Matrices and Linear transformations

We have been thinking of matrices in connection with solutions to linear systems of equations like $Ax = b$. It is time to broaden our horizons a bit and start thinking of matrices as functions. In particular, if $A$ is $m \times n$, we can use $A$ to define a function $f_A$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ which sends $v \in \mathbb{R}^n$ to $Av \in \mathbb{R}^m$. That is, $f_A(v) = Av$.

Example: Let

$$A_{2 \times 3} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$ 

If

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,$$

then

$$f_A(v) = Av = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 4x + 5y + 6z \end{pmatrix}$$

sends the vector $v \in \mathbb{R}^3$ to $Av \in \mathbb{R}^2$.

Definition: A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be linear if

- $f(v_1 + v_2) = f(v_1) + f(v_2)$, and
- $f(cv) = cf(v)$

for all $v_1, v_2 \in \mathbb{R}^n$ and for all scalars $c$.

A linear function $f$ is also known as a linear transformation.
Examples:

- Define \( f : \mathbb{R}^3 \to \mathbb{R} \) by
  \[
  f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3x - 2y + z.
  \]

  Then \( f \) is linear because for any
  \[
  v_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \quad \text{and} \quad v_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix},
  \]
  we have
  \[
  f(v_1 + v_2) = f \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} = 3(x_1 + x_2) - 2(y_1 + y_2) + (z_1 + z_2).
  \]

  And the right hand side can be rewritten as \((3x_1 - 2y_1 + z_1) + (3x_2 - 2y_2 + z_2)\), which is the same as \(f(v_1) + f(v_2)\). So the first property holds. So does the second, since \(f(cv) = 3cx - 2cy + cz = c(3x - 2y + z) = cf(v)\).

- Notice that the function \( f \) is actually \( f_A \) for the right \( A \): if \( A_{1 \times 3} = (3, -2, 1) \), then \( f(v) = Av \).

- If \( A_{m \times n} \) is a matrix, then \( f_A : \mathbb{R}^n \to \mathbb{R}^m \) is a linear transformation because \( f_A(v_1 + v_2) = A(v_1 + v_2) = Av_1 + Av_2 \) (this is a fundamental property of matrix multiplication) \( = f_A(v_1) + f_A(v_2) \). And \( A(cv) = cAv \Rightarrow f_A(cv) = cf_A(v) \).

- Although we don't give the proof, it can be shown that any linear transformation can be written as \( f_A \) for a suitable matrix \( A \).
There are many other examples of linear transformations; some of the most interesting do not go from \( \mathbb{R}^n \) to \( \mathbb{R}^m \):

1. If \( f \) and \( g \) are differentiable functions, then

\[
\frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx}, \quad \text{and} \quad \frac{d}{dx}(cf) = c\frac{df}{dx}.
\]

The function \( D(f) = \frac{df}{dx} \) is linear.

2. If \( f \) is continuous, then we can define

\[
I f(x) = \int_0^x f(s) \, ds,
\]

and \( I \) is linear, by well-known properties of the integral.

3. The Laplace operator, \( \Delta \), defined before, is linear.

4. Let \( y \) be twice continuously differentiable and define

\[
L(y) = y'' - 2y' - 3y.
\]

Then \( L \) is linear, as you can (and should!) verify.

Linear transformations on functions, like the above, are generally known as linear operators. They're a bit more complicated than matrix multiplication operators, but they have the same essential property of linearity.

**Exercises:**

1. Give an example of a function from \( \mathbb{R}^2 \) to itself which is not linear.

2. Identify all the linear transformations from \( \mathbb{R} \) to \( \mathbb{R} \).
3. If $f: \mathbb{R}^n \to \mathbb{R}^m$ is linear then

$$\text{Ker}(f) := \{ v \in \mathbb{R}^n \text{ such that } f(v) = 0 \}$$

is a subspace of $\mathbb{R}^n$, called the kernel of $f$.

4. If $f: \mathbb{R}^n \to \mathbb{R}^m$ is linear, then

$$\text{Range}(f) = \{ y \in \mathbb{R}^m \text{ such that } y = f(v) \text{ for some } v \}$$

is a subspace of $\mathbb{R}^m$ called the range of $f$.

Remark: Everything we've been doing regarding the solution of linear systems of equations can be recast in the framework of linear transformations. In particular, if $f_A$ is multiplication by some matrix $A$, then the range of $f_A$ is just the set of all $y$ such that the linear system $Av = y$ has a solution. And the kernel of $f_A$ is the set of all solutions to the homogeneous equation $Av = 0$.

We now want to study square matrices, regarding an $n \times n$ matrix $A$ as a linear transformation from $\mathbb{R}^n$ to itself. We'll just write $Av$ for $f_A(v)$ to simplify the notation, and to keep things really simple, we'll just talk about $2 \times 2$ matrices -- all the problems that exist in higher dimensions are present in $\mathbb{R}^2$.

There are several questions that present themselves:

- Can we visualize the linear transformation $x \to Ax$? One thing we can't do is draw a graph! Why not?

- Connected with the first question is: can we choose a better coordinate system in which to view the problem?
The answer is not an unequivocal "yes" to either of these, but we can generally do some useful things.

To pick up at the end of the last lecture, note that when we write \( f_A(v) = y = Av \), we are actually using the coordinate vector of \( v \) in the standard basis. Suppose we change the basis to some other basis \( \{e_1, e_2\} \) using the invertible matrix \( E \). Then we can rewrite the equation in the new coordinates and basis:

We have \( v = E v_e \), and \( y = E y_e \), so

\[
\begin{align*}
y &= Av \\
Ey_e &= AEv_e, \text{ and} \\
y_e &= E^{-1}AEv_e
\end{align*}
\]

That is, the matrix equation \( y = Av \) is given in the new basis by the equation

\[
y_e = E^{-1} AE v_e \text{ or } y_e = A_e v_e, \text{ where } A_e = E^{-1} AE.
\]

We can now restate the second question: Can we find a nonsingular matrix \( E \) so that \( E^{-1} AE \) is particularly "nice"? The primary example of such a "nice" matrix is a diagonal matrix.

**Definition:** The matrix \( A \) is **diagonal** if the only nonzero entries lie on the main diagonal. That is, \( a_{ij} = 0 \) if \( i \neq j \).

**Example:**

\[
A = \begin{pmatrix}
4 & 0 \\
0 & -3
\end{pmatrix}
\]

is diagonal. This is nice because we can (partially) visualize the linear transformation corresponding to multiplication by \( A \): a vector \( v \) lying along the first coordinate
axis mapped to $4\mathbf{v}$, a multiple of itself. A vector $\mathbf{w}$ lying along the second coordinate axis is also mapped to a multiple of itself: $A\mathbf{w} = -3\mathbf{w}$. It's length is tripled, and its direction is reversed.

Now it turns out that we can find vectors like $\mathbf{v}$ and $\mathbf{w}$, which are mapped to multiples of themselves, without first finding the matrix $E$. This is the subject of the next few sections.

**Eigenvalues and eigenvectors**

**Definitions:** If a vector $\mathbf{v} \neq \mathbf{0}$ satisfies the equation $A\mathbf{v} = \lambda \mathbf{v}$, for some real number $\lambda$, then $\lambda$ is said to be an *eigenvalue* of the matrix $A$, and $\mathbf{v}$ is said to be an *eigenvector* of $A$ corresponding to $\lambda$.

**Example:** If

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}, \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

then

$$A\mathbf{v} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5\mathbf{v}.$$  

So $\lambda = 5$ is an eigenvalue of $A$, and $\mathbf{v}$ an eigenvector corresponding to this eigenvalue.

**Remark:** Note that the definition of eigenvector requires that $\mathbf{v} \neq \mathbf{0}$. The reason for this is that if $\mathbf{v} = \mathbf{0}$ were allowed, then any number $\lambda$ would be an eigenvalue since the statement $A\mathbf{0} = \lambda\mathbf{0}$ holds for any $\lambda$. On the other hand, we can have $\lambda = 0$. See the exercise below.
Exercises:

• Show that

\[
\begin{pmatrix}
1 \\
-1
\end{pmatrix}
\]

is also an eigenvector of the matrix \( A \) above. What’s the eigenvalue?

• Eigenvectors are not unique. In particular, if \( v \) is an eigenvector for \( A \), then so is \( cv \), for any real number \( c \neq 0 \).

• Suppose \( \lambda \) is an eigenvalue of \( A \). Define

\[
E_\lambda = \{ v \in \mathbb{R}^n \text{ such that } Av = \lambda v \}
\]

Show that \( E_\lambda \) is a subspace of \( \mathbb{R}^n \). (N.b: the definition of \( E_\lambda \) does not require \( v \) to be an eigenvector of \( A \), so \( v = 0 \) is allowed; otherwise, it wouldn’t be a subspace.)

• \( E_0 = \text{Ker}(f_A) \) is just the null space of the matrix \( A \).

Example: The matrix

\[
A = \begin{pmatrix}
0 & -1 \\
1 & 0
\end{pmatrix} = \begin{pmatrix}
\cos(\pi/2) & -\sin(\pi/2) \\
\sin(\pi/2) & \cos(\pi/2)
\end{pmatrix}
\]

represents a counterclockwise rotation through the angle \( \pi/2 \). Apart from 0, there is no vector which is mapped by \( A \) to a multiple of itself. So not every matrix has eigenvectors.