Recap | Lecture V | Aug 17

* **MMSE:** Given $x_1, x_2 \ldots x_n$.

$$y = c_1x_1 + c_2x_2 + \ldots c_nx_n.$$

M.M.S.E. $\boxed{\hat{c} = C_{xx}^{-1} C_{yx}}$; $\boxed{\hat{y} = \hat{c} \vec{X}}$

$$\boxed{\min \|e\|^2 = \sigma_y^2 - \hat{c}^T C_{xx}^{-1} C_{yx}}$$

* Eigen Values & Eigen Vectors

$$A\vec{x} = \lambda \vec{x}$$

$\vec{x}$ = eigen vector
λ = eigen value

Eigenvector $\in N(A - \lambda I)$

Characteristics Polynomial $P(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda)$.

$\det(A) = \prod_{i=1}^{n} \lambda_i$ = Product of eigenvalues.

* Eigenspace $\Rightarrow E_\lambda = N(A - \lambda I)$ is called eigenspace corresponding to λ

$\dim(E_\lambda) = \text{Nullity}(A - \lambda I)$.

* Diagonalization $A = T \Lambda T^{-1}$

A is diagonalizable if Eigenvectors of A are L.I. and $T = [v_1 \; v_2 \; \ldots v_n]$

* Linearly dependent EV

Ex Consider $A = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix}$

$\det(A - \lambda I) = (\lambda + 1)^2 = 0$

$\Rightarrow \boxed{\lambda = -1}$ and algebraic multiplicity is 2.

EVs $\Rightarrow (A + I)x = 0$

$\Rightarrow \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\Rightarrow \boxed{x_2 = 0}$

$\therefore$ Eigenvectors corresponding to $\boxed{\lambda = -1}$ are $\text{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$, geometric multiplicity $= 1$

Can you diagonalize this matrix? No.

$\star$ Consider: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\lambda = 1$; algebraic multiplicity $= 2$

$(A - \lambda I)\vec{x} = 0 \Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} x = 0$

$\Rightarrow x$ can be any vector.

$\Rightarrow x \in \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

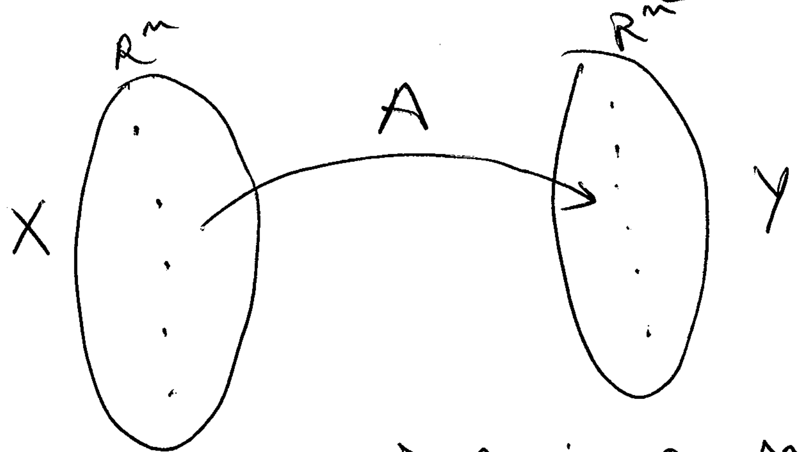
geometric multiplicity $= 2$

$\star$ ~~L.D. eigenvectors is not the right thing to.~~

Understanding Eigen space is more imp. than E.v. itself. You can choose E.V. from ES

* Invertibility of Matrix $\Leftrightarrow$ L.I. of column vectors.

If $A: R^n \rightarrow R^n$ is an $[n \times n]$ matrix.



Since A is a L.T. $\Rightarrow$ A is a mapping from $R^n \rightarrow R^n$

Let us assume that A is invertible

$\Rightarrow$ A is one to one and onto (By definition of invertible mapping)

$\Rightarrow$ ~~$\forall$ x~~ if $y = Ax$ then $\forall x \in R^n$ then exist a unique $y \in R^n$.

$$\boxed{y = x_1 a_1 + x_2 a_2 + \dots x_n a_n}$$

y is a L.C. of 'n' column vectors of 'A'

and $y \in R^n \Rightarrow \{a_1, \dots a_n\}$ should span $\{R^n\}$. We know that for n vectors to span a vector space of dimension 'n' they should be L.I.

Hence Proved

~~(5)~~ * Theorem: Trace of an $[n \times n]$ matrix A is the sum of its eigenvalues :-

Note that: $Tr(A) = a_{11} + a_{22} + \dots a_{nn}$
(Sum of diagonal)

If A & B are two $(n \times n)$ matrices then

$$Tr(AB) = Tr(BA)$$

$$Tr(AB) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij} = Tr(BA)$$

$AB \neq BA$ ~~&~~ $AB(i,j) \neq BA(i,j)$

only ~~&~~ $Tr(AB) = Tr(BA)$.

Now Let $A = T \Lambda T^{-1}$

$$\Rightarrow \quad \Lambda = T^{-1} A T$$

$$Trace(\Lambda) = \sum_{i=1}^{n} \lambda_i$$

$$Trace(T^{-1} A T) = Trace(A T T^{-1}) = Trace(A).$$

$$\Rightarrow \boxed{Trace(A) = \lambda_1 + \lambda_2 + \dots \lambda_n}$$

★ Theorem: If $\lambda_1, \dots \lambda_m$ are distinct eigenvalues of A and $\vec{v}_1, \dots \vec{v}_m$ are corresponding eigenvectors then $\{\vec{v}_1, \dots \vec{v}_m\}$ are L.I.

Corollary: If $A \in \mathbb{C}^{n\times n}$ has distinct eigenvalues $\lambda_1, \lambda_2 \dots \lambda_n$ then A is diagonalizable.

Proof: Let $\lambda_1, \lambda_2 \dots \lambda_n$ be n distinct eigenvalues of A. $\{v_i\}$ = eigenvectors

$$\Rightarrow \quad A\vec{v}_i = \lambda_i \vec{v}_i$$

Consider the L.C. of eigenvectors

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \dots c_n\vec{v}_n = 0 \qquad -①$$

If we can show that eq^n ① holds only when $c_i = 0 \; \forall i$ that means all v_i's are L.I. $\Rightarrow$ A is diagonalizable

To show that $c_1 = 0$ let us multiply both sides of ① by $(A-\lambda_2 I)(A-\lambda_3 I)\dots(A-\lambda_n I)$

Since $(A-\lambda_k I)\vec{v}_k = 0$, all terms disappear except for the first term, which is

$$c_1(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)\dots(\lambda_1-\lambda_n)v_1 = 0$$

$\because$ All λ_i are distinct (assumption) and $v_1 \neq 0$, we conclude that $c_1 = 0$

Similarly we can prove that all c_i's are also 0

Hence Proved

* Gram - Schmidt ~~Orthogonalization~~ Orthonormalisation.

If we have 'M' L.I. vectors $\{p_1, p_2 \ldots p_M\}$ that spans a vector space V, we can find 'M' ortho~~gonal~~normal basis vectors that spans V

~~Steps~~
Solution: Let $\{\vec{p}_1, \vec{p}_2 \ldots \vec{p}_M\}$ be M L.I. vectors $p_i \in V$

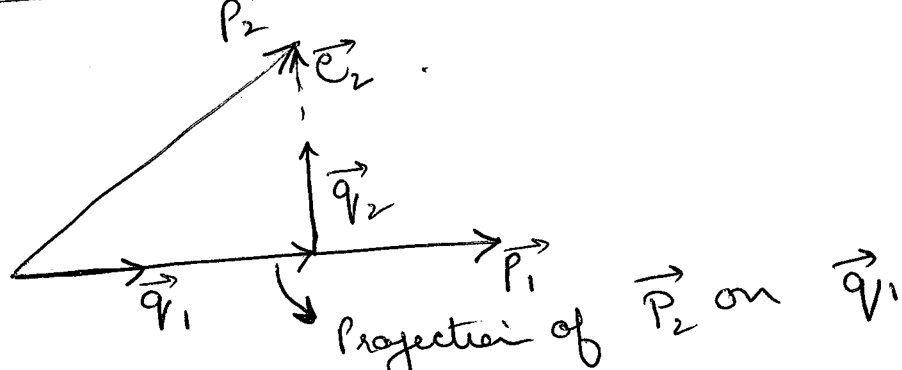
Step 1: Normalize first vector

$$\vec{q}_1 = \frac{\vec{p}_1}{\|\vec{p}_1\|}$$

Step 2: Compute the difference between the projection of $\vec{p}_2$ onto $\vec{q}_1$ and $\vec{p}_2$. By the orthogonality theorem, this is orthogonal to p_1:

$$\vec{e}_2 = \vec{p}_2 - \frac{\langle \vec{p}_2, \vec{q}_1 \rangle}{\|\vec{q}_1\|^2} \vec{q}_1$$

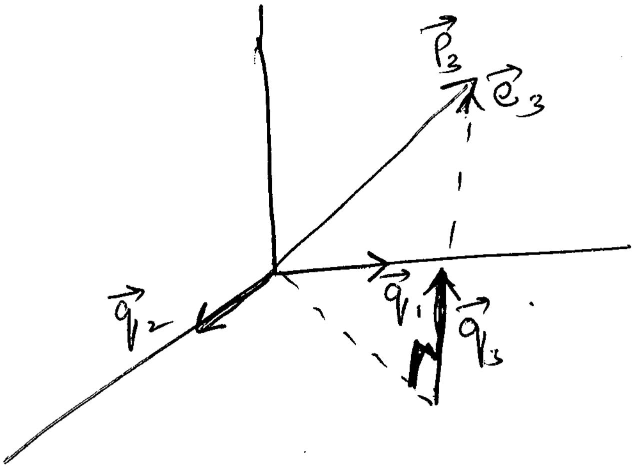
$$\boxed{\vec{e}_2 = \vec{p}_2 - \langle \vec{p}_2, \vec{q}_1 \rangle \vec{q}_1}$$



$$\therefore \quad \vec{q}_2 = \frac{\vec{e}_2}{\|\vec{e}_2\|^2}$$

Step 3: A vector orthogonal to $\vec{q}_1$ & $\vec{q}_2$ obtained from the error between $\vec{p}_3$ & its projection onto span $(\vec{q}_1, \vec{q}_2)$.

$$\vec{e}_3 = \vec{p}_3 - \langle \vec{p}_3, \vec{q}_1 \rangle \vec{q}_1 - \langle \vec{p}_3, \vec{q}_2 \rangle \vec{q}_2$$



$$\Rightarrow \quad \vec{q}_3 = \frac{\vec{e}_3}{\|\vec{e}_3\|}$$

4: Similarly: $$\vec{e}_k = \vec{p}_k - \sum_{i=1}^{k-1} \langle p_k, q_i \rangle q_i$$

$$\vec{q}_k = \frac{\vec{e}_k}{\|\vec{e}_k\|}$$

* The Spectral Theorem (Self-Adjoint Matrices)

This is an extremely important and powerful result on the diagonalizability of Hermitian / symmetric matrices $\Rightarrow$ (Symmetric interactions)

Hermitian Matrix $\Rightarrow$ $A = A^H$
Symmetric " $\Rightarrow$ $A = A^T$
$\Big\} \Rightarrow \langle Ax, y\rangle = \langle x, Ay\rangle$

<u>Proposition</u> $\Rightarrow$ If A is an $n \times n$ Hermitian matrix

(i) The eigenvalues of A are real. } <u>Lemma</u>

(ii) Eigenvectors corresponding to distinct eigenvalues are orthogonal. } <u>Lemma</u>

Proof: Suppose $A\vec{v} = \lambda \vec{v}$ $\quad v \neq 0$

(i)
$$\lambda \|v\|^2 = \vec{v}^H A \vec{v} = \vec{v}^H A^H \vec{v} = (A\vec{v})^H \vec{v} = (\lambda \vec{v})^H \vec{v} = \bar{\lambda}\, \vec{v}^H \vec{v} = \bar{\lambda}\, \|\vec{v}\|^2$$

$\Rightarrow \lambda = \bar{\lambda} \Rightarrow \boxed{\lambda \text{ is real}}$

(ii) Suppose $A\vec{v}_1 = \lambda \vec{v}_1$, $A\vec{v}_2 = \lambda \vec{v}_2$ for $\lambda_1 \neq \lambda_2$

$\langle A\vec{v}_1, \vec{v}_2\rangle = \langle \vec{v}_1, A\vec{v}_2\rangle = \langle v_1, \lambda_2 \vec{v}_2\rangle = \lambda_2 \langle v_1, v_2\rangle$

$\langle A\vec{v}_1, \vec{v}_2\rangle = \langle \lambda_1 \vec{v}_1, \vec{v}_2\rangle = \lambda_1 \langle \vec{v}_1, \vec{v}_2\rangle$

$\Rightarrow \boxed{(\lambda_1 - \lambda_2)\langle \vec{v}_1, \vec{v}_2\rangle = 0} \Rightarrow \langle \vec{v}_1, \vec{v}_2\rangle = 0 \Rightarrow \boxed{\vec{v}_1 \perp \vec{v}_2}$

$\because \lambda_1 \neq \lambda_2 \Rightarrow$

Therefore if eigenvalues of A (Hermitian/symmetric) matrix are distinct then the eigenvectors of A form a set of orthogonal basis / orthonormal of $\mathbb{R}^n$

Spectral Theorem

Statement :- We can decompose any symmetric matrix A with the eigen value decomposition

$$A = U \Lambda U^T$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n)$

and $U = [\vec{u}_1, \vec{u}_2 \dots \vec{u}_n]$ is orthogonal matrix i.e. $U^T U = U U^T = I_{[n \times n]}$ or $U^T = U^{-1}$

and $\{\vec{u}_1, \vec{u}_2 \dots \vec{u}_n\}$ are eigenvectors of A

Similarly $A = U \Lambda U^H$

where U is unitary matrix s.t.

$$U U^H = U^H U = I_{[n \times n]}$$

Both unitary and Orthonormal matrices are norm preserving

$$\begin{aligned} \|Ux - Uy\|^2 &= (Ux - Uy)^H (Ux - Uy) \\ &= (x-y)^H \underbrace{U^H U}_{I} (x-y) \\ &= \|x-y\|^2 \end{aligned}$$

Note: The proof of spectral theorem is straight forward when A has distinct eigenvalues. What if eigenvalues are not distinct?

* When A has repeated eigenvalues
Since A is diagonalizable the eigenvectors of A are L.I. Suppose the algebraic multiplicity of one eigenvalue λ is k.

$\Rightarrow \dim\{N(A-\lambda I)\} = k$ } this should be true otherwise A is not diagonizable

$E_\lambda = N(A-\lambda I)$

We can use Gram-Schmidt orthogonalizatn to find k orthogonal eigenvectors that span E_λ

Corollary : The geometric multiplicity of eigenvalue of A is equal to algebraic multiplicity

---

Projection interpretation

$$A\vec{x} = U\Lambda U^H \vec{x} = U\Lambda \begin{bmatrix} \vec{u}_1^H \vec{x} \\ \vec{u}_2^H \vec{x} \\ \vdots \\ \vec{u}_n^H \vec{x} \end{bmatrix} = U \begin{bmatrix} \lambda_1 u_1^H x \\ \lambda_2 u_2^H x \\ \vdots \\ \lambda_n u_n^H x \end{bmatrix}$$

$$= \sum_{i=1}^{n} \vec{u}_i \lambda_i \vec{u}_i^H \vec{x}$$

$$= \left(\sum_{i=1}^{n} \lambda_i \vec{u}_i \vec{u}_i^H \right) \vec{x}$$

$$= \lambda_1 \underbrace{\vec{u}_1 \vec{u}_1^H}_{\text{Projection onto } \vec{u}_1} \vec{x} + \lambda_2 u_2 u_2^H x + \ldots$$

$$= \lambda_1 \langle \vec{u}_1, \vec{x} \rangle \vec{u}_1 + \ldots$$

★ Rotation Interpretation :-

U is an orthogonal Matrix $U^TU = UU^T = I$

Let $U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for some $\theta \in [0, 2\pi]$

So $A = U \Lambda U^T$

$=$ rotate, scale, rotate back

★ Positive (Semi)-definite matrices

$x^H A x > 0 \quad \forall x \neq 0 \quad$ PD

$x^H A x \geq 0 \quad \forall x \quad$ PSD.

also denoted as $A > 0$, $A \geq 0$

Theorem: A is Hermitian and PD (PSD) $\Longleftrightarrow$ the eigenvalues of A are all +ve (non-negative).

Proof: Suppose $x \neq 0$

$$x^H A x = x^H \left(\sum_{i=1}^{N} \lambda_i u_i u_i^H \right) x$$

$$= \sum_{i=1}^{N} \lambda_i \, x^H u_i u_i^H x.$$

$$= \sum_{i=1}^{N} \lambda_i \, \| u_i^H x \|^2$$

> 0 if $\lambda_i > 0 \;\; \forall i$

≥ 0 if $\lambda \geq 0 \;\; \forall i$

★ Facts :- If A is Symmetric / Hermitian.

① A is PSD $\Rightarrow$ $\det(A) = \prod_{i=1}^{n} \lambda_i \geq 0$

② A is PD $\Rightarrow$ $\det(A) > 0$

③ A is PD $\Rightarrow$ A is invertible ($\cdot^{-1}$)

④ If A is PD, $A = U \Lambda U^T$, $\Lambda = \operatorname{diag}(\lambda_1 \ldots \lambda_n)$
$\lambda_i > 0 \quad \forall i$ then

$$A^{-1} = U \Lambda^{-1} U^T, \quad \Lambda^{-1} = \operatorname{diag}(\lambda_1^{-1} \ldots \lambda_n^{-1})$$

$$U \Lambda U^T U \Lambda^{-1} U^T = U \Lambda \Lambda^{-1} U^T$$
$$= U U^T$$
$$= I$$

$\Rightarrow$ A^{-1} is also PD.

★ Application $\Rightarrow$ Quadratic form : (Maximum Principle) For a PSD, (self adjoint) Hermitian matrix A with $Q_A(\vec{x}) = \langle A\vec{x}, \vec{x} \rangle = \vec{x}^H A \vec{x}$ is the ~~maximum~~ $\max Q_A(\vec{x})$ constraint to $\|\vec{x}\|^2 = 1$

λ_1 the largest eigenvalue of A and the maximizing $\vec{x}$ is $\vec{x}_1$ the eigenvector corresponding to λ_1

Proof: Consider the constrained optimization using Lagrange multipliers

$$J(\vec{x}) = \vec{x}^H A \vec{x} - \lambda(\vec{x}^H \vec{x} - 1)$$

$$\frac{\partial J}{\partial x} = A\vec{x} - \lambda \vec{x}$$

Maximizing solution must satisfy $A\vec{x} = \lambda \vec{x}$