

Cours d'analyse numérique de licence de mathématiques

Roland Masson

Année 2015-2016

1 Introduction

- Objectifs
- Plan du cours
- Exemples d'applications du calcul scientifique
- Débouchés
- Calendrier du cours
- Evaluation

2 Quelques rappels d'algèbre linéaire en dimension finie

- Espaces vectoriels
- Applications linéaires
- Matrices
- Transposition de matrices et matrices symétriques
- Déterminants
- Normes matricielles

3 Méthodes itératives

Analyse numérique: objectifs

- Analyse numérique: conçoit et analyse mathématiquement les algorithmes à la base des simulations numériques de la physique
- Objectifs du cours
 - Introduction à quelques algorithmes de bases en calcul scientifique
 - Fondements mathématiques (complexité, stabilité, convergence, consistance, ...)
 - Exemples d'applications et mise en oeuvre informatique sous scilab (TDs et TPs)

Plan du cours

- Résolution des systèmes linéaires $Ax = b$, $A \in \mathcal{M}_n$, $b, x \in \mathbb{R}^n$
 - Méthodes itératives: méthodes de Richardson, de Jacobi, de Gauss Seidel
 - Méthodes directes: méthode d'élimination de Gauss, factorisation LU, factorisation de Choleski
- Résolution des systèmes non linéaires $f(x) = 0$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$
 - Méthodes de Newton
- Algorithmes d'optimisation: $x = \operatorname{argmin}_{y \in \mathbb{R}^n} f(y)$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$
 - Méthodes de descente selon le gradient
- Résolution des équations différentielles ordinaires (EDO): $y' = f(y, t)$, $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$
 - Schémas explicites
 - Schéma d'Euler implicite

Références

Site web: <http://math.unice.fr/~massonr/ANL3/20152016/ANL3.html>

Cours d'analyse numérique de Raphaèle Herbin (Université de Provence):
<http://www.cmi.univ-mrs.fr/~herbin/anamat.html>

Livre de P.G. Ciarlet: Introduction à l'analyse numérique matricielle et à l'optimisation.

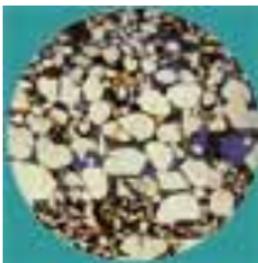
Livre de Quateroni et al: "Méthodes numériques, algorithmes, analyse et applications", Springer, 2007.

Domaines d'applications du calcul scientifique

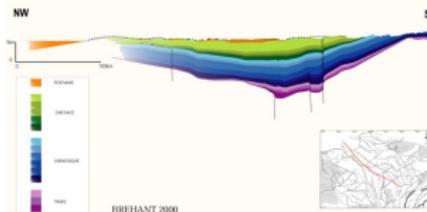
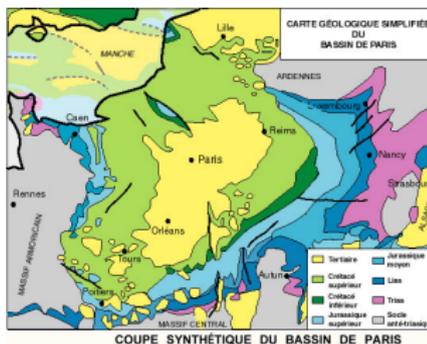
- Énergie
 - Nucléaire
 - Pétrole
 - Fusion nucléaire
 - Eolien, hydrolien, solaire, ...
- Transport
 - Aéronautique
 - Spatial
 - Automobile
- Environnement
 - Météorologie
 - Hydrologie
 - Géophysique
 - Climatologie
- Finance, Biologie, Santé, Télécommunications, Chimie, matériaux, ...

Exemple de la simulation des réservoirs pétroliers

- Pétrole = huile de pierre

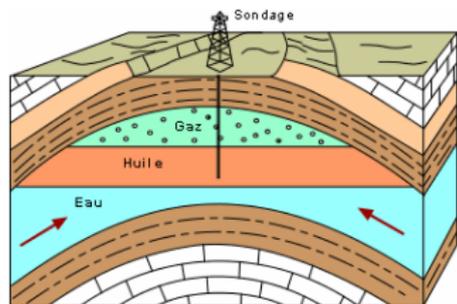


■ Bassin de paris

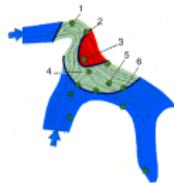


Exemple de la simulation des réservoirs pétroliers

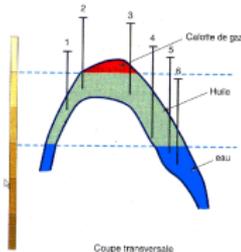
- Réservoir: piège géologique rempli d'hydrocarbures



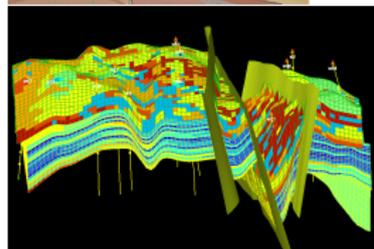
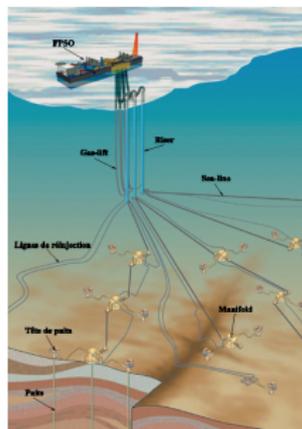
Piège structural: anticlinal



Dispositif vu du dessus et implantation des puits



Coupe transversale



Exemple de la simulation des réservoirs pétroliers

- Enjeux de la simulation
 - Prédiction de la production
 - Optimisation de la production (ou du rendement économique)
 - Intégration des données
 - Evaluation des incertitudes sur la production

■ Compétences

- Analyse numérique
- Modélisation
- Informatique

■ Métiers

- Développements de codes de calculs scientifiques
- Etudes en modélisation numérique
- Ingénieur de recherches en calcul scientifique
- Chercheur académique en mathématiques appliquées

■ Employeurs

- SSII en calcul scientifique
- EPIC: CEA, ONERA, IFPEN, BRGM, IFREMER, INRA, CERFACS, ...
- Industrie: EDF, EADS, Dassault, Michelin, Areva, Total, CGGVeritas, Thales, Safran, Veolia, Rhodia, ...
- Académique: Universités, CNRS, INRIA, Ecoles d'ingénieurs, ...

Calendrier du cours

- Cours en M32 le vendredi de 9h45 à 11h15
- TP en PV201 et PV202 deux groupes en binôme avec Francesca Rapetti le vendredi de 13h15 à 14h45
- TD en M32 le vendredi de 15h à 16h30

Calendrier du cours

sem 1: 11/09

sem 2: 18/9

sem 3: 25/09

sem 4: 02/10

sem 5: 09/10

sem 6: 16/10

sem 7: 23/10: Partiel

pause pédagogique

sem 8: 6/11

sem 9: 13/11

sem10: 20/11

sem11: 27/11: contrôle TP

sem12: 4/12

sem13: 11/12

Evaluation

- Un examen partiel semaine 7 le 23/10: note *Partiel*
- Un examen de TP semaine 11 le 27/11 en TP: note *TP*
- Un examen final: note *EF* (du 4 au 15 janvier)

Note finale: $0.4EF + 0.3Partiel + 0.3TP$

Espaces vectoriels

- Définition d'un e.v. sur $K = \mathbb{R}$ ou \mathbb{C} : ensemble E muni d'une loi de composition interne notée $+$ et d'une loi d'action de K sur E notée \cdot tels que:
 - $(E, +)$ est un groupe commutatif
 - $1 \cdot \mathbf{x} = \mathbf{x}$, $(\alpha\beta) \cdot \mathbf{x} = \alpha \cdot (\beta \cdot \mathbf{x})$ (associativité)
 - $(\alpha + \beta) \cdot \mathbf{x} = \alpha \cdot \mathbf{x} + \beta \cdot \mathbf{x}$, $\alpha \cdot (\mathbf{x} + \mathbf{y}) = \alpha \cdot \mathbf{x} + \alpha \cdot \mathbf{y}$ (distributivité)
- Exemple: \mathbb{R}^n e.v. sur \mathbb{R} (\mathbb{C}^n e.v. sur \mathbb{C}):

$$\bullet \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \cdot \mathbf{x} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

Familles libres, génératrices, base, dimension

- Famille libre de m vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_m$ de E :
 - $\sum_{i=1}^m \lambda_i \mathbf{v}_i = 0 \Rightarrow \lambda_i = 0 \forall i = 1, \dots, m$
- Famille génératrice de m vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_m$ de E :
 - $E = \text{Vect}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$
- Base: famille libre et génératrice
 - Dimension (supposée finie): toutes les bases ont même dimension appelée dimension de l'espace vectoriel E notée n
 - Une famille libre de n vecteurs est génératrice, c'est une base
 - Une famille génératrice de n vecteurs est libre, c'est une base

Espaces vectoriels normés

- Définition: e.v. muni d'une norme, ie une application de $E \rightarrow \mathbb{R}^+$, notée $\mathbf{x} \rightarrow \|\mathbf{x}\|$ satisfaisant les propriétés suivantes
 - $\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = 0$
 - $\|\lambda \cdot \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$
 - $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
- Une norme définit sur E une topologie d'espace métrique avec $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$
 - Limite de suite de vecteurs: $\lim_{k \rightarrow +\infty} \mathbf{v}_k = \mathbf{v} \Leftrightarrow \lim_{k \rightarrow +\infty} \|\mathbf{v}_k - \mathbf{v}\| = 0$
- Exemples de normes sur \mathbb{R}^n
 - $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$, $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$, $\|\mathbf{x}\|_\infty = \max_{i=1, \dots, n} |x_i|$.
- En dimension finie toutes les normes sont équivalentes ie il existe $c, C > 0$ telles que $c\|\mathbf{x}\| \leq \|\mathbf{x}\|_* \leq C\|\mathbf{x}\|$ (attention c et C dépendent de n).

Espaces vectoriels euclidiens

- e.v. muni d'un produit scalaire ie une forme bilinéaire symétrique définie positive notée $\langle \cdot, \cdot \rangle$
- Sur \mathbb{R}^n le produit scalaire canonique est $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$
- $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ est une norme appelée norme euclidienne

Applications linéaires

- $f : E \rightarrow F$, $f(\lambda \cdot \mathbf{x}) = \lambda \cdot f(\mathbf{x})$, $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$
- $\mathcal{L}(E, F)$ espace vectoriel des applications linéaires de E dans F
- $\mathcal{L}(E)$ espace vectoriel des applications linéaires de E dans E ou endomorphismes de E
- $(\mathcal{L}(E), +, \cdot, \circ)$ algèbre non commutative munie de la loi de composition des applications $f \circ g(\mathbf{x}) = f(g(\mathbf{x}))$
- Noyau de f , $\text{Ker}(f) = \{\mathbf{x} \in E \text{ tels que } f(\mathbf{x}) = 0\}$ (sous e.v. de E)
- Image de f , $\text{Im}(f) = \{f(\mathbf{x}), \mathbf{x} \in E\}$ (sous e.v. de F)
- Endomorphismes de E inversibles:
 - Application bijective ssi il existe $f^{-1} \in \mathcal{L}(E)$ telle que $f \circ f^{-1} = f^{-1} \circ f = \text{Id}$
 - f bijective $\Leftrightarrow f$ injective: $\text{Ker}(f) = \{0\}$
 - f bijective $\Leftrightarrow f$ surjective: $\text{Im}(f) = E$

Matrice d'une application linéaire

- Bases $(\mathbf{e}_j, j = 1, \dots, n)$ de E et $(\mathbf{f}_i, i = 1, \dots, m)$ de F
- $f \in \mathcal{L}(E, F)$ telle que $f(\mathbf{e}_j) = \sum_{i=1}^m A_{i,j} \mathbf{f}_i$
- $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j \in E$
- $\mathbf{y} = f(\mathbf{x}) = \sum_{i=1}^m \left(\sum_{j=1}^n A_{i,j} x_j \right) \mathbf{f}_i$
- $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m, \quad Y = AX$
- Retenir que les n colonnes j de A sont données par les images $f(\mathbf{e}_j)$
- Espace vectoriel des matrices de dimension m, n : $\mathcal{M}_{m,n}$ (à coefficients dans $K = \mathbb{R}$ ou \mathbb{C})
- Matrices remarquables: diagonale, symétrique, triangulaires inférieure ou supérieure

Exercice: produit de matrices versus composition d'applications linéaires

- Soient E, F, G des e.v de dimensions resp. n, m, p , $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$
- Des bases étant données, f a pour matrice $A \in \mathcal{M}_{m,n}$ et g a pour matrice $B \in \mathcal{M}_{p,m}$
- $g \circ f$ a pour matrice le produit $BA \in \mathcal{M}_{p,n}$ tel que

$$(BA)_{i,j} = \sum_{k=1}^m B_{i,k} A_{k,j}$$

- Produit de matrices: $\mathcal{M}_{p,m} \times \mathcal{M}_{m,n} \rightarrow \mathcal{M}_{p,n}$
 - produit matrice vecteur: $\mathcal{M}_{m,n} \times \mathcal{M}_{n,1} \rightarrow \mathcal{M}_{m,1}$
 - produit scalaire de deux vecteurs: ligne . colonne $\mathcal{M}_{1,n} \times \mathcal{M}_{n,1} \rightarrow \mathcal{M}_{1,1}$
 - produit tensoriel de deux vecteurs: colonne. ligne $\mathcal{M}_{n,1} \times \mathcal{M}_{1,n} \rightarrow \mathcal{M}_{n,n}$

Exercice: changements de base pour les vecteurs et les matrices

- P : matrice de passage d'une base dans une autre $\tilde{e}_j = \sum_{k=1}^n P_{k,j} \mathbf{e}_k$
(colonnes de la nouvelle base dans l'ancienne)
- Changement de base pour les coordonnées des vecteurs: $X = P\tilde{X}$.
- Changement de base pour les matrices des applications linéaires: $X = P\tilde{X}$,
 $Y = Q\tilde{Y}$ et $\tilde{Y} = \tilde{A}\tilde{X}$, $Y = AX$ implique que

$$\tilde{A} = Q^{-1}AP.$$

Matrices carrés inversibles

- $A \in \mathcal{M}_{n,n} = \mathcal{M}_n$ est inversible ssi l'une des propriétés suivantes est vérifiée
 - Il existe $A^{-1} \in \mathcal{M}_{n,n}$ tel que $AA^{-1} = A^{-1}A = I$
 - A est injective ie $AX = 0 \Rightarrow X = 0$
 - A est surjective ie $\text{Im}(A) = \{AX, X \in \mathbb{R}^n\} = \mathbb{R}^n$

- $A, B \in \mathcal{M}_n$ inversibles

$$(AB)^{-1} = B^{-1}A^{-1}$$

Transposition de matrices

- $A \in \mathcal{M}_{m,n}$, on définit $A^t \in \mathcal{M}_{n,m}$ par

$$(A^t)_{i,j} = A_{j,i} \text{ pour tous } i = 1, \dots, n, j = 1, \dots, m$$

- Produit scalaire canonique de deux vecteurs (colonnes) $X, Y \in \mathbb{R}^n$:

$$X^t Y = \sum_{i=1}^n X_i Y_i$$

- Matrice carrée $A \in \mathcal{M}_n$ est symétrique ssi

$$A^t = A$$

Diagonalisation d'une matrice carrée symétrique $A \in \mathcal{M}_n$

- Les valeurs propres sur \mathbb{C} d'une matrice réelle symétrique A sont réelles et il existe une base orthonormée de vecteurs propres $F^i \in \mathbb{R}^n$, $i = 1, \dots, n$ telle que

$$AF^i = \lambda_i F^i \text{ et } (F^i)^t F^j = \delta_{i,j} \text{ pour tous } i, j = 1, \dots, n$$

- Si P est la matrice de passage de la base canonique dans la base F^i , $i = 1, \dots, n$, alors on a

$$P^{-1} = P^t$$

et

$$P^t A P = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Déterminants de n vecteurs dans un e.v. E de dimension n pour une base donnée

- Unique forme n -linéaire alternée sur E valant 1 sur la base

- $\text{Det}(\mathbf{v}_1, \dots, \mathbf{v}, \dots, \mathbf{v}, \dots, \mathbf{v}_n) = 0$ (alternée)
- Antisymétrie:

$$\text{Det}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -\text{Det}(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

- On a donc aussi pour toute permutation σ de $\{1, \dots, n\}$,

$$\text{Det}(\mathbf{v}_1, \dots, \mathbf{v}_n) = \text{sign}(\sigma) \text{Det}(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(n)})$$

- Déterminant d'une matrice carrée $A =$ déterminant des vecteurs colonnes

$$\text{Det}(A) = \text{Det}(A_{.,1}, \dots, A_{.,n}) = \sum_{\sigma \in \Sigma_n} \prod_{i=1}^n \text{sign}(\sigma) A_{\sigma(i),i}$$

Propriétés du déterminant

- Les vecteurs colonnes de A sont libres ssi $\text{Det}(A) \neq 0$
- Donc A est inversible ssi $\text{Det}(A) \neq 0$
- $\text{Det}(AB) = \text{Det}(A)\text{Det}(B) = \text{Det}(BA)$
- $\text{Det}(A^t) = \text{Det}(A)$
- Développement par rapport aux lignes ou aux colonnes

$$\text{Det}(A) = \sum_{i=1}^n (-1)^{i+j} A_{i,j} \text{Det}(A^{(i,j)}) = \sum_{j=1}^n (-1)^{i+j} A_{i,j} \text{Det}(A^{(i,j)})$$

Normes matricielles

- Une norme matricielle sur l'e.v. \mathcal{M}_n est une norme telle que

$$\|AB\| \leq \|A\|\|B\|$$

- Une norme matricielle induite par une norme $\|\cdot\|$ sur \mathbb{R}^n est la norme matricielle définie par

$$\|A\| = \sup_{X \neq 0} \frac{\|AX\|}{\|X\|}$$

- On a pour une norme matricielle induite: $\|AX\| \leq \|A\|\|X\|$ pour tout $X \in \mathbb{R}^n$

Exercice: exemples de normes induites

- $\|A\|_\infty = \text{Sup}_{X \neq 0} \frac{\|AX\|_\infty}{\|X\|_\infty} = \max_{j=1, \dots, n} \sum_{i=1}^n |A_{i,j}|$
- $\|A\|_1 = \text{Sup}_{X \neq 0} \frac{\|AX\|_1}{\|X\|_1} = \max_{j=1, \dots, n} \sum_{i=1}^n |A_{i,j}|$
- $\|A\|_2 = \text{Sup}_{X \neq 0} \frac{\|AX\|_2}{\|X\|_2} = \rho({}^tAA)^{1/2}$

Norme et rayon spectral

- Spectre de A :
 $\text{spec}(A) = \{\lambda \in \mathbb{C} \text{ tel qu'il existe } u \in \mathbb{C}^n \text{ non nul tel que } Au = \lambda u\}$
- Rayon spectral $\rho(A) = \max_{\lambda \in \text{spec}(A)} |\lambda| \in \mathbb{C}$.

On admettra le lemme suivant qui montre la proximité entre norme et rayon spectral (voir preuve dans le livre de Ciarlet ou le cours de Raphaële Herbin)

- $\rho(A) \leq \|A\|$ quel que soit la norme matricielle sur \mathcal{M}_n
- Quels que soient $A \in \mathcal{M}_n$ et $\epsilon > 0$ il existe une norme induite $\|\cdot\|_{A,\epsilon}$ telle que

$$\|A\|_{A,\epsilon} \leq \rho(A) + \epsilon.$$

Théorème: convergence de la suite A^k pour $A \in \mathcal{M}_n$

- Les propositions suivantes sont équivalentes
 - (1) $\lim_{k \rightarrow +\infty} A^k = 0$
 - (2) $\lim_{k \rightarrow +\infty} A^k x = 0$ quel que soit $x \in \mathbb{R}^n$
 - (3) $\rho(A) < 1$
 - (4) Il existe une norme induite $\|\cdot\|$ telle que $\|A\| < 1$.

Théorème: convergence de la suite A^k pour $A \in \mathcal{M}_n$

Preuve:

- (1) \Rightarrow (2): car $\|A^k x\| \leq \|A^k\| \|x\|$
- (2) \Rightarrow (3): par la contraposée non (3) \Rightarrow non (2): $\rho(A) \geq 1$ implique qu'il existe un $u \in \mathbb{C}^n$ et $\lambda \in \mathbb{C}$ tels que $Au = \lambda u$ et $|\lambda| \geq 1$. Ceci implique que $A^k u = \lambda^k u$ et donc (2) ne peut être vérifié.
- (3) \Rightarrow (4): découle du lemme ci-dessus
- (4) \Rightarrow (1): on a $\|A^k\| \leq \|A\|^k$ avec $\|A\| < 1$ d'où (1)

Corollaire: suite A^k et rayon spectral

- Rayon spectral et suite des puissances d'une matrice: on démontre la propriété suivante: $\rho(A) = \lim_{k \rightarrow +\infty} \|A^k\|^{1/k}$ quel que soit la norme sur \mathcal{M}_n

Preuve: on a tout d'abord $\rho(A)^k = \rho(A^k) \leq \|A^k\|$ d'où $\rho(A) \leq \|A^k\|^{1/k}$
Soit $\epsilon > 0$. On considère $B_\epsilon = \frac{A}{\rho(A) + \epsilon}$, telle que $\rho(B_\epsilon) < 1$. D'après le théorème précédent $\lim_{k \rightarrow +\infty} (B_\epsilon)^k = 0$, donc il existe N tel que pour tout $k \geq N$ on a

$$\|(B_\epsilon)^k\| = \frac{\|A^k\|}{(\rho(A) + \epsilon)^k} \leq 1.$$

On déduit que $\rho(A) \leq \|A^k\|^{1/k} \leq \rho(A) + \epsilon$ ce qui prouve le corollaire.

Matrices de la forme $I + A$ ou $I - A$

- Si $\rho(A) < 1$ alors les matrices $I + A$ et $I - A$ sont inversibles
- La série de terme général A^k converge (vers $(I - A)^{-1}$ ssi $\rho(A) < 1$
 - Preuve: $\sum_{k=0}^N A^k (I - A) = I - A^{N+1}$ et utiliser le lemme précédent
- Si $\|A\| < 1$ pour une norme matricielle, alors $I - A$ est inversible et on a $\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$ (idem pour $I + A$)

Conditionnement

Soit $\|\cdot\|$ la norme induite dans \mathcal{M}_n par une norme $\|\cdot\|$ sur \mathbb{R}^n

Conditionnement de A inversible : $\text{Cond}(A) = \|A\| \|A^{-1}\|$

$$\begin{cases} \text{Cond}(A) \geq 1 \\ \text{Cond}(\alpha A) = \text{Cond}(A) \\ \text{Cond}(AB) \leq \text{Cond}(A)\text{Cond}(B) \end{cases}$$

Soit A inversible et σ_1, σ_n les vp min et max de $A^t A$, on a pour la norme $\|\cdot\|_2$

$$\text{Cond}_2(A) = \left(\frac{\sigma_n}{\sigma_1}\right)^{1/2}$$

On en déduit que $\text{Cond}_2(A) = 1$ ssi $A = \alpha Q$ où Q matrice orthogonale

Pour A SDP de vp min et max λ_1 et λ_n , on a pour la norme $\|\cdot\|_2$

$$\text{Cond}_2(A) = \frac{\lambda_n}{\lambda_1}$$

Erreur d'arrondi

Soit A une matrice inversible, on cherche à estimer l'influence sur la solution du système $Ax = b$ d'une erreur d'arrondi sur le second membre b

$$\begin{cases} Ax = b \\ A(x + \delta x) = b + \delta b \end{cases}$$

implique

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta b\|}{\|b\|}$$

Méthodes itératives: motivations

- Matrice creuse: $\mathcal{O}(n)$ termes non nuls
- Méthodes itératives: calcul d'une suite x^n faisant intervenir que des produits matrice - vecteur
- Pour une matrice creuse, une itération coûte $\mathcal{O}(n)$ opérations flottantes
- Le problème à résoudre est la maîtrise de la convergence de la suite x^n vers x et du nombre d'itérations
- Coût pour une matrice creuse et une convergence en nit itérations

$$\mathcal{O}(n.nit)$$

matrices creuses: exemple du Laplacien 2D sur maillage Cartésien

Maillage cartésien uniforme $(n+1) \times (n+1)$ du carré $\Omega = (0, 1) \times (0, 1)$ de pas $\Delta x = \frac{1}{n+1}$ Laplacien avec conditions limites homogènes:

$$\begin{cases} -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f \text{ sur } \Omega \\ u(x) = 0 \text{ sur } \partial\Omega \end{cases}$$

Discrétisation:

$$\begin{cases} \frac{1}{(\Delta x)^2} (4u_{i,j} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1}) = f_{i,j} \text{ pour } i, j = 1, \dots, n \\ u_{i,j} = 0 \text{ pour } i = 0, n+1, j = 0, \dots, n+1 \text{ et } j = 0, n+1, i = 0, \dots, n+1 \end{cases}$$

Système linéaire: numérotation des inconnues $k = i + (j-1)n$ de 1 à $N = n^2$

$AU = F$ avec A matrice pentadiagonale de largeur de bande $q = n$

Coût d'une méthode directe LU : $2Nn^2 = 2N^2$

Coût d'une méthode itérative convergente en nit itérations: $10N \cdot nit$

Méthode de Richardson

Soit $A \in \mathcal{M}_n$ inversible et $b \in \mathbb{R}^n$. Soit $x \in \mathbb{R}^n$ solution de

$$Ax = b.$$

Méthode de Richardson: soit $\alpha \in \mathbb{R}$, on construit une suite de solution x^k de la forme

$$x^{k+1} = x^k + \alpha(b - Ax^k).$$

On a donc

$$(x^{k+1} - x) = (I - \alpha A)(x^k - x),$$

$$(x^{k+1} - x) = (I - \alpha A)^k(x^1 - x),$$

$$B = (I - \alpha A)$$

- Convergence: ssi $\rho(B) < 1$
- Taux de convergence: $\|B^k\|^{1/k} \rightarrow \rho(I - \alpha A)$

Méthode de Richardson pour les matrices SDP

Soit $A \in \mathcal{M}_n$ SDP (Symétrique Définie Positive) et $\lambda_i > 0, i = 1, \dots, n$ ses valeurs propres par ordre croissant

$$\rho(I - \alpha A) = \max(|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n|)$$

$$\alpha_{opt} = \operatorname{argmin}_{\alpha \in \mathbb{R}} \max(|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n|)$$

$$\alpha_{opt} = \frac{2}{\lambda_1 + \lambda_n}, \quad \rho(I - \alpha_{opt}A) = \frac{\lambda_n - \lambda_1}{\lambda_1 + \lambda_n} = \frac{\kappa - 1}{\kappa + 1} < 1$$

Problème: on ne connaît pas les valeurs propres de A

Voir Exercice: Méthode de Richardson à pas variable

Méthode de Richardson pour les matrices SDP

Nombre d'itérations pour une précision fixée:

$$\|x^{k+1} - x\|_2 \leq \left(\rho(I - \alpha_{opt}A)\right)^k \|x^1 - x\|_2,$$

$$\|x^{k+1} - x\|_2 \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|x^1 - x\|_2,$$

On cherche le nb d'itération k pour atteindre une précision ϵ , ie

$$\left(\frac{\kappa - 1}{\kappa + 1}\right)^k \leq \epsilon,$$

$$k \geq \frac{\ln(\frac{1}{\epsilon})}{\ln(\frac{1+\frac{1}{\kappa}}{1-\frac{1}{\kappa}})}$$

Pour κ grand:

$$k \gtrsim \frac{\kappa}{2} \ln\left(\frac{1}{\epsilon}\right)$$

Méthode de Richardson à pas variable pour les matrices SDP

Soit $A \in \mathcal{M}_n$ SDP (Symétrique Définie Positive).

On pose $e^k = x - x^k$, $r^k = Ae^k = b - Ax^k$,

et on considère l'algorithme itératif: x^1 donné et pour $k = 1, \dots$

$$\begin{cases} \alpha^k = \frac{(r^k, r^k)}{(Ar^k, r^k)}, \\ x^{k+1} = x^k + \alpha^k r^k. \end{cases}$$

On montre que

$$\alpha^k = \operatorname{Argmin}_{\alpha \in \mathbb{R}} (Ae^{k+1}, e^{k+1}) = \alpha^2 (Ar^k, r^k) - 2\alpha (r^k, r^k) + (Ae^k, e^k),$$

et

$$(Ae^{k+1}, e^{k+1}) = \left(1 - \frac{(r^k, r^k)^2}{(Ar^k, r^k)(A^{-1}r^k, r^k)}\right) (Ae^k, e^k),$$

d'où

$$(Ae^{k+1}, e^{k+1}) \leq \left(1 - \frac{1}{\operatorname{Cond}_2(A)}\right) (Ae^k, e^k)$$

Méthode de Richardson à pas variable pour les matrices SDP: algorithme

$$Ax = b \text{ avec } A \text{ matrice SDP}$$

- Choix de la précision ϵ sur le résidu relatif
- Initialisation: $x^1, r^1 = b - Ax^1, nr = nr^0 = \|r^1\|$
- Itérer tant que $\frac{nr}{nr^0} \geq \epsilon$
 - $p^k = Ar^k$
 - $\alpha^k = \frac{(r^k, r^k)}{(p^k, r^k)}$
 - $x^{k+1} = x^k + \alpha^k r^k$
 - $r^{k+1} = r^k - \alpha^k p^k$
 - $nr = \|r^{k+1}\|$

Méthode de Richardson préconditionnée

Préconditionnement: matrice $C \in \mathcal{M}_n$ inversible

$$x^{k+1} = x^k + \alpha C^{-1}(b - Ax^k)$$

$$(x^{k+1} - x) = (I - \alpha C^{-1}A)(x^k - x)$$

$$B = (I - \alpha C^{-1}A)$$

On cherche un preconditionnement C tel que

- $C \sim \alpha A$ ie $\rho(I - \alpha C^{-1}A) \ll 1$
- le système $Cy = r$ est peu coûteux à résoudre

Exemple des matrices et préconditionnements SDP

$A, C \in \mathcal{M}_n$ symétriques définies positives.

Soient $y = C^{1/2}x$, $y^k = C^{1/2}x^k$, $c = C^{-1/2}b$ on a

$$\left(C^{-1/2}AC^{-1/2}\right)y = c,$$

La matrice $C^{-1/2}AC^{-1/2}$ est SDP, et

$$y^{k+1} = y^k + \alpha \left(c - C^{-1/2}AC^{-1/2}y^k\right)$$

Convergence ssi $\rho(I - \alpha C^{-1/2}AC^{-1/2}) < 1$

$$\alpha_{opt} = \frac{2}{\lambda_{min}(C^{-1/2}AC^{-1/2}) + \lambda_{max}(C^{-1/2}AC^{-1/2})}$$

$$\rho(I - \alpha_{opt} C^{-1/2}AC^{-1/2}) = \frac{\lambda_{max}(C^{-1/2}AC^{-1/2}) - \lambda_{min}(C^{-1/2}AC^{-1/2})}{\lambda_{min}(C^{-1/2}AC^{-1/2}) + \lambda_{max}(C^{-1/2}AC^{-1/2})}$$

Exemple des matrices et préconditionnements SDP

$A, C \in \mathcal{M}_n$ symétriques définies positives.

Soient $y = C^{1/2}x$, $y^k = C^{1/2}x^k$, $c = C^{-1/2}b$ on a

$$\left(C^{-1/2}AC^{-1/2}\right)y = c,$$

La matrice $C^{-1/2}AC^{-1/2}$ est SDP, et

$$y^{k+1} = y^k + \alpha \left(c - C^{-1/2}AC^{-1/2}y^k\right)$$

Convergence ssi $\rho(I - \alpha C^{-1/2}AC^{-1/2}) < 1$

$$\alpha_{opt} = \frac{2}{\lambda_{min}(C^{-1/2}AC^{-1/2}) + \lambda_{max}(C^{-1/2}AC^{-1/2})}$$

$$\rho(I - \alpha_{opt} C^{-1/2}AC^{-1/2}) = \frac{\lambda_{max}(C^{-1/2}AC^{-1/2}) - \lambda_{min}(C^{-1/2}AC^{-1/2})}{\lambda_{min}(C^{-1/2}AC^{-1/2}) + \lambda_{max}(C^{-1/2}AC^{-1/2})}$$

Méthode de Richardson préconditionnée à pas variable pour les matrices et préconditionnements SDP: algorithme

Soient A et C SDP et le système $Ax = b$.

On applique l'algorithme de Richardson à pas variable au système

$$C^{-1/2}AC^{-1/2}y = C^{-1/2}b.$$

Il se formule comme précédemment avec la matrice $\tilde{A} = C^{-1/2}AC^{-1/2}$, le second membre $c = C^{-1/2}b$, les itérés $y^k = C^{1/2}x^k$ et les résidus $\tilde{r}^k = C^{-1/2}r^k$.

En repassant à A , x , r on obtient:

- Choix de la précision ϵ sur le résidu relatif
- Initialisation: $x^1, r^1 = b - Ax^1, nr = nr^0 = \|r^1\|$
- Itérer tant que $\frac{nr}{nr^0} \geq \epsilon$
 - $q^k = C^{-1}r^k$
 - $p^k = Aq^k$
 - $\alpha^k = \frac{(q^k, r^k)}{(p^k, q^k)}$
 - $x^{k+1} = x^k + \alpha^k q^k$
 - $r^{k+1} = r^k - \alpha^k p^k$
 - $nr = \|r^{k+1}\|$

Exemples de préconditionnements

$$A = D - E - F$$

avec D diagonale de A (supposée inversible), $D - E = \text{tril}(A)$, $D - F = \text{triu}(A)$

- Jacobi:

$$C = D$$

- Gauss Seidel

$$C = D - E \text{ ou } C = D - F$$

- SOR (Successive over relaxation) $\omega \in (0, 2)$

$$C = \frac{D}{\omega} - E$$

- SSOR (Symmetric Successive over relaxation) $\omega \in (0, 2)$

$$C = \left(\frac{D}{\omega} - E\right)D^{-1}\left(\frac{D}{\omega} - F\right)$$

Jacobi

$C = D$ et $\alpha = 1$

$$x^{k+1} = x^k + D^{-1}(b - Ax^k)$$

$$Dx^{k+1} = b - (A - D)x^k$$

Pour $i = 1, \dots, n$:

$$A_{i,i}x_i^{k+1} = b_i - \sum_{j \neq i} A_{i,j}x_j^k$$

Gauss Seidel $A = D - E - F$

$$C = D - E \text{ et } \alpha = 1$$

$$x^{k+1} = x^k + (D - E)^{-1}(b - Ax^k)$$

$$(D - E)x^{k+1} = b + Fx^k$$

Pour $i = 1, \dots, n$:

$$A_{i,i}x_i^{k+1} = b_i - \sum_{j < i} A_{i,j}x_j^{k+1} - \sum_{j > i} A_{i,j}x_j^k$$

SOR $A = D - E - F$

Pour $i = 1, \dots, n$:

$$\begin{cases} A_{i,i}\tilde{x}_i^{k+1} = b_i - \sum_{j<i} A_{i,j}x_j^{k+1} - \sum_{j>i} A_{i,j}x_j^k \\ x_i^{k+1} = \omega\tilde{x}_i^{k+1} + (1-\omega)x_i^k \end{cases}$$

Vérification de $C = \frac{D}{\omega} - E$.

$$A_{i,i}\left(\frac{x_i^{k+1}}{\omega} - \frac{(1-\omega)}{\omega}x_i^k\right) = b_i - \sum_{j<i} A_{i,j}x_j^{k+1} - \sum_{j>i} A_{i,j}x_j^k$$

$$\left(\frac{D}{\omega} - E\right)x^{k+1} = b - \left(-F - \frac{(1-\omega)}{\omega}D\right)x^k = \left(\frac{D}{\omega} - E\right)x^k + (b - Ax^k)$$

Convergence de SOR pour $\alpha = 1$

Si A est une matrice SDP, alors $\rho\left(I - \left(\frac{D}{\omega} - E\right)^{-1}A\right) < 1$ pour $\omega \in]0, 2[$ et la méthode SOR converge pour $\omega \in]0, 2[$.

On va montrer que pour tous A SDP et M inversible telle que $(M^t + M - A)$ SDP alors

$$\rho(I - M^{-1}A) < 1$$

Preuve: on considère la norme sur \mathbb{R}^n $\|x\|_A^2 = (Ax, x)$ et on va montrer que

$$\|I - M^{-1}A\|_A^2 = \sup_{x \neq 0} \frac{\|(I - M^{-1}A)x\|_A^2}{\|x\|_A^2} < 1$$

Soit $x \neq 0$, on définit $y \neq 0$ tel que $Ax = My$

$$\begin{aligned}\|(I - M^{-1}A)x\|_A^2 &= (A(x - y), (x - y)) \\ &= \|x\|_A^2 + (Ay, y) - 2(Ax, y) \\ &= \|x\|_A^2 + (Ay, y) - 2(My, y) \\ &= \|x\|_A^2 - ((M^t + M - A)y, y)\end{aligned}$$

Convergence de SOR pour $\alpha = 1$ (suite)

En utilisant $M + M^t - A$ SDP et l'équivalence des normes $\|\cdot\|_2$ et $\|\cdot\|_* = \|A^{-1}M\|_A$, on déduit qu'il existe $C > 0$ telle que

$$((M^t + M - A)y, y) \geq (1 - C)\|x\|_A^2$$

et donc $\|I - M^{-1}A\|_A^2 \leq 1 - C < 1$.

On conclut pour la convergence de SOR avec

$$M^t + M - A = \frac{D}{\omega} - E + \frac{D}{\omega} - F - D + E + F = \left(\frac{2}{\omega} - 1\right)D$$

SDP ssi $\omega \in]0, 2[$.

Rq: SOR coincide avec GS pour $\omega = 1$.