Chapter 1

Rappel d'algèbre linéaire matricielle

1.1 Introduction

L'algèbre linéaire numérique fait partie des connaissances fondamentales que doivent maîtriser les scientifiques qui font des calculs: La résolution des systèmes linéaires $Ax = b$ ou bien du problème de valeurs propres $Ax = \lambda x$ est quasiment incontournable dans la pratique de l'ingénieur .
La recherche en algèbre linéaire numérique reste très active car elle dèpend étroitement de l'évolution rapide de l'architecture des ordinateurs. De plus elle présente un intêret économique immédiat pour les industriels.
La résolution des grands systèmes (linéaires et non linéaires) est pratique et courante de nos jours, spécialement en génie mécanique, génie civil, génie électrique et de façon générale dans tous les domaines où l'on s'intèresse à la résolution numérique des équations aux dérivées partielles. La résolution numérique de ces équations conduit à l'inversion des systèmes de grande taille.
Dans ce chapitre on va rappeler les notions d'algèbre linéaire dont on aura besoin pour la résolution des systèmes linéaires.

1.2 Notations et généralités:

Dans tout ce qui suit, on se placera dans l'espace vectoriel $E = K^n$ où K représentera suivant le cas Rou C et n un entier naturel non nul.
La base canonique de E sera toujours notée $(e_1, \ldots, e_n)$; le vecteur v de coordonnées v_1,v_2,. . .,v_n,i.e., $v = \sum_{i=1}^{n} v_i e_i$ sera représenté en notation matricielle par un vecteur colonne

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Le vecteur ligne dont les composantes conjuguées des composantes v^t est noté $v^* = (\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n)$.

Une matrice A dont l'élément placé sur la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne est a_{ij} est notée $A = (a_{ij})$. Une matrice dont les seuls éléments non nuls sont les a_{ij} tel que i=j est dite diagonale et est parfois notée $diag(a_{11}, \ldots, a_{nn})$

1.2.1 Changement de base:

Soient B et B' deux bases de E. Soit u un vecteur de E dont les composantes sont représentées par le vecteur X dans la base B et par X' dans la base B'.
Si P est la matrice de passage de B' à B alors $X = PX'$.
Soit f un endomorphisme de E et, $B_1 = (v_1, \ldots, v_n)$ et $B'_1 = (w_1, \ldots, w_n)$ deux bases de E. La matrice de f de la base B_1 dans la base B'_1 est notée $\mathcal{M}_{B_1,B'_1}(f)$, c'est la matrice dont les vecteurs colonnes sont formés des coordonnées des vecteurs $(f(v_1), \ldots, f(v_n))$ dans la base B'_1.
Si B_2 et B'_2 sont deux nouvelles bases de E et si l'on note:

$$S = \mathcal{M}_{B_1 B_2} \overset{def}{=} \mathcal{M}_{B_1,B_2}(I)$$
$$T = \mathcal{M}_{B'_1 B'_2} = \mathcal{M}_{B'_1,B'_2}(I)$$

alors les matrices de l'endomorphisme f de B_1 dans B'_1 et celle de B_2 dans B'_2 vérifient la relation

$$\mathcal{M}_{B_2,B'_2}(f) = T^{-1}\mathcal{M}_{B_1,B'_1}(f)S$$

Plus généralement deux matrices A et B telles qu'il existe deux matrices inversibles T et S telles que $B = T^{-1}AS$ sont dites équivalentes.
Si de plus $S = T$ elles sont dites semblables.

1.2.2 Produit scalaire, produit hermitien:

Le produit canonique sur E est l'application (.,.) de $E \times E$ dans K définie par :
$(u, v) = v^t u = \sum_{i=1}^{n} u_i v_i$ si K=R (produit scalaire)
$(u, v) = v^* u = \sum_{i=1}^{n} u_i \overline{v}_i$ si K=C (produit hermitien).
Deux vecteurs u et v sont dits orthogonaux si $(u, v) = 0$.

1.2.3 Matrice transposée et matrice adjointe:

La transposée de la matrice A est notée A^t. $(A^t)_{ij} = a_{ji}$ et A^* désigne le conjugué de la transposée de A.
En d'autres termes la matrice adjointe de A est $(A^*)_{ij} = \overline{a}_{ji}$.
Ainsi, si A est une matrice $n \times n$, pour tout $(u, v) \in E \times E$ alors

$$(Au, v) = (u, A^t v) \quad \text{si K=}\mathbf{R}$$

$$(Au, v) = (u, A^* v) \quad \text{si K=C}$$

Si A est une matrice quelconque alors:

$$\begin{array}{ll}
(A^t)^t = A. & (A^*)^* = A. \\
(A+B)^t = A^t + B^t. & (A+B)^* = A^* + B^*. \\
(\lambda A)^t = \lambda A^t. & (\lambda A)^* = \overline{\lambda} A^* \quad \lambda \in K. \\
(AB)^t = B^t A^t. & (AB)^* = B^* A^*.
\end{array}$$

Une matrice A est dite :

- Symètrique si elle est à coefficients réels et si $A = A^t$.
- Hermitienne si $A = A^*$.
- Orthogonale si elle est à coefficients réels et si $A^t A = A A^t = I$.
- Unitaire si $A^* A = A A^* = I$.
- Normale si $A A^* = A^* A$.

1.2.4 Valeurs propres, valeurs singulières:

Les valeurs propres d'une matrice A sont des scalaires $\lambda \in K$ racine du polynôme caractéristique de A ,i.e. $det\ (A - \lambda I) = 0$.
Un vecteur propre associé à la valeur propre $\lambda \in K$ est le vecteur u non nul de E tel que :

$$Au = \lambda u$$

Le rayon spectrale $\rho(A)$ de la matrice A est le module de la valeur propre de A de plus grand module:

$$\rho(A) = \sup_{1 \leq i \leq n} |\lambda_i|$$

Les valeurs singulières de la matrice A sont les racines carrées des valeurs propres de $A^* A$ (qui sont toujours réelles et positives).

théorème Soit A une matrice d'ordre n. Les valeurs propres de A sont situées dans la réunion des disques suivants:

$$\{|z - a_{ii}| \leq r_i = \sum_{1 \leq j \leq n, j \neq i} |a_{ij}|\}$$

Preuve

1.3 Norme de vecteurs et de matrices:

1.3.1 Définition:

Une norme sur l'espace vectoriel E est une application de E dans R_+ notée $||.||$ vérifiant les trois propriétées suivantes:

- $\forall x \in E, \quad ||x|| = 0 \quad \Longleftrightarrow \quad x = 0.$
- $\forall x \in E, \quad \forall \lambda \in K \quad ||\lambda x|| = |\lambda| ||x||.$
- $\forall x, y \in E \quad ||x + y|| \leq ||x|| + ||y||.$

Les trois normes les plus couramment utilisées dans la pratique sont les suivantes:

- $||x||_1 = \sum_{i=1}^{n} |x_i|.$
- $||x||_2 = \sqrt{(\sum_{i=1}^{n} |x_i|^2)}.$ (norme euclidienne)
- $||x||_\infty = \max_{1 \leq i \leq n} (|x_i|).$

Les deux premiers exemples sont des cas particuliers de la norme $||.||_p$ définie pour tout nombre réel $p \geq 1$ par:

$$||x||_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$$

.

1.3.2 Définition:

L'ensemble $\mathcal{M}_n(K)$ des matrices carrées d'ordre n sur le corps K est un espace vectoriel de dimension n^2.
Une norme matricielle est, par définition, une application de $\mathcal{M}_n(K)$ dans R_+ satisfaisant les trois propriétés des normes vectorielles.

- $\forall A \in \mathcal{M}_n(k), \quad ||A|| = 0 \quad \Longleftrightarrow \quad A = 0.$
- $\forall A \in \mathcal{M}_n(K), \quad \forall \lambda \in K, \quad ||\lambda A|| = |\lambda| ||A||.$
- $\forall A, B \in \mathcal{M}_n(K), \quad ||A + B|| \leq ||A|| + ||B||.$

ainsi que la propriétée supplémentaire suivante:

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

La connaissance d'une norme vectorielle sur E permet de construire une norme matricielle sur $\mathcal{M}_n(K)$ de la manière suivante:

$$||A|| = \sup_{x \neq 0} \frac{||Ax||}{||x||} = \sup_{||x|| \leq 1} ||Ax|| = \sup_{||x||=1} ||Ax||$$

1.3.3 Norme matricielle et rayon spectrale:

Théorème:

Soit A une matrice carrée quelconque.

1. Si $||.||$ est une norme matricielle quelconque :
$$||A|| \geq \rho(A)$$
2. $\forall \epsilon > 0$, $\exists$ une norme (qui peut dépendre de ϵ et de A) telle que $||A|| \leq \rho(A) + \epsilon$.
3. Si A est une matrice normale, alors:

$$||A||_2 = \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} = \sqrt{\rho(AA^*)} = \rho(A)$$

Preuve
. Soit λ une valeur propre de A et v un vecteur propre associé :

$$Av = \lambda v => |\lambda|||v|| = ||Av|| \leq ||A||||v||$$

soit pour toute valeur propre λ de la matrice $A, |\lambda| \leq ||A||$ donc $\rho(A) \leq ||A||$
Remarque :
Toute matrice A peut s'ecrire sous la forme A=QTQ*, où Q est une matrice unitaire et T une matrice triangulaire supérieure dont la diagonale est formée des valeurs propres de la matrice A

$$T = \begin{pmatrix} \lambda_1 & . & . & . & . & . \\ & \lambda_2 & & & & . \\ & & & & & . \\ & & & & & . \\ & & & & & . \\ & & & & & \lambda_n \end{pmatrix}$$

Soit δ un nombre réel strictement positif, et soit $D \in IR^{n \times n}$ la matrice diagonale

$$D = \begin{pmatrix} \delta & 0 & & & 0 \\ 0 & \delta^2 & & & \\ & & . & & \\ & & & . & \\ & & & . & \\ 0 & & & 0 & \delta^n \end{pmatrix}$$

On vérifie que

$$(QD)^{-1}A(QD) = D^{-1}TD = \begin{pmatrix} \lambda_1 & \delta T_{1,2} & \delta^2 T_{1,3} & ... & ... & \delta^{n-1}T_{1,n} \\ & \lambda_2 & \delta T_{2,3} & & & \delta^{n-2}T_{2,n} \\ & & & & & . \\ & & & & & \delta^2 T_{n-2,n} \\ & & & & & \delta T_{n-1,n} \\ & & & & & \lambda_n \end{pmatrix}$$

Pour toute matrice A et tout réel ε donné, on peut trouver $\delta \in IR$ tel que

$$max_{i<j} \left| \delta^{j-i} T_{i,j} \right| < \varepsilon / n$$

Alors la norme matricielle (dépend de A et ε) définie pour toute matrice B par

$$\|B\|_{A,\varepsilon} = \left\|(QD)^{-1}B(QD)\right\|_{\infty}$$

vérifie l'inégalité

$$\|A\|_{A,\varepsilon} \leq \max_i |\lambda_i| + \varepsilon$$

Par ailleurs, cette norme matricielle subodonnée à la norme vectorielle

$$v \mapsto \left\|(QD)^{-1}v\right\|_{\infty}$$

car

$$\left\|(QD)^{-1}B(QD)\right\|_{\infty} = \max_{y\neq 0}\frac{\|(QD)^{-1}B(QD)y\|_{\infty}}{\|y\|_{\infty}} = \max_{x\neq 0}\frac{\|(QD)^{-1}Bx\|_{\infty}}{\|(QD)^{-1}x\|_{\infty}}$$

pour un réel $\varepsilon > 0$ donné, on a associé à la matrice A une norme matricielle qui vérifie

$$\|A\| - \varepsilon \leq \rho(A)$$

On a $\rho(A) \leq \|A\|_2$; de plus par définition

$$\|A\|_2 = \max_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} = \max_{x\neq 0}(\frac{(Ax, Ax)}{(x,x)})^{\frac{1}{2}} = \max_{x\neq 0}(\frac{(A^*Ax, x)}{(x,x)})^{\frac{1}{2}} = \rho^{\frac{1}{2}}(A^*A)$$

En effet A^*A est une matrice hermitienne, qui admet une base de vecteurs propres orthogonaux $\{v_1, v_2, ..., v_n\}$, en rangeant les valeurs propres associées $\lambda_i(A^*A) \in IR^+$

$$\lambda_n(A^*A) \leq \lambda_{n-1}(A^*A) \leq ... \leq \lambda_2(A^*A) \leq \lambda_1(A^*A)$$

on obtient pour tout vecteur $u \in E$

$$u = \sum_{i=1}^{n}\alpha_i v_i \quad \text{et} \quad A^*Au = \sum_{i=1}^{n}\alpha_i\lambda_i(A^*A)v_i$$

Ainsi

$$\frac{(Au, Au)}{(u,u)} = \frac{\sum_{i=1}^{n}|\alpha_i|^2|\lambda_i(A^*A)|(v_i, v_i)}{\sum_{i=1}^{n}|\alpha_i^2|(v_i, v_i)} \leq \max_i\lambda_i(A^*A) = \rho(A^*A)$$

Dans le cas d'une matrice A hermitienne ($A = A^*$), on écrit de même

$$\frac{(Au,Au)}{(u,u)} = \sum_{i=1}^{n} |\alpha_i|^2 |\lambda_i(A)|^2 (v_i,v_i) / \sum_{i=1}^{n} \left|\alpha_i^2\right| (v_i,v_i) \leq \max_i |\lambda_i(A)|^2 = \rho^2(A)$$

On vient de rappeler que pour toute matrice A, la matrice A^*A est hermitienne, et ses valeurs propres $\lambda(A^*A)$ sont réelles positives.

1.4 Condionnement d'un système linéaire:

Considérons l'exemple suivant:
soit:

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \quad \text{La solution de } Ax = b \text{ est } x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Étudions maintenant le système légèrement perturbé suivant: $Ay = b'$ avec $b' = \begin{pmatrix} 32.01 \\ 22.99 \\ 33.01 \\ 30.99 \end{pmatrix}$.

La solution de ce système est $y = \begin{pmatrix} 1.82 \\ -0.36 \\ 1.35 \\ 0.79 \end{pmatrix}$.

À une perturbation relative du second membre de 0.01, correspond une perturbation relative de l'ordre de 1 sur la solution.
De même une légère perturbation de la matrice A peut entraîner des variations considŕables de la solution $A'z = b$

où $A' = \begin{pmatrix} 10 & 7 & 8.1 & 7.2 \\ 7.08 & 5.04 & 6 & 5 \\ 8 & 5.91 & 9.89 & 9 \\ 6.99 & 4.99 & 9 & 9.98 \end{pmatrix}$

a pour solution $z = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix}$.

tout ceci peut paraître très étrange, d'autant plus que $Ax = b$ semble tout à fait inoffensif.
Ceci est très gênant lorsqu'on s'intéresse à la résolution numérique de ce système : si chaque erreur d'arrondi faite par un ordinateur lors de l'écriture d'un système linéaire est susceptible d'être amplifiée 200 fois, on imagine que le résultat fourni après même un très faible nombre d'inversions de systèmes de ce genre risque de n'avoir aucun rapport avec le résultat exacte .
Il faut donc être capable de prévoir à priori de combien une erreur sur les coefficients du système peut être amplifiée.
On introduit alors la notion de conditionnement d'une matrice inversible .

1.4.1 Définition:

Soit $||.||$ une norme matricielle subordonnée.
On appelle conditionnement de la matrice inversible A (relativement à cette norme), le réel

$$\mu(A) = ||A|| \left|\left|A^{-1}\right|\right|$$

La connaissance de $\mu(A)$ fournit une majoration des perturbations relatives des solutions de l'équation $Ax = b$ en fonction de la perturbation relative du second membre.

1.4.2 Théorème:

Soit A une matrice inversible. Soient x et $x + \Delta x$ les solutions des deux systèmes linéaires $Ax = b$ et $Ax = b + \Delta b$.
On a la majoration suivante:

$$\frac{||\Delta x||}{||x||} \leq \mu(A) \frac{||\Delta b||}{||b||}$$

Preuve:

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

$\Rightarrow A\Delta x = \Delta b \Rightarrow \Delta x = A^{-1}\Delta b$ Donc

$$||\Delta x|| = \left||A^{-1}\Delta b\right|| \leq \left||A^{-1}\right|| ||\Delta b||$$

D'autre part

$$||b|| \leq ||A|| ||x||$$

Et donc

$$\frac{1}{||x||} \leq \frac{||A||}{||b||} \Rightarrow \frac{||\Delta x||}{||x||} \leq ||A|| ||A^{-1}|| \frac{||\Delta b||}{||b||} = \mu(A)\frac{||\Delta b||}{||b||}$$

1.4.3 Théorème:

Soit A une matrice inversible.
Soient x et $x+\Delta x$ deux solutions des systèmes linéaires $Ay = b$ et $(A+\Delta A)y = b$. On a la majoration suivante:

$$\frac{||\Delta x||}{||\Delta x + x||} \leq \mu(A)\frac{||\Delta A||}{||A||}$$

Preuve:

$$\begin{array}{rcl} Ax & = & b \\ & = & (A+\Delta A)(x+\Delta x) \\ & = & Ax + A\Delta x + \Delta A(x+\Delta x) \\ & \Rightarrow & A\Delta x + \Delta A(x+\Delta x) = 0. \\ & \Rightarrow & \Delta x = -A^{-1}\Delta A(x+\Delta x). \\ & \Rightarrow & ||\Delta x|| \leq ||A^{-1}|| ||\Delta A|| ||\Delta x + x||. \\ & \Rightarrow & \frac{||\Delta x||}{||x+\Delta x||} \leq ||A^{-1}|| ||\Delta A|| = \mu(A)\frac{||\Delta A||}{||A||} \end{array}$$

1.4.4 Théorème:

Soit A une matrice inversible.

- $\mu(A) \geq 1.$ $\quad \mu(A) = \mu(A^{-1}).$
  $\mu(\alpha A) = \mu(A) \quad \forall \alpha \in R^*.$ $\quad \mu_2(A) = \frac{\lambda_+(A)}{\lambda_-(A)}.$ Où λ_+ et λ_- sont réspectivement la plus grande

  et la plus petite valeur singulière de la matrice A.

- Si A est normale:
  $\mu_2(A) = \frac{|\lambda_+(A)|}{|\lambda_-(A)|}.$

  Où $\lambda_+(A)$ et $\lambda_-(A)$ sont réspectivement la plus grande et la plus petite valeur propre de la matrice A.

- Si A est unitaire ou orthogonale:
  $\mu_2(A) = 1.$

- Si U est une matrice unitaire:
  $\mu_2(A) = \mu_2(AU) = \mu_2(UA) = \mu_2(U^*AU).$

Preuve:

- Les trois premièrs propriètées sont évidentes.

- $\mu_2(A) = ||A||_2 ||A^{-1}||_2.$
  $||A||_2{}^2 = \rho(A^*A) = \lambda_+{}^2(A).$

  $||A^{-1}||_2{}^2 = \rho(A^{-1}(A^{-1})^*) = \max_i |\lambda_i(A^{-1}(A^{-1})^*)| = \max_i \left|\lambda_i{}^2(A^{-1})\right|$

  $= \max_i \left|\frac{1}{\lambda_i{}^2(A)}\right| = \frac{1}{min_i\left|\lambda_i{}^2(A)\right|} = \frac{1}{\lambda_-{}^2(A)}.$

  donc $\mu_2(A) = \frac{\lambda_+(A)}{\lambda_-(A)}.$

- A normale $\Rightarrow A^{-1}$ normale.
  $\mu_2(A) = \rho(A)\rho(A^{-1}) = \frac{|\lambda_+(A)|}{|\lambda_-(A)|}.$

- Si A est unitaire, toutes les valeurs propres sont de module égal à 1.
  Donc $\mu_2(A) = 1.$

- U unitaire, la norme matricielle est invariante par transformation unitaire,
  donc $||A||_2 = ||AU||_2.$

Chapter 2

Résolution des systèmes linéaires

Soit le système $Ax = b$ où A est une matrice carrée $n \times n$ inversible à coefficients dans K.
Le but de ce chapitre est l'étude des méthodes pratiques de résolution d'un tel système (pour n grand).
On dispose de deux types de méthodes:

- *Les méthodes directes: x est obtenu exactement au bout d'un nombre fini d'opérations.*
- *Les méthodes itératives: x est obtenu comme la limite d'une suite infinie convergente, (i.e. $x = \lim_{m\to\infty} x^m$).*

2.1 Méthodes directes:

2.1.1 Méthode de Gauss:

Exemple:
Soit à résoudre le système:

$$(S)\left\{\begin{array}{cc} 2x + 3y + z = 5 & (1^1) \\ 5x - y + z = 9 & (2^1) \\ \frac{1}{2}x + 4y + 3z = -1 & (3^1) \end{array}\right.$$

1. *On élimine x dans les équations* (2^1) *et* (3^1) *par combinaison linéaire avec l'équation* (1^1).

$$\begin{array}{ccc} (2^1) - \frac{5}{2}(1^1) & \Longleftrightarrow & (-1 - \frac{15}{2})y + (1 - \frac{5}{2})z = 9 - \frac{25}{2} \\ & \Longleftrightarrow & -\frac{17}{2}y - \frac{3}{2}z = -\frac{7}{2}. \\ (3^1) - \frac{\frac{1}{2}}{2}(1^1) & \Longleftrightarrow & (4 - \frac{3}{4})y + (3 - \frac{1}{4}z = -1 - \frac{5}{4} \\ & \Longleftrightarrow & \frac{13}{4}y + \frac{11}{4}z = -\frac{9}{4}. \end{array}$$

Le système (S) *devient:*

$$\left\{\begin{array}{cc} 2x+3y+z=5 & (1^2) \\ -\frac{17}{2}y-\frac{3}{2}z=-\frac{7}{2} & (2^2) \\ \frac{13}{4}y+\frac{11}{4}z=-\frac{9}{4} & (3^2) \end{array}\right.$$

2. *On élimine* y *dans l'équation* (3^2).

   $$(3^2)-\frac{\frac{13}{4}}{\frac{-17}{2}}(2^2) \quad \Longleftrightarrow \quad \frac{148}{68}z=-\frac{244}{68}.$$

   Enfin le système (S) *s'écrit:*

   $$\left\{\begin{array}{c} 2x+3y+z=5 \\ 0-\frac{17}{2}y-\frac{3}{2}z=-\frac{7}{2} \\ 0+0+\frac{148}{68}z=-\frac{244}{68} \end{array}\right.$$

3. *La dérnière équation donne* z, *l'équation* (2^2) *donne* y *puis l'équation* (1^2) *donne* x.

2.1.2 Description de la méthode de Gauss:

La méthode de Gauss se fait en deux étapes:
La première (les points 1 et 2 dans l'exemple) est la triangulation du système, la deuxième (le point 3 de l'exemple) est la résolution du système triangulaire.

1. *Triangulation:*
   La méthode de Gauss ramène la résolution de $Ax=b$ *à la résolution d'un système dont la matrice est triangulaire. Pour cela, on cherche une matrice inversible* L *telle que* LA *soit triangulaire, alors:*

   $$Ax=b \quad \Longleftrightarrow \quad LAx=Lb$$

   En pratique, on détermine L *comme produit de matrices élémentaires*
   $L=L^{n-1}L^{n-2}\dots L^2L^1$. *avec:*

$$L^1 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & 0 & 0 & \dots & 1 \end{pmatrix}, \quad L^2 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & l_{n2} & 0 & \dots & 1 \end{pmatrix} \ \dots$$

$$\dots, L^{n-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & l_{nn-1} & 1 \end{pmatrix}$$

Ces matrices sont triangulaires inférieures de Det égale à 1.
Nous supposons que a_{11} de la matrice A est non nul.

La 1 ère étape de la méthode de Gauss consiste à multiplier $A = A^{(1)}$ par la matrice $L^{(1)}$.
Avec

$$l_{21} = -\frac{a_{21}}{a_{11}}, l_{31} = -\frac{a_{31}}{a_{11}}, \dots, l_{n1} = -\frac{a_{n1}}{a_{11}}$$

On obtient une matrice $A^{(2)} = L^{(1)}A^{(1)}$.

$$A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}{}^{(2)} & \dots & a_{2n}{}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}{}^{(2)} & \dots & a_{nn}{}^{(2)} \end{pmatrix}$$

Supposons que $a_{22}{}^{(2)}$ est non nul, la 2 ème étape consiste à multiplier $A^{(2)}$ par $L^{(2)}$.

$$A^{(3)} = L^{(2)}A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22}{}^{(2)} & a_{23}{}^{(2)} & \dots & a_{2n}{}^{(2)} \\ 0 & 0 & a_{33}{}^{(3)} & \dots & a_{3n}{}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}{}^{(3)} & \dots & a_{nn}{}^{(3)} \end{pmatrix}$$

Le processus se poursuit et on obtient à la fin de la $(n-1)^{\text{ème}}$ *étape une matrice* $A^{(n)}$.

$$A^{(n)} = L^{(n-1)}L^{(n-2)}\dots L^{(1)}A^{(1)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22}{}^{(2)} & a_{23}{}^{(2)} & \dots & a_{2n}{}^{(2)} \\ 0 & 0 & a_{33}{}^{(3)} & \dots & a_{3n}{}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn}{}^{(n)} \end{pmatrix}$$

et :

$$L = L^{(n-1)}L^{(n-2)}\dots L^{(1)}$$

Les matrices $L^{(n-1)}$, ..., $L^{(1)}$ *sont inversibles donc L est inversible.*

$$Ax = b \quad \Longleftrightarrow \quad LAx = Lb \quad \Longleftrightarrow \quad A^{(n)}x = b^{(n)}$$

2. *Résolution du Système Triangulaire:*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ 0 + a_{22}{}^{(2)}x_2 + \dots + a_{2n}{}^{(2)}x_n = b_2{}^{(2)} \\ \vdots \\ 0 + \dots + 0 + a_{nn}{}^{(n)}x_n = b_n{}^{(n)} \end{cases}$$

Ce qui se fait en remontant:

$x_n = \frac{b_n{}^{(n)}}{a_{nn}{}^{(n)}}$.

$x_i = \frac{1}{a_{ii}{}^{(i)}}(b_i{}^{(i)} - \sum_{j=i+1}^{n} a_{ij}{}^{(i)}x_j)$ *pour* $i = n-1, \dots, 1$.

On note $a_{kk}{}^{(k)}$ *le Pivot de Gauss.*

- **Cas d'un Pivot nul:**
Comme nous l'avons déjà remarqué, la procédure ne peut continuer que si le Pivot $a_{kk}{}^{(k)}$ *est*

non nul. Plus exactement, le processus se poursuit si parmis tous les ${a_{ik}}^{(k)}$ *pour* $i = k, \ldots, n$, *il y en a un non nul qui puisse servir de pivot après échange de l'ordre des équations restantes*

Exemple :

$$\begin{cases} x + y + z + t & = & 2 \\ x + y + 2z - 4t & = & -1 \\ x - y - \frac{1}{2}z + 2t & = & 1 \\ 2x + y - \frac{1}{2}t & = & 3 \end{cases}$$

$a_{11} = 1 \neq 0$, *la première opération donne*

$$\begin{cases} x + y + z + t & = & 2 \\ 0 + z - 5t & = & -3 \\ -2y - \frac{3}{2}z + t & = & -1 \\ -y + 2z - \frac{5}{2}t & = & -1 \end{cases}$$

$a_{22}^{(1)} = 0$ *mais* $a_{32}^{(2)} = -2 \neq 0$ *et on peut éliminer* y *de la dernière équation et permuter la 2ème et la troisième équation, ce qui donne :*

$$\begin{cases} x + y + z + t & = & 2 \\ 0 - 2y - \frac{3}{2}z + t & = & -1 \\ 0 \quad 0 \quad z - 5t & = & -3 \\ 0 \quad 0 \quad -\frac{1}{4}z - 3t & = & -\frac{1}{2} \end{cases}$$

Changement de pivot :

Si après la $(k-1)^{ème}$ *étape* $a_{kk}^{(k)} = 0$, $A^{(k)}$ *est inversible, puisque* $A^{(k)}$ *possède des* 0 *sous la diagonale dans les* $(k-1)$ *premières colonnes, il existe* $j(j > k)$ *telque* $a_{jk}^{(k)} \neq 0$ *on échange alors les lignes* j *et* k.

Autrement dit, on pose :

$A^{(k+1)} = \; L^{(k)} \sum^{(j,k)} A^{(k)}$ *où* $\sum^{(j,k)}$ *est la matrice de permutation*

$$\sum^{(j,k)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ . & . & . & 1 & . & . & . \\ & ... & 1 & 0 & 0 & 0 & 0 \\ & ... & 0 & 0 & 0 & 1 & 0 \\ 0 & & & 0 & 0 & 0 & 1 \end{pmatrix}$$

la matrice $\sum^{(j,k)}$ *est inversible de det -1*

Stratégie des pivots :

Il est préferable même lorsque le pivot n'est pas nul de choisir un pivot aussi grand que possible en valeur absolue (afin d'attenuer l'importance des erreurs d'arrondi).

Donc, à la $k^{ième}$ *étape du calcul on amène en position pivot (c'est à dire l'intersection de la* $k^{ième}$ *ligne et la* $k^{ième}$ *colonne) le terme* $a_{jk}^{(k)}$ *de module maximum, tel que :*

$$\left|a_{ik}^{(k)}\right| \leq \left|a_{ik}^{(k)}\right| pour i \geq k$$

S'il existe plusieurs indices j tels que $\left|a_{jk}^{(k)}\right|$ *soit le plus grand de* $\left|a_{ik}^{(k)}\right|$ $(i \geq k)$ *on prend le plus petit de ces indices i. on échange alors les lignes i et k.*

Traitement du second membre :

Les élements de b doivent subir les mêmes combinaisons que les lignes de la matrice.

Conclusion :

Une fois l'échange des lignes éffectué, on obtient $A^{(k+1)}$ *par les formules :*

$a_{ij}^{k+1} = a_{ij}^{k} - \frac{a_{ik}^{k}}{a_{kk}^{k}} a_{kj}^{k}$ *si* $k+1 \leq i \leq n$ *et* $k+1 \leq j \leq n$

$b_i^{(k+1)} = b_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \cdot b_k^{(k)}$

$a_{ij}^{k+1} = a_{ij}^{(k)}, b_i^{(k+1)} = b_i^{(k)} \quad si\ 1 \leq i \leq k\ \ et\ 1 \leq j \leq k$

$a_{ij}^{k+1} = 0\ \ si\ k+1 \leq i \leq n\ \ et\ 1 \leq j \leq k$

Nombre d'opétations

à la $k^{ième}$ étape, nous devons calculer $(n-k+1)(n-k)$ termes, et le calcul de chaque terme nécessite 3 opérations (1 soustracion, 1 multiplication, 1 division). Le nombre total d'opération est donc :

$$3 \sum_{k=1}^{(n-1)} (n-k+1)(n-k) = 3(\sum_{k=1}^{(n-1)} (n-k)^2 + \sum_{k=1}^{(n-1)} (n-k)$$

$$= 3(\frac{n(n-1)(2n-1)}{6} + \frac{3n(n-1)}{6})$$

$$= n(n-1)(n+\frac{1}{2})$$

Il faut compter en plus la résolution du système triangulaire, ce qui demande n^2 opérations. Le nombre total d'opérations est de l'ordre de n^3, pour n grand.

Comparaison avec la formule de Cramer :

Dans la formule de Cramer, chaque inconnue est calculée comme un quotient de 2 déterminants de matrices $n \times n$(*) un tel déterminant est une somme de $n!$ termes dont chacun est le produit n coefficients. Il y a en tout $(n+1)$ déterminant à calculer, donc $n(n+1).n!$ opérations à effectuer.

(*) $\left[\begin{array}{lcl} x_i & = & \det(A_{x_i}) / \det(A) \\ \det(A) & = & \sum_{\sigma \in S_n} \varepsilon(\sigma) \Pi_{i=1}^{n} a_{i\sigma(i)} \end{array} \right.$ *avec $\varepsilon(\sigma)$ est la signature de la permutation σ*

Si n est grand la formule de Stirling donne $n! \sim (\frac{n}{e})^n \sqrt{2\Pi n}$. Donc la résolution du système par la formule de Cramer nécessite $n^{n+2} e^{-n} \sqrt{2\Pi n}$ opérations, ce qui est très grand que celui nécessité par la méthode de Gauss.

Exemple : *Pour $n = 100$ en utilisant un ordinateur qui fait $10^8 op/s$*

	nb d'opération	*temps*
Gauss	$6, 6.105$	$10^{-2}s$
Cramer	10^{161}	$10^{153}s$

2.1.3 Décomposition de la forme : $A = L.U$

Lorsqu'on résoud le système $Ax = b$ par la méthode de Gauss sans permutation de lignes, c'est à dire lorsque tous les pivots sont non nuls on écrit :

$$L^{(n-1)}....L^{(1)}Ax = L^{(n-1)}....L^{(1)}b$$

$$A^{(n)}x = L^{(n-1)}...L^{(1)}A$$

où $A^{(n)}$ est triangulaire supérieure

on pose

$$L = L^{(n-1)}...L^{(1)}$$

on a

$$A = L^{-1}A^{(n)}$$

L étant un produit de matrices triangulaires inferieures, est aussi triangulaire inferieure. L^{-1} l'est aussi.

$A^{(n)}$ est triangulaire supérieure. Dans ce cas, on a donc

$$A = L^{-1}A^{(n)}$$

où L^{-1} est triangulaire inferieure et $A^{(n)}$ est triangulaire supérieure.

Définition :

Soit A une matrice carrée d'ordre n. on dit que A admet une décomposition du type $A = L.U$. S'il existe une matrice L triangulaire inférieure et une matrice U triangulaire supérieure telles que $A = L.U$

Définition :

Soit A une matrice d'ordre n, soit p inférieure ou égale à n, on appelle mineur fondamental d'ordre p de A, le déterminant de A_p ou $A_p = (a_{ij})_{1 \leq i.j \leq p}$

Proposition : *Soit A une matrice d'ordre n, non singulière. Alors A admet une décomposition du type $A = L.U$ si et seulement si les mineurs fondamentaux de A sont non nuls. Dans ce cas la décomposition est unique si on impose la valeur de L ou U sur la diagonale.*

Preuve :

1) Condition nécessaire

$$\left\{ \begin{array}{lcl} \det\ A & \neq & 0 \\ A & = & L.U \end{array} \right.$$

$$\det\ A \neq 0 \Rightarrow \det\ L.\det U \neq 0 \Rightarrow \left\{ \begin{array}{lcl} \det\ L & \neq & 0 \\ \det\ U & \neq & 0 \end{array} \right.$$

$a_{ij} = \sum_{k=1,\ldots,n} L_{ik} U_{kj}$

$a_{ij} = \sum_{k=1,\ldots,i} L_{ik} U_{kj}$ *puisque L est triangulaire inférieure*

Soit p: $1 \leq p \leq n$. Alors si i et j sont $\leq p$ on a

$$a_{ij} = \sum_{k=1,\ldots,p} L_{ik} U_{kj}$$

et donc

$$A_p = L_p.U_p$$

det $A_p = \det L_p.\det U_p = (\Pi_{i=1,\ldots,p} L_{ii})(\Pi_{i=1,\ldots,p} U_{ii})$

$$\left\{ \begin{array}{lcl} \det\ L \neq 0 & \Rightarrow & \det\ L_p \neq 0 \\ \det\ U \neq 0 & \Rightarrow & \det\ U_p \neq 0 \end{array} \right\} \Rightarrow \det\ A_p \neq 0$$

2) Condition suffisante :

On suppose que c'est vérifié jusqu'à l'ordre $n-1$. Soit A une matrice d'ordre n, non singulière, dont les mineurs fondamentaux sont non nuls.

$$A = \begin{pmatrix} A_{n-1} & d \\ c & a_{nn} \end{pmatrix}$$

Les mineurs feondamentaux de A sont non nuls, donc les mineurs fondamentaux de A_{n-1} sont non nuls. d'après l'hypothèse de récurrence A_{n-1} admet une décomposition du type

$$A_{n-1} = L_{n-1}.U_{n-1}$$

On cherche alors une décomposition $A = L.U$ avec

$$L = \begin{pmatrix} L_{n-1} & 0 \\ x & l_{nn} \end{pmatrix} etU = \begin{pmatrix} U_{n-1} & y \\ 0 & u_{nn} \end{pmatrix}$$

On procède par identification de bloc ce qui donne

$$\begin{cases} A_{n-1} & = & L_{n-1}U_{n-1} \\ c & = & x.U_{n-1} \\ d & = & L_{n-1}y \\ a_{nn} & = & xy + l_{nn}u_{nn} \end{cases}$$

L_{n-1} et U_{n-1}sont inversibles donc :

on obtient x et y par la deuxième et la troisième équation, puis la quatrème équation donne l_{nn} et u_{nn} non uniques.

Si on fixe la valeur de l_{nn}, on obtient u_{nn} de manière unique.

Remarque :

Lorsque A=L.U on a .

$$Ax = b \Leftrightarrow \begin{cases} Ly & = & b \\ Ux & = & y \end{cases}$$

2.1.4 Méthode de Scholeski

Cette méthode ne s'adresse qu'à des matrices symétriques définies positives.

Lemme :

Si A est symétrique, définie positive, il existe une matrice $S = (S_{ij})$ réelle triangulaire supérieure telle que : $A = S^t S$ et $S_{ii} > 0$ $(i = 1, ..., n)$

Preuve :

Raisonnons par récurrence sur l'ordre n de la matrice A.

- Pour n=1, le lemme est évident.

- Supposons la proposition vraie jusqu'à l'ordre $(n-1)$ et écrivons la matrice A (qui est supposée symétrique) sous la forme :

$$A = \begin{pmatrix} A_{n-1} & c \\ c^t & a_{nn} \end{pmatrix}$$

Puisque A est symétrique, définie positive dans IR^n, A_{n-1} est symétrique définie positive dans IR^{n-1}, par hypothèse de récurrence il existe une matrice S_{n-1} triangulaire supérieure, d'ordre $(n-1)$, telleque:

$$A_{n-1} = S_{n-1}^t . S_{n-1}$$

Posons :

$$S_n = \begin{pmatrix} S_{n-1} & z \\ 0 & S_{nn} \end{pmatrix}$$

Nous avons

$$S_n^T . S_n = \begin{pmatrix} S_{n-1}^T S_{n-1} & S_{n-1}^T z \\ z^T S_{n-1} & z^T . z + S_{nn}^2 \end{pmatrix}$$

Cherchons à identifier $S_n^T . S_n$ à la matrice A:

$$c = S_{n-1}^T . z \quad \text{et} \quad a_{nn} = z^T . z + S_{nn}^2$$

Nous devons donc prendre $z = (S_{n-1}^T)^{-1} . c$ et montrons qu'alors $a_{nn} - z^T . z > 0$

Soit S_{nn} *l'une des racines carrés de* $(a_{nn} - z^T.z)$ *(qui, à priori, peuvent être imaginaires)*

par hypothèse : $S_{nn} = \begin{pmatrix} S_{n-1} & z \\ 0 & S_{nn} \end{pmatrix}$ *et nous voulons avoir* $A = S_n^T.S_n$

alors $\det A = (\det S_n)^2 \ = (\det(S_{n-1}))^2.S_{nn}^2$

puisque A *est définie positive :* $\det A > 0$, *donc* $a_{nn} - z^T.z = S_{nn}^2 > 0$ *et* S_{nn} *réel.*

Nous choisissons alors pour $S_{nn} = \sqrt{a_{nn} - z^T.z}$

le lemme est démontré, avec $S_n = S$

Déterminons la matrice $S = \begin{pmatrix} S_{11} & . & . & S_{1n} \\ & & & \\ 0 & & & S_{nn} \end{pmatrix}$

Identifions A *et* $S^T.S$

$\forall$ *i et* $j, a_{ij} = \sum_{k=1}^{n} S_{ki}.S_{kj}$ *il suffit d'étudier les termes* a_{ij} *tels que* $i \leq j$

- *Si* $i = j = 1$ *:* $a_{11} = S_{11}^2$, $S_{11} = \sqrt{a_{11}}$
- *Si* $i = 1$, $j > 1$ *:* $a_{1j} = S_{11}S_{1j}$, *donc* $S_{1j} = \frac{a_{1j}}{S_{11}}$
- *Si* $i > 1, j = i$ *:* $a_{ii} = \sum_{k=1}^{i} S_{ki}^2$ $\ S_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} S_{ki}^2}$
- *Si* $i > 1$, $i < n$ *et* $i < j \leq n$, $S_{ij} = [a_{ij} - \sum_{k=1}^{i-1} S_{ki}S_{kj}]/S_{ii}$

Nombre d'opérations:

- *Multiplications :* $\sum_{i=2}^{n}(i-1)(n-i+1) = \frac{n(n^2-1)}{6} \sim \frac{n^3}{6}$
- *Division :* $\sum_{i=1}^{n}(n-i) = \frac{n(n-1)}{2}$
- *Addition :* $\sum_{i=1}^{n}(n-i+1) = \frac{n(n-1)}{2}$

2.1.5 Méthode de Givens

Présentation de la méthode

Considèrons une matrice $(n \times n)$ *de la forme :*

$$Q(\theta) = \begin{pmatrix} cos(\theta) & sin(\theta) & 0 \\ -sin(\theta) & cos(\theta) & 0 \\ 0 & ... & I_{n-2} \end{pmatrix}$$

$Q(\theta)$ est la matrice de rotation d'angle θ dans le plan $(x_3 = ... = x_n = 0)$

La matrice $Q(\theta)$ est orthogonale, et $detQ(\theta) = 1$

si C est matrice (n, q) $(q \in IN)$ et si $C' = Q(\theta)C$

soient L_i $(1 \leq i \leq n)$ les lignes de C, L'_i $\quad (1 \leq i \leq n)$ les lignes de C':

$$L'_1 = \cos\theta L_1 + \sin\theta L_2$$

$$L'_2 = -\sin\theta L_1 + \cos\theta L_2$$

$$L_i^{'} = L_i \qquad i \neq 1 \ , \ i \neq 2$$

considérons la matrice :

$$A = \begin{pmatrix} a_{11}...a_{nn} \\ a_{n1}...a_{nn} \end{pmatrix}$$

On cherche à mener des 0 sous la diagonale de A, et d'abord dans la première colonne, en multipliant à gauche la matrice A par des matrices convenables

- si $a_{11} = a_{21} = 0$: on passe à la troisième

- si a_{11}et a_{21} ne sont pas simultanément nuls

posons :

$$Q^{(1,1)} = \begin{pmatrix} cos(\theta) & sin(\theta) & 0 & \\ -sin(\theta) & cos(\theta) & 0 & \\ & & & \\ & 0 & & I_{n-2} \end{pmatrix}$$

avec

$$\cos\theta = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}}\ ,\ \sin\theta = \frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}}$$

alors $A^{(2,1)} = Q^{(1,1)}A$ *possède un zéro à la place de* a_{21}*:*

$$A^{(2,1)} = \begin{pmatrix} a_{11}^{(2,1)} & ...... & & & a_{1n}^{(2,1)} \\ 0 & a_{22}^{(2,1)} & & .... & a_{2n}^{(2,1)} \\ a_{31} & a_{32} & a_{33} & .... & a_{3n} \\ a_{n1} & .... & & & a_{nn} \end{pmatrix}$$

Pour faire apparaître un zéro à la place de a_{31} *(s'il n'y en a pas déjà un), posons :*

$$Q^{(2,1)} = \begin{pmatrix} \cos\theta & 0 & \sin\theta & \\ 0 & 1 & & 0 \\ -\sin\theta & 0 & \cos\theta & \\ & & 0 & I_{n-3} \end{pmatrix}$$

avec

$$\cos\theta = \frac{a_{11}^{(2,1)}}{\sqrt{(a_{11}^{(2,1)})^2 + a_{31}^2}}\ ,\ \sin\theta = \frac{a_{31}}{\sqrt{(a_{11}^{(2,1)})^2 + a_{31}^2}}$$

et

$$A^{(3,1)} = Q^{(2,1)}A^{(2,1)}$$

Le processus se poursuit jusqu'à la dernière ligne.

Nous obtenons :

$$A^{(n,1)} = Q^{(n-1,1)}...Q^{(2,1)}Q^{(1,1)}A,$$

où $A^{(n,1)}$ *possède des zéros sous sa diagonale en première colonne, et où les matrices* $Q^{(k,1)}$ *sont orthogonales de determinant* 1.

On passe ensuite à la deuxième colonne où l'on fait apparaître des zéros sous la diagonale en multipliant $A^{(n,1)}$ à gauche par des matrices de rotation de la forme :

$$\begin{pmatrix} 1 & 0 & 0 & . & . & . & 0 \\ 0 & cos(\theta) & 0 & sin(\theta) & 0 & 0 & \\ . & 0 & 1 & 0 & . & . & 0 \\ . & . & 0 & 1 & 0 & . & 0 \\ 0 & -sin(\theta) & 0 & cos(\theta) & 0 & 0 & \\ . & . & . & . & 0 & 1 & 0 \\ . & . & . & . & . & 0 & 1 \end{pmatrix}$$

Le processus se poursuit, colonne par colonne, jusqu'à obtention d'une matrice $A^{(n,n)}$ triangulaire supérieure.

$$A^{(n,n)} = Q.A$$

où Q est un produit de matrices de rotation :

$$Q = \Pi_{j=1}^{n-1} \Pi_{i=j+1}^{n} Q_{(\theta)}^{(i,j)}$$

La matrice Q est orthogonale, $detQ = 1$

Exercice

Ecrire l'algorithme de cette méthode puis calculer le nombre d'opération nécessaire pour résoudre un système linéaire de dimension n.

2.1.6 Méthode de Haushölder

Il s'agit encore de faire apparaître des zéros sous la diagonale, d'abord en première colonne, puis en seconde colonne, etc,..., jusqu'à la $(n-1)^{ième}$ colonne.

Dans une première étape on cherche une matrice $B^{(1)}$ telle que :

$$B^{(1)} A_{.1}^{(1)}$$

soit parallèle à

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$$

On note $A_{.1}^{(1)}$ *la première colonne de* $A^{(1)} = A$.

Lemme 1 :

Soit $u \in IR^n$, $\|u\| = \sqrt{\sum u_i^2} = 1$ *la matrice* $H = I - 2uu^T$ *est orthogonale et elle représente la symétrie par rapport à l'yperplan* P*, orthogonal à* u.

Preuve

On a H *est symétrique :* $H^T = (I - 2uu^T)^T = H$

$H^T.H = (I - 2uu^T)(I - 2uu^T)$

$= I - 2uu^T - 2uu^T + 4uu^T.u.u^T$

$= I$

Donc H *est orthogonale*

$$\forall\ x \in IR^n, Hx = x - 2u.(u^T.x)$$

Et $u^T x$ *est en fait le produit scalaire* $< u, x >$: *la composante de* x *suivant* u *est changée de signe (multipliée par -1). La composante suivant* P *est inchangée.*

Lemme 2

Soit $a \in IR^n$;*non parallèle à* e_1*il existe* $u \in IR^n$,*il existe* $\alpha \in IR$*, tels que*

$Ha = \alpha e_1$ *avec* $H = I - 2uu^T$

Preuve

On a $\|Ha\| = \|\alpha e_1\| = |\alpha|$

H *est orthogonale* $\Rightarrow \|Ha\| = \|a\|$

On prend donc $\alpha = \pm \|a\|$

Posons $\mu = u^T.a$

$a - 2uu^T a = \alpha e_1 \Rightarrow a^t.a - 2a^T uu^T a = \alpha a^T e_1$

$\Rightarrow \|a\|^2 - 2\mu^2 = \pm \|a\| a_1 \quad (aveca_1 = a^T e_1)$

$\Rightarrow \mu^2 = \frac{\|a\|^2 \pm \|a\| a_1}{2} = \frac{\|a\|(\|a\| \pm a_1)}{2}$

Vérifions que $\frac{\|a\|(\|a\| \pm a_1)}{2} \geq 0$

a n'est pas parallèle à e_1 : $\exists\, i \neq 1$ *tel que* $a^T e_i \neq 0$

$\pm a_1 \leq |a_1| < \|a\|$

Donc

$$\begin{cases} \mu = & \pm\sqrt{\frac{\|a\|(\|a\| \pm a_1}{2}} \\ et & \\ u \neq & 0 \end{cases}$$

$Ha = \alpha e_1 \Leftrightarrow a - 2u\mu = \alpha e_1$

Donc $u = \frac{a - \alpha e_1}{2\mu}$

La première étape du calcul consiste à former

$$A^2 = H^{(1)} A^1$$

où $H^{(1)}$ *est la matrice de symétrie donnée par le Lemme 2 avec* $a = A^{(1)}_{.1}$, *la première colonne de* A *.*

$$A^{(2)} = H^{(1)} A^{(1)} = \begin{pmatrix} \pm \left\| A^{(1)}_{.1} \right\| & A^{(2)}_{12} \\ & \\ 0 & A^{(2)}_{n-1,n-1} \end{pmatrix} = (I - 2uu^T) A^{(1)}$$

On applique la même technique à la matrice $A^{(2)}_{n-1,n-1}$ *pour faire apparaître des* 0 *sous la diagonale dans la deuxième colonne. Le processus se poursuit jusqu'à triangulation de la matrice* A.

Exercice

Ecrire un programme pour résoudre un système linéaire $Ax = b$ *par la méthode de Houshölder.*

2.1.7 Méthodes Itératives :

Introduction :

Résoudre le système linéaire $Ax = b$ *par une méthode itérative c'est construire une suite de vecteurs* $x^{(m)}$ *telle que* $\lim_{m \to +\infty} x^{(m)} = x$ *et* $Ax = b$.

Pratiquement on arrêtera évidement les calculs dès que l'on aura atteint la précision désirée (ou imposée par la machine)

Toutes les méthodes itératives classiques sont du type :

$$x^{(m+1)} = Bx^{(m)} + c \qquad (1)$$

où B *est une matrice carrée (n,n) et* $c \in IR^n$ *ou* $\mathcal{C}\backslash$.

Ces méthodes se construisent toutes de la manière suivante :

on pose $A = M - N$

où M *est inversible et, en pratique, facile à inverser (en particulier matrices diagonales ou triangulaires).*

On détermine la suite $x^{(m)}$ *par la relation de récurrence :*

$$Mx^{(m+1)} = Nx^{(m)} + b$$

c'est-à-dire

$$x^{(m+1)} = M^{-1}Nx^{(m)} + M^{-1}b$$

donc $x^{(m+1)} = Bx^{(m)} + c$ *avec* $\begin{cases} B = M^{-1}N \\ c = M^{-1}b \end{cases}$

Convergence des méthodes itératives :

Cherchons à quelles conditions sur B *et* c *la suite* $x^{(m)}$ *définit par la relation (1) soit convergente.*

On appelle x *solution exacte :* $x = A^{-1}b$

$$x = Bx + (I - B)x$$

d'ou

$$x = Bx + (I - B)A^{-1}b$$

$$x^{(m+1)} = Bx^{(m)} + c$$

Retranchons membre à membre ces deux égalités :

$$x^{(m+1)} - x = B(x^{(m)} - x) + [c - (I - B)A^{-1}b]$$

Définition :

On dit que la méthode (1) est consistante si, lorsque la limite de la suite $x^{(m)}$ existe, cette limite est la solution x.

Supposons que $x^{(m)}$ converge , soit $y = \lim_{m+\infty} x^{(m)}$ alors $y - x = B(y - x) + [c - (I - B)A^{-1}b]$ pour que la méthode (1) soit consistante, il faut et il suffit que :

$$(I - B)(y - x) = c - (I - B)A^{-1}b \Rightarrow y - x = 0$$

donc (1) consistante si et seulement si $\begin{cases} c = (I - B)A^{-1}b \\ I - B \; estinversible \end{cases}$

Définition :

Une méthode itérative est dite convergente si quelque soit le choix de vecteur initial $x^{(0)}$,

$$\lim_{m\to\infty} x^{(m)} = A^{-1}b = x$$

Lemme :

On suppose que la méthode (1), associée à B *et* c *est consistante. Alors cette méthode est convergente si, et seulement si :* $\lim_{m\to\infty} B^{(m)} = 0$

Preuve :

Puisque la méthode est consistante : $c = (I - B)A^{-1}b$ *et* $\forall m, x^{(m+1)} - x = B(x^{(m)} - x)$*, par récurrence, on obtient :*

$$x^{(m)} - x = B^{(m)}(x^{(0)} - x)$$

la méthode est convergente si et seulement si :

$$B^{(m)}(x^{(0)} - x) \to 0, \forall x^{(0)}$$

c'est-à-dire si et seulement si : $B^m z \to 0$ *quand* $m \to +\infty$ $\forall z$

ceci est équivalent à $B^n \to 0$ *quand* $m \to +\infty$

En effet, si $B^m \to 0, \forall z \, \|B^m z\| \leq \|B^m\| \,.\, \|z\| \Rightarrow B^m z \to 0$

Réciproquement : si $\forall z$*,* $B^m z \to_{m\to\infty} 0$ *, alors prenons successivement par* z *les valeurs*

de la base canonique $e_j (j = 1, ..., n)$

$B^m e_j \to 0$ *et* $B^m e_j$ *est le* $j^{\text{ième}}$ *vecteur colonne de* B^m *, donc* $B^m \to 0$

Proposition

Soit B *une matrice* (n, n)*; alors :*$\lim_{m\to\infty} B^{(m)} = 0 \Leftrightarrow \rho(B) < 1$

a) Condition suffisante :

si $\rho(B) < 1 :$ $\exists \eta > 0$ *telque* $\rho(B) + \eta < 1$ *alors d'après le théorème cité, il existe*

une norme $\|.\|$ *telle que :* $\|B\| \leq \rho(B) + \eta < 1$ *d'ou* $\|B\|^m \to_{m\to+\infty} 0$

comme $0 \leq \|B^m\| \leq \|B\|^m$*, on a* $\|B^m\| \to 0$

b) Condition nécessaire :

Supposons que $\rho(B) \geq 1$*. Alors* B *possède au moins une valeur propre* λ*,*

de module :$|\lambda| \geq 1$*, soit* x *un vecteur propre de* B *correspondant à* λ*:*

$Bx = \lambda x$*, d'où* $B^m x = \lambda^m x \; \forall m$

le vecteur x *est non nul et* $|\lambda^m|$ *tend vers 1 ou* $+\infty$ *quand* $m \to +\infty$*, donc* $B^m x$ *ne tend pas vers* 0 *et* B^m *ne tend pas vers* 0*.*

2.1.8 Principales méthodes itératives :

Nous allons voir trois méthodes applicables dans le cas où tous les élements de la diagonale de A sont non nuls.

a) Méthode de Jacobi

On pose pour cette méthode :

$M = D$, $N = E + F$ *avec* $D = (dij)$ *est diagonale*

$$\begin{cases} d_{ii} = a_{ii} & si \quad i = j \\ d_{ij} = 0 & si \quad i \neq j \end{cases}$$

$E = (e_{ij})$ *est strictement triangulaire inférieure*

$$\begin{cases} e_{ij} = 0 & si \quad i \leq j \\ e_{ij} = -a_{ij} & si \quad i > j \end{cases}$$

$F = (f_{ij})$ *est strictement triangulaire supérieure*

$$\begin{cases} f_{ij} = 0 & si \quad i \geq j \\ f_{ij} = -a_{ij} & si \quad i < j \end{cases}$$

on a : $Dx^{(m+1)} = (E + F)x^{(m)} + b$

D *est inversible puisque tous les* a_{ii} *sont non nuls.*

Soit

$$x_i^{m+1} = -\sum\nolimits_{j \neq i} \frac{a_{ij}}{a_{ii}} x_j^{(m)} + \frac{b_i}{a_{ii}} \ \ \forall i \ 1 \leq i \leq n$$

Par la méthode de Jacobi, on obtient les composantes de $x^{(m+1)}$ *à partir de celles de* $x^{(m)}$.

On appelle $B = M^{-1}N$ *la matrice de l'itération :*

Soit $B = D^{-1}(E + F)$

b) Méthode de Gauss-Seidel

On pose ici $M = D - E$, $N = F$ et on écrit la formule de récurrence :

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

Soit

$$\textstyle\sum_{j=1}^{i} a_{ij}x_j^{m+1} = -\sum_{j=i+1}^{n} a_{ij}x_j^{m} + b_j \ \forall\ i\ 1 \leq i \leq n$$

$$a_{ii}x_i^{m+1} = -\sum_{j=1}^{i-1} a_{ij}x_j^{m+1} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + b_j$$

$$x_i^{(m+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}}x_j^{(m+1)} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}}x_j^{(m)} + \frac{b_i}{a_{ii}} \ \forall\ i\ 1 \leq i \leq n$$

Remarque:

Dans la méthode de Jacobi, la $i^{ème}$ composante de $x^{(m+1)}$ est calculée en fonction des (n-1) composantes de $x^{(m)}$ qui doivent donc être conservées en mémoire jusqu'au calcul de $x_n^{(m+1)}$. Par contre la méthode de Gauss -Seidel donne la $i^{ème}$ composante de $x^{(m+1)}$ en fonction des $(i-1)$ premières composantes de $x^{(m+1)}$ et des (n-i-1) dérnières composantes de $x^{(m)}$. Elle ne necessite pas la conservation en mémoire des autres coposantes de $x^{(m)}$, ce qui est un avantage importantsur la méthode de Jacobi.

Pour la méthode de Gauss-Seidel, on a la matrice de l'itération

$$B = M^{-1}N = (D - E)^{-1}F$$

c) Méthode de Relaxation

Pour cette méthode on pose :

$$M = \frac{1}{w}D - E\ ,\ N = (\frac{1}{w} - 1)D + F$$

où w est un nombre réel non nul, (on verra en suite, d'autres limitations pour w). Il faut en fait que $w \in]0, 2[$.

On a la formule de récurrence :

$$(\frac{1}{w}D - E)x^{(m+1)} = ((\frac{1}{w} - 1)D + F)x^{(m)} + b$$

Soit :

$$(D - wE)x^{(m+1)} = ((1-w)D + wF)x^{(m)} + wb$$

$$x_i^{(m+1)} = (1-w)x_i^m - w\sum_{j=i+1}^{n}\frac{a_{ij}}{a_{ii}}x_j^{(m)} - w\sum_{j=1}^{i-1}\frac{a_{ij}}{a_{ii}}x_j^{(m+1)} + w\frac{b_i}{a_{ii}} \quad 1 \le i \le n$$

Remarque

1/ si $w = 1$ on retrouve la méthode de Gauss-Seidel

2) on dit qu'il y a sous ou sur-relaxation suivant que $w < 1$ ou $w > 1$

Convergence des méthodes de Jacobi et Gauss-Seidel

Proposition :

Soit A une matrice à diagonale strictement dominante, alors les méthodes de Gauss-Seidel et de Jacobi sont convergentes.

Preuve :

a) Montrons que la méthode de Jacobi est convergente.

La matrice de l'itération B est égale à $D^{-1}(E+F)$

Soit $B = B_{ij} \begin{cases} B_{ij} = 0 & si \quad i = j \\ B_{ij} = -\frac{a_{ij}}{a_{ii}} & si \quad i \neq j \end{cases}$

A étant à diagonale strictement dominante on a :

$\sum_{i\neq j}|B_{ij}| = \sum_{i\neq j}\left|\frac{a_{ij}}{a_{ii}}\right| < 1 \quad \forall\, i$

et les valeurs propres de B sont donc dans le disque de centre 0 et de rayon strictement inférieur à 1.

Donc toutes les valeurs propres de B ont un module strictement inférieur à 1 et $\rho(B) < 1$

Donc la méthode de Jacobi converge.

b) Montrons que la méthode de Gauss-Seidel est convergente.

Il suffit de verifier que si λ valeur propre de la matrice d'itération $B = (D-E)^{-1}F$, $|\lambda| < 1$

Soit $y \neq 0$, $y \in IR^n$ un vecteur propre associé à λ alors :

$$(D-E)^{-1}Fy = \lambda y$$

ou encore

$$Fy = \lambda(D-E)y$$

$$-\lambda\sum_{j=1}^{i-1}\frac{a_{ij}}{a_{ii}}y_j - \sum_{j=i+1}^{n}\frac{a_{ij}}{a_{ii}}y_j = \lambda y_i \quad i = 1,...,n$$

y étant non nul , $\exists\ i_0$ tel que: $|y_{i_0}| \geq |y_j| \quad j = 1,...,n$ et $y_{i_0} \neq 0$

$$\lambda y_{i_0} = -\lambda\sum_{j=1}^{i_0-1}\frac{a_{i_0j}}{a_{i_0i_0}}y_j - \sum_{j=i_0+1}^{n}\frac{a_{i_0j}}{a_{i_0i_0}}y_j$$

alors

$$|\lambda| \leq |\lambda|\sum_{j<i_0}\frac{|a_{i_0j}|}{|a_{i_0i_0}|} + \sum_{j>i_0}\frac{|a_{i_0j}|}{|a_{i_0i_0}|}$$

d'ou

$$|\lambda|\,(1-\sum_{j<i_0}\frac{|a_{i_0j}|}{|a_{i_0i_0}|}) \leq \sum_{j>i_0}\left|\frac{a_{i_0j}}{a_{i_0i_0}}\right| < (1-\sum_{j<i_0}\left|\frac{a_{i_0j}}{a_{i_0i_0}}\right|)$$

d'après la diagonale dominance.

Donc $|\lambda| < 1$

Convergence de la méthode de relaxation :

La matrice d'itération de la méthode de relaxation est $B = (D - wE)^{-1}((1-w)D + wF)$

Théorème

Si la méthode de relaxation de paramètre w *converge alors* $w \in \,]0, 2[$

Preuve :

$$detB = det((D - wE)^{-1}((1-w)D + wF))$$

$$\det B = \frac{\det((1-w)D + wF)}{\det(D - wE)} = (1-w)^n$$

et on a

$$|\det B| = |\Pi_{i=1}^{n} \lambda_i|$$

où λ_i *est une valeur propre de* B

si la méthode converge alors $\rho(B) < 1$

$$|1 - w|^n \leq (\rho(B))^n \Rightarrow |1 - w| \leq \rho(B)$$

et donc

$$|1 - w| < 1 \Rightarrow w \in \,]0, 2[$$

Théorème :

Soit A *une matrice carrée d'ordre n, hérmitienne non singulière. Si la matrice* D *est définie positive, la méthode de relaxation pour* A *est convergente ssi* $w \in \,]0, 2[$ *et si* A *est définie positive.*

Preuve :

a) Préliminaires :

Soit $x \in IR^n$ $y = Bx$, *et* $z = x - y$
on a

$$y = (D - wE)^{-1}((1-w)D + wE^*)x$$

$$(D - wE)y = ((1-w)D + wE^*)x$$

Soit

$$(D - wE)z = wDx - wE^*x - wE^*x = wAx$$

et

$$Dy - wEy + wDy - wE^*y = (1-w)Dz + Dy + wE^*z$$

$$d'ou \left\{ \begin{array}{l} (D - wE)z = wAx \\ (1-w)Dz + wE^*z = wAy \end{array} \right.$$

$\Rightarrow w < x, Ax > -w < y.Ay >=< x, (D - wE)z > - < y, (1-w)Dz + wE^*z >$
$\Rightarrow w < x, Ax > -w < y.Ay >=< x, (D - wE)z > - < y, (1-w)Dz > - < y, wE^*z >$
$= < z, Dz > -w < Ez, x > +w < Dz, y > -w < y, E^*z >$
D étant réelle positive
$\Rightarrow w < x, Ax) - w < y, Ay >=< Dz, z > -w < z, E^*x > +w < Dy, z) - w(z, Ey)$
$= < Dz, z) - w < z, E^*x + Ey > +w < Dy, z >$
or on a :

$$(1-w)Dx + wE^*x + wEy = Dy$$

donc

$$w < x, Ax > -w < y - Ay >=< Dz, z > - < z, Dy - (1-w)Dx) + w < z, Dy >$$

$$= (2-w) \ < z, Dz >$$

on a donc le résultat suivant

$$* \left\{ \begin{array}{lcl} (2-w) < Dz, z > & = & w < x, Ax > -w < y, Ay > \\ où\ x \in IR^n, y = Bx & et & z = x - y \end{array} \right.$$

b) supposons que A est définie positive et $w \in]0,2[$

Montrons que la méthode de relaxation est convergente,

soit λ valeur propre de B, soit x un vecteur propre de B associé à λ

$$y = Bx = \lambda x$$

(*) donne

$$(2-w)\left|1-\lambda\right|^2 < x, Dx >= w(1-\left|\lambda\right|^2) < x, Ax >$$

$\lambda \neq 1$ car si $\lambda = 1$ la relation $(D - wE)z = wAx$

devient $wAx = 0$ $w \neq 0 \Rightarrow Ax = 0 \Rightarrow A$ singulière

on a

$$(2-w)\left|1-\lambda\right|^2 < x, Dx > > 0$$

A étant définie positive, donc

$$1 - \left|\lambda\right|^2 > 0 \Rightarrow \left|\lambda\right| < 1$$

donc $\rho(B) < 1 \Rightarrow$ la méthode converge

Réciproquement, supposons que la méthode de la relaxation converge et que A n'est pas définie positive on prend le vecteur initial de l'itération tel que l'erreur initiale e^0 vérifie

$$< Ae^0, e^0) < 0$$

(*) donne

$$(2-w) < D(e^0 - e^1), e^0 - e^1 >= w < e^0, Ae^0 > -w < e^1, Ae^1 >$$

$e^1 = Be^0$ et 1 n'est pas valeur propre de B

donc $e^1 = Be^0$ et 1 n'est pas valeur propre de B

donc $e^1 \neq e^0$

on a

$$(2-w) < D(e^0 - e^1), e^0 - e^1 >> 0$$

donc

$$< e^0, Ae^0 > > < e^1, Ae^1 >$$

on vérifie de même que :

$$< e^m, Ae^m > > < e^{m+1}, Ae^{m+1} > \quad \forall\ m$$

ce qui est contradictoire du fait que $e^{(m)} \to_{m\to+\infty} 0$

et donc

$$< \ e^m, Ae^m > \to_{m\to+\infty} 0$$

donc A est définie positive

Remarque :

Si la matrice A est décomposée en blocs A_{ij} *de taille* $n_i \times n_j$ *(les blocs diagonaux* A_i *sont donc carrés) on peut décomposer A sous la même forme* $A = D - E - F$*, mais où cette fois D est la matrice diagonale par bloc dont les blocs diagonaux sont les blocs* A_{ii}*,* $-E$ *est la matrice composée des blocs situés sous les blocs diagonaux de A et* $-F$ *des blocs situés au dessus des blocs diagonaux de A. On définit alors comme précedement les méthodes.*

2.1.9 Méthode de Projection

Introduction

On cherche à résoudre $Ax = b$ *où* A *régilière*

Soit $z \in IR^n$.

Posons $\Pi_z = \{x \in IR^n / < z, Ax - b >= 0\}$

Propriètè :

Le plan Π_z *contient* $x^* = A^{-1}b$ *et il est normale à* $A^T z$

Preuve :

$$< A^T z, x_1 - x_2 >=< z, Ax_1 - Ax_2 >$$

$$=< z, Ax_1 - b > - < z, Ax_2 - b >$$

$$= 0 \Rightarrow x_1 - x_2 \perp A^T z$$

Principe de la méthode:

Soit $x^{(p)}$ *une approximation de* x^* *comme* $x^* \in \Pi_{z^{(p)}}$ *il est naturel de chercher* $x^{(p+1)}$ *dans* $\Pi_{z^{(p)}}$. *On prendra* $x^{(p+1)}$ *comme étant la projection orthogonal de* $x^{(p)}$ *sur* $\Pi_{z^{(p)}}$ *parallelement à* $A^t {z^{(p)}}^T$.

Soit donc $x^{(p+1)} = x^{(p)} + \lambda_p A^t z^{(p)}$

Calcul de λ_p:

$$x^{(p+1)} \in \Pi_{z^{(p)}} \Leftrightarrow < z^p, Ax^{p+1} - b >= 0$$

$$\Leftrightarrow < z^p, Ax^p - b + \lambda_p A A^t z^p >= 0$$

$$\Leftrightarrow < z^p, r^p + \lambda_p A A^t z^p >= 0$$

$$\Leftrightarrow \lambda_p = \frac{-<z^p.r^p>}{<A^tz^p, A^tz^p>} = -\frac{<z^p, r^p>}{\|A^Tz^p\|^2}$$

d'ou la méthode

$$\begin{cases} x^{(p)} & q.c.q \\ x^{p+1} & = \quad x^p - \frac{<z^p,r^p>}{\|A^tz^p\|} \; A^tz^p \end{cases}$$

Propriètés:

1) $\|x^{p+1} - x^*\|^2 = \left\|x^p - \frac{<z^p,r^p>}{\|A^tz^p\|}A^tz^p - x^*\right\|^2$

$= \|x^p - x^*\|^2 - 2\frac{<z^p,r^p>}{\|A^tz^p\|^2} < A^tz^p, x^p - x^* > + \frac{<z^p,r^p>^2}{\|A^tz^p\|^4} < A^tz^p, A^tz^p >$

$= \|x^p - x^*\|^2 - \frac{<z^p,r^p>^2}{\|A^tz^p\|^2}$

2) $\begin{cases} x^{p+1} \ etx^* \in \Pi_{z^p} \\ etA^tz^p \perp \Pi_{z^p} \end{cases} \Rightarrow < x^{p+1} - x^*, A^tz^p >= 0$

Remarque :

La propriété 1) montre que $\|x^p - x^*\|$ *est décroissante comme elle est minorée par 0 elle est convergente. Il existe* $\lambda > 0$ *tq* $\|x^p - x^*\| \rightarrow_{p\rightarrow+\infty} \lambda$

Théorème :

si $\alpha_p = \cos(r^p, z^p) =< \frac{r^p}{\|r^p\|}, \frac{z^p}{\|z^p\|} >$ *ne tend pas vers 0 alors* $\lim_{p\rightarrow+\infty} x^{(p)} = x^*$ *de plus on a :*

$\|e^{p+1}\|^2 \leq \|e^p\|^2 \left[1 - \frac{\alpha_p^2}{\|A^T\|^2\|A^{-1}\|^2}\right]$

Preuve :

$$e^{(p)} = x^{(p)} - x^*$$

$$\left\|e^{p+1}\right\|^2 = \|e^p\|^2 \left[1 - \frac{< z^p, r^p >^2}{\|e^p\|^2 . \|A^t z^p\|^2}\right]$$

$$\frac{< z^p, r^p >^2}{\|e^p\|^2 . \|A^t z^p\|^2} = \frac{< z^p, r^p >^2}{\|z^p\|^2 \|r^p\|^2} . \frac{\|z^p\|^2}{\|A^t z^p\|^2} . \frac{\|r^p\|^2}{\|e^p\|^2}$$

$$= \alpha_p^2 . \frac{\|z^p\|^2}{\|A^t z^p\|^2} . \frac{\|r^p\|^2}{\|e^p\|^2}$$

$$r^p = Ax^p - b = A(x^p - x^*) = Ae^p$$

$$\Rightarrow \frac{< z^p, r^p >^2}{\|e^p\|^2 . \|A^t z^p\|^2} = \alpha_p^2 . \frac{\|z^p\|^2}{\|A^t z^p\|^2} . \frac{\|r^p\|^2}{\|A^{-1} r^p\|^2}$$

$$\geq \alpha_p^2 \frac{1}{\|A^T\|^2} . \frac{1}{\|A^{-1}\|^2}$$

$$\Rightarrow \left\|e^{p+1}\right\| \leq \|e^p\|^2 \left[1 - \frac{\alpha_p^2}{\|A^t z^p\|^2 . \|A^{-1}\|^2}\right]$$

$Posons$

$$\beta_p = \frac{\alpha_p^2}{\|A^t z^p\|^2 . \|A^{-1}\|^2}$$

$$\left\|e^{p+1}\right\|^2 \leq \|e^p\|^2 (1 - \beta_p)$$

$$\leq \left\|e^{p-1}\right\|^2 (1 - \beta_{p-1})(1 - \beta_p)$$

$$\leq \left\|e^0\right\|^2 (1 - \beta_0)...(1 - \beta_p)$$

$$\leq \left\|e^0\right\|^2 \exp(-\sum_{i=0}^{p} \beta_i)$$

$$\beta_i \geq 0 \ , \ \beta_i \rightarrow_{i\rightarrow\infty} 0 \Rightarrow \sum_{i=0}^{p} \beta_i \rightarrow_{p\rightarrow+\infty} +\infty$$

$$\Rightarrow \|e^p\| \rightarrow_{p\rightarrow+\infty} 0$$

Donc pour avoir la convergence, il suffit que

$$\frac{< z^p, r^p >^2}{\|e^p\|^2 . \|A^t z^p\|^2} \rightarrow_{p\rightarrow+\infty} 0$$

on prend

$$z^p = r^p \Rightarrow \alpha_p = 1$$

$$\begin{cases} x^{(0)} & qlq \\ x^{p+1} & = \quad x^p - \frac{\|r^p\|^2}{\|A^T r^p\|^2} A^T r^p \end{cases}$$

Remarque :

Si A est symetrique

$$\left\|e^{p+1}\right\|^2 \leq \|e^p\|^2 \left[1 - \frac{\alpha_p}{cond(A^2)}\right]$$

Donc plus le cond(A) est grand plus la méthode est lente.

Généralisation :

On a

$$\forall \ z \in IR^n, \ x^* = A^{-1}b \in \Pi_z \Rightarrow x^* \in \sqcap_{i=1}^{k} \Pi_z$$

au lieu de considérer une seule direction $z^{(p)}$ on va considérer k directions lineairement indépendantes.

Soit

$$v^p = (v_1^{(p)}, ..., v_k^{(p)}) \;\; (v_i^{(p)})_{i=1,\cdots,k}$$

e.i

$$\Pi_{v_i^{(p)}} = \left\{ x \in IR^n / < v_i^{(p)},\, Ax - b >= 0 \right\}$$

on a

$$x^* \in \sqcap_{i=1}^{k} \Pi_{v_i^{(p)}}$$

Cherchons

$$x^{(p+1)} \in \sqcap_{i=1}^{k} \Pi_{v_i^{(p)}}$$

$$\sqcap_{i=1}^{k} \Pi_{v_i^{(p)}} = \Pi_{v^{(p)}} = \{ x \in IR^n / v^{(p)T}(Ax - b) = 0)$$

on a

$$A^t v_i^{(p)} \perp \Pi_{v_i^{(p)}} \Rightarrow A^t v_i^{(p)} \perp \sqcap_{i=1}^{k} \Pi_{v_i^{(p)}} = \Pi_{v^{(p)}}$$

$$\Rightarrow \;\; A^T V^{(p)} \perp \Pi_{v^{(p)}}$$

On peut chercher $x^{p+1} \in \Pi_{v^{(p)}}$ *tq* $x^{p+1} - x^p$ *soit porté par* $A^T V^{(p)}$
Soit donc

$$x^{p+1} = x^p + A^T V^{(p)}.\lambda \; avec \lambda \in IR^k$$

$$x^{p+1} \in \Pi_{v^{(p)}} \Leftrightarrow V^{p^T}(Ax^{p+1} - b) = 0$$

$$\Leftrightarrow V^{p^T}(Ax^p - b + AA^T V^{(p)}.\lambda) = 0$$

$$\Leftrightarrow V^{p^T}.r^p + V^{p^T} AA^T V^{(p)}.\lambda = 0$$

$$\Leftrightarrow \lambda = -(V^T AA^T V^p)^{-1}.V^{(p)^T} r^p$$

d'ou la méthode :

$$x^{p+1} = x^p - A^T V^{(p)} (V^{p^T} A.A^T V^p)^{-1}.V^{(p)} r^{(p)}$$

On pose

$$V^{p^T} A \Rightarrow x^{p+1} = x^p - w^T (w.w^T)^{-1} V^{\Pi}.r^p$$

Cas des Systèmes Symétriques

Théorème :

Si A est symétrique définie positive, le point $x^* = A^{-1}b$ est centre de symétrie des éllipsoide Q_k d'équation $\frac{1}{2}x^t Ax - x^t b = k$

Preuve :

$x = x_0 + tu$ *rencontre* Q_k *ssi* $\frac{1}{2}x^t Ax - x^t b = k$

$$\Leftrightarrow \frac{1}{2}(x_0 + tu)^t A(x_0 + tu) - (x_0 + tu)^t b = k$$

$$\Leftrightarrow \frac{1}{2}(x_0^t Ax_0 + tx_0^t Au + tu^t Ax_0 + t^2 u^t Au) - x_0^t b - tu^t b = k$$

$$\Leftrightarrow \frac{1}{2}x_0^t Ax_0 - x_0^t b + \frac{1}{2}t(x_0^t Au + u^t Ax_0) - tu^t b + \frac{1}{2}t^2 u^t Au = k$$

Posons $x_0 = x^*$

$$\Rightarrow t(\frac{1}{2}(x^{*t}Au - u^t b)) + \frac{1}{2}t^2 u^t Au = k - F(x^*)$$

$$\frac{1}{2}t^2 u^t Au = k - F(x^*) \quad F(x^*) = -\frac{1}{2}x^* Ax^* < 0$$

$$t^2 = 2\frac{K - F(x^*)}{u^t Au} \Leftrightarrow t = \pm\sqrt{2\frac{k - F(x^*)}{u^t Au}}$$

Donc on a deux racines réelles.

Théorème

Soit $x \in IR^n$ q.c.q. Le résidu $r(x) = Ax - b$ est dirrigé suivant la normale à l'ellipsoîde passant par x

Preuve :

On pose

$$F(x) = \frac{1}{2}x^t A x - x^t b$$

$$\overrightarrow{grad} F(x) = Ax - b = r(x)$$

$$\Rightarrow r(x) \ est \perp Q_k$$

Théorème :

La solution du système $Ax = b$ est le point pour lequel $F(x) = \frac{1}{2}x^t Ax - x^t b$ atteint son minimum

Preuve :

$$\left\{ \begin{array}{l} Trouver \ x \in IR^n \\ Ax = b \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} min \ F(x) \\ x \in IR \end{array} \right.$$

$$F(x^* + y) = \frac{1}{2}(x^* + y)^t A (x^* + y)^t - (x^* + y)^t b$$

$$= \ \frac{1}{2}x^{*t} A x^* - x^{*t} b + \frac{1}{2} y^t A x^* + \frac{1}{2} x^{*^t} A y + \frac{1}{2} y^t A y - y^t b$$

$$= F(x^*) + \frac{1}{2} y^t A y \geq F(x^*)$$

$$F(x^*) \leq F(x) \ \forall x \in IR^n$$

$$F(x^* + y) = F(x^*) \Leftrightarrow y = 0$$

Donc $F(x^*)$ est le min de $F(x) = \frac{1}{2}x^tAx - x^tb$

Construction d'une méthode générale:

Définition :

Soit $u, v \in IR^n$ u et v sont dits conjugués ssi $< u, Av >= 0$

$$\forall\ u \in IR^n,\ \|u\|_A =< u, Au >^{\frac{1}{2}}$$

Soit x^p une approximation de x^* on construit $x^{(p+1)}$ de la forme $x^{p+1} = x^p + \lambda_p u^p$ tel que

$$F(x^{p+1}) = min(F(x^p + \lambda_p u^p)\ \ \lambda_p \in IR)$$

$$F(x^{p+1}) = \frac{1}{2}\lambda_p^2 u^{p^t} Au^p + \lambda_p u^{pt} r^p + F(x^p) = g(\lambda_p)$$

g est strictement convexe donc le minimum est atteint en

$$g'(\lambda_p) = 0 \Leftrightarrow \lambda_p = \frac{-< u^p, r^p >}{< u^p, Au^p >}$$

d'ou la méthode

$$\begin{cases} x^0 & q.c.q \\ x^{p+1} & = \quad x^p - \frac{-<u^p,r^p>}{<u^p,Au^p>}u^p \end{cases}$$

Théorème :

Si A est S.D.P et si cos $(u^p, r^p) \rightarrow_{p\rightarrow+\infty} 0$ alors la méthode converge.

Preuve :

De même que le théorème précédent, sauf il faut utiliser $(u, Au) = \|u\|_A^2$

Méthodes particulières :

a) Méthode de la plus forte descente :

on a $r(x) = Ax - b$ *est dirrigé suivant la normale à l'ellipsiode de centre de symetrie* x *on prend donc* $r(x)$ *comme direction c'est la direction de plus profonde décroissance de* $F(x)$. *d'ou la méthode*

$$\begin{cases} x^0 & q.c.q \\ x^{p+1} & = \quad x^p - \frac{\|r^p\|^2}{<r^p, Ar^p>} r^p \end{cases}$$

c'est la méthode de plus profonde déscente

b) Méthode de gradient conjugué :

Par cette méthode, u^p *est construit à partir de* u^{p-1} *de telle sorte que* u^p *et* u^{p+1} *soit conjugués c'est à dire* $< u^p,\ Au^{p+1} >= 0$

Le choix de $u^p = r^p$ *de la p.p.d est remplacé par* $u^p = r^p + \beta_p u^{p-1}$

Le premier par défaut le même que la p.p.d $u^0 = r^0$

Calcul de β_p:

on a

$$< Au^p, u^{p-1} >= 0 \Leftrightarrow < Ar^p + \beta p Au^{p-1}, u^{p-1} >= 0$$

$$\Leftrightarrow \beta_p = -\frac{< r^p, Au^{p-1} >}{< u^{p-1}, Au^{p-1} >}$$

1) on a λ_p *est caractérisé par* $< u^p, r^{p+1} >= 0$

$$x^{p+1} = x^p + \lambda_p u^p \Rightarrow r^{p+1} = r^p + \lambda_p Au^p$$

$$< r^{p+1}, r^p >=< r^p, r^p > -\frac{< u^p, r^p >}{< u^p, Au^p >} < Au^p, r^p >$$

$$= < r^p, r^p > -\frac{u^{p+1}.r^p.r^{pt}.Au^p}{< u^p, Au^p >} =< r^p, r^p > - < r^p, r^p >$$

$$= 0$$

2)

$$< u^p, r^p > = < r^p, r^p > + \beta_p < u^{p-1}, r^p >$$

$$= < r^p, r^p >$$

donc

$$\beta_p = -\frac{< r^p, \frac{r^p - r^{p-1}}{\lambda_{p-1}} >}{< u^{p-1}, \frac{r^p - r^{p-1}}{\lambda_{p-1}} >} = -\frac{< r^p, r^p >}{< r^{p-1}, r^{p-1} >} = -\frac{\|r^p\|^2}{\|r^{p-1}\|^2}$$

d'ou la méthode

$$x^0 \ q.c.q \ r^0 = Ax^0 - b \ \ u^0 = r^0$$

$$\begin{cases} \lambda_p & = & -\frac{\|r^p\|^2}{<u^p, Au^p>} \\ x^{p+1} & = & x^p + \lambda_p u^p \\ r^{p+1} & = & Ax^{p+1} + b \\ u^{p+1} & = & r^{p+1} - \frac{\|r^{p+1}\|^2}{\|r^p\|^2} u^p \end{cases}$$

Propriétés fondamentales et convergence :

La méthode de G.C permet de construire les suites x^p, u^p *en partant de* x^0, r^0 *avec les propriétés locales :*

$< u^p, Au^{p-1} >= 0, < r^{p+1}, r^p >= 0$, x^{p+1} *minimise* $F(x) = \frac{1}{2}x^t Ax - x^t b$ *sur la droite* $\{x^p + \lambda u^p : \lambda \in IR\}$*, il se trouve que ces suites ont les propriétés globales*

$$< u^i, \ Au^j >= 0 \ \forall \ 0 \leq j < i \leq N$$

$$< r^i, \ r^j >= 0 \ \forall \ 0 \leq j < i \leq N$$

x^{p+1} *minimise* $F(x)$ *sur la variété* $\{x^0 + \sum_{i=0}^{p} \lambda_i u^i, \lambda_i \in IR\}$

Théorème :

Pour tout vecteur $u_0 \in IR^N$*, il existe un plus petit entier* l *telque* $u^l = 0$*, les propriétés suivantes sont vérifiées:*

1) $x^l = x^*$

2) $< r^i,\ r^j >= 0 \ \forall\ 0 \leq j < i \leq l$

3) $< r^i,\ u^j >= 0 \ \forall\ 0 \leq j < i \leq l$

4) $< u^i,\ Au^j >= 0 \ \forall\ 0 \leq j < i \leq l$

5) $F(x^{p+1}) = \min\{F(x_0 + \sum_{i=0}^{p} \lambda_i u^i) :\ \lambda_0, ..., \lambda_p \in IR\}$

Preuve :

** l'existence de l'entier naturel* l *découle de la propriété 4)*

*** $u^l = 0 \ \Leftrightarrow x^l = x^*$

** pour démontrer les propriétés 2)-4) on procède par récurrence*

Soit

$$P_k = \left\{ \begin{array}{lcl} * \ < r^i,\ r^j > & = & 0 \ \forall\ j < i \leq k \\ * \ < r^i,\ u^j > & = & 0 \ \forall\ j < i \leq k \\ * \ < r^i,\ u^i > & = & < r^i,\ r^i > \\ * \ < u^i,\ Au^j > & = & 0 \ \forall\ j < i \leq k \end{array} \right\}$$

Supposons que ces propriétés sont vraies jusqu'à k*. Montrons qu'elles restent vraies pour* $k+1$*.*

$$< r^{k+1},\ r^j >=< r^k + \lambda_k Au^k, r^j >$$

$$=< r^k,\ r^j > +\lambda_k < Au^k, r^j >$$

$$= 0 - \frac{< u^k,\ r^k >}{< u^k,\ Au^k >} < Au^k, r^j >$$

$$= \ -\frac{< u^k, Au^k >}{< u^k,\ Au^k >} < r^k,\ r^j >= 0$$

$$< r^{k+1},\ u^j >=< r^k, u^j) + \lambda_k < Au^k, u^j >= 0$$

$$< r^{k+1}, u^j >= (r^k, u^j) + \lambda_k < Au^k, u^j >= 0$$

$$< r^i, u^i >=< r^i, r^i + \beta_i u^{i-1} >= < r^i, r^i >$$

$$< u^{k+1}, Au^j >=< r^k + \beta_k u^k, \frac{r^{j+1} - r^j}{\lambda_j} >$$

$$= \frac{1}{\lambda_j}(< r^k,\ r^{j+1} > - < r^k, r^j > +\beta_k < u^k, r^{j+1} > -\beta_k < u^k, r^j >)$$

$$= \frac{1}{\lambda_j}(< r^k, r^j + \lambda_j Au^j > - < r^u, r^j > +\beta_k < u^k, r^j + \lambda_j Au^j > -\beta_k < u^k, r^j >)$$

$$= \frac{1}{\beta_j}(< r^k, r^j > +\lambda_j < r^k, Au^j > - < r^k, rj > +\beta_k < u^k, Au^j >) = 0$$

Pour le problème de minimisation dans la variété de dimension k+1 on a :

$$g(\lambda) = F(x_0 + \sum_{i=0}^{k} \lambda_i u^i) = F(x_0 + U_k \lambda)$$

où u_k est la matrice de N lignes dont les $k+1$ colonne sont $u^0, u^1, ..., u^k$ d'ou

$$g(\lambda) = F(x_0) + \lambda^t u_k^t r^0 + \frac{1}{2}\lambda^t u_k^t A u_k \lambda$$

La solution λ minimisant $g(\lambda)$ est donc la solution du système à matrice S.D.P

$$u_k^t A u_k \lambda = -u_k^t r^0$$

La matrice étant digonale d'après la propriété 4) *donc*

$$\lambda_i = \frac{- < u^i, r^0 >}{< u^i, Au^i >} \qquad i = 0, ..., k$$

et comme

$$r^i = r^0 + A(\sum_{i=0}^{i-1} \lambda_i u^i)$$

on a

$$< u^i, r^0 >=< u^i, r^i >$$

La solution du problème de minimisation est donc $x_0 + \sum_{i=0}^{k} \alpha_i u^i$ α_i *de la méthode G.C*

Théorème

Dans la méthode du G.C, la suite $e_k = x_k - x^*$ *vérifie:*
$\|e_{k+1}\|_A = \min\{\|p(A)e_0\|_A : p \in IP^1_{k+1}$
ou IP^1_{k+1} *désigne l'ensemble de polynôme* p *de degré* $k+1$ *au plus tels que* $p(0) = 1$

Preuve :

$$e_{k+1} = x_{k+1} - x^*$$

$$= x_0 + \sum_{i=0}^{k} \lambda_i u^i - x^*$$

$$= \|e_{k+1}\|_A = \min\{\left\| e_0 + \sum_{i=0}^{k} \lambda_i u^i \right\|_A : \lambda_0, ..., \lambda_u \in IR\}$$

On montre par récurrence que u^i *vérifie*

$$u^i = P_i(A).r_0$$

$$P_i \in IP_i$$

$$* \; u^0 = r^0$$

$$* u^{i+1} = r^{i+1} + \beta_{i+1} u^i$$

$$r^{i+1} = r^0 + A(\sum_{j=0}^{i-1} \lambda_i u^j) \Rightarrow u^{i+1} = r^0 + A(\sum_{j=0}^{i} \lambda_j u^j) + \beta_{i+i} u^i$$

$$\Rightarrow u^{i+1} = r^0 + A(\sum_{j=0}^{i} \lambda_j P_j(A) r^0) + \beta_{i+1} p(A) r^0$$

$$= r^0 + (\sum_{j=0}^{i} \lambda_j P_{j+1}(A) r^0) + \beta_{i+1} P_i(A).r^0$$

$$= P_{i+1}(A) r^0$$

donc

$$\forall_i \quad u^i = P_i(A) r^0$$

Les polynômes $P_i, \; i = 0, ..., k$ *forment une base de* IP_k *leurs indépendance résultant de celle des* u^i , $i = 0, ..., k$ *qui sont orthogonaux par :*

$$\sum_{i=0}^{k} \alpha_i P_i = 0 \Rightarrow \sum_{i=0}^{k} \alpha_i P_i(A) r_0 \Leftrightarrow \sum_{i=0}^{k} \alpha_i u^i = 0 \Leftrightarrow \alpha_i = 0 \;\; \forall \; i$$

on a

$$\min\{\left\| e_0 + \sum_{i=0}^{k} \lambda_i u^i \right\|_A, \lambda_0, ..., \lambda_n \in IR\} = \min\{\left\| e_0 + \sum_{i=0}^{k} \lambda_i P_i(A) r^0 \right\|_A, \lambda_0, ..., \lambda_n \in IR\}$$

$$= \min\{\left\|e_0 + \widetilde{p}r^0\right\|_A ,\ \widetilde{p} \in IP_k\}$$

$$= \min\{\|e_0 + \widetilde{p}(A).Ae_0\|_A ,\ \widetilde{p} \in IP_k\}$$

$$= \min\{\|p(A)e_0\|_A , p \in IP^1_{k+1}\}$$

puisque

$$p \in IP^1_{k+1} \Leftrightarrow p(x) = 1 + x\widetilde{p}(x) \ \ \widetilde{p} \in IP_k$$

donc

$$\|e_{k+1}\|_A = \min\{\|p(A)e_0\|_A , p \in IP^1_{x+1}\}$$

Théorème :

Dans la méthode de G.C, l'erreur au pas k vérifie la majoration:

$$\|e_k\|_A \leq 2(\frac{\sqrt{cond(A)} - 1}{\sqrt{cond(A)} + 1})^k \|e_0\|_A$$

Preuve :

Soit $(v_1, ..., v_N)$ *la famille orthonormée de vecteurs propres de la matrice A.*

$Av_i = \lambda_i v_i \qquad v_i^t v_i = \delta_{ij}$

Cette famille est A-orthogonale $< v_i, Av_j >= \lambda_j v_i^t v_j = \lambda_j \delta_{ij}$ *de plus* $\|v_i\|_A^2 = \lambda_i$

on écrit

$$e^0 = \sum_{i=1}^{N} \varepsilon_i v_i$$

$$p(A)e_0 = \sum_i \varepsilon_i p(A).v_i = \sum_i \varepsilon_i p(\lambda_i).v_i$$

$$\|e_0\|_A^2 = \sum_i \varepsilon_i^2 \|v_i\|_A^2 = \sum_i \varepsilon_i^2 \lambda_i$$

et

$$\|p(A)e_0\|_A^2 \leq (\sum_i \lambda_i \varepsilon_i^2)\max_i p^2(\lambda_i) = \|e_0\|_A^2 \max_i p^2(\lambda_i)$$

on a

$$\|e_{k+1}\|_A = \min\{\|p(A)e_0\|_A \,, p \in IP_{k+1}^1\}$$

$$= \min\{\left\|\sum\nolimits_i \varepsilon_i p(\lambda_i) v_i\right\|_A , p \in IP_{k+1}^1\}$$

$$\leq \min\{\max_i |p(\lambda_i)| \,, IP_{k+1}^1\}. \|e_0\|_A$$

$max|p(\lambda_i)| \leq \max |p(\lambda)|_{\lambda \in (\lambda_1, \lambda_N)}$

donc

$$\|e_{k+1}\|_A \leq \min\{\|p\|_\infty \,, p \in IP_{k+1}^1\}. \|e_0\|_A$$

le min est égal à

$$\frac{1}{T_k(\frac{\lambda_N+\lambda_1}{\lambda_N-\lambda_1})}$$

et est atteint pour le polynôme

$$p(\lambda) = T_k(\frac{\lambda_N + \lambda_1 - 2\lambda}{\lambda_N - \lambda_1})/T_k(\frac{\lambda_N + \lambda_1}{\lambda_N - \lambda_1})$$

où T_k est le polynôme de tchebychev de premier espèce de degré k.

on a en fin

$$T_k(\frac{\lambda_N + \lambda_1}{\lambda_N - \lambda_1}) = T_k(\frac{cond(A)+1}{cond(A)-1}) > \frac{1}{2}(\frac{\sqrt{cond(A)}+1}{\sqrt{cond(A)}-1})^k$$

d'où

$$\|e_{k+1}\|_A \leq 2(\frac{\sqrt{cond(A)}-1}{\sqrt{cond(A)}+1})^k \|e_0\|_A$$

Méthode du G.C préconditionnée (G.C.P)

l'estimation ci-dessus donnant par la méthode du G.C., un nombre de pas proportionnel à $\sqrt{cond(A)}$, une possibilité d'obtenir une convergence plus rapide pour la résolution du système $Ax = b$ est d'appliquer la méthode du G.C à une forme préconditionnée du système $\overline{A}\overline{x} = \overline{b}$ avec $cond(\overline{A}) < cond(A)$

Le passage se fait par un changement de varuiable, c'est-à-dire: $\overline{x} = B^t x$ ou $x = -B^t\overline{x}$

où B est une matrice régulière. Par ce changement de variable, la fonctionnelle $J(x) = \frac{1}{2}x^t Ax - x^t b$ devient

$$J(x) = \overline{J}(\overline{x}) = \frac{1}{2}(B^{-t}\overline{x})^t A(B^{-t}\overline{x}) - (B^{-t}\overline{x})^t b$$

$$= \frac{1}{2}x^{-t}(B^{-1}AB^{-t})\overline{x} - x^{-t}(B^{-1}b)$$

$$= \frac{1}{2}x^{-t}\overline{A}\overline{x} - x^{-t}b$$

avec $\overline{A} = B^{-1}AB^{-t}$ et $\overline{b} = B^{-1}b$

$\overline{x} = \overline{A}^{-1}b$ minimise $J(x)$ dans IR^n

on applique la méthode itérative à $\overline{A}\overline{x} = \overline{b}$

construit $(x^{-k})_{k \in IN}$ et on obtient x^k par $x^k = B^{-t}x^{-k}$ qui converge vers $x^A = A^{-1}b$

les erreurs $e_k = x^k - x^A$ et $\overline{e}_k = \overline{x}^k - \overline{x}^A$ sont liées par :

$$\|e_k\|_A = \|\overline{e}_k\|_{\overline{A}} \quad \forall\, k \geq 0$$

Pour la méthode du G.C appliquée à $\overline{A}\overline{x} = \overline{b}$, on aura

$$\|e_k\|_A \leq 2(\frac{\sqrt{cond(\overline{A})} + 1}{\sqrt{cond(\overline{A})} - 1})^k \|e_0\|_A \quad etcond(\overline{A}) < cond(A)$$

donc $\overline{x}^k$ *converge plus rapidement que* x^k, *il est plus commode d'introduire la matrice S.D.P* C *définie par :* $C = BB^t$ *qui est dite matrice de préconditionnement.*

On a

$$C^{-1}A = B^{-t}B^{-1}A$$
$$= B^{-t}B^{-1}AB^{-t}B^t = B^{-t}\overline{A}B^t$$

$C^{-1}A$ *est semblable à* $\overline{A}$

on appelle parfois système précondotionné $CAx = C^{-1}b$

Algorithme du G.C.P

l'algorithme du G.C.P appliquée à $\overline{A}\overline{x} = \overline{b}$ *s'écrit :*

$$\overline{r}_0 = \overline{A}x^0 - \overline{b} \ , \ \ \overline{u}_0 = \overline{r}_0$$

$$\begin{cases} \overline{\alpha}_k & = & \frac{<\overline{r}_k, \overline{r}_k>}{<\overline{u}_k, \overline{A}\overline{u}_k>} \\ \overline{x}_{k+1} & = & \overline{x}^k + \overline{\alpha}_k \overline{u}_k \\ \overline{r}^{k+1} & = & \overline{r}^k + \overline{\alpha}_k A \overline{u}_k \\ \overline{\beta}_k & = & -\frac{<\overline{r}_{k+1}, \overline{r}_{k+1}}{<\overline{r}_k, \overline{r}_k>} \\ \overline{u}_{k+1} & = & \overline{r}_{k+1} + \overline{\beta}_k \overline{u}_k \end{cases}$$

Remarque :

Cette première version, bienque permettant de construire x_k, *n'est pas bien adapté à la résolution du système* $Ax = b$ *car elle necessite notamment le calcul et le stockage de la matrice* $\overline{A}$. *l'algorithme du G.C.P consiste à travailler directement avec* A, b, x_k, r_k...

Pour travailler avec A, b, x_k, r_k, u_k...

on a : $u_k = B^{-t}\overline{u}_k \Rightarrow x^{k+1} = x^k + \overline{\alpha}_k u_k$

on a $< \overline{u}_k, \overline{A}\overline{u}_k >$

$\overline{r}_k = \overline{A}\overline{x}^k - \overline{b} = B^{-1}AB^{-t}\overline{x}^{-k} - B^{-t}b = B^{-1}r^k$

d'où $< \overline{r}^k, \overline{r}^k > = < B^{-1}r^k, Br^k > = < r^k, B^{-t}B^{-1}r^k >$

$= < r^k, c^{-1}r^k >$

$a_{k+1} = B^{-t}u_{k+1} = -c^{-1}r_{k+1} + \overline{B}_x u_k$

En notant $g_x = c^{-1}r_k$, on peut écrire l'algorithme du G.C.P pour le système $Ax = b$ comme suit :

$$r_0 = Ax^0 - b \ , g_0 = c^{-1}r_0 \ , \ u_0 = -g_0$$

$$\alpha_k = \frac{< r_k, g_k >}{< u_k, Au_k >}$$

$$x^{k+1} = x^k - \alpha_k u_k$$

$$r_{k+1} = \alpha_k - \alpha A u_k$$

$$g_{k+1} = c^{-1}r_{k+1}$$

$$\beta_x = -\frac{< r_{k+1}, g_{k+1} >}{< r_k, g_k >}$$

$$u_{k+1} = r_{k+1} + \beta_k u_k$$

Remarque :

la seule complication par rapport à l'algorithme du G.C est la construction de la suite $g_k = C^{-1}r_k$ c'est-à-dire la résolution à chaque pas d'un système de matrice c: $Cg_k = r_k$

pour conserver le même coût ue celui de G.C, le choix de la matrice de préconditionnement C est limité à des matrices du type : $C = BB^t$, avec B triangulaire inferieure.

Résolution des systèmes non symétriques

Introduction :

Pour résoudre le système lineaire $Ax = b$ avec A non symétrique, des méthodes ont été developpées pour résoudre le système $A^tAx = A^tb$.

ou A^tA est S.D.P.

Ces méthodes sont limités par le fait que $cond(A^tA) > cond(A)$ malgré les techniques de précondition.

En remarquant des systèmes symétriques, plusieurs travaux sont menés dans le sens de généraliser cette méthode au cas des systèmes non symétriques.

Soit $V = [v_1, ..., v_k]$ *une base de* K

$W_k = [w_1, ..., w_k]$ *une base de* L_k

$$on\ a \left\{ \begin{array}{c} trouver\ x = x^0 + z \quad tq \\ < Ax - b,\ \forall w \in L_k \end{array} \right. \Leftrightarrow \left\{ \begin{array}{c} trouver\ z \in k_x\ tq \\ < Az - r^0, w >= 0\ \forall w \in L_k \end{array} \right.$$

avec $r^0 = b - Ax^0$

$z \in K_m \Rightarrow z \in V_k y$

$< A_2 - r^0, w >= 0\ \forall w \in L_k \Leftrightarrow W_k^t(AV_k y - r^0) = 0$

$\Leftrightarrow w_k^t AV_k.y = W_k^t r^0$

donc $x = x^0 + V_k y = x^0 + V_k(W_x^t AV_k)^{-1}W_k^t)r^0$

La méthode de Petrov-Galertin est bien définie si $W_x^t AV_k$ *est inversible*

- *Si* $L_k = K_k$ *c'est la méthode de Galerkin*

- *Si* $L_k = AK_k$ *c'est la méthode least-squares*

Théorème :

Si $L_k = AK_k$, $\overline{x}$ *est une approximation de la solution du système donnée par PG ssi* $\overline{x}$ *minimise* $\|b - AX\|_2\ \forall x \in x_0 + K_k$

Preuve :

Soit $x = x^0 + z \quad z \in K_k$

$\|Ax - b\|^2 = \|A[x - \overline{x}) + \overline{x}] - b\|^2 =< A(x - \overline{x} + \overline{x}) - b, A(x - \overline{x} + \overline{x}) - b >$

$= < Ax - b, A\overline{x} - b > -2 < A\overline{x} - b, A(x - \overline{x}) > + < A(x - \overline{x}), A(x - \overline{x}) >$

on a $x - \overline{x} \in K_k \Rightarrow A(x - \overline{x}) \in L_k$ *donc* $< A\overline{x} - b, A(x - \overline{x}) >= 0$

$\|Ax - b\|^2 = \|A\overline{x} - b\|^2 + \|A(x - \overline{x})\|^2$

$\Rightarrow \forall x \in K_x, \|A\overline{x} - b\|^2 \leq \|Ax - b\|^2$

inversement

si $\overline{x} = x^0 + 2 \in x^0 + K_k$ *minimise* $\|Ax - b\|^2\ \forall x \in K_x$

alors la forme quadratique $Q(\alpha) = \|r_0 - A\overline{z} + \alpha Az\|^2$ *est minimale pour* $\alpha = 0$

$\frac{dQ(\alpha)}{d\alpha} = 0 \Leftrightarrow \alpha = 0 \; Q(0) = \|r_0 - A\overline{z}\|^2 = \|b - A\overline{x}\|^2$

Remarque :

- une condition d'optimalité n'est vérifiée pour la méthode de Galerkin $L_k = K_n$ *que si* A *est S.D.P cette condition est* $\|A\overline{x} - b\|_{A^{-1}} = \min_{x \in K_k} \|Ax - b\|_{A^{-1}}$

- plusieurs difficultés sont liées à cette méthode à savoir

1) si k *est grand, la matrice* $W_k^t A V_k$ *est pleine et*

2)$W_k^t A V_k$ *est male conditionnée*

3) la construction de la solution approchée $x = x^0 + 2$ *necessite le stockage des vecteurs* $v_i, i = 1, 2, ..., k$ *pour détourner ces difficultés plusieurs méthodes ont été développées. Ces méthodes se généralisent sous le nom "méthode de Petrov-Galerkin-krylav"*

Algorithmes de Petrov-Galerkin-krylav

Soit b(.,.) une forme bilineaire sur IR *et* θ *une fonction croissante à valeurs entières définie sur* IN *tq* $0 \leq \theta(i) \leq i + 1$ *on définie une classe générale d'algorithmes.*

Algorithme PGK(b, θ)

1) $U_0 = r_0 = b - Ax_0$

2) $K = 0, 1, 2, ...$ *jusqu'à convergence*

$x^{k+1} = x^0 + \sum_{i=0}^{k} \lambda_i u^i$

$r^{k+1} = r^0 - \sum^k{}_{i=0} \lambda_i u^i$

les coefficients λ_i *sont tq :*$\begin{cases} < r^{k+1}, u^i >= 0 \; \forall \; i \leq k \\ ou \\ < r^{k+1}, Au^i >= 0 \; \forall \; i \leq k \end{cases}$

$$\begin{cases} u^{k+1} = r^{k+1} + \sum^k{}_{i=\theta(x)} \beta_i^x u^i \\ ou \\ \beta_{k+1}^k u^{k+1} = Au^k + \sum_{i=\theta(x)}^k \beta_i^x u^i \end{cases}$$

les $\{\beta_i^x\}$ *sont tq :*$b(u^{k+1}, u^i) = 0 \;\; \theta(k) \leq i \leq k$

Algorithme : $TPGK(b, \theta, \phi)$

le même algorithme que PGX sauf que $\sum_{i=0}^{x} \lambda_i u^i$ *est trouvée sous la forme* $\sum_{i=\theta(k)}^{x} \lambda_i u^i$ *ou* ϕ *vérifie les mêmes hypothèses que* θ

Cas particuliers de P.G.K

a) Méthode du Residu conjugué Géneralisé G.C.R

C'est une méthode de forme $PGK(b,\theta)$ *avec* $\theta(i) = 0 \ \forall\ i,$

$< r_{i+1}, Ap_i >= 0 \ u_{i+1} = r_{i+1} + \sum_{i=0}^{x} \beta_i^{(i)} u_j \ \ b(u,v) =< Au, Av)$

Chapter 3

calcul des valeurs propres et des vecteurs d'une matrice

3.1 Introduction :

La recherche des valeurs propres et vecteurs propres d'une matrice se rencontre dans des problèmes assez divers telsque : la recherche du rayon spectrale de la methode de Jaccobi pour calculer le paramètre optimale de relation mais les complications qui se présentent dans la détermination numérique des élements propres d'une matrice sont bien plus grandes que pour la résolution des systèmes linéaires.

1) Détermination du polynôme caractéristique d'une matrice A :

Soit A une matrice réelle, de type (n,n) dont on cherche les valeurs propres.

Pour trouver les éléments propres de A, la méthode la plus naturelle consiste à déterminer le polynôme caractéristique, à chercher se zéros $\lambda_1, ..., \lambda_n$; puis à résoudre les systèmes linéaires : $Ax = \lambda_i x \ \ (i = 1, ..., n)$.

Ce schéma n'est pas trés satisfaisant pour les applications numériques en effet la détermination du polynôme caractéristique necessite un trés grand nombre d'opérations, ce qui signifie une accumulation des erreurs d'arrandi, et une précision faible sur les coéfficients du polynôme.

Toutefois, comme la connaissance du polynôme caractéristique peut être utile en elle-même, indépendamment de la recherchedes valeurs propres, nous allons donner une des méthode numérique de ce polynôme.

$P_A(n) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n$ $(a_0 \neq 0)$

En pratique, cette méthode ne sera utilisable que pour de petites valeurs de n, $n \leq 10)$

3.2 MéthodedeLeverrier:

Soient $x_1, ..., x_n$ *les valeurs propres, c'est à dire les zéros du polynôme caractéristique* P_A.

Soit $\tau_1 = x_1 + ... + x_n$

$\tau_2 = \sum_{i>j} x_i x_j$

$\tau_3 = \sum_{i>j>k} x_i x_j x_k$

.

.

$\tau_n = x_1 + ... + x_n$

on a les τ_i *sont liés aux coéfficients* a_i *du polynôme* P_A *par les relations :*

$\tau_i = (-1)^i \frac{a_i}{a_0}$

Considérons les sommes de Newton des x_i

$S_1 = x_1 + ... + x_n$

$S_2 = x_1^2 + ... + x_n^2$

.

.

$S_k = x_1^k + ... + x_n^k$ $(k \in IN)$

les sommes S_k *sont reliées aux coefficients* a_i *par les formules dites de Newton :*

$a_i = -a_0 S_1$

$a_1 S_1 + 2a_2 = -a_0 S_2$

$a_1 S_2 + a_2 S_1 + 3a_3 = -a_0 S_3$

.

.

.

$a_1 S_{k-1} + a_2 S_{k-2} + ka_k = -a_0 S_k$

.

.

$a_1 S_{n-1} + a_{n-1} S_1 + n a_n \quad = -a_0 S_n$

Connaissant les sommes S_k, il suffit de résoudre un système linéaire triangulaire pour calculer les coefficients $\frac{a_i}{a_0}$

les S_k sont facilement accessibles dès que l'on connait les puissances A^k de la matrice A.

En effet : $S_1 = x_1 + ... + x_n$ est égale à la trace de A.

$S_1 = \sum_{i=1}^{n} a_{ii}$

S_k n'est autre que la trace de A^k : car si $x_1 + ... + x_n$ sont les valeurs propres de A, $x_1^k + ... + x_n^k$ sont les valeurs propres de A^k

Cette méthode donne exactement les coefficients du polynôme caractéristique, mais les nombres d'opérations necessaire est très grand qui est de l'ordre de n^k multiple

Méthode de Leverrier améliorée :

Cette méthode donne les coéfficients du polynômes caractéristiques sans passer par la résolution d'un système linéaire: mais il faut quand même calculer $(n-1)$ produits de matrice (n, n).

En effet $A_i, B_i \quad i = 1, ..., n$ est les n nombre p_i définis par :

$A_1 = A \;\; p_1 = trace(A_1), \;\; B_1 = A_1 - p_1 I$

$A_2 = B_1 A, \;\; P_2 = \frac{1}{2} tr(A_2); B_2 = A_2 - p_2 I$

.

.

$A_n = B_{n-1} A, \;\; p_n = \frac{1}{n} tr(A_n), \;\; B_n = A_n - p_n I$

Proposition :

Les polynômes caractéristiques de A est (à un facteur près)

$P_A(\lambda) = \lambda^n - p_1 \lambda^{n-1} - ... - p_{n-1}\lambda - p_n$

et $B_n = 0$

Preuve : Soit $p_a(\lambda)$ le plolynôme caractéristique de A :

$P_A(\lambda) = \det(\lambda I - A) = \lambda^n - b_1 \lambda^{n-1} - ... - b_{n-1}\lambda - b_n$

par recurrence sur k nous alons montrer que b_k est égale à P_k pour tout k $(1 \leq k \leq n)$

pour $k = 1$: $b_1 = tr(A) = tr(A_1) = p_1$

supposons que $p_i = b_i \;\; 1 \leq i \leq k-1$

montrons que $p_k = b_k$

$A_k = B_{k-1}A = (A_{k-1} - p_{k-1}I)A = A_{k-1}A - p_{k-1}A$

$\quad = B_{k-2}A^2 - p_{k-1}A = A_{k-2}A^2 - p_{k-2}A^2 - p_{k-1}A$

par recurrence on verifie que :

$A_k = A_{k-h}A^h - p_{k-h}A^h - p_{k-h+1}A^{h-1}... - p_{k-1}A \qquad 0 \leq h \leq k-1$

d'ou pour $h = k-1$

$A_k = A_1A^{k-1} - p_1A^{k-1} - p_2A^{k-2} - ... - p_{k-1}A$

$= A^k - p_1A^{k-1} - ... - p_{k-1}A$

par définition de p_k

$P_k = \frac{1}{k}tr(A_k) = \frac{1}{k}tr[A^k - p_1A^{k-1} - ... - p_{k-1}A]$

$= \frac{1}{k}[S_k - p_1S_{k-1}... - k_{k-1}S_1$

ou les S_i *sont les sommes de Newton des racines de* P_A

utilisent l'hypothèse de récurrence : $p_i = b_i \qquad 1 \leq i \leq k-1$

on a $p_k = \frac{1}{k}[S_k - b_1S_{k-1} - ... - b_{k-1}S_1] = b_k$ *d'après la nieme forme de Newton .*

Donc $p_k = b_k$

Montrons que $B_1 = 0$

$B_n = A_n - p_nI$

$\quad = A^n - p_1A^{n-1} - p_2A^{n-1}... - p_nI$

$B_n = p_A(A)$

on a $B_n = 0$ *(Théorème de Cayley-Hamilton)*

Remarque :

par cette méthode on obtient indirectement l'inverse de A, lorsque A est inversible:

$B_1 = 0 \Leftrightarrow A_n = p_nI$

$\qquad \Leftrightarrow B_{n-1}A = p_nI$

Si A est régulière : $p_n = \det(A) \neq 0$ *et* $A^{-1} = \frac{B_{n-1}}{p_n}$

2) calcul de la valeur propre de A de module maximum:

Soient $\lambda_1, ..., \lambda_n$ *les valeurs propres de A , on peut supposer que leurs modules sont. rangés dans l'ordre décroissant :* $|\lambda_1| \geq ... \geq |\lambda_n|$

il s'agit de calculer λ_1*et éventuellement les autres valeurs propres de module égale à* $|\lambda_1|$, *ainsi que les vecteurs propres associés.*

on suppose que les λ_i *sont simples.*

1er cas : $|\lambda_1| > |\lambda_2|$

Supposons qu'il existe dans IR^n *une base* $(x_1, ..., x_n)$ *de vecteurs propres de la matrice* A, *associées respectivement à* $(\lambda_1, ..., \lambda_n)$ *pour tout vecteur* $y^0 \in IR^n$ $(y^0 \neq 0)$, *soient* $a_1, ..., a_n$ *les coordonnées de* y^0 *dans la base* $(x_1, ..., x_n)$: $y^0 = a_1x_1 + ... + a_nx_n$

on construit la suite $y^{(k)}$ *par :* $y^k = Ay^{k-1} = A^k y^0$

puisque $x_1, ..., x_n$ *sont des vecteurs propres :*

$y^{(k)} = a_1\lambda_1^k x_1 + ... + a_n\lambda_n^k x_n$

Si $a_1 \neq 0$

$$y^{(k)} = a_1\lambda_1^k[x_1 + \tfrac{a_2}{a_1}\tfrac{\lambda_2^k}{\lambda_1^k}x_2 + ... + \tfrac{a_n}{a_1}\tfrac{\lambda_n^k}{\lambda_1^k}x_n]$$

on a : $\left|\frac{\lambda_n}{\lambda_1}\right| \leq \left|\frac{\lambda_{n-1}}{\lambda_1}\right| \leq ... \leq \left|\frac{\lambda_2}{\lambda_1}\right| < 1$

$\Rightarrow$ *on a* $\left\|\frac{a_2}{a_1}\frac{\lambda_2^k}{\lambda_1^k}x_2 + ... + \frac{a_n}{a_1}\frac{\lambda_n^k}{\lambda_1^k}x_n\right\| \leq \left|\frac{\lambda_2^k}{\lambda_1^k}\right| \sum_{i=2}^n \left|\frac{a_i}{a_1}\right| \|x_i\|$

$\Rightarrow y^k = a_1\lambda_1^k[x_1 + \theta(\frac{\lambda_2^k}{\lambda_1^k})]$

Soit $(\rho_j)_{j=1,...n}$ *la base canonique de* IR^n, $\exists$ *j tq la* j^{ieme} *composante de* x_1*soit non nulle (puisque* x_0*)*

$$\Rightarrow \left|\frac{\lambda_j^{k+1}}{\lambda_j^k}\right| = \lambda_1 \frac{(x_1)_j + \theta(\frac{\lambda_2^k}{\lambda_1^k})}{(x_1)_j + \theta} \to k \to +\infty\lambda_1$$

donc si $a_1 \neq 0$, *alors* $\frac{\lambda_j^{k+1}}{\lambda_j^k} \to \lambda_1 \ \ \forall j$ *tq* $(x_1)_j \neq 0$

Remarques :

1) Si $a_1 = 0$ *: le terme prépandérant de* y^k *est alors* $a_2\lambda_2^k x_2$, *et le résultat de* *convergence est faux. Si a_1, est petit, il se peut qu'il faille un grand nombre de'iteration vant que le terme $a_2\lambda_2^k x_2$ ne devienne négligeable devant $a_1\lambda_1^k x_1$, dans ce cas, il est préférable de recommencer l'itération avec un autre vecteur initial y^0 (même pour $a_0 = 0$)*

2) Si A n'admet pas de base de vecteurs propres, on utilise la base $(x_1, ...x_n)$ dans laquel $A = J$ (matrice de Jordan) et on construit la suite $y^k = J^k y^{(0)}$:

$$y^{(0)} = \begin{pmatrix} a_1 \\ . \\ a_n \end{pmatrix}$$

$$y^k = J^k \begin{pmatrix} a_1 \\ . \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda_1^k & & \\ & J_2^k & \\ & & \\ & & J_p^k \end{pmatrix} \begin{pmatrix} a_1 \\ . \\ . \\ a_n \end{pmatrix}$$

(λ_1 est une valeur propre simple, donc le bloc de Jordan est réduit à l'ordre 1)

$$y^k = \lambda_1^k \begin{pmatrix} 1 & & \\ & \theta_1(\frac{\lambda_2^k}{\lambda_1^k}) & \\ & & \\ & & \theta_1(\frac{\lambda_2^k}{\lambda_1^k}) \end{pmatrix} \begin{pmatrix} a_1 \\ . \\ . \\ . \\ a_n \end{pmatrix}$$

il est possible de tenir les mêmes raisonnement que dans le cas ou il existe une base de vecteur propres de A.

Deuxième cas : $|\lambda_1| = |\lambda_2|$

On suppose encore qu'il existe une base de vecteurs propres $x_1, ..., x_n$, correspondantrespective aux valeurs propres : $\lambda_1, ..., \lambda_n$

si $\lambda_1 = -\lambda_2$: λ_1 réelle

Soit $y^{(0)} = a_1 x_{1+...+} a_n x_n$

on pose $y^{(k)} = Ay^{(k-1)} = A^{(k)} y^0$

$y^x = a_1\lambda_1^k x_1 + a_2\lambda_2^k x_2 + ... + a_n\lambda_n^k x_n$

$y^{2k} = a_1\lambda_1^{2k} x_1 + a_2\lambda_2^{2k} x_2 + ... + a_n\lambda_n^{2k} x_n$

$= \lambda_1^{2k}[a_1x_1 + a_2x_2 + \theta(\frac{\lambda_3^{2k}}{\lambda_1^{2k}})]$

$y^{2k+1} = \ \lambda_1^{2k+1}[a_1x_1 - a_2x_2 + \theta((\frac{\lambda_3^{2k}}{\lambda_1^{2k}})^{2k+1})]$

Soit $(e_i)_{j=1-n}$ *la base canonique de* IR^n
$\forall\ j\ \ tq\ (a_1x_1 + a_2x_2)_j \neq 0$ *on a :*

$\frac{y_j^{2k+2}}{y_j^{2k}} \to \lambda_1^2$ *quand* $k \to +\infty$

si $\lambda_1 = \lambda_2$ *on utilise le même raisonnement et on aura* $\frac{\lambda_j^{2k+1}}{\lambda_j^{2k}} \to \lambda_1$ *quand* $k \to +\infty$

si $\lambda_1 = \overline{\lambda_2}$: λ_1 *est complexe; posons* $\lambda_1 = re^{i\theta},\ \lambda_2 = re^{-i\theta}$
$\theta \neq 0\,[2\pi]$
on pose $p = -r(\lambda_1 + \lambda_2) = -2r\cos\theta$
les valeurs propres λ_1 *et* λ_2 *seront déterminées par la connaissance de* r *et de* $\cos\theta$ *donc par celle de* r *et* p.
Soit $y^0 = a_1x_1 + ... + a_nx_n$
$y^k = r^k[a_1e^{ik\theta} + a_2e^{-ik\theta}x_2 + \theta(\frac{\lambda_3}{r})^k]$

$\forall j \quad y_j^{(k-1)} +\ y_j^{(k-2)} -\ y_j^{(k)}\ y_j^{(k+1)} = a_1a_2(x_1)_j(x_2)_jr^{2k+1}(e^{-3i\theta} + e^{3i\theta} - e^{i\theta} - e^{-i\theta})$

$+\ \theta((\frac{\lambda_3}{r})^{2k+1})$

$y_j^{(k-1)} +\ y_j^{(k-1)} -\ (y_j^k)^2 = a_1a_2(x_1)_j(x_2)_jr^{2k}(e^{-3i\theta} + e^{3i\theta} - 2) + \theta(\frac{\lambda_3}{r})^{2k}$
or $e^{3i\theta} + e^{-3i\theta} - e^{i\theta} - e^{i\theta} = 2\cos\theta(e^{2i\theta} + e^{-2i\theta} - 2)$
donc , si $a_1a_2 \neq 0,\ \forall j\ \ tq\ (x_1) \neq 0$ *et* $(x_2) \neq 0$ *on aura*

$\lim_{k\to+\infty}\frac{y_j^{(k-1)} + y_j^{(k+1)} - y_j^k + y_j^{k+1}}{y_j^{(k-1)} + y_j^{(k+1)} - (y_j^k)^2} = -p$

De même :
$y_j^k +\ (y_j^{k+2}) = a_1a_2(x_1)_j(x_2)_jr^{2k+2}(e^{2i\theta} + e^{-2i\theta} - 2 + \theta(\frac{\lambda_3}{r})^{k+2})$

$y_j^{k-1} + \ y_j^{k+1} - (y_j^k)^2 = a_1 a_2 (x_1)_j (x_2)_j r^{2k} (e^{2i\theta} + e^{-2i\theta} - 2 + \theta(\frac{\lambda_3}{r})^{2k+2}))$

$\lim_{k\to\infty} \frac{y_j^{(k)} + y_j^{(k+2)} - (\ y_j^{(k+1)})^2}{y_j^{(k-1)} + y_j^{(k+1)} - (y_j^k)^2} = r^2$

3.3 Algorithme L.R :

Dans ce paragraphe on cherche à construire une suite (A_k) de matrice semblable à A, donc qui ont les même valeurs propre de A, et qui donne à la limite $(k \to +\infty)$ des informations sur les valeurs propres de A. On note R_i les matrices triangulaires superieures et L_i les matrices triangulaires inférieures qui ont la valeur1 sur la diagonale.

Supposons que A est factorisable sous le forme $A = L.U$ on pose :

$A_1 = A = L_i R_1$

$A_2 = R_1 L_1 = L_1^{-1} A L_1$

Si A_2 est factorisable sous la forme LU, on pose :

$A_2 = L_2 R_2$

$A_3 = R_2 L_2 = L_2^{-1} A_2 L_2$

.

.

.

$A_k = L_k.R_k$ *si* A_k *est factorisable L.U*

$A_{k+1} = R_k L_k = L_k^{-1} A_k L_k$

par récurrence on a :

$A_{k+1} = (L_k^{-1} L_{k-1}^{-1} .... L_1^{-1}) A (L_1 ... L_{k-1} L_x)$

Soit $M_k = L_1 ... L_k$, M_k *est triangulaire inférieure avec la valeur 1 sur la diagonale*

on a alors $A_{k+1} = M_k^{-1} A M_k \Rightarrow M_k A_{K+1} = A M_k$

Posons $S_k = R_k ... R_1$; S_k *triangulaire supérieure*

on a $M_k S_k = L_1 ... L_k R_k ... R_1$

$\quad\quad = L_1 ... L_{k-1} A_k R_{k-1} ... R_1$

$\quad\quad = A.M_{k-1} S_{k-1}$

d'ou $M_k S_k = A^k$

3.4 Convergence de l'algorithme L.R :

Soit A une matrice $n \times n$. On suppose que A admet n valeur propres distinctes en module :

$|\lambda_1| > ... > |\lambda_n|$. *On suppose on outre que A est toutes les A_k sont factorisables sous la forme $L.U$ (condition pour que l'algorithme soit applicable).*

Soit $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$, *il X telle* $A = XDX^{-1} \ = \ XDY$ *avec* $Y = X^{-1}$

alors on a : $A^k = XD^kY$

a) On suppose que X et Y sont décomposable sous la forme $L.L'$ avec L triangulaire inférieur avec 1 sur la diagonale et u triangulaire supérieure.

$X = L_x U_x$ *et* $X = Ly.Uy$

Alors : $A^x = L_x U_x D^x L_y.U_y$

$L_x U_x (D^k L_y D^{-k}) D^k U_y$

Soit $d_{ij} = (D^k L_y D^{-k})_{ij}$ *et* $l_{ij} = (L_y)_{ij}$ *on a :*

$$d_{ij} = \left\{ \begin{array}{c} 0 \ \ si \ \ i < j \\ 1 \quad si \quad i = j \\ l_{ij}(\frac{\lambda_i}{\lambda_j})^k \ si \ \ i > j \end{array} \right\}$$

or $\left|\frac{\lambda_i}{\lambda_j}\right| < 1$ *pour* $i > j$ *donc* $lim_{k\to\infty} \left|\frac{\lambda_i}{\lambda_j}\right|^k = 0$

par conséquent $D^k L_y D^{-k} \rightarrow_{k\to+\infty} I$ *et on peut écrire* $D^k L_y D^{-k} = I + E_k$ *avec* $E_k \rightarrow_{k\to+\infty} 0$

donc $A^k = L_x U_x (I + Ex) D^k U_y$

$\qquad = L_x (U_x + U_x E_x) D^k U_y$

$\qquad = L_x (I + U_x E_x U_x^{-1}) U_x D^k L_y$

La matrice $I + U_x E_x U_x^{-1} \rightarrow_{k\to+\infty} I \Rightarrow (I + U_x E_x U_x^{-1}) = \widetilde{L}_k \widetilde{U}_k$ *de plus* $\widetilde{L}_k \to I$ *et* $\widetilde{U}_k \to I$

$A^x = \ L_x \widetilde{L}_k \widetilde{U}_k U_x D^x L_y$

La décomposition de A^x sous la forme $L.U$ et unique lorsque L est triangulaire inférieure avec 1 sur la diagonale et U triangulaire supérieure par identification on obtient :

$L_X \widetilde{L}_k = M_k$ *et* $\widetilde{U}_k U_x D^k L_y = S_k$

donc $M_k \rightarrow_{k\to+\infty} L_k$ *et* $A_{k+1} = M_k^{-1} A M_k \rightarrow_{k\to+\infty} L_x^{-1} A L_x$

on a d'autre part :

$L_x^{-1}AL_x = L_x^{-1}XDYL_x$

$= L_x^{-1}L_xU_xDU_x^{-1}L_x^{-1}L_x$

$= U_xDU_k^{-1}$ *triangulaire supérieure*

donc $A_{k+1} \rightarrow U_xDU_x^{-1}$

or $U_xDU_x^{-1}$*est demblable à* D *et de la forme* $\begin{pmatrix} \lambda_1 & \\ 0 & \lambda_n \end{pmatrix}$

on a donc la conclusion suivante : (A_k) *est convergente et la diagonale de* A_k *tend vers* $(\ \lambda_1 - \lambda_n)$

b) Si Y *n'est pas factorisable sous la forme L.U il existe une matrice de permutation* P_y *telle que* $P_y.Y = L_y.U_y$ *et* $P_y^T L_y P_y$ *triangulaire inférieure avec* 1 *sur la diagonale*

$A^k = XD^kY = XD^kP_y^TL_yU_y = XP_y^T(P_yD^kP_y^T)L_y.U_y$

$P_yD^kP_y^T$ *est diagonale*

Si on pose $P_yDP_y^T = D_0 = \begin{pmatrix} \lambda_{i_1} & 0 \\ 0 & \lambda_{i_n} \end{pmatrix}$

on a $P_yD^kP_y^T = D_0^k$

A^k *s'écrit alors :* $A^k = XP_y^T(D_0^kL_yD_0^{-k})D_0^kU_y$

et $D_0^kL_yD_0^{-k} = P_yD^kP_y^TL_yP_y^TD^{-k}P_y^T = P_yD^k\widetilde{L}_yD^{-k}P_y^T$

Soient $\widetilde{d}_{ij} = (D^k\widetilde{L}_yD^{-k})_{ij}$ *et* $\widetilde{l}_{ij} = (\widetilde{L})_{ij}$

on a alors :

$$\widetilde{d}_{ij} = \begin{pmatrix} 0 & i > i \\ 1 & i = 0 \\ \widetilde{l}_{ij}(\frac{\lambda_i}{\lambda_j})^k & j < i \end{pmatrix}$$

donc $D^k\widetilde{L}_yD^{-k} \rightarrow_{k\rightarrow+\infty} I$

On peut donc écrire $D^k\widetilde{L}_yD^{-k} = I + E_k$ *avec* $E_k \rightarrow 0$

$A^k = XP_y^T(I+E_k)D_0^kU_y$

On suppose que $XP_y^T = L_xU_x$

Alors $A^k = L_x(U_x + U_xE_k)D_0^kU_y$

$\qquad = L_x(I + U_xE_kU_x^{-1})U_xD_0^kU_y$

La matrice $I + U_xE_kU_x^{-1} \rightarrow_{k\rightarrow+\infty} I \Rightarrow I + U_xE_kU_x^{-1} = \widehat{L}_x.\widehat{U}_k$

avec $\widehat{L}_k \rightarrow I$ *et* $\widehat{U}_k \rightarrow I$

donc $A^k = L_x\widehat{L}_k\widehat{U}_kU_xD_0^kU_y = M_kS_k$

Par identification $M_k = L_x\widehat{L}_k$ *et* $S_k = \widehat{U}_kU_xD_0^kU_y$

donc $M_k \rightarrow_{k\rightarrow+\infty} L_x$

d'ou $A_{x+1} = M_k^{-1}AM_k \rightarrow_{k\rightarrow+\infty} L_x^{-1}AL_x$

d'autre part $L_x^{-1}AL_x = U_xD_0U_x^{-1}$

donc $A_{k+1} \rightarrow U_xD_0U_x^{-1}$ *qui est la forme* $\begin{pmatrix} \lambda_{i_1} & 0 \\ 0 & \lambda_{i_n} \end{pmatrix}$

Proposition : Sous les synthèses suivantes :

i) $|\lambda_1| > ... > |\lambda_n|$

ii) Les matrices A_k *sont factorisables sous la forme L.U*

iii) X *et* $Y = X^{-1}$ *sont factorisable L.U ou*

iii)bis) XP_y^T *est factorisable L.U* (P_y *matrice de permutation associée à* Y*)*

alors :

L'algorithme L.R est constructible; les matrices A_k *convergent vers une matrice triangulaire supérieure et la diagonale de* A_k *tend vers :*

$(\lambda_1, ..., \lambda_n)$ *dans le cas iii)*

$(\lambda_{i_1}, ..., \lambda_{i_n})$*dans le cas iii) bis)*

$\qquad = perm(\lambda_1, ..., \lambda_n)$

3.5 Algorithme QR :

Notation : on note $(sup, +)$ *l'ensemble des matrices carrées, triangulaires supérieures, dont les élements diagonaux dont striclement positifs.*

Dans ce paragraphe, on utilise la décomposition des matrices sous forme $Q.U$ avec Q orthogonale et $U \in (\sup, +)$ cette décomposition existe est unique (voir méthode de Hansholder) on construit la suite de matrice A_k de la façon suivante :

on pose

$A_1 = A = Q_1.R_1$

$A_2 = R_1 Q_1 = Q_1^T Q_1 R_1 Q_1$

$\quad = Q_1^T A_1 Q_1$

A_2 est semblable à A_1 et non singulière

On suppose que l'on obtient A_k non singulière et semblable à A.

Alors on factorise $A_x : Q_k.R_k$

et on pose $A_{x+1} = R_{k+1} Q_k$

$\quad = Q_k^T A_k Q_k$

$\quad = Q_k^T Q_{k-1}^T ... Q_1^T A Q_1 ... Q_k$

Posons $M_k = Q_k^T ... Q_1^T$

$\quad S_k = Q_1 ... Q_k$

Avec ces notations

$A_{k+1} = M_x^T A M_k$

on a d'autre part :

$\quad M_k S_k = A M_{k+1} S_{k-1}$

d'ou $\quad M_k S_k = A^x$

Convergence de l'algorithme QR

1) cas ou $|\lambda_1| > ... > |\lambda_n|$

Soit $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$ $\exists X$ *telle que* $A = XDX^{-1}$ *on pose* $Y = X^{-1}$

a) On suppose que $Y = L_y U_y$ ou L_y est triangulaire inférieure avec 1 sur la diagonale, U_y triannngulaire supérieure.

$A^k = XD^kY = XD^kL_yU_y = X(D^kL_yD^{-k})D^kU_y$

$$(D^kL_yD^{-k})_{ij} = \begin{cases} 0 & si \;\; i<j \\ 1 & si \;\; i=j \\ (L_y)_{ij}(\frac{\lambda_i}{\lambda_j})^k & si \;\; i<j \end{cases}$$

$$lim(\frac{\lambda_i}{\lambda_j})^k = 0 \Rightarrow D^kL_yD^{-k} \rightarrow_{k\rightarrow+\infty} I$$

$\Rightarrow D^kL_yD^{-k} = I + E_k \;\; ou \;\; E_k \rightarrow_{k\rightarrow+\infty} 0$

$A^k = X(I+E_k)D^kU_y$

$X \;\; admet \; une \; d\acute{e}comosition \; QU \;\; ou \; U \in ou \; U \in (\sup, +)$

$X = Q_x.R_x$

$\Rightarrow A^k = Q_xR_x(I+E_k)D^kU_y$

$\quad\quad = Q_x(I+R_xE_kR_x^{-1})R_xD^kU_y$

$I + R_xE_kR_x^{-1} \rightarrow_{k\rightarrow+\infty} I$

$I + R_xE_kR_x^{-1} = \widehat{Q}_k\widehat{R}_k \rightarrow_{k\rightarrow+\infty} I \;\; et \; \widehat{Q}_k \rightarrow I$

$\quad\quad\quad\quad\quad\quad\quad\quad \widehat{R}_k \rightarrow I$

$$\begin{cases} A^k & = & (\widehat{Q}_x\widehat{Q}_k)\widehat{R}_kR_xD^kU_y \\ et & & \\ A^k & = & M_kS_k \end{cases} \Rightarrow \begin{cases} M_k & = & Q_x\widehat{Q}_k \\ S_k & = & \widehat{R}_kR_xD^kU_y \end{cases}$$

$donc \; M_k \rightarrow_{k\rightarrow+\infty} Q_x$

$on \; a \; A_{k+1} = M_k^TAM_k \Rightarrow A_k \rightarrow_{k\rightarrow+\infty} Q_x^TAQ_x$

$or \; Q_x^TAQ_x = Q_x^TXDX^{-1}Q_x$

$\quad\quad = Q_x^TQ_xR_xDR_x^{-1}Q_x^TQ_x$

$$= R_xDR_x^{-1} = \begin{pmatrix} \lambda_1 & \\ 0 & \lambda_n \end{pmatrix}$$

Proposition :

Sous l'hypothèse que toutes les valeurs propres sont distinctes en module et que Y *admet la factorisation* LU *alors* A_x *converge vers* $R_xDR_x^{-1}$

b) cas ou Y *n'admet pas de factorisation* LU *il existe une matrice de permutation* P_y *telle que* $P_yY = L_yU_y$*, ou* $\overline{L_y} = P_y^TL_yP_y$ *est triangulaire inférieure avec 1 sur la diagonale.*

$A^k = XD^k y = XD^k P_y^T P_y Y = XD^k P_y^T L_y U_y$
$= XP_y^T P_y D^k P_y^T L_y U_y$

$= XP_y^T P_y D^k \overline{L_y} P_y^T U_y$

$= XP_y^T P_y (D^k \overline{L_y} D^{-k}) P_y^T P_y D^k P_y^T U_y$

$$D^k \overline{L_y} D^{-k} = \begin{cases} 0 & si \quad i < j \\ 1 & si \quad i = j \\ (\overline{L_y})_{ij} (\frac{\lambda_i}{\lambda_j})^k & si \quad i > j \end{cases}$$

$lim_{k \to +\infty} (\frac{\lambda_i}{\lambda_j})^k = 0 \quad et \;\; D^k \overline{L_y} D^{-k} \to I$

donc on peut écrire $P_y (D^k \overline{L_y} D^{-k}) P_y^T = I + E_k$ *avec* $E_k \rightarrow_{k \to +\infty} 0$
$A^k = XP_y^T (I + E_k) P_y D^k P_y^T U_y$
on pose $D_0 = P_y D^k P_y^T$

$$D_1 = \begin{pmatrix} \lambda_{i_1} & \\ & \lambda_{i_n} \end{pmatrix} \quad avec \quad (\lambda_{i_1}, ..., \lambda_{i_n}) = \{\lambda_1, ..., \lambda_n\}$$

$\Rightarrow A^k = XP_y^T (I + E_k) D_0^k U_y$
XP_y^T *admet une factorisation* QU *avec* $ut(\sup, +)$
$XP_y^T = Q_x R_x$
$A^x = Q_x R_x (I + E_k) D_0^k U_y$
$\quad = \; Q_x (I + R_x E_k R_x^{-1}) R_x D_0^k U_y$
$I + R_x E_k R_x^{-1} = \widehat{Q}_k \widehat{R}_k \rightarrow_{k \to +\infty} I \quad \Rightarrow \widehat{Q}_k \rightarrow_{k \to +\infty} I$
$\qquad\qquad\qquad\qquad\qquad\qquad \widehat{R}_k \rightarrow_{k \to +\infty} I$

donc

$$\begin{cases} A^k & = & (Q_x \widehat{Q}_k) \widehat{R}_k R_x D_0^k U_y \\ A^n & = & M_k S_k \end{cases} \Rightarrow \begin{cases} M_k & = & Q_x \widehat{Q}_k \\ S_k & = & \widehat{R}_k R_x D_0^k U_y \end{cases}$$

d'ou $M_k \rightarrow_{k \to +\infty} Q_x$
$\Rightarrow A_x \rightarrow_{k \to +\infty} Q_x^T A Q_x$

$or \quad Q_x^T A Q_x = Q_x^T (Q_x R_x P_y)^{-1}$

$$= R_x D_0 R_x^{-1} = \begin{pmatrix} \lambda_{i_1} & \\ & \lambda_{i_n} \end{pmatrix}$$

2) Cas ou quelques valeurs propres sont égales

$$a)\ on\ suppose \begin{cases} |\lambda_1| > |\lambda_2| ... > |\lambda_r| = |\lambda_{r+1}| = ... = |\lambda_p| > ... > |\lambda_n| \\ \lambda_r = ... = \lambda_p \\ et l'espace propre de \lambda_r est de dimension p - r + 1 \end{cases}$$

α) Supposons que y admet une factorisation LU :

$Y = L_y.U_y$

$A^x = XD^kY = XD^kL_yU_y = XD^kL_yD^{-k}D^kU_y$

$$(D^kL_yD^{-k})_{ij} \begin{cases} 0 & si \quad i < j \\ 1 & si \quad i = j \\ (L_y)_{ij}(\frac{\lambda_i}{\lambda_j})^k & si \quad i > j \end{cases}$$

. $Si\ i\ ou\ j \notin \{r, r+1, ..., r+p\}\ (L_y)_{ij}(\frac{\lambda_i}{\lambda_j})^k \rightarrow_{k \to +\infty} 0$

. $Si\ i\ et\ j \in \{r, r+1, ..., r+p\}\ (L_y)_{ij}(\frac{\lambda_i}{\lambda_j})^k = (L_y)_{ij}$

$$D^kL_yD^{-k} \rightarrow_{k \to +\infty} \widetilde{L} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{pmatrix}$$

$On\ pose\ E_k = D^kL_yD^{-k} - \widetilde{L}$

$E_k \rightarrow_{k \to +\infty} 0$

$A^k = X(\widetilde{L} + E_k)D^kU_y = X\widetilde{L}^{-1}(I + \widetilde{L}^{-1}E_k)D^kU_y$

$X\widetilde{L}$ *se décompose comme* $X\widetilde{L} = Q_xR_x$

$donc\ A^k = Q_xR_x(I + \widetilde{L}^{-1}E_k)D^kU_y$

$$= Q_x(I + R_x\widetilde{L}^{-1}E_kR_x^{-1})R_xD^kU_y$$

$$= (Q_x\widehat{Q}_k)\widehat{R}_kR_xD^kU_y \quad avec\ \widehat{R}_k \in (\sup, +)$$

$$\widehat{R}_k \rightarrow I \quad et\ \widehat{Q}_k \rightarrow I$$

d'autre part $A^k = M_k S_k$

donc $M_k = Q_x \widehat{Q}_k$ *et* $M_k \rightarrow_{k \rightarrow +\infty} Q_x$

d'ou $A_k \rightarrow_{k \rightarrow +\infty} Q_x^T A Q_x$

or $Q_x^T A Q_x = Q_x^T (Q_x R_x \widetilde{L}^{-1}) D (Q_x R_x \widetilde{L}^{-1}) Q_x$

$= R_x (\widetilde{L}^{-1} D \widetilde{L}) R_x^{-1}$ *est une matrice triangulaire supérieure de type* $\begin{pmatrix} \lambda_1 & \\ 0 & \end{pmatrix}$

B) Si Y *n'est pas factorisable sur la forme* $L.U$ *il existe une permutation* P_y *tq* $P_y = L_y U_y$ *la démonstration est sensiblement la même qu'avant.*