5.3. Diagonalization of square matrices

- Diagonali matrices are "the simplest" ones: they behave as n numbers acting in n directions.
- Diagonalization of A means A is presented as PDP^{-1} for some diagonal matrix D.
- Matrix A can be diagonalized iff there is a basis consisting of A eigenvectors $v_1, \dots, v_n$. Then $P = [v_1 \dots v_n]$ the matrix whose columns are the eigenvectors $v_1, \dots, v_n$. Also,

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$ where $Av_i = \lambda_i v_i$.

Procedure:
- Find all eigenvalues.
- For each eigenvalue λ find a basis B_λ of the eigenspace V_λ.
- If the size of B_λ is the multiplicity m_λ of λ, then the union of all B_λ for all λ is a basis B of eigenvectors

§5.3 Diagonalization:

D1

A. A diagonalization of a matrix A is a matrix D which is diagonal and an invertible matrix P, such that $\boxed{A = PDP^{-1}}$.

{ P-conjugate of D
A, D similar.

A1. { Motivation: calculation of powers }

example

$A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$

claim:

$\boxed{A = PDP^{-1}}$ for $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ & $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$

Q. $\boxed{A^{1000} = \underbrace{A \cdot A \cdots A}_{1000}} = ?$

n times

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

$$= P \cdot \overbrace{D \cdots D}^{n} \cdot P^{-1} = P(D^n)P^{-1}.$$

$$D^n = D^{1000} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}^{1000} = \begin{bmatrix} 5^{1000} & \\ & 3^{1000} \end{bmatrix}$$

$n = 1000$

Rule

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix} = \begin{bmatrix} ap & 0 \\ 0 & bq \end{bmatrix}$$

diagonal entries get multiplied!

$$A^{1000} = \underbrace{\begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}}_{P} \cdot \underbrace{\begin{bmatrix} 5^{1000} & \\ & 3^{1000} \end{bmatrix}}_{D^{1000}} \cdot \underbrace{\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}}_{P^{-1}}$$

$\Rightarrow$ $\boxed{\text{Calculation of powers!}}$ easy!

A2. Powers of matrices represent repeated processes:

A matrix A describes a lin. transform $T(x) = Ax$

In applications: T is a description of some process. Then $T^n = \underbrace{T \cdots T\,T\,T}_{n}$ describes the process repeated n times. Here $T^n(x) = A^n x$.

Of interest is how this {iteration of a process} behaves for large n.

A3. Ex. Dynamical systems.

Consider an ecological system in woods:

x = # squirrels,
y = # wolves;

$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} = \{$ the same but one year later $\}$

Then $T^n\begin{pmatrix} x \\ y \end{pmatrix}$ = #s (population) after n years.

- population after a long time:
- stable cute: does not change much
- catastrophy: {die out, explode} >

A4. Conclusion: • A diagonalization of a matrix A means presenting A as PDP^{-1} for some matrix P and a diagonal matrix D.

• This is useful because calculations with D are much simpler!

B. Which matrices A can be diagonalized?

Th. (a) A can be diagonalized if $\mathbb{R}^n$ has a basis $\mathcal{B}=\{b_1,\dots,b_n\}$ of eigenvectors of A, i.e., $Ab_i=\lambda_i b_i$ for some λ_i.

(b) Then we can choose D as $D=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$

and $P=P_{\mathcal{B}}=[b_1 \cdots b_n]$ columns in $\mathcal{B}$.

Proof. $\boxed{Ab_i=\lambda_i b_i}$ and also

Claim (*): $\boxed{PDP^{-1}b_i \overset{(*)}{=} \lambda_i b_i}$. (as we will see!).

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

(*) $\left\{\begin{array}{l} A \text{ and } B=PDP^{-1} \text{ act the same on some basis } \mathcal{B}, \\ \text{they must be the same matrices!} \end{array}\right.$

i.e. $A=B$ $(=PDP^{-1})$

[Pf of (*): • Matrices A, B act the same on all linear combinations of vectors b_i

• These are all vectors! (since $\mathcal{B}$ is a basis)

• In particular $Ae_i=Be_i$.

• However, these are the i^{th} column A_i and B_i of A & B.

• So, $A_i=B_i$ hence $A=B$!]

So, it suffices to check that
$$PDP^{-1}\,b_i = \lambda_i\, b_i.$$

However, Pe_i = i^{th} column of $P = b_i$.

hence $P^{-1}b_i = e_i$!

Therefore:
$$PDP^{-1}b_i = PDe_i$$

and since De_i = i^{th} column of D = $\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i$ $= \lambda_i e_i$,

For a diagonal D, e_i's are eigenvectors!

we have:
$$PDP^{-1}b_i = PDe_i = P\lambda_i e_i =$$
$$= \lambda_i Pe_i = \lambda_i b_i. \qquad \square$$

C. For which A is there a basis of eigenvectors?

T2. If an $n \times n$ matrix A has n different eigenvalues $\lambda_1, \ldots, \lambda_n$ then A can be diagonalized! (Because it has a basis of eigenvectors!)

Pf. Let v_i be an eigenvector for A, with eigenvalue λ_i, so $Av_i = \lambda_i v_i$.

Then $v_1, \ldots, v_n$ are independent!!!

But then this is a basis (since we have n independent vectors in $\mathbb{R}^n$!)

Now we know diagonalization by T1. $\square$

D) Example: We will find a basis of eigenvectors & diagonalize.

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

We want to **Diagonalize!**

Find a basis of eigenvectors

Step 1. Calculate

$\det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4$

Factor it into linear terms:

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

So, eigenvalues are $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = -2$.

Step 2. Find eigenvectors for eigenvalues:

$v_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $v_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

for $\lambda_1 = 1$, for $\lambda_2 = -2$, for $\lambda_3 = -2$.

Remark: Here, the three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ were not all different but we were lucky and we still found 3 independent eigenvectors!

then it is a basis of $\mathbb{R}^3$.

Step 3. $P = [v_1\ v_2\ v_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

use T1:

$$D = \begin{bmatrix} 1 & & \\ & -2 & \\ & & 2 \end{bmatrix}$$ eigenvalues of A

and

$A = PDP^{-1}$. by T1. ! $\square$

E2. $A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$

$\det(A - \lambda I) = -(\lambda - 1)(\lambda + 2)^2$

- V_{λ} for $\lambda = \lambda_1 = 1$ has basis $v_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$
- V_{λ} for $\lambda = \lambda_2 = \lambda_3 = -2$ has basis $v_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

There are only two independent eigenvectors. So, no eigenvector basis.

So: A can not be diagonalized! $\square$

E3. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, eigenvalues:

(triangular!) $\lambda_1 = 1$, $\lambda_2 = 1$.

Eigenvectors $v = \begin{bmatrix} x \\ y \end{bmatrix}$ satisfies $Av = 1 \cdot v$

i.e. $Av = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$

$v = \begin{bmatrix} x \\ y \end{bmatrix}$

So: $Av = v$ means: $\begin{bmatrix} x+y \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

i.e. $x = 0$,

So: $v = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} = y \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\mathbb{R}^2 \supseteq V_1 = \{v \in V ; Av = v\} = \mathbb{R} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

So, no basis of eigenvectors.

So, A does not diagonalize. $\square$

E. Criterion for diagonalizability:

T3. Let A be $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$ which appear with multiplicities $m_1, m_2, \dots, m_p$
{ So, $m_1 + m_2 + \dots + m_p = n$ } $n = \deg[\det(A - \lambda I)]$.

(a) $\dim\{V_{\lambda_i}\} \leq m_i$

dimension of λ-eigenspace is $\leq$ multiplicity of λ.

(b) A is diagonalizable iff all these inequalities are equalities.
[For each eigenvalue λ any basis of V_λ has the number of elements the same as multiplicity of λ.]

(c) Then, we choose bases $\mathcal{B}_1$ for V_{λ_1}, $\mathcal{B}_2$ for V_2 etc, and together they form a basis $\mathcal{B}$ of $\mathbb{R}^n$: $\boxed{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_p}$.

This is a basis consisting of eigenvectors of A so:

$$A = PDP^{-1}, \quad A_{\mathcal{B}} = P, \quad D = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_1 & & \\ & & & \ddots & \\ & & & & \lambda_p \\ & & & & & \ddots \\ & & & & & & \lambda_p \end{bmatrix}$$

m_1 times

m_p times