MT09 : Chapitre 2
Méthodes directes de résolution des systèmes linéaires

MT09
Vincent.Martin@utc.fr

UTC
Compiègne, France

UTC, A2020

Plan

Plan

Introduction et motivation

- Pourquoi résoudre des systèmes linéaires ?
  - Pour résoudre des systèmes linéaires $Ax = b$ (A matrice $n \times n$) et non-linéaires $f(x) = 0$ (f fonction $\mathbb{R}^n \to \mathbb{R}^n$).
    - dans tous les domaines. Ex : éléments finis (NF04)...
  - Très gros systèmes *creux*
    - $n \geq 100\,000$ lignes : typique.
    - très peu de termes non-nuls par ligne.
  - Systèmes pleins de taille modérée

Introduction et motivation

- Pourquoi résoudre des systèmes linéaires ?
  - Pour résoudre des systèmes linéaires $Ax = b$ (A matrice $n \times n$) et non-linéaires $f(x) = 0$ (f fonction $\mathbb{R}^n \to \mathbb{R}^n$).
    - dans tous les domaines. Ex : éléments finis (NF04)...
  - Très gros systèmes *creux*
    - $n \geq 100\,000$ lignes : typique.
    - très peu de termes non-nuls par ligne.
  - Systèmes pleins de taille modérée
- 2 types de méthodes :

Méthodes directes

$\Longrightarrow$ Chap. 2.
Algorithme : on passe par toutes les lignes de A.
$\mathcal{O}(n)$ opérations fixe au départ.

Méthodes itératives

$\Longrightarrow$ Chap. 4.
Algorithme : suite d'itérés $x^{(k)}$ qui converge vers x.
Nombre infini (?) d'itérations.

Introduction et motivation

- Pourquoi résoudre des systèmes linéaires ?
  - Pour résoudre des systèmes linéaires $Ax = b$ (A matrice $n \times n$) et non-linéaires $f(x) = 0$ (f fonction $\mathbb{R}^n \to \mathbb{R}^n$).
    - dans tous les domaines. Ex : éléments finis (NF04)...
  - Très gros systèmes *creux*
    - $n \geq 100\,000$ lignes : typique.
    - très peu de termes non-nuls par ligne.
  - Systèmes pleins de taille modérée
- 2 types de méthodes :

Méthodes directes

$\Longrightarrow$ Chap. 2.
Algorithme : on passe par toutes les lignes de A.
$\mathcal{O}(n)$ opérations fixe au départ.

Méthodes itératives

$\Longrightarrow$ Chap. 4.
Algorithme : suite d'itérés $x^{(k)}$ qui converge vers x.
Nombre infini (?) d'itérations.

- Les 2 ont leurs avantages... Voir plus tard (Chap 4).

Méthode directe et factorisation

Résoudre $Ax = b$.

- Algorithme de Gauss : voir Section 2.
  - $\Longrightarrow$ obtient une factorisation LU : voir Section 3.

- Factorisation $A = LU$:
  - produit de L triangulaire inférieure (“lower”, “tri-inf”) et U triangulaire supérieure (“upper”, “tri-sup”).
  - Attention : ce n’est PAS un changement de base !
  - $\Longrightarrow$ se ramener à des systèmes plus faciles à résoudre : systèmes triangulaires $Lx = b$ et $Ux = b$

$$Ax = b \iff L(\underbrace{Ux}_{y}) = b \iff \begin{cases} Ly = b \\ Ux = y \end{cases}$$

Comment résoudre $Lx = b$? (L inversible $\Leftrightarrow l_{ii} \neq 0\ \forall i$).

Note : tous les vecteurs de MT09 sont des vecteurs colonne (dans $\mathcal{M}_{n1}(\mathbb{R})$).

$$Lx = b \quad \Leftrightarrow \quad \begin{pmatrix} l_{1,1} & 0 & 0 & \cdots & 0 \\ l_{2,1} & l_{2,2} & 0 & \cdots & 0 \\ l_{3,1} & l_{3,2} & l_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ l_{n,1} & l_{n,2} & l_{n,3} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

$$\Leftrightarrow \quad \left\{ \begin{array}{lcl} l_{1,1}x_1 & = & b_1 \\ l_{2,1}x_1 + l_{2,2}x_2 & = & b_2 \\ l_{3,1}x_1 + l_{3,2}x_2 + l_{3,3}x_3 & = & b_3 \\ \cdots & & \\ l_{n,1}x_1 + l_{n,2}x_2 + l_{n,3}x_3 + \cdots + l_{nn}x_n & = & b_n \end{array} \right.$$

$$\Leftrightarrow \quad \forall i = 1, \ldots, n \quad \sum_{j=1}^{i-1} l_{i,j}x_j + l_{i,i}x_i = b_i$$

$$\Leftrightarrow \quad \forall i = 1, \ldots, n \quad x_i = \frac{1}{l_{i,i}} \left(b_i - \sum_{j=1}^{i-1} l_{i,j}x_j \right), \quad \text{si } l_{i,i} \neq 0.$$

Résolution des systèmes triangulaires $Lx = b$

- Maths :

$$Lx = b \quad \Leftrightarrow \quad \forall i = 1, \dots, n \qquad x_i = \frac{1}{l_{i,i}} \left(b_i - \sum_{j=1}^{i-1} l_{i,j} x_j \right), \quad \text{si } l_{i,i} \neq 0.$$

- Récurrence :
  1. Ligne $i = 1$: $x_1 = \frac{b_1}{l_{1,1}}$ ($\Longrightarrow x_1$ connu)
  2. Ligne $i = 2$: $x_2 = \frac{1}{l_{1,1}}(b_2 - l_{2,1}x_1)$ ($\Longrightarrow x_1, x_2$ connus)
  3. Ligne i : si $x_1, x_2, \dots x_{i-1}$ connus, alors $x_i = \frac{1}{l_{i,i}} \left(b_i - \sum_{j=1}^{i-1} l_{i,j} x_j \right)$
     ($\Longrightarrow x_1, x_2, \dots, x_i$ connus)

Résolution des systèmes triangulaires $Lx = b$

- Maths :

$$Lx = b \quad \Leftrightarrow \quad \forall i = 1, \dots, n \quad x_i = \frac{1}{l_{i,i}} \left(b_i - \sum_{j=1}^{i-1} l_{i,j} x_j \right), \quad \text{si } l_{i,i} \neq 0.$$

Algorithme

```
Pour i=1 à n
    Si L(i,i) = 0
        Sortir, erreur : matrice L non inversible
    Fin Si
    s ⟵ 0
    Pour j=1 à i-1
        s ⟵ s + L(i,j) × x(j)
    Fin Pour
    x(i) ⟵ (b(i) − s) / L(i,i)
Fin Pour
```

Résolution des systèmes triangulaires $Lx = b$

- Maths :

$$Lx = b \quad \Leftrightarrow \quad \forall i = 1, \ldots, n \qquad x_i = \frac{1}{l_{i,i}} \left(b_i - \sum_{j=1}^{i-1} l_{i,j} x_j \right), \qquad \text{si } l_{i,i} \neq 0.$$

Algorithme

```
Pour i=1 à n
    Si L(i,i) = 0   (Attention: flottants!)
       Sortir, erreur : matrice L non inversible
    Fin Si
    s ⟵ 0
    Pour j=1 à i-1
       s ⟵ s + L(i,j) × x(j)
    Fin Pour
    x(i) ⟵ (b(i) − s) / L(i,i)
Fin Pour
```

Factorisation et déterminant

Comment calculer $\det(A)$?

Factorisation et déterminant

Comment calculer $\det(A)$?

- Calcul direct : $\det(A) = \sum_\sigma (-1)^{\varepsilon(\sigma)} \Pi_{i=1}^n \det A_{|i,\sigma(i)|}$:
  - σ : permutations de $\{1, \dots, n\}$, $A_{|i,\sigma(i)|}$, matrice sans ligne i et colonne $\sigma(i)$
  - $\Longrightarrow$ coût en $\mathcal{O}(n!)$
  - matrice 100×100 : calculateur à 100 PetaFlops (10^{17} opérations / s) $\Longrightarrow$ $10^{141} s = 3\ 10^{134}$ ans.
    (âge de l'univers $\approx 14\ 10^9$ ans. . .)
  - calcul analytique : impossible en pratique
    - et en plus : peu précis (accumulation d'erreurs) !

Factorisation et déterminant

Comment calculer $\det(A)$?

- Calcul direct : $\det(A) = \sum_\sigma (-1)^{\varepsilon(\sigma)} \Pi_{i=1}^n \det A_{|i,\sigma(i)|}$:
  - σ : permutations de $\{1, \ldots, n\}$, $A_{|i,\sigma(i)|}$, matrice sans ligne i et colonne $\sigma(i)$
  - $\Longrightarrow$ coût en $\mathcal{O}(n!)$
  - matrice 100×100 : calculateur à 100 PetaFlops (10^{17} opérations / s) $\Longrightarrow$ $10^{141} s = 3\ 10^{134}$ ans.
    (âge de l'univers $\approx 14\ 10^9$ ans...)
  - calcul analytique : impossible en pratique
    - et en plus : peu précis (accumulation d'erreurs) !
- Par la factorisation $A = LU$

$$A = LU = \begin{pmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ \star & l_{22} & 0 & \cdots & 0 \\ \star & \star & l_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \star & \star & \star & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} u_{11} & \star & \star & \cdots & \star \\ 0 & u_{22} & \star & \cdots & \star \\ 0 & 0 & u_{33} & \cdots & \star \\ \vdots & \vdots & \vdots & \ddots & \star \\ 0 & 0 & 0 & \cdots & u_{nn} \end{pmatrix}$$

  - Calculer la factorisation : $\Longrightarrow$ coût en $\mathcal{O}(n^3)$ (voir plus loin)
  - $\det(A) = \det(L) \times \det(U) = \Pi_{i=1}^n l_{ii} u_{ii}$.

Remarques

- Calcul du déterminant : un enseignement fondamental
  - calcul “analytique” : parfois ne marche pas en pratique !
  - méthodes “naïves” souvent vouées à l’échec
  - $\Longrightarrow$ besoin de déterminer le coût (taille mémoire, nombre d’opérations), la stabilité et la robustesse d’un algorithme.

Dicton

En théorie, la pratique et la théorie donnent le même résultat.
En pratique : non.

Plan

Élimination de Gauss : objectifs et notations

Hypothèse : A matrice carrée $n \times n$ inversible, $b \in \mathbb{R}^n$.

Objectif

- écrire l'algorithme : maths et informatique.
- voir les conditions de fonctionnement
- construire la factorisation $A = LU$

Notations

- $\underline{A}_i$: ligne i de A.
- A_j : colonne j de A.
- si $C = AB$, alors $C_j = AB_j$ et $\underline{C}_i = \underline{A}_i B$.

Exemple

$A \in \mathcal{M}_{1,3}$ et $B \in \mathcal{M}_{3,2}$

$$\begin{aligned} AB &= \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,2} \end{pmatrix} \\ &= \left(a_{1,1}b_{1,1} + a_{1,2}b_{2,1} + a_{1,3}b_{3,1} | a_{1,1}b_{1,2} + a_{1,2}b_{2,2} + a_{1,3}b_{3,2} \right) \\ &= a_{1,1}\left(b_{1,1} | b_{1,2} \right) + a_{1,2}\left(b_{2,1} | b_{2,2} \right) + a_{1,3}\left(b_{3,1} | b_{3,2} \right) \\ &= a_{1,1}\underline{B}_1 + a_{1,2}\underline{B}_2 + a_{1,3}\underline{B}_3. \end{aligned}$$

Conclusion :

$$\underline{C}_i = \underline{A}_i B = \sum_{k=1}^{n} a_{ik}\underline{B}_k$$

Lien vers *LU* .

Élimination de Gauss : exemple 3×3 $Ax = b$

$$Ax = b \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix}$$

Élimination de Gauss : exemple 3×3 $Ax = b$

$$A^{(1)} = A.$$

$$Ax = b \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix} \qquad \left\{ \begin{array}{lcl} \underline{A}_1^{(2)} & = & \underline{A}_1^{(1)} \\ \underline{A}_2^{(2)} & = & \underline{A}_2^{(1)} - \frac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \\ \underline{A}_3^{(2)} & = & \underline{A}_3^{(1)} - \frac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \end{array} \right.$$

Élimination de Gauss : exemple 3×3 $Ax = b$

$$Ax = b \Leftrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix}$$

$$\begin{cases} \underline{A}_1^{(2)} = \underline{A}_1^{(1)} \\ \underline{A}_2^{(2)} = \underline{A}_2^{(1)} - \dfrac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \\ \underline{A}_3^{(2)} = \underline{A}_3^{(1)} - \dfrac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \end{cases}$$

$$\Leftrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

Élimination de Gauss : exemple 3×3 $Ax = b$

$$Ax = b \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix}$$

$$\left\{ \begin{array}{lll} \underline{A}_1^{(2)} & = & \underline{A}_1^{(1)} \\ \underline{A}_2^{(2)} & = & \underline{A}_2^{(1)} - \frac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \\ \underline{A}_3^{(2)} & = & \underline{A}_3^{(1)} - \frac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \end{array} \right.$$

$$\Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

$$\left\{ \begin{array}{lll} \underline{A}_1^{(3)} & = & \underline{A}_1^{(2)} \\ \underline{A}_2^{(3)} & = & \underline{A}_2^{(2)} \\ \underline{A}_3^{(3)} & = & \underline{A}_3^{(2)} - \frac{a_{3,2}^{(2)}}{a_{2,2}^{(2)}} \underline{A}_2^{(2)} \end{array} \right.$$

Élimination de Gauss : exemple 3×3 $Ax = b$

$$Ax = b \iff \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix}$$

$$\left\{ \begin{array}{lcl} \underline{A}_1^{(2)} & = & \underline{A}_1^{(1)} \\ \underline{A}_2^{(2)} & = & \underline{A}_2^{(1)} - \dfrac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \\ \underline{A}_3^{(2)} & = & \underline{A}_3^{(1)} - \dfrac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \end{array} \right.$$

$$\iff \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

$$\left\{ \begin{array}{lcl} \underline{A}_1^{(3)} & = & \underline{A}_1^{(2)} \\ \underline{A}_2^{(3)} & = & \underline{A}_2^{(2)} \\ \underline{A}_3^{(3)} & = & \underline{A}_3^{(2)} - \dfrac{a_{3,2}^{(2)}}{a_{2,2}^{(2)}} \underline{A}_2^{(2)} \end{array} \right.$$

$$\iff \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ -9 \end{bmatrix}$$

Élimination de Gauss : exemple 3×3 $Ax = b$

$$Ax = b \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ 3 \end{bmatrix}$$

$$\begin{cases} \underline{A}_1^{(2)} = \underline{A}_1^{(1)} \\ \underline{A}_2^{(2)} = \underline{A}_2^{(1)} - \frac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \\ \underline{A}_3^{(2)} = \underline{A}_3^{(1)} - \frac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} \underline{A}_1^{(1)} \end{cases}$$

$$\Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

$$\begin{cases} \underline{A}_1^{(3)} = \underline{A}_1^{(2)} \\ \underline{A}_2^{(3)} = \underline{A}_2^{(2)} \\ \underline{A}_3^{(3)} = \underline{A}_3^{(2)} - \frac{a_{3,2}^{(2)}}{a_{2,2}^{(2)}} \underline{A}_2^{(2)} \end{cases}$$

$$\Leftrightarrow \quad \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ -9 \end{bmatrix} \quad \Leftrightarrow \quad A^{(3)} x = b^{(3)}$$

$$A = LU \text{ avec } U = A^{(3)} \text{ et } L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{a_{2,1}^{(1)}}{a_{1,1}^{(1)}} & 1 & 0 \\ \frac{a_{3,1}^{(1)}}{a_{1,1}^{(1)}} & \frac{a_{3,2}^{(2)}}{a_{2,2}^{(2)}} & 1 \end{bmatrix}$$

Élimination de Gauss : cas général

$$A^{(k)} = \begin{bmatrix}
a_{1,1}^{(1)} & a_{1,2}^{(1)} & a_{1,3}^{(1)} & \dots & a_{1,k-1}^{(1)} & a_{1,k}^{(1)} & a_{1,k+1}^{(1)} & a_{1,k+2}^{(1)} & \dots & a_{1,n}^{(1)} \\
0 & a_{2,2}^{(2)} & a_{2,3}^{(2)} & \dots & a_{2,k-1}^{(2)} & a_{2,k}^{(2)} & a_{2,k+1}^{(2)} & a_{2,k+2}^{(2)} & \dots & a_{2,n}^{(2)} \\
0 & 0 & a_{3,3}^{(3)} & \dots & a_{3,k-1}^{(3)} & a_{3,k}^{(3)} & a_{3,k+1}^{(3)} & a_{3,k+2}^{(3)} & \dots & a_{3,n}^{(3)} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & a_{k-1,k+1}^{(k-1)} & a_{k-1,k+2}^{(k-1)} & \dots & a_{k-1,n}^{(k-1)} \\
0 & 0 & 0 & \dots & 0 & a_{k,k}^{(k)} & a_{k,k+1}^{(k)} & a_{k,k+2}^{(k)} & \dots & a_{k,n}^{(k)} \\
0 & 0 & 0 & \dots & 0 & a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & a_{k+1,k+2}^{(k)} & \dots & a_{k+1,n}^{(k)} \\
0 & 0 & 0 & \dots & 0 & a_{k+2,k}^{(k)} & a_{k+2,k+1}^{(k)} & a_{k+2,k+2}^{(k)} & \dots & a_{k+2,n}^{(k)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & 0 & a_{n,k}^{(k)} & a_{n,k+1}^{(k)} & a_{n,k+2}^{(k)} & \dots & a_{n,n}^{(k)}
\end{bmatrix}$$

Élimination de Gauss : cas général

$$
A^{(k+1)} = \begin{bmatrix}
a_{1,1}^{(1)} & a_{1,2}^{(1)} & a_{1,3}^{(1)} & \dots & a_{1,k-1}^{(1)} & a_{1,k}^{(1)} & a_{1,k+1}^{(1)} & a_{1,k+2}^{(1)} & \dots & a_{1,n}^{(1)} \\
0 & a_{2,2}^{(2)} & a_{2,3}^{(2)} & \dots & a_{2,k-1}^{(2)} & a_{2,k}^{(2)} & a_{2,k+1}^{(2)} & a_{2,k+2}^{(2)} & \dots & a_{2,n}^{(2)} \\
0 & 0 & a_{3,3}^{(3)} & \dots & a_{3,k-1}^{(3)} & a_{3,k}^{(3)} & a_{3,k+1}^{(3)} & a_{3,k+2}^{(3)} & \dots & a_{3,n}^{(3)} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & a_{k-1,k+1}^{(k-1)} & a_{k-1,k+2}^{(k-1)} & \dots & a_{k-1,n}^{(k-1)} \\
0 & 0 & 0 & \dots & 0 & a_{k,k}^{(k)} & a_{k,k+1}^{(k)} & a_{k,k+2}^{(k)} & \dots & a_{k,n}^{(k)} \\
0 & 0 & 0 & \dots & 0 & 0 & a_{k+1,k+1}^{(k+1)} & a_{k+1,k+2}^{(k+1)} & \dots & a_{k+1,n}^{(k+1)} \\
0 & 0 & 0 & \dots & 0 & 0 & a_{k+2,k+1}^{(k+1)} & a_{k+2,k+2}^{(k+1)} & \dots & a_{k+2,n}^{(k+1)} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & 0 & 0 & a_{n,k+1}^{(k+1)} & a_{n,k+2}^{(k+1)} & \dots & a_{n,n}^{(k+1)}
\end{bmatrix}
$$

Élimination de Gauss : cas général

$$A^{(k+1)} = \begin{bmatrix}
a^{(1)}_{1,1} & a^{(1)}_{1,2} & a^{(1)}_{1,3} & \dots & a^{(1)}_{1,k-1} & a^{(1)}_{1,k} & a^{(1)}_{1,k+1} & a^{(1)}_{1,k+2} & \dots & a^{(1)}_{1,n} \\
0 & a^{(2)}_{2,2} & a^{(2)}_{2,3} & \dots & a^{(2)}_{2,k-1} & a^{(2)}_{2,k} & a^{(2)}_{2,k+1} & a^{(2)}_{2,k+2} & \dots & a^{(2)}_{2,n} \\
0 & 0 & a^{(3)}_{3,3} & \dots & a^{(3)}_{3,k-1} & a^{(3)}_{3,k} & a^{(3)}_{3,k+1} & a^{(3)}_{3,k+2} & \dots & a^{(3)}_{3,n} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & a^{(k-1)}_{k-1,k-1} & a^{(k-1)}_{k-1,k} & a^{(k-1)}_{k-1,k+1} & a^{(k-1)}_{k-1,k+2} & \dots & a^{(k-1)}_{k-1,n} \\
0 & 0 & 0 & \dots & 0 & a^{(k)}_{k,k} & a^{(k)}_{k,k+1} & a^{(k)}_{k,k+2} & \dots & a^{(k)}_{k,n} \\
0 & 0 & 0 & \dots & 0 & 0 & a^{(k+1)}_{k+1,k+1} & a^{(k+1)}_{k+1,k+2} & \dots & a^{(k+1)}_{k+1,n} \\
0 & 0 & 0 & \dots & 0 & 0 & a^{(k+1)}_{k+2,k+1} & a^{(k+1)}_{k+2,k+2} & \dots & a^{(k+1)}_{k+2,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\
0 & 0 & 0 & \dots & 0 & 0 & a^{(k+1)}_{n,k+1} & a^{(k+1)}_{n,k+2} & \dots & a^{(k+1)}_{n,n}
\end{bmatrix}$$

$$A^{(k)} \longrightarrow A^{(k+1)} : \quad \left\{ \begin{array}{lcl}
\underline{A}^{(k+1)}_1 & = & \underline{A}^{(k)}_1 \\
\underline{A}^{(k+1)}_2 & = & \underline{A}^{(k)}_2 \\
\dots & & \\
\underline{A}^{(k+1)}_{k-1} & = & \underline{A}^{(k)}_{k-1} \\
\underline{A}^{(k+1)}_k & = & \underline{A}^{(k)}_k
\end{array} \right. \quad
\left\{ \begin{array}{lcl}
\underline{A}^{(k+1)}_{k+1} & = & \underline{A}^{(k)}_{k+1} - \dfrac{a^{(k)}_{k+1,k}}{a^{(k)}_{k,k}} \underline{A}^{(k)}_k \\
\underline{A}^{(k+1)}_{k+2} & = & \underline{A}^{(k)}_{k+2} - \dfrac{a^{(k)}_{k+2,k}}{a^{(k)}_{k,k}} \underline{A}^{(k)}_k \\
\dots & & \\
\underline{A}^{(k+1)}_n & = & \underline{A}^{(k)}_n - \dfrac{a^{(k)}_{n,k}}{a^{(k)}_{k,k}} \underline{A}^{(k)}_k
\end{array} \right.$$

Élimination de Gauss : méthode

$$A^{(k)} \longrightarrow A^{(k+1)} : \left\{ \begin{array}{lcl} \underline{A}_1^{(k+1)} & = & \underline{A}_1^{(k)} \\ \underline{A}_2^{(k+1)} & = & \underline{A}_2^{(k)} \\ \dots & & \\ \underline{A}_{k-1}^{(k+1)} & = & \underline{A}_{k-1}^{(k)} \\ \underline{A}_k^{(k+1)} & = & \underline{A}_k^{(k)} \end{array} \right. \quad \left\{ \begin{array}{lcl} \underline{A}_{k+1}^{(k+1)} & = & \underline{A}_{k+1}^{(k)} - \frac{a_{k+1,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \\ \underline{A}_{k+2}^{(k+1)} & = & \underline{A}_{k+2}^{(k)} - \frac{a_{k+2,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \\ \dots & & \\ \underline{A}_n^{(k+1)} & = & \underline{A}_n^{(k)} - \frac{a_{n,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \end{array} \right.$$

Donc : on pose $A^{(1)} = A$, puis on calcule

$$\forall k = 1, \dots, n-1 \ : \left\{ \begin{array}{lcll} \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} & \forall i = 1, \dots, k \\ \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} - \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} & \forall i = k+1, \dots, n \end{array} \right.$$

On obtient $U = A^{(n)}$, triangulaire supérieure. (On fait pareil pour b).

$$A^{(k)} \longrightarrow A^{(k+1)} : \left\{\begin{array}{ccc} \underline{A}_1^{(k+1)} & = & \underline{A}_1^{(k)} \\ \underline{A}_2^{(k+1)} & = & \underline{A}_2^{(k)} \\ \dots & & \\ \underline{A}_{k-1}^{(k+1)} & = & \underline{A}_{k-1}^{(k)} \\ \underline{A}_k^{(k+1)} & = & \underline{A}_k^{(k)} \end{array}\right. \quad \left\{\begin{array}{ccc} \underline{A}_{k+1}^{(k+1)} & = & \underline{A}_{k+1}^{(k)} - \frac{a_{k+1,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \\ \underline{A}_{k+2}^{(k+1)} & = & \underline{A}_{k+2}^{(k)} - \frac{a_{k+2,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \\ \dots & & \\ \underline{A}_n^{(k+1)} & = & \underline{A}_n^{(k)} - \frac{a_{n,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} \end{array}\right.$$

Donc : on pose $A^{(1)} = A$, puis on calcule

$$\forall k = 1, \dots, n-1 \ : \left\{\begin{array}{cccl} \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} & \forall i = 1, \dots, k \\ \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} - \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}} \underline{A}_k^{(k)} & \forall i = k+1, \dots, n \end{array}\right.$$

On obtient $U = A^{(n)}$, triangulaire supérieure. (On fait pareil pour b).

Notation

On pose $l_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$ pour $i = k+1, \dots, n$.

Élimination de Gauss : algorithme

Algorithme d'élimination de Gauss

```
Pour k=1 jusqu'à n-1
    Si |A(k,k)|< tol
        Sortir, erreur : pivot nul
   Sinon
        Pour i=k+1 jusqu'à n
            c ⟵ A(i,k) / A(k,k)
            b(i) ⟵ b(i) - c × b(k)
            A(i,k) ⟵ 0
            Pour j=k+1 jusqu'à n
                A(i,j) ⟵ A(i,j) - c × A(k,j)
            Fin Pour
        Fin Pour
   Fin Si
Fin Pour
Si |A(n,n)| < tol
    Sortir, erreur : pivot nul
Fin Si
```

Élimination de Gauss : algorithme, coût

Calcul du nombre de multiplications et divisions :
Calcul similaire pour les additions.
Commencer par les boucles les plus internes.

- 2ème boucle interne (en j) : $n-k$ multiplications.
- 1ère boucle interne (en i) : $(n-k)^2+2(n-k)$ multiplications et divisions.
- boucle externe (en k) : $\sum_{k=1}^{n-1}(n-k)^2+2(n-k)$ multiplications et divisions.

Calcul du nombre de multiplications et divisions :

Calcul similaire pour les additions.
Commencer par les boucles les plus internes.

- 2ème boucle interne (en j) : $n-k$ multiplications.
- 1ère boucle interne (en i) : $(n-k)^2+2(n-k)$ multiplications et divisions.
- boucle externe (en k) : $\sum_{k=1}^{n-1}(n-k)^2+2(n-k)$ multiplications et divisions.

Coût total :

$$C=\sum_{k=1}^{n-1}(n-k)^2+2(n-k)=\sum_{p=1}^{n-1}p^2+2p=\frac{(n-1)n(2n-1)}{6}+n(n-1)\sim\frac{n^3}{3}.$$

(Rappel : $\displaystyle\sum_{p=1}^{n}p^2=\frac{n(n+1)(2n+1)}{6}$ et $\displaystyle\sum_{p=1}^{n}p=\frac{n(n+1)}{2}$.)

Calcul du nombre de multiplications et divisions :
Calcul similaire pour les additions.
Commencer par les boucles les plus internes.

- 2ème boucle interne (en j) : $n-k$ multiplications.
- 1ère boucle interne (en i) : $(n-k)^2+2(n-k)$ multiplications et divisions.
- boucle externe (en k) : $\sum_{k=1}^{n-1}(n-k)^2+2(n-k)$ multiplications et divisions.

Coût total :

$$C=\sum_{k=1}^{n-1}(n-k)^2+2(n-k)=\sum_{p=1}^{n-1}p^2+2p=\frac{(n-1)n(2n-1)}{6}+n(n-1)\sim\frac{n^3}{3}.$$

(Rappel : $\sum_{p=1}^{n}p^2=\frac{n(n+1)(2n+1)}{6}$ et $\sum_{p=1}^{n}p=\frac{n(n+1)}{2}$.)

Théorème (Coût computationnel de Gauss)

Le coût de l'algorithme de Gauss est en $\frac{n^3}{3}$ multiplications.

Élimination de Gauss : factorisation $A = LU$

On pose $A^{(1)} = A$, puis on calcule

$$\forall k = 1, \ldots, n-1 \ : \left\{ \begin{array}{lll} \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} \qquad \forall i = 1, \ldots, k \\ \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} - l_{i,k}\underline{A}_k^{(k)} \qquad \forall i = k+1, \ldots, n \end{array} \right. \qquad \text{avec } l_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$$

On fixe une ligne i et on regarde les modifications :

$$\left\{ \begin{array}{lll} \underline{A}_i^{(1)} & = & \underline{A}_i \\ \underline{A}_i^{(2)} & = & \underline{A}_i^{(1)} - l_{i,1}\underline{A}_1^{(1)} \\ \underline{A}_i^{(3)} & = & \underline{A}_i^{(2)} - l_{i,2}\underline{A}_2^{(2)} \\ \ldots & & \\ \underline{A}_i^{(i-1)} & = & \underline{A}_i^{(i-2)} - l_{i,i-2}\underline{A}_{i-2}^{(i-2)} \\ \underline{A}_i^{(i)} & = & \underline{A}_i^{(i-1)} - l_{i,i-1}\underline{A}_{i-1}^{(i-1)} \\ \underline{A}_i^{(i+1)} & = & \underline{A}_i^{(i)} \\ \underline{A}_i^{(i+2)} & = & \underline{A}_i^{(i+1)} \\ \ldots & & \\ \underline{A}_i^{(n-1)} & = & \underline{A}_i^{(n-2)} \\ \underline{A}_i^{(n)} & = & \underline{A}_i^{(n-1)} \end{array} \right.$$

Élimination de Gauss : factorisation $A = LU$

On pose $A^{(1)} = A$, puis on calcule

$$\forall k = 1, \ldots, n-1 \ : \left\{ \begin{array}{lcll} \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} & \forall i = 1, \ldots, k \\ \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} - l_{i,k}\underline{A}_k^{(k)} & \forall i = k+1, \ldots, n \end{array} \right. \qquad \text{avec } l_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$$

On fixe une ligne i et on regarde les modifications :

$$\left\{ \begin{array}{lcl} \underline{A}_i^{(1)} & = & \underline{A}_i \\ \underline{A}_i^{(2)} & = & \underline{A}_i^{(1)} - l_{i,1}\underline{A}_1^{(1)} \\ \underline{A}_i^{(3)} & = & \underline{A}_i^{(2)} - l_{i,2}\underline{A}_2^{(2)} \\ \ldots & & \\ \underline{A}_i^{(i-1)} & = & \underline{A}_i^{(i-2)} - l_{i,i-2}\underline{A}_{i-2}^{(i-2)} \\ \underline{A}_i^{(i)} & = & \underline{A}_i^{(i-1)} - l_{i,i-1}\underline{A}_{i-1}^{(i-1)} \\ \underline{A}_i^{(i+1)} & = & \underline{A}_i^{(i)} \\ \underline{A}_i^{(i+2)} & = & \underline{A}_i^{(i+1)} \\ \ldots & & \\ \underline{A}_i^{(n-1)} & = & \underline{A}_i^{(n-2)} \\ \underline{A}_i^{(n)} & = & \underline{A}_i^{(n-1)} \end{array} \right.$$

On somme et on élimine les termes :

$$\left\{ \begin{array}{lcl} \underline{A}_i^{(n)} & = & \underline{A}_i - l_{i,1}\underline{A}_1^{(1)} - l_{i,2}\underline{A}_2^{(2)} \\ & & \ldots \\ & & -l_{i,i-2}\underline{A}_{i-2}^{(i-2)} - l_{i,i-1}\underline{A}_{i-1}^{(i-1)} \end{array} \right.$$

Élimination de Gauss : factorisation $A = LU$

On pose $A^{(1)} = A$, puis on calcule

$$\forall k = 1, \dots, n-1 \ : \left\{ \begin{array}{lll} \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} \\ \underline{A}_i^{(k+1)} & = & \underline{A}_i^{(k)} - l_{i,k}\underline{A}_k^{(k)} \end{array} \right. \begin{array}{l} \forall i = 1, \dots, k \\ \forall i = k+1, \dots, n \end{array} \qquad \text{avec } l_{i,k} = \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$$

On fixe une ligne i et on regarde les modifications :

$$\left\{ \begin{array}{lll} \underline{A}_i^{(1)} & = & \underline{A}_i \\ \underline{A}_i^{(2)} & = & \underline{A}_i^{(1)} - l_{i,1}\underline{A}_1^{(1)} \\ \underline{A}_i^{(3)} & = & \underline{A}_i^{(2)} - l_{i,2}\underline{A}_2^{(2)} \\ \dots & & \\ \underline{A}_i^{(i-1)} & = & \underline{A}_i^{(i-2)} - l_{i,i-2}\underline{A}_{i-2}^{(i-2)} \\ \underline{A}_i^{(i)} & = & \underline{A}_i^{(i-1)} - l_{i,i-1}\underline{A}_{i-1}^{(i-1)} \\ \underline{A}_i^{(i+1)} & = & \underline{A}_i^{(i)} \\ \underline{A}_i^{(i+2)} & = & \underline{A}_i^{(i+1)} \\ \dots & & \\ \underline{A}_i^{(n-1)} & = & \underline{A}_i^{(n-2)} \\ \underline{A}_i^{(n)} & = & \underline{A}_i^{(n-1)} \end{array} \right.$$

On somme et on élimine les termes :

$$\left\{ \begin{array}{lll} \underline{A}_i^{(n)} & = & \underline{A}_i - l_{i,1}\underline{A}_1^{(1)} - l_{i,2}\underline{A}_2^{(2)} \\ & & \dots \\ & & -l_{i,i-2}\underline{A}_{i-2}^{(i-2)} - l_{i,i-1}\underline{A}_{i-1}^{(i-1)} \end{array} \right.$$

ce qui donne

$$\begin{array}{lll} \underline{A}_i & = & \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(j)} \\ & = & \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(n)} \end{array}$$

car $\underline{A}_j^{(j)} = \underline{A}_j^{(j+1)} = \dots = \underline{A}_j^{(n)}$.

Élimination de Gauss : factorisation $A = LU$

On a pour toute ligne i :

$$\underline{A}_i = \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(n)} = l_{i,1}\underline{A}_1^{(n)} + l_{i,2}\underline{A}_2^{(n)} + \cdots + l_{i,i-1}\underline{A}_{i-1}^{(n)} + \underline{A}_i^{(n)}$$

Élimination de Gauss : factorisation $A = LU$

On a pour toute ligne i :

$$\begin{aligned}
\underline{A}_i &= \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(n)} = l_{i,1}\underline{A}_1^{(n)} + l_{i,2}\underline{A}_2^{(n)} + \cdots + l_{i,i-1}\underline{A}_{i-1}^{(n)} + \underline{A}_i^{(n)} \\
&= \begin{bmatrix} l_{i,1} & l_{i,2} & \ldots & l_{i,i-1} & 1 \end{bmatrix} \times \begin{bmatrix} \underline{A}_1^{(n)} \\ \underline{A}_2^{(n)} \\ \ldots \\ \underline{A}_{i-1}^{(n)} \\ \underline{A}_i^{(n)} \end{bmatrix}
\end{aligned}$$

Élimination de Gauss : factorisation $A = LU$

On a pour toute ligne i :

$$\begin{aligned}
\underline{A}_i &= \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(n)} = l_{i,1}\underline{A}_1^{(n)} + l_{i,2}\underline{A}_2^{(n)} + \cdots + l_{i,i-1}\underline{A}_{i-1}^{(n)} + \underline{A}_i^{(n)} \\
&= \begin{bmatrix} l_{i,1} & l_{i,2} & \dots & l_{i,i-1} & 1 \end{bmatrix} \times \begin{bmatrix} \underline{A}_1^{(n)} \\ \underline{A}_2^{(n)} \\ \dots \\ \underline{A}_{i-1}^{(n)} \\ \underline{A}_i^{(n)} \end{bmatrix} \\
&= \begin{bmatrix} l_{i,1} & l_{i,2} & \dots & l_{i,i-1} & 1 & 0 & \dots & 0 \end{bmatrix} \times \begin{bmatrix} \underline{A}_1^{(n)} \\ \underline{A}_2^{(n)} \\ \dots \\ \underline{A}_{i-1}^{(n)} \\ \underline{A}_i^{(n)} \\ \underline{A}_{i+1}^{(n)} \\ \dots \\ \underline{A}_n^{(n)} \end{bmatrix}
\end{aligned}$$

Élimination de Gauss : factorisation $A = LU$

On a pour toute ligne i :

$$\begin{aligned}
\underline{A}_i &= \underline{A}_i^{(n)} + \sum_{j=1}^{i-1} l_{i,j}\underline{A}_j^{(n)} = l_{i,1}\underline{A}_1^{(n)} + l_{i,2}\underline{A}_2^{(n)} + \cdots + l_{i,i-1}\underline{A}_{i-1}^{(n)} + \underline{A}_i^{(n)} \\
&= \begin{bmatrix} l_{i,1} & l_{i,2} & \dots & l_{i,i-1} & 1 \end{bmatrix} \times \begin{bmatrix} \underline{A}_1^{(n)} \\ \underline{A}_2^{(n)} \\ \dots \\ \underline{A}_{i-1}^{(n)} \\ \underline{A}_i^{(n)} \end{bmatrix} \\
&= \begin{bmatrix} l_{i,1} & l_{i,2} & \dots & l_{i,i-1} & 1 & 0 & \dots & 0 \end{bmatrix} \times \begin{bmatrix} \underline{A}_1^{(n)} \\ \underline{A}_2^{(n)} \\ \dots \\ \underline{A}_{i-1}^{(n)} \\ \underline{A}_i^{(n)} \\ \underline{A}_{i+1}^{(n)} \\ \dots \\ \underline{A}_n^{(n)} \end{bmatrix} \\
&= \underline{L}_i \times A^{(n)} \qquad \text{et } \underline{L}_i \text{ nulle à droite du terme de la diagonale...}
\end{aligned}$$

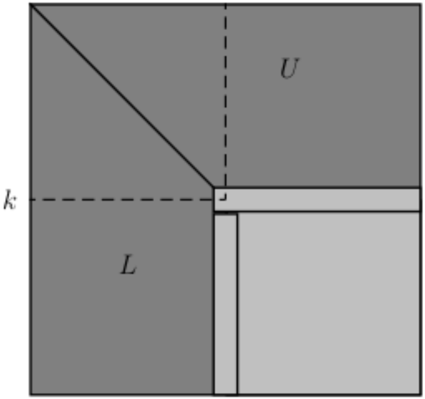
Élimination de Gauss : factorisation $A = LU$

Conclusion

$$\underline{A}_i = \underline{L}_i \times A^n \quad \forall i = 1, \ldots, n \qquad \Longleftrightarrow \qquad A = L \times A^n = LU,$$

à condition que les pivots $a_{k,k}^{(k)}$ soient tous non nuls.
La matrice L est triangulaire inférieure avec des 1 sur la diagonale.
La matrice U est triangulaire supérieure avec les pivots $a_{k,k}^{(k)}$ sur la diagonale.
Attention : les pivots ne sont pas les termes diagonaux : $a_{k,k}^{(k)} \neq a_{k,k}$.

Élimination de Gauss : aspects pratiques



Note : en pratique, on ne calcule pas les termes qui sont nuls. On ne modifie que la partie carrée en bas à droite de taille $n - k$.

En pratique, on n'a besoin que d'une seule matrice (on utilise l'espace mémoire de A, qui est écrasée). On stocke L dans la partie inférieure stricte (les 1 sur la diagonale ne sont pas stockés) et on stocke U dans la partie supérieure.

Élimination de Gauss : existence et unicité de $A = LU$

Théorème (Existence et unicité de la factorisation LU)

Soit A inversible. Si les pivots sont tous non-nuls pendant l'élimination de Gauss, alors il existe une matrice L triangulaire inférieure et une matrice U triangulaire supérieure inversibles, telles que $A = LU$. De plus, si la diagonale de L est constituée de 1, alors la factorisation est unique.

Démonstration : existence : faite jusqu'ici.
Unicité : supposons que $A = LU = \widetilde{L}\widetilde{U}$, telles que L et $\widetilde{L}$ sont tri-inf avec $\mathrm{diag}(L) = \mathrm{diag}(\widetilde{L}) = 1$, et U et $\widetilde{U}$ sont tri-sup.
L, $\widetilde{L}$, U et $\widetilde{U}$ sont inversibles car $\det(A)$ est non nul.
Donc $LU = \widetilde{L}\widetilde{U} \iff U = L^{-1}\widetilde{L}\widetilde{U} \iff U\widetilde{U}^{-1} = L^{-1}\widetilde{L}$.
(Attention, le produit matriciel n'est pas commutatif !)
$U\widetilde{U}^{-1}$ est tri-sup, $L^{-1}\widetilde{L}$ est tri-inf, donc $U\widetilde{U}^{-1} = L^{-1}\widetilde{L}$ est diagonale.
De plus $\mathrm{diag}(L^{-1}) = \mathrm{diag}(\widetilde{L}) = 1$, donc $L^{-1}\widetilde{L} = I$ et $L = \widetilde{L}$. Et donc $U = \widetilde{U}$.
Il y a unicité.

Remarques sur les matrices triangulaires

Lemme

Soient A et B deux matrices triangulaires supérieures.
Alors le produit $C = AB$ est triangulaire supérieur.
De plus $c_{i,i} = a_{i,i} b_{i,i}$ pour $i = 1, \ldots, n$.
Si A est inversible, alors l'inverse $D = A^{-1}$ est triangulaire supérieure.
De plus, $d_{i,i} = \dfrac{1}{a_{i,i}}$ pour $i = 1, \ldots, n$.

La proposition similaire pour les matrices tri-inf est également vraie.

Remarques sur les matrices triangulaires

Lemme

Soient A et B deux matrices triangulaires supérieures.
Alors le produit $C = AB$ est triangulaire supérieur.
De plus $c_{i,i} = a_{i,i}b_{i,i}$ pour $i = 1, \dots, n$.
Si A est inversible, alors l'inverse $D = A^{-1}$ est triangulaire supérieure.
De plus, $d_{i,i} = \dfrac{1}{a_{i,i}}$ pour $i = 1, \dots, n$.

La proposition similaire pour les matrices tri-inf est également vraie.
Voir l'exercice Chapitre 0, TD 11 (Exercice B.1.11 TD0-Exercice11).
Montrer que si A et B sont tri-sup, alors :

$$C_{i,j} = \sum_{k=i}^{j} A_{i,k} B_{k,j} \; \forall i, j = 1, \dots, n, \quad (\text{et } C_{i,j} = 0 \text{ si } i > j.)$$

Convention : une somme faite sur un ensemble d'indice vide est nulle.

Trouver la formule similaire pour A et B tri-inf.

Plan

Factorisation LU directe : principe

Unicité $\Longrightarrow$ calcul direct par identification.
Si possible, c'est **LA** factorisation $A = LU$.

Lien vers prod mat.

$$A = LU \quad \Longleftrightarrow \quad \left\{ \; \underline{A}_i = \underline{L}_i U \quad \forall i = 1, \dots, n \right.$$

$$\Longleftrightarrow \quad \left\{ \; A_j = L U_j \quad \forall j = 1, \dots, n \right.$$

$$\Longleftrightarrow \quad \left\{ \; A_{i,j} = \sum_{k=1}^{\min(i,j)} L_{i,k} U_{k,j} \quad \forall i, j = 1, \dots, n \right.$$

Factorisation LU directe : principe

Unicité $\Longrightarrow$ calcul direct par identification.
Si possible, c'est **LA** factorisation $A = LU$.

Lien vers prod mat.

$$A = LU \iff \left\{ \underline{A}_i = \underline{L}_i U \quad \forall i = 1, \dots, n \right. \\ \iff \left\{ A_j = L U_j \quad \forall j = 1, \dots, n \right. \\ \iff \left\{ A_{i,j} = \sum_{k=1}^{\min(i,j)} L_{i,k} U_{k,j} \quad \forall i,j = 1, \dots, n \right.$$

Donc principe : identifier en alternant les lignes et les colonnes de A.

Calcul matriciel ligne / colonne

Si $A \in \mathcal{M}_{p,n}$, $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^p$, on a :

$$\begin{aligned} Ax &= A_1 x_1 + A_2 x_2 + \cdots + A_n x_n \\ y^T A &= y_1 \underline{A}_1 + y_2 \underline{A}_2 + \cdots + y_p \underline{A}_p \end{aligned}$$

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow$ $\underline{U}_1 = \underline{A}_1$ connu.

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow \underline{U}_1 = \underline{A}_1$ connu.
- Colonne $j = 1$: $A_1 = LU_1 = L \begin{bmatrix} u_{1,1} \\ 0 \\ 0 \end{bmatrix} = L_1 u_{1,1}$ $\Longrightarrow L_1 = \dfrac{A_1}{u_{1,1}}$ connu.

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow$ $\underline{U}_1 = \underline{A}_1$ connu.
- Colonne $j = 1$: $A_1 = LU_1 = L \begin{bmatrix} u_{1,1} \\ 0 \\ 0 \end{bmatrix} = L_1 u_{1,1}$ $\Longrightarrow$ $L_1 = \dfrac{A_1}{u_{1,1}}$ connu.

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

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow \underline{U}_1 = \underline{A}_1$ connu.
- Colonne $j = 1$: $A_1 = LU_1 = L \begin{bmatrix} u_{1,1} \\ 0 \\ 0 \end{bmatrix} = L_1 u_{1,1}$ $\Longrightarrow L_1 = \dfrac{A_1}{u_{1,1}}$ connu.

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

- Ligne $i = 2$: $\underline{A}_2 = \underline{L}_2 U = \begin{bmatrix} l_{2,1} & 1 & 0 \end{bmatrix} U = l_{2,1} \underline{U}_1 + \underline{U}_2$
  $\Longrightarrow \underline{U}_2 = \underline{A}_2 - l_{2,1} \underline{U}_1$ connu.

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow \underline{U}_1 = \underline{A}_1$ connu.
- Colonne $j = 1$: $A_1 = LU_1 = L \begin{bmatrix} u_{1,1} \\ 0 \\ 0 \end{bmatrix} = L_1 u_{1,1}$ $\Longrightarrow L_1 = \dfrac{A_1}{u_{1,1}}$ connu.

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

- Ligne $i = 2$: $\underline{A}_2 = \underline{L}_2 U = \begin{bmatrix} l_{2,1} & 1 & 0 \end{bmatrix} U = l_{2,1}\underline{U}_1 + \underline{U}_2$ $\Longrightarrow \underline{U}_2 = \underline{A}_2 - l_{2,1}\underline{U}_1$ connu.
- Colonne $j = 2$: $A_2 = LU_2 = L \begin{bmatrix} u_{1,2} \\ u_{2,2} \\ 0 \end{bmatrix} = L_1 u_{1,2} + L_2 u_{2,2}$

$$\Longrightarrow L_2 = \frac{1}{u_{2,2}} \left(A_2 - L_1 u_{1,2} \right) \text{ connu.}$$

Factorisation LU directe : exemple 3×3

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \star & 1 & 0 \\ \star & \star & 1 \end{bmatrix} \times \begin{bmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{bmatrix} = LU???$$

- Ligne $i = 1$: identifier $\underline{A}_1 = \underline{L}_1 U = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} U = \underline{U}_1$ $\Longrightarrow$ $\underline{U}_1 = \underline{A}_1$ connu.
- Colonne $j = 1$: $A_1 = LU_1 = L \begin{bmatrix} u_{1,1} \\ 0 \\ 0 \end{bmatrix} = L_1 u_{1,1}$ $\Longrightarrow$ $L_1 = \dfrac{A_1}{u_{1,1}}$ connu.

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

- Ligne $i = 2$: $\underline{A}_2 = \underline{L}_2 U = \begin{bmatrix} l_{2,1} & 1 & 0 \end{bmatrix} U = l_{2,1} \underline{U}_1 + \underline{U}_2$ $\Longrightarrow$ $\underline{U}_2 = \underline{A}_2 - l_{2,1} \underline{U}_1$ connu.
- Colonne $j = 2$: $A_2 = LU_2 = L \begin{bmatrix} u_{1,2} \\ u_{2,2} \\ 0 \end{bmatrix} = L_1 u_{1,2} + L_2 u_{2,2}$

$\Longrightarrow$ $L_2 = \dfrac{1}{u_{2,2}} \left(A_2 - L_1 u_{1,2} \right)$ connu.

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

Factorisation LU directe : exemple 3×3

- Ligne $i = 3$: $\underline{A}_3 = \begin{bmatrix} l_{3,1} & l_{3,2} & 1 \end{bmatrix} U = l_{3,1}\underline{U}_1 + l_{3,2}\underline{U}_2 + \underline{U}_3$

  $\Longrightarrow \underline{U}_3 = \underline{A}_3 - l_{3,1}\underline{U}_1 - l_{3,2}\underline{U}_2$ connu.

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & -9 \end{bmatrix}$$

Note : en pratique, on ne calcule pas les termes qui doivent être nuls (parties triangulaires nulles de L et de U) ou égaux à 1 (diagonale de L).

Pour que les calculs soient possibles, il faut que les $u_{i,i}$ soient non nuls. Ce sont les pivots de l'élimination de Gauss.

Factorisation LU directe : cas général

Récurrence :

Supposons que $\underline{U}_1, L_1, \underline{U}_2, L_2, \ldots, \underline{U}_{k-1}$ et L_{k-1} sont connus. Alors

- Ligne $i = k$:

$$\begin{aligned} \underline{A}_k &= \underline{L}_k U = \begin{bmatrix} l_{k,1} & \ldots & l_{k,k-1} & 1 & 0 & \ldots & 0 \end{bmatrix} U \\ &= l_{k,1}\underline{U}_1 + \cdots + l_{k,k-1}\underline{U}_{k-1} + \underline{U}_k \end{aligned}$$

$$\Longrightarrow \underline{U}_k = \underline{A}_k - \sum_{j=1}^{k-1} l_{k,j}\underline{U}_j \quad \text{connu}$$

On ne modifie que les termes $j = k$ à n (U est tri-sup)..

Factorisation LU directe : cas général

Récurrence :

Supposons que $\underline{U}_1, L_1, \underline{U}_2, L_2, \ldots, \underline{U}_{k-1}$ et L_{k-1} sont connus. Alors

- Ligne $i = k$:

$$\begin{aligned} \underline{A}_k &= \underline{L}_k U = \begin{bmatrix} l_{k,1} & \ldots & l_{k,k-1} & 1 & 0 & \ldots & 0 \end{bmatrix} U \\ &= l_{k,1}\underline{U}_1 + \cdots + l_{k,k-1}\underline{U}_{k-1} + \underline{U}_k \\ &\Longrightarrow \underline{U}_k = \underline{A}_k - \sum_{j=1}^{k-1} l_{k,j}\underline{U}_j \quad \text{connu} \end{aligned}$$

On ne modifie que les termes $j = k$ à n (U est tri-sup)..

- Colonne $j = k$: si $u_{k,k} \neq 0$

$$\begin{aligned} A_k &= LU_k = L\begin{bmatrix} u_{1,k} & \ldots & u_{k-1,k} & u_{k,k} & 0 & \ldots & 0 \end{bmatrix}^\top \\ &= L_1 u_{1,k} + \cdots + L_{k-1}u_{k-1,k} + L_k u_{k,k} \\ &\Longrightarrow L_k = \frac{1}{u_{k,k}}\left(A_k - \sum_{j=1}^{k-1} L_j u_{j,k}\right) \quad \text{connu} \end{aligned}$$

Ne modifier que les termes $i = k + 1$ à n (L est tri-inf et $L_{i,i} = 1$).

Factorisation LU directe : cas général

Récurrence :

Supposons que $\underline{U}_1, L_1, \underline{U}_2, L_2, \dots, \underline{U}_{k-1}$ et L_{k-1} sont connus. Alors

- Ligne $i = k$:

$$\begin{aligned} \underline{A}_k &= \underline{L}_k U = \left[\begin{array}{cccccc} l_{k,1} & \dots & l_{k,k-1} & 1 & 0 & \dots & 0 \end{array}\right] U \\ &= l_{k,1}\underline{U}_1 + \dots + l_{k,k-1}\underline{U}_{k-1} + \underline{U}_k \\ & \Longrightarrow \underline{U}_k = \underline{A}_k - \sum_{j=1}^{k-1} l_{k,j}\underline{U}_j \quad \text{connu} \end{aligned}$$

On ne modifie que les termes $j = k$ à n (U est tri-sup)..

- Colonne $j = k$: si $u_{k,k} \neq 0$

$$\begin{aligned} A_k &= LU_k = L\left[\begin{array}{ccccccc} u_{1,k} & \dots & u_{k-1,k} & u_{k,k} & 0 & \dots & 0 \end{array}\right]^\top \\ &= L_1 u_{1,k} + \dots + L_{k-1}u_{k-1,k} + L_k u_{k,k} \\ & \Longrightarrow L_k = \frac{1}{u_{k,k}}\left(A_k - \sum_{j=1}^{k-1} L_j u_{j,k}\right) \quad \text{connu} \end{aligned}$$

Ne modifier que les termes $i = k+1$ à n (L est tri-inf et $L_{i,i} = 1$).

Récurrence achevée.

Factorisation LU : algorithme de Doolittle

Algorithme de Doolittle (sans écraser A)

```
Initialiser L à I_n et U à 0 dans M_{n,n}.
Pour k=1 jusqu'à n
    Pour j=k jusqu'à n
        U(k,j) ⟵ A(k,j) - Σ_{p=1}^{k-1} L(k,p) × U(p,j)
    Fin Pour
    Si |U(k,k)|< tol
        Sortir, erreur : pivot nul
    Fin Si
    Pour i=k+1 jusqu'à n        (sauté pour i = n)
        L(i,k) ⟵ 1/U(k,k) × ( A(i,k) - Σ_{p=1}^{k-1} L(i,p) × U(p,k))
    Fin Pour
Fin Pour
```

$$\text{U(k,j)} \longleftarrow \text{A(k,j)} - \sum_{p=1}^{k-1} \text{L(k,p)} \times \text{U(p,j)}$$

$$\text{L(i,k)} \longleftarrow 1/\text{U(k,k)} \times \left(\text{A(i,k)} - \sum_{p=1}^{k-1} \text{L(i,p)} \times \text{U(p,k)}\right)$$

Coût identique à Gauss : $\sim \dfrac{n^3}{3}$ multiplications.

(À montrer, Chap 2, Exo TD n^o 3.).

Plan

Pivot nul : ?

Que faire si un pivot est nul ? Exemple : $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ inversible, mais $a_{1,1}^{(1)} = a_{1,1} = 0$!

$\longrightarrow$ permuter des lignes de A (ou des colonnes, inconnues x).

Définition

Soit $A \in \mathcal{M}_{n,n}$.
La sous matrice principale d'ordre $k \leq n$ de A notée $[A]_k$, est définie par :

$$[A]_k = (a_{i,j})_{i=1,\ldots,k;j=1,\ldots,k}$$

Exemple

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & -9 \end{bmatrix} \quad \left\{ \begin{array}{l} \det([A]_1) = 1 \\ \det([A]_2) = 1, \\ \det([A]_3) = -9. \end{array} \right.$$

Remarques sur les matrices par blocs...

Lemma (matrices par blocs)

Soit une matrice carrée, triangulaire supérieure par blocs. Alors son déterminant vaut :

$$\det\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) = \det(A) \times \det(C), \ \textit{et} \ \det\left(\begin{bmatrix} D & 0 \\ E & F \end{bmatrix}\right) = \det(D) \times \det(F).$$

Démonstration : extension de Chap 0, exo TD n^o 12.

Remarque : tri-inf par bloc $\neq$ tri-inf. (idem pour tri-sup). Prenez un exemple !

Remarques sur les matrices par blocs...

Lemma (matrices par blocs)

Soit une matrice carrée, triangulaire supérieure par blocs. Alors son déterminant vaut :

$$\det\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) = \det(A) \times \det(C), \ \textit{et} \ \det\left(\begin{bmatrix} D & 0 \\ E & F \end{bmatrix}\right) = \det(D) \times \det(F).$$

Démonstration : extension de Chap 0, exo TD n^o 12.
Remarque : tri-inf par bloc $\neq$ tri-inf. (idem pour tri-sup). Prenez un exemple !

$$\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right)$$

Remarques sur les matrices par blocs...

Lemma (matrices par blocs)

Soit une matrice carrée, triangulaire supérieure par blocs*. Alors son déterminant vaut :*

$$\det\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) = \det(A) \times \det(C), \ \textit{et} \ \det\left(\begin{bmatrix} D & 0 \\ E & F \end{bmatrix}\right) = \det(D) \times \det(F).$$

Démonstration : extension de Chap 0, exo TD n^o 12.
Remarque : tri-inf par bloc $\neq$ tri-inf. (idem pour tri-sup). Prenez un exemple !

$$\det\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) \neq \det(A)\det(D) - \det(C)\det(B)$$

Attention : ARCHI faux ! car B et C sont rectangulaires en général ! !

Pivot nul : CNS pour l'élimination de Gauss

Théorème (2.4.2)

Les trois propositions suivantes sont équivalentes :

- *Proposition 1 :*

$$\text{Tous les pivots } a_{kk}^{(k)} \text{ sont définis et } a_{kk}^{(k)} \neq 0, \ k = 1, \ldots, n.$$

- *Proposition 2 :*

$$A \text{ admet une factorisation } A = LU \text{ avec } U \text{ inversible .}$$

- *Proposition 3 :*

$$[A]_1, \ [A]_2, \ \ldots, [A]_n \text{ inversibles.}$$

Pivot nul : CNS pour l'élimination de Gauss

Théorème (2.4.2)

Les trois propositions suivantes sont équivalentes :

- *Proposition 1 :*

  Tous les pivots $a_{kk}^{(k)}$ *sont définis et* $a_{kk}^{(k)} \neq 0, \; k = 1, \dots, n.$

- *Proposition 2 :*

  A admet une factorisation $A = LU$ *avec* U *inversible .*

- *Proposition 3 :*

  $[A]_1, \; [A]_2, \; \dots, [A]_n$ *inversibles.*

Prop. 3 : critère théorique
$\longrightarrow$ Prop. 3 vraie pour certaines matrices : symétrique définie positive (SDP), à diagonale strictement dominante (voir plus tard et Exo5 TD)
$\longrightarrow$ En pratique : on ne calcule pas les $\det([A]_k)$, on calcule les pivots et on vérifie au fur et à mesure s'ils sont nuls ou pas.

Pivot nul : CNS

Démonstration :

- On a déjà prouvé Prop. 1 $\iff$ Prop. 2 :
  élmination de Gauss faisable jusqu'au bout
  $\iff$ les pivots existent et sont tous non nuls
  $\iff$ factorisation $A = LU$ faisable (et $\det(U) = \Pi_{i=1}^{n} u_{i,i} = \Pi_{i=1}^{n} a_{i,i}^{(i)} \neq 0$).

Pivot nul : CNS

Démonstration :

- On a déjà prouvé Prop. 1 $\iff$ Prop. 2 :
  élmination de Gauss faisable jusqu'au bout
  $\iff$ les pivots existent et sont tous non nuls
  $\iff$ factorisation $A = LU$ faisable (et $\det(U) = \Pi_{i=1}^{n} u_{i,i} = \Pi_{i=1}^{n} a_{i,i}^{(i)} \neq 0$).
- Prop. 2 $\Longrightarrow$ Prop. 3 : soit $1 \leq k \leq n$. Décomposition de $A = LU$ en blocs de tailles k et $(n-k)$:

$$A = \begin{bmatrix} [A]_k & B \\ C & D \end{bmatrix} = \begin{bmatrix} [L]_k & 0 \\ E & F \end{bmatrix} \times \begin{bmatrix} [U]_k & G \\ 0 & H \end{bmatrix}$$

Démonstration :

- On a déjà prouvé Prop. 1 $\Longleftrightarrow$ Prop. 2 :
  élmination de Gauss faisable jusqu'au bout
  $\Longleftrightarrow$ les pivots existent et sont tous non nuls
  $\Longleftrightarrow$ factorisation $A = LU$ faisable (et $\det(U) = \Pi_{i=1}^{n} u_{i,i} = \Pi_{i=1}^{n} a_{i,i}^{(i)} \neq 0$).
- Prop. 2 $\Longrightarrow$ Prop. 3 : soit $1 \leq k \leq n$. Décomposition de $A = LU$ en blocs de tailles k et $(n-k)$:

$$A = \begin{bmatrix} [A]_k & B \\ C & D \end{bmatrix} = \begin{bmatrix} [L]_k & 0 \\ E & F \end{bmatrix} \times \begin{bmatrix} [U]_k & G \\ 0 & H \end{bmatrix}$$
$$= \begin{bmatrix} [L]_k[U]_k & [L]_k G \\ E[U]_k & EG + FH \end{bmatrix}$$

En identifiant : $[A]_k = [L]_k[U]_k$. Comme $0 \neq \det(U) = \det([U]_k)\det(H)$ et $1 = \det(L) = \det([L]_k)\det(F)$, on a $\det([A]_k) \neq 0$.

Donc $[A]_k$ est inversible.

On a $\det([A]_k) = 1 \times \det([U]_k) = \Pi_{i=1}^{k} u_{i,i}$ (produit des k premiers pivots.)

Pivot nul : CNS

Démonstration :

- On a déjà prouvé Prop. 1 $\iff$ Prop. 2 :
  élmination de Gauss faisable jusqu'au bout
  $\iff$ les pivots existent et sont tous non nuls
  $\iff$ factorisation $A = LU$ faisable (et $\det(U) = \Pi_{i=1}^{n} u_{i,i} = \Pi_{i=1}^{n} a_{i,i}^{(i)} \neq 0$).
- Prop. 2 $\implies$ Prop. 3 : soit $1 \leq k \leq n$. Décomposition de $A = LU$ en blocs de tailles k et $(n-k)$:

$$A = \begin{bmatrix} [A]_k & B \\ C & D \end{bmatrix} = \begin{bmatrix} [L]_k & 0 \\ E & F \end{bmatrix} \times \begin{bmatrix} [U]_k & G \\ 0 & H \end{bmatrix}$$
$$= \begin{bmatrix} [L]_k[U]_k & [L]_k G \\ E[U]_k & EG + FH \end{bmatrix}$$

  En identifiant : $[A]_k = [L]_k[U]_k$. Comme $0 \neq \det(U) = \det([U]_k)\det(H)$ et $1 = \det(L) = \det([L]_k)\det(F)$, on a $\det([A]_k) \neq 0$.
  Donc $[A]_k$ est inversible.
  On a $\det([A]_k) = 1 \times \det([U]_k) = \Pi_{i=1}^{k} u_{i,i}$ (produit des k premiers pivots.)
- Prop. 3 $\implies$ Prop. 1 : voir le poly (Doc. B.1.1).

Permutation de ligne si pivot nul.

Soit A inversible.
Si $(k-1)$ étapes faisables sans permutations ($a_{i,i}^{(i)} \neq 0,\ i = 1, \ldots, k-1$) :

$$A^{(k)} = \begin{bmatrix} a_{1,1}^{(1)} & a_{1,2}^{(1)} & \ldots & a_{1,k-1}^{(1)} & a_{1,k}^{(1)} & a_{1,k+1}^{(1)} & \ldots & a_{1,n}^{(1)} \\ 0 & a_{2,2}^{(2)} & \ldots & a_{2,k-1}^{(2)} & a_{2,k}^{(2)} & a_{2,k+1}^{(2)} & \ldots & a_{2,n}^{(2)} \\ \vdots & \vdots & & \ddots & \ddots & \vdots & & \vdots \\ 0 & 0 & \ldots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & a_{k-1,k+1}^{(k-1)} & \ldots & a_{k-1,n}^{(k-1)} \\ 0 & 0 & \ldots & 0 & a_{k,k}^{(k)} & a_{k,k+1}^{(k)} & \ldots & a_{k,n}^{(k)} \\ 0 & 0 & \ldots & 0 & a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & \ldots & a_{k+1,n}^{(k)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\ 0 & 0 & \ldots & 0 & a_{n,k}^{(k)} & a_{n,k+1}^{(k)} & \ldots & a_{n,n}^{(k)} \end{bmatrix}$$

Écriture par blocs : $A^{(k)} = \begin{bmatrix} [U]_{k-1} & D^{(k)} \\ 0 & S^{(k)} \end{bmatrix}$ avec $S^{(k)} \in \mathcal{M}_{n-k+1,n-k+1}$

Pas de permutation, donc $0 \neq \det(A) = \det(A^{(k)}) = \Pi_{i=1}^{k-1} a_{i,i}^{(i)} \times \det(S^{(k)})$.

Donc $S^{(k)}$ inversible, donc colonne $S_1^{(k)}$ non nulle :

$\Longrightarrow$ il existe une ligne $i_0 \in \{k, k+1, \ldots, n\}$ telle que $a_{i_0,k}^{(k)} \neq 0$.

Permutation de lignes si pivot nul.

Après $(k-1)$ étapes, il existe une ligne $i_0 \in \{k, k+1, \ldots, n\}$ telle que $a_{i_0,k}^{(k)} \neq 0$.

Donc si $a_{k,k}^{(k)} = 0$, on permute les lignes i_0 et k : $A^{(k)} \longrightarrow A^{(k)\prime}$

$\Longrightarrow$ $a_{k,k}^{(k)\prime} = a_{i_0,k}^{(k)} \neq 0$ nouveau pivot

$\Longrightarrow$ $\det(A^{(k)\prime}) = -\det(A^{(k)})$ (car 1 permutation).

Proposition (2.4.1)

Soit A une matrice inversible. À toute étape k de la méthode d'élimination de Gauss, il existe au moins un élément non nul dans la colonne k et situé sur la ligne i avec $i \geq k$.

Permutation de lignes : factorisation $PA = LU$.

Théorème (2.4.4)

Soit A une matrice inversible. Alors il existe une matrice triangulaire inférieure L, une matrice triangulaire supérieure U, et une matrice de permutation P telles que

$$PA = LU.$$

- P matrice de permutation : que des 0, un seul 1 par ligne et par colonne. $\det(P) = \pm 1$.
- La décomposition $PA = LU$ n'est pas unique ! (choix de plusieurs i_0 en général.)
- Démonstration (hors programme) : truc "magique" dans l'échange de lignes.

Permutation de lignes : factorisation $PA = LU$.

Théorème (2.4.4)

Soit A une matrice inversible. Alors il existe une matrice triangulaire inférieure L, une matrice triangulaire supérieure U, et une matrice de permutation P telles que

$$PA = LU.$$

- P matrice de permutation : que des 0, un seul 1 par ligne et par colonne. $\det(P) = \pm 1$.
- La décomposition $PA = LU$ n'est pas unique ! (choix de plusieurs i_0 en général.)
- Démonstration (hors programme) : truc "magique" dans l'échange de lignes.
- En pratique : on fait des permutations à chaque étape k pour améliorer la précision.
  On choisit le plus grand pivot possible en valeur absolue.
  $\longrightarrow$ prendre $i_0 \in \{k, k+1, \ldots, n\}$ tel que $|a^{(k)}_{i_0,k}| \geq |a^{(k)}_{i,k}|$ pour $i \in \{k, k+1, \ldots, n\}$.
- Algorithme de Gauss **avec** permutation : **stable** par rapport aux erreurs d'arrondis (erreurs croissent pas trop vite). (Non démontré).
- Algorithme de Gauss **sans** permutation : **pas stable**.
  Voir exemple, TD2-Exercice 6 (Exo C.2.6).

Instabilité de Gauss sans permutation : exemple

Algorithme de Gauss **sans** permutation : **instable**.

Exemple : **sans** permutation avec un pivot “petit” :

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \overset{\Longrightarrow}{\text{arithm. exacte}} \quad x = \begin{bmatrix} \dfrac{1}{1-\varepsilon} \\ \dfrac{1-2\varepsilon}{1-\varepsilon} \end{bmatrix} \sim \begin{bmatrix} 1+\varepsilon \\ 1-\varepsilon \end{bmatrix}$$

$$\overset{\Longrightarrow}{\text{arithm. flottante}} \quad \tilde{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{si } fl(1-\varepsilon) = 1.$$

Résultat très mauvais.

Instabilité de Gauss sans permutation : exemple

Algorithme de Gauss **sans** permutation : **instable**.

Exemple : **sans** permutation avec un pivot “petit” :

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \underset{\text{arithm. exacte}}{\Longrightarrow} \quad x = \begin{bmatrix} \dfrac{1}{1-\varepsilon} \\ \dfrac{1-2\varepsilon}{1-\varepsilon} \end{bmatrix} \sim \begin{bmatrix} 1+\varepsilon \\ 1-\varepsilon \end{bmatrix}$$

$$\underset{\text{arithm. flottante}}{\Longrightarrow} \quad \tilde{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{si } fl(1-\varepsilon) = 1.$$

Résultat très mauvais.

Avec permutation :

$$\begin{bmatrix} 1 & 1 \\ \varepsilon & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \underset{\text{arithm. flottante}}{\Longrightarrow} \quad \tilde{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{si } fl(1-\varepsilon) = fl(1-2\varepsilon) = 1.$$

Résultat précis.

Instabilité de Gauss sans permutation : exemple

Algorithme de Gauss **sans** permutation : **instable**.

Exemple : **sans** permutation avec un pivot “petit” :

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \overset{\Longrightarrow}{\text{arithm. exacte}} \quad x = \begin{bmatrix} \dfrac{1}{1-\varepsilon} \\ \dfrac{1-2\varepsilon}{1-\varepsilon} \end{bmatrix} \sim \begin{bmatrix} 1+\varepsilon \\ 1-\varepsilon \end{bmatrix}$$

$$\overset{\Longrightarrow}{\text{arithm. flottante}} \quad \tilde{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{si } fl(1-\varepsilon) = 1.$$

Résultat très mauvais.

Avec permutation :

$$\begin{bmatrix} 1 & 1 \\ \varepsilon & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \overset{\Longrightarrow}{\text{arithm. flottante}} \quad \tilde{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{si } fl(1-\varepsilon) = fl(1-2\varepsilon) = 1.$$

Résultat précis.

Conclusion : en pratique, il faut toujours permuter ($\longrightarrow$ grands pivots !)

Plan

Matrices quelconques : factorisation LDV

Proposition

Soit A une matrice dont toutes les sous-matrices principales sont régulières. Alors elle admet une unique factorisation sous la forme

$$A = LDV,$$

où

- *L est une matrice triangulaire inférieure à diagonale unité,*
- *D une matrice diagonale régulière*
- *V est une matrice triangulaire supérieure à diagonale unité.*

Matrices quelconques : factorisation LDV

Proposition

Soit A une matrice dont toutes les sous-matrices principales sont régulières. Alors elle admet une unique factorisation sous la forme

$$A = LDV,$$

où

- *L est une matrice triangulaire inférieure à diagonale unité,*
- *D une matrice diagonale régulière*
- *V est une matrice triangulaire supérieure à diagonale unité.*

Schéma de démonstration.

1. **Existence :** $A = LU$ existe.
   $D = \mathrm{diag}(U) = \mathrm{diag}\left((u_{i,i})_{i=1,\dots,n}\right)$ inversible car pivots non nuls.
   $D^{-1} = \mathrm{diag}\left((\frac{1}{u_{i,i}})_{i=1,\dots,n}\right)$, donc $V = D^{-1}U$ est tri-sup à diagonale unité.

Matrices quelconques : factorisation *LDV*

Proposition

Soit A une matrice dont toutes les sous-matrices principales sont régulières. Alors elle admet une unique factorisation sous la forme

$$A = LDV,$$

où

- L *est une matrice triangulaire inférieure à diagonale unité,*
- D *une matrice diagonale régulière*
- V *est une matrice triangulaire supérieure à diagonale unité.*

Schéma de démonstration.

1. **Existence :** $A = LU$ existe.
   $D = \mathrm{diag}(U) = \mathrm{diag}\left((u_{i,i})_{i=1,\dots,n}\right)$ inversible car pivots non nuls.
   $D^{-1} = \mathrm{diag}\left((\frac{1}{u_{i,i}})_{i=1,\dots,n}\right)$, donc $V = D^{-1}U$ est tri-sup à diagonale unité.
2. **Unicité :** découle de l'unicité de $A = LU$. Prendre $A = LDV = \tilde{L}\tilde{D}\tilde{V}$, et prouver que les matrices sont égales.

Matrices symétriques : factorisation LDL^T

Proposition (2.5.1)

Soit A une matrice **symétrique** *dont toutes les sous-matrices principales sont régulières.*
Alors elle admet une unique factorisation sous la forme

$$A = LDL^T,$$

où

- *L est une matrice triangulaire inférieure à diagonale unité,*
- *D une matrice diagonale régulière.*

Matrices symétriques : factorisation LDL^T

Proposition (2.5.1)

Soit A une matrice **symétrique** *dont toutes les sous-matrices principales sont régulières.*
Alors elle admet une unique factorisation sous la forme

$$A = LDL^T,$$

où

- *L est une matrice triangulaire inférieure à diagonale unité,*
- *D une matrice diagonale régulière.*

Démonstration. Factorisation $A = LDV$ existe et est unique.
A symétrique donc : $A = A^T \iff LDV = (LDV)^T \iff LDV = V^T D^T L^T = V^T D L^T$.

Par unicité $L = V^T$.

Matrices symétriques : factorisation LDL^T

Proposition (2.5.1)

Soit A une matrice **symétrique** *dont toutes les sous-matrices principales sont régulières.*
Alors elle admet une unique factorisation sous la forme

$$A = LDL^T,$$

où

- *L est une matrice triangulaire inférieure à diagonale unité,*
- *D une matrice diagonale régulière.*

Démonstration. Factorisation $A = LDV$ existe et est unique.
A symétrique donc : $A = A^T \iff LDV = (LDV)^T \iff LDV = V^T D^T L^T = V^T D L^T$.
Par unicité $L = V^T$.

On ne stocke que la partie utile (tri-inf par exemple).
D : ne contient pas les valeurs propres de A !!
$A = LDL^T$ n'est pas une diagonalisation de A !!

Matrices SDP : définition

Définition (matrices symétriques (semi-)définies positives)

Soit $A \in \mathcal{M}_{n,n}$ une matrice symétrique ($A^T = A$). Alors

$$A \ \text{SDP} \quad \Longleftrightarrow \quad \begin{cases} \forall x \in \mathbb{R}^n : & x^T A x \geq 0 \\ \forall x \in \mathbb{R}^n : & x^T A x = 0 \Longrightarrow x = 0 \end{cases}$$

$$\Longleftrightarrow \quad \left\{ \ \forall x \in \mathbb{R}^n \setminus \{0\} : \quad x^T A x > 0 \right.$$

Et :

$$A \ \text{semi DP} \quad \Longleftrightarrow \quad \left\{ \ \forall x \in \mathbb{R}^n : \quad x^T A x \geq 0 \right.$$

Matrices SDP : définition

Définition (matrices symétriques (semi-)définies positives)

Soit $A \in \mathcal{M}_{n,n}$ une matrice symétrique ($A^T = A$). Alors

$$A \ \text{SDP} \quad \Longleftrightarrow \quad \begin{cases} \forall x \in \mathbb{R}^n : & x^T A x \geq 0 \\ \forall x \in \mathbb{R}^n : & x^T A x = 0 \Longrightarrow x = 0 \end{cases}$$

$$\Longleftrightarrow \quad \left\{ \ \forall x \in \mathbb{R}^n \setminus \{0\} : \quad x^T A x > 0 \right.$$

Et :

$$A \ \text{semi DP} \quad \Longleftrightarrow \quad \left\{ \ \forall x \in \mathbb{R}^n : \quad x^T A x \geq 0 \right.$$

Rappels :

- A symétrique $\Longrightarrow$ A diagonalisable dans une BON. Il existe P orthogonale ($P^T = P^{-1}$, P matrice de changement de BON) et Λ diagonale (contenant les vp réelles de A), telles que $\Lambda = P^T A P = P^{-1} A P$.
- A semi-DP : $\Longrightarrow$ les vp de A sont ≥ 0, A SDP : $\Longrightarrow$ les vp de A sont > 0.
- Décomposition en carrés de Gauss d'une forme quadratique.
- Voir exemple.

Formes quadratiques : rappels

Rappels :

- soit $A \in \mathcal{M}_{n,p}$, $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^p$. Alors

$$x^T Ay = \sum_{i=1}^{n} \sum_{j=1}^{p} x_i A_{i,j} y_j = \sum_{j=1}^{p} \sum_{i=1}^{n} y_j A_{i,j} x_i = y^T A^T x \qquad \in \mathbb{R}.$$

- soit $A \in \mathcal{M}_{n,n}$ **symétrique**, x et $y \in \mathbb{R}^n$. Alors

$$x^T Ay = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{i,j} y_j = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{j,i} y_j = \sum_{k=1}^{n} \sum_{l=1}^{n} y_k A_{k,l} x_l = y^T Ax \qquad \in \mathbb{R}.$$

et on définit une forme quadratique $\Phi \ : \mathbb{R}^n \longrightarrow \mathbb{R}$ par :

$$\Phi(x) = x^T Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{i,j} x_j = \sum_{i=1}^{n} A_{i,i} x_i^2 + 2 \sum_{1 \leq i < j \leq n} A_{i,j} x_i x_j.$$

- A semi-DP $\iff$ $\Phi(x) \geq 0 \ \forall x \in \mathbb{R}^n$.
- A SDP $\iff$ $\Phi(x) > 0 \ \forall x \in \mathbb{R}^n \setminus \{0\}$.

Formes quadratiques : exemple 3×3

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 9 \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \Phi(x) &= x^\top A x = \left\{ \begin{array}{r} 1x_1^2 + 1x_1x_2 + 2x_1x_3 \\ +1x_2x_1 + 2x_2^2 + 1x_2x_3 \\ +2x_3x_1 + 1x_3x_2 + 9x_3^2 \end{array} \right. \\ &= 1x_1^2 + 2x_2^2 + 9x_3^2 \\ &\quad +2\,(1x_1x_2 + 2x_1x_3 + 1x_2x_3) \end{aligned}$$

Formes quadratiques : exemple 3×3

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 9 \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \Phi(x) &= x^\top A x = \left\{ \begin{array}{r} 1x_1^2 + 1x_1x_2 + 2x_1x_3 \\ +1x_2x_1 + 2x_2^2 + 1x_2x_3 \\ +2x_3x_1 + 1x_3x_2 + 9x_3^2 \end{array} \right. \\ &= 1x_1^2 + 2x_2^2 + 9x_3^2 \\ &\quad +2\left(1x_1x_2 + 2x_1x_3 + 1x_2x_3\right) \end{aligned}$$

Décomposition de Gauss en carrés :

$$\boxed{\left\{ \begin{array}{rcl} x_1^2 + 2\alpha x_1 y &=& (x_1 + \alpha y)^2 - \alpha^2 y^2 \\ x_1 x_2 &=& \frac{1}{4}\left((x_1 + x_2)^2 - (x_1 - x_2)^2\right). \end{array} \right.}$$

Formes quadratiques : exemple 3 × 3

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 9 \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \Phi(x) &= x^\top A x = \left\{ \begin{array}{l} 1x_1^2 + 1x_1x_2 + 2x_1x_3 \\ +1x_2x_1 + 2x_2^2 + 1x_2x_3 \\ +2x_3x_1 + 1x_3x_2 + 9x_3^2 \end{array} \right. \\ &= 1x_1^2 + 2x_2^2 + 9x_3^2 \\ &\quad +2\,(1x_1x_2 + 2x_1x_3 + 1x_2x_3) \end{aligned}$$

Décomposition de Gauss en carrés :

$$\boxed{\left\{ \begin{array}{rcl} x_1^2 + 2\alpha x_1 y &=& (x_1 + \alpha y)^2 - \alpha^2 y^2 \\ x_1 x_2 &=& \frac{1}{4}\left((x_1 + x_2)^2 - (x_1 - x_2)^2\right). \end{array} \right.}$$

On trouve

$$\begin{aligned} \Phi(x) = x^\top A x &= 1(\underbrace{1x_1 + 1x_2 + 2x_3}_{y_1})^2 + 1(\underbrace{1x_2 - 1x_3}_{y_2})^2 + 4(\underbrace{1x_3}_{y_3})^2 \\ &= 1y_1^2 + 1y_2^2 + 4y_3^2 \quad \geq 0. \end{aligned}$$

Formes quadratiques : exemple 3×3

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 9 \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \Phi(x) &= x^\top A x = \left\{ \begin{array}{l} \quad 1x_1^2 + 1x_1x_2 + 2x_1x_3 \\ +1x_2x_1 + 2x_2^2 + 1x_2x_3 \\ +2x_3x_1 + 1x_3x_2 + 9x_3^2 \end{array} \right. \\ &= 1x_1^2 + 2x_2^2 + 9x_3^2 \\ &\quad +2\left(1x_1x_2 + 2x_1x_3 + 1x_2x_3\right) \end{aligned}$$

Décomposition de Gauss en carrés :

$$\boxed{\left\{ \begin{array}{rcl} x_1^2 + 2\alpha x_1 y &=& (x_1 + \alpha y)^2 - \alpha^2 y^2 \\ x_1 x_2 &=& \frac{1}{4}\left((x_1 + x_2)^2 - (x_1 - x_2)^2\right). \end{array} \right.}$$

On trouve

$$\begin{aligned} \Phi(x) = x^\top A x &= 1(\underbrace{1x_1 + 1x_2 + 2x_3}_{y_1})^2 + 1(\underbrace{1x_2 - 1x_3}_{y_2})^2 + 4(\underbrace{1x_3}_{y_3})^2 \\ &= 1y_1^2 + 1y_2^2 + 4y_3^2 \quad \geq 0. \end{aligned}$$

Donc pour tout $x \in \mathbb{R}^3$, $\Phi(x) \geq 0$ et

$$\begin{aligned} \Phi(x) = 0 \quad &\Longleftrightarrow \quad y_3 = y_2 = y_1 = 0 \\ &\Longleftrightarrow \quad x_3 = x_2 = x_1 = 0 \quad \Longleftrightarrow \quad x = 0 \end{aligned}$$

Donc Φ DP.

Formes quadratiques : exemple 3×3

Décomposition de Gauss en carrés :

$$\begin{aligned}\Phi(x) = x^\top Ax &= (\underbrace{1x_1 + 1x_2 + 2x_3}_{y_1})^2 + (\underbrace{1x_2 - 1x_3}_{y_2})^2 + 4(\underbrace{1x_3}_{y_3})^2 \\ &= 1y_1^2 + 1y_2^2 + 4y_3^2 = y^\top Dy\end{aligned}$$

avec $y = L^\top x$, où $L^\top = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, et $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. Donc

$$\Phi(x) = x^\top Ax = y^\top Dy = \left(L^\top x\right)^\top D \left(L^\top x\right) = x^\top LDL^\top x$$

donc

$$A = LDL^\top .$$

Formes quadratiques : exemple 3×3

Décomposition de Gauss en carrés :

$$\begin{aligned}
\Phi(x) = x^\top A x &= (\underbrace{1x_1 + 1x_2 + 2x_3}_{y_1})^2 + (\underbrace{1x_2 - 1x_3}_{y_2})^2 + 4(\underbrace{1x_3}_{y_3})^2 \\
&= 1y_1^2 + 1y_2^2 + 4y_3^2 = y^\top D y
\end{aligned}$$

avec $y = L^\top x$, où $L^\top = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, et $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. Donc

$$\Phi(x) = x^\top A x = y^\top D y = \left(L^\top x\right)^\top D \left(L^\top x\right) = x^\top L D L^\top x$$

donc

$$A = LDL^\top.$$

De plus, ici, $D_{i,i} > 0$, donc il existe G diagonale tq. $D = G^2$ et $G_{i,i} = +\sqrt{D_{i,i}} > 0$.

$$A = LDL^\top = (LG)(GL^\top) = (LG)(G^\top L^\top) = (LG)(LG)^\top = CC^\top$$

et

$$C^\top = G^\top L^\top = \begin{bmatrix} \sqrt{1} & 0 & 0 \\ 0 & \sqrt{1} & 0 \\ 0 & 0 & \sqrt{4} \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Lemme (2.5.2)

Soit A matrice SDP.
Alors toutes ses sous-matrices principales $[A]_k$ sont régulières.

Démonstration : voir poly.

$\Longrightarrow$ donc A SDP admet une factorisation : $A = LDL^T$.

Matrices symétriques définies positives : factorisation de Cholesky

Théorème (2.5.3 : factorisation de Cholesky)

Soit A matrice SDP.
Alors elle admet une unique factorisation de Cholesky sous la forme

$$A = CC^T, \quad \textit{où}$$

- C *est triangulaire inférieure, et* $C_{i,i} > 0$, $\forall i = 1, \ldots, n$.

Théorème (2.5.3 : factorisation de Cholesky)

Soit A matrice SDP.
Alors elle admet une unique factorisation de Cholesky sous la forme

$$A = CC^T, \quad \text{où}$$

- C *est triangulaire inférieure, et* $C_{i,i} > 0,\ \forall i = 1, \ldots, n.$

Démonstration : *Existence*. A SDP : donc $\forall k$ les $[A]_k$ sont SDP donc inversibles. Donc il existe une unique factorisation $A = LDL^\top$.
De plus, $\forall x$, $x^\top Ax = x^\top LDL^\top x = (L^\top x)^\top D(L^\top x) = y^\top Dy = \sum_{i=1}^n D_{i,i}y_i^2$, avec $y = L^\top x$.
Comme $L^\top$ est tri-sup avec diagonale unité, elle est inversible. Donc pour $i = 1, \ldots, n$,
$y = e_i \iff x = L^{-\top}e_i \neq 0$. Donc pour $x = L^{-\top}e_i \neq 0$, on a $x^\top Ax = (e_i)^\top De_i = D_{i,i} > 0$
Donc il existe G diagonale tq. $D = G^2$ et $G_{i,i} = +\sqrt{D_{i,i}} > 0$. On pose $C = LG$ et

$$A = LDL^\top = (LG)(GL^\top) = (LG)(G^\top L^\top) = (LG)(LG)^\top = CC^\top.$$

Matrices symétriques définies positives : factorisation de Cholesky

Théorème (2.5.3 : factorisation de Cholesky)

Soit A matrice SDP.
Alors elle admet une unique factorisation de Cholesky sous la forme

$$A = CC^T, \quad \text{où}$$

- C *est triangulaire inférieure, et* $C_{i,i} > 0,\ \forall i = 1, \dots, n.$

Démonstration : *Existence.* A SDP : donc $\forall k$ les $[A]_k$ sont SDP donc inversibles. Donc il existe une unique factorisation $A = LDL^\top$.
De plus, $\forall x$, $x^\top Ax = x^\top LDL^\top x = (L^\top x)^\top D(L^\top x) = y^\top Dy = \sum_{i=1}^n D_{i,i}y_i^2$, avec $y = L^\top x$.
Comme $L^\top$ est tri-sup avec diagonale unité, elle est inversible. Donc pour $i = 1, \dots, n$,
$y = e_i \iff x = L^{-\top}e_i \neq 0$. Donc pour $x = L^{-\top}e_i \neq 0$, on a $x^\top Ax = (e_i)^\top De_i = D_{i,i} > 0$
Donc il existe G diagonale tq. $D = G^2$ et $G_{i,i} = +\sqrt{D_{i,i}} > 0$. On pose $C = LG$ et

$$A = LDL^\top = (LG)(GL^\top) = (LG)(G^\top L^\top) = (LG)(LG)^\top = CC^\top.$$

Démonstration : *Unicité.* On suppose que $A = CC^\top = \tilde{C}\tilde{C}^\top$. Comme C est tri-inf et $C_{i,i} > 0$ et $\tilde{C}^\top$ tri-sup et $\tilde{C}_{i,i} > 0$, les matrices C et $\tilde{C}^\top$ sont inversibles (Voir Chap. 0, exo TD 11). Donc $D = C^\top\tilde{C}^{-\top} = C^{-1}\tilde{C}$: matrice tri-sup et tri-inf donc diagonale, et $D_{i,i} = C_{i,i}/\tilde{C}_{i,i} = \tilde{C}_{i,i}/C_{i,i}$.
Donc $C_{i,i}^2 = \tilde{C}_{i,i}^2 \Leftrightarrow C_{i,i} = \tilde{C}_{i,i}$ car > 0. Donc $D = I$ et $C = \tilde{C}$. Unicité.

Matrices SDP : principe de la factorisation de Cholesky : $A = CC^\top$

$$
\begin{bmatrix}
c_{1,1} & 0 & 0 & \cdots & 0 & 0 \\
c_{2,1} & c_{2,2} & 0 & \cdots & 0 & 0 \\
c_{3,1} & c_{3,2} & c_{3,3} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
c_{n-1,1} & c_{n-1,2} & c_{n-1,3} & \cdots & c_{n-1,n-1} & 0 \\
c_{n,1} & c_{n,2} & c_{n,3} & \cdots & c_{n,n-1} & c_{n,n}
\end{bmatrix}
\begin{bmatrix}
c_{1,1} & c_{2,1} & c_{3,1} & \cdots & c_{n-1,1} & c_{n,1} \\
0 & c_{2,2} & c_{3,2} & \cdots & c_{n-1,2} & c_{n,2} \\
0 & 0 & c_{3,3} & \cdots & c_{n-1,3} & c_{n,3} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & c_{n-1,n-1} & c_{n-1,n} \\
0 & 0 & 0 & \cdots & 0 & c_{n,n}
\end{bmatrix}
$$

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).

Matrices SDP : principe de la factorisation de Cholesky : $A = CC^\top$

$$\begin{bmatrix} C_{1,1} & 0 & 0 & \cdots & 0 & 0 \\ C_{2,1} & C_{2,2} & 0 & \cdots & 0 & 0 \\ C_{3,1} & C_{3,2} & C_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n-1,1} & C_{n-1,2} & C_{n-1,3} & \cdots & C_{n-1,n-1} & 0 \\ C_{n,1} & C_{n,2} & C_{n,3} & \cdots & C_{n,n-1} & C_{n,n} \end{bmatrix} \begin{bmatrix} C_{1,1} & C_{2,1} & C_{3,1} & \cdots & C_{n-1,1} & C_{n,1} \\ 0 & C_{2,2} & C_{3,2} & \cdots & C_{n-1,2} & C_{n,2} \\ 0 & 0 & C_{3,3} & \cdots & C_{n-1,3} & C_{n,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & C_{n-1,n-1} & C_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & C_{n,n} \end{bmatrix}$$

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).

- $A_{1,1} = C_{1,1}C_{1,1} \qquad \Longrightarrow C_{1,1} = \sqrt{A_{1,1}}$

  Si $i > 1$: $A_{i,1} = C_{i,1}C_{1,1} \qquad \Longrightarrow C_{i,1} = \dfrac{A_{i,1}}{C_{1,1}}$

Matrices SDP : principe de la factorisation de Cholesky : $A = CC^\top$

$$\begin{bmatrix} C_{1,1} & 0 & 0 & \cdots & 0 & 0 \\ C_{2,1} & C_{2,2} & 0 & \cdots & 0 & 0 \\ C_{3,1} & C_{3,2} & C_{3,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n-1,1} & C_{n-1,2} & C_{n-1,3} & \cdots & C_{n-1,n-1} & 0 \\ C_{n,1} & C_{n,2} & C_{n,3} & \cdots & C_{n,n-1} & C_{n,n} \end{bmatrix} \begin{bmatrix} C_{1,1} & C_{2,1} & C_{3,1} & \cdots & C_{n-1,1} & C_{n,1} \\ 0 & C_{2,2} & C_{3,2} & \cdots & C_{n-1,2} & C_{n,2} \\ 0 & 0 & C_{3,3} & \cdots & C_{n-1,3} & C_{n,3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & C_{n-1,n-1} & C_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & C_{n,n} \end{bmatrix}$$

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).

- $A_{1,1} = C_{1,1} C_{1,1} \qquad \Longrightarrow C_{1,1} = \sqrt{A_{1,1}}$

  Si $i > 1$: $A_{i,1} = C_{i,1} C_{1,1} \qquad \Longrightarrow C_{i,1} = \dfrac{A_{i,1}}{C_{1,1}}$

- $A_{2,2} = C_{2,1} C_{2,1} + C_{2,2} C_{2,2} \qquad \Longrightarrow C_{2,2} = \sqrt{A_{2,2} - C_{2,1}^2}$

  Si $i > 2$: $A_{i,2} = C_{i,1} C_{2,1} + C_{i,2} C_{2,2} \qquad \Longrightarrow C_{i,2} = \dfrac{1}{C_{2,2}} \left(A_{i,2} - C_{i,1} C_{2,1} \right)$

Matrices SDP : principe de la factorisation de Cholesky

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).
Pour $i, j = 1, \ldots, n$, en utilisant le fait que C est tri-inf :

$$A = CC^\top \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}(C^\top)_{k,j} \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}C_{j,k}$$

Matrices SDP : principe de la factorisation de Cholesky

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).
Pour $i, j = 1, \ldots, n$, en utilisant le fait que C est tri-inf :

$$A = CC^\top \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}(C^\top)_{k,j} \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}C_{j,k}$$

- $A_{1,1} = C_{1,1}C_{1,1} \Longrightarrow C_{1,1} = \sqrt{A_{1,1}}$

  Si $i > 1$: $A_{i,1} = C_{i,1}C_{1,1} \Longrightarrow C_{i,1} = \dfrac{A_{i,1}}{C_{1,1}}$

- $A_{2,2} = C_{2,1}C_{2,1} + C_{2,2}C_{2,2} \Longrightarrow C_{2,2} = \sqrt{A_{2,2} - C_{2,1}^2}$

  Si $i > 2$: $A_{i,2} = C_{i,1}C_{2,1} + C_{i,2}C_{2,2} \Longrightarrow C_{i,2} = \dfrac{1}{C_{2,2}}\left(A_{i,2} - C_{i,1}C_{2,1}\right)$

Matrices SDP : principe de la factorisation de Cholesky

Unicité de la factorisation : $\Longrightarrow$ travaille par identification (diagonale /ligne).
Pour $i, j = 1, \dots, n$, en utilisant le fait que C est tri-inf :

$$A = CC^\top \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}(C^\top)_{k,j} \iff A_{i,j} = \sum_{k=1}^{\min(i,j)} C_{i,k}C_{j,k}$$

- $A_{1,1} = C_{1,1}C_{1,1} \quad \Longrightarrow C_{1,1} = \sqrt{A_{1,1}}$

  Si $i > 1$: $A_{i,1} = C_{i,1}C_{1,1} \quad \Longrightarrow C_{i,1} = \dfrac{A_{i,1}}{C_{1,1}}$

- $A_{2,2} = C_{2,1}C_{2,1} + C_{2,2}C_{2,2} \quad \Longrightarrow C_{2,2} = \sqrt{A_{2,2} - C_{2,1}^2}$

  Si $i > 2$: $A_{i,2} = C_{i,1}C_{2,1} + C_{i,2}C_{2,2} \quad \Longrightarrow C_{i,2} = \dfrac{1}{C_{2,2}}\left(A_{i,2} - C_{i,1}C_{2,1}\right)$

- Pour $j > 1$ donné :

$$A_{j,j} = C_{j,1}^2 + C_{j,2}^2 + \ldots + C_{j,j-1}^2 + C_{j,j}^2 \quad \Longrightarrow C_{j,j} = \sqrt{A_{j,j} - \sum_{k=1}^{j-1} C_{j,k}^2}$$

Si $i > j$: $A_{i,j} = C_{i,1}C_{j,1} + C_{i,2}C_{j,2} + \ldots + C_{i,j-1}C_{j,j-1} + C_{i,j}C_{j,j}$

$$\Longrightarrow C_{i,j} = \frac{1}{C_{j,j}}\left(A_{i,j} - \sum_{k=1}^{j-1} C_{i,k}C_{j,k}\right)$$

Matrices SDP : algorithme de Cholesky

Algorithme de Cholesky

```
Pour jj = 1 jusqu'à n
    s ⟵ A(jj, jj) - Σ_{kk=1}^{jj-1} C(jj, kk)^2
    Si s < tol
        Sortir, erreur : A non SDP
    Fin Si
    C(jj, jj) ⟵ sqrt(s)
    Pour ii = jj+1 jusqu'à n   (sauté pour jj=n)
        C(ii, jj) ⟵ 1 / C(jj, jj) ×
                    ( A(ii, jj) - Σ_{kk=1}^{jj-1} C(ii, kk) × C(jj, kk))
    Fin Pour
Fin Pour
```

$$s \longleftarrow A(jj,jj) - \sum_{kk=1}^{jj-1} C(jj,kk)^2$$

$$C(ii,jj) \longleftarrow 1 / C(jj,jj) \times \left(A(ii,jj) - \sum_{kk=1}^{jj-1} C(ii,kk) \times C(jj,kk) \right)$$

Exercice

Calculer la factorisation $A = CC^\top$ avec l'algorithme de Cholesky pour

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 9 \end{bmatrix}$$

Montrer que l'on retrouve le C obtenu par la décomposition en carrés de Gauss.

Plan

Normes vectorielles

Définition (2.6.1 norme vectorielle)

Soit E un espace vectoriel sur K ($K = \mathbb{R}$ ou $\mathbb{C}$).
On appelle norme sur E une application de E dans $\mathbb{R}$ notée

$$x \mapsto \|x\|$$

possédant les propriétés suivantes : quel que soit $x \in E$, $y \in E$, $\lambda \in K$:

1. $\|x\| \geq 0$
2. $\|x\| = 0 \iff x = 0$
3. $\|\lambda x\| = |\lambda| \, \|x\|$ (la norme est positivement homogène)
4. $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire)

Pour $a \in \mathbb{C}$ ou $\mathbb{R}$, $|a| =$ module ou valeur absolue de a ($|a| = \sqrt{a\bar{a}}$).

Normes vectorielles : exemples

Exemple (Normes vectorielles usuelles)

1. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_1 = \sum_{i=1}^{n} |x_i|,$
2. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x|_i^2},$ (norme euclidienne),
3. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$

Normes vectorielles : exemples

Exemple (Normes vectorielles usuelles)

1. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_1 = \sum_{i=1}^{n} |x_i|,$
2. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x|_i^2},$ (norme euclidienne),
3. $E = \mathbb{R}^n$ ou $\mathbb{C}^n$: $\quad \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$

Théorème (Normes vectorielles équivalentes sur $\mathbb{R}^n$)

Toutes les normes vectorielles sont équivalentes sur $\mathbb{R}^n$. Soit N une norme sur $\mathbb{R}^n$, alors

$$\exists \alpha, \beta > 0 \ : \quad \forall x \in \mathbb{R}^n \quad \alpha N(x) \leq \|x\|_\infty \leq \beta N(x).$$

Pour les normes usuelles, on a :

$$\|x\|_\infty \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty.$$

seconde inégalité découle de Cauchy-Schwarz.

Normes vectorielles : autres exemples

Exemple (Normes vectorielles sur des espaces de dimension infinie)

1. $E = \mathcal{C}_0([0,1])$: $\quad \|f\|_{L^1} = \displaystyle\int_0^1 |f(t)|\mathrm{d}t,$
2. $E = \mathcal{C}_0([0,1])$: $\quad \|f\|_{L^2} = \sqrt{\displaystyle\int_0^1 f^2(t)\mathrm{d}t},$
3. $E = \mathcal{C}_0([0,1])$: $\quad \|f\|_{L^\infty} = \max\limits_{t\in[0,1]} |f(t)|,$
4. ...

Espaces euclidiens (norme euclidienne)

Théorème (Rappels sur les espaces euclidiens)

Si $E = \mathbb{R}^n$ *(résultats similaires adaptés pour* $\mathbb{C}^n$*), pour* $x, y \in \mathbb{R}^n$ *:*

1. *Norme euclidienne associée à un produit scalaire :*

$$\|x\|_2 = \sqrt{\langle x, x\rangle}, \quad \text{avec } \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

2. *Inégalité de Cauchy-Schwarz (cf. Chap. 2 Exo C.1.8) :*

$$|\langle x, y\rangle| \leq \|x\|_2 \, \|y\|_2 \iff \left|\sum_{i=1}^{n} x_i y_i\right| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}.$$

Égalité dans CS $\iff$ x *et* y *sont liés.*

3. *Théorème de Pythagore :*

$$\|x + y\|_2^2 = \|x\|_2^2 + 2\langle x, y\rangle + \|y\|_2^2.$$

4. *Voir Chap 3...*

Normes matricielles

Définition (2.6.2 norme matricielle)

On appelle **norme matricielle** une application de $\mathcal{M}_{mn}(\mathbb{R})$ dans $\mathbb{R}$ notée

$$A \mapsto \|A\|,$$

vérifiant :

1. $\|A\| \geq 0,$
2. $\|A\| = 0 \iff A = 0,$
3. $\|\lambda A\| = |\lambda| \, \|A\|, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{R}), \forall \lambda \in \mathbb{R},$
4. $\|A + B\| \leq \|A\| + \|B\|, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{R}), \forall B \in \mathcal{M}_{mn}(\mathbb{R}),$
5. $\|AB\| \leq \|A\| \, \|B\|, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{R}), \forall B \in \mathcal{M}_{np}(\mathbb{R}).$

On déduit :

$$\|B^k\| \leq \|B\|^k, \quad \forall B \in \mathcal{M}_n(\mathbb{R}).$$

Rappels sur sup et max

Soit une fonction $f : X \subset \mathbb{R} \to \mathbb{R}$.

Definition

Le sup, quand il existe, est le plus petit des majorants de f sur X :

$$\forall x \in X \ : \quad f(x) \leq \sup_{x \in X} f(x), \qquad \forall \varepsilon > 0 \ \exists x_\varepsilon \in X \ : \quad \sup_{x \in X} f(x) \leq f(x_\varepsilon) + \varepsilon.$$

Le sup est un max quand il est atteint, c'est-à-dire si

$$\exists \tilde{x} \in X \ : \quad f(\tilde{x}) = \sup_{x \in X} f(x) = \max_{x \in X} f(x).$$

Rappels sur sup et max

Soit une fonction $f : X \subset \mathbb{R} \to \mathbb{R}$.

Definition

Le sup, quand il existe, est le plus petit des majorants de f sur X :

$$\forall x \in X \ : \quad f(x) \leq \sup_{x \in X} f(x), \qquad \forall \varepsilon > 0 \ \exists x_\varepsilon \in X \ : \quad \sup_{x \in X} f(x) \leq f(x_\varepsilon) + \varepsilon.$$

Le sup est un max quand il est atteint, c'est-à-dire si

$$\exists \tilde{x} \in X \ : \quad f(\tilde{x}) = \sup_{x \in X} f(x) = \max_{x \in X} f(x).$$

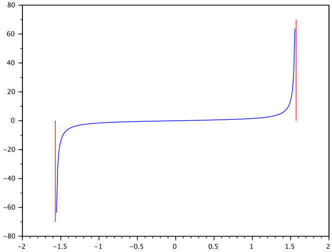
On a

Lemme

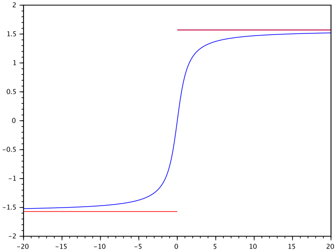
1. Si f est majorée ($\exists M \in \mathbb{R}, \forall x \in X \ : f(x) \leq M$), alors le sup existe dans $\mathbb{R}$ et $\sup_{x \in X} f(x) \leq M$. (propriété de $\mathbb{R}$, fausse dans $\mathbb{Q}$.)
2. Si $\lambda > 0$: $\sup_{x \in X}(\lambda f(x)) = \lambda \sup_{x \in X} f(x)$
3. $\sup_{x \in X} (f(x) + g(x)) \leq \sup_{x \in X} f(x) + \sup_{x \in X} g(x)$

Exemples de sup et max

$]-\frac{\pi}{2}, \frac{\pi}{2}[\ni x \to \tan(x) \in \mathbb{R}$



$\mathbb{R} \ni x \to \arctan(x) \in]-\frac{\pi}{2}, \frac{\pi}{2}[$



- Sur $\mathbb{R}$, $\arctan$ admet une infinité de majorants : $M = 2$ est un majorant.
- Sur $\mathbb{R}$, $\arctan$ admet un $\sup$ (le plus petit des majorants) : $\frac{\pi}{2}$, mais pas de $\max$ (le $\sup$ n'est pas atteint).
- Sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, $\tan$ n'admet pas de $\sup$.
- Sur $[-\frac{\pi}{4}, \frac{\pi}{4}]$, $\tan$ admet un $\sup$ (qui vaut 1), et qui est atteint en $\frac{\pi}{4}$: ce $\sup$ est un $\max$.
- Sur $]-\frac{\pi}{4}, \frac{\pi}{4}[$, $\tan$ admet un $\sup$ (qui vaut encore 1), mais qui n'est pas atteint sur l'intervalle : pas de $\max$.

Normes matricielles subordonnées

Définition (2.6.3 norme matricielle subordonnée)

Soient deux normes vectorielles sur $\mathbb{R}^n$ et sur $\mathbb{R}^m$, notées $\|\cdot\|$. On appelle **norme matricielle subordonnée** l'application de $\mathcal{M}_{mn}(\mathbb{R})$ dans $\mathbb{R}$ définie par

$$A \mapsto |||A||| = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

Normes matricielles subordonnées

Théorème (norme matricielle subordonnée)

Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$. On a :

1. *L'application $A \mapsto |||A|||$ est bien définie.*
2. *L'application $A \mapsto |||A|||$ est une norme matricielle.*
3. *Dans la définition, le $\sup$ est atteint, c'est un $\max$.*
4. *On a de plus pour toute norme subordonnée :*

$$\|Ax\| \leq |||A|||\, \|x\|, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{R}), \forall x \in \mathbb{R}^n.$$

5. *Pour $I \in \mathcal{M}_{nn}(\mathbb{R})$, si on prend la même norme vectorielle pour $\mathbb{R}^n$: $|||I||| = 1$.*

Normes matricielles subordonnées

Théorème (norme matricielle subordonnée)

Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$. On a :

1. *L'application $A \mapsto |||A|||$ est bien définie.*
2. *L'application $A \mapsto |||A|||$ est une norme matricielle.*
3. *Dans la définition, le $\sup$ est atteint, c'est un $\max$.*
4. *On a de plus pour toute norme subordonnée :*

$$\|Ax\| \leq |||A|||\,\|x\|, \quad \forall A \in \mathcal{M}_{mn}(\mathbb{R}), \forall x \in \mathbb{R}^n.$$

5. *Pour $I \in \mathcal{M}_{nn}(\mathbb{R})$, si on prend la même norme vectorielle pour $\mathbb{R}^n$: $|||I||| = 1$.*

Démonstrations (partielles) : voir polycopié et notes de cours.

1. bien définie : soit $x \in \mathbb{R}^n$ décomposé dans la base canonique de $\mathbb{R}^n$: $x = \sum_{j=1}^n x_j e_j$. Alors en utilisant l'inégalité triangulaire et l'équivalence avec la norme infinie ($\|x\|_\infty \leq \beta \|x\|$), il vient :
$\|Ax\| \leq \sum_{j=1}^n |x_j|\,\|Ae_j\| \leq \|x\|_\infty \sum_{j=1}^n \|A_j\| \leq (\beta \sum_{j=1}^n \|A_j\|)\|x\|$.
$\Longrightarrow$ si $x \neq 0$: division par $\|x\|$ puis passage au $\sup$: $|||A|||$ existe et $|||A||| \leq \beta \sum_{j=1}^n \|A_j\|$.

Normes matricielles subordonnées usuelles : valeurs

Proposition (Normes matricielles subordonnées)

1. *avec* $\|\cdot\|_1$ *pour* $\mathbb{R}^n$ *et* $\mathbb{R}^m$ ($|||A|||_1 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_1}{\|x\|_1}$), *on montre :*

$$|||A|||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{i,j}| = \max_{1\leq j\leq n} \|A_j\|_1.$$

2. *avec* $\|\cdot\|_\infty$ *pour* $\mathbb{R}^n$ *et* $\mathbb{R}^m$ ($|||A|||_\infty = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}$), *on montre :*

$$|||A|||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |A_{i,j}| = \max_{1\leq j\leq n} \|(\underline{A}_i)^\top\|_1.$$

3. $|||A|||_2$ *: voir plus loin...*

Normes matricielles subordonnées usuelles : valeurs

Proposition (Normes matricielles subordonnées)

1. *avec* $\|\cdot\|_1$ *pour* $\mathbb{R}^n$ *et* $\mathbb{R}^m$ ($|||A|||_1 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_1}{\|x\|_1}$), *on montre :*

$$|||A|||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{i,j}| = \max_{1\leq j\leq n} \|A_j\|_1.$$

2. *avec* $\|\cdot\|_\infty$ *pour* $\mathbb{R}^n$ *et* $\mathbb{R}^m$ ($|||A|||_\infty = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}$), *on montre :*

$$|||A|||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |A_{i,j}| = \max_{1\leq j\leq n} \|(\underline{A}_i)^\top\|_1.$$

3. $|||A|||_2$ *: voir plus loin...*

Note : il existe des normes matricielles qui ne sont **pas** subordonnées.

Exemple : la norme de Frobenius : $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |A_{i,j}|^2} = \sqrt{\mathrm{trace}(A^\top A)}$.

Exercice : montrer que c'est une norme matricielle. Puis calculer $\|I\|_F$ et conclure que ce n'est pas une norme subordonnée.

Norme matricielle subordonnée à la norme 1

On pose $N(A) = \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{i,j}| = \max_{j=1,\ldots,n}(\|A_j\|_1)$. On montre $|||A|||_1 = N(A)$.

1. **Majoration : on montre que $|||A|||_1 \leq N(A)$.**
   Soit $x \neq 0$ dans $\mathbb{R}^n$. On a $Ax = \sum_{j=1}^{n} A_j x_j \in \mathbb{R}^m$. On prend la norme et on majore :

$$\|Ax\|_1 \leq \sum_{j=1}^{n} |x_j|\, \|A_j\|_1 \leq \max_{j=1,\ldots,n}(\|A_j\|_1)(\sum_{j=1}^{n} |x_j|) = N(A)\|x\|_1.$$

   On divise par $\|x\|_1 > 0$ et on obtient :

$$\forall x \neq 0 \quad \frac{\|Ax\|_1}{\|x\|_1} \leq N(A),$$

   donc $N(A)$ est un majorant, le $\sup$ existe et $|||A|||_1 \leq N(A)$.

Norme matricielle subordonnée à la norme 1

On pose $N(A) = \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{i,j}| = \max_{j=1,\ldots,n}(\|A_j\|_1)$. On montre $|||A|||_1 = N(A)$.

1. **Majoration : on montre que $|||A|||_1 \leq N(A)$.**
   Soit $x \neq 0$ dans $\mathbb{R}^n$. On a $Ax = \sum_{j=1}^{n} A_j x_j \in \mathbb{R}^m$. On prend la norme et on majore :

$$\|Ax\|_1 \leq \sum_{j=1}^{n} |x_j|\,\|A_j\|_1 \leq \max_{j=1,\ldots,n}(\|A_j\|_1)(\sum_{j=1}^{n}|x_j|) = N(A)\|x\|_1.$$

   On divise par $\|x\|_1 > 0$ et on obtient :

$$\forall x \neq 0 \quad \frac{\|Ax\|_1}{\|x\|_1} \leq N(A),$$

   donc $N(A)$ est un majorant, le $\sup$ existe et $|||A|||_1 \leq N(A)$.

2. **Minoration : on cherche un $\tilde{x}$ particulier pour lequel $\frac{\|A\tilde{x}\|_1}{\|\tilde{x}\|_1}$ vaut $N(A)$.**
   Soit j_0 tel que $N(A) = \sum_{i=1}^{n} |A_{i,j_0}| = \|A_{j_0}\|_1$ ($\max_j$ atteint en un j_0).
   On prend $\tilde{x} = e_{j_0}$, qui vérifie $Ae_{j_0} = A_{j_0}$ et $\|e_{j_0}\|_1 = 1$: $\|Ae_{j_0}\|_1 = \|A_{j_0}\|_1 = N(A)$.
   $\Longrightarrow \; N(A) = \frac{\|Ae_{j_0}\|_1}{\|e_{j_0}\|_1} \leq \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_1}{\|x\|_1} = |||A|||_1$. Il y a donc égalité.

Norme matricielle subordonnée à la norme 2

Théorème (2.6.1)

Si $A \in \mathcal{M}_{nn}(\mathbb{R})$ est symétrique alors

$$\max_{1 \le k \le n} \lambda_k = \max_{x \in \mathbb{R}^n, x \ne 0} \frac{x^\top A x}{x^\top x} = \max_{x \in \mathbb{R}^n, x \ne 0} \frac{\langle x, Ax \rangle}{\langle x, x \rangle},$$

où $(\lambda_k)_{1 \le k \le n}$ sont les n valeurs propres réelles, distinctes ou non, de A.

Norme matricielle subordonnée à la norme 2

Démonstration : A symétrique réelle : $\implies$ admet une base de vecteurs propres orthonormés $(Y_i)_{1\leq i\leq n}$ associés aux valeurs propres réelles ordonnées : $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. On a :
$\langle Y_i, Y_j\rangle = Y_i^\top Y_j = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$ (et donc $\|Y_i\|_2 = 1$) et $AY_i = \lambda_i Y_i$.

Norme matricielle subordonnée à la norme 2

Démonstration : A symétrique réelle : $\Longrightarrow$ admet une base de vecteurs propres orthonormés $(Y_i)_{1\leq i\leq n}$ associés aux valeurs propres réelles ordonnées : $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. On a :

$\langle Y_i, Y_j\rangle = Y_i^\top Y_j = \delta_{ij} = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{sinon} \end{cases}$ (et donc $\|Y_i\|_2 = 1$) et $AY_i = \lambda_i Y_i$.

1. **Majoration :** Soit $x \in \mathbb{R}^n$ écrit dans la BON $(Y_i)_i$: $x = \sum_{i=1}^n \xi_i Y_i$. On déduit $Ax = \sum_{i=1}^n \xi_i \lambda_i Y_i$. Il vient en utilisant l'orthonormalité des $(Y_i)_i$:

$$\|x\|_2^2 = \left(\sum_{i=1}^n \xi_i Y_i^\top\right)\left(\sum_{j=1}^n \xi_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j Y_i^\top Y_j = \sum_{i=1}^n \xi_i^2$$

$$x^\top Ax = \left(\sum_{i=1}^n \xi_i Y_i^\top\right)\left(\sum_{j=1}^n \xi_j \lambda_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^n \xi_i \xi_j \lambda_j Y_i^\top Y_j = \sum_{i=1}^n \lambda_i \xi_i^2.$$

Donc : $x^\top Ax \leq \lambda_1 \sum_{i=1}^n \xi_i^2 = \lambda_1 \|x\|_2^2$. D'où $\forall x \neq 0$ $\dfrac{x^\top Ax}{x^\top x} \leq \lambda_1$, donc λ_1 est un majorant, le sup existe et $\sup\limits_{x\in\mathbb{R}^n\setminus\{0\}} \dfrac{x^\top Ax}{x^\top x} \leq \lambda_1 = \max\limits_{1\leq k\leq n} \lambda_k$.

Norme matricielle subordonnée à la norme 2

Démonstration : A symétrique réelle : $\Longrightarrow$ admet une base de vecteurs propres orthonormés $(Y_i)_{1\leq i\leq n}$ associés aux valeurs propres réelles ordonnées : $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. On a :
$\langle Y_i, Y_j\rangle = Y_i^\top Y_j = \delta_{ij} = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{sinon} \end{cases}$ (et donc $\|Y_i\|_2 = 1$) et $AY_i = \lambda_i Y_i$.

1. **Majoration :** Soit $x \in \mathbb{R}^n$ écrit dans la BON $(Y_i)_i$: $x = \sum_{i=1}^n \xi_i Y_i$. On déduit $Ax = \sum_{i=1}^n \xi_i \lambda_i Y_i$. Il vient en utilisant l'orthonormalité des $(Y_i)_i$:

$$\begin{aligned}
\|x\|_2^2 &= \left(\sum_{i=1}^n \xi_i Y_i^\top\right)\left(\sum_{j=1}^n \xi_j Y_j\right) = \sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j Y_i^\top Y_j = \sum_{i=1}^n \xi_i^2\\
x^\top Ax &= \left(\sum_{i=1}^n \xi_i Y_i^\top\right)\left(\sum_{j=1}^n \xi_j\lambda_j Y_j\right) = \sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j\lambda_j Y_i^\top Y_j = \sum_{i=1}^n \lambda_i\xi_i^2.
\end{aligned}$$

Donc : $x^\top Ax \leq \lambda_1 \sum_{i=1}^n \xi_i^2 = \lambda_1\|x\|_2^2$. D'où $\forall x \neq 0$ $\dfrac{x^\top Ax}{x^\top x} \leq \lambda_1$, donc λ_1 est un majorant, le sup existe et $\sup\limits_{x\in\mathbb{R}^n\setminus\{0\}} \dfrac{x^\top Ax}{x^\top x} \leq \lambda_1 = \max\limits_{1\leq k\leq n} \lambda_k$.

2. **Minoration :** Pour obtenir l'égalité il suffit de remarquer que $(Y_1)^\top AY_1 = \lambda_1\|Y_1\|_2^2 = \lambda_1$. Donc

$$\lambda_1 = \frac{(Y_1)^\top AY_1}{\|Y_1\|_2^2} \leq \sup_{x\in\mathbb{R}^n\setminus\{0\}} \frac{x^\top Ax}{x^\top x},$$

donc il y a égalité, le sup est atteint, c'est un max.

Norme matricielle subordonnée à la norme 2

Définition (Spectre et rayon spectral)

Soit $A \in \mathcal{M}_{nn}(\mathbb{C})$. On appelle spectre de A l'ensemble

$$\mathrm{Sp}(A) = \{\lambda_i(A) \mid 1 \leq i \leq n\} \subset \mathbb{C},$$

où les $\lambda_i(A)$ sont les valeurs propres de A. On appelle rayon spectral de A le réel

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i(A)| \in \mathbb{R}.$$

Proposition (Propriétés du rayon spectral)

On a :

1. $\mathrm{Sp}(AB) \setminus \{0\} = \mathrm{Sp}(BA) \setminus \{0\}$ *et* $\rho(AB) = \rho(BA)$, $\forall A \in \mathcal{M}_{np}(\mathbb{C})$ *et* $B \in \mathcal{M}_{pn}(\mathbb{C})$.
2. $\rho(A) \leq |||A|||$, *pour toute norme subordonnée* $|||\cdot|||$ *et* $\forall A \in \mathcal{M}_{nn}(\mathbb{C})$.
3. $\rho(A^2) = \rho(A)^2$, $\forall A \in \mathcal{M}_{nn}(\mathbb{C})$.

Démonstration : (1) et (3, cas symétrique) voir Chapitre 2, exercices TD n^o 11 et 12.

Norme matricielle subordonnée à la norme 2

Théorème (2.6.2 : norme euclidienne subordonnée)

1. *Soit $A \in \mathcal{M}_{np}(\mathbb{R})$, alors*

$$\begin{aligned} |||A|||_2^2 &= \rho(AA^\top) = \rho(A^\top A), \\ |||A|||_2^2 &= \left|\left|\left|AA^\top\right|\right|\right|_2 = \left|\left|\left|A^\top A\right|\right|\right|_2. \end{aligned}$$

2. *Si la matrice C est symétrique, alors*

$$|||C|||_2 = \rho(C).$$

Norme matricielle subordonnée à la norme 2

Théorème (2.6.2 : norme euclidienne subordonnée)

1. *Soit $A \in \mathcal{M}_{np}(\mathbb{R})$, alors*

$$
\begin{aligned}
|||A|||_2^2 &= \rho(AA^\top) = \rho(A^\top A), \\
|||A|||_2^2 &= \left|\left|\left| AA^\top \right|\right|\right|_2 = \left|\left|\left| A^\top A \right|\right|\right|_2 .
\end{aligned}
$$

2. *Si la matrice C est symétrique, alors*

$$|||C|||_2 = \rho(C).$$

Démonstration : (2) voir Chapitre 2, exercices C.1.23 et TD n^o 11.
On montre $|||A|||_2^2 = \rho(AA^\top)$. ($\rho(A^\top A) = \rho(A^\top A)$ car $\rho(AB) = \rho(BA)$.)

Norme matricielle subordonnée à la norme 2

Théorème (2.6.2 : norme euclidienne subordonnée)

1. *Soit $A \in \mathcal{M}_{np}(\mathbb{R})$, alors*

$$\begin{aligned} |||A|||_2^2 &= \rho(AA^\top) = \rho(A^\top A), \\ |||A|||_2^2 &= \left|\left|\left| AA^\top \right|\right|\right|_2 = \left|\left|\left| A^\top A \right|\right|\right|_2. \end{aligned}$$

2. *Si la matrice C est symétrique, alors*

$$|||C|||_2 = \rho(C).$$

Démonstration : (2) voir Chapitre 2, exercices C.1.23 et TD n^o 11.
On montre $|||A|||_2^2 = \rho(AA^\top)$. ($\rho(A^\top A) = \rho(A^\top A)$ car $\rho(AB) = \rho(BA)$.)
On note : $|||A|||_2^2 = \left(\sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}\right)^2 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2^2}{\|x\|_2^2}$, car sup de termes ≥ 0.

Norme matricielle subordonnée à la norme 2

Théorème (2.6.2 : norme euclidienne subordonnée)

1. *Soit $A \in \mathcal{M}_{np}(\mathbb{R})$, alors*

$$\begin{aligned} |||A|||_2^2 &= \rho(AA^\top) = \rho(A^\top A), \\ |||A|||_2^2 &= \left|\left|\left|AA^\top\right|\right|\right|_2 = \left|\left|\left|A^\top A\right|\right|\right|_2. \end{aligned}$$

2. *Si la matrice C est symétrique, alors*

$$|||C|||_2 = \rho(C).$$

Démonstration : (2) voir Chapitre 2, exercices C.1.23 et TD n^o 11.
On montre $|||A|||_2^2 = \rho(AA^\top)$. ($\rho(A^\top A) = \rho(A^\top A)$ car $\rho(AB) = \rho(BA)$.)
On note : $|||A|||_2^2 = \left(\sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}\right)^2 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2^2}{\|x\|_2^2}$, car sup de termes ≥ 0.
On a : $\|Ax\|_2^2 = (Ax)^\top(Ax) = x^\top A^\top Ax = x^\top(A^\top A)x = x^\top Bx$ où l'on a noté $B = A^\top A \in \mathcal{M}_{nn}$.
B est symétrique semi-DP (car $x^\top Bx = \|Ax\|_2^2 \geq 0$), donc ses n valeurs propres $\mu_k(B)$ sont ≥ 0.

Norme matricielle subordonnée à la norme 2

Théorème (2.6.2 : norme euclidienne subordonnée)

1. *Soit $A \in \mathcal{M}_{np}(\mathbb{R})$, alors*

$$\begin{aligned} |||A|||_2^2 &= \rho(AA^\top) = \rho(A^\top A), \\ |||A|||_2^2 &= \left|\left|\left|AA^\top\right|\right|\right|_2 = \left|\left|\left|A^\top A\right|\right|\right|_2 . \end{aligned}$$

2. *Si la matrice C est symétrique, alors*

$$|||C|||_2 = \rho(C).$$

Démonstration : (2) voir Chapitre 2, exercices C.1.23 et TD n^o 11.
On montre $|||A|||_2^2 = \rho(AA^\top)$. ($\rho(A^\top A) = \rho(A^\top A)$ car $\rho(AB) = \rho(BA)$.)
On note : $|||A|||_2^2 = \left(\sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}\right)^2 = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2^2}{\|x\|_2^2}$, car sup de termes ≥ 0.
On a : $\|Ax\|_2^2 = (Ax)^\top(Ax) = x^\top A^\top Ax = x^\top(A^\top A)x = x^\top Bx$ où l'on a noté $B = A^\top A \in \mathcal{M}_{nn}$.
B est symétrique semi-DP (car $x^\top Bx = \|Ax\|_2^2 \geq 0$), donc ses n valeurs propres $\mu_k(B)$ sont ≥ 0.
$\Longrightarrow$ Théorème (2.6.1) : $\rho(A^\top A) = \rho(B) = \max_{1\leq k\leq n} |\mu_k(B)| = \max_{1\leq k\leq n} (\mu_k(B)) = \max_{x\in\mathbb{R}^n, x\neq 0} \frac{x^\top Bx}{x^\top x} = \max_{x\in\mathbb{R}^n, x\neq 0} \frac{x^\top A^\top Ax}{x^\top x} = \max_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|_2^2}{\|x\|_2^2} = |||A|||_2^2$.

Norme matricielle subordonnée à la norme 2

Exemple (norme euclidienne subordonnée)

$\rho(A) \neq |||A|||_2$ en général.

Soit : $A = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$. $\mathrm{Sp}(A) = \{0\}$, donc $\rho(A) = 0$, mais $A \neq 0$, donc $|||A|||_2 \neq 0$! !

$A^\top A = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$, donc $\mathrm{Sp}(A^\top A) = \{0, 4\}$ et $\rho(A^\top A) = 4$, donc $|||A|||_2 = 2$! !

Plan

Conditionnement : définition

Définition (Conditionnement)

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ inversible. On appelle conditionnement de A relatif à la norme $|||\cdot|||$ le réel

$$\chi(A) = |||A||| \left|\left|\left|A^{-1}\right|\right|\right| \in \mathbb{R}.$$

On a : $1 = |||I||| = \left|\left|\left|AA^{-1}\right|\right|\right| \leq |||A||| \left|\left|\left|A^{-1}\right|\right|\right| = \chi(A)$. Donc :

Conditionnement : définition

Définition (Conditionnement)

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ inversible. On appelle conditionnement de A relatif à la norme $|||\cdot|||$ le réel

$$\chi(A) = |||A||| \left|\left|\left| A^{-1} \right|\right|\right| \in \mathbb{R}.$$

On a : $1 = |||I||| = \left|\left|\left| AA^{-1} \right|\right|\right| \leq |||A||| \left|\left|\left| A^{-1} \right|\right|\right| = \chi(A)$. Donc :

Proposition

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ inversible. Alors $\chi(A) \geq 1$.

Conditionnement : définition

Définition (Conditionnement)

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ inversible. On appelle conditionnement de A relatif à la norme $|||\cdot|||$ le réel

$$\chi(A) = |||A||| \left|\left|\left|A^{-1}\right|\right|\right| \in \mathbb{R}.$$

On a : $1 = |||I||| = \left|\left|\left|AA^{-1}\right|\right|\right| \leq |||A||| \left|\left|\left|A^{-1}\right|\right|\right| = \chi(A)$. Donc :

Proposition

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ inversible. Alors $\chi(A) \geq 1$.

Proposition (Conditionnement χ_2, cas symétrique)

Soit $C \in \mathcal{M}_{nn}(\mathbb{R})$ symétrique et inversible. Alors

$$\chi_2(C) = \rho(C)\rho(C^{-1}) = \frac{\max\limits_{i=1,\ldots,n} |\lambda_i(C)|}{\min\limits_{i=1,\ldots,n} |\lambda_i(C)|}$$

Conditionnement : exemple

Soit le système $Ax = b$ dans $\mathbb{R}^4$:

$$A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}, \; b = \begin{bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{bmatrix} \quad \Longrightarrow \text{solution } x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

On perturbe le second membre : $b \to b + \delta b$

$$b + \delta b = \begin{bmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{bmatrix} \quad \Longrightarrow \text{solution } x + \delta x = \begin{bmatrix} 9.2 \\ -12.6 \\ 4.5 \\ 1.1 \end{bmatrix}$$

Écarts en norme infinie : $\frac{\|\delta b\|_\infty}{\|b\|_\infty} \approx 3.\ 10^{-3}$, $\quad \frac{\|\delta x\|_\infty}{\|x\|_\infty} = 13.6$. Et donc : $\frac{\frac{\|\delta x\|_\infty}{\|x\|_\infty}}{\frac{\|\delta b\|_\infty}{\|b\|_\infty}} \approx 4500$.

On calcule $|||A|||_\infty = 33$ et $\left|\left|\left|A^{-1}\right|\right|\right|_\infty = 136$, avec

$$A^{-1} = \begin{bmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{bmatrix}, \quad \text{et donc : } \chi_\infty(A) = |||A|||_\infty \left|\left|\left|A^{-1}\right|\right|\right|_\infty = 4488.$$

$\Longrightarrow$ $Ax = b$: système mal conditionné.