Etude expérimentale d'une méthode itérative de calcul des valeurs propres d'une matrices. Représentation graphique de la convergence de la méthode.

PRADALIER Cédric   ROBERT Audrey

Chapitre 1

Avant-propos

1.1 Materiel utilisé et choix logiciel

L'ensemble des résultats ont été calculé sur le sun pendant les accès en libre service grâce à des bibliothèques de calcul matriciel et de représentation graphique développée en Ada (et C pour les accès aux librairies X11). De plus, les tests statistiques ont été réalisés sur un PC personnel : un 486DX2 66MHz à partir de programmes écrits en Turbo Pascal 7.0.
Pour la rédaction du TP, nous avons utilisé LaTeX; nous avons de plus réalisé quelques graphiques statistiques sous Excel©

1.2 Méthode utilisée pour le calcul et le rendu des matrices

1.2.1 Décomposition QR : algorithme de Householder-Businger-Golub

On cherche tout d'abord une matrice H_1 orthogonale telle que

$$H_1 A = \begin{pmatrix} \alpha_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$$

Cette matrice est donnée par la relation suivante :

$$\begin{array}{rrcl} & H_1 & = & I - 2uu^T \\ \text{où} & u & = & C(A_1 - \alpha_1 e_1) \end{array}$$

où

$$\begin{cases} C \text{ est déterminé par } \|u\| = 1 \\ \alpha_1 = -sign(a_{1,1}) \cdot \|A1\|_2 \end{cases}$$

En appliquant ensuite récursivement cette methode a la matrice C, on obtient :

$$\underbrace{H_n \cdot \ldots \cdot H_1}_{Q^T} A = R$$

Ceci donne la décomposition :

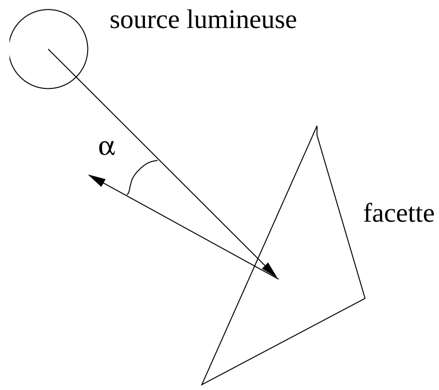
$$A = QR \tag{1.1}$$

A partir de cette décomposition, on peut appliquer la méthode décrite dans la section suivante. Le critére de convergence est basé sur un seuil : on considère que la matrice est diagonale lorsque tout ses coefficients non-diagonaux sont supérieurs à un seuil que l'on notera $\mathcal{S}$.

1.2.2 rendu des matrices

Les matrices sont représentées par des matrices de pyramides en 3 dimensions dont la hauteur est proportionnelle au coefficient correspondant de la matrice. Les pyramides sont representés grâce à 3 méthodes :

- Un tri des faces avant l'affichage qui permet de masquer les faces cachées.
- Le remplissage n'est fait que si la pyramide représente un nombre supérieur au seuil $\mathcal{S}$, ce qui permet de rendre transparentes les pyramides trop petites et de visualiser les pyramides négatives.
- Enfin, le choix de la couleur de remplissage est une couleur dont la luminosité est proportionnelle à l'angle α formé par le rayon lumineux venant d'une source lumineuse donnée et la normale à la face considérée.



Chapitre 2

Eléments théoriques

2.1 Première approche : décomposition QR sans translation

Soit n entier, dans tout ce qui va suivre, on étudiera des matrices carrées (n,n). On considère A_0 donné, et on construit de la manière suivante la suite de matrices $(A_n)_{n\geq 0}$:

$$(M)\left\{\begin{array}{rcl} A_r & = & Q_r R_r \\ A_{r+1} & = & R_r Q_r \end{array}\right.$$

2.1.1 Toute les matrices A_r sont semblables á A_0

En effet, on a pour tout r entier, Q_r orthogonale par définition de la décomposition QR. i.e.

$$Q^T.Q = I_n$$

d'où , au rang 0, On a bien A_0 semblable à A_0!
Supposons que, pour p entier fixé, on ait, pour tout $j \leq p$, A_j semblable à A_0, alors

$$\begin{array}{rcl} A_{p+1} & = & R_p.Q_p \\ \Rightarrow A_{p+1} & = & Q_p^T.Q_p.R_p.Q_p \\ \Rightarrow A_{p+1} & = & Q_p^T.A_p.Q_p \end{array}$$

Or, $A_p = S^T.A_0.S$ avec S orthogonale, par hypothèse.
D'où, $A_{p+1} = (S.Q_p)^T.A_0.S.Q_p$, et A_{p+1} semblable à A_0.
Par récurrence, on a donc $\forall p, A_p$ est semblable à A_0. Les matrices $(A_n)_{n\geq 0}$ ont donc toutes les mêmes valeurs propres, donc si $(A_n)_{n\geq 0}$ converge vers une matrice diagonale, on aura sur la diagonale de cette matrice les valeurs propres de A_0.

2.1.2 Si A_0 symétrique réelle, $(A_n)_{n\geq 0}$ est une suite de matrices symétriques réelles.

Tout d'abord, il est clair que $\forall r$ entier, si A_r est réelle, alors A_{r+1} l'est aussi donc par récurrence, $(A_n)_{n\geq 0}$ est une suite de matrices symétriques réelles.
Pour la symétrie, on sait que A_0 est symétrique donc :

$$\forall (X,Y)\ X^T A_0 Y = Y^T A_0 X$$

De plus, $\forall r$, A_r est semblable à A_0 donc :

$$\forall r, \exists P_r \text{ orthogonale telle que } A_r = P_r^T A_0 P_r$$

On a donc :

$$\begin{array}{rcll} \forall (X,Y),\ X^T A_r Y & = & X^T P_r^T A_0 P_r Y & \\ & = & (P_r X)^T A_0 (P_r Y) & \\ & = & (P_r Y)^T A_0 (P_r X) & \text{Par symtrie de } A_0 \\ & = & Y^T P_r^T A_0 P_r X & \\ & = & Y^T A_r X & \end{array}$$

D'où pour tout r, A_r est symétrique.

2.1.3 Relation entre $\mathcal{Q}_r$, $\mathcal{R}_r$ et $(A_n)_{n\geq 0}$

On pose :

$$\begin{cases} \mathcal{Q}_r &= Q_0 \dots Q_r \\ \mathcal{R}_r &= R_r \dots Q_0 \end{cases}$$

On veut montrer :

$$\begin{aligned} \mathcal{Q}_r A_{r+1} &= A_0 \mathcal{Q}_\nabla \\ \mathcal{Q}_r \mathcal{R}_r &= A_0^{r+1} \end{aligned}$$

Tout d'abord, on a :

$$\begin{aligned} \mathcal{Q}_0 A_1 &= Q_0 A_1 \\ &= Q_0 R_0 Q_0 \\ &= A_0 \mathcal{Q}_0 \end{aligned}$$

Et ,

$$\begin{aligned} \mathcal{Q}_0 \mathcal{R}_0 &= Q_0 R_0 \\ &= A_0 \\ &= A_0^1 \end{aligned}$$

La relation est donc bien vérifiée au rang 0.
Soit r entier, supposons que pour tout $k < r$, la propriété soit vérifiée. Alors :

$$\begin{aligned} \mathcal{Q}_r A_{r+1} &= \mathcal{Q}_{r-1} Q_r R_r Q_r \\ &= \mathcal{Q}_{r-1} A_r Q_r \\ &= A_0 \mathcal{Q}_{r-1} Q_r \\ &= A_0 \mathcal{A}_r \qquad \text{Par hypothèse de récurrence} \end{aligned}$$

De même :

$$\begin{aligned} \mathcal{Q}_r \mathcal{R}_r &= \mathcal{Q}_{r-1} Q_r R_r \mathcal{R}_{r-1} \\ &= \mathcal{Q}_{r-1} A_r \mathcal{R}_{r-1} \\ &= A_0 \mathcal{Q}_{r-1} \mathcal{R}_{r-1} \\ &= A_0 A_0^r \qquad \text{Par hypothèse de récurrence} \\ &= A_0^{r+1} \end{aligned}$$

On a donc bien par recurrence le résultat demandé.

2.2 Seconde approche : décomposition QR avec translation

Soit n entier. Toutes les matrices éudiées seront de taille nxn.

Soit A_0 une matrice réele symétrique.
On crée une suite de matrices $(A_n)_{n\geq 0}$ de la manère suivante :

$$\begin{cases} A_r - \alpha_r I &= Q_r . R_r \\ A - r + 1 &= R_r . Q_r + \alpha_r . I \end{cases}$$

où α_r est un réel à choisir à chaque pas (on retrouve (M) pour $\alpha_r = 0$).

2.2.1 Toute les matrices A_r sont semblables á A_0

En effet Q_r étant orthogonale pour tout r, par définition de la décomposition QR, on a Q_r inversible et $Q_r^{-1} = Q_r^T$.

Au rang 0, on a bien A_0 semblable à A_0!

Supposons que, pour p entier fixé, on ait, pour tout $j \leq p$, A_j semblable à A_0, alors

$$\begin{aligned} A_{p+1} &= R_p.Q_p + \alpha_p.I \\ \Rightarrow A_{p+1} &= (Q_p^T.A_p - \alpha_p.Q_0^T).Q_p + \alpha_p.I \\ \Rightarrow A_{p+1} &= Q_p^T.A_p.Q_p \end{aligned}$$

Or, il existe S orthogonale telle que $A_p = S^T.A_0.S$, par hypothèse.
D'où, $A_{p+1} = (S.Q_p)^T.A_0.S.Q_p$, et A_{p+1} semblable à A_0.
Par récurrence, on a donc $\forall p$,A_p est semblable à A_0.

2.2.2 Si A_0 symétrique réelle, $(A_n)_{n\geq 0}$ est une suite de matrices symétriques réelles.

Au rang 0, on a bien A_0 matrice symétrique.

Supposons que, pour p entier fixé, on ait, pour tout $j \leq p$, A_j symétrique.
Alors, étant donné que A_{p+1} est semblable à A_0, il existe P orthogonale telle que $A_{p+1} = P^T A_0 P$.
Donc, $A_{p+1}^T = P^T A_0^T P$.
Or $A_0^T = A_0$, d'où $A_{p+1}^T = A_{p+1}$.
La propriété étant vérifiée au rang $p+1$,elle est vraie pour tout positif.

Chapitre 3

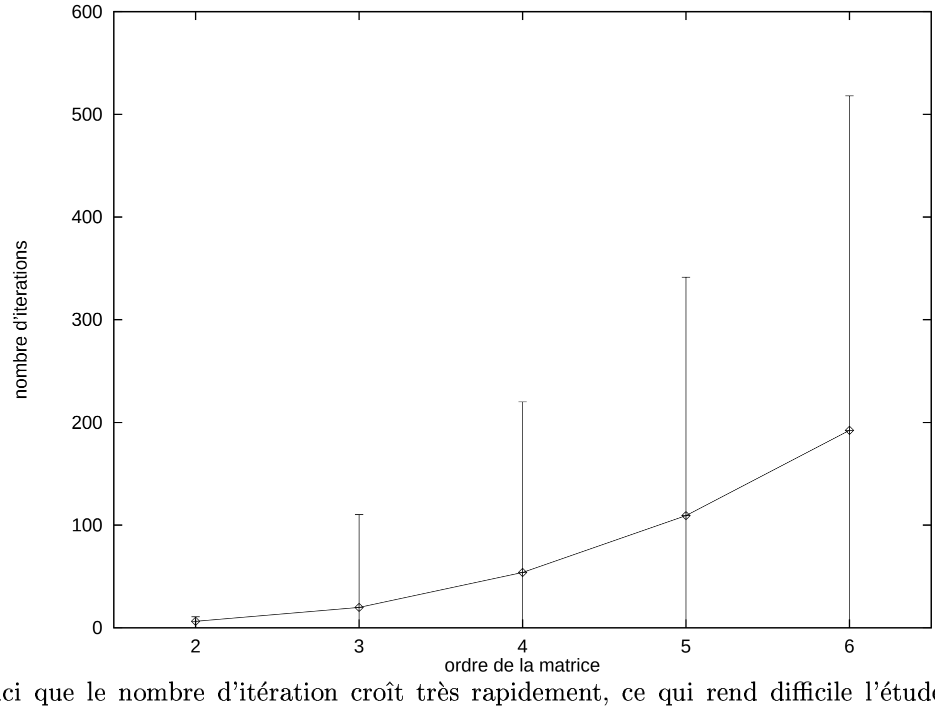
Etude statistiques de la convergence

3.1 Méthode de base

3.1.1 Nombre moyen d'itérations pour atteindre la convergence

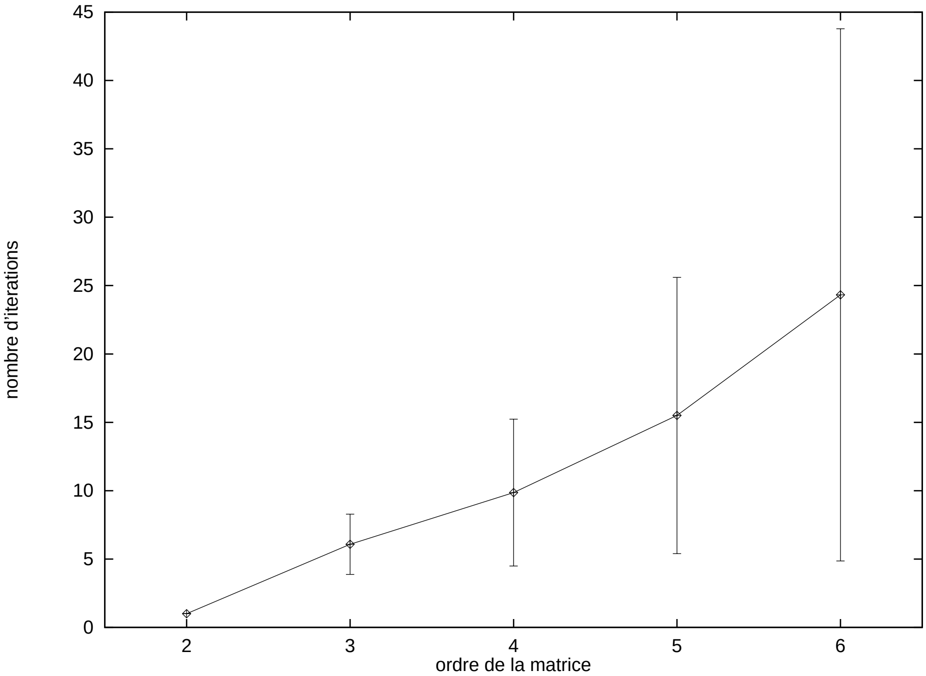
Une première approche consiste à prendre comme critère de convergence, la norme infinie de la matrice privée de sa diagonale : Si $A = D + M$, avec D diagonale alors on considère que A est diagonale si $\|M\|_\infty < \mathcal{S}$. On prend ici $\mathcal{S} = \infty /^{-\ni}$ et des matrices dont les coefficients sont gˆenérés uniformément dans $[-1{,}1]$
Avec ce critère, on obtient les moyennes suivantes :

Taille de la matrice	Nombre moyen d'itération	Ecart-type	Nombre d'experience
2	6.428	4.362	837
3	19.90	89.37	610
4	53.86	166.1	213
5	109.3	231.4	267
6	192.3	325.7	86



On constate ici que le nombre d'itération croît très rapidement, ce qui rend difficile l'étude statistique sur des matrices de grande taille. Cependant, on constate en représentant les matrices que l'algorithme converge très rapidement vers des matrices diagonales par blocs : ces blocs étant de dimension 2x2, on peut calculer théoriquement les valeurs propres, il n'est donc pas forcément nécessaire de continuer l'algorithme.
Avec ce critère d'arrêt, on obtient :

Taille de la matrice	Nombre moyen d'itération	Ecart-type	Nombre d'experience
2	1.000	0.000	837
3	6.081	2.198	533
4	9.861	5.365	531
5	15.5	10.1	4608
6	24.32	19.46	151



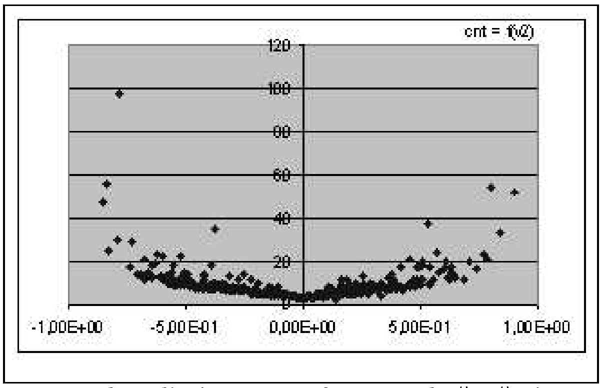
Il apparaît nettement que ce critère entraîne une convergence plus rapide et une croissance moindre. En effet, en évaluant rapidement l'évolution du nombre d'itérations, on obtient :

ordre de la matrice :	$3 = ordre0$	$\frac{5}{3} \cdot ordre0$	$2 \cdot ordre0$
nb itér : critère 1	20	$109 \approx \frac{5}{3}^{3.3} \cdot 20$	$192 \approx 2^{3.3} \cdot 20$
nb itér : critère 2	6	$15.5\ approx \frac{5}{3}^{2} \cdot 6$	$24.3 \approx 2^2 \cdot 6$

Une première approximation donnerait pour la première méthode une convergence en $O(n^{3.3})$ et en $O(n^2)$ pour le second critère.

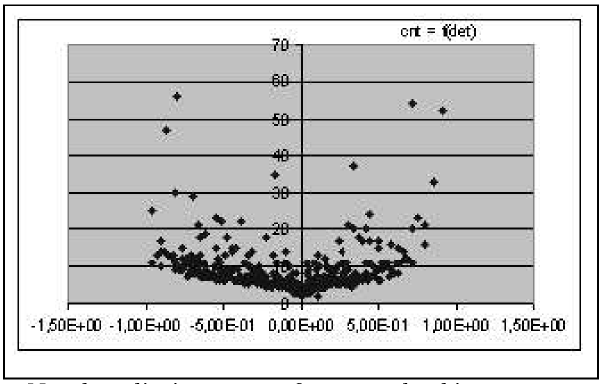
3.1.2 Facteurs de convergence

La convergence des matrices 2x2 étant très rapide et le calcul du déterminant très facile, on peut essayer d'étudier les relations entre les grandeurs caractéristiques des matrices : trace, déterminant, valeurs propres, coefficients, écart en valeur absolue des valeurs propres... et le nombre d'itérations pour atteindre le critère d'arrêt (on utilise ici le critère de diagonalité réelle).
On obtient alors les résultats suivants :



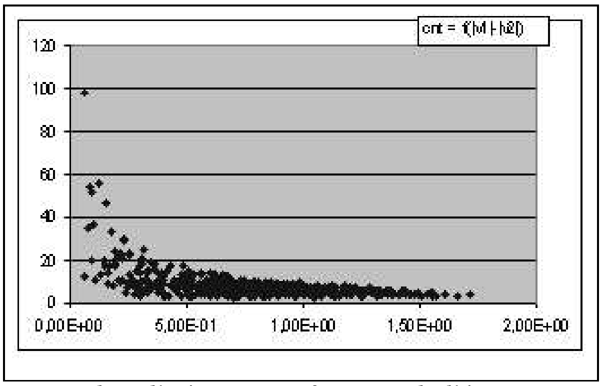
Nombre d'itération en fonction de $||v2||$ où
v2 est la plus petite valeur propre de la matrice.

On constate ici que plus $||v2||$ est proche de zéro plus la convergence est rapide.



Nombre d'itération en fonction du déterminant
de la matrice.

Plus le déterminant est petit (proche de zéro), plus la convergence est rapide. On pourrait donc essayer de diviser la matrice par $\frac{1}{10}$, ce qui correspondrait à diviser son déterminant par $\frac{1}{10^n}$. Cependant, cela revient aussi à rapprocher tout les coefficients de la matrice du seuil $\mathcal{S}$. la convergence est plus rapide mais on perd en précision : en moyenne, **l'erreur est multipliée par 100**, mais sur certain cas, on passe d'une erreur sur la $8^{ème}$ décimale par rapport à la vraie valeur propre à une erreur sur la $1^{ère}$ décimale.



nombre d'itération en fonction de l'écart en norme des valeurs propres.

Plus l'écart entre les deux valeurs propres est grand, plus la convergence est longue.