Last time: ORTHOGONAL MATRICES

$A \in M_n(\mathbb{R})$ with $\underline{\underline{A^{-1} = A^T}}$.

Facts: If A is orthogonal, then: • $A^TA = AA^T = I_n$.

- Rows and columns of A are orthonormal bases of $\mathbb{R}^n$
- $\|Ax\| = \|x\|$ and $Ax \cdot Ay = x \cdot y$ for all x,y in $\mathbb{R}^n$ (ie A preserves lengths of vectors (hence distances/angles)

ORTHOGONAL DIAGONALIZATION

— Making process of diagonalization simpler

If $A = PDP^{-1}$ with D diagonal,

then $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ for $\lambda_1, \ldots, \lambda_n$ eigenvalues of A

and $P = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ has columns the corresponding eigenvectors of A

since (rearranging) $AP = PD$ (so $Av_i = \lambda_i v_i$ for each i)

Let's start with real matrices.

<u>Defⁿ</u> We say that A (real matrix) is orthogonally diagonalizable if $A = PDP^{-1}$ for D (real) diagonal & P orthogonal (i.e. $P^{-1} = P^T$ so $A = PDP^T$).

So what the reasoning above tells us is

<u>Theorem</u> $A \in M_n(\mathbb{R})$, A is orthogonally diagonalizable exactly when A has an orthonormal set of n eigenvectors.

When does this happen?

Where can we hope to find A with an orthogonal diagonalization?

First suppose $A \in M_n(\mathbb{R})$ has an orthogonal diag.

Then $A = PDP^T$ for P orthogonal, D diagonal.

What is $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PD^T P^T = PDP^T$
$= A$

i.e. A symmetric. In fact:

Theorem (Spectral Theorem for Symmetric Matrices)

$A \in M_n(\mathbb{R})$ is orthogonally diagonalizable exactly when A is symmetric

Goal: Find a procedure for finding an orthogonal diagonalization of a symmetric matrix A.

To do this we'll need:

Theorem If $A \in M_n(\mathbb{R})$ is symmetric, then

(i) the eigenvalues of A are all real numbers;

(ii) eigenvectors from different eigenspaces (i.e. eigenvectors corresponding to different eigenvalues) are orthogonal to one another.

[Proof] To prove this we need to enter the world of complex matrices/vectors:

Recall: $A = [a_{ij}]_{i,j}$ then $\overline{A} = [\overline{a_{ij}}]_{i,j}$

$\uparrow$ The conjugate of A

(This will play a big role later!!)

You should verify that $\overline{(AB)} = \overline{A}\,\overline{B}$

the conjugate transpose of $A \longrightarrow \overline{A}^T = \overline{A^T}$

In $\mathbb{R}^n$ $x \cdot y = x^T y$. In $\mathbb{C}^n$ $x \cdot y = x^T \bar{y}$.

Here comes the proof:

(i) Let λ be an eigenvalue of A. We'll prove $\lambda = \bar{\lambda}$.

We know $Av = \lambda v$ for some $v \in \mathbb{C}^n$.

We'll solve for λ by multiplying the equation by $\bar{v}^T$!
↳ taking dot products with v

$$\bar{v}^T A v = \bar{v}^T \lambda v$$
$$= \lambda (\bar{v}^T v)$$
$$= \lambda (v \cdot v) = \lambda \|v\|^2 \qquad \text{So } \lambda = \frac{\bar{v}^T A v}{\underset{\text{real}}{\|v\|^2}}$$

So now, enough to show $\bar{v}^T A v$ is real.
i.e. $\overline{(\bar{v}^T A v)} = \bar{v}^T A v$.

$$\overline{(\bar{v}^T A v)} = \bar{\bar{v}}^T \bar{A} \bar{v} = v^T A \bar{v} \qquad \text{(as } A \text{ real)}$$
$$\left.\begin{aligned} &= (\bar{v}^T A^T v)^T \\ &= \bar{v}^T A^T v \end{aligned}\right\} \text{ just a number, so equal to its own transpose}$$
$$= \bar{v}^T A v \qquad \text{as } A \text{ symmetric.}$$

(ii) Take v_1 and v_2 eigenvectors of distinct eigenvalues $\lambda_1 \neq \lambda_2$ (i.e. $Av_i = \lambda_i v_i$, $i = 1, 2$)

We want to show v_1 & v_2 orthogonal i.e. $v_1 \cdot v_2 = 0$.

$$(Av_1)\cdot v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T A v_2$$

$$= v_1 \cdot (\lambda_2 v_2) = \lambda_2 (v_1 \cdot v_2)$$

$$\Vert$$

$$(\lambda_1 v_1)\cdot v_2 = \lambda_1 (v_1 \cdot v_2)$$

So $\underbrace{(\lambda_2 - \lambda_1)}_{\neq 0}(v_1 \cdot v_2) = 0$

So $v_1 \cdot v_2 = 0. /\!\!/$

Recap: How to diagonalize a matrix $A \in M_n(\mathbb{R})$?

1. Find eigenvalues of A (solve characteristic eq. $\det(A - \lambda I) = 0$)
2. Find eigenvectors corresponding to eigenvalues $v_1, \dots, v_m$

Q: Are there n linearly independent eigenvectors?

If yes, then

3. $A = PDP^{-1}$ where $P = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ and $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ for λ_i eigenvalue corresponding to v_i

So our procedure for finding an orthogonal diagonalization of a symmetric matrix A is as follows:

Step 1 Find the eigenvalues of A.

Step 2 Find a basis of eigenvectors for each eigenspace.

[So far this is exactly the same procedure as diagonalization. Here, since A symmetric ($n \times n$ say) by Theorem, we definitely will get n linearly independent eigenvectors. The issue is, we need to "orthonormalize" them i.e. find an orthonormal set of n eigenvectors.]

Step 3 For each eigenspace/eigenvalue use Gram—Schmidt Process to find an orthonormal basis for each eigenspace in turn.

[By the above theorem, this is enough. Within eigenspaces, the basis vectors are orthogonal by Gram—Schmidt, and basis vectors from distinct eigenspaces are orthogonal by the theorem.]

Step 4 Form P with columns the vectors found from Step 3. [P is orthogonal & $A = PDP^{-1}$ with eigenvalues of A on the diagonal of D.]

Example Find an orthogonal diagonalization of

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \leftarrow \text{A symmetric.}$$

Solution (skeleton)

Step 1 Find eigenvalues i.e. solve $\det(A - \lambda I) = 0$

i.e. $0 = \det\begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} = \ldots = -(\lambda+1)^2(\lambda-2)$

So eigenvalues are $\lambda = -1, 2$

Step 2 Find eigenvectors for each eigenvalue.

i.e. solve system $(A - \lambda I)v = 0$ for each λ.

$\lambda = -1$

i.e. solve $(A+I)v = 0$

Augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \xrightarrow[\text{red.}]{\text{row}} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So $v_1 + v_2 + v_3 = 0$. If $v_2 = s$, $v_3 = t$

then eigenvectors have form $\begin{pmatrix} -s-t \\ s \\ t \end{pmatrix}$

$$= s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

So $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for the eigenspace of $\lambda = -1$

$\lambda = 2$

i.e. solve $(A - 2I)v = 0$

$$\left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right] \xrightarrow[\text{red.}]{\text{row}}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So $v_1 = v_2 = v_3$; so if $v_3 = s$

then eigenvectors have the form $\begin{pmatrix} s \\ s \\ s \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

So $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ is a basis for the eigenspace of $\lambda = 2$.

Step 3 "Orthonormalize" the basis for each eigenspace.

Run Gram–Schmidt on $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ . . .

get $\left\{ \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2}\sqrt{3} \\ -1/\sqrt{2}\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{pmatrix} \right\}$

Run Gram–Schmidt on $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ i.e. normalize $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

. . .

get $\left\{ \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}$.

Notice (or check!) that all $\underline{\underline{3}}$ vectors are orthogonal

Step 4 $P = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2}\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2}\sqrt{3} & 1/\sqrt{3} \\ 0 & \sqrt{2}/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$. Then $P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Notice When an eigenspace is spanned by a single vector Step 3 ("run Gram–Schmidt") is really easy – just normalize the basis vector.

In particular, this will happen for every eigenspace if $\underset{n \times n}{A \text{ (symmetric)}}$ has $\underline{\boxed{n} \text{ distinct eigenvalues}}$.

Analogue (of orthogonal matrices) for Complex matrices

If $A \in M_n(\mathbb{R})$ then A orthogonal i.e. $A^{-1} = A^T$ has orthonormal columns & rows because of how the real dot product is defined.

(For columns $v_1, \dots, v_n$, say, $\|v_i\| = 1$, $v_i \cdot v_j = 0, i \neq j$.)

But in $\mathbb{C}^n$, $x \cdot y = x_1\bar{y}_1 + \dots + x_n\bar{y}_n$ so the analogue to orthogonal matrix is

Def^n A in $M_n(\mathbb{C})$ is unitary if $A^{-1} = \underbrace{\bar{A}^T}_{\text{conjugate transpose of } A} =: A^*$

Theorem If A, B unitary then (i) A^* (ii) A^{-1} (iii) AB are all unitary.

(iv) $\|\det(A)\| = 1$.

Theorem For A in $M_n(\mathbb{C})$ the following equivalent

(1) A unitary

(2) rows of A form an orthonormal basis of $\mathbb{C}^n$ (wrt complex dot product)

(3) columns of A " " " " " " — " —

(4) $\|Ax\| = \|x\|$ for any x in $\mathbb{C}^n$

(5) $Ax \cdot Ay = x \cdot y$ for any x, y in $\mathbb{C}^n$ (complex dot product)

Theorem $[I]_{B',B}$ ← Any transition matrix between orthonormal bases B, B' of a complex inner product ~~vector~~ space is unitary.

We also have here the concept of unitary diagonalization

Def^n A in $M_n(\mathbb{C})$ is said to unitary diagonalizable if $A = PDP^{-1}$ where D diagonal & P is unitary ($P^{-1} = \overline{P}^T = P^*$ so $A = PDP^*$).

As in the orthogonal case, by the theorem above:

Theorem $A \in M_n(\mathbb{C})$ has a unitary diagonalization exactly when A has orthonormal eigenvectors.

When do we have unitary diagonalization? Just as in real case let's look at what $A \in M_n(\mathbb{C})$ has to look like if A has a unitary diagonalization:

$$A = PDP^{-1} \quad \text{So}$$

$$\overline{\overline{P}}^T = P^T$$

$$A^* = \overline{A}^T = \overline{(PDP^{-1})}^T = (\overline{P}\,\overline{D}\,\overline{P^{-1}})^T$$

$$= P\overline{D}\,\overline{P}^T \qquad (\text{as } \overline{D} = \overline{D}^T)$$

$$\uparrow \text{diagonal}$$

$$= P\overline{D}P^* = P\overline{D}P^{-1}$$

If all eigenvalues of A are real, $D = \overline{D}$ and then we know $\underbrace{A = A^* = \bar{A}^T}_{}$.

Important Definition : $A \in M_n(\mathbb{C})$ is called Hermitian if $A = A^*$.

Aside on Def^ns	Real	Complex
	A orthogonal : $A^{-1} = A^T$	A unitary $A^{-1} = A^*$ $= \bar{A}^T$
	A symmetric : $A = A^T$	A Hermitian $A = A^*$

$\nearrow$ Almost analogue of symmetric. Why "almost"?

Theorem If A is unitary diagonalizable, then A is Hermitian if all eigenvalues are real

(By argument above.)

But we do have:

Theorem If A is Hermitian, then A is unitary diagonalizable.

Observation about identifying Hermitian matrices:

$A = \bar{A}^T$ so "mirrored" entries are conjugate to one another and in particular diagonal entries are real. e.g. $A = \begin{pmatrix} 3 & 2-i \\ 2+i & 5 \end{pmatrix}$

Go back argument above. Suppose A is unitary diagonalizable. We showed (from $A = PDP^*$) that

$$A^* = P\bar{D}P^*$$

Now compute
$$\begin{aligned} AA^* &= (PDP^*)(P\bar{D}P^*) \\ &= (PDP^{-1})(P\bar{D}P^{-1}) \\ &= PD\bar{D}P^{-1} \end{aligned}$$

and $D\bar{D} = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix}$.

Now compute
$$\begin{aligned} A^*A &= (P\bar{D}P^*)(PDP^*) \\ &= (P\bar{D}P^{-1})(PDP^{-1}) \\ &= P\bar{D}DP^{-1} \end{aligned}$$

but $\bar{D}D = D\bar{D}$ so we get $\underbrace{AA^* = A^*A}$.

<u>Def^n</u> Matrices A with $AA^* = A^*A$ are called <u>normal</u>.

In fact:
<u>Theorem</u> A is unitary diagonalizable exactly when A is normal.

Procedure for finding a unitary diagonalization of a Hermitian matrix is <u>same</u> as for finding an orthogonal diagonalization of a symmetric matrix:—

Because of this theorem:

<u>Theorem</u> If A is Hermitian, then eigenvectors from different eigenspaces are orthogonal (don't forget: with respect to the complex dot product)