LECTURE 14

# 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 upper 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 14.1. 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

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

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

$$\begin{aligned} (\mathbf{A} - (3)\mathbf{I})\,\mathbf{x} &= \mathbf{0} \quad \Rightarrow \quad \left[\begin{array}{cc} -1 & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \quad \Rightarrow \quad -x_1 = 0 \\ &\Rightarrow \quad \mathbf{x} \in span\left(\left[\begin{array}{c} 0 \\ 1 \end{array}\right]\right) \end{aligned}$$

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 14.2. *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* $\lambda_j$ *.*

*Proof.* Under the hypotheses

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

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 14.3. *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 14.4. *An* $n \times n$ *matrix* $\mathbf{A}$ *is diagonalizable if 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 column 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 14.5. Find the matrix that diagonalizes

$$\mathbf{A} = \left[\begin{array}{cc} 2 & 6 \\ 0 & -1 \end{array}\right]$$

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

  $$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\left[\begin{array}{cc} 2-\lambda & 6 \\ 0 & -1-\lambda \end{array}\right] = (2-\lambda)(-1-\lambda) \quad \Rightarrow \quad \lambda = 2, -1$$

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

  $$\left[\begin{array}{cc} 0 & 6 \\ 0 & -3 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \quad \Rightarrow \quad \begin{array}{c} 6x_2 = 0 \\ -3x_2 = 0 \end{array} \quad \Rightarrow \quad x_2 = 0 \quad \Rightarrow \quad \mathbf{x} = r\left[\begin{array}{c} 1 \\ 0 \end{array}\right]$$

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

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

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,

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

## 1. Criteria for Diagonalizability

EXAMPLE 14.6. Recall that in the preceding lecture we found the eigenvalues and eigenvectors for

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

What we found were, two possible eigenvalues and two corresponding eigenvectors (actually two possible eigenspaces).

$$\begin{aligned} \lambda &= -2 \quad \Rightarrow \quad \mathbf{v}_1 = \left[-2, -\frac{2}{3}, 1\right] \\ \lambda &= 1 \quad \Rightarrow \quad \mathbf{v}_2 = [0,1,0] \end{aligned}$$

In order to diagonalize $\mathbf{A}$ we need to construct an invertible $3 \times 3$ matrix $\mathbf{C}$ using the eigenvectors of $\mathbf{A}$ as the columns. However, we have only two linearly independent eigenvectors - so this construction is not going to work. In fact, $\mathbf{A}$ is not diagonalizable.

So an $n \times n$ matrix need not be diagonalizable. Nevertheless,

THEOREM 14.7. *Suppose the characteristic equation* $\det(\mathbf{A} - \lambda \mathbf{I}) = \mathbf{0}$ *for an* $n \times n$ *matrix* $\mathbf{A}$ *has* $n$ *distinct roots. Then* $\mathbf{A}$ *is diagonalizable.*

THEOREM 14.8. *Suppose* $\mathbf{A}$ *is a symmetric* $n \times n$ *matrix. Then each root of the characteristic equation for* $\mathbf{A}$ *is real and moreover* $\mathbf{A}$ *is diagonalizable.*

## 2. Applications

**2.1. Principal Axes of Inertia.** Consider a rigid body $B$ whose center of mass lies at the origin $\mathbf{0} \in \mathbb{R}^3$ and which is rotating with constant angular velocity $\boldsymbol{\omega}$. (The magitude of $\boldsymbol{\omega} \in \mathbb{R}^3$ is the speed of the rotation in radians/sec and the direction of $\boldsymbol{\omega}$ corresponds to the axes of rotation, oriented according to the right hand rule.) The rotational kinetic energy for such a system is given by

$$T_{rot} = \sum_{i=1}^{3} \sum_{j=1}^{3} I_{ij} \omega_i \omega_j$$

where the $I_{ij}$ are the components of the inertia tensor:

$$I_{ij} = \int_B \rho(\mathbf{x}) \left[\delta_{ij} \|\mathbf{x}\|^2 - x_i x_j\right] d\mathbf{x}$$

(here $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$). The angular momentum $\mathbf{L}$ of the body is given

$$L_i = \sum_{j=1}^{3} I_{ij}\omega_j$$

Interpreting the inertia tensor as a $3 \times 3$ matrix, and $\mathbf{L}$ and $\boldsymbol{\omega}$ as vectors in $\mathbb{R}^3$, we can write

$$\mathbf{L} = \mathbf{I}\boldsymbol{\omega}$$

where the multiplication on the right is just the multiplication of a vector in $\mathbb{R}^3$ by a $3 \times 3$ matrix.

Question 14.9. *When is* $\mathbf{L}$ *parallel to* $\boldsymbol{\omega}$*?*

A physical motivation for this question is as follows. Rotational motions can in general be rather complicated. For a body can be simultaneously spinning about one axes and tumbling about another axis (indeed, in gymnastics one often sees sommersaults combined with twists). And all of these motions will contribute to the total angular momentum. However, if the angular momentum is parallel to the angular velocity, the motion will be particularly simple; and mathematically this will happen whenever

$$\mathbf{I}\boldsymbol{\omega} = \mathbf{L} = \lambda\boldsymbol{\omega}$$

i.e., whenever $\boldsymbol{\omega}$ is an eigenvector of $\mathbf{I}$. The correponding directions are referred to as **principal axes of inertia**.

Example 14.10. Find the principle axes for a body whose inertia tensor is given by

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

, eigenvalues: $3, 1, 6$

- First we find the eigenvalues of $\mathbf{A}$:

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda\mathbf{I}) = (2-\lambda)(5-\lambda)(3-\lambda) - (2)(2)(3) \\ &= -\lambda^3 + 10\lambda^2 - 31\lambda + 18 \\ &= -(\lambda-1)(\lambda-3)(\lambda-6) \\ \lambda &= 1, 3, 6 \end{aligned}$$

- Next we find the eigenvectors:

$\lambda = 1:$

$$\begin{aligned} \mathbf{A} - (1)\mathbf{I} &= \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix} \rightarrow R.R.E.F.(\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ &\Rightarrow \left.\begin{array}{l} x_1 + 2x_2 = 0 \\ x_3 = 0 \\ 0 = 0 \end{array}\right\} \Rightarrow \mathbf{v}_{\lambda=1} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

$\lambda = 3$

$$\begin{aligned} \mathbf{A} - (3)\mathbf{I} &= \begin{bmatrix} -1 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow R.R.E.F.(\mathbf{A} - 3\mathbf{I}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \end{bmatrix} \\ &\Rightarrow \left.\begin{array}{l} x_1 = 0 \\ x_2 = 0 \\ 0 = 0 \end{array}\right\} \Rightarrow \mathbf{v}_{\lambda=3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

$\lambda = 6$

$$\mathbf{A} - (6)\,\mathbf{I} = \begin{bmatrix} -4 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \rightarrow R.R.E.F.\,(\mathbf{A} - \mathbf{I}) = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \left.\begin{array}{l} x_1 - \frac{1}{2}x_2 = 0 \\ x_3 = 0 \\ 0 = 0 \end{array}\right\} \Rightarrow \mathbf{v}_{\lambda=6} = \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}$$

- Each eigenvector $\mathbf{v}_\lambda$ of $\mathbf{A}$ corresponds to a principle axis of inertia, and the corresponding eigenvalue is the corresponding moment of inertia about that axis: Thus

| principle axis | moment of inertia |
|---|---|
| $[-2,1,0]$ | 1 |
| $[0,0,1]$ | 3 |
| $\left[\frac{1}{2},1,0\right]$ | 6 |

Notice that

$$\mathbf{v}_{\lambda_i} \cdot \mathbf{v}_{\lambda_j} = 0 \quad \text{if } i \neq j$$

so these three principle axes are perpendicular to each other. To see that this is always the case, suppose $\lambda \neq \lambda'$ and consider the triple product

$$\mathbf{v}_\lambda^T \mathbf{A} \mathbf{v}_{\lambda'}$$

Now we can evaluate this expression two ways: First of all

$$\mathbf{v}_\lambda^T \mathbf{A} \mathbf{v}_{\lambda'} = \mathbf{v}_\lambda^T (\mathbf{A}\mathbf{v}_{\lambda'}) = \mathbf{v}_\lambda^T \left(\lambda' \mathbf{v}_{\lambda'}\right) = \lambda' (\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'})$$

Alternatively, since $\mathbf{A} = \mathbf{A}^T$, we have

$$\mathbf{v}_\lambda^T \mathbf{A} = \left(\mathbf{A}^T \mathbf{v}_\lambda\right)^T = (\mathbf{A}\mathbf{v}_\lambda)^T = (\lambda \mathbf{v}_\lambda)^T$$

and so

$$\mathbf{v}_\lambda^T \mathbf{A} \mathbf{v}_{\lambda'} = \lambda (\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'})$$

But then we have

$$\lambda (\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'}) = \mathbf{v}_\lambda^T \mathbf{A} \mathbf{v}_{\lambda'} = \lambda' (\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'})$$

or

$$\left(\lambda - \lambda'\right)(\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'}) = 0$$

Since, by hypothesis $\lambda \neq \lambda'$, we must have $\mathbf{v}_\lambda \cdot \mathbf{v}_{\lambda'} = 0$

**2.2. Systems of ODEs.** Our study of eigenvalues/eigenvectors and diagonalization has another very useful application to the solution of systems of ordinary differential equations. In what follows below, we'll consider a system of two ordinary differential equations in two unknown functions; but it should be easy to see how to generalize this techniques to systems of $n$ differential equations in $n$ unknown functions.

Consider the following general system of first order ODEs:

$$\begin{aligned} \frac{dx_1}{dt}(t) &= a_{11}x_1(t) + a_{12}x_2(t) \\ \frac{dx_2}{dt}(t) &= a_{21}x_1(t) + a_{22}x_2(t) \end{aligned}$$

Such systems occur in a number of disparate contexts

- Chemistry. The rate at which the concentration of a reactant changes is proportional to its concentration and the concentration of another reactant.
- Biology. The rate at which a predator and prey populations changes is related to the populations of predators and prey.
- Physics. Coupled oscillators
- Electrical Engineering. Simple passive element circuits

If we have such a system, we can reformulate it as a matrix differential equation. To do this, we set

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad , \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

so that we can write the system as matrix equation.

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) \tag{1}$$

A particularly easy case is when $\mathbf{A}$ is diagonal.

$$\mathbf{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{2}$$

In this case, we say that the system is *decoupled*; because the differentiaion equations for such a system are of the form

$$\begin{aligned} \frac{dx_1}{dt} &= \lambda_1 x_1 + 0 \\ \frac{dx_2}{dt} &= 0 + \lambda_2 x_2 \end{aligned}$$

Such equations are easily solved, one-at-a-time,

$$x_1(t) = c_1 e^{\lambda_1 t} \tag{14.1}$$
$$x_2(t) = c_2 e^{\lambda_2 t} \tag{14.2}$$

It is the general (non-diagonal) case that we want to solve. This will do by using the diagonalization process to convert the problem to the easier solvable case of diagonal matrices.

Thus, suppose we have found the eigenvalues $\lambda_1, \lambda_2$ and eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ of the coefficient matrix $\mathbf{A}$, as well as the matrices $\mathbf{C}$ and $\mathbf{D}$ such that

$$\mathbf{C} = \begin{pmatrix} \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 \\ \downarrow & \downarrow \end{pmatrix} \quad , \quad \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

with

$$\mathbf{D} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C} \quad \Longleftrightarrow \quad \mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$$

Now consider the related system

$$\begin{pmatrix} \frac{dy_1}{dt} \\ \frac{dy_2}{dt} \end{pmatrix} = \frac{d}{dt}\mathbf{y}(t) = \mathbf{D}\mathbf{y}(t) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 y_1 \\ \lambda_2 y_2 \end{pmatrix} \tag{4}$$

This is a decoupled system and we have as its general solution

$$\mathbf{y}(t) = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \end{pmatrix} \tag{5}$$

Now consider

$$\mathbf{x}(t) = \mathbf{C}\mathbf{y}(t) \tag{6}$$

This vector function will satisfy

$$\begin{aligned}
\frac{d}{dt}\mathbf{x}(t) &= \frac{d}{dt}(\mathbf{C}\mathbf{y}(t)) \\
&= \mathbf{C}\frac{d}{dt}\mathbf{y}(t) \qquad \text{since } \mathbf{C} \text{ is a constant matrix} \\
&= \mathbf{C}(\mathbf{D}\mathbf{y}(t)) \\
&= \mathbf{C}\mathbf{D}\mathbf{C}^{-1}\mathbf{C}\mathbf{y}(t) \\
&= \left(\mathbf{C}\mathbf{D}\mathbf{C}^{-1}\right)(\mathbf{C}\mathbf{y}(t)) \\
&= \mathbf{A}\mathbf{x}(t)
\end{aligned}$$

That is to say,

$$\mathbf{x}(t) = \mathbf{C}\begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \end{pmatrix} \tag{7}$$

will satisy the original system of coupled ODEs (in fact, it will be the general solution).

In summary (and in more generality). one can solve a system of coupled ODEs

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1(t) + \cdots + a_{nn}x_n(t)
\end{aligned}$$

by carrying out the following steps:

- Form the coefficient matrix

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

- Find the eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbf{A}$, and use them to form the diagonal matrix $\mathbf{D}$ and the diagonalizing matrix $\mathbf{C}$

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad , \quad \mathbf{C} = \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & & \downarrow \end{pmatrix}$$

- Solve the decoupled system (easy)

$$\frac{d\mathbf{y}}{dt} = \mathbf{D}\mathbf{y}(t) \quad \Rightarrow \quad \mathbf{y}(t) = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix}$$

- Transform the decoupled solutions back to solutions $\mathbf{x}(t)$ of the original system

$$\mathbf{x}(t) = \mathbf{C}\mathbf{y}(t)$$