Incertitudes et calcul de valeurs propres de matrices

Bernard Philippe

INRIA-IRISA
Rennes, FRANCE

FIABLE : Action coopérative de l'Inria

Paris, 22-24 septembre 1999

Plan

- Introduction
- Quelques ennuis classiques dans le calcul de valeurs propres
- Conditionnements d'une valeur propre et d'un vecteur propre
- Les pseudo-spectres
- Dénombrement de valeurs propres
- Arithmétique d'intervalles
- Conclusion

Exemples de localisation de valeurs propres

Les études de stabilité de dispositifs ou de procédés physiques conduisent souvent à la localisation de valeurs propres :

– Le système différentiel :

$$\frac{dx}{dt}(t) = A\ x(t) + b$$

est stable lorsque $\sigma(A) \subset \mathbb{C}_-$

– L'itération vectorielle :

$$x_{n+1} = Ax_n + b$$

est stable lorsque $\sigma(A) \subset D(O,1)$

On recherche des réponses sûres

Même LAPACK peut échouer !

Exemple dans Matlab (gallery(5)) :

$$A = \begin{pmatrix} -9 & 11 & -21 & 63 & -252 \\ 70 & -69 & 141 & -421 & 1684 \\ -575 & 575 & -1149 & 3451 & -13801 \\ 3891 & -3891 & 7782 & -23345 & 93365 \\ 1024 & -1024 & 2048 & -6144 & 24572 \end{pmatrix}$$

est nilpotente : $A^5 = 0$ et pourtant :

```
d=eig(A)
d = -0.0337 + 0.0250i ; -0.0337 - 0.0250i ; 0.0134 + 0.0389i ;
0.0134 - 0.0389i ; 0.0407
```

La bibliothèque Lapack est la bibliothèque la plus fiable pour le calcul des valeurs propres de matrices pleines.

Mais les ennuis sont plus fréquents avec les méthodes adaptées aux grandes matrices

Pour les grandes matrices : Arnoldi avec transformation spectrale.

Par exemple, pour le problème généralisé

$$Ax = \lambda Bx$$

rechercher les valeurs propres voisines d'un nombre complexe donné σ est équivalent à rechercher les valeurs propres de plus grand module du problème standard

$$(A - \sigma B)^{-1} Bx = \frac{1}{\lambda - \sigma} x.$$

L'itération d'Arnoldi

Choose an orthonormal set of vectors $V_1 = [v_1,v_2,\ldots,v_l]$ for the initial block and a dimension $m \ll n$: order of the problem

for $j = 1, \cdots m$, **do**
1. $W_j = BV_j$
2. solve the systems $(A - \sigma B)U_j = W_j$
3. **for** $i = 1, \ldots j$ **do**
 $H_{ij} = V_i^H U_j$
 $U_j = U_j - V_i H_{ij}$
 end for
4. $Q_j R_j = U_j \quad$ *QR decomposition*
5. $V_{j+1} = Q_j \quad ; \quad H_{j+1,j} = R_j$

end for

Normalement les valeurs propres sont obtenues dans l'ordre des modules décroissants, mais...

la matrice

$$A = \begin{pmatrix} 1+i & 0 & 0 & -2i & -7+2i & 0 \\ 0 & 4+2i & 0 & -1 & 0 & 0 \\ 3-5i & 0 & 1+3i & 0 & 0 & -1+7i \\ -1+7i & 0 & 1+2i & 0 & 3+4i & 0 \\ 0 & -3+7i & 0 & 2+9i & 0 & 8+4i \\ 0 & -11+4i & 0 & -7i & 9 & 0 \end{pmatrix}$$

avec une taille de bloc égale à 2, une dimension de base égale à 4, pour le bloc initial

$$u = \begin{pmatrix} 0.1551 & 0.0333 & 0.4809 & 0.4812 & 0.6621 & 0.2716 \\ 0.3730 & 0.7293 & -0.2654 & -0.2486 & 0.0696 & 0.4381 \end{pmatrix}^H$$

et après 49 redémarrages, la deuxième et la troisième valeurs propres sont calculées à une précision de 10^{-8}. La première est manquante...

Diagonaliser est souvent dangereux

Numériquement, toute matrice est presque sûrement diagonalisable et même à valeurs propres distinctes,

mais lorsque la matrice est proche d'une matrice défective, les vecteurs propres sont presque dépendants.

Il peut alors être dangereux de l'utiliser pour un autre calcul :

Calcul non sûr d'exponentielle

par exemple : $A = XDX^{-1} \Rightarrow \exp(A) = X \exp(D) X^{-1}$

```
n=10; rand('state',0); X=ones(n)-2*rand(n) ;
v=10*ones(1,n-1);
A= X*(diag([1:10])+diag(v,1))/X ;
(Z,D) = eig(A);
Ecalc = Z*diag(exp(diag(D)))  Z ;
Eex = expm(A) ;
errel=norm(Ecalc-Eex)/norm(Eex)
```

donne une erreur de : 0.6 (aucun chiffre significatif).

Remarque : Il faut considérer la factorisation de Schur (A est orthogonalement semblable à une matrice quasi-triangulaire)

Stabilité asymptotique $\neq$ stabilité

Le système différentiel :

$$\frac{dx}{dt} = \begin{pmatrix} -1 & 10 & 0 & 0 & 0 \\ 0 & -1.1 & 10 & 0 & 0 \\ 0 & 0 & -1.2 & 10 & 0 \\ 0 & 0 & 0 & -1.3 & 10 \\ 0 & 0 & 0 & 0 & -1.4 \end{pmatrix} x$$

est asymptotiquement stable, mais $\|x(3.5)\| \approx 10^3$ (C.I : $x(0) = (0,0,0,0,1)^t$).

Si la matrice n'est pas exactement connue

La matrice $A = X * D * X^{-1}$

où $D = diag([0.999, 0.99, 0.9, 0.5, 0.1])$
et X définie par $X_{ij} = 1/(i+j)$

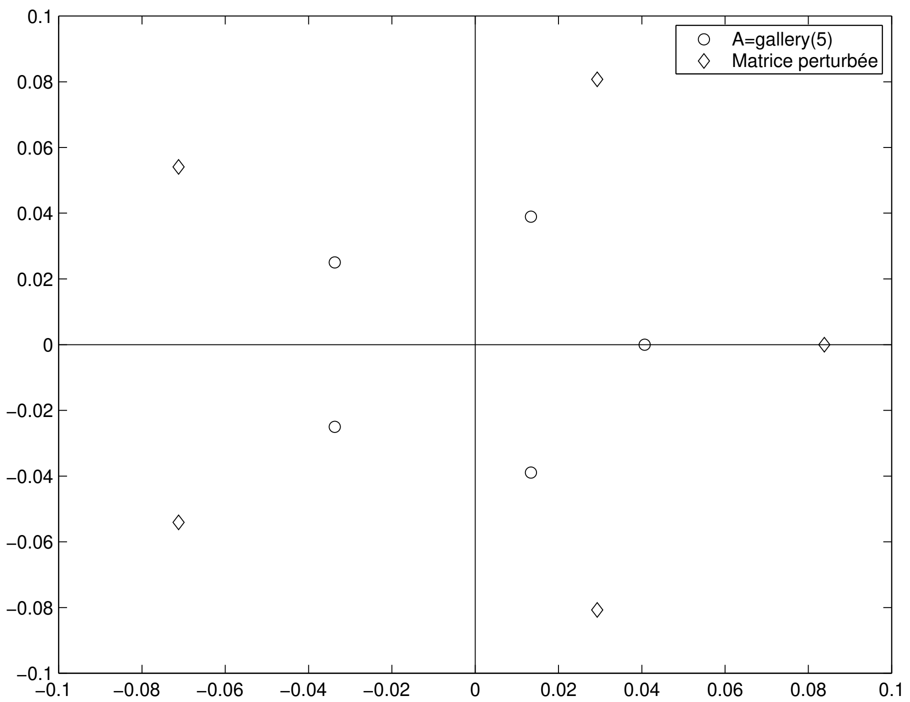
est stable : $\lim_{k->+\infty} A^k = 0$.

Une erreur relative de 2×10^{-4} sur A_{15} donne comme valeurs propres :
$[12.36, \ -9.423, \ -1.503, 1.057, 1.001]$:

La matrice perturbée est instable : $\lim_{k->+\infty} \|\tilde{A}^k\| = \infty$.

Deuxième exemple de perturbation de spectre

```
» A=gallery(5) ; P=ones(5)-2*rand(5);
» tol=eps*norm(A);
» D1=sort(eig(A)); D2=sort(eig(A+tol*P));
```



Perturbation des valeurs propres de gallery(5)

Incertitude contrôlée par le résidu

Théorème de Bauer-Fike : Soit $A = XDX^{-1}$ **diagonalisable** avec $(\tilde{\lambda},\tilde{u})$ une paire approchée de valeur et vecteur propres dont le résidu est : $r = A\tilde{u} - \tilde{\lambda}\tilde{u}$.

Alors il existe une valeur propre λ de A telle que

$$|\lambda - \tilde{\lambda}| \ \leq \ \chi_2(X) \ \|r\|_2$$

.

Cas hermitien : Soit A hermitienne avec $(\tilde{\lambda},\tilde{u})$ une paire approchée de valeur et vecteur propres dont le résidu est : $r = A\tilde{u} - \tilde{\lambda}\tilde{u}$.

Alors il existe une valeur propre λ de A telle que

$$|\lambda - \tilde{\lambda}| \ \leq \ \|r\|_2$$

.

Théorèmes généraux de perturbation

Continuité : Soit λ une valeur propre de A de multiplicité algébrique m.
Alors pour toute norme et $\epsilon > 0$ assez petit, il existe $\delta > 0$ tel que

$\|E\| < \delta \Rightarrow D(\lambda,\epsilon)$ contient exactement m valeurs propres de $A + E$.

Différentiabilité : Soit λ une **valeur propre simple** de A de vecteurs propres gauche y et droite x et E une perturbation de A.
Alors il existe une unique valeur propre $\tilde{\lambda}$ de $A + E$ telle que

$$\tilde{\lambda} = \lambda + \frac{y^H E x}{y^H x} + O(\|E\|^2).$$

Perturbation du vecteur propre

Différentiabilité : Soit (λ,u) un couple propre simple de la matrice A.

Soit P le projecteur spectral associé à λ et $S(\lambda) = (A - \lambda I)^{(*)}$ la résolvente généralisée en λ.

On considère la matrice $A(t) = A + tE$ pour $t \in [0,\delta]$ et le couple propre $(\lambda(t),u(t))$ tel que $\lambda(0) = \lambda$, $u(0) = u$ et $Pu(t) = u$ (condition de normalisation).

Alors

$$\frac{du}{dt}(0) = -S(\lambda)Eu$$

Définitions des conditionnements

Conditionnement d'une valeur propre simple :

$$Cond(\lambda) = \frac{1}{\cos(x,y)}$$

Conditionnement du vecteur propre associé :

$$Cond(u) = \|S(\lambda)\|.$$

Si la matrice est hermitienne :

$$Cond(\lambda) = 1 \text{ et } Cond(u) = \frac{1}{dist(\lambda,\ \sigma(A) - \{\lambda\})}.$$

Dans la bibliothèque LAPACK

Consulter le paragraphe : Accuracy and Stability à l'adresse :

`http://www.netlib.org/lapack/lug/node5.html`

- Error Bounds for the Symmetric Eigenproblem
- Error Bounds for the Nonsymmetric Eigenproblem (Overview, Balancing and Conditioning, Computing s and sep)
- Error Bounds for the Generalized Symmetric Definite Eigenproblem
- Error Bounds for the Generalized Nonsymmetric Eigenproblem

Voisinage de matrices défini par une norme matricielle

Les perturbations de A sont définies par $A + \Delta$ avec

$$\Delta \in \{\Delta \mid \ \|\Delta\| \leq \epsilon \|A\|,\}$$

où ϵ est un paramètre de précision.

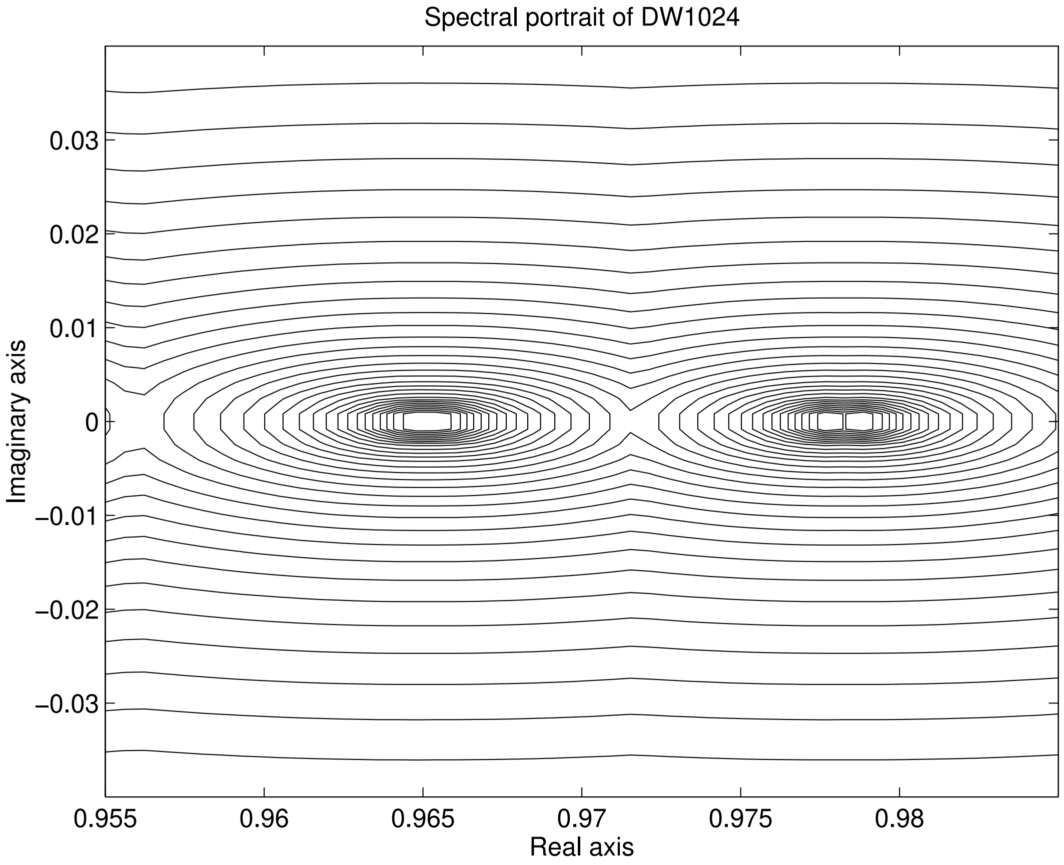
On définit un pseudo-spectre (Trefethen, Godunov) par

$$\Lambda_\epsilon(A) = \{z \in \mathbb{C} : z \text{ est une valeur propre de } A + \Delta; \ \|\Delta\| \leq \epsilon \|A\|\}$$

Sa construction repose sur la propriété :

$$(\ \exists \Delta \quad \|\Delta\| = \tau, \quad \lambda \in \Lambda(A + \Delta)\ ) \Longleftrightarrow \sigma_{\text{min}}(A - \lambda I) \leq \tau.$$

Exemple de pseudo-spectre



D'après la caractérisation, un portrait spectral est construit en déterminant les lignes de niveaux de la fonction

$$f : z \longrightarrow \|(A - zI)^{-1}\| = 1/\sigma_{\min}(A - zI)$$

La plus simple procédure consiste à calculer une Décomposition aux Valeurs Singulières (SVD) en tous les points d'une grille discrétisant la région considérée dans le plan complexe.

Le coût de calcul devient prohibitif pour les grandes matrices :

Une SVD complexe $\approx 40\ n^3$ flops ; une petite grille : 50×50 ...

Deux directions possibles :

- Parallélisation des méthodes (Projets Européens PORTRAIT et STABLE)
- Méthodes de suivi des lignes de niveaux par homotopie (Bruel, Beckas & Gallopoulos)

Méthode parallèle pour une matrice pleine (sur une grille)

Le meilleur algorithme est celui qui consiste à :

- réduire unitairement la matrice A sous forme de Hessenberg H
- pour chaque point z, calculer la valeur propre dominante de
$$\tilde{B}(z) = \begin{pmatrix} 0 & (H - zI)^{-H} \\ (H - zI)^{-1} & 0 \end{pmatrix}$$
par l'algorithme de Lanczos.

Idée : Toh-Trefethen ; parallélisation : Robert
Code existant dans le style SCALAPACK.

Performances : pour une matrice (Tolosa) de dimension 1000, sur 35 processeurs de la Paragon : 150 s pour 50 points. Un a deux ordres de grandeur d'accélération / procédure classique de SVD complète parallélisée.

Méthode parallèle pour une matrice pleine (suivi de lignes de niveaux)

Deux phases : Prédiction - Correction

Parallélisation : COBRA

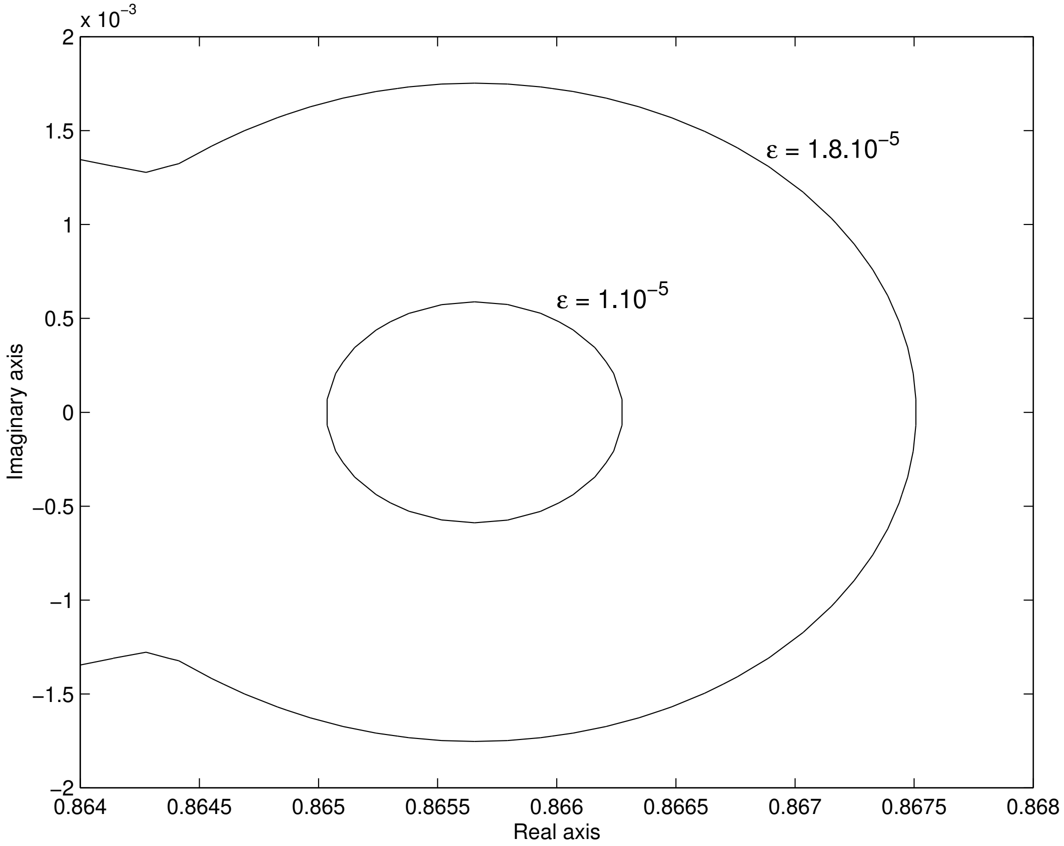
Davidson pour une matrice creuse (sur une grille)

$A \in \mathbb{C}^{n*n}$ and l denotes the block-size.
Choose an initial orthonormal matrix $V_1 = [v_1, \cdots, v_l] \in \mathbb{C}^{n*l}$;
For $k = 1, \cdots$

1: Compute the matrix $U_k = AV_k$;
2: Compute the matrix $W_k = A^H U_k$;
3: Compute the Rayleigh matrix $H_k = V_k^H W_k$;
4: Compute the l smallest eigenpairs $(\nu_{k,i}^2, y_{k,i})_{1 \le i \le l}$ of H_k;
5: Compute the vectors $x_{k,i} = V_k y_{k,i} \quad i = 1, \cdots, l$;
Let $X_k = [x_{k,1}, \cdots, x_{k,l}]$;
6: Compute the residuals $r_{k,i} = W_k y_{k,i} - \nu_{k,i}^2 x_{k,i}$ for $i = 1, \cdots, l$;
If convergence then **exit**
7: Compute the new directions $t_{k,i} = C_{k,i}(r_{k,i}) \quad i = 1, \cdots, l$;
8: **If** $dim(V_k) \le m - l$
then $V_{k+1} = [V_k, t_{k,1}, \cdots, t_{k,l}]$
where $[t_{k,1}, \cdots, t_{k,l}] = MGS((I - V_k^H V_k)[t_{k,1}, \cdots, t_{k,l}])$;
else $V_{k+1} = [X_k, t_{k,1}, \cdots, t_{k,l}]$
where $[t_{k,1}, \cdots, t_{k,l}] = MGS((I - X_k^H X_k)[t_{k,1}, \cdots, t_{k,l}])$;

endfor

Pseudo-spectre de DW8192



Cette taille de matrice correspond à la limite actuelle de faisabilité : en certains points de la grille, la plus petite valeur singulière n'a pas été correctement calculée.

Analyse complexe

– Intégration sur $[z, z + \Delta z]$:

$$\frac{1}{2i\pi}\int_{z}^{z+\Delta z}\frac{\frac{d}{dz}f(z)}{f(z)}dz = \frac{1}{2i\pi}(G(f(z+\Delta z)) - G(f(z)))$$

où $G(z) = ln|z| + i\,Arg(z)$ est une détermination du logarithme complexe.

– L'integration sur une courbe fermée annule la partie réelle de G:

$$\frac{1}{2i\pi}\int_{z_0}^{z_{end}}\frac{\frac{d}{dz}f(z)}{f(z)}dz = \frac{1}{2i\pi}(Arg(f(z_{end})) - Arg(f(z_0)))$$

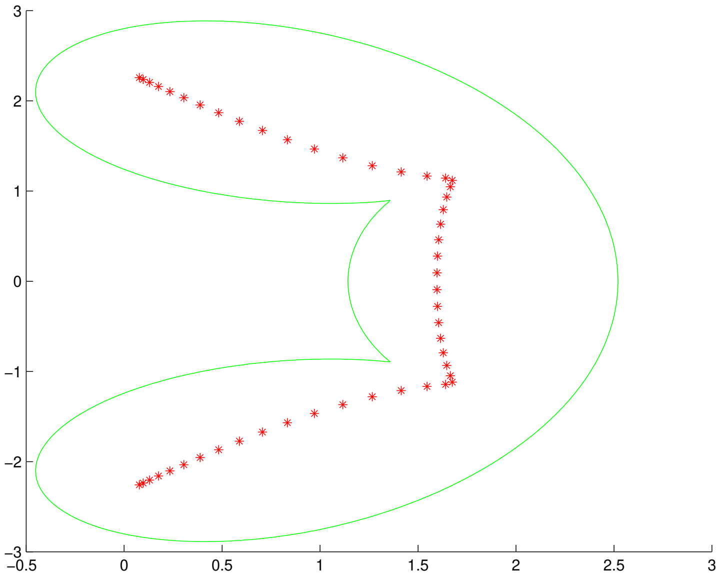
Quatre algorithmes programmés (O. Bertrand)

- par développement du logarithme
- intégration par méthode du trapèze à pas fixe (propriété de convergence exponentielle)
- intégration par méthode des trapèzes à pas variable (possibilité d'adaptation aux fortes variations)
- intégration par méthode d'Adams à pas variable (méthode d'ordre supérieur à la méthode des trapèzes)

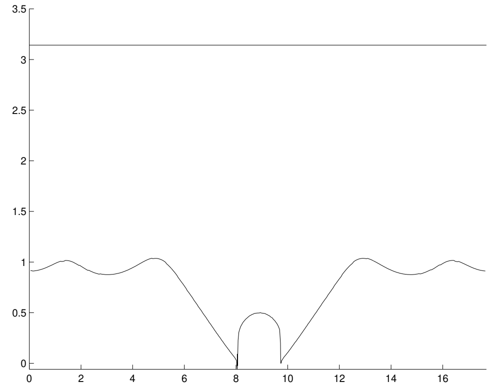
La première est adaptée au cas du suivi de lignes de niveau de la norme de la résolvante. Les autres, sont utilisables pour des courbes prédéfinies.

La limite des méthodes : calcul de l'argument du déterminant (factorisation LU).

Développement du logarithme



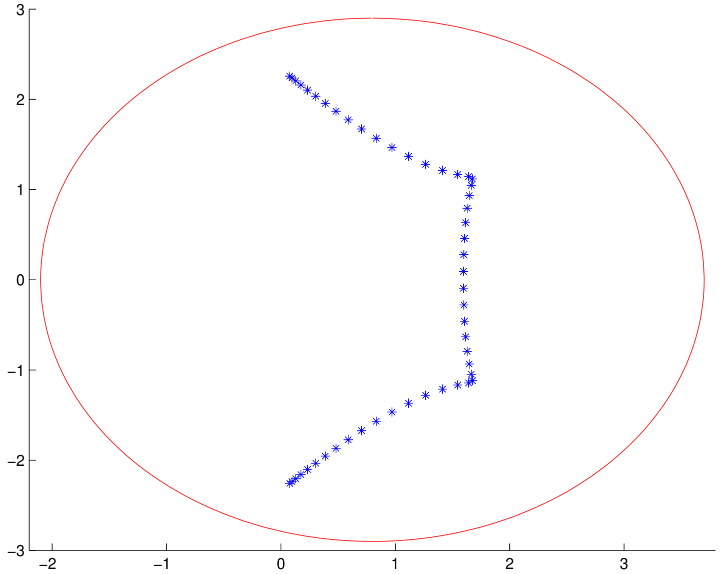
Suivi de contour de pseudo-spectre



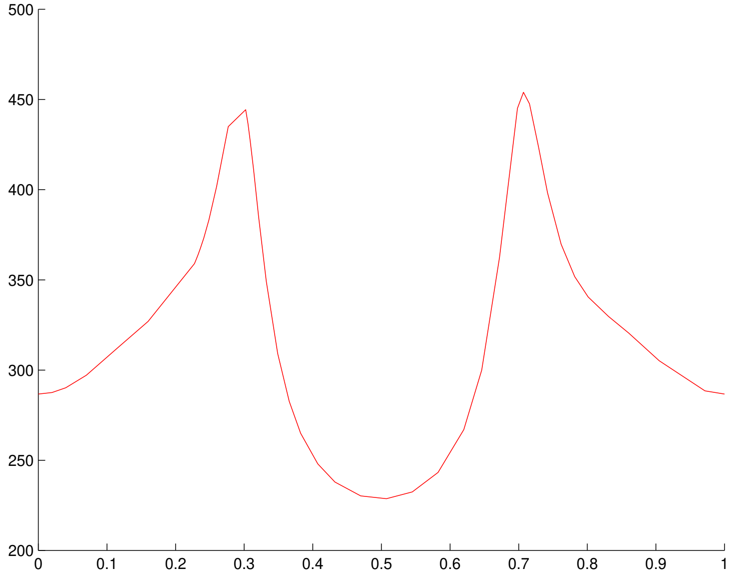
Evolution de la quantité qui doit rester inférieure à π

$\sigma_{min}(A - zI) = 3.10^{-2}$
424 étapes
424 évaluations de déterminants avec la matrice Grcar d'ordre 50.

Contour prédéfini entourant toutes les valeurs propres



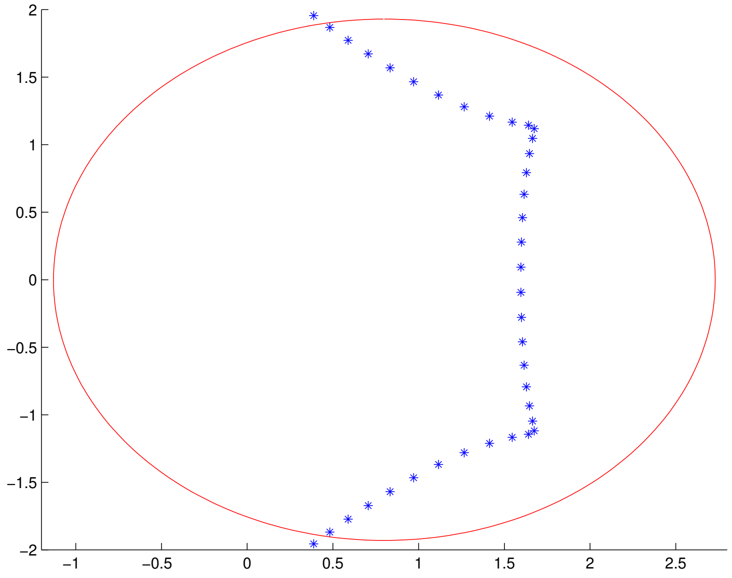
Contour



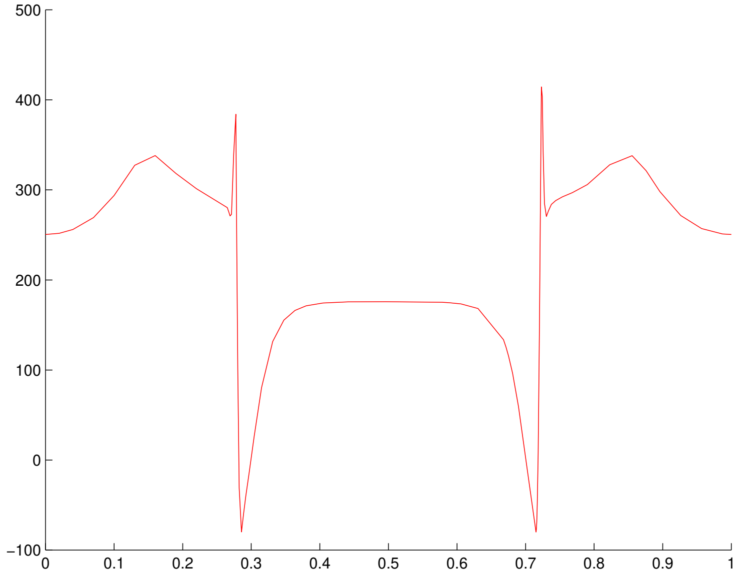
Fonction à intégrer

Trapèzes à pas constant : 32 pas - 64 déterminants
Trapèzes à pas variable : 45 pas - 120 déterminants
Méthode d'Adams (Ordre 4) : 59 pas - 128 déterminants

Courbe traversant le spectre



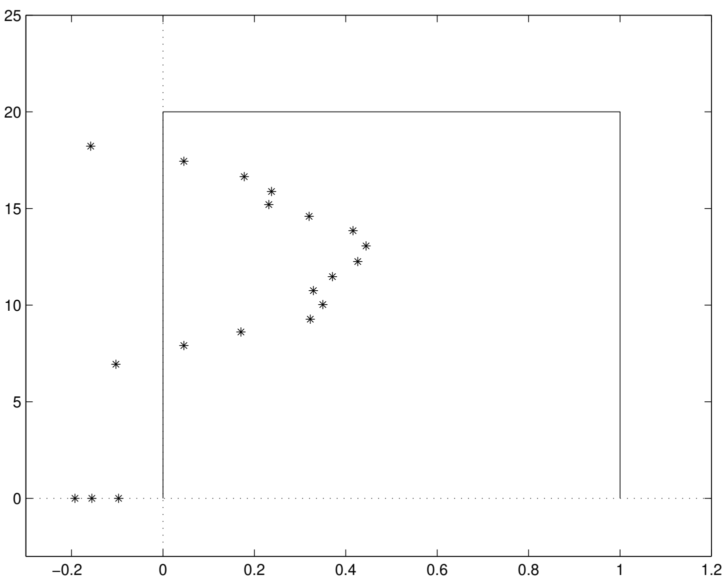
Contour



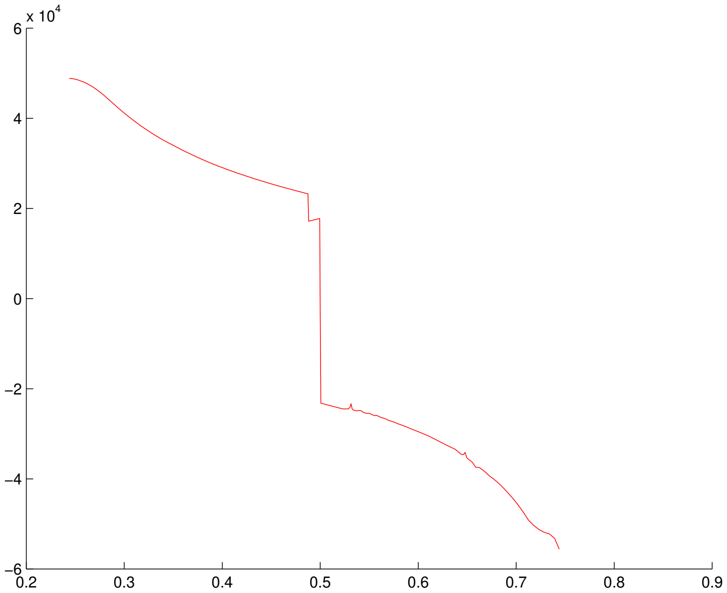
Fonction à intégrer

Trapèzes à pas constant : 512 pas - 1024 déterminants
Trapèzes à pas variable : 87 pas - 236 déterminants
Méthode d'Adams (Ordre 4) : 80 pas - 176 déterminants

Problème d'ordre 16002



Position des valeurs propres



Fonction à intégrer

Trapèzes à pas variable : 126 pas - 296 déterminants

Matrice issue de la stabilité d'un état stationnaire d'écoulement de fluide : 14 valeurs propres trouvées.

Définition d'un voisinage par composantes

Soit $\Delta A = \epsilon|A|$; l'ensemble des perturbations est alors :

$$\Delta \in [-\Delta A, \Delta A]$$

On recherche alors des algorithmes d'encadrement.

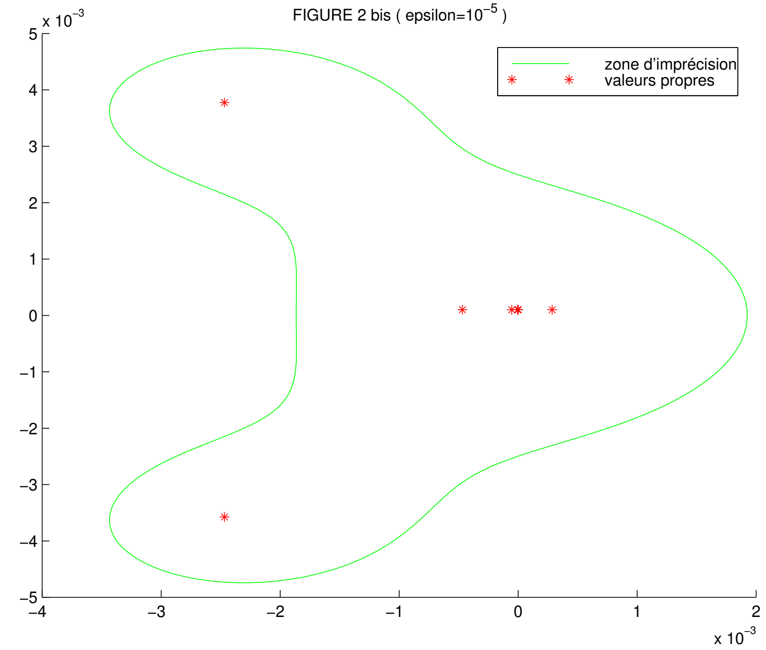
En général, ils consistent à raffiner un encadrement grossier.

Pour les valeurs propres, on ne sait le faire que si les valeurs propres sont assez séparées. Ce n'est plus vrai : voir exposé de S. Rump !

Exemple : Cinétique chimique

Pseudo-spectre autour de zéro, obtenu sur une matrice creuse d'ordre 18 (79 éléments non nuls). $\epsilon = 10^{-5}$.

(Provenance EDF)



Par les intervalles , on ne perturbe pas les composantes nulles et on ne considère que les perturbations réelles.
$\Rightarrow$ on trouve des encadrements 1000 fois plus petits.

Conclusion

La localisation des valeurs propres doit tenir compte de la sensibilité du problème aux perturbations.

Des outils mathématiques existent ; d'autres se développent encore (ex. dichotomie spectrale [Godunov-Malyshev-Sadkane]).

Des logiciels se développent ; le principal obstacle à leur emploi est leur coût en calcul.

On peut espérer des progrès dans ce domaine mais il faut aussi accepter de payer la fiabilité !