Review

- Eigenvectors & Eigenvalues. A $n \times n$

$$A\underline{v} = \lambda \underline{v}, \quad \underline{v} \neq \underline{0}.$$

- Eigenspace $E_\lambda = \text{Null}(A - \lambda I) = $ soln set to $(A - \lambda I)\underline{x} = \underline{0}$.

- Characteristic Polynomial

$$\text{CP of } A = \det(A - tI_n) = (-1)^n t^n + \cdots$$

- λ is an eigenvalue of A iff λ is a root of the CP.

- multiplicity of λ.

- Thm 5.1 $\quad 1 \leq \dim E_\lambda \leq$ multiplicity of λ.

Sec. 5.3. Diagonalization of Matrices.

Defn a matrix A is said to be diagonalizable (diagz) if there is an invertible matrix P and a diagonal matrix D such that

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

- A is diagz iff A is similar to a diagonal matrix D.

- If A is diagz, then $A = PDP^{-1}$

Compute A^k, k: positive integer

$$A^2 = A \cdot A = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1}.$$

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

$$D^k = \begin{bmatrix} d_{11}^k & & \text{O} \\ & d_{22}^k & \\ \text{O} & \ddots & d_{nn}^k \end{bmatrix}$$

- If A is invertible, $A^{-1} = PD^{-1}P^{-1}$.

- Find a matrix B such that $B^3 = A$.

Take $B = PD^{1/3}P^{-1}$, $D^{1/3} = \begin{bmatrix} d_{11}^{1/3} & & \text{O} \\ & d_{22}^{1/3} & \\ \text{O} & \ddots & d_{nn}^{1/3} \end{bmatrix}$.

Are all matrices diagz? No.

e.g. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

If A were diagz, then $A = PDP^{-1}$.

$A^2 = PD^2P^{-1} = \mathcal{O} \Rightarrow D^2 = \mathcal{O} \Rightarrow d_{11}^2 = d_{22}^2 = 0 \Rightarrow D = \mathcal{O}$

$\Rightarrow A = PDP^{-1} = \mathcal{O}$ a contradiction!

Thm 5.2. An $n \times n$ matrix A is diagz iff there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

Moreover $A = PDP^{-1}$ iff the cols of P are eigenvectors of A and the diagonal entries of D are eigenvalues of A.

$$P = [\underline{p}_1 \ \underline{p}_2 \ \cdots \ \underline{p}_n] \quad \text{and} \quad D = \begin{bmatrix} d_{11} & & \mathcal{O} \\ & d_{22} & \\ \mathcal{O} & & \ddots & \\ & & & d_{nn} \end{bmatrix}.$$

<u>Pf</u>: Notice that

$$[\underline{p}_1 \ \underline{p}_2 \ \cdots \ \underline{p}_n] \cdot \begin{bmatrix} d_{11} & & \mathcal{O} \\ & d_{22} & \\ \mathcal{O} & & \ddots \ d_{nn} \end{bmatrix} = P\left[d_{11}\underline{e}_1 \ \ d_{22}\underline{e}_2 \ \cdots \ d_{nn}\underline{e}_n\right] = \left[d_{11}\underline{p}_1 \ \ d_{22}\underline{p}_2 \ \cdots \ d_{nn}\underline{p}_n\right] \quad (1)$$

Suppose A is diagz. i.e. $A = PDP^{-1}$.

$\Leftrightarrow AP = PD$. using (1)

$$[A\underline{p}_1 \ A\underline{p}_2 \ \cdots \ A\underline{p}_n] = [d_{11}\underline{p}_1 \ \ d_{22}\underline{p}_2 \ \cdots \ d_{nn}\underline{p}_n]$$

$$A\underline{p}_1 = d_{11}\underline{p}_1, \quad A\underline{p}_2 = d_{22}\underline{p}_2, \quad \ldots \quad A\underline{p}_n = d_{nn}\underline{p}_n$$

$\therefore$ P is invertible, $\underline{P_i} \neq \underline{0}$.

$\{\underline{P_1}, \underline{P_2}, \cdots, \underline{P_n}\}$ forms a basis for $\mathbb{R}^n$.

and $\underline{P_i}$ are eigenvectors of A.

Suppose $\{\underline{P_1}, \underline{P_2}, \cdots, \underline{P_n}\}$ forms a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

$$A\underline{P_1} = d_{11}\underline{P_1}, \quad \cdots \quad A\underline{P_n} = d_{nn}\underline{P_n}.$$

$$[A\underline{P_1} \; A\underline{P_2} \; \cdots \; A\underline{P_n}] = [d_{11}\underline{P_1} \; d_{22}\underline{P_2} \; \cdots \; d_{nn}\underline{P_n}].$$

$$AP = PD$$ using (1).

$\therefore$ $\{\underline{P_1}, \cdots, \underline{P_n}\}$ forms a basis for $\mathbb{R}^n$, P is invertible

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

$\#$

Ex. $A = \begin{bmatrix} 0.85 & 0.03 \\ 0.15 & 0.97 \end{bmatrix}$ $\quad CP = \det(A - tI_2) = (t-0.82)(t-1)$.

$\lambda_1 = 0.82 \qquad \lambda_2 = 1$

$m_1 = 1 \qquad m_2 = 1$.

$\underline{E_{\lambda_1}}$ $(A - \lambda_1 I)\underline{X} = \underline{0}$ $\cdots\cdots$ $\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$ is a basis for E_{λ_1}.

$\underline{E_{\lambda_2}}$ $(A - \lambda_2 I)\underline{x} = \underline{0}$ $\cdots$ $\left\{ \begin{bmatrix} 1 \\ 5 \end{bmatrix} \right\}$ is a basis for E_{λ_2}.

$\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \end{bmatrix} \right\}$ forms a basis for $\mathbb{R}^2$.

$\Rightarrow$ A is diagz.

$P = \begin{bmatrix} -1 & 1 \\ 1 & 5 \end{bmatrix}$, $\quad D = \begin{bmatrix} 0.82 & 0 \\ 0 & 1 \end{bmatrix}$

$P^{-1} = \frac{1}{6} \begin{bmatrix} -5 & 1 \\ 1 & 1 \end{bmatrix}$. $\quad$ verify $PDP^{-1} = \begin{bmatrix} 0.85 & 0.03 \\ 0.15 & 0.97 \end{bmatrix} = A$.

$$\lim_{k \to \infty} A^k = \lim_{k \to \infty} P D^k P^{-1} = P \left(\lim_{k \to \infty} D^k \right) P^{-1}.$$

$$= P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \frac{1}{6} \begin{bmatrix} 1 & 1 \\ 5 & 5 \end{bmatrix}.$$

Thm 5.3. a set of eigenvectors corresponding to distinct eigenvalues are l. indep.

$\lambda_1, \lambda_2, \ldots, \lambda_k$ distinct eigenvalues of A. $\quad \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k\}$ is

$\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k$ corresponding eigenvectors. $\quad$ l. indep.

Pf. Suppose $\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_k}\}$ is l. dep.

$\underline{v_1} \neq \underline{0}$.

From Thm 1.9, we know that some $\underline{v_i}$ is a l.c. of its preceding vectors.

Let $\underline{v_l}$ be the first vector that can be written as a l.c. of its preceding vectors.

i.e. $\underline{v_l} = c_1 \underline{v_1} + c_2 \underline{v_2} + \cdots + c_{l-1} \underline{v_{l-1}}$ —— (2)

and $\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_{l-1}}\}$ is l. indep ($\because$ no vectors is a l.c. of its preceding vectors)

Multiplying (2) by A, we get.

$$\lambda_l \underline{v_l} = c_1 \lambda_1 \underline{v_1} + c_2 \lambda_2 \underline{v_2} + \cdots + c_{l-1} \lambda_{l-1} \underline{v_{l-1}} \text{ —— } (3)$$

(3) $- \lambda_l \cdot$ (2), we have

$$\underline{0} = c_1(\lambda_1 - \lambda_l)\underline{v_1} + c_2(\lambda_2 - \lambda_l)\underline{v_2} + \cdots + c_{l-1}(\lambda_{l-1} - \lambda_l)\underline{v_{l-1}}$$

$c_1(\lambda_1 - \lambda_l) = 0,\ c_2(\lambda_2 - \lambda_l) = 0,\ \cdots,\ c_{l-1}(\lambda_{l-1} - \lambda_l) = 0$ ($\because \underline{v_1} \cdots v_{l-1}$ are l. indep)

Because λ_i are distinct, $c_1 = 0,\ c_2 = 0,\ \cdots,\ c_{l-1} = 0$.

From (2), we have $\underline{v_l} = \underline{0}$, a contradiction! $\because \underline{v_l}$ is an eigenvector. #

- Using Thm 5.3, we can conclude that if $A_{(n\times n)}$ has n distinct eigenvalues, then A is diagz.

  $\because$ if A has n distinct eigenvalues, we can find a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

$\lambda_1, \lambda_2, \cdots, \lambda_k$ distinct eigenvalues of A.

$E_{\lambda_1}, E_{\lambda_2}, \cdots, E_{\lambda_k}$.

$\mathcal{L}_1 \quad \mathcal{L}_2 \quad \mathcal{L}_k$ l. indep subsets of E_{λ_i}

Is $\mathcal{L}_1 \cup \mathcal{L}_2 \cup \cdots \mathcal{L}_k$ l. indep? Yes.

Thm 5.3a (Extension of Thm 5.3).

Let $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues of A.

Let $\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_k$ be l. indep. subsets of $E_{\lambda_1}, E_{\lambda_2}, \cdots, E_{\lambda_k}$ respectively.

Then $\mathcal{L}_1 \cup \mathcal{L}_2 \cdots \cup \mathcal{L}_k$ is l. indep.

Defn. An $n \times n$ matrix is called a Vandermonde matrix if it has the form

$$A = \begin{bmatrix} 1 & a_1 & a_1^2 & a_1^3 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & a_2^3 & & a_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & & a_n^{n-1} \end{bmatrix} \quad n \times n.$$

- A is invertible iff $a_1, a_2, \dots, a_n$ are distinct.

(see P75 of Sec. 3.2, 3×3 Vandermonde).

Proof of Thm 5.3a.

Let $\lambda_1, \lambda_2, \lambda_3$ be distinct eigenvalues of A.

Let $\mathcal{L}_1 = \{\underline{u_1}, \underline{u_2}, \dots, \underline{u_j}\}$ be a l. indep. subset of E_{λ_1}.

$\mathcal{L}_2 = \{\underline{v_1}, \underline{v_2}, \dots, \underline{v_\ell}\}$ - - - - - - - - E_{λ_2}

$\mathcal{L}_3 = \{\underline{w_1}, \underline{w_2}, \dots, \underline{w_m}\}$ - - - - - - - - E_{λ_3}.

Suppose $\underbrace{(a_1\underline{u_1} + a_2\underline{u_2} + \dots + a_j\underline{u_j})}_{\underline{U}} + \underbrace{(b_1\underline{v_1} + b_2\underline{v_2} + \dots + b_\ell\underline{v_\ell})}_{\underline{V}} + \underbrace{(c_1\underline{w_1} + \dots + c_m\underline{w_m})}_{\underline{W}} = \underline{0}$ — ①

Multiplying the above by A, we get

$$\lambda_1 \underline{u} + \lambda_2 \underline{v} + \lambda_3 \underline{w} = \underline{0}. \quad \text{—} (2)$$

Multiplying by A again, we get.

$$\lambda_1^2 \underline{u} + \lambda_2^2 \underline{v} + \lambda_3^2 \underline{w} = \underline{0}. \quad \text{—} (3).$$

$$\begin{bmatrix} \underline{u} & \underline{v} & \underline{w} \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix} = \begin{bmatrix} \underline{0} & \underline{0} & \underline{0} \end{bmatrix}.$$

invertible ($\because$ λ_i are distinct)

$$\Rightarrow \begin{bmatrix} \underline{u} & \underline{v} & \underline{w} \end{bmatrix} = \begin{bmatrix} \underline{0} & \underline{0} & \underline{0} \end{bmatrix}.$$

$$\left.\begin{aligned} \underline{u} &= a_1 \underline{u_1} + a_2 \underline{u_2} + \cdots + a_j \underline{u_j} = \underline{0} \\ \underline{v} &= b_1 \underline{v_1} + b_2 \underline{v_2} + \cdots + b_l \underline{v_l} = \underline{0} \\ \underline{w} &= c_1 \underline{w_1} + \cdots + c_m \underline{w_m} = \underline{0} \end{aligned}\right\} \text{ all the coefficients } a\text{'s, } b\text{'s, and } c\text{'s are all zeros.}$$

$$\Rightarrow \mathcal{L}_1 \cup \mathcal{L}_2 \cup \mathcal{L}_3 \text{ is } l.\text{ indep}$$

$\#$.

If A is diagz, how to find P and D?

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

- Compute CP of A and find all eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$.
- find a basis for each eigenspace E_{λ_i}.
- Collect all the vectors in all these bases, and form

$$P = [\underline{p}_1 \ \underline{p}_2 \ \cdots \ \underline{p}_n], \quad D = \begin{bmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{bmatrix}$$ corresponding eigenvalues

---

Ex. $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ CP of $A = \det(A - tI_3) = -(t-3)(t+1)^2$

$\lambda_1 = 3 \qquad \lambda_2 = -1$

$m_1 = 1 \qquad m_2 = 2$.

$\underline{E_{\lambda_1}}$ $(A - \lambda_1 I)\underline{x} = \underline{0}$ .... $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for E_{λ_1}.

$\underline{E_{\lambda_2}}$ $(A - \lambda_2 I)\underline{x} = \underline{0}$ .... $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$ - - .... E_{λ_2}.

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

- P and D are not unique. e.g. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$ is another basis for E_{λ_2}

$$P' = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Test for Diagz Matrices.

Let the eigenvalues of A be $\lambda_1, \lambda_2, \cdots, \lambda_k$ (distinct) with multiplicities $m_1, m_2, \cdots, m_k$ respectively.

$n \times n$. A is diagz iff it satisfies.

① $\overbrace{m_1 + m_2 + \cdots + m_k}^{\text{total number of eigenvalues}} = n$.

② $\dim E_{\lambda_i} = m_i$ for each eigenspace E_{λ_i}.

---

recall from Thm 5.1, $1 \leq \dim E_{\lambda_i} \leq m_i$.

Then $\dim E_{\lambda_i} = \underline{1}$ for those $m_i = 1$.

- only need to verify ② for those $m_i > \underline{1}$.

Ex. $\underset{3\times 3}{A} = \begin{bmatrix} 0 & 2 & 1 \\ -2 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix}$. CP of $A = -(t+1)(t^2+4)$

$\lambda_1 = -1$

number of eigenvalues $= m_1 = \underline{1} \neq 3$ $\quad m_1 = \underline{1}$.

A is not diagz.

Ex. $M = \begin{bmatrix} -3 & 2 & 1 \\ 3 & -4 & -3 \\ -8 & 8 & 6 \end{bmatrix}$ CP of $M = -(t+1)(t^2-4)$

$= -(t+1)(t+2)(t-2)$,

M has 3 distinct eigenvalues

$\Rightarrow$ M is diagz.

Ex. $B = \begin{bmatrix} -7 & -3 & -6 \\ 0 & -4 & 0 \\ 3 & 3 & 2 \end{bmatrix}$ $CP = -(t+1)(t+4)^2$.

$\lambda_1 = -1, \quad \lambda_2 = -4.$

$m_1 = 1, \quad m_2 = 2.$

① $m_1 + m_2 = 3$ ✓

② $\dim E_{\lambda_2} = 2$? ✓

$\underline{E_{\lambda_2}}$ $[A - \lambda_2 I] \xrightarrow{\cdots}$ rref. rank $= \underline{1}$, nullity $= 2$.

$\Downarrow$

B is diagz. $\quad \dim E_{\lambda_2} = 2 = m_2$

Ex. $C = \begin{bmatrix} -6 & -3 & 1 \\ 5 & 2 & -1 \\ 2 & 3 & 5 \end{bmatrix}$. $CP = -(t+1)(t+4)^2$.

$\lambda_1 = -1$ $\lambda_2 = -4$

(1) $m_1 + m_2 = 3$ ✓ $m_1 = 1$ $m_2 = 2$.

(2) $\dim E_{\lambda_2} = m_2$? X

$[C - \lambda_2 I] \rightsquigarrow$ rref. rank = 2 nullity = 1.

$\dim E_{\lambda_2} = \underline{1} \neq m_2$. C is not diagz.

Claim If A and B are similar, then A is diagz iff B is diagz.

$A = P^{-1} B P$.

Explanation of the test.

- notice all eigenvectors of A are in $E_{\lambda_1}, E_{\lambda_2}, \cdots, E_{\lambda_k}$.
  let $l_i = \dim E_{\lambda_i}$.

- There are at most l_i l. indep eigenvectors in E_{λ_i}.
- There are at most $l_1 + l_2 + \cdots + l_k$ l. indep eigenvectors in all eigenspaces.
- In order for A to be diagz, we need to have n l. indep eigenvectors.

$$\underbrace{l_1 + l_2 + \cdots + l_k}_{\substack{\text{at most this many} \\ \text{l. indep eigenvectors}}} \le m_1 + m_2 + \cdots + m_k \le n.$$

$(l_i \le m_i$ Thm 5.1$)$.

We conclude that $l_1 + \cdots + l_k = n$ iff
① $m_1 + \cdots + m_k = n$
② $l_i = m_i$ for each i.

Sec. 5.3. T&F, 9, 10, 20, 56, 60, 61, 70, 80~89.

Review T&F, 25, 32, 39~42.

Cayley - Hamilton Theorem. A: $n \times n$.

Let CP of $A = f(t) = (-1)^n t^n + p_{n-1} t^{n-1} + \cdots + p_1 t + p_0 t^0$

Then $f(A) = (-1)^n A^n + p_{n-1} A^{n-1} + \cdots + p_1 A + p_0 I = \Theta$ for all A.

- In Sec. 5.3, P88 prove that the theorem is true if A is diagz.

P89, If A is diagz, (not necessarily diagz) and let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be eigenvalues of A

then ① $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$.

② $\operatorname{trace} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

$\operatorname{trace} A = a_{11} + a_{22} + \cdots + a_{nn}$.

Sec. 5.5* Applications

differential Equations, Markov process.

Google Search.