Lecture 22: Special matrices

Reading: Meyer 7.5 Normal matrices
7.6 Positive semi-definite matrices

Last time:

THEOREM: A has a complete, orthogonal set of eigenvectors

$$\Updownarrow$$

$$A^\dagger A = A A^\dagger . \qquad \text{("A is normal")}$$

$$\Rightarrow A = \underset{\text{unitary w/ e-vector columns}}{U} \; \underset{\text{diagonal w/ e-values}}{D} \; U^\dagger$$

Examples:

- Every diagonal matrix is normal ($U = I$ above)
- No upper- or lower-triangular matrix is normal, unless it is diagonal!

```
>> n = 4;
>> A = randn(n,n);
>> A = A + A';        % make the matrix symmetric (and hence normal)
>> [U, D] = eigs(A)

U =
```

$\vec{v}_1$	$\vec{v}_2$	$\vec{v}_3$	$\vec{v}_4$
0.6207	-0.5958	0.1753	-0.4785
0.6988	0.1282	-0.1866	0.6786
-0.1355	-0.3777	-0.9151	-0.0408
-0.3287	-0.6971	0.3116	0.5558

```
D =
```

λ_1	λ_2	λ_3	λ_4
-7.2760	0	0	0
0	3.7712	0	0
0	0	1.8551	0
0	0	0	-0.3472

```
>> A * U - U * D

ans =

   1.0e-14 *
```

$A\vec{v}_j = \lambda_j \vec{v}_j$

0.3553	-0.1776	0.0056	-0.1832
-0.0888	-0.0722	0.1554	-0.3608
-0.1221	-0.0444	0.1110	0.0073
0.0444	0	-0.0888	-0.0416

the e-vectors are orthonormal:

```
>> U' * U

ans =
```

1.0000	0	0.0000	0.0000
0	1.0000	0.0000	0.0000
0.0000	0.0000	1.0000	-0.0000
0.0000	0.0000	-0.0000	1.0000

Proposition: A normal $\Rightarrow \|A\| = \max_{\text{eigenvalues } \lambda} |\lambda|$.

Proof:

Recall $\|A\| = \max_{x : \|x\|=1} \|Ax\|$. Since unitaries don't change lengths,

$$\begin{aligned} \|A\| &= \|UDU^{-1}\| \\ &= \|D\| \\ &= \max_i |D_{i,i}| \checkmark \end{aligned}$$

$\square$

This proposition is **false** for non-normal matrices, e.g., both eigenvalues of $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ are 0, but $\left\|\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\| = 1 \neq 0$. ✓

Today: More examples and applications.

APPLICATIONS of unitary diagonalizability of normal matrices

I. Eigenvalues of orthogonal and unitary matrices

Recall:

Definition:

$n \times n$ **real** matrix w/ orthonormal columns = orthogonal $A^T = A^{-1}$

$n \times n$ **complex** matrix w/ orthonormal columns = unitary $A^\dagger = A^{-1}$

$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ an orthonormal basis for $\mathbb{R}^n$ or $\mathbb{C}^n$ $\Longleftrightarrow$ $A = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{pmatrix}$ is orthogonal or unitary

Example: 2D rotation $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is orthogonal and unitary

$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$ is unitary.

Every orthogonal matrix is unitary.

Theorem: Any orthogonal or unitary matrix is diagonalizable.

Examples:
$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}$$
$\omega = e^{2\pi i/3}$

Proof: U unitary $\Rightarrow U^\dagger = U^{-1}$
$\Rightarrow UU^\dagger = I = U^\dagger U$
$\Rightarrow U$ is normal ✓ $\square$

Claim: All eigenvalues of an orthogonal or unitary matrix have magnitude 1.

Proof:

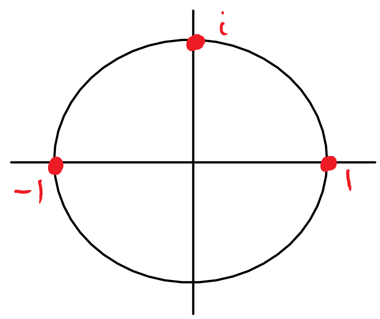
Say $U^\dagger = U^{-1}$, and $\vec{v} \neq \vec{0}$ is an e-vector, $U\vec{v} = \lambda\vec{v}$.

$$\Rightarrow \|\vec{v}\|^2 = v^\dagger v = v^\dagger \overbrace{U^\dagger U}^{I} v$$
$$= \|Uv\|^2$$
$$= |\lambda|^2 \times \|v\|^2 \qquad \Rightarrow |\lambda| = 1 \qquad \square$$

Examples:

1) Fourier transform mod 3:

```
>> omega = exp(2*pi*i/3);
C = (1/sqrt(3)) * [1 1 1; 1 omega omega^2; 1 omega^2 omega]
eig(C)

C =

   0.5774 + 0.0000i   0.5774 + 0.0000i   0.5774 + 0.0000i
   0.5774 + 0.0000i  -0.2887 + 0.5000i  -0.2887 - 0.5000i
   0.5774 + 0.0000i  -0.2887 - 0.5000i  -0.2887 + 0.5000i


ans =

   1.0000 + 0.0000i
  -1.0000 + 0.0000i
   0.0000 + 1.0000i

>> abs(eig(C))

ans =

    1.0000
    1.0000
    1.0000
```



Fourier transform mod n:

```
>> n = 100;
>> C = dftmtx(n) / sqrt(n);
>> sum(sum(abs(  C^4 - eye(n)  )))

ans =

  2.2902e-13
```

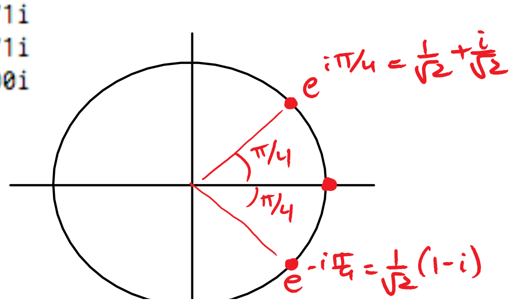
$C^4 = I$

$\Rightarrow$ all e-values must be in $\{\pm 1, \pm i\}$.

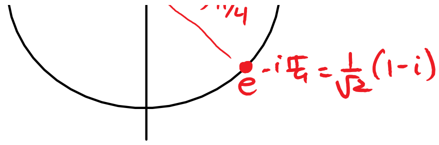
2)
```
>> B = (1/sqrt(2)) * [1 0 1; 0 sqrt(2) 0; -1 0 1]
eig(B)
abs(eig(B))

B =

    0.7071         0    0.7071
         0    1.0000         0
   -0.7071         0    0.7071
```

$\leftarrow$ rotation about the y-axis in $\mathbb{R}^3$

```
ans =

   0.7071 + 0.7071i
   0.7071 - 0.7071i
   1.0000 + 0.0000i


ans =

    1.0000
    1.0000
```



1.0000
1.0000
1.0000



Observe: The eigenvalues are paired up:

$\lambda = e^{i\theta}$ an e-value $\Rightarrow$ so is $\lambda^* = e^{-i\theta}$ (complex conjugate)

Why?

Theorem: If A is any real matrix, then

- λ an e-value of A $\Rightarrow$ so is λ^*

w/ corr. e-vector $\vec{v}$ w/ corr. e-vector v^*

Proof:

$$A v = \lambda v$$

$\Rightarrow A^* v^* = \lambda^* v^*$ (taking complex conjugates of everything)

$\Rightarrow A v^* = \lambda^* v^*$ (since A is real, $A^* = A$). ✓ $\square$

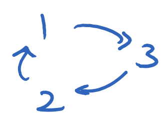
Exercise: What are the eigenvalues of

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} ?$$

Answer: A is orthogonal, so its eigenvalues lie on the unit circle.

A is a permutation matrix:

$$A e_1 = e_3$$
$$A e_2 = e_1$$
$$A e_3 = e_2$$



$\Rightarrow A^3 = I$

$\Rightarrow$ each eigenvalue λ must satisfy $\lambda^3 = 1$

$\Rightarrow$ each $\lambda \in \{1, e^{2\pi i/3}, e^{-2\pi i/3}\}$

Since $A \neq I$, and A is diagonalizable, the e-values aren't all 1.

$\Rightarrow$ Either $e^{2\pi i/3}$ or $e^{-2\pi i/3}$ must be an e-value

$\Rightarrow$ Both $e^{2\pi i/3}$ and $e^{-2\pi i/3}$ are e-values

(since A is real, complex e-values are paired up)

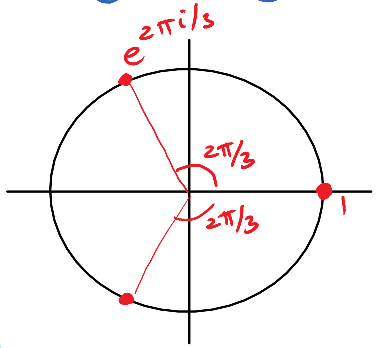
Only one dimension remains, so there isn't room for another complex e-value $\Rightarrow$ It must be real

$\Rightarrow$ It is 1

Answer: $\boxed{1,\ e^{2\pi i/3},\ e^{-2\pi i/3}}$

Check: $\mathrm{Trace}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = 0 + 0 + 0 = 0$

and $1 + e^{2\pi i/3} + e^{-2\pi i/3} = 0$



Check in Matlab:

```
>> A = [0 1 0; 0 0 1; 1 0 0]
eig(A)

A =

     0     1     0
     0     0     1
     1     0     0


ans =

  -0.5000 + 0.8660i
  -0.5000 - 0.8660i
   1.0000 + 0.0000i

>> abs(eig(A))

ans =

    1.0000
    1.0000
    1.0000
```

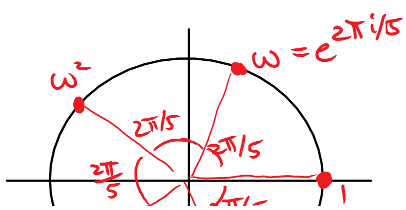
Exercise: Similarly, what are the eigenvalues of

$$A_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} ?$$

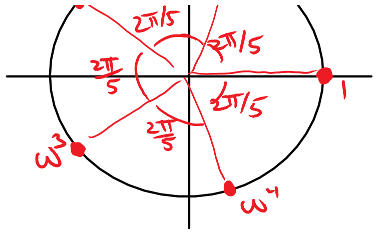
Answer: $A^5 = I$, so e-values must lie in:

$\{1, \omega, \omega^2, \omega^3, \omega^4\}$

where $\omega = e^{2\pi i/5}$.



Since the dimension is odd



Since the dimension is odd and the complex e-values come in pairs, 1 is an e-value.

$$\mathrm{Tr}(A)=0 \Rightarrow \boxed{\{1, \omega, \omega^4, \omega^2, \omega^3\}}$$

Problem: If all the eigenvalues of A have the same magnitude, then how can we use the power method to find them?

$$\vec{x} = (v_1\cdot x)\vec{v}_1 + (v_2\cdot x)\vec{v}_2 + \cdots + (v_n\cdot x)\vec{v}_n$$

$$\Rightarrow A\vec{x} = e^{i\theta_1}(v_1\cdot x)\vec{v}_1 + e^{i\theta_2}(v_2\cdot x)\vec{v}_2 + \cdots + e^{i\theta_n}(v_n\cdot x)\vec{v}_n$$

$$\Rightarrow A^k\vec{x} = e^{ik\theta_1}(v_1\cdot x)\vec{v}_1 + e^{ik\theta_2}(v_2\cdot x)\vec{v}_2 + \cdots + e^{ik\theta_n}(v_n\cdot x)\vec{v}_n$$

The magnitudes of the coefficients don't change

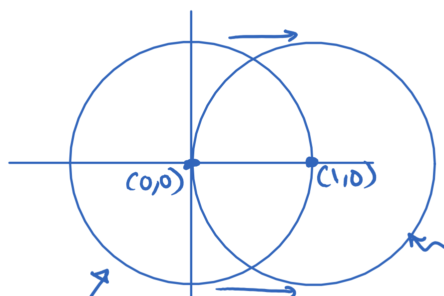
$$|e^{ik\theta_j}(v_j\cdot x)| = |v_j\cdot x|$$

so no term dominates!

Possible answer:

Shift the matrix!

E.g., work with $B = A + I$ instead of with A; B's eigenvalues are $e^{i\theta_j}+1$, no longer all lying on the unit circle.



II. Eigenvalues of real-orthogonal and Hermitian matrices

Definition: Matrix A is

symmetric	if	$A^T = A$	← useful for real matrices
Hermitian	if	$A^\dagger = A$	← useful for complex matrices

symmetric	if	$A^T = A$	← useful for real matrices
Hermitian	if	$A^\dagger = A$	← useful for complex matrices

adjoint = conjugate-transpose

(For real matrices, symmetric = Hermitian.)

Examples:

	Symmetric?	Hermitian?
$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$	✓	✓
$\begin{pmatrix} 1 & i \\ i & 0 \end{pmatrix}$	✓	✗
$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & 0 \end{pmatrix}$	✓	only if $\frac{\theta}{\pi} \in \mathbb{Z}$
$\begin{pmatrix} 1 & a+bi \\ a-bi & 2 \end{pmatrix}$	✗	✓

Example: $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$

basis for 2×2 symmetric matrices

$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \right\}$

basis over $\mathbb{R}$ for 2×2 Hermitian matrices

Note: $\{n \times n$ Hermitian matrices$\}$ is only a vector space over $\mathbb{R}$.
It is not closed under multiplication by complex numbers.

E.g., $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is Hermitian,

$\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = i \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is not.

Theorem: Any real symmetric matrix is diagonalizable.

Proof: Let A be a real symmetric matrix: $A^T = A$.

Since A is real, $A^\dagger = A^T$.

$\Rightarrow A^T A = A^2 = A A^T$

$\Rightarrow A$ is normal $\Rightarrow$ diagonalizable $\square$

Theorem: Any complex Hermitian matrix is diagonalizable.
$(A^\dagger = A)$

Same proof.

Claim: Any real symmetric or complex Hermitian matrix has all real eigenvalues.

Proof: Let $\vec{v}$ be an eigenvector of A, with e.v. λ.

$$v^\dagger A v$$

$$(v^\dagger A) v \quad = \quad v^\dagger (A v)$$

$$= (A^\dagger v)^\dagger v \qquad\qquad = \lambda \cdot v^\dagger v = \lambda \|v\|^2$$

$$= (A v)^\dagger v$$

$$= (\lambda v)^\dagger v = \lambda^* \|v\|^2 \qquad \Rightarrow \lambda = \lambda^*$$

i.e., λ is real. $\square$

Furthermore, a real symmetric matrix has an orthonormal basis of real eigenvectors.

Proof: All the eigenvalues are real.

For each eigenvalue λ, we can compute an orthonormal basis for the eigenspace

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

using Gaussian elimination and the Gram-Schmidt procedure, using only real numbers. ✓

Examples:

-
```
>> n = 10;
A = -2 * diag(ones(n,1)) + diag(ones(n-1,1),1) + diag(ones(n-1,1),-1)

A =

    -2     1     0     0     0     0     0     0     0     0
     1    -2     1     0     0     0     0     0     0     0
     0     1    -2     1     0     0     0     0     0     0
     0     0     1    -2     1     0     0     0     0     0
     0     0     0     1    -2     1     0     0     0     0
     0     0     0     0     1    -2     1     0     0     0
     0     0     0     0     0     1    -2     1     0     0
     0     0     0     0     0     0     1    -2     1     0
     0     0     0     0     0     0     0     1    -2     1
     0     0     0     0     0     0     0     0     1    -2

>> eig(A)

ans =

   -3.9190
   -3.6825
```

symmetric $A = A^T$

```
-3.9190
-3.6825
-3.3097
-2.8308
-2.2846
-1.7154
-1.1692
-0.6903
-0.3175
-0.0810
```

all real

- How to create a random symmetric matrix?

  Of course it depends on what distribution you want.
  If you don't care, then $A+A^T$ is symmetric for any real matrix A. (Since $(A+A^T)^T = A^T + A = A + A^T$ ✓)
  If you want every entry to follow the same distⁿ, eg., uniform on $[0,1]$, use something like

```
>> n = 6;
A = rand(n,n)
A = diag(diag(A)) + triu(A,1) + triu(A,1)'
eig(A)
```

picks out diagonal; terms above diagonal

```
A =

    0.4211    0.5710    0.3736    0.3474    0.8006    0.5301
    0.1841    0.1769    0.0875    0.6606    0.7458    0.2751
    0.7258    0.9574    0.6401    0.3839    0.8131    0.2486
    0.3704    0.2653    0.1806    0.6273    0.3833    0.4516
    0.8416    0.9246    0.0451    0.0216    0.6173    0.2277
    0.7342    0.2238    0.7232    0.9106    0.5755    0.8044


A =

    0.4211    0.5710    0.3736    0.3474    0.8006    0.5301
    0.5710    0.1769    0.0875    0.6606    0.7458    0.2751
    0.3736    0.0875    0.6401    0.3839    0.8131    0.2486
    0.3474    0.6606    0.3839    0.6273    0.3833    0.4516
    0.8006    0.7458    0.8131    0.3833    0.6173    0.2277
    0.5301    0.2751    0.2486    0.4516    0.2277    0.8044


ans =

   -0.6673
   -0.3633
    0.2743
    0.4173
    0.7283
    2.8979
```

- What about an anti-symmetric matrix $(A^T = -A)$?

  $A - A^T$ is always antisymmetric

(since $(A-A^T)^T = A^T - A = -(A-A^T)$. ✓)

```
>> n = 6;
A = rand(n,n);
A = triu(A,1) - triu(A,1)'
eig(A)

A =

         0    0.0669    0.8854    0.0418    0.9843    0.6624
   -0.0669         0    0.8990    0.1069    0.9456    0.2442
   -0.8854   -0.8990         0    0.6164    0.6766    0.2955
   -0.0418   -0.1069   -0.6164         0    0.9883    0.6802
   -0.9843   -0.9456   -0.6766   -0.9883         0    0.5278
   -0.6624   -0.2442   -0.2955   -0.6802   -0.5278         0


ans =

   0.0000 + 2.3224i
   0.0000 - 2.3224i
   0.0000 + 1.0839i
   0.0000 - 1.0839i
  -0.0000 + 0.2142i
  -0.0000 - 0.2142i
```

Question: Why are the eigenvalues purely imaginary?

Answer:

If A is a real matrix with $A^T = -A$,

$\Rightarrow$ $(iA)^+ = -i A^T$

$= iA$

$\Rightarrow$ iA is Hermitian

$\Rightarrow$ Eigenvalues of iA are real

$\Rightarrow$ Eigenvalues of A are purely imaginary. ✓

II. Positive Semi-Definite Matrices

Definition: A Hermitian (or real-symmetric) matrix A with all eigenvalues ≥ 0 is called

$\boxed{\text{positive semi-definite}}$

denoted $A \succeq 0$.

Key property:

Theorem: A Hermitian matrix A is positive semidefinite

$$\Longleftrightarrow$$

$x^\dagger A x \geq 0$ for all vectors x.

Proof:

$\Rightarrow$: If $A \succeq 0$,

let $v_1, \ldots, v_n$ be an orthonormal basis of eigenvectors, with corresponding e-values $\lambda_1, \ldots, \lambda_n \geq 0$.

Any vector x can be expanded as

$$\vec{x} = \sum_{j=1}^{n} \alpha_j \vec{v}_j$$

where $\alpha_j = v_j \cdot \vec{x}$.

$$\Rightarrow x^\dagger A x = \sum_{i,j} \alpha_i^* \underbrace{v_i^\dagger A \vec{v}_j}_{\lambda_j v_i^\dagger v_j} \alpha_j$$

$$\lambda_j v_i^\dagger v_j = \begin{cases} \lambda_j, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

$$= \sum_{j=1}^{n} |\alpha_j|^2 \cdot \lambda_j$$

$$\geq 0 \quad ✓$$

$\Leftarrow$: If $A \not\succeq 0$, i.e., some eigenvalue $\lambda < 0$, let $\vec{x}$ be a corresponding e-vector.

$$\Rightarrow x^\dagger A x = \lambda x^\dagger x$$

$$= \lambda \|x\|^2$$

$$< 0. \quad ✓ \qquad \square$$

Example 1: For any real matrix A, $A^T A \succeq 0$.

Proof: $x^T A^T A x = (Ax)^T (Ax)$

$$= \|Ax\|^2$$

$$\geq 0$$

$\Rightarrow$ by the theorem, $A^T A \succeq 0$. ✓ $\square$

Singular-value decomposition vs. Spectral decomposition

$A = U S V^\dagger$ — S: diagonal; U, V: unitary

$A = T D T^{-1}$ — D: diagonal

What is it?

$A = U \Sigma V^\dagger$ — Σ: diagonal; U, V: unitary

$= \sum_i \sigma_i \vec{u}_i \vec{v}_i^\dagger$

σ_i: sing. values ≥ 0; $\vec{u}_i$: left, $\vec{v}_i$: right singular vectors; $\{u_i\}$, $\{v_i\}$ orthonormal

$A = T D T^{-1}$ — D: diagonal

if A is normal:

$A = T D T^\dagger$ — same unitary!

$= \sum_i \lambda_i \vec{t}_i \vec{t}_i^\dagger$

λ_i: e-values; $\vec{t}_i$: e-vectors (orthonormal)

When?

all matrices!

square, diagonalizable matrices

$\mathbb{R}$ or $\mathbb{C}$?

real if A is real

complex even if A is real

e.g. $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \to e^{\pm i\theta}$

real if A is real & symmetric

Other properties

$\|A\| = \max \text{sing. value}(A)$

$A \to A + I$ does not shift sing. values in general, e.g. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

svs 1, 1; evs 1, −1 → svs 2, 0; evs 2, 0

$\|A\| \neq \max |\text{eigenvalue}|$ in general

e.g. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has evalues 0, 0

$A \to A + \alpha I$ shifts e-values by α

Example:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \underbrace{1}_{\sigma_1} \cdot \underbrace{\begin{pmatrix} c \\ s \end{pmatrix}}_{\vec{u}_1} \underbrace{\begin{pmatrix} 1 & 0 \end{pmatrix}}_{\vec{v}_1^\dagger} + \underbrace{1}_{\sigma_2} \cdot \underbrace{\begin{pmatrix} -s \\ c \end{pmatrix}}_{\vec{u}_2} \underbrace{\begin{pmatrix} 0 & 1 \end{pmatrix}}_{\vec{v}_2^\dagger} = \underbrace{e^{-i\theta}}_{\lambda_1} \cdot \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}}_{\vec{v}_1} \underbrace{\frac{\begin{pmatrix} 1 & -i \end{pmatrix}}{\sqrt{2}}}_{\vec{v}_1^\dagger} + \underbrace{e^{i\theta}}_{\lambda_2} \cdot \underbrace{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}}_{\vec{v}_2} \underbrace{\frac{\begin{pmatrix} 1 & i \end{pmatrix}}{\sqrt{2}}}_{\vec{v}_2^\dagger}$$

They can't be entirely different. What's the connection?

Recall: For any $m \times n$ matrix A,

$$A^\dagger A \succeq 0.$$

In particular, $A^\dagger A$ is Hermitian (and thus normal), and its eigenvalues are all real and ≥ 0.

Proposition: Let A be an $m \times n$ real or complex matrix.

Proposition: Let A be an $m \times n$ real or complex matrix.

Then,

- The nonzero singular values of A are the square-roots of the eigenvalues of $A^\dagger A$ (same as the square-roots of the e-values of $AA^\dagger$).
- Eigenvectors of $A^\dagger A$ are right singular vectors of A.
- Eigenvectors of $AA^\dagger$ are left singular vectors of A.

Corollary: $AA^\dagger$ and $A^\dagger A$ have the same nonzero eigenvalues.

Corollary: How to compute the SVD of a matrix?

Answer: Prove this proposition!

Proof:

Let $A = \sum_i \sigma_i \vec{u}_i \vec{v}_i^T$ be the SVD of A.

(We don't know what it is, maybe, but it exists!)

$$\Rightarrow A^T A = \Big(\sum_i \sigma_i \vec{v}_i \vec{u}_i^T\Big)\Big(\sum_j \sigma_j \vec{u}_j \vec{v}_j^T\Big)$$

$$= \sum_i \sigma_i^2 \vec{v}_i \vec{v}_i^T \qquad \text{since } u_i^T u_j = u_i \cdot u_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\Rightarrow A^T A \vec{v}_j = \sigma_j^2 \vec{v}_j$$

i.e., the right singular vector $\vec{v}_j$ is an e-vector of $A^T A$, with e-value σ_j^2. ✓

$\Rightarrow$ By diagonalizing $A^T A$, we now know the singular values σ_i and right singular vectors $\vec{v}_i$.

To get $\vec{u}_i$, note $\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i$. ✓ $\square$

Corollary: For any A,

$$\|A\| = \sqrt{\text{max e-value of } A^T A}$$

$$= \sqrt{\text{max e-value of } AA^T}.$$

For normal matrices, the relationship between singular values and eigenvalues is even simpler.

A normal

$$\Rightarrow \quad A = \underset{\text{unitary}}{U} \underset{\text{diag. w/ e-values}}{D} U^\dagger$$

$$= \sum_i \lambda_i \vec{u}_i \vec{u}_i^\dagger$$

e-vectors (columns of U)

$$= \sum_i |\lambda_i| \cdot \left(\frac{\lambda_i}{|\lambda_i|} \vec{u}_i\right) \cdot \vec{u}_i^\dagger$$

This is an SVD for A! sing. values left/right sing. vectors

$\Downarrow$

Claim: If A is normal, then its singular values = absolute values of eigenvalues.

Corollary: If A is normal,

$$\|A\| = \max|\text{eigenvalue of } A|.$$

(Again, this is generally false for non-normal matrices.)

Example: If A is orthogonal or unitary, all its singular values are 1.

(Of course, we already knew this; an SVD for A is $A = \sum_i (Ae_i) e_i^T$, since the set $\{Ae_i\}$ is orthonormal.)