

Lecture 36: Diagonalization of Symmetric Matrices

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

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

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

Ex 2 / $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \rightarrow \det(A - \lambda \mathbb{I}_2) = \lambda^2 - 6\lambda + 8 = (\lambda - 4)(\lambda - 2)$

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

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

Proposition 1: 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 .

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

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

□

(Notice that the equality

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

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 $\vec{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n$ be a nonzero (complex) eigenvector of A , w/ eigenvalue $\lambda \in \mathbb{C}$.

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

$$\begin{aligned} \lambda\vec{v} \cdot \overline{\vec{v}} &= (A\vec{v}) \cdot \overline{\vec{v}} = \vec{v} \cdot (A\overline{\vec{v}}) = \vec{v} \cdot \overline{(A\vec{v})} = \vec{v} \cdot \overline{(\lambda\vec{v})} \\ &= \overline{\lambda} \vec{v} \cdot \overline{\vec{v}} \quad \Rightarrow \quad (\lambda - \overline{\lambda}) \vec{v} \cdot \overline{\vec{v}} = 0. \end{aligned}$$

(For denotes complex conjugation) \uparrow () \uparrow A real $\Rightarrow A = \overline{A}$*

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

Remarks: (i) Proposition 1 \Rightarrow 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 $A = P D P^{-1} = P D P^T$.

\uparrow
P orthogonal

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

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

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

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 / $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightarrow \det(A - \lambda I_3) = \lambda^2(\lambda - 3)$
 \rightarrow eigenvalues $\begin{cases} 0 & (\text{multiplicity } 2) \\ 3 & (\text{multiplicity } 1) \end{cases}$

Eigenspaces $E_0 = \text{Nul}(A) = \text{Span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$
 $E_3 = \text{Nul}(A - 3I) = \text{Span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$. \perp

To get an o.n. eigenbasis, apply Gram-Schmidt to the basis of E_0 : $\vec{v}_1 := \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}}{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$
 $\Rightarrow \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$.

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

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 is every symmetric matrix diagonalizable?

The Spectral Theorem: YES.

Proof: 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

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

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

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

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

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

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

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

$$\begin{aligned}
 &= \varphi'(0) \cdot A \varphi(0) + \varphi(0) \cdot A \varphi'(0) \\
 &= \varphi'(0) \cdot A \varphi(0) + A \varphi(0) \cdot \varphi'(0)
 \end{aligned}$$

since A is symmetric $(*)$

$$\begin{aligned}
 &= 2(A \varphi(0)) \cdot \varphi'(0) \\
 &= 2(A \vec{w}_1) \cdot \left(\cos'(0) \vec{w}_2 + \sin'(0) \frac{\vec{z}}{\|\vec{z}\|} \right)
 \end{aligned}$$

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

$$\vec{z} \perp \vec{w}_1 \implies 2 \frac{\vec{z} \cdot \vec{z}}{\|\vec{z}\|} = 2 \|\vec{z}\| \implies \|\vec{z}\| = 0,$$

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

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

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

$$\mathcal{S}_1 = \{ \vec{v} \in W^\perp \mid \|\vec{v}\| = 1 \}$$

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

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

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

Step 2: Find bases for each eigenspace $E_{\lambda_0} = \text{Nul}(A - \lambda_0 I)$.

Step 3: 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)

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

