More on linear independence

Last class we saw that a set of vectors $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \ldots,\ \mathbf{v}_k\}$ is linearly dependent if there is a nontrivial dependence relation $r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \ldots + r_k\mathbf{v}_k = \mathbf{0}$ among them. Otherwise the set is linearly independent.

Theorem. A nonzero set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors is linearly dependent if and only if, for some index j, the vector $\mathbf{v}_j$ is a linear combination of the vectors $\mathbf{v}_1$, …, $\mathbf{v}_{j-1}$.

Theorem. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, then $k \leq n$.

Example. We know that the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} \right\}$$

is linearly dependent. What are all possible dependence relations among this set of vectors?

Linear transformations

In order to understand the definition of a linear transformation, let's start with some examples of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. (As we shall see, **not all of these examples are linear transformations.**)

Examples: functions $f : \mathbb{R} \to \mathbb{R}$

1. $f_1(x) = 2x$
2. $f_2(x) = 2x + 1$
3. $f_3(x) = x^2$
4. $f_4(x) = \cos x$

Examples: functions $g : \mathbb{R}^2 \to \mathbb{R}^2$

1. $g_1(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2)$
2. $g_2(x_1, x_2) = (\cos(x_1 + x_2), x_1 + x_2^2)$

Examples: functions h defined on $\mathbb{R}^3$

1. $h_1(x_1, x_2, x_3) = (x_1 + x_3, x_1 - x_2 + x_3)$
2. $h_2(x_1, x_2, x_3) = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Definition. Given a function (transformation) $T : \mathbb{R}^n \to \mathbb{R}^m$, we say that T is *linear* if

1. $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ for all $\mathbf{v}_1$ and $\mathbf{v}_2$ in $\mathbb{R}^n$, and
2. $T(r\mathbf{v}) = rT(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$ and all r in $\mathbb{R}$.

Terminology: Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$:

- $\mathbb{R}^n$ is the *domain* of T.
- Our textbook says that $\mathbb{R}^m$ is the *codomain* of T.
- The *image* or *range* of T is the set of vectors

$$\{\mathbf{w} \in \mathbb{R}^m \mid T(\mathbf{v}) = \mathbf{w} \text{ for some } \mathbf{v} \in \mathbb{R}^n\}.$$

Example. Consider the function $T(x_1, x_2, x_3) = (x_1, 0)$. Its domain is $\mathbb{R}^3$, its codomain is $\mathbb{R}^2$, and its image is the x_1-axis in $\mathbb{R}^2$. Note that its image is not $\mathbb{R}^1$.

Basic facts about linear transformations T

1. $T(\mathbf{0}) = \mathbf{0}$
2. $T(r_1\mathbf{v}_1 + \ldots + r_k\mathbf{v}_k) = r_1T(\mathbf{v}_1) + \ldots + r_kT(\mathbf{v}_k)$

Which of the functions given above are linear?

One way to show that a transformation is linear is to verify the two conditions of linearity directly, but there is an easier way to see that these transformations are linear.

Important class of examples: Given an $m \times n$ matrix $\mathbf{A}$, then we can define a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ by the equation

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x}.$$

We know that T is a linear transformation because the matrix-vector product satisfies the necessary conditions.

Example. Let

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

Then

$$\mathbf{H} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

On the course web page, there is a link to a java program called `Matrix Machine` written by Hu Hohn. It lets you investigate the mapping properties of various 2×2 matrix transformations. Try the following three matrices:

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

There are also links to two applets by David Austin at Grand Valley State University. These two applets are particularly useful when you want to understand the mapping properties of linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Theorem. Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be written as

$$T(\mathbf{x}) = \mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ is an $m \times n$ matrix. This matrix $\mathbf{A}$ is called the *standard matrix representation* of T.

Why? Let's make two observations using the "standard basis" of $\mathbb{R}^n$.

Definition. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in $\mathbb{R}^n$ are the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \ldots \qquad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Observation 1: If we know the images $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$ for all of the standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$, then we can calculate $T(\mathbf{x})$ for any $\mathbf{x}$ in $\mathbb{R}^n$.

Observation 2: If

$$\mathbf{A} = \left[\begin{array}{c|c|c|c} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_n) \end{array} \right],$$

then

$$\mathbf{A}\mathbf{x} = x_1 T(\mathbf{e}_1) + \ldots + x_n T(\mathbf{e}_n).$$

Example. What is the matrix representation of the linear transformation that rotates $\mathbb{R}^2$ by 45° ($\pi/4$ radians) counterclockwise around the origin?