LECTURE 32

Systems of First Order Linear ODEs

In this lecture we will consider first order ordinary differential equations in which more than one unknown function occurs. Let's begin with a definition

DEFINITION 32.1. *An $n \times n$* ***system of first order linear ODEs*** *is a set of n differential equations involving n unknown functions $y_1, \ldots, y_n$ of the form*

$$\begin{aligned}
\frac{dy_1}{dt} + a_{11}(t)\,y_1(t) + a_{12}(t)\,y_2(t) + \cdots + a_{1n}(t)\,y_n(t) &= g_1(t) \\
\frac{dy_2}{dt} + a_{21}(t)\,y_1(t) + a_{22}(t)\,y_2(t) + \cdots + a_{2n}(t)\,y_n(t) &= g_2(t) \\
&\vdots \\
\frac{dy_n}{dt} + a_{n1}(t)\,y_1(t) + a_{n2}(t)\,y_2(t) + \cdots + a_{nn}(t)\,y_n(t) &= g_n(t)
\end{aligned}$$

Differential equations are crucial to modeling a variety of physical phenomena, where the rate of growth of one quantity depends on other quantities as well as itself. For example, the rate at which a population of rabbits changes could depend not only on the number of rabbits (reproducing) but also on the population of foxes that prey on the rabbits.

In this lecture we will concentrate on a particularly simple case where the coefficient functions a_{ij} are constants, and the "driving functions" $g_i(t)$ are all zero. Thus, we'll be considering differential equations of the form

$$\begin{aligned}
\frac{dy_1}{dt} &= A_{11}y_1 + \cdots + A_{1n}y_n \qquad (32.1) \\
\frac{dy_2}{dt} &= A_{21}y_1 + \cdots + A_{2n}y_n \\
&\vdots \qquad (32.2) \\
\frac{dy_n}{dt} &= A_{n1}y_1 + \cdots + A_{nn}y_n \qquad (32.3)
\end{aligned}$$

This simple case, nevertheless, will be sufficient to illustrate the general technique for larger systems of ODEs.

1. Matrix Formulation

From the unknown functions $y_1, y_2, \ldots, y_n$ and their derivatives $\frac{dy_1}{dt}, \frac{dy_2}{dt}$ we form two vectors (depending on the underlying variable t)

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad , \quad \dot{\mathbf{y}} = \begin{bmatrix} \frac{dy_1}{dt} \\ \vdots \\ \frac{dy_n}{dt} \end{bmatrix}$$

and we arrange the coefficients A_{ij} as the entries of a $n \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

Then in the language of linear algebra the set of n differential equations (32.1) can be rewritten as a single matrix equation

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y}$$

2. An Easy Case

Suppose the matrix $\mathbf{A}$ corresponding to an $n \times n$ system of first order differential equations is diagonal; that is to say, of the form

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

with non-zero entries only along the diagonal running from the upper left to the lower right. For such a matrix the corresponding set of first order ODEs is easy to solve. The corresponding equations will be

$$\begin{aligned} \frac{dy_1}{dt} &= \lambda_1 y_1 \\ \frac{dy_2}{dt} &= \lambda_2 y_2 \\ &\vdots \\ \frac{dy_n}{dt} &= \lambda_n y_n \end{aligned}$$

Notice that each of these equations depends on only one of the unknown functions and its derivative; and indeed each of the differential equations is just the differential equation for an exponential function. Solving these equations one by one we get

$$\begin{aligned} \frac{dy_1}{dt} &= \lambda_1 y_1 && \Rightarrow && y_1 = C_1 e^{\lambda_1 t} \\ \frac{dy_2}{dt} &= \lambda_2 y_2 && \Rightarrow && y_2 = C_2 e^{\lambda_2 t} \\ &\vdots \\ \frac{dy_n}{dt} &= \lambda_n y_n && \Rightarrow && y_n = C_n e^{\lambda_n t} \end{aligned}$$

and so we arrive at a complete solution involving n arbitrary constants $C_1, \ldots, C_n$ (which can be interpreted as the initial values of $y_1, \ldots, y_n$ at time $t = 0$).

3. The General Case

What's beautiful about systems of differential equations of the form (32.1) is that after a little linear algebra, solving even a complicated set of differential equations can be reduced to easy case solved above; where the coefficient matrix is diagonal.

Here's a sketch of how this will work. Suppose we had an invertible matrix $\mathbf{C}$ such that

$$\mathbf{C}^{-1}\mathbf{AC} = \mathbf{D} = \left[\begin{array}{cccc} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{array}\right]$$

$\mathbf{D}$ bein a diagonal matrix. Then as above we could simple write down the general solution of

$$\dot{\mathbf{z}} = \mathbf{Dz} \qquad \Rightarrow \qquad \left\{\begin{array}{ll} z_1 & = C_1 e^{\lambda_1 t} \\ z_2 & = C_2 e^{\lambda_2 t} \\ \vdots & \\ z_n & = C_n e^{\lambda_n t} \end{array}\right.$$

Let $\mathbf{z}(t)$ be such a solution, and consider the vector $\mathbf{y}(t)$ obtained by multiplying $\mathbf{z}(t)$ from the left by the matrix $\mathbf{C}$

$$\mathbf{y} = \mathbf{Cz}$$

Then

$$\dot{\mathbf{y}} = \frac{d}{dt}(\mathbf{Cz}) = \mathbf{C}\frac{d}{dt}\mathbf{z} = \mathbf{C}\dot{\mathbf{z}} = \mathbf{C}(\mathbf{Dz}) = \mathbf{C}\left(\mathbf{C}^{-1}\mathbf{AC}\right)\mathbf{z} = \left(\mathbf{CC}^{-1}\right)\mathbf{ACz} = \mathbf{IACz} = \mathbf{A}(\mathbf{Cz}) = \mathbf{Ay}$$

In other words, $\mathbf{y} = \mathbf{Cz}$ will be a solution of our original differential equation.

$$\dot{\mathbf{y}} = \mathbf{Ay}$$

Thus, systems of the form (32.1) can be easily solved if we can find an invertible matrix $\mathbf{C}$ such that $\mathbf{C}^{-1}\mathbf{AC}$ is a diagonal matrix.

4. Diagonalization of Matrices

Recall that a **diagonal matrix** is a square $n \times n$ matrix with non-zero entries only along the diagonal from the under left to the lower right (the *main diagonal*).

Diagonal matrices are particularly convenient for eigenvalue problems since the eigenvalues of a diagonal matrix

$$\mathbf{A} = \left[\begin{array}{cccc} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{array}\right]$$

coincide with the diagonal entries $\{a_{ii}\}$ and the eigenvector corresponding the eigenvalue a_{ii} is just the i^{th} coordinate vector.

Example 32.2. Find the eigenvalues and eigenvectors of

$$\mathbf{A} = \left[\begin{array}{cc} 2 & 0 \\ 0 & 3 \end{array}\right]$$

- The characteristic polynomial is

  $$P_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\left[\begin{array}{cc} 2-\lambda & 0 \\ 0 & 3-\lambda \end{array}\right] = (2-\lambda)(3-\lambda)$$

  Evidently $P_{\mathbf{A}}(\lambda)$ has roots at $\lambda = 2, 3$. The eigenvectors corresponding to the eigenvalue $\lambda = 2$ are solutions of

  $$(\mathbf{A} - (2)\mathbf{I})\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \quad \Rightarrow \quad x_2 = 0$$

  $$\Rightarrow \quad \mathbf{x} \in span\left(\left[\begin{array}{c} 1 \\ 0 \end{array}\right]\right)$$

The eigenvectors corresponding to the eigenvalue $\lambda = 3$ are solutions of

$$(\mathbf{A} - (3)\mathbf{I})\,\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad -x_1 = 0$$

$$\Rightarrow \quad \mathbf{x} \in span\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

This property (that the eigenvalues of a diagonal matrix coincide with its diagonal entries and the eigenvectors corresponds to the corresponding coordinate vectors) is so useful and important that in practice one often tries to make a change of coordinates just so that this will happen. Unfortunately, this is not always possible; however, if it is possible to make a change of coordinates so that a matrix becomes diagonal we say that the matrix is *diagonalizable*. More formally,

LEMMA 32.3. *Let* $\mathbf{A}$ *be a real (or complex)* $n \times n$ *matrix, let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be a set of* n *real (respectively, complex) scalars, and let* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *be a set of* n *vectors in* $\mathbb{R}^n$ *(respectively,* $\mathbb{C}^n$ *). Let* $\mathbf{C}$ *be the* $n \times n$ *matrix formed by using* $\mathbf{v}_j$ *for* j^{th} *column vector, and let* $\mathbf{D}$ *be the* $n \times n$ *diagonal matrix whose diagonal entries are* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *. Then*

$$\mathbf{AC} = \mathbf{CD}$$

if and only if $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are the eigenvalues of* $\mathbf{A}$ *and each* $\mathbf{v}_j$ *is an eigenvector of* $\mathbf{A}$ *correponding the eigenvalue* λ_j *.*

Proof. Under the hypotheses

$$\begin{aligned} \mathbf{AC} &= \mathbf{A} \begin{bmatrix} | & \cdots & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & \cdots & | \\ \mathbf{A}\mathbf{v}_1 & \cdots & \mathbf{A}\mathbf{v}_n \\ | & \cdots & | \end{bmatrix} \\ \mathbf{CD} &= \begin{bmatrix} | & \cdots & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = \left[\begin{bmatrix} | & \cdots & | \\ \lambda_1\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n \\ | & \cdots & | \end{bmatrix} \right] \end{aligned}$$

and so $\mathbf{AC} = \mathbf{CD}$ implies

$$\begin{aligned} \mathbf{A}\mathbf{v}_1 &= \lambda_1\mathbf{v}_1 \\ &\vdots \\ \mathbf{A}\mathbf{v}_n &= \lambda_n\mathbf{v}_n \end{aligned}$$

and vice-versa. □

Now suppose $\mathbf{AC} = \mathbf{CD}$, and the matrix $\mathbf{C}$ is invertible. Then we can write

$$\mathbf{D} = \mathbf{C}^{-1}\mathbf{AC}.$$

And so we can think of the matrix $\mathbf{C}$ as converting $\mathbf{A}$ into a diagonal matrix.

DEFINITION 32.4. *An* $n \times n$ *matrix* $\mathbf{A}$ *is* ***diagonalizable*** *if there is an invertible* $n \times n$ *matrix* $\mathbf{C}$ *such that* $\mathbf{C}^{-1}\mathbf{AC}$ *is a diagonal matrix. The matrix* $\mathbf{C}$ *is said to* ***diagonalize*** $\mathbf{A}$*.*

THEOREM 32.5. *An* $n \times n$ *matrix* $\mathbf{A}$ *is diagonalizable iff and only if it has* n *linearly independent eigenvectors.*

Proof. The argument here is very simple. Suppose $\mathbf{A}$ has n linearly independent eigenvectors. Then the matrix $\mathbf{C}$ formed by using these eigenvectors as column vectors will be invertible (since the rank of $\mathbf{C}$ will be equal to n). On the other hand, if $\mathbf{A}$ is diagonalizable then, by definition, there must be an invertible matrix $\mathbf{C}$ such that $\mathbf{D} = \mathbf{C}^{-1}\mathbf{AC}$ is diagonal. But then the preceding lemma says that the columns vectors of $\mathbf{C}$ must coincide with the eigenvectors of $\mathbf{A}$. Since $\mathbf{C}$ is invertible, these n column vectors must be linearly independent. Hence, $\mathbf{A}$ has n linearly independent eigenvectors. □

EXAMPLE 32.6. Find the matrix that diagonalizes

$$\mathbf{A} = \begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix}$$

- First we'll find the eigenvalues and eigenvectors of $\mathbf{A}$.

  $$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{bmatrix} 2-\lambda & 6 \\ 0 & -1-\lambda \end{bmatrix} = (2-\lambda)(-1-\lambda) \quad \Rightarrow \quad \lambda = 2, -1$$

  The eigenvectors corresponding to the eigenvalue $\lambda = 2$ are solutions of $(\mathbf{A} - (2)\mathbf{I})\,\mathbf{x} = \mathbf{0}$ or

  $$\begin{bmatrix} 0 & 6 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{c} 6x_2 = 0 \\ -3x_2 = 0 \end{array} \quad \Rightarrow \quad x_2 = 0 \quad \Rightarrow \quad \mathbf{x} = r\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

  The eigenvectors corresponding to the eigenvalue $\lambda = -1$ are solutions of $(\mathbf{A} - (-1)\mathbf{I})\,\mathbf{x} = \mathbf{0}$ or

  $$\begin{bmatrix} 3 & 6 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{c} 3x_1 + 6x_2 = 0 \\ 0 = 0 \end{array} \quad \Rightarrow \quad x_1 = -2x_2 \quad \Rightarrow \quad \mathbf{x} = r\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

  So the vectors $\mathbf{v}_1 = [1, 0]$ and $\mathbf{v}_2 = [-2, 1]$ will be eigenvectors of $\mathbf{A}$. We now arrange these two vectors as the column vectors of the matrix $\mathbf{C}$.

  $$\mathbf{C} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

  In order to compute the diagonalization of $\mathbf{A}$ we also need $\mathbf{C}^{-1}$. This we compute using the technique of Section 1.5:

  $$\left[\begin{array}{cc|cc} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \to R_1 + 2R_2} \left[\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 \end{array}\right] \quad \Rightarrow \quad \mathbf{C}^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

  Finally,

  $$\begin{aligned} \mathbf{D} &= \mathbf{C}^{-1}\mathbf{A}\mathbf{C} = \mathbf{C}^{-1}(\mathbf{A}\mathbf{C}) \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\left(\begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\right) \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

EXAMPLE 32.7. Find the general solution of the following system of differential equations

$$\begin{aligned} \frac{dy_1}{dt} &= y_1 + 4y_2 \\ \frac{dy_2}{dt} &= y_1 + y_2 \end{aligned}$$

The matrix formulation of this problem would be

$$\begin{bmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

And so we'll begin by finding a matrix $\mathbf{C}$ that diagonalizes $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}$

FIrst we find the eigenvalues of $\mathbf{A}$

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda\mathbf{I}) = (1-\lambda)^2 - 4 = 1 - 2\lambda + \lambda^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda+1)(\lambda-3) \\ \Rightarrow &\quad \lambda = -1, 3 \end{aligned}$$

Next, we find the corresponding eigenvectors

$\lambda = -1:$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (\mathbf{A} - (-1)\,\mathbf{I})\,\mathbf{v} = \begin{bmatrix} 1-(-1) & 4 \\ 1 & 1-(-1) \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 2v_1 + 4v_2 \\ v_1 + 2v_2 \end{bmatrix}$$

$$\Rightarrow \quad v_1 + 2v_2 = 0 \quad \Rightarrow \quad v_1 = -2v_2 \quad \Rightarrow \quad \mathbf{v} = v_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$\lambda = 3:$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (\mathbf{A} - (3)\,\mathbf{I})\,\mathbf{v} = \begin{bmatrix} 1-(3) & 4 \\ 1 & 1-(3) \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -2v_1 + 4v_2 \\ v_1 - 2v_2 \end{bmatrix}$$

$$\Rightarrow \quad v_1 - 2v_2 = 0 \quad \Rightarrow \quad v_1 = 2v_2 \quad \Rightarrow \quad \mathbf{v} = v_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Having found the eigenvectors and eigenvalues of $\mathbf{A}$ we can now write down the matrices $\mathbf{C}$ and $\mathbf{D}$

$$\mathbf{C} = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \quad , \quad \mathbf{D} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$$

The general solution of

$$\frac{d\mathbf{z}}{dt} = \mathbf{D}\mathbf{z}$$

will be

$$\mathbf{z} = \begin{bmatrix} c_1 e^{-t} \\ c_2 e^{3t} \end{bmatrix}$$

And so the general solution of $\frac{d\mathbf{y}}{dt} = \mathbf{A}\mathbf{y}$ will be

$$\mathbf{y} = \mathbf{C}\mathbf{z} = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 e^{-t} \\ c_2 e^{3t} \end{bmatrix} =: \begin{bmatrix} -2c_1 e^{-t} + 2c_2 e^{3t} \\ c_1 e^{-t} + c_2 e^{3t} \end{bmatrix}$$