Linear Algebra & Differential Equations

Elif Uskuplu

If you see any mistake, please email me (euskuplu@usc.edu).

---

11 Week 11

For the motivation of our new problem, let's consider a basic differential equation

$$F'(x) = aF(x)$$

It has a solution

$$F(x) = ce^{ax}$$

When we generalize this to

$$\begin{bmatrix} F_1'(x) \\ \vdots \\ F_n'(x) \end{bmatrix} = A \begin{bmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{bmatrix}$$

where A is an $n \times n$ matrix.

If we can make the equation like before, namely

$$\begin{bmatrix} F_1'(x) \\ \vdots \\ F_n'(x) \end{bmatrix} = \lambda \begin{bmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{bmatrix}$$

where λ is a scalar, we have a similar solution

$$\begin{bmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{bmatrix} = \begin{bmatrix} c_1 e^{\lambda x} \\ \vdots \\ c_n e^{\lambda x} \end{bmatrix}$$

So we want to make $A\vec{x}$ as the same as $\lambda\vec{x}$.

Also, whenever we try to solve $A\vec{x} = \vec{b}$, the simpler A is the easier the solution is. So, if we make A "simpler" to a diagonal matrix, it would make the elimination easier.

The next sections will give necessary tools for these.

11.1 Eigenvalue & Eigenvector problem

Definition 11.1. *Let A be an $n \times n$ matrix. Any values of λ for which*

$$A\vec{v} = \lambda\vec{v}$$

has nontrivial solutions $\vec{v}$ are called eigenvalues of A. The corresponding nonzero vectors $\vec{v}$ are called eigenvectors of A.

Example. $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$. Then $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. So $\lambda = 4$ is an eigenvalue of A, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to 4.

However, $A \begin{bmatrix} 1 \\ -1 \end{bmatrix} \neq \lambda \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ for any λ. So $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is not an eigenvector of A.

Also, $A \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} = -1 \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. So $\lambda = -1$ is an eigenvalue of A, and $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ is an eigenvector of A corresponding to -1.

How to solve eigenvalue/eigenvector problem?

First, we want that $A\vec{v} = \lambda\vec{v}$ has a nontrivial solution. On the other hand, we have

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

The last one is a usual homogeneous system, and it has a nontrivial solution if and only if $\det(A - \lambda I) = 0$. Therefore, to solve the problem,

1) Solve $\det(A - \lambda I) = 0$

2) If λ_1, λ_k are solutions, then solve

$$(A - \lambda_i I)\vec{v} = \vec{0}$$

Definition 11.2. *For a given $n \times n$ matrix A, the polynomial*

$$p(\lambda) = \det(A - \lambda I)$$

*is called the **characteristic polynomial** of A, and the equation*

$$p(\lambda) = 0$$

*is called the **characteristic equation** of A.*

Examples.

1. To find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 7 & 1 \\ 6 & 2 \end{bmatrix},$$

we proceed as follows:

1) Solve the characteristic equation $\det(A - \lambda I) = 0$:

The characteristic equation for matrix A is:

$$\lambda^2 - 9\lambda + 8 = 0.$$

Solving this quadratic equation gives us the eigenvalues:

$$\lambda_1 = 1,$$

$$\lambda_2 = 8.$$

2) Solve $(A - \lambda_i I)\vec{v} = \vec{0}$ for each λ_i:

For $\lambda_1 = 1$, the eigenvector $\vec{v_1}$ is found by solving the system:

$$(A - \lambda_1 I)\,\vec{v} = \vec{0}$$

which simplifies to:

$$\begin{bmatrix} 6 & 1 \\ 6 & 1 \end{bmatrix} \vec{v_1} = \vec{0}.$$

The solution to this system gives the eigenvector:

$$\vec{v_1} = \begin{bmatrix} x \\ -6x \end{bmatrix}.$$

For $\lambda_2 = 8$, the eigenvector $\vec{v_2}$ is found by solving the system:

$$(A - \lambda_2 I)\,\vec{v} = \vec{0}$$

which simplifies to:

$$\begin{bmatrix} -1 & 1 \\ 6 & -6 \end{bmatrix} \vec{v_2} = \vec{0}.$$

The solution to this system gives the eigenvector:

$$\vec{v_2} = \begin{bmatrix} x \\ x \end{bmatrix}.$$

Thus, we have found the eigenvalues and eigenvectors for the matrix A.

2. To find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 7 & 9 \\ -1 & 3 \end{bmatrix},$$

we proceed as follows:

1) Solve the characteristic equation $\det(A - \lambda I) = 0$:

The characteristic equation is obtained by calculating the determinant of $A - \lambda I$, which leads to:

$$\lambda^2 - 10\lambda + 30 = 0.$$

Solving this quadratic equation gives us the eigenvalues:

$$\lambda_1 = 5 - \sqrt{5}i,$$

$$\lambda_2 = 5 + \sqrt{5}i.$$

2) Solve $(A - \lambda_i I)\vec{v} = \vec{0}$ for each λ_i:

For $\lambda_1 = 5 - \sqrt{5}i$, we find the eigenvector $\vec{v_1}$ by solving the system:

$$(A - \lambda_1 I)\,\vec{v} = \vec{0}$$

which simplifies to:

$$\begin{bmatrix} 2 + \sqrt{5}i & 9 \\ -1 & -2 + \sqrt{5}i \end{bmatrix} \vec{v_1} = \vec{0}.$$

The solution to this system gives the eigenvector:

$$\vec{v_1} = \begin{bmatrix} (-2 + \sqrt{5}i)y \\ y \end{bmatrix}.$$

For $\lambda_2 = 5 + \sqrt{5}i$, we solve the system:

$$(A - \lambda_2 I)\,\vec{v} = \vec{0}$$

which simplifies to:

$$\begin{bmatrix} 2 - \sqrt{5}i & 9 \\ -1 & -2 - \sqrt{5}i \end{bmatrix} \vec{v_2} = \vec{0}.$$

The solution to this system gives the eigenvector:

$$\vec{v_2} = \begin{bmatrix} (-2 - \sqrt{5}i)y \\ y \end{bmatrix}.$$

Thus, we have found the eigenvalues and eigenvectors for the matrix A.

3. To find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 10 & -12 & 8 \\ 0 & 2 & 0 \\ -8 & 12 & -6 \end{bmatrix},$$

we follow these steps:

1) Solve the characteristic equation $\det(A - \lambda I) = 0$:

The characteristic polynomial for matrix A is:

$$-\lambda^3 + 6\lambda^2 - 12\lambda + 8 = -(\lambda - 2)^3.$$

Solving this cubic equation, we find the eigenvalue to be:

$$\lambda = 2.$$

2) Solve $(A - \lambda I)\vec{v} = \vec{0}$ for $\lambda = 2$:

We set up the system:

$$(A - 2I)\,\vec{v} = \vec{0},$$

which results in the following matrix equation after substituting $\lambda = 2$:

$$\begin{bmatrix} 8 & -12 & 8 \\ 0 & 0 & 0 \\ -8 & 12 & -8 \end{bmatrix} \vec{v} = \vec{0}.$$

Solving this system gives us the eigenvectors:

$$\vec{v} = \begin{bmatrix} \frac{3}{2}y - z \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ 1 \\ 0 \end{bmatrix} y + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} z$$

For any different y, z, we have a different eigenvectors. These eigenvectors correspond to the eigenvalue $\lambda = 2$ for the matrix A.

11.2 Non-defective & Defective matrices

Let consider an $n \times n$ matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

Its characteristic polynomial is

$$p(\lambda) = \det(\lambda I - A) = \det \begin{bmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{bmatrix},$$

and this can be written as follows:

$$p(\lambda) = (-1)^n \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0.$$

Fact 1. *Fundamental Theorem of Algebra* says that $p(\lambda) = 0$ has always a solution in $\mathbb{C}$. In other words, a matrix A can have complex eigenvalues. We can compute the complex eigenvectors using the following theorem.

Theorem 11.1. *Let A be an $n \times n$ matrix with real elements. If λ is a complex eigenvalue of A with corresponding eigenvector v, then $\bar{\lambda}$ is also an eigenvalue of A with corresponding eigenvector $\bar{v}$.*

Proof. If $Av = \lambda v$ then $\bar{A}\bar{v} = \bar{\lambda}\bar{v}$. Since A has real coefficients, we get $A\bar{v} = \bar{\lambda}\bar{v}$. □

Fact 2. If $\lambda_1, \ldots, \lambda_k$ are roots of $p(\lambda)$, we can also write $p(\lambda)$ as

$$p(\lambda) = (-1)^n(\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$$

where $m_1 + \cdots + m_k = n$. The number m_i is called **the algebraic multiplicity of** λ_i.

Fact 3. For a given eigenvalue λ of $A \in M_n(\mathbb{C})$, let E_λ be the set of all vectors v satisfying

$$Av = \lambda v.$$

Then E_λ is called the **eigenspace** of A corresponding to the eigenvalue λ. In other words,

$$E_\lambda = \text{nullspace}(A - \lambda I_n).$$

Then $\dim(E_\lambda)$ is called the **geometric multiplicity of** λ**.** Note that we have E_λ is a subspace of $\mathbb{C}^n$ as a complex vector space.

Fact 4. Eigenvectors corresponding to distinct eigenvalues are linearly independent.[1]

Now, there are some natural questions: what if there are less than n eigenvalues, or the total number of linearly independent eigenvectors is less than n?

Definition 11.3. *A $n \times n$ matrix A that has n linearly independent eigenvectors is called* ***non-defective****, or A has less than n linearly independent eigenvectors, A is called* ***defective****.*

Now, we have a new problem!!!

For given $n \times n$ matrix A, determine whether A is non-defective.

We have an effective solution method for such a problem.

1. Find the eigenvalues of A.
    - If there are n distinct eigenvalues, then A is non-defective because they produce n linearly independent eigenvectors.
    - Suppose there are less than n distinct eigenvalues. Then we need to go the next step.
2. Find the corresponding eigenvectors.

[1] The proof is written on the textbook.

- If geometric multiplicity equals algebraic multiplicity for each eigenvalue, then A is non-defective because $m_1 + m_2 + \ldots + m_n = n$ and we have n independent eigenvectors.
- Otherwise, we can decide A is defective.

Example. Let $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & -3 \\ 0 & -2 & 1 \end{bmatrix}$. We'll determine whether A is non-defective. First we need to solve the characteristic equation to find its eigenvalues.

$$\begin{aligned} \det \begin{bmatrix} 4-\lambda & 0 & 0 \\ 0 & 2-\lambda & -3 \\ 0 & -2 & 1-\lambda \end{bmatrix} &= (4-\lambda)\left[(2-\lambda)(1-\lambda) - 6\right] \\ &= (4-\lambda)\left[\lambda^2 - 3\lambda + 2 - 6\right] \\ &= (4-\lambda)(\lambda^2 - 3\lambda - 4) \\ &= (4-\lambda)(\lambda-4)(\lambda+1) = -(4-\lambda)^2(\lambda+1) \end{aligned}$$

So we have two eigenvalues: $\lambda_1 = 4$ with algebraic multiplicity 2 and $\lambda_2 = -1$ with algebraic multiplicity 1.

Maybe A is defective, we need further check.

If $\lambda_4 = 4$, we need to solve the following system.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & -3 \\ 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

It means x can be arbitrary and $-2y - 3z = 0 \Rightarrow y = -\frac{3}{2}z$. So all eigenvectors are of the form $\mathbf{v} = (x, -\frac{3}{2}z, z)$. In other words,

$$E_4 = \text{span}\{(1,0,0), (0, -\frac{3}{2}, 1)\}.$$

This means that the geometric multiplicity is 2 for λ_2.

If $\lambda_2 = -1$, we need to solve the following system.

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & -3 \\ 0 & -2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

It means $x = 0$, and $y = z$.So all eigenvectors are of the form $\mathbf{v} = (0, y, y)$. In other words,

$$E_2 = \text{span}\{(0,1,1)\}.$$

This means that the geometric multiplicity is 1 for λ_2.

As a result, we achieve that A is non-defective.

Examples. Let's take the examples in the previous section. Read the discussions about their eigenvalue/eigenvector solutions. Then we get

1. $A = \begin{bmatrix} 7 & 1 \\ 6 & 2 \end{bmatrix}$ is non-defective.
2. $A = \begin{bmatrix} 7 & 9 \\ -1 & 3 \end{bmatrix}$ is non-defective.
3. $A = \begin{bmatrix} 10 & -12 & 8 \\ 0 & 2 & 0 \\ -8 & 12 & -6 \end{bmatrix}$ is defective.

11.3 Diagonalization

Definition 11.4. *Let A and B be $n \times n$ matrices. We say A is similar to B if there exists an invertible matrix S such that*

$$S^{-1}AS = B.$$

Example. $A = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 56 & -95 \\ 33 & -56 \end{bmatrix}$ are similar because if $S = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$, then we get

$$S^{-1}AS = B.$$

Theorem 11.2. *If A and B are similar, then they have the same eigenvalues.*

Proof. On the textbook. □

First of all, the converse is not true. For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$, then they have the same eigenvalue which is 1, but if there is an invertible S such that

$$S^{-1}AS = B$$

then we get

$$A = SBS^{-1} = SS^{-1} = B$$

but $A \neq B$. So the theorem is not biconditional. However, we can consider a special case.

Suppose A has eigenvalues $\lambda_1, \ldots, \lambda_n$ (they are not necessarily distinct). Then A and

$$\operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

have the same eigenvalues. We want to achieve a condition that makes A and $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ similar.

Theorem 11.3. *An $n \times n$ matrix A is similar to a diagonal matrix if and only if A is nondefective. In this case, A is called **diagonalizable**.*

In such case, $S = [v_1, \dots, v_n]$ where $\{v_1, \dots, v_n\}$ is the set of n linearly independent eigenvectors and

$$S^{-1}AS = \textit{diag}(\lambda_1, \dots, \lambda_n)$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A (not necessarily distinct).

The proof indeed gives the necessary algorithm. Instead of giving the whole proof, we will apply the same idea in the following example:

Example 1.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -7 \\ 1 & 1 & -3 \end{bmatrix}$$

Then the characteristic equation of A is

$$p(\lambda) = (\lambda - 1)(\lambda - 4)(\lambda + 4).$$

Since A has three distinct eigenvalues, A is nondefective, so A is diagonalizable. We need to find S to achieve similarity. The details are exercises but we have $\lambda_1 = 1, \lambda_2 = 4, \lambda_3 = -4$ and

$$E_1 = \text{span}\{(-15, -7, 2)\},$$
$$E_2 = \text{span}\{(0, 7, 1)\},$$
$$E_3 = \text{span}\{(0, -1, 1)\}.$$

Now we have

$$A \begin{bmatrix} -15 \\ -7 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} -15 \\ -7 \\ 2 \end{bmatrix},$$
$$A \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix},$$
$$A \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = -4 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

In other words,

$$\left[A \begin{bmatrix} -15 \\ -7 \\ 2 \end{bmatrix}, \; A \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix}, \; A \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right] = \left[\begin{bmatrix} -15 \\ -7 \\ 2 \end{bmatrix}, \; 4 \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix}, \; -4 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right],$$

$$\Rightarrow A \begin{bmatrix} -15 & 0 & 0 \\ -7 & 7 & -1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -15 & 0 & 0 \\ -7 & 7 & -1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -4 \end{bmatrix}.$$

Let's take $S = \begin{bmatrix} -15 & 0 & 0 \\ -7 & 7 & -1 \\ 2 & 1 & 1 \end{bmatrix}$. Since the columns are independent, S is invertible and we get

$$AS = S \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -4 \end{bmatrix},$$

and hence

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -4 \end{bmatrix}.$$

So A is diagonalizable. For the reverse direction in the theorem, we apply the proof idea in the following example.

Example 2.

$$A = \begin{bmatrix} -2 & 1 & 4 \\ -2 & 1 & 4 \\ -2 & 1 & 4 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We have

$$S^{-1}AS = B.$$

We can deduce A and B are similar, so they have the same eigenvalues, which are 3 and 0. Also, we have $AS = SB$ namely they have equal columns. This means

$$AS_1 = 3S_1, \quad AS_2 = 0S_2, \quad AS_3 = 0S_3.$$

In other words, the eigenspace corresponded to 3 is span$\{S_1\}$, and the eigenspace corresponded to 0 is span$\{S_2, S_3\}$. This means A is nondefective.

To sum up, in order to determine whether A is diagonalizable or not, we should determine whether A is nondefective or not. If A is nondefective,

1) A is similar to diag$(\lambda_1, \ldots, \lambda_n)$ where $\lambda_1, \ldots, \lambda_n$ are eigenvalues (not necessarily distinct).

2) A is similar to diag$(\lambda_1, \ldots, \lambda_n)$ by S where the columns of S are the independent corresponded eigenvectors of A. When you take the columns, be careful about the order. The columns should be eigenvectors for λ_1, eigenvectors for λ_2, ..., and eigenvectors for λ_n.

Example.

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

It has characteristic polynomial $p(\lambda) = (\lambda + 1)(\lambda - 3)^2$. Therefore, A has two eigenvalues $\lambda_1 = -1, \lambda_2 = 3$.

For $\lambda_1 = -1$:

$$\begin{bmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \vec{v} = \vec{0} \Rightarrow \text{Then we have } \vec{v} = (x, x, 0) \text{ for } x \in \mathbb{R}.$$

This means $E_1 = \text{span}\{(1, 1, 0)\}$.

For $\lambda_2 = 3$:

$$\begin{bmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{v} = \vec{0} \Rightarrow \text{Then we have } \vec{v} = (x, -x, z) \text{ for } x, z \in \mathbb{R}.$$

This means $E_2 = \text{span}\{(1, -1, 0), (0, 0, 1)\}$.

Since algebraic multiplicity = geometric multiplicity for each eigenvalue of A, A is non-defective. Also, A is diagonalizable. Take $S = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and then we get

$$S^{-1}AS = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$