RÉPUBLIQUE
FRANÇAISE
Liberté
Égalité
Fraternité

ONERA

THE FRENCH AEROSPACE LAB

Randomized flexible GMRES with singular vectors based deflated restarting

Yongseok Jang[1,*], Cédric Content[1], Emeric Martin[1], Laura Grigori[2]

[1] ONERA, *the French Aerospace Lab*. [2] Inria, Sorbonne Université, CNRS, Laboratoire Jacques-Louis Lions.

Fast linear solvers

Krylov subspace methods

Aim

We seek a fast and reliable linear solver for non-symmetric systems arising in CFD simulations in the context of aerodynamics.

Krylov subspace methods, precisely GMRES, are of our particular interest.

A new proposed variant of GMRES

Randomized GMRES MDR

Motivation

- Flexible GMRES [Saad, 1993, Giraud et al., 2010]
- Flexible GMRES with deflated restarting [Morgan, 2002, Giraud et al., 2010]
- Randomized GMRES [Balabanov and Grigori, 2020]
- SVD based deflation [Daas et al., 2021]

We propose a randomized FGMRES with deflated restarts based on either the harmonic Ritz pairs or singular vectors.

Krylov subspace methods

FGMRES

Krylov subspace

Let us consider $A\boldsymbol{x} = \boldsymbol{b}$ where $A \in \mathbb{C}^{n\times n}$ and $\boldsymbol{b} \in \mathbb{C}^n$. We approximate the solution in the **Krylov subspace** $\mathcal{K}_m(A, \boldsymbol{b})$ is defined by

$$\mathcal{K}_m(A, \boldsymbol{b}) = \mathrm{span}\{\boldsymbol{b}, A\boldsymbol{b}, A^2\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}\}.$$

GMRES (*Generalized minimal residual method*) is one of Krylov subspace method finding the approximate solution in the Krylov subspace with minimal residual $\boldsymbol{r}_m := \boldsymbol{b} - A\boldsymbol{x}_m$.

Randomized Gram-Schmidt (RGS)

RGS flexible Arnoldi

RGS

To construct a basis of $\mathcal{K}_m(A, \boldsymbol{b})$, **Gram-Schmidt** (CGS/MGS) algorithms are commonly used. Instead of l_2 norm sense orthonomality, RGS algorithm proposes **random sketched** norm sense. This method leads to less computational costs.

The Q factor (basis vectors) of RGS is more stable than CGS and as stable as MGS with less complexity.

For more details, please refer to [Balabanov and Grigori, 2020].

Randomized Gram-Schmidt (RGS)

Remark

Dimension reduction technique

With a **random sketching matrix** $\Theta \in \mathbb{C}^{t \times n}$ for $t \ll n$, we can construct the Θ orthonormal basis vectors in the context of l_2 subspace embedding from the high dimension n to the low dimension t.

$$(\boldsymbol{v}, \boldsymbol{w}) \approx (\boldsymbol{v}, \boldsymbol{w})_\Theta = (\Theta\boldsymbol{v}, \Theta\boldsymbol{w}) \qquad \text{for any } \boldsymbol{v}, \boldsymbol{w} \in \mathbb{C}^n.$$

Θ can be defined with Gaussian distribution, Radmacher distribution, (P-)SRHT, etc.

RGS flexible Arnoldi

Algorithm

Algorithm 1: Flexible RGS Arnoldi iteration

Data: matrix $A \in \mathbb{C}^{n\times n}$, vector $\boldsymbol{v} \in \mathbb{C}^n$, sketching matrix $\Theta \in \mathbb{R}^{t\times n}$, number of iterations m, and variable preconditioner M_j for $j = 1, \dots, m$

Result: $V_{m+1} \in \mathbb{C}^{n\times(m+1)}$, $Z_m \in \mathbb{C}^{n\times m}$ and Hessenberg matrix $H_m \in \mathbb{C}^{(m+1)\times m}$ (optionally $S_{m+1} \in \mathbb{C}^{t\times(m+1)}$)

$\boldsymbol{w} = \boldsymbol{v}$; $\boldsymbol{p} = \Theta\boldsymbol{w}$; $\boldsymbol{q} = \boldsymbol{w}$; $\boldsymbol{s} = \boldsymbol{p}$; $h = \|\boldsymbol{s}\|$;

$\boldsymbol{s}_1 = \boldsymbol{s}/h$; $\boldsymbol{v}_1 = \boldsymbol{q}/h$;

$S_1 = [\boldsymbol{s}_1]$; $V_1 = [\boldsymbol{v}_1]$;

for $j = 1 : m$ **do**

- Apply preconditioner M_j onto $\boldsymbol{v}_j$: $\boldsymbol{z}_j = \mathrm{apply}(M_j, \boldsymbol{v_j})$;
- $\boldsymbol{w} = A\boldsymbol{z}_j$;
- Sketch: $\boldsymbol{p} = \Theta\boldsymbol{w}$;
- Solve $t \times j$ least squares problem: $H_m(1:j, j) = \arg\min\limits_{\boldsymbol{y}\in\mathbb{C}^j} \|S_j\boldsymbol{y} - \boldsymbol{p}\|$;
- $\boldsymbol{q} = \boldsymbol{w} - V_j H_m(1:j, j)$;
- Sketch: $\boldsymbol{s} = \Theta\boldsymbol{q}$;
- $h = \|\boldsymbol{s}\|$; $\boldsymbol{s}_{j+1} = \boldsymbol{s}/h$; $\boldsymbol{v}_{j+1} = \boldsymbol{q}/h$;
- $S_{j+1} = [\boldsymbol{s}_1, \dots, \boldsymbol{s}_{j+1}]$; $V_{j+1} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_{j+1}]$; $Z_j = [\boldsymbol{z}_1, \dots, \boldsymbol{z}_j]$; $H_m(j+1, j) = h$;

end

While CGS/MGS requires $2nm^2$ flops, RGS needs only nm^2 flops.

FGMRES

Minimizing the residual norm

As in the classical manner, we want to minimize the residual vector $\boldsymbol{r}_m \in \mathrm{range}(V_{m+1})$. Note that since the approximate solution $\boldsymbol{x}_m$ resides in $\boldsymbol{x}_0 + \mathrm{range}(Z_m)$, we have

$$\boldsymbol{r}_m = \boldsymbol{b} - A\boldsymbol{x}_m = \boldsymbol{r}_0 - AZ_m\boldsymbol{y} = \boldsymbol{r}_0 - V_{m+1}H_m\boldsymbol{y},$$

by the Arnoldi identity $AZ_m = V_{m+1}H_m$.

RGS FGMRES

Minimizing the sketched residual norm

Unlike CGS/MGS, V_{m+1} of RGS is not l_2 orthogonal but Θ orthonormal. Thus, we shall minimize the sketched norm of the residual by

$$\text{minimize } \|\boldsymbol{r}_m\|_\Theta \Leftrightarrow \text{solve } \boldsymbol{y}^* = \arg\min_{\boldsymbol{y}\in\mathbb{C}^m} \|\boldsymbol{c} - H_m\boldsymbol{y}\|,$$

where $\boldsymbol{c} = V_{m+1}^H \Theta^H \Theta \boldsymbol{r}_0$.

$\boldsymbol{c}$ can be also simplified as

$$\boldsymbol{c} = V_{m+1}^H \Theta^H \Theta \boldsymbol{r}_0 = [\|\boldsymbol{r}_0\|_\Theta, 0, \ldots, 0]^T. \tag{1}$$

Norm equivalence

l_2 norm and Θ norm

For a Θ equipped with l_2 subspace embedding property,

$$\sigma_{\min}(\Theta)\|\boldsymbol{v}\| \leq \|\boldsymbol{v}\|_{\Theta} \leq \|\Theta\|\|\boldsymbol{v}\|,$$

where $\sigma_{\min}(\Theta)$ is the minimum singular value of Θ. Thus, taking into account the sketched norm of the residual is reasonable.

The approximate solution $\boldsymbol{x}_m$ by RGS GMRES is a quasi-optimal minimizer of the residual norm.

Randomized Rayleigh-Ritz method

Approximation of Ritz pairs

Definition

Let $B \in \mathbb{C}^{n \times n}$ and $\mathcal{S}$ be a k dimensional subspace of $\mathbb{C}^n$. Then a Ritz pair $(\boldsymbol{y}, \lambda) \in \mathbb{C}^n \times \mathbb{C}$ of B with respect to $\mathcal{S}$ satisfies

$$\boldsymbol{y} \in \mathcal{S} \qquad \text{and} \qquad B\boldsymbol{y} - \lambda \boldsymbol{y} \perp \mathcal{S}.$$

Here, we call $\boldsymbol{y}$ a *Ritz vector* associated with a *Ritz value* λ. If $B = A^{-1}$ and $\mathcal{S} = A\mathcal{K}_m(A, \boldsymbol{b})$, the pair is called ***harmonic* Ritz pairs** of A.

The randomized Rayleigh-Ritz method [Balabanov and Grigori, 2021] allows us to approximate eigenpairs $(\boldsymbol{v}, \mu)$ of A by

$$(\boldsymbol{v}, \mu) \approx (\boldsymbol{u}, \lambda) = (V_m \boldsymbol{y}, \lambda) \qquad \text{where } V_m \text{ is defined by RGS.}$$

Computation of deflated vectors

Harmonic Ritz eigenpairs

Harmonic Ritz vectors

We want to deflate the eigenvectors associated with the smallest eigenvalues. To approximate them, we solve the following eigenvalue problem:

$$H_m^H H_m \boldsymbol{g} = \lambda \hat{H}_m^H \boldsymbol{g} \qquad \text{where } \hat{H}_m = H_m(1:m, 1:m).$$

We can rewrite the eigenvalue problem as

$$\left(\hat{H}_m + h^2 \hat{H}_m^{-H} \boldsymbol{e}_m \boldsymbol{e}_m^T\right) \boldsymbol{g} = \lambda \boldsymbol{g},$$

where $h = H_m(m+1, m)$ and $\boldsymbol{e}_m$ is the m-th standard basis of $\mathbb{C}^m$.

Computation of deflated vectors

Singular vectors

Consider $B = A^H A$ and $\mathcal{S} = \mathcal{K}_m(A, \boldsymbol{b})$ for the (randomized) Rayleigh-Ritz method.

Singular vectors

In a similar way with the harmonic Ritz pairs, we can compute the singular vectors associated with the smallest singular values by solving the following eigenvalue problem:

$$H_m^H H_m \boldsymbol{g} = \lambda \boldsymbol{g}.$$

Deflation Krylov Subspaces

Randomized FGMRES with (modified) deflated restarts

RGS FGMRES DR

As [Morgan, 2002, Giraud et al., 2010] proposed, we introduce the randomized FGMRES DR with the augmented form of deflation vectors (**harmonic Ritz pairs**).

RGS FGMRES MDR

Using the split of the SVD decomposition, we derive the randomized FGMRES MDR (deflating **singular vectors**).

RGS FGMRES MDR

GMRES and GCRO

Morgan used the fact of the harmonic residual vectors such that

$$AZ_m \boldsymbol{g} - \lambda V_m \boldsymbol{g} \in \mathrm{range}(V_{m+1})$$

in GMRES DR, but it does not hold with the singular based Ritz residual vectors.

Thus, SVD based deflation is not applicable in the context of GMRES-DR but GCRO-DR.

- GMRES: $AZ_k = V_{k+1} H_k$ and $V_{k+1}^H V_{k+1} = I_{k+1}$.
- GCRO (*Generalized Conjugated Residual Orthogonalization* method): $AZ_k = V_k$ and $V_k^H V_k = I_k$.

GMRES and GCRO are algebraically equivalent.

Algorithm

RGS FGMRES

Data: matrix $A \in \mathbb{C}^{n\times n}$, non-zero vector $\boldsymbol{b} \in \mathbb{C}^n$, sketching matrix $\Theta \in \mathbb{R}^{t\times n}$, number of restarts m, variable preconditioner M_j for $j = 1, \dots, m$, tolerance $tol > 0$, and initial vector $\boldsymbol{x}_0 \in \mathbb{C}^n$

Result: approximate solution $\boldsymbol{x}$ for $A\boldsymbol{x} = \boldsymbol{b}$

1 $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$; $\beta = \|\boldsymbol{r}_0\|$; $\boldsymbol{c} = [\|\Theta \boldsymbol{r}_0\|, O_{1\times m}]^T$; $\boldsymbol{e}_m =$ the mth Cartesian basis vector of $\mathbb{C}^m$;

2 Initial RGS Arnoldi process to get $V_{m+1},\ Z_m$ and H_m with the starting vector $\boldsymbol{r}_0$ such that satisfying

$$AZ_m = V_{m+1}H_m \qquad \text{and} \qquad V_{m+1}^H \Theta^H \Theta V_{m+1} = I_{m+1}$$

by Algorithm 1.

3 $\boldsymbol{y}^* = \arg\min_{\boldsymbol{y}\in\mathbb{C}^m} \|\boldsymbol{c} - H_m \boldsymbol{y}\|$; $\boldsymbol{x}_m = \boldsymbol{x}_0 + Z_m \boldsymbol{y}^*$; $\boldsymbol{x}_0 = \boldsymbol{x}_m$; $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$; $\beta = \|\boldsymbol{r}_0\|$; $\boldsymbol{c} = V_{m+1}^H \Theta^H \Theta \boldsymbol{r}_0$;

Repeat this until converge.

Algorithm

RGS FGMRES DR

After one cycle of RGS FGMRES,

while $\beta / \|\boldsymbol{b}\| > tol$ **do**

$h = H_m(m+1, m)$; $\hat{H}_m = H_m(1:m, 1:m)$;

Compute k harmonic Ritz vectors by solving the eigenvalue problem:

$$\left(\hat{H}_m + h^2 \hat{H}_m^{-H} \boldsymbol{e}_m \boldsymbol{e}_m^T\right) \boldsymbol{g}_i = \lambda_i \boldsymbol{g}_i \qquad \text{for} \qquad i = 1, \ldots, k$$

and set $G_k = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_k] \in \mathbb{C}^{m \times k}$;

$G_{k+1} = \left[\begin{bmatrix} G_k \\ O_{1\times k} \end{bmatrix}, \boldsymbol{c} - H_m \boldsymbol{y}^*\right]$;

Perform QR decomposition of G_{k+1}: $G_{k+1} = Q_{k+1} R_{k+1}$;

Define $V_{k+1} = V_{m+1} Q_{k+1}$, $Z_k = Z_m Q_{k+1}(1:m, 1:k)$ and $H_k = Q_{k+1}^H H_m Q_{k+1}(1:m, 1:k)$ satisfying $AZ_k = V_{k+1} H_k$;

Perform $(m-k)$ steps of RGS Arnoldi process to construct $V_{m+1},\ Z_m$ and H_m with $V_{k+1},\ Z_k$ and H_k such that satisfying

$$AZ_m = V_{m+1} H_m \qquad \text{and} \qquad V_{m+1}^H \Theta^H \Theta V_{m+1} = I_{m+1}$$

Update $\boldsymbol{y}^* = \arg\min_{\boldsymbol{y} \in \mathbb{C}^m} \|\boldsymbol{c} - H_m \boldsymbol{y}\|$, $\boldsymbol{x}_m = \boldsymbol{x}_0 + Z_m \boldsymbol{y}^*$, $\boldsymbol{x}_0 = \boldsymbol{x}_m$, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\beta = \|\boldsymbol{r}_0\|$ and $\boldsymbol{c} = V_{m+1}^H \Theta^H \Theta \boldsymbol{r}_0$;

end

$\boldsymbol{x} = \boldsymbol{x}_m$;

Algorithm

RGS FGMRES MDR

After one cycle of RGS FGMRES,

while $\beta / \|\boldsymbol{b}\| > tol$ **do**

Compute k singular vectors by solving the eigenvalue problem:

$$H_m^H H_m \boldsymbol{g}_i = \lambda_i \boldsymbol{g}_i \qquad \text{for} \qquad i = 1, \dots, k$$

and set $G_k = [\boldsymbol{g}_1, \dots, \boldsymbol{g}_k] \in \mathbb{C}^{m \times k}$;

Perform QR decomposition of $H_m G_k$: $H_m G_k = Q_k R_k$;

Define $V_k = V_{m+1} Q_k$ and $Z_k = Z_m G_k R_k^{-1}$ satisfying $A Z_k = V_k$;

Perform $(m-k)$ steps of RGS Arnoldi process to construct V_{m-k+1}, Z_{m-k} and H_{m-k} with the starting vector $\boldsymbol{r}_0$ with respect to the linear operator $(I_n - V_k V_k^H \Theta^H \Theta) A$ such that satisfying

$$(I_n - V_k V_k^H \Theta^H \Theta) A Z_{m-k} = V_{m-k+1} H_{m-k} \qquad \text{with} \qquad V_{m-k+1}^H \Theta^H \Theta V_{m-k+1} = I_{m-k+1}$$

based on Algorithm $\boxed{1}$;

Define $V_{m+1} = [V_k, V_{m-k+1}]$, $Z_m = [Z_k, Z_{m-k}]$ and $H_m = \begin{bmatrix} I_k & V_k^H \Theta^H \Theta A Z_{m-k} \\ O_{(m-k+1) \times k} & H_{m-k} \end{bmatrix}$ yielding $A Z_m = V_{m+1} H_m$ and $V_{m+1}^H \Theta^H \Theta V_{m+1} = I_{m+1}$;

Update $\boldsymbol{y}^* = \arg\min_{\boldsymbol{y} \in \mathbb{C}^m} \|\boldsymbol{c} - H_m \boldsymbol{y}\|$, $\boldsymbol{x}_m = \boldsymbol{x}_0 + Z_m \boldsymbol{y}^*$, $\boldsymbol{x}_0 = \boldsymbol{x}_m$, $\boldsymbol{r}_0 = \boldsymbol{b} - A\boldsymbol{x}_0$, $\beta = \|\boldsymbol{r}_0\|$ and $\boldsymbol{c} = V_{m+1}^H \Theta^H \Theta \boldsymbol{r}_0$;

end

$\boldsymbol{x} = \boldsymbol{x}_m$;

Reorthogonalization

RGS2

l_2 orthogonal RGS

After the RGS Arnoldi iteration, additional l_2 QR leads us to l_2 orthonormal basis. For instance, performing QR on $V_{m+1} = QR$ gives the new

$$\bar{V}_{m+1} = Q \qquad \text{and} \qquad \bar{H}_m = RH_m,$$

satisfying Arnoldi identity $AZ_m = \bar{V}_{m+1}\bar{H}_m$ and l_2 orthogonality.

It seems equivalent to re-orthogonalization like CGS2 and MGS2 [Giraud et al., 2005], so that we call this implementation **RGS2**. Obviously, RGS2 has less complexity than CGS2/MGS2.

Numerical comparison of Gram-Schmidt process

Condition numbers

Consider a condition number of the orthonormal matrix Q_m arising in QR decomposition with respect to Gram-Schmidt procedures. For instance, we will compare

1. Classical Gram-Schmidt (CGS)
2. Modified Gram-Schmidt (MGS)
3. Randomized Gram-Schmidt (RGS)

RGS QR

Construction of an orthogonal basis for synthetic functions

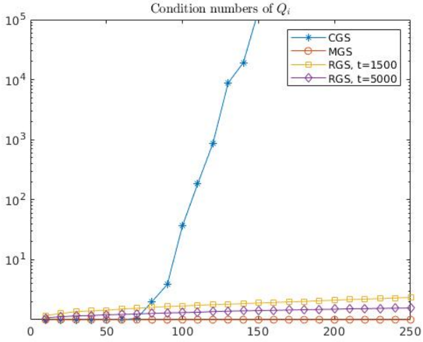
Matrix W

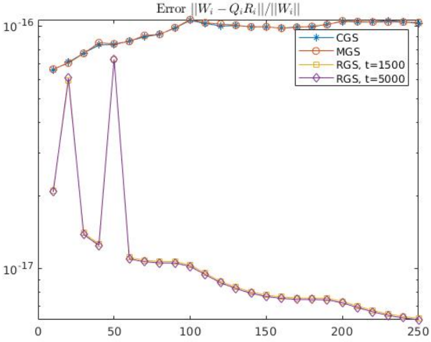
We refer to [Balabanov and Grigori, 2020, Section 5.1] for the construction of W.

- $W \in \mathbb{R}^{10^5 \times 300}$.
- $\mathrm{cond}(W) = O(10^{14})$.
- By the corresponding GS process, we have $W_i = Q_i R_i$ where W_i is the submatrix of W consisting the first to i-th columns for the iteration i, and Q_i and R_i are defined by QR factorizations.

RGS QR

Numerical result





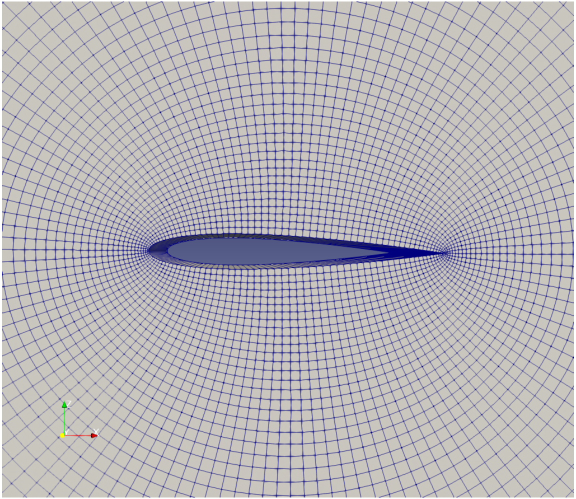
NACA12

Consider turbulent compressible CFD (*negative SA model* [Allmaras and Johnson, 2012]) in airfoil shapes for aircraft wings developed by the National Advisory Committee for Aeronautics (NACA).

- The system of a matrix generated in elsA.
- Solve $A\mathbf{x} = \mathbf{b}$ with ILU(0) preconditioner.
- A is not symmetric positive definite.
- $n = 81,920$ and $\mathrm{cond}(A) = O(10^6)$.
- Compare the numerical performance w.r.t. GS process.

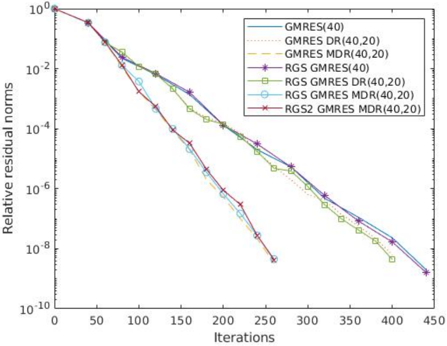
NACA12

Mesh



NACA 12

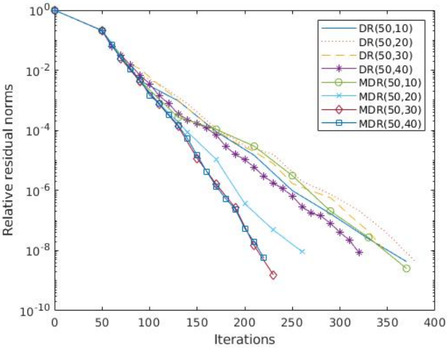
Residuals



- Compare the convergence with respect to deflation and GS process.
- SVD deflation works better.

NACA 12

The number of deflated vectors *k*



- ◇ With fixed m, perform RGS GMRES (M)DR varying with k.
- ◇ SVD deflation still works better.
- ◇ Increasing k improves the performance.

Challenging problem

LS89

Consider the LS89 test case is one of the challenging turbine test cases which is frequently used to benchmark CFD tools [Arts and De Rouvroit, 1990].

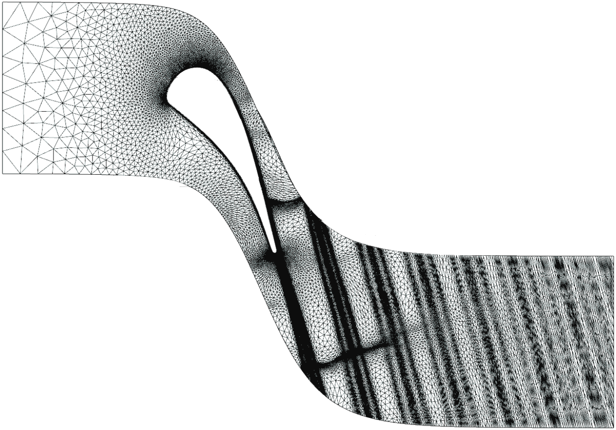
A turbine-blade cascade profile (2D)

Employing a second order scheme in space on RANS equations leads to the linear system on the unstructured mesh consisting of about 80,000 cells.

- Solve $A\mathbf{x} = \mathbf{b}$ with outer-inner GMRES with RAS preconditioning as the preconditioner of inner GMRES.
- A is not symmetric positive definite.
- $n = 115,368$ and $\mathrm{cond}(A) = O(10^{14})$.

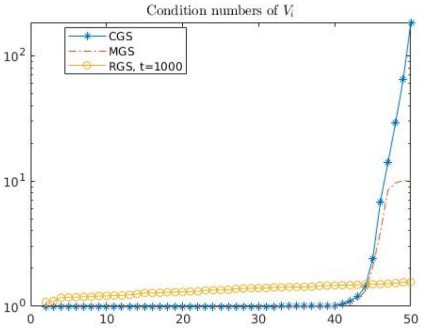
LS89

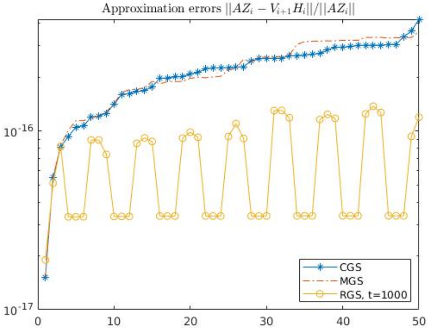
Mesh



LS89

Stability

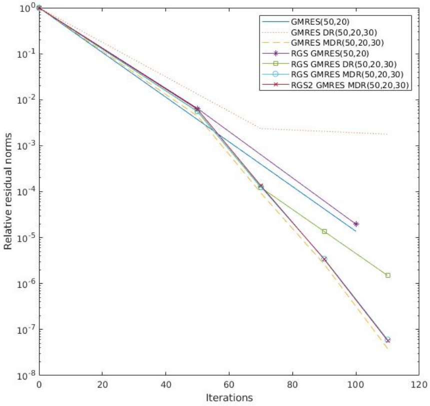




For large m, CGS/MGS will significantly degrade, whereas RGS will provide stable Krylov basis with $\text{cond}(V_{m+1}) = 1 + O(\epsilon)$.

LS89

Relative residuals



- ◇ GMRES DR failed with CGS, while deflation with RGS improved the convergence.
- ◇ MDR showed the best for any GSs.

Conclusion

Summary

- RGS based GMRES variants showed high quality of performance.
- We observed that RGS is more stable than CGS/MGS with less computational costs.
- Singular vectors based deflation performed better than Harmonic Ritz vectors based deflation.

Discussion

- Easy parallel implementation.
- The optimal deflation numbers k is still ambiguous. But we can introduce some adaptive strategy to define variable k. E.g. for each i cycle, choosing k_i vectors when those associated eigenvalues (or singular values) are sufficiently close to the origin.
- Preprint available soon...

References

[Allmaras and Johnson, 2012] Allmaras, S. R. and Johnson, F. T. (2012).
Modifications and clarifications for the implementation of the spalart-allmaras turbulence model.
In *Seventh international conference on computational fluid dynamics (ICCFD7)*, volume 1902. Big Island, HI.

[Arts and De Rouvroit, 1990] Arts, T. and De Rouvroit, M. L. (1990).
Aero-thermal performance of a two dimensional highly loaded transonic turbine nozzle guide vane: A test case for inviscid and viscous flow computations.
In *Turbo Expo: Power for Land, Sea, and Air*, volume 79047, page V001T01A106. American Society of Mechanical Engineers.

[Balabanov and Grigori, 2020] Balabanov, O. and Grigori, L. (2020).
Randomized gram-schmidt process with application to gmres.
arXiv preprint arXiv:2011.05090.

[Balabanov and Grigori, 2021] Balabanov, O. and Grigori, L. (2021).
Randomized block Gram-Schmidt process for solution of linear systems and eigenvalue problems.
arXiv preprint arXiv:2111.14641.

[Daas et al., 2021] Daas, H. A., Grigori, L., Hénon, P., and Ricoux, P. (2021).
Recycling krylov subspaces and truncating deflation subspaces for solving sequence of linear systems.
ACM Transactions on Mathematical Software (TOMS), 47(2):1–30.

[Giraud et al., 2010] Giraud, L., Gratton, S., Pinel, X., and Vasseur, X. (2010).
Flexible gmres with deflated restarting.
SIAM Journal on Scientific Computing, 32(4):1858–1878.

[Giraud et al., 2005] Giraud, L., Langou, J., and Rozloznik, M. (2005).
The loss of orthogonality in the gram-schmidt orthogonalization process.
Computers & Mathematics with Applications, 50(7):1069–1075.

[Morgan, 2002] Morgan, R. B. (2002).
Gmres with deflated restarting.
SIAM Journal on Scientific Computing, 24(1):20–37.

[Saad, 1993] Saad, Y. (1993).
A flexible inner-outer preconditioned gmres algorithm.
SIAM Journal on Scientific Computing, 14(2):461–469.

Thanks for your attention!

Merci de votre attention!

www.onera.fr