Estimations de la norme de l'erreur dans les méthodes de Krylov

Gérard MEURANT

Janvier 2011

1. Introduction
2. Le cas symétrique
3. Relation avec la discrétisation
4. CG en précision finie
5. Bibliographie CG
6. FOM et GMRES
7. Exemples FOM
8. Exemples GMRES
9. Conclusions

Introduction

Supposons que l'on modélise un phénomène physique par une équation aux dérivées partielles linéaire que l'on discrétise par différences finies ou éléments finis

Il existe plusieurs causes d'erreur :

1- Modélisation

2- Erreur de discrétisation

3- Erreur due au solveur linéaire

4- Erreurs d'arrondi

Cet exposé est centré sur le point 3, un petit peu 2 et une pincée de 4

On résout

$$Ax = b$$

par méthodes itératives de Krylov :

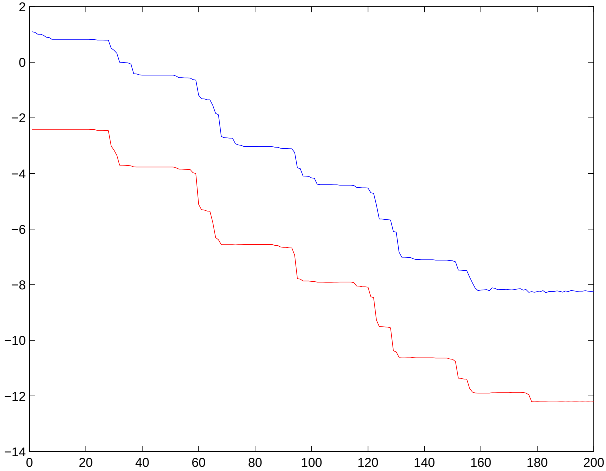
- Gradient Conjugué (CG) si A est symétrique définie positive
- FOM ou GMRES si A est non symétrique

La plupart des codes utilisent

$$\|r^k\| \leq \epsilon \|r^0\|$$

comme critère d'arrêt où r^k est le résidu (en théorie $r^k = b - Ax^k$)
D'autres utilisent la norme de l'erreur inverse (backward error)

MAIS, la norme du résidu peut être très différente de la norme de l'erreur. Exemple :



Steam2 : $n = 600, \kappa = 3.78\ 10^6$, GMRES, norme l_2,
Résidu, Erreur

Ce qu'on voudrait faire, c'est obtenir des estimations (ou des bornes) de la norme de l'erreur durant les itérations

$$Ax = b$$

Supposons que l'on connaisse une solution approchée $\hat{x}$
L'erreur est $e = x - \hat{x}$ et le résidu $r = b - A\hat{x}$, r est calculable, mais pas e
On a la relation fondamentale

$$Ae = A(x - \hat{x}) = b - A\hat{x} = r$$

Bien entendu, on ne veut pas résoudre ce système. Mais

$$\|e\|^2 = e^T e = (A^{-1}r)^T A^{-1} r = r^T A^{-2} r$$

Si A est définie positive on s'intéresse à $\|e\|_A^2 = e^T Ae$ et l'on a

$$\|e\|_A^2 = r^T A^{-1} r$$

Le cas symétrique

On s'intéresse à

$$u^T f(A) u$$

où u est un vecteur donné et f une fonction régulière ($1/x$ ou $1/x^2$)

$$A = Q \Lambda Q^T$$

avec Q orthonormale (vecteurs propres de A) et Λ diagonale (valeurs propres λ_i)

$$f(A) = Q\, f(\Lambda)\, Q^T$$

C'est, en fait, la définition de $f(A)$ lorsque A est symétrique
Bien sur, en général, on ne connaît pas Q et Λ

$$
\begin{aligned}
u^T f(A) u &= u^T Q f(\Lambda) Q^T u \\
&= \gamma^T f(\Lambda) \gamma \\
&= \sum_{i=1}^{n} f(\lambda_i) \gamma_i^2
\end{aligned}
$$

Cette somme peut être considérée comme une intégrale de Riemann–Stieltjes

$$
I[f] = u^T f(A) u = \int_a^c f(\lambda)\, d\alpha(\lambda)
$$

où la mesure α est constante par morceaux et définie par

$$
\alpha(\lambda) = \begin{cases} 0 & \text{if } \lambda < a = \lambda_1 \\ \sum_{j=1}^{i} \gamma_j^2 & \text{if } \lambda_i \leq \lambda < \lambda_{i+1} \\ \sum_{j=1}^{n} \gamma_j^2 & \text{if } c = \lambda_n \leq \lambda \end{cases}
$$

Quadratures de Gauss et variantes

L'idée est d'utiliser des quadratures de Gauss pour obtenir des bornes de l'intégrale

La formule générale pour une intégrale de Riemann–Stieltjes est

$$I[f] = \int_a^c f(\lambda)\ d\alpha(\lambda) = \sum_{j=1}^{N} w_j f(t_j) + \sum_{k=1}^{M} v_k f(z_k) + R[f],$$

où les poids $[w_j]_{j=1}^N, [v_k]_{k=1}^M$ et les nœuds $[t_j]_{j=1}^N$ sont inconnus et les nœuds $[z_k]_{k=1}^M$ sont prescrits

Voir Davis and Rabinowitz ; Gautschi ; Golub and Welsch

- Si $M = 0$, c'est Gauss sans nœuds prescrits
- Si $M = 1$ et $z_1 = a$ ou $z_1 = c$ c'est Gauss–Radau
- Si $M = 2$ et $z_1 = a, z_2 = c$, c'est Gauss–Lobatto

Le terme $R[f]$ est le reste. En général on ne peut pas le calculer, mais si la mesure α est positive non décroissante

$$R[f] = \frac{f^{(2N+M)}(\eta)}{(2N+M)!} \int_a^c \prod_{k=1}^{M} (\lambda - z_k) \left[\prod_{j=1}^{N} (\lambda - t_j) \right]^2 d\alpha(\lambda), \quad a < \eta < c$$

Pour la quadrature de Gauss ($M = 0$), le reste $R[f]$ à le signe de $f^{(2N)}(\eta)$

Voir Stoer and Bulirsch

Quadrature de Gauss

Comment calcule t-on les nœuds t_j et les poids w_j ?

On utilise les polynômes orthogonaux associés à la mesure α

$$\int_a^b p_i(\lambda)p_j(\lambda)\ d\alpha(\lambda) = \delta_{i,j}$$

Ces polynômes satisfont une récurrence à 3 termes

$$P(\lambda) = [p_0(\lambda)\ p_1(\lambda) \cdots\ p_{N-1}(\lambda)]^T, \quad e^N = (0\ 0\ \cdots\ 0\ 1)^T$$

$$\lambda P(\lambda) = J_N P(\lambda) + \gamma_N p_N(\lambda) e^N$$

$$J_N = \begin{pmatrix} \omega_1 & \gamma_1 & & & \\ \gamma_1 & \omega_2 & \gamma_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{N-2} & \omega_{N-1} & \gamma_{N-1} \\ & & & \gamma_{N-1} & \omega_N \end{pmatrix}$$

J_N est une matrice de Jacobi, ses valeurs propres sont réelles, simples et dans l'intervalle $[a, c]$

Les valeurs propres de J_N sont les nœuds t_j de la quadrature de Gauss

Les poids w_j sont les carrés des premiers éléments des vecteurs propres normalisés de J_N

Dans les cas qui nous intéressent, on connaît le signe des dérivées de f et on obtient des bornes pour $I[f]$

Pour Gauss-Radau et Gauss-Lobatto, on modifie des éléments de J_N pour avoir une ou deux valeurs propres prescrites

Calcul de l'intégrale

Il n'est pas toujours nécessaire de calculer les nœuds et les poids

On a

$$\sum_{l=1}^{N} w_l f(t_l) = (e^1)^T f(J_N) e^1$$

Pour $f(x) = 1/x$ il faut calculer

$$(J_N^{-1})_{1,1}$$

pour une matrice tridiagonale symétrique J_N

Comment calcule t-on la matrice de Jacobi (les coefficients de la récurrence des polynômes orthogonaux) pour la mesure α ?

La réponse est d'utiliser la méthode de Lanczos

La méthode de Lanczos

Partant d'un vecteur $\tilde{v}^1 = v/\|v\|$

$$\alpha_1 = (Av^1, v^1), \tilde{v}^2 = Av^1 - \alpha_1 v^1$$

et pour $k = 2, 3, \ldots$

$$\eta_{k-1} = \|\tilde{v}^k\|$$

$$v^k = \frac{\tilde{v}^k}{\eta_{k-1}}$$

$$\alpha_k = (v^k, Av^k) = (v^k)^T Av^k$$

$$\tilde{v}^{k+1} = Av^k - \alpha_k v^k - \eta_{k-1} v^{k-1}$$

Ca définit une matrice tridiagonale J_k

Soit $\chi_k(\lambda)$ le déterminant de $J_k - \lambda I$,

$$v^k = p_k(A)v^1, \quad p_k(\lambda) = (-1)^{k-1}\frac{\chi_{k-1}(\lambda)}{\eta_1 \cdots \eta_{k-1}}$$

Les polynômes p_k satisfont une récurrence à 3 termes

$$\eta_k p_{k+1}(\lambda) = (\lambda - \alpha_k)p_k(\lambda) - \eta_{k-1}p_{k-1}(\lambda), \;\; k = 1, 2, \ldots$$

avec $p_0 \equiv 0, \; p_1 \equiv 1$

Il existe une mesure β telle que

$$(v^k, v^l) = \langle p_k, p_l \rangle = \int_a^c p_k(\lambda)p_l(\lambda)d\beta(\lambda)$$

où $a \leq \lambda_1 = \lambda_{min}$ et $c \geq \lambda_n = \lambda_{max}$, λ_{min} et λ_{max} sont les plus petite et plus grande valeurs propres de A

Si $\hat{v} = Q^T v^1$

$$\alpha(\lambda) = \begin{cases} 0 & \text{si } \lambda < \lambda_1 \\ \sum_{j=1}^{i} [\hat{v}_j]^2 & \text{si } \lambda_i \leq \lambda < \lambda_{i+1} \\ \sum_{j=1}^{n} [\hat{v}_j]^2 & \text{if } \lambda_n \leq \lambda \end{cases}$$

Donc

$$\beta \equiv \alpha$$

Calcul de $u^T f(A) u$

- on normalise u si nécessaire $\rightarrow v^1$
- on fait k itérations de la méthode de Lanczos avec A en partant de v^1, ce qui donne la matrice de Jacobi J_k
- pour Gauss–Radau ou Gauss–Lobatto on modifie $J_k \rightarrow \tilde{J}_k$. Pour Gauss on a $\tilde{J}_k = J_k$
- si c'est possible, on calcule $(e^1)^T f(\tilde{J}_k) e^1$. Sinon, on calcule les valeurs propres et les premières composantes des vecteurs propres avec l'algorithme de Golub et Welsch pour obtenir les approximations données par les méthodes de quadrature

Si $\|u\| = 1$

$$u^T f(A) u = (e^1)^T f(J_n) e^1$$

et

$$R[f] = (e^1)^T f(J_n) e^1 - (e^1)^T f(J_k) e^1$$

La convergence de la quadrature de Gauss vers la valeur de l'intégrale dépend de la convergence des valeurs de Ritz (valeurs propres de J_k) vers les valeurs propres de A

Gradient Conjugué

Pour résoudre les problèmes SPD on utilise CG. Mais

$$\text{CG} \equiv \text{Lanczos}$$

Il serait stupide d'utiliser Lanczos à partir des itérés de CG

Que peut-on faire ?

La A–norme de l'erreur dans CG

Le carré de la A–norme de l'erreur à l'itération k est

$$\|\epsilon^k\|_A^2 = \|r^0\|^2[(J_n^{-1}e^1, e^1) - (J_k^{-1}e^1, e^1)]$$

où n est l'ordre de la matrice A et J_k est la matrice de Jacobi des coefficients de Lanczos qui peuvent être calculés à partir de ceux de CG. De plus

$$\|\epsilon^k\|_A^2 = \|r^0\|^2 \left[\sum_{j=1}^{n} \frac{[(z_{(n)}^j)_1]^2}{\lambda_j} - \sum_{j=1}^{k} \frac{[(z_{(k)}^j)_1]^2}{\theta_j^{(k)}}\right]$$

où $z_{(k)}^j$ est le $j$ième vecteur propre normalisé de J_k correspondant à la valeur propre $\theta_j^{(k)}$

Ces formules montrent le lien entre CG et les quadratures de Gauss

Le carré de la A–norme de l'erreur est le reste d'une quadrature de Gauss pour calculer l'intégrale $(A^{-1}r^0, r^0)$

Notons que la norme de l'erreur à l'itération k dépend de l'itération n. Ceci est du au fait que CG (et les méthodes de Krylov en général) sont des méthodes directes utilisées comme méthodes itératives
Dans CG il n'y a pas de notion de vitesse asymptotique de convergence. La vitesse de convergence dépend de la convergence des valeurs de Ritz

Comment utilise t-on cette formule ?

A l'itération k de CG on ne connaît pas $(J_n^{-1})_{1,1}$!

L'avenir est inconnu, mais le présent est l'avenir du passé !

Soit d un entier (> 0) donné, l'approximation de la A–norme de l'erreur à l'itération $k-d$ est donnée par

$$\|\epsilon^{k-d}\|_A^2 \approx \|r^0\|^2((J_k^{-1})_{(1,1)} - (J_{k-d}^{-1})_{(1,1)})$$

– La justification de cette approximation est que l'on a

$$\|\epsilon^{k-d}\|_A^2 - \|\epsilon^k\|_A^2 = \|r^0\|^2((J_k^{-1})_{(1,1)} - (J_{k-d}^{-1})_{(1,1)})$$

et l'on suppose que $\|\epsilon^k\|_A$ est négligeable devant $\|\epsilon^{k-d}\|_A$

– Une autre interprétation est qu'ayant une quadrature de Gauss avec $k-d$ nœuds à l'itération $k-d$, on utilise une quadrature plus précise avec k nœuds pour estimer l'erreur

Il faut faire attention à la façon de calculer $(J_k^{-1})_{(1,1)} - (J_{k-d}^{-1})_{(1,1)}$

On peut montrer que

$$\|\epsilon^{k-d}\|_A^2 \approx \sum_{j=k-d}^{k-1} \gamma_j \|r^j\|^2$$

avec

$$\gamma_j = \frac{(r^j, r^j)}{(p^j, Ap^j)}$$

un des 2 coefficients de CG

Gauss donne une borne inférieure. D'autres bornes peuvent être obtenues avec Gauss–Radau et Gauss–Lobatto

Gauss–Radau donne une borne supérieure de la norme de l'erreur si on connaît une borne inférieure de la plus petite valeur propre

CG préconditionné

Avec un préconditionnement la formule devient

$$\|\epsilon^k\|_A^2 = (z^0, r^0)((J_n^{-1})_{1,1} - (J_k^{-1})_{1,1})$$

où $Mz^0 = r^0$, M est le préconditionnement, une matrice symétrique définie positive choisie pour accélérer la convergence

La borne donnée par la quadrature de Gauss est

$$\|\epsilon^{k-d}\|_A^2 \approx \sum_{j=k-d}^{k-1} \gamma_j (z^j, r^j)$$

avec

$$Mz^j = r^j$$

Exemple

$$-\mathrm{div}(\lambda(x,y)\nabla u) = f, \quad u|_\Gamma = 0$$

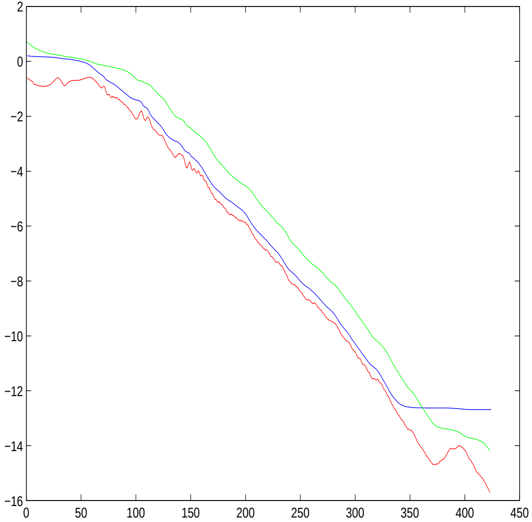
Différences finies dans le carré unité

$$\lambda(x,y) = \frac{1}{(2 + p\sin\frac{x}{\eta})(2 + p\sin\frac{y}{\eta})}$$

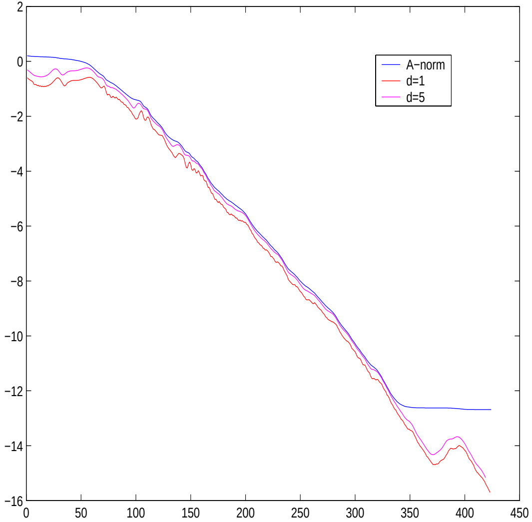
Paramètres : $p = 1.8$ and $\eta = 0.1$

On calcule f telle que la solution soit $u(x,y) = \sin(\pi x)\sin(\pi y)$

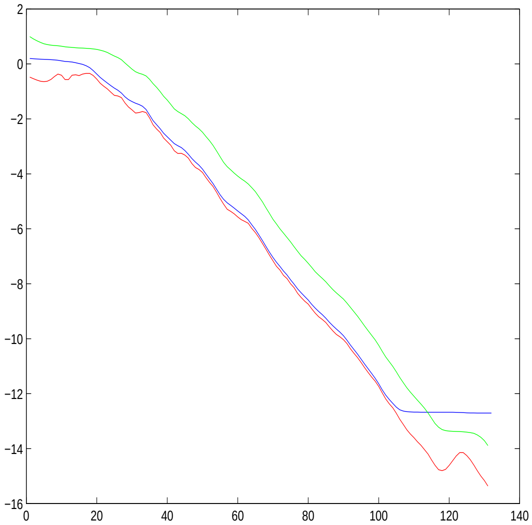
CG avec préconditionnement diagonal et IC(0)



CG, $d = 1$, $n = 10000$, $\log_{10}$ A–norme (bleu), Gauss (rouge), Gauss–Radau (vert), $a = 10^{-4}$, prec. diagonal



CG, $n = 10000$, $\log_{10}$ A–norme (bleu), Gauss d=1 (rouge), Gauss d=5 (magenta), prec. diagonal



CG, $d = 1$, $n = 10000$, $\log_{10}$ A–norme (bleu), Gauss (rouge), Gauss–Radau (vert), $a = 10^{-4}$, prec. IC(0)

Relation avec la discrétisation

Supposons que l'on veuille résoudre

$$\mathcal{L}u = f \quad \text{in } \Omega$$

Ω ouvert borné, avec des conditions limites appropriées sur Γ, frontière de Ω

Exemple simple :

$$-\Delta u = f, \quad u|_\Gamma = 0$$

Solution dans $H_0^1(\Omega)$

$$a(u, v) = (f, v), \ \forall v \in V = H_0^1(\Omega)$$

$a(u, v)$ forme bilinéaire auto-adjointe

Problème discret :

$$a(u_h, v_h) = (f, v_h), \ \forall v_h \in V_h$$

Le solveur linéaire donne $x^k \rightarrow u_h^k$
On voudrait

$$\|u - u_h^k\|_a \approx \|u - u_h\|_a$$

On a

$$\|u_h^k\|_a = \|x^k\|_A$$

Critère d'arrêt proposé par Arioli :
Soit ζ_k une estimation de $\|\varepsilon^k\|_A^2$,

$$\zeta_k \leq \eta^2((x^k)^T r^0 + b^T x^0)$$

Le paramètre η est h ou équivalent

Critère d'arrêt

Pour les différences finies, on a multiplié le second membre par h^2, on modifie un peu le critère d'Arioli

$$\text{Si } \zeta_k \leq 0.1(1/n)^2((x^k)^T r^0 + c^T x^0) \text{ alors stop}$$

où ζ_k est la borne inférieure $\|\epsilon^k\|_A^2$ donnée par Gauss

Avec $n = 10000$, la A–norme de la différence entre la solution "exacte" du système linéaire et la discrétisation de u est $n_u = 1.2682\ 10^{-3}$

Avec un préconditionnement diagonal, on fait 163 itérations et l'on a $n_x = \|u_h^k - u\|_a = 1.2755\ 10^{-3}$

Avec Cholesky incomplet IC(0) on fait 53 itérations et $n_x = 1.2668\ 10^{-3}$

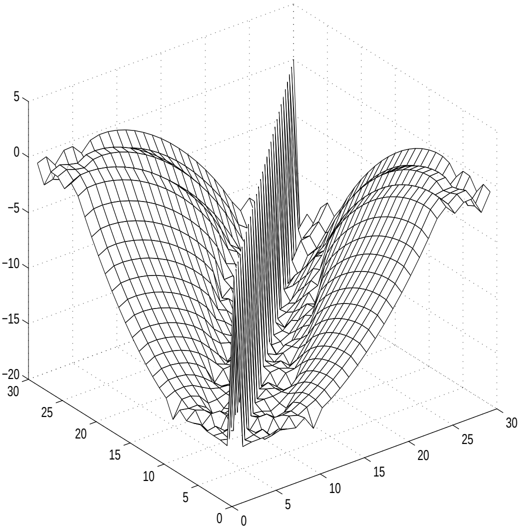
Voir aussi Jiranek, Z. Strakoš et Vohralik

CG en précision finie

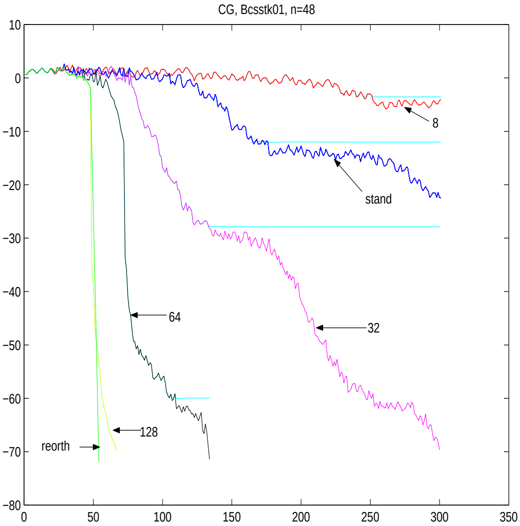
Les bornes de la norme de l'erreur fonctionnent encore en précision finie malgré les erreurs d'arrondi et la perte d'orthogonalité des résidus, voir Strakoš et Tichý (2002)

Les erreurs d'arrondi ont 3 effets sur les méthodes de Krylov :

- perte d'orthogonalité des vecteurs de base (résidus pour CG)
- augmentation du nombre d'itérations
- limitation de la précision maximum atteignable pour $\|b - Ax^k\|$



CG, Strakos30, perte d'orthogonalité



CG, BCSSTK01, $n = 48$, différentes précisions, $\log_{10}$ de la norme du résidu

Bibliographie CG

M. Arioli, *Stopping criterion for the conjugate gradient algorithm in a finite element method framework*, Numer. Math., v 97 (2004), pp 1–24

M. Arioli, D. Loghin and A. Wathen, *Stopping criteria for iterations in finite element methods*, Numer. Math., v 99, (2005), pp 381–410

G.H. Golub and G. Meurant, *Matrices, moments and quadrature*, in Numerical Analysis 1993, D.F. Griffiths and G.A. Watson eds., Pitman Research Notes in Mathematics, v 303, (1994), pp 105–156

G.H. Golub and G. Meurant, *Matrices, moments and quadrature II or how to compute the norm of the error in iterative methods*, BIT, v 37 n 3, (1997), pp 687–705

G.H. Golub and G. Meurant, *Matrices, moments and quadrature with applications*, Princeton University Press, (2010) G.H. Golub and Z. Strakoš, *Estimates in quadratic formulas*, Numer. Algo., v 8, n II–IV, (1994)

G.H. Golub and J.H. Welsch, *Calculation of Gauss quadrature rules*, Math. Comp., v 23, (1969), pp 221–230

P. Jiranek, Z. Strakoš and M. Vohralik, *A posteriori error estimates including algebraic error and stopping criteria for iterative solvers*, SIAM J. Sci. Comp., v 32, (2010), pp 1567–1590

G. Meurant, *The computation of bounds for the norm of the error in the conjugate gradient algorithm*, Numer. Algo., v 16, (1997), pp 77–87

G. Meurant, *Numerical experiments in computing bounds for the norm of the error in the preconditioned conjugate gradient algorithm*, Numer. Algo., v 22, (1999), pp 353–365

G. Meurant, *Estimates of the l_2 norm of the error in the conjugate gradient algorithm*, Numer. Algo., v 40 n 2, (2005), pp 157–169

G. Meurant, *The Lanczos and Conjugate Gradient algorithms, from theory to finite precision computations*, SIAM, (2006)

Z. Strakoš, *Model reduction using the Vorobyev moment problem*, Numer. Algo., v 51, (2009), pp 363–379

Z. Strakoš and P. Tichý, *On error estimates in the conjugate gradient method and why it works in finite precision computations*, Elec. Trans. Numer. Anal., v 13, (2002), pp 56–80

Z. Strakoš and P. Tichý, *Error estimation in preconditioned conjugate gradients*, BIT Numerical Mathematics, v 45, (2005), pp 789–817

Problèmes non symétriques

Le but est de faire pour FOM et GMRES la même chose que pour CG :

1) Obtenir des formules pour la norme l_2 de l'erreur

2) Utiliser ces formules pour calculer des estimations de la norme de l'erreur d itérations "en arrière"

Pour le cas non symétrique, on n'a pas de correspondance avec des formules de quadrature

FOM et GMRES

Soit A réelle non singulière

V_k matrice des vecteurs de base orthonormaux v^j, $j = 1, \ldots, k$ de l'espace de Krylov $\mathcal{K}_k(r^0, A)$ générés par le procédé d'Arnoldi (avec MGS)

$$x^k = x^0 + V_k z^k$$

$$AV_k = V_k H_k + h_{k+1,k} v^{k+1} (e^k)^T$$

où H_k est une matrice de Hessenberg d'ordre k et d'éléments $h_{i,j}$

On a $H_k = V_k^T A V_k$ et $AV_n = V_n H_n$ si on peut aller jusqu'à la fin

Pour FOM on demande que le résidu soit orthogonal à l'espace de Krylov

$$H_k z^k = V_k^T r^0 = \|r^0\| e^1$$

Soit

$$H_k^{(e)} = \begin{pmatrix} H_k \\ h_{k+1,k}(e^k)^T \end{pmatrix}$$

Pour GMRES on minimise la norme du résidu

$$\min_z \| \, \|r^0\| e^1 - H_k^{(e)} z \|$$

La norme de l'erreur dans FOM

On suppose que tous les H_k sont non singuliers

Le carré de la norme l_2 de l'erreur est

$$\begin{aligned} \|\epsilon^k\|^2 &= \|r^0\|^2 \left[(H_n^{-1}e^1, H_n^{-1}e^1) - (H_k^{-1}e^1, H_k^{-1}e^)) \right. \\ &+ \left. 2h_{k+1,k}(H_k^{-1}e^1, e^k)((H_n^{-1}e^{k+1})^k, H_k^{-1}e^1) \right] \end{aligned}$$

où on note $(H_n^{-1}e^{k+1})^k$ les k premières composantes de la colonne $k+1$ de l'inverse de H_n

Cette formule est intéressante mais pas très adaptée au calcul

1) Le calcul de $(H_n^{-1}e^1, H_n^{-1}e^1) - (H_k^{-1}e^1, H_k^{-1}e^{)})$ peut être problématique

2) On ne connaît pas le signe des autres termes

Il est donc utile de transformer cette formule

“Simplification” de la formule pour FOM

L’idée est d’exprimer H_n^{-1} en fonction de H_k^{-1}
Soit

$$H_n = \begin{pmatrix} H_k & W_k \\ Y_k^T & \tilde{H}_k \end{pmatrix}$$

et $w^k = W_k \tilde{H}_k^{-1} e^1$,

$$\gamma_k = \frac{h_{k+1,k}(e^k, H_k^{-1} e^1)}{1 - h_{k+1,k}(e^k, H_k^{-1} w^k)}$$

$\|\epsilon^k\|^2 / \|r^0\|^2$ est donnée par

$$\{h_{k+1,k}\,(\,(e^k, H_k^{-1} e^1) + \gamma_k (e^k, H_k^{-1} w^k))\,\}^2 \|\tilde{H}_k^{-1} e^1\|^2 + \gamma_k^2 \|H_k^{-1} w^k\|^2$$

Dans cette version on a la somme de deux termes positifs
Evidemment, on ne connaît pas $\tilde{H}_k$ et W_k

Relations avec le résidu

En supposant H_k non singulier,

$$\|r^k\|^2 = \|r^0\|^2 h_{k+1,k}^2 (H_k^{-1} e^1, e^k)^2$$

Avec les notations précédentes, on a

$$\|\epsilon^k\|^2 = \|r^k\|^2 \frac{\|\tilde{H}_k^{-1} e^1\|^2 + \|H_k^{-1} w^k\|^2}{[1 - h_{k+1,k}(e^k, H_k^{-1} w^k)]^2}$$

Estimation de la norme de l'erreur dans FOM

$$H_k = \begin{pmatrix} H_{k-d} & W_{k-d} \\ Y_{k-d}^T & \tilde{H}_{k-d} \end{pmatrix}$$

(Attention au changement de notations)

A l'itération k on obtient une approximation de la norme de l'erreur $\|\epsilon^{k-d}\|^2$ a l'itération $k-d$ $(d \geq 1)$ par

$$\begin{aligned} & \|r^0\|^2 \,[\, \{ h_{k-d+1,k-d} \,(\, (e^{k-d}, H_{k-d}^{-1} e^1) \\ + \quad & \gamma_{k-d}(e^{k-d}, H_{k-d}^{-1} w^{k-d})) \,\}^2 \|\tilde{H}_{k-d}^{-1} e^1\|^2 + \gamma_{k-d}^2 \|H_{k-d}^{-1} w^{k-d}\|^2 \,] \end{aligned}$$

où $w^{k-d} = W_{k-d} \tilde{H}_{k-d}^{-1} e^1$ et

$$\gamma_{k-d} = \frac{h_{k-d+1,k-d}(e^{k-d}, H_{k-d}^{-1} e^1)}{1 - h_{k-d+1,k-d}(e^{k-d}, H_{k-d}^{-1} w^{k-d})}$$

Exemples

Matrices du Matrix Market et discrétisation d'une EDP :

E05r0500 : dynamique des fluides, Reynolds $Re = 500$, $n = 236, \kappa = 1.16\ 10^6$, valeurs propres complexes, parties réelles négatives

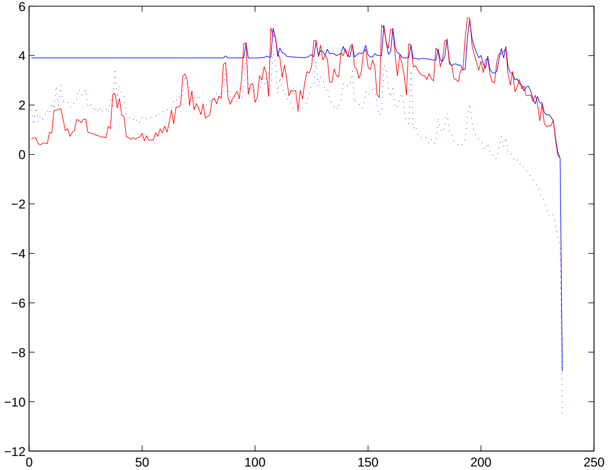
Steam1 : modèle 3D de réservoir pétrolier, $n = 240, \kappa = 2.82\ 10^7$, valeurs propres réelles et négatives

Steam2 : modèle 3D de réservoir pétrolier, $n = 600, \kappa = 3.78\ 10^6$, valeurs propres réelles et négatives

Convection-diffusion :

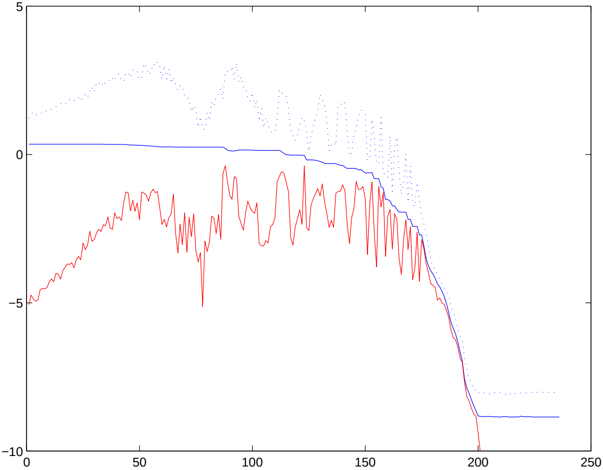
$$-\Delta u + 2e^{2(x^2+y^2)}\frac{\partial u}{\partial x} = f$$

carré unité, schéma upwind, $n = 2500, \kappa = 1360$



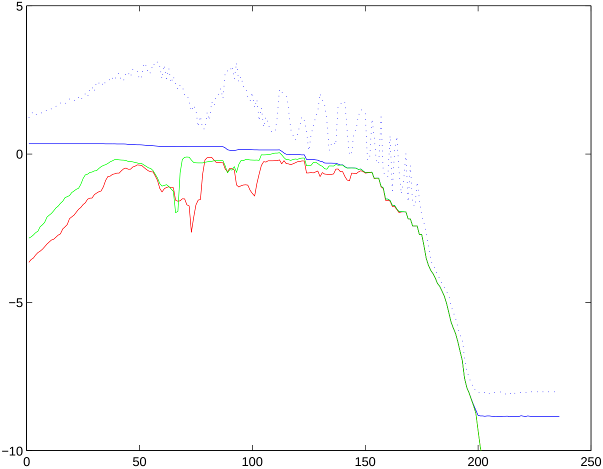
E05r0500 : $n = 236, \kappa = 1.16\ 10^6$, FOM, d=1

Erreur (bleu), résidu (pointillé), estimation (rouge)

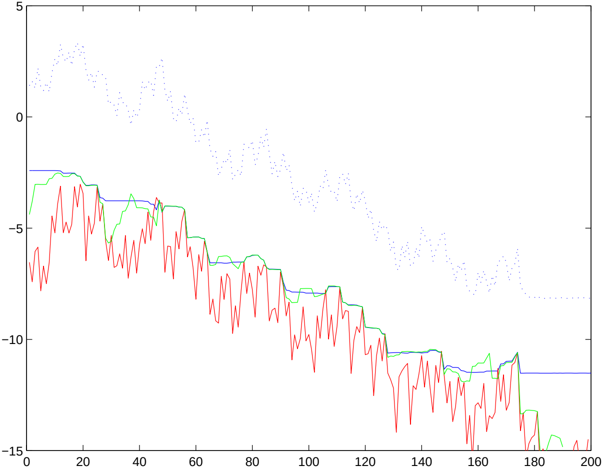


Steam1 : $n = 240, \kappa = 2.82\ 10^7$, FOM, d=1

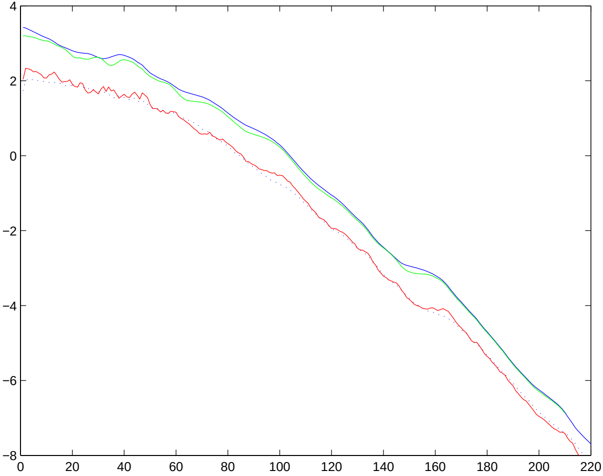
Erreur (bleu), résidu (pointillé), estimation (rouge)



Steam1 : $n = 240, \kappa = 2.82\ 10^7$, FOM
Erreur (bleu), résidu (pointillé), estimation d=10 (rouge), d=20 (vert)



Steam2 : $n = 600, \kappa = 3.78\ 10^6$, FOM
Erreur (bleu), résidu (pointillé), estimation d=1 (rouge), d=10 (vert)



Convection-diffusion : $n = 2500, \kappa = 1360$, FOM
Erreur (bleu), résidu (pointillé), estimation d=1 (rouge), d=10 (vert)

Relations entre FOM et GMRES

On note FOM (F) et GMRES (G)

Soit t^k la dernière colonne de $(H_k^T H_k)^{-1}$,

$$\delta_{k+1} = \frac{h_{k+1,k}^2}{1 + h_{k+1,k}^2 t_k^k}$$

et $u^k = \delta_{k+1} t^k$

$$x_G^k = x_F^k - (z_F^k)_k V_k u^k$$

$$\epsilon_G^k = \epsilon_F^k + (z_F^k)_k V_k u^k$$

où z_F^k est le vecteur des coordonnées de FOM donnée par $z_F^k = \|r^0\| H_k^{-1} e^1$

La norme de l'erreur dans GMRES

Avec les mêmes notations que pour FOM $w^k = W_k \tilde{H}_k^{-1} e^1$

$$\gamma_k = \frac{h_{k+1,k}(e^k, H_k^{-1} e^1)}{1 - h_{k+1,k}(e^k, H_k^{-1} w^k)}$$

Soit t^k la dernière colonne de $(H_k^T H_k)^{-1}$ et

$$\delta_{k+1} = \frac{h_{k+1,k}^2}{1 + h_{k+1,k}^2 t_k^k}$$

avec $u^k = \delta_{k+1} t^k$

$$\|\epsilon_G^k\|^2 = \|\epsilon_F^k\|^2 + \|r^0\|^2 [2\gamma_k (H_k^{-1} e^1, e^k)(H_k^{-1} w^k, u^k) + (H_k^{-1} e^1, e^k)^2 \|u^k\|^2]$$

Estimation de la norme de l'erreur dans GMRES

On utilise la formule pour $\|\epsilon_F^{k-d}\|^2$

Le terme additionnel est approché par

$$\|r^0\|^2[2\gamma_{k-d}(H_{k-d}^{-1}e^1, e^{k-d})(H_{k-d}^{-1}w^{k-d}, u^{k-d})$$

$$+(H_{k-d}^{-1}e^1, e^{k-d})^2\|u^{k-d}\|^2]$$

Exemples

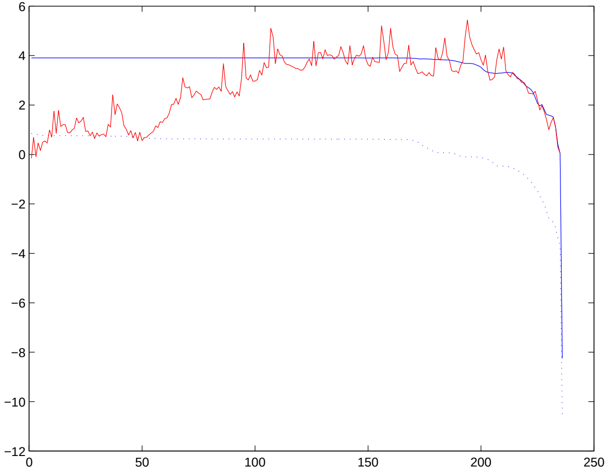
On utilise les mêmes exemples que pour FOM

Remarques :

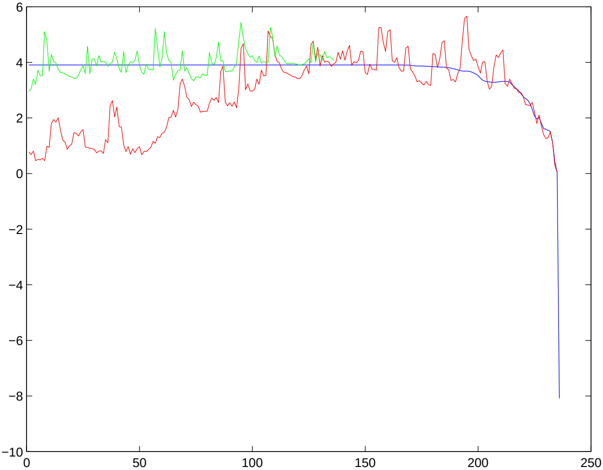
1) Les résultats sont un peu meilleurs

2) Les courbes oscillent moins (voir Steam1)

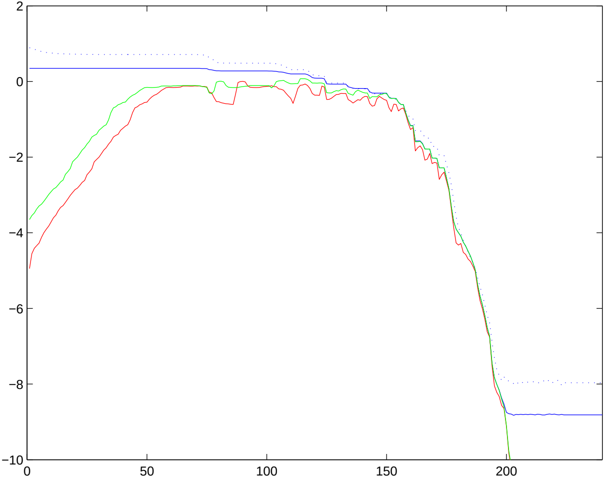
3) Augmenter le délai d améliore les approximations



E05r0500 : $n = 236, \kappa = 1.16\ 10^6$, GMRES, d=1
Erreur (bleu), résidu (pointillé), estimation (rouge)

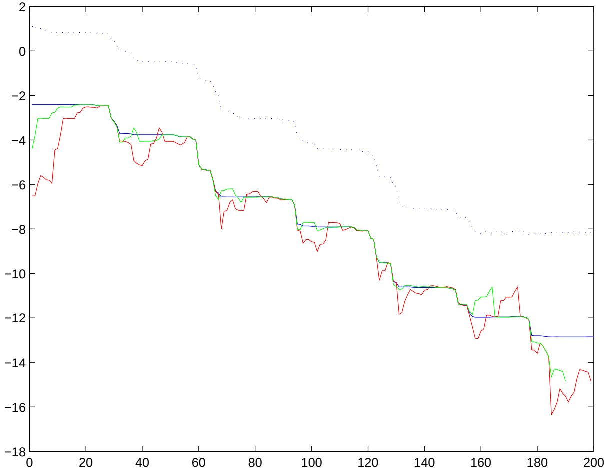


E05r0500 : $n = 236, \kappa = 1.16\ 10^6$, GMRES,
Erreur (bleu), résidu (pointillé), estimation $d = 1$ (rouge), $d = 100$ (vert)

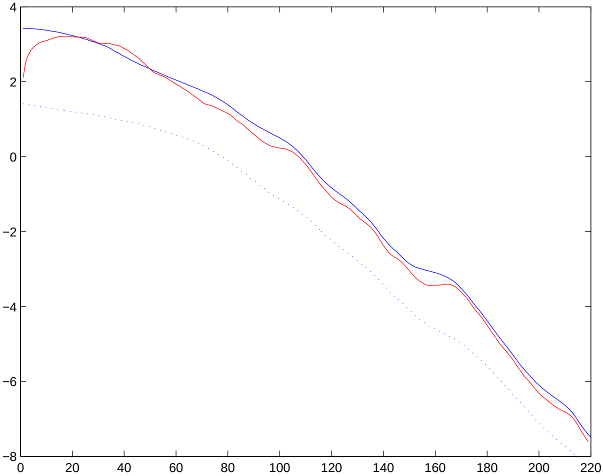


Steam1 : $n = 240, \kappa = 2.82\ 10^7$, GMRES

Erreur (bleu), résidu (pointillé), estimation d=1 (rouge), d=10 (vert)



Steam2 : $n = 600, \kappa = 3.78\ 10^6$, GMRES
Erreur (bleu), résidu (pointillé), estimation d=1 (rouge), d=10 (vert)



Convection-diffusion : $n = 2500, \kappa = 1360$, GMRES, d=1

Erreur (bleu), résidu (pointillé), estimation (rouge)

Il existe d'autres possibilités d'estimation de la norme de l'erreur

Voir les articles de Brezinski et al.

Conclusions

Pour les méthodes de Kylov utilisant des bases orthogonales on sait écrire des formules (simples !) donnant la norme de l'erreur

Ces formules sont utilisées pour calculer des estimations (ou des bornes pour CG) de la norme de l'erreur au cours des itérations

Ces estimations peuvent être utilisées pour obtenir des critères d'arrêt des itérations plus surs, reliés à l'erreur de discrétisation pour les problèmes venant de la discrétisation d'EDP