Section 5.2 Diagonalization

Objectives.

- Define similarity transformations and identify some properties of similar matrices.
- Introduce the idea of diagonalizing a matrix.
- Use diagonalization to compute powers of a matrix efficiently.

Let A and P be $n \times n$ matrices with P invertible. The transformation that sends A to the matrix product $P^{-1}AP$ is called a similarity transformation.

More generally, if A and B are $n \times n$ matrices then we say that B is similar to A if there is an invertible matrix P such that $B = P^{-1}AP$.

Example 1. Suppose that B is similar to A. Show that A is similar to B.

Because B is similar to A, there is an invertible matrix P such that $B = P^{-1}AP$.

Then $PBP^{-1} = P(P^{-1}AP)P^{-1} = (PP^{-1})A(PP^{-1}) = IAI = A$.

That is $A = Q^{-1}BQ$ where $Q = P^{-1}$. Thus A is similar to B.

(Notice that the previous example allows us to say that A and B are similar if one is similar to the other.)

Similar matrices share several important properties. In particular, if A and B are similar then A and B have the same ...

determinant, rank, nullity, trace, characteristic polynomial, eigenvalues, ...

note: similar matrices represent the same linear transformation with respect to different bases.

Example 2. Suppose that A and B are similar matrices. Show that $\det(A) = \det(B)$.

Because A and B are similar, there is an invertible matrix P such that $B = P^{-1}AP$.

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P)$$
$$= \frac{1}{\det(P)}\det(A)\det(P) = \det(A).$$

An $n \times n$ matrix A is diagonalizable if it is similar to a diagonal matrix. That is, if there is an invertible matrix P such that $P^{-1}AP$ is diagonal, in which case we say that P diagonalizes A.

Example 3. Consider the 2×2 matrices $A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

(a) Show that P diagonalizes A.

$$P^{-1} = \frac{1}{\det(P)} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2-1} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

$$P^{-1}AP = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}.$$

Thus P ~~diagonal~~ diagonalizes A.

(b) What are the eigenvalues of A?

$P^{-1}AP$ has eigenvalues $\lambda = 4, 5$, so A also has eigenvalues $\lambda = 4, 5$.

The key ingredient for diagonalizing a matrix is the set of eigenvectors of the matrix.

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

Theorem. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of a matrix A, and $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are corresponding eigenvectors, then the set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly independent.

It follows from the previous two theorems that an $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Why? each eigenvalue corresponds to (at least) one eigenvector, so n distinct eigenvalues gives us n linearly independent eigenvectors.

Thus these n eigenvectors are a basis for $\mathbb{R}^n$.

Strategy. To find a matrix that diagonalizes A:

- find the eigenvalues and corresponding eigenvectors $\vec{v}_1, \ldots, \vec{v}_k$ of A.
- if you find n eigenvectors, then $P = [\vec{v}_1 \mid \vec{v}_2 \mid \cdots \mid \vec{v}_n]$ diagonalizes A.

Example 4. Find a matrix P that diagonalizes $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$.

From Ex. 6, Section 5.1, the eigenvalues of A are $\lambda = 1$ (with eigenvector $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$) and $\lambda = 2$ (with eigenvectors $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$). These three eigenvectors are linearly independent, so $P = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ diagonalizes A.

eigenvalues of A!!

check: $P^{-1}AP = \begin{bmatrix} -1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

note: if we chose $P = \begin{bmatrix} 0 & -2 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, then $P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Example 5. Show that the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is not diagonalizable.

A is triangular, so the eigenvalues are $\lambda = 1$ (repeated) and $\lambda = 4$.

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda-1 & -1 & 0 \\ 0 & \lambda-1 & -1 \\ 0 & 0 & \lambda-4 \end{bmatrix} = (\lambda-1)^2(\lambda-4). \Rightarrow \lambda = 1, \lambda = 4.$$

$\lambda = 1$: $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = t,\ x_2 = 0,\ x_3 = 0.$

The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ is a basis for this eigenspace.

$\lambda = 4$: $\begin{bmatrix} 3 & -1 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = t,\ x_2 = 3t,\ x_3 = 9t.$

The set $\left\{ \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} \right\}$ is a basis for the eigenspace.

Because A has only two linearly independent eigenvectors, we cannot diagonalize A.

Example 6. Explain why the matrix $A = \begin{bmatrix} 2 & 1 & -3 & 5 \\ 0 & 4 & -1 & 2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ is diagonalizable.

The eigenvalues of A are $\lambda = 1, 2, 4, 5$. These are distinct, so A has four linearly independent eigenvectors. Therefore A is diagonalizable.

One application of diagonalization is finding powers of a matrix. Recall that if D is a diagonal matrix, then D^k can be found by raising each diagonal entry to the power k.

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \Rightarrow D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix} \quad (\text{for } k > 0).$$

Suppose that A is similar to a diagonal matrix D, so that $A = P^{-1}DP$ where P is invertible. Then:

i.e. A is diagonalizable.

$$A^k = (P^{-1}DP)^k = (P^{-1}DP)(P^{-1}DP)\cdots(P^{-1}DP)$$
$$= P^{-1}D(PP^{-1})D(PP^{-1})D\cdots DP = P^{-1}D^kP.$$

Example 7. Compute A^5 for the matrix $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$ in Example 4.

$A = PDP^{-1}$ where $P = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Thus: $A^5 = (PDP^{-1})^5 = PD^5P^{-1} = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 32 & 0 \\ 0 & 0 & 32 \end{bmatrix} \begin{bmatrix} -1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

$$= \cdots = \begin{bmatrix} -30 & 0 & -62 \\ 31 & 32 & 31 \\ 31 & 0 & 63 \end{bmatrix}.$$