Review

For $n \times n$ matrices A and C,

C is similar to A if $\exists$ invertible P s.t. $C = P^{-1}AP$

<Similarity invariants>

(a) determinant (b) rank (c) nullity (d) trace

(e) characteristic polynomial; eigenvalues with algebraic multiplicities

Let $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Then $\det(A) = \det(C) = 0$, $\text{rank}(A) = \text{rank}(C) = 2$, $tr(A) = tr(C) = 0$

$\det(\lambda I - A) = \det(\lambda I - C) = \lambda^4$.

However, A and C are not similar since A^2 and C^2 are not similar.

※ If A and C are similar, then A^n and C^n are similar for $n \in \mathbb{N}$.

Thm 8.2.4 Similar matrices have the same eigenvalues and those eigenvalues have the same algebraic and gemetric multiplicities for both matrices.

pf) Assume that $C = P^{-1}AP$, P is invertible.

Then $\det(\lambda I - A) = \det(\lambda I - C)$. Hence A and C have the same eigenvalues and those eigenvalues have the same algebraic multiplicities.

$\Rightarrow$ geom. multip. of λ_0 for C

$= \dim(\text{sol. sp. of } (\lambda_0 I - C)x = 0)$

$= \text{nullity}(\lambda_0 I - C) = \text{nullity}(P^{-1}(\lambda_0 I - A)P)$

$= \text{nullity}(\lambda_0 I - A) = $ geom. multip. of λ_0 for A. ∎

Thm 8.2.5 Assume $C = P^{-1}AP$ and λ is an eigenvalue of A and C.

(a) If $\mathbf{x}$ is an eigenvector of C corr. to λ, then $P\mathbf{x}$ is an eigenvector of A corr. to λ.

(b) If $\mathbf{x}$ is an eigenvector of A corr. to λ, then $P^{-1}\mathbf{x}$ is an eigenvector of C corr. to λ.

pf) By def. $\mathbf{x} \neq \mathbf{0}$ and $C\mathbf{x} = \lambda\mathbf{x}$.

Then $P^{-1}AP\mathbf{x} = \lambda\mathbf{x} \Rightarrow AP\mathbf{x} = P(\lambda\mathbf{x}) = \lambda P\mathbf{x}$

Since P is invertible, $P\mathbf{x} \neq \mathbf{0}$. This proves (a).

Since $A = (P^{-1})^{-1} C (P^{-1})$, (b) follows from (a). ∎

< Diagonalizability >

Def. If a matrix A is similar to some diagonal matrix, we say that A is diagonalizable. If $P^{-1}AP$ is diagonal, we say that P diagonalize A.

Question When is a square matrix A diagonalizable?

Thm 8.2.6 An $n \times n$ matrix A is diagonalizable iff A has n linearly independent eigenvectors.

pf) ($\Rightarrow$) Set $P = [\mathbf{p}_1 \ \mathbf{p}_2 \cdots \mathbf{p}_n]$ and $D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$.

If $P^{-1}AP = D$, then

$$AP = [A\mathbf{p}_1 \ A\mathbf{p}_2 \ \cdots \ A\mathbf{p}_n]$$
$$\parallel$$
$$PD = [\lambda_1\mathbf{p}_1 \ \lambda_2\mathbf{p}_2 \ \cdots \ \lambda_n\mathbf{p}_n]$$

$\therefore \ A\mathbf{p}_i = \lambda_i \mathbf{p}_i, \quad i = 1, \cdots, n$

($\Leftarrow$) Assume that A has n linearly indep. eigenvectors $\mathbb{p}_1, \cdots, \mathbb{p}_n$ and that the corr. eigenvalues are $\lambda_1, \cdots, \lambda_n$.

Then $A\mathbb{p}_i = \lambda_i \mathbb{p}_i$. Letting $P = [\mathbb{p}_1 \cdots \mathbb{p}_n]$, we can see that

$$AP = PD, \quad D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad P: \text{invertible}.$$

$$\therefore \quad D = P^{-1}AP.$$

$\blacksquare$

Method for diagonalizing a matrix

Step 1 Find n lin. inedp. eigenvectors of A, say $\mathbb{p}_1, \cdots, \mathbb{p}_n$

Step 2 Form the matrix $P = [\mathbb{p}_1 \cdots \mathbb{p}_n]$

Step 3 The matrix $P^{-1}AP$ will be diagonal;

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad \text{with} \quad A\mathbb{p}_i = \lambda_i \mathbb{p}_i$$

Thm 8.2.7 If $\mathbb{v}_1, \cdots, \mathbb{v}_k$ are eigenvectors corr. to distict eigenvalues $\lambda_1, \cdots, \lambda_k$, then $\{\mathbb{v}_1, \cdots, \mathbb{v}_k\}$ is linearly independent.

pf) Suppose that $\mathbb{v}_1, \cdots, \mathbb{v}_k$ are linearly dependent.

Then $\exists\, \mathbb{v}_{r+1}$ s.t. $\mathbb{v}_{r+1} = c_1\mathbb{v}_1 + \cdots + c_r\mathbb{v}_r$ for some $r < k$,

where $\mathbb{v}_1, \cdots, \mathbb{v}_r$ are linearly indep.

$$A\mathbb{v}_{r+1} = c_1 A\mathbb{v}_1 + \cdots + c_r A\mathbb{v}_r = (c_1\lambda_1\mathbb{v}_1 + \cdots + c_r\lambda_r\mathbb{v}_r)$$

(not all $c_1, \cdots, c_r$ are zero)

$$\begin{aligned} &\| \\ &\lambda_{r+1}\mathbb{v}_{r+1} \\ &\| \\ &\lambda_{r+1}(c_1\mathbb{v}_1 + \cdots + c_r\mathbb{v}_r) \end{aligned}$$

$$\therefore \quad c_1(\underbrace{\lambda_1 - \lambda_{r+1}}_{\neq 0})\mathbb{v}_1 + c_2(\underbrace{\lambda_2 - \lambda_{r+1}}_{\neq 0})\mathbb{v}_2 + \cdots + c_r(\underbrace{\lambda_r - \lambda_{r+1}}_{\neq 0})\mathbb{v}_r = 0$$

This is a contradiction that $\mathbb{v}_1, \cdots, \mathbb{v}_r$ are linearly indep.

$\blacksquare$

Thm 8.2.8 An $n \times n$ matrix with n distinct "real" eigenvalues is diagonalizable.

Rmk The converse is false. $\left(\text{e.g. } A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \right)$

Thm 8.2.9

An $n \times n$ matrix A is diagonalizable iff the sum of the geometric multiplicities of its eigenvalues is n.

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

Let $E_1, \cdots, E_k$ denote the eigenspaces corr. to $\lambda_1, \cdots, \lambda_k$, resp.

Let $B_1, \cdots, B_k$ be bases for these eigenspaces, resp.

Now let $B = B_1 \cup B_2 \cup \cdots \cup B_k$.

i) B is linearly independent.

Set $B_i = \{ w_1^i, \cdots, w_{m_i}^i \}$ for $i = 1, \cdots, k$.

If $\sum_{i=1}^{k} \sum_{j=1}^{m_i} a_j^i w_j^i = 0$ for $a_j^i \in \mathbb{R}$, then $a_j^i = 0$ since

$\sum_{j=1}^{m_i} a_j^i w_j^i \in E_i$ for each i.

ii) A is diagonalizable $\Leftrightarrow m_1 + \cdots + m_k = n$.

$\because$ A : diagonalizable $\Leftrightarrow \exists$ a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

$\Leftrightarrow$ B is a basis for $\mathbb{R}^n$

$\Leftrightarrow$ B has n linearly independent eigenvectors

Proof is outlined in Exercise 6.

$\Leftrightarrow m_1 + \cdots + m_k = n$. ▩

Thm 8.2.10 A: a square matrix

(a) geometric multiplicity of an eigenvalue $\leq$ its algebraic multiplicity

(b) A is diagonalizable $\Leftrightarrow$ geometric multiplicity of each eigenvalue $=$ its algebraic multiplicity.

§8.3. Orthogonal diagonalizability ; Functions of a matrix

Def. Let A and C be $n \times n$ matrices.

- C is orthogonally similar to A if $C = P^T A P$, P: orthogonal.
- A is orthogonally diagonalizable if $P^T A P$ is a diagonal matrix, where P is an orthogonal matrix.

Thm 8.3.2 A and C are orthogonally similar iff there exist orthonormal bases B and B' s.t. $A = [T]_B$ and $C = [T]_{B'}$ for some linear operator T.

(Proof is similar to the proof of Theorem 8.2.2.)

Thm 8.3.3 + 8.3.4 (a)

Let A be an $n \times n$ matrix. The following are equivalent.

(1) A is orthogonally diagonalizable.

(2) There is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

(3) A is symmetric.

Thm 8.3.4 (b) If A is a symmetric matrix, then eigenvectors from different eigenvalues are orthogonal.

pf) Assume $A w_1 = \lambda_1 w_1$ & $A w_2 = \lambda_2 w_2$, $\lambda_1 \neq \lambda_2$.

Then
$$\lambda_1 (w_1 \cdot w_2) = \lambda_1 w_1^T w_2 = (A w_1)^T w_2 = w_1^T A^T w_2$$
$$= w_1^T A w_2 = w_1^T (\lambda_2 w_2) = \lambda_2 (w_1 \cdot w_2)$$

Since $\lambda_1 \neq \lambda_2$, $w_1 \cdot w_2 = 0$.

(proof)

<u>(1) ⇒ (2)</u> Let $P = [\mathbb{p}_1\ \mathbb{p}_2 \cdots \mathbb{p}_n]$ and $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ s.t. $D = P^T A P$.

Then $PD = AP = \begin{bmatrix} | & | & & | \\ A\mathbb{p}_1 & A\mathbb{p}_2 & \cdots & A\mathbb{p}_n \\ | & | & & | \end{bmatrix}$. Hence $A\mathbb{p}_i = \lambda_i \mathbb{p}_i \quad \forall i = 1, \cdots, n.$

$$PD = [\lambda_1\mathbb{p}_1\ \lambda_2\mathbb{p}_2 \cdots \lambda_n\mathbb{p}_n]$$

Since P is orthogonal, $\{\mathbb{p}_1, \cdots, \mathbb{p}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

<u>(2) ⇒ (1)</u> If $B = \{\mathbb{p}_1, \mathbb{p}_2, \cdots, \mathbb{p}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A s.t. $A\mathbb{p}_i = \lambda_i \mathbb{p}_i$ for $i = 1, \cdots, n$, then for $P = [\mathbb{p}_1 \cdots \mathbb{p}_n]$ & $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, P is orthogonal and $P^T A P = D$.

<u>(1) ⇒ (3)</u> $P^T A P = D \iff A = PDP^T$

$\therefore A^T = (PDP^T)^T = PD^TP^T = PDP^T = A.$

<u>(3) ⇒ (2)</u> Let $\lambda_1, \cdots, \lambda_k$ be the distinct eigenvalues of A.

Let $\{\mathbb{p}_1^i, \cdots, \mathbb{p}_{a_i}^i\}$ be an basis for the eigenspace corr. to λ_i $(i = 1, \cdots, n)$.

Using Gram-Schmidt process, we can find an orthonormal basis $\{\mathbb{q}_1^i, \cdots, \mathbb{q}_{a_i}^i\}$ for the eigenspace corr. to λ_i. Then $\bigcup_{i=1}^{k} \{\mathbb{q}_1^i, \cdots, \mathbb{q}_{a_i}^i\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A by Thm 8.3.4 (2). ◪

This is the way to orthogonally diagonalizing a symmetric matrix.