LECTURE 16

# 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

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

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

EXAMPLE 16.1. Find the eigenvalues and eigenvectors of

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

- The characteristic polynomial is

$$P_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \det \begin{bmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} = (2-\lambda)(3-\lambda)$$

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

$$\begin{aligned} (\mathbf{A} - (2)\mathbf{I})\mathbf{x} &= \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad x_2 = 0 \\ &\Rightarrow \quad \mathbf{x} \in span\left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \end{aligned}$$

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

$$\begin{aligned} (\mathbf{A} - (3)\mathbf{I})\mathbf{x} &= \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad -x_1 = 0 \\ &\Rightarrow \quad \mathbf{x} \in span\left( \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \end{aligned}$$

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 16.2. *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

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

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

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

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 16.3. *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 16.4. *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 16.5. 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

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

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

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

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,

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