Lecture 9

Eigenvector bases

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

(a) If $u_1, \dots, u_k$ are eigenvectors, with eigenvalues $\lambda_1, \dots, \lambda_k$, and these eigenvalues are all different, then the vectors $u_1, \dots, u_k$ are ______.

(b) The eigenvalues are the roots of $\chi_A(t)$, which is a polynomial of degree n. Thus, there are at most n different eigenvalues.

(c) Suppose there are exactly n distinct eigenvalues, say $\lambda_1, \dots, \lambda_n$. We can then choose an eigenvector u_i for each eigenvalue λ_i, and part (a) says that the list $\mathcal{U} = u_1, \dots, u_n$ is ______. As $\mathcal{U}$ is an ______ list of n vectors in $\mathbb{R}^n$, it is in fact a ______.

Eigenvector basis example

Consider $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, so $\chi_A(t) = \det(A - tI) = (\_\_\_)(\_\_\_)(\_\_\_)$,
so the eigenvalues are _, _ and _. Suppose we have eigenvectors u_1, u_2 and u_3, where u_k has eigenvalue k. By the previous slide: the list u_1, u_2, u_3 is a basis for $\mathbb{R}^3$. We can find the eigenvectors explicitly by row-reduction:

$$\_\_\_ = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\_\_\_ = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\_\_\_ = \begin{bmatrix} -2 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -3/2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \qquad u_3 = \begin{bmatrix} 3/2 \\ 2 \\ 1 \end{bmatrix}.$$

We can check more directly that the u_i form a basis :

$$[u_1|u_2|u_3] = \begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

Eigenvector basis example

Consider $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, so $\chi_A(t) = \det \begin{bmatrix} -t & 1 \\ -1 & -t \end{bmatrix} = t^2 + 1$.
For all $t \in \mathbb{R}$ we have $t^2 + 1 \geq 1 > 0$, so the characteristic polynomial has no real roots, so there are no real eigenvalues or eigenvectors.

However, there are complex eigenvalues i and $-i$, with corresponding eigenvectors $u_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ and $u_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$, which form a basis for $\mathbb{C}^2$.

This example and the previous one are typical. If we pick an $n \times n$ matrix at random, it will usually have n different eigenvalues (some of which will usually be complex), and so the corresponding eigenvectors will form a basis for $\mathbb{C}^n$. However, there are some exceptions, as we will see soon. Such exceptions usually arise because of some symmetry or other interesting feature of the problem that gives rise to the matrix.

Eigenvector basis example

Consider $A = \begin{bmatrix} 5 & 5 & 0 \\ 0 & 5 & 5 \\ 0 & 0 & 5 \end{bmatrix}$, so $\chi_A(t) = (5-t)^3$, so the only eigenvalue is 5.

The eigenvectors are the solutions of $\begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, which reduces

to $5y = 5z = 0$ so $y = z = 0$, so the eigenvectors are the multiples of $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

This means that any two eigenvectors are multiples of each other, and so are linearly dependent. Thus, we cannot find a basis consisting of eigenvectors.

Eigenvector basis example

Consider $A = \begin{bmatrix} 0 & 0 & 5 \\ 0 & 5 & 0 \\ 5 & 0 & 0 \end{bmatrix}$. The characteristic polynomial is

$$\chi_A(t) = \det \begin{bmatrix} -t & 0 & 5 \\ 0 & 5-t & 0 \\ 5 & 0 & -t \end{bmatrix} = -t \det \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} + 5 \det \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}$$

$$= t^2(5-t) - 25(5-t) = -(t-5)(t^2-25) = -(t-5)(t-5)(t+5)$$

$$= -(t-5)^2(t+5).$$

The only eigenvalues (even in $\mathbb{C}$) are and . As there are only two distinct eigenvalues, we do not *automatically* have a basis of eigenvectors. However, it turns out that there is a basis of eigenvectors anyway. Indeed, we can take

$$u_1 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T \qquad u_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \qquad u_3 = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T.$$

We can check that $Au_1 = 5u_1$ and $Au_2 = 5u_2$ and $Au_3 = -5u_3$, so the u_i are eigenvectors with eigenvalues 5, 5 and -5 respectively. We can also check that the u_i form a basis, either by row-reducing $[u_1|u_2|u_3]$ or using the formula

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (x+z)/2 \\ 0 \\ (x+z)/2 \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} (x-z)/2 \\ 0 \\ (z-x)/2 \end{bmatrix} = \frac{x+z}{2} u_1 + y u_2 + \frac{x-z}{2} u_3.$$

Diagonalisation

Definition 14.1: We write $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ for the $n \times n$ matrix such that the entries on the diagonal are $\lambda_1, \dots, \lambda_n$ and the entries off the diagonal are zero.

Example 14.2: $\mathrm{diag}(\ ,\ ,\ ,\ ) = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}$

Definition 14.3: Let A be an $n \times n$ matrix.

- To *diagonalise* A means to give an invertible matrix U and a diagonal matrix D such that $U^{-1}AU = D$ (or equivalently $A = UDU^{-1}$).
- We say that A is *diagonalisable* if there exist matrices U and D with these properties.

Diagonalisation and eigenvectors

Proposition 14.4: Suppose we have a basis $u_1, \dots, u_n$ for $\mathbb{R}^n$ such that each vector u_i is an eigenvector for A, with eigenvalue λ_i say.

Put $U = [u_1|\cdots|u_n]$ and $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Then $U^{-1}AU = D$, so we have a diagonalisation of A.

Moreover, every diagonalisation of A occurs in this way.
The proof will be given after a lemma.

A matrix multiplication lemma

Lemma 14.5: Let A and U be $n \times n$ matrices, let $\lambda_1, \dots, \lambda_n$ be real numbers, and put $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Let $u_1, \dots, u_n$ be the columns of U. Then

$$AU = \left[\begin{array}{c|c|c} Au_1 & \cdots & Au_n \end{array}\right] \qquad UD = \left[\begin{array}{c|c|c} \lambda_1 u_1 & \cdots & \lambda_n u_n \end{array}\right].$$

Proof: Let the rows of A be $a_1^T, \dots, a_n^T$. By the definition of matrix multiplication, we have

$$AU = \begin{bmatrix} a_1.u_1 & \cdots & a_1.u_n \\ \cdots & \cdots & \cdots \\ a_n.u_1 & \cdots & a_n.u_n \end{bmatrix}$$

The p'th column is $\begin{bmatrix} a_1.u_p \\ \vdots \\ a_n.u_p \end{bmatrix}$, and this is just the definition of Au_p. In other words, we have

$$AU = \left[\begin{array}{c|c|c} Au_1 & \cdots & Au_n \end{array}\right]$$

A matrix multiplication lemma

Lemma 14.5: Let A and U be $n \times n$ matrices, let $\lambda_1, \dots, \lambda_n$ be real numbers, and put $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Let $u_1, \dots, u_n$ be the columns of U. Then

$$AU = \left[\begin{array}{c|c|c} \quad & \cdots & \quad \end{array}\right] \qquad UD = \left[\begin{array}{c|c|c} \quad & \cdots & \quad \end{array}\right].$$

Proof continued: For the second claim, we just do the 3×3 case:

$$UD = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 a & \lambda_2 b & \lambda_3 c \\ \quad & \quad & \quad \\ \lambda_1 g & \lambda_2 h & \lambda_3 i \end{bmatrix}$$

Everything in the first column gets multiplied by λ_1, everything in the second column gets multiplied by λ_2 and everything in the third column gets multiplied by λ_3. In other words, we have

$$\left[\begin{array}{c|c|c} u_1 & u_2 & u_3 \end{array}\right] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \left[\begin{array}{c|c|c} \lambda_1 u_1 & \lambda_2 u_2 & \lambda_3 u_3 \end{array}\right]$$

as claimed.

Diagonalisation and eigenvectors

Proposition 14.4: Suppose we have a basis $u_1, \dots, u_n$ for $\mathbb{R}^n$ such that each vector u_i is an eigenvector for A, with eigenvalue λ_i say. Put $U = [u_1 | \cdots | u_n]$ and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Then $U^{-1}AU = D$, so we have a diagonalisation of A. Moreover, every diagonalisation arises in this way.

Proof.

- The columns u_i of U form a basis for $\mathbb{R}^n$, so U is invertible.
- First half of the lemma: $AU = [Au_1 | \cdots | Au_n]$. But u_i is an eigenvector of eigenvalue λ_i, so $Au_i = \lambda_i u_i$, so $AU = [\lambda_1 u_1 | \cdots | \lambda_n u_n]$.
- Second half of the lemma: $UD = [\lambda_1 u_1 | \cdots | \lambda_n u_n]$. So $AU = UD$.
- It follows that $U^{-1}AU = U^{-1}UD = D$ and $UDU^{-1} = AUU^{-1} = A$.

Conversely: suppose we have an invertible matrix U and a diagonal matrix $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ such that $U^{-1}AU = D$. Let $u_1, \dots, u_n$ be the columns of U. By reversing the above steps: u_i is an eigenvector of eigenvalue λ_i, and $u_1, \dots, u_n$ is a basis for $\mathbb{R}^n$. □

Diagonalisation example

Example 13.23: the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$ and $\lambda_3 = 3$; and eigenvectors

$$u_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \qquad u_2 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T \qquad u_3 = \begin{bmatrix} 3/2 & 2 & 1 \end{bmatrix}^T.$$

It follows that $A = UDU^{-1}$, where

$$U = \left[\begin{array}{c|c|c} u_1 & u_2 & u_3 \end{array}\right] = \begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{bmatrix}; \qquad U^{-1} = \begin{bmatrix} 1 & \quad & \quad \\ 0 & 1 & \quad \\ 0 & 0 & 1 \end{bmatrix}.$$

We thus have

$$DU^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{bmatrix}$$

$$UDU^{-1} = \begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1/2 \\ 0 & 2 & -4 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix} = A.$$

Diagonalisation example

In Example 13.24 we showed that the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ does not have any real eigenvalues or eigenvectors, but that over the complex numbers we have eigenvectors $u_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ and $u_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$ with eigenvalues $\lambda_1 = i$ and $\lambda_2 = -i$.
We thus have a diagonalisation $A = UDU^{-1}$ with

$$U = [u_1 | u_2] = \begin{bmatrix} & \\ & \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix}.$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad U^{-1} = \frac{1}{-2i} \begin{bmatrix} -i & -1 \\ -i & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{bmatrix}.$$

This gives

$$UDU^{-1} = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{bmatrix} = \begin{bmatrix} i & -i \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

As expected, this is the same as A.

Non-diagonalisation example

Consider the matrix $A = \begin{bmatrix} 5 & 5 & 0 \\ 0 & 5 & 5 \\ 0 & 0 & 5 \end{bmatrix}$.

The characteristic poly is $(t-5)^3$, so the only eigenvalue is $\lambda = 5$.

Any eigenvector $u = \begin{bmatrix} x & y & z \end{bmatrix}^T$ must satisfy $(A - 5I_3)u = 0$ so

$$\begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{aligned} 5y &= 0 \\ 5z &= 0 \\ 0 &= 0 \end{aligned}$$

so $u = \begin{bmatrix} x & 0 & 0 \end{bmatrix}^T$.

It follows that there is no basis of eigenvectors, so A is not diagonalisable.

It is possible to understand non-diagonalisable matrices using the theory of "Jordan blocks". However, we will not cover Jordan blocks in this course.