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.

- A is *diagonalizable* if $A = PDP^{-1}$ for some invertible matrix P and 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.

- The Fibonacci numbers are defined by $f_0 = 0$, $f_1 = 1$, and $f_n = f_{n-2} + f_{n-1}$ for $n \geq 2$.

  The ability to diagonalize a matrix lets us derive the exact formula

$$f_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \approx 0.447\,(1.618^n - (-0.618)^n)$$

- 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 the matrix $D = P^{-1}AP$ 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 $\lambda \in \mathbb{R}$.

Recall that v an *eigenvector* for A with *eigenvalue* λ if $0 \neq v \in \mathrm{Nul}(A - \lambda I)$, which means that $Av = \lambda v$.

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 $A = PBP^{-1}$.

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

Similar matrices have the same eigenvalues but usually different eigenvectors.

However, matrices may have the same eigenvalues but not be similar.

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.

A matrix is *diagonal* if all of its nonzero entries appear in diagonal positions $(1, 1), (2, 2), \dots,$ 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 important properties hold:

- The numbers $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A.
- The columns of P are a basis for $\mathbb{R}^n$ of eigenvectors of A.
- Specifically, if $P = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$ then $Av_i = \lambda_i v_i$ for each $i = 1, 2, \dots, n$.

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.

Theorem. If A is an $n \times n$ matrix with n distinct eigenvalues then A 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.

2 Diagonalization and Fibonacci numbers

Knowing how to diagonalize matrices will let us prove an exact formula for the *Fibonacci numbers*.

The sequence f_n of Fibonacci numbers starts as

$$f_0 = 0, \quad f_1 = 1, \quad f_2 = 1, \quad f_3 = 2, \quad f_4 = 3, \quad f_5 = 5, \quad f_6 = 8, \quad f_7 = 13 \quad \dots$$

For $n \geq 2$, the sequence is defined by $f_n = f_{n-2} + f_{n-1}$.

We have $f_{10} = 55$ and $f_{100} = 354224848179261915075$.

Define $a_n = f_{2n}$ and $b_n = f_{2n+1}$ for $n \geq 0$.

If $n > 0$ then $a_n = f_{2n} = f_{2n-2} + f_{2n-1} = a_{n-1} + b_{n-1}$.

Similarly, if $n > 0$ then $b_n = f_{2n+1} = f_{2n-1} + f_{2n} = b_{n-1} + a_n = a_{n-1} + 2b_{n-1}$.

We can put these two equations together into one matrix equation:

$$\begin{bmatrix} a_n \\ b_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ b_{n-1} \end{bmatrix}.$$

Since this holds for all $n > 0$, we have

$$\begin{bmatrix} a_n \\ b_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ b_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^2 \begin{bmatrix} a_{n-2} \\ b_{n-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^3 \begin{bmatrix} a_{n-3} \\ b_{n-3} \end{bmatrix} = \dots = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}.$$

In other words, $\begin{bmatrix} a_n \\ b_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Thus if we could get an exact formula for the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^n$ then we could derive a formula for $a_n = f_{2n}$ and $b_n = f_{2n+1}$, which would determine f_n for all n.

The best way we know to compute A^n for large values of n is to *diagonalize* A, that is, to find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$, since then $A^n = PD^nP^{-1}$.

Define the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

To determine if A is diagonalizable, our first step is to compute its eigenvalues, which are solutions to

$$0 = \det(A - xI) = \det \begin{bmatrix} 1-x & 1 \\ 1 & 2-x \end{bmatrix} = (1-x)(2-x) - 1 = x^2 - 3x + 1.$$

By the quadratic formula, the eigenvalues of A are $\alpha = \frac{3+\sqrt{5}}{2}$ and $\beta = \frac{3-\sqrt{5}}{2}$.

Since $\alpha - \beta = \sqrt{5} \neq 0$, these eigenvalues are distinct so A is diagonalizable. Note that

$$\alpha\beta = (3-\sqrt{5})(3+\sqrt{5})/4 = (9-5)/4 = 1.$$

Our next step is to find bases for the α- and β-eigenspaces of A.

To find an eigenvector for A with eigenvalue α, we row reduce

$$A - \alpha I = \begin{bmatrix} 1-\alpha & 1 \\ 1 & 2-\alpha \end{bmatrix} \sim \begin{bmatrix} 1 & 2-\alpha \\ 1-\alpha & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2-\alpha \\ 0 & 1-(2-\alpha)(1-\alpha) \end{bmatrix} = \begin{bmatrix} 1 & 2-\alpha \\ 0 & 0 \end{bmatrix} = \mathsf{RREF}(A - \alpha I).$$

The second equality holds since $(2-\alpha)(1-\alpha) = (1-\sqrt{5})(-1-\sqrt{5})/4 = (-1+5)/4 = 1$.

This computation shows that $x \in \mathrm{Nul}(A - \alpha I)$ if and only if $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $x_1 + (2-\alpha)x_2 = 0$, so

$$v = \begin{bmatrix} \alpha - 2 \\ 1 \end{bmatrix}$$

is an eigenvector for A with $Av = \alpha v$.

To find an eigenvector for A with eigenvalue β, we similarly row reduce

$$A - \beta I = \begin{bmatrix} 1-\beta & 1 \\ 1 & 2-\beta \end{bmatrix} \sim \begin{bmatrix} 1 & 2-\beta \\ 1-\beta & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2-\beta \\ 0 & 1-(2-\beta)(1-\beta) \end{bmatrix} = \begin{bmatrix} 1 & 2-\beta \\ 0 & 0 \end{bmatrix} = \mathsf{RREF}(A - \beta I).$$

The second equality holds since also $(2-\beta)(1-\beta) = 1$.

By algebra identical to the previous case, we deduce that

$$w = \begin{bmatrix} \beta - 2 \\ 1 \end{bmatrix}$$

is an eigenvector for A with $Aw = \beta w$.

This means that for

$$P = \begin{bmatrix} v & w \end{bmatrix} = \begin{bmatrix} \alpha - 2 & \beta - 2 \\ 1 & 1 \end{bmatrix} \qquad \text{and} \qquad D = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$$

we have $A = PDP^{-1}$. Since P is 2×2 with $\det P = (\alpha - 2) - (\beta - 2) = \alpha - \beta = \sqrt{5}$, we have

$$D^n = \begin{bmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{bmatrix} \qquad \text{and} \qquad P^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2-\beta \\ -1 & \alpha - 2 \end{bmatrix}.$$

We therefore have

$$\begin{bmatrix} a_n \\ b_n \end{bmatrix} = A^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} = PD^nP^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha - 2 & \beta - 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{bmatrix} \begin{bmatrix} 1 & 2-\beta \\ -1 & \alpha - 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Before computing anything further, it helps to make a few simplifications. Note that

$$\alpha - 2 = \frac{-1+\sqrt{5}}{2} = 1 - \beta \qquad \text{and} \qquad \beta - 2 = \frac{-1-\sqrt{5}}{2} = 1 - \alpha.$$

Hence

$$\begin{aligned}
\begin{bmatrix} a_n \\ b_n \end{bmatrix} &= \frac{1}{\sqrt{5}} \begin{bmatrix} 1-\beta & 1-\alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{bmatrix} \begin{bmatrix} 1 & \alpha - 1 \\ -1 & 1-\beta \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&= \frac{1}{\sqrt{5}} \begin{bmatrix} 1-\beta & 1-\alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{bmatrix} \begin{bmatrix} \alpha - 1 \\ 1-\beta \end{bmatrix} \\
&= \frac{1}{\sqrt{5}} \begin{bmatrix} 1-\beta & 1-\alpha \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (\alpha-1)\alpha^n \\ -(\beta-1)\beta^n \end{bmatrix} \\
&= \frac{1}{\sqrt{5}} \begin{bmatrix} (\alpha-1)(\beta-1)\,(\beta^n - \alpha^n) \\ (\alpha-1)\alpha^n - (\beta-1)\beta^n \end{bmatrix}.
\end{aligned}$$

Since

$$(\alpha - 1)(\beta - 1) = \frac{(1-\sqrt{5})(1+\sqrt{5})}{4} = \frac{1-4}{4} = -1,$$

rewriting this matrix equation gives

$$f_{2n} = a_n = \tfrac{1}{\sqrt{5}}\left(\alpha^n - \beta^n\right) \qquad \text{and} \qquad f_{2n+1} = b_n = \tfrac{1}{\sqrt{5}}\left((\alpha-1)\alpha^n - (\beta-1)\beta^n\right). \tag{*}$$

We now make one more unexpected observation:

$$(\alpha-1)^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{1+2\sqrt{5}+5}{4} = \frac{3+\sqrt{5}}{2} = \alpha$$

and

$$(\beta-1)^2 = \left(\frac{1-\sqrt{5}}{2}\right)^2 = \frac{1-2\sqrt{5}+5}{4} = \frac{3-\sqrt{5}}{2} = \beta.$$

Thus (*) become

$$f_{2n} = \tfrac{1}{\sqrt{5}}\left((\alpha-1)^{2n} - (\beta-1)^{2n}\right) \qquad \text{and} \qquad f_{2n+1} = \tfrac{1}{\sqrt{5}}\left((\alpha-1)^{2n+1} - (\beta-1)^{2n+1}\right). \tag{**}$$

Now we combine the identities in (**). Since $\alpha - 1 = \frac{1+\sqrt{5}}{2}$ and $\beta - 1 = \frac{1-\sqrt{5}}{2}$, we get:

Theorem. For all integers $n \geq 0$ it holds that

$$\boxed{f_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right) \approx 0.447\left(1.618^n - (-0.618)^n\right)}$$

Remark. Since $\frac{1-\sqrt{5}}{2} = -0.618\ldots$, if n is large then $f_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$.

Fun fact. The first few Fibonacci numbers are

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

If we add up all the decimal numbers

$$\begin{array}{l} 0.0 \\ 0.01 \\ 0.001 \\ 0.0002 \\ 0.00003 \\ 0.000005 \\ 0.0000008 \\ 0.00000013 \\ 0.000000021 \\ 0.0000000034 \\ 0.00000000055 \\ 0.000000000089 \\ 0.0000000000144 \\ \vdots \end{array}$$

then we get exactly 1/89. More precisely:

$$\boxed{\frac{1}{89} = \sum_{n=0}^{\infty} \frac{f_n}{10^{n+1}}.}$$

Proof. If $x \neq 1$ then $\sum_{n=0}^{N-1} x^n = \frac{1-x^N}{1-x}$ since

$$(1-x)\sum_{n=0}^{N-1} x^n = (1 + x + x^2 + \cdots + x^{N-1}) - (x + x^2 + x^3 + \cdots + x^N) = 1 - x^N.$$

It follows that if $|x| < 1$ so that $x^N \to 0$ as $N \to \infty$ then $\sum_{n=0}^{\infty} x^n = \lim_{N\to\infty} \sum_{n=0}^{N} x^n = \frac{1}{1-x}$. Now

$$\sum_{n=0}^{\infty} \frac{f_n}{10^{n+1}} = \frac{1}{10\sqrt{5}} \sum_{n=0}^{\infty} \left(\left(\frac{1+\sqrt{5}}{20}\right)^n - \left(\frac{1-\sqrt{5}}{20}\right)^n \right).$$

We have both $|\frac{1+\sqrt{5}}{20}| < 1$ and $|\frac{1-\sqrt{5}}{20}| < 1$ so

$$\sum_{n=0}^{\infty} \left(\left(\tfrac{1+\sqrt{5}}{20}\right)^n - \left(\tfrac{1-\sqrt{5}}{20}\right)^n \right) = \sum_{n=0}^{\infty} \left(\tfrac{1+\sqrt{5}}{20}\right)^n - \sum_{n=0}^{\infty} \left(\tfrac{1-\sqrt{5}}{20}\right)^n = \frac{1}{1 - \frac{1+\sqrt{5}}{20}} - \frac{1}{1 - \frac{1-\sqrt{5}}{20}}.$$

The last expression can be simplified a lot:

$$\frac{1}{1 - \frac{1+\sqrt{5}}{20}} - \frac{1}{1 - \frac{1-\sqrt{5}}{20}} = \frac{20}{19-\sqrt{5}} - \frac{20}{19+\sqrt{5}} = \frac{20(19+\sqrt{5}) - 20(19-\sqrt{5})}{(19-\sqrt{5})(19+\sqrt{5})} = \frac{40\sqrt{5}}{19^2 - 5} = \frac{40\sqrt{5}}{356} = \frac{10\sqrt{5}}{89}.$$

Substituting this above gives $\sum_{n=0}^{\infty} \frac{f_n}{10^{n+1}} = \frac{1}{10\sqrt{5}} \sum_{n=0}^{\infty} \left(\left(\frac{1+\sqrt{5}}{20}\right)^n - \left(\frac{1-\sqrt{5}}{20}\right)^n \right) = \frac{1}{10\sqrt{5}} \frac{10\sqrt{5}}{89} = \frac{1}{89}$. □

3 Diagonalizing matrices whose eigenvalues are not distinct

If an $n \times n$ matrix A has n distinct eigenvalues with corresponding eigenvectors $v_1, v_2, \ldots, v_n$, then the matrix $P = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}$ 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.

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 the matrix $D = P^{-1}AP$ 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 $\det(A - xI) = (5-x)^2(-3-x)^2$.

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

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

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

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & & & \\ & 5 & & \\ & & -3 & \\ & & & -3 \end{bmatrix}$$

then $A = PDP^{-1}$.

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 $c_i^j \in \mathbb{R}$. 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. □