5.3 Similarity and Diagonalization

- Diagonalization Revisited

Thm. [A] $A : n \times n$ *matrix.*
A *is diagonalizable if and only if it has eigenvectors* $X_1, X_2, \cdots, X_n$ *s.t.* $P = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$ *is invertible.* *In this case,* $P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, *where* λ_i *is the eigenvalue of* A *corresponding to* X_i.

Thm. [A'] $A : n \times n$ *matrix.*
A *is diagonalizable if and only if* $\mathbf{F}^n$ *has a basis* $\{X_1, X_2, \cdots, X_n\}$ *of eigenvectors of* A.

Thm. [B] *Let $X_1, X_2, \cdots, X_k$ be eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$ of A. Then $\{X_1, X_2, \cdots, X_k\}$ is linearly independent.*

Proof. Assume that $\{X_1, X_2, \cdots, X_k\}$ is linearly dependent. We can find j s.t. $\{X_1, X_2, \cdots, X_{j-1}\}$ is linearly independent, and $\{X_1, X_2, \cdots, X_j\}$ is linearly dependent. Then we have

$$(*) \qquad a_1X_1 + a_2X_2 + \cdots + a_jX_j = O,$$

where not all a_i's are zero, and in particular $a_j \neq 0$. Multiplying $(*)$ by A from the left, we get

$$a_1\lambda_1X_1 + a_2\lambda_2X_2 + \cdots + a_j\lambda_jX_j = O.$$

On the other hand, multiplying $(*)$ by λ_j, we obtain

$$a_1\lambda_j X_1 + a_2\lambda_j X_2 + \cdots + a_j\lambda_j X_j = O.$$

Subtracting two equations, we have

$$a_1(\lambda_1 - \lambda_j)X_1 + a_2(\lambda_2 - \lambda_j)X_2 + \cdots + a_{j-1}(\lambda_{j-1} - \lambda_j)X_{j-1} = O,$$

and $a_1(\lambda_1 - \lambda_j) = a_2(\lambda_2 - \lambda_j) = \cdots = a_{j-1}(\lambda_{j-1} - \lambda_j) = 0$.
Since λ_i's are distinct, we have

$$a_1 = a_2 = \cdots = a_{j-1} = 0,$$

and $a_jX_j = 0$, $a_j = 0$, a contradiction.
Therefore, $\{X_1, X_2, \cdots, X_k\}$ is linearly independent. $\square$

Cor. [B'] *If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.*

Fact. *If one chooses linearly independent sets of eigenvectors corresponding to distinct eigenvalues, and combines them into a single set, then that combined set will be linearly independent.*

Def. *An eigenvalue λ of A is said to have multiplicity m if it occurs m times as a root of $c_A(x)$.*

Def. *The set*

$$E_\lambda(A) = \{X \in \mathbf{F}^n | AX = \lambda X\}$$

of λ-eigenvectors is a subspace of $\mathbf{F}^n$ called the eigenspace of A corresponding to λ.

Note that an eigenspace $E_\lambda(A)$ is merely the null space of $\lambda I - A$.

Thm. [C] *A : $n \times n$ matrix.*
A is diagonalizable if and only if $\dim E_\lambda(A)$ is equal to the multiplicity of λ for every eigenvalue λ of A.

Proof. $(\Rightarrow)$ We omit it.
$(\Leftarrow)$ Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues. Assume that $\dim E_{\lambda_i}(A)$ is equal to the multiplicity of λ_i for each $i = 1, 2, \cdots, k$. Choose a basis B_i of $E_{\lambda_i}(A)$ for each λ_i. Let $B = B_1 \cup B_2 \cup \cdots \cup B_k$. Then $|B| = n$ and B is linearly independent from **Fact**. Thus B is a basis of $\mathbf{F}^n$, and A is diagonalizable by **Thm A'**. $\square$

Thm. [C'] A : $n \times n$ *matrix.*
A *is diagonalizable if and only if every eigenvalue* λ *of multiplicity* m *yields* m *basic solutions of the equation*

$$(\lambda I - A)X = O.$$

Fact. *Let* λ *be an eigenvalue of multiplicity of* m *of* A. *Then*

$$\dim E_\lambda(A) \le m.$$

- Diagonalization Algorithm

  Let A be an $n \times n$ matrix.

  1. Find all the eigenvalues λ of A.
  2. For each λ, compute the basic solutions of $(\lambda I - A)X = O$. If there are n basic solutions in total, A is diagonalizable.
  3. Construct the matrix P whose columns are (scalar multiples of) basic solutions.
  4. $P^{-1}AP$ is diagonal. (P is invertible.)

Eg. $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$, $c_A(x) = x(x-1)^2$.

For $\lambda = 1$,

$$\lambda I - A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

A *is not diagonalizable.*

$B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ *is diagonalizable.*

- Similar Matrices

Def. A, B: $n \times n$ *matrices*
We say that A *and* B *are similar if* $B = P^{-1}AP$ *for some invertible* P. *We will write* $A \sim B$ *for similar matrices* A *and* B.

Eg. $\begin{bmatrix} 1 & -2 \\ -3 & 2 \end{bmatrix}$ *and* $\begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$ *are similar.*
Indeed, for $P = \begin{bmatrix} \frac{2}{3} & 1 \\ -1 & 1 \end{bmatrix}$, *we have*

$$P^{-1} \begin{bmatrix} 1 & -2 \\ -3 & 2 \end{bmatrix} P = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}.$$

Observations :

1. A is diagonalizable if and only if A is similar to a diagonal matrix.
2. Assume that A and B are similar. Then $A^{-1} \sim B^{-1}$, $A^T \sim B^T$, $A^k \sim B^k$. If one of A and B is diagonalizable, then the other is also diagonalizable.
3. If A is diagonalizable, then A^{-1}, A^T and A^k are also diagonalizable.

Def. *Let* $A = [a_{ij}]$. *The trace of an* $n \times n$ *matrix* A *is defined by*

$$\mathrm{tr}A = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Prop.

1. $\operatorname{tr}(A+B) = \operatorname{tr}A + \operatorname{tr}B$,
2. $\operatorname{tr}(kA) = k\operatorname{tr}A$,
3. $\operatorname{tr}(A^T) = \operatorname{tr}A$,
4. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Proof. $A = [a_{ij}]$, $B = [b_{ij}]$, $AB = [c_{ij}]$, and $BA = [d_{ij}]$.

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \sum_{k=1}^{n} d_{kk} = \operatorname{tr}(BA). \end{aligned}$$

□

Thm. *If $A \sim B$, then A and B have the same determinant, rank, trace, characteristic polynomial, and eigenvalues.*

Proof. Let $B = P^{-1}AP$ for some invertible P.

$$\det B = \det(P^{-1}AP) = \det P^{-1} \det A \det P = \det A.$$

$$\mathrm{tr}B = \mathrm{tr}(P^{-1}AP) = \mathrm{tr}((AP)P^{-1}) = \mathrm{tr}A.$$

$$\begin{aligned} c_B(x) &= \det(xI - B) = \det(P^{-1}xIP - P^{-1}AP) \\ &= \det[P^{-1}(xI - A)P] = \det(xI - A) = c_A(x). \end{aligned}$$

$$\mathrm{rank}B = \mathrm{rank}(P^{-1}AP) = \mathrm{rank}(AP) = \mathrm{rank}A.$$

□