Lecture 12

Diagonalization

A square matrix A is called <u>diagonalizable</u> if there exists P s.t. $P^{-1}AP$ is diagonal.

Suppose

$$P^{-1}AP = D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

where $P = [P_1, P_2, \dots P_n]$. Then $AP = PD$:

$$A[P_1 P_2 \dots P_n] = [P_1 \dots P_n] \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} =$$

$$= [\lambda_1 P_1 \;\; \lambda_2 P_2 \;\; \dots \;\; \lambda_n P_n]$$

In other words;

$AP_1 = \lambda_1 P_1$, $AP_2 = \lambda_2 P_2$, ..., $AP_n = \lambda_n P_n$

with lin. independent $\{P_1, P_2, \dots P_n\} \subseteq \mathbb{R}^{n\times 1}$.

(Df) Let R be a square matrix.
$\lambda \in \mathbb{R}$ is called an **eigenvalue** of R if the equation $RX = \lambda X$ has a nonzero solution.
A vector X is called an **eigenvector** of R if $X \neq 0$ and there exists $\lambda \in \mathbb{R}$ s.t. $R \cdot X = \lambda X$.

Theorem An $n \times n$-matrix R is diagonalizable if it possesses n linearly independent eigenvectors (= basis for $M^{n\times 1}$).

Proof If R is diagonalizable then

$$P^{-1} R P = D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \text{ and } R P_k = \lambda_k P_k,$$

where $P = [P_1 \; P_2 \; \dots \; P_n]$.
Then $\{P_1, P_2, \dots, P_n\}$ is a basis for $M^{n\times 1}$.
If $B = \{v_1, v_2, \dots, v_n\}$ is a basis consisting of eigenvectors, i.e. $Rv_1 = \lambda_1 v_1$, $Rv_2 = \lambda_2 v_2$, $\dots$, $Rv_n = \lambda_n v_n$.
Then $[R v_k]_B = \begin{bmatrix} 0 \\ \vdots \\ \lambda_k \\ \vdots \\ 0 \end{bmatrix}$, and hence the coordinate matrix A_R of R relative to bases (B, B) is

$$A_R = \Big[[Rv_1]_B \, [Rv_2]_B \dots [Rv_n]_B \Big] = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

λ is an eigenvalue iff
$(\lambda I - A)X = 0$ has a nonzero solution iff
$(\lambda I - A)$ is not invertible iff
$\det(\lambda I - A) = 0.$

Consider the polynomial

$$\Delta_A(t) = \det(tI - A) \in \mathcal{P}_n(\mathbb{R}).$$

$\Delta_A(t)$ is called the **characteristic polynomial** of A, and the equation

$$\Delta_A(t) = 0$$

is called the **characteristic equation** of A.

Theorem $\lambda \in \mathbb{R}$ is an eigenvalue of A iff λ is a root of its characteristic equation.

Example Let $A = \begin{bmatrix} -11 & -5 & -3 \\ 12 & 7 & 2 \\ 12 & 5 & 4 \end{bmatrix}$.

Find P and $\mathcal{D}$ such that $P^{-1}AP = \mathcal{D}$ is diagonal if the characteristic polynomial $\det(\lambda I - A) = (\lambda - 1)(\lambda - 2)(\lambda + 3)$ is given.

Solution

$$\Delta_A(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda + 3) = 0.$$

The roots are $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = -3$.
Thus there exists a basis of eigenvectors.
Let us find this basis

$$\lambda_1 I - A = \begin{bmatrix} 12 & 5 & 3 \\ -12 & -6 & -2 \\ -12 & -5 & -3 \end{bmatrix} \xrightarrow[R_1+R_3]{R_1+R_2} \begin{bmatrix} 12 & 5 & 3 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} 12x + 5y + 3z &= 0 \\ -y + z &= 0 \\ 0 \cdot z &= 0 \end{aligned} \Rightarrow \quad y = z, \quad 12x = -8z \Rightarrow x = -\frac{2}{3}z$$

$$\begin{bmatrix} -\frac{2}{3}z \\ z \\ z \end{bmatrix} = z \cdot \begin{bmatrix} -\frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$ is the general solution.

The fundamental solution is $\begin{bmatrix} -\frac{2}{3} \\ 1 \\ 1 \end{bmatrix} \in M^{3\times 1}$

Thus $P_1 = \begin{bmatrix} -\frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$ satisfies $A \cdot P_1 = \lambda_1 \cdot P_1$ and is an eigenvector.

$$\lambda_2 I - A = \begin{bmatrix} 13 & 5 & 3 \\ -12 & -5 & -2 \\ -12 & -5 & -2 \end{bmatrix} \xrightarrow[R_2+R_3]{R_2+R_1} \begin{bmatrix} 1 & 0 & 1 \\ -12 & -5 & -2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{12R_1+R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -5 & 10 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{(-\frac{1}{5})R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x + z &= 0 \\ y - 2z &= 0 \end{aligned} \qquad \begin{aligned} x &= -z \\ y &= 2z \end{aligned} \qquad P_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

$$\lambda_3 I - A = \begin{bmatrix} 8 & 5 & 3 \\ -12 & -10 & -2 \\ -12 & -5 & -7 \end{bmatrix} \xrightarrow[R_1+R_3]{R_1+R_2} \begin{bmatrix} 8 & 5 & 3 \\ -4 & -5 & 1 \\ -4 & 0 & -4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} -4 & -5 & 1 \\ 8 & 5 & 3 \\ -4 & 0 & -4 \end{bmatrix}$$

$$\xrightarrow{R_1+R_2} \begin{bmatrix} -4 & -5 & 1 \\ 4 & 0 & 4 \\ -4 & 0 & -4 \end{bmatrix} \xrightarrow[R_2+R_1]{R_2+R_3} \begin{bmatrix} 0 & -5 & 5 \\ 4 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow[\frac{1}{4}\cdot R_2]{\frac{1}{5}\cdot R_1} \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} -y + z = 0 \\ x + z = 0 \\ 0 \cdot z = 0 \end{cases} \Rightarrow \quad \begin{aligned} y &= z \\ x &= -z \end{aligned} \qquad P_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Thus $P = \big[[P_1] \, [P_2] \, [P_3] \big] = \begin{bmatrix} -\frac{2}{3} & -1 & -1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Example $A = \begin{bmatrix} 5 & 0 & 4 \\ -6 & 1 & -1 \\ -6 & 0 & -5 \end{bmatrix}$

The characteristic equation $\Delta_A(\lambda) = |\lambda I - A| = 0$ gives us

$$\begin{vmatrix} \lambda-5 & 0 & -4 \\ 6 & \lambda-1 & 1 \\ 6 & 0 & \lambda+5 \end{vmatrix} = 0.$$

$$\begin{vmatrix} \lambda-5 & 0 & -4 \\ 6 & \lambda-1 & 1 \\ 6 & 0 & \lambda+5 \end{vmatrix} = (\lambda-1)(-1)^{2+2} \begin{vmatrix} \lambda-5 & -4 \\ 6 & \lambda+5 \end{vmatrix} = (\lambda-1)\left[(\lambda-5)(\lambda+5)+24\right] =$$

$$= (\lambda-1)(\lambda^2-25+24) = (\lambda-1)(\lambda^2-1) = (\lambda-1)^2(\lambda+1).$$

The roots of $\Delta_A = 0$ are $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = -1$.

$\lambda_1 = \lambda_2 = 1$:

$$\lambda_1 I - A = I - A = \begin{bmatrix} -4 & 0 & -4 \\ 6 & 0 & 1 \\ 6 & 0 & 6 \end{bmatrix} \xrightarrow[\frac{1}{6}\cdot R_3]{\left(-\frac{1}{4}\right)R_1} \begin{bmatrix} 1 & 0 & 1 \\ 6 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow[-R_1+R_3]{(-6)R_1+R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\left.\begin{aligned} x + z &= 0 \\ -5z &= 0 \\ y &= y. \end{aligned}\right\} \Rightarrow x = z = 0. \qquad X = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} = y \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

The fundamental solution is $P_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = P_2$.

$$\lambda_3 I - A = -I - A = \begin{bmatrix} -6 & 0 & -4 \\ 6 & -2 & 1 \\ 6 & 0 & 4 \end{bmatrix} \xrightarrow[R_1+R_3]{R_1+R_2} \begin{bmatrix} -6 & 0 & -4 \\ 0 & -2 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} -6x - 4z &= 0 \\ -2y - 3z &= 0 \\ z &= z \end{aligned} \qquad \begin{aligned} x &= -\tfrac{2}{3}z \\ y &= -\tfrac{3}{2}z \end{aligned} \qquad P_3 = \begin{bmatrix} -\frac{2}{3} \\ -\frac{3}{2} \\ 1 \end{bmatrix}$$

The set $\{P_1, P_2, P_3\} = \{P_1, P_3\}$ is <u>not</u> a basis for $M^{3\times 1}$ since it consists of $2 < 3$ vectors!

Therefore the matrix A is <u>not</u> diagonalizable.

Exercises Find the characteristic polynomial, the eigenvalues, and the associated eigenvectors for each of the following matrices.

a) $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$.

$$|\lambda I - A| = \begin{vmatrix} \lambda-1 & -1 \\ 2 & \lambda-4 \end{vmatrix} = (\lambda-1)(\lambda-4)+2 = \lambda^2-5\lambda+6 = 0$$

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{5 \pm \sqrt{25-24}}{2} = \frac{5 \pm 1}{2}$$

$\lambda_1 = 3$, $\lambda_2 = 2$

$\lambda_1 = 3$. $3I - A = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \xrightarrow{R_1+R_2} \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$ $\begin{cases} 2x - y = 0 \\ y = y \end{cases}$

$X = \begin{bmatrix} \frac{1}{2}y \\ y \end{bmatrix} = y\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$; $P_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$

$\lambda_2 = 2$. $2I - A = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \xrightarrow{(-2)R_1+R_2} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ $\begin{cases} x - y = 0 \\ y = y \end{cases}$

$X = y\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $P_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$P = [[P_1][P_2]]$; $P = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$.

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

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

$$|\lambda I - A| = (\lambda-2)\begin{vmatrix} \lambda-2 & 3 \\ 2 & \lambda-1 \end{vmatrix} - \begin{vmatrix} 2 & 3 \\ 2 & \lambda-1 \end{vmatrix} - \begin{vmatrix} 2 & \lambda-2 \\ 2 & 2 \end{vmatrix} =$$

$$= (\lambda-2)((\lambda-2)(\lambda-1)-6) - (2(\lambda-1)-6) - (4-(\lambda-2)\cdot 2) =$$

$$= (\lambda-2)(\lambda^2-3\lambda-4) - (2\lambda-8) - (8-2\lambda) = (\lambda-2)(\lambda^2-3\lambda-4) =$$

$$= (\lambda-2)(\lambda-4)(\lambda+1).$$

$\lambda^2-3\lambda-4=0$ $\lambda = \frac{3 \pm \sqrt{9+16}}{2} = \frac{3 \pm \sqrt{25}}{2} = \frac{3 \pm 5}{2}$

$\lambda_1 = 2 \quad , \quad \lambda_2 = 4, \quad \lambda_3 = -1$

$$\lambda_1 I - A = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 0 & 3 \\ 2 & 2 & 1 \end{bmatrix} \xrightarrow{-R_2 + R_3} \begin{bmatrix} 0 & 1 & -1 \\ 2 & 0 & 3 \\ 0 & 2 & -2 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{(-2)R_1 + R_3} \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 0 \\ z &= z \end{aligned} \qquad \begin{aligned} x &= -\tfrac{3}{2} z \\ y &= z \\ z &= z \end{aligned} \qquad X = \begin{bmatrix} -\frac{3}{2} \\ 1 \\ 1 \end{bmatrix} z \qquad P_1 = \begin{bmatrix} -\frac{3}{2} \\ 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 I - A = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 2 & 3 \\ 2 & 2 & 3 \end{bmatrix} \xrightarrow{-R_2 + R_3} \begin{bmatrix} 2 & 1 & -1 \\ 2 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-R_1 + R_2} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} 2x + y - z = 0 \\ y + 4z = 0 \\ z = z \end{cases} \qquad X = \begin{bmatrix} \frac{5}{2} \\ -4 \\ 1 \end{bmatrix} z \qquad P_2 = \begin{bmatrix} \frac{5}{2} \\ -4 \\ 1 \end{bmatrix}$$

$$\lambda_3 I - A = \begin{bmatrix} -3 & 1 & -1 \\ 2 & -3 & 3 \\ 2 & 2 & -2 \end{bmatrix} \xrightarrow[\frac{1}{2} \cdot R_2 \;\; \frac{1}{2} \cdot R_3]{\frac{1}{3} \cdot R_1} \begin{bmatrix} -1 & \frac{1}{3} & -\frac{1}{3} \\ 1 & -\frac{3}{2} & \frac{3}{2} \\ 1 & 1 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -\frac{3}{2} & \frac{3}{2} \\ -1 & \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

$$\xrightarrow[R_1 + R_3]{-R_1 + R_2} \begin{bmatrix} 1 & 1 & -1 \\ 0 & -\frac{5}{2} & \frac{5}{2} \\ 0 & \frac{4}{3} & -\frac{4}{3} \end{bmatrix} \xrightarrow[\frac{3}{4} \cdot R_3]{\frac{2}{5} \cdot R_2} \begin{bmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \xrightarrow[R_2 + R_3]{R_2 + R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} x = 0 \\ -y + z = 0 \\ z = z \end{cases} \qquad X = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} z \qquad P_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$P = \Big[[P_1] \; [P_2] \; [P_3] \Big] = \begin{bmatrix} -\frac{3}{2} & \frac{5}{2} & 0 \\ 1 & -4 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Theorem (Cayley - Hamilton)

Every square matrix satisfies its characteristic equation.

Corollary If $A \in M^{n\times n}$ and $a_0 + a_1\lambda + \dots + a_n\lambda^n$ is the characteristic polynomial of A s.t. $a_0 \neq 0$ then A is invertible and

$$A^{-1} = -\frac{a_1}{a_0} I - \frac{a_2}{a_0} A - \dots - \frac{a_n}{a_0} A^{n-1}$$

Proof

$$a_0 I + a_1 A + \dots + a_n A^n = 0$$

$$I = -\frac{a_1}{a_0} A - \frac{a_2}{a_0} A^2 - \dots - \frac{a_n}{a_0} A^n =$$

$$= \underbrace{\left(-\frac{a_1}{a_0} I - \frac{a_2}{a_0} A - \dots - \frac{a_n}{a_0} A^{n-1}\right)}_{A^{-1}} A$$

Exercise Use the Cayley-Hamilton theorem to compute the inverse of A in terms of powers A, where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$.

Solution $[\lambda I - A] = \begin{bmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & +1 & \lambda \end{bmatrix}$

$$|\lambda I - A| = (\lambda - 1)(-1)^{1+1}\begin{vmatrix} \lambda & -1 \\ 1 & \lambda \end{vmatrix} + (-1)\cdot(-1)^{1+2}\begin{vmatrix} -1 & -1 \\ -1 & \lambda \end{vmatrix} + (-1)\cdot(-1)^{1+3}\begin{vmatrix} -1 & \lambda \\ -1 & 1 \end{vmatrix} =$$

$$= (\lambda - 1)(\lambda^2 + 1) + (-\lambda - 1) - (-1 + \lambda) = \lambda^3 - \lambda^2 + \lambda - 1 - 2\lambda = \lambda^3 - \lambda^2 - \lambda - 1$$

Therefore, by Cayley - Hamilton theorem

$$A^3 - A^2 - A - I = 0 \quad \text{and} \quad I = A(A^2 - A - I).$$

Hence $A^{-1} = A^2 - A - I$

Exercise

Find all 2×2-matrices X satisfying:

$$X^2 + X = I \quad \text{and} \quad |X| \neq \pm 1.$$

Solution By Cayley-Hamilton theorem, X a root of its characteristic polynomial. But $X^2+X-I=0$, and hence the characteristic polinomial is $p(x) = x^2+x-1$.

Roots of $p(x)=0$ are; $x_{1,2} = \frac{-1\pm\sqrt{1+4}}{2} = \frac{-1\pm\sqrt{5}}{2}$.

Matrix solutions are:

a) $$R^{-1}\begin{bmatrix} \frac{-1\pm\sqrt{5}}{2} & 0 \\ 0 & \frac{-1\pm\sqrt{5}}{2} \end{bmatrix} R = \begin{bmatrix} \frac{-1\pm\sqrt{5}}{2} & 0 \\ 0 & \frac{-1\pm\sqrt{5}}{2} \end{bmatrix}$$

and

b) $$S^{-1}\begin{bmatrix} \frac{-1+\sqrt{5}}{2} & 0 \\ 0 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} S,$$

for invertible matrices R and S.

The solution b) is impossible since its determinant

$$|S^{-1}(\ldots)S| = \begin{vmatrix} \frac{-1+\sqrt{5}}{2} & 0 \\ 0 & \frac{-1-\sqrt{5}}{2} \end{vmatrix} = \frac{(-1)^2-5}{2^2} = -1.$$

Answer: $X_{1,2} = \frac{-1\pm\sqrt{5}}{2}\cdot\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.