Lecture 36 : Diagonalization of Symmetric Matrices

Recall that an $n \times n$ matrix A is called <u>symmetric</u> if $A = A^T$.

Ex 1/ $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}$ //

(<u>Non</u>-examples include rotation matrices, non-diagonal upper-triangular matrices, most orthogonal matrices.)

Ex 2/ $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \rightsquigarrow \det(A - \lambda I_2) = \lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2)$

$\mathrm{Nul}(A - 4I) = \mathrm{Nul}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \mathrm{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$, $\mathrm{Nul}(A - 2I) = \mathrm{Nul}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \mathrm{span}\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$. //

Notice that in Ex. 2 the eigenvalues are real and the eigenvectors orthogonal. This is no accident.

<u>Proposition 1</u>: Suppose $\vec{v}_1$ & $\vec{v}_2$ are (real) eigenvectors of a (real) symmetric matrix A, with distinct eigenvalues $\lambda_1, \lambda_2 \in \mathbb{R}$. Then

$$\vec{v}_1 \perp \vec{v}_2$$

under the dot product on $\mathbb{R}^n$.

<u>Proof</u>:
$$\lambda_1 \vec{v}_1 \cdot \vec{v}_2 = (A\vec{v}_1) \cdot \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$
$$\| \leftarrow A \text{ symmetric}$$
$$\lambda_2 \vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot (\lambda_2 \vec{v}_2) = \vec{v}_1 \cdot (A\vec{v}_2) = \vec{v}_1^T A \vec{v}_2$$

$$\Rightarrow \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0 \ (\lambda_1, \lambda_2 \text{ distinct})} \vec{v}_1 \cdot \vec{v}_2 = 0 \Rightarrow \vec{v}_1 \cdot \vec{v}_2 = 0.$$

$\square$

(Notice that the equality

$$(*) \qquad \boxed{(A\vec{v}_1)\cdot\vec{v}_2 = \vec{v}_1\cdot(A\vec{v}_2)}$$

for symmetric matrices that occurs in this proof has nothing to do with the fact that $\vec{v}_1$ & $\vec{v}_2$ are eigenvectors.)

Proposition 2: Eigenvalues of symmetric (real) matrices are real.

Proof: Let $\vec{v} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in \mathbb{C}^n$ be a nonzero (complex) eigenvector of A, w/eigenvalue $\lambda \in \mathbb{C}$.

Then $A\vec{v} = \lambda\vec{v}$, so

$$\lambda\vec{v}\cdot\overline{\vec{v}} = (A\vec{v})\cdot\overline{\vec{v}} \underset{(*)}{=} \vec{v}\cdot(A\overline{\vec{v}}) \underset{A \text{ real} \Rightarrow A = \overline{A}}{=} \vec{v}\cdot\overline{(A\vec{v})} = \vec{v}\cdot\overline{(\lambda\vec{v})}$$

(bar denotes complex conjugation)

$$= \overline{\lambda}\,\vec{v}\cdot\overline{\vec{v}} \qquad \Longrightarrow \qquad (\lambda - \overline{\lambda})\,\vec{v}\cdot\overline{\vec{v}} = 0.$$

Since $\vec{v}\cdot\overline{\vec{v}} = \sum_{i=1}^{n} \gamma_i\overline{\gamma_i} = \sum_{i=1}^{n} |\gamma_i|^2 > 0$, we therefore have $\lambda - \overline{\lambda} = 0$, or $\lambda = \overline{\lambda}$. That is, $\lambda \in \mathbb{R}$. $\square$

Remarks: (i) Proposition 1 $\Longrightarrow$ eigenspaces of a symmetric matrix are $\perp$

(ii) The simplest matrices with (non-real) complex eigenvalues are the rotation-dilation matrices $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b \neq 0$, which are not symmetric.

(iii) The simplest non-diagonalizable matrices are $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, $a \neq 0$. Clearly also not symmetric.

Orthogonal Diagonalization

The above two properties of symmetric matrices amount to the statement that if A is diagonalizable, then there is an orthonormal eigenbasis. Hence we can use an orthogonal matrix P to write $A = PDP^{-1} \underset{\substack{\uparrow \\ P \text{ orthogonal}}}{=} PDP^T$.

Ex 3/ $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Know $\lambda_1 = 4$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
$\lambda_2 = 2$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Normalizing gives $\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

So $A = PDP^T = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$. //

Why do we care? Suppose you have a discrete dynamical system

$$\vec{x}_{k+1} = A\vec{x}_k$$

with initial state vector $\vec{x}_0 = \vec{y}$, and A is symmetric w/ o.n. eigenbasis $\vec{u}_1, \dots, \vec{u}_n$ / eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Then

$$\vec{y} = \sum_{i=1}^{n} (\vec{y}\cdot\vec{u}_i)\vec{u}_i \implies \underbrace{\vec{x}_k = A^k\vec{y} = \sum_{i=1}^{n} (\vec{y}\cdot\vec{u}_i)\lambda_i^k\vec{u}_i}$$

is about the simplest solution we've seen so far.

Moreover, it is conceptually important to know that the magical eigenbasis is obtained by (essentially) a rotation of the standard basis $\vec{e}_1, \dots, \vec{e}_n$, which is what orthogonal diagonalization is telling you!

Ex 4 / $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightsquigarrow \det(A - \lambda I_3) = \lambda^2(\lambda - 3)$

$\rightsquigarrow$ eigenvalues $\begin{cases} 0 & \text{(multiplicity 2)} \\ 3 & \text{(multiplicity 1)} \end{cases}$

Eigenspaces $E_0 = \mathrm{Nul}(A) = \mathrm{Span}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

$E_3 = \mathrm{Nul}(A - 3I) = \mathrm{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$. $\quad\perp$

To get an o.n. eigenbasis, apply Gram-Schmidt to the basis of E_0: $\vec{v}_1 := \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \dfrac{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix}$

$\Rightarrow \vec{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$.

Then apply to the basis of $E_3 \rightsquigarrow \vec{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

So we get $D = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 3 \end{pmatrix}$ and $P = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$. //

We now know that if A is symmetric and diagonalizable, it is orthogonally diagonalizable. But <u>is every symmetric matrix diagonalizable</u>?

<u>The Spectral Theorem</u>: YES.

<u>Proof</u>: Idea is to find one eigenvector, then argue that you can find more (in its orthogonal complement) in the same way.

So consider the higher-dimensional "sphere" of unit vectors

$$\mathcal{S} := \{ \vec{v} \in \mathbb{R}^n \mid \|\vec{v}\| = 1 \}$$

and define a continuous function

$$f : \mathcal{S} \to \mathbb{R}$$

by $f(\vec{v}) := \vec{v} \cdot A\vec{v}$. By the "maximum principle" from Calculus, or by intuition, there is a $\vec{w}_1 \in \mathcal{S}$ such that

$$f(\vec{w}_1) \geq f(\vec{v}) \quad \text{for all } \vec{v} \in \mathcal{S}.$$

Let $W = \text{span}\{\vec{w}_1\}$ and decompose

$$\begin{aligned} A\vec{w}_1 &= \text{proj}_W(A\vec{w}_1) + (A\vec{w}_1 - \text{proj}_W(A\vec{w}_1)) \\ &= \lambda_1 \vec{w}_1 + \vec{z} \quad , \quad \text{where } \vec{z} \perp \vec{w}_1 . \end{aligned}$$

I claim that $\vec{z} = \vec{0}$, so that $\vec{w}_1$ is an eigenvector.
If instead $\vec{z} \neq \vec{0}$, we could divide by $\|\vec{z}\|$, and consider the path

$$\begin{cases} \varphi : \mathbb{R} \to \mathcal{S} \\ t \longmapsto (\cos t)\vec{w}_1 + (\sin t)\dfrac{\vec{z}}{\|\vec{z}\|} \end{cases}$$

this is a unit vector (why?)

passing through $\vec{w}_1$ at "time" $t = 0$.
Since f has maximum at $\vec{w}_1$, the composition

$$f \circ \varphi : \mathbb{R} \to \mathbb{R}$$

has a maximum at $t = 0$. Calculus tells us that

$$0 = (f \circ \varphi)'(0) = \left(\frac{d}{dt} \{ \varphi(t) \cdot A\varphi(t) \} \right)(0)$$

$$= \varphi'(0)\cdot A\varphi(0) + \varphi(0)\cdot A\varphi'(0)$$

since A is symmetric uses (*)

$$= \varphi'(0)\cdot A\varphi(0) + A\varphi(0)\cdot\varphi'(0)$$

$$= 2(A\varphi(0))\cdot\varphi'(0)$$

$$= 2(A\vec{w}_1)\cdot\left(\underbrace{\cancel{\cos'(0)}}_{0}\,\vec{w} + \underbrace{\cancel{\sin'(0)}}_{1}\,\frac{\vec{z}}{\|\vec{z}\|}\right)$$

$$= 2(\lambda\vec{w}_1 + \vec{z})\cdot\frac{\vec{z}}{\|\vec{z}\|}$$

$\vec{z}\perp\vec{w}_1$

$$= 2\,\frac{\vec{z}\cdot\vec{z}}{\|\vec{z}\|} = 2\|\vec{z}\| \quad\Longrightarrow\quad \|\vec{z}\| = 0,$$

in contradiction to our assumption $\vec{z}\neq\vec{0}$. So in fact our claim must be true, i.e. $\vec{z}=\vec{0}$, and

$$A\vec{w}_1 = \lambda_1\vec{w}_1 .$$

Now what? If $\vec{x}\in W^{\perp}$, then

$$\vec{w}_1\cdot A\vec{x} = A\vec{w}_1\cdot\vec{x} = \lambda_1\vec{w}_1\cdot\vec{x} = 0$$

$\Longrightarrow A\vec{x}\in W^{\perp}$. Apply the above argument to

$$S_1 = \{\vec{v}\in W^{\perp} \mid \|\vec{v}\| = 1\}$$

to get a new eigenvector w_2 of A (in $W^{\perp}$), with eigenvalue λ_2. Now take the orthogonal complement of $\operatorname{span}\{\vec{w}_1,\vec{w}_2\}$, and keep going. Eventually you get n (orthogonal) eigenvectors and you're done. $\square$

By the spectral theorem, we have the "Spectral decomposition"

$$A = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{u}_1 & \cdots & \vec{u}_n \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} \leftarrow \vec{u}_1^T \rightarrow \\ \vdots \\ \leftarrow \vec{u}_n^T \rightarrow \end{pmatrix}$$

$$= (\vec{u}_1, \dots, \vec{u}_n) \begin{pmatrix} \lambda_1 \vec{u}_1^T \\ \vdots \\ \lambda_n \vec{u}_n^T \end{pmatrix} = \lambda_1 \underbrace{\vec{u}_1 \vec{u}_1^T}_{\substack{(n\times 1)\cdot(1\times n) \\ = n\times n \text{ projection matrix} \\ \text{to span}\{\vec{u}_1\}}} + \dots + \lambda_n \vec{u}_n \vec{u}_n^T$$

which is another way of getting

$$A\vec{y} = \lambda_1 \vec{u}_1 \underbrace{\vec{u}_1^T \vec{y}}_{\vec{u}_1 \cdot \vec{y}} + \dots + \lambda_n \vec{u}_n \underbrace{\vec{u}_n^T \vec{y}}_{\vec{u}_n \cdot \vec{y}} = \lambda_1 (\vec{y}\cdot\vec{u}_1)\vec{u}_1 + \dots + \lambda_n (\vec{y}\cdot\vec{u}_n)\vec{u}_n.$$

What's nice about the spectral decomposition is that it breaks A up into a sum of orthogonal projections to eigenspaces, weighted by A's eigenvalues.

Ex 5 / Let's see how this looks for $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$:

$$A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \lambda_2 \vec{u}_2 \vec{u}_2^T = 4 \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} + 2 \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Notice the effects of each matrix on, say,

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

//

To sum up: given an $n \times n$ symmetric real matrix A, we can always orthogonally diagonalize it as follows:

<u>Step 1</u>: Find the eigenvalues using $0 = \det(A - \lambda I)$.

<u>Step 2</u>: Find bases for each eigenspace $E_{\lambda_0} = \mathrm{Nul}(A - \lambda_0 I)$.

<u>Step 3</u>: Apply Gram-Schmidt (with respect to the dot product) to each of the bases from Step 2. If E_{λ_0} is 1-dim'l, just normalize the vector. (By the way, here I mean "full" G-S)

<u>Step 4</u>: Form D & P and write $A = PDP^T$.