§5.3 Diagonalization

Recall a square matrix is <u>diagonal</u> if all of its non-diagonal entries are zero.

These matrices are "nice" since they are easy to work with computationally. They have the best echelon forms (lots of zeroes) and powers of them are easy to compute.

<u>Example</u>

Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$. Compute D^2 and D^3. Come up with a guess for D^{100}

<u>Solution</u>

$$D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

$$D^3 = D^2 \cdot D = \begin{bmatrix} 25 & 0 \\ 0 & 9 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 125 & 0 \\ 0 & 27 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

Guess: $D^{100} = \begin{bmatrix} 5^{100} & 0 \\ 0 & 3^{100} \end{bmatrix}$

As it turns out, this happens for all diagonal matrices!

Theorem

Let $D = \begin{bmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{bmatrix}$ be a ~~diag~~ diagonal $n \times n$ matrix.

Then for any integer $k \geq 1$

$$D^k = \begin{bmatrix} a_1^k & & & 0 \\ & a_2^k & & \\ & & \ddots & \\ 0 & & & a_n^k \end{bmatrix}$$

Ideally we'd like nice formulas like this for any non-diagonal square matrix A. However, such formulas only exist for matrices "close" to diagonal matrices.

Defn

A square matrix A is diagonalizable if it is similar to a diagonal matrix D. In other words, there's an invertible matrix P such that

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

Diagonalizable matrices have a similar property to compute their powers.

Theorem

Let A be a square diagonalizable matrix and write $A = PDP^{-1}$ with D diagonal and P invertible. Then for any integer $k \geq 1$

$$A^k = PD^kP^{-1}$$

(Recall D^k is easy to compute!)

Proof

$A = PDP^{-1}$ so

$$A^k = \underbrace{(PDP^{-1})\cdot(PDP^{-1})\cdot \cdots\cdots (PDP^{-1})}_{k \text{ times}}$$

$$= PD\underbrace{(P^{-1}P)}_{I}D\underbrace{(P^{-1}P)}_{I}D\underbrace{(P^{-1}\cdot)}_{I}\cdots\underbrace{(P)}_{I}DP^{-1}$$

$$= PD^kP^{-1}$$

Pros

- Cool formula
- Easy to compute

Cons

- How do we know if A is diagonalizable
- Even if we know A is diagonalizable, how do we find D and P?

Theorem (Diagonalization Theorem)

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

In this case $A = PDP^{-1}$ where the columns of P are the n linearly independent eigenvectors of A. Also the diagonal entries of D are the eigenvalues corresponding to the ~~eigenvectors~~ columns of P (in the same order!).

Example

From the §5.1 lesson we know the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$ are $\lambda_1 = 3$ and $\lambda_2 = -7$ with corresponding eigenvectors $v_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

Clearly v_1 and v_2 are linearly independent so A is diagonalizable.

Moreover,

$$A = \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & -3 \end{bmatrix}^{-1}$$ Check this!

you can swith the orders of things as long as you do it everywhere. For instance

$$A = \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} -7 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}^{-1}$$ check this!

Here we switched the columns of P AND the diagonal entries of D.

<u>Theorem</u>

If A is an $n \times n$ matrix and $v_1, \ldots, v_r$ are eigenvectors corresponding to <u>distinct</u> eigenvalues $\lambda_1, \ldots, \lambda_r$ then $\{v_1, \ldots, v_r\}$ is linearly independent.

<u>Corollary</u>

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

<u>Warning!</u>

This is not an equivalence! For example I_n is clearly diagonalizable, but only has one eigenvalue!

Defn

The process of expressing A as $A = PDP^{-1}$ with D diagonal is called diagonalization of A.

Theorem

Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_p$.

a) The dimension of the eigenspace of each λ_i ($\mathrm{Null}(A - \lambda_i I)$) is at most the multiplicity of λ_i.

b) Let $E_1, \ldots, E_p$ denote the eigenspaces. A is diagonalizable if and only if

$$\dim E_1 + \dim E_2 + \cdots + \dim E_p = n$$

Notice that (b) is actually easy to check when n is small (2, 3, 4 like most of our examples).

Again

$$\begin{aligned}\dim E_i &= \dim \mathrm{Null}(A - \lambda_i I) \\ &= \#\text{ non-pivots of } A - \lambda_i I\end{aligned}$$

usually quick to get into echelon form.