5.3 Diagonalization

The goal here is to develop a useful factorization $A = PDP^{-1}$, when A is $n \times n$. We can use this to compute A^k quickly for large k.

The matrix D is a *diagonal* matrix (i.e. entries off the main diagonal are all zeros).

D^k is trivial to compute as the following example illustrates.

EXAMPLE: Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$. Compute D^2 and D^3. In general, what is D^k, where k is a positive integer?

Solution:

$$D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} & 0 \\ 0 & \end{bmatrix}$$

$$D^3 = D^2 D = \begin{bmatrix} 5^2 & 0 \\ 0 & 4^2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} & 0 \\ 0 & \end{bmatrix}$$

and in general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}$$

EXAMPLE: Let $A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$. Find a formula for A^k given that $A = PDP^{-1}$ where $P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$.

Solution:

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDDP^{-1} = PD^2P^{-1}$$

Again,

$$A^3 = A^2A = (PD^2P^{-1})(PDP^{-1}) = PD^2(P^{-1}P)DP^{-1} = PD^3P^{-1}$$

In general,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 4^k & -5^k + 4^k \\ 2 \cdot 5^k - 2 \cdot 4^k & -5^k + 2 \cdot 4^k \end{bmatrix}.$$

A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, i.e. if $A = PDP^{-1}$ where P is invertible and D is a diagonal matrix.

When is A diagonalizable? (The answer lies in examining the eigenvalues and eigenvectors of A.)

Note that

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Altogether

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix}$$

Equivalently,

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \_\_\_ & 0 \\ 0 & \_\_\_ \end{bmatrix}$$

or

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$

In general,

$$A\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

and if $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$ is invertible, A equals

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}^{-1}$$

THEOREM 5 The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

EXAMPLE: Diagonalize the following matrix, if possible.

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

Step 1. Find the eigenvalues of A.

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

Eigenvalues of A: $\lambda = 1$ and $\lambda = 2$.

Step 2. Find three linearly independent eigenvectors of A.

By solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$, for each value of λ, we obtain the following:

Basis for $\lambda = 1$: $\mathbf{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

Basis for $\lambda = 2$: $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

Step 3: Construct P from the vectors in step 2.

$$P = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Step 4: Construct D from the corresponding eigenvalues.

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Step 5: Check your work by verifying that AP = PD

$$\boldsymbol{AP} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\boldsymbol{PD} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

EXAMPLE: Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{bmatrix}.$$

Since this matrix is triangular, the eigenvalues are $\lambda = 2$ and $\lambda = 4$. By solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$ for each eigenvalue, we would find the following:

$$\text{Basis for } \lambda = 2: \qquad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{Basis for } \lambda = 4: \qquad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$$

Every eigenvector of A is a multiple of $\mathbf{v}_1$ or $\mathbf{v}_2$ which means there are not three linearly independent eigenvectors of A and by Theorem 5, A is not diagonalizable.

EXAMPLE: Why is $A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ 3 & 2 & 1 \end{bmatrix}$ diagonalizable?

Solution: Since A has three eigenvalues ($\lambda_1 =$ ____, $\lambda_2 =$ ____, $\lambda_3 =$ ____) and since eigenvectors corresponding to distinct eigenvalues are linearly independent, A has three linearly independent eigenvectors and it is therefore diagonalizable.

THEOREM 6 An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

EXAMPLE: Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 24 & -12 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Solution: Eigenvalues: -2 and 2 (each with multiplicity 2).

Solving $(A - \lambda I)\mathbf{x} = \mathbf{0}$ yields the following eigenspace basis sets.

Basis for $\lambda = -2$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -6 \\ 0 \end{bmatrix}$ $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 0 \end{bmatrix}$

Basis for $\lambda = 2$: $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent

$\Rightarrow P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{bmatrix}$ is invertible

$\Rightarrow A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -6 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

THEOREM 7 Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.

a. For $1 \leq k \leq p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the distinct eigenspaces equals n, and this happens if and only if the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and β_k is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\beta_1, \ldots, \beta_p$ forms an eigenvector basis for $\mathbf{R}^n$.