Linear Algebra (part 3) : Linear Transformations (by Evan Dummit, 2021, v. 3.00)

Contents

3 Linear Transformations

In this chapter we will study linear transformations, which are structure-preserving maps between vector spaces. Such maps are a generalization of the idea of a linear function, and have many of the same properties as linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, whose properties we emphasize. We begin by studying linear transformations in general, with a discussion of some geometric aspects of linear transformations on the Euclidean plane $\mathbb{R}^2$, and then discuss two important subspaces associated to a linear transformation: the kernel and the image. Next, we study a special class of linear transformations known as isomorphisms, and prove the rather stunning result that, in a very precise sense, any finite-dimensional vector space has the same structure as the vector space $\mathbb{R}^n$.

We then explore the various relationships between linear transformations and matrices, and use our understanding of bases to give a concrete "matrix representation" of a linear transformation in the finite-dimensional case. We will study linear transformations using matrices (and vice versa), and in particular discuss the idea of a change of basis in a vector space, and its relation to similarity of matrices.

3.1 Linear Transformations

- Now that we have a reasonably good idea of what the structure of a vector space is, the next natural question is: what do maps from one vector space to another look like?
- It turns out that we don't want to ask about arbitrary functions, but about functions from one vector space to another that preserve the structure (namely, addition and scalar multiplication) of the vector space.

3.1.1 Definition, Examples, and Properties

- Definition: If V and W are vector spaces, we say a function T from V to W (denoted $T : V \to W$) is a linear transformation if the following two properties hold:

  [T1] The map respects addition of vectors: $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ for any vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V.

  [T2] The map respects scalar multiplication: $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$ for any vector $\mathbf{v}$ in V and any scalar α.

- Warning: It is important to note that in the statement $T(\mathbf{v}_1+\mathbf{v}_2) = T(\mathbf{v}_1)+T(\mathbf{v}_2)$, the addition on the left-hand side is taking place inside V, whereas the addition on the right-hand side is taking place inside W. Likewise, in the statement $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$, the scalar multiplication on the left-hand side is in V while the scalar multiplication on the right-hand side is in W.

- Example: If $V = W = \mathbb{R}^2$, show that the map T defined[1] by $T(x, y) = \langle x, x+y\rangle$ is a linear transformation from V to W .
  - We simply check the two parts of the definition.
  - Let $\mathbf{v} = \langle x, y\rangle$, $\mathbf{v}_1 = \langle x_1, y_1\rangle$, and $\mathbf{v}_2 = \langle x_2, y_2\rangle$, so that $\mathbf{v}_1 + \mathbf{v}_2 = \langle x_1 + x_2, y_1 + y_2\rangle$.
  - [T1]: We have $T(\mathbf{v}_1+\mathbf{v}_2) = \langle x_1 + x_2, x_1 + x_2 + y_1 + y_2\rangle = \langle x_1, x_1 + y_1\rangle+\langle x_2, x_2 + y_2\rangle = T(\mathbf{v}_1)+T(\mathbf{v}_2)$.
  - [T2]: We have $T(\alpha\mathbf{v}) = \langle \alpha x, \alpha x + \alpha y\rangle = \alpha \langle x, x + y\rangle = \alpha T(\mathbf{v})$.

- We can substantially generalize the example above:

- Example: If $V = \mathbb{R}^n$ (thought of as $n \times 1$ matrices) and $W = \mathbb{R}^m$ (thought of as $m \times 1$ matrices) and A is any $m \times n$ matrix, show that the map T with $T(\mathbf{v}) = A\mathbf{v}$ is a linear transformation.
  - The verification is exactly the same as in the previous example.
  - [T1]: We have $T(\mathbf{v}_1 + \mathbf{v}_2) = A(\mathbf{v}_1 + \mathbf{v}_2) = A\mathbf{v}_1 + A\mathbf{v}_2 = T(\mathbf{v}_1) + T(\mathbf{v}_2)$.
  - [T2]: Also, $T(\alpha\mathbf{v}) = A(\alpha\mathbf{v}) = \alpha(A\mathbf{v}) = \alpha T(\mathbf{v})$.

- Example: If $V = M_{2\times2}(\mathbb{R})$ and $W = \mathbb{R}$, determine whether the trace map is a linear transformation from V to W .
  - Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $M_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$, $M_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$ so $M_1 + M_2 = \begin{bmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{bmatrix}$.
  - [T1]: We have $\text{tr}(M_1 + M_2) = (a_1 + a_2) + (d_1 + d_2) = (a_1 + d_1) + (a_2 + d_2) = \text{tr}(M_1) + \text{tr}(M_2)$.
  - [T2]: We have $\text{tr}(\alpha \cdot M) = \alpha a + \alpha d = \alpha \cdot (a + d) = \alpha \cdot \text{tr}(M)$.
  - Both parts of the definition are satisfied, so the trace $\boxed{\text{is a linear transformation}}$.

- Example: If $V = M_{2\times2}(\mathbb{R})$ and $W = \mathbb{R}$, determine whether the determinant map $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ is a linear transformation from V to W.
  - Let $M_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$, $M_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix}$ so $M_1 + M_2 = \begin{bmatrix} a_1+a_2 & b_1+b_2 \\ c_1+c_2 & d_1+d_2 \end{bmatrix}$.
  - [T1]: We have $\det(M_1 + M_2) = (a_1 + a_2)(d_1 + d_2) - (b_1 + b_2)(c_1 + c_2)$, while $\det(M_1) + \det(M_2) = (a_1d_1 - b_1c_1) + (a_2d_2 - b_2c_2)$.
  - When we expand out the products in $\det(M_1 + M_2)$ we will quickly see that the expression is not the same as $\det(M_1) + \det(M_2)$.
  - An explicit example is $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $M_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$: $\det(M_1) = \det(M_2) = 0$, while $M_1 + M_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ has determinant 1.
  - The first part of the definition does not hold, so this function $\boxed{\text{is not a linear transformation}}$. (In fact, the condition [T2] fails as well.)

- Example: If $V = P_4(\mathbb{R})$ and $W = \mathbb{R}^2$, determine whether the map $T(p) = \langle p(1), p'(1)\rangle$ is a linear transformation from V to W.
  - [T1]: We have $T(p_1+p_2) = \langle (p_1 + p_2)(1), (p_1 + p_2)'(1)\rangle = \langle p_1(1), p_1'(1)\rangle+\langle p_2(1), p_2'(1)\rangle = T(p_1)+T(p_2)$
  - [T2]: We have $T(\alpha p) = \langle (\alpha p)(1), (\alpha p)'(1)\rangle = \langle \alpha p(1), \alpha p'(1)\rangle = \alpha \langle p(1), p'(1)\rangle = \alpha T(p)$.

[1] In principle here we should actually write $T(\langle x, y\rangle) = \langle x, x + y\rangle$, but this notation looks rather ugly, so we will suppress the vector brackets inside the function notation when writing linear transformations on vectors in $\mathbb{R}^n$.

- Both parts of the definition hold, so this function <u>is a linear transformation</u>.

- Here are a few additional examples of linear transformations:
  - If V is the vector space of differentiable functions and W is the vector space of real-valued functions, the derivative map D sending a function to its derivative is a linear transformation from V to W.
  - If V is the vector space of all continuous functions on $[a, b]$, then the map $T(f) = \int_a^b f(x)\,dx$ is a linear transformation from V to $\mathbb{R}$.
  - The transpose map is a linear transformation from $M_{m\times n}(\mathbb{F})$ to $M_{n\times m}(\mathbb{F})$ for any field $\mathbb{F}$ and any positive integers m, n.
  - If V and W are any vector spaces, the <u>zero map</u> sending all elements of V to the zero vector in W is a linear transformation from V to W.
  - If V is any vector space, the <u>identity map</u> sending all elements of V to themselves is a linear transformation from V to V.

- Here are a few simple algebraic properties of linear transformations:

- <u>Proposition</u> (Properties of Linear Transformations): Suppose $T : V \to W$ is a linear transformation. Then the following hold:
  1. The linear transformation $T : V \to W$ sends the zero vector of V to the zero vector of W.
     - <u>Proof</u>: Let $\mathbf{v}$ be any vector in V. Since $0\mathbf{v} = \mathbf{0}_V$ from basic properties, applying [T2] yields $0T(\mathbf{v}) = T(\mathbf{0}_V)$.
     - But $0T(\mathbf{v}) = \mathbf{0}_W$ since scaling any vector of W by 0 gives the zero vector of W.
     - Combining these two statements gives $T(\mathbf{0}_V) = 0T(\mathbf{v}) = \mathbf{0}_W$, so $T(\mathbf{0}_V) = \mathbf{0}_W$ as claimed.
  2. For any vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ and any scalars $a_1, \dots, a_n$, we have $T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1T(\mathbf{v}_1) + \cdots + a_nT(\mathbf{v}_n)$.
     - This result says that linear transformations can be moved through linear combinations.
     - <u>Proof</u>: By a trivial induction using [T1], we see that $T(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n) = T(a_1\mathbf{v}_1)+\cdots+T(a_n\mathbf{v}_n)$.
     - Then by [T2], we have $T(a_i\mathbf{v}_i) = a_iT(\mathbf{v}_i)$ for each $1 \le i \le n$.
     - Plugging these relations into the first equation gives $T(a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n) = a_1T(\mathbf{v}_1)+\cdots+a_nT(\mathbf{v}_n)$ as required.
  3. The map $T : V \to W$ is a linear transformation if and only if for any $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and any scalar α, $T(\mathbf{v}_1 + \alpha\mathbf{v}_2) = T(\mathbf{v}_1) + \alpha T(\mathbf{v}_2)$.
     - <u>Proof</u>: If T is linear, then by [T1] and [T2], $T(\mathbf{v}_1 + \alpha\mathbf{v}_2) = T(\mathbf{v}_1) + T(\alpha\mathbf{v}_2) = T(\mathbf{v}_1) + \alpha T(\mathbf{v}_2)$.
     - Conversely, suppose that $T(\mathbf{v}_1 + \alpha\mathbf{v}_2) = T(\mathbf{v}_1) + \alpha T(\mathbf{v}_2)$. Setting $\alpha = 1$ produces $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ so T satisfies [T1].
     - Then taking $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{0}$ and $c = \alpha$ yields $T(\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, so $T(\mathbf{0}) = \mathbf{0}$.
     - Finally, setting $\mathbf{v}_1 = \mathbf{0}$ yields $T(\alpha\mathbf{v}_2) = T(\mathbf{0}) + \alpha T(\mathbf{v}_2) = \alpha T(\mathbf{v}_2)$ so T satisfies [T2].

- A linear transformation is completely determined by its values on a basis:

- <u>Theorem</u> (Linear Transformations and Bases): Any linear transformation $T : V \to W$ is characterized by its values on a basis of V. Conversely, for any basis $B = \{\mathbf{v}_i\}$ of V and any vectors $\{\mathbf{w}_i\}$ in W, there exists a unique linear transformation $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i.
  - This theorem holds even if the sets are infinite. However, since the notation is cumbersome, we will give the proof only in the finite case.
  - <u>Proof</u> (Finite Case): For the first statement, let $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis of V. Then any vector $\mathbf{v}$ in V can be written as $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ for unique scalars $a_1, \dots, a_n$.
  - By (2) in the previous proposition, $T(\mathbf{v}) = a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \cdots + a_nT(\mathbf{v}_n)$, so the value of T is determined by the values $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$, ... , $T(\mathbf{v}_n)$.

- Conversely, suppose that we are given the values $T(\mathbf{v}_i) = \mathbf{w}_i$ for each $\mathbf{v}_i$ in B. We claim that the map $T : V \to W$ defined by setting $T(b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n) = b_1\mathbf{w}_1 + b_2\mathbf{w}_2 + \cdots + b_n\mathbf{w}_n$ is a well-defined linear transformation from V to W.
- Notice that every vector in V can be written in exactly one way as a linear combination of the basis vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, so there is no ambiguity about how T is defined on any vector in V.
- For [T1], if $\mathbf{x} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$ and $\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, then $T(\mathbf{x}+\mathbf{y}) = (b_1+c_1)\mathbf{w}_1 + \cdots + (b_n+c_n)\mathbf{w}_n = T(\mathbf{x}) + T(\mathbf{y})$.
- For [T2], if $\mathbf{x} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$ then $T(\alpha\mathbf{x}) = \alpha b_1\mathbf{v}_1 + \cdots + \alpha b_n\mathbf{w}_n = \alpha T(\mathbf{x})$.
- Finally, suppose that there were some other linear transformation S with $S(\mathbf{v}_i) = \mathbf{w}_i = T(\mathbf{v}_i)$ for each i.
- If $\mathbf{v}$ is any vector in V, we can write $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$, and then by our previous proposition, we see $S(\mathbf{v}) = S(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) = a_1 S(\mathbf{v}_1) + \cdots + a_n S(\mathbf{v}_n) = a_1 T(\mathbf{v}_1) + \cdots + a_n T(\mathbf{v}_n) = T(\mathbf{v})$. Thus, S and T are the same.

- Using the result above, we can "reconstruct" the entirety of a linear transformation given its values on a basis.

- Example: If V is the vector space of polynomials of degree ≤ 2 and $T : V \to \mathbb{R}$ is the linear transformation such that $T(1) = 5$, $T(1+x) = 4$, and $T(2+x^2) = 3$, find $T(5+2x+2x^2)$.
  - We simply need to express $5 + 2x + 2x^2$ in terms of the basis $\{1, 1+x, 2+x^2\}$ for V.
  - A straightforward calculation shows $5 + 2x + 2x^2 = -1(1) + 2(1+x) + 2(2+x^2)$.
  - Thus, $T(5+2x+2x^2) = -T(1) + 2T(1+x) + 2T(2+x^2) = -1(5) + 2(4) + 2(3) = \boxed{9}$.

- As a final remark, we observe that we can do many algebraic operations with linear transformations. For example, the sum of two linear transformations is also a linear transformation, as is any scalar multiple of a linear transformation. Indeed, these facts imply that the collection of linear transformations from V to W itself has a vector space structure:

- Theorem (Space of Linear Transformations): Let V and W be vector spaces. Then the set $\mathcal{L}(V, W)$ of all linear transformations from V to W is a subspace of the space of functions from V to W.
  - Proof: We verify the subspace criterion.
  - [S1]: The zero map is a linear transformation.
  - [S2]: Suppose that T_1 and T_2 be linear transformations: we must show that $T_1 + T_2$ is also a linear transformation. This follows from the observations that

$$\begin{aligned}(T_1+T_2)(\mathbf{v}_1+\mathbf{v}_2) &= T_1(\mathbf{v}_1+\mathbf{v}_2) + T_2(\mathbf{v}_1+\mathbf{v}_2) = [T_1(\mathbf{v}_1) + T_1(\mathbf{v}_2)] + [T_2(\mathbf{v}_1) + T_2(\mathbf{v}_2)] \\ &= [T_1(\mathbf{v}_1) + T_2(\mathbf{v}_1)] + [T_1(\mathbf{v}_2) + T_2(\mathbf{v}_2)] = (T_1+T_2)(\mathbf{v}_1) + (T_1+T_2)(\mathbf{v}_2)\end{aligned}$$

    and that

$$(T_1+T_2)(\alpha\mathbf{v}) = T_1(\alpha\mathbf{v}) + T_2(\alpha\mathbf{v}) = \alpha T_1(\mathbf{v}) + \alpha T_2(\mathbf{v}) = \alpha[(T_1+T_2)(\mathbf{v})].$$

  - [S3]: Suppose that T is a linear transformation: we must show that cT is also a linear transformation. This follows from the observations that

$$(cT)(\mathbf{v}_1+\mathbf{v}_2) = cT(\mathbf{v}_1+\mathbf{v}_2) = c\,[T(\mathbf{v}_1) + T(\mathbf{v}_2)] = cT(\mathbf{v}_1) + cT(\mathbf{v}_2)$$

    and

$$(cT)(\alpha\mathbf{v}) = cT(\alpha\mathbf{v}) = c[\alpha T(\mathbf{v})] = \alpha[cT(\mathbf{v})].$$

3.1.2 Kernel and Image

- We will now study a pair of important subspaces associated to a linear transformation.
- Definition: If $T : V \to W$ is a linear transformation, then the kernel of T, denoted $\ker(T)$, is the set of elements $\mathbf{v}$ in V with $T(\mathbf{v}) = \mathbf{0}$.
  - In the event that $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by a matrix A, then a vector $\mathbf{x}$ is in the kernel precisely when $A\mathbf{x} = \mathbf{0}$: in other words, the kernel of T is the nullspace of the matrix A.
  - Thus, we see that the kernel is a generalization of the nullspace to arbitrary linear transformations.
- Definition: If $T : V \to W$ is a linear transformation, then the image of T (often also called the range of T), denoted $\mathrm{im}(T)$, is the set of elements $\mathbf{w}$ in W such that there exists a $\mathbf{v}$ in V with $T(\mathbf{v}) = \mathbf{w}$.
  - The image is the elements in W which can be obtained as output from T. If $\mathrm{im}(T) = W$, we say T is onto (or surjective).
  - Even though they mean the same thing, we use the word "image" with linear transformations (rather than "range") to emphasize the additional structure that the image of a linear transformation possesses, relative to the range of a general function.
  - In the event that $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by a matrix A, then a vector $\mathbf{b}$ is in the image precisely when there is a solution $\mathbf{x}$ to the matrix equation $A\mathbf{x} = \mathbf{b}$: in other words, the image of T is the column space of the matrix A.
- Example: If $T : \mathbb{R}^3 \to \mathbb{R}^3$ has $T(x, y, z) = \langle x + y, z, x + y\rangle$, find the kernel and image of T.
  - For the kernel, we want to find all (x, y, z) such that $T(x, y, z) = \langle 0, 0, 0\rangle$, so we obtain the three equations $x + y = 0$, $z = 0$, $x + y = 0$. These equations collectively say $y = -x$, so we see that the kernel is the set of vectors of the form $\langle x, -x, 0\rangle$.
  - For the image, one possible answer is simply "the set of vectors of the form $\langle x + y, z, x + y\rangle$". A slightly more useful description would be "the vectors of the form $\langle a, b, a\rangle$" since the first and second coordinates can be arbitrary, but the third is always equal to the first.
- The kernel and image are subspaces of V and W respectively:
- Proposition: The kernel is a subspace of V.
  - [S1] We have $T(\mathbf{0}) = \mathbf{0}$, by simple properties of linear transformations.
  - [S2] If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in the kernel, then $T(\mathbf{v}_1) = \mathbf{0}$ and $T(\mathbf{v}_2) = \mathbf{0}$. Therefore, $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.
  - [S3] If $\mathbf{v}$ is in the kernel, then $T(\mathbf{v}) = \mathbf{0}$. Hence $T(\alpha \cdot \mathbf{v}) = \alpha \cdot T(\mathbf{v}) = \alpha \cdot \mathbf{0} = \mathbf{0}$.
- Proposition: The image is a subspace of W.
  - [S1] We have $T(\mathbf{0}) = \mathbf{0}$, by simple properties of linear transformations.
  - [S2] If $\mathbf{w}_1$ and $\mathbf{w}_2$ are in the image, then there exist $\mathbf{v}_1$ and $\mathbf{v}_2$ such that $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. Then $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$, so that $\mathbf{w}_1 + \mathbf{w}_2$ is also in the image.
  - [S3] If $\mathbf{w}$ is in the image, then there exists $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{w}$. Then $T(\alpha \cdot \mathbf{v}) = \alpha \cdot T(\mathbf{v}) = \alpha \cdot \mathbf{w}$, so $\alpha \cdot \mathbf{w}$ is also in the image.
- There is a straightforward way to find a spanning set for the image of a linear transformation:
- Proposition: If $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis for V and $T : V \to W$ is linear, then $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$ spans the image of T.
  - Note that in general the vectors $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$ are not necessarily a basis for the image since they need not be linearly independent. (But we have already discussed methods for converting a spanning set into a basis, so it is not hard to find an actual basis for the image.)

- Proof: Suppose $\mathbf{w}$ is in the image of T. Then by hypothesis, $\mathbf{w} = T(\mathbf{v})$ for some vector $\mathbf{v}$.
- Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis for V, there are scalars $a_1, \ldots, a_n$ such that $\mathbf{v} = a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n$.
- Then $\mathbf{w} = T(\mathbf{v}) = a_1 \cdot T(\mathbf{v}_1) + \cdots + a_n \cdot T(\mathbf{v}_n)$ is a linear combination of $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$, so it lies in their span. This is true for any $\mathbf{w}$ in the image of T, so $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)$ spans the image of T as claimed.
- Remark: It is natural to wonder whether there is an equally simple way to find a spanning set for the kernel of a general linear transformation: unfortunately, there is not. For matrix maps, however, the kernel is the same as the nullspace, so we can compute it using row reductions.

- Example: If $T : \mathbb{R}^2 \to \mathbb{R}^3$ is the linear transformation with $T(x, y) = (x + y, 0, 2x + 2y)$, find a basis for the kernel and for the image of T.
  - For the kernel, we want to find all (x, y, z) such that $T(x, y) = (0, 0, 0)$, so we obtain the three equations $x + y = 0$, $0 = 0$, $2x + 2y = 0$. These equations collectively say $y = -x$, so we see that the kernel is the set of vectors of the form $\langle x, -x\rangle = x \cdot \langle 1, -1\rangle$, so a basis for the kernel is given by the single vector $\boxed{\langle 1, -1\rangle}$.
  - For the image, by the proposition above it is enough simply to find the span of $T(\mathbf{v}_1)$, $T(\mathbf{v}_2)$ where $\mathbf{v}_1$ and $\mathbf{v}_2$ are a basis for $\mathbb{R}^2$. Using the standard basis, we compute $T(1, 0) = \langle 1, 0, 2\rangle$ and $T(0, 1) = \langle 1, 0, 2\rangle$, so a basis for the image is given by the single vector $\boxed{\langle 1, 0, 2\rangle}$.

- Example: If $T : P_2(\mathbb{R}) \to \mathbb{R}^2$ is the linear transformation with $T(p) = \langle p(1), p'(1)\rangle$, find a basis for the kernel and for the image of T.
  - Notice that $T(a + bx + cx^2) = \langle a + b + c, b + 2c\rangle$.
  - For the kernel, we want to find all p such that $T(p) = \langle 0, 0\rangle$, which is equivalent to requiring $a+b+c = 0$ and $b + 2c = 0$, so that $b = -2c$ and $a = c$. Thus, the kernel is the set of polynomials of the form $p(x) = c - 2cx + cx^2$, which is spanned by the polynomial $\boxed{1 - 2x + x^2}$.
  - The image is spanned by $T(1) = \langle 1, 0\rangle$, $T(x) = \langle 1, 1\rangle$, $T(x^2) = \langle 1, 2\rangle$. Since these vectors clearly span $\mathbb{R}^2$, we can take any basis for $\mathbb{R}^2$, such as $\boxed{\langle 1, 0\rangle, \langle 1, 1\rangle}$.

- The kernel of a linear transformation is closely tied to whether it is one-to-one:
  - A one-to-one linear transformation sends different vectors in V to different vectors in W. A one-to-one function of a real variable is one that passes the "vertical line test", and thus has an inverse function f^{-1}.

- Proposition (Kernel and One-to-One Maps): For any linear transformation $T : V \to W$, the kernel $\ker(T)$ consists of only the zero vector if and only if the map T is one-to-one: that is, if $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$.
  - Proof: If T is one-to-one, then (at most) one element of V maps to $\mathbf{0}$. But since the zero vector of V is taken to the zero vector of W, we see that T cannot send anything else to $\mathbf{0}$. Thus $\ker(T) = \{\mathbf{0}\}$.
  - Conversely, if $\ker(T)$ is only the zero vector, then since T is a linear transformation, the statement $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ is equivalent to the statement that $T(\mathbf{v}_1) - T(\mathbf{v}_2) = T(\mathbf{v}_1 - \mathbf{v}_2)$ is the zero vector.
  - But, by the definition of the kernel, $T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$ precisely when $\mathbf{v}_1 - \mathbf{v}_2$ is in the kernel. However, this means $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, so $\mathbf{v}_1 = \mathbf{v}_2$. Hence $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$, which means T is one-to-one.

- We can give some intuitive explanations for what the kernel and image are measuring.
  - The image of a linear transformation measures how close the map is to giving all of W as output: a linear transformation with a large image hits most of W, while a linear transformation with a small image misses most of W.
  - The kernel of a linear transformation measures how close the map is to being the zero map: a linear transformation with a large kernel sends many vectors to zero, while a linear transformation with a small kernel sends few vectors to zero.

- We can quantify these notions of "large" and "small" using dimension:

- Definitions: The dimension of $\ker(T)$ is called the nullity of T, and the dimension of $\text{im}(T)$ is called the rank of T.
  - A linear transformation with a large nullity has a large kernel, which means it sends many elements to zero (hence "nullity").

- There is a very important relationship between the rank and the nullity of a linear transformation:

- Theorem (Nullity-Rank): For any linear transformation $T : V \to W$, $\dim(\ker(T)) + \dim(\text{im}(T)) = \dim(V)$. In words, the nullity plus the rank is equal to the dimension of V.
  - We will prove this theorem in the case where V is finite-dimensional. (The proof in the infinite-dimensional case is essentially the same, but the notation tends to obscure the ideas more.)
  - Proof (Finite-dimensional case): Let $\mathbf{w}_1, \ldots, \mathbf{w}_k$ be a basis for $\text{im}(T)$ in W.
  - Then by the definition of the image, there exist $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in V such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each $1 \le i \le k$.
  - Also let $\mathbf{a}_1, \ldots, \mathbf{a}_l$ be a basis for $\ker(T)$. We claim that the set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{a}_1, \ldots, \mathbf{a}_l\}$ is a basis for V.
  - To see that S spans V, let $\mathbf{v}$ be an element of V.
    - Since $T(\mathbf{v})$ lies in $\text{im}(T)$, there exist scalars $b_1, \ldots, b_k$ such that $T(\mathbf{v}) = \sum_{j=1}^{k} b_j \mathbf{w}_j$.
    - By properties of linear transformations, we then can write

$$T\left(\mathbf{v} - \sum_{j=1}^{k} b_j \mathbf{v}_j\right) = T(\mathbf{v}) - \sum_{j=1}^{k} b_j T(\mathbf{v}_j) = \sum_{j=1}^{k} b_j \mathbf{w}_j - \sum_{j=1}^{k} b_j \mathbf{w}_j = \mathbf{0}.$$

    - Therefore, $\mathbf{v} - \sum_{j=1}^{k} b_j \mathbf{v}_j$ is in $\ker(T)$, so it can be written as a sum $\sum_{i=1}^{l} c_i \mathbf{a}_i$ for unique scalars c_i.
    - Thus, $\mathbf{v} = \sum_{j=1}^{k} b_j \mathbf{v}_j + \sum_{i=1}^{l} c_i \mathbf{a}_i$ for scalars b_j and c_i, so S spans V.
  - To see that S is linearly independent, suppose we had a dependence $\mathbf{0} = \sum_{j=1}^{k} b_j \mathbf{v}_j + \sum_{i=1}^{l} c_i \mathbf{a}_i$.
    - Applying T to both sides yields $\mathbf{0} = T(\mathbf{0}) = \sum_{j=1}^{k} b_j T(\mathbf{v}_j) + \sum_{i=1}^{l} c_i T(\mathbf{a}_i) = \sum_{j=1}^{k} b_j \mathbf{w}_j$.
    - Since the $\mathbf{w}_j$ are linearly independent, we conclude that all the coefficients b_j must be zero.
    - We then obtain the relation $\mathbf{0} = \sum_{i=1}^{l} c_i \mathbf{a}_i$, but now since the $\mathbf{a}_i$ are linearly independent, we conclude that all the coefficients c_i must also be zero.

- In the event that the linear transformation is multiplication by a matrix, the nullity-rank theorem reduces to a fact we already knew.
  - Explicitly, if A is an $m \times n$ matrix, the kernel of the multiplication-by-A map is the solution space to the homogeneous system $A\mathbf{x} = \mathbf{0}$ (i.e., the nullspace), and the image is the set of vectors $\mathbf{c}$ such that there exists a solution to $A\mathbf{x} = \mathbf{c}$ (i.e., the column space).
  - The value of $\dim(\ker(T))$ is the dimension of the nullspace, which we know is the number of nonpivotal columns in the reduced row-echelon form of A.
  - Also, the value of $\dim(\text{im}(T))$ is the dimension of the column space, which is the number of pivotal columns in the reduced row-echelon form of A.

- Therefore, the sum of these two numbers is the number of columns of the matrix A (every column is either pivotal or nonpivotal), which is simply n, the dimension of the domain space.
- Incidentally, we also see that the use of the word "rank" for the the dimension of $\text{im}(T)$ is consistent with our use of the word "rank" to refer to the rank of a matrix (since the rank of a matrix is the same as the number of pivot elements in its row-echelon form).

- Example: If $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ is the trace map $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d$, find the nullity and the rank of T and verify the nullity-rank theorem.
  - First, we compute the kernel: we see that $T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0$ when $d = -a$, so the elements of the kernel have the form $\begin{bmatrix} a & b \\ c & -a \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$.
  - So the kernel has a basis given by the three matrices $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, meaning that the nullity is $\boxed{3}$.
  - For the image, we can clearly obtain any value in $\mathbb{R}$, since $T\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} = a$ for any a. So the image is 1-dimensional, meaning that the rank is $\boxed{1}$.
  - Then the rank plus the nullity is 4, which (per the theorem) is indeed equal to the dimension of the space of 2×2 matrices.

3.1.3 Isomorphisms of Vector Spaces

- We will now discuss an important notion of equivalence of vector spaces.
- Definition: A linear transformation $T : V \to W$ is called an isomorphism if T is one-to-one and onto. Equivalently, T is an isomorphism if $\ker(T) = \{\mathbf{0}\}$ and $\text{im}(T) = W$. We say that two vector spaces are isomorphic if there exists an isomorphism between them.
  - Saying that two spaces are isomorphic is a very strong statement, as we will see: it says that the vector spaces V and W have exactly the same structure.
  - More specifically, saying that $T : V \to W$ is an isomorphism means that we can use T to relabel the elements of V to have the same names as the elements of W, and that (if we do so) we cannot tell V and W apart at all.
- Example: Show that the map $T : \mathbb{R}^4 \to M_{2\times 2}(\mathbb{R})$ given by $T(x_1, x_2, x_3, x_4) = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$ is an isomorphism.
  - This map is a linear transformation; it clearly is additive and respects scalar multiplication.
  - Also, $\ker(T) = 0$ since the only element mapping to the zero matrix is $(0, 0, 0, 0)$. And it is also clear that $\text{im}(T) = M_{2\times 2}$.
  - Thus T is an isomorphism.
- Example: Show that the map $T : \mathbb{R}^3 \to P_2(\mathbb{R})$ given by $T(a, b, c) = (a + b) + (a + c)x + (b + c)x^2$ is an isomorphism.
  - This map is a linear transformation; it clearly is additive and respects scalar multiplication.
  - Also, $\ker(T) = 0$ since $T(a, b, c) = 0$ requires $a + b = a + c = b + c = 0$, and the only solution to this system is $a = b = c = 0$.
  - Finally, a brief calculation will show that $T\left(\frac{a_0 + a_1 - a_2}{2}, \frac{a_0 + a_2 - a_1}{2}, \frac{a_1 + a_2 - a_0}{2}\right) = a_0 + a_1 x + a_2 x^2$, so $\text{im}(T) = P_2(\mathbb{R})$.

- Thus T is an isomorphism.
- Remark: Alternatively, after computing $\ker(T) = \{\mathbf{0}\}$, we could have used the nullity-rank theorem to conclude that the dimension of $\text{im}(T)$ was $3 - 0 = 3$, hence necessarily all of $P_2(\mathbb{R})$.

- Isomorphisms preserve many properties, such as linear independence:

- Proposition (Isomorphisms and Independence): If $T : V \to W$ is an isomorphism, the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ in V are linearly independent if and only if $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$ are linearly independent in W.
  - Proof: Because T is a linear transformation, we have $a_1 \cdot T(\mathbf{v}_1) + \cdots + a_n \cdot T(\mathbf{v}_n) = T(a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n)$ for any scalars $a_1, \dots, a_n$.
  - To see that $\mathbf{v}_1, \dots, \mathbf{v}_n$ independent implies $T(\mathbf{v}_1), \cdots, T(\mathbf{v}_n)$ independent:
    * If $a_1 \cdot T(\mathbf{v}_1) + \cdots + a_n \cdot T(\mathbf{v}_n) = \mathbf{0}$, then by the above we have $T(a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n) = \mathbf{0}$.
    * But now since $\ker(T) = 0$, we get $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$, and independence of $\mathbf{v}_1, \dots, \mathbf{v}_n$ then gives $a_1 = \cdots = a_n = 0$.
    * So $T(\mathbf{v}_1), \cdots, T(\mathbf{v}_n)$ are linearly independent.
  - To see that $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$ independent implies $\mathbf{v}_1, \dots, \mathbf{v}_n$ independent:
    * If $a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n = \mathbf{0}$, then $a_1 \cdot T(\mathbf{v}_1) + \cdots + a_n \cdot T(\mathbf{v}_n) = T(a_1 \cdot \mathbf{v}_1 + \cdots + a_n \cdot \mathbf{v}_n) = T(\mathbf{0}) = \mathbf{0}$.
    * But now the linear independence of $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$ gives $a_1 = \cdots = a_n = 0$, so $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.

- Any isomorphism necessarily possesses an inverse map:

- Proposition (Isomorphisms and Inverse Maps): If T is an isomorphism, then there exists an inverse function $T^{-1} : W \to V$, with $T^{-1}(T(\mathbf{v})) = \mathbf{v}$ and $T(T^{-1}(\mathbf{w})) = \mathbf{w}$ for any $\mathbf{v}$ in V and $\mathbf{w}$ in W. This inverse map T^{-1} is also a linear transformation.
  - Proof: The fact that there is an inverse function $T^{-1} : W \to V$ follows immediately because T is one-to-one and onto.
  - Specifically, for any $\mathbf{w}$ in W, by the assumption that T is onto there exists a $\mathbf{v}$ in V with $T(\mathbf{v}) = \mathbf{w}$, and because T is one-to-one, this vector $\mathbf{v}$ is unique. We then define $T^{-1}(\mathbf{w}) = \mathbf{v}$.
  - Now we check the two properties of a linear transformation:
    * [T1] If $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$, then because $T(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$, we have $T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v}_1 + \mathbf{v}_2 = T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2)$.
    * [T2] If $T(\mathbf{v}) = \mathbf{w}$, then because $T(\alpha \cdot \mathbf{v}) = \alpha \cdot \mathbf{w}$, we have $T^{-1}(\alpha \cdot \mathbf{w}) = \alpha \cdot \mathbf{v} = \alpha \cdot T^{-1}(\mathbf{w})$.

- It may seem that isomorphisms are hard to find, but this is not the case.

- Theorem (Isomorphism and Dimension): Two vector spaces V and W are isomorphic if and only if they have the same dimension. In particular, any finite-dimensional vector space is isomorphic to $\mathbb{R}^n$, where $n = \dim(V)$.
  - Proof (Finite-Dimensional Case): Isomorphisms preserve linear independence, so two vector spaces can only be isomorphic if they have the same dimension.
  - For the other direction, choose a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ for V and a basis $\mathbf{w}_1, \dots, \mathbf{w}_n$ for W. We claim the map T defined by $T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \cdots + a_1\mathbf{w}_n$ is an isomorphism between V and W.
  - We need to check five things: that T is well-defined, that T respects addition, that T respects scalar multiplication, that T is one-to-one, and that T is onto.
  - T is well-defined: The description above defines T on every element $\mathbf{v}$ of V because $\mathbf{v}_1, \dots, \mathbf{v}_n$ spans V, and the definition is unique because there is only way of writing $\mathbf{v}$ as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_n$ (because $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent).
  - T respects addition: If $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_1\mathbf{v}_n$ and $\tilde{\mathbf{v}} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$, then $T(\mathbf{v} + \tilde{\mathbf{v}}) = (a_1 + b_1)\mathbf{w}_1 + \cdots + (a_n + b_n)\mathbf{w}_n = T(\mathbf{v}) + T(\tilde{\mathbf{v}})$ by the distributive law.
  - T respects scalar multiplication: For any scalar β we have $T(\beta\mathbf{v}) = (\beta a_1)\mathbf{w}_1 + \cdots + (\beta a_n)\mathbf{w}_n = \beta T(\mathbf{v})$.

- T is one-to-one: Since $\mathbf{w}_1, \dots, \mathbf{w}_n$ are linearly independent, the only way that $a_1\mathbf{w}_1 + \cdots + a_1\mathbf{w}_n$ can be the zero vector is if $a_1 = a_2 = \cdots = a_n = 0$, which means $\ker(T) = 0$.
- T is onto: Since $\mathbf{w}_1, \dots, \mathbf{w}_n$ span W, every element $\mathbf{w}$ in W can be written as $\mathbf{w} = a_1\mathbf{w}_1 + \cdots + a_1\mathbf{w}_n$ for some scalars $a_1, \cdots a_n$. Then for $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_1\mathbf{v}_n$, we have $T(\mathbf{v}) = \mathbf{w}$.

- The result above should be rather unexpected: it certainly doesn't seem obvious, just from the eight axioms of a vector space, that all finite-dimensional vector spaces are essentially "the same" as $\mathbb{R}^n$ for some n. But that is precisely the result established above!

3.2 Matrices Associated to Linear Transformations

- So far, we have studied linear transformations $T : V \to W$ in a fairly generic way, without much reference to the structure of V or W.
  - If we choose a basis for V and a basis for W, however, we can describe the behavior of T with respect to this basis, and it turns out that T behaves exactly like multiplication by a matrix[2], at least when V and W are finite-dimensional.
- To illustrate the idea, consider the map from $T : \mathbb{R}^3 \to \mathbb{R}^2$ with $T(x, y, z) = \langle 2x - y + z, 3x + 4y - 5z\rangle$.
  - Let us choose the standard basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ for V and the standard basis $\{\mathbf{w}_1, \mathbf{w}_2\} = \{\langle 1,0\rangle, \langle 0,1\rangle\}$ for W.
  - Then $T(\mathbf{v}_1) = 2\mathbf{w}_1 + 3\mathbf{w}_2$, $T(\mathbf{v}_2) = -\mathbf{w}_1 + 4\mathbf{w}_2$, and $T(\mathbf{v}_3) = \mathbf{w}_1 - 5\mathbf{w}_2$.
  - We can summarize this by saying that $T(a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3) = (2a - b + c)\mathbf{w}_1 + (3a + 4b - 5c)\mathbf{w}_2$.
  - Notice that the coefficients of $\mathbf{w}_1$ and $\mathbf{w}_2$ are given by the entries in the matrix product $\begin{bmatrix} 2 & -1 & 1 \\ 3 & 4 & -5 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.
  - Furthermore, as we proved earlier, the behavior of T on V is completely characterized by its behavior on a basis of V, and by the definition of a basis, any vector in W is completely characterized by the coefficients when it is written as a linear combination of the basis elements of W.
  - In other words, the entries in the matrix $\begin{bmatrix} 2 & -1 & 1 \\ 3 & 4 & -5 \end{bmatrix}$ completely characterize the behavior of the linear transformation T, once we have chosen the bases $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for V and $\{\mathbf{w}_1, \mathbf{w}_2\}$ for W.
  - Observe that the columns of this matrix are simply the coefficients of the basis elements of V in terms of the basis elements of W.
- By choosing particular bases for V and for W, we obtain a correspondence between linear transformations from V to W and matrices: this will allow us to analyze both types of objects together, and to study each one using our understanding of the other.
  - For example, by using properties of linear transformations, it is possible to provide almost trivial proofs of some of the algebraic properties of matrix multiplication which are hard to prove by direct computation.
  - Conversely, we will be able to prove a number of things about linear transformations by using properties of matrix arithmetic.
  - This correspondence explains how matrices and vector spaces, which initially seem like they have almost nothing to do with one another, are in fact very closely related: matrices describe the linear transformations from one vector space to another.

[2]In fact, the correspondence between linear transformations and matrix multiplication is the reason that matrix multiplication is defined the way it is.

3.2.1 The Matrix Associated to a Linear Transformation

- To define matrices associated to linear transformations, we first need to define the objects we will use for the construction:
- Definition: If V is a finite-dimensional vector space, an ordered basis for V is a basis of V equipped with a particular ordering.
  - We will write an ordered basis in the same way as we write a generic set, and it is to be taken for granted the fact that when we write an expression like $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, we intend β to be an ordered basis unless specifically stated otherwise.
  - Example: The pairs $\beta_1 = \{\langle 1,0\rangle, \langle 0,1\rangle\}$, $\beta_2 = \{\langle 1,1\rangle, \langle 0,2\rangle\}$, and $\beta_3 = \{\langle 0,2\rangle, \langle 1,1\rangle\}$ are three different ordered bases of $\mathbb{R}^2$. (Note that $\beta_2 \neq \beta_3$ because the ordering is different.)
- Definition: Let V be a finite-dimensional vector space with an ordered basis $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. For a vector $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$, we define the coordinate vector of $\mathbf{v}$ relative to β to be the vector $[\mathbf{v}]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ in $\mathbb{R}^n$. (Note that because β is a basis of V, the coefficients $a_1, a_2, \dots, a_n$ exist and are unique.)
  - Example: If V is the space of polynomials of degree ≤ 2 with ordered basis $\beta = \{1, x, x^2\}$, then the coordinate vectors of $3 - 4x + x^2$ and $-x$ are $\begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ respectively.
  - Example: If $V = \mathbb{R}^2$ with ordered basis $\beta = \{\langle 1,1\rangle, \langle 0,2\rangle\}$, then the coordinate vectors of $\langle 1,1\rangle$, $\langle 1,5\rangle$, and $\langle 4,-2\rangle$ relative to β are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} 4 \\ -3 \end{bmatrix}$ respectively.
- By working with coordinate vectors, we can essentially transport our discussion from the general vector space V into the more concrete setting of $\mathbb{R}^n$. Explicitly:
- Proposition: Let V be a finite-dimensional vector space and let β be an ordered basis of V. Then the map $\varphi : V \to \mathbb{R}^n$ defined by $\varphi(\mathbf{v}) = [\mathbf{v}]_\beta$ is an isomorphism.
  - Proof: It is easy to see that φ is linear, since $[\mathbf{v}+\mathbf{w}]_\beta = [\mathbf{v}]_\beta + [\mathbf{w}]_\beta$ and $[c\mathbf{v}]_\beta = c[\mathbf{v}]_\beta$.
  - Furthermore, since β is linearly independent, the only vector $\mathbf{v}$ whose coordinate vector is the zero vector is $\mathbf{v} = \mathbf{0}$, so φ is one-to-one. Finally, since β spans V, the map φ is onto.
- Given a linear transformation $T : V \to W$, if we choose ordered bases $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for V and $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ for W, we can represent the behavior of T by writing down the coordinate vectors for the elements $T(\mathbf{v}_j)$ with respect to the vectors $\mathbf{w}_i$.
- Definition: Let V and W be finite-dimensional vector spaces with ordered bases $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and $\gamma = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ respectively. If $T : V \to W$ is a linear transformation, for each $1 \leq j \leq n$ and $1 \leq i \leq m$ there exist unique scalars $a_{i,j}$ such that $T(\mathbf{v}_j) = \sum_{i=1}^{m} a_{i,j}\mathbf{w}_i$ for each $1 \leq j \leq n$. The $m \times n$ matrix $[T]_\beta^\gamma$ whose (i,j)-entry is $a_{i,j}$ is called the matrix representation of T with respect to the ordered bases β and γ.
  - The definition is rather lengthy, but the basic idea is the same as the one we described above: the jth column of the matrix $[T]_\beta^\gamma$ is $[T(\mathbf{v}_j)]_\gamma$, the coordinate vector of $T(\mathbf{v}_j)$ with respect to the basis γ (of W).
- Example: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ with $T(x,y,z) = \langle 2x - y + z, 3x + 4y - 5z\rangle$. Find the matrix associated to T with respect to the standard bases $\beta = \{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ and $\gamma = \{\langle 1,0\rangle, \langle 0,1\rangle\}$ of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively.
  - We have $T(1,0,0) = 2\langle 1,0\rangle + 3\langle 0,1\rangle$, $T(0,1,0) = -1\langle 1,0\rangle + 4\langle 0,1\rangle$, and $T(0,0,1) = 1\langle 1,0\rangle - 5\langle 0,1\rangle$.

- Therefore, the matrix associated to T is $[T]_\beta^\gamma = \boxed{\begin{bmatrix} 2 & -1 & 1 \\ 3 & 4 & -5 \end{bmatrix}}$.

- Example: Let $T : \mathbb{R}^3 \to P_2(\mathbb{R})$ be defined by $T(a, b, c) = (a + b) + (a - 2c)x + (a + b + c)x^2$. Find the matrix associated to T with respect to the standard bases $\beta = \{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ and $\gamma = \{1, x, x^2\}$ of $\mathbb{R}^3$ and $P_2(\mathbb{R})$ respectively.
  - We have $T(1, 0, 0) = 1 + x + x^2$, $T(0, 1, 0) = 1 + x^2$, and $T(0, 0, 1) = -2x + x^2$.
  - Therefore, the matrix associated to T is $[T]_\beta^\gamma = \boxed{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}}$.

- Example: Let $T : P_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ be defined by $T(p) = \begin{bmatrix} p(0) & p(1) \\ p'(0) & p'(1) \end{bmatrix}$. Find the matrix associated to T with respect to the standard bases $\beta = \{1, x, x^2\}$ and $\gamma = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ of $P_2(\mathbb{R})$ and $M_{2\times 2}(\mathbb{R})$ respectively.
  - We have $T(1) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 1e_{1,1} + 1e_{1,2} + 0e_{2,1} + 0e_{2,2}$, $T(x) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = 0e_{1,1} + 1e_{1,2} + 1e_{2,1} + 1e_{2,2}$, and $T(x^2) = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = 0e_{1,1} + 1e_{1,2} + 0e_{2,1} + 2e_{2,2}$.
  - Therefore, the matrix associated to T is $[T]_\beta^\gamma = \boxed{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}}$.

- Example: Let $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$ be defined by $T(p) = \dfrac{12}{x-1} \int_1^x p(t)\, dt$. Find the matrix associated to T with respect to the standard basis $\beta = \gamma = \{1, x, x^2, x^3\}$.
  - We have $T(1) = 12$, $T(x) = 6 + 6x$, $T(x^2) = 4 + 4x + 4x^2$, and $T(x^3) = 3 + 3x + 3x^2 + 3x^3$.
  - Therefore, the matrix associated to T is $[T]_\beta^\gamma = \boxed{\begin{bmatrix} 12 & 6 & 4 & 3 \\ 0 & 6 & 4 & 3 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 3 \end{bmatrix}}$.

- We note in particular that if we use different bases, the same linear transformation will generally have different associated matrices:

- Example: Let $I : \mathbb{R}^2 \to \mathbb{R}^2$ be the identity transformation $I(a, b) = \langle a, b\rangle$. Find the matrix associated to I with respect to the standard basis $\beta_1 = \gamma_1 = \{\langle 1,0\rangle, \langle 0,1\rangle\}$ of $\mathbb{R}^2$.
  - We have $I(1, 0) = 1\langle 1,0\rangle + 0\langle 0,1\rangle$ and $I(0, 1) = 1\langle 1,0\rangle + 0\langle 0,1\rangle$.
  - Therefore, the matrix associated to I is $[I]_{\beta_1}^{\gamma_1} = \boxed{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$, the 2×2 identity matrix.

- Example: Let $I : \mathbb{R}^2 \to \mathbb{R}^2$ be the identity transformation $I(a, b) = \langle a, b\rangle$. Find the matrix associated to I with respect to the bases $\beta_2 = \{\langle 2,-2\rangle, \langle 3,1\rangle\}$ and $\gamma_2 = \{\langle 1,-1\rangle, \langle 1,1\rangle\}$ of $\mathbb{R}^2$.
  - We have $I(2, -2) = 2\langle 1,-1\rangle + 0\langle 1,1\rangle$ and $I(3, 1) = 1\langle 1,-1\rangle + 2\langle 1,1\rangle$.
  - Therefore, the matrix associated to I is $[I]_{\beta_2}^{\gamma_2} = \boxed{\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}}$.
  - Note that the matrix for this linear transformation is different from the one given above: this should not be surprising, since we are using different bases.

3.2.2 Algebraic Properties of Matrices Associated to Linear Transformations

- We can use the matrix associated to a linear transformation to evaluate the linear transformation on arbitrary vectors, using matrix multiplication.
  - Recall that if A is an $m \times n$ matrix and B is an $n \times q$ matrix, then the matrix product AB is the $m \times q$ matrix whose (i, j)-entry is the sum $(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$.
- <u>Proposition</u> (Associated Matrix Action): Suppose that $\dim(V) = n$ with an ordered basis $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, that $\dim(W) = m$ with an ordered basis $\gamma = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$, and that $T : V \to W$ is linear. If $M = [T]_\beta^\gamma$ and $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$ is a vector in V, then $T(\mathbf{v}) = y_1\mathbf{w}_1 + \cdots + y_m\mathbf{w}_m$, where $y_i = \sum_{k=1}^{n} m_{i,k} x_k$. Equivalently, the coordinate vector $[T(\mathbf{v})]_\gamma$ is given by the matrix product $M\,[\mathbf{v}]_\beta$.
  - Roughly speaking, this proposition says that the linear transformation T acts as left-multiplication by its associated matrix $[T]_\beta^\gamma$, when considered on the level of coordinate vectors.
  - <u>Proof</u>: By properties of linear transformations and the fact that $T(\mathbf{v}_i) = \sum_{j=1}^{n} m_{i,k}\mathbf{w}_i$, we can write $T(\mathbf{v}) = T(\sum_{k=1}^{n} x_i\mathbf{v}_i) = \sum_{k=1}^{n} x_i T(\mathbf{v}_i) = \sum_{k=1}^{n} x_i \left[\sum_{i=1}^{n} m_{i,k}\mathbf{w}_i\right] = \sum_{i=1}^{n} \left[\sum_{k=1}^{n} m_{i,k} x_i\right] \mathbf{w}_i$ from which we see that $y_i = \sum_{k=1}^{n} m_{i,k} x_k$ as claimed.
- <u>Example</u>: For $T : \mathbb{R}^3 \to \mathbb{R}^2$ with $T(x, y, z) = \langle 2x - y + z, 3x + 4y - 5z\rangle$, and with the standard bases β and γ of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively, verify that $[T(\mathbf{v})]_\gamma = [T]_\beta^\gamma\,[\mathbf{v}]_\beta$ for $\mathbf{v} = \langle 2, 3, 5\rangle$.
  - We computed earlier that $[T]_\beta^\gamma = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 4 & -5 \end{bmatrix}$ for this transformation.
  - For $\mathbf{v} = \langle 2, 3, 5\rangle$, we have $[\mathbf{v}]_\beta = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, so $[T]_\beta^\gamma\,[\mathbf{v}]_\beta = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 4 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ -7 \end{bmatrix}$.
  - Since $T(\mathbf{v}) = T(2, 3, 5) = \langle 6, -7\rangle$, we indeed see that $[T(\mathbf{v})]_\gamma = \begin{bmatrix} 6 \\ -7 \end{bmatrix} = [T]_\beta^\gamma\,[\mathbf{v}]_\beta$.
- By applying this result to a composition of linear transformations, we can deduce that the matrix associated to a composition of linear transformations is the matrix product of the associated matrices.
  - Indeed, the fact that matrix multiplication models the composition of linear transformations is precisely the reason that matrix multiplication is defined the way it is!
- <u>Corollary</u> (Linear Transformations and Matrix Multiplication): Suppose that U, V, and W are finite-dimensional and have ordered bases α, β, and γ respectively, and that $T : U \to V$ and $S : V \to W$ are linear transformations. Then $[ST]_\alpha^\gamma = [S]_\beta^\gamma [T]_\alpha^\beta$, or, in words, the matrix associated to ST is the product of the matrix associated to S with the matrix associated to T.
  - <u>Proof</u>: Let $\mathbf{v}$ be any vector in U. Then by the previous proposition, $[ST]_\alpha^\gamma[\mathbf{v}]_\alpha = [ST(\mathbf{v})]_\gamma$, while $[S]_\beta^\gamma[T]_\alpha^\beta[\mathbf{v}]_\alpha = [S]_\beta^\gamma[T(\mathbf{v})]_\beta = [ST(\mathbf{v})]_\gamma$.
  - Since these two expressions are equal for every vector $\mathbf{v}$ in U, the matrices $[ST]_\alpha^\gamma$ and $[S]_\beta^\gamma[T]_\alpha^\beta$ are equal.
- <u>Example</u>: Let $T : \mathbb{R}^3 \to P_2(\mathbb{R})$ be defined by $T(a, b, c) = (a + b) + (a - 2c)x + (a + b + c)x^2$ and $S : P_2(\mathbb{R}) \to M_{2\times2}(\mathbb{R})$ be defined by $S(p) = \begin{bmatrix} p(0) & p(1) \\ p'(0) & p'(1) \end{bmatrix}$. For the standard bases $\alpha = \{\langle 1, 0, 0\rangle, \langle 0, 1, 0\rangle, \langle 0, 0, 1\rangle\}$ of $\mathbb{R}^3$, $\beta = \{1, x, x^2\}$ of $P_2(\mathbb{R})$, and $\gamma = \{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$ of $M_{2\times2}(\mathbb{R})$, verify that $[ST]_\alpha^\gamma = [S]_\beta^\gamma[T]_\alpha^\beta$.

- We computed earlier that $[T]_\alpha^\beta = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$ and that $[S]_\beta^\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$.
- Thus, $[S]_\beta^\gamma [T]_\alpha^\beta = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & -1 \\ 1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix}$.
- We can also see that $ST(a,b,c) = \begin{bmatrix} a+b & 3a+2b-c \\ a-2c & 3a+2b \end{bmatrix}$ from a direct calculation, so $[ST]_\alpha^\gamma = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & -1 \\ 1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix}$. This is indeed equal to $[S]_\beta^\gamma [T]_\alpha^\beta$.

- Once we choose ordered bases β and γ for V and W, we can in fact view linear transformations $T : V \to W$ completely interchangeably with their associated matrices. More explicitly:
- Theorem (Matrices and Linear Spaces): Suppose that $\dim(V) = n$ with V having an ordered basis $\beta = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and that $\dim(W) = m$ with W having an ordered basis $\gamma = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$. Then the map $\Phi : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{R})$ defined by $\Phi(T) = [T]_\beta^\gamma$ is an isomorphism.
  - This theorem says that the space $\mathcal{L}(V,W)$ of linear transformations from V to W is isomorphic to the space of $m \times n$ matrices, where the correspondence is given by writing down the associated matrix with respect to the fixed ordered bases β and γ.
  - Proof: By the definition of isomorphism, we must show that Φ is linear, one-to-one, and onto.
  - It is a straightforward check using the definitions of the respective quantities that $[S+T]_\beta^\gamma = [S]_\beta^\gamma + [T]_\beta^\gamma$ and that $[cT]_\beta^\gamma = c[T]_\beta^\gamma$, for any linear transformations S and L in $\mathcal{L}(V,W)$ and scalar c.
  - Next, Φ is one-to-one: if $[T]_\beta^\gamma$ is the zero matrix, then for any $\mathbf{v}$ in V, the coordinate vector $[T(\mathbf{v})]_\gamma$ is the zero vector. Thus, $T(\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$ in V, so T is the zero transformation.
  - Finally, Φ is onto: for any matrix M in $M_{m\times n}(\mathbb{R})$, the linear transformation T specified by choosing $T(\mathbf{v}_j) = \sum_{i=1}^{m} m_{i,j}\mathbf{w}_i$ for each $1 \le j \le n$ has $[T]_\beta^\gamma = M$.
- Corollary: If $\dim(V) = n$ and $\dim(W) = m$, then the dimension of $\mathcal{L}(V,W)$ is mn.
  - Proof: Isomorphisms preserve dimension; the theorem above says that $\mathcal{L}(V,W)$ is isomorphic to $M_{m\times n}(\mathbb{R})$, and the latter has dimension mn.
- Finally, observe that there is a simple criterion involving the associated matrix for whether a linear transformation $T : V \to W$ is an isomorphism:
- Proposition (Isomorphisms and Associated Matrices): If V and W are finite-dimensional vector spaces with respective ordered bases β and γ, and $T : V \to W$ is linear, then T is an isomorphism if and only if $[T]_\beta^\gamma$ is an invertible matrix.
  - Proof: If T is an isomorphism, then T possesses an inverse map T^{-1} with the property that $T^{-1}(T(\mathbf{v})) = \mathbf{v}$ for every $\mathbf{v}$ in V, which is to say that $T^{-1}T$ is the identity transformation on V.
  - Then $[T^{-1}]_\gamma^\beta [T]_\beta^\gamma = [T^{-1}T]_\beta^\beta = [I]_\beta^\beta = I_n$, so $[T]_\beta^\gamma$ is invertible with inverse matrix $[T^{-1}]_\gamma^\beta$.
  - Conversely, suppose $[T]_\beta^\gamma$ is an invertible matrix. Then T is one-to-one, since $T(\mathbf{v}) = \mathbf{0} \iff [T(\mathbf{v})]_\gamma = \mathbf{0}$ $\iff [T]_\beta^\gamma [\mathbf{v}]_\beta = \mathbf{0} \overset{[T]_\beta^\gamma \text{ invertible}}{\iff} [\mathbf{v}]_\beta = \mathbf{0} \iff \mathbf{v} = \mathbf{0}$.
  - Furthermore, if $\mathbf{w}$ is any vector in W, then the vector $\mathbf{v}$ with $[\mathbf{v}]_\beta = ([T]_\beta^\gamma)^{-1}[\mathbf{w}]_\gamma$ has the property that $[T(\mathbf{v})]_\gamma = [T]_\beta^\gamma [\mathbf{v}]_\beta = [T]_\beta^\gamma ([T]_\beta^\gamma)^{-1}[\mathbf{w}]_\gamma = [\mathbf{w}]_\gamma$, so that $T(\mathbf{v}) = \mathbf{w}$. Thus, T is also onto, hence is an isomorphism.

- By using this result in tandem with our prior results about invertible matrices, we can give a quite lengthy list of equivalent criteria for invertibility:

- Proposition (Invertible Matrices): If V is an n-dimensional vector space with ordered basis β, and $T : V \to V$ has associated matrix $A = [T]_\beta^\beta$, the following are equivalent:

  1. The matrix A is invertible: there exists an $n \times n$ matrix B with $AB = I_n = BA$.
  2. The matrix A has a right inverse: there exists an $n \times n$ matrix B with $AB = I_n$.
  3. The matrix A has a left inverse: there exists an $n \times n$ matrix B with $CA = I_n$.
  4. The determinant of A is nonzero.
  5. The linear system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.
  6. The linear system $A\mathbf{x} = \mathbf{c}$ has exactly one solution $\mathbf{x}$, for any $\mathbf{c}$.
  7. The matrix A is row-equivalent to the identity matrix.
  8. The linear transformation T is an isomorphism.
  9. The kernel of T consists only of the zero vector, which is to say, T is one-to-one.
  10. The rank of T (equivalently, the rank of A) is equal to n, which is to say, T is onto.

  - Proof: We already showed that (1)-(7) are equivalent to one another during our study of matrices and determinants.
  - By the nullity-rank theorem, we have $\dim(\ker T) + \dim(\mathrm{im} T) = \dim(V)$, so the dimension of $\ker(T)$ is 0 if and only if the dimension of $\mathrm{im}(T)$ is equal to $\dim(V)$. In other words, V is one-to-one if and only if it is onto, so (9) and (10) are equivalent.
  - Since (9) and (10) together are equivalent to (8), this means that (8)-(10) are equivalent.
  - Finally, the proposition above establishes that (1) is equivalent to (8), so (1)-(10) are all equivalent.

- As a final warning, we will note that some of the equivalent conditions above break down in infinite-dimensional spaces:

- Example: If V is the vector space of infinite sequences of real numbers, let $L : V \to V$ be the "left shift" operator $L(a_1, a_2, a_3, a_4, \dots) = (a_2, a_3, a_4, \dots)$ and $R : V \to V$ be the "right shift" operator $R(a_1, a_2, a_3, \dots) = (0, a_1, a_2, a_3, \dots)$.

  - Observe that $LR(a_1, a_2, a_3, a_4, \dots) = L(0, a_1, a_2, a_3, a_4, \dots) = (a_1, a_2, a_3, a_4, \dots)$, so that LR is the identity operator.
  - However, $RL(a_1, a_2, a_3, a_4, \dots) = R(a_2, a_3, a_4, \dots) = (0, a_2, a_3, a_4, \dots)$ is not the identity operator.
  - In this case, we see that L has a right inverse (namely, R) and R has a left inverse (namely, L), but neither L nor R has a two-sided inverse.
  - Furthermore, observe that L is onto, but that it is not one-to-one, since the sequences of the form $(a_1, 0, 0, 0, \dots)$ are in $\ker(L)$.
  - Inversely, R is one-to-one but not onto, since $\mathrm{im}(R)$ consists of the sequences whose first entry is zero.

3.2.3 Geometry of Linear Transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$

- We can give concrete geometric descriptions of various classes of linear transformations on the Euclidean plane (i.e., linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$).

  - As noted above, these are the linear transformations of the form $T(\mathbf{v}) = A\mathbf{v}$ for a 2×2 matrix A.
  - Since a linear transformation is uniquely determined by its values on a basis, we can determine the matrix A by calculating the action of A on the standard basis $\{\langle 1, 0\rangle, \langle 0, 1\rangle\}$ of $\mathbb{R}^2$: if $T(1, 0) = \langle a, b\rangle$ and $T(0, 1) = \langle c, d\rangle$, then the matrix A is $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$.
  - Unless otherwise noted, we will always write our matrices with respect to the standard basis here.

- The simplest transformations are <u>scalings</u>: these transformations simply stretch in the coordinate directions by an appropriate amount.
  - We can see that the linear transformation that scales by a factor of a in the x-direction and a factor of b in the y-direction maps $\mathbf{e}_1$ to $a\mathbf{e}_1$ and $\mathbf{e}_2$ to $b\mathbf{e}_2$, so the underlying matrix is $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.
  - <u>Example</u>: The matrix $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ corresponds to a scaling by a factor of 2 in the x-direction and a factor of 3 in the y-direction.
  - Note that a negative scaling factor corresponds to a reflection, so for example the matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ corresponds to reflection through the origin.
- Another simple class of transformations are the <u>rotations</u>: these transformations rotate by an angle θ counterclockwise about the origin.
  - It follows from a straightforward calculation using polar coordinates (which is ultimately just using the definition of sine and cosine) that a rotation of θ radians counterclockwise maps the vector $\langle 1, 0\rangle$ to $\langle \cos\theta, \sin\theta\rangle$ and maps $\langle 0, 1\rangle$ to $\langle -\sin\theta, \cos\theta\rangle$.
  - Therefore, the underlying matrix for a counterclockwise rotation of θ radians is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.
  - <u>Example</u>: The matrix $\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$ corresponds to a counterclockwise rotation of $\pi/4$ radians.
- A more complicated transformation is a <u>projection</u> onto a line.
  - We will discuss orthogonal projections in greater detail later, but for the case of $\mathbb{R}^2$, the basic idea to generalize the two observations that the linear transformation T_x with $T_x(x, y) = \langle x, 0\rangle$ gives the component of a vector along the x-axis and the linear transformation T_y with $T_y(x, y) = \langle 0, y\rangle$ gives the component along the y-axis.
  - To generalize this, first note that for a nonzero vector $\mathbf{v} = \langle a, b\rangle$, the vector $\mathbf{v}^\perp = \langle -b, a\rangle$ is perpendicular to $\mathbf{v}$.
  - Then, since $\{\mathbf{v}, \mathbf{v}^\perp\}$ is a basis for $\mathbb{R}^2$, the linear transformation $P_\mathbf{v}$ representing projection into the direction of $\mathbf{v}$ has $P_\mathbf{v}(\mathbf{v}) = \mathbf{v}$ and $P_\mathbf{v}(\mathbf{v}^\perp) = \mathbf{0}$, since the component of $\mathbf{v}$ in the direction of itself is clearly $\mathbf{v}$, and the component of the vector perpendicular to $\mathbf{v}$ is $\mathbf{0}$.
  - Explicitly, we require $P_\mathbf{v}(a, b) = \langle a, b\rangle$ and $P_\mathbf{v}(-b, a) = \langle 0, 0\rangle$. It is then not hard to verify that the associated matrix for P with respect to the standard basis is then $\dfrac{1}{a^2 + b^2}\begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$.
  - In particular, if we take $\mathbf{v} = \langle 1, m\rangle$, we obtain the orthogonal projection onto the line $y = mx$.
  - <u>Example</u>: With $\mathbf{v} = \langle 1, 1\rangle$, we see that the matrix $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ corresponds to an orthogonal projection onto the line $y = x$.
  - <u>Example</u>: With $\mathbf{v} = \langle 1, 2\rangle$, we see that the matrix $\begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix}$ corresponds to an orthogonal projection onto the line $y = 2x$.
- Using our calculations with projections, we can also describe <u>reflection</u> across a line.
  - Explicitly, suppose we wish to write down the linear transformation R_ℓ characterizing reflection across a line ℓ through the origin.
  - If $\mathbf{w}$ is the initial vector and $R_\ell(\mathbf{w})$ is its reflection across the line ℓ, then the midpoint of $\mathbf{w}$ and $R_\ell(\mathbf{w})$ is the vector projection of $\mathbf{w}$ onto ℓ.
  - This says $(\mathbf{w} + R_\ell(\mathbf{w}))/2 = P_\ell(\mathbf{w})$, so we obtain the formula $R_\ell(\mathbf{w}) = 2P_l(\mathbf{w}) - \mathbf{w}$.

- If ℓ is spanned by $\mathbf{v} = \langle a, b \rangle$, then by the calculation above, the matrix associated to reflection across the line is the matrix $2\frac{1}{a^2+b^2}\begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix} - I_2 = \frac{1}{a^2+b^2}\begin{bmatrix} a^2-b^2 & 2ab \\ 2ab & b^2-a^2 \end{bmatrix}$.
- Example: With $\mathbf{v} = \langle 1, 1 \rangle$, we see that the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ corresponds to reflection across the line $y = x$.
- Example: With $\mathbf{v} = \langle 1, 2 \rangle$, we see that the matrix $\begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}$ corresponds to reflection across the line $y = 2x$.

- By composing various transformations with one another, we can generate other transformations of the plane, such as scaled rotations.
  - We remark that transformations applied in a sequence are composed right-to-left, since they are functions.
  - For example, the composition of scaling both coordinates by a factor of r and then rotating counterclockwise by θ radians corresponds to the matrix product $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix} = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$ where $x = r\cos\theta$ and $y = r\sin\theta$ in polar coordinates.

- Example: Find the matrix associated to the transformation obtained by first doubling the x-coordinate and quadrupling the y-coordinate, then rotating counterclockwise by $\pi/2$ radians, and finally projecting onto the line $y = x$.
  - The first transformation has associated matrix $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, the second transformation has associated matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, and the last transformation has associated matrix $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$.
  - The transformations are composed right-to-left (just as functions are), so the composition of these three transformations corresponds to the product $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} = \boxed{\begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix}}$.

- Example: Find the matrix associated to the transformation obtained by first reflecting across the line $y = 2x$ and then reflecting across the line $y = 3x$.
  - The first transformation has associated matrix $\begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}$ and the second transformation has associated matrix $\begin{bmatrix} -4/5 & 3/5 \\ 3/5 & 4/5 \end{bmatrix}$.
  - The desired composition is then $\begin{bmatrix} -4/5 & 3/5 \\ 3/5 & 4/5 \end{bmatrix}\begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix} = \boxed{\begin{bmatrix} 24/25 & -7/25 \\ 7/25 & 24/25 \end{bmatrix}}$.
  - Remark: Observe that the composition of these two reflections is a rotation (in this case, counterclockwise rotation by the angle θ with $\cos\theta = 24/25$ and $\sin\theta = 7/25$). Interestingly, if we compose these two reflections in the reverse order, we will obtain a clockwise rotation by this angle θ.

3.2.4 Change of Basis and Similarity

- Next we will discuss the idea of change of coordinates in the context of vector spaces.
  - As motivation, consider the graph of the equation $6x^2 + 4xy + 9y^2 = 1$ in the plane. Without modifying the equation, it is difficult to determine the shape of the graph of this curve.
  - If, however, we define new variables $s = \frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y$ and $t = \frac{2}{\sqrt{5}}x - \frac{1}{\sqrt{5}}y$, a short computation will show that the equation $6x^2 + 4xy + 9y^2 = 1$ is equivalent to $2s^2 + t^2 = 5$.

- Since the vectors $\mathbf{s} = \frac{1}{\sqrt{5}} \langle 1, 2 \rangle$ and $\mathbf{t} = \frac{1}{\sqrt{5}} \langle 2, -1 \rangle$ are orthogonal and have length 1 in $\mathbb{R}^2$, we can see that the equation $2s^2 + t^2 = 5$ therefore represents an ellipse whose two axes have lengths $2\sqrt{5}$ (in the $\mathbf{t}$-direction) and $\sqrt{10}$ (in the $\mathbf{s}$-direction).
- By using the basis $\{\mathbf{s}, \mathbf{t}\}$ for $\mathbb{R}^2$ rather than the standard basis $\{\langle 1, 0 \rangle, \langle 0, 1 \rangle\}$, we obtain a more useful description of the curve $6x^2 + 4xy + 9y^2 = 1$.
- We would like to describe, in general, how the coordinates of vectors change when we write them in terms of a new basis.

- Definition: Suppose V is a finite-dimensional vector space with two ordered bases β and γ. The associated change-of-basis matrix from β to γ is defined to be $Q = [I]_\beta^\gamma$, where I is the identity transformation on V.
  - Note that $[I]_\beta^\gamma$ is the matrix whose columns represent the vectors in β as linear combinations of the vectors in γ.

- Proposition (Change of Basis and Inverses): Suppose β and γ are two ordered bases of the finite-dimensional vector space V. Then the change-of-basis matrix $[I]_\beta^\gamma$ is invertible with inverse $[I]_\gamma^\beta$, and for any vector $\mathbf{v}$ in V we have $[\mathbf{v}]_\gamma = [I]_\beta^\gamma [\mathbf{v}]_\beta$.
  - Proof: Observe that $[I]_\beta^\gamma [I]_\gamma^\beta = [I]_\gamma^\gamma$ is the identity matrix, and likewise $[I]_\gamma^\beta [I]_\beta^\gamma = [I]_\beta^\beta$ is also the identity matrix. Therefore, $[I]_\beta^\gamma$ is invertible and its inverse is $[I]_\gamma^\beta$.
  - Furthermore, by our proposition about the associated matrix action, we have $[I]_\beta^\gamma [\mathbf{v}]_\beta = [I\mathbf{v}]_\gamma = [\mathbf{v}]_\gamma$.

- Example: In $\mathbb{R}^3$, let $\beta = \{\langle 2, 1, 2 \rangle, \langle -1, 1, 0 \rangle, \langle 3, 1, 3 \rangle\}$ and $\gamma = \{\langle 1, 0, 0 \rangle, \langle 1, 1, 0 \rangle, \langle 1, 1, 1 \rangle\}$. Find the change-of-basis matrix $[I]_\beta^\gamma$ and verify that $[\mathbf{v}]_\gamma = [I]_\beta^\gamma [\mathbf{v}]_\beta$ for $\mathbf{v} = \langle 13, 9, 16 \rangle$.
  - We compute $\langle 2, 1, 2 \rangle = 1 \langle 1, 0, 0 \rangle - 1 \langle 1, 1, 0 \rangle + 2 \langle 1, 1, 1 \rangle$, $\langle -1, 1, 0 \rangle = -2 \langle 1, 0, 0 \rangle + 1 \langle 1, 1, 0 \rangle + 0 \langle 1, 1, 1 \rangle$, and $\langle 3, 1, 3 \rangle = 2 \langle 1, 0, 0 \rangle - 2 \langle 1, 1, 0 \rangle + 3 \langle 1, 1, 1 \rangle$.
  - Thus, the change-of-basis matrix is $[I]_\beta^\gamma = \boxed{\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & -2 \\ 2 & 0 & 3 \end{bmatrix}}$.
  - We also calculate $\langle 13, 9, 16 \rangle = 2 \langle 2, 1, 2 \rangle + 3 \langle -1, 1, 0 \rangle + 4 \langle 3, 1, 3 \rangle = 4 \langle 1, 0, 0 \rangle - 7 \langle 1, 1, 0 \rangle + 16 \langle 1, 1, 1 \rangle$.
  - Then $[I]_\beta^\gamma [\mathbf{v}]_\beta = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & -2 \\ 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \\ 16 \end{bmatrix} = [\mathbf{v}]_\gamma$, as required.

- If we have a linear transformation $T : V \to W$, we can change basis in both V and W to obtain a new matrix associated to T. This matrix is a product of the original matrix with the appropriate change-of-basis matrices in a natural way:

- Proposition (Change of Basis): Suppose α and β are ordered bases of the finite-dimensional vector space V, that γ and δ are ordered bases of the finite-dimensional vector space W, and that $T : V \to W$ is linear. Then $[T]_\gamma^\delta = P^{-1} [T]_\alpha^\beta Q$, where $P = [I]_\delta^\beta$ and $Q = [I]_\gamma^\alpha$ are the change of basis matrices from δ to β and γ to α respectively.
  - Proof: By the previous proposition on the change of basis matrix, $P^{-1} = [I]_\beta^\delta$. Then $P^{-1} [T]_\alpha^\beta Q = [I]_\beta^\delta [T]_\alpha^\beta [I]_\gamma^\alpha = [ITI]_\gamma^\delta = [T]_\gamma^\delta$, as claimed.

- Corollary: Suppose α and β are ordered bases of the finite-dimensional vector space V and $T : V \to V$ is linear. Then $[T]_\gamma^\gamma = Q^{-1} [T]_\beta^\beta Q$ where $Q = [I]_\gamma^\beta$ is the change of basis matrix.
  - Proof: Apply the previous result when $\beta = \alpha$ and $\delta = \gamma$.

- Example: In $P_1(\mathbb{R})$, let $\beta = \{1, 1 - x\}$ and $\gamma = \{1 + x, x\}$. For $T(p) = p(1) + xp'(x)$, find $[T]_\beta^\beta$ and $[T]_\gamma^\gamma$ and verify that $[T]_\gamma^\gamma = Q^{-1} [T]_\beta^\beta Q$ where $Q = [I]_\gamma^\beta$.

- We have $T(1) = 1$ and $T(1-x) = -1 + (1-x)$, so $[T]_\beta^\beta = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$.
- Also, $T(1+x) = 2(1+x) - x$ and $T(x) = 1 + x$, so $[T]_\gamma^\gamma = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$.
- Since $1 + x = 2 - (1-x)$ and $x = 1 - (1-x)$, we have $Q = [I]_\gamma^\beta = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$.
- Inversely, since $1 = (1+x) - x$ and $1 - x = (1+x) - 2x$, $Q^{-1} = [I]_\beta^\gamma = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$.
- Then $Q^{-1}[T]_\beta^\beta Q = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} = [T]_\gamma^\gamma$ as claimed.

- For a variety of reasons, we will be interested in studying classes of matrices which represent the same linear transformation in different bases. Such matrices have a particular name:
- Definition: We say two $n \times n$ matrices A and B are similar (or conjugate) if there exists an invertible $n \times n$ matrix Q such that $B = Q^{-1}AQ$. (We refer to $Q^{-1}AQ$ as the conjugate of A by Q.)
  - Example: The matrices $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 \\ 2 & -1 \end{bmatrix}$ are similar with $Q = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$. Explicitly, we can compute that $Q^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$, and then see $\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$, so that $Q^{-1}AQ = B$.
  - Remark: The matrix $Q = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$ also has $B = Q^{-1}AQ$. In general, if two matrices A and B are similar, then there can be many different matrices Q with $B = Q^{-1}AQ$.
- Proposition (Similar Matrices): If A and B are similar $n \times n$ matrices, then there exists a linear transformation $T : V \to V$ on an n-dimensional vector space and ordered bases α and β of V such that $A = [T]_\alpha^\alpha$ and $B = [T]_\beta^\beta$.
  - More simply: If A and B are similar $n \times n$ matrices, then A and B are the associated matrices to some shared linear transformation.
  - Proof: Suppose $B = Q^{-1}AQ$ for some Q. Choose any n-dimensional vector space V with ordered basis α, and let $T : V \to V$ be the linear transformation with $A = [T]_\alpha^\alpha$.
  - Take β to be the ordered basis such that $Q = [I]_\beta^\alpha$: in other words, with $\beta_j = \sum_{i=1}^{n} Q_{i,j}\alpha_i$. (Note that since Q is invertible, these β_j are actually a basis for V.)
  - Then $B = Q^{-1}AQ = Q^{-1}[T]_\alpha^\alpha Q = [T]_\beta^\beta$ by our results above.
- Similar matrices, owing to the fact that they represent the same linear transformation in different bases, share many algebraic properties. If $B = Q^{-1}AQ$ and $D = Q^{-1}CQ$, then we have the following:
  - The sum of the conjugates is the conjugate of the sum: $B + D = Q^{-1}AQ + Q^{-1}CQ = Q^{-1}(A + C)Q$.
  - The product of the conjugates is the conjugate of the product: $BD = Q^{-1}AQ \cdot Q^{-1}CQ = Q^{-1}(AC)Q$.
  - The inverse of the conjugate is the conjugate of the inverse: A^{-1} exists if and only if B^{-1} exists, and $B^{-1} = Q^{-1}A^{-1}Q$.
- We will return to study similar matrices, and in particular make substantial progress toward answering the following question: given a matrix A, what is the simplest matrix B that A is similar to?

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