11/26/2019. Lecture 24

Diagonalization: $n \times n$ A $\lambda_1, \dots, \lambda_n$ eigenvalues
$v_1, \dots, v_n$ corresponding eigenvectors
linearly independent.
eigenbasis.

$$Av_j = \lambda_j v_j.$$

$$S = [v_1 | \cdots | v_n] \quad \text{invertible.}$$

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Then, $A = SDS^{-1}$.

Definition: A is diagonalizable if such S and D exist.

Note: If A is diagonalizable with $A = SDS^{-1}$, then necessarily the columns of S are eigenvectors of A and the elements on the diagonal of D are the corresp. eigenvalues.

Why is diagonalization useful?

Suppose A is $n \times n$; $A = SDS^{-1}$.

(1). Determinant and trace.

Determinant: Recall that $\det(BC) = \det(B)\det(C)$

$$\Rightarrow \det(B^{-1}) = \frac{1}{\det(B)}.$$

Thus, $\det(A) = \det(S)\det(D)\det(S^{-1})$
$= \det(D)$

$$\Rightarrow \boxed{\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.}$$

Trace: For $n \times n$ B, $\operatorname{trace}(B) = \operatorname{tr}(B) = \sum_{i=1}^{n} B_{ii}$.

Fact: $\operatorname{tr}(BC) = \operatorname{tr}(CB)$ $\forall\, \underset{n\times k}{B}, \underset{k\times n}{C}$.

So: $tr(A) = tr(SDS^{-1}) = tr(S^{-1}SD) = tr(D) = \lambda_1 + \dots + \lambda_n$

$$\boxed{tr(A) = \lambda_1 + \dots + \lambda_n}.$$

(2). Power of diagonalisable matrices.

$A = SDS^{-1}$

$A^2 = SDS^{-1}SDS^{-1} = SD^2S^{-1}$

$\vdots$

$$\boxed{A^k = SD^kS^{-1}} \quad \forall k \in \mathbb{N}.$$

Note that $D^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}$ diagonal as well.

$\Rightarrow$ (*)
- $A^k = SD^kS^{-1}$ is the diagonalization of A^k.
- $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ are the eigenvalues of A^k.
- $v_1, v_2, \dots, v_n$ are the corresp. eigenvectors.

$\forall k \in \mathbb{N}$

In fact (*) holds $\forall k \in \mathbb{Z}$ as long as $\lambda_j \neq 0 \ \forall j$.

Hermitian matrices

Definition: A square matrix A is Hermitian if

$A^* = A$ (i.e. $\bar{A}^T = A$).

Examples:
- (all real symmetric matrices, i.e. A real $A = A^T$.)

  e.g. $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$

- $\begin{bmatrix} 1 & 1+i \\ 1-i & 5 \end{bmatrix}$.

Remarks: • all diagonal entries of a Hermitian matrix are real.

• a Hermitian matrix that is real is also called symmetric.

Properties:

① A is Hermitian iff $\langle v, Aw\rangle = \underset{= \langle A^* v, w\rangle}{\langle Av, w\rangle}$ $\forall v, w$.

② Eigenvalues of Hermitian matrices are real

Proof: Let $A = A^*$ be Hermitian, and suppose $Av = \lambda v$ for $v \neq \vec{0}$.

Then, $\langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle = \lambda \|v\|_2^2$.

$ \| $

$\langle Av, v\rangle = \langle \lambda v, v\rangle = \bar{\lambda}\langle v, v\rangle = \bar{\lambda}\|v\|_2^2$

$\Rightarrow \lambda = \bar{\lambda}$, i.e. $\lambda \in \mathbb{R}$.

③ Eigenvectors of Hermitian matrices corresponding to distinct eigenvalues are orthogonal.

i.e. If $Av_1 = \lambda_1 v_1$, $Av_2 = \lambda_2 v_2$ and $\lambda_1 \neq \lambda_2 \Rightarrow v_1 \perp v_2$.

Proof: See the typed notes.

Direct implication: If A is $n\times n$ Hermitian with distinct eigenvalues, then the corresp eigenbasis $\{v_1, \dots, v_n\}$ can be chosen to be an ONB!

diagonalization: $S = [v_1 | \dots | v_n]$ is unitary!

and $A = SDS^{-1} = SDS^*$

In this case, we say that $\boxed{A \text{ is } \underline{\text{unitarily decomposable}}}$

④. In fact, $\forall$ Hermitian matrix is unitarily decomposable, i.e. $A = SDS^*$, where S is unitary, even if A has repeated eigenvalues.

The operator norm of a Hermitian matrix:

Suppose $A = UDU^*$ and U is unitary
D diagonal.

Then, $\|A\|_{op} = \max\limits_{\|x\|_2=1} \|Ax\|_2 = \max\limits_{\|x\|_2=1} \|UD\underbrace{U^*x}_{y}\|_2$

$\|U^*x\|_2 = \|x\|_2$.
U^* is unitary

$= \max\limits_{\|y\|_2=1} \underset{\substack{\| \\ \|Dy\|_2 \\ U \text{ is unitary}}}{\|UDy\|_2} = \max\limits_{\|y\|_2=1} \|Dy\|_2 = \|D\|_{op}$

$= \max\{|\lambda_j| : j=1,\dots,n\}$

$$\|A\|_{op} = \max\{|\lambda_j| : j=1,\dots,n\}.$$

Condition number of a Hermitian matrix: suppose A is invertible

$\text{cond}(A) = \|A\|_{op}\|A^{-1}\|_{op} = \max\{|\lambda_j| : j=1,\dots,n\} \max\{|\frac{1}{\lambda_j}| : j=1,\dots,n\}$

$= \dfrac{\max\{|\lambda_j|\}}{\min\{|\lambda_j|\}}.$

Example: if $A = \begin{bmatrix} \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} & -\frac{1}{\sqrt2} \end{bmatrix} \begin{bmatrix} -2 & \\ & 5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt2} & -\frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}$

Then $\text{cond}(A) = \dfrac{5}{2}$.

IV.3. Power method for computing eigenvalues:

Suppose A is $n\times n$ with eigenvalues $\lambda_1,\dots,\lambda_n$ and corresp eigenvectors $v_1,\dots,v_n$ s.t

(i). $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$

(ii). $\{v_1,\dots,v_n\}$ is an eigenbasis.

(iii). $\|v_j\|_2 = 1$ $\forall j=1,\dots,n$

For simplicity, we also assume that $\lambda_j \in \mathbb{R}$, which is the case e.g. when A is Hermitian.

Goal: Find λ_1 and v_1 (the dominant eigenvalue and corresponding eigenvector).

Observe: Pick an arbitrary $x_0 \in \mathbb{R}^n$, $x_0 \neq \vec{0}$.

Then, $\exists c_1, \dots, c_n$ s.t.

$$x_0 = c_1 v_1 + c_2 v_2 + \dots + c_n v_n.$$

Then,
$$Ax_0 = c_1 A v_1 + c_2 A v_2 + \dots + c_n A v_n$$
$$= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + c_n \lambda_n v_n$$
$$A^2 x_0 = c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + \dots + c_n \lambda_n^2 v_n.$$
$$\vdots$$
$$A^k x_0 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \dots + c_n \lambda_n^k v_n.$$
$$= \lambda_1^k \Big(c_1 v_1 + \underbrace{c_2 \underbrace{\left(\frac{\lambda_2}{\lambda_1}\right)^k}_{\to 0 \text{ as } k \to \infty} v_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n}_{= \varepsilon_k} \Big).$$

Then: $A^k x_0 = \lambda_1^k (c_1 v_1 + \varepsilon_k)$

where $\varepsilon_k \to \vec{0}$ as $k \to \infty$.

$$\frac{A^k x_0}{\|A^k x_0\|} = \underbrace{\frac{\lambda_1^k}{|\lambda_1|^k}}_{= \pm 1 \text{ since } \lambda_1 \in \mathbb{R}} \frac{(c_1 v_1 + \varepsilon_k)}{\|c_1 v_1 + \varepsilon_k\|_2} = \pm \frac{c_1 v_1 + \varepsilon_k}{\|c_1 v_1 + \varepsilon_k\|_2} \longrightarrow \pm v_1 \text{ as } k \to \infty.$$

$\Rightarrow$ In other words, $\dfrac{A^k x_0}{\|A^k x_0\|} \to \pm v_1$ as $k \to \infty$

$\| \quad$ eigenvector corresponding to top eigenvalue λ_1.

So, we found v_1.
Next, to find λ_1, compute: $\langle v_1, A v_1 \rangle = \lambda_1 \|v_1\|^2 = \lambda_1$.