1.7 Linear Independence

Definition An indexed set of vectors $\{\bar{v}_1, \dots, \bar{v}_p\} \subset \mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\bar{v}_1 + \cdots + x_p\bar{v}_p = \bar{0}$$

has only the trivial solution, i.e. if the only solution is $(x_1, \dots, x_p) = (0, \dots, 0)$. Likewise, the set $\{\bar{v}_1, \dots, \bar{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zero, such that

$$c_1\bar{v}_1 + \cdots + c_p\bar{v}_p = \bar{0}.$$

We call such an equation a **linear dependence relation** when the weights are not all zero.

Note A set of vectors cannot be both linearly independent and linearly dependent, but it must be one of them!

Example Determining if a set of vectors is linearly independent is tantamount to solving the matrix equation

$$A\bar{x} = \bar{0},$$

where the columns of A are given by the vectors. The set of vectors is linearly independent if and only if the only solution is $\bar{x} = \bar{0}$, . Otherwise, there is some linear dependence relation. We determine if the following vectors in $\mathbb{R}^3$ are linearly independent:

$$\bar{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \qquad \bar{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \bar{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

We begin by reducing the corresponding augmented matrix to echelon form:

$$\begin{aligned} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 \end{bmatrix} &\sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -1 & -1 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \end{aligned}$$

from which we see that there is a solution. Now we want to know if the solution is the trivial solution $(0, 0, 0)$. Normally, we would continue row operations until we reach reduced echelon form, but we can be smarter about this. Notice first of all that there are no bad rows (so there is at least one solution), which we expect since a homogeneous system always has at least the trivial solution. Since there are no free variables, we see there is *exactly* one solution. This tells us that the solution must be

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hence the set of vectors $\{\bar{v}_1, \bar{v}_2, \bar{v}_3\}$ is linearly independent.

In general, it is very useful to determine if some set of vectors is linearly independent, so it is good to have some theorems to handle this problem quickly in certain special cases.

Theorem If a set $S = \{\bar{v}_1, \dots, \bar{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Proof By reindexing the vectors (if necessary), we may assume that $\bar{v}_1 = \bar{0}$. Then

$$1\bar{v}_1 + 0\bar{v}_2 + \cdots + 0\bar{v}_p = \bar{0},$$

which shows that S is linearly dependent.

Theorem A set of two vectors in $\mathbb{R}^n$ is linearly independent if and only if neither vector is a multiple of the other.

Proof For the forward direction of the proof, take two vectors $\bar{v}_1$ and $\bar{v}_2$ in $\mathbb{R}^n$ that are linearly independent. By the preceding theorem, neither vector is zero. If one was a multiple of the other, say $\bar{v}_1 = c\bar{v}_2$, then we could rewrite the equation as

$$\bar{v}_1 - c\bar{v}_2 = \bar{0},$$

which would be a linear dependence relation between $\bar{v}_1$ and $\bar{v}_2$. This contradicts the assumption that they are linearly independent. This tells us that if $\bar{v}_1$ and $\bar{v}_2$ are linearly independent, then one cannot be a multiple of the other.

For the reverse direction, we suppose the vectors $\bar{v}_1$ and $\bar{v}_2$ are not multiples of each other. If there was a linear dependence relation between them, we would have

$$a\bar{v}_1 + b\bar{v}_2 = \bar{0}$$

with at least one of a or b not equal to zero. Suppose $a \neq 0$. Then we could solve the above equation to see

$$\begin{aligned} a\bar{v}_1 &= -b\bar{v}_2 \\ \bar{v}_1 &= -\frac{b}{a}\bar{v}_2, \end{aligned}$$

which tells us that $\bar{v}_1$ is a multiple of $\bar{v}_2$. But since we have assumed that $\bar{v}_1$ is *not* a multiple of $\bar{v}_2$, this is a contradiction. This tells us that if $\bar{v}_1$ and $\bar{v}_2$ are not multiples of each other, then they must be linearly independent. This completes the proof.

Theorem Any set $\{\bar{v}_1, \ldots, \bar{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

Proof Let $A = [\bar{v}_1 \ldots \bar{v}_p]$, so A is an $n \times p$ matrix. We consider the matrix equation $A\bar{x} = \bar{0}$, which corresponds to a homogeneous system of n equations in p variables. If $p > n$, this means there are more variables than there are equations, so there must be at least one free variable. This tells us that $A\bar{x} = \bar{0}$ has a nontrivial solution, which implies that the columns of A are linearly dependent.

Example Without doing any computation, state why the following sets of vectors are linearly dependent:

$$\text{(a)} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \right\} \quad \text{(b)} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\} \quad \text{(c)} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -4 \\ -8 \\ -12 \\ -16 \end{bmatrix} \right\}.$$

For (a), there are four vectors having three entries each. For (b), the set includes the zero vector. For (c), one vector is a multiple of the other.

Characterizing Linearly Dependent Sets

Theorem An indexed set $S = \{\bar{v}_1, \ldots, \bar{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. If S is linearly dependent and $\bar{v}_1 \neq 0$, then some $\bar{v}_j$ (with $1 < j \leq p$) is a linear combination of the preceding vectors $\bar{v}_1, \ldots, \bar{v}_{j-1}$.

Note This does not mean that *every* vector in S is a linear combination of the others, but only that *at least one* vector in a linearly dependent set is a linear combination of others.

Example Let

$$\bar{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \text{ and } \bar{v} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

We describe the set spanned by $\bar{u}$ and $\bar{v}$, and we explain why a vector $\bar{w}$ is in $\mathrm{Span}\{\bar{u}, \bar{v}\}$ if and only if $\{\bar{u}, \bar{v}, \bar{w}\}$ is linearly dependent. Since $\bar{u}$ and $\bar{v}$ are not multiples of each other, the set $\{\bar{u}, \bar{v}\}$ is linearly independent, so these vectors span a plane in $\mathbb{R}^3$. If $\bar{w}$ is a linear combination of $\bar{u}$ and $\bar{v}$, then $\{\bar{u}, \bar{v}, \bar{w}\}$ is linearly *dependent* (by the above theorem). Reasoning in the other direction, if we suppose $\{\bar{u}, \bar{v}, \bar{w}\}$ is linearly *dependent*, then some vector in $\{\bar{u}, \bar{v}, \bar{w}\}$ is a linear combination of the others. Since $\bar{u} \neq 0$ and $\bar{u}$ is not a multiple of $\bar{v}$, the only remaining possibility is that $\bar{w}$ must be a linear combination of $\bar{u}$ and $\bar{v}$.

Examples The following problem is to be worked on in class (time-permitting):
Let

$$\bar{u} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \qquad \bar{v} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}, \qquad \bar{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \bar{z} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(a) Is any pair of these vectors (e.g. $\{\bar{u}, \bar{v}\}$) linearly dependent? Explain.
(b) Does the answer to part (a) tell us that $\{\bar{u}, \bar{v}, \bar{w}, \bar{z}\}$ is linearly independent?
(c) Is $\{\bar{u}, \bar{v}, \bar{w}, \bar{z}\}$ linearly dependent? You should be able to answer this question without any computation.

1.8 Introduction to Linear Transformations

Definition A **transformation** (or a **function** or **map**) T from $\mathbb{R}^n$ to $\mathbb{R}^m$, denoted

$$T : \mathbb{R}^n \to \mathbb{R}^m,$$

is a rule that assigns to each vector $\bar{x} \in \mathbb{R}^n$ a vector $T(\bar{x}) \in \mathbb{R}^m$. We call $\mathbb{R}^n$ the **domain** of T, and we call $\mathbb{R}^m$ the **codomain** of T. For all $\bar{x} \in \mathbb{R}^n$, we call $T(\bar{x})$ the **image** of $\bar{x}$ under T. The set of all images $T(\bar{x})$ is called the **range** (or **image**) of T.

We will focus mainly on *matrix transformations*, or transformations $T : \mathbb{R}^n \to \mathbb{R}^m$ given by $T(\bar{x}) = A\bar{x}$ for some $m \times n$ matrix A. In this case, the domain of T is $\mathbb{R}^n$, the codomain is $\mathbb{R}^m$, and the range is all linear combinations of the columns of A.

Definition A transformation T is called a **linear transformation** if:
(i) $T(\bar{u} + \bar{v}) = T(\bar{u}) + T(\bar{v})$, and
(ii) $T(c\bar{u}) = cT(\bar{u})$.

In words these conditions mean that linear transformations preserve the operations of addition and scalar multiplication. That is, you can add the vectors $\bar{u}$ and $\bar{v}$ in $\mathbb{R}^n$ first, then apply T, or you can apply T to each vector separately and then add the resulting vectors in $\mathbb{R}^m$, and either way you wind up with the same thing. A parallel statement holds for scalar multiplication. It may help to see both an example and a non-example.

Example Let $T : \mathbb{R} \to \mathbb{R}$ be given by $T(x) = 2x$. Then T is a linear transformation since for $a, b, c \in \mathbb{R}$ we have

$$\begin{aligned} T(a+b) &= 2(a+b) = 2a + 2b = T(a) + T(b), \text{ and} \\ T(ca) &= 2(ca) = c(2a) = cT(a). \end{aligned}$$

Non-example Let $T : \mathbb{R} \to \mathbb{R}$ be given by $T(x) = x^2$. Then T is *not* a linear transformation, because

$$4 = 2^2 = T(2) = T(1+1) \neq T(1) + T(1) = 1^2 + 1^2 = 2.$$

Notice that the two conditions for a map to be linear are very similar to the theorem at the end of Section 1.4. Namely, for an $m \times n$ matrix A, vectors $\bar{u}, \bar{v} \in \mathbb{R}^n$, and scalars $c \in \mathbb{R}$, we know that

$$\begin{aligned} A(\bar{u} + \bar{v}) &= A\bar{u} + A\bar{v}, \text{ and} \\ A(c\bar{u}) &= cA(\bar{u}). \end{aligned}$$

This tells us that any matrix transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

given by

$$T(\bar{x}) = A\bar{x}$$

is a linear transformation.

Properties Some additional properties of linear transformations include the following:
(i) $T(\bar{0}) = \bar{0}$,
(ii) $T(c\bar{u} + d\bar{v}) = cT(\bar{u}) + dT(\bar{v})$, and more generally

$$T(c_1\bar{v}_1 + \cdots + c_p\bar{v}_p) = c_1T(\bar{v}_1) + \cdots + c_pT(\bar{v}_p).$$

Examples of Linear Transformations and Their Geometric Interpretations Here we give some examples of linear transformations.

Projection For a vector in $\mathbb{R}^m$, a map $T : \mathbb{R}^m \to \mathbb{R}^m$ that zeroes out one or more of the coordinates is a linear map that we call a *projection*. For example, the map

$$T : \mathbb{R}^3 \to \mathbb{R}^3$$

given by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

is an example of projection. Specifically, it takes a vector in $\mathbb{R}^3$ and projects it onto the x_1x_2-plane. This same map could also be given as a matrix transformation. Let

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

Then $T(\bar{x}) = A\bar{x}$ is the same mapping defined above, since

$$T(\bar{x}) = A\bar{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}.$$

Shear Shear transformations are those that take a square and turn it into a parallelogram. For example, the map $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(\bar{x}) = A\bar{x}$$

where A is the matrix $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ is a shear transformation. We will see how this looks in class.

Scaling Given a scalar $r \in \mathbb{R}$, the map $T : \mathbb{R}^m \to \mathbb{R}^m$ given by $T(\bar{x}) = r\bar{x}$ is called a **contraction** if $0 \leq r \leq 1$, and it is called a **dilation** if $r > 1$. We think of such a map as stretching or shrinking the size of a vector.

Example The following is to be worked on in class (time-permitting):

(Practice Problem #2 on Page 68)
Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and give a geometric description of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(\bar{x}) = A\bar{x}$.

1.9 The Matrix of a Linear Transformation

Notation We let $\bar{e}_i$ represent the vector that has 1 in the ith place, and 0 everywhere else. For example, if we are in $\mathbb{R}^4$, then

$$\bar{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Definition We call the matrix

$$I_n = [\ \bar{e}_1 \quad \bar{e}_2 \quad \dots \quad \bar{e}_n\ ] = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & & & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

having 1 in each diagonal entry and 0 elsewhere the **identity matrix** of $\mathbb{R}^n$.

Note We call I_n the identity matrix because $I_n\bar{x} = \bar{x}$ for all $\bar{x} \in \mathbb{R}^n$.

For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, we can completely describe the linear transformation in terms of its effect on the vectors $\bar{e}_i$. This is made formal by the following theorem.

Theorem Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is given uniquely as a matrix transformation by the $m \times n$ matrix $A = [\ T(\bar{e}_1) \quad T(\bar{e}_2) \quad \dots \quad T(\bar{e}_n)\ ]$. That is

$$T(\bar{x}) = A\bar{x}$$

for all $\bar{x} \in \mathbb{R}^n$.

Proof We prove that $T(\bar{x}) = A\bar{x}$, but not the uniqueness part. By the note above, we can write $\bar{x} = I_n\bar{x} = x_1\bar{e}_1 + \cdots + x_n\bar{e}_n$ for each $\bar{x} \in \mathbb{R}^n$. Then

$$\begin{aligned} T(\bar{x}) &= T(x_1\bar{e}_1 + \cdots + x_n\bar{e}_n) \\ &= x_1T(\bar{e}_1) + \cdots + x_nT(\bar{e}_n) \\ &= [\ T(\bar{e}_1) \quad T(\bar{e}_2) \quad \dots \quad T(\bar{e}_n)\ ] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= A\bar{x}. \end{aligned}$$

The second equality follows by linearity of T, and the last two equalities follow by definition.

Definition We call the matrix A mentioned in the preceding theorem the **standard matrix for the linear transformation** T.

Example Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that rotates each point in $\mathbb{R}^2$ about the origin by an angle θ. We take for granted that such a transformation is linear (there's a picture on page 67 of your textbook that provides a geometric reason why this should be true). From the geometry of the unit circle, we see that

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

so the standard matrix for this linear transformation is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Example During class, we will talk about reflections, scalings, shears, and projections, giving examples of each as they are described in the tables on pp. 73-75 of the textbook.

Definition A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** if the range of T is $\mathbb{R}^m$. In other words, each $\bar{b} \in \mathbb{R}^m$ is the image of *at least one* $\bar{x} \in \mathbb{R}^n$.

Definition A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each $\bar{b} \in \mathbb{R}^m$ is the image of *at most one* $\bar{x} \in \mathbb{R}^n$.

Example Let T be the linear transformation whose standard matrix is

$$A = \begin{bmatrix} 2 & 7 & 3 & 0 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

So T is a map from $\mathbb{R}^4 \to \mathbb{R}^3$. Is T onto? Is it one-to-one?

For the onto question, note that A is in echelon form, and there's a pivot in each row, so by a theorem from section 1.4 we know that the columns of A span $\mathbb{R}^3$, i.e. the equation $A\bar{x} = \bar{b}$ is consistent for each $\bar{b} \in \mathbb{R}^3$. This means exactly that each element of $\mathbb{R}^3$ is in the range of T, so T is onto.

To answer the one-to-oneness question, observe that the equation $A\bar{x} = \bar{b}$ has a free variable (look at column 3). Hence there are infinitely many solutions to this equation for each $\bar{b} \in \mathbb{R}^3$, which is to say that each element of $\mathbb{R}^3$ is the image of many points in $\mathbb{R}^4$ under T. Hence T is *not* one-to-one.

When is T one-to-one? For a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A, this is tantamount to figuring out if the homogeneous system $A\bar{x} = \bar{0}$ has only the trivial solution.

Theorem Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with standard matrix A. Then T is one-to-one if and only if $T(\bar{x}) = \bar{0}$ has *only* the trivial solution. In other words, T is one-to-one if and only if the columns of A are linearly independent.

Similarly, it is useful to have a condition for determining when a linear transformation T is onto. The following theorem is one way to settle this question.

Theorem Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with A its standard matrix. Then T is onto if and only if the columns of A span $\mathbb{R}^m$.

Example Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be given by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_2 \\ x_1 - x_2 \end{bmatrix}.$$

Is T one-to-one? Is it onto? Rewriting the right-hand side of the equation, we see

$$\begin{bmatrix} x_1 + x_2 \\ x_2 \\ x_1 - x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 0x_1 + x_2 \\ x_1 - x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

from which we conclude that the standard matrix of T is

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

Since the two columns are not multiples of each other, the columns are linearly independent. Hence T is one-to-one. As for onto, recall that the columns of an $m \times n$ matrix span $\mathbb{R}^m$ only if there is a pivot position in every row. In this case, we have 3 rows but only 2 columns, so it is not possible to have pivots in every row. Hence T is *not* onto.

Example The following example is to be worked on during class (time-permitting):

(Inspired by the Practice Problem on Page 77)
Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that first performs a horizontal shear sending $\bar{e}_1$ to itself, and $\bar{e}_2$ to $\bar{e}_2 + \frac{1}{2}\bar{e}_1$, and then reflects the result over the x_1 axis. Assuming T is a linear transformation, write the standard matrix of T.