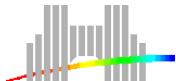


Résoudre et certifier la solution d'un système linéaire

NGUYEN Hong Diep, Nathalie Revol

RAIM 2008

2 juin 2008



Objectif :

- 1 Résoudre un système linéaire
 - $A \in \mathbb{R}^{n \times n}$
 - $b \in \mathbb{R}^n$
 - Trouver un vecteur $\tilde{x} \in \mathbb{R}^n$ tel que : $A * \tilde{x} \approx b$
- 2 Borner l'erreur de la solution trouvée en même temps en utilisant l'arithmétique par intervalle.
 - $\Delta x = x^* - \tilde{x}$
 - Trouver un petit intervalle \mathbf{e} contenant Δx .

Méthode classique de raffinement itératif

Wilkinson, Li & Demmel, Higham, ...

Algorithme 1 (Méthode classique de raffinement itératif)

Entrées : $A \in \mathbb{F}^{n \times n}$, $b \in \mathbb{F}^n$

$x = A \setminus b$;

while(test d'arrêt)

$r = b - A * x$;

$e = A \setminus r$;

$x = x + e$;

end

Passer à l'arithmétique par intervalle

Notations : x^* la solution exacte, \tilde{x} une approximation flottante, \mathbf{x} un intervalle contenant la solution

Algorithme 2 (Méthode de raffinement itératif par intervalle)

Entrées : $A \in \mathbb{F}^{n \times n}$, $b \in \mathbb{F}^n$

$\tilde{x} = A \setminus b$;

while(test d'arrêt)

$\mathbf{r} = [b - A * \tilde{x}]$; % $A * (x^* - \tilde{x}) \in \mathbf{r}$

$\mathbf{e} = A \setminus \mathbf{r}$; % $x^* - \tilde{x} \in \mathbf{e}$

$\tilde{x} = \tilde{x} + \text{mid}(\mathbf{e})$;

end

$$A * \mathbf{e} = \mathbf{r}$$

\mathbf{r} doit être calculé avec une précision étendue (précision doublée)

Méthode : Raffinement itératif par intervalle

- ne chercher pas à résoudre directement un système par intervalle.
- considérer ce système comme des contraintes pour raffiner la borne d'erreur

$$\begin{aligned} \mathbf{e}' &= \mathbf{e} \cap (A \setminus \mathbf{r}) \\ &= \{ \tilde{\mathbf{e}} \in \mathbf{e} \mid \exists \tilde{\mathbf{r}} \in \mathbf{r} : A * \tilde{\mathbf{e}} = \tilde{\mathbf{r}} \} \end{aligned}$$

- 1 Calculer la décomposition LU de A ,
- 2 Calculer une approximation flottante \tilde{x} du système,
- 3 Calculer une matrice R telle que $L * R * U \simeq I$,
- 4 Calculer $K = R * A$,
- 5 Tester si K est une H-matrice
Si K n'est pas une H-matrice, alors échec de certification de la solution,

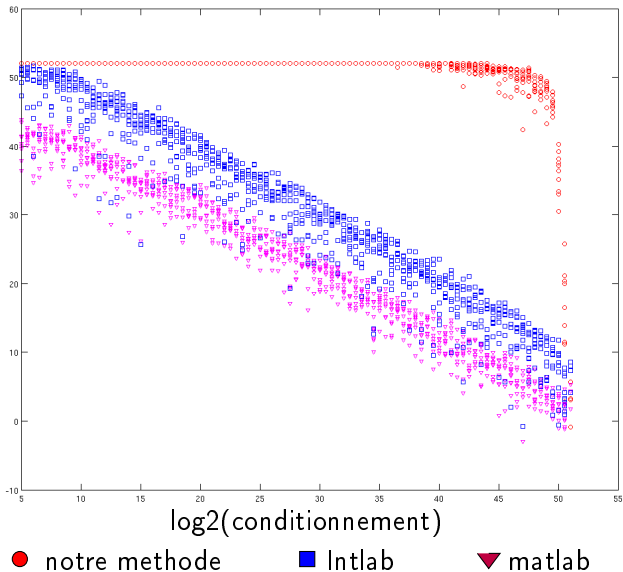
- 6 Calculer le résidu : $\mathbf{r} = \mathbf{b} - \mathbf{A} * \tilde{\mathbf{x}}$ avec une précision doublée,
- 7 Calculer : $\mathbf{z} = \mathbf{R} * \mathbf{r}$,
- 8 Calculer une approximation initiale de l'erreur $\mathbf{e0}$,
- 9 while (not convergence)
 - Appliquer 3 itérations de Gauss-Seidel sur $\mathbf{K}, \mathbf{z}, \mathbf{e0}$,
 - Mettre à jour l'approximation flottante et la borne d'erreur :

$$\tilde{\mathbf{x}} = \tilde{\mathbf{x}} + \text{mid}(\mathbf{e0}) \quad \mathbf{e0} = \mathbf{e0} - \text{mid}(\mathbf{e0})$$

- Recalculer le résidu.

Résultats expérimentaux

$-\log_2(\text{erreur relative})$



- 1 Comment trouver une solution initiale
- 2 Qu'est-ce que la méthode de Gauss-Seidel
- 3 Quand arrêter les itérations

- 1 Comment trouver une solution initiale
- 2 Qu'est-ce que la méthode de Gauss-Seidel
- 3 Quand arrêter les itérations

Principe :

Si la matrice A a de bonnes propriétés : elle est facilement inversible
On peut la préconditionner par C et obtenir une matrice proche de l'Identité : diagonale positive et dominante par exemple.

→ alors on peut trouver un encadrement de la solution.

Solution initiale (2)

Proposition simplifiée de Neumaier¹ : Soient $\mathbf{A} \in \mathbb{R}^{n \times n}$ et $\mathbf{C} \in \mathbb{R}^{n \times n}$.
Si $\langle \mathbf{CA} \rangle u \geq v > 0$ pour $u \geq 0$ alors

$$\mathbf{A}^H b \subseteq \|Cb\|_v * [-u, u].$$

Soient :

- $\|Cb\|_v = \max\{|Cb_i|/v_i \mid i = 1, \dots, n\}$
- \mathbf{A}^H une enveloppe convexe de l'inverse de la matrice \mathbf{A}
- $\langle \mathbf{CA} \rangle$ la matrice de comparaison de \mathbf{CA}

$$\begin{aligned}\langle \mathbf{CA} \rangle_{i,i} &= \min(|\mathbf{CA}_{i,j}|) \\ \langle \mathbf{CA} \rangle_{i,j \neq i} &= -\max(|\mathbf{CA}_{i,j}|)\end{aligned}$$

¹page 121, Arnold Neumaier, *Interval Methods for Systems of Equations*

Solution initiale (3)

Soit R l'inverse approximative de A .

Choisir : $C = R \Rightarrow CAC' = RA \ (\approx I)$

Hypothèse : RA est une H-matrice.

Il existe $u \geq 0$ tel que $\langle RA \rangle * u = v > 0$.

Approximation initiale pour la borne d'erreur :

$$\begin{aligned} \mathbf{e} &\subseteq A^H \mathbf{r} \\ &\subseteq \underbrace{\|R * \mathbf{r}\|_v * [-u, u]}_{\mathbf{e}_0} \end{aligned}$$

Pré-conditionner : $\mathbf{K} = [R * A] \quad \mathbf{z} = R * \mathbf{r}$

$$\begin{aligned} A * \mathbf{e} &= \mathbf{r} \\ \Rightarrow \mathbf{K} * \mathbf{e} &= \mathbf{z} \end{aligned}$$

- 1 Comment trouver une solution initiale
- 2 Qu'est-ce que la méthode de Gauss-Seidel
- 3 Quand arrêter les itérations

Méthodes :

- Krawczyk
- Gauss-Seidel
- Élimination de Gauss par intervalle

La méthode de Gauss-Seidel est prouvée converger plus rapidement dans le cas où la matrice \mathbf{K} est une H-matrice (cf. Théorème de Neumaier ²)

²page 138, Arnold Neumaier, *Interval Methods for Systems of Equations*

La méthode de Gauss-Seidel (1)

$$Ke = z$$
$$\sum_{j=1}^n K_{ij}e_j = z_i \quad (i = 1, \dots, n)$$

la i^e ligne s'écrit, en développant selon les composantes :

$$K_{i,1}e_1 + \dots + K_{i,i-1}e_{i-1} + K_{i,i}e_i + K_{i,i+1}e_{i+1} + \dots + K_{i,n}e_n = z_i$$

$$e_i = \left(z_i - \sum_{j=1}^{i-1} K_{i,j}e_j - \sum_{j=i+1}^n K_{i,j}e_j \right) / K_{i,i}$$

La méthode de Gauss-Seidel (1)

$$Ke = z$$
$$\sum_{j=1}^n K_{ij}e_j = z_i \quad (i = 1, \dots, n)$$

la i^e ligne s'écrit, en développant selon les composantes :

$$K_{i,1}e_1 + \dots + K_{i,i-1}e_{i-1} + K_{i,i}e_i + K_{i,i+1}e_{i+1} + \dots + K_{i,n}e_n = z_i$$

$$e'_i = \left(z_i - \sum_{j=1}^{i-1} K_{i,j}e_j - \sum_{j=i+1}^n K_{i,j}e_j \right) / K_{i,i} \quad (\text{Jacobi})$$

La méthode de Gauss-Seidel (1)

$$Ke = z$$
$$\sum_{j=1}^n K_{ij}e_j = z_i \quad (i = 1, \dots, n)$$

la i^e ligne s'écrit, en développant selon les composantes :

$$K_{i,1}e_1 + \dots + K_{i,i-1}e_{i-1} + K_{i,i}e_i + K_{i,i+1}e_{i+1} + \dots + K_{i,n}e_n = z_i$$

$$e'_i = \left(z_i - \sum_{j=1}^{i-1} K_{i,j}e_j - \sum_{j=i+1}^n K_{i,j}e_j \right) / K_{i,i} \quad (\text{Jacobi})$$

$$e'_i = \left(z_i - \sum_{j=1}^{i-1} K_{i,j}e'_j - \sum_{j=i+1}^n K_{i,j}e_j \right) / K_{i,i} \quad (\text{Gauss - Seidel})$$

La méthode de Gauss-Seidel (2)

$$(\mathbf{K} \setminus \mathbf{z}) \cap \mathbf{e} = \{ \tilde{\mathbf{e}} \in \mathbf{e} \mid \exists \tilde{\mathbf{K}} \in \mathbf{K}, \exists \tilde{\mathbf{z}} \in \mathbf{z} : \tilde{\mathbf{K}} * \tilde{\mathbf{e}} = \tilde{\mathbf{z}} \}$$

$$\tilde{K}_{i,1}\tilde{e}_1 + \dots + \tilde{K}_{i,i-1}\tilde{e}_{i-1} + \tilde{K}_{i,i}\tilde{e}_i + \tilde{K}_{i,i+1}\tilde{e}_{i+1} + \dots + \tilde{K}_{i,n}\tilde{e}_n = z_i$$

$$\tilde{e}_i = \left(z_i - \sum_{j=1}^{i-1} \tilde{K}_{i,j}\tilde{e}_j - \sum_{j=i+1}^n \tilde{K}_{i,j}\tilde{e}_j \right) / \tilde{K}_{i,i}$$

$$e'_i \subseteq \left(z_i - \sum_{j=1}^{i-1} K_{i,j}e'_j - \sum_{j=i+1}^n K_{i,j}e_j \right) / K_{i,i} \cap e_i$$

La méthode de Gauss-Seidel (2)

$$(\mathbf{K} \setminus \mathbf{z}) \cap \mathbf{e} = \{\tilde{\mathbf{e}} \in \mathbf{e} \mid \exists \tilde{\mathbf{K}} \in \mathbf{K}, \exists \tilde{\mathbf{z}} \in \mathbf{z} : \tilde{\mathbf{K}} * \tilde{\mathbf{e}} = \tilde{\mathbf{z}}\}$$

$$\tilde{K}_{i,1}\tilde{e}_1 + \dots + \tilde{K}_{i,i-1}\tilde{e}_{i-1} + \tilde{K}_{i,i}\tilde{e}_i + \tilde{K}_{i,i+1}\tilde{e}_{i+1} + \dots + \tilde{K}_{i,n}\tilde{e}_n = z_i$$

$$\tilde{e}_i = \left(z_i - \sum_{j=1}^{i-1} \tilde{K}_{i,j}\tilde{e}_j - \sum_{j=i+1}^n \tilde{K}_{i,j}\tilde{e}_j \right) / \tilde{K}_{i,i}$$

$$\mathbf{e}'_i \subseteq \left(z_i - \sum_{j=1}^{i-1} \mathbf{K}_{i,j}\mathbf{e}'_j - \sum_{j=i+1}^n \mathbf{K}_{i,j}\mathbf{e}_j \right) / \mathbf{K}_{i,i} \cap \mathbf{e}_i$$

La méthode de Gauss-Seidel (2)

$$(\mathbf{K} \setminus \mathbf{z}) \cap \mathbf{e} = \{ \tilde{\mathbf{e}} \in \mathbf{e} \mid \exists \tilde{\mathbf{K}} \in \mathbf{K}, \exists \tilde{\mathbf{z}} \in \mathbf{z} : \tilde{\mathbf{K}} * \tilde{\mathbf{e}} = \tilde{\mathbf{z}} \}$$

$$\tilde{K}_{i,1}\tilde{e}_1 + \dots + \tilde{K}_{i,i-1}\tilde{e}_{i-1} + \tilde{K}_{i,i}\tilde{e}_i + \tilde{K}_{i,i+1}\tilde{e}_{i+1} + \dots + \tilde{K}_{i,n}\tilde{e}_n = z_i$$

$$\tilde{e}_i = \left(z_i - \sum_{j=1}^{i-1} \tilde{K}_{i,j}\tilde{e}_j - \sum_{j=i+1}^n \tilde{K}_{i,j}\tilde{e}_j \right) / \tilde{K}_{i,i}$$

$$\mathbf{e}'_i \subseteq \left(z_i - \sum_{j=1}^{i-1} \mathbf{K}_{i,j}\mathbf{e}'_j - \sum_{j=i+1}^n \mathbf{K}_{i,j}\mathbf{e}_j \right) / \mathbf{K}_{i,i} \cap \mathbf{e}_i$$

- 1 Comment trouver une solution initiale
- 2 Qu'est-ce que la méthode de Gauss-Seidel
- 3 Quand arrêter les itérations

Hypothèse (rappel) : le problème de départ est de calculer $x | Ax = b$ $A \in \mathbb{F}^{n \times n}$, $b \in \mathbb{F}^n$ ← système sans coefficient intervalle.

- 1 Nombre de bits corrects demandés :

$$nb_bits = -\log_2(\max(\text{diam}(e_j)/\text{abs}(\tilde{x}_j))) \quad (j = 1, \dots, n)$$


- 2 $\text{diam}(e^0), \text{diam}(e^1), \dots$ est une suite positive décroissante

$$\text{diam}(e^i) - \text{diam}(e^{i+1}) < \epsilon * \text{diam}(e^i)$$

- Notre méthode utilise deux niveaux de raffinement :
 - ① Un raffinement flottant pour l'approximation flottante de la solution,
 - ② Un raffinement par intervalle pour la borne d'erreur sur cette approximation.
- Notre méthode donne des résultats très fins quand le conditionnement de la matrice des coefficients est plus petit que 2^p (p est la précision courante).
- Perspectives :
 - Étudier l'effet de la précision utilisée pour le calcul du résidu sur la qualité du résultat.
 - Adapter les méthodes de Rump pour inverser les matrices mal conditionnées

 Arnold Neumaier
Interval Methods for Systems of Equations
Cambridge University Press, 1990.

 Siegfried M. Rump,
Computer-assisted Proofs and Self-validating Methods
MicroHandbook on Accuracy and Reliability in Scientific Computation
Bo Einarsson, (p.195-240).

 Nicholas J. Higham,
Accuracy and Stability of Numerical Algorithms
SIAM, 2002.



J. H. Wilkinson,
Rounding Errors in Algebraic Processes
Notes on Applied Science No. 32
Her Majesty's Stationery Office, London, 1963.



James Demmel, Yozo Hida, W. Kahan, Xiaoye S.Li, Soni Mukherjee,
E. Jason Riedy
Error Bounds from Extra Precise Iterative Refinement
Report No. UCB/CSD-04-1344
Computer Science Division (EECS), University of California, 2004.