Linear Algebra (part 1) : Vector Spaces (by Evan Dummit, 2022, v. 3.00)

Contents

1 Vector Spaces

In this chapter we will introduce the notion of a vector space, which generalizes the notion of vectors in 2- and 3-dimensional space. We introduce vector spaces from an axiomatic perspective, deriving a number of basic properties using only the axioms, and then develop the general theory of vector spaces. Specifically, we discuss subspaces, span, linear dependence and independence, and bases, in particular showing that every vector space possesses a basis (a linearly independent spanning set) and giving methods for finding bases of vector spaces.

1.1 The Formal Definition of a Vector Space

- The two operations of addition and scalar multiplication (and the various algebraic properties they satisfy) are the key properties of vectors in $\mathbb{R}^n$ and of matrices. We would like to investigate other collections of things which possess those same properties.

- Definition: Let F be a field, and refer to the elements of F as scalars. A vector space over F is a triple $(V, +, \cdot)$ of a collection V of elements called vectors, together with two binary operations[1], addition of vectors $(+)$ and scalar multiplication of a vector by a scalar $(\cdot)$, satisfying the following axioms:

  [V1] Addition is commutative: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ for any vectors $\mathbf{v}$ and $\mathbf{w}$.

  [V2] Addition is associative: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

  [V3] There exists a zero vector $\mathbf{0}$, with $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$ for any vector $\mathbf{v}$.

  [V4] Every vector $\mathbf{v}$ has an additive inverse $-\mathbf{v}$, with $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$.

  [V5] 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 $\mathbf{v}$.

  [V6] Addition of scalars distributes: $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$ for any scalars α, β and vector $\mathbf{v}$.

  [V7] 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 $\mathbf{w}$.

  [V8] The scalar 1 acts like the identity on vectors: $1 \cdot \mathbf{v} = \mathbf{v}$ for any vector $\mathbf{v}$.

[1] The result of adding vectors $\mathbf{v}$ and $\mathbf{w}$ is denoted as $\mathbf{v} + \mathbf{w}$, and the result of scalar-multiplying $\mathbf{v}$ by α is denoted as $\alpha \cdot \mathbf{v}$ (or often simply $\alpha\mathbf{v}$). The definition of "binary operation" means that $\mathbf{v} + \mathbf{w}$ and $\alpha \cdot \mathbf{v}$ are also vectors in V.

- We will primarily consider vector spaces where the collection of scalars (namely, the field F) is either the set of real numbers or the set of complex numbers: we refer to such vector spaces as real vector spaces or complex vector spaces, respectively.
  - However, all of the general theory of vector spaces will hold over any field. Some complications can arise in certain kinds of fields (such as the two-element field $\mathbb{F}_2 = \{0, 1\}$ where $1 + 1 = 0$) where adding 1 a finite number of times to itself yields 0; we will generally seek to gloss over such complications.
- Here are some examples of vector spaces:
- Example: The vectors in $\mathbb{R}^n$ are a real vector space, for any $n > 0$.
  - 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 $\alpha \cdot \langle x, y\rangle = \langle \alpha x, \alpha y\rangle$. (The dot here is *not* the dot product!)
  - [V1]: We have $\langle x_1, y_1\rangle + \langle x_2, y_2\rangle = \langle x_1 + x_2, y_1 + y_2\rangle = \langle x_2, y_2\rangle + \langle x_1, y_1\rangle$.
  - [V2]: We have $(\langle x_1, y_1\rangle + \langle x_2, y_2\rangle) + \langle x_3, y_3\rangle = \langle x_1 + x_2 + x_3, y_1 + y_2 + y_3\rangle = \langle x_1, y_1\rangle + (\langle x_2, y_2\rangle + \langle x_3, y_3\rangle)$.
  - [V3]: The zero vector is $\langle 0, 0\rangle$, and clearly $\langle x, y\rangle + \langle 0, 0\rangle = \langle x, y\rangle$.
  - [V4]: The additive inverse of $\langle x, y\rangle$ is $\langle -x, -y\rangle$, since $\langle x, y\rangle + \langle -x, -y\rangle = \langle 0, 0\rangle$.
  - [V5]: We have $\alpha_1 \cdot (\alpha_2 \cdot \langle x, y\rangle) = \langle \alpha_1\alpha_2 x, \alpha_1\alpha_2 y\rangle = (\alpha_1\alpha_2) \cdot \langle x, y\rangle$.
  - [V6]: We have $(\alpha_1 + \alpha_2) \cdot \langle x, y\rangle = \langle (\alpha_1 + \alpha_2)x, (\alpha_1 + \alpha_2)y\rangle = \alpha_1 \cdot \langle x, y\rangle + \alpha_2 \cdot \langle x, y\rangle$.
  - [V7]: We have $\alpha \cdot (\langle x_1, y_1\rangle + \langle x_2, y_2\rangle) = \langle \alpha(x_1 + x_2), \alpha(y_1 + y_2)\rangle = \alpha \cdot \langle x_1, y_1\rangle + \alpha \cdot \langle x_2, y_2\rangle$.
  - [V8]: Finally, we have $1 \cdot \langle x, y\rangle = \langle x, y\rangle$.
- Example: The set $M_{m\times n}(F)$ of $m \times n$ matrices, for any fixed m and n, forms a vector space over F.
  - The various algebraic properties we know about matrix addition give [V1] and [V2] along with [V5], [V6], [V7], and [V8].
  - The "zero vector" in this vector space is the zero matrix (all entries zero), and [V3] and [V4] 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 are a real vector space under normal addition and multiplication.
  - The axioms all follow from the standard properties of complex numbers: the "zero vector" is $0 = 0 + 0i$, and the additive inverse of $a + bi$ is $-a - bi$.
  - 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: If F is any field and S is any set, the collection of all functions from S to F is a vector space over F, where we define the sum of two functions as $(f + g)(x) = f(x) + g(x)$ for every x, and scalar multiplication via $(\alpha \cdot f)(x) = \alpha\, f(x)$.
  - 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 the properties of the field F: we simply verify that each axiom holds at every value x in S. The "zero vector" in this space is the zero function; namely, the function 0 which has $0(x) = 0$ for every x.
  - For example (to demonstrate a few of the axioms), for any value x in S and any functions f and g,
    * [V1]: $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$.
    * [V6]: $\alpha \cdot (f + g)(x) = \alpha\, f(x) + \alpha\, g(x) = (\alpha f)(x) + (\alpha g)(x)$.
    * [V8]: $(1 \cdot f)(x) = 1 \cdot f(x) = f(x)$.

- Example: If F is any field, the space $F[x]$ of polynomials in x with coefficients in F is a vector space over F.
  - This follows in the same way as the verification for general functions.
- Example: If F is any field, 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 over F.
  - 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 [V2] and [V3], 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 [V1]-[V4] 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 [V3], $\mathbf{v}=\mathbf{0}+\mathbf{v}$, 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 [V4], $\mathbf{0}=(-\mathbf{v})+\mathbf{v}$, 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 [V6] and [V8] 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 [V3], 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 [V5] and [V8] we have $\alpha\cdot\mathbf{0}=\alpha\cdot(\mathbf{0}+\mathbf{0})=\alpha\cdot\mathbf{0}+\alpha\cdot\mathbf{0}$.
     - Thus, by [V3], 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 [V6]-[V8] 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 (6) twice and [V7]-[V8], $-(-\mathbf{v})=(-1)\cdot(-\mathbf{v})=(-1)\cdot[(-1)\cdot\mathbf{v}]=(-1)^2\cdot\mathbf{v}=1\cdot\mathbf{v}=\mathbf{v}$.
  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 an element of a field, it has a multiplicative inverse α^{-1}.
     - Then by property (5) and [V5], [V8], 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 [V3], $\mathbf{0}-\mathbf{v}=\mathbf{0}+(-\mathbf{v})=-\mathbf{v}$.
  10. Negation distributes over addition: $-(\mathbf{v}+\mathbf{w})=(-\mathbf{v})+(-\mathbf{w})=-\mathbf{v}-\mathbf{w}$.

   - Proof: By property (6) and [V7], $-(\mathbf{v}+\mathbf{w}) = (-1)\cdot(\mathbf{v}+\mathbf{w}) = (-1)\cdot\mathbf{v} + (-1)\cdot\mathbf{w} = (-\mathbf{v}) + (-\mathbf{w})$.
   - 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.
   - (Outline): Induct on the number of terms. The base cases follow from the axioms [V1] and [V2].
   - 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.

1.2 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.
- 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.
- 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:
  - [V1] Addition is commutative: $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} + \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
  - [V3] The zero element is the zero matrix, since $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
  - [V4] The additive inverse of $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ is $\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$ since $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
  - [V5] Scalar multiplication is consistent: $p \cdot q \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} pqa & 0 \\ 0 & pqb \end{bmatrix} = pq \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
- It is very time-consuming to verify all of the axioms for a subspace, and much of the work seems to be redundant. Conveniently, we can 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 $\mathbf{0}_W$: then we can write $\mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W = \mathbf{0}_V + \mathbf{0}_W$ by [V3] in V and [V3] in W respectively, so by cancellation we get $\mathbf{0}_W = \mathbf{0}_V$ meaning that 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 [V1]-[V2] and [V5]-[V8] will hold in W because they hold in V. Axiom [V3] for W follows from [S1] since $\mathbf{0}_W = \mathbf{0}_V$ as noted above. Finally, for [V4], 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.

- 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 $n \times n$ matrices of trace zero is a subspace of the space of all $n \times n$ matrices.
  - [S1]: The zero matrix has trace zero.
  - [S2]: Since $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$, we see that if A and B have trace zero then so does $A + B$.
  - [S3]: Since $\mathrm{tr}(\alpha A) = \alpha \mathrm{tr}(A)$, 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 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 $2t_1 t_2 = 0$.)
  - 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 $\langle x, y, z \rangle$ where $x, y, z \geq 0$ forms a subspace of $\mathbb{R}^2$.
  - It is $\boxed{\text{not a subspace}}$: 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.

- There are a few more general subspaces that serve as important examples.
- Example: For any interval $[a, b]$, show that the collection of continuous functions on $[a, b]$ is a subspace of the space of all functions on $[a, b]$, as is the set of n-times differentiable functions on $[a, b]$ for any positive integer n.
  - We show each of these sets is a subspace of the collection of all (real-valued) functions on the interval $[a, b]$, which we already know is a vector space.
  - For the first statement, observe that the zero function is continuous, that the sum of two continuous functions is continuous, and that any scalar multiple of a continuous function is continuous.
  - The second statement follows in the same way: the zero function is also n-times differentiable, as is the sum of two n-times differentiable functions and any scalar multiple of an n-times differentiable function.
- Example: Show that the real-valued 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 real-valued 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. (But for completeness, the general solution is $y = Ae^{-x} + Be^{-5x}$ for arbitrary constants A and B.)
- We can use the subspace criterion to give easier proofs of a number of results about subspaces, such as the following:
- Proposition (Intersection of Subspaces): If V is a vector space, the intersection of any collection of subspaces of V is also a subspace of V.
  - Proof: Let $\mathcal{S}$ be a collection of subspaces of V and take $I = \bigcap_{W\in\mathcal{S}} W$ to be the intersection of the subspaces in $\mathcal{S}$. By the subspace criterion, the zero vector of V is in each subspace in $\mathcal{S}$, so it also is contained in I.
  - Now let $\mathbf{w}_1$ and $\mathbf{w}_2$ be any vectors in I, and α be any scalar. By the definition of I, the vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ are in each subspace W in S.
  - So by the subspace criterion, $\mathbf{w}_1 + \mathbf{w}_2$ and $\alpha \cdot \mathbf{w}_1$ are also in each subspace W in S: but this means both $\mathbf{w}_1 + \mathbf{w}_2$ and $\alpha \cdot \mathbf{w}_1$ are in I.
  - Thus, I satisfies each component of the subspace criterion, so it is a subspace of V.
  - Remark: Unlike the intersection of subspaces, the union of two subspaces will not generally be a subspace.
- 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.

1.3 Linear Combinations and Span

- Definition: Given a set $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of vectors in a vector space V, we say a vector $\mathbf{w}$ in V is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ 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 finite set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in V, denoted $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, to be the set of all vectors which are linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. In other words, the span is the set of vectors of the form $a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n$, for scalars $a_1, \ldots, a_n$. For an infinite collection of vectors, we define the span to be the set of all linear combinations of finitely many of the vectors.
  - Note: 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: The span of the polynomials $\{1, x, x^2, x^3\}$ is the set of polynomials of degree at most 3.

- Example: Determine whether the vectors $\langle 2,3,3\rangle$ and $\langle 4,-1,3\rangle$ are in $\text{span}(\mathbf{v},\mathbf{w})$, 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 $2 = a + 2b$, $3 = -a + b$, $3 = 2a - b$.
  - Adding the first two equations yields $5 = 3b$ so that $b = 5/3$. The second equation then yields $a = -4/3$. However, this does not satisfy the third equation. So there are no such a and b, meaning that $\boxed{\langle 2,3,3\rangle \text{ is not in the span}}$.
  - Similarly, for $\langle 4,-1,3\rangle$ we want to solve $\langle 4,-1,3\rangle = c\cdot\langle 1,-1,2\rangle + d\cdot\langle 2,1,-1\rangle$, or $4 = c + 2d$, $-1 = -c + d$, $3 = 2c - d$.
  - Using a similar procedure as above shows that $d = 1$, $c = 2$ is a solution: thus, we have $\langle 4,-1,3\rangle = 2\cdot\langle 1,-1,2\rangle + 1\cdot\langle 2,1,-1\rangle$, meaning 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 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in V, if W is any subspace of V that contains $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then W contains $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. In other words, the span is the smallest subspace containing the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
  - Proof: Consider any element $\mathbf{w}$ in $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$: 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 span(S) is a subspace of span(T).
  - Proof: Since the span is always a subspace, we know that span(T) is a subspace of V containing all the vectors in S. By the previous proposition, span(T) therefore contains every linear combination of vectors from S, which is to say, span(T) contains span(S).

- Here are two results showing how span interacts with adjoining a new vector to a set:

- Proposition (Span and Unions 1): If S is any set of vectors in V and $T = S \cup \{\mathbf{w}\}$ for some vector $\mathbf{w}$ in V, then span$(T) =$ span(S) if and only if $\mathbf{w}$ is in span(S).
  - Proof: By the previous corollary, since T contains S, span(T) contains span(S).
  - If span$(T) =$ span(S), then since $\mathbf{w}$ is in T (hence in span(T)) we conclude $\mathbf{w}$ is in span(S).
  - Conversely, if $\mathbf{w}$ is in span(S), then we can eliminate $\mathbf{w}$ from any linear combination of vectors in T to obtain a linear combination of vectors only in S.
  - Explicitly: suppose $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$ where $\mathbf{v}_1, \dots, \mathbf{v}_n$ are in S. Then any $\mathbf{x}$ in span(T) is some linear combination $\mathbf{x} = c \cdot \mathbf{w} + b_1 \cdot \mathbf{v}_1 + \cdots + b_n \cdot \mathbf{v}_n + b_{n+1} \cdot \mathbf{v}_{n+1} + \cdots + b_m \cdot \mathbf{v}_m$ for some $\mathbf{v}_1, \dots, \mathbf{v}_m$ in S.
  - But then $\mathbf{x} = (b_1 + ca_1) \cdot \mathbf{v}_1 + \cdots + (b_n + ca_n) \cdot \mathbf{v}_n + b_{n+1} \cdot \mathbf{v}_{n+1} + \cdots + b_m \cdot \mathbf{v}_m$ can be written as a linear combination only involving vectors in S, so $\mathbf{x}$ is in span(S). Thus, span$(S) =$ span(T).

- Proposition (Span and Unions 2): If S and T are any subsets of V with span$(S) =$ span(T) and $\mathbf{w}$ is any vector in V, then span$(S \cup \{\mathbf{w}\}) =$ span$(T \cup \{\mathbf{w}\})$.
  - Proof: By hypothesis, every vector in S lies in span$(S) =$ span(T), and since span(T) is contained in span$(T \cup \{\mathbf{w}\})$, every vector in S is contained in span$(T \cup \{\mathbf{w}\})$.
  - Since $\mathbf{w}$ is also in span$(T \cup \{\mathbf{w}\})$, $S \cup \{\mathbf{w}\}$ is in span$(T \cup \{\mathbf{w}\})$. But now since span$(T \cup \{\mathbf{w}\})$ is a subspace of V, it contains span$(S \cup \{\mathbf{w}\})$.
  - Therefore, span$(S \cup \{\mathbf{w}\}) \subseteq$ span$(T \cup \{\mathbf{w}\})$. By the same argument with S and T interchanged, span$(T \cup \{\mathbf{w}\}) \subseteq$ span$(S \cup \{\mathbf{w}\})$. Therefore, we must have equality: span$(S \cup \{\mathbf{w}\}) =$ span$(T \cup \{\mathbf{w}\})$.

- Sets whose span is the entire space have a special name:

- Definition: Given a set S of vectors in a vector space V, if span$(S) = V$ then we say that S is a spanning set (or generating set) for 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 $a_1, \dots, a_n$ and some vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in S.

- 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.

- <u>Example</u>: Show that the matrices $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ also span the vector space of 2×2 matrices of trace zero.
    - We can write $\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + (c-a) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.
    - This set of matrices is different from the spanning set in the previous example, which underlines an important point: any given vector space may have many different spanning sets.

- <u>Example</u>: Determine whether the polynomials 1, $1 + x^2$, x^4, $1 + x^2 + x^4$ span the space W of polynomials with complex coefficients having degree at most 4 and satisfying $p(x) = p(-x)$.
    - It is straightforward to verify that this set of polynomials is a subspace of the polynomials with complex coefficients.
    - 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 complex numbers 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 illustrates 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.

- <u>Example</u>: 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 $\boxed{\text{do span } \mathbb{R}^2}$.

- <u>Example</u>: 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

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

    - Now, if $r + 2q - p \neq 0$, the last row will be a contradictory equation. 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 $\boxed{\text{do not span } \mathbb{R}^3}$.

1.4 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.

- 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 $\boxed{\text{linearly dependent}}$.

- 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 (Independence of Two Vectors): 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 $2\cdot\langle 1,0,2,2\rangle + (-1)\cdot\langle 2,-2,0,3\rangle + (-2)\cdot\langle 0,3,3,1\rangle + 1\cdot\langle 0,4,2,1\rangle = \langle 0,0,0,0\rangle$. This is a nontrivial linear combination giving the zero vector, so these vectors are $\boxed{\text{linearly dependent}}$.

- 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):

- Proposition (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.
    - Proof: 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.

- Example: 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 $a\cdot\langle 1,-1\rangle + b\cdot\langle 2,2\rangle + c\cdot\langle 2,1\rangle = \langle 0,0\rangle$.
    - By row-reducing the appropriate matrix we can find the solution $2\cdot\langle 1,-1\rangle + 3\cdot\langle 2,2\rangle - 4\cdot\langle 2,1\rangle = \langle 0,0\rangle$.
    - By rearranging we can then write $\boxed{\langle 1,-1\rangle = -\dfrac{3}{2}\cdot\langle 2,2\rangle + 2\cdot\langle 2,1\rangle}$. (Of course, this is not the only possible answer: any of the vectors can be written as a linear combination of the other two, since all of the coefficients in the linear dependence are nonzero.)

- Linear independence and span are related in a number of ways, such as the following:

- Theorem (Independence and Span): Let S be a linearly independent subset of the vector space V, and $\mathbf{v}$ be any vector of V not in S. Then the set $S\cup\{\mathbf{v}\}$ is linearly dependent if and only if $\mathbf{v}$ is in $\text{span}(S)$.

    - Proof: 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 $\mathbf{v}$ is zero, then we would obtain a nontrivial dependence among the vectors in S (impossible, since S is linearly independent), so 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 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S.
- Then $\mathbf{v} = (-\frac{b_1}{a}) \cdot \mathbf{v}_1 + \cdots + (-\frac{b_n}{a}) \cdot \mathbf{v}_n$ is a linear combination of the vectors in S, so $\mathbf{v}$ is in $\text{span}(S)$.

- We can also characterize linear independence using the span:
- Theorem (Characterization of Linear Independence): A set S of vectors is linearly independent if and only if every vector $\mathbf{w}$ in $\text{span}(S)$ may be *uniquely* written as a sum $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$ for unique scalars $a_1, a_2, \ldots, a_n$ and unique vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S.
  - Proof: First suppose the decomposition is always unique: then for any $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S, $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$ 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 we see that 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 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 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 $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_n = b_n$, which is to say that the two decompositions are actually the same.

1.5 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.

1.5.1 Definition and Basic Properties of Bases

- Definition: A linearly independent set of vectors that 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 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 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 $a_1, a_2, \dots, a_n$.
  - In particular, this proposition says that if we have a basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ for V, then we can describe all of the other vectors in V in a particularly simple way (as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$) that is *unique*. A useful way to interpret this idea is to think of the basis vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ as "coordinate directions" and the coefficients $a_1, a_2, \dots, a_n$ as "coordinates".
  - Proof: Suppose $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is a basis of V. Then by definition, the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 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 $a_1, a_2, \dots, a_n$.
  - Furthermore, since the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ 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 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ span V.
  - Furthermore, by our earlier proposition, because every vector in $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ can be uniquely written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are linearly independent: thus, they are a linearly independent spanning set for V, so they form a basis.

- If we have a good 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 real polynomials $p(x)$ of degree ≤ 3 such that $p(1) = 0$.
  - Notice that W is a subspace of the vector space V of all polynomials with real coefficients, 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.

1.5.2 Existence and Construction of Bases

- A basis for a vector space can be obtained from a spanning set:
- Theorem (Spanning Sets and Bases): If V is a vector space, then any set spanning 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[2]), so we will only treat the case of a finite spanning set consisting of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.
  - Proof (finite spanning set case): Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 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 $1 \leq k \leq n$, perform the following procedure:
    1. 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.)
    2. 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 show that S_n is linearly independent, we use induction on k to show that S_k is linearly independent for each $0 \leq k \leq n$.
    - For the base case we take $k = 0$: clearly, S_0 (the empty set) is linearly independent.
    - For the inductive step, suppose $k \geq 1$ and that S_{k-1} is linearly independent.
    - If $\mathbf{v}_k$ is in the span of S_{k-1}, then $S_k = S_{k-1}$ is linearly independent.
    - If $\mathbf{v}_k$ is not in the span of S_{k-1}, then $S_k = S_{k-1} \cup \{\mathbf{v}_k\}$ is linearly independent by our proposition about span and linear independence.
    - In both cases, S_k is linearly independent, so by induction, S_n is linearly independent.
  - To show that S_n spans V, let $T_k = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. We use induction on k to show that $\text{span}(S_k) = \text{span}(T_k)$ for each $0 \leq k \leq n$.
    - For the base case we take $k = 0$: clearly, $\text{span}(S_0) = \text{span}(T_0)$, since both S_0 and T_0 are empty.
    - For the inductive step, suppose $k \geq 1$ and that $\text{span}(S_{k-1}) = \text{span}(T_{k-1})$.
    - If $\mathbf{v}_k$ is in the span of S_{k-1}, then $S_{k-1} = S_k$ so $\text{span}(S_{k-1}) = \text{span}(S_k)$. By the inductive hypothesis, the span of S_{k-1} is the same as the span of T_{k-1}, so $\mathbf{v}_k$ is in the span of T_{k-1}. But now by one of our propositions about span, we see that $\text{span}(T_k) = \text{span}(T_{k-1})$, so $\text{span}(T_k) = \text{span}(S_k)$ as claimed.
    - If $\mathbf{v}_k$ is not in the span of S_{k-1}, then by our proposition on spans and adjoining a vector, since $\text{span}(S_{k-1}) = \text{span}(T_{k-1})$, we have $\text{span}(S_{k-1} \cup \{\mathbf{v}_k\}) = \text{span}(S_{k-1} \cup \{\mathbf{v}_k\})$, which is the same as saying $\text{span}(S_k) = \text{span}(T_k)$.
    - In both cases, $\text{span}(S_k) = \text{span}(T_k)$, so by induction, $\text{span}(S_n) = \text{span}(T_n) = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$.
- 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.

[2] It has been proven that the statement "every vector space has a basis" is actually equivalent to the axiom of choice (under the Zermelo-Frankel axioms of set theory), so in fact appealing to the axiom of choice, or equivalently Zorn's lemma, is necessary here!

- 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 (Replacement Theorem): Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for V and $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is a linearly independent subset of V. Then there is a reordering of the basis S, say $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ such that for each $1 \le k \le m$, the set $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n\}$ is a basis for V. Equivalently, the elements $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ can be used to successively replace the elements of the basis, with each replacement remaining a basis of V.
  - Proof: We prove the result by induction on k. For the base case $k = 0$, there is nothing to prove.
  - For the inductive step, suppose that $B_k = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n\}$ is a basis for V: we must show that we can remove one of the vectors $\mathbf{a}_i$ and reorder the others to produce a basis $B_{k+1} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n\}$ for V.
  - By hypothesis, since B_k spans V, we can write $\mathbf{w}_{k+1} = c_1 \cdot \mathbf{w}_1 + \cdots + c_k \cdot \mathbf{w}_k + d_{k+1} \cdot \mathbf{a}_{k+1} + \cdots + d_n \cdot \mathbf{a}_n$ for some scalars c_i and d_i.
  - If all of the d_i were zero, then $\mathbf{w}_{k+1}$ would be a linear combination of $\mathbf{w}_1, \dots, \mathbf{w}_k$, contradicting the assumption that $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is a linearly independent set of vectors.
  - Thus, at least one d_i is not zero. Rearrange the vectors $\mathbf{a}_i$ so that $d_{k+1} \ne 0$: then $\mathbf{w}_{k+1} = c_1 \cdot \mathbf{w}_1 + \cdots + c_k \cdot \mathbf{w}_k + d'_{k+1} \cdot \mathbf{a}'_{k+1} + \cdots + d'_n \cdot \mathbf{a}'_n$.
  - We claim now that $B_{k+1} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n\}$ is a basis for V.
  - To see that B_{k+1} spans V, since $d_{k+1} \ne 0$, we can solve for $\mathbf{a}'_{k+1}$ as a linear combination of the vectors $\mathbf{w}_1, \dots, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n$. (The exact expression is cumbersome, and the only fact we require is to note that the coefficient of $\mathbf{w}_{k+1}$ is not zero.)
  - If $\mathbf{x}$ is any vector in V, since B_k spans V we can write $\mathbf{x} = e_1 \cdot \mathbf{w}_1 + \cdots + e_k \cdot \mathbf{w}_k + e_{k+1} \cdot \mathbf{a}'_{k+1} + \cdots + e_n \cdot \mathbf{a}'_n$.
  - Plugging in the expression for $\mathbf{a}'_{k+1}$ in terms of $\mathbf{w}_1, \dots, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n$ then shows that $\mathbf{x}$ is a linear combination of $\mathbf{w}_1, \dots, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n$.
  - To see that B_{k+1} is linearly independent, suppose we had a dependence $\mathbf{0} = f_1 \cdot \mathbf{w}_1 + \cdots + f_k \cdot \mathbf{w}_k + f_{k+1} \cdot \mathbf{a}'_{k+1} + \cdots + f_n \cdot \mathbf{a}'_n$.
  - Now plug in the expression for $\mathbf{a}'_{k+1}$ in terms of $\mathbf{w}_1, \dots, \mathbf{w}_{k+1}, \mathbf{a}'_{k+2}, \dots, \mathbf{a}'_n$: all of the coefficients must be zero because B_k is linearly independent. But the coefficient of $\mathbf{w}_{k+1}$ is f_{k+1} times a nonzero scalar, so $f_{k+1} = 0$.
  - But this implies $\mathbf{0} = f_1 \cdot \mathbf{w}_1 + \cdots + f_k \cdot \mathbf{w}_k + f_{k+2} \cdot \mathbf{a}'_{k+2} + \cdots + f_n \cdot \mathbf{a}'_n$, and this is a dependence involving the vectors in B_k. Since B_k is (again) linearly independent, all of the coefficients are zero. Thus $f_1 = f_2 = \cdots = f_n = 0$, and so B_{k+1} is linearly independent.
  - Finally, since we have shown B_{k+1} is linearly independent and spans V, it is a basis for V. By induction, we have the desired result for all $1 \le k \le m$.

- Although the proof of the Replacement Theorem is cumbersome, we obtain several useful corollaries.

- Corollary: Suppose V has a basis with n elements. If $m > n$, then any set of m vectors of V is linearly dependent.
  - Proof: Suppose otherwise, so that $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is a linearly independent subset of V.
  - Apply the Replacement Theorem with the given basis of V: at the nth step we have replaced all the elements of the original basis with those in our new set, so by the conclusion of the theorem we see that $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ is a basis for V.

- Then $\mathbf{w}_{n+1}$ is necessarily a linear combination of $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$, meaning that $\{\mathbf{w}_1, \dots, \mathbf{w}_n, \mathbf{w}_{n+1}\}$ is linearly dependent. Thus $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ is linearly dependent.

- Corollary: Any two bases of a vector space have the same number of elements.
  - Proof: If every basis is infinite, we are already done, so now suppose that V has some finite basis, and choose B to be a basis of minimal size[3].
  - Suppose B has n elements, and consider any other basis B' of V. By the previous corollary, if B' contains more than n vectors, it would be linearly dependent (impossible). Thus, B' also has n elements, so every basis of V has n elements.

- 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 (finite basis case): Let S be a set of linearly independent vectors and let B be any basis of V (we have already shown that V has a basis). Apply the Replacement Theorem to B and S: this produces a new basis of V containing S.
  - Remark: Although we appealed to the Replacement Theorem here, we can also give a slightly different, more constructive argument like the one we gave for obtaining a basis from a spanning set.
    1. Start with a linearly independent set S of vectors in V. If S spans V, then we are done.
    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 this procedure will always terminate 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.

1.5.3 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 $\dim(V)$. If the dimension of V is a finite number, we say that V is finite-dimensional; otherwise, we say V is infinite-dimensional.
  - 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 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 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 zero space is 0, because the empty set (containing 0 elements) is a basis.
  - Example: Over the field of real numbers, the vector space of complex numbers has dimension 2, since the set $\{1, i\}$ forms a basis.
  - Example: Over the field of complex numbers, the vector space of complex numbers has dimension 1, since the set $\{1\}$ forms a basis.

- As the last two examples indicate, the dimension of a vector space depends on the field we are using.

[3] The size of a basis is either a nonnegative integer or ∞. The fact that a basis of smallest size must exist follows from an axiom (the well-ordering principle) that any nonempty set of nonnegative integers has a smallest element.

- To avoid ambiguities, it is best to indicate which field we are using when we discuss dimensions: we often do this by writing a subscript to indicate the field, so that $\dim_F V$ denotes the dimension of V as a vector space over the field F.
- Thus, $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ while $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.
- When the field is implied by context (or not relevant to the discussion) we will omit it.

- Here are a few basic properties of dimension that follow from our previous results:
- Proposition (Properties of Dimension): Suppose V and W are vector spaces. Then the following hold:
  1. 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.
  2. If $\dim(V) = n$, then any linearly independent set of vectors has at most n elements.
     - Proof: This result was a corollary to the Replacement Theorem.
  3. If $\dim(V) = n$, then any linearly independent set of n vectors is a basis for V.
     - Proof: This follows immediately from the Replacement Theorem.
  4. If $\dim(V) = n$, then any spanning set of V has at least n elements.
     - Proof: As we showed, any spanning set contains a basis.
  5. If $\dim(V) = n$, then any spanning set of V having exactly n elements is a basis for V.
     - Proof: The spanning set contains a basis, but since the basis must also have n elements, the basis is the entire spanning set.
  6. If $\dim(V) = n$, a subset of V having exactly n vectors is a basis if and only if it spans V if and only if it is linearly independent.
     - Proof: This follows by combining the results of (3) and (5) above.

- The simplest way to find the dimension of a vector space is to write down an explicit basis:
- Example: Find the dimension of the complex vector space 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 $h = f$.
  - Thus, $A = \begin{bmatrix} 0 & b & c \\ -b & 0 & f \\ -c & -f & 0 \end{bmatrix} = b \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c \cdot \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + f \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.
  - 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, so the dimension is $\boxed{3}$.

- 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.

1.5.4 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: We showed all of these results above.
- 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 is a subspace. It is also easy to see 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 $a_1 \cdot \mathbf{r}_1 + \cdots + a_r \cdot \mathbf{r}_r = \mathbf{0}$ 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 $a_1 = \cdots = a_r = 0$.
  - 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 $M\mathbf{x} = a_1 \cdot \mathbf{c}_1 + \cdots + a_n \cdot \mathbf{c}_n$ and $E\mathbf{x} = a_1 \cdot \mathbf{e}_1 + \cdots + a_n \cdot \mathbf{e}_n$ 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 $a_1 \cdot \mathbf{c}_1 + \cdots + a_n \cdot \mathbf{c}_n = \mathbf{0}$ if and only if $a_1 \cdot \mathbf{e}_1 + \cdots + a_n \cdot \mathbf{e}_n = \mathbf{0}$.
  - 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.
- Example: Find bases for the row space, column space, and nullspace of $M = \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ -2 & 2 & 0 & -3 & 1 \\ 1 & -1 & 0 & 3 & 8 \end{bmatrix}$.

- First, row-reduce M: $\begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ -2 & 2 & 0 & -3 & 1 \\ 1 & -1 & 0 & 3 & 8 \end{bmatrix} \xrightarrow[R_3-R_1]{R_2+2R_1} \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 2 & 7 \end{bmatrix} \xrightarrow{R_3-2R_2} \begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.
- The row space has a basis given by the rows $\boxed{\langle 1,-1,0,2,1\rangle, \langle 0,0,0,1,3\rangle, \langle 0,0,0,0,1\rangle}$.
- Since there are pivots in columns 1, 4, and 5, the column space has a basis $\boxed{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}}$.
- For the nullspace, solving the linear system $M\mathbf{x} = \mathbf{0}$ (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 the nullspace has a basis $\boxed{\langle 1,1,0,0,0\rangle, \langle 0,0,1,0,0\rangle}$.

- 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).

- <u>Example</u>: If $S = \{\langle 1,0,1,2\rangle, \langle 3,0,3,6\rangle, \langle 2,1,2,1\rangle, \langle 3,1,3,3\rangle\}$, find a subset of S that is a basis for 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 span(S).

- <u>Example</u>: Extend the set $S = \{1-2x^2, 2+x\}$ to a basis for the vector space V of polynomials of degree ≤ 3 with real coefficients.
  - 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 coefficient 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 V.

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