TLDR

Quick summary of today's notes. Lecture starts on next page.

- Let A be an $n \times n$ matrix. Let $I = I_n$ be the $n \times n$ identity matrix.

  Let λ be a number and suppose $0 \neq v \in \mathbb{R}^n$.

  If $Av = \lambda v$ then we say that v is an *eigenvector* for A and that λ is an *eigenvalue* for A.

- If A and B are $n \times n$ matrices and there exists an invertible $n \times n$ matrix P with

  $$A = PBP^{-1}$$

  then we say that A is *similar* to B and that B is *similar* to A.

- Similar matrices have the same characteristics equations and same eigenvalues.

- A is *diagonalizable* if A is similar to a diagonal matrix D.

  An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

  An $n \times n$ matrix with n distinct eigenvalues is always diagonalizable.

- Suppose an $n \times n$ matrix A has $p \leq n$ distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$.

  Then A is diagonalizable if and only if

  $$\dim \mathrm{Nul}(A - \lambda_1 I) + \dim \mathrm{Nul}(A - \lambda_2 I) + \cdots + \dim \mathrm{Nul}(A - \lambda_p I) = n.$$

  Assume this holds. Suppose $\mathcal{B}_i$ is a basis for $\mathrm{Nul}(A - \lambda_i I)$.

  Then the union $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_p$ is a set of n linearly independent eigenvectors for A.

  If the elements of this union are the vectors $v_1, v_2, \dots, v_n$ then the matrix

  $$P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

  is invertible and $D = P^{-1}AP$ is diagonal, and $A = PDP^{-1}$.

1 Last time: similar and diagonalizable matrices

Let n be a positive integer. Suppose A is an $n \times n$ matrix, $v \in \mathbb{R}^n$, and $\lambda \in \mathbb{R}$.

Let I be the $n \times n$ identity matrix.

Recall that v an *eigenvector* for A with *eigenvalue* λ if $v \neq 0$ and $Av = \lambda v$, or equivalently if v is a nonzero element of $\mathrm{Nul}(A - \lambda I)$. The number λ is an eigenvalue of A if there exists some eigenvector with this eigenvalue.

If the nullspace $\mathrm{Nul}(A - \lambda I)$ is nonzero, then it is called the λ-*eigenspace* of A.

The eigenvalues of A are the solutions to the polynomial equation $\det(A - xI) = 0$.

Important fact. Any set of eigenvectors of A with all distinct eigenvalues is linearly independent.

Two $n \times n$ matrices A and B are *similar* if there is an invertible $n \times n$ matrix P such that $A = PBP^{-1}$.

Fact. Similar matrices have the same eigenvalues but usually different eigenvectors.

Proof. If $A = PBP^{-1}$ then $A - xI = PBP^{-1} - xI = P(B - xI)P^{-1}$ since $PIP^{-1} = I$.

Note that $\det(P)\det(P^{-1}) = \det(PP^{-1}) = \det(I) = 1$.

Therefore $\det(A - xI) = \det(P)\det(B - xI)\det(P^{-1}) = \det(B - xI)$. □

Example. The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is similar to $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$.

Caution. Matrices may have the same eigenvalues but not be similar.

The implication goes in one direction only: *similar* $\Rightarrow$ *same eigenvalues.*

For example, the matrices

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

both have eigenvalues $2, 2$ but are not similar. This holds because $A = 2I$ so we have

$$PAP^{-1} = 2PIP^{-1} = 2PP^{-1} = 2I = A \neq B$$

for all invertible 2×2 matrices P.

Caution. Row equivalence of matrices $\neq$ similarity of matrices.

Row operations usually change eigenvalues, whereas similar matrices always have the same eignenvalues.

2 Diagonalizable matrices

A matrix is *diagonal* if all of its nonzero entries appear in diagonal positions $(1, 1), (2, 2), \ldots,$ or (n, n).

A matrix A is *diagonalizable* if it is similar to a diagonal matrix.

In other words, A is diagonalizable if we can write $A = PDP^{-1}$ where

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

is a diagonal matrix. In this case some stronger properties hold:

- The diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$ of D are the eigenvalues of A.
- If $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ are the columns of $P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$ then $Av_i = \lambda_i v_i$.
- In particular, the columns of P are a basis for $\mathbb{R}^n$ of eigenvectors of A.

Theorem. An $n \times n$ matrix A is diagonalizable if and only if $\mathbb{R}^n$ has a basis whose elements are all eigenvectors of A.

Proof. We have just seen that if $A = PDP^{-1}$ where D is diagonal then the columns of P are a basis for $\mathbb{R}^n$ consisting of eigenvectors for A. Conversely, suppose A has n linearly independent eigenvectors $v_1, v_2, \dots, v_n$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Define

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$$

and write $e_1, e_2, \dots, e_n$ for the standard basis of $\mathbb{R}^n$. Since $Pe_i = v_i$ and $P^{-1}v_i = e_i$, we have

$$P^{-1}APe_i = P^{-1}Av_i = P^{-1}(\lambda_i v_i) = \lambda_i P^{-1} v_i = \lambda_i e_i.$$

This calculates the ith column of $P^{-1}AP$. Since $\lambda_i e_i$ is also the i column of the diagonal matrix D, we deduce that $P^{-1}AP = D$. Therefore $A = P(P^{-1}AP)P^{-1} = PDP^{-1}$ is diagonalizable. □

Not every matrix is diagonalizable. It takes some work to decide if a given matrix is diagonalizable. Here is one easy criterion, which is sufficient but not necessary:

Corollary. An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Proof. Suppose A has n distinct eigenvalues. Any choice of eigenvectors for A corresponding to these eigenvalues will be linearly independent, so A will have n linearly independent eigenvectors.

These eigenvectors are a basis for $\mathbb{R}^n$ since any set of n linearly independent vectors in $\mathbb{R}^n$ is a basis. □

Example. The matrix $A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$ is triangular so has eigenvalues $5, 0, -2$.

These are distinct numbers, so A is diagonalizable.

3 Diagonalizing matrices whose eigenvalues are not distinct

If an $n \times n$ matrix A has n distinct eigenvalues with corresponding eigenvectors $v_1, v_2, \dots, v_n$, then the matrix $P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$ is automatically invertible since its columns are linearly independent, and the matrix $D = P^{-1}AP$ is diagonal such that $A = PDP^{-1}$.

When A is diagonalizable but has fewer than n distinct eigenvalues, we can still build up P in such a way that P is automatically invertible and $P^{-1}AP$ is automatically diagonal.

Recall that if λ is an eigenvalue of A then $\mathrm{Nul}(A - \lambda I)$ is the λ-*eigenspace* of A.

The *(algebraic) multiplicity* of the eigenvalue λ is the largest integer $m \geq 1$ such that we can write the characteristic polynomial of A as the product $\det(A - xI) = (\lambda - x)^m p(x)$ for some polynomial $p(x)$.

For example, if $A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$ then

$$\det(A - xI) = \det \begin{bmatrix} -x & -1 \\ 1 & 2-x \end{bmatrix} = (-x)(2-x) + 1 = x^2 - 2x + 1 = (x-1)^2$$

so 1 is an eigenvalue of A with multiplicity 2.

Theorem. Let A be an $n \times n$ matrix. Suppose A has distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$ where $p \leq n$. The following properties then hold:

(a) For each $i = 1, 2, \dots, p$, it holds that $\dim \mathrm{Nul}(A - \lambda_i I)$ is at most the multiplicity of λ_i.

(b) A is diagonalizable if and only if the sum of the dimensions of the eigenspaces of A is n, i.e.:

$$\dim \mathrm{Nul}(A - \lambda_1 I) + \dim \mathrm{Nul}(A - \lambda_2 I) + \dots + \dim \mathrm{Nul}(A - \lambda_p I) = n.$$

(c) Suppose A is diagonalizable and $\mathcal{B}_i$ is a basis for the λ_i-eigenspace.

Then the union $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_p$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

If the elements of this union are the vectors $v_1, v_2, \dots, v_n$ then the matrix

$$P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

is invertible and $D = P^{-1}AP$ is diagonal, and $A = PDP^{-1}$.

Before giving the proof, we illustrate the result through an example.

Example. Consider the lower-triangular matrix

$$A = \begin{bmatrix} 5 & & & \\ 0 & 5 & & \\ 1 & 4 & -3 & \\ -1 & -2 & 0 & -3 \end{bmatrix}.$$

Its characteristic polynomial is $\det(A - xI) = (5-x)^2(-x-3)^2$.

The eigenvalues of A are therefore 5 and -3, each with multiplicity 2. Since

$$A - 5I = \begin{bmatrix} 0 & & & \\ 0 & 0 & & \\ 1 & 4 & -8 & \\ -1 & -2 & 0 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 8 & 16 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{RREF}(A - 5I)$$

it follows that $x \in \mathrm{Nul}(A - 5I)$ if and only if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -8x_3 - 16x_4 \\ 4x_3 + 4x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

so

$$\boxed{\begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix} \text{ is a basis for } \mathrm{Nul}(A-5I).}$$

Since

$$A-(-3)I = A+3I = \begin{bmatrix} 8 & & & \\ 0 & 8 & & \\ 1 & 4 & 0 & \\ -1 & -2 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{RREF}(A+3I)$$

it follows that $x \in \mathrm{Nul}(A+3I)$ if and only if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

so

$$\boxed{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \text{ is a basis for } \mathrm{Nul}(A+3I).}$$

Each eigenspace has dimension 2, so the sum of the dimensions of the eigenspaces of A is $2+2=4=n$.

Thus A is diagonalizable. In particular, if

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

then P is invertible and $A = PDP^{-1}$ where

$$D = P^{-1}AP = \begin{bmatrix} 5 & & & \\ & 5 & & \\ & & -3 & \\ & & & -3 \end{bmatrix}.$$

Proof of theorem. Fix an index $i \in \{1, 2, \dots, p\}$.

Let $\lambda = \lambda_i$ and suppose λ has multiplicity m and $\mathrm{Nul}(A - \lambda I)$ has dimension d.

Let $v_1, v_2, \dots, v_d$ be a basis for $\mathrm{Nul}(A - \lambda I)$.

One of the corollaries we saw for the dimension theorem is that it is always possible to choose vectors $v_{d+1}, v_{d+2}, \dots, v_n \in \mathbb{R}^n$ such that $v_1, v_2, \dots, v_d, v_{d+1}, v_{d+2}, \dots, v_n$ is a basis for $\mathbb{R}^n$.

Define $Q = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$. The columns of this matrix are linearly independent, so Q is invertible with $Qe_j = v_j$ and $Q^{-1}v_j = e_j$ for all $j = 1, 2, \dots, n$. Define $B = Q^{-1}AQ$.

If $j \in \{1, 2, \dots, d\}$ then the jth column of B is $Be_j = Q^{-1}AQe_j = Q^{-1}Av_j = \lambda Q^{-1}v_j = \lambda e_j$.

This means that the *first d columns* of B are

$$\begin{bmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

so B has the block-triangular form

$$B = \begin{bmatrix} \lambda & & & & * & * & \dots & * \\ & \lambda & & & * & * & \dots & * \\ & & \ddots & & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & * & \dots & * \\ 0 & 0 & \dots & 0 & * & * & \dots & * \\ \vdots & & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & * & * & \dots & * \end{bmatrix} = \left[\begin{array}{c|c} \lambda I_d & Y \\ \hline 0 & Z \end{array}\right]$$

where Y is an arbitrary $d \times (n-d)$ matrix and Z is an arbitrary $(n-d) \times (n-d)$ matrix.

Now, we want to deduce that

$$\det(B - xI) = (\lambda - x)^d \det(Z - xI).$$

Since $\det(A - xI) = \det(B - xI)$ as A and B are similar, and since $\det(Z - xI)$ is a polynomial in x, we see that $\det(A - xI)$ can be written as $(\lambda - x)^d p(x)$ for some polynomial $p(x)$. Since m is maximal such that $\det(A - xI) = (\lambda - x)^m p(x)$, it must hold that $d \leq m$. This proves part (a).

To prove parts (b) and (c), suppose $v_i^1, v_i^2, \dots, v_i^{\ell_i}$ is a basis for the λ_i-eigenspace of A for each $i = 1, 2, \dots, p$. Let $\mathcal{B}_i = \{v_i^1, v_i^2, \dots, v_i^{\ell_i}\}$. We claim that $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \mathcal{B}_p$ is a linearly independent set.

To prove this, suppose $\sum_{i=1}^p \sum_{j=1}^{\ell_i} c_i^j v_i^j = 0$ for some coefficients $c_i^j \in \mathbb{R}$.

It suffices to show that every $c_i^j = 0$.

Let $w_i = \sum_{j=1}^{\ell_i} c_i^j v_i^j \in \mathbb{R}^n$. We then have $w_1 + w_2 + \dots + w_p = 0$.

Each w_i is either zero or an eigenvector of A with eigenvalue λ_i. (Why?)

Since eigenvectors of A with distinct eigenvalues are linearly independent, we must have

$$w_1 = w_2 = \dots = w_p = 0.$$

But since each set $\mathcal{B}_i$ is linearly independent, this implies that $c_i^j = 0$ for all i, j.

We conclude that $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \mathcal{B}_p$ is a linearly independent set.

If the sum of the dimensions of the eigenspaces of A is n then $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_p$ is a set of n linearly independent eigenvectors of A, so A is diagonalizable.

If A is diagonalizable then A has n linearly independent eigenvectors. Among these vectors, the number that can belong to any particular eigenspace of A is necessarily the dimension of that eigenspace, so it follows that sum of the dimensions of the eigenspaces of A at least n. This sum cannot be more than n since the sum is the size of the linearly independent set $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_p \subset \mathbb{R}^n$. This proves part (b).

To prove part (c), note that if A is diagonalizable then $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_p$ is a set of n linearly independent vectors in $\mathbb{R}^n$, so is a basis for $\mathbb{R}^n$. The last assertion in part (c) is something we discussed at the beginning of this lecture. □