CME 200 WORKSHOP: EIGENVECTORS, EIGENVALUES, & DIAGONALIZABILITY

LAURA LYMAN

Contents

1. Eigenvectors & Eigenvalues

Definition 1.1 (Eigenvector, eigenvalue)**.** Recall that an *eigenvector* of an $n \times n$ matrix A is a nonzero vector $\vec{y}$ such that $A\vec{y} = \lambda\vec{y}$ for some scalar $\lambda \in \mathbb{R}$ (or $\mathbb{C}$). The associated scalar λ is called the *eigenvalue* associated with eigenvector $\vec{y}$ of A.

Example 1.2. Say we have $\begin{pmatrix} 1 & 5 \\ 0 & 6 \end{pmatrix}$. Then observe that

$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 0 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 6\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

So $\vec{y} = (1,1)^T$ is an eigenvector of A with eigenvalue $\lambda = 6$. Similarly, suppose

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix} = (-i)\begin{pmatrix} 1 \\ -i \end{pmatrix},$$

meaning $\vec{y} = (1,-i)^T$ is an eigenvector of A with eigenvalue $\lambda = -i$.

For each eigenvector $\vec{y}$, there exists a *unique* eigenvalue λ associated with $\vec{y}$.

Exercise for Audience 1.3. What about the other way around? That is, does a single eigenvalue correspond to only one eigenvector?

Answer: No! Take $A = I$. Then for any $y \in \mathbb{R}^n$(or $\mathbb{C}^n$), $A\vec{y} = \vec{y} = (1)\vec{y}$, so *every nonzero vector* $\vec{y}$ is an eigenvector of A with eigenvalue $\lambda = 1$. That is, $\lambda = 1$ corresponds to infinitely many eigenvectors.

So for each λ, there may be many eigenvectors $\vec{y}$ with λ as an eigenvalue.

How do we go about finding the eigenvalues of some matrix A? As was discussed in Margot's last lecture,

$$\lambda \text{ is an eigenvalue of } A \Leftrightarrow \det(A - \lambda I) = 0.$$

Exercise for Audience 1.4. Find the eigenvalues of $A = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}$.

Answer: We compute

$$\det(A - \lambda I) = \det\left(\begin{pmatrix} 2-\lambda & 2 \\ 3 & 1-\lambda \end{pmatrix}\right) = (2-\lambda)(1-\lambda) - 6$$
$$= \lambda^2 - 3\lambda + 2 - 4 = (\lambda - 4)(\lambda + 1).$$

By setting $\det(A - \lambda I) = 0$, we have the solutions $\lambda = 4$ and $\lambda = -1$. The eigenvalues happen to be distinct here, but they *do not have to be.*

Exercise for Audience 1.5. Can you have an eigenvalue of a rectangular matrix? Say A is $m \times n$ with $m \neq n$.

Answer: No! For A an $m \times n$ matrix, $A\vec{y} = \lambda\vec{y}$ would look like

$$\begin{bmatrix} & & \\ & & \end{bmatrix}_{m\times n} \begin{bmatrix} \vec{y} \\ \end{bmatrix}_{n\times 1} = \lambda \begin{bmatrix} \vec{y} \\ \end{bmatrix}_{m\times 1}$$

meaning $\vec{y}$ would have inconsistent dimension. That is, if A is $m \times n$, then A acting on $\vec{y}$ changes the dimension of $\vec{y}$, so left-multiplication by A cannot just be scaling its input vector. Therefore, *only square matrices can have eigenvalues/eigenvectors.*

As a sanity-check, we see that this is consistent with taking $\det(A-\lambda I)$; we can only take the determinants of *square matrices*, so A must be square. But where does this determinant calculation come from?

Theorem 1.6. *For $n \times n$ matrix A,*

$$\lambda \textit{ is an eigenvalue of } A \Leftrightarrow \det(A - \lambda I) = 0.$$

Proof. ($\Rightarrow$) Suppose λ is an eigenvalue of A with eigenvector $\vec{y}$. Then $A\vec{y} = \lambda\vec{y}$ for some $\vec{y} \neq 0$. So

$$\vec{0} = A\vec{y} - \lambda\vec{y} = (A - \lambda I)\vec{y},$$

where $\vec{y} \neq 0$. What can we say about $A - \lambda I$ from this? We notice that $\vec{y} \in \mathcal{N}(A - \lambda I)$ (the nullspace of $A - \lambda I$) since $(A - \lambda I)\vec{y} = 0$. However, since $\vec{y} \neq 0$ we have that $A - \lambda I$ is *singular*. (If you do not remember why that is, we will review the proof after this current proof is completed.) Since a square matrix is singular if and only if its determinant equals zero, $\det(A - \lambda I) = 0$.

($\Leftarrow$) This direction is nearly identical. Suppose $\det(A-\lambda I) = 0$. Then $A-\lambda I$ is singular, so $\exists\, \vec{y} \in R^n$ (or $\mathbb{C}^n$) such that $(A - \lambda I)\vec{y} = 0$. (That is, the nullspace of $A - \lambda I$ must have a nonzero element.) Then

$$\vec{0} = (A - \lambda I)\vec{y} = A\vec{y} - \lambda\vec{y}$$

and so $A\vec{y} = \lambda\vec{y}$. Therefore, λ is defined as an eigenvalue of A (with $\vec{y}$ as one of its eigenvectors). □

As promised, we can provide some review for why A is singular $\Leftrightarrow$ A has a nonzero element in its nullspace. This is a very useful fact to have on hand when proving statements about matrices (or linear mappings in general), so you should understand why it is true.

Recall that for any matrix $A(\vec{0}) = \vec{0}$, so the nullspace of A *always* contains $\vec{y} = \vec{0}$; this is true regardless of whether A is singular or not. **The nullspace of a matrix is never empty.** We say a nullspace $\mathcal{N}(A)$ is *trivial* if $\mathcal{N}(A) = \{0\}$, which is to say that the only vector A maps to $\vec{0}$ is the zero vector itself. We will show that

- A is invertible (not singular) if and only if the vector $\vec{y}$ such that $A\vec{y} = 0$ can *only* be $\vec{y} = \vec{0}$ (i.e. $\mathcal{N}(A) = \{\vec{0}\}$).

The logical equivalent of this statement (via the contrapositive) is that

- A is singular if and only if there is a *nonzero* $\vec{y}$ such that $A\vec{y} = 0$ (i.e. $\mathcal{N}(A) \neq \{0\}$).

Note that we are not requiring A to be square in this proof.

Theorem 1.7 (Review)**.** *An $m \times n$ matrix A is invertible (not singular) $\Leftrightarrow \mathcal{N}(A) = \{y \in \mathbb{R}^n$ (or $\mathbb{C}^n$) $\mid Ay = \vec{0}\} = \{0\}$, which is to say that the only $\vec{y}$ such that $A\vec{y} = 0$ is $\vec{y} = \vec{0}$.*

Proof. ($\Rightarrow$) Suppose $A\vec{y} = \vec{0}$. Since A is invertible, we can multiply both sides by A^{-1} to get $\vec{y} = A^{-1}(\vec{0}) = \vec{0}$. Thus, $\mathcal{N}(A) = \{\vec{0}\}$.

($\Leftarrow$) Let $\mathcal{N}(A) = \{\vec{0}\}$. For any $\vec{y}$, by matrix multiplication we have that

$$A\vec{y} = \left(\begin{array}{c|c|c} | & & | \\ \vec{a_1} & \cdots & \vec{a_n} \\ | & & | \end{array} \right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = y_1\vec{a_1} + \cdots + y_n\vec{a_n}.$$

Since $\mathcal{N}(A) = \{\vec{0}\}$, we know that $A\vec{y} = \vec{0} \Leftrightarrow \vec{y} = \vec{0}$. Thus, $y_1\vec{a_1} + \cdots + y_n\vec{a_n} = \vec{0} \Leftrightarrow y_1 = \cdots = y_n = 0$, meaning the columns $\vec{a_1}, \ldots, \vec{a_n}$ of A are *linearly independent.* Since A has full column rank, A must be invertible. □

Theorem 1.8. *The eigenvectors of an orthogonal $n \times n$ matrix Q are $\lambda = \pm 1$.*

Proof. Take $Q\vec{y} = \lambda\vec{y}$, so $Q^TQ\vec{y} = \lambda Q^T\vec{y} = \vec{y}$. Multiplying through by y^T yields

$$\|y\|_2^2 = \vec{y}^T Q^T \vec{y} = (Q\vec{y})^T\vec{y} = \lambda\vec{y}^T\vec{y} = \lambda\|y\|_2^2.$$

Since $\vec{y}$ is an eigenvector, by definition $\vec{y} \neq \vec{0}$. Thus $\|\vec{y}\|_2^2 \neq 0$, so we can divide through by this scalar value $\|\vec{y}\|_2^2$ to get $\lambda^2 = 1$. Therefore $\lambda = \pm 1$. □

2. Diagonalizable Matrices

Definition 2.1 (Diagonalizable)**.** We say that an $n \times n$ matrix A is *diagonalizable* if $A = Y\Lambda Y^{-1}$ for some $n \times n$ matrix Y and diagonal matrix Λ.

If A is diagonalizable, we call the factorization $A = Y\Lambda Y^{-1}$ the *eigendecomposition* of matrix A.

Example 2.2. Let $A = \begin{pmatrix} 2 & 6 \\ 0 & 1 \end{pmatrix}$. Then

$$A = \begin{pmatrix} 2 & 6 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1}$$

so A is diagonalizable.

Why is such a decomposition useful? Suppose we could write $A = Y\Lambda Y^{-1}$. Then

$$A^k = (Y\Lambda Y^{-1})^k = (Y\Lambda Y^{-1})(Y\Lambda Y^{-1})\cdots(Y\Lambda Y^{-1}) = Y\Lambda^k Y^{-1}.$$

Since Λ is diagonal,

$$\Lambda^k = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}.$$

Using naive matrix multiplication, the cost of computing A^2 is $O(n^3)$. So to get A^k, we have $(k-1)O(n^3) = O(n^3 k)$, which is very slow. But this is faster.

Why else is such a decomposition useful? Suppose A can be diagonalized, so $A = Y\Lambda Y^{-1}$ for some invertible $n \times n$ matrix Y. Then

$$AY = Y\Lambda = Y \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Therefore,

$$A \left(\begin{array}{c|c|c} | & & | \\ \vec{y_1} & \cdots & \vec{y_n} \\ | & & | \end{array} \right) = \left(\begin{array}{c|c|c} | & & | \\ A\vec{y_1} & \cdots & A\vec{y_n} \\ | & & | \end{array} \right) = Y \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \left(\begin{array}{c|c|c} | & & | \\ \lambda_1\vec{y_1} & \cdots & \lambda_n\vec{y_n} \\ | & & | \end{array} \right).$$

By comparing columns, we see that $A\vec{y_i} = \lambda_i \vec{y_i}$ for $i = 1, \dots, n$. So the *diagonal entries of Λ are eigenvalues whose eigenvectors are the corresponding columns of Y*. That is another reason why the eigendecomposition is important. In addition to its computation convenience, the factorization encodes essential information about the matrix A.

Example 2.3. Let us practice diagonalizing A by hand. Take $A = \begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix}$. Then $\det(A - \lambda I) = \det\left(\begin{pmatrix} 2-\lambda & 6 \\ 0 & -1-\lambda \end{pmatrix}\right) = (2-\lambda)(-1-\lambda) = 0 \Rightarrow \lambda_1 = 2, \lambda_2 = -1$. Then we solve

$$\vec{0} = (A - 2I)\vec{y} = \begin{pmatrix} 0 & 6 \\ 0 & -3 \end{pmatrix} \vec{y} \Longrightarrow y_2 = 0.$$

So $E_2(A) = \{y_1(1,0)^T\}$. Let us pick $\vec{y_1} = (1,0)^T$. Similarly, $\vec{y_2} = (-2,1)^T$, so one possibility is

$$Y = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Then Y is invertible $Y^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Then we can check that

$$A = Y\Lambda Y^{-1}.$$

Exercise for Audience 2.4. Find the eigendecomposition of $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$.

Answer: For the sake of pedagogy, we go through the computation

$$\det(A-\lambda I) = \det \begin{pmatrix} 2-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ -1 & 0 & 1-\lambda \end{pmatrix} = (2-\lambda)^2(1-\lambda)+0+0-(0+0+0) = (2-\lambda)^2(1-\lambda).$$

Then setting $\det(A-\lambda I) = 0$ yields $\lambda_1 = 2$ and $\lambda_2 = 1$. An eigenvector x for λ_1 solves

$$(A-2I)x = 0$$

or equivalently

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

So $x_1 = -x_3$ and x_2, x_3 are free. Hence, we can pick $(1,0,-1)^T$ and $(0,1,0)^T$ as our representative two eigenvectors corresponding to λ_1. That is,

$$\begin{aligned} E_{\lambda_1}(A) &= \{(-x_3,0,x_3)^T, (0,x_2,0)^T \mid x_2, x_3 \in \mathbb{R}\} = \{x_3(-1,0,1)^T, x_2(0,1,0)^T \mid x_2, x_3 \in \mathbb{R}\} \\ &= \text{span}\{(-1,0,1)^T, (0,1,0)^T\} = \text{span}\{(1,0,-1)^T, (0,1,0)^T\}. \end{aligned}$$

An eigenvector y for $\lambda_2 = 1$ will satisfy $(A-I)y = 0$, meaning

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Hence, $y_1 = 0$ and $y_3 = -y_2$ with y_2 free. Then

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ y_2 \\ -y_2 \end{pmatrix} = y_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

for any $y_2 \in \mathbb{R}$. We can pick $(0,1,-1)^T$ as the representative eigenvector for $\lambda_2 = 1$. Therefore, one possible eigendecomposition for A is

$$A = Y\Lambda Y^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}^{-1}.$$

How can we quickly compute Y^{-1}? Gaussian elimination is probably the fastest way:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right).$$

Hence,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix}$$

as desired.

Exercise for Audience 2.5. Do all matrices have an eigendecomposition?

Answer: No! Take $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. No matter what we do, we cannot find an invertible matrix Y such that $A = Y \operatorname{diag}(\lambda_1(A), \lambda_2(A)) Y^{-1}$. Why is that?

Theorem 2.6. *An $n \times n$ matrix A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors.*

Proof. ($\Rightarrow$) Suppose A is diagonalizable, so $A = Y \Lambda Y^{-1}$. By the previous computation, we showed that the columns $\vec{y_1}, \dots, \vec{y_n}$ of Y are eigenvectors of A. Since Y is invertible, what does this mean?

Since Y is invertible, its columns form a linearly independent set (by a previous HW), so the columns$\vec{y_1}, \dots, \vec{y_n}$ of Y are independent. But this precisely means that A has n linearly independent eigenvectors.

($\Leftarrow$) Suppose A has n linearly independent eigenvectors $\vec{y_1}, \dots, \vec{y_n}$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Then by writing

$$Y = \left(\begin{array}{c|c|c} | & & | \\ \vec{y_1} & \cdots & \vec{y_n} \\ | & & | \end{array}\right) \qquad \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

we have that $AY = Y\Lambda$. However, since the columns $\vec{y_1}, \dots, \vec{y_n}$ are linearly independent, Y is invertible. So $A = Y \Lambda Y^{-1}$. ☐

So what went wrong in the previous example? We had $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ with eigenvalues $\lambda_1 = \lambda_2 = 1$. So the eigenvectors of A are found by solving $A\vec{y} = \lambda \vec{y}$. Then

$$(A - \lambda I)\vec{y} = (A - (1)I)\vec{y} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \vec{y} = \vec{0}.$$

So $y_2 = 1$, and y_1 can be anything. Thus, the eigenspace $E_1(A) = \{(y_1, 0)^T \mid y_1 \in \mathbb{R} \text{ (or } \mathbb{C})\} = \text{span}\{(1,0)^T\}$. Let us pick $v_1 = (1,0)^t$ for the eigenvector of λ_1. However, since $\lambda_2 = 1$ we also have $v_2 \in \text{span}\{(1,0)^T\}$, so no matter what v_2 we try to pick we will not have $\{v_1, v_2\}$ be independent.

It is also true that *symmetric matrices are always diagonalizable.* This will be proved later in the course using the Schur decomposition.

Exercise for Audience 2.7. Suppose A has an eigendecomposition. Is this decomposition unique? What about when A has distinct eigenvalues?

Answer: No! Take $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then setting $\det(A - \lambda I) = 0$ yields $\lambda = \pm 1$. Let $\lambda_1 = 1$ and $\lambda_2 = -1$. Then the corresponding eigenspaces are

$$E_{\lambda=1}(A) = \text{span}\{(1,1)^T\} \qquad E_{\lambda=-1}(A) = \text{span}\{(1,-1)^T\}.$$

So if we pick $Y = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, we have

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}.$$

However, if we instead pick representative eigenvalues $(20,20)^T \in \text{span}\{(1,1)^T\} = E_{\lambda=1}(A)$ and $(5,-5)^T \in \text{span}\{(1,-1)^T\} = E_{\lambda=-1}(A)$, we can write

$$A = \begin{pmatrix} 20 & 5 \\ 20 & -5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 20 & 5 \\ 20 & -5 \end{pmatrix}^{-1}.$$

However, notice that the diagonal matrix Λ is unchanged. If you order the diagonal elements of Λ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, then Λ is unique. (We need to pick a consistent way to order the λ_i, or else we could get "different" factorizations by swapping diagonal elements.)

Exercise for Audience 2.8. Does A diagonalizable $\Rightarrow$ A invertible?

Answer: No! Take $A = (0)_{n \times n}$.

Exercise for Audience 2.9. Does A invertible $\Rightarrow$ A diagonalizable?

Answer: No! Take $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ again. We have $\det(A) = 1 \neq 0$, so A is invertible but not diagonalizable.

Exercise for Audience 2.10. Can a nonsingular matrix with repeated eigenvalues be diagonalizable?

Answer: Yes! Take $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

The point of all these examples is that *there is no connection between invertibility and diagonalizability. The number of linearly independent eigenvectors is not a reflection of the rank of a matrix.*

3. Application to a PDEs Problem

For an $n \times n$ matrix A, consider PDEs of the form

$$\frac{d\vec{y}}{dt} = A\vec{y}$$

where $\vec{y} = (y_1(t), \ldots, y_n(t))^T$ and $\frac{d\vec{y}}{dt} = (\frac{dy_1(t)}{dt}, \ldots, \frac{dy_n(t)}{dt})^T$. As you saw in HW and lecture, PDEs of this form have solutions $\vec{y}(t) = e^{At}\vec{C}_0$, where $C_0 = \vec{y}(0)$ is the vector of initial conditions.

From Taylor expanding, we know

$$e^{At} = I + (At) + \frac{1}{2}(At)^2 + \frac{1}{3!}(At^3) + \cdots$$

and if we are lucky, the higher powers of A will be zero, so this approximation can be truncated. But that does not always happen. How do we know how many terms to keep in the truncation? How is this computation improved when A is diagonalizable?

If $A = V\Lambda V^{-1}$ then

$$\begin{aligned} e^{At} = e^{V\Lambda V^{-1}t} &= I + (V\Lambda V^{-1}t) + \frac{1}{2}(V\Lambda V^{-1}t)^2 + \frac{1}{3!}(V\Lambda V^{-1}t)^3 + \cdots \\ &= I + (V\Lambda V^{-1}) + \frac{1}{2}(V(\lambda t)^2 V^{-1}) + \frac{1}{3!}(V(\Lambda t)^3 V^{-1}) + \cdots \\ &= V(I + \Lambda t + \frac{1}{2}(\Lambda t)^2 + \frac{1}{3!}(\Lambda t)^3 + \cdots)V^{-1} \\ &= V(e^{\Lambda t})V^{-1}. \end{aligned}$$

Furthermore,

$$(e^{\Lambda t}) = \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix}$$

and so

$$e^{At} = V \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} V^{-1}$$

when A is diagonalizable. This computation is exact and *often easier to manage.*

Example 3.1. Let us try $Y = \begin{pmatrix} 1 & 1 \\ 0 & -3 \end{pmatrix}$ and $\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}$. This yields

$$e^{At} = \begin{pmatrix} 1 & 1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{-2t} \end{pmatrix} = \begin{pmatrix} e^t & \frac{1}{3}e^t - \frac{1}{3}e^{-2t} \\ 0 & e^{-2t} \end{pmatrix}.$$