Eigenvectors
Lay 5.1

November 4, 2013

1 From last time

Remember from last time we were interested in the stationary distribution $\mathbf{q}$ of the Markov chain transition matrix P:

$$P\mathbf{q} = \mathbf{q}. \tag{1}$$

Notice that the action of P on $\mathbf{q}$ is very simple. Today's philosophy will be: given a matrix A, can we find vectors that A acts on in a "simple" way?

2 Eigenvectors

Let A be an $n \times n$ matrix.

Definition 2.1. We say a nonzero vector $\mathbf{x}$ is an eigenvector of A if there is some scalar λ such that $A\mathbf{x} = \lambda x$. We call λ an eigenvalue of A. Then $\mathbf{x}$ is said to be the eigenvector corresponding to the eigenvalue λ.

Note that in the above definition, it is important that A be square.

Example 2.2. Let

$$A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}.$$

Then 3 and -2 are eigenvalues of A. Indeed, an eigenvector corresponding to 3 is

$$\begin{bmatrix} -4 \\ 1 \end{bmatrix},$$

and an eigenvector corresponding to -2 is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Check this!

Notice that any (nonzero) scalar multiple of an eigenvector is an eigenvector corresponding to the same eigenvalue.

2.1 Finding eigenvectors

We will talk next time about how to find the eigenvalues of a matrix. Today we will talk about a less ambitious goal: if you are given an eigenvalue of the matrix, how do you find an eigenvector?

The answer is as follows. If λ is an eigenvalue of A, then the corresponding eigenvector $\mathbf{x}$ satisfies

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Subtracting from both sides gives

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

So if you know that λ is an eigenvalue, you can find the eigenvectors corresponding to λ by solving the above linear system.

Example 2.3. Consider A from the last example. We know that 3 is an eigenvalue. To find the eigenvectors corresponding to this eigenvalue, we find the vectors $\mathbf{x}$ satisfying

$$\begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix} \mathbf{x} = \mathbf{0}.$$

It is simple to check that all $\mathbf{x}$ satisfying this equation must be multiples of $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$.

2.2 Eigenspaces

We have noted already that any multiple of an eigenvector corresponding to an eigenvalue λ is also an eigenvector corresponding to λ. In fact, more is true:

Theorem 2.4. *Let A be $n \times n$. If λ is an eigenvalue of A, then the union of the zero vector with the set of eigenvectors corresponding to λ forms a subspace, the "eigenspace of A corresponding to λ".*

Proof. This is immediate, because the set of eigenvectors corresponding to λ is the nullspace of $A - \lambda I$, excluding the zero vector. So adding the zero vector to this set of eigenvectors gives us $\text{nul}(A - \lambda I)$, which is a subspace.

□

We have so far only seen one-dimensional eigenspaces. Of course, we can have bigger ones, as shown by the following trivial example:

Example 2.5. Let

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then A has the eigenvalues 2 and 1. A basis for the eigenspace corresponding to 2 is $\{\mathbf{e}_1, \mathbf{e}_2\}$.

Example 2.6. Consider the matrix

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Find a basis for the eigenspace corresponding to the eigenvalue 1.

We subtract off $1I = I$ from the above matrix, and use the standard method for finding the basis of a nullspace. The augmented matrix is

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

The RREF of the above matrix is

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

So the nullspace is the set

$$\left\{ \begin{bmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} : x_2 \in \mathbb{R},\ x_3 \in \mathbb{R} \right\}$$

and so a basis is given by

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Please, if you don't understand how the basis was found, look back in Lay to the section where he describes finding the basis for a null/column space!

3 Eigenvalues of triangular matrices

Theorem 3.1. *The eigenvectors of an upper- or lower-triangular matrix are the entries along the main diagonal.*

Proof. See Lay for a full proof. The main idea is to look at the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$, which has a nontrivial solution if and only if the equation has a free variable, which happens if and only if one of the diagonal entries of $A - \lambda I$ is zero.

□

4 Eigenvectors and independence

The main point of the following theorem is that no eigenvector corresponding to an eigenvalue λ can be written as a linear combination of eigenvectors corresponding to eigenvalues other than λ:

Theorem 4.1. *If $\mathbf{x}_1, \ldots \mathbf{x}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of the $n \times n$ matrix A, then the set*

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$$

is linearly independent.

Proof. Assume (for the sake of contradiction) that the set is linearly dependent. Then there exists some largest number p such that the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ is linearly independent. Then there are scalars $c_1, \ldots, c_{p+1}$ not all zero such that

$$c_1\mathbf{x}_1 + \ldots + c_{p+1}\mathbf{x}_{p+1} = \mathbf{0}. \tag{2}$$

Multiplying both sides by A gives

$$c_1\lambda_1\mathbf{x}_1 + \ldots + c_{p+1}\lambda_{p+1}\mathbf{x}_{p+1} = \mathbf{0}.$$

Multiply (2) by $-\lambda_{p+1}$ and add it to the above to see

$$c_1(\lambda_1 - \lambda_{p+1}\mathbf{x}_1 + \ldots + c_p(\lambda_p - \lambda_{p+1})\mathbf{x}_p = \mathbf{0}.$$

Since each $\lambda_i - \lambda_{p+1}$ appearing above is nonzero (by distrinct eigenvalues), linear independence of $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ gives that $c_1 = 0, c_2 = 0, \ldots, c_p = 0$. But then (2) says that $\mathbf{x}_{p+1} = \mathbf{0}$, a contradiction (since it is an eigenvector, it is nonzero).

□

4.1 What is to come

I will spend a couple minutes talking about the power of eigenvectors. Basically, if A is an $n \times n$ matrix whose eigenvectors form a basis for $\mathbb{R}^n$, we can describe easily its action...this will become a major theme.