Differential Eqns and Linear Algebra

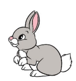
Textbook: *Differential Equations and Linear Algebra* by Edward and Penney

Previous Lecture

- Eigenvalues, eigenvectors: $\mathbf{A}\vec{v} = \lambda\vec{v}$
- Characteristic Equation: $|\mathbf{A} - \lambda\mathbf{I}| = 0$
- Eigenspaces: $span\{\vec{v}_1, \vec{v}_2, \dots \vec{v}_m\}$

6.2: Diagonals of Matrices

Motivation $\vec{x} = \begin{bmatrix} 2,000 \\ 1,500 \\ 10,000 \end{bmatrix}$



In many applications (like population models), one can discover a **transition matrix A**, that will transition a vector $\vec{x}$ (e.g., containing the populations of different animals in a region) from one state $\vec{x}_0$ to another $\vec{x}_1$ (e.g., from the populations in some year, to the populations in the next year).

They're also used in onboard aviation software, satellite orbit maintenance, various statistics applications, and many other fields.

Transitioning to the next state is done by simply multiplying: $\mathbf{A}\vec{x}_0 = \vec{x}_1$.

However, we're usually interested in the long-term behavior of $\vec{x}$ (e.g., the population), so perhaps what $\vec{x}_{1000}$ will be. ...

But this would require us to calculate $\overset{\text{(1000 times!)}}{\mathbf{AA}\dots\mathbf{A}}\ \vec{x}_0 = \mathbf{A}^{1000}\vec{x}_0 = \vec{x}_{1000}$.

But $\mathbf{A}^{1000}$ is tedious to calculate without some better technique. Diagonalization is that technique.

Diagonal Matrix: $\begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$, "Zeros off of the diagonal."

So, $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is diagonal, but so is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$!!

If we can characterize $\mathbf{A}$ as $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix, and $\mathbf{P}, \mathbf{P}^{-1}$ are invertible matrices, then notice that:

$$\mathbf{A}^3 = (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1})$$

$$= \mathbf{PD}(\mathbf{P}^{-1}\mathbf{P})\mathbf{D}(\mathbf{P}^{-1}\mathbf{P})\mathbf{DP}^{-1}$$

$$= \mathbf{PD}(\mathbf{I})\mathbf{D}(\mathbf{I})\mathbf{DP}^{-1} = \mathbf{PDDDP}^{-1} = \mathbf{PD}^3\mathbf{P}^{-1}.$$

Also note that for any diagonal matrix: $\begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{bmatrix}^n = \begin{bmatrix} a_1^n & 0 & \dots & 0 \\ 0 & a_2^n & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & a_n^n \end{bmatrix}$.

Example: $\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}^3$ ...

$$= \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 9 & 0 \\ 0 & 25 \end{bmatrix} = \begin{bmatrix} 27 & 0 \\ 0 & 125 \end{bmatrix} = \begin{bmatrix} 3^3 & 0 \\ 0 & 5^3 \end{bmatrix}.$$

As a result, to calculate $\mathbf{A}^{1000}$, we need only calculate $\mathbf{D}^{1000}$, by raising each component to the 1000*th* power, and then compute $\mathbf{PD}^{1000}\mathbf{P}^{-1}$ (two matrix multiplications instead of 1000).

But how can we transform $\mathbf{A}$ into $\mathbf{PDP}^{-1}$? This transformation is called **diagonalizing**.

Diagonalizing Criteria

The $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are called **similar**, if there exists an invertible matrix $\mathbf{P}$, such that: $\mathbf{B} = \mathbf{PAP}^{-1}$.

Diagonalizable: $\mathbf{A}^{n \times n}$ is called diagonalizable if it's *similar* to a diagonal matrix $\mathbf{D}$.
That is, there exists a diagonal matrix $\mathbf{D}$ and an invertible matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$, and so $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

Criteria for Diagonalizability: $\mathbf{A}^{n \times n}$ is **diagonalizable** *iff* it has n linearly independent e-vecs $\vec{v}_i$
(note that this may be possible even if you have less than n distinct e-vals λ_i).

Proof: Due to the "if and only if" (*iff*), there will be two parts to our proof.
First we'll show ($\Leftarrow$), that if we have n linearly independent e-vecs, then $\mathbf{A}$ is diagonalizable.

That is, we must show the existence of $\mathbf{D}, \mathbf{P}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$.

Suppose our e-vals are $\lambda_1, \dots, \lambda_n$ (perhaps not all unique) corresponding to the

n independent e-vecs $\vec{v}_1, \dots, \vec{v}_n$, and let $\mathbf{P} := \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix}$.

Then, $\mathbf{AP} = \mathbf{A}\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{A}\vec{v}_1 & \mathbf{A}\vec{v}_2 & \dots & \mathbf{A}\vec{v}_n \\ | & | & & | \end{bmatrix}$

$$= \begin{bmatrix} | & | & & | \\ \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \dots & \lambda_n\vec{v}_n \\ | & | & & | \end{bmatrix}. \qquad \text{(by definition of e-vec)}$$

Now consider the diagonal matrix $\mathbf{D} := \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$.

So, $\mathbf{PD} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \dots & \lambda_n\vec{v}_n \\ | & | & & | \end{bmatrix}$.

And note that above we have shown that $\mathbf{AP} = \mathbf{PD}$.

And since we know that $\mathbf{P}$ is invertible (having n linearly independent column vectors), we can multiply on the right by $\mathbf{P}^{-1}$, to obtain: $\mathbf{A} = \mathbf{PDP}^{-1}$. So we have now shown ($\Leftarrow$).

Next we will show ($\Rightarrow$). That if $\mathbf{A}$ is diagonalizable, then we have n linearly independent e-vecs.

Suppose that $\mathbf{A}$ is similar to $\mathbf{D} := \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}$, and let $\mathbf{P} = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix}$ be an invertible matrix such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$ or equivalently $\mathbf{AP} = \mathbf{PD}$ (this is what it means to be diagonalizable).

We must show that $\vec{v}_1, \dots, \vec{v}_n$ are e-vecs and linearly independent.

We calculate, $\mathbf{AP} = \mathbf{A}\begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}\vec{v}_1 & \dots & \mathbf{A}\vec{v}_n \end{bmatrix}$,

and $\mathbf{PD} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \vdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix} = \begin{bmatrix} d_1\vec{v}_1 & \dots & d_n\vec{v}_n \end{bmatrix}$.

Comparing $\mathbf{AP} = \mathbf{PD}$ component-wise, it follows that $\mathbf{A}\vec{v}_j = d_j\vec{v}_j$ for $j = 1, 2, \dots, n$.

But this defines what it is to be in e-vec/e-val. Thus the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are e-vecs of $\mathbf{A}$ associated with the e-vals $d_1, d_2, \dots, d_n$, respectively.

And it follows from previous theorems that these n e-vecs of the matrix $\mathbf{A}$ are linearly independent, because they are the column vectors of the invertible matrix $\mathbf{P}$. ■

So we have proven the claim that: "$\mathbf{A}^{n\times n}$ is **diagonalizable** *iff* it has n linearly independent e-vecs $\vec{v}_i$."

Thus our ability to diagonalize depends upon **A** having n linearly independent e-vecs!

The following theorem is helpful in this regard:

E-vecs Associated w/Distinct E-vals Thm: Suppose that the k e-vecs $\vec{v}_1, \ldots, \vec{v}_k$ are associated with the k distinct e-vals $\lambda_1, \ldots, \lambda_k$ of the matrix **A**. Then these k e-vecs are linearly independent.

Proof: Using induction on k, the thm is obviously true when $k = 1$, so this satisfies our base case.

Our next task is to show that if any set of $k-1$ e-vecs associated with distinct e-vals is linearly independent, then any set of k e-vecs $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{k-1}, \vec{v}_k\}$ is also linearly independent.

In other words, that $c_1\vec{v}_1 + c_2\vec{v}_2, + \ldots + c_k\vec{v}_k = \vec{0}$, $\qquad (*)$

requires $c_1 = c_2 = \ldots = c_k = 0$. First note that $(\mathbf{A} - \lambda_j\mathbf{I})\vec{v}_j = 0$, for all j since these are e-vecs.

For our next calculation we observe that $(\mathbf{A} - \lambda_j\mathbf{I})\vec{v}_i = \mathbf{A}\vec{v}_i - \lambda_j\vec{v}_i = \lambda_i\vec{v}_i - \lambda_j\vec{v}_i = (\lambda_i - \lambda_j)\vec{v}_i$, for all i,j.

Therefore, if we multiply $(*)$ by $(\mathbf{A} - \lambda_1\mathbf{I})$, we get $0 + (\lambda_2 - \lambda_1)c_2\vec{v}_2, + \ldots + (\lambda_k - \lambda_1)c_k\vec{v}_k = \vec{0}$.

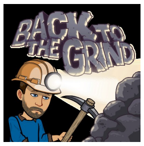
Since we assumed the e-vals were distinct, and since we assumed that any set of $k-1$ e-vecs is linearly independent, this requires $c_2 = c_3 = \ldots = c_k = 0$. If we substitute these into $(*)$, the remaining equation is:

$c_1\vec{v}_1 = 0$, but since we know that e-vecs are nontrivial, it must be that $c_1 = 0$. Having now shown that all $c_i = 0$, we can conclude that the k e-vecs are linearly independent, and by induction the theorem follows. ■

Conclusion: So if we find n such distinct e-vals, our matrix is diagonalizable.

However, this does **NOT** mean that if you find $k < n$ distinct e-vals that your matrix is un-diagonalizable!

Rather, it means you must calculate your e-vecs to see if your k e-vals *nonetheless* generate n e-vecs.



Diagonalizing Algorithm

Upon calculating the e-vals and e-vecs of **A**, if you find n linearly independent e-vecs, then below I lay out how you can construct $\mathbf{A} = \mathbf{PDP}^{-1}$ where **P** is an invertible matrix, and **D** is a diagonal matrix (which is apparently similar to **A**).

♦ Arrange the e-vals along the principal diagonal of an otherwise zero matrix (including non-distinct λ_ks if any).

For example, $\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$.

♦ Arrange the corresponding e-vecs vertically as columns in a new matrix (in the same order as you did the λ_ks):

$$\mathbf{P} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}.$$

♦ Calculate the inverse $\mathbf{P}^{-1}$. Now you have: $\mathbf{A} = \mathbf{PDP}^{-1}$, done!

Example: Determine whether $\mathbf{A} = \begin{bmatrix} 5 & -4 & 2 \\ 4 & -3 & 2 \\ 2 & -2 & 4 \end{bmatrix}$ is diagonalizable.

If it is, find a diagonalizing matrix **P** and a diagonal matrix **D** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$.

Solution: Must use: $|\mathbf{A} - \lambda\mathbf{I}| = 0$, $\qquad \begin{vmatrix} 5-\lambda & -4 & 2 \\ 4 & -3-\lambda & 2 \\ 2 & -2 & 4-\lambda \end{vmatrix} = 0$

$$= \begin{vmatrix} 5-\lambda & -4 & 2 \\ 0 & 1-\lambda & -6+2\lambda \\ 2 & -2 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 5-\lambda & 1-\lambda & 2 \\ 0 & 1-\lambda & -6+2\lambda \\ 2 & 0 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 5-\lambda & 0 & 8-2\lambda \\ 0 & 1-\lambda & -6+2\lambda \\ 2 & 0 & 4-\lambda \end{vmatrix}$$

$$= 2(0 - (8-2\lambda)(1-\lambda)) + (4-\lambda)((5-\lambda)(1-\lambda) - 0)$$

$$= (1-\lambda)(-2(8-2\lambda) + (4-\lambda)(5-\lambda)) = (1-\lambda)((-16+4\lambda) + (\lambda^2 - 9\lambda + 20))$$

$$= (1-\lambda)(\lambda^2 - 5\lambda + 4) = (1-\lambda)(\lambda-4)(\lambda-1) = -(\lambda-4)(\lambda-1)^2.$$

Eigenvalues: $\lambda_1 = 1,\ \lambda_2 = 1,\ \lambda_3 = 4$.

Recall: if $\mathbf{A}^{n\times n}$ has n distinct e-vals, then it's diagonalizable. But 1 has multiplicity two!
So we don't know yet, we must check the e-vecs.

$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. $\qquad$ Must use: $(\mathbf{A} - \lambda_k\mathbf{I})\vec{v} = \vec{0}$.

With $\lambda_{1,2} = 1$: $\begin{bmatrix} 5-1 & -4 & 2 \\ 4 & -3-1 & 2 \\ 2 & -2 & 4-1 \end{bmatrix} = \begin{bmatrix} 4 & -4 & 2 \\ 4 & -4 & 2 \\ 2 & -2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 4 & -4 & 2 \\ 2 & -2 & 3 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 2 & -2 & 1 \\ 2 & -2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, so only one free variable.

One free variable will only produce one e-vec. And since $\lambda_3 = 4$ (mult 1) can produce, at most, one e-vec, we'll not have sufficient e-vecs (we need three) to diagonalize.

Example: Determine whether $\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is diagonalizable.

If it is, find a diagonalizing matrix **P** and a diagonal matrix **D** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$.

Solution: Must use: $|\mathbf{A} - \lambda\mathbf{I}| = 0$, $\begin{vmatrix} 3-\lambda & 1 & 0 \\ 1 & 3-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0.$

$$\begin{vmatrix} 3-\lambda & 1 & 0 \\ 1 & 3-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)((3-\lambda)^2 - 1) = (2-\lambda)(\lambda^2 - 6\lambda + 8) = (2-\lambda)(\lambda-2)(\lambda-4).$$

Eigenvalues: $\lambda_1 = 2,\ \lambda_2 = 2,\ \lambda_3 = 4.$

With $\lambda_{1,2} = 2$: $\begin{bmatrix} 3-2 & 1 & 0 \\ 1 & 3-2 & 0 \\ 0 & 0 & 2-2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 0 \end{bmatrix},$

Two free variables, so we'll be able to diagonalize! Let's set $z = s,\ \ y = t$, and so $x = -y = -t$.

$(x,y,z) = (-t,t,s) = t(-1,1,0) + s(0,0,1)$. So $\vec{v}_1 = (-1,1,0)^T$, and $\vec{v}_2 = (0,0,1)^T$, when $t = s = 1$.

With $\lambda_3 = 4$: $\begin{bmatrix} 3-4 & 1 & 0 \\ 1 & 3-4 & 0 \\ 0 & 0 & 2-4 \end{bmatrix} \Rightarrow \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix},\ \ y = s.$

$\Rightarrow y = s, \qquad z = 0, \qquad x = y = s. \qquad (x,y,z) = (s,s,0) = s(1,1,0).$

$\vec{v}_3 = (-1,1,0)^T$, when $s = 1$.

So, $\mathbf{P} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. And, $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

Other Tidbits

Similarity is Different than Row Equivalence

Note from above, that the e-vals λ_i of **A** were necessarily the e-vals of the similar matrix **D**.

Proof: $|\mathbf{D} - \lambda\mathbf{I}| = \begin{vmatrix} \lambda_1 - \lambda & 0 \\ 0 & \lambda_2 - \lambda \end{vmatrix} = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = 0 \quad \Rightarrow \quad \lambda \in \{\lambda_1, \lambda_2\}.$ ■

However, consider $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Observe that you get **B** from **A** by subtracting the 1st row of **A** from the 2nd row, so **they are row equivalent**.

However, $|\mathbf{A} - \lambda\mathbf{I}| = \lambda^2 - 3\lambda + 1 \Rightarrow \lambda_{1,2} = \frac{1}{2}\left(3 \pm \sqrt{5}\right)$ (thus diagonalizable), and $|\mathbf{B} - \lambda\mathbf{I}| = (1-\lambda)^2 \Rightarrow \lambda_{1,2} = 1$.

So if they *were* similar, we would have: $\mathbf{A} = \mathbf{PDP}^{-1}$, and $\mathbf{B} = \mathbf{LAL}^{-1}$.

And thus: $\mathbf{B} = \mathbf{L}(\mathbf{PDP^{-1}})\mathbf{L}^{-1} = (\mathbf{LP})\mathbf{D}\left(\mathbf{P^{-1}L}^{-1}\right)$, where $(\mathbf{LP})$ and $\left(\mathbf{P^{-1}L}^{-1}\right)$, are inverses.

In other words, $\mathbf{B}$ would be similar to $\mathbf{D}$, and have e-vals $\frac{1}{2}\left(3 \pm \sqrt{5}\right)$. Which it doesn't.

Thus they are not similar.

Test for Linear Independence of Vectors

If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ are k e-vecs of k **distinct** e-vals $\lambda_1, \lambda_2, \ldots, \lambda_k$ of $\mathbf{A}$ (w/mult 1), then **these e-vecs are linearly independent** (see proof above).

Also, if $\lambda_1, \lambda_2, \lambda_3$ are distinct e-vals of $\mathbf{A}^{n\times n}$, and $B_{\lambda_1} = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$, $B_{\lambda_2} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_s\}$, $B_{\lambda_3} = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t\}$ are the bases of the associated eigenspaces, then their union $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\} \cup \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_s\} \cup \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_t\}$ is a linearly independent set of e-vecs of $\mathbf{A}$.

Now you are ready to predict those populations over time!

Exercises 6.2

What did we learn?

- Diagonals of Matrices: "Zeros off of the diagonal"
- Diagonalizing Criteria: similar matrices $\mathbf{B} = \mathbf{PAP}^{-1}$
- Diagonalizing Algorithm