A two-level preconditioning framework based on a Richardson iterative process

Thomas Dufaud

1 Introduction

We focus on the solution of a general linear system $Au = f$ by a Krylov type iterative method, where $A \in \mathbb{R}^{m \times m}$ is non-singular, $u, f \in \mathbb{R}^m$. The major drawback of the GCR (Generalized Conjugate Residual) [7] and the GMRES (General Minimum Residual) [8] methods is their convergence rate that depends on the conditioning number $\kappa(A) = \|A\| \, \|A^{-1}\|$.

The convergence rate of these techniques decreases while κ increases and the use of such methods needs preconditioning. In the following we consider left preconditioning. The goal is to solve $M^{-1}Au = M^{-1}f$ with M^{-1} such that $\kappa(M^{-1}A) \ll \kappa(A)$.

Preconditioning can be enhanced by multilevel techniques. Multilevel techniques are known to be robust for scalar elliptic Partial Differential Equations with standard discretization and to enhance the scalability of domain decomposition method such as Restricted Additive Schwarz preconditioning techniques. An issue is their application to linear system encountered in industrial applications which can be derived from non-elliptic PDEs. Moreover, the building of coarse levels algebraically becomes an issue since the only known information is contained in the operator to inverse.

One can consider a coarse space as a space to represent an approximated solution of a smaller dimension than the leading dimension of the system. It is possible to build a coarse level based on a coarse representation of the solution. Drawing our inspiration from the Aitken-SVD methodology [9] dedicated to Schwarz methods, we proposed to construct an approximation space by computing the Singular Value Decomposition of a set of iterated solutions of the Richardson process associated to a given preconditioner.

From a preconditioner M^{-1} associated to a Richardson process:

$$u^k = u^{k-1} + \alpha M^{-1}\left(f - Au^{k-1}\right) \text{ with } \alpha \in \mathbb{R} \tag{1}$$

We propose to build a two-level additive preconditioner M_{2L}^{-1}:

$$M_{2L}^{-1} = M^{-1} + M_c^{-1} \tag{2}$$

where for a basis $\mathbb{U}_q \in \mathbb{R}^{m \times q}$, $M_c^{-1} = \mathbb{U}_q(\mathbb{U}_q^T A \mathbb{U}_q)^{-1}\mathbb{U}_q^T$.

Thomas Dufaud
INRIA Rennes - Bretagne Atlantique, Campus universitaire de Beaulieu, 35042 Rennes Cedex FRANCE,
e-mail: thomas.dufaud@inria.fr

The plan of the paper is the following. Section 2 describes the methodology to compute an algebraic coarse level from successive iterations of a Richardson process. Numerical investigations with the RAS preconditioner built for real non-symmetric indefinite operator, are performed in section 3. Section 4 concludes the study.

2 Methodology

The idea is to compute a coarse representation of the solution. In [9] a fully algebraic computation of a coarse space is proposed to perform an Aitken acceleration of vectorial sequence generated with an iterative domain decomposition method. In [6] Aitken-SVD Schwarz algorithms were derived for the Aitken Restricted Additive Schwarz preconditioning technique [5].

The choice of constructing the coarse space with the SVD is based on the following properties. Let $G \in \mathbb{R}^{m \times l}$ Assume that the values $\sigma_k, 1 \leq k \leq l$ are ordered in decreasing order and there exists a q such that $\sigma_q > 0$ while $\sigma_q + 1 = 0$. Then G can be decomposed in a dyadic decomposition:

$$G = \sigma_1 U_1 V_1^* + \sigma_2 U_2 V_2^* + \ldots + \sigma_q U_q V_q^*. \tag{3}$$

This means that SVD provides a way to find optimal lower dimensional approximations of a given series of data. More precisely, it produces an orthonormal base for representing the data series in a certain least squares optimal sense. This can be summarized by the theorem of Schmidt-Eckart-Young-Mirsky:

Theorem 1. *A non-unique minimizer X_* of the problem $\min_{X, rankX=q} \|G - X\|_2 = \sigma_{q+1}(G)$, provided that $\sigma_q > \sigma_{q+1}$, is obtained by truncating the dyadic decomposition of 3 to contain the first q terms: $X_* = \sigma_1 U_1 V_1^* + \sigma_2 U_2 V_2^* + \ldots + \sigma_q U_q V_q^*$*

Moreover, the SVD of a matrix is well-conditioned with respect to perturbations of its entries. Consider the matrix $G, B \in \mathbb{R}^{m \times l}$, the Fan inequalities write $\sigma_{q+s+1}(G + B) \leq \sigma_{q+1}(G) + \sigma_{s+1}(B)$ with $q, s \geq 0, q + s + 1 \leq l$. Considering the perturbation matrix E such that $\|E\| = O(\varepsilon)$, then $|\sigma_k(G + E) - \sigma_k(G)| \leq \sigma_1(E) = \|E\|_2, \forall k$. This property does not hold for eigenvalues decomposition where small perturbations in the matrix entries can cause a large change in the eigenvalues.

These properties allow us to search an approximation of the solution in the base linked to the SVD of a sequence of vectors obtained by iterating a linearly convergent iterative process.

Here, we propose a general framework which enables to compute algebraically a two-level additive preconditioner from any preconditioner that can be used in a Richardson iterative process. Algorithm 1 shows the steps to compute M_{2L}^{-1} that way. In step 1, we compute the SVD of l successive iterations stored in a matrix

$G \in \mathbb{R}^{m \times l}$ of a Richardson process (1) having a linear convergence, *i.e.* we compute a dyadic decomposition of G, as $G = \mathbb{U}_l \Sigma_l \mathbb{V}_l^T$, with $\mathbb{U}_l \in \mathbb{R}^{m \times l}$ $\Sigma_l \in \mathbb{R}^{l \times l}$ and $\mathbb{V}_l \in \mathbb{R}^{m \times l}$. In step 2, $\mathbb{U}_q$ is made of the first q columns of $\mathbb{U}_l$ with respect to the decreasing of the singular values $\Sigma_{i,i}$, such that $\mathbb{U}_q$ is full rank. This selection is done according to Theorem 1 where $X_q \in \mathbb{R}^{m \times q}$ is a non-unique minimizer of the problem $\min_{X, rankX=q} \|G - X\|_2 = \sigma_{q+1}(G)$, such that $X_q = \mathbb{U}_q \Sigma_q \mathbb{V}_q^T$ and $rk(\mathbb{U}_q) = q$, with $\mathbb{U}_q \in \mathbb{R}^{m \times q}$, $\Sigma_q \in \mathbb{R}^{q \times q}$ and $\mathbb{V}_q \in \mathbb{R}^{m \times q}$. Once this basis of the coarse space is defined, one can compute the coarse operator (step 3) and solve the coarse problem (step 4).

Algorithm 1 Computation of M_{2L}^{-1} with SVD of solutions of a Richardson process

---

Require: $(u^k)_{0 \le k \le l-1}$, l successive iterates satisfaying $u^{k+1} - u^\infty = (I - \alpha M^{-1} A)(u^k - u^\infty)$ starting from any initial guess u^0
1: Compute the Singular Value Decomposition of the snapshots $G = [u^0, \ldots, u^{l-1}] = \mathbb{U}_l \Sigma_l \mathbb{V}_l^T$
2: Set the index q such that $q = \max_{0 \le i \le l-1} \{\Sigma(i,i) > tol\}$, to define the full rank matrix $\mathbb{U}_q = [U_0, U_1, \ldots, U_q]$ {ex.: $tol = 10^{-12}$.}
3: Define the coarse operator $A_c \in \mathbb{R}^{q \times q}$ such that $A_c = \mathbb{U}_q^T A \mathbb{U}_q$
4: Define the two-level additive preconditioner $M_{2L}^{-1} = M^{-1} + \mathbb{U}_q A_c^{-1} \mathbb{U}_q^T$

It is possible to see this approach as a way to approximate a Krylov subspace. Basically, the solution of the linear system $Au = f$ defined in Section 1 consists on minimizing $F(u^k) = (f - Au^k, f - Au^k)$ on a Krylov space $K_l(A, r^0) = \{r^0, Ar^0, \ldots, A^{l-1}r^0\} = \{d^0, \ldots, d^{l-1}\}$, where from an arbitrary initial solution $u^0 \in \mathbb{R}^m$, $r^0 = f - Au^0$.

Let choose $u^0 = 0$. Each iterate u^k of the Richardson process can be written in a Krylov subspace:

$$u^k = \sum_{i=0}^{k} \beta_i (M^{-1}A)^i M^{-1} f , \quad \beta_i \ne 0$$

Following Algorithm 1, we can write that

$$span(U_0, \ldots, U_{q-1}) \subset span(U_0, \ldots, U_{l-1})$$

Then the solution of the coarse problem is an approximation of a solution in $span(K_l(M^{-1}A, M^{-1}f))$.

This link enable us to choose a good initial guess for the Krylov method preconditioned by this two-level preconditioning approach by computing the solution $u_c \in \mathbb{R}^q$ of the coarse linear system: $A_c u_c = f_c$, with $f_c = \mathbb{U}_q f$.

Then we can set the initial guess for the Krylov method such that,

$$u^0 = \mathbb{U}_q u_c$$

3 Numerical experiments

In this section we propose to apply the methodology for a RAS preconditioner for the solution of CFD problems. The considered matrices A are real, non-symmetric, indefinite and possibly not positive.

The Additive Schwarz (AS) preconditioning is built from the adjacency graph $G = (W, E)$ of A, where $W = \{1, 2, ..., m\}$ and $E = \{(i, j) : a_{ij} \neq 0\}$ are the edges and vertices of G. Starting with a non-overlapping partition $W = \cup_{i=1}^{p} W_{i,0}$ and $\delta \geq 0$ given, the overlapping partition $\{W_{i,\delta}\}$ is obtained defining p partitions $W_{i,\delta} \supset W_{i,\delta-1}$ by including all the immediate neighbouring vertices of the vertices in the partition $W_{i,\delta-1}$. Then the restriction operator $R_{i,\delta} : W \to W_{i,\delta}$ defines the local operator $A_{i,\delta} = R_{i,\delta} A R_{i,\delta}^{T}, A_{i,\delta} \in \mathbb{R}^{m_{i,\delta} \times m_{i,\delta}}$ on $W_{i,\delta}$. The AS preconditioning writes: $M_{AS,\delta}^{-1} = \sum_{i=1}^{p} R_{i,\delta}^{T} A_{i,\delta}^{-1} R_{i,\delta}$. Introducing $\tilde{R}_{i,\delta}$ the restriction matrix on a non-overlapping subdomain $W_{i,0}$, the Restricted Additive Schwarz (RAS) iterative process [2] writes:

$$u^k = u^{k-1} + M_{RAS,\delta}^{-1} \left(f - A u^{k-1} \right), \text{with} M_{RAS,\delta}^{-1} = \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} R_{i,\delta} \qquad (4)$$

When the number of subdomains increases the convergence rate of RAS decreases. When it is applied to linear problems, the RAS has a pure linear rate of convergence.

First we study the robustness and scalability of the preconditioner on a $2D$ driven cavity problem. Second we propose a test of the quality of our coarse space on an $2D$ industrial problem.

3.1 Robustness

Here, we want to study the numerical scalability of the method for the domain decomposition preconditioner chosen. We fix the number of Richardson iteration to perform while decreasing the convergence rate of the preconditioner, *i.e.* we set a coarse space size l and increases the number of partitions p.

We consider a test case called $e30r2000$ coming from modeling 2D fluid flow in a driven cavity proposed in the Matrix Market data collection [3] referenced under the name DRIVCAV. The flow is modeled using the incompressible Navier Stokes equations discretized using Finite Element Method and linearized using Newton's method. The unit square on what the problem is solved is dicretized by 30 elements on the edges. The Reynolds number is set to 2000.

The matrix A is real, non-symmetric and indefinite of size $m = 9661$ and has 306356 entries. The estimated condition number given by the condest function of MATLAB is $\kappa_\infty(A) = 6.77e + 11$.

We partition the operator with the METIS software for partitioning graphs with a multilevel recursive-bisection algorithm, in $p = \{4, 8, 12\}$ partitions. We compute $l = 60$ iterations of a RAS iterative process starting from an initial guess $u^0 = 0$, and perform the SVD of the corresponding sequence of vectors.

Figure 1 (top) shows the singular values profile. When p increases the spectrum coverage decreases which implies a decreasing of the quality of the solution on the coarse space.

Figure 1 (bottom) shows the convergence to the solution of a GCR method preconditioned by a RAS preconditioning technique on the left and enhanced by the given algebraic two-level approach with initialisation of the Krylov method by the solution of the coarse system written on $\mathbb{R}^m$. The convergence rate of the RAS method is reduced for each choice of partitioning. For $p = 4$ and $p = 8$ the initialization by the coarse solution is efficient and we observe an enhancement about 8 and 2 orders of convergence at the initialization respectively. For all partitioning the accuracy is better than for the RAS, *i.e.* the GCR reaches greater convergences and, although there is still a plateau due to the bad conditioning of the system, the convergence to the solution for $p = 12$ can reach 10^{-7} instead of 10^{-5}.

3.2 Quality

Here, we want to observe the influence of the quality of the coarse space on the convergence rate of the preconditioned solution method.

We apply our technique on the case GT01R proposed by a CFD company called FLUOREM, on [1], which deals with steady flow parametrization. From a steady RANS simulation (compressible Navier-Stokes equations) on a reference configuration they obtain linear systems with real, square and indefinite matrices. Those matrices, generated through automatic differentiation of the flow solver around a steady state, correspond to the Jacobian with respect to the conservative fluid variables of the discretized governing equations (finite-volume discretization). The right hand side represents the derivative of the equations with respect to a parameter (of operation or shape).

The CASE_004 GT01 operator comes from a $2D$ inviscid case in the context of a linear cascade turbine. The solution of the discrete system is defined over five variables per node. The discretisation is done among 1596 nodes, describing one inter-blade channel. The stencil involved by the convective scheme uses nine nodes. Thus, there are nine non-zero blocks for each node in the matrix. The peculiarity is that the computational domain is periodic, which introduces some non-zero elements far away from the diagonal. The resulting matrix is real, non-symmetric and not positive definite, of size $m = 7980$.

Figure 2 shows the singular values (left) obtained after 20, 40 or 60 iterations of a RAS iterative process with $p = 8$. For $l = 60$, σ covers 15 orders of magnitude, while it covers 10 orders of magnitude for $l = 40$ and 5 for $l = 20$. For each we
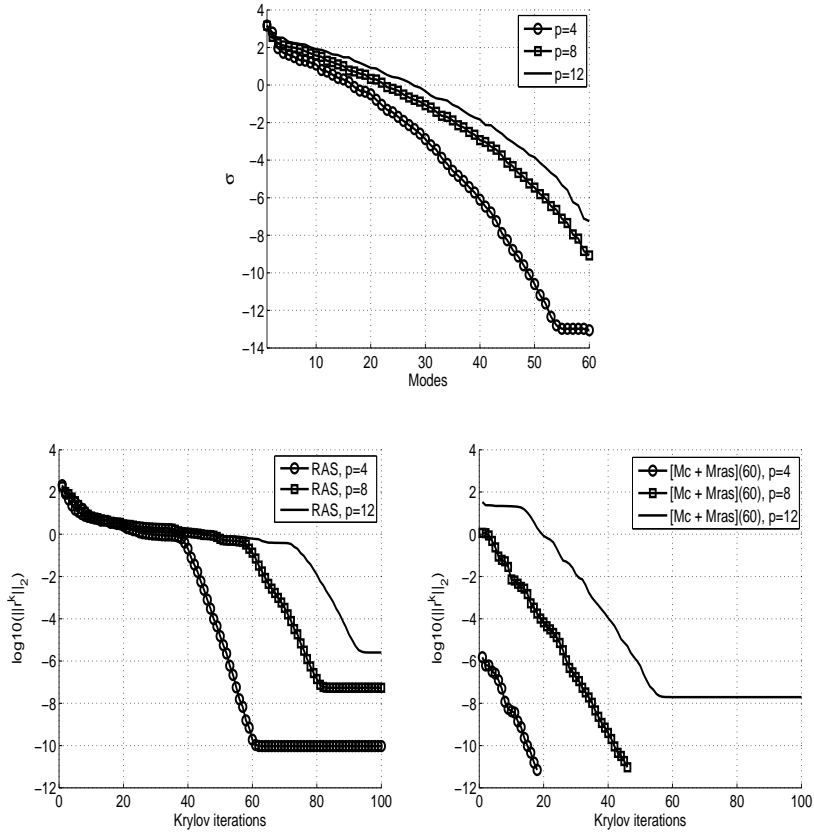
Fig. 1 Solving 2D driven cavity, Re=2000, n=9661, with GCR preconditioned by RAS (left) and ML_RAS_svd(60) (right). Singular values of RAS solutions to compute M_c^{-1} for p = 4, 8, 12 (top).

choose $l = q$. As expected, the convergence of the GMRES (right) is better when q increases. Nevertheless, the convergence plots for 20 and 40 singular values kept are similar.

Table 1 shows the coarse solution accuracy compared to a solution given using *LU* factorization. The greater the number of iterations of a Richardson process is, the better the coarse solution accuracy is.

Those results shows that, although the quality of the coarse space is increasing with the number of Richardson iterations, it is not necessary to compute a lot of singular values to enhance the convergence with this technique.
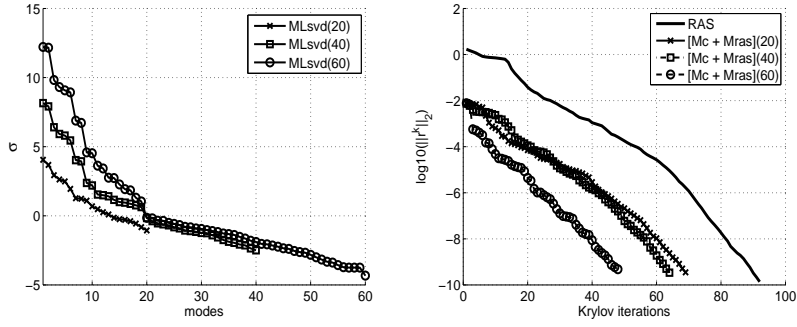
Fig. 2 Solving 2D case, GT01, n=7814, with GMRES preconditioned by RAS and ML_RAS_svd(q), p = 8 (right), Singular values of RAS solutions to compute M_c^{-1} (left)

modes	20	40	60
$\|u_{ex} - \mathbb{U}_q^T u_c\|$	7.23 e-01	5.99 e-02	8.39 e-03

Table 1 Coarse solution accuracy for the GT01 case, compared to a solution given using *LU* factorization.

4 Conclusion

As in [9] and [6] the principle of using the SVD of successive solutions of an iterative process enables to compute a coarse solution without the knowledge of the underlying equations but it not used to accelerate a sequence of vectors but to construct a Krylov subspace. Then it can also be used to construct algebraic coarse levels for a two-level preconditioning technique based on any preconditioner which can be used in an iterative Richardson process.

Numerical results have been shown for the RAS preconditioning technique on two fluid flow problems. The algebraic framewok enables to deal with real, non-symmetric and not positive definite operators. The two-level preconditioners produced are numerically scalable for domain decomposition technique such as RAS and the coarse space enables to compute an approximation of the solution which is used to initialize the chosen Krylov method.

Further work concerns the study of the non-singularity of the coarse operators built with this approach [4]. Moreover, a discussion about the choice of the SVD algorithms and the quality of the coarse space produced should be studied.

References

1. Boisvert, R.F., Pozo, R., Remington, K., Barrett, R.F., Dongarra, J.J.: Matrix market: A web resource for test matrix collections. In: The Quality of Numerical Software: Assessment and Enhancement, pp. 125–137. Chapman & Hall (1997). URL http://math.nist.gov/ MatrixMarket
2. Cai, X.C., Sarkis, M.: A restricted additive Schwarz preconditioner for general sparse linear systems. SIAM J. Sci. Comput. **21**(2), 792–797 (electronic) (1999). DOI 10.1137/ S106482759732678X. URL http://dx.doi.org/10.1137/S106482759732678X
3. Davis, T.A., Hu, Y.: The university of florida sparse matrix collection. ACM Transactions on Mathematical Software (to appear) (20YY). URL http://www.cise.ufl.edu/ research/sparse/matrices
4. Dufaud, T.: An algebraic two-level methodology based on information of a richardson iterative process for constructing preconditioners (20YY). In preparation
5. Dufaud, T., Tromeur-Dervout, D.: Aitken's acceleration of the restricted additive Schwarz preconditioning using coarse approximations on the interface. C. R. Math. Acad. Sci. Paris **348**(13-14), 821–824 (2010). DOI 10.1016/j.crma.2010.06.021. URL http://dx.doi.org/10. 1016/j.crma.2010.06.021
6. Dufaud, T., Tromeur-Dervout, D.: Efficient parallel implementation of the fully algebraic preconditioning technique aras2. Advances in Engineering Software (20YY). Submitted
7. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for nonsymmetric systems of linear equations. SIAM J. Numer. Anal. **20**(2), 345–357 (1983). DOI 10.1137/ 0720023. URL http://dx.doi.org/10.1137/0720023
8. Saad, Y., Schultz, M.H.: GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Statist. Comput. **7**(3), 856–869 (1986). DOI 10.1137/0907058. URL http://dx.doi.org/10.1137/0907058
9. Tromeur-Dervout, D.: Meshfree adaptative aitken-schwarz domain decomposition with application to darcy flow. In: B. Topping, P. Ivanyi (eds.) PARALLEL, DISTRIBUTED AND GRID COMPUTING FOR ENGINEERING, *Computational Science Engineering and Technology Series*, vol. 21, pp. 217–250. SAXE-COBURG PUBLICATIONS (2009)