Session 20
April 22, 2020

4.5 Diagonalization

Let us come back to our original motivation. If possible, we would like to express a given matrix in a different basis, such that it has a very simple form, namely diagonal.

In order to do that for an $n\times n$ matrix A, we need that A has n linearly independent eigenvectors. E.g., last time we proved that we always have that if all n eigenvalues are distinct (but this is not a necessary condition, e.g., $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ has only one eigenvalue 1, but, e.g., $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are two linearly independent eigenvectors).

So let us now assume that the $n\times n$ matrix A has eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding linearly independent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$.

Let us build a matrix with the eigenvectors as columns: $V = \left(\vec{v}_1 \middle| \vec{v}_2 \middle| \cdots \middle| \vec{v}_n\right)$.

Then we find
$$AV = \left(A\vec{v}_1 \middle| A\vec{v}_2 \middle| \cdots \middle| A\vec{v}_n\right)$$
$$= \left(\lambda_1\vec{v}_1 \middle| \lambda_2\vec{v}_2 \middle| \cdots \middle| \lambda_n\vec{v}_n\right)$$
$$= \left(\vec{v}_1 \middle| \vec{v}_2 \middle| \cdots \middle| \vec{v}_n\right) \underbrace{\begin{pmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ & & \ddots & \\ 0 & & & \lambda_n\end{pmatrix}}_{=:\Lambda}$$

$=:\Lambda$ = matrix with the eigenvalues as diagonal and zeroes otherwise

$$= V\Lambda$$

(in components: $(AV)_{ij} = \sum_{k=1}^{n} A_{ik} V_{kj}$
$= \sum_{k=1}^{n} A_{ik} (\vec{v}_j)_k$
$= (A\vec{v}_j)_i = \lambda_j (\vec{v}_j)_i$)

And we constructed V with linearly independent columns, so V has full rank $= n$, and it is thus invertible ($\exists V^{-1}$ s.t. $V^{-1}V = VV^{-1} = I :=$ identity).

Therefore, we could write $AV = V\Lambda$ as $\underbrace{\boxed{V^{-1}AV = \Lambda}}$.

This is called diagonalization of A.
We say V diagonalizes A.

Every matrix A for which this works is called diagonalizable, i.e.:

Definition: A matrix A is called diagonalizable if there exists an invertible matrix V such that $V^{-1}AV$ is diagonal.

Note that the matrix V (with full rank) can be regarded as a change of basis.

Let us summarize our findings:

Theorem: An $n\times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$. In this case $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = V^{-1}AV$ and $V = (\vec{v}_1 | \dots | \vec{v}_n)$.

Equivalent formulation: A is diagonalizable if and only if the sum of the geometric multiplicities is n (because: if the sum is not $= n$ there aren't n lin. indep. eigenvectors; if the sum is n then we just choose some basis in each $E_\lambda(A)$). Since the sum of the algebraic multiplicities is by definition always n we have:

A is diagonalizable if and only if for all eigenvalues the algebraic equals the geometric multiplicity.

And we already know: If all eigenvalues of A are distinct, then A is diagonalizable.

(but not necessarily the other way around)

Examples (that we already discussed before):

- $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has one eigenvalue $\lambda = 1$ (alg. mult. $= 2$) with $E_\lambda(A) = \mathbb{R}^2$, i.e., geom. mult. $= 2$. A is already diagonal, so it is of course diagonalizable

- $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has two distinct eigenvalues: $\lambda_\pm = \pm 1$. So it is diagonalizable, and we already computed that $V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. One easily computes that $V^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ as well.

  Check: $V^{-1} A V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{= \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}}$. ✓

- $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{pmatrix}$ has 3 distinct eigenvalues, so it is diagonalizable

- $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has one eigenvalue $\lambda = 1$, but its geometric multiplicity is just 1. So A does not have n linearly independent eigenvectors, so it is not diagonalizable.

Diagonalization has numerous applications in the sciences, as you will see in your major-related classes.

One simple application is taking powers of a diagonalizable matrix A:

$\hookrightarrow$ Get $V^{-1} A V = \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, then $A = V \Lambda V^{-1}$.

Then $A^k = (V\Lambda V^{-1})^k = V\Lambda \underbrace{V^{-1} V}_{=I}\Lambda V^{-1} \dots V\Lambda V^{-1} = V\Lambda^k V^{-1}$, where

$$\Lambda^k = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}.$$

So $A^k = V \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} V^{-1}$ i.e., instead of multiplying a matrix k times (computationally very costly if k or n are large), we only need to multiply three matrices to compute A^k!

So for a diagonalizable matrix A we could even define $e^A = \exp(A)$ by its Taylor series:

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{(V\Lambda V^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{V\Lambda^k V^{-1}}{k!} = V\left(\sum_{k=0}^{\infty} \frac{\Lambda^k}{k!}\right)V^{-1}$$

$$= V\left(\sum_{k=0}^{\infty} \begin{pmatrix} \frac{\lambda_1^k}{k!} & & & \\ & \frac{\lambda_2^k}{k!} & & \\ & & \ddots & \\ & & & \frac{\lambda_n^k}{k!} \end{pmatrix}\right)V^{-1} = V \begin{pmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{pmatrix} V^{-1}$$