# Lecture 23: Special matrices

Thursday, November 12, 2015 9:30 AM

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

**Last time:**

**THEOREM:** $A$ has a complete, orthogonal set of eigenvectors

$$\Updownarrow$$

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

$$\Rightarrow A = U D U^\dagger$$

$D$: diagonal w/ e-values

$U$: unitary w/ e-vector columns

**Examples:**

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

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

U =
    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 =
   -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 $\Rightarrow \|A\| = \max_{\text{eigenvalues } \lambda} |\lambda|$.

Proof:

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

$$\|A\| = \|UDU^{-1}\| = \|D\| = \max_i |D_{i,i}| \checkmark$$ □

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

## More about the power method

```
>> n = 10^6;
>> density = 10/n;
>> A = sprandsym(n, density);
>> tic; eigs(A); toc
Elapsed time is 192.303369 seconds.
>> tic; eigs(A, 1); toc
Elapsed time is 80.018051 seconds.
```

find all eigenvalues

find largest magnitude eigenvalue

```
>> help sprandsym
 sprandsym Sparse random symmetric matrix.
    R = sprandsym(S) is a symmetric random matrix whose lower triangle
        and diagonal have the same structure as S.  The elements are
        normally distributed, with mean 0 and variance 1.

    R = sprandsym(n,density) is a symmetric random, n-by-n, sparse
        matrix with approximately density*n*n nonzeros; each entry is
        the sum of one or more normally distributed random samples.
```

`eigs(A, 1, 'SM')` finds the smallest magnitude e-value

**Exercise:** Write Matlab code to find the e-value closest to 10.

Answer 1: `eigs(A-10*eye(n), 1, 'SM')`

Answer 2: Using the power method:

```
B = A - 10*eye(n);
x = randn(n, 1);
for j = 1:10000
    x = B \ x;
    x = x/norm(x);
end
```

← it would be better to check for convergence...

← compute $B^{-1}x$

it would be faster to precompute the LU decomposition of $B$

**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** $\quad A^T = A^{-1}$

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

$\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ 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 ✓ □

**Claim:** All eigenvalues of an orthogonal or unitary matrix

have magnitude 1.

Proof:

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

$$\Rightarrow \|\vec{v}\|^2 = v^+ v = v^+ \underbrace{U^+ U}_{I} v$$
$$= \|Uv\|^2$$
$$= |\lambda|^2 \times \|v\|^2 \qquad \Rightarrow |\lambda| = 1 \qquad \square$$

**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

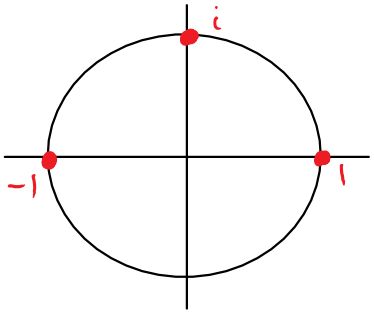

ans =

   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 =
```

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

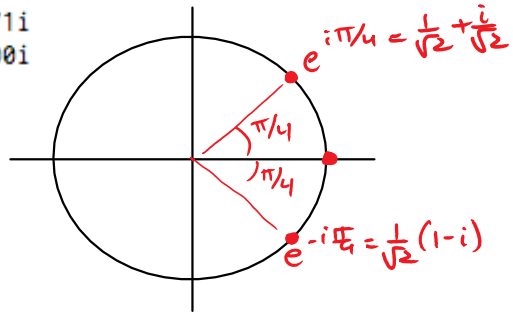

ans =

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


ans =

    1.0000
    1.0000
    1.0000
```

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




**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

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

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

Proof:

$$Av = \lambda v$$
$$\Rightarrow A^* v^* = \lambda^* v^* \quad \text{(taking complex conjugates of everything)}$$
$$\Rightarrow A v^* = \lambda^* v^* \quad \text{(since } A \text{ is real, } A^* = A\text{).} \checkmark \square$$

**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:

$Ae_1 = e_3$

$Ae_2 = e_1$

$Ae_3 = e_2$

$\Rightarrow A^3 = I$

$\Rightarrow$ each eigenvalue $\lambda$ 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)

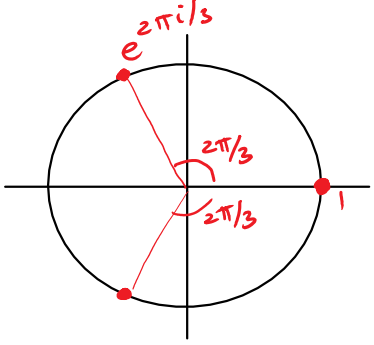
Only one dimension remains, so there isn't room for another complex e-value $\Rightarrow$ H must be real

$\Rightarrow$ H is 1

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

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

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 =
```

```
ans =

    1.0000
    1.0000
    1.0000
```

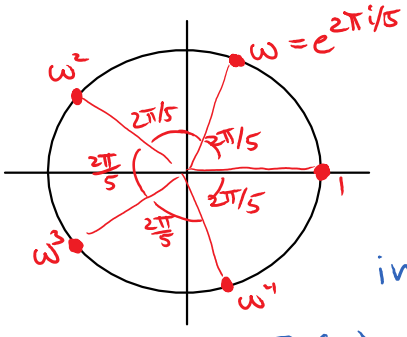
**Exercise:** Similarly, what are the eigenvalues of

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

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

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

**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

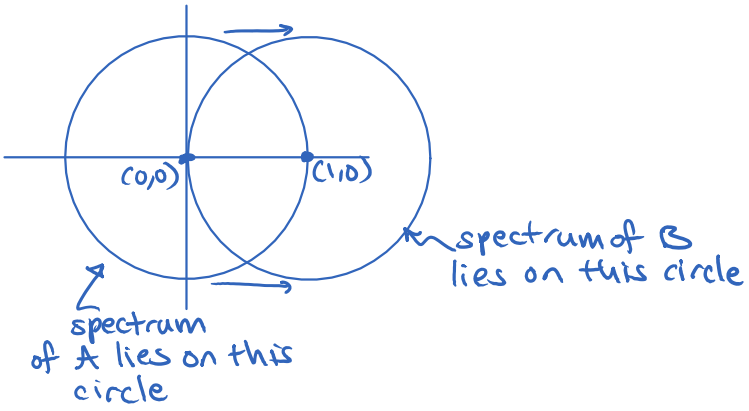
$$|e^{ik\theta_j}(v_j \cdot x)| = |v_j \cdot x|$$

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.




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

**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 $2 \times 2$ 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 $2 \times 2$ 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.

It is not closed under multiplication by complex numbers.

Eg.: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 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. $\lambda$.

$v^\dagger A v$

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

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

$\Rightarrow \lambda = \lambda^*$

ie., $\lambda$ is real. $\square$

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

Proof: All the eigenvalues are real.

For each eigenvalue $\lambda$, 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
     -3.3097
     -2.8308
     -2.2846
     -1.7154
     -1.1692
     -0.6903
     -0.3175
     -0.0810
  ```

  symmetric $A = A^T$

  } 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. ✓

## III. Positive Semi-Definite Matrices

**Definition:** A Hermitian (or real-symmetric) matrix $A$ with all eigenvalues $\geq 0$ is called

**positive semi-definite**

denoted $A \succeq 0$.

**Key property:**

**Theorem:** A Hermitian matrix $A$ is positive semidefinite

$\Leftrightarrow$

$x^T A x \geq 0$ 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 $\alpha_j = v_j \cdot x$.

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

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

$$\geq 0 \quad \checkmark$$

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

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

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

$$< 0. \quad \checkmark$$

□

**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, $A^T A \succeq 0$. ✓ □

**Example 2:** A simple iterative solver for $Ax = b$

Gaussian elimination solves for $x$ in $Ax = b$ using $O(n^3)$ basic operations, when $A$ is $n \times n$.

**Goal:** Approximately solve $Ax = b$ faster than G.E.

**Assume:** $A \succeq 0$ (i.e. $A^T = A$ and eigenvalues $\geq 0$)

(Otherwise, solve the normal equations $A^T A x = A^T b$.)

Assume: $A \succ 0$ (i.e. $A = A^T$ and eigenvalues $> 0$)

(Otherwise, solve the normal equations $\underbrace{A^TA}_{\succ 0}x = A^Tb$.)

**Algorithm:**

0. Choose initial guess $\vec{x}^0$, and some $\alpha > 0$.
1. Repeat for $t = 1, 2, \ldots$:

$$x^t = x^{t-1} - \alpha(\underbrace{Ax^{t-1} - b}_{\text{error}})$$

**Complexity:**

time $\sim$ (# of iterations) $\cdot$ (time for matrix multiplication $=$ # nonzero entries in $A$ $\leq O(n^2)$)

**Convergence analysis:**

Let $x^*$ solve $Ax^* = b$.

$$\Rightarrow x^t - x^* = (I - \alpha A)x^{t-1} + \alpha \underbrace{b}_{Ax^*} - x^*$$

$$= (I - \alpha A)x^{t-1} - (I - \alpha A)x^*$$

$$= (I - \alpha A)(x^{t-1} - x^*)$$

$\Rightarrow$ error magnitude

$$\|x^t - x^*\| \leq \|I - \alpha A\| \cdot \|x^{t-1} - x^*\|$$

How should we choose $\alpha$?

If $A$'s eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, the eigenvalues of $I - \alpha A$ are $1 - \alpha\lambda_1 \leq \cdots \leq 1 - \alpha\lambda_n$.

$$\Rightarrow \|I - \alpha A\| = \max_i |1 - \alpha\lambda_i|$$

The maximum is achieved at one of the endpoints, $i = 1$ or $i = n$, and to make it as small as possible we should **ideally** choose $\alpha$ to equalize them:

$$-(1 - \alpha\lambda_1) = 1 - \alpha\lambda_n$$

$$\Rightarrow \alpha = \frac{2}{\lambda_1 + \lambda_n}$$

$$\Rightarrow \|I - \alpha A\| = 1 - \underbrace{\frac{2\lambda_n}{\lambda_1 + \lambda_n}}_{\leq 2\lambda_1} \leq 1 - \frac{\lambda_n}{\lambda_1}$$

$$= 1 - \frac{1}{\kappa}$$

$\kappa$: condition number

$$\Rightarrow \|x^t - x^*\| \le \|I - \alpha A\|^t \cdot \|x^0 - x^*\|$$

$$\le \left(1 - \frac{1}{\kappa}\right)^t \cdot \|x^0 - x^*\|$$

$$\Rightarrow t = \kappa\left(\log\frac{1}{\varepsilon} + \log\frac{1}{\|x^0 - x^*\|}\right) \quad \text{since } \left(1-\frac{1}{\kappa}\right)^k \le e^{-1}$$

ensures $\|x^t - x^*\| \le \varepsilon$.

**Remark:** The conjugate gradient algorithm has $O(\sqrt{\kappa})$ dependence instead of $O(\kappa)$ dependence, but this is simpler!

**Note:** $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ might be an even better example of clever diagonalization

- normal