Linear Algebra (part 2) : Vector Spaces (by Evan Dummit, 2017, v. 2.60)

Contents

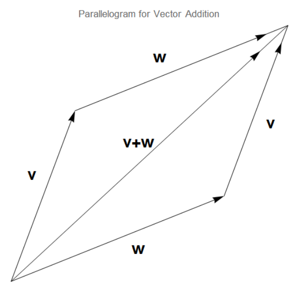
2 Vector Spaces

In this chapter we will introduce the notion of an abstract vector space, which is, ultimately, a generalization of the ideas inherent in studying vectors in 2- or 3-dimensional space (which we discuss first, as motivation). We will study vector spaces from an axiomatic perspective, discuss the notions of span and linear independence, and ultimately explain why every vector space possesses a linearly independent spanning set called a "basis". We then close with a discussion of computational aspects of vector spaces arising in the context of matrices: the row space, the column space, and the nullspace.

2.1 Vectors in $\mathbb{R}^n$

- A vector, as we typically think of it, is a quantity which has both a magnitude and a direction. This is in contrast to a scalar, which carries only a magnitude.
  - Real-valued vectors are extremely useful in just about every aspect of the physical sciences, since just about everything in Newtonian physics is a vector: position, velocity, acceleration, forces, etc. There is also "vector calculus" (namely, calculus in the context of vector fields) which is typically part of a multivariable calculus course; it has many applications to physics as well.
- We often think of vectors geometrically, as a directed line segment (having a starting point and an endpoint).
- Algebraically, we write a vector as an ordered tuple of coordinates: we denote the n-dimensional vector from the origin to the point $(a_1, a_2, \cdots, a_n)$ as $\mathbf{v} = \langle a_1, a_2, \cdots, a_n \rangle$, where the a_i are real-number scalars.
  - Some vectors: $\langle 1, 2\rangle$, $\langle 3, 5, -1\rangle$, $\left\langle -\pi, e^2, 27, 3, \frac{4}{3}, 0, 0, -1 \right\rangle$.
  - Notation: We use angle brackets $\langle \cdot \rangle$ rather than parentheses $(\cdot)$ so as to draw a visual distinction between a vector and the coordinates of a point in space. We also draw arrows above vectors (as $\vec{v}$) or typeset them in boldface (as $\mathbf{v}$) in order to set them apart from scalars. Boldface is hard to produce without a computer, so it is highly recommended to use the arrow notation $\vec{v}$ when writing by hand.

- We often think of, and draw, vectors as directed line segments.
  - However, technically speaking, vectors are slightly different from directed segments, because we don't care where a vector starts: we only care about the difference between the starting and ending positions. Thus: the directed segment whose start is $(0, 0)$ and end is $(1, 1)$ and the segment starting at $(1, 1)$ and ending at $(2, 2)$ represent the *same vector* $\langle 1, 1\rangle$.
- We can add vectors (provided they are of the same length!) in the obvious way, one component at a time: if $\mathbf{v} = \langle a_1, \cdots, a_n\rangle$ and $\mathbf{w} = \langle b_1, \cdots, b_n\rangle$ then $\mathbf{v} + \mathbf{w} = \langle a_1 + b_1, \cdots, a_n + b_n\rangle$.
  - We can justify this using our geometric idea of what a vector does: $\mathbf{v}$ moves us from the origin to the point $(a_1, \cdots, a_n)$. Then $\mathbf{w}$ tells us to add $\langle b_1, \cdots, b_n\rangle$ to the coordinates of our current position, and so $\mathbf{w}$ moves us from $(a_1, \cdots, a_n)$ to $(a_1+b_1, \cdots, a_n+b_n)$. So the net result is that the sum vector $\mathbf{v}+\mathbf{w}$ moves us from the origin to $(a_1 + b_1, \cdots, a_n + b_n)$, meaning that it is just the vector $\langle a_1 + b_1, \cdots, a_n + b_n\rangle$.
  - Geometrically, we can think of vector addition using a parallelogram whose pairs of parallel sides are $\mathbf{v}$ and $\mathbf{w}$ and whose diagonal is $\mathbf{v} + \mathbf{w}$:



- We can also 'scale' a vector by a scalar, one component at a time: if r is a scalar, then we have $r\,\mathbf{v} = \langle ra_1, \cdots, ra_n\rangle$.
  - Again, we can justify this by our geometric idea of what a vector does: if $\mathbf{v}$ moves us some amount in a direction, then $\dfrac{1}{2}\mathbf{v}$ should move us half as far in that direction. Analogously, $2\mathbf{v}$ should move us twice as far in that direction, while $-\mathbf{v}$ should move us exactly as far, but in the opposite direction.
- Example: If $\mathbf{v} = \langle -1, 2, 2\rangle$ and $\mathbf{w} = \langle 3, 0, -4\rangle$ then $2\mathbf{w} = \boxed{\langle 6, 0, -8\rangle}$, and $\mathbf{v}+\mathbf{w} = \boxed{\langle 2, 2, -2\rangle}$. Furthermore, $\mathbf{v}-2\mathbf{w} = \boxed{\langle -7, 2, 10\rangle}$.
- The arithmetic of vectors in $\mathbb{R}^n$ satisfies several algebraic properties that follow more or less directly from the definition:
  - Addition of vectors is commutative ($\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$) and associative ($\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$).
  - There is a zero vector $\mathbf{0}$ (namely, the vector with all entries zero), and every vector $\mathbf{v}$ has an additive inverse $-\mathbf{v}$ with $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.
  - Scalar multiplication distributes over addition of vectors ($r(\mathbf{v} + \mathbf{w}) = r\mathbf{v} + r\mathbf{w}$) and scalars ($(r + s)\mathbf{v} = r\mathbf{v} + s\mathbf{v}$).
- With vectors in $\mathbb{R}^n$, we also have a quantity that resembles a product, the dot product:
- Definition: The dot product of two vectors $\mathbf{v}_1 = \langle a_1, \ldots, a_n\rangle$ and $\mathbf{v}_2 = \langle b_1, \ldots, b_n\rangle$ is defined to be the scalar $\boxed{\mathbf{v}_1 \cdot \mathbf{v}_2 = a_1b_1 + a_2b_2 + \cdots + a_nb_n}$.
  - Note: The dot product of two vectors is a scalar, *not* a vector! (For this reason, the dot product is sometimes called the "scalar product" of two vectors.)
  - Example: The dot product $\langle 1, 2\rangle \cdot \langle 3, 4\rangle$ is $(1)(3) + (2)(4) = \boxed{11}$.
  - Example: The dot product $\langle -1, 2, 2\rangle \cdot \langle 3, 0, -4\rangle$ is $(-1)(3) + (2)(0) + (2)(-4) = \boxed{-11}$.

- The dot product obeys several very nice properties reminiscent of standard multiplication. For any vectors $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}$, and any scalar r, we can verify the following properties directly from the definition:
    - The dot product is commutative: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.
    - The dot product distributes over addition: $(\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{w} = (\mathbf{v}_1 \cdot \mathbf{w}) + (\mathbf{v}_2 \cdot \mathbf{w})$.
    - Scalars can be factored out of a dot product: $(r\,\mathbf{v}) \cdot \mathbf{w} = r\,(\mathbf{v} \cdot \mathbf{w})$.
- One of the chief motivations of the dot product is to provide a way to measure angles between vectors. First, we must define the length of a vector:
- Definition: The norm (length, magnitude) of the vector $\mathbf{v} = \langle a_1, \ldots, a_n \rangle$ is $||\mathbf{v}|| = \sqrt{(a_1)^2 + \cdots + (a_n)^2}$.
    - This is an application of the distance formula: the norm of the vector $\langle a_1, \ldots, a_n \rangle$ is just the length of the line segment joining the origin $(0, \ldots, 0)$ to the point $(a_1, \ldots, a_n)$.
    - Example: For $\mathbf{v} = \langle -1, 2, 2 \rangle$ and $\mathbf{w} = \langle 3, 0, -4, 5 \rangle$, we have $||\mathbf{v}|| = \sqrt{(-1)^2 + 2^2 + 2^2} = \boxed{3}$, and $||\mathbf{w}|| = \sqrt{3^2 + 0^2 + (-4)^2 + 5^2} = \boxed{\sqrt{50}}$.
    - If r is a scalar, we can see immediately from the definition that $||r\,\mathbf{v}|| = |r|\;||\mathbf{v}||$, since we can just factor $\sqrt{r^2} = |r|$ from each term under the square root.
    - Observe also that the dot product of a vector with itself is the square of the norm: $\mathbf{v} \cdot \mathbf{v} = ||\mathbf{v}||^2$.
- Using the dot product, we can give a formula for the angle between two vectors:
- Theorem (Dot Product): For vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ forming an angle θ between them, $\mathbf{v}_1 \cdot \mathbf{v}_2 = ||\mathbf{v}_1||\;||\mathbf{v}_2||\,\cos(\theta)$.
    - Proof: Consider the triangle formed by $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_2 - \mathbf{v}_1$: applying the Law of Cosines in this triangle yields
    $$||\mathbf{v}_2 - \mathbf{v}_1||^2 = ||\mathbf{v}_1||^2 + ||\mathbf{v}_2||^2 - 2\,||\mathbf{v}_1||\,||\mathbf{v}_2||\cos(\theta).$$
    Since the square of the norm is the dot product of a vector with itself, and the dot product is distributive, we can write
    $$\begin{aligned} ||\mathbf{v}_2 - \mathbf{v}_1||^2 &= (\mathbf{v}_2 - \mathbf{v}_1) \cdot (\mathbf{v}_2 - \mathbf{v}_1) \\ &= (\mathbf{v}_2 \cdot \mathbf{v}_2) - (\mathbf{v}_1 \cdot \mathbf{v}_2) - (\mathbf{v}_2 \cdot \mathbf{v}_1) + (\mathbf{v}_1 \cdot \mathbf{v}_1) \\ &= ||\mathbf{v}_2||^2 - 2(\mathbf{v}_1 \cdot \mathbf{v}_2) + ||\mathbf{v}_1||^2. \end{aligned}$$
    Now by comparing to the Law of Cosines expression, we must have $||\mathbf{v}_1||\,||\mathbf{v}_2||\cos(\theta) = \mathbf{v}_1 \cdot \mathbf{v}_2$, as claimed.
- Using the Dot Product Theorem, we can compute the angle between two vectors.
- Example: Compute the angle between the vectors $\mathbf{v} = \langle 2\sqrt{2}, 1, \sqrt{3} \rangle$ and $\mathbf{w} = \langle 0, \sqrt{3}, 1 \rangle$.
    - We compute $\mathbf{v} \cdot \mathbf{w} = (2\sqrt{2})(0) + (1)(\sqrt{3}) + (\sqrt{3})(1) = 2\sqrt{3}$, and $||\mathbf{v}|| = \sqrt{(2\sqrt{2})^2 + 1^2 + (\sqrt{3})^2} = 2\sqrt{3}$ and $||\mathbf{w}|| = \sqrt{(\sqrt{3})^2 + 0^2 + 1^2} = 2$.
    - Then by the Dot Product Theorem, the angle θ between the vectors satisfies $2\sqrt{3} = 2 \cdot 2\sqrt{3} \cdot \cos(\theta)$, meaning that $\theta = \cos^{-1}\left(\frac{1}{2}\right) = \boxed{\frac{\pi}{3}}$.
- Example: Compute the angle between the vectors $\mathbf{v} = \langle 0, 2, 1, 2 \rangle$ and $\mathbf{w} = \langle 3, 4, 0, -12 \rangle$.
    - We have $\mathbf{v} \cdot \mathbf{w} = 0 + 8 + 0 - 24 = -16$, $||\mathbf{v}|| = \sqrt{0^2 + 2^2 + 1^2 + 2^2} = 3$, and $||\mathbf{w}|| = \sqrt{3^2 + 4^2 + 0^2 + (-12)^2} = 13$.
    - Then by the Dot Product Theorem, the angle θ between the vectors satisfies $-16 = 3 \cdot 13 \cdot \cos(\theta)$, so $\theta = \boxed{\cos^{-1}\left(-\frac{16}{39}\right)}$.

2.2 The Formal Definition of a Vector Space

- The two operations of addition and scalar multiplication (and the various algebraic properties they satisfy) are the most fundamental properties of vectors in $\mathbb{R}^n$. We would like to investigate other collections of things which possess those same properties.
- Definition: A (real) vector space is a collection V of vectors together with two binary operations, addition of vectors ($+$) and scalar multiplication of a vector by a real number ($\cdot$), satisfying the following axioms:

  [A1] Addition is commutative: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ for any vectors $\mathbf{v}$ and $\mathbf{w}$.

  [A2] Addition is associative: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

  [A3] There exists a zero vector $\mathbf{0}$, with $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$ for any vector $\mathbf{v}$.

  [A4] Every vector $\mathbf{v}$ has an additive inverse $-\mathbf{v}$, with $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$.

  [M1] Scalar multiplication is consistent with regular multiplication: $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}$ for any scalars α, β and vector $\mathbf{v}$.

  [M2] Addition of scalars distributes: $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$ for any scalars α, β and vector $\mathbf{v}$.

  [M3] Addition of vectors distributes: $\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w}$ for any scalar α and vectors $\mathbf{v}$ and $\mathbf{w}$.

  [M4] The scalar 1 acts like the identity on vectors: $1 \cdot \mathbf{v} = \mathbf{v}$ for any vector $\mathbf{v}$.

- Remark: One may also consider vector spaces where the collection of scalars is something other than the real numbers: for example, there exists an equally important notion of a complex vector space, whose scalars are the complex numbers. (The axioms are the same, except we allow the scalars to be complex numbers.)
  - We will primarily work with real vector spaces, in which the scalars are the real numbers.
  - The most general notion of a vector space involves scalars from a field, which is a collection of numbers which possess addition and multiplication operations which are commutative, associative, and distributive, with an additive identity 0 and multiplicative identity 1, such that every element has an additive inverse and every nonzero element has a multiplicative inverse.
  - Aside from the real and complex numbers, another example of a field is the rational numbers ("fractions").
  - One can formulate an equally interesting theory of vector spaces over any field.
- Here are some examples of vector spaces:
- Example: The vectors in $\mathbb{R}^n$ are a vector space, for any $n > 0$. (This had better be true!)
  - For simplicity we will demonstrate all of the axioms for vectors in $\mathbb{R}^2$; there, the vectors are of the form $\langle x, y \rangle$ and scalar multiplication is defined as $\alpha \cdot \langle x, y \rangle = \langle \alpha x, \alpha y \rangle$.
  - [A1]: We have $\langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle = \langle x_1 + x_2, y_1 + y_2 \rangle = \langle x_2, y_2 \rangle + \langle x_1, y_1 \rangle$.
  - [A2]: We have $(\langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle) + \langle x_3, y_3 \rangle = \langle x_1 + x_2 + x_3, y_1 + y_2 + y_3 \rangle = \langle x_1, y_1 \rangle + (\langle x_2, y_2 \rangle + \langle x_3, y_3 \rangle)$.
  - [A3]: The zero vector is $\langle 0, 0 \rangle$, and clearly $\langle x, y \rangle + \langle 0, 0 \rangle = \langle x, y \rangle$.
  - [A4]: The additive inverse of $\langle x, y \rangle$ is $\langle -x, -y \rangle$, since $\langle x, y \rangle + \langle -x, -y \rangle = \langle 0, 0 \rangle$.
  - [M1]: We have $\alpha_1 \cdot (\alpha_2 \cdot \langle x, y \rangle) = \langle \alpha_1\alpha_2 x, \alpha_1\alpha_2 y \rangle = (\alpha_1\alpha_2) \cdot \langle x, y \rangle$.
  - [M2]: We have $(\alpha_1 + \alpha_2) \cdot \langle x, y \rangle = \langle (\alpha_1 + \alpha_2)x, (\alpha_1 + \alpha_2)y \rangle = \alpha_1 \cdot \langle x, y \rangle + \alpha_2 \cdot \langle x, y \rangle$.
  - [M3]: We have $\alpha \cdot (\langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle) = \langle \alpha(x_1 + x_2), \alpha(y_1 + y_2) \rangle = \alpha \cdot \langle x_1, y_1 \rangle + \alpha \cdot \langle x_2, y_2 \rangle$.
  - [M4]: Finally, we have $1 \cdot \langle x, y \rangle = \langle x, y \rangle$.
- Example: The set of $m \times n$ matrices for any m and any n, forms a vector space.
  - The various algebraic properties we know about matrix addition give [A1] and [A2] along with [M1], [M2], [M3], and [M4].
  - The "zero vector" in this vector space is the zero matrix (with all entries zero), and [A3] and [A4] follow easily.

- Note of course that in some cases we can also multiply matrices by other matrices. However, the requirements for being a vector space don't care that we can multiply matrices by other matrices! (All we need to be able to do is add them and multiply them by scalars.)

- Example: The complex numbers (the numbers of the form $a + bi$ for real a and b, and where $i^2 = -1$) are a vector space.
  - The axioms all follow from the standard properties of complex numbers. As might be expected, the "zero vector" is just the complex number $0 = 0 + 0i$.
  - Again, note that the complex numbers have "more structure" to them, because we can also multiply two complex numbers, and the multiplication is also commutative, associative, and distributive over addition. However, the requirements for being a vector space don't care that the complex numbers have these additional properties.

- Example: The collection of all real-valued functions on any part of the real line is a vector space, where we define the "sum" of two functions as $(f+g)(x) = f(x) + g(x)$ for every x, and "scalar multiplication" as $(\alpha \cdot f)(x) = \alpha\, f(x)$.
  - To illustrate: if $f(x) = x$ and $g(x) = x^2$, then $f + g$ is the function with $(f+g)(x) = x + x^2$, and $2f$ is the function with $(2f)(x) = 2x$.
  - The axioms follow from the properties of functions and real numbers. The "zero vector" in this space is the zero function; namely, the function z which has $z(x) = 0$ for every x.
  - For example (just to demonstrate a few of the axioms), for any value x in $[a, b]$ and any functions f and g, we have
    - [A1]: $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$.
    - [M2]: $\alpha \cdot (f+g)(x) = \alpha\, f(x) + \alpha\, g(x) = (\alpha f)(x) + (\alpha g)(x)$.
    - [M4]: $(1 \cdot f)(x) = f(x)$.

- Example: The "zero space" with a single element $\mathbf{0}$, with $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $\alpha \cdot \mathbf{0} = \mathbf{0}$ for every α, is a vector space.
  - All of the axioms in this case eventually boil down to $\mathbf{0} = \mathbf{0}$.
  - This space is rather boring: since it only contains one element, there's really not much to say about it.

- Purely for ease of notation, it will be useful to define subtraction:

- Definition: The difference of two vectors $\mathbf{v}, \mathbf{w}$ in a vector space V is defined to be $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$.
  - The difference has the fundamental property we would expect: by axioms [A2] and [A3], we can write $(\mathbf{v} - \mathbf{w}) + \mathbf{w} = (\mathbf{v} + (-\mathbf{w})) + \mathbf{w} = \mathbf{v} + ((-\mathbf{w}) + \mathbf{w}) = \mathbf{v} + \mathbf{0} = \mathbf{v}$.

- There are many simple algebraic properties that can be derived from the axioms which (therefore) hold in every vector space.

- Theorem (Basic Properties of Vector Spaces): In any vector space V, the following are true:
  1. Addition has a cancellation law: for any vector $\mathbf{v}$, if $\mathbf{a} + \mathbf{v} = \mathbf{b} + \mathbf{v}$ then $\mathbf{a} = \mathbf{b}$.
     - Proof: By [A1]-[A4] we have $(\mathbf{a} + \mathbf{v}) + (-\mathbf{v}) = \mathbf{a} + (\mathbf{v} + (-\mathbf{v})) = \mathbf{a} + \mathbf{0} = \mathbf{a}$.
     - Similarly we also have $(\mathbf{b} + \mathbf{v}) + (-\mathbf{v}) = \mathbf{b} + (\mathbf{v} + (-\mathbf{v})) = \mathbf{b} + \mathbf{0} = \mathbf{b}$.
     - Finally, since $\mathbf{a} + \mathbf{v} = \mathbf{b} + \mathbf{v}$ then $\mathbf{a} = (\mathbf{a} + \mathbf{v}) + (-\mathbf{v}) = (\mathbf{b} + \mathbf{v}) + (-\mathbf{v}) = \mathbf{b}$ so $\mathbf{a} = \mathbf{b}$.
  2. The zero vector is unique: if $\mathbf{a} + \mathbf{v} = \mathbf{v}$ for some vector $\mathbf{v}$, then $\mathbf{a} = \mathbf{0}$.
     - Proof: By [A3], $\mathbf{v} = \mathbf{0} + \mathbf{v}$, so we have $\mathbf{a} + \mathbf{v} = \mathbf{0} + \mathbf{v}$. Then by property (1) we conclude $\mathbf{a} = \mathbf{0}$.
  3. The additive inverse is unique: for any vector $\mathbf{v}$, if $\mathbf{a} + \mathbf{v} = \mathbf{0}$ then $\mathbf{a} = -\mathbf{v}$.
     - Proof: By [A4], $\mathbf{0} = (-\mathbf{v}) + \mathbf{v}$, so $\mathbf{a} + \mathbf{v} = (-\mathbf{v}) + \mathbf{v}$. Then by property (1) we conclude $\mathbf{a} = -\mathbf{v}$.
  4. The scalar 0 times any vector gives the zero vector: $0 \cdot \mathbf{v} = \mathbf{0}$ for any vector $\mathbf{v}$.

- Proof: By [M2] and [M4] we have $\mathbf{v} = 1 \cdot \mathbf{v} = (0+1) \cdot \mathbf{v} = 0 \cdot \mathbf{v} + 1 \cdot \mathbf{v} = 0 \cdot \mathbf{v} + \mathbf{v}$.
    - Thus, by [A3], we have $\mathbf{0} + \mathbf{v} = 0 \cdot \mathbf{v} + \mathbf{v}$ so by property (1) we conclude $\mathbf{0} = 0 \cdot \mathbf{v}$.

5. Any scalar times the zero vector is the zero vector: $\alpha \cdot \mathbf{0} = \mathbf{0}$ for any scalar α.
    - Proof: By [M1] and [M4] we have $\alpha \cdot \mathbf{0} = \alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0}$.
    - Thus, by [A3], we have $\mathbf{0} + \alpha \cdot \mathbf{0} = \alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0}$, so by property (1) we conclude $\mathbf{0} = \alpha \cdot \mathbf{0}$.
6. The scalar -1 times any vector gives the additive inverse: $(-1) \cdot \mathbf{v} = -\mathbf{v}$ for any vector $\mathbf{v}$.
    - Proof: By property (4) and [M2]-[M4] we have $\mathbf{v} + (-1) \cdot \mathbf{v} = 1 \cdot \mathbf{v} + (-1) \cdot \mathbf{v} = (1 + (-1)) \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{0}$.
    - But now by property (3), since $\mathbf{v} + (-1) \cdot \mathbf{v} = \mathbf{0}$, we see that $(-1) \cdot \mathbf{v} = -\mathbf{v}$.
7. The additive inverse of the additive inverse is the original vector: $-(-\mathbf{v}) = \mathbf{v}$ for any vector $\mathbf{v}$.
    - Proof: By property (5) twice and [M3]-[M4], $-(-\mathbf{v}) = (-1) \cdot (-\mathbf{v}) = (-1) \cdot [(-1) \cdot \mathbf{v}] = (-1)^2 \cdot \mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v}$.
8. The only scalar multiples equal to the zero vector are the trivial ones: if $\alpha \cdot \mathbf{v} = \mathbf{0}$, then either $\alpha = 0$ or $\mathbf{v} = \mathbf{0}$.
    - Proof: If $\alpha = 0$ then we are done. Otherwise, if $\alpha \neq 0$, then since α is a real number, it has a multiplicative inverse α^{-1}.
    - Then by property (5) and [M1], [M4], we have $\mathbf{0} = \alpha^{-1} \cdot \mathbf{0} = \alpha^{-1} \cdot (\alpha \cdot \mathbf{v}) = (\alpha^{-1}\alpha) \cdot \mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v}$.
9. The additive inverse is obtained by subtraction from the zero vector: $-\mathbf{v} = \mathbf{0} - \mathbf{v}$ for any vector $\mathbf{v}$.
    - Proof: By the definition of subtraction and [A3], $\mathbf{0} - \mathbf{v} = \mathbf{0} + (-\mathbf{v}) = -\mathbf{v}$.
10. Negation distributes over addition: $-(\mathbf{v} + \mathbf{w}) = (-\mathbf{v}) + (-\mathbf{w}) = -\mathbf{v} - \mathbf{w}$.
    - Proof: By property (6) and [M3], $-(\mathbf{v} + \mathbf{w}) = (-1) \cdot (\mathbf{v} + \mathbf{w}) = (-1) \cdot \mathbf{v} + (-1) \cdot \mathbf{w} = (-\mathbf{v}) + (-\mathbf{w})$.
    - Also, by the definition of subtraction, $-\mathbf{v} - \mathbf{w} = (-\mathbf{v}) + (-\mathbf{w})$. So all three quantities are equal.
11. Any sum of vectors can be associated or rearranged in any order without changing the sum.
    - The precise details of this argument are technical and we will omit them. However, this result allows us to freely rearrange sums of vectors.

- The results above are useful, and at the very least they suggest that the notation for vector spaces is sensible: for example, the scalar multiple $(-1) \cdot \mathbf{v}$ is in fact the same as the additive inverse $-\mathbf{v}$, as the notation very strongly suggests should be true. However, we do not seem to have gotten very far.
    - It might seem that the axioms we have imposed do not really impose much structure aside from rather simple properties like the ones listed above: after all, each individual axiom does not say very much on its own.
    - But in fact, we will show that the axioms taken collectively force V to have a very strong and regular structure. In particular, we will be able to describe all of the elements of any vector space in a precise and simple way.

2.3 Subspaces

- Definition: A subspace W of a vector space V is a subset of the vector space V which, under the same addition and scalar multiplication operations as V, is itself a vector space.
- Example: Show that the set of diagonal 2×2 matrices is a subspace of the vector space of all 2×2 matrices.
    - To do this directly from the definition, we need to verify that all of the vector space axioms hold for the matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ for some a, b.
    - First we need to check that the addition operation and scalar multiplication operations actually make sense: we see that $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a+c & 0 \\ 0 & b+d \end{bmatrix}$ is also a diagonal matrix, and $p \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} pa & 0 \\ 0 & pb \end{bmatrix}$ is a diagonal matrix too, so the sum and scalar multiplication operations are well-defined.

- Then we have to check the axioms, which is rather tedious. Here are some of the verifications:
- [A1] Addition is commutative: $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} + \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
- [A3] The zero element is the zero matrix, since $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
- [A4] The additive inverse of $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ is $\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ since $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
- [M1] Scalar multiplication is consistent with regular multiplication: $p \cdot q \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} pqa & 0 \\ 0 & pqb \end{bmatrix} = pq \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.

- It is very time-consuming to verify all of the axioms for a subspace, and much of the work seems to be redundant. It would be convenient if we could clean up the repetitive nature of the verifications:

- Theorem (Subspace Criterion): A subset W of a vector space V is a subspace of V if and only if W has the following three properties:

  [S1] W contains the zero vector of V.

  [S2] W is closed under addition: for any $\mathbf{w}_1$ and $\mathbf{w}_2$ in W, the vector $\mathbf{w}_1 + \mathbf{w}_2$ is also in W.

  [S3] W is closed under scalar multiplication: for any scalar α and $\mathbf{w}$ in W, the vector $\alpha \cdot \mathbf{w}$ is also in W.

  - Proof: Each of these conditions is necessary for W to be a subspace: the definition of binary operation requires [S2] and [S3] to hold, because when we add or scalar-multiply elements of W, we must obtain a result that is in W. For [S1], W must contain a zero vector $\mathbf{0}_W$, and then we can write $\mathbf{0}_V = \mathbf{0}_V + \mathbf{0}_W = \mathbf{0}_W$, so W contains the zero vector of V.
  - Now suppose each of [S1]-[S3] holds for W. Since all of the operations are therefore defined, axioms [A1]-[A2] and [M1]-[M4] will hold in W because they hold in V. Axiom [A3] for W follows from [S1] since $\mathbf{0}_W = \mathbf{0}_V$. Finally, for [A4], for any vector $\mathbf{w}$ in W, by our basic properties we know that $(-1) \cdot \mathbf{w} = -\mathbf{w}$, so since $(-1) \cdot \mathbf{w}$ is in W by [S3], we see that $-\mathbf{w}$ is in W.

- Any vector space automatically has two subspaces: the entire space V, and the "trivial" subspace consisting only of the zero vector.
  - These examples are rather uninteresting, since we already know V is a vector space, and the subspace consisting only of the zero vector has very little structure.

- Very often, if we want to check that something is a vector space, it is often much easier to verify that it is a subspace of something else we already know is a vector space, which is easily done using the subspace criterion. In order to show that a subset is *not* a subspace, it is sufficient to find a single example in which any one of the three criteria fails.

- Example: Determine whether the set of vectors of the form $\langle t, t, t \rangle$ forms a subspace of $\mathbb{R}^3$.
  - We check the parts of the subspace criterion.
  - [S1]: The zero vector is of this form: take $t = 0$.
  - [S2]: We have $\langle t_1, t_1, t_1 \rangle + \langle t_2, t_2, t_2 \rangle = \langle t_1 + t_2, t_1 + t_2, t_1 + t_2 \rangle$, which is again of the same form if we take $t = t_1 + t_2$.
  - [S3]: We have $\alpha \cdot \langle t_1, t_1, t_1 \rangle = \langle \alpha t_1, \alpha t_1, \alpha t_1 \rangle$, which is again of the same form if we take $t = \alpha t_1$.
  - All three parts are satisfied, so this subset $\boxed{\text{is a subspace}}$.

- Example: Determine whether the set of vectors of the form $\langle t, t^2 \rangle$ forms a subspace of $\mathbb{R}^2$.
  - We try checking the parts of the subspace criterion.
  - [S1]: The zero vector is of this form: take $t = 0$.

- [S2]: For this criterion we try to write $\langle t_1, t_1^2 \rangle + \langle t_2, t_2^2 \rangle = \langle t_1 + t_2, t_1^2 + t_2^2 \rangle$, but this does not have the correct form, because in general $t_1^2 + t_2^2 \neq (t_1 + t_2)^2$. (These quantities are only equal if $2t_1t_2 = 0$.)
- From here we can find a specific counterexample: the vectors $\langle 1, 1 \rangle$ and $\langle 2, 4 \rangle$ are in the subset, but their sum $\langle 3, 5 \rangle$ is not. Thus, this subset is $\boxed{\text{not a subspace}}$.
- Note that all we actually needed to do here was find a single counterexample, of which there are many. Had we noticed earlier that $\langle 1, 1 \rangle$ and $\langle 2, 4 \rangle$ were in the subset but their sum $\langle 3, 5 \rangle$ was not, that would have been sufficient to conclude that the given set was not a subspace.

- Example: Determine whether the set of vectors of the form $\langle s, t, 0 \rangle$ forms a subspace of $\mathbb{R}^3$.
  - We check the parts of the subspace criterion.
  - [S1]: The zero vector is of this form: take $s = t = 0$.
  - [S2]: We have $\langle s_1, t_1, 0 \rangle + \langle s_2, t_2, 0 \rangle = \langle s_1 + s_2, t_1 + t_2, 0 \rangle$, which is again of the same form, if we take $s = s_1 + s_2$ and $t = t_1 + t_2$.
  - [S3]: We have $\alpha \cdot \langle s_1, t_1, 0 \rangle = \langle \alpha s_1, \alpha t_1, 0 \rangle$, which is again of the same form, if we take $s = \alpha s_1$ and $t = \alpha t_1$.
  - All three parts are satisfied, so this subset $\boxed{\text{is a subspace}}$.

- Example: Determine whether the set of vectors of the form $\langle x, y, z \rangle$ where $2x - y + z = 0$ forms a subspace of $\mathbb{R}^3$.
  - [S1]: The zero vector is of this form, since $2(0) - 0 + 0 = 0$.
  - [S2]: If $\langle x_1, y_1, z_1 \rangle$ and $\langle x_2, y_2, z_2 \rangle$ have $2x_1 - y_1 + z_1 = 0$ and $2x_2 - y_2 + z_2 = 0$ then adding the equations shows that the sum $\langle x_1 + x_2, y_1 + y_2, z_1 + z_2 \rangle$ also lies in the space.
  - [S3]: If $\langle x_1, y_1, z_1 \rangle$ has $2x_1 - y_1 + z_1 = 0$ then scaling the equation by α shows that $\langle \alpha x_1, \alpha x_2, \alpha x_3 \rangle$ also lies in the space.
  - All three parts are satisfied, so this subset $\boxed{\text{is a subspace}}$.

- Example: Determine whether the set of vectors of the form $\langle x, y, z \rangle$ where $x, y, z \geq 0$ forms a subspace of $\mathbb{R}^3$.
  - [S1]: The zero vector is of this form: take $t = 0$.
  - [S2]: If $\langle x_1, y_1, z_1 \rangle$ and $\langle x_2, y_2, z_2 \rangle$ have $x_1, y_1, z_1 \geq 0$ and $x_2, y_2, z_2 \geq 0$, then $x_1 + x_2 \geq 0$, $y_1 + y_2 \geq 0$, and $z_1 + z_2 \geq 0$, so $\langle x_1 + x_2, y_1 + y_2, z_1 + z_2 \rangle$ also lies in the space.
  - [S3]: If $\langle x, y, z \rangle$ has $x, y, z \geq 0$, then it is not necessarily true that $\alpha x, \alpha y, \alpha z \geq 0$: specifically, this is not true when $\alpha = -1$.
  - From here we can find a specific counterexample: the vector $\langle 1, 1, 1 \rangle$ is in the subset, but the scalar multiple $-1 \cdot \langle 1, 1, 1 \rangle = \langle -1, -1, -1 \rangle$ is not in the subset. Thus, this subset is $\boxed{\text{not a subspace}}$.

- Example: Determine whether the set of 2×2 matrices of trace zero is a subspace of the space of all 2×2 matrices.
  - [S1]: The zero matrix has trace zero.
  - [S2]: Since $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$, we see that if A and B have trace zero then so does $A + B$.
  - [S3]: Since $\text{tr}(\alpha A) = \alpha \text{tr}(A)$, we see that if A has trace zero then so does αA.
  - All three parts are satisfied, so this subset $\boxed{\text{is a subspace}}$.

- Example: Determine whether the set of 2×2 matrices of determinant zero is a subspace of the space of all 2×2 matrices.
  - [S1]: The zero matrix has determinant zero.
  - [S3]: Since $\det(\alpha A) = \alpha^2 \det(A)$ when A is a 2×2 matrix, we see that if A has determinant zero then so does αA.

- [S2]: If A and B have determinant zero, then there does not appear to be a nice way to compute the determinant of $A+B$ in general.
- We can in fact find a counterexample: if $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ then $\det(A) = \det(B) = 0$, but $\det(A+B) = 1$. Thus, this subset is $\boxed{\text{not a subspace}}$.

- Here are a few more examples of subspaces of vector spaces which will be of interest to us:
- Example: The collection of solution vectors $\langle x_1, \cdots, x_n \rangle$ to any homogeneous system of linear equations forms a subspace of $\mathbb{R}^n$.
  - It is possible to check this directly by working with equations. But it is much easier to use matrices: write the system in matrix form, as $A\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a solution vector.
  - [S1]: We have $A\mathbf{0} = \mathbf{0}$, by the properties of the zero vector.
  - [S2]: If $\mathbf{x}$ and $\mathbf{y}$ are two solutions, the properties of matrix arithmetic imply $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$ so that $\mathbf{x}+\mathbf{y}$ is also a solution.
  - [S3]: If α is a scalar and $\mathbf{x}$ is a solution, then $A(\alpha \cdot \mathbf{x}) = \alpha \cdot (A\mathbf{x}) = \alpha \cdot \mathbf{0} = \mathbf{0}$, so that $\alpha \cdot \mathbf{x}$ is also a solution.
- Example: The collection of continuous functions on $[a, b]$ is a subspace of the space of all functions on $[a, b]$.
  - [S1]: The zero function is continuous.
  - [S2]: The sum of two continuous functions is continuous, from basic calculus.
  - [S3]: The product of continuous functions is continuous, so in particular a constant times a continuous function is continuous.
- Example: The collection of n-times differentiable functions on $[a, b]$ is a subspace of the space of continuous functions on $[a, b]$, for any positive integer n.
  - The zero function is differentiable, as are the sum and product of any two functions which are differentiable n times.
- Example: The collection of all polynomials is a vector space.
  - Observe that polynomials are functions on the entire real line. Therefore, it is sufficient to verify the subspace criteria.
  - The zero function is a polynomial, as is the sum of two polynomials, and any scalar multiple of a polynomial.
- Example: The collection of all polynomials of degree at most n, denoted $P_n(\mathbb{R})$, is a vector space.
  - From above, we know that the collection of all polynomials (of any degree) is a vector space. So we only need to verify the subspace criteria.
  - The zero function has degree at most n, as does the sum of two polynomials of degree at most n, and any scalar multiple of a polynomial of a polynomial of degree at most n.
- Example: The collection of solutions to the (homogeneous, linear) differential equation $y'' + 6y' + 5y = 0$ form a vector space.
  - We show this by verifying that the solutions form a subspace of the space of all functions.
  - [S1]: The zero function is a solution.
  - [S2]: If y_1 and y_2 are solutions, then $y_1'' + 6y_1' + 5y_1 = 0$ and $y_2'' + 6y_2' + 5y_2 = 0$, so adding and using properties of derivatives shows that $(y_1+y_2)'' + 6(y_1+y_2)' + 5(y_1+y_2) = 0$, so y_1+y_2 is also a solution.

- [S3]: If α is a scalar and y_1 is a solution, then scaling $y_1'' + 6y_1' + 5y_1 = 0$ by α and using properties of derivatives shows that $(\alpha y_1)'' + 6(\alpha y_1)' + 5(\alpha y_1) = 0$, so αy_1 is also a solution.
- Note that we did not need to know how to solve the differential equation to answer the question.
- For completeness, the solutions are $y = Ae^{-x} + Be^{-5x}$ for any constants A and B. (From this description, if we wanted to, we could directly verify that such functions form a vector space.)

- This last example should help explain how the study of vector spaces and linear algebra is useful for the study of differential equations: namely, because the solutions to the given homogeneous linear differential equation form a vector space.
  - It is true more generally that the solutions to an arbitrary homogeneous linear differential equation $y^{(n)} + P_n(x) \cdot y^{(n-1)} + \cdots + P_1(x) \cdot y = 0$ will form a vector space.
  - Most of the time we cannot explicitly write down the solutions to this differential equation; nevertheless, if we can understand the structure of a general vector space, we can still say something about what the solutions look like.
  - The reason this is a useful idea is that, once we know some more facts about vector spaces, they will automatically apply to the set of solutions to a homogeneous linear differential equation. Thus, we will not need to "reinvent the wheel" (so to speak) by proving these properties separately, because they are automatic from the vector space structure of the solution set.

- One thing we would like to know, now that we have the definition of a vector space and a subspace, is what else we can say about elements of a vector space: that is, we would like to know what kind of structure the elements of a vector space have.
  - In some of the earlier examples we saw that, in $\mathbb{R}^n$ and a few other vector spaces, subspaces could all be written down in terms of one or more parameters. We will develop this idea in the next few sections.

2.4 Linear Combinations and Span

- Definition: Given a set $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ of vectors in a vector space V, we say a vector $\mathbf{w}$ in V is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ if there exist scalars $a_1, \cdots, a_n$ such that $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$.
  - Example: In $\mathbb{R}^2$, the vector $\langle 1, 1\rangle$ is a linear combination of $\langle 1, 0\rangle$ and $\langle 0, 1\rangle$, because $\langle 1, 1\rangle = 1 \cdot \langle 1, 0\rangle + 1 \cdot \langle 0, 1\rangle$.
  - Example: In $\mathbb{R}^4$, the vector $\langle 4, 0, 5, 9\rangle$ is a linear combination of $\langle 1, 0, 0, 1\rangle$, $\langle 0, 1, 0, 0\rangle$, and $\langle 1, 1, 1, 2\rangle$, because $\langle 4, 0, 5, 9\rangle = 1 \cdot \langle 1, -1, 2, 3\rangle - 2 \cdot \langle 0, 1, 0, 0\rangle + 3 \cdot \langle 1, 1, 1, 2\rangle$.
  - Non-Example: In $\mathbb{R}^3$, the vector $\langle 0, 0, 1\rangle$ is not a linear combination of $\langle 1, 1, 0\rangle$ and $\langle 0, 1, 1\rangle$ because there exist no scalars a_1 and a_2 for which $a_1 \cdot \langle 1, 1, 0\rangle + a_2 \cdot \langle 0, 1, 1\rangle = \langle 0, 0, 1\rangle$: this would require a common solution to the three equations $a_1 = 0$, $a_1 + a_2 = 0$, and $a_2 = 1$, and this system has no solution.

- Definition: We define the span of a collection of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in V, denoted $\mathrm{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$, to be the set W of all vectors which are linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Explicitly, the span is the set of vectors of the form $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$, for some scalars $a_1, \cdots, a_n$.
  - For technical reasons, we define the span of the empty set to be the zero vector.

- Example: The span of the vectors $\langle 1, 0, 0\rangle$ and $\langle 0, 1, 0\rangle$ in $\mathbb{R}^3$ is the set of vectors of the form $a \cdot \langle 1, 0, 0\rangle + b \cdot \langle 0, 1, 0\rangle = \langle a, b, 0\rangle$.
  - Equivalently, the span of these vectors is the set of vectors whose z-coordinate is zero, which (geometrically) forms the plane $z = 0$.

- Example: Determine whether the vectors $\langle 2, 3, 3\rangle$ and $\langle 4, -1, 3\rangle$ are in $\mathrm{span}(\mathbf{v}, \mathbf{w})$, where $\mathbf{v} = \langle 1, -1, 2\rangle$ and $\mathbf{w} = \langle 2, 1, -1\rangle$.
  - For $\langle 2, 3, 3\rangle$ we must determine whether it is possible to write $\langle 2, 3, 3\rangle = a \cdot \langle 1, -1, 2\rangle + b \cdot \langle 2, 1, -1\rangle$ for some a and b.

- Equivalently, we want to solve the system $2 = a + 2b$, $3 = -a + b$, $3 = 2a - b$.
- Row-reducing the associated coefficient matrix gives

$$\left[\begin{array}{cc|c} 1 & 2 & 2 \\ -1 & 1 & 3 \\ 2 & -1 & 3 \end{array}\right] \xrightarrow[R_3-2R_1]{R_2+R_1} \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & 3 & 5 \\ 0 & -5 & -3 \end{array}\right] \xrightarrow{R_3+\frac{5}{3}R_1} \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & 3 & 5 \\ 0 & 0 & 16/3 \end{array}\right]$$

  and we obtain a contradiction. Thus, $\boxed{\langle 2,3,3\rangle \text{ is not in the span}}$.
- Similarly, for $\langle 4,-1,3\rangle$ we want to solve $\langle 4,-1,3\rangle = c\cdot\langle 1,-1,2\rangle + d\cdot\langle 2,1,-1\rangle$.
- Row-reducing the associated coefficient matrix gives

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

  from which we can easily obtain the solution $d = 1$, $c = 2$.
- Since $\langle 4,-1,3\rangle = 2\cdot\langle 1,-1,2\rangle + 1\cdot\langle 2,1,-1\rangle$ we see that $\boxed{\langle 4,-1,3\rangle \text{ is in the span}}$.

- Here are some basic properties of the span:
- Proposition (Span is a Subspace): For any set S of vectors in V, the set $\text{span}(S)$ is a subspace of V.
  - Proof: We check the subspace criterion. If S is empty, then by definition $\text{span}(S) = \{\mathbf{0}\}$ and $\{\mathbf{0}\}$ is a subspace of V.
  - Now assume S is not empty. Let $\mathbf{v}$ be any vector in S: then $0\cdot\mathbf{v} = \mathbf{0}$ is in $\text{span}(S)$.
  - The span is closed under addition because we can write the sum of any two linear combinations as another linear combination: $(a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n) + (b_1\cdot\mathbf{v}_1 + \cdots + b_n\cdot\mathbf{v}_n) = (a_1+b_1)\cdot\mathbf{v}_1 + \cdots + (a_n+b_n)\cdot\mathbf{v}_n$.
  - Finally, we can write any scalar multiple of a linear combination as a linear combination: $\alpha\cdot(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = (\alpha a_1)\cdot\mathbf{v}_1 + \cdots + (\alpha a_n)\cdot\mathbf{v}_n$.
- Proposition (Minimality of Span): For any vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$ in V, if W is any subspace of V that contains $\mathbf{v}_1,\dots,\mathbf{v}_n$, then W contains $\text{span}(\mathbf{v}_1,\dots,\mathbf{v}_n)$. In other words, the span is the smallest subspace containing the vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$.
  - Proof: Consider any element $\mathbf{w}$ in $\text{span}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n)$: by definition, we can write $\mathbf{w} = a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n$ for some scalars $a_1,\cdots,a_n$.
  - Because W is a subspace, it is closed under scalar multiplication, so each of $a_1\cdot\mathbf{v}_1$, $\cdots$, $a_n\cdot\mathbf{v}_n$ lies in W.
  - Furthermore, also because W is a subspace, it is closed under addition. Thus, the sum $a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n$ lies in W.
  - Thus, every element of the span lies in W, as claimed.
- Corollary: If S and T are two sets of vectors in V with $S \subseteq T$, then $\text{span}(S)$ is a subspace of $\text{span}(T)$.
  - Proof: Since the span is always a subspace, we know that $\text{span}(T)$ is a subspace of V containing all the vectors in S. By the previous proposition, $\text{span}(T)$ therefore contains every linear combination of vectors from S, which is to say, $\text{span}(T)$ contains $\text{span}(S)$.
- Sets whose span is the entire space have a special name:
- Definition: Given a set S of vectors in a vector space V, if $\text{span}(S) = V$ then we say that S is a spanning set (or generating set) for V.
  - There are a number of different phrases we use for this idea: we also say that the vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$ span or generate the vector space V.
  - Spanning sets are very useful because they allow us to describe every vector in V in terms of the vectors in S.

* Explicitly, every vector in V is a linear combination of the vectors in S, which is to say, every vector $\mathbf{w}$ in V can be written in the form $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$ for some scalars $a_1, \ldots, a_n$ and some vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S.

- Example: Show that the vectors $\langle 1,0,0 \rangle$, $\langle 0,1,0 \rangle$, and $\langle 0,0,1 \rangle$ span $\mathbb{R}^3$.
  * For any vector $\langle a,b,c \rangle$, we can write $\langle a,b,c \rangle = a \cdot \langle 1,0,0 \rangle + b \cdot \langle 0,1,0 \rangle + c \cdot \langle 0,0,1 \rangle$, so it is a linear combination of the three given vectors.

- Example: Show that the matrices $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ span the vector space of 2×2 matrices of trace zero.
  * Recall that we showed earlier that the space of matrices of trace zero is a vector space (since it is a subspace of the vector space of all 2×2 matrices).
  * A 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has trace zero when $a + d = 0$, or equivalently when $d = -a$.
  * So any matrix of trace zero has the form $\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.
  * Since any matrix of trace zero is therefore a linear combination of the matrices $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, we conclude that they are a spanning set.

- Example: Determine whether the vectors $\langle 1,2 \rangle$, $\langle 2,4 \rangle$, $\langle 3,1 \rangle$ span $\mathbb{R}^2$.
  * For any vector $\langle p,q \rangle$, we want to determine whether there exist some scalars a,b,c such that $\langle p,q \rangle = a \cdot \langle 1,2 \rangle + b \cdot \langle 2,4 \rangle + c \cdot \langle 3,1 \rangle$.
  * Equivalently, we want to check whether the system $p = a + 2b + 3c$, $q = 2a + 4b + c$ has solutions for any p,q.
  * Row-reducing the associated coefficient matrix gives

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

    and since this system is non-contradictory, there is always a solution: indeed, there are infinitely many. (One solution is $c = \frac{2}{5}p - \frac{1}{5}q$, $b = 0$, $a = -\frac{1}{5}p + \frac{3}{5}q$.)
  * Since there is always a solution for any p,q, we conclude that these vectors $\boxed{\text{do span } \mathbb{R}^2}$.

- Example: Determine whether the vectors $\langle 1,-1,3 \rangle$, $\langle 2,2,-1 \rangle$, $\langle 3,4,7 \rangle$ span $\mathbb{R}^3$.
  * For any vector $\langle p,q,r \rangle$, we want to determine whether there exist some scalars a,b,c such that $\langle p,q,r \rangle = a \cdot \langle 1,-1,3 \rangle + b \cdot \langle 2,2,-1 \rangle + c \cdot \langle 3,1,2 \rangle$.
  * Row-reducing the associated coefficient matrix gives

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

  * Now, if $r + 2q - p \neq 0$, the final column will have a pivot and the system will be contradictory. This can certainly occur: for example, we could take $r = 1$ and $p = q = 0$.
  * Since there is no way to write an arbitrary vector in $\mathbb{R}^3$ as a linear combination of the given vectors, we conclude that these vectors $\boxed{\text{do not span } \mathbb{R}^3}$.

- We can generalize the idea in the above examples to give a method for determining whether a collection of vectors in $\mathbb{R}^n$ will span $\mathbb{R}^n$.

- Theorem (Spanning Sets in $\mathbb{R}^n$): A collection of k vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ will span $\mathbb{R}^n$ if and only if, for every vector $\mathbf{b}$, there is at least one vector $\mathbf{x}$ satisfying the matrix equation $M\mathbf{x} = \mathbf{b}$, where M is the matrix whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$. Such a solution $\mathbf{x}$ exists for any $\mathbf{b}$ if and only if M has rank n: that is, when a row-echelon form of M has n pivotal columns.
    - Proof: Write each $\mathbf{v}_i = \begin{bmatrix} m_{1,i} \\ \vdots \\ m_{n,i} \end{bmatrix}$ as a column matrix.
    - Then $a_1 \cdot \mathbf{v}_1 + \cdots + a_k \cdot \mathbf{v}_k = \begin{bmatrix} m_{1,1} \\ \vdots \\ m_{n,1} \end{bmatrix} a_1 + \cdots + \begin{bmatrix} m_{1,k} \\ \vdots \\ m_{n,k} \end{bmatrix} a_k = M \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$, where $M = \begin{bmatrix} m_{1,1} & \cdots & m_{1,k} \\ \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,k} \end{bmatrix}$ is the matrix whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.
    - So the statement that, for any $\mathbf{b}$, there exist scalars $a_1, \dots, a_k$ such that $a_1 \cdot \mathbf{v}_1 + \cdots + a_k \cdot \mathbf{v}_k = \mathbf{b}$ is equivalent to the statement that there is a solution $\mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$ to the matrix equation $M\mathbf{x} = \mathbf{b}$.
    - For the second part of the theorem, consider the matrix equation $M\mathbf{x} = \mathbf{b}$, and perform row operations to put M in row-echelon form.
    - By our theorems on systems of linear equations, this system will have at least one solution precisely when there is no pivot in the augmented column of coefficients.
    - Since $\mathbf{b}$ can be chosen arbitrarily, so can the column of constants in the augmented matrix once we put it in row-echelon form.
    - Since the augmented matrix has at most n pivots (since it has n rows), the only way we can prevent having a pivot in the column of constants is to have all the pivots in the matrix M itself: thus, M must have n pivots. From the definition of rank, this is equivalent to saying M has rank n.

- Example: Determine whether the vectors $\langle 1, 0, 3, 2 \rangle$, $\langle 2, 2, -1, 1 \rangle$, $\langle 3, 4, 3, 2 \rangle$, $\langle -1, 2, 6, -1 \rangle$ span $\mathbb{R}^4$.
    - By the theorem above, we simply need to row-reduce the matrix whose columns are the given vectors:

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 2 & 4 & 2 \\ 3 & -1 & 3 & 6 \\ 2 & 1 & 2 & -1 \end{bmatrix} \underset{R_4 - 2R_1}{\xrightarrow{R_3 - 3R_1}} \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 2 & 4 & 2 \\ 0 & -7 & -6 & 9 \\ 0 & -3 & -4 & 1 \end{bmatrix} \longrightarrow \cdots \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

    - From this reduced row-echelon form, we see that the matrix has rank 3, so the vectors $\boxed{\text{do not span } \mathbb{R}^4}$.

- For other vector spaces, we can apply the same sorts of ideas to determine whether a particular set of vectors is a spanning set.

- Example: Determine whether the polynomials 1, $1 + x^2$, x^4, $1 + x^2 + x^4$ span the space W of polynomials of degree at most 4 satisfying $p(x) = p(-x)$.
    - It is straightforward to verify that this set W is a subspace of the vector space of polynomials. (We omit this verification.)
    - A polynomial of degree at most 4 has the form $p(x) = a + bx + cx^2 + dx^3 + ex^4$, and having $p(x) = p(-x)$ requires $a - bx + cx^2 - dx^3 + ex^4 = a + bx + cx^2 + dx^3 + ex^4$, or equivalently $b = d = 0$.
    - Thus, the desired polynomials are those of the form $p(x) = a + cx^2 + ex^4$ for arbitrary coefficients a, c, and e.
    - Since we can write $a + cx^2 + ex^4 = (a - c) \cdot 1 + c \cdot (1 + x^2) + e \cdot x^4 + 0 \cdot (1 + x^2 + x^4)$, the given polynomials $\boxed{\text{do span } W}$.
    - Note that we could also have written $a + cx^2 + ex^4 = (a - c) \cdot 1 + (c - e) \cdot (1 + x^2) + 0 \cdot x^4 + e \cdot (1 + x^2 + x^4)$, so the polynomials in W can be written as a linear combination of the vectors in the spanning set in more than one way. (In fact, they can be written as a linear combination in infinitely many ways.)
    - This example underlines another important point: if $\text{span}(S) = V$, it is possible that any given vector in V can be written as a linear combination of vectors in S in many different ways.

2.5 Linear Independence and Linear Dependence

- Definition: We say a finite set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent if $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$ implies $a_1 = \cdots = a_n = 0$. Otherwise, we say the collection is linearly dependent. (The empty set of vectors is by definition linearly independent.)
  - In other words, $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent precisely when the only way to form the zero vector as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$ is to have all the scalars equal to zero (the "trivial" linear combination). If there is a nontrivial linear combination giving the zero vector, then $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly dependent.
  - Note: For an infinite set of vectors, we say it is linearly independent if every finite subset is linearly independent, per the definition above. Otherwise, if some finite subset displays a dependence, we say the infinite set is dependent.

2.5.1 Examples and Basic Properties

- Example: The matrices $\begin{bmatrix} 2 & 3 \\ 2 & -4 \end{bmatrix}$, $\begin{bmatrix} -1 & -1 \\ -1 & 2 \end{bmatrix}$, and $\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$ are linearly dependent, because $3 \cdot \begin{bmatrix} 2 & 3 \\ 2 & -4 \end{bmatrix} + 6 \cdot \begin{bmatrix} -1 & -1 \\ -1 & 2 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
- Example: Determine whether the vectors $\langle 1, 1, 0 \rangle$, $\langle 0, 2, 1 \rangle$ in $\mathbb{R}^3$ are linearly dependent or linearly independent.
  - Suppose that we had scalars a and b with $a \cdot \langle 1, 1, 0 \rangle + b \cdot \langle 0, 2, 1 \rangle = \langle 0, 0, 0 \rangle$.
  - Comparing the two sides requires $a = 0$, $a + 2b = 0$, $b = 0$, which has only the solution $a = b = 0$.
  - Thus, by definition, these vectors are $\boxed{\text{linearly independent}}$.
- Example: Determine whether the vectors $\langle 1, 1, 0 \rangle$, $\langle 2, 2, 0 \rangle$ in $\mathbb{R}^3$ are linearly dependent or linearly independent.
  - Suppose that we had scalars a and b with $a \cdot \langle 1, 1, 0 \rangle + b \cdot \langle 2, 2, 0 \rangle = \langle 0, 0, 0 \rangle$.
  - Comparing the two sides requires $a + 2b = 0$, $a + 2b = 0$, $0 = 0$, which has (for example) the nontrivial solution $a = 1$, $b = -2$.
  - Thus, we see that we can write $2 \cdot \langle 1, 1, 0 \rangle + (-1) \cdot \langle 2, 2, 0 \rangle = \langle 0, 0, 0 \rangle$, and this is a nontrivial linear combination giving the zero vector meaning that these vectors are $\boxed{\text{linearly dependent}}$.
- Example: Determine whether the polynomials $1 + x^2$, $2 - x + x^2$, and $1 + x + 2x^2$ are linearly dependent or linearly independent.
  - Suppose that we had scalars a, b, and c with $a(1 + x^2) + b(2 - x + x^2) + c(1 + x + 2x^2) = 0$.
  - Equivalently, this says $(a + 2b + c) + (-b + c)x + (a + b + 2c)x^2 = 0$. This will be true precisely when each coefficient is zero, which requires $a + 2b + c = 0$, $-b + c = 0$, and $a + b + 2c = 0$.
  - Solving this system reveals a nonzero solution with $a = -3$ and $b = c = 1$: thus, the polynomials are $\boxed{\text{linearly dependent}}$. Explicitly, $-3(1 + x^2) + 1(2 - x + x^2) + 1(1 + x + 2x^2) = 0$.
- Here are a few basic properties of linear dependence and independence that follow from the definition:
  - Any set containing the zero vector is linearly dependent. (Choose zero coefficients for the other vectors, and a nonzero coefficient for the zero vector.)
  - Any set containing a linearly dependent set is linearly dependent. (Any dependence in the smaller set gives a dependence in the larger set.)
  - Any subset of a linearly independent set is linearly independent. (Any dependence in the smaller set would also give a dependence in the larger set.)
  - Any set containing a single nonzero vector is linearly independent. (If $a \neq 0$ and $a \cdot \mathbf{v} = \mathbf{0}$, then scalar-multiplying by $1/a$ yields $\mathbf{v} = \mathbf{0}$. Thus, no nonzero multiple of a nonzero vector can be the zero vector.)

- The case of a set with two vectors is nearly as simple:
- Proposition: In any vector space V, the two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if one is a scalar multiple of the other, and they are linearly independent otherwise.
  - Proof: If $\mathbf{v}_1 = \alpha \cdot \mathbf{v}_2$ then we can write $1 \cdot \mathbf{v}_1 + (-\alpha) \cdot \mathbf{v}_2 = \mathbf{0}$, and similarly if $\mathbf{v}_2 = \alpha \cdot \mathbf{v}_1$ then we can write $(-\alpha) \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 = \mathbf{0}$. In either case the vectors are linearly dependent.
  - If the vectors are dependent, then suppose $a \cdot \mathbf{v}_1 + b \cdot \mathbf{v}_2 = \mathbf{0}$ where a, b are not both zero. If $a \neq 0$ then we can write $\mathbf{v}_1 = (-b/a) \cdot \mathbf{v}_2$, and if $b \neq 0$ then we can write $\mathbf{v}_2 = (-a/b) \cdot \mathbf{v}_1$. At least one of these cases must occur, so one of the vectors is a scalar multiple of the other as claimed.
- It is more a delicate problem to determine whether a larger set of vectors is linearly independent. Typically, answering this question will reduce to determining whether a set of linear equations has a solution.
- Example: Determine whether the vectors $\langle 1, 0, 2, 2\rangle$, $\langle 2, -2, 3, 0\rangle$, $\langle 0, 3, 1, 3\rangle$, and $\langle 0, 4, 1, 2\rangle$ in $\mathbb{R}^4$ are linearly dependent or linearly independent.
  - Suppose that we had scalars a,b,c,d with $a \cdot \langle 1, 0, 2, 2\rangle + b \cdot \langle 2, -2, 3, 0\rangle + c \cdot \langle 0, 3, 1, 3\rangle + d \cdot \langle 0, 4, 1, 2\rangle = \langle 0, 0, 0, 0\rangle$.
  - This is equivalent to saying $a + 2b = 0$, $-2b + 3c + 4d = 0$, $2a + 3b + c + d = 0$, and $2a + 3c + 2d = 0$.
  - To search for solutions we can convert this system into matrix form and then row-reduce it:

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 0 \\ 0 & -2 & 3 & 4 & 0 \\ 2 & 3 & 1 & 1 & 0 \\ 2 & 0 & 3 & 2 & 0 \end{array}\right] \xrightarrow[R_4-2R_1]{R_3-2R_1} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 0 \\ 0 & -2 & 3 & 4 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & -4 & 3 & 2 & 0 \end{array}\right] \longrightarrow \cdots \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

    from which we can obtain a nonzero solution $d = 1$, $c = -2$, $b = -1$, $a = 2$.
  - So we see $2 \cdot \langle 1, 0, 2, 2\rangle + (-1) \cdot \langle 2, -2, 0, 3\rangle + (-2) \cdot \langle 0, 3, 3, 1\rangle + 1 \cdot \langle 0, 4, 2, 1\rangle = \langle 0, 0, 0, 0\rangle$, and this is a nontrivial linear combination giving the zero vector meaning that these vectors are $\boxed{\text{linearly dependent}}$.
- We can generalize the idea in the above example to give a method for determining whether a collection of vectors in $\mathbb{R}^n$ is linearly independent:
- Theorem (Dependence of Vectors in $\mathbb{R}^n$): A collection of k vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is linearly dependent if and only if there is a nonzero vector $\mathbf{x}$ satisfying the matrix equation $M\mathbf{x} = \mathbf{0}$, where M is the matrix whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.
  - Proof: Write each $\mathbf{v}_i = \begin{bmatrix} m_{1,i} \\ \vdots \\ m_{n,i} \end{bmatrix}$ as a column matrix.
  - Then $a_1 \cdot \mathbf{v}_1 + \cdots + a_k \cdot \mathbf{v}_k = \begin{bmatrix} m_{1,1} \\ \vdots \\ m_{n,1} \end{bmatrix} a_1 + \cdots + \begin{bmatrix} m_{1,k} \\ \vdots \\ m_{n,k} \end{bmatrix} a_k = M \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$, where $M = \begin{bmatrix} m_{1,1} & \cdots & m_{1,k} \\ \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,k} \end{bmatrix}$ is the matrix whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$.
  - So the linear combination $a_1 \cdot \mathbf{v}_1 + \cdots + a_k \cdot \mathbf{v}_k$ is the zero vector precisely when the matrix product $M\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$.
  - By definition, the vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ will be linearly dependent when there is a nonzero $\mathbf{x}$ satisfying this matrix equation, and they will be linearly independent when the only solution is $\mathbf{x} = \mathbf{0}$.
- The terminology of "linear dependence" arises from the fact that if a set of vectors is linearly dependent, one of the vectors is necessarily a linear combination of the others (i.e., it "depends" on the others):

- <u>Proposition</u> (Dependence and Linear Combinations): A set S of vectors is linearly dependent if and only if one of the vectors is a linear combination of (some of) the others.
  - To avoid trivialities, we remark here that if $S = \{\mathbf{0}\}$ then the result is still correct, since the set of linear combinations (i.e., the span) of the empty set is the zero vector.
  - <u>Proof</u>: If $\mathbf{v}$ is a linear combination of other vectors in S, say $\mathbf{v} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$, then we have a nontrivial linear combination yielding the zero vector, namely $(-1) \cdot \mathbf{v} + a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$.
  - Conversely, suppose there is a nontrivial linear combination of vectors in S giving the zero vector, say, $b_1 \cdot \mathbf{v}_1 + b_2 \cdot \mathbf{v}_2 + \cdots + b_n \cdot \mathbf{v}_n = \mathbf{0}$. Since the linear combination is nontrivial, at least one of the coefficients is nonzero, say, b_i.
  - Then $b_i \cdot \mathbf{v}_i = (-b_1) \cdot \mathbf{v}_1 + \cdots + (-b_{i-1}) \cdot \mathbf{v}_{i-1} + (-b_{i+1}) \cdot \mathbf{v}_{i+1} + \cdots + (-b_n) \cdot \mathbf{v}_n$, and by scalar-multiplying both sides by $\dfrac{1}{b_i}$ (which exists because $b_i \neq 0$ by assumption) we see $\mathbf{v}_i = (-\dfrac{b_1}{b_i}) \cdot \mathbf{v}_1 + \cdots + (-\dfrac{b_{i-1}}{b_i}) \cdot \mathbf{v}_{i-1} + (-\dfrac{b_{i+1}}{b_i}) \cdot \mathbf{v}_{i+1} + \cdots + (-\dfrac{b_n}{b_i}) \cdot \mathbf{v}_n$.
  - Thus, one of the vectors is a linear combination of the others, as claimed.
- <u>Example</u>: Write one of the linearly dependent vectors $\langle 1, -1 \rangle$, $\langle 2, 2 \rangle$, $\langle 2, 1 \rangle$ as a linear combination of the others.
  - If we search for a linear dependence, we require $a \langle 1, -1 \rangle + b \langle 2, 2 \rangle + c \langle 2, 1 \rangle = \langle 0, 0 \rangle$.
  - By row-reducing the appropriate matrix we can find the solution $2 \langle 1, -1 \rangle + 3 \langle 2, 2 \rangle - 4 \langle 2, 1 \rangle = \langle 0, 0 \rangle$.
  - By rearranging we can then write $\boxed{\langle 1, -1 \rangle = -\dfrac{3}{2} \langle 2, 2 \rangle + 2 \langle 2, 1 \rangle}$.
  - We could have solved for any of the vectors that had a nonzero coefficient in the linear dependence, so we could alternatively have written $\boxed{\langle 2, 2 \rangle = -\dfrac{2}{3} \langle 1, -1 \rangle + \dfrac{4}{3} \langle 2, 1 \rangle}$, or $\boxed{\langle 2, 1 \rangle = \dfrac{1}{2} \langle 1, -1 \rangle + \dfrac{3}{4} \langle 2, 2 \rangle}$.
- Linear independence and span are related in a number of other ways as well. Here are two:
- <u>Theorem</u> (Independence and Span): Let S be a linearly independent subset of the vector space V, and $\mathbf{v}$ be any vector in V not in S. Then the set $S \cup \{\mathbf{v}\}$ is linearly dependent if and only if $\mathbf{v}$ is in $\text{span}(S)$.
  - <u>Proof</u>: If $\mathbf{v}$ is in $\text{span}(S)$, then one vector (namely $\mathbf{v}$) in $S \cup \{\mathbf{v}\}$ can be written as a linear combination of the other vectors (namely, the vectors in S). So by our earlier proposition, $S \cup \{\mathbf{v}\}$ is linearly dependent.
  - Conversely, suppose that $S \cup \{\mathbf{v}\}$ is linearly dependent, and consider a nontrivial dependence. If the coefficient of $\mathbf{v}$ were zero, then we would obtain a nontrivial dependence among the vectors in S (impossible, since S is linearly independent).
  - Thus, the coefficient of $\mathbf{v}$ is not zero: say, $a \cdot \mathbf{v} + b_1 \cdot \mathbf{v}_1 + \cdots + b_n \cdot \mathbf{v}_n = 0$ with $a \neq 0$ and for some $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in S.
  - Then $\mathbf{v} = (-\dfrac{b_1}{a}) \cdot \mathbf{v}_1 + \cdots + (-\dfrac{b_n}{a}) \cdot \mathbf{v}_n$ is a linear combination of the vectors in S, so $\mathbf{v}$ is in $\text{span}(S)$.
- <u>Theorem</u> (Characterization of Linear Independence): The set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent if and only if every vector $\mathbf{w}$ in $\text{span}(S)$ may be written as a linear combination $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$ for unique scalars $a_1, a_2, \dots, a_n$.
  - <u>Proof</u>: First suppose the decomposition is always unique. Then $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$ implies $a_1 = \cdots = a_n = 0$, because $0 \cdot \mathbf{v}_1 + \cdots + 0 \cdot \mathbf{v}_n = \mathbf{0}$ is by assumption the only decomposition of $\mathbf{0}$, so the vectors are linearly independent.
  - Now suppose that we had two ways of decomposing a vector $\mathbf{w}$, say as $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$ and as $\mathbf{w} = b_1 \cdot \mathbf{v}_1 + \cdots + b_n \cdot \mathbf{v}_n$.
  - By subtracting, we obtain $(a_1 - b_1) \cdot \mathbf{v}_1 + \cdots + (a_n - b_n) \cdot \mathbf{v}_n = \mathbf{w} - \mathbf{w} = \mathbf{0}$.
  - But now because $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent, we see that all of the scalar coefficients $a_1 - b_1, \cdots, a_n - b_n$ are zero. But this says $a_1 = b_1$, $a_2 = b_2$, $\dots$, $a_n = b_n$, meaning that the two decompositions are actually the same.

2.5.2 Linear Independence of Functions

- We can also say a few things about linear independence of general functions:
- Example: Determine whether the functions e^x and e^{2x} are linearly independent in the vector space of all real-valued functions.
  - Suppose that we had scalars a and b with $a \cdot e^x + b \cdot e^{2x} = 0$ for all x.
  - Taking the derivative of both sides with respect to x yields $a \cdot e^x + b \cdot 2e^{2x} = 0$.
  - Subtracting the original equation from this one produces $b \cdot e^{2x} = 0$, and since x is a variable we must have $b = 0$.
  - The first equation then gives $a \cdot e^x = 0$ so it must also be true that $a = 0$.
  - Thus, by definition, these functions are $\boxed{\text{linearly independent}}$.
- We can generalize the idea in the above example to give a method for determining whether a collection of functions is linearly independent:
- Definition: For n functions $y_1(x)$, $y_2(x)$, $\dots$, $y_n(x)$ which are each differentiable $n-1$ times, their Wronskian is defined to be $W(y_1, y_2, \cdots, y_n) = \det \begin{bmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{bmatrix}$. (Note that the Wronskian will also be a function of x.)
  - Example: The Wronskian of e^x, e^{2x} is $W(e^x, e^{2x}) = \begin{vmatrix} e^x & e^{2x} \\ e^x & 2e^{2x} \end{vmatrix} = e^{3x}$.
- Theorem (Linear Independence of Functions): Suppose that n functions $y_1, y_2, \dots, y_n$ which are each differentiable $n-1$ times have a Wronskian that is not the zero function. Then the functions are linearly independent in the vector space of real-valued functions.
  - Proof: Suppose that the functions are linearly dependent with $\sum_{j=1}^{n} a_j y_j = 0$, then by differentiating the appropriate number of times we see that $\sum_{j=1}^{n} a_j y_j^{(i)} = 0$ for any $0 \le i \le n$.
  - Hence, in particular, the rows of the Wronskian matrix are linearly dependent (as vectors), and so the determinant of the matrix is zero.
  - Therefore, if the Wronskian determinant is *not* zero, the functions cannot be dependent.
  - Remark: The theorem becomes an if-and-only if statement (i.e., the functions are linearly independent if and only if the Wronskian is nonzero) if we know that the functions $y_1, y_2, \cdots, y_n$ are infinitely differentiable. The proof of the other direction is significantly more difficult and we will omit it.
- Example: Show that the functions $\sin(x)$ and $\cos(x)$ are linearly independent using the Wronskian.
  - We compute $W(\sin(x), \cos(x)) = \begin{vmatrix} \sin(x) & \cos(x) \\ \cos(x) & -\sin(x) \end{vmatrix} = -1$, which is not the zero function.
- Example: Determine whether the functions $1+x$, $2-x$, and $3+4x$ are linearly dependent or linearly independent.
  - We compute $W(1, x, 1+x) = \begin{vmatrix} 1+x & 2-x & 3+4x \\ 1 & -1 & 4 \\ 0 & 0 & 0 \end{vmatrix} = 0$, by expanding along the bottom row.
  - Because these functions are infinitely differentiable and have Wronskian equal to the zero function, they are $\boxed{\text{linearly dependent}}$.

- A little searching will produce the explicit linear dependence $-11(1+x)+(2-x)+3(3+4x)=0$.

- Example: Compute the Wronskian of the functions x^2 and $x\,|x|$. Are they linearly dependent or linearly independent?

  - We compute $W(x^2, x\,|x|) = \begin{vmatrix} x^2 & x\,|x| \\ 2x & 2\,|x| \end{vmatrix} = 2x^2\,|x| - 2x^2\,|x| = 0$, which is the zero function. (One can verify using the definition of the derivative that $x\,|x|$ is differentiable everywhere and that its derivative is $2\,|x|$.)
  - This suggests these functions are linearly dependent. But in fact, they are linearly independent: if $a \cdot x^2 + b \cdot x\,|x| = 0$, then setting $x = 1$ produces $a + b = 0$ and setting $x = -1$ produces $a - b = 0$, and the only solution is $a = b = 0$.
  - Ultimately the issue is that $x\,|x|$ is not infinitely differentiable (its first derivative exists everywhere, but its second derivative does not exist at $x = 0$).

2.6 Bases and Dimension

- We will now combine the ideas of a spanning set and a linearly independent set, and use the resulting objects to study the structure of vector spaces.

2.6.1 Definition and Basic Properties of Bases

- Definition: A linearly independent set of vectors which spans V is called a basis for V.

  - Terminology Note: The plural form of the (singular) word "basis" is "bases".

- Example: Show that the vectors $\langle 1,0,0\rangle$, $\langle 0,1,0\rangle$, $\langle 0,0,1\rangle$ form a basis for $\mathbb{R}^3$.

  - The vectors certainly span $\mathbb{R}^3$, since we can write any vector $\langle a,b,c\rangle = a\cdot\langle 1,0,0\rangle + b\cdot\langle 0,1,0\rangle + c\cdot\langle 0,0,1\rangle$ as a linear combination of these vectors.
  - Furthermore, the vectors are linearly independent, because $a\cdot\langle 1,0,0\rangle + b\cdot\langle 0,1,0\rangle + c\cdot\langle 0,0,1\rangle = \langle a,b,c\rangle$ is the zero vector only when $a = b = c = 0$.
  - Thus, these three vectors are a linearly independent spanning set for $\mathbb{R}^3$, so they form a basis.

- A particular vector space can have several different bases:

- Example: Show that the vectors $\langle 1,1,1\rangle$, $\langle 2,-1,1\rangle$, $\langle 1,2,1\rangle$ also form a basis for $\mathbb{R}^3$.

  - Solving the system of linear equations determined by $x\cdot\langle 1,1,1\rangle + y\cdot\langle 2,-1,1\rangle + z\cdot\langle 1,2,1\rangle = \langle a,b,c\rangle$ for x,y,z will yield the solution $x = -3a-b+5c$, $y = a-c$, $z = 2a+b-3c$.
  - Therefore, $\langle a,b,c\rangle = (-3a-b+5c)\cdot\langle 1,1,1\rangle + (a-c)\cdot\langle 2,-1,1\rangle + (2a+b-3c)\cdot\langle 1,2,1\rangle$, so these three vectors span $\mathbb{R}^3$.
  - Furthermore, solving the system $x\cdot\langle 1,1,1\rangle + y\cdot\langle 2,-1,1\rangle + z\cdot\langle 1,2,1\rangle = \langle 0,0,0\rangle$ yields only the solution $x = y = z = 0$, so these three vectors are also linearly independent.
  - So $\langle 1,1,1\rangle$, $\langle 2,-1,1\rangle$, $\langle 1,2,1\rangle$ are a linearly independent spanning set for $\mathbb{R}^3$, meaning that they form a basis.

- Example: Find a basis for the vector space of 2×3 (real) matrices.

  - A general 2×3 matrix has the form $\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = a\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} + e\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + f\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
  - This decomposition suggests that we can take the set of six matrices $\boxed{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$ as a basis.

- Indeed, they certainly span the space of all 2×3 matrices, and they are also linearly independent, because the only linear combination giving the zero matrix is the one with $a = b = c = d = e = f = 0$.

- Non-Example: Show that the vectors $\langle 1, 1, 0 \rangle$ and $\langle 1, 1, 1 \rangle$ are not a basis for $\mathbb{R}^3$.
  - These vectors are linearly independent, since neither is a scalar multiple of the other.
  - However, they do not span $\mathbb{R}^3$ since, for example, it is not possible to obtain the vector $\langle 1, 0, 0 \rangle$ as a linear combination of $\langle 1, 1, 0 \rangle$ and $\langle 1, 1, 1 \rangle$.
  - Explicitly, since $a \cdot \langle 1, 1, 0 \rangle + b \cdot \langle 1, 1, 1 \rangle = \langle a + b,\ a + b,\ b \rangle$, there are no possible a, b for which this vector can equal $\langle 1, 0, 0 \rangle$, since this would require $a + b = 1$ and $a + b = 0$ simultaneously.

- Non-Example: Show that the vectors $\langle 1, 0, 0 \rangle$, $\langle 0, 1, 0 \rangle$, $\langle 0, 0, 1 \rangle$, $\langle 1, 1, 1 \rangle$ are not a basis for $\mathbb{R}^3$.
  - These vectors do span $\mathbb{R}^3$, since we can write any vector $\langle a, b, c \rangle = a \cdot \langle 1, 0, 0 \rangle + b \cdot \langle 0, 1, 0 \rangle + c \cdot \langle 0, 0, 1 \rangle + 0 \cdot \langle 1, 1, 1 \rangle$.
  - However, these vectors are not linearly independent, since we have the explicit linear dependence $1 \cdot \langle 1, 0, 0 \rangle + 1 \cdot \langle 0, 1, 0 \rangle + 1 \cdot \langle 0, 0, 1 \rangle + (-1) \cdot \langle 1, 1, 1 \rangle = \langle 0, 0, 0 \rangle$.

- Having a basis allows us to describe all the elements of a vector space in a particularly convenient way:

- Proposition (Characterization of Bases): The set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ forms a basis of the vector space V if and only if every vector $\mathbf{w}$ in V can be written in the form $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$ for *unique* scalars $a_1, a_2, \ldots, a_n$.
  - This proposition says that if we have a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ for V, then we can describe all of the other vectors in V in a particularly simple way (as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$) that is *unique*. A useful way to interpret this idea is to think of the basis vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ as "coordinate directions" and the coefficients $a_1, a_2, \ldots, a_n$ as "coordinates".
  - Proof: Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis of V. Then by definition, the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span the vector space V: every vector $\mathbf{w}$ in V can be written in the form $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$ for some scalars $a_1, a_2, \ldots, a_n$.
  - Furthermore, since the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent, by our earlier proposition every vector $\mathbf{w}$ in their span (which is to say, every vector in V) can be uniquely written in the form $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$, as claimed.
  - Conversely, suppose every vector $\mathbf{w}$ in V can be uniquely written in the form $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$. Then by definition, the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ span V.
  - Furthermore, by our earlier proposition, because every vector in $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ can be uniquely written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent: thus, they are a linearly independent spanning set for V, so they form a basis.

- If we have a description of the elements of a vector space, we can often extract a basis by direct analysis.

- Example: Find a basis for the space W of polynomials $p(x)$ of degree ≤ 3 such that $p(1) = 0$.
  - We remark that W is a subspace of the vector space V of polynomials, as it satisfies the subspace criterion. (We omit the verification.)
  - A polynomial of degree ≤ 3 has the form $p(x) = ax^3 + bx^2 + cx + d$ for constants a, b, c, d.
  - Since $p(1) = a + b + c + d$, the condition $p(1) = 0$ gives $a + b + c + d = 0$, so $d = -a - b - c$.
  - Thus, we can write $p(x) = ax^3 + bx^2 + cx + (-a - b - c) = a(x^3 - 1) + b(x^2 - 1) + c(x - 1)$, and conversely, any such polynomial has $p(1) = 0$.
  - Since every polynomial in W can be uniquely written as $a(x^3 - 1) + b(x^2 - 1) + c(x - 1)$, we conclude that the set $\boxed{\{x^3 - 1,\ x^2 - 1,\ x - 1\}}$ is a basis of W.

2.6.2 Existence of Bases

- We now turn our attention to the question of constructing bases for general vector spaces.
- Theorem (Spanning Sets and Bases): If V is a vector space, then any spanning set for V contains a basis of V.
  - In the event that the spanning set is infinite, the argument is rather delicate and technical (and requires an ingredient known as the axiom of choice), so we will only treat the case of a finite spanning set consisting of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.
  - Proof (finite spanning set case): Suppose $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ spans V. We construct an explicit subset that is a basis for V.
  - Start with an empty collection S_0 of elements. Now, for each $1 \le k \le n$, perform the following procedure:
    * Check whether the vector $\mathbf{v}_k$ is contained in the span of S_{k-1}. (Note that the span of the empty set is the zero vector.)
    * If $\mathbf{v}_k$ is not in the span of S_{k-1}, let $S_k = S_{k-1} \cup \{\mathbf{v}_k\}$. Otherwise, let $S_k = S_{k-1}$.
  - We claim that the set S_n is a basis for V. Roughly speaking, the idea is that the collection of elements which we have not thrown away will still be a generating set (since removing a dependent element will not change the span), but the collection will also now be linearly independent (since we threw away elements which were dependent).
  - To see that S is linearly independent, observe that if $\mathbf{v}_k$ is included in S_k, then $\mathbf{v}_k$ is linearly independent from the vectors already in S_{k-1} (as it is not in the span of S_{k-1}). Thus, each time we add a new vector, we preserve the linear independence, so when the procedure terminates, S_n will be linearly independent.
  - To see that S_n spans V, the idea is to observe that the span of S is the same as the span of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.
  - Explicitly, consider any vector $\mathbf{v}_k$ that is not in S_n: it was not included in S_k during the construction, so it must have been contained in the span of the vectors already in S_{k-1}. Therefore, adding $\mathbf{v}_k$ to S_n will not change the span. Doing this for each vector $\mathbf{v}_k$ not in S will not change the span and will yield the set $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, so we conclude $\text{span}(S) = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = V$.
- By removing elements from a spanning set, we can construct a basis for any vector space.
- Theorem (Bases of Vector Spaces): Every vector space V has a basis.
  - Proof: Let S to be any spanning set for V. (For example, we could take S to be the set of all vectors in V.) Then since S spans V, it contains a basis for V.
  - Remark: That a basis always exists is incredibly helpful, and is without a doubt the most useful fact about vector spaces. Vector spaces in the abstract are very hard to think about, but a vector space with a basis is something very concrete, since the existence of a basis allows us to describe all the vectors in a precise and regular form.
- The above procedure allows us to construct a basis for a vector space by "dropping down" by removing linearly dependent vectors from a spanning set. We can also construct bases for vector spaces by "building up" from a linearly independent set.
- Theorem (Building-Up Theorem): Given any linearly independent set of vectors in V, there exists a basis of V containing those vectors. In short, any linearly independent set of vectors can be extended to a basis.
  - Proof: Let S be a set of linearly independent vectors. (In any vector space, the empty set is always linearly independent.)
    1. If S spans V, then we are done, because then S is a linearly independent generating set; i.e., a basis.
    2. If S does not span V, there is an element $\mathbf{v}$ in V which is not in the span of S. Put $\mathbf{v}$ in S: then by hypothesis, the new S will still be linearly independent.
    3. Repeat the above two steps until S spans V.
  - If V is "finite-dimensional" (see below), then we will always be able to construct a basis in a finite number of steps. In the case where V is "infinite-dimensional", matters are trickier, and we will omit the very delicate technical details required to deal with this case.

- Now that we have several ways of constructing bases, we can use them to study linear independence and span:
- Theorem (Bases and Dependence): Suppose V has a basis with n elements. If $m > n$, then any set of m vectors of V is linearly dependent.
  - Proof: Suppose B is a basis with n elements and A is a set of m vectors with $m > n$. Then since B is a basis, we can write every element $\mathbf{a}_i$ in A as a linear combination of the elements of B, say as $\mathbf{a}_i = \sum_{j=1}^{n} c_{i,j} \cdot \mathbf{b}_j$ for $1 \le i \le m$.
  - We would like to see that there is some choice of scalars d_k, not all zero, such that $\sum_{k=1}^{n} d_k \cdot \mathbf{a}_k = \mathbf{0}$: this will show that the vectors $\mathbf{a}_i$ are linearly dependent.
  - So consider a linear combination $\sum_{k=1}^{n} d_k \cdot \mathbf{a}_k = \mathbf{0}$: if we substitute in for the vectors in B, then we obtain a linear combination of the elements of B equalling the zero vector. Since B is a basis, this means each coefficient of $\mathbf{b}_j$ in the resulting expression must be zero.
  - If we tabulate the resulting system, we can check that it is equivalent to the matrix equation $C\mathbf{d} = \mathbf{0}$, where C is the $m \times n$ matrix of coefficients with entries $c_{i,j}$, and $\mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$ is the $n \times 1$ matrix with entries the scalars d_k.
  - Since C is a matrix which has more rows than columns, by the assumption that $m > n$, we see that the homogeneous system $C\mathbf{d} = \mathbf{0}$ has a nonzero solution for $\mathbf{d}$. (There is at most one pivot per column, and so there must be at least one row that does not have a pivot.)
  - But then we have $\sum_{k=1}^{n} d_k \cdot \mathbf{a}_k = \mathbf{0}$ for some scalars d_k not all of which are zero, meaning that the set A is linearly dependent.
- Using the result above, we can deduce the fundamental fact that any two bases of a vector space must have the same size:
- Theorem (Bases of Vector Spaces): Every vector space has a basis, and any two bases of V contain the same number of elements.
  - Proof: We already showed above that every vector space has a basis.
  - For the other statement, if every basis of V is infinite, we are done. Now suppose that V has some finite basis, and choose B to be a basis of minimal size.
  - Suppose B has n elements, and consider any other basis B' of V. By the previous theorem about bases and linear independence, if B' contains more than n vectors, it would be linearly dependent, which is impossible.
  - Thus, B' has at most n elements, but since B had minimal size, B' cannot have fewer than n elements. Thus B' has exactly n elements; since B' was arbitrary, this means every basis of V has n elements.

2.6.3 Dimension

- Definition: If V is a vector space, the number of elements in any basis of V is called the dimension of V and is denoted $\dim(V)$.
  - Our results above assure us that the dimension of a vector space is always well-defined: every vector space has a basis, and any other basis will have the same number of elements.
- Here are a few examples:

- Example: The dimension of $\mathbb{R}^n$ is n, since the n standard unit vectors form a basis. (This at least suggests that the term "dimension" is reasonable, since it is the same as our usual notion of dimension.)
- Example: The dimension of the vector space $M_{m\times n}(\mathbb{R})$ of $m \times n$ matrices is mn, because there is a basis consisting of the mn matrices $E_{i,j}$, where $E_{i,j}$ is the matrix with a 1 in the (i,j)-entry and 0s elsewhere.
- Example: The dimension of the vector space $P(\mathbb{R})$ of all polynomials is ∞, because the (infinite list of) polynomials $1, x, x^2, x^3, \cdots$ are a basis for the space.
- Example: The dimension of the vector space $P_n(\mathbb{R})$ of polynomials of degree at most n is $n+1$, because $\{1, x, x^2, \dots, x^n\}$ is a basis for the space.
- Example: The dimension of the zero space is 0, because the empty set (containing 0 elements) is a basis.
- Example: The dimension of the space of complex numbers is 2, since the set $\{1, i\}$ forms a basis.

- Proposition: If W is a subspace of V, then $\dim(W) \leq \dim(V)$.
  - Proof: Choose any basis of W: it is a linearly independent set of vectors in V, so it is contained in some basis of V by the Building-Up Theorem.

- The result above tells us that a subspace of a finite-dimensional vector space will also be finite-dimensional, and thus have a finite basis. The most direct way of computing the dimension of a vector space or subspace is simply to find a basis explicitly.

- Example: Find the dimension of the vector space V of 3×3 matrices A satisfying $A^T = -A$.
  - If $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is such a matrix, then $A^T = -A$ requires $\begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} = -\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, so that $a = e = i = 0$, $b = d$, $c = g$, and $h = f$.
  - Then $A = \begin{bmatrix} 0 & b & c \\ -b & 0 & f \\ -c & -f & 0 \end{bmatrix} = b \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c \cdot \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + f \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.
  - Thus, the three matrices $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ form a basis for the space (the calculation above shows that they span it, and they are clearly linearly independent), so $\dim(V) = 3$.

- In general, finite-dimensional vector spaces are much better-behaved than infinite-dimensional vector spaces. We will therefore usually focus our attention on finite-dimensional spaces, since infinite-dimensional spaces can have occasional counterintuitive properties. For example:

- Example: The dimension of the vector space of all real-valued functions on the interval $[0,1]$ is ∞, because it contains the infinite-dimensional vector space of polynomials.
  - We have not actually written down a basis for the vector space of all real-valued functions on the interval $[0,1]$, although (per our earlier results) this vector space does have a basis.
  - There is a good reason for this: it is not possible to give a simple description of such a basis.
  - The set of functions $f_a(x) = \begin{cases} 1 & \text{if } x = a \\ 0 & \text{if } x \neq a \end{cases}$, for real numbers a, does not form a basis for the space of real-valued functions: although this infinite set of vectors is linearly independent, it does not span the space, since (for example) the constant function $f(x) = 1$ cannot be written as a finite linear combination of these functions.

2.6.4 Finding Bases for $\mathbb{R}^n$, Row Spaces, Column Spaces, and Nullspaces

- The fact that every vector space has a basis is extremely useful from a theoretical standpoint. We will now discuss some practical methods for finding bases for particular vector spaces that often arise in computational applications of linear algebra.
  - Our results provide two different methods for constructing a basis for a given vector space.
  - One way is to "build" a linearly independent set of vectors into a basis by adding new vectors one at a time (choosing a vector not in the span of the previous vectors) until a basis is obtained.
  - Another way is to "reduce" a spanning set by removing linearly dependent vectors one at a time (finding and removing a vector that is a linear combination of the others) until a basis is obtained.
- Proposition (Bases, Span, Dependence): If V is an n-dimensional vector space, then any set of fewer than n vectors cannot span V, and any set of more than n vectors is linearly dependent. Furthermore, a set of exactly n vectors is a basis if and only if it spans V, if and only if it is linearly independent.
  - Proof: Suppose first that S is a set of fewer than n vectors in V.
  - Then since S spans $\text{span}(S)$ by definition, S contains a basis T for $\text{span}(S)$, and T is a linearly independent set of fewer than n vectors in V.
  - Thus, we can extend T to a basis of V, which necessarily contains n elements, strictly more than in T. So there is some vector $\mathbf{v}$ in this extended basis that is not in T: then $\mathbf{v}$ is not in $\text{span}(S)$, so S does not span V.
  - Now suppose that S is a set of more than n vectors in V that is linearly independent. We would then be able to extend S to a basis of V, but this is impossible because any basis contains only n elements.
  - Finally, suppose S contains exactly n vectors. If S is a basis, it is by definition a spanning set and linearly independent, so it remains to show that if S spans V then it is a basis, and if S is linearly independent then it is a basis.
  - If S spans V, then S contains a basis: but since the basis must have n elements, the basis is the entire set S. If S is linearly independent, then S is contained in a basis of V: but we cannot add any more vectors without making the set linearly dependent, so S must already be a basis.
- Example: Determine whether the vectors $\langle 1,2,2,1\rangle$, $\langle 3,-1,2,0\rangle$, $\langle -3,2,1,1\rangle$ span $\mathbb{R}^4$.
  - They $\boxed{\text{do not span}}$: since $\mathbb{R}^4$ is a 4-dimensional space, any spanning set must contain at least 4 vectors.
- Example: Determine whether the vectors $\langle 1,2,1\rangle$, $\langle 1,0,3\rangle$, $\langle -3,2,1\rangle$, $\langle 1,1,4\rangle$ are linearly independent.
  - They $\boxed{\text{are not linearly independent}}$: since $\mathbb{R}^3$ is a 3-dimensional space, any 4 vectors in $\mathbb{R}^3$ are automatically linearly dependent.
- We can also characterize bases of $\mathbb{R}^n$:
- Theorem (Bases of $\mathbb{R}^n$): A collection of k vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is a basis if and only if $k = n$ and the $n \times n$ matrix M, whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, is an invertible matrix.
  - Remark: The statement that B is invertible is equivalent to saying that $\det(M) \neq 0$. This gives a rapid computational method for determining whether a given set of vectors forms a basis.
  - Proof: Since $\mathbb{R}^n$ has a basis with n elements, any basis must have n elements by our earlier results, so $k = n$.
  - Now suppose $\mathbf{v}_1, \dots, \mathbf{v}_n$ are vectors in $\mathbb{R}^n$. For any vector $\mathbf{w}$ in $\mathbb{R}^n$, consider the problem of finding scalars $a_1, \cdots, a_n$ such that $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{w}$.
  - This vector equation is the same as the matrix equation $M\mathbf{a} = \mathbf{w}$, where M is the matrix whose columns are the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, $\mathbf{a}$ is the column vector whose entries are the scalars $a_1, \dots, a_n$, and $\mathbf{w}$ is thought of as a column vector.

- By our earlier results, $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$ precisely when the scalars $a_1 \dots, a_n$ are unique. In turn this is equivalent to the statement that $M\mathbf{a} = \mathbf{w}$ has a unique solution $\mathbf{a}$ for any $\mathbf{w}$.
- From our study of matrix equations, this equation has a unique solution precisely when M is an invertible matrix, as claimed.

- Example: Determine whether the vectors $\langle 1, 2, 1\rangle$, $\langle 2, -1, 2\rangle$, $\langle 3, 3, 1\rangle$ form a basis of $\mathbb{R}^3$.
  - By the theorem, we only need to determine whether the matrix $M = \begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 3 \\ 1 & 2 & 1 \end{bmatrix}$ is invertible.
  - We compute $\det(M) = 1 \begin{vmatrix} -1 & 3 \\ 2 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} = 10$ which is nonzero.
  - Thus, M is invertible, so these vectors $\boxed{\text{do form a basis of } \mathbb{R}^3}$.

- Associated to any matrix M are three spaces that often arise when doing matrix algebra and studying the solutions to systems of linear equations.

- Definition: If M is an $m \times n$ matrix, the row space of M is the subspace of $\mathbb{R}^n$ spanned by the rows of M, the column space of M is the subspace of $\mathbb{R}^m$ spanned by the columns of M, and the nullspace of M is the set of vectors $\mathbf{x}$ in $\mathbb{R}^n$ for which $M\mathbf{x} = \mathbf{0}$.
  - By definition the row space and column spaces are subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, since the span of any set of vectors is a subspace.
  - It is also simple to verify that the nullspace is a subspace of $\mathbb{R}^m$ via the subspace criterion.

- We have already studied in detail the procedure for solving a matrix equation $M\mathbf{x} = \mathbf{0}$, which requires row-reducing the matrix M. It turns out that we can obtain a basis for the row and column spaces from a row-echelon form of M as well:

- Theorem (Bases for Row and Column Spaces): If M is an $m \times n$ matrix, let E be any row-echelon form of M. If r is the number of pivots in E, then the row space and column space are both r-dimensional and the nullspace is $(n - r)$-dimensional. Furthermore, a basis for the row space is given by the nonzero rows of E, while a basis for the column space is given by the columns of M that correspond to the pivotal columns of E.
  - For the column space, we also remark that another option would be to row-reduce the transpose matrix M^T, since the columns of M are the rows of M^T. This will produce a basis that is easier to work with, but it is not actually necessary to perform the extra calculations.
  - Proof: First consider the row space, which by definition is spanned by the rows of M.
  - Observe that each elementary row operation does not change the span of the rows of M: for any vectors $\mathbf{v}_i$ and $\mathbf{v}_j$, we have $\text{span}(\mathbf{v}_i, \mathbf{v}_j) = \text{span}(\mathbf{v}_j, \mathbf{v}_i)$, $\text{span}(c\mathbf{v}) = \text{span}(\mathbf{v})$ for any nonzero c, and $\text{span}(\mathbf{v}_i, \mathbf{v}_j) = \text{span}(\mathbf{v}_i + c\mathbf{v}_j, \mathbf{v}_j)$ for any c.
  - So we may put M into a row-echelon form E without altering the span. Now we claim that the nonzero rows $\mathbf{r}_1, \dots, \mathbf{r}_r$ of E are linearly independent. Ultimately, this is because of the presence of the pivot elements: if $a_1 \cdot \mathbf{r}_1 + \dots + a_r \cdot \mathbf{r}_r = \mathbf{0}$ then each of the vectors $\mathbf{r}_1$, ... , $\mathbf{r}_r$ will have a leading coefficient in an entry that is zero in all of the subsequent vectors, so the only solution to the associated system of linear equations is $a_1 = \dots = a_r = 0$.
  - Now consider the column space. Observe first that the set of solutions $\mathbf{x}$ to the matrix equation $M\mathbf{x} = \mathbf{0}$ is the same as the set of solutions to the equation $E\mathbf{x} = \mathbf{0}$, by our analysis of row-operations.
  - Now if we write $\mathbf{x} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ and expand out each matrix product in terms of the columns of M and E, we will see that $M\mathbf{x} = a_1 \cdot \mathbf{c}_1 + \dots + a_n \cdot \mathbf{c}_n$ and $E\mathbf{x} = a_1 \cdot \mathbf{e}_1 + \dots + a_n \cdot \mathbf{e}_n$ where the $\mathbf{c}_i$ are the columns of M and the $\mathbf{e}_i$ are the columns of E.
  - Combining these two observations shows that, for any scalars $a_1, \dots, a_n$, we have $a_1 \cdot \mathbf{c}_1 + \dots + a_n \cdot \mathbf{c}_n = \mathbf{0}$ if and only if $a_1 \cdot \mathbf{e}_1 + \dots + a_n \cdot \mathbf{e}_n = \mathbf{0}$.

- What this means is that any linear dependence between the columns of M gives a linear dependence between the corresponding columns of E (with the same coefficients), and vice versa. So it is enough to determine a basis for the column space of E: then a basis for the column space of M is simply the corresponding columns in M.
- All that remains is to observe that the set of pivotal columns for E forms a basis for the column space of E: the pivotal columns are linearly independent by the same argument given above for rows, and every other column lies in their span (specifically, any column lies in the span of the pivotal columns to its left, since each row has a pivotal element).
- Finally, the statement about the dimensions of the row and column spaces follows immediately from our descriptions, and the statement about the dimension of the nullspace follows by observing that the matrix equation $M\mathbf{x} = \mathbf{0}$ has r bound variables and $n - r$ free variables.

- <u>Example</u>: Find a basis for the row space, the column space, and the nullspace of the matrix $M = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}$, as well as the dimension of each space.
  - We begin by row-reducing the matrix M:
  $$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{R_3 - R_1} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & -1 & 2 \end{bmatrix} \xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$
  - The row space is spanned by the two vectors $\boxed{\langle 1, 0, 2, 1 \rangle, \langle 0, 1, -1, 2 \rangle}$.
  - Since columns 1 and 2 have pivots, the first two columns of M give a basis for the column space: $\boxed{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}$.
  - For the nullspace, there are two free variables corresponding to columns 3 and 4. Solving the corresponding system (with variables x_1, x_2, x_3, x_4 and free parameters a, b) yields the solution set $\langle x_1, x_2, x_3, x_4 \rangle = \langle -2a - b, a - 2b, a, b \rangle = a \langle -2, 1, 1, 0 \rangle + b \langle -1, -2, 0, 1 \rangle$.
  - Thus, a basis for the nullspace is given by $\boxed{\langle -2, 1, 1, 0 \rangle, \langle -1, -2, 0, 1 \rangle}$.
  - The row space, column space, and nullspace all have dimension $\boxed{2}$.

- <u>Example</u>: Find a basis for the row space, the column space, and the nullspace of $M = \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ -2 & 2 & 0 & -3 & 1 \\ 1 & -1 & 0 & 3 & 8 \end{bmatrix}$, as well as the dimension of each space.
  - We begin by row-reducing the matrix M:
  $$\begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ -2 & 2 & 0 & -3 & 1 \\ 1 & -1 & 0 & 3 & 8 \end{bmatrix} \underset{R_3 - R_1}{\xrightarrow{R_2 + 2R_1}} \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 2 & 7 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$
  - The row space is spanned by the three vectors $\boxed{\langle 1, -1, 0, 2, 1 \rangle, \langle 0, 0, 0, 1, 3 \rangle, \langle 0, 0, 0, 0, 1 \rangle}$.
  - Since there are pivots in columns 1, 4, and 5, those columns of M give a basis for the column space: $\boxed{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}}$.
  - For the nullspace, there are two free variables corresponding to columns 2 and 3. Solving the corresponding system (with variables x_1, x_2, x_3, x_4, x_5 and free parameters a, b) yields the solution set $\langle x_1, x_2, x_3, x_4, x_5 \rangle = \langle a, a, b, 0, 0 \rangle = a \langle 1, 1, 0, 0, 0 \rangle + b \langle 0, 0, 1, 0, 0 \rangle$, so a basis for the nullspace is given by $\boxed{\langle 1, 1, 0, 0, 0 \rangle, \langle 0, 0, 1, 0, 0 \rangle}$.

- As particular applications, we can use these ideas to give algorithms for reducing a spanning set to a basis and for building a basis from a linearly independent set.
  - To reduce a spanning set to a basis, we write down the associated matrix (whose columns are the elements of the spanning set) and then row-reduce it: the columns corresponding to pivotal columns will then be a basis for the column space.
  - To build a linearly independent set S into a basis, we first find additional vectors so that the resulting set spans the space, and then (listing the vectors in S first) reduce this spanning set to a basis using the procedure above.
  - Note that using either of these procedures will require us to have chosen a particular basis for the space already (since we need to work with the coefficient vectors for the elements of our spanning set).

- Example: Given $S = \{\langle 1,0,1,2\rangle, \langle 3,0,3,6\rangle, \langle 2,1,2,1\rangle, \langle 3,1,3,3\rangle\}$, find a subset of S that is a basis for $\text{span}(S)$.
  - We simply row-reduce the matrix whose columns are the vectors in S:

$$\begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & 3 & 2 & 3 \\ 2 & 6 & 1 & 3 \end{bmatrix} \xrightarrow[R_4-2R_2]{R_3-R_1} \begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & -3 \end{bmatrix} \xrightarrow{R_4+3R_2} \begin{bmatrix} 1 & 3 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

  - Since the first and third columns are pivotal, we conclude that the vectors $\boxed{\langle 1,0,1,2\rangle, \langle 2,1,2,1\rangle}$ are a basis for the column space, which is the same as $\text{span}(S)$.

- Example: Extend the set $S = \{1-2x^2, 2+x\}$ to a basis for $P_3(\mathbb{R})$.
  - We extend S to a spanning set, and then reduce the result to a basis: the easiest way to do this is simply to append some other basis to S. Let us append the standard basis $\{1, x, x^2, x^3\}$: we therefore want to reduce $S' = \{1-2x^2, 2+x, 1, x, x^2, x^3\}$ to a basis.
  - To do this, row-reduce the matrix whose columns are the coordinate vectors of the elements of S':

$$\begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3+2R_1} \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 4 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3-4R_2} \begin{bmatrix} 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

  - Since columns 1, 2, 3, and 6 are pivotal, we conclude that $\boxed{\{1-2x^2, 2+x, 1, x^3\}}$ is a basis for $P_3(\mathbb{R})$.

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