Lineær algebra – 5
Egenværdiproblemer og lineære transformationer

Lineær algebra og dynamiske systemer

Troels B. Sørensen /Gilberto Berardinelli

Linear transformations

A linear transformation is a mapping F from vector space X to space Y

$$\bar{x} \in X \qquad \bar{y} = F(\bar{x}) \qquad \bar{y} \in Y$$

"image"

$$\bar{\bar{A}}\,\bar{x} = \bar{y}$$

$\bar{x} = x_1 \bar{e}_1 + \dots x_n \bar{e}_n$

$\bar{e}_i$ $(n \times 1)$ basis

$F(\bar{x}) = x_1 F(\bar{e}_1) + \dots x_n F(\bar{e}_n)$

$\bar{\varepsilon}_i = F(\bar{e}_i)$ $(m \times 1)$ basis

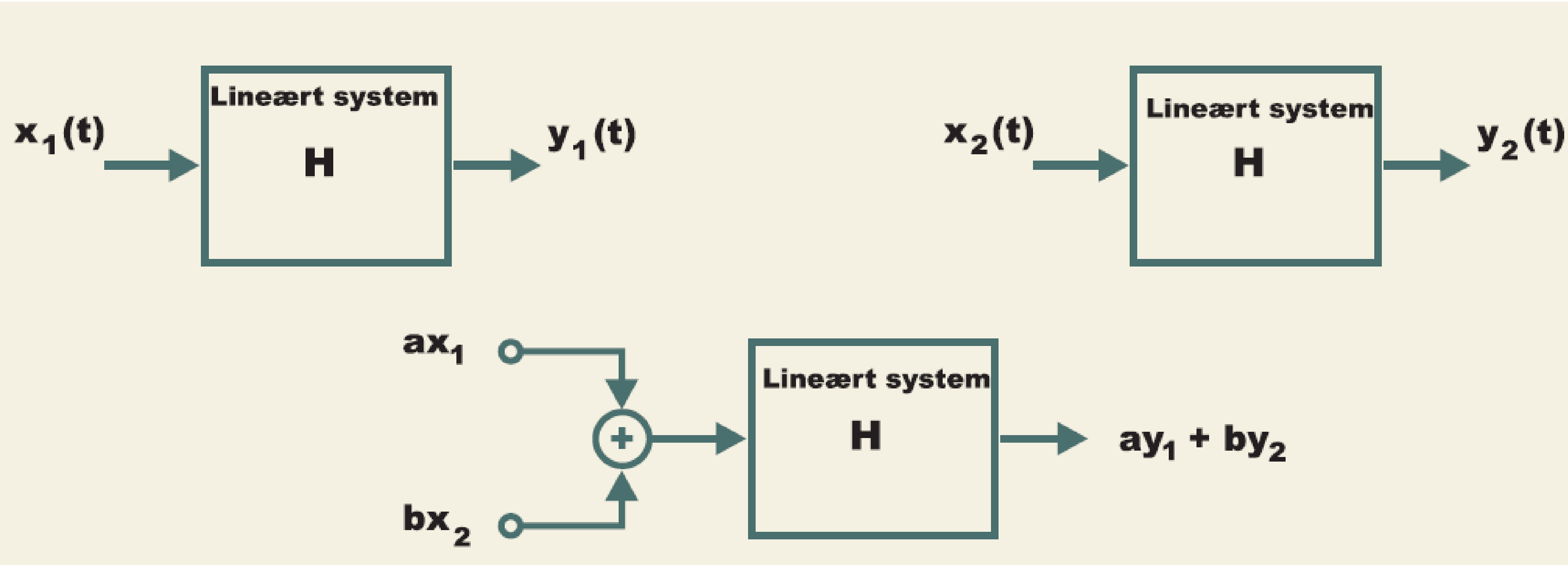
Linear transformation (cont.)

On the left we have $\bar{x}$ specified in terms of the basis $\bar{e}_i$, e.g. standard basis, in the vector space X, and on the right on the basis $\bar{\varepsilon}_i$ in vector space Y given by the columns of $\bar{\bar{A}}$.

We replace one object of study, $\bar{x}$, with another, $\bar{y}$, on another basis

Linear systems (general)

Superposition and Homogeneity
(proportionality)



Linear transformations

If $\bar{\bar{A}}$ is square $(n \times n)$ and non-singular, $\bar{\bar{A}}^{-1}$ is defined, thus

$$\bar{x} = F^{-1}(\bar{y}) = \bar{\bar{A}}^{-1}\bar{y}$$

An orthogonal $(n \times n)$ matrix is one such, $\bar{\bar{A}}^{-1} = \bar{\bar{A}}^{T}$, i.e. a rotation in 2D and 3D Euclidean space. It preserves the inner product (length).

Eigenvalue problems

One special and important linear transformation is : $\bar{A}$ $(n \times n)$

$$\bar{\bar{A}}\,\bar{x} = \lambda\,\bar{x}$$

where the image is a scaled version of the input!

This arises in many engineering problems, known as the matrix eigenvalue problem

The eigenvalue problem

$$\bar{\bar{A}}\,\bar{x} = \lambda\,\bar{x} \quad , \quad \bar{x} \neq \bar{0} \ !$$

unknown eigenvector (characteristic vector)

unknown eigenvalue (scalar) (characteristic value, latent value)

$\{\lambda_i\}$: spectrum of $\bar{\bar{A}}$

$\max|\lambda_i|$: spectral radius of $\bar{\bar{A}}$

$\{\bar{x}_i, \bar{0}\}_{\lambda_i}$: eigenspace of $\bar{\bar{A}}$ corresponding to λ_i

The eigenvalue problem (cont.)

We want to find solutions of

$$\bar{\bar{A}}\,\bar{x} = \lambda\,\bar{x} \iff (\bar{\bar{A}} - \lambda\bar{\bar{I}})\,\bar{x} = \bar{0}$$

a homogeneous linear system of equ.'s

For this to have solutions $(\bar{\bar{A}} - \lambda\bar{\bar{I}})$ needs to be singular, i.e., $\det(\bar{\bar{A}} - \lambda\bar{\bar{I}})$ needs to be zero:

$$D(\lambda) = \det(\bar{\bar{A}} - \lambda\bar{\bar{I}}) = 0$$

characteristic equation

characteristic determinant / polynomial (of degree n)

Eigenvalues

The roots of

$$D(\lambda) = \det(\bar{\bar{A}} - \lambda \bar{\bar{I}}) = 0$$

are the eigenvalues of $\bar{\bar{A}}$ $(n \times n)$

There is at least 1 and at most n numerically different eigenvalues.

The order m_{λ_i} of an eigenvalue λ_i is called the Algebraic multiplicity of λ_i

!

$$\sum_i m_{\lambda_i} = n$$

Eigenvectors

For each λ_i we need to find those x_i's

$$(\bar{\bar{A}} - \lambda_i \bar{\bar{I}})\, \bar{x}_i = \bar{0}$$

Gauss elimination !

We know that there are infinitely many (non-trivial) solutions, lying in the null-space of $\bar{\bar{A}} - \lambda \bar{\bar{I}}$ (whose dimension is the nullity).

When solving the system, we will therefore have "free variables" we can assign freely to obtain the essential vectors (a basis) !

Eigenvectors (cont.)

The number of linearly independent vectors (corresponding to the "free assignments", m_{λ_i}, is the Geometric multiplicity of λ_i

AND the dimension of the eigenspace of $\bar{\bar{A}}$ corresponding to λ_i: $\{\bar{x}_i\}_{\lambda_i}$

- $m_{\lambda_i} \leq M_{\lambda_i}$
- if $\{\bar{x}_i\}_{\lambda_i}$ are eigenvectors for λ_i, then are their linear combinations $\rightarrow \{\bar{x}_i, \bar{0}\}_{\lambda_i}$

Recipe – eigenvalue problem

1. **Dan den karakteristiske matrix:**

$$\cancel{M} = A - \lambda I$$

2. **Find den karakteriske determinant**

$$\Delta \cancel{M} = |A - \lambda I|$$

polynomium / ligning

3. **Dette giver den karakteriske ligning:**

$$a\lambda^{n} + b\lambda^{n-1} + c\lambda^{n-2} + \cdots \quad = 0$$

4. **Løsningerne til den karakteristiske ligning giver spektret:**

Spektral radius

$$S = \{\lambda_1, \lambda_2, \lambda_3, \cdots\}$$

5. **Dette indeholder mellem 1 og n egenværdier**

   Bemærk den algebraiske multiplicitet for λ_k (ordenen af roden).

6. **Indsæt hver λ_k i den karakteristiske matrix M og løs det homogene system (fx vha. Gaussisk elimination):**

$$A - \lambda I = 0$$

   Løsningerne er egenvektorerne.

7. **Bemærk den geometriske multiplicitet (antallet af lineært uafhængige egenvektorer for den givne λ_k)**

Number of eigenvectors

$$\bar{\bar{X}} = \bar{\bar{A}} - \lambda \bar{\bar{I}} \quad \text{in} \quad \bar{\bar{X}}\bar{x} = \bar{0}$$

We know that:

$$\text{rank}(\bar{\bar{X}}) \leq n-1$$

$\Updownarrow$ $(\times(-1))$ to have solutions $x \neq \bar{0}$

$$-\text{rank}(\bar{\bar{X}}) \geq -(n-1)$$

$\Updownarrow$ $(+n)$

$$\underline{\text{nullity}(\bar{\bar{X}})} = n - \text{rank}(\bar{\bar{X}}) \; \underline{\geq} \; n-(n-1) = \underline{1}$$

The dimension of the eigenspace is minimum 1

Number of eigenvectors (cont.)

Using the definition of $\bar{\bar{X}}$

$$\mathrm{rank}(-\lambda \bar{\bar{I}}) = \mathrm{rank}(\bar{\bar{X}} - \bar{\bar{A}})$$

- Rank is subadditive, which means

$$\mathrm{rank}(-\lambda \bar{\bar{I}}) = n \leq \mathrm{rank}(\bar{\bar{X}}) + \mathrm{rank}(\bar{\bar{A}})$$

$$\Updownarrow$$

$$\mathrm{nullity}(\bar{\bar{X}}) = n - \mathrm{rank}(\bar{\bar{X}}) \leq \mathrm{rank}(\bar{\bar{A}})$$

The dimension of the eigenspace is maximum n (for $\bar{\bar{A}}$ $n \times n$)

Eigenvalues (square matrices)

Hermetian C^n $\qquad \bar{A}^{*T} = \bar{A}$
(symmetric, R^n)

λ_i real (0)

Skew-Hermetian C^n $\qquad \bar{A}^{*T} = -\bar{A}$
(skew-symmetric, R^n)

λ_i imaginary (0)

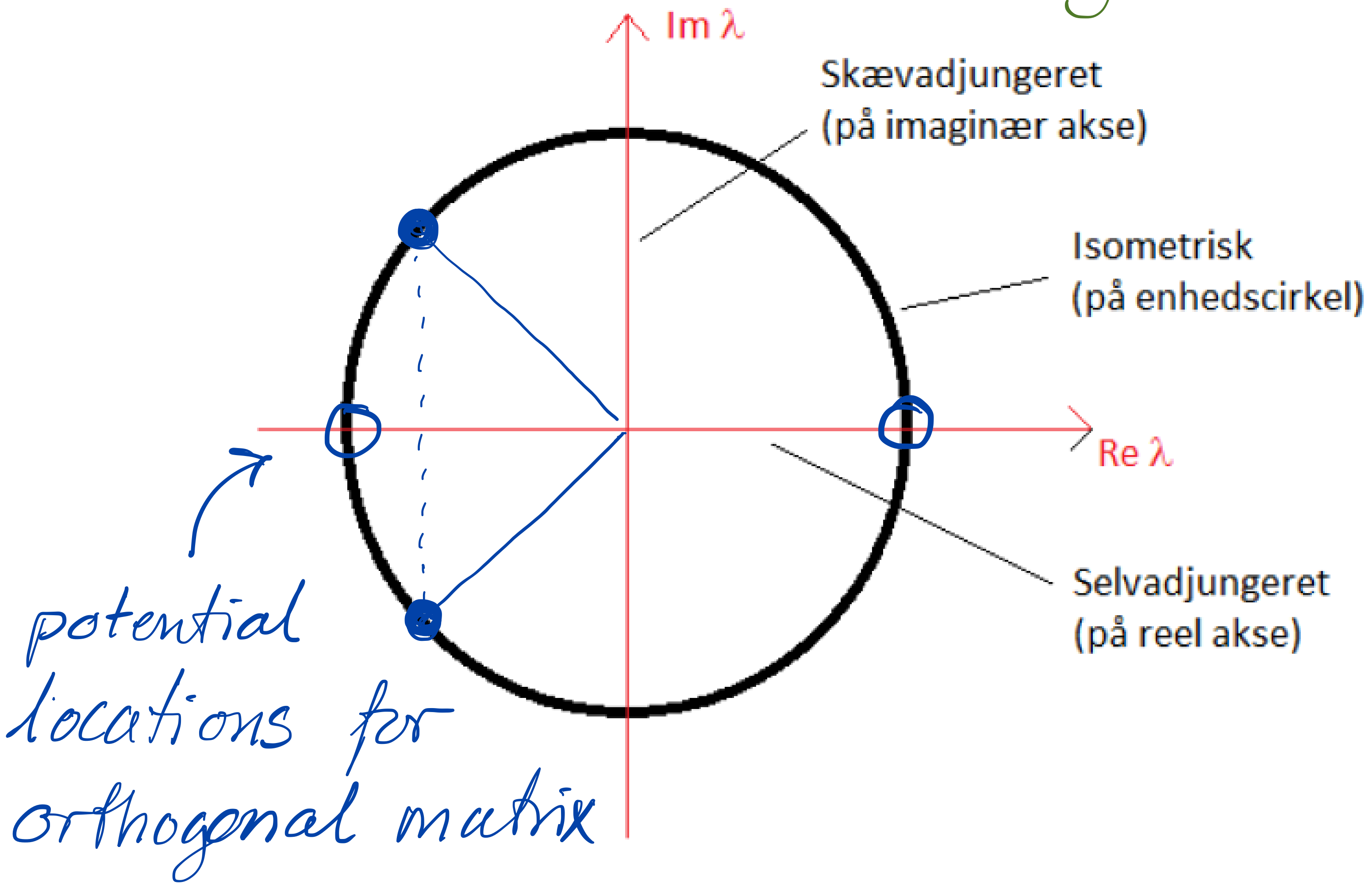
Unitary C^n $\qquad \bar{A}^{*T} = \bar{A}^{-1}$
(orthogonal, R^n)

$|\lambda_i| = 1$

$*$ complex conjugate

Eigenvalues in complex plane

Argand diagram



Taxonomy for normal matrices

Reel matrix	Kompleks matrix	Generel	Normal[2]	Egenværdier
Symmetrisk $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$	Hermitesk (diagonal er reel) $\boldsymbol{A}^{*\mathrm{T}} = \boldsymbol{A}$	Selvadjungeret[1]	Ja	Reelle (inkl. 0)
Skævsymmetrisk (diagonal = 0) $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$	Skævhermitesk (diagonal imaginær eller 0) $\boldsymbol{A}^{*\mathrm{T}} = -\boldsymbol{A}$	Skævadjungeret	Ja	Imaginære (inkl. 0)
Ortogonal ($\Delta\boldsymbol{A} = \pm 1$) $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}^{-1}$	Unitær ($\lvert\Delta\boldsymbol{A}\rvert = 1$) $\boldsymbol{A}^{*\mathrm{T}} = \boldsymbol{A}^{-1}$	Isometrisk	Ja	Absolut værdi 1

(1) Den (komplekst) konjugerede transponerede, $\boldsymbol{A}^{*\mathrm{T}}$, kaldes for den adjungerede til $\boldsymbol{A}$ (og deraf betegnelsen selvadjungeret i dette tilfælde); mere formelt kaldes den komplekst konjugerede transponerede for den hermitesk adjungerede, hvor hermitesk adjungering er analogt til kompleks konjugering.

(2) En normal matrix er en (generelt) kompleks kvadratisk matrix der kommuterer med sin adjungerede, dvs. opfylder $\boldsymbol{A}^{*\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{*\mathrm{T}}$.

$$\Delta \bar{\bar{A}} = \det(\bar{\bar{A}})$$

Basis of eigenvectors

Do the eigenvectors of $\bar{\bar{A}}$ $(n \times n)$ form a basis (for a vector space)?

Yes, in many cases, specifically:

- when $\bar{\bar{A}}$ has $\underline{n}$ distinct eigenvalues
- when $\bar{\bar{A}}$ is Hermetian, skew-Hermetian or Unitary, C^n (R^n)

! We call such a basis an <u>eigenbasis</u>

Basis of eigenvectors (cont.)

If an eigenbasis for $\bar{\bar{A}}$ exists $\{\bar{x}_i\}$:

- Eigendecomposition $\qquad \bar{\bar{A}}\bar{x}_i = \lambda_i \bar{x}_i$

$$\bar{y} = \bar{\bar{A}}\bar{x} = \bar{\bar{A}}(c_1\bar{x}_1 + \cdots c_n\bar{x}_n)$$
$$= c_1\bar{\bar{A}}\bar{x}_1 + \cdots c_n\bar{\bar{A}}\bar{x}_n$$
$$= c_1\lambda_1\bar{x}_1 + \cdots c_n\lambda_n\bar{x}_n$$

- Diagonalization

$$\bar{\bar{D}} = \bar{\bar{X}}^{-1}\bar{\bar{A}}\,\bar{\bar{X}} \longrightarrow$$

Diagonalization

$$\bar{\bar{D}} = \bar{\bar{X}}^{-1}\,\bar{\bar{A}}\,\bar{\bar{X}}$$

$$\begin{bmatrix} \lambda_1 & & \bar{0} \\ & \ddots & \\ \bar{0} & & \lambda_n \end{bmatrix}$$ (diagonal) matrix of eigenvalues

$$[\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n]$$ matrix of eigenvectors

$\bar{\bar{X}}$ is said to diagonalize $\bar{\bar{A}}$

$\bar{\bar{D}}$ and $\bar{\bar{A}}$ are called similar* in that they have the same eigenvalues

(* generally, any non-singular matrix $\bar{\bar{X}}$ sufficise

Diagonalization (cont.)

To see this: $\bar{\bar{A}}\bar{x} = \lambda \bar{x}$, $\bar{x} \neq 0$

$$\bar{\bar{X}}^{-1}(\bar{\bar{A}}\bar{x}) = \lambda \bar{\bar{X}}^{-1}\bar{x}$$

$$\Updownarrow$$

$$\bar{\bar{X}}^{-1}\bar{\bar{A}}(\bar{\bar{X}}\,\bar{\bar{X}}^{-1})\bar{x} = \lambda \bar{\bar{X}}^{-1}\bar{x}$$

$$\Updownarrow$$

$$(\bar{\bar{X}}^{-1}\bar{\bar{A}}\bar{\bar{X}})(\bar{\bar{X}}^{-1}\bar{x}) = \lambda(\bar{\bar{X}}^{-1}\bar{x})$$

So, λ is (also) an eigenvalue of $\bar{\bar{X}}^{-1}\bar{\bar{A}}\bar{\bar{X}} = \bar{\bar{D}}$,

BUT corresponding eigenvector $\hat{\bar{x}} = \bar{\bar{X}}^{-1}\bar{x}$

$\left(\bar{x} = \bar{\bar{X}}\,\bar{\bar{X}}^{-1}\bar{x} = \bar{\bar{X}}\bar{0} = \bar{0}, \text{ so } (\bar{\bar{X}}^{-1}\bar{x}) \neq \bar{0} \right)$

Relations for trace and det.

From the similarity, e.g. from $\bar{\bar{D}}$, it follows that for an invertible matrix $\bar{\bar{A}}$ $(n \times n)$:

$$\operatorname{trace}(\bar{\bar{A}}) = \sum_i a_{ii} = \sum_i \lambda_i \quad \text{trace}(\bar{\bar{X}}\,\bar{\bar{D}}\,\bar{\bar{X}}^{-1})$$

$$\det(\bar{\bar{A}}) = \prod_i \lambda_i \quad \det(\bar{\bar{X}}\,\bar{\bar{D}}\,\bar{\bar{X}}^{-1})$$

! This can be a shortcut to find the eigenvalues

! (Also, $\bar{\bar{A}}^n = \bar{\bar{X}}\,\bar{\bar{D}}^n\,\bar{\bar{X}}^{-1}$)

Numerical methods / limits

When n is large we might need numerical methods for "locating" the eigenvalues.

<u>Gerschgorin's theorem:</u>

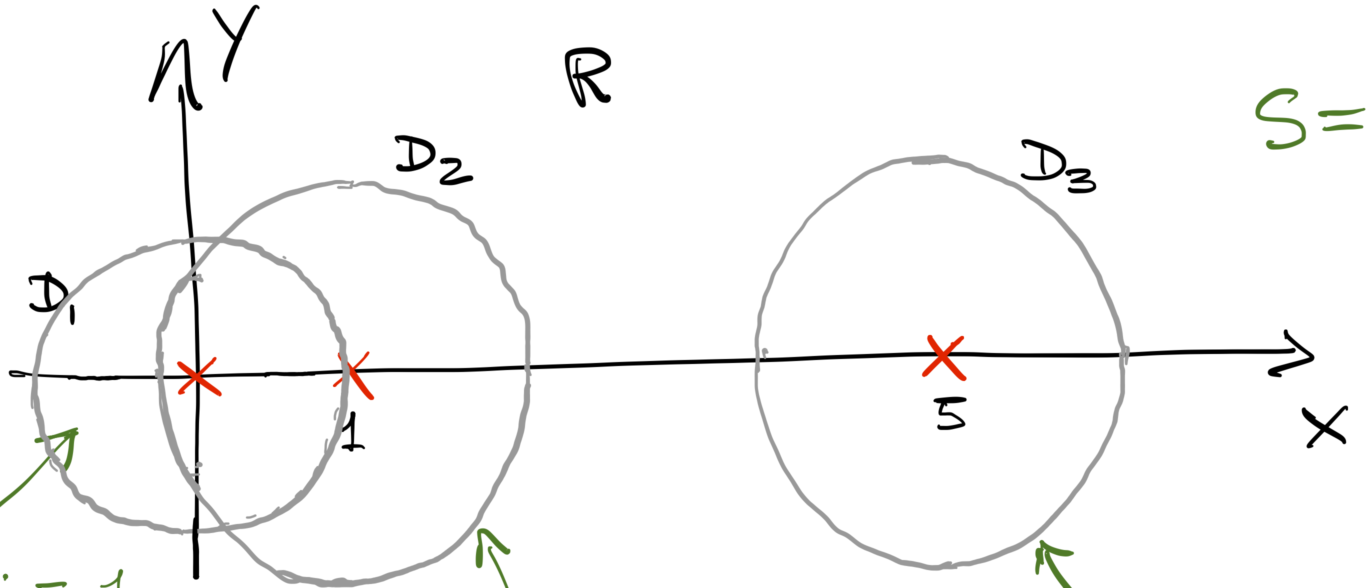
If λ is an eigenvalue of $\bar{\bar{A}}$ $(n \times n)$, then for $1 \le i \le n$ (some i)

$$|a_{ii} - \lambda| \le \sum_{j \ne i} |a_{ij}|$$

Disk inclusion (ex.)

$$\bar{\bar{A}} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 5 & 1 \\ 1/2 & 1 & 1 \end{bmatrix} \qquad \bar{\bar{A}}^T = \bar{\bar{A}} \Rightarrow \lambda_i \text{ real}$$

disk centers



$S = \{D_1, D_2\} \cap D_3 = \emptyset$

$\Downarrow$

S contains precisely $|S|$ λ_i's !

$\sum_{j \neq 1} a_{1j} = 1$

$\sum_{j \neq 2} a_{2j} = 1.5$

$\sum_{j \neq 3} a_{3j} = 1.5$

Limit bounds (ex.)

Schur's theorem (inequality)

If λ is an eigenvalue of $\bar{\bar{A}}$ $(n \times n)$, then

$$|\lambda_m|^2 \leq \sum_i |\lambda_i|^2 \leq \sum_i \sum_j |a_{ij}|^2 \quad , 1 \leq m \leq n$$

(=) holds if and only if $\bar{\bar{A}}$ is normal

With $\bar{\bar{A}}$ on previous slide:

$$|\lambda_m| \leq 5.315$$

! compared to $5 + 1.5 = 6.5$ obtained with disk inclusion

Power method

Power algorithm for dominant eigenvalue of $\bar{\bar{A}}$ (* symmetric $(n \times n)$)

$i = 0$, $\bar{x}^{(0)} = \bar{b}_0 \neq \bar{0}$, ε (stop criteria)

do

$$\bar{x}^{(i+1)} = \bar{\bar{A}}\,\bar{x}^{(i)}$$

$$\left.\begin{aligned} s &= \max_j |x_j^{(i+1)}| \\ \bar{x}^{(i+1)} &= \bar{x}^{(i+1)}/s \end{aligned}\right\}$$

scaling by absolutely largest element in vector

$$m_0 = (\bar{x}^{(i)})^T \bar{x}^{(i)} \,;\; m_1 = (\bar{x}^{(i)})^T \bar{x}^{(i+1)} \,;\; m_2 = (\bar{x}^{(i+1)})^T \bar{x}^{(i+1)}$$

$$\tilde{\lambda} = m_1/m_0$$

$$^{*}\ \delta = \sqrt{m_2/m_0 - \tilde{\lambda}^2}$$

while $(\delta > \varepsilon)$

$\boxed{\underline{\text{OUTPUT}}: \tilde{\lambda},\ \bar{x}^{(i+1)} \quad \forall\ \bar{x}^{(i+1)} = \bar{\bar{A}}^{i+1}\bar{b}_0}$

Power method (cont.)

- Applies in general when there is a dominant eigenvalue ($|\lambda_k| > |\lambda_i|$, $1 \leq i \leq n$)

  To make sure there's a dominant eigenvalue it is possible to apply a spectral shift:

$$\tilde{\bar{\bar{A}}} = \bar{\bar{A}} - k \cdot \bar{\bar{I}} \implies \tilde{\lambda}_i = \lambda_i - k$$

  to help improve convergence (speed).

- Applies when the dominant eigenvalue is of interest, and when $\bar{\bar{A}}$ is "large" and sparse

Power method (cont.)

Assume that $\bar{\bar{A}}$ is diagonizable, then from before

doesn't need to be!

(eigendecomposition)

$$\bar{x}^{(i+1)} = \bar{\bar{A}}^i \bar{b}_0 = c_1 \bar{\bar{A}}^i \bar{x}_1 + c_2 \bar{\bar{A}}^i \bar{x}_2 + \ldots\ldots c_n \bar{\bar{A}}^i \bar{x}_n$$

$$= c_1 \lambda_1^i \bar{x}_1 + \ldots\ldots\ldots\ldots\ldots c_n \lambda_n^i \bar{x}_n$$

$$= c_1 \lambda_1^i \left(\bar{x}_1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^i \bar{x}_2 + \ldots\ldots \frac{c_n}{c_1}\left(\frac{\lambda_n}{\lambda_1}\right)^i \bar{x}_n \right)$$

$$= \Big|_{i\to\infty} c_1 \lambda_1^i \bar{x}_1 \quad \propto \bar{x}_1 \quad \underline{\text{when } |\lambda_1| > |\lambda_i|}$$

hence,

$$\frac{m_1}{m_0} = \frac{(\bar{x}^{(i)})^T \bar{x}^{(i+1)}}{(\bar{x}^{(i)})^T \bar{x}^{(i)}} = \Big|_{i\to\infty} \frac{\bar{x}_1^T \bar{\bar{A}} \bar{x}_1}{\bar{x}_1^T \bar{x}_1} = \frac{\bar{x}_1^T \lambda_1 \bar{x}_1}{\bar{x}_1^T \bar{x}_1} = \lambda_1$$