Review of the convergence of some Krylov subspace methods for solving linear systems of equations with one or several right hand sides

Hassane SADOK

NASCA18, Kalamata, Jully 2018

BLOCK SUBSPACE KRYLOV METHODS

We consider the linear systems of equations with multiple right-hand sides

$$AX = F, \quad A \in \mathbb{C}^{n\times n}, \quad F \in \mathbb{C}^{n\times s}, \quad X \in \mathbb{C}^{n\times s}, \quad 1 \le s \ll n. \tag{1}$$

Which is equivalent to the s-linear systems of equations

$$Ax^{(i)} = f^{(i)}, \quad \text{for} \quad i = 1, \ldots, s, \tag{2}$$

➠ To solve problem (1), several block methods have been developed.

- Standard GMRES (Classical GMRES)
- Block GMRES
- Global GMRES

Reference

- L. Elbouyahyaoui, A. Messaoudi and H. Sadok, Algebraic Properties of Block Arnoldi algorithm and Block GMRES method, Elect. Trans. Num. Anal., 33 (2009) pp. 207–220.

definition

Consider the linear system of equations

$$A\,x^{(i)} = f^{(i)}$$

.

- For the GMRES method, the iterates $\{x_k^{(i)}\}$ are defined by the following conditions

Standard-GMRES

$$x_k^{(i)} - x_0^{(i)} \in \mathcal{K}_k(A, r_0^{(i)}),$$

$$(A^j\, r_0^{(i)}, r_k^{(i)}) = 0 \qquad \text{for} \quad j = 1, \dots, k,$$

where $r_0^{(i)} = f^{(i)} - A\,x_0^{(i)}$

reference

Y. Saad, Iterative methods for sparse linear systems, PWS. Publishing Company (1996).

definition : Matrix Krylov subspace

Let

$$\mathbf{K}_k^G(A,U) = span\{U, A\,U, \ldots, A^{k-1}\,U\} \subset \mathbb{C}^{n\times s}$$

denote the matrix Krylov subspace spanned by the matrices $U, AU, \ldots, A^{k-1}U$, where U is an $n \times s$ matrix. Note that $Z \in \mathbf{K}_k^G(A,U)$ implies that

$$Z \;=\; \sum_{j=1}^{k} \alpha_j A^{j-1} U, \quad \alpha_j \in \mathbb{C}, \quad j = 1, \ldots, k.$$

Reference

- K. Jbilou, A. Messaoudi and H. Sadok, Global GMRES algorithm for solving nonsymmetric linear systems of equations with multiple right-hand sides. Applied Num. Math., 31 (1999) pp. 49–63.

Global-GMRES

The global GMRES method constructs, at step k, the approximation X_k satisfying the following two relations

Global-GMRES

$$X_k - X_0 \in \mathbf{K}_k^G(A, R_0) \text{ and } \langle A^j R_0, R_k \rangle_F = 0, \qquad \text{for } j = 1, \ldots, k,$$

where we define the inner product $\langle Y, Z \rangle_F = \text{trace}(Y^H Z)$ (where Y^H denotes the conjugate transpose of Y). The associated norm is the Frobenius norm $\| \,.\, \|_F$.
We can also use the following inner product

$$\langle Y, Z \rangle = \sum_{1 \leq i,j \leq s} Y_i^H Z_j$$

Convergence

If d is the degree of the minimal polynomial of A with respect to R_0, then

$$R_d = 0.$$

,

Global-GMRES

The residual $R_k = F - AX_k$ satisfies the minimization property

$$\|R_k\|_F = \min_{Z \in \mathbf{K}_k^G(A, R_0)} \|R_0 - AZ\|_F. \tag{3}$$

Definition of BGMRES method

$\mathbb{K}_k(A, R_0)$ denote the matrix Krylov subspace defined as

$$\begin{aligned} \mathbb{K}_k(A, R_0) &= blockspan\{R_0, AR_0, \ldots, A^{k-1}R_0\} && (4) \\ &= \left\{ \sum_{i=1}^{k} A^k R_0 \Omega_i / \Omega_i \in \mathbb{C}^{s\times s} \right\}. && (5) \end{aligned}$$

Let $\mathcal{B}_k(A, R_0) = \sum_{i=1}^{s} \mathcal{K}_k(A, r_0^{(i)}),$

☞ For $i = 1, \ldots, s$ we have $x_k^{(i)} \in x_0^{(i)} + \mathcal{B}_k(A, R_0)$.

reference

M.H. Gutknecht, Block Krylov space methods for linear systems with multiple right-hand sides : an introduction, in Modern Mathematical Models, Methods and Algorithms for Real World Systems, A. H. Siddiqi, I. S. Duff, and O. Christensen, eds., Anamaya Publishers, New Delhi, 2005, pp. 420–447.

For nonsymmetric problems, the BGMRES [Vital 1990] generate an approximate solution X_k over the matrix krylov subspace $\mathbb{K}_k(A, R_0)$ with orthogonality condition on the residual R_k.

Block GMRES

$$\begin{cases} X_k - X_0 \in \mathbb{K}_k(A, R_0) \\ R_k = F - AX_k \perp A\mathbb{K}_k(A, R_0), \end{cases}$$

Reference

- B. Vital Etude de quelques méthodes de résolution de problèmes linéaires de grande taille sur multiprocesseur, Ph.D. thesis, Université de Rennes, Rennes, France, 1990.
- V. Simoncini and E. Gallopoulos, *Convergence properties of block GMRES and matrix polynomials*, Linear Algebra Appl., 247(1996), pp. 97-119.
- V. Simoncini and E. Gallopoulos, *An Iterative Method for Nonsymmetric Systems with Multiple Right-hand Sides*, SIAM J. Sci. Comp., 16(1995), pp. 917-933.

Block GMRES Algorithm

The definition of BGMRES is equivalent to

$$X_k = X_0 + Z_k$$
$$\| R_0 - AZ_k \|_F = \min_{Z \in \mathbb{K}_k(A, R_0)} \| R_0 - AZ \|_F \ . \qquad (6)$$

Thus BGMRES method proceeds as follows

ALGORITHM (BGMRES)

1. *Choose* $X_0 \in \mathbb{C}^{N \times s}$ *and compute* $R_0 = F - AX_0$.
2. $R_0 = V_1 H_{1,0}$*(The QR factorisation of* R_0*);*
3. *For j=1, ... ,k, do*
   construct V_j *and* $\widetilde{\mathbb{H}}_j$ *by block Arnoldi.*
4. *Solve the least squares problem :*

$$Y_k = arg \min_{Y \in \mathbb{C}^{ks \times s}} \| R_0 - A\mathbb{V}_k Y \|_F,$$

5. *The approximate solution is* $X_k = X_0 + \mathbb{V}_k Y_k$.

Let U be an $N \times s$ matrix. An Arnoldi-type algorithm constructs a basis $\{V_1^\bullet, \ldots, V_k^\bullet\}$ of $\mathbf{K}_k^B(A, U)$, which satisfies an orthogonal property. Moreover the block matrix $\mathcal{V}_k^\bullet = [V_1^\bullet, \ldots, V_k^\bullet]$ is such that the matrix $\mathcal{H}_k^\bullet = \mathcal{V}^\bullet{}_k^H A \mathcal{V}^\bullet{}_k$ is an upper block Hessenberg.

We examine three possibly choices of orthogonality. Let $\Phi^\bullet$ be the map : $\mathbb{C}^{n\times s} \times \in \mathbb{C}^{n\times s} \longrightarrow \mathbb{C}^{s\times s}$ defined for $\bullet \in \{B, S, G\}$ by

$$\left\{ \begin{array}{lcl} \Phi^B(X,Y) & = & X^H Y, \\ \Phi^S(X,Y) & = & \text{the diagonal of the matrix } (X^H Y), \\ \Phi^G(X,Y) & = & \text{trace}(X^H Y) I_s = \langle X, Y \rangle_F I_s, \end{array} \right.$$

$\forall X \in \mathbb{C}^{n\times s}$ and $\forall Y \mathbb{C}^{n\times s}$.

If $\Phi^B(X,Y) = X^H Y = 0$, then the block-vectors X, Y are called block-orthogonal by Gutknecht. Similarly X is called block-normalized if $X^H X = I_s$. Of course the vector space of $\mathbb{C}^{n\times s}$ of block vectors is an Euclidean space, which is a finite-dimensional inner product space with the inner product : $\langle X, Y \rangle_F = \text{trace}(X^H Y)$. If $\langle X, Y \rangle_F = 0$, then X and Y are called F-orthogonal. If $\Phi^S(X,Y) = diag(X^H Y) = 0$, then X and Y can be called diagonally orthogonal.

The Block Arnoldi-type algorithm

1. Let U be an $n \times s$ matrix.
2. Compute $V_1^\bullet \in \mathbb{C}^{n\times s}$ by determining the factorization of U : $U = V_1^\bullet H_{1,0}^\bullet, H_{1,0}^\bullet \in \mathbb{C}^{s\times s}$, such that $H_{1,0}^\bullet = \Phi^\bullet(V_1^\bullet, U)$ and $\Phi^\bullet(V_1^\bullet, V_1^\bullet) = I_s$.
3. for i=1,...,k do
   - Compute $W = AV_i^\bullet$.
   - for j=1,...,i do
     1. $H_{j,i}^\bullet = \Phi^\bullet(V_j^\bullet, W)$
     2. $W = W - V_j^\bullet H_{j,i}^\bullet$
   - End
   - Compute $H_{i+1,i}^\bullet$ by determining the decomposition of W : $W = V_{i+1}^\bullet\, H_{i+1,i}^\bullet$, such that $H_{i+1,i}^\bullet = \Phi^\bullet(V_{i+1}^\bullet, W)$ and $\Phi^\bullet(V_{i+1}^\bullet, V_{i+1}^\bullet) = I_s$.
4. End

For $\Phi^{\bullet}(X,Y) = \Phi^{B}(X,Y) = X^{H}Y$, the preceding algorithm reduces to Block Arnoldi, which builds an orthonormal basis $\{V_1^B, \ldots, V_k^B\}$ such that the block matrix $\mathcal{V}_k^B = [V_1^B, \ldots, V_k^B]$ satisfies $(\mathcal{V}_k^B)^H \mathcal{V}_k^B = I_{ks}$.
It is well known that

$$A\mathcal{V}_k^B = \mathcal{V}_k^B \mathcal{H}_k^B + V_{k+1}^B H_{k+1,k}^B E_k^T, \tag{7}$$

where $E_k^T = [0_s, \ldots, O_s, I_s] \in \mathbb{R}^{s \times ms}$.

$$\widetilde{\mathbb{H}}_k = \begin{pmatrix} H_{1,1} & H_{1,2} & \ldots & H_{1,k-1} & H_{1,k} \\ H_{2,1} & H_{2,2} & \ldots & H_{2,k-1} & H_{2,k} \\ & H_{3,2} & \ldots & H_{3,k-1} & H_{3,k} \\ & & \ddots & \vdots & \vdots \\ & & & H_{k,k-1} & H_{k,k} \\ & & & & H_{k+1,k} \end{pmatrix}$$

The Block Arnoldi algorithm

ALGORITHM

The Block Arnoldi algorithm

1. *Let* U *be an* $N \times s$ *matrix.*
2. *Compute the* $N \times s$ $V_1 \in \mathcal{R}^{N\times s}$ *by finding the QR factorization of* U : $U = V_1 R, R \in \mathcal{R}^{s\times s}$
3. *for i=1,...,k do*
    - *Compute* $W = AV_i$.
    - *for j=1,...,i do*
        1. $H_{j,i} = V_j^\top W$
        2. $W = W - V_j H_{j,i}$
    - *End*
    - *Compute* $H_{i+1,i}$ *by finding the QR decomposition of* $W : W = V_{i+1}\, H_{i+1,i}$
4. *End*

The Global-Arnoldi algorithm

The Global Arnoldi algorithm

1. `Set` $V_1 = V/\|V\|_F$.

2. `For` $j = 1, \ldots, k$. `do`

$\tilde{V} = A\,V_j$,

`for` $i = 1, \ldots, j$. `do`

$h_{i,j} = \langle V_i, \tilde{V} \rangle_F$,

$\tilde{V} = \tilde{V} - h_{i,j} V_i$,

`endfor`

$h_{j+1,j} = \| \tilde{V} \|_F$,

$V_{j+1} = \tilde{V} / h_{j+1,j}$.

`EndFor`.

Block GMRES Algorithm

The definition of BGMRES is equivalent to

$$X_k = X_0 + Z_k$$
$$\| R_0 - AZ_k \|_F = \min_{Z \in \mathbb{K}_k(A,R_0)} \| R_0 - AZ \|_F \ . \tag{8}$$

Thus BGMRES method proceeds as follows

ALGORITHM (BGMRES)

1. *Choose* $X_0 \in \mathbb{C}^{N \times s}$ *and compute* $R_0 = B - AX_0$.
2. $R_0 = V_1 H_{1,0}$ *(The QR factorisation of* R_0*);*
3. *For j=1, ... ,k, do*
   construct V_j *and* $\widetilde{\mathbb{H}}_j$ *by block Arnoldi.*
4. *Solve the least squares problem :*

$$Y_k = arg \min_{Y \in \mathbb{C}^{ks \times s}} \| R_0 - A\mathbb{V}_k Y \|_F,$$

5. *The approximate solution is* $X_k = X_0 + \mathbb{V}_k Y_k$.

GMRES polynomial

If we denote by K_k the Krylov matrix $K_k = [r_0, \ldots, A^{k-1}\, r_0]$, then the GMRES polynomial is defined by

$$p_k^{GMRES}(\xi) = \frac{\begin{vmatrix} 1 & \xi & \ldots & \xi^k \\ (A\, r_0, r_0) & (A\, r_0, A\, r_0) & \ldots & (A\, r_0, A^k\, r_0) \\ \vdots & \vdots & \ldots & \vdots \\ (A^k\, r_0, r_0) & (A^k\, r_0, A r_0) & \ldots & (A^k\, r_0, A^k\, r_0 \end{vmatrix}}{\det(K_k^T A^T A\, K_k)}.$$

GMRES

- We have $p_k^{GMRES}(0) = 1$ and $r_k^{GMRES} = p_k^{GMRES}(A) r_0$.
- $p_k^{GMRES}(\xi) = 1 - (\xi \quad \xi^2 \quad \ldots \quad \xi^k)(K_k^T A^T A\, K_k)^{-1} K_k^T A^T r_0$.

Convergence results for GMRES

Diagonalizable matrices

If the matrix A is diagonalizable $A = X\,\Lambda\,X^{-1}$, and $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$. It is well known that

$$\|r_k\| \leq \|r_0\| \|X\|\,\|X^{-1}\| \min_{p \in P_k, p(0)=1} \max_{i=1,\ldots,n} |p(\lambda_i)|,$$

where λ_i is an eigenvalue of A.
If $r_k = p(A)\,r_0$, then

$$\|r_k\| = \|X\,p(\Lambda)\,X^{-1} r_0\| \leq \kappa(X)\,\|r_0\|\,\|p(\Lambda)\|.$$

where $\kappa(X) = \|X\|\,\|X^{-1}\|$ is the condition number of X.

However since the matrix X is not uniquely defined. We can write

$$\frac{\|r_k\|}{\|r_0\|} \leq \inf_X \kappa(X) \min_{p \in P_k, p(0)=1} \max_{i=1,\ldots,n} |p(\lambda_i)|,$$

Diagonalizable matrix

Let us study the possible simplest case ($n = 2$, $k = 1$). Then $A = X\, diag(\lambda_1, \lambda_2)\, X^{-1}$ and

$$X^H X = \begin{pmatrix} 1 & c_1 \\ \overline{c_1} & 1 \end{pmatrix}, \quad \text{with} \quad |c_1| < 1.$$

The exact residual norms is given by

$$||r_1||^2 = \frac{(1 - |c_1|^2)\, |\alpha_1|^2\, |\alpha_2|^2\, |\lambda_2 - \lambda_1|^2}{|\alpha_1|^2 |\lambda_1|^2 + 2\Re e(c_1 \overline{\alpha_1 \lambda_1} \alpha_2 \lambda_2) + |\alpha_2|^2 |\lambda_2|^2},$$

with $r_0 = X(\alpha_1, \alpha_2)^T$

Convergence results for GMRES : Example

In order to obtain the optimal bound for $\frac{||r_1||}{||r_0||}$ we have to solve an optimization problem.

Diagonalizable matrices : Example

If we assume that all the parameters are reals and $\lambda_1\lambda_2 > 0$, we obtain

$$\frac{||r_1||}{||r_0||} \leq \sqrt{1-|c_1|^2}\,\frac{|\lambda_2-\lambda_1|}{|\lambda_1-2c_1\sqrt{\lambda_1\lambda_2}+\lambda_2|}.$$

It is obvious that this bound refines the classical ones.

$$\frac{||r_1||}{||r_0||} \leq \frac{\sqrt{1+|c_1|}}{\sqrt{1-|c_1|}}\,\frac{|\lambda_2-\lambda_1|}{|\lambda_1|+|\lambda_2|}.$$

If $c_1 = 0$, the

$$\frac{||r_1||}{||r_0||} \leq \frac{|\lambda_2-\lambda_1|}{|\lambda_1|+|\lambda_2|}.$$

Normal matrices ($n \geq 2,\ k = 1$)

Then $A = X\, diag(\lambda_1, \lambda_2, \ldots, \lambda_n)\, X^{-1}$ and $X^H X = I$

$$||r_1||^2 = \frac{\sum_{1\leq i<j\leq n} |\alpha_i|^2\, |\alpha_j|^2\, |\lambda_i - \lambda_j|^2}{|\alpha_1|^2|\lambda_1|^2 + \ldots + |\alpha_n|^2|\lambda_n|^2},$$

with $r_0 = X(\alpha_1, \ldots, \alpha_n)^T$.

$$\frac{||r_1||^2}{||r_0||^2} = \frac{\sum_{1\leq i<j\leq n} \beta_i\, \beta_j\, |\lambda_i - \lambda_j|^2}{\sum_{i=1}^{n} \beta_i |\lambda_i|^2} = F_1(\beta_1, \ldots, \beta_n),$$

where $\beta_i = \dfrac{|\alpha_i|^2}{\sum_{j=1}^{n} |\alpha_j|^2}$.

We have $0 \leq \beta_i \leq 1$ and $\sum_{j=1}^{n} \beta_j = 1$.

$(n \geq 2,\ k = 1)$

$$\frac{||r_1||^2}{||r_0||^2} \leq F_1(\beta^*),$$

$$F_1(\beta^*) = \max_{\substack{\sum_{i=1}^{n} \beta_i = 1 \\ \beta_i \geq 0 \\ i=1,\dots,n}} f_1(\beta)$$

If all eigenvalues are reals, we have

$$F_1(\beta^*) = \left(\frac{|\lambda_{i_2} - \lambda_{i_1}|}{|\lambda_{i_1}| + |\lambda_{i_2}|} \right)^2 = \delta,$$

and $\beta_j^* = 0$ if $j \notin \{i_1, i_2\}$,

$$\beta_{i_1}^* = \frac{1}{2}(1 - \sqrt{\delta}) \quad \text{and} \quad \beta_{i_2}^* = \frac{1}{2}(1 + \sqrt{\delta}).$$

Consequently $r_2 = 0$.

$(n \geq 2,\, k = 1)$

If one of the eigenvalues is complex, we have $C \leq \frac{4}{\pi}$?,

$$F_1(\beta^*) = \left(\frac{|\lambda_{i_2} - \lambda_{i_1}|}{|e^{\imath\theta_{i_1}}\lambda_{i_2} - e^{\imath\theta_{i_2}}\lambda_{i_1}|} \right)^2 \leq C^2 \left(\frac{|\lambda_{i_2} - \lambda_{i_1}|}{|\lambda_{i_1}| + |\lambda_{i_2}|} \right)^2 .$$

Exemple : Let us consider the following matrix, $\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2+\imath & 0 \\ 0 & 0 & 3 \end{pmatrix}$,

1. $\beta_1^* = \dfrac{3}{13}$, $\beta_2^* = \dfrac{6}{13}$ and $\beta_3^* = \dfrac{4}{13}$.
2. The optimal choice is given by $e^{\imath\,\theta_1} = \dfrac{3+2\,\imath}{\sqrt{13}}$, $e^{\imath\,\theta_2} = \dfrac{2-3\,\imath}{\sqrt{13}}$, and $e^{\imath\,\theta_3} = \dfrac{-2+3\,\imath}{\sqrt{13}}$. We have also

$$\sqrt{\delta^*} = \frac{\begin{vmatrix} e^{\imath\theta_1} & \lambda_1 \\ e^{\imath\theta_2} & \lambda_2 \end{vmatrix}}{\begin{vmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{vmatrix}} = \frac{\begin{vmatrix} e^{\imath\theta_1} & \lambda_1 \\ e^{\imath\theta_3} & \lambda_3 \end{vmatrix}}{\begin{vmatrix} 1 & \lambda_1 \\ 1 & \lambda_3 \end{vmatrix}} = \frac{\begin{vmatrix} e^{\imath\theta_2} & \lambda_2 \\ e^{\imath\theta_3} & \lambda_3 \end{vmatrix}}{\begin{vmatrix} 1 & \lambda_2 \\ 1 & \lambda_3 \end{vmatrix}} . \tag{9}$$

3. We have $\frac{\|r_1\|}{\|r_0\|} \leq \dfrac{1}{\sqrt{13}}$ and $r_2 \neq 0$

$(n \geq 2,\ k = 2)$

$k = 2$, we obtain

$$\frac{||r_2||^2}{||r_0||^2} = \frac{\sum_{i,j,k} \beta_i\, \beta_j\, \beta_k\, |\lambda_j - \lambda_i|^2\, |\lambda_k - \lambda_j|^2\, |\lambda_k - \lambda_i|^2}{\sum_{i,j} \beta_i\, \beta_j\, |\lambda_i|^2 |\lambda_j|^2 |\lambda_j - \lambda_i|^2}.$$

We assume that $\{1, 2, 3, 4\} \subset Sp(A) \subset [1, 2] \cup [3, 4]$

1. We have $\beta_1^* = \frac{3}{5} - \beta_4^*,\ \beta_2^* = \frac{3}{5} - 3\beta_4^*,\ \beta_3^* = -\frac{1}{5} + 3\,\beta_4^*$, and $\beta_4 \in \left[\frac{1}{15}, \frac{1}{5}\right]$.
2. $$\frac{\|r_1\|}{\|r_0\|} \leq \frac{3}{5}$$
3. $$\frac{\|r_2\|}{\|r_0\|} \leq \frac{1}{5}$$

$r_3 \neq 0$, but $r_4 = 0$.

Influence of the initial residual

Theorem

Let us assume that the columns of X are normalized i.e. $\|X_i\| = 1$ where $X = [X_1, \ldots, X_n]$. If we expand r_0 in the eigen-basis $r_0 = X\,\alpha$, then

$$\|r_k\| \leq \left(\sum_{i=1}^{n} |\alpha_i|\right) \min_{p \in \widetilde{\mathcal{P}}_k} \max_{\lambda \in \sigma(A)} |p(\lambda)|\,.$$

If the matrix A is normal ($X^H\,X = I$), then we have

$$\|r_k\| \leq \left(\sqrt{\sum_{i=1}^{n} |\alpha_i|^2}\right) \min_{p \in \widetilde{\mathcal{P}}_k} \max_{\lambda \in \sigma(A)} |p(\lambda)|\,.$$

where $\widetilde{\mathcal{P}}_k$ is the set of polynomials of degree less or equal to k, such that $p(0) = 1$.

Proof

Let K_k be the Krylov matrix whose columns are $r_0, A\, r_0, \ldots, A^{k-1}\, r_0$, we have :

[HS, Habilitation Thesis]

If $\|r_k\| \neq 0$ then $\|r_k\|^2 = \dfrac{\det(K_{k+1}^H K_{k+1})}{\det(K_k^H A^H A K_k)} = \dfrac{1}{e_1^T\,(K_{k+1}^H K_{k+1})^{-1}\, e_1}$.

- [1] I. Ipsen, Expressions and bounds for the Gmres Residual, BIT, 38 (1998) 101-104 $\|r_k\| = \dfrac{1}{\|e_1^T (K_{k+1}^{\dagger})\|}$.
- It is not obvious how the expressions for normal matrices in [1] compares to existing polynomial bounds (Min-Max).

Ipsen Decomposition

$$\|r_k\| = \frac{1}{\|((X\, D_\alpha V_{k+1})^\dagger)^H \; e_1\|},$$

$$\|r_k\|^2 = \frac{1}{e_1^T \; \left(V_{k+1}^H D_\alpha^H X^H X D_\alpha \, V_{k+1}\right)^{-1} \; e_1},$$

where

$$D_\alpha = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n), \qquad V_k = \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \ldots & \lambda_2^{k-1} \\ \vdots & \vdots & \ldots & \vdots \\ 1 & \lambda_n & \ldots & \lambda_n^{k-1} \end{pmatrix}. \tag{10}$$

Comparison

Theorem

Introduce the function

$$F_k(t) = \frac{1}{e_1^H \left(V_{k+1}^H D_t V_{j+1}\right)^{-1} e_1}, \tag{11}$$

where $t = (t_1, \ldots, t_n)^T$.

If $\rho = (\rho_1, \ldots, \rho_n)^T$ where $\rho_i = \dfrac{|\alpha_i|}{\sum_{j=1}^n |\alpha_j|}$, then

$$\|r_k\|^2 \leq (\sum_{i=1}^n |\alpha_i|)^2 \quad F_k(\rho).$$

Let the matrix A in addition be normal, then

$$\|r_k\|^2 = (\sum_{i=1}^n |\alpha_i|^2) F_k(\beta),$$

where $\beta = (\beta_1, \ldots, \beta_n)^T$ and $\beta_i = \frac{|\alpha_i|^2}{\sum_{j=1}^n |\alpha_j|^2}$.

the saddle point problem with multiple right-hand sides

Many problems in Science and Engineering require the solution of the saddle point problem with multiple right-hand sides.

$$\underbrace{\begin{pmatrix} A & B^T \\ \epsilon B & O \end{pmatrix}}_{\mathcal{A}} \underbrace{\begin{pmatrix} X \\ Y \end{pmatrix}}_{\mathcal{X}} = \underbrace{\begin{pmatrix} F \\ \epsilon G \end{pmatrix}}_{\mathcal{B}}, \tag{12}$$

Where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite matrix and $B^T \in \mathbb{R}^{n\times m}$ has full column rank, with $X \in \mathbb{R}^{n\times s}$, $Y \in \mathbb{R}^{m\times s}$ and $F \in \mathbb{R}^{n\times s}$, $G \in \mathbb{R}^{m\times s}$.

Convergence analysis of the global GMRES method

In this subsection, we recall some convergence results for the global GMRES method. Let $\mathcal{A} = \mathcal{Z}\mathcal{D}\mathcal{Z}^{-1}$, where $\mathcal{D}$ is the diagonal matrix whose elements are the eigenvalues $\lambda_1, \ldots, \lambda_{n+m}$, and $\mathcal{Z}$ is the eigenvector matrix.
Let the initial residual R_0 be decomposed as $R_0 = \mathcal{Z}\beta$ where β is an $(n+m) \times s$ matrix whose columns are denoted by $\beta^{(1)}, \ldots, \beta^{(s)}$. Let $R_k = \mathcal{B} - \mathcal{A}\mathcal{X}_k$ be the kth residual obtained by the global GMRES when applied to (12). Then we have

$$\|R_k\|_F^2 \leq \frac{\|\mathcal{Z}\|_2^2}{e_1^T (V_{k+1}^T \widetilde{\mathcal{D}} V_{k+1})^{-1} e_1}, \tag{13}$$

where

$$\widetilde{\mathcal{D}} = \begin{pmatrix} \sum_{i=1}^{s} \mid {\beta_1}^{(i)} \mid^2 & & \\ & \ddots & \\ & & \sum_{i=1}^{s} \mid {\beta_{n+m}}^{(i)} \mid^2 \end{pmatrix} \quad \text{and} \quad V_{k+1} = \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^k \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_{n+m} & \ldots & \lambda_{n+m}^k \end{pmatrix} \tag{14}$$

The coefficients $\beta_1^{(i)}, \ldots, \beta_{n+m}^{(i)}$ are the components of the vector $\beta^{(i)}$ and e_1 is the first unit vector of $\mathbb{R}^{k+1}$.

Preconditioning

In this following section we present the preconditioner $\mathcal{P}_p$, for solving saddle point problems with multiple right-hand sides (12). Now we propose the preconditioner $\mathcal{P}_p$ for solving saddle point problems with multiple right-hand sides (12)

$$\mathcal{P}_p = \begin{pmatrix} A & B^T \\ \epsilon B & \alpha Q \end{pmatrix}, \text{ with } \mathcal{P}_p^{-1}\mathcal{A}\mathcal{X} = \mathcal{P}_p^{-1}\mathcal{B}. \tag{15}$$

Where $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix, $B \in \mathbb{R}^{m\times n}$ has a full row rank and Q is an approximation of Shur complement $S = -BA^{-1}B^T$ and $\alpha > 0$.

Preconditioner factorization

The preconditioner has the block-triangular factorization

$$\mathcal{P}_p = \begin{pmatrix} A & B^T \\ \epsilon B & \alpha Q \end{pmatrix} = \begin{pmatrix} I & O \\ \epsilon B A^{-1} & I \end{pmatrix} \begin{pmatrix} A & O \\ O & \tilde{S} \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ O & I \end{pmatrix}, \tag{16}$$

where $\tilde{S} = \left(\alpha Q - \epsilon B A^{-1} B^T\right)$.
If $\epsilon = 1$, the block $\tilde{S}$ is positive definite matrix for all $\alpha \lambda_{\min}(Q) > \lambda_{\max}(BA^{-1}B^T)$.
Thus the inverse of the preconditioned matrix $\mathcal{P}_p$ is given by the following equality

$$\mathcal{P}_p^{-1} = \begin{pmatrix} A & B^T \\ \epsilon B & \alpha Q \end{pmatrix}^{-1} = \begin{pmatrix} I & -A^{-1}B^T \\ O & I \end{pmatrix} \begin{pmatrix} A^{-1} & O \\ O & \tilde{S}^{-1} \end{pmatrix} \begin{pmatrix} I & O \\ BA^{-1} & I \end{pmatrix}. \tag{17}$$

Preconditioned matrix

If $\epsilon = -1$, the preconditioned matrix $\mathcal{P}_p^{-1}\mathcal{A}$ can be rewritten as follow

$$\mathcal{P}_p^{-1}\mathcal{A} = \begin{pmatrix} A & B^T \\ -B & \alpha Q \end{pmatrix}^{-1} \begin{pmatrix} A & B^T \\ -B & O \end{pmatrix} = \begin{pmatrix} I & K_1 \\ O & K_2 \end{pmatrix}, \tag{18}$$

where $\tilde{S} = \left(\alpha Q + BA^{-1}B^T\right)$, $K_1 = A^{-1}B^T - A^{-1}B^T\tilde{S}^{-1}BA^{-1}B^T$ and $K_2 = \tilde{S}^{-1}BA^{-1}B^T$.

The precondioned Global GMRES

ALGORITHM

Algorithm 2 : The precondioned Global GMRES

---

1 : $\mathcal{P}_p V_1 = R_0,\ V_1 = V_1/\|V_1\|_F$
2 : for $j = 1, 2, ..., k$ *do ;*
3 : $\mathcal{P}_p W := \mathcal{A} V_j$ *;*
4 : for $i = 1, 2, ..., j$ *do ;*
5 : $H_{i,j} =< W, V_i >_F$ *;*
6 :$W = W - H_{ij} V_i$;
7 : end ;
8 : $H_{j+1,j} = ||W||_F$;
9 : $V_{j+1} = W/H_{j+1,j}$;
10 : Solve the linear system $H_{k,k} y = \beta e_1$ *for* y *;*
11 : Set $\mathcal{X}_k = \mathcal{X}_0 + V_k \diamond y$ *and* $R_k = \mathcal{B} - \mathcal{A}\mathcal{X}_k$ *;*
12 : end for

The precondioned Global GMRES

At each step of applying the preconditioner $\mathcal{P}_p$ inside the GMRES algorithm, we need to solve the system 1 and 3 of algorithm 2. For a given matrix $V = [V_1; V_2]$ where $V_1 \in \mathbb{R}^{n\times s}$ and $V_2 \in \mathbb{R}^{m\times s}$. Let $Z = [Z_1; Z_2]$, where $Z_1 \in \mathbb{R}^{n\times s}$ and $Z_2 \in \mathbb{R}^{m\times s}$.

$$\underbrace{\begin{pmatrix} A & B^T \\ -B & \alpha Q \end{pmatrix}}_{\mathcal{P}_p} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}. \tag{19}$$

We can solve (19) by using the following algorithm.

Algorithm 3 :

---

1 : Solve $\underbrace{\left(A + \frac{1}{\alpha} B^T Q^{-1} B\right)}_{A_\alpha} Z_1 = \underbrace{V_1 - \frac{1}{\alpha} B^T Q^{-1} V_2}_{J}$;

2 : Compute $Z_2 = \frac{1}{\alpha} Q^{-1}(V_2 + B Z_1)$;

The matrix A_α is symmetric positive definite. Therefore, we can solve the system with the coefficient matrix A_α by the preconditioned global CG method or by the preconditioned global MINRES method inexactly.

Numerical resultS

Table 3 : Numerical results for multiple right-hand sides with global approach .

α	$\mathcal{P}_p$		$\mathcal{P}_{VPSS}$		$\mathcal{P}_T$		$\mathcal{P}$
10^{-5}	IT	4	IT	72	IT	83	IT
	CPU	1.46	CPU	4.85	CPU	2.81	CP
	RES	1.25e-04	RES	1.08e-05	RES	9.10e-07	RE
	ERR	6.06e-03	ERR	6.27e-05	ERR	7.68e-06	ER
10^{-4}	IT	8	IT	75	IT	94	IT
	CPU	1.83	CPU	5.28	CPU	3.24	CP
	RES	3.93e-06	RES	1.25e-06	RES	1.09e-06	RE
	ERR	3.43e-05	ERR	8.73e-06	ERR	9.23e-06	ER
10^{-3}	IT	14	IT	74	IT	87	IT
	CPU	1.58	CPU	5.03	CPU	3.64	CP
	RES	1.17e-07	RES	8.59e-07	RES	1.03e-06	RE
	ERR	1.79e-06	ERR	6.27e-06	ERR	7.72e-06	ER
10^{-2}	IT	27	IT	74	IT	78	IT
	CPU	1.54	CPU	5.09	CPU	3.65	CP
	RES	1.18e-08	RES	1.10e-07	RES	1.22e-06	RE
	ERR	1.69e-07	ERR	1.45e-06	ERR	9.58e-06	ER
10^{-1}	IT	40	IT	55	IT	67	IT
	CPU	2.15	CPU	2.50	CPU	4.35	CP

Thank you for your attention