5.2. Diagonalization

Definition: An $n \times n$ matrix A is called *diagonalizable* if there is a basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of A. For example, the matrices in the two examples of last chapter are diagonalizable.

Why are they called "diagonalizable"?

Let us write the eigenvector equation for each vector:

$$A\vec{v}_1 = \lambda_1\vec{v}_1, \quad A\vec{v}_2 = \lambda_2\vec{v}_2, \quad \ldots, \quad A\vec{v}_n = \lambda_n\vec{v}_n$$

We can write all these equations together by constructing a matrix P whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$:

$$A\left[\ \vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_n\ \right] = \left[\ \lambda_1\vec{v}_1 \mid \lambda_2\vec{v}_2 \mid \ldots \mid \lambda_n\vec{v}_n\ \right]$$

The right-hand side can be written as:

$$\left[\ \lambda_1\vec{v}_1 \mid \lambda_2\vec{v}_2 \mid \ldots \mid \lambda_n\vec{v}_n\ \right] = \left[\ \vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_n\ \right] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

so we have

$$A\left[\ \vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_n\ \right] = \left[\ \vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_n\ \right] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

Let

$$D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

Then the eigenvector equations for all the eigenvectors read:

$$AP = PD$$

Since the columns of P are linearly independent, P is invertible, so we can multiply on the left (or on the right) by P^{-1} to get:

$$P^{-1}AP = D \quad (\text{or } A = PDP^{-1}).$$

This means that: A

After multiplying A by P^{-1} on the left and P on the right, it becomes a diagonal matrix.

Theorem: A is diagonalizable if and only if it can be written as

$$A = PDP^{-1}$$

where the columns of P are a basis of eigenvectors of A, and D is a diagonal matrix whose diagonal entries are the eigenvalues corresponding to the eigenvectors of A, in the same order as the corresponding eigenvectors in P.

Example: Let us find the eigenvalues and eigenvectors of the matrix

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

The characteristic polynomial of A is

$$\chi_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} -1-\lambda & -2 & 6 \\ -2 & -1-\lambda & 6 \\ -2 & -2 & 7-\lambda \end{bmatrix}.$$

Computing this determinant, one finds

$$\chi_A(\lambda) = (\lambda - 1)^2(\lambda - 3) = 0.$$

Thus the eigenvalues are

$$\lambda_1 = 1 \quad \text{(with algebraic multiplicity 2)}, \qquad \lambda_2 = 3.$$

Now for each λ, we find the eigenvectors:

Eigenvalue $\lambda_2 = 3$. We have to solve

$$(A - 3I)\vec{x} = \begin{bmatrix} -4 & -2 & 6 \\ -2 & -4 & 6 \\ -2 & -2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This system gives the condition $x = y = z$. Hence an eigenvector corresponding to the eigenvalue $\lambda = 3$ is

$$\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Eigenvalue $\lambda_1 = 1$. We have to solve

$$(A - 1I)\vec{x} = \begin{bmatrix} -2 & -2 & 6 \\ -2 & -2 & 6 \\ -2 & -2 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

All rows give the single equation

$$-2x - 2y + 6z = 0 \quad \Longleftrightarrow \quad x = -y = 3z.$$

Thus the eigenspace has dimension 2. Choosing convenient solutions:

- If $(y, z) = (1, 0)$, then $x = -1$, giving

$$\vec{v}_1^{(1)} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

- If $(y, z) = (0, 1)$, then $x = 3$, giving

$$\vec{v}_1^{(2)} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

These two vectors span the eigenspace corresponding to $\lambda = 1$. Thus if we write

$$P = \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

then we have that the diagonal form of A is D. You can check that

$$PDP^{-1} = \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & -2 \\ -1 & -1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & -2 & 6 \\ -2 & -1 & 6 \\ -2 & -2 & 7 \end{bmatrix} = A.$$

Similar Matrices

Definition: Two $n \times n$ matrices A, B are called similar if there is a matrix P such that

$$A = PBP^{-1}.$$

(Note that this is the same as saying that $B = P^{-1}A(P^{-1})^{-1}$.)

Similar matrices have the same determinant, rank, nullity, eigenvalues, and the eigenspace dimensions are also the same (the eigenspaces themselves may be different).

Thus, we can say:

$$A \text{ is diagonalizable} \iff \text{it is similar to a diagonal matrix.}$$

How do we know when a matrix is diagonalizable, without actually diagonalizing it? The following theorem gives some help (although it does not give the whole answer).

Theorem: Let A be an $n \times n$ matrix.

(a) If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues, and if $\vec{v}_1, \ldots, \vec{v}_k$ are the corresponding eigenvectors, then $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent.

(b) If A has n distinct eigenvalues, then A is diagonalizable.

Note that not all matrices are diagonalizable. For example

Example: Let us find the eigenvalues and eigenvectors of the matrix

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

The characteristic polynomial of A is

$$\chi_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 2 & -2 \\ 0 & 3-\lambda & -1 \\ 1 & 2 & -2-\lambda \end{bmatrix}.$$

Computing this determinant gives

$$\chi_A(\lambda) = -(\lambda - 2)^2(\lambda + 1) = 0.$$

Thus the eigenvalues are

$$\lambda_1 = 2 \text{ (algebraic multiplicity 2)}, \qquad \lambda_2 = -1 \text{ (algebraic multiplicity 1)}.$$

Now for each λ, we find the eigenvectors:

Eigenvalue $\lambda_1 = 2$. We solve

$$(A - 2I)\vec{x} = \begin{bmatrix} 0 & 2 & -2 \\ 0 & 1 & -1 \\ 1 & 2 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The second row gives $y - z = 0$, so $y = z$. The third row gives $x + 2y - 4z = 0$, which with $y = z$ becomes $x - 2y = 0$, so $x = 2y$. Hence all eigenvectors have the form

$$\begin{bmatrix} 2y \\ y \\ y \end{bmatrix} = y \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \qquad y \neq 0.$$

A convenient choice is

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Eigenvalue $\lambda_2 = -1$. We solve

$$(A + I)\vec{x} = \begin{bmatrix} 3 & 2 & -2 \\ 0 & 4 & -1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

From the second row $4y - z = 0$ so $z = 4y$. From the third row $x + 2y - z = 0$ so $x = 2y$. Thus the eigenvectors are

$$\begin{bmatrix} 2y \\ y \\ 4y \end{bmatrix} = y \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \qquad y \neq 0.$$

A convenient choice is

$$\vec{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}.$$

Note: The eigenvalue $\lambda = 2$ has algebraic multiplicity 2 but its eigenspace is one-dimensional (spanned by $\vec{v}_1$). The full list of eigenpairs (up to scaling) is

$$\lambda = 2: \ \vec{v} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}; \qquad \lambda = -1: \ \vec{v} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}.$$

An application: powers of matrices

Example: Let

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

Compute A^5.

Too long!!

But let's think. . .

$$A = PDP^{-1} \quad \text{(from example above) where}$$

$$P = \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & -2 \\ -1 & -1 & 3 \end{bmatrix}.$$

Now, we have

$$\begin{aligned} A^5 &= A \cdot A \cdot A \cdot A \cdot A \\ &= PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1}PDP^{-1} \\ &= PD^5P^{-1} \end{aligned}$$

Now,

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

and D^5 is easy to compute:

$$D^5 = D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 243 \end{bmatrix}$$

Thus:

$$\begin{aligned} A^5 = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 243 \end{bmatrix} P^{-1} &= \begin{bmatrix} -1 & 3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 243 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & -2 \\ -1 & -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 3 & 243 \\ 1 & 0 & 243 \\ 0 & 1 & 243 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 1 & 1 & -2 \\ -1 & -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} -241 & -242 & 726 \\ -242 & -241 & 726 \\ -242 & -242 & 727 \end{bmatrix}. \end{aligned}$$

Done!

Algebraic and geometric multiplicity

Definition: Let A be an $n \times n$ matrix, and λ an eigenvalue of A.

- The *algebraic multiplicity* of λ is the multiplicity of λ as a root of the characteristic polynomial. That is, if you factor
$$\chi_A(t) = (t - \lambda)^k(\dots)$$
then k is the algebraic multiplicity of λ.
- The *geometric multiplicity* of λ is the dimension of the eigenspace corresponding to λ.

Theorem

1. The algebraic multiplicity of an eigenvalue is always greater than or equal to its geometric multiplicity.
2. The algebraic multiplicity of every eigenvalue is equal to its geometric multiplicity if and only if A is diagonalizable.

For example, above we found that for the matrix

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

the characteristic polynomial was

$$\chi_A(\lambda) = -(\lambda - 2)^2(\lambda + 1)$$

Thus, the algebraic multiplicity of $\lambda = 2$ is 2, and the algebraic multiplicity of $\lambda = -1$ is 1.

On the other hand, the only eigenvector corresponding to $\lambda = 2$ was

$$\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

so the geometric multiplicity of $\lambda = 2$ is 1.

The geometric multiplicity of $\lambda = -1$ is 1, since the only eigenvector was

$$\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

.

Thus, the matrix A is not diagonalizable.

Note also that some matrices do not have real eigenvalues at all.

Example: Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix}$.

The characteristic polynomial is

$$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & -2 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) + 2 = \lambda^2 - 1 + 2 = \lambda^2 + 1.$$

Note that the characteristic equation $\lambda^2 + 1 = 0$ has no real solutions. If we consider that our set of scalars is the complex numbers instead of the real numbers, then we could diagonalize it using complex vectors, but otherwise there is no diagonal form.

However, for matrices like that, and using only real numbers, there are other *standard* forms that are quite useful. We will not see them in this course, however.