Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 Eigenvectors and eigenvalues 1

2 The characteristic equation 3

1 Eigenvectors and eigenvalues

Lay–Lay–McDonald §5.1 pp. 268 – 273

Although a transformation $\mathbf{x} \mapsto A\mathbf{x}$ may transform vectors in a number of directions, it often happens that there are special vectors on which the action of A is simple. Our discussion in this chapter will be about square matrices.

Example 1. *Let* $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. *Compute* $A\mathbf{u}$ *and* $A\mathbf{v}$.

$Au = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}$; $Av = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Today we will study equations of the form $A\mathbf{x} = \lambda\mathbf{x}$ where special vectors are transformed by A into scalar multiples of themselves.

Definition 2. *An eigenvector of an* $n \times n$ *matrix* A *is a nonzero vector* $\mathbf{x}$ *such that* $A\mathbf{x} = \lambda\mathbf{x}$ *for some scalar* λ. *A scalar* λ *is called an eigenvalue of* A *if there is a nontrivial solution* $\mathbf{x}$ *of* $A\mathbf{x} = \lambda\mathbf{x}$. *Such a* $\mathbf{x}$ *is called an eigenvector corresponding to* λ.

Example 3. *Let* $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. *Are* $\mathbf{u}$ *and* $\mathbf{v}$ *eigenvectors of* A*?*

$Au = \lambda u$ or $Av = \lambda v$

$Au = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 6 \\ -5 \end{pmatrix} = \begin{pmatrix} -24 \\ 20 \end{pmatrix} = -4\begin{pmatrix} 6 \\ -5 \end{pmatrix} \Rightarrow u$ is an eigenvector of A (with eigenvalue -4)

$Av = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -9 \\ 11 \end{pmatrix} \neq \lambda v \Rightarrow v$ is not an eigenvector of A.

Example 4. *Show that 7 is an eigenvalue of the matrix* A *in the previous example, and find the corresponding eigenvectors.*

$Ax = 7x$

$\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 7\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{matrix} x_1 + 6x_2 = 7x_1 \\ 5x_1 + 2x_2 = 7x_2 \end{matrix} \Rightarrow \begin{matrix} -6x_1 + 6x_2 = 0 \\ 5x_1 - 5x_2 = 0 \end{matrix}$

$\Rightarrow \begin{matrix} x_1 - x_2 = 0 \\ x_1 - x_2 = 0 \end{matrix}$

$\begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow x_1 = x_2 \Rightarrow \begin{pmatrix} x_1 \\ x_1 \end{pmatrix}$

all solutions look like $\begin{pmatrix} x_1 \\ x_1 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\Rightarrow$ any vector of this form is an eigenvector.

Remark 5. *Note that while row reduction can be used to find eigenvectors, it cannot be used to find eigenvalues. An echelon form of a matrix A usually does not display the eigenvalues of A.*

Note that λ is an eigenvalue of an $n \times n$ matrix A if and only if the equation

$$(A - \lambda I)\mathbf{x} = \mathbf{0} \iff Ax - \lambda I x = 0 \iff Ax = \lambda x \tag{1}$$

has a nontrivial solution. The set of all solutions of (1) is just the null space of the matrix $A - \lambda I$. So this is a *subspace* of $\mathbb{R}^n$ and is called the *eigenspace of A corresponding to λ.*

Example 6. *Let* $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. *An eigenvalue of A is 2. Find a basis for the corresponding eigenspace.*

Look at $A-2I$ and compute nullspace:

$$A - 2I = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

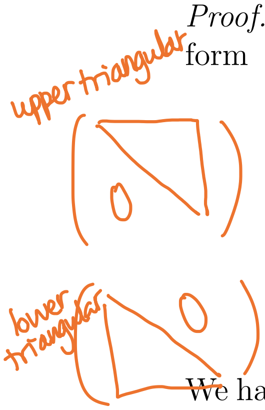
$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \text{pivot} & \text{free} & \end{array}$$

$2x_1 - x_2 + 6x_3 = 0$

$x_1 = \frac{x_2}{2} - 3x_3 \Rightarrow \begin{pmatrix} \frac{x_2}{2} - 3x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \Rightarrow$ basis is given $\left\{ \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Here is one situation where it's easy to compute eigenvalues:

Theorem 7. *The eigenvalues of a triangular matrix are the entries on its main diagonal.*



Proof. For simplicity, consider the 3×3 case. If A is upper-triangular, then $A - \lambda I$ has the form

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}.$$

We have that λ is an eigenvalue of A if and only if the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution, i.e., if and only if the equation has a free variable. We see that $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a free variable if and only if at least one of the entries on the diagonal of $A - \lambda I$ is zero. This happens if and only if λ is one of the entries a_{11}, a_{22}, a_{33} in A.
We leave the case of lower-triangular matrices as an exercise. □

Example 8. *Let* $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ *and* $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. *What are the eigenvalues of A? What are the eigenvalues of B?*

eigenvalues of A: 3, 0, 2

eigenvalues of B: 1, 4

What does it mean for a matrix A to have an eigenvalue of 0? This happens if and only if the equation $A\mathbf{x} = 0\mathbf{x}$ has a nontrivial solution. But this is equivalent to $A\mathbf{x} = \mathbf{0}$, which has a nontrivial solution if and only if A is not invertible. *Thus 0 is an eigenvalue of A if and only if A is not invertible.*

Here is an important result about eigenvectors that we will record for later use:

Theorem 9. *If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.*

2 The characteristic equation

Lay–Lay–McDonald §5.2 pp. 276 – 278

Now we practice finding eigenvalues:

Example 10. *Find the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.* want to solve for λ s.t. $(A-\lambda I)x = 0$ has nontrivial solutions

By Invertible Matrix Thm means that $(A-\lambda I)$ is not invertible $\Rightarrow \det(A-\lambda I) = 0$.

So $\det(A-\lambda I) = \begin{vmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix} = (2-\lambda)(-6-\lambda) - 9$

$= -12 + 6\lambda - 2\lambda + \lambda^2 - 9$

$= \lambda^2 + 4\lambda - 21 = (\lambda-3)(\lambda+7) \Rightarrow \lambda = 3$ or -7.

$\Rightarrow$ eigenvalues are $3, -7$.

We have the following important result: *a scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation*

$$\det(A - \lambda I) = 0.$$

Example 11. *Find the characteristic equation of $A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.*

$\det(A-\lambda I) = 0$

$= \begin{vmatrix} 5-\lambda & & * & \\ & 3-\lambda & & \\ & 0 & 5-\lambda & \\ & & & 1-\lambda \end{vmatrix} = \boxed{(5-\lambda)^2(3-\lambda)(1-\lambda) = 0.}$

(det of upper triangular matrix is product of diagonal entries)

It can be shown that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n called the *characteristic polynomial* of A.

Example 12. *The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.*

$\lambda^6 - 4\lambda^5 - 12\lambda^4 = 0$

$\lambda^4(\lambda^2 - 4\lambda - 12) = 0$

$\lambda^4(\lambda - 6)(\lambda + 2) = 0 \Rightarrow$ $\lambda = 0$ (multiplicity 4)
$\lambda = 6$ (multiplicity 1)
$\lambda = -2$ (multiplicity 1)

($\lambda = 0$ tells you the matrix is not invertible)

The next theorem presents one use of the characteristic polynomial and is helpful for iterative methods that approximate eigenvalues. We begin with some terminology. If A and B are $n \times n$ matrices, then we say that A is *similar to* B if there is an invertible matrix P such that

$$P^{-1}AP = B.$$

Writing $Q := P^{-1}$, we also have

$$Q^{-1}BQ = A.$$

So B is also similar to A, and we say that A *and* B *are similar.*

Theorem 13. *If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues with the same multiplicities.*

Proof. If $B = P^{-1}AP$ then

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P.$$

We compute

$$\begin{aligned} \det(B - \lambda I) &= \det(P^{-1}(A - \lambda I)P) \\ &= \det(P^{-1})\det(A - \lambda I)\det(P). \end{aligned}$$

Since $\det(P^{-1})\det(P) = \det(P^{-1}P) = \det(I) = 1$, we see that $\det(B - \lambda I) = \det(A - \lambda I)$. □

Remark 14. *Note that matrices that have the same eigenvalues might not be similar: for instance, the matrices $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ have the same eigenvalues but are not similar.*

Remark 15. *Similarity is not the same as row equivalence. (If A is row equivalent to B, then $B = EA$ for some invertible matrix E.) Row operations on a matrix usually change its eigenvalues.*

We can use eigenvalues and eigenvectors to analyze the evolution of a dynamical system.

Example 16. *Let* $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. *Analyze the long-term behavior of the dynamical system defined by* $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ *with* $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

$x_1 = Ax_0$

$x_2 = Ax_1 \Rightarrow x_2 = A(Ax_0) = A^2 x_0$

$\vdots$

$x_k = A^k x_0$

First step: compute eigenvalues of A:

Idea: If $Av = \lambda v$, then can calculate $\lim_{k\to\infty} A^k v = \lim_{k\to\infty} \lambda^k v$.

$$\det(A - \lambda I) = \begin{vmatrix} 0.95-\lambda & 0.03 \\ 0.05 & 0.97-\lambda \end{vmatrix} = (0.95-\lambda)(0.97-\lambda) - 0.03(0.05)$$
$$= \lambda^2 - 1.92\lambda + 0.92 = (\lambda - 1)(\lambda - 0.92) \Rightarrow \lambda = 1, 0.92.$$

For each eigenvalue, find eigenvectors:

$$\begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} v_1 = \underset{\lambda_1}{1} v_1$$

$$\Rightarrow \begin{pmatrix} -0.05 & 0.03 \\ 0.05 & -0.03 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 0.95 & 0.03 \\ 0.05 & 0.97 \end{pmatrix} v_2 = \underset{\lambda_2}{0.92} v_2$$

$$\begin{pmatrix} 0.03 & 0.03 \\ 0.05 & 0.05 \end{pmatrix} \Rightarrow v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

in terms of eigenvectors

Now write $x_0 = av_1 + bv_2$ (since we understand how A transforms v_1, v_2, but Ax_0 is more difficult to directly analyze)

$$\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = a\begin{pmatrix} 3 \\ 5 \end{pmatrix} + b\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = \frac{1}{-8} \cdot \begin{pmatrix} -1 & -1 \\ -5 & 3 \end{pmatrix}\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = -\frac{1}{8}\begin{pmatrix} -1 \\ -1.8 \end{pmatrix}$$
$$= \begin{pmatrix} 0.125 \\ 0.225 \end{pmatrix}$$

So $x_0 = 0.125 v_1 + 0.225 v_2$

$$\Rightarrow x_1 = Ax_0 = A(0.125 v_1 + 0.225 v_2)$$
$$= 0.125\, Av_1 + 0.225\, Av_2$$
$$= 0.125 \lambda_1 v_1 + 0.225 \lambda_2 v_2$$

$$\Rightarrow x_2 = Ax_1 = A(0.125\lambda_1 v_1 + 0.225 \lambda_2 v_2)$$
$$= 0.125 \lambda_1 Av_1 + 0.225 \lambda_2 Av_2 = 0.125 \lambda_1^2 v_1 + 0.225 \lambda_2^2 v_2$$

$\vdots$

$$\Rightarrow x_k = 0.125 \lambda_1^k v_1 + 0.225 \lambda_2^k v_2$$

$$\boxed{x_k = 0.125 v_1 + 0.225 (0.92)^k v_2}$$

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty} (0.125 v_1 + \cancel{0.225 (0.92)^k v_2}) = 0.125 v_1.$$

$$= 0.125\begin{pmatrix} 3 \\ 5 \end{pmatrix}.$$