$

$$-2x_1 - x_2 + 2x_3 = 0$$

$$-x_1 + x_3 = 0.$$

26. Consider the subspace of $\mathcal{R}^4$ consisting of all vectors (a, b, c, d) for which $d = a + b$ and $c = a - b$. Provide a basis for this subspace.

Exercise Solutions

1. Find $\text{proj}(\overline{a}, \overline{b})$ for the following vectors.

   a. $(\frac{12}{13}, -\frac{8}{13})$
   b. $(0, 0)$
   c. $(-\frac{80}{13}, 0, -\frac{16}{13})$
   d. $(\frac{32}{89}, \frac{12}{89}, \frac{16}{89})$

2. The translation distance is $\frac{b}{|\overline{w}|}$, where $b = 7$ and $\overline{w} = (-2, 5, 3)$. Therefore, the translation distance is $7/\sqrt{38}$.

3. Given the plane equation $-2x + 5y - 3z = 7$ for plane P, determine where each of the following points lie with respect to P (i.e. positive space, on P, negative space, or semi-negative space).

   a. $(1, 2, 1/3)$: on the plane
   b. $(2, 1, 3)$ negative space
   c. $(2, 3, 0)$: positive space
   d. $(0, 1, 0)$: semi-negative space

4. a. $5/3$
   b. $9/7$
   c. $137/21$

5. Find the solution set for $6x - 7y + z = 3$. Set $y = s$, $z = t$, where s and t can be any real numbers. Then $x = 1/2 + 7s/6 - t/6$.

6. The augmented matrix is

$$\begin{pmatrix} 1 & -2 & 0 \\ 3 & 4 & -1 \\ 2 & -1 & 3 \end{pmatrix}.$$

7. The following linear system of equations corresponds with the augmented matrix.

$$3x - z = 1$$

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

$$y + 3z = 3$$

8. The augmented matrix for the linear system system is

$$\begin{pmatrix} 1 & 1 & 2 & 8 \\ -1 & -2 & 1 & 1 \\ 3 & -7 & 4 & 10 \end{pmatrix}.$$

   Transform into reduced-row echelon form.

$r_1 + r_2$, $-3r_1 + r_3$:

$$\begin{pmatrix} 1 & 1 & 2 & 8 \\ 0 & -1 & 3 & 9 \\ 0 & -10 & -2 & -14 \end{pmatrix}.$$

$-r_2$, $10r_2 + r_3$:

$$\begin{pmatrix} 1 & 1 & 2 & 8 \\ 0 & 1 & -3 & -9 \\ 0 & 0 & -32 & -104 \end{pmatrix}.$$

$-1/32r_3$:

$$\begin{pmatrix} 1 & 1 & 2 & 8 \\ 0 & 1 & -3 & -9 \\ 0 & 0 & 1 & 13/4 \end{pmatrix}.$$

$-r_2 + r_1$:

$$\begin{pmatrix} 1 & 0 & 5 & 17 \\ 0 & 1 & -3 & -9 \\ 0 & 0 & 1 & 13/4 \end{pmatrix}.$$

$3r_3 + r_2, -5r_3 + r_1$:

$$\begin{pmatrix} 1 & 0 & 0 & 3/4 \\ 0 & 1 & 0 & 3/4 \\ 0 & 0 & 1 & 13/4 \end{pmatrix}.$$

Solution: $x_1 = 3/4$, $x_2 = 3/4$, $x_3 = 13/4$.

9. The augmented matrix for the linear system system is

$$\begin{pmatrix} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{pmatrix}.$$

Transform into reduced-row echelon form.

$-2r_1 + r_2$, $r_1 + r_3$, $-3r_1 + r_4$:

$$\begin{pmatrix} 1 & -1 & 2 & -1 & -1 \\ 0 & 3 & -6 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -6 & 0 & 0 \end{pmatrix}.$$

$r_2 \leftrightarrow r_3$, $-3r_2 + r_3$, $-3r_2 + r_4$:

$$\begin{pmatrix} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$r_2 + r_1$:

$$\begin{pmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Solution: $x_1 = t - 1$, $x_2 = 2s$, $x_3 = s$, $x_4 = t$.

10. Solution: $x = -3t$, $y = -4t$, $z = 7t$.

11. No solution. The system is inconsistent.

12. The augmented matrix for the linear system system is

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 1 & 0 & 1 & b \\ 2 & 1 & 3 & c \end{pmatrix}.$$

Transform into reduced-row echelon form.

$-r_1 + r_2$, $-2r_1 + r_3$:

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 0 & -1 & -1 & b-a \\ 0 & -1 & -1 & c-2a \end{pmatrix}.$$

$-r_2$, $r_2 + r_3$:

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 0 & 1 & 1 & a-b \\ 0 & 0 & 0 & c-a-b \end{pmatrix}.$$

Thus, the system will be consistent provided row 3 does not represent a fallacy; i.e. provided $-a - b + c = 0$.

13. a. Suppose $\overline{x}$ satisfies the linear equation $\overline{a} \cdot \overline{x} = b$. Then

$$(k\overline{a}) \cdot \overline{x} = k(\overline{a} \cdot \overline{x}) = kb,$$

where the first equality is due to the *scalar-associativity* property of the dot product: $(k\overline{u}) \cdot \overline{v} = k(\overline{u} \cdot \overline{v})$.

b. Now assume $\overline{x}$ satisfies the linear equations $\overline{a} \cdot \overline{x} = b_1$ and $\overline{d} \cdot \overline{x} = b_2$. Then

$$(\overline{a} + \overline{d}) \cdot \overline{x} = \overline{a} \cdot \overline{x} + \overline{d} \cdot \overline{x} = b_1 + b_2,$$

where the first equality is due to the *additivity* property of the dot product:

$$(\overline{u} + \overline{v}) \cdot \overline{w} = \overline{u} \cdot \overline{w} + \overline{v} \cdot \overline{w}.$$

c. Now assume $\overline{x}$ satisfies the linear equations $\overline{a} \cdot \overline{x} = b_1$ and $(\overline{a} + \overline{d}) \cdot \overline{x} = b_1 + b_2$. Then,

$$\overline{d} \cdot \overline{x} = ((\overline{a} - \overline{a}) + \overline{d}) \cdot \overline{x} = ((\overline{a} + \overline{d}) - \overline{a}) \cdot \overline{x} =$$

$$(\overline{a} + \overline{d}) \cdot \overline{x} - \overline{a} \cdot \overline{x} = (b_1 + b_2) - b_1 = b2.$$

14. Let $\overline{u} = (a, b)$, $\overline{v} = (c, d)$, and $\overline{w} = (e, , f)$

Associative

$$(\overline{u}+\overline{v})+\overline{w} = ((a,b)+(c,d))+(e,f)) = (a+c,b+d)+(e,f) = ((a+c)+e,(b+d)+f) =$$

$$(a+(c+e),b+(d+f)) = (a,b)+(c+e,d+f) = (a,b)+((c,d)+(e,f)) = \overline{u}+(\overline{v}+\overline{w}),$$

where the fourth equality is due to the associativity of real numbers.

Commutative

$$\overline{u}+\overline{v} = (a,b)+(c,d) = (a+c,b+d) = (c+a,d+b) = (c,d)+(a,b) = \overline{v}+\overline{u},$$

where the third equality is due to the commutativity of real numbers.

Zero Vector

$$\overline{v}+\overline{0} = (c,d)+(0,0) = (c+0,d+0) = (c,d) = (0+c,0+d) = (0,0)+(c,d) = \overline{0}+\overline{v} = \overline{v}.$$

Additive Inverse Define $-\overline{v}$ as $-\overline{v} = (-c,-d)$. Then

$$\overline{v}+(-\overline{v}) = (c,d)+(-c,-d) = (c+(-c),d+(-d)) = (0,0) = \overline{0}.$$

15. **Unit Scalar** Let $\overline{v} = (a, b)$. Then

$$1\overline{v} = 1(a,b) = (1a,1b) = (a,b) = \overline{v}.$$

Notice that, in $(1a, 1b)$, $1a$ refers to real-number multiplication, and *not* the scaling operation. The same is true for $1b$.

Associative

$$r(s\overline{v}) = r(s(a,b)) = r(sa,sb) = (r(sa),r(sb)) =$$
$$((rs)a,(rs)b) = (rs)(a,b) = (rs)\overline{v},$$

where the fourth equality is from the associative property of real-number multiplication.

Distributive over $\mathcal{V}$ Now let $\overline{u} = (c, d)$. Then

$$r(\overline{u}+\overline{v}) = r((a,b)+(c,d)) = r(a+c,b+d) = (r(a+c),r(b+d)) = (ra+rc,rb+rd) =$$

$$(ra,rb)+(rc,rd) = r(a,b)+r(c,d) = r\overline{u}+r\overline{v},$$

where the fourth equality is from the distributive property of real-number multiplication.

Distributive over $\mathcal{R}$

$$(r+s)\overline{v} = (r+s)(a,b) = ((r+s)a,(r+s)b) = (ra+sa,rb+sb) =$$

$$(ra,rb)+(sa,sb) = r(a,b)+s(a,b) = r\overline{v}+s\overline{v},$$

where the third equality is from the distributive property of real-number multiplication.

16. a.

$$0\overline{v} = (0+0)\overline{v} = 0\overline{v} + 0\overline{v},$$

by the distributivity over scalars axiom. But for any vector $\overline{u}$ that satisfies $\overline{u} = \overline{u} + \overline{u}$, we can add $-\overline{u}$ to both sides of the equation to get

$$\overline{u} + (-\overline{u}) = \overline{0} = (\overline{u} + \overline{u}) + (-\overline{u}) = \overline{u} + (\overline{u} + (-\overline{u})) = \overline{u} + \overline{0} = \overline{u},$$

which implies that $\overline{u} = \overline{0}$. Therefore, since $0\overline{v} = 0\overline{v} + 0\overline{v}$, it follows that $0\overline{v} = \overline{0}$.

b. $r\overline{0} = r(\overline{0} + \overline{0}) = r\overline{0} + r\overline{0}$, by the distributivity over vectors axiom. Then subtracting $r\overline{0}$ from both sides yields $\overline{0} = r\overline{0}$, or $r\overline{0} = \overline{0}$.

c. By Part a), $\overline{0} = 0\overline{v} = (1 + (-1))\overline{v} = v + (-1)\overline{v}$, by the distributivity over scalars axiom. But this means that $(-1)\overline{v}$ is the additive inverse of $\overline{v}$. In other words, $(-1)\overline{v} = -\overline{v}$.

d. If $r = 0$, then we are done. So suppose $r \neq 0$. Then, since $r\overline{v} = \overline{0}$, we must have $(1/r)r\overline{v} = (1/r)\overline{0} = \overline{0}$ by Part b. But by associativity, $(1/r)r\overline{v} = (1/r \cdot r)\overline{v} = 1\overline{v} = \overline{v}$. Therefore, $\overline{v} = \overline{0}$.

17. To see if vector (a, b, c) is a linear combination of $\overline{u} = (1, -1, 3)$ and $\overline{v} = (2, 4, 0)$, we must find c_1 and c_2 so that the following system of equations has a solution.

$$c_1 + 2c_2 = a$$

$$-c_1 + 4c_2 = b$$

$$3c_1 = c$$

The third equation gives $c_1 = c/3$, while adding the first two equations gives $c_2 = (a+b)/6$. For $(a, b, c) = (3, 3, 3)$, this implies $c_1 = c_2 = 1$, while $(a, b, c) = (4, 2, 6)$ implies $c_1 = 2$ and $c_2 = 1$. Check that these coefficients are correct for both a) and b). On the other hand, $(a, b, c) = (1, 5, 6)$ also implies $c_1 = 2$ and $c_2 = 1$, but $2\overline{u} + \overline{v} \neq (1, 5, 6)$. Finally, $0\overline{u} + 0\overline{v} = (0, 0, 0)$.

18. We must find constants c_1, c_2, c_3 such that $c_1\overline{u} + c_2\overline{v} + c_3\overline{w} = (2, 2, 3)$. In other words, we must solve the system of equations

$$2c_1 + c_2 + 3c_3 = 2$$

$$c_1 - c_2 + 2c_3 = 2$$

$$4c_1 + 3c_2 + 5c_3 = 4$$

Solving the system gives $c_1 = 3$, $c_2 = c_3 = -1$.

19. Yes.

20. No.

21. The linear-independence equation

$$c_1(2, -1, 4) + c_2(3, 6, 2) + c_3(2, 10, -4) = 0$$

implies the homogeneous system

$$2c_1 + 3c_2 + 2c_3 = 0$$

$$-c_1 + 6c_2 + 10c_3 = 0$$
$$4c_1 + 2c_2 - 4c_3 = 0.$$

Moreover, the reduced-row-echelon form of the coefficient matrix equals the identity matrix. Thus, the system only has the trivial solution. Therefore, the vectors are linearly independent. Moreover, by Theorem 8, the vectors from a basis for $\mathcal{R}^3$ since $\mathcal{R}^3$ is a 3-dimensional vector space.

22. If $\overline{u} = c_1\overline{v}_1 + \cdots + c_n\overline{v}_n$, then

$$1\overline{u} - c_1\overline{v}_1 - \cdots - c_n\overline{v}_n = \overline{0}.$$

Therefore, $\overline{u}, \overline{v}_1, \ldots, \overline{v}_n$ are linearly dependent, since the coefficients $(1, -c_1, \ldots, -c_n)$ are not all zero.

23. Suppose S is a linearly dependent set of two or more vectors. If S includes the zero vector, and if $\overline{u}$ is another vector in S, then $\overline{0} = 0\overline{u}$, which proves the statement. So assume $\overline{0} \notin S$. Then, since S is linearly dependent, there exist vectors $\overline{v}_1, \ldots, \overline{v}_n \in S$, $n \geq 2$, for which

$$c_1\overline{v}_1 + \cdots + c_n\overline{v}_n = \overline{0},$$

where the coefficients $c_1, \ldots, c_n$ are not all zero. Without loss of generality, assume that $c_1 \neq 0$. Then, solving for $\overline{v}_1$ gives

$$\overline{v}_1 = -\frac{c_2}{c_1}\overline{v}_2 - \cdots - \frac{c_n}{c_1}\overline{v}_n,$$

showing that $\overline{v}_1$ is a linear combination of the other $n-1$ vectors.

24. Suppose $\overline{v} \in S$ can be written as $c_1\overline{u}_1 + \cdots + c_n\overline{u}_n$, where $\overline{u}_1, \ldots, \overline{u}_n$ are vectors in S. Let $\overline{w}$ be in the span of S and suppose $\overline{w} = d_1\overline{x}_1 + \cdots + d_{m-1}\overline{x}_{m-1} + d_m\overline{v}$, where $\overline{x}_1, \ldots, \overline{x}_{m-1}$ are vectors in S. In other words, $\overline{w}$ is a linear combination of vectors from S, including vector $\overline{v}$. Then substituting $c_1\overline{u}_1 + \cdots + c_n\overline{u}_n$ for $\overline{v}$ gives

$$\overline{w} = d_1\overline{x}_1 + \cdots + d_{m-1}\overline{x}_{m-1} + d_m(c_1\overline{u}_1 + \cdots + c_n\overline{u}_n) =$$

$$d_1\overline{x}_1 + \cdots + d_{m-1}\overline{x}_{m-1} + (d_mc_1)\overline{u}_1 + \cdots + (d_mc_n)\overline{u}_n.$$

Thus $\overline{w}$ is a linear combination of vectors from S that do not include $\overline{v}$. Therefore, removing $\overline{v}$ from S does not change the span of S.

25. Basis: $(1, 0, 1)$ (or any multiple of this vector).

26. Notice that a and b can be independently assigned values, say $a = s$ and $b = t$, then $d = s + t$ and $c = s - t$. In other words,

$$(a, b, c, d) = s(1, 0, 1, 1) + t(0, 1, -1, 1).$$

Therefore, $(1, 0, 1, 1)$ and $(0, 1, -1, 1)$ is a basis for this subspace.