Linear Independence of Eigenvalues

Recall from last time: to diagonalize an n×n matrix A:

1. Compute \( p(\lambda) = \det(A - \lambda I_n) \)
2. Solve \( p(\lambda) = 0 \) to find the eigenvalues
3. Find a basis for each eigenspace
4. Combine all these bases.

\* If you end up with \( n \) vectors, they’re LI
  - Otherwise \( A \) is not diagonalizable

In \* we need to justify why the eigenvectors are LI.

Fact: If \( \mathbf{w}_1, \ldots, \mathbf{w}_p \) are eigenvectors of \( A \) with different eigenvalues then \( \{ \mathbf{w}_1, \ldots, \mathbf{w}_p \} \) is LI.

Here’s how the Fact implies \*: Suppose

- \( \{ \mathbf{w}_1, \mathbf{w}_2 \} \) is a basis for the \( \lambda_1 \)-eigenspace
- \( \mathbf{w}_3 \) is a basis for the \( \lambda_2 \)-eigenspace.

I claim \( \{ \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 \} \) is LI.

Suppose \( x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2 + x_3 \mathbf{w}_3 = \mathbf{0} \). We need \( x_1 = x_2 = x_3 = 0 \).

- \( x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2 \) is in the \( \lambda_1 \)-eigenspace
- Since \( (x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2) + x_3 \mathbf{w}_3 = \mathbf{0} \), the Fact implies \( x_1 \mathbf{w}_1 + x_2 \mathbf{w}_2 = \mathbf{0} \) and \( x_3 \mathbf{w}_3 = \mathbf{0} \) (so \( x_3 = 0 \))
- Since \( \{ \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 \} \) is LI, this implies \( x_1 = x_2 = 0 \)
Proof of the Fact: Say $A w_i = \lambda_i w_i$ and all of the $\lambda_i, \ldots, \lambda_p$ are distinct. Suppose $\{w_1, \ldots, w_p\}$ is LD. Then for some $i$, $\{w_1, \ldots, w_i\}$ is LI but $w_{i+1} \in \text{Span } \{w_1, \ldots, w_i\}$, so

$$w_{i+1} = x_i w_1 + \cdots + x_i w_i$$

$$\Rightarrow A w_{i+1} = A(x_i w_1 + \cdots + x_i w_i)$$

$$\Rightarrow \lambda_i w_{i+1} = \lambda_i x_i w_1 + \cdots + \lambda_i x_i w_i$$

If $\lambda_i = 0$ then $\lambda_i x_i w_1 + \cdots + \lambda_i x_i w_i = 0 \implies x_i = \cdots = x_i = 0$ (because $\lambda_1, \ldots, \lambda_i \neq 0$), so $w_{i+1} = 0$, which can't happen because $w_{i+1}$ is an eigenvector.

If $\lambda_i \neq 0$ then

$$w_{i+1} = \frac{\lambda_i}{x_i} x_i w_1 + \cdots + \frac{\lambda_i}{x_i} x_i w_i$$

Subtract $w_{i+1} = x_i w_1 + \cdots + x_i w_i$:

$$0 = \left(\frac{\lambda_i}{x_i} - 1\right) x_i w_1 + \cdots + \left(\frac{\lambda_i}{x_i} - 1\right) x_i w_i$$

But $\lambda_j \neq \lambda_i$ for $j \leq i$, so $\frac{\lambda_i}{x_i} - 1 \neq 0$

$$\Rightarrow x_i = \cdots = x_i = 0$$

which is impossible, as before.
Consequence: If $A$ has $n$ (different) eigenvalues then $A$ is diagonalizable.

Indeed, if $\lambda_1, \ldots, \lambda_n$ are eigenvalues and $A\mathbf{w}_1 = \lambda_1 \mathbf{w}_1$, $\ldots$, $A\mathbf{w}_n = \lambda_n \mathbf{w}_n$ then $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is an eigenbasis by the Fact.

We’ll give a more general criterion (AM/GM) next time.
Matrix Form of Diagonalization

Thm: $A$ is diagonalizable $\iff$ there exists an invertible matrix $C$ and a diagonal matrix $D$ such that

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

In this case the columns of $C$ form an eigenbasis & the diagonal entries of $D$ are the corresponding eigenvalues.

$$C = \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix} \quad D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad A\omega_i = \lambda_i \omega_i$$

same order: $\omega_i \prec \lambda_i$

Eg: $A = \begin{pmatrix} 0 & 13 & 12 \\ 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} \Rightarrow A = CDC^{-1}$ for

$$C = \begin{pmatrix} 32 & 2 & 18 \\ 4 & -1 & -3 \\ 1 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -3/2 \end{pmatrix}$$

Eg: $A = \begin{pmatrix} 14 & -18 & -33 \\ -12 & 20 & 33 \\ 12 & -18 & -31 \end{pmatrix} \Rightarrow A = CDC^{-1}$ for

$$C = \begin{pmatrix} 3 & 11 & 1 \\ 2 & 4 & -1 \\ 0 & 1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
**Proof:** \( C \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right) = x_1 \omega_1 + \cdots + x_n \omega_n \)

\[ \Rightarrow C^{-1}(x_1 \omega_1 + \cdots + x_n \omega_n) = \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right) \]

Any vector has the form \( v = x_1 \omega_1 + \cdots + x_n \omega_n \), and two matrices are equal if they act the same on every vector. So check:

\[
CDC^{-1}v = CDC^{-1}(x_1 \omega_1 + \cdots + x_n \omega_n) = CDC^{-1}\left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right)
\]

\[
= C \left( \begin{array}{ccc} x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n \end{array} \right) \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right)
\]

\[
= \left( \begin{array}{ccc} \omega_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \omega_n \end{array} \right) \left( \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} \right) = x_1 \omega_1 + \cdots + x_n \omega_n
\]

\[
= A(x_1 \omega_1 + \cdots + x_n \omega_n) = Av
\]

**NB:** If \( A = CDC^{-1} \) then

\[
A^k = (CDC^{-1})^k = (CDC^{-1})(CDC^{-1})\cdots(CDC^{-1})
\]

\[
= CD^kC^{-1} = C \left( \begin{array}{ccc} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{array} \right) C^{-1}
\]

This is a **closed form expression** for \( A^k \) in terms of \( k \): much easier to compute!

\[
A^k = CD^kC^{-1}
\]

\[ \text{← this matrix has } n^2 \text{ entries that are functions of } k \]
Compare: \( A^k(x_1\omega_1 + \cdots + x_n\omega_n) = \lambda_1^k x_1\omega_1 + \cdots + \lambda_n^k x_n\omega_n \)

(vector form of the same identity).

Eq: \( A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \) is diagonal:

\( A e_1 = 2 e_1 \quad A e_2 = 3 e_2 \quad A e_3 = 4 e_3 \)

So \( \{e_1, e_2, e_3\} \) is an eigenbasis – so can take \( C = I_3 \), so the diagonalization is

\( A = I_3 A I_3 \)

Q: What if we take \( e_2 \) to be our first eigenvector?

NB: A matrix is diagonal \( \iff \) the unit coordinate vectors \( e_1 \ldots e_n \) are eigenvectors.
Geometry of Diagonalizable Matrices

When $A$ is diagonalizable, every vector can be written as a linear combination of eigenvectors, so multiplication by $A$ is reduced to scalar multiplication:

$$A(x_1w_1 + \ldots + x_nw_n) = \lambda_1x_1w_1 + \ldots + \lambda_nx_nw_n.$$ 

What does this mean geometrically?

→ Expanding in an eigenbasis and scalar multiplication can both be formulated geometrically!

NB: "Visualizing" a matrix means understanding how $x$ relates to $Ax$: think of $A$ as a function

$$x \mapsto Ax$$

input output

Eg: $D = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ so $D(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} 2x \\ \frac{1}{2}y \end{pmatrix}$

* scales the $x$-direction by 2
* scales the $y$-direction by $\frac{1}{2}$

[demo] $\frac{1}{2}$-eigenspace $2$-eigenspace
Eg: \[ A = \frac{1}{10} \begin{pmatrix} 11 & 6 \\ 9 & 14 \end{pmatrix} \]
\[ p(\lambda) = \lambda^2 - \frac{5}{2} \lambda + 1 = (\lambda - 2)(\lambda - \frac{1}{2}) \]
\[ \lambda_1 = 2, \quad \omega_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \lambda_2 = \frac{1}{2}, \quad \omega_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \]

Expand in the eigenbasis!
(think in terms of LCs of \( \omega_1, \omega_2 \))
\[ A(\omega_1 + x_2 \omega_2) = 2 \omega_1 + \frac{1}{2} x_2 \omega_2 \]
- scales the \( \omega_1 \)-direction by 2
- scales the \( \omega_2 \)-direction by \( \frac{1}{2} \)

This is the vector form. In matrix form,
\[ A = CDC^{-1} \]
\[ C = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} \]
\[ D = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \]

Then \( Av = CDC^{-1}v \)
- first multiply \( v \) by \( C^{-1} \)
- then multiply by the diagonal matrix \( D \)
- then multiply by \( C \) again

Note \( C(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = x_1 \omega_1 + x_2 \omega_2 \iff C^{-1}(x_1 \omega_1 + x_2 \omega_2) = (\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) \)
E_8: D = \begin{pmatrix} 1 & 0 \\ 0 & \nu_2 \end{pmatrix} \quad D(\vec{x}) = \begin{pmatrix} \nu_2 & \nu_2 \end{pmatrix} \vec{x} \\

- scales the \text{x-direction} by 1
- scales the \text{y-direction} by \nu_2

\begin{align*}
\text{1-eigenspace} & \quad 1 \times \text{v} = (1) \\
\text{2-eigenspace} & \quad \nu_2 \times \text{v} = (\nu_2) \\
\text{3-eigenspace} & \quad \frac{1}{2} \times \text{v} = (\frac{1}{2})
\end{align*}
\[ E_0: A = \frac{1}{6} \left( \begin{array}{cc} 5 & 4 \\ 2 & 3 \end{array} \right) \quad \rho(\lambda) = \lambda^2 - \frac{3}{2} \lambda + \frac{1}{2} = (\lambda - 1)(\lambda - \frac{1}{2}) \]

\[ \lambda_1 = 1 \quad \mathbf{u}_1 = (1) \quad \lambda_2 = \frac{1}{2} \quad \mathbf{u}_2 = \left( \begin{array}{c} 1 \\ 2 \end{array} \right) \]

Expand in the eigenbasis!

\[ A(\mathbf{x}_1 \mathbf{u}_1 + \mathbf{x}_2 \mathbf{u}_2) = 1 \mathbf{x}_1 \mathbf{u}_1 + \frac{1}{2} \mathbf{x}_2 \mathbf{u}_2 \]

- Scales the \( \mathbf{u}_1 \)-direction by 1
- Scales the \( \mathbf{u}_2 \)-direction by \( \frac{1}{2} \) 

Matrix Form: \[ A = CDC^{-1} \quad C = \left( \begin{array}{cc} 1 & 0 \\ 0 & 1/2 \end{array} \right) \quad D = \left( \begin{array}{cc} 1 & 0 \\ 0 & \frac{1}{2} \end{array} \right) \]
Eg: \[ A = \frac{1}{580} \begin{pmatrix} 503 & 73 & 269 \\ 207 & 1137 & -441 \\ 270 & -30 & 680 \end{pmatrix} \] has eigenbasis

\[ \mathbf{u}_1 = \begin{pmatrix} -7 \\ 2 \\ 5 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} -1 \\ -9 \\ 0 \end{pmatrix} \quad \mathbf{u}_3 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \]

and eigenvalues

\[ \lambda_1 = 1/2 \quad \lambda_2 = 2 \quad \lambda_3 = 3/2 \]

Expand in the eigenbasis!

\[ A(x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + x_3 \mathbf{u}_3) = \frac{1}{2} x_1 \mathbf{u}_1 + 2 x_2 \mathbf{u}_2 + \frac{3}{2} x_3 \mathbf{u}_3 \]

- scales the \( \mathbf{u}_1 \)-direction by \( \frac{1}{2} \) [demo]
- scales the \( \mathbf{u}_2 \)-direction by \( 2 \)
- scales the \( \mathbf{u}_3 \)-direction by \( \frac{3}{2} \)
Algebraic & Geometric Multiplicity

Next time we will discuss a criterion for diagonalizability.

(We like diagonalizable matrices because we can solve difference equations.)

Recall: If $\lambda$ is a root of a polynomial $p(x)$, its multiplicity $m$ is the largest power of $(x-\lambda)$ dividing $p$:

$$p(x) = (x-\lambda)^m \cdot (\text{other factors})$$

Def: Let $A$ be an $n \times n$ matrix with eigenvalue $\lambda$.

1. The algebraic multiplicity (AM) of $\lambda$ is its multiplicity as a root of the characteristic polynomial $p(\lambda)$.

2. The geometric multiplicity (GM) of $\lambda$ is the dimension of the $\lambda$-eigenspace:

$$GM(\lambda) = \dim \text{Null}(A-\lambda I_n)$$

$$= \# \text{free variables in } A-\lambda I_n.$$  

$$= \# \text{linearly independent } \lambda\text{-eigenvectors}$$
Example: 

\[ A = \begin{pmatrix} -7 & 3 & 5 \\ -6 & 5 & 6 \\ -9 & 3 & 7 \end{pmatrix} \quad p(\lambda) = -(\lambda-2)^2(\lambda-1) \]

So the eigenvalues are 1 & 2.

- \( \lambda = 1 \): \( \text{AM} = 1 \).
  
  \[ \text{Null}(A-1I_3) = \text{Span} \{(1)\} \]
  
  \[ \rightarrow \text{this is a line: GM = 1} \]

- \( \lambda = 2 \): \( \text{AM} = 2 \).
  
  \[ \text{Null}(A-2I_3) = \text{Span} \{(\frac{3}{4})\} \]
  
  \[ \rightarrow \text{this is a line: GM = 1} \]

This matrix is not diagonalizable:

only two linearly independent eigenvectors.

[demo]
Eg: $B = \begin{pmatrix} -4 & 3 & 2 \\ -6 & 5 & 2 \\ -6 & 3 & 4 \end{pmatrix}$ $p(\lambda) = -(\lambda - 2)^2(\lambda - 1)$

So the eigenvalues are 1 & 2.

- $\lambda = 1$: $AM = 1$
  
  $\text{Null}(B - 1I_3) = \text{Span} \begin{Bmatrix} (1) \end{Bmatrix}$
  
  $\rightarrow$ this is a line: $GM = 1$

- $\lambda = 2$: $AM = 2$
  
  $\text{Null}(B - 2I_3) = \text{Span} \begin{Bmatrix} (\frac{3}{4}), (\frac{1}{2}) \end{Bmatrix}$

  $\rightarrow$ this is a plane: $GM = 2$

This matrix is diagonalizable: an eigenbasis is

$\begin{Bmatrix} (1), (\frac{3}{4}), (\frac{1}{2}) \end{Bmatrix}$

Both matrices have only 2 eigenvalues.

The difference is that $B$ had $AM = GM = 2$ LI 2-eigenvectors and $A$ had one.