

# Méthodes itératives de Krylov II

Gérard MEURANT

Novembre 2014

① Généralités sur Q-OR et Q-MR

② FOM et GMRES

③ Exemples GMRES

④ Q-OR et Q-MR

⑤ Exemples Q-OR/Q-MR

# Systèmes non symétriques

Méthode la plus simple : les équations normales

$$Ax = b$$

$$A^T Ax = A^T b$$

et on utilise CG, mais

$$\kappa(A^T A) = \kappa(A)^2$$

Variante :

$$AA^T y = b, \quad x = A^T y$$

De nombreuses méthodes de Krylov ont été proposées depuis 1950 pour résoudre les systèmes linéaires

La plupart d'entre elles peuvent être classées comme quasi-orthogonalisation (Q-OR) ou comme quasi-résidu minimum (Q-MR)

Q-OR: FOM, BiCG, Hessenberg, ...

Q-MR: GMRES, truncated GMRES, QMR, CMRH, ...

Quelle que soit leur définition ces méthodes partagent de nombreuses propriétés fondamentales

Voir le bel article de [M. Eiermann et O.G. Ernst](#), *Geometric aspects in the theory of Krylov subspace methods*, Acta Numerica, v 10 n 10 (2001), pp. 251–312

La différence principale vient de la base de l'espace de [Krylov](#) qui est utilisée :

- orthogonale pour [FOM/GMRES](#) (véritables méthodes OR/MR)
- bi-orthogonale pour [BiCG/QMR](#)
- basée sur une factorisation LU de la matrice de Krylov pour [Hessenberg/CMRH](#)

# Rappel

Calcul de la base orthogonale : méthode d'Arnoldi

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

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

$$AV_k = V_{k+1} \underline{H}_k$$

avec

$$\underline{H}_k = \begin{pmatrix} H_k \\ h_{k+1,k} (e_k)^T \end{pmatrix}$$

On a aussi

$$V_k^T AV_k = H_k$$

Les itérés s'écrivent  $x_k = x_0 + V_k y_k$

GMRES (Saad and Schulz 1986) minimise la norme  $\ell_2$  du résidu  $b - Ax_k$  soit

$$\| \|r_0\| e_1 - \underline{H}_k y_k \|$$

On résout en calculant incrémentalement une décomposition QR de  $\underline{H}_k$  avec des rotations de Givens

Dans FOM on résout des systèmes linéaires

$$H_k y_k = \|r_0\| e_1$$

avec des décompositions QR

**Problème:** le stockage croît avec le nombre d'itérations puisqu'il faut stocker les vecteurs de base  $v_j$

**Solution :** on redémarre toutes les  $m$  itérations  $\rightarrow$  GMRES( $m$ )

Pb : choix de  $m$  ?



# Résolution du pb de moindres carrés

Décomposition QR de  $\underline{H}_k$

On utilise des rotations de Givens pour annihiler les éléments sous-diagonaux

$$\begin{pmatrix} x & x & \dots & x \\ & \ddots & & \vdots \\ & & \ddots & \\ & & & x & x \\ & & & 0 & r \\ & & & 0 & h \end{pmatrix}$$

On multiplie par une matrice de rotation (orthogonale)

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c_j & -s_j & \\ & & & s_j & c_j & \end{pmatrix}$$

$$c_j = \frac{r}{\sqrt{r^2 + h^2}}, \quad s_j = -\frac{h}{\sqrt{r^2 + h^2}}$$

$$Q_k \underline{H}_k = R_k$$

$$r_{k+1,k+1} = c_k r - s_k h = (r^2 + h^2) / \sqrt{r^2 + h^2} = \sqrt{r^2 + h^2} > 0$$

# Relations entre FOM et GMRES

Il y a des relations entre les résidus de FOM  $r_k^F$  et ceux de GMRES

$r_k^G$

Soit  $s_k, c_k$  la rotation entre les lignes  $k$  et  $k + 1$

$$\frac{\|r_k^G\|}{\|r_{k-1}^G\|} = |s_k|$$

$$\|r_k^G\| = |c_k| \|r_k^F\|$$

$$\frac{1}{\|r_k^F\|^2} = \frac{1}{\|r_k^G\|^2} - \frac{1}{\|r_{k-1}^G\|^2}$$

Si  $s_k$  est petit, GMRES réduit beaucoup la norme du résidu et c'est la même chose pour FOM

Si FOM dégénère ( $H_k$  singulière), GMRES stagne

Si  $A$  est diagonalisable ( $A = X\Lambda X^{-1}$ )

$$\|r_k\| \leq \epsilon_k \|X\| \|X^{-1}\| \|r_0\|$$

où  $\epsilon_k$  est le minimum de  $\max_{\lambda} |p_k(\lambda)|$  pour tous les polynômes  $p_k$  de degré  $k$  qui satisfont  $p_k(0) = 1$  et les  $\lambda$  sont les valeurs propres de  $A$

Pour bien comprendre la convergence de FOM et GMRES, il faut se rappeler que toutes ces méthodes de Krylov sont (en arithmétique exacte) des méthodes *directes*

Elles s'arrêtent avec un résidu nul avant ou à l'itération  $n$

La notion de vitesse de convergence asymptotique n'a pas de sens pour ces méthodes (contrairement aux méthodes classiques - Jacobi, Gauss-Seidel, ...)

# Que sait-on à propos de GMRES?

La question principale est : de quelles quantités la convergence de **GMRES** dépend-elle ?

Soit

$$K = (b \quad Ab \quad A^2b \quad \dots \quad A^{n-1}b)$$

la matrice de **Krylov** que l'on suppose de rang  $n$ . Alors

$$K = VU$$

où  $V$  est orthogonale (ou unitaire) et  $U$  est triangulaire supérieure avec des éléments diagonaux réels et positifs

Comme on sait, la matrice  $H = V^*AV$  est Hessenberg supérieure

On a

$$H = UCU^{-1}$$

où  $C$  est la matrice compagnon des valeurs propres de  $A$

[C'est une conséquence de  $AK = KC$ , Th de Cayley-Hamilton]

Sans perte de généralité supposons que  $x_0 = 0$  et  $\|b\| = 1$

On sait que

- toute courbe de convergence non croissante est possible pour la norme des résidus de GMRES
- on sait construire des matrices  $A$  avec un spectre donné et des second membres  $b$  tels que GMRES donne une courbe prescrite de convergence décroissante pour les normes des résidus. De plus on peut prescrire les valeurs de Ritz (valeurs propres de  $H_k$ ) à chaque itération
- on connaît deux paramétrisations de cette classe de matrices et de second membres

Pour ces propriétés, voir

A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent advances in iterative methods, G.H. Golub, A. Greenbaum and M. Luskin, eds., Springer, (1994), pp. 95–118

A. Greenbaum, V. Pták and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., v 17 (1996), pp. 465–469

M. Arioli, V. Pták and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT Numerical Mathematics, v 38 n 4 (1998), pp. 636–643

J. Duintjer Tebbens and G. Meurant, *Any Ritz value behavior is possible for Arnoldi and for GMRES*, SIAM J. Matrix Anal. Appl., v 33 n 3 (2012), pp. 958–978

# La paramétrisation APS

Soient  $n$  nombres positifs donnés tels que

$$1 = f_0 \geq f_1 \geq \cdots \geq f_{n-1} > 0$$

et  $n$  nombres complexes non nuls  $\lambda_1, \dots, \lambda_n$ . Soit  $A$  une matrice d'ordre  $n$  and  $b$  un vecteur de longueur  $n$ . Les assertions suivantes sont équivalentes :

- 1- Le spectre de  $A$  est  $\{\lambda_1, \dots, \lambda_n\}$  et GMRES appliqué à  $A$  et  $b$  donne des résidus  $r_j^G$ ,  $j = 0, \dots, n-1$  tels que

$$\|r_j^G\| = f_j, \quad j = 0, \dots, n-1$$



- 2- La matrice  $A$  est de la forme  $A = WYCY^{-1}W^*$  et  $b = Wh$ , où  $W$  est n'importe quelles matrice unitaire,  $Y$  est donné par

$$Y = \begin{pmatrix} h & R \\ & 0 \end{pmatrix}$$

$R$  étant une matrice triangulaire supérieure non singulière quelconque d'ordre  $n - 1$ ,  $h$  est un vecteur tel que

$$h = (\eta_1, \dots, \eta_n)^T, \quad \eta_j = (f_{j-1}^2 - f_j^2)^{1/2}$$

et  $C$  est la matrice compagnon correspondant au polynôme  $q$ ,

$$q(z) = (z - \lambda_1) \cdots (z - \lambda_n) = z^n + \sum_{j=0}^{n-1} \alpha_j z^j$$

$$C = \begin{pmatrix} 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \dots & -\alpha_2 \\ & & \ddots & \ddots & \vdots \\ & & & 1 & -\alpha_{n-1} \end{pmatrix}$$

Ces résultats ont conduit de nombreuses personnes à penser (et à écrire) que “pour les matrices non normales, la convergence de GMRES ne dépend pas des valeurs propres de  $A$ ”

Nous allons voir que cette affirmation n'est pas correcte

En fait : “pour les matrices non normales, la convergence de GMRES ne dépend pas **seulement** des valeurs propres”

# Une autre paramétrisation, GM-JDT

On remplace [2] dans APS par

- 2- La matrice  $A$  est de la forme  $A = VUCU^{-1}V^*$  et  $b = Ve_1$ , où  $V$  est n'importe quelle matrice unitaire,  $U$  est non singulière et triangulaire supérieure telle que

$$U_{1,1}^{-1} = 1, \quad U_{1,j}^{-1} = \left( \frac{1}{f_{j-1}^2} - \frac{1}{f_{j-2}^2} \right)^{\frac{1}{2}}, \quad j = 2, \dots, n$$

les autres éléments étant quelconques et  $C$  est la matrice compagnon des valeurs propres prescrites

De plus

- $|(U^{-1})_{1,k}| = 1/\|r_{k-1}^F\|$
- $\|r_k^G\|^2 = 1/(M_{k+1}^{-1})_{1,1}$  with  $M_{k+1} = U_{k+1}^* U_{k+1}$

Ce dernier résultat a été démontré par de nombreux chercheurs : Stewart, Zitko, Ipsen, Liesen, Rozložník et Strakoš, et Sadok

En suivant des idées proposées par H. Sadok, on peut calculer  $(M_{k+1}^{-1})_{1,1}$  en utilisant deux outils “simples” :

- ▶ la formule de **Cramer** (1750 mais connue auparavant)
- ▶ la formule de **Cauchy-Binet** (1812) pour  $\det(AB)$  avec  $A$  and  $B$  rectangulaires



G. Cramer  
1704-1752



A.L. Cauchy  
(1789-1857)



J. Binet  
(1786-1856)

# Normes des résidus pour les matrices diagonalisables

Soit  $A = X\Lambda X^{-1}$  et  $c = X^{-1}b$ . La matrice de Krylov est

$$K = X \begin{pmatrix} c & \Lambda c & \cdots & \Lambda^{n-1}c \end{pmatrix}$$

Donc puisque  $M = U^*U$  et  $K = VU$

$$M = K^*K = \begin{pmatrix} c & \Lambda c & \cdots & \Lambda^{n-1}c \end{pmatrix}^* X^*X \begin{pmatrix} c & \Lambda c & \cdots & \Lambda^{n-1}c \end{pmatrix}$$

et

$$M_{k+1} = \mathcal{V}_{k+1}^* D_c^* X^* X D_c \mathcal{V}_{k+1}$$

$D_c$  est diagonale avec  $c_i$  sur la diagonale et ...

$$\mathcal{V}_{k+1} = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^k \\ 1 & \lambda_2 & \cdots & \lambda_2^k \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^k \end{pmatrix}$$

une matrice de **Vandermonde**  $n \times (k+1)$

En utilisant la formule de **Cramer** on a

$$\begin{aligned} (M_{k+1}^{-1})_{1,1} &= \frac{1}{\det(M_{k+1})} \det \begin{pmatrix} 1 & m_{1,2} & \cdots & m_{1,k} \\ 0 & & & \\ \vdots & & M_{2:k+1,2:k+1} & \\ 0 & & & \end{pmatrix} \\ &= \frac{\det(M_{2:k+1,2:k+1})}{\det(M_{k+1})} \end{aligned}$$



Soit  $F = XD_c \mathcal{V}_{k+1}$ , une matrice  $n \times (k+1)$ . Alors  $M_{k+1} = F^* F$  et par la formule de Cauchy-Binet

$$\det(M_{k+1}) = \sum_{I_{k+1}} |\det(F_{I_{k+1},:})|^2$$

la somme est sur tout les ensembles de  $k+1$  indices de ligne  $I_{k+1}$   $(i_1, i_2, \dots, i_{k+1})$  tels que  $1 \leq i_1 < \dots < i_{k+1} \leq n$  et  $F_{I_{k+1},:}$  est la sous-matrice de  $F$  dont les indices de ligne appartiennent à  $I_{k+1}$

$$F_{I_{k+1},:} = (XD_c)_{I_{k+1},:} \mathcal{V}_{k+1}, \quad [(k+1) \times n] * [n \times (k+1)]$$

On peut appliquer la formule de Cauchy-Binet une deuxième fois

$$\det(F_{I_{k+1},:}) = \sum_{J_{k+1}} \det(X_{I_{k+1},J_{k+1}}) c_{j_1} \cdots c_{j_{k+1}} \det(\mathcal{V}(\lambda_{j_1}, \dots, \lambda_{j_{k+1}}))$$

avec la matrice (carrée) de **Vandermonde**

$$\mathcal{V}(\lambda_{j_1}, \dots, \lambda_{j_{k+1}}) = \begin{pmatrix} 1 & \lambda_{j_1} & \cdots & \lambda_{j_1}^k \\ 1 & \lambda_{j_2} & \cdots & \lambda_{j_2}^k \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_{j_{k+1}} & \cdots & \lambda_{j_{k+1}}^k \end{pmatrix}$$

Mais  $\det(\mathcal{V}(\lambda_{j_1}, \dots, \lambda_{j_{k+1}}))$  est connu

On applique la même technique pour calculer  $\det(M_{2:k+1,2:k+1})$

Soit  $A$  une matrice diagonalisable avec  $A = X \Lambda X^{-1}$ ,  $c = X^{-1}b$ .

Alors

$$\|r_k^G\|^2 = \sigma_{k+1}^N / \sigma_k^D$$

avec

$$\sigma_{k+1}^N = \sum_{I_{k+1}} \left| \sum_{J_{k+1}} \det(X_{I_{k+1}, J_{k+1}}) c_{j_1} \cdots c_{j_{k+1}} \prod_{j_1 \leq j_l < j_p \leq j_{k+1}} (\lambda_{j_p} - \lambda_{j_l}) \right|^2$$

$$\sigma_1^D = \sum_{i=1}^n \left| \sum_{j=1}^n X_{i,j} c_j \lambda_j \right|^2$$

et

$$\sigma_k^D = \sum_{I_k} \left| \sum_{J_k} \det(X_{I_k, J_k}) c_{j_1} \cdots c_{j_k} \lambda_{j_1} \cdots \lambda_{j_k} \prod_{j_1 \leq j_l < j_p \leq j_k} (\lambda_{j_p} - \lambda_{j_l}) \right|^2, \quad k > 1$$

Si la matrice  $A$  est normale, on a  $X^*X = I$ ,  $c = X^*b$  et des formules plus simples

$$\sigma_{k+1}^N = \sum_{l_{k+1}} |c_{j_1}|^2 \cdots |c_{j_{k+1}}|^2 \prod_{j_1 \leq j_l < j_p \leq j_{k+1}} |(\lambda_{j_p} - \lambda_{j_l})|^2$$

$$\sigma_1^D = \sum_{i=1}^n |c_j|^2 |\lambda_j|^2$$

et

$$\sigma_k^D = \sum_{l_k} |c_{j_1}|^2 \cdots |c_{j_k}|^2 |\lambda_{j_1}|^2 \cdots |\lambda_{j_k}|^2 \prod_{j_1 \leq j_l < j_p \leq j_k} |(\lambda_{j_p} - \lambda_{j_l})|^2, \quad k > 1$$

Voir J. Duintjer Tebbens, GM, H. Sadok and Z. Strakoš, LAA v 450 (2014)

On peut étendre en partie ces résultats aux matrices non diagonalisables en utilisant la forme canonique de [Jordan](#)

# Bornes pour les normes des résidus (matrices diagonalisables)

En utilisant un résultat de [Bellalij](#), [Jbilou](#) et [Sadok](#) on peut montrer que

$$\|r_k\|^2 \leq \frac{\|X\|^2}{e_1^T (\mathcal{V}_{k+1}^* D_c^* D_c \mathcal{V}_{k+1})^{-1} e_1}$$

On obtient

$$\|r_k\|^2 \leq \|X\|^2 \frac{\sum_{1 \leq i_1 < \dots < i_{k+1} \leq n} \prod_{j=1}^{k+1} \omega_{ij} \prod_{i_1 \leq i_\ell < i_p \leq i_{k+1}} |\lambda_{i_p} - \lambda_{i_\ell}|^2}{\sum_{1 \leq i_1 < \dots < i_k \leq n} \prod_{j=1}^k \omega_{ij} |\lambda_{ij}|^2 \prod_{i_1 \leq i_\ell < i_p \leq i_k} |\lambda_{i_p} - \lambda_{i_\ell}|^2}$$

où  $\omega_j = |c_j|^2$  avec  $c = X^{-1}b$

On peut aussi obtenir une borne inférieure similaire où le facteur multiplicatif est  $\sigma_{\min}(X)^2$  (min des valeurs singulières)

Ainsi, les normes des résidus dépendent des (différences des) valeurs propres, des vecteurs propres et du second membre (par  $X^{-1}b$ )

Si  $\sigma_{\min}(X) \approx \|X\|$ , la convergence dépend essentiellement des valeurs propres (et du second membre) comme pour les matrices normales

Si on prescrit la courbe de convergence et que l'on veut changer les valeurs propres, il faut aussi modifier les vecteurs propres

# Extension de GMRES

Flexible GMRES (Saad) avec un préconditionnement pouvant varier à chaque itération

$x_0$  donné,  $r_0 = b - Ax_0$

$$v_1 = \frac{r_0}{\|r_0\|}, \quad f = \|r_0\| e_1$$

for  $k = 1, \dots$

$$M_k z_k = v_k, \quad w = Az_k$$

for  $i = 1, \dots, k$

$$h_{i,k} = (w, v_i), \quad w = w - h_{i,k} v_i$$

end  $i$

$$h_{k+1,k} = \|w\|, \quad v_{k+1} = \frac{w}{h_{k+1,k}}$$

On applique les rotations des itérations 1 à  $k - 1$  sur

$$(h_{1,k} \dots h_{k+1,k})^T$$

On calcule la rotation  $R_{k+1,k}$  pour éliminer  $h_{k+1,k}$ ,  $f = R_{k+1,k}f$ ,  
on résout le système triangulaire pour  $y_k$  et

$$x_k = x_0 + Z_k y_k$$

avec  $Z_k = [z_1 \dots z_k]$

end  $k$

Notons qu'il faut stocker les vecteurs  $z_k$

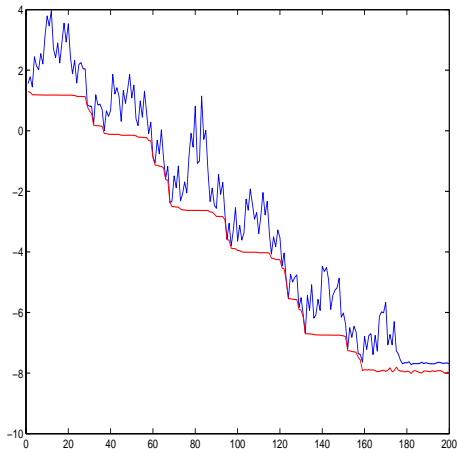


# GMRES en précision finie

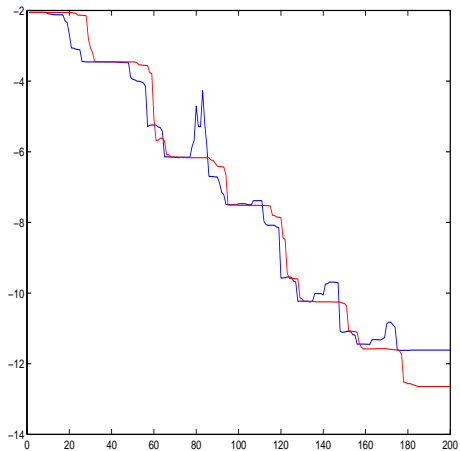
(MGS)-GMRES possède la stabilité inverse ([Paige, Rozloznik et Strakoš](#), SIMAX 2006)

## Exemples

Matrice Steam2 : modèle de réservoir pétrolier,  $n = 600$ ,  
 $\kappa(A) = 3.78 \cdot 10^6$

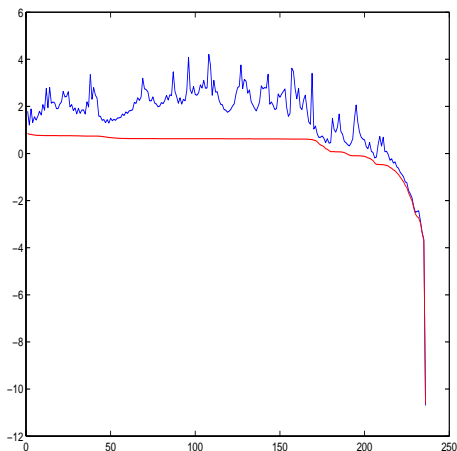


Steam2 :  $\|r_k^F\|$  (bleu) et  $\|r_k^G\|$  (rouge)

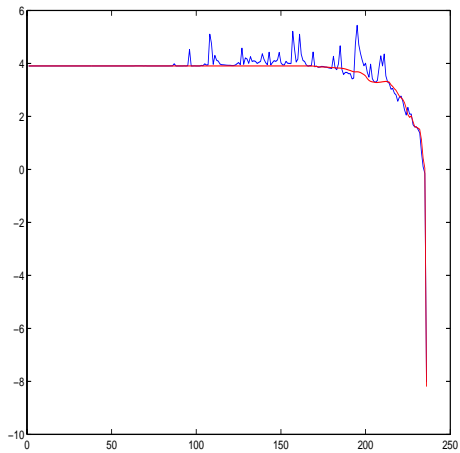


Steam2 :  $\|\varepsilon_k^F\|$  ( bleu ) et  $\|\varepsilon_k^G\|$  ( rouge )

Matrice e05r0500 : driven cavity flow, Reynolds 500,  $n = 236$ ,  
 $\kappa(A) = 1.16 \cdot 10^6$



e05r0500 :  $\|r_k^F\|$  (bleu) et  $\|r_k^G\|$  (rouge)



e05r0500 :  $\|\varepsilon_k^F\|$  (bleu) et  $\|\varepsilon_k^G\|$  (rouge)

$$-\nu \Delta u + w \cdot \nabla u = 0$$

avec  $w = [0, 1]^T$  dans  $\Omega = (0, 1)^2$ , cond. lim. Dirichlet  $u = g$  sur  $\partial\Omega$

Méthode de Petrov–Galerkin SUPG avec éléments finis bilinéaires sur un maillage cartésien régulier

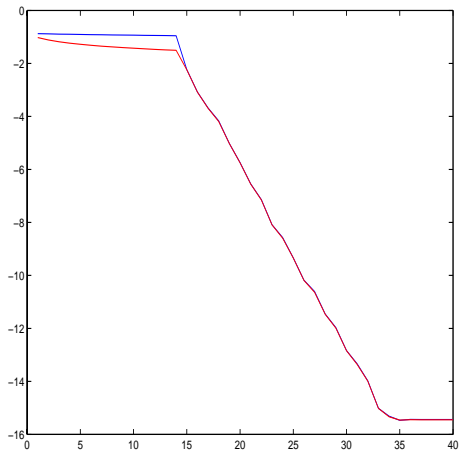
$$A = \nu K \otimes M + M \otimes ((\nu + \delta h)K + C)$$

avec  $\delta$  paramètre de régularisation, pas du maillage  $h$  et

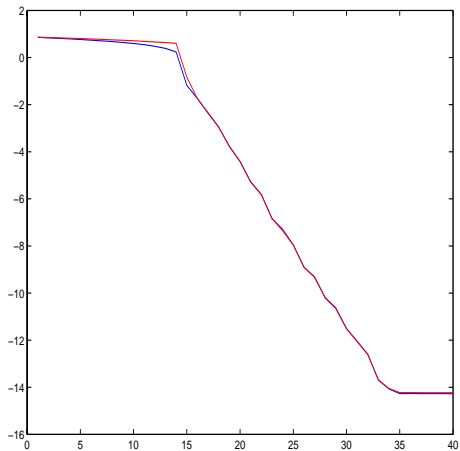
$$M = \frac{h}{6} \text{tridiag}(1, 4, 1)$$

$$K = \frac{1}{h} \text{tridiag}(-1, 2, -1), \quad C = \frac{1}{2} \text{tridiag}(-1, 0, -1)$$

$h = 1/16$ ,  $\nu = 0.01$  et  $\delta = 0.34$ . Dimension du système 225



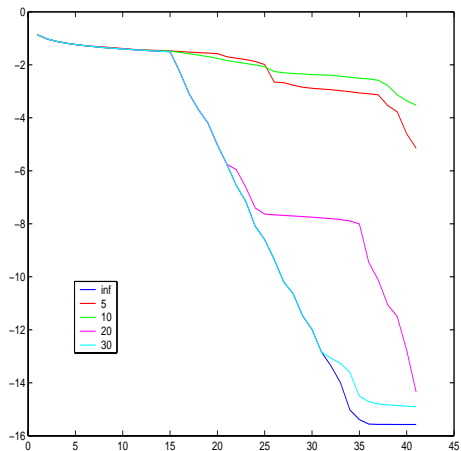
Supg :  $\|r_k^F\|$  (bleu) et  $\|r_k^G\|$  (rouge)



Supg :  $\|\varepsilon_k^F\|$  (bleu) et  $\|\varepsilon_k^G\|$  (rouge)



# GMRES(m)

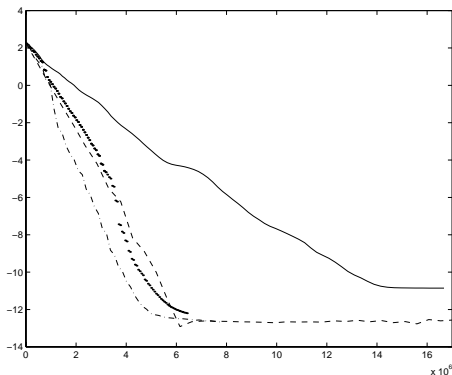


Supg :  $\|r_k^G\|$  pour différentes valeurs de  $m$

Mais, il faut regarder le temps de calcul

## Autre problème

Equation de convection-diffusion  $-\Delta u + 2e^{2(x^2+y^2)} \frac{\partial u}{\partial x} = f$ , schéma upwind



nb de flops, GMRES (plein), GMRES(20) (tirets), GMRES(10) (point-tirets) and GMRES(5) (points)

# Méthodes Q-OR et Q-MR

Peut-on étendre ces résultats aux méthodes Q-OR et Q-MR ?

Supposons que nous ayons une base ascendante  $V$  de l'espace de Krylov (avec des colonnes de norme 1) telle que  $K = VU$  avec  $V$  non singulière et  $U$  triangulaire supérieure

On définit  $H = UCU^{-1}$ . En conséquence  $AV = VH$ . Les itérés (avec  $x_0 = 0$ ) sont

$$x_k = V_k y_k$$

où  $V_k$  est la matrice des  $k$  premières colonnes de  $V$ . Le résidu  $r_k$  est

$$V_k e_1 - AV_k y_k = V_k (e_1 - H_k y_k) - h_{k+1,k} (y_k)_k v_{k+1} = V_{k+1} (e_1 - \underline{H}_k y_k)$$

La méthode Q-OR est définie (pourvu que  $H_k$  soit non singulière) par

$$H_k y_k^O = e_1$$

où  $H_k$  est la sous-matrice principale d'ordre  $k$ . Ce choix annule le premier terme du résidu

Dans la méthode Q-MR le vecteur  $y_k^M$  est calculé comme solution du problème de moindres carrés

$$\min_y \|e_1 - \underline{H}_k y\|$$

où  $\underline{H}_k$  est  $(k+1) \times k$

Le vecteur  $z_k^M = e_1 - \underline{H}_k y_k^M$  s'appelle le quasi-résidu

Le résidu est  $r_k^M = V_{k+1} z_k^M$

On résout les deux problèmes avec des rotations de **Givens** de sinus  $s_j$ . On sait que

$$\|z_k^M\| = |s_1 s_2 \cdots s_k|$$

De plus on a une relation entre les normes des résidus de **Q-OR** et les normes des quasi-résidus de **Q-MR**

$$\frac{1}{\|r_k^O\|^2} = \frac{1}{\|z_k^M\|^2} - \frac{1}{\|z_{k-1}^M\|^2}$$

Voir Eiermann et Ernst (2001), Freund et Nachtigal (1991)

On montre que

$$|(U^{-1})_{1,k}| = \frac{1}{\|r_{k-1}^O\|}$$

Soit  $M_{k+1} = U_{k+1}^* U_{k+1}$ . Puisque  $M_{k+1}^{-1} = U_{k+1}^{-1} U_{k+1}^{-*}$  et d'après la première ligne de  $U^{-1}$ , une conséquence de ce résultat est que

$$\|z_k^M\|^2 = \frac{1}{(M_{k+1}^{-1})_{1,1}}$$

La différence avec **GMRES** est que l'on a seulement une caractérisation de la norme du quasi-résidu

On obtient ainsi des expressions pour les normes des quasi-résidus. Il suffit de remplacer  $X$  par  $Z = V^{-1}X$

# Lanczos non symétrique

Comment calcule t-on d'autres bases de l'espace de Krylov ?

Pour les matrices non symétriques :

minimisation  $\neq$  orthogonalisation

réurrences courtes  $\longrightarrow$  pas d'orthogonalité des résidus

## Méthode de Lanczos non symétrique

on choisit deux vecteurs  $v_1$  et  $\tilde{v}_1$  tels que  $(v_1, \tilde{v}_1) = 1$

$v_{-1} = \tilde{v}_{-1} = 0$

# Algorithme Lanczos non symétrique

for  $k = 0, 1, \dots$

$$z_k = Av_k - \delta_k v_k - \eta_k v_{k-1}, \quad w_k = A^T \tilde{v}_k - \delta_k \tilde{v}_k - \tilde{\eta}_k \tilde{v}_{k-1}$$

$$\delta_k = (\tilde{v}_k, Av_k)$$

$$\eta_{k+1} \tilde{\eta}_{k+1} = (z_k, w_k)$$

$$v_{k+1} = \frac{z_k}{\tilde{\eta}_{k+1}}, \quad \tilde{v}_{k+1} = \frac{w_k}{\eta_{k+1}}$$

Les vecteurs  $v_k$  et  $\tilde{v}_k$  sont bi-orthogonaux



$$T_k = \begin{pmatrix} \delta_1 & \eta_2 & & & \\ \tilde{\eta}_2 & \delta_2 & \eta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \tilde{\eta}_{k-1} & \delta_{k-1} & \eta_k \\ & & & \tilde{\eta}_k & \delta_k \end{pmatrix}$$

$$V_k = [v_1 \ \cdots \ v_k], \quad \tilde{V}_k = [\tilde{v}_1 \ \cdots \ \tilde{v}_k]$$

$$AV_k - V_k T_k = \tilde{\eta}_{k+1} v_{k+1} (e_k)^T$$

$$A^T \tilde{V}_k - \tilde{V}_k T_k^T = \eta_{k+1} \tilde{v}_{k+1} (e_k)^T$$

# Gradient BiConjugué (BiCG)

On utilise la base bi-orthogonale en construisant des résidus bi-orthogonaux  $r_k$  et  $\tilde{r}_k$

$$r_{k+1} = r_k - \alpha_k A p_k$$

$$\tilde{r}_{k+1} = \tilde{r}_k - \alpha_k A^T \tilde{p}_k$$

$$p_{k+1} = r_k + \beta_{k+1} p_k$$

$$\tilde{p}_{k+1} = \tilde{r}_k + \beta_{k+1} \tilde{p}_k$$

$$\alpha_k = \frac{(\tilde{r}_k, r_k)}{(\tilde{p}_k, A p_k)}, \quad \beta_{k+1} = \frac{(\tilde{r}_{k+1}, r_{k+1})}{(\tilde{r}_k, r_k)}$$

$$\implies (\tilde{r}_k, r_\ell) = 0, \quad (\tilde{p}_k, A p_\ell) = 0, \quad k \neq \ell$$

# Problèmes avec BiCG

- ▶ dégénérescences  $(\tilde{r}_k, r_k) = 0, \tilde{r}_k \neq 0, r_k \neq 0$
- ▶ Remèdes : algorithmes de Look Ahead (Freund, Brezinski)
- ▶ comportement erratique de la norme des résidus

# Quasi-résidu minimum (QMR)

On utilise l'algorithme de Lanczos non symétrique

$$AV_k = V_{k+1}\underline{T}_k$$

$$\begin{aligned}\|r_k\| &= \|b - Ax_k\| \\ &= \|b - AV_k y_k\| \\ &= \|\|b\| V_{k+1} e_1 - V_{k+1} \underline{T}_k y_k\|\end{aligned}$$

Malheureusement  $V_{k+1}$  n'est pas orthogonale  
QMR minimise

$$\|\|b\| e_1 - \underline{T}_k y\|$$

# Construction de systèmes avec une courbe de convergence prescrite

Peut-on construire des systèmes linéaires avec une courbe de convergence prescrite pour les méthodes Q-OR et Q-MR ?

Pour FOM/GMRES c'est facile puisqu'il suffit de construire une matrice triangulaire supérieure  $U^{-1}$  avec les inverses des normes des résidus de FOM (obtenus avec les normes de GMRES) sur la première ligne. Ensuite

$$A = VUCU^{-1}V^*, \quad b = Ve_1$$

où  $C$  est la matrice compagnon des valeurs propres prescrites et  $V$  est une matrice unitaire quelconque

Les choses sont plus difficiles pour d'autres méthodes Q-OR/Q-MR parce qu'il peut être nécessaire d'imposer une structure donnée à  $H$

# BiCG

Peut-on construire des systèmes linéaires avec une courbe de convergence prescrite pour BiCG ?

On veut construire des matrices  $H$  tridiagonale (avec un spectre donné) et  $U$  triangulaire supérieure telles que

$$H = \begin{pmatrix} \gamma_1 & \beta_2 & 0 & 0 & 0 \\ \rho_2 & \gamma_2 & \beta_3 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \rho_{n-1} & \gamma_{n-1} & \beta_n \\ 0 & 0 & 0 & \rho_n & \gamma_n \end{pmatrix} = UCU^{-1}$$

et la première ligne de  $U^{-1}$  est prescrite :  $(1 \quad g_1 \quad \cdots \quad g_{n-1})$   
avec  $g_j \neq 0$

Soient  $\omega_2, \dots, \omega_n$  les valeurs choisies arbitrairement de la dernière colonne de  $U^{-1}$  avec  $\omega_n \neq 0$ ,  $\omega_1 = g_{n-1}$  et  $-\alpha_0, \dots, -\alpha_{n-1}$  les composantes de la dernière colonne de  $C$  que l'on connaît d'après le spectre

On calcule  $U^{-1}$  et  $H$  colonne par colonne

La dernière colonne de  $U^{-1}H = CU^{-1}$  donne

$$\begin{pmatrix} g_{n-2}\beta_n + g_{n-1}\gamma_n \\ \vdots \\ \omega_n\gamma_n \end{pmatrix} = \begin{pmatrix} -\alpha_0\omega_n \\ \vdots \\ \omega_{n-1} - \alpha_{n-1}\omega_n \end{pmatrix}$$

On utilise la première et la dernière équations

$$\begin{pmatrix} g_{n-2} & g_{n-1} \\ 0 & \omega_n \end{pmatrix} \begin{pmatrix} \beta_n \\ \gamma_n \end{pmatrix} = \begin{pmatrix} -\alpha_0\omega_n \\ \omega_{n-1} - \alpha_{n-1}\omega_n \end{pmatrix}$$

La solution de ce système  $2 \times 2$  non singulier donne  $\gamma_n, \beta_n$

Avec les autres équations on peut calculer les inconnues  $\nu_{j,n-1}$  de la colonne  $n-1$  de  $U^{-1}$

On considère ensuite la colonne  $n - 1$

On a trois inconnues  $\beta_{n-1}, \gamma_{n-1}$  et  $\rho_n$

On prend la première et les deux dernières équations qui donnent le système linéaire

$$\begin{pmatrix} g_{n-3} & g_{n-2} & g_{n-1} \\ 0 & \nu_{n-1,n-1} & \omega_{n-1} \\ 0 & 0 & \omega_n \end{pmatrix} \begin{pmatrix} \beta_{n-1} \\ \gamma_{n-1} \\ \rho_n \end{pmatrix} = \begin{pmatrix} 0 \\ \nu_{n-2,n-1} \\ \nu_{n-1,n-1} \end{pmatrix}$$

Et ainsi de suite... On construit  $A = VHV^{-1}$  et  $b = Ve_1$  pour une matrice  $V$  bien choisie

Pour l'instant on ne sait pas traiter le cas où l'on veut imposer des zéros sur la première ligne de  $U^{-1}$

On peut étendre ce procédé avec une plus grande largeur de bande

On peut donc prescrire pour BiCG n'importe quelle courbe de convergence (finie) pour les normes des résidus (ou pour les normes des quasi-résidus dans le cas de QMR)



# Améliorations de BiCG

BiCG présente des inconvénients : oscillations des normes des résidus et nécessité de multiplier par  $A^T$

Evolution: CGS (Sonneveld) (élimine le second problème)

Améliorations: BiCGstab, BiCGstab( $\ell$ ) (Van Der Vorst and co)  
(éliminent -en partie- le premier problème)

Notons que BiCGstab n'entre pas dans le cadre des méthodes Q-OR/Q-MR utilisé précédemment

Pour BiCG

$$p_k = r_{k-1} - \beta_k p_{k-1}, \quad x_{k+1} = x_k + \alpha_k p_k, \quad r_{k+1} = r_k - \alpha_k A p_k$$

avec  $r_k$  et  $A p_k$  orthogonaux à  $\mathcal{K}_k(A^T, \tilde{r}_0)$

Pour tout polynôme  $\psi_k$  de degré  $k$  et de coefficient principal  $\theta_k$

$$\beta_k = \frac{\theta_{k-1}}{\theta_k} \frac{\rho_k}{\sigma_{k-1}}, \quad \alpha_k = \frac{\rho_k}{\sigma_k}$$

où

$$\rho_k = (r_k, \psi_k(A^T) \tilde{r}_0) = (y_k, \tilde{r}_0)$$

avec  $y_k = \psi_k(A) r_k$

$$\sigma_k = (Ap_k, \psi_k(A^T)\tilde{r}_0)$$

On peut choisir librement le polynôme  $\psi_k$  à condition que

$$\psi_0(A^T)\tilde{r}_0, \dots, \psi_{k-1}(A^T)\tilde{r}_0$$

engendre l'espace de Krylov  $\mathcal{K}_k(A^T, \tilde{r}_0)$

H. Van der Vorst considère des résidus de la forme

$$r_k = \psi_k(A)\phi_k(A)r_0$$

où  $\phi$  est le polynôme de BiCG

Pour  $\psi$  on choisit le produit de polynômes de degré 1 de résidu minimum

$$\psi_{k+1}(t) = (1 - \omega_k t) \cdots (1 - \omega_1 t)$$

le polynôme  $\psi_k$  peut être calculé par récurrence

$$\psi_{k+1}(t) = (1 - \omega_k t)\psi_k(t)$$

On a des récurrences pour  $\psi_k \phi_k$  et  $\psi_k \theta_k$

$$r_k = \psi_k(A) \phi_k(A) r_0$$

$$p_k = \psi_k(A) \theta_k(A) r_0$$

Les paramètres  $\omega_k$  sont choisis pour minimiser la norme du résidu

On note que l'on a deux multiplications matrice-vecteur par itération

# BiCGstab

Soit  $x_0$  donné,  $r_0 = b - Ax_0$ ,  $p_0 = r_0$ ,  $\tilde{r}_0$  arbitraire,  
for  $k = 0, 1, \dots$

$$\alpha_k = \frac{(r_k, \tilde{r}_0)}{(Ap_k, \tilde{r}_0)}$$

$$s_k = r_k - \alpha_k Ap_k$$

$$\omega_k = \frac{(As_k, s_k)}{(As_k, As_k)}$$

$$x_{k+1} = x_k + \alpha_k p_k + \omega_k s_k$$

$$r_{k+1} = s_k - \omega_k As_k$$

$$\beta_{k+1} = \frac{(r_{k+1}, \tilde{r}_0)}{(r_k, \tilde{r}_0)} \frac{\alpha_k}{\omega_k}$$

$$p_{k+1} = r_{k+1} + \beta_{k+1}(p_k - \omega_k Ap_k)$$

# IDR

Dernier avatar : IDR(s), Induced Dimension Reduction (Sonneveld et Van Gijzen, 2007)

basé sur une idée de Sonneveld (1980) et sur le résultat suivant :

Soient  $A$  d'ordre  $n$  et  $\mathcal{G}_0 = \mathcal{K}_n(A, v_0)$  espace de Krylov d'ordre  $n$ ,  $\mathcal{S}$  un sous espace tel que  $\mathcal{S}$  et  $\mathcal{G}_0$  ne contiennent pas un même sous espace invariant de  $A$ . On définit

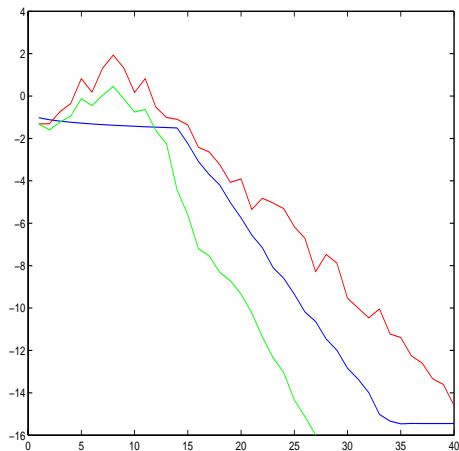
$$\mathcal{G}_j = (I - \omega_j A)(\mathcal{G}_{j-1} \cap \mathcal{S})$$

Alors

$$\mathcal{G}_j \subset \mathcal{G}_{j-1}, j = 1, 2, \dots$$

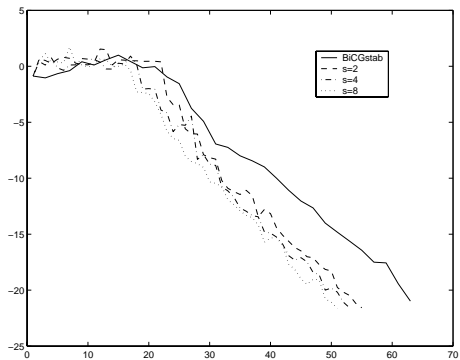
et il existe un indice  $j$  tel que  $\mathcal{G}_j = \{0\}$ ,  $j \leq n$

# SUPG

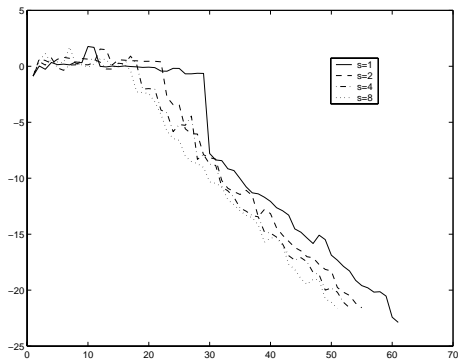


SUPG :  $\log_{10}(\|r_k\|)$ , GMRES (bleu), BiCG (rouge), BiCGStab (vert)

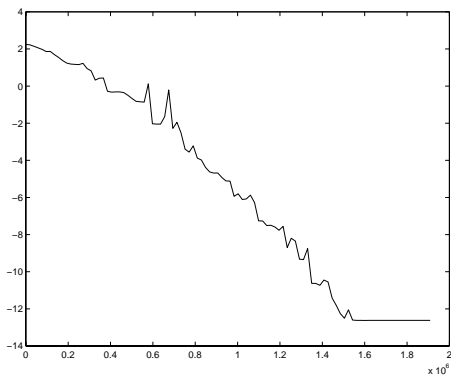




SUPG :  $\log_{10}(\|r_k\|)$ , BiCGstab et IDR(s)  
 nb de produits matrice-vecteur



SUPG :  $\log_{10}(\|r_k\|)$ , IDR(s)  
 nb de produits matrice-vecteur



Pb de convection-diffusion, BiCGstab  
nb de flops