Lecture 22: Special matrices

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

Admin: Midterm 2 on Thursday
open notes, Matlab calculator
no internet

Last time:

THEOREM: A has a complete, orthogonal set of eigenvectors

$$\Updownarrow$$

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

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

Jordan decomposition

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \end{pmatrix}$$

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 =
      v1        v2        v3        v4
    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 =
      λ1        λ2        λ3        λ4
   -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 *

    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
```

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

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 $\Rightarrow \|A\| = \max\limits_{\text{eigenvalues } \lambda} |\lambda|$.

Proof:

A normal $\Rightarrow A = UDU^{-1}$ with U unitary

$$\|A\| = \|UDU^{-1}\| = \|D\| = \max|\lambda|. \checkmark$$

This proposition is **false** for non-normal matrices,

e.g., both eigenvalues of $\left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\| = 1$

e.g., both eigenvalues of $\left\|\begin{pmatrix}0&1\\0&0\end{pmatrix}\right\| = 1$

$\left\|\begin{pmatrix}1&1\\0&1\end{pmatrix}\right\| > \sqrt{2}$

Key point:

Normal matrix $\Rightarrow$ Different eigenspaces are orthogonal

Because the nullspace (e-value 0 e-space) satisfies

$N(A) = N(A^\dagger) = R(A)^\perp$

and all other e-spaces are in $R(A)$

Power method:

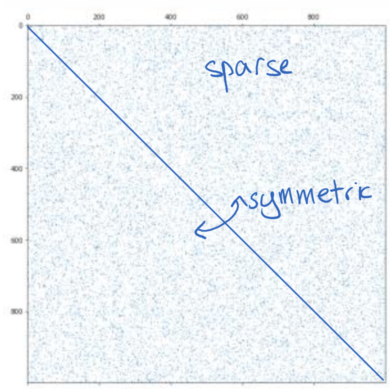
```
import numpy as np
from scipy import sparse

n = 1000
s = .01
m = int(s * n * n)

rows = np.random.randint(n, size=m)
cols = np.random.randint(n, size=m)
data = np.random.randn(m)

A = sparse.coo_matrix((data, (rows,cols)), shape=(n,n))
A = A + A.T
A = A.tocsr()

import matplotlib.pyplot as plt
plt.figure(figsize=(10,10))
plt.spy(A, markersize=.3)
plt.show()
```



for Hermitian matrices only

```
from scipy.sparse.linalg import eigsh

k = 6

eigsh(A, k)[0]
```

k largest magnitude e-values

```
array([-9.59771023, -9.57258027, -9.40408838,  9.47442673,  9.6965802 ,
        9.80762864])
```

```
x = np.random.randn(n, k)
x, _ = np.linalg.qr(x)
trials = 20000
for t in range(trials):
  x = A.dot(x)
  x, _ = np.linalg.qr(x)
ratios = A.dot(x) / x
(ratios.T)[:,:4]
```

```
array([[ 9.80762864,  9.80762864,  9.80762864,  9.80762864],
       [ 9.6965802 ,  9.6965802 ,  9.6965802 ,  9.6965802 ],
       [-9.59771023, -9.59771023, -9.59771023, -9.59771023],
       [-9.57258027, -9.57258027, -9.57258027, -9.57258027],
       [ 9.47442673,  9.47442673,  9.47442673,  9.47442673],
       [-9.40408838, -9.40408838, -9.40408838, -9.40408838]])
```

symmetric $A = A^T \Rightarrow$ normal

Today: More examples and applications.

APPLICATIONS of unitary diagonalizability of normal matrices

I. Eigenvalues of orthogonal and unitary matrices

Isometry

$V = $ ($n \times k$)

$V^\dagger V = I$

$VV^\dagger = P_{R(V)}$, $\dim = k$

rank k $n\times n$ matrix

$\mathrm{Tr}(VV^\dagger) = \mathrm{Tr}(V^\dagger V) = k$

Orthogonal / Unitary

$U = $ ($n \times n$) square isometry

$U^\dagger U = I$

$UU^\dagger = I$

Recall:
Definition:

rank k

$\mathrm{Tr}(VV^\dagger) = \mathrm{Tr}(V^\dagger V) = k$

Definition:

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

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

$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ an orthonormal basis for $\mathbb{R}^n$ or $\mathbb{C}^n$ $\iff$ $A = \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{pmatrix}$ 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: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$, $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}$

$\omega = e^{2\pi i/3}$

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: Let $\vec{v}$ be an e-vector of unitary U. e-value λ

$$\|\vec{v}\|^2 = v \cdot v = v^\dagger v = v^\dagger I v = v^\dagger U^\dagger U v = \|Uv\|^2 = \|\lambda v\|^2 = |\lambda|^2 \|v\|^2$$

$\Rightarrow |\lambda| = 1.$ ✓

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
```

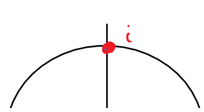
$$C = \frac{1}{\sqrt{3}} \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 1 & 1 & 1 \\ 1 & 1 & \omega & \omega^2 \\ 2 & 1 & \omega^2 & \omega \end{array}$$

$\omega = \omega^4$

$\omega = e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$

$\omega^3 = 1$

$\mathrm{Tr}\, C = \frac{\cancel{\sqrt{3}}\, i}{\cancel{\sqrt{3}}} = i$

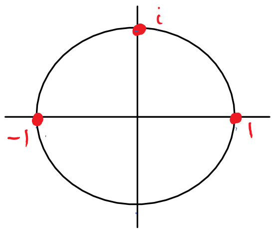


```
 1.0000 + 0.0000i
-1.0000 + 0.0000i
 0.0000 + 1.0000i

>> abs(eig(C))

ans =

    1.0000
    1.0000
    1.0000
```



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

$\Rightarrow A^k \vec{v} = A^{k-1} A\vec{v} = \lambda A^{k-1}\vec{v}$

$= \cdots = \lambda^k \vec{x}$

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 =
```

$$\begin{pmatrix} 0.7071 & 0 & 0.7071 \\ 0 & 1.0000 & 0 \\ -0.7071 & 0 & 0.7071 \end{pmatrix}$$

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

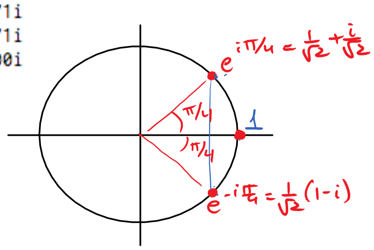
$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \sim e^{\pm i \frac{\pi}{4}}$

```
ans =

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

ans =

    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: If $A\vec{v} = \lambda\vec{v} \Rightarrow A^* x^* = \lambda^* v^*$

$\Rightarrow A x^* = \lambda^* v^*$ ✓

$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ 1, 1, −1

$C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Exercise: What are the eigenvalues of

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

~~symmetric~~ ~~diagonal~~

permutation $1 \to 3 \to 2$

isometry

unitary $\Rightarrow |\lambda| = 1$

$\Downarrow$

normal (unitarily diagonalizable)

Answer: $\mathrm{Tr} A = 0 = \lambda_1 + \lambda_2 + \lambda_3$

$A^3 = I$

$\Downarrow$

row-stochastic

$A^3 = I \Longleftarrow$ row-stochastic $\Downarrow$ $A\begin{pmatrix}1\\1\\1\end{pmatrix} = \boxed{1}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}$

$\Downarrow$ normal (unitarily diagonalizable)

$\Downarrow$

$\lambda^3 = 1$

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

real $\Rightarrow$ evs in c.c. pairs λ, λ^*

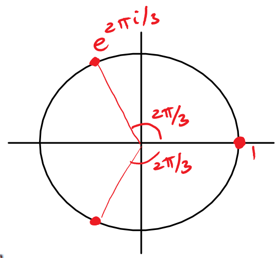
$A \neq I \Rightarrow$ not all e-values $=1$

$\Rightarrow$ E-values	E-vectors
1	$(1,1,1)$
$\omega = e^{2\pi i/3}$	$(1,\omega,\omega^2)$
$\omega^2 = \omega^{-1}$	$(1,\omega^2,\omega)$

$A\begin{pmatrix}1\\a\\b\end{pmatrix} = \begin{pmatrix}a\\b\\1\end{pmatrix} = \omega \begin{pmatrix}a\\b\end{pmatrix}$

$\Rightarrow a = \omega, b = \omega^2$

(**Remark**: These are Fourier vectors)



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: What are the eigenvalues of

$$A = \left(\begin{array}{c|c} 0 & \begin{smallmatrix} 0 & & & 1 \\ & 1 & 1 & \\ & & & \\ 1 & & & 0\end{smallmatrix} \\ \hline \begin{smallmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{smallmatrix} & 0 \end{array}\right)_{8\times 8} \; ? \quad B = \begin{array}{c|c} & \begin{matrix} 1 & 4 & 5 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 4 \\ 5 \\ 8 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}_{4\times 4} \end{array}$$

$1 \to 5 \to 4 \to 8$

permutation $A^4 = I$

$\operatorname{Tr} A = 0 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$

unitary $\quad \lambda \in \{1, i, -i, -1\}$

A: E-values	E-vectors
1	$(1,0,0,1,1,0,0,1)$
1	$(0,1,1,0,0,1,1,0)$
i	$(1\ 0\ 0\ i^2\ i^3\ 0\ 0\ i)$
i	$(0\ 1\ i^2\ 0\ 0\ i^3\ i\ 0)$

$\operatorname{Tr} A = 0 = \lambda_1+\lambda_2+\lambda_3+\lambda_4$

unitary $\qquad \lambda \in \{1, i, -i, -1\}$

real (λ & λ^*)

$A^2 \neq I \Rightarrow$ some e-value has to be i or $-i$

$i \quad i \quad -i \quad -i \quad -1 \quad -1$

$(0\ 1\ i^2\ 0\ 0\ i^3\ i\ 0)$

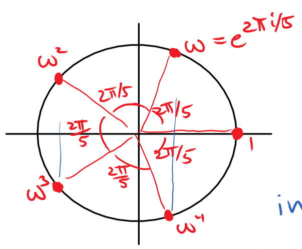
	E-values	E-vectors
B:	1	(1, 1, 1, 1)
	i	$(1, i^2, i^3, i)$
	$-i$	$(1, (-i)^2, (-i)^3, -i)$
	-1	$(1, 1, -1, -1)$

$$B\begin{pmatrix}1\\a\\b\\c\end{pmatrix} = \begin{pmatrix}c\\1\\a\\b\end{pmatrix} = \lambda\begin{pmatrix}1\\a\\b\\c\end{pmatrix} \Rightarrow \begin{pmatrix}1\\a\\b\\c\end{pmatrix} = \begin{pmatrix}1\\\lambda^2\\\lambda^3\\\lambda\end{pmatrix}$$

Exercise: Similarly, what are the eigenvalues of

$$A = \begin{pmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\1&0&0&0&0\end{pmatrix} ? \qquad A^5 = I$$

Answer:



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

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

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

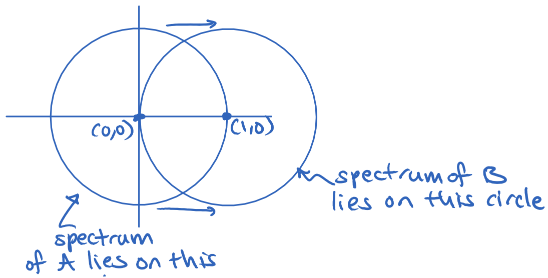
$$|e^{ik\theta_j}(v_j\cdot x)| = |v_j\cdot x|$$

so no term dominates!

Possible answer:

Shift the matrix!

Eg., 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.



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

$A^\dagger$: 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 $\frac{\theta}{\pi} \in \mathbb{Z}$
$\begin{pmatrix} 1 & a+bi \\ a-bi & 2 \end{pmatrix}$	✗	✓

$$\begin{pmatrix} a & d+ei & f+gi \\ d-ei & b & h+ji \\ f-gi & h-ji & c \end{pmatrix}_{3\times 3}$$

$a, b, c, d, e, f, g, h, j \in \mathbb{R}$

$$iI = \begin{pmatrix} i & & 0 \\ & i & \\ 0 & & i \end{pmatrix}$$

↑ not Hermitian

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:

$$AA^T = A^2 = A^TA \quad \text{normal.} ✓$$

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

Same proof. $$AA^\dagger = A^2 = A^\dagger A \quad \text{normal} ✓$$

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.

$$\underbrace{v^\dagger A v}_{} = v \cdot (\lambda v) = \lambda \|v\|^2$$

$$\underbrace{v^\dagger A v}_{} = v \cdot (\lambda v) = \lambda \|v\|^2$$

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

$$\Rightarrow \lambda = \lambda^* \quad ✓$$

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

Proof: E-space for e-value λ $= N(A - \lambda I)$ ✓ (A: real, λ: real)

GE. only needs 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
     -3.3097
     -2.8308
     -2.2846
     -1.7154
     -1.1692
     -0.6903
     -0.3175
     -0.0810
  ```

  symmetric $A = A^T$

  all real; sum is -4 why?

$B = A + 2I$

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & & \\ 2 & 1 & 0 & 1 & \\ 3 & & 1 & 0 & 1 \\ 4 & & & 1 & 0 \end{array}$$

e-values are paired as $\lambda, -\lambda$

$$\begin{array}{c|cc|cc} & 1 & 3 & 2 & 4 \\ \hline 1 & \multicolumn{2}{c|}{\multirow{2}{*}{0}} & 1 & 0 \\ 3 & & & 1 & 1 \\ \hline 2 & 1 & 1 & \multicolumn{2}{c}{\multirow{2}{*}{0}} \\ 4 & 0 & 1 & & \end{array} = PBP^{-1}$$

bipartite

Def.: A matrix is bipartite if up to a permutation of the rows and columns (the same permutation!), it breaks up as

$$A = \left(\begin{array}{c|c} 0 & B \\ \hline C & 0 \end{array}\right)$$

Claim: If A is bipartite, then λ e-value $\Rightarrow -\lambda$ e-value.

Proof:

$$A\begin{pmatrix} v \\ w \end{pmatrix} = \lambda \begin{pmatrix} v \\ w \end{pmatrix}$$

$$\begin{pmatrix} Bw \\ Cv \end{pmatrix} \Rightarrow Bw = \lambda v, \quad Cv = \lambda w$$

$$\Rightarrow A\begin{pmatrix} v \\ -w \end{pmatrix} = \begin{pmatrix} -Bw \\ Cv \end{pmatrix} = \begin{pmatrix} -\lambda v \\ \lambda w \end{pmatrix} = -\lambda \begin{pmatrix} v \\ -w \end{pmatrix} \quad \square$$

- 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, e.g., 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)

  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
  ```

  diag(diag(A)): picks out diagonal; triu(A,1): terms above diagonal
  ```

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

picks out diagonal — terms above diagonal

$$\begin{pmatrix} Bw \\ Cv \end{pmatrix} \Rightarrow Bw = \lambda v, \quad Cv = \lambda w$$

$$\Rightarrow A\begin{pmatrix} v \\ -w \end{pmatrix} = \begin{pmatrix} -Bw \\ Cv \end{pmatrix} = \begin{pmatrix} -\lambda v \\ \lambda w \end{pmatrix} = -\lambda \begin{pmatrix} v \\ -w \end{pmatrix} \quad \square$$

```
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 real 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
```

$B = iA$ is Hermitian

$B^\dagger = -iA^T = iA = B$

Question: Why are the eigenvalues purely imaginary?

Answer:

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 $\iff$ $x^\dagger A x \geq 0$ 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 $\alpha_j = v_j \cdot \vec{x}$.

$$\Rightarrow x^\dagger A x = \sum_{i,j} \alpha_i^* \underbrace{v_i^\dagger A \vec{v}_j}_{\lambda_j v_i^\dagger v_j \,=\, \begin{cases} \lambda_j, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}} \alpha_j$$

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

$$\geq 0 \quad ✓$$

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

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

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

$$< 0 . \quad ✓$$

□

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, $A^T A \succeq 0$. ✓ □

Singular-value decomposition vs. Spectral decomposition

What is it?

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

$= \sum_j \sigma_j \vec{v}_j \vec{u}_j^\dagger$

sing. values $\sigma_j \geq 0$; $\vec{v}_j$: left, $\vec{u}_j$: right singular vectors

$A = T D T^{-1}$ (D: diagonal)

if A is normal:

$A = T D T^\dagger$ (same unitary!)

What

$\sum_j \sigma_j v_j u_j$ — sing. values ≥ 0; left, right singular vectors; $\{u_i\}, \{v_i\}$ orthonormal

$A = T D T^\dagger$ — same unitary!

$= \sum_j \lambda_j \vec{t}_j \vec{t}_j^\dagger$ — e-values; e-vectors (orthonormal)

$\vec{t}_k \to \lambda_k \vec{t}_k$

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 → svs 2, 0

evs 1, −1 → 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:

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

unitary

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

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

$A^\dagger A \succeq 0$. Pf: $(A^\dagger A)^\dagger = A^\dagger A^{\dagger\dagger} = A^\dagger A$ ✓

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.

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.

$A = \sum_j \sigma_j v_j u_j^\dagger$

$A A^\dagger = \sum_j \sigma_j^2 v_j v_j^\dagger$

$(AA^\dagger)\vec{v}_j = \sigma_j^2 \vec{v}_j$, $\sigma_j^2 \geq 0$

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!)

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

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

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

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,

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

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

$$A \text{ normal} \Rightarrow A = \underset{\text{unitary}}{U} D U^{-1}$$

$$\|A\| = \|U D U^{-1}\| = \|D\| = \max_j |\lambda_j|$$

spectral decomposition

$$= \sum_j \lambda_j \vec{t}_j \vec{t}_j^{\,\dagger} = \sum_j \underbrace{|\lambda_j|}_{\sigma_j \text{ sing. value}} \cdot \underbrace{\left(\frac{\lambda_j}{|\lambda_j|} \vec{t}_j\right)}_{\text{left s.v.}} \cdot \underbrace{\vec{t}_j^{\,\dagger}}_{\text{right s.v}} \quad \leftarrow \text{SVD!}$$

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.

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.)

Diagonalizable

in basis of e-vectors
$$\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \lambda_3 \\ 0 & & & \ddots \end{pmatrix}$$

$$\boxed{\text{rank} = \#\text{ nonzero e-values}}$$

Non-diagonalizable

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{all e-values } 0, 0$$

rank = 1

$$\boxed{\underset{n\times n}{\text{rank}(A)} = n - \underbrace{\dim N(A)}_{\text{multiplicity of e-value } 0} \geq \#\text{ nonzero e-values}}$$

Application: Systems of differential equations

Warmup:

$$\frac{d}{dt}x(t) = x(t) \longrightarrow x(t) = \underset{\substack{\shortparallel \\ x(t=0)}}{C}e^t$$

What about

$$\dot{x} = 4x - 5y$$
$$\dot{y} = 2x - 3y$$

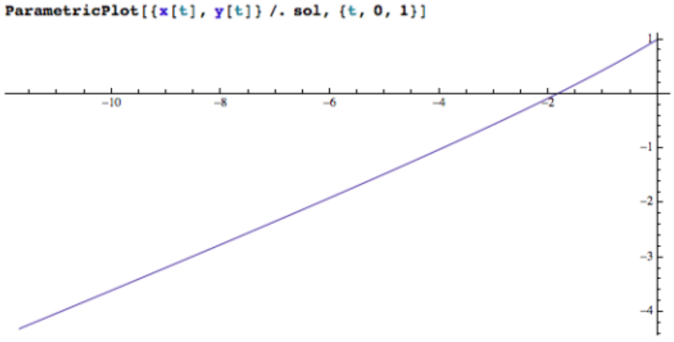
Numerical solution (Mathematica):

```
In[2]:= sol = NDSolve[{
          D[x[t], t] == 4 x[t] - 5 y[t],
          D[y[t], t] == 2 x[t] - 3 y[t],
          x[0] == 0,
          y[0] == 1
        }, {x, y}, {t, 0, 1}];
        ParametricPlot[{x[t], y[t]} /. sol, {t, 0, 1}]
```

```
In[1]:= ?NDSolve
```

NDSolve[*eqns*, *u*, {*t*, t_{min}, t_{max}}] finds a numerical solution to the ordinary differential equations *eqns* for the function *u* with the independent variable *t* in the range t_{min} to t_{max}.
NDSolve[*eqns*, *u*, {*t*, t_{min}, t_{max}}, {*x*, x_{min}, x_{max}}] finds a numerical solution to the partial differential equations *eqns*.
NDSolve[*eqns*, {u_1, u_2, ...}, {*t*, t_{min}, t_{max}}] finds numerical solutions for the functions u_i. »

Out[3]=



Exact solution:

"Rule": When you see multiple equations, try to vectorize them!

$\dot{x} = 4x - 5y$
$\dot{y} = 2x - 3y$

Problem: Solve

$$\frac{d}{dt}\vec{v}(t) = A\,\vec{v}(t)$$

subject to $\vec{v}(t=0) = \vec{v}_0$ initial conditions

Answer:

① In terms of eigenvalues/eigenvectors:

- Say $A\vec{u}_1 = \lambda_1\vec{u}_1$
  and $\vec{v}_0 = \vec{u}_1$

- Say $A\vec{u}_1 = \lambda_1 u_1$
  $A\vec{u}_2 = \lambda_2\vec{u}_2$
  and $\vec{v}_0 = \vec{u}_1 - 2\vec{u}_2$

- If A is diagonalizable, then its eigenvectors form a basis,

If A can be diagonalized,

$A = UDU^{-1}$

$\Rightarrow \frac{d}{dt}v = UDU^{-1}v$

The eigenvalues of A, λ_1 and λ_2, determine the exponents (growth rates).

We can actually simplify even further:

$$v(t) = \mathcal{U} \cdot \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} u(0)$$

$$= \mathcal{U} \cdot \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \mathcal{U}^{-1} v(0)$$

2 In terms of matrices:

$$v(t) = e^{At} \vec{v}_0$$

$$e^{At} \equiv I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots$$

$$= \sum_{j=0}^{\infty} \frac{1}{j!} (At)^j$$

Proof.

Functions of diagonalizable matrices:

Definition: If A is a diagonalizable matrix

$$A = \mathcal{U} \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \mathcal{U}^{-1}$$

and $f: \mathbb{C} \to \mathbb{C}$ any function, define

$$f(A) = \mathcal{U} \begin{pmatrix} f(\lambda_1) & & & 0 \\ & f(\lambda_2) & & \\ & & \ddots & \\ 0 & & & f(\lambda_n) \end{pmatrix} \mathcal{U}^{-1}.$$

Example: $f(x) = x^2$

$\Rightarrow f(A) = U D^2 U^{-1} = A^2$, as you'd expect ✓

Example: $f(x) = e^{tx}$ exponential

$$\Rightarrow f(A) = U \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} U^{-1}$$

So,
$$v(t) = \underbrace{U \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} U^{-1}}_{e^{At}} v(0)$$

$\Rightarrow$ The solution to
$$\frac{d}{dt} v(t) = A\, v(t)$$
is $v(t) = e^{At} v(0)$.

This looks just like the single-variable case.

You can also show this using Taylor series:

$$e^{xt} = 1 + xt + \frac{(xt)^2}{2} + \cdots$$
$$= \sum_{j=0}^{\infty} \frac{(xt)^j}{j!}$$

and $e^{At} = \sum_{j=0}^{\infty} \frac{(At)^j}{j!}$ where $A^0 = I$ identity matrix.

Extension: Higher-order differential equations with constant coefficients.

To solve a 2nd-order diff. eq. like

$$\ddot{x} = -\dot{y} + 2x$$
$$\ddot{y} = \dot{x} + \dot{y} - 3y$$