Partial Differential Equations

Chapter VI - Numerical Linear Algebra

John Cagnol, Pauline Lafitte

The Engineering Program of CentraleSupélec

Lecture 8 – February 24th 2020

VI.1. Introduction

We have explained how to approximate the solution to a PDE using the Finite Element Method.

We have to solve the linear system $A_h u_h = b_h$.

We have explained how to approximate the solution to a PDE using the Finite Element Method.

We have to solve the linear system $A_h u_h = b_h$.

The Gaussian elimination has an arithmetic complexity in n^3.

We have explained how to approximate the solution to a PDE using the Finite Element Method.

We have to solve the linear system $A_h u_h = b_h$.

The Gaussian elimination has an arithmetic complexity in n^3.

The matrix A_h can be very large:
Solving $A_h u_h = b_h$ can be an issue.

This requires developing strategies.

VI.2. Norms on a matrix

Norms

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.
$E = \mathcal{M}_{q \times p}(\mathbb{K})$ is a $\mathbb{K}$-linear space.

Norms

Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.
$E = \mathcal{M}_{q\times p}(\mathbb{K})$ is a $\mathbb{K}$-linear space.

Recall the definition of a norm from the first lecture of CIP:

Definition VI.2.1

Let E be a $\mathbb{K}$-linear space. A **norm** N: $E \to \mathbb{R}_+$ *is a mapping satisfying:*

- $N(x) = 0 \Leftrightarrow x = 0$ **(separation)**;
- $\forall \lambda \in \mathbb{K}, \forall x \in E,\ N(\lambda x) = |\lambda| N(x)$ **(homogeneity)**;
- $\forall x, y \in E,\ N(x+y) \leq N(x) + N(y)$ **(triangle inequality)**.

Definition VI.2.2 (Norm)

The mapping $\|\cdot\| : \mathcal{M}_{q\times p} \to \mathbb{R}^+$ *is* **norm** *if it satisfies:*

- $\forall A, B \in \mathcal{M}_{q\times p},\ \|A\| = 0 \Leftrightarrow A = 0$ **(separation)**
- $\forall \lambda \in \mathbb{K}, \forall A \in \mathcal{M}_{q\times p},\ \|\lambda A\| = |\lambda|\|A\|$ **(homogeneity)**
- $\forall A, B \in \mathcal{M}_{q\times p},\ \|A + B\| \leq \|A\| + \|B\|$ **(triangle ineq.)**.

Definition VI.2.2 (Norm)

The mapping $\|\cdot\| : \mathcal{M}_{q\times p} \to \mathbb{R}^+$ *is* **norm** *if it satisfies:*

- $\forall A, B \in \mathcal{M}_{q\times p},\ \ \|A\| = 0 \Leftrightarrow A = 0$ **(separation)**
- $\forall \lambda \in \mathbb{K}, \forall A \in \mathcal{M}_{q\times p},\ \ \|\lambda A\| = |\lambda|\|A\|$ **(homogeneity)**
- $\forall A, B \in \mathcal{M}_{q\times p},\ \ \|A + B\| \leq \|A\| + \|B\|$ **(triangle ineq.)**.

Definition VI.2.3 (Subordinance)

A norm $\|\cdot\|$ *on* $\mathcal{M}_{q\times p}$ *is* **subordinated** *if*

$$\forall A \in \mathcal{M}_{q\times p}, ||Ax|| \leq \|A\|||x||$$

Definition VI.2.2 (Norm)

The mapping $\|\cdot\| : \mathcal{M}_{q\times p} \to \mathbb{R}^+$ *is* **norm** *if it satisfies:*

- $\forall A, B \in \mathcal{M}_{q\times p},\ \|A\| = 0 \Leftrightarrow A = 0$ **(separation)**
- $\forall \lambda \in \mathbb{K}, \forall A \in \mathcal{M}_{q\times p},\ \|\lambda A\| = |\lambda| \|A\|$ **(homogeneity)**
- $\forall A, B \in \mathcal{M}_{q\times p},\ \|A + B\| \le \|A\| + \|B\|$ **(triangle ineq.)**.

Definition VI.2.3 (Subordinance)

A norm $\|\cdot\|$ *on* $\mathcal{M}_{q\times p}$ *is* **subordinated** *if*

$$\forall A \in \mathcal{M}_{q\times p}, ||Ax|| \le \|A\| ||x||$$

Definition VI.2.4 (Submultiplicativity)

A norm $\|\cdot\|$ *on* $\mathcal{M}_{q\times p}$ *is* **submultiplicative** *if*

$$\forall A, B \in \mathcal{M}_{q\times p},\ \|AB\| \le \|A\| \|B\|$$

Induced norm

Consider $\mathbb{K}^p$ and $\mathbb{K}^q$ equipped with a norm $\|\cdot\|$.

Induced norm

Consider $\mathbb{K}^p$ and $\mathbb{K}^q$ equipped with a norm $\|\cdot\|$.

Let $A \in \mathcal{M}_{q\times p}(\mathbb{K})$ be a matrix.
It represents a linear application from $\mathbb{K}^p$ to $\mathbb{K}^q$.

Induced norm

Consider $\mathbb{K}^p$ and $\mathbb{K}^q$ equipped with a norm $\|\cdot\|$.

Let $A \in \mathcal{M}_{q\times p}(\mathbb{K})$ be a matrix.
It represents a linear application from $\mathbb{K}^p$ to $\mathbb{K}^q$.

For $x \in \mathbb{K}^p \setminus \{0\}$, the non-negative number

$$\frac{\|Ax\|}{\|x\|}$$

gives the gain (or amplification) of A in the direction of x.

Induced norm

Consider $\mathbb{K}^p$ and $\mathbb{K}^q$ equipped with a norm $\|\cdot\|$.

Let $A \in \mathcal{M}_{q\times p}(\mathbb{K})$ be a matrix.
It represents a linear application from $\mathbb{K}^p$ to $\mathbb{K}^q$.

For $x \in \mathbb{K}^p \setminus \{0\}$, the non-negative number

$$\frac{\|Ax\|}{\|x\|}$$

gives the gain (or amplification) of A in the direction of x.

The amplification factor will vary with the direction of x.

Induced norm

Consider $\mathbb{K}^p$ and $\mathbb{K}^q$ equipped with a norm $\|\cdot\|$.

Let $A \in \mathcal{M}_{q\times p}(\mathbb{K})$ be a matrix.
It represents a linear application from $\mathbb{K}^p$ to $\mathbb{K}^q$.

For $x \in \mathbb{K}^p \setminus \{0\}$, the non-negative number

$$\frac{\|Ax\|}{\|x\|}$$

gives the gain (or amplification) of A in the direction of x.

The amplification factor will vary with the direction of x.
What is the maximum possible gain of A?

Induced norm

Definition-Proposition VI.2.5

The maximum gain of A

$$\sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

is a norm on the linear space $\mathcal{M}_{q\times p}(\mathbb{K})$ *called the* **induced norm**. *It depends on the norm defined on the linear space* $\mathbb{K}$*: different norms on* $\mathbb{K}$ *will produce different induced norms on* $\mathcal{M}_{q\times p}(\mathbb{K})$

Remark VI.2.6

The induced normed is subordinated: $\|Ax\| \leq \|\!|A|\!\| \|x\|$.

Proposition VI.2.7

The induced norm is submultiplicative.

Induced norms

Let $q = p$ and $\mathbb{K} = \mathbb{R}$.

Example

Consider the norm $\|\cdot\|_1$ *on* $\mathbb{R}^q$*:*
For $x = (x_1, \ldots, x_q) \in \mathbb{R}^q$, $\|x\| = \sum_{i=1}^{q} |x_i|$.

The induced matrix norm $\|\!|\cdot|\!\|_1$ *is given by*

$$\|\!|A|\!\|_1 = \max_{1 \le j \le q} \sum_{i=1}^{q} |a_{ij}|$$

which is the maximum absolute column sum of the matrix.

Induced norms

Let $q = p$ and $\mathbb{K} = \mathbb{R}$.

Example

Consider the norm $\|\cdot\|_\infty$ *on* $\mathbb{R}^q$*:*
For $x = (x_1, \ldots, x_q) \in \mathbb{R}^q$, $\|x\|_\infty = \max_{1\leq n\leq q} |x_i|$.

The induced matrix norm $\|\cdot\|_\infty$ *is given by*

$$\|A\|_\infty = \max_{1\leq i\leq q} \sum_{j=1}^{q} |a_{ij}|$$

which is the maximum absolute row sum of the matrix.

Induced norms

Let $q = p$ and $\mathbb{K} = \mathbb{R}$.

Example

Consider the norm $\|\cdot\|_2$ on $\mathbb{R}^q$:

For $x = (x_1, \dots, x_q) \in \mathbb{R}^q$, $\|x\|_2 = \sqrt{\sum_{i=1}^q x_i^2}$.

The induced matrix norm $\|\cdot\|$ is given by

$$\|A\|_2 = \sigma_{\max}(A)$$

which is the largest singular value of the matrix A.

(Reminder: the singular values of A are the square roots of the eigenvalues of $A^* A$).

The Frobenius norm

Definition VI.2.8 (The Frobenius inner product)

$\mathcal{M}_{q\times p}(\mathbb{R})$ *can be endowed with the inner product:*

$$\langle A, B\rangle_F = \sum_{i=1}^{p}\sum_{j=1}^{q} A_{ij}B_{ij}$$

It is called the **Frobenius inner product**.

It can also be defined on $\mathcal{M}_{q\times p}(\mathbb{C})$ *by conjugating* A_{ij}.

The Frobenius norm

Definition VI.2.9 (The Frobenius norm)

The norm deriving from the Frobenius inner product is the **Frobenius norm**. *It is defined by*

$$\|A\|_F = \sqrt{\sum_{i=1}^{p} \sum_{j=1}^{q} A_{ij}^2}$$

It can also be defined on $\mathcal{M}_{q\times p}(\mathbb{C})$ by replacing A_{ij}^2 by $|A_{ij}|^2$.

Proposition VI.2.10

The Frobenius norm is submultiplicative.

The Frobenius norm

Proposition VI.2.11

Let $\|\!|\cdot|\!\|_2$ be the norm induced by the norm $\|\cdot\|_2$ on $\mathbb{R}^p$ and $\mathbb{R}^q$,
Let $\|\!|\cdot|\!\|_F$ be the Frobenius norm,
Then

$$\|\!|\cdot|\!\|_2 \leq \|\!|\cdot|\!\|_F$$

The Frobenius norm is easy to compute. This gives a simple way to estimate a bound of $\|Ax\|_2$.

$$\|Ax\|_2 \leq \|\!|A|\!\|_F \|x\|_2$$

Conditioning

Definition VI.2.12 (Condition number)

Let $\|\cdot\|$ be a norm on $\mathbb{K}^q$ and $\|\!|\cdot|\!\|$ be the induced norm on $\mathcal{M}_{q\times p}(\mathbb{K})$.

Let $A \in \mathcal{M}_{q\times p}(\mathbb{K})$ be a non-singular matrix.

The **condition number** *of A relative to $\|\!|\cdot|\!\|$ is defined by*

$$\kappa(A) = |\!|\!|A|\!|\!| \, |\!|\!|A^{-1}|\!|\!|$$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\begin{bmatrix} 41 & 28 \\ 97 & 66 \end{bmatrix}$$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\left[\begin{array}{cc} 41 & 28 \\ 97 & 66 \end{array}\right]$$

The eigenvalues of $A^* A$ *are*

$$\begin{aligned} \lambda_1 &= \frac{1623 - 5\sqrt{105365}}{20} \\ \lambda_2 &= \frac{1623 + 5\sqrt{105365}}{20} \end{aligned}$$

The largest eigenvalue of $A^* A$ *is* λ_2.
Thus $\|A\| = \sqrt{\lambda_2}$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\begin{bmatrix} 41 & 28 \\ 97 & 66 \end{bmatrix}$$

The eigenvalues of $A^* A$ *are approximately*

$$\begin{aligned} \lambda_1 &\simeq 6\,10^{-5} \\ \lambda_2 &\simeq 162 \end{aligned}$$

The largest eigenvalue of $A^* A$ *is* λ_2.
Thus $\quad \|A\| \simeq 12.7$.

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\begin{bmatrix} 41 & 28 \\ 97 & 66 \end{bmatrix}$$

$\|A\| \simeq 12.7.$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\begin{bmatrix} 41 & 28 \\ 97 & 66 \end{bmatrix}$$

$\|A\| \simeq 12.7$.
The eigenvalues of $(A^{-1})^* A^{-1}$ *are approximately*

$$\begin{aligned} \mu_1 &= 8115 - 25\sqrt{105365} \\ \mu_2 &= 8115 + 25\sqrt{105365} \end{aligned}$$

The largest eigenvalue of $(A^{-1})^* A^{-1}$ *is* μ_2.
Thus $\|A^{-1}\| = \sqrt{\mu_2}$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\left[\begin{array}{cc} 41 & 28 \\ 97 & 66 \end{array}\right]$$

$|||A||| \simeq 12.7.$
The eigenvalues of $(A^{-1})^* A^{-1}$ *are approximately*

$$\begin{array}{rcl} \mu_1 & \simeq & 6\ 10^{-3} \\ \mu_2 & \simeq & 16227 \end{array}$$

The largest eigenvalue of $(A^{-1})^* A^{-1}$ *is* μ_2.
Thus $|||A^{-1}||| \simeq 127.4.$

Conditioning

Example

Let $\|\cdot\| = \|\cdot\|_2$ *and*

$$A = \frac{1}{10}\begin{bmatrix} 41 & 28 \\ 97 & 66 \end{bmatrix}$$

$\|A\| \simeq 12.7.$
$\|A^{-1}\| \simeq 127.4.$

Therefore $\kappa(A) \simeq 1618$

Conditioning

Proposition VI.2.13

Consider $A \in \mathcal{M}_{q\times p}(\mathbb{R})$ a non-singular matrix.
Let b_1 and b_2 be two vectors in $\mathbb{R}^q$.
Let $x_1 = A^{-1}b_1$ and $x_2 = A^{-1}b_2$.

Conditioning

Proposition VI.2.13

Consider $A \in \mathcal{M}_{q\times p}(\mathbb{R})$ *a non-singular matrix.*
Let b_1 *and* b_2 *be two vectors in* $\mathbb{R}^q$*.*
Let $x_1 = A^{-1}b_1$ *and* $x_2 = A^{-1}b_2$*. Then*

$$\frac{\|x_2 - x_1\|}{\|x_1\|} \leq \kappa(A)\,\frac{\|b_2 - b_1\|}{\|b_1\|}$$

Where κ *is the condition number relative to the matrix norm induced by* $\|\cdot\|$*.*

Conditioning

Proposition VI.2.13

Consider $A \in \mathcal{M}_{q\times p}(\mathbb{R})$ a non-singular matrix.
Let b_1 and b_2 be two vectors in $\mathbb{R}^q$.
Let $x_1 = A^{-1}b_1$ and $x_2 = A^{-1}b_2$. Then

$$\frac{\|x_2 - x_1\|}{\|x_1\|} \leq \kappa(A)\,\frac{\|b_2 - b_1\|}{\|b_1\|}$$

Where κ is the condition number relative to the matrix norm induced by $\|\cdot\|$.

In other words, the error on solution x to $Ax = b$ is of the same order of magnitude as the error on b multiplied by $\kappa(A)$.

Conditioning

Example

With the matrix A from the previous example, a relative variation of 0.1 on b can lead to a relative variation of up to 161.8 on the solution.

The condition number is always greater or equal to 1.

When solving a linear system of equation, we hope for the condition number to be as small as possible (as close to 1 as possible).

Conditioning

Definition VI.2.14

A **preconditioner** *P of a matrix A is a matrix such that*

$$\kappa(PA) < \kappa(A)$$

Conditioning

Definition VI.2.14

A **preconditioner** *P of a matrix A is a matrix such that*

$$\kappa(PA) < \kappa(A)$$

For a non-singular matrix A, the best preconditioner is $P = A^{-1}$ but if A^{-1} is known, the problem was solved in the first place.

Finding a suitable P is often a trade-off.

Conditioning

Proposition VI.2.15

Let $A \in \mathcal{M}_q(\mathbb{R})$ be a positive definite symmetric matrix with eigenvalues $0 < \lambda_1 \leq \ldots \leq \lambda_q$.

The condition number κ_2 relative to the matrix norm induced by $\|\cdot\|_2$ is given by

$$\kappa_2(A) = \frac{\lambda_q}{\lambda_1}$$

VI.3. Spectral Properties

Eigenvalues, Eigenvectors

As you know, every matrix $A \in \mathcal{M}_q$ has q eigenvalues in $\mathbb{C}$.

The **spectrum of** A, denoted $\sigma(A)$, is the set of the eigenvalues of A in $\mathbb{C}$.

Eigenvalues, Eigenvectors

As you know, every matrix $A \in \mathcal{M}_q$ has q eigenvalues in $\mathbb{C}$.

The **spectrum of** A, denoted $\sigma(A)$, is the set of the eigenvalues of A in $\mathbb{C}$.

Definition VI.3.1 (Schur)

Let $A \in \mathcal{M}_q$. Then, there exist

- *an upper triangular matrix T*
- *a unitary matrix U*

such that $A = UTU^{-1}$.

(Reminder: U is unitary iff $U^* = U^{-1}$)

Finding the spectrum

Computing the eigenvalues of $A \in \mathcal{M}_q$ has a complexity in q^3.

Recent algorithms can bring down the complexity to $q^{2.3}$.

A "brute force" direct computation is rarely achieved on a large matrix.

It is a difficult problem.

Finding the spectrum: The Gershgörin circle theorem

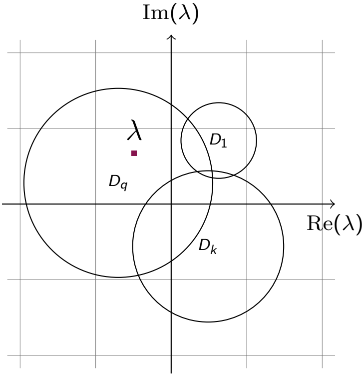
Let $A \in \mathcal{M}_q(\mathbb{C})$.
We denote A_{ij} the components of A.

Definition VI.3.2 (Gershgörin discs)

For $i \in \{1, \ldots, q\}$, let $R_i = \sum_{j \neq i} |Aij|$. The closed disc $D_i = B(a_{ii}, R_i) \subseteq \mathbb{C}$ is called a Gershgorin disc.

Theorem VI.3.3 (Gershgörin)

$\mathrm{Sp}(A) \subset \bigcup_{k=1}^{q} D_k$



Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$A_{11} = 5 \qquad R_1 = 2 + 0 + 1$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$A_{11} = 5 \qquad R_1 = 3$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{aligned} A_{11} &= 5 \qquad & R_1 &= 3 \\ A_{22} &= 3 \qquad & R_2 &= 1 + 2 + 1 \end{aligned}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \end{array}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 + 0 + 0 \end{array}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{aligned} A_{11} &= 5 \qquad & R_1 &= 3 \\ A_{22} &= 3 & R_2 &= 4 \\ A_{33} &= -2 & R_3 &= 0 \end{aligned}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 1 + 1 + 0 \end{array}$$

Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$

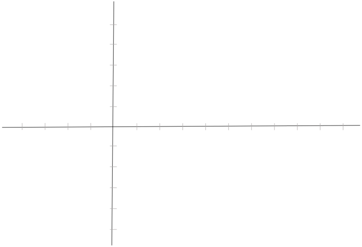
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$



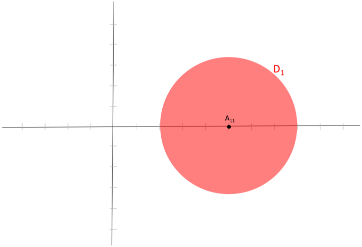
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{aligned} A_{11} &= 5 & R_1 &= 3 \\ A_{22} &= 3 & R_2 &= 4 \\ A_{33} &= -2 & R_3 &= 0 \\ A_{44} &= 8 & R_4 &= 2 \end{aligned}$$



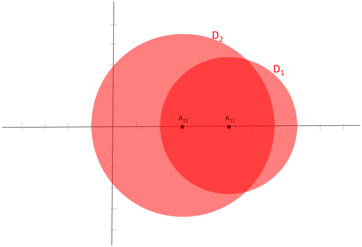
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$



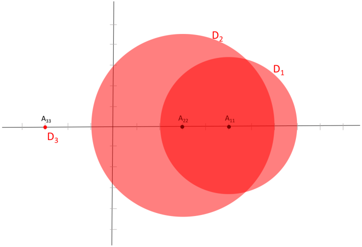
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$



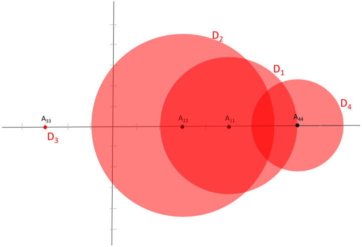
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$



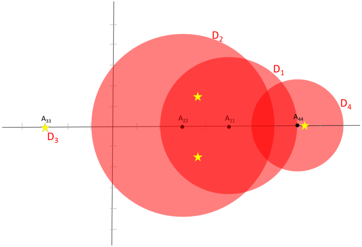
Example

Consider

$$A = \begin{bmatrix} 5 & 2 & 0 & 1 \\ -1 & 3 & 2 & 1 \\ 0 & 0 & -2 & 0 \\ 1 & -1 & 0 & 8 \end{bmatrix}$$

we have

$$\begin{array}{ll} A_{11} = 5 & R_1 = 3 \\ A_{22} = 3 & R_2 = 4 \\ A_{33} = -2 & R_3 = 0 \\ A_{44} = 8 & R_4 = 2 \end{array}$$



An interesting application of the Gershgörin circle theorem

When looking for a preconditioner P for matrix A, the eigenvalues of PA should all be close to 1.

An interesting application of the Gershgörin circle theorem

When looking for a preconditioner P for matrix A, the eigenvalues of PA should all be close to 1.

The Gershgorin circle theorem yields that every eigenvalue of PA lies within a known area.

An interesting application of the Gershgörin circle theorem

When looking for a preconditioner P for matrix A, the eigenvalues of PA should all be close to 1.

The Gershgorin circle theorem yields that every eigenvalue of PA lies within a known area.

We can get an estimate of how good our choice of P is.

Spectral radius

Definition VI.3.4

Let $A \in \mathcal{M}_q(\mathbb{C})$. The **spectral radius** *of A is the non-negative number*

$$\rho(A) = \max\{|\lambda|,\ \lambda \in \sigma(A)\}.$$

Spectral radius

Definition VI.3.4

Let $A \in \mathcal{M}_q(\mathbb{C})$. The **spectral radius** *of A is the non-negative number*

$$\rho(A) = \max\{|\lambda|,\ \lambda \in \sigma(A)\}.$$

Proposition VI.3.5

Let $\|\cdot\|$ be an induced norm on $\mathcal{M}_q(\mathbb{C})$, then

$$\forall A \in \mathcal{M}_q(\mathbb{C}),\ \rho(A) \leq \|A\|$$

Spectral radius

Definition VI.3.4

Let $A \in \mathcal{M}_q(\mathbb{C})$. The **spectral radius** *of A is the non-negative number*

$$\rho(A) = \max\{|\lambda|,\ \lambda \in \sigma(A)\}.$$

Proposition VI.3.5

Let $\|\cdot\|$ be an induced norm on $\mathcal{M}_q(\mathbb{C})$, then

$$\forall A \in \mathcal{M}_q(\mathbb{C}),\ \rho(A) \leq \|A\|$$

Remark VI.3.6

Let $A \in \mathcal{M}_q(\mathbb{C})$, then $\|A\|_2 = \sqrt{\rho(A^* A)}$
Furthermore if A is symmetric positive definite then $\|A\| = \rho(A)$.

Spectral radius: The Power Method

Theorem VI.3.7 (The Power Method)

Let $A \in \mathcal{M}_q$*, and* $\|\cdot\|$ *be a norm on* $\mathbb{K}^q$*,*
Let $\lambda_1, \ldots, \lambda_q$ *be the eigenvalue of* A *with* $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_q|$
Let e_1 *be an eigenvector associated to* λ_1 *and* $F = \mathrm{Im}(A - \lambda_1 I)$*.*
Let $x_0 = \mu e_1 + f$*, with* $\mu \neq 0$ *and* $f \in F$*.*
Define the sequence $(x_n)_{n \in \mathbb{N}}$ *by*

$$x_{n+1} = \frac{Ax_n}{\|Ax_n\|}$$

Then $\lim_{n\to\infty} \|Ax_n\| = \rho(A)$*.*

Spectral radius: The Power Method

Theorem VI.3.7 (The Power Method)

Let $A \in \mathcal{M}_q$, and $\|\cdot\|$ be a norm on $\mathbb{K}^q$,
Let $\lambda_1, \dots, \lambda_q$ be the eigenvalue of A with $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_q|$
Let e_1 be an eigenvector associated to λ_1 and $F = \mathrm{Im}(A - \lambda_1 I)$.
Let $x_0 = \mu e_1 + f$, with $\mu \neq 0$ and $f \in F$.
Define the sequence $(x_n)_{n\in\mathbb{N}}$ by

$$x_{n+1} = \frac{Ax_n}{\|Ax_n\|}$$

Then $\lim_{n\to\infty} \|Ax_n\| = \rho(A)$.

We start with a vector x which has a non-zero component in the direction of an eigenvector associated e_1. We apply A to x and we normalize it. Intuitively, the dominant eigenvalue λ_1 will "pull" x toward the direction of e_1.

Spectral radius: The Power Method (example)

```
import numpy as np

A = np.array([
    [ 5, 2, 0, 1],
    [-1, 3, 2, 1],
    [ 0, 0,-2, 0],
    [ 1,-1, 1, 8]])

x = np.random.rand(A.shape[1])

nb_iterations = 10

for n in range(nb_iterations):
    Ax = np.dot(A, x)
    Ax_norm = np.linalg.norm(Ax)
    x = Ax / Ax_norm;
    print ("||Ax(%d)|| = %f"%(n, Ax_norm));
```

Spectral radius: The Power Method (example)

Example

With the previous matrix A

$$
\begin{aligned}
\|Ax(0)\| &= 3.383218\\
\|Ax(1)\| &= 7.280632\\
\|Ax(2)\| &= 8.182650\\
\|Ax(3)\| &= 8.184859\\
\|Ax(4)\| &= 8.249779\\
\|Ax(5)\| &= 8.258322\\
\|Ax(6)\| &= 8.263641\\
\|Ax(7)\| &= 8.263286\\
\|Ax(8)\| &= 8.262592\\
\|Ax(9)\| &= 8.261866
\end{aligned}
$$

VI.4. Iteration Matrix

Linear Recurrence Relation

Definition VI.4.1 (Linear recurrence relation)

Let $M \in \mathcal{M}_q(\mathbb{K})$ and $b, x_0 \in \mathbb{K}^q$. The sequence $(x_n)_{n\in\mathbb{N}}$ defined by

$$x_{n+1} = M\, x_n + b$$

is called a **linear recurrence relation**.
The matrix M is called the **iteration matrix**.

Linear Recurrence Relation

Definition VI.4.2

Consider a linear recurrence relation with the iteration matrix M and a non-empty set of possible initial conditions $C \subset \mathbb{K}^q$.

$$\begin{cases} x_0 \in C \\ \forall n \geq 0, \quad x_{n+1} = M\, x_n + b \end{cases}$$

The sequence is a **convergent numerical method** *if for all $b \in \mathbb{K}^q$, $(x_n)_{n\in\mathbb{N}}$ converges.*

Convergence error

Definition VI.4.3

Consider a convergent numerical method $(x_n)_{n\in\mathbb{N}}$.
Let $x = \lim x_n$ *and* $e_n = x_n - x$ *pour tout* $n \geq 0$.

The **convergence rate** *of* $(x_n)_{n\in\mathbb{N}}$ *is a measure of the decline of* $(e_n)_{n\in\mathbb{N}}$ *toward 0.*

Linear Recurrence Relation

Example

Let $q = 1$ *and* $m \in \mathbb{C}$.
m *is a* 1×1 *matrix.*
Consider

$$\begin{cases} x_0 \in \mathbb{C} \\ \forall n \geq 0, \quad x_{n+1} = m x_n. \end{cases}$$

This numerical method converges iff $|m| < 1$ *or* $m = 1$.
If $|m| < 1$, *we have* $e_n = m^n x_0$.
The convergence rate is $|m|$.

Convergence of numerical methods

Theorem VI.4.4

Let $M \in \mathcal{M}_q(\mathbb{K})$.
These statements are equivalent:

1. $\forall x \in \mathbb{K}^q$, *the sequence* $(x_n)_{n\in\mathbb{N}}$ *defined by* $x_0 = x$ *and* $\forall n \geq 0$ $x_{n+1} = Mx_n$ *converges toward* 0
2. $\lim_{n\to\infty} M^n = 0$
3. $\rho(M) < 1$
4. *There exists an induced norm* $\|\!| \cdot |\!\|$ *s.t.* $\|\!| M |\!\| < 1$

Applications

Two main applications of iteration matrices will be considered:

- Solving linear systems (next section)
- Parabolic PDEs (chapter VIII)

VI.5. Solving Linear Systems

VI.5.1. Introduction

There are two main types of methods to solve $Ax = b$.

- **The direct methods:**

  A is decomposed in $A = BC$
  $Ax = b$ is replaced by $Cy = b$ and $Bx = y$
  Examples: LU and QR

- **The iterative methods:**

  We build a sequence $(x_n)_{n\in\mathbb{N}}$ that converges toward x.
  Examples: Jacobi, Gauß-Seidel, Conjugate gradient

VI.5.2. Direct methods

LU decomposition

Definition VI.5.1

Let $A \in GL_q(\mathbb{K})$.
A has a **LU decomposition** *if there exist*

- *a lower triangular matrix L with 1 on the diagonal*
- *an upper triangular matrix U*

such that $A = LU$

This works using the Gaussian elimination.
If you are not familiar with the method watch the video by Gilbert Strang at cagnol.link/fact

LU decomposition

- Pros:
  - Very useful if we have to solve $Ax = b$ for several b.
  - Sparse matrices stay sparse in the process.
- Cons: the complexity is q^3.

Remark VI.5.2

We solve the system $Ax = b$. We do not inverse A.

Theorem VI.5.3

There exists a LU decomposition if the gauss elimination works out. In this case it is unique.

Cholesky decomposition

Theorem VI.5.4

Let $A \in GL_q(\mathbb{R})$ symmetric positive definite.
There exists a lower triangular matrix B with with positive diagonal terms such that $B\,B^* = A$.

This applies to the QR decomposition (Gram-Schmidt)

Theorem VI.5.5 (Decomposition QR)

Every matrix $A \in GL_q(\mathbb{R})$ can be uniquely decomposed

$$A = QR$$

where $Q \in O_q(\mathbb{R})$ and R are upper triangular matrices with positive terms on the diagonal.

VI.5.3. Iterative Methods

Definition VI.5.6

Solving a linear system using an iterative method consists of

- *decomposing A in $A = M - N$ where M is a non-singular matrix.*
- *considering the numerical method*

$$\begin{cases} x_0 \in \mathbb{K}^q \\ \forall n \geq 0, \quad Mx_{n+1} = Nx_n + b. \end{cases}$$

Remark VI.5.7

At each iteration, one needs to solve a system. The matrix M needs to be chosen so this system is simple.

Theorem VI.5.8

If $\rho(M^{-1}N) < 1$, the numerical method converges.

Jacobi

Definition VI.5.9 (Jacobi)

Consider $M = diag(A)$ *and* $N = -A + diag(A)$

Gauß-Seidel

Definition VI.5.10 (Gauß-Seidel)

For

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Define

$$M = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \qquad N = -\begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Conjugate Gradient

Definition VI.5.11

Let $A \in \mathcal{M}_q$ be symmetric and positive definite. Consider

- *The sequence $(x_n)_{n\in\mathbb{N}}$ that will converge toward the solution*
- *The sequence $(r_n)_{n\in\mathbb{N}}$ of "residuals" ($r_n = b - Ax_n$)*
- *The sequence $(p_n)_{n\in\mathbb{N}}$ of "directions"*

We initialize with any $x_0 \in \mathbb{R}^q$ (if possible, close to the solution to be found) and $r_0 = b - Ax_0$ and $p_0 = b - Ax_0$. Then, for $n \geq 0$:

- $x_{n+1} = x_n + \alpha_n p_n$
- $r_{n+1} = r_n - \alpha_n A p_n$
- $p_{n+1} = r_{n+1} + \beta_n p_n$

Where α_n and β_n are numbers: $\alpha_n = \frac{\|r_n\|^2}{Ap_n \cdot p_n}$ and $\beta_n = \frac{\|r_{n+1}\|^2}{\|r_n\|^2}$.